On estimating the phase scintillation index using TEC provided by ISM and IGS professional GNSS receivers and machine learning

Available online at www.sciencedirect.com

ScienceDirect

Advances in Space Research 73 (2024) 3753–3771

ADVANCES IN SPACE RESEARCH (a COSPAR publication)

www.elsevier.com/locate/asr

On estimating the phase scintillation index using TEC provided by ISM and IGS professional GNSS receivers and machine learning

Rayan Imam [a,*], Lucilla Alfonsi [a], Luca Spogli [a,b], Claudio Cesaroni [a], Fabio Dovis [c]

[a] Upper Atmosphere Physics and Radiopropagation Unit, Istituto Nazionale di Geofisica e Vulcanologia (INGV), Via di Vigna Murata 605, 00142 Rome, Italy

[b] SpacEarth Technology, Via di Vigna Murata 605, 00142 Rome, Italy

[c] Department of Electronics and Telecommunications, Politecnico di Torino, Corso Castelfidardo, 39, 10129 Torino, Italy

Received 23 February 2023; received in revised form 11 July 2023; accepted 18 July 2023
Available online 26 July 2023

Abstract

Amplitude and phase scintillation indexes (S4 and σ_ϕ) provided by Ionospheric Scintillation Monitoring (ISM) receivers are the most used GNSS-based indicators of the signal fluctuations induced by the presence of ionospheric irregularities. These indexes are available only from ISM receivers which are not as abundant as other types of professional GNSS receivers, resulting in limited geographic distribution. This makes the scintillation indexes measurements rare and sparse compared to other types of ionospheric measurements available from GNSS receivers. Total Electron Content (TEC), on the other hand, is an ionospheric parameter available from a wide range of multi-frequency GNSS receivers. Many efforts have worked on establishing scintillation indicators based on TEC, and geodetic receivers in general, introducing various metrics, including the Rate of TEC change (ROT) and ROT Index (ROTI). However, a possible relationship between TEC and its variation, and the corresponding scintillation index that an Ionospheric Scintillation Monitor (ISM) receiver would estimate is not trivial. In principle, TEC can be retrieved from carrier phase measurements of the GNSS receiver, as σ_ϕ. We investigate how to estimate σ_ϕ from time series of TEC and ROT measurements from an ISM in Ny-Ålesund (Svalbard) using Machine Learning (ML). To evaluate its usability to estimate σ_ϕ from geodetic receivers, the model is tested using TEC data provided by a quasi-co-located geodetic receiver belonging to the International GNSS Service (IGS) network. It is shown that the model performance when TEC from the IGS receiver is used gives comparable results to the model performance when TEC from the ISM receiver is utilised. The model's ability to infer the exact value of the scintillation index is bound to Mean Square Error (MSE) = 0.1 radians^2 when $\sigma_\phi \leq 0.8$ radians. For $\sigma_\phi > 0.8$ the MSE reaches 0.18 and 0.45 radians^2 in operative testing using ISM and IGS measurements, respectively. However, the model's ability to detect phase scintillation from IGS TEC measurements is comparable to expert visual inspection. Such a model has potential in alerting against phase fluctuations resulting in enhanced σ_ϕ, especially in locations where ISM receivers are not available, but other types of dual-frequency GNSS receivers are present.

Keywords: Ionospheric irregularities; Geodetic receivers; Artificial intelligence; Space weather

1. Introduction

Ionospheric scintillations are rapid random fluctuations in the amplitude and phase of electromagnetic wave signals as observed by a radio receiver. It happens when the signals pass through irregularities in the electron density distribution of the ionosphere. Scintillation affects radio wave propagation including satellite communications, astronomy observations, radar remote sensing, and Global Navigation Satellite Systems (GNSS) signals.

* Corresponding author.
E-mail address: rayan.imam@ingv.it (R. Imam).

https://doi.org/10.1016/j.asr.2023.07.039
0273-1177/

Ionospheric irregularities are concentrated near the magnetic equator, in the auroral zone, and the polar region. The mechanism behind the low- and high-latitude irregularities are different, and their effect on radio signals is also different (Wernik et al., 2003) leading to a distinction between low and high-latitude scintillation. The former is mainly driven by irregularities in the electron density that form after local sunset when the main source of ionization, the Sun, disappears. Under geomagnetically perturbed conditions, the formation of low-latitude irregularities can increase or suppress (Balan et al., 2018). The latter is associated with geomagnetic storms when batches of the ionosphere are pushed from the day side over the polar caps into the night side creating irregularities in the ionosphere.

Ionospheric Scintillation Monitoring (ISM) receivers are equipped with special firmware designed to process GNSS signals under disturbed ionospheric conditions and extract scintillation monitoring parameters (Bougard et al., 2011). These are the widely adopted phase (σ_ϕ) and amplitude (S4) indices (Fremouw et al., 1978), measuring the entity of the signal fluctuations in a given time window, which is typically of 1 min for ground-based GNSS observations (van Dierendonck et al., 1993). High-latitude scintillations are usually observed as fluctuations in the phase of the signal and rarely in the amplitude, while for low-latitude scintillations both amplitude and phase fluctuations are usually observed (Jiao and Morton, 2015).

The sparse distribution of ISM receivers, relative to other types of professional GNSS receivers, has pushed the scientific community to investigate alternative GNSS-based ionospheric monitoring capabilities to complement (or replace) the ISM receivers. Dual Frequency geodetic receivers' networks have always been a good candidate for these studies because of their spatial and temporal data availability that well exceeds the ISM receivers (see, e.g., Scherliess et al., 2019). This has not been a trivial task because ISM receivers implement resilient tracking loops, low-phase noise oscillators and stable clocks with advanced signal processing techniques to cope with the signal dynamics in case of severe signal fluctuations due to the presence of ionospheric irregularities. In fact, these receivers need to guarantee that the effects from the local oscillator are way less than the effects of the ionosphere. This is not the case for geodetic receivers that are usually noisier than the ISM receivers' oscillators (Nguyen et al., 2019). The latter is the main reason why geodetic receivers are not directly implemented as scintillation monitoring receivers.

Another difference is that the geodetic receivers provide the end users with pseudorange measurements from the Position, Velocity and Time (PVT) unit typically at rates that are 1 Hz or lower. Differently, computation of the scintillation indexes requires the signal intensity and phase measurements, i.e., correlators level measurements, at a high rate of 50 Hz or higher. This prevents computing the scintillation metrics in terms of amplitude and phase scintillation and thus limits the integration of scintillation observations from geodetic receivers directly with ISM receivers. Furthermore, the firmware of commercial ISM receivers has been designed to compute the scintillation indexes and other parameters in quasi-real-time (typically a few minutes latency, depending on how the monitoring station that is usually installed in a remote location is configured to operate), making them a precious tool to monitor ionospheric weather and its impact on GNSS-based services (see, e.g. Kauristie et al., 2021).

Many methods to facilitate using the common geodetic receivers for scintillation monitoring have been proposed. The rate of TEC change (ROT) was introduced in the nineties (Wanninger, 1993) and has been used in many scintillation studies to characterize irregularities having scale sizes from few to tens of kilometres (see, e.g. Alfonsi et al., 2011, de Franceschi et al., 2019). Contrary to scintillation indexes that utilise single frequency measurements, ROT uses dual frequency data to observe fluctuations in the GNSS phase measurements. Then, (Pi et al., 1997) introduced the GPS phase fluctuation index, known as ROT index (ROTI), which is considered today the metric for proxying scintillation from geodetic receivers' measurements. Enhancements of phase scintillation index are associated with electron density gradients, and, indeed, TEC distribution indicates where gradients in the electron density are present. ROTI is then widely used in the recent literature addressing GNSS signal fluctuations of ionospheric origin at both low (see, e.g. Alfonsi et al., 2021; Carrano et al., 2019; Ma and Maruyama, 2006; Yang and Liu, 2016; Yizengaw and Groves, 2018) and high-latitudes (see, e.g., Cherniak and Zakharenkova, 2017; Kotulak et al., 2020; Shagimuratov et al., 2012; Sokolova et al., 2023; Zhao et al., 2022) and for monitoring purposes (Cherniak et al., 2018). However, the detection of scintillation based on ROTI alone is still not reliable and unclear, with debate among the scientific community on the actual physical value estimated by ROTI w.r.t the scintillation indexes (Li et al., 2022). From a geodetic GNSS receiver point of view, low latitude scintillations reduce the signal amplitude and increase the phase noise (see Juan José Miguel et al., 2018 and references within) while the high latitude ionospheric activity acts as an additional dynamic stress on the receiver tracking loops.

In recent years, the introduction of the 1 Hz IGS geodetic receivers (formally known as the high-rate IGS receivers) has attracted efforts to estimate the scintillation indexes themselves from geodetic receivers (Juan et al., 2017; Mrak et al., 2020; Nguyen et al., 2019). For example, (Nguyen et al., 2019) proposed computing the phase scintillation index from the noisy geodetic receivers' phase measurements. The authors demonstrated the possibility of modelling the oscillator noise and eventually removing its noise from phase measurements and estimating the phase scintillation index.

In this work, we alternatively propose using ML techniques to achieve this estimation. We rely on the fact that

TEC measured by geodetic receivers under ionospheric disturbances incorporate high-order terms that are due to the ionospheric perturbances (see Section 2). Since these terms change rapidly under scintillation conditions, as has been already observed and utilized by ROT and ROTI, we rely on the ML model's ability to learn the pattern of TEC and dTEC time series variation under scintillation. Since the effects from the local oscillator are present in all cases, we expect the ML algorithm to learn to differentiate between these time series when only variations due to the local oscillator are present and when scintillation is present above that. Such types of noisy signals are suitable for ML techniques that can learn how scintillation affected TEC time series evolve, rather than learning what are the exact values of TEC and dTEC under disturbed ionospheric conditions.

The objective of this work is to use TEC measurements from geodetic receivers to detect phase scintillation with the help of ML techniques, for possible application in scintillation alerting systems when/where ISM receives are not available. The focus of this work is only on high-latitude scintillation where phase scintillation is dominant and the phase-without-amplitude scintillation phenomenon is observed. There are three main reasons for limiting this investigation to high latitudes. First, the different physics beneath low and high latitude irregularities formation and the resulting scintillation effects which lead to TEC, dTEC and σ_ϕ values at low and high latitudes to take different values. Second, and a direct consequence of the previous sentence, to train and/or validate the model to work both for low and high latitudes scintillation, at leat one ISM receiver co-located with an IGS receiver is needed from each region. To the best of our knowledge, there is no IGS receiver quasi-co-located with one of the low-latitude scintillation receivers available on the eSWua scintillation repository (see Section 3). Finally, the lack of low-latitude geodetic data that well represents all the range of σ_ϕ values. Although the ML model might learn the pattern of strong scintillations from the TEC values provided by ISM receivers under such conditions, the availability of IGS measurements to test the model under the same conditions is questionable. IGS receivers easily loose tracking the signal under strong scintillation, and thus their measurements will not be available for the ranges of σ_ϕ for strong scintillation. For these reasons, we limit the model training and testing to high-latitude phase scintillation.

This paper is organised as follows. In Section 2, the theoretical background of this research is covered. The theoretical TEC formula is recalled and compared to the actual measured TEC by GNSS receivers. Scintillation detection and the phase scintillation index are introduced. Also, a brief introduction to ML models' development is given. The methodology is detailed in Section 3 followed by the results in Section 4. Results are discussed and summarized in Section 5. The conclusions are given in Section 6.

2. Background

2.1. What is included in the Geometry-Free (GF) combination

TEC is one of the parameters that are widely used to characterise the ionosphere. TEC is defined as the number of electrons in a cylinder of 1 m^2 cross-section from a receiver to a satellite. It is measured in TEC units (TECU) where $1 TECU = 1 \times 10^{16}$ electrons m^{-2}. Theoretical TEC is given by the integral of the electron density (n_e) along a ray path from the satellite S to the receiver R:

$$TEC_T = \int_S^R n_e ds \tag{1}$$

Dual frequency GNSS receivers estimate TEC by taking advantage of the fact that the ionosphere is a dispersive medium, i.e., the effects induced by the ionosphere are frequency dependent. Recalling the expression of GNSS phase pseudo-range between satellite S and a receiver R at time t ($\phi_L^{S,R}$):

$$\phi_L^{S,R}(f,t) = \rho^{S,R}(t) + c\delta t^{S,R}(t) + T^{S,R}(t) - I_L^{S,R}(f,t) + \lambda N(t_0) + \epsilon_L^{S,R}(f,t) \tag{2}$$

where $\rho^{S,R}$ is the true range between the satellite and the receiver, $\delta t^{S,R}$ is the difference between the satellite and receiver clocks, $T^{S,R}$ is the tropospheric delay, $I_L^{S,R}$ is the ionospheric phase advance, λN is the ambiguity term (resolved by geodetic receivers using many techniques that are out of the scope of this work), and $\epsilon_L^{S,R}$ is the term that includes all the errors including the receiver noise and multipath effects.

