The Space Weather impact on Arctic Navigation (SWAN) project Development of operational forecasts of space weather impacts on GNSS

Knut Stanley Jacobsen, Vincent Fabbro, Aymeric Mainvis, Igor Zakharov, Matthew Angling, Giorgio Savastano, Pel Lundahl Thomsen, Michael Linden-Vornle, Arnlaug Hogas Skjaeveland, Jonas Bregnhoj Nielsen, et al.

The Space Weather impact on Arctic Navigation (SWAN) project

Development of operational forecasts of space weather impacts on GNSS

Knut Stanley Jacobsen, Vincent Fabbro, Aymeric Mainvis, Igor Zakharov, Matthew Angling, Giorgio Savastano, Per Lundahl Thomsen, Michael Linden-Vørnle, Arnlaug Høgås Skjæveland, Jonas Bregnhøj Nielsen, Mohammed Ouassou

Abstract – The forecast of Space Weather events is becoming more and more important as the reliance on space-based systems is increased. Reliable and timely forecasts can allow mitigating actions to be taken in order to reduce disruptive or harmful consequences of space weather disturbances. The objective of the Space Weather impact on Arctic Navigation (SWAN) project is the development of forecasts for impacts of space weather disturbances on satellite navigation services in the Arctic region. The SWAN project is part of the efforts of the European Space Agency (ESA) to provide monitoring and forecasts of space weather effects in Europe. Two approaches have been considered for the development of operational forecasts; An empirical approach based on prior work by project partners, and a novel machine learning approach. This paper reports on the developments performed during the project, the results obtained from the validation of the developed approaches to forecasts, and the forecast service which is the main result of the project.

Introduction

There are several factors which influence the accuracy of a position estimated by a Global Navigation Satellite System (GNSS) device [1]. However, during strong ionospheric activity, the ionosphere is the dominant error source for GNSS signals.

Ionospheric scintillation refers to the random amplitude and phase fluctuations observed in radio signals propagating through electron density irregularities in the ionospheric plasma and most commonly occurs in equatorial, auroral, and polar regions [2], [3]. The scintillations are caused by the presence of a wide range of scale size irregularities in the ionosphere. The irregularities causing scintillation may be considered as an irregular, spatial distribution that is drifting but also evolving in time. The amplitude scintillations come from irregularities having a scale size of the first Fresnel zone, which depends on the signal wavelength and average height of the irregularity layer [4], [5]. For GPS, this scale is about 370-400 m. Large-scale plasma density structures in the ionosphere with scale sizes greater than the first Fresnel zone do not cause significant amplitude scintillation but may still cause fluctuations in the signal phase [5]. Under such conditions, refraction effects must be taken into account and the phase fluctuations are due to the optical path changes of a radio wave [6]. The low-frequency GPS phase fluctuations may be directly due to electron density changes along the radio ray path.

The occurrence of scintillation at high latitudes is related to the auroral oval, cusp, and polar-cap patches, through the formation of small-scale plasma structures [7]–[17]. It has been observed that phase scintillation occurs more often than amplitude scintillation at high latitudes, and that scintillation is more common on geomagnetically disturbed days in the auroral oval region and close to noon and midnight [12], [13], [16]–[19]. Phase scintillations at high latitudes are associated with particle precipitation and plasma patches [20]–[23] and the scale of the processes (i.e. the spatial extent of the region of irregularities) can be substantially enlarged during geomagnetic storms and substorms [24], [25].

Calculating disturbance indices from GNSS measurements is a proven method for the detection and near real-time monitoring of the occurrence and dynamics of ionospheric irregularities. High-rate (sampling rates of 50 Hz or more) GNSS receivers can be used to calculate the variations of the signal phase and amplitude, quantifying them by the phase and amplitude scintillation indices (σ_φ and S_4 [26]. The Rate Of Total electron content Index (ROTI) is another measure of disturbances in the phase measurements and can be calculated based on lower-rate data. Thus it can be based on data from normal geodetic GNSS receiver networks, of which there exists substantial data archives. ROTI has been shown to be a useful tool for characterizing the ionospheric disturbance level of GNSS measurements [27]–[31].

Tracking errors imposed in the presence of ionospheric scintillation can exceed receiver thresholds and result in signal loss of lock. Loss of satellite signals will degrade the positioning solution. In severe cases, the receiver may be unable to maintain lock on at least four GNSS satellites and a temporary loss of positioning capabilities can occur. The duration and frequency of such outages depend on the duration and severity of the disturbances, the geometry of the satellites in view and the signal reacquisition time [32]. Even if the signal is not completely lost rapid phase variation may cause cycle slips, which are discontinuities in a receiver's phase lock on a satellite's signal.

[33] quantified the global-scale impacts of ionospheric disturbances on GPS Precise Point Positioning (PPP) solutions during an extreme space weather event on the 17^{th} of March 2015 by taking advantage of more than 5,500 GNSS stations installed worldwide. The overall impact was more severe at high latitudes. Specifically, the results show that kinematic PPP solutions degraded following an intensified auroral particle precipitation during the storm's main phase (06-23 UTC) when up to $\sim 70\%$ of the high-latitude stations experienced degraded position solutions in the range of several meters at 16-18 UTC. Around magnetic noon and midnight, the storm-induced plasma irregularities caused notable PPP errors (> 10 m) at high latitudes.

[34] demonstrated how the PPP 3D position error correlates with ROTI at high latitudes. The risk of having several satellites observing enhanced ROTI values simultaneously is greater at higher latitudes. Several receivers (FOLC, VEGS, TRO1, HAMC, LYRS, and NYA1) were analyzed and 3D position error up to 15 cm was observed.

Occurrence of scintillation is difficult to predict and model due to the variability of its numerous influencing factors, which include solar activity, interplanetary magnetic field orientation, local electric field and conductivity, ionospheric convection processes, and wave interactions [35].

The ESA Space Weather Service Network (https://swe.ssa.esa.int/) provides several space weather monitoring services. Available types of data that may be relevant for GNSS users include TEC maps, ROTI maps, scintillation indices, and limited forecasted TEC maps. In addition, other data are available that could be used as an input for algorithms to estimate impacts on PNT/GNSS systems (Positioning, Navigation and Timing GNSS systems), broadly categorized as solar monitoring data, solar wind monitoring data, geomagnetic conditions monitoring data and solar flare forecasts (in the form of current probability of occurrence).

Examples of other space weather services providing global and/or local monitoring are the US Space Weather Prediction Center (SWPC) (https://www.swpc.noaa.gov/) , Norwegian Centre for Space Weather (NOSWE) (https://site.uit.no/spaceweather/) , Space Weather Canada (https://spaceweather.gc.ca/index-en.php) , Russian space weather services (http://ipg.geospace.ru/) , the UK Met Office Space Weather Operations Centre (MOSWOC) (https://www.metoffice.gov.uk/weather/specialist-forecasts/space-weather) , International Space Environment Service (ISES) (http://www.spaceweather.org/). The types of ionospheric monitoring data are similar for many of these services, but as they are based on different measurement networks their coverage and resolution varies.

In these operationally available services there is little focus on measures of impact on GNSS end users. Also, forecasts of ionospheric variations (σ_φ, S_4, $ROTI$ or similar) are very limited.

Consequently, the main objective of the SWAN project was to "*... develop and demonstrate new techniques, which are expected to improve on the current state-of-the-art Space Weather forecasting techniques and products related to Space Weather effects for Positioning, Navigation and Timing (PNT) systems in the Arctic, with particular attention to the Greenland area.*". To achieve this, two approaches have been investigated:

- Empirical modelling of ROTI statistics, by modifying and extending the HAPEE model [31].
- Forecasting ROTI time series from solar wind inputs, through a machine learning algorithm.

At the end of the project, the machine learning approach is still considered as not mature enough to be used for operational services. Therefore, a forecast service was established using the HAPEE model.

Data sources

This section describes the data sources and the steps taken to check and prepare the data before using it for algorithm development.

Solar wind data

High-resolution solar wind magnetic field, density and velocity data was extracted from NASA/GSFC's OMNI data set through OMNIWeb [36].

