Aerospace
Research
in Bulgaria

36

Bulgarian Academy of Sciences

Space Research and Technology Institute

2024

Aerospace Research in Bulgaria

36
Sofia, 2024

C o n t e n t s

С Ъ Д Ъ Р Ж А Н И Е

SPECTRA OF ANOMALOUS COSMIC RAYS IN THE ATMOSPHERE. SINGLY IONIZED AND MULTIPLY CHARGED ACR COMPONENTS

Peter Velinov, Simeon Asenovski, Lachezar Mateev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: PeterIYVelinov@gmail.com; asenovski@gmail.com

Abstract

The Anomalous Cosmic Ray (ACR) spectra and intensity in the middle atmosphere are determined using the CORSIMA (COsmic Ray Spectra and Intensity in Middle Atmosphere) model. ACR spectra are presented for various atmospheric altitudes within the range of 40–50 km, with the lower boundary of the ionosphere at approximately 50 km. Experimental satellite measurements are utilized for the main ACR constituents Helium and Oxygen nuclei. It is found that the influence of ACRs extends to the polar cap regions above 65°–70° geomagnetic latitude, and certain ACR ionization rate values in these regions are comparable to Galactic Cosmic Ray (GCR) ionization rates. Our studies also consider Multiply Charged Anomalous Cosmic Rays (MCACRs), which exhibit similar differential spectra to the singly ionized ACR component4.

Introduction

In 1972 and 1973 several groups observed during solar quiet times enhanced fluxes of helium, oxygen, and nitrogen at energies of ≈10 MeV/nuc (N, O) and ≈ <50 MeV/nuc (He), respectively. These flux increases showed peculiar elemental abundances and energy spectra, e.g. a C/O ratio ≤ 0.1 at ≈10 MeV/nuc, significantly different from the abundances of Solar Cosmic Rays (SCR) and Galactic Cosmic Rays (GCR). Since then, this "Anomalous" Cosmic Ray (ACR) component has been studied extensively and several elements have been found (He, N, O, Ne, Ar, and, to a lesser extent, H and C) whose energy spectra show anomalous increases above the quiet time solar and galactic energetic particle spectrum. There have been several models proposed to explain the ACR component. The presently most plausible theory for the origin of ACR ions identifies neutral interstellar gas as the source material. After penetration into the inner heliosphere, the neutral particles are ionized by solar UV radiation and by charge exchange reactions with the solar wind protons. After ionization, the now singly charged ions are picked up by the interplanetary magnetic field and are convected with the solar wind to the outer solar system.

There 0the ions are accelerated to high energies, possibly at the solar wind termination shock, and then propagate back into the inner heliosphere. A unique prediction of this model is that ACR ions should be singly ionized. Meanwhile, several predictions of this model have been verified, e.g. low energy pickup ions have been detected and the single charge of ACR ions in the energy range of ≈10 MeV/nucleon has been observed [1, 2, 3]. Later, Multiply Charged Anomalous Cosmic Rays (MCACRs) were discovered.

In this paper, the spectra and intensities of ACRs will be calculated using our CORSIMA (COsmic Ray Spectra and Intensity in Middle Atmosphere) model. The energy spectrum of ACRs in the Earth's atmosphere at altitudes ranging from 40 to 50 km is an area that has not been extensively studied, primarily because ACRs are typically measured in space. However, this altitude range is significant because it corresponds to the lower boundary of the ionosphere, which is located around 50 km above the Earth's surface [4].

While the energy spectrum of ACRs remains relatively unchanged at higher altitudes where the atmosphere has minimal influence, as ACRs penetrate deeper into the atmosphere, they interact with air molecules, leading to energy loss and the possible production of secondary particles [5]. Consequently, the energy spectrum of ACRs undergoes modification as they traverse the Earth's atmosphere.

At higher altitudes, ACRs experience minimal effects from the Earth's atmosphere, and their energy spectrum closely resembles that observed in space. However, as ACRs descend and interact with air molecules in the atmosphere, their energy can be attenuated, and secondary particle production may occur [5]. These interactions with the atmosphere contribute to a modification of the energy spectrum of ACRs. The specific shape of the energy spectrum of ACRs in the Earth's atmosphere within the 40-50 km altitude range is influenced by various factors, including the initial energy and composition of the ACRs, as well as the atmospheric conditions at these altitudes. It is expected that the interactions with the atmosphere will lead to a distinct modification in the shape and intensity of the energy spectrum compared to observations made in space [6]. There is a significant difference here from galactic CRs [7, 8].

The primary objective of the present study is to determine the energy spectra of ACRs in the middle atmosphere around the lower boundary of the ionosphere, specifically within the altitude range of 40-50 km. This research aims to shed light on the energy distribution of ACRs at these specific altitudes, considering the interactions they undergo with the Earth's atmosphere. By investigating the energy spectra of ACRs in this region, valuable insights can be gained regarding the behavior of ACRs as they penetrate the Earth's atmosphere and interact with air molecules. This information is crucial for understanding the complex dynamics of ACRs in the middle atmosphere and their implications for the ionization processes and atmospheric conditions within this altitude range.

Model approximations

We introduce five main characteristic energy intervals in the approximation of ionization losses (MeV.g^{-1}.cm^2) according to the Bohr-Bethe-Bloch function using experimental data [6–8]. This approximation for protons and singe charge particles ($Z = 1$) has the form [9]:

$$(1) \quad -\frac{1}{\rho}\frac{dE}{dh} = \begin{cases} 2.57 \times 10^3 E^{0.5} & \text{if } kT \le E \le 0.15 \text{ MeV} & \text{interval 1} \\ 231 E^{-0.77} & \text{if } 0.15 \le E \le 200 \text{ MeV} & \text{interval 2} \\ 68 E^{-0.53} & \text{if } 200 \le E \le 850 \text{ MeV} & \text{interval 3} \\ 1.91 & \text{if } 850 \le E \le 5 \times 10^3 \text{ MeV} & \text{interval 4} \\ 0.66 E^{0.123} & \text{if } 5 \times 10^3 \le E \le 5 \times 10^6 \text{ MeV} & \text{interval 5} \end{cases}$$

E is the kinetic energy of the penetrating particles.

By introducing multiple characteristic energy intervals in our model, we enhance the accuracy of the obtained results compared to previous approximations with fewer intervals [10, 11].

Our model is designed to analyze the contributions of different types of cosmic rays (CRs), including galactic CRs, solar CRs, and anomalous CRs (ACRs), to the ionization in the ionosphere-middle atmosphere. Each submodel within our framework focuses on evaluating the specific contributions of these CR types and considers the distinct characteristic energy intervals in the total ionization process. To investigate the impact of random differential spectrum energy intervals on ionization in the middle atmosphere, we utilize satellite measurements of differential spectra, with a particular emphasis on ACR spectra in this study. By decomposing the ACR spectra into different groups of ACR nuclei and characteristic energy intervals, we gain insights into their properties and examine their effects on the ionization losses function boundaries.

Our newly developed code, CORSIMA (**CO**smic **R**ay **S**pectra and **I**ntensity in **M**iddle **A**tmosphere), builds upon the results and advancements of our previous model, CORIMIA (**CO**smic **R**ay **I**onization **M**odel for **I**onosphere and **A**tmosphere) [6, 12]. By leveraging the capabilities of CORSIMA and incorporating the refined analysis of energy intervals, we aim to provide a comprehensive and improved understanding of ACR spectra and their influence on ionization in the middle atmosphere.

Model description for single ionized ACRs

The submodel for the calculation of the ACR ionization rate profiles has different properties in comparison with GCR or SCR submodels. In the presented calculation ACR constituents are singly charged. That is why we don't introduce the charge decrease interval [12] but we consider the influence of atomic weight A.

When considering the penetration of ACRs into the atmosphere, we calculate the electron production rate within two intervals that pertain to the low-energy range of ionization losses (MeV.g^{-1}.cm^2) following the Bohr-Bethe-Bloch formula. These intervals, which we consider in our model, are as follows [6]:

$$(2) \qquad -\frac{1}{\rho}\frac{dE}{dh} = \begin{cases} \dfrac{2.57 \times 10^3}{A} E^{0.5} & \text{if } kT \leq E \leq 0.15 \text{ MeV/nucl} \qquad \text{interval 1} \\[2ex] \dfrac{231}{A} E^{-0.77} & \text{if } 0.15 \leq E \leq 200 \text{ MeV/nucl} \qquad \text{interval 2} \end{cases}$$

where A is the atomic weight of ACR particles, $\rho(h)$ (g.cm^{-3}) is the neutral density of the Earth's atmosphere. E is the kinetic energy of ACRs in MeV/nucl.

These intervals described the part of the ionization losses function where ACR spectra are acting. Based on this statement we derived the following expression for the ACR ionization rate submodel:

$$(3) \qquad q(h) = \frac{\rho(h)}{Q}\Big\{ 2.57 \times 10^3 \int_{E_{\min}}^{0.15} D(E)\{[E_1(h)]^{0.5} +$$

$$2.57 \times 10^3 \int_{0.15}^{E_{0.15;2}(h)} D(E)[E_{21}(h)]^{1/2}\, dE +$$

$$231 \int_{E_{0.15;2}(h)}^{200} D(E)[E_2(h)]^{-.077}\, dE \Big\}$$

where $Q = 35$ eV is the energy necessary for the formation of one electron-proton pair. $E_1(h)$, $E_2(h)$ and $E_{21}(h)$ are corresponding interval's energy decrease laws. $D(E)$ is the differential spectrum in (cm^{-2}.s^{-1}.sr^{-1}.MeV^{-1}). E_{min} is energy cut off which is determined in (1). $E_{0.15;\ 2}(h)$ is the initial energy of particles (before entering of spectrum in the atmosphere), which have energy $E(h) = 0.15$ MeV at altitude h (km).

$$(4) \qquad E_{0.15;2}(h) = \left(408.87\tilde{h} + 0.15^{1.77}\right)^{0.56}$$

The contribution of the ACR differential spectrum to electron production is determined by the geomagnetic and atmospheric cut-offs. Its lower boundary for the given point in the Earth's atmosphere is calculated with equation [6, 9, 12]:

$$(5) \qquad E_{\min} = \max\{E_R(\lambda_m), E_A(\tilde{h})\}$$

8

$E_R(\lambda_m)$ is the geomagnetic cut-off in GeV which depends on geomagnetic latitude λ_m as follows [6,9]:

$$(6) \qquad E_R(\lambda_m) = \left(14.9 \left(Cos\left[\frac{\pi\lambda_m}{180}\right]^4 \right)^2 + 0.88 \right)^{1/2} - 0.938$$

where 0.938 is the rest energy of the proton. $E_A(\tilde{h})$ is an atmospheric cut-off which depends on the traveling substance path $\tilde{h}$ (g.cm^{-2}) [4, 5]. For the first interval in (1) it has the form:

$$(7) \qquad E_{A1}(h) = \left((kT)^{0.5} + 1285\tilde{h}\right)^2$$

The kinetic energy transformations described in equation (3) are as follows: $E_1(h)$ represents the kinetic energy decrease within interval 1, $E_{21}(h)$ represents the kinetic energy decrease when crossing the boundary between interval 1 and interval 2, and $E_2(h)$ represents the kinetic energy decrease within interval 2. These transformations are applicable for the specific height h in the Earth's atmosphere.

Computer code and mathematical program

We have implemented a computer code for the CORSIMA model using advanced computational techniques [13, 14]. The code utilizes numerical methods to solve integration problems arising from the mathematical expressions involved in the model. This operational model allows for interactive computation, where users can input the required data and obtain computational results for different altitudes and specified geomagnetic latitudes. The code has been designed to be user-friendly, providing an intuitive interface for users to easily navigate and interact with the model.

Results

Evidence that the anomalous component is singly ionized is given in [15]. Figures 1 and 2 present the spectra of the main ACR species, namely the singly charged particles: Helium (Fig. 1) and Oxygen (Fig. 2). The spectra are specifically calculated for a geomagnetic latitude of $\lambda_m = 90^\circ$. In the lower portion of the profiles, the helium spectra are primarily affected by the atmospheric cut-offs. Below 40 km, the helium spectra decrease due to the influence of the atmospheric cut-off (Fig. 1). Notably, the helium and oxygen spectra exhibit the most prominent intensities among the ACR species.

These Figures represent the variation in the intensity of the ACR spectra with altitude, specifically between 50 and 40 km in the Earth's atmosphere. Figures 1 and 2 clearly demonstrate the difference in intensity between these two altitudes, with a

higher intensity observed at 50 km compared to 40 km. This change in the spectra provides significant insights into the dynamics of the Earth's upper atmosphere and the interactions between cosmic rays and atmospheric particles at different altitudes.
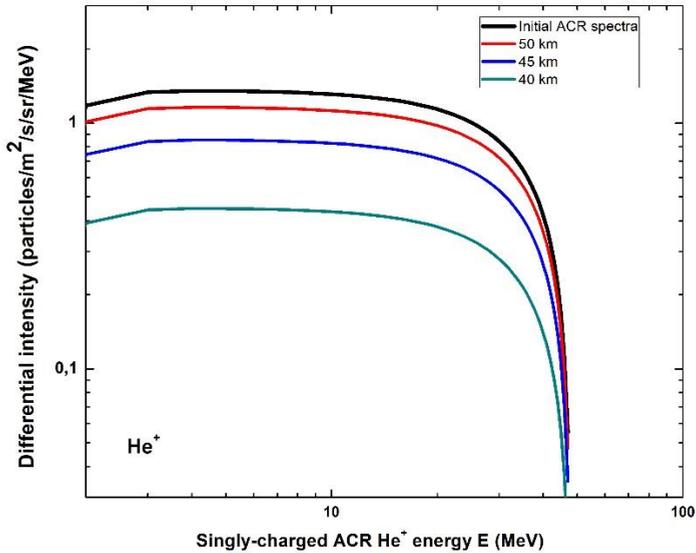
Fig. 1. Energy Spectrum of Singly-charged ACR He+ Particles at Different Altitudes
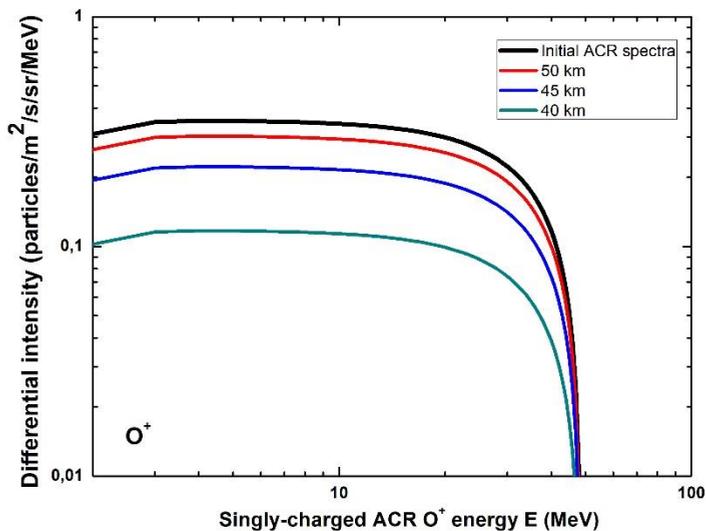
Fig. 2. Energy Spectrum of Singly-charged ACR O+ Particles at Different Altitudes

Model description for multiply charged ACRs

The submodel for the calculation of the ACR ionization rate profiles has different properties in comparison with GCR or SCR submodels. In the presented calculation ACR constituents are multiply charged, i.e. we have the case of MCACRs [16]. That is why we introduce the charge decrease interval [12]. We consider also the influence of atomic weight A.

When considering the penetration of MCACRs into the atmosphere, we calculate the electron production rate within three intervals that pertain to the low-energy range of ionization losses (MeV.g^{-1}.cm^2) following the Bohr-Bethe-Bloch formula. Unlike the case of single ionized particles, now we include the new interval dependencies. In this way, the expression of the ionization loss function becomes more general than the case $Z = 1$ [10]

$$(8) \quad -\frac{1}{\rho}\frac{dE}{dh} = \begin{cases} 2.57 \cdot 10^3 E^{0.5} & \text{if } kT \leq E \leq 0.15 \text{ MeV/n,} & \text{interval 1} \\ 1540 E^{0.23} & \text{if } 0.15 \leq E \leq E_a = 0.15 Z^2 \text{ MeV/n,} & \text{interval 2} \\ 231 \cdot Z^2 E^{-0.77} & \text{if } E_a \leq E \leq 200 \text{ MeV/n,} & \text{interval 2} \end{cases}$$

The electron production rate expression is characterized by the following main terms:

$$q(h) = \frac{\rho(h)}{Q}\left\{ 2.57 \cdot 10^3 \int_{E_{min}}^{0.15} D(E)[E_1(h)]^{0.5} dE \right.$$

$$(9) \quad + 2.57 \cdot 10^3 \int_{0.15}^{E_{0.15;2}(h)} D(E)\,[E_{21}(h)]^{0.5} dE + 1.54 \cdot 10^3 \int_{0.15;2(h)}^{E_a} D(E)\,[E_2(h)]^{0.23} dE$$

$$\left. + 1.54 \cdot 10^3 \int_{E_a}^{E_{a;3}(h)} D(E)\,[E_{32}(h)]^{0.23} dE + 231 \cdot Z^2 \int_{E_{a;3}}^{200} D(E)\,[E_3(h)]^{-0.77} dE \right\}$$

where $Q = 35$ eV is the energy necessary to form one electron-proton pair [xxx]; $\rho(h)$ is the atmospheric density at height h; $E_1(h)$, $E_2(h)$, $E_3(h)$, $E_{21}(h)$ and $E_{32}(h)$ are the corresponding interval energy decrease laws. $D(E)$ is the differential spectrum in (cm^{-2}.s^{-1}. st^{-1}.MeV^{-1}). E_{min} is the energy cut-off. $E_{0.15;2}(h)$ and $E_{a;3}$ are the initial energies of particles (before entering the spectrum in the atmosphere), which have energy $E(h) = 0.15$ and $E(h) = E_a$ (MeV) at altitude h (km) respectively [12]. We present now the initial energies of the interval boundaries which are effective for the electron production rate by MCACR expressed in formula (10). It depends namely on the depth of penetration in the atmosphere. The initial energy for $E_{min} = kT$, when it is situated in interval 2 from (9) is the following:

$$(10) \quad E_{A2}(h) = \left(0.15^{1.77} + 408.87\tilde{h} - 0.318(0.15^{0.5} - (kT)^{0.5}) \right)^{0.56}$$

This initial energy is situated in interval 1 from (9), it is given by formula (8).

The initial energy of the boundary E_a before the spectrum penetration in the atmosphere is given by formula (12):

$$(11) \qquad E_{a,3}(h) = \left(200^{1.53} + \frac{104.04 Z^2}{A} \tilde{h} - 0.25(200^{1.77} - E_a^{1.77}) \right)^{0.65}$$

Conclusion

In this study, we have investigated the spectra of ACRs and MCACRs in the polar cap regions of the Earth's atmosphere at an altitude of 40-50 km. Our analysis focused on the main ACR species, including Helium and Oxygen. We utilized the CORSIMA mathematical program to determine the ACR spectra, taking into account the processes of ionization and possible scattering.

At an altitude of 40-50 km in the Earth's atmosphere, the primary physical process that affects the ACR spectrum is ionization. As ACR particles collide with atmospheric molecules, they can ionize them and produce secondary particles such as electrons and other ions. This process leads to energy loss and the redistribution of ACR particles, thereby modifying their energy spectrum.

Scattering is also a possible process that can affect the ACR spectrum at this altitude range, but its contribution is likely to be smaller compared to ionization. Adiabatic cooling and nuclear interactions are not expected to have a significant impact on the ACR spectrum in this altitude range.

Our results demonstrate that the ACR impact is limited to the polar cap region above a geomagnetic latitude of $\lambda_m = 62^\circ$-63°. The obtained ACR spectra for different species show distinct characteristics, with variations in intensity and shape.

The mentioned models CORSIMA and CORIMIA, as also CORSIKA can be applied in space physics for the complex study of the solar-terrestrial connections and space weather [17–21].

References

1. Giacalone, J., Fahr, H., Fichtner, H., Florinski, V., Heber, B. et al. (2022)Anomalous cosmic rays and heliospheric energetic particles, Space Sci. Rev., 218 (4), article id.22, Doi:10.1007/s11214-022-00890-7
2. Cummings, A. C., Stone, E. C. (2007) Composition of Anomalous Cosmic Rays, Space Sci. Rev., 130 (1–4), 389–399. Doi:10.1007/s11214-007-9161-y.
3. Cummings, A.C., E.C. Stone, C.D. Steenberg (2002) Composition of anomalous cosmic rays and other heliospheric ions, Astrophys. J. 578, 194–210. Doi:10.1086/342427
4. Velinov, P. I. Y., G. Nestorov, L. Dorman. (1974) Cosmic Ray Influence On The Ionosphere And On Radiowave Propagation. Sofia, BAS Publ. House, 314 p., ISBN:4897
5. Dorman, L. (2004) Cosmic Rays in the Earth's Atmosphere and Underground, Dordrecht, Kluwer Academic Publishers. Doi:10.1007/978-1-4020-2113-8

6. Velinov, P. I. Y., Asenovski, S., Mateev L. (2012) Ionization of Anomalous Cosmic Rays in ionosphere and middle atmosphere simulated by CORIMIA code. C. R. Acad. Bulg. Sci., 65 (9), 1261–1268.

7. Velinov, P. I. Y. (1966) An expression for ionospheric electron production rate by cosmic rays. C. R. Acad. Bulg. Sci., 19 (2), 109–112.

8. Velinov, P. I. Y. (1968) On ionization of the ionospheric D-region by galactic and solar cosmic rays. J. Atmos. Terr. Phys., 30 (11), 1891–1905.

9. Velinov, P.I.Y., L. Mateev (2007) Ionization model for cosmic ray protons in ionosphere and atmosphere with 5 interval approximation of the ionization losses function. C. R. Acad. Bulg. Sci., 60 (8), 839–844.

10. Usoskin, I. G., L. Desorgher (2009) Ionization of the Earth's atmosphere by solar and galactic cosmic rays. Acta Geophysica, 57 (1), 88–101.

11. Usoskin, I. G., L. Desorgher (2008) 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, A. Belehaki, M. Messerotti, R. Vainio, J. Watermann, S. Poedts), COST Office, Luxemburg, 2008, ISBN:978-92-898-0044-0, pp. 124–132.

12. Velinov, P. I. Y., Asenovski, S., Mateev, L. (2012) Improved Cosmic Ray Ionization Model for the Ionosphere and Atmosphere (CORIMIA) with account of 6 characteristic intervals. C. R. Acad. Bulg. Sci., 65 (8), 1137–1144.

13. Wolfram Research Inc. (2008), Mathematica: Modern Technical Computing, Version 7.0, Champaign, Illinois. https://www.wolfram.com/ mathematica/

14. Fox William P. (2011) Mathematical Modeling with Maple, Version 14, (A comprehensive e-book), ISBN: 9780495109419, Publisher: Brooks Cole

15. Cummings, A. C., E. C. Stone, W. R. Webber (1984) Evidence that the anomalous component is singly ionized. Astrophysical Journal (Letters), 287, 99–103.

16. Mewaldt, R. A., R. S. Selesnick, J. R. Cummings, E. C. Stone (1996) Evidence for Multiply Charged Anomalous Cosmic Rays. The Astrophys. J., 466, L43–L46.

17. Mishev, A., P. I. Y. Velinov (2011) Normalized ionization yield function for various nuclei obtained with full Monte Carlo simulations. Adv. Space Res., 48, 19–24.

18. Eroshenko, E. (2010) Relationships between Neutron Fluxes and Rain Flows. Adv. Space Res., 46 (1), 637–641. https://doi.org/10.1016/j.asr.2010.04.016

19. Gronoff, G., Mertens C., Lilensten J., Desorgher L., Flueckiger E. et al. (2011) Ionization processes in the atmosphere of Titan. III - Ionization by high-Z cosmic rays. Astron. & Astrophys. (A&A), 529 (5), A143-A146. Doi:10.1051/0004-6361/20101567

20. Tsagouri, I., Belehaki A. (2013) Progress in space weather modeling in an operational environment, J. Space Weath. & Space Clim., 3, A17, 1–72. DoI:http://dx.doi.org/10.1051/swsc/2013037

21. Velinov, P. I. Y. I., Asenovski, S., Mateev, L. (2023) Cosmic Ray Spectrum and Intensity in Middle Atmosphere (CORSIMA) Model. Use and Application for Solar Cosmic Rays. Aerospace Res. Bulg., 35, 5–15.

22. Tomova, D., Velinov P. I. Y., Tassev, Y. (2017) Comparison between extreme solar activity during periods March 15-17, 2015 and September 4-10, 2017 at different phases of solar cycle 24. Aerospace Res. Bulg., 29, 3–29.

23. Velinov, P. I. Y. (2016) Extended categorisation of solar energetic particle events rising to ground level enhancements of cosmic rays. Aerospace Res. Bulg., 28, 3–20.

24. Mishev, A., Velinov, P. I. Y. I. (2024) Altitude profiles of ion production and Regener-Pfotzer region ionization during peculiar GLE 71 on May 17, Bulg. Astron. J., 40, 36–45, Institute of Astronomy and Rozhen NAO, 2024, ISSN:1314-5592

25. Velinov, P. I. Y. (2014) New Analitical Aproach for Cosmic Ray Ionization Modeling in Planetary Environments by Using the Yield Functions. Aerospace Res. Bulg., 26, 266–290.

26. Velinov, P. I. Y., Asenovski, S., Mateev L., Vashenyk E., Mishev A. (2013) Investigation of Middle Atmosphere Ionization During GLE 70 Event from December 2006 by Means of CORIMIA Model and Normalized CR Spectra. Aerospace Res. Bulg., 25, 62–69.

27. Abunina, M., Abunin A., Belov A., Gaidash S. (2013) Geoeffectivity of Solar Coronal Holes with Different Magnetic Field Polarity.. Aerospace Res. Bulg., 25, 70–77.

28. Velinov, P. I. Y., Mateev L., Ruder H. (2008) Atmospheric Cut-offs in the Generalized Model of Ionization Profiles Due to the Cosmic Ray Charged Particles in Planetary Ionospheres and Atmospheres with 5 Energy Interval Approximation of the Ionization Losses Function. Aerospace Res. Bulg., 22, 24–36.

СПЕКТРЫ АНОМАЛЬНЫХ КОСМИЧЕСКИХ ЛУЧЕЙ В АТМОСФЕРЕ. ОДНОИОНИЗОВАННЫЕ И МНОГОЗАРЯДНЫЕ КОМПОНЕНТЫ АКЛ

П. Велинов, С. Асеновский, Л. Матеев

Абстракт

Спектры и интенсивности Аномальных Космических Лучей (АКЛ) в средней атмосфере определяются с использованием модели CORSIMA (COsmic Ray Spectra and Intensity in Middle Atmosphere). Спектры АКЛ представлены для различных высот атмосферы в диапазоне 40–50 км, с нижней границей ионосферы примерно на 50 км. Экспериментальные спутниковые измерения используются для определения основных компонентов ACR: ядер гелия и кислорода. Установлено, что влияние АКЛ распространяется на области полярной шапки выше 65°-70° геомагнитной широты, а некоторые значения скорости ионизации АКЛ в этих регионах сравнимы со скоростями ионизации Галактических Космических Лучей (ГКЛ). В наших исследованиях также рассматриваются многозарядные аномальные космические лучи (MCACR), которые имеют дифференциальные спектры, аналогичные одноионизованному компоненту АКЛ.

DEVELOPING SITUATIONAL CONDITIONS AND PROGRAM CODES FOR PARALLEL SITUATIONAL ANALYSIS SOLVER BASED ON CONIC UMBRA/SUNLIT MODELS

Atanas Atanassov

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: At_M_Atanassov@yahoo.com

Abstract

The search for optimal time intervals for satellite operations performance is based on verifying specific conditions of a geometric or physical nature. Several conditions are combined in a situational problem. To be able to solve different situational problems, it is necessary to develop program codes for verifying different situational conditions in advance. Two situational conditions for determining if the satellite is in the sunlit zone or the umbra based on the conical models of the earth's shadow are presented. Necessary (but not sufficient) conditions are introduced to locate the satellite in the umbra or the sunlit zone, respectively. These conditions simplify the calculations and increase the computational efficiency for a large part of the satellite orbit. Problems with one or more conditions are solved by a developed parallel solver for situational analysis. Conditions checking the position of a satellite relative to the Earth's shadow can be combined with other situational conditions. A model of description of situational conditions is shown.

Introduction

Current trends in the development of satellite technologies lead, on the one hand, to the use of smaller satellites and, on the other hand, to multi-satellite missions instead of large multi-purpose satellites [1, 2] The process of space missions' analysis and design is related to the clarification and optimization of many and different parameters linked to scientific instruments, satellite subsystems, orbital parameters or templates of multi-satellite systems.

Computer simulations are playing an increasing role at all stages of preparation and operational implementation of multi-satellite missions. The application of parallel algorithms and calculations is crucial in solving large-scale problems involving hundreds and thousands of satellites.

An important part of the space missions' analysis is the so-called situational analysis. The situational analysis deals with the determination of optimal time

intervals suitable for the execution of satellite operations, depending on various geometric and physical conditions. This type of analysis is applied to different stages of mission preparation- starting with the conceptual study and preliminary analysis, going through mission definition, design and development, and finishing with implementation. The determination of the initial and final moments of the passage of a satellite through the shadow of the Earth as well as through the sunlit zone has wide practical applications. The evaluation of the energy produced by solar panels and its use in executing various operations is an example of such an application. These moments are necessary to perform measurements in the Earth's shadow or the illuminated part of a satellite's orbit.

Wertz and Larson [3] emphasize that mission analysis algorithms must be simple and effective enough to allow multiple runs, collect statistical data, and explore various scenarios and design options. Developing effective methods and computer programs for situational analysis is very important when many problems need to be solved, each involving more than one situation.

Algorithms and calculation tools for space mission analysis and design are under development at the Space Research and Technology Institute at the Bulgarian Academy of Sciences. One such tool is a parallel solver for situational analysis [4]. Algorithms and program realization of the conical model of the Earth's shadow are discussed in the present work. Algorithms and program codes suitable for multi-satellite situational analysis are proposed.

Concept of situation analysis

As pointed out above, situational analysis solves problems related to satellite operation optimization according to necessary restrictive conditions. Each situational problem SP is composed of one or a conjunction of several situational conditions sc_i of different types:

(1) $\qquad SP = sc_1 \wedge sc_2 \wedge \dots \wedge sc_n$

The conditions themselves are predicate functions accepting values of one or zero. They can be generally represented as:

$$sc_i = sc_i\left(\{\vec{R}\}, \{\alpha\}, \{\beta\}, t\right)$$

where $\{\vec{R}\}$ is a set of radius vectors of objects, $\{\alpha\}$ is a set of parameters of the mathematical model describing the situational condition and $\{\beta\}$ is a set of constraints that are specific to the conditions. The calculation of the SP function is done by sequentially checking the conditions sc_i. The application of Horner's rule allows cancelling the verification of the conditions when an unfulfilled condition sc_l is found:

(2) $\qquad SP = (\dots(\dots(sc_1 \wedge sc_2) \wedge \dots) \wedge sc_l \dots \wedge sc_{n-1}) \wedge sc_n$

16

It is important to keep in mind that each condition sc_i is satisfied within the time interval $T_{i,k}$ and is not fulfilled for the next adjacent interval $T_{i,k}^*$. The only requirement to $T_{i,k}$ and $T_{i,k}^*$, which will be taken into account, are $T_{i,k}, T_{i,k}^* \gg \triangle t$, where $\triangle t$ is a step in simulation time. The index **k** denotes the serial number of the interval within the time horizon.

Parallel solving of situational problems

A simulation of multi-satellite missions with many objects of observation requires solving many situational problems. Repeated simulations to optimize different parameters also lead to the need for multiple situational analyses. The functionality and reliability of multi-satellite systems result from the space environment's impact on the subsystems and the scientific instruments over a long time horizon. The computer simulations clarifying changes in the functionality and the reliability require computational efficiency of situational analysis. The application of parallel calculations is a step in the right direction.

Advances in computer technologies provide new opportunities for compiling even more complex simulation models but also place demands on their development. The article comments on solving situational problems in the context of parallel calculations (using multiprocessor and multicore systems). This requires special approaches and program models in the development of computational algorithms. The reasons for applying parallel calculations are the following:

1. complex mathematical models are used for the calculation of different situational conditions when searching for orbital events;
2. simultaneous solving of many situation problems is included in a simulation model;
3. the simulation time intervals (t_{begin}, t_{end}) are large;
4. Repeatedly solving of situation model based on situational analysis for space mission parameters optimization (concerning satellites, instruments, ground-based stations' and objects of investigations).

All compiled situational problems can be applied to all members in a multi-satellite system. The number of situational problems needed to be solved may increase according to the number of satellites and objects for observation and scientific problems for investigation.

Different situational conditions based on specific computational models give rise to a different number of computational operations. Such models are irregular, and an imbalance occurs when the calculations are parallelized. This is due to the uneven distribution of computational operations between the available processors. In poor distribution, some processors complete their problems and are free, while others continue their calculations. The "Pool of threads" model copes with this problem [5].

Parallel Situational Analysis Solver

A parallel solver for situational analysis was developed for this purpose. It is a processing program that consistently checks the feasibility of the conditions in a particular situational problem. The parallelization is based on computational threads organized in a variant of the program model "pool of threads" [6]. In this variant, the threads are synchronized with each otherwhile receiving situational problems to solve (race condition). This excludes the solution of one situational problem by more than one thread. The threads are also synchronized with the parent thread, which initiates the calculations and waits for them to complete at each simulated time step. Each thread takes one or more situational problems for processing according to the specified parameter value known as granularity.

Situational problems designer

A dialogue editor of situational problems is developed as an auxiliary tool. The compilation of a situational problem is initiated, situational conditions are successively selected, and their respective parameters and restrictions are set with the help of dialogue controls. An already compiled situational problem can be rejected or approved and saved as a template for future use. An optimization method can be selected too. One already assembled situational problem can be related to a particular satellite or all satellites of the respective multi-satellite system.

Conic Earth shadow model

Models of the earth's shadow have been developed and used for various purposes long ago. Ferraz-Mello [6] describes the shadow of the Earth with a cylindrical model. Bordovitzina et al. [7] apply a conical model of the Earth's shadow with penumbra. Determining the moments of entry and exit from the shadow is discussed in [8, 9]. Recently Srivastava has considered models of Earth's shadow [10]. Here, we will present detailed considerations of the conic models (also of the light zone) due to their use in developing program code for situational analysis purposes.

A geometric model that is used for calculating the parameters of the Earth's shadow is shown in Fig. 1. The circles (O_s, R_s) and (O_E, R_E) represent the Sun and the Earth in a plane passing through the line $\overleftrightarrow{l}$ and point S where the satellite is located, respectively. The radii of the Earth R_E and the Sun R_S are assumed to be known. The radius vector $\vec{R}_{SE}(t)$ of the Sun in a geo-equatorial coordinate system is determined based on methods of astrodynamics [11]. The radius vector of the satellite $\vec{r}_{sat}(t)$ is also determined based on numerical or analytical methods [11].

A first case – umbra

The lines $\overleftrightarrow{t_1}$ and $\overleftrightarrow{t_2}$ are tangent at the points $(T_{S,1}, T_{S,2})$ and $(T_{E,1}, T_{E,2})$ to the Sun and the Earth respectively in a plane defined by the vectors $\vec{R}_{SE}$ and $\vec{r}_{sat}$. From the similarity of the triangles $\Delta O_S T_{S,2} O_H$ and $\Delta O_E T_{E,2} O_H$ a formula can be derived for the height of the conical shadow of the Earth:

(3)
$$H_E = \frac{R_E \cdot |\vec{R}_{SE}|}{R_S - R_E}$$
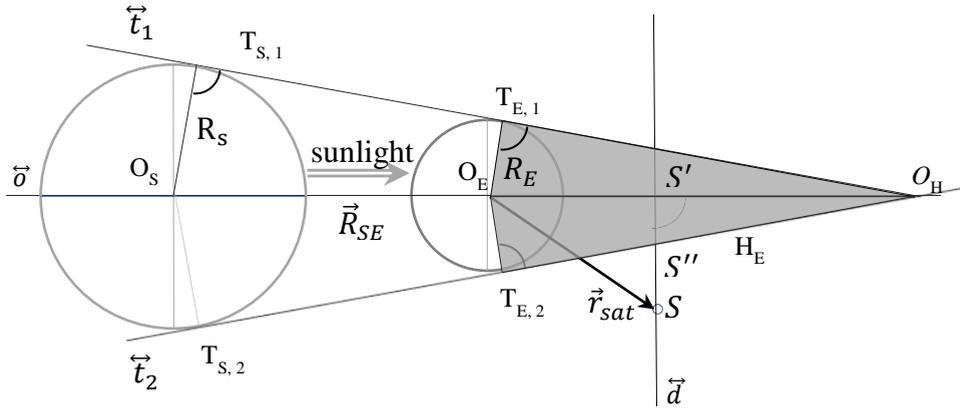
Fig. 1. Sun-Earth's umbra cone geometry

In this formula, R_E denotes the Earth's radius, the distance between the Sun and the Earth (the magnitude of the vector $\vec{R}_{SE}$) varies during the year in the interval from 147 099 760 km to 152 104 285 km. The height of the Earth's shadow H_E can be assumed to be a slowly changing quantity.

The next step is to determine whether the satellite is in the shadow of the Earth. One condition for this is that the sub-satellite point is located in the unlit part of the earth's surface. It is equivalent to checking if $\angle O_H O_E T_{E,2} < \angle O_H O_E S$:

(4)
$$\frac{\vec{R}_{SE} \cdot \vec{r}_{sat}}{|\vec{R}_{SE}| \cdot |\vec{r}_{sat}|} > \frac{\overline{O_E T_{E,2}}}{H_E},$$

because it compares the cosines of the indicated angles on both sides of the inequality without the use of arccosines.

To determine whether the satellite is in the shadow of the Earth, it remains to compare the segments $\overline{S'S''}$ and $\overline{S'S}$ on the line $\overleftrightarrow{d}$, perpendicular to the line $\overleftrightarrow{o}$. From the similarity of the triangles $\Delta O_H S' S''$ and $\Delta O_H T_{E,2} O_E$ it follows:

$$\overline{S'S''} = R_E.\frac{H_E - \overline{O_E S'}}{\overline{T_{E,2} O_H}} = R_E.\frac{H_E - |\vec{r}_{sat}|.\cos(\sphericalangle O_H O_E S)}{\sqrt{H_E{}^2 - R_E{}^2}}$$

The length of the line segment $\overline{T_{E,2} O_H}$, a leg in the right triangle $\Delta O_H T_{E,2} O_E$, is determined based on the length of the Earth's shadow H_E and the magnitude of the radius of the Earth R_E or finally:

$$\overline{S'S''} = R_E.\frac{H_E - r_{sat}.\dfrac{\vec{R}_{SE}.\vec{r}_{sat}}{R_{SE} \cdot r_{sat}}}{\sqrt{H_E{}^2 - R_E{}^2}}$$

or

(5) $$\overline{S'S''} = R_E.\frac{H_E - \vec{e}_{SE}.\vec{r}_{sat}}{\sqrt{H_E{}^2 - R_E{}^2}}$$

where $\vec{e}_{SE}$ is the unit radius vector of the Sun.

The length of $\overline{S'S}$ is determined by $\Delta O_E S' S$:
$$\overline{S'S} = |\vec{r}_{sat}|.\sin\sphericalangle S' O_E S$$

or

(6) $$\overline{S'S} = r_{sat}.\sqrt{1 - \left(\frac{\vec{R}_{SE}.\vec{r}_{sat}}{|\vec{R}_{SE}| .|\vec{r}_{sat}|}\right)^2} \quad)$$

It is important not to forget that determining the length of $\overline{S'S''}$ makes sense only if $\overline{O_E S'} < H_E$. Otherwise, there is a third case where the satellite is in the so-called antumbra.

Second case – penumbra
In like manner, we can determine the conditions where a satellite falls in the penumbra of the Earth. The conditions for the penumbra are determined by the satellite-Earth configuration, in which the Sun is partially obscured by the Earth, viewed from the position of the satellite (Figure 2). The lines $\overleftrightarrow{t_1'}$ and $\overleftrightarrow{t_2'}$ are tangent at the points $(T_{S,1}', T_{S,2}')$ and $(T_{E,1}', T_{E,2}')$ to the Sun and the Earth respectively in a plane defined by the vectors $\vec{R}_{SE}$ and $\vec{r}_{sat}$. From the similarity of the triangles $\Delta O_S T_{S,2} O_H^*$ and $\Delta O_E T_{E,2} O_H^*$ it follows:

$$\frac{H_E^*}{|\vec{R}_{SE}| - H_E^*} = \frac{R_E}{R_S}$$

or

20

(7)
$$H_E^* = \frac{R_E \cdot |\vec{R}_{SZ}|}{R_S + R_E}$$

where H_E^* denotes the line segment $\overline{O_H^* O_E}$. (H_E^* is analogous to the height of the conical umbra)

Fig. 2. Sun-Earth's penumbra geometry

The next step, after determining the distance H_E^*, is to check if a satellite is in the Earth's penumbra. The first condition checks whether the satellite is outside the penumbra (and shadow) zone. For this, it is sufficient to compare the angles $\sphericalangle O_S O_E T_{E,1}$ and $\sphericalangle O_S O_E S$ (or their cosines). Further calculations make sense if $\sphericalangle O_S O_E T_{E,1} < \sphericalangle O_S O_E S$:

(8)
$$-\frac{\vec{R}_{SE} \cdot \vec{r}_{sat}}{|\vec{R}_{SE}| \cdot |\vec{r}_{sat}|} > \frac{\overline{O_E T_{E,1}}}{H_E^*},$$

because it compares the cosines of the indicated angles on both sides of the inequality without the use of arccosines. To determine whether the satellite is in the Earth's penumbra, it remains to compare the segments $\overline{S'S'''}$ and $\overline{S'S}$ on line $\overleftrightarrow{d}$, perpendicular to the line $\overleftrightarrow{o}$ (Figure 2). From the similarity of the triangles $\Delta O_H^* S' S'''$ and $\Delta O_H^* T_{E,1} O_E$ it follows:

$$\overline{S'S'''} = \frac{R_E}{\sqrt{H_E^{*\,2} - R_E^{\,2}}} \cdot \left(H_E^* + r. \frac{\vec{R}_{SE}.\vec{r}_{sat}}{|\vec{R}_{SE}| . |\vec{r}_{sat}|} \right)$$

For the size of the line segment $\overline{S'S}$ we have again:

$$\overline{S'S} = r_{sat}. \sqrt{1 - \left(\frac{\vec{R}_{SE}.\vec{r}_{sat}}{|\vec{R}_{SE}| . |\vec{r}_{sat}|} \right)^2}$$

In the penumbra zone, the Sun is partially obscured by the Earth, reducing the flow of light reaching the satellite. This reduction depends on the part of the solar disk covered by the Earth's disk. At low orbits, the time for satellites to pass through the penumbra is short and insignificant compared to the times for passing through the sunlit and umbra zones. However, at higher orbits, this time is longer and could be important to be considered.

Program realization

A situational problem description model
A descriptor of situational problems is a one-dimensional array whose elements are derived types containing the values of different attributes (parameters and constraints as well as results) of the conditions comprised in the problem. The first (zero) element of the descriptor contains control information and results about the entire situational problem. The following elements contain the values of different attributes (parameters and constraints as well as results) of the situational conditions.

Figure 3 illustrates a general template in Fortran language for the creation of situational problem descriptor objects. The descriptors of different situational problems are combined into a two-dimensional array with dimensions $K \times N$, where N is the number of problems and K is the maximum number of situational conditions among all problems.

- **SitCond** is a derived type that describes different situational conditions. It contains various attributes, which are common for each situational condition. Some of them are:**sit_code** – identification code of situation condition
- **sat_num** - satellite number to which the condition is associated
- **duration** – time interval when the condition is met
- **dt_sit** – local parameter- accumulates duration of condition before ending duration

The **UNION** statement defines groups of fields that share memory among different situational conditions. For the conditions that are discussed here, the following map … end map contains three logical variables **umbra**, **penumbra,** and

sunlit. When the satellite is in a part of the orbit falling in the shadow, penumbra, or illuminated zone, the corresponding variable is assigned a value of true. These variables have control functions. The **UNION** statement defines an array of storage. The **UNION** operator specifies an area of memory that is used polymorphically by each of the functions computing different situational conditions.

```
MODULE
  type    SitCond
    integer   sit_code        ! Code of the situation condition; every situation has identification code
    integer     sat_num       ! Which satellite concern this situational problem
    logical     flag          ! Satisfaction of situational condition: .false. or .true.
    logical   begin_sit       ! If is true – the beginning of situational condition satisfaction
    logical   fl_rezults      ! If .true. - flag for end of situational conditional's interval
    real*8        t12(2,3)    ! Determine the last time interval where the condition is satisfied
    real        duration      ! Duration of a current situational condition (event)
    real        dt_sit        ! Local parameter- accumulates duration of condition before ending duration
    real        t_cond_total  ! Local parameter- accumulates total durations for the whole time horizon
  union
    map …    ! Other situational conditions
    end  map ! Other situational conditions
    map                       ! Sit_78/79: satellite in conic shadow
      real      num_sat_79
      logical    umbra        ! If      umbra.EQ..true. - penumbra=.false.,  sunlit=.false.
      logical  penumbra       ! If penumbra.EQ..true. -      sunlit=.false., umbra=.false.
      logical  sunlit         ! If      sunlit.EQ..true. - penumbra=.false.,  umbra=.false.
    end map
    map …
    end  map ! For other situational conditions
  end union
  end type  SitCond

  type    sit_problem
  union
    map                        ! Only for solving control- contains the number of situation conditions
      integer    SP_code       ! Contains the serial number of the situational problem
      integer    SP_type       ! Contains a unique code of situational problems
      integer    max_cond      ! The number of situational conditions for the current problem
      logical    requirement   ! Satisfaction of situational problem: .false. or .true.
      integer    opt_level     ! Optimization algorithm: 0- none, 1/2/3
      logical    begin_sit
      logical    fl_results    ! If .true. - flag for end of situation interval end ready results
      real*8     t1,t2         ! The last time interval where the situational conditions are satisfied
      real      duration,dt_sit ! Duration of time interval when all conditions are satisfied
      real      t_ problem_total   ! Accumulates total durations for the whole simulated period
      integer    problem_code     ! A code of satellite operation related to the situational problem
    end map
    map
      type (SitCond) sit_cond
    end map
  end union
  end type  sit_ problem
END MODULE
```

Fig. Derived types for situational problems compilation

The **sit_problem** is a derived type that describes different situational problems. The UNION statement defines groups of two MAP blocks that describe the elements of one situational problem. The first MAP block describes the zero element of the situational problem descriptor, which contains control parameters and the problem's attributes. The second MAP block allows the inheritance of the properties of each of the situational conditions in the situational problem. The Situational solver interprets situational condition attributes according to the identification code **sit_code**. Each function corresponding to a given situational condition interprets the attributes in a specific way corresponding to a corresponding MAP block.

Design of the computer subroutines for the umbra/sunlit situational conditions

Two computer subroutine functions **Sit__78** and **Sit__79** have been developed to check if the satellite is in the shadow or in the sunlight zone respectively (Appendix A). The subroutines are realized in the Fortran language. When carrying out a situational analysis, which is related to the simulation of multi-satellite systems, some tasks refer to observations in the Earth's shadow and others to observations in the sunlit zone. Some calculations about the geometry of the shadow are the same for all satellites. For this reason, the subroutine named **Preliminary_Calculations** is added to increase computational efficiency. It performs these common calculations related to the umbra and light zone, which depend only on the Sun-Earth distance-the height of the conical shadow of the Earth H_E.(expression (3)), analogous quantity H_E^* and some others. This subroutine, as well as the **Sit__78**, is an additional entry point to the **Sit__79** subroutine. The results of the calculations in **Preliminary_Calculations** are contained in static local variables that are available within the **Sit__79** and **Sit__78** subroutines. When these subroutines are used within the parallel processor for situational analysis (Atanassov, 2016), these variables are common to all computational threads. Some variables (described in operator "automatic") need to be declared as dynamic to be in the local memory for each thread. The calculations for different situational tasks are not influenced by each other. Besides, the code equal to the two subroutines **Sit__79** and **Sit__78** is differentiated within the internal subroutine **If_flag**. This subroutine has control functions for the situation analysis processor and calculates the values of the variables specific for each situational condition stored in a generic structured variable **SitCond** (Appendix A). The approach allows situational conditions' subroutines to be reentrant and protected when multi-thread parallelization is applied. **If_Flag** internal function operates with the dummy arguments of **Sit__78** and **Sit__79** functions.

The main purpose of the subroutines **Sit__79** is to check the feasibility of the relevant conditions. The first check through operator **a1:IF** is related to inequality (4). This check expresses the necessary condition to seek a satellite in the

Earth's umbra. If the necessary condition is fulfilled, then distances $S'S''$ and $S'S$ according to the expressions (5) and (6) are calculated. Let us recall that when calculating $S'S''$, preliminary calculations performed in the subroutine **Preliminary_Calculations** are used. Checking whether the satellite is in the umbra is performed in the operator **a2:IF**. If the satellite is not in the umbra, the operator **a3:IF** checks based on the expressions (6) and (9) whether it is in the penumbra (checking with the **a1:IF** operator doesn't guarantee either shadow or penumbra).

Subroutine **Sit__78** deals with the light zone analogously to **Sit__79**. The quantity H_E^* (analogous to the height of the Earth's umbra H_E) is presented by expression (7) and calculated in subroutine **Preliminary_Calculations** as variable H_penu. The **a5:IF** operator checks the necessary condition (8). The operator **a6:IF** checks whether the satellite is in the light zone. The operator **a7:IF** checks based on expressions (5) and (6) whether it is also in the penumbra (regardless of the condition (8)) in case it is not in the light zone.

The determination of the time intervals, when conditions are met, is performed by the internal subroutine **If-Flag**. The interval begins when the value of the condition changes from false to true. The end of the interval is taken to be the moment when the value of the situational condition changes from true to false.

Conclusion and outlook

Situational conditions related to the models of both, the Earth's shadow and the lit zone are examples of conditions participating in several different situational problems. Such situational problems arise from numerous scientific problems and involve the use of data from diverse instruments located on satellites, either of one or more space missions. The attention is directed to using these models in the frame of **Situational Analysis Solver** in the analysis of multi-satellite space missions.

Performing an analysis related to solving a large number of situational problems, each with several conditions can be time-consuming. Apart from the computational models' optimization of the situational conditions and the application of parallel calculations, the search and research for optimization methods of the situational analysis are an interesting challenge and necessary continuation of the work.

Acknowledgment

Appendix A. Source code of the subroutines checking the situational conditions

```fortran
!********************************************************************
! Sit__78: conic shadow, checks if the satellite is in the sunlit zone, out of umbra
! Sit__79: conic shadow, checks if the satellite is in the penumbra
! Preliminary_Calculations : common for all satellites
!                 If_Flag  : internal subroutine
!      Dist_sun   – Sun-Earth distance
!         H_shad – a height of the conical shadow of the Earth
!         Re       – radius of the Earth
!         Rm_sun– solar radius
!            r_sat – modulus of the radius vector of satellite in GeKS
!..................................................................
FUNCTION      Sit__79  (t,dt,xv,umbra,penumbra,fl_rezults,duration,begin_sit,dt_sit,t12)
 USE DFlib
   logical      Sit__79,Sit__78,umbra,penumbra,fl_rezults,         begin_sit
   real                                             duration,       t12*8(2,3)
   real*8      t,dt,xv(3)
   real*8       r_sat
   real*8      Re/6378.D3/,Rm_Sun/695510.D3/ ! [m]  /1.D9/
   real*8       S1S2,S1S,SpSss,SpS
   real*8      H_shad, H_penu, D_shad, D_penu, cos_OhOeTe1, cos_OhOeTe2
   real*8      rS,rSun(3)
   common   /cSun_vek/rS,rSun ! Modulus of vector and directional cosines
   logical       flag
   automatic    flag,S1S2,S1S3,S1S,SpSss,SpS,cos_OhOeS
    r_sat    = SQRT(xv(1)**2 + xv(2)**2 + xv(3)**2) ! |radius - vector| of the satellite - modulus
  cos_OhOeS= -(xv(1)*rSun(1)   + xv(2)*rSun(2)   + xv(3)*rSun(3)  )/(r_sat) !- from scalar product
a1:IF(cos_OhOeS.GT.cos_OhOeTe2) THEN ! It makes sense to check for umbra
    S1S2= D_shad*(H_shad - r_sat*cos_OhOeS  )
    S1S  = r_sat*SQRT((1.D0 - cos_OhOeS)*(1.D0 + cos_OhOeS));
   a2:IF(S1S.LE.S1S2) THEN ! The satellite is in umbra
                       flag=.true.; umbra=.true.
                     ELSE ! The satellite is in penumbra eventually
                       flag=.false.;
         S1S3= D_penu*(H_penu + r_sat*cos_OhOeS  )
     a3:IF(S1S.LT.S1S3) THEN ! Checking for penumbra
                        penumbra=.true.;
                       ELSE
                        penumbra=.false.
          ENDIF a3
       ENDIF a2
                  ELSE ! No sense to check in the sunlit zone
                   flag=.false.
  ENDIF a1
    CALL If_Flag(Sit__79)
RETURN
ENTRY              Sit__78(t,dt,xv,umbra,penumbra,fl_rezults,duration,begin_sit,dt_sit,t12)

    r_sat    = SQRT(xv(1)**2 + xv(2)**2 + xv(3)**2) ! |radius - vector| of satellite - modulus
  cos_OhOeS= -(xv(1)*rSun(1)   + xv(2)*rSun(2)   + xv(3)*rSun(3)  )/(r_sat) !- from scalar product

a5:IF(-cos_OhOeS.GT.cos_OhOeTe1) THEN ! The satellite is in sunlit zone
      flag=.true.
                  ELSE ! No sense to check about umbra and penumbra
      S1S3= D_penu*(H_penu + r_sat*cos_OhOeS  )
      S1S  = r_sat*SQRT((1.D0 - cos_OhOeS)*(1.D0 + cos_OhOeS))
   a6:IF(S1S.GE.S1S3) THEN ! Satellite is in sunlit zone!!!
```

```fortran
                        flag=.true.
                    ELSE ! the satellite is in penumbra eventually
                        flag=.false.
              S1S2= D_shad*(H_shad - r_sat*cos_OhOeS)
      a7:IF(S1S.GT.S1S2) THEN ! Checking for penumbra
                              penumbra=.true.; umbra=.false. ; flag=.false.
                          ELSE ! outside the penumbra
                              penumbra=.false.; umbra=.true. ; flag=.false.
            ENDIF a7;
          ENDIF a6
        ENDIF a5
        CALL  If_Flag(Sit__78)
RETURN
ENTRY        Preliminary_Calculations()

      H_shad= Re*rS/(Rm_Sun  - Re);        cos_OhOeTe2= Re/H_shad
      H_penu= Re*rS/(Rm_Sun + Re);        cos_OhOeTe1= Re/H_penu
      D_shad= Re/SQRT((H_shad  - Re)*(H_shad + Re))
      D_penu= Re/SQRT((H_penu  - Re)*(H_penu + Re))
 RETURN
CONTAINS
  SUBROUTINE        If_Flag(Sit_cod)
    Logical               Sit_cod
  IF(flag) THEN
        IF(.NOT.begin_sit) THEN ! Beginning of the interval
                    begin_sit=.true.
                t12(1,1)= t     ! stores the start time
                  dt_sit=.0; fl_rezults=.false.
                        ELSE
                t12(2,1)= t     ! stores the final  time
          ENDIF
              dt_sit= dt_sit + dt;
          Sit__cod=.true. ;
    ELSE ! Satellite isn't visible
        IF(begin_sit) THEN      ! the situational interval continues
                    duration= dt_sit - dt; fl_rezults=.true.
                      dt_sit=.0;          begin_sit=.false.
                    ELSE
                     duration=.0;
          ENDIF
        Sit_cod=.false.
    ENDIF
  END SUBROUTINE  If_Flag
END FUNCTION  Sit__79
```

References

1. Poghosyan, A., Lluch, I., Matevosyan, H., Lamb, A., Moreno, C. A., et al. Unified classification for distributed satellite systems. In *4th International Federated and Fractionated Satellite Systems Workshop, 10-11 oct, Rome, Italy,* 2016.
2. Selva, D., A., Golkar, O. Korobova, I. L. I. Cruz, P. Collopy, O.L. de Weck Distributed earth satellite systems: What is needed to move forward?. *Journal of Aerospace Information Systems*, 2017, *14*(8), pp. 412–438.
3. Wertz, J. R., W. J. Larson Space Mission Analysis and Design. Microcosm Press, Kluwer Academic Publishers, third ed. 1999, 892 p.

4. Atanassov, A.M. (2016) Parallel satellite orbital situational problems solver for space missions design and control, *Advances in Space Research*, v. 58, 9, 2016, pp. 1819–1826.

5. Rauber, T., Rünger G. Parallel Programming: For Multicore and Cluster Systems. Springer. 2010, 455 p; DOI 10.1007/978-3-642-04818-0

6. Ferraz-Mello, S. Sur le problém de la pression de radiation dans la théorie des Satellites Artificiels. C.R.Acad. Sc. Paris, 1965, 258, p. 463.

7. Bordovitzina, T. V., Bykova, L. E., Kardash, A. V., Fedyaev, Yu. A., Sharkovskii, N. A.,. Efficient algorithms for numerical simulation of the motion of earth satellites, Russian Physics Journal, 1992, Vol. 35, № 8, pp. 735–742.

8. Eremenko, R. P. Tochnoe reshenie uravneniya teni. Bjul. Instituta teoreticheskoi astronomii, 1965, 6 (119), p. 446. (rus.)

9. Neta, B., D. Vallado On Satellite Umbra/Penumbra Entry and Exit Positions, *Journal of the Astronautical Sciences*, 1998, 46, No. 1, pp. 91–104.

10. Srivastava, V. K., Pitchaimani, M., & Chandrasekhar, B. S.,. Eclipse prediction methods for LEO satellites with cylindrical and cone geometries: a comparative study of ECSM and ESCM to IRS satellites. *Astronomy and Computing*, 2013, 2, 11–17.

11. Vallado, D. A., Fundamentals of Astrodynamics and Applications. Microcosm Press, fourth edition, 2013, 1108 p.

РАЗРАБОТКА СИТУАЦИОННЫХ УСЛОВИЙ И ПОДПРОГРАММ ДЛЯ ПАРАЛЛЕЛЬНОГО ВЫЧИСЛИТЕЛЬНОГО ИНСТРУМЕНТА СИТУАЦИОННОГО АНАЛИЗА С ИСПОЛЬЗОВАНИЕМ КОНИЧЕСКОЙ МОДЕЛИ ТЕНИ ЗЕМЛИ

А. Атанасов

Аннотация

Поиск оптимальных временных интервалов выполнения операций спутников основан на проверке конкретных условий геометрического или физического характера. В ситуационной задаче сочетаются несколько условий. Для решения различных ситуационных задач необходимо заранее разработать программные коды для проверки различных ситуационных условий. Представлены два ситуационных условия для определения того, находится ли спутник в освещенной солнцем зоне или в тени на основе конических моделей земной тени. Вводятся необходимые (но не достаточные) условия для нахождения спутника соответственно в тени или освещенной солнцем зоне упрощающие расчеты и повышают вычислительную эф-фективность для большей части орбиты спутника. Задачи с одним или несколькими условиями решаются с помощью разработанного параллельного решателя для ситуационного анализа. Условия проверки положения спутника относительно тени Земли могут сочетаться с другими ситуационными условиями.

DETERMINE COVERAGE CHARACTERISTICS OF EARTH SATELLITE

Assen Marinov

Aviation Faculty, Bulgarian Air Force Academy
e-mail: asen_aerodynamics@abv.bg

Abstract

    Threats posed by counterspace capabilities are directed against space systems, their supporting ground infrastructure, and data links between space systems and ground infrastructure. Space countermeasures include direct attack and co-orbital anti-satellite systems, cyber attacks, electronic warfare and directed energy. Earth, as seen from space, is a key visual element in planning operations. This necessitates a careful selection of continuous review of certain areas of operational and strategic interest. Satellite coverage planning covers the entire process, from the idea of a new satellite system to final in-orbit testing. It is a multidisciplinary activity that ranges from defining the areas of interest in the relevant geographic areas, designing the appropriate orbit, and arriving at the determination of the necessary sensors that will meet the mission's need.

Introduction

The modern security environment requires more and more complete interaction between the types of armed forces and the branches of troops, based on a unified information space, which creates conditions for situational awareness and information superiority. This cannot be achieved without adequate infrastructure, mobile and secure logistics, command and control systems, communications, computers, cyber defense, combat systems, and intelligence (the result of surveillance and intelligence combined with other data), surveillance and intelligence (collection of data to solve a specific military issue) {C6ISR}, electronic warfare, drones that are directly dependent on continuous communication with satellites orbiting in outer space.

Depending on their purpose, the missions of some satellites require continuous coverage of a certain area of the Earth or the ability to communicate simultaneously with any point on the Earth. Given the shape of the Earth, it is clear that a single satellite cannot provide simultaneous communication with every point. This requires the satellites to be placed in the same or different orbits to provide the necessary coverage. For this reason, it is very important to correctly determine the

coverage of each individual satellite. The most famous example is the Global Positioning System (GPS), whose mission requires that every point on Earth be visible to at least four GPS satellites at any given time. To ensure the mission, the GPS constellation contains 24 satellites working together to provide continuous coverage of the world. Placing satellites in a higher orbit requires a larger launch vehicle and greater cost. However, the height of the orbit essentially depends on how much of the earth's surface the satellite's sensors can cover. Logically, the higher the orbit, the larger the total area the corresponding satellite can cover. A satellite's field of view is defined as the cone of visibility for a particular sensor. Depending on the field of view of the satellite and the height of its orbit, the total area of the earth's surface that is constantly covered is determined as the corresponding linear width or diameter of this area is called coverage width (Fig. 1) [1–2].
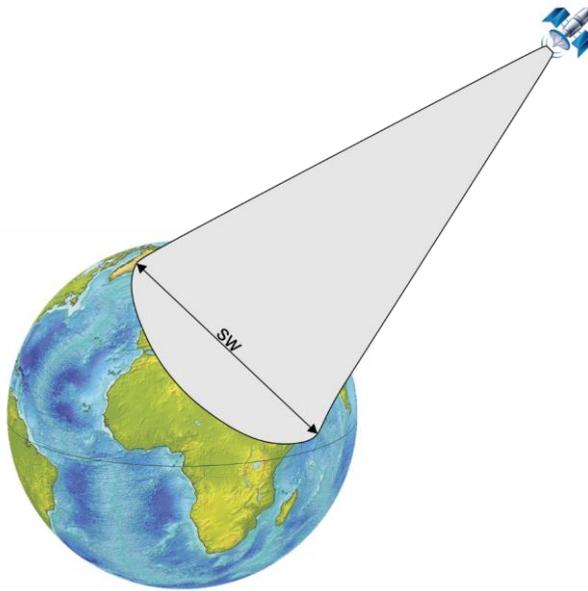
Fig. 1. Coverage on Earth's surface (swath width)

Depending on the desired coverage of the Earth, necessary for the implementation of the set mission of the satellite, the main elements of the orbit in which it will be placed are determined. For example, if the satellite is to survey the entire surface during the mission, it needs a near-polar inclination of about 90°.

