



CAN UNCLASSIFIED



DRDC | RDDC  
technologysciencetechnologie

# The Assimilation Canadian High Arctic Ionospheric Model (A-CHAIM)

## *Implementation, Workflow, and Performance*

David R. Themens

Benjamin Reid; Anthony M. McCaffrey; P.T. Jayachandran

Prepared by:  
University of New Brunswick  
3 Bailey Drive  
P.O. Box 4400  
Fredericton, New Brunswick  
Canada  
E3B 5A3

PSPC Contract Number: W7714-186507/001/SS  
Technical Authority: Thayanathan (Thaya) Thayaparan, Defence Scientist

Contractor's date of publication: March 2020

The body of this CAN UNCLASSIFIED document does not contain the required security banners according to DND security standards. However, it must be treated as CAN UNCLASSIFIED and protected appropriately based on the terms and conditions specified on the covering page.

**Defence Research and Development Canada**

**Contract Report**

DRDC-RDDC-2020-C045

March 2020

CAN UNCLASSIFIED

**Canada**

## CAN UNCLASSIFIED

### IMPORTANT INFORMATIVE STATEMENTS

This document was reviewed for Controlled Goods by Defence Research and Development Canada using the Schedule to the *Defence Production Act*.

Disclaimer: This document is not published by the Editorial Office of Defence Research and Development Canada, an agency of the Department of National Defence of Canada but is to be catalogued in the Canadian Defence Information System (CANDIS), the national repository for Defence S&T documents. Her Majesty the Queen in Right of Canada (Department of National Defence) makes no representations or warranties, expressed or implied, of any kind whatsoever, and assumes no liability for the accuracy, reliability, completeness, currency or usefulness of any information, product, process or material included in this document. Nothing in this document should be interpreted as an endorsement for the specific use of any tool, technique or process examined in it. Any reliance on, or use of, any information, product, process or material included in this document is at the sole risk of the person so using it or relying on it. Canada does not assume any liability in respect of any damages or losses arising out of or in connection with the use of, or reliance on, any information, product, process or material included in this document.

- © Her Majesty the Queen in Right of Canada (Department of National Defence), 2020
- © Sa Majesté la Reine en droit du Canada (Ministère de la Défense nationale), 2020

CAN UNCLASSIFIED

# **The Assimilation Canadian High Arctic Ionospheric Model (A-CHAIM)**

*Implementation, Workflow, and Performance*

David R. Themens  
University of New Brunswick

Benjamin Reid, Anthony M. McCaffrey, P.T. Jayachandran  
University of New Brunswick

Prepared For:  
Department of National Defence  
DRDC-Ottawa, BLDG 29, 3701 Carling Ave.  
Ottawa, Ontario, K1A0Z4, Canada  
Contract # W7714-186507/001/SS

The following work is supported through the All Domain Situational Awareness (ADSA) Science & Technology program, contract number W7714-186507/001/SS.

- © Her Majesty the Queen in Right of Canada, as represented by the Minister of National Defence, [2020]
- © Sa Majesté la Reine (en droit du Canada), telle que représentée par le ministre de la Défense nationale, [2020]
- © The University of New Brunswick, [2020]

## Abstract

---

The high latitude ionosphere represents a challenging environment for space weather modeling. This region of the ionosphere is very dynamic and undersampled by available observations; thus, models struggle adequately represent this region for operational purposes, such as for use with Over-The-Horizon-Radar (OTHR). It has been shown [e.g. Shim et al., 2011; 2018] that physics-based models of the ionosphere are not capable of matching the performance of empirical climatologic models, such as the International Reference Ionosphere (IRI) and Empirical Canadian High Arctic Ionospheric Model (E-CHAIM) [Themens et al., 2017; 2018a; 2019], regardless of the region tested; however, by design, climatological models cannot adequately capture the short-term variability of the ionosphere and thus are limited in their capacity to represent high latitude regions. While empirical storm adjustments to these models have been implemented to some success, the fact remains that, in order to satisfy the needs of operational users, data assimilation is necessary [Themens et al, 2018b]. To be able to provide the best possible understanding of the current state of the ionosphere, it is necessary to augment traditional empirical models with near-real-time data sources, primarily ionosondes and Global Navigation Satellite System (GNSS) receivers; however, these instruments are both sparsely and unevenly distributed at high latitudes, which creates challenges for traditional assimilation approaches. In this study, we present the first data assimilation system specifically designed to mitigate these limitations for implementation at high latitudes.