We applied a 10-minute running window average to the solar wind data, in order to reduce the impact of rapid solar wind variations and data gaps. Then we calculated the coupling function $\frac{d\Phi_{MP}}{dt}$ as described in [37]. The value of this coupling function will be used as the input to the forecast algorithms, as a proxy index of space weather activity in the ionosphere at high latitudes.

ROTI from GNSS receivers

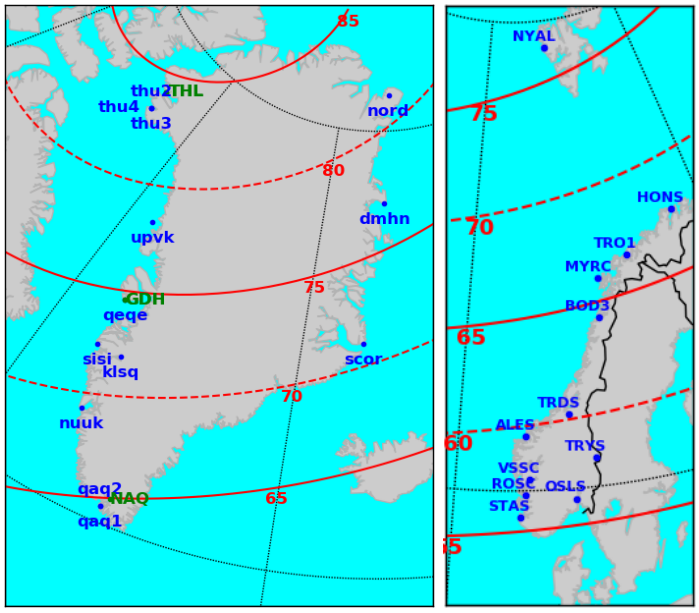


Figure 1: Maps showing the location of the GNSS receivers used in this work, on Greenland (GNET) and in Norway (NMA). The red lines show the geomagnetic latitude.

Figure 1 shows the locations of the GNSS receivers used to provide ROTI data for the project, which are subsets of the Norwegian Mapping Authority (NMA) network and the Greenland GNSS Network (GNET). The receivers are geodetic quality, multi-frequency, and sampling observations at 1 Hz. The ROTI data from NMA used in the framework of this project corresponds to the period [2007-2019], covering the 24^{th} solar cycle. The ROTI data from GNET used in the framework of this project corresponds to the period [2010-2020], covering most of the 24^{th} solar cycle.

After quality checks, including clock jump and cycle slip detection, ROTI has been calculated every 5 minutes for each observed satellite. To detect cycle slips, the change of the Melbourne-Wübbena linear combination [38], [39] from the previous to the current timestep is compared to a threshold. If the threshold is exceeded, it is considered as a cycle slip. A description of this cycle slip detection method can be found in [40].

The elevation angle of a ROTI data record is defined as the average elevation angle of the satellite during the time period used for the ROTI calculation. Only measurements with elevation above 30^{o} have been kept for the model derivation, as the data from satellites at low elevation is often contaminated with e.g. multipath effects or other disturbances from vegetation and structures on the ground. The remaining ROTI data have been corrected for other elevation-dependent effects based on the theoretical behavior proposed by [41], assuming an isotropic ionosphere.

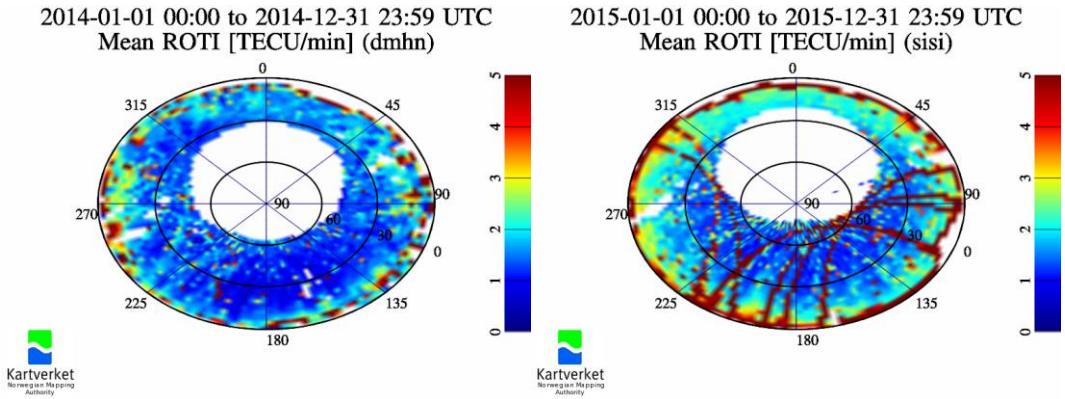


Figure 2: Example of skyplot of mean ROTI, used to identify multipath or other persistent disruptions. On the left is "dmhn" station in 2014. On the right is "sisi" station in 2015, which has signs of bad data.

The data quality has been carefully checked by analyzing the skyplots of each station per year. Examples of skyplots for the GNET stations "dmhn" and "sisi" for the years 2014 and 2015, respectively, are shown in Figure 2. Using such an analysis, persistent unwanted disturbances that are not related to space weather (e.g. multipath) can be removed from the data. In the skyplot graphs, the mean value of ROTI is plotted as a function of azimuth and elevation angles. The skyplots illustrate the need to exclude low elevation data. This filtering allows removing the multipath visible at low elevations for station "dmhn". However, other errors can be observed for station "sisi", tracing large arcs related to some satellite paths. This type of disturbance, which is likely caused by measurement or processing errors in the receiver, can be difficult to remove when it is too dominant. For those cases we have removed the entire year of data. This analysis and filtering has been applied to the whole database to generate the final database that will be used in the continuation of the project.

Another check of ROTI data has been performed analyzing the CDF (Cumulative Distribution Function) of the ROTI index for each station, per year. We expect higher levels of ROTI for years close to the maximum of the solar cycle (2013-2015) and lower values close to the minimum (2008-2009 or 2019-2020). Also, the CDF should be smooth (i.e. without any significant jumps or breaks in the graph). These points have been taken into account and years with too high level of noise or unphysical breaks in the graph have not been included in the final database.

The final ROTI database contained more than 95 million ROTI data records.

Empirical forecasts - HAPEE extension and validation

HAPEE extension

The High lAtitude disturbances Positioning Error Estimator (HAPEE) model is described in [31]. It is an empirical, statistical model that can forecast the probability distribution of ROTI for the near future, using solar wind measurements as its real-time input. The original HAPEE model described in the paper is based on data from GNSS receivers in Norway and provides a forecast for the coming hour. It is parameterized by the magnetic latitude, magnetic longitude, and the value of the Newell coupling function $\frac{d\Phi_{MP}}{dt}$ ([37]). An optional extension of the algorithm transforms the output from ROTI distributions to distributions of PPP positioning error, using an empirical model.

In the frame of the SWAN project, the HAPEE model was modified/extended in two ways. As a 3-hour forecast was requested, all statistical regression had to be re-run with a different configuration. Also, the

database used to build the model was extended by the addition of data from the GNET network in Greenland. A new realization of the HAPEE model was regressed as described in [31], using the extended dataset and providing parameters for 3-hour forecasts.

Validation of ROTI forecasts

For the validation of the HAPEE model a statistical approach has been used. The ROTI datasets consist of the data measured by NMA and GNET networks measured during period [2007-2019].

During the regressions of the HAPEE model, 80% of the whole database of ROTI data has been considered for the regression, while 20% of the database was kept for the validation.

3 different thresholds of ROTI index (T_{ROTI}) have been defined for low (0.75 TECU/min), moderate (1.5 TECU/min) and high (3.0 TECU/min) levels of ionosphere disturbance. The probability of exceeding these thresholds, $P(ROTI > T_{ROTI})$, has been investigated by comparing forecasts from the HAPEE model with the actual observations in the validation dataset.