Geometry of satellite coverage

Satellite coverage depends on and can be calculated from the following satellite orbital parameters (Fig. 2) [3]:

- perigee;
- apogee;
- extreme northern latitude (for Bulgaria - the mouth of Timok in the Danube River, Vidin region, coordinates 44.214555°N and 22.67459°E);
- extreme southern latitude (for Bulgaria – Veikata mountain, Kardzhali region, 41.236022°N and 25.288167°E);
- true anomaly;
- latitude of the satellite;
- nadir (na) (Nadir is a term used to designate a point on the celestial sphere opposite the zenith, or more precisely, the point with a slope of −90°, located in the direction "down" from the observer);
- central angle of the Earth (ca);
- elevation angle (ea);
- slope range (sr);
- geocentric radius of the satellite $\left(R_s = R_E + H_s \right)$.
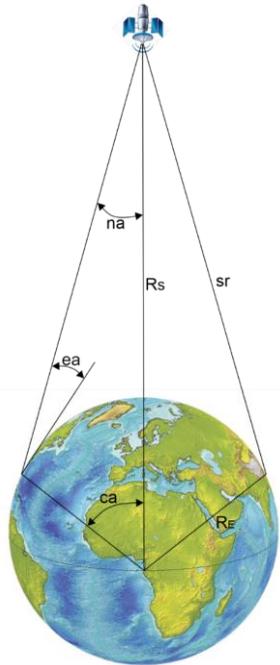
Fig. 2. Coverage geometry of a satellite orbiting a spherical planet with a nadir-pointing conical sensor

The relationship between the nadir, central angle, and elevation angle can be written as follows:

(1) $\quad na + ca + ea = 90^o$;

(2) $\quad sr.\cos(ea) = R_s \sin(ca)$;

(3) $\quad sr.\sin(na) = R_E \sin(ca)$;

Therefore:

(4) $\quad na_s = \arcsin\left(\dfrac{R_E}{R_s}\right)$;

The nadir, central angle, and elevation angle can also be represented as a function of slope range sr:

(5) $\quad na = \arccos\left(\dfrac{R_s^2 - R_E^2}{2.sr.R_s} + \dfrac{sr}{2R_s}\right)$;

(6) $\quad ca = \arccos\dfrac{R_s^2 + R_E^2 - sr}{2.sr.R_E}$;

(7) $\quad ea = \arcsin\left(\dfrac{R_s^2 - R_E^2}{2.sr.R_E} + \dfrac{sr}{2R_E}\right)$.

The nadir, the central angle, and the slope range can also be represented as a function of the elevation angle:

(8) $\quad na = \arcsin\left(\dfrac{R_E}{R_s}\cos(ea)\right)$;

(9) $\quad ca = \arccos\left(\dfrac{R_E}{R_s}\cos(ea)\right) - ea$;

(10) $\quad sr = \sqrt{R_s^2 - R_E^2 \cos^2(ea)} - R_E \sin(ea)$;

The relationship between the central angle, the elevation angle, and the slope range can be represented as a function of the nadir:

(11) $\quad ca = \arcsin\left(\dfrac{R_s}{R_E}\sin(na)\right) - na$;

$$(12) \quad ea = \arccos\left(\frac{R_s}{R_E}\sin(na)\right);$$

$$(13) \quad sr = R_s\cos(na) - \sqrt{R_E^2 - R_s^2\sin^2(na)}.$$

The relationship between the nadir, elevation angle, and slope range as a function of the central angle can be represented as follows:

$$(14) \quad na = \arcsin\left(\frac{R_E\sin(ca)}{sr}\right);$$

$$(15) \quad ea = \arctan\left(\frac{R_s\cos(ca) - R_E}{R_s\sin(ca)}\right);$$

$$(16) \quad sr = \sqrt{R_s^2 - R_E^2 - 2R_sR_E\cos(ca)}.$$

The width of coverage (sw) can be calculated as follows:

$$(17) \quad sw = 2R_E.ca;$$

The observed land surface percentage under these conditions is $50(1 - \cos(ca))$, and the covered surface (SA):

$$(18) \quad SA = 2\pi R_E^2(1 - \cos(ca)).$$

Results

The calculation sequence is presented in Fig. 3, and the results are in Tables 1, 2, 3 and 4.
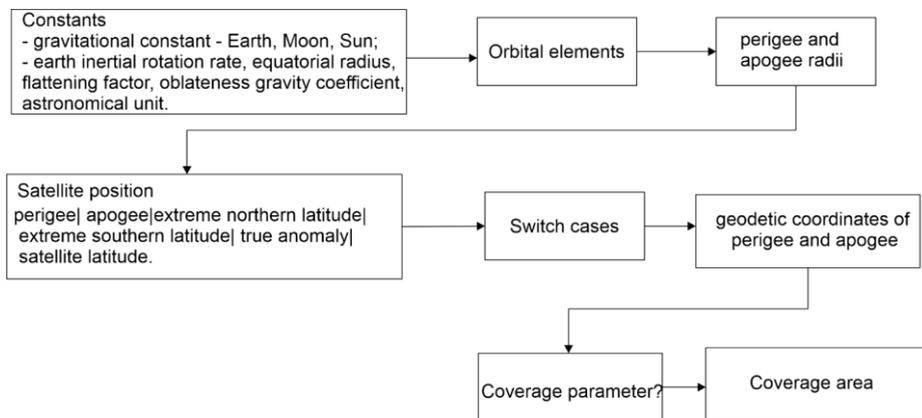
Fig.3. Calculation sequence

Table 1. Calculation results -1

Input data	Output data
✓ semimajor axis -10000 km; ✓ inclination – 0º; ✓ satellite's position - extreme northern latitude; ✓ coverage constraint - elevation angle = 7º.	• satellite altitude - 3621.8637 km; • nadir - 39.2762º; • central angle - 43.7238º; • slant range - 6963.7324 km; • coverage area - 70884025.0530 km^2 = 13.8660 %; • view latitude 1 – (- 43.7238º); • view latitude 2 – (43.7238º); • distance of arc - 4867.3099 km; • true anomaly – 0º.

Table 2. Calculation results -2

Input data	Output data
✓ semimajor axis -10000 km; ✓ inclination – 30º; ✓ satellite's position - extreme northern latitude; ✓ coverage constraint - elevation angle = 7º.	• satellite altitude - 3627.2203km; • nadir - 39.2762º; • central angle - 43.7238º; • slant range - 6963.7324 km; • coverage area - 70884025.0530 km^2 = 13.8660 %; • view latitude 1 – (-13.7238º); • view latitude 2 – (73.7238º); • distance of arc - 4867.3099 km; • true anomaly – 90º.

Table 3. Calculation results -3

Input data	Output data
✓ semimajor axis -10000 km; ✓ inclination – 60º; ✓ satellite's position - extreme northern latitude; ✓ coverage constraint - elevation angle = 7º.	• satellite altitude - 3637.9127 km; • nadir - 39.2762º; • central angle - 43.7238º; • slant range - 6963.7324 km; • coverage area - 70884025.0530 km^2 = 13.8660 %; • view latitude 1 – (16.2762º); • view latitude 2 – (76.2762º); • distance of arc - 4867.3099 km; • true anomaly – 90º; • view latitude is over the pole.

Table 4. Calculation results -4

Input data	Output data
✓ semimajor axis -17893 km; ✓ inclination – 0°; ✓ satellite's position - extreme northern latitude; ✓ coverage constraint - elevation angle = 7°.	• satellite altitude - 11514.8637 km; • nadir - 20.7201°; • central angle - 62.2799°; • slant range - 15958.3818 km; • coverage area - 136709100.0687 km^2 = 26.7424 %; • view latitude 1 – (-62.2799°); • view latitude 2 – (62.2799°); • distance of arc - 6932.9668 km; • true anomaly – 0°.

Conclusions

There are many factors that influence the coverage of a satellite or constellation of satellites. Of primary importance is the ability to steadily collect images depending on the elements of the orbit. Most moderate and low-resolution sensors possess an intelligence plan that covers the globe in a regular and repeating pattern of activity. Information assurance for military operations requires constellations of satellites to achieve global coverage at moderately high resolution. Current orbital plan requirements for individual satellites provide high-resolution coverage that does not affect active imaging. Even if the sensors are constantly switched on and transmit data, it is possible that the optical images are unusable over areas with dense clouds covering them. Unlike commercial targets, where there is less interest in much of the global south, for military targets, all parts of the earth's surface are of real interest. This places various increased demands on the sensor's ability to collect imagery over a given area. A satellite can carry a suite of sensors and collect images from space, just like drones, aerostats, and aircraft, but they do it from a much lower altitude. Satellites provide high temporal resolution over large areas of the Earth's surface, periodically collecting information over a specific area, depending on orbit and satellite coverage over a very long period of time, which sets the stage for large-scale reconnaissance. At the same time, airborne vehicles provide images with very high resolution - up to 1 cm per pixel. This makes them suitable when a specific task is to be performed.

References

1. Lambeth, B., Next steps in the military uses of space, United States Air Force, 2003.
2. Wright, D., Grego L., The Physics of Space Security, American Academy of Arts and Sciences, 2005.
3. Yuan, J., Cheng Y., Orbital Dynamics and Mission Trajectory Desig, Science Press, Beijing, China, 2019.

ОПРЕДЕЛЯНЕ ХАРАКТЕРИСТИКИТЕ НА ПОКРИТИЕТО
НА САТЕЛИТ В ОКОЛОЗЕМНА ОРБИТА

А. Маринов

Резюме

В зависимост от предназначението си, мисиите на някои спътници изискват непрекъснато покритие на определен район от Земята или възможност за едновременна комуникация с всяка точка на Земята. Предвид формата на Земята е ясно, че един сателит не може да осигури едновременна комуникация с всяка точка. Това налага спътниците да бъдат разположени на еднакви или различни орбити, за да осигурят необходимото покритие. Поради тази причина е много важно да се определи правилно покритието на всеки отделен спътник. В настоящата статия е определено покритието на сателит в околоземна орбита, като е взета предвид крайна северна ширина и ъгъла на издигане, на базата на което са определени и характеристиките на покритието.

DIFFERENT ASPECTS OF THE USE OF SMALL SATELLITES FOR OBSERVATION OF THE MODERN STAGE

Pavel Penev

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

Abstract: The reasons for the appearance of small satellites have been revealed. The evolution of the military and civilian small satellites created in the 21st century for surveillance and intelligence, including dual-purpose ones, is reviewed. Three groups of tasks solved by small observation satellites are defined - applied, technological, and scientific, as well as their growing importance and areas of use. The need is substantiated, and steps are proposed in the Republic of Bulgaria, respectively in SRTI-BAS, to activate the activity of creating and using small satellites for observation for the needs of the economy, ecology, humanitarian, and other areas.

Introduction

The modern development of world cosmonautics is inextricably linked to the use of small space vehicles (SSVs), in the most general case small satellites (satellites) solving a wide range of tasks of a different nature. For example, 94% of the number of satellites launched in 2020 are small satellites [1].

At the end of the last century, the Arianespace Company proposed a conditional classification of small satellites according to their mass, as follows: minisatellite 500-100 kg; microsatellite 100-10 kg; nanosatellite 10-1 kg; picosatellite - under 1 kg. Recently, the mass of picosatellites (0.1-1 kg) has been refined. With the development of high technologies in the XXI century, the above classification is supplemented [2] with a new category of MCLA - femtosatellites with a mass of 10 to 100 grams, implemented on a crystal or in an integrated circuit.

Several reasons can be given for the appearance of small satellites. A large satellite is created in 2-5 years, while the SSVs take 6-9 months. Therefore, several small satellites can be prepared for the time of creation of a large satellite.

Small satellites' relative simplicity and short development time usually result in lower costs. The shorter time it takes to create small satellites allows them to be renewed when they are in low Earth orbit (LEO) because the term of their

operation in space is significantly shorter (from one to several years) than that of large satellites. This also allows for the periodic introduction of new technologies.

The cost of creating a small satellite (for example, a nanosatellite) and putting it into Earth orbit is significantly less than that of a large satellite. For example, the Maxar company's 2,500 kg WorldView-4 commercial optoelectronic observation (OEO) satellite costs $850 million to build and launch, while the cost to build and launch a nanosatellite is in the hundreds of thousands or tens of thousands of dollars.

The increasingly high degree of miniaturization of electronic elements, reduced electrical consumption, advances in mechanical systems, batteries, and sensors are also prerequisites for creating and efficiently using small satellites.

The rapid development of the element base of small satellites predetermines their commercial use now and in the future.

In modern conditions, the leading space countries are increasingly launching SSVs, whose mass does not exceed 10 kg, into Earth orbit. An important role in this is played by the "CubeSat" standard (Cube Satellite - cubic satellite, CubSat), by which these miniature satellites with a modular cubic shape - mainly nanosatellites and picosatellites - are built. More specifically, the "CubeSat" standard (Fig. 1) was created in 1999 by scientists and specialists at the Stanford and California Universities in the USA to provide the possibility of launching small payloads (PL) into space.

In the "CubeSat" standard, several sizes of satellites are defined, the smallest of which is denoted by 1U (from unit) and has the shape of a cube with an edge of 10 cm. Specifically, the "CubeSat" standard is a format of small satellites for space exploration, having dimensions of 10×10×10 cm and a mass of no more than 1.33 kg.

The "CubeSat" standard allows the unification of two or three cubes in the composition of one satellite with designations 2U and 3U, respectively, having dimensions of 10x10x20 cm and 10x10x30 cm, respectively. Currently, links from nanosatellites and picosatellites are increasingly being launched into Earth orbit, respectively, according to the standards "CubeSat 6U", "CubeSat 12U" and "CubeSat 16U".
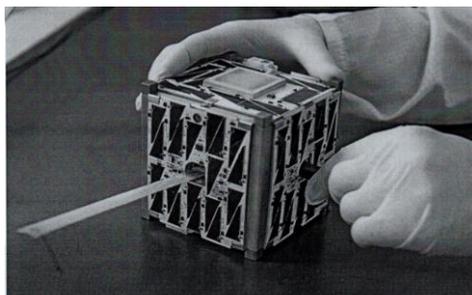
Fig. 1. General view of a small satellite of the "CubeSat" standard

After the emergence of the "CubeSat" standard, the even smaller "PocketCube" format appears with a mass of up to 250 grams and dimensions of 5x5x5 cm, and this size is designated as 1p.

Structurally, the satellite, according to the "CubeSat" standard, is an aluminum skeleton (Fig. 1), inside which the following elements are most often placed: PT; central processing unit; antenna-feeder devices; radio channel; power system including battery and charge controller; solar cells (surface or dissolving); satellite positioning system; satellite position correction system.

The CubeSat standard satellite is placed in a specialized container, usually located in the corresponding launch vehicle (LV). To launch the nanosatellite into Earth orbit, it is placed in a so-called "deployment device", which is a rectangular container. In this container, attached to the LV, the satellites of the "CubeSat" standard are placed.

CubeSat small satellites are currently launched into Earth orbit by the dozens, even over 100 SSVs at a time, either via RNs or aboard manned or automated cargo spacecraft (SC) and the International Space Station (ISS).

Problem status and research area

At the current stage, creating nanosatellites according to the "CubeSat" standard is significantly easier compared to the recent past, but putting them into orbit around the Earth is not within the power of the creators of these satellites alone. The delivery of nanosatellites into space requires the conclusion of a contract with a licensed rocket and space organization (NASA, ESA, Roskosmos, etc.).

The evolution of SSVs over the last ten years is closely related to the main producers of nanosatellite constellations in the USA - the private companies "Planet Labs" and "Spaceflight Industries". As a typical example of a satellite with a mass not exceeding 10 kg, a nanosatellite from the "Dove" series of the American private company "Planet Labs" can be cited (Fig. 2.).

Fig. 2. General view of a Dove-series nanosatellite in space flight

These nanosatellites were used in the second decade of the XXI century; the mass of each of them is 4.7 kg, and the spatial resolution of the received images is 3.7 m in four spectral channels. (no panchromatic channel). Each of the "Dove" series nanosatellites is equipped with a telescope, digital camera, and software to image different areas of the Earth. They are intended for operational monitoring of deforestation, agricultural territories, harvests, urbanization processes, natural disasters, etc. Their term of active existence is three years. The nanosatellites of the "Dove" series, put into orbit with American and Russian RNs in 2013, provide images from heights of 250 - 580 km, and the dimensions at their launch are 10x10x32 cm (CubeSat 3U).

Since early 2014, Planet Labs has begun deploying an orbital configuration of first-generation Flock nanosatellites, using the Dove series nanosatellites as their prototypes. Nanosatellites of the "Flock-1" series are built according to the "CubeSat 3U" standard and do not have propulsion units (RU). They are characterized by a mass of 4 kg, dimensions of 10x10x30 cm (length, width, and height), an orbit height of 400 km, an inclination $\alpha=52°$, a spatial resolution of 3-5 m, and a period of active existence of 2 years. In the following years of the second decade of the 21st century, the company "Planet Labs" continues to simultaneously launch dozens of nanosatellites of the "Flock-2", "Flock-3" and "Flock-4" series. Thanks to their low cost and the high efficiency of obtaining the data, these KLA are highly suitable for monitoring the environment, crops, forest areas, and emergencies.

By creating a Flock constellation of nanosatellites, Planet Labs intends to obtain operational images from any region of the globe to solve humanitarian, environmental, and commercial tasks. At the same time, the evolution of generations of nanosatellites from the "Flock" family takes place very quickly - practically in a few months. As a result, based on nanosatellites of the "Dove" series, a unique constellation of several hundred MCLAs of the "Flock" series was built, flying in a sun-synchronous orbit (SSO) and providing for one-day imaging of more than 200 million km^2 of the earth's (sea) surface. By 2021, the number of Flock series nanosatellites is approaching 400.

As a result of the modernization of the "Doves/Flock" series nanosatellites, 36 "SuperDoves Flock 4u 1/36" nanosatellites were also included among the 114 SSVs launched by the USA on 01/03/2023 in Earth orbit with the "Falcon-9FT Block-5" launch vehicle" according to the standard "CubeSat 3U" of the company "Planet Labs.", each of which has a mass of 5 kg. These nanosatellites include a telescope and a CCD-sensor camera in their onboard equipment, providing 8-band multispectral imaging and obtaining images with a spatial resolution of up to 50 cm/pixel.

In April 2017, the American multinational company "Google" acquired a stake in the company "Planet Labs" and concluded a multi-year agreement for the purchase of images from the "SkySat" satellites. The "SkySat" family of satellites serves to monitor the Earth's surface in the optical range with sub-meter spatial

resolution and are dual-purpose. They are based on the "CubeSat" concept, but in size and mass exceed those of the "CubeSat 3U" standard (Fig. 3). Their length is about 80 cm, and the mass is not less than 100 kg. Their onboard equipment includes panchromatic, multispectral, and video sensors, with a spatial resolution of 0.9 m, 2 m in four areas of the spectrum and 1.1 m, respectively. For the period 2013–2022, 21 mini-satellites of the "SkySat" family, located in four orbital planes, were launched into Earth orbit.

Fig. 3. General view of the "SkySat-1" optoelectronic observation minisatellite

With the help of these minisatellites, monitoring is carried out in the economic and humanitarian fields.

Another well-known American private operator for the production of small satellites is the company "Spaceflight Industries", whose microsatellites for the UEN "Black Sky" (Fig. 4.) have a mass of 55 kg and a period of operation of 3 years. Through these dual-purpose microsatellites, optical images are obtained initially with a spatial resolution of 50-90 cm from an orbit height of 430 km and an inclination $\alpha=42°$, and later, the spatial resolution reaches up to 30 cm per pixel. As of March 2023, their orbital constellation includes 16 satellites.
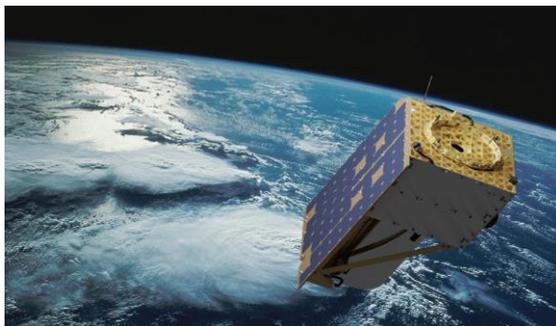

Fig. 4. General view of the Black Sky optoelectronic surveillance microsatellite

According to mass media (mass media), the orbital configuration of the Black Sky SSVs is characterized by the highest frequency of reshooting in the world - 15 times a day. The number of these satellites for the UN is expected to reach 60 SSVs. Monitoring in the economic and humanitarian fields is also carried out with the help of the OEO microsatellites of the Black Sky family.

Along with the discussed small satellites with civil and/or dual purposes, since the 90s of the 20th century and the first decade of the 21st century, the Pentagon has been creating and launching into Earth orbit experimental SSVs, based on which to build space systems with small satellites with a tactical purpose. The management of the program is entrusted to the "ORS" (Operationally Responsive Space Office) management, and the developed "ORS" program aims to accelerate the implementation of innovative technical solutions to shorten the terms for developing and putting tactical KLA into orbit. As of 2012, within the framework of the ORS program and platform, two mini-satellites for tactical optoelectronic reconnaissance (OER) - TacSat-3 and ORS-1 - have been launched into orbit and are in use. More specifically, the TacSat-3 tactical air defense system is equipped with two optoelectronic cameras - hyperspectral and panchromatic. At the same time, the hyperspectral camera allows the detection of masked and hard-to-detect ground objects in the infrared part of the spectrum, and the panchromatic camera is characterized by a maximum image resolution of 1 m in the visible part of the spectrum, which ensures the detection and identification of tactical targets.

Another system for tactical OER is "Kestrel Eye", which includes dozens of small satellites of the same name. According to [3], an orbital configuration of 40 microsatellites of the system can provide the relevant battlefield commander with an interval of 10 minutes and fewer images of the areas of interest with a spatial resolution of 1.5 m in panchromatic mode. The main merits of the MKLA of the "Kestrel Eye" series are their small mass (50 kg), compact dimensions, low cost, and the ability to provide images of the battlefield on mobile devices of servicemen in real-time scale (RTS). The "Kestrel Eye" project was stopped in 2018 at the "prototype" stage, and three new programs were launched on its basis - "Gunsmoke", "Lonestar", and "Polaris". In this way, the efforts to develop new MCLAs for the OEO on the battlefield in the Russian Federation continue.

For observation from space of land and sea objects in any weather and day conditions, the most suitable are the MKLA with on-board radar stations (BRLS) with a synthesized aperture (SA) for radar surveillance (RLN) and radar reconnaissance (RLS). Among them, the American minisatellites and microsatellites for the "Capella" and "Umbra" RLSs, as well as the Finnish microsatellites for the "ICEYE" RLD, stand out at the moment. The listed SSVs are dual-purpose.

The American company "Capella Space" realized its intention to replace the traditional heavy and expensive satellites for air defense and anti-aircraft defense, equipped with radar SA (for example, "Lacrosse") with significantly smaller mass and dimensions MKLA with a relatively low value. For this purpose, the space radar

system of the Capella Space company was created, solving the tasks of ensuring a high periodicity of observation from space and operational delivery of the relevant radar images to users, primarily in the interest of defense and security. This is realized with a phased deployment of an orbital configuration consisting of 36 satellites equipped with a SA radar, which is why the considered space radar system acquires the name "Capella 36", as well as through the possibility of "retargeting" the corresponding satellite with a high-speed SA radar from one shooting object to another object. The beginning of operational use of the Capella 36 space radar system was laid in August 2020, when the Capella 2 satellite with a SA radar with a mass of 107 kg was launched into an Earth orbit at an altitude of 525 km (Fig. 5). In the period January 2021 - January 2022, the "Capella 3, 4, 5, 6, 7 and 8" satellites, which are in operational use, enter Earth orbit. The satellites' orbit inclination ($\alpha=45°$) provides high-frequency monitoring of the Middle East, DPRK, PRC, Japan, and USA regions.

The Capella 36 space radar system allows for obtaining radar images with a spatial resolution of 50 cm. According to SMI, the system can also provide a spatial resolution of 25 cm, but American legislation currently blocks this possibility.

Capella Space is the first commercial operator of SAR satellite imagery in the United States. It is assumed that with its full orbital configuration of 36 satellites (expected by the end of 2023), it is possible to obtain a radar image of any part of the Earth at an interval of no more than one hour.
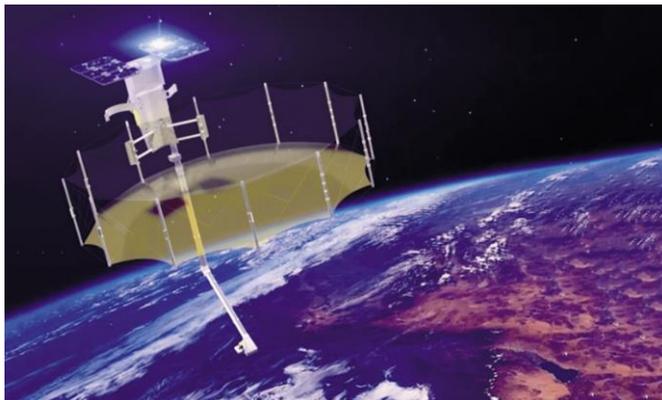
Fig. 5. General view of the Capella 2 radar surveillance satellite (Sequoia)

The American company "Umbra Lab" is creating and launching into Earth orbit a constellation of commercial microsatellites "Umbra", with surface-to-air radar (SAR) operating in the X-band. As of June 2023, three "Umbra" microsatellites with a 70 kg SAR have been launched into Earth orbit, which are dual-purpose. In 2018, this company received a license for permission for a spatial resolution of

25 cm per pixel from the satellites of the same name flying on the SSO at an altitude of 515 km and plans to increase the spatial resolution of radar images to 15 cm per pixel.

The Finnish start-up company "ICEYE" is creating a constellation of "ICEYE" series microsatellites equipped with X-band SA radars. The orbital configuration created by the SSVs mentioned above is intended for the RLS of the Earth, and the resulting radar images have a spatial resolution better than 1 m and are dual-purpose. After the first launch into Earth orbit on 12.01.2018 of a microsatellite with a mass of less than 100 kg, as of January 2023, the orbital configuration of microsatellites of the "ICEYE" series includes more than 20 SSVs.

Research method

The analysis of crises and conflicts in the 21st century from the point of view of their information provision shows that the military mentioned above and civilian SSVs, including dual-purpose ones, successfully complement the strategic space systems for OEE and RLR of the USA, Russia, China, and other leading space states, because the latter provides no more than 30% of the entire volume of intelligence-related information [3].

The theory of systems analysis is a major method of studying the modern use of small satellites for observation.

The analysis of the experience of using SSVs for observation in the XXI century for various applications and their creation concepts shows that the corresponding small satellites can be successfully used to solve mainly three groups of tasks (Fig. 6) - applied, technological, and scientific.

The tasks that the MKLA applies for surveillance can, in turn, be divided into three subgroups - military, civil, and educational.

```
          ┌─────────────────────────────────────┐
          │   GROUPS OF OBSERVATION TASKS        │
          │  SOLVED BY THE SMALL SATELLITES      │
          └─────────────────────────────────────┘
```

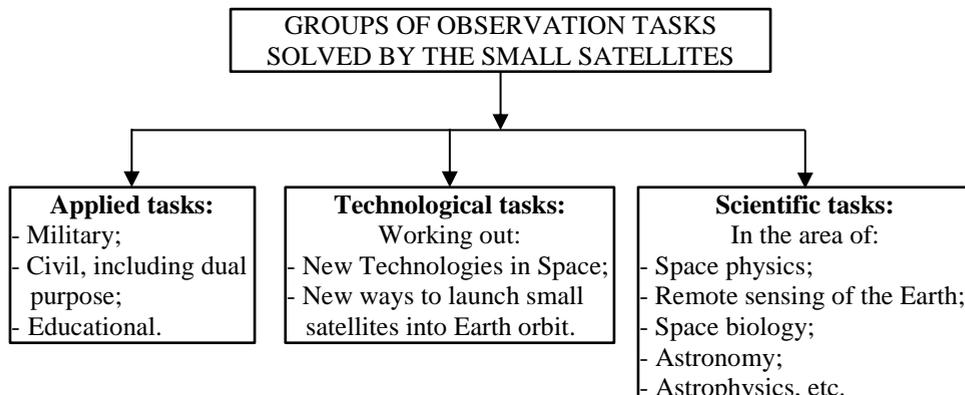| Applied tasks:	Technological tasks:	Scientific tasks:
- Military;	Working out:	In the area of:
- Civil, including dual purpose;	- New Technologies in Space;	- Space physics;
- Educational.	- New ways to launch small satellites into Earth orbit.	- Remote sensing of the Earth;
		- Space biology;
		- Astronomy;
		- Astrophysics, etc.

Fig. 6. Groups of tasks solved by small observation satellites

The military tasks of SSVs for surveillance as part of the applied ones are solved by small satellites for tactical OEO and RLS, as well as commercial small satellites for OEO and dual-purpose RLS.

One of the most important new directions in the development of small satellites is their use as satellite-inspectors [4, 5], since a significant part of them are MCLA, as well as for OER [6, 7].

Several innovative military surveillance ICBMs continue to operate [8], some of which fall into the microsatellite (100–10 kg) and nanosatellite (10–1 kg) classifications. As noted, the SSVs' military surveillance tasks are also handled by commercial small satellites for OEO and dual-purpose RLS. For example, the "Black Sky" minisatellites for the UN discussed above (Fig. 4), in addition to services for private companies, also perform tasks in the interests of the Pentagon. Through these satellites, the US discovered a secret anti-satellite laser weapons test site in China - specifically, two laser installations in hangars with removable roofs. Through "Black Sky" satellites, two bases for launching high-altitude balloons in China were discovered - on Hainan Island and North China.

A major trend in the military space activities of the USA, China, and Russia at present is the development of multi-satellite systems based on SSVs with different purposes and the formation of corresponding orbital groupings (OG). Essential reasons for this are that, according to the Pentagon, large multi-ton satellites (e.g., Lacrosse) are vulnerable to the anti-satellite means of a likely adversary and are economically unprofitable. Against the backdrop of US successes in the use of commercial small satellites, the country's Ministry of Defense (MoD) is considering their use [8] within multi-satellite space systems (clusters of small satellites) for intelligence, communication, and control for early warning of missile attack, for reconnaissance, as well as for active anti-satellite combat (Fig. 7).

Fig. 7. Orbital grouping of small satellites for military purposes

The main reasons for this are that, due to the short production time and relatively low price of SSVs, in the event of a military-political crisis and military conflict, it is possible to quickly replace them in orbit to launch new ones for specific tasks, to restore damaged ones, as well as to increase the number of small satellites on a specific orbit to increase the observation time. In this way, OG can be restored or improved, increasing the resistance of clusters of small satellites to external influences. Along with the stated motives, clusters of a significant number of small satellites provide the necessary periodicity (frequency) in obtaining images, facilitating decision-makers and shortening the process duration.

To implement the above considerations, the US Space Development Agency "SDA" (Space Development Agency) of the Ministry of Defense and the Defense Advanced Research and Projects Agency (DARPA) are developing the "Blackjack" program, which envisages putting a large number of small satellites into orbit around the Earth solving various military tasks - from monitoring and communication to early warning of a missile attack. The program basis is the unification of a single sensor network of small satellites to perform combat tasks, where each SSV performs a specific task.

It should be noted that the well-known private satellite operators Black Sky, Maxar Technologies, and Planet, which have SSVs, entered into a contract in May 2022 with the National Reconnaissance Office (NRO) of the USA for a 10-year delivery of optoelectronic images from its satellites with ultra-high image resolution. More specifically, the American company "Maxar Technologies", also profiled in Earth observation, periodically fulfills orders of the American government, especially the Ministry of Defense, to photograph various crisis and conflict areas (Fig. 8).

*Fig. 8. Image of Russian military equipment involved in the war in Ukraine
in a deployment area, obtained by the company "Maxar Technologies"*

Civil surveillance tasks of SSVs as part of applied tasks are broad-spectrum and are generally aimed at space monitoring in various areas of society's life and

46

activity. So, for example, the following tasks are solved with the help of minisatellites for OEO from the "SkySat" family (Fig. 3): agricultural monitoring, oil infrastructure monitoring, monitoring of sites from the extractive industry; monitoring of natural disasters (detection of victims, coordination of rescue activities, planning of restoration works). Some of the civilian tasks are closely related to the military tasks and have dual purposes.

The ICEYE series of RLS microsatellites provide an effective toolkit for land and sea surface observation in thematic areas such as monitoring land, ocean, and atmosphere; mitigating the consequences of natural disasters and man-made accidents; ensuring safety and assessing the impacts of climate change in other areas as well.

The educational tasks solved by the SSVs for observation are aimed at the student community and, recently also, at the students of the upper classes of high schools. Along with the cognitive nature and the acquisition of new knowledge, the creation of empathy for space technology and technology. These tasks also aim to form the trainee's practical habits in the construction or assembly of the SSVs. In this way, for example, students get the opportunity for 1-2 years to participate in the entire cycle of activities on an actual space project related to the creation and use of the SSVs.

The technological tasks of the SSVs for observation allow the development of new technologies in space, the testing of systems and assemblies of rocket-space technology, including new methods of launching small satellites into Earth orbit (Fig. 6). At the same time, it is the SSVs that make possible the implementation of these technologies and approaches, which is impossible for traditional spacecraft with their mass-dimensional characteristics.

Of the new methods for launching small satellites into Earth orbit, the most common in recent years are three: through the implementation of the "Air Launch" program, from a specialized module on board the ISS, from the `SC board for multiple uses.

Of the modern "Air Launch" programs for putting small satellites into orbit using aerospace systems, the American "Pegasus" and the related program of the American company "Virgin Orbit" [9] are functioning. The Virgin Orbit program continues today, using a new ultralight liquid-propellant rocket.

It should be noted that in 2020, SRTI-BAS developed and adopted a concept for implementing the "Air Launch" project for launching a small satellite (small satellites) into Earth orbit from a carrier aircraft (MiG-29 UB) in the conditions of the Republic of Bulgaria [10–11].

The second method of launching small satellites into Earth orbit was implemented for the first time with the prototype of the microsatellite for tactical OEO "Kestrel Eye" located on the ISS since 08/14/2017. After completing the relevant technological time and preparatory activities, on 10/24/2017, the prototype

microsatellite "Kestrel Eye" was launched into Earth orbit by the Japanese module of the ISS, after which its tests began.

In some cases, cosmonauts or astronauts from the ISS, during a planned exit into open space, manually "launch" small satellites of the "CubeSat" standard into space.

The third way to launch small satellites into Earth orbit was implemented on board the X-37B OTV unmanned minishuttle [12], which performed six orbital missions of increasing duration from 2010 to 2022. According to the press, from the board of this orbital plane, SSVs with different purposes were taken into open space under conditions of secrecy.

Currently, the SSVs also solve several other technological tasks for monitoring.

The scientific tasks solved by the SSVs for observation can be related to conducting research in the field of space physics, remote studies of the Earth, space biology, astronomy, astrophysics, etc. (Fig. 6). For this purpose, most satellites of the "CubeSat" standard include in their PL at least one or two scientific instruments.

Due to its specificity, each of the listed scientific fields imposes its own requirements on the PL of the SSV. The aspects of using small satellites for observation are not the only ones discussed.

Conclusion. Need to use small satellites from the Republic of Bulgaria.

In our country, respectively in SRTI-BAS, it is imperative to study and evaluate the benefits and costs of creating and using small satellites. In case of positive results, the purpose and areas of their application should be formulated for the conditions of the Republic of Bulgaria, young and promising staff should be identified for work on this issue, and their training should be carried out at home and/or abroad (up to the level of assembly of small satellite), to create the necessary laboratory base and to accelerate the development of national and international projects in the various fields of application of small satellites.

Small satellites have become increasingly popular in recent years and dominate the total number of spacecraft launched into Earth orbit, and this trend is currently accelerating.

References

1. Palavenis, D., The growing importance of small satellites in modern warfare: what are the options for small countries?. The Space Review. 12/05/2022. https://www.thespacereview.com/article/4492/1.
2. Romanov, A., A. Romanov. Fundamentals of space information systems. Monograph. M., SRI-RAS, 2019, 292 p., ISBN 978-5-978-5-00015-034-4 (in Russian).
3. Vasilyev, V. et al. Analysis of space reconnaissance capabilities for information support for aviation control when performing fire missions, Aerospace Forces. Theory and practice, March 2021, book. No. 17, pp. 47–56 (in Russian).
4. US military space satellites. Review. VI 2023 (in Russian).
5. Zagorski, N., Military presence of China in space, International Conference on Advanced Research and Technology for Defence ARTDef-2021. 29-30 June 2021. Nikola Vaptsarov Naval Academy – Varna. pp. I-116–I-123. ISSN 2815-2581 (in Bulgarian).
6. Zagorski, N., Space intelligence system in China. International Conference on Advanced Research and Technology for Defence ARTDef-2021. 29-30 June 2021. Nikola Vaptsarov Naval Academy – Varna. pp. I-109–I-115. ISSN 2815-2581 (in Bulgarian).
7. Ivanov, V. et al. Main trends in military space activities at the present stage, M., Aerospace Sphere No. 3 (104) 2020, pp. 72–81 (in Russian).
8. Penev, P., N. Zagorski. Aerospace systems „Aerial launch" at modern stage. Aerospace Research in Bulgaria, S., 2022, Volume 34, pp. 80–93.
9. Yordanov, D., P. Penev. Possibilities for delivery nanosatellites with MiG-29, SRTI-BAS, Space Ecology Safety, Proceeding SES 2020, pp. 124–131 (in Bulgarian).
10. Penev, P., D. Yordanov. The concept for the implementation of the "Aerial launch" project to carry out a small satellite (small satellites) into low-Earth orbit from the carrier aircraft in the conditions of the Republic of Bulgaria, S., SRTI-BAS, ACS section, 2020, 43 p. (in Bulgarian).
11. Penev, P. Space and security. Monograph. Veliko Tarnovo., Publishing House "Vital", 2014. 390 p. ISBN 978-954-8259-94-1 (in Bulgarian).
12. CHALLENGES TO SECURITY IN SPACE, DIA, 2022.
13. AVIATION WEEK & SPACE TECHNOLOGY, 2018-2023.
14. https://spacenews.com
15. https://flickr.com/space_foundation
16. https://aboutspacejornal.net
17. https://vpk.name/
18. https://topwar.ru
19. https://vz.ru

РАЗЛИЧНИ АСПЕКТИ НА ИЗПОЛЗВАНЕТО НА МАЛКИ СПЪТНИЦИ ЗА НАБЛЮДЕНИЕ НА СЪВРЕМЕННИЯ ЕТАП

П. Пенев

Резюме

Разкрити са причините за възникване на малките спътници. Разгледана е еволюцията на създаваните през XXI век военни и граждански малки спътници за наблюдение и разузнаване, включително с двойно предназначение. Дефинирани са три групи задачи, решавани от малките спътници за наблюдение – приложни, технологични и научни, както и нарастващото им значение и сфери на използване. Обоснована е необходимостта и са предложени стъпки в Република България, респективно в Института за космически изследвания и технологии към Българската академия на науките, да се активира дейността по създаването и използването на малки спътници за наблюдение за нуждите на икономиката, екологията, хуманитарната и в други области.

ON THE DELAY OF THE ADDITIONAL MORTALITY LINKED TO THE GEOMAGNETIC DISTURBANCES

Tsvetan Georgiev, Siyka Simeonova, Luba Dankova[1]

[1]Institute of Astronomy and National Astronomical Observatory –
Bulgarian Academy of Sciences
e-mail: tsgeorg@astro.bas.bg

Keywords: *Geomagnetic disturbances, Sun- mortality relationship*

Abstract

Geomagnetic disturbances, mainly geomagnetic storms (GMSs) but also low-frequency resonances, touch some people susceptible to cerebrovascular diseases (CVDs). Sometimes, the geomagnetic effect is overestimated speculatively. Against this concept, we compare the changes of geomagnetic indexes (GMIs) with the changes of the additional mortality rate (AMR). We compared employing cross-correlation functions (CCFs) and use the Wolf number (WN) as a referent time scale. We suspect that strong GMSs, like these in 2003, increase the relative common MR 3–4 years later by up to 4×10^{-5}. Otherwise, the typical GMS-linked AMR seems less than 10^{-5}. Even if these values are overestimated, generally, they are small. Analyzing data about Bulgaria and five of its regions for the last Solar cycles, we confirm that the lag of the maxima of the GMS-linked ANR behind the WN maxima is ~5 years. We also confirm that the lag of the GMSs maxima behind the WN maximum is 1–2 years. We found that the lag of the maxima of the CVD-linked AMR behind the maxima of the GMSs is 3–4 years. So, we consider the 5-year lag of the AMR linked to the GMSs behind the WN maximum appears a sum of two delays mentioned above, 1–2 years and 3–4 years. In principle, the typical duration of CVDs may be derived if the beginnings are known. In medicine, they are usually unknown. However, suspecting the GMSs as triggers of a part of the CWD-linked AMR, we should suppose that these CWDs finish with lethal outcomes after 3–4 years.

Introduction

Usually, the moments of geomagnetic activity are referred to as the time scale of the Wolf number (WN). The WN *W* is defined as a relative number of sunspots. The number and the intensities of the high-energetic solar processes that affect the Earth correlate with the WN.

The changes of the speed and density of the solar wind (due to the flares and coronal mass ejections) cause geomagnetic disturbances (GMDs) for up to several hours. The GMDs, especially the geomagnetic storms (GMSs), affect many processes on Earth, including human health. The GMDs are characterized by geomagnetic indexes (GMIs, Section 2). The GMS maxima lags behind the WN

maxima by 1–2 years ([1], Figs. 17, 19; [2], Figs. 5, 6; [3], Fig. 1; [4], Fig. 3). In this work the lag of the GMS maxima behind the WN maxima is found to be also 1–2 years, see Figs. 4, right, and 6.

The proton concentration Np above the Earth's atmosphere is due mainly to the galactic cosmic rays varies. While the solar activity is high, the solar wind suppresses the galactic cosmic rays, and Np is low (effect of Forbush). When Np is high and variable, it creates low-frequency electromagnetic resonances (LFRs) in the chamber between the Earth's surface and the ionosphere (Schumann resonances). When the resonance frequency is very low, 1–2 Hz, it may be somewhat dangerous for the heart rhythm of some people [5, 6]. The LFR maxima lag behind the WN maxima, depending on the solar wind intensity, is 4–7 years. It takes place over and after the WN minima, see Figs. 2a and 6d.

The GMDs are linked to some health outcomes, connected mainly to the cerebrovascular diseases (CVDs) – coronary heart diseases, myocardial infarction (MI), brain stroke (BS), etc. GMDs are also linked to neurological system diseases, behavioral diseases, etc. In principle, CVDs cause half of the common mortality rate (MR) worldwide. However, in this paper, we concentrate on the additional MR (AMR) caused suggestively by the GMDs.

Usually, the studies concentrate on the correlation between solar activity and MIs and/or BSs. Many pieces of evidence exist about the negative influence of GMDs and LFRs on physiological and psychological human health [7–10]. For example, during days with GMS, the additional BSs and MIs suffer in Moscow, grow by 7.5% and 13%, respectively. (See the references of the Russian studies in [11]). It is also established that the GMDs caused by solar magnetic clouds are related to the increase in MI. The last-mentioned connection is higher than the GMDs caused by high-speed solar wind streams and on days with quiet geomagnetic activity [12].

Ionosphere and geomagnetic changes influence mortality from circulatory diseases. The CVDs' response to the changes in solar activity and abnormal solar events indirectly influence the concentration of electrical charges in the Earth's environment [13]. The different patterns in daily numbers of deaths during the quiet periods of solar activity are examined later. It is shown that there is a connection between the daily number of deaths and all indices of solar and geomagnetic activity in periods of low solar activity, in contrast to periods of strong solar storms [14].

The relationships between GMDs and the time course and lags of autonomic nervous system responses have been examined in [14]. It is confirmed that the daily nervous system activity responds to GMDs. The response is initiated at different times after the changes in the various environmental factors and persists over varying time periods. An increase in the solar wind, cosmic rays, solar radio flux, and Schumann resonance power was associated with increased heart rate and parasympathetic activity, interpreted as a biological stress response. The people are

52

affected in different ways depending on their sensitivity, health status, and capacity for self-regulation. The impact of short exposure to GMDs on total and cause-specific MR in 263 US cities has been investigated recently [3]. The GMDs and LFRs lead to an increase in city-specific and season-stratified common MR in all cities. The effects on total deaths were found in all seasons and on CVD and MI deaths – more in spring and autumn. These results may be explained through the direct impact of environmental electric and magnetic fields produced during GMDs and LFRs on the human autonomic nervous system.

In a review of the health effects of GMDs, Palmer et al. (2006) [16] reported five definite conclusions: (1) GMDs have a greater effect on humans at higher geomagnetic latitudes. (2) Unusually high geomagnetic activity seems to have a negative effect on human cardiovascular health. (3) Unusually low values of geomagnetic activity seem to have a negative effect on human health. (4) Only 10–15% of the people are negatively affected by GMDs, and (5) heart rhythm variations are negatively correlated with GMD. In this paper, we confirm that the lag of the AMR maxima behind the maxima of the WN is about five years, see Fig. 2. We propose an explanation of this "paradox". See our Summary.

Earlier, we found correlations of the cause-specific CVD AMR linked with GMSs for the Smolyan region of Bulgaria [11]. We found that concerning the years with low CMDs (1993, 1995, 1996, 1999), in the years with strong GMSs (2000, 2001, 2003–2005), the AMR is higher with 20–30% and the MR related to CVD is higher with 30–40% ([11], Figs.10–13). We also noted that the time delay of the maximum of the common and CVD AMRs in 2007–2008 takes place about 3–4 years behind the maximum of the strong storms in 2003–2005, see Fig. 1a. The increase in AMR and in common MR is about 50% and 5%, respectively. This result suggests that the influence of the GMSs on the AMR may manifest itself 3–4 years later. This is the motivation of the present work.

This paper reveals time delays of the AMR maxima by maxima positions of cross-correlation functions (CCFs). The CCF measures the similarity between the structures of two time series. It is a function of their relative lag time t_L, see Fig. 2 and 3. The large-scale trend in the series is an obstacle and must be removed preliminary. Fortunately, all time series used here pose linear large-scale trends. For example, in Fig. 2, we juxtapose the WN, W, and a few kinds of MR, M, after removing their linear fits. Thus, the CCF uses the deviations, for example, $\Delta W = W - W_F$ and $\Delta M = M - M_F$, where W_F and M_F are the relevant linear fits of the time series. The deviation ΔM is just the AMR.

So, in this work, the maxima positions of the CCFs are used to determine the time lags of the time series Θ, see Fig. 6 to 11. The CCF maximum is characterized by its value C and standard error σ_C. The CCF maxima in this paper are not high, often about 0.6, but their standard errors are relatively small. Then, the Student ratio $R = C/\sigma_C$ is usually high, giving evidence that the CCF maxima are significant, see Fig. 4 to 11.

Note that a graphical representation of the Student criterion for such cases as in [17], Fig. 4, shows how the threshold increases when the data number decreases. So, 20–10 data the 95% threshold is 0.6–1.1 and the 99% threshold is 0.9–1.6. Sometimes, our ratio R overcomes the threshold. Note also that the GMSs, LFRs, and CVDs have a highly complex origin and nature, which is out of the subject of this work. We are interested mainly in the cross-correlations between the deviations from the GMI fits (reasons) and the deviations from the MR fits (results), regarded as AMR.

Used abbreviations follow.

AMR – additional mortality rate;
BS – brain stroke;
CCF – cross-correlation function;
CVD – cerebrovascular disease;
GMD – geomagnetic disturbance;
GMI – geomagnetic index;
GM – geomagnetic storm;
LFR – low-frequency resonance;
MI – myocardial infarction;
MR – mortality rate;
NI – NASA (planetary) indexes (Section 2: *B, Kp, Ap, Np;* Fig. 5 and 6);
PI – Panagyurishte (local) indexes (Section 2: *Sa, Sb, Sc, Sn*; Fig. 4);
WN – Wolf number of the sunspots.

1. Data about the MR. Lags of the ANRs behind the WNs.

Figure 1 shows the behaviour of the habitant numbers N (circles) and MRM (dots) over the years. The numbers N are expressed in specific (implemented) units. The numbers M are always expressed per mile (in 10^{-3} N). Hereafter, the straight lines show the fits, while g and s are the gradient (slope coefficient) and standard deviation of the fit.

Figure 1a shows the initial data about the Smolyan region, Rodopi Mountain, between 1988 and 2008 [11] (Sun cycles 22+23). These data are valuable because they contain only the common MR M_1 and the medically confirmed CVD MR M_0. The data content follows:

N_1 – habitant number;
M_1 – common MR;
M_0 – MR linked medically with CVD;
M_D – residual MR, $M_D = M_1 - M_0$.

The residual MR M_D corresponds to deaths caused by other reasons, including LFRs.

In the time episode of Fig.1a, N_1 decreases 1.3 fold, mainly due to young people's migration. Simultaneously, M_1 increases 1.5 fold, mostly because the

54

population gets older. The increases of M_0 and M_D are about 1.4 and 1.3 fold, respectively. In Fig.1a, two squares show an extraordinary maximum of the CVD MR M_0. Two large dots show respective deep minima in the residual mortality M_D. (The data marked by squares and dots do not participate in the relevant fits.)
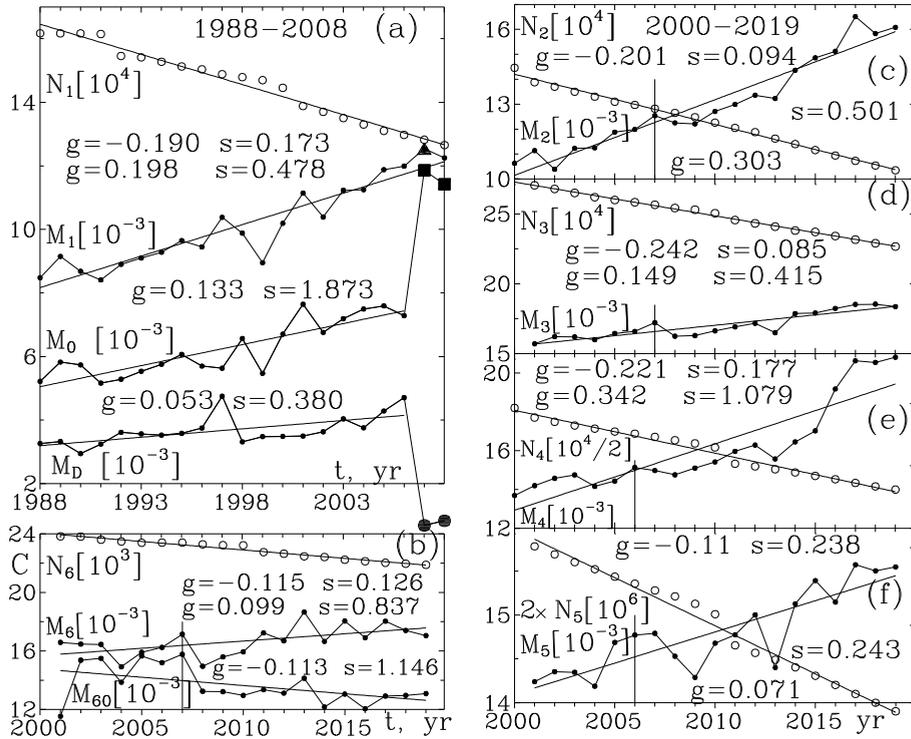
Fig. 1. Annual data about the habitants N and mortalities M for five Bulgarian regions plus Bulgaria as a whole. See the text at the beginning of Section 1.

Figures 1b–1f show five MR data systems for the time episode from 2000 to 2019 (Sun cycles are 23+24). The data source is the National Statistical Institute of Bulgaria [18]. The data content follows:

N_2, M_2 – again for Smolyan region, Fig. 1c;

N_3, M_3 – for Sofia suburb (without Sofia city), Fig. 1d;

N_4, M_4 – for the region of Dobrich plus Silistra together, Fig. 1e;

N_5, M_5 – for Bulgaria as a whole, Fig. 1f;

N_6, M_6, M_{60} -- for the eastern part of the Sofia suburb, Fig. 1b.

In the time episode of Figs. 1b–1f all habitant numbers N_2–N_6 decrease and all common MRs M_2–M_6 increase. The reason is the same as in Fig. 1a. The region in Fig. 1b (namely Elin Pelin) covers about 1/10 of the habitants of the Sofia

suburb, but it is also valuable here. It contains two kinds of common MR. M_{60} is recorded only inside the territory of this region, and it decreases. M_6 is the common MR, containing M_{60} plus the number of deaths of habitants of this region, but recorded in the nearby big hospital in Sofia. As expected, M_6 increases.

Let us return to Fig. 1. There, we may estimate the extreme and the ordinary AMR linked with the GMS. The extraordinary CVD MR M_0 in 2006–2007, after the strong GMSs in 2003, exceeded the local MR M_0 by ~50% (Fig. 1a). The relevant small peak in the common M_1 exceeds the local MR M_1 with ~4%. The vertical segments in Figs. 1b–1f mark the respective small local peaks of the common MR in 2006–2007 yr. The height of these peaks, including for Bulgaria as a whole, is up to ~4% above the local MR.

So, the strong GMSs (in this single case) seem to trigger an increase of the common MR up to $\sim \Delta M = 0.04 \times 10^{-5} = 4 \times 10^{-5}$, with a time lag of 3–4 years. Otherwise, the typical CVD AMR, linked with GMSs, seems to be up to 1×10^{-5} per year. Because of unknown random AMR contributions, these AMRs seem to be overestimated, though. These AMR values seem to be negligible. For comparison, the number of deaths by car accidents for Bulgaria as a whole in 2017 is $\sim 10 \times 10^{-5}$. Both values are overestimated and need justification.

The deviations of the MRs from the linear fits in Fig. 1a, ΔM_1, ΔM_0, or ΔM_D, are the AMR. After fit removal, these deviations participate in 3 important CCFs in Fig. 2a. They distinct the supposed contributions of CVD linked MR, common MR and common MR minus CVD linked NR LFR MR. The deviations of the MRs, $\Delta M_1 - \Delta M_5$, clearly show the delay of the AMR linked with the GMDs with respect to the WN maxima.
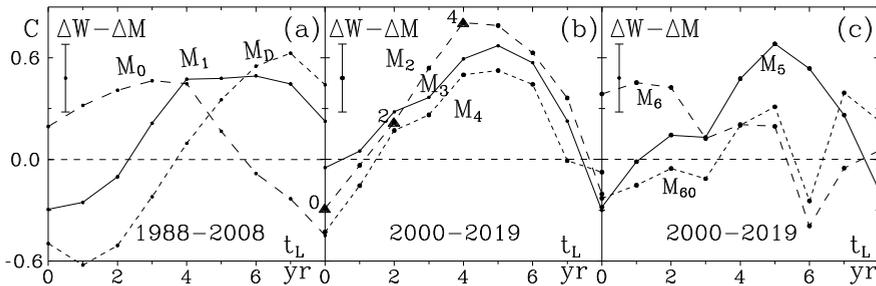
Fig. 2. CCFs between the changes of the WNs and fluctuations of the AMRs. The maxima position shows the delay of the AMR. See the text in Section 1.

Fig. 2 juxtaposes CCFs $C(t_L)$ between the WNs (shown in Fig. 5) and AMRs (shown in Fig. 1) over the time lag t_L. The CCF maxima mark the time lag of the AMR maxima behind the WN maxima. Note that linear fits of the compared time series are always removed. Vertical segments show the typical error bars of

the CCF values. Triangles in (b) mark the CCF values whose value deriving are illustrated in Fig. 3. Details about these graphs are shown in Figs. 7–11.

Fig. 2a shows the CCFs of WNs with the data M_1, M_0, and M_D from Fig. 1a. The CCFs have similar shapes. They show maxima lags behind the WNs of 3, 5, and 7 years, respectively. At about five years, the middle maximum corresponds to the common AMR. However, the maxima at 3 and 7 years may be linked to displays of CVD ANRs and LFR AMRs. In both cases, some additions of deaths for other reasons are present undoubtedly.

Fig. 2b shows the CCFs for the M_2, M_3, and M_4 data in Fig.1c–1e. The CCFs have similar shapes again. Their maxima show lags behind the WN for about five years. Hints of humps in the left parts of the CCFs, about a lag of two years, seem linked with CVD AMR. Triangles mark the CCF values derived as the coefficient of correlation, as illustrated in Fig. 3.

Fig. 2c shows by solid lines the CCF for the AMR of Bulgaria as a whole, M_5 from Fig. 1c. The shape is similar to the mentioned shapes of CCFs, with a peak lag of five years behind the WN. This CCF shows a local convexity at a lag of about two years, which ought to be linked with the contribution of CVD AMR. Figure 2c shows, by dashed broken lines, the CCFs for the common AMRs M_6 and the territory-bounded AMR M_{60}. These curves are very different. The left part of CCF for M_{60} is flat, as if CVD AMRs are missing. Obviously, a significant number of CVS AMP happen out of the territory, in the nearby big hospital. A remarkable hump is present in the left part of the CCF, only in the common AMR M_6. It seems CVD AMR dominates in this region.

In Fig. 2, the humps at lags of about five years behind the WRs contain contributions from the CVD, LFR, and other AMRs. It may be seen well in M_0, Fig. 2a, with a 2- to 4-year lag. In the other cases suspected, in M_2–M_5 with a lag of about two years and in M_6 with a 1- to 2-year lag.

So, if the strong GMDs are regarded as triggers of a part of the CVDs, with postponed lethal outcomes, then the lags of the GMD AMR behind the WNs, as well as the lags of AMRs behind the GMDs, may be revealed. In Section 2, we derived lags of GMIs behind the WNs. In Section 3, we showed details of deriving the CCFs, as shown in Fig. 2.

Fig. 3 shows the derivations of the CCF values at t_L-lags of 0, 2, and 4 years, marked in Fig. 2b by triangles. The top panels show the shifts of the shape of the WNs (thick broken lines) over the shape of AMR (thin broken lines). The linear fits of the compared series are removed. Here, n is the number of currently used points. Dashed broken lines show the useless edges of the time series after the shifts of WNs. The bottom panels show the respective correlation diagrams and CCF values C. Solid and dashed lines represent direct and reverse linear fits. Note that because of the large range of the WN, the compatibility of the graphs in the top panels is difficult. For this reason, the WN values $W'=W^{0.5}/20$ are used. This is

admissible because the values of the CCFs are dimensionless. Note that the values of C in the case (c$_2$), after a suitable mutual shift, become significant.
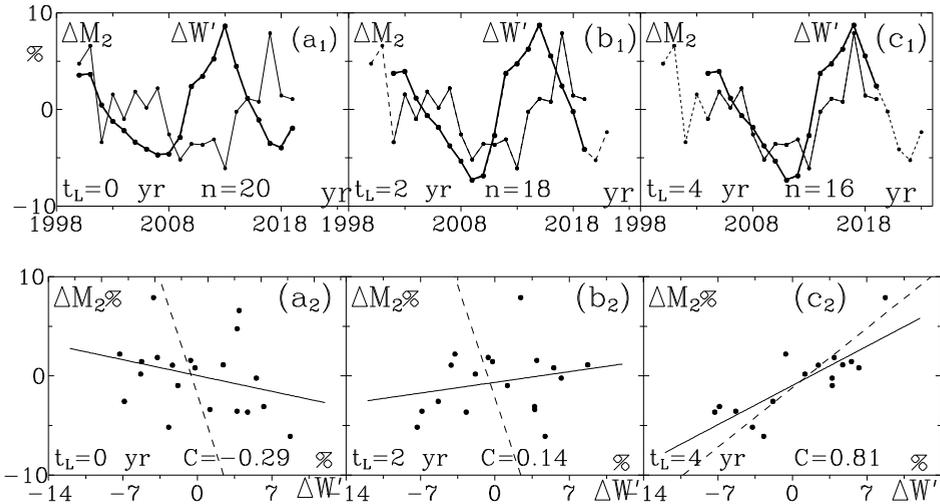

Fig. 3. illustrates the action of the CCF and the sources of the CCF values at points 0, 2, and 4 in Fig. 2b

2. Indexes about the GMDs. Lags of the GMDs behind the WNs

Initially, local GMIs in 1988-2008 were acquired from the Panagyurishte Geomagnetic Observatory of Bulgaria [11]. The Panagyurishte indexes (PIs) used here are derivatives of the measured vertical component H of the local geomagnetic field. The used annual PIs are:

Sa – average amplitude of all storms, in nT;

Sb – average amplitude of the moderate and strong storms, for $H>120$ nT;

Sc – average amplitude of Sb-type storms, but with sudden onset, in nT;

Sn – number of all storms.

Fig. 4 shows the behaviour of the PIs and their CCF with WNs. Figure 4, left graphs, represents the behaviour of the PIs and their trends over the years. Hereafter, g and s are the fits' gradient and standard error. The graphs cover the Sun cycles 22+23. In this episode, the everyday solar activity decreases (see Fig. 4d$_1$ and 5a$_1$–5e$_1$), but the large-scale trends of Sa and Sb are slightly positive. The powerful GMSs in 2003–2005 caused high peaks in the graphs of Sa–Sc. Fig. 4, the right graphs, show the CCFs $C(t_L)$ of the PIs with the WNs over the time lag t_L. Hereafter, Θ and C are the delay and the value of the CCF maximum. $R = C/\sigma_C$ is the Student ratio. The CCF maxima are relatively low and blunt. The time lags Θ of the CCFs maxima behind the WN maximum are 2, 1, 2, and 0 years,

respectively. Note that the number of all GMSs *Sn* reflects the general decrease of the solar activity and obeys the behaviour of the WN trend (Fig. 4d$_1$, 4d$_2$). The PI *Sn* is useless in the study of AMR delays. Therefore, we may consider the lag of the GMSs behind the WN maxima to be roughly two years.
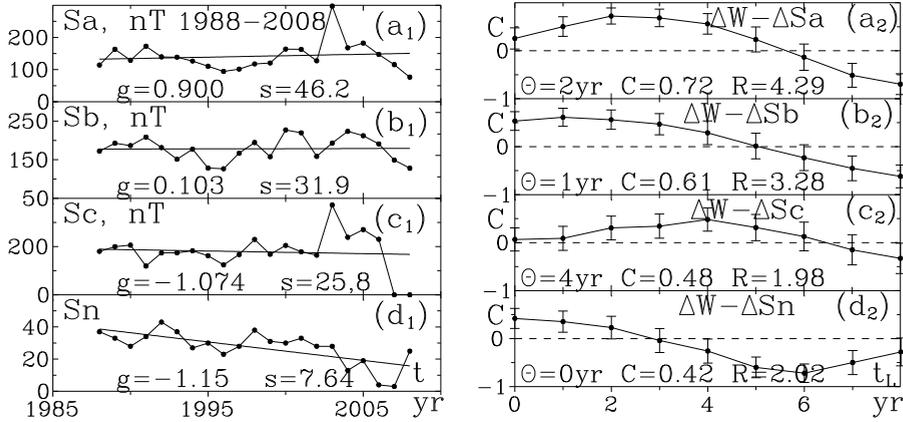

Fig. 4. PIs in Solar cycles 22+23 and their CCFs with the WNs. See the text.

We also used GMIs for 1988-2008 and 2000-2019 from the NASA website [19]. The NASA (planetary) annual indexes (NIs) used here are:

W – Wolf number of the sunspots;

Bm – the scalar value of the Earth's magnetic field, in nT;

Kp – GMI that characterizes the fluctuations of the electromagnetic field due to the GMSs;

Ap – GMI like *Kp* and approximately proportional to log *Kp*;

Np – proton concentration above the Earth's atmosphere, in cm^{-1}.

The NIs *Kp* and *Ap* indicate indirectly the powers of the GMSs. See [20].

Fig. 5 represents the behaviour of the NIs and their trends over 1988–2008 (cycles 22+23, left graphs) and 1998–2019 (cycles 23+24, right graphs). The right graphs show that the decrease of the common solar activity continues, but with decreasing Np in Fig. 5d2, it even increases weakly. The power storms in 2003 are observed as peaks in *Kp* and *Ap*.