The ionospheric phase advance ($I_L^{S,R}(f,t)$) can be estimated by integrating the index of refraction along the signal path from the satellite S to the receiver R (Datta-Barua et al., 2008; Hoque and Jakowski, 2007):

$$I_L(f) = 40.3\frac{TEC_T}{f^2} + \frac{1}{2}2.26 \times \frac{10^{12}}{f^3}\int_S^R Bcos\theta n_e ds + \frac{1}{3} \times \frac{2439.42}{f^4}\int_S^R n_e^2 ds \tag{3}$$

TEC is estimated from the GF combination of GNSS measurements at two frequencies f_1 and f_2 as:

$$TEC_{GF}(t) = \frac{1}{40.3}\frac{f_1^2 f_2^2}{f_2^2 - f_1^2}(\phi(f_1,t) - \phi(f_2,t)) = TEC_T(t) + \Sigma\beta_k n_e^k(f_{GF},t) + \epsilon_{GF}^{s,r}(t) \tag{4}$$

where $\epsilon_{GF}^{s,r}(f_{GF},t)$ includes all the error terms, and $\Sigma\beta_k n_e^k(f_{GF},t)$ includes all the terms of n_e that do not follow $1/f^2$. In general, $\epsilon_{GF}^{s,r}(f_{GF},t)$ does not change rapidly over short time windows and it can be removed by calibration. The calibration of TEC is subject to a whole literature (Cesaroni et al., 2021; Tornatore et al., 2021). The

calibration method proposed by Ciraolo et al. will be referred to as GG Calibration (Ciraolo et al., 2007), was implemented in this study to remove the error terms $\epsilon_{GF}^{s,r}$ and obtain:

$$TEC_{GG}(t) = TEC_T(t) + \Sigma\beta_k n_e^k(f_{GF}, t) \quad (5)$$

However, the high-order terms $\Sigma\beta_k n_e^k(f_{GF}, t)$, although are negligible for quiet ionosphere, do not follow the $1/f^2$ and can vary rapidly over short time windows (see for example (Fritsche et al., 2005; Prikryl et al., 2021)). We leverage the ML model to learn the scintillation index from the time evolution of these terms. The change in TEC over short time windows has been utilised in the literature for characterising the ionosphere, especially in disturbed conditions. The rate of change of TEC (ROT) (Wanninger, 1993), here referred to as dTEC, and ROTI (Pi et al., 1997) are given by:

$$dTEC(t) = \frac{TEC(t) - TEC(t - t_0)}{t_0} \quad (6)$$

$$ROTI(t) = \sqrt{\frac{1}{N}\Sigma_{j=t-N}^{t}\left(dTEC(j) - \bar{dTEC}\right)^2} \quad (7)$$

where t_0 is the interval of dTEC, $\bar{dTEC}$ is the average dTEC over the ROTI time window, and N is the number of dTEC values in the ROTI time window. ROT and ROTI are typically computed over 1- and 5-minute time windows, respectively. A common value for the time between consecutive TEC measurements from IGS data is 30 s. For the high-rate IGS data available today from some IGS receivers, TEC can be computed with a rate of up to 1 Hz and ROTI can be defined also on shorter time windows, e.g., 1 min to be comparable with ISM receivers' data.

In summary, TEC measurements provided by GNSS receivers include residuals that carry information about scintillation. TEC is computed from the GF combination of GNSS measurements and then calibrated to remove the noise factors.

2.2. Scintillation detection with phase scintillation index

The phase scintillation index σ_ϕ is defined as the variance of the detrended phase over a given time interval, which is usually 1 min for ISM receivers. Phase detrending is a topic of debate among the scientific community since the early 2000s, especially for high-latitude scintillations (Forte and Radicella, 2002; Spogli, Ghobadi, et al., 2021), and it is implemented by receivers in different ways (For example, (Crowley et al., 2011)). It is not the objective of this work to investigate detrending, so we will stick to the method most used in the literature because this is how most phase scintillation data are presented. This means that the detrending of 50 Hz phase measurements is accomplished with a sixth-order Butterworth filter and a cut-off frequency of 0.1 Hz. This makes σ_ϕ sensitive to irregularities covering the full range of scale sizes at all latitudes. Additionally, the authors are aware that this choice affects the capability of distinguishing refractive and diffractive contribution to scintillation, which is an open field in scintillation research (Ghobadi et al., 2020; McCaffrey and Jayachandran, 2019; Zheng Yuhao et al., 2022).

Historically, the method for scintillation detection relies on thresholds on the scintillation index. Scintillation is declared present if the scintillation index exceeds a predefined threshold. The thresholds on the indexes are not fixed among the scientific community, with various studies adopting different thresholds on the amplitude and phase scintillation indexes for the different scintillation monitoring receivers and for the different scintillation severity (Alfonsi et al., 2011; Spogli et al., 2009). Acceptable threshold on σ_ϕ can be defined as (Vilà-Valls et al., 2020):

$\sigma_\phi > 0.25$ strong scintillation.
$\sigma_\phi > 0.15$ and <0.25 moderate scintillation.
$\sigma_\phi < 0.15$ weak-to-no scintillation.

Scintillation detection using a threshold on the scintillation index is not optimum (Taylor et al., 2012). To illustrate this, Fig. 1 shows an example of a phase scintillation event on L1 frequency with manual labelling and threshold labelling. The data was taken from PRN25 observed by an ISM receiver in Ny-Ålesund (Svalbard, Norway) on September 8th, 2017. Data are part of the GNSS scintillation data collection (Upper atmosphere physics and radiopropagation Working Group, 2020) available through the electronic Space Weather upper atmosphere (eSWua, eswua.ingv.it) data portal managed by Istituto Nazionale di Geofisica and Vulcanologia (INGV, Italy). The red box shows the samples that are manually labelled as scintillation by an expert inspector. The red dashed line shows a threshold of $\sigma_\phi = 0.15$ radians. The grey box masks the samples that are not considered scintillation by the threshold. It can be seen from the latter that a good portion of the scintillation window is cut by the threshold. In other words, although the threshold detects the existence of scintillation, it doesn't detect the whole scintillation window. This will be very important when the ML model results are evaluated because it means we can not rely on the number of scintillation samples that exceed the threshold as the exact true label, but rather as an indication of scintillation time windows that are identified by the threshold. This threshold labelling can be enhanced, for example, by using a data series check for the proximity of events (Linty et al., 2019).

2.3. Model development with machine learning

ML has gained success in a wide range of applications including remote sensing, telecommunications, and scientific fields. Its ability to learn without prior knowledge about the underlying physical phenomenon attracted its use in science and engineering especially when it is complex

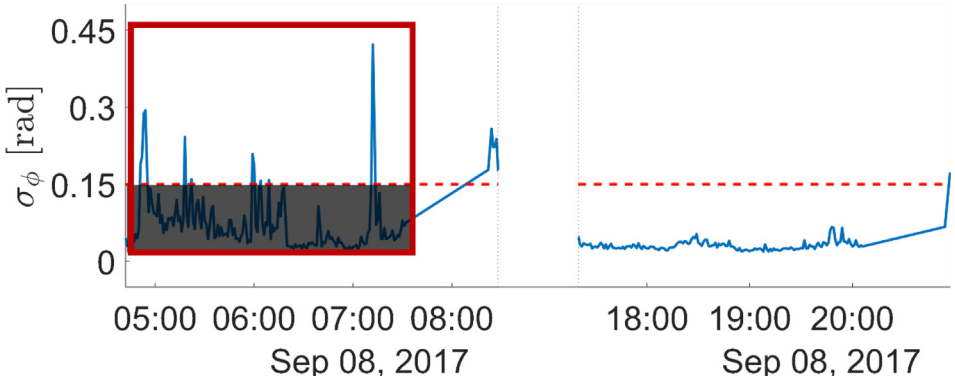

Fig. 1. Doubts about using a threshold on scintillation indexes as labelling methodology. σ_ϕ samples from September 8th, 2017 PRN25 are shown. The red box shows the samples that were labelled as scintillation by expert inspection. The grey box masks the samples that are not labelled as scintillation by a threshold of$\sigma_\phi = 0.15$. A good portion of the scintillation event falls below the threshold.

to model the physics that generated the data. A ML model can discover patterns in the data or predict values of some variables based on measurements and previous data.

To develop a ML model, certain steps must be followed before the model is declared operational. These steps are illustrated in Fig. 2 (adapted from Kulin et al., 2021) and explained in the following paragraphs.

The bulk work of ML modelling lies in data preparation. Feature engineering (also known as feature discovery and feature extraction) is the process of selecting, transforming and manipulating features (also known as attributes) from the raw data using domain knowledge, in order to improve the learning of the models.

In data cleaning, corrupted samples and outliers are removed from the data set. Since the ML model is data-driven, the quality of the model relies on the quality of the data set. For example, in GNSS scintillation data, cycle slips and oscillator anomalies might be considered outliers and thus removed from the data. However, they can also be considered valuable inputs if the objective of the model is to detect outliers in the measurements. A careful definition of the objective of the model and the problem statement is stressed here because of its importance in developing a model able to accomplish the task.

Selecting the ML model is guided at the beginning by the nature of the data set, and later by the model performance among the tested algorithms (trial and error) (Badillo et al., 2020; Kulin et al., 2021; Zhang et al., 2021). Depending on the amount of labelled data provided to the ML algorithm, ML techniques can be divided into supervised, semi-supervised, reinforcement and unsupervised learning. In supervised learning, the ML algorithm is informed about the target outputs it is expected to predict. Thus, the model learns how to fit the data and obtain the desired output. Supervised ML predicts the class corresponding to the input data if the output is categorical data or performs regression if it has to predict a numerical value. In unsupervised learning, on the other hand, the ML model explores the data and understands how they can be grouped. In this work, the ML model is to perform regression and we will only consider bagged regression trees (Breiman, 1996).

Decision tree learning is based on tree structures, defined by recursively partitioning the input space. The decision tree is an acyclic graph in which at each branching node a decision is made by examining a specific feature vector and depending on the decision a branch is followed (Burkov, 2019). In other words, the learning takes place along the branches and nodes employing applied functions for the decision criteria in each node (Linty et al., 2019).

ML methods sometimes are not able to obtain adequate performances when dealing with complex data, like noisy imbalanced high-dimensional data. In such cases, these methods might fail to capture the underlying multiple characteristics and structures of the data (Dong et al., 2020). Ensemble learning methods exploit multiple ML

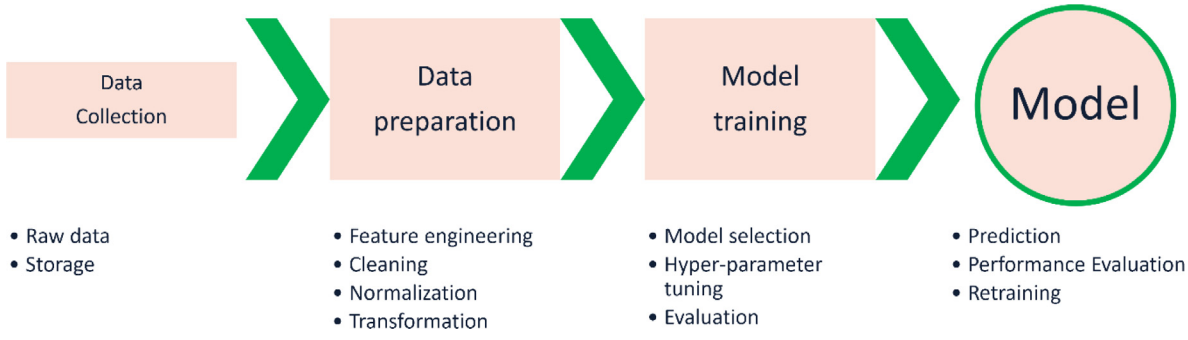

Fig. 2. Machine Learning Modelling Cycle.

algorithms to produce many predictive models based on diverse features, and fuse results of individual models using a selected voting mechanism. By doing this, ensemble learning aims to achieve better performances than that obtained by the individual child algorithms. Increasing the complexity of the ensemble model decreases the model error until reaching a certain complexity, after which the error just increases (Dong et al., 2020). Thus, increasing the complexity of the ensemble is not always the way to achieve better models, but a balance between bias and variance is what we search for. The widely used ensemble classification methods include bagged trees, AdaBoost, gradient boosting, random forest and random subspace.

