A Collaborative Data Science Platform for the Geospace Community: Assimilative Mapping of Geospace Observations (AMGeO) v1.0.0

AMGeO Collaboration

December 7, 2019

Website: https://amgeo.colorado.edu
Email: amgeo@colorado.edu

1 Purpose

The purpose of Assimilative Mapping of Geospace Observations (AMGeO) is to make the latest geospace data science tool accessible to scientists and students. AMGeO is a collaborative data science platform for bringing together a diverse set of heterogeneous geospace observations from NSF-funded facility programs and individual community users to obtain complete maps of high-latitude ionospheric electrodynamics for scientific discovery and space weather research. The platform is made of the AMGeO open-source software and web application service that facilitate the data acquisition and pre-processing steps that are otherwise prohibitively labor-intensive. It is developed at the University of Colorado Boulder by the AMGeO Collaboration, with support from the NSF Earth Cube program (ICER 1928403).

2 Functionality

The AMGeO software is designed to streamline data pre-processing, quality control steps, and data assimilation analysis steps in open-source package to support accessible, reproducible and transparent data science practices in the geospace science community. The assimilative mapping technique implemented in the AMGeO software modernizes and improves on the Assimilative Mapping of Ionospheric Electrodynamics [Richmond and Kamide, 1988] as summarised in Matsuo [2020]. The AMGeO web application service, which is the API web application software served on the University of Colorado Boulder's cloud computing resources, automates data retrieval from third-party data services, including SuperDARN (http://supermag.jhuapl.edu), SuperMAG (http://supermag.jhuapl.edu), and NASA's Space Physics Data Facility (SPDF) (https://spdf.gsfc.nasa.gov).

2.1 AMGeO Software

Figure 1 shows the entire AMGeO v1.0.0 software workflow. This flowchart demonstrates how a typical AMGeO spatial prediction of high-latitude electrodynamics is conducted. While the AMGeO web application service handles the details of data collection, there are several additional data pre-processing steps performed by the AMGeO software. The AMGeO v1.0.0 software pre-processes SuperDARN plasma drift data and ground-based magnetometer data distributed from the SuperMAG data service for use in assimilative mapping described in Matsuo [2020].

2.1.1 SuperDARN Data Pre-processing

The AMGeO software uses the SuperDARN grid data product [Chisham et al., 2007], which provides ion drift velocity measured along the line-of-sight for all available radars at given time. Each individual record