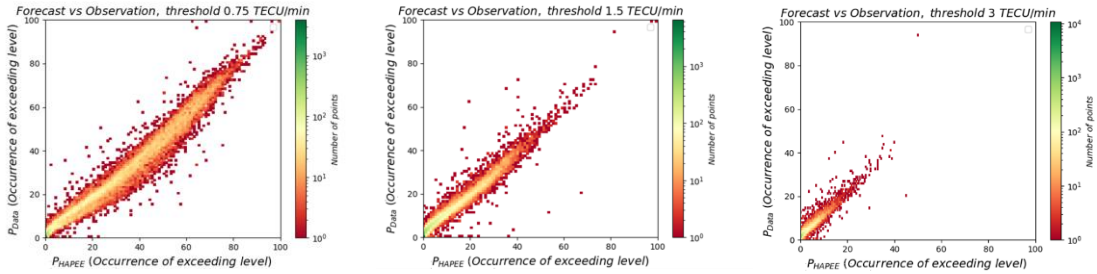


Figure 3: Comparison of forecasted (abscissa axis, HAPEE model) and measured (ordinate axis, validation dataset) probability to exceed the low (left), moderate (middle) and high (right) thresholds.

The problem is binned into different intervals of solar wind conditions (i.e. coupling function), MLT time and Magnetic latitudes. 41 intervals in MLAT have been defined (from 50 to 90^o of magnetic latitudes, sampled per degree, and an initial bin [0,50] for potential low latitude data), 12 intervals of coupling function levels (from 0 to 19000 Wb/s and a last bin for all upper solar wind conditions), and 48 MLT time intervals from 0 to 24 hours. In each interval combination, $P(ROTI > T_{ROTI})$ has been compared between validation dataset and HAPEE model. This way HAPEE model is tested for all levels of coupling function (i.e. solar wind condition), all possible MLT time and Magnetic latitudes.

The results of correspondence between modeled and measured $P(ROTI > T_{ROTI})$ are presented in Figure 3. In those plots, the percentages computed by HAPEE model and obtained from the data have been binned every percent. The colors represent the number of points in each 1 x 1 % bin. This representation shows the spread between modeled and observed percentages. An ideal model would provide a perfect straight line from the origin to the point 100% in abscissa and ordinate. The model results closely follow the actual observations, with the forecasted occurrence rates generally staying within 10 percentage points of the actual occurrence.

One can observe a decay in the percentage correspondence whereas the ROTI threshold is increasing. This phenomenon is due to the lower percentages of occurence to exceed higher values of threshold. The number of points is also lower, as the high levels of scintillation become rarer.

Machine Learning forecasts

Adapting Machine Learning to the problem

In parallel to the statistical approach, based on classification and regressions, Machine Learning techniques have been investigated. Machine learning is normally defined as the group of algorithms trained using an example database in order to find rules, which in time allow them to perform the same task for new data. Three types of machine learning tasks can be classified based on the feedback signal used: supervised learning, unsupervised learning and reinforcement learning. In the case of supervised learning, a database including examples of inputs and the corresponding desired output is presented to the machine learning model with the objective to learn the mapping between the inputs and the outputs. On the contrary, when unsupervised learning is performed the inputs given to the model are not labelled and its task is to find hidden patterns or structures in the data. Another type of machine learning approach is called reinforcement learning, in which the main idea is to maximize a reward in a particular situation by interacting with the environment to solve a problem in a game-like way.

Recurrent Neural Networks (RNN) are a class of artificial neural networks that has applications in all three types of machine learning tasks. RNN are specifically designed to process data for which the temporal behavior is important. In fact, they create a memory of the system by processing each element of the sequence and conserving information regarding what it has encountered up to that time epoch. [42] used RNN to forecast the Dst index. In this project we implemented a variant of the RNN, called Long Short Term Memory (LSTM) [43], in order to forecast ground-based ROTI from solar wind parameters at L1 through the Newell coupling function [37]. The LSTM is designed to be able to retain past information for many time steps and has previously been successfully applied in the field of space weather (See e.g. [44] , [45],[46]). [47] implemented a LSTM NN to predict the am index. By improving the internal structure of basic RNN, the LSTM avoids the gradient issue that can occur with RNN and has the ability to retain long term memory for significant information. Important properties of LSTM NN are:

- It captures the long-term dependencies.
- It models the time series temporal structure.
- It allows variable length inputs and/or outputs.
- It can address multivariate problems and multi-step prediction.

In addition to the processing of the ROTI, for each epoch at each receiver a super observation is created by taking the median of ROTI from the observed satellites as it has been done previously in [48]. Gaps in the solar wind coupling function time series have been filled by Nearest-neighbor interpolation, as it is important for the machine learning algorithm to have continuous time series. As the ROTI data have a 5-minute time resolution, the solar wind coupling function data is resampled to 5-minutes in order to have a matching timeline for all data. The dataset is then split into three parts:

- 60% of the total dataset is used to train the LSTM NN, this is the "Train" dataset.
- 20% of the total dataset is used to evaluate the loss function of the LSTM NN during the training phase, this is the "Validation" dataset.
- 20% of the total dataset is used to test the LSTM NN, this is the "Test" dataset.

The first step for building the machine learning forecaster is to provide local time series of ROTI in an appropriate magnetic latitude interval that is more likely to contain the auroral oval at different magnetic local times. In order to select this relevant interval, we analyzed the geographical distribution of ROTI

with respect to the magnetic local time. Based on our data, the interval of 70^o to 85^o MLAT is the most frequently disturbed and so it is chosen.

The length of the history of the input features has been set to twelve samples, corresponding to one hour of data with a time step of five minutes. The predicted value is set at twelve steps ahead, as it has been done previously by [48]. In summary, this model predicts the value of the ionospheric scintillation index one-hour in advance based on the values of the solar wind coupling function at Earth's bow shock nose during the past hour and the MLT position of the IPP.

The LSTM NN was implemented using the Python libraries "Tensorflow", "Keras", "Pandas" and "Scikit-learn". To obtain the architecture of the LSTM NN model used in this study, several sets of hyperparameters have been tried. Thus the number of LSTM NN layer used to extract representative feature from historical input data is 512. Those layers are then connected to a 3 fully connected layers with 128, 32 and 1 nodes respectively. The batch size used during the training step is 288. The LSTM NN layers are combined with Dropout sequence to prevent overfitting. A patience of 100 epochs has been added to further reduce the overfitting, while the maximum number of training epochs is 1000. The loss function corresponds to the mean square error.

Validation of Machine Learning forecasts

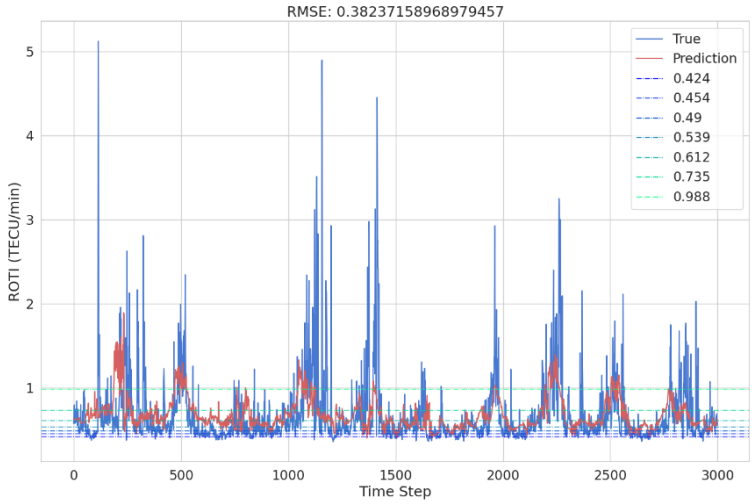


Figure 4: Observed ROTI vs predicted ROTI from the LSTM NN model

Figure 4 shows a comparison of the observed ROTI the with 1-hour predicted ROTI from the LSTM NN model. Note that the figure shows only a part of the validation dataset, and thus must be considered as an example. The LSTM NN forecasting correctly reacts to the disturbed periods, generating forecasted ROTI values larger during these periods. The main difference is that the LSTM NN forecasts time series present lower values than the amplitudes of the measured peaks.

