ISSN 1313 - 0927

Space Research and Technology Institute
Bulgarian Academy of Sciences

bl.1, *Acad. G. Bonchev* St.,
1113 Sofia, Bulgaria
tel./fax: (+359 2) 988 35 03,
http://www.space.bas.bg

Aerospace Research in Bulgaria

33

Bulgarian Academy of Sciences
Space Research and Technology Institute

2021

AEROSPACE RESEARCH IN BULGARIA

Founded in 1978

Volume 33, Sofia, 2021
Space Research and Technology Institute
Bulgarian Academy of Sciences

Address

AEROSPACE RESEARCH IN BULGARIA
Space Research and Technology Institute
Acad. Georgi Bonchev St., block 1, 1113 Sofia, Bulgaria

web-site: http://journal.space.bas.bg
e-mail: journal@space.bas.bg

Pre-Publication Processing
Tsveta Srebrova

p-ISSN 1313 – 0927 e-ISSN 2367 – 9522

This issue of *Aerospace Research in Bulgaria* is published with financial support from the Bulgarian National Science Fund – Contract No КП-06-НП2/48 – 07.12.2020

Bulgarian Academy of Sciences **ISSN 1313 – 0927**

Aerospace Research in Bulgaria

33

Sofia, 2021

C o n t e n t s

С Ъ Д Ъ Р Ж А Н И Е

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e01

DESCRIPTION OF THE "LIULIN TEN-KOH" CHARGED PARTICLES SPECTROMETER FOR THE JAPANESE "TEN-KOH" SATELLITE

Tsvetan Dachev[1], Plamen Dimitrov[1], Borislav Tomov[1], Yuri Matviichuk[1] Premkumar Saganti[2], Doug Holland[3], Kei-Ichi Okuyama[4], Isai Tapia[4]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences
[2]Radiation Institute for Science and Engineering Prairie View A&M University, Prairie View, TX, USA
[3]Holland-Space LLC, Houston, TX, USA
[4]Kyushu Institute of Technology, Graduate School of Engineering, Kyushu, Japan
e-mail: tdachev@bas.bg

Keywords: *Spectrometer, Charged Particle Detector, Liulin Ten-Koh*

Abstract

On 29 October 2018 at 13:08 Japanese Standard Time was successfully launched the Greenhouse gas Observing SATellite (GOSAT-2) from the JAXA Tanegashima Space Centre. Piggy-back with the GOSAT-2 satellite in a circular (623 km), polar synchronous orbit, was launched the 22 kilogram mass satellite Ten-Koh (http://kit-okuyama-lab.com/en/ten-koh/). It was developed in Kyushu Institute of Technology by Prof. K. Okuyama, Chief Scientist of the Spacecraft. Ten-Koh satellite is observing Low Earth Environment (LEO). The primary purpose is to provide valuable data for future development of satellites for operation. Ten-Koh's primary science instrument is the Charged Particle Detector (CPD) developed at the Prairie View A&M University, and NASA Johnson's Space Centre of Houston, TX, USA. Principal Scientist of this payload is Prof. P. Saganti (https://www.pvamu.edu/raise/space-payload/charged-particle-detector-2018/). Principal Engineer of the CPD project is S. D. Holland (formerly with NASA-JSC and currently with Holland-Space LLC, Houston, TX, USA). SRTI-BAS received a request from Prof. Saganti to develop a Liulin type instrument to be part of the CPD payload. Scientists from SRTI's Solar-Terrestrial Physics Department have developed and handed three units per request (engineering, flight, and operational models) of the instrument named "Liulin Ten-Koh". These instruments are similar to the RADOM instrument, which worked in 2008–2009 on the Indian Moon satellite Chadrayaan-1 [12]. This paper describes the flight model, "Liulin Ten-Koh Saganti" instrument and standard sources radiation tests, which were performed during the calibrations in the laboratory of SRTI-BAS. As of this writing, Ten-Koh spacecraft is making a polar orbit passes as expected at about 623 km altitude and at 98 degree inclination with healthy telemetry data as received by several ground stations across the world. The first received data from "Liulin Ten-Koh Saganti" instrument of the Ten-Koh spacecraft are presented. The available at this moment Galactic Cosmic Rays (GCR) L-value profiles of the dose rate and the dose to flux ratio (D/F) from 11 December 2018 are compared with the R3DE profile at the International Space Station (ISS). Additionally, the integral "Liulin Ten-Koh Saganti" instrument LET spectrum from 11 December 2018 is compared with spectra from other instruments, measured in and out of the Earth magnetosphere.

1. Introduction and “Liulin Ten-Koh” instrument description

The most notable difference between occupational radiation exposures that occur on Earth and those in space is that astronauts experience a persistent low background field of radiations of mixed biological effectiveness, including energetic electrons, high-energy heavy-ion component of galactic cosmic rays (GCR), secondary neutrons, and densely ionising low-energy secondary ions and energy-degraded primary ions [1].

The major risk of space travel is cancer from GCR, while circulatory diseases in suggested in some but not all epidemiology studies at modest doses (<1 Gy) and detriments in cognition are suggested by rodent studies following acute irradiation with moderate doses of heavy ions. The GCR are not easily shielded since they consist of high-energy protons, heavy ions and secondary radiation produced in shielding and tissue. Furthermore, heavy ions are more effective per unit dose in causing solid cancers compared to gamma rays. Additionally, non-targeted effects (NTEs) are suggested by most low dose radiobiology studies to increase the biological effectiveness for low doses of high linear energy transfer (LET) radiation [2]. Therefore, precise measurements of radiation sources doses and their distribution in the low earth orbits (LEO) are of great importance.

The “Liulin Ten-Koh” instrument is a Liulin-type deposited energy spectrometer (DES) instrument, which was successfully flown: (1) in the US Laboratory module of the ISS in May–August 1991 [3–9]; (2) inside ESA Biopan-5/6 facilities on Foton M2/M3 satellites [10–11]; (3) Indian Chandrayaan-1 satellite [12] and (4) in-side the ESA EXPOSE-E/R/R2 facilities outside the Columbus/Zvezda modules of the ISS in 2008–2016 [13–15].

Fig. 1 shows the external view of the 3 “Liulin Ten-Koh” instruments. The first picture to the left presents the technological model “Liulin Ten-Koh”. In the middle is the flight model (“Liulin Ten-Koh Saganti”, while the spare model (“Liulin Ten-Koh Saganti 2F”, is at the right side of the figure. With the instruments are shown the necessary for laboratory use with computer external power supplies, RS-232 bridges and cables.

The “Liulin Ten-Koh” spectrometer main purpose is to measure the spectrum (in 256 channels) of the deposited energy from primary and secondary particles in the Ten-Koh satellite radiation environment.

The “Liulin Ten-Koh” spectrometer is designed as spectrometer-dosimeter for continuous monitoring of the satellite radiation environment, which can consist of GCR particle, Inner radiation belt (IRB) protons, outer radiation belt (ORB) relativistic electrons and energetic protons from solar origin (SEP). The last sources selection procedure was developed and published recently in [15].

Fig. 1. External view of the 3 “Liulin Ten-Koh” instruments. The instruments are presented with the necessary: external power supply, RS-232 bridges and cables.

After switching on, the “Liulin Ten-Koh” spectrometer performs internal test and wait to obtain external command. On a receipt of command, it starts to accumulate in 256 channels the spectrum of the deposited energy used further to calculate the dose and flux of particles in the silicon detector. The exposition time of one spectrum depends on the frequency of the external command issue to measure and can be between 5 and 2100 sec with 1 sec resolution.

The “Liulin Ten-Koh” spectrometer contains: one semiconductor detector silicon PIN diode of Hamamatsu S2744-08 type with 2 cm^2 in area and 0.3 mm thickness, one low noise hybrid charge-sensitive preamplifier A225F type of AMPTEK Inc., a fast 12 channel ADC, 2 microcontrollers and buffer memory. The pulse high analysis technique is used for measurement of the energy losses in the detector through specially developed firmware. A block schema of “Liulin Ten-Koh” portable spectrometer-dosimeter is presented in Fig. 2.

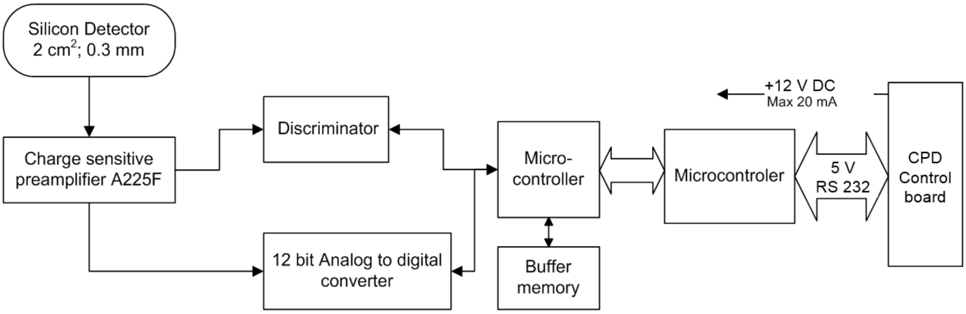


Fig. 2. Block-scheme of the “Liulin Ten-Koh” spectrometer

The main measurement unit in the "Liulin Ten-Koh" instrument is the amplitude of the pulse after the preamplifier, generated by particles or quanta, hitting the detector [3]. The amplitude of the pulse is proportional to a factor of 240 mV MeV^{-1} to the energy loss in the detector and, respectively, to the dose. By 12-bit analogue to digital converter (ADC) these amplitudes are digitised and organised in a 256-channel deposited energy spectrum. The dose in the silicon detector D_{Si} [Gy] by definition in System international (SI) is one Joule deposited in a 1-kg matter. The absorbed dose is calculated by dividing the summarised in 256 channels energy depositions in the spectrum in Joules to the mass of the detector in kilogrammes.

The semiconductor detector of the "Liulin Ten-Koh" instrument is mounted approximately 2 mm below the 0.3 mm thick aluminium cover plate. Furthermore, there is shielding from 0.07 mm copper and 0.2 mm plastic, which provided 0.3 g cm^{-2} of total shielding from the front side. The "Liulin Ten-Koh" instrument is additionally shielded by 2 FR-4 (glass epoxy) plates with a total thickness of 3.15 mm and by 5.0 mm carbon fiber reinforced polymer (CFRP) The calculated required kinetic energy of particles penetrating all the shielding's perpendicular to the detector is 2.6 MeV for electrons and 62.5 MeV for protons (https://physics.nist.gov/PhysRefData/Star/Text/PSTAR.html). The obtained kinetic energy values are approximatively because in the cited above tables are not listed exactly the FR-4 and CFRP materials.

This means that only electrons and protons with energies exceeding the values listed above can cross the "Ten-Koh" satellite and "Liulin Ten-Koh" instrument shielding materials and reach the detector surface. The detector shielding, being larger from the sides and from behind, stops less energetic ORB relativistic electrons, attenuates the lover energy IRB and SEP protons, but does not change the flux of the primary GCR particles.

2. The Charged Particle Detector (CPD) and the "Ten-Koh" satellite

The "Liulin Ten-Koh" instrument is mounted on the top (Fig. 3) of the Charged Particle Detector (CPD) developed at the Prairie View A&M University, and NASA Johnson's Space Centre of Houston, TX, USA. Principal Scientist of this payload is Prof. P. Saganti (https://www.pvamu.edu/raise/space-payload/charged-particle-detector-2018/). Principal Engineer of the CPD project is S. D. Holland (formerly with NASA-JSC and currently with Holland-Space LLC, Houston, TX, USA (www.holland-space.com).

The CPD manages the measurements with the following sensors: Liulin Spectrometer, 2 Open Sensors for Ambient Radiation Measurements, 2 Polyethylene Covered Sensors for Shielding Assessment, 2 Polyethylene Covered Sensors for Skin Dose Assessment, 2 X-ray Detectors. CPD also communication with In-Flight programing capacity with "Ten-Koh" satellite (Fig. 3).

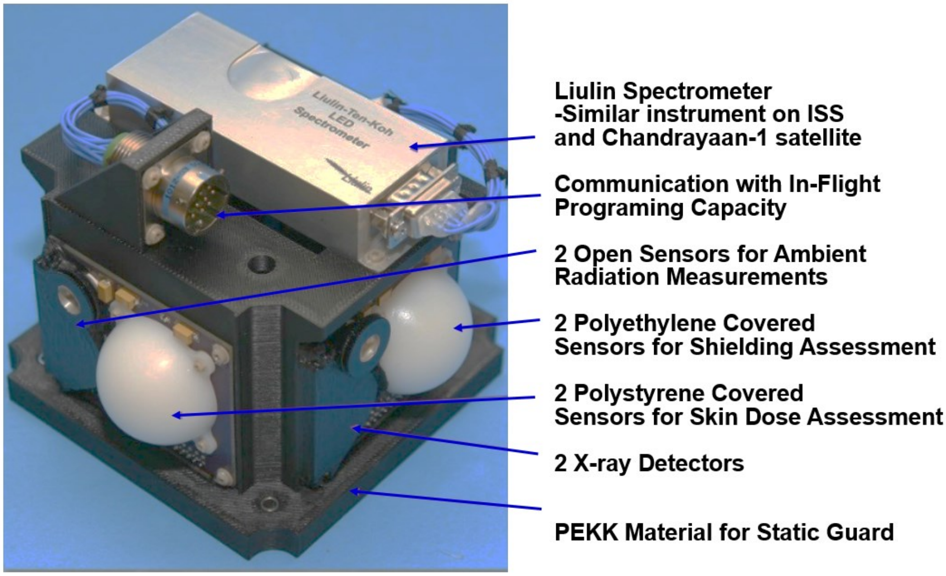


Fig. 3. The CPD with mounted above "Liulin Ten-Koh" instrument

The Ten-Koh satellite was developed at Kyushu Institute of Technology by Prof. K. Okuyama, Chief Scientist (http://kit-okuyama-lab.com/en/ten-koh/). The satellite shape is quasi-spherical (Fig. 4), which diameter is about 500 mm and surrounded by solar cells. This satellite is constituted of structure, thermal control, attitude control, telecommunication, Bus power system and payload power system. The exact sizes are H465×W500×D500 mm and the mass is about 22.0 kg. The Ten-Koh satellite houses the CPD behind the top solar panel and 2 FR-4 plates.

The first purpose of the Ten-Koh satellite missions is to observe LEO environment and initial unveiling of the data. Recently, as the development of microsatellites by universities and other institutes has become active, breakdowns of them on LEO 2000 km below have often occurred. Most of these causes are cosmic radiation, which has various energy levels and come from the ORB and Sun. To prevent these breakdowns, the "Ten-Koh will measure various LEO environment parameters and unveil the data via the internet initially. The second purpose of the Ten-Koh satellite is to measure the degradation of advanced material, which can use in the future.

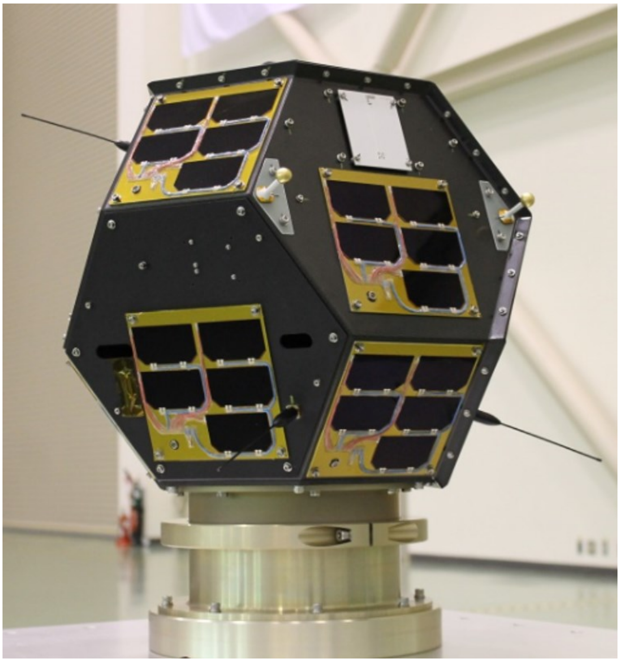


Fig. 4. The “Ten-Koh” satellite

On 29 October 2018 at 13:08 Japanese Standard Time was successfully launched the Greenhouse Gas Observing SATellite (GOSAT-2) from the JAXA Tanegashima Space Centre. Piggy-pack with the GOSAT-2 satellite in a circular (623 km), polar synchronous orbit was launched the “Ten-Koh” satellite.

3. The “Liulin-Ten-Koh 2.exe” software product

The “Liulin-Ten-Koh 2.exe” software product is used for managing the “Liulin-Ten-Koh” spectrometer and for express analysis of the results when is used directly with PC in laboratory tests and measurements. The software includes subprograms for data listing and data visualisations.

The “Liulin-Ten-Koh 2.exe” software was developed in the Windows environment. The purpose of the software is to manage the Liulin spectrometer performance and to analyse and visualise the data during the laboratory tests. At the PC it automatically creates the subdirectory "Data" in the directory in which “Liulin-Ten-Koh.exe” is located. In the subdirectory, "Data" one binary file with extension “LTF” is automatically created at the end of measurements when reading the data from “Liulin-Ten-Koh 2.exe” is performed. The binary file of “Liulin Ten-Koh Saganti 2F” instrument is named automatically and contains in the name the string “YYMMDDhhmm”. The extension of the binary file is “LTF”.

"YYMMDDhhmm" is the date and time of the moment when the measurements begin. This file contains the rough binary data and is for permanent storage of data from the instrument because it is with minimal volume. Three ASCII files are automatically created in the same subdirectory "Data", from the binary file, when "Select HEX data" button in "Liulin-Ten-Koh 2.exe" is activated. The names of the ASCII files contain the same "YYMMDDhhmm" string as the binary file. For example, the ASCII files with extensions of type "d2F", "s2F", and "y2F" contain the "D"ose, "S"pectrum, "Y" (pure spectrum) data, respectively, from the Liulin Ten-Koh Saganti 2F instrument.

4. Standard radiation sources tests and calibrations

The calibration procedures, which were performed using analogical to "Liulin Ten-Koh" instruments, are described in [2, 9 and 15]. The response curve of the "Liulin Ten-Koh" instrument is expected to be similar to that published by Uchihori et al. (2002) [16] because all Liulin DES instruments were manufactured using the same electronic parts and schematic. In a specific example [16] of the calibration was performed by Dr. Yukio Uchihori, it was found that the linear coefficient of the response curve, obtained during the calibrations with protons, He^+ and Ne^+ ions, was equal to 81.3 keV, whereas the Liulin DES instruments predicted value was 81.4 keV.

The first procedure during the calibration process of the "Liulin Ten-Koh Saganti" instrument in the laboratory of IKIT-BAS was to adjust the position of the first spectrometric channel using the ^{241}Am 60 keV gamma line. Furthermore, the linearity is controlled by electronic methods, as described by AMPTEK INC. In the A225 preamplifier-operating notes (http://www.amptek.com/pdf/a225.pdf).

Fig. 5 was obtained during the tests of 'Liulin-Ten-Koh Saganti" (flight) instrument in the SRTI-BAS laboratory on 12/02/2018. The exposition time is 15 sec. The dose curve is with red points and lines, while the flux curve is with blue points and lines. The low doses in the left, middle and right part of the picture correspond to natural background radiation that is why the calculated average dose between the two vertical lines in the centre of Fig. 5 is 0.1296 μGy h^{-1}. This value is close to the natural background radiation but a little higher because of small exposition time and, respectively, small statistics. The first stage (step) in doses and fluxes in Fig. 5 was generated using 60 keV gamma line from ^{241}Am source, while the second one corresponds to much higher doses (~300 μGy h^{-1}) and fluxes (120 cm^{-2} s^{-1}) generated by ^{137}Cs source. The calibrations with ^{137}Cs source are the second procedure in the calibration process, which shows that the instrument works stable at high count and dose rates expected in space.

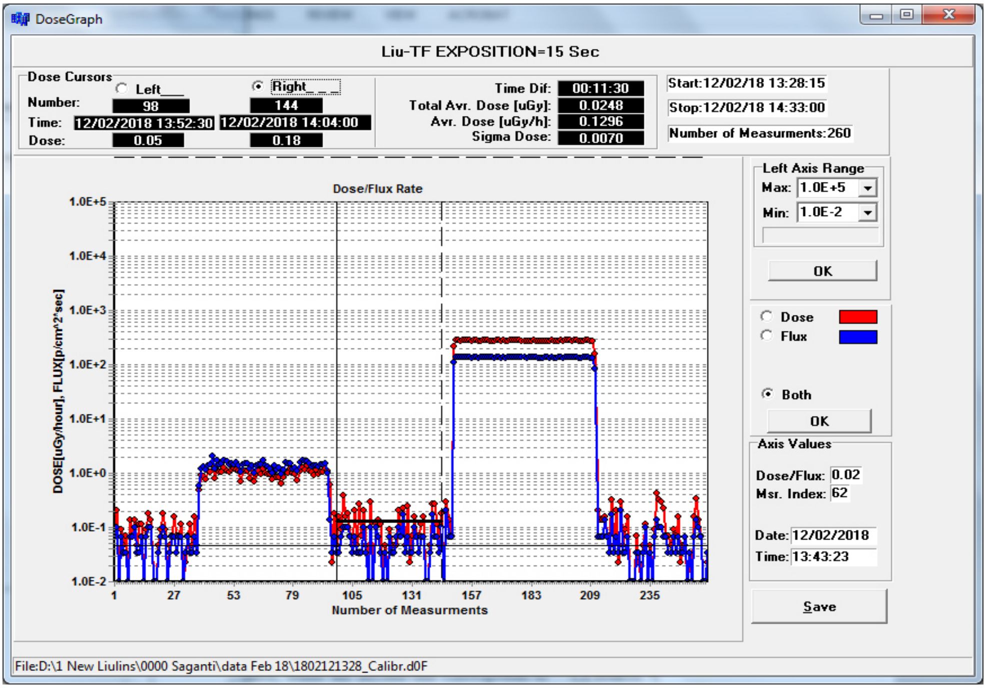


Fig. 5. Calibrations with ^{241}Am and ^{137}Cs sources

Fig. 6 is the colour-coded Energy/Time diagram of the same tests as in Fig. 5 of "Liulin Ten-Koh Saganti". It is well seen in the upper panel of Fig. 6 that the ^{241}Am 60 keV gamma line source produces high-count rate only in the first channel of the "Liulin Ten-Koh Saganti" spectrometer what is expected because the energy loss in this channel is between 40 and 120 keV. The ^{137}Cs source produces a wider spectrum (up to 10th channel) with a maximum count rate in the second channel. In the lover panel of Fig. 6 the sum of the events in each spectra is displayed.

The last procedure during the calibration process is long-term background measurements, which must contain events with high-energy depositions in high channel number depositing high dose rates. The latter confirm whole calibration process and the ability of the instrument to work in real space conditions

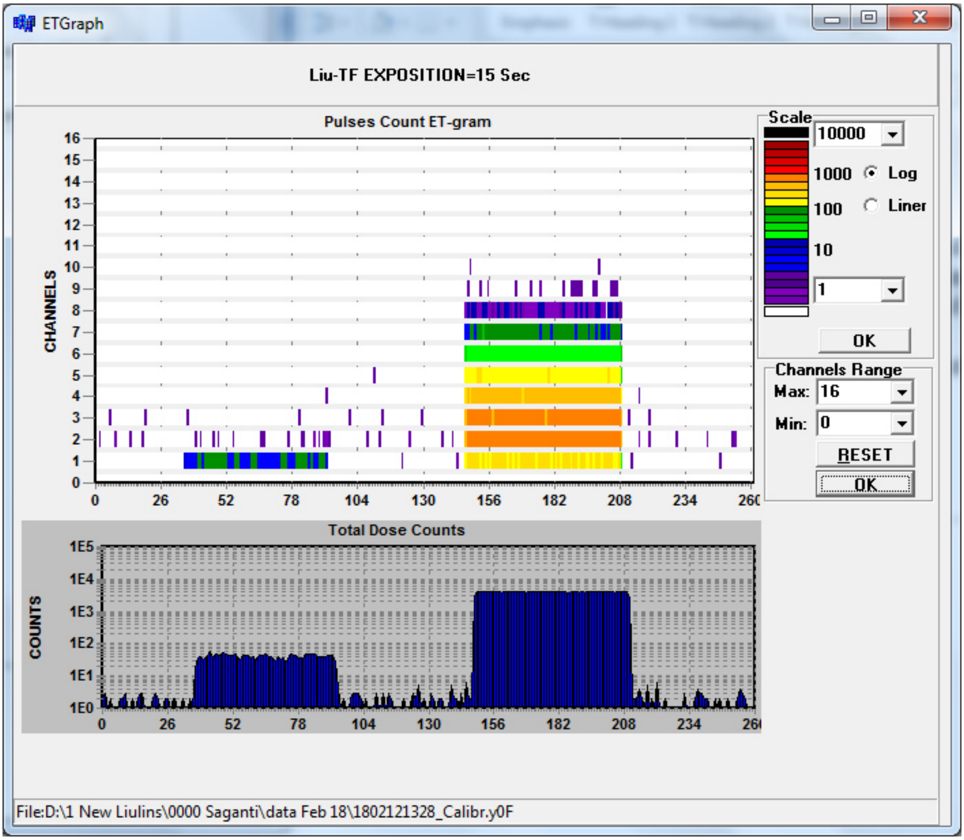


Fig. 6. Same as in Fig. 5 but colour-coded

Fig. 7 shows the dose/flux graphic obtained from the long-term measurements performed using “Liulin Ten-Koh Saganti” instrument. It is seen that the start was on 09/02/18 at 19:04:00 LT. 485 measurements with 300 seconds exposition were performed and the stop time was on 11/02/18 at 11:24:00 LT. This information is automatically calculated by the software and presented in the black filled table cells in the right part of Fig. 7. From the right part black filled table cells, it is seen that for 40:20:00 hours the total dose is 4.9752 μGy, while the average dose rate is 0.1234 μGy h^{-1}. Sigma of the dose is 0.0006. Very high dose events are seen in the left-hand part of Fig. 7. The highest dose rate is ~ 0.42 μGy h^{-1} seen in magenta circles of Fig. 7.

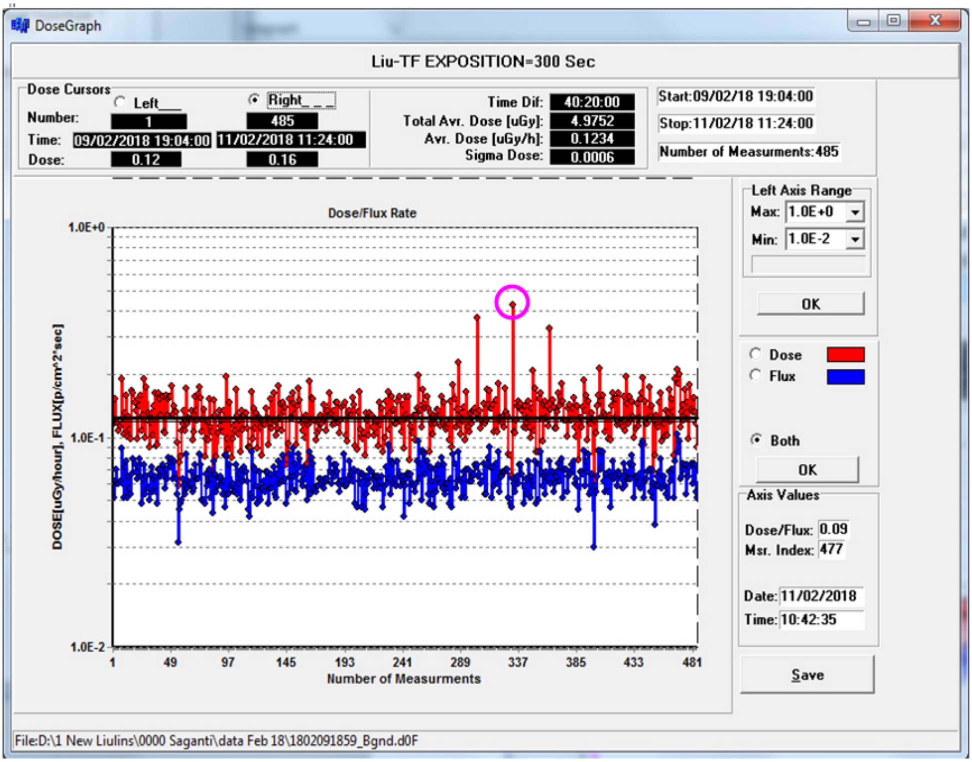


Fig. 7. Long-term background measurements

Fig. 8 shows the spectrum graphic obtained from the long-term measurements performed using “Liulin Ten-Koh Saganti” instrument. The red points present the separate counts in each channel from 1 to 256, while the blue points and line present the summarised number of counts in each channel. It is seen that the major number of points are distributed normally in the channel range from 1 to 16 with maximum counts in the 2nd channel. Accidentally, high-energy (dose) depositions are observed in channel numbers of the spectrometer around 146th, 206th and 256th channels (Emphasized with the magenta circle.). These high-energy depositions once again confirm that the “Liulin Ten-Koh Saganti” spectrometer covers the whole energy range.

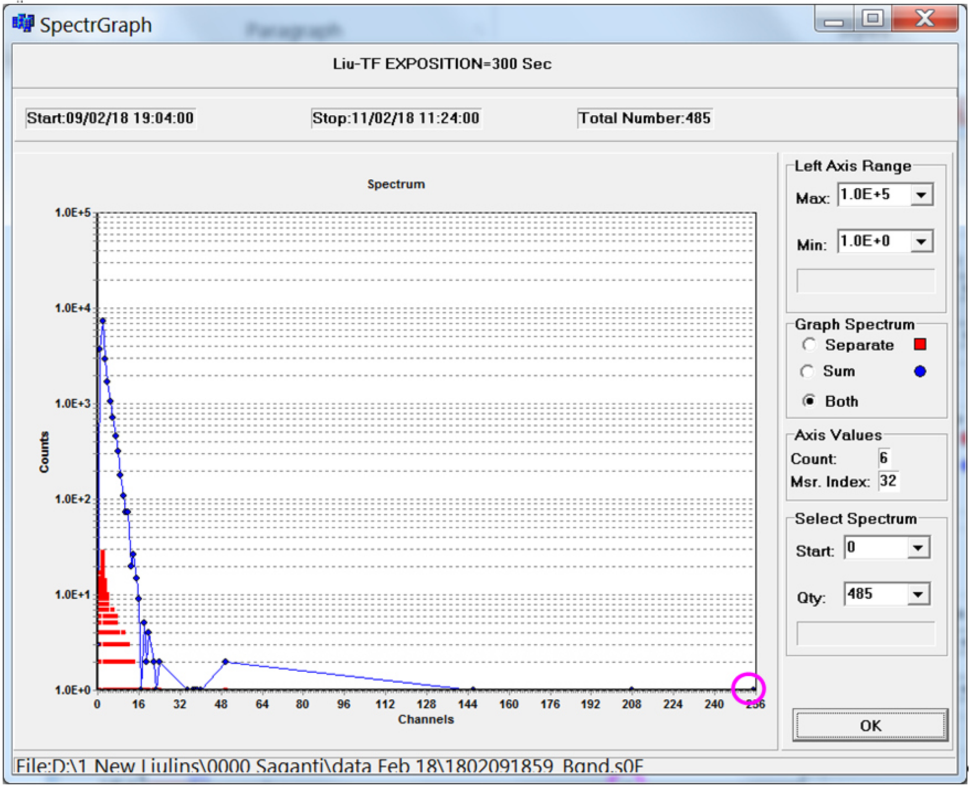


Fig. 8. The obtained spectrum during the long-term background measurements

5. Analysis of the first data obtained from "Liulin Ten-Koh Saganti" instrument

The first command transmission and data downlink with Ten-Koh satellite happen over KIT, Japan on 31 Oct 2018. The first "Liulin Ten-Koh Saganti" spectrum was received on 08 November 2018 at 13:17:00 Japanese Standard Time. The full number of 5 measurements is presented in Table 1.

From Table 1 it is seen that the "Liulin Ten-Koh Saganti" average dose rate is 1.72 μGy h^{-1}, while the average flux is 0.28 cm^{-2} s^{-1}. The values are typical for GCR at low latitude ~34°N at the longitude of about ~140°E. To confirm these values, we take data from the R3DR2 instrument on the ISS in the time interval 21–30 June 2015. The GCR L-value profile for this time interval can be seen at Fig 3b of [15]. The obtained R3DR2 average dose rate and flux values from 46 points inside a rectangular area with coordinates between 139° and 141° geographic longitude and between 30° and 40° geographic latitude are 1.57 μGy h^{-1} and

0.45 cm^{-2} s^{-1}. These values are close to the “Liulin Ten-Koh Saganti” values and confirms our expectations that GCR particles are registered.

Table 1

Day/Time	Exposition [sec]	Counts	Flux [cm^{-2} s^{-1}]	Dose rate [μGy h^{-1}]
08/11/2018 13:17:00	15.286	10	0.33	2.13
12/11/2018 13:59:00	15.286	11	0.39	2.35
12/11/2018	15.286	6	0.19	0.39
18/11/2018 13:33:53	15.286	8	0.26	2.79
18/11/2018	15.286	7	0.23	0.92
Aver. (Liulin Ten-Koh)			**0.28**	**1.72**
Aver. (R3DR2)			**0.46**	**1.57**

The number of 9 “Liulin Ten-Koh Saganti” spectra was received on 11 December 2018. The date, time and geographic coordinates of Ten-Koh satellite normal mode execution is characterised by Mission starting time: 09:02:31 JST (11-12-2018 00:02:31 UTC) and mission-ending time: 09:07:05 JST (11-12-2018 00:07:05 UTC); The starting latitude is 47.3239°N, while the starting longitude is 6.3512°W.

Keeping in mind the amount of more than 2 MeV energetic electron flux in the L-shell plot from the MagEIS instrument aboard the Van Allen Probes (https://www.swpc.noaa.gov/products/van-allen-probes-radiation-belt-plots) for 11th December 2018 our first concern was to establish in the best way the type of predominant radiation source (GCR particles or outer radiation belt (ORB) relativistic electrons) measured. Recently, in [15] it was shown that the best analysis of the radiation source type could be made from the deposited energy spectrum shape.

Figure 9 illustrates the different shapes of the deposited energy spectra (GCR, IRB, ORB and SEP) as obtained from the R3DR2 instrument from 21 to 30 June 2015 (data seen in figures 3a and 4 of [15]). The deposited dose rate is the area between the abscissa and the curve of the deposited energy spectrum. The “GCR” and “IRB,” spectra obtained in the period 21–30 June 2015 is shown in Figure 9 only to validate the R3DR2 data spectrum shapes against those described in [17] that are why we will not analyse them further. The “ORB” spectrum is clearly divided into 2 parts. The low-energy part up to 2 MeV deposited energy is populated by relativistic electrons from ORB, while the high-energy part up to 20.83 MeV is populated by GCR particles, observed at high latitude region. Therefore, this part of the spectrum looks similar to the GCR spectrum but elevated to higher values because smaller geomagnetic shielding. (Please, read more in the Fig. 10-description.)

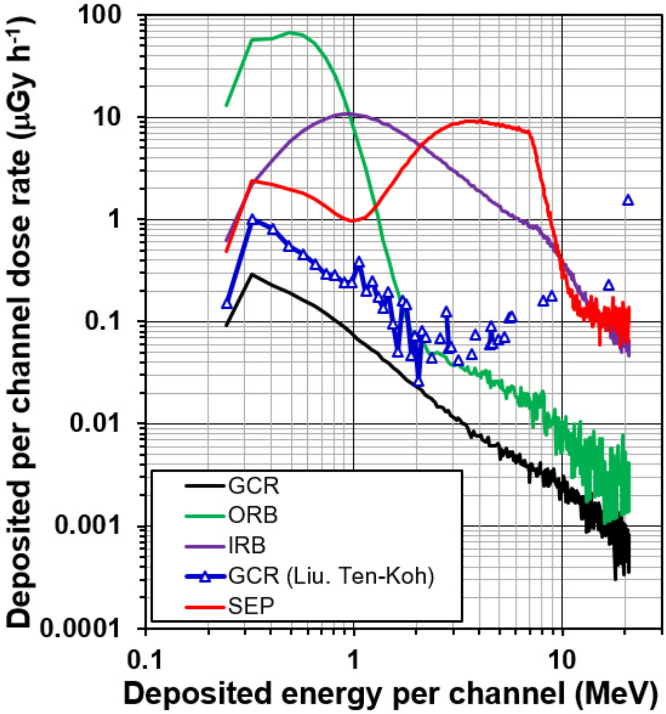


Fig. 9. Comparison of the "Liulin Ten-Koh Saganti" deposited energy spectrum shape with spectra, obtained on ISS with R3DR2 instrument for the period 21–30 June 2015

The "Liulin Ten-Koh Saganti" average spectrum is presented with blue triangles and lines. Nevertheless that, the values of this spectrum are higher than the GCR R3DR2 instrument spectrum (black line), the shape of this spectrum looks similar to the R3DR2 spectrum up to deposited energy of 3 MeV. Further, the Liulin Ten-Koh spectrum points, being only single points, depend strongly by the energy depositions in the channel (i.e. channel number) and their values are far above the R3DR2 instrument GCR spectrum values. Generally, the "Liulin Ten-Koh" spectrum is above the R3DR2 spectrum because

- The R3DR2 spectrum represents the global GCR dose rate, which is in average 2.78 μGy h^{-1}, while the Liulin Ten-Koh average spectrum represents only high latitude doses with an average value of 10.52 μGy h^{-1}. This is confirmed with the shape and position of the R3DR2 ORB spectrum above 3 MeV. This part of the ORB spectrum, obtained at high latitudes, contain GCR particles [15] that is why this part looks like natural continuation of the Liulin Ten-Koh average spectrum above 3 MeV;

- The higher altitude of the Ten-Koh satellite (~620 km) than ISS (415 km);

- The lover solar activity, which increases the GCR flux in the near Earth radiation environment. The small elevation of the first 2 points of the Liulin Ten-Koh spectrum can be interpreted as the presence of small, non-possible to distinguish, number of ORB relativistic electrons.

Using data from a rectangle with a-bit wider geographic region of BION-M No1 satellite [18], we calculate the end point L-coordinates of the 9 Ten-Koh satellite points. Fig. 10a shows the L-value [19] profile of the obtained 9 "Liulin Ten-Koh Saganti" dose rates and dose to flux (D/F) ratio.

Two variables are presented in Fig. 10a. The heavy red points present the measured dose rates, while the heavy blue points present the calculated dose to flux ratio. From previous experiments [15], we know that the D/F parameter could not fully characterise the type of predominant radiation source in the spectrum but can give significant information. The fact that the D/F ratio in Fig. 10a have values close to 1 nGy cm^2 particle^{-1} allow us to conclude that the predominant source are GCR particles. The lack of values greater than 2 nGy cm^2 particle^{-1} and smaller than 0.7 nGy cm^2 particle^{-1} allow us to conclude [15] that spectra with predominant IRB and ORB sources are not observed in the data. The comparison with the D/F data in Fig. 10b also confirms this conclusion.

Geomagnetic shielding [20] is the reason for reduced GCR dose rates of the first 2 points at low L values in Figure 10a and for the slightly rising dose rates towards L values of 2.5. At these increasing L values the vertical cut-off rigidity decreases, and the major number of low-energy GCR spectra penetrate down to the ten-Koh orbit. At higher L values, up to L = 4.5, the dose rate has fixed values because the small increase in the high-energy flux of the primary GCR flux does not affect it. The similarity of the polynomial asymptote (black curve) through the "Liulin Ten-Koh Saganti" dose rates shown in Fig. 10a with the moving average over 50 points black curve in Fig. 10b also confirms the consideration that the observed 9 spectra contain mainly GCR source particles. As in Fig. 9 the "Liulin Ten-Koh Saganti" dose rates are higher even the R3DE dose rates, obtained at the previous solar minimum period in 2009. The reasons are the same as already discussed in the previous paragraphs.

As GCR pass through a target, many electromagnetic and nuclear interactions cause the incident particles to deposit some of their kinetic energy into the target material. The energy is deposited primarily in the form of ionisation of atoms in the target. The rate at which the incident particle deposits its energy in the target is termed linear energy transfer (LET = –dE/dx), energy deposited per unit path length) [21]. The LET spectrum and its evolution through the human body are essential ingredients in understanding and mitigating the potential radiation risk posed by energetic particles [22].

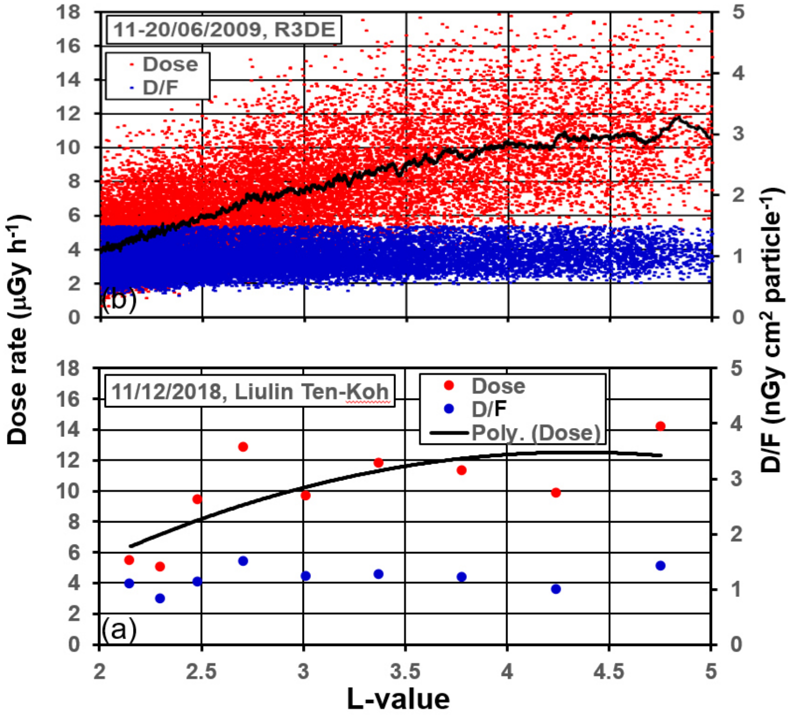


Fig. 10. Comparison of the L-Value profiles, obtained from the "Liulin Ten-Koh Saganti" spectrometer with R3DE data at ISS

Fig. 11 [23] compare:

- 2012–2013 silicon LET spectrum (large black points) from RAD instrument on Curiosity rower at the surface of Mars behind ~21 g cm^{-2} CO_2 shielding;

- 2009–2011 silicon LET spectra from CRaTER instrument [24] on board the Lunar Reconnaissance Orbiter (LRO), which orbited the Moon in a 50 km (average) polar orbit with a period of about 100 min. The data were obtained behind 0.2 g cm^{-2} (CRaTER D1/D2, small black points), 6 g cm^{-2} (CRaTER D3/D4, small red points), and 9 g cm^{-2} (CRaTER D5/D6, small blue points);

- 2008 silicon LET GCR spectrum (magenta points) from RADOM instrument [12], obtained behind 0.3 g cm^{-2} shielding in the interplanetary space between Earth and Moon in the period 29/10/2008 09:46:12-08/11/2008 00:00:00 UT. 52,687 10 s spectra were averaged;

- 2014–2016 silicon LET GCR global distribution spectrum (dark green points) from R3DR2 instrument, obtained behind 0.3 g cm^{-2} shielding. The spectrum was obtained by averaging 3,393,592 10 s GCR spectra in the period 23/10/2014 10:31:43-10/01/2016 23:59:56 UT.

- 2018 silicon LET GCR spectrum (blue triangles) from “Liulin Ten-Koh Saganti” instrument, obtained behind more than 0.3 g cm^{-2} shielding on 11 December 2018. The exact shielding is unknown but it is expected to be more than the self-shielding of the instrument because it is situated deeper in the construction of the Ten-Koh satellite and behind solar cell panel. The spectrum was obtained by averaging four 25.165824 s GCR spectra and five 29.61984 s spectra.

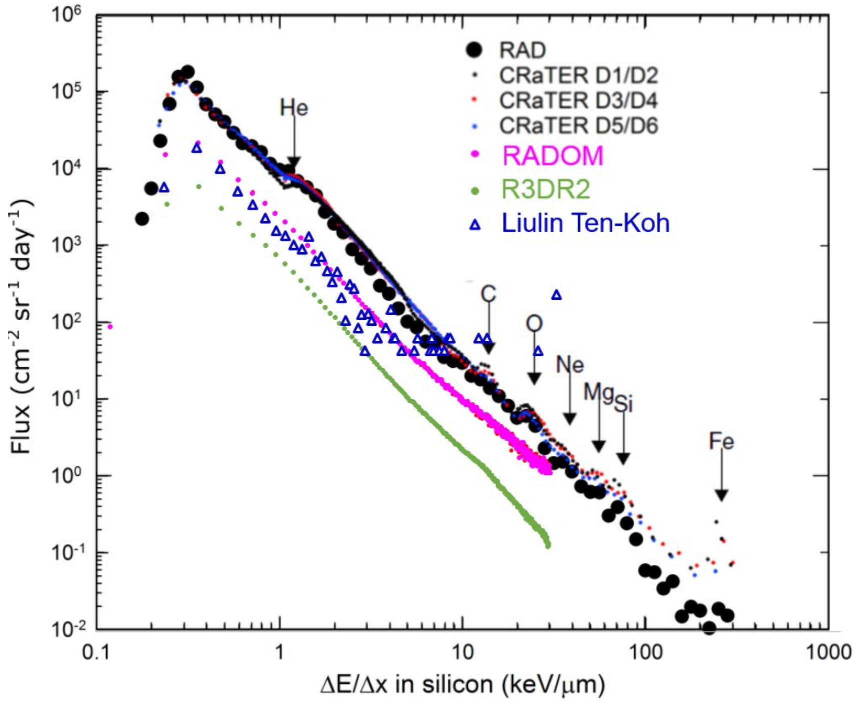


Fig. 11. Comparison of LET spectra, obtained by different instruments at different carriers inside and outside the Earth magnetosphere

Fig. 11 shows relatively good agreement between the shapes of the different spectra. The RADOM, R3DR2 and “Liulin Ten-Koh Saganti” spectra are shorter than the RAD and CRaTER spectra because they were obtained using single silicon detector, which covers only the LET range between 0.233 and 29.8 keV μm^{-1}. These spectra include the energy depositions of Neon (Ne^{+}) ions as obtained by Dr Y. Uchihori [16] (see Fig. 11 there) but in Fig. 11, this is not seen, probably because unclear position of Neon maximum in the CRaTER spectra. The RADOM spectrum is below the RAD and CRaTER spectra because smaller shielding. The R3DR2 GCR spectrum is obtained inside the Earth magnetosphere and upper atmosphere that is why the number of CGR particles building the spectrum is smaller than the RAD, CRaTER and RADOM spectra. The "Liulin Ten-Koh Saganti" spectrum is a little bit below the RADOM spectrum because these data are obtained inside the magnetosphere in relatively small geographic region at middle latitudes.

6. Conclusions

This paper reveals that the authors and colleagues had successfully fulfilled the main task, i.e. to develop, construct and calibrate an engineering, flight, and operational models of a Liulin type particle spectrometer. After the standard sources radiation tests, which were performed during the calibrations, the flight model of the instrument was incorporated in the CPD instrument of Ten-Koh spacecraft. Nevertheless that the first amount of data from "Liulin Ten-Koh Saganti" spectrometer is relatively limited the comparisons show that they are correct and can be used for further analysis of the LEO environment.

References

1. Cucinotta, F. A, M. R. Shavers, P. Saganti, and J. Miller, Radiation Protection Studies of International, Space Station Extravehicular Activity Space Suits, NASA/TP-2003-212051. https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20040031719.pdf.
2. Cucinotta, F. A., Cacao, E., Kim, M. H. Y., and Saganti, P. B., Cancer and circulatory disease risks for a human mission to Mars: Private mission considerations. Acta Astronautica, 2018. https://doi.org/10.1016/j.actaastro.2018.08.022.
3. Dachev, Ts., B. Tomov, Yu. Matviichuk, Pl. Dimitrov, J. Lemaire, Gh. Gregoire, M. Cyamukungu, H. Schmitz, K. Fujitaka, Y. Uchihori, H. Kitamura, G. Reitz, R. Beaujean, V. Petrov, V. Shurshakov, V. Benghin, F. Spurny, Calibration Results Obtained With Liulin-4 Type Dosimeters. Adv. Space Res. 2002, 30, 4, 917–25. http://dx.doi.org/10.1016/S0273-1177(02)00411-8.
4. Reitz, G., R. Beaujean, E. Benton, S. Burmeister, Ts. Dachev, S. Deme, M. Luszik-Bhadra, and P. Olko, Space radiation measurements on-board ISS-the DOSMAP experiment. Radiat Prot Dosimetry. 2005, 116, 374–79. http://rpd.oxfordjournals.org/cgi/content/abstract/116/1-4/374.
5. Nealy, J. E., F. A. Cucinotta, J. W. Wilson, F. F. Badavi, N. Zapp, T. Dachev, B. T. Tomov, E. Semones, S. A. Walker, G. de Angelis, S. R. Blattnig, and W. Atwell, Pre-engineering spaceflight validation of environmental models and the 2005 HZETRN simulation code, Adv. Space Res., 2007, 40, 11, 1593–1610. http://dx.doi.org/10.1016/j.asr.2006.12.030.
6. Wilson, J. W., J. E. Nealy, T. Dachev, B. T. Tomov, F. A. Cucinotta, F. F. Badavi, G. de Angelis, N. Leutke, and W. Atwell, Time serial analysis of the induced LEO environment within the ISS 6A, Adv. Space Res., 2007, 40, 11, 1562–70. http://dx.doi.org/10.1016/j.asr.2006.12.030.

7. Slaba, T. C., S. R. Blattnig, F. F. Badavi, N. N. Stoffle, R. D. Rutledge, K. T. Lee, E. N. Zappe, T.P. Dachev, and B.T. Tomov, Statistical Validation of HZETRN as a Function of Vertical Cutoff Rigidity using ISS Measurements, Adv. Space Res., 2011, 47, 600–10. http://dx.doi.org/10.1016/j.asr.2010.10.021
8. Badavi, F. F., Validation of the new trapped environment AE9/AP9/SPM at low Earth orbit. Adv. Space Res. 2014, 54, 917–28. DOI:10.1016/j.asr. 2014.05.010.
9. Dachev, T. P., J. V. Semkova, B. T. Tomov, Yu. N. Matviichuk, Pl. G. S. Maltchev, R. Koleva, Pl., Dimitrov, N. G. Bankov, V. V., Shurshakov, V. V., Benghin, E. N., Yarmanova, O. A. Ivanova, D.-P. Häder, M. T. Schuster, G. Reitz, G. Horneck, Y. Uchihori, H. Kitamura, O. Ploc, J. Kubancak, and I. Nikolaev, Overview of the Liulin type instruments for space radiation measurement and their scientific results, 2015, 92–114. http://dx.doi.org/10.1016/j.lssr.2015.01.005.
10. Häder, D. P., P. Richter, M. Schuster, Ts. Dachev, B. Tomov, Pl. Georgiev, and Yu. Matviichuk, R3D-B2 - Measurement of ionizing and solar radiation in open space in the BIOPAN 5 facility outside the FOTON M2 satellite, Adv. Space Res. 2009, 43, 8, 1200–11. http://dx.doi.org/10.1016/j.asr. 2009.01.021.
11. Damasso, M., Ts. Dachev, G. Falzetta, Giardi M. T., Rea G., and Zanini A., The radiation environment observed by Liulin-Photo and R3D-B3 spectrum-dosimeters inside and outside Foton-M3 spacecraft, Radiation Measurements, 2009, 44, 3, 263–72. http://dx.doi.org/10.1016/j.radmeas.2009.03.007.
12. Dachev, Ts. P., B. T. Tomov, Yu. N. Matviichuk, Pl .G. Dimitrov, Vadawale, S. V., J. N. Goswami, V. Girish, and G. de Angelis, An overview of RADOM results for Earth and Moon Radiation Environment on Chandrayyan-1 Satellite, Adv. Space Res., 2011, 48, 5, and 779–91. http://dx.doi.org/10.1016/j.asr.2011.05.009.
13. Dachev, Ts., G. Horneck, D.-P. Häder, M. Lebert, P. Richter, M. Schuster, and R. Demets, Time profile of cosmic radiation exposure during the EXPOSE-E mission: the R3D instrument. Journal of Astrobiology, 2012, 12, 5, 403–11. http://eea.spaceflight.esa.int/attachments/spacestations/ID501800a9c26c2.pdf.
14. Dachev, Ts., G. Horneck, D.-P. Häder, M. Schuster, and M. Lebert, EXPOSE-R cosmic radiation time profile, Journal of Astrobiology, 2015, 14, 17–25. http://dx.doi.org/ 10.1017/S1473550414000093.
15. Dachev, T. P., Bankov, N. G., Tomov, B. T., Matviichuk, Y. N., Dimitrov, P. G., Häder, D.-P., and Horneck, G. Overview of the ISS radiation environment observed during the ESA EXPOSE-R2 mission in 2014–2016. Space Weather, 2017, 15, 1475–89. https://doi.org/10.1002/2016SW001580.
16. Uchihori, Y., H. Kitamura, K. Fujitaka, Ts. P. Dachev, B. T. Tomov, P. G. Dimitrov, and Y. Matviichuk, Analysis of the calibration results obtained with Liulin-4J spectrometer-dosimeter on protons and heavy ions, Radiation Measurements, 2002, 35, 127–34. http://dx.doi.org/10.1016/S1350-4487(01)00286-4
17. Dachev, Ts., Characterization of near Earth radiation environment by Liulin type instruments, Adv. Space Res., 2009, 44, 1441–49. DOI:10.1016/ j.asr.2009.08.007
18. Dachev, Ts., B. Tomov, Yu. Matviichuk, Pl. Dimitrov, N. Bankov, V. Shurshakov, O. Ivanova, D.-P. Häder, M. T. Schuster, G. Reitz, and G. Horneck, “BION-M” №1 spacecraft radiation environment as observed by the RD3-B3 radiometer-dosimeter in April-May 2013, Journal of Atmospheric and Solar-Terrestrial Physics. 2015, 123, 82–91. http://dx.doi.org/10.1016/j.jastp.2014.12.011

19. McIlwain, C. E. Coordinates for mapping the distribution of magnetically trapped particles. Journal of Geophysical Research, 1961, 66, 3681–91. https://doi.org/10.1029/JZ066i011p03681
20. Shea, M. A., Smart, D. F., Gentile, and L.C., The use of the McIlwain L-parameter to estimate cosmic ray vertical cutoff rigidities for different epochs of the geomagnetic field. In NASA. Goddard Space Flight Center 19th Intern. Cosmic Ray Conf. 1985, 5, 332–35. http://articles.adsabs. harvard.edu//full/1985icrc....5..332s/E000332.000.htmlSimpson
21. Case, A.W., J. C. Kasper, H. E. Spence, C. J. Zeitlin, M. D. Looper, M. J. Golightly, N. A. Schwadron, L. W. Townsend, J. E. Mazur, J. B. Blake, and Y. Iwata., The deep space galactic cosmic ray lineal energy spectrum at solar minimum, Space Weather, 2013, 11, 361–68. doi:10.1002/swe.20051
22. Cucinotta, F. A., and M. Durante. Cancer risk from exposure to galactic cosmic rays: Implications for space exploration by human beings, The Lancet Oncology, 7, 2006. doi:10.1016/S1470-2045(06)70695-
23. Zeitlin, C., Results from the MSL-RAD Experiment on the Curiosity Mars Rover, Nineteenth WRMISS, Krakow, Poland, 9–11 September 2014. http://wrmiss.org/workshops/nineteenth/Zeitlin_MSL-RAD.pdf
24. Chin, G., Brylow, S., Foote, M. et al., Lunar reconnaissance orbiter overview: The instrument suite and mission. Space Science Reviews, 2007, 129, 4, 391–419. https://doi.org/10.1007/s11214-007-9153-y

ОПИСАНИЕ НА СПЕКТРОМЕТЪРА НА ЗАРЕДЕНИ ЧАСТИЦИ “LIULIN TEN-KOH” ЗА ЯПОНСКИЯ СПЪТНИК “TEN-KOH”

Ц. Дачев, П. Димитров, Б. Томов, Ю. Матвийчук, П. Саганти, Д. Холанд, К.-И. Окуяма

Резюме

На 29 октомври 2018 г. в 13:08 часа по японско стандартно време успешно беше изстрелян в орбита спътника за наблюдение на парникови газове (GOSAT-2) от космическия център на JAXA Tanegashima. Заедно със спътника GOSAT-2 в кръгова (623 км) полярна синхронна орбита стартира и 22 килограмовият спътник Ten-Koh (http://kit-okuyama-lab.com/en/ten-koh/, разработен в Kyushu Institute of Technology от проф. К. Окуяма. Ten-Koh ще наблюдава околоземното пространство и основната цел е да се осигурят данни за бъдещо развитие на спътниците и на тяхната работа. Основният научен експеримент на спътника Ten-Koh е Детектор за заредени частици (ДЗЧ), разработен в университета Prairie View A&M, Тексас, САЩ и Джонсъновият космическия център на НАСА. Главен изследовател на експеримента е проф. П. Саганти (https://www.pvamu.edu/raise/space-payload/charged-particle-detector-2018). Отговорен инженер е Д. Холанд. ИКИТ-БАН

получи искане от проф. Саганти за разработване на инструмент от типа „Люлин“, който да бъде част от ДЗЧ. Учени от секцията по „Слънчево и земна физика“ на ИКИТ разработиха и предадоха на университета Prairie View A&M технологичен, летателен и резервен модели на инструмента “Liulin Ten-Koh”. Последният е подобен на прибора RADOM, който работи през 2008-2009 г. на индийския спътник на луната Chadrayaan-1. Тази статия описва прибора "Liulin Ten-Koh" и тестовете със стандартни радиационни източници, които са проведени по време на калибровките. Спътникът Ten-Koh е на полярна орбита на около 623 километра надморска височина и наклон на орбитата от 98°. Получени са стабилни телеметрични данни от няколко наземни станции по целия свят. В статията са представени и първите получени данни от инструмента "Liulin Ten-Koh Saganti" на космическия кораб Ten-Koh. Наличният до този момент интегрален LET спектър от прибора "Люлин Ten-Koh Saganti" е сравнен със спектри от други прибори, измерени във и извън земната магнитосфера.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e02

INVESTIGATION OF THE MOTION OF A SATELLITE, ACCORDING TO GENERAL THEORY OF RELATIVITY

Kostadin Sheyretski

University of National and World Economy
e-mail: sheyretski@unwe.bg

Keywords: *Asymptotic methods, Oscillation theory, General relativity, Nonlinear dynamics, Celestial mechanics*

Abstract

In this paper, we investigate the energy integral, which is obtained in the problem of the motion of a material point in a central field within the frame of the general theory of relativity. Applied is the method of the small parameter in a combination with the balance method. Derived is a compact formula, describing the trajectory of the motion. This formula gives a correct quantitative description of the basic relativistic effects. We prove the shortening of the major axis of the orbit compared with the case where we do not consider relativistic effects. This result can be useful for analysing the structure of planet systems around massive stars.

Introduction

The motion of a satellite in a central gravitational field is one of the milestone problems, whose solution imposes the use of the general theory of relativity. Its solution exactly defines the precession angle of Mercury, a problem, which had engaged the theorists for a long time until the arrival of Einstein's theory [1]. After this success, to a great extent the interest for this problem decreases as it is assumed that the major goals for investigating this problem have already been achieved. The fact that the problem is defined using a nonlinear differential equation gives an opportunity for work in this direction. On one hand, in the literature the problem is solved using sensible physical assumptions for the weak influence of a term in the differential equation on the solution, as the results of such an interpretation are being justified [1–2]. In this work we mathematically motivate the application of the small parameter method [3]. We use a method in which the analytical technique for describing the perturbed behaviour is applied to the energy integral. This was we directly define the link between the orbital parameters of the motion and we derive a compact formula for the solution. A result in such representation is that the major axis of the orbit is being shortened in comparison with the case when we do not take into account relativistic effects. In this work, we follow a mathematical framework set in the problem of Zelmanov and Agakov [2], but we use a different asymptotic method for finding the solution, which aims mathematically rigorous conclusions.

Mathematical Framework of the Problem

Consider the motion of a planet *T* in the radial field of a star *S.* We assume that the central object has a spherical shape and the planet can be approximated as a material point. Let note the distance between the two fields with *R,* and the polar angle with φ: Fig. 1.

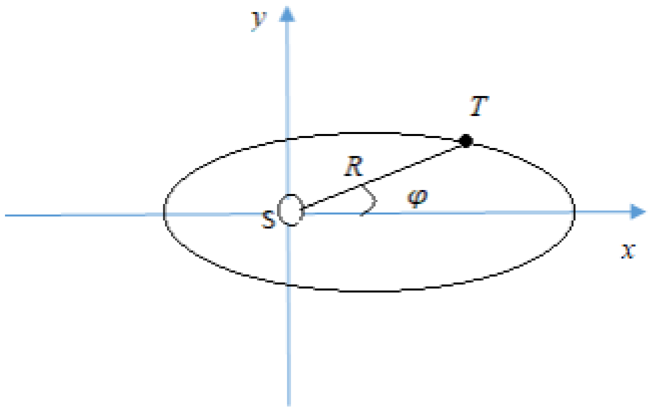


Fig. 1. The motion of a material point in a radial gravitational field

Let assume the mass of the material point to be unity, as well as being negligible compared to the mass μ of the body in whose field it is moving. Then, the differential equation of motion of the material point has the following form [2]:

(1) $$\frac{d^2u}{d\varphi^2} + u = \frac{\gamma\mu}{h^2} + 3\frac{\gamma\mu}{C^2}u^2.$$

In this formula: $u = \frac{1}{R}$, C is the speed of light in vacuum, γ is the universal gravitational constant, *h* is a constant defining the angular momentum, and φ is the angular parameter. We substitute

(2) $$\frac{1}{P} = \frac{\gamma\mu}{h^2}.$$

Changing the variables:

(3) $\xi = uP.$

We finally get

(4) $\frac{d^2\xi}{d\varphi^2} + \xi = 1 + 3\varepsilon\xi^2..$

The variable

(5) $\varepsilon = \frac{\gamma\mu}{PC^2}$

can be interpreted as a small parameter in our further analytical investigations.

Integrating Equation (4) we get

(5) $\left(\frac{d\xi}{d\varphi}\right)^2 + \xi^2 - 2\xi - 2\varepsilon\xi^3 = 2\epsilon, \epsilon = \frac{EP}{\gamma\mu},$

E is the energy of the system.

Finding and asymptotic solution

We are looking for a solution in a series [4]:

(6) $\xi = \xi_0 C_0 + \varepsilon\xi_1 + \varepsilon C_1,$

where ξ_0 and ξ_1 are functions of φ, C_0 and C_1 are constants. We make the guess that for the angular variable we have:

(7) $\varphi = \varphi_0 + \varepsilon\varphi_1.$

In further calculations we will use the formula:

(8) $\frac{d}{d\varphi} = \frac{d}{d\varphi_0} - \varepsilon\frac{d\varphi_1}{d\varphi_0}\frac{d}{d\varphi_0}.$

In addition, we write the integral constant as

(9) $\epsilon = \epsilon_0 + \varepsilon\epsilon_1.$

For Equation (5) to zeroth order we get

(10) $\left(\frac{d\xi_0}{d\varphi_0}\right)^2 + {\xi_0}^2 + {C_0}^2 + 2\xi_0 C_0 - 2\xi_0 - 2C_0 = 2\epsilon_0.$

We are looking for a solution in the form

(11) $\xi_0 = e\cos\varphi_0.$

It is easy to see that $C_0 = 1$. Then for the constant *e* we get

(12) $e^2 = 2\epsilon_0 + 1.$

The parameter *e* is the *eccentricity* of the orbit, and *P* – the *focal parameter*. For the orbit to be an ellipse, we should have the following condition: $-1 < 2\epsilon_0 < 0$.

For the first approximation of the equation we get:

(13) $\xi_1 = eB\cos\varphi_0.$

We look for ξ_1 in the form:

(14) $\xi_1 = eB\cos\varphi_0.$

After some calculations we equate the sum of all constants to be zero. The same procedure is carried out for the sum of the coefficients in front of the periodic functions $\cos\varphi_0$ and $\cos 2\varphi_0$. The calculations are carried out with precision up to $\cos 2\varphi_0$. Then for the parameters we get:

$$\epsilon_1 = -1,$$
$$\frac{d\varphi_1}{d\varphi_0} = 3,$$
$$B = 3,$$
$$C_1 = 3\left(1 + \frac{e^2}{4}\right).$$

Then, we substitute:

(15) $n = 1 - \frac{3\gamma\mu}{C^2 P}.$

The final equation for ξ is:

(16) $\xi = \frac{1}{n} + \frac{e}{n}\cos n\varphi + O\left(\frac{3\gamma\mu}{C^2 P} e^2\right).$

We can readily write the equation for the trajectory of the satellite:

(17) $R = \frac{\tilde{P}}{1+e\cos n\varphi}, \tilde{P} = Pn.$

From this formula it is clear, that the motion on the orbit includes precession but we can also note the shortening of the major axes of the orbits of space objects, which does not depend on the distance between the planet and the central object, but only on the gravity radius of the central object: Table 1.

Table 1. Relativistic effects

Precession: $\Delta\omega$	Shortening of the major axis of the elliptical trajectory: ΔP
$\frac{6\pi\gamma\mu}{C^2 P}$	$\frac{3\gamma\mu}{C^2}$
Shortening of the major axes of the orbits for all planets in the Solar system	4.44 10^3 m

The value we find for the precession coincides with the value derived analytically using other methods.

We can easily find the equation for energy, including the orbital elements [5]:

(18) $E = -\frac{\gamma\mu}{2a} - \frac{h^2}{(1-e^2)^3 C^2}\frac{\gamma\mu}{a^3}, P = a(1-e^2).$

Conclusion

The nonlinear character of the problem at hand supposes to obtain asymptotic solutions which can be mathematically devised differently. Although all such solutions describe the motion in the same manner, some conclusions can be made to depend on the structures of the equations obtained [6–7]. In this work it is shown that adopting the method presented we obtain a formula, which reflects not only the angle of precession but also the shortening of the major axis of the orbit, in comparison with calculations based on Newtonian mechanics. For massive stars such shortening would be of greater importance for the existence of planet systems.

References

1. Utiyama, R. General Relativity. Atomizdat, Moscow, 1979.
2. Zelmanov, A., V. Agakov. Elements of General Relativity. Nauka, Moscow, 1989.
3. Hayashi, C. Nonlinear oscillations in physical systems. McGraw-Hill book company, New York, 1964.
4. Naifehn, A. Introduction to Perturbation Techniques. John Wiley & Sons, New York, 1981.
5. Lichtenberg, A., M. Liberman. Regular and Stochastic Motion. Mir, Moscow, 1984.
6. Zaslavskii, G., R. Sagdeev. Introduction to Nonlinear Physics. Nauka, Moscow, 1988.
7. Moiseev, N. Asymptotic methods in nonlinear mechanics. Nauka, Moscow, 1969.

ИЗСЛЕДВАНЕ ДВИЖЕНИЕТО НА СПЪТНИК, СЪГЛАСНО ОБЩАТА ТЕОРИЯ НА ОТНОСИТЕЛНОСТТА

Костадин Шейретски

Резюме

В статията се изследва интеграла на енергията, който се получава в задачата за движение на материална точка в централно поле спрямо Обща теория на относителността. Приложени са метод на малкия параметър в комбинация с метод на хармоничния баланс. Изведена е компактна формула, описваща траекторията на движение. Формулата дава правилно количествено описание на основните релативистки ефекти. Доказано е скъсяване на главната полуос на отрбитата, в сравнение със случая когато не се отчитат релати-вистките ефекти. Този резултат може да бъде полезен при анализиране структурата на планетни системи, образувани около масивни звезди.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e03

ULTRAVIOLET RADIATION LEVELS OVER BULGARIAN HIGH MOUNTAINS

Rolf Werner[1], Veneta Guineva[1], Atanas Atanassov[1], Dimitar Valev[1], Dimitar Danov[1], Boyan Petkov[2], Andrey Kirillov[3,4]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences
[2]Institute of Atmospheric Sciences and Climate, CNR, Bologna, Italy
[3]Polar Geophysical Institute, Apatity, Russia
[4]ESRI Bulgaria Ltd.
e-mail: rolwer52l@yahoo.co.uk

Keywords: *Ultraviolet radiation, UV-index, Total ozone content*

Abstract

The UV-index (UVI) is a measure of the erythemally effective solar radiation reaching the Earth surface and it was introduced to alert people about the need for Sun protection. This study applies a model that estimates the UVI over the high Bulgarian mountains for clear sky conditions considering the Total Ozone Content (TOC), which was taken from satellite measurements. The results show that during the periods from May to August at altitudes above 2 000 m a.s.l. high UVI's (greater than 8) were observed for more than 18 days per month. The UVI values were high for every day of July at altitudes higher than 1 500 m. Extremely high UVI result from episodes with TOC lower than 290 DU during June and July at the highest mountain parts with elevations greater than 2 500 m. High radiation risks were observed during April, especially when the preceding polar vortex was strong and the mountains were snow covered.

Introduction

The solar ultraviolet (UV) radiation affects the Earth's atmosphere by heating, exciting, dissociating and ionising its main constituents. UV-rays are important for its impact on Earth's surface, atmosphere, ionosphere and exosphere [1]. The ionosphere is created by the most penetrating radiations, which are very sensitive to solar activity. An analysis of the conditions of absorption of solar radiation shows that the UV lines Lyman-Alpha (Ly-α), Lyman-Beta (Ly-β) and UV continuum, along with X-rays and cosmic rays, are factors for the formation of ionosphere of the Earth, planets and their satellites in the solar system [1–4].

Ultraviolet irradiance reaching the Earth surface impacts the human's health (induces erythema, cataract, provokes skin cancer, DNA damages and others). The UV-index (UVI) was developed from Canadian scientists in 1992 as a simple measure of the risk from unprotected sun exposure. It represents a dimensionless measure of the erythemally effective solar radiation reaching the Earth surface, defined as an integral over the radiation reaching the Earth surface at a certain time, expressed in W/(m^2*nm), weighted by the erythemal action spectrum and multiplied with 40 m^2/W (see e.g. WHO [5]).

The UVI can be determined by highly precise spectrometric measurements, by multichannel filter instruments or by UV photometers of solar irradiance both, both ground-based and satellite born [6, 7].

In many countries, local UVI monitoring networks were established. In Europe, currently 160 stations in 25 countries deliver online values to the public *via* the Internet [8]. UVI daily forecast maps for Bulgaria and forecasts for some stations based on multi-yearly averages were constructed and empirical models were developed in National Institute of Geophysics, Geodesy and Geography-BAS [9] (http://data.niggg.bas.bg/uv_index/uv_index_bg.php).

To study the day-to-day variations of the harmful solar irradiance, the UVI is usually determined for clear sky conditions at solar noon, when the risk of harmful sunburn is the highest. Under clear weather the UVI depends strongly on TOC and varies with the surface albedo and with the sun elevation change. The main goal of the presented here paper is to estimate the UVI for a multi-year period in the Bulgarian high mountains.

Data used and description

For estimations of the ozone over the Bulgarian high mountain region, long-time TM3DAM-OMI overpass data for Sofia and Thessaloniki were used (http://temis.nl/protocols/o3field/overpass_omi.html). The data are provided in ASCII format. We have used the ozone data for Sofia and Thessaloniki at noon. The time series are almost gapless, and the data are available for the time interval from 10.01.2004 up to now.

Stratospheric ozone

Stratospheric ozone is produced by solar ultraviolet irradiance reaching the Earth atmosphere mainly in the tropics during summer. However, a significant amount of ozone is generated at mid-latitudes as well [10]. Ozone is transported poleward by the Brewer-Dobson circulation forced by temperature difference over the tropics and the polar atmosphere. The meridional temperature gradients and hence the dynamical processes are the strongest in winter. At mid-latitudes the stratospheric ozone has maximal values during spring. In each winter, the polar vortex is built up with very low temperatures inside it. By heterogenic reaction ozone can be destroyed catalytically in spring. In the Northern Hemisphere, strong planetary waves induced by the orography frequently disturb the polar vortex. In addition, other atmospheric circulations, short-term and mid-term variations up to a few months (e.g., advection and upwelling of the air masses) contribute to the ozone variability.

After strong Vulcan eruptions TOC increases are observed worldwide. However, the causes of the increase are complex [11] and the net effect on UVI must be carefully investigated.

Ultraviolet radiation

The energy emitted by the Sun that arrives at the Earth is about 1361 W/m^2 and is almost constant (the solar constant) slightly varying by about 0.1% during the solar cycles [12, 13]. In the near UV (300–400 nm), the variations are somewhat greater (probably they are of the order of 0.5%–1%). Today, short-term UV variations are the object of intensive research. As the Total Solar Irradiance (TSI), the near UV radiation decreases with the increase in the Sun-Earth distance. Passing through the Earth atmosphere the UV radiation is strongly absorbed by ozone and scattered by air molecules and aerosols. The received UV radiation at a horizontal Earth surface varies with the local zenith angle and hence with the day time, season, and geographic location. For the public, the UVI forecast is usually given as its maximum over the day (at solar noon) for cloudless weather taking in such a way into account only the absorption of UV irradiance by ozone. However, the UVI depends on the albedo of the underlying surface, in particular in the case of ice and snow. At mid-latitudes, the UVI mainly varies between 1 and 10 and the radiation risk of UVI values exceeding 8 is considered as very high, while for values greater than 11 it is extremely high.

Method for investigation

The UVI variations caused by the ozone deviations from the seasonal mean for the Bulgarian high mountains will be the subject of the paper. A part of the Bulgarian mountains higher than 2500 m is located between Sofia (42.817°N, 23.383°E) and Thessaloniki (40.520°N, 22.970°E) stations at latitudes from about 41.6°N to 42.2° N. For describing the ozone series variations over the Bulgarian high mountains the average values from the Sofia and Thessaloniki time series were calculated with a weight of 0.6 for the first station and 0.4 for the second one. These weights correspond to the distances of the mountains from Sofia and Thessaloniki, respectively. The seasonal multi-year means of the resulting series were determined by a Fourier series of second order with the basic period of one year, consisting of 365.25 days [14]. The time series is shown in Fig. 1 together with the sum of its trend and seasonal components in a wide range. The daily ozone values vary in a wide range from about −70 DU up to 100 DU around the seasonal mean with the correspondingly weaker or stronger solar UV absorption. On the assumption, that about 10% of TOC is contained in the troposphere and ozone in this layer is nearly uniformly distributed, the ozone values w ere recalculated to the sea level. Taking the TOC values for clear sky, UVI at noon were determined by applying the fast empirical algorithm developed by Allaart et al. [15], where besides the TOC Two more parameters are necessary: the solar zenith angle (*SZA*) and the Sun-Earth distance (*D*).

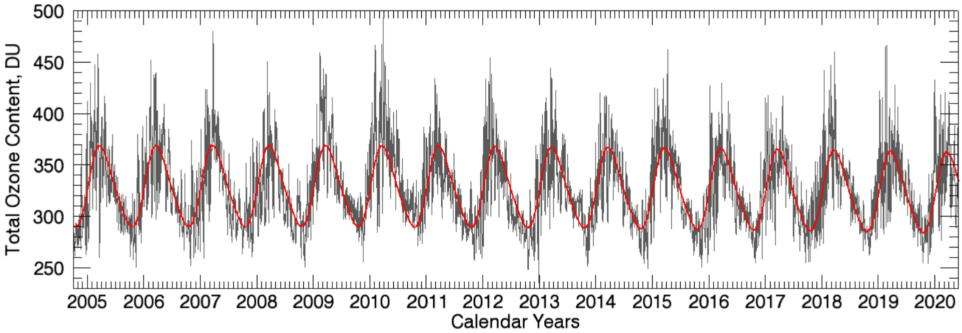


Fig. 1. TOC time series for the Bulgarian High Mountain region recalculated to sea level by using weighted OMI overpass ozone data for Sofia and Thessaloniki. The red line shows the sum of the trend and seasonal components.

$$UVI = \left[\left(\frac{D_0}{D}\right)^2 \cdot S_0 \cdot \mu_x \cdot exp\left(-\frac{\tau}{\mu_x}\right)\right] \cdot \left[F \cdot X^G + \frac{H}{TOC} + J\right]$$

(1) $$S = 1.24\,Wm^{-2}nm^{-1} \quad \tau = 0.58 \quad \mu_x = \mu_0 \cdot (1-\varepsilon) \cdot \varepsilon$$

$$\mu_0 = \cos(SZA) \qquad \varepsilon = 0.17 \quad F = 2.0 \qquad X = 1000 \cdot \frac{\mu_0}{TOC}$$

$$G = 1.62 \quad H = 280.0 \quad J = 1.4 \quad .$$

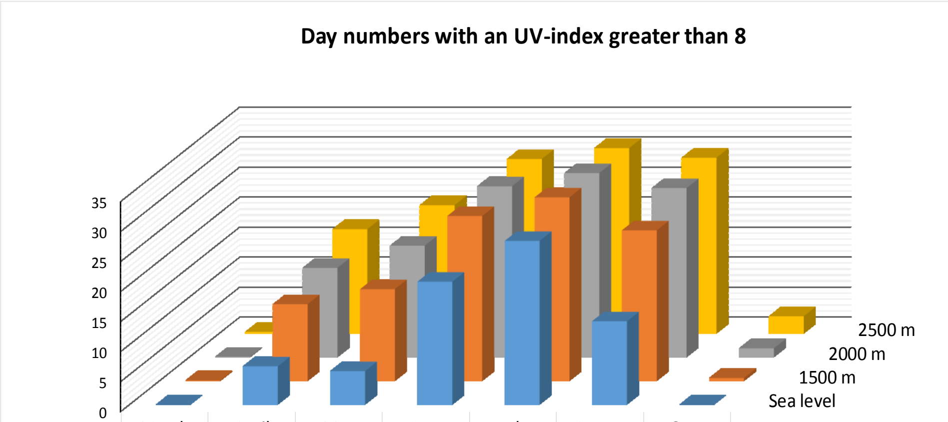
SZA and *D* were calculated from the implementation of an astronomical algorithm [16].

Using the TOC at sea level in the empirical model the UVI for the elevation h=0 m is obtained. By help of the Tropospheric Ultraviolet and Visible (TUV) radiative model [17] we found an UVI increase of 6% per km with elevation increasing, which is in very good agreement with the value of 6%–8% per km given in [18,19].

The UVI behaviour was studied for the March-April and May-September (hereinafter referred to as spring and summer, respectively). During the spring the polar vortex is already developed. At the same time, the *SZA* becomes low and the High Mountains are usually snow covered. The UVI for snow covered surfaces were corrected by a factor of 1.25 corresponding to a mean snow albedo of 0.8. [20, 21]. During the summer the Sun culmination is high that corresponds to elevated irradiation risk. The obtained UVI time series allows estimation of the number of days with very high UVI (greater than 8) in the time intervals under study.

Results Satellite data processing

The UVI values corresponding to the TOC time series determined above were calculated for all days of the time interval under study and the monthly means were determined for four altitudes – sea level, 1500 m, 2 000 m and 2 500 m. The results are presented as column chart and as table in Fig. 2. (For the days of spring the UVI's were snow corrected.) During the period from April to August in the high mountains (above 2 000 m altitude) more than 18 days per month with very high UVI'-s were observed. At altitude higher than 1 500 m in practice all the days of July are characterized by very high values. However, the monthly variations in the number of days with UVI greater than 8 from year to year are of the order of 30% (not shown here), especially in spring. The lowest number of days with UVI > 8 was observed in the winters 2006/2007 and 2008/2009, when the polar vortex was weak and the TOC exceeded the average values in March – April with some exceptions (Fig. 3).