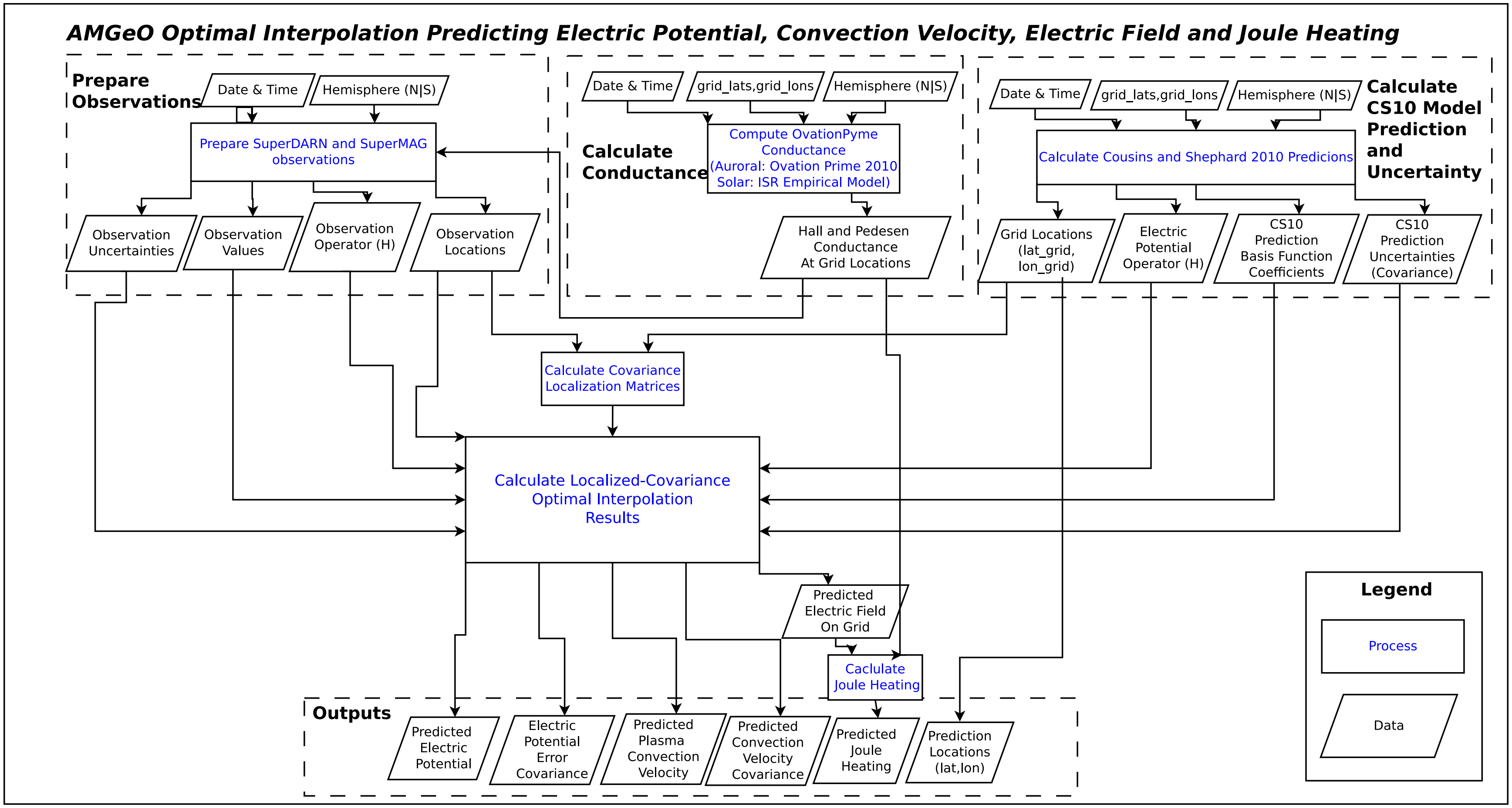

Figure 1: Flowchart of the entire AMGeO v1.0.0 software workflow. Rectangles with blue text represent processing steps, while parallelograms with black text represent data which processing steps take as inputs and produce as outputs. AMGeO starts by retrieving three main elements: (1) the prior mean and error covariance from an empirical model (top-right), (2) auroral height-integrated conductivity from the Ovation Prime empirical model [Newell, Sotirelis and Wing, 2009] (top-center), and (3) SuperDARN and SuperMAG data and associated error covariances (top-left). The prior covariance is localized around the observation locations (center). The localized covariance, model predictions and observations are combined to yield estimates of ionospheric electrodynamic states (e.g., electrostatic potential, electric field, and plasma convection) and analysis error covariances for each predicted quantity. The predicted electric fields are combined with the empirical conductance to calculate Joule heating rate (bottom-center).

in the grid data files contains the ion velocity measurements during a two-minute window of time, averaged into a regular latitude and longitude grid. A ion drift velocity magnitude, an uncertainty, and the radar line-of-sight angle (with respect to local magnetic north) are provided for each grid cell. Figure 2 shows all pre-processing steps of SuperDARN data for use in AMGeO. AMGeO converts SuperDARN line-of-sight plasma velocity to electric field ($\vec{E}$), which is related to ion drift velocity ($\vec{v}$) and the geomagnetic main field ($\vec{B}$) as $\vec{E} = -\vec{v} \times \vec{B}$. This calculation is simplified by assuming the ion drift velocity has no upward component and the geomagnetic field orientation is radial in high-latitudes. The measurement uncertainty is propagated from line-of-sight plasma velocity to electric field in a similar manner and scaled accordingly with the number of observations as described in Cousins, Matsuo and Richmond [2015].