To illustrate this in another way, we compute the standard skill scores using the entire Test dataset. A confusion matrix has been used to evaluate the quality of the prediction system. To classify the ROTI measurements, 8 intervals containing the same number of measurements have been defined. The limits corresponding to these intervals are [0, 0.424, 0.454, 0.49, 0.539, 0.612, 0.735, 0.988, +inf], in units of TECU/minute. Figure 4 shows these ROTI thresholds, represented as dashed horizontal lines. The last of the bins ([0.988, +inf]) is the most important for predicting events. From the confusion matrix, the different skill scores used are the Probability of Detection (POD) which shows what fraction of "true" events is correctly predicted, the False Alarm Rate (FAR) which shows the fraction of the predicted event

that actually did not occur, and the True Skill Score (TSS) which accounts for random chance. Other skill scores used are the Accuracy (ACC) and the Root-Mean-Square Error (RMSE).

Table 1 shows the skill scores when considering the condition ROTI > 0.988 TECU/minute (i.e. the detection of disturbed ionosphere). A perfect model would have POD = 1 and FAR = 0. As can be seen in the table, the current ML model has a substantial number of failed detections and false positives when tested in this way. While the model has some prediction skill, it needs to be improved before it can be considered for operational use. For this reason the forecast service has been based on the empirical statistical approach.

POD	FAR	ACC	TSS	RMSE
0.19	0.61	0.86	0.15	0.38

Table 1: ML model prediction skill scores, considering the condition ROTI > 0.988.

Forecast service

A major objective of the project was to demonstrate usage of the forecast method by establishing a prototype of a forecast service. The service was built on the forecasts supplied by the extended HAPEE algorithm, using archive data to provide forecasts for previous times (hindcasts), but also designed in a way that allows integration of real-time data feeds in the future.

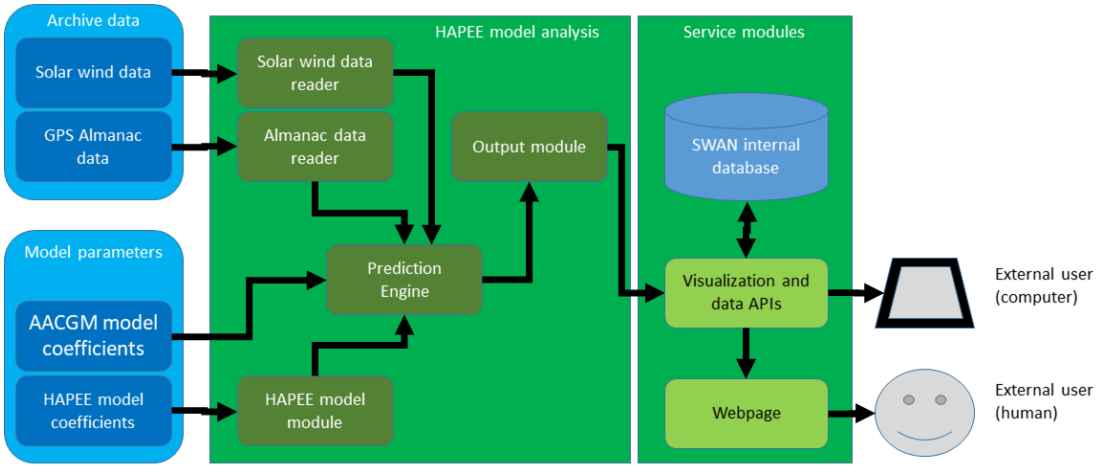


Figure 5: Overview of service components and data flow

Figure 5 shows an overview of the components of the service and the data flows. Solar wind data and GNSS almanacs are the dynamic inputs to the processing, while the model parameters are static sets of coefficients. The AACGM coefficients are a part of the AACGM library [49]. The HAPEE model coefficients define a realization of the HAPEE model. The HAPEE model analysis modules reads the various input data, perform the calculations to provide forecasts, and transmits the output to the database of the forecast service. The model is forecasting the percentage of occurence to exceed a given threshold of ROTI during the following three hours. Note that it also proposes the percentage of occurence to exceed a given PPP error, applying the approach described in [31]. The service modules comprise a database which holds the forecasts, data transformation code to transform the output into a format suitable for visualization, APIs for both internal and external data access, and a webpage. External users can connect to the human interface (webpage), or directly to data APIs (intended for machine-to-machine communication). For a future real-time service, real-time solar wind measurements and almanac data would be processed on a periodic basis. The processing time for a single set of data is less than one

minute. We note that solar wind data is available in near-real-time from the NOAA SWPC data service (currently available at https://services.swpc.noaa.gov/products/solar-wind/).

The Prediction Engine module uses the input modules to acquire input data, and to perform some calculations (notably, calculating geographic IPP coordinates). For the calculation of geomagnetic coordinates, it uses the AACGM library. The Prediction Engine module will, given a date&time and solar wind parameters, calculate maps of forecasted probabilities to exceed thresholds as follows:

- From Solar Wind Data, calculate coupling function value [37]
- For each grid point of the map:
  - Calculate geomagnetic coordinates for IPPs seen from that grid point
  - Retrieve HAPEE model parameters for the current coupling function and set of geomagnetic coordinates
  - For each defined threshold, calculate the probability to exceed it
- Package results and send it to the output module

The dataflow and major processing steps of the Prediction Engine are shown in Figure 6.

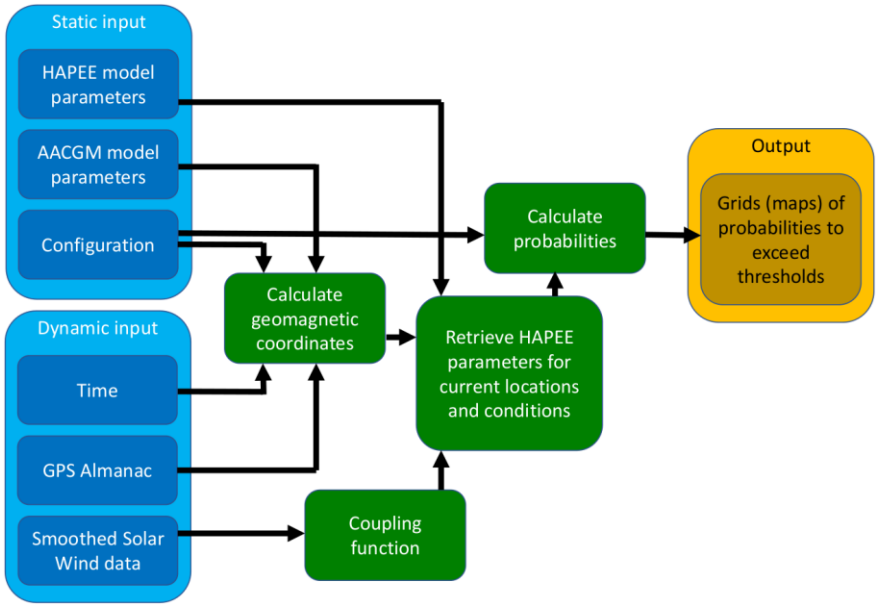


Figure 6: Overview of major processing steps inside the Prediction Engine

The output module transmits data in NetCDF format (See https://www.unidata.ucar.edu/software/netcdf/) via http to an API endpoint that ingests the data into a PostgreSQL database (See https://www.postgresql.org/). Note that this is purely an internal data flow (not exposed externally to users).

In order to best represent the data structure, we can consider it from the point of view of the database. Figure 7 describes three tables, related by foreign key relationships where the data array is stored in the "rast" column in the "coverage" table. Each "coverage" represents one forecast map. In short, one metadata entry, which contains the timestamp, can be related to many coverages, which in turn can be one of many use-cases. This allows us to keep track of the use-cases and uniquely identify the correct

coverage (data array) for a given timestamp, use-case and layer (threshold type). Note that the external users do not directly access the database.