Fig. 6 shows the CCFs of the Nis and the WNs. The CCFs pose again blunt maxima. The lags Θ of the NIs behind the WNs are 0, 0, 0, and 6 years in the left graphs and 1, 2, 2, and 4 years in the right graphs. The lags of the NIs *Bm, Kp,* and *Ap* behind the WNs in both cases may be considered to be between 0 and 1 year or 1 and 3 years. Further, the contributions of the AMRs, linked to different GMIs, will be regarded separately.
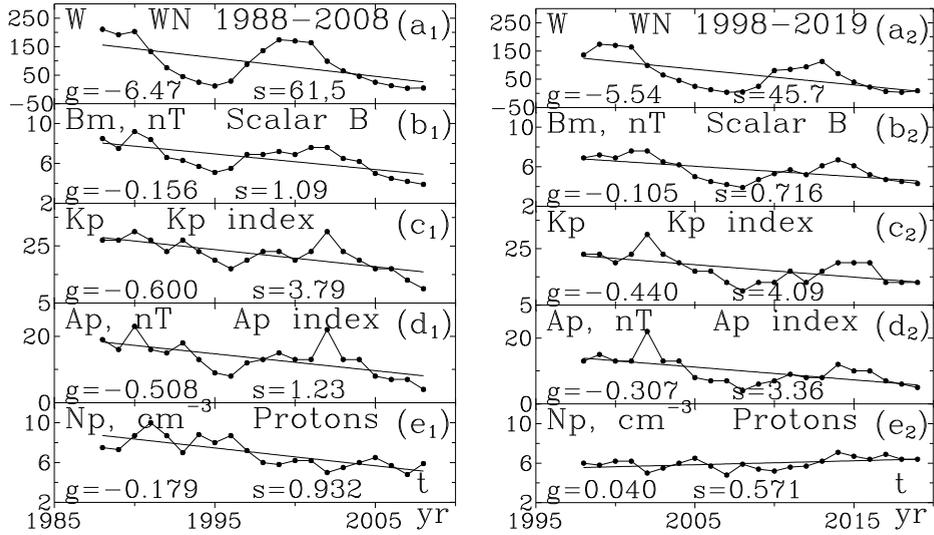
*Fig. 5. NIs in Solar cycles 22+23 (left graphs) and 23+24 (right graphs).
See Fig. 4, left graph, and the text.*
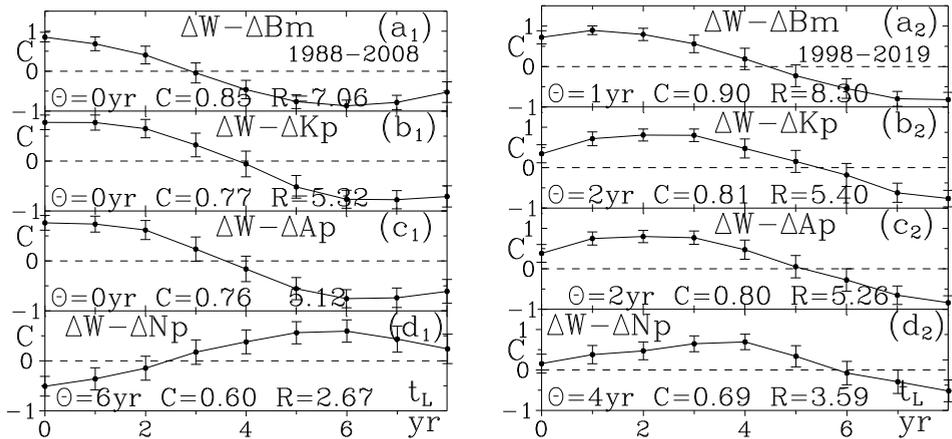
*Fig. 6. CCFs between NIs and the WNs for 1988–2008 (left graphs) and for
1998–2019. See also Fig. 4, right graphs and the text.*

The proton concentration Np is very interesting. Its maxima lags in Fig. $6d_1$ or Fig .$6d_2$ behind the WN are from 5 to 7 years or 3 to 4 years, respectively. In Fig. $6d_1$, this is the effect of higher solar activity (and higher Forbush effect). Then,

the LFR linked AMR contribution is distinct, as in Fig. 2a. In Fig. 6d$_2$, this is the effect of lower solar activity (and lower Forbush effect). Then, the LFR linked AMR contributes to the ordinary CCF maxima in Fig. 2a–2c, i.e., in the region of WN minima. For this reason, the maximum in M_2 in Fig. 2b is higher than the maximum in M_1 in Fig. 2a.

3. Lags of the AMRs behind the GMIs. Explanation of Fig. 2.

Fig. 7a to 7d show the CCFs between the common AMR M_1 and four GMIs for 1988–2006. The lags Θ of the AMR maxima are 4, 3, 0, and 5 years (left graphs with PIs) or 5, 5, 5, and 2 years (right graphs with NIs). Fig. 7e$_1$ and 7e$_2$ show the CCF between M_1 and WN (thick broken curve), as in Fig. 2a. Here, the beginning parts of the CCFs in (a)–(d) (thin dashed broken curves) are implemented. Note that the added curves are shifted to the right in accordance with their specific lags behind the GMI maxima in Figs.4b and 6a. So, the systems of added curves describe approximately the hump of the "main" CCFs (thick broken curves), better in 7e$_2$. The maximum of the common AMR is situated at ~5 yr behind the WNs maximum.
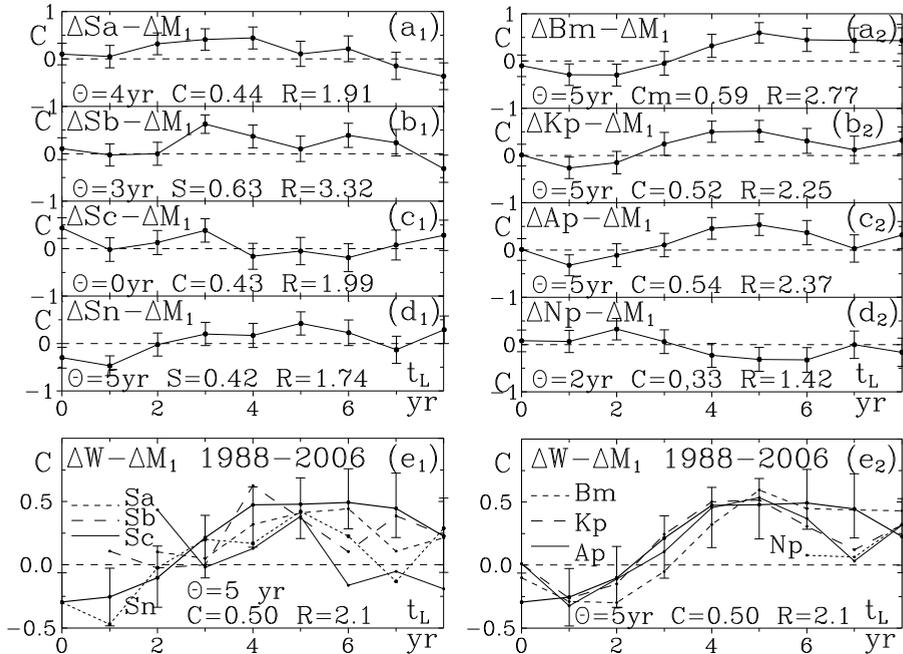

Fig. 7. CCFs for the common AMRs M_1 in Fig. 1a, with PIs (left graphs),
NIs (right graphs) and WN (bottom graphs). See Fig. 4, right graphs and text.

Figures 8a–8d show the CCFs between the CVD linked AMR M_0 and 4 GMIs for 1988–2006. The lags Θ of the maxima are 0, 0, 0, and 3 years (left graphs, with PIs) or 3, 2, 2, and 7 years (right graphs with NIs). Fig. 8e$_1$ and 8e$_2$ show the CCFs between M_{10} and WN (thick broken curves), as in Fig. 2a. Similar to Figs. 7, the shifted beginning parts of the CCFs in (a)–(d) are implemented in (e) (thin and dashed broken curves). The systems of added curves again describe approximately the hump of the "main" CCFs (thick broken curves), better in 8e$_2$. The maximum of the CVD linked AMR is situated at ~three years behind the WNs maximum. It seems that this hump contains a significant contribution from CVD linked AMRs.

Fig. 8. CCFs for the CVD linked AMRs M_0 in Fig. 1a, with PIs (left graphs), NIs (right graphs) and WN (bottom graphs). See Fig. 4, right graphs and text.

Figures 9a–9d show the CCFs between the residual AMR M_D in Fig. 1a and four GMIs for 1988–2006. The lags Θ of the maxima are 3, 6, 3, 5 yr (left graphs) or 6, 4, 6, 7 (right graphs). Figures 9e$_1$ and 9e$_2$ show the CCFs between M_D and WN (thick broken curves), as in Fig. 2a. Similar to Figs. 7 and 8, the shifted beginning parts of the CCFs in (a)–(d) are implemented (thin and dashed broken curves). Again, the systems of shifted AMRs describe approximately the position of the hump in the "main" CCFs (thick broken curves), better in (e$_2$). The LFR

linked AMR's maximum is ~seven years behind the WNs maximum. It seems that this hump contains a significant contribution from LFR linked AMRs.
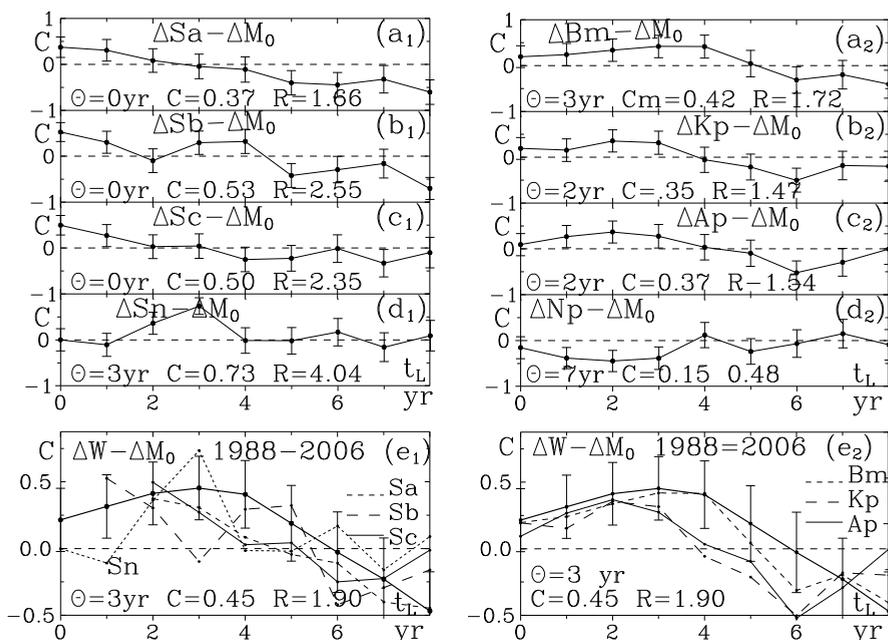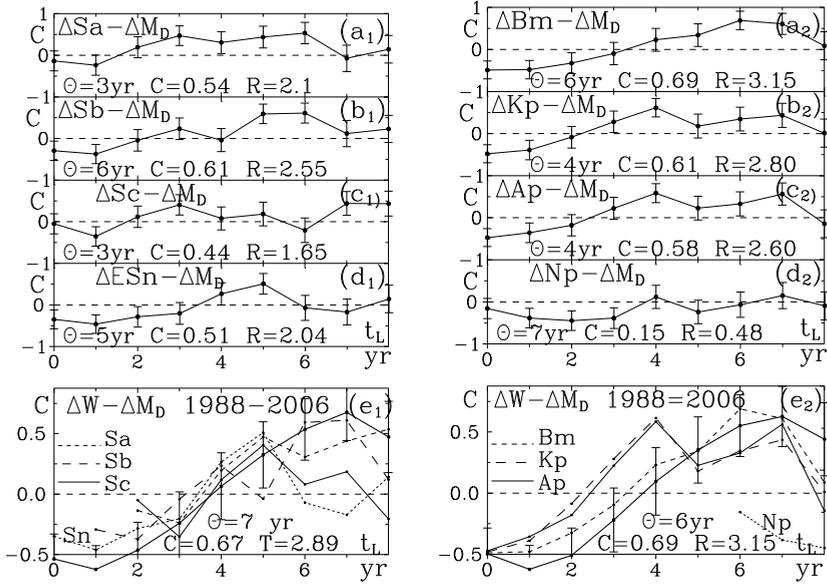

Fig. 9. CCFs for the LFR linked AMRs M_0 in Fig. 1a, with PIs (left graphs), NIs (right graphs) and WN (bottom graphs). See Fig. 4, right graphs and text.

Fig. 10 shows the CCFs between the common AMRs M_2 and M_3 with 4 NIs for 2000–2018. The lags Θ of the maxima behind the WN maxima are 3, 3, 3, and 2 (left graphs) and 3, 1, 3, and 2 (right graphs). The typical lag is ~three years. Again, the graphs in 10a–10d are implemented in 10e, shifted in respect to the lags behind their NIs in Fig. 6b. Obviously, the systems of added curves describe well the hump of the "main" CCFs (thick broken curves). The common AMRs' maxima in Figs.10e are 4 to 5 years behind the WN maximum.

Fig. 11 shows the CCFs between the common AMR M_4 and M_5 with four NIs for 2000–2018. The lags Θ of the maxima are 4, 3, 3, and 2 (left graphs) and 4, 4, 3, and 2 (right graphs). The typical lag is ~ three years. Again, the graphs in 11a–11d are implemented in 11e, shifted in respect to the lags behind their NIs in Fig. 6b. Obviously, the systems of added curves describe well the hump of the "main" CCFs (thick broken curves). The common AMRs' maximum is about five years behind the WN maximum.

The humps of the CCFs in Figs. 7e–11e1 which are explained here, are shown together in Fig. 2. These examples show that the lag of the AMR is about five years behind the maximum of the WN. Simultaneously, this lag may be regarded as a sum of two lags: a 1- to 2-year lag of PI or NI with respect to WN

plus a 3- to 4-year lag of the CVD linked AMR behind the CMIs. The hypothesis appears that the GMDs may be triggers of a part of the CVDs with a postponed lethal outcomes.
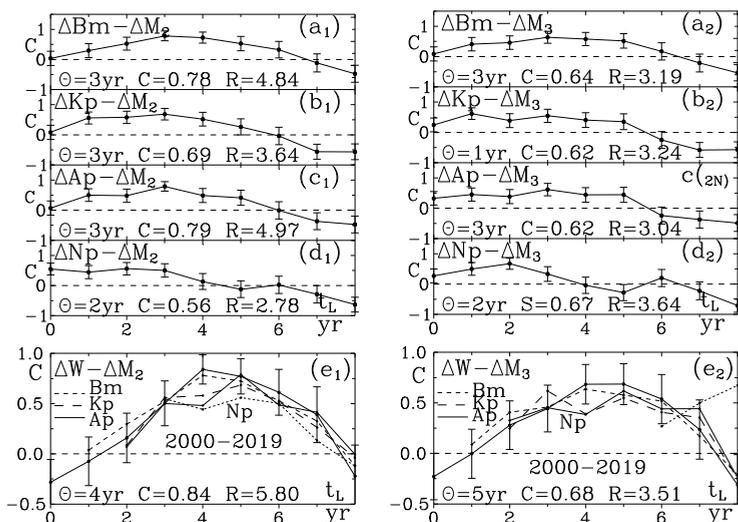

Fig. 10. CCFs for the common AMRs M_2 and M_3 in Fig. 1b, with NIs (left and right graphs), NI, and WN (bottom graphs). See Fig. 6, right graphs and text.
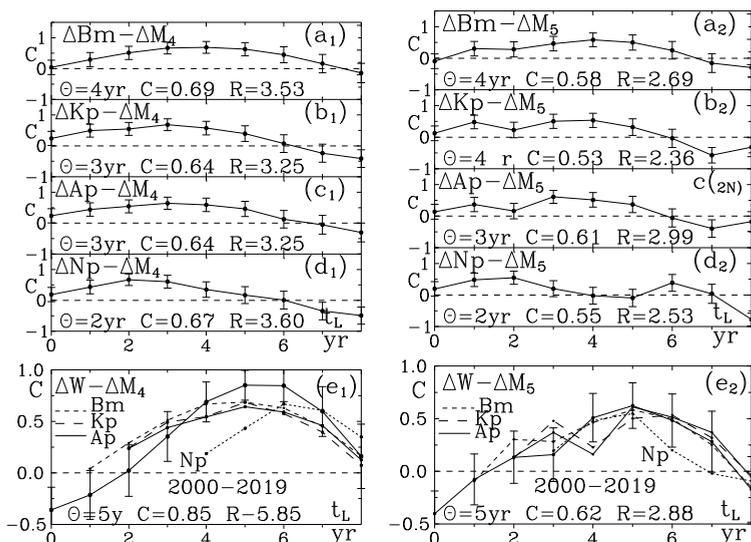

Fig. 11. CCFs for the common AMRs M_4 and M_5 in Fig. 1b and 1c, with NIs (left and right graphs) and WN (bottom graphs). See Fig. 6, right graphs and text.

64

4. Summary

The main results follow.

1. We confirm our suggestion in [11] that the strong GMSs in 2003 caused 3–4 years later, in 2006-2007, an annual increase of the CVD linked AMR with ~50% (Fig. 1a). The local increase in the common MRs is ~4% (Figs.1b–1f). Therefore, the strongest GMS may increase the common MR by about 4×10^{-5} with a lag of 3–4 years. Otherwise, the typical CVD linked AMR seems to be up to 10^{-5}.

2. Based on common AMR data M_1–M_5, for five regions, including Bulgaria as a whole, we showed that the shapes of the CCFs between WN and common AMR are very similar (Fig. 2). The lag of the maxima behind the WN maxima is about five years. So, the common AMR maxima fall on the WN minima, confirming consideration No.3 in [16] (see introduction).

3. We confirm that the lag of the GMSs maxima behind the WN maxima is 1–2 yr (Figs. 4b, 6). We find also that the lag of the CVD ANR maxima or LFR AMR maxima behind the WN maxima is 3–4 years or 4–7 years, respectively (Fig. 2 and 7–9). Therefore, consider the lag of the maxima of the common ANR behind the WN maxima, typically five years, which may be explained as a sum of the above-mentioned lags of 1–2 and 3–4 years.

4. Considering GMSs as triggers for CVDs, we estimated that the duration of the CVDs before the lethal outcome is typically 3–4 years. In medicine, the causes of CVDs are not known, and such direct estimation is impossible.

Acknowledgments

References

1. Hathaway, D. H., (2010), Reviews in Solar Physics, Volume 7,
    The Solar Cycle (Springer).
2. Zerbo, J.-L., C. Amory-mazaudier, F. Ouattara, (2013),
    Adv. Res. 4/3, 265–274.
3. Zili Vieira, C. L., D. Alvares, A. Blomberg, et al., (2019),
    Environmental Health 18, No. 8   3.
4. Abe, O. E., M. O. Fakomiti,, W. N. Igboama, et al., (2023),
    Adv. Sp. Res. 71, 2240–2251.
5. Southwood, D., (1974), Planet. Sp. Sci. 22, 483–491.
6. Thayer, J. F., A. L. Hansen, E. Saus-Rose, et al., (2009)
    Ann. Behav. Med. 37, 141–153.
7. Friedman, H., R. O. Becker, C. H. Bachman, (1967), Nature 213, 949.
8. Gnevyshev, M. N., K. F. Novikova, (1972),
    J. Interdiscipl. Cycle Res. 3, 99.
9. Lipa, B. J, P. A. Sturrock, E. Rogot, (1976), Nature 259, 302.

10. Malin, S.R., C. Malin,, B. J. Srivastava, (1979), Nature, 277, 646
11. Simeonova, S. G., R. C. Georgieva,., B. H. Dimitrova, et al., (2010),
        Bulg. Astron. J. 14, 144–161
12. Dimitrova, S., I. Stoilova, K. Georgieva., et al., (2009),
        Fund. Sp. Res., HELIOBIOLOGY 161–165
13. Podolská, K., (2018), Journal of Stroke and Cerebrovascular
        Diseases} 27/2, 404–427.
14. Podolská, K., (2022), Atmosphere 13(1), 13.
15. Alabdulgader, A., R. McCraty, M. Atkinson, et al., (2018),
        Scientific Reports} 8, 1–14
16. Palmer, S. J., M. J. Rycroft, M. Cermack, (2006),
        Surv. Geophys. 27, 557–593.
17. Georgiev, Ts. B., (2014), Bulg. Astron. J., 20, 14–25.
18. https://www.nsi.bg/bg/content/2920/
19. https://omniweb.gsfc.nasa.gov/form/dx1.html.
20. https://www.ngdc.noaa.gov>stp>geomag>kp_ap

ВЪРХУ ЗАКЪСНЕНИЕТО НА ДОПЪЛНИТЕЛНАТА СМЪРТНОСТ, СВЪРЗВАНА С ГЕОМАГНИТНИТЕ СМУЩЕНИЯ

Ц. Георгиев, С. Симеонова, Л. Данкова

Резюме

Геомагнитните смущения, главно геомагнитни бури (ГМБи), но и нискочестотни резонанси, засягат част от хората, предразположени към мозъчни и съдови болести (МСБи). Понякога геомагнитният ефект се надценява спекулативно. Срещу това ние сравняваме изменения на геомагнитни индекси (ГМИи) и изменения на добавъчна смъртност (ДС). Ние правим това чрез кроскорелационни функции (ККФи), използвайки числото на Волф (ЧВ) за референтна времева скала. Ние подозираме, че силни ГМБи, като тези през 2003 г., увеличават 3–4 г. по-късно относителната обща смъртност с до 4×10^{-5}. Иначе, типичната ДС, свързана с ГМБи, е под $\sim 10^{-5}$ годишно. Даже ако тези наши величини са преувеличени, общо взето те са малки. Анализирайки данни за България и 5 нейни региони за последните 3 слънчеви цикли, ние потвърждаваме, че максимумът на ДС, свързвана с ГМБ, закъснява след максимума на ЧВ с ~5 г. Ние потвърждаваме и че максимумът на ГМБ закъснява след максимума на ЧВ с 1–2 г. Ние намираме специално, че максимумът на ДС, свързвана с ГМБи, закъснява след максимума на ГМИи с 3–4 г. По принцип, типичните продължителности на МСБи могат да бъдат определени, ако началата им са известни. В медицината началата са обикновено неизвестни. Обаче, подозирайки, че ГМБ са тригери на част от МСБи, ние следва да предположим, че тези МСБи завършват с летален изход след 3–4 г.

OBSERVATION OF SPECTRAL INDICES PERFORMANCE FOR POST-FIRE FOREST MONITORING

Daniela Avetisyan, Nataliya Stankova

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

Abstract

Monitoring post-fire forest disturbances and subsequent recovery is vital for the management and preservation purposes of the forest ecosystems. This study aimed to assess forests' damages and regrowth dynamics after fire using remotely sensed data and to compare its reliability for post-fire monitoring in different forest environments. This paper compared forest regrowth dynamics applying selected spectral indices – Differenced Normalized Difference Vegetation Index (dNDVI), Differenced Normalized Burn Ratio (dNBR), and Differenced Disturbance Index (dDI). The post-fire environmental impact and recovery processes were performed on the territory of the three fires in Bulgaria – Ardino, Bistrishko branishte, and Perperek.

Introduction

Due to global climate change the number of wildfires is increasing resulting in disturbances to forest ecosystems. Monitoring post-fire forest disturbances and subsequent regrowth processes is of great importance for arranging activities for forest ecosystem preservation. Aerospace remote sensing methods are a high-tech tool for reliable and large-scale monitoring of recovery processes occurring in forest ecosystems after a fire [1, 2]. Many researchers apply spectral vegetation indices (VIs) to monitor forest regrowth dynamics [3, 4]. Normalized Difference Vegetation Index (NDVI) [5] uses spectral reflectance characteristics (SRC) of vegetation in Red and Near-infrared (NIR) bands and Normalized Burn Ratio (NBR) [6] – NIR and Short-wave infrared (SWIR) bands. Disturbance Index (DI) [7] uses the linear orthogonal transformation of multispectral satellite images – Tasseled cap transformation (TCT) [8, 9], that increases the degree of identification of the main landscape components changing during a fire – soil, vegetation, and moisture/water.

The purpose of the present study was to observe forest disturbances and regrowth dynamics after a fire in three different forest environments, using dNDVI, dNBR, and dDI.

Study area

Post-fire forest disturbances and regrowth monitoring were performed on the territory of three study fires in Bulgaria: Ardino, Bistrishko branishte, and Perperek (Fig. 1). The test sites were initially described in a previous study assessing the performance of selected spectral VIs for post-fire monitoring [10].
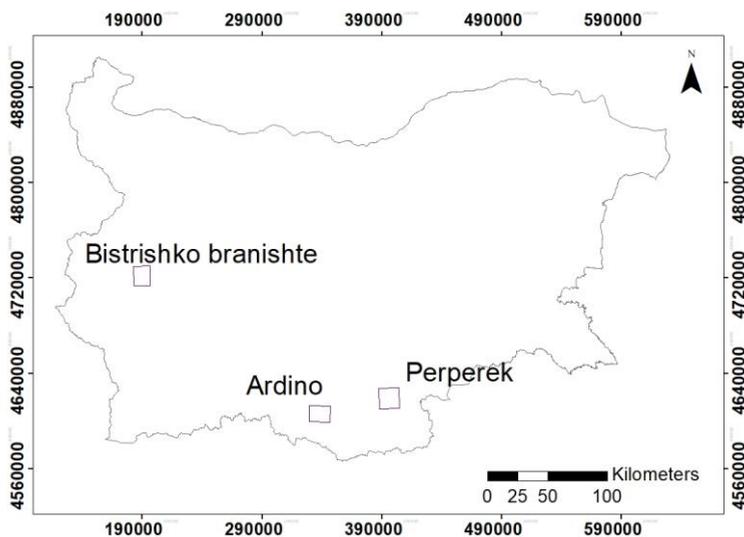
Fig. 1. Location of the study areas on the territory of Bulgaria

A fire broke out on 29 July, 2016 near Ardino town (Fig. 1), located in the southeastern part of Rhodope Mountains, Bulgaria. The fire affected 100 ha of coniferous forests. The climate in the area is Continental-Mediterranean, with mountainous elements. The slopes are mainly with east, southeast, and south exposures, which determine warm and dry conditions for vegetation regeneration. After the fire, the damaged forest stands were removed by sanitary logging in 2018.

The second fire occurred on July 1, 2012, in the nature reserve Bistrishko branishte, situated on the northeastern slope of Vitosha Mountain, next to Sofia, Bulgaria (Fig. 1). The fire affected 70 ha mostly dry and dead spruce forests because of the tornado that occurred in 2001 and bark beetle spots in 2004. Due to its preservation status as a nature reserve, no sanitary loggings were conducted in

the area. The climate is mountainous, and the slopes are mostly with northwest and north exposures, determining cold and wet conditions for vegetation regrowth.

A fire occurred on November 21, 2015, in the northeastern part of the Rhodope Mountains, near Perperek village, Bulgaria (Fig. 1). The fire burned 30 ha of coniferous forests. The damaged forest stands were removed by sanitary loggings in 2017. The climate in this area is Continental-Mediterranean and the slopes have northeastern exposure. The character of the relief and gentler slopes provide more favorable conditions for vegetation development.

Data and methods

Forest disturbances were assessed, and regrowth monitoring was conducted for the study period – 2012-2021, using Landsat (ETM+ and OLI) and Sentinel 2 (A and B) satellite imageries (Table 1). Sentinel 2 images were downloaded through Copernicus Open Access Hub [11], and Landsat images – from the US Geological Survey – Earth Explorer [12].

Table 1. Satellite imageries used for the calculation of dNDVI, dNBR, and dDI

Bistrishko	branishte	Ardino		Perperek	
Date	*Sensor*	*Date*	*Sensor*	*Date*	*Sensor*
29/06/2012	*Landsat ETM+*	11/07/2016	*Sentinel 2A*	07/11/2015	*Landsat OLI*
15/07/2012	*Landsat ETM+*	05/08/2016	*Sentinel 2A*	25/12/2015	*Landsat OLI*
19/08/2013	*Landsat ETM+*	15/07/2017	*Sentinel 2A*	21/08/2016	*Sentinel 2A*
05/07/2014	*Landsat ETM+*	24/08/2018	*Sentinel 2A*	15/07/2017	*Sentinel 2A*
08/07/2015	*Landsat ETM+*	29/08/2019	*Sentinel 2A*	29/08/2018	*Sentinel 2A*
13/07/2016	*Sentinel 2A*	28/08/2020	*Sentinel 2A*	24/08/2019	*Sentinel 2A*
27/08/2017	*Sentinel 2A*	23/08/2021	*Sentinel 2A*	23/08/2020	*Sentinel 2A*
01/09/2018	*Sentinel 2A*			18/08/2021	*Sentinel 2A*
12/08/2019	*Sentinel 2A*				
05/09/2020	*Sentinel 2A*				
01/08/2021	*Sentinel 2A*				

The observation of forest disturbances and post-fire regrowth monitoring were performed using selected spectral indices – dNDVI, dNBR, and dDI. Table 2 presents the indices formulas.

Table 2. Spectral indices used for the regrowth monitoring

Index	Abbreviation	Formula
Normalized Difference Vegetation Index	NDVI	$\dfrac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$
Differenced Normalized Difference Vegetation Index	dNDVI	$\text{NDVI}_{\text{post-fire}} - \text{NDVI}_{\text{pre-fire}}$
Normalized Burn Ratio	NBR	$\dfrac{\rho_{NIR} - \rho_{SWIR}}{\rho_{NIR} + \rho_{SWIR}}$
Differenced Normalized Burn Ratio	dNBR	$\text{NBR}_{\text{pre-fire}} - \text{NBR}_{\text{post-fire}}$
Disturbance Index	DI	$nBR - (nGR + nW)$
Differenced Disturbance Index	dDI	$\text{DI}_{\text{post-fire}} - \text{DI}_{\text{pre-fire}}$

The proposed methodology using the selected spectral indices for post-fire regrowth monitoring was validated in a previous study with the help of a method involving the delineation of dynamic boundaries for spatial accuracy assessment [10]. That previous study used VHR satellite data, including World View (2/3) and GeoEye (1) sensors, for validation.

Post-fire forest monitoring

For the post-fire monitoring, dNDVI, dNBR, and dDI rasters were generated on a yearly basis and compared with the values from the previous year. Fig. 2 summarizes the mean values of differenced indices on the territory of Ardino, Bistrishko branishte, and Perperek test sites. The negative mean dNDVI and the high positive mean dDI and dNBR values for all test areas indicate high disturbances immediately after the fire, corresponding to the left-most points in each plot in Fig. 2. The high dNDVI and the low dDI and dNBR mean values for all test sites indicate high recovery rates at the beginning of the study period – one year after the fire. In the following years, the indices values show the dynamics in the forests' regrowth, indicating clearly its interruption due to sanitary logging in Ardino (2018) and Perperek (2017) (Fig. 2).

a)

b)

c)
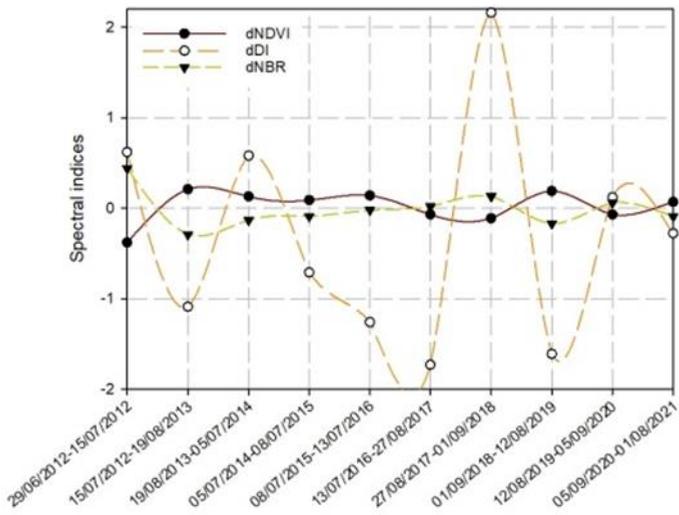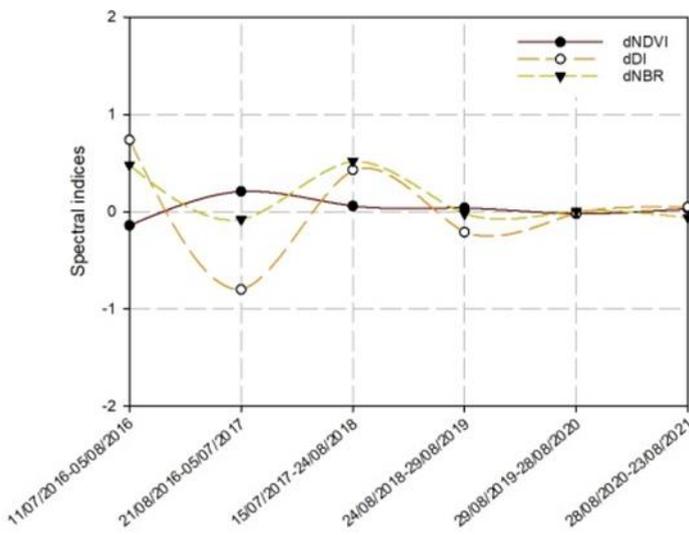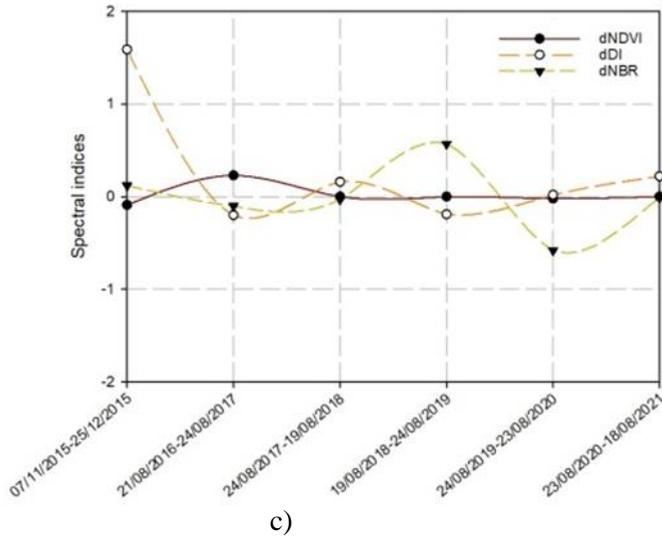
*Fig. 2. dNDVI, dNBR, and dDI mean values for Bistritsa (a), Ardino (b),
and Perperek (c) test sites*

Fig. 3 plots the standard deviations (SD) of each of the indices, calculated based on minimum, maximum, and mean values for each test site for five years, starting from the fire event. The SD values exhibit the dynamics of the natural environment. The highest SD values are observed during the fire event years and the years of the sanitary logging carried out afterward. This dependence is particularly noticeable in the SD values of dDI and dNBR, indicating increasing spectral reflection of bare soils. Amongst the test sites, Ardino seems to have the highest SD values of dNBR in the fire and logging years. Ardino has a larger share of the cut forest in the second year after the fire (compared to Perperek). With the progress of vegetation regrowth processes, the SD values decreased. That is noticeable in the SD values of dDI and dNBR for Ardino and Perperek. In Bistritsa, this performance was pronounced less. The SD values for this test site remain high for dDI. Bistritsa is the test site with the lowest post-fire vegetation regrowth (Fig. 3).
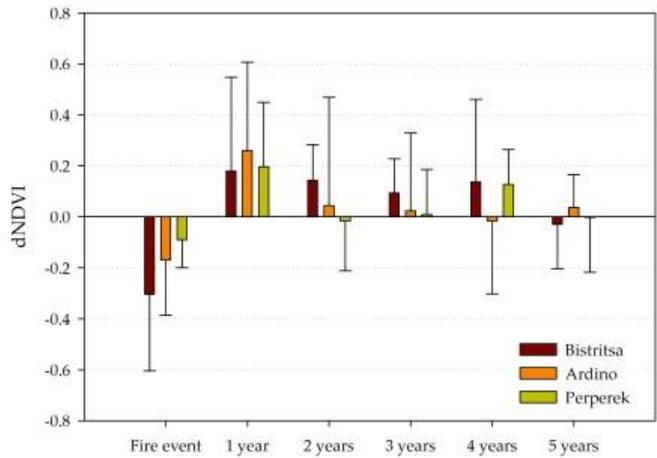
a)
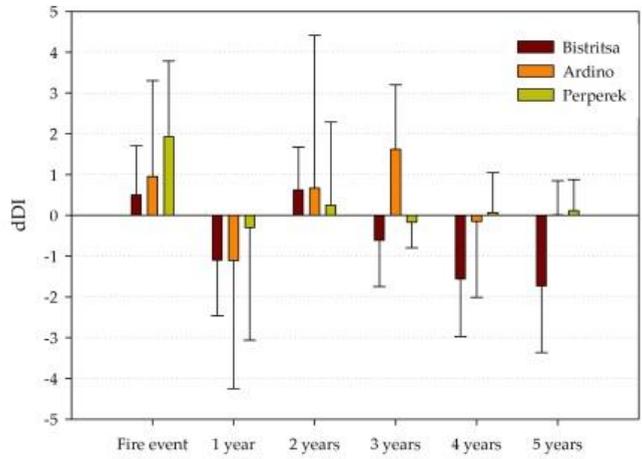
b)
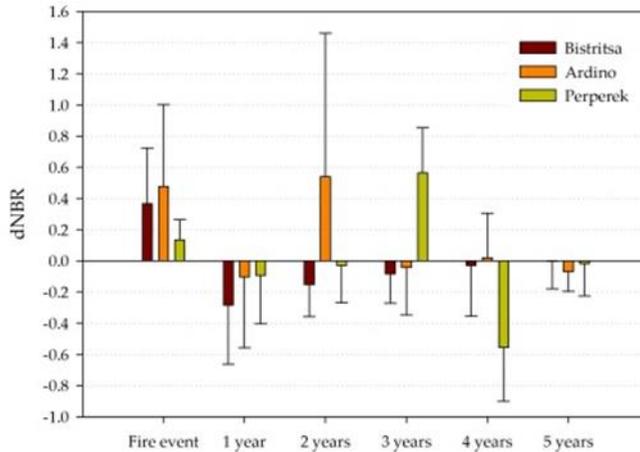
b)

Fig. 3. SD error bars, using minimum, maximum, and mean values of dNDVI (a), dDI (b), and dNBR (c) for each of the test sites for a five-year period, starting from the fire event

Discussion

Post-fire actual state of the test sites

Due to different forest fire intensities and the influence of topographical and climatic factors on forest ecological recovery, the process of post-fire vegetation regrowth is complicated and needs accurate and in-depth studying [13]. The post-fire environment impact on forest regrowth in the three test sites was demonstrated by the performance of dNDVI, dNBR, and dDI.

The three test sites are distinguished by both environmental conditions and management practices. Bistritrishko branishte had the slowest post-fire vegetation regrowth among the three test sites. That was induced by the landscape characteristics of the area: the wetter and colder habitat, steep slopes, and shallow soils, where erosion processes are more pronounced. Even years after the fire, vegetation regrowth is slow in this test site. The vegetation is mainly represented by annual herbaceous species, whose phenological development strongly influences the indices values. This is a reason for the highly pronounced dynamics in the indices mean values in Bistritrishko branishte, associated with the seasonality of the vegetation. dNDVI had the highest mean values (0.19) in mid-summers (Fig. 2a). dNDVI mean values below zero were recorded late in the summer and early in September when herbaceous vegetation senescence and lose large part of its moisture content (Fig. 2a). This vegetation type is also strongly influenced by environmental conditions during vegetation seasons. It is less resistant to drought and other anomalies related to temperature and humidity

74

[14,15]. Water-limited ecosystems with low gross primary productivity, such as grassland ecosystems, show higher dependency on hydro-climatic variations, influencing vegetation greenness. They are characterized by substantial productivity decreases under drought stress, which influence their SRC [15]. This was the reason for the higher dynamics of the three studied indices in Bistrishko branishte and especially of dNDVI. Studying different grassland types, Chen et. al., 2022 [13] also confirmed a greater dependency of NDVI values on the moisture content in ecosystems and precipitation fluctuations.

Perperek test site was distinguished with the optimal condition after the fire and with the lowest damage. The sanitary logging in this test site was not significant. Burnt trees in a small territory have been removed. This fact, as well as the characteristics of the terrain, does not favor the development of intensive erosion processes, and the vegetation has better conditions to recover. Amongst the three test sites, Perperek had the optimal environmental condition for vegetation regrowth.

The Ardino test site was significantly affected by the fire, which was also the reason for the sanitary logging of a large part of the forest vegetation. This fact, as well as the landscape characteristics (steep terrain and soil erosion, slopes exposure, heat-moisture ratio), determine less favorable conditions for vegetation regrowth than those in Perperek, yet better than in Bistrishko branishte.

The results demonstrating the influence of the post-fire environment on the actual state of the forest's regrowth confirm the results of Chen et. al., 2022 [13]. Using DI as a factor removing phenological interference within the area of observation, they assessed the impact of local forest ecology on the post-fire vegetation regrowth and found a clear correlation between the index and various topographic and climatic factors. Amongst the studied factors, elevation, and slope exposure, through their influence on the heat-moisture ratio, stand out as factors with the highest impact on the forest vegetation regrowth. In mountainous areas, colder habitats are distinguished with lower recovery rates than warmer ones [13].

Performance of dNDVI, dNBR, and dDI for post-fire monitoring

Various post-fire monitoring studies have reported differences in spectral vegetation indices performance that depend on the vegetation state in the observed ecosystems. The differences in post-fire vegetation state are determined primarily by the temporal pattern of the observation (immediate post-fire observation, one and/or several years of post-fire observations), vegetation type pattern, seasonal differences, and environmental conditions. These results confirm the conclusions obtained in the present study.

Amongst the studied indices, dNDVI and dDI clearly showed disturbances in all three test sites soon after the fires (Fig. 2). The mean values of dNDVI were lowest (-0.38 in Bistrishko branishte, -0.14 in Ardino, and -0.09 in Perperek)

(Fig. 2) in these observation periods, and the values of dDI were amongst the highest (0.62 in Bistrishko branishte, 0.74 in Ardino, and 1.59 in Perperek) (Fig. 2). The only exception was dDI in Bistritsa, which recorded the highest mean values (2.16) years after the fire (01/09/2018 – 27/08/2017) (Fig. 2a). dNBR recorded its highest mean values immediately after the sanitary logging in Ardino and Perperek (0.52 in Ardino and 0.57 in Perperek) (Fig. 2b and 2c). The indices involving the SWIR band in their calculation are superior to using only red and NIR bands for monitoring forest disturbances. The logging activities and the increasing bare soils areas significantly impacted the SRC in both test sites, increasing the spectral reflectance in SWIR. The dry, bare soils and burned territories have similar signature profiles. Hence, the dNBR performance was expected. In Bistrishko branishte, such logging was not carried out, due to the protected area status.

The mean dNBR value after logging in Ardino was 0.52, and after the fire, it was slightly lower – 0.48, whereas, in Perperek, the difference between the dNBR after these two events was significant. After the logging, the mean dNBR was 0.57, and after the fire, it was barely 0.12 (Fig. 2c). The low difference between the two values for Ardino, and the large one for Perperek, is determined by the characteristics of the two fires, which differ in their intensity and damage caused. In the Ardino test site, the fire occurred in summer, and the characteristics of the terrain and weather conditions caused more significant forest damage. That led to clear-cutting in a large part of the area. In Perperek, on the other hand, the fire occurred in winter, and the terrain and wet and cool conditions supported its rapid suppression and less forest damage. The fire has affected less the forest vegetation than in the other test sites. As a result, selective logging was performed in Perperek. That influenced greater dNDVI and dDI mean values than dNBR (Fig. 2).

As expected, the highest values of dNDVI were recorded one year after the fire. At that moment, dNBR values were also the lowest (Fig. 2). That is associated with the initial rapid growth of grasses, covering the burned areas with vegetation, and the maximum change in vegetation's chlorophyll content. These results confirm the results obtained in other post-fire monitoring studies. In the following years, the indices dynamics weren't significant, and the differences between the individual test sites were determined by the differences in the ecological conditions [13] and the way the territory was regulated. dNDVI was distinguished with mean values close to zero in Perperek and Ardino test sites (Fig. 2b and 2c), which was determined firstly by the drastic reduction of vegetation after the sanitary logging, and secondly by the drier habitats. In Bistrishko branishte, where there was no sanitary logging, the values of the dNDVI showed higher dynamics (Fig. 2a), and dDI showed a greater dependence on the disturbances caused most probably by changes in environmental conditions. After the logging, the dDI values in Ardino slightly increased, while in Perperek, where the logging was significantly smaller, the index values were almost unaffected (Fig. 6b and 6c).

References

1. Мардиросян, Г., Природни бедствия и екологични катастрофи – изучаване, превенция, защита, София, Акад. Издателство „Проф. Марин Дринов", 2009, ISBN 978-954-322-350-3.

2. Stankova, N., Post-fire recovery monitoring using remote sensing: A review, Aerospace Research in Bulgaria, 2023, 35, 192-200. https://doi.org/10.3897/arb.v35.e19.

3. Stankova, N., R. Nedkov, Monitoring forest regrowth with different burn severity using aerial and Landsat data, Proceedings of Geoscience and Remote Sensing Symposium (IGARSS), Milano, Italy, 26-31 July 2015, 2166–2169. doi 10.1109/IGARSS.2015.7326233.

4. Kim, Y., M. H. Jeong, M. Youm, J. Kim, J. Kim, Recovery of Forest Vegetation in a Burnt Area in the Republic of Korea: A Perspective Based on Sentinel-2 Data, Appl. Sci. 2021, 11, 2570. https://doi.org/ 10.3390/app11062570.

5. Rouse, J., R. Haas, J. Schell, D. Deering, Monitoring Vegetation Systems in the Great Plains with ERTS, Proceedings of Third Earth Resources Technology Satellite-1 Symposium, Greenbelt, MD, USA, 1973, 309-317.

6. Key, C., N. Benson, The Normalized Burn Ratio (NBR): A Landsat TM Radiometric Index of Burn Severity, 1999. URL = http://www.nrmsc.usgs.gov/ research/ndbr.htm visited.

7. Healey, S., W. Cohen, Z. Yang, O. Krankina, Comparison of Tasseled Cap-based Landsat data structures for use in forest disturbance detection, Remote Sensing of Environment, 2005, 97, 301-310. doi: http://dx.doi.org/10.1016/ j.rse.2005.05.009.

8. Kauth, R., G. Thomas, The Tasseled Cap – a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat, Proceedings second ann. symp. machine processing of remotely sensed data, West Lafayette' Purdue University Lab. App. Remote Sensing, 1976.

9. Spasova T., A. Stoyanov, Open data and remote sensing in flood monitoring in the municipality of Karlovo, Bulgaria, CMDR COE Proceedings 2022, 2022, 12–21. https://cmdrcoe.org/fls/pubs/CMDR_COE_Proceedings8_-_2022.pdf.

10. Avetisyan, D., N. Stankova, Z. Dimitrov, Assessment of Spectral Vegetation Indices Performance for Post-Fire Monitoring of Different Forest Environments, Fire, 2023, 6, 290. https://doi.org/10.3390/fire6080290.

11. Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu/.

12. USGS. Earth Explorer. Available online: https://earthexplorer.usgs.gov/.

13. Chen, X., W. Chen, M. Xu, Remote-Sensing Monitoring of Postfire Vegetation Dynamics in the Greater Hinggan Mountain Range Based on Long Time-Series Data: Analysis of the Effects of Six Topographic and Climatic Factors, Remote Sens., 2022, 14, 2958. doi: https://doi.org/10.3390/rs14132958.

14. Cartwright, J. M., C. E. Littlefield, J. L. Michalak, J. J. Lawler, S. Z. Dobrowski, Topographic, soil, and climate drivers of drought sensitivity in forests and shrublands of the Pacific Northwest, USA. Sci. Rep., 2020, 10, 1–13.

15. Avetisyan, D., D. Borisova, E. Velizarova, Integrated Evaluation of Vegetation Drought Stress through Satellite Remote Sensing. Forests, 2021, 12, 974. doi: https://doi.org/10.3390/ f12080974.

НАБЛЮДЕНИЕ НА ПРИЛОЖЕНИЕТО НА СПЕКТРАЛНИ ИНДЕКСИ ЗА МОНИТОРИНГ НА ГОРИТЕ СЛЕД ПОЖАР

Д. Аветисян, Н. Станкова

Резюме

Мониторингът на нарушенията на горите след пожар и последващото възстановяване е от жизненоважно значение за целите на управлението и опазването на горските екосистеми. Това проучване има за цел да оцени щетите в горските екосистеми след пожар, както и динамиката на протичането на възстановителните процеси, използвайки данни от дистанционно наблюдение и да сравни неговата надеждност за мониторинг след пожар в различни горски екосистеми. Тази статия сравнява динамиката на възстановяване на горите, като прилага избрани спектрални индекси – Differenced Normalized Difference Vegetation Index (dNDVI), Differenced Normalized Burn Ratio (dNBR) и Differenced Disturbance Index (dDI). Направената оценка на въздействието на пожарите върху околната среда, както и процесите на възстановяване, са извършени на територията на три тестови пожара в България – Ардино, Бистришко бранище и Перперек.

TESTING OPTICAL SPECTRAL INDICES FOR ASSESSMENT OF SURFACE CHANGES DUE TO PERMAFROST MELTING ON LIVINGSTON ISLAND, ANTARCTICA

Nadya Yanakieva[1,2], Daniela Avetisyan[1]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences
[2]Bulgarian Antarctic Institute
e-mail: nadya.yanakieva@mail.space.bas.bg, davetisyan@mail.space.bas.bg

Abstract

The present study aims to examine the potential of optical satellite data and spectral indices to assess surface changes induced by permafrost melting. Surface changes related to permafrost melting on Livingston Island, Antarctica, were examined using optical satellite data from Sentinel-2 sensors of the European Space Agency (ESA). The study area coincides with previous field studies by electrical resistivity tomographic profiles made to establish and visualize the presence of permafrost. Utilizing the advantages of remote sensing methods and calculation of optical indices, it was tracked whether and to what extent there was a surface change and melting of the permafrost in the study area. The observation period encompasses the astral summer seasons from 2016 to 2023.

The results show that the combination of different optical indices gives a better understanding of changes in the terrain. The combined use of the Normalized Difference Glacier Index (NDGI), Normalized Difference Snow Index (NDSI), Normalized Difference Snow and Ice Index (NDSII), Normalized Difference Water Index (NDWI), Normalized Difference Vegetation Index (NDVI), and Moisture Stress Index (MSI) indicates for a pronounced trend of melting of the active layer of the permafrost periglacial area of research in March 2016 and 2017, and from January to mid of March 2023.

Introduction

Permafrost is spread among frozen soil, rocks, or underwater sediment, which continuously remains below 0°C for two years or more. It typically exists beneath the so-called active layer, which freezes and thaws annually, and so can support plant growth, as the roots can only take hold in the soil that is thawed [1].

Permafrost is widespread in Polar Regions or high mountains and can be found in the ice-free areas of rocks and soil. It is an object of research in the periglacial areas. In addition, it has a key role in ecosystems, hydrology, and geomorphological dynamics. For example, only in the last decade, it was possible to have a more accurate overview of the thermal state of permafrost and active layer

dynamics in Antarctica. This requires the installation of new GTN-P (Global Terrestrial Network for Permafrost) boreholes and Circumpolar Active Layer Monitoring sites (CALM) as part of the Scientific Committee on Antarctic Research (SCAR) expert groups' projects. However, there is still a lot to be understood about the Antarctic permafrost, the active layer, and mainly about their relationships to other environmental variables [2].

Polar Regions are among the most vulnerable territories on the planet and are very sensitive to global climate change. Remote sensing methods are very useful for studying the changes and processes that occur in these hardly reachable places. Satellites provide the ability to measure and monitor elements of the Cryosphere continuously and with better spatial coverage than field or in situ measurements. Copernicus Sentinel-2 carries an innovative wide-swath high-resolution multispectral sensor with 13 spectral bands. It is providing high-resolution optical imagery with global coverage of Earth's land surface every five days [3]. Remote sensing is used to monitor the glacier and periglacial territories. For example, permafrost cannot be directly observed from space, but different types of satellite data, along with ground measurements and modeling, allow scientists to picture permafrost ground conditions.

This research aims to test the abilities of remote sensing optical images to track the snow cover and thawing changes in the area of previously studied permafrost profiles close to the Bulgarian Antarctic Base on Livingstone Island, South Shetland Islands, Antarctica. The expectations were for lower snow coverage through the years and more wet ground because of the thawing active layer.

Study area, scientific background, and related research

The study area is situated on the South Shetland Islands, one of the Earth's regions where warming has been more significant in the last 50 years [4]. Field activities had been focused on Livingston Island, Antarctica (62º39'S, 60º21'W) (*Fig. 1*).

High mountain relief distinguishes Livingston Island. The terrain is 90% covered by glaciers. The island's geological setting provides an excellent site for studying and monitoring the relationships between permafrost, geomorphodynamics, and climate. Thermal anomalies occur at several localities and the island is a very good site for studying the interactions between volcanic eruptions near Deception Island, its ash transportation and deposition on Livingston Island, geomorphodynamics, and permafrost.

Fig. 1. Livingston Island, Sought Shetland Islands, Maritime Antarctica. ©ESA

The study area coincides with previous geophysical field studies performed by the deployment of electrical resistivity tomographic profiles to establish and visualize the changes in the active layer of the permafrost. Boreholes (deep 6 to 25 m) for permafrost temperature monitoring (GTN-P) and sites for active layer monitoring (CALM-S) were installed. The main aims were identifying permafrost characteristics and spatial distribution, identifying the climate controls on permafrost temperatures and their sensitivity to climate change, and modeling permafrost distribution and temperature in space and time to assess the potential effects of climate change [5–7].

Fig. 2. Electro-tomographic profiles A and B with their coordinates. Geographic map of the area close to Bulgarian Antarctic Base, Livingstone Island, Antarctica. M 1:2000, map sheet №4. ©Bulgarian Antarctic Institute, ©Agency for geodesy, cartography and cadaster, military geographic service. 2017.

Compared with the Arctic, very little is known about the distribution, thickness, and properties of the permafrost in Antarctica. The previous geophysical research was part of the so-called PERMANTAR project focusing on Permafrost and Climate Change in the Maritime Antarctica. The project contributes to the Global scientific effort to bridge the gap in the knowledge of Antarctic permafrost characteristics, sensitivity, and implications for climate change.

a)   b)
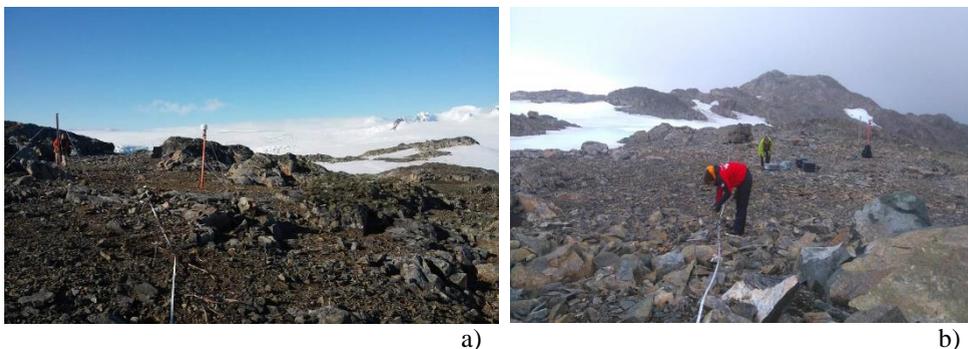
Fig. 3. Electro-tomography profiles A and B (see Fig. 2) of the periglacial terrain, close to the Bulgarian Antarctic Base, Livingstone Island, Antarctica. 2017.
*©Nadya Yanakieva*

PERMANTAR involves 3 Portuguese research centers, a Spanish and Argentinean research group, and the Bulgarian Antarctic Institute, which contributes with the logistic support of its Antarctic base and scientists, who were involved in the realization of the project. The project is interdisciplinary and with multinational collaboration [8].

The last electro-tomography, as part of the PERMANTAR project, was made during the Austral summer of 2017 to repeat for the fourth time some of the profiles of permafrost research. The terrain conditions during February 2017 were without snow cover, which serves as an isolator of the ground from the direct influence of meteorological factors. The results showed that in the active layer, there is no solid permafrost, but just separate patches of frozen ground, that depleted in time [5–7].

Methods

The present research is based on the application of remote sensing methods using optical multispectral images from Copernicus satellites Sentinel-2. The Sentinel-2 images encompass the period of eight years (from 2016 until 2023) during the Austral summer (January to April). The Polar Regions, and in particular Livingstone Island in Antarctica, have been covered by clouds very often during most of the year. This is one of the obstacles to studying that part of the world using optical satellite data. Therefore, for the present research, we used only nine cloud-free images over the study area for the chosen period.

Sentinel-2 provides high-resolution images in the visible and infrared wavelengths, to monitor vegetation, soil and water cover, inland waterways, and coastal areas. Data are available globally from June 2015 onwards. The spatial resolution is 10 m, 20 m, and 60 m, depending on the wavelength and the revisit time is a maximum of 5 days to revisit the same area [9].

a) 28.03.2016   b) 30.03.2017   c) 04.04.2019
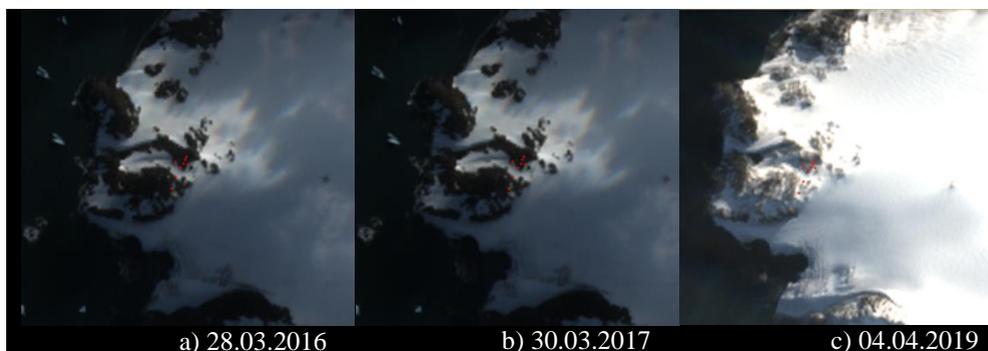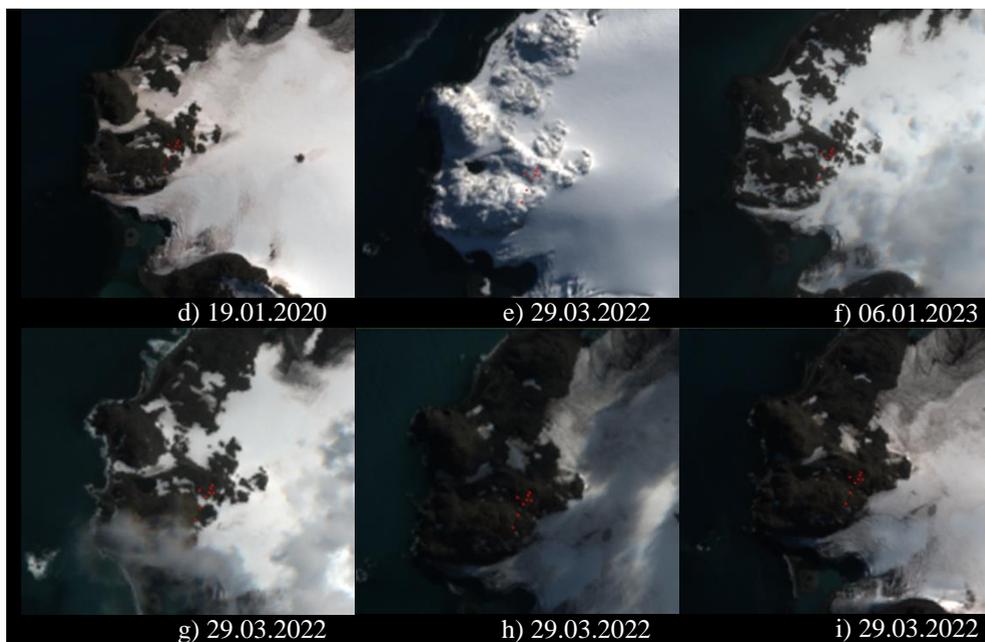
Fig. 4. Optical satellite images from Sentinel-2 of the permafrost study area, close to the Bulgarian Antarctic Base on Livingston Island, Antarctica, for six different years. ©ESA

The Sentinel-2 images (*Fig. 4 a)–i)*) were downloaded from the Copernicus data hub [10] and were transformed into composite images of 13 bands using ERDAS IMAGINE 2014. It is the software of Intergraph Corporation – an American software development and services company, which now forms part of Hexagon AB – the world's leading geospatial data authoring system [11].

The next step was to select and apply mathematical models of different indices. They provide information to define the specific characteristics of the terrain cover, such as water, snow, ice, rocks, or plants (*Table 1*) [12, 13].

The classifications of the individual indices enable comparative assessments between the individual dates in the period of observation. They are consistent with the ranges of the indices data and correspond to different land cover types – water, snow, ice, rocks, or plants (*Fig. 5*).

For the catalog of the visualization of the indices, a polygon of the study area close to the electro-tomographic profiles A and B with their GPS coordinates was drawn out. The size of the pixel is 10 m. The size of the area is around 2 square km. The different ranges of the indices' values are characterized by different colors of the visualization.

Table 1. Optical indices used in the study and description of their characteristics

Index and Range	Formula and Sentinel 2 Bands, used for its calculation	Application
MSI Moisture Stress Index 0 – 3	(1) $MSI = \dfrac{MidIR}{NIR}$ $MSI = \dfrac{B11}{B8}$	Moisture Stress Index (MSI) is used for canopy stress analysis, productivity prediction, and biophysical modeling. Interpretation of the MSI is inverted relative to other water vegetation indices; thus, higher values of the index indicate greater plant water stress and in inference, less soil moisture content. The values of this index range from 0 to more than 3 with the common range for green vegetation being 0.2 to 2 (Welikhe et al., 2017).
NDGI Normalized Difference Glacier Index -1 – 1	(2) $NDGI = \dfrac{Green - Red}{Green + Red}$ $NDGI = \dfrac{B3 - B4}{B3 + B4}$	Normalized Difference Glacier Index (NDGI) is used to detect and monitor glaciers by application of the Green and Red spectral bands. This equation is commonly used in glacier detection and glacier monitoring applications (Bluemarblegeo, 2019).
NDSI Normalized Difference Snow Index 0 – 1	(3) $NDSI = \dfrac{Green - SWIR}{Green + SWIR}$ $NDSI = \dfrac{B3 - B11}{B3 + B11}$	The Normalized Difference Snow Index (NDSI) is a numerical indicator that shows snow cover over land areas. The Green and short wave infrared (SWIR) spectral bands are used within this formula to map the snow cover. Since snow absorbs most of the incident, radiation in the SWIR while clouds do not, this enables NDSI to distinguish snow from clouds. This formula is commonly used in snow/ice cover mapping applications as well as glacier monitoring (Bluemarblegeo, 2019).
NDSII Normalized Difference Snow and Ice Index 0 – 1	(4) $NDSII = \dfrac{Green - SWIR}{Green + SWIR}$ $NDSII = \dfrac{B3 - B11}{B3 + B11}$	The Normalized Difference Snow and Ice Index (NDSII) is a numerical indicator that shows snow cover over land areas. The Green and Short Wave Infrared (SWIR) spectral bands are used within this formula to map the snow cover. Since snow absorbs most of the incident, radiation in the SWIR while clouds do not, this enables NDSI to distinguish snow

		from clouds. This formula is commonly used in snow/ice cover mapping applications as well as glacier monitoring (Bluemarblegeo, 2019).
NDWI Normalized Difference Water Index -1 – 1	(5) $$NDWI = \frac{Green - NIR}{Gren + NIR}$$ $$NDWI = \frac{B3 - B8}{B3 + B8}$$	Normalize Difference Water Index (NDWI) is used for water bodies analysis. The index uses Green and Near-Infrared (NIR) bands of remote sensing images. The NDWI can enhance water information efficiently in most cases. It is sensitive to build-up land and results in over-estimated water bodies. The NDWI products can be used in conjunction with NDVI change products to assess the context of apparent change areas (McFeeters, 1996).
NDVI Normalized Difference Vegetation Index -1 – 1	(6) $$NDVI = \frac{NIR - RED}{NIR + RED}$$ $$NDVI = \frac{B8 - B4}{B8 + B4}$$	The Normalized Difference Vegetation Index (NDVI) is a numerical indicator that uses the red and near-infrared spectral bands. NDVI is highly associated with vegetation content. High NDVI values correspond to areas that reflect more in the Near-Infrared spectrum. Higher reflectance in the Near-Infrared (NIR) corresponds to denser and healthier vegetation (GU, 2019).

Attributive tables were generated for each of the raster images and were used in diagrams to show the changes in the different points of the two electro-tomographic permafrost profiles A and B during the six years (*Fig. 6*).

Results

The optical images used for the study area are from the austral summer seasons. It is very difficult to select cloud-free optical satellite images appropriate for continuous monitoring of the same place because Livingston Island is covered mostly by clouds. For that reason, there are no images for the years 2018 and 2021.

The images used for the testing of different optical indices for assessment of surface changes due to permafrost melting on Livingston Island are from 28/03/2016, 30/03/2017, 04/04/2019, 19/01/2020, 29/03/2022, 16/01/2023, 14/03/2023, 17/03/2023 and 27/03/2023.