2.1.2 SuperMAG Data Pre-processing

The SuperMAG data service aggregates data from a global network of ground-based magnetometers from individual magnetometer operators and consortia, rotates each measurement vector into a consistent coordinate frame, and removes signals not related to ionospheric currents [Gjerloev, 2012]. The AMGeO software uses the Baseline Subtracted SuperMAG data product. The default data product has one measurement per magnetometer per minute. Figure 3 shows all pre-processing steps of SuperMAG data for use in AMGeO. After applying a correction for partial ring current effects, the magnetic perturbation vectors are transformed from geographic coordinates to Modified Magnetic Apex coordinates [Richmond, 1995]. Ground-level magnetic perturbation observations are then converted to the ionospheric equivalent currents. In order to build in the relationship between the ionospheric equivalent currents and electric field according the the Ohm's law in the forward operator, the estimate of the height-integrated ionospheric conductivity is required.

2.2 AMGeO Web Application Service

As shown in Figure 4, the AMGeO web application service works in tandem with the AMGeO software to route users' data requests to the third-party data providers. The AMGeO software and web application service are configured to ensure that traffic from the AMGeO web application service appears to geospace community data providers as though it were coming directly from AMGeO users. This is important because data providers often require that each user creates an account on their website before downloading data to collect their own user statistics. By automating this interface, the AMGeO web application service provides seamless access while supporting the needs of the data providers. They also help mitigate the burden of supporting individual AMGeO users, and count the number and nature of data requests, which can be used to gauge community adoption of the AMGeO software and fulfill our funding reporting obligations.

3 License and Attribution

The AMGeO software is written in Python and is available for researchers and students to use and modify upon registering through the AMGeO website and agreeing to the terms and conditions of the license agreement and the privacy and data policy. Under this licence agreement, no commercial use is permitted. Commercial use will require contacting the AMGeO Collaboration for a separate licence agreement. In any publication of the results of research performed using the AMGeO software or a modified version of the AMGeO software, attribution to the AMGeO Collaboration is required by citing the appropriate digital object identifier for the version of the AMGeO software used. For the AMGeO v1.0.0, this technical report (doi:10.5281/zenodo.3564914) should be referenced.