	March	April	May	June	July	August	Sep-tember
Sea level	0	6,4	5,6	20,5	27,3	13,9	0
1500 m	0,1	12,8	15,3	27,5	30,6	25,1	0,5
2000 m	0,1	14,9	18,6	28,5	30,7	28,2	1,5
2500 m	0,3	17,4	21,4	29,1	30,9	29,3	2,9

Fig. 2. Mean monthly number of days with an UVI greater than 8 for some altitudes during the period 2005–2020 in the Bulgarian high mountain region

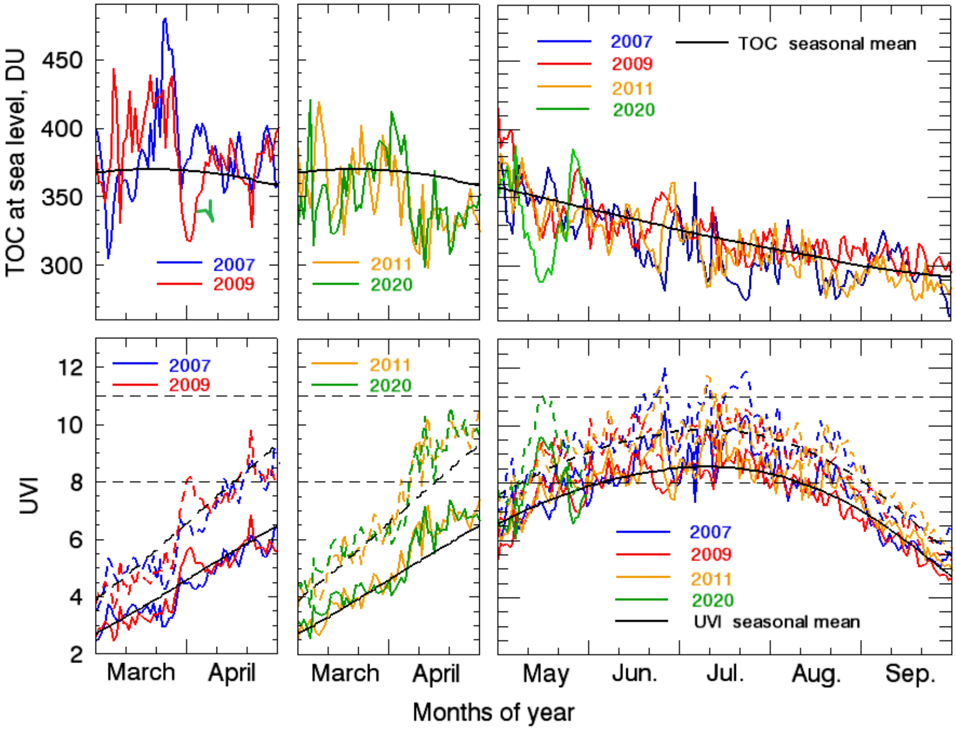


Fig. 3. At the top row TOC series are shown. At the bottom row UVI series are presented calculated from the TOC. The solid lines are for the sea level, and the dashed ones are for the altitude of 2 500 m. The UVI at 2 500 m is snow corrected for the time periods of March – April. The horizontal black dashed lines represent the UVI = 8 and UVI = 11 limits.

The observed lower TOC values in March had not a significant impact on UVI because of the low Sun culmination. However, at the end of March/April 2009 low TOC caused two UVI peaks higher than eight, but only for some days. For 2011 and 2020, when the vortexes were strong, air masses with low TOC values were located over the snow covered Bulgarian high mountains in April, that lead to continuously very high UVI values during the whole period with peak values of more than 10. For the summer period the TOC values did not show annual differences, depending on the vortex strength in winter/spring. An exception was the vortex 2019/2020, which persisted up to the beginning of May – the latest persistence up to now. Low ozone events caused by strong planetary wave activities were observed up to the mid of May. During the same time a part of the tropospheric jet stream transported Sahara dust, which reduced the surface UV radiation importantly.

On the other hand, the highest parts of the mountains with an elevation greater than 2 500 m were snow covered during the mid of May 2020. In period from May to August the UVI was very high. For some days in June and July when the Sun culmination is highest the TOC was relatively low (lower than 290 DU) that lead to extremely high UVI (greater than 11) in the mountains at altitudes higher than 2 500 m. At the same time, the UVI at sea level was very high from June to July, as well.

Summary and conclusions

Using the clear sky UVI as a measure of the harmful UV radiation, in days with high UVI values one could choose appropriate clothes and means with higher protection. Clouds absorb UV radiation and the real UVI is usually lower than that of clear sky. However, in some cases, the UV can be higher due to reflections from clouds. To take into consideration the fast changing cloudiness for the prognosis of UVI, especially in mountains, requires complex weather forecast models. Thus, the easiest way that provides a realistic assessment is the use of a clear sky UVI.

Based on the multiyear ozone values, the clear sky UVI's were estimated and corrected for altitude and snow albedo. The paper reports results about the UVI in the high mountain regions in Bulgaria. It was shown, that in the mountains, due to the altitude and reflections by snow, the UVI increases to very high levels at the end of spring, particularly in years when the polar vortex was very strong. A very high radiation risk in the Bulgarian high mountains was found in summer as well.

The highest UVI were observed during June and July, as it was expected. At 2 500 m elevation in the mountains, the UVI is higher than the UVI at sea level by a factor of about 1.15. In limited cases when the TOC is below 290 DU in summer, extremely high radiation risk (UVI > 11) can be achieved. But very high UVI levels in the high mountains can also be occurred in spring, when the TOC values are below the mean in April and the terrains are snow covered.

We would like to remind, that the World Meteorological Organisation recommends staying indoor during midday hours, when the radiation risk is very high (UVI > 8). Since above 2 500 m, the alpine grasslands dominate, the probability to find shade at these altitudes for minimising the radiation risk is very limited, both at the end of spring and in summer. When there are very high radiation risk levels during the summer and winter seasons in the high mountain regions, people has to be alarmed in their accommodations or online about the risk.

The results obtained in this work will help to improve the environmental and solar-terrestrial models and provide the input for planetary modelling of helio-biological and space weather processes in quiet and disturbed conditions [22].

References

1. Serafimov, K., Physics of Middle Ionosphere, Sofia, BAS Publishers, 1970, 275 p.
2. Usoskin, I., L. Desorgher, P. Velinov et al. Ionization of the Earth's atmosphere by solar and galactic cosmic rays, Acta Geophysica, 2009, 57, 1, 88–101. DOI:10.2478/s11600-008-0019-9.
3. Velinov, P. I. Y. and L. Mateev, Ionization of cosmic rays and particles in the ionosphere and atmosphere of Mars, C. R. Acad. Bulg. Sci., 1991, 44, 1, 31–34.
4. Gronoff, G., C. Mertens, J. Lilensten, et al., Ionization processes in the atmosphere of Titan, Astron. & Astroph., 2011, 529, 5, A143–A146, DOI:10.1051/0004-6361/201015675.
5. Global Solar UV Index: A Practical Guide, WHO, Geneva, Switzerland, 2002, 28 p.
6. Dahlback, A., Measurements of biological effective UV doses, total ozone abundances, and cloud effects with multichannel, moderate bandwidth filter, Appl. Opt., 1996, 30, 33, 6514–21, DOI:10.1364/AO.35.006514.
7. Fioletov, V. E., J. B. Kerr, L. J. B. McArthur, D. I. Wardle, and T. W. Mathews, Estimating UV index climatology over Canada, J. Appl. Meteorol., 2003, 42, 417–33.
8. Schmalwieser, A., J. Gröbner, M. Blumthaler, B. Klotz, H. de Backer et al., UV Index monitoring in Europe. Photochemical and Photobiological Science, 2017, 16, 1349–70, DOI:10.1039/C7PP00178A.
9. Bojilova, R., P. Mukhtarov, N. Miloshev, Climatology of the index of the biologically active ultraviolet radiation for Sofia. An empirical forecast model from prediction the UV-Index, Compt. Rend. Acad. Bulg. Sci., 2020, 73, 4, 531–38. DOI:10.7546/CRABS.2020.04.12.
10. Grewe, V., Impact of climate variability on tropospheric ozone, Science of the Environment, 2007, 374, 167–81. DOI:10.1016/ j.scitotenv.2007.01.032.
11. Komitov, B., K. Stoychev, Stratospherizone, solar activity and volcanism, Bulg. Astron. J., 2011, 17, 118–25.
12. Kopp, G., J. L. Lean, A new, lower value of total solar irradiance: Evidence and climate significance, Geophysical Research Letters, 2011, 38, L01706, DOI: 10.1029/2010GL045777.
13. DeLand M. T. and R. P. Cebula, Solar UV Variations During the Decline of Cycle 23, 2011, https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20110023422.pdf
14. Kaleyna, P., P. Mukhtarov, and N. Miloshev, Seasonal variations of the total column ozone over Bulgaria in the period 1996-2012, Comptes rendus de l'Académie bulgare des Sciences, 2014, 67, 7, 979–86.
15. Allaart, M., M. Van Weele, P. Fortuin, H. Kelder, An empirical model to predict the UV-index based on solar zenith angles and total ozone, Meteorol. Appl., 2004, 11, 59–65. DOI:10.1017/S1350482703001130
16. Meeus, J., Astronomical Algorithms, 2nd ed., Va.,Willmann-Bell, 1998, 477 p.
17. Madronich, S. UV radiation in the natural and perturbed atmosphere, in Environmental Effects of UV Radiation, in M. Tevini, Ed., UV-B Radiation and Ozone Depletion, Lewis Publishers, London, 1993, pp. 17–69.

18. Vanicek, K., T. Frei, Z. Litynska, and A. Schmalwieser, UV-Index for Public, A guide for publication and interpretation of solar UV Index forecasts for the public, prepared by the Worging Group 4 of the COST-713 Action “UVB Forecasting”, Brüssel, 1999, 27 p.
19. Dahlback, A., N. Gelsor, J. J. Stamnes, Y. Gjessing, UV measurements in the 3000–5000 m altitude region in Tibet. J. Geophys. Res., 2007, 112, D09, DOI:10.1029/2006JD007700.
20. Schmalwieser, A.W. and G. Schauberger, Validation of the Austrian forecast model for solar, biologically effective UV radiation-UV index for Vienna, J. Geophys. Res., 2000, 105, D21, 26 661-667, DOI:10.1029/2000JD900135.
21. Werner, R., B. Petkov, D. Valev, V. Guineva, A. Atanassov, D. Danov, and A. Kirillov. Forecast scheme of the local UV-Index over Bulgaria – First test results, Proceedings of the Fifteenth International Scientific Conference “Space Ecology Safety”, Ed. Petar Getsov and Garo Mardirossian, Sapce Research Institute and technology - BAS, Sofia, 2019, pp. 46–51.
22. Tsagouri, I., A. Belehaki, N. Bergeot, C. Cid, V. Delouille et al. Progress in space weather modelling in an operational environment. J. Space Weath. & Space Clim., 2013, 3, A17, 1–72. DOI:10.1051/swsc/2013037.

НИВА НА УЛТРАВИОЛЕТОВА РАДИАЦИЯ НАД ВИСОКИТЕ ПЛАНИНИ В БЪЛГАРИЯ

Р. Вернер, В. Гинева, А. Атанасов, Д. Вълев, Д. Данов, Б. Петков, А. Кирилов

Резюме

Ултравиолетовият индекс (UV-индекс, UVI) е мярка за еритемално ефективната слънчева, достигаща земната повърхност, и той беше въведен за да предупреждава хората за необходимостта за предпазване от слънцето. В настоящото изследване се прилага модел, който оценява UV-индекса над високите планини в България при условия на чисто небе, като се отчита общото съдържание на озон (TOC), получено от спътникови измервания. Резултатите показват, че през периодите от май до август при височини над 2 000 m са наблюдавани много високи UV-индекси (по-високи от 8) за повече от 18 дни на месец. На практика стойностите на UV-индекса са много високи за всеки ден на юли за височини над 1500 m. Екстремно високите UV-индекси са резултат от случаи на TOC по-ниско от 290 DU през юни и юли, и се отнасят за най-високите части на планините, над 2 500 m. Висок риск от слънчева радиация е наблюдаван през април, особено когато предшестващият полярен вихър е силен и планините са покрити със сняг.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e04

CROP TYPE MAPPING IN BULGARIA USING SENTINEL-1/2 DATA

Petar Dimitrov, Lachezar Filchev, Eugenia Roumenina, Georgi Jelev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: petar.dimitrov@space.bas.bg; lachezarhf@space.bas.bg;
roumenina@space.bas.bg; gjelev@space.bas.bg

Keywords: *Crop Type Mapping, Sentinel, Google Earth Engine, Random Forest*

Abstract

Advanced possibilities have emerged in recent years for semi-automatic crop type mapping at the national level due to the availability of Sentinel-1 and -2 satellite data. In this study, 14 crop type classes were mapped over Bulgaria using three bi-*monthly composite image mosaics for 2019 generated in the Google Earth Engine (GEE) cloud computing platform. The overall accuracy, when both Sentinel-1 and -2 mosaics were used, was 78%, while the accuracy was slightly less when only Sentinel-2 data were used (75%). The accuracy was highest for "Cereals", "Maize", "Sunflower", "Winter rapeseed", and "Rice" – over 80% for both user's and producer's. However, the accuracy for classes such as "Vegetables", "Technical crops", "Forage crops", "Fallow", etc. was low. These classes represent categories suitable for agricultural practice and statistics, but are too general and difficult to distinguish using satellite data. It was also found that accuracy tends to be higher for larger parcels. Using composites with higher frequency and adapting the legend classes to include only crops similar in phenology and morphology are suggested as possible ways forward.*

Introduction

The potential of Sentinel-2 for crop type mapping has been demonstrated in recent years by numerous studies. The high temporal resolution (5 days when both Sentinel-2 A and B satellites are used) is one of the key characteristics of the Sentinel-2 imagery, which makes it particularly useful for crop mapping because it provides multiple snapshots of crop development during the growing season. These benefits are clearly demonstrated by the multi-date approach, where (all) available cloud-free images during the season are used for classification, e.g. [1, 2]. While this approach is relevant for relatively small study areas, national scale or large area applications, e.g. [3–5], should deal with images from different orbits (thus different date), large volumes of data, and the cloud cover. Cloud storage and computing facilities, such as Google Earth Engine (GEE) [6] facilitate significantly such applications. Additionally, Sentinel-1 Synthetic Aperture Radar (SAR) data have also been used to map crop types [7, 8]. Van Tricht et al. [3] found that combining radar and optical data for crop classification led to increasing classification accuracies compared to optical-only classification.

A previous case study from Bulgaria [9] demonstrated the utility of Sentinel-2 imagery for crop type mapping in small regions using selected cloud-free scenes. The authors suggested that future studies should address the problems related to mapping at national scale and also to integrate Sentinel-1 data in the classification in an attempt to increase the accuracy. This study, therefore, tries to build on past results and its aim is to produce and evaluate a national scale crop type map of Bulgaria based on Sentinel-1 and Sentinel-2 imagery. The results are also compared with those obtained using only Sentinel-2 imagery. Finally, the map accuracy is analysed with respect to the field size.

Data and methods

The CORINE Land Cover (CLC) 2018 dataset is used to define the area of interest, which includes only agricultural areas. Thus, only the regions with code 200 according to the CLC level 1 are further considered for the image classification.

Based on the agroclimatic zoning of Bulgaria [10] the country is divided into four agroclimatic regions: 1. cool and moderately cool, wet region (mountain areas); 2. Moderately warm and warm region, less liable to droughts (most of the Danube plain and the basins and low mountain parts in southern Bulgaria); 3. moderately hot and hot region, liable to droughts (the northernmost part of the Danube plain, the Upper Thracian lowland, and lowland of Burgas); and 4. Hot, arid region (the lower part of Struma valley). The data pre-processing and classification are repeated for each region and the final crop map is obtained after merging the maps of the individual regions.

Data about parcel borders and the crop sown in each parcel in 2018/2019 agricultural year are available in a vector format from State Fund "Agriculture" (SFA). The data are based on declarations made by farmers who apply for aid under Common Agricultural Policy (CAP) and national programmes and have complete coverage of the country's agricultural area. These data are collected as part of the Integrated Administration and Control System (IACS). A special nomenclature of crops is used in this dataset, which, at the lowest level, includes more than 200 crops, which are aggregated in groups (e.g. technical crops) and subgroups (e.g. industrial crops, oil crops, etc.) at the higher levels. For this study, the crops were aggregated in a customised legend, including some important individual crops and some wider classes based on the groups and subgroups of the original nomenclature. The 14 classes are as follows: "Cereals", "Maize", "Grain legumes", "Technical crops", "Sunflower", "Winter rapeseed", "Forage crops", "Meadows and pastures", "Alfalfa", "Vegetables", "Fallow", "Vineyards and orchards", "Perennial medical and aromatic crops", and "Rice". The legend is constructed to be maximally close to the nomenclature used by the state authorities in the agricultural sector. Fig. 1 presents the major stages in the phenological cycle of some crops and crop types. The parcels are divided by random in two parts: for training and for the validation of the classification algorithm (70:30), and are converted to raster format.

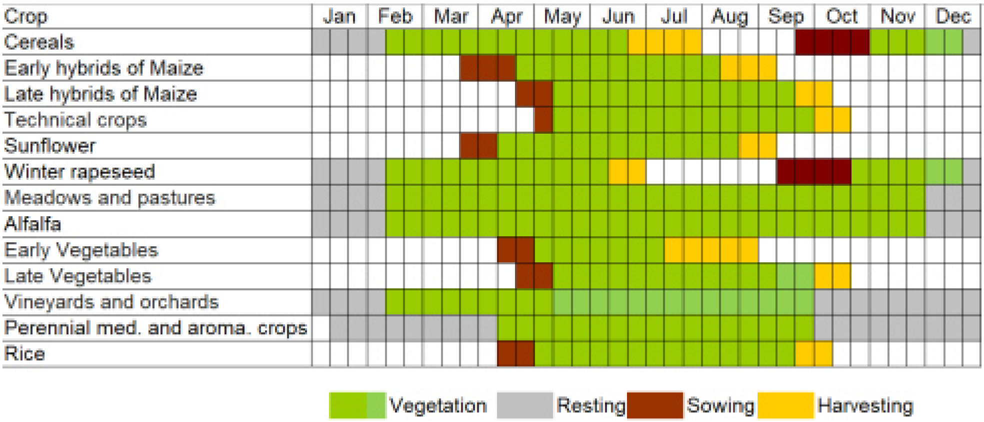


Fig. 1. Crop calendar of some crops and crop types in Bulgaria for 2019

The next steps, including satellite image pre-processing, training of the classifier, and classification, are performed in Google Earth Engine (GEE). The following GEE collections are used: “COPERNICUS/S2” consisting of Sentinel-2 A&B scenes at level 1C, and “COPERNICUS/S1_GRD” consisting of Sentinel-1 Ground Range Detected (GRD) scenes. The pre-processing steps for the Sentinel-2 imagery include 1. Selection of scenes with cloud cover lower than 20%; 2. Applying the cloud and cirrus masks, which are part of the dataset (band QA60); 3. Generating three multiband temporal composite images using the median compositing rule, each containing bands B02-B08, B11, and B12: March-April, June-July, and August-September 2019. The pre-processing steps for the Sentinel-1 imagery include 1. Filter the scenes by orbit type and selecting only “ascending” imagery; 2. Clipping the edge of the scenes to remove bad pixels (an inland buffer); 3. Apply a function (provided by Kristof Van Tricht, VITO) to make sure all acquisitions in one pixel result from the same relative orbit; 4. Generating three multiband temporal composite images using the median compositing rule, each containing VV and VH polarisations: March-April, June-July, and August-September 2019. Note that May is omitted from the compositing periods due to the frequent cloud cover this month. Two datasets were constructed from the imagery. The first has 27 bands and includes the Sentinal-2 composites. The second has 33 bands and includes both Sentinal-2 and Sentinel-1 composites.

The Random Forest (RF) classification algorithm [11, 12] as implemented in GEE is used to classify the satellite image datasets. The raster with the parcels designated for training is imported in GEE and a stratified random sampling is performed with 1000 pixels per class (note that in some of the agroclimatic regions this number cannot be attained for some classes, which have limited distribution). The values of the image bands are extracted for each training pixel. These data are then used to train the RF classifier. The number of trees is set up to 100, which is considered good compromise between accuracy and computational time [13]. All other parameter values are left by default. The final map is exported from GEE in GeoTiff format with 10 m pixel size.

The final crop map obtained after the four agroclimatic regions have been merged is "smoothed" by eliminating patches smaller than 10 pixels. This is performed using the Sieve tool of QGIS. Accuracy assessment is also performed in QGIS. For that purpose, the raster with the parcels designated for validation and the crop map raster are compared pixel by pixel and a confusion matrix is generated. Over 100 million pixels are used for this validation. Overall accuracy and class-wise accuracies (User's and Producer's) are calculated. Additionally, We repeated the same validation procedure several times but using only parcels with specific size: less than 0.5 ha, 0.5-1ha, 1–3 ha, 3–5 ha, and over 5ha.

Results and discussion

The overall accuracy of the map based solely on Sentinel-2 data is 74.8%, while the overall accuracy of the map based on a combination of Sentinel-1 and -2 data is 78.1%. This confirms the added value of SAR data in crop type mapping. All results and discussions further on concern the map based on the combination of Sentinel-1 and -2 data which is shown in Fig. 2. A visual examination of the map shows that the agricultural land use pattern is well portrayed in most of the territory where large parcel sizes dominate. For example, Fig. 3A shows a map excerpt representing a small area near the town of Knezha in the Danube plain. Here, parcel borders and shapes are realistically represented and within-field heterogeneities caused by errors in the classification are rate. More importantly, in most parcels the crop type is accurately determined by the RF classifier if we compare it with the IACS dataset used in this study as a reference. In other parts of the country, however, the classification results are characterised with much noise. A typical example is to be found in the Upper Thracian lowland near Plovdiv, where parcel sizes are much smaller (Fig. 3B). The post-processing (i.e. the smoothing with the "sieve" tool) reduces noise but due to small parcel size, it resulted in disturbance of the parcels' shape. Also, compliance with the IACS dataset is poorer.

The performance is not constant among classes and the accuracy varies for the different crop types. Both user's and producer's accuracy are above 80% for the classes “Cereals”, “Maize”, “Sunflower”, “Winter rapeseed”, and “Rice” (Fig. 4). Cereals, (which include mostly winter wheat and winter barley), maize, and sunflower represent the most important crops in the country in terms of area. Rice is particularly well classified, which is due to its specific method of cultivation. The class “Meadows and pastures” is mapped with moderate accuracy (70% and 75% for producer’s and user’s accuracy respectively; Fig. 4). The accuracy for the other classes is lower. In particular, their user’s accuracies are low, which indicates that their occurrence is overestimated. For example, most of the pixels belonging to class “Vegetables” in the map are actually other crop types. The classes for which the RF classification has low accuracy are rare classes, which mean they represent a small part of the arable land in Bulgaria. This can be seen in the area distribution shown in Fig. 5.

The most important misclassifications are as follows: 1. “Alfalfa” is overestimated at the expense of “Cereals” and “Meadows and pastures”; 2. “Technical crops” and “Grain legumes” are overestimated at the expense of “Sunflower” and “Cereals”; 3. “Forage crops” is overestimated at the expense of “Maize” and “Cereals”; 4. “Vegetables” is overestimated at the expense of “Sunflower” and “Fallow”; 5. “Vineyards and orchards” is overestimated at the expense of “Meadows and pastures”; 6. “Perennial medical and aromatic crops” and “Fallow” are mixed with many of the other classes. Most of the mixtures are with “Cereals”, “Sunflower”, and “Maize”, which can partially be explained by the fact that these are the most widespread classes. The similarity of classes in terms of crop phenology and/or physiognomy also plays a part. For example, “Alfalfa” is mixed with “Meadows and pastures”, both classes representing low herbaceous plants with continuous cover and similar phenological cycle (Fig. 1). Another reason for the errors in the classification is that some classes are too general and include crops which are not similar in their spectral characteristics but in their usage. For example, the “Forage crops” class includes, among others, crops as different as clover and corn for silage. This can partly explain the mixture with the “Maize” class. Fig. 6 shows the overall accuracy calculated for different parcel sizes. As the visual inspection of the map suggested the parcel size is related to the accuracy. The accuracy increased from below 60% for the smallest parcels to over 80% for those larger than 5 ha. While the smallest parcels (<0.5 ha) are the most numerous, they account for only 3% of the area of all parcels designated for validation. The largest parcel category (> 5 ha) constitutes by far the largest area (76 %). These results can be explained with the fact that smaller fields have more border pixels, which represent a mixture of land uses.

The application of the national crop map based on Sentinel data could be the calculation of areas of different crops for statistical purposes. To check the accuracy of the calculated areas they are compared with the areas from the IACS dataset (Fig. 5). To guarantee that the areas are comparable the Sentinel-based crop map is clipped to the extent of the IACS dataset. In general, the magnitude of the class area differences is well reproduced using the Sentinel-based map. For example, “Meadows and pastures” has roughly half the area of “Sunflower” according to both datasets. However, the area of the three largest classes is somewhat underestimated with the Sentinel-based map data (the difference with IACS areas is 11–14%). The small-area classes are, as a rule, overestimated, this being the most severe for “Vineyards and orchards”, “Vegetables”, “Grain legumes”, and “Forage crops” where the difference from IACS data is more than 100%. The most accurate are the areas of “Winter rapeseed” and “Rice”, which are within 3% and 9% of the IACS data, respectively.

The accuracy of the Sentinel derived crop maps reported in the literature vary depending on the input data, methods and study area specifics. Very high accuracy (95–96%) was reported for example by Vuolo et al. [1], but they used a large number of cloud-free Sentinel-2 images, instead of composites, and mapped small region. In a study, similar to this presented here, Griffiths et al. [4] mapped 12 crop and land cover classes over Germany with 81% overall accuracy. In another national scale exercise Van Tricht et al. [3] classified dense time series of Sentinel-2 NDVI and Sentinel-1 backscatter data to map 12 crops and land cover types in Belgium, achieving overall accuracy of 82%. These results are similar to the accuracy reported here.

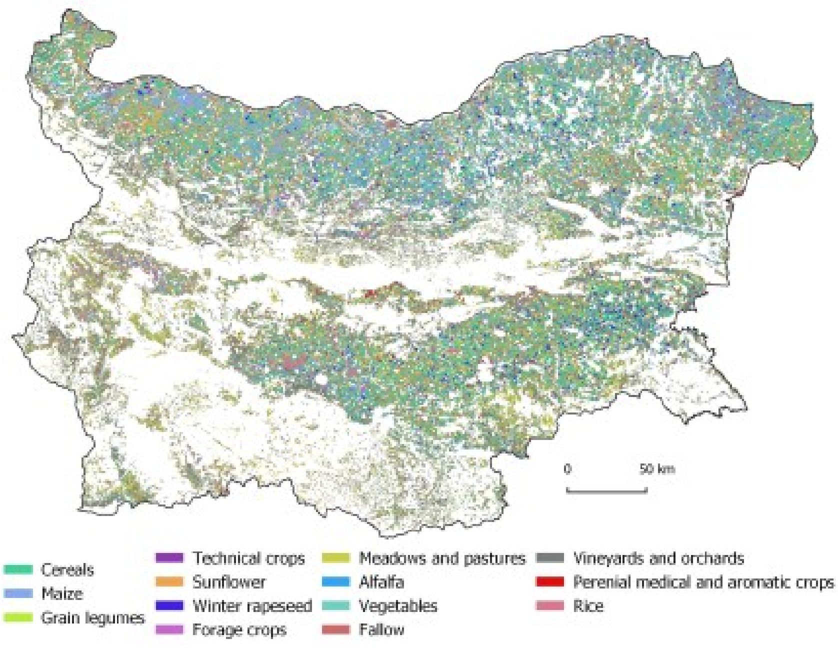


Fig. 2. Crop type map of Bulgaria for 2018/2019 agricultural year derived from Sentinel-1 *and -2 data. White areas are non-agricultural land*

The two-month compositing interval used in this study is relatively long to allow fine phenological differences between crops and be captured (Fig. 1). However, it ensured cloud-free Sentinel-2 mosaics over the entire study area with a negligible cloud contamination according to the visual inspection. Other studies have successfully applied shorter compositing periods for Sentinel-2, e.g. 10-day or month, but this may require smoothing and gap-filling the time series or even ingestion of Landsat observations [4, 5]. Griffiths et al. [4] showed that using 10-day composites resulted in higher accuracy for most classes than longer compositing periods. These developments may increase mapping accuracy in the Bulgarian context as well and should be examined in future studies.

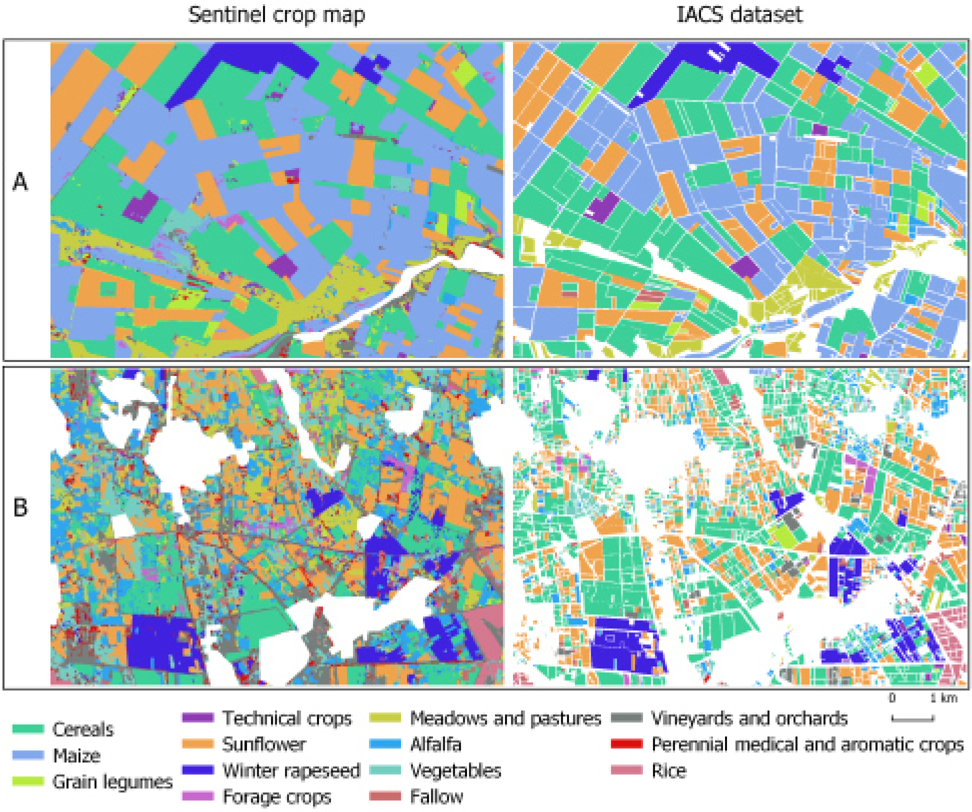


Fig. 3. Comparison of the crop type map of Bulgaria for 2019 derived from Sentinel-1 *and -2 data with the IACS dataset for selected regions: (A) Danube plain near Knezha and (B) Upper Thracian lowland near Plovdiv*

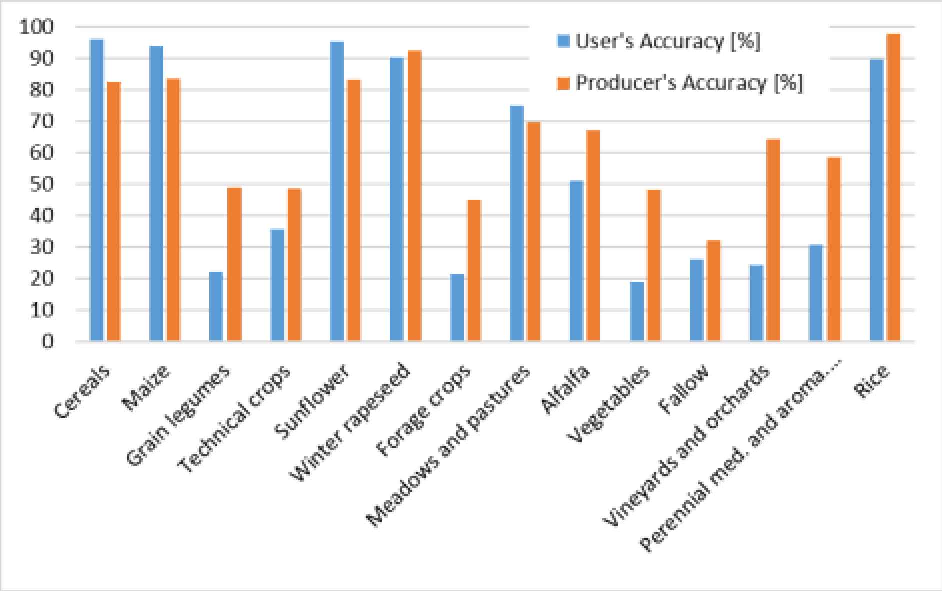


Fig. 4. User's and Producer's accuracy (%) of the 14 classes of the crop map

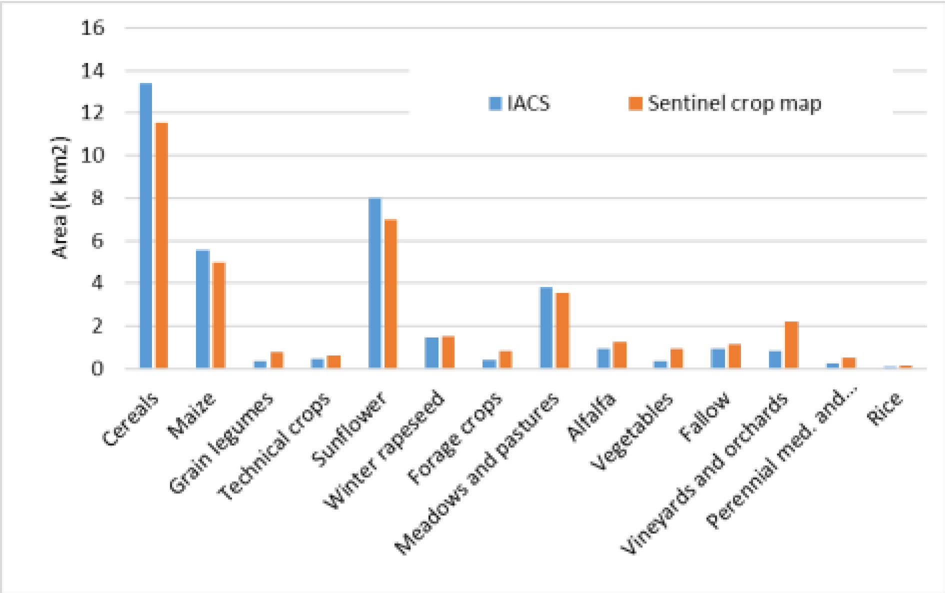


Fig. 5. Comparison of the areas of the 14 classes according to the Sentinel-based national crop map and the IACS dataset

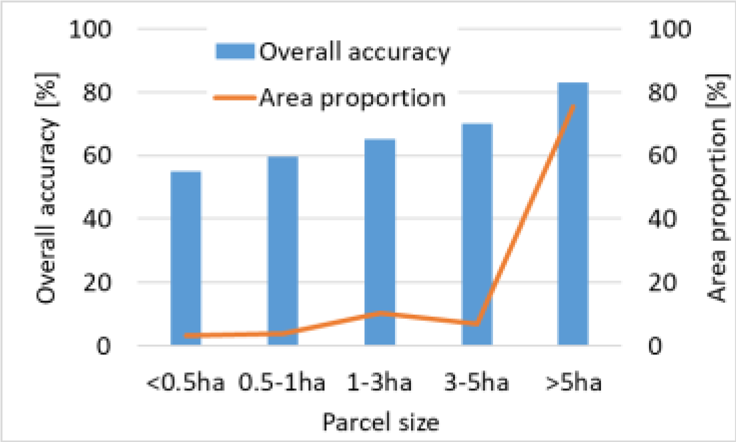


Fig. 6. Overall accuracy as a function of the parcel size. Area proportions are based on the validation dataset

Conclusions

This study is, to our knowledge, the first attempt to map crop types over the entire Bulgarian territory using Sentinel satellite imagery. A moderate overall accuracy of 78% is achieved, but results are better for the most important crops and crop types – "Cereals", "Sunflower", and "Maize". Problem for the classification is the recognition of some classes, which are too heterogeneous, e.g. "Vegetables" and "Forage crops". Such classes are included in the legend to comply with the existing nomenclature of crops used in the country, but the poor accuracy suggests that their usage is impossible in the context of the semi-automated remote sensing-based mapping. Higher overall accuracy was achieved with a combination of Sentinel-1 and -2 data than using only optical imagery. This confirms that SAR data derive important information for crop discrimination. It was also found that accuracy tends to be higher for larger parcels. Future studies should concentrate on the adjustment of the definitions of the classes. Mapping only individual crops, instead of groups of crops, is another approach but this would require a more computational resources. Further improvement of results may require testing of other classification algorithms and/or using composites with higher frequency.

Acknowledgements

This study is implemented under Action No.: 2018-1-86 "Copernicus Awareness Raising Programme for Bulgaria - COPE4BG" from FPCUP project with implementing partners SRTI-BAS and CASTRA. The FPCUP project is financed by the European Commission under the FPA no.: 275/G/GRO/COPE/17/10042.

The contents of the publication do not represent the official views of the European Commission nor do they constitute a commitment of any kind on its behalf. The IACS data were kindly provided by State Fund “Agriculture” of the Bulgarian Ministry of Agriculture, Food and Forestry.

References

1. Vuolo, F., M. Neuwirth, M. Immitzer, C. Atzberger, and W. Ng. How much does multi-temporal Sentinel-2 data improve crop type classification?, Int. J. Appl. Earth Obs. Geoinformation, 2018, 72, 122–30. DOI:10.1016/j.jag.2018.06.007.
2. Maponya, M. G., A. van Niekerk, and Z. E. Mashimbye. Pre-harvest classification of crop types using a Sentinel-2 time-series and machine learning, Computers and Electronics in Agriculture, 2020, 169, 105164. DOI:10.1016/j.compag. 2019.105164.
3. Van Tricht, K., A. Gobin, S. Gilliams, and I. Piccard. Synergistic use of radar Sentinel-1 and optical Sentinel-2 imagery for crop mapping: a case study for Belgium, Remote Sens., 2018, 10, 1642. DOI:10.3390/rs10101642.
4. Griffiths, P., C. Nendel, and P. Hostert. Intra-annual reflectance composites from Sentinel-2 and Landsat for national scale crop and land cover mapping, Remote Sens. Env., 2019, 220, 135–51. DOI:10.1016/j.rse.2018.10.031.
5. You, N. and J. Dong. Examining earliest identifiable timing of crops using all available Sentinel 1/2 imagery and Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161, 109–23. DOI:10.1016/j.isprsjprs. 2020.01.001.
6. Gorelick, N., M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, and R. Moore. Google Earth Engine: Planetary-scale geospatial analysis for everyone. Remote Sensing of Environment, 2017, 202, 18–27. DOI:10.1016/j.rse.2017.06.031.
7. Xu, L., H. Zhang, C. Wang, B. Zhang, and M. Liu. Crop Classification Based on Temporal Information Using Sentinel-1 SAR Time-Series Data, Remote Sens., 2019, 11, 53. DOI:10.3390/rs11010053.
8. Tomppo, E., O. Antropov, and J. Praks. Cropland Classification Using Sentinel-1 Time Series: Methodological Performance and Prediction Uncertainty Assessment, Remote Sens., 2019, 11, 21, 2480. DOI:10.3390/rs11212480.
9. Gikov, A., P. Dimitrov, L. Filchev, E. Roumenina, and G. Jelev. Crop type mapping using multi-date imagery from the Sentinel-2 satellites, C. R. Acad. Bulg. Sci., 2019, 72(6), 787–795. DOI:10.7546/CRABS.2019.06.11.
10. Hershkovich, E. Agroclimatic resources of Bulgaria. Sofia, Publishing house of the Bulgarian Academy of Sciences, 1984, 115 p. (in Bulgarian)
11. Breiman, L. Random forest. Mach. Learn., 2001, 45, 5–32. DOI:10.1023/ A:1010933404324.
12. Belgiu, M. and L. Drăguţ. Random forest in remote sensing: A review of applications and future directions. ISPRS J. Photogramm. Remote Sens., 2016, 114, 24–31. DOI:10.1016/j.isprsjprs.2016.01.011.
13. Pelletier, C., S. Valero, J. Inglada, N. Champion, and G. Dedieu. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas, Remote Sens. Env., 2016, 187, 156–68. DOI:10.1016/ j.rse.2016.10.010.

КАРТОГРАФИРАНЕ НА ЗЕМЕДЕЛСКИТЕ КУЛТУРИ В БЪЛГАРИЯ ЧРЕЗ ДАННИ ОТ SENTINEL-1/2

П. Димитров, Л. Филчев, Е. Руменина, Г. Желев

Резюме

През последните години, благодарение на достъпа до сателитни данни от Sentinel-1 и -2, се появиха нови възможности за полуавтоматично картографиране на земеделските култури на национално ниво. В това изследване са картографирани 14 земеделски култури и групи от култури на територията на България използвайки три двумесечни композитни изображения за 2019 година, генерирани в облачната платформа Google Earth Engine (GEE). Общата точност, когато се използват изображения както от Sentinel-1, така и от Sentinel-2 е 78%, докато точността е малко по-ниска, когато се използват само данни от Sentinel-2 (75%). Точността е най-висока за класовете "Зърнено-житни култури", "Царевица", "Слънчоглед", "Зимна рапица" и "Ориз" – над 80%. Точността при класове като "Зеленчуци", "Технически култури", "Фуражни култури", "Угар" и др. обаче е по-ниска. Тези класове представляват категории, подходящи за използване в земеделската практика и статистика, но са твърде общи и трудни за отличаване чрез сателитни данни. Беше установено също така, че точността е по-висока за парцелите с по-големи размери. Като възможни пътища за подобряване на резултатите са посочени използването на серия от композитни изображения с по-голяма честота и адаптирането на класовете от класификационната система, така че да включват култури, които са сходни по фенология и морфология.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e05

LAND COVER AND LAND USE CHANGE IN KARST REGION DEVETASHKO PLATEAU

Georgi Jelev[1], Dilyana Stefanova[2], Petar Stefanov[2]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences
[2]National Institute of Geophysics, Geodesy and Geography – Bulgarian Academy of Sciences
e-mail: g_jelev@abv.bg, dili_stefanova@abv.bg, psgeo@abv.bg

Keywords: *Land Cover, Land Use, Karst, Karst Geosystems, Devetashko Plateau, Corine, Land Cover*

Abstract

The Corine Land Cover (CLC) is a digital data about land cover, which is distributed into 44 classes, whereas for the territory of Bulgaria the CLC classes are 36. The minimal mapping unit is 25 hectares (for 2D objects and 100 m for linear objects). Data sets for the years 1990, 2000, 2006, 2012, and 2018 are available, as well as for the changes , which have occurred between each couple of years (1990–2000, 2000–2006, 2006–2012, and 2012–2018).

The great data sets provide to track over a nearly 30-year period of land cover changes in model karst regions, which are strongly vulnerable to anthropogenic and natural influences.

This paper considers the changes in the land cover types on the Devetashko plateau – a typical karst plateau in North Bulgaria. Land cover and land use changes directly affect the processes of modern karst-genesis, the soil-vegetation cover, the quantity and quality of underground karst waters.

Introduction

Karst is a widely spread natural phenomenon with which the lives of millions of people are directly or indirectly related [1]. Karst territories are made of karst geosystems, which are characterised by the fact that they consist of clearly expressed underground (caves) and above-ground part that are closely related genetically and dynamically [2, 3]. Karst geosystems feature high vulnerability and enhanced risk of influences, especially against the background of the increasingly expanding global changes. In Bulgaria, karst territories comprise ¼ of its area [4] and from specific living and economic activity environment for people who must consider its peculiarities on a daily basis. In addition, this puts forward on the agenda the urgent need of good knowledge of and compliance with karst specifics. This is the only way to achieve sustainable development of karst territories and to resolve successfully the topical social and economic problems during their handling and management.

The active processes occurring on a global and European scale since the end of the 20th century affect the social and economic development of Bulgaria, as well. After 1989, many structural (with respect to ownership and production), technological and organizational reforms are ongoing, which affects all spheres of economic activity: industry, agriculture, services. After the passing of the Restoration of Land Ownership Act (1991) and with the start of the reform in 1992, minor owners and unemployed became predominant in the field of agriculture (as a result of the destroyal of the production structures existing before). The new users of the former publicly owned land became owners of agricultural land, tenants, the hired workforce, self-employed people, i. e. differentiation of economic interests occurred. Subsequently, access to European subsidies from the pre-accession funds and mostly, from the structural European funds in the field of agriculture, resulted in new changes in the use of agricultural land and the grown agricultural crops. These changes affect especially strongly the vulnerable karst territories. Changes in the development of agriculture and cattle breeding also occurred after the adoption of Natura 2000 and the application of the relevant environmental protection law.

Changes are also observed in the use of building material quarries, which are typical for karst territories. Some quarries are abandoned, but others are let on concession. New quarries are also developed.

The sequence of changes in land use strongly affects the fragile balance in karst geosystems. The role of anthropogenic pressure is of varying nature. For instance, incorrect agricultural practices in karst territories result in accelerated erosion and pollution of underground karst waters. This requires prevention and undertaking of urgent measures for the sustainable development of karst territories. These are also indispensable because of the depopulation of karst regions, which makes them socially unstable. At the same time, during the recent years, external users (tenants) are widely applying deep ploughing with heavy machines of the arable lands, including the lands that were neglected during the transition. The economic objectives of these land users are not always related to the ecological standards imposed by the karst specifics of the regions [5]. The consequences of the increasingly active manifestations of global climatic changes and the increasing number of extreme climatic events, especially torrential rains and quick snow melting, are not accounted for, either. On the karst terrains with shallow soils and numerous whirlpools and pot-holes, they cause accelerated erosion, especially for arable lands. This results in import and deposition of great quantities of solid (clayish) deposits in the underground cave systems (Fig. 3).

During recent years, karst territories with attractive forms are becoming an object of increasing interest for tourists and are turning into sports and leisure sites. [6–9]. They attract both organised, as well as unorganised tourists. Regretfully, tourist infrastructure in karst terrains is insignificant or badly designed and built, including the ecopaths to popular karst objects. These pose a number of risks for visiting tourists, with one lethal accident already recorded (Krushunska ecopath). The major reason is that the available ecopaths lack clear status and their handling and maintenance is not regulated by law. No control on the visits is exercised, including along the ecopaths in protected karst territories. This has often negative effect on karst objects, incl. pollution, knocking-off, scratching. However, karst territories will continue to be attractive tourist destinations with positive regional and local economic effects. The latter, however, is still weak because of the insufficient experience of local communities in tourist industry, on the one hand, and on the other hand – their insufficient knowledge of karst specifics.

All this requires, especially at the managerial level, a better understanding of the relations between the major activities in karst terrains, such as agriculture, mining, forestry, tourism and more. The long–term and short-term effects of these activities might have and have on karst environment and karst heritage should also be accounted for. It is also important to observe the changes in agriculture and to assess and forecast the consequences thereof in view of rational management and minimising the negative effects on biodiversity, health, and life of the population. In this relation, land cover changes are an important information source [10]. Within the earth surface monitoring service of the EU Copernicus Programme and under the guidance of the European Environment Agency under the Corine, land cover (CLC) Project, the vector layers with land cover types and their change with time are maintained and updated on a regular based on [11]. The major thesis of this publication is to determine the extent to which this information may be useful to analyse the changes in land use in a typical karst region. It tracks the land cover changes on the Devetashko plateau in Northern Bulgaria, which is a model region for multi-annual studies by the Experimental Laboratory of Karst Studies of the National Institute of Geophysics, Geodesy and Geography of the Bulgarian Academy of Sciences (NIGGG – BAS).

Materials and Methods

Using geographic information systems (GIS), the spatial distribution and quantitative characteristics of the land cover and land use (LCLU) classes of the Devetashko plateau are analysed.

During the conduct of this study, a GIS database was composed (KARST.gdb), with a set of vector and raster layers containing administrative boundaries (of districts, municipalities, lands/settlements' land-use areas (SLUAs)[1], populated places and state boundary), water objects (lakes, dams, and river network), contour of the model regions, protected territories under Natura 2000

[11] *Totality of the land properties belonging to a given settlement (i.e. the land properties both in the settlement and in the settlement's adjacent territory).*

(according to the Directive on birds, Directive on habitats), Corine land cover (CLC) under the Copernicus Programme, containing vector layers for 1990, 2000, 206, 2012, 2018, and the change of land cover (Corine, Land Cover Change, LCC) over the subperiods: 1990–2000, 2000–2006, 2006–2012, 2012–2018, digital elevation model (altitude belts) – ASTER Global Digital Elevation Model (GDEM) with spatial resolution of 30 m.

Table 2. Nomenclature „Corine, land cover", Level 1 ÷ 3, Fig. 2.÷5.

<table>
<tr><th>Level 1</th><th>Level 2</th><th>Level 3</th></tr>
<tr><td rowspan="3">1 Artificial surfaces</td><td>11 Urban fabric</td><td>112 Discontinuous urban fabric</td></tr>
<tr><td>12 Industrial, commercial and transport units</td><td>121 Industrial or commercial units</td></tr>
<tr><td>13 Mine, dump and construction sites</td><td>131 Mineral extraction sites</td></tr>
<tr><td rowspan="6">2 Agricultural areas</td><td>21 Arable land</td><td>211 Non-irrigated arable land</td></tr>
<tr><td rowspan="2">22 Permanent crops</td><td>221 Vineyards</td></tr>
<tr><td>222 Fruit trees and berry plantations</td></tr>
<tr><td>23 Pastures</td><td>231 Pastures</td></tr>
<tr><td rowspan="2">24 Heterogeneous agricultural areas</td><td>242 Complex cultivation patterns</td></tr>
<tr><td>243 Land principally occupied by agriculture, with significant areas of natural vegetation</td></tr>
<tr><td rowspan="6">3 Forest and semi natural areas</td><td rowspan="3">31 Forests</td><td>311 Broad-leaved forest</td></tr>
<tr><td>312 Coniferous forest</td></tr>
<tr><td>313 Mixed forest</td></tr>
<tr><td rowspan="2">32 Scrub and/or herbaceous vegetation associations</td><td>321 Natural grasslands</td></tr>
<tr><td>324 Transitional woodland-shrub</td></tr>
<tr><td>33 Open spaces with little or no vegetation</td><td>332 Bare rocks</td></tr>
<tr><td>5 Water bodies</td><td>51 Inland waters</td><td>512 Water bodies</td></tr>
</table>

Source: I.1., Elaboration: Georgi Jelev

The methods of spatial analysis and statistics in the medium of a geographic information system (GIS) were used. The data were processed within the boundaries of the Devetashko plateau and were matched at SLUAs level. The data are organised in a Pivot table and the spatial relationships between the land cover/land use types and their changes over the period 1990–2018 were studied. The period after 1990 was studied when the change of ownership and the restitution of agricultural land caused exceptionally serious changes in the organization of agriculture, too.

The spatial distribution and the quantitative characteristics of the presented land cover and land use classes were analysed at two levels: for the entire region of the Devetashko plateau and for groups of lands/SLUAs, falling within parts of municipalities which are included within the plateau's boundaries. The analyses were conducted at level 1 and level 3 of Corine land cover. Comparative analysis was also used. The transition from one land cover class to another over subperiods was also tracked.

Devetashko Karst Plateau

The Devetashko plateau (343.08 km^2) is located between the valleys of the rivers Rositsa and Osam in the southern periphery of the Danube (Moesian) plane in Bulgaria. Its northern foot features altitude of 100÷150 m a.s.l. and is outlined by a clearly expressed fault area in the relief, and its southern foot features altitude of 350 ÷ 400 m a.s.l. and is not expressed clearly. The plateau-shaped ridge part varies between 350 and 450 m a.s.l. and is inclined to the north. The highest point of the plateau is the Chukata peak (558 m a.s.l.).

The Devetashko plateau is composed mainly of organogenic limes of Lower Cretaceous age (Apt-Urgon). They make up subhorizontal layers with a slight inclination to the north. The prevailing development of limes is the reason for which, on the plateau, classical karst of autochthonous type has developed, after [12], without permanently running river waters. In the relief, morphologic karst complexes of classical forms – surface and underground, dominate, with prevailing whirlpools and hollows/fif/ek their total number being several thousand. They are elements of well-differentiated karst geosystems whose exits are mostly on the northern slopes and the foot of the Plateau, which are marked by high-capacity karst springs and large entrances of spring caves. One of the most typical karst geosystems is the Kurshunska one (43 km^2, located in the eastern part of the plateau), which since 1990 has been the model karst region of the Experimental Laboratory of Karst Studies of the NIGGG–BAS (Fig. 3).

Many of the plateau's pot-holes have been reshaped into precipices and precipice caves. Another part of the pot-holes has been tamponed (incl. by humans), with karst lakes and bogs formed around – one of the most typical in Bulgaria [13, 14].

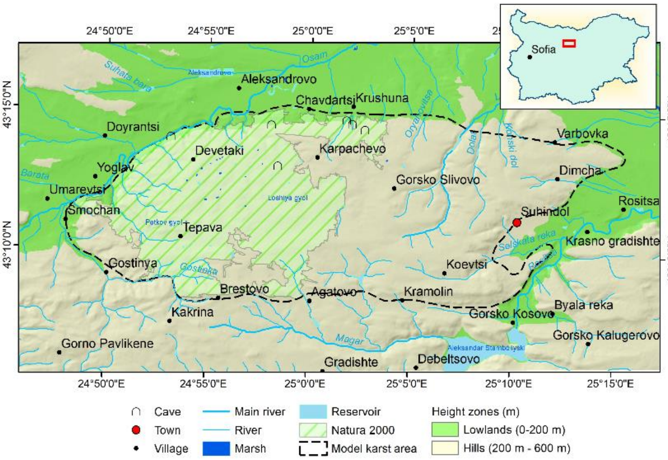


Fig. 1. Devetashko Plateau

Sources: *Information System for protected areas according to ecological network NATURA 2000 [I.2.], Register of protected areas in Bulgaria, Executive Environment Agency – MOEW, Bulgaria: Digital Elevation and administrative data: The Study of Integrated Water Management in the Republic of Bulgaria. JICA; Field mapping of caves and dolines: P. Stefanov – unpublished paper „Comparative geodynamics of karst geosystems in the Predbalkan and Western Rhodopes", Sofia, 1993, Institute of Geography –BAS. Drawn by Georgi Jelev*

On the Devetashko plateau, 68 karst caves with a total length of about 18 km have been studied. The longest one is the water Boninska Cave (Popova cave) by the Krushuna village – 4 530 m, which is connected with the spring cave Vodopada (1 995 m) – one of the longest water cave systems being studied in Bulgaria (above 6.5 km). Some caves on the plateau are former objects of economic activity (the Chavdarska cave/Mandrata and the Devetashka cave – a former object of the army) or have been used for water supply. As tourism in the Devetashko plateau became more popular, 7 caves are already objects of non-regulated tourist activity. The most frequently visited one is the Devetashka cave where the greatest cave hall in Bulgaria has formed.

Fig. 2. Devetashko plateau – Landscape in the vicinity of Gorsko Slivovo village
Photo: P. Stefanov

The climate of the Devetashko plateau is moderately continental, with an average annual temperature about 10.5 ºC. In winter, the plateau-shaped part features temperature, which is by about 2 ºC higher than the temperature of the north foot. During recent years, abrupt heatings have been observed even in February, which causes quick snow melting in the ridge parts (Fig. 3). The average annual amount of precipitation is between 650 and 700 mm, with spring-to-summer maximum (May–June). About 33% of the precipitation form underground water outflow, which is a typical karst. The biggest karst springs are concentrated in the northern periphery of the plateau (near-fault drainage system). Their waters are fresh, hydrocarbon-calcium. A typical feature of the underground karst waters is their strong vulnerability to pollution. Its sources are the non-regulated landfills in the widely spread whirlpools with active pot-holes, as well as the continuous anthropogenic loading (mainly agricultural) on the plateau's territory. The settlements located on the plateau lack sewerage systems and wastewater is not treated. Therefore, the abounding underground karst waters on the plateau are strongly polluted and the greatest part of them is not potable.

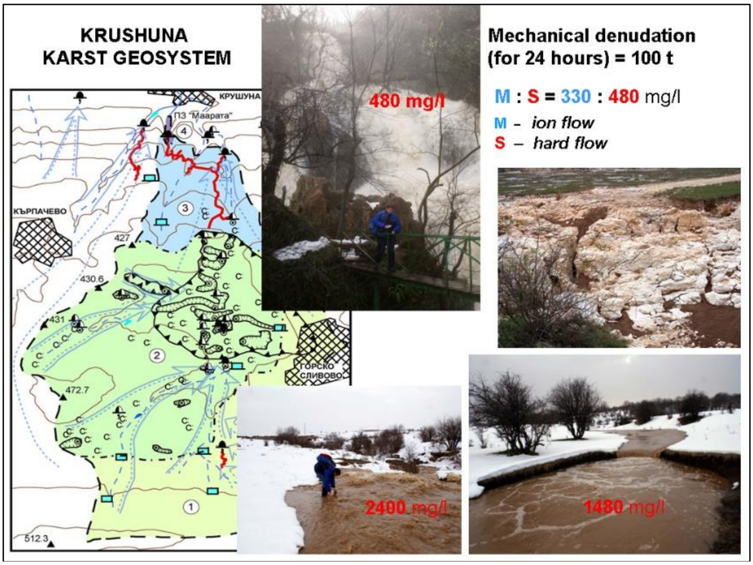


Fig. 3. Active mechanical denudation in the Krushuna karst geosystems in the Devetashko plateau resulting from intensive snow melting (February, 2012) (archive of the Experimental Laboratory of Karst Studies of the NIGGG–BAS, Drawn by P. Stefanov)

The soil cover of the plateau is made of typical-for-Bulgarian-karst soil types [15], with prevailing luvisols (LV), rendzik, (LPk), and dystric (CLd) soils, developed in the foot of the Plateau and in the uvalas, hollows, and whirlpools.

Albeit the karst terrain, the slightly sloped ridge parts are covered with a thick soil layer (mostly luvisols) where traditional agriculture is developing. The eroded arable lands and the slopes of the negative karst forms with karren fields are used as pastures. Cattle breeding in the Devetashko plateau has deep historical roots and with it, the artificial tamponing of whirlpools' pot-holes in which karst bogs and lakes where cattle drinks water have formed, is related. Calcic (GLk) soils have also formed there. During the dating of the depositions in Irmanov gyol, located westward of Gorno Slivovo village, it was established that the age of the bog is about

Three centuries [16]. The results from the palinologic studies show that grass vegetation with rich diversity of ruderal and anthropophyte types was dominant. They evidence of active anthropopressure – the palaeoecological circumstances have been favourable for development of agriculture and cattle breeding in the region.

The prevailing natural vegetation on the Devetashko plateau is broad-leaved, mostly oak and Carpinus orientalis. Some limited areas have been artificially planted with coniferous forests. Part of the plateau's territory has been included in the ecological network Natura 2000. Two natural landmarks and six protected countryside with an area of 55.64 ha have been announced.

The arable land, which occupies a great percentage of the plateau's ridge part is attacked by accelerated erosion during torrential rains and especially, during active snow melting, when arable lands are most vulnerable to erosion. In addition, it is particularly active in the ploughed periphery of the whirlpools and hollows into which the generated snow water flows out. The ploughing applied by the tenants does not comply with the specifics of the karst relief and is carried out up to the edge of the whirlpools and hollows falling within the boundaries of the rented land. As a result, part of the soil is carried away through the underground karst systems and the springs that drain them. A typical example is the Krushunska karst geosystem (Fig. 3). During the observed active snow melting in February 2012, within 24 h only, about 100 tonnes of deposits (mostly at the expense of eroded ploughed soils) were carried away through the Vodopada karst spring, whereas 10 tonnes of them were deposited along the Krushunska travertine cascade proclaimed as natural landmark under the name „Maarata".

The area of the Devetashko plateau includes, in whole or in part, 20 lands pertaining to five municipalities and three administrative regions (Fig. 4). On the territory of the Devetashko plateau and along its periphery, 15 settlements are located [2] (Fig. 4). The administrative partitioning of the Devetashko plateau poses significant risks for its management from the viewpoint of sustainable development. The reason for this lies in the lack of an integrated regional and sector policies, incl. in the field of agriculture, tourism, and other activities exercised on the plateau's territory.

The population of the Devetashko plateau amounts to 5 858 (2018). Over the period 1990–2018, the region has lost 45.50% of its population (Fig. 5). Ageing and depopulation are typical demographic processes of the plateau, which account for the old residential buildings, some of which have been abandoned and are crumbling down.

The economically active population is less than one-third of the overall active population. In some villages (Tepava, Devetaki, Brestovo, Gostinya and more), the economically inactive population exceeds 90%.

Fig. 4. Territorial and administrative location of the Devetashko plateau. Drawn by Georgi Jelev

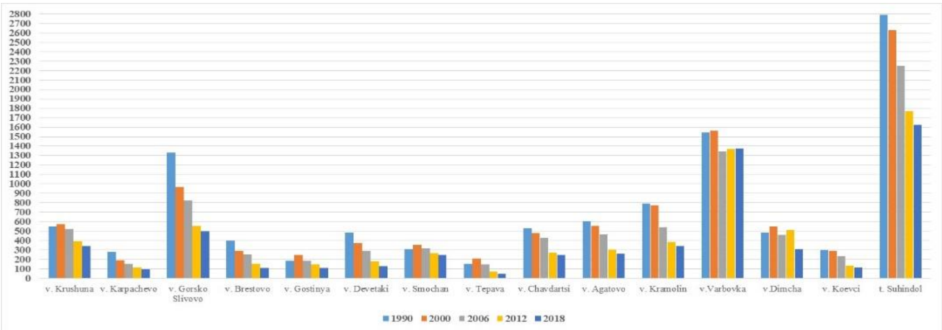


Fig. 5. Movement of the population by settlements on the territory of the Devetashko plateau over the period 1990–2018. Source: NSI – Bulgaria, Elaboration: D. Stefanova

On the territory of the Devetashko plateau, there are no industrial enterprises. Agriculture is the main developing economic activity which is exercised by tenant farmers and, to a lesser degree, by local users of agricultural land. EU subsidies for direct payments stimulate farmers to increase the area of the used agricultural land (UAL). Active activity is ongoing, aimed at changing the manner of the agricultural land's lasting use – more areas with cereals and less with orchards compared to the period up to 1989. The widest grown cereals are wheat and barley; in the gardens, they plant maize, sunflower, as well as plums, cherries, and peaches. Vines are traditionally typical for Suhindol and, to a lesser degree, for Krushuna. Because of the established system of middlemen of forest fruits, the growing of strawberries, raspberries, and blackberries is increasing steadily. A tendency for the restoration of the old orchards is also observed, due to the possibility to obtain additional agroecological payments [5].

Cattle's breeding is mostly related to the growing of cows, sheep, and goats, using the pastures in the karst terrains. It pertains to the small, mostly family sector, which provides no conditions to breed a large number of animals.

On the forest territory, apart from logging, a number of auxiliary uses are available – pasture of large and small cattle, collecting of hay from the bare areas, collection of leaf fodder. Herbs, forest fruits, mushrooms and nuts are also collected.

The Devetashko plateau provides possibilities for developing various forms of sport and tourism: ecotourism, educational tourism, and cultural tourism. Karst objects – caves and waterfalls, are attractive. Because of the low economic level of the settlements located on the Devetashko plateau and along its periphery, tourism in karst terrains is a good prospect for improving the life quality of the local communities. From this perspective, it is also possible to establish leisure and recreation territories.

The active social and economic changes in the Devetashko Plateau have also been reflected as land cover and land use changes, which, because of the specific karst territory, require analysis in three aspects – economic, social, and ecological.

Results and Discussions

The overall review of Corine, land cover and land use on the territory of the Devetashko plateau

On the territory of the Devetashko plateau, four land cover and land use classes have been observed after the CORINE nomenclature, level 1: artificial surfaces (1[2]), agricultural areas (2), forest and semi natural areas (3), and water bodies (5) over the whole observation period, 1990–2018 (Fig. 6). Of them,

[2] *after the Corine, land cover nomenclature (Table 2)*

Artificial surfaces (1) feature the greatest area. This account for a little bit less than 60% of the plateau's area. Over the whole period, their share has been declining within the range 59.29% – 58.76%, i.e. a tendency for a minor decrease by 0.53% is available. On the contrary, with *Forest and semi natural areas* (3), which come second by area, a tendency for a minor increase from 37.11% to 38.33% is observed. Ranking third by their share presentation are *Artificial surfaces* (1). With this class, the tendency is for a slight decrease, which starts from 2006, their share being preserved during the reporting years afterwards. *Artificial surfaces* (1) are mostly formed by class *Discontinuous urban fabric* (112) and, to an insignificant degree – by *Industrial or commercial units* (121) and *Mineral extraction sites* (131), (Fig. 6, Fig. 7). The identified change the identified change is mostly due to the depopulation of the Devetashko plateau and the scrambling of the abandoned anthropogenic infrastructure. The share of *Water bodies* (5) is less than one percent.

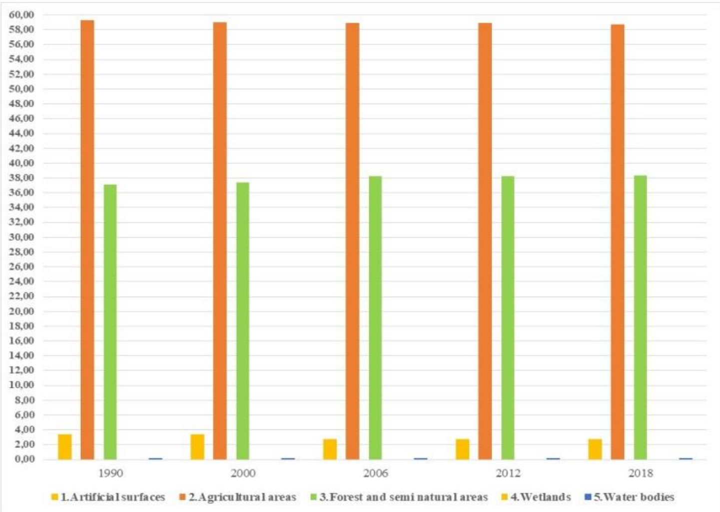


Fig. 6. Distribution and dynamics of Corine, land cover level 1
on the territory of the Devetashko plateau
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

The analysis of Corine land cover level 1 demonstrates insignificant dynamics of the four major classes identified on the plateau. This means that it is mandatory to track the changes at lower levels so as to not ignore changes with effects of significance for the karst territory.

On the territory of the Devetashko plateau, four classes of Corine land cover and land use level 3 are observed (Fig. 7, Fig. 8). They are included in the above-mentioned classes at level 3 and are tracked for the years 1990, 2000, 2006, 2012, and 2018. The review at this level reveals in greater detail the change, which is important to observe when accounting for potential negative impacts.

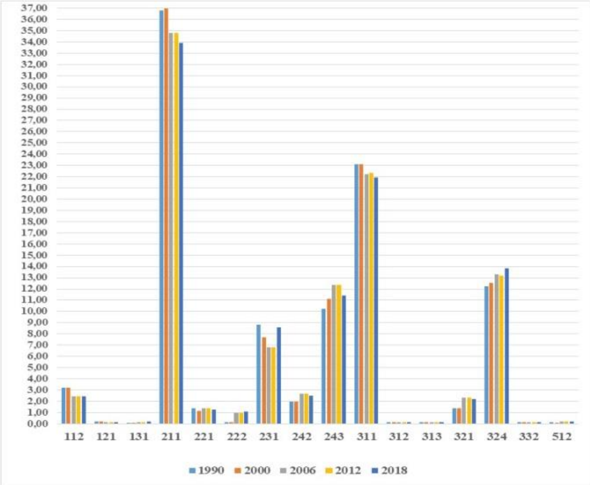
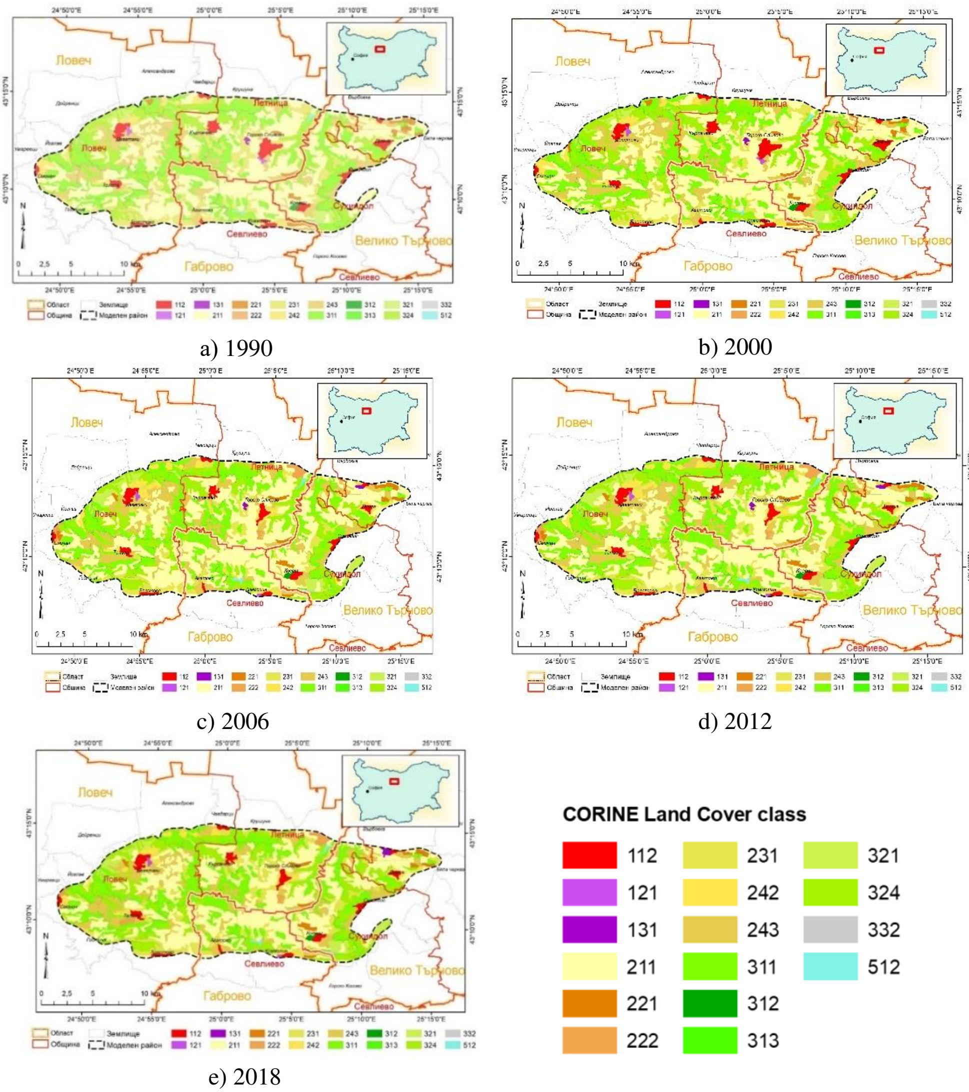


Fig. 7. Distribution and dynamics of Corine, land cover level 3 on the territory of the Devetashko plateau
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

During the five observed years, *Non-irrigated arable land* (211) features greatest shares (above 33%) of the plateau's territory. Irrespective of preserving its greatest shares with respect to the other classes, a tendency for declining by less than 3% during the last year (2018 – 33.90%) compared to the first year (1990 – 36.78%) (Fig. 7) is observed. Second, come *Broad-leaved forests* (311), which drop from 23.13% (1990) to 21.94% (2018) (Fig. 7). As of 2018, these two classes of Corine land cover level 3 occupy a bit more than half of the territory of the plateau, which makes them exceptionally important from the viewpoint of impact on the karst. Ranking third is the share of *Transitional woodland-shrub* (324) whereas, during the years, a lasting tendency for the shares' increase from 12.23% to 13.81% (Fig. 7) is observed. Featuring shares close to class (324) is class *Land principally occupied by agriculture, which* ranks third, with significant areas of *Natural vegetation* (243). It increases from 10.23% (1990) to 12.37% (2006, 2012), after which it decreases to 11.39% (2018) (Fig. 7).

The shares of class *Pastures* (231) are also significant, which change over the years within the range 7%–9% (Fig. 7). Their change is closely related to the development of cattle breeding on the territory of the plateau, which, after 1989, was strongly affected by the social and economic changes in the country and, more specifically, with those in agriculture. The shares of class *Complex cultivation patterns* (242) (Fig. 7) also increase, which is related again with the changes in agriculture and the reinstitution of land in actual boundaries to its owners. The other presented classes (Fig. 7) feature insignificant shares, but irrespective of this, the change therein may also produce some negative impacts on karst.

From the viewpoint of the performed analysis, the changes related to the transition of one class into another in Corine land cover (Fig. 9), are interesting, as well.



a) 1990 b) 2000 c) 2006 d) 2012 e) 2018

Fig. 8. Territorial distribution of Corine, land cover (CLC) level 3 (classes, Table 1) of the Devetashko plateau Source: Drawn by G. Jelev

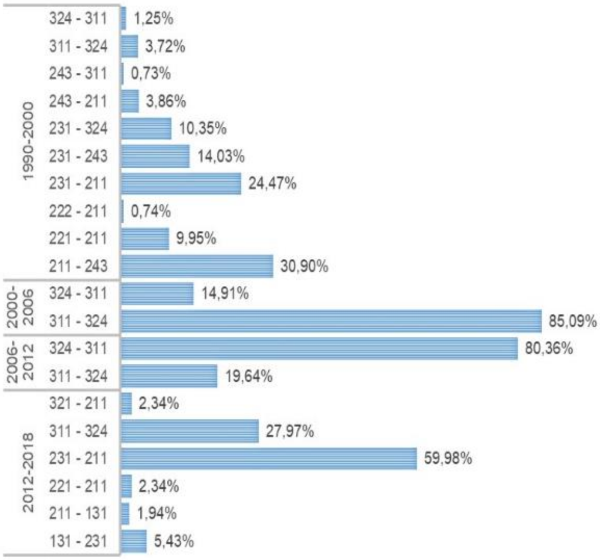


Fig. 9. Land cover change for the considered subperiods
Source: [I.1.], Elaboration: G. Jelev

The greatest number of changes of one class into another are effected in ***sub-period 1990–2000***, predominantly for 7 classes of agricultural areas (2) (Fig. 9, Table 1), logically following the economic changes in the country. A significant share of the changes over this time interval (30.90%) is related to the transition of Non-irrigated arable land (211) into Land principally occupied by agriculture, to *significant areas of natural vegetation* (243). Another nearly one-fourth of the changes are affected by *Pastures* (231) turning into *Non-irrigated arable land* (211). Changes from *Pastures* (231) to *Land principally occupied by agriculture, with significant areas of natural vegetation* (243) and T*ransitional woodland-shrub* (324) are also observed. Overall, most strongly affected are *Pastures*, of which a total of 44.85% have been transformed into other classes over this period. *Vineyards* (221) and *Fruit trees and berry plantations* (222) also lose areas, passing to *Non-irrigated arable land* (211). In conclusion, this sub-period is characterised by exceptionally great dynamics within *Agricultural areas* (2) (Fig. 9, Table 1).

During the next sub-period, 2000–2006, *the changes are only in Forest and semi natural areas* (3) and they are related to transition of *Broad-leaved forest* (311) into *Transitional woodland-shrub* (324), which accounts for 85.90% of the change. The other 24.91% of the changes over the period are in the reverse direction – from (324) to (311) (Fig. 9, Table 1).

During the third sub-period, 2006–2012, the changes are again in *Forest and semi natural areas* (3) between classes (311) and (324), but the difference lies in the fact that now the greater share of the change (80.36%) is oppositely directed, from *Transitional woodland-shrub* (324) to *Broad-leaved forest* (311) (Fig. 9, Table 1).

During the last sub-period, 2012–2018, the greatest share of the changes (64.26%) comprises three classes of A*gricultural areas* of Corine, land cover (Fig. 9, Table 1). Most significant are the changes in *Pastures* (231), the greatest share being occupied by the transition into *Non-irrigated arable land* (211). The changes in *Forest and semi natural areas* (3) are also significant. They comprise *Broad-leaved forest* (311) and *Natural grasslands* (321). A greater share for the subperiod is characteristic of the transition of *Broad-leaved forest* (311) into *Transitional woodland-shrub* (324) – 27.97%.

In relation to the karst plateau, the changes related to the transition of *Non-irrigated arable land* (211) into *Mineral extraction sites* (131) deserve particular attention.

Review of CORINE, land cover and land use by types of SLUAs within the boundaries of municipalities falling within the territory of the Devetashko Plateau

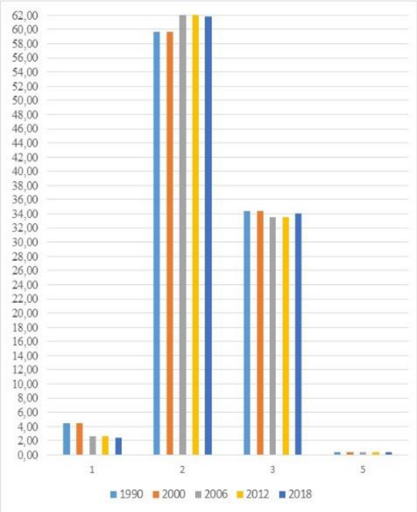
The Devetashko karst plateau is managed by different municipal administrations (Fig. 4). Because of the specifics of karst geosystems, the handling and use of the land require integrated management approach. In this sense, the observation of changes in land cover/land use types by groups of the SLUAs within the boundaries of the municipalities falling within the territory of the Devetashko plateau provides information which may be used in the development of ecological assessments and policies for local development and planning.

Lands from the Municipality of Letnitsa (Lovech District)

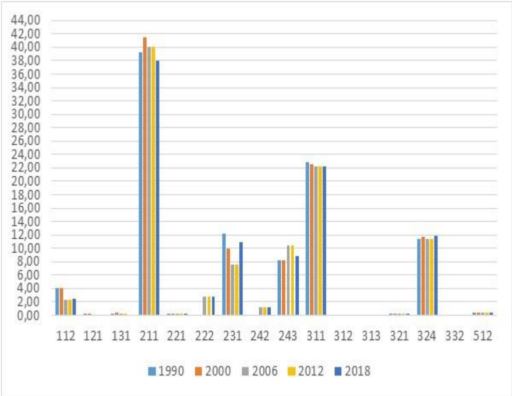
The lands of the villages from the Municipality of Letnitsa (Lovech District), which fall within the plateau's boundaries, are located mostly in the planar ridge part (Fig. 4). The SLUAs of Gorsko Slivovo and Karpachevo villages are included in whole, and the land of Krushuna village is included in part. Their total area amounts to 84.53 km^2 or 24.64% of the plateau's territory. The population of the three settlements amounts to 934 people (2018). The drop in the population compared to 1990 (2158 people) is 56.72% (Fig. 5). The major economic activities are farming (agriculture and cattle breeding) and tourism.

This part of the plateau houses one of the most typical karst geosystems – the Krushunska geosystem (area of 43 km^2) (Fig. 3). Of the surface karst forms, whirlpools prevail, their number only in the karst geosystem being 379.