Ensemble methods are considered the state-of-the-art solution for many ML challenges. They compensate for the errors of a single learner by other learners, and they reach an overall better prediction performance of the ensemble with respect to the single inducers. These improvements are attributed to their overfitting avoidance, extended search space and local minimum avoidance (Sagi and Rokach, 2018).

2.4. ML models evaluation

There are various methods to visualise and evaluate the performance of ML models. The confusion matrix is a popular method to visualise ML models' performance, showing the number of True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) samples. The following metrics easily visualizable from the confusion matrix are used in this paper:

- Accuracy: the ratio of the number of correctly classified samples to the total number of samples.

$$Accuracy = (TP + TN)/(TP + TN + FP + FN)$$

- True positive rate (TPR) or sensitivity: ratio of the detected scintillation samples to the total number of true scintillation samples.

$$TPR = (TP)/(TP + FN)$$

- False Negative Rate (FNR) or Miss detection: the ratio of the number of scintillation samples wrongly classified to the total number of scintillation samples.

$$FNR = (FN)/(TP + FN)$$

- False Positive Rate (FPR) or False alarm: is the ratio of the number of samples wrongly classified as scintillation to the total number of samples classified as scintillation.

$$FPR = FP/(TP + FP)$$

Another performance visualisation tool, that is also quite used for evaluating regression models, is the Receiver Operating Characteristic (ROC) curve (Gneiting and Vogel, 2022; Gönen, 2006). ROC is a graphical tool, originally proposed in signal detection theory, that analyses the performance of a classification model by varying the detection threshold. ROC trades off between TPR and FPR and visualises the comparison in a two-dimensional. The ideal threshold value, the highest TPR and lowest FPR, is located in the upper left corner of the plot. In many cases, this is not achievable and thus a trade-off based on the targeted TPR and the accepted FPR is made. The area under the curve (AUC) is a measure of the discriminatory power of the model, with AUC = 1 indicating an ideal model and AUC = 0.5 indicating the model performance is as good as a random guess.

3. Methodology: Scintillation detection with machine learning

There are two approaches from a supervised ML point of view to implement the objective of this work. The first one is to train a classifier that outputs a discrete value representing the existence of scintillation or not, and the second is to train a regressor which outputs a continuous value that represents the level of scintillation. The former requires manually labelling the training, validation and testing data sets as scintillation or no scintillation, while the second approach requires knowledge of the continuous value output value, and in this case, it is the phase scintillation index value corresponding to the input TEC time series.

Since the objective of this work is to issue alerts when scintillation exceeds certain levels, on first look it seems that we need to train a classifier that outputs scintillation or no scintillation, based on TEC measurements. However, manually labelling the data as scintillation or no-scintillation is labour-intensive (see Section 2.4). We went around this by training a regression model that infers σ_ϕ itself from TEC measurements and thus avoiding the manual labelling task. The outcome of the ML training task is thus a data-driven model that predicts the outcomes (σ_ϕ) from the input data (TEC). In the implementation phase, this ML model is followed by a detection threshold on the estimated σ_ϕ ($\widehat{\sigma}_\phi$). As discussed in Section 2.4, a threshold on σ_ϕ is not the optimum way to detect scintillation, but rather a way to identify the presence of scintillation windows. We use the visual inspection as the final and definitive method for evaluating the developed ML model performance, which is easier than manually labelling the training data sets. The threshold was selected following ROC analysis that trades off between true scintillation detection and false scintillation alarms. The model is evaluated at two stages: the accuracy of estimating σ_ϕ and its ability to detect scintillation. Fig. 3 shows the methodology block diagram.

3.1. Data selection

For training the model, a multi-frequency ISM receiver is needed. For testing the model, a high-rate IGS receiver co-located with an ISM receiver is needed, to obtain the TEC measurements from the former and the corresponding σ_ϕ from the latter. For this, we choose the IGS receiver (NYA2) in Ny-Ålesund and the nearby Septentrio PolaRx5S ISM receivers (NYA0P and NYA1P) belonging to INGV (Upper atmosphere physics and radiopropagation Working Group, 2020). NYA2 was selected because it provides the high-rate observable (1 Hz rate) which allows computing TEC at the 15 s cadence needed for this work. The choice of Nyalesund is backed up by the unlimited access to the ISM data available from INGV's receiver closely located near the IGS receiver. The distance between NYA2 and NYA1P is approximately 226 m, while NYA2 and NYA0P are approximately 1.6 km apart. The IGS receiver is observing the same ionosphere as the ISM receiver, thus the σ_ϕ that these two receivers would estimate are assumed to be equal. Certainly, some degree of spatial decorrelation in the σ_ϕ values is expected. However, the scintillation metric is available with a 1-minute resolution. Such resolution does not allow visualizing the differences resulting from the 1.6 km distance which would be in second scale latency between the scintillation indexes that the IGS would have estimated if capable. Indeed, it would have been beneficial to use co-located ISG/ISM receivers for this purpose, which to the best of our knowledge, do not exist for the eSWua scintillation repository.

Data from 2017 to 2021 were utilised to train and test the model. In particular, the months reported in Table 1, with some known geomagnetic storms (Linty et al., 2018; Spogli, Sabbagh, et al., 2021), were selected to ensure including enough scintillation data in the data set.

The training and testing data sets (middle column of Table 1) were balanced to ensure that the number of samples with $\sigma_\phi > 0.4$ radians are well represented. A total of 2370 samples were obtained with $\sigma_\phi \geq 0.4$ radians and thus a similar number of samples were selected among the samples with $\sigma_\phi < 0.4$ radians for the training and testing phase. This dataset was randomly divided into 70/30 for training and testing.

The data from April 2021 were reserved for testing the model in an operational scenario using ISM data. For testing the model in an operative scenario using IGS TEC, the data of NYA2 for the whole month of September 2017 were selected.

3.2. Data preparation

ISM TEC and dTEC measurements were extracted from the ISM records as provided by the receiver manufacturer. IGS TEC measurements were calculated from the observation and navigation RINEX files available from the CDDIS GNSS data and products archive. *TEC* was calculated from the carrier-phase measurements applying Eq. (4) and calibrated using the GG method to obtain calibrated TEC_{GG} from Eq. (5). Then, the rate of change in TEC (dTEC) was calculated using (6). In this study, we use TEC and dTEC at 15 s cadence because TEC in Septentrio PolaRx5S ISM records are available at this rate. Indeed, higher rate measurements could be obtained from the Septentrio PolaRx5S ISM receivers by processing the high-rate observables of the receiver, which are beyond the scope of this paper.

The samples are prepared in blocks of 3 min that include TEC, dTEC and the satellite elevation (θ_{EL}). This choice of the input features set was inspired by a previous work by the authors (Imam et al., 2021) where this feature set gave the best performance among other sets that were derived

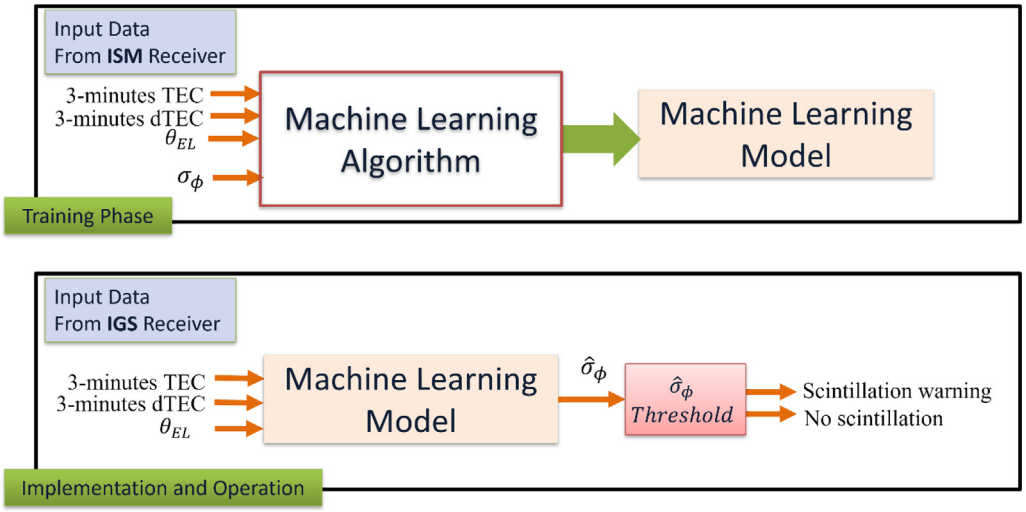

Fig. 3. Model Training, Testing and deployment flow diagram.

Table 1
Selected days for training, testing and validating the model.

Station	Training and Testing Phase	Operational Testing Phase
Nya0p	2017.09, 2018.08, 2019.09, 2020.09	
nya1p	2020.09, 2021.02, 2021.03	2021.04
nya2		2017.09

from the ISM record. For the sake of focusing the work on the regression task, we follow the same methodology from the previous work. The model predicts σ_ϕ at the end of each 3-minutes data block with a cadence of 1 min. The ground truth label is given by σ_ϕ index available from the ISM record.

Samples with cycle slip giving $\sigma_\phi \geq 10$ radians were removed from the dataset. Samples with $1 < \sigma_\phi < 10$ radians were limited and given the value $\sigma_\phi = 1$ radian. The reason behind the latter is that these samples show phase scintillation index inflations that are not corresponding to the scintillation intensity alone. On one side, these cycle slips are often associated with strong scintillations where the receiver fails to track the signal (Nguyen et al., 2019). On the other hand, these σ_ϕ inflations can be due to receiver and satellite anomalies (Liu and Morton, 2020). In this work, we chose to keep them in the training sample and consider them as indicators of strong scintillation.

3.3. Model training and testing

A bagged regression tree with 30 learners was trained to predict $\widehat{\sigma}_\phi$ from TEC measurements provided by the ISM receivers. The model and hyperparameters were selected based on previous work by the authors, where this model gave the best performance among a selected group of supervised ML techniques (Imam et al., 2021). In this work, and in favour of focusing the investigation on the regression task, we do not discuss ML algorithms performance comparison and hyper-parameters optimization. Indeed, these could be the topic of follow-up work.

IGS data can provide TEC values but not the σ_ϕ index, while ISM receivers provide both. We opt for using ISM receivers' data both for the inputs and outputs of the ML model training. This choice was motivated by a need to investigate if it is even possible to train a ML model to infer the value of σ_ϕ in the first place. Furthermore, by using only the data from ISM receivers, the effect of the local oscillator noise between the receiver that calculated σ_ϕ and the receiver that is providing TEC is not present.

The testing of the model, however, was carried out using TEC from both ISM and IGS receivers in operative scenarios, i.e., in the same way the data is expected to be provided to the model when it is deployed. The testing using IGS data is indeed the main result of this work, which explores the feasibility of this proposed methodology in estimating $\widehat{\sigma}_\phi$ using TEC from IGS receivers.

4. Results and discussion

In this section results of training and testing the model are shown. First, the performance of the model during the training and testing phase is shown. Next, the model is tested in an operative scenario using 1 month of ISM TEC data. Such a test is important to assess the model performance when the same type of data used for the training phase is used for testing the model on a large scale. This is also important to confirm that the model is not overfitted to the training data set. Then the model testing using data from the IGS receivers is shown. The latter, if successful, could be the beginning of a new era of reliable scintillation alerts using geodetic receivers in areas where ISM receivers are not present.

4.1. Model training performance: Estimating σ_ϕ from ISM TEC

In this section, we report the results of training and testing the ML model in retrieving the phase scintillation index. Here, TEC was calculated using the same receiver that produced σ_ϕ.

Fig. 4 shows the MSE reported by the model training (a) and testing (b). The histogram on the left panel shows the true σ_ϕ distribution. The middle panel shows the predicted $\widehat{\sigma}_\phi$ distribution. The right panel shows the MSE between σ_ϕ and $\widehat{\sigma}_\phi$, grouped by σ_ϕ range. The MSE reported by the model training is below 0.05 for all σ_ϕ ranges, with an overall MSE = 0.0089 radians^2. The overall MSE obtained when testing the model however is higher (0.0198 radians^2) and reaches approx. 0.1 radians^2 for a couple of σ_ϕ ranges. The balancing of the classes to include an equal number of samples above and below $\sigma_\phi = 0.4$ radians shows evident flaws. Looking at the distribution of the true σ_ϕ, the samples with $\sigma_\phi \approx 0.4$ radians are obviously not well represented. We decide to keep investigating the model in its current state, without further enhancing the training set.