The values of two snow and ice indices (NDSI and NDSII), one glacier index (NDGI), one moisture stress index (MSI), one water index (NDWI), and one

vegetation index (NDVI) were compared to confirm the results of each of the indices, taking into consideration their typical range for a given object of observation.

The two snow-ice indices (NDSI and NDSII) show that if their values are above 0.4, usually the object of observation is snow. The range between 0.75 and 1 indicates thick snow (as it was on 04.04.2019), and thin snow is between 0.5 – 0.7 (as it was on 28.03.2016, 30.03.2017 and 29.03.2022) (*Fig. 5 a) and b)*).

When the values are between 0.2 – 0.5, it means that there is a melting snow area. That could be seen partially on 29.03.2022 in the west part of the study area, close to the shore; on 30.03.2017 in the middle of the study area, where the permafrost electro-tomographic profile A and partially profile B were situated; and on 04.04.2019, which is distinguishable only by NDSII calculation (*Fig. 5 b)*). These three days correspond mainly with thick or thin snow cover, but at the same time some parts probably are partially melted and this could be confirmed by using some of the other indices like the moisture (MSI) and water (NDWI) indices.

The other days of observation (19.01.2020, 06.01.2023, 16.01.2023, 14.03.2023, 17.03.2023, 27.03.2023) were with results below 0.2, indicating the lack of snow covering the terrain (*Fig. 5 a) and b)*). However, during these days, an increased amount of water on the terrain can be observed, which could be due to melting processes or another, meteorological factor (*Fig. 5 d) and f)*).

a)

b)

87
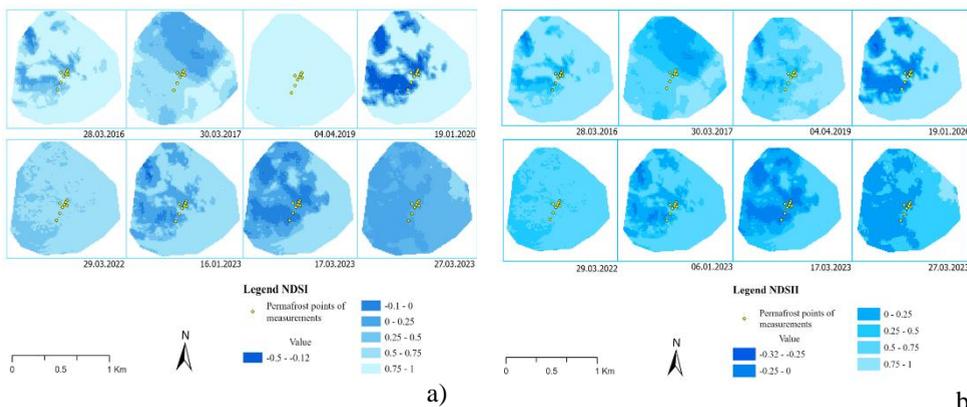
Fig. 5. Dynamics of the optical indices during the Austral summers of six different years (2016, 2017, 2019, 2020, 2022, 2023) in the permafrost study area: a) NDSI, b) NDSII, c) NDGI, d) NDWI, e) NDVI and f) MSI

The NDGI gives information on whether the terrain was covered by snow-ice when the results are above >0.45 [14]. NDGI values below <0.45 indicate ice-mixed debris where snow is more prevalent than ice. As a result of this experiment, the NDGI values could be divided into two main groups: above 0 to 0.1, corresponding to no snow cover, and below <0 – for most snowy terrain (as it was on 04.04.2019 and 29.03.2022) (*Fig. 5 c)*). In the days of observation, there are no indications for ice to cover the terrain.

The NDVI could be used not only for analyzing the state of vegetation but also for detecting snow, rocks, or water bodies. NDVI values below <-0.1 to -1 correspond to water bodies, values between -0.1 to 0 – to rocks, from 0 to 0.2 – to snow, NDVI between 0.2 to 0.5 corresponds with shrubs and grassland, and from 0.6 to 1 – with dense vegetation or rainforest [12]. The NDVI values at 04.04.2019, 29.03.2022, and 27.03.2023, for the permafrost study area on Livingston Island, are

in the range between 0 and 0.2 (*Fig. 5 e)*), indicating snow conditions, which corresponds to the values of the NDGI, NDSI, and NDSII indices. For the other days of observation, the NDVI values are below 0 indicating mainly rocky terrain.

The NDWI is used mostly for detecting water bodies, but also for humid or dry surfaces. If the results are between 0.2 and 1, that corresponds with the water surface, between 0 and 0.2 – with flooding or humidity, from -0.3 to 0 – it is about moderately dry, non-aqueous surfaces, and between -1 to -0.3 – for dry, non-aqueous areas [12]. In this research, the highest NDWI values were approximately 0.25, indicating temporary water bodies that could be formed due to the melting of permafrost or recent rainfalls. Such NDVI values are typical for 28.03.2016 and 30.03.2017. Most of the other images correspond with moderate humidity between 0 to 0.2 and with moderately dry surfaces. Non-aqueous surfaces are the three mostly snowy days – 04.04.2019, 29.03.2022, and 27.03.2023 (*Fig. 5 d)*). As a main conclusion, it could be seen that the less the snow, the higher the detected humidity.

The MSI is used for canopy stress analysis. The higher values indicate greater plant water stress and, in inference, less soil moisture content. MSI and Soil Moisture Classes below <0.2 correspond to very wet conditions, values between 0.2 and 0.7 - very moist, from 0.7 to 1.2 – moist, between 1.2 and 1.7 – slightly moist, from 1.7 to 2.2 – slightly dry, and above >2.2 – dry soil [13]. The MSI values between 0.2 and 0.7 in the previous research correspond to very moist conditions. Such conditions are typical at 28.03.2016, and partially observed at 30.03.2017, 04.04.2019, 29.03.2022, and 27.03.2023 (*Fig. 5 f)*). The MSI values for 06.01.2023 and partially for 30.03.2017 and 27.03.2023 indicate moist conditions, and for the other days – slightly moist to slightly dry conditions (*Fig. 5 f)*).

The raster images in Fig. 5 give information for the whole study area. The specific differences between the two electro-tomographic profiles (A and B) could be tracked through the diagrams in Fig. 6.

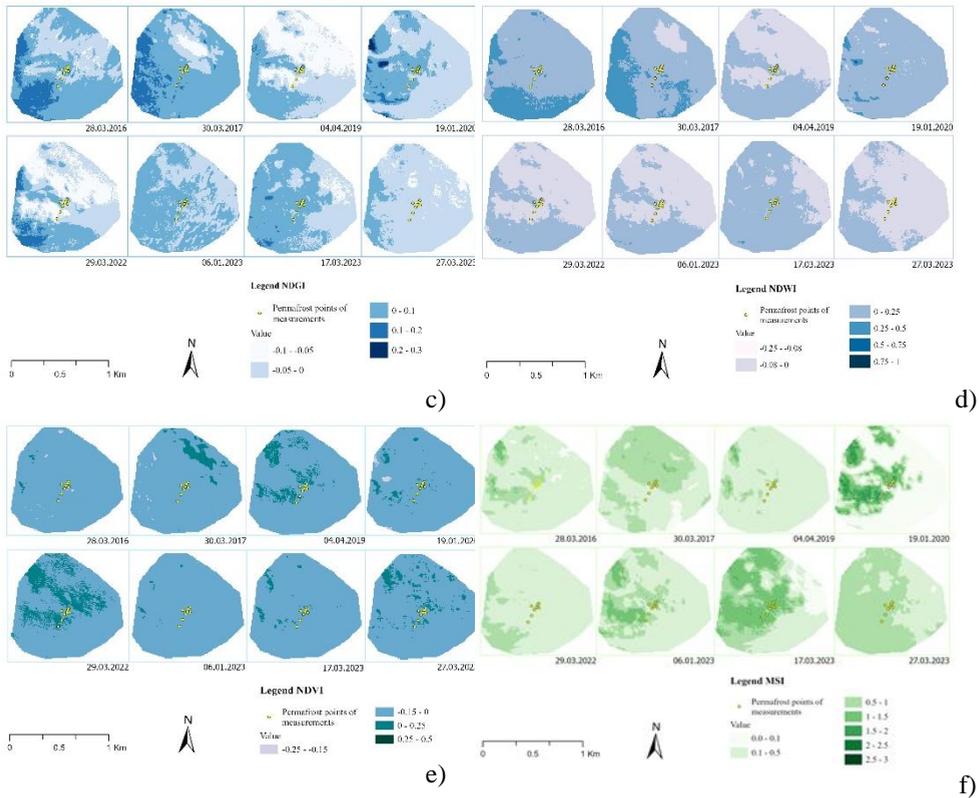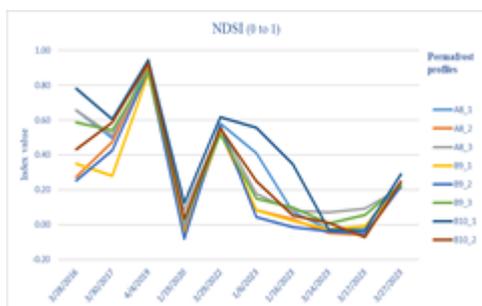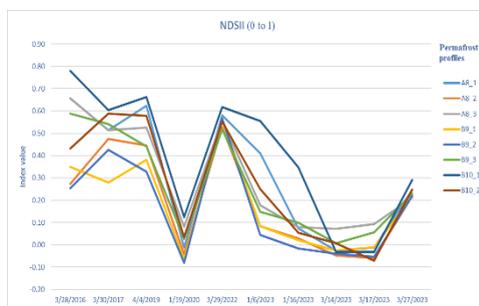a)   b)
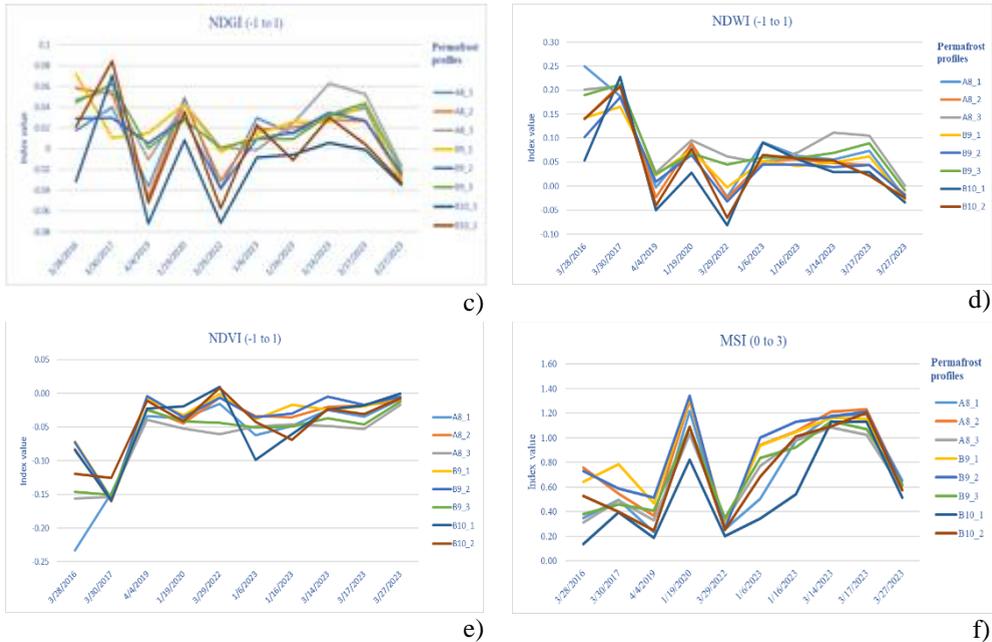
Fig. 6. Dynamics of the optical indices during the Austral summers of six different years (2016, 2017, 2019, 2020, 2022, 2023) in the two electro-tomographic profiles (A and B): a) NDSI, b) NDSII, c) NDGI, d) NDWI, e) NDVI and f) MSI

The diagrams confirm the main trend of the dynamics of the tested indices.

The two snow-ice indices NDSI and NDSII, and the moisture index MSI show almost the same trend of the two profiles (A and B) with deviation only in point 10_1 of profile B in January 2023, characterized by a higher presence of snow and moisture than the other parts of the profiles. The NDSI values indicate melting snow but, at the same time, there was no snow cover in that area during that period. So, that could be an indicator of the thawing process of the active layer of the permafrost area.

The water index NDWI shows the highest humidity in March 2016, 2017, and 2022 when the snow-ice indices values (NDSI and NDSII) indicate melting snow. The combined interpretation of those three indices confirms that the observed terrain was wet. However, these observations should be combined with meteo data, so it could be said for sure that the snow/permafrost in the observed area was melted during that period.

The moisture index MSI directly corresponds with the two snow-ice indices (NDSI and NDSII). The lower the MSI is, the lower the snow indices' values, but it does not correspond with NDWI, which gives more information about the humidity of the terrain, and also does not correspond with the NDSI data for melting snow.

The vegetation index NDVI could not detect any presence of vegetation across the study area, but showed the presence of some temporary water bodies with values between -0.1 and -0.23 in 2016 and 2017. The highest NDVI values were observed in the area of point 8_1 from profile A and also in point 10_3 from profile B in January 2023. That trend is another confirmation of the probable melting process on the surface of the study area.

Conclusions

The periglacial areas have higher dynamics in the conditions and changes of the ground than the glacial areas. Snow is the main isolator of the permafrost and its complete absence would be a reason for the thawing of the active layer of permafrost. In addition, the rocky terrain facilitates the snow and ice melting and, for that reason, they are very short-lived.

The results showed that the combination of different spectral indices gives a better understanding of the dynamics in the condition of the terrain. The combined use of the indices serves as some kind of verification of the results obtained. The values of NDSI, NDSII, NDWI, NDVI, and MSI indices show a trend of intense melting of the active layer of the permafrost periglacial area in March 2016 and 2017, and in the period between January and mid-March 2023. The observed trend for the latter period could be seen mostly in point 10_1 from profile B of the electro-tomographic profiles.

For a better understanding of the processes, it is necessary to conduct frequent observations in February, because this is the month with the highest temperatures during the austral summer in Antarctica. The main disadvantage of the study is the lack of cloud-free optical images from February during all of the six studied years, and most of all from February 2017 when the electro-tomographic profiles A and B were installed. This disadvantage could be overcome by adding Synthetic Aperture radar (SAR) data from the Sentinel-1 sensor in the research [15]. In addition, to better distinguish the possible factors for the observed high amount of water content on the terrain of the studied area we need daily meteorological data. One possible factor is the melting process of the snow cover or permafrost, but this water could be a result of recent rainfalls.

The integration of SAR and daily meteorological data is a required asset in our future research.

References

1. International Permafrost Association, (2023), What is Permafrost?
2. https://www.scar.org/science/antpas/about/#:~:text=Permafrost%20in%20the%20Antarcti c%20is,layer%20dynamics%20in%20the%20Antarctic
3. https://www.esa.int/Enabling_Support/Operations/Sentinel-2_operations

4. Bockheim, James, G. Vieira, M. Ramos, J. López-Martínez, (2013), Climate warming and permafrost dynamics in the Antarctic Peninsula region, p. 215–223, Global and Planetary Change 100

5. Correia, A., N. Yanakieva, G. Vieira, (2017), Three years of electrical resistivity tomography monitoring at the Papagal site near the Bulgarian Antarctic base St. Kliment Ohridski, Livingston Island, Maritime Antarctica, p. 92–93, Periglacial environments: advances in their study, patrimonial valuation and associated risks

6. Yanakieva, N., António Correia, Gonçalo Vieira, (2017), Remote Sensing of Surface Melt on Antarctica: Opportunities and Challenges, Visiones de Ciencia Antártica, Libro de Resúmenes, IX Congreso Latinoamericano de Ciencias Antártica, Punta Arenas, Chile, Publicación del Instituto Antártico Chileno, ISBN: 978-956-7046-11-9

7. Correia, A., N. Yanakieva, G. Vieira, (2017), Quatro anos de monitorização por tomografias de resistividade eléctrica no sítio calm da base antárctica búlgara St. Kliment Ohridski, Ilha Livingston, Antárctida Marítima, p. 91–92, Periglacial environments: advances in their study, patrimonial valuation and associated risks

8. http://www.bai-bg.net/projects.html

9. https://sentinel.esa.int/web/sentinel/missions/sentinel-2/overview

10. https://scihub.copernicus.eu/

11. https://hexagon.com/products/erdas-imagine

12. Bahadur Kshetri, T., (2018), NDVI, NDBI & NDWI Calculation Using Landsat 7, 8, Asian Institute of Technology

13. Welikhe, P., J. Quansah, S. Fall et al. (2017), Estimation of Soil Moisture Percentage Using LANDSAT-based Moisture Stress Index, Journal of Remote Sensing & GIS, DOI: 10.4172/2469-4134.1000200

14. Shukla, A., I. Ali, (2016), A hierarchical knowledge-based classification for glacier terrain mapping: a case study from Kolahoi Glacier, Kashmir Himalaya, Annals of Glaciology 57(71)

15. Spasova, T., (2018), Monitoring of short-lived snow coverage by SAR data around Livingston Island, South Shetland Islands in Antarctica, Adaptation Futures 2018 – 5th International Climate Change Adaptation Conference: Cape Town, South Africa, https://doi.org/10.15641/0-7992-2543-3

ИЗПИТВАНЕ НА ОПТИЧНИ СПЕКТРАЛНИ ИНДЕКСИ ЗА ОЦЕНКА НА ПОВЪРХНОСТНИ ПРОМЕНИ, ПОРОДЕНИ ОТ ТОПЕНЕТО НА ПЕРМАФРОСТА НА ОСТРОВ ЛИВИНГСТЪН, АНТАРКТИДА

Н. Янакиева, Д. Аветисян

Резюме

Целта на настоящото изследване е да проучи потенциала на оптичните сателитни данни и спектралните индекси за оценка на промените на земната повърхност, предизвикани от топенето на пермафроста. Повърхностните про-мени, свързани с топенето на пермафроста на остров Ливингстън, Антарктика,

бяха изследвани с помощта на оптични сателитни данни от сензори Sentinel–2 на Европейската космическа агенция (ESA). Районът на изследването съвпада с предишни теренни проучвания чрез томографски профили на електрическо съпротивление, направени за установяване и визуализиране на наличието на пермафрост. Използвайки предимствата на методите за дистанционно наблюдение и изчисляване на оптични индекси, беше проследено дали и в каква степен се наблюдават изменения на повърхността и има ли топене на пермафроста в района на изследване. Периодът на наблюдение обхваща астралния летен сезон от 2016 до 2023 г.

Резултатите показаха, че комбинацията от различни спектрални индекси дава по-добро разбиране на промените в терена и се валидират един друг като вид проверка на информацията. Според комбинацията от индекси – нормализиран разликов индекс за сняг (NDSI), нормализиран разликов индекс за сняг и лед (NDSII), нормализиран разликов воден индекс (NDWI), нормализиран разликов вегетационен индекс (NDVI) и индекс на стрес от влага (MSI) – се вижда, че има висока тенденция за вероятно топене на активния слой на перигляциалната зона на изследване на пермафроста през март 2016 и 2017 г. и от януари до средата на март 2023 г.

COMPARATIVE ANALYSES OF SECONDARY ECOLOGICAL SUCCESSION FOLLOWING WILDFIRES IN THREE DISTINCT FOREST TYPES. A STUDY CASE FROM MOGUER, SPAIN

Deyan Moni, Daniela Avetisyan

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: deyanmoni@gmail.com, davetisyan@space.bas.bg

Abstract

Wildfires have become increasingly prevalent and destructive in forest ecosystems worldwide, necessitating a comprehensive understanding of post-fire recovery dynamics for effective conservation and management. Remote sensing technology, coupled with vegetation indices such as Normalized Burn Ratio (NBR), Normalized Difference Vegetation Index (NDVI), Green Red Vegetation Index (GRVI), and Red Vegetation Index (RVI), offers a powerful means to investigate these processes. In this study, we utilize remote sensing techniques to conduct a comparative analysis of secondary ecological succession following wildfires in three distinct forest types (Coniferous, Sclerophyll, and Mixed) of a forest affected by fire near Moguer, Spain. Through the acquisition and analysis of multispectral satellite imagery, we monitored changes in vegetation health and recovery across the region of interest. The NBR index allowed us to assess the severity and extent of wildfire damage, while NDVI quantified vegetation greenness and regrowth. GRVI and RVI provided insights into subtle variations in vegetation composition and health. We identified distinct temporal and spatial patterns in post-fire recovery among the different forest types by applying these indices for the period between 2017 and 2021. Our findings underscore the significance of understanding the diverse responses of these ecosystems to wildfires. While common recovery patterns emerged, such as an initial decrease in NDVI followed by regeneration, variations were observed in the timing and magnitude of recovery. These distinctions are attributed to differences in species composition, fire adaptations, and ecological processes specific to each forest type. In conclusion, the utilization of NBR, NDVI, GRVI, and RVI indices allows for a more nuanced evaluation of post-fire recovery dynamics.

Introduction

Wildfires have become an increasingly prevalent and destructive force in forest ecosystems worldwide. These events have grown in both frequency and intensity [1]. To effectively address these challenges, it is essential to gain a comprehensive understanding of the dynamics of post-fire recovery. In this context, remote sensing technology has emerged as an invaluable tool, allowing us to observe and analyze these events from a broader perspective. Remote sensing

techniques enable us to acquire data through multispectral satellite imagery, facilitating a more detailed examination of post-fire recovery processes [2].

In this study, we aim to employ remote sensing technology and vegetation indices to conduct a comparative analysis of secondary ecological succession following wildfires in *coniferous, sclerophyllous,* and mixed forest types. By utilizing the Normalized Burn Ratio (NBR)[3], Normalized Difference Vegetation Index (NDVI)[4], Green Red Vegetation Index (GRVI)[5], and Red Vegetation Index (RVI)[6], we can assess the severity of wildfire damage and measure vegetation greenness and regrowth, providing critical insights into the impact and recovery of these ecosystems.

Normalized Burn Ratio (NBR) is commonly used to assess the severity of burn scars and monitor post-fire vegetation recovery [7].

Normalized Difference Vegetation Index (NDVI) is widely utilized to evaluate and monitor vegetation health and density, aiding in the assessment of ecosystem dynamics [8].

Green-Red Vegetation Index (GRVI) is specifically designed to emphasize the presence of green vegetation and is valuable in distinguishing plant vigor and stress levels [6].

Ratio Vegetation Index (RVI) is often employed to measure the density and vigor of vegetation cover, particularly in agricultural plots [6].

In this research, we are utilizing these indices to analyze the impact of a recent wildfire on a local forest ecosystem's vegetation cover and health. By integrating these indices, we can gain a comprehensive understanding of the post-fire recovery process, the changes in vegetation density, and the overall ecosystem resilience. The combination of these indices enables us to capture a holistic picture of the complex vegetation dynamics, ensuring a more accurate assessment of the recovery progress and facilitating informed management decisions for ecological restoration.

The relevance and importance of this study are underscored by the urgent need to comprehend how different forest types respond to contemporary wildfires. Each forest type may exhibit unique patterns of post-fire recovery, influenced by species composition, fire adaptations, and ecological processes specific to their environment. This knowledge is critical for developing tailored strategies to conserve and manage these ecosystems effectively.

Study area

The study area is situated close to Moguer town, in the province of Huelva, within the autonomous community of Andalucia (Fig. 1). The fire was declared on June 24, 2017, in Mazagón, within the area corresponding to the municipal district of Moguer, Huelva. It lasted a total of 10 days, and it was extinguished on July 4th. It covers an area of 203.5 km² (20,350 ha) and is positioned at an elevation of 49 meters above sea level. Due to its location along the Huelva coastal area, the

study area experiences a Mediterranean climate with influences from the Atlantic, resulting in a maritime climate pattern. The annual average temperature is 19.2°C (Moguer climate station). During summer, the average air temperature is 26°C and the maximum values reach up to 30°C.

Moguer has dry periods in June, July, and August. On average, December is the wettest month with 72 mm of precipitation. On average, July is the driest month with 1 mm of precipitation. The average amount of annual precipitation is 251 mm (https://www.eltiempo.es/moguer.html).
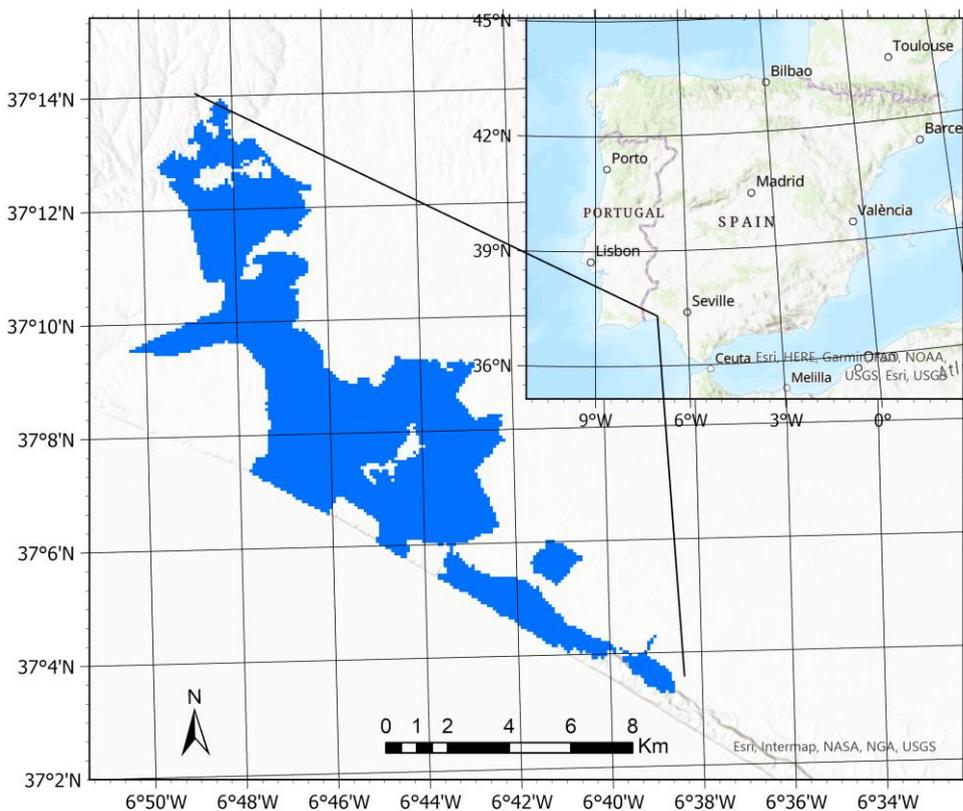

Fig. 1. Map depicting the study area affected by the fire declared on June 24, 2017, in Mazagón

Data and Methods

Data and data processing

A series of multispectral satellite images were acquired from the Landsat 8 satellite for post-wildfire dates, with a spatial resolution of 60 meters. The satellite image pre-processing steps include georeferencing, layer stacking, subseting, etc.

After the pre-processing, spectral vegetation indices were calculated (Table 1). They were used to assess the post-fire vegetation dynamics and to analyze the secondary ecological succession in individual forest types.

Table 1. Formulas for calculating the spectral vegetation indices, used in the present study

Spectral index	Abbreviation	Formula	References
Normalized Difference Vegetation Index	NDVI	$NDVI = \dfrac{NIR - RED}{NIR + RED}$	[4]
Differenced Normalized Difference Vegetation Index	dNDVI	$NDVI_{post-fire} - NDVI_{pre-fire}$	[9]
Normalized Burn Ratio	NBR	$NBR = \dfrac{NIR - SWIR}{NIR + SWIR}$	[3]
Differenced Normalized Burn Ratio	dNBR	$NBR_{pre-fire} - NBR_{post-fire}$	[9]
Green-Red Vegetation Index	GRVI	$GRVI = \dfrac{GREEN - RED}{GREEN + RED}$	[5]
Differenced Green-Red Vegetation Index	dGRVI	$GRVI_{post-fire} - GRVI_{pre-fire}$	-
Ratio Vegetation Index	RVI	$RVI = \dfrac{RED}{NIR}$	[6]
Differenced Ratio Vegetation Index	dRVI	$RVI_{post-fire} - RVI_{pre-fire}$	-

Where NIR represents near-infrared reflectance, SWIR is shortwave infrared reflectance, Red is red band reflectance, and Green denotes green band reflectance.

Analysis of Ecological Succession

Temporal trend analysis was conducted to examine the changes in vegetation over time using the computed dNBR, dNDVI, dGRVI, and dRVI indices. Spatial pattern recognition techniques, including GIS-based analysis, were employed to identify the spatial distribution of different successional stages in the post-fire forest ecosystem.

Software and Tools

Data processing and analysis were carried out using the ERDAS IMAGINE software 2014 (https://hexagon.com/products/erdas-imagine) for image preprocessing and the ArcGIS Pro platform [10] for spatial analysis and visualization.

This research adhered to the ethical guidelines outlined by the International Society for Photogrammetry and Remote Sensing (ISPRS). The limitations of this study include the reliance on satellite data with a 60-meter resolution, which may not capture fine-scale changes in vegetation.

Results

The results are structured to provide a comprehensive understanding of the recovery dynamics, utilizing data from a four-year period (2018-2022) and differential indices, including NBR, NDVI, GRVI, and RVI. The results are organized according to each forest type (*coniferous, sclerophyllous*, and mixed forests), highlighting the respective recovery rates, spatial and temporal variations on the regeneration process.

When examining the distribution of dNDVI, dNBR, dGRVI, and dRVI values across *coniferous, sclerophyll,* and mixed forests over multiple years (2018-19 to 2021-22), distinct patterns emerge, shedding light on the vegetation health and density within these ecosystems.

The *coniferous* forest consistently flaunted a dNDVI range of 0.1 to 0.2, signifying robust, healthy vegetation recovery. Higher dNDVI values were sporadic. Contrasting this, the mixed forest showed fluctuations, ranging from low to moderate dNDVI values. In recent years, a notable shift towards lower values has been seen, hinting at changing vegetation health. Meanwhile, the *sclerophyll* forest remained steady, maintaining moderate dNDVI values consistently. Its narrative whispered of unwavering stability in vegetation health over time.

The dNDVI for *coniferous* forests predominantly favored the 0.1 to 0.2 dNDVI range throughout, suggesting a consistent presence of healthy vegetation, while the higher dNDVI ranges remained marginal (Fig. 2).

2018 - 2019

2019 - 2020

2020 - 2021

2021 - 2022

Fig. 2. Classification Results of Four Differential Indexes Four Years Post-Wildfire. The figure illustrates the distinct categorization and evolution of four key differential indexes post-wildfire over a four-year period.

The dNDVI for mixed forest type varied between 0 to 0.2 dNDVI values, showcasing a dynamic response to environmental changes, especially with a significant shift towards lower dNDVI values in recent years (Fig. 3). The dNDVI for *sclerophyll* forest type maintained a consistent presence within moderate dNDVI values, hinting at stable v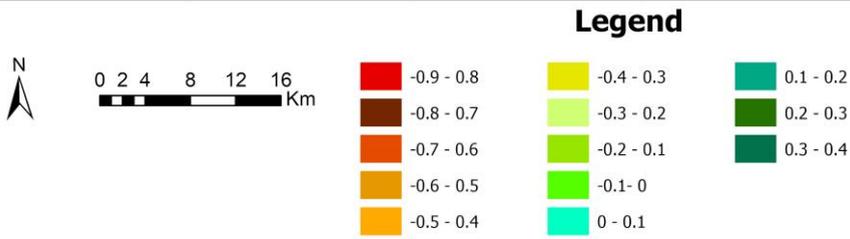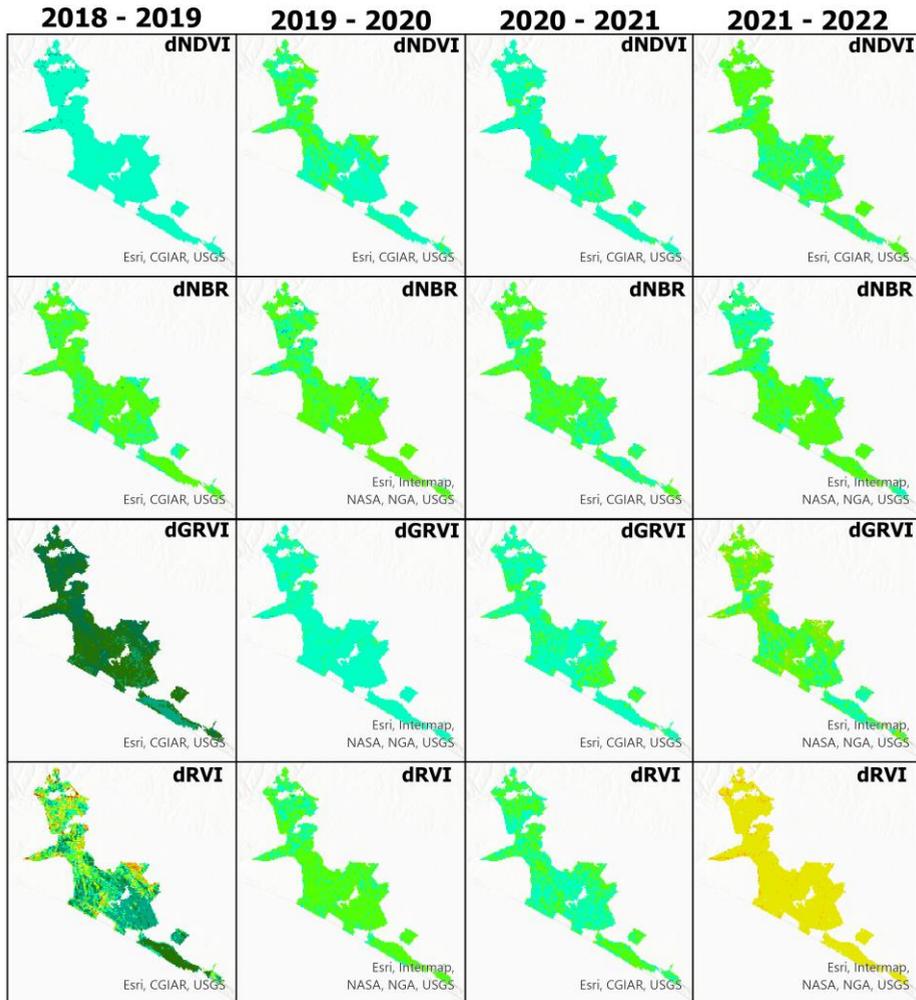egetation health trends over the observed years (Fig. 4). Figure 2. Classification Results of Four Differential Indexes Four Years Post-Wildfire. The figure illustrates the distinct categorization and evolution of four key differential indexes post-wildfire over a four-year period.
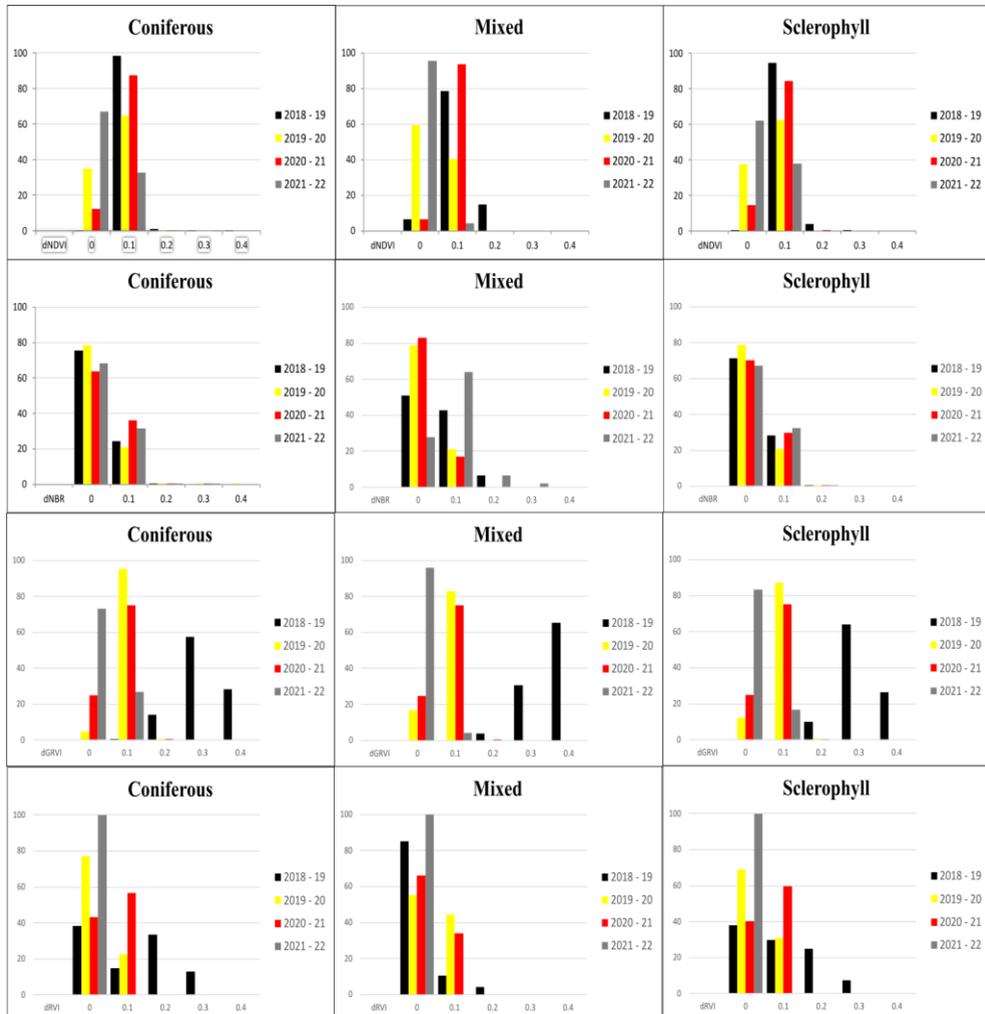
Fig. 3. Percentage covered area per value for every index used for the study

The dNBR for the *coniferous* forest showcased stability in burn severity, primarily residing within the 0 to 0.1 and 0.1 to 0.2 dNBR ranges, indicating minimal changes in post-fire recovery or severity (Fig. 3). The dNBR for mixed forest type notably fluctuated between 0 to 0.3, indicating varied recovery stages or potential environmental stressors influencing the forest (Fig. 2).

The dNBR for the *sclerophyll* forest type showed stability, predominantly residing within 0 to 0.2 ranges, indicating consistent post-fire recovery or limited variations (Fig. 3).
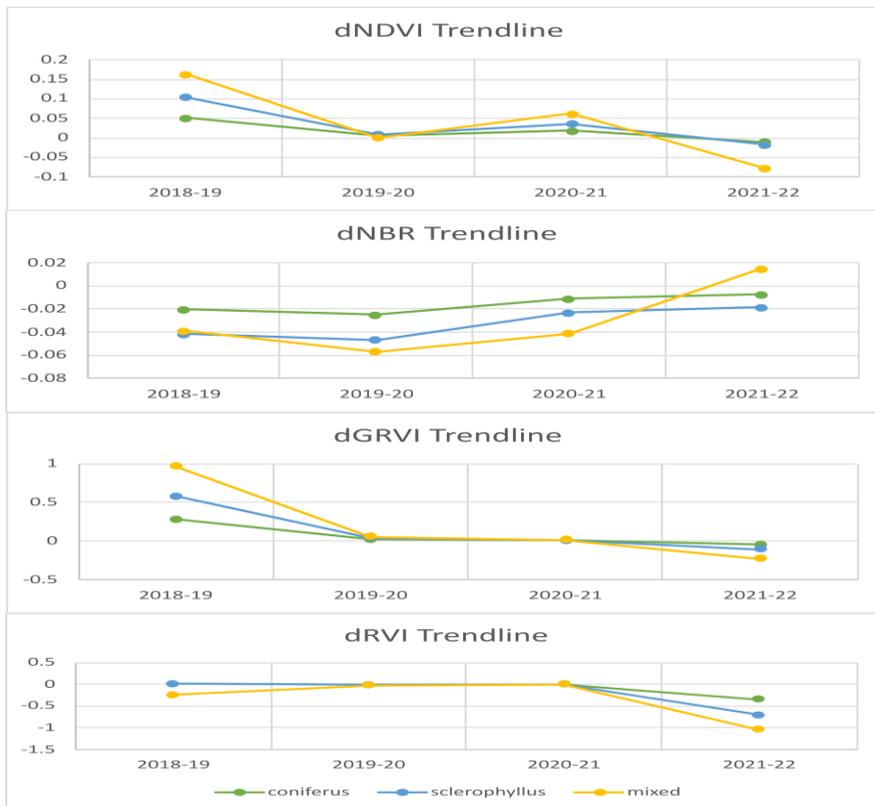
Fig. 4. Trendlines through the years from 2018 to 2022 of the differencial indices dNDVI, dNBR, dGRVI and dRVI

The dGRVI measurements for the three types of forests demonstrated fluctuating patterns, showcasing 0.2 to 0.4 values initially, indicative of robust vegetation recovery, particularly in grasses, within the first year. The subsequent years displayed consistent but comparatively lower recovery rates, characterized by

fluctuating dGRVI values, signifying a sustained yet moderated vegetation health and density within this specialized primary growth environment (Fig. 2).

The initial high values of dRVI in certain sites during the first year indicated robust recovery post-disturbance. However, subsequent years revealed declining trends, with dRVI values decreasing over time and eventually reaching negative values by the fourth year.
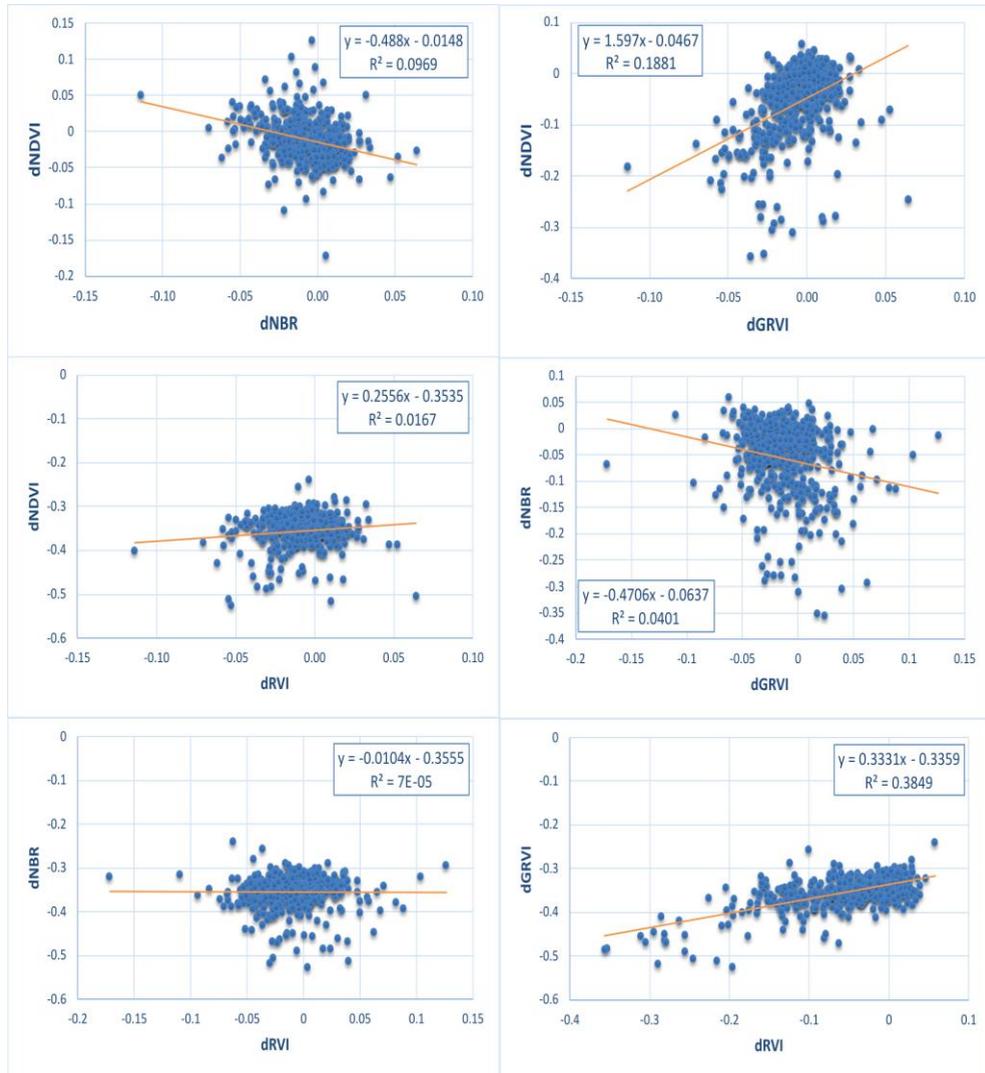
Fig. 5. Correlations between the indices used in this study using linear regression

This diminishing trajectory suggests an initial promising recovery that was not sustained over the long term (Fig. 2).

The integration of dNDVI, dNBR, and dGRVI across the *coniferous*, Mixed, and *sclerophyll* forests presents a multi-faceted understanding of these ecosystems. While the *coniferous* and *sclerophyll* forests displayed relatively consistent trends in vegetation health and recovery, the Mixed Forest stood out with its pronounced variability, indicating a more dynamic response to environmental factors or disturbances (Fig. 4).

The correlation analysis among the indices reveals distinct relationships: a weak negative correlation (r = -0.311) between the change in dNDVI and dNBR, hinting at a slight inverse trend between vegetation health and burn severity. Meanwhile, a moderately positive correlation (r = 0.434) emerges between dNDVI and dGRVI, suggesting parallel movements and potentially similar trends in vegetation health and density changes.

The moderately positive correlation between these indices implies a more aligned relationship between both, hinting that they may respond similarly to changes in vegetation health.

This suggests that dGRVI might serve as a complementary indicator to dNDVI in monitoring vegetation recovery, potentially capturing different types of vegetation response after disturbances. However, the correlation between dNDVI and dRVI is weak (r = 0.129), indicating a lesser association in measuring vegetation recovery. Additionally, a weak negative correlation (r = -0.200) between dNBR and dGRVI implies a subtle inverse relationship between burn severity and fluctuations in vegetation health, while the correlation between dNBR and dRVI is negligible (r = -0.008), suggesting a lack of meaningful relationship between burn severity and the RVI differential index in assessing recovery. This might suggest that the dRVI measures aspects of recovery unrelated to burn severity, emphasizing other factors influencing vegetation dynamics post-fire. Notably, a relatively strong positive correlation (r = 0.620) between dGRVI and dRVI signifies a notable tendency for these indices to move together, potentially reflecting similar patterns in assessing vegetation health and density changes within the study area (Fig. 5). This suggests that dGRVI and dRVI might capture similar aspects of vegetation recovery, potentially providing redundant information or reinforcing each other's assessments.

Conclusions

The comprehensive analysis spanning a four-year period from 2018 to 2022 has unearthed nuanced insights into the recovery dynamics of *coniferous*, *sclerophyllous*, and mixed forests. Through the meticulous examination of differential indices - NBR, NDVI, GRVI, and RVI - distinct patterns emerged,

painting a vivid picture of vegetation health, density, and recovery post-disturbance.

The *coniferous* forests boasted a consistent dNDVI range, reflecting robust and sustained vegetation recovery, while sporadic higher values hinted at localized vigor. Contrasting this, the mixed forests exhibited fluctuations, especially towards lower dNDVI values in recent years, indicating evolving vegetation health. Meanwhile, the *sclerophyllous* forests remained steadfast, maintaining moderate dNDVI values, suggesting unwavering stability in vegetation health over time.

Similarly, the analysis of dNBR, dGRVI, and dRVI offered intriguing insights. *Coniferous* forests showcased stability in burn severity, while the mixed forests displayed fluctuations, hinting at varied recovery stages or environmental stressors. *Sclerophyllous* forests demonstrated consistent post-fire recovery.

The correlations among these indices highlighted intriguing relationships. A weak negative correlation between dNDVI and dNBR hinted at an inverse trend between vegetation health and burn severity. Conversely, a moderately positive correlation emerged between dNDVI and dGRVI, suggesting parallel movements in vegetation health and density changes. Interestingly, the correlation between dGRVI and dRVI indicated a notable tendency for these indices to move together, potentially reflecting similar patterns in assessing vegetation recovery.

Looking ahead, future studies could benefit from delving into the influence of various environmental factors, especially the dynamics of climate elements during the research period. Accounting for these factors may provide a more holistic understanding of vegetation dynamics post-disturbance. Conducting such an analysis, while keeping other variables constant, could offer valuable insights into the intricate interplay between environmental factors and vegetation recovery.

References

1. Duane, M. Castellnou, and L. Brotons, "Towards a comprehensive look at global drivers of novel extreme wildfire events," Clim Change, vol. 165, no. 3–4, Apr. 2021, doi: 10.1007/s10584-021-03066-4.
2. E. Kurbanov et al., "Remote Sensing of Forest Burnt Area, Burn Severity, and Post-Fire Recovery: A Review," Remote Sensing, vol. 14, no. 19. MDPI, Oct. 01, 2022. doi: 10.3390/rs14194714.
3. McKenna, P., S. Phinn, and P. D. Erskine, "Fire severity and vegetation recovery on mine site rehabilitation using worldview-3 imagery," Fire, vol. 1, no. 2, pp. 1–28, Sep. 2018, doi: 10.3390/fire1020022.
4. J. W. Rouse, R. H. Haas, J. A. Schell, D. W. Deering, and others, "Monitoring vegetation systems in the Great Plains with ERTS," NASA Spec. Publ, vol. 351, no. 1, p. 309, 1974.
5. Motohka, T., K. N. Nasahara, H. Oguma, and S. Tsuchida, "Applicability of Green-Red Vegetation Index for remote sensing of vegetation phenology," Remote Sens (Basel), vol. 2, no. 10, pp. 2369–2387, Oct. 2010, doi: 10.3390/rs2102369.

6. A. Krtalic, A. Prodan, and I. Racetin, "Analysis of burned vegetation recovery by means of vegetation indices," in International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management, SGEM, International Multidisciplinary Scientific Geoconference, 2019, pp. 449–456. doi: 10.5593/sgem2019/2.2/S10.055.
7. N. Stankova, "Post-Fire Recovery Monitoring Using Remote Sensing: A Review," Aerospace Research in Bulgaria, vol. 35, pp. 192–200, 2023, doi: 10.3897/arb.v35.e19.
8. I. Ivanova, T. Spasova, and N. Stankova, "Using Sentinel-2 data for efficient monitoring and modeling of wetland protected areas," SPIE-Intl Soc Optical Eng, Sep. 2023, p. 42. doi: 10.1117/12.2681790.
9. C. H. Key and N. C. Benson, "LA-1 Landscape Assessment (LA) Sampling and Analysis Methods."
10. M. Law and A. Collins, "Getting to know ArcGIS PRO," (No Title), 2019.

СРАВНИТЕЛНИ АНАЛИЗИ НА ВТОРИЧНА ЕКОЛОГИЧНА СУКЦЕСИЯ СЛЕД ГОРСКИ ПОЖАРИ В ТРИ ОТДЕЛНИ ТИПА ГОРИ. ПРОУЧВАН СЛУЧАЙ ОТ МОГЕР, ИСПАНИЯ

Д. Мони, Д. Аветисян

Резюме

Горските пожари стават все по-разпространени и разрушителни в горските екосистеми по света, което налага цялостно разбиране на динамиката на възстановяване след пожар за ефективно опазване и управление. Технологията за дистанционно наблюдение, съчетана с индекси на растителността като нормализиран коефициент на изгаряне (NBR), нормализиран индекс на разликите в растителността (NDVI), зелен-червен индекс на растителност (GRVI) и червен индекс на растителност (RVI), предлага мощно средство за изследване на тези процеси. В това проучване ние използваме техники за дистанционно наблюдение, за да проведем сравнителен анализ на вторичната екологична сукцесия след горски пожари в три различни типа гори (иглолистни, широколистни и смесени) на гора, засегната от пожар близо до Могер, Испания. Чрез придобиването и анализа на мултиспектрални сателитни изображения ние наблюдавахме промените във възстановяването на растителността в интересуващия ни регион. Индексът NBR ни позволи да оценим тежестта и степента на щетите от горски пожари, докато NDVI количествено определи зелеността и повторния растеж на растителността. GRVI и RVI предоставиха представа за фините вариации в състава и здравето на растителността. Ние идентифицирахме различни времеви и пространствени модели при възстановяването след пожар сред различните типове

гори, като приложихме тези индекси за периода между 2017 г. и 2021 г. Нашите открития подчертават значението на разбирането на разнообразните реакции на тези екосистеми към горските пожари. Докато се появяват общи модели на възстановяване, като първоначално намаляване на NDVI, последвано от регенерация, се наблюдават вариации във времето и степента на възстановяване. Тези разлики се дължат на разликите във видовия състав, адаптациите към пожар и екологичните процеси, специфични за всеки тип гора. В заключение, използването на индекси NBR, NDVI, GRVI и RVI позволява по-нюансирана оценка на динамиката на възстановяване след пожар.

APPLICATION OF SATELLITE DATA FOR MONITORING THE RETREAT OF MOUNTAIN GLACIER *MORTERATSCH*, SWISS ALPS, OVER A PERIOD OF 51 YEARS

Andrey Stoyanov

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

Abstract
    The presented study aims to apply a method for monitoring the retreat of mountain glaciers in the Alpine region of the Swiss Alps, which consists of using optical satellite imagery and their spectral capabilities to observe snow and ice objects on the earth's surface. Satellite imagery with its big legacy archive dating back to 1972 for Landsat imagery, can be of big help to track and monitor the alpine glaciers retreat for long periods and serve as a database for modelling and predicting the glaciers retreat in the future. By combining different satellite data processing approaches, results have been obtained on the spatial distribution and dynamics of the Morteratsch Glacier over a period of 51 years. The focus of the study is to track the changes that have occurred along the positions of the ice front (terminus) in the glacier's ablation zone to obtain information about the glacier's dynamics during the study period.

Introduction

The negative effects of climate change and global warming are rapidly increasing every year worldwide, including natural phenomena such as heat waves, droughts, extreme rainfalls, floods, rockfalls, etc. These consequences of Climate Change are happening twice as fast, especially in the Alpine region [1], where events such as glacier recession, reduction of snow cover, and rockfalls (driven by the permafrost thaw) are observed.

The retreat of glaciers is well documented and is one of the major negative effects of climate change. It provides evidence for the rise in global temperatures since the late 19th century. Glacier retreat impacts the availability of freshwater supplies for irrigation and domestic use, river runoffs, mountain recreation, and animals and plants depending on glacier melt [2].

Because of the long history of optical satellite imagery of the Earth, satellite remote sensing offers a wide variety of opportunities for mapping glacier recession. With repeated images over time satellite remote sensing allows for the

regular monitoring of glacier surface elevation, velocity, area, length, equilibrium line altitude, terminus position, and more [3].

Area of interest

The Morteratsch Glacier [Fig. 1] is the largest glacier by area situated in the Bernina Range of the Bündner Alps, Switzerland, at an altitude starting from 2040 m and reaching up to 4048 m of Piz Bernina peak. It is located in the strongly glaciated Bernina Range on the border with Italy [4]. Morteratsch is a typical valley glacier with a distinct *terminus* and its *accumulation zone* lies between the peaks of Morteratsch, Bernina, Zupò, and Bellavista. Yearly length change measurements have been recorded since 1878, when its maximum length was 8,63 km, while in 2023, it was 5,59 km [5, 6].
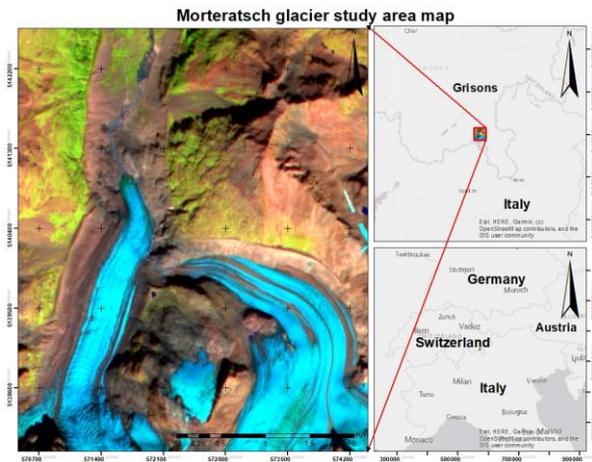
Fig. 1. Area of interest

Method and processing steps

The proposed approach in the present study involves processing of collected optical cloudless satellite datasets within the selected time frame of several missions: Landsat 1-2 MSS [8], Landsat 5 TM[8], Landsat 8 OLI [8] and Sentinel-2 MSI [9]; creating RGB pseudo-composites [10] by using the Green, NIR and MIR bands of each satellite image (except Landsat 1-2); applying of Normalized Differential Snow Index (NDSI) as reference and measurements, through GIS techniques, focused on the changes that occurred in the *terminus* position of Morteratsch glacier over the period of 51 years. Additionally, the changes that occurred in the size, length, and width of its *ablation zone* are calculated using GIS tools. Tracking the change in the location of a glacier *terminus* is a proof method of monitoring a glacier's movement based on the

differences in positions of the terminus measured from fixed positions at different time intervals [7].

Data use

All the temporal points of the chosen temporal period are collected during the late-summer time around the last decade of August and the first two decades of September, with the exception of the year 1972. The criteria for this selection of the year are due to the past maximum solar radiation and new and fresh snowfall typical of this altitude, which is yet to come. All the collected satellite images within the timeframe (Table 1) are processed in a GIS environment.

Table 1. Satellite datasets for the time frame of the study

Satellite, sensor	Acquisition date	Spectral bands	Resolution (m)
Landsat 1-2 MSS	07.10.1972 13.09.1975 04.09.1980	visible, NIR	60x60
Landsat 5 TM	20.08.1985 12.09.1990 16.08.1995 22.08.2000 05.09.2005 25.08.2010	NIR, SWIR, Green	30x30
Landsat 8 OLI	10.08.2013	NIR, SWIR, Green	30x30
Sentinel-2 MSI	26.08.2015 09.09.2016 30.08.2017 09.09.2018 04.09.2019 08.09.2020 13.09.2021 24.09.2022 08.09.2023	NIR, SWIR, Green	10x10 20x20

Results

Landsat 1-2 MSS RGB pseudo-composite 3D maps with a combination of NIR-NIR-Green bands delineating the glacier territory are shown in Fig. 2, where the temporal point of 1972 year serves as a starting point (and as fixed position of subsequent measurements) on the selected study period. The retreat of the glacier ice terminus for the 1972–1980 period is estimated at about 258 m.

Fig. 2. Landsat 1-2 MSS RGB composites from 1972, 1975 and 1980 year

Landsat 5 TM RGB pseudo-composites 3D maps with band combination SWIR-NIR-Green for the 1990–2000 period are shown in Fig.3. The change of the glacier terminus of the period 1980–1990 is estimated approximately at 128 m and for the 1990-2000 period at 146 m, respectively.
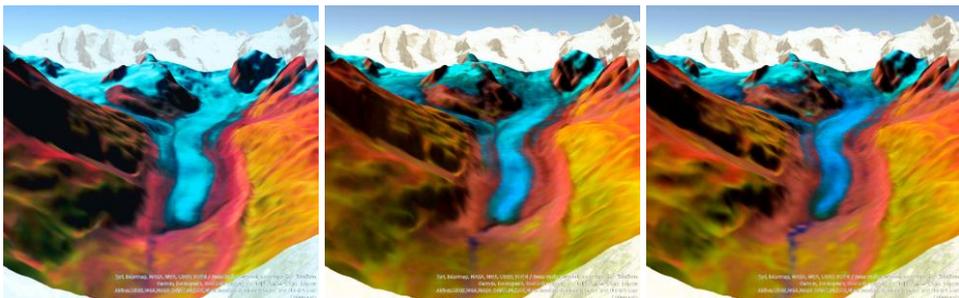
Fig. 3. Landsat 5 TM RGB composites from 1990, 1995 and 2000 year

Landsat 5 TM and Landsat 8 OLI RGB pseudo-composites 3D maps with band combination SWIR-NIR-Green (2005–2013) are shown in Fig. 4. The change of the glacier terminus for the decade 2000–2010 is estimated at 459 m, with a maximum loss of 260 m for the period 2005–2010.
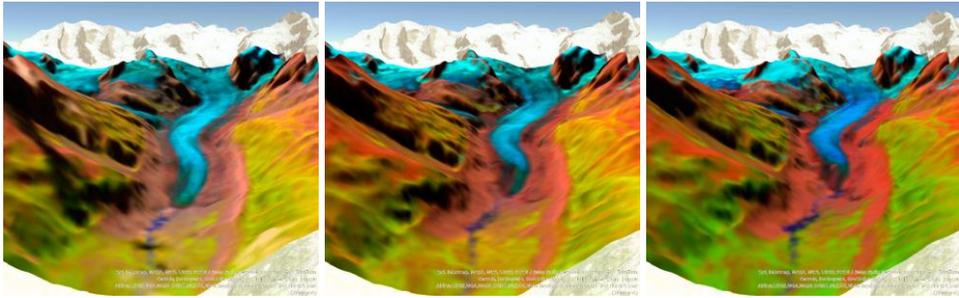
Fig. 4. Landsat 5 TM RGB composites from 2005, 2010 and Landsat 8 OLI - 2013 year

Sentinel-2 MSI RGB pseudo-composites 3D maps with band combination SWIR-NIR-Green for the years 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, and 2023 are shown in Fig. 5, 6, 7.
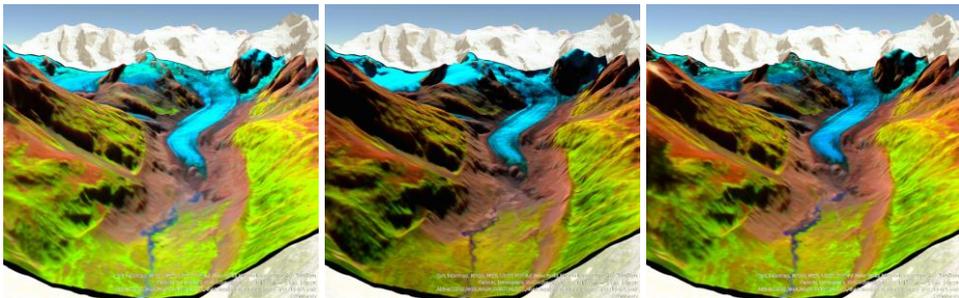
Fig. 5. Sentinel 2 MSI RGB composites from 2015, 2016 and 2017 year
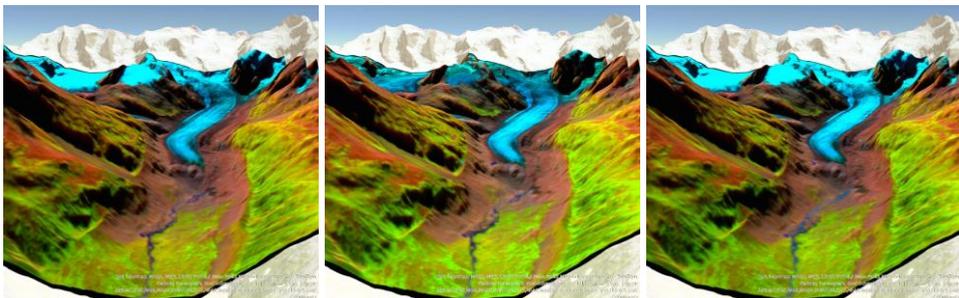
Fig. 6. Sentinel 2 MSI RGB composites from 2018, 2019 and 2020 year

The retreat of the glacier ice front for the period 2010–2015 is estimated at 385 m, for the period 2015–2019 is 240 m, and for the period 2019–2023 is 109 m.
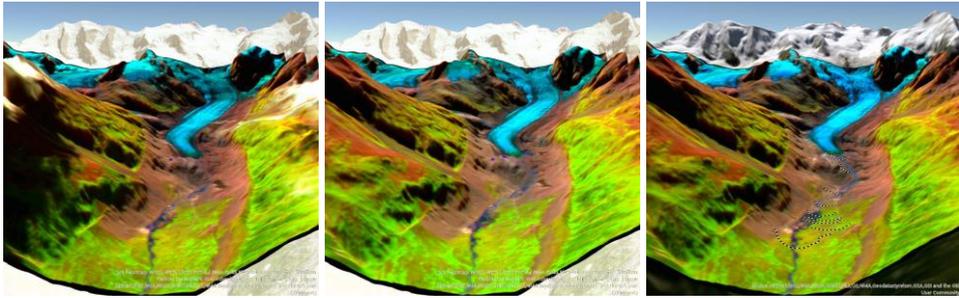
Fig. 7. Sentinel 2 MSI RGB composites from 2021, 2022 and 2023 year

Fig. 8 marks and graphically presents the positions of the glacier terminus and the changes that occurred in the length of Morteratsch from 1972 to 2023. The major retreat at a distance of about 844 m between 2000 and 2015 is easily visible.