Following the development of the Empirical Canadian High Arctic Ionospheric Model (E-CHAIM), we have been tasked with the development of a near-real-time (NRT) data assimilation system based on E-CHAIM. In the following work, we present the methodology, workflow, data composition, and validation of the Assimilation CHAIM (A-CHAIM). This assimilation system follows a sequential maximum likelihood methodology, implemented using a particle filter (e.g. a Monte Carlo-based system), to update E-CHAIM characteristics with NRT measurements from ground-based Global Navigation Satellite System (GNSS) receivers, ionosondes, satellite-borne altimeters, and Incoherent Scatter Radars (ISR).

## Significance for Defence and Security

---

This work includes the development of an NRT data assimilation system designed to augment the best available empirical electron density model (E-CHAIM) for operational use at high latitudes. Project outcomes include the provision of code to operate the assimilation system locally, web services that support the distribution and interpretation of electron density products generated by A-CHAIM run at the UNB Radio Physics Laboratory, and this report detailing the approach and performance of the system. This assimilation model will enable higher fidelity representation of the high latitude ionosphere for operational purposes. The resulting A-CHAIM system will be crucial to the operation and deployment of OTHR systems in the Canadian Arctic, will act as an improved ionospheric environment for communications forecasting, and will improve significantly on current ionospheric specification capacity in this region.

## Résumé

---

L'ionosphère supérieure constitue un environnement complexe pour la modélisation de la météorologie spatiale. Cette région de l'ionosphère en est une très dynamique qui a fait l'objet de très peu d'observations; ainsi, les modèles ont de la difficulté à représenter adéquatement cette région à des fins opérationnelles, notamment pour une utilisation avec le radar transhorizon (OTHR). Il a été démontré [p. ex. par Shim et coll., 2011; 2018] que les modèles de l'ionosphère fondés sur la physique ont un rendement inférieur à celui des modèles climatologiques empiriques, comme le Modèle de l'ionosphère internationale de référence (IIR) et le Modèle empirique ionosphérique de l'Extrême-Arctique canadien (E-CHAIM) [Themens et coll., 2017; 2018a; 2019], sans égard à la région échantillonnée. Cependant, de par leur conception, les modèles climatologiques ne peuvent pas détecter convenablement la variabilité à court terme de l'ionosphère, ce qui explique leur capacité limitée à représenter les régions situées en hautes latitudes. Bien que des ajustements empiriques relatifs aux tempêtes ont été mis en œuvre avec un certain succès dans ces modèles, il est nécessaire d'assimiler les données afin de satisfaire les besoins des utilisateurs opérationnels [Themens et coll., 2018b]. Afin de comprendre le mieux possible l'état actuel de l'ionosphère, il faut bonifier les modèles empiriques traditionnels avec des sources de données en temps quasi réel, principalement des ionosondes et des récepteurs du Système mondial de navigation par satellite (GNSS). Toutefois, ces instruments sont répartis de façon éparse et inégale dans les hautes latitudes, ce qui crée des défis pour les méthodes traditionnelles d'assimilation. Dans la présente étude, nous présentons le premier système d'assimilation de données spécifiquement conçu pour atténuer ces contraintes d'assimilation dans les hautes latitudes.

À la suite de l'élaboration du Modèle empirique ionosphérique de l'Extrême-Arctique canadien (E-CHAIM), nous avons eu pour mandat d'élaborer un système d'assimilation de données en temps quasi réel (TQR) fondé sur le modèle E-CHAIM. Dans la présente étude, nous présentons la méthode, le flux de travail, la composition des données et la validation du modèle CHAIM d'assimilation (A-CHAIM). Ce système d'assimilation repose sur une méthode d'estimation séquentielle de vraisemblance maximale, laquelle est mise en œuvre au moyen d'un filtre à particules (p. ex. un système de type Monte Carlo), afin de mettre à jour les caractéristiques du modèle E-CHAIM avec des mesures en TQR provenant de récepteurs GNSS au sol, d'ionosondes, d'altimètres embarqués sur des satellites et de radars à diffusion incohérente.

# Table of contents

---

Abstract.....	i
Significance for Defence and Security .....	i
Résumé.....	ii
Table of contents .....	iii
List of figures .....	iv
List of tables .....	vii
Acknowledgements .....	viii
1 Introduction.....	1
2 Data Acquisition and Processing .....	2
2.1 Ground-based GNSS Data.....	3
2.1.1 Data Acquisition .....	3
2.1.2 Post-Processing.....	4
2.2 Ionosondes.....	6
2.3 Altimeter Data .....	7
2.4 Satellite-borne GNSS Data.....	8
3 Assimilation System .....	9
3.1 Introduction .....	9
3.2 Parameterization of the System .....	9
3.3 Inverse Modelling.....	12
3.4 Assimilation Workflow .....	14
3.5 User Interfacing with Output Data .....	15
4 Performance.....	16
4.1 Simulations.....	16
4.2 Further Simulations .....	20
4.3 Preliminary Assessment with Real Data .....	23
4.4 Further Testing with Real Data .....	27
4.4.1 Overall Performance of fitting dataset.....	27
4.4.2 Performance against CHAIN assets.....	29
4.4.3 Preliminary performance against assimilation delay .....	31
5 Conclusions.....	33
6 What to look forward to in Phase III: R-CHAIM .....	34
7 Appendix.....	35
References/Bibliography.....	38
List of symbols/abbreviations/acronyms/initialisms .....	40