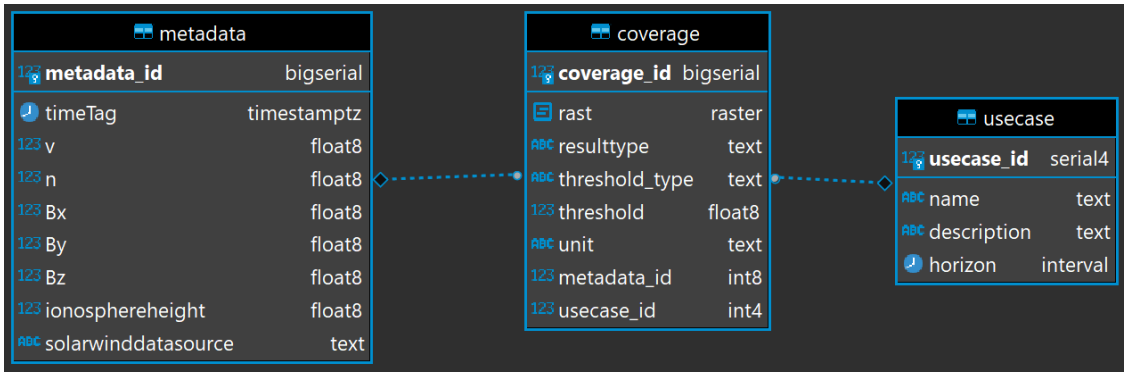


Figure 7: Database Entity Relationship diagram of the normalized data model

The external users can request data via an API. An example of an API call is:

```
https://swan.kartverket.no/api/collections/PPP01/area?parameter_name=pROTI&datetime=2015-01-01T00
```

In the URL for the API call we have specified the use case (for example "PPP01" for Precise Point Positioning user type), the parameter "pROTI" (corresponding to the ROTI probability to exceed 1 TECu/mn) and the time "2015-01-01T00". The API returns the forecast map data to the user in the "CoverageJSON" format (See https://covjson.org). The data access APIs are used to provide content to the service webpage.

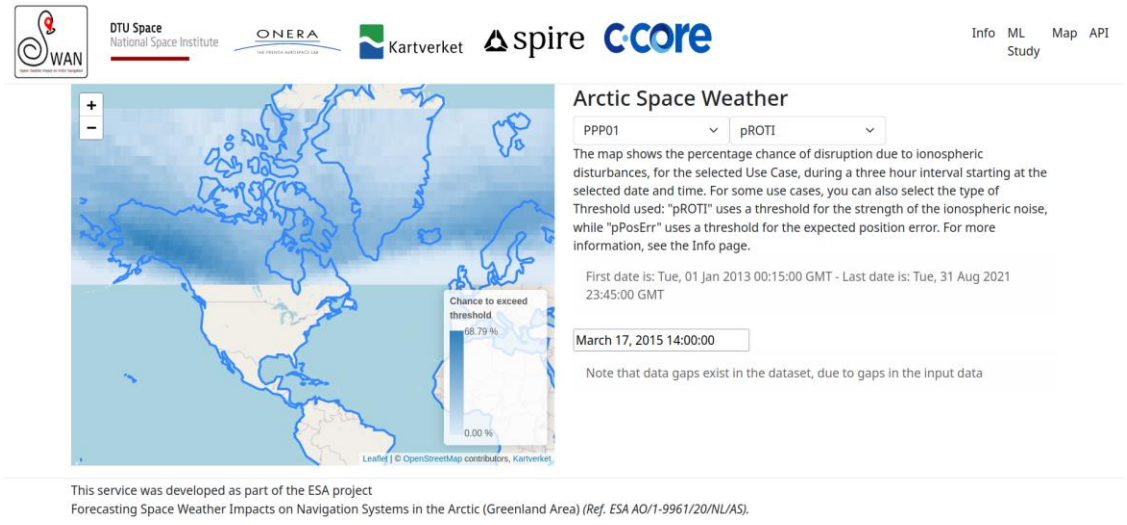


Figure 8: Screenshot of the SWAN service

Figure 8 is a screenshot of the prototype forecast service webpage. The user selects the use case ("PPP01") and threshold type ("pROTI") in the dropdown boxes, and the date and time, and a forecast map is displayed to the left. The map is dynamic, supporting panning and zooming. At the top right are links to information pages about the service itself, the machine learning study, and the APIs for direct data access.

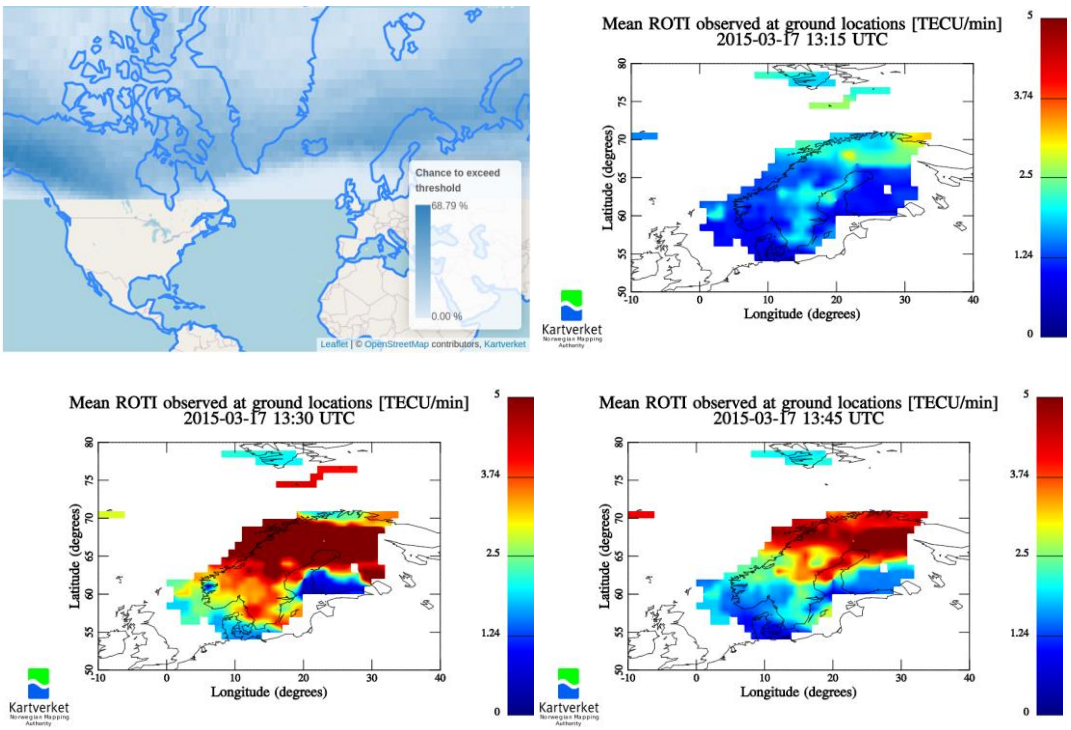


Figure 9: Forecast of chance to exceed ROTI threshold (top left), issued at 2015-03-17 13:00 UTC, and ground ROTI maps for Fennoscandia at three points in time during the forecasted time period.