The total change length of the Morteratsch terminus during the study period was approximately 1704 m. In addition, the width of the glacier measured in conjunction with Pers glacier has decreased by 582 m, and the area of the ablation zone (measured below the altitude line of 2600 m near the equilibrium line) of the glacier has seized approximately 2,37 km^2.
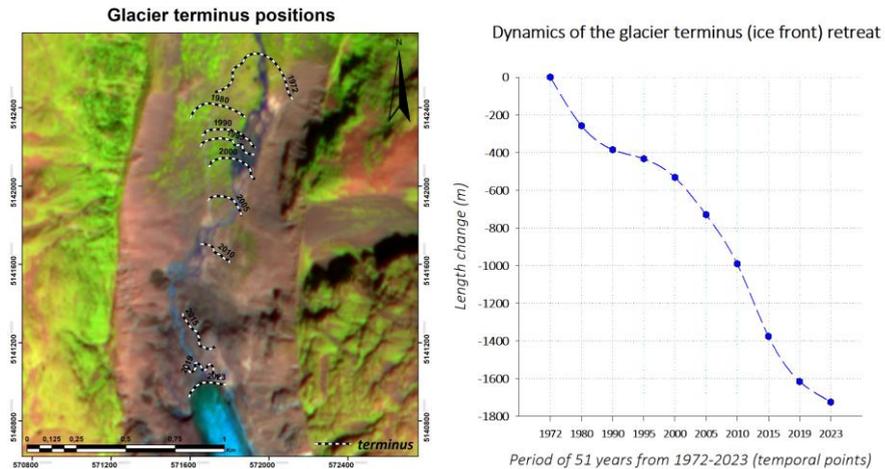
Fig. 8. Glacier terminus positions during the period 1972-2023 and graphic of its length change in m

The area size of the glacier ablation zone that decreased throughout the study period according to the processed data is presented as follows in Table 2.

Table 2. Changes in the area of the Morteratsch ablation zone during the study period

Year of the study period	Area in km^2
1972	3,41
1990	2,16
2000	1,62
2010	1,11
2020	1,06
2023	1,04

Conclusion

During the entire study, the temporal trend of the glacier dynamics was negative, with a change length of -1704 m. Until the year 1980, the retreat of the terminus was 258 m, showing a high melting rate, while for 1980–1990, the rate was slowing to 128 m per decade. For the period 1990–2000, the dynamics of the retreat were still keeping a slow movement of 146 m per decade. During the decade 2000-2010, the dynamics had a huge movement of 459 m, and the negative trend increased to nearly 630 m for the decade 2010–2020. Of the total length of 1704 m of the Morteratsch terminus retreat in a period of 51 years, nearly 1200 m of it has melted in the last 23 years, showing a trend of 52 m per year on average and around 550 m per decade.

Acknowledgments

References

1. https://creamontblanc.org/en/climate-change-and-its-impacts-alps/
2. Lee, E., Carrivick, J.L., Quincey, D.J. *et al.* (2021), Accelerated mass loss of Himalayan glaciers since the Little Ice Age. *Sci Rep* 11, 24284. https://doi.org/ 10.1038/s41598-021-03805-8
3. https://www.antarcticglaciers.org/
4. https://www.glamos.ch/en/factsheet#/E22-03
5. GLAMOS 1881-2023, The Swiss Glaciers 1880-2022/23, Glaciological Reports No 1-142, Yearbooks of the Cryospheric Commission of the Swiss Academy of Sciences (SCNAT), published since 1964 by VAW / ETH Zurich, doi:10.18752/glrep_series.
6. https://www.swisseduc.ch/glaciers/morteratsch/2005-2006/index-en.html
7.https://web.archive.org/web/20060903035209/http://www.glaciers.pdx.edu/MRNP/Res00 .html5. http://www.codeblocks.org/

8. https://earthexplorer.usgs.gov/
9. https://dataspace.copernicus.eu/
10. https://eos.com/

ПРИЛОЖЕНИЕ НА САТЕЛИТНИ ДАННИ ЗА МОНИТОРИНГ НА ПЛАНИНСКИЯ ЛЕДНИК MORTERATSCH, ШВЕЙЦАРСКИ АЛПИ, ЗА ПЕРИОД ОТ 51 ГОДИНИ

А. Стоянов

Резюме

Представеното изследване има за цел да приложи метод за наблюдение на намаляване територията на планински ледник в алпийския регион на Швейцарските Алпи, който се състои в използване на оптични сателитни изображения и техните спектрални възможности за наблюдение на снежни и ледени обекти на земната повърхност. Сателитните изображения със своя голям натрупан архив, датиращ от 1972 г. за изображенията от Landsat, могат да бъдат от голяма помощ за проследяване и наблюдение на намаляването на алпийските ледници за дълги периоди и да служат като база данни за моделиране и прогнозиране на динамиката на ледниците в бъдеще. Чрез комбиниране на различни подходи за обработка на сателитни данни са получени резултати за пространственото разпределение и динамиката на ледника Morteratsch за период от 51 години. Фокусът на изследването е да се проследят промените, които са настъпили по протежение на ледниковия език в зоната на *топене* на ледника, за да се придобие информация относно динамиката на ледника за изследвания период.

SYSTEM IDENTIFICATION OF AIRCRAFT LONGITUDINAL MOTION UTILIZING FLIGHT SIMULATOR DATA

Konstantin Metodiev

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

Abstract

In the presented article, an exemplary problem of identifying the parameters of an aircraft's motion within the vertical plane of symmetry is solved. The data was collected by FlightGear simulator for a Pilatus PC-9M aircraft flight case. After submitting a sequence of commands to pitch elevator and engine thrust, the simulator records five parameters of longitudinal motion at a certain frequency. The collected data are further used to build a state-space model of the flight through least-squares estimation. The obtained numerical results are compared with the experimental ones and are overlaid graphically. Charts depicting elevator angle to trim are also derived. A source code developed in GNU Octave was applied, and the problem was solved.

Introduction

The proposed study aims to develop a mathematical model of a linear time-invariant dynamic system by measuring input and output signals. The obtained model can be further used to examine system response to a broader range of input "stimuli." In this line of thought, a relatively easy-to-implement identification algorithm has been applied to a flight simulator model of Pilatus PC-9M in order to study the aircraft's longitudinal motion subject to control inputs of elevator angle and throttle percentage. A straightforward choice of aircraft properties to look into includes but is not limited to dynamic stability. Another important property is stick-fixed stability, which yields the elevator angular deflection to trim the aircraft at certain velocities and mass center locations.

The insufficiency of both analytical and experimental data with regard to PC-9M flying and maneuvering characteristics is somewhat significant and understandable. Presumably, one reason might be the military objectives the aircraft is meant to fulfill. The paper goes a short way towards filling that gap provided the model borrowed from FlightGear flight simulator, [1] is plausible enough for the obtained results to be credible and true.

System identification

The objective of system identification is to build up a mathematical model of a linear time-invariant system on the basis of observations. The model could be derived in both time and frequency domains and might be further used to look into the system response to different inputs. In the current study, preference is given to time domain analysis for simplicity reasons. In order to keep away from underlying complex theory, the employed system identification algorithm is solely described.

Consider the following state-space model of aircraft motion:

$$\dot{x} = Ax + Bu$$

$$(1) \quad \int_{t_0}^{t_1} \dot{x}\,dt = x(t_1) - x(t_0) = A\int_{t_0}^{t_1} x\,dt + B\int_{t_0}^{t_1} u\,dt$$

State variables x are gathered during the numerical experiment from time $t_0 = 0$ to t_1. Control inputs u have been assigned in advance. In eq. (1), matrices A and B are unknown. Sufficient measurements are available for all system states for the augmented matrix $//A\ B//$ to be estimated through the least squares approach as demonstrated in the thesis [2], for instance.

$$Y = \bar{A}.X = \|A \quad B\|.X$$

$$(2) \quad Y = \|x(\Delta t) \quad x(2\Delta t) \quad ... \quad x(n\Delta t)\| \quad X = \left\| \begin{matrix} \int_0^{\Delta t} x\,dt & \int_0^{2\Delta t} x\,dt & ... & \int_0^{n\Delta t} x\,dt \\ \int_0^{\Delta t} u\,dt & \int_0^{2\Delta t} u\,dt & ... & \int_0^{n\Delta t} u\,dt \end{matrix} \right\|$$

$$\bar{A} = YX^T\left(XX^T\right)^{-1}$$

In Fig. 1, a multistep control input to the system is depicted. Pulses with alternating widths, according to rule 3-2-1-1, are passed to the elevator. The throttle receives two-sided pulses with equal widths. The sampling frequency is 20 Hz and the total integration time is 20 sec. The weather conditions are set to the simplest possible. The aircraft is brought to a steady horizontal flight using the FlightGear native autopilot. Only then can the measurement begin. The control inputs are stored in a text file and passed to the simulator utilizing Yet Another Terminal, [3], a serial monitoring utility. A delay of 1000 ms is added to each line. The connection between the simulator and the serial monitor utility is established by UDP protocol.

116
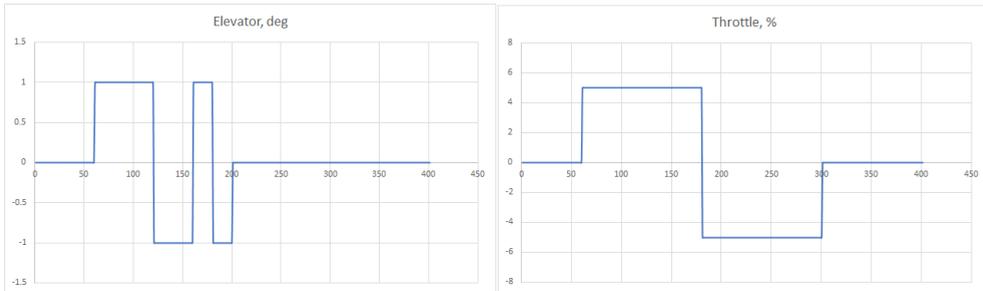
Fig. 1. Control inputs: elevator (left) and throttle

Elevator angle to trim

The PC-9M model has been produced using the Aeromatic++ utility, which comes with every JSBSim distribution [4]. A *.param file containing some basic data is required in advance for the utility to generate project files. The outcome is elaborate *.xml files containing data about airplane and engine performance, mass and balance, flight dynamics, control systems, etc. For example, the lift coefficient data shown in Fig. 2 might also be found in the main configuration PC-9M.xml file (<function name="aero/force/Lift_alpha">). In Fig. 2, a small fraction of the lift coefficient chart is shown according to angle of attack variation, deg. The lift coefficient data have been suggested by Holfman and Culp [5] and implemented in the configuration PC-9M.xml file in the JSBSim/aircraft directory. The curve slope was determined to be 5.218235 rad^{-1}. Savov and Marinov quote similar experimental results in paper [6] as follows: 5.01 rad^{-1} @ M = 0.4 and 6.2 rad^{-1} @ M = 0.6, which agrees well with the aforementioned data.
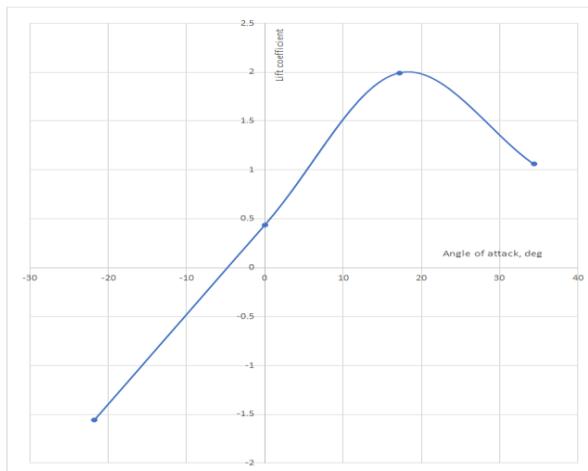
Fig. 2. Lift coefficient vs. angle of attack (deg), PC-9M.xml model

117

An input *.xml script is passed to the JSBSim executable as a command line option. Among other important lines, the script contains the so-called events for the program to execute. An exemplary event to point out is basic aircraft trimming carried out after setting the property "simulation / do_simple_trim" to a specified step value, exponential approach, or ramp. For example, consider the following line <set name="simulation/do_simple_trim" value="1"/>. JSBSim works out the elevator angle to trim the aircraft (stick fixed) at a given altitude and airspeed. Exemplary scripts are available in the JSBSim installation directory. A good tutorial on the matter still exists and might be found in the link [7]. An extensive description of xml schema used by JSBSim is available in the link [8].

It is somewhat important to mention that JSBSim succeeded in solving the assigned task for velocity values of 180 kts or less.

FlightGear photo realism

A project called FlightGear PhotoScenery, uploaded in link [9], allows the user to overlay existing scenery textures with realistic satellite orthophotos whenever available at the aircraft current longitude and latitude. A Python script downloads the requested scenery tile from specific ArcGIS servers [10]. In the FlightGear splash screen, tiles location on the drive must be specified in the relevant text box (Add-ons/Additional scenery folders). Photo realistic ability is invoked in the flight simulator environment by the View/Rendering Options/Satellite Photoscenery dialog. In Fig. 3, PC-9M flying over LB41 (ICAO ID) airstrip is shown in order to make a comparison between realistic and default textures. In the left half, credible environmental features, such as the SuperMoto circuit "Dolna Mitropolia – Pleven," are clearly distinguishable.

Fig. 3. PC-9M flying over LB41. The left half depicts a photo realistic scenery.

Results

In Fig. 4, results obtained after system identification completion are shown. The identified data are computed using a script developed in the GNU Octave environment [11]. The state vector is $\|V, \alpha, q, \theta, H\|$, i.e., total velocity, angle of attack, pitch rate, pitch angle, and altitude. The control input is $\|\delta e, \delta th\|$, i.e., elevator and throttle, according to Fig. 1. The script used to work out the numerical solution might be found in Appendix 1 alongside an extract of the input data file, Appendix 2. Initial conditions could be read in the first line of the file containing gathered data. An extract of the input file (raw data) can be seen in Appendix 2.
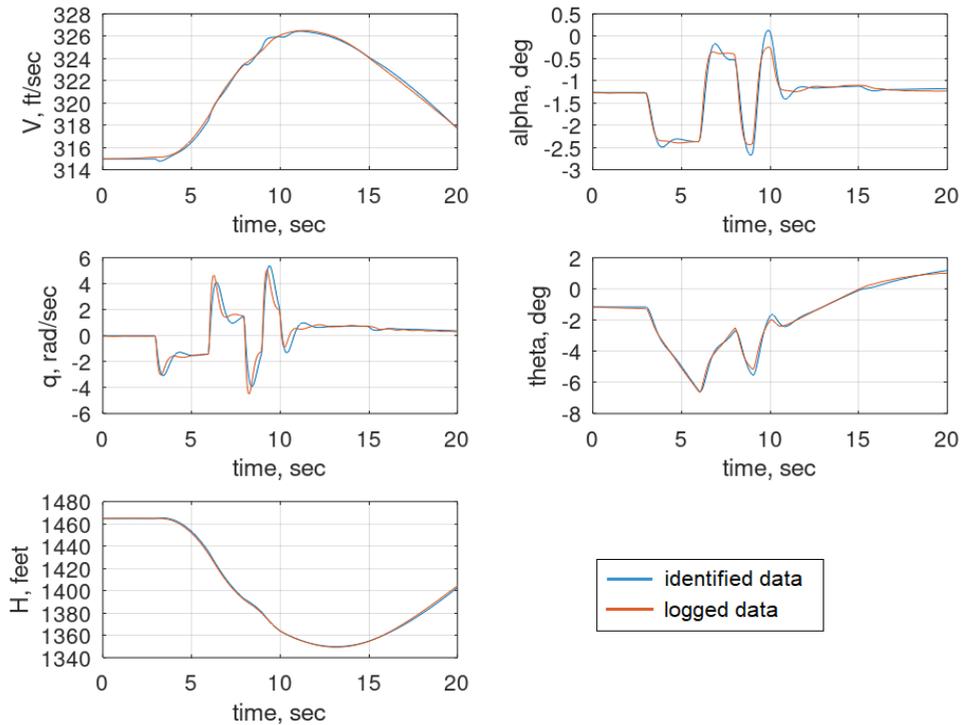
Fig. 4. Recorded and identified data

The number of samples acquired is ~400, i.e., the sampling frequency is 20 Hz and total integration time is 20 sec, as it was already mentioned. It yields an integration interval long enough for the identified data to match the collected ones with sufficient precision.

In Fig. 5, a screenshot taken from the GNU Octave console window shows elements of the augmented matrix $\overline{A} = \|A\ B\|$ according to eq. (2).

```
>> sys

sys.a =
              x1            x2            x3            x4            x5
   x1    -0.06688        -9.227       -29.86       -35.75     -0.005529
   x2    9.589e-05        -1.588        1.021      0.02475     8.999e-05
   x3    -0.004134       -11.76        -2.246       0.1361    -0.0006117
   x4     0.001439     -0.008806        1.139      0.04919     0.0001737
   x5      0.2025        -289.2         32.65         329       0.01836

sys.b =
            u1          u2
   x1    -99.88       1.465
   x2   -0.08787    0.005515
   x3    -13.95      -0.5265
   x4     0.2332     0.01632
   x5      124.3       9.932

sys.c =
        x1   x2   x3   x4   x5
   y1    1    0    0    0    0

sys.d =
        u1   u2
   y1    0    0

Continuous-time model.
>>
```

Fig. 5. Results reported by GNU Octave

In Fig. 6, results reported by JSBSim after doing a simple trim are shown. Curves are drawn for various values of flight speed and center of gravity location.
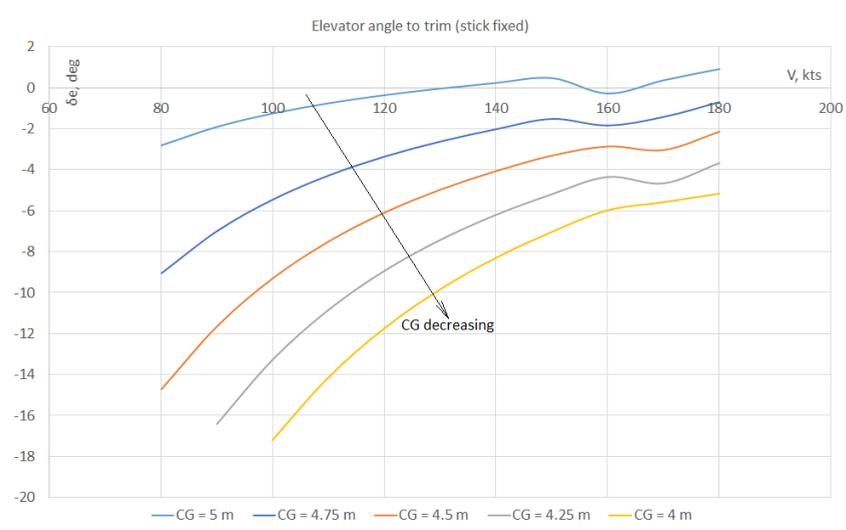
Fig. 6. Elevator angle to trim (stick fixed)

Conclusion

The identified model precision might be improved further by tuning both the integration step and interval (data gathering time). It takes numerous experiments to guess the right settings, though. The "trial and error" approach is a last resort for no evident rule exists with regard to what the best choice of settings might be. Additional experiments are to be carried out with derived model to verify its feasibility including test flights onboard the real machine.

References

1. https://www.flightgear.org/
2. Cetin, E., System Identification and Control of a Fixed Wing Aircraft by Using Flight Data Obtained From X-Plane Flight Simulator, M.Sc. thesis, Middle East Technical University, 2018
3. https://sourceforge.net/projects/y-a-terminal/
4. https://github.com/JSBSim-Team/jsbsim/releases
5. http://mirrors.ibiblio.org/flightgear/ftp/Aircraft/PC-9M.zip
6. Savov, V., A. Marinov, Aerodynamic Characteristics Calculation of Pilatus-9M Aircraft during Longitudinal and Lateral Motion, Proceedings of Scientific Conference "40 Years since a Man Stepped on the Moon and 30 Years since the First Bulgarian Cosmonaut Flight," p.p. 66-71, vol. 1, Air Force Academy, Dolna Mitropolia, 2009, in Bulgarian
7. http://www.holycows.net/JSBSim/index.htm
8. http://jsbsim.sf.net/JSBSimScript.xsd
9. https://github.com/nathanielwarner/flightgear-photoscenery
10. https://services.arcgisonline.com/arcgis/rest/services/World_Imagery/MapServer/
11. https://octave.org/

Appendix 1. Source code in GNU Octave for system identification

```
% >> pkg load control
clear;
t = 0.:.05:20.;
[r, c] = size(t);
xfile = dlmread('in.csv', ',', 1, 0); % CSV file with data
x = zeros(7,c); % Fill in data
for i=1:7
  x(i,:) = xfile(:,i);
end
% Elevator and throttle inputs
x(6,:) = x(6,:)*pi/180.; x(7,:) = x(7,:)/100.;
x0 = zeros(1, 5); x0 = x(:,1); % Vector of initial conditions
for i=1:5 % V (ft/s), α (rad), q (rad/s), θ (rad), H asl (ft)
  x(i,:) = x(i,:) - x(i,1);
end
```

```matlab
X = zeros(7,c);
% Trapezoidal numerical integration
for j = 1:7
  for i = 1:c
    X(j,i) = trapz(t(1:i),x(j,1:i));
  end
end


Y = zeros(5,c); Y = x(1:5,:);


a = Y*X'*inv(X*X');


A = zeros(5,5); A(:,:) = a(:,[1:5]);
B = zeros(5,2); B(:,:) = a(:,[6:7]);
C = [1., 0., 0., 0., 0.]; D = [0., 0.];


sys = ss(A,B,C,D);
[y,t,z] = lsim(sys,x([6,7],:)',t);

% Revert initial conditions
for i=1:5
  if i == 2 || i == 3 || i == 4
    z(:,i) = z(:,i)*180./pi;
    x(i,:) = x(i,:)*180./pi;
  end
  z(:,i) = z(:,i)+x0(i);
  x(i,:) = x(i,:)+x0(i);
end
```

Appendix 2. A few lines extracted from the input file

	A	B	C	D	E	F	G
1	V_{Total} (ft/s)	Alpha (rad)	Q (rad/s)	Theta (rad)	Altitude ASL (ft)	Elevator, deg	Throttle, %
2	314.978696	-0.022108261	-0.000324947	-0.020274211	1464.838392	0	0
3	314.9776457	-0.022108982	-0.000369167	-0.020290665	1464.852321	0	0
4	314.9766597	-0.022111421	-0.000404946	-0.020309048	1464.866002	0	0
5	314.9758003	-0.022115363	-0.000444505	-0.020329195	1464.879432	0	0
6	314.9751434	-0.022121083	-0.000495619	-0.020351454	1464.892604	0	0
7	314.9747671	-0.022129122	-0.000562202	-0.020376484	1464.905511	0	0
8	314.9747416	-0.022140088	-0.000644654	-0.020405068	1464.918141	0	0
9	314.9750814	-0.022154479	-0.000739169	-0.020437945	1464.93048	0	0
10	314.9757056	-0.022172445	-0.000837242	-0.020475539	1464.942509	0	0

in ⊕

ИДЕНТИФИКАЦИЯ НА ПАРАМЕТРИТЕ НА НАДЛЪЖНОТО ДВИЖЕНИЕ НА САМОЛЕТ ПО ДАННИ ОТ ФЛАЙТ СИМУЛАТОР

К. Методиев

Резюме

В настоящата статия е решена примерна задача за идентификация на параметрите на движение на самолет във вертикалната равнина на симетрия. Данните са събрани от симулатор FlightGear за случай на полет на самолет Pilatus PC-9M. След подаване на последователност от команди към кормилото за височина и тягата на двигателя, симулаторът регистрира с определена честота пет параметъра на надлъжното движение. Събраните данни се използват за построяване на модел на полета в пространство на състоянията чрез оценка по метода на най-малките квадрати. Получените числени резултати са сверени с експерименталните и са показани в графичен вид. Балансировъчните криви на кормилото за височина са допълнително определени. Приложен е сорс-код на GNU Octave, с помощта на който е решена поставената задача.

HYBRID MULTICOPTER DESIGNS

Svetoslav Zabunov, Garo Mardirossian

Space Research and Technology Institute — Bulgarian Academy of Sciences

Abstract
  Hybrid unmanned aerial vehicles are vehicles that employ more than one type of energy delivery system for powered flight. Hybrid drones are not restricted to using electrical power. Nevertheless, most such drones employ electric propulsion by implementing brushless three-phase electrical motors in their rotors due to the inherent high efficiency and ease of control of such devices.
  Electrical energy for the motors is supplied by various sources such as batteries, hydrogen fuel cells, internal combustion engine electrical generators, etc. The current battery technology offers flight times from 10 minutes to 2 hours, depending on the payload and drone design. For longer flight times, a hybrid power system is required. The common hybrid drone design involves a combination of batteries and an internal combustion engine electrical generator.
  The current article introduces two novel hybrid unmanned aerial vehicle designs. The first exhibits improved efficiency due to the implementation of a hybrid power system powering two horizontal tandem counter-rotating direct driven rotors and at least four electrically driven smaller horizontal attitude control rotors. The second presented herein invention offers the addition of a horizontal rotor or rotors, rendering the design a compound multicopter. The vertical rotor or rotors are also direct driven improving efficiency.

Introduction

Hybrid unmanned aerial vehicles (UAVs) are drones that utilize more than a single type of energy delivery system to enable the powered flight of the vehicle.

Although hybrid UAVs are not limited to electrical propulsion systems, most such vehicles rely solely on electric propulsion based on brushless three-phase electric motors. Brushless electric motors exhibit high efficiency and are easy to control.

The electrical energy supply for the motors is delivered from different sources such as batteries, hydrogen fuel cells, internal combustion engine electrical generators, etc. The current battery technology, most commonly used, is Li-ion or Li-poly batteries, securing energy densities in the range of 200–250 Wh/kg and flight times from 10 minutes to 2 hours, depending on the payload and drone design.

The most common hybrid drone combines batteries with an internal combustion engine generator [1–3]. The second most common variant employs a hydrogen fuel cell [4–5].

Internal combustion engine-based generators are cheaper in comparison with hydrogen fuel cells. Hence, these are implemented in the overwhelming number of hybrid drone cases. Two-stroke and four-stroke engines are common in such scenarios, although the four-stroke solutions are heavier. Despite their higher energy efficiency, they offer shorter flight times than the two-stroke designs. By running the internal combustion engine, the coupled electrical generator produces electrical energy that powers the electrical motors. The excess electrical energy produced by the generator is stored in a buffer on-board battery. During power surges, high currents are drawn from the battery. The latter is also employed in cases of emergency when the generator of the internal combustion engine fails.

It should be mentioned that besides hybrid unmanned aerial vehicles, land and sea-based drones powered by hybrid power supplies exist [6].

The current article introduces novel hybrid unmanned aerial vehicles employing improved efficiency due to not only a hybrid power system but also a hybrid propulsion system.
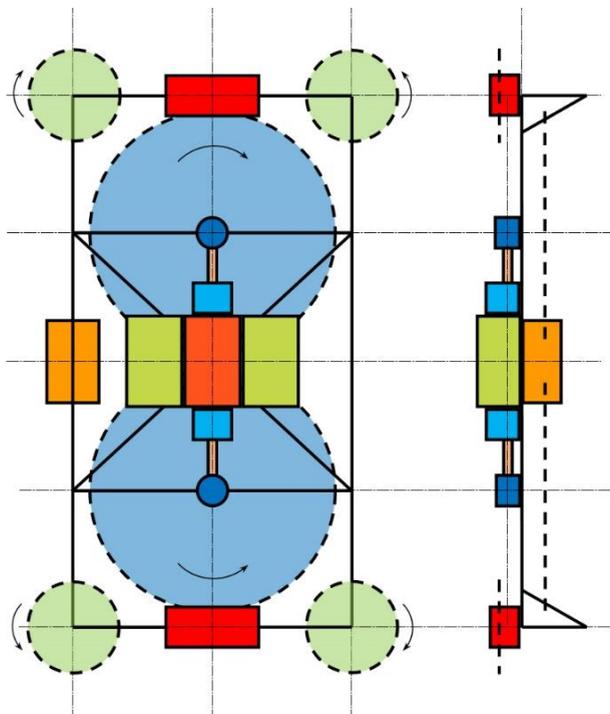

Fig. 1. Double-hybrid unmanned aerial vehicle, model XZ-15

Double-hybrid UAVs

A hybrid drone has many power sources and a single propulsion, namely brushless electrical motors. A double-hybrid adds a multitude of propulsion systems. The current material introduces two novel hybrid drone designs offering not only hybrid power sources but also hybrid propeller propulsion systems involving both propellers driven by brushless electrical motors and direct drive propellers. The efficiency of the power output of the directly driven propellers is significantly higher than that of the electrical motors coupled propellers due to the lack of power conversion through generators and motors, but their controllability is poor. On the other hand, the control of the brushless electrical motors is superior. By combining both types of propulsion and implementing the brushless electrical motor-driven propellers only for attitude control of the drone, we achieve higher flight efficiency, which translates into longer flight times, heavier payloads, longer ranges, etc.

Herein, we present two innovations: a hybrid multicopter having a tandem of direct-driven rotors and four electrically driven rotors [7] (see Fig. 1). The second invention is a similar design but with added vertical rotors for high-speed flight, thus creating the so-called compound multirotor [8] (see Fig. 2).

Double-hybrid hexa-rotor

The double-hybrid design shown in Fig. 1 is the model XZ-15 of the XZ unmanned aerial vehicle series developed at the Space Research and Technology Institute – Bulgarian Academy of Sciences in 2019. The model is a hexa-rotor multicopter having two major rotors positioned in the middle of the fuselage in counter-rotating co-planar tandem and four smaller rotors mounted at the airframe's corners. The large centre rotors are driven directly by the on-board internal combustion engine. This engine may be any type, but for higher efficiency, a two-stroke diesel engine or a four-stroke petrol variant may be employed. The XZ-15 UAV may be used for camera observations, remote sensing, cargo delivery, etc.

Another innovation is the positioning of the two large rotors, which are mounted under the fuselage, thus ensuring higher flight efficiency. This higher efficiency is due to unobstructed by the fuselage's high-speed airflow accelerated by the two major rotors. The two major rotors create the majority of the lift. They employ fixed-pitch propellers for simplicity, reliability, and weight savings and thus are unable to offer attitude control. The latter is secured by adding a minimum of four electrically driven rotors (in Fig. 1, there are four additional rotors). These electrically driven rotors are powered by the hybrid drone's power system. An electrical generator is coupled to the internal combustion engine and generates electrical power. The electrically driven rotors, using brushless three-phase motors, offer fast response and appropriate attitude control of the aerial vehicle.

Earlier inventions of double-hybrid UAVs do exist, such as the hybrid unmanned aerial vehicle having a single major rotor directly driven by an internal

combustion engine and an additional four smaller electrically driven rotors [9]. This design is inferior to XZ-15 due to the implementation of a single major rotor instead of a counter-rotating tandem of rotors. This configuration offers lower efficiency. Another drawback of the previous designs is the standard mounting of the rotors over the fuselage, which yields lower efficiency.

Another earlier invention of a hybrid power plant for UAVs is presented in [10], where the concrete drone design is not disclosed.

The major drawback of previous designs is the lower efficiency due to the mounting of rotors over the airframe. Such an approach predisposes the high-speed rotor airflow to create drag in a downward direction when meeting the fuselage. Thus, the drone exhibits increased dynamic pressure, and consequently, higher power is required to sustain flight. A number of drawbacks follow, such as shorter range and flight times, less payload weight available, etc. The previous design's single major rotor approach creates a reactive moment that needs to be counteracted by the smaller electrically driven rotors, but their efficiency is much lower; thus, power is wasted. The tandem counter-rotating design creates virtually no reactive moment in the vertical axis.

XZ-15 (Fig. 1) is based on a rectangular airframe constructed in the horizontal plane. All rotors are horizontal. The major tandem rotors are directly driven through shafts from the internal combustion engine. The latter also drives two electrical generators. Two generators are proposed, instead of one, for redundancy and reliability. The generators provide electrical power for the four corner-mounted rotors and also to charge a buffer secondary battery. The drone also carries a payload.

Double-hybrid compound multicopter

In Fig. 2 is presented the double-hybrid compound multicopter design XZ-15C of the XZ unmanned aerial vehicle series. This model was developed in 2023 again at Space Research and Technology Institute – Bulgarian Academy of Sciences. The model is a septa-rotor multicopter design that differs from the XZ-15 in that it only has an additional seventh vertical rotor mounted at the tail of the fuselage. This additional rotor is directly driven from the internal combustion engine through a shaft and a clutch. The benefit of mounting vertical rotors in multicopters is that such multicopters achieve higher horizontal speeds. This type of multicopters is called compound multicopters. Another benefit is that the efficiency of horizontal flight increases due to lower aerodynamic drag of the fuselage and increased efficiency of the production of horizontal aerodynamic propulsion forces.
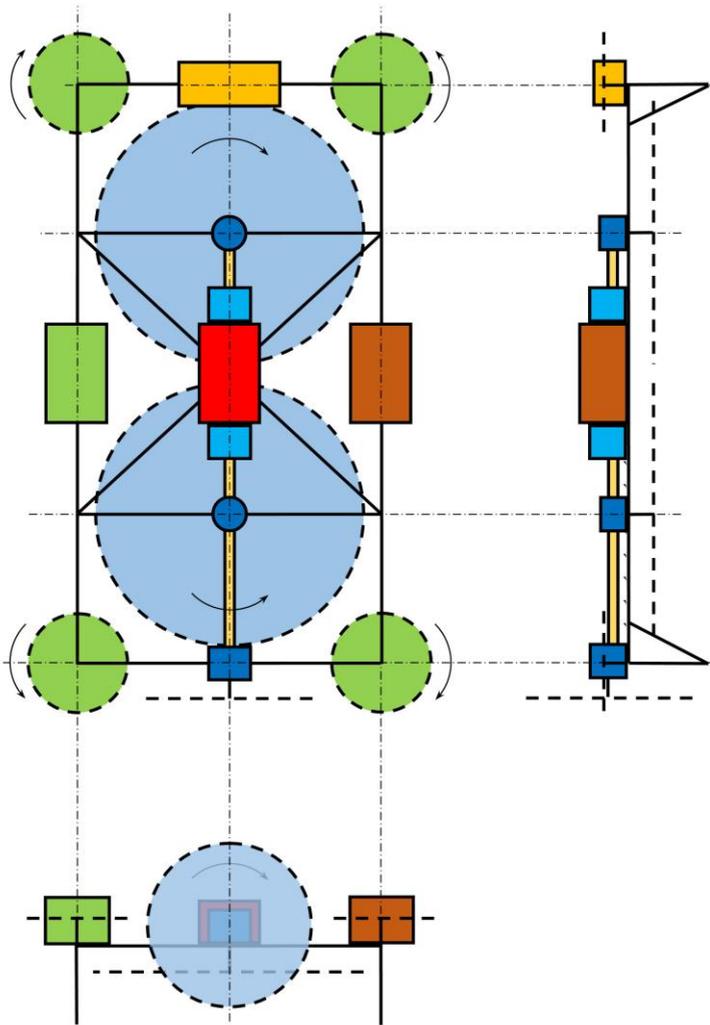
Fig. 2. Double-hybrid compound unmanned aerial vehicle XZ-15C

Conclusions

The introduction of directly driven tandem horizontal rotors in hybrid multicopters increases the efficiency of flight and thus prolongs flight time, extends range, and makes possible the transport of larger and heavier payloads. Further, by implementing vertical directly driven rotors in hybrid multicopters established the class of hybrid compound multicopters exhibiting superior performance to any other hybrid multicopter in terms of horizontal speed and efficiency of the horizontal flight.

References

1. Gang, B., K.-B. Kim. (2022). Performance evaluation of a compact designed electric power system composed of a two-stroke gasoline engine, generator, and battery for high-endurance unmanned aerial vehicles. *Energy Science & Engineering*. vol. 10, iss. 10, pp. 4071–4084.
2. Valencia, E., D. Jimenez, V. Alulema, I. Roumeliotis, J. Montalvan, M. Pozo, E. Cando. (2020). Modeling of a Series Hybrid Propulsion UAV Used for Monitoring in the Galapagos Islands. *AIAA Propulsion and Energy 2020 Forum,* (2020-3960), Session: Unmanned Aircraft Systems II
3. Yinka-Banjo, C., O. Ajayi. (2019). Sky Farmers: Applications of Unmanned Aerial Vehicle (UAV) in Agriculture. Chapter in book *Autonomous Vehicles*. IntechOpen. doi: 10.5772/intechopen.73376
4. Liu, H., Y. Yao, J. Wang, T. Yang, T. Li. (2022). Energy management and system design for fuel cell hybrid unmanned aerial vehicles. *Energy Science & Engineering*. vol. 10, iss. 10, pp. 3987–4006.
5. Boukoberine, M., Z. Muhammad, M. Benbouzid, Z. Zhou, T. Donateo. (2021). Hybrid Fuel Cell Powered Drones Energy Management Strategy Improvement and Hydrogen Saving using Real Flight Test Data. *Energy Conversion and Management*. vol. 236, Article 113987.
6. Bui, T., T. Dinh, J. Marco, C. Watts. (2021). Development and Real-Time Performance Evaluation of Energy Management Strategy for a Dynamic Positioning Hybrid Electric Marine Vessel. *Electronics*. vol. 10, iss. 11, pp. 1280–1308.
7. Zabunov, S., G. Mardirossian, R. Nedkov (2019). Hybrid Multicopter. *Patent office of Republic of Bulgaria*, utility model reg. no. 3175 / 28.05.2019.
8. Zabunov, S., G. Mardirossian, G. Jelev (2023). Hybrid Compound Multicopter. *Patent office of Republic of Bulgaria*, utility model reg. no. 4476 / 28.06.2023
9. Yasuda, K., T. Yamanashi. (2014). Parallel hybrid multi-rotor aircraft, *Japan patent,* JP2015137092A
10. De Soto, J., J.F. Miranda. (2017). Hybrid Power Supply For Electric Multirotor Rotorcraft, *US Patent Office*, US 2017/0008627 A1

КОНСТРУКЦИИ НА ХИБРИДНИ МУЛТИКОПТЕРИ

С. Забунов, Г. Мардиросян

Резюме

Хибридните безпилотни летателни апарати използват повече от един вид източници на мощност за осигуряване на полета. Хибридните дронове не са ограничени до използване само на електрическа мощност. Въпреки това повечето такива дронове разчитат на електрическа енергия и безколекторни трифазни електродвигатели за задвижване на роторите, поради високия коефициент на полезно действие на този вид двигатели и не сложното им управление. Електрическа мощност за задвижване на моторите се предоставя

от разнообразни източници като батерии, водородни горивни клетки, двигатели с вътрешно горене, свързани с електрически генератори и др. Текущите технологии на електрически мултикоптери предлагат полетно време от 10 минути до 2 часа, в зависимост от дизайна на летателния апарат. За постигане на по-дълго полетно време се налага използването на хибридна енергийна система. Стандартният дизайн на хибриден дрон включва двигател с вътрешно горене, задвижващ електрически генератор.

Настоящата статия представя два иновативни хибридни мултикоптера. Първото изобретение гарантира подобрен коефициент на полезно действие поради използване на хибридна енергийна система и два хоризонтални ротора, въртящи се в противоположни посоки и задвижвани директно от двигателя с вътрешно горене. Управлението на полета се извършва от най-малко четири допълнителни ротора, задвижвани от електродвигатели. Второто изобретение предлага допълнителен вертикален ротор и така превръща конструкцията в хибриден съставен мултикоптер. Вертикалният ротор е също директно задвижван от двигателя с вътрешно горене.

INTELLIGENT MONITORING AND PROTECTION SYSTEM OF CRITICAL INFRASTRUCTURE BASED ON MOBILE COMMUNICATION-INFORMATION SYSTEM WITH ELEMENTS OF ARTIFICIAL INTELLIGENCE

Evgeni Hubenov, Zoya Chiflidjanova

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

Abstract

The paper proposes innovative modern solutions for an integrated mobile monitoring system with elements of artificial intelligence in the Internet environment based on a mobile communication-information system for collecting, aggregating, processing, and presenting in real-time streams of information objects. It presents the management of communication networks to optimize the transport and processing of information objects based on discrete-event data flow representation and modeling with a hierarchical structure. Qualitative aspects are considered for networks designed for monitoring and notification of geophysical, climatic, and other natural phenomena as well as for anthropogenic systems. A comparison of centralized and decentralized management capabilities is made, as well as the state of the art of network technologies and the possibilities for practical implementation in different network architectures.

Introduction

Critical infrastructure (CI), considered a complex hierarchical system structured in elements, connections, and relationships between them, has a systemic goal (desired systemic property, result) of building and improving its management and protection [1]. The three main components of CI management are organization, communication and information support, and information. Management support encompasses all personnel, systems, and force resources that support the delivery, transport, and processing of information flows provided by specific system functions of the information system. The primary purpose of management support is to enhance the ability to make and implement decisions. Important elements of management support are the information systems, equipment, software, and infrastructure that enable the management of the system and the identified resources. These systems further help the manager monitor and

influence his resources through the hierarchical chain. The CI protection and management system functions as a set of subsystems with unified management, providing specific objectives with their own information space (ISp). The ISp is a set of information resources and infrastructures for information access with structuring, navigation, and transport capabilities.

The critical infrastructure monitoring information system (CIMIS), which provides the information necessary to achieve the system goals, is an information space in which all the elements support the management processes functions. Its components have a certain structure and hierarchy, which provide a connection with one or more information subsystems for CI protection. An information system (IS) is a set of interconnected means, methods, and personnel that are used to store, process, and provide information, knowledge, and digital services [2] to achieve a set goal. In a narrow sense of the concept, IS is considered an information cluster, which is a set of information blocks with signs of a grouping of elements.

The management of the monitoring information system (MIS) must be carried out uniquely within the monitoring system structure, and the results must be presented in other subsystems with the necessary level of security to protect the information. If information from other sources outside the MIS (e.g., satellite information or open sources) is required for CI protection purposes, it should be in the form of an information service [3].

Access to reliable, accurate, and timely information at all levels of society is critical just before, during, and after a disaster. Without information, people and institutions are often forced to make critical decisions based on fragmented, contradictory messages or on the basis of "guesswork". Information on disaster risk and subsequent events should also be made available to the general public as one of the stakeholders in the disaster risk management process. Information and communication technologies (ICTs) have advantages in information dissemination and management that can and should be used to improve disaster and emergency risk management.

Communication-Information Monitoring System

The system-information approach focuses on the processes of receiving, transporting, processing, and presenting information in the system and its interaction with the external environment and critical infrastructure protection subsystems. This information interaction implies the construction of monitoring as a communication and information system (CIS) [4]. An organizational-structural approach defines a CIS as a set of technical (including communication means, border protection devices, cryptographic means, and signal distribution medium within the system boundaries) and programmatic means, methods, procedures, and personnel organized to perform one or several of the functions of creating, processing, using, storing, and exchanging (classified) information in electronic

form. CIS is a complex distributed spatio-temporal discrete-event system that includes in its structure two interconnected subsystems (communication and information), which are created for the purpose of transporting and processing flows of information objects [5]. The information object streams are of different subjects and enter the CIS system from different sources in or related to the CI areas. The purpose of CIMIS is to detect, classify, identify, and monitor CI-related events in a timely manner and provide real-time situational awareness to support CI protection decision-making.

Event-oriented CIS include hardware and software components that operate simultaneously in different domains of interconnected communication environments, with the use of events as the main object to organize the dynamic communication between components and adapt the structure to the data flow parameters. The dynamics of these systems are related to the occurrence of physical events at previously unknown, irregular moments of time. In general, an event is referred to as a change of state of the system in the discrete state space. Various information flows are formed in a CIS for monitoring and control purposes. For the users of the system, changes in monitoring parameters directly related to CI protection or events related to disasters, accidents, and catastrophes are of interest. For the purpose of maintaining the functional characteristics of the CIS and for the system operator, the information flows resulting from the interaction and interdependence of the CIS with other ISp elements and directly related to the management of the CIS are important. Only information flows related to CI monitoring will be the subject of this paper.

The architecture of data stream processing systems links the communication environment and network transport to information processing, including its aggregation for complex event processing in large systems. These systems are communication and information systems (CIS) insofar as the viewpoint and approach are from the network layer side of the communication model. From the perspective of the application layer of the network model (OSI, TCP/IP), the formation, movement, transformation, processing, and presentation of information objects give rise to the simultaneous use of the term information-communication system (ICS). Data flow can be viewed as information about events or things that have happened within an external system or domain. The same term is also used to refer to the object that represents what has happened in the ICS. Event-driven ICSs include hardware and software components that operate simultaneously across domains in interconnected communication environments, with the use of events as the primary object to organize dynamic communication between components and adapt the structure to the parameters of the data stream.

Cognitive (rational-mental segment) of the monitoring system

The need for diverse, timely, accurate, and adequate information on the state of CI and for making timely management decisions related to the prevention of the possible consequences of various contingencies leads to the need for the use of information systems (IS) that monitor the various possible states of CI, the various influences on it, as well as its behavior patterns. Moreover, the more different information-gathering systems there are, the greater the reliability of the information obtained.

Creating an IS for decision support based on emergency forecasting that combines problem solving for all types of natural and man-made emergencies [6] and allows for rapid response is an important and labor-intensive task. At the same time, the use of different monitoring systems necessitates the development of a conceptual approach to form a unified information space (UISp) of CI.

In this case, the UISp is a set of information tools and resources integrated into a single system, namely:

- its own information resources (document sets, databases and data banks, archives of all kinds, etc. containing information recorded on appropriate media);
- network and special software;
- telecommunications network (distributed corporate computer networks, telecommunications networks and systems for special and general use, data transmission networks and channels, means of switching and managing information flows).

Principally, IS is implemented on two levels:

- As an automated workplace (AWP) using a computer.
- In the form of a local computing network, connecting two to several computers (workstations) and peripherals within the CI to access common resources and exchange information.

In practice, it is possible to distinguish between natural, technogenic, and natural-technogenic emergencies. For all three classes of emergencies, it is first and foremost necessary to address the problems of early forecasting. When emergencies occur, it is necessary to make forecasts of their development and consequences. It is recommended that the whole IS structure be divided into three subsystems, each of which performs its functions:

The emergency prediction subsystem meets the following functional requirements:

- Provide analysis of monitoring and forecast information on the sources of accidents;
- Develop forecasts for the occurrence and development of emergencies;
- Provide for the creation and maintenance of a database of forecasts of the occurrence and development of emergencies and the data to substantiate them;

134

• Ensure monitoring and forecast data processing to identify new, more effective predictive relationships between emergency sources, their causes, conditions, and development parameters.

The subsystem for forecasting the consequences of man-made accidents meets the requirements:

• Provide predictive analysis of consequence assessment data and predictions;

• Develop forecasts of the consequences of man-made accidents;

• Ensure the establishment and maintenance of a database of forecasts of the consequences of man-made accidents and their degree of substantiation;

• Ensure the processing of monitoring and forecast data to establish new, more effective predictive relationships between the parameters and conditions for the occurrence, development, and progression of man-made accidents and their consequences.

The compliant decision support subsystem:

• Present forecasts of emergency occurrence and development in the form of forecast bulletins for management review and approval;

• Ensure the preparation of regulatory documentation for sites;

• Ensure the production of documents required by the rapid response services.

The following methods may be used to assess the consequences of emergencies:

• Assessment of the consequences of accidents in fire- and explosion-hazardous facilities;

• Prediction and assessment of the medical consequences of accidents in explosive and fire-hazardous installations;

• Prediction of the magnitude of contamination with highly toxic substances in accidents (destruction) of chemically hazardous facilities and transport;

• Prediction of possible accidents, catastrophes, and natural disasters;

• Damage assessment of emergencies of technogenic, natural, and terrorist nature, as well as classification and reporting of emergencies.

Figure 1 shows a typical functional diagram of a decision support system based on emergency monitoring and forecasting, which includes developed base software.
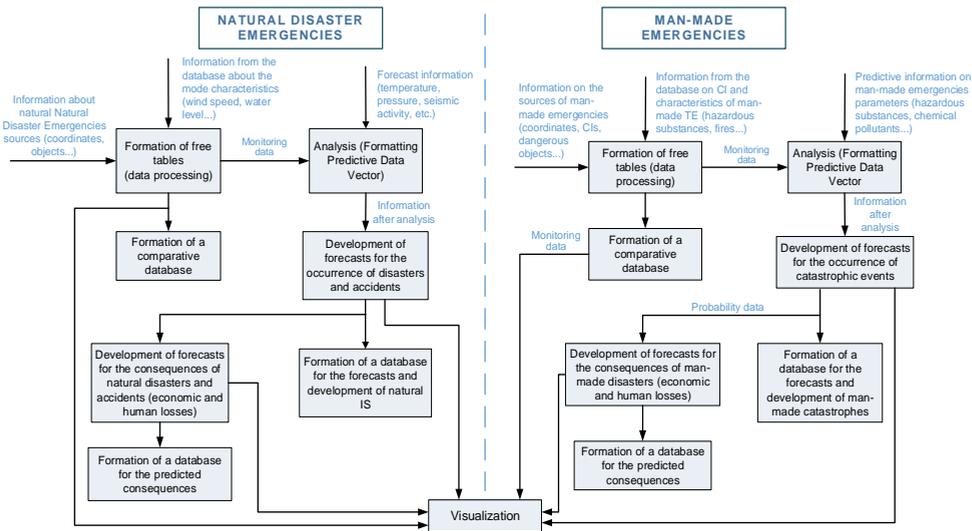
Fig. 1. Functional diagram of a decision support system information

The state must have a disaster information strategy to provide information that can be used for pre-disaster preparation, emergency response during a disaster, and damage and loss assessment, recovery, and reconstruction after a disaster [7]. Such baseline information can be gathered through careful mapping of risk areas and assessment of regions prone to major disasters.

Data are measurements or observations with variable magnitude (e.g., population size), classes (e.g., ethnic groups), or images (e.g., photographs). After analysis, the original data are converted into information that, through useful information extraction operations, is used for decision-making and action.

Information is translated into disaster risk knowledge through a learning process, and the timely and correct application of knowledge is translated into practical activities on the ground. Practice, in turn, produces new data that can be collected and analyzed. Thus, the entire information management cycle is not a linear process; rather, it is the information management cycle that continuously moves in a circle.

The development of intelligent management systems is one of the most important tasks nowadays, when computerization and intellectualization of vast areas of our lives are the basis for solving modern management problems. Modern automated control systems are essentially intelligent systems. In these systems, decision-making is a central process at all levels of human information processing. Problems are associated with the choice of decisions under conditions of incomplete information, arising when modeling the work of the human operator who perceives signals from the screens of indicators.

Intelligent technologies and AI systems to support decision-making

An intelligent decision support system (IDSS) is a decision support system that makes extensive use of artificial intelligence (AI) techniques. In general, intelligence is the ability to think, understand, and make decisions instead of doing something instinctively or automatically. The main ideas for creating artificial intelligence relate to the study of human thought processes, the representation and duplication of these processes using machines (computers, robots, etc.), and the study of behavior using a machine but performed by a human. The creation and development of artificial intelligence aim to make computers do things that humans now do to replicate some of the intelligent behaviors in a computer system.

The decision-making process today is complex, supported by computerized systems, and involves the following steps:

1. Problem definition. This is the basic stage. It provides decision-makers with a basis on which they can make assumptions, collect and analyze data, and evaluate alternatives. Problem definition begins with the recognition that a problem exists. A problem exists when

- There is a gap between what is expected and what is delivered;
- There is a deviation from the usual;
- The action taken is not justified;

The IDSS defines the problem and the complexities associated with the matched results.

2. Identification of the decision-maker. Depending on the nature of the problem, it is sent to the right person. A poorly structured problem will go to upper management, a difficult problem will go to managers, and repetitive problems will be sent to an employee at a lower hierarchical level.

3. Collection of information. Once the problem has been sent to the right person, the affected person can begin to collect data and identify the factors influencing the situation. Without DSS, data collection and analysis will take too long. DSS processes the data in a matter of seconds.

4. Evaluation of alternatives and decisions. This stage involves analyzing all possible courses of action and determining the most appropriate one by evaluating the pros and cons of each alternative. IDSS helps justify a particular choice.

5. Implementation and control. Once a decision has been made, it is necessary to move forward. Implementation requires planning. Monitoring is also important to determine the usefulness of a decision in achieving goals. This may require some adjustments or lead to a new problem. In the latter case, the entire process may need to be repeated.

The increased use of computer-based decision support systems is perceived as shifting the emphasis of decision-making to programs. The example in Fig. 2 of

such a system [8] consists of the following components: data management, model management, user interface management, and decision support system architecture.
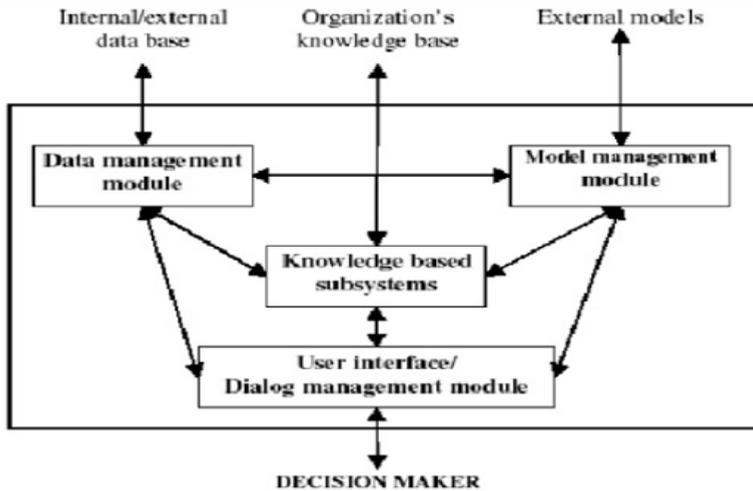

Fig. 2. Decision support system

The intellectual capabilities and behaviors integrated with the computer system create an intelligent machine. The machine must act as an assistant in decision-making, information search, the management of complex objects, and, finally, in understanding the meaning of words. To develop an intelligent computer system, it is necessary to gather, organize, and utilize human expertise and knowledge in various fields to improve the computing power of the system's brain with sophisticated algorithms using sensory processing, world modeling, behavior generation, value estimation, and global communication.

Mobile wireless networks have undergone a significant transformation in recent decades. At the same time, the generation of mobile wireless networks is usually determined based on such metrics as speed, technology, frequency, data volume, delay, user density, etc. Each generation has some features, standards, different capabilities, new methods, and new characteristics that make it different from the previous generation.

Wireless mobile technologies have evolved and improved significantly over the years as a result of intensive research and innovation. Now is the time when we can connect different wireless technologies, networks, and applications simultaneously through the latest technology, 5G (5th generation mobile communication technology). The new revolution in the mobile communication market is changing the use of mobile phones with very high bandwidth. This has transformed the network planning process from designing only for mobile devices to designing systems that connect different devices at high speeds [9].

138

Artificial intelligence and algorithms simplify the decision-making process. Using data processing systems has several undeniable advantages:

- Speed of decision-making: artificial intelligence can find optimal solutions in the enterprise much faster than a human;
- Minimizing errors and human influence: algorithms are not influenced by emotions and impulses when making decisions;
- Increasing the scope of the data analyzed: e.g., cross-analyses of geopolitical, economic, technical, etc. data;
- Expanding the possibilities and perspectives in problem-solving: applying new solutions and looking for different options.

In contemporary publications, various aspects of the intellectual activity of the human operator (HO) are considered, in particular expediency, the ability to acquire, complete, reproduce, and use knowledge, the ability to pose and solve problems, the ability to anticipate the unknown, the ability to generalize and make associations, etc. The intellectual activity of a person is associated with the search for solutions (actions and regularities) in new, non-standard situations when the solution is a priori unknown. In this case, the solution to the problem is understood as any activity (human or machine) related to the creation of plans and actions necessary to reach a certain goal, as well as the corresponding new conclusions and regularities.

Knowledge is the useful information accumulated by an individual, and intelligence is his ability to use this accumulated information in some useful and purposeful form. The intellectual (cognitive) functions of the living intellect are perception, intuition, creativity, association, induction (generalization), syllogism, prediction, planning, deduction, classification, and also search and selection, comparison, identification, and calculation. At present, the following functions have been analyzed and formalized in detail: search, choice, calculation, comparison, and deduction. Attempts to implement higher-level intellectual abilities on the computer have so far yielded no practical result. Thus, the full realization of intellectual capabilities related to decision-making, planning, forecasting, and effective management, as well as intellectual decision support systems, should be based on the use of the latest AI-based technologies and expert systems.

Nature and composition of the critical infrastructure monitoring system

By its nature, an CIMIS from a system-information perspective should be an event-driven CIS (ICS) with the provision of a continuous flow of sensor data processed and presented with AI technologies and tools as the CI that performs real-time information assurance and decision-making for users.

In organizational and technical terms, the CIMIS should be a network-centric system that includes a communication subsystem (a common transport network environment), a sensor area (sensors for conversion of physical parameters into data and a sensor network for access), and a technical management area including other management subsystems (Unmanned Aircraft Systems (UAS), network, and application management). The users of the network-centric system should be provided with a common communication environment with a high degree of intelligence and interactivity and with capabilities to access, update, and reuse the CI from other information subsystems [10, 11].

The CIMIS must be integrated and provide and present both information from its sensor networks so that, in the application layer, it is open and can include and integrate data from other segments and subsystems (e.g., space, specialized IS of government and public organizations).

Integrated mobile monitoring systems in Internet environment

Modern integrated mobile monitoring systems (IMMS) are CIS-inherently distributed computing resource systems that secure important sectors of the national economy, security, ecology, and people's business activities. The convergence of information objects and network structure in information-centric networks allows for increasing the speed of real-time information exchange in the composition of their information structure, including in their geographical boundaries and hard-to-reach areas. In addition to communication services with guaranteed availability and quality, the IMMC provides access to dynamic information in real-time with information object formation, data analysis, and presentation appropriate for the system, which gives grounds for its analysis as a CIS.

Mobility nodes for forming, collecting, and aggregating information from fixed or mobile sensors deployed on unmanned aerial vehicles (UAVs) are aggregated into CI with information integrated from other information systems, including satellite systems. The latter, from a cybernetic point of view, are large (complex) systems and, in contrast to UAVs, are characterized by specific properties such as globality of use, multi-purpose, and multi-functionality. Aerospace technology is an important modern preventive factor ensuring timely detection, identification, and scoping of various crisis processes. Space- and air-based technologies overcome a significant portion of the shortcomings and limitations of traditional ground-based technologies, mainly related to monitoring, navigation, and communication.

Modern IMMS for monitoring have an information and communication resource, providing a wide class of tasks for events, the solution of which exceeds the capacity of the system to serve the usual activities, ensuring business activity and protecting society. The observation and monitoring based on such technical solutions are necessary when it is necessary to supply and analyze data from

urgently installed sensors in connection with a specific situation, such as fires and their containment in remote areas; chemical pollution with high dynamics and danger to people's lives and health; in the event of an urgent inspection and expert assessment of damage in hard-to-reach objects; finding and rescuing people; monitoring for telemedicine purposes; and many others. Integration into the monitoring system of satellite information and data from other information systems provides additional opportunities for forecasting and analysis of the monitoring data obtained as a result and their processing to achieve the system's goals. In addition, the space monitoring system is a complex system with a space and ground segment, collecting, processing, archiving, and providing users with information about objects, phenomena, and processes on the earth's surface.

Trends in the formation of a global information infrastructure in modern societies and the development of the needs of citizens, businesses, and society imply the simultaneous use of monitoring systems, modern mobile data transmission technologies (4G/5G), and export opportunities of information on the Internet with appropriate formatting and presentation.
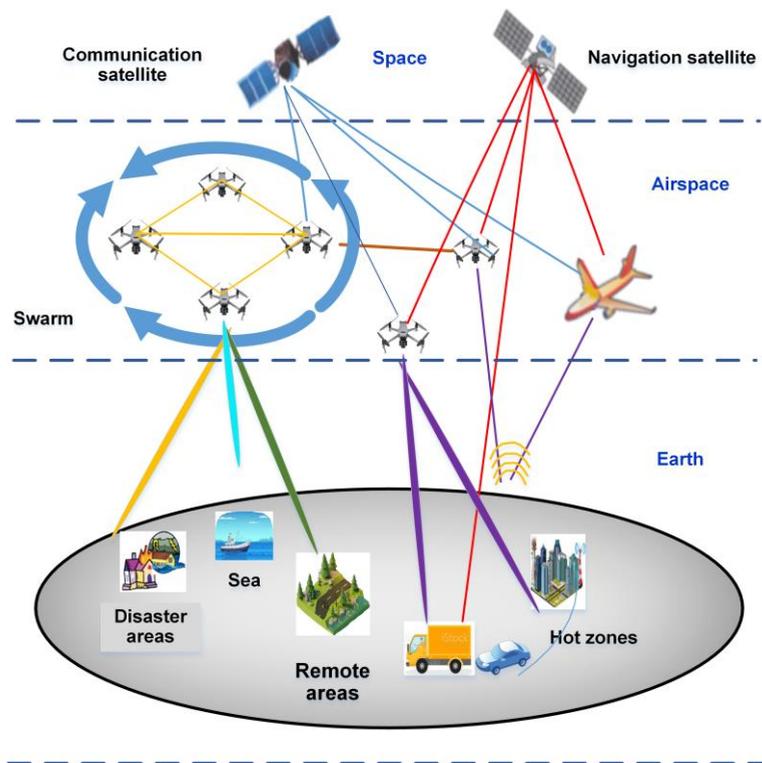

Fig. 3. Integrated monitoring system

The adoption of Internet technologies (the TCP/IP architecture as a network model and protocol stack) was justified as the basis for building a network-centric CIMIS. Its structure in generalized form consisted of functionally complete information-communication clusters (ICCs). The communication architecture of the backbone network should ensure the operation of mobile and stationary information clusters, integration with other ISs (space segment, national networks related to CI protection), and the presentation of results on the Internet. The requirements to be network-centric require the use of virtual private network technologies and transport over the Internet. For mobile ICS, transport must be over Internet protocol over public mobile data networks (5G, 4G, 3G, and 2G). In geographic areas where there is no coverage by national mobile operators, satellite Internet can be used, or Internet protocol transport can be built over a radio channel. A necessary requirement for all ICs is that they are connected to the IP backbone and operate in a network-centric environment.
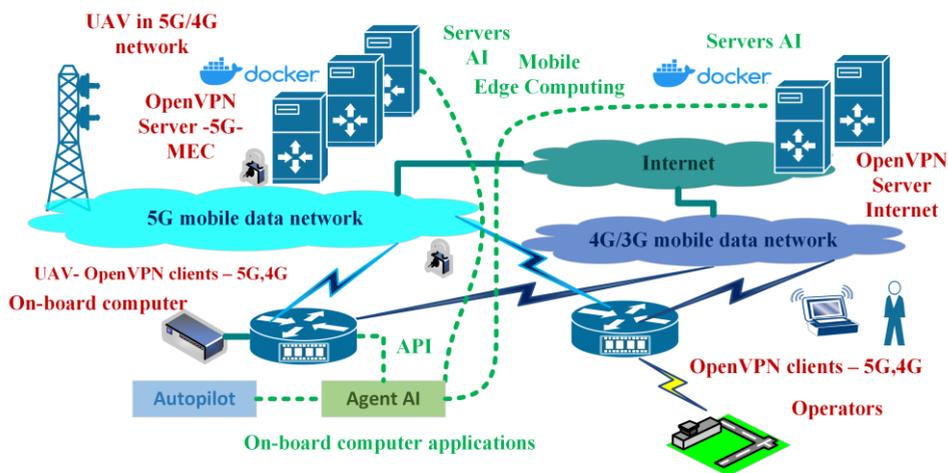

Fig. 4. Clusters of CIMIS in a virtual private network [5]

The server part of the IC (the application layer in the network model) can also be viewed as an ICC or a group of clusters that are located in the network address space [5].