Based on the performed analysis, in this group of SLUAs at CORINE level 1, *Agricultural areas* (about and over 60%), *Forest and semi natural areas* (about 1/3 of the total area), and with insignificant shares – *Artificial surfaces* and *Water bodies* (Fig. 10a) are identified. The share distribution is preserved until 2000, and since 2006, changes occur in the first three classes at this level. *Artificial surfaces* have dropped by about 2%. *Agricultural areas* display a tendency of gradual weak increase during the next years, and *Forest and semi natural areas* display the opposite tendency – insignificant drop during the next two observed years followed by an insignificant increase during the last year, 2018.



a)



b)

Fig. 10. Land cover changes (classes, Table 1) in a group of SLUAs from the Municipality of Letnitsa falling within the territory of the Devetashko plateau (%):
a) (Level 1); b) (Level 3)
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

At CORINE level 3 on the lands of the Municipality of Letnitsa, changes in the number of presented classes are observed over the years. For 1990 and 2000 they are 11, in 2006 and 2012 they increase to 12, and in 2018 they are 11 again (Fig. 10b). With respect to the plateau as a whole, on these lands, the classes *Coniferous forest* (312), *Mixed forest* (313), and *Bare rocks* (332) are lacking – all of them pertaining to the *Forest and semi natural areas*. In the beginning of the period, *Fruit trees and berry plantations* (222) and *Complex cultivation patterns* (242) from *Agricultural areas* are also lacking, which is closely related to the changes in agriculture and their specific impact on these territories. The other absent classes at the end of the period, and more specifically, *Industrial or* *commercial units* (121) and *Mineral extraction sites* (131) refer to *Artificial surfaces*.

The typical thing about the lands of the Municipality of Letnitsa is the exclusively great shares of *Non-irrigated arable land* (211), followed by *Broad-leaved forest* (311) (Fig. 10b). *Transitional woodland-shrub* (324), *Land principally occupied by agriculture, with significant areas of natural vegetation* (243) and *Pastures* (231) feature approximately equal shares within the limits of up to 12%. Another characteristic is the dynamics of the changes in the shares within *Agricultural areas* where decreases and increase alternate.

Lands from the Municipality of Lovech (Lovech District)

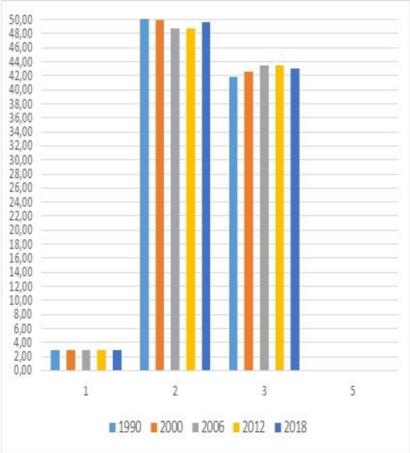
The SLUAs from the Municipality of Lovech (Lovech District), which falls within the boundaries of the plateau, include in whole the SLUAs of Devetaki and Tepava villages and partially, the lands of another seven villages (Fig. 4). Their total area amounts to 126.93 km^2, or 37% of the plateau's area. The population numbers 891 people (2018). The drop of the population compared to 1990 (2072 people) is 57% (Fig. 5). This is the most depopulated part of the plateau. In economic aspect, mainly agriculture and forestry are presented. In the northern part, the Devetashka cave is located, which is a landmark and has been turned into tourist attraction.

On the SLUAs from this Municipality, in contrast to the Municipality of Letnitsa, relatively close shares of *Agricultural areas* and *Forest and semi natural areas* (Fig. 11a) are observed. The changes in both classes at level 1 are insignificant, with alternating drops and increases. Class *Water bodies* (512) is absent, and class *Artificial surfaces* retain exclusively low values during the whole observed period.

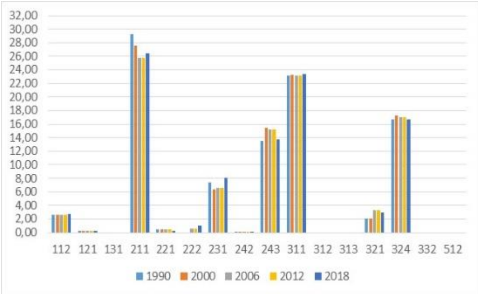
At CORINE level 3, changes in the number of presented classes over the years are also observed. For 1990 and 2000 they are 10, and after 2006 they increase to 11 (Fig. 11b). The classes permanently lacking on the plateau as a whole are *Mineral extraction sites* (131), *Coniferous forest* (312), *Mixed forest* (313) and *Bare rocks* (332), all of them pertaining to *Forest and semi natural areas*. For the beginning of the period 1990–2000, absence of areas with *Fruit trees and berry plantations* (222) is identified.

About 50% of the areas are distributed between *Non-irrigated arable land* (211) and *Broad-leaved forest* (311), with their values tending to become equal as of the end of the period (Fig. 11b). *Agricultural areas* are supplemented by *Land principally occupied by agriculture, with significant areas of natural vegetation* (243) and *Pastures* (231), with shares of varying dynamics ranking between 6% and 8%. *Forest and semi natural areas* supplements its shares mainly by *Transitional woodland-shrub* (324), which features almost constant values

(Fig. 11b). The changes in the classes pertaining to *Agricultural areas* are more obvious.



a)



b)

Fig. 11. Land cover changes (classes, Table 1) in a group of SLUAs from the Municipality of Lovech (Lovech District) falling within the territory of the Devetashko plateau (%): a) (Level 1); b) (Level 3)
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

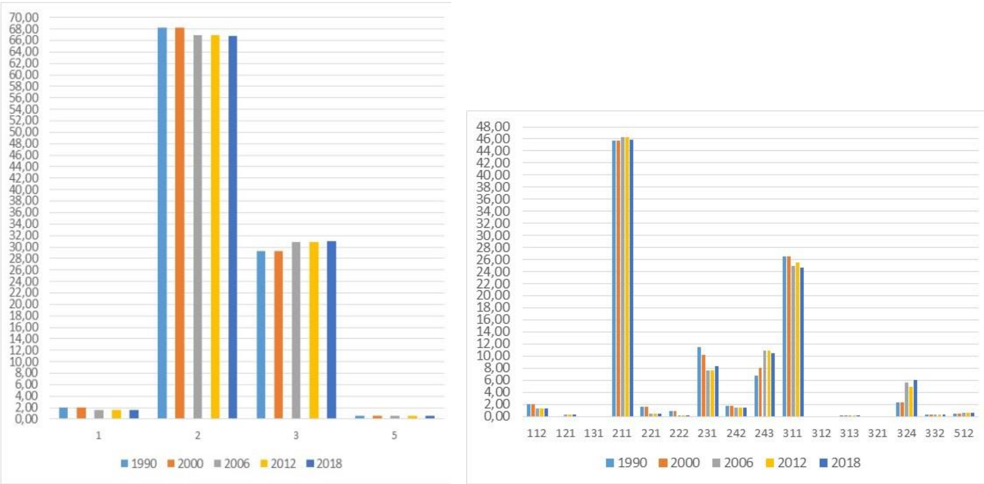
Lands from the Municipality of Sevlievo (Gabrovo District)

The SLUAs from the Municipality of Sevlievo (Gabrovo District), which fall within the limits of the plateau include in whole the SLUA of Agatovo village and partially that of Kramolin village (Fig. 4). Their total area amounts to 49.12 km^2 or 14.32% of the plateau's area. The population in both settlements numbers 604 people (2018). The drop with respect to 1990 (1 398 people) is 56.80% (Fig. 5).

On the SLUAs from the Municipality of Sevlievo, similar to the Municipality of Letnitsa, at level 1, the shares of *Agricultural areas* are significantly greater (more than 2 times) then those of *Forest and semi natural areas* (Fig. 12a). At this CORINE, land cover level, and insignificant changes are recorded over the years. The situation at level 3, however, is different. More essential changes take place there in both classes from level 1.

The number of presented classes also manifests changes and difference with respect to the plateau as a whole. In 1990 and 2000, they are 12, and from 2006 to 2018, the classes are 13 (Fig. 12b). *Mineral extraction sites* (131) from *Artificial surfaces* and *Coniferous forest* (312) and *Natural grasslands* (321) from *Forest and semi natural areas* are permanently lacking over the entire period. In the beginning of the period (1990 and 2000), lack of areas of *Industrial or commercial units* (121) is also identified.

Non-irrigated arable land (211) demonstrates exclusively high shares, retaining relatively close values over the entire period (Fig. 12b). The other classes from level 3 of *Agricultural areas*, which feature significantly lower shares display-varying dynamics. *Broad-leaved forest* (311) is the second ranking land cover, with lasting tendency for dropping down. With *Transitional woodland-shrub* (324), an increase is observed after 2000, but on the overall it is exceptionally weak.



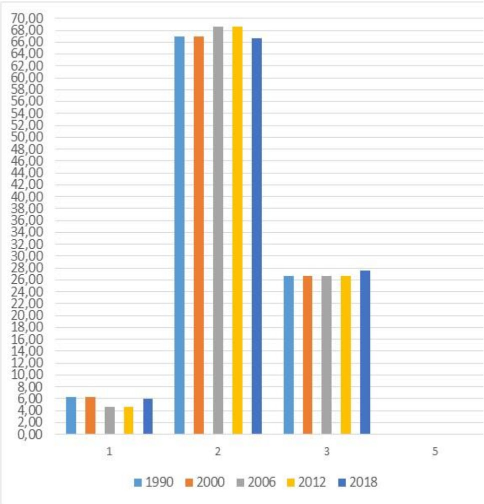
a) b)

Fig. 12. Land cover changes (classes, Table 1) in a group of SLUAs from the Municipality of Sevlievo (Lovech District) falling within the territory of the Devetashko plateau (%): a) (Level 1); b) (Level 3)
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

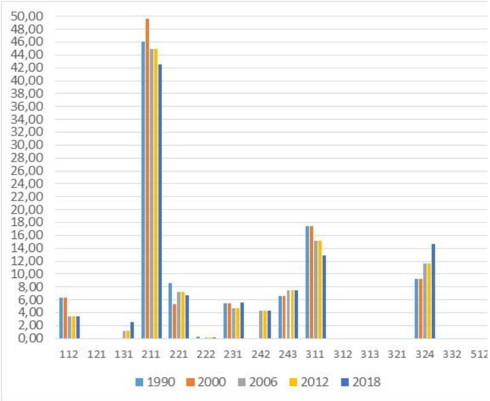
Lands from the Municipality of Pavlikeni (Veliko Tarnovo District)

The SLUAs from the Municipality of Pavlikeni (Veliko Tarnovo District), which falls within the boundaries of the plateau, include in whole the SLUAs of Dimcha village, and in part – of Varbovka village and the town of Byala Cherkva (Fig. 4). Their total area amounts to 22.83 km^2 or 6.65% of the plateau's area. The population of both settlements falling entirely within the plateau's periphery numbers 1 687 people (2018). The drop with respect to 1990 (2 034 people) is 17.06% (Fig. 5).

Similarly, with the SLUAs from the Municipality of Pavlikeni, as well as with those from the Municipality of Sevlievo, *Agricultural areas* are over twice greater than *Forest and semi natural areas* (Fig. 13a.). Over the entire observed period, they feature about 66%, whereas the changes during the years are insignificant. During the first four years, *Forest and semi natural areas* retain constant shares of 26.70% and only in 2018; they increase insignificantly to 27.49%.



a)



b)

Fig. 13. Land cover changes (classes, Table 1) in a group of SLUAs from the Municipality of Pavlikeni (Lovech District) falling within the territory of the Devetashko plateau (%): a. (Level 1); b. (Level 3)
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

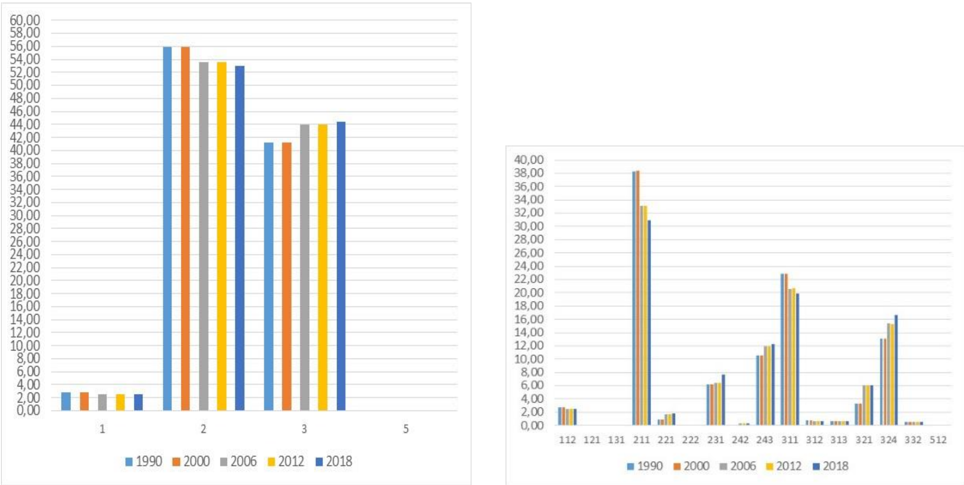
Over the entire observed period, six of the classes at level 3, characteristic of the whole plateau, are not present on the SLUAs of the Municipality of Pavlikeni. These are mainly classes from *Forest and semi natural areas*, namely, *Coniferous forest* (312), *Mixed forest* (313), *Natural grasslands* (321) and *Bare rocks* (332). No areas from the classes *Industrial or commercial units* (121) and *Water bodies* (512) are available, either.

Characteristic of this Municipality are the high shares of *Non-irrigated arable land* (211), which in 2000 reach 49.59%, irrespective of the subsequent drop in 2018 to 42.50%. The other classes from level 3 of *Agricultural areas* experience continuous changes in their shares. With *Forest and semi natural areas* and *Broad-leaved forest* (311), the drop is lasting, and vice versa, with *Transitional woodland-shrub* it displays a tendency for increase.

Lands of the Municipality of Suhindol (Veliko Tarnovo District)

The SLUAs from the Municipality of Suhindol (Veliko Tarnovo District), which falls within the boundaries of the plateau, include in whole the SLUA of Koevci village and in part – of the town of Suhindol and Gorsko Kosovo village (Fig. 4). Their total area amounts to 59.67 km^2, or 17.39% of the plateau's area. The population in both settlement numbers 1 742 people for 2018. The drop with respect to 1990 (2 034 people) is 43.55%.

In the SLUAs from the Municipality of Suhindol, the shares of *Agricultural areas* are approximately close to those of *Forest and semi natural areas*. This is the second municipality after Lovech, where such ratio between these two classes at level 1 is observed. The third class at level 1, which is presented by very small shares during the whole period, is *Artificial surfaces* (Fig. 14a). As a tendency, with *Agricultural areas,* there is a decrease of shares in 2006 and maintenance of constant shares during the years afterwards. With *Forest and semi natural areas,* it is vice versa – increase in shares in 2006 and maintenance of constant shares during the years afterwards.



a) b)

Fig. 14. Land cover changes (classes, Table 1) in a group of SLUAs from the Municipality of Suhindol (Lovech District) falling within the territory of the Devetashko plateau (%): a) (Level 1); b) (Level 3)
Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

The charactistic thing for the number of classes at level 3 from this subregion is the absence of one class from each class *Artificial surfaces* (incl. *Mineral extraction sites* (131), *Agricultural areas* (incl. *Fruit trees and berry plantations* (222)), and *Water bodies* (incl. Water bodies (512)) (Fig. 14b).

The observation of the shares at level 3 shows that *Non-irrigated arable land* (211) comes first with the greatest share. Second comes *Broad-leaved forest* (311), followed by *Transitional woodland-shrub* (324). The drop with *Non-irrigated arable land* is accompanied by an increase in *Vineyards* (221), *Pastures* (231) and *Land principally occupied by agriculture, with significant areas of natural vegetation* (243). Another regularity which is observed within *Forest and semi natural areas* is drop of *Broad-leaved forest* (311) and increase in *Natural grasslands* (321) and *Transitional woodland-shrub* (324) and retaining of the values of *Coniferous forest* (312) and *Mixed forest* (313) (Fig. 14b).

Change of the land cover classes

The changes in sub-period 1990–2000 comprise 14 SLUAs from five municipalities on the territory of the Devetashko plateau with total area of 7.62 km^2 (Table 2, fig. 15). For this period, eight groups of transformation of areas from one class into another at CORINE, land cover level 3 may be differentiated, most of them being the transformations in *Agricultural areas*.

The transformation from *Non-irrigated arable land* (211) into *Land principally occupied by agriculture, with significant areas of natural vegetation* (243) with an area of 2.41 km^2 comprises the SLUAs of strongly depopulated villages and the SLUAs of Aleksandrovo village, which is outside the plateau's territory. All of them pertain to the Municipality of Lovech.

Table 2. Corine, land cover change in Model karst area "Devetashko plateau" by classes and lands for the subperiods 1990–2000, 2000–2006, 2006–2012, 2012–2018. Source: [I.1.], Elaboration: D. Stefanova, G. Jelev

1990-2000		
Total area - 7,62 km2		
SLUAs	CLCC	Area (km2)
Aleksandrovo	211-243	0,00
Gostinya	211-243	0,00
Devetaki	211-243	0,71
Tepava	211-243	1,69
Varbovka	222-211	0,04
Dimcha	222-211	0,63
Suhindol	222-211	0,03
Dimcha	222-211	0,06
Gorsko Slivovo	231-211	1,89
Karpachevo	231-211	0,01
Agatovo	231-243	0,62
Chavdartsi	231-243	0,47
Brestovo	231-324	0,39
Tepava	231-324	0,42
Devetaki	243-211	0,30
Smochan	243-311	0,06
Krushuna	311-324	0,07
Karpachevo	311-324	0,22
Chavdartsi	311-324	0,01

2000-2006		
Total area - 3,84 km2		
SLUAs	CLCC	Area (km2)
Agatovo	311-324	0,69
Brestovo	311-324	0,02
Gostinya	311-324	0,11
Devetaki	311-324	0,79
Dimcha	311-324	0,30
Kramolin	311-324	0,42
Karpachevo	311-324	0,00
Suhindol	311-324	0,76
Tepava	311-324	0,08
Чавдарци	311-324	0,01
Agatovo	324-311	0,09
Devetaki	324-311	0,26
Krushuna	324-311	0,07
Karpachevo	324-311	0,06
Smochan	324-311	0,03
Chavdartsi	324-311	0,01

2006-2012		
Total area - 0,88 km2		
SLUAs	CLCC	Area (km2)
Kramolin	311-324	0,17
Kramolin	324-311	0,46
Suhindol	324-311	0,16
Koevci	324-311	0,08

2012-2018		
Total area - 4,82 km2		
SLUAs	CLCC	Area (km2)
Gorsko Slivovo	131-231	0,26
Suhindol	211-131	0,09
Varbovka	221-211	0,11
Agatovo	231-211	0,21
Brestovo	231-211	0,07
Varbovka	231-211	0,29
Gorsko Slivovo	231-211	0,00
Devetaki	231-211	1,30
Dimcha	231-211	0,00
Koevci	231-211	0,41
Karpachevo	231-211	0,49
Suhindol	231-211	0,12
Agatovo	311-324	0,19
Gorsko Slivovo	311-324	0,50
Gostinya	311-324	0,01
Devetaki	311-324	0,07
Kramolin	311-324	0,37
Karpachevo	311-324	0,01
Suhindol	311-324	0,12
Tepava	311-324	0,09
Devetaki	321-211	0,11

Another type of transformation during this period is from *Fruit trees and berry plantations* (222) into *Non-irrigated arable land* (211) with an area of 0.77 km^2, which comprise the SLUAs from the Municipality of Pavlikeni (Varbovka and Dimcha villages) and from the Municipality of Suhindol (town of Suhindol). When the Agrarian–Industrial Complexes (AICs) were liquidated, the orchards were abandoned.

The liquidation of cattle breeding during the period 1990–2000 resulted in the transformation of pastures with an area of 4.10 km^2. This process affects the SLUAs of Gorsko Slivovo and Karpachevo villages, with the transformation into *Non-irrigated arable land* (211), of Agatovo and Chavdartsi villages, with the transformation into *Land principally occupied by agriculture, with significant areas of natural vegetation* (243), and of Brestovo and Tepava villages with transformation into *Transitional woodland-shrub* (324).

Insignificant transformation of the area of 0.29 km^2 comprises *Broad-leaved forest* (311), which has been transformed into *Transitional woodland-shrub* (324) on the SLUAs of Chavdartsi, Krushuna, and Karpachevo villages.

The changes in sub-period 2000–2006 comprise 12 SLUAs from five municipalities on the territory of the Devetashko plateau with total area of 3.84 km^2 (Table 2, Fig. 15). The transformations have been differentiated into two groups: the transformation of *Broad-leaved forest* (311) into *Transitional woodland-shrub* (324), which comprises the SLUAs of 10 settlements (Agatovo, Brestovo, Gostinya, Devetaki, Dimcha, Kramolin, Karpachevo, Suhindol, Tepava, and Chavdartsi villages); reverse transformation from *Transitional woodland-shrub* (324) into *Broad-leaved forest* (311) (on the SLUAs of Agatovo, Devetaki, Krushuna, Karpachevo, Smochan, and Chavdartsi villages).

Sub-period 2006–2012 comprises 3 SLUAs from two municipalities – Kramolin village from the Municipality of Sevlievo and Koevci village and town of Suhindol from the Municipality of Suhindol, with total area of 0.88 km^2 (Table 2, Fig. 15). The transformations are from *Transitional woodland-shrub* (324) into *Broad-leaved forest* (311) and vice versa for Kramolin village – from *Broad-leaved forest* (311) into *Transitional woodland-shrub* (324).

The last sub-period 2012–2018 comprises 12 SLUAs from five municipalities with total area of 4.82 km^2 (Table 2, Fig. 15). The transformations have been differentiated into 5 groups, whereas in 2 of them the changes are more essential and comprise more SLUAs. On 9 SLUAs, *Pastures* (231) are transformed into *Non-irrigated arable land* (211), and on other 8 SLUAs – from *Broad-leaved forest* (311) into *Transitional woodland-shrub* (324).

The results from the performed analyses and the assessment of the land cover and land use of the Devetashko plateau after the CORINE nomenclature have been summarised, as follows:

- the land cover on the territory of the Devetashko plateau is presented with the greatest share by *Agricultural areas*, followed by *Forest and semi natural* *areas*, the difference between the two being about 20%. The values of *Artificial surfaces* are exceptionally low and those of *Water bodies* are minor. At this level, the observation of the share distribution by classes displays insignificant changes, which are in the direction of increase for *Forest and semi natural areas*, and vice versa – decrease for *Agricultural areas* (Fig. 6);

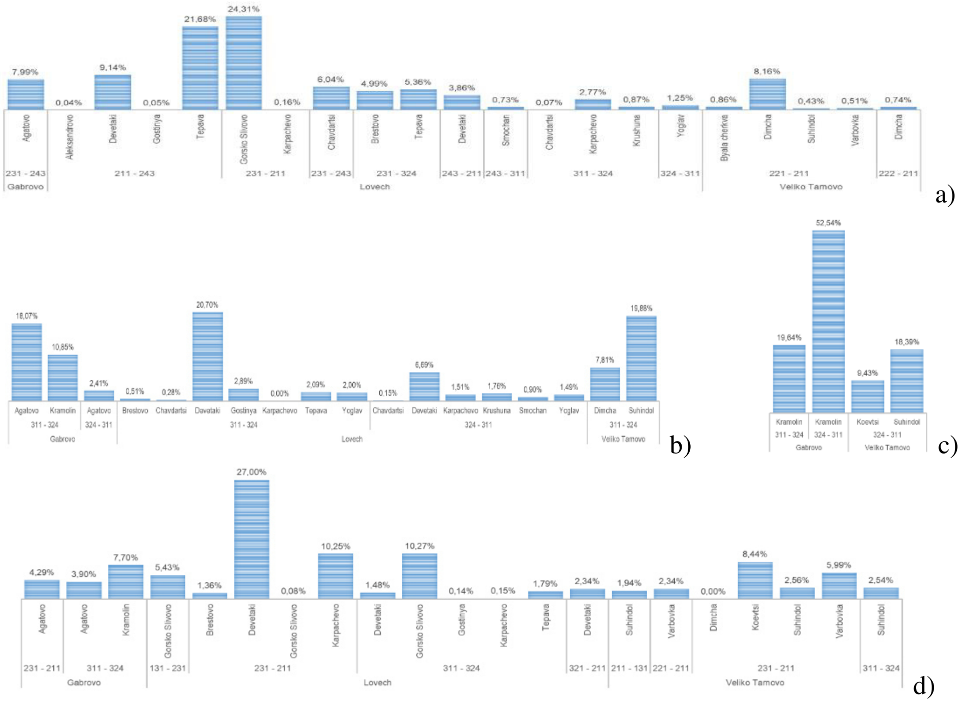


Fig. 15. Corine, land cover change in Model karst area "Devetashko plateau";
a) 1990–2000, b) 2000–2006, c) 2006–2012, d) 2012–2018)
Source: [I.1.], Elaboration: Georgi Jelev

- at CORINE, land cover level 3 on the territory of the Devetashko plateau, 16 land cover and land use classes are observed. *Non-irrigated arable land* (211) and *Broad-leaved forest* (311) feature the greatest relative shares, whereas the tendency with them is for the weak decrease. Another characteristic is that the changes are mainly concentrated in classes within *Agricultural areas* and *Forest and semi natural areas*;

- over the entire period, the changes of one class into another comprise 5% of the plateau's territory;

- with time, the changes in land cover follow the social and economic changes in society – from the restitution of land and the liquidation of agricultural farms to provide subsidies in agriculture, which again triggered changes in the share of the individual classes of *Agricultural areas*. Their specific manifestation is expressed in the fact that, during the early years of the „transition“, transformation of *Pastures* (231), *Non-irrigated arable land* (211), *Fruit trees and berry plantations* (222) is mostly observed, resulting from collapse in cattle breeding and abandonment of arable land. A change in the areas of *Broad-leaved forest* (311) is also observed resulting from intensified logging. Recent years witness the reverse tendency – with the development of large-scale agriculture, pastures and natural grass areas are transformed into agricultural land;

- at the level of a group of SLUAs, located within the boundaries of territorial and administrative units (municipalities) on the territory of the Devetashko plateau, two types of combinations of land cover are observed. On three municipal lands (Letnitsa, Sevlievo, and Pavlikeni), the shares of *Agricultural areas* outdistance significantly (almost two-fold with respect to *Forest and semi natural areas*). With the other two municipalities (Lovech and Suhindol), the shares of these two classes from level 1 are approximately equal. These means that, when implementing current and future economic activities, the municipal administrations and the persons handling the lands must keep a track of the current changes and forecast the upcoming changes and possible transformations by land cover classes. But, minding the specifics of these territories, they should also comply with karst's high vulnerability to anthropogenic impacts and climatic changes.

Conclusion

The land cover and land use types and their change as a result of anthropogenic pressure or natural intervention have a strong impact on karst geosystems. Relief character and a special distribution of karst regions influence economic activity and it, on its part, influences karst geosystems.

The social and economic factor also affects land use and thence, land cover change. The pressure on karst systems is further aggravated by various conflicts of interest arising during the exercising of different economic activities, such as farming (agriculture and cattle breeding), forestry, mining, settlement development, industry, protected territories. The new circumstances of transition and continuous changes require adequate management complying with the karst specifics of the territory, among others. The absence of such management would result in serious ecological problems. In this sense, observation of land use in the highly vulnerable karst territories and the possibilities for its planning require special attention. The studies based on land cover and land use, especially in sensitive territories, such as the karst ones, may provide additional valuable information, which might improve land management and assist the design of strategic development policies. Because of human activity, incl. traditional agricultural activity, which prevails on the Devetashko plateau results in the transformation of karst types [17] and structural and functional changes in karst geosystems [3]. These changes, on their part, reflect, sometimes unexpectedly, on the anthropogenic activity and social environment in karst territories.

Acknowledgement

This study was funded by the research project „Current impacts of global changes on the evolution of karst (based on the integrated monitoring of model karst geosystems in Bulgaria)", Science Research Fund, Grant No. ДН14/10/20.12.2017.

References

1. Ford, D., P. Williams. Karst Hydrogeology and Geomorphology. John Wiley and Sons Ltd., 2007, 562 p.
2. Andreychuk, V., P. Stefanov. Karst geosystems and principles of protection of Karst territories. GEOgraphy`, 2006, 21, 1, 5–11. (in Bulgarian)
3. Andrejczuk, W., P. Stefanow. Podstawowe zasady ochrony krajobrazu krasowego i jaskiń. Acta Geographica Silesiana, WNoZ UŚ, Sosnowiec, 2017, 11/1, 25, 5–10 (in Polish)
4. Popov, V. Razprostranenie na karsta v B"lgarija i njakoi negovi osobenosti. Izv. Geogr.. inst., BAN, T. HIII (HHIII), Sofija, 1970, 5–19. (in Bulgarian)
5. Zapletalova, J., Stefanova, D., Vaishar, A., Stefanov, P., Dvořák, P., and Tcherkezova, E. Social development of ecologically sensitive rural areas: Case studies of the Moravian Karst (Czechia) and the Devetashko Plato (Bulgaria). Problems of Geography, 2016, 3–4, 64–84.
6. Kostova, B., R. Berberova, R. Ergina, and R. Gjurov. Rodopi – nauchno-obrazovatelen marshrut za geolozhki turiz"m, Sbornik s dokladi Osmi mezhdunaroden simpozium „Ekologija – ustojchivo razvitie 2012". 2013, SU-Vraca. 117–127. (in Bulgarian)
7. Kostova, B., and R. Berberova. Geolozhki turiz"m po Isk"rski prolom, B"lgarija. Godishnik na departament „Prirodni nauki". 2015, NBU, 2016, 63–70. ISSN 2367–6302 (online) (in Bulgarian)
8. Kostova, B., and R. Berberova. Geolozhki turiz"m v rajona na Karlukovo, B"lgarija. Godishnik na departament „Prirodni nauki" 2016. NBU. 61–69. ISSN 2367–6302 (online) (in Bulgarian)
9. Kostova, B., R. Berberova, and R. Gjurov. Geolozhki osobenosti na Iztochnite Rodopi i vlijanieto im v"rhu njakoi kulturni aspekti. Sbornik s dokladi Mezhdunarodna nauchna konferencija „Ekologizacija 2012". NBU. 2018, I-VII. ISBN 978-619-233-048-4 (CD/DVD) (in Bulgarian)

10. Fan, Z., J. Li, T. Yue, X. Zhou, and A. Lan. 2015. Scenarios of land cover in Karst area of Southwestern China. Environ. Earth Sci. 2015, 74, 6407–20. Springer-Verlag Berlin Heidelberg, DOI:10.1007/s12665-015-4223-z
11. Dimitrov, V., Ju. Tepeliev, R. Koleva, Ju. Krumova, N. Ilieva, T. Ljubenov, B. Borisova. 2019. Monitoring na zemnata pov"rhnost po programata „Kopernik" prez 2017–2018 g. za B"lgarija - osnovni rezultati. Proceedings of the Fifteenth International Scientific Conference SES 2019, 146–51. e-ISSN 2603–3321
12. Jakucs, Laslo 1979. Morphogenetics of karst regions, Moskva, 244–46 (in Russian)
13. Petrov, Pavel. Karstovite blata po Devetashkoto plato. Izv. b"lg. geogr. d-vo, 1933, 1, 287–91. (in Bulgarian)
14. Popov, V., P. Penchev, and L. Zjapkov. Morfologija i hidrologija na karsta v severnata chast na Predbalkana mezhdu rekite Jantra i Os"m. Izv. Geogr. inst., 1965, 9, 69–91. (in Bulgarian)
15. Ninov, N. Soils. In: Geograpfy of Bulgaria, GI-BAS, „ForCom", Sofia, 2002, 277–315. (in Bulgarian)
16. Svetlik, I., J. Mizera, P. Stefanov, and M. Lazarova: Application of radio-carbon dating in tracing land-use on karst terrains (as exemplified with the karst bog of Irmanov gyol, the Devetaki plateau, North Bulgaria). – Abstracts of the International Scientific Conference "Protected Karst Territories: Condition, Problems and Perspectives", October 2005, Shumen, 7–8.
17. Gvozdeckij, N. A. Problemy izuchenija karsta i praktika. Moskva, Mysl', 1972, 391. (in Russian)

Internet sources

I.1. Corine, Land Cover (CLC) for 1990, 2000, 2006, 2012, 2018 and Land Cover Changes (LCC) 1990–2000, 2000–2006, 2006–2012, 2012–2018. Available at https://land.copernicus.eu/pan-european/corine-land-cover [Accessed: 01 November 2019]
I.2. Informacionna sistema za zashhiteni zoni ot ekologichna mrezha Natura 2000. Available at http://natura2000.moew.government.bg/Home/Natura2000Protected Sites [Accessed: 01 November 2019] (in Bulgarian)
I.3. ASTER Global Digital Elevation Map Announcement. Available at https://asterweb.jpl.nasa.gov/gdem.asp. [Accessed: 01 November 2019]

ПРОМЕНИ В ЗЕМНОТО ПОКРИТИЕ И ЗЕМЕПОЛЗВАНЕТО НА КАРСТОВ РАЙОН ДЕВЕТАШКОТО ПЛАТО

Г. Желев, Д. Стефанова, П. Стефанов

Резюме

Данните на CORINE Land Cover (CLC) представляват цифрови данни за земното покритие, разпределено в 44 класа, като за територията на България класовете са 36. Минималната картируема единица е 25 ha (за площни обекти) и 100 m за линейни обекти. Налични са набори от данни за годините 1990, 2000, 2006, 2012 и 2018 г., както и за промените настъпили между всяка двойка години (1990–2000, 2000–2006, 2006–2012 и 2012–2018).

Големият набор от данни дава възможност да се проследи за период от близо 30 години как се е променяло земното покритие в моделни карстови райони, които са силно уязвими на антропогенни и природни въздействия.

Настоящата статия разглежда промените в типовете земно покритие в Devetashko plateau – типично карстово плато в Северна България. Промените в земното покритие и в земеползването пряко влияят върху процесите на съвременния карстогенезис, почвено-растителната покривка, количеството и качеството на подземните карстови води.

[2] The settlements located on the other five lands included in the Devetashko plateau are at a great distance from the plateau and their population has not been included in the population of the plateau. The tendencies with them are the same.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e06

SEASONAL CHANGES OF SAHARA DESERT DUST TRANSPORT OVER BALKANS

Maria Dimitrova

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: maria@space.bas.bg

Keywords: *Remote Sensing, Air Pollution, Dust Pollution, Ecology*

Abstract

This article presents an investigation of seasonal behaviour of the Sahara desert dust transport over the Balkans. The data used are satellite measurements of monthly averaged Absorption Aerosol Index (AAI) value. The research period is from June 1995 to the end of 2019. The data used is from four space instruments onboard five satellites. The area of interest is a rectangle with corners 23 E 43 N and 245 E 35 N. The data from different sources are compared and discussed.

Introduction

Mineral dust is the second largest source of natural aerosols. North African deserts emit most of the dust particles released into the atmosphere worldwide.

The smallest dust particles strongly affect human health (mainly the respiratory system).

The physics, chemistry and biology of marine atmosphere and biogeochemical cycles in seawater are strongly influenced by the atmospheric transport and deposition of mineral aerosol particles, which are massively exported from the desert areas. Sahara and the peripheral regions are the main source of soil-derived aerosols over the Atlantic and Mediterranean.

Desert dust transports minerals, bacteria and other small pollutants. Cristal aerosols influence the atmospheric radiative balance through scattering and absorption, and by acting as cloud condensation nuclei when sulphation and nitration occur.

It is now well known that satellite observations of the sunlight reflected by the Earth-atmosphere system enable relatively accurate retrieval of the vertically integrated dust aerosol content over the ocean in terms of optical thickness [1, 2]. Remote sensing by operational meteorological onboard sensors is suitable for monitoring large desert dust plumes, which exhibit high temporal and spatial variability.

These dust plumes are spread over thousands of kilometers, and can persist for many days. They have extremely deleterious effects on the air quality near the African continent but are also important on a global scale.

Some authors during the last decades investigate dust transport over the Mediterranean [1–6], more often in direction towards Italy and Spain [7] and even more scientists paid attention to dust transport over the Atlantic [8–11]. Results indicate that during "African dust-events", the numbers of cultivatable airborne microorganisms can be 2 to 3 times more than these found during "clear atmospheric conditions" [12, 13]. There are no exhaustive investigations done on the influence of the Sahara dust over the Balkans.

In this paper, we use monthly averaged satellite data of AAI (Absorption Aerosol Index), which has a close to linear dependence with optical depth from five satellites for the period 1995 until the end of 2019 to investigate seasonal behaviour over the Balkans in nine areas in direction North – South.

Area of interest, satellite data and investigation methods

In this work we research seasonal behaviour (and its variations) of dust atmospheric pollution from Africa in North – South direction over the Balkans. invested area is rectangle with corners respectively 23 E 43 N and 245 E 35 N.

We choose nine areas (cells of 1 × 1 degrees). Points indicate the center of each area (Fig. 1).

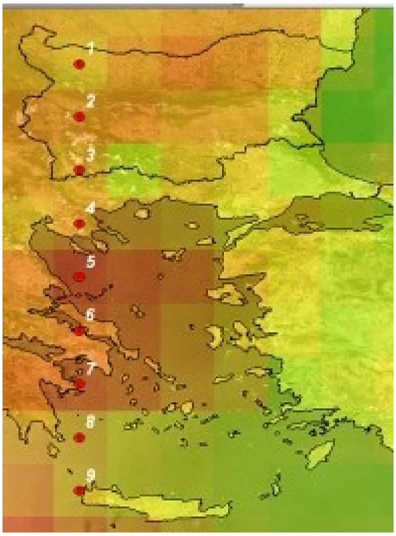


Fig. 1. Area of interest. On the left – optical image from MODIS with points. On the right – combined image AAI from GOME-2 over the same optical picture and points.

We use satellite data from four space instruments onboard five satellites with similar space and temporal resolution as follows:

- GOME onboard ERS-2 – from 6.1995 till 6.2003 (40 × 40 km)
- SCIAMACHY onboard ENVISAT – from 9.2002 till 4.2012 (30 × 60 km)
- OMI onboard AURA – from 9.2004 till now (40 × 40 km)
- GOME-2 onboard MetOp A – from 1.2007 till now (40 × 40 km)
- GOME-2 onboard MetOp B – from 12.2012 till now (40 × 80 km)

On fig. 2 it the working duration of each one of these instruments is shown.

In the figure, we show GOME-2 instrument onboard MetOp C satellite. We do not include data from them later on because data are available only for one year.

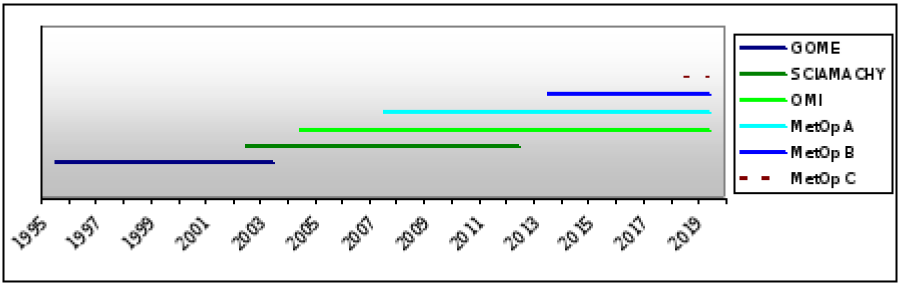


Fig. 2. The working duration of chosen space instruments

We choose to use monthly averaged data for AAI in "*.nc" format (Longitude: 360 bins centered on 179.5 W to 179.5 E (1 degree steps), Latitude: 180 bins centered on 89.5 S to 89.5 N (1 degree steps) from ESA - TEMIS (Tropospheric Emission Monitoring Internet Service) [14].

There is free data for almost every day during the work time of each one of the chosen instruments.

The Absorbing Aerosol Index (AAI) indicates the presence of elevated absorbing aerosols in the Earth's atmosphere. The aerosol types that are mostly seen in the AAI are desert dust and biomass-burning aerosols. The AAI provided here are derived from the reflectance measured by GOME-1, SCIAMACHY and GOME-2 at 340 and 380 nm, and from the reflectance measured by OMI at 354 and 388 nm.

Results

For every instrument and each year it's working we built graphics of dependence AAI – month of year. Then, we obtained an average value of AAI for each area and month.

We do not show results for the 5th and the 8th areas, because these areas are fully over sea regions. Our interest is focused on the spatial dust transport distribution over land areas. Results we show in Figures 3–7.

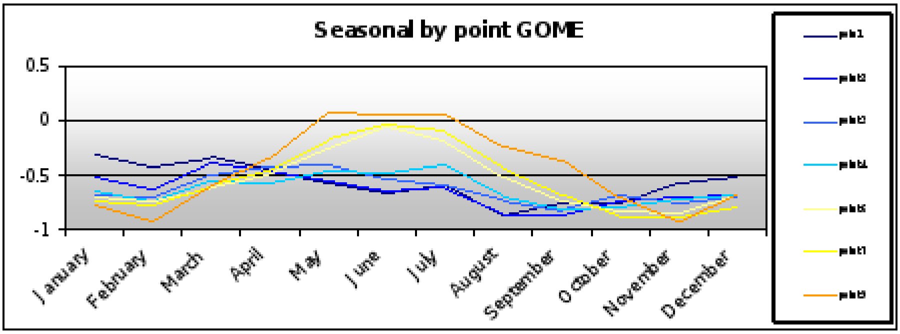


Fig. 3. Seasonal behaviour of averaged AAI by points from GOME instrument

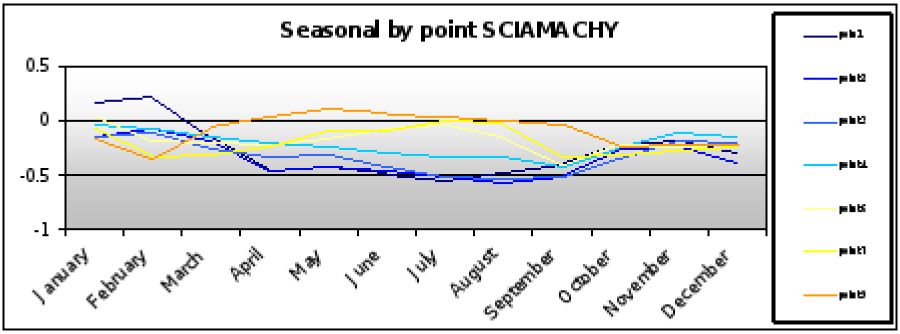


Fig. 4. Seasonal behaviour of averaged AAI by points from SCIAMACHY instrument

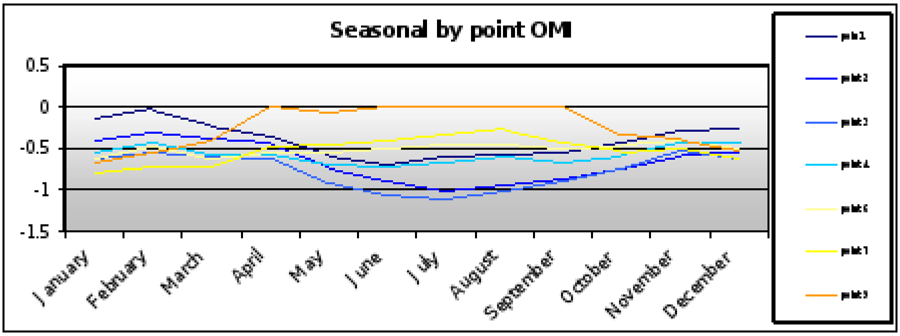


Fig. 5. Seasonal behaviour of averaged AAI by points from OMI instrument

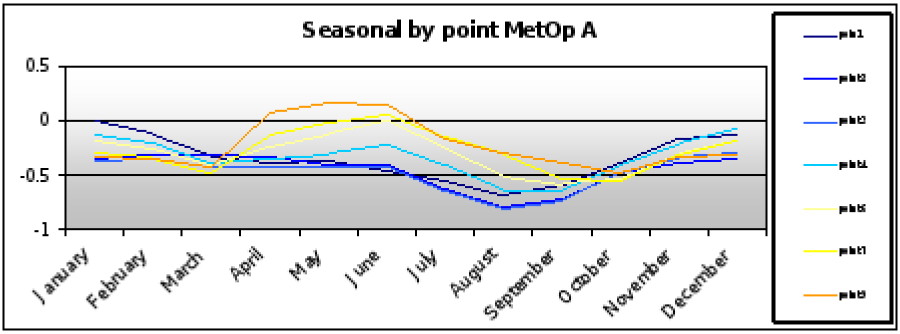


Fig. 6. Seasonal behaviour of averaged AAI by points from GOME-2 instrument onboard MetOp A satellite

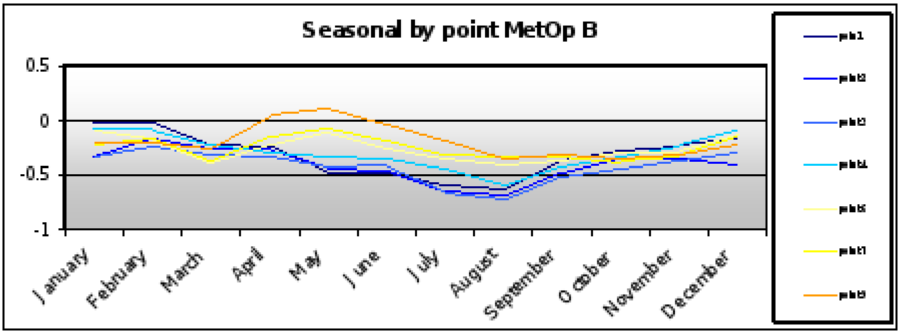


Fig. 7. Seasonal behaviour of averaged AAI by points from GOME-2 instrument onboard MetOp B satellite

As it is seen from the above figures, seasonal AAI behaviour is different for different areas. Areas from 1st to 4th show an increase of AAI for winter months, while areas 6th, 7th and 9th show significant increases in AAI during summer. Moreover, – AAI increase with an increase in the point number. As our previous investigations show [15], sand storms from Africa during the last 15 years are mainly in winter months (from February till June). Results from the older instruments (GOME and SCIAMACHY) show a longer period with increase – till October.

As the area number increases with the decrease of Latitude, it shows increasing of desert dust with Latitude decreases. For illustrating above conclusion and comparison of data from different instruments, on fig. 8 we show averaged AAI value for each area.

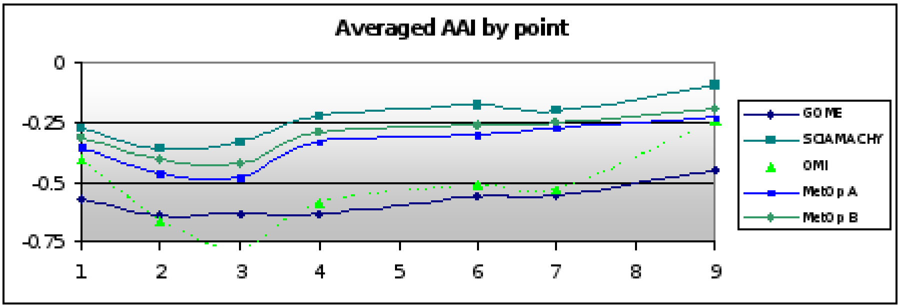


Fig. 8. averaged AAI value for each instrument and different areas

As it is seen from the last above graphic - averaged AAI for last area has the highest value. AAI differences in areas are similar for GOME-2 and SCIAMACHY instruments. Data from GOME show similar behaviour but lower values.

It seems that the combination of data from OMI with data from other mentioned instruments is not corrected. The reason for it may be longer and non-stable work of the satellite and the different spectral diapason of AAI measurement of this instrument.

Conclusions

As a result of this research, we can conclude that satellite data for AAI show the expected increase in south direction. It means that the dust transport affects the southern Balkan areas (here Crete) stronger then the northern Bulgarian regions. The seasonal behaviour shows changes during the last 15 years. While earlier sand transport continues from February until October, during the last decade it shows increase only from March to July with outstanding maximum around May.

References

1. Moulin, C. and F. Guillard, Long-term daily monitoring of Saharan dust load over the ocean using Meteosat ISCCP-B2 data 1. Methodology and preliminary results for 1983–1994 in the Mediterranean, Journal of Geophysical Research, 1997, 102, D14, 16947–958.
2. Moulin, C., C. E. Lambert, U. Dayan, T. M V. Masson, M. Ramonet, P. Bousquet, M. Legrand, Y. J. Balkanski, W. Guelle, B. Marticorena, G. Bergametti, and F. Dula, Satellite climatology of African dust transport in the Mediterranean atmosphere, Journal of Geophysical Research, 1998, 103, Dll, 13137–144.
3. Kubilay, N., Sl. Nickovic, C. Moulin, and F. Dulac, An illustration of the transport and deposition of mineral dust onto the eastern Mediterranean, Atmospheric Environment, 2000, 34, 1293–1303.
4. Querol, X., J. Peya, M. Pandolfi, A. Alastuey, M. Cusack, N. Peŕ rez, T. Moreno, M. Viana, N. Mihalopoulos, G. Kallos, and S. Kleanthous, African dust contributions to mean ambient PM10 mass-levels across the Mediterranean Basin, Atmospheric Environment, 2009, 43, 4266–77.
5. Pey, J., X. Querol, A. Alastuey, F. Forastiere, and M. Stafoggia, African dust outbreaks over the Mediterranean Basin during 2001–2011: PM10 concentrations, phenomenology and trends, and its relation with synoptic and mesoscale meteorology, Atmos. Chem. Phys., 2013, 13, 1395–1410.
6. Salvador, P., S. Alonso-Pérez, J. Pey, B. Artíñano, J. J. de Bustos, A. Alastuey, and X. Querol, African dust outbreaks over the western Mediterranean Basin: 11-year characterization of atmospheric circulation patterns and dust source areas, Atmos. Chem. Phys., 2014, 14, 6759–75.
7. Blanc, F. De Tomasi, E. Filippo, D. Manno, M. R. Perrone, A. Serra, A. M. Tafuro, and A. Tepore, Characterization of African dust over southern Italy, Atmos. Chem. Phys., 2003, 3, 2147–59.
8. Kevin, D. Perry, Thomas A. Cahill, Robert A. Eldred, and Dabrina D. Dutcher, Long-range transport of North African dust to the eastern United States, Journal of Geophysical Research, 1997, 102, D10, 11225–238.
9. Joseph, M. Prospero, Long-term measurements of the transport of African mineral dust to the southeastern United States: Implications for regional air quality, Journal of Geophysical Research, 1999, 104, DI3, 15, 917–27.
10. Sebastian Engelstaedter, Ina Tegen, Richard Washington, North African dust emissions and transport, Earth-Science Reviews 2006, 79, 73–100.
11. Y. Ben-Ami, I. Koren, and O. Altaratz, Patterns of North African dust transport over the Atlantic: winter vs. summer, based on CALIPSO first year data, Atmos. Chem. Phys., 2009, 9, 7867–75.
12. Dale, W. Griffin, Virginia H. Garrison, Jay R. Herman & Eugene A. Shinn, African desert dust in the Caribbean atmosphere: Microbiology and public health, Aerobiologia 2001, 17, 203–13.
13. Griffin, D., W., D. L. Westphal, and M. A. Gray, Airborne microorganisms in the African desert dust corridor over the mid-Atlantic ridge, Ocean Drilling Program, Leg 209, Aerobiologia, 2006, 22, 211–26.
14. ESA – TEMIS - http://www.temis.nl/index.php

15. Dimitrova, M., P. Trenchev, E. Georgieva, N. Neykova, R. Neykova, R. Nedkov, D. Gochev, D. Syrakov, B. Veleva, D. Atanassov, and T. Spassova, Seasonal changes of aerosol pollutants over Bulgaria, Proceedings of the Fifteenth International Scientific Conference Space, Ecology, Safety, SES 2019, (Print ISSN 2603–3313; Online ISSN 2603–3321), 2019, 241–52.

СЕЗОННО ИЗМЕНЕНИНЕ НА РАЗПРОСТРАНЕНИЕТО НА ПУСТИНЕН ПЯСЪК ОТ САХАРА НАД БАЛКАНИТЕ

М. Димитрова

Резюме

Статията представя едно изследване на сезонното поведение на пустинен пясък от Сахара, който се разпространява над Балканите. Използвани са спътникови данни за месечните стойности на AAI. Данните са от четири инструмента на борда на пет спътника за периода от юни 1995-та до края на 2019-та. Изследване е правоъгълна област с ъгли съответно 23 E 43 N and 245 E 35 N. Данните от различните инструменти са сравнени и дискутирани.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e07

APPLICATION OF GNSS AND SAR DATA IN LANDSLIDE MONITORING ALONG THE BLACK SEA COAST OF BULGARIA

Mila Atanasova[1], Hristo Nikolov[2], Keranka Vassileva[3]

[1]*National Institute of Geophysics, Geodesy and Geography – Bulgarian Academy of Sciences*
[2]*Space Research and Technology Institute – Bulgarian Academy of Sciences*
[3]*Faculty of Physics, Sofia University "St. Kliment Ohridski",*
e-mail: mila_at_zl@abv.bg

Keywords: *GNSS Data, SAR Data, Black Sea Coast, Landslide Processes*

Abstract

Landslide processes are considered the major part of the natural hazards occurring on the northern part of the Bulgarian sea side. Their monitoring can be done with high precision using GNSS data. The objective of this study is to provide solid grounds for monitoring of the landslide processes using GNSS and SAR data. This goal will be achieved by the implementation of the following: 1) establishment a verified methodology for extracting high-quality information from SAR images aimed at continuous monitoring of landslide areas integrating InterFerometric Images (IFI) and GNSS data and 2) creation of a working prototype of an information system for monitoring and prevention of the effects of earth crust movements (landslides, falls, etc.) based on freely accessible data provided by ESA and national sources. One of the scientific tasks to be solved includes the development of methodological approaches for comparison of the results from combined processing of interferometric images from SAR, measurements at permanent GNSS stations of the national NIGGG network in the area of study and geodetic measurements of a newly established test network covering a specific area on the Northern Black Sea coast of Bulgaria with active landslide processes.

Introduction

The main objective of this research is to monitor the ongoing landslide processes by complementary use of Synthesized Aperture Radar (SAR) and Global Navigation Satellite Systems (GNSS) data. It will be achieved by means of proved methodology for continuous monitoring of landslide areas by integrating information from interferometric images (IFIs) and GNSS data from permanent and local geodetic networks. The outcome of this study will deliver reliable information for the ongoing risky geo-processes for the region of the Northeastern Bulgaria, which is known with several large active landslides. These results could provide a better understanding of the origin and dynamics of current landslide processes as well as assessing the resulting hazards by creating thematic maps. The selection of the mentioned area is based on information provided by the national authority responsible for the monitoring of the landslides according to which their number has almost doubled over the past two years.

In this study, the selected landslides are located at the coastal zone of NE Bulgarian Black Sea. The selected region is based on the fact that until now more than 120 landslide events have been registered in the districts of Varna and Dobrich, and 80% of landslides affect the coastal line. As a contributing factor provoking the interest of this specific area is the weak stability of its slopes, which can be impaired even by single events such as abrasion, erosion, prolonged rainfalls, seismic, etc.

Fig. 1. Map of active, stabilised and potential landslides in NE Bulgaria (according to GIS of the Ministry of Regional Development and Public Works (MRDPW) (http://gis.mrrb.government.bg/)

At present, the authors have investigated a short strip along the Black Sea starting from city of Varna and ending at *Kaliakra* cape. The area is about 80 km long, where large numbers (more than 120) of irregularly distributed landslides have been registered and monitored by the competent authorities of MRDPW (Fig. 1). From the geological perspective, most of the landslides located in this region are of complex nature [1, 2].

Obtaining reliable data and delivery of an adequate information about the ongoing risky geo-processes of the Earth's surface is a key factor in tracking the origin and dynamics of landslide processes as well as assessing the resulting threats to the population and infrastructure. This is the main reason for the development and implementation of a rapid and accurate method for their operational monitoring. In this paper proposed is a possible solution for estimation of the ground motions caused by landslides based on information derived from differential interferometric processing of SAR data. It was proved that the Earth's crust deformations can be registered with the magnitude of centimetres by this method [3]. Results from SAR data processing unambiguously indicate these movements in certain zones of the investigated areas, but GNSS measurements are also required to obtain validated information. The main objective of this research is formulated as follow – to study of the landslide processes by using innovative methods. It was achieved through the extraction of high-quality and reliable information from SAR images focused on regular monitoring of landslide areas and integrating IFIs and GNSS data. Based on the freely available SAR data and the software provided by ESA, as well as data from national sources, a prototype of an information system was elaborated on the basis of archive IFIs produced for monitoring the movements of the Earth's crust (landslides, subsidence, etc.).

In the past decades the northeastern Black Sea coast of Bulgaria has been the subject of scientific research and geodetic surveying by the former “Laboratory of Geotechnics” BAS and "Central Laboratory of Higher Geodesy" today, the Department of Geodesy. By the mid-1990s a geodynamic network was built for monitoring the landslide processes around the town of Balchik [4]. Due to intensive construction activities of the coastal area in the last 20 years and the lack of funding for research and its maintenance, some of the points of the mentioned network (pillars for precise instrument positioning) were destroyed and do not exist. Nevertheless, an extensive and comprehensive methodology for geodetic investigations of landslides was developed [5]. In it underlined is the significance of the type, size and form of the landslide, the velocity of the displacements, the availability of stable areas in the vicinity of the study object, a type of equipment to be used according to the necessary precision.

Materials and Methods

Synthetic Aperture Radar (SAR) Data and DInSAR Processing

In this section, a short description of the SAR data and the processing steps used to produce interferometric images (IFIs) that were combined with the geodetic data in studying landslides is presented. The SAR data that have been processed are from a constellation comprised two satellites (A and B) of the Sentinel-1 mission freely distributed by ESA and were obtained from Scientific Data Hub maintained by ESA [6]. Each satellite has revisiting time of 12 days, which means that one and the same area is imaged every 6 days by a satellite. The SAR instrument onboard the satellite is C-band radar (corresponding to a wavelength of 5.56 cm) with a right-looking line of sight (LOS) regardless of the orbit direction. It operates in four acquisition modes stripmap, interferometric wide swath (IW), extra-wide swath and wave mode. For interferometric processing single look complex (SLC) data obtained in IW mode must be used since in it not only the amplitude of the backscattered signal is available, but its phase signal too. The phase signal is of crucial importance since after appropriate processing it delivers information about the changes in the distance to the objects on the Earth's surface between two satellite overpasses [3].

The DInSAR (differential interferometric SAR data processing) is a method that uses SLC SAR data to produce topographic and surface motion maps based on information contained in the interferometric phase. It is based on the acquisition of complex-valued data over the same area at different times and uses the difference found in the phase signal to detect the horizontal/vertical changes caused by ground deformations. Since in the measured phase, there are two components – one corresponding to the distance of a single object from the surface and the next reflecting the phase changes caused by the environment – a measure of quality, known as coherence, is introduced, which is an estimate of the noise level present in the phase signal. It is widely accepted that for the single pixel from the phase band of the produced IFI to be considered reliable the same pixel in the coherence band should have value above 0.3. Low values of this parameter are due to many external factors such as the state of the troposphere at the time of acquisition, the position of the satellites in their "orbital tube", which defines perpendicular baseline, presence of vegetation in the area of study, etc.

Fig. 2. Shape of the SAR images from the ascending (A) and descending orbits (B)

It needs to be underlined that the information provided by the phase signal is a relative concerning one of the SAR images, often called "master", and concerning a point on the ground, which is assumed to be stable. One more thing that needs to be addressed here is that all detected deformations are measured along the LOS of the SAR instrument and for this reason additional calculations are needed for properly combining GNNS and SAR data to obtain vertical ground deformations.

A drawback of the SAR data that need mention is that they cannot detect ground changes along the track of the satellite, which results in better registering movements in the east-west direction, then in north-south. This can be overcome by combining information derived from IFIs from both ascending and descending orbits (Fig. 2). An advantage offered using the DInSAR method is the possibility to register ground changes over large or difficult-to-access areas, thus delivering more information than by in situ GNSS acquisitions. This does not mean that it can completely substitute terrain measurements, but rather to provide details on the surface movements for larger areas in the investigated region.

The processing of SAR data for the production of a deformation maps includes the following steps – precise co-registration based on the orbital data of the two images used in the IFI, formation of the interferogram, filtering and speckle reduction, phase unwrapping, and geocoding. The most important step in this procedure is the phase unwrapping since only after it the information contained in the phase signal of |the IFI is converted into ground displacements. At this step, by integrating the phase difference between neighbouring pixels at every 2π cycles the difference in altitude in LOS is generated after any integer number of altitudes of ambiguity has been deleted.

GNSS for landslide monitoring

Geodynamic networks established for landslides monitoring generally consist of two types of points – reference or fixed points located on geologically stable terrain and survey points located within the investigated landslide.

To accomplish the objectives of this study geodetic data from these two types of points are necessary. Data from the stable points situated in non-deformable zone are provided by the permanent GNSS network. The newly established points that from geodynamic networks located inside the specific landslide will be measured in few GNNS acquisition cycles. The deformation analysis of those geodynamic networks will be done after the third measurement cycle by applying an appropriate approach [7].

First results and discussion

The first step to achieve the main objective was to create a local archive with Sentinel-1A/B images for the region of Northeastern Bulgaria, consisting of about 300 SLC images. For mapping deformations in the region of interest interferometric images at intervals of 4 and 8 months were produced.

The stated time intervals were used, since one of the main factors affecting the quality of the produced IFIs is the vegetation and for this reason only autumn and spring scenes were processed. Also most of the landslides activations occur in the said seasons caused by underground water-level change. Another factor that should be accounted before producing IFI is the presence of snow – used data are from days with no snow coverage [8].

In Fig. 3 presented are the displacements found in the IFI produced from image pair from dates Nov 26th, 2014 and Dec 27th, 2015. On it the detected displacements are colored in purpura and coincide with the landslide regions where most of the registered landslides are located (Fig. 5).

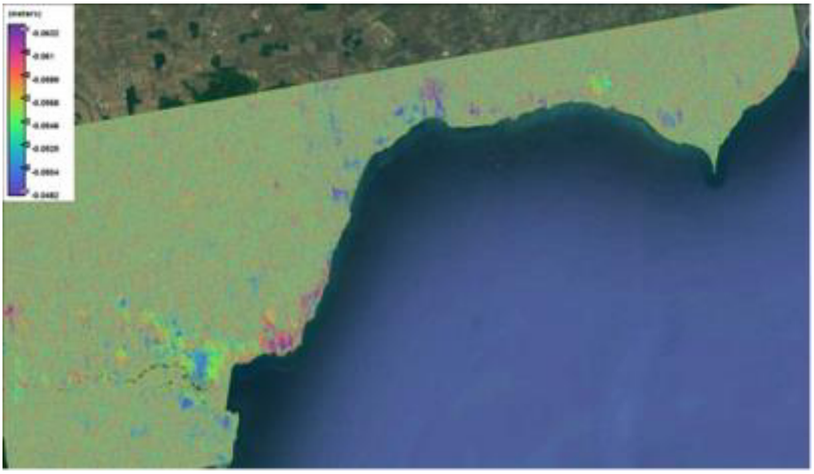


Fig. 3. Displacements obtained from IFI 26Nov2014-27Dec2015

In Fig. 4 are shown the displacements obtained from the IFI Jan 1st, 2015 – Dec 21st, 2016. The color of the pixel represents the surface movement in metric units for the investigated period ranging from dark blue to purpura. Particularly, vulnerable landslide areas are shown in purpura, less vulnerable in yellow and green. The landslide activity assessment resulting from this research shows (Fig. 4) that the subsidence range from –48 mm to –69 mm.

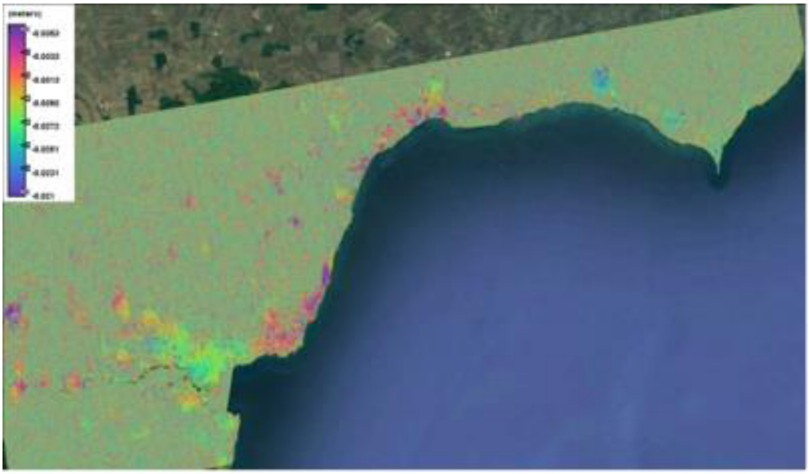


Fig. 4. Ground displacements obtained from the IFI Jan 1st 2015 – Dec 21st 2016

Fig. 5. Raster heat map based on displacement values at the points of landslides registered

The area of interest of this study is the one marked by red quadrate in which the concentration of ground deformations has been observed (Fig. 5). This area is called "*Dalgiya yar*" - a landslide circus in which several active landslides are located. Since their boundaries overlap, it is difficult to differentiate them from each other. Even for some of the investigated landslides located in this area a smaller landslide could be delineated inside them. This phenomenon can be seen in Fig. 6. In this figure the registration codes of every landslide are shown as they appear in the landslide register maintained by MRDPW and the boundaries of the separate landslides are shown in different colours.

The landslide "*Fara*" located between the village of Kranevo and the touristic resort "*Panorama*" covers only the low stage of the circus (Fig. 6). This landslide assigned identification number VAR 02.54145-01-17 in the register of landslides in the Republic of Bulgaria activated on October 13th 2012, destroying the lighthouse and villas. The landslides and collapses in this area that activated in 2013 are largely attributed to the human activities that took place in the last

Twenty years mainly the illegal construction, as well as to the fact that the requirements for civil engineering were not respected. For example, instead of building small bungalow houses two- and three-storey buildings were erected. In some of them, swimming pools have been built whose waters flow down the slope of the landslide. Those flows had very serious impact, as the water from the said pools flows down the slope where there is no drainage. Water supply network accidents often occur there, because landslides that tear water mains slip and, in turn, water – regardless of its origin (from the water supply system, rains, or pools) leads to activation and development of landslide and collapse processes.

Fig. 6. ID codes from MRDPW register and the boundaries of the investigated landslides in the area of "Fara" landslide

In this research, the authors set their attention on a landslide located in the investigated area – "Dalgiya yar" – “Fara”. This specific object was selected since only for it data from previous geodetic measurements and geological observations were made available. This fact made possible the comparison between the in situ data and data obtained from satellites. For landslide “Fara” two measurement cycles were carried out in years 2013 and 2018 of a network consisting of 8 points located on the road I-9 above the landslide. The geodetic data of these points were obtained from Geozashtita Varna Ltd. According to the Contract № RD-02-29-372/11.11.2013 under the Law on Public Procurement for Execution assigned by the Ministry of Regional Development to the contractor "Survey Group" Ltd. with subject "Geodetic survey of landslides between bus stop Fara and bus stop Obzor, and Kranevo village" (Fig. 7) [9–11].

Fig. 7. Geodetic network used by Survey Group Ltd.
for determination of displacements along the road (colored in blue)

Fig. 8. Points of the new geodynamic network in and around the landslide "Dalgiya yar"

For this specific study a geodynamic network covering the landslide area "*Dalgiya yar*" – "*Fara*" (Fig. 8) was established by the authors. It consists of 30 stabilised points with some being metal pipes 35 cm long, while other are metal bolts nailed in the rock. The GNSS measurements were carried with 2 receivers of type CHC i80 GNSS with horizontal precision 2.5 mm + 0.1 ppm RMS and vertical 3.5 mm + 0.4 ppm RMS and 1 receiver - P3E GNSS sensor used for reference station. The static mode was applied for the GNSS measurements. The two stations of the permanent GNSS network (Varna, Krushartsy) maintained by the private company GeoVara Ltd. and located in the non-deformable zone of the landslide were used as stable points. The newly established geodynamic network (Fig. 8) located inside the landslide will be measured once a year. In the mentioned geodynamic network "*Dalgiya yar*" all old six points that were found on the terrain from the network used to monitor deformations along the road (blue colored in Fig. 7 and Fig. 8) are included. The first measurement cycle of the geodynamic network was carried out on June 19–23, 2019. Thus, it is the third cycle of measuring the deformations along the road. The results of the GNSS measurements were processed using "CHC Geomatics Office 2" software in the coordinate system WGS84. Since the point coordinates from the measurements made in 2013 and in 2018 are in the national geodetic system 1970, K-7, the software "*BGSTrans*" was used to transform the points obtained in 2019 into it. This process was done in order to able to use the data in the next processing.

Fig. 9. Horizontal displacements for the period 2018–2013

Fig. 9 illustrates the horizontal displacements of 8 points along the road I-9. The largest displacements for the area under consideration were found in the southwest part of the site. The horizontal displacements shown in Table 1 were obtained from the measurements conducted in years 2013, 2018 and 2019.

Table 1. Horizontal (ΔD) and vertical (ΔH) displacements for the periods 2018–2013 and 2019–2018

Point number		***2018 – 2013***		***2019 - 2018***	
old	*new*	*ΔD, [m]*	*ΔH, [m]*	*ΔD, [m]*	*ΔH, [m]*
2008	*101*	*0.607*	*–0.088*	*0.085*	*–0.056*
375	*102*	*1.034*	*–0.465*	*0.094*	*–0.072*
2002	*103*	*0.648*	*–0.375*	*0.148*	*–0.054*
1004	*104*	*0.242*	*–0.068*	*0.105*	*–0.074*
2004	*105*			*0.092*	*0.019*
1005		*0.164*	*–0.038*		
1003		*0.159*	*0.017*		
1001	*106*	*0.139*	*0.054*	*0.087*	*–0.037*

Horizontal deformations obtained for the period 2013–2018 along the road I-9 are in the range from 0.61 to 1.04 m and vertical ones are in the range between -0.088 up to 0.465 m. It was established that the overall movement of the terrain located in the southwestern part of the area after a turn of the road I-9 (points 2002, 375, and 2008). The maximum displacements are at the point 375 – horizontal 1.034 m and vertical –0.465 m. In the northern part of the road I-9 (points 1001, 1003, 1004) slight movements of the terrain were obtained that are in the range from 0.139 m to 0.242 m in the horizontal plane and from 0.05 to –0.07 m in height.

For the IInd period (2018–2019), which is only 10 months long, the deformations have lower values ranging from 0.085 to –0.148 m in the horizontal plane and in vertical the maximum is 0.074 m.

The obtained first results of the landslides investigation presented in this paper can be summarised as follows:

1. Extensive research was performed on the recent activations of landslides and old and new geodetic data concerning the deformations were selected;
2. A local image archive of Sentinel-1 satellites was created for the region of Northeastern Bulgaria;

3. A set of interferometric images was created at fixed intervals - monthly, every 4 months, 8 months, a year;

4. Thematic interferometric images used in mapping deformations for the region of Northern Black Sea coast are generated,

5. The relationship between geodetic and satellite derived information concerning ongoing landslide processes is confirmed.

Conclusions

On the basis of the obtained results from processing GNSS and SAR data, it can be concluded that both used data sources lead to similar results (the registered displacements are in the range of centimeters) and they confirm the overall behaviour of the landslides under study. The differences between them could be explained by the large number of external factors affecting SAR data such as vegetation and temporal decorrelation. When comparing the two methods, it should also be taken into account that the values of the IFIs elements correspond to a much larger area (15 × 15 m), while the GNSS refers to point measurements. Nevertheless, the results of this study are encouraging and the authors will continue their research on the investigation of the landslide zones using SAR data. Another factor that supports the usage of IFIs in landslides investigation is that the price and the man effort are much lower than those necessary for GNSS measurements.

Acknowledgements

This paper has been made available with the financial support provided by the National Science Fund, call identifier "Financial support for basic research projects on societal challenges – 2018" Project number КП-06-ОПР 06/1. The authors would like to express their gratitude to GeoVARA Ltd, Geozastita Varna Ltd., Ministry of Regional Development and Public Works и Survey Group Ltd. for the provided data.

References

1. Berov, B., Pl. Ivanov, N. Dobrev, R. Nankin, and M. Krastanov. State of the Art for Landslides Along the North Bulgarian Black Sea Coast, Landslide Science and Practice, Volume 5: Complex Environment, Springer, Editors: Claudio Margottini, Paolo Canuti, Kyoji Sassa, 2013, 97–102.
2. Bruchev, I., N. Dobrev, G. Frangov, Pl. Ivanov et al. The landslides in Bulgaria – factors and distribution, Geologica Balcanica, 2007, 36, 3–4, 3–12.
3. Luis, V. 2016. TOPS Interferometry Tutorial, http://sentinel1.s3.amazonaws.com/docs/S1TBX%20TOPSAR%20Interferometry%20with%20Sentinel-1%20Tutorial.pdf (accessed January 2019)

4. Tsenkov, T. Assessment of the use of geodetic methods in the study of landslides, abrasion and erosion. Journal of the Bulgarian Geological Society, Laboratory of Geotechnics - BAS, 1993, LIV, 1, (in Bulgarian)
5. Milev, G., T. Tsenkov, M. Avdzhiev et al. Instruction for the study of landslides by geodetic methods, Sofia, GUKK, 143 p., 1987 (in Bulgarian).
6. Copernicus Open Access Hub. https://scihub.copernicus.eu/dhus/#/home (accessed January 2019)
7. Vassileva, K., G. Valev, and M.Atanasova-Zlatareva. Deformation analysis of the Balkan Peninsula from GPS data 2011–2016. Coordinates, XIII, 12, Thomson Press (India), 2017, 31–34, ISSN: 0973–2136.
8. Stringmeteo. https://www.stringmeteo.com/ (accessed January 2019)
9. Geozastita – Varna Ltd., Annual Report under Contract No РД-02-29-198 / 02.05.2018 and Additional Agreement № РД-02-29-198 (1) / 06.11.2018 Preventive activities related to registration and monitoring of the landslide areas on the territory of the Republic of Bulgaria and areas with abrasive processes along the Black Sea coast in the regions of Dobrich, Shumen, Varna, Burgas (in Bulgarian), 2018
10. Ministry of Regional Development and Public Works, Map of landslide, http://gis.mrrb.government.bg/
11. Ministry of Regional Development Contract № RD-02-29-372/11.11.2013 "Geodetic survey of landslides between bus stop Fara and bus stop Obzor, and Kranevo village". Contractor "Survey Group" Ltd.

ПРИЛОЖЕНИЕ НА ДАННИ ОТ ГНСС И РСА ЗА МОНИТОРИНГ НА СВЛАЧИЩА ПО ЧЕРНОМОРСКОТО КРАЙБРАЖИЕ НА БЪЛГАРИЯ

М. Атанасова, Х. Николов, К. Василева

Резюме

Свлачищните процеси се считат за основна част от природните опасности, възникващи в северната част на българското черноморско крайбрежие. Техният мониторинг може да се реализира с висока точност, като се използват данни от глбални навигационни спътникови системи (ГНСС). Целта на това изследване е да предостави надеждни резултати при мониторинг на свлачищните процеси, използвайки данни от ГНСС и радари със синтезирана апертура (РСА). Тя ще бъде постигната чрез прилагане на: 1) създаване на доказана методика за извличане на висококачествена информация от РСА изображения, насочена към непрекъснат мониторинг на свлачищните зони, интегриращи интерферометрични изображения (ИФИ) и данни от ГНСС, и 2) създаване на работещ прототип на информационна система за наблюдение и предотвратяване на ефектите от движението на земната кора (свлачища,

пропадания и др.) базирана на свободно достъпни данни предоставени от ЕКА и национални източници. Една от научните задачи, които трябва да бъдат решени включва разработването на методологични подходи за сравнение на резултатите от комбинирана обработка на интерферометрични изображения от РСА, измервания от националната мрежа перманентни ГНСС станции на Националния институт по геофизика, геодезия и география, намиращи се в изследвания район, както и геодезически измервания на новосъздадена тесто-ва геодинамична мрежа, обхващаща конкретна зона по Северното Черноморие на България, в която са проявени активни свлачищни процеси.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e08

INTERFEROMETRIC PROCESSING OF TANDEM-X BISTATIC SAR DATA USING GAMMA – IMPLEMENTATION OUTLINE

Zlatomir Dimitrov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: zlatomir.dimitrov@space.bas.bg

Keywords: *GAMMA Software, SAR, Interferometry, Interferometry Processing*

Abstract

This article represents a showcase of two coding approaches with GAMMA, used to calculate topographic and differential phases from high resolution TanDEM-X bistatic data, provided by DLR. The first implementation approach comprises "BASH" scripting in Linux environment, having direct control of the GAMMA executables. The second approach is the utilisation of the PyroSAR framework, via GAMMA-API, in Python environment. Two spatial resolution scales are used – of 4 and 12 metres, to test the feasibility of TanDEM-X InSAR output products in mountainous forests in the rugged region. The first approach allowed thorough processing with abundant GAMMA output, whereas the high scale PyroSAR framework via GAMMA-API showed a fast implementation. Comparison over 4 and 12m spatial resolution products showed good feasibility with strong influence from topography. Intense multi-looking resolved better connection of coherence amplitude to the volume decorrelation in canopy, despite preserving high resolution reveals plenty of details in land cover. The differential height calculation, without phase unwrapping, showed its significance in data processing over mountainous regions. Intensities normalisation and terrain flattening showed good performance in both resolution scales. Finally, utilisation of GAMMA in InSAR processing of high resolution TanDEM-X bistatic SAR data showed good feasibility and flexibility to derive interferometric products.

Introduction

Applying SAR Interferometry is promising in survey natural environment and has been widely assessed in scientific studies, pointing out Earth crust deformation after hazards, derivation of DEMs in various areas, studying forest canopy in respect to different environmental conditions [1, 2] Recent years, considering forestry domain, plausible results in respect to the microwave band are provided from the Japanese L-band system ALOS PALSAR for assessing forest biomass and canopy height via POLinSAR modeling [3, 4]; via InSAR coherence on large scale boreal forest [5]; providing time series acquisitions to estimate large scale mapping of growing stocks either via L-band [6] or via C-band by ESA ENVISAT [7]; and the DLR's X-band bi-satellite system TanDEM-X, which overcomes the limitation of the temporal decorrelation, where the bistatic coherence is directly linked with the forest biomass [8, 9]. Also, this allows forest modelling using POLinSAR-scattering models, via model inversion with two baselines [10]. Considering SAR remote sensing software, various graphical based products, both freeware and commercial, have been developed in recent years – such as ESA *SNAP*, ENVI *SARScape©*, ERDAS *Radar Ortho Suite©*, and PCI Geomatics *SAR polarimetry workstation©*. In those software products, various InSAR processing workflows are implemented, comprising - direct calculation of DEM, Coherence change detection, or particular DInSAR workflows, providing more or less ease of use. Aiming at the SAR data and systems, the desired results from interferometric processing via provided tool-boxes, are not always as accurate as are needed. An example from the practice is the impossibility to adjust the particular intermediate parameter during processing, to react more adequate to the physical and environmental conditions of the observed area. In the sake of that, in particular to remote sensing studies useful approach could be provided by software GAMMA©, developed by Swiss Company "*GAMMA Remote Sensing and Research AG*", and available in Friedrich-Schiller-Universität-JENA, Institut für Geographie, Lehrstuhl für Fernerkundung. Coding with GAMMA offers professional attitude to the SAR processing tasks, where executable binaries from supported toolbox could be implemented in Linux-shell environment (BASH), or embedded in Python code.