4.2. Testing the model using 1 month of ISM TEC data

In this section, the model is tested in an operative scenario using TEC data from the same scintillation receivers that were used for training the model. This is similar to deploying the model in an operative phase. The ML model is provided with all the data of all the satellites for a certain period and the predicted $\widehat{\sigma}_\phi$ is evaluated.

The data from April 2021 were reserved for this testing, i.e., they were not part of the training or testing sets. In Fig. 5, sample results for PRN02, PRN17 and PRN01 observed on the 2nd, 16th, and 19th of the month, respectively, are shown. The top left plot shows σ_ϕ and $\widehat{\sigma}_\phi$ in red and blue respectively. The ROC curve is shown in the top right plot. The bottom left plot shows the detected scintil-

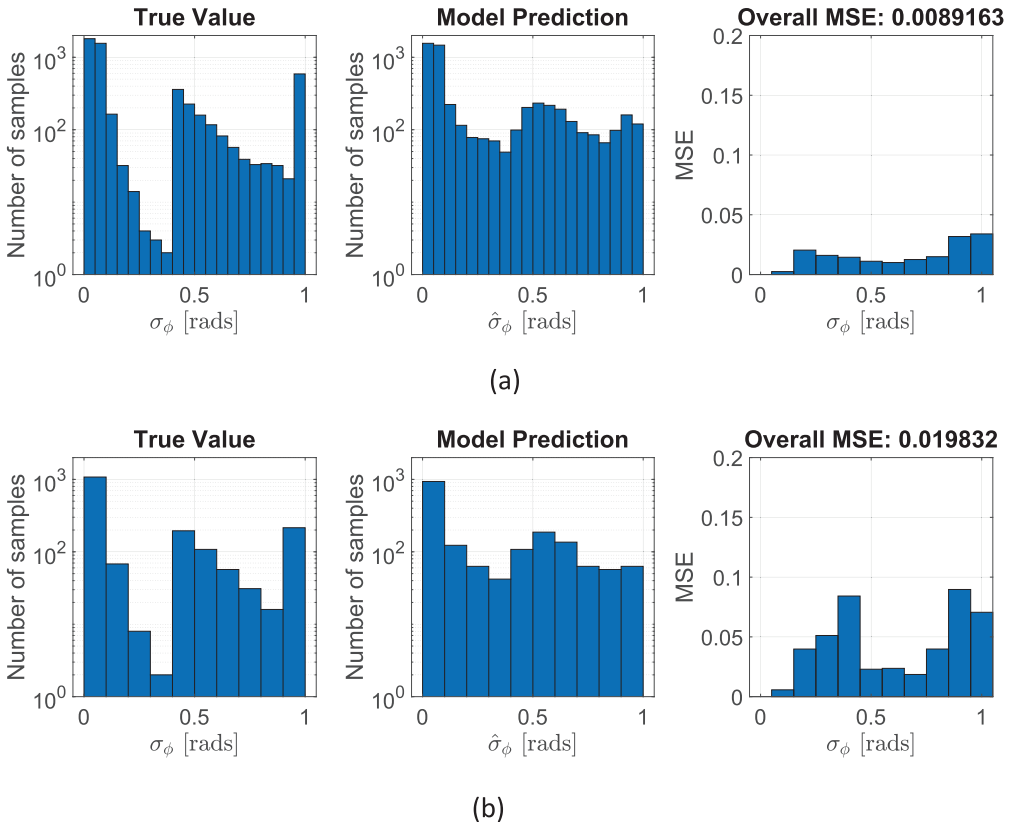

Fig. 4. Model training (a) and testing (b) performances. The left and middle panels show the histograms of the true σ_ϕ and predicted $\hat{\sigma}_\phi$, respectively. The right panel shows the MSE between σ_ϕ and $\hat{\sigma}_\phi$ binned by true σ_ϕ range.

lation instances, and the bottom right plot shows the corresponding confusion matrix. Focusing on Fig. 5(a), it can be seen that $\hat{\sigma}_\phi$ is inflated around the same values that the true σ_ϕ is inflated. However, the error in estimating the exact σ_ϕ value, as anticipated by the MSE analysis, is evident. This indicated that the model's ability to detect scintillation occurrence might be high, although its ability to estimate the exact σ_ϕ is low. To quantify this detection ability, the model is further evaluated using the ROC. Here the true label was set using a threshold of $\sigma_\phi = 0.15$ radians, to test the model's ability to alert about scintillation when σ_ϕ exceed this threshold. Indeed, the disadvantages of using such a detection mechanism are understood as discussed in Section 2. From the ROC analysis, it is found that by setting a threshold of $\hat{\sigma}_\phi \approx 0.2$ radians, the TPR and FPR corresponding to the red circle in the ROC curves are obtained. The detected scintillation instances using this threshold on $\hat{\sigma}_\phi$ are shown in the bottom left plot, and the corresponding confusion matrix is shown in the bottom right plot. The red dots show the true σ_ϕ, the red dashed line is the alert threshold of $\sigma_\phi = 0.15$ radians, and the blue dots show the instance of scintillation as identified by the ML model and the $\hat{\sigma}_\phi$ threshold. Although only a couple of samples in this data series exceeded the alert threshold, the ML model identifies the time window between 12:30 and 13:00 as scintillation. This is indeed what would be manually labelled as scintillation by visual inspection. This also explains the high number of FP due to all the samples that are part of the scintillation window but fall below the alert threshold. This brings back our concerns about using a threshold on σ_ϕ as a strict scintillation detection mechanism. Nevertheless, the ML model learnt to detect scintillation windows with performance closer to manual labelling than the threshold.

Similar results can be observed for PRN17 Fig. 5(b) and PRN01 Fig. 5(c), where the detection threshold on $\hat{\sigma}_\phi$ slightly differs for each satellite, as it is extracted from the relevant ROC curves. However, they are all within $\hat{\sigma}_\phi \approx 0.2$ radians.

To evaluate the overall model performance for the whole month of April 2021, the model performance in terms of MSE and ROC is reported. Fig. 6(a) reports the distribution of σ_ϕ (left), the distribution of $\hat{\sigma}_\phi$(middle) and the MSE between the two values (right). The number of samples is much greater than the ones reported in the training and testing phase because here all the samples of 1 month are included. The samples with $\sigma_\phi < 0.15$ radians are abundant while the samples with $\sigma_\phi > 0.4$ radians are hardly present. This is a typical scenario for scintillation, and it shows why although we used approximately 6 months of data for training the model, the number of samples was at the end in the order of 5,000 samples.

In the same Figure, the MSE for $\sigma_\phi > 0.8$ radians is noticeably high. This is due to the samples with cycle slips

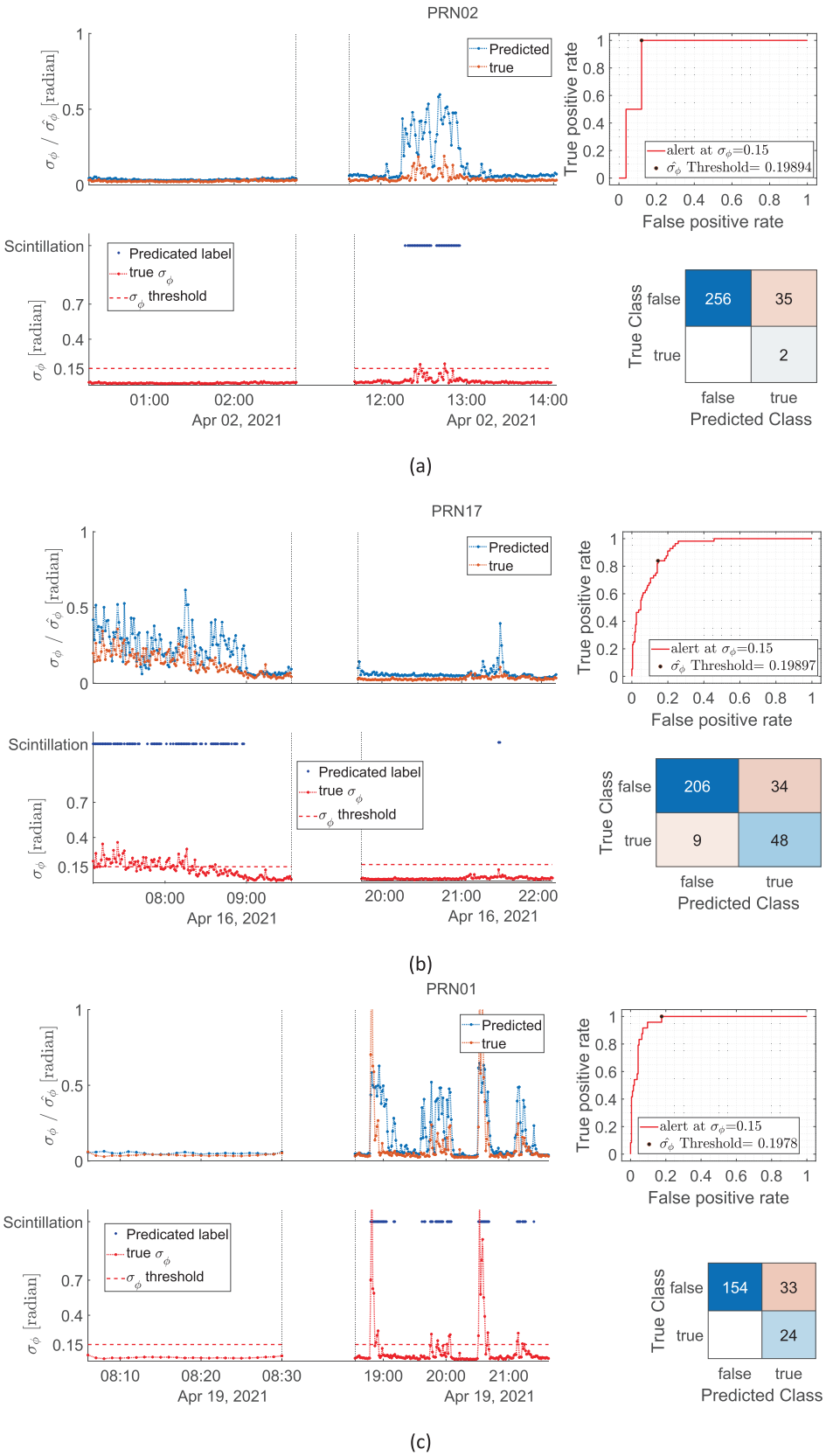

Fig. 5. Sample results of testing the model in an operative scenario using ISM TEC. The figures show model testing results using data of PRN02 (a) PRN17 (b) and PRN01 (c) on the 2nd, 16th and 19th of April 2021 respectively. The upper left plot shows σ_ϕ (red) and $\hat{\sigma}_\phi$(blue). The upper right plot shows the ROC evaluated for alerts when $\sigma_\phi > 0.15$ radian. The red circle shows the threshold on $\hat{\sigma}_\phi$ and the corresponding FPR and TPR. The lower left plot shows the detection using this threshold on $\hat{\sigma}_\phi$ and the corresponding confusion matrix is shown in the lower right plot.

that we chose to include in the training set. Indeed the σ_ϕ associated with these samples do not correspond to the scintillation activity (as mentioned in Section 2) and the model indeed did not learn how to associate $\hat{\sigma}_\phi$ to these values. In fact, the model did not provide any prediction with $\hat{\sigma}_\phi > 0.7$ radians, which explains the high MSE for $\sigma_\phi > 0.7$ radians, and confirms the limitations of using the model in its current condition for estimating the exact value of the scintillation index.

Next, we focus on the model's ability to detect scintillation from TEC. The ROC curve for the data of the whole month is shown in Fig. 6. Here we show the results when the true scintillation threshold was set to$\sigma_\phi = 0.15$ rad (red), σ_ϕ=0.25 rad (green) and$\sigma_\phi = 0.4$ rad (blue) radian, i.e., setting different scintillation alert levels. It can be noticed that the bow of the curve moves toward the upper left corner of the plot (the performance of a perfect classifier) as the threshold on the true σ_ϕ increases. For the alert threshold of $\sigma_\phi = 0.15$ radians in particular, the model can achieve approx. 95% true scintillation detection, accepting 10% false scintillation alarm. These are close to the values shown in Fig. 5 where the high false positives agreed with visual inspection. Hence, if the objective is to alert about scintillation with $\sigma_\phi > 0.15$ radians, then the model is able to achieve this task with TP and FP rates >95% and <10%, respectively, and detection capability close to visual inspection. However, if the objective is to achieve the exact value of σ_ϕ, then the model wrongly estimates the index with up to MSE = 0.18 radians^2 when σ_ϕ is above 0.8 rad, according to the 1-month operational scenario testing. Next, this model will be further tested using TEC from IGS receivers.

