Herausgeber et al. (Hrsg.): Datenbanksysteme für Business, Technologie und Web (BTW 2023),
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2023 11

Meduse: Interactive and Visual Exploration of Ionospheric Data

Joshua Reibert[1], Arne Osterthun[2], Marcus Paradies[3]

Abstract: Spatio-temporal models of ionospheric data are important for atmospheric research and the evaluation of their impact on satellite communications. However, researchers lack tools to visually and interactively analyze these rapidly growing multi-dimensional datasets that cannot be entirely loaded into main memory. Existing tools for large-scale multi-dimensional scientific data visualization and exploration rely on slow, file-based data management support and simplistic client-server interaction that fetches all data to the client side for rendering.

In this paper we present our data management and interactive data exploration and visualization system Meduse. We demonstrate the initial implementation of the interactive data exploration and visualization component that enables domain scientists to visualize and interactively explore multi-dimensional ionospheric data. Use-case-specific visualizations additionally allow the analysis of such data along satellite trajectories to accommodate domain-specific analyses of the impact on data collected by satellites such as for global navigation satellite systems and earth observation.

Keywords: Exploratory Data Analysis; Ionospheric Data; Data Cubes

1 Introduction

The Earth's ionosphere is observed by an ever growing amount of sensors–both ground- and space-based–which results in a giant corpus of raw data [Ca20]. The research of this space weather is important since space weather events (e.g., solar storms) can affect the quality of satellite communication which in turn can impact services like global navigation satellite systems (GNSS) such as the widely used global positioning system (GPS) [SJH19]. Hence, scientists have taken on harmonizing and integrating the raw measurement data into dense models of the ionosphere to relate noise in satellite-based data back to space weather.

The resulting spatio-temporal model data quickly becomes unwieldy as the data volume grows, especially in the temporal dimension. Hence, researchers oftentimes work on large sets of small files and resort to static plots of small regions of interest which is cumbersome and limits exploration. Interactive visualization allows users to visually explore the data

[1] German Aerospace Center (DLR), Institute of Data Science, Mälzerstr. 3 – 5, 07745 Jena, Germany joshua.reibert@dlr.de
[2] German Aerospace Center (DLR), Institute of Data Science, Mälzerstr. 3 – 5, 07745 Jena, Germany arne.osterthun@dlr.de
[3] German Aerospace Center (DLR), Institute of Data Science, Mälzerstr. 3 – 5, 07745 Jena, Germany marcus.paradies@dlr.de

and quickly gain new insights but it highly depends on the system's latency which impacts analysis performance [LH14]. As a consequence, adequate data management technologies to store and access the spatio-temporal model data are required to support scientists in their tasks. Besides providing fast access to the data in the data management system, well-known techniques to reduce latency, e.g., client-side caching, binary & compressed data transfers, progressive visualization, and execution of native code on the client side, reduce the overall execution time of interactive exploration queries. To the best of our knowledge, currently no such data management and interactive exploration system for spatio-temporal models exists. VirES[4] provides some of the features but is focused on specific satellite products and is hard to extend to model data.

In this paper we present an initial design of the MEDUSE eco system consisting of a data management backend exposing a data cube data model and an interactive data exploration and visualization component specifically designed to interactively and visually explore and analyze ionospheric model data. MEDUSE builds on a custom data backend to query metadata and provide fast access to the underlying multi-dimensional data. The data backend decouples the physical data layout, i.e. files in specific data formats such as NetCDF, TileDB, or Zarr, from the logical layout and data model based on the concept of *data cubes* [Ba17]. The model data is visualized along the spatial and temporal dimensions–each visualizing individual slices of the data. This allows users to iteratively navigate the data which is adaptively loaded to reduce latency and provide fluid interaction. Specialized features like showing data along satellite trajectories are tailored to expert domain users that want to analyze the ionosphere's impact on satellite communications.

2 Background

The ionospheric models in our use case are developed by domain experts from the atmospheric sciences [HJP22]. They are based on ground- and space-based measurements to compute a dense grid of earth-centric electron density data. In addition to the two common spatial dimensions of longitude and latitude, this also includes the vertical altitude dimension and a temporal dimension, thereby resulting in a four-dimensional dataset.

The model includes the *electron density* (Ne) in the ionosphere with these four dimensions and three additional, derived variables which omit the altitude dimension and are hence only three-dimensional: The *vertical total electron content* (VTEC) is the integral of the electron density along the altitude dimension. The maximum electron density along the altitude dimension is the *peak density* (NmF2) and the corresponding altitude the *peak density height* (hmF2). The spatial resolutions are relatively low with 72 values for longitude and latitude and 112 altitude steps. However, the data can be computed for up to 5-minute intervals which quickly adds up when larger time spans are processed. While the data for a single point in time makes up 2.4 MiB, a year worth of data amounts to 250 GiB. The data

[4] `https://earth.esa.int/eogateway/tools/vires-for-aeolus`

size is then again multiplied by the number of models that are considered. Currently, we consider two models in our application.