## List of figures

---

Figure 1. Process diagram for the A-CHAIM NRT data acquisition, processing, and assimilation system. ....	2
Figure 2. Distribution of ground-based GNSS stations available in NRT. Red: Stations within the E-CHAIM domain. Blue: All remaining stations above 30°N. ....	5
Figure 3. Map of the global ionosonde NRT networks currently in operation above 30°N. Purple: GIRO or NOAA MIRION ionosondes providing scaled electron density profiles in NRT. Orange: Roshydromet ionosondes providing only foF2 and hmF2 data in NRT. Red: CHAIN ionosondes (no auto-scaled data available). Blue: Russian Geophysical Observatory stations (website currently down for maintenance, anticipated return to availability is August 2019). ....	6
Figure 4. Example of JASON-3 vTEC data within the E-CHAIM domain for a two-hour orbit file between 7:17 UT and 9:28 UT on June 5 <sup>th</sup> , 2016. ....	8
Figure 5. Example of the assimilation workflow on September 19 <sup>th</sup> , 2019. The dots reflect periods where various datasets are available, where blue represented GNSS data, green represents ionosonde data retrieved from NOAA, black represents ionosonde data retrieved from GIRO, purple represents JASON altimeter data, and red represents ISR data. Vertical lines are used to represent the time the assimilation program was called (blue dotted line), the time the assimilation run began from (black dotted line), and the time that the assimilation was run until (black dashed line). ....	14
Figure 6. Right: An example of the filled in Madrigal TEC map. Left: 4 <sup>th</sup> Order spherical cap harmonic fit of the example in the map on the left. ....	16
Figure 7. Right: The small scale TEC variability remaining after removing Figure 6 (right) from Figure 6 (left) . Left: 14 <sup>th</sup> Order spherical cap harmonic fit of the map on the left. ....	17
Figure 8. Small scale variability superimposed on the background E-CHAIM state for each A-CHAIM assimilation parameter. ....	17
Figure 9. Real measured sTEC (top) at Eureka on September 4 <sup>th</sup> , 2017, with the corresponding simulated measurements in the middle plot and the E-CHAIM unperturbed sTEC in the bottom plot. ....	18
Figure 10. RMS errors of the particle filter assimilation scheme (red), an EnKF scheme (pink), the background model (black), and the theoretical best fit state (blue) for the period between September 7 <sup>th</sup> and September 10 <sup>th</sup> , 2017. The top plot corresponds to the performance in representing the measured sTEC, the second plot corresponds to the errors in representing vTEC, with respect to the truth state, the third plot is that for NmF2 (in e/m <sup>3</sup> ), the fourth is for hmF2 (in km), the fifth is for H <sub>Bot</sub> (in km), and the bottom plot is the corresponding RMS errors in H <sub>Top</sub> (in km). ....	19
Figure 11. Example ionospheric electron density profiles at Eureka (left) and Dourbes (right) during the disturbed periods leading up to the September 8 <sup>th</sup> geomagnetic storm.	

The black dotted lines correspond to a random sample of the particle filter ensemble members, the solid black lines correspond to the background model profile, the blue dashed lines correspond to the true ionospheric state, and the red lines correspond to the particle filter inverted state. .... 20

Figure 12. Overall RMS errors in the electron density at 10 altitudes spanning from 100km to 600km in 50km steps from the assimilation (red) and background (black) using E-CHAIM as the background model in the assimilation. The  $A_p$  index for the period is presented in blue..... 21

Figure 13. Overall RMS errors in the electron density at 10 altitudes spanning from 100km to 600km in 50km steps from the assimilation (red) and background (black) using the IRI as the background model in the assimilation. The  $A_p$  index for the period is presented in blue..... 21

Figure 14. Vertical profiles of RMS errors in plasma frequency over the course of the simulation for the assimilation system (top) and the background model (bottom)..... 22

Figure 15. Difference between background and assimilation errors in plasma frequency with respect to altitude (blue implies improvement by the assimilation)..... 23