4.3. Testing the model using IGS TEC data

The model validated above uses TEC data from ISM receivers. In this section, the model is tested with TEC data

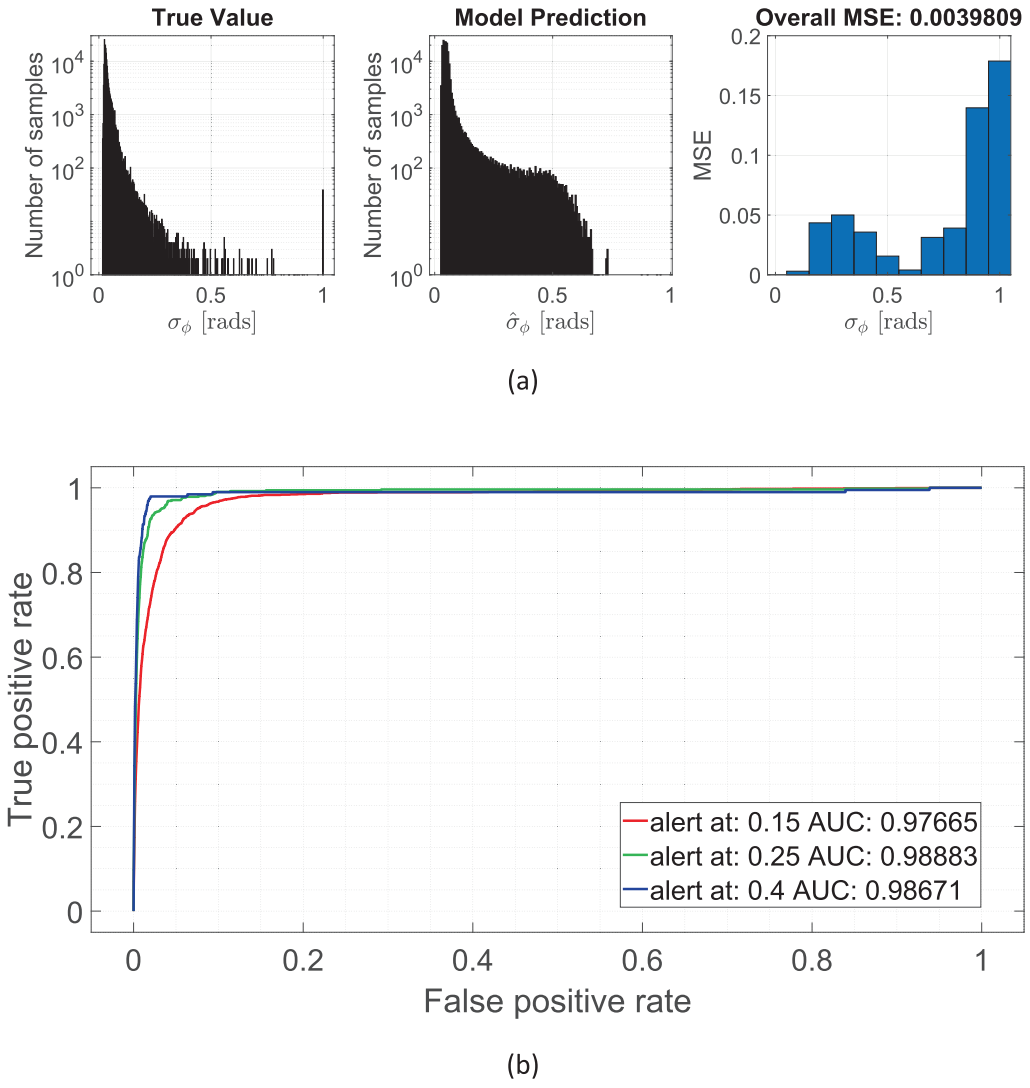

Fig. 6. Overall model performance in an operative scenario using 1 month of ISM TEC data. (a) The left and middle panels show the histograms of the true σ_ϕ and predicted $\hat{\sigma}_\phi$, respectively. The right panel shows the MSE between σ_ϕ and $\hat{\sigma}_\phi$ binned by σ_ϕ range. (b) ROC for three levels of alert thresholds of $\sigma_\phi = \{0.15, 0.25, 0.4\}$ radians.

provided by the IGS receiver. First, we have a look at sample correlation plots between TEC measurements from IGS and ISM receivers. Then we show three selected results of testing the ML model on IGS data. Finally, we show statistics of testing the model with 1 month of data.

Before evaluating the performance of the ML model using IGS TEC data, we investigate the correlation between ISM and IGS TEC (Fig. 7(a)), and the correlation between ISM and IGS dTEC (Fig. 7(b)). Four hours of data from PRN01 on September 3rd, 2017 were used for producing these correlations. The two plots also show the histograms of the 4 values (TEC_{ISM}, TEC_{IGS}, $dTEC_{ISM}$ and $dTEC_{IGS}$) for the sake of a complete picture. The correlation between TEC_{ISM} and TEC_{IGS} is low (0.46) while the correlation between $dTEC_{ISM}$ and $dTEC_{IGS}$ is high (0.95). This was anticipated in Section 2 when discussing TEC calibration. TEC_{ISM} is calibrated using the GG method, while TEC_{IGS} is provided by the scintillation receiver manufacturer. The temporal series of the same data are shown in Fig. 7(c). The top plot shows the true σ_ϕ(blue and left y-axis) and the satellite elevation (red and right y-axis). The middle plot shows TEC provided by the ISM receiver (red) and the one estimated from the IGS data (blue). The bottom plot shows dTEC for the ISM (red) and IGS (blue) receivers. The TEC values from the ISM receiver are noisier than the values from the IGS receiver, however, the dTEC values are almost identical. Thus, if the model is relying on dTEC more than TEC as an indicator of scintillation, it is expected to give a similar performance to the one reported in the previous Section. This can be further investigated by evaluating how each input feature is contributing to the model performance (i.e. feature importance analysis).

The ML model prediction results on sample satellites with PRNs 08 (strong scintillation), 10 (moderate scintillation) and 26 (weak scintillation) measured on the 8th, 10th and 26th of the month, are shown in Fig. 8(a)-(c), respectively. The upper left plot shows in red and blue dots, respectively, the true σ_ϕ and the corresponding $\widehat{\sigma}_\phi$ predicted by the ML model. The x-axis is the time in UT. The upper right plot is the ROC curve. The lower left plot shows the scintillation detection results using the ML model (blue dots) and the corresponding true σ_ϕ. The lower right plot shows the confusion matrix for the detection results. Focusing on Fig. 8(a), σ_ϕ and $\widehat{\sigma}_\phi$ values are inflated around the same time window. However, as observed in Section 4.2, the ML model is not inferring the exact σ_ϕ value measured by the ISM receiver. This further confirms that the model in its current state is not suitable for inferring σ_ϕ. The detection ability of the model using IGS data is, however, interesting. Again, the ROC curve was obtained by setting the threshold on the true σ_ϕ to 0.15 rad. For these data, and with a threshold of $\widehat{\sigma}_\phi \approx 0.2$ radians, it is possible to detect the time window of scintillation from IGS measurements as shown in the bottom left plot. The confusion matrix (bottom right) confirms that many of the samples in the scintillation window fall below the 0.15 rad alert threshold, however, they have been detected as a scintillation event by the ML model.

The same result is observed for PRN10 (middle panel) and PRN26 (bottom panel). The ML model correctly detects most of the instances that the true label using visual inspection will consider as scintillation. This confirms that although the model is using TEC from IGS receivers as inputs, it can provide the same performance that it achieved when TEC from ISM receivers were considered. In particular, the model's ability to accurately detect the scintillation windows is impressive. The latter can be implemented for scintillation alerts from non-ISM receivers' measurements, and thus issue reliable scintillation alerts from non-ISM receivers.

The visual inspection included 200 figures in total, from September 2017. Each figure represents one day of one satellite data, similar to the panels of Fig. 8. The ML label agrees with visual inspection in 193 of these figures (95%). The ML model missed to detect only 2 scintillation events (1%), and 8 figures were wrongly labelled by the ML model as scintillation (4%).

To finalise this results section, we show in Fig. 9 the MSE and the ROC for the whole month of September 2017, similar to the figures shown in Section 4.2. In panel (a), the left and middle plots show the histograms of the true σ_ϕ and predicted $\widehat{\sigma}_\phi$, respectively. The right panel shows the MSE between σ_ϕ and $\widehat{\sigma}_\phi$ binned by σ_ϕ range. Panel (b) shows the ROC curves for three levels of alert thresholds of σ_ϕ, 0.15, 0.25 and 0.4 rad. Surely, TEC values are taken from the IGS receiver while the ground truth σ_ϕ is taken from the ISM receiver. Looking at the histogram of σ_ϕ, there are not many samples with $\sigma_\phi > 0.4$ radians and all the values with $\sigma_\phi = 1.0$ radians are related to cycle slips explained in Section 3. The ML model estimated that a good number of samples have $\widehat{\sigma}_\phi > 0.4$ radian, and none of the samples have $\widehat{\sigma}_\phi = 1$ radians. This explains the high MSE in the right plot, where the samples with true $\sigma_\phi = 1$ radians were all "wrongly" estimated by the ML model. Similar behaviour of the model was observed when it was tested using ISM TEC data for 1 month where the model never assigned $\widehat{\sigma}_\phi > 0.7$ radians. This further confirms that the model testing with IGS data gives similar results to testing the model with ISM data.

The ROC curve in panel (b) is also similar to the ROC in Fig. 6 that was obtained when testing the model using ISM TEC data. The AUC is slightly lower than AUC reported in Fig. 6. However, since the detection threshold on $\widehat{\sigma}_\phi$ in Fig. 8 were similar to those observed in Fig. 5, and the false positives agree with the manual labelling using visual inspection, this detection capability of the model using IGS TEC measurements is considered equivalent to the model performance using ISM TEC data. This detection accuracy is to be compared with other metrics for scintillation detection from non-ISM receivers, for example, ROTI. Also, the agreement between the model's detection and

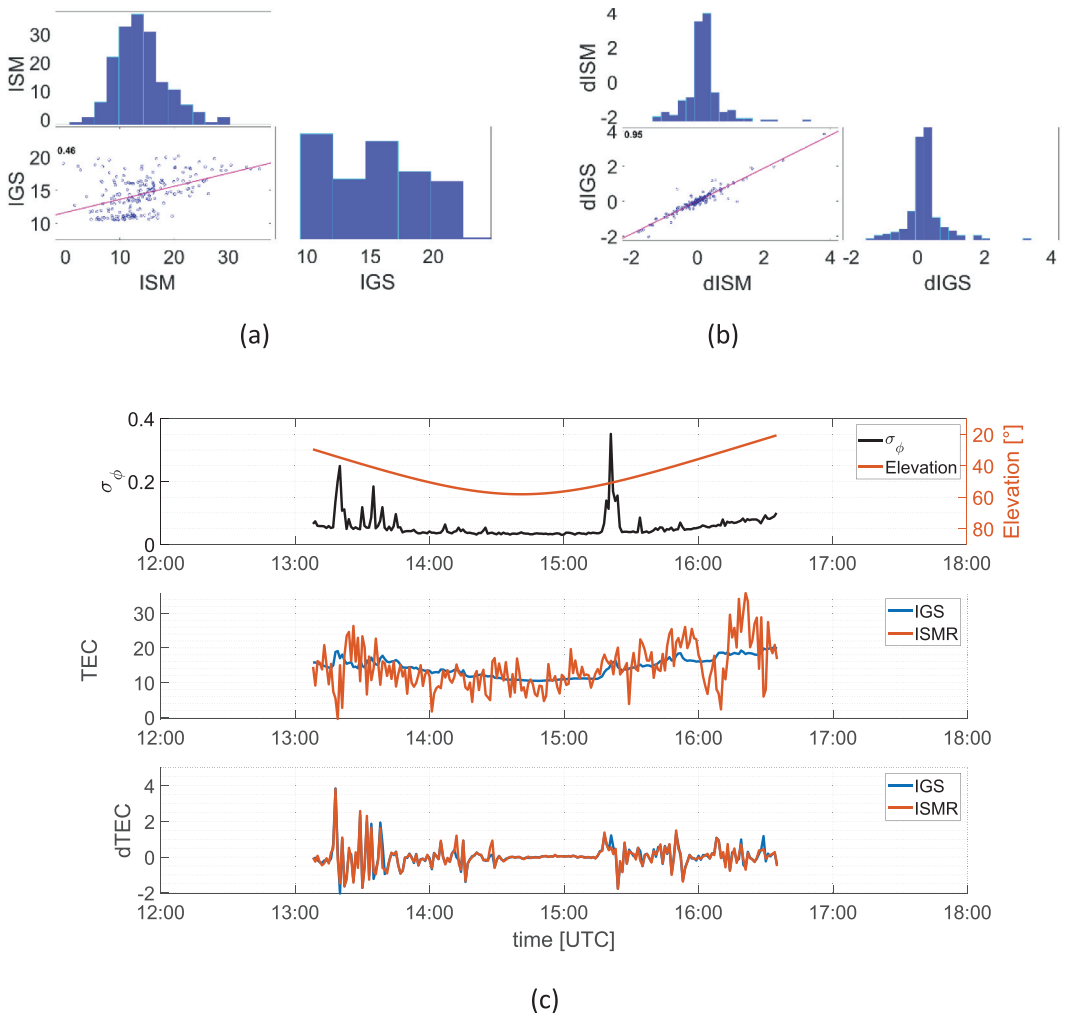

Fig. 7. Correlation between TEC and dTEC estimated by the ISM and IGS receivers. Four hours of data from PRN01 on September, 3rd 2017 were used in these plots (a) the histograms of IGS and ISM TEC and their correlation (b) the histograms of IGS and ISM dTEC and their correlation (c) the time series of σ_ϕ and the satellite elevation angle (top), slant TEC from IGS and ISM data (middle), and dTEC from the IGS and ISM data (bottom).

manual labelling with visual inspection is still to be quantified.