3 Interactive Visual Analysis of Ionospheric Models in MEDUSE

Figure 1 depicts the overall architecture of MEDUSE consisting of a web-based client and the data management backend storing all the model data.

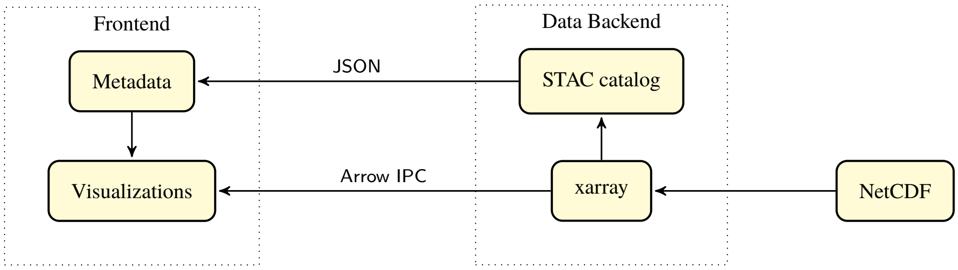

Fig. 1: MEDUSE architecture overview.

Data Backend

The data for the ionospheric models are provided in the NetCDF format from which metadata is generated in the spatio-temporal asset catalogs (STAC)[5] format in a pre-processing step. It encompasses data dimensions, axis types, coordinates, extreme values, and free-form metadata, such as textual descriptions. A custom backend provides endpoints to query this STAC metadata, queries for dicing and slicing using dimension ranges as well as point queries for specific dimension value tuples. All of the queries are validated against the metadata. The actual data access is performed using xarray[6] and results are sent over the network in the inter-process communication (IPC) format of Apache Arrow[7] via HTTP. Furthermore, aggregations can be queried directly by specifying the operation to compute and the dimensions to apply them on. To support querying along satellite trajectories, point queries interpolate the model data linearly.

Frontend

We developed a web application that builds on the metadata and the data cubes to enable analysts and other domain users to quickly and interactively explore the ionospheric data.

[5] `https://stacspec.org`
[6] `https://docs.xarray.dev/`
[7] `https://arrow.apache.org/`

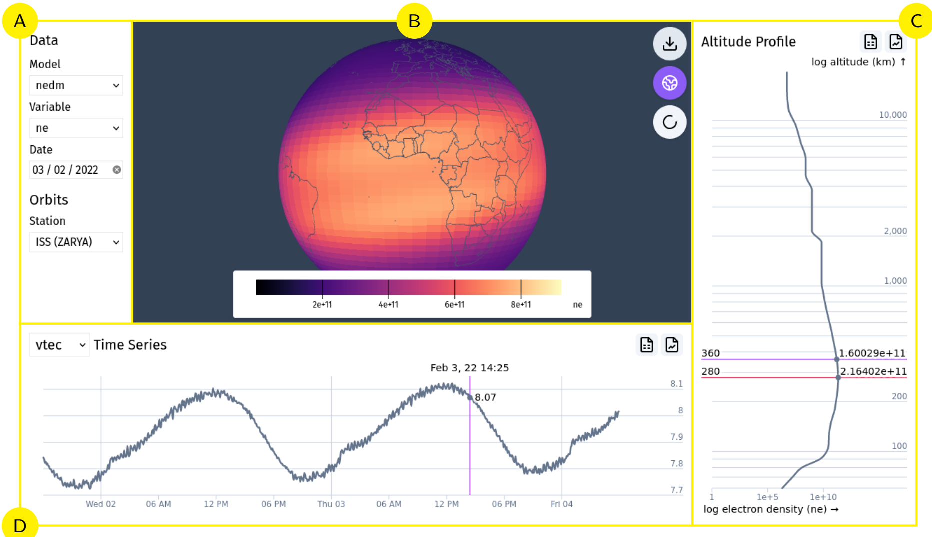

Fig. 2: The web application visualizes the ionospheric data along the different dimensions. Different models, variables, and dates in that dataset can be selected (A). The map visualizes the selected variable along the spatial dimensions (B), the altitude profile visualizes the electron density along the vertical altitude dimension (C), and the time-series chart visualizes a variable along the temporal dimension (D).

Figure 2 depicts a screenshot of the interactive data exploration and visualization tool of Meduse. It features a map or globe visualization for the horizontal spatial dimensions as well as an altitude profile for the vertical spatial dimensions, and a time-series chart for the temporal dimension. All of these visualizations are linked such that users can select specific dimension values in them, for example a certain timestamp in the time-series, and the other visualizations will update accordingly. As a result, users can directly and interactively select dimension values in the visualizations to explore regions of interest in the data instead of relying on external widgets.

