Near-earth space as a habitat

David Waring Dunham¹, Vladimir Petrovich Kulagin², Victor Yakovlevich Tsvetkov³

¹Member of the International Academy of Astronautics and of the Russian Academy of Natural Sciences, Head of the Laboratory for Space Science, Technology, Systems and Processes Higher School of Economics, Moscow, Russia
²Member of the Russian Academy of Natural Sciences, Moscow, Russia, Co-director of the Laboratory for Space Science, Technology, Systems and Processes Higher School of Economics, Moscow, Russia
³Member of the Russian Academy of Space Named after E. K. Tsiolkovsky, Moscow, Russia

Email address:
david.dunham@kinetx.com (D. W. Dunham), vkulagin@hse.ru (V. P. Kulagin), cvj2@list.ru (V. Y. Tsvetkov)

To cite this article:

Abstract: The article analyzes the near-Earth space as a future habitat for humankind. This article investigates the factors affecting the location in this environment. We estimate the boundaries of space and related space. The article highlights the main features of the near-Earth space as a human-friendly environment.

Keywords: Astronomy, Space Exploration, Remote Sensing, Near Space, Planetology

1. Introduction

The development of humankind drives it toward the search for new habitats [1], and makes them, in turn, an object of scientific research. Near-Earth Space (NES) has become an object of study for such research and a possible new habitat [2]. This term has been used consistently since the launch of the first artificial satellite, but is vague in meaning as a geographical category. [3]

The boundaries of this space must be identified to solve the problems of global monitoring [4], the connection of this space with the Earth itself [5], for navigation and other tasks

2. Analysis of the Near-Earth Space

Currently there is no clear definition of the boundaries of Near-Earth Space (NES). To determine the boundaries we need to peg this area against the physical fields of the Earth and the human presence in this space.

The analysis of the human factor binds the border to the height of geostationary satellites at approximately 36,000 km. However, the higher elliptical orbits stretch as far as 47,000 km. This gives ground to fix the upper limit value of NES at 50,000 km from the Earth's surface.

The lower limit of NES is determined by boundary between airspace and outer space. This value is set at 100 km above the Earth's surface.

It would be far more convenient to define the boundaries of NES in Earth radii (R = 6371 km) in order to compare it with the size of the Earth. In this case, the far edge of NES is defined as 8Re (Re - earth radii) above the surface of Earth (ca. 51,000 km). In this respect, the far edge of NES (relative to the Earth) is defined by a sphere centered at the center of the Earth with a radius equal to 9Re therefrom or 8Re from the surface of the Earth. The nature of this boundary is dictated by human activity. We can call this zone Near Earth Space 1 or NES1.

Space can be divided into near-Earth, cis-lunar and trans-lunar space. Cis-lunar Outer Space (CLOS) stretches to the orbit of the moon (384,400 km or about 60Re). Trans-lunar Outer Space (TLOS) is determined by one astronomical unit (23481Re).

The vast majority of space resources is concentrated in NES1, more specifically, in the lower orbit area (about 60%).

Satellites in lower orbits (700-1500 km) offer several advantages over other spacecraft in terms of the energy performance, however, they look much worse when it comes to the duration of the communication sessions, as well as their overall service life. The period in which the satellite makes a full circle is an average of 100 minutes, whereby about 30% of the time it stays on the shady side of the planet. On-board rechargeable battery go through around 5000 charge/discharge cycles per year, as a result their term of service does not exceed more than 5-8 years.

Medium-altitude orbits are located between the first and second radiation belts, that is, at an altitude of 5,000-15,000
km. These units are weaker than geostationary ones, so to fully cover the surface of the Earth an orbital group of 8-12 satellites is required (e.g., Spaceway NGSO, ICO, “Rostelesat”), whereby each satellite is within the range of ground stations for quite a short period of time – about 1.5-2 hours.