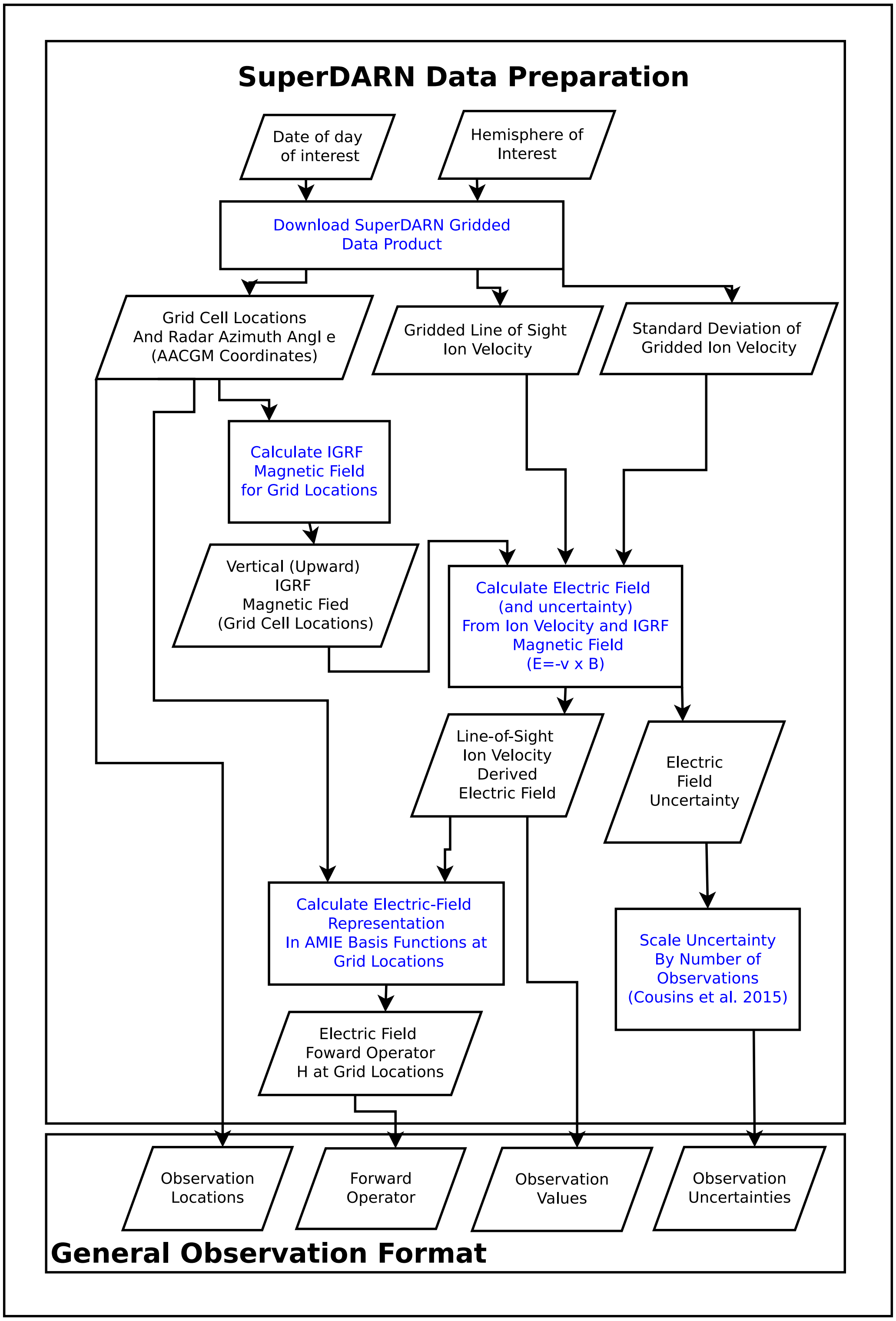

Figure 2: Flowchart of the SuperDARN data pre-processing workflow. Rectangles with blue text represent processing steps, while parallelograms with black text represent data which processing steps take as inputs and produce as outputs. Beginning at the top of the figure, one day of data for either the northern or southern hemisphere are downloaded. The International Geomagnetic Reference Field model [Thébault et al., 2015] is used to calculate the geomagnetic field at the center of each grid cell. The SuperDARN line-of-sight plasma velocity is converted to electric field, and the observation uncertainty is computed. Finally, the forward operator for electric field observations is computed.

