Model of the global distribution of the total electron content based on deep dense convolutional autoencoder

Stanislav V Khristoforov1, Vladimir V Bochkarev2

1,2Kazan Federal University, Institute of Physics, Russian Federation

E-mail: 1 stnslv91@gmail.com

Abstract. Nowadays the prediction of ionospheric parameters is an important and acute problem in the field of ensuring stable operation of radio communication and radio navigation facilities. The network of two-frequency GPS receivers data is used for monitoring the ionospheric condition. Based on these data, a number of laboratories are building global maps of total electron content (TEC). There are strong spatial and temporal correlations in the TEC maps. As a result, in order to successfully solve the problem of TEC prediction, it is advisable to perform preliminary processing of maps data with dimensionality reduction. In this paper, the problem of constructing a low-dimensional model of global distribution of the TEC is solved. In addition, the model of global distribution of the TEC can be useful for the ionosphere dynamics investigation. In this paper, it is proposed to use dense convolutional autoencoders as a base element of the model. This architecture allows us to speed up the neural network learning process and avoid the gradient-vanishing problem in error backpropagation algorithm.

1. Introduction
At present, modelling of the ionosphere is an acute task in the field of providing stability for radio communication. A high-quality model should take into account all possible states of the ionosphere. This is an important condition, since the presence of ionospheric disturbances can significantly affect the performance quality of various radio receivers. For example, in the case of a GPS system, the ionosphere is a constant source of delays for the signal arriving at receivers. These measurement errors can cause uncertainty of coordinates, which is undesirable.

Therefore, to ensure high-quality operation of the radio navigation equipment, it is necessary to make corrections associated with conditions of signal propagation in the ionosphere in due time. This task is associated with the task of predicting its state.

Due to strong spatial and temporal correlations of TEC values, it is possible to represent the state of the ionosphere as a vector with a much smaller dimension than in the original data. This will allow us to generalize information on the evolution of the ionosphere state, and also reduce the need for computational resources during the construction and use of the model. Therefore, an important element in the construction of the ionospheric TEC model is using the data compression method. One of the classic methods of data compression is the principal components analysis (PCA) [1]. Its main feature is that the algorithm tries to approximate the data by hyperplanes. Therefore, in the case of complex nonlinear dependences, its efficiency can be greatly reduced. Other variants of compression algorithms of dimension, for example, UMAP [2] or T-SNE [3], try to save distances to the nearest neighbours during compression. Despite high operation efficiency, even in the case of strong
nonlinearities, their main disadvantage is the absence of an inverse transformation of any vector of the low-dimensional representation into the original one, which greatly complicates the possibility of using them in prediction systems. Another possibility is the use of artificial neural networks, namely deep autoencoders (AE) [4]. Direct and inverse transformation is created in a natural way, by dividing the network in the place with the minimum dimension of the layer. In addition, AE make it possible to represent very complex nonlinear data spaces, providing a high compression ratio. Therefore, in this paper, AE is selected as a method of map compression. All the obtained results will be compared with the results obtained using PCA.

2. Data description

One of the side effects of the GPS satellite system is the measurement of the total electron content of the ionosphere. These data allow us to improve operation of radio communication devices to ensure greater stability of their work. On the other hand, these data are widely used in studies of the ionosphere. The number of electrons in a vertical column is the most convenient for interpretation. To assess the vertical total electronic content, GIM (Global Ionospheric Maps) technology was developed, based on the processing and interpolation to a rectangular grid of data obtained by the world network of IGS receivers. In this paper, we used global TEC maps obtained by the JPL laboratory due to two factors: availability of a sufficiently large database (we used a set of data obtained in 1998 - 2017), as well as the reproducibility of the equatorial anomaly on the resulting maps.

These data are the global TEC maps. They represent a set of maps, which size is 72 × 72 points (from -180 °C to 180 °C with 5 °C increment in longitude and from -90 °C to 90 °C with 2.5 °C increment in latitude). The maps are updated in 2-hour intervals.

To provide a better simulation of TEC, it is required to perform primary data preprocessing. The first stage is correction of these maps, taking into account the curvature of the Earth. Since the cells on the TEC map correspond to different areas on the Earth’s surface, depending on the latitude, the values for each point of the map were weighed in proportion to the cosine of its latitude, analogous to [5].

The maps also were logarithmized. As the dynamic range of the TEC maps is several orders of magnitude, taking the logarithm will ensure their compactification.

To simplify the analysis of maps and better accounting of the daily periodicity, a transition to a rotating coordinate system was carried out in such a way that the subsolar point was always in the centre of the map. This is equivalent to trying to subtract the 24-hour periodicity in the data.

3. Model Description

It seems reasonable to select convolutional layers as the main transforming and storage elements for the autoencoder. The use of convolutional layers is useful in the case of equal probability of detecting an elementary feature (for example, a sharp change in signal intensity along an axis) anywhere in the image under study. Despite this limitation, neural networks of this type show state-of-the-art results in problems of classification and processing of images, audio and video data ([6], [7], [8]). Thus, several cascades of convolutional layers will form the main part of an ANN encoder. Decoding is provided by a neural network equivalent in power, which sequentially performs inverse transforms until an image with the original size is obtained. In this paper, mean square error between the original and decoded map was selected as the loss function.