Higher orbits include geostationary and elliptical orbits. Geostationary orbits (also known as circular orbits) are most often used to accommodate the spacecraft. They offer significant advantages: the communication is constant and round-the-clock and the frequency shift is virtually nonexistent. Geostationary satellites are at an altitude of approximately 36,000 km above the Earth's surface rotating synchronously with it. They "hover" at a particular point – the "sub-satellite point." However, in fact, the position of the satellite is not fixed. It experiences a certain "drift" due to a number of factors, so the orbit itself shifts over time.

A High elliptical orbit (HEO) is a type of elliptical orbit where the height at the apogee (ca. 50,000 km) is many times greater than the height at perigee (ca. 500 km).

Let us consider the physical factors. The basic physical properties and natural features of the Near Earth Space (NES):

- gravitational, magnetic and electric field of the Earth;
- Processes in the Earth's ionosphere;
- a deep vacuum of space;
- thermal radiation;
- cosmic rays and solar radiation;
- Earth's radiation belts.

Let us now consider the effect of these factors.

The influence of the gravitational field of the Earth. The shape of the Earth is a rather complex surface called a geoid. Its shape has so far not been determined completely, so the task of further refinement is the subject of the gravimetric theory and the theory of the earth's figure. Perturbing acceleration caused by the gravitational field of non-spherical Earth is 1/656 part of g. For low and high elliptical orbits this factor plays a significant role, as the spacecraft may enter and exit the zone of the Earth's significantly perturbing gravitation (ca. 930 km).

The analysis of the impact of perturbations on the motion of the spacecraft (SC) in an elliptical orbit shows that the average orbit does not change its size and shape, that is, its semi-major axis, the focal parameter, eccentricity and linear eccentricity, apogee and perigee remain almost unchanged. However, it gradually turns around the axis coincident with the axis of rotation of the Earth, which leads to the rotation of the apogee - perigee line, which significantly affects the ballistic stability of the spacecraft.

When forming groups of spacecraft orbiting the Earth, the comprising elements should be chosen in such a way as to minimize the effect of the gravitational field of the Earth on the motion of the spacecraft. This would ultimately provide groups with ballistic stability for a long time and minimize the on-board fuel stocks required for angle correction.

The influence of the magnetic field of the Earth. The magnetosphere of the Earth has a complex shape (see Fig. 1).

The boundary of the magnetosphere, or the magneto, from the side of the Sun is shaped like a shell, and is at a distance of roughly 15 Re. The night side of the Earth's magnetosphere extends as a long cylindrical tail (the magnetic tail) with a radius of about 20-25 Re. The tail stretches to a considerable distance of over 200 Re. The physical fields (the magnetosphere) seem to push back the boundary of the NES to 15 Re or 95565 km. This zone can be identified as Near Earth Space 2 or NES2.

The sources of magnetism of the Earth are mainly found in three components of the planet: in the nucleus, the mantle, and the upper atmosphere. The magnetic field undergoes certain changes and disturbances over time. Regular changes, that are subject to certain laws and algorithms, are called unperturbed variations. These include the annual lunar-diurnal and solar-diurnal unperturbed variations. Sporadic changes in Earth's magnetic field are called magnetic disturbances.

The boundary of the magnetosphere, or the magneto, from the side of the Sun is shaped like a shell, and is at a distance of roughly 15 Re. The night side of the Earth's magnetosphere extends as a long cylindrical tail (the magnetic tail) with a radius of about 20-25 Re. The tail stretches to a considerable distance of over 200 Re. The physical fields (the magnetosphere) seem to push back the boundary of the NES to 15 Re or 95565 km. This zone can be identified as Near Earth Space 2 or NES2.

The sources of magnetism of the Earth are mainly found in three components of the planet: in the nucleus, the mantle, and the upper atmosphere. The magnetic field undergoes certain changes and disturbances over time. Regular changes, that are subject to certain laws and algorithms, are called unperturbed variations. These include the annual lunar-diurnal and solar-diurnal unperturbed variations. Sporadic changes in Earth's magnetic field are called magnetic disturbances.
high latitudes. Magnetic storms are caused by solar plasma from the active regions of the Sun intensified by the quiet solar wind. Magnetic storms are dangerous for the crews of manned spacecraft and the functioning of on-board equipment.