OpenVPN technology is chosen as the backbone of the virtual private network. The advantages and features provided by the technology will be described in detail for a mobile ICC of UAS and connectivity to a processing and presentation cluster. The application of OpenVPN technology provides the integration of the ICC into a common environment, scalability, manageability, redundancy, migration, reliability, and evolution of the backbone network. On this

basis, flows of information objects (IO) can be transported and a functional ICC structure can be built.

The network-centric environment performs its protection and monitoring functions in four main logical network architectures: centralized, on-demand (ad hoc), swarm, and combined.

Centralized network-centric architecture

In the centralized architecture, a single OpenVPN server is used to provide network access to all clusters through routers running as OpenVPN clients. The availability and survivability of the configuration are related to having a single vulnerable point in the structure at all levels of the network model. The advantages of the centralized architecture are in terms of synergy effects in terms of management and coordination of the entities and sites in the CIMIS. It provides coordination of the work with the different ICCs and efficiency greater than the efficiency of the clusters individually. A significant disadvantage is the complexity of maintaining, administering, and coordinating a large number of nodes in the network, for example, from the CI control center level down to the UAV or gateway level in a sensor radio network. The information flows in the CIMIS have to be pre-formed as IO-type flows, and their direction is not always directly to the logical center of the network, which is fixed in a centralized architecture. For CIMIS, a centralized network-centric architecture can be provided by OpenVPN technology, but it is only applicable to relatively small systems.

Swarm architecture

The chosen technology for building a virtual private network, OpenVPN, enables the most complex and most promising architecture in a network-centric environment: swarm architecture. A swarm is a combination of homogeneous elements with limited and fixed resources of their own that must exchange information with each other for the purposes of self-synchronization, self-organization, efficiency gains, and achieving multiplicative effects. A swarm can be considered a group of UAVs, one of which has the role of an information hub, an ICC of mobile monitoring and control stations, united and linked to one of them, which has a satellite Internet connection and provides the network access of the others and the coordination between them. There are different types of network-centric swarm architectures: managed, hierarchical, distributed, and mixed.

The chosen virtual private network technology, OpenVPN, with centralized routing management for full connectivity (everyone to everyone), independence from a fixed port, and the possibility of simultaneous participation of a network node as a server or client, allows the implementation of an arbitrary logical structure in a network-centric environment.

Conclusions

Modern integrated mobile monitoring systems are inherently communication and information systems (CIS) with distributed computing resources, which provide important branches of the national economy, security, ecology, and business activity of people. The convergence of information objects and network structure in information-centered networks allows for increased speed of real-time information exchange in the composition of their information structure, including in their geographical boundaries and hard-to-reach areas. In addition to communication services with guaranteed availability and quality, the monitoring system provides access to dynamic information in real-time with the formation of information objects, data analysis, and presentation appropriate for the purposes of the system, which gives grounds for its analysis as an information and communication system.

The integrated system incorporates both real-time geographic data from sensors in hard-to-reach areas and mobile asset-based sensors (UAS), as well as data for forecasting or overcoming disasters and critical situations obtained through electro-optical, infrared, and radar sensors on surveillance satellites. Solving problems related to the integration of data from different systems, management of information flows, presentation on the Internet, export to other systems, and provision of information security is relevant and significant in modern monitoring systems. The combination of local and cloud technologies, tailored to the capabilities of mobile data networks (4G/5G), and the optimization with regard to the required speed performance are contemporary issues in the context of the development of modern automated systems for the transport of people and goods.

The pace and direction of modern scientific and technical progress require new state-of-the-art ideas on the methodology of building integrated systems for forecasting, containment, and eradication of environmental incidents, search for people, and rescue of human life. They must also be compatible with new generations of information and communication systems designed to ensure user mobility and innovation in the use of modern technologies. This is one of the main objectives of this project.

References

1. Димитров, Н., Системен подход към критичната инфраструктура, ВВМУ "Н. Й. Вапцаров", Варна, 2019.
2. Додонов, А., Д. Ландэ, Живучесть информационных систем. К.: Наук. думка, 2011.
3. Сотиров, Г., Е. Хубенов, З. Чифлиджанова, Агрегиране на услуги, базирани на изкуствен интелект, в интегрирани мобилни системи за мониторинг в интернет среда, Eighteenth International Scientific Conference, SES2022, PROCEEDINGS SES`22, pp. 127–133, Sofia. http://space.bas.bg/SES/archive/SES%202022_DOKLADI/2_Aerospace%20Technologies/7_Sotirov.pdf
4. The Future Internet, Editor: Domingue J., Galis A., and ect., 2011, Publisher: Springer Berlin Heidelberg, Lecture Notes in Computer Science, DOI:10.1007/978-3-642-20898 -0   , https://www.researchgate.net/publication/230634079_The_Future_Internet
5. Хубенов, Е., Г. Сотиров, З. Хубенова, Безпилотни авиационни системи с елементи на изкуствен интелект в интернет среда, SES`2021, Seventeenth International Scientific Conference „Space, Ecology, Safety", PROCEEDINGS SES`21, pp. 121–127, Sofia. http://space.bas.bg/SES/archive/SES%202021_DOKLADI/PROCEEDINGS%20SES%202021.pd
6. Мардиросян, Г., Природни бедствия и екологични катастрофи АИ "Проф. Марин Дринов", 2009 г.
7. ДИРЕКТИВА 2008/114/ЕО НА СЪВЕТА, относно установяването и означаването на европейски критични инфраструктури и оценката на необходимостта от подобряване на тяхната защита, https://eur-lex.europa.eu/legal-content/BG/TXT/PDF/?uri=CELEX:32008L0114&from=HR
8. Scott Morton M. S. Management Decision Systems: Computer-based Support for Decision Making.—Boston: Harvard University, 1971.
9. Munir Md. S., S. F. Abedin, C. S. Hong, Artificial Intelligence-based Service Aggregation for Mobile-Agent in Edge Computing, Conference: The 20th Asia-Pacific Network Operations and Management Symposium, Japan, 2019.
10. Department of Defense Net-Centric Services Strategy Strategy for a Net-Centric, Service Oriented DoD Enterprise March 2007 Prepared by the DoD CIO
11. Net-Centric Environment Joint Functional Concept Department of Defence USA, version 1.0

ИНТЕЛИГЕНТНА СИСТЕМА ЗА МОНИТОРИНГ И ЗАЩИТА НА КРИТИЧНАТА ИНФРАСТРУКТУРА НА БАЗА МОБИЛНА КОМУНИКАЦИОННА-ИНФОРМАЦИОННА СИСТЕМА С ЕЛЕМЕНТИ НА ИЗКУСТВЕН ИНТЕЛЕКТ

Е. Хубенов, З. Чифлиджанова

Резюме

В статията се предлагат съвременни иновационни решения за интегрирана мобилна система за мониторинг с елементи на изкуствен интелект в Интернет среда на база мобилна комуникационна-информационна система за събиране, агрегиране, обработка и презентиране в реално време на потоци от информационни обекти. Представено е управлението на комуникационни мрежи с цел оптимизация на транспорта и обработката на информационни обекти на основа на дискретно-събитийно представяне на потока от данни и моделиране с йерархична структура. Разгледани са качествени аспекти както за мрежи, предназначени за мониторинг и известяване на геофизически, климатични и други природни феномени, така и за антропогенни системи. Направено е сравнение на възможностите на централизираното и децентрализирано управление, както и състоянието на съвременните мрежови технологии и възможностите за практическа реализация в различни мрежови архитектури.

PERFORMANCE ANALYSIS OF CURVED SHAPE ON THE INLET GUIDE VANES IN CENTRIFUGAL BLOWERS

Teodor Stănescu[1,2], Daniel Uşeriu[1,3]

[1]National Research and Development Institute for Gas Turbines COMOTI, Romania
[2]Faculty of Energy Engineering, University Politehnica of Bucharest, Romania
[3]Faculty of Aerospace Engineering, University Politehnica of Bucharest, Romania
e-mail: daniel.useriu@comoti.ro

Abstract

This study aims to analyze the performance of two types of stator blades used in centrifugal blowers: forward-curved blades and backward-curved blades (concave and convex shapes), under various operating conditions and tilt angles. The primary goal of the research is to evaluate compression performance by comparing the results obtained for these two blade types. Ultimately, the objective of the study is to provide clearer insights for choosing the correct shape of the inlet guide vane in the specific field of application for centrifugal blowers.

Introduction

The development and improvement of centrifugal blowers hold significant importance in the realm of mechanical engineering. Extensive research has been conducted in this domain, exploring various facets such as impeller aerodynamic performance analysis [1], volute analysis [2], stator analysis [3], optimization of rotor and stator geometry [4], assessment of operational conditions on efficiency, enhancement of acoustic characteristics, vibration reduction, and more. For instance, C. Ji et al. [5] demonstrated that by accurately defining the flow channel, it is possible to eliminate the impact of tip clearance on Inlet Guide Vane (IGV) blades, thereby expanding the control range of mass flow rate.

Optimizing the profile selection of IGV contributes to enhancing the performance of compressors and an efficient flow control. In the same vein, H. Cao et al. [6] examined the efficiencies achieved by three different airflow control systems in a centrifugal fan. Furthermore, D. A. Nguyen et al. [7] concluded that the direction of the inlet flow into the impeller influences the performance of an axial pump, and adjusting the IGV blade angle can improve energy efficiency. The optimal positioning of IGV blades relative to the centrifugal impeller is a critical

requirement for designers of centrifugal compressors, directly impacting efficiency [8]. Additionally, excessive clearances within turbomachinery result in diminished performance when chosen with excessively high values [9].

To enhance the compression process, J. Xin et al. [10] investigated the effect of introducing slots in the aerodynamic blades at various placement angles. Their study aimed to modify the geometry of the Inlet Guide Vane (IGV) blade, with the goal of improving overall compression efficiency.

CFD analyses have gained significant momentum in recent years due to their high accuracy in reproducing fluid behavior, particularly within turbomachinery [11], [12]. Through these simulations, S. Rabet et al. [13] highlighted the fluid behavior in a centrifugal pump when changing the number of blades on the impeller. Additionally, the selection of the optimal number of blades on the IGV plays a crucial role in centrifugal compressors [14]. According to the conclusions brought by O. Dumitrescu et al. [15], it has been demonstrated that the size of the discretization grid and the choice of turbulence model have a significant impact on the performance resulting from CFD numerical simulations. Through CFD analyses, J. Fang et al. [16] demonstrate the potential for improving the isentropic efficiency of a centrifugal blower by modifying the blade curvature.

Description

This article focuses on the performance analysis of curved-shaped inlet guide vanes in centrifugal blowers, involving numerical simulations with two types of stator blades (Type I - concave shape and Type II - convex shape) used at three different tilt angles (-15 degrees, 0 degrees, and +30 degrees). The setup employed for this study is illustrated in Fig. 1.
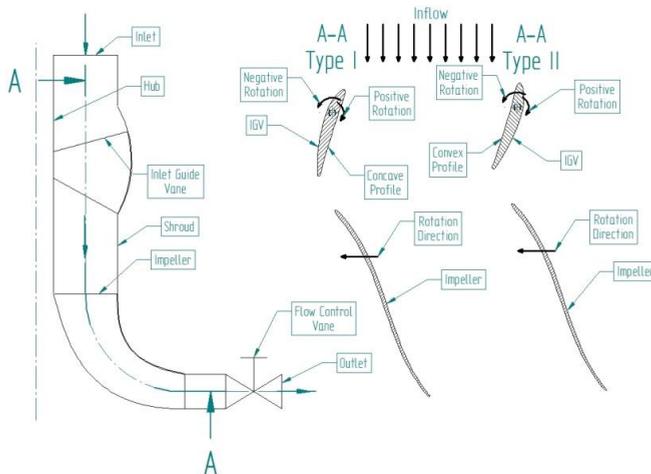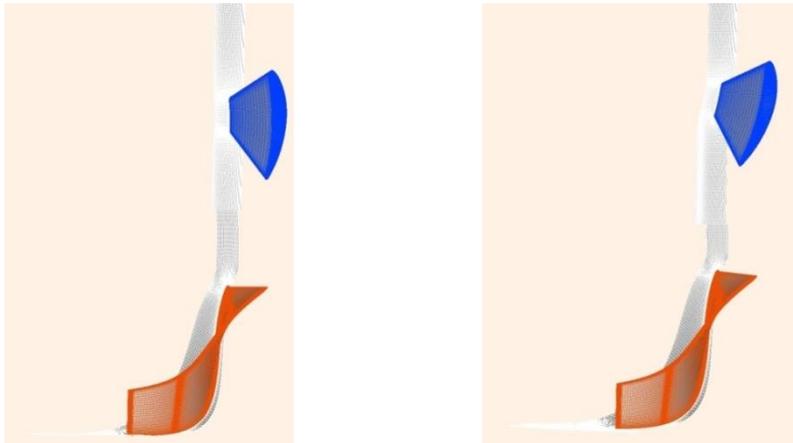

Fig. 1. The schematic representation for comparing those two types of IGV configurations

148

The geometry for the impeller and flow channel was created using the Ansys CCD software, while the geometry of the IGV blade was defined using the NACA 6412 airfoil profile.

a) IGV type II, tilted at 0 degrees   *b) IGV type II, tilted at -15 degrees*

Fig. 2. The resulting mesh for a single flow channel at a different angle of the IGV

The structured mesh for a single flow channel (Fig. 2) was generated using the Numeca Autogrid 5 software. For this purpose, two different IGV shapes were defined, each positioned at three angles (-15, 0 and +30 degrees). As a result, the mesh consists of two subdomains: the impeller domain, comprising a total of ~975k cells (Fig. 4) and the stator domain, consisting of ~900k cells (Fig. 3), which vary slightly depending on the blade placement angle within the grid.
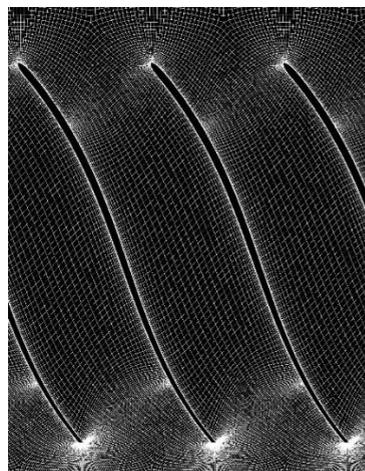

Fig. 3. B2B mesh for stator domain   *Fig. 4. B2B mesh for impeller domain*

The boundary conditions employed in the simulations, for a steady-state case, are outlined in the table provided below (Table 1):

Table 1. CFD Simulations: Boundary and Initial Conditions Overview

Inlet (IGV) Total pressure	101 325 Pa
Inlet (IGV) Total temperature	273.15 K
Outlet (impeller) Mass flow rate	1.7 kg/s* at nominal point
Rotational speed (impeller)	20 800 rpm
Turbulence model	SST k-omega
Number of blades – IGV / impeller	8 / 11 blades

*To generate the characteristic curves, the imposed flow rate at the impeller outlet varied from the surge to the choke point.

In order to validate the grid refinement, an analysis was conducted on the $y+$ term, which describes the accuracy of the results obtained when using the turbulence model in the current study. The resulting $y+$ value from the simulations was a maximum of 2.3 (Fig. 5), indicating that the influence of the mesh on the results can be considered negligible.
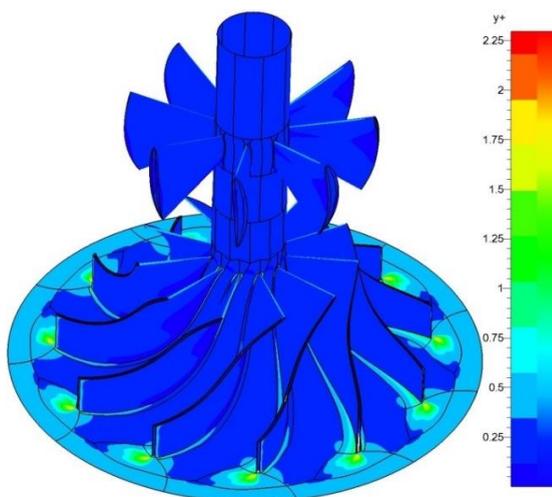

Fig. 5. The y+ values reported for the hub and blades

Numerical results

Due to the 0° position imposed on the IGV blades, they are arranged to guide the fluid in order to achieve a 90° angle for the α_1 term (the absolute angle at the inlet) at the entry into the impeller domain. It can be observed that both, the concave and convex profiles, exhibit the same characteristic curves, indicating that the obtained values (pressure rise or efficiency) are close or identical. The maximum differences recorded are at most 0.31% in terms of pressure increase and 0.24% in terms of efficiency (according to Fig. 6). These errors are primarily caused by small angular deviations in the positioning of the stator blades.

150
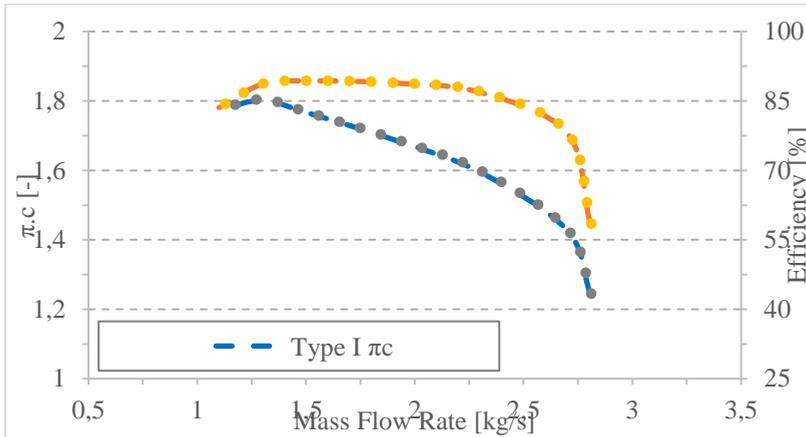
Fig. 6. Pressure ratio and Efficiency vs Mass Flow Rate with IGV Tilt Angle at 0°

Through simulations, the streamlines and aerodynamic forces developed along the profiles have been obtained, with the only visible differences resulting from the reversal of the two surfaces of the profile, namely the pressure side and suction side (Fig. 7).

a) type I (concave shape)   *b) type II (convex shape)*
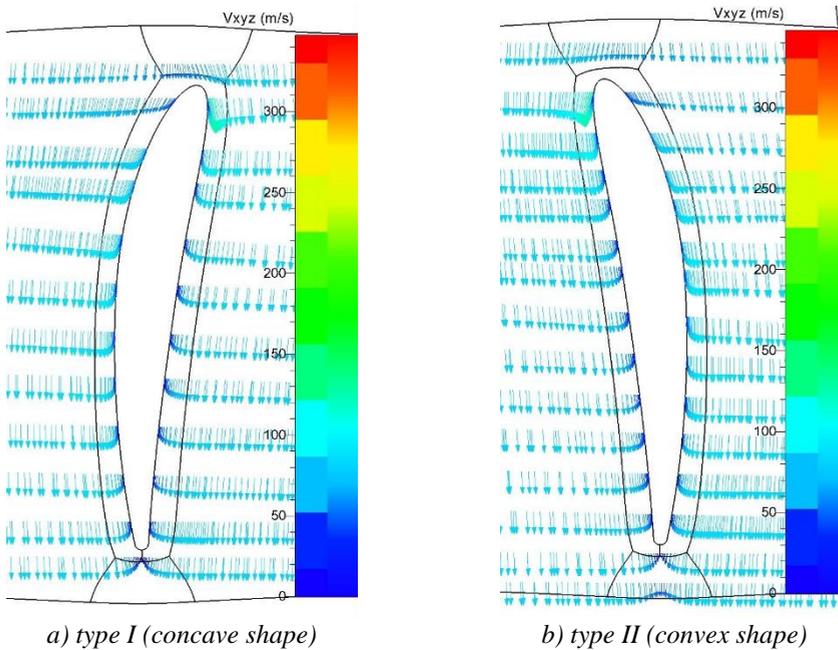
Fig. 7. Velocity vectors of the IGV blades positioned at 0°

In contrast to the previous case, as a result of simulations conducted for blade-tilted angles of -15 degrees (Fig. 8) or +30 degrees (Fig. 9), the characteristic curves no longer overlapped. It is observed that the concave shape enhances pressure rise and performance at negative rotation angles of the IGV blades, while the convex shape of the blades (Type II) is favorable at positive rotation angles. Moreover, it is noted that the shapes of the curves exhibit similarities: "Type I π_c" at -15 degrees with "Type II π_c" at +30 degrees; "Type I Efficiency" at -15 degrees with "Type II Efficiency" at +30 degrees, and so forth.
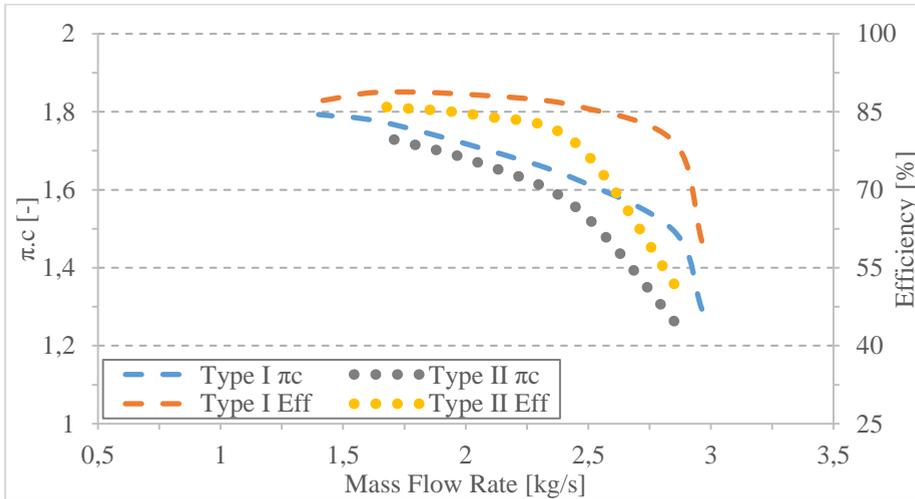
Fig. 8. Pressure ratio and Efficiency vs Mass Flow Rate with IGV Tilt Angle at -15°
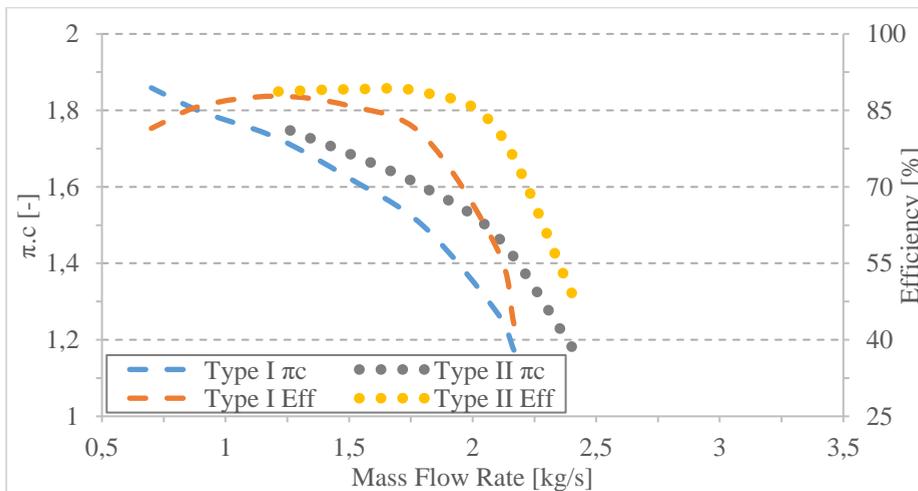
Fig. 9. Pressure ratio and Efficiency vs Mass Flow Rate with IGV Tilt Angle at +30°

a) type I (concave shape) *b) type II (convex shape)*
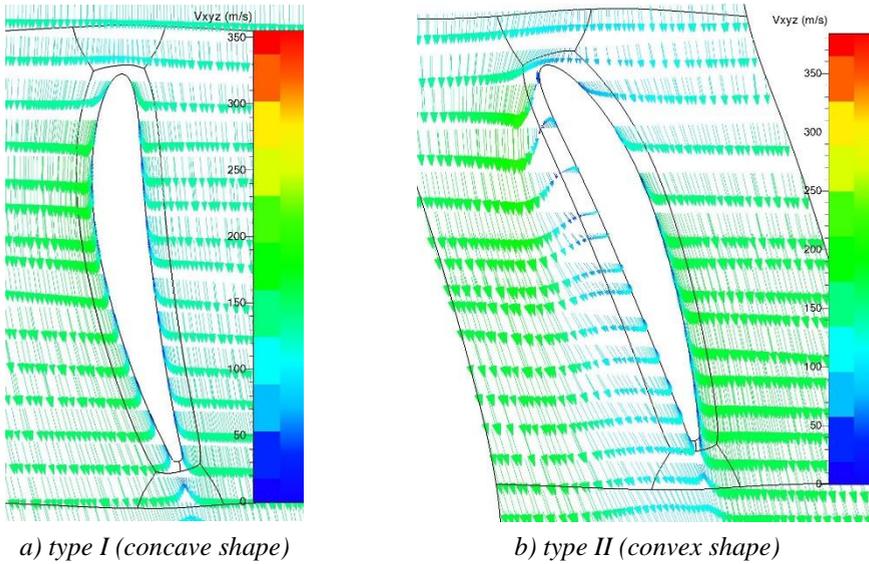
Fig. 10. Velocity vectors of the IGV blades positioned at -15°

The high performance achieved by the concave shape at negative incidences of the IGV blades is attributed to the smooth flow along the aerodynamic profile (Fig. 10 a.). In the case of the convex shape (type II), separation or detachment is observed on the pressure side (Fig. 10 b.), resulting in penalties of 8.4% in terms of efficiency and 4.4% in terms of pressure rise at a flow rate of 2.45 kg/s.

a) type I (concave shape) *b) type II (convex shape)*
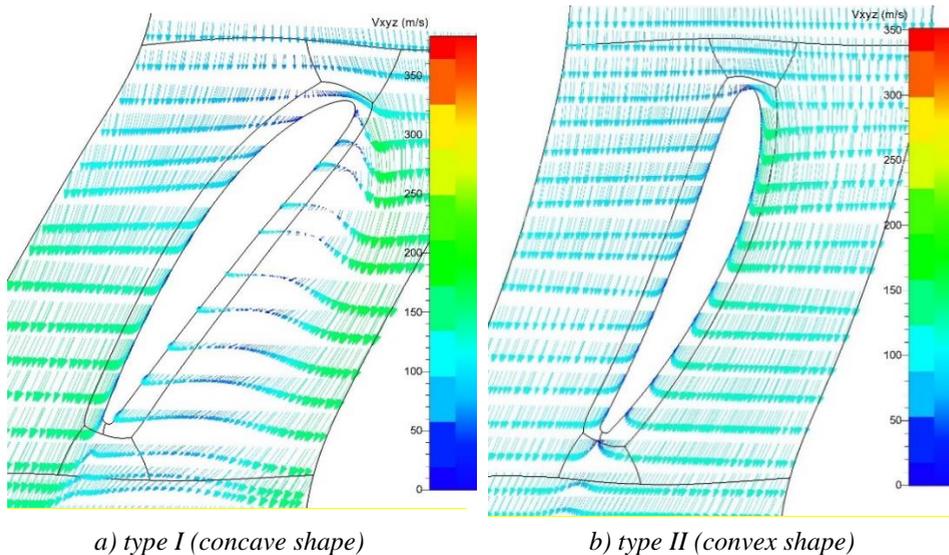
Fig. 11. Velocity vectors of the IGV blades positioned at +30°

Similar to the previous case, when the stator blades with a concave shape are rotated in a positive direction, detachments are observed on the pressure side (Fig. 11 a.), resulting in 10.6% lower efficiencies and 6.8% lower pressure rises compared to the convex shape at a flow rate of 1.8 kg/s.

Conclusions

In this studied case, a 27% efficiency difference was observed between these two blade shapes (concave and convex) at an IGV blade inclination of -15 degrees and a 23% difference at an IGV blade inclination of +30 degrees. Both points where efficiency values showed significant differences were near the choke point, where the flow rate was high, and the influence of separations/detachments on the IGV aerodynamic profile resulted in significant losses.

On the one hand, the concave shape (Type I) is favorable for higher flow rates and pressures compared to the nominal point or for negative inclinations of the IGV blades, while the convex shape (Type II) is more favorable for lower flow rates and pressures compared to the nominal point or for positive inclinations. On the other hand, for a turbomachinery operating under variable conditions (with the flow rate varying above or below the nominal point), the radial shape is the most advantageous, achieving a compromise between these two shapes.

Additionally, the use of the morphing blade can lead to higher efficiency compared to the radial-shaped blade. However, it is important to note that the manufacturing and control of the morphing blade are much more complex.

Acknowledgment

References

1. Tiainen, J., A. Jaatinen-Värri, A. Grönman et al. (2016), Numerical Study of the Reynolds Number Effect on the Centrifugal Compressor Performance and Losses.
2. Dumitrescu, O., G. Fetea, G. B. Gherman (2017), Influence of the volute design on performances of a centrifugal compressor.
3. Poujol, N., I. Trebinjac, P. Duquesne (2021), "Effects of Inlet Guide Vanes on the Performance and Stability of an Aeronautical Centrifugal Compressor," J. Turbomach., vol. 143, pp. 1–12, doi: 10.1115/1.4050944.
4. Karabanova, V., A. Vanyashov, V. Yusha (2020), "The influence of blade profile shape of inlet guide vane on performance map of centrifugal compressor stage," Omsk Sci. Bull. Ser. Aviat. Power Eng., vol. 4, pp. 41–48, doi: 10.25206/2588-0373-2020-4-4-41-48.

5. Ji, C., Q. Sun, Z. Fan et al. (2016), "Study of High Efficiency Flow Regulation of VIGV in Centrifugal Compressor," Int. J. Rotating Mach., vol. 2016, p. 6097823, doi: 10.1155/2016/6097823.

6. Cao, H., Z. Zhou, J. Zhang (2018), "Aerodynamic design schemes of the inlet guide vane in a core-driven fan stage," Adv. Mech. Eng., vol. 10, p. 168781401774770, doi: 10.1177/1687814017747707.

7. Nguyen, D.-A., S.-B. Ma, S. Kim et al. (2023), "Influence of inflow directions and setting angle of inlet guide vane on hydraulic performance of an axial-flow pump," Sci. Rep., vol. 13, doi: 10.1038/s41598-023-30511-4.

8. Gherman, G. B., M. Gall, V. Popa et al. (2018), "IGV position optimization for centrifugal blower," IOP Conf. Ser. Mater. Sci. Eng., vol. 400, p. 42025, doi: 10.1088/1757-899X/400/4/042025.

9. Dragan, V., G. B. Gherman, I. Malael et al. (2015), "Global Analysis of Tip Clearance Influence on Centrifugal Compressor Performance," Appl. Mech. Mater., vol. 811, pp. 128–132, doi: 10.4028/www.scientific.net/AMM.811.128.

10. Xin, J., X. Liu, L. Zhao et al. (2022), "Impact of Slots on the Aerodynamic Performance of the Variable Inlet Guide Vane Cascade of a Centrifugal Compressor," J. Appl. Fluid Mech., vol. 16, no. 2, pp. 245–256, doi: 10.47176/jafm.16.02.1381.

11. Li, D., C. Yang, M. Zhou et al. (2011), "Numerical and experimental research on different inlet configurations of high speed centrifugal compressor," Sci. China Technol. Sci., vol. 55, doi: 10.1007/s11431-011-4635-2.

12. Barhatte, S., M. Chopade, A. Walimbe (2021), "Performance Enhancement of a Centrifugal Blower using Experimental and Computational Techniques," p. 1.

13. Rabet, S., S. Abdollahi, F. Ranjbar (2023), CFD Simulation of Flow in a Centrifugal Pump.

14. Anbarsooz, M., M. Amiri, A. Erfanian et al. (2021), "Effects of number of inlet guide vanes on the aerodynamic performance of a centrifugal compressor," Proc. Inst. Mech. Eng. Part A J. Power Energy, vol. 236, p. 095765092110268, doi: 10.1177/09576509211026824.

15. Dumitrescu, O., V. Dragan, I. Porumbel et al. (2020), "Numerical assessment of a very high pressure ratio centrifugal impeller," IOP Conf. Ser. Mater. Sci. Eng., vol. 916, p. 12035, doi: 10.1088/1757-899X/916/1/012035.

16. Fang, J., P. Yan, J. Qin et al. (2023), "Research on Centrifugal Blower Design and Optimization Method Based on Mean Camber Line Optimization and CFD," J. Phys. Conf. Ser., vol. 2419, p. 12062, doi: 10.1088/1742-6596/2419/1/012062.

АНАЛИЗ НА РАБОТАТА НА ЦЕНТРОБЕЖНИ КОМПРЕСОРИ В ЗАВИСИМОСТ ОТ ЛОПАТКИТЕ С ИЗВИТА ФОРМА НА ВХОДНИЯ НАПРАВЛЯВАЩ АПАРАТ

Т. Стънеску, Д. Ушериу

Резюме

Това проучване има за цел да анализира ефективността на два типа статорни лопатки, използвани в направляващия апарат на центробежни компресори: лопатки с предно извит профил и лопатки със задно извита форма на профила (вдлъбнати и изпъкнали форми), при различни работни условия и ъгли на завъртане на лопатките. Основната цел на изследването е да се оцени производителността на компресора чрез сравняване на резултатите, получени за тези два типа лопатки. В крайна сметка, целта на проучването е да предостави по-ясна представа за избора на правилната форма на лопатката на входния направляващ апарат в специфичната област на приложение на центробежните вентилатори.

REUSABLE AND LOW-COST SPACE ROCKET ENGINE WITH HIGH-EFFICIENT PROPELLANT PUMP

Abbasali Saboktakin[1], Mohammad Monjezi

[1]Ass. Professor, Department of Aerospace Engineering, İzmir Economic University, Turkey
e-mail: abbasali.saboktakin@ieu.edu.tr, Alaptakin@gmail.com

Keywords: Engine, Rocket, Piston Pump, Propellant, Finite element modeling.

Abstract

Space propulsion mainly uses reciprocating pumps that are fully integrated with small-scale applications. In this paper, the specific specifications and designing factors that should be met by rocket engine fuel pumps are demonstrated, and a comparative study is formed of the suitableness of all the necessary kinds of pumps to be used with rocket engines and their applications. Furthermore, it examines a piston pump intended to refuel the liquid fuel nozzles. Due to the performance of this pump, different parts and its performance have been evaluated. The design of the pump is discussed such that after completing the design steps according to the pump performance, a numerical solution of the stress and strain applied on the inner wall of the cylinder is performed according to its performance under internal pressure. Then, considering the consistency of the body material and the same pressure and heat applied to the cylinder and piston, a sample of the cylinder and piston is analyzed in finite element modeling technique and the modeling results obtained from the simulation are presented.

Introduction

The development efforts of the main engine for the space vehicle were initiated in 1971 by NASA. Rockwell's Rocket dyne division as the prime contractor with NASA, after several years of development and testing, three space shuttles delivered to the space transportation system [1]. All space vehicles included high performance, perfect thrust, and high reliability and reusability systems. The simplest manner to differentiate rocket engines is to categorize them based on their technique of propellant pressurization and delivery [2]. All rocket engines can generally be divided into two categories: pressure-fed and pump-fed. While small-pressure engines use pressurized tanks for propellant delivery, the majority of rocket engines use turbo pumps that allow the propellants to be delivered to the desired pressure level. The pressure-fed engine is a self-pressurization such that self-pressurization is usually carried out through mono-propellant rocket engines and is obtained through the thermal decomposition of the liquid propellants or its vaporization. Pressure-fed system engines usually use high-pressure helium bottles.

In any case, the thrust stage of pressure-fed engines is confined through the tank technology [3]. An instance of these engines is the ARIANE 5G and AESTUS engines. Pump-fed engines use a turbo pump to increase the propellant pressure. In fact, some of the propellants are fed into a gas generator, which typically works at a necessarily high-pressure level. Modern liquid rocket engines have required pumping systems to transfer the propellant to the rocket engine [4, 5]. These pumps decrease the mass and size of other hardware by using lightweight high-pressure thrust chambers while decreasing the pressure of the liquid tank and minimizing the storage of inert gas. Figure 1 shows typical rocket engines that use high-performance pumping systems.
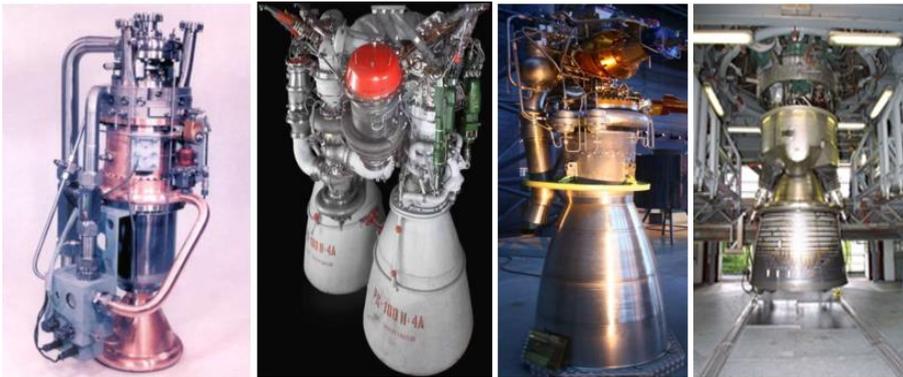

Fig. 1. Modern liquid rocket engines with high reliability and reusability. From left to right: P111 engine, RD-180, Viking engine, Vulcan 2 engine [2, 3].

The diverse use of piston pumps for different liquid rocket propulsion systems may be found as shown in Fig. 2. The most adaptable usage of a piston pump is on a satellite that has to make multiple and massive ΔV maneuvers. Any polar orbit satellite and geostationary satellite could have enormous ΔV maneuver requirements [6]. Since polar satellites and geostationary satellites are commonly high-priced and need to perform for several years, they constitute likelihood applicants for performance improvements [7]. However, it must be referred to that the performance supplied using a solid booster could, in a few cases, exceeds the overall performance of a bipropellant liquid rocket. The simplest improvement is that the piston pump can offer on/off functionality and throttling as well [8, 9]. Other applications are probes to the moon, interplanetary probes, and near-earth objects that commonly require enormous ΔV maneuvers. These specific missions have mass budgets and may benefit from any weight reduction and financial savings while overall performance stays high. Finally, missions to or from the other planets surface, moons will also benefit from overall performance enhancements.
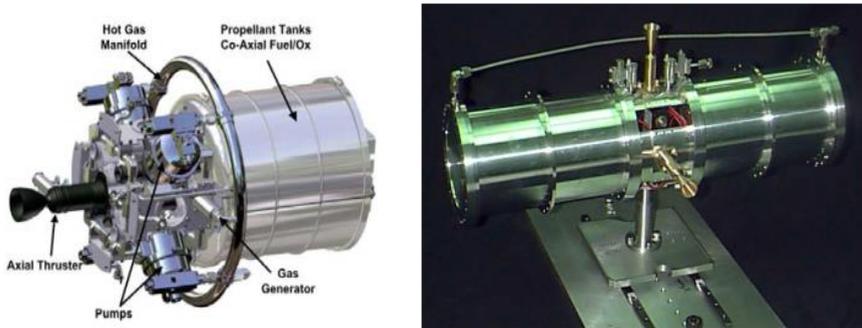
Fig. 2. Pump-fed propulsion system in satellite [8]

The design of the pump, as well as the representation of the fluid, inside the pump, is very difficult, time consuming and costly [10–13]. Compact Fluid Analysis (CFD) is the best available tool for analyzing flow patterns inside the piston pump and gas generator to predict their behavior under different operating conditions. It also helps optimize pump design parameters by providing the most correct flow patterns along with more efficient pump operation [14,15].

The objective of this paper is to survey the status of research in the areas of pumps for liquid rocket engines in order to provide a comprehensive review of the state of the art and understanding of important challenges. In this article, the emphasis will be on a reciprocating pump and the pump pistons are stimulated by gas. We use a gas generator to convert the liquid fuel of the peroxide to the gas with the pressure and temperature. We analyzed the pump and also the generator in the finite element modeling software, and the results are expressed to give complete awareness of how to use it in different situations. The research will concentrate on the design and development of a 300-gram pump capable of delivering fuel and oxidizer at 5 MPa for a 1000-N engine with remaining tank pressure at 0.35 MPa.

Pump design

The pump contains four cylinders and pistons that are joined together, as shown in Fig. 3. It has a central section that provides the pump with a liquid inlet from the propellant tank and directions reflect the flows of the inlet and outlet depicted in Fig. 3a. The liquid propellant's inlet port is the large port at the middle of Fig. 3b, and a separate outlet hole is on the opposite side of the piston pump that is evident in Fig. 3a. The gas is distributed and operated by valves to the outer cylinders. Liquid cylinders are smaller in diameter than gas cylinders. The area ratio allows the reciprocating pump to be operated by means of delivered propellants. At the end of every cylinder, the gas cylinder is larger in diameter than the liquid cylinder and the gas entry point. A piston separating the fuel and gas chambers is

located among them. As seen in the pump schematic, it is understood that no shaft or other rotating components are required to apply gas pressure to the rocket liquid. The gas inlet valves are retained to cancel the mass effect of movement, so that the opposite movement of pistons is towards or away from each other. Furthermore, because of the existing pistons in pump chambers, the control scheme is considered to compensate for the pressure loss. According to Fig. 3a, cylinder numbers 1 and 3 reach the end of the stroke, while cylinders 2 and 4 have been refilled with the propellant. A control mechanism provides the continuation of the flow such that cylinders 2 and 4 are pressurized with gas before reaching pistons 1 and 3 to their limit and venting their cylinders.

a)   *b)*   *c)*
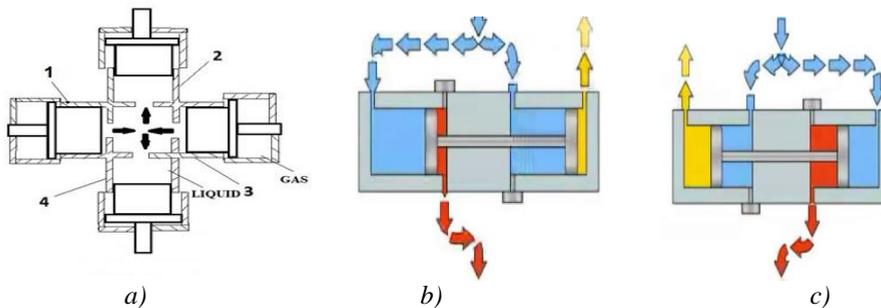
Fig. 3. a) Primary schematics of the piston pump; b)&c) Preparation of fuel injection from propellant tank to engine chamber occurring within the reciprocating pump.

This pump weighs 470 grams, and the blocks and cylinders are all made of aluminum, with the exception of the valves and tubes. The role of curved pipes is to send gas to the inlet valves shown in Figure 4 are used in the design of light alloy metals. The largest hole located in the centre of the block is the fuel entry site, and the exit point of the four small holes is situated in the middle of the pump and on the other side. Each cylinder has a perfect displacement of 8 cc between the piston stops, or 32 cc per pump cycle. The valve opening time for fuel input is also less than 10ms, and this is independent of the piston speed. In order to provide the required force to drive the pistons, a key aspect previously mentioned is that the diameter of the gas cylinder is greater than that of the liquid cylinder. The gas and liquid cylinders are not combined but are connected by screws. This pump works at a far higher pressure than those pumps used in conventional launchers.

Fig. 4. The primary design of the reciprocating pump

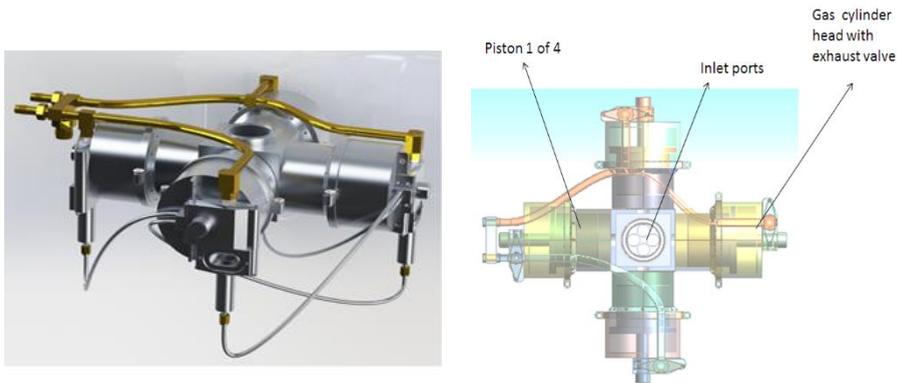As mentioned earlier, this pump does not use any shafts to move the pistons and power transmission. However, in order to move the pistons, we used a pneumatic valve to control the inlet and outlet valves. As shown in Fig. 4, a tap is located at the bottom of each piston. In fact, this valve is driven by the gas used in the system to open and close the gas inlet and outlet. This will control the movement of the pistons in the system and cause the pistons to always move against each other. In addition, such controls reduce friction and damage in the system. Because it causes injection before the piston reaches the end of the motor cycle to prevent its collision with the body. Each cylinder has a standard displacement of 7 cm between the start of motion and the piston stroke per pump cycle. The opening time for the fuel valve is about 9 ms, which is independent of the piston speed.

Investigation of stress and strain of cylinder and piston

One of the important things to keep in mind is that the pump is subject to high internal pressure, which requires careful consideration in its design due to the low thickness of the cylinder wall. Here, we examine the stresses and strains of the cylinder and piston based on the applied pressure and then simulate it using finite element modeling. In Fig. 5, there is a cylinder and a piston. One side is the fuel inlet and outlet, and the other is the gas inlet and outlet valves. The fuel inlet pressure enters the cylinder and moves the cylinder upwards. When the piston reaches a high point, the valve opens and the pressure gas enters the cylinder, causing the piston to move downward and leave the fuel outlet.

161
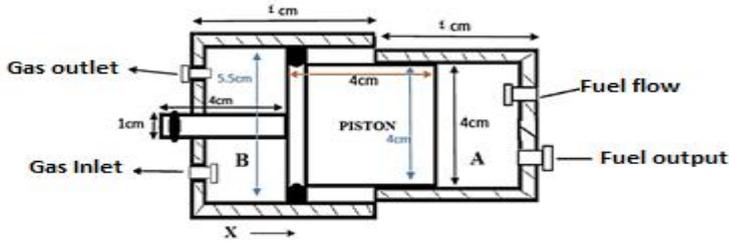
Fig. 5. Pressure cylinder and piston

We also use the specifications given in Table 2 to solve the problem.

Table 1. Pump specification

Fuel inlet pressure	1 MPa
Inlet gas pressure	4 MPa
Output fuel pressure	4 MPa
The thickness of the cylinder walls	3 mm
Speed of opening and closing of valves	9 ms
Diameter of inlet and outlet valves	5 mm

Numerical Problem Solving:

The first step in calculating the stress and strain of the cylinder is to examine its thin or thick wall. If we call the cylinder thickness t and the inner diameter of the cylinder din, we classify the cylinders according to the value [16]:

(1) $a = \dfrac{t}{d_{in}}$

It will be done. If the value a is less than 20.1, we call the cylinder a thin wall, otherwise the cylinder is a thick wall. There are different opinions on the boundary value of a. It can be larger than 20.1 (such as 0.5 or 0.2), but it should be noted that the smaller a is, the more accurate the calculations for the thin wall cylinder are. A condition of 1/20 for the value of a gives the results with acceptable accuracy. So, the first step in calculating the stress and strain of the cylinder is to check whether the wall is thin or wound [17].

According to the formula given for both sides of the cylinder is obtained.

$a_A = \dfrac{0.003}{0.04} = 0.075;$

$a_B = \dfrac{0.003}{0.055} = 0.054.$

162

So, the first and second parts of the thin wall cylinder are assumed. We now turn to problem calculations.

Longitudinal Stress in Cylinders

To obtain the longitudinal stress of the cylinder, we use the following formula:

(2) $\qquad \sum F_X = 0;$

(3) $\qquad \partial_x (2\pi rt) = P (\pi r^2) \rightarrow \partial_x = \partial_L = \frac{Pr}{2t} = \frac{Pd}{4t}.$

According to the above formulas, the longitudinal stress for both sides of the cylinder is calculated as follows.

$$\delta_{lA} = \frac{5000000 \times 0.04}{4 \times 0.003} = 150 \text{ Pa};$$

$$\delta_{lB} = \frac{4000000 \times 0.055}{4 \times 0.003} = 165 \text{ Pa}.$$

Environmental Stress in Cylinders

The environmental stress of the cylinder is obtained from the following relationship.

(4) $\qquad \sum F_y = 0;$

(5) $\qquad \partial_y (2Lt) = P (2rL) \rightarrow \partial_y = \partial_h = \frac{Pr}{t} = \frac{Pd}{2t}.$

The environmental stresses obtained are as follows.

$$\delta_{hA} = \frac{5000000 \times 0.04}{2 \times 0.003} = 300 \text{ Pa};$$

$$\delta_{hB} = \frac{4000000 \times 0.055}{2 \times 0.003} = 330 \text{ Pa}.$$

The radial stress in the first and second parts of the cylinder is assumed to be zero due to its thin wall.

Shear Stress at 45 and 60 Degrees:

We use the following formula to obtain the stress at angles of 45 and 60 degrees on both sides of the cylinder.

(6) $\tau_\theta = \frac{1}{2}(\partial_x - \partial_y)\sin 2\theta.$

   Part A cylinder with a degree angle.

   $\tau_\theta = 0.5 \times (|150 - 300|)\sin 2 \times 45 = 0.5 \times 150 = 75$ Pa.

   Part A cylinder with a 60-degree angle.

   $\tau_\theta = 0.5 \times (|150 - 300|)\sin 2 \times 60 = 0.5 \times 150 \times 0.866 = 64.95$ Pa.

   Part B with a 45-degree angle.

   $\tau_\theta = 0.5 \times (|165 - 330|)\sin 2 \times 45 = 0.5 \times 165 = 82.5$ Pa.

   Part B cylinder with a 60-degree angle.

   $\tau_\theta = 0.5 \times (|165 - 330|)\sin 2 \times 60 = 0.5 \times 165 \times 0.866 = 71.445$ Pa.

The strain on the cylinder

The longitudinal strain for both sides of the cylinder is obtained from the following relation:

(7) $\varepsilon_x = \frac{1}{E}(\delta_X - \vartheta\delta_y).$

E is a modulus of aluminum elasticity, which is 69 GPa, which after conversion to Pagal is considered to be 69×10^9 and is a fixed number.
Therefore, the longitudinal strain for Part A of the cylinder is equal to

   $\varepsilon_{xA} = \frac{1}{69\times10^9}(150 - 0.32 \times 300) = 14 \times 10^{-10}.$

Also, for Part B cylinder is equal to

   $\varepsilon_{xB} = \frac{1}{69\times10^9}(165 - 0.32 \times 330) = 23 \times 10^{-10}.$

Cylinder peripheral strain

The strain created in the cylinder environment is obtained using the following equation:

(8) $\varepsilon_y = \frac{1}{E}(\delta_y - \vartheta\delta_x)$

According to the stated relation, the strain created in Part A is a cylinder.

164

$$\varepsilon_{yA} = \frac{1}{69 \times 10^9}(300 - 0.32 \times 150) = 3.64 \times 10^{-9}.$$

And for Part B cylinders too

$$\varepsilon_{y1} = \frac{1}{69 \times 10^9}(330 - 0.32 \times 165) = 4 \times 10^{-9}.$$

The radial strain is zero because the thin cylinder wall is assumed.

The stress on the piston

(9) $\sigma_A = \frac{F_2}{A_2} = P_2 = 5$ MPa;

(10) $\sigma_B = \frac{F_1}{A_1} = P_1 = 4$ MPa.

The strain on the piston

$$\varepsilon_A = 5 \times 10^6 \times 69 \times 10^9 = 345 \times 10^{15};$$

$$\varepsilon_B = 4 \times 10^6 \times 69 \times 10^9 = 276 \times 10^{15}.$$

Modeling results

The results for the stress and strain of the cylinder and piston according to the pressure exerted on the inner wall of the cylinder using finite element modeling are as follows.

By taking the stress and strain outputs from a node and drawing the diagram for it, the strain diagram of the cylinder can be obtained. Fig. 6 depicts the strain diagram of a node in directions X and Y.

a)   *b)*
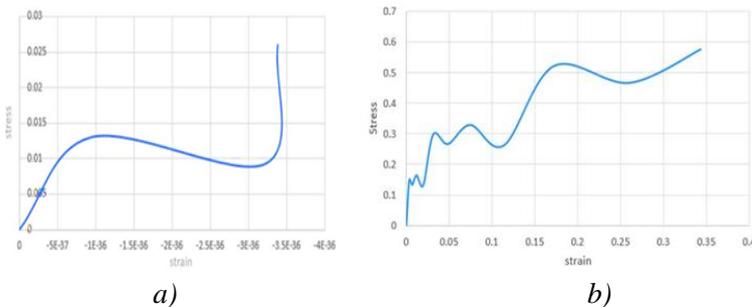
Fig. 6. Stress and strain diagram of the cylinder in direction of: a) X and b) Y

As shown in Fig. 7, the highest and lowest stresses on the cylinder can be observed in the finite element results. The maximum stress in the body is 1136 MPa.

Fig. 7. Stress on the inner wall of the cylinder

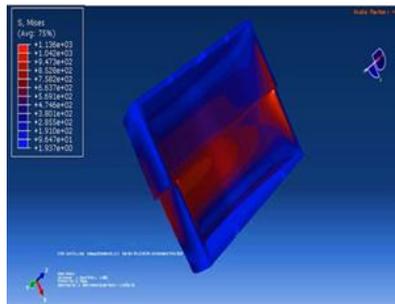The same steps are then repeated for the piston and the results are shown in Fig. 8.

a)   *b)*

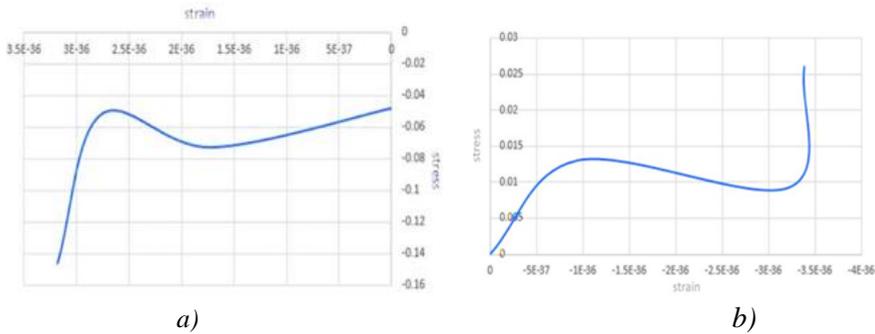*Fig. 8. Stress and strain diagram of the piston in direction: a) X and b) Y*

The maximum stress and strain in the piston is as shown in Fig. 9.
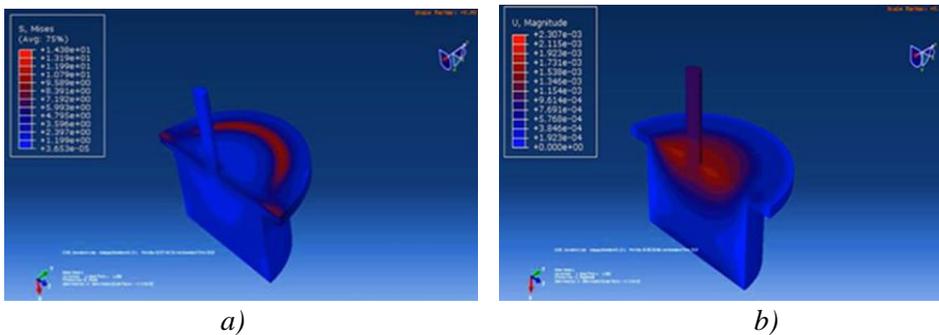

a)   *b)*

*Fig. 9. a) Maximum and minimum stresses on the piston b) The strain applied
to the piston*

Conclusion

The space rocket consists of a number of subsystems and has been designed to maximize safety, simplicity, and redundancy in all respects. By keeping the fuel injection using the reciprocating pump as an important device, the results indicate that there are fewer modes of failure than with other pumps. A series of stress and flow analyses for the pump were performed using ANSYS, and stress and pressure distributions were recognized. In order to confirm the accuracy of our modeling results, comparisons with experimental results were performed, and a very good agreement was found. Future development work will be performed through other pump design factors of importance will be factored in such as the materials' properties and determine the heating of various pump components, as well as any subsequent thermal stresses.

References

1. Wofford, S., J. Fred Jue, J. Cook, Space Shuttle Main Engine Design Evolution, 57th Joint Army-Navy-NASA-Air Force Propulsion Meeting / 5th Liquid Propulsion Subcommittee Meeting. Colorado Springs, CO. May 3-7, 2010.
2. R. Biggs, Space Shuttle Main Engine: The First Twenty Years and Beyond. American Astronautically Society History Series, Volume 29. Univelt for American Astronautical Society, San Diego, CA, 2008.
3. Betts, E., R. Frederick, R. A., Jr.: A Historical Systems Study of Liquid Rocket Engine Throttling Capabilities, AIAA Paper 2010-6541, 2010.
4. U. Barske, High Pressure Pumps for Rocket Motors. Rocket Propulsion Department, Wescott. August 2016.
5. R. Albat, G. Langel, O. Haidn, Handbuch der Raumfahrttechnik, Hanser Verlag, 2007.
6. EOPortal Directory: All-Electric Propulsion and Multi-Satellite Launch System for Communication Satellites, 2020.
7. D. Lev, et al., The technological and commercial expansion of electric propulsion in the past 24 years. In Proceedings of the 35th International Electric Propulsion Conference, Atlanta, GA, USA, 8–12 October 2017.
8. Z. Rosenberg, Xcor tests piston pump-fed rocket engine, August 2016.
9. B. Campen, Liquid oxygen piston pump ready for reusable space flight, April 2017.
10. J. C. Whitehead, Propulsion Engineering Study for Small-Scale Mars Missions, UCRL-CR-122442, Univ. of Calif. Lawrence Livermore National Lab, September1995.
11. J. C. Whitehead, United States Patent, Patent Number: 5,026,259 , Date of Patent: Jun. 25, 1991.
12. J. C. Whitehead, Mars ascent propulsion options for small sample return vehicle , May 12, 1997.
13. M. A. Corbo and Stearns, Ch. F., Practical design against pump pulsations, Proc. 22nd International Pump Users Symposium, 2005.
14. J. C. Whitehead, Hydrogen Peroxide Gas Generator Cycle with a Reciprocating Pump, AIAA Paper 2002-3702, July 2002.

15. M. Ventura, Novel Concepts for an Advanced Non-Toxic Gas Generator, 42nd AIAA/ASME/ SAE/ASEE Joint Propulsion Conference & Exhibit Sacramento, CA, July 9-12, 2006.
16. E. J. Hearn, Mechanics of Materials 2: The mechanics of elastic and plastic deformation of solids and structural materials, Butterworth-Heinemann, 1997.
17. G. P. Sutton and O. Biblarz, Rocket Propulsion Elements, Wiley, 2001.

КОСМИЧЕСКИ РАКЕТЕН ДВИГАТЕЛ ЗА МНОГОКРАТНО ИЗПОЛЗВАНЕ НА НИСКА ЦЕНА, ОБОРУДВАН С ВИСОКОЕФЕКТИВНИ ПОМПИ

А. Сабоктакин, М. Монджези

Резюме

Космическото задвижване използва главно бутални помпи, които са напълно интегрирани с малките приложения. В тази статия са демонстрирани специфичните спецификации и конструктивни фактори, на които трябва да отговарят горивните помпи за ракетни двигатели, и е направено сравнително проучване на пригодността на всички необходими видове помпи, които да се използват с ракетни двигатели и техните приложения. Освен това се изследва бутална помпа, предназначена за зареждане на дюзите за течно гориво. Поради производителността на тази помпа са оценени различни части и нейната производителност. Дизайнът на помпата се обсъжда така, че след завършване на проектните стъпки в съответствие с производителността на помпата, численото решение на напрежението и деформацията, приложени върху вътрешната стена на цилиндъра, се извършва в съответствие с нейната производителност при вътрешно налягане. След това, като се има предвид плътността на материала на тялото и същото налягане и топлина, приложени към цилиндъра и буталото, проба от цилиндъра и буталото се анализира с техника за моделиране с крайни елементи и се представят резултатите от моделирането, получени от симулацията.

AFM ANALYSIS OF GLASSY CARBON COATINGS AFTER AN EXTENDED STAY ON THE INTERNATIONAL SPACE STATION (ISS)

Anna Bouzekova-Penkova[1], Silviya Simeonova[2], Dimiter Teodosiev[1]

[1]Space Research and Technology Institute – Bulgarian Academy of Sciences
[2]Department Physical Chemistry, Faculty of Chemistry and Pharmacy, Sofia University St. Kliment Ohridski
e-mail: a_bouzekova@space.bas.bg, fhsss@chem.uni-sofia.bg, dteod@space.bas.bg

Abstract

In this paper we present the results of atomic force microscopy (AFM), which is applied to characterize the surface morphology of graphite samples with glassy carbon coatings is developed, stored for 28 months under different conditions. One sample was stored on Earth at room temperature and second one was mounted on the outer side of the International Space Station.

Introduction

In the course of the development of the technique, various methods and technologies are developed and used to prepare sensors for measuring direct (DC) and alternating (AC) electric fields in space plasma. The study and knowledge of the characteristics of materials for creating sensors, including materials for creating Legmuir sensors, presumption the use of technologies for creating materials with certain qualities of the working surfaces of sensitive elements of the sensors.

The main requirement for these materials is determined by the need to increase the accuracy and sensitivity of measurements. The accuracy of the electric field measurements, by the double probe method [1–6], strongly depends on variations, the separation work function for electrons on the working surfaces of the probes, as a result of the working environment characterized by radiation exposure, charged particle flows, sudden changes in temperature, high vacuum, the concentration of the ambient plasma, bombardment of their working surfaces by micrometeorites, vibration and shock loads, etc.

Glassy carbon is a preferred material for making sensors. It is a black glass-like material with a shiny surface and fragility, earning it the name vitreous or glassy carbon. GC is easy to process and can be produced in various shapes, sizes, and cross-sections [7–10]. Electron work function measurements on sample surfaces

with glassy carbon (CG) coatings or monolithic glassy carbon coatings show the highest value, which suggests little photoemission and provides minimal point-to-point variation across the surface of the probes [11–13].

Materials and Methods
Sample Types

Graphite samples coated with glassy carbon stored in different environments for a period of two years and four months were analyzed. One of the samples was stored in Earth's environmental conditions, "reference", and the other sample was mounted outside of the International Space Station for a period of more than two years, "space" [14]. In outer space, the samples were exposed to radiation and a two-hour cyclic temperature change in the range of ~ 300°C. The purpose of the experiment was to obtain new data on the influence of space factors on the physico-chemical characteristics of materials used in space research and clarify the aging processes of these materials, as a result of their long stay in the conditions of outer space.

The method of coating graphite with glassy carbon is an original Bulgarian technology developed and implemented by a team of scientists of the Bulgarian Academy of Sciences (IKIT-BAS and IMCTH) and is protected by a copyright certificate [15].

For the studies presented below, samples with dimensions of $15 \times 0.5 \times 0.3$ mm were cut, according to the requirements of the AFM apparatus. Measurements were made on three samples of "reference" R1, R2, and R3 and three samples of "space" S1, S2, and S3. The analyzed surfaces for both types of samples are carried out at five points: end, periphery, center, periphery and end, along the diagonal of the sample for the purpose of reproducibility of the results. The "space" samples were investigated from the front side of the sample - direct contact with outer space, and from the back side of the sample from the side of the block in which they were located [1].

Characterization methods

AFM imaging was performed on the Nano Scope V system (Bruker Ltd, Germany) operating in tapping mode in the air at a room temperature. We used silicon cantilevers (Tap 300Al-G, Budget Sensors, Innovative solutions Ltd. Bulgaria) with 30 nm thick aluminum reflex coatings. According to the producer's specifications, the cantilever spring constant and the resonance frequency are in the range of 1.5 to 15 N/m and 150 ± 75 kHz, respectively. The radius of tip curvature was less than 10 nm. The scanning rate was set at 1 Hz and the images were taken in the highest possible resolution mode of the AFM 512×512 pixels in JEPG format. The NanoScope software was used for the section analysis and roughness of all images.

Results and discussion

Atomic force microscopy (AFM) is widely used to study the topography of materials with great accuracy and precision. By measuring the surface of the material at the nano level, the roughness of the investigated sample is quantified.