Users can initially select one of the provided models, a variable, and a date of interest (see Figure 2 (A)). The central visualization shows the color-encoded data along the horizontal spatial dimensions mapped to Earth with country borders to give a frame of reference (see Figure 2 (B)). It can be toggled between a 3D globe view and 2D map with a dedicated button and uses WebGL for real-time rendering and interaction. Users can select individual grid cells to focus on that spatial region. The altitude profile visualizes the electron density for the selected timestamp along the altitude dimension (see Figure 2 (C)). It highlights the peak density value as well as the respective altitude and can be used to interactively inspect altitude and density values by hovering, and to select a different altitude value. The time-series visualizes how a selected variable changes over time in a 60 hours window around the selected date (see Figure 2 (D)). Again, specific values and the respective timestamp are

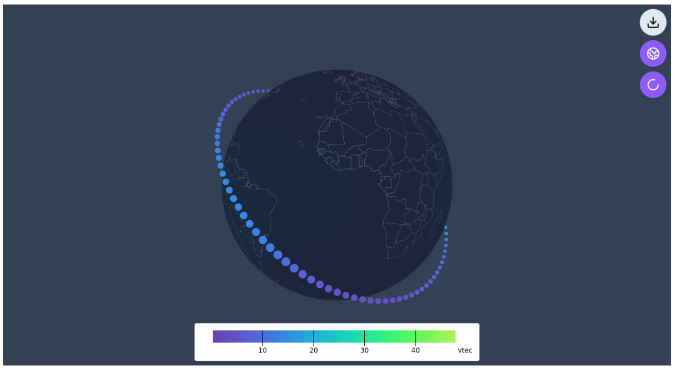

Fig. 3: The satellite trajectory visualization shows the trajectory of a selected satellite and color-codes variable values along it. The time-series chart simultaneously plots the data as a line chart.

shown when hovering the chart and also allow interactively selecting a different timestamp that is propagated to the other visualizations.

Since the electron density affects the quality of satellite communications, satellite operators and users of GNSS or earth observation are often interested in investigating electron density along satellite paths. We accommodate this use case by including a set of satellites and allowing users to select one. Satellite data is provided as two-line element set (TLE) which is used to compute the earth-centric 3D position from given timestamps. The resulting set of values for all four dimensions is used to query the data from the data cube. The satellite trajectory can then be toggled to be shown on the map or the globe and again color-coding data values along its path (see Figure 3).

The application aims to allow researchers working with ionospheric data in validating their models and investigating effects of the electron density on collected data from satellites. However, this is often just a first step to discover relations and anomalies that entail further, more focused analyses. Hence, we also provide the functionality to export the underlying data of visualizations for further inspection and the charts as images for presentation.

4 Demonstration Outline

For the demonstration, we will assume the role of a researcher working with earth observation data who has encountered unusual levels of noise in their data. We will use the web application to interactively explore the ionospheric models to look for unusual ionospheric activity. Finally, we will take a look at the satellite's path to analyze the electron density along its path and identify anomalies.

5 Summary & Outlook

In this paper we presented the initial design of MEDUSE, a data management and interactive exploration and visualization system for spatio-temporal model data from the ionosphere. In the future, we want to continue to improve our prototype to reduce latency, e. g. by using web assembly for client-side data cube access and computations as well as WebSockets to reduce network latency. Furthermore, we want to generalize the data backend to be more generally applicable to data-intensive web applications working with multi-dimensional data beyond atmospheric sciences.

References

[Ba17] Baumann, P.: Standardizing big earth datacubes. In (Nie, J.; Obradovic, Z.; Suzumura, T.; Ghosh, R.; Nambiar, R.; Wang, C.; Zang, H.; Baeza-Yates, R.; Hu, X.; Kepner, J.; Cuzzocrea, A.; Tang, J.; Toyoda, M., eds.): 2017 IEEE International Conference on Big Data (IEEE BigData 2017), Boston, MA, USA, December 11-14, 2017. IEEE Computer Society, pp. 67–73, 2017.

[Ca20] Carley, E. P.; Baldovin, C.; Benthem, P.; Bisi, M. M.; Fallows, R. A.; Gallagher, P. T.; Olberg, M.; Rothkaehl, H.; Vermeulen, R.; Vilmer, N.; Barnes, D.; the LOFAR4SW Consortium: Radio observatories and instrumentation used in space weather science and operations. J. Space Weather Space Clim. 10/7, 2020.

[HJP22] Hoque, M. M.; Jakowski, N.; Prol, F. S.: A new climatological electron density model for supporting space weather services. J. Space Weather Space Clim. 12/1, 2022.

[LH14] Liu, Z.; Heer, J.: The Effects of Interactive Latency on Exploratory Visual Analysis. IEEE Trans. Vis. Comput. Graph. 20/12, pp. 2122–2131, 2014.

[SJH19] Su, K.; Jin, S.; Hoque, M. M.: Evaluation of Ionospheric Delay Effects on Multi-GNSS Positioning Performance. Remote. Sens. 11/2, p. 171, 2019.