Figure 16. Left) An example  $vTEC$  contour map generated from the assimilation system with station measured  $vTEC$  indicated by the filled circles. Right) RMS errors in  $sTEC$  for each station. The thin lines reflect the  $vTEC$  contours of the figure on the left. .... 24

Figure 17. RMS errors in foF2 (top), hmF2 (middle), and sTEC (bottom) for the background E-CHAIM model (black) and the assimilation system (red)..... 25

Figure 18. Example foF2 at Qikiqtarjuaq from CADI (blue), E-CHAIM (grey), and A-CHAIM (orange)..... 26

Figure 19. Same as Figure 18 but for Cambridge Bay. .... 26

Figure 20. RMS errors in foF2 (top), hmF2 (middle), and sTEC (bottom) using E-CHAIM (black) and A-CHAIM (red), calculated against the A-CHAIM assimilation dataset..... 27

Figure 21. foF2 at Boulder from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period. .... 28

Figure 22. foF2 at Juliusruh from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period. .... 29

Figure 23. foF2 at Qikiqtarjuaq from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period. .... 30

Figure 24. foF2 at Resolute from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period. .... 30

Figure 25. foF2 at Sachs Harbour from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period. .... 31

Figure 26. Top: Ionogram using the ARTIST autoscaling program, demonstrating an early-stopping error with 100% quality score. Bottom: Ionogram with corrected manual

scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile. .... 35

Figure 27. Top: Ionogram using the ARTIST autoscaling program, demonstrating an early-stopping error with 80% quality score and multi-hop traces. Bottom: Ionogram with corrected manual scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile. .... 36

Figure 28. Top: Ionogram using the ARTIST autoscaling program, erroneously scaling the E-region as the F-region with 100% quality score. Bottom: Ionogram with corrected manual scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile. Note: Even with the correct manual scaling, the inverted profile is unphysical as the inversion software's choice of valley is exaggerated. In the case of the manually scaled profile, only the F2-peak parameters would be used by the assimilation system. 37

## List of tables

---

Table 1. List of the data sources accessed for the NRT data assimilation system.....	3
Table 2. An example of how the assimilation state could be defined, with different number of state variables available for each profile parameter of interest. ....	11
Table 3. Overall foF2 RMS errors for the period for the September test period using E-CHAIM and A-CHAIM with various assimilation delay times. ....	32

## Acknowledgements

---

Infrastructure funding for CHAIN was provided by the Canadian Foundation for Innovation and the New Brunswick Innovation Foundation. Science funding is provided by the Natural Sciences and Engineering Research Council of Canada. The authors would like to thank the many ionosonde operators who provided data to this project. This paper uses ionospheric data from the USAF NEXION Digisonde network, the NEXION Program Manager is Mark Leahy. This publication makes use of data from the Qaanaaq and Nord ionosondes, owned by the U.S. Air Force Research Laboratory Space Vehicles Directorate and supported in part by the Air Force Office of Scientific Research. The authors thank Svend Erik Ascanius of the Danish Meteorological Institute and Denmark's Arctic Command for the operation of these ionosondes. This paper uses data from the Juliusruh Ionosonde which is owned by the Leibniz Institute of Atmospheric Physics Kuehlungsborn. The responsible Operations Manager is Jens Mielich. The authors are grateful to Konstantin Ratovski for the operation of the Irkutsk ionosonde. The Tromsø ionosonde is funded in part by QinetiQ.