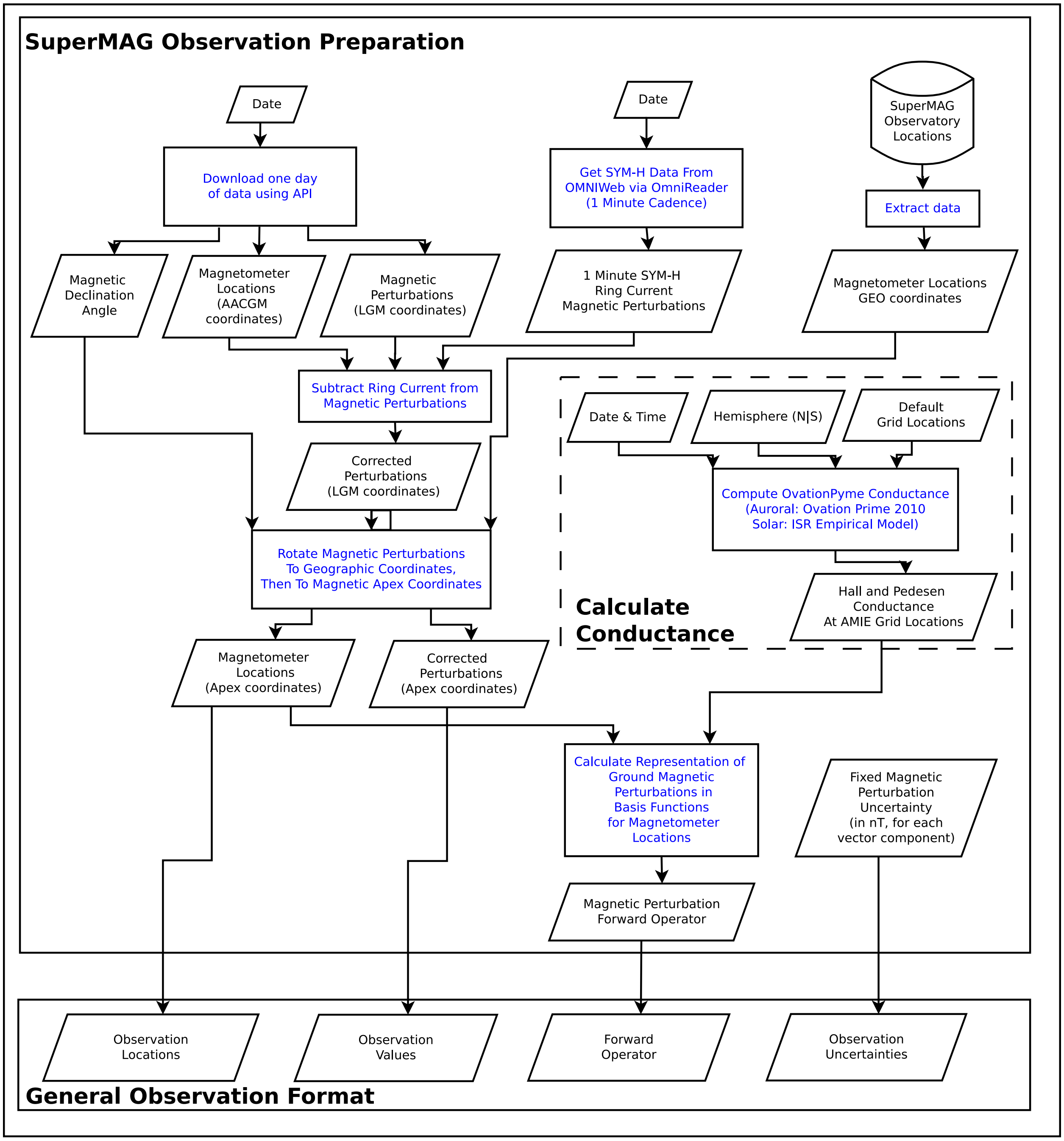

Figure 3: Flowchart of the SuperMAG data pre-processing workflow. Rectangles with blue text represent processing steps, while parallelograms with black text represent data which processing steps take as inputs and produce as outputs. First, the Baseline Subtracted SuperMAG data product is downloaded for both hemispheres for one day. The partial ring current effects are first removed using the geomagnetic declination angle for each magnetometer location and the SYM-H index obtained from NASA SPDF OMNIWeb. The magnetic perturbation vectors are then transformed in Modified Magnetic Apex coordinates [Richmond, 1995]. This step requires the geographic location of each magnetometer, which is read in from the SuperMAG magnetometer stations file. The estimate of the height-integrated ionospheric conductivity caused by auroral particle precipitation and solar EUV radiation is required for building in the relationship between the magnetic perturbation vectors and the electric field in the forward operator. Finally, the forward operator for magnetic perturbation observations is computed, and an empirically chosen uncertainty of 50 nT is assigned to each vector component of the measurement.