This study demonstrates the use of GAMMA software for InSAR processing according to the use of two different method approaches. This constitutes of practical utilization of GAMMA using InSAR technology on provided SAR data from TanDEM-X, implemented in shell BASH scripts and open-source Python framework *PyroSAR*. Motivation is to utilise both software implementation and to test for the first time before the Bulgarian scientific community the interferometric processing of high resolution TanDEM-X bistatic data.

The general task is to derive the topographic and differential phases from bistatic X-band SAR data from TanDEM-X over mountainous forest in Bulgaria, along the presence of topography and different types of land cover. Therefore, following tasks are formulated in this study:

- Elaboration of optimal implementation chains using GAMMA software package, for interferometric processing of TanDEM-X data,
- Clarification of the InSAR workflow via specific implementation chains,
- Utilize *PyroSAR* framework, especially the GAMMA-API, in TanDEM-X bistatic data processing
- Calculate terrain flattened intensities by elimination of topography [11];
- Calculate bistatic Coherence from TanDEM-X data;
- Calculate Topographic phase, with high spatial resolution;

- Calculate Differential phase, related to relevant phase center height from the referent DEM.

Data and study area

According to formulated aims and tasks the high resolution data from German satellite system – TanDEM-X provided by German Space Agency (DLR) was used in this study. Milestone is that provided SAR data are bistatic - simultaneous acquisition, thus resulting elimination of the temporal decorrelation from InSAR pair. In the contribution to the scope of the study, high resolution with unprecedented accuracy Digital Elevation Model (DEM) from TanDEM-X - *"TanDEM-X DEM12"* was used, provided by DLR.

SAR data

TanDEM-X bi-satellite system

The primary objective of TanDEM-X mission is to map terrain topography by generating a digital elevation model (DEM) on a global scale, with unprecedented accuracy employing bistatic acquisitions [12]. This is achieved by a couple of twin satellites – TerraSAR-X and TanDEM-X flying in a Sun-synchronous orbit in close formation, with varying baseline separation in between. A key feature of TanDEM-X is the bistatic acquisition that employs to eliminate the temporal decorrelation factor from interferometric coherence. Operational modes of acquisition that are currently operated by TanDEM-X bistatic system, are - bistatic, monostatic and alternating bistatic, which comes in combination with imaging modes of – ScanSAR, Stripmap and Spotlight with respect to resolution.

Used TanDEM-X bistatic SAR data

A single-pol TanDEM-X SAR dataset over the test site – TS21, provided by DLR in terms of *CoSSC data proposal*, was used to test outlined methodology. The operating mode is Alternating Bistatic, which constitutes two bistatic acquisitions – 'ALT1' and 'ALT2', and one pursuit monostatic acquisition – 'MONO', packed in three separate *dims*-catalogs; it is spring acquisition, with specifications as follows:

Table 1. Specifications of the TanDEM-X SAR data acquisition

AOI	TS21	**Average Coherence**	0.7506
Product	TDM1_ COS_SM	**Along-track Baseline (ǁ)** [m]	-184.9802
Orbit	DESCENDING	**Effective Baseline (⊥)** [m]	150.9871
Date & time	2018-05-02, T04:36	**Ambiguity height** [m]	48.3917
Acquisition	Alternating Bistatic	**Incidence Angle** [deg]	39.2733
Polarisation	HHHH	**Range pixel spacing** [m]	1.3641
Band, freq. [GHz]	"X", 5.589	**Azimuth pixel spacing** [m]	2.4277

Each dataset is compiled in tar archive as a *dims*-catalog, with structure as follows:

- *The first level* – compulsory hierarchical tree with common supporting data of the bistatic acquisition, with the two acquisitions from TerraSAR-X/TanDEM-X satellites, where in bistatic configuration first is transmitting (BTX1) and the second one is receiving only (BRX2).
- *Second level* – tree structure of the particular satellite's measurement, where:
    - Annotation xml-files with various measurement metadata,
    - Various auxiliary raster of previews for *GoogleEarth*,
    - Measurements as a COS-file, with respect to polarisation,
    - Various preview raster and other supporting information.

Table 2. Structure levels of the DIMS-catalog, with compiled TanDEM-X data

<table>
<tr><th colspan="3">The first level structure</th></tr>
<tr><td rowspan="6">dims_op_oc_dfd2_xxx
(TDM.SAR.COSSC)
(spc.number)</td><td rowspan="6">TDM1_SAR__COS_ALT1_SM_S_SRA_20180502T043648_20180502T043651</td><td>COMMON_ANNOTATION</td></tr>
<tr><td>COMMON_AUXRASTER</td></tr>
<tr><td>COMMON_IMAGEDATA</td></tr>
<tr><td>COMMON_PREVIEW</td></tr>
<tr><td>TDX1_SAR__SSC_BRX2_SM_S_SRA_20180502T043648_20180502T043651</td></tr>
<tr><td>TSX1_SAR__SSC_BTX1_SM_S_SRA_20180502T043648_20180502T043651</td></tr>
<tr><th colspan="3">Second level structure:</th></tr>
<tr><td rowspan="5">TDX1_SAR__SSC_BRX2_SM_S_SRA_20180502T043648_20180502T043651</td><td>ANNOTATION</td><td>*.xml</td></tr>
<tr><td>AUXRASTER</td><td>*.tif, *tif.kml</td></tr>
<tr><td>IMAGEDATA</td><td>IMAGE_HH_SRA_stripFar_013.cos</td></tr>
<tr><td>PREVIEW</td><td>*.tif</td></tr>
<tr><td>SUPPORT</td><td>*.xsd</td></tr>
</table>

Both acquisitions from TerraSAR-X/TanDEM-X during pre-processing in DLR are precisely co-registered in slant-range geometry, thus formulating the CoSSC data format. Each acquisition is structured in a binary file in COSAR file format, formulated as 32-bit floating point Big-Endian data, organised in row order starting from the first to the last one. Relative to orbit the acquired image are normally flipped on North-South or East-West direction. A raw overview of the interferometric pair is also provided from the TanDEM-X pre-processor at DLR, including – intensity, coherence and interferometric phase along the scene, as shown on Fig. 1.

(A) **Intensity** (B) **Coherence** (C) **Interferometric phase**

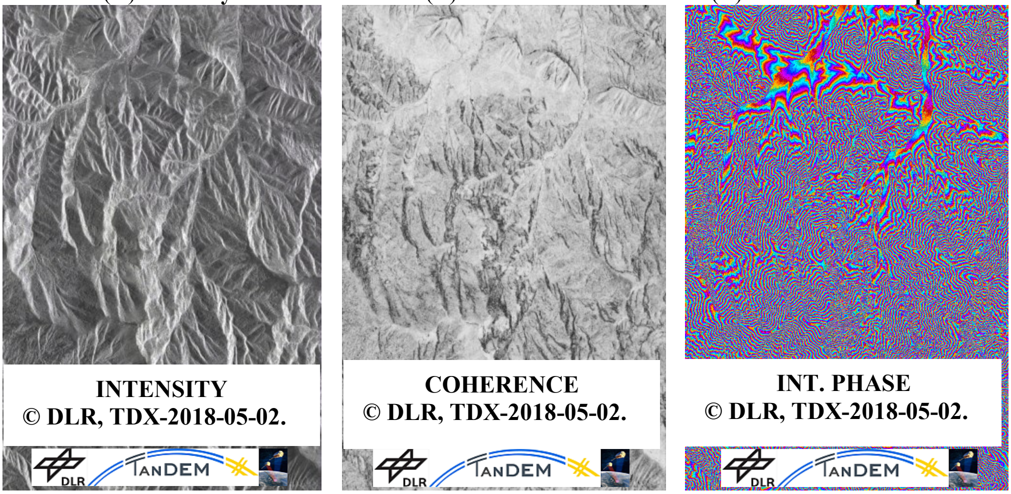


Fig. 1. Provided quick-look images of the interferometric pair of the bistatic acquisition, from TanDEM-X; ©DLR

DEM data

High resolution "TanDEM-X World DEM12" reference DEM product

The *TanDEM-X World DEM12©* was elaborated according to conducted from DLR World DEM campaign, from 2010 to 2014, using a particular set of parameters with respect to topography and land cover [12, 13]. It is acquired in the bistatic InSAR stripmap operational acquisition mode in single - HH polarisation. During whole four global acquisitions trade-off in values of ambiguity height and therefore effective baseline is made, map with sufficient accuracy for whole types of topography on earth. Most acquisitions with varying ambiguity heights and combination of ascending / descending nodes are over mountainous areas and tropical forests, held in the last global acquisitions. The final DEM product achieves absolute vertical accuracy of less than 10 [m] for 90% of the Earth's surface, and relative vertical accuracy less than 2 [m] for 10% of the Earth's surface (*HRTI-3 DEM definition*) [14].

Used "TanDEM-X DEM12" data

The innovative high accuracy *TanDEM-X DEM12* product was used as a reference DEM, in the scope of a DLR's DEM-proposal, provided from the German Space Agency (DLR) in support to the study with the delivered TanDEM-X bistatic data. An example is shown below of the DEM tile over North-West Bulgaria over mountain massif of "Stara Planina". Each tile constitutes at least –

35 TanDEM-X acquisitions, with ascending/descending orbits, with a spatial resolution of 0.4 [arcsec] – or c.a. 12 m in UTM projection, shown in Fig. 2.

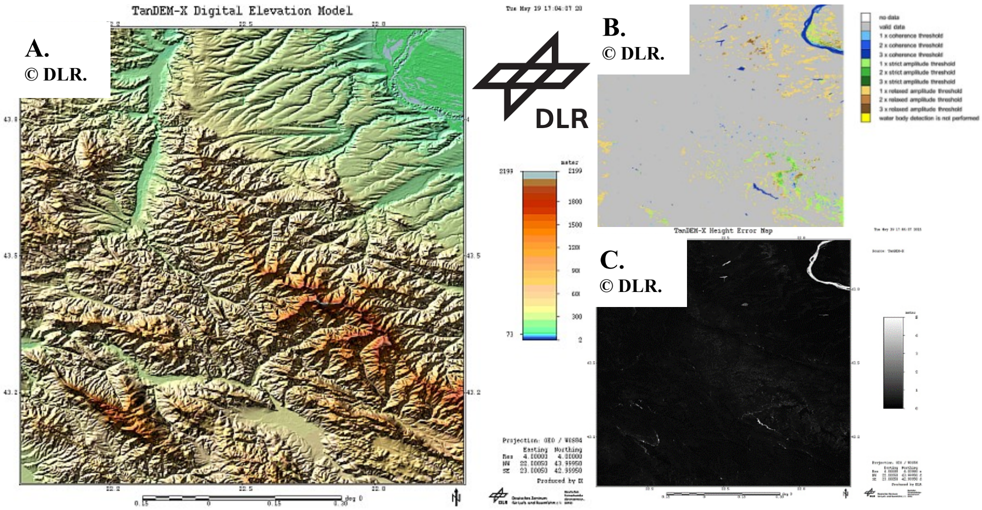


Fig. 2. Tile of the TanDEM-X DEM12 Product (A), height error map (B) and coherence thresholds, showing inconsistencies (C) – preview of provided DEM tile, located in NW Bulgaria over Test sites – 20/21/22, in SFE "Chuprene", SFE "Mijur" and SFE "Govezhda" respectively. Source – ©DLR

Study area

The selected test site is part from the PhD study of the author, with index – TS21 "Chuprene". The test site is located in the Bulgarian mountainous temperate forests in North-West part of "Stara Planina" mountain massif, bounded to the national border with Serbia on South.

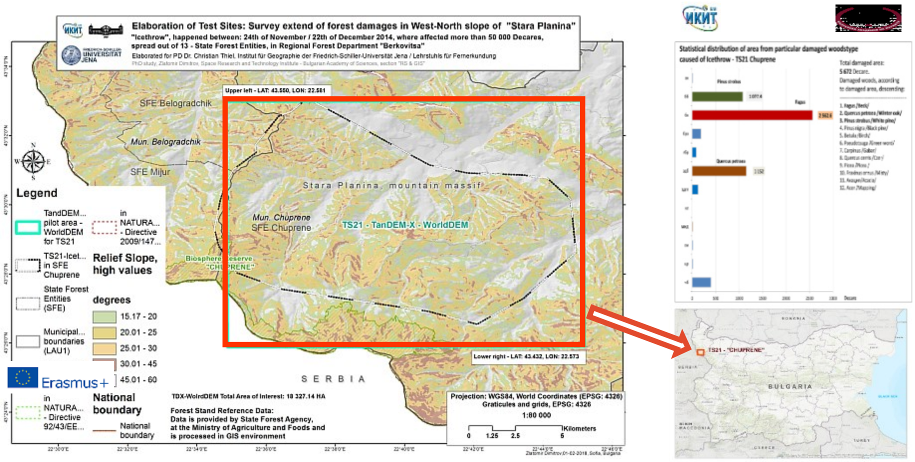


Fig. 3. The TS21 at SFE "Chuprene"; disturbed forest species with representative statistic.

The test site is specifically selected to offer various environmental conditions, which comprise hilly and rugged terrain with strong topography and forest diversity - in the aspect of density and type. Also, the test site includes disturbed forest area with different percentages of damages on forest (from 20 up to 100%), caused by severe Icethrow natural disaster happened at the end of 2014. Deciduous forests are the main forest type, where main species diversity is presented by Winter Oak, Beck, Carpinus and other, showed on figure-3 (see, statistics chart).

Meteo-situation:

The meteorological condition at the time of acquisition is showing shows clear calm weather, with low winds; previous precipitation, relative humidity ab. 80% representing the mean value on vertical stratification; mean temperatures in the early morning was ab. 15 °C. That highlights almost perfect weather conditions.

Methods

Package based software GAMMA© is collection of programs for interferometric and differential-interferometric processing, with a high scale of output products, such as DEM or displacement maps [15, 16]. In that study, following GAMMA modules are concerned – the *Interferometry, Differential Interferometry and Geocoding* (ISP / DIFF & GEO) modules, *Land Application Tools* (LAT) module and *Display* (DISP) module. Modularity architecture of the software allows building of complex processing workflows, and accurate tracking of processing steps and debugging. The GAMMA release from 2018 was available at the Uni-Jena's cluster; release supports bistatic InSAR processing needed for TanDEM-X bistatic, and programme tools for proper import of digital elevation models (DEM) in GAMMA format. In particular, for accessing GAMMA executables, it was considered to develop two implementation lines with different method approaches. Both implementation lines use the same GAMMA executable tools. Therefore, the calculation of the components from the interferometric phase via both approach results in the same output. In spite, the methodology intends to test calculation of the phase components separately, along both implementation lines.

The first approach comprises the implementation of the GAMMA executables in Bourne Shell Linux environment (BASH), as a set of shell scripts. Primal aim here is to calculate the Topographic phase component derived from interferometric measurement. Considered benefits from using this approach are related to: flexibility related to environmental variables in shell-scripts, with the direct output in terminal from processing GAMMA executables, useful in error tracking.

The second approach comprises the implementation of GAMMA functionality, using GAMMA – API in *PyroSAR*, which is a Python-based framework, developed by MSc John Truckenbrodt, at Friedrich-Schiller-Inoversität-JENA, Lehrstuhl für Fernerkundung. It is open source software framework for large-scale SAR based satellite data processing, with extensive functional capabilities [17]. The primary aim pursued in that approach is calculation of the Differential phase component, which allows the differential height between referent DEM and SAR measurement to be derived. Considered benefits from using this approach via *PyroSAR* GAMMA-API are related to: fast implementation in terms of object-oriented Python environment and fast access to image specifications in metadata. The idea of the both line approaches is illustrated in Fig. 4A, down below.

Processing with GAMMA

Whole processing flow is conducted via two implementation lines, as mentioned above, to test different approaches in accessing GAMMA package to process SAR data, illustrated in Fig. 4A. Both implementation lines include method approaches via BASH shell scripts, and via Python scripts using PyroSAR GAMMA-API. A wide range of scripts has been developed by the author in Linux/Shell environment and Python 2.7, according to the scope of the two implementation lines. Coding aims to automate the processing chain with GAMMA binaries. The processing flow begins with data import of the TanDEM-X bistatic SAR data in GAMMA format. Afterwards, both implementations include initial and advanced data processing using InSAR-advanced techniques. Output products from both lines are scaled into 4 m and 12 m pixel spacing.

1. Data import in GAMMA

On the first instance – *first level processing*, the raw TanDEM-X CoSSC bistatic data are imported into GAMMA file format; it consists of two files, where:

- The binary file, which contains the COSAR – 32-bit floating point Big-Endian complex image data (FCOMPLEX), organised in row order starting from the first to the last one;
- ASCII parametric files (PAR) providing general CoSSC product annotation, such as sensor type, date, range samples and azimuth lines, range and azimuth pixel spacing in metres, incidence angle image center (ellipsoid), and Orbit State Vectors (OSV), and other.

Data import is performed using *PyroSAR*, GAMMA-API. On the first instance function – *finder()* is used to retrieve desired TDX - datasets. Afterwards, datasets are identified and imported as PyroSAR objects, where function –

convert2gamma() is used to import the datasets in GAMMA format; scheme is drawn on Fig. 4B.

Afterwards, the reference DEM is imported in GAMMA. Non-additional correction is preliminary needed for the *TanDEM-X DEM12,* in terms of vertical datum (WGS84) or axis offsets (*which is not the case with SRTM for instance*); thus, the input GeoTIFF is directly imported in GAMMA. The delivered output is a 32-bit floating point binary file, and annotation constituted in a PAR-file giving information for the – lines/rows, datum, offsets, data format (e.g. REAL), and posting a resolution on latitude/longitude; the DEM-projection is according to the assumption in GAMMA – EQA coordinates [18].

4-A.

4-B.

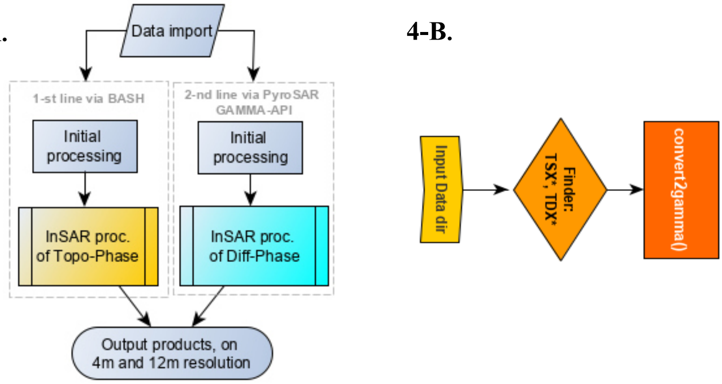


Fig. 4-A: (left) Basic scheme of the two-lines implementation, first via BASH scripting, and second via PyroSAR GAMMA-API in Python. 4-B: (right) Import of TanDEM-X datasets in GAMMA format, with PyroSAR/PyroSAR GAMMA API.

2. Initial processing

Initial processing workflow is executed prior to interferometric processing of SAR data. It aims pre-processing of imported bistatic SLC's for transmit/receive, in particular, as follows:

Calculate Offset – In the proposed processing chain offsets are calculated in purpose of the accuracy testing of the co-registration. This step generates a polynomial of offsets on range (Rg) and azimuth (Az), showed extreme accuracy with error of less than 1/10th of a pixel.

Resample Slave on Master geometry – Considering the bistatic acquisitions this resampling could be omitted. Besides, resampling is performed as assurance against possible inaccuracies (e.g. presence of rugged terrain).

Calculate multi-look factors, with respect to target resolution – Multi-look factors are calculated in a separately coded function, according to target-resolution, incidence angle, and Rg/Az pixel spacing in particular TanDEM-X imagery. Varieties of target-resolution values with ML-factors are tested to produce balanced solution of pixel spacing on Rg and Az. In target-resolution of 4 m corresponding to ML-factors of 2, resulted in ground range pixel spacing on Rg = 3.864 m, and on Az = 4.794 m, where Rg-Az – delta = –0.930 m. This is close to the critical delta difference, where higher values were found to introduce problematic geocoding. A good solution is derived aiming target-resolution of 12 m, where derived ML-factors are giving:

Rg = 11.988 m, and on Az = 11.581 m, with the delta = 0.406 m.

Multi-Looking – Higher multi-looking is better to reduce multiplicative speckle noise which bias the InSAR measurements, with respect to resolution. Multi-looking is performed according to the calculated ML-factors and the target-resolution. The produced outputs are the Multi-Looked Intensities (MLI).

Calculate Over-Sampling Factors for LUT – Over-sampling factors (OSF) are used as scaling factors for the DEM-pixels, in the sake of resampling them to a size of resolution cell in SAR image. The postings on LAT/LON provided in the DEM-PAR file are used.

Calculation of LUT, refinement (optional) – The generation of Look-up table is essential for geocoding process, and it is generated with the provided DEM and OSV of SAR (in PAR file), along the calculated OSF. The necessary refinement of the LUT as an optional step is done in the first-line implementation via BASH.

Resample DEM to Range-Doppler Coordinates (RDC) – In this step, the transformation of the corresponding AOI from the DEM, from map to radar geometry (RDC) is performed, using LUT.

Radiometric calibration – Both MLI are radiometric calibrated to sigma-nought (σ^0), which express radar-cross section for each pixel, physically measured in dB, derived from the geometry of the acquisition held in OSV in PAR file. In purpose of normalisation, the ellipsoid-based pixel reference area in ground range is outputted.

Normalization and Terrain Flattening – The final step comprises the normalisation of radar cross-section (γ^0), and compensation of the local topography based on method by David Small [11]. After the calculation of pixel-scattering area - representing the backscattered energy from the ground with respect to local topography, the actual Normalisation and Terrain-flattening are calculated, whilst perform eliminating the effect of topography. Terrain-flattening enhance the analysis based on intensities. For this purpose, the layover, shadow and incidence angle maps are used. Please, refer to Figures-5 and 6, down below.

Implementation via BASH – The first line implementation of the initial workflow is performed via BASH scripts, including whole steps above. The direct scripting approach here allowed LUT-refinement – to overcome spatial misalignment, because of problems found after geocoding. An important step in BASH scripting during Multi-Look was storing the sizes of the MLI images as BASH variables, needed further in processing. The workflow is presented on following Fig. 5:

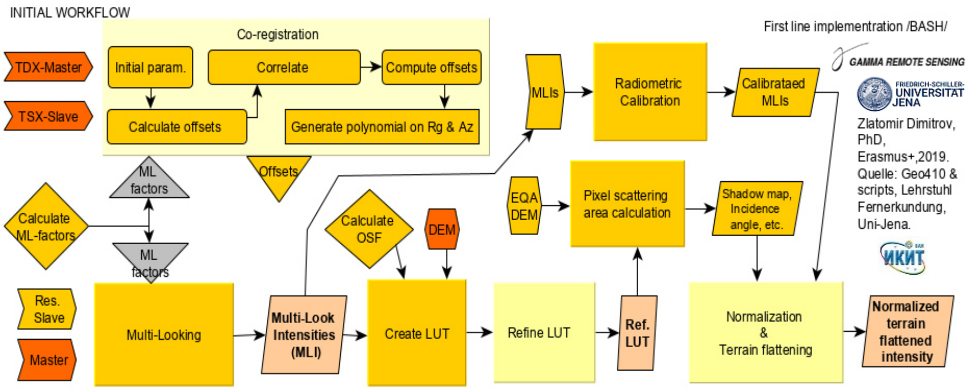


Fig. 5. Initial workflow at first line implementation constituted as Linux BASH scripts, which control directly GAMMA software binaries. Content – reddish inputs: TDX products; - yellow inputs: resampled slave over master / MLI or CMLI / DEM in SAR geometry; - yellow diamond: functions; - yellow triangles: intermediate parameters; - greyish triangles: input BASH variables; yellowish rectangles: processing steps; light yellow – complex processing steps; rose objects: outputs from processing in 4 or 12 m resolution scale.

Implementation via PyroSAR GAMMA-API:

The second line implementation of the initial workflow is performed via Python scripts accessing GAMMA-API in *PyroSAR* and additional functions available in the framework.

It differs from the first line implementation, where LUT refinement is excluded, due to unavailability within GAMMA-API. Workflow is compact and performed faster, than the first one; object-oriented Python environment allowed easy transition of the input/output function parameters and access to GAMMA-API. Easy access to SAR imagery metadata was provided by *ISPar* class in *PyroSAR*. Also, coding of workflow was facilitated by the build-in functions in Python. Nonetheless, the general negative issue was found, resulting in the misalignment of the geocoded output products, due to possible lack of LUT refinement. Please, refer to Fig. 6.

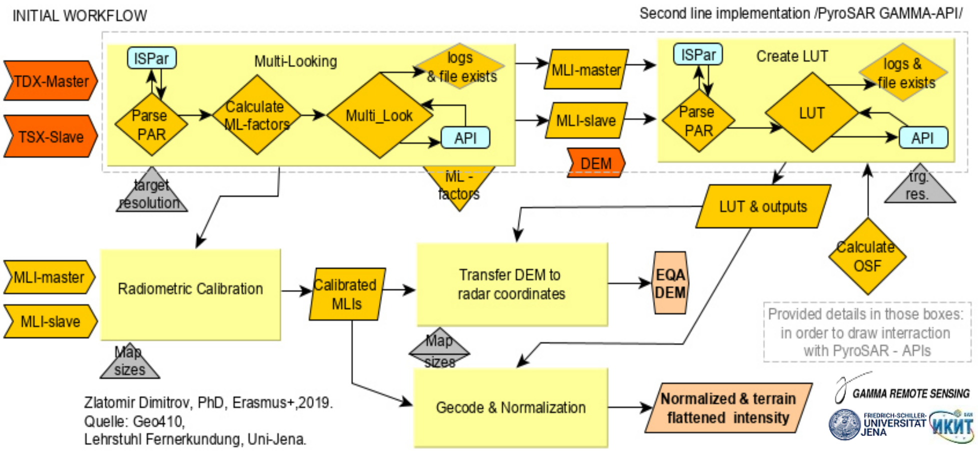


*Fig. 6. Initial processing at the second implementation line constituted as Python scripts using PyroSAR GAMMA-API, which creates access to GAMMA executable binaries. Content: - bluish rectangles: GAMMA API/PyroSAR functions; - yellowish diamonds: Module functions; - light yellowish rectangles: processing steps *in dashed contour - details provided; - yellowish trapezoid: function output as a Python object; - rose objects: outputs.*

3. InSAR Processing

Interferometric processing of the CoSSC bistatic pair derives the Interferometric phase components - topographic and differential phases, where height products along radar beam to be calculated. Approaches in GAMMA for baseline, coherence and phase-to-height calculations are similar, but rather different in the scope of interferometric calculations. On the other side, interferometric workflows in both implementation lines also differ each other. Therefore, workflows are presented separately according to both implementation lines.

InSAR processing of Topographic phase – implementation via BASH:

For the Topographic phase InSAR processing several BASH scripts were developed, constituted of:

Raw-Interferogram calculation – In this step, interfering with the bistatic pair (CoSSC) is performed using Master and resampled Slave, using OSV, offsets, and multi-look factors on Rg and Az. The raw Interferogram is strongly modulated by the range-dependent fringes - Flat-earth phase component from the interferometric phase. Results are displayed from GAMMA using MLI overlay.

Baseline estimation – For Baseline calculation two methods are used – by using OSV (less accurate), and based on the fringe-frequency method (FFT) from the raw Interferogram (more accurate). Both are compared.

Interferogram flattening – Removal of the Flat-earth phase component is performed, using OSV and baseline, where the range-dependent fringes are removed. Results analysis showed a strong modulation, apparent over flat areas.

Coherence estimation – The interferometric correlation is systematically high, due to the bistatic scenario acquisition, whereas main decorrelation in behind is presented by the volumetric decorrelation. This is so, due to the baseline decorrelation, resulting from the different location of the active phase centers, within the resolution cell. Coherence is calculated using adaptive approach and fixed window of max 9 pixels.

Adaptive filtering – Adaptive filtering of the complex Interferogram is performed, in order to filter out the presence of speckle noise. Results analysis showed well performance, with reduced phase noise also.

Phase unwrapping – Phase unwrapping is performed on three steps, solving ambiguities of discrete distribution of the phase, ranging within the interval [0, 2π]. Firstly, areas with low coherence are masked out – those are mountainous terrain AND areas with a high value of biomass. As observed over the test site, due to geometric distortions, such areas are very common. Secondly, the minimum-cost-flow approach is applied over the filtered complex Interferogram, using triangular-mesh method in GAMMA, by provided validity mask. Finally, the 2D-filtering is executed to fill the rest small gaps.

Phase-to-Height – To estimate the height from the unwrapped Interferometric phase, two essential steps are performed in advance: Improving the baseline using ground control points, and inversion of the phase-to-height. Due to high scale reference DEM precise GCPs in terms of location and height are available, allowing the extraction of those GCPs from the reference TanDEM-X DEM. Afterwards, improving the Baseline is performed firstly by providing the unwrapped interferogram, and secondly, by providing GCPs and Baseline. The conversion of the unwrapped phase-to-height is finally performed considering geometry along radar beam, which outputs the height map in metres.

Geocoding & export to GeoTIFF – Final steps from InSAR workflow are geocoding and exporting results to GeoTIFFs. Three general outputs are considered to be finally geocoded and exported – Multi-Looked-Intensities (MLI) of master TDX image, Coherence, and Height map from the topographic phase – high resolution DEM.

Important moment in geocoding concerns LUT accuracy used to transfer as accurately as possible the Range-Doppler Coordinates to Map-geometry. In proposed processing the geodetic world coordinate system (EPSG:4326) with vertical datum WGS 84 is used, stored in 24-bit GeoTIFF files. Whole datasets are displayed and overlaid in GIS software examined for consistency after InSAR processing. Please, refer to Fig. 7, as follows.

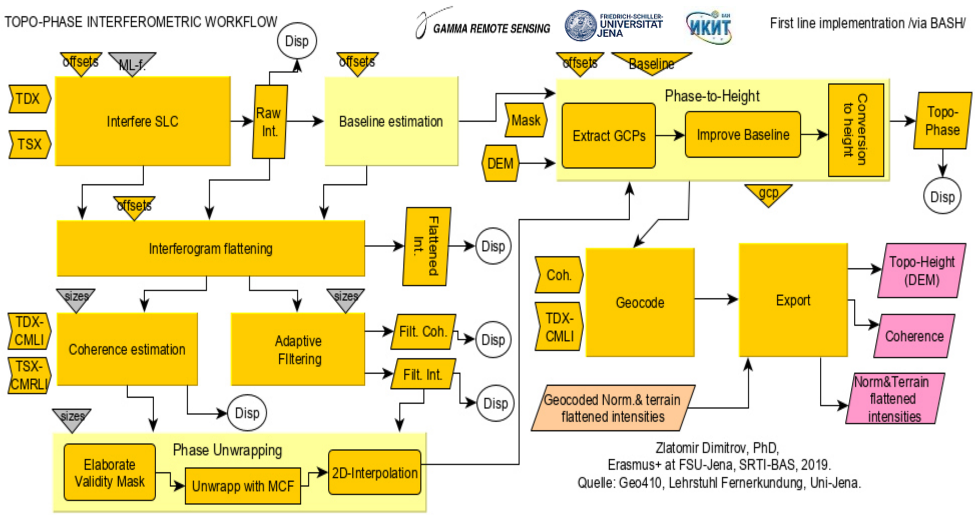


Fig. 7. InSAR-workflow at first line implementation constituted as Linux BASH scripts. Content – reddish/yellow inputs: TDX products/Calibrated MLI / DEM in SAR geometry, Low coherence mask, Coherence, Filtered coherence, Filtered Interferogram etc.; - yellow triangles: intermediate parameters; - greyish triangles: input BASH variables, MLI-width/height; yellowish rectangles: processing steps; light yellow – complex processing steps; magenta objects: final outputs from InSAR processing in 4 or 12 m resolution scale in GeoTIFF. White circles represent intermediate display via DISP-GAMMA module, from X-server in Linux environment.

InSAR Processing of Differential Phase - implementation via PyroSAR GAMMA-API

The purpose for calculation of Differential phase originally derives differential height, which could be linked directly with the canopy height, in case of the bistatic data acquisitions from TanDEM-X [19]. The main idea of this approach considers existence of LiDAR DTM, from which the ground phase to be simulated and then subtracted from the interferometric phase. Thus, calculated differential height is directly related to the forest canopy height, determined by uncertainties of the acquisition!

Despite the differential phase calculation, the second objective in this implementation line considers the powerful GAMMA-API from *PyroSAR* framework to be tested and utilised for TanDEM-X interferometric processing. Therefore, a project in a Python environment was developed, including one main script with a module, where embedding the *PyroSAR* GAMMA-API functionalities. The, SAR imagery and supporting data are Python objects, thus allowing marvellous flexibility considering processing, manipulation and accessing metadata parameters. The InSAR processing in the second line implementation includes some of the processing steps already described in the above, but includes the following steps:

Baseline estimation – estimation differs here, based only on the FFT method;

Calculate InSAR parameters – A special function in Python was developed, where to calculate the specific parameters relevant to the InSAR pair - ambiguity height (*HOA* [m]) and vertical wave number (*kz* [rad/m]);

Calculate raw interferogram – Intentionally for reference purposes and analysis;

Simulate unwrapped InSAR phase – High-scale unwrapped phase simulated from the reference TanDEM-X DEM, to be used in the calculation of the differential interferogram from TanDEM-X bistatic acquisition;

Calculate differential interferogram – The unwrapped differential interferogram is directly calculated in GAMMA functionality via GAMMA-API, using simulated unwrapped phase. Thus, differential interferogram is representable to differences between the vertical position of the phase centers and height from the reference TanDEM-X DEM. It also includes atmospheric components and phase noise – a matter of further filtering.

Coherence estimation and adaptive filtering, Phase-to-Height, Geocoding & export to GeoTIFF – Coherence estimation, adaptive filtering, conversion of phase to height along the radar beam and geocode and export is the same as stated in the first line, but here coding is facilitated by the ease of object orientated Python coding and benefits from the GAMMA-API. Results are successfully exported to Geo-Tiffs. Please, refer to Fig. 8, as follows.

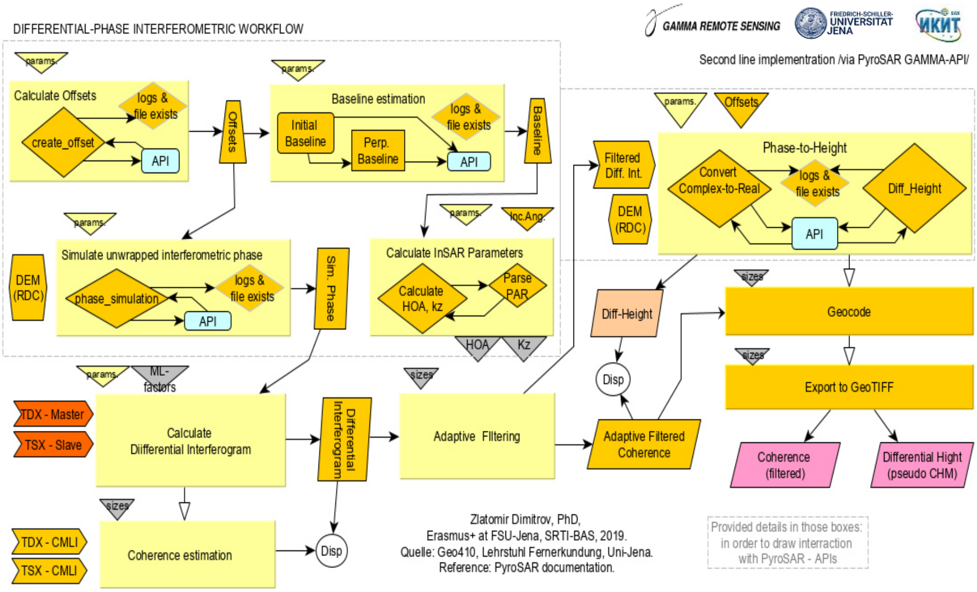


*Fig. 8. Interferometric workflow of the Differential-Phase calculation with PyroSAR GAMMA-API, where: bright yellow boxes– processing steps *in dashed contour - details provided; yellow diamonds (borders) – functions of main script/module in Python environment; blue box – reference to PyroSAR GAMMA-API / other functions; yellow diamonds (grey border) – set of Module script functions; grey rectangles - output/input variables from/to functions; red inputs – input TDX/TSX CoSSC in GAMMA; white ellipse – display functionality via GAMMA display tools within Python; - yellowish boxes: function output as a Python object; - Magenta rhomboid: outputs geocoded products in GeoTIFF.*

Results and analysis

Calculated output products have been processed in two scale of spatial resolution – 4 m and 12 m. The first spatial resolution scale is determined from the desire for high resolution InSAR products from TanDEM-X, whilst second one is especially considered to have an overlapping of the resolution cells, between MLI and the reference DEM from TanDEM-X. It is intentionally for differential phase analysis. During processing monitoring of intermediate results is performed in the first line implementation by providing graphical output in BASH via GAMMA, by the display tools, such as - *dishgt*, *dis2hgt*, *dis2rmg*, *dis2mph*, etc., for educational and developing purposes, and for adjusting the variety of additional input parameters.

Thus, the GAMMA output consists of the following dataset – *Coherence*, *Filtered Coherence* with ADF, *Differential-Height* (pseudo Canopy Height Model - CHM) calculated from Differential phase, *Differntial-phase*, *Topographic-Height*

(DEM), *Normalized & terrain-flattened Intensity* and *Averaged Intensity* between Master and Slave.

Analysis on output products with resolution scale of – 4 [m]

Coherence – Both interferometric coherence amplitudes are showing strong dependency from the topography, where areas with geometric distortions, coherence is low, nonetheless filtering is applied. Also, the relative dependency from volumetric decorrelation is also present, thus showing forest areas.

Topographic Height – Comparing the results from the calculated Topo-phase, shows conformity along the TanDEM-X DEM, where the observed differences could be mainly related to errors in phase unwrapping, and differences in heights of the phase centers. Here, must be aware, that the TanDEM-X DEM is exclusively advanced product with very high accuracy, comprising more than 38-acquisitions in a point of Earth. This is a result of covering various seasonality and meteorological conditions over the canopy, different HOA, different orbit and look angles, thus so to cope with the complicity of processes of the backscatter from natural targets. It should be noted, that the reference height from the DEM from TanDEM-X data is closer to the canopy height, than the reference DEM.

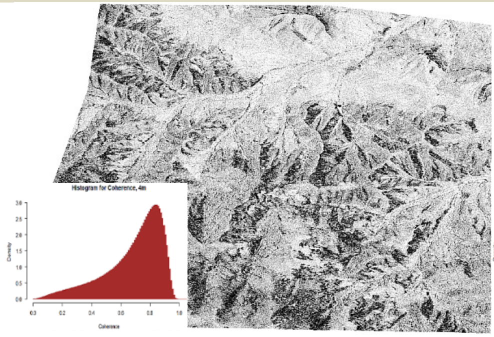
Differential Height – The output product of Differential height held within the second implementation line via PyroSAR GAMMA-API, is not representative of CHM due to no ground phase is known, but rather some mean reference canopy height is used as the reference ground phase; thus, this product is much to support the analysis on the canopy, and to tell what is the difference between referent DEM and the height of the TanDEM-X scattering phase centers from particular acquisition. Moreover, phase unwrapping is obeyed, because of the proposed assumption in the literature [20], aiming to improve accuracy over mountainous areas. Values of the Differential height are showing inconsistency with the canopy height, rather, as already stated a measure of differences. Most negative values reside in areas where the interferometric phase is noisy and coherence is low – e.g., geometric distortions. Nonetheless, resulted product shows conformity with the type of canopy, where differential height is homogenous non-forest areas, and rather heterogeneous over forest areas.

Normalised and Terrain Flattened Intensity – Terrain flattened and normalised intensity fosters amplitude analysis in the rugged terrain. Topography influence is almost eliminated, were intensity values stayed almost invariant to the slope orientation along the radar beam. As expected, a drop in backscatter is observed over forest areas. Contrary, high backscatter is observed over some bare fields and of course over urban areas (e.g. houses in Chuprene).

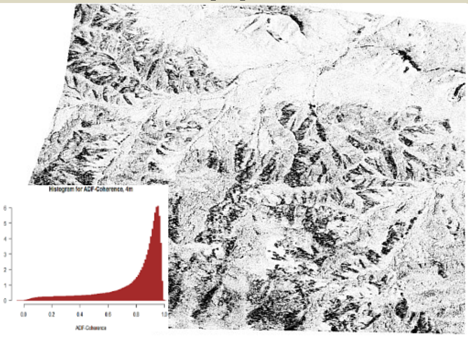
Averaged de-Speckled Intensity – The averaged intensity with applied Frost-speckle filter with a window of 7 × 7 showed good elimination of speckle, and allowed consistent amplitude analysis. Radar shadows frustrate analysis, but represents the real influence of the mountainous terrain. The GAMMA output product set is presented as follows:

Table 2. Output products from GAMMA (both implementation lines), with a spatial resolution of 4 m

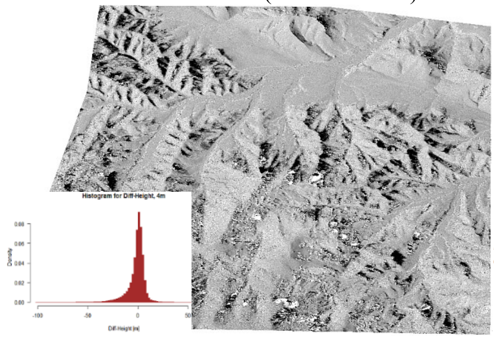
Output products from both implementation lines, **with - 4 [m] resolution cell**



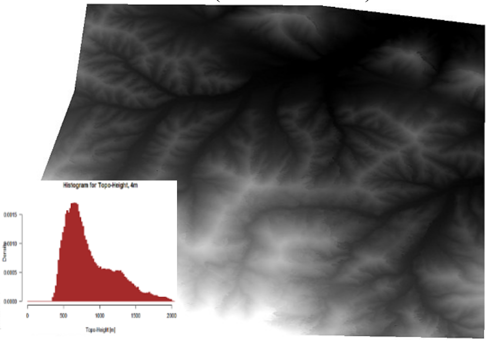
A) Coherence
values ∈ (0.177 ≑ 0.929)



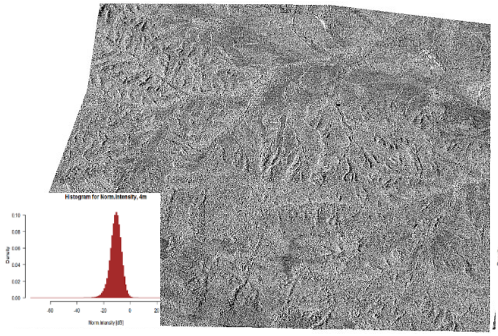
D) Filtered Coherence with ADF
values ∈ (0.134 ≑ 0.976)



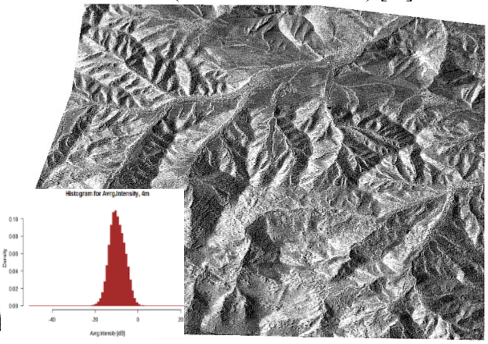
B) Differential-Height
values ∈ (-27.125 ≑ 11.379) [m]



E) Topographic-Height
values ∈ (421.958 ≑ 1726.19) [m]



C) Normalized and Terrain flattened Intensity
values ∈ (-20.816 ≑ -3.127) [dB]



F) Averaged de-Speckled Intensity
values ∈ (-16.601 ≑ -2.438) [dB]

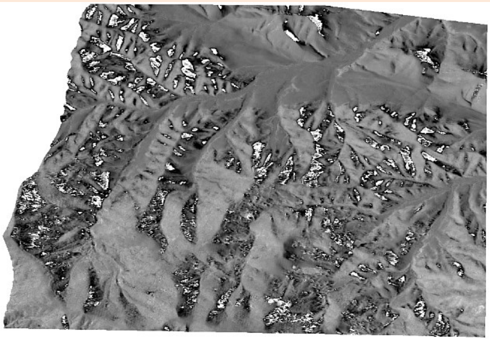
Analysis on products with resolution scale – 12 [m]

The output products are showing almost the same dependencies, as stated above, with small differences with respect to spatial resolution and phase unwrapping for differential height. Coherence amplitude shows better relation to the volumetric decorrelation over canopy. Geometric distortions are almost masked out with homogenous low values. Agricultural areas are well distinguished, also man-made objects. Analysis of the Differential height shows relative improvement with respect to the 4 m – product. Sensitivity to forest areas is better, with a smaller range of values showing better consistency with the assumption this to be measure for differences between reference height and height of the phase centers, which could be related to problems in forest areas – such as forest disturbances. In spite, at the 12 m – products discrete values showing phase jumps are observed, because phase unwrapping is omitted.

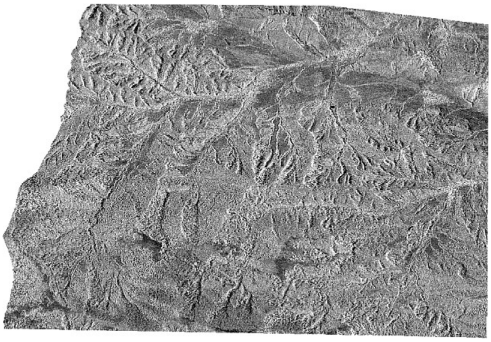
Table 3. Output products from GAMMA (second implementation line), with a spatial resolution of 12 [m]

Output products from secondonly implementation line, **with - 12 [m] resolution cell**

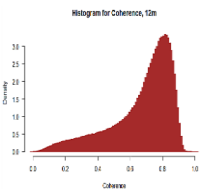
G) Coherence
values ∈ (0.158 ≑ 0.893)



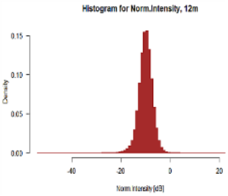
I) Differential-Height
values ∈ (-19.113 ≑ 12.672) [m]



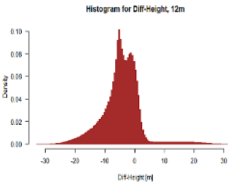
H) Normalized and Terrain flattened Intensity
values ∈ (-16.476≑ - -4.449) [dB]



Hist. of Coherence,12m



Hist. of Intensity,12m



Hist.of Diff.Height,12m

Discussion

Distinctly, most advantage of GAMMA software package is that it offers various functionalities giving possibility to produce different components of the Interferometric phase. Nonetheless, small discussion is useful to be made considering frustration and difficulties on both implementation approaches. Main difficulties at first approach (first implementation line), are related to the enormous parameters that should be passed to the GAMMA executables. This expects punctuality and precision during coding; in spite, freedom to any functional needs is available. Considering second approach (second implementation line), difficulties are much reduced, with respect to the benefits from object-oriented nature of the *PyroSAR* framework, that resides in Python environment. The only minus here is that customisation of functionality needs coding, which puts inspiration to further contribute to this useful framework, where to expand its functionality.

During the utilisation of the TanDEM-X bistatic data over mountainous forest areas, analysis in GIS environment of the output interferometric products showed particular inaccuracies in respect of terrain slope and the type of land cover. There are two main reasons, why products derived from topo- and diff-phases from this processing flow were not completely accurate. First – considering topographic phase, data acquisitions in X-band are preliminary backscattered from smaller objects comparable to the wavelength, thus resulting in backscatter from the top of the forest canopy. This obliges filtering after phase unwrapping, because of increased error probability density. Second – considering the differential phase, concerns the available reference *TanDEM-X DEM12*, which is influenced by the canopy height due to the property of short microwaves (e.g. X-band) , which saturates in the higher biomass, thus extinction is higher and therefore it does not represent the ground phase. Moreover, the bistatic acquisitions are showing big sensitivity to geometric distortions in rugged terrain, thus those areas are ambitious to be interpreted.

It is considered further development of this study in case of available LiDAR reference acquisition providing the ground phase, or the topography of the forest floor (e.g. DTM). Thus, the differential phase method approach (second implementation line) could derive the realistic canopy height based on differential phase, which is truthfully utilization of differential SAR interferometry based on X-band SAR data from bistatic acquisition.

Conclusion

This study demonstrated utilization of interferometric processing of TanDEM-X bistatic SAR data acquisitions with software GAMMA©. It is based on two different coding approaches, resulting in interferometric outputs on two different resolution scales of 4 and 12 m.

First implementation line via BASH scripts allowed better track of intermediate results and detailed scripting capabilities. Whilst, utilisation of TanDEM-X bistatic processing via *PyroSAR*, GAMMA-API showed very good feasibility and fast scripting, via the object-based environment in Python. Also, the powerful *PyroSAR* framework with the GAMMA-API allowed ease of coding into very common Python environment, allowing object-based approach in processing of SAR data. A comparison of both implementation lines (e.g. via BASH and via *PyroSAR* GAMMA-API) showed that they complement to each other, in terms of the expected result and processing chain.

Analysis of interferometric outputs on both resolution scales showed various performances. The interferometric processing of topographic phase showed high fidelity in a 4 m resolution scale, and utilised well the bistatic acquisitions from TanDEM-X in rugged terrain, with the presence of various forest canopy. Multi-looking of up to 12 m of the high-resolution TanDEM-X bistatic data revealed better sustainability of the interferometric products, in mountainous forest areas with strong topography. Considering differences in reference height of the high scale TanDEM-X DEM12 and the vertical position of the active phase centers from the TanDEM-X acquisition, omitting phase unwrapping gave better results, than high-resolution products of 4 [m] with phase-unwrapping applied.

Finally, could be concluded that utilisation showed that the GAMMA package offers extreme flexibility in solving SAR remote sensing tasks offering professional approach in applying interferometry on complex land cover, especially over mountainous forest areas with the presence of topography. It is least likely such performance to be achieved via other available graphical based software, based on author’s experience.

Acknowledgement

The TanDEM-X CoSSC data was provided from DLR in the frame of *CoSSC data proposal*, where the *TanDEM-X DEM12©* high resolution DEM product in the frame of a *DEM proposal*.

Elaborated scripts by the author follow up the scope of the course GEO410 ‘Programmierung in der Fernerkundung (GAMMA)’, held in the frame of Erasmus+ exchange at Friedrich-Schiller-Universität – JENA, Lehrstuhl für Fernerkundung.

During development, I wish sincerely to thank John Truckenbrodt, for the precious help over *PyroSAR*, installation and usage of GAMMA-API, and the help of the Python debris. I wish to thank to Dr. Christian Berger from the Lehrstuhl, for providing me with their available base scripts that facilitated the development of the BASH scripts.

Also, mine sincere thanks to the Bulgarian Academy of Sciences, and ERASMUS+ programme that gave the chance to visit FSU – JENA, Lehrstuhl für Fernerkundung, and to Dr. Christian Thiel who invited me, prof. Christiane

Schmullius the head of the Lehrstuhl; the TanDEM-X Science Team from DLR - the Microwave and Radar Institute for providing me with this extraordinary TanDEM-X data. Finally to my tutor prof. Eugenia Roumenina, and my father prof. Dobri Dimitrov, and to all scientists I have met in Jena, who showed a friendly support and where always ready to answer my questions.

References

1. Bamler and Hartl. Synthetic Aperture Radar Interferometry. Inverse Problems, 14. 1998.
2. Richards, M. A. Beginner's Guide to Interferometric SAR Concepts and Signal Processing. IEEE A&E Systems Magazine, 2007, 22, 9.
3. Hajnsek, I., et al. Tropical-Forest-Parameter Estimation by Means of Pol-InSAR: The INDREX-II Campaign. IEEE, 2009, 47, 2.
4. Dimitrov, Z. Potential of POLinSAR in survey of forest ecosystems. Conference paper, In: Proceedings of 12th Scientific Conference with International Participation-SES2016, Nov-2016, URL: http://space.bas.bg/SES/archive/SES%202016_DOKLADI/3_Remote%20Sensing/15_Dimitrov.pdf. (In Bulgarian)
5. Thiel, C. and Schmullius, C. The potential of ALOS PALSAR backscatter and InSAR coherence for forest growing stock volume estimation in Central Siberia. FSU-Jena, Elsevier, RSE-173, 2016.
6. Wilhem, S., et al. Large Area Mapping of Boreal Growing Stock Volume on an Annual and Multi-Temporal Level Using PALSAR L-Band Backscatter Mosaics. MDPI, 2014, 5.
7. Santoro, M., et al. Forest growing stock volume of the northern hemisphere: Spatially explicit estimates for 2010 derived from Envisat ASAR. Elsevier, RSE-168, 2015.
8. Martone, M., et al. Volume Decorrelation Effects in TanDEM-X Interferometric SAR Data. IEEE, 2016, 13, 12.
9. Caicoya, A. T., et al. Large-Scale Biomass Classification in Boreal Forests With TanDEM-X Data. IEEE, 2016, 54, 10.
10. Kugler, F., K. Papathanassiou, and I. Hajnsek. Pol-InSAR Techniques Forest Characterisation with TanDEM-X. HR Institute at DLR, POLinSAR-2013 ESA, Frascatti, 29 January 2013.
11. Small, D. Flattening Gamma: Radiometric Terrain Correction for SAR Imagery. IEEE, 2011, 49, 8.
12. Krieger, G., A. Moreira, H. Fiedler, I. Hajnsek et al. TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry. IEEE, 2007, 45, 11.
13. Shulze, D., M. Zink, G. Krieger, J. Böer, and A. Moreira. TanDEM-X Mission Concept and Status. DLR, 2008.
14. Krieger, G., et al. TanDEM-X - A Satellite Formation for High Resolution SAR Interferometry. DLR, Microwave and Radar Institute, 2005.
15. GAMMA Software Information, v1.4, 22-Feb-2019, GAMMA Remote Sensing Research and Consulting AG, 2019.
16. GAMMA SAR AND INTERFEROMETRY SOFTWARE, v2.1, TS, UW, CW, GAMMA Remote Sensing Research and Consulting AG, Sep. 2006.
17. Truckenbrodt, J., F. Cremer, I. Baris, J. Eberle. PyroSAR - A Framework for Large-Scale SAR Satellite Data Processing. FSU-Jena, Conference paper, DOI: 10.13140/ RG.2.2.16424.83206, May 2019.

18. Santoro, Maurizio, (Cartus, Oliver). Course GEO 410, Uni-JENA, SS 2019. "Lection-1. Geocoding with the DIFF&GEO module". Slide 10/12. Gamma Remote Sensing AG, 17.Nov.2017.
19. Santoro, M., (Cartus, Oliver). Course GEO 410, Uni-JENA, SS 2019. "Lection-5. Interferometric processing of TanDEM-X with ISP/DIFF&GEO/LAT". Slides 37-39. Gamma Remote Sensing AG, 17.Nov.2017.
20. Santoro, M., (Cartus, Oliver). Course GEO 410, Uni-JENA, SS 2019. „Lection-2. Interferometric processing of ALOS PALSAR for generating a Digital Elevation Model". Slides 37–39. Gamma Remote Sensing AG, 17.Nov.2017.

ИНТЕРФЕРОМЕТРИЧНА ОБРАБОТКА ЧРЕЗ СОФТУЕРА"GAMMA" НА РАДИОЛОКАЦИОННИ ИЗОБРАЖЕНИЯ ОТ SAR, ОТ НЕМСКАТА СПЪТНИКОВА БИСТАТИЧНА СИСТЕМА - TANDEM-X

З. Димитров

Резюме

Интерферометричният анализ (InSAR) на радиолокационни изображения от SAR с цел пресмятане на различните компоненти на интерферометрич-ната фаза е сложна задача в дистанционните изследвания. Софтуерът GAMMA предлага най-добрата гъвкавост и професионален подход. Тази статия пред-ставя пилотно приложение на софтуера GAMMA посредством две работни вериги, за пресмятане на топографската и диференциалната фази от радарни изображения с бистатична конфигурация на заснемане, в микровълнов канал – "X" от немската спътникова система – TanDEM-X, на немската космическа агенция – DLR. Използван е тестови район в северозападния дял на Стара Планина. Изследвани са два подхода на обработка – чрез „BASH"- скриптове с директен достъп до команди от GAMMA; – чрез многофункционалния пакет „*PyroSAR*" и съответния му приложен програмен интерфейс – GAMMA-API. Изходните продукти са представени с 4 m и 12 m пространствена разделителна способност (ПРС), като показват различни предимства при условия на релеф и висока надземна биомаса. Едно от заключенията на изследването е, че из-ползването на GAMMA дава най-добрата гъвкавост и точност в изчисляването на различни компоненти на интерферометричната фаза, особено в планински територии.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e09

MATHEMATIC- SYSTEMIC RESEARCHING WITH THE INVOLVEMENT OF FRACTALS WHEN PROCESSING OF SPACE SURVEILLANCE SYSTEMS ON RECOGNITION OF WASTE DISPOSAL OBJECTS

Maretta Kazaryan

North Ossetian State Medical Academy
e-mail: marettak@bk.ru

Keywords: *Waste Disposal Objects, Satellite Imagery Data, Involvement of Fractals*

Abstract

The article analyzes the factors that stimulate research and development of an automated system for processing and analyzing satellite imagery data. In particular, the study of satellite imagery for the presence of waste disposal objects is considered as an application. This problem is of particular interest when conducting scientific research. The development of automated control systems (ACS) in the field of space image (SI) recognition is characterized primarily by the approach of interest, i.e. for what purpose and for solving of which particular problem of the subject area the system is created. Obviously, the system is a set of certain information technologies and the main factor in the development of an automated system is and remains the study of peculiarities of creating an appropriate information system (IS). The article proposes a methodological approach to the processing and analysis of data by the IS. The approach is based on the conceptual idea of fractal sets. This enhances the detection of abnormal signals in difficult phono-target environments. Testing is conducted on the detection of WDS under conditions of a low signal-to-noise ratio.

Introduction

Modern realities require a new approach for the presence of certain problems when monitoring that are directly related to the directions of development of the national economy of our country. This includes hydrometeorological conditions, ecological conditions, frequent forest fires, etc. With the advent of open global networks, satellite information has become more accessible to scientists, researchers and workers in all areas of the country's economy. The works [1–3, 26–27] ware devoted to this issue. Namely, the need to create an information system for monitoring territories for a certain range of tasks, in our case, we investigate the presence of the waste disposal facilities (WDF) and industrial waste (IW).

The article discusses the problems of waste disposal facilities and the creation of monitoring systems, carrying out the system-mathematical aspect of the problem under study. Issues related to the architecture of building an information system that monitors territories for the presence of WDF, were discussed in articles [26–27]. This paper continues the study of this problem.

The problem of identifying unauthorized landfills and monitoring the correct operation of the existing landfills in accordance with the current regulatory documents is actual, and the solution to this problem would be very relevant in order to draw attention to the issues of environmental development and biodiversity conservation. In his interview "On the rubbish disaster of Russia", Russian President Vladimir Putin acknowledged that Russia does not have a complete waste recycling system, there is no separate waste collection system, this could face an environmental catastrophe for some areas. He noted that only half of all waste in Russia is disposed of, the rest is buried. "And this process is in most cases uncontrolled and criminalized," he said (https://lenta.ru/news/2016/04/14/musor/).

This problem is of interest from a scientific point of view. The development of automated control systems (ACS) in the field of space image (SI) recognition is characterized primarily by the approach of interest, i.e. for what purpose and for solving of which particular problem of the subject area the system is created. Obviously, the system is a set of certain information technologies and the main factor in the development of an automated system is and remains the study of features when creating an appropriate information system (IS).

Using the example of processing and recognition of a SI for the presence of an WDF with a small signal-to-noise ratio, we will conduct some mathematical research and in support of the obtained results we will conduct an experiment or approbation.

The research part

One of the leading directions in the analysis of aerospace monitoring data is the development of the theory of digital multispectral and hyperspectral images of the Earth and the atmospheric surface layer processing to solve the problem of automatic (without human participation) recognition of useful signals localized in time and space randomly in the background of additive and multiplicative noise.

First of all, the study of the task of deciphering of SI is the analysis of images and of applied tasks. The image is a multi-dimensional signal. The basis of information technology (IT) of image processing includes mathematical principles, the nature of which is determined by the range of tasks. The following studies are known in the field of image processing for the creation of IS and their mathematical analysis in certain areas. Thus, on the problems of integrating system-wide approaches to image analysis [13], on developing mathematical models of images that have a stochastic nature [14], on image analysis using an adaptive filtering method [15], on recovering signals based on regularization of inverse problems [16], the synthesis of regenerative filters for the case of a small number of observations [17], by statistical texture analysis of images [18], by automatic classification and recognition of images based on statistical clustering [19], by algebraic by the traditional methods of image analysis in the form of a set of primitives having a geometric shape [20], by non-parametric image analysis methods [21], by neural network image processing technologies [22], by system structural analysis based on Peano-Hilbert type quasi-continuous sweeps [23], by system structural analysis based on field analysis having a structurally redundant quasiperiodic character [24], according to system structural analysis based on fractal signatures [25].

Analyzing the problem of detecting and evaluating signals of a random nature, it is possible to identify factors that contribute to the creation of an automated geo-ecological system for detecting WDF. The brightness fields form the background in the space monitoring task. The first factor can be formulated as follows: the background is non-stationary in the one-dimensional and non-uniform in the n-dimensional case. The second factor – the brightness fields have the property of non-Gaussianity. The third factor – the background environment is a non-Markov process. Consequently, it is necessary to take into account the correlation of long-range order, limiting the use of classical correlation methods. The fourth factor - the background is a singular process, i.e. it is non-differentiable and this causes mathematical difficulties in data processing. The fifth factor is that the signal-to-noise ratio is rather small and this leads to certain problems in signal detection and evaluation. The sixth factor is the prior uncertainty of signal knowledge, i.e. form uncertainty. The seventh factor – the spatial-temporal resolution of the means of observation is rather limited.

So, the basis of the developed information system for the processing of satellite imagery is the creation of a special information technology (IT) to extract information about WDS from the space image.

The theoretical and conceptual approach to this problem can be characterized as follows. The core of the space observation systems (SOS) for the recognition of WDS is the integration of certain concepts that are directly related to the concept of multi-scale representation of signals. In system integration, all concepts have a certain dominant role. One of those concepts is the study of fractal sets. The dominant function in this concept is the simulation of phono-target environment based on large-scale self-similarity. Further research will be devoted to the study of the remaining concepts that have a role in the formation of IT, used for thematic interpretation of the SI as part of an automated information system.

The mathematical model of the information processing system for information processing in the monitoring of territories for the presence of WDS is a stochastic process with a fractal structure.

Why namely the theory of fractals attracts our attention. As is known [2], when modeling the space image of territories, it is necessary to investigate the physical processes of dispersion, radiation of electromagnetic waves from the surface, in our case, the WDS.

If to consider the statistical characteristics of the WDS, then, as a rule, they are determined by the statistical characteristics of the irregularities of the surfaces that make up the WDS. In any WDS surface irregularities are connected and formed under the influence of certain factors - thermal, chemical, etc. There are different methods and approaches for describing the scattering and radiating surfaces of the objects under study. You can consider single-scale irregularities using a single-scale correlation radius and, accordingly, multiscale irregularities using a multi-scale correlation radius [4–8]. With an increase in the order of multi-scale, the multi-scale of the correlation radius also increases and this leads to the complication of the description of the mathematical model of the investigated space image. That is why it makes sense to pass to the fractal representation of the image, specifically the space image in the monitoring of territories, the presence of the WDS. Let us take preliminary studies related to the theory of fractals.

According to the theory of fractals [4–8], any natural object has a fractal self-similar organization structure. Each image has a scan – let's denote it by $\xi(x)$.

Definition 1. Let there be a set of hierarchical levels *i.* Each level is represented by a certain scale and radius of correlation. Then the superposition of these levels is defined as follows:

$$\xi(x) = \sum_{i=1}^{\infty} \xi_i(x)$$

Definition 2. Each level *i* have a statistical ordering defined by a correlation function $\langle \xi_i(x_1)\, \xi_i(x_2) \rangle$.

Let us assume that the correlation of this level is Gaussian, i.e.

$$\langle \xi_i(x_1)\, \xi_i(x_2) \rangle = \sigma_i \exp\left[-\frac{(x_1 - x_2)^2}{\rho_i^2} \right]$$

Here σ_i is the variance, ρ_i is the correlation radius. Each component ξ_i *(x)* has a certain zone of influence, the scale of which is given by the radius of correlation - ρ_i.

Let us suppose that the quantities σ_i and ρ_i obey scaling laws - this follows from the scale invariance of multilevel relaxation processes, i.e.

$$\sigma_k = \frac{\sigma_0}{a^k} \quad \text{and} \quad \rho_k = \rho_0 \cdot b^k$$

For hierarchically organized processes the delimitation of zones of influence at different levels, is characteristic, i.e. there is a relationship: $\rho_{i-1} \ll \rho_i \ll \rho_{i+1}$

Independence of different-scale correlations follows from this relation:

$$\langle \xi_i(x_1)\, \xi_i(x_2) \rangle = 0, \quad i \neq j$$

From here, the correlation function will have the following form:

$$\langle \xi_i(x_1)\, \xi_i(x_2) \rangle = \sigma_i \exp\left[-\frac{(x_1 - x_2)^2}{\rho_i^2} \right]$$

Taking into account the scaling properties of σ_i and ρ_i the parameters and we will have:

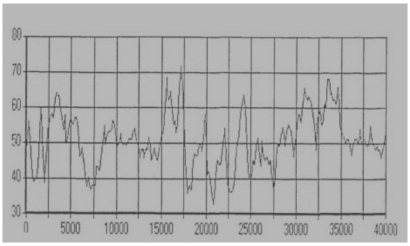
$$\langle \xi_i(x_1)\, \xi_i(x_2) \rangle \cong \sigma_0 \int_0^\infty \exp(-xp) \exp\left[-\frac{(x_1 - x_2)^2}{\rho_i^2} \exp(-xy) \right] dx$$

Let's give the asymptotic estimate of this integral with $x_1 - x_2 = \rho \to \infty$:

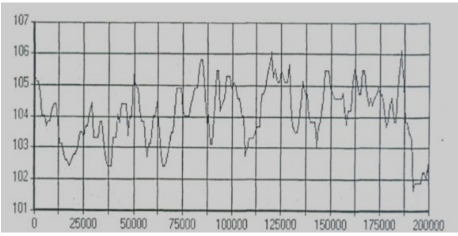
$$\langle \xi_i(x_1)\, \xi_i(x_2) \rangle_{|x_1 - x_2| \to \infty} \approx \left(\frac{\rho}{\rho_0} \right)^{-\alpha}$$

here - $\alpha = \ln a / \ln b$.

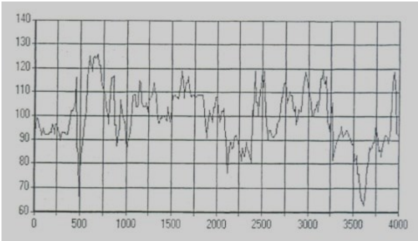
Let's consider as a structural object (set) the totality of brightness, forming a row (scan) of the satellite image when scanning the earth surface by the satellite of space observation (SSO).



a) MSU-SK data R = 160 m



b) NOAA data, AVHRR, R = 1100 m



c) SPOT data, VHR, R = 10 m
Fig. 1. Scanning of space images as structural objects with fractal properties

Fig. 1 gives a set of brightness, which together form an irregular curve associated with a random process. The scan of the image line in Figure 1 (b) is the spatial resolution of the observation system R = 1100 m., i.e. – this is a discrete sampling with an interval of 1100 m, obtained by the AVHRR scanner of the NOAA satellite system. Fig. 1 (c) – scan of the image line from the SPH satellite based on VHR, spatial resolution R = 10 m, i.e. we have a sample at intervals of 10 m.

From Fig. 1 it is clear that the sample scale does not change the appearance of the curves. An image scan is a structural object that, in a mathematical sense, can be represented as a random process with the property of scale self-similarity.

So, when developing a mathematical model of a space image, we will proceed from the fact that a space image is a stochastic process with a fractal structure [3].

Why exactly the theory of fractals is so interesting for us when building a mathematical model of a satellite image (SI) – once again we'll dwell on this issue as applied to the WDS. As is known [2–3], when modeling these image data, it is necessary to investigate the physical processes – scattering, radiation of electromagnetic waves from the surface, in our case, WDS.

If we consider the statistical characteristics of the WDS, then they, as a rule, are determined by the statistical characteristics of the irregularities of those surfaces that are part of the WDS.

In a WDS, surface irregularities are connected and formed under the influence of certain factors - thermal, chemical, etc. [2].

There are different methods and approaches for the description of the scattering and radiating surfaces of the investigated objects of a WDS. You can consider single-scale irregularities using a single-scale correlation radius and multiscale irregularities using a multi-scale correlation radius [4–7]. With an increase in the order of multi-scale, the multi-scale of the correlation radius increases and this leads to the complication of the description of the mathematical model of the given space image. Therefore, it is advisable to go to the fractal representation of the image, specifically the space image, while monitoring for the presence of a WDS.

Any natural object has a fractally self-similar organization structure [4]. As is known [3–4], a texture is a matrix or a fragment of the spatial properties of sections of images with uniform statistical characteristics. The method of describing image textures is the relationship with the spatial distributions and interdependence of the brightness values of the local area of the image or block. Textural signs refer to probabilistic signs. Random values of textural attributes are distributed to all classes of natural objects. Therefore, the concept of a texture signature is introduced. A texture signature is the distribution of the total population of measurements for a given texture in scenes of the same type as this one.

Now let's turn to the concept of fractal. The fractal dimension characterizes the degree of filling of the space in which the fractal system exists. The theory of fractals considers fractional instead of integer measures and uses new quantitative indicators, namely the fractional dimensions of D and the corresponding fractal signatures [4, 9–10, 12]. The theory of fractals and nonlinearity constitute the geometry of chaos. The contours of all natural objects are dynamic processes, frozen in physical forms and combining chaos and stability.

One of the most important issues of fractal geometry is the relationship of fractals and textures and this can be successfully applied in the tasks of detecting and recognizing low-contrast targets against the background of earthly covers.

According to Mandelbrot [4]: "A fractal is a functional mapping or a set obtained by an infinite recursive process and having the following properties 1) self-similarity, 2) their dimension (Hausdorff dimension) is fractional and strictly more than topological dimension, 3) non-differentiable and operating with fractional derivatives and integrals.

Fractal processing of low-contrast images is an unconventional way to solve a similar problem.

The establishment of the invariance of the fractal dimensions of D images from natural formations, in this case the WDS; of their brightness is the main prerequisite for the transition to fractal processing of low-contrast images.

The main idea of further research is as follows. Any deterministic object in certain territorial limits of observation has the same characteristics - size and area. When you change the scale of the terrain image, the area of fractal formations also changes. The appearance of an artificial object on any image changes the fractal dimension D of the complex image as a whole.

The scale can be varied by frame filtering the source image A into image A'.

Let $f_{i\,j}$ be elements of the image $f_H\ (x,\ y)$ of size $Nx \times Ny$., which is subjected to frame filtering, f'_{mn} – the elements of the resulting image $f'_H\ (x,\ y)$ of size $Nx \times Ny$. Then for filtering with a frame or a window *(2M +1) (2N + 1)* elements, we have:

(1) $$A \to A': f'_{mn} /\forall ij = \frac{1}{(2M+1)(2N+1)} \sum_{i=m-M}^{m+M} \sum_{j=n-N}^{n+N} f_{ij}$$

The objective presence of fractional dimensionality and scale invariance in fractals makes it possible to investigate the problem of detecting a WDS from a new point of view for small signal-to-noise ratios [5–7].
The article presents the initial experimental studies on this topic on a WDS.

The purpose of the study is the possibility of analyzing optical observation systems, as a result of which we have a SI to determine a WDS using, and constructing the field of fractal dimensions (FFD) and choosing the parameters for constructing FFD.

Experimental part

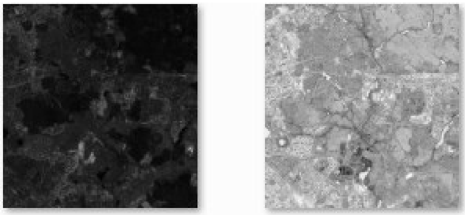
The construction of the FFD is carried out according to the following algorithm:

- image scanning by "window" with the following parameters: window size $a \times b$ pixels, movement step s (for $s = 1$ "window" sliding, for $s > 1$ window jumping);
- at each step, the numerical value of the dimension $D_{i,j}$ in the "window" is determined and written into the matrix D. This matrix will be called the "field of fractal dimensions".

The formula for calculating the number of rows and columns of a matrix depends on the number of rows and columns of the analyzed image.
M, N – the number of rows and columns of the matrix *D;*
m, n – the number of rows and columns of the analyzed image;
a (b) – the size of the "window";
s – the value of the "jump" "window";
] [– procedure for taking the whole part.

(2) $$M(N) = \left]\frac{m(n) - a(b)}{s}\right[+ 1$$

The image of the WDS at the SI is considered as the initial information allocated for the study. The WDS is displayed on the earth's surface, namely, the Kuchino SW landfill (August 2011) is given; similarly, you can further consider the object of household waste on the water surface [2, 11].



a *b*

Fig. 2: a – the image on the SI has a size 270 × 270.); b - field of fractals dimensions (270 × 270 pixels)

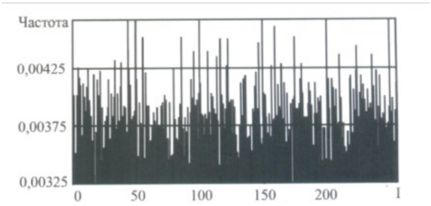


Fig. 3. Histogram of brightness

It is assumed that no pre-processing is carried out with the original image, i.e. we want to test the fractal method of processing the original signal. Let's consider the basic parameters of the field of fractal dimension. A test image is an image whose brightness has a uniform distribution law in the range from 0 to 1 and their fractal dimension $D \approx 2{,}5$.

Step 1: generation of the field of fractal dimensions. Figure 2 (b) gives the field of fractal dimensions, constructed using the "sliding window" of 17 × 17 pixels. To calculate the fractal dimension is used – the prism method [8–10]. A visual analysis of the FFD of the original image shows that the surface along the river has some features that are not visible on the SI. The type and speed of building a FFD are influenced by the size of the "window" and the displacement step. The test image is given in Figure 2 (a). The areas of location of natural objects on the fractal binary image have zero brightness, and the areas of location of WDS - maximum brightness. The total number of bright points of the object specified by the binary image is stable and obtained experimentally from the reference image of the object. To detect all areas of the image, which assumes the presence of WDS, you must scan it using a sliding window; the window size corresponds to the size of the object.

Step 2: Construction of a histogram of distribution *I* in the interval [0, 255].

Step 3: Let's construct the dependence of the fractal dimension of the image on the range of gradations of brightness.

The smaller the size of the test image, i.e. "Windows", the faster the value of the fractal dimension tends to a value of 2.5.

The error modulus δD of determining the fractal dimension in the "window" depends on the brightness range ΔI as follows:

$$\delta D = |2.5 - B(\Delta I)|$$

Fig. 4 shows that at small brightness ranges, i.e. – this is a low contrast image, it is more reasonable to use "windows" not exceeding 17 × 17 pixels.

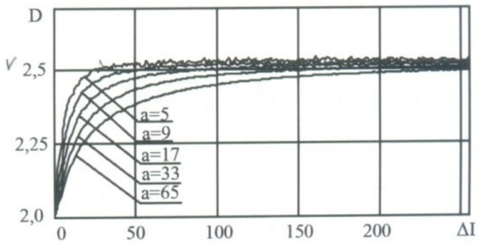


Fig. 4. The dependence of the fractal dimension with different sizes of the "window" on the range of brightness gradations.

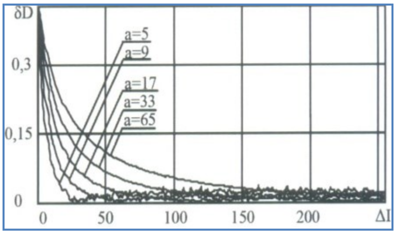


Fig. 5. Connection of the error of determining the fractal dimension of the range of brightness gradations

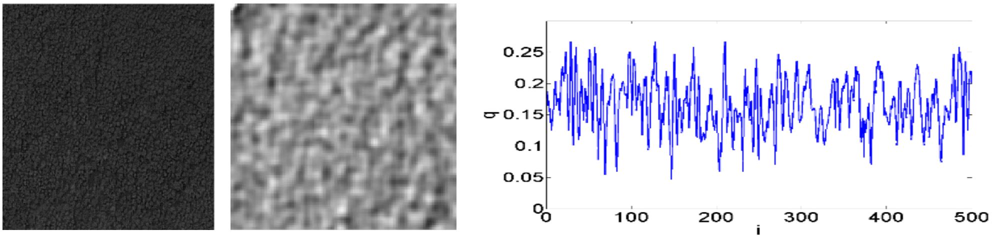
Table 1 shows the field values of the fractal dimension (FFD) when processing the original image with a "sliding" window of various sizes (a × a), s – window moving step.

Window moving step	Window size			
	5 × 5	9 × 9	17 × 17	33 × 33
s = 1	896	892	884	868
s = a	180	100	53	27

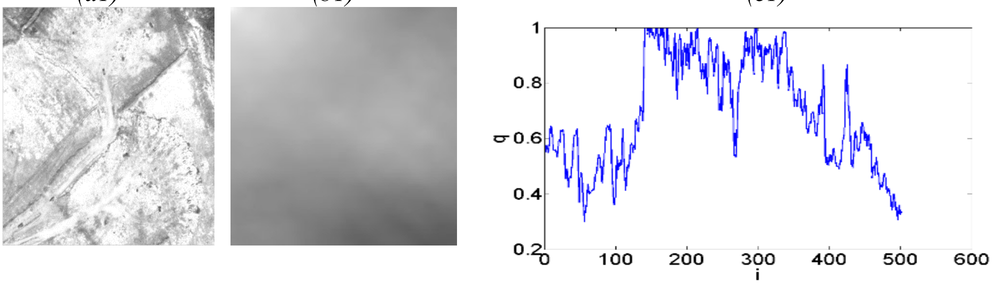
Now let's conduct a number of further studies in terms of image texture.

From fig. 6a we can see that the littering texture is visible, as well as the texture of the forest has *regrularity.* Some coatings by littering, as well as forest coverings, have the same regularity in all *directions*, others - in some directions are much larger than in others. This can be observed in the correlation images of C textures (b), which have *bands* responsible for the preferred direction of regularity. The elements of *C (x, y)* show the *correlation coefficient* of two-dimensional signals from a point (1.1) of constant length d, smaller than the size of the original image, and the shifted signals by x to the right and y down. So CML histograms (c) along vertical, horizontal and any other sections for forests are approximately the same in terms of fractal characteristics, and for littering, more regular along certain directions. In this regard, *the fractal signs* of littering are "more mobile" than forests.

WDS is an object of random but regular form and a CML; therefore it can be detected by fractal signs. *Fractals* are objects that have self-similarity and are measured by different parameters: *self-similarity density*, fractal dimension and its *field*, *accuracy of determination* (construction), etc.



(a1) (b1) (c1)



(a2) (b2) (c2)

Fig. 6. Texture patterns (blue channel) (a), patterns of correlation images (b), values of the spectral brightness coefficient (CML) in one of the sections; 1) forest, 2) littering

Color has little effect on the *value of the fractal dimension*, so its estimate can be given by one channel. Some of the “fractal” trash coverings have the same texture as the forest (see fig. 7a), and most of them refer to the typical trash texture, but with a difference in the composition of colors. So areas covered by littering at different times of the year and time of day have close fractal dimensions.