# 1 Introduction

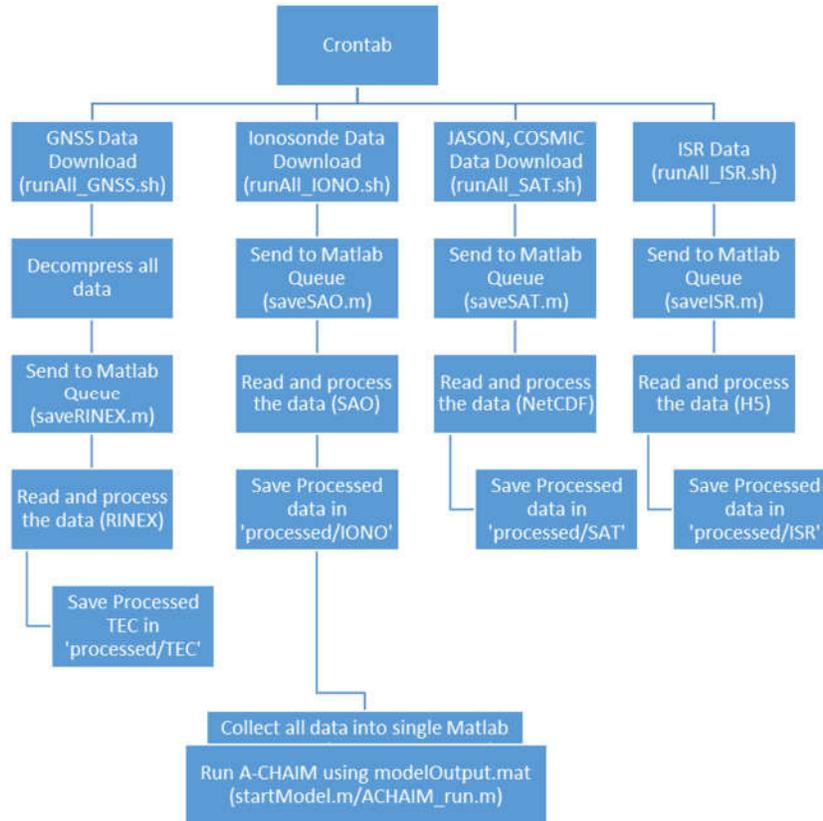
---

The following report details the workflow, development, and performance of a high latitude electron density assimilation system based on the Empirical Canadian High Arctic Ionospheric Model (E-CHAIM) and developed for use by DRDC. All references to this assimilation system herein will refer to the assimilation system as the Assimilation CHAIM (A-CHAIM).

In Section 2, we detail the production of routines to gather, process, and calibrate ionospheric observations from ground-based ionosondes and Global Positioning System (GPS) receivers. In Section 3, we detail the near-real-time data assimilation framework before discussing the performance of the assimilation in Section 4.

## 2 Data Acquisition and Processing

Perhaps the most challenging component of an NRT data assimilation system is the gathering and processing of the required data in an accurate and timely manner. This requires careful tracking of data through the pipeline and intricate timing of the required processes. For all data sources, servers are available to download the data. In most cases these are publicly available. A logical diagram for the download and processing of the NRT data is provided in Figure 1.



**Figure 1. Process diagram for the A-CHAIM NRT data acquisition, processing, and assimilation system.**

There are four groups of data that are being downloaded in near-real time. These are GPS observables (RINEX format), COSMIC podTEC (NetCDF format), JASON altimeter data (NetCDF), and Ionosonde observables (SAO format). We obtain all available data within the last 3 full hours, for all stations located within the E-CHAIM boundary (where applicable). The required data is distributed across twelve different databases, listed in Table 1.

**Table 1. List of the data sources accessed for the NRT data assimilation system.**

Source	Data Type	Service Type
German Federal Agency for Cartography and Geodesy (BKG)	Ground-based GNSS	ftp
Canadian High Arctic Ionospheric Network (CHAIN)	Ground-based GNSS	ftp
Crustal Dynamics Data Information System (CDDIS)	Ground-based GNSS	ftp
Hexagon Geosystems SmartNet	Ground-based GNSS	ftp
NOAA National Oceanographic Data Center (NODC)	Altimeter	ftp
UCAR COSMIC Data Analysis and Archive Center (CDAAC)	Space-borne GNSS	ftp
NOAA National Geodetic Survey (NGS)	Ground-based GNSS	ftp
NOAA National Geophysical Data Center (NGDC)	Ionosonde	ftp
California Spatial Reference Center (CSRC) GARNER GPS Archive	Ground-based GNSS	ftp
Natural Resources Canada (NRCan)	Ground-based GNSS	ftp
Swedish Institute of Space Physics (IRF)	Ground-based GNSS	http
Global Ionospheric Radio Observatory (GIRO)	Ionosonde	https

## 2.1 Ground-based GNSS Data

Ground-based GNSS data is perhaps the most widely used dataset in ionospheric modeling. GNSS receivers are able to determine the total electron content (TEC), e.g. the integrated electron density, along the receiver to satellite ray path. With each receiver being able to log data from 6-12 satellites at a given time, a single receiver can cover a significant spatial area. Its reliable availability, low latency, and spatial coverage make it ideal for ionospheric data assimilation; however, the fact that these are path-integrated measurements requires that tomographic or other advanced inversion approaches be used in order to apply these measurements to infer information on the vertical structure of the ionosphere.

### 2.1.1 Data Acquisition

Ground-based GNSS data is downloaded from eight different sources. These RINEX data providers are polled three times per hour, at 20-minute intervals, starting at 10 minutes past the hour. To limit the number of times we poll a server, we have created an approximate estimate of the RINEX data latency, for each station, based on 7 days of data uploads. We obtain the mean and standard deviation of the latency. If the current time of any given attempt to download data is less than the mean plus one standard deviation of the latency, we do not check for new data. Otherwise, a check for new data for a desired station is performed and if new data is available it is downloaded.