A summary of these results and discussion is given in the next Section.

5. Summary and discussion

The ML model trained in this article is a regressor that predicts the value of σ_ϕ, hence outputs $\widehat{\sigma}_\phi$. A threshold set on $\widehat{\sigma}_\phi$ was implemented as the scintillation detection mechanism. The optimal threshold was selected following ROC analysis. The model performance in terms of the MSE between σ_ϕ and $\widehat{\sigma}_\phi$ was evaluated. Also, the scintillation alert capabilities from $\widehat{\sigma}_\phi$ was investigated. Two ISM receivers in Ny-Ålesund were utilised for training the ML model. A quasi-co-located IGS receiver was utilised for testing the model when TEC from a non-ISM receiver is provided as input to the model.

The model was evaluated in two stages. In the first stage, 1 month of TEC measurements from the ISM receivers were used to infer the value of σ_ϕ. It was found that the model is not able to infer the exact value of σ_ϕ, however $\widehat{\sigma}_\phi$ from the ML model is inflated around the same time windows as the true σ_ϕ. This agreement was evaluated in terms of ROC and detection of scintillation from $\widehat{\sigma}_\phi$. It was found that detecting scintillation from $\widehat{\sigma}_\phi$ estimated by the ML shows a good correlation with scintillation detection using visual inspection. Sample results of this are shown in Fig. 5. The model's ability to reliably detect scintillation from ISM TEC measurements is further confirmed.

In the second stage of testing, TEC measurements from an IGS receiver were provided as inputs to the model. The MSE between σ_ϕ and $\widehat{\sigma}_\phi$ (Fig. 6) is around 0.1, except for $\sigma_\phi \approx 0.9 - 1.0$ radians where the MSE reached 0.4 radians^2, as was the case for the ISM TEC testing. However, detecting scintillation from this ML inferred $\widehat{\sigma}_\phi$ shows great agreement with scintillation detection using visual inspection (Fig. 9), as was the case for the TEC from ISM receivers. In particular, the MSE (panel (a) in Fig. 6 and Fig. 9) and the ROC curves (panel (b) in Fig. 6 and Fig. 9) are very similar in both test scenarios. This indicates

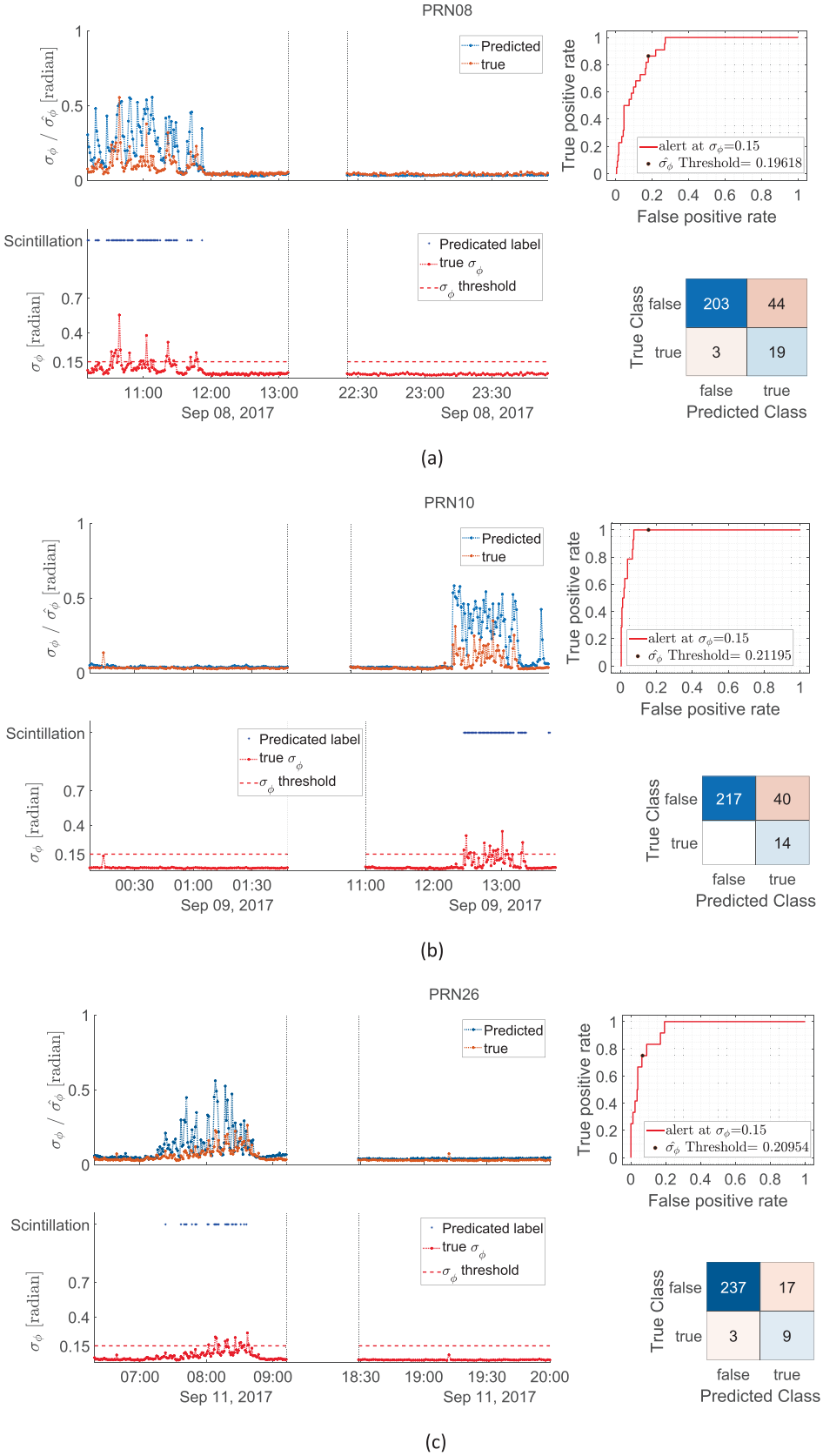

Fig. 8. Sample results of testing the model in an operative scenario using IGS TEC. The Panels show model testing results using data of (a) PRN08 (strong scintillation), (b) PRN10 (moderate scintillation) and (c) PRN26 (weak scintillation) measured on the 8th, 10th and 26th of September 2017 respectively. The upper left plot shows σ_ϕ (red) and $\widehat{\sigma}_\phi$(blue). The upper right plot shows the ROC evaluated for alert when $\sigma_\phi > 0.15$ radian. The lower left plot shows the detection using this threshold on $\widehat{\sigma}_\phi$ and the corresponding confusion matrix is shown in the lower right plot.

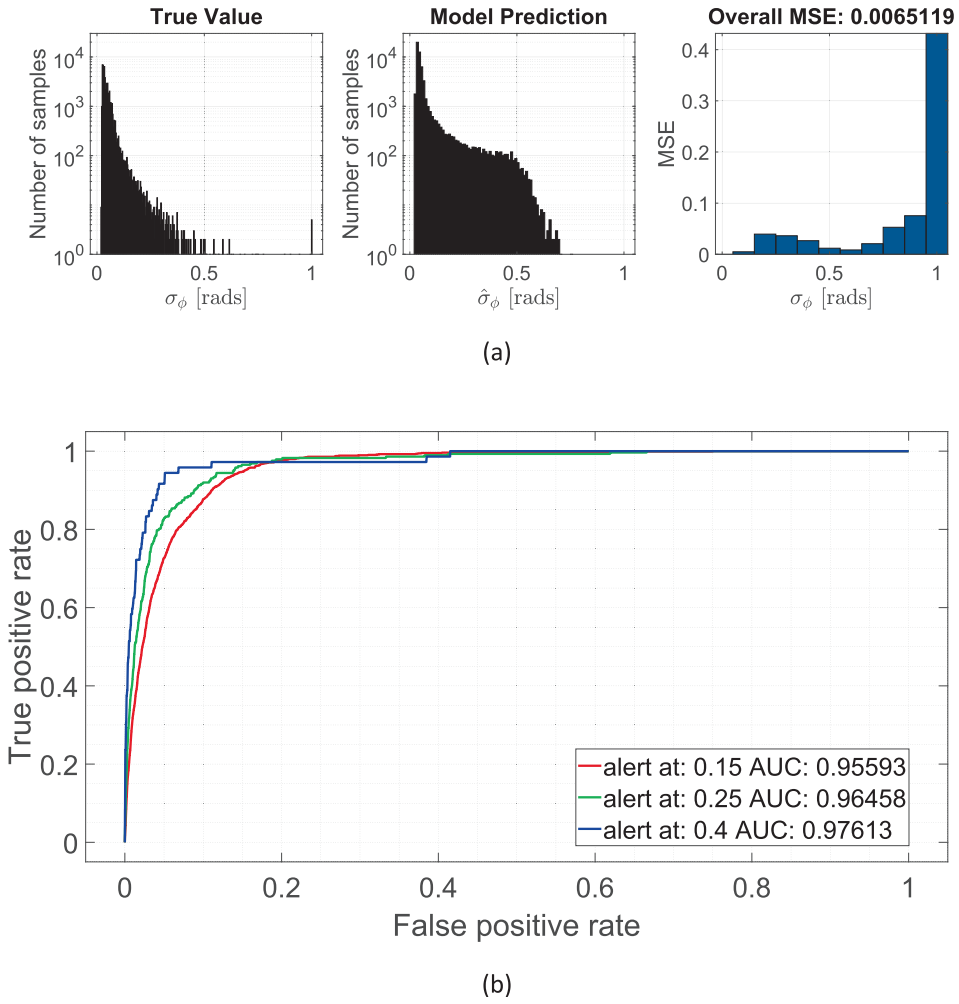

Fig. 9. Overall model performance in an operative scenario using 1 month of IGS TEC data. (a) The left and middle panels show the histograms of the true σ_ϕ and predicted $\widehat{\sigma}_\phi$, respectively. The right panel shows the MSE between σ_ϕ and $\widehat{\sigma}_\phi$ binned by true σ_ϕ range. (b) ROC for three levels of alert thresholds of $\sigma_\phi = \{0.15, 0.25, 0.4\}$ radian.

that indeed TEC from IGS can be used for detecting scintillation with performance comparable to scintillation detection from an ISM TEC using this ML model. This detection capability based on dTEC is in line with the literature on detecting scintillation from geodetic receivers using ROT and ROTI for example.

The TEC values from the ISM receivers are noisier than the calibrated TEC from IGS receivers while the dTEC from the two receivers are almost identical. Moreover, the ML model performance when TEC and dTEC from IGS receivers were used as inputs gave a comparable performance to the ML testing using TEC and dTEC from ISM receivers. An investigation into how much the ML model relies on TEC and dTEC values will give insight into this aspect. In particular, the model seems to rely on dTEC more than TEC and thus it was able to achieve this detection when the type of receiver changes from ISM to IGS. Indeed, such reliance on dTEC is desirable because it is possible to obtain almost identical dTEC values from ISM and IGS receivers, without doing calibration. In principle, the geometry-free combination (i.e. $(\phi(f_1,t) - \phi(f_2,t))$ in Eq. (4)) can be argued to give similar results without the need to calculate TEC and calibrate it.