The encoder architecture is borrowed from the Dense Convolutions model [9]. In the original article, block cascade connections of convolutional layers were used for the problem of image classification, however it is possible to modify the model architecture to perform compression and decompression of data. Cascade connections make it possible to practically avoid occurrence of the gradient vanishing problem and provide rapid training to all layers of the model. In the model presented in the current work, it was decided to use 3 consecutive blocks of dense convolutions for 4 convolutional layers in each block both for the encoder and decoder.

As an activation function for all convolutional layers, the function LeakyReLU [10] is used. It provides a sufficiently high degree of sparse activation. When using this function, the gradients never
go to zero, so a function of this kind preserves from gradient vanishing \([11]\), which makes it possible to concatenate a large number of layers.

The last layer of the encoder, as well as the first layer of the decoder, are fully connected with linear activation functions. Dimension of this hidden layer \(D\) varied from 64 to 121. All the results presented in this article will be shown for \(D = 81\). Also, due to the technical features of the convolution layers, it is convenient that \(D\) be the square of an integer. In addition, 81 components in the case of map decomposition using the PCA provide an average accuracy of the encoding-decoding process of 95%.

Training of the auto-encoder was carried out using stochastic gradient descent based on the Adam method \([12]\). The batch size of examples, after which the update of the network weights is performed, is 16 pieces. The total number of examples in the training sample is about 64,000, which means that there are about 4000 updates of all weights of the neural network during one learning epoch. To control the occurrence of retraining, the entire sample of data was randomly divided into two subsamples: training and validation at the ratio 9:1. In addition, since a decrease in the learning rate parameter can give a short-term gain in benefits from 2 to 10 times during the network training \([13]\), the validation sample was also used to monitor the reduction of the learning rate. The following tactic was used: the learning rate was halved, in the case when the error on the validation sample did not decrease within at least 6 epochs.

![Figure 1. Comparison of PCA and deep autoencoder results](image1)

![Figure 2. Probability plot of encoding-decoding process errors](image2)
4. Results
To assess the quality of the constructed low-dimensional model, one can compare the results obtained using the proposed technique and the linear PCA.

An example of the map obtained after the compression-decompression process using PCA is shown in Fig. 1. The decoding error in this case turned out to be equal to 7%. The same map, reconstructed with an auto-encoder, is also shown in Fig. 1. It can be seen that the reproduction quality of the equatorial anomaly on the reconstructed map is much higher, if it is compressed by deep neural network.

It is possible to construct the distribution of errors for all maps from the training and test samples for the both presented methods. Here, Fig. 2 shows that a deep autoencoder has an advantage over the linear PCA in all examples.

5. Conclusion
In this paper, a model of global ionospheric TEC maps based on deep neural networks was proposed. Cascade connections of convolutional layers, together with algorithms of stochastic gradient descent, allowed us to obtain a model that exceeds the PCA by an accuracy of up to 2 times at the same dimensionality of the latent space. In addition, the number of variable network parameters proved to be less than in the case of using PCA. Another important feature is the reproduction of the presence of an equatorial anomaly on the decoded maps by the neural network model. Thus, the presented model can be used in forecasting systems of global TEC maps.

6. Acknowledgments
This research was financially supported by the Russian Government Program of Competitive Growth of Kazan Federal University, state assignment of Ministry of Education and Science, grant agreement № 34.5517.2017/6.7.

References
[1] Pearson K 1901 On Lines and Planes of Closest Fit to Systems of Points in Space Philosophical Magazine. 2(11): 559-572
[2] McInnes L, Healy J 2018 UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction arXiv:1802.03426v1
[3] Sanjeev A, Wei H, Pravesh K 2018 An Analysis of the t-SNE Algorithm for Data Visualization arXiv:1803.01768v2
[4] Charté D, Charté F, García S, María J, Francisco H 2018 A practical tutorial on autoencoders for nonlinear feature fusion: Taxonomy, models, software and guidelines arXiv:1801.01586
[5] Bochkarev V, Maslennikova Y. 2014 Principal component analysis of global maps of the total electronic content Geomagnetism and Aeronomy 54(2) 216-223
[6] Hyeonwoo N, Seunghoon H, Bohyung H 2015 Learning Deconvolution Network for Semantic Segmentation arXiv:1505.04366v1
[7] van den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior A, Kavukcuoglu K 2016 wavenet: a generative model for raw audio. arXiv:1609.03499v2,
[8] Kim Y, Jernite Y, Sontag D, Rush A. 2015 Character-Aware Neural Language Models, arXiv:1508.06615
[9] Huang G, Liu Z, van der Maaten L, Weinberger K. 2018 Densely Connected Convolutional Networks, arXiv:1608.0693v5
[10] Xu B, Wang N, Chen T, Li M. 2015 Empirical Evaluation of Rectified Activations in Convolutional Network. arXiv:1505.00853
[11] Razvan P, Tomas M, Yoshua B. 2012 On the difficulty of training Recurrent Neural Networks. arXiv:1211.5063
[12] Kingma D, Ba J. 2015 Adam: a Method for Stochastic Optimization. International Conference on Learning Representations, 1-13.

[13] Smith L. 2018 a disciplined approach to neural network hyper-parameters: part 1 - learning rate, batch size, momentum, and weight decay arXiv:1803.09820v2