The topography of the investigated samples "space" - face, "space" - back and "reference" are shown in Fig. 1 and Fig. 2 in 2D and 3D format at a single point of the performed analysis with a scan area of 5 μm × 5 μm. The morphology of the "space" – face, "space" – back samples have a smoother surface in comparison with the morphology of the "reference" samples with the same scan area (Fig. 1 and Fig. 2). In the AFM images of the "reference" samples, structures with different diameters of the order between 0.5 μm and 1 μm are observed. The surface of the three graphite samples coated with glassy carbon "space" samples – back side S1(*BS*), S2(*BS*), S3(*BS*) is smoother in comparison to the surface of "space" samples – front S1(*F*), S2(*F*), S3(*F*). The difference in R_q values for the two types of samples is 0.34 nm for the second S2, 1.56 nm for the first S1 and 1.84 nm for the third S3 samples, respectively. These differences in the roughness values of the material on the front and back sides are due to the influence of sudden temperature changes, radiation, meteors, etc. in space.
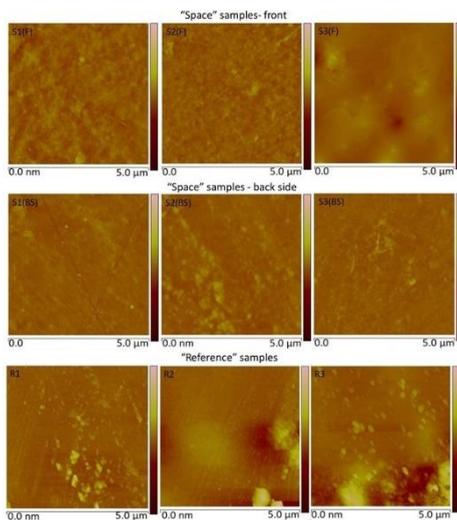
Fig. 1. 2D AFM images 5 μm × 5 μm of the surface of the three graphite samples coated with glassy carbon – "space" samples – front (S1(F), S2(F), S3(F)), "space" samples – back side (S1(BS), S2(BS), S3(BS)), and "reference" samples (R1, R2, R3)
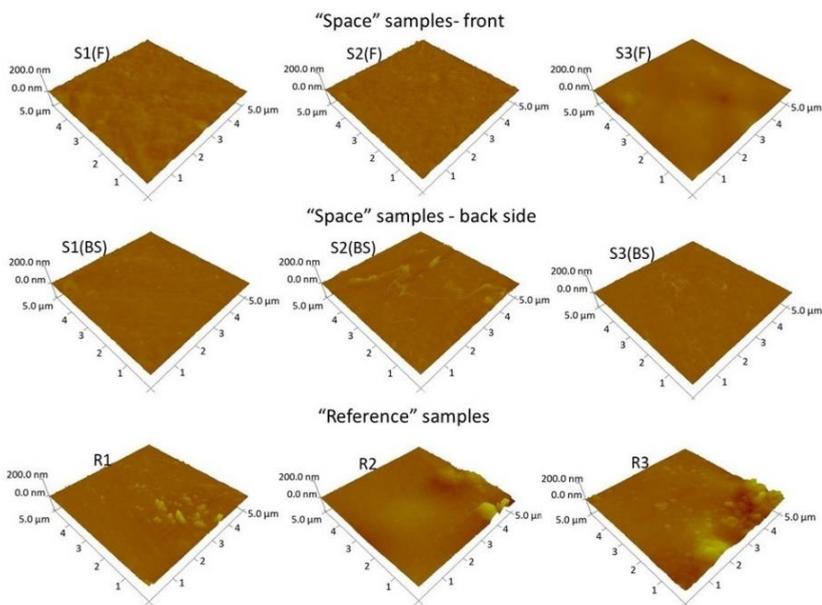
Fig. 2. 3D AFM images 5 μm × 5 μm of the surface of the three graphite samples coated with glassy carbon – "space" samples – front (S1(F), S2(F), S3(F)), "space" samples – back side (S1(BS), S2(BS), S3(BS)), and "reference" samples (R1, R2, R3)

The roughness analysis gives the value R_a which is the arithmetic mean of the absolute values Z_i of the surface height deviations measured from the mean plane, i.e.

(1) $R_a = \frac{1}{N} \sum_{i=1}^{N} |Z_i|$

while R_q is the root-mean-square value of the height deviations taken from the plane of the average images date [16].

(2) $R_q = \sqrt{\frac{1}{N} \sum_{i=1}^{N} Z_i^2}$

Fig. 3 presents the average values obtained from the five analysis points for each of the investigated samples for R_a and R_q.
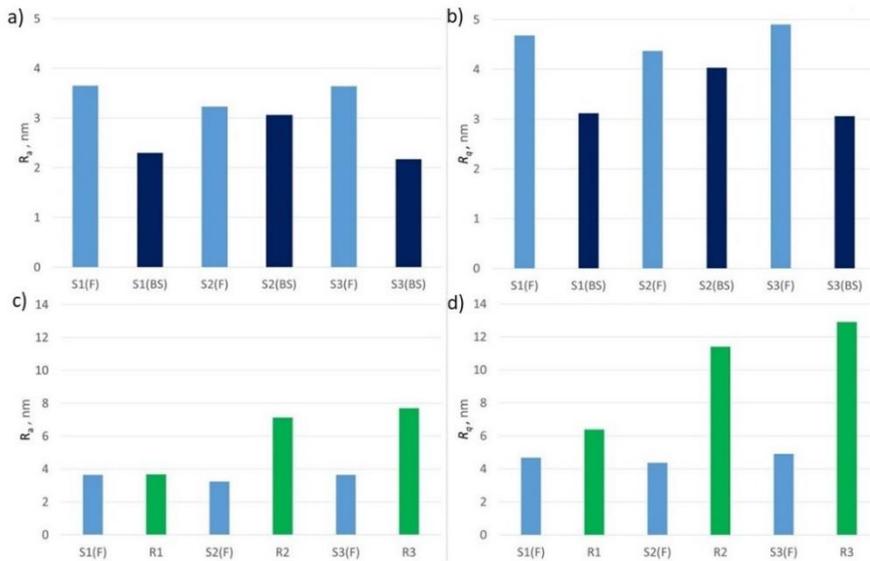
Fig. 3. Roughness analysis of AFM images 5 μm × 5 μm for the three graphite samples coated with glassy carbon – a) R_a for "space" samples – front and "space" samples – back side; b) R_q for "space" samples – front and "space" samples – back side; c) R_a for "space" samples – front and "reference" samples; d) R_q for "space" samples – front and "reference" samples

The morphology and roughness of graphite samples with glassy carbon coatings were compared and presented in Table 1 and Fig. 3. The "space" samples - face "space" samples - front (S1(F), S2(F), S3(F)) have a rougher surface for the values of R_a and R_q compared to the surface of "space" samples - back side (S1(BS), S2(BS), S3(BS)).

The differences in the roughness value for R_a of S1(F) versus S1(BS) is 3.65 nm to 2.3 nm and for the R_q value - 4.68 nm to 3.12 nm, respectively. For sample S2, the differences in the roughness value for R_a of S2(F) versus S2(BS) are smaller, 3.23 nm to 3.06 nm, respectively, and for the value for R_q – 4.37 nm to 4.03 nm, respectively. The difference in the roughness value for R_a of sample S3(F) is approximated as sample S1(F). The roughness value for R_a of S1(F) vs. S1(BS) is respectively 3.64 nm to 2.17 nm and for the value of R_q – 4.90 nm to 3.06 nm, respectively for "space" – face, "space" – back. The roughness values for R_a and R_q of "reference" samples R-1, R-2, and R-3 are higher than those of "space" samples - face and "space" samples - back side.
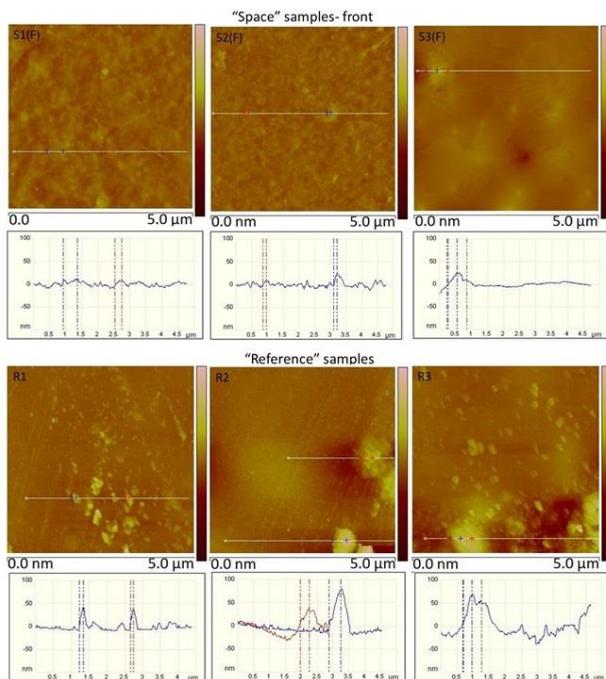
*Fig. 4. 2D AFM images and section analysis of the surface of "space" samples –
front and "reference" samples*

Atomic force microscope data provide valuable information on the homogeneity of the glassy carbon layer for "space" samples. The AFM images of both types of "reference" and "space" samples were compared and presented in Fig. 4. The morphology of the "reference" samples with a scanning area of 5×5 µm^2 (R-1, R-2, R-3) is smooth with the presence of places with "spherical" structures in comparison with the morphology of the "space" samples –front (S1(F), S2(F), S3(F)). From the cross-section and surface roughness determined for both the "reference" and "space" samples, it was found that the "reference" samples have a rougher surface.

Conclusions

The surface topography of the glassy carbon coatings was studied using Atomic force microscopy. The surface morphology of the surface of graphite samples with glassy carbon coatings, stored for 28 months in Cosmos "space" - face, "space" - back is smooth and homogeneous. The surface morphology of graphite samples with glassy carbon coatings stored for 28 months on Earth at room temperature "reference" is rougher with the presence of structures with diameters

from 0.5 μm to 1.5 μm. The roughness values for Ra and Rq of "reference" samples R-1, R-2, and R-3 are higher than those of "space" samples - face S1(F), S2(F), S3(F). The presence of "spherical" structures on the surface of the "reference" samples is a result of the storage conditions on Earth. The surface of the deposited glassy carbon coatings on a graphite substrate of both types of samples – "reference" and "space" is homogeneous with small differences in roughness values.

References

1. Bonnet, I., Soonckindt L., Lassabatere L.,The Kelvin probe method for work function topographies: technical problems and solutions, Vacuum, 1984, 34, 693–698.
2. Boyd, R. L., Measurement of electric field in the ionosphere and magnetosphere, Space Sci.Rev., 1967, 7, 230.
3. Fahleson, U., Theory of electric field measurements conducted in the magnetosphere with electric probes, Space Sci.Rev., 1967, 7, 238.
4. Fahleson, U. V., M. C. Kelly and F. S. Mozer, Investigation of the operation of a DC electric field detector, Planet.Space Sci., 1970, 18, 1551–1561.
5. Maynard, N. C., E. A. Bielecki, H. F. Burdick, Instrumentation for vector electric field measurements from DE-B, Space Science Instruments, 1981, 5, 523.
6. Mozer, F. S., Analyses of techniques for measuring DC and AC electric fields in the magnetosphere, Space Sci. Rev., 1973, 14 (2), 272–313.
7. Uskoković, V., A historical review of glassy carbon: Synthesis, structure, properties and applications. Carbon Trends. 2021, 5, 100116.
8. Fitzer, E., K.-H. Köchling, H. P. Böhm, H. Marsh, Recommended Terminology for the Description of Carbon as a Solid (©1995 IUPAC), Pure and Applied Chemistry, 1995, 67(3), 473–506.
9. Harris, P. J. Carbon, Nanotubes and Related Structures, Cambridge, Cambridge University Press, 1999, 356.
10. Miteva, A., Some applications of glassy carbon – an overview, Proceedings of the Annual University Scientific Conference 2023, Vasil Levski National Military University, 08 – 09.06.2023, V. Turnovo, Bulgaria, 2023, 2, 1459–1466.
11. Mozer, F. S., Analysis of techniques for measuring DC and AC electric fields in the magnetosphere, Space Sci.Rev., 1973, 14(2), 272.
12. Bonnet, J., Soonckindt L., Lassabatere L., The Kelvin method for work function topographies: technical problems and solutions, Vacuum, 1984, 34, 693–698.
13. Teodosiev, D., N. Isaev., Physicochemical Properties of Probes for Measuring Electric Fields in the Near-Earth Plasma, Chemical Physics Reports, 20011, 9(6), 1165–1175.
14. Bouzekova-Penkova, Anna, Stanislav Klimov, Valery Grushin, Olga Lapshinova, Denis Novikov, Dimitar Teodosiev, Space Experiment "Obstanovka (1 - stage)", block DP - PM of the Russian segment of the International Space Station (ISS), Aerospace Research in Bulgaria, 2023, 35, 156–164.

15. Теодосиев, Д., Печеняков, И., Георгиев, Й., Петров, П., Добрев, Р., Вълов, Р., Станев, Г., "Метод за уплътняване и покриване на порьозни огнеупорни материали със стъкловъглерод" Авторско свидетелство № 31847, Май 11, 1982. (in Bulgarian).
16. Eaton, P., West P., Atomic Force Microscopy, 1st ed.; OXFORD University Press Inc.: New York, USA, 2010.

АФМ АНАЛИЗ НА СТЪКЛОВЪГЛЕРОДНИ ПОКРИТИЯ СЛЕД ПРОДЪЛЖИТЕЛЕН ПРЕСТОЙ НА МЕЖДУНАРОДНАТА КОСМИЧЕСКА СТАНЦИЯ (МКС)

А. Бузекова-Пенкова, С. Симеонова, Д. Теодосиев

Резюме

В тази статия представяме резултатите от охарактеризирането на повърхностна морфология с Атомно-силовата микроскопия (АФМ) на графитни проби със стъкловъглеродни покрития, съхранявани в продължение на 28 месеца при различни условия. Едни проби бяха съхранявани на Земята при стайна температура, а вторите бяха монтирани от външната страна на Международната космическа станция.

XPS INVESTIGATION OF THE WORK FUNCTION OF GLASSY CARBON COATINGS AFTER EXTENDED STAY ON THE INTERNATIONAL SPACE STATION (ISS)

Aleksandar Tsanev[1], Korneli Grigorov[2], Hristo Kolev[3],
Anna Bouzekova-Penkova[2], Boyko Tsyntsarski[4], Peter Tsvetkov[1],
Dimitar Teodosiev[2]

[1]Institute of General and Inorganic Chemistry – Bulgarian Academy of Sciences
[2]Space Research and Technology Institute – Bulgarian Academy of Sciences
[3]Institute of Catalysis – Bulgarian Academy of Sciences
[4]Institute of Organic Chemistry with Centre of Phytochemistry – Bulgarian Academy of Sciences

e-mail: tsanew@abv.bg, kgrigoro@abv.bg, a_bouzekova@space.bas.bg,
Boyko.Tsyntsarski@orgchm.bas.bg, p-tzvetkov@gmx.net, dteod@space.bas.bg

__Keywords:__ X-ray photoelectron microscopy (XPS), Glassy carbon, Cosmic ray and gamma radiation.

Abstract

Using XPS, the electron work function values on the surface of glassy carbon coatings of graphite samples, after an extended stay on a board of the ISS in open space conditions, were investigated. The results were compared with the characteristics of glassy carbon coatings of samples left on the ground for the same period. It was found that the electron work function does not change significantly, and this proves the possibilities for the successful application of these coatings, obtained by original Bulgarian technology, for space experiments on the board of satellites for measuring electric fields in the ionospheric-magnetospheric plasma. The minimal observed variations in the values of the electron work function are explained by small differences in the content of traces of different chemical elements on the surface of the coatings. It has been established that glassy carbon coatings have stable characteristics after a long stay in space, despite the small fluctuations in the values of the electron work function. The results show that glassy carbon coatings are chemically and mechanically stable. The results from this original technological experiment are unique for the development of sensitive elements, such as sensors for measuring weak electric fields in cosmic plasma.

1. Introduction

From the beginning of the space age until today, the double probe method [1–3] has been established as a method that measures the electrical potential difference between two points in space. Variations of the electron work function on

the surfaces of each probe affect the current value in a complex way, especially in the case of inhomogeneous plasmas such as ionospheric-magnetospheric plasmas, and are particularly important for experiments aboard rotating satellites. Differences in the floating potentials for the individual probes, relative to the surrounding plasma, can lead to additional electrical signals, many times exceeding the investigated ones. This fact determines the need for maximum identity of the physico-chemical and geometrical characteristics of the Sensitive Elements of the Probes (SPE). Obviously, a central factor that determines the stable performance of the SPE is the material choice and its surface properties.

During the development of the signal acquisition methods of constant and variable fields, various probe materials have been considered in order to gain better sensitivity and higher signal accuracy registration. The accuracy of electric field measurements, by the double probe method, strongly depends on the variations of the Work Function (WF) of the electrons on the probe surface due to the solar wind (irradiation) in outer space. The electron WF, on the surfaces of samples with Glassy-Carbon (GC) coatings, shows the highest value, about 5.05eV, which suggests lower photoemission compared to the other used materials such as silver, gold, aquadag, etc. Due to these properties of GC materials, spherical probes (Langmuir probes) made of GC have been used in the last three decades in practically all experiments for measuring electric fields. Spherical sensors made of monolithic GC have also been used in other satellite experiments such as GEOS, S3-3, ISEE-1,2, WIND, POLAR, CLUSTER – 1,2,3, 4, FREJA, THEMIS, CRRES, GEOS, Viking, and GEOTAIL [4-9].

An original method has been established [10] for GC synthesis on pyrolytic graphite, a technology that meets severe test requirements, and double SPE, has been installed on satellites, such as IC-Bulgaria 1300, IC-24 Active, IC-25 APEX, INTERBALL-2, the Magion-2-5 sub satellites, as well as the International Space Station (ISS) [10], for the period from 1981 to 2023. In order to study the influence of outer space on the physico-chemical parameters of GC coatings, in the period 2013–2015, an international project "Obstanovka 1-step" was planned and successfully implemented, which included the implementation of a technological experiment block "DP-PM" aboard the International Space Station (ISS) [10].

2. Materials and Methods

Sample Types

The analysed samples with dimensions of $30 \times 15 \times 5$ mm of spectrally pure graphite, compacted and covered with GC, were analyzed and they stayed at different conditions for 28 months. Some samples have been in open space mounted on the Russian module of the ISS [10] and conventionally were called "space", while the others have been stored in terrestrial conditions and named as "reference". To achieve high reliability in the analysis of the obtained data for the "space" samples,

they are compared with "reference". The purpose of the experiment is to investigate the influence of outer space on the physico-chemical and structural characteristics of the GC coatings, after their long stay in the space. The most important characteristic for the stability and durability requirements of these coatings, we used the value and WF variations of the electrons on its surface. In open space, samples were directly exposed to various impacts on the front surface of the sample, which we conventionally call *S-front*, and those that were not exposed to direct contact, *S-back side*.

All samples were analyzed by surface sensitive technique – X-Ray Photoelectron Spectroscopy (XPS). Later the surface of as prepared samples is labeled as "ras", the samples cleaned with isopropyl alcohol are labeled as "rclean", as well as scribed surface, using a P2500 silicon carbide sandpaper, in order to get results more related to the bulk, labelled "rscribed". The cleaning of the samples was done using an ultrasonic bath for 5 min in order to remove the fats and contamination residuals after sample preparation.

Characterization methods

X-ray photoelectron measurements have been carried out on the ESCALAB MkII (VG Scientific, now Thermo-Scientific) electron spectrometer with a base pressure in the analysis chamber of 5.10^{-10} mbar (9.10^{-8} mbar during the measurements), equipped with twin anode MgKα/AlKα non-monochromatic X-ray source used excitation energies of 1253.6 and 1486.6 eV, respectively. The measurements are provided only with an AlKα non-monochromatic X-ray source (1486.6 eV). The instrumental resolution is measured as the full width at a half maximum (FWHM) of the $Ag3d_{5/2}$, photoelectron peak is about 1 eV. The data was analyzed using SpecsLab2, Casa-XPS (Casa Software Ltd), and Origin 11 software. The processing of the measured spectra includes subtraction of X-ray satellites and a Shirley-type background [11]. The peak positions and areas are evaluated by a symmetrical Gaussian-Lorentzian curve fitting [11]. The relative concentrations of the different chemical species are determined based on the normalization of the peak areas to their photoionization cross-sections, calculated by Scofield [12].

Work function determination using the XPS technique

One of the most reliable and widely used spectroscopic techniques for surface analysis is X-ray Photoelectron Spectroscopy (XPS). It is a surface-sensitive, non-destructive, and element-specific technique that provides quantitative information on the chemical state of the elements of the top 10 atomic layers of the surface of the analyzed sample. Moreover, it gives information about the chemical state of the analyzed elements, which is based on the change of the binding energy and the line shape. Furthermore, XPS can easily be adapted to measure the electron

work function of the surfaces without any additional equipment modification or additional sampling procedure. That is because the equipment possesses a well-developed high-resolution electron analyzer, usually with a hemispherical shape and a well-defined X-ray source. However, there are some requirements that have to be taken into account. Samples need to be stable inorganic materials that are resistant to X-rays. It is generally necessary to bias the sample with a small potential voltage (e.g. - 5 V) in order to avoid effects that are a result of scattered electrons in the spectrometer's detection system. Therefore, the WF of insulating/organic samples is difficult to measure by XPS . The electron Work-Function (WF) is dependent very much on the surface condition; even minor modification to the surface brings a dramatic change in its value. The measured value strongly depends on the surface impurities and composition homogeneity. Even after cleaning the surface under ultrahigh vacuum (UHV) condition, the WF changes are often observed over time due to the contamination from residual gases within the chamber. Figure 1 illustrates the procedure of WF measurement for GC probes. The first calibration of the spectrometer is performed by the C1s element at 284.5 eV, and hence, the Fermi level (E_F) could be established. At that point, we only need to define the secondary electron cut-off position. Then, the WF is defined as the difference between X-ray excitation energy and secondary electron cut-off energy.
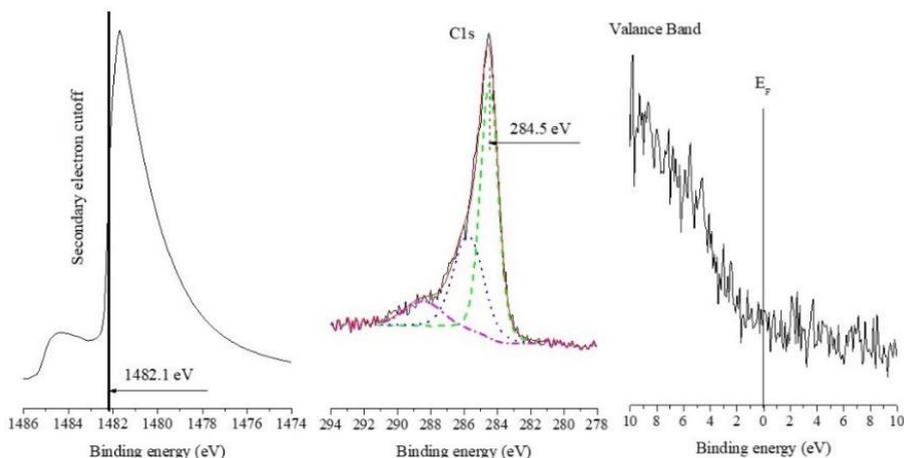

Fig. 1. WF measurement of GC material using spectrometer by XPS

3. Results and Discussion

In our approach, we use the method of WF measurement, generally explained by Kim & Kim [13], where the total photoelectron spectrum is displayed in the range between X-ray energy and the sample value of WF. Practically, two

values have to be obtained in order to calculate WF, namely, secondary electron cut-off and the maximum kinetic energy of photoelectrons emitted from a metal, defined as Fermi level (EF), which appears at 0 eV binding energy in the XP spectrum (see Fig. 1). Thus, the WF is obtained by the difference between the of X-ray photons energy (1486.6 eV) and the secondary electron cut-off energy. One can conclude that these energy values, secondary electron cut-off, and Fermi level EF at 0 eV, have to be measured very precisely. Therefore, calibration of the spectrometer by measuring noble metal surface after its cleaning, e.g. palladium, is crucial. The performed measurement of the well-defined surface of the Pd-metal gave a result of 5.2 eV, as expected, equal to the book values (5.22 eV) of Pd-metal WF. In this way, our measurements of glassy carbon surfaces can be considered reliable. Moreover, we have used additional parameter which has been followed to assure the calibration of the binding energy scale of spectra, namely C1s core level. Regardless of the presence of impurities and contaminations, we are investigating surfaces that consist basically of glassy carbon. Therefore, we can assume that the carbon 1s core level will have predominantly binding energy at 284.5 eV [14]. This gave us a chance to lower the error by defining EF = 0 eV of the investigated surface, which is challenging because of the low intensity and noisy spectrum of the valence band. Additionally, to determine it more precisely, the slope of the valence band was modeled using a high-degree polynomial and then the model curve was differentiated. The center of the resulting peak coincides with the Fermi level. Figure 1 demonstrates the shapes of the corresponding measured peaks of Secondary electron cutoff, C1s glassy carbon core level with additional sub-peaks resulting from C - O and C = O bonds impurities and contaminations, as well as the valence band slope and Fermi level.

Several surfaces of the sample were measured: various contamination has been found on the as prepared surface, labeled as (ras); surfaced cleaned with isopropyl alcohol ultrasonic bath for 5 min. in order to remove fats and contamination residuals after sample preparation (rclean); as well as scribed surface, using a P2500 silicon carbide sandpaper, in order to get results more related to the bulk (rscribed). The ras surface shows various discolorations on different points. Therefore, we performed XPS measurements on several points on the surface to study this discoloration.

The results summarized in Table 3 show differences in the concentration of elements presented on the surface as well as variations in the composition. As a consequence, we can expect different WF values for these surfaces. Indeed, values between 3.3 - 4.8 eV of WF have been obtained for the sample ras and rclean for all studied points. For the scribed GC sample, we have observed a relatively homogeneous surface with one discoloration and fewer impurities. Impurities on the glassy carbon surfaces like silicone, oxide, nitrogen, chlorine, and others have also been reported by other authors [15–17]. Close or even the same values are obtained by other authors published in scientific literature, for instance, by Ivey [18].

As expected, we obtained WF with close values for all the measured points from 4.3 to 4.6 eV. Small deviations of the WF values are observed for the exposed samples (S1, S2), rather than for the "reference" one (R). These results are summarized in Table 1, where S1-2 and S2-2 are noted as repeated measurements on samples S1 and S2. As expected, we obtained close WF values of 4.8 eV and 5.0 eV, respectively. These values are equal within the error bars and fully consistent with the close predominant carbon and oxygen atom concentrations of both surfaces.

For the "space" sample S2, the electron work function value for the face is 4.8 eV and the work function value for the back is 5.0 eV. These values are again very close and fully consistent with the close oxygen concentrations on the two surfaces. All discussed results are graphically expressed in Fig. 2 and Fig. 3, and the results are shown in Table 1.

Table 1. Fermi level and work-function of referent and exposed (space) samples

Sample:	Side	Fermi Level [eV]	HBE* [eV]	Work Function [eV]
R	face	-28.2	1454.0	4.8
	back	-29.6	1452.6	4.8
S1	face	-31.1	1451.1	4.4
	back	-30.4	1451.4	4.8
S2	face	-30.9	1450.3	5.4
	back	-30.2	1450.8	5.6
S1-2	face	-30.0	1451.4	5.2
	back	-29.5	1451.9	5.2
S2-2	face	-30.9	1450.9	4.8
	back	-30.8	1450.8	5.0

A standard sample of monolithic glassy carbon, protected by copyright [20], shows a very homogeneous surface with only one surface colour - dark grey. Nevertheless, the XPS study also detects some impurities on the surface of this sample. The surface atomic concentrations of the monolithic glassy carbon sample [20] are summarized in Table 2. The solutions used for the preparation of glassy carbon can be the origin of these impurities. The obtained WF of this sample is 4.5 eV, which is equal, within the error bar, to the value reported in the scientific literature [16–18].

Table 2. Surface atomic concentration of homogeneous monolithic glassy carbon [20]

Sample	C [at.%]	O [at.%]	N [at.%]	Si [at.%]	S [at.%]	Ca [at.%]	Cl [at.%]	Al [at.%]
f	79.10	14.03	2.57	0.78	0.82	0.85	0.31	1.55

Table 3. Surface atomic concentrations obtained for "ras", "rclean", and "rscribed"
surfaces of the "reference" GC samples

Sample	C [at.%]	O [at.%]	N [at.%]	Na [at.%]	Zn [at.%]	Si [at.%]	S [at.%]	Ca [at.%]	Cl [at.%]	F [at.%]	Al [at.%]	Mg [at.%]
ras Point 1	85.40	11.20	2.70	0.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ras Point 2	80.70	13.30	2.80	0.70	0.20	1.50	0.80	0.00	0.00	0.00	0.00	0.00
rclean Point 1	71.25	19.15	2.94	0.22	0.00	2.87	0.00	0.95	0.36	0.66	1.60	0.00
rclean Point 2	69.41	18.61	3.39	0.34	0.16	3.76	0.00	0.99	0.00	0.61	2.51	0.23
rclean Point 3	67.36	18.79	2.92	0.63	0.49	4.46	0.00	1.78	0.00	1.17	2.16	0.23
rclean Point 4	79.52	14.38	2.17	0.00	0.00	2.30	0.00	1.34	0.00	0.00	0.29	0.00
rclean Point 5	89.68	7.98	1.70	0.00	0.00	0.00	0.00	0.63	0.00	0.00	0.00	0.00
rscribed Point 1	93.20	4.51	0.00	0.00	0.00	1.90	0.00	0.00	0.00	0.39	0.00	0.00
rscribed Point 2	92.04	6.23	0.00	0.00	0.00	1.42	0.00	0.00	0.00	0.30	0.00	0.00
rscribed Point 3	93.15	4.71	0.00	0.00	0.00	1.70	0.00	0.00	0.00	0.45	0.00	0.00

Furthermore, we adopt the following notations for the studied samples: for the "reference" samples, face and back – R-front and R-back side and for the "space" samples, S1-front and S1-back side; S2-front and S2-back side, respectively. One can expect that there will be differences between both sides of the sample, namely, the side that is exposed directly to cosmic radiation (front) and the sample backside (back side). Therefore, both sides are studied by means of the XPS technique. The calculated WF values for sample S1-front and S1-back side are 4.4 eV and 4.8 eV, respectively. The small variation of these WF values can be explained by a significant difference measured for oxygen from 13 at. % to 7.6 at. % and presence of silicon, sulfur, and calcium atoms on the surface only on the S1-front. The sample S2 shows similar results with some minor variation in the surface atomic concentrations compared to both the S2-front and the S2-back side surfaces. The surface atomic concentrations of both samples – "reference" and "space" are displayed in Table 3.
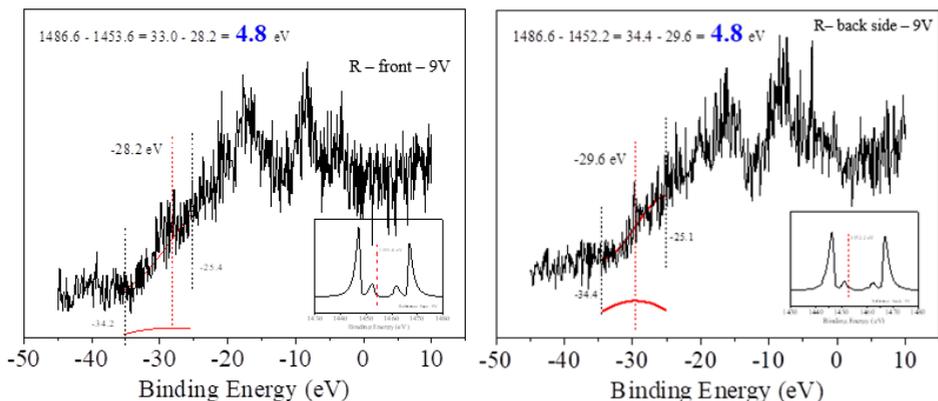
1486.6 - 1453.6 = 33.0 - 28.2 = **4.8** eV

R – front – 9V

-28.2 eV

-25.4

-34.2

1486.6 - 1452.2 = 34.4 - 29.6 = **4.8** eV

R– back side – 9V

-29.6 eV

-25.1

-34.4

Fig. 2. XPS valence band spectrum for the "reference" samples R-front (a)
and R-back side (b)

Summarizing and comparing the results for the WF regarding the "reference" samples R- front and R-back side of the "space" S1 and S2, we can conclude that the values between 4.4 and 4.8 are close-range values, and their insignificant difference lies in the impurities and contaminations of the surfaces. In this sense, and due to the importance of material WF used as a major parameter in the design of devices, we can conclude that the investigated glassy carbon can be used as a stable, sensitive element in electronic detectors.

In Table 3, one of the main components of the near-surface layers is the oxygen, whose concentration ranges over a wide range from approximately 8 at. % to about 15 at.%. Many other chemical elements are present on the surface, among which, the main ones, are nitrogen, sodium, silicon, fluorine, and aluminum. In the reference sample, there is also, although insignificant, a difference between the amount of oxygen on the face and the back of the sample - on the face, it is 11.2 at. %, while on the back, it is 13.3 at. %, i.e., the difference is about 2 at. %. At the same time, the concentration of nitrogen and sodium is approximately equivalent on both sides. Unlike the face, silicon and minor amounts of sulfur were found on the back of this specimen. Much more significant is the difference in oxygen content between the face and the back of sample *S1*. On the face of the sample, the oxygen coverage of this sample is 13 at. %, while on the back, it is 7.6 at.%.
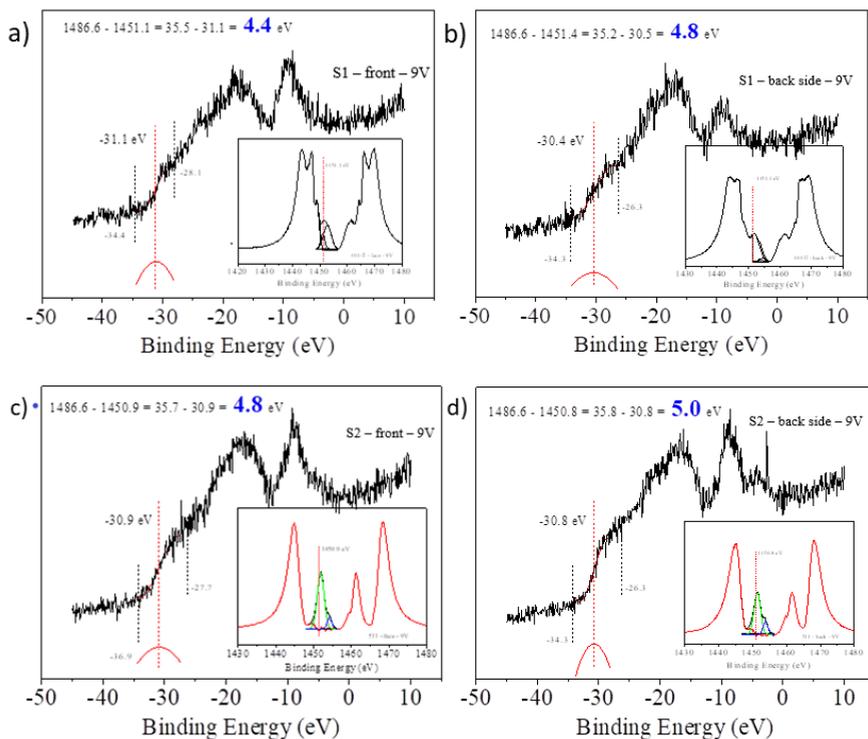
*Fig. 3. XPS valence band spectrum for samples S1-front (a) and S1-back side (b)
and S2 – front (c) and S2-back side (d)*

To verify the reproducibility of the data from the experiments, a second measurement of the chemical element concentration of the GC coating on the face and back was made for this sample. The results displayed in Table 3 show that the reproducibility of the measurements is very good. The oxygen concentration in sample S2 is also a bit different. On the face, the oxygen is 9.7 at. %, while on the back, it is significantly less – 10.8 at.%. What is interesting about this specimen is that significant amounts of silicon (Si) were also observed on the face, more so than the other specimens. On this sample, XPS analysis was performed at various points, on the face and back, with the size of each analyzed spot being 700 μm × 300 μm.

The analyses that were conducted show that the chemical elements do not significantly differ at different points of measurement. Most of the detected chemical elements on the sample surfaces on both sides have an accidental origin and lead to less to slow influence on the required values of the electron work function. The latter is also due to the spherical forms of the sensor element.

185

Table 3. Surface atomic concentrations obtained by XPS for samples R, S1, and S2

Sample	C [at.%]	O [at.%]	N [at.%]	Na [at.%]	Zn [at.%]	Si [at.%]	S [at.%]	Ca [at.%]	Al [at.%]	Ni [at.%]
R-front	85.4	11.2	2.7	0.8	0.0	0.0	0.0	0.0	0.0	0.0
R-back	80.7	13.3	2.8	0.7	0.2	1.5	0.8	0.0	0.0	0.0
S1-front	84.30	13.0	0.80	0.20	0.10	0.90	0.70	0.10	0.00	0.00
S1-back	90.70	7.60	1.00	0.60	0.20	0.00	0.00	0.00	0.00	0.00
S2-front	87.60	9.70	1.40	0.20	0.10	0.40	0.00	0.60	0.00	0.00
S2-back	85.70	10.8	0.80	0.30	0.10	0.00	0.00	0.50	1.20	0.20

Summarizing and comparing the results for the WF (table 1) regarding the "reference" samples R- front and R-back side of the "space" S1 and S2, we can conclude that the values between 4.4 and 4.8 are close-range values and their insignificant difference lays on the impurities and contaminations of the surfaces. In this sense, and due to the importance of material WF used as a major parameter in the design of devices, we can conclude that the investigated glassy carbon can be used as a stable, sensitive element in electronic detectors.

4. Conclusions

Detailed investigations on the electron work-function variations due to ionospheric-magnetospheric plasmas field have been carried out. The tested materials were glassy-carbon coatings obtained by the original method. A specific approach has been used to calculate these values for exposed and non-exposed samples. All measurements state similar values of the WF being between 4.4 eV and 4.8 eV. The discrepancy suggests that the GC coatings are stable against aging processes induced by energetic irradiations, which generally lead to severe crystalline defects. Moreover, the surface contaminations do not lead to noticeable WF fluctuations. Probes of glassy-carbon performed as chemically and mechanically stable and sensitive elements are used as double probes for satellite electric field measurements. The higher native values of the WF up to some 5 eV are a prerequisite for lower photoemission. The latter phenomenon increases the coating's electrical conductivity, which results in higher sensitivity to low-intensity electromagnetic fields.

References:

1. Fahleson, U., Theory of electric field measurements conducted in the magnetosphere with electric probes, Space Science Reviews, 1967, 7, 238.
2. Boyd, R. L., Measurement of electric field in the ionosphere and magnetosphere, Space Science Reviews, 1967, 7, 230.
3. Fahleson, U. V., M. C. Kelly and F. S. Mozer, Investigation of the operation of a DC electric field detector, Planet. Space Sci., 1970, 18, 1551–1561.
4. Pedersenl, A., C. A. Cattell, C.-G. falthammar, V. Formisano, P.-A. Lindqvist, F. Mozer, R. Torbert, Quasistatic electric field measurements with spherical double probes on the geos and isee satellites, Space Science Reviews, 1984, 37, 269–312.
5. Tsuruda, K., H. Hayakawa, M. Nakamura, T. Okada, A. Matsuoka, F. S. Mozer and R. Schmidt, Electric Field Measurements on the GEOTAIL Satellite, J. Geomag. Geoelectr, 1994, 46, 693–711.
6. Harvey, P., Mozer, F. S., Pankow, D., Wygant, J., Maynard, N. C., Singer, H., Sullivan, W., Anderson, P. B., Pfaff, R., Aggson, T., Pedersen, A., F¨althammar, C.-G., and Tanskanen, P., The electric field instrument on the Polar satellite, Space Science Reviews, 1995, 71, 583–596.
7. Eriksson, A. I., M. Andre, B. Klecker, H. Laakso, P.-A. Lindqvist, F. Mozer, G. Paschmann, A. Pedersen, J. Quinn, R. Torbert, K. Torkar, and H. Vaith, Electric field measurements on Cluster: comparing the double-probe and electron drift techniques, Annales Geophysicae, 2006, 24, 275–289.
8. Berthelier, J. J., M. Godefroy, F. Leblanc, M. Malingre, M. Menvielle, D. Lagoutte, J. Y. Brochot, F. Colin, F. Elie, C. Legendre, P. Zamora, D. Benoist, Y. Chapuis, J. Artru, R. Pfaff, ICE, the electric field experiment on DEMETER, Planetary and Space Science, 2006, 54, 456–471.
9. Bonnell, J. W., F. S. Mozer, G. T. Delory, A. J. Hull, R. E. Ergun, C. M. Cully, V. Angelopoulos, P. R. Harvey, The Electric Field Instrument (EFI) for THEMIS, Space Science Reviews, 2008, 141, 303–341. DOI 10.1007/s11214-008-9469-2
10. Bouzekova-Penkova, A., S. Klimov, V. Grushin, O. Lapshinova, D. Novikov, D. Teodosiev, Space Experiment "Obstanovka (1 - stage)", block DP - PM of the Russian segment of the International Space Station (ISS), Aerospace Research in Bulgaria, 2023, 35, 156–164, ISSN 1313-0927. https://doi.org/10.3897/arb.v35.e15
11. Shirley, D., High-Resolution X-Ray Photoemission Spectrum of the Valence Bands of Gold, Phys. Rev. B, 1972, 5, 4709–4714.
12. Scofield, J. H. Hartree-Slater subshell photoionization cross-sections at 1254 and 1487 eV. J. Electron Spectrosc. Relat. Phenom. 1976, 8(2), 129–137.
13. Jeong Won Kim, Ansoon Kim, Absolute work function measurement by using photoelectron spectroscopy, Current Applied Physics, 2021, 31, 52–59. https://doi.org/10.1016/j.cap.2021.07.018
14. Kelemen, S. R., & Freund, HXPS characterization of glassy-carbon surfaces oxidized by O_2, CO_2, and HNO_3. Energy & fuels, 1988, 2(2), 111–118.
15. Brunetti, B., De Giglio, E., Cafagna, D., & Desimoni, E., XPS analysis of glassy carbon electrodes chemically modified with 8-hydroxyquinoline-5-sulphonic acid, Surface and interface analysis, 2012, 44(4), 491–496.

16. Skriver, H. L., N. M. Rosengaard, Surface energy and work function of elemental metals, Phys. Rev. B, 1992, 46, 7157–7168.
17. Methfessel, M., D. Hennig, M. Scheffler, Trends of the surface relaxations, surface energies, and work functions of the 4d transition metals, Phys. Rev. B, 1992, 46, 4816–4829.
18. Ivey, H. F., Phys. Rev., 1949, 76, 567.
19. Singh-Miller, N. E., N. Marzari, Surface energies, work functions, and surface relaxations of low-index metallic surfaces from first principles, Phys. Rev. B, 2009, 80, 235407.
20. Teodosiev, D. K., I. D. Pechenykov, J. S. Georgiev, P. A. Petrov, R. P. Dobrev, R.V. Valov, G.A. Stanev, Method for impregnation and coating of porous refractory materials with glass-carbon, Bulgarian Certificate of Authenticity № 31847, МКП С 04 35/84, 1982.

XPS ИЗСЛЕДВАНЕ НА ОТДЕЛИТЕЛНАТА РАБОТА НА СТЪКЛОВЪГЛЕРОДНИ ПОКРИТИЯ СЛЕД ПРОДЪЛЖИТЕЛЕН ПРЕСТОЙ НА МЕЖДУНАРОДНАТА КОСМИЧЕСКА СТАНЦИЯ (МКС)

А. Цанев, К. Григоров, Х. Колев, А. Бузекова-Пенкова, Б. Цинцарски, П. Цветков, Д. Теодосиев

Резюме

С помощта на XPS са изследвани стойностите на отделителната работа на повърхността на стъкловъглеродни покрития върху образци след продължителен престой на борда на МКС в условията на открития Космос. Резултатите са сравнени с характеристиките на стъкловъглеродни покрития върху образци, престояли на Земята за същия период. Установено е, че отделителната работа не се променя съществено, и това доказва и потвърждава възможностите за успешно прилагане на тези стъкловъглеродни покрития, получени по оригинална българска технология, за космически експерименти на борда на спътници за измерване на електрични полета в йоносферно-магнитосферната плазма. Минималните вариации на стойностите на отдели-телната работа се обяснява с малка разлика в съдържанието на следи от различни химични елементи на повърхността на покритията. Установено е, че стъкловъглеродните покритията са със стабилни характеристики след дълъг престой в Космоса, въпреки малките колебания в стойностите на отдели-телната работа на електрона. Резултатите показват, че стъкловъглеродните покрития са химически и механически стабилни. Получените резултати на този оригинален технологичен експеримент са уникални за създаване на чувствителни елементи – сензори за измерване на слаби електрични полета в околоземната плазма.

188

IMPACT OF SPACECRAFT AND ISS ENGINES ON EXPERIMENTAL SAMPLES OF VARIOUS MATERIALS MOUNTED ON THE OUTER SURFACE OF THE "ZVEZDA" MODULE

*Peter Getsov, Anna Bouzekova-Penkova, Nikolay Zagorski,
Garo Mardirossian, Dimo Zafirov*

*Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: getsovp@space.bas.bg, a_bouzekova@space.bas.bg*

Keywords: Spaceships, Engines, Gas jet, Experimental samples

Abstract
The work presents a study of the influence of the spacecraft and ISS engines during the correction of the station's orbit. The zones of action of the gas jets and their impact on the environment and the surfaces of the experimental samples of various materials mounted outside the "Zvezda" module were determined.

Introduction

The development of scientific equipment for the exploration of outer space on board satellites and space stations requires the creation and testing of new materials that can withstand the stresses of operating in the conditions of open space, such as high and low temperatures, vacuum, exposure to ionizing radiation, microparticle flows, etc. All these outer space factors adversely affect the materials from which spacecraft and scientific equipment are made. As a result, various physico-chemical and mechanical processes occur in the materials used, leading to a deterioration of their parameters. Depending on the nature of the processes caused by the impact of the space environment, near the board of satellites and orbital stations, with scientific equipment mounted on them, the surface properties of materials can change, leading to reversible and irreversible processes.

To acquire complete and real data on the effects of these processes, a technological experiment was planned and carried out in cooperation with a leading country in this industry - the Russian module of the International Space Station (ISS), in whose division the team with the block "DP- PM" as part of the "Obstanovka 1-step" experiment. A "DP-PM" block is a container measuring 185 mm × 70 mm × 10 mm (Fig.1) [1].

Method

The experiment to study the influence of the space environment on the structure and properties of various samples was conducted from 04.19.2013 until 10.08.2015. The operation of the "DP-PM" block from its placement on the outside of the Zvezda module until its delivery to Earth is carried out following the documents of Product 17KS #12801. P42422.

Fig.1. Block "DP-PM"

The experiment's purpose is to study surface degradation and structural changes occurring in graphite samples densified and coated with glassy carbon (GC) during prolonged stays in outer space conditions [1,2].

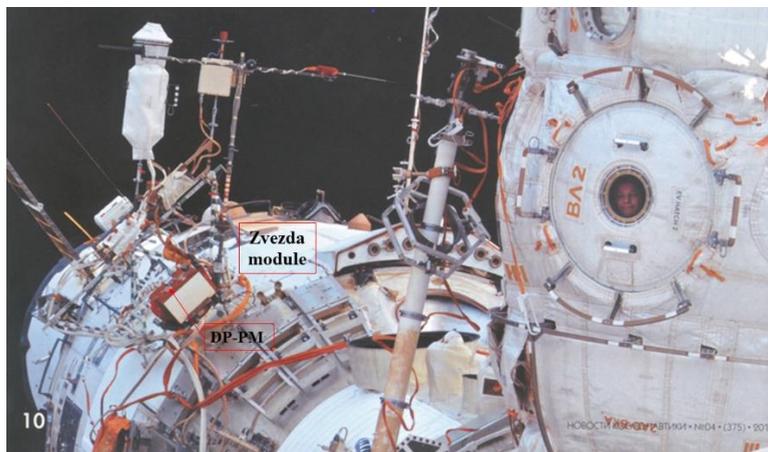Fig. 2 shows the exact location of the "DP-PM" block on the surface of the ISS [3].

Fig. 2. The exact location of the "DP-PM" block on the surface of the ISS [3]

After completion of the planned period, the "DP-PM" block is dismantled from the outer surface of the ISS and the cosmonauts place it in a work bag (Fig. 3a). Then, in the transition compartment of the Zvezda module, the cosmonauts put it in a special soft container (Fig. 3b). In preparation for returning to Earth with the Soyuz-TMA spacecraft No. 716, the container was additionally placed in an airtight transport package (Fig. 3c). In this type, the "DP-PM" block was handed over for conducting research at the Space Research and Technologies Institute at the Bulgarian Academy of Sciences (SRTI–BAS).
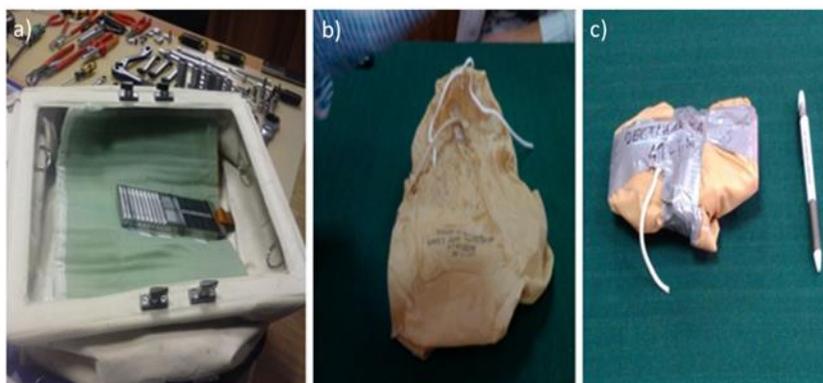
Fig.3. Different types of packaging for transporting a DP-PM block:
a) Work bag; b) Special soft container; c) airtight transport package

Upon external inspection of the samples (Fig. 4), it was found that their surfaces are heavily blackened, which suggests an atypical impact on them in the conditions of the space vacuum and weightlessness of this orbit. Our working hypothesis was that the condition of the surface of the samples was due to the impact of the gas jets from the various engines used in the control and correction of the ISS, both during the docking of the ships coming from Earth and their separation from the ISS.

Fig. 4. Appearance of the samples from block "DP-PM" during research at SRTI–BAS

After March 2011, a total of 15 pressurized modules were added to the ISS: Zarya, Zvezda, Destiny, Unity (Node1), Harmony (Node2), Tranquility (Node3), Columbus, Leonardo, Japanese, Quest, Cupola, Poisk, Rassvet and Pierce. To these modules we can add all three pressurized docking adapters. We note that after the completion of the flights of the "Space Shuttle" program, the orbit of the station was raised from 350 km to 390–420 km.

In order to study the influence of the gas jets of the engines on the samples, in addition to their location, also the modes of docking (de-docking), control and corrections of the ISS by the ships and the Zvezda module were taken into account.

The scheme presented below (Fig. 5) depicts the main components of the station and the ships arriving (departing) to it, as well as the location of the main and correction engines (red arrows) [4].
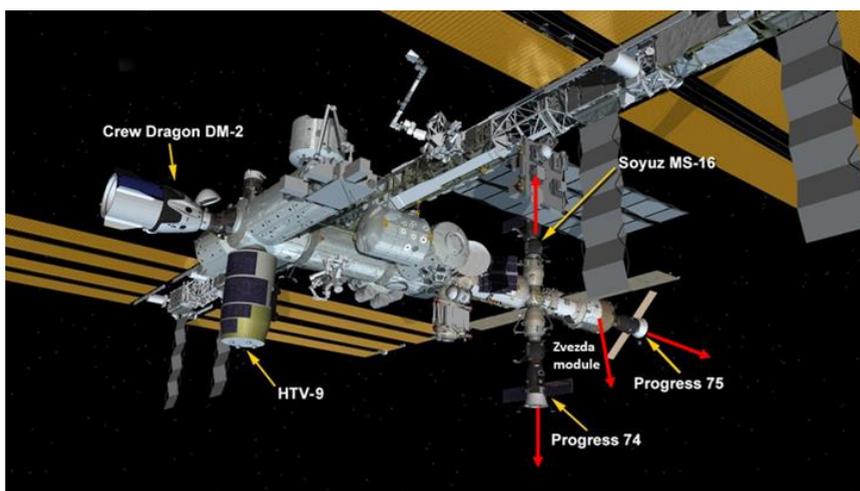

Fig. 5. Schematic of the main components of the ISS and the spacecraft docked to it [4]

The height of the station's orbit is constantly changing due to the influence of solar activity and friction with the rarefied atmosphere, which leads to the gradual deceleration of the movement and loss of height. Atmospheric drag lowers the altitude on average by about 2 km per month [5].

The graph of the change in the height of the ISS, including the period of the experiment, is shown in Fig. 6 [6].

The ISS's orbit is adjusted several times a year to compensate for friction, to avoid larger and larger pieces of space junk, and for other reasons. Orbit correction is done using the ISS's own engines. Until 2000, the engines of the Zarya functional cargo model were used for this purpose, and then - of the Zvezda service model. The engines of the arriving transport ships are also used for orbit correction, which

also refuel the ISS [6]. Up to a certain period of time, the correction in the orbit only compensates for the decrease in altitude.
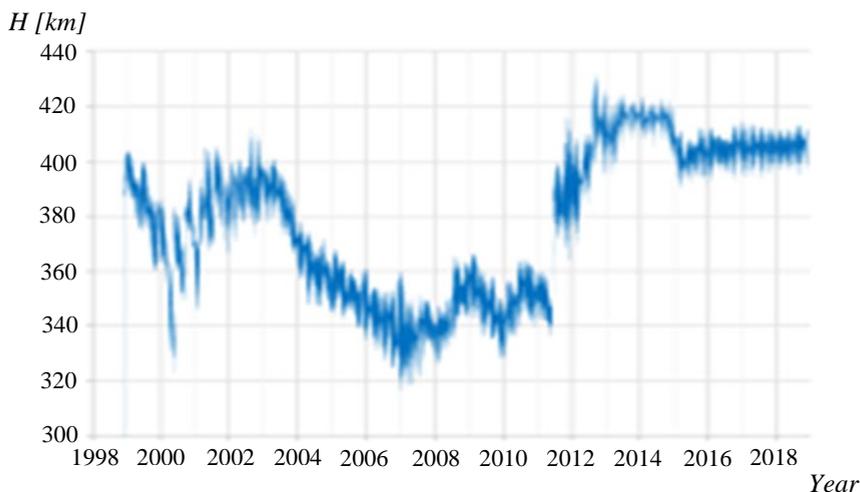

Fig. 6. Diagram showing the change in height of the ISS
for the period from 1998 to 2018 [6].

Research shows that using Space Shuttle craft to maintain a 350 km high orbit of the ISS required 8 600 kg of fuel for one year. When increasing the height of the station's orbit to 400 km, the required fuel is 3 600 kg. During this period, the fuel for control and orientation was 1 926 kg [7]. The increased altitude allows for a substantial reduction in the amount of fuel required to maintain this orbit. At the same time, it enables increased supplies of water and food products, as well as other payloads. The ISS completes a complete orbit around the Earth in about 92 minutes. From launch to February 5, 2023, the station has completed 138 805 complete orbits.

During the experiment, the following spacecraft were flown to the ISS and docked with the station: Soyuz, Progress, and Cygnus. The results are shown in Tables 1, 2, and 3.

Table 1. Soyuz spacecraft flights

By №	Mission Name	Board number	Launch date	Date of landing
141	Soyuz TMA-09M	709	05/29/2013	11/11/2013
142	Soyuz TMA-10M	710	09/26/2013	03/11/2014
143	Soyuz TMA-11M	711	11/07/2013	05/14/2014
144	Soyuz TMA-12M	712	03/26/2014	09/11/2014
145	Soyuz TMA-13M	713	05/28/2014	11/10/2014
146	Soyuz TMA-14M	714	09/26/2014	03/12/2015
147	Soyuz TMA-15M	715	11/24/2014	06/11/2015

Table 2. Progress spacecraft flights

By №	Mission Name	Board number	Launch date	Date of landing
142	Progress M-19M	419	04/24/2013	06/19/2013
143	Progress M-20M	420	07/28/2013	02/11/2014
144	Progress M-21M	421	11/26/2013	06/09/2014
145	Progress M-22M	422	02/05/2014	04/18/2014
146	Progress M-23M	423	04/09/2014	08/01/2014
147	Progress M-24M	424	07/24/2014	10/27/2014
148	Progress M-25M	425	10/29/2014	04/26/2015
149	Progress M-26M	425	02/17/2015	08/14/2015
150	Progress M-27M	426	04/28/2015	05/08/2015
151	Progress M-28M	428	07/03/2015	12/19/2015

Table 3. Cygnus spacecraft flights

By №	Mission Name	Spacecraft name	Launch date	Note
1	Cygnus Orb-D1	George Lowe	09/18/2013	First Cygnus spacecraft mission. The first Cygnus docking mission with the ISS was delayed due to a computer problem but was successful.
2	Cygnus Orb-CRS-1	Charles Fullerton	01/09/2014	First logistics mission of a Cygnus spacecraft.
3	Cygnus Orb-CRS-2	Janice Voss	07/13/2014	
4	Cygnus Orb-CRS-3	Donald Slayton	10/28/2014	During the launch attempt, there is a catastrophic anomaly, and the launch vehicle explodes.
5	Cygnus Orb-CRS-4		11/19/2015	The first mission of an enhanced version of Cygnus.

Figure 7 shows the structure of the Soyuz transport ship and the location of the engine bay [8].

The propulsion module of the Progress spacecraft, like that of the Soyuz, is located at the rear of the spacecraft. The module consists of the ship's main engine and navigation engines used for automatic docking with the ISS and for correcting the station's orbit when the ship is docked with it. On these ships, only one set of orientation engines type OE (8 units) with a thrust of 1.5 kgf, a KTDU-35 brake-engine unit with a thrust of 417 kgf and 14 engines for docking and orientation (EDO) with a thrust of 10 kgf were installed. The fuel for KTDU-35 is asymmetric dimethylhydrazine (ADMH) and oxidizer-nitric acid. EDO and OE work with hydrogen peroxide [8].
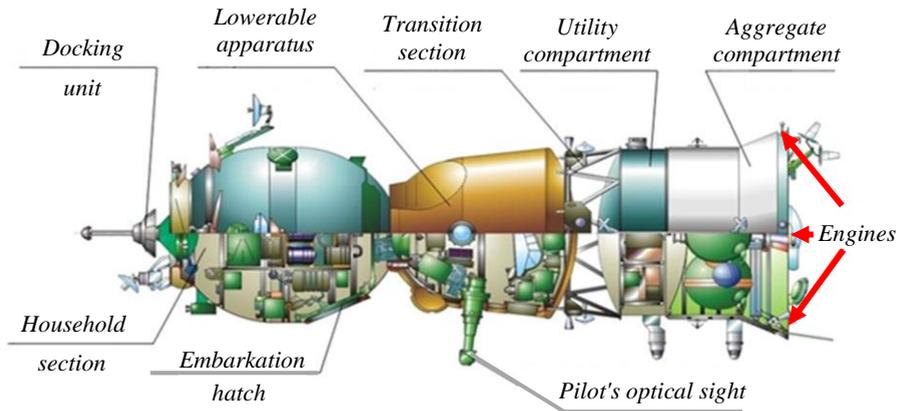
194

Fig. 7. Construction of the "Soyuz" transport spacecraft [9]

The combined propulsion system, which is used in many cases for orbit correction and control of the ISS [9] consists of two 2×312 kgf thrusters, 32×13.3 kgf thrusters for orientation, 302 kg of propellant (ADMH) and 558 kg of oxidizer (nitrogen tetraoxide).

The Cygnus spacecraft consists of two main components - a service module and a cargo module. The service module has a gross weight of about 1 800 kg and is powered by Japanese rocket engines (BT-4 of the Aerospace company) running on ethylhydrazine and dinitrogen tetroxide [10]. Each motor develops a thrust of 46 kgf (Fig. 8).

Fig. 8. Cygnus spacecraft approaching the International Space Station [10]

After detaching aboard the Antares rocket, Cygnus approaches the ISS, and when within a few meters of the station, the craft is engaged by the Canadarm-2 mechanized arm and docked with the Harmony module. Cygnus

does not have the ability to return cargo to Earth, but it can be filled with debris to burn up with it as it enters Earth's atmosphere.

From the beginning of the station's creation until October 2022, the ISS's orbit has been corrected 327 times, 176 times with the engines of the transport spacecraft Progress, which shows the importance and frequency of the correction.

For the period of the positioning of the samples on the outer surface of the Zvezda module (28 months), a total of about 4 550 kg was used (spacecraft Progress – 3 248 kg; modules Zvezda, Cygnus, etc. about 1 302 kg) for correcting the orbit of the ISS for control necessary for avoiding space debris and preparing the station for docking and undocking of arriving and departing spacecraft.
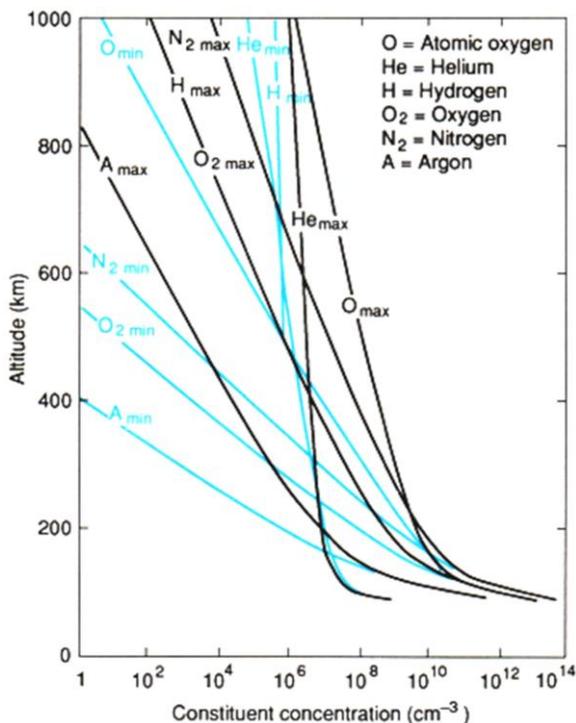
Fig. 9. Concentration of the components of the space environment
at altitudes from 100 to 1000 km

As noted, the fuel used is unsymmetrical dimethylhydrazine ADMH, codenamed "heptyl", which is a high-boiling component of rocket fuel (boiling point above 0°C). Nitrogen tetroxide (AT) is used as an oxidizer of ADMH, either pure or in a mixture with nitric acid. Cases using pure acid and liquid oxygen are known. It can be used in a mixture with hydrazine, known as aerosine [9,11] to improve its properties. It self-ignites upon contact with nitric acid and dinitrogen tetroxide-based

oxidizers, which simplifies construction and provides easy starting and re-engagement of rocket engines. ADMH is thermally stable up to +350°C.

In the range from +350 to +1000°C, dimethylhydrazine decomposes into ammonia, amines, hydrocyanic acid, hydrogen, nitrogen, methane, ethane, resinous, and other substances. At increased temperature, it decomposes into nitrogen and hydrocarbons with the release of heat, an increase in the volume, and the creation of jet force from the nozzles.

Taking into account that the impact process on the surface of the spacecraft takes place even with very small amounts of atoms of the elements at different altitudes of the trajectory of the spacecraft (Fig. 9), (well systematized and summarized in [12], where cited and the results of the team from SRTI–BAS [2,13]), we can conclude that in the given case there are far more intense impact processes taking place, albeit short-lived, as a result of the spread of gases from the engines.