Predicting magnetic storms is possible within a three day period only. During this time you can take any practical action with the spacecraft and crew to reduce (or exclude) the impact of the flow of high-energy particles. During magnetic storms, the parameters of the layers of the ionosphere, reflecting and absorbing the radio waves significantly change. The result is a significant interference in the short-wave communication.

The influence of the electric field of the Earth. The electric field of the Earth (ca. 36,000 km) gets into the NES1 zone. The natural electric field of the earth is observed in the solid earth, in the seas, in the atmosphere and magnetosphere. It is caused by a complex set of geophysical phenomena. The potential distribution of the field is a source of some information about the structure of the Earth, about the processes occurring in the lower atmosphere, the ionosphere, magnetosphere, as well as near interplanetary space and the Sun. The existence of an electric field in the Earth's atmosphere is mainly due to the air ionization processes, and the space separation arising from ionization of the positive and negative electric charges exploded arising. Ionization occurs under the influence of cosmic rays, ultraviolet radiation from the sun, radiation, radioactive substances present on the surface of the earth and in the air, electrical discharges in the atmosphere and other factors. Many atmospheric phenomena and processes: convection, cloud formation, precipitation, and others - lead to a partial separation of opposite electric charges, and the emergence of atmospheric electric fields. Earth is negatively charged relative to the atmosphere.

The existing electric fields of the atmosphere lead to the appearance of currents unloading the "atmosphere - the surface of the earth" "capacitor". At the same time, precipitation plays a significant role in the exchange of charges.

The processes occurring in the upper atmosphere and magnetosphere cause the electric fields in the ionosphere. Tidal movement of air masses, winds, turbulence are all sources of electric field generation in the ionosphere. Its intensity depends on the location, time of day, the general state of the magnetosphere and ionosphere, and solar activity. Amperage can reach hundreds of thousands of amperes.

One of the immediate sources of the electric field in the magnetosphere is the solar wind. During magnetic storms and 'northern lights' the electric fields and currents in the magnetosphere and ionosphere experience significant changes.

Studies suggest that the effect of the electric field of the Earth on the operation of spacecraft appears in the form of external surfaces electrification; inappropriate functioning of on-board systems, and communication interference.

The Influence of the Earth's Ionosphere

The ionosphere extends from the mesosphere for around 1,000 km. It is entirely a part of NES1. The ionosphere is a natural formation of discharged weakly ionized plasma in a magnetic field of the Earth and due to its high electrical conductivity has specific properties that determine the nature of the radio waves distribution in it and various disturbances. When the radio waves enter in the ionosphere, they undergo reflection, double refraction, dispersion and nonlinear effects.

Depending on the density of the charged particles in the ionosphere layers D (60-90 km), E (90-120 km) and F (above 130-140 km) can be identified [3]. Getting into the ionosphere (anisotropic environment) the wave undergoes birefringence, in other words it is split into two waves with different speed and direction, absorption and polarization. With the spread of the two a phase shift accumulates between the waves, so the polarization of the resulting wave is preserved in certain circumstances, but the field of polarization rotates during the distribution. In general, the polarization of the two waves is elliptical.

In addition to the regular dependence (regular ionosphere) of the electron density on the height in the ionosphere, accidental changes in the concentration also occur. The ionospheric layer contains a large number of sporadic formations of various sizes, which are in constant motion and change, dissolving and emerging again. It results in the random change of the received signal transmitted. The existence of heterogeneous formations leads to the scattering of radio waves at frequencies much higher than the maximum frequency of the reflection from the regular ionosphere. Heterogeneous formations occur in the ionosphere also upon the passage of meteors therethrough.

There are peculiarities of the distribution of the different range of radio waves in the ionosphere. In particular, they are limited to low frequency absorption. Therefore, communication is carried out in a range of short waves during the day and at night in the medium wave band. Long and extra-long waves practically do not penetrate the ionosphere, bouncing off the bottom border.