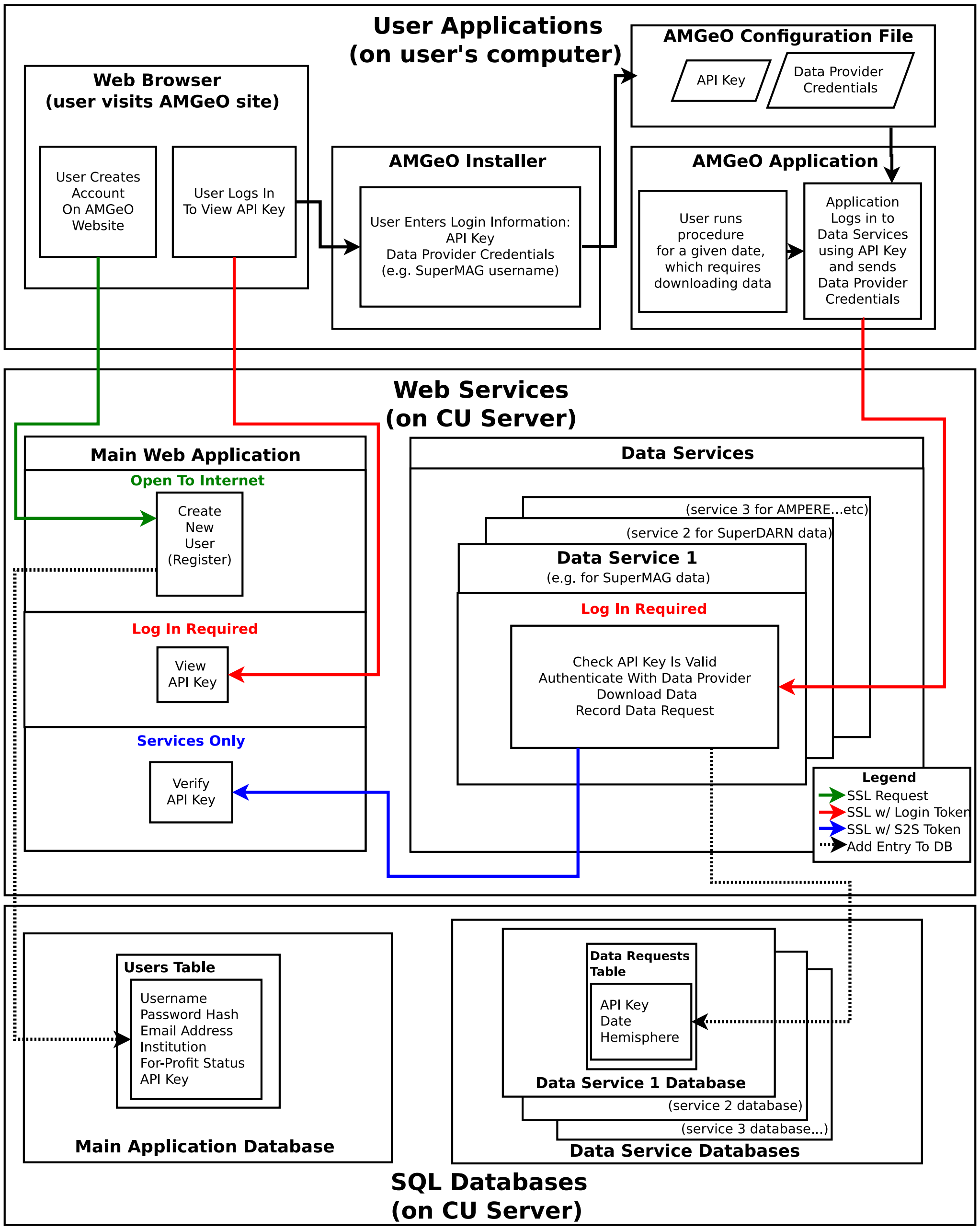

Figure 4: Flowchart of the AMGeO web application service's workflow. The topmost panel represents the actions performed by a user running the AMGeO software, whereas the center and bottom panels represent what happens on the AMGeO web application service on the University of Colorado Boulder's cloud computing resources. Beginning from the top-left, a user who wants to use the AMGeO software first visits the AMGeO website and registers for an account. Once they have finished the registration process, they can log into their account and to receive an API key (a long random string which is unique for each user). Then, when the user installs the AMGeO software, they will be prompted to enter their API key, along with any other credentials required by the Third Party data providers (currently only a username for the SuperMAG website). The API key and credentials are stored in a configuration file and sent to the web service when a user runs the AMGeO software. When the web service receives a request for data, it first checks the API key to see if it belongs to a registered user. If this check is successful, it then sends any necessary credentials to the data provider along with the specifics of what data the user requested. Finally, the web service routes the data file from the provider to the user. After the data request has completed successfully, the data service records the request in its database, identifying the user only by their API key.