(a) (b) (c) (d) (e)

Fig. 7. Littering textures of different chromaticity, but close in fractal dimension: a) f=1.9249; b) f=1.8729; c) f=2.0952; d) f=2.0426; e) f=2.0726

The fractal dimension *f* is calculated as the slope of the regression line of the reports y*(x)*, where:

(3) $$x = \ln\frac{1}{r},\ y = \ln N,\quad N(r) = \sum_{k=1}^{l}\frac{b_k}{a_k}$$

a_k and b_k - the minimum and maximum CML of the k-th cube of the original image when it is cut into cubes of size r. Fig. 8 shows an example of fractal calculation dimensions for the standard in fig. 7a.

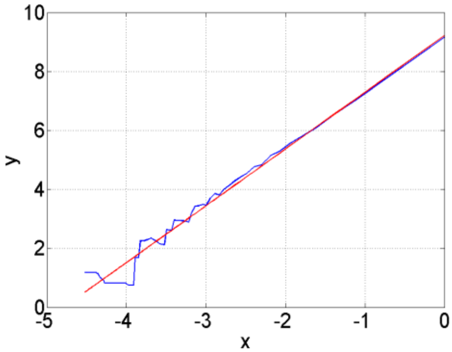


Fig. 8. Calculation of the fractal dimension

Another formula for calculating *N*:

(4) $$N(r) = \sum_{k=1}^{l}(b_k - a_k)$$

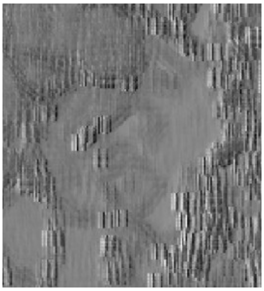
As shown by calculations using model (3), the fractal dimensions of the littering plots lie in the range of 1.85 ÷ 2.1 ($f_{min} = 1.85$, $f_{max} = 2.1$) and turn out to be close to the dimension of the ideal uniform distribution of 2.5 (see Fig. 7). For comparison, the dimension of the forest massif, which, unlike littering, has an ordered *cell structure*, according to this model, is 1.5 ÷ 1.6. Turns, spatial resolution, image size, scale of its increase or decrease do not affect the fractal dimension (Fig. 9)

(a) (b) (c)

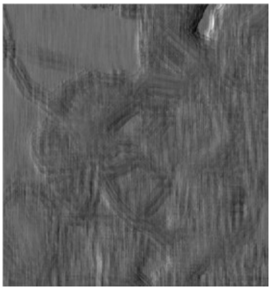
Fig. 9. Textures of littering: a) a large area (f = 1.8); b) with an increase in its area (f = 1.8405); c) at higher magnification and rotation (f = 1.8148) [Google Earth]

The trash areas are detected by the values of the fractal dimensions, according to the scheme: 1) the original image is scanned with *a square window* with a side-by-side focus on each pixel; 2) the fractal dimension is calculated by the model (3) in a square neighborhood r of each point, i.e. the field of fractal dimensions is calculated; 3) the threshold filtering method is used to allocate a detection area with fractal dimensions for littering 1.85 ÷ 2.1.

(a)



(b)



(c)

Fig. 10. Original image (a); field of fractal dimensions by model: b) (1), c) (2)

Fig. 10 shows an example of normalized fields of fractal dimensions for the image (a), calculated by models (1) and (2).

More precisely, WDS are detected by the dependences *f (q),* where *f* is the fractal dimensions of binary images *J (q)* obtained from *I* the original *I* (texture image) by threshold filtering on the lower level of the CML q. The functions *f (q)* exponentially decreases from the maximum fmax to the minimum fmin characteristic of this texture I. For the surfaces of different texture patterns, these dependences differ (Fig. 11).

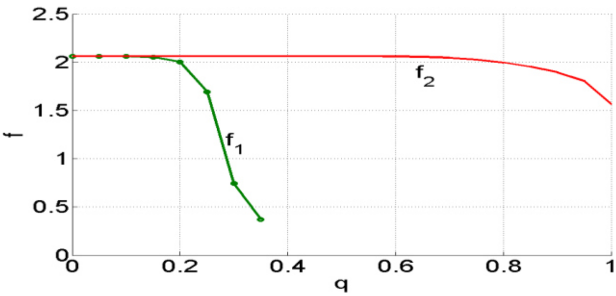


Fig. 11. Dependencies of the fractal dimensions f on the threshold CML q for covering: f_1 is a forest, f_2 is littering

The proposed method of fractals will allow overcoming the problems of processing the SPE. This is due primarily to the incompleteness of knowledge about signals, non-stationarity, non-Markovism, noise singularity based on preliminary information about the spatial scales of the detected signals.

Conclusion

One of the leading directions in the analysis of aerospace monitoring data is the development of the theory of processing digital multispectral and hyperspectral images of the Earth and the atmospheric surface layer to solve the problem of automatic (without human participation) recognition of useful signals localized in time and space randomly in the background of additive and multiplicative noise. The experiment leads to the following conclusions:

- FFD can be used to analyze SI with a WDS image.
- The size of the “window” depends on the range of brightness values of the image.
- Performing detailed segmentation of the image requires the presence of "sliding windows".
- The speed of building a FFD is increased by using “jumping windows”.

Thus theresearches of image texture and the construction of FFD were held.

References

1. Arunyants, G. G., and Kazaryan M. L. Class of systems of orthogonal Haar functions based on modified Rademacher complex functions // Vladikavkaz Mathematical Journal. 2005, 7, 1, 16–24.
2. Kazaryan, M. L., Richter A. A., Shakhramanyan M. A., and Nedkov R. Space monitoring of solid household waste and industrial waste disposal facilities (MSW and software): theoretical, methodological and socio-economic aspects: monograph / - M .: INFRA-M, 2019, 280 p. (Scientific thought).
3. Schovengerdt, R. A. Remote sensing. Methods and models of image processing Moscow: Technosphere, 2010. 560 p.
4. Mandelbrot Century. Century. The Fractal Geometry of Nature. W., H. Freeman, New York, 1983. 468 p.
5. Chen, S., Keller J., Crownover M. On the Calculation of Fractal Features from Images. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1993, 15, JSFO 10, 1087–90.
6. De Cola, L. Fractal analysis of a classified Landsat scene. Photogrammetric Engineering and Remote Sensing. 1989, 55, 5, 601–10.
7. Stein, M. C. Fractal image-models and object detection. Visual Communications and Image Processing 11, Proceedings of SPIE. 1987, 845, 293–300.
8. Peli, T. Multiscale fractal theory and object characterization. Journal of the Optical Society of America. 1990, 7, 6, 1101–12.

9. Pentland, A. R. Shading into Texture. Artificial Intelligence. 1986, 29, 147170.
10. Carr, J. R., Numerical analysis for the geological sciences: Prentice Hall, 1995.
11. Kazaryan, M., Shahramanian M., Voronin V. "The automated space monitoring system of waste disposal sites," Proc. SPIE 10793, Remote Sensing Technologies and Applications for Urban Environments III, 1079318 (9 October 2018); DOI:10.1117 / 12.2500059.
12. Sun, W. Fractal analysis of remotely sensed images: A review of methods and applications. W. Sun, G. Xu, P. Gong, S. Liang. Int. Remote Sens. 2006, 27, 22, 4963–90.
13. Pratt, E. Digital image processing, Book. 1.: Trans. from English M .: Mir, 1982. 312 p.
14. Methods of computer image processing / Ed. Soifer V.A. 2nd ed., Corr.
15. Yaroslavsky, L. P. Digital signal processing in optics and holography: Introduction to digital optics / M .: Radio and Communications, 1987, 296 p.
16. Tikhonov, A. N., Arsenin V. Ya. Methods for solving incorrect problems. 2nd ed. M .: Science: Main edition of the physical and mathematical literature, 1979.285 p.
17. Fursov, V. A. Analysis of the accuracy and construction of identification algorithms for a small number of observations. Izv. Academy of Sciences of the USSR, Techn. Cybernetics, 1991, 6.
18. Haralick, R. M. Statistical and structural approaches to texture. Proceedings of the IEEE, 1979, 67, 5, 768–804.
19. Duda, R., Hart P. Pattern Recognition and Scene Analysis: Trans. from English, M.: Mir, 1976, 512 p.
20. Zhuravlev, Yu. I. On the algebraic approach to solving problems of recognition or classification // Problems of Cybernetics, "Science", Moscow, vol. 33, 5-68 (1978)
21. Lapko, A. V. Non-parametric methods of classification and their application // Novosibirsk: IN Science. 1993. 152 p.
22. Bel'chanskiy, G., V. A. Yeremeyev, I. N. Mordvintsev, and N. G. Platonov. Comparative Analysis of the Sea of Ice Concentrations Determined Using the Okean-O1, SSM / I, and Radarsat Satellite Data. Mapping Sciences and Remote Sensing. V.H.Winston & Son, Inc. / Bellwether Publishing Ltd. Columbia MD. USA. 2002, 39, 1, 56–65.
23. Sergeev, V. V. Image processing using the Hilbert-Peano scan // Novosibirsk: Nauka. - 1984.
24. Hramov, A. G. The method of the field of directions in the analysis and interpretation of diagnostic images // Authors. bottom. Samara. 2006. 32 p.
25. Potapov, A. A, Fractals in radiophysics and radiolocation: Sampling topology // M .: University book. 2005. 848 p.
26. Lavrova, O. Yu., Loupian E. A., Mityagina M. I., Uvarov I. A., Bocharova T. Yu. Multi-User Guidelines for Satellite Remote Sensing Data // Bollettino di Geofisica teorica ed applicata. An International Journal of Earth Sciences. 2013, 54, 146–47.
27. Loupian, E. A, Mazurov A. A, Flitman E. V., Ershov D. V., Korovin G. N., Novik V. P., Abushenko N. A., Altyntsev D. A., Koshelev V. V., Tashchilin S. A., Tatarnikov A.V., Csiszar I., Loboda T., Sukhinin A. I., Ponomarev E. I., fonin S. V., Belov V.V., Matvienko G. G. Satellite Monitoring for Global Change. 2006, 11, 1, 113–45.

ПРОВЕДЕНИЕ МАТЕМАТИЧЕСКИХ – СИСТЕМНЫХ ИССЛЕДОВАНИЙ С ПРИВЛЕЧЕНИЕМ ФРАКТАЛОВ ПРИ АВТОМАТИЗАЦИИ ОБРАБОТКИ КОСМИЧЕСКИХ СИСТЕМ НАБЛЮДЕНИЯ ПО РАСПОЗНАВАНИЮ ОБЪЕКТОВ ЗАХОРОНЕНИЯ ОТХОДОВ

М. Казарян

Аннотация

В статье дан анализ факторов, стимулирующих исследования и разработку автоматизированной системы по обработке и анализу данных космических снимков. В частности, в качестве приложения рассматривается исследование космических снимков на наличие объектов захоронения отходов. Данная проблема представляет особый интерес при проведении научных изысканий. Разработка автоматизированных систем управления (АСУ) в области распознавания космических снимков (КС) характеризуется, прежде всего, интересуемым подходом, т.е. для какой цели и для решения какой именно задачи предметной области создается система. Очевидно, система это совокупность определенных информационных технологий и главным фактором при разработке автоматизированной системы есть и остается изучение особенностей при создании соответствующей информационной системы (ИТ). В статье предлагается методологический подход к обработке и анализу данных КС. Подход основывается на концептуальной идее фрактальных множеств. Это повышает обнаружение аномальных сигналов в сложной фоноцелевой обстановке. Проводится апробация по обнаружению ОЗО в условиях малого отношения сигнала к шуму.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e10

ON THE DESIGN OF THE AUTOMATED MONITORING SYSTEM OF WASTE DISPOSAL OBJECTS WITH THE APPLICATION OF EARTH REMOTE SENSING TECHNOLOGIES

Maretta Kazaryan

North Ossetian State Medical Academy
e-mail: marettak@bk.ru

Keywords: *Earth remote sensing, satellite monitoring, WDA, software, spatial data processing.*

Abstract

Currently, geographic information systems have taken strong positions in economics, politics and in almost all spheres of human activity. The result of human activity in urbanized areas is the emergence and further spread of unauthorized landfills and industrial waste. In order to timely intervene and stop the spread of foci of infection and changes in the soil composition, it is necessary to have an automated system for monitoring the disposal facilities and industrial waste.

The purpose of the work is to design a model of an automated system for monitoring waste disposal sites, including industrial sites, using remote sensing technologies.

General research methodology: The paper uses modern methods of information systems design, using the theory of databases, data warehouses, and geographic bases data.

Scientific novelty: The paper suggests a model of an automated space system monitoring for the presence of waste disposal facilities. A general methodology for constructing a geo-information model that monitors territories for the presence of waste disposal sites and also a methodology for the development of geodatabase are given. Landfill geo-databases are a structure of attribute and geographic data obtained in an automated mode. An analysis of current perspectives in the creation and operation of such automated geo-information systems is given using specific knowledge in the field of very large, distributed and open data archives. As a realization of theoretical reasoning, the practical part of the formation of software, technical and mathematical support of the remote monitoring system of the waste disposal area (WDA) and industrial waste (IW) landfill is given. In general, the work is survey-research in nature.

Introduction

Let us consider the ideology of creating a system for remote monitoring of waste disposal facilities (WDF), including industrial (IW) – (SRM_WDF-IW). It should be noted that such studies are actively conducted abroad [5–12, 30–32], as well as in the leading specialized research institutes of our country [3–5] and the same topic is often published in mass media including the Internet.

The relevance of research in the field of application of Earth remote sensing technologies is obvious. This is due to the range of problems created by the emergence of unauthorized WDAs and software, as well as the resulting consequences in all spheres of human activity.

By the order of the Government of the Russian Federation of August 21, 2006, N 1157-p the Concept of creation and development of the spatial data infrastructure of the Russian Federation is approved [1]. The concept defines that the creation and development of the spatial data infrastructure of the Russian Federation is due to the objective needs of citizens, organizations, government bodies and local governments in the effective use of reliable, operational and relevant spatial data. By spatial data here is meant some digital information (form, location, properties) about WDA and software defined in the coordinate-time frame of reference. Under the basic spatial data we will understand the spatial data that are allowed for open publication in mass media; they have a steady spatial position in the coordinate-time frame of reference. This makes it possible to orient other features to them. Metadata is information that describes the volume, content, and other characteristics of spatial data.

The concept notes: «The currently existing systems for identifying spatial objects by their address description, including registers, inventories, registers maintained by federal executive authorities, do not allow the integration and sharing of spatial data obtained from various sources. The result was the absence in the Russian Federation of a single system for identifying spatial objects; this fact prevents the use of spatial data as a universal communication element for various databases and makes it impossible to build a single information space of the country. The currently existing systems for identifying spatial objects by their address description, including registers, inventories, which are maintained by federal executive authorities, do not allow for the integration and sharing of spatial data obtained from various sources. The result was the absence in the Russian Federation of a unified system of identification of spatial objects, which prevents the use of spatial data as a universal communication element of various databases and makes it impossible to build a single information space of the country».

The creation of spatial data infrastructure will allow providing the following:

1. improvement of the quality and efficiency of management at the state and municipal levels due to the wide use of information resources of spatial data when making management decisions and monitoring their implementation;

2. provision of the up-to-date and reliable information on basic spatial data to consumers according to uniform rules and tariffs;

3. reduction of budget expenditures on the creation of spatial data in general, improving their quality by eliminating duplication of work on the creation of spatial data;

4. encouragement of investments in the creation of spatial data and related information services.

As it is noted in the article by B. A. Dvorkin [2]: “The rapid development of computer technologies, the emergence of the Internet, the emergence of desktop geo-information applications, and the active introduction of location services (so-called LBS) into the daily life based on satellite navigation systems led to the fact that spatial data is increasingly becoming a constant attribute of our daily life. If relatively recently, the main source of spatial data was topographical survey, now space imagery, aerial photography from unmanned aerial vehicles, laser scanning, etc. are coming to the fore. Now any organization, any private user, has had the opportunity to create their own maps and applications thanks to easy access to spatial data, as well as to a wide range of image processing, analysis and visualization programs. ”

From the reasoning given in article [2], we can draw the following conclusions:

1. there is an objective need for the creation and development of an information infrastructure of spatial data, both at the federal, regional and municipal levels;

2. spatial data, the infrastructure of their creation and development are one of the most important links in the digital transformation of the country's economy.

Particular attention should be paid to the problem of identifying unauthorized storage sites for solid household waste and industrial waste (SHW and IW), monitoring the correct operation of existing landfills for SHW and IW in accordance with current legislation, and assessing their impact on the environment. This problem has become so acute for Russia that Russian President Putin V.V. drew attention to it, calling it a “garbage disaster” of Russia https://lenta.ru/news/2016/04/14/musor/.

There is a great need for environmental protection systems to use spatial data based on the technologies of remote sensing of the Earth from space [14–29].

Below are examples of the application of the technology of remote sensing of the Earth from space to meet some of the needs of the environmental protection system in the area of storage of solid waste and industrial waste.

Fig. 1. Detection of unauthorized landfills

Fig. 2. Main dumpsites of the Crimea

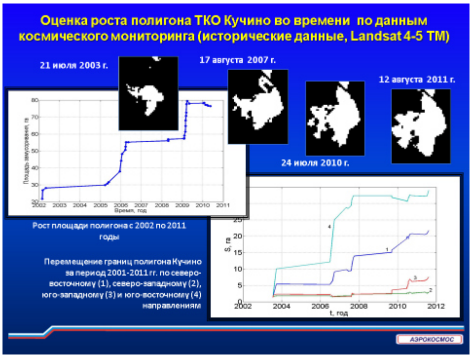


Fig. 3. Detection of unauthorized solid landfills in Sevastopol

Fig. 4. Assessment of the growth of the landfill of municipal waste in Kuchino

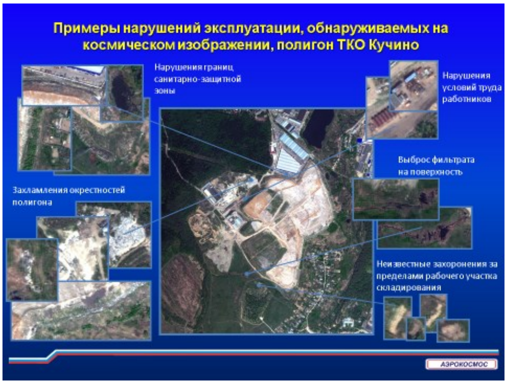


Fig. 5. Monitoring violations storage of waste

Fig. 6. Queuingcontrol from the operation of the SW landfill in Kuchino

Fig. 7. Execution control of requirements for sanitary zones

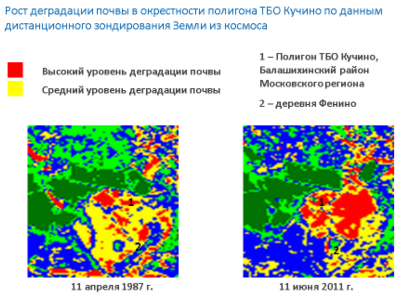


Fig. 8. Control of soil degradation in the area Kuchino landfill

The geography of our country and, in particular, of the North Caucasus is characteristic of lands that are difficult to access for field studies. To solve the problems associated with the detection of unauthorized waste disposal sites (WDS) and industrial waste (IW), it is necessary to conduct the remote sensing methods and then field studies confirming the results. The interest of the authorities responsible for the timely prevention of accumulations of solid household waste (SHW), which have inevitable consequences both in the field of sanitary – epidemiological and other areas of the socio-economic and medical-biological state of the region, is also obvious.

It should be noted that in recent years the number of spacecraft (SC) has increased in the world. This contributes to an increase in the frequency and volume of information necessary for further processing. For our task, namely, environmental monitoring for the presence of WDS and IW – means the possibility of detecting some negative consequences at the early stages of the appearance of unauthorized dumps, and this is essential.

Information received from the space station, which was previously classified as t6his sphere was actively used and is now used in the military-defense industry, became more accessible. It is quite possible to receive space images (SI) in open access on the Internet, using various technologies. The availability of information has also increased on the subject of research – WDS and IW, which contributes to the use of satellite information without creating specialized additional centers for the accumulation of primary information and this is an additional cost savings and increase in the profitability of the system. This fact may lead to the creation of its own very large archives of satellite information on the subject of WDS and IW.

Research part

Creating of a monitoring system for waste disposal facilities as well as industrial waste requires the implementation of certain requirements - these are the basic elements of a monitoring system (MS). Regardless of the type of MS, these elements are represented in the performance of certain actions with space information. The typical tasks that are part of the system for processing, storing and using space information are the following [29]:

1. receiving and collecting of space information in the automated mode;
2. storing information in archives, respectively, supporting and organizing them;
3. information processing in an automated mode;
4. transfer of information to users, including remotely;
5. use of the processed space information of a specific MS, in our case, MS_WDS and IW;
6. management and control of the relevant units MS_WDS and IW.

In carrying out all of the above tasks, software systems with a high level of automation, the ability to use distributed network technologies, etc. are used.

Let us consider the issues related to the development of software for the organization of automated reception and collection of satellite data.

We are investigating the task of creating software for the monitoring system of WDS and IW according to remote sensing data; first of all, we study the situation with the collection and reception of space information. Based on modern realities, namely, on open (relatively) access to the storage bases of the SI through the Internet, it makes it possible to create specialized centers for receiving and processing SI. All information on the SI from everywhere, including various funds of research projects, will be sent to these centers. To solve such a problem, you must have multifunctional utilities and programs for processing various SI and obtaining information for processing and monitoring territories for the presence of WDS and IW.

Let us consider the task of creating an automated system for archiving satellite data. Any automated system of this type should have a unit for maintaining satellite data archives, the main purpose of which is to store and provide the SI to users in local and remote operation.

To automate this monitoring system block, it is necessary to use a DBMS with embedded SQL language that allows you to work in software mode, having a friendly computer program interface, or using a SQL query system to implement a range of tasks related to searching and retrieving the necessary SI series. The information storage unit should have the following properties:

a. adaptability to all types of satellite information, which can be modified with time;

b. users work in remote access mode;

c. the maximum automation of the preliminary storage stages of the SI is the annotations, archiving and the storage process itself in the database;

d. this system should be open in terms of replenishment with new SIs and, accordingly, modification of technical means (servers, etc.) allocated for archives;

e. and, as a rule, separate geographically distributed centers should exist in the country as information repositories and, accordingly, certain nodes of the open Internet system, physical information carriers (CDs, etc.).

Fig. 9 shows the architecture model of the automated SI archive.

We give some comments on each block of this model.

The first block of the archiving system includes databases on all types of information, i.e. by data directory, file storage and long-term file storage. The archives of WDS and IW space data are presented in the form of a software package, which is updated depending on the tasks of interest to users and, accordingly, researchers.

The second block – the archive administration system includes archiving utilities, work with data, as well as the selection of necessary information for processing and supporting users' orders. In the management plan of the unit, a group of operators responsible for the front of the system is attached.

The third block is the availability of archiving systems via the Internet, i.e. relevant websites. This system includes user web-interfaces, as well as service web-interfaces. The work on managing this block is performed by operators and users who are competent in terms of information culture.

The fourth block is system control. It includes checking web-interfaces, email notification of various emergency situations, checking the technical condition of computers and automatic launching of special utilities – tests of the storage system of the CS as a whole. This system is open, i.e. utilities are replenished with a variety of application programs that improve the quality of the QS and modify the methods of storing them in the information repositories.

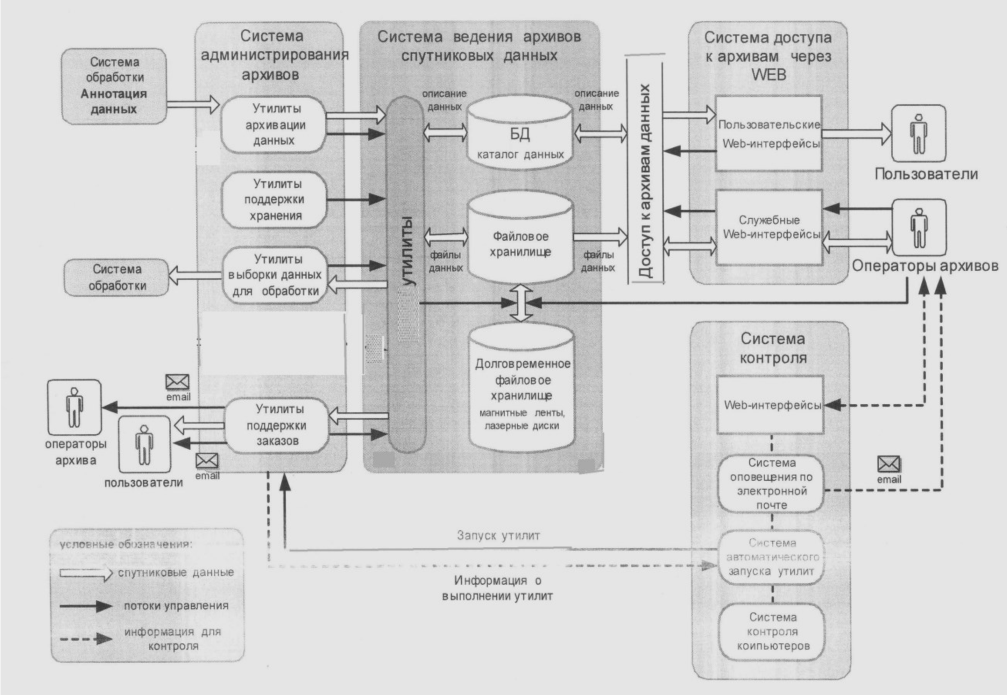


Fig. 9. Model of the system of automated archiving of the SIs when monitoring territories for WDS and IW

So, we have a description of the model of WDS and IW remote monitoring using information systems that ensure the work with super-large, distributed and replenished storages of satellite images.

Let us investigate the issue related to the space information processing system. Processing of space information is represented in the form of primary and thematic data processing. The implementation of this work is carried out in two modes - automatic and automated.

In the first case, a software package is developed; a system of macros is used, for writing of which an interface is provided. It provides a mode of remote access to information.

In the second case (automated), a sequential chain of information processing is created, forming data streams and parallelizing the processing of space data. At the same time, tasks for workstations are formed; their work is monitored, which leads to an increase in production capacity. At any time you can increase the number of workstations, etc.

Now let us investigate the issue related to the integration of the results of the processing of space data with the diverse information used in the WDS and IW monitoring system.

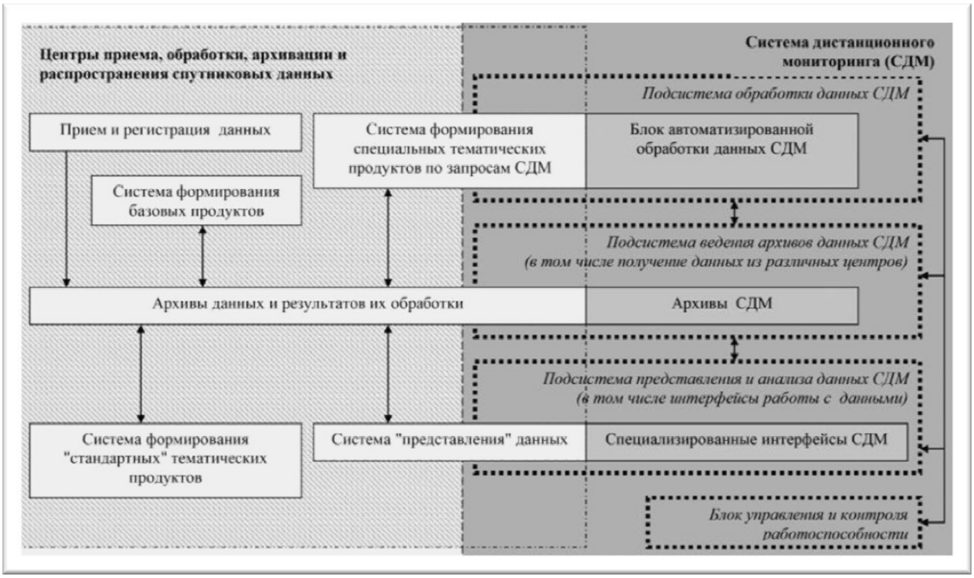


Fig. 10. The basic scheme of RMS WDS-IW formation

Provides the GIS. This option of work demands additional expenses for resources. This method is used when the monitoring system is not created from scratch. An additional unit with satellite information is inserted into the system. Therefore, a mechanism is needed that allows satellite data to be exported to GIS and inserted into its archives.

Let us consider the second approach. In the GIS, inquiries are made to obtain the necessary information from the corresponding archives of space data. In GIS, interfaces are implemented that allow combining heterogeneous information.

This approach, which is widely used nowadays, simplifies issues related to maintaining distributed SIs archives and updating information received from a spacecraft.

Let us consider the basic principles of modern remote monitoring systems (RMS), in particular WDS-IW.

Fig. 10 proposes a basic scheme for building RMS WDS-IW.

From this figure it follows that the RMS includes the following main blocks: data processing; archiving; data presentation and analysis; management and performance monitoring.

Let us consider the operation of the main blocks of the system, presented in Fig. 10. Modern realities in the field of Earth remote sensing are such that the preliminary processing of space information goes into a new category – the formation of basic products. Basic information products have a number of qualitative properties that allow for thematic processing of a SI without serious consequences – a stable radiometric calibration, illumination, state of the atmosphere, geographical and temporal reference, etc. The modern user of the SI is interested in obtaining information in the form of basic products. In exceptional cases, to obtain thematic products used in our case when monitoring territories for the presence of WDS-IW, data processing systems are created.

Recently, a definite situation has developed regarding the implementation of the RMS and the work of its individual units. The equipment necessary for the operation of these units is provided to space information receiving and processing centers. In such situations, the speed of information processing increases, since there is no need to transfer data in large volumes from the centers to the RMS.

There is a tendency to maximally automate the processing of information during the implementation of RMS, and accordingly, based on this, the work of the thematic processing unit for SI in various information systems is organized in a certain way.

We have to deal with situations related to the fact that the implementation of various processes for the processing of space information uses a variety of software environments, and perhaps also different operating systems. It is necessary to organize the interaction of these procedures and control over their implementation.

The introduction of distributed computing resources requires different approaches for their correct application in the processing of remote sensing data.

The data archiving subsystem is one of the key elements of any RMS. Above, we have described this block in detail and its principle of operation (see Fig. 9).

The information presentation and analysis subsystem is undoubtedly one of the most important elements of any RMS.

Let us consider the main factors shaping this subsystem.

The first factor affecting the formation of a system is the provision of a distributed user of space information and tools for its analysis in order to ensure remote monitoring. To perform this function, it is necessary to have a web-interface - this is a modern trend in the development of information systems. The advantages of these interfaces are that there is no need to purchase and maintain a significant number of licenses, as is done for desktop applications, the same GIS.

The second factor is the involvement of various Internet technologies for the development of complex tools in the implementation of distributed data analysis in RMS.

The third factor is the possibility of online access to data from external information systems, as well as to the resources of suppliers of remote sensing data.

And finally, speaking of the RMS control and operability control unit, it is important to note that an increase in the level of automation of health monitoring processes is required, as well as the creation of technologies for the automated detection and diagnosis of faulty situations.

Practical part

As a realization of the above theoretical considerations, we present the following practical part on the formation of software, technical support of the remote monitoring system of the WDS and IW.

The complex for receiving satellite information assumes the presence of the following equipment:

- UniScan – 36 (ScanEx manufacturer);
- Data processing server (2 pcs.) by HP;
- Cluster data storage systems (2 pcs.);
- Work stations (4 pcs.);
- Local area network (10 G; speed 10 Gbit / C);

The software of the system also includes:

- data reception (software developed by ScanEx);
- preprocessing software for different satellite systems (TERRA / AQUA for MODIS; SuomiNPP / NOAA-20 for NIIRS, CrIS, ATMS; Fengyun – 3A / B / C for MERSI). Linux / CentOS operating system on one server and every 15 minutes run snapshot processing scripts;
- for thematic processing using two servers. Each TERRA / AQuA received by the SI is generated into a set of multiscale multichannel images and this information is immediately displayed on the web page. Next, the application programs are performed to decrypt the image for the presence of WDS and IW (NDVI calculation, cloud masking, etc.).

The presence of high-performance computing systems (clusters) is assumed:

ArmCluster computing system (http://www.cluster.am): computing field of 128 cores (64 × 2 Intel Xeon 3.06GHz, 2 GB RAM), computing network – Myrinet, control network – Gigabit Ethernet, computing system ArmGrid (http://www.grid.am): computational field of 368 cores (30 × 2 Intel Quad Core Xeon E5420 2.5GHz, RAM 8 Gb), computer network - Gigabit Ethernet, control network - Gigabit Ethernet + 32 × 2 Intel Quad Core Xeon E5405 2.0 GHz, RAM 8 Gb), computer network – Infiniband, control network – Gigabit Ethernet, PhiCompute computing system: computing field 48 CPU cores (2 × 2 Intel Xeon E5-2680 v3, 128 GB RAM) + 244 Phi cores (2 × 2 Intel Xeon Phi 7120), 10 GbE computing network managing I have a network-Gigabit Ethernet.

Servers:

Application servers: 2 × HP DL380p Gen8, 2 x Intel Xeon E5-2620v2, 128 GB RAM, 8 × 600 GB 6G SAS 10k HDD.

GPGPU server: HP DL380p Gen8, 2 × Intel Xeon E5-2620v2, 128 GB RAM, 3 × 600 GB 6G SAS 10k HDD, NVIDIA Tesla k40,

management server: HP DL380p Gen8, Intel Xeon E5-2620v2, 32GB RAM, 3 × 600GB 6G SAS 10k HDD.

Data Warehouses:

QNAP TS-809U-RP network storage, storage capacity up to 48 TB

IBM x3650 network media storage, storage capacity up to 12 TB.

HP MSL2024 tape library, storage capacity up to 360Tb.

Workstations:

Dell T5500 Workstation with NVIDIA Tesla C1060 GPGPU

Dell T3500 Workstation graphics station with two ATI FirePro V5900 graphics cards.

When working on the creation of RMS WDS-IW, free or free software will be used, as well as proprietary software systems.

Consider some of the development in relation to software that is developed in the basic utilities of the system.

1. The adaptation of existing image processing algorithms to the problem of research [14–29]. These algorithms include:

- affine transformations over the image (rotation, shift, scaling);
- image enhancement;
- selection of objects;
- filtering according to the spectrum and size of objects, dilatation, erosion, opening, closing, morphological processing of images (selection of boundaries, filling of areas, selection of connected components);
- clustering and classification by methods of K-intragroup averages and ISODATA, etc.

These algorithms are adapted to the characteristics of the WDS: size, distribution, component composition, spectral composition, characteristics of the environment.

2. Development of special image processing algorithms for the detection and analysis of WDS [14–29]. These algorithms include:

- algorithms for splitting a time series of images into a time series of sections of these images;
- algorithms for the detection of texture components in the composition of the WDS and the environment;
- algorithms for assessing the morphological composition of the components of WDS;
- special training algorithms to detect WDS;
- algorithms for obtaining component landfills as part of the WDS and its surrounding natural environment;
- algorithms for obtaining relief WDS on a pair of images;

- algorithms for assessing the dynamics of changes in the surface and body of the WDS and its main geometric characteristics;
- algorithms for assessing chemical processes occurring in landfills;
- algorithms for the classification of WDS according to the data of space images;
- algorithms for calculating the linearity of polygonal objects;
- algorithms for evaluating component, temperature, vegetative, technological, and other characteristics of the WDS.

3. Study of the state of the soil and vegetation [14–29]. Algorithms for assessing the degree of soil degradation. The basis of these algorithms is the calculation. To study the state of the soil and vegetation, special indices of vegetation response have been developed depending on those or other influencing factors. Factors, vegetation reaction indices and the degree of soil degradation are presented in normalized form. At the same time, the factors themselves are ranked according to the degree of importance. This method of assessing the degree of soil degradation is the cheapest because does not require additional archival data, like most other methods for assessing soil degradation. In addition, it is fast-acting, although less accurate, and allows us to estimate the tendency of the soil not only to degradation, but also to recovery. The state of the soil is also determined by assessing the yield of the soil, since in most cases, landfills are located in close proximity to rural areas. To assess the yield of the soil, many algorithms based on multiple regression have been developed and implemented.

4. Evaluation of the geometric characteristics of the WDS [14–29].

The algorithms for estimating the geometric characteristics of the WDS are divided into algorithms for planar and spatial characteristics. A single image is used to estimate the plane characteristics, and a stereo pair is used to estimate the spatial characteristics. To assess changes in geometric characteristics, a time series of images of one territory is used. Planar geometrical characteristics include area, perimeter, polygonal region, contour and contour traversal, linearity, ellipticity, center of mass, scattering coefficient, planar parameters of components. The spatial geometrical characteristics include the volume, spatial parameters of the components, and the body of the WDS, the relief of the surface, the slope coefficient, the average height of the WDS, the number of tiers. The change in geometric parameters is divided into average and direction. The direction is given by a vector on a plane for plane parameters and a vector in space for spatial parameters. The change in parameters is characterized by speed, i.e. increment parameter per unit of time. By changing the parameters, you can make a short-term forecast of future values of geometric parameters.

5. Development of attribute databases [14–29]. Attribute databases are developed in database management systems. Structurally, they consist of a set of tables connected to each other through special link tables. Different tables characterize objects of different types. Attribute databases have been developed: on

general information of the WDS, on individuals and legal entities associated with them, on administrative-territorial units, on the component composition of the WDS, on the spectral characteristics of the WDS, on space images and metadata snapshots, etc. Each database has its own data scheme, and all developed databases are combined into a data warehouse. To manage each database, special programs have been developed in which the procedures of creation, opening, closing, updating, updating, etc. are implemented. To integrate databases into a data warehouse and to manage this storage, programs are also implemented in special software environments. Each the database and the data warehouse itself have their own interface, which allows managing databases interactively. Descriptions of databases and data storage are developed. The source data for reading the database are normalized. Some of them were obtained using programs from textual information taken from various sources. The other part is calculated according to the algorithms for detection and analysis of WDS.

Conclusion

The paper proposes the design and further use of an information system for the implementation of remote monitoring using remote sensing methods and technologies. Here we consider the features and trends in the development of modern geographic information systems in certain thematic areas, in our case, we consider WDS-IW. The article is of an overview and methodological nature. An analysis of current perspectives in the creation and operation of such automated geo-information systems is given using specific knowledge in the field of very large, distributed and open data archives. As a realization of theoretical reasoning, the practical part of the formation of software, technical and mathematical support of the remote monitoring system of the WDS-IW is given.

References

1. Order of the Government of the Russian Federation of 21.08.2006, N 1157-p.
2. Dvorkin, B. A. Spatial data infrastructure: a regional aspect. Journal of Geomatics, 2014, 1, 17–22.
3. Tolpin, V. A., Balashov I. V., Efremov V. Yu., Lupyan E. A., Proshin A. A., Uvarov I. A., and Flitman E. V. Creating interfaces for working with data of modern remote monitoring systems (GEOSMIS system). Modern problems of remote sensing of the Earth from space. 2011, 8, 3, 93–108.
4. Tolpin, V. A., Lupyan E. A., Bartalev S. A., Plotnikov D. E., and Matveev A. M. Opportunities for analyzing the state of agricultural vegetation using the VEGA satellite service. Optics of the Atmosphere and the Ocean. 2014, 27, 7, 306, 581–86.
5. Uvarov, I. A., Matveev, A. M., Burtsev, M. A., Lupyan, EA, Mazurov, A. A, Proshin, A. A., Savorsky, V. P., and Sudneva, O. A. Organization of distributed work with

satellite hyperspectral data for solving scientific and applied problems. Modern problems of remote sensing of the Earth from space. 2014, 11, 1, 322–33.
6. Acker, J. G., and Leptoukh G. Online analysis of the use of NASA earth science data. Eos, Transactions American Geophysical Union, 2007, 88, 2, 14–17.
7. Asrar, G., and Greenstone, R., 1995. MTPE EOS Reference Handbook. NASA Goddard Space Flight Center.
8. Budget Activity: National Environmental Satellite, Data, and Information Service. 2012. http://www.corporateservices.noaa.gov/nbo/fy13_presidents_budget/7_NESDIS.pdf
9. Interoperable Catalog System Valids, CEOS / WGISS / ICS / Valids, Issue: 1.2, April 2005, 55 p.
10. Lavrova, O. Yu., Loupian E. A., Mityagina M. I., Uvarov I. A., and Bocharova T. Yu. Multi-User Guidelines for Satellite Remote Sensing Data. Bollettino di Geofisica teorica ed applicata. An International Journal of Earth Sciences. 2013, 54, 146–47.
11. Loupian, E. A, Mazurov A. A, Flitman E. V., Ershov D. V., Korovin G. N., Novik V. P., Abushenko N. A., Altyntsev D. A., Koshelev V. V., Tashchilin S. A., Tatarnikov A.V., Csiszar I., Loboda T., Sukhinin A. I., Ponomarev E. I., Fonin S. V., Belov V. V., and Matvienko G. G. Satellite Monitoring for Global Change. 2006, 11, 1, 113–45.
12. Moore, R. T., Hansen M. C. Google Earth Engine: a new cloud-computing platform for global-scale earth observation observation and analysis. AGU Fall Meeting Abstracts, 2011, 1, 2 p.
13. Ramapriyan, H. K. Hazard Science Data. CENDI / NFAIS Workshop on Repositories in Science & Technology: Preserving Access to the Record of Science November 2011, 30.
14. Shakhramanyan, M. A., Kazaryan M. L, etc. Space monitoring of the state of the environment for the development of a system of additional environmental education". Modern problems of science and education. 2015.– №3.
15. Shakhramanyan, M. A., Richter, A. A., and Kazaryan, M. L., Automation of the Detection and Disposal of Waste Disposal Objects, Fundamental Research, 2015, 8 (Part 2), 281–86.
16. Shakhramanyan, M. A., Richter, A. A., and Kazaryan, M. L Cosmic image processing using multimedia and its application in space monitoring of waste disposal sites. Basic Research, No. 8 (Part 2) 2015, 328–32.
17. Richter, A. A., Shakhramanyan M. A., and Kazaryan M. L. The method of visual detection in the task of space monitoring of waste disposal facilities. Bulletin of the Vladikavkaz Scientific Center. 2015, 15, 3, 61–67.
18. Shakhramanyan, M. A., Richter, A. A., and Kazaryan, M. L., Estimation of Geometric Parameters of Littering Areas from Multispectral Space Images, Fundamental Research, 2015, 2, (part 13), 2866–70.
19. Shakhramanyan, M. A., Richter A. A., Kazaryan M. L. Development of a method for assessing the degree of soil degradation based on long-term observation data. Basic Research, 2015, 2 (part 14), 3095–99.
20. Shakhramanyan, M. A., Kazaryan M. L., Richter A.A. Identification of foci of littering according to low-resolution satellite images Landsat. Information and space, 2016, 3, 91–96.

21. Shakhramanyan, M. A., A. A. Richter, and M. L. Kazaryan. The method of automating the assessment of underlying surface indices and their changes over time using space images and its application in assessing the state of the environment in the vicinity of solid waste landfills. Bulletin of Tomsk Polytechnic University. Engineering georesources. 2016, 327, 8, 52–58.
22. Shakhramanyan, M. A., Richter A. A., and Kazaryan, M. L. Methodology for automatic detection of components of waste disposal facilities from space images. News of the Tomsk Polytechnic University, Engin. of georesources, 2017, 328, 3, 46–5.
23. Kazaryan, M. L., M. A. Schahramanian, and A. A. Richter, Future Space monitoring of the Earth and the Haar wavelet transform.
24. Schahramanian, M. A., A. A. Richter, and M. L. Kazaryan Evaluation of chemical process parameters in waste disposal sites by satellite images. Ecological Engineering and Environment Protection, 2017, 1, 22–28.
25. Schahramanian, M. A., A. A. Richter, and M. L. Kazaryan Pattern recognition algorithm using descriptors combined radio and visible spectra. Proc. SPIE Vol. 10221, Mobile Multimedia/Image Processing, Security, and Applications, 1022107 (2017) DOI:10.1117/12.226287.
26. Schahramanian, M. A., A. A. Richter, and M. L. Kazaryan. Information modeling of waste disposal sites. Ecol.l Engineering and Envir. Protection, 2017, 1, 15–21.
27. Schahramanian, M. A., A. A. Richter, M. L. Kazaryan Space Monitoring of the Earth on the Presence of Solid Domestic Wastes Using a Discrete Orthogonal Transforms. Serbian journal of electrical engineering. 2017, 14, 3, 343–64.
28. Kazaryan, M. L., Shakhramanyan M. A, and Voronin V. The automated space-monitoring system of waste disposal sites. Proc. SPIE Remote Sensing, Proceedings Volume 10793, Remote Sensing Technologies and Applications in Urban Environments III, 1079318 (9 October 2018); DOI:10.1117/12.2500059 Event: SPIE Remote Sensing, 2018, Berlin, Germany.
29. Kazaryan, M. L. Space monitoring of solid domestic waste and industrial waste disposal facilities (MSW and SW): theoretical, methodological and socio-economic aspects: monograph / M. L. Kazaryan, A. A. Richter, M. A. Shahramanyan, R. D. Nedkov. INFRA-M, 2019. 278 p. + Add. Materials. URL: http://www.znanium.com. DOI:10.12737/monography_5c4efa771779a4.
30. Astsatryan, H., W. Narsisian, and Sh. Asmaryan, SWAT Hydrological Model as a DaaS Cloud Service, Springer Earth Science Informatics, 2016, 1–7, DOI:10.1007/s12145-016-0254-6.
31. Astsatryan, H., W. Narsisian, A. Hayrapetyan, A. Saribekyan, S. Asmaryan, V. Muradyan, Y. Guigoz, G. Giuliani, and N. Ray, An interoperable web portal for parallel geoprocessing of satellite image vegetation indices, Springer Earth Science Informatics, 2015, 8, 2, 453–60, DOI:10.1007/s12145-014-0165-3.
32. Astsatryan, H., W. Narsisian, and G. da Costa, SaaS for Energy Efficient Utilization of HPC Resources of Linear Algebra Calculations, Scalable Computing: Practice and Experience, 2017, 18, 2, 145–50, DOI:10.12694/scpe.v18i2.1284.

О ПРОЕКТИРОВАНИИ АВТОМАТИЗИРОВАННОЙ СИСТЕМЫ МОНИТОРИНГА ОБЪЕКТОВ ЗАХОРОНЕНИЯ ОТХОДОВ И В ТОМ ЧИСЛЕ ПРОМЫШЛЕННЫХ С ПРИМЕНЕНИЕМ ТЕХНОЛОГИЙ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ

М. Казарян

Аннотация

В настоящее время геоинформационные системы заняли прочные позиции в области экономики, политики и, практически, всех сфер жизнедеятельности человека. Как следствие человеческой деятельности на урбанизированных территориях – это появление и дальнейшее распространение не санкционированных мусорных свалок и промышленных отходов. Чтобы вовремя вмешаться и остановить распространение очагов инфекции и изменений в почвенном составе территорий и т.д. необходимо иметь автоматизированную систему объектов захоронения и промышленных отходов.

Цель работы – проектирование модели автоматизированной системы мониторинга объектов захоронения отходов и в том числе промышленных с применением технологий дистанционного зондирования земли.

Общая методика исследований. В работе используются современные методы проектирования информационных систем с применением теории баз данных, хранилищ данных, геобаз данных.

Научная новизна. Предлагается модель автоматизированной системы космического мониторинга на наличие объектов захоронения отходов. Дается общая методология построения геоинформационной модели, осуществляю-щей мониторинг территорий на наличие объектов захоронения отходов. Общая методика разработки геобаз данных. Геобазы данных ОЗО представляют собой структуру атрибутивных и географических данных, получаемых в автоматизированном режиме. Дается анализ современных перспектив в создании и работе подобных автоматизированных геоинформационных систем с применением определенных знаний в области сверхбольших, распределенных и открытых архивов данных. В качестве реализации теоретических рассуждений приводится практическая часть по формированию программного, технического и математического обеспечения системы дистанционного мониторинга ОЗО и ПО. Работа в целом носит обзорно-исследовательский характер.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e11

APPLYING THE CONSTANT STRENGTH DOUBLET METHOD TO RESOLVE A STEADY FLOW AROUND AN AIRFOIL

Konstantin Metodiev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: komet@space.bas.bg

Key words: *Panel Method, Airfoil, Lower Upper Decomposition, Tiny C, Mugin*

Abstract

In the paper hereby, a numerical (panel) method is applied to analyze steady two-dimensional flow of ideal gas around an airfoil. Initially, the airfoil is divided into a finite number of panels. Then the panels are replaced by doublets with constant strength. In addition, a wake panel is added to fulfill Kutta condition at the airfoil trailing edge.

In order to implement this, a numerical realization is developed and built by means of Tiny C Compiler. To work out a solution to the linear non-homogeneous algebraic system, direct schemes for lower-upper factorization / decomposition of matrix of coefficients were applied, namely Crout, Doolittle, and Cholesky.

The obtained results are validated against exact solution and shown for various values of angle of attack and Reynolds number.

Introduction

For couple of years, the personnel of Department of Aerospace Control Systems have been utilizing the Chinese Mugin-3 unmanned aerial vehicle, [1], Fig. 1. Unfortunately, the airplane is deficient in some important flying and maneuvering characteristics which are supposed to have been computed by the manufacturer. By reason of making up for the missing data, the proposed study aims at retrieving aerodynamic characteristics of Mugin-3M UAV wing airfoil, such as the static pressure coefficient. The initial data are airfoil geometry, Reynolds number, and angle of attack. The airfoil geometry was politely submitted to the author by a Mugin Ltd. correspondence clerk.

The proposed study follows a computational algorithm thoroughly described by Katz and Plotkin in their famous textbook [2] and implemented further by a computational code developed in ForTran. In order that infringement of proprietary rights can be avoided, alternative source code has been developed in

C sticking rigidly to the same algorithm of computing the induced velocities and proposing an alternative for working out a solution to the linear algebraic system.

Fig. 1. Mugin-3 3220 mm UAV V-tail platform frame kit, [1]

Method

A few major stages during solution workflow are outlined below.

Discretization of geometry

At first, the airfoil is divided into finite number of panels in accordance with what is depicted in Fig. 2. A circle is drawn with center placed at the chord middle and radii of half a chord. Copies of the radii are evenly distributed in a circular pattern around the center point so as to obtain the so-called polar array, [3]. Each copy of the radii intersects the circle at a point which is further connected to the mirror one across the abscissa by a vertical line (dash). Wherever a vertical line intersects the airfoil contour, the panel end point is defined. This method makes the panel end points close up (concentrate) in the vicinity of both leading and trailing edges where the flow parameters change intensively.

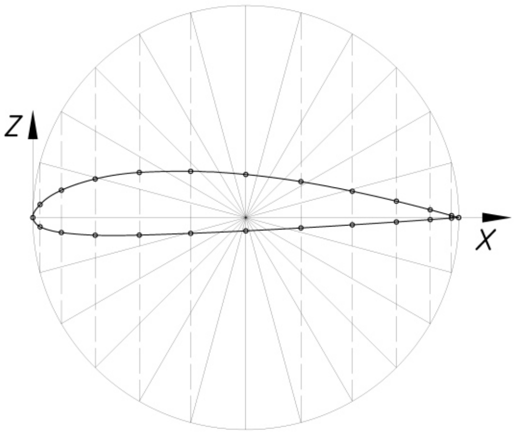


Fig. 2. Dividing an airfoil into panels

Computation of influence coefficients

A boundary condition imposed on the problem under consideration implies that normal flow component is not allowed through the airfoil contour. The normal velocity component at each collocation point C_i, Fig. 3, could be divided into a self-induced and a free-stream part, [2]. The former quantity is computed by

$$(1)\qquad \begin{aligned} u_p &= \frac{1}{2\pi}\left[\frac{z}{(x-x_1)^2+z^2}-\frac{z}{(x-x_2)^2+z^2}\right] \\ v_p &= \frac{-1}{2\pi}\left[\frac{x-x_1}{(x-x_1)^2+z^2}-\frac{x-x_2}{(x-x_2)^2+z^2}\right] \end{aligned}$$

where u_p and v_p denote induced velocities computed in a local (panel) coordinate system, x_1, x_2 denote panel end abscissas ($z_1 = z_2 = 0$), x, y denote collocation point coordinates. In Fig. 3, the local coordinate system (ξ, ζ), aligned along the panel j, is visible, so is the collocation point C_i (at panel middle). A nested loop is required to compute velocities induced at all collocation points by all panels, thus deriving a system of linear equations. The quantities u_p and v_p must be remapped back to the global coordinate system (X, Z) prior to computing the influence coefficient

$$(2)\qquad a_{ij} = (u,v)_{ij}\,\mathbf{n}_i$$

which is essentially an element of left-hand side of the system equations. The latter quantity (free stream) is a dot product of free stream velocity and normal vector in global coordinate system

$$(3)\qquad RHS_i = -(U_\infty, V_\infty)\,\mathbf{n}_i$$

which in turn belongs to the system right-hand side.

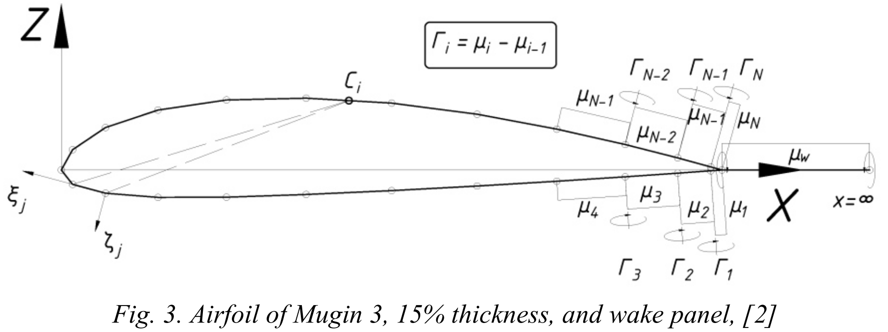


Fig. 3. Airfoil of Mugin 3, 15% thickness, and wake panel, [2]

After computing the aforementioned quantities, following linear non-homogenous system of equations is obtained:

$$
\left\| \begin{matrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N-1} & a_{1,N} & a_{1,W} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,N-1} & a_{2,N} & a_{2,W} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{N-1,1} & a_{N-1,2} & \cdots & a_{N-1,N-1} & a_{N-1,N} & a_{N-1,W} \\ a_{N,1} & a_{N,2} & \cdots & a_{N,N-1} & a_{N,N} & a_{N,W} \\ 1 & 0 & \cdots & & -1 & 1 \end{matrix} \right\| \left\| \begin{matrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{N-1} \\ \mu_N \\ \mu_W \end{matrix} \right\| = \left\| \begin{matrix} RHS_1 \\ RHS_2 \\ \vdots \\ RHS_{N-1} \\ RHS_N \\ 0 \end{matrix} \right\| \tag{4}
$$

System (4) is augmented further by Kutta condition implying that circulation at the trailing edge should be zero. However, according to Fig. 3, the vortex strength at the trailing edge is found to be $-\Gamma = \mu_l - \mu_N$, [2], hence a wake panel is required to meet the condition requirements, i.e.

$$
(\mu_1 - \mu_N) + \mu_W = 0 \tag{5}
$$

Equation (5) is added to the last row of system (4). In addition, a column is added to the matrix of coefficients denoting velocities generated by the wake panel (index W) at the collocation point (index i). The system is said to be well defined and stable in terms of numerical solution, [2].

Solving a linear algebraic system

Having computed the influence coefficients, a non-homogenous linear algebraic system (4) is obtained in terms of doublet strength μ distribution along the airfoil contour. In the current study, a direct method of compact lower – upper (LU) factorization / decomposition of matrix A was employed to work out solution to the system (1). In general, there are three types of factorization thoroughly described in textbook [4]:

- Doolittle if U matrix has 1s along its diagonal. In this case, the matrices elements might be solved for by means of following formulae:

$$
\begin{aligned} u_{kj} &= a_{kj} - \sum_{p=1}^{k-1} l_{kp} u_{pj}, \quad j = k, k+1, \ldots, n \\ l_{ik} &= \frac{a_{ik} - \sum_{p=1}^{k-1} l_{ip} u_{pk}}{u_{kk}}, \quad j = k+1, \ldots, n \end{aligned} \tag{6}
$$

- Crout if L matrix has 1s along its diagonal. For the matrices elements at step k it follows:

(7)
$$l_{ik} = a_{ik} - \sum_{p=1}^{k-1} l_{ip} u_{pk}, \quad j = k, k+1, ..., n$$
$$u_{kj} = \frac{a_{kj} - \sum_{p=1}^{k-1} l_{kp} u_{pj}}{l_{kk}}, \quad j = k+1, ..., n$$

- Cholesky if U = L^T or L = U^T

(8)
$$u_{kj} = \frac{a_{kj} - \sum_{p=1}^{k-1} l_{kp} u_{pj}}{l_{jj}}, \quad j = 1, 2, ..., k-1$$
$$u_{kk} = \sqrt{a_{kk} - \sum_{p=1}^{k-1} l_{kp}^2}$$

The Cholesky decomposition is solely applicable to a Hermitian (self-adjoint), positive-definite matrix.

The basic concept of the LU decomposition is constructing lower and upper triangular matrices for the following equation to be true:

(9) $$A = LU \therefore LU\boldsymbol{\mu} = b$$

The aforementioned decomposition might be used to solve for the unknown vector U**μ** = **y** first, [5], i.e.

(10) $$A\boldsymbol{\mu} = L(U\boldsymbol{\mu}) = L\mathbf{y} = b$$

Then, the obtained intermediate vector **y** is to be used to solve for the vector **μ**

(11) $$U\boldsymbol{\mu} = \mathbf{y}$$

The LU factorization is only possible if inverted matrix exists, i.e. det(A) = 0

(12) $$A^{-1} = \frac{1}{\det(A)} adj(A)$$

Secondary quantities

Having computed the doublet distribution along the airfoil, it is possible to work out values of some secondary quantities, [2], as follows:

- Perturbation tangential velocity in terms of induced velocities

$$(13) \qquad q_{t,i} = \sum_{j=1}^{N+1} \left(u, w\right)_{i,j} \mathbf{t}_i$$

- Pressure coefficient

$$(14) \qquad C_{p,i} = 1 - \frac{\left(Q_{t\infty} + q_t\right)_i^2}{Q_\infty^2}$$

In formula (14), $Q_{t\infty}$ stands for the dot product of free stream velocity and panel i unit tangent vector.

Source code description and results validation

The source code was built by Tiny C Compiler, [6]. In Fig. 4, the header file included in the solution might be seen. Apart from the function prototypes, an alias of a structure is defined. The structure contains main geometric quantities of a panel. These include pointers to dynamically allocated arrays containing panel end points, normal and tangent vectors, and collocation point. The header file also contains a function-like macro IX(i, j) taking row and column indexes (i, j) as arguments. Throughout the source, one-dimensional arrays are solely used even in case of matrix of coefficients (left-hand side) of system (4).

Briefly, the software work sequence is following. An array of type myPanel is dynamically allocated containing as many panels as necessary (function void *malloc). All fields within a single panel are initialized by means of a text file containing airfoil panels coordinates. For the airfoil to be divided according to algorithm depicted in Fig 2, using a CAD software is recommended, for instance AutoCAD. Then the algorithm proceeds to computing each panel geometrical quantities, function geom(). The next step consists of two nested loops intended to compute velocities at current collocation point induced by current panel, function influenceDueToDoublet(). As it was mentioned earlier, the collocation point is placed at the panel center. Having computed the influence coefficients, a method of working out a solution to the non-homogenous linear system is invoked, function Doolittle(), Crout(). Eventually, results are exported to an stdout device and all allocated structures and arrays are freed, function free().

Broadly speaking, the header file shown in Fig. 4 is self-explanatory.

```c
#ifndef DEFS_H
#define DEFS_H

// number of panels
#define I 80
#define PI 4. * atan(1.)
// Matrix dimension (includes wake)
#define COLS (I+1)
#define IX(i, j) (i) * COLS + j

typedef float real;
// All in global coordinates
typedef struct panel {
        real *x, *y;     // end panel points, [2] each
        real *n, *t, S;  // normal[2], tangent[2], area
        real *c, *p      // collocation point[2], panel centroid[2]
} myPanel;

myPanel* createPanels(int N);
int deletePanels(myPanel *foo, int N);
real* make1Darray(int N);
int delete1Darray(real *foo);
int geom(myPanel *foo, char *type);
void influenceDueToDoublet(myPanel *foo, real x, real y, int i, int j,
real *u, real *v);
real* Doolittle(int N, real *a, real *b);
real* Crout(int N, real *a, real *b);
real* Cholesky(int N, real *a, real *b);

#endif // DEFS_H
```

Fig. 4. A header file included in the developed software

In order to estimate the program ability to work out a solution in advance, a validation case was carried out about an ideal flow around a cylinder, 2D. The cylinder had been previously divided into 80 panels. It is widely known that the distribution of coefficient of pressure along the cylinder surface might be estimated according to following formula, [7]

(15) $$C_p = 1 - 4\sin^2\theta$$

where θ is a polar angle of which the coefficient C_p in (15) is solely dependent. Formula (15) is derived from Bernoulli's equation for total energy conservation of ideal gas flow, i.e. in case of inviscid, incompressible, irrotational flow.

In Fig. 5, the solution of system (4), i.e. doublet distribution, is shown, so is the static pressure coefficient distribution, Fig. 6, computed by means of formula (14). For this particular test case, $\alpha = 0$ deg, $Q_\infty = 1$, cylinder diameter (chord) = 1. Evidently, both numerical (14) and exact (15) solutions coincide, Fig. 6. The numerical solution does not include a zone of flow separation for an obvious reason, i.e. a boundary layer does not emerge in case of ideal gas flow.

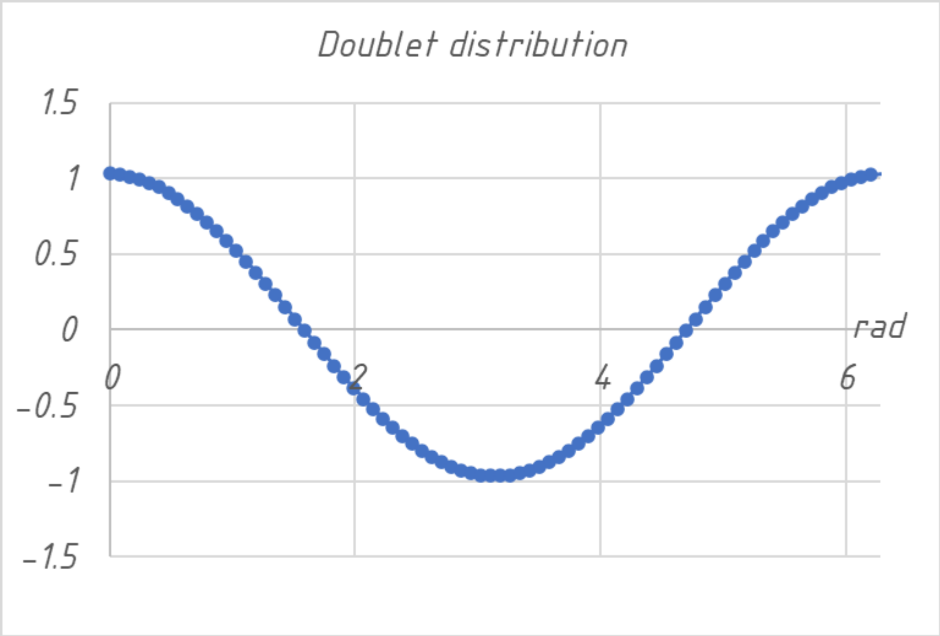


Fig. 5. Doublet distribution, circle, α = 0 deg, Re = 66177

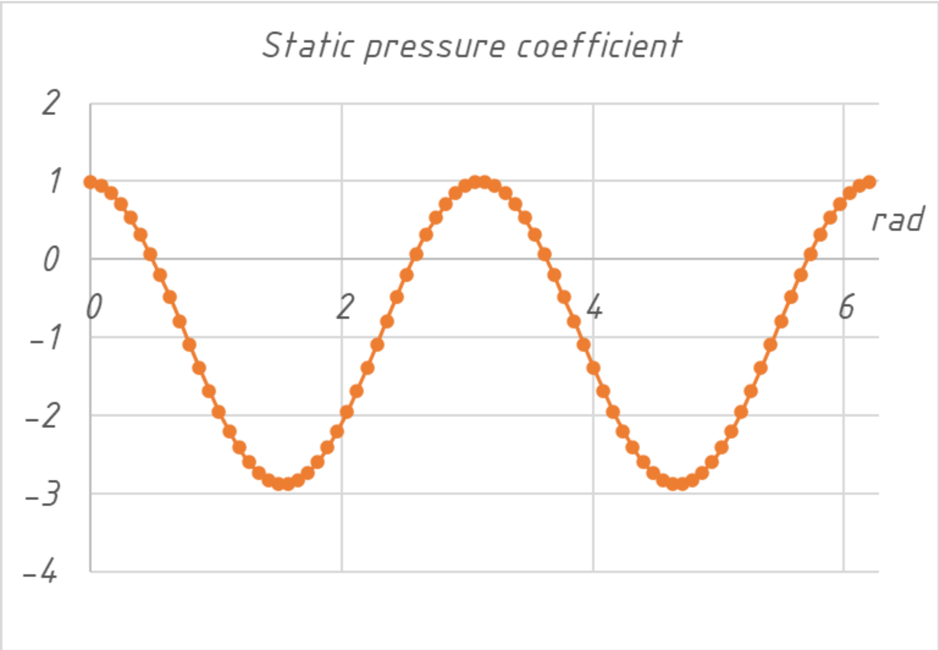


Fig. 6. Static pressure coefficient, circle, α = 0 deg, Re = 66177

Numerical results

In Fig. 7, numerical solution of system (4) regarding Mugin wingfoil is shown. Test case details are outlined in the caption below the figure.

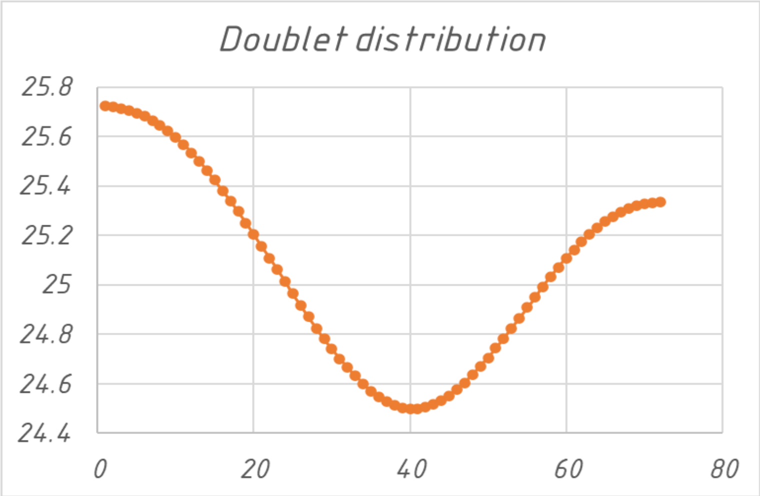


Fig. 7. Numerical solution of (4), Mugin 72 panels, Crout LU, α = 10 deg, Re = 66177

In Fig. 8, the static pressure coefficient is shown regarding the same test case.

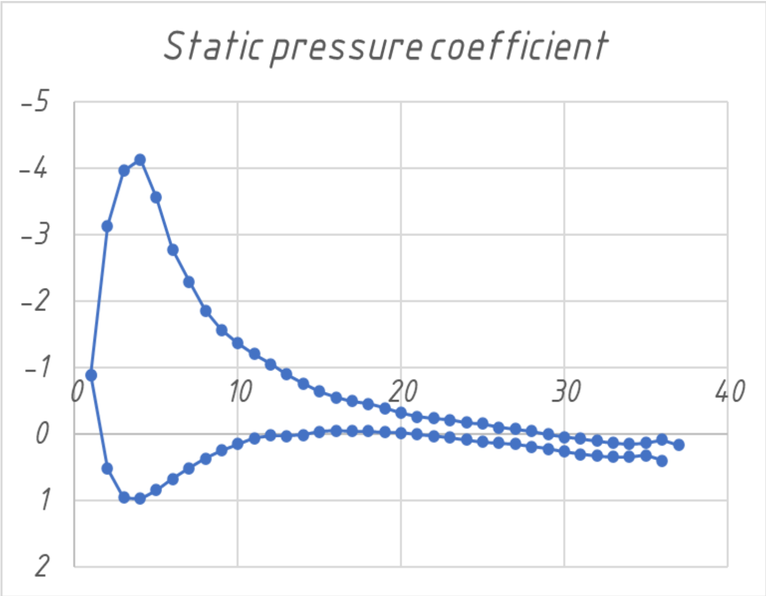


Fig. 8. Static pressure coefficient, Mugin 72 panels, α = 10 deg, Re = 66177

Concluding notes

Apart from static pressure coefficient, other secondary quantities might be computed, such as total lift and moment. Recalling the Kelvin's theorem, the total lift might be computed by taking the wake doublet strength, obtained after solving system (4). Bearing in mind small initial quantities $Q_\infty = 1$ m/s, $\rho = 1.2$ kg/m^3 $\mu_W = -0.386$ m^2/s, l = 1 m, then, following Joukowski theorem $L = -\rho Q_\infty \mu_W l = 0.465$ N.

References

1. https://www.muginuav.com/product/new-mugin-3m-uav-v-tail-platform-frame-kit/
2. Katz, J., A. Plotkin, *Low – Speed Aerodynamics*, Cambridge University Press, NY, 2001, p.p. 283, 340, ISBN 978-0-521-66219-2, ISBN 978-0-521-66552-0
3. ARRAYPOLAR (Command), Autodesk© Inc., Help file, 2020
4. Dahlquist, G., A. Bjorck, *Numerical Methods*, Dover Publications, 2003, p. 157, ISBN 0-486-42807-9
5. Press, W., S. Teukolsky, W. Vetterling, B., Flannery, *Numerical Recipes, The Art of Scientific Computing*, Third edition, Cambridge University Press, 2007, ISBN-13 978-0-521-88068-8, ISBN-10 0-521-88068-8
6. TCC. https://bellard.org/tcc/
7. Bertin, J., R. Cummings, *Aerodynamics for Engineers*, Sixth edition, Pearson Education Ltd., 2014, p.138, ISBN 10: 0-273-79327-6, ISBN 13: 978-0-273-79327-4

ПРИЛОЖЕНИЕ НА МЕТОДА НА ДИСКРЕТНИТЕ ДИПОЛИ С ПОСТОЯНА ИНТЕНЗИВНОСТ ЗА АНАЛИЗ НА СТАЦИОНАРНО ТЕЧЕНИЕ ОКОЛО КРИЛЕН ПРОФИЛ

К. Методиев

Резюме

В настоящата статия е приложен числен (панелен) метод за анализ на двумерно течение на идеален газ около крилен профил. Отначало профилът беше разделен на краен брой панели. Впоследствие панелите бяха заменени от диполи с постоянна интензивност. Добавен беше и панел (следа) към изходящия ръб на профила, за да се удовлетвори условието на Кута.

За да се реализира този алгоритъм, числена реализация беше разработена и компилирана на Tiny C Compiler. За да се реши числено получената линейна нехомогенна система уравнения, беше приложен директен метод на LU-декомпозиция на матрицата коефициенти, а именно Crout, Doolittle, Cholesky.

Получените резултати са валидирани с точно решение и показани за различни стойности на ъгъла на атака и числото на Рейнолдс.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e12

A NEW DESIGN METHOD FOR HIGH CONDITIONS APPLIED TO MINIMUM LENGTH NOZZLES

Mohamed Roudane[1, 3, 4], Merouane Salhi[1, 3, 4], Ahmed Boucherit[2, 3, 4]

[1] Mechanical Engineering Department, Faculty of Technology
[2] Chemical Process Engineering Department, Faculty of Technology
[3] Fundamental and applied physics laboratory, FUNDAPL
[4] University of Blida 1, BP 270 Blida 09000, Algeria

Keywords: *Rocket Nozzle, Perfect Gas, High Temperature Model*

Abstract

This present work focused on new nozzles design method, based on the characteristics method, which is a technique method to reduce a partial differential equation to linear differential equations along which the solution can be integrated from initial conditions. The latter is developed under the real gas theory, because when the both pressure and temperature of a gas increases, the specific heat and their ratio do not remain constant anymore and start to vary with the gas parameters. The gas doesn't stay perfect, and it becomes a real gas. The presented equations of the characteristics remain valid whatever area or field of study. With the assumptions that Berthelot's state equation accounts for molecular size and intermolecular force effects, expressions are developed for analyzing the supersonic flow for thermally and calorically imperfect gas. The resolution has been made by the finite differences method using the corrector predictor algorithm. As result, the developed mathematical model used to design 2D minimum length nozzles under effect of the stagnation parameters of fluid flow. A comparison for air with the perfect gas PG and high temperature HT models on the one hand and our results by the real gas theory on the other of nozzles are made. An important gain of length and weight can rise up to 40% and 20% respectively. It is in this context that Minimum Length Nozzle (MLN) nozzles for aerospace engines based on real gas theory were developed to achieve maximum thrust with the smallest possible nozzle weight (minimum length).

Introduction

In mathematics, the characteristics method is a technique for solving partial differential equations. Particularly suitable for transport problems, it is used in many fields such as fluid mechanics or particle transport [1]. In some particular cases, the characteristics method may allow the purely analytical resolution of the partial derivative equation. In more complex cases (encountered for example in modeling of physical systems), the characteristics method can be used as a method for the numerical resolution of the problem. For a first order partial differential equation, the characteristics method looks for curves called characteristic lines or more simply characteristics along which the partial differential equation is reduced to a simple ordinary differential equation. The resolution of the ordinary differential equation along a characteristic makes it possible to find the solution of the original problem.

The need for improving the performance of supersonic nozzles plays a very important role in the field of propulsion and aerospace engineering. The supersonic nozzle contains a very important weight in an engine, including missiles and supersonic aircraft [2]. The weight of the nozzle can reach 80% of the total weight of the engine [3, 4]. So it's interesting to find physical solutions generally to the different problems for improving performances. The performance of a supersonic nozzle is usually the mass of the structure, the thrust coefficient, the exit Mach number delivered, since the design of the nozzle is made on the basis of a non-viscous fluid.

Among several known types of nozzles in aerospace industry. it is expected that there are about twenty forms of nozzles; one is interested in our work with the nozzle with centered expansion or by the Minimum Length Nozzle (MLN) nozzle [5, 6], for reason that the international construction of several aerospace projects currently use this type of nozzle [7]. The resolution of the conservation equations is done in the first step by the characteristic method in order to transform these equations to simplified coupled nonlinear algebraic equations according to privileged directions called the characteristics. In this case the equations obtained are considerably simplified, but the mesh calculation becomes very complex [8]. The mesh generation is done in parallel with the calculation of the parameters of the flow.

After a literature search, it has been noted that the majority of published work in the field of supersonic nozzle design is based on the use of two models which are either the perfect gas model with constant specific heat CP [9, 10] or the high temperature model when CP is a function of temperature [11]. These assumptions will not take into account the real behavior of the gas when the stagnation pressure is high. In this case, the mathematical model of the calculation changes and must be completely revised.

With the advent of space propulsion, engine manufacturers were constrained by a specification limiting the weight and length of the diverging part of a nozzle, to be defined, according to the optimum of the sections ratio and the weight, while trying to minimize thrust losses compared to the ideal nozzle. The problem encountered in aerospace applications is that the use of nozzles designed on the basis of the perfect gas assumptions degrades the performance desired by this nozzle [12]. If we take measurements by experience, we will find values different from those determined by calculation; especially if the stagnation temperature and pressure of the combustion chamber are high. As current and future rocket engines have to be adapted to high pressures and temperatures, the concept of an ideal nozzle is excluded because it would lead to a length and a weight of the divergent which are prohibitive [13]. Engine manufacturers constrained by a specification limiting the weight and length of a space engine nozzle are led to seek an optimum of the thrust-to-weight ratio for a fixed sections ratio.

In this work, we will present the method of design and dimensioning of two-dimensional centered expansion type nozzles or MLN nozzles using the new form of method of the characteristics. We have added the effect of the gaseous imperfections on all parameters and then the method becomes a function of the temperature and the density, and strongly stays valid when the stagnation temperature and pressure of the combustion chamber are high, lower than the dissociation threshold of the molecules.

Materials and Methods

The application of the minimum length nozzle with straight sonic line is used for hypersonic wind tunnels as well as rocket motors [14]. The study is limited for the case of the two dimensional minimum length nozzles.

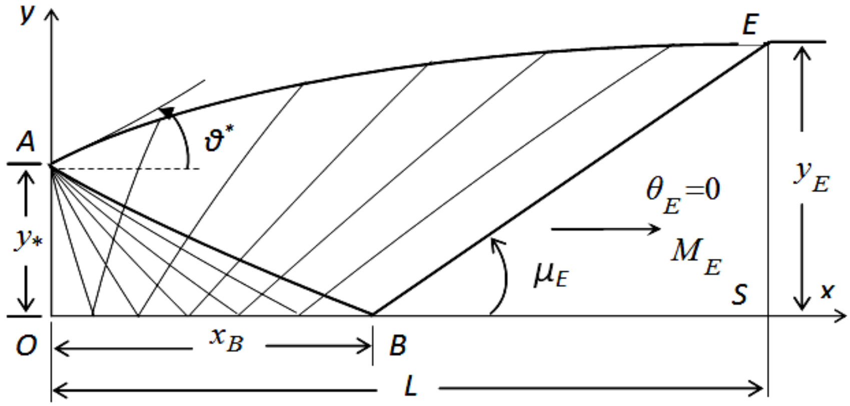


Fig. 1. Presentation of the flow field in the bidimensional centered expansion nozzle

Fig. 1 illustrates the general scheme of the minimum length nozzle with straight sonic line and represents the characteristics of the flow field in different regions [5]. This nozzle is called a centric expansion nozzle (MLN). The flow between the throat OA and the uniform region BES is divided into two regions. The OAB region, called by Kernel region, is a region of non-simple waves. The triangular region BES is a uniform flow region with exit Mach number M_E. In this contest, the wall, at the throat, is inclined at an angle θ^*.

For a supersonic, irrotational, adiabatic flow, the characteristic method gives the following equations, called equations of characteristics and compatibilities [6]:

According to ξ (1-3):

$$(1)\qquad \begin{cases} d(\nu+\theta)=\delta\dfrac{\sin\theta.\sin\mu}{y}d\xi \\ \dfrac{dy}{dx}=\tan(\theta-\mu) \end{cases}$$

According to η (2-3):

$$(2)\qquad \begin{cases} d(\nu-\theta)=\delta\dfrac{\sin\theta.\sin\mu}{y}d\eta \\ \dfrac{dy}{dx}=\tan(\theta+\mu) \end{cases}$$

Equations (1) and (2) are valid for C^- and C^+, respectively, as shown in Fig. 2. In the real case, the characteristics are curved, and if the mesh is fine so that the points are close to each other, we can bring the curvature by a straight line, the calculation will be on the lines of Mach named ξ on characteristic C^- and η on characteristic C^+ as shown in Fig. 2.

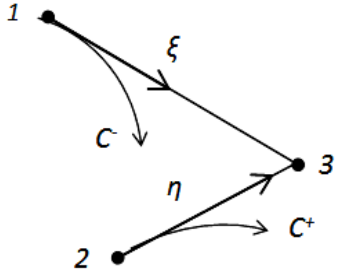


Fig. 2. Illustration of characteristic lines and Mach lines.