Phase 1 processing involves the unpacking of compressed files. In this case, only the RINEX files come compressed, with various forms of compression needing to be dealt with. The files can be zipped (.zip compressed), gzipped (.gz), or Unix compressed (.Z). When a file is downloaded with any of these compressions applied, it is immediately decompressed using the appropriate

utility. The resulting RINEX files may be further compressed using Hatanaka compression. In these cases, the files are decompressed using the Hatanaka decompression utility provided by SOPAC.

As soon as decompressed files become available (either after decompressing, in the case of the RINEX data, or after downloading for all other data types), the files are read and saved to Matlab binary files as is; the files are read and parsed into structures and arrays, but no data manipulation or post-processing takes place. For some data types, this is the final step (like COSMIC podTEC for example) and these binary files will be used by the model during data ingestion.

For the RINEX data, RINEX readers/parsers have been written in-house, based on the RINEX v2.11 and RINEX v3.03 format specifications.

## 2.1.2 Post-Processing

Slant Total Electron Content (sTEC) is calculated the traditional way [Komjathy, 1997]

$$sTEC = 9.52(\Delta\Phi + N + DCB_s + DCB_r)$$

where  $\Delta\Phi$  is the difference of the L1 and L2 carrier phase observables (in meters),  $N$  is the integer ambiguity, and  $DCB_s$  is the satellite Differential Code Bias (DCB), and  $DCB_r$  is the receiver DCB. All satellite DCBs are provided by CDDIS; some receiver DCBs are provided as well. We use the most up-to-date bias file that is available at the given processing time to remove all satellite DCBs and, when available, the receiver DCBs as well. In these cases, the receiver DCB is used as an initial guess, since receiver biases are solved within the assimilation solution itself.

The integer ambiguity is a result of the ambiguous nature of the GPS carrier phase observable. This can be resolved with the pseudorange-derived sTEC, which is not ambiguous, but is considerably noisier than the phase-derived sTEC. To solve for the integer ambiguity, an elevation-weighted difference between the phase-derived and pseudorange-derived sTEC is calculated for each arc. This technique is known as phase-levelling. Here we define an arc as any continuous set of sTEC measurements. If a hole is encountered in the data, a new arc is started. Although this is unrealistic for the typical definition of an arc, this serves as a cycle slip correction method. Phase-levelling is performed using the procedure outlined in [Carrano and Groves, 2009]. As previously mentioned, the integer ambiguity is defined as the difference between the phase-derived and pseudorange-derived sTEC:

$$n = sTEC_\Phi - sTEC_{PSR}$$

and is weighted according to the elevation using:

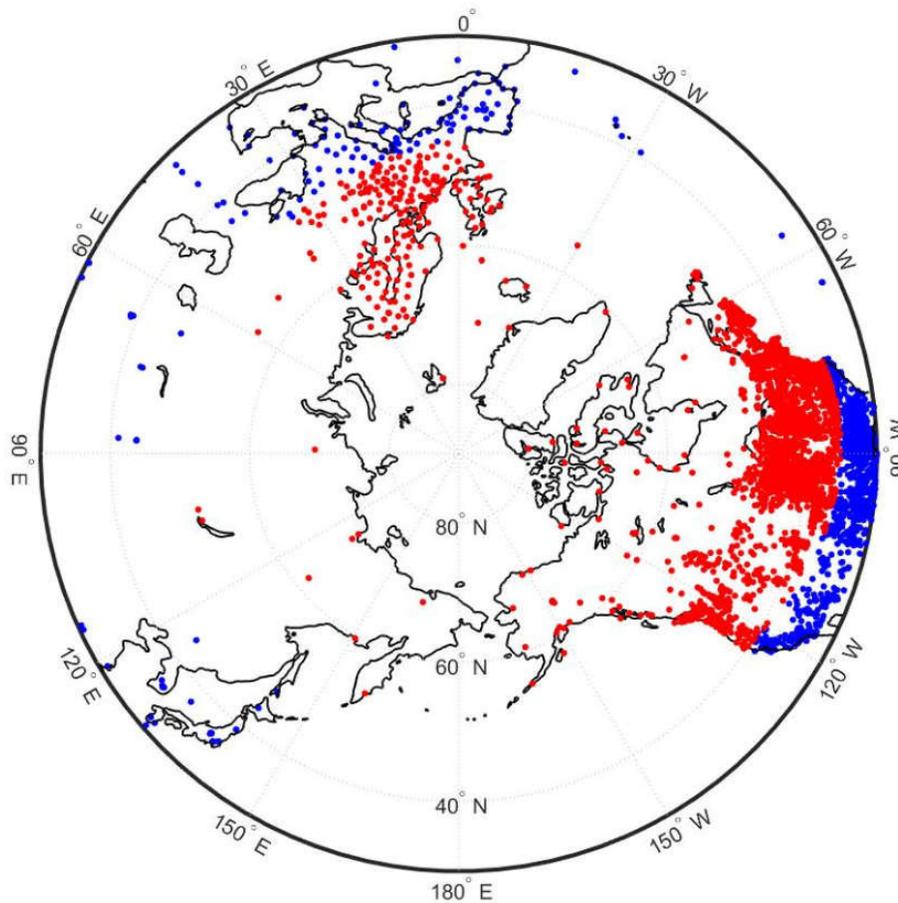
$$N = \frac{\sum_i w_i n_i}{\sum_i w_i}$$