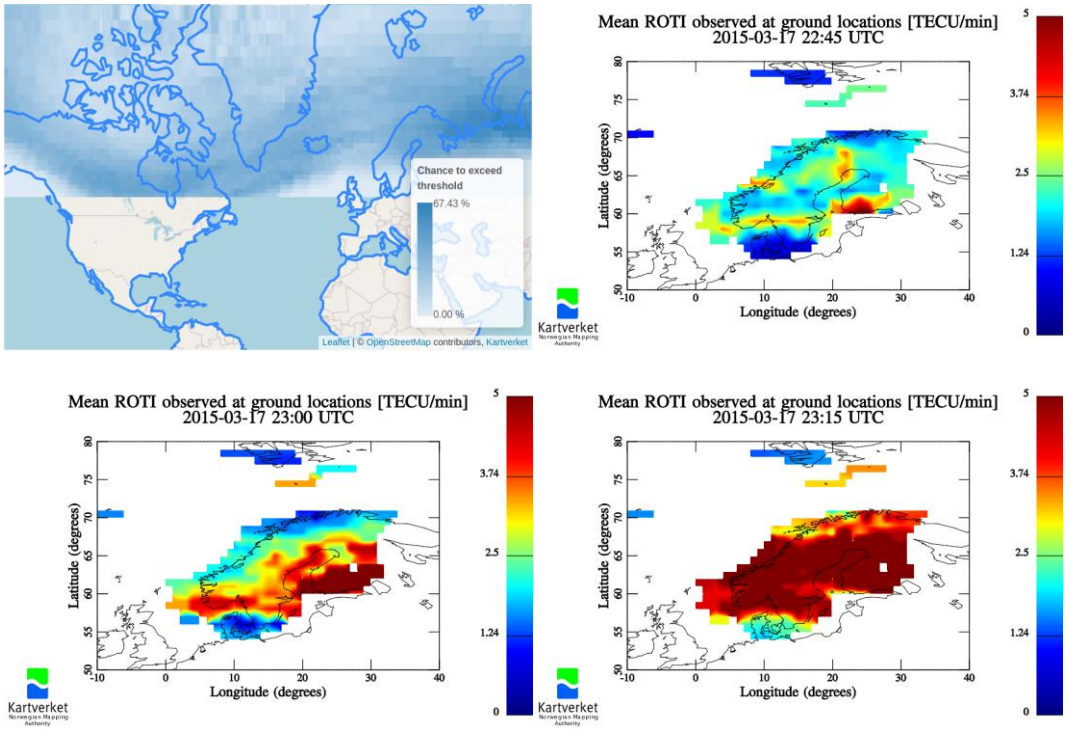


Figure 10: Forecast of chance to exceed ROTI threshold (top left), issued at 2015-03-17 22:30 UTC, and ground ROTI maps for Fennoscandia at three points in time during the forecasted time period.

Figure 9 and Figure 10 show examples of forecast maps and actual observation maps, during a space weather event in 2015 known as the "St. Patrick's day storm" [28], [50]–[52]. The forecasting provided by the service are represented in the top left panel, with probability of occurence to exceed 1 TECu/mn represented in colorscale. The NMA network ROTI measurements are represented by the mean ROTI value at IPP position, and are reported in the other panels for different time corresponding to the interval of the forecasting. The position of the auroral oval appears to be well reproduced by the forecasting.

Summary and Road ahead

The SWAN project has successfully implemented a statistical method (HAPEE) for predicting the occurrence of GNSS disturbances in the polar region. Furthermore, an alternative method using a more sophisticated machine learning (ML) model has been investigated.

The ROTI database was extended by the addition of data from the GNET GNSS receiver network in Greenland (i.e. 13 stations, measuring during the period [2010-2020]. The output of the HAPEE model is probability distributions of ROTI or PPP position error, which can be combined with a threshold to quantify the risk of exceeding that threshold within a period of time. A realization of the HAPEE model for a 3-hour forecast was generated, and a prototype of an operational service based on this approach was created. The service provides forecasts both via a machine-to-machine API, and as a graphical view on a webpage. The provided forecasts were validated, with very good results for the ROTI forecasts.

In parallel to the HAPEE statistical model, a machine learning approach using a LSTM NN was investigated. This method is complementary to the statistical one, providing one-hour forecasted time series of ROTI index, valid for a limited area of MLAT. The results are very promising regarding the timing of the disturbed periods, but are struggling to correctly predict the amplitudes of the disturbances. Further work on this topic is expected to improve the performance.

Future work on the HAPEE-based service

The prototype service is a candidate for inclusion into the ESA Space Weather service network. The model and service are capable of real-time operation, by providing a reliable real-time data flow of solar wind and almanac data and setting up a continuously operating processing server.

Currently, it provides ROTI index forecasts for several use cases but position error forecasts only for the PPP use case. To define position error forecasts for other use cases requires studies to characterize their response to the ionospheric disturbance. More work is needed on validation of position error forecasts.

The spatial coverage of the service may be further extended to fully cover the northern polar cap, and potentially to also cover the southern hemisphere auroral region if sufficient data from that region can be acquired. However, due to the way the model is built it is not suitable for modeling equatorial disturbances.

Future work on machine learning

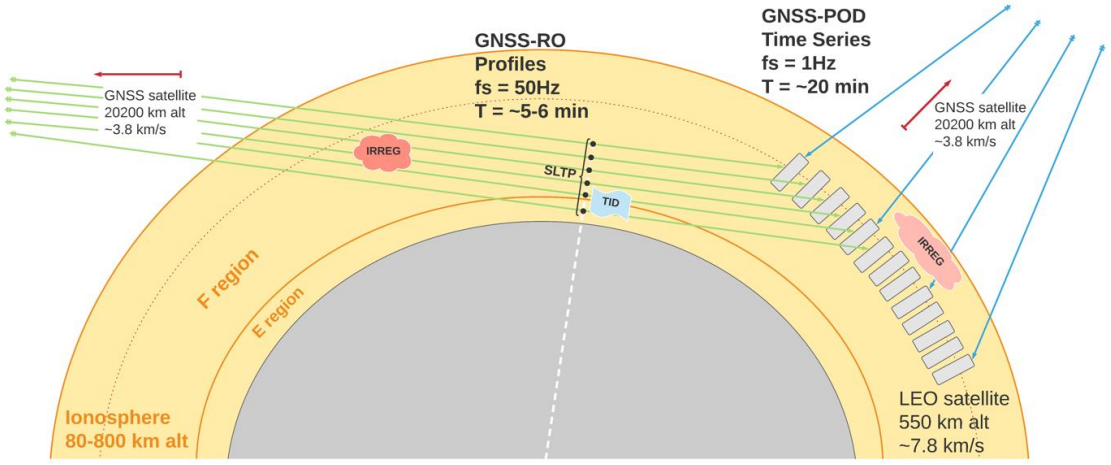
The Machine Learning approach is promising and its performances have been quantified. The model paves the way for a new kind of operational service based on data driven methods. Even with a simple approach linking solar wind data, magnetic local time and the ROTI, the LSTM NN should be able to blindly learn a coarse regular pattern, which is the strength of this approach. One of the main improvements that can be imagined is to consider additional data sources as inputs for the model to extend the environment description.

Additional data sources for future work

The current SWAN system provides a statistical link between the solar wind at L1 and the ROTI (rate of TEC index) or position error experienced by a GNSS receiver on the ground. However, in the future it may be beneficial to include additional ground- and space-based measurements which can improve the forecasts by providing information about intermediate parts of the physical system (i.e. physical processes taking place after the L1 solar wind measurements but before the ground based ROTI measurements). In particular, space-based GNSS measurements can provide information on the state of small scale irregularities in the ionosphere. Furthermore, space-based magnetometers can provide information on the large scale features of the high latitude ionosphere (i.e. the position of the auroral zone) that have a direct impact on the likelihood of scintillation occurring.

Space-based measurements of the high latitude ionosphere can be provided by low cost LEO satellites in sun-synchronous near polar orbits. One way to use these platforms is to deploy GNSS receivers that can be used to sense the ionosphere in a variety of geometries (Figure 11):

- Towards zenith, usually making use of a receiver designed for precise orbit determination (GNSS-POD). The measurements are generally made with 1 Hz sampling and at positive elevation angles (i.e., >30°), effectively sensing only the topside region of the ionosphere above the satellite's orbit [53], [54].
- In a radio occultation (GNSS-RO) geometry (i.e. pseudo horizontally). As the LEO satellite moves in its orbit, the GNSS satellites are seen to rise above or set below the horizon. Thus, the GNSS signals traverse the atmosphere at a range of heights and the direction of the line integral of the electron density (i.e., the sTEC) has mainly an horizontal component. The high rate (50Hz) data can be used to detect and characterise phase and amplitude scintillation [55].



- *Figure 11: GNSS-RO and GNSS-POD observation geometries. Fs is the measurement sampling frequency; T is the measurement duration; SLTP is the straight line tangent point.*

Space based magnetometers in LEO can also be used to help determine the location of the Auroral boundaries through the effects of field aligned currents [56].