The relationships in the system of equations (1) and (2) are developed for our model in previous work:

The new form of Prandtl Meyer function is given by [15]:

(3) $$d\upsilon(T,\rho) = -\left(\frac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)} .\, C_P(T,\rho) \right) dT - \left(\frac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)} .C_T(T,\rho) \right) d\rho$$

The specific heats ratios at constant pressure and volume [16]:

(4) $$c_p(T,\rho) = c_{p(gp)} \left[1 + \frac{\gamma_{(gp)}-1}{\gamma_{(gp)}} \left\{ \left(\frac{\theta}{T^2} \right) \frac{e^{\frac{\theta}{T}}}{\left(1-e^{\frac{\theta}{T}} \right)^2} + \frac{2a\rho}{RT^2} \left[1 + \frac{\left(\frac{2-b\rho}{1-b\rho} + \frac{a\rho}{2RT^2} \right)}{\frac{1}{(1-b\rho)^2} - \frac{2a\rho}{RT^2}} \right] \right\} \right]$$

(5) $$C_T(T,\rho) = \left(\frac{3ab^2\rho^2 - 6ab\rho - RT^2 b + 3a}{2T\rho b - Tb^2\rho^2 - T} \right)$$

(6) $$c_v(T,\rho) = c_{v_{gp}} \left\{ 1 + (\gamma_{gp}-1) \left(\left(\frac{\theta}{T} \right)^2 \frac{e^{\left(\frac{\theta}{T}\right)}}{\left(1-e^{\left(\frac{\theta}{T}\right)} \right)^2} + \frac{2a\rho}{RT^2} \right) \right\}$$

(7) $$\gamma(T,\rho) = \left[C_p(T,\rho) / c_v(T,\rho) \right]$$

The flow and sound velocities [16, 17]

(8) $$V^2(T,\rho) = 2\left\{ c_{v_{gp}} \left(T_0 - T \right) + R\theta \left(\frac{1}{\left(1-e^{\left(\frac{\theta}{T}\right)} \right)} - \frac{1}{\left(1-e^{\left(\frac{\theta}{T_0}\right)} \right)} \right) + 4a\left(\frac{\rho}{T} - \frac{\rho_0}{T_0} \right) + \left(\frac{p_0}{\rho_0} - \frac{p}{\rho} \right) \right\}$$

(9) $$a^2(T,\rho) = \frac{RT}{(1-b\rho)^2} - \frac{2a\rho}{T} + \frac{\rho^2 T \left(\frac{a}{T^2} + \frac{R}{\rho(1-b\rho)} \right)^2}{c_{v_{gp}} . \left[1 + (\gamma_{gp}-1) \left\{ \left(\frac{\theta}{T} \right)^2 \frac{e^{\left(\frac{\theta}{T}\right)}}{\left(1-e^{\left(\frac{\theta}{T}\right)} \right)^2} + \frac{2a\rho}{RT^2} \right\} \right]}$$

The Mach number and the Mach angle are given by [10]:

(10) $$M^2(T,\rho) = \left(V^2(T,\rho) / C_S^2(T,\rho)\right)$$

(11) $$\mu\ (T,\rho) = arcsin\left(1/M(T,\rho)\right)$$

The integration of the systems (1) and (2) gives:

According to ξ (1-3):

(12) $$\begin{cases} \int_{T_1}^{T_3}\left(-\dfrac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)}.\, C_P(T,\rho)\right) dT \ + \int_{\rho_1}^{\rho_3}\left(-\dfrac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)}.C_T(T,\rho)\right) d\rho + \left(\theta_3-\theta_1\right) = \\ \int_{x_1}^{x_3}\delta\dfrac{\sin\theta.\sin\mu}{y.\cos\left(\theta-\mu\right)}dx \\ y_3 - y_1 = \int_{x_1}^{x_3}\tan\left(\theta-\mu\right)dx \end{cases}$$

According to η (2-3):

(13) $$\begin{cases} \int_{T_2}^{T_3}\left(-\dfrac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)}.\, C_P(T,\rho)\right) dT \ + \int_{\rho_2}^{\rho_3}\left(-\dfrac{\sqrt{M^2(T,\rho)-1}}{V^2(T,\rho)}.C_T(T,\rho)\right) d\rho - \left(\theta_3-\theta_2\right) = \\ \int_{y_2}^{y_3}\delta\dfrac{\sin\theta.\sin\mu}{y.\sin\left(\theta+\mu\right)}dy \\ y_3 - y_2 = \int_{x_2}^{x_3}\tan\left(\theta+\mu\right)dx \end{cases}$$

When our systems have five unknowns (x, y, θ, T and ρ), we need to add another equation presented as follows:

According to ξ (1-3):

(14) $$\int_{\rho_1}^{\rho_3}\left[\frac{1}{\rho} - \frac{C_T(T,\rho)}{a^2(T,\rho)}\right] d\rho = \int_{T_1}^{T_3}\left[\frac{C_P(T,\rho)}{a^2(T,\rho)}\right] dT$$

According to η (2-3):

$$\int_{\rho_2}^{\rho_3}\left[\frac{1}{\rho}-\frac{C_T(T,\rho)}{a^2(T,\rho)}\right]\,d\rho=\int_{T_2}^{T_3}\left[\frac{C_P(T,\rho)}{a^2(T,\rho)}\right]\,dT \tag{15}$$

As mentioned, the characteristics C- and C + are curves, the application of the method of the characteristics obliges to introduce a fine mesh in order to approximate each characteristic between two points by segments of straight line [18]. The properties (x, y, T, θ, ρ, P) at a point of flow field can be determined from those of the two points connected with the point considered by the characteristic lines which precede it. For example the properties in point 3 of Fig. 2 can be determined from those of points 1 and 2 which connect them.

Error of Perfect Gas and High Temperature Models

The mathematical perfect gas model is developed on the basis to regarding the specific heat C_P and ratio ϒ as constants, which gives acceptable results for low temperature. According to this study, we can notice a difference on the given results between the perfect gas model and the developed model. The error given by the PG model compared to our RG (real gas) model can be calculated for each parameter. Then, for each value (P_0, T_0, M), the ε error can be evaluated by the following relationship [19, 20]:

$$\varepsilon_{Parameter}(\%)=\left|1-\frac{Parameter_{PG\ or\ HT}\left(P_0,T_0,M\right)}{Parameter_{RG}\left(P_0,T_0,M\right)}\right|\times 100 \tag{16}$$

As a rule for the aerodynamic application, the error should be lower than 5%.

Results and Comments

Effect Of Discretization And Mesh Refinement On The Convergence

Fig. 3, shows the steps of the mesh by the insertion of the additional characteristics Ni between the sonic line and the first descending characteristic as well as the injection of a condensation function Δ between the two last descending characteristics, the final quality of the mesh of the 2D MLN nozzle for Δθ = 0.6°, Ni = 10 and Δ = 8. It can be said that the number of N_C points found on the last C- depends on the exit Mach number, the stagnation temperature T_0, the stagnation pressure P_0, the step Δθ, the number of inserted characteristics Ni, the coefficient of condensation Δ. Note that it is very interesting to refine the mesh on the wall in the vicinity of the neck, because the nature of a supersonic flow determines the properties at a point as a function of two points which are upstream. Then, a poor presentation of the wall at the thoat will propagate and enlarge the errors at the exit, and therefore, we will determine a bad pace of the wall. The control of the results is done by the use of the ratio of the sections which remains always valid since the flow at the exit of the nozzle is uniform and parallel.

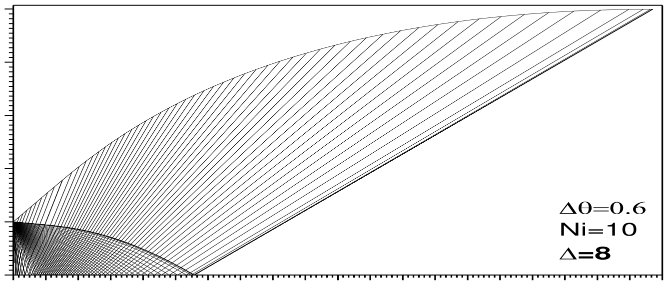
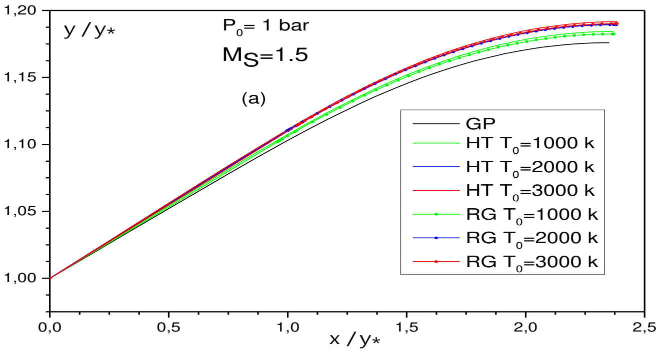
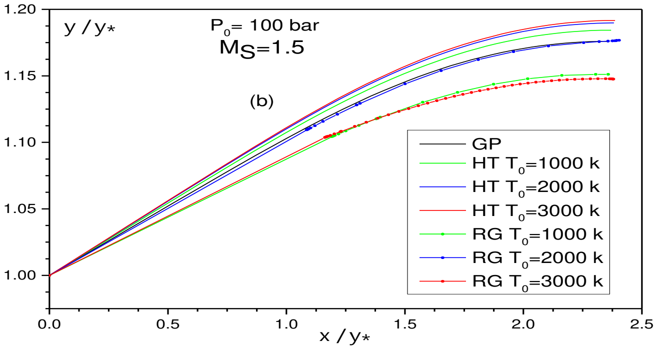


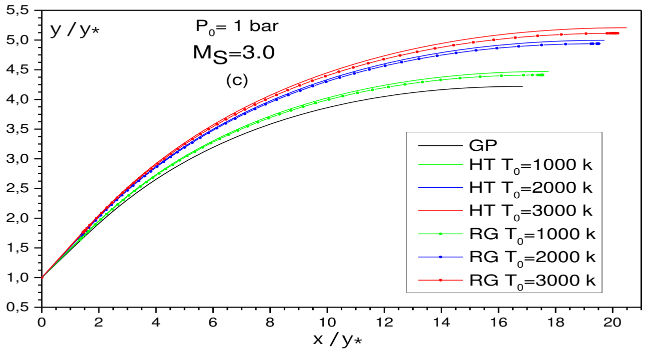
Fig. 3. Refinement of the mesh in the Kernel zone of the MLN 2D nozzle with the effects of the insertion of additional characteristics Ni and the condensation coefficient Δ

Effect Of The Stagnation Conditions T_0 And P_0 On The Wall Shape

Fig. 4, represent the variation of the nozzles shape y / y* obtained when the exit Mach number M_S is equal to 1.50, 3.0 and 6.0 given respectively by the stagnation temperatures T_0 = 1 000 K, 2 000 K, and 3 000 K and for the stagnation pressures P_0 = 1 bar, 10 bar, and 100 bar, as a function of the abscissa number x / y*, of the RG model compared to the PG and HT models. We note that the increase of the ratio x / y* for different models leads to an increase of y / y *, we also note that the variation of the stagnation temperature for the values P_0 = 1 bar, 10 bar and 100 bar, influences on the ratio y / y*, the shape decrease when P_0 increases, which is not the case for the PG and HT models, this reduction is more important and remarkable when the stagnation pressure P_0 and the exit Mach number are high (see Fig. 4.f), hence the need to use the RG model to correct the results, and to show the effect of the stagnation pressure P_0 on the design. Between the figures presented, we can say that if the stagnation pressure P_0 increases, the difference between the GP, HT models and our RG model enlarged and becomes considerable, independently of the exit Machnumber M_S, or from $M_S > 2.00$ for any pressure P_0. This limit can be found if we choose an ε error less than 5%.


y / y*
P0= 1 bar
MS=1.5
(a)
GP
HT T0=1000 k
HT T0=2000 k
HT T0=3000 k
RG T0=1000 k
RG T0=2000 k
RG T0=3000 k
x / y*
1,20
1,15
1,10
1,05
1,00
0,0
0,5
1,0
1,5
2,0
2,5


y / y*
P0= 100 bar
MS=1.5
(b)
GP
HT T0=1000 k
HT T0=2000 k
HT T0=3000 k
RG T0=1000 k
RG T0=2000 k
RG T0=3000 k
x / y*
1.20
1.15
1.10
1.05
1.00
0.0
0.5
1.0
1.5
2.0
2.5


y / y*
P0= 1 bar
MS=3.0
(c)
GP
HT T0=1000 k
HT T0=2000 k
HT T0=3000 k
RG T0=1000 k
RG T0=2000 k
RG T0=3000 k
x / y*
5,5
5,0
4,5
4,0
3,5
3,0
2,5
2,0
1,5
1,0
0,5
0
2
4
6
8
10
12
14
16
18
20

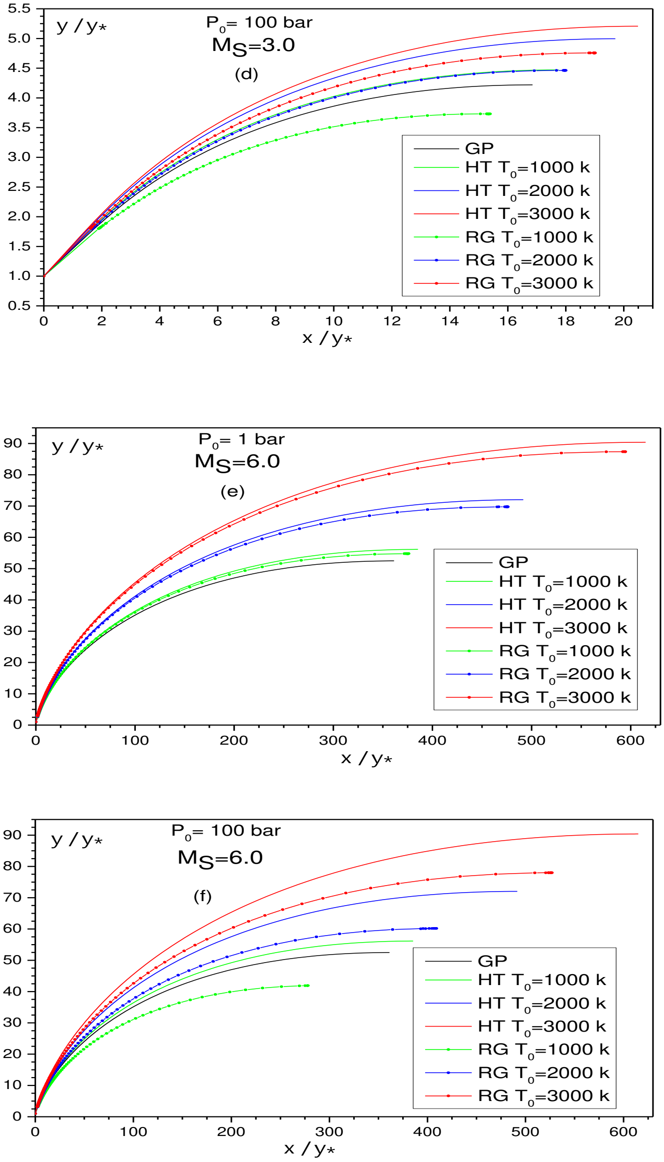


Fig. 4. 2D MLN nozzle shapes for some exit Mach number values

Variation Of The Parameters Through The Nozzle

Fig. 5 represent the variation of the Mach number along the wall of the nozzle as a function of the ratio x / y*, for the HT, GP and RG models. We can clearly remark that the increase in the Mach number for M = 1 at the collar at M = M* just after the expansion, then at $M = M_S$ at the exit section of the nozzle.

The example taken here is for $M_s = 6.00$. There is an uniform Mach number constant along the wall near the throat regardless of T_0 and P_0, which is interpreted by the existence of a nearly uniform flow zone in this region for this type of nozzle. The increase in the Mach number through the wall is interpreted by the expansion of the gas to the exit section.

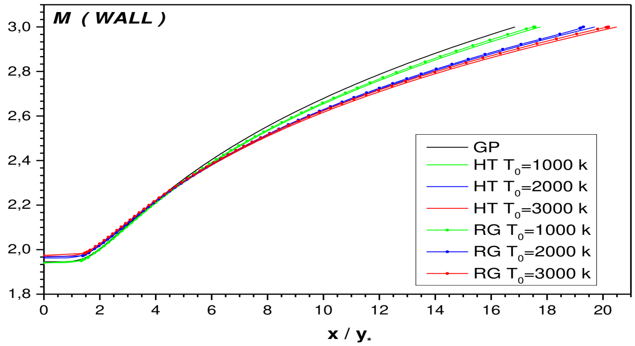


Fig. 5. Variation of the Mach number along the wall of the nozzle for $M_E = 3.0$

Conception Parameters

Figs. 6–9 present, the variation of the Kernel zone length, the total length, the nozzle structure weight and the exit section area of the nozzle respectively, as a function of Mach number M_E and T_0, for GP, HT and RG models, it is noted that the more the nozzle delivers a high exit Mach number, the higher these results become important. The purpose of the presentation of this variation is that, from the length of this zone, one can deduce the length of the nozzle directly without making the calculation of the flow in the transition zone. Always note that the curves are confused at low Mach number up to about $M_E = 2.00$. From this value, the curves start to differentiate, and the results obtained by our RG model are away from those obtained by the HT model when the temperature T_0 increases, we can say that the perfect gas theory gives good results if this condition is verified. From these results it can be said that there are significant gains in the length and weight of the nozzle.

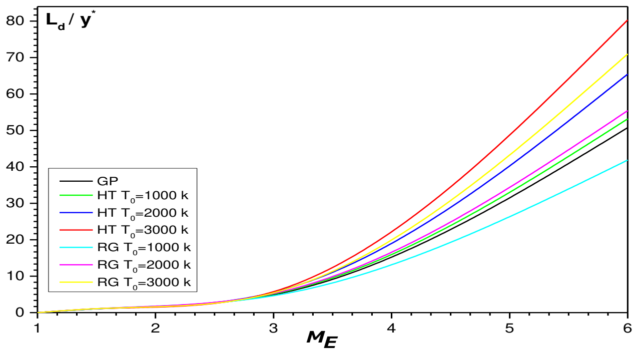


Fig. 6. Variation of the Kernel zone length of the nozzle

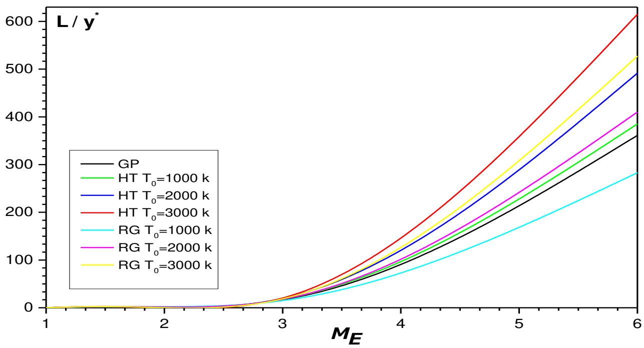


Fig. 7. Variation of the total length of the nozzle

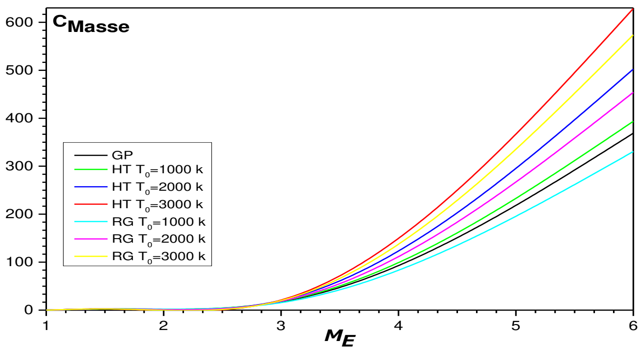


Fig. 8. Variation of the nozzle structure weight

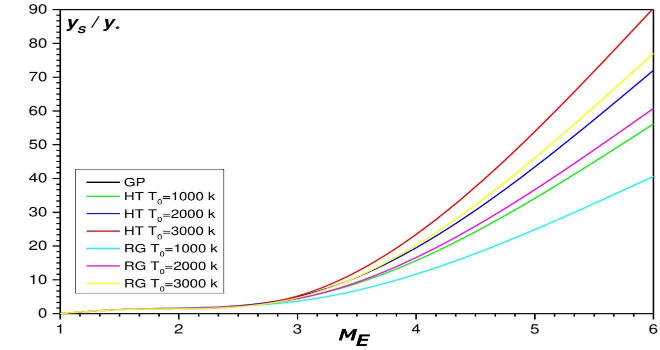


Fig. 9. Variation of the exit section area of the nozzle

Fig. 10 shows, the variation of the thrust coefficient C_F versus exit Mach number, for *HT*, *PG* and *RG* models, we note that the increasing the Mach number for different models leads to an increase of thrust coefficients, and the variation of the stagnation temperature, e.g. for T_0 = 1 000 K, 2 000 K, 3 000 K, influences the C_F values. Therefore, the C_F coefficient increases when T_0 increases, which is not the case for the perfect gas model. Otherwise, if the exit Mach number is less then M_E = 2.0, we note that the three *PG*, *HT* and *RG* models are confounded, and moving away when M_S increases. For instance, if M_s = 5.0 and T_0 = 3 000K, C_F = 1.65234 for the *PG* model, C_F = 1.73381 for the *HT* model, C_F = 1.73004 for *RG* model, with a relative error between the *HT* model and for our *RG* model equal to ε = 1.22%. Therefore, when the stagnation parameters increases the thrust coefficient C_F obtained by the *RG* model moves away from those obtained by *HT* model, which shows the effect of the stagnation parameters on the nozzle performances.

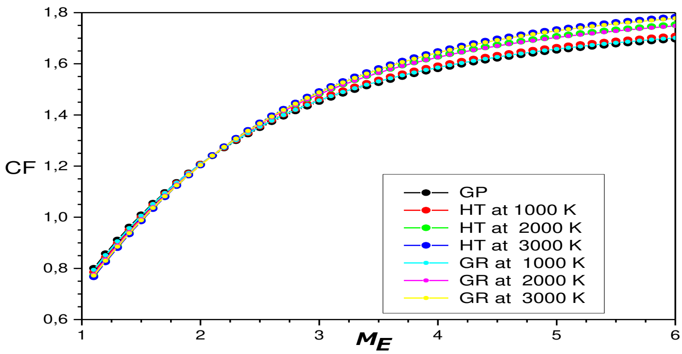


Fig. 10. Variation of the pressure force exerted on the nozzle wall

Results On The Conception Errors

Fig. 11 presents, the variation of the relative error given by the length and the exit section area of the nozzle given respectively by the generating temperatures T_0 = 1000 K, 2000 K, and 3000 K, it is clear that the error depends on the values of T_0 and M_E , and increases if M_E increases, and decreases if the temperature T_0 increases. For example, if T_0 = 1000 K and M_S = 6.00, the use of the HT model will give us a relative error equal to ε = 36.03%, the latter will decreases up to ε = 15.56% when temperature T_0 = 3000 K for the length of the nozzle. It is clear that if we choose an error for example lower than 5%, the *PG* model may be used, if T_0 is less than 1000 K for any value of the Mach number. If an author accepts an error greater than 5%, he can use the PG model in moderate interval of M, P_0 and T_0.

It may be noted that, at low values of M_E, the error ε is small. In these figures we find the error below 5%. This position is interpreted by the possibility of using the *PG* model for the aeronautical applications, if we accept an error less than 5%. Otherwise, if the temperature T_0 is low, the error increases progressively, in this case, we can use the *PG* model independently to the temperature T_0, when the Mach number does not exceed M = 2.0 with an error of about 5.5%.

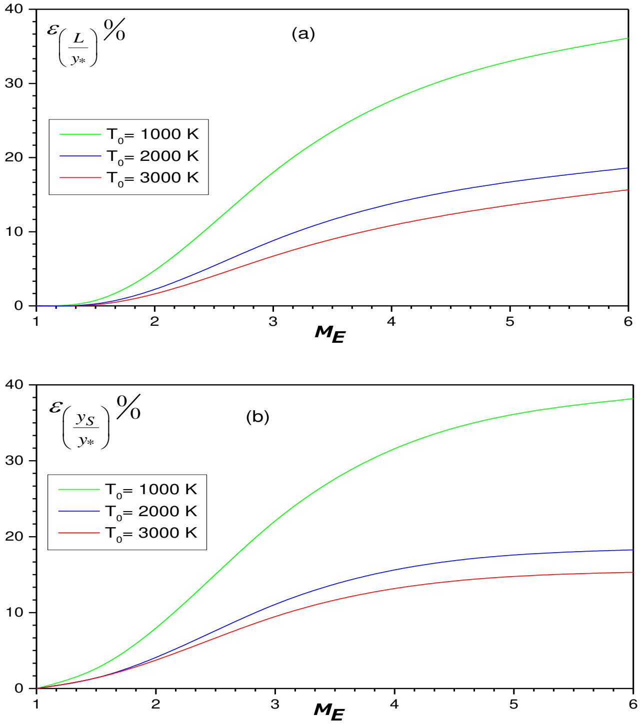


Fig. 11. Variation of the relative error given by the length and the exit section area of the nozzle

Conclusion

From this study, we can highlight the following points:

If we accept an error lower than 5%, which is generally the case for aerodynamic applications, we can study a supersonic flow using the relationships of a perfect gas, if T_0 is less than 1000 K for any value of the Mach number, with a stagnation pressure well definite and moderate.

The *PG* model is represented by explicit and simple relations and does not require much time to make calculation, which is not the case for our *RG* model, where it is presented by solving a nonlinear algebraic equations, solving a nonlinear algebraic equation system formed by three equations and derivation and integration of a complex analytical functions require and take more for calculation time and numerical programming and data processing.

At low temperature, the difference in results obtained between the *PG* and RG models is small, which gives the opportunity to study *RG* flow using the *PG* relations, especially when P_0 increases. By cons, when T_0 increases, the *PG* theory starts to give results moving away progressively from the real cases, where we need to use the *RG* model.

At low pressure, the difference in results obtained between the HT and RG models is small, which gives the opportunity to study *RG* flow using the relations *HT*, especially when T_0 increases. Otherwise, when P_0 increases, the *HT* theory starts to give results moving away progressively from the real cases, where we need to use the *RG* model.

If the MLN 2D nozzle of the *RG* model all deliver the same exit Mach number M_E as delivered by the MLN 2D nozzle for *HT* and *GP* models, they all however have a mass lower than it. This gain in mass can increases up to 20%, which is very significant in aerospace applications. But their thrust coefficients remain constant because they have the same exit Mach number M_E.

Significant gains was found on Kernel length and the total nozzle length, as well as the nozzle exit section area and the nozzle structure weight, then an improvement in nozzle performance by reducing the volume occupied by the nozzle and its mass, this variation of the weight can replace the increase of the payload of the apparatus.

Since the flow at the exit section is horizontal, the nozzle may be truncated to a section having a velocity deviation of one or two degrees. In this case, we make a large gain in mass and reduce the weight. The flow at the exit of the truncated nozzle becomes inclined in the vicinity of the wall, and in this case the pressure force exerted on the inner wall of the nozzle will decrease slightly.

References

1. AI-Ajlouni, M. An Automatic Method for Creating the Profile of Supersonic Convergent-Divergent Nozzle, Jordan J Journal of Mechanical and Industrial Engineering.,2010, 4, 3, 404–11.
2. Malina, F. J. Characteristics of the rocket motor based on the theory of perfect gases, J. Franklin Inst.,1940, 230, 433–50.
3. Annamalai, K. K., Ishwar, and P. Milind, A. J. Advanced Thermodynamics engineering. USA, 2nd edition, CRC press: Taylor and Francis Group, 2011, 1142.
4. Peterson, C. R., and Hill, P. G. Mechanics and Thermodynamics of Propulsion. New York, Addition-Wesley Publishing Company Inc, 1965, 754.
5. Zebbiche, T., and Youbi, Z. Supersonic Two-Dimensionnel Minimum Length Nozzle Design at High Temperature. Application for Air, Chinese Journal of Aeronautics., 2007, 20, 01, 29–39. https://doi.org/10.1016/S1000-9361(07)60004-1.
6. Zebbiche, T., and Youbi, Z. Supersonic Two-Dimernsional Minimum Length Nozzle Design at High Temperature. Application for Air. 42st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Sacramento, California, 9 - 12 July 2006, AIAA-2006-4599.
7. Sutton, G. P. Rocket Propulsion Elements. 8th edition, New Jersey, USA: John Wiley & Sons Inc., Hoboken, 2010, 792.
8. Serra, R. A. Determination of Internal Gas Flows by a Transient Numerical Technique, AIAA Journal., 1972, 10, 5, 603–11.
9. Anderson, J. D. Fundamentals of Aerodynamics. McGraw-Hill Book company, New York, 1988, 1152.
10. Anderson, J. D. Modern Compressible Flow With Historical Perspective. 2nd edition, New York, USA: McGraw-Hill Book Company, 1992, 466.
11. Zebbiche, T., and Youbi, Z. Effect of Stagnation Temperature on the Supersonic Flow Parameters with Application for Air in Nozzles, The Aeronautical Journal., 2007, 111, 1115, 31–40. https://doi.org/10.1017/S0001924000001731.
12. Zebbiche, T., and Youbi, Z. Design of Two-Dimensional Supersonic Minimum Length Nozzle at Higth Temperature, Application for Air. German Aerospace Congress, Friendrichshafen, Germany, 26-29 Sep. 2005, DGLR 2005-257.
13. Argrow, B. M., and Emanuel, G. Computational Analysis of the Transonic Flow field of Two-Dimensional Minimum Length Nozzles. 20th Fluid Dynamics, Plasma Dynamics and Lasers Conference, Buffalo, NY, 12-14 June 1989, AIAA-1989-1822.
14. Argrow, B. M., and Emanuel, G. Comparison of Minimum Length Nozzles, Journal of Fluids Engineering., 1988, 110, 283–88. https://doi.org/10.1115/1.3243546.
15. Salhi, M., and Zebbiche, T. A New General Form of Prandtl Meyer Function,Application on Air. 8th AIAA Theoretical Fluid Mechanics Conference, Denver, Colorado, USA, 5-9 June 2017, 2017. AIAA 2017- 2621484, APA-11.
16. Salhi, M., Zebbiche, T., and Mehalem, A. Stagnation pressure effect on the supersonic flow parameters with application for air in nozzles, The Aeronautical Journal., 2016, 120, 1224, 313–54. DOI: https://doi.org/10.1017/aer.2015.13.

17. Salhi, M., and Roudane, M. Numerical Investigation of the Thermal Caloric Imperfections on entropy Enhancement across Normal Shock Waves, High Temperatures High pressures Journal., 2019, 48, 4, 285–308.
18. Salhi, M., and Zebbiche, T. Gaseous imperfections effect on the supersonic flow parameters for air in nozzles. 3rd International Conference and Exhibition on Mechanical & Aerospace Engineering, San Francisco, USA,05-07 Oct 2015,74.
19. Salhi, M., and Kirad, A. Roudane, M. Méthode des éléments finis en génie mécanique, Simulations et applications. European Universities Publishing, ISBN : 978-613-9-52187-6, 2019, 432.
20. Zebbiche, T. New Generalized Form of the Prandtl Meyer Function. Application for Air at High Temperature. 25th Applied Aerodynamics Conference, San Francisco, 25-28 June 2007. AIAA-2006-3674.

НОВЫЙ МЕТОД ДИЗАЙНА ДЛЯ ВЫСОКИХ УСЛОВИЙ, ПРИМЕНЯЕМЫЙ НА СОПЛАХ МИНИМАЛЬНОЙ ДЛИНЫ

М. Рудан, М. Салхи, А. Бушерит

Аннотация

Резюме, настоящая работа сосредоточена на новом методе проектирования сопел, основанном на методе характеристик, который представляет собой технический метод для сведения уравнения в частных производных к линейным дифференциальным уравнениям, по которым решение может быть интегрировано из начальных условий. Последний был разработан в соответствии с теорией реального газа, потому что, когда давление и температура газа увеличиваются, удельная теплоемкость и их соотношение больше не остаются постоянными и начинают изменяться в зависимости от параметров газа. Газ не остается идеальным, и он становится настоящим газом. Представленные уравнения характеристик остаются в силе независимо от области или области исследования. Исходя из предположения, что уравнение состояния Бертло учитывает размер молекул и эффекты межмолекулярных сил, разработаны выражения для анализа сверхзвукового потока для термически и калорически несовершенного газа. Разрешение было выполнено методом конечных разностей с использованием алгоритма корректора-предсказателя. В результате на основе разработанной математической модели спроектированы двухмерные сопла минимальной длины с учетом параметров торможения потока жидкости. Сделано сравнение для воздуха с идеальным газом PG и высокотемпературными моделями HT, с одной стороны, и наши результаты по теории реального газа, с другой стороны, для сопел. Значительный прирост длины и веса может достигать 40% и 20% соответственно.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e13

AERODYNAMIC PERFORMANCE OF CHEVRON GEOMETRY TRAILING EDGE FOR A TRANSONIC AXIAL COMPRESSOR IMPELLER

***Ioana Octavia Bucur*[1], *Ion Malael*[1], *Valeriu Dragan*[1]**

[1]National Research and Development Institute for Gas Turbines COMOTI, 220D Iuliu Maniu Bd., sector 6, cod 061126, OP 76, CP174, Bucharest, Romania.
e-mail: ioana.bucur@comoti.ro

Key words: *Attenuate Noise Level, Transonic Axial Compressor, Chevron Geometry*

Abstract

Several studies present how serrated rotor blades mix wakes in order to attenuate noise levels. The current paper analyses how this geometry, applied on the trailing edge, affects the global parameters of a transonic axial compressor impeller. Innovative solutions tackling the rotor-stator interaction mechanism in an axial compressor for noise reduction include serrated trailing edges. Inspired by chevron nozzles, serrations can be transferred to the open-rotor concept in order to reduce tonal noise. Throughout the study we will be focusing on aerodynamic loss estimation while being mindful of the mechanisms which lead to rotor-stator interaction noise, without assessing its per-se effectiveness for noise mitigation. Owing to its qualitative experimental data availability, NASA's Rotor 37 was chosen as a baseline. A set of fully viscous 3D simulations, using the SST k-omega turbulence model and RANS, was carried out to this effect. Spatial discretization was made using a fully structured pre-mesh in order to optimize resolution and accelerate convergence. Full-factorial samples were generated for the geometric variations in order to capture the aerodynamic implications of this concept. Overall, the analysed case provides promising perspectives, pending optimization studies and experimental tests thereof.

Introduction

Current concerns regarding the state of the art referring to transonic axial compressors mainly relates to optimization solutions and development ideas that can help in engine downsizing, without adversely impacting performances [1]. Understanding flow phenomena through the axial compressor is of vital importance in the context of modern growth, as most optimization methods rely on flow control through various methods (for example design modifications) [2].

Blade optimization represents a problem of interest in the aerospace engineering field as it may consistently improve overall performances of the propulsion system and its components. Hence, there are multiple innovative approaches considering blade improvement that can prevent loss and malfunction in axial compressors [3] and many recent studies concentrate on blade shape optimization using CFD methods.

A well-known and largely used transonic axial compressor test case for optimization is NASA's Rotor 37. First, it was tested as a single stage by Lonnie Reid and Royce D. Moore in 1978 [4], then again in 1980, this time as an isolated component [5] and so its geometry began to be identified by NASA and other developers in the field as Rotor 37 [6].

This test case is still a present topic and represents an important evaluation base in terms of overall performances and parameters for studies using CFD codes and other modern computational methods. Some researches based on the Rotor 37 case are [6, 7]. Centralized data for Rotor 37, including overall performances and parameters of interest, is presented in Tab. 1.

Table 1. Design Overall performance parameters for Rotor 37, [5]

Parameters	Value
Rotor total pressure ratio	2.106
Rotor total temperature ratio	1.270
Rotor adiabatic efficiency	0.877
Rotor polytropic efficiency	0.889
Rotor head rise coefficient	0.333
Flow coefficient	0.453
Airflow per unit frontal area	100.950
Airflow per unit annulus area	200.549
Airflow [kg/s]	20.188
RPM	17 188.700
Tip speed [m/s]	454.136
Hub-tip radius ratio	0.70
Rotor aspect ratio	1.19
Number of rotor blades	36

In paper [8] is presented an attempt to create an optimized design for the Rotor 37s blade. The proposed shape for the blade aimed to slightly improve aerodynamic performances compared to the results provided by NASA's technical reports mentioned above.

In the context of modern development, studies have been carried out concerning optimization solutions for rotor blades. Those address both overall performances of the rotor and also mitigation of tonal sound or broadband sound that occurs in the compressor's rotor [9].

General Electric brings an innovative proposal that consists of using serrated fan blades in order to attenuate noise levels. This technique should not have a negative impact on the effectiveness of the engine or its performances, they claim [10]. For a better understanding of their invention, Fig. 1 is provided from the original patent.

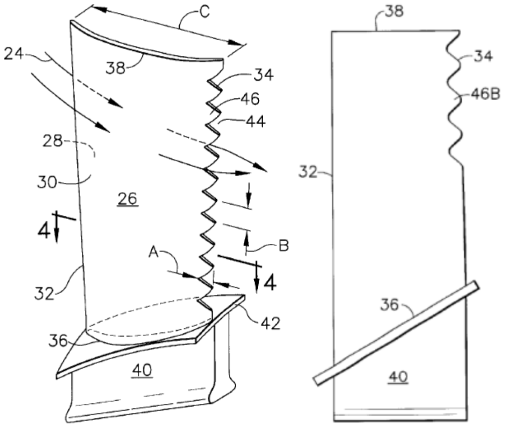


Fig. 1. Isometric and side view of an isolated fan blade embodiment [10]

The flow through an axial fan stage is analyzed by J. M. Fernández et al. in paper [11], using a LES based simulation to assess better understanding of rotor-stator interaction in the fan stage.

Jet engines have benefited from chevron technology for noise reduction ever since its advent. Since the successful implementation of chevrons for turbofan engines, other jet related applications arose such as jet pumps for ventilation [12, 13].

Hence, several studies exploited chevron patterns in order to obtain solutions regarding noise attenuation for a sustainable development.

For example, trailing edge serrations were used in the improvement of an open-rotor. Their purpose was to mitigate tonal interaction noise. The results of steady-state RANS simulations concluded that this approach was successful in providing acoustical optimization [14].

Another recent study analysed the effects of leading edge serrations, taking into consideration overall aerodynamic performances and the impact on noise levels. For the tests conducted a wind tunnel was used. This research concluded that leading edge serrations can provide the attenuation of turbulence-interaction broadband noise when they are carefully optimized. Furthermore, there is introduced the concept of curved-serrations. The authors claim that their study can serve as a baseline in future attempts to develop new techniques in the area of noise control [15].

In the current paper, the main purpose is to study the influence of trailing edge blade serrations on a rotor impeller. To this extent NASA's Rotor 37 is used as a baseline, as its geometry is available for such test cases. 3D viscous simulations are carried out on the baseline geometry and three modified configurations, using the commercial CFD software Ansys. Details regarding the mesh and the used scheme are presented in the next section. Regarding the chevron characteristics, they are distributed along one third of the upper side of the blade and several configurations are studied: with two, three and four chevrons, distributed equally on the mentioned region. This approach allows assessing insightful information about how the number of serrations can affect the global performances of the rotor.

CFD Setup

Two batteries of CFD cases are carried out in order to assess chevron influence on aerodynamic behavior of the high performance transonic axial compressor. Four geometries were considered: the baseline plus three altered chevroned trailing edge blades. A minimum of two chevrons were cut from the original blade progressively, offset from the tip of the blade, being placed on one third of the upper side of the trailing edge. The other two geometries altered only the named section, but this time with three and four chevrons.

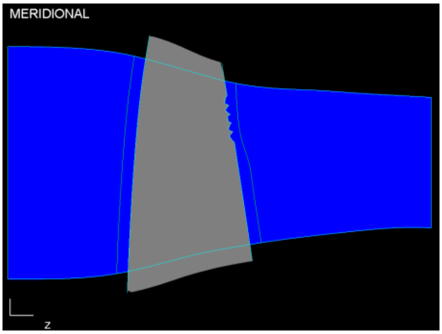
In order to make the cases comparable, the same mesh topology (Fig. 2 - right) was used for each geometry, hence a cutoff trailing edge was opted for. Also, in order to eliminate as much of the secondary flow patterns in that region as possible, the geometry was considered without the tip gap – hence the rotor was treated as a shrouded compressor.

Care was taken so that the first element near the wall is consistent with the requirements for k-omega SST turbulence model used. Also, quality wise, the skewness as well as aspect and growth ratios were evaluated and made to fit within the best practice guidelines. The adopted value for y+ was 1 and the growth ratio was no more than 1.1. Regarding the skewness, a minimum of 20º was set.

The second order upwind scheme was employed in all cases to insure the greatest accuracy of the model. Air was considered an ideal fluid and the inlet conditions were ISA total pressure for all simulations.

The first battery consisted of a greater than optimal outlet constant mass flow of 20.38 kg/s (compared to the 20.188 kg/s value given by NASA's technical report in Tab. 1). The decision was taken to study this regime because it will bring more information to the overall essay regarding a sensitive region of the compressor map. Functionally, the compressor operates with a nominal static back-pressure; hence the second battery of tests was carried out with static outlet pressure of 1.1 atm.

In Fig. 2, the meridional view of the blade with four chevrons is presented on the left. The baseline and the other two configurations (with two and three chevrons) are not illustrated, as the geometry with four chevrons was considered to be the most relevant in serving the purpose of highlighting the area of interest.



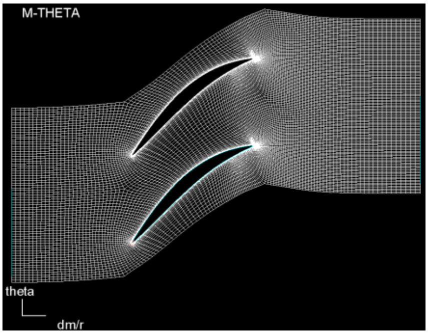


Fig. 2. Meridional view (left) and blade to blade topology of the mesh used (right)

Results

Regarding the 20.38 kg/s batch, for the baseline the isentropic efficiency is 76.54% and the pressure ratio 1.644 compared to the values given in Table 1 by NASA of 87.7% for the isentropic efficiency, 2.106 for the pressure ratio respectively.

After synthesizing the constant mass flow cases, we have observed a clear decrease in isentropic efficiency of the modified stages along with a less dramatic decrease in total pressure ratio. In Fig. 3 one can see that the slight decrease in power consumption of the compressor is in spite of a dramatic decrease in isentropic efficiency and due to a decrease in delivered mechanical work.

An oddity is that the case with 4 chevrons appears to behave better than the rest of the modified geometries. This indicates that optimal placement and span of chevrons is not intuitive and must be tested further.

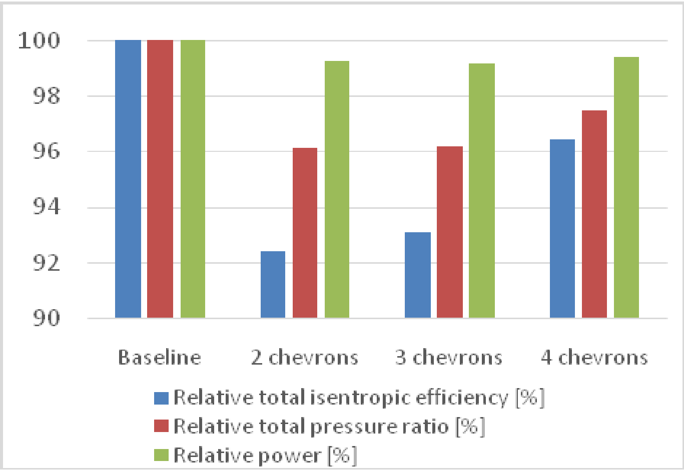
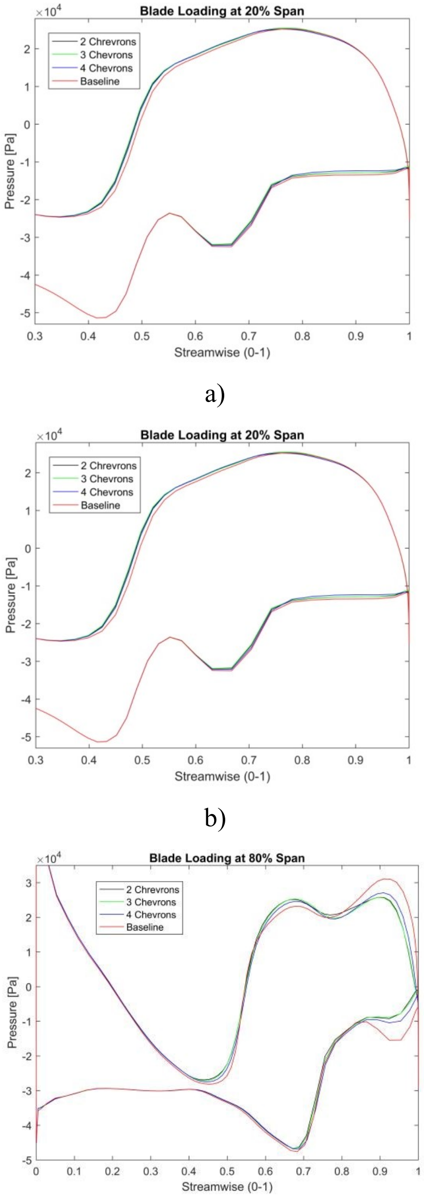


Fig. 3. Comparison between the functional parameters of the 20.38kg/s batch

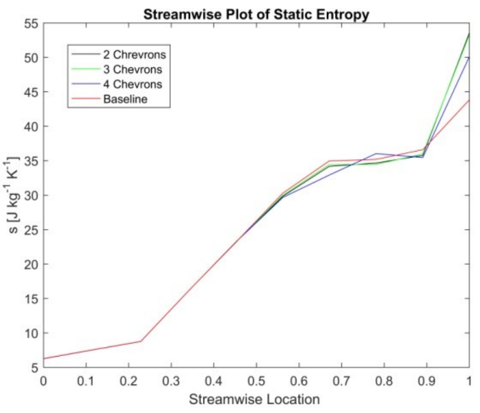


c)

Fig. 4. Blade loading across the span of the blade: a) 20% Span, b) 50% Span, c) 80% Span

In the plots of Fig. 4, one can see the blade loading at 20%, 50% and 80%, for all cases at every span, the loading is slightly lower than the baseline. It is interesting that even near the hub, where the chevrons are not present, the loading is lower.

On the other hand, there is no lowering of the overall efficiency as seen when analyzing the downstream entropy generation (Fig. 5 – left) or the evolution of isentropic efficiency across streamwise locations (Fig. 5 – right). Due to the vortical structures induced by the chevrons, part of the total pressure is converted to heat.



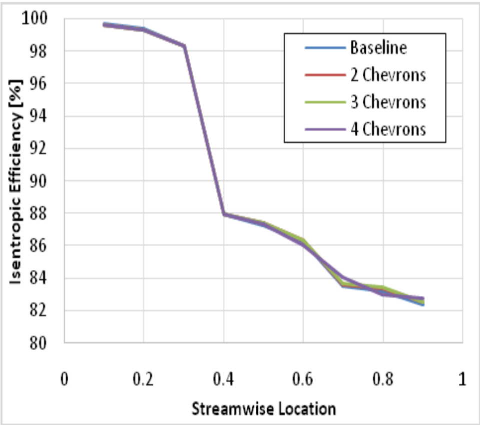


Fig. 5. Entropy across the blade passage (left) and isentropic efficiency evolution (right)

Note that the measurements on the outlet boundary, at streamwise location 1, is not relevant because of boundary effects and should not be considered for analysis.

The second batch of tests refers to the static pressure outlet of 1.1 atm. In this case, when comparing the values of the isentropic efficiency (84.77%) and pressure ratio (1.89 kg/s) for the baseline with the data provided in Tab. 1 by NASA (87.7%, 2.106 respectively), the accuracy is better than in the previous discussed case.

One can see in Fig. 6 that the relative power consumption is no longer linear with the number of chevrons. This is because all parameters vary independently; therefore each case has a different mass flow, total pressure ratio and total efficiency.

Although difficult to interpret under these circumstances, the main conclusion is that the ability to compress the fluid is diminished proportionately to the span of the chevroned portion. This has little or no bearing on the efficiency with which the compression is done.

Furthermore, even though the performances no longer present a linear dependency with the number of chevrons, the best global results are, as in the previous case, given by the configuration with four chevrons.

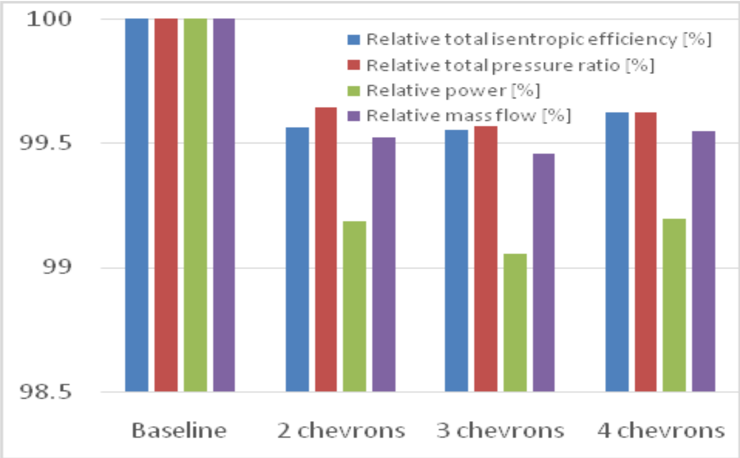
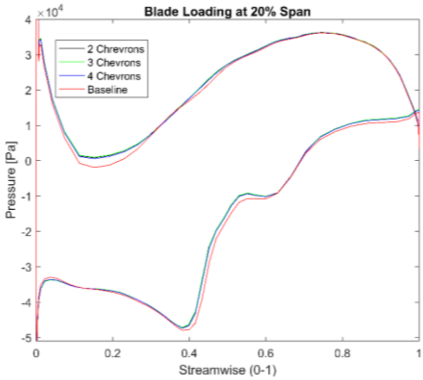


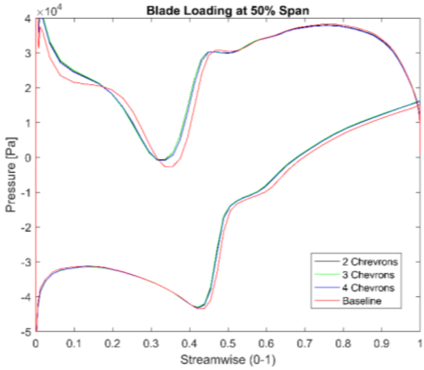
Fig. 6. Comparison between the functional parameters of the constant backpressure batch

In Fig. 7, the pressure distribution is presented, for the 20%, 50% and 80% span locations. For all chevron cases, the pressure loading increases slightly, particularly towards the leading edge on the pressure side. This trend is however compensated by the lower loading of the trailing edge side and the fact that overall, the suction side has a higher pressure (lower suction, per se).

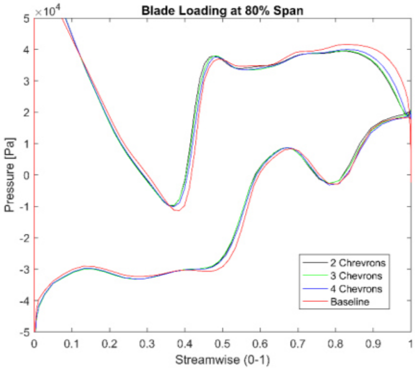
Therefore, placing chevrons on the trailing edge shifts the pressure map to slightly lower pressure ratios. On the contrary, the efficiency map appears to be unaltered, as seen in the constant mass flow batch.



a)



b)



c)

Fig. 7. Blade loading across the span of the blade for the constant backpressure batch at: a) 20% Span, b) 50% Span, c) 80% Span

Conclusions

The current paper explored the use of chevrons on the trailing edge of a highly loaded, high performance axial fan. NASA's Rotor 37 was selected and modified by cutting two, three and four chevrons into the trailing edge portion of the rotor near the tip of the blade. For simplicity, the rotor was considered to be shrouded in order to eliminate tip gap secondary flows that may interfere with the chevron induced secondary flow.

Two batches were considered, one with constant mass flow and one functional, with constant backpressure. To validate the computational methodology the results were compared with those given by NASA's technical report discussed in the introduction, with the data presented in Table 1. The second test case proved itself to be more accurate.

The presence of chevrons influences loading across the entire blade span, even if the effect is more visible closer to the chevrons themselves.

On all cases pressure ratio dropped slightly, along with the power consumption. Counterintuitive, the efficiency was not altered in any of the cases, meaning that the small Reynolds number at which the chevrons function helps keep in check the pressure losses.

In the end, the use of chevrons appears to have a minimal impact on the functionality of the compressor, keeping the efficiency at a constant level while decreasing the pressure ratio by unloading the blades slightly. A more in-depth CFD study, covering the entire speedline would be a logical next step, along with the use of more sophisticated models such as DES or LES.

Also, in this paper the main emphasis is on the influence of chevron number and future work should concentrate on assessing the impact the geometrical characteristics of the chevrons have.

To conclude with, further work may also include a chevron-like trailing edge treatment such as the incorporation of serrated Gurney flaps which has been shown to increase pressure ratios of the rotor [16]. This would compensate the loading drop seen in the current paper. Also, chevron efficiency can be improved with unconventional devices such as sinusoidal chevrons [17] or rhino chevrons [18].

References

1. Biollo, R. and E. Benini. State-of-Art of Transonic Axial Compressors. Advances in Gas Trubine Technology, Italy, 2011. DOI: 10.5772/25257.
2. Dixon, S. L. Fluid mechanics, thermodynamics of turbomachinery. The University of Michigan, Pergamon Press, 1978, 263 p.
3. Samad, A., K. Y. Kim, T. Goel, R. T. Haftka, and W. Shyy. Multiple Surrogate Modeling for Axial Compressor Blade Shape Optimization. American Institute of Aeronautics and Astronautics: Jorunal of Propulsion and Power, 2008, 24, 2, 302–10. DOI: 10.2514/1.28999.
4. Reid, L. and R. D. Moore. Design and Overall Performance of Four Highly-Loaded, High-Speed Inlet Stages for an Advanced, High-Pressure-Ratio Core Compressor. Technical Report, NASA-TP-1337, 1978.
5. Reid, L. and R. D. Moore. Experimental Study of Low Aspect Ratio Compressor Blading. Technical Report, NASA-TM-79280, 1979.
6. Boretti, A. Experimental and Computational Analysis of a Transonic Compressor Rotor. Proc. of the 17th Australasian Fluid Mechanics Conference, 2010, Auckland, New Zealand.
7. Chen, N., H. Zhang, Y. Xu, and W. Huang. Blade Parametrization and Aerodynamic Design Optimization for a 3D Transonic Compressor Rotor. Jorunal of Thermal Science, 2007, 16, 2, 105–14.

8. Ahn, C. S. and K. Y. Kim. Aerodynamic Design Optimization of an Axial Flow Compressor Rotor. Proc. of ASME Turbo Expo 2002: Power for Land, Sea, and Air, 2002, 813–19. DOI:10.1115/GT2002-30445.
9. Dost, M., R. Jensen, and D. Meer. Fighting Noise in Gas Turbines. Noise reduction techniques in gas turbines. Fact sheet, 2016, AviationsFacts.eu.
10. Gliebe, P. R. Serrated Fan Blade. General Electric Company, 2004. United States Patent, US 6, 733, 240 B2.
11. Oro, J. M. F., A. Meana-Fernández, M. G. Vega, B. Pereiras and J. G. Pérez. LES-based simulation of the time-resolved flow for rotor-stator interactions in axial fan stages. International Journal of Numerical Mehtods for Heat & Fluid Flow, 2019, 29, 2, 657–81.
12. Florean, F. G., A. C. Petcu, I. Porumbel and G. Dediu. PIV measurements in Low Noise Optimized Air Jet Pump Demonstrators. International Journal of Energy, 2016, 10, 33–43.
13. Gherman, B. G., I. Mălăel, M. Mihăescu and I. Porumbel. Jet Pump Optimization through Reynolds Averaged Navier-Stokes Simulation Analysis. AIAA 22nd Computational Fluid Dynamics Conference, Dallas, TX, USA, 2015.
14. Jaron, R., A. Moreau, S. Guérin and R. Schnell. Optimization of Trailing-Edge Serrations to Reduce Open-Rotor Tonal Interaction Noise. Journal of Fluids Engineering, 2018, 140, 2. DOI:10.1115/1.4037981.
15. Juknevicius, A. and T. P. Chong. On the leading edge noise and aerodynamics of thin aerofoil subjected to the straight and curved serrations. Journal of Sound and Vibration, 2018, 425, 324–43. DOI:10.1016/j.jsv.2018.02.038.
16. Mălăel, I., H. Dumitrescu and A. Dumitrache. Methods for Improve the Performance of the Turbomachines Using the Flow Control. Proc. of the 9th International Conference of Numerical Analysis and Applied Mathematics, 2011, 1389, 1. DOI:10.1063/1.3637913.
17. Chernem, F. J., J. E. Hammerberg, M. J. Andrews, V. Karkhanis and P. Ramaprabhu. On shock driven jetting of liquid from non-sinusoidal surfaces into a vacuum. Journal of Applied Physics, 2015, 118, 18. DOI:10.1063/ 1.4934645.
18. Drăgan, V. Contributions regarding the development of a highly aerodynamically efficient chevron for use in turbofan noise mitigation – the rhino chevron. Proc. of Caius Iacob 2011 International Conference, Bucharest, 2011.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e14

SOME AEROSPACE APPLICATIONS OF FUNCTIONALLY GRADED MATERIALS

***Adelina Miteva*[1], *Anna Bouzekova-Penkova*[1*]**

[1]*Space Research and Technology Institute – Bulgarian Academy of Sciences*
e-mail: ad.miteva@gmail.com, a_bouzekova@space.bas.bg

Keywords: *Functionally Graded Materials, Properties of Functionally Graded Materials, Composites, Areas of Applications of Functionally Graded Materials, Aerospace Applications of Functionally Graded Materials*

Abstract

Functionally graded materials (FGMs) are currently the subject of great and ever-growing interest from industry and science, and are widely used due to their advantages. These advantages are due to their unique properties and, therefore, their many real and potential applications in various fields of industry, science and everyday life. In this literature review paper, we will briefly focus on some of the properties of FGMs and on some of the existing and expanding future applications of FGM in aerospace and related industries. A critical discussion is presented. Possible future expansion of work in this area is being considered.

Introduction

The present review paper is motivated by a huge interest in the rapidly developing field of materials science, namely, functionally graded (or functionally gradient) materials (FGMs). This interest is due to the ability to produce materials with tailored properties which are suitable candidates for numerous high tech applications such as aerospace, bioengineering and nuclear industries. Over the past two decades, the number of publications in the field of FGMs research is growing exponentially. The information provided has been compiled from the existing literature on FGMs and it may be considered as a brief introduction review to the subject mentioned above. Functionally Graded Material (FGM) is an advanced material that exhibits a gradual change in material properties in at least one spatial direction. FGMs can be designed for specific properties, functions and applications.

[*] Corresponding author e-mail: a_bouzekova@space.bas.bg

Overall properties of FMG are unique and different from any of the individual material that forms it. FGMs are now recognized, by many people, as important composite materials throughout the world. But they are not only a special class of composite materials, although some of them today can be considered as the next step in the development of composite materials. There is a wide range of FGMs and a variety of their applications, and their number in the future is expected to increase. In this paper, an overview of some of the most promising properties and aerospace applications of FGMs is presented.

Definition, origin, classification, methods of production of FGMs

FGMs are an advanced and artificial class of engineered materials characterized by gradual variation in properties depending on the spatial position in the material. Functional change occurs in at least one of their properties. The FGMs area is still under development and the precise definition, classification and properties of this relatively new advanced class of materials are still not generally accepted and are constantly being supplemented. FGMs occur in nature as: bones, teeth, human skin, wood, bamboo etc. The property gradient in the material is caused by a position-dependent chemical composition, microstructure, porosity, density, grain size, refractive index, lattice constant, atomic order, etc. The gradient change in FGMs can be of various types, but it is always continuous and smooth, and it is never abrupt in all directions of the gradient. In FGMs the composition and/or microstructure vary in space, following a predetermined law, in one or in more than one specific space direction.

FGMs are microscopically inhomogeneous materials, such as: composites, single-phase materials, two-phase materials, multiphase materials or nanostructures. Functional properties of FGMs change uniformly at least in one dimension of the particle, film, joint, coating, nanostructure or a bulk sample.

Due to the continuous gradient in their microstructure, FGMs do not possess well distinguished specific interfaces or boundaries between their different regions, as in the case of conventional composite or inhomogeneous materials. Because of this, such materials posses good chances of reducing mechanical and thermal stress concentration in many structural elements, which can be developed for specific applications.

For the first time, in 1970, the general idea for theoretical applications of graded structure composite and polymeric materials was suggested as a concept by Bever et al. [1]. However, these works had only limited impact, probably due to a lack of suitable production methods and technologies for FGMs at that time.

The necessity to bring into practice new materials appears crucial with space vehicles: on the surface side and the skin plates should have very good heat-resistance; on the inside however, – high mechanical qualities (e.g. toughness) were needed. The problem was successfully solved in Japan in the mid of 1980s by manufacturing specific composite: metallic matrix and ceramic particles with graded distribution of these particles. So, the scientific term “functionally graded material” was first introduced in Japan in 1984 for development and implication of thermal barrier materials, being developed for the reusable rocket engine [2–4].

The structural unit of a FGM is referred to as an *element* or a *material ingredient* [2]. FGMs can be composed of various *elements*. *Element* is a conceptual unit for constructing a FGM that includes various aspects of its chemical composition, physical state, and geometrical configuration. The term *element*, probably expresses the overall FGM concept best. *Elements* can resemble biological units such as cells and tissues. As example: bamboo, tooth, and bone all have graded structures consisting of biological *elements*. Examples of typical *elements*, of which FGMs can be composed, are listed below:

- chemical – inorganic, organic, ceramic, metal, polymer, alloys, semiconducting materials, nonmetallic materials;
- physical – electronic state, ionic state, crystalline state, dipole moment, magnetic moment, band gap, potential well, barrier, quantum well;
- geometrical – granule, rod, needle, fiber, platelet, sheet, pore, texture, orientation;
- biological – complex macromolecule, organelle, cell, tissue.

Depending on the geometry and cross-sectional area of the produced material, FGMs can be divided into two main groups; thin and bulk FGMs [5,6]. Thin FGMs are usually in the form of relatively thin sections or thin surface coatings and have thin cross sections. Bulk FGMs are those with thickness greater than 1 mm and whose functional properties vary with respect to the gradient profile of the material.

A separate group of very thin FGMs are low-dimensional nanostructured FGMs (or functionally graded low dimensional nanostructures) [7–9]. Examples are semiconductor graded gap quantum wells (QWs) [10] and semiconductor graded gap superlattices, which electronic structure may tailor for specific applications. Most often this happens in their structure by graded compositional and doping changes in the growth direction. These properties enable one to use these nanostructured FGMs for high-performance room temperature optoelectronic devices. Moreover, to improve the performance of these optical devices, band structure modifications have also been investigated. The modification of the well potential shape (functionally graded QW region) can create different optical properties and thus optimize nanostructure-based devices compared to conventional rectangular QWs. In Fig. 1 is presented a schematic band diagram of: (a) conventional rectangular QW (without grading); (b) linear analog graded-gap QW; (c) parabolic graded-gap QW [10]. The $Al_xGa_{1-x}As$ alloy composition x in homogenious rectangular QW is constant, equals zero, i.e. the well is made of pure GaAs. The alloy composition $Al_xGa_{1-x}As$ in the linear well region varies linearly from x = 0.3 to x = 0.0 (from the left to the right side of the QW; see Fig. 1b). The aluminium concentration x for $Al_xGa_{1-x}As$ alloy in parabolic quantum well varies parabolically (Fig. 1c). In these three QWs cases we have pure AlAs at the barriers and the arrow above shows the growth and grading directions.

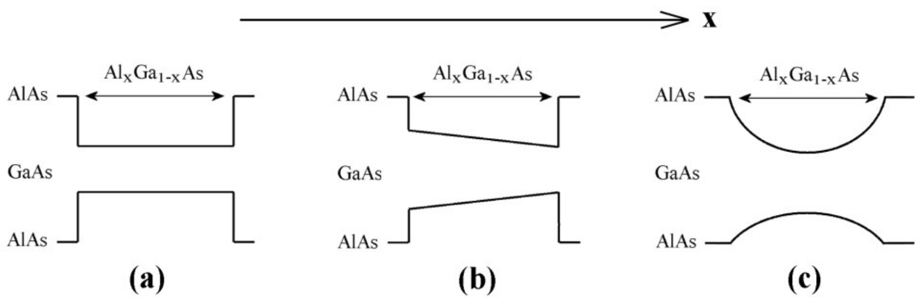


Fig. 1. Schematic QW band diagram: (a) rectangular QW - without grading, i.e. constant composition x; (b) linear graded-gap QW; (c) parabolic QW

Depending on the number of directions in which the properties have changed, we can distinguish one-dimensional, two-dimensional or three-dimensional FGMs.

Different types of functionally graded composites, depending upon the nature of gradient, are the following: fraction gradient; shape gradient; orientation gradient, size gradient type and dispersed phase gradient. [6, 11].

Some schemes of FGM structures are shown in Fig. 2. In Fig. 2a for comparison are shown two nongraded homogenous materials with abrupt interface. In one of the simplest FGMs, two different material elements change gradually from one to the other side as illustrated in Fig. 2b. In Fig. 2c two different material elements continuously change gradually from the center to the end (outer side). The elements can also change in a stepwise gradation illustrated in Fig. 2d. Fig. 2e shows the functionally graded joint/transition between two materials. Fig. 2f shows the continuously functionally graded outer surface or coating. The x arrow in Fig. 2 shows the grading direction.

Physical properties of FGMs can be described by material function f(x). This function shows the change/gradient of a particular property (for example such as composition) along one space direction – x, always following a predetermined law. In homogenous materials this function is constant (see Fig. 1a). In FGM (see Fig. 1b,c and Fig. 2-b,c,d,e,f). The only requirement for material function f(x) is that it be continuous or quasi-continuous, and smooth.

In Fig. 3 are presented material functions f(x) in the x direction of the corresponding material structures, shown in Fig. 2.

In the case of continuous graded structure, the change in composition and microstructure occurs continuously depending on the position. On the other hand, in the case of stepwise grading (Fig. 2d and Fig. 3d), composition and microstructure feature changes in stepwise manner, giving rise to a multilayered structure with interface existing between discrete layers. In some materials, if the width of the step is small enough (two or several atomic layers in a given direction, as with semiconductor hetrostructures and graded gap quantum wells, and superlattices), there are no obvious interfaces between layers which is achieved by current epitaxial methods of crystal growth.

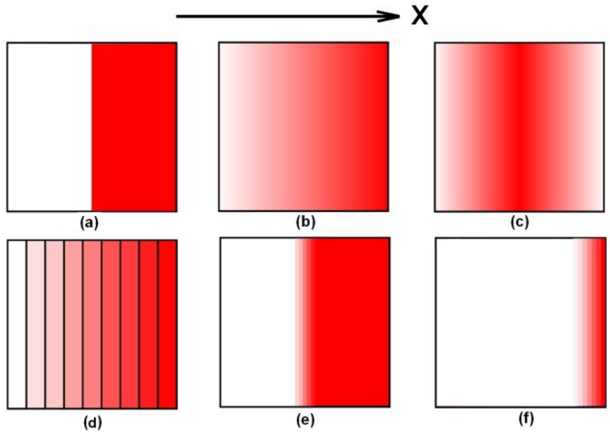


Fig. 2.: Type of FGM structure: (a) – structure without grading; graded structures - (b), (c), (e), (f) – continuous - linear; (d) - stepwise graded structure. Arrow x - is the grading direction.

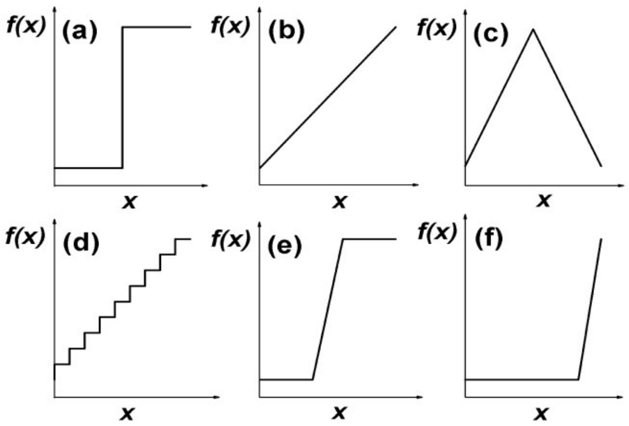


Fig. 3.: Material function f(x) in x direction (a) - structure without grading; (b), (c), (e), (f) – continuous- linear graded; (d) - stepwise graded structure.

A gradual distribution of pores, from the interior to the surface, is also a FGM. Even if the gradation of the *elements* is limited to a specific location in the material such as the interface, a joint (Fig. 2e), or a surface as shown in Fig. 2f, the material can be considered to be an FGM because it incorporates the FGM concept.

The fabrication process is one of the most important fields in FGM research. There are many different methods for producing FGMs [2, 5, 12–14]. Most of the processes for FGM production are based on a variation of conventional processing methods which are already well established. These methods are physical and chemical in form. There is a lot of literature available on the subject of their production. The choice of FGM manufacturing technology is mainly influenced by: the desired application, the FGM group of material to be produced, the material combination, geometry of the desired component, available manufacturing equipment for FGMs.

Major thin FGMs production techniques are various deposition processes such as physical or chemical vapor deposition, atomic layer deposition, spray deposition, electrodeposition, laser deposition, plasma spraying, self-propagating high temperature synthesis, ion beam assisted deposition, electrophoretic deposition, surface deposition method and coating processes. The choice of deposition is dependent on the performace required of the material.

Processing of bulk FGMs are usually energy intensive and slow, and cannot be produced using deposition techniques, such as vapor deposition [4, 15]. The process of manufacturing bulk FGM is generally grouped into two: the gradation process and the consolidation process. The gradation process comprises the constitutive, homogenizing and the segregation processes. The bulk FGMs consolidation process follows the gradation process. This process involves the sintering and solidification of the powder material. Processing conditions for the material are chosen such that their gradient structure is not altered while unequal shrinkage is also mitigated. Bulk manufacturing processes include powder metallurgy, metal casting, solid freeform fabrication techniques [5, 7], sequential casting, infiltration process, frictional stir casting, centrifugal casting method, melting processes.

A special group of very thin FGMs – functionally graded low dimensional nanostructures (such as compositionally graded gap QWs and superlattices) [7] are produced mainly by two epitaxial methods: molecular beam epitaxy and metalorganic vapor phase epitaxy. These modern crystal growth techniques make it possible to control the purity, size and directions at the atomic level.

Application areas of FGMs

The concept of FGMs, applicable to almost all material fields, is described as a systematic process of bringing incompatible functions such as thermal, wear and corrosion resistance, toughness and machinability into a single part [2, 15].

This has expanded the application of FGMs in many sectors. There is a big amount of real and potential applications of FGMs. The main, but not the only area of usage of FGMs, are materials and devices operating in extreme conditions (high temperature gradients, mechanical loads, etc.). In aerospace and nuclear energy applications, reliability rather than cost is the key issue. But in applications such as cutting tools, high temperature rollers, and engine components, which require wear, heat, mechanical shock, and corrosion resistance; the key issues are the cost/performance ratio and reliability.

Some examples of various applications of FGMs in different sectors are given below:

- aerospace (rocket nozzle, heat exchange panels, solar panels, turbine wheels, space plane nose, combustion chamber protective layer, body components, rocket engine components, reflectors, camera housing, caps and leading edge of missiles and space shuttle. etc);
- automobile (combustion chambers, diesel engine pistons, racing car brakes, crown of piston, cylinder liners, exhaust valves and valve seating);
- medicine (medical implants, teeth and bone replacement, artificial skin, drug delivery system);
- construction (roads – e.g. pavement life can be increased by 59% as a result of the grading of the asphalt concrete and base layers together);
- commodities (car body, sport goods, window glass, etc.);
- energy (energy conversion devices, nuclear energy, thermionic and thermoelectric converters, fuel cells, solar batteries, fuel pellet, etc.);
- electronics (graded band semiconductor, substrate, sensors);
- optoelectronic (antireflective layers, fibres, GRIN lenses);
- nanotechnology (nanostructured devices and systems etc);
- chemical plants (heat exchanger, heat pipe, slurry pump, reaction vessel);
- optics (optical fibre, lens);
- nuclear (first wall of fusion reaction, fuel pellets);
- energy conversion (solar cell, fuel cell, thermoelectric generator, thermionic convertor);
- engineering (cutting tools, machine parts, engine components, turbine blade, roller, shaft, etc.) etc.

Some aerospace applications of FGM structures

A special class of FGMs is metal porous materials, which appears as gas reinforced metal matrix composites with graded control of pore shape and orientation [3]. Like all porous metals, the ordered porosity metals have a wide range of applications including filters, catalysis, silencers, flame arresters, heat exchangers, fuel cells, electrolytic cells, and fluid substance separators, ionic rocket engine parts, self-lubricating bearings, thermal screens and vibration dampers. The most attractive applications are as components of rocket combustion chambers and oil high pressure filters. Gasar stainless steel is a promising material for medical applications, especially for artificial bones or joints with good corrosion resistance.

Thin shells, as very common structural elements, occupy a leadership position in civil, mechanical, architectural, aeronautical and marine engineering, since they give rise to optimum conditions for dynamic behavior, strength and stability. In other words, these structures support applied external forces efficiently by virtue of their geometrical shape. An important aspect in the successful applications of these structures is fact that shells cover large pans. As for many other shape kinds, conical, cylindrical shells and shells of complex geometry are very common structural elements. In [16] a layered epitrochoidal shell element has been used for the modeling and analysis of functionally graded composite shell structures. Obtained results have shown the strengthening role of functionally graded materials for shell structure. Static analysis of the functionally graded epitrochoidal shell under self-weight and thermal loading have been done and validated with the published results.

The problems of modeling and vibrations of FGM thin walled rotating blades that could be used in helicopters and turbomachinery applications and the associated subject of spinning circular cylindrical beams were considered [2]. In these studies, the blade was modeled as a pretwisted thin-walled shear-deformable beam of a nonuniform cross section with material properties varying in the thickness direction according to a power law. The thermal field was assumed steady state, and the rotation velocity was constant. The effects of material grading and blade taper ratios on the natural frequencies were elucidated. Besides free vibrations, the divergence and flutter instability of the blades, accounting for gyroscopic forces, were analyzed where both the natural frequencies and the spinning speed corresponding to the blade instability were significantly increased by appropriate variations in grading.

Thermal Barrier Coatings (TBCs) are typically used in applications where it becomes necessary to protect metallic or composite components in military and commercial aeroengines, aircraft engines, for gas turbine engines for automobiles, helicopters, marine vehicles, and electric power generators, combustion chambers from excessive temperature. The problems that have to be addressed to design TBC include processing technology, heat transfer during and immediately after the processing, microstructure formation, residual thermal stresses, micromechanics of TBC, and thermomechanical response during the lifetime [2, 4]. FGM TBCs are attractive due to the potential for a reduction in thermal stresses, avoiding delamination and spallation tendencies, prevention of oxidation. As for FGM TBC, the heat transfer problem was simplified due to the fact that it is typically one dimensional, occurring in the thickness direction. For example, if a FGM coating consists of a number of layers with constant volume fractions of individual components, the total thermal resistance can be evaluated as the sum of individual layer resistances. According to the authors [2, 4] if the coating will be sprayed as a multilayered system with a compositional gradient varying from pure metal adjacent to the substrate to a ceramic exposed layer, then obvious interfaces between the layers of the coating could be eliminated. Thermal fatigue may occur in TBCs subject to periodic temperature variations during their lifetime. As was shown, using a FGM coating may result in a five-time increase in the resistance to thermal fatigue compared to a conventional counterpart. Additionally, the oxidation resistance was improved as a result of the grading, as reported in [2].

Turbine blades are one of the most highly stressed rotating parts in gas turbines. In order to increase the efficiency and performance of turboengines, gas inlet temperatures in the high pressure turbines must be increased, and component cooling must be decreased. Here, ceramic TBCs with a low thermal conductivity applied on turbine components play a key role. Ceramic TBCs are connected to components by thin metallic bond coats, which also protect the components from hot corrosion and oxidation. Conventional bond coats are single layers of MCrAIY (where M = Ni or NiCo) or Pt-AI based materials. The interface between the bond coat and the ceramic top coat is the most critical region with respect to the lifetime of the both TBCs and bond coat (the MCrAIY layers). Increasing their lifetimes was achieved by grading the composition (with optimal concentration distribution) across the coating thickness ([12, 17] in [2]). Graded TBC systems are potentially advantageous compared with TBC systems that have ungraded layers. The use of FGMs to join high temperature materials is being actively investigated.

Space vehicles flying at hypersonic speeds experience extremely high temperatures from aerodynamic heating due to friction between the vehicle surface and the atmosphere. They are of two types: vehicles, that are launched vertically into space by a rocket propulsion system (like the U.S. space shuttle and the capsules used for the Apollo missions); and fully reusable spacecrafts, that are based on a horizontal takeoff either from a ground-based runway or from a horizontally flying carrier (as the U.S. National Aerospace Plane) [2, 17].

In the first type, after sufficient acceleration the rocket system completely separates from the space vehicle. During reentry at velocities greater than 11 km/s, rapid heating of the leading edge, where the heat protection shield is located, occurs at altitudes between 120 and 50 km, and maximum temperatures (the radiant equilibrium temperature) above 2 500 °C develop. Because the relatively flat heat protection shield is exposed to the extreme heat of reentry for just a few minutes and is used only one time, it can be fabricated from ablative materials. The reentry velocity of the U.S. space shuttle at an altitude of 120 km is below 8 km/s, and the maximum temperature experienced is about 1 500 °C for a few minutes. Structural components that experience the maximum exposure to heat such as the nose cone, the leading edges, the rudder, and the flapperons need to be made of non-metallic carbon/carbon composites (C/C) with adequate oxidation protection coatings.

Horizontally launched space planes that are accelerated by air-breathing engines (e.g., jet engines) fly in the atmosphere at hypersonic speeds for a longer time than vehicles launched vertically by rockets. Therefore, the space plane experiences its maximum exposure to heat during its launch into space. Initially, one of the main objectives of investigating FGMs deposited by chemical vapor deposition (CVD) was to develop thermal barrier coatings for a space plane. In a comparison test, models of the components of a nose cone (hemispherical C/C composites 50 mm in diameter) were coated with an ungraded 100 μm thick protective layer of SiC. Similar C/C composite models were coated first with a graded SiC/C FGM by penetrative CVD followed by deposition of the 100 μm thick SiC protective layer. All the coated nose cone models were subjected for 1 minute at 1 900 °C to a supersonic gas flow containing an amount of oxygen approximately equal to a standard atmosphere. The nose cones with the SiC/C FGM intermediate layer showed no discernible change in structure even after 10 cycles. In contrast, those without the intermediate SiC/C layer between the C/C substrate and the ungraded SiC coating deteriorated after the first cycle [2]. Sheets of SiC/C FGMs produced by CVD provide excellent thermal stability and thermal insulation at 1 227 °C, as well as excellent thermal fatigue properties and resistance to thermal shock [2].

Most rocket and scramjet engines use TBC materials that have been previously developed for turbine engine applications. The heat flux in the path of the hot gases is much greater in rocket engines than in turbine engines. Here the TBCs are exposed to a hostile environment, that is higher temperatures and more severe thermal transients, but for shorter mission cycles. Hence, the TBCs are mainly deposited as thin structures (< 0.2 mm thick) to reduce the probability of coating failure.