Audio-frequency radio waves can seep through the ionosphere along the magnetic field lines of the Earth. Spreading, a wave goes over a distance of several earth radii, and then returns to the conjugate point located in the other hemisphere. For subsonic radio waves with a frequency lower than the frequency of the ions, the ionosphere acts as a neutral fluid, the motion of which is described by the equations of hydrodynamics. Thus, the influence of the ionosphere shows through effects on the communication lines in the form of noise and signal change.

The Van Allen Radiation Belts

The Van Allen radiation belts are regions of space filled with charged particles held in place by the Earth's magnetic field. The intensity of the flow of charged particles in the Earth's radiation belt exceeds by several orders of magnitude.
the intensity of cosmic rays. A spacecraft, operating in high and stationary orbits, passes through the two radiation belts around the Earth. Their impact affects the operation of electronic equipment and the state of the crew.

The Van Allen radiation belts were discovered by an American scientist (James Van Allen) hence the name. It is a toroid, comprised of two areas: inner radiation belt at an altitude of ~ 4000 km, consisting mainly of protons with energies of tens of MeV; outer radiation belt at an altitude of about 17 000 km, consisting primarily of electrons with energies of tens of keV.

The dependence of the position of the lower boundary of the radiation belt is dictated by the geographical longitude. An increase in intensity over the Atlantic starts at an altitude of 500 km and over Indonesia at an altitude of 1300 km.

Between the inner and outer radiation belts, there is a free zone, located in the range of 2 to 3 Re. Particle fluxes are greater in the outer belt than in the inner. The composition is also different: particles in the inner belt are protons and electrons whereas in the outer there are electrons. It can be noted that the radiation belt of the Earth is entirely a part of NES1.

**Belt of Debris**

In a broad sense, "space debris" is understood as "man-made" garbage, as well as natural space objects or parts that fall into the sphere of attraction of the Earth and near-Earth space. Components of space debris are a dangerous factor affecting the operation of a spacecraft, especially manned ones.

Originally, two physical zones were identified: low-orbit 800-1500 km and in the geostationary orbits 36,000 km.

Besides the physical debris a belt can be formed as a result of high-altitude nuclear explosions and radioactive fumes from nuclear power plants. Since the early 1990s, the compaction of fine particles was revealed (with the help of infrared imaging) in orbits close to 100 - 900 km with 65-degree inclination [6]. All the debris belts fall into NES1.

Legal aspects of the use of NES. Legal particularities of NES1 use deserve special attention. It does not lend itself to division through natural or man-made boundaries. At present, international law does not have a unified definition of the lower bounds of near-Earth space. The space and air space are separated by a diffusion layer with a thickness of about 35 km, where an orbital flight is not yet possible and aerodynamic controlled flight is no longer possible.

3. Conclusions

The following are the main peculiarities of NES1.

1. Near Earth Space is a natural environment that is gradually emerging as a geotechnical system.

2. We can distinguish two zones NES1 and NES2. NES1 is conditioned by human activities, and includes most of the physical fields of the Earth. NES2 is conditioned by the presence of the magnetosphere. The outer boundary of NES1 tends to get closer to NES2.

3. Functioning in NES1 and NES2, specific factors have an effect on technical systems that cannot be said of other environments. The following are the more evident effects: evaporation and permanent erasure of the surface layers of outer lining materials, change in the optical properties of thermal control coatings and optical materials, transformation of lubricants into abrasive material, the electronic discharges on the surface, breakdown of thin-walled shells of objects. These circumstances lead to a decrease in the reliability of different technical systems and the reduction of their lifetime in orbit. This suggests that space requires specific means adapted to function effectively in this environment.

4. Near Earth Space does not lend itself to division by nationality; it is purely extraterritorial. On the one hand, the recognition of the full or partial sovereignty of the states in the space jeopardizes the implementation of the programs of space exploration because of protests about the violation of their sovereignty, on the other hand, the extra-territorial nature of space allows influencing the events occurring in any part of the world.

5. Near-Earth space is currently a specific sphere of human activity where military action can be carried out in order to protect national interests.

6. The study of NES raises a number of new scientific problems, the solution of which will determine the progress of humanity and its future.

References