where

$$w_i = \sin(e_i)$$

where  $i$  refers to the current epoch, and  $e$  is the satellite elevation. Note that data associated with elevation angles less than  $20^\circ$  are not used in calculating the integer ambiguity.

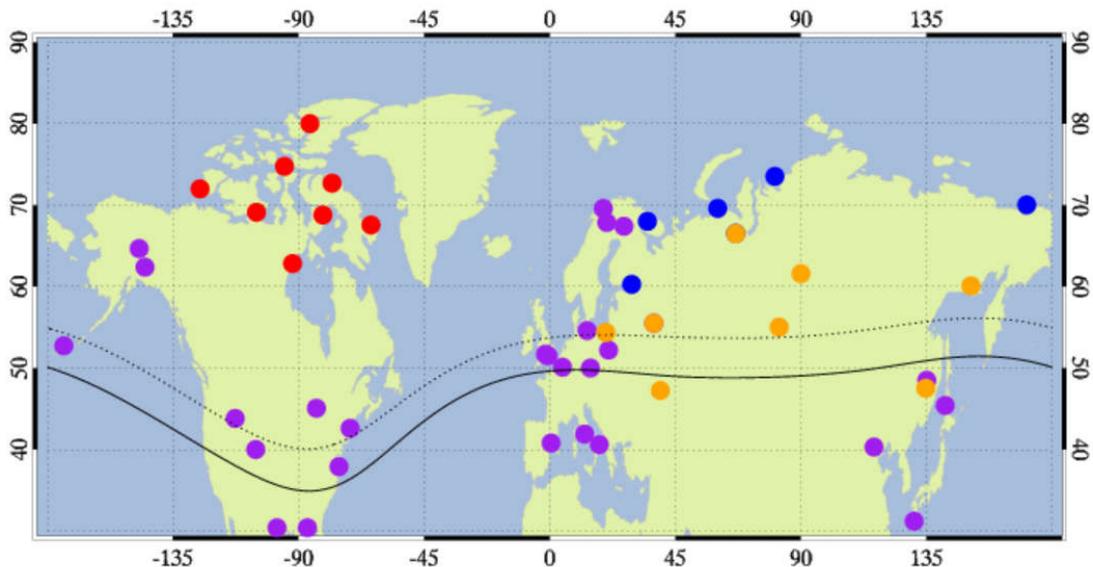
Quality control for the phase-levelling routine is performed using a standard deviation of the difference between the phase-derived sTEC (after removing the integer ambiguity) and the pseudorange-derived sTEC. Arcs with standard deviations greater than 3.5 TECU are not ingested into the model.



**Figure 2. Distribution of ground-based GNSS stations available in NRT. Red: Stations within the E-CHAIM domain. Blue: All remaining stations above  $30^\circ$ N.**

## 2.2 Ionosondes

Ionosondes are vertically pointed HF radars capable of providing the vertical electron density profile up to the height of the peak density of the ionosphere (hmF2). These instruments have been used for ionospheric specification since the discovery of the ionosphere and formed an important component of the dataset used to build the Empirical Canadian High Arctic Ionospheric Model (E-CHAIM). The use of ionosonde data can, however, be challenging. The process of deriving electron density information from ionosonde echoes requires the use of an inversion process to correct the effect of ionospheric delays on the group range of the echoes. Also, ionograms, images of the group range of echoes for an HF frequency sweep, must be processed to extract the correct ionospheric echoes. This process, generally referred to as “scaling”, is a challenging one, as interference, ionospheric irregularities, and sporadic ionospheric structures can significantly degrade the quality of ionograms. Automated routines developed for this process (autoscalers) are somewhat reliable at midlatitudes but can suffer significant errors at high latitudes, where ionospheric irregularities and sporadic structures are frequently observed. For the A-CHAIM NRT system, we make use of data autoscaled by ARTIST, developed by the University of Massachusetts Lowell, and Autoscala, developed by the Italian National Institute of Geophysics and Volcanology (INGV). ARTIST data is given by the purple dots in Figure 3, while Autoscala data correspond to the orange dots.