For example, protective CVD-SiC/C FGMs produced for rocket combustors have undergone critical tests. The maximum outer wall temperature of these model combustors was 1376 °C to 1 527 °C, while the inner wall temperature reached 1 677 °C to 2 027 °C. No damage to the combustors was observed after two test cycles [2].

In large combustion chambers, TBCs are not commonly used, as the heat cannot be dissipated quickly enough to avoid local hot spots and coating damage [2, 6]. Here longlife polymeric graded TBCs are potentially applicable in preventing coating failures.

In large liquid propellant rocket engines, TBCs are mainly used in the high pressure hydrogen and oxidizer turbopumps ([6] in [3]). TBCs have been used as liners for the spark igniters and pre burners, for turbo housing liners, for turbine blade shanks (located between the blade platform and root), and for vane shrouds. Experimental coatings have been used on the turbine blade platforms and vane airfoils. In addition, graded TBCs have potential applications in the upper part of the main combustion chamber as coatings on the interpropellant plate, spark igniter, and injector baffle tips.

Graded TBCs have been considered also for other rocket engines such as small regeneratively cooled thrust chambers in orbital maneuvering systems. These zirconia/nickel ($Zr0_2Ni$) FGM chambers are prepared by a combination of galvanoforming and plasma spraying. The graded layer is first deposited (up to 25% $Zr0_2$ on a Ni metal chamber) by galvanoforming and subsequently coated to 100% $Zr0_2$ by plasma spraying. No delamination of $Zr0_2$ was observed after 550 s of combustion. In order to assure the reliability of the $Zr0_2$/Ni FGM, it was necessary to engineer the microstructure to form strong layers as well as to further optimize the graded structures, and also to control the reaction with a propellant. As noted above, graded TBCs are potentially applicable for engine and airframe structures in reusable hypersonic vehicles ([7–9] in [2]).

Stealthiness is now a required specification for modern weapons. Parts made of specific materials can be used in stealth missiles to absorb the emitted electromagnetic energy to minimize waves reflected in the direction of the enemy radar receiver. In some applications, e.g. high velocity missiles, the materials can be subjected to high thermomechanical stress. For these applications, the most promising new materials are ceramic matrix composites reinforced with ceramic woven fabrics. The use of long, continuous ceramic fibers embedded in a refractory ceramic matrix creates a composite material with much greater toughness than monolithic ceramics, which have an intrinsic inability to tolerate mechanical damage without brittle rupture.

The conducting properties of these ceramic composites depend on the fibers, the matrix, the interfaces, and other parameters such as the topology of the arrangement of the various phases. Nicalon® SiC fibers, which have semiconducting properties, and Nextel® mullite ($3Al_20_3$_$2Si0_2$_0.1 B_20_3) fibers, which are completely dielectric, are used in the preparation of oxide matrix ceramic composites. Nasicon matrix composites reinforced with long semiconducting and/or dielectric fibers can have mechanical and electrical properties, ranging from dc to microwave frequencies [4]. The Nasicon solid solution, structural formula $Na_{1+x}Zr_2Si_xP_{3-x}0_{12}$ ($0 \leq x \leq 3$), which has a graded electrical conductivity that varies by four orders of magnitude as a function of x, is a useful system for investigating the preparation and properties of graded ceramic matrix composites with tailored microwave properties.

FGMs are used also in aerospace equipment, such as solar panels and solar cells. Gallium arsenide (GaAs) is the current market leader for solar cells deployed for extra-terrestrial applications, but it is prohibitively expensive for terrestrial applications. To reduce the cost of the GaAs solar cell, in [18] was carried out an optoelectronic optimization for an $Al_xGa_{1-x}As$ (AlGaAs) solar cell containing an n-AlGaAs absorber layer with a graded bandgap. The bandgap of the absorber layer was varied either sinusoidally or linearly. Sinusoidal grading of the bandgap was predicted to enhance the solar cell maximum effiency to 34:5%. An effciency of 33:1% was predicted with linearly graded bandgap. Thus, grading the bandgap of the absorber layer can help realize ultrathin and high-efficient AlGaAs solar cells that will be cheaper and already suitable for both extra-terrestrial and terrestrial applications.

Extensive understandings of FGMs mechanical properties permit one to design and manufacture graded composite structures with deserved functionality. So accurately modelling of FGMs mechanical properties is greatly needed. Modelling their mechanical properties is central for engineers and scientists to accurately predict their mechanical behaviors under different and/or extreme conditions and obtaining new spectacular functionalities for various aerospace applications [6–7, 19].

Future work

Due to the broad and rapidly developing field of FGMs, all their properties and aerospace applications cannot be encompassed in this paper. Nevertheless, here are some of the observations of the authors based on the published research and their own analysis of the subject.

We, in our Space Materials Science department at SRTI-BAS, are working on obtaining new FGMs based on semiconductor graded gap QWs in the $Al_xGa_{1-x}As$ system [10].

At SRTI-BAS, we are also working on new FGMs, based on reinforced Al metal matrix composites and special Al alloy materials, with high hardness, high density and high thermal stability, for space applications. SiC particles and nanodiamond particles were used as reinforcing additives.

Such materials will be of considerable interest for their application in space technology and in particular for aerospace instrumentation [10, 20]. The work is in progress in this direction. Moreover, in these fields we have some research experience in obtaining and investigating:

- Quantum confined Stark effect in compositionally graded gap $Al_xGa_{1-x}As$ QWs [10];
- an Al alloy – B95 with certain additions of tungsten and nanodiamonds (This Al alloy was a part of the international outer space experiment "Obstanovka", carried out in the Russian sector of the International Space Station, and was exposed to outer space for 2 and a half years.) [20].

We will use theoretical methods of bandgap engineering (tight binding and ab initio methods) and the powder metallurgy methods.

Conclusion

Today FGMs are the new top trend in all aerospace materials research and development.

This paper provides a brief overview of FGMs with a focus on their aerospace applications.

FGMs represent a rapidly developing area of science and engineering with numerous practical applications. Due to functional variation of FGM-materials, their physical/chemical properties (e.g. stability, hardness, conductivity, reactivity, optical sensitivity, melting point, etc.) can be manipulated to improve the overall properties of conventional materials. Although so far the practical applications of FGMs have not yet been fully implemented and used. The research needs in this area are uniquely numerous and diverse, but FGMs promise significant potential benefits that fully justify the necessary effort.

There are still more to be done in terms of research to improve the performance of manufacturing processes of FGMs in order to make them more cost effective.

FGMs are perspective materials for modern optoelectronic devices, which are irreplaceable and vital part of aerospace equipment and structures.

The emergence of FGMs has revolutionized the aerospace and aerocraft industry. FGMs used initially as thermal barrier materials for aerospace structural applications and fusion reactors are now developed for general use as structural components in high temperature environments.

The list of applications is endless, and it will increase with improving of the processing technology, cost of production, and properties of FGMs. The development of FGMs needs an integration of multidisciplinary domains to work together in designing, manufacturing, and exploring more areas of applications.

Last but not least, both experimental and theoretical studies of the FGMs are very important and need to be developed in order to look for unknown and possible properties of FGMs for new potential applications. Future applications require materials with exceptional mechanical, electronic, and thermal properties that can withstand various environmental conditions and are readily available at reasonable prices.

References

1. Bever, M. B. and P. E. Duwez, On gradient composites, ARPA Materials Summer Conference, 1979, 117–40.
2. Miyamoto, Y., W. A. Kaysser, B. H. Rabin, A. Kawasaki and R. G. Ford (Eds.), Functionally graded materials: design, processing and applications, Vol. 5, Springer Science & Business Media, 2013, 339.
3. Sobczak, J. and L. B. Drenchev, Metal Based Functionally Graded Materials, Bentham Science Publishers, 2010, 79.

4. Mahamood, R. M.and E. T. Akinlabi, Functionally Graded Materials. Topics in Mining, Metallurgy and Materials Engineering, Springer, Cham, 2017, 103.
5. Owoputi, A.O., F. L. Inambao and W. S. Ebhota, A Review of Functionally Graded Materials: Fabrication Processes and Applications, International Journal of Applied Engineering Research, 2018, 13, 16141–51.
6. Bouzekova-Penkova, A. and A. Miteva, Aluminium-Based Functionally Graded Materials, Proceedings of the fourth national conference with international participation "Material science, hydro- and aerodynamics and national security'2014", 23-24 October 2014, Sofia, IMETHC-BAS, 2014, 152–56.
7. Bharti, I., N. Gupta and K. M. Gupta, Novel Applications of Functionally Graded Nano, Optoelectronic and Thermoelectric Materials, Int. J. of Materials, Mechanics and Manufacturing, 2013, 1, 221–24.
8. Ou, Q., X. Bao, Y. Zhang et. al., Band structure engineering in metal halide perovskite nanostructures for optoelectronic applications, Nano Materials Science, 2019, 1, 268–87.
9. Daikh, A .A., A. Drai, I. Bensaid, M. S. A. Houari and A. Tounsi, On vibration of functionally graded sandwich nanoplates in the thermal environment, Journal of Sandwich Structures & Materials, 2020, 1099636220909790.
10. Miteva, A. M., Digital versus analog graded-gap quantum well in the presence of applied constant electric field, Machines. Technologies. Materials, 2017, 11, 236–39.
11. El-Wazery, M.S. and A. R. El-Desouky, A review on functionally graded ceramic-metal materials, Journal of Materials and Environmental Science, 2015, 6, 1369–76.
12. Naebe, M. and K. Shirvanimoghaddam, Functionally graded materials: A review of fabrication and properties, Applied Materials Today, 2016, 5, 223–45.
13. Tripathy, A., S. K. Sarangi and A. Chaubey, A review of solid state processes in manufacture of functionally graded materials, Int J Eng Technol, 2018, 7, 1–5.
14. Mahamood, R. M., E. T. Akinlabi, M. Shukla and S. Pityana, Functionally Graded Material, An Overview, Proceedings of the World Congress on Engineering, III, London, U.K., 2012.
15. Miteva A., An overview of the functionally graded materials, Machines. Technologies. Materials, 2014; 8, 13–16.
16. Mathieu, G. O. and I. T. Farhan, Using FGM for cyclic shell structures, Structural Mechanics of Engineering Constructions and Buildings, 2016, 4, 14–20.
17. Mardirossian, G. Fundamentals of remote aerospace technology, (Book in Bulgarian) NBU, Sofia, 2015, ISBN: 978-954-535-882-1, 240.
18. Ahmad, F., A. Lakhtakia and P. B. Monk, Optoelectronic optimization of graded-bandgap thin-film AlGaAs solar cells, Applied Optics, 2020, 59, 1018–27.
19. Xu, X.J. and J. M. Meng, A model for functionally graded materials, Composites Part B: Engineering, 2018, 145, 70–80.
20. Bouzekova-Penkova, A., K. Grigorov, M. Datcheva and C. A. Cunha, Influence of the outer space on nanohardness properties of Al-based alloy, Comptes rendus de l'Acad´emie bulgare des Sciences, 2016, 69, 1351–54.

НЯКОИ АЕРОКОСМИЧЕСКИ ПРИЛОЖЕНИЯ НА ФУНКЦИОНАЛНО ГРАДИЕНТНИТЕ МАТЕРИАЛИ

А. Митева, А. Бузекова-Пенкова

Резюме

Функционално градиентните материали (ФГМ) понастоящем са обект на голям и непрекъснато нарастващ интерес от индустрията и науката и са широко използвани поради своите предимства. Тези предимства се дължат на техните уникални свойства и следователно многото им действителни и потенциални приложения в различни области на индустрията, науката и ежедневието. В този литературен обзор ще се спрем накратко върху някои от свойствата на ФГМ и върху някои от съществуващите и разширяващи се бъдещи приложения на ФГМ в аерокосмическата и свързаните с нея индустрии. Представена е критична дискусия. Разглеждат се възможните бъдещи разширения на работата в тази област.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e15

HUMAN RELIABILITY INCREASING APPROACH IN AIRCRAFT MAINTENANCE TECHNICIAN'S TRAINING PROCESS

Angel Tanev

Sensata Technologies Bulgaria Ltd.
e-mail:ang_tanev@abv.bg

Keywords: *human reliability, aircraft maintenance, optimal control, utility function.*

Abstract

This article presents an approach for increasing the maintenance technician's reliability by considering the human fault rate as utility function's maximization problem in the technician's training process. Adequately trained technicians are capable to perform a maintenance and manage the reliability of their assigned assets within the complex aircraft systems. In general, a degradation of aircraft reliability, due to maintenance tecnician's competency, typically leads to significant, undesirable safety and economic consequences. In this article, an optimal control theory is applied on the purpose of finding of a fault rate reduction series in the training process which leads to highest technician's reliability in the maintenance process of the complex aircraft systems.

Introduction

Aircrafts equipment and systems are becoming more complex, as well as the associated cost is significantly increasing due to loss of operation in case of failure. In aviation domain, four aspects are considered – reliability, maintainability, availability and safety (RAMS). From theoretical background point of view and practical observations, the reliability level is decreasing by accumulating the time in the field. Therefore, it is very important to keep an aircraft reliability above a critical (lower) level to not compromising the aviation safety. Aviation safety is a domain which is being threated by many agents [1]. One should note that, on one hand, errors related to maintenance can be more difficult to detect and, on other hand, they have the potential to affect the safe operation of aircraft for some time period. Technical/maintenance failure emerged as the leading cause of airline accidents and fatalities [2]. Improperly trained maintenance technicians is one of the contributing factors to aircraft accidents [4]. Some authors describe about US National Transport Safety Board (NTSB) reports related to deficient maintenance of 50% – 7 out of 14 airline accidents [2]. Another analysis on the accidents in the period 1990–2006, done by EASA, shows the major cause was maintenance [5]. International Air Transport Association (IATA) safety report stated on average about 10% were maintenance events which led to aircraft accidents (2009–2013) [6]. For the same time period, maintenance operations together with training systems were highlighted to be a latent cause for 27 out of 338 non-fatal accidents [6]. Another study shows that maintenance factors (6%) take the third rank [7]. IATA (2003–2008) stated that incorrectly performed maintenance was a primary cause for 30% (on average) of the registered worldwide accidents with aircrafts [8]. Another study from Boeing shows that 20% of the accidents contained maintenance or inspection action [3]. In United Kingdom, Civil Aviation Authority (CAA) has reported that 10% of recorded events are maintenance related. For 10 years period (1996–2006): 51.1% were assigned to incorrect maintenance actions, 26.2% to ineffective maintenance control and 20.7% to incomplete maintenance [9]. Some other studies focused on fatality of maintenance related accidents. For the time period between 1999 and 2008, 26.7% of all fatal accidents were maintenance related [4].

Maintenance errors do not only cause safety issues but also have significant economic impact and they are very costly to the industry. Maintenance errors may cause, for example, aircraft unavailability, in-flight shut downs, maintenance rework, maintenance equipment damage and injury to maintenance personnel. Some estimations show a cost of USD 500 000 (per engine in-flight shutdown) [1].

The growing demand for maintenance personnel will require highly qualified technicians. For example, the need for Europe region will be approximately 130 000 new technicians who should be available to maintain the new aircrafts during the next 20 years [1].

One can conclude that the human's reliability is a very important part of aircraft/aerospace systems reliability and safety, for example, see [10–17]. However, many studies in that area do not consider utility function and dynamic optimization in their modelling. The main goal of the proposed study is to find an approach how to increase the technician's reliability by considering the fault rate as utility function's maximization problem in the technician's training process.

Theoretical Background

Suppose a maintenance technician's full working capability restoration after a dedicated technical training where the probability of fault-free operating for a given time t is defined by the following expression [19]:

$$R_0(t,\tau) = e^{-\int_0^t \lambda(z,\varepsilon)dz . e^{-\int_0^\tau \upsilon(l,\xi)dl.}} \quad (1)$$

By analogy to the reliability theory, the term under the first exponent is associated with the technician's reliability [19]:

(2) $$R(t,\varepsilon;\tau,\xi)=\int_0^t \lambda(z,\varepsilon)dz.e^{-\int_0^\tau \upsilon(l,\xi)dl.}$$

In (2) the first term $\int_0^t \lambda(z,\varepsilon)dz = q(t,\varepsilon)$ is called an exhausted reliability for a time t under conditions ε [19]. The second term $\gamma(\tau,\xi)=\int_0^\tau \upsilon(l,\xi)dl$ is called a restored reliability which is obtained in the technician's training process for a time τ under conditions ξ. The technician's reliability can be then expressed in the following simplified way [19]:

(3) $$R_0(t,\tau)=e^{-q.e^{-\gamma}}=R_0(q,\gamma)$$

After reviewing (3), one can make a conclusion that the technician's reliability $R_0(q,\gamma)$ is decreasing when the exhausted reliability q is higher, and $R_0(q,\gamma)$ is increasing when the restored reliability γ is higher.

Let's consider in our case study the following problem – the restored reliability term $\gamma(\tau,\xi)=\int_0^\tau \upsilon(l,\xi)dl$ where the technician's fault rate υ to be considered as utility function which needs to be maximized on the purpose of obtaining as high as possible technician's reliability in the training process. An optimal control theory (dynamic optimization) will be further applied to solve this problem.

Nowadays, the applied mathematical modeling (e.g. applied optimization) is widely used in many research areas – for example, see [18, 20].

First, suppose an optimal control task defined over the following frame with periods: 0, 1, 2, ...T [18]. The general consideration is that the state variable x_t is measured at the beginning of each period t and the control variable u_t is applied during this period t. Fig. 1 shows this problem statement:

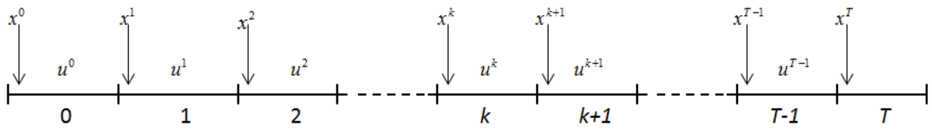


Fig. 1. Discrete time optimal control problem

with some functions which are continuously differentiable:

$$f: E^n \times E^m \times \Theta \to E^n, F: E^n \times E^m \times \Theta \to E^1, g: E^m \times \Theta \to E^s, S: E^m \times \Theta \cup \{T\} \to E^1$$

The modeling of the problem stated above can be done in the following way [20]:

$$\max \left\{ J = \sum_{t=0}^{T} F(x_t, u_t, t) + S(x_T, T) \right\} \tag{4}$$

where: the above expression (4) is a subject to the following constraints:

$$\Delta x_t = x_{t+1} - x_t = f(x_t, u_t, t), \ t = 0,1,...,T-1 \tag{5}$$

x_0 – given

$g(u_t, t) \geq b_t, \ t = 0,1,...,T-1$

However, in the real practice, due to some feasibility reasons, it is impossible to have records in continuous time. Therefore, having a daily information, the time series of fault rate control can be expressed as [20]:

$$\vec{u} = \{ u_0, u_1, u_2, ..., u_T \} \tag{6}$$

Next, supposing the total time period in our case study T=10 days (i.e. the scheduled techncian's training period is 10 days), the objective function in our optimization problem can be expressed by [20]:

$$V(x_0, \vec{u}) = \sum_{t=0} p \ C(u_t) \tag{7}$$

where: $C(u_t)$ – cash flow for *t-th* day;

$\beta = 1/(1+r)$ – discounting factor;

r – the interest rate.

From a financial theoretical background point of view, (7) can be considered as net present value (NPV) and in many practical tasks can be modeled with a power function [20]:

$$V(x_0, \vec{u}) = \sum_{t=0} p \ u_t^{\alpha} \tag{8}$$

To maximize the objective function (8), the qualified trainer/instructor can decide to apply the time series of fault rate control in (6) with a constraint shown on equation (9) [20]:

$$x_{t+1} - x_t = -u_t \qquad (9)$$

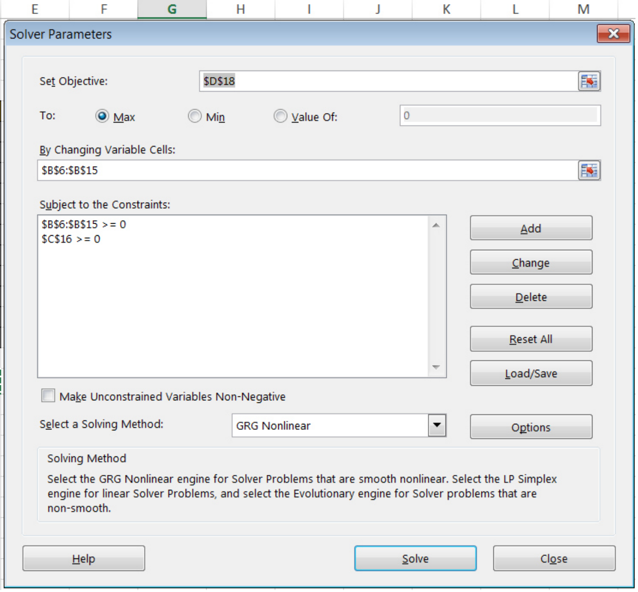
To proceed further with the task, a numerical solution to a problem related to dynamic optimization requires two endpoint conditions. Suppose the technician's initial fault rate is 2.5 [1/day] and the trainer/instructor has planned to deliver the training for time period of 10 days. Then in this case study are assumed: the initial condition $x_0 = 2.5$ [1/day] and final one $x_T = x_{10} \geq 0$.

Results and discussion

The task described above is solved by using Microsoft Excel® software where expressions (1) – (9) are introduced in a spreadsheet format. In Table 1 are shown the inputs together with the equations listed in the theoretical chapter:

Table 1. Inputs for our case study

	A	B	C	D
1	**Alpha**	0.8		
2	**r**	0.08		
3	**Beta**	0.926		
4				
5	**time(day)**	**u(t)**	**x(t)**	**DCF(t)**
6	0	0.25	2.50	0.330
7	1	0.25	2.25	0.305
8	2	0.25	2.00	0.283
9	3	0.25	1.75	0.262
10	4	0.25	1.50	0.242
11	5	0.25	1.25	0.225
12	6	0.25	1.00	0.208
13	7	0.25	0.75	0.192
14	8	0.25	0.50	0.178
15	9	0.25	0.25	0.165
16	10		0	
17				
18			**PV**	**2.39**



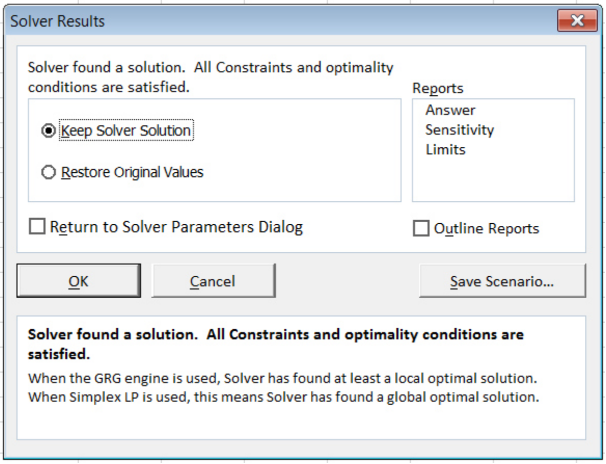
As a starting point, the inputs in our study: α (see (8), note: $0<\alpha<1$) and the interest rate r are assumed to be 0.8 and 8% respectively. The cell B3 in Table 1 is dedicated on computing the discounting factor in (7). The equation (9) is implemented in column C which represents the time series of the technician's fault rate. At the beginning of each day, the fault rate is equal to the fault rate of the beginning of previous day minus the reduced fault rate in previous day (see (9)). The initial fault rate is x_0 = 2.5 [1/day] which is shown in cell C6. Applying “Solver” algorithm introduced by Microsoft Excel® which requires initial guess values of the optimal time series of fault rate. And at 1st iteration, suppose that the fault rate is controlled uniformly each day, for example, with 0.25 [1/day] (cells B6:B15 in Table 1). In the column D in spreadsheet is shown the DCF(t) which stands for discounted cash flow gained during day t, i.e. $\beta^t u_t^\alpha$. The cell D18 shows the net present value denoted by (8).

The NPV given by (8) depends on the initial fault rate x_0 and on the fault rate control $\vec{u}$, i.e. $V(x_0,\vec{u})$=2.39 if x_0=2.5 and uniform fault rate series of 0.25 [1/day]. One may conclude in that case the uniform fault rate series is not optimal when the future cash flows are subject to discounting factor. Applying the optimization algorithm in that case, the optimal fault rate series $\vec{u}$ can be found which maximizes the NPV in (8). The non-negative constraints on control and state variable can be seen in the “Solver’s dialog box (see Table 1).

The results are now shown in Table 2. Then the optimal fault rate reduction series is: 0.82 [1/day] in day t = 0; 0.56 [1/day] in day t = 1; 0.38 [1/day] in day t = 2; etc. In that case of the optimal fault rate series is applied, and then the global objective function (cell D18) increases from 2.39 up to 2.60. It is interesting to highlight that the optimal fault rate series is with a slope that is declining due to the fact that the discounting factor accelerates the technician’s fault rate reduction.

Table 2. Optimal solution for fault rate reduction

	A	B	C	D
1	**Alpha**	0.8		
2	**r**	0.08		
3	**Beta**	0.926		
4				
5	**time(day)**	**u(t)**	**x(t)**	**DCF(t)**
6	0	0.82	2.50	0.850
7	1	0.56	1.68	0.579
8	2	0.38	1.13	0.394
9	3	0.26	0.75	0.268
10	4	0.18	0.49	0.182
11	5	0.12	0.32	0.124
12	6	0.08	0.20	0.084
13	7	0.06	0.12	0.057
14	8	0.04	0.06	0.039
15	9	0.03	0.03	0.027
16	10		0	
17				
18			**PV**	**2.60**

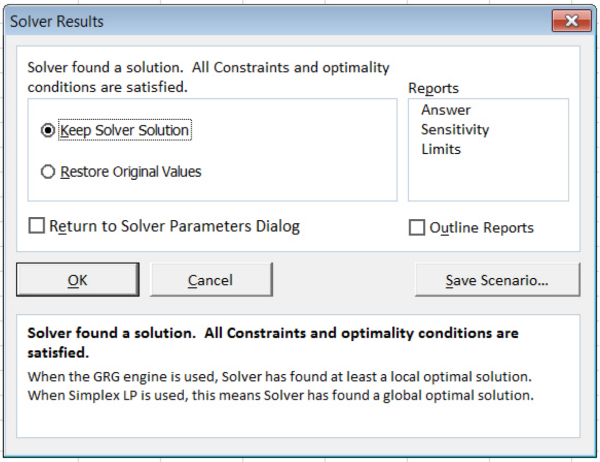


Let's now perform some sensitivity analysis, i.e. to highlight the influence of the interest rate over the optimal solution- going back to Table 1 to change the interest rate (cell B2), and then again running "Solver" algorithm. Table 3 shows the new optimal solution.

By choosing an increased interest rate- this accelerates faster fault rate reduction: 0.96 [1/day] for day $t = 0$; 0.59 [1/day] in day $t = 1$; then smaller reduction with respect to the modified interest rate (0.1). The new optimal fault rate series is with higher slope since the future cash flows are discounted with higher value. The NPV increases up to 2.52 with the new optimal series. In this case, the higher discounting value is impacting the DCF value which is 2.52, even with optimal values.

Table 3. Case with modified interest rate

	A	B	C	D
1	**Alpha**	0.8		
2	**r**	0.1		
3	**Beta**	0.909		
4				
5	**time(day)**	**u(t)**	**x(t)**	**DCF(t)**
6	0	0.96	2.50	0.964
7	1	0.59	1.54	0.599
8	2	0.37	0.95	0.372
9	3	0.23	0.58	0.231
10	4	0.14	0.35	0.143
11	5	0.09	0.21	0.089
12	6	0.05	0.12	0.055
13	7	0.03	0.07	0.034
14	8	0.02	0.03	0.021
15	9	0.01	0.01	0.013
16	10		0	
17				
18			**PV**	**2.52**

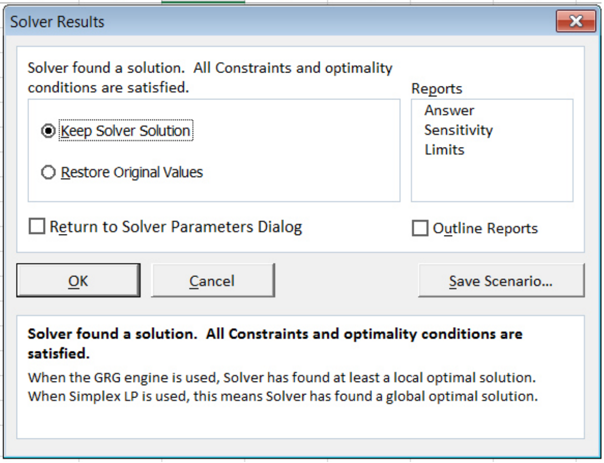


We need to highlight that significant factors impacting the optimal fault rate control are the discounting factor which induces that the fault rate has to be reduced faster. However, this effect is suppressed by the decreasing returns of the daily cash flow.

Reviewing now Table 4 which is showing the optimal fault rate series for the case when the input α is modified: from 0.8 to 0.85 and the "Solver" algorithm is started again- here the new optimal fault rate series is even steeper: 1 [1/day] for day t = 0; 0.6 [1/day] for day t = 1; 0.36 [1/day] for day t = 2; etc. It should be noted that approximately 80% from the (initial) fault rate is expected to be reduced over the first 4 days of the technician's training.

Table 4. Case with modified "alpha" parameter

	A	B	C	D
1	**Alpha**	0.85		
2	**r**	0.08		
3	**Beta**	0.926		
4				
5	**time(day)**	**u(t)**	**x(t)**	**DCF(t)**
6	0	1.01	2.50	1.008
7	1	0.60	1.49	0.603
8	2	0.36	0.89	0.361
9	3	0.22	0.52	0.216
10	4	0.13	0.31	0.129
11	5	0.08	0.18	0.078
12	6	0.05	0.10	0.046
13	7	0.03	0.05	0.028
14	8	0.02	0.03	0.017
15	9	0.01	0.01	0.010
16	10		0	
17				
18			**PV**	**2.50**



Let's consider a numerical example of a restored reliability computation-see second term $\gamma(\tau,\xi)=\int_0^{\tau}\upsilon(l,\xi)dl$ of (3). Consider the results summarized in Tables 1 and 2: uniform series of fault rate reduction vs optimal fault rate reduction series-the fault rate vs training day is plotted in Fig. 2:

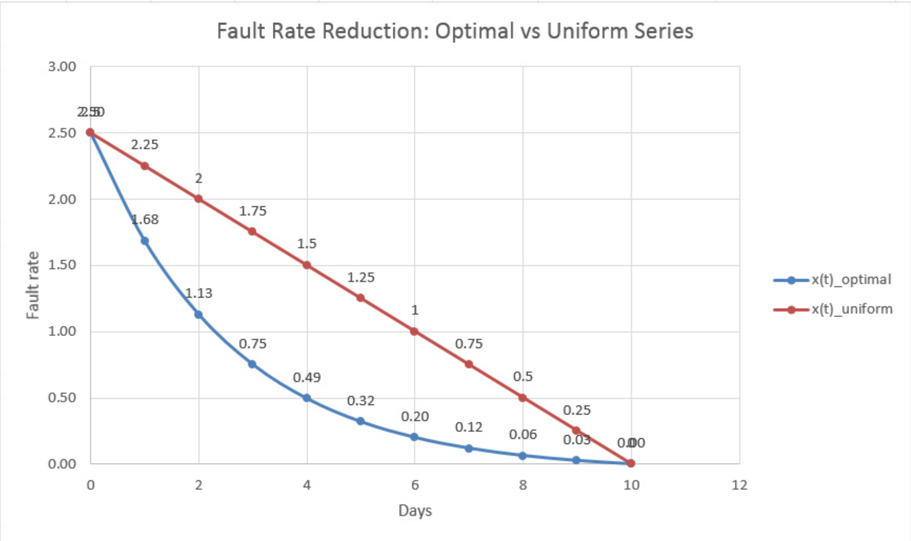


Fig. 2. Fault rate reduction vs training days

The restored reliability (exp(-υt), where: υ-fault rate[1/day] at day *t*) of uniform vs optimal fault rate series is plotted in Fig. 3. One can conclude that in case of optimal fault rate reduction series the restored reliability has a steeper slope (i.e. increasing faster) than the case with uniform fault rate reduction series:

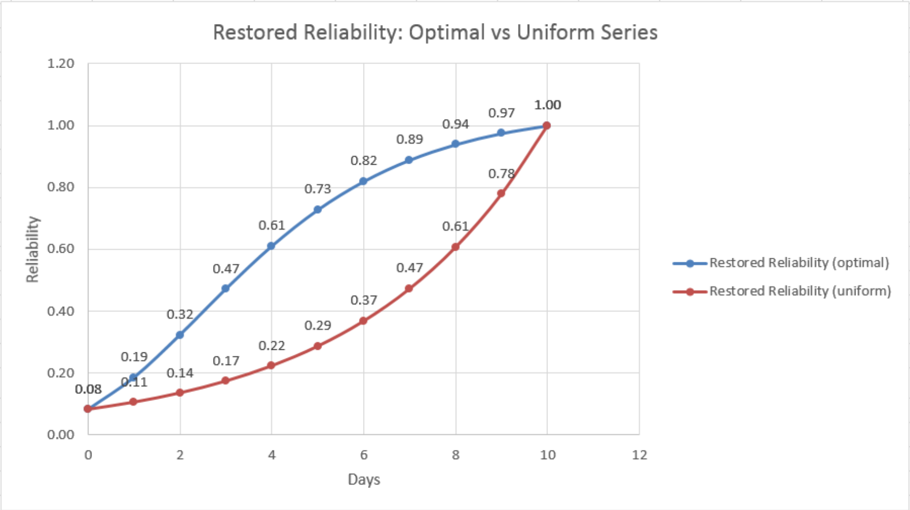


Fig. 3. Restored reliability vs training days

Conclusion

In this section, I would like to summarize some of the most important outcomes obtained by the proposed research study which are:

Maximized technician's utility function (8) requires an optimal fault rate control series $\vec{u}$ during the considered technician's training period.

The increasing in the technician's restored reliability shows steeper slope for the case of optimal fault rate reduction series compared to uniform fault rate reduction series (Fig. 3). For example, 63% of restored reliability in case of an optimal fault rate reduction series is obtained at about t = 4 days while in case of an uniform fault rate reduction series this value is obtained at about t = 8 days.

The influence of the interest rate over the optimal solution has been analyzed via performing a sensitivity study. Faster fault rate reduction can be accelerated by increasing the interest rate (Table 4): approximately 80% from the (initial) fault rate is expected to be reduced during the first 4 days of the technician's training.

The proposed study suggests an overlapping between today's very important and modern subjects like financial modeling, applied optimization and human reliability.

References

1. Dalkilic, S., Improving aircraft safety and reliability by aircraft maintenance technician training, Journal Engineering Failure Analysis, 2017, 82, 687–94, Elsevier.
2. Hobbs, A., An Overview of Human Factors in Aviation Maintenance. Australian Transport Safety Bureau, (December, 2008) (ISBN 978-1-921490-93-4).
3. Safety Regulation Group, CAA, CAP 716 Aviation Maintenance Human Factors (EASA/JAR145 Approved Organisations).http://www.caa.co.uk/docs/33/C AP716.pdf, (2003) (accessed 09.10.15).
4. Marais, K. B., M. R. Robichaud, Analysis of trends in aviation maintenance risk: an empirical approach, Reliab. Eng. Syst. Saf. 2012, 106, 104–18.
5. EASA, Guidance on Safety Hazard Identification,http://www.easa.europa.eu/essi/docu ments/ECASTSMSWG-GuidanceonHazardIdentification.pdf, (2009)
6. Safety Report 2013, International Air Transport Association, Montreal, 2013 (ISBN 978-92-9252-349-7).
7. Federal Aviation Administration, Introduction to Maintenance Error Analysis. http://www.faa.gov/about/initiatives/maintenance_hf/library/documents/medi a/hfa cs/1_intro-duction.pdf, (2014).
8. IATA Safety Report, 2013. http://www.iata.org/publications/Documents/iata-safety-report-2013.pdf,.
9. Safety Regulation Group, CAA, CAP 2009/05, Aircraft Maintenance Incident Analysis. http://www.caa.co.uk/docs/33/2009_05.pdf, (2009).
10. Erjavac, A. J., R. Iammartino, and J. M. Fossaceca, "Evaluation of preconditions affecting symptomatic human error in general aviation and air carrier aviation accidents," Reliability Engineering & System Safety, vol. 178, pp. 156–163, 2018.
11. Calhoun, J., C. Savoie, M. Randolph-Gips, and I. Bozkurt, "Human reliability analysis in spaceflight applications," Quality and Reliability Engineering International, vol. 29, no. 6, pp. 869–882, 2013.
12. Getsov, P., Yordanov D., (2018). "Modeling the emergency situation with a failure of the engine and asymmetric thrust on a transport aircraft with two engines", International scientific conference of "Aviation Faculty", D. Mitropoliya, 19–20 April 2018, 97–102. (In Bulgarian)
13. Getsov, P., Mardirosyan G., (2018), "Aspects and application of space technology", Proceedings of scientific conference of NMU "Vasil Levski, V. Tarnovo, 14–15 June 2018, 42–53. (In Bulgarian)
14. Getsov, P., (2018). "Bulgarian contribution to space research", International scientific conference of "Aviation Faculty", D. Mitropoliya, 19–20 April 2018, 17–25. (In Bulgarian)
15. Getsov, P., Hubenova Z., "Reliability study of operators within a complex ergatic system", Journal scientific and applied research, 2014, 5, 7–14.
16. Getzov, P., Z. Hubenova, D. Yordanov, W. Popov. Modeling of the human –operator in a complex system functioning under extreme conditions, Journal Aerospace Research in Bulgaria. 2013, 25, 206–27, Sofia.
17. Gindev, E. *Introduction of Reliability Engineering*. Sofia: Marin Drinov, 2001.
18. Suresh, P. Sethi, Gerald L. Thompson. *Optimal control theory: applications to management science and economics*. Berlin: Springer, 2006.

19. Smagin, V. A., Modely nadezhnosti zhivogo organizma (operatora), http://sir35.narod.ru/Cmagin/MODEL_04032.htm#Beg (in Bulgarian)
20. Petrov, N., Tanev A. Optimal Control of Natural Resources in Mining Industry, International Journal of Mining Science and Technology, 2015, 25, 2, 193–98, Elsevier.

ПОДХОД ЗА ПОВИШАВАНЕ НА НАДЕЖДНОСТТА НА ЧОВЕКА ПРИ ТРЕНИРОВЪЧНИЯ ПРОЦЕС НА ТЕХНИЦИТЕ ЕКСПЛОАТИРАЩИ АВИАЦИОННА ТЕХНИКА

Ангел Танев

Резюме

Статията представя подход за повишаване на надеждността на техника по експлоатация на авиационна техника, като разглежда интензивността на грешките на човека като проблем на оптимизация на функция на полезност в процеса на обучение на техника. Адекватно обучените техници са в състояние да извършват поддръжка и да управляват надеждността на възложените им активи. Като цяло, влошаването на надеждността на авиационната техника поради некомпетентността на техника за техническо обслужване обикновено води до значителни, нежелани последици свързани с безопасността и икономически загуби. В тази статия е приложена оптимална теория за управление с цел намиране на серия на редуциране на грешките в процеса на обучение, което води до най-висока надеждност на техника при експлоатация на сложните авиационни системи.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e16

USAGE OF EYE TRACKER TECHNOLOGY IN EXAMINING ATTENTION DISTRIBUTION OF OPERATORS OF UNMANNED AERIAL VEHICLES

***Zoya Hubenova*[1], *Konstantin Metodiev*[1], *Svetla Dimitrova*[1], *Liubomir Alexiev*[2]**

[1]*Space Research and Technology Institute – Bulgarian Academy of Sciences*
[2]*Military Medical Academy, Sofia, Bulgaria*
e-mail: komet@space.bas.bg

Key words: *Flight Simulator, Eye Tracker, Computer Vision, Unmanned Aerial Vehicle*

Abstract

This article proposes yet another approach towards looking into causes for attention distribution of an operator of unmanned aerial vehicle. During examination, the operator is being tested at dedicated flight simulator while data are gathered and visualized through a mobile eye tracker. Two work stages are considered sequentially, i.e. building a geometric 2D transformation of region of interest (homography) within an image, and overlaying a dynamic heatmap as well. In the former stage, spontaneous movements of the operator's head, recorded by the video, are eliminated thus enabling the operator to use the mobile eye tracker instead of a desktop-based one. During the latter stage, the distribution of operator's attention over time is displayed.

In order to implement the current research, a source code has been developed in C++ for some features readily available in OpenCV library to be used. In addition, data gathered after carrying out flight session are processed and discussed thoroughly.

Introduction

Currently, an increasing number of countries around the world are making use of Unmanned Aerial Vehicles (UAV) for the reason that these vehicles are able to keep expenditures low and maintain multi role functionalities simultaneously. Possible applications of UAV, beyond the military one, are quite diverse. These include monitoring of oil pipelines and dams, protection of forests, monitoring of hazardous sites, etc. To control the UAV, it takes the operator a wide range of hands-on experience largely identical to the conventional pilot's skills. From that point of view, UAV operators' training done by means of flight simulators seems to be a pivotal stage in developing even better airplane control habits.

The growing number of UAVs, starting from the smallest and getting to the largest, make both public and relevant authorities in a number of countries pay increasing attention to training a new type of pilots. These pilots, in particular, neither are subject to high load factor nor exposed to danger.

The operator makes all subsequent decisions based on analysis of the situation. Therefore, it becomes clear that this process is the most critical one both in terms of reliability of making optimal decisions (within the system of detected events, objects, processes) and requirements to be met in order to provide for real-time functionality. The human factor, for example in aviation, is still the main cause of most accidents and fatalities; hence, the problem of objective evaluation of operator training progress is quite acute.

Human role in ensuring proper operation of unmanned aerial systems is essential. It involves organization of process planning, maintenance and usage of the control system, evaluating UAV effectiveness and achieved results. What is more, the “Operator – UAV – Computer” interface within the control system architecture is considered fundamental. Choice and certification of such an interface as well as organization of the relevant operator training is further bounded up with developing special programs and centers.

The accompanying software developed for the study needs is a replicated solution to the so-called “Heatmap Explorer”, developed in Python and thoroughly described in [1]. Although the source code is freely available to everybody for downloading (Bitbucket), it is incompatible with new releases of Pupil software, and deprived from cross-platform functionality. The proposed alternative code was developed in C++ starting from the very scratch, borrowing the same idea, and simplifying the implementation to some extent. In addition, some functionality available in OpenCV library, [2], was applied in order to facilitate the development process further.

Method

Training UAV operators by means of simulators appears to be one of the most important factors in ensuring safe operation of the UAV. Training process itself will minimize negative impact of the human factor by diminishing possibility of erroneous actions made by the UAV crew. In general, there are two main concepts in preparing and training UAV operators: training at dedicated simulators (specifically designed for this purpose), and providing real ground control station with a training mode.

Research is carried out in laboratory for selection, training and control of unmanned aerial vehicle operators at Space Research Institute, Bulgarian Academy of Sciences. The laboratory is equipped with simulator C-Star developed by Israeli firm SimLat, [3]. The simulator offers a virtual environment that largely resembles the real one. In addition, the simulator is able to register and store dynamics of interaction between crew members. It is also possible for the instructor to train each crew member (either pilot or payload operator) separately. In this mode, the instructor replaces the absent crew member and does her/his duties. For example, in case of training the payload operator, it is the instructor who controls the airplane, thus completing the relevant pilot task. On the other hand, whenever pilot is being trained, the instructor handles the payload. The user interface is intuitive, user friendly. Every trainee who is somewhat skillful in handling conventional computer peripherals will succeed in mastering the simulator too, Fig. 1.

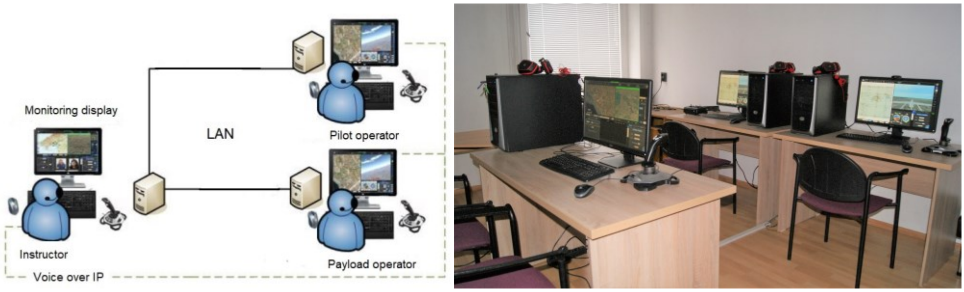


Fig. 1. The C-Star simulator layout (left) and our equipped laboratory

The simulator basic configuration is composed of working stations for instructor and trainees (pilot and payload operator), one station each. The pilot is trained for controlling and monitoring the UAV which in turn may fly in either of modes: manual or automatic. The trainee monitors not only the airplane and onboard systems but also the flight task. The payload operator is trained in handling the payload (EO / IR camera), identifying and tracking targets in initially designed scenario. The simulator has an intercom for communication between the instructor and the trainees.

In order to examine distribution of the operator's visual attention, we make use of the eye-tracking technology so as to record and evaluate eye movement. The main task of the eye tracking research is recording and transmitting data in real time regarding oculomotor activity. Eye-tracking (oculography) is a research method employing a device to track the trainee’s gaze point coordinates or eye movements during task performance. Currently, the most popular eye trackers use a video-based method for non-invasive and non-contact measurement of the eye movements to be made possible.

Eye movement tracking indicators can be applied to various aspects of oculomotor behavior, depending on type of analysis. The most common metrics is based on fixations and / or saccades. The fixations are presented as discrete portions of near-stable points the eye stares at. The saccades are defined as eye movements between fixations. In order to record the aforementioned eye movement parameters, the research team benefited from a mobile eye tracker.

The mobile eye tracker used in the current study is Pupil Labs Core, [4], Fig. 2.

Fig. 2. Pupil Labs Core binocular headset

The head set consists of a diode Smartled® SFH4050 IR, emitting light at 850 nm wavelength, an environmental camera, and two removable lenses with fixed focal length and field of view 60 and 100 deg. It was a binocular version of the eye tracker that research team used in the present study, Fig. 2. The binocular implementation is said to produce more robust data in terms of both calibration and measurement. What is more, to carry out the calibration stage is considered highly desirable, not obligatory. The accompanying software provides a readily available calibration values by default. It is also possible for the user to collect data by means of an independent mobile device, such as a tablet or a smartphone, in order to make the process of data gathering smooth. The obtained video is processed afterwards for the accompanying software to be capable of discovering the calibration markers (if any) and mapping the gaze point offline.

Broadly speaking, a heatmap is computed as a Kernel Density Approximation meant for estimating the probability density function of gaze points. There are many examples of symmetrical kernels, such as Epanechnikov, Gaussian, triangle, rectangle, etc. A unit density of Gauss kernel is estimated according to following formula, [5]:

(1) $$K(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-t^2/2\right)$$

A one-dimensional Gauss kernel is shown in Fig. 3.

The bivariate estimate, in its most widespread form, depends upon two variables. It is presented as a sum of kernel products in direction of both axes:

(2) $$\hat{f}(x,y)=\frac{1}{Nh_x h_y}\sum_{i=1}^{N}K\left(\frac{x_i-x}{h_x}\right)K\left(\frac{y_i-y}{h_y}\right)$$

In eq. (2) $\{x_i, y_i\}$ are sample (gaze) points, $\{x, y\}$ are pixel coordinates, h_x and h_y are smoothing coefficients, and N is the sample size.

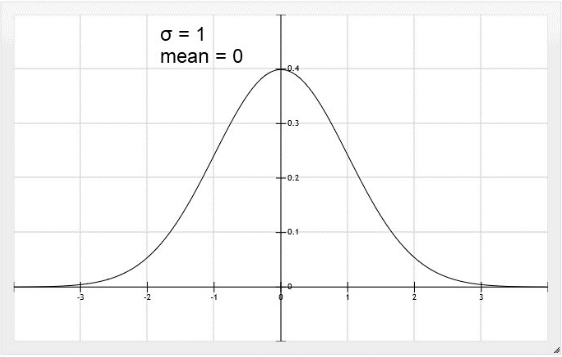


Fig. 3. Gauss kernel, mean = 0, standard deviation = 1

For each image pixel, the algorithm calculates exponentially declining gaze intensity relevant to a center of distribution (point of gaze) with coordinates $\{x_i, y_i\}$ according to formula (2). A pixel is not to be processed if it lies outside $\pm 3\sigma$ interval. At this point, the algorithm assumes zero gaze intensity to speed up the computational process. After "scanning" all pixels, the algorithm superimposes the heatmap on the image so that the kernels are visible “from above.” Finally, the complete heatmap is remapped within the interval 0 ... 1.

The so-called Short Rainbow color map has been used to represent gaze intensity at given frame, Fig. 4.

Fig. 4. Short rainbow color map

The heatmap color depends upon its normalized value. Values closer to one are represented by red color while values closer to zero are depicted in blue. In addition, two intermediate colors are visible in Fig. 4, i.e. cyan RGB(0, 255, 255) and yellow RGB(255, 255, 0). It is possible to compute any intermediate color by applying a linear interpolation to selected channels (Red, Green, Blue) separately. In order to work out the transitional color, in terms of two end colors and a weighing variable, it is sufficient to make use of following simple formula:

(3) $$Lerp(start, end, weight) = (end - start) * weight + start$$
$$0 \leq weight \leq 1 \quad 0 \leq start \leq 255 \quad 0 \leq end \leq 255$$

During color computation, four transitional zones might be figured out with regard to the heatmap value. Depending on zone, some channels are held constant whilst others undergo interpolation. Prior to interpolation, the heatmap value should be remapped within 0 … 1 interval so as to get usable quantity of the weighing coefficient in formula (3). Transitional zones and color channel definitions are shown in Table 1.

Table 1

Transitional Zone	Heatmap Value, f	Color channels		
		Red	Green	Blue
Blue to Cyan	[0; 0.25)	0	Lerp(0, 255, f)	255
Cyan to Green	[0.25; 0.5)	0	255	Lerp(0, 255, 1–f)
Green to Yellow	[0.5; 0.75)	Lerp(0, 255, f)	255	0
Yellow to Red	[0.75, 1]	255	Lerp(0, 255, 1–f)	0

The algorithm in Table 1 is implemented by means of if … else if … else branching. In this way, a smooth transition between two arbitrary end colors is achieved.

The transparency level is commonly set within 0 ... 255 range and subsequently applied to each color channel. For example, given pixel RGB channels of (180, 195, 192) and alpha channel of 40% transparency. In order to overlay a red pixel (255, 0, 0), derived from the heatmap, the old pixel color and the red one have to be “merged.” The new color might be computed according to expression: 0.4*(180, 195, 192) + 0.6*(255, 0, 0) = (225, 78, 77). In this way, the resulting color channels do not exceed amount of 255 each.

In Fig. 5, an example of dynamic heatmap is visible, so are the normalized Gaussian kernels, right half. Several normalized kernels are shown at a time. The red zone depicts the highest value of gaze intensity whilst the blue one denotes the lowest. The heatmap maximum value is always equal to 1 (read peak). Eventually, the picture to the left is overlaid with the heatmap layer set to a certain transparency level beforehand.

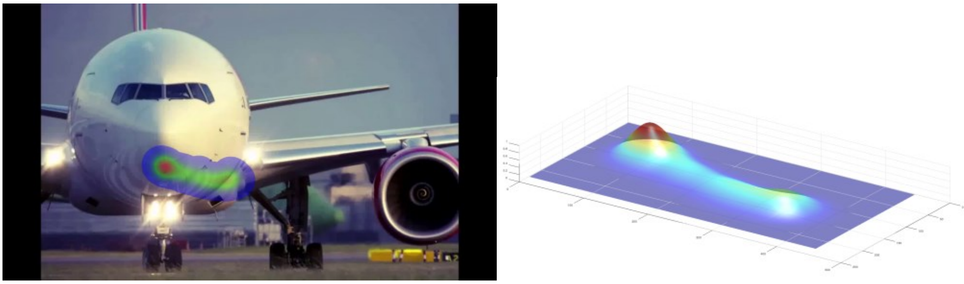


Fig. 5. A heatmap consisting of several Gaussians: top view (left) and perspective

During the video recording stage, the eye tracker user is not expected to be standing still which is the reason why objects appear with varying coordinates on the environmental video. The region of interest is not exception to the rule; hence, the bounding rectangle is to be measured by the algorithm frame by frame. One way of doing it is making use of four fiducial markers, either Aruco, [6], [7], or AprilTag, placed at the region corners. OpenCV is capable of detecting the markers and computing their coordinates on the screen. Four markers define a quadratic region of interest for it to undergo a perspective alignment later on (cv::warpPerspective method). Whenever a key is pressed, the region of interest is stretched so that it can fill up the viewing surface, i.e. a screen. In this way, part of the frame outside the region of interest is removed. The operator's spontaneous head shake is compensated. Exemplary Aruco markers used in the current study are shown in Fig. 6, id 105, 106, 107, 108, dictionary id 16.

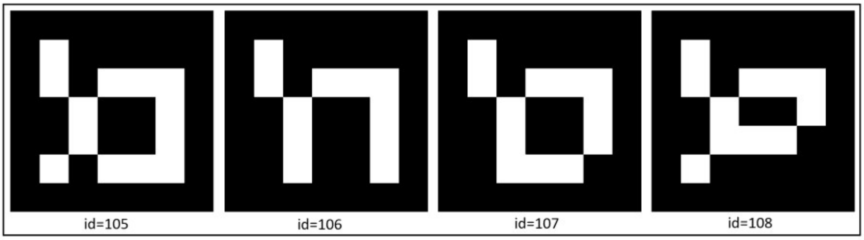


Fig. 6. Aruco markers used in the presented study

In addition, exemplary AprilTag markers are shown in Fig. 7, id 0, 14, 22, 17, dictionary id 36h11

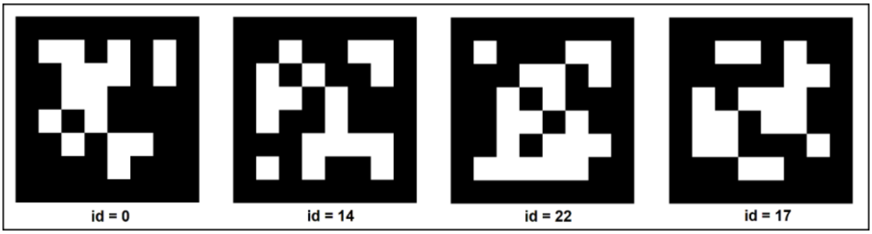


Fig. 7. AprilTag markers used in the presented study

Due to short focal length of lens, the video frame appears distorted with barrel distortion dominating. In order to make up for it, a compensation must be carried out (cv::undistort method) preceded by a compulsory calibration. In the current study, a readily available example in Aruco module has been used to carry out the camera calibration stage, yielding in turn the so-called camera matrix and distortion coefficients vector.

In Fig. 8, the C-Star pilot station screen is shown. For the purpose of this experiment, the screen has been divided into four groups (information sources) according to what type of information each group outputs. Broadly speaking, following groups might be distinguished: Group 1 – navigation map; Group 2 – video channel; Group 3 – instrumentation panel; Group 4 – onboard sensors data. The number of fixations is to be counted for each information group separately.

Fig. 8. Information sources available at the pilot station

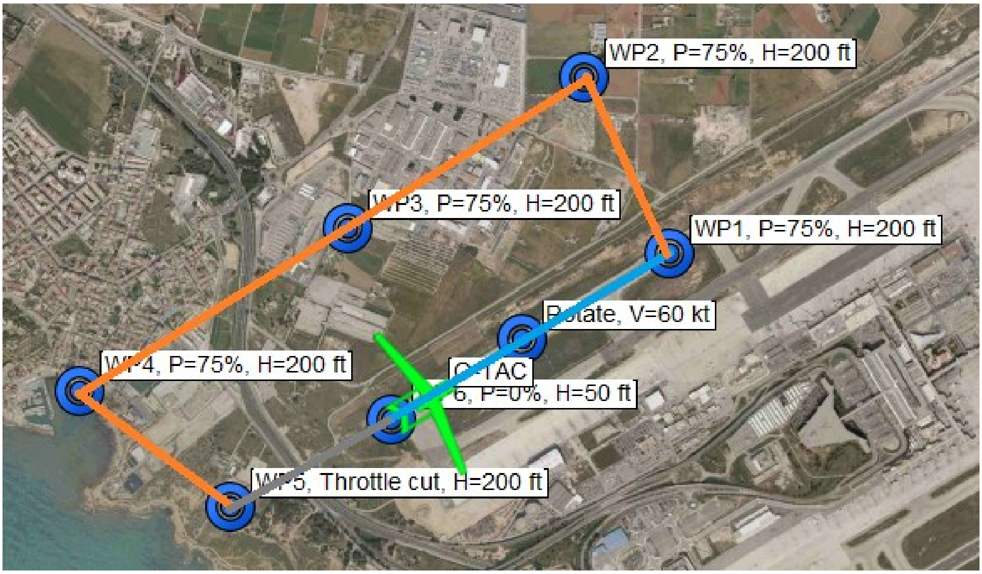


Fig. 9. Flight stages

In Fig. 9 three flight stages are shown any flight is formally divided into: 1) from the start to the first turn; 2) up to the 4th turn; 3) to stop. In this case, the distribution of fixations is determined for each flight stage for each information source separately.

Results

In Fig. 10, a frame taken by the environmental camera is shown.

Four fiducial Aruco markers are attached to corners of the monitor to define a region of interest and measure the relevant coordinates. OpenCV draws a green square, indicating that the method cv::aruco::detectMarkers succeeded in recognizing the current marker. Only then can the geometric center of each marker (red dots) be calculated. In addition, the barrel distortion is no longer observable because numerical compensation has been applied to each frame. Finally, a dynamic heatmap, which is built in advance, is superimposed on each frame.

In Fig. 11, the transformed (cv::undistort method) region of interest at exactly the same frame is shown. Evidently, that part of the image surrounding the region has been removed after the perspective transformation. The process resembles image cropping. In this way, spontaneous movements of the operator's head are eliminated. The mobile eye tracker works like a desktop based one.

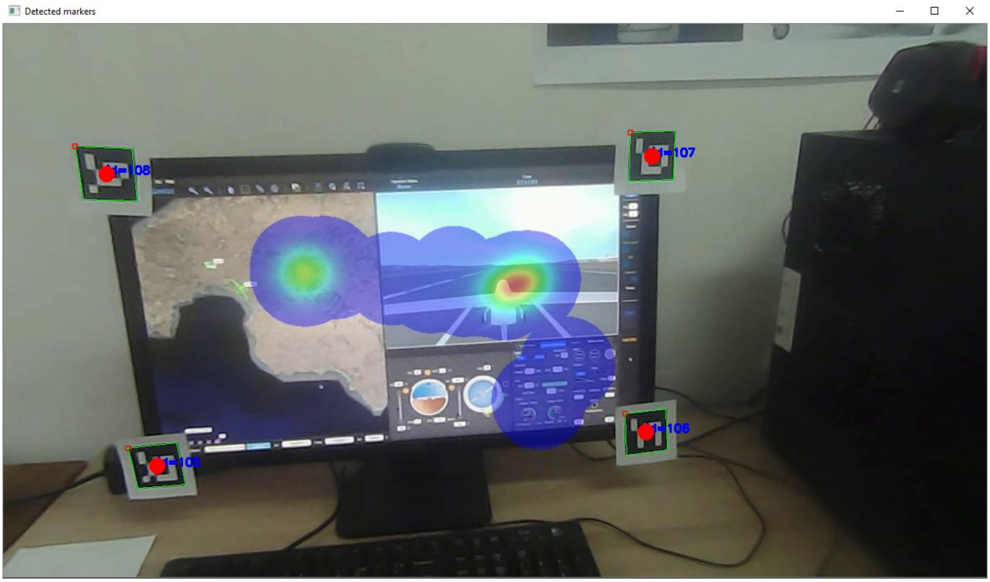


Fig. 10. C-Star pilot station with recognized Aruco markers, imposed dynamic heatmap and compensated barrel distortion

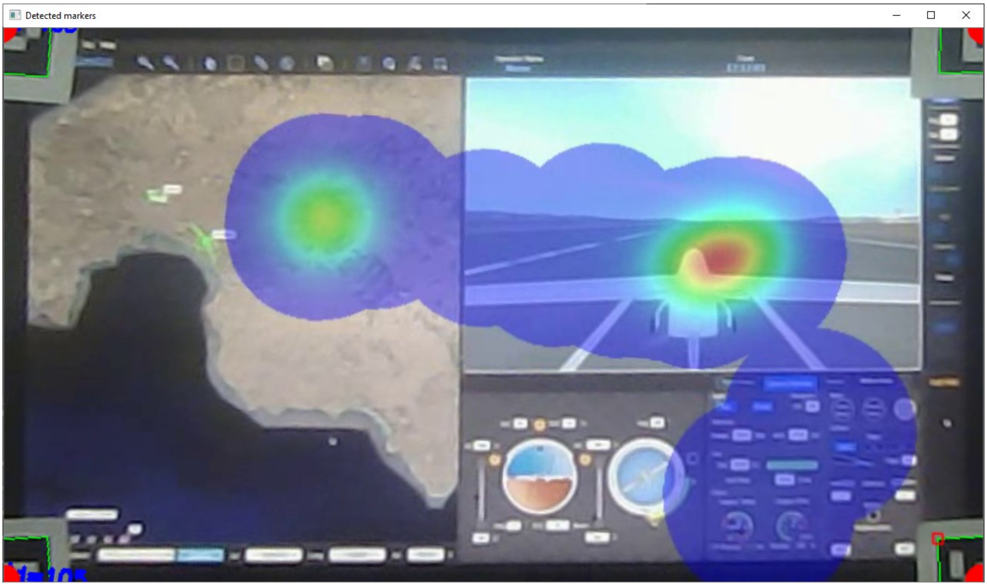


Fig. 11. Transformed region of interest

In Fig. 12, a detected fixation within the video channel is shown shortly after take-off. The fixation duration is 1 316 ms. Severe meteorological conditions are excluded in this very scenario, so are equipment failures. The flight task is the easiest possible, including stages take-off, flight along the approach pattern, landing approach and landing. The pilot has a lot of experience.

It is necessary for the algorithm to determine the number of fixations both within individual areas of interest, Fig. 8, and flight stages, Fig. 9. In Fig. 12, the current number of fixations is shown in the upper right corner of each information source. For the video channel, for example (second information source), the counted fixations from the start to the “Rotate” point are nineteen.

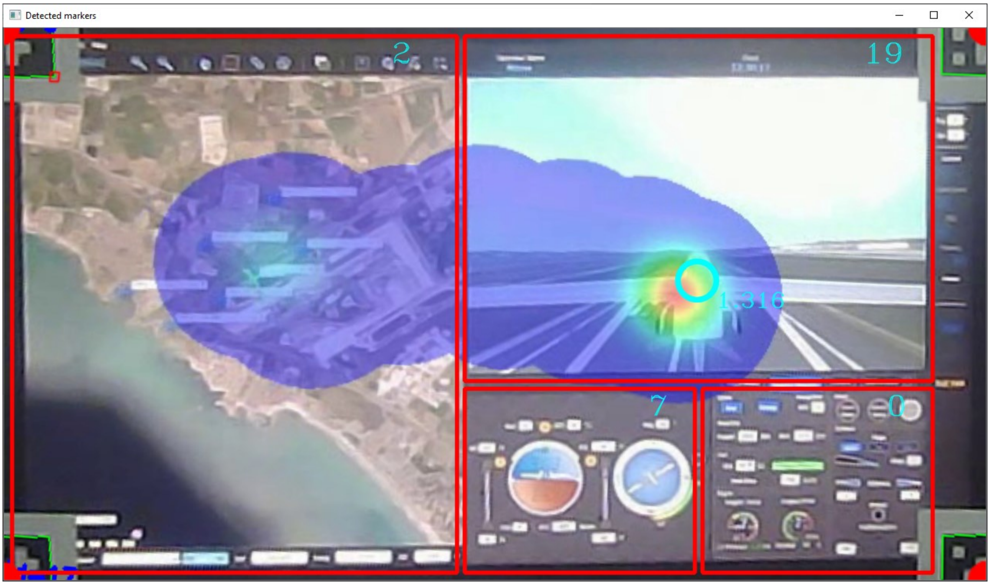


Fig. 12. A fixation lasting 1316 ms found within the second information source

In Table 2, the number of fixations and derived parameters for the three fight stages are shown without taking into account distribution by information sources, i.e. for full screen. The measurement was carried out by a Pupil Labs Core eye tracker on a C-Star simulator, SimLat.

Table 2

Stage / Parameter	Fixations number	Fixations per second, sec^{-1}	Mean duration, ms	Standard error, ms
Stage 1, 41 sec	75	1.828020	543.600	68.391
Stage 2, 95 sec	214	2.245495	420.243	27.796
Stage 3, 38 sec	36	0.966699	1 030.441	216.606

The fixation durations for each flight stage are shown in Fig. 13.

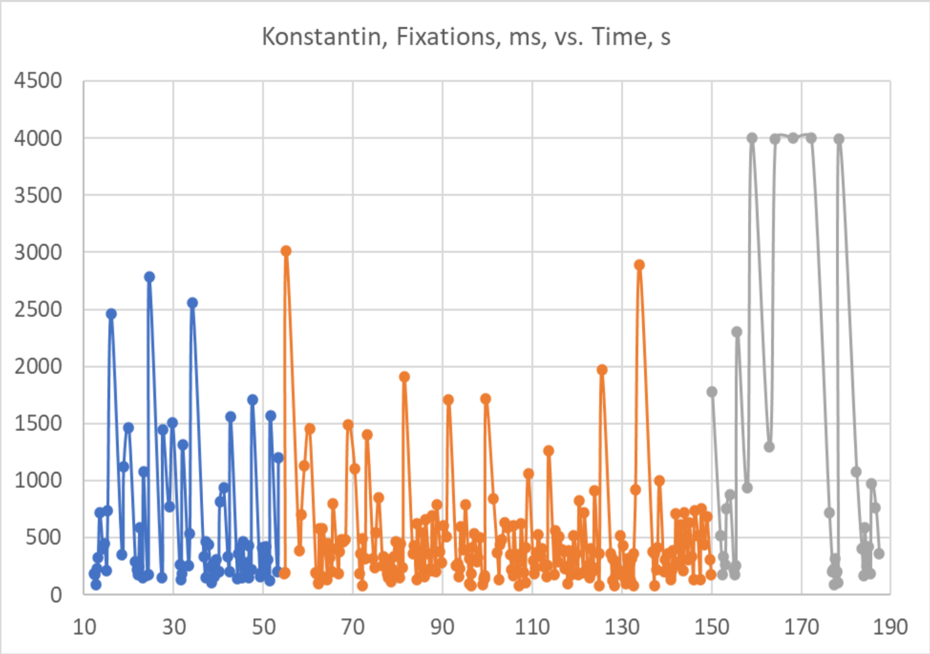


Fig. 13. Fixations distribution among the flight stages

In Table 3, distribution of the fixations is shown among both flight stages, Fig. 9 and information sources, Fig. 8. At this point, a check can be made for correctness of the obtained data. The total sum of fixations in each row of Table 2 must be equal to the number of fixations in column 2 of table 3 for the corresponding stage.

Table 3

Stage \ Information source	**1**	**2**	**3**	**4**
1 (up to 1st turn), 00:12 – 00:53	10	43	20	2
2 (up to 4th turn), 00:54 – 02:29	96	85	33	0
3 (up to stop), 02:30 – 03:08	2	30	4	0

4. Discussion

Image distortion compensation seems to be quite laborious as compared to other stages of the presented study. As a whole, there are two types of image distortions: radial and tangential. Typically, the former distortion is predominant owing to lens geometric parameters. The latter distortion is due to the imperfect lens alignment as parallel to the sensing element as possible. Although OpenCV provides readily available methods for compensating image distortion, analytical approaches toward rectifying image exist which are worth mentioning and recommended to further project development. These are Brown–Conrady model and the Division model, [8, 9]. In both models, it is required for the distortion center to be estimated in advance.

Low resolution of the transformed region of interest might be pointed out as a minor drawback. Obviously, the wider the lens' field of view, the worse the transformed image quality. However, for the transformation to be successful, it is necessary all four fiducial markers to be visible and discoverable within the image frame. It is easier to meet this certain requirement by means of a lens with shorter focal length.

The obtained results indicate that the proposed solution might be successfully applied to any kind of monitors and imaging screens. Unlike the desktop-based eye trackers, the mobile device functionality is not limited to the screen size. What is more, the displayed stimulus does not depend upon the window manager (DirectX, OpenGL) and screen resolution. Therefore, the stimulus may run in either of two modes: full screen or windowed. This seems to be in a more advantageous position over desktop-based eye trackers. Some of them fail to output a record if the stimulus, a game for instance, is displayed in full screen mode. What is more, the mobile eye tracker can be plugged in to an alternative device (PC, tablet, smartphone, etc.) to make the gathering data task less demanding.

According to obtained results, an experienced pilot switches their attention between information sources less. The reason is flight scenario simplicity. It is only sufficient for the pilot to look at the video channel straight so as to acquire necessary information and accomplish flight task. Furthermore, the Euler angles (pitch, roll) are limited by the C-star software enabling the pilot to perform basic flight maneuvering well. Almost all fixations are caught into the second information source, i.e. the video channel. In addition, as few as seven blinks were registered during entire flight.

Our experience in using the Pupil Labs mobile tracker led us to the following conclusions:

1) The presented complex of equipment will let us study the visual attention distribution of UAV operators during training at a flight simulator.

2) Studying peculiarities of UAV operators' visual attention distribution by means of eye-tracker will justify future creation of a science-based methodology for training students in drone control in different conditions and flight modes.

The data analysis derived from the registered fixations, with respect to the three conditionally separated flight stages, shows that:

- The fixation records are qualitative and admit processing and analysis.

- The flight conditional division allows for analysis of visual attention in terms of number and duration of the individual fixations.
- Fixations distributed on information sources are more difficult to register, especially on the instrument panel due to the specifics of the screen of the simulator itself.

A data interpretation derived from the registered fixations follows.

- Approximately identical depicted models of fixations throughout the flight is established for all test operators.
- The first stage – take-off, is characterized by an increased frequency of fixations at constant duration.
- Among all observed areas of interest, the operators focus their attention on the external environment more than the instruments.
- The second stage – cruise flight, is distinguished by a constant and slightly changing nature of the fixations for all operators.
- Landing features much more diverse fixations for individual operators, though, as long as all operators are concerned, an extended fixation time is observable as well as a decreased number of fixed informational sources.

From the obtained results and observation, the following conclusions can be derived, which, due to the limited number of experiments, cannot be presented as conclusions with a certain objectivity and reliability.

- The methodology at this stage for studying the cognitive functions of UAV operators by means of oculographic methods and developing an experimental flight model, including both visual control of the environment and certain parameters for maintaining flight on the available instruments, i.e. objective performance of the task might be considered developed and can be improved for the purposes of the study.
- There is an initial evidence that oculographic research reveals opportunities in studying effectiveness of the distribution of visual attention during training and improving the UAV operator skills.
- It is impressive from the fixations data that the landing stage appears most difficult to perform, i.e. efforts in the preparation and development of reflexes should be concentrated, contributing to fast and accurate perception and quality cognitive functions, which is a prerequisite for safe and effective operator work.

Source code developed to meet the presented research needs might be downloaded for free by following link [10].

Acknowledgements

The current study has been funded by the Scientific Research Fund of the Republic of Bulgaria according to research contract No. КП-06/Н27-10, 11th of December, 2018. The project title is “Human Factor in Remotely Controlled Aerial Systems – Analysis, Estimation, and Control.”

References

1. Diaz Tula, A., A. T. N. Kurauchi, F. L. Coutinho, and C. H. Morimoto, Heatmap Explorer: An Interactive Gaze Data Visualization Tool for the Evaluation of Computer Interfaces, Proceedings of the 15th Brazilian Symposium on Human Factors in Computing Systems (IHC'16), October, 2016. https://www.ime.usp.br/~kurauchi/publication/heatmap-explorer/
2. Bradski, G., The OpenCV Library, Dr. Dobb's Journal of Software Tools, article id=2236121, 2000. https://www.drdobbs.com/open-source/the-opencv-library/184404319
3. Simlat, 2021. https://www.simlat.com/
4. Kassner, M., W. Patera, and A. Bulling, Pupil: An Open Source Platform for Pervasive Eye Tracking and Mobile Gaze-Based Interaction, Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication, 2014, 1151–60. New York: ACM. DOI:10.1145/2638728.2641695 arXiv:1405.0006. https://arxiv.org/pdf/1405.0006.pdf
5. Węglarczyk, S, Kernel Density Estimation and its Application, XLVIII Seminar of Applied Mathematics, ITM Web of Conferences. 2018, 23, 00037.
6. Romero-Ramirez, F. J., R. Muñoz-Salinas, R. Medina-Carnicer, Speeded up Detection of Squared Fiducial Markers, Image and Vision Computing, 2018, 76, 38–47.
7. Garrido-Jurado, S., R. Muñoz Salinas, F. J. Madrid-Cuevas, R. Medina-Carnicer, Generation of Fiducial Marker Dictionaries Using Mixed Integer Linear Programming, Pattern Recognition, 2016, 51, 481–91.
8. Calib3D, OpenCV. https://docs.opencv.org/2.4/modules/calib3d/doc/camera_calibration_and_3d_reconstruction.html
9. https://www.imatest.com/support/docs/pre-5-2/geometric-calibration/distortion-models/
10. https://github.com/samolet4e

ИЗПОЛЗВАНЕ НА ОКУЛОГРАФ ЗА ИЗСЛЕДВАНЕ НА РАЗПРЕДЕЛЕНИЕ НА ВНИМАНИЕТО НА ОПЕРАТОРИ НА БЕЗПИЛОТНИ ЛЕТАТЕЛНИ АПАРАТИ

З. Хубенова, К. Методиев, С. Димитрова, Л. Алексиев

Резюме

В настоящата статия се предлага подход за изследване на причини за разпределение на вниманието на оператор на безпилотен летателен апарат. По време на изследването операторът е тестван на полетен симулатор и данни са събирани и визуализирани с помощта на мобилен окулограф. Два етапа на работа са разгледани последователно: построяване на 2D трансформация на регион на интерес (хомография) в рамките на изображение и наслагване на динамична топлинна карта. По време на първия етап се елиминират спонтанни движения на главата на оператора, което позволява мобилният окулограф да се използва като настолен. По време на втория етап се изобразява разпределението на вниманието на оператора.

За да се реализира настоящето изследване бе разработен код на C++ с използване на методи, налични в библиотека OpenCV. Събраните данни от изследването са обработени и обсъдени подробно.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e17

ANALYSIS OF THE MILITARY APPLICATION OF UNMANNED AIRCRAFT AND MAIN DIRECTION FOR THEIR DEVELOPMENT

Nikolay Zagorski

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: nzagorski@space.bas.bg

Keywords: *Unmanned Aircraft, Military Conflicts, Military UAV*

Abstract

Modern military unmanned aerial vehicles (UAVs) are assigned a wide range of functions, for the implementation of which they perform many tasks in various military conflicts. The results of the analysis give them the opportunity to reveal the problems in the use of UAVs, make changes in their functions and tasks and identify areas for further development. At present, this requires the introduction of the achievements of artificial intelligence, the introduction of expert systems and microelectronics on board UAVs, as well as their integration with various other means of conducting armed struggle. At the same time, some of the technological solutions for the creation and improvement of UAVs for military purposes can be applied in the civilian sector.

Introduction

One of the most characteristic features of modern military conflicts is the integrated process of using real-time combat information at all hierarchical levels responsible for decision-making.

Military experts from the world's leading aerospace countries are of the opinion [3] that in today's combat environment, reconnaissance unmanned aerial vehicles (UAVs) can solve tactical air reconnaissance and electronic warfare (EW) tasks with greater efficiency, effectiveness and efficiency, than manned aircraft (MA). The advantages of UAVs are manifested to a significant extent in the functions of targeting and correcting artillery fire, in combat control and retransmission (transmission) of signals, in biological, chemical, radiation and meteorological intelligence, without risk to personnel involved in these activities. The term for providing the intelligence received from the UAV to the end user is also shortened, especially in real time.

The basis for the widespread use of UAVs for intelligence was laid in the Vietnam War [3], and their combat use began in 1964, with the beginning of large-scale US hostilities in Vietnam. In particular, as a result of the significant combat losses suffered in US pilots and aircraft, air reconnaissance was launched with the

AQM-34L Firebee UAV, with equipment for photo, infrared and electronic reconnaissance. As a result, in the period from 1964 to 1975, various modifications of this type of UAV performed over the territory of Vietnam and South China a total of 3 435 battlefields, of which 2 873 flights (84%) were successful, with only 4 combat losses %. Not a single US serviceman was killed in all the UAV combat missions.

Study area

Modern military UAVs perform a wide range of functions, for the implementation of which they perform many tasks over the territory of hostilities and in the depth of the enemy's position. At the same time, as a result of the development of high technologies and innovations in military affairs, the functions and tasks of UAVs often change, in real time. At present, there is no unified definition and classification of UAVs, but such have the United States, Russia and other developed countries in terms of aviation. For example, in the European classification UAVs are divided into the following classes: micro and mini UAVs; for close action; for action at low altitudes; for action at medium heights; for high altitude action; for long-range operation, for medium-altitude operation with long flight duration; for operation at high altitudes with a long flight duration; shock UAVs [3].

Depending on the range of military UAVs can be divided into tactical, operational-tactical, operational and strategic, and by purpose can be reconnaissance; percussion; multi-purpose (intelligence-strike), for targeting; for correction of artillery fire; for signal retransmission; for REB; for radiochemical intelligence; such as air targets, false targets, etc. [1,3].

According to the mass, UAVs are subdivided into: micro UAVs (<5 kg); mini UAV (5 ÷ 200 kg); medium UAVs (<2 000 kg); large (<5 000 kg) and heavy (> 5 000 kg) [2].

UAVs, on the other hand, are a component of an unmanned aerial system (UAS), which is essentially a target system comprising several UAVs, a ground control and management station, and security facilities [3].

Due to the fact that there is no pilot on board the UAV, the value of this device is significantly lower than that of manned aircraft (ten times or more). According to US experts, the preparation of each pilot for combat aircraft to reach the level of "combat use" costs millions of dollars (7–8 and more) and is commensurate with the value of his aircraft. For this reason, compared in terms of “efficiency/value”, UAVs are significantly superior to manned aircraft.

Research method

The material proposes an approach derived from the theory of systems analysis [6], showing the relationship between the functions of UAVs, the tasks associated with the implementation of each of these functions, the results obtained after the tasks, the analysis of results, problems the tasks, the proposed solutions, as well as the feedback in the logical block diagram with these functions and tasks (Fig. 1). The proposed approach can be defined as functional.