The gas jets coming out of the engines' nozzles spread almost instantaneously in the three-dimensional space, which is determined by the practical vacuum around the station (Fig. 10). Depending on the spatial position of the station and the direction of the change in the speed of the ISS, as well as from the engines that are used, the studied samples are subjected to an impact, causing processes (oxidation) on their surfaces.
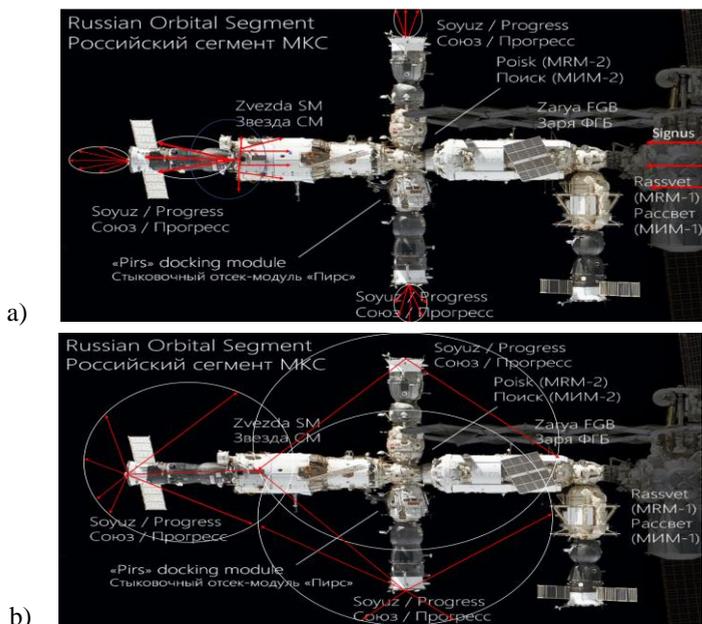

a)

b)

Fig. 10. Propagation of the gas jets from the engine nozzles
in the three-dimensional space around the ISS [14]

In fact, contrary to the initial impression that the gases from the exhaust gases from the engine nozzles always remain behind the board of the ISS, thanks to their relative motion, they spread partially on the surface of the station, which is also the reason for the condition of the surfaces of the examined samples (Fig. 11).
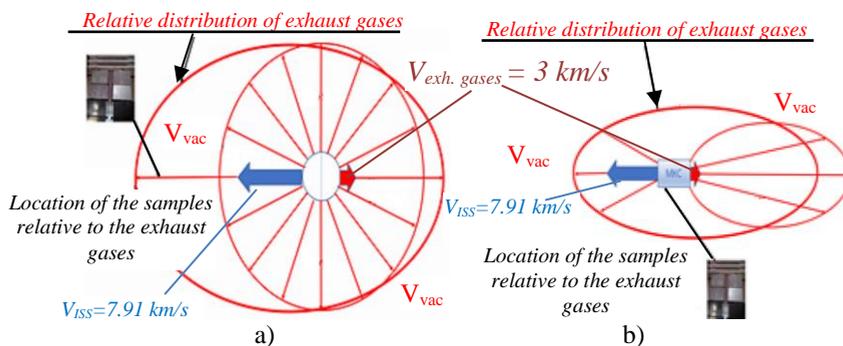

Fig. 11. Different flow of the samples, depending on the position
of the motors used for correction

Conclusion

It can be assumed that the minimal changes on the surface of GC samples, "front", is due precisely to the large amount of fuel used during the experiment, which turned into gases from the nozzles of the engines of the ships Progress, Soyuz and Cygnus during their docking (undocking) and correction and control of the Zvezda module in ISS orbit. These assumptions are also confirmed by the conducted physico-chemical studies of the surfaces of GC samples (respectively "front" – "backside"). Differences are observed between the "front" and "back-side" glassy carbon coating. Thinning in the glassy carbon "front" coating layer is due to the different impact and oxidation conditions from the gas jets determined by their location.

References

1. Bouzekova-Penkova, Anna, Stanislav Klimov, Valery Grushin, et al. Space Experiment "Obstanovka (1 - stage)", block DP - PM of the Russian segment of the International Space Station (ISS), Aerospace Research in Bulgaria, 2023, 35, 156–164. DOI: https://doi.org/10.3897/arb.v35.e15

2. Teodosiev, D., Anna Bouzekova-Penkova, Korneli Grigorov, et al. Structural and mechanical properties of glass-carbon coatings after an extended stay on the International Space Station (ISS), Comptes rendus de l'Acade'mie bulgare des Sciences, 2021, 74, 2, 197–206. DOI:10.7546/CRABS.2021.02.05

3. Cosmonautics News, the magazine is published under the auspices of the Federal Space Agency and Aerospace Defense Forces, 2014, 4(375), 10. (in Russian)

4. https://habr.com/en/news/516036/

5. https://eol.jsc.nasa.gov/Tools/orbitTutorial.htm

6. https://en.wikipedia.org/wiki/International_Space_Station#Orbit

7. Kauderer Amiko. Higher altitude improves staton's fuel economy NASA (14 february 2011)

8. https://triptonkosti.ru/23-foto/soyuz-kosmicheskij-korabl-shema.html

9. https://ru.wikipedia.org/wiki/%D0%9C%D0%B5%D0%B6%D0%B4%D1%83%D0%BD%D0%B0%D1%80%D0%BE%D0%B4%D0%BD%D0%B0%D1%8F_%D0%BA%D0%BE%D1%81%D0%BC%D0%B8%D1%87%D0%B5%D1%81%D0%BA%D0%B0%D1%8F_%D1%81%D1%82%D0%B0%D0%BD%D1%86%D0%B8%D1%8F#%D0%9A%D0%BE%D1%80%D1%80%D0%B5%D0%BA%D1%86%D0%B8%D1%8F_%D0%B2%D1%8B%D1%81%D0%BE%D1%82%D1%8B_%D0%BE%D1%80%D0%B1%D0%B8%D1%82%D1%8B

10. https://en.wikipedia.org/wiki/Cygnus_%28spacecraft%29

11. https://ru.wikipedia.org/wiki/1,1%D0%94%D0%B8%D0%BC%D0%B5%D1%82%D0%B8%D0%BB%D0%B3%D0%B8%D0%B4%D1%80%D0%B0%D0%B7%D0%B8%D0%BD#cite_ref-ecosm_5-0

12. Charles E. Vest the effects of the space environment on spacecraft surfaces, Papers Space Materials 07 2022.zip.

13. Tsyntsarski, Boyko, Anna Bouzekova-Penkova, Urszula Szeluga, et al. SEM and EDX study of glassy carbon coatings after an extende stay on the international space station (ISS), Aerospace Research in Bulgaria, 2023, 35, 165-173. DOI: https://doi.org/10.3897/arb.v35.e16

14. https://commons.wikimedia.org/wiki/File:Russian_Orbital_Segment.png

ВЛИЯНИЕ НА РАБОТАТА НА ДВИГАТЕЛИТЕ НА КОСМИЧЕСКИТЕ КОРАБИ И НА МКС ВЪРХУ ОПИТНИТЕ ОБРАЗЦИ ОТ РАЗЛИЧНИ МАТЕРИАЛИ, МОНТИРАНИ ВЪРХУ ВЪНШНАТА ПОВЪРХНОСТ НА МОДУЛА „ЗВЕЗДА"

П. Гецов, А. Бузекова-Пенкова, Н. Загорски, Г. Мардиросян, Д. Зафиров

Резюме

В работата е представено изследване на влиянието на двигателите на космическите кораби и Международната космическа станция при корекция на височината на орбитата на станцията. Определени са зоните на действие на газовите струи и тяхното въздействие на околната среда и повърхностите на експерименталните образци от различни материали, монтирани на външната страна на модула „Звезда".

__Reviews__

REVIEW OF THE APPLICATIONS OF SATELLITE REMOTE SENSING IN ORGANIC FARMING – PART II

Milen Chanev, Lachezar Filchev

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

Abstract
The use of remote sensing methods for monitoring, managing, and decision support in agriculture is increasingly intensifying. With the advancement of technologies, they become more accessible, while the quality and security of the obtained data are improving. To improve the quality of the environment and its preservation, expanding the areas occupied by organic farming will allow us to achieve these goals. At the same time, this type of agriculture provides healthy and safe food. For this reason, it is essential to start applying satellite data in organic farming as quickly as possible. In Part II of the "Review of the applications of satellite remote sensing in organic farming," we examine the various areas of satellite data application in organic farming. Five areas of satellite data application in organic farming have been identified, including satellite remote sensing monitoring of weeds, remote sensing of crop stress and irrigation needs, yield forecasting using remote sensing methods, and remote sensing monitoring of plant nutrition. From the conducted review, we found that satellite data can significantly support and facilitate the transition to organic farming, adequate fertilization, application in phytosanitary monitoring of crops, and assessment of crop stress.

Introduction

In the first part of the "Review of the applications of satellite remote sensing in organic farming," published in issue 35, we examined the main applications of remote methods, the use of satellite data, and the potential of uncrewed aerial vehicles (UAVs) platforms in agriculture, and their possible application in organic farming. We analyzed publications on the topic published in specialized databases. Our analysis revealed that the top 25 most cited publications, indexed in CrossRef and parsed by the Scite_ platform (https://scite.ai accessed on 06.02.2024), were published by Elsevier, followed by AAAS, Wiley, and MDPI. Further analysis of the literature on the issue showed that the most cited publications are from 2010, 2011, and 2013, indicating that the interest in satellite data applications in organic

farming is relatively recent, dating back approximately 14 years. We also analyzed using the open web tool WordItOut (https://worditout.com accessed on 06.02.2024). The word cloud mapped 141 words out of 301 unique words. Other commonly occurring words include "image," "plant," "vision," "resolution," "spectral," "surface," and "grain," which more or less reveal the main objectives of the studies in question.

Materials and methods

This study provides an overview and a comprehensive analysis of articles, reports, and materials published online in the scientific databases Scopus, ResearchGate, and Google Scholar. A combination of keywords with logical queries was used when searching the scientific and specialized database from the beginning of space remote sensing from the late 1980s to 2021. The main keywords that we used are: "organic farming" AND "remote sensing", "organic farming" AND "satellite data". The results obtained were categorized into five application categories of satellite data in organic farming. The separated application categories are satellite remote sensing monitoring of weeds, remote sensing of crop stress and the need for irrigation, forecasting yields using remote sensing methods, and remote sensing monitoring of plant nutrition.

Results and discussion

Satellite remote sensing monitoring of weeds

Weed control on organic farms is one of the main problems facing organic farming. Remote sensing monitoring systems can help overcome the problems caused by weeds. This can be achieved by integrating drones carrying different sensors and satellite images that include hyperspectral, multispectral, and RGB in combination with artificial intelligence, guaranteeing a better result in weed management [1].

Information on the distribution of weeds in the field is necessary to compile an assessment map of crops to determine the achievement of their biological threshold of harmfulness. Perez et al. (2000) [2] propose two approaches to automatic weed monitoring:

- Rough identification of weeds in the observed areas by remote sensing.

- Fine identification using proximal methods, such as video imaging and image analysis, should confirm the location and allow the most appropriate local treatment of the crop to be selected.

A review of the potential of remote sensing techniques for crop protection suggests that one way to differentiate between weeds and crops is by studying temporal patterns of plant indices during the growing season [2, 3]. In addition, using remote sensing methods, only a few species of weeds in different phenophases can

usually be distinguished. The spectral characteristics of weeds should be taken from populations in groups of weeds identified shortly before the detection process, as characteristics are highly variable and depend on the phenophase of weeds or weed associations [2, 4]. Two approaches are usually used for automatic weed monitoring. One is to establish geometric differences in leaf litter between crops and weeds [5–10]. The second approach is based on differences in spectral reflectivity. It is possible to look for differences in the location of the crop and the weeds in the crop [2, 11]. Guyer [5] explored the possibility of using leaf shape for plant identification. The team led by Franz [12] used local spectral characteristics of plant leaves to distinguish different weeds. Zhang and Chaisattapagon (1995) [13] use machine vision to identify weeds in wheat fields. They apply three approaches to distinguish them from cultivated plants: color analysis, shape analysis, and texture analysis. In the laboratory, they use black-and-white digital images with various color filters. They found that red and green filters effectively detect reddish stems in some weeds. Leaf-blade parameters are adequate for distinguishing deciduous weeds from cereal leaves. Another way to differentiate young crops and weeds is by analyzing the spectral reflectivity using specific wavelengths ranging from 200 to 2000 nm [14, 15]. We can say that the opportunities offered by RS are significant for weed management in organic farms, as the use of herbicides is prohibited there.

Remote sensing of crops' stress and the need for irrigation

The RS provides an excellent opportunity to assess stress, and these methods are also used to calculate different vegetation indices that estimate different crop parameters [16-18]. Remote sensing methods in agriculture can be a powerful technique for visualizing, diagnosing, and quantifying the crop response to stress caused by abiotic, biotic factors or improper management decisions. In most cases, stress leads to deviations in plants' pigmentation, which can be used to diagnose stress in crops using RS methods [19, 20]. The main problem facing conventional and organic farming is the quantification of crop water consumption and the water stress they experience. Using satellite-based images and the computational processing of satellite images is an opportunity to manage water stress on crops in organic farming [21, 22]. Sharma [23] proposed that the irrigation problem in organic farms be managed using RS and data processing using a Support Vector Machine (SVM). Data is stored in computer software sent to SVM to determine the status of crops on organic farms. The information is sent to the user interface, where the farmer receives farm information using Google's mobile communication module. This image-processing method allows farmers to take preventive action to save crops. Solaiman and Salaheen [22] note that there has been an increase in interest in using satellite and RS for organic farming, with the primary goal being to establish the health status of farms. Vroege [24] proposed that RS technologies be used to compile drought risk assessment maps based on satellite soil moisture data. These cards are used to limit financial risk for farmers affected by drought.

Forecasting of yields using remote sensing methods

Crop production is perhaps the most critical information for crop management in precision agriculture. A big problem is that yield data are obtained after the season, while issues such as nutrient deficiencies, water stress, or pest infestation must be managed during the growing season. Images from satellites or UAVs obtained during the growing season have the potential for post-season and in-season management. In addition, yield maps composed of images obtained by DM can be used as an alternative when yield monitor data are unavailable [19, 25]. Because yield does not affect the reflectivity of the crop, it is derived indirectly from other biophysical parameters of the crop [25, 26]. It was ultimately found that this relationship can only be explained implicitly by biophysical and biochemical characteristics, even though it may suggest that the determination of cereal yield directly from reflection spectra is statistically feasible [26–30]. From RS data, chlorophyll content (CHL) and leaf area index (LAI) may be accurately determined and linked to yield [26, 31–33]. In addition, soil mineral nitrogen naturally has a strong effect on plant development and, thus, on chlorophyll production and leaf area, so these parameters can be considered indicators of nitrogen uptake in plants [34]. LAI is one of the most important parameters for describing plant conditions in agriculture. It can be used, for example, to obtain information on biomass, nutrient supply, growth stage, and yield assessment [35]. Many studies have been published on assessing LAI using hyperspectral data from remote sensing [26]. According to [26], LAI and CHL forecasts from UAV-based hyperspectral data for yield forecasting are promising.

Remote sensing monitoring of plant nutrition

Efficient food production requires a balance between minimizing environmental damage and maximizing yields [36]. From the point of view of the agricultural producer, the most critical economic parameter is the achieved yields. Generous application of nitrogen fertilizers within legal limits leads to higher costs without added value in additional yield. In addition, new concepts for monitoring these effects during vegetative growth allow the development of precision farming applications specifically designed for efficient N fertilization [37]. RS will support decision-making regarding plant nutrition in organic farming by collecting and analyzing data, formulating specific management recommendations, and implementing management practices to correct factors that limit crop growth, productivity, and quality [37-41]. Sozzi [42] found that it is most cost-effective to use the vegetation index NDVI generated from satellite images with medium-resolution satellite data with good optical properties and high-resolution satellites with lower optical quality to determine the needs of H fertilization in the fields. In combination with GIS and using different vegetation indices such as NDVI, remote

sensing methods make it possible to map different fertilization rates in the field [37, 43].

Conclusions

In conclusion, it should be said that remote sensing methods can significantly support and facilitate the transition to organic farming. RS can adequately manage fertilization in organic agriculture. RS can signal the phytosanitary state promptly and assess crop stress on organic farms and in weed management. Using the capabilities of RS, organic farmers can forecast yields, which helps them to plan their costs and profits better and thus increase the sustainability of their organic farm. With the help of RS, biodiversity in agriculture can be monitored. They provide a perfect opportunity to remotely distinguish biological fields from conventional fields, which can serve state organizations and certification bodies to control organic farmers to comply with the rules and norms of organic farming.

Acknowledgments

Conflicts of interest

The authors reported no potential conflicts of interest.

References

1. Roslim MHM, Juraimi AS, Che'Ya NN, Sulaiman N, Manaf MNHA, Ramli Z, Motmainna M. (2021). Using Remote Sensing and an Unmanned Aerial System for Weed Management in Agricultural Crops: A Review. Agronomy. 2021; 11(9):1809. https://doi.org/10.3390/agronomy11091809
2. Perez, A., Lo´pez F., Benlloch J. V., and Christensen S. (2000). Colour and shape analysis techniques for weed detection in cereal fields. Computers and Electronics in Agriculture. 25: 197–212.
3. Hatfield, P. L. and Pinter P. (1993). Remote sensing for crop protection. Crop Protection. 12(6): 403–413. doi:10.1016/0261-2194(93)90001-y.
4. Brown, R.W., Steckler J. P. G. A., and Anderson G. W. (1994). Remote sensing for identification of weeds in no-till corn. Trans. ASAE. 37(1): 297–302.
5. Guyer, D. E., Miles G. E., Schreiber M. M., Mitchell O. R., and Vanderbilt V. C. (1986). Machine vision and image processing for plant identification, Trans. ASAE. 29(6): 1500–1507. doi:10.13031/2013.30344
6. Shearer, S. A. and Holmes R. G. (1990). Plant identification using color co-occurrence matrices. Trans. of ASAE. 33(6): 2037–2044.

7. Meyer, G. E., Mehta T., Kocher M. F., Mortensen D. A., and Samal A. (1998). Textural imaging and discriminant analysis for distinguishing weeds for spot spraying. Transactions of the ASAE. 41(4): 1189–1197.

8. Ahmad, U., Kondo N., Arima S., Monta M., and Mohri K. (1999). Weed detection in lawn field using machine vision: Utilization of textural features in segmented area. J. of the Japanese Society of Agricultural Machinery. 61(2): 61–69.

9. Burks, T. F., Shearer S. A., Gates R. S., and Donohue K. D. (2000). Backpropagation neural network design and evaluation for classifying weed species using color image texture. Transactions of the ASAE. 43(4): 1029–1037.

10. Mao, W., Wang Y., and Wang Y. (2003). Real-time Detection of Between-row Weeds Using Machine Vision. Las Vegas, NV. July 27-30, 2003. doi:10.13031/2013.15381.

11. Thompson, J. F., Stafford J. V., and Ambler B. (1990). Weed Detection in Cereal Crops. ASAE paper. 90–1629.

12. Franz, E., Gebhardt M. R., and Unklesbay K. B., (1991). The use of local spectral properties of leaves as an aid for identifying weed seedlings in digital images. Trans. ASAE. 34(2): 682–687.

13. Zhang, N. and Chaisattapagon C. (1995). Effective criteria for weed identification in wheat fields using machine vision. Trans. ASAE. 38(3): 965–974.

14. Vrindts, E. and De Baerdemaeker J. (1997). Optical discrimination of crop, weed and soil for on-line weed detection. In: J. Stafford (Ed.). Proc. of the 1st European Conference on Precision Agriculture. Warwick, UK. Bios Scientific Publishers Ltd, Oxford, UK, 537–544.

15. Noble, S. D., Brown R. B., & Crowe T. G. (2002). The use of spectral properties for weed detection and identification-a review. In AIC meeting Saskatoon, Canada.

16. Bendig, J., Bolten A., and Bareth G. (2013). UAV-based Imaging for Multi-Temporal, very high Resolution Crop Surface Models to monitor Crop Growth Variability. Photogrammetrie – Fernerkundung – Geoinformation. (6): 551–562. doi:10.1127/1432-8364/2013/0200.

17. Thenkabail, P. S., Smith R. B., and Pauw E. D. (2000). Hyperspectral vegetation indices and their relationship with agricultural crop characteristics. Remote Sensing of Environment. 71(2): 152–182.

18. Hansen, P. M. and Schjoerring J. K. (2003). Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. Rem. Sens. of Envir. 86(4): 542–553.

19. Usha, K. and B. Singh. (2013). Potential applications of remote sensing in horticulture - A review. Scientia Horticulturae. 153: 71–83. doi:10.1016/j.scienta.2013.01.008

20. Jones, H. G., & Schofield P. (2008). Thermal and other remote sensing of plant stress. General and Applied Plant Physiology, 34(1-2), 19–32.

21. Allen, R., Irmak A., Trezza R., Hendrickx J. M. H., Bastiaanssen W., & Kjaersgaard J. (2011). Satellite-based ET estimation in agriculture using SEBAL and METRIC. Hydrological Processes, 25(26), 4011–4027. doi:10.1002/hyp.8408

22. Solaiman, S., & Salaheen S. (2019). Future of Organic Farming: Bringing Technological Marvels to the Field. Safety and Practice for Organic Food, 291–303. doi:10.1016/b978-0-12-812060-6.00014-3

206

23. Sharma, S, M., Singh A., Solanki S.P., Sethuramalingam V., S., & Singh S.P. (2021). SVM-based compliance discrepancies detection using remote sensing for organic farms. Arabian Journal of Geosciences, 14, 1334. https://doi.org/10.1007/s12517-021-07700-4

24. Vroege, W., Bucheli, J., Dalhaus, T., Hirschi, M., & Finger, R. (2021). Insuring crops from space: the potential of satellite-retrieved soil moisture to reduce farmers' drought risk exposure. European Review of Agricultural Economics, 48(2), 266–314. doi: 10.1093/erae/jbab010

25. Li, G., Wan S., Zhou J., Yang Z., and Qin P. (2010). Leaf chlorophyll fluorescence, hyper spectral reflectance, pigments content, malondialdehyde and proline accumulation responses of castor bean (Ricinus communis L.) seedlings to salt stress levels. Ind. Crops Prod. 31: 13–19.

26 Kanning, M., Kühling I., Trautz D., & Jarmer T. (2018). High-Resolution UAV-Based Hyperspectral Imagery for LAI and Chlorophyll Estimations from Wheat for Yield Prediction. Remote Sensing, 10(12), 2000. doi:10.3390/rs10122000

27. Ma, B. L., Dwyer L. M., Costa C., Cober E. R., and Morrison M. J. (2001). Early prediction of soybean yield from canopy reflectance measurements, Agron. J. 93: 1227–1234.

28. Oøvergaard, S. I., Isaksson T., Kvaalc K., and Korsaeth A. (2010). Comparisons of two hand-held, multispectral field radiometers and a hyperspectral airborne imager in terms of predicting spring wheat grain yield and quality by means of powered partial least squares regression. J. Near Infrared Spectrosc. 18: 247–261.

29. Weber, V. S., Araus J. L., Cairns J. E., Sanchez C., Melchinger A.E., and Orsini E. (2012). Prediction of grain yield using reflectance spectra of canopy and leaves in maize plants grown under different water regimes. F. Crop. Res. 128: 82–90.

30. Herrmann, I., Vosberg S. K., Ravindran P., Singh A., Chang H. X., Chilvers M. I., Conley S. P., and Townsend P. A. (2018). Leaf and canopy level detection of Fusarium virguliforme (sudden death syndrome) in soybean. Rem. Sens. 10: 426.

31. Gitelson, A. Remote sensing estimation of crop biophysical characteristics at various scales. In: Hyperspectral Remote Sensing of Vegetation, Thenkabail, P. S., Lyon, J. G., Huete, A., Eds.; CRC Press. USA, 2012. ISBN 978- 1-4398-4537-0.

32. Liu, X., Zhang K., Zhang Z., Cao Q., Lv Z., Yuan Z., Tian Y., Cao W., and Zhu Y. (2017). Canopy chlorophyll density-based index for estimating nitrogen status and predicting grain yield in rice. Front. Plant Sci. 8: 1–12.

33. Singh, R., (2012). Crop Yield Estimation and Forecasting Using Remote Sensing, Indian Agricultural Statistics Research Institute: New Delhi, India.

34. Muñoz-Huerta, R. F., Guevara-Gonzalez R. G., Contreras-Medina L. M., Torres-Pacheco I., Prado-Olivarez J., & Ocampo-Velazquez R. V. (2013). A review of methods for sensing the nitrogen status in plants: advantages, disadvantages and recent advances. sensors, 13(8), 10823–10843. https://doi.org/10.3390/s130810823

35. Herrmann, I., Pimstein A., Karnieli A., Cohen Y., Alchanatis V., and Bonfil D. J. (2011). LAI assessment of wheat and potato crops by VENμS and Sentinel-2 bands. Rem. Sens. Environ. 115: 2141–2151

36. Tilman, D. (1999). Global environmental impacts of agricultural expansion: The need for sustainable and efficient practices, Proc. Natl. Acad. Sci. USA, 96: 5995–6000.

37. Diacono, M., De Benedetto D., Castrignanò A., Rubino P., Vitti C., Abdelrahman H.M., Sollitto D., Cocozza C., Ventrella D., (). A combined approach of geostatistics and geographical clustering for delineating homogeneous zones in a durum wheat field in organic farming, NJAS – Wageningen Journal of Life Sciences, 2013, Vol. 64–65, 47–57, ISSN 1573–5214, https://doi.org/10.1016/j.njas.2013.03.001.
38. Ferreira, S., Sanchéz J.M., Gonçalves J.M., (2021). Can Organic Farming development be driven by Remote Sensing technology?, in Proceedings of the 1st International Electronic Conference on Agronomy, 3–17 May 2021, MDPI: Basel, Switzerland, doi:10.3390/IECAG2021-10014
39. Taylor, J. A., McBratney A. B., & Whelan B. M. (2007). Establishing Management Classes for Broadacre Agricultural Production. Agronomy Journal, 99(5), 1366. doi:10.2134/agronj2007.007
40. Ammann, K. (2009). Why farming with high tech methods should integrate elements of organic agriculture. New Biotechnology, 25(6), 378–388. doi:10.1016/j.nbt.2009.06.933
41. Salima, Y., Peira J. F. M., Horra G., Rincon de la, & Ablanque, P. V. M. (2019). Remote Sensing Data: Useful Way for the Precision Agriculture. 2019 Sixth International Conference on Internet of Things: Systems, Management and Security (IOTSMS). doi:10.1109/IOTSMS48152.2019.8939229
42. Sozzi, M., Kayad, A., Gobbo, S., Cogato, A., Sartori, L., & Marinello, F. (2021). Economic comparison of satellite, plane and uav-acquired NDVI images for site-specific nitrogen application: Observations from Italy. *Agronomy*, *11*(11), 2098.
43. Castrignanò, A., Giugliarini L., Risaliti R., & Martinelli N. (2000). Study of spatial relationships among some soil physico-chemical properties of a field in central Italy using multivariate geostatistics. Geoderma, 97(1–2), 39–60. https://doi.org/10.1016/S0016-7061(00)00025-2

ПРИЛОЖЕНИЕ НА СПЪТНИЦИТЕ В БИОЛОГИЧНОТО ЗЕМЕДЕЛИЕ
(Част II)

М. Чанев, Л. Филчев

Резюме

Употребата на дистанционни методи за мониторинг, управление и подпомагане на вземането на решения в земеделието се засилва все повече. С развитието на технологиите те стават все по-достъпни, като същевременно се подобрява качеството и сигурността на получените данни. Стремейки се към подобряване на качеството на околната среда и нейното опазване разширяването на заетите площи с органично земеделие ще ни позволи да постигнем тези цел. Същевременно с това този тип земеделие предоставя здравословни и безопасни храни. Именно поради тази причина е от голямо значение

максимално бързо да се започне с прилагане на сателитни данни при органич-ното земеделие. В част II на „Review of the applications of satellite remote sensing in organic farming" разглеждаме различните области на приложение на сателитните данни в органичното земеделие. Отделени са пет различни области на приложение на сателитните данни в органичното земеделие, които са: сателитен мониторинг на плевелите, дистанционно наблюдение на стреса на културите и необходимостта от напояване, прогнозиране на добивите с помощта на дистанционни методи, дистанционен мониторинг на храненето на растенията. От направения преглед установихме, че сателитните данни могат значително да подпомогнат и улеснят прехода към биологично земеделие, да се извършва адекватно торене, намира приложение при фитосанитарния мониторинг на посевите и оценка на стреса, който изпитват културите.

IN-FLIGHT AND ON-ORBIT CALIBRATION OF LUNAR IMAGING SPECTROMETERS – A REVIEW

Iskren Ivanov, Lachezar Filchev

Space Research and Technology Institute – Bulgarian Academy of Sciences
e-mail: *iso.ivanov@space.bas.bg, lachezarhf@space.bas.bg*

Abstract

Lunar imaging spectrometers play a leading role in studying the Moon's mineral composition. The accuracy and reliability of the acquired data depend on the calibration process. Key stages of it include laboratory calibration, in-flight validation, on-orbit calibration, and cross-calibration. During these stages, various techniques and methods are used for calibration to achieve higher radiometric accuracy when recording the spectral reflective characteristics of materials in scenes from the lunar surface. These methods include capturing well-known calibration astronomical targets and calibration sites and comparing data from previous lunar surface studies obtained from orbital devices or ground-based telescopes. Other methods are capturing the Earth's atmosphere and utilization of on-board sources, such as lamps with a standardized emission spectrum. This paper reviews the techniques and methods utilized for in-flight and on-orbit calibration of lunar imaging spectrometers, drawing from an extensive overview of referenced science papers.

Introduction

Spectroscopy is an essential analytical method used to investigate material composition and related phenomena by observing the interactions between light and matter [1–3]. A notable benefit is its ability to ascertain composition remotely without requiring direct physical contact [4, 5]. The primary role of an imaging spectrometer is to identify materials or terrestrial features based on their spectral signature. The imaging component essentially acts as a map that displays the spatial location of these spectra, enabling comprehensive mapping and analysis of planetary surfaces [6]; it also finds wide application in Earth observation research, ranging from agricultural studies to wildlife population observation [7, 8]. To ensure precise spectroscopic measurements, it is essential to have a high signal-to-noise ratio (SNR), high calibration accuracy, and high response uniformity [9].

This article specifically focuses on the methods applied for in-flight, on-orbit, and on-board validation of the laboratory calibration of imaging spectrometers used in the SMART-1 (ESA), KAGUYA (SELENE) (JAXA), Chang'E-1 (CNSA),

Chandrayaan-1 (ISRO), and Chandrayaan-2 (ISRO) Moon missions. The specific instruments discussed include SIR (SMART-1 Infra-Red Spectrometer) [10], SP (Spectral Profiler onboard KAGUYA (SELENE) [11], IIM (Interference Imaging Spectroradiometer onboard Chang'E-1) [12], HySI (Hyper Spectral Imager onboard Chandrayaan-1) [13], SIR-2 (Near Infrared Spectrometer onboard Chandrayaan-1) [14], M3 (The Moon Mineralogy Mapper onboard Chandrayaan-1) [9], and IIRS (Imaging Infrared Spectrometer onboard Chandryaan-2 orbiter) [15]. Each of these spectrometers has its advantages and disadvantages, leading to specific characteristics in the acquired dataset. These datasets undergo specific processing and calibration before being prepared for use. In some cases, there were disruptions in the acquired spectral data, which imposed limitations on their usage. The science teams, dedicated to each instrument developed custom calibration methods and mathematical algorithms to normalize the registered output data and better understand the effects and characteristics of the imaging spectrometers.

The success of an imaging spectrometer relies on its design, built-in components, alignment, calibration, and stability during and after launch in the operational space environment [16]. These factors contribute to enhanced performance in terms of spectral, radiometric, spatial, and uniformity characteristics, which are critical for the calibration. It is paramount to verify the instrument characteristics measured during laboratory calibration after launch [17]. This is achieved by utilizing pre-launch measurements for characterization and observing well-known astronomical targets during flight to ascertain the effects of the space environment on imaging spectrometers [18]. The calibration process includes laboratory characterization and techniques for in-flight, on-orbit, and on-board characterization and validation. The survey aims to compile and present different calibration targets and methods employed in the reviewed imaging spectrometers' in-flight and on-orbit calibration process.

Development, alignment, testing, and calibration

The imaging spectrometer components are aligned and tested to ensure that its opto-mechanical, thermal, and electronics subsystems meet the science measurement requirements. The laboratory testing and calibration process is improved by utilizing a thermos-vacuum chamber, which simulates space vacuum conditions and a wide range of temperature fluctuations. This chamber allows for evaluating the instrument's performance under challenging low Moon orbit conditions, where temperatures can range from 400 K to 70 K when not illuminated and up to 400 K under direct illumination [9, 14, 15].

To accomplish this, a series of full imaging spectroscopic light measurements are conducted through calibration cold cycles in a thermal vacuum chamber for alignment and calibration purposes. Within this thermal vacuum chamber, the imaging spectrometers observe spectral, radiometric, and spatial

illumination sources that can be traced back to absolute standards at operational temperatures. Laboratory measurements and calibrations are performed to understand the intrinsic effects of the spectrometer and to offset these effects during subsequent calibration processes. Thorough lab analysis leads to the development of algorithms that can effectively correct a sensor's output and enhance its performance [19].

In-flight, on-orbit, and onboard confirmation of calibration and measurements

Human-made devices deployed in space are exposed to intense cosmic and solar radiation, potentially damaging the equipment. Fortunately, the Earth's atmosphere and geomagnetic field serve as protective shields against these harmful effects. However, devices in space must still endure severe vibrations and the vacuum environment to reach their designated research destinations. Once in space, these devices face many challenges, including wide temperature variations, heightened galactic and solar radiation, and the constant risk of colliding with high-energy particles, micrometeoroids, and space debris. Unlike Earth, the Moon lacks an atmosphere and a robust magnetic field, leaving devices in lunar orbit more vulnerable to these hazardous conditions. As a result, the harsh conditions in space can disrupt the components of these instruments, hinder their standard functionality, and cause damage [20].

Given the challenging launch process, the harsh environment of space, and the lengthy journey to the Moon, it is essential to calibrate imaging spectrometers by comparing their in-flight and on-orbit data with the preflight calibration results. The preflight calibration is a benchmark for the spectrometer's properties, while in-flight and on-orbit activities help update these calibrations to ensure accurate measurements of the object's surface [21].

The process of obtaining data from Earth observations (Fig. 1.) while the device is in-flight or in orbit is crucial for the precise validation and calibration of an imaging spectrometer's spectral performance and dark data (Fig. 2.). To achieve this, observations of Earth are conducted to record the spectral signature of the Earth's atmosphere, which is influenced by various gases such as O_2, CO_2, and water vapor. These absorption features in the atmosphere are then utilized to evaluate the spectral calibration of the imaging spectrometer while in orbit, ensuring consistency with the pre-launch laboratory spectral calibration. The accuracy of the imaging spectrometer's measurements can be validated using the MODTRAN (MODerate resolution atmospheric TRANsmission) [22] radiative transfer code, which involves comparing the calibrated spectra from the imaging spectrometer with the modeled spectra from MODTRAN. MODTRAN, developed through a partnership between Spectral Sciences Inc. and the Air Force Research Laboratory, is integrated into various operational systems and research sensors to process multi- and hyperspectral

remote sensing data requiring atmospheric correction [23]. This goes back to the HITRAN database of atmospheric gas absorption lines [24] and the laser calibration sphere, which has known laser wavelengths. HITRAN (High-Resolution Transmission Molecular Absorption Database) is a collection of spectroscopic parameters used by several computer programs to simulate the transmission and emission of light in the Earth's atmosphere. It was developed by the Atomic and Molecular Physics Division at the Harvard-Smithsonian Center for Astrophysics and is accessible to users on HITRANonline [25]. Aligning the positions of the wavelengths of the detected spectral absorption bands in Earth's atmosphere with those of the modeled ones provides crucial confirmation of the spectrometer's in-flight spectral calibration [9, 15].
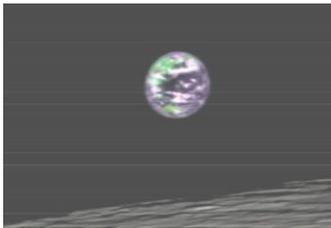

Fig. 1. Image of Earth taken from the Moon's orbit for M3 spectral calibration validation [26]
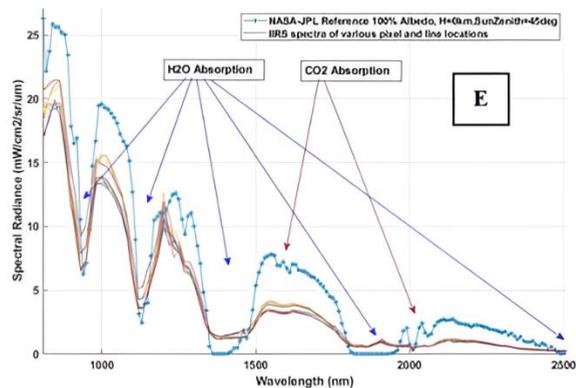
Fig. 2. Spectral radiance signature is graphed in conjunction with reference or standard spectra to compare the positions of spectral absorption bands [15]

To perform in-flight calibration, spectra can be captured from celestial bodies such as Jupiter (Fig. 3) and bright stars. Additionally, dark frames of unilluminated surfaces and areas of the dark sky can be utilized to test and confirm the calibration carried out on the ground [10]. The uniformity of the spectral Instantaneous Field of View (IFOV), which refers to the position of the IFOV relative to wavelength, is evaluated by analyzing bright targets in shadowed Polar Regions of datasets. These profiles, see Fig. 2, are normalized to a high radiance value, and their alignment across different spectral regions as they intersect with a brightly illuminated sample is used to validate the spectral IFOV uniformity of imaging spectrometers in lunar orbit [9].
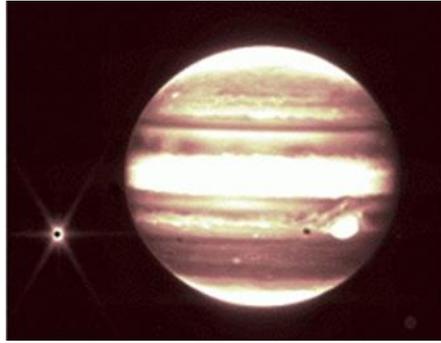
Fig. 3. The James Webb Space Telescope's NIRCam instrument, using a 2.12 micron filter, captures images of Jupiter, positioned in the center, and its moon Europa, to the left [Image credit: NASA, ESA, CSA, and B. Holler and J. Stansberry (STScI), 27]

During flight, radiation exposure can affect the dark current in detectors. While the overall rate of dark current remains constant, there is an increase in the number of pixels that exhibit a significant rise in responsivity (hot pixels) and multi-stable responsivity (RTS pixels). To account for this, bad pixel maps can be continuously updated to reflect the growing count of these pixel populations. The noise level of the imaging spectrometer is assessed by examining every pixel of the focal-plane array for new dead or damaged pixels. This is done using nighttime data from multiple images with lower exposure rather than capturing a single image with a long exposure time [28]. In addition, calibration characteristics obtained during the laboratory phase are utilized to evaluate the wear of the sensor and identify any inconsistencies in the collected data. This is achieved by recording calibration sites sequentially under consistent lighting conditions to track data changes and highlight any degradation in the sensor and its calibration. Regular observations conducted before and after each lunar day during the night [11] are used to make corrections for dark-level noise. Moreover, observations from deep space, nighttime [13, 29], and the dark side of the Moon can also contribute to the dark calibration process [26]. These observations aid in removing or correcting dark offset in the data.

The process of lunar calibration involves using the Moon's surface as a reference point for calibration. The reflectance of the Moon's surface is based on measurements obtained by Earth-based telescopes and previous lunar missions. Over a billion years, the Moon's surface reflectance has exhibited less than 1% variance, indicating high stability [30]. It is assumed that the Moon's phase of active volcanism concluded around 1.2 billion years ago during the Copernican period, and thus, the basic layout of albedo units on the Moon's surface is believed to have remained constant since then. With its high stability, the Moon's surface reflectance serves as a reliable photometric benchmark. Notably, the Moon's surface exhibits significant brightness variation at a phase angle $|\alpha| \leq 7°$, which is attributed to the strong

214

backscattering or brightness opposition effect. This brightness dependency on the phase angle provides valuable insights into the surface's composition and microstructure [31].
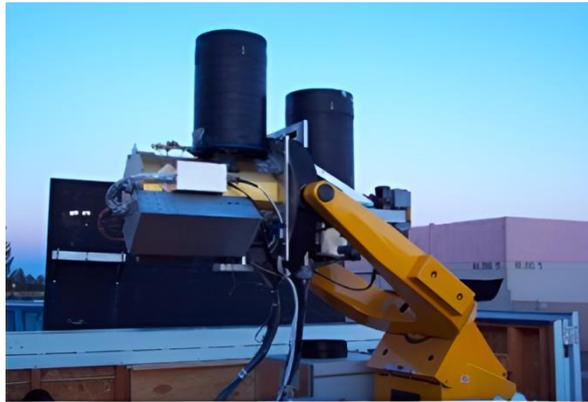

Fig. 4. RObotic Lunar Observatory (ROLO) at U.S. Geological Survey at the Flagstaff Science Campus in Flagstaff, Arizona [32]

To enable the Moon to serve as a radiometric calibration standard for spacecraft instruments in Earth's orbit and to develop a radiometric model of the Moon, NASA funded the U.S. Geological Survey at the Flagstaff Science Campus in Flagstaff, Arizona. This funding led to the establishment of the ground-based RObotic Lunar Observatory (ROLO) (Fig. 4.). The ROLO's lunar spectral irradiance specification is based on a database of spatially resolved radiance images of the Moon. These images were obtained from an observatory specifically designed and constructed for this project. Over more than six years, ROLO collected observational data, covering phase angles from near eclipse to typically 90 degrees before and after a full Moon, and encompassing a broad spectrum of observable libration angles. The radiometric model of the RObotic Lunar Observatory (ROLO) can estimate the Moon's brightness (irradiance) with an approximate precision of 1% over a broad phase range. This precision is advantageous for calibrating imaging spectrometers that are in orbit around the Moon [33]. The validation of spectral calibration while in orbit is achieved by comparing the location of absorption lines of pyroxene and olivine, as measured in the spectra of the imaging spectrometer. The validation of radiometric calibration is done by comparing measurements from imaging spectrometers with those from ROLO [9, 34]. The data sets are searched to find measurements with similar illumination and observational conditions for comparison. Data from an internal calibration lamp, which offers radiometric calibration details independent of lunar surface calibration sites, assists in tracking the radiometric sensitivity and spectral positioning of each pixel. This method aids

in distinguishing variations in sensitivity between data collected during flight and prior to flight. It also allows for updating the master flats (reference image) and the responsivity of each pixel throughout the mission [11, 28 and 35]. Lunar observations made by lunar on-orbit imaging spectrometers can be validated by comparing them with data from other instruments that observe the moon, including Earth-based telescopes.

The imaging spectrometer calibration can be validated by monitoring specific lunar calibration sites, such as Apollo 16 (Fig. 5), Mare Serenitatis 2 (Fig. 6), and other areas on the lunar surface. These sites have recognized near-infrared spectra derived from soil samples or telescope observations that have been adjusted for atmospheric influences [11, 13].
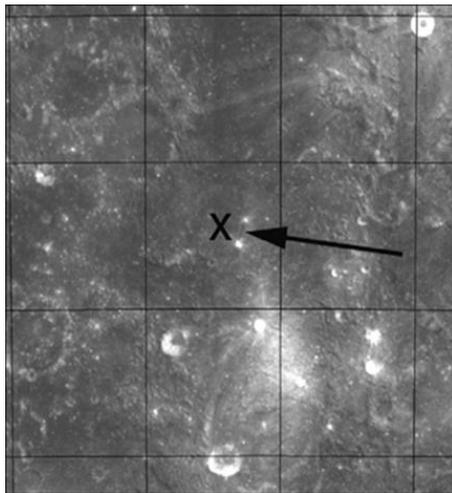

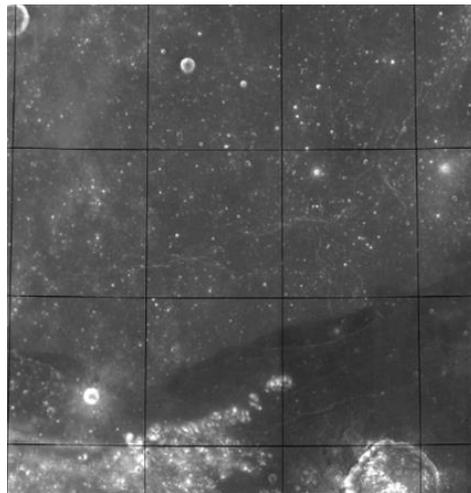Fig. 5. Image of Apollo 16 Central Nearside Highlands [35]

Fig. 6. Image of Mare Serenitatis [35]

Following global discussions, at scientific forums, under the COSPAR (Committee on Space Research) program in Beijing, and at the 8th International Conference on Exploration and Utilization of the Moon, the Lunar International Scientific Calibration/Coordination Targets (L-ISCT) (Fig. 7.) [35, 36], were proposed. This list of eight lunar calibration sites aims to enable cross-calibration of various multinational instruments, with the goal of creating a globally calibrated dataset for comparison with other instruments. Five of these targets were identified and discussed in depth at the 8th International Conference on Exploration and Utilization of the Moon, following their presentation at the COSPAR meeting in Beijing. This concept received international endorsement and was included in the Lunar Beijing Declaration [37]. The Apollo 16 site was selected first on the list of lunar calibration sites as part of the international collaboration and coordination

effort. Given the inevitable limitations of spacecraft resources, the site was chosen for its optimal calibration characteristics, which allow calibration of instruments performing orbital imaging, UV-Vis-NIR, gamma-ray spectroscopy, X-ray and neutron spectroscopy, altimetry, thermal, radar, and microwave imaging. The Apollo 16 site is a large area of relatively uniform feldspathic highlands on the lunar nearside. This knowledge is derived from the analysis of returned lunar samples. The site has become crucial for calibrating lunar spectroscopic data, with spectra of representative lunar soil sample 62231 [38, 39] collected and returned to Earth by the Apollo 16 crew. These samples were analyzed under laboratory conditions by the Reflectance Experiment Laboratory (RELAB) [40] and included in the Lunar Sample Compendium (https://curator.jsc.nasa.gov/lunar/lsc/ index.cfm) serving as ''ground truth'' for data obtained by remote sensing [37]. The remaining calibration targets on the list include: #2 Lichtenberg crater; #3 Apollo 15 Hadley Rille (Fig. 8.); #4 South Pole-Aitken Basin Th-anomalies; #5 Tycho crater; #6 Polar Region with shadows; #7 North Schrodinger; #8 Mare Serenitatis. Each of these proposed calibration targets is distinguished by its unique features and is linked to significant unresolved scientific inquiries.

This calibration method leads to better results than previous methods, such as observing stars or star patterns. By observing calibration sites, data from remote sensing of the Moon can be calibrated using laboratory-tested spectral reflectance characteristics of lunar surface samples. Aligning this data with data from previous lunar missions and data recorded by ground-based telescopes greatly aids consistent interpretation and analysis of lunar surface mineralogy.
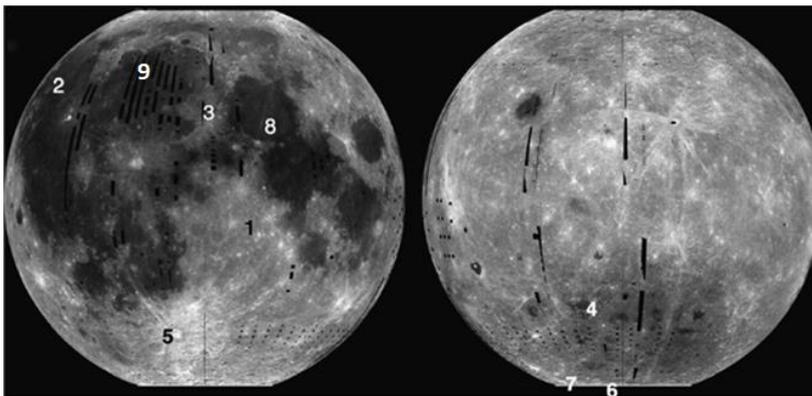

Fig. 7. A lunar map highlighting the eight suggested ISCT areas, each marked by a number, along with the CE-3 landing site, which is proposed as a new calibration site and labeled as #9 [35]
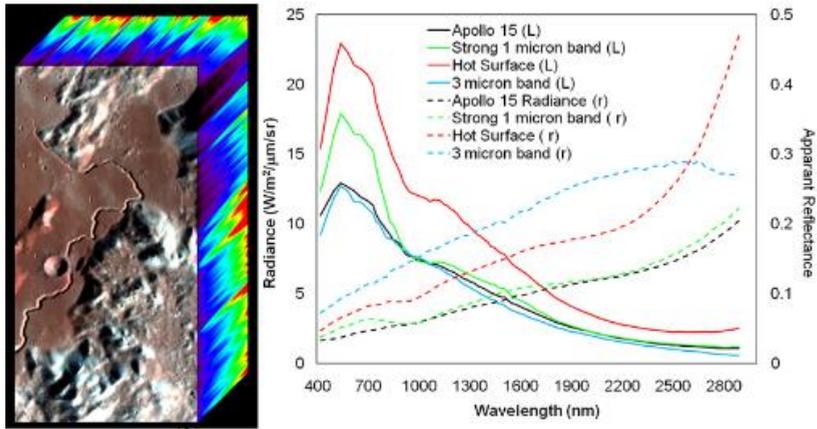
Fig. 8. M3 image of the Apollo 15 landing area, including the Hadley Rille [26]

The Visible-Near Infrared Spectrometer (VNIS) on China's Chang'E-3 (CE-3) mission's "Yutu" rover recorded the Moon's first in situ reflectance. The landing site of CE-3 is proposed as a new calibration site (Fig. 9.). The VNIS in situ reflectance indicates that the CE-3 landing site has a very low absolute reflectance, implying a high concentration of FeO and TiO2. The VNIS measurements fall between those from the Lunar Reconnaissance Orbiter Camera Wide Angle Camera (LROC WAC) and Spectral Profiler (SP), and those from the Moon Mineralogy Mapper (M3) and Imaging Infrared Mapper (IIM). Compared to commonly used calibration sites like MS-2 and Apollo 16 Highlands [41], the CE-3 calibration site is much younger and less impacted.

Fig. 9. Image of CE-3 calibration site [41]

Data from the Clementine mission can also be utilized. This can be achieved either by using Clementine's five channels that fall within the instrument's spectral range or by using Clementine UVVIS data for consistency checks [42].

The calibration models and laboratory tests primarily focus on the structures of the imaging spectrometers, neglecting the stray light from other parts of the spacecraft. However, images from in-flight system health checkouts have shown that, under specific illumination geometries, imaging spectrometers are susceptible to stray light reflecting off other instruments and structures on the spacecraft. A thorough stray light model of the entire spacecraft is utilized to mitigate these effects and examine various observational conditions. Such analyses enable the determination of spacecraft orientations relative to the Sun that are most beneficial for reducing stray light during critical observations throughout the mission. These predictions are validated in flight by capturing images with the spacecraft positioned in a range of orientations surrounding the anticipated optimal conditions [28].

Conclusion

This study underscores the critical role of comprehensive calibration in guaranteeing the accuracy of data acquired by lunar imaging spectrometers. Our key findings reveal that custom-developed corrections effectively address performance variations during flight and orbit. Earth's atmosphere spectral signature observations, lunar models, and dedicated calibration targets facilitate data calibration. Additionally, internal calibration lamps, Clementine mission data, and telescope observations offer valuable validation sources. Importantly, in-flight and on-orbit calibration strategies tackle instrument stability, spectral registration, stray light, and environmental effects. Challenges such as limited access to calibration standards, variations in lunar surface composition, stray light contamination, and the visibility of calibration targets persist. Continuous advancements in calibration methodologies are essential for enhancing the accuracy and reliability of lunar imaging spectrometers, ultimately paving the way for gaining deeper insights into the properties, composition, and geological history of the Moon.

Acknowledgments

References

1. Lucey, P. G. Radiative transfer modeling of the effect of mineralogy on some empirical methods for estimating iron concentration from multispectral imaging of the Moon, *Journal of Geophysical Research: Planets*, 2006, 111(E8). DOI: https://doi.org/10.1029/2005JE002661.
2. Mardirossian, G. Osnovi na distantsionnite aerokosmicheski tehnologii. Sofiya, *Nov balgarski universitet*, 2015. (In Bulgarian).
3. Mishev, D. Remote sensing of the Earth. Sofia, *Publishing house of the Bulgarian academy of sciences*, 1981, 206 p.
4. Mardirossian, G., D. Dimitrov, I. Ivanova et. al. Informatsionni tehnologii za nablyudenie na zemyata ot kosmosa. Praktichesko rakovodstvo, chast 1. Sofiya, *Institut za kosmicheski izsledvaniya i tehnologii – BAN*, 2020. (In Bulgarian).
5. Mardirossian, G. Vavedenie v kosmonavtikata. Sofiya, *Institut za kosmicheski izsledvaniya i tehnologii – BAN*, 2012, 261 p. (In Bulgarian).
6. Hu, X., T. Jiang, P. Ma et al. A Spectral Library Study of Mixtures of Common Lunar Minerals and Glass. *Remote Sensing*, 2023, 15(8), Article 8. DOI: https://doi.org/10.3390/rs15082195.
7. Chanev, M., L. Filchev. Review of the Applications of Satellite Remote Sensing in Organic Farming (Part I), *Aerospace Research in Bulgaria*, 2023, 35, 183–191. DOI: https://doi.org/10.3897/arb.v35.e18.
8. Chanev, M., N. Dolapchiev, I. Kamenova, et. al. Application of remote sensing methods for monitoring wild life populations: a review, Proceedings of the 9th International Conference on Remote Sensing and Geoinformation of the Environment, 21 September 2023, SPIE 12786, (RSCy2023), 1278622. DOI: https://doi.org/10.1117/12.2681760.
9. Green, R., C. Pieters, P. Mouroulis, et al. The Moon Mineralogy Mapper (M(3)) imaging spectrometer for lunar science: Instrument description, calibration, on-orbit measurements, science data calibration and on-orbit validation, *Journal of Geophysical Research-Planets*, 2011, 116. DOI: https://doi.org/10.1029/2011JE003797.
10. Basilevsky, A., H. Keller, Nathues, et al. Scientific objectives and selection of targets for the SMART-1 Infrared Spectrometer (SIR), *Planetary and Space Science*, 2004, 52(14), 1261–1285. DOI: https://doi.org/10.1016/j.pss.2004.09.002.
11. Matsunaga, T., M. Ohtake, J. Haruyama, et al. KAGUYA (SELENE) / Spectral Profiler : In-flight performance and future plan. *Lunar and Planetary Science,* 2008, XXXIX.
12. Wu, Y., X. Xu, Z. Xie, et al. Absolute calibration of the Chang'E-1 IIM camera and its preliminary application. *Sci China Ser G*, 2009, 52(12), 1842–1848. DOI: https://doi:10.1007/s11433-009-0282-z.
13. Seelin, K., R. Kumar, Chowdhury, et al. The Hyper Spectral Imager Instrument on Chandrayaan-1, 2009, 1589.
14. Mall, U., M. Banaszkiewicz, K. Bronstad, et al. Near Infrared Spectrometer SIR-2 on Chandrayaan-1, *Current Science*, 2009, 96, 506–511.

15. Kumar, P., A. Bhattacharya, S. Sampat, et al. IIRS (Imaging Infrared Spectrometer onboard Chandryaan-2 orbiter) Onboard spectral calibration and validation operations, *54th Lunar and Planetary Science Conference*, 2023.

16. Atanassov, V., G. Jelev & L. Kraleva. Imaging spectrometer data correction, *Space Ecology Safety*, 2005, Varna, Book I, SRI-BAS ISBN 954-438-484-7, 221–226.

17. Mouroulis, P., R. Green, Review of high fidelity imaging spectrometer design for remote sensing, *Optical Engineering*, 2018, 57(4).
    DOI: https://doi.org/10.1117/1.OE.57.4.040901.

18. Filacchione, G., F. Capaccioni, E. Ammannito, et al. Calibration pipeline of VIS-NIR imaging spectrometers for planetary exploration: The rosetta VIRTIS-M case, *First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, 2009, 1–4.
    DOI: https://doi.org/10.1109/WHISPERS.2009.5289050.

19. Woodward, J. T., S. W. Brown, Smith, et al. Hyperspectral imager characterization and calibration, *2009 IEEE International Geoscience and Remote Sensing Symposium*, 2009, II-77–II-80. DOI: https://doi.org/10.1109/IGARSS.2009.5418005.

20. Lu, Y., Q. Shao & H. Yue. A Review of the Space Environment Effects on Spacecraft in Different Orbits. *IEEE Access*, 2019, PP, 1–1.
    DOI: https://doi.org/10.1109/ACCESS.2019.2927811.

21. Vilenius, E., U. Mall & V. Kaydash. In-flight calibration of SIR near infrared spectrometer onboard SMART-1, 2006, 139.
    DOI: https://ui.adsabs.harvard.edu/abs/2006epsc.conf..139V.

22. Berk, A., G. Anderson, L. Bernstein, et al. MODTRAN4 Radiative Transfer Modeling for Atmospheric Correction, *Proceedings of SPIE - The International Society for Optical Engineering*, 1999, 3756. DOI: https://doi.org/10.1117/12.366388.

23. MODTRAN – Radiative Transfer. Spectral sciences Inc., https://www.spectral.com/our-software/modtran/ (last date visited 25 October 2023).

24. The HITRAN2020 molecular spectroscopic database. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 277,
    DOI: https://doi.org/10.1016/j.jqsrt.2021.107949.

25. Rothman, L. S. History of the HITRAN Database, *Nature Reviews Physics*, 2021, 3(5), 302–304. DOI: https://doi.org/10.1038/s42254-021-00309-2.

26. Green, R. O. Assessment of the Complete Moon Mineralogy Mapper Data Set and On-Orbit Validation of the Spectral Calibration, *41st Lunar and Planetary Science Conference*, 2010, 1533, 2190.

27. Fisher, A. Webb Images of Jupiter and More Now Available in Commissioning Data, NASA, ESA, CSA, and B. Holler and J. Stansberry from (STScI), https://blogs.nasa.gov/webb/2022/07/14/webb-images-of-jupiter-and-more-now-available-in-commissioning-data/ (date on-line on July 14, 2022).

28. Golish. Ground and In-Flight Calibration of the OSIRIS-REx Camera Suite, *Space Science Reviews*, 2020, 216. DOI: https://doi:10.1007/s11214-019-0626-6.

29. Chowdhury, A. R., A. Banerjee, S. Joshi, et al. Imaging Infrared Spectrometer onboard Chandrayaan-2 Orbiter. *Current Science*, 2020, 118, 368–375.
    DOI: https://doi.org/10.18520/cs/v118/i3/368-375.

30. Kieffer, H.H. Photometric Stability of the Lunar Surface. *Icarus*, 1997, 130, 323–327.
    DOI: https://doi:10.1006/icar.1997.5822.

31. Kaydash, V. Pieters, C. Shkuratov, et al. Lunar Opposition Effect as Inferred from Chandrayaan-1 M3 Data. *J. Geophys. Res. Planets*. 2013, 118, 1221–1232. DOI: https://doi:10.1002/jgre.20098.

32. Stone. T. C. Current Status of ROLO and Future Development, *GSICS/IVOS 2nd Lunar Calibration Workshop*, Xi'an, China, 13-16 November 2017, https://slideplayer.com/slide/15885306/.

33. Kieffer, H. H., Stone, T. C. The Spectral Irradiance of the Moon. *Astron. J.* 2005, 129, 2887. DOI: https://doi:10.1086/430185.

34. Stone, T. C., & Kieffer, H. H. Absolute Irradiance of the Moon for On-orbit Calibration, *Proceedings of SPIE - The International Society for Optical Engineering*, 2002. DOI: https://doi:4814. 10.1117/12.451694.

35. Pieters, C. M., J. W. Head, P. Isaacson, et al. Lunar international science coordination/calibration targets (L-ISCT). *Advances in Space Research*, 2008, 42(2), 248–258. DOI: https://doi.org/10.1016/j.asr.2007.05.038.

36. Ivanov, I., Filchev, L. Calibration sites and samples for optical remote sensing of the Moon, Sofia, In: Nineteenth International Scientific Conference, Space Ecology Safety (SES), 2023, 213–218, http://space.bas.bg/SES/archive/ SES%202023_DOKLADI/3_Remote%20Sensing/10_Ivanov.pdf, (In Bulgarian).

37. Foing, B. H., & J. Wu. Beijing Lunar Declaration 2006. *Advances in Space Research*, 2008, 42(2), 244–245. DOI: https://doi.org/10.1016/j.asr.2008.03.015.

38. Meyer, C. Carbon content and maturity index for 62241. Lunar Sample Compendium, *NASA STI/Recon Technical Report*, 2010, 6, 11039. https://curator.jsc.nasa.gov/lunar/lsc/index.cfm (last date visited 25 October 2023).

39. Pieters. C. M., The moon as a spectral calibration standard enabled by lunar samples: the Clementine example. New Views of the Moon 2: Understanding the Moon Through the Integration of Diverse Datasets, 1999, 47.

40. Pieters, C., & T. Hiroi. RELAB (Reflectance Experiment Laboratory): A NASA multiuser spectroscopy facility, Lunar and Planetary Science XXXV, 2004, 35.

41. Yunzhao, W., Wang, et al. The Absolute Reflectance and New Calibration Site of the Moon, *The Astronomical Journal*, 2018, 213. DOI: https://doi.org/10.3847/1538-3881/aabaf5.

42. Hillier, J. K., B. J. Buratti & K. Hill. Multispectral Photometry of the Moon and Absolute Calibration of the Clementine UV/Vis Camera, *Icarus*, 1999, Volume 141, Issue 2, 205–225. DOI: https://doi.org/10.1006/icar.1999.6184.

КАЛИБРИРАНЕ НА ОПТИЧЕСКИ СПЕКТРОМЕТРИ ПО ВРЕМЕ НА ПОЛЕТ И В ОКОЛОЛУННА ОРБИТА - ОБЗОР

И. Иванов, Л. Филчев

Резюме

Оптическите спектрометри имат водеща роля в изследването на минералния състав на Луната. Точността и надеждността на данните, регистрирани от тези инструменти, зависят от процеса на калибриране. Ключови етапи от него заема лабораторното калибриране, валидирането по време на полет, в орбита и насрещното калибриране. По време на тези етапи се използват различни техники и методи за калибриране с цел постигане на по-висока радиометрична точност. Тези методи включват заснемане на земната атмосфера, на добре известни астрономически цели и площадки за калибриране. Съпоставка с данни от предишни изследвания, получени от орбитални апарати или земни телескопи. Приложение намират и бордови източници за калибриране, като лампи със стандартизиран спектър на излъчване. Обзорът представя преглед на техники и методи, използвани за калибриране на оптически спектрометри, в полет към Луната и в лунна орбита, подготвен въз основа на обширен преглед на цитираната реферирана и нереферирана научна литература.

At the end of 2023, the Bulgarian Academy of Sciences Publishing House "Prof. Marin Drinov" published the second revised and supplemented edition of the book "Introduction to Cosmonautics" by Prof. Doctor of Technical Sciences and Doctor of Physics Eng. Garo Mardirossian. The author is well known to space research and technology specialists for his ten books, over 150 scientific publications and reports, more than 60 patents, and hundreds of popular science articles in print and electronic media.

The book "Introduction to Cosmonautics" is dedicated to an interesting and complex interdisciplinary field of scientific and applied science – space research and technology.

After a brief synopsis of outer space, the Earth-orbiting Cosmos, and their study, the reader is introduced to the individual elements of cosmonautics. The spacecraft, their orbits for various purposes and the ground-based space complexes are examined. The main ideas in cosmonautics follow – construction of space apparatuses, space physics, space meteorology, remote sensing of the Earth, space materials science and materials production, space communication systems, space navigation and geodetic systems, space biology, and medicine.

Brief information on the application of space technologies in military affairs and security is also presented. The last part of the book is devoted to studying some of the solar system planets.

Naturally, this second revised and supplemented edition also includes the most significant successes and achievements of world cosmonautics in the period from the first edition in 2012 of the book until now. It also provides a brief overview of the historical development and more significant successes of space research and technology in Bulgaria.

The book also includes some of the results obtained by the author during his more than 40-year work at the Space Research and Technologies Institute (SRTI) at the Bulgarian Academy of Sciences.

The book is intended for a wide readership – researchers in the field of space studies, Earth sciences (geography, geophysics, geology, hydrology, meteorology, etc.), natural sciences, aviation and cosmonautics, as well as teachers and students in secondary schools and higher civilian and military educational institutions.

Corr. Member DSc Eng. Petar Getsov