References

Chisham, G., M. Lester, S. E. Milan, M. P. Freeman, W. A. Bristow, A. Grocott, K. A. McWilliams, J. M. Ruohoniemi, T. K. Yeoman, P. L. Dyson, R. A. Greenwald, T. Kikuchi, M. Pinnock, J. P. S. Rash, N. Sato, G. J. Sofko, J.-P. Villain and A. D. M. Walker. 2007. "A decade of the Super Dual Auroral Radar Network (SuperDARN): scientific achievements, new techniques and future directions." *Surveys in Geophysics* 28(1):33–109.
URL: *https://doi.org/10.1007/s10712-007-9017-8*

Cousins, Matsuo and Richmond. 2015. "Mapping high-latitude ionospheric electrodynamics with SuperDARN and AMPERE." *Journal of Geophysical Research: Space Physics* 120(7):5854–5870.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2014JA020463*

Gjerloev, J. W. 2012. "The SuperMAG data processing technique." *Journal of Geophysical Research: Space Physics* 117(A9).
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2012JA017683*

Matsuo, Tomoko. 2020. Recent Progress on Inverse and Data Assimilation Procedure for High-Latitude Ionospheric Electrodynamics. In *Ionospheric Multi-Spacecraft Analysis Tools: Approaches for Deriving Ionospheric Parameters*, ed. Malcolm Wray Dunlop and Hermann Lühr. ISSI Scientific Report Series Cham: Springer International Publishing pp. 219–232.
URL: *https://doi.org/10.1007/978-3-030-26732-*$2_1$0

Newell, P. T., T. Sotirelis and S. Wing. 2009. "Diffuse, monoenergetic, and broadband aurora: The global precipitation budget." *Journal of Geophysical Research: Space Physics* 114(A9):A09207.
URL: *http://onlinelibrary.wiley.com/doi/10.1029/2009JA014326/abstract*

Richmond, A. D. 1995. "Ionospheric Electrodynamics Using Magnetic Apex Coordinates." *Journal of geomagnetism and geoelectricity* 47(2):191–212.

Richmond, A. D. and Y. Kamide. 1988. "Mapping electrodynamic features of the high-latitude ionosphere from localized observations: Technique." *Journal of Geophysical Research: Space Physics* 93(A6):5741–5759.
URL: *https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JA093iA06p05741*

Thébault, Erwan, Christopher C. Finlay, Ciarán D. Beggan, Patrick Alken, Julien Aubert, Olivier Barrois, Francois Bertrand, Tatiana Bondar, Axel Boness, Laura Brocco, Elisabeth Canet, Aude Chambodut, Arnaud Chulliat, Pierdavide Coïsson, François Civet, Aimin Du, Alexandre Fournier, Isabelle Fratter, Nicolas Gillet, Brian Hamilton, Mohamed Hamoudi, Gauthier Hulot, Thomas Jager, Monika Korte, Weijia Kuang, Xavier Lalanne, Benoit Langlais, Jean-Michel Léger, Vincent Lesur, Frank J. Lowes, Susan Macmillan, Mioara Mandea, Chandrasekharan Manoj, Stefan Maus, Nils Olsen, Valeriy Petrov, Victoria Ridley, Martin Rother, Terence J. Sabaka, Diana Saturnino, Reyko Schachtschneider, Olivier Sirol, Andrew Tangborn, Alan Thomson, Lars Tøffner-Clausen, Pierre Vigneron, Ingo Wardinski and Tatiana Zvereva. 2015. "International Geomagnetic Reference Field: the 12th generation." *Earth, Planets and Space* 67(1):79.
URL: *https://doi.org/10.1186/s40623-015-0228-9*