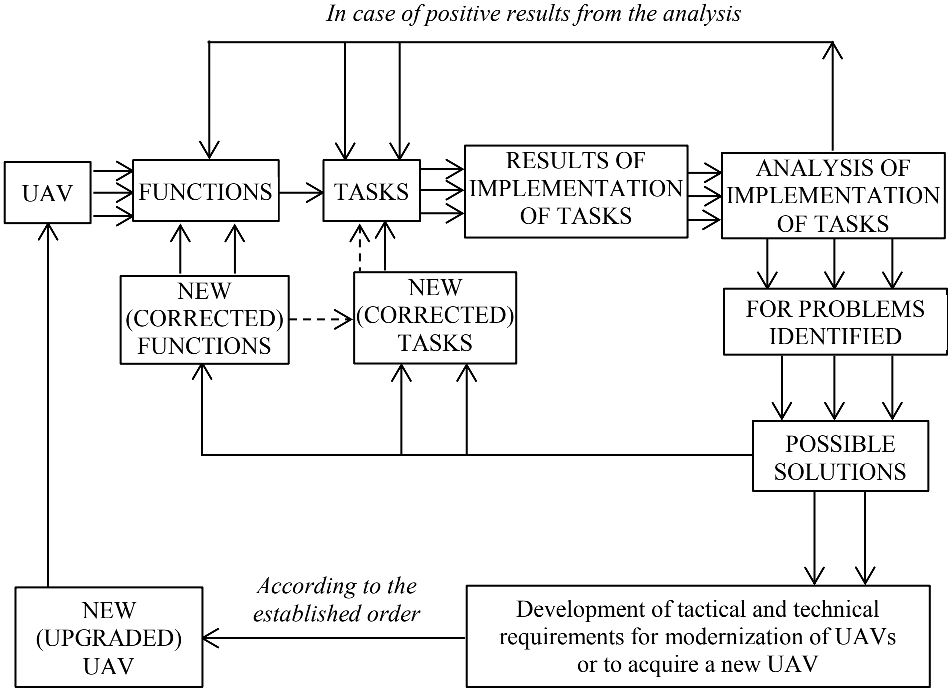


Fig. 1. Block diagram of application of system analysis for determining the functions and tasks of UAVs

As can be seen from Fig. 1, in case of positive results from the UAV actions, as a result of analysis, if necessary, the functions and tasks indicated in the upper left corner of this figure are repeated. Otherwise, after an analysis, the problems that have arisen and the possible solutions that have arisen are defined (lower right corner of the figure). In this situation, there are two possibilities. The first is the adjustment of the functions and tasks of UAVs or the formulation of new ones, and the second is the development by a team of diverse specialists of tactical and technical requirements for modernization of UAVs or the purchase (creation) of new UAVs.

The results of the solved tasks with the help of UAVs are evaluated on the following three-point scale:

- complete success: the task is completed;
- partial success: the task is partially completed;
- without success: the task is not completed.

Military conflicts with the active participation of UAVs

This material is the study of the battle using UAVs in the war in Lebanon in 1982, in “Operation Desert Storm” in Iraq in 1991, in “Operation Allied Force” in Yugoslavia in 1999, “Operation Enduring Freedom” in Afghanistan in 2001, the military action in Syria (2014 ÷ 2020) and the military action in Nagorno-Karabakh in the autumn of 2020.

The Lebanon War, 1982. During the Syrian-Israeli conflict in Lebanon, both countries actively used UAVs, mainly to solve air reconnaissance tasks. During the war, a new class of UAVs appeared – small UAVs, which due to their size are weakly vulnerable to missile air defense systems (ADS). New ways of conducting combat operations, respectively, new functions of UAVs appear, for each of which specific tasks are formulated and set.

The results of the implementation of the functions and the tasks solved by the UAV in Lebanon are shown in Table 1.

Table 1. Functions and tasks solved during the fighting in Lebanon by the Israeli UAV (1982)

№	Function	Tasks solved	Results	Problems
1.	Air reconnaissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows to the enemy.	Complete success	Strong enemy air defenses
2.	Targeting	- transmission from the UAV board in real time of data for the purposes.	Complete success	Strong enemy air defenses
3.	Artillery fire adjustment	- control of strikes on enemy targets and adjustment of artillery fire.	Complete success	Strong enemy air defenses
4.	***<u>New</u>:*** Complicating the air situation	- imitation by UAV on a flight of combat aircraft; - use of new tactics - small UAVs in “Birds in a flock”.	Complete success	Strong enemy air defenses
5.	***<u>New</u>:*** Demonstrative (distracting)	- to mislead operators of radar stations (RS); - radar stations to go into combat mode; - unmasking of air defense funds; - lifting the enemy's fighters into the air, as a result of which they waste their fuel aimlessly.	Complete success	Need for modeling air attack on enemy sites

During the operation on June 10, 1982, the Israeli aviation used the small-scale Scout and Mastiff UAVs on a large scale, using the “Birds in a flock” tactic.

In this tactic, UAVs fly in large groups at low altitudes, thereby misleading radar station (radar) operators that manned aircraft are approaching them. These groups of UAVs complicate the air situation for air defense and aviation [1, 3].

The intelligence information from the UAV board arrives at the land management points in real time and with great accuracy for the coordinates of the targets. These data are extremely important for the implementation of functions № 1, 2 and 3 of Table 1.

Following the implementation of function № 4, as a result of the demonstrative (distracting) function of the Israeli UAVs (function № 5 in Table 1), the Syrian operators switched the radar statin to combat mode, whereby the stations and Anti-Aircraft Missile Complex (AAMC) became "visible" to the Israeli aviation. The Syrian fighter jets raised in the air meet no enemy and, after running out of fuel, return and land at the airports for bases. In the next moment, the Israeli strike aircraft, in the absence of resistance from the Syrian fighters, took to the air, bombing and rocketing aircraft. During this operation, 18 AAMC batteries and 86 aircraft were destroyed [4, 5], which changed the outcome of the war.

Operation Desert Storm, Iraq, 1991. For the first time, the Multinational Force Command (MNC) uses different types of UAVs to conduct air reconnaissance – the American Pioneer, Pointer, Shadow-600 and the French Mart. At the same time, their implementation is considered limited for reasons of a different nature. The results of the implementation of the functions and the tasks solved by the UAV of the MNC in Operation Desert Storm are shown in Table 2.

Table 2. Functions and tasks solved by the MNC UAV in Operation Desert Storm (1991)

№	Function	Tasks solved	Results	Problems
1.	Air reconnaissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows to the enemy.	Complete success	- extremely large amount of intelligence information; - difficult coordination of data from different intelligence sources; - low reliability of UAV sensors.
2.	Artillery fire adjustment	- control of strikes on enemy targets and adjustment of artillery fire.	Complete success	
3.	Complicating the air situation	- launching "false targets"; - change of heights for UAV use.	Complete success	
4.	**New:** Strike at ground targets	- destruction of a mobile radar station by a Predator UAV armed with a radar warhead.	Incidentally	Simultaneous solving of a reconnaissance-strike task by one UAV

* Incidentally: The task is performed for the first time or experimental.

In the course of hostilities against Iraq, in the interests of the ground forces and the Marines, UAVs were deployed and used on foreign territory, each of which included 4 to 5 RQ-2 Pioneer UAVs.

The analysis of the fighting in Iraq shows the active use by the French armed forces of the “Mart” UAV, based on the territory of Saudi Arabia. They mainly perform tasks of monitoring, targeting, correcting artillery fire, conducting air reconnaissance at the operational and tactical level. At the beginning of the offensive operation, the French UAVs were operating at an altitude of 300 m. What is new in the process of combat operations is the change of the echelons for the use of “Mart” UAVs, and in order to avoid a collision with the strike aircraft, the flight altitude of the UAV is reduced to 150 m (functions № 1, 2 and 3 of Table 2).

Operation Allied Force, Yugoslavia, 1999. Different types of UAVs are involved in this operation in five of the participating countries. For a long time, low clouds remained in the area of hostilities (March and April), which did not allow the optoelectronic reconnaissance satellites in orbit to provide the necessary information. This necessitates the successive launch and use of several UAVs to ensure a continuous flow of intelligence.

In Kosovo, UAVs are used at medium altitudes, provide a constant flow of information for the movement of Serbian armored vehicles and solve the following main tasks: reconnaissance of the group of troops in the plains and the routes for their movement; detection and tracking of mobile AAMCs and radars; reconnaissance of the results of air strikes on sites; targeting, for which three Predator UAVs are equipped with laser targeting and guidance systems.

The results of the performance of the functions and the tasks solved by the UAV during the fighting in Yugoslavia in 1999 are shown in Table 3.

In the course of the operation, failures also occur as a result of impaired interaction and incomplete training of the operators. In particular, an operation is planned (function № 4 in Table 3), in which a US UAV must perform aerial reconnaissance over the territory of Kosovo, transmitting the information to a British UAV flying over the territory of Macedonia, and from there to a ground center for management. The actions of the American and British UAV operators, due to the heavily rugged terrain and the complex meteorological situation, do not ensure the flow of information to the ground control center.

In Operation Allied Force, about 500 flights with a total duration of 3 800 h were completed. Depending on the intensity of the strikes, there are from one to four UAVs in the air over the territory of Yugoslavia at the same time. During the entire period of hostilities, 27 UAVs were lost, six of which suffered accidents [1, 5]. The US and NATO commands acknowledge the high efficiency of the use of UAVs in Operation Allied Force, in contrast to satellite surveillance systems.

Table 3. Functions and tasks solved during the hostilities in Yugoslavia by the UAVs of the opposing states (1999)

№	Function	Tasks solved	Results	Problems
1.	Air recon-naissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows to the enemy;	Complete success	- bad weather (low clouds); - icing of UAVs; - rugged terrain of the battle zone; - high humidity.
		- detection and tracking of mobile AAMCs and radars.	Partial success	
2.	Targeting	- transmission from the UAV in real time of data for the purposes, including through laser devices.	Partial success	
3.	Weather recon-naissance	- gathering specialized information over the enemy′s territory.	Partial success	
4.	Retransmit data	- planning of 2 UAVs in one mission, respectively over Kosovo and over Macedonia, for retransmission of information to a ground control center.	Failure	- insufficient training of operators.

Operation Enduring Freedom in Afghanistan, 2001. This anti-terrorist operation operation uses, for the first time, the control of a Global Hawk UAV by satellites in orbit around the Earth. The operations of this UAV in the operation are limited to several fields and are related to ensuring the combat operations of the US Armed Forces in Afghanistan. The analysis of the use of the Global Hawk UAV in the operation shows that the main complexity lies in the management of their flight, which requires the presence of three free satellite communication lines. [2, 4].

In the course of the operation, the RQ-1 Predator UAV was used for the first time in a reconnaissance strike, directly striking ground targets with remote-controlled anti-tank missiles, air-to-surface class, Hellfire-C and Hellfire-K. The first target is a moving truck of terrorists, as the operator remotely directs the missile from the ground. Thus, including in conjunction with US Air Force strike aircraft, several dozen strikes were inflicted on mobile and stationary targets with high accuracy of hits.

A new UAV counter-terrorism function is also emerging in the fighting in Afghanistan, including the protection of convoys and the warning of terrorist attacks.

An analysis of the experience of using the Predator UAV in Afghanistan shows that its combat capabilities are limited at both high and low ambient temperatures. Low temperatures cause freezing and icing of the aircraft, and at high temperatures, for example, at $t° = +\ 38°C$ of ambient air, the start of the UAV should be done in the next $5 \div 10$ min. Otherwise, the UAV is not able to perform the combat task, as the temperature of the air inside it reaches t° = + 66 °C, and at such a temperature the reconnaissance equipment cannot work normally [5, 7].

An analysis of UAV operations in Afghanistan shows that of the three lost Predator UAVs, at least two are due to icing. Such a problem was already known during Operation Allied Force in Yugoslavia in 1999 and during the fighting in the Bosnian region. Table 4 shows the functions and tasks solved by American UAVs in anti-terrorist operation in Afghanistan in 2001.

Table 4. Functions and tasks solved during the actions in Afghanistan by American UAVs (2001)

<table>
<tr><th>№</th><th>Function</th><th>Tasks solved</th><th>Results</th><th>Problems</th></tr>
<tr><td>1.</td><td>Air recon-naissance</td><td>- collection of information on objects of the defense infrastructure;
- control of the results of the blows to the enemy;</td><td>Partial success</td><td rowspan="4">- bad weather (low clouds);
- high humidity;
- extremely high and low ambient temperatures;
- conditions for icing of UAVs;
- significant resource required (free channels) for satellite communication with the Global Hawk UAV;
- insufficiently good training of UAV operators.</td></tr>
<tr><td>2.</td><td>Targeting</td><td>- transmission from the UAV in real time of data for the purposes, including through laser devices.</td><td>Partial success</td></tr>
<tr><td>3.</td><td>Weather recon-naissance</td><td>- gathering specialized information over the enemy's territory.</td><td>Complete success</td></tr>
<tr><td>4.</td><td><u>New</u>:
Anti-terrorist</td><td>- hitting terrorist targets on the ground using guided anti-tank missiles;
- protection of convoys from terrorist actions;
- a warning of an attack by terrorists.</td><td>Partial success</td></tr>
</table>

Several other Predator UAVs were lost in Afghanistan due to insufficient operator training. Landing of UAVs of this type turns out to be a rather complex task and is performed by the operator manually, as UAVs during this period still do not have an automatic takeoff and landing system. This imposes the conclusion of the need for high professional training of UAV operators in peacetime.

Military action in Syria, 2014 ÷ 2020. According to the Chief of the General Staff of the Armed Forces of the Russian Federation, 60 ÷ 70 flights of Russian UAVs are performed daily in the skies over Syria. The main functions performed by UAVs are: air reconnaissance, electronic countermeasures (ECM), targeting and correcting artillery fire and air strikes.

With the help of UAVs, the so-called "reconnaissance and strike contours" are created, with the help of which control over the territory of Syria is ensured [10]. This control is also carried out by the deployed satellite group of Russia, and in order to clarify the characteristics of the detected targets by the satellites, the Russian UAVs conduct post-satellite reconnaissance.

Table 5. Functions and tasks solved by Russian and Turkish UAVs, as well as UAVs of Islamists during the fighting in Syria (2014 ÷ 2020)

№	Function	Tasks solved	Results	Problems
	Russia			
1.	Air recon-naissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows.	Complete success	Difficulties in aerial identification of terrorist structures
2.	Targeting	- transmission from the UAV in real time of data;	Complete success	
		- creation of "reconnaissance-strike contours".	Partial success	
3.	Artillery fire adjustment	- control of strikes on enemy targets and adjustment of artillery fire.	Partial success	
4.	**<u>New</u>:** Anti-terrorist	- collecting information on the movement of terrorist structures; - liquidation of terrorist leaders and manpower; - ECA of terrorists′ UAVs, control of their control and deviation from their targets; - destruction of UAVs by terrorists from AAMCs, covering relevant sites.	Complete success	
5.	**<u>New</u>:** Post-satellite recon-naissance	- specification of parameters of detected targets by intelligence satellites; - clarification of the location and type of mobile and stationary targets detected by the satellites.	Complete success	
	Turkey			
1.	Air recon-naissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows.	Complete success	Difficulties in aerial identification of terrorist structures
2.	Targeting	- real-time transmission of data for purposes, including with laser devices.	Complete success	
3.	**<u>New</u>:** Anti-terrorist	- collecting information on the movement of terrorists; - defeating mobile targets with terrorist leaders and manpower; - defeat armored targets.	Complete success	
	Islamic State terrorists			
1.	Strike at ground targets	- destruction of ground targets, including "Birds in a flock".	Partial success	Strong air defense and ECM
2.	Targeting	- directs a suicide bomber; - records the events.	Partial success	Overcoming ECM

Turkish UAVs are constantly patrolling the skies over northern Syria, with the Bayraktar TB-2 being particularly active. On a daily basis, one of these UAVs either hits a target or refines its location and directs F-16 or strike helicopters.

Table 5 shows the functions and tasks of UAVs in Syria. Terrorists in Syria use hand-made or modernized UAVs and apply modern targeting and control technologies (groups of 10 ÷ 15 UAVs, including "Birds in a flock"). In this sense, of interest is [8, 10] the massive use by terrorists against an air base and a naval checkpoint of the Russian Air Force in Syria of 13 UAVs with ammunition on board on the night of the 5th against the 6th of January 2018. As a result, these air targets were detected by the air defense system, and 7 of them were destroyed by the Panzer C1 AAMC. The remaining 6 UAVs were taken under control, as 3 were forcibly landed outside the territory of the concealed sites, and the other 3 detonated in a collision with the ground.

Recently, Islamic State (ISIL) operators have used a scheme in which a UAV indicates the target and directs a suicide bomber to the target vehicle, in which the drone controls the process and records events for propaganda use [7, 10].

The military operations in Nagorno-Karabakh in the autumn of 2020. In these actions, Azerbaijan widely uses strike UAVs and with their help a new model of combat operations in a local military conflict is being tested.

The main approaches to UAV operations in Nagorno-Karabakh by the Azerbaijani army are diverse. In the first place, strong and prolonged strikes by various types of UAVs are methodically applied, together with artillery and surface-to-surface missiles. Secondly, unmanned barrage ammunition (single-action) is used, as well as reconnaissance and strike UAVs for tactical purposes, equipped with high-precision ammunition with small dimensions.

Azerbaijan′s main air strike force consists of the Bayraktar TB-2 UAV, presumably operated by Turkish operators with significant combat experience in Libya and Syria. Used in conjunction with barrage ammunition, these UAVs have become an effective means of destroying tanks, vehicles and artillery and air defense equipment of the Armenian army. According to some data, the Bayraktar TB-2 UAV destroyed more than 60 T-72 tanks, others about 20 armored vehicles, 11 self-propelled howitzers, 5 Radar stations, 20 pcs. AAMC launchers, separate C-300 AAMC launchers, etc. [10].

The combat use of all types of UAVs in the operations in Nagorno-Karabakh clearly proves that the detection and destruction of targets with a low effective reflecting surface, such as unmanned barrage ammunition, is a difficult task for modern air defense systems.

On the other hand, Bayraktar TB-2 reconnaissance UAVs demonstrate medium-altitude tactics with small-scale high-precision weapons fired outside the air defense units of the ground forces. In practice, this allows UAVs to wage a "contactless" war, without human casualties, at relatively modest financial costs.

The success of the Azeri UAVs has been achieved mainly in the relatively flat areas. With the onset of autumn fogs and low clouds, the use of UAVs is severely limited in the mountainous and semi-mountainous and forested areas of Nagorno-Karabakh.

Table 6 shows the functions and tasks of UAVs in Nagorno-Karabakh (autumn, 2020).

Table 6. Functions and tasks solved by Azeri and Armenian UAVs during the fighting in Nagorno Karabakh (autumn 2020)

№	Function	Tasks solved	Results	Problems
	Azerbaijan			
1.	Air recon-naissance	- collection of information on objects of the defense infrastructure; - control of the results of the blows to the enemy.	Complete success in the plane part	Complex meteorological situation in the mountains
2.	Targeting	- targeting high-precision weapons with small dimensions; -management of unmanned barrage ammunition.	Complete success in the plane part	Overcoming the enemy′s air defenses
3.	Recon-naissance & strike at ground targets	- reconnaissance and execution of effec-tive strikes within one flight of a UAV; - use of UAVs in conjunction with artillery and surface-to-surface missiles; - launching UAV weapons outside the air defense zone.	Complete success in the plane part	
	Armenia			
1.	Air recon-naissance	- collecting information to move the enemy; - control of the results of the blows to the enemy.	Without success	Unable to fight barraging ammunition

While Azerbaijan uses mainly Israeli and Turkish UAVs, as well as its own production under an Israeli license, Armenia opposes several dozen UAVs, entirely its own development. They are far behind the Azeris, both in terms of technology and their number in the area of hostilities. The incidental use of Armenian UAVs is limited to a few reconnaissance fields.

The analysis of the fighting in Nagorno-Karabakh in the autumn of 2020 confirms the reconnaissance and strike function of the UAV with the ensuing relevant tasks. In particular, UAVs are increasingly performing integrated in-flight reconnaissance and strike functions, capable of carrying a variety of weapons or being used as “guided missiles”. After being used for the first time in Syria, the use of “kamikaze drones” is also confirmed here.

Main directions for UAV development

The analysis of the combat operations with the participation of UAVs in the military conflicts discussed above provides an opportunity to formulate the main directions for their development. These should be related to the introduction of the achievements of robotics, artificial intelligence, on-board expert systems, algorithms for “Birds in a flock”, as well as integration with other various means of armed struggle [2], in the context of ongoing miniaturization of UAVs.

The main directions for the development of military UAVs are specified according to their purpose.

UAV for reconnaissance and targeting. First of all, their development requires the implementation of complex reconnaissance with various sensors, including on-board radars with synthesized aperture, as well as improving the quality of algorithms in the program for flight control of UAVs in the respective modes, towards increasing autonomy and successful targeting.

Secondly, it is necessary to work on improving the programs and methods for coding radio signals, to create a “chaotic change of frequencies”, to use different modulations of the signal, etc., in order to maximize the difficulty of the opposing side.

The next direction is aimed at the development and improvement of data exchange systems with other aircraft and ground control points, the surveillance cameras used by UAVs and the image recognition algorithms. It is believed that UAVs will be able, observing the gathering of many people, to detect even suspicious persons and the presence of explosives, small arms and more dangerous objects [8].

UAV for strikes on targets. First of all, it is necessary to continue improving and developing the management of ammunition and high-precision UAV weapons used for strikes on land and sea targets. The modernization of the technologies and the element base will allow reducing their mass, to create light enough “smart” ammunition for UAVs with self-targeting systems.

Secondly, specialized UAVs (for example, the deck X-47A) with a high degree of invisibility should be used to break through the air defense system, which, after hitting the targets, should transmit information in a protected from interference mode to the following strike aircraft. To increase the strike capabilities of UAVs, some manufacturers are increasing the amount of ammunition on board. For example, China is building the world′s largest UAV, capable of carrying up to 24 air-to-ground missiles and staying in the air for up to 60 hours [10].

UAV for electronic warfare. It is necessary to improve the methods and techniques for electronic warfare in complicated air conditions, as well as the creation of false targets in the process of using UAVs.

Use of UAVs as air control points. Experts believe that equipping UAVs with computer systems with artificial intelligence (expert systems) will improve their security [2]. It is considered that in case of disruption or disconnection of the operator, such UAVs will be able to independently detect and identify objects, use the means of destruction and perform other independent actions.

Use of UAVs for logistics deliveries. The development of UAVs shows that a new UAV function can be established. For example, in modern military conflicts, the importance of small reconnaissance and sabotage groups of Special Forces, which periodically need to supply ammunition and material resources, is growing.

The listed UAVs for different purposes can be configured as convertibles.

Use of aerospace UAVs. Based on the X-37B aerospace UAV with a high degree of robotics, which has successfully completed several flights in the current decade, in the period until 2035 it is planned to produce and accept weapons in the US Air Force of several X-37B unmanned spacecraft [9]. They provide fast access to space and solve a wide range of tasks, among which the leading ones are reconnaissance and satellite inspection.

The analysis of the military applications of UAVs allows to formulate the main directions for the development of unmanned aerial vehicles for civil (commercial) purposes, such as: monitoring (surveillance and control) of the critical infrastructure of the country, mainly energy and transport; early detection of natural disasters (floods, fires, earthquakes, etc.) and man-made accidents, as well as damage assessment; independent control of the radiation situation in the Nuclear power plant area (s); monitoring of marine areas; assessment of environmental losses from various pollution and cataclysms; information provision of search and rescue operations; delivery of goods in hard-to-reach places; control of traffic on highways; cadastral survey of regions on the territory of the country; medical care, etc. The most serious problem in these cases is of a legal nature. In order to obtain the status of civil aircraft, each of the types of UAVs must pass the certification procedures in accordance with the airworthiness standards in force in the country.

Conclusion

The new functions and tasks of the UAV analyzed in the material require the respective scientific and engineering-technical researches and substantiations. According to some authors, in the coming decades, UAVs will replace manned aircraft. Obviously, a balance should be sought between these two types of aircraft.

References

1. Karemov, St. Air and space intelligence in achieving information superiority, Monograph, NMU, Faculty of Aviation, Dolna Mitropolia, 2019, 168 p.
2. Karemov, St., and R. Dimitrov, Unmanned aerial vehicles, status and prospects, MPH, Sofia, 2009, 135 p.
3. Penev, P., Fundamentals of the application of air power, Tutorial, Military Academy, Sofia, 2012, 182 p.
4. Vasilin, N., Unmanned aerial vehicles, LLC "Potpurri", Minsk, 2003, 269 p.
5. Mosov, S., Unmanned reconnaissance aircraft of the countries of the world, Monograph, Publishing house "Rumb", Kiev, 2008, 159 p.
6. Peregudov, F., and F. Tarasenko, Introduction to systems analysis, Higher School, Moscow, 1989, 367 p.
7. Poltavsky, A., S. Simeonov, and A. Burba, N. Phuong, Modeling of systems and objects of multifunctional robotic complexes of unmanned aircraft, publishing house of JSC "PSTM", Moscow, 2019, 407 p.
8. Sovenko, An., XXI century. In the sky - robots, Aviation and Time, Kiev, 2016, 3, 26–27, 45–46; 2016, 4, 22–23, 32–33; 2016, 5, 31–34; 2016, 6, 45–49.
9. Khabarov, E., E. Gagarin, and A. Terekhin, Development in the USA of aerospace unmanned aerial vehicles, Foreign Military Review, 2018, 3, 64–68.

Internet resources:

www.bbc.com>news-42596160
www.rg.ru/2018/01/08
www.fraza.com/analytics/294318
www.nvo.ru>1_1114_karabakh
www.vz.ru>world>2020/10/8/1063885

АНАЛИЗ НА ВОЕННОТО ПРИЛОЖЕНИЕ НА БЕЗПИЛОТНИТЕ ЛЕТАТЕЛНИ АПАРАТИ И ОСНОВНИ НАПРАВЛЕНИЯ ЗА РАЗВИТИЕТО ИМ

Н. Загорски

Резюме

На съвременните военни безпилотни летателни апарати (БЛА) се възлагат широк спектър от функции, за реализацията на които те изпълняват множество задачи в различни военни конфликти. Резултатите от анализа им дават възможност да се разкрият проблемите по използването на БЛА, внесат изменения във функциите и задачите им и се набележат направления за по-нататъшно развитие. Понастоящем това налага въвеждане на постиженията на изкуствения интелект, внедряването на експертни системи и микроелектро-ника на борда на БЛА, както и интегрирането им с други разнообразни средства за водене на въоръжена борба. Едновременно с това част от техно-логичните решения за създаване и усъвършенстване на БЛА за военни цели могат да бъдат приложени и в гражданския сектор.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e18

A REVIEW OF EARTH OBSERVATION RESOURCES FOR SECONDARY SCHOOL EDUCATION – PART II

Lachezar Filchev[1], Ioannis Manakos[2], Rainer Reuter[3], Garo Mardirossian[1], Tsveta Srebrova[1], Lubomira Kraleva[1], Dimitar Dimitrov[4], Kalliroi Marini[2], Andreas Rienow[5]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences;
e-mail: lachezarhf@space.bas.bg, G.Mardirossian@space.bas.bg, lkraleva@eeobss.space, Tsveta_Srebrova@space.bas.bg
[2]Information Technologies Institute of the Centre for Research and Technology Hellas, Thessaloniki, Greece; e-mail: imanakos@iti.gr, marini@iti.gr
[3]Carl von Ossietzky Universität Oldenburg, Germany, Germany;
e-mail: rainer.reuter@uni-oldenburg.de
[4]National Institute in Geophysics, Geodesy and Geography – Bulgarian Academy of Sciences; e-mail: clgdimi@abv.bg
[5]Ruhr-University, Bochum, Germany;
e-mail: andreas.rienow@ruhr-uni-bochum.de

Keywords: *Secondary school, science education, Earth Observation education, STEM, EO, e-Learning.*

Abstract

This article is a continuation of an overview of the contemporary resources for Earth Observation (EO) education for secondary schools. The themes covered by the sequel are the main EO education initiatives supported by international, EU and national organisations, outreach activities, citizen scientists' projects and free and open source software (FOSS) EO tools. The article elaborates on future prospects of EO resource developments in the education system its relevance for society and its connection with STEM subjects.

Introduction

In the first part of this review, published in Vol. 32 of Aerospace Research in Bulgaria, we covered the worldwide web and e-Learning resources for Earth Observation (EO) education for secondary schools and the main EO education initiatives supported by international, EU and national organisations. It was made clear also by other authors, that despite the developments in the EO for school education [1, 2] it is still underutilized [3].

The aim of this review, which is Part II of the sequel, is to provide an overview of the available international, regional and national education resources in EO for secondary schools, to identify subjects who are not yet well represented in the curricula but would benefit from using remote sensing in the classroom, and to address several didactic aspects, which could help establish Earth Observation in school.

Data and methods

Similar to Part I of the review, Part II integrate the material presented at Frascati EO Education Workshop 2014 with other information obtained in the ESA LearnEO! Project and ESA EO Open Science 2016. It is also based on the experience obtained in project *Science Education through Earth Observation for High Schools* (SEOS)[1] supported in the 6^{th} FP-EU, FIS, and from the ongoing EO4GEO project (ERASMUS+, EC). The experience gained during EEOBSS project implementation also served as a basis for the review.

Earth observation resources for secondary school education

Part II of the review we keep on following the logic of Part I when presenting the resources. Thus, the main review findings are presented in alphabetical order of the material producer/provider. In Part II of the review article, EO education activities and resources are organised in the following categories: 1) outreach activities, 2) citizen science, and 3) software and tools.

Outreach activities

DLR

The *DLR_School_Lab Oberpfaffenhofen* [4] is an extracurricular science lab, designed to attract secondary school students to Mathematics, Informatics, Natural Sciences, and Technology (MINT). It has been developed and operated since 2003 and offers thirteen hands-on experiments for secondary school classes, as well as advanced teacher trainings in Physics and Geography.

ESA

ESA Space Education Resource Offices (ESERO) – so far such offices are established in the following ESA member-states: Belgium, the Czech Republic, Germany, Ireland, Netherlands, Nordic (Denmark, Finland, Sweden, and Norway), Poland, Portugal, Romania, and the United Kingdom. One of these offices is located in Belgium, which has its own website URL: http://www.esero.be and a Facebook page: https://www.facebook.com/ESERO.BELGIE. One of the most

1 SEOS project. http://www.seos-project.eu (last date visited 18 March 2021)

interesting and active offices is the ESERO-UK , which has a dedicated website to the Principia mission [5] of the UK member of the International Space Station (ISS) Tim Peake. The space education resource office has its own project activities related to ESA research priorities, among which is EO [6].

The collaboration between ESA and the German Aerospace Centre, DLR (Deutsches Zentrum für Luft- und Raumfahrt) was set up in 2012 to produce 3D videos that will be used in the virtual theatre facilities at ESA. Building on this and the DLR School Lab for schools, a School Lab was organised as part of the IEEE International Geoscience and Remote Sensing Symposium 2012 in Munich. Building on these activities and the education resources developed in the UK, a similar School Lab took place at the ESA Living Planet Symposium in 2013, in collaboration between ESA, DLR, the UK Space Agency (UKSA), the National Centre for Earth Observation (NCEO) and the National Space Centre in Leicester, UK. Similar resources and activities for similar outreach events may be included in future EO symposia and conferences.

New videos for schools and the general public have been created in ESA/ESRIN. This includes an educational video on EO and Volcanoes for schools, produced jointly with the Italian National Institute of Geophysics and Volcanology, INGV (Istituto Nazionale di Geofisica e Vulcanologia).

NASA

NICE – Innovations in Climate Education (NICE) were a competitive project to promote climate and Earth system science literacy among underrepresented minority groups to science careers and educational opportunities. The three-year project was implemented in partnership with Jackson State University (JSU) and MSU under the cooperative agreement "Strengthening Global Climate Change education through Remote Sensing Application in Coastal Environment using NASA Satellite Data and Models". The goal was to increase the number of high-school and undergraduate students at JSU, a Historically Black University, who are prepared to pursue higher academic degrees and careers in STEM fields [7].

University of Berlin (Freie Universität)

From 26 February to 22 May 2008, the German Space Agency (DLR) and the Freie Universität Berlin (FU) presented an exhibition on “The new picture of our neighbour Mars” at the university campus of Dahlem in Berlin. Results of the European Mars mission “Mars Express” were presented, in particular images taken by the High Resolution Stereo Camera (HRSC). A guided tour through the images and movies on display was presented to several groups of youngsters: starting from 4 to 15 years. To give a real 3-D impression of the Mars surface, the staff from “the Planetary Sciences and Remote Sensing Group” presented special fly-over animations by using a GeoWall display and by working with anaglyph movies.

Especially for HRSC data, make use of the press-release data and modify it appropriately so that children of all ages can extract the most important information.

Citizen Scientists (CS) Projects

ESA is funding *four citizen science (CS) projects* to test and explore the potential and limits of CS in EO [8]. These pilots look at agriculture support, forest monitoring, land-use classification and air quality.

The *Geo-wiki Project* [9] is a global network of volunteers who wish to help improve the quality of global land cover maps by adding information on speciesdistribution, habitat, ecosystems and other scientific landuse information.

A *Land Cover validation game*, developed as a web-based game, was developed by the VGI & Citizen Science – GEOlab, Polytechnico di Milano (Polymy) [10]. The web-application was successfully tested at a mapping party during the FOSS4G Europe 2015 conference.

EO Software and Tools

During the past years, numerous software packages and tools have been released, which can be used for EO education in class. Most of them are free and open source software (FOSS), which allows schools and teachers who don't have a dedicated budget for software licences to use free satellite data for their classes. Although their graphical user interface GUI) is developed with a varying level of complexity, they can still be used to perform alone or in combination of basic and even advanced image processing tasks.

The *Bilko* software [11] was first developed for UNESCO in 1987 to provide free image processing capability for education use. The software has since been updated regularly to keep pace with the developments in remote sensing technology. The LearnEO! The project has extended Bilko's capabilities to include support for data from ESA satellites such as ERS, SMOS and CryoSat, as well as along-track altimeter data from Envisat and Jason-2.

Erdas ER viewer (Hexagon Geospatial ©)**,** available free of charge, can be used to open raster image formats such as TIFF, GeoTiff, IMG, ECW, ERS, ALG, DAT and many more, and are capable of handling large file sizes [12].

ArcGIS for Schools is a recently announced programme for schools by ESRI Inc. The educational programme of the world renown GIS leader was altogether redesigned during past years, including licencing and the professionalisation of the offered courses, through academia and school education. One of the main benefits that ESRI provides during a pandemic is the free use of ArcGIS for schools, which is part of the company social commitment and responsibility. The popular cornerstone product ArcGIS of ESRI is offered for free, including the new image processing (including EO data processing) functionality, which is already a part of the suite [13].

ESA toolboxes [14] are designed for professional users of EO data, but are also extensively used in the training of more advanced users, for example through computer exercises for university courses and shorter advanced courses and summer schools. The ESA toolboxes (mostly open source) are freely available from the web and include the *Basic ERS & Envisat (A)ATSR and MERIS Toolbox (BEAM), the Next ESA SAR Toolbox (NEST*), a toolbox for the scientific exploitation of polarimetric SAR data (*POLSARPRO)*, the *Basic Radar Altimetry Toolbox (BRAT*), the GOCE User Toolbox (*GUT*) and many more. The new satellite data from Sentinel-1, -2 and -3 will be processed using the developed SNAP with Sentinel Toolboxes. Even though more complex for education , they can be easily adopted in class due to its GUI similarities with the ESA education software LEOWorks.

Google Earth [15] has become one of the widely used EO data visualisation tools. Along with the API, the platform allowed many scientists and educators around the world to visualise their own data on a 3D surface. The features, such as street view, geo-tagged imagery, underwater relief attracted many individuals and companies to use the platform to build upon their own ideas. As of the past few years the Google Earth Project released the Professional version free of charge that increased its popularity.

In Europe, one of the earliest and most widely used software tools for image processing with the support of numerous satellite datasets is the *Integrated Land and Water Information System (ILWIS)* [16]. In late 1984, ITC was awarded a grant from the Dutch Ministry of Foreign Affairs. By the end of 1988, the project resulted in the official release of the DOS version 1.0 of ILWIS. The ILWIS was launched commercially two years later. As of July 1, 2007 ILWIS Open 3.X is available as 52°North free and open source software (GNU GPL).

Two software packages are indicated by ESA for general education purposes: LEOWorks and Bilko. The *LEOWorks* [17] is a didactic tool with extensive help pages and an all-inclusive tutorial, allowing students to process satellite imagery and combine them with other geospatial information. The new version is developed in Java, is platform-independent (Windows, MacOS, Linux) and will be released under a General Public License (GPL)[2]. It will include advanced GIS functionality and optical and SAR image processing in a user-friendly environment.

One of the honourable mentions for the EO education legacy software used intensively in the early 2000s is *MicroMSI*. The MicroMSI for Windows is a remote sensing imagery analysis programme designed for use in introductory courses in remote sensing, developed by the National Geospatial-Intelligence

[2] GPL. http://www.opensource.org/licenses/gpl-license.php (last date visited 18 March 2021)

Agency. MicroMSI for Windows is a "public domain programme and can be freely redistributed for non-commercial purposes", after modern terminology freeware. Although not supported, it still features some advanced image processing multispectral image analysis techniques and it supports hypercube visualisation HYDICE and basic spectral analysis [18].

MultiSpec© is being developed at the Purdue University, West Lafayette, IN, by David Landgrebe and Larry Biehl from the School of Electrical and Computer Engineering, ITaP and LARS. It results from an on-going multiyear research effort, which is intended to define a robust and fundamentally based technology for analysing multispectral and hyperspectral image data, and to transfer this technology to the user community in as rapid a manner as possible. The results of the research are implemented into MultiSpec and made available to the user community via the download pages. MultiSpec© with its documentation is distributed free of charge as a desktop online system on the following web-page [19].

One of the first EO software tools used in University education in Asia was Dragon/ips® software developed and marketed by Goldin-Rudahl Systems, Inc. As of 2010, Goldin-Rudahl Systems has agreed that the *Open Dragon* software, based on Dragon version 5, will be open source for non-commercial use. The Project is currently providing an Open version of the Dragon software which is used worldwide by schools and universities [20]. One of the Open Dragon Project is free lecture materials and hands-on developed and distributed software.

Orfeo ToolBox (OTB) is an open-source project of CNES for a state-of-the-art remote sensing. Although a complex one, its free distribution enables educators to use it even in class. It can process high-resolution optical, multispectral and radar images at the terabyte scale. A wide variety of applications are: from ortho-rectification or pansharpening, all the way to the classification, SAR processing, and more. It includes a fast image viewer, apps callable from Bash, Python or Monteverdi, QGIS, and a powerful C++ API [21].

Quantum GIS (QGIS) is a user friendly free and open source software (FOSS) Geographic Information System (GIS) licenced under the GNU General Public License. The QGIS is an official project of the Open Source Geospatial Foundation (OSGeo). It runs on Linux, Unix, Mac OSX, Windows, and Android and supports numerous vector, raster, and database formats and functionalities [22]. During the last years many educational projects and courses have been intensively relying on QGIS that is due its multifunctionality including GIS and EO capabilities. The wide user community, which produces scripts and extension, leverage its capabilities further, which makes the FOSS GIS and EO software a true alternative for education.

Discussions and conclusions

In the Part II of the review we provided a brief overview of outreach activities as well as the citizen science projects and software tools for EO education. With this we hope that we will enable the educators, scientists and student alike, to reach out much easier to the tools and resources needed to jumpstart with EO. However, some of the most pronounced key issues and perspectives with regard to secondary school education were outlined in Part I, which we will not restate here. Rather, we conclude with the self-imposing conclusion of the review that due to many external factors to the secondary school educational system worldwide the change from a non-systemic to a systemic approach in EO education is about to turn into a reality soon. This is mainly due to the existing critical mass of resources, tools, software, as well as projects and initiatives that all lead to a better realisation in the EO community and the policy-makers for the professionalisation of the EO. It is well known that the career choice is commonly made during the school years. This is becoming a well-understood fact not only to educators but also to scientists and university staff who always feel the pressure of the need of the well-educated students who are capable of undertaking more and more complex tasks involving EO data and resources. The process of transition is manifold and it also involves many actors on different levels but will inevitably lead to the formulation either of a common education framework for EO education in schools or in the worst case scenario – inclusion of the EO technology in various STEM disciplines taught in class. This is a necessity for times, which also witness an unprecedented inflow of EO data as well as galloping new technology developments and transformations that will build a bridge between education and practice.

Acknowledgements

The work was carried out as part of "Education in Earth observation for Bulgarian secondary schools – EEOBSS" Project (ESA-PECS for Bulgaria), Contract No. 4000117592/16/NL/NDe concluded between SRTI-BAS and ESA. The views expressed herein do not in any way reflect the official position of the European Space Agency.

References

1. Merry, C. J., and Stockman, S. Educational outreach activities for Landsat-7. Remote Sens. Environ., 2001, 78, 217–220. DOI:10.1016/S0034-4257(01)00261-9
2. Naumann, S., Siegmund, A., Ditter, R., Haspel, M., Jahn, M., and Siegmund, A. Remote sensing in school – Theoretical concept and practical implementation. In E-Learning Tools, Techniques and Applications; König, G., Lehmann, H., Eds.; ISPRS: Potsdam, Germany, 2009.

3. Dziob, D., Krupiński, M., Woźniak, E., and Gabryszewski, R. Interdisciplinary Teaching Using Satellite Images as a Way to Introduce Remote Sensing in Secondary School. Remote Sens. 2020, 12, 2868. DOI:10.3390/rs12182868
4. DLR_School_Lab Online. http://www.dlr.de/schoollab/en/desktopdefault.aspx/tabid-1708/ (last date visited 18 March 2021)
5. ESERO-UK. https://principia.org.uk/ (last date visited 18 March 2021)
6. European Space Education Resource Office. http://www.esa.int/Education/Teachers_Corner/European_Space_Education_Resource_Office (last date visited 18 March 2021)
7. Reddy, S.~R. Climate Literacy: STEM and Climate Change Education and Remote Sensing Applications. AGU Fall Meeting Abstracts. 2015. Abstract id. ED33A-0934
8. Educeo. http://educeo.net/ (last date visited 18 March 2021)
9. Geo-Wiki: Earth Observation & Citizen Science. IIASA. http://www.geo-wiki.org/ (last date visited 18 March 2021)
10. http://landcover.como.polimi.it/landcover/#/ (last date visited 18 March 2018)
11. Bilko Software. http://bilko.org/software.php (last date visited 18 March 2021)
12. Other Producer Products. Hexagon. http://www.hexagongeospatial.com/products/power-portfolio/other-producer-products/erdas-er-viewer (last date visited 18 March 2021)
13. Schools Mapping Software Bundle. ESRI. 2021. https://www.esri.com/en-us/industries/education/schools/schools-mapping-software-bundle (last date visited 18 March 2021)
14. TOOLS & TOOLBOXES. Earth Online. ESA. https://earth.esa.int/web/guest/pi-community/toolboxes (last date visited 18 March 2021)
15. Google Earth. https://www.google.com/earth/ (last date visited 18 March 2021)
16. ILWIS. http://52north.org/communities/ilwis (last date visited 18 March 2021)
17. LeoWorks. http://leoworks.asrc.ro (last date visited 18 March 2021)
18. Loomer, S. A. Remote sensing education with MicroMSI, In: Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004, 3, 1963–64. IEEE International
19. MultiSpec. https://engineering.purdue.edu/~biehl/MultiSpec/ (last date visited 18 March 2021)
20. Open Dragon. http://www.open-dragon.org/ (last date visited 18 March 2021)
21. Orfeo Toolbox. https://www.orfeo-toolbox.org/ (last date visited 18 March 2021)
22. QGIS. http://www.qgis.org/en/site/ (last date visited 18 March 2021)

ОБЗОР НА РЕСУРСИТЕ ПО НАБЛЮДЕНИЕ НА ЗЕМЯТА ЗА СРЕДНОТО ОБРАЗОВАНИЕ – ЧАСТ 2

Л. Филчев, Й. Манакос, Р. Ройтер, Г. Мардиросян, Ц. Среброва, Л. Кралева, Д. Димитров, К. Марини, А. Рийнов

Резюме

Тази статия е продължение на обзора (Част 1) на съвременните ресурси за наблюдение на Земята за средните училища. Темите, обхванати от продължението, са основните образователни инициативи, подкрепяни от международни, европейски и национални организации, дейности по разпространение на резултати, проекти свързани с гражданска наука и инструменти за работа (включително софтуер) със свободен и отворен код (FOSS). В статията се разглежда бъдещето на развитието на ресурсите за наблюдение на Земята в системата на средното образование, което е от значение както за обществото, така и за връзката на наблюдението на Земята със STEM дисциплините.

Bulgarian Academy of Sciences. Space Research and Technology Institute.
Aerospace Research in Bulgaria. 33, 2021, Sofia

DOI: https://doi.org/10.3897/arb.v33.e19

This article is written in honour of Yuri Gagarin,
Celebrated around the world every year on April 12.

ADVANCES IN SPACE SCIENCE AND TECHNOLOGY IN CONNECTION WITH 60−TH ANNIVERSARY OF FIRST HUMAN SPACEFLIGHT

Peter Velinov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: pvelinov@bas.bg

Keywords: *Yuri Gagarin, First Human Spaceflight, Space Science, Space Exploration*

Abstract

On April 12, 1961, Yuri Gagarin proclaimed the arrival of a new space age. The rapid advances in the different space sciences and technologies began after the first human spaceflight. Then, fundamentally new sciences and technologies appeared. At present, space science covers a broad range of disciplines. The following outline is provided as an overview and topical guide to space sciences:

Astronomy and Space Astronomy, Cosmology, Astrophysics, Space Physics, Solar-Terrestrial Physics, Aeronomy, Solar physics, Heliospheric Physics, Cosmic Ray Physics, Space Weather and Space Climate (Earth-Space Climatology), Space Dosimetry, Space Chemistry or Cosmochemistry, Remote Sensing of the Earth and Planets, Planetary Science, Planetary Geology, Astrogeology or Exogeology, Exoplanetology or Exoplanetary Science (Science for Extrasolar Planetary Systems), Intelligent Life in the Universe, Astronautics (or Cosmonautics), Orbital mechanics or Astrodynamics, Space life sciences: Bioastronautics, Space Medicine, Space Neuroscience, Space Biology, Radiation Biology, Biotechnology, Space Botany or Astrobotany, Microgravity Environment Research; Archaeoastronomy, Space Anthropology, Xeno-anthropology (Exo-anthropology), Space Law, Space Technology, Space Navigation, Space Communications, Space Architecture, Space Logistics, Space Robotics, Space Robotic Colonies, Space Colonization (also called Space Settlement or Extraterrestrial Colonization), Planetary Habitability, Space Manufacturing, Space Materials Science, Satellite Industry, Space Business, Space Tourism, Space Hardware, Space Industry, and Space Ecology.

With the help of these advanced space sciences humankind confidently began the exploration of space. But these studies led also to numerous new technologies and applications to improve people's lives.

Finally, we mention again Yuri Gagarin and his cosmic heritage. He left behind an inspirational legacy, which even today still continues to motivate millions of people worldwide.

Introduction

On October 4, 1957, the first human-made Earth satellite Sputnik I was launched into outer space, thus opening the way for space exploration. On April 12, 1961, Yuri Gagarin became the first human to orbit the Earth, opening a new chapter of human endeavour in outer space.

The Declaration of General Assembly of United Nations (UN) of April 7, 2011 recalls "the amazing history of human presence in outer space and the remarkable achievements since the first human spaceflight, in particular Valentina Tereshkova becoming the first woman to orbit the Earth on June 16, 1963, Neil Armstrong becoming the first human to set foot upon the surface of the Moon on July 20, 1969, and the docking of the Apollo and Soyuz spacecrafts on July 17, 1975, being the first international human mission in space, and recall that for the past decade humanity has maintained a multinational permanent human presence in outer space aboard the International Space Station."

The beginning of the Space Era for mankind

The 12 April 1961 was the date of the first human space flight, carried out by Yuri Gagarin (Figs. 1, 2), a Soviet citizen. This historic event opened the way for space exploration for the benefit of all humanity.

Fig. 1. Gagarin's name was immortalised when he made a single orbit of Earth on April 12, 1961. (Science reporters, BBC News, https://www.bbc.com/news/science-environment-12460720)

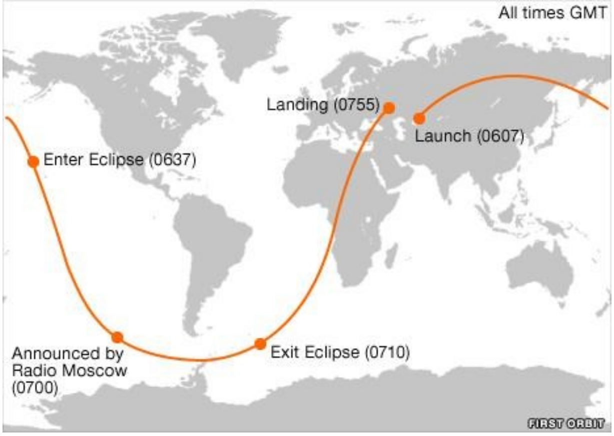


Fig. 2. The trajectory of the spacecraft „Vostok-1“. Gagarin went into darkness behind the Earth over the Pacific. He saw the Sun rise as he was moving over the South Atlantic. (Science reporters, BBC News, https://www.bbc.com/news/science-environment-12460720)

The General Assembly of UN, in its resolution A/RES/65/271 of April 7, 2011, further declared 12 April as the International Day of Human Space Flight “to celebrate each year at the international level the beginning of the space era for mankind, reaffirming the important contribution of space science and technology in achieving sustainable development goals and increasing the well-being of States and peoples, as well as ensuring the realisation of their aspiration to maintain outer space for peaceful purposes.”

The General Assembly expressed its deep conviction of the common interest of mankind in promoting and expanding the exploration and use of outer space, as the province of all mankind, for peaceful purposes and in continuing efforts to extend to all States the benefits derived there from.

UN and Space

From the very beginning of the Space Age, the United Nations recognised that outer space added a new dimension to humanity's existence. The United Nations family strives continuously to utilise the unique benefits of outer space for bettering all humankind.

Today, the United Nations work for international cooperation in peaceful uses of outer space with the United Nations Committee on the Peaceful Uses of Outer Space (COPUOS).

This Committee is also responsible for implementing the Secretary-General's responsibilities under international space law and maintaining the United Nations Register of Objects Launched into Outer Space.

The United Nations has approved for the international community the following periods related to space:

- International Day of Human Space Flight (12 April);
- World Space Week (4 – 10 October), associated with the celebration of the first artificial satellite Sputnik I;
- International Asteroid Day (30 June);
- International Space Year ISY (1992).

Space Science and Exploration

Space science makes us look outwards from our planet, to the stars and beyond. It's a subject who strives to answer the ultimate questions: How did our Earth and our Solar System form and evolve? What is our place in the Universe? Where are we going? Where did life come , and are we alone?

By studying the other planets in our own Solar System, we can try placing Earth in context. The space states, organizations and institutes, for instance: NASA, RosCosmos, CNSA (China National Space Administration), ESA, JAXA (Japan Aerospace Exploration Agency), etc. have already sent the spacecraft to Earth's nearest planetary neighbours – Mars and Venus – to understand why they evolved so differently, and in the next decade we'll be unlocking the secrets of the innermost planet, Mercury, too.

The gas giants, and in particular Jupiter with its four large moons – some of which may harbour underground oceans – is also a key to piece together the Solar System evolution. Unravelling the behaviour of our parent star, the Sun, is another crucial element to decipher our cosmic origins. And as more and more planets are found orbiting other stars, understanding our own cosmic neighbourhood has never been important.

What about the origin of the Universe itself? If we could 'see' microwaves, the night sky would glow with the first light ever released into space. This is the relic radiation of the Big Bang – the event that set the beginning to the Universe itself. With our space science missions we can tease out the details of the Universe's earliest moments, seek out the first stars and galaxies, and learn about the fabric of space and time.

These incredible space observatories can also give us clues to the future destiny of our Milky Way galaxy, and the Universe itself. The space science missions provide access to the largest science laboratory we have ever known: our Universe.

Overview of space sciences

Thousands of years ago, our remote ancestors looked up and wondered about their place between Earth and sky. Like them, we ask the same profound questions, such as how did the universe begin?

Today, we are beginning to answer these questions. Using tools of science that range from abstract mathematics and computer modelling to laboratories and observatories, humans are filling in the details of the amazing storey of the universe.

In the last 60 years, space probes and space observatories have played a central role in this fascinating process, and RosCosmos, ESA, NASA, CNSA, JAXA, etc. Space Science agencies will continue to address these four profound questions:

1) How did the universe begin and evolve? We seek to, explain the earliest moments of the universe, how stars and galaxies formed, and how matter and energy are entwined on the grandest scales.

2) How did we get here? We investigate how the chemical elements necessary for life have been built up and dispersed throughout the cosmos, evidence about how the Sun affects Earth, similarities between Earth and other planets, and how comets and asteroids in our solar system affect Earth.

3) Where are we going? Our ultimate place in the cosmos is wrapped up in the fate of the universe. Humanity has taken its first steps off our home world, and we will make it safe to travel throughout the solar system.

4) Are we alone? Beyond astrophysics and cosmology, there lies the central human question: Are we on Earth because of an improbable accident of nature? Or is life, perhaps even intelligent life, scattered throughout the cosmos?

Now, in support of the new vision of space exploration, orbiting observatories and planetary probes will be joined by human explorers in seeking answers to these questions. Robotic scouts will blaze the trail, reconnoitring the planets, moons, asteroids, and comets of the solar system in advance of human expeditions, as observatories monitor the sun and its effects on its planetary retinue.

The Space Science will work with new exploration systems to develop and deploy new technologies, first on automated spacecraft and then on human missions.

Branches of Space sciences

Space science covers a broad range of disciplines, from meteorology and geology, to lunar, solar, and planetary science, to astronomy and astrophysics, to the life sciences. Space science encompasses all of the scientific disciplines that involve space exploration and study natural phenomena and physical bodies occurring in outer space, such as space medicine and space biology [1, 2].

The following outline is provided as an overview and topical guide to space science:

<u>Astronomy</u> - sciences that study the laws of the stars are a natural science that studies celestial objects and phenomena. It uses <u>mathematics</u>, <u>physics</u>, and

chemistry to explain their origin and evolution. Objects of interest include sun, moon, planets, stars, nebulae, galaxies, comets, gas, dust and other non-Earthly bodies and phenomena. Relevant phenomena include supernova explosions, gamma ray bursts, quasars, blazars, pulsars, and cosmic microwave background radiation. More generally, astronomy studies everything that originates outside Earth's atmosphere. Cosmology is a branch of astronomy. It studies the Universe as a whole.

Space Astronomy is engaged now in observational research from instruments for the next generation and the study and launch of new improved space missions. Observations are conducted over a wide range of wavelengths, including radio submillimeter infra-red, X-ray and gamma-rays. Main targets include galaxy clusters, active galactic nuclei, galaxies, stars star-forming regions and protostars, supernova remnants, interstellar matter, exoplanets.

Astrophysics – study of the physics of the universe; of extraterrestrial objects and interstitial spaces. Astrophysics refers to the branch of astronomy dealing with the behaviour, physical properties, and dynamic processes of celestial objects and phenomena. For astrophysics are especially important results and experiments in space astronomy.

Space Physics studies physical processes and phenomena in outer space. Space physics is distinct from astrophysical plasma and the field of astrophysics, which studies similar plasma phenomena beyond the Solar System. Space physics utilises in situ measurements from rockets and spacecraft [1–3] in contrast to astrophysical plasma that relies on the deduction of theory and astronomical observation.

Solar-Terrestrial Physics, is the study of plasmas as they occur naturally in the Earth's environment - upper atmosphere (aeronomy) and within the Solar System [3].

Space physics and solar-terrestrial physics encompass a far-ranging number of topics, such as heliophysics, which includes the solar physics of the Sun [4]: the solar wind, planetary magnetospheres and ionosphere, galactic and solar cosmic rays [5–8], auroras, high energetic particles and fields, cosmic electrodynamics and synchrotron radiation [9–15].

Heliospheric physics research into the heliospheric structure and its relation to the solar boundary, also behavior of the heliospheric magnetic field and the solar wind. The heliospheric community search a progress towards a predictive model describing the connections between the Sun and its space environment, between the closed corona and the open corona extending to the planets. This requires an understanding of the basic processes heating the corona and transporting open magnetic field [1].

Space physics, space geophysics and solar-terrestrial physics are a fundamental part of the study of space weather and space climate (Earth-Space climatology) and has important implications in not only to understanding the universe, but also for practical everyday life, including the biomedical problems related to human health, environmental radiation, geomagnetic storms, operations of communications and weather satellites [16].

Cosmic ray physics - cosmic ray story begins at the beginning of XX century (they are discovered in 1912 by Victor Hess). More than 100 years later, most of the main issues are still open questions, as sources, acceleration mechanism, propagation and composition [17–21].

The search for a theory of the origin of cosmic rays that may be considered a standard agreeable model is still ongoing. However, much circumstantial evidence exists of the fact that supernovae in our Galaxy play a crucial role in producing the bulk of cosmic rays observed on Earth. Whether the maximum energy of light nuclei is as high as 1000 TeV has important consequences on the crucial issue of the transition from Galactic to extragalactic cosmic rays.

Space Dosimetry. Cosmonauts (astronauts) in space are exposed to a radiation environment that can have deleterious health consequences. This environment is both complex (trapped electrons and protons, galactic cosmic ray GCR – ions, secondary charged fragments, and neutrons) and dynamic (changing in time and orbital location). The probabilities of adverse health effects – the health risks of space radiation – cannot be measured directly, but must be calculated [22, 23].

Space chemistry or cosmochemistry is a relatively recent area of specialisation compared to astrophysics. Astrochemistry is the study of the abundance and reactions of molecules in the Universe, and their interaction with radiation. The discipline is an overlaps of astronomy and chemistry [3, 21].

Remote Sensing of the Earth and Planets is the acquisition of information about an object or phenomenon without making physical contact with the object and thus is in contrast to the on-site observation. The term is applied especially to acquiring information about the Earth. Remote sensing is used in numerous fields, including geography, land surveying and most Earth science disciplines (for example, hydrology, ecology, meteorology, oceanography, glaciology, geology); it also has a military, intelligence, commercial, economic, planning, and humanitarian applications, among others.

Planetary Science is the study of the assemblage of planets, moons, dwarf planets, comets, asteroids, and other bodies orbiting the Sun, as well as extrasolar planets. The Solar System has been relatively well-studied, initially through telescopes and then later by spacecrafts. This has provided a good overall understanding of the formation and evolution of the Sun's planetary system, although many new discoveries are still being made.

Planetary Geology, alternatively known as Astrogeology or Exogeology, is a planetary science discipline concerned with the geology of the celestial bodies such as the planets and their moons.

Exoplanetology, or Exoplanetary Science (Science for Extrasolar Planetary Systems), is an integrated field of space science dedicated to the search for and study of exoplanets (extrasolar planets). It employs an interdisciplinary approach that includes astronomy, astrophysics and space physics, cosmochemistry, space geology, geochemistry, space biology, and planetary science.

Astronautics (or Cosmonautics) are the theory and practise of travel beyond Earth's atmosphere into outer space. Spaceflight is one of its main applications and space science its overarching field. Actual astronautics is the science and engineering of spacefaring and spaceflight, a subset of Aerospace engineering (which includes atmospheric flight) [1, 2].

Orbital mechanics or Astrodynamics, are the application of ballistics and celestial mechanics to the practical problems concerning the motion of rockets and other spacecraft. The motion of these objects is usually calculated from Newton's laws of motion and law of universal gravitation. Orbital mechanics are a core discipline within space-mission design and control.

Space life sciences:

Bioastronautics is a specialty area of biological and astronautical research , which encompasses numerous aspects of biological, behavioral, and medical concern governing humans and other living organisms in a space flight environment; and includes the design of payloads, space habitats, and life-support systems. In short, it spans the study and support of life in space.

Space Medicine is the practice of medicine on astronauts in outer space whereas astronautical hygiene is the application of science and technology to the prevention or control of exposure to the hazards that may cause astronaut ill health. Both these sciences work together to ensure that astronauts work in a safe environment. The main objective is to discover how well and for how long people can survive the extreme conditions in space [1, 22].

Space Neuroscience is the scientific study of the central nervous system (CNS) functions during spaceflight. Living systems can integrate the inputs from the senses to navigate in their environment and to coordinate posture, locomotion, and eye movements. Gravity plays a fundamental role in controlling these functions. In weightlessness during spaceflight, integrating the sensory inputs and coordinating motor responses is harder to do because gravity is no longer sensed during free-fall [1].

Space Biology Research in Space Biology is aimed at addressing the basic questions regarding the extent to which gravity plays a role in growth, morphology, and function of cells in the space environment (Cell Biology), and from the early development of animals and plants to several life cycles (Developmental Biology). A more applied aspects of Space Biology research also include the biological effects of space radiation and radiation standards (Radiation Biology) and the production of cells for medically valuable proteins (Biotechnology) [1].

Space Botany or Astrobotany is an applied sub-discipline of botany that is the study of plants in space environments. It is a branch of astrobiology and botany. It has been a subject of study that plants may be grown in outer space typically in a weightless but pressurised controlled environment in specific space gardens. In the context of human spaceflight, they can be consumed as food and/or provide a refreshing atmosphere. Plants can metabolise carbon dioxide in the air to produce valuable oxygen, and can help control cabin humidity. Growing plants in space may provide a psychological benefit to human spaceflight crews.

Micro-g Environment Research. The term micro-g environment (also µg, often referred to by the term Microgravity) is more or less synonymous with the terms weightlessness. The most commonly known microgravity environment can be found aboard the International Space Station (ISS) which is located in low-earth orbit at an altitude of around 400km, orbiting Earth around 15 times per day in what is considered free fall [1].

Archaeoastronomy is the interdisciplinary or multidisciplinary study of how people in the past have understood the phenomena in the sky, how they used these phenomena and what role the sky played in their cultures.

Space Anthropology is a sub discipline of anthropology that looks at all human responses to and interactions with Space; from Archaeo-astrology to Xeno-anthropology (Exo-anthropology) we bring together all different areas of study surrounding humans and Space to better understand where we might be heading in the future.

Space Law is the body of law governing space-related activities, encompassing both international and domestic agreements, rules, and principles.[1] Parameters of space law include space exploration, liability for damage, weapons use, rescue efforts, environmental preservation, information sharing, new technologies, and ethics.

Space Technology is a technology developed by space science for use in astronautics, for purposes such as spaceflight or space exploration. Space technology includes spacecraft, satellites, space stations, and support infrastructure equipment, and procedures and space warfare. Space is such a novel environment that attempting to work in it requires new tools and techniques. Many common everyday services such as weather forecasting, remote sensing, satellite navigation systems, satellite television, and some long-distance communication systems critically rely on space infrastructure.

Of the sciences, astronomy and Earth science benefit from space technology [1]. New technologies originating with or accelerated by space-related endeavours are often subsequently exploited in other economic activities.

Space Navigation - today's spacecraft demand complex guidance and navigation control solutions that are safe, low risk, low cost and reliable.

Space navigators had to invent a new science of space navigation, using star sightings, precise timing, and radio communications. The great distance spacecraft had to travel called for even greater precision in timing and positioning than ever before.

A Deep-Space Positioning System (DPS) has been developed for this purpose. DPS is a single device that provides a spacecraft's position and velocity in interplanetary space.

Space Communications to and from space is a challenging endeavour [24-28]. Fortunately, we have the experience and expertise to get space data to the ground. The Space Communications and Navigation (SCaN) programme enables this data exchange, whether it's with cosmonauts aboard the International Space Station, rovers on Mars, or the missions to the Moon.

Man's greatest achievement in this regard is the Voyager spacecraft launched in 1977, that 44 years later is still communicating with Earth from more than 180 billion km away and has far outlived even the most optimistic projections of longevity.

Space Architecture is the theory and practice of designing and building inhabited environments in outer space.[1] The architectural approach to spacecraft design addresses the total built environment. It is mainly based on the field of engineering (especially aerospace engineering), but also involves diverse disciplines such as space physiology, psychology, and sociology. Like architecture on Earth, the attempt is to go beyond the component elements and systems and gain a broad understanding of the issues that affect design success.

Space Logistics is the theory and practice of driving space system design for operability and of managing the flow of material, services, and information needed throughout a space system lifecycle.

However, this definition in its larger sense includes terrestrial logistics in support of space travel, including any additional "design and development, acquisition, storage, movement, distribution, maintenance, evacuation, and disposition of space materiel", movement of people in space (both routine and for medical and other emergencies), and contracting and supplying any required support services for maintaining space travel.

Space Robotics - In the space community, any unmanned spacecraft can be called a robotic spacecraft. However, *Space Robots* are deemed more capable devices that can facilitate manipulation, assembling, or servicing functions in orbit as assistants to astronauts, or to extend the areas and abilities of exploration on remote planets as surrogates for human explorers.

Space Robotic Colonies are perhaps one of the most advanced and exciting areas of today's space robotic research and technology development. A robotic colony provides a permanent, continuous operational presence elsewhere in our solar system. Established and operated by robots, possibly visited by humans –

eventually perhaps populated by both—a colony is established, expanded and resupplied using resources from Earth, and native, *in situ* resources.

Space Colonization (also called Space Settlement or Extraterrestrial Colonization) is the hypothetical permanent habitation and exploitation of natural resources from outside planet Earth. As such it is a form of human presence in space, beyond human spaceflight or operating space outposts.

Planetary Habitability is the measure of a planet's or a natural satellite's potential to develop and maintain environments hospitable to life. Life may be generated directly on a planet or satellite endogenously or be transferred to it from another body, through a hypothetical process known as panspermia.

Space Manufacturing - In-Space Manufacturing (ISM) involves a comprehensive set of processes aimed at the production of manufactured goods in the space environment. ISM is also often used interchangeably with the term in-orbit manufacturing given that current production capabilities are limited to low Earth orbit.

There are some rationales supporting in-space manufacturing, f.e.: the space environment, in particular the effects of microgravity and vacuum, enable the research of and production of goods that could otherwise not be manufactured on Earth.

Space Materials Science studies microstructure formation during casting of technical alloys under diffusive and magnetically controlled convective conditions. The experimental results with parametric studies using numerical simulations can be used to optimise industrial casting processes [1].

Space Industry refers to economic activities related to manufacturing components that go into Earth's orbit or beyond, delivering them to those regions, and related services. Owing to the prominence of the satellite-related activities, some sources use the term Satellite Industry interchangeably with the term space industry. The term Space Business has also been used. A narrow definition encompasses only hardware providers (primarily related to launch vehicles and satellites). This definition does not exclude certain activities, such as Space Tourism. Thus, more broadly, space industry can be described as the companies involved in the Space Economy, and providing goods and services related to space. Space economy has been defined as "all public and private actors in developing and providing space-enabled products and services. It comprises a long value-added chaining, starting with research and development actors and manufacturers of Space Hardware and ending with the providers of space-enabled products and services to final users.

Yuri Gagarin and his cosmic heritage

One of the most inspiring figures in spaceflight history is Yuri Gagarin - a Russian pilot and cosmonaut, the first human in space.

Yuri Gagarin left behind an inspirational legacy, which still continues to motivate millions to this day. Yuri's Night is celebrated every year on his launch date, the 12th of April, worldwide to honour the milestone Yuri Gagarin set as being the first human in space and inspiring those who came after him. Yuri's Night was founded in 2001 (on occasion of the 40-th anniversary of first human spaceflight) and attracting thousands of celebrants each year.

NASA's Apollo 11, the first mission to put people on the moon, landed in July 1969, and the crew left behind a commemorative medallion - memorial plaque bearing Gagarin's name (Fig. 3).

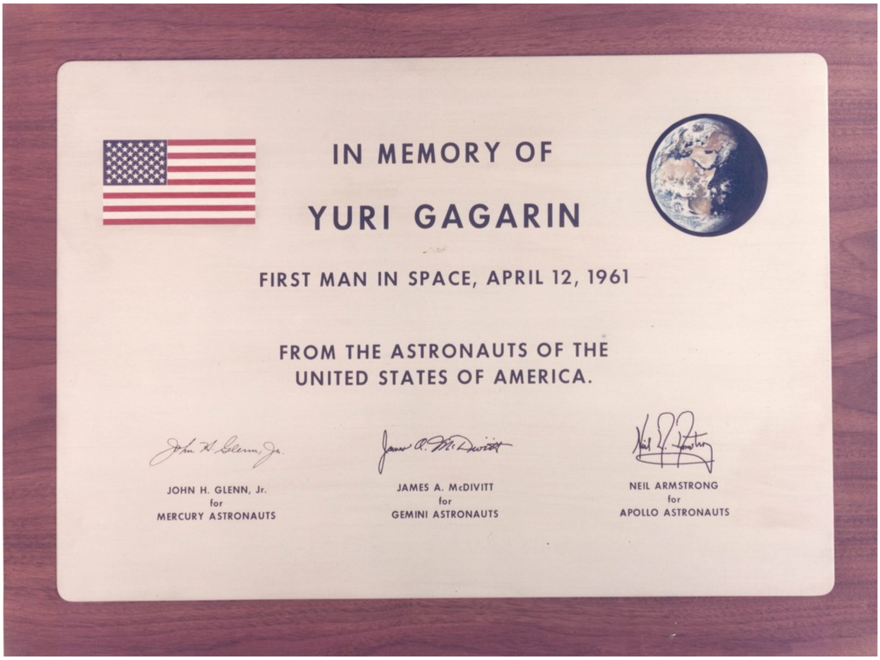


Fig. 3. A memorial plaque commemorating Yuri Gagarin as the first human in space.
Image source: Flickr.
The plaque was signed by astronauts of USA: Neil Armstrong (the first man landing on the Moon, July 20, 1969), for the astronauts of the Apollo spacecraft; James A. McDivitt - for astronauts from the Gemini spacecraft (1965-1966); and by John H. Glenn (the first American astronaut to orbit the Earth February 20, 1962) - for the crews of the Mercury spacecrafts.

The plaque was signed by astronauts of USA: Neil Armstrong (the first man landing on the Moon, July 20, 1969), for the astronauts of the Apollo spacecraft; James A. McDivitt – for astronauts from the Gemini spacecraft (1965–

1966); and by John H. Glenn (the first American astronaut to orbit the Earth February 20, 1962) – for the crews of the Mercury spacecraft (Fig. 3).

Over time, the U.S. and the Soviet Union began to work together in their spaceflight endeavours. The first joint U.S.-Soviet spaceflight was in 1975, called Apollo-Soyuz. Following that, NASA sent several space shuttle astronauts to Soviet/Russian space station Mir. The shuttle-Mir collaboration paved the way for NASA and the Russian space agency (RosCosmos) to become major partners in the International Space Station programme, which first launched modules in 1998 and continues research today.

Gagarin's importance in the Russian space programme continues. Crews using the Soyuz spacecraft participate in a number of prelaunch traditions prior to climbing on to the spacecraft, to follow in the footsteps of Gagarin's historic flight. Beyond that, Gagarin is often held up as an example of character and heroism to younger children in Russia.

The date of Gagarin's space flight, 12 April, has been commemorated. Since 1962, it has been celebrated in the USSR and most of its former territories as Cosmonautics Day.

There are statues of Gagarin and monuments to him in a number of places around the world (including the United States and Great Britain).

In 2011 (on occasion of the 50-th anniversary of first human spaceflight), a statue of Gagarin was unveiled at the Admiralty Arch in The Mall in London, opposite the permanent sculpture of James Cook. In 2013, the statue was moved to a permanent location outside the Royal Observatory, Greenwich (Fig. 4).

Fig. 4. Yuri Gagarin statue at the Royal Greenwich Observatory in London.
https://www.rmg.co.uk/royal-observatory/attractions/yuri-gagarin-statue

In 2012, a new statue of Gagarin was unveiled at the site of NASA's original spaceflight headquarters in Houston. In April 2018, a bust of Gagarin erected on the street in Belgrade, Serbia, that bears his name.

Numerous streets, boulevards and squares around the world are also named after Gagarin.

References

1. Serafimov, K., Space Research in Bulgaria, Sofia, BAS Publishing House, 1979, 221 p.
2. Mishev, D., History of Space Research in Bulgaria, Sofia, BAS Publishing House, 2004, 234 p.
3. Velinov, P. I. Y., G. Nestorov, L. I. Dorman, Cosmic Ray Influence on the Ionosphere and on Radiowave Propagation, Sofia, BAS Publishers, 1974, ISBN:4897, 314 p.
4. Velinov, P. I. Y., Nestorov G., Effect of solar flares on the low ionosphere. C. R. Acad. Bulg. Sci., 1967, 20, 293–96.
5. Velinov, P. I. Y., A formula derivation for electron production rate in theionosphere under the influence of cosmic rays, C. R. Acad. Bulg. Sci., 1966, 19, 109–12.
6. Velinov, P. I. Y, Low ionosphere ionization by cosmic rays, C. R. Acad.Bulg. Sci., 1966, 19, 281–84.
7. Velinov, P. I. Y., Contribution of cosmic rays to the ionization of the lowerionosphere, C. R. Acad. Bulg. Sci., 1966, 19, 889–92.
8. Nestorov, G., Velinov P. I. Y., Effect of solar cosmic rays on the lower ionosphere. C. R. Acad. Bulg. Sci., 1966, 19, 1011–14.
9. Velinov, P. I. Y., On electron production rates in the polar cap ionosphere due to solar cosmic rays. C. R. Acad. Bulg. Sci., 1967, 20, 1275–78.
10. Velinov, P. I. Y., On the planetary distribution and energy ballance of the Cosmic Ray (CR layer in lower ionosphere, Proc. Geophys. Inst., 1967, 11, 87–102, BAS Publishers, Sofia.
11. Velinov, P. I. Y., Some results of the rate of electron production in the cosmic layer of low ionosphere. C. R. Acad. Bulg. Sci., 1967, 20, 1141–44.
12. Velinov, P. I. Y., On ionization of the ionospheric D-region by galactic andsolar cosmic rays, J. Atmos. Terr. Phys.,1968, 30, 1891–1905.
13. Velinov, P. I. Y., Solar cosmic ray ionization in the lower ionosphere, J.Atmos. Terr. Phys., 1970, 32, 139–47.
14. Velinov, P. I. Y., On variations of the cosmic ray (CR) layer in the lower ionosphere, J. Atmos. Terr. Phys., 1971, 33, 429–36.
15. Velinov, P. I. Y., Cosmic ray ionization rates in the planetary atmospheres,J. Atmos. Terr. Phys., 1974, 36, 359–62.
16. Tsagouri, I., Belehaki A., Viljanen A. et al., Progress in space weather modeling in an operational environment, Journal of Space Weather and Space Climate, 2013, 3, A17, 1–72. DOI:10.1051/swsc/2013037
17. Velinov, P. I. Y. Cosmic Ray Effects on the Lower Ionosphere, PhD Thesis, Geophys. Inst. BAS, Sofia, 1969, 161 p.
18. Velinov, P. I. Y. Study on Geophysical and Astrophysical Aspects of the Problem of Substance Ionization by Energetic Particles, DSc Thesis, Geophys, Inst. BAS, Sofia, 1975, 280 p.

19. Usoskin, I., and L. Desorgher, Ionization of the Earth's atmosphere by solar and galactic cosmic rays. Acta Geophysica, 2009, 57, 1/March, VERSITA, Warsaw, 88–101.
20. Usoskin, I., and L. Desorgher, Solar and Galactic Cosmic Rays in the Earth's Atmosphere. In the Book: Developing the scientific basis for monitoring, modelling and predicting Space Weather, COST 724 final report (eds. J. Lilensten e), COST Office, Luxemburg, 2008, ISBN:978-92-898-0044-0, 124–32.
21. Gronoff, G., Mertens C., Lilensten J., Desorgher L., Flueckiger et al., Ionization processes in the atmosphere of Titan. III - Ionization by high-Z cosmic rays, Astronomy and Astrophysics (A&A), 2011, 529, 5, DOI:10.1051/0004-6361/201015675, A143-A146.
22. Velinov, P. I. Y. On the protection from cosmic rays and internal radiation belt in the space flights. In: Space exploration and applications. Proc. of First UN Conf. on Exploration and Peaceful Uses of Outer Space, Vienna, Austria, 14-27 August 1968, A/CONF. Report 34/IV, B.4, UN Publishers, New York, https://digitallibrary.un.org/record/files/A_7285-EN, 1968, 1–21.
23. Mishev, A., and Velinov P. I. Y., Ionization effect in the Earth's atmosphere during the sequence of October–November 2003 Halloween GLE events. J. Atmos. Solar-Terr. Phys., 2020, 211, 105484, Elsevier, DOI:10.1016/j.jastp.2020. 105484, 1–7.
24. Velinov, P. I. Y., Altitude variations of the frequences and electron density by reflexion from the D-region. C. R. Acad. Bulg. Sci., 1965, 18, 1115–18.
25. Nestorov, G., and P. I. Y. Velinov, Electron concentration variations by long wave reflection from D-region. C. R. Acad. Bulg. Sci., 1965, 18, 1111–14.
26. Velinov, P. I. Y., Electromagnetic Field Variations of Long Radiowave Propagation in the Quiet and Disturbed Ionosphere, MSc Thesis, Geophys. Inst. BAS, Sofia, 1965, 95 p.
27. Shinev, C., and Velinov P. I. Y. On the Theory of the asymmetrical antennae. Proc. Techn. Univ., Sofia, 1965, XVIII, 1, 63–72.
28. Shinev, C., and Velinov P. I. Y. On the error correction of linse antenna. Proc. Techn. Univ., Sofia, 1965, XVIII, 1, 73–80.

URL addresses:

https://www.un.org/en/observances/human-spaceflight-day
https://www.esa.int/Science_Exploration/Space_Science/About_Space_Science
https://en.wikipedia.org/wiki/Outline_of_space_science
https://en.wikipedia.org/wiki/Astronomy#Scientific_revolution
https://www.nasa.gov/pdf/55388main_05 Space Science.pdf

ПРОГРЕСЪТ В КОСМИЧЕСКИТЕ НАУКИ И ТЕХНОЛОГИИ ВЪВ ВРЪЗКА С 60-ТА ГОДИНИНА НА ПЪРВИЯ ПОЛЕТ НА ЧОВЕКА В КОСМОСА

П. Велинов

Резюме

На 12 април 1961 г. Юрий Гагарин възвести идването на нова космическа ера. Бързият прогрес в различните космически науки и технологии започна веднага след първия полет на човека в Космоса. Тогава се появяват принципно нови науки и технологии. В момента космическата наука обхваща широк спектър от дисциплини. Тук е представен един общ преглед и класификация на космическите науки:

Астрономия и космическа астрономия, космология, астрофизика, космическа физика, слънчево-земна физика, аерономия, слънчева физика, хелиосферна физика, физика на космическите лъчи, космическо време и космически климат (климата на системата Земя-Космоса), космическа дозиметрия, кос-мическа химия (или космохимия), дистанционни изследвания на Земята и планетите, планетни науки, планетна геология, астрогеология или екзогеология, екзопланетология или екзопланетна наука (наука за извънслънчевите планетни системи), интелигентен живот във Вселената; астронавтика (или космонавтика), орбитална механика или астродинамика, космически науки за живота: биоастронавтика, космическа медицина, космическа неврология, космическа биология, биотехнологии, космическа ботаника или астроботаника, микрогравитационни изследвания при космическите полети; археоастрономия, космическа антропология, ксеноантропология (екзоантропология), космическо право, космически технологии, космическа навигация, космически комуникации, космическа архитектура, космическа логистика, космическа роботика, космически роботизирани колонии, космическа колонизация (наричана още космическо заселване или извънземна колонизация), планетна обитаемост, космическо производство, космически материали, спътникова индустрия, космически бизнес, космически туризъм, космически хардуер, космическа индустрия, космическа екология.

С помощта на тези прогресиращи космически науки човечеството започна уверено да изследва Космоса. Но тези изследвания доведоха и до множество нови технологии и приложения за подобряване на живота на хората.

Накрая отново се споменава за Юрий Гагарин и неговото космическо наследство. Той остави след себе си вдъхновяващо наследство, което и до днес все още продължава да мотивира милиони хора по целия свят.