This model performance is not a surprise. High latitude σ_ϕ is known to be affected by the ionospheric refractive effects, which are the base for calculating TEC. However, the model performance in detecting the diffractive effects has not been investigated. In particular, it would be interesting to evaluate the model performance in detecting the scintillation events identified by (Ghobadi et al., 2020; Spogli, Ghobadi, et al., 2021) using adaptive phase detrending techniques.

On a different note, taking into account the differences between low- and high-latitude scintillations, the model performance in detecting equatorial phase scintillations using TEC from equatorial ISM and IGS receivers could be investigated. A new ML model that predicts $\widehat{\sigma}_\phi$ or $\widehat{S}_4$ in this case, might be needed because of the differences between high- and low-latitude scintillations discussed in Section 2.

During data preparation, strong decisions were made regarding cleaning and balancing the training set which can be summarized as follows:

- In the training phase, the data set was balanced by including a comparable number of samples with $\sigma_\phi \geq 0.4$ radians and $\sigma_\phi < 0.4$ radians. This was done by removing a large portion of the noise samples. However, this resulted in a change in the shape of the distribution of σ_ϕ as shown in Fig. 4. In particular, the samples with σ_ϕ slightly less than 0.4 rad were not well represented in the training set. Since this balancing was implemented using a random selector, this misrepresentation of samples with $\sigma_\phi \approx 0.4$ radians was evident and expected. Future works are recommended to balance the training set using techniques that are suitable for imbalanced non-Gaussian data.
- Cycle slips by ISM receivers are usually indications of strong scintillation where the receiver fails to cope with the signal. Therefore, σ_ϕ in these cases show inflated values that are not proportional to the scintillation intensity. On the other hand, ISM receivers' σ_ϕ>1 rad inflations can be due to non-ionospheric related sources, e.g. from the receivers and satellites oscillator anomalies. In this work, they were considered valuable samples and were limited to $\sigma_\phi\left(1 < \sigma_\phi < 10\right) = 1$ assuming all these inflations are due to strong scintillation. This choice, although guarantees more scintillation samples associated with strong scintillations, needs to be further studied to understand the effect of including and excluding ISM receivers' cycle slips on the final testing when IGS receivers are used.

Finally, in developing this model, several choices were made by the authors during the model training. These are summarized next and recommendations for future implementations are listed:

- First of all, the ML model was trained with TEC values provided by the ISM receivers as inputs because, since we are proposing to use TEC to infer σ_ϕ, it was important to investigate in the first place if the ML will be able to learn the relationship between TEC and σ_ϕ that were estimated by the same receiver. Later, the model was tested in a situation where the receiver estimating TEC and the one estimating σ_ϕ are different. Given the promising results achieved in this paper, future models can attempt to directly train the model using measurements from the geodetic receivers.
- When testing the model using IGS data, TEC was calibrated using the GG method as explained in Section 2. Looking at the relationship between TEC from the two receivers (Fig. 7) and the overall performance of the model when IGS-calibrated TEC values were utilized (Fig. 9), it can be argued the calibration was not important and GF combination could have led to the same results while simplifying the data preparation. This can be verified by testing the model using uncalibrated TEC values or by using other calibration methods. Another approach would be to calibrate TEC from the ISM receivers using the same GG calibration method that was used for IGS TEC, or even to calculate the GF from both ISM and IGS receivers' data and use that for training and testing the model.
- The ML algorithm choice and the feature set selection were inspired by a previous work by the authors (Imam et al., 2021). Neither the optimal ML model nor the input features were investigated in this paper; in order to focus the work on the regression task. However, since the trained regression model in this work shows good results, we recommend future works to consider ML models that are tailored for time-series inputs, including deep learning models, to improve the model's ability to infer σ_ϕ value with high accuracy.
- Lastly, this paper considers TEC at 15 s cadence. However, since IGS high-rate data are provided at a rate of 1 Hz, time series with denser samples can be considered as inputs for the ML model. On the other hand, if the model is to be deployed on a wide scale, and since many of the IGS receivers today provide measurements at 30 s cadence, a ML model that takes as inputs TEC at a lower sampling rate might be needed. This is important also for the possible processing of historic IGS data for scintillation detection using the ML model.

In summary, the results of detecting phase scintillation from non-ISM receivers' TEC measurements and ML techniques are satisfying and show comparable results to scintillation detection from TEC provided by ISM receivers. Also, the feasibility of inferring phase scintillation level is promising. The testing of the model with TEC data from IGS and ISM receivers gave comparable results, meaning that when it comes to TEC, ISM receivers have no advantage over geodetic receivers. It is applicability in alerting against scintillation in areas where ISM receivers are not available is comparable to visual inspection and is to be validated against other scintillation detection mechanisms from non-ISM receivers, like the ROTI index. However, for inferring the value of σ_ϕ from geodetic receivers' TEC measurements, this ML model is not capable of this because of the choice of the ML algorithm and the utilized training data set. To keep investigating the latter, the paper gives recommendations on the next steps.

6. Conclusion

This work investigates detecting scintillation from TEC measurements using ML techniques. The choice of TEC as input to the ML model is interesting because TEC can be obtained from a wide range of GNSS receivers, including the IGS network. Utilising the IGS receivers' data for inferring the value of σ_ϕ combined with ML capabilities is

foreseen to enhance reliable scintillation alerts using geodetic receivers in areas where ISM receivers are not available.

A bagged regression tree was trained using temporal series of TEC and dTEC measurements provided by two ISM receivers in Ny-Ålesund, Svalbard. Data from 2017 to 2021 were used for this training. Data from both ISM and IGS receivers were considered when testing the model.

The model was tested in two stages. First, one month of TEC measurements from the same ISM receivers that were used for training the model was used. The model's ability to infer the exact value of σ_ϕ showed MSE of up to 0.18 radians^2. However, its ability to detect scintillation, which was evaluated using ROC curves, is impressive and its performance is close to expert scintillation detection using visual inspection.

In the second stage of testing, TEC measurements from an IGS receiver were provided to the ML model. The model's ability to infer the value of σ_ϕ from IGS data showed the same MSE as its performance when TEC from ISM receivers were used, except for the cases where cycle slips on the ISM receiver were suspected. Hence, the model's ability to infer the exact σ_ϕ was not achieved. More importantly, the model's ability to detect scintillation from $\widehat{\sigma}_\phi$ inferred using IGS TEC measurements is found to be similar to the previous test scenario using ISM TEC, and thus it shows agreement with scintillation detection by expert visual inspection. This detection ability is in line with the objectives of this work in alerting of scintillation using non-ISM receivers' measurements with potential application in scintillation alerts in regions where ISM receivers are not available.

The next steps in this work can be divided into two tracks. The first one is to compare the model's detection capability from IGS TEC measurements to other metrics, for example, ROTI. The second track is to work on enhancing the ML model training. This includes investigating ML algorithms other than the bagged regression trees, investigating the optimal input features, and enhancing the data set with better balancing and less non-ionospheric σ_ϕ inflations.

In conclusion, the work on inferring the value of σ_ϕ from TEC measurements is promising. However, the scintillation detection from $\widehat{\sigma}_\phi$ is to be improved. Further performance evaluation for the latter is needed to assess its capacity compared to the other metrics for detecting scintillation from geodetic receivers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The scintillation receiver's data are available from eSWua website (Upper atmosphere physics and radiopropagation Working Group, 2020) (http://www.eswua.ingv.it/index.php/data-access-and-info/download-tool). **https://doi.org/10.13127/ESWUA/GNSS.** We thank the CNR for hosting the GNSS receivers in "Dirigibile Italia" station at Ny-Ålesund**.** The IGS data provided by the International GNSS Service is available from the CDDIS GNSS data and products archive https://cddis.nasa.gov/Data_and_Derived_Products/GNSS/high-rate_data.html. Rayan Imam's current research fellowship is funded by the Swarm Space Weather Variability of Ionospheric Plasma (Swarm VIP) project, that has been funded by the European Space Agency, contract 4000130562/20/I-DT with the title "Swarm + 4D Ionosphere.

References

Alfonsi, L., Spogli, L., de Franceschi, G., Romano, V., Aquino, M., Dodson, A., Mitchell, C.N., 2011. Bipolar climatology of GPS ionospheric scintillation at solar minimum. Radio Sci. 46 (3), RS0D05. https://doi.org/10.1029/2010RS004571.

Alfonsi, L., Cesaroni, C., Spogli, L., Regi, M., Paul, A., Ray, S., Lepidi, S., di Mauro, D., Haralambous, H., Oikonomou, C., Shreedevi, P.R., Sinha, A.K., 2021. Ionospheric disturbances over the indian sector during 8 September 2017 geomagnetic storm: plasma structuring and propagation e2020SW002607. Space Weather 19 (3). https://doi.org/10.1029/2020SW002607.

Badillo, S., Banfai, B., Birzele, F., Davydov, I.I., Hutchinson, L., Kam-Thong, T., Siebourg-Polster, J., Steiert, B., Zhang, J.D., 2020. An introduction to machine learning. Clin. Pharmacol. Ther. 107 (4), 871–885. https://doi.org/10.1002/CPT.1796.

Balan, N., Souza, J., Bailey, G.J., 2018. Recent developments in the understanding of equatorial ionization anomaly: A review. J. Atmos. Sol. Terr. Phys. 171, 3–11. https://doi.org/10.1016/J.JASTP.2017.06.020.

Bougard, B., Sleewaegen, J.-M., Spogli, L., Veettil, S. V., Monico, J.F.G., 2011. CIGALA: Challenging the Solar Maximum in Brazil with PolaRxS. In: Proceedings of the 24th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2011), pp. 2572–2579. http://www.ion.org/publications/abstract.cfm?jp=p&articleID=9810.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24 (2), 123–140. https://doi.org/10.1007/BF00058655.

Burkov, A., 2019. The Hundred-page Machine Learning Book. Andriy Burkov, Quebec City, QC, Canada.

Carrano, C.S., Groves, K.M., Rino, C.L., 2019. On the relationship between the rate of change of total electron content index (ROTI), irregularity strength (C k L), and the scintillation index (S4). J. Geophys. Res. Space Phys. 124 (3), 2099–2112. https://doi.org/10.1029/2018JA026353.

Cesaroni, C., Spogli, L., de Franceschi, G., 2021. IONORING: real-time monitoring of the total electron content over Italy. Remote Sens. (Basel) 13 (16), 3290. https://doi.org/10.3390/rs13163290.

Cherniak, I., Krankowski, A., Zakharenkova, I., 2018. ROTI Maps: a new IGS ionospheric product characterizing the ionospheric irregularities occurrence. GPS Solut. 22, 69. https://doi.org/10.1007/s10291-018-0730-1.

Cherniak, I., Zakharenkova, I., 2017. New advantages of the combined GPS and GLONASS observations for high-latitude ionospheric irregularities monitoring: case study of June 2015 geomagnetic storm. Earth Planets Space 69 (1), 66. https://doi.org/10.1186/s40623-017-0652-0.

Ciraolo, L., Azpilicueta, F., Brunini, C., Meza, A., Radicella, S.M., Azpilicueta, F., Brunini, C., Meza, A., Radicella, S.M., 2007. Calibration errors on experimental slant total electron content (TEC) determined with GPS. J Geod 81, 111–120. https://doi.org/10.1007/s00190-006-0093-1.

Crowley, G., Bust, G. S., Reynolds, A., Azeem, I., Wilder, R., O'Hanlon, B. W., Psiaki, M. L., Powell, S., Humphreyes, T. E., Bhatti, J.A., 2011. CASES: A Novel Low-Cost Ground-based Dual-Frequency GPS Software Receiver and Space Weather Monitor. In: Proceedings of the 24th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS 2011), Portland, pp. 1437–1446. http://www.ion.org/publications/abstract.cfm?jp=p&articleID=9702.

Datta-Barua, S., Walter, T., Blanch, J., Enge, P., 2008. Bounding higher order ionosphere errors for the dual frequency GPS user. Radio Sci. 43 (5), RS5010. https://doi.org/10.1029/2007rs003772.

de Franceschi, G., Spogli, L., Alfonsi, L., Romano, V., Cesaroni, C., Hunstad, I., 2019. the ionospheric irregularities climatology over Svalbard from solar cycle 23. Sci. Rep. 9 (1), 9232. https://doi.org/10.1038/s41598-019-44829-5.

Dong, X., Yu, Z., Cao, W., Shi, Y., Ma, Q., 2020. A survey on ensemble learning. Front. Comp. Sci. 14 (2), 241–258. https://doi.org/10.1007/s11704-019-8208-z.

Forte, B., Radicella, S.M., 2002. Problems in data treatment for ionospheric scintillation measurements. Radio Sci. 37 (6), 1096. https://doi.org/10.1029/2001rs002508.