Given such diverse measurement regions and types, and the naturally irregular sampling that comes from LEO space-based measurements, it is likely that a combined data assimilation (DA) and machine learning (ML) approach will be required. One possible, multi-step approach might be to

- predict the Auroral boundary using L1 data (ML)

- Update the boundaries using data (DA)
- Update an irregularity model using data (DA)
- Produce a short term forecast of scintillation (ML)

The use of an ionospheric irregularity model (i.e. using height-integrated strength of irregularities, $C_k L$) [57] rather than a ROTI model allows data at different frequencies, with different geometries, and with different sampling rates to be used self-consistently.

Acknowledgements

This work was funded by the ESA project "Forecasting Space Weather Impacts on Navigation Systems in the Arctic (Greenland Area)" (Ref. ESA AO/1-9961/20/NL/AS).

We acknowledge use of NASA/GSFC's Space Physics Data Facility's OMNIWeb service (https://omniweb.gsfc.nasa.gov/), and OMNI data.

The original HAPEE model was developed in a cooperation between NMA, ONERA, CNES (Centre National d'Études Spatiales) and NOSA (Norwegian Space Agency).

For GNSS data from Norway, contact NMA. NMAs website is https://kartverket.no/en.

For GNSS data from Greenland, contact DTU Space. See the GNET website http://go-gnet.org/ for information.

References

[1] C. Jeffrey, *An introduction to gnss: GPS, glonass, galileo and other global navigation satellite systems*. NovAtel, 2010.

[2] J. Aarons, "Global morphology of ionospheric scintillations," *Proceedings of the IEEE*, vol. 70, no. 4, pp. 360–378, 1982.

[3] J. Aarons and S. Basu, "Ionospheric amplitude and phase fluctuations at the gps frequencies," in *Proceedings of the 7th international technical meeting of the satellite division of the institute of navigation (ion gps 1994)*, 1994, pp. 1569–1578.

[4] I. I. Shagimuratov, A. Krankowski, I. Ephishov, Y. Cherniak, P. Wielgosz, and I. Zakharenkova, "High latitude tec fluctuations and irregularity oval during geomagnetic storms," *Earth Planet Sp*, vol. 64, pp. 521–529, 2012.

[5] K. C. Yeh and C.-H. Liu, "Radio wave scintillations in the ionosphere," *Proceedings of the IEEE*, vol. 70, no. 4, p. 324, 1982.

[6] X. Pi, A. J. Mannucci, U. J. Lindqwister, and C. M. Ho, "Monitoring of global ionospheric irregularities using the worldwide gps network," *Geophysical Research Letters*, vol. 24, no. 18, pp. 2283–2286, 1997.

[7] J. Aarons, "Global positioning system phase fluctuations at auroral latitudes," *Journal of Geophysical Research: Space Physics*, vol. 102, no. A8, pp. 17219–17231, 1997.

[8] L. Kersley, C. D. Russell, and D. L. Rice, “Phase scintillation and irregularities in the northern polar ionosphere,” *Radio Science*, vol. 30, no. 3, pp. 619–629, 1995.

[9] A. Krankowski, I. Shagimuratov, L. Baran, I. Ephishov, and N. Tepenitzyna, “The occurrence of polar cap patches in TEC fluctuations detected using GPS measurements in southern hemisphere,” *Advances in Space Research*, vol. 38, no. 11, pp. 2601–2609, 2006.

[10] E. J. Weber, J. A. Klobuchar, J. Buchau, H. C. Carlson, R. C. Livingston, O. de la Beaujardiere, M. McCready, J. G. Moore, and G. J. Bishop, “Polar cap f layer patches: Structure and dynamics,” *Journal of Geophysical Research: Space Physics*, vol. 91, no. A11, pp. 12121–12129, 1986.

[11] S. Skone, M. Feng, R. Tiwari, and A. Coster, “"Characterizing ionospheric irregularities for auroral scintillations"," in *Proceedings of the 22nd international technical meeting of the satellite division of the institute of navigation (ion gnss 2009)*, 2009, pp. 2551–2558.

[12] R. Tiwari, F. Ghafoori, O. Al-Fanek, O. Haddad, and S. Skone, “"Investigation of high latitude ionospheric scintillations observed in the canadian region"," in *Proceedings of the 23rd international technical meeting of the satellite division of the institute of navigation (ion gnss 2010)*, 2010, pp. 349–360.

[13] M. Aquino, F. S. Rodrigues, J. Souter, T. Moore, A. Dodson, and S. Waugh, “Ionospheric scintillation and impact on gnss users in northern europe: Results of a 3 year study,” *Space Communications*, vol. 20, nos. 1-2, pp. 17–29, Jan. 2005.

[14] Ö. Kivanc and R. A. Heelis, “Structures in ionospheric number density and velocity associated with polar cap ionization patches,” *Journal of Geophysical Research: Space Physics*, vol. 102, no. A1, pp. 307–318, 1997.

[15] R. Burston, I. Astin, C. Mitchell, L. Alfonsi, T. Pedersen, and S. Skone, “Turbulent times in the northern polar ionosphere?” *Journal of Geophysical Research: Space Physics*, vol. 115, no. A4, 2010.

[16] L. Spogli, L. Alfonsi, G. De Franceschi, V. Romano, M. H. O. Aquino, and A. Dodson, “Climatology of gps ionospheric scintillations over high and mid-latitude european regions,” *Annales Geophysicae*, vol. 27, no. 9, pp. 3429–3437, 2009.

[17] P. Prikryl, R. Ghoddousi-Fard, B. S. R. Kunduri, E. G. Thomas, A. J. Coster, P. T. Jayachandran, E. Spanswick, and D. W. Danskin, “GPS phase scintillation and proxy index at high latitudes during a moderate geomagnetic storm,” *Annales Geophysicae*, vol. 31, no. 5, pp. 805–816, 2013.

[18] Moen, Jøran, Oksavik, Kjellmar, Alfonsi, Lucilla, Daabakk, Yvonne, Romano, Vineenzo, and Spogli, Luca, “Space weather challenges of the polar cap ionosphere,” *Journal of Space Weather and Space Climate*, vol. 3, p. A02, 2013.

[19] P. Prikryl, P. T. Jayachandran, S. C. Mushini, D. Pokhotelov, J. W. MacDougall, E. Donovan, E. Spanswick, and J.-P. St.-Maurice, “GPS tec, scintillation and cycle slips observed at high latitudes during solar minimum,” *Annales Geophysicae*, vol. 28, no. 6, pp. 1307–1316, 2010.

[20] Jin, Yaqi, Moen, Jøran I., Oksavik, Kjellmar, Spicher, Andres, Clausen, Lasse B.N., and Miloch, Wojciech J., “GPS scintillations associated with cusp dynamics and polar cap patches,” *J. Space Weather Space Clim.*, vol. 7, p. A23, 2017.

[21] Y. Jin, J. I. Moen, and W. J. Miloch, “GPS scintillation effects associated with polar cap patches and substorm auroral activity: Direct comparison,” *J. Space Weather Space Clim.*, vol. 4, p. A23, 2014.

[22] C. van der Meeren, K. Oksavik, D. A. Lorentzen, M. T. Rietveld, and L. B. N. Clausen, “Severe and localized gnss scintillation at the poleward edge of the nightside auroral oval during intense substorm aurora,” *Journal of Geophysical Research: Space Physics*, vol. 120, no. 12, pp. 10, 607–10, 621, 2015.

[23] L. B. N. Clausen, J. I. Moen, K. Hosokawa, and J. M. Holmes, “GPS scintillations in the high latitudes during periods of dayside and nightside reconnection,” *Journal of Geophysical Research: Space Physics*, vol. 121, no. 4, pp. 3293–3309, 2016.

[24] A. Krankowski, I. Shagimuratov, L. Baran, and I. Ephishov, “Study of tec fluctuations in antarctic ionosphere during storm using gps observations,” *Acta Geophysica Polonica*, vols. Vol. 53, nr 2, pp. 205–218, 2005.

[25] C. Watson, P. T. Jayachandran, E. Spanswick, E. F. Donovan, and D. W. Danskin, “GPS tec technique for observation of the evolution of substorm particle precipitation,” *Journal of Geophysical Research: Space Physics*, vol. 116, no. A10, 2011.

[26] P. M. Kintner, B. M. Ledvina, and E. R. de Paula, “GPS and ionospheric scintillations,” *Space Weather*, vol. 5, no. 9, 2007.

[27] I. Cherniak, I. Zakharenkova, and R. J. Redmon, “Dynamics of the high-latitude ionospheric irregularities during the 17 march 2015 st. Patrick’s day storm: Ground-based gps measurements,” *Space Weather*, vol. 13, no. 9, pp. 585–597, 2015.

[28] K. S. Jacobsen and Y. L. Andalsvik, “Overview of the 2015 st. Patrick's day storm and its consequences for rtk and ppp positioning in norway,” *J. Space Weather Space Clim.*, vol. 6, p. A9, 2016.

[29] O. E. Abe, X. Otero Villamide, C. Paparini, R. H. Ngaya, S. M. Radicella, and B. Nava, “Signature of ionospheric irregularities under different geophysical conditions on sbas performance in the western african low-latitude region,” *Ann. Geophys.*, vol. 35, pp. 1–9, 2017.

[30] J. Berdermann, M. Kriegel, D. Banyś, F. Heymann, M. M. Hoque, V. Wilken, C. Borries, A. Heßelbarth, and N. Jakowski, “Ionospheric response to the x9.3 flare on 6 september 2017 and its implication for navigation services over europe,” *Space Weather*, vol. 16, no. 10, pp. 1604–1615, 2018.

[31] V. Fabbro, K. S. Jacobsen, Y. L. Andalsvik, and S. Rougerie, “GNSS positioning error forecasting in the arctic: ROTI and precise point positioning error forecasting from solar wind measurements,” *J. Space Weather Space Clim.*, vol. 11, p. 43, 2021.

[32] M. Najmafshar, “Modeling high-latitude ionospheric scintillations for radio occultation gps receiver performance analysis.” PRISM, 2017.

[33] Z. Yang, Y. T. J. Morton, I. Zakharenkova, I. Cherniak, S. Song, and W. Li, “Global view of ionospheric disturbance impacts on kinematic gps positioning solutions during the 2015 st. Patrick’s day storm,” *Journal of Geophysical Research: Space Physics*, vol. 125, no. 7, p. e2019JA027681, 2020.

[34] K. S. Jacobsen and M. Dähnn, “Statistics of ionospheric disturbances and their correlation with gnss positioning errors at high latitudes,” *J. Space Weather Space Clim.*, vol. 4, p. A27, 2014.

[35] Y. Jiao and Y. T. Morton, “Comparison of the effect of high-latitude and equatorial ionospheric scintillation on gps signals during the maximum of solar cycle 24,” *Radio Science*, vol. 50, no. 9, pp. 886–903, 2015.

[36] J. H. King and N. E. Papitashvili, “Solar wind spatial scales in and comparisons of hourly wind and ace plasma and magnetic field data,” *Journal of Geophysical Research: Space Physics*, vol. 110, no. A2, 2005.

[37] P. T. Newell, T. Sotirelis, K. Liou, C.-I. Meng, and F. J. Rich, “A nearly universal solar wind-magnetosphere coupling function inferred from 10 magnetospheric state variables,” *Journal of Geophysical Research: Space Physics*, vol. 112, no. A1, 2007.

[38] W. Melbourne, “The case for ranging in gps-based geodetic systems,” *Proc. 1st int. symp. on precise positioning with GPS*, pp. 373–386, 1985.

[39] G. Wübbena, “Software developments for geodetic positioning with gps using ti 4100 code and carrier measurements,” *Proceedings 1st International Symposium on Precise Positioning with the Global Positioning System*, pp. 403–412, 1985.

[40] Z. Liu, “A new automated cycle slip detection and repair method for a single dual-frequency gps receiver,” *Journal of Geodesy*, vol. 85, 2011.

[41] C. S. Carrano, K. M. Groves, and C. L. Rino, “On the relationship between the rate of change of total electron content index (roti), irregularity strength (ckl), and the scintillation index (s4),” *Journal of Geophysical Research: Space Physics*, vol. 124, no. 3, pp. 2099–2112, 2019.

[42] H. Lundstedt, H. Gleisner, and P. Wintoft, “Operational forecasts of the geomagnetic dst index,” *Geophysical Research Letters*, vol. 29, no. 24, pp. 34–1–34–4, 2002.

[43] S. Hochreiter and J. Schmidhuber, “Long Short-Term Memory,” *Neural Computation*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[44] Y. Tan, Q. Hu, Z. Wang, and Q. Zhong, “Geomagnetic index kp forecasting with lstm,” *Space Weather*, vol. 16, no. 4, pp. 406–416, 2018.

[45] I. Srivani, G. Siva Vara Prasad, and D. Venkata Ratnam, “A deep learning-based approach to forecast ionospheric delays for gps signals,” *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 8, pp. 1180–1184, 2019.

[46] L. Liu, S. Zou, Y. Yao, and Z. Wang, “Forecasting global ionospheric tec using deep learning approach,” *Space Weather*, vol. 18, no. 11, p. e2020SW002501, 2020.

[47] M. Gruet, “Intelligence artificielle et prévision de l’impact de l’activité solaire sur l’environnement magnétique terrestre,” Theses, UNIVERSITE DE TOULOUSE, 2018.

[48] R. M. McGranaghan, A. J. Mannucci, B. Wilson, C. A. Mattmann, and R. Chadwick, “New capabilities for prediction of high-latitude ionospheric scintillation: A novel approach with machine learning,” *Space Weather*, vol. 16, no. 11, pp. 1817–1846, 2018.

[49] S. G. Shepherd, “Altitude-adjusted corrected geomagnetic coordinates: Definition and functional approximations,” *Journal of Geophysical Research: Space Physics*, vol. 119, no. 9, pp. 7501–7521, 2014.

[50] P. Prikryl, R. Ghoddousi-Fard, J. M. Weygand, A. Viljanen, M. Connors, D. W. Danskin, P. T. Jayachandran, K. S. Jacobsen, Y. L. Andalsvik, E. G. Thomas, J. M. Ruohoniemi, T. Durgonics, K. Oksavik, Y. Zhang, E. Spanswick, M. Aquino, and V. Sreeja, “GPS phase scintillation at high latitudes during the geomagnetic storm of 17–18 march 2015,” *Journal of Geophysical Research: Space Physics*, vol. 121, no. 10, pp. 10, 448–10, 465, 2016.

[51] Y. Jin and K. Oksavik, “GPS scintillations and losses of signal lock at high latitudes during the 2015 st. Patrick’s day storm,” *Journal of Geophysical Research: Space Physics*, vol. 123, no. 9, pp. 7943–7957, 2018.

[52] I. Zakharenkova, I. Cherniak, and A. Krankowski, “Features of storm-induced ionospheric irregularities from ground-based and spaceborne gps observations during the 2015 st. Patrick’s day storm,” *Journal of Geophysical Research: Space Physics*, vol. 124, no. 12, pp. 10728–10748, 2019.

[53] I. Zakharenkova, E. Astafyeva, and I. Cherniak, “GPS and in situ swarm observations of the equatorial plasma density irregularities in the topside ionosphere,” *Earth, Planets and Space*, vol. 68, 2016.

[54] I. Zakharenkova and E. Astafyeva, “Topside ionospheric irregularities as seen from multisatellite observations,” *Journal of Geophysical Research: Space Physics*, vol. 120, no. 1, pp. 807–824, 2015.

[55] M. J. Angling, O. Nogués-Correig, V. Nguyen, S. Vetra-Carvalho, F.-X. Bocquet, K. Nordstrom, S. E. Melville, G. Savastano, S. Mohanty, and D. Masters, “Sensing the ionosphere with the spire radio occultation constellation,” *J. Space Weather Space Clim.*, vol. 11, p. 56, 2021.

[56] A. G. Burrell, G. Chisham, S. E. Milan, L. Kilcommons, Y.-J. Chen, E. G. Thomas, and B. Anderson, “AMPERE polar cap boundaries,” *Annales Geophysicae*, vol. 38, no. 2, pp. 481–490, 2020.

[57] J. Secan and R. Bussey, “An improved model of high-latitude f-region scintillation (wbmod version 13),” p. 93, Aug. 1994.