Fremouw, E.J., Leadabrand, R.L., Livingston, R.C., Cousins, M.D., Rino, C.L., Fair, B.C., Long, R.A., 1978. Early results from the DNA Wideband satellite experiment-Complex-signal scintillation. Radio Sci. 13 (1), 167–187. https://doi.org/10.1029/RS013i001p00167.

Fritsche, M., Dietrich, R., Knöfel, C., Rülke, A., Vey, S., Rothacher, M., Steigenberger, P., 2005. Impact of higher-order ionospheric terms on. GPS estimatesGeophys. Res. Lett. 32 (23), L23311. https://doi.org/10.1029/2005GL024342.

Ghobadi, H., Spogli, L., Alfonsi, L., Cesaroni, C., Cicone, A., Linty, N., Romano, V., Cafaro, M., 2020. Disentangling ionospheric refraction and diffraction effects in GNSS raw phase through fast iterative filtering technique. GPS Solut. 24, 85. https://doi.org/10.1007/s10291-020-01001-1.

Gneiting, T., Vogel, P., 2022. Receiver operating characteristic (ROC) curves: equivalences, beta model, and minimum distance estimation. Mach. Learn. 111 (6), 2147–2159. https://doi.org/10.1007/S10994-021-06115-2/FIGURES/1.

Gönen, M., 2006. Receiver operating characteristic (ROC) curves. SAS Users Group Int. (SUGI) 31, 210–231.

Hoque, M.M., Jakowski, N., 2007. Higher order ionospheric effects in precise GNSS positioning. J. Geod. 81 (4), 259–268. https://doi.org/10.1007/s00190-006-0106-0.

Imam, R., Savas, C., Dovis, F., 2021. Detecting phase scintillation at high latitudes using ionospheric scintillation monitoring records and machine learning techniques. In: 2021 IEEE International Conference on Wireless for Space and Extreme Environments (WiSEE), pp. 37–42. https://doi.org/10.1109/wisee50203.2021.9613840.

Jiao, Y., Morton, Y.T., 2015. Comparison of the effect of high-latitude and equatorial ionospheric scintillation on GPS signals during the maximum of solar cycle 24. Radio Sci. 50 (9), 886–903. https://doi.org/10.1002/2015RS005719.

Juan, J.M., Aragon-Angel, A., Sanz, J., González-Casado, G., Rovira-Garcia, A., 2017. A method for scintillation characterization using geodetic receivers operating at 1 Hz. J. Geod. 91 (11), 1383–1397. https://doi.org/10.1007/S00190-017-1031-0/FIGURES/15.

Kauristie, K., Andries, J., Beck, P., Berdermann, J., Berghmans, D., Cesaroni, C., de Donder, E., de Patoul, J., Dierckxsens, M., Doornbos, E., Gibbs, M., Hammond, K., Haralambous, H., Harri, A.-M., Henley, E., Kriegel, M., Laitinen, T., Latocha, M., Maneva, Y., Österberg, K., 2021. Space weather services for civil aviation—challenges and solutions. Remote Sens. (Basel) 13 (18), 3685. https://doi.org/10.3390/rs13183685.

Kotulak, K., Zakharenkova, I., Krankowski, A., Cherniak, I., Wang, N., Fron, A., 2020. Climatology characteristics of ionospheric irregularities described with GNSS ROTI. Remote Sens. (Basel) 12 (16), 2634. https://doi.org/10.3390/rs12162634.

Kulin, M., Kazaz, T., de Poorter, E., Moerman, I., 2021. A survey on machine learning-based performance improvement of wireless networks: PHY, MAC and network layer. Electronics (Switzerland) 10 (3), 1–64. https://doi.org/10.3390/electronics10030318.

Li, W., Song, S., Jin, X., 2022. Ionospheric scintillation monitoring With ROTI from geodetic receiver: limitations and performance evaluation e2021RS007420. Radio Sci. 57 (5). https://doi.org/10.1029/2021RS007420.

Linty, N., Minetto, A., Dovis, F., Spogli, L., 2018. Effects of phase scintillation on the GNSS positioning error during the September 2017 storm at Svalbard. Space Weather 16 (9), 1317–1329. https://doi.org/10.1029/2018SW001940.

Linty, N., Farasin, A., Favenza, A., Dovis, F., 2019. Detection of GNSS ionospheric scintillations based on machine learning decision tree. IEEE Trans. Aerosp. Electron. Syst. 55 (1), 303–317. https://doi.org/10.1109/TAES.2018.2850385.

Liu, Y., Morton, Y.J., 2020. Automatic detection of ionospheric scintillation-like GNSS satellite oscillator anomaly using a machine-learning algorithm. Navigation 67 (3), 651–662. https://doi.org/10.1002/navi.385.

Ma, G., Maruyama, T., 2006. A super bubble detected by dense GPS network at east Asian longitudes. Geophys. Res. Lett. 33 (21), 21103. https://doi.org/10.1029/2006GL027512.

McCaffrey, A.M., Jayachandran, P.T., 2019. Determination of the refractive contribution to GPS Phase "scintillation". J. Geophys. Res. Space Phys. 124 (2), 1454–1469. https://doi.org/10.1029/2018JA025759.

Miguel, J.J., Jaume, S., Guillermo, G.-C., Adrià, R.-G., Adriano, C., Jaume, R., José, B., Estefania, B., David, A., Raul, O., 2018. Feasibility of precise navigation in high and low latitude regions under scintillation conditions. J. Space Weather Space Clim. 8, A05. https://doi.org/10.1051/swsc/2017047.

Mrak, S., Semeter, J., Nishimura, Y., Rodrigues, F.S., Coster, A.J., Groves, K., 2020. Leveraging geodetic GPS receivers for ionospheric scintillation science e2020RS007131. Radio Sci. 55 (11). https://doi.org/10.1029/2020RS007131.

Nguyen, V.K., Rovira-Garcia, A., Juan, J.M., Sanz, J., González-Casado, G., La, T.V., Ta, T.H., 2019. Measuring phase scintillation at different frequencies with conventional GNSS receivers operating at 1 Hz. J. Geod. 93 (10), 1985–2001. https://doi.org/10.1007/s00190-019-01297-z.

Pi, X., Mannucci, A.J., Lindqwister, U.J., Ho, C.M., 1997. Monitoring of global ionospheric irregularities using the worldwide GPS network. Geophys. Res. Lett. 24 (18), 2283–2286. https://doi.org/10.1029/97GL02273.

Prikryl, P., Weygand, J.M., Ghoddousi-Fard, R., Jayachandran, P.T., Themens, D.R., McCaffrey, A.M., Kunduri, B.S.R., Nikitina, L., 2021. Temporal and spatial variations of GPS TEC and phase during auroral substorms and breakups. Polar Sci. 28. https://doi.org/10.1016/J.POLAR.2020.100602 100602.

Sagi, O., Rokach, L., 2018. Ensemble learning: A survey. In: Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, vol. 8, no. 4, pp. e1249. https://doi.org/10.1002/widm.1249.

Scherliess, L., Tsagouri, I., Yizengaw, E., Bruinsma, S., Shim, J.S., Coster, A., Retterer, J.M., 2019. The international community coordinated modeling center space weather modeling capabilities assessment: overview of ionosphere/thermosphere activities. Space Weather 17 (4), 527–538. https://doi.org/10.1029/2018SW002036.

Shagimuratov, I.I., Krankowski, A., Ephishov, I., Cherniak, Y.u., Wielgosz, P., Zakharenkova, I., 2012. High latitude TEC fluctuations and irregularity oval during geomagnetic storms. Earth Planets Space 64 (6), 521–529. https://doi.org/10.5047/eps.2011.10.015.

Sokolova, N., Morrison, A., Jacobsen, K.S., 2023. High latitude ionospheric gradient observation results from a multi-scale network. Sensors 23 (4), 2062. https://doi.org/10.3390/s23042062.

Spogli, L., Alfonsi, L., de Franceschi, G., Romano, V., Aquino, M.H.O., Dodson, A., 2009. Climatology of GPS ionospheric scintillations over high and mid-latitude European regions. Ann. Geophys. 27 (9), 3429–3437. https://doi.org/10.5194/ANGEO-27-3429-2009.

Spogli, L., Sabbagh, D., Regi, M., Cesaroni, C., Perrone, L., Alfonsi, L., di Mauro, D., Lepidi, S., Campuzano, S.A., Marchetti, D., de Santis, A., Malagnini, A., Scotto, C., Cianchini, G., Shen, X.H., Piscini, A., Ippolito, A., 2021a. Ionospheric Response Over Brazil to the August 2018 Geomagnetic Storm as Probed by CSES-01 and Swarm Satellites and by Local Ground-Based Observations e2020JA028368. J. Geophys. Res.: Space Phys. 126 (2). https://doi.org/10.1029/2020JA028368.

Spogli, L., Ghobadi, H., Cicone, A., Alfonsi, L., Cesaroni, C., Linty, N., Romano, V., Cafaro, M., 2021b. Adaptive phase detrending for GNSS scintillation detection: A case study over Antarctica. IEEE Geosci. Remote Sens. Lett. 19, 8009905. https://doi.org/10.1109/LGRS.2021.3067727.

Taylor, S., Morton, Y., Jiao, Y., Triplett, J., Pelgrum, W., 2012. An improved ionosphere scintillation event detection and automatic trigger for GNSS data collection systems. In: Proceedings of the 2012 International Technical Meeting of The Institute of Navigation, Newport Beach, CA, pp. 1563–1569. http://www.ion.org/publications/abstract.cfm?jp=p&articleID=10034.

Tornatore, V., Cesaroni, C., Pezzopane, M., Alizadeh, M.M., Schuh, H., 2021. Performance evaluation of VTEC GIMs for regional applications during different solar activity periods, using RING TEC values. Remote Sens. 13 (8), 1470. https://doi.org/10.3390/RS13081470.

Upper atmosphere physics and radiopropagation Working Group, 2020. Electronic Space Weather upper atmosphere database (eSWua) - GNSS scintillation data (1.0). Istituto Nazionale di Geofisica e Vulcanologia (INGV). https://doi.org/10.13127/eswua/gnss.

van Dierendonck, A.J., Klobuchar, J., Hua, Q., 1993. Ionospheric scintillation monitoring using commercial single frequency C/A code receivers. Proc. ION GPS 93, 1333–1342.

Vilà-Valls, J., Linty, N., Closas, P., Dovis, F., Curran, J.T., 2020. Survey on signal processing for GNSS under ionospheric scintillation: Detection, monitoring, and mitigation. Navigat. J. Instit. Navigat. 67 (3), 511–536. https://doi.org/10.1002/navi.379.

Wanninger, L., 1993. The occurrence of ionospheric disturbances above Japan and their effects on precise GPS positioning. Proc. CRCM 93, 175–179.

Wernik, A.W., Secan, J.A., Fremouw, E.J., 2003. Ionospheric irregularities and scintillation. Adv. Space Res. 31 (4), 971–981. https://doi.org/10.1016/S0273-1177(02)00795-0.

Yang, Z., Liu, Z., 2016. Correlation between ROTI and Ionospheric Scintillation Indices using Hong Kong low-latitude GPS data. GPS Solut. 20 (4), 815–824. https://doi.org/10.1007/S10291-015-0492-Y/FIGURES/9.

Yizengaw, E., Groves, K.M., 2018. Longitudinal and seasonal variability of equatorial ionospheric irregularities and electrodynamics. Space Weather 16 (8), 946–968. https://doi.org/10.1029/2018SW001980.

Zhang, Q., Barri, K., Jiao, P., Salehi, H., Amir, Alavi, H., Alavi, A.H., Jiao, P., 2021. Genetic programming in civil engineering: advent, applications and future trends. Artif. Intell. Rev. 54, 1863–1885. https://doi.org/10.1007/s10462-020-09894-7.

Zhao, D., Li, W., Li, C., Hancock, C.M., Roberts, G.W., Wang, Q., 2022. Analysis on the ionospheric scintillation monitoring performance of ROTI extracted from GNSS observations in high-latitude regions. Adv. Space Res. 69 (1), 142–158. https://doi.org/10.1016/J.ASR.2021.09.026.

Zheng, Y., Xiong, C., Jin, Y., Liu, D., Oksavik, K., Xu, C., Zhu, Y., Gao, S., Wang, F., Wang, H., Yin, F., 2022. The refractive and diffractive contributions to GPS signal scintillation at high latitudes during the geomagnetic storm on 7–8 September 2017. J. Space Weather Space Clim. 12, 40. https://doi.org/10.1051/swsc/2022036.