**Figure 3. Map of the global ionosonde NRT networks currently in operation above 30°N. Purple: GIRO or NOAA MIRION ionosondes providing scaled electron density profiles in NRT. Orange: Roshydromet ionosondes providing only foF2 and hmF2 data in NRT. Red: CHAIN ionosondes (no auto-scaled data available). Blue: Russian Geophysical Observatory stations (website currently down for maintenance, anticipated return to availability is August 2019).**

In order to filter out unreliable data, outlier removal is performed. Firstly, any data associated with hmF2 measurements that are unavailable or that are outside the range of 175 – 450 km are deemed unphysical and are removed. Scalings with hmF2 outside of this range within the E-CHAIM domain are generally associated with the accidental scaling of sporadic-E as an F-region trace, with the scaling of the F1 trace as the F2 trace, or with the scaling of a secondary hop as the first hop trace. Secondly, outlier detection is performed using the Inter-Quartile Range (IQR), executed by the Matlab provided ‘isoutlier’ function. Outlier detection uses all available data from the current day and the 5 previous days. If the data is considered an outlier, it is removed. This accounts for the common error of autoscalers quitting mid-way up the F-region trace or failing to scale the F2 trace in the presence of an F1 layer. The remaining minor autoscaling errors can be handled within the assimilation system itself through the use of overlapping, complementary measurements. Examples of common autoscaling errors are provided in the Appendix in Section 7.

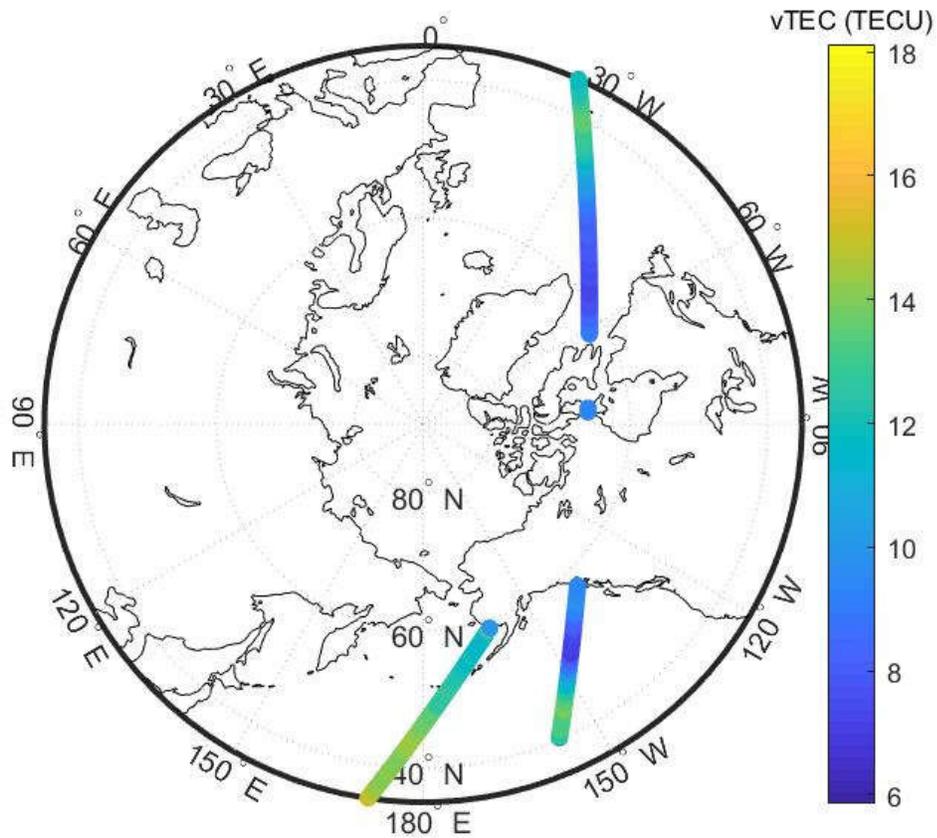
Ionosonde data is polled every 10 minutes, on the hour. All new data, for stations of interest, is downloaded. The ionosonde SAO files are read using an in-house reader, following the SAO v4 format specification.

## 2.3 Altimeter Data

The NRT system also makes use of space-borne altimeter data from the Jason-2 and Jason-3 satellite missions. As a byproduct of the altimeter solution for sea-surface height, vertical ionospheric TEC above the ocean can be inferred. This is done following the same concept as GNSS TEC products, where Jason’s Ku band antenna excess phase can be directly related to the TEC along the ray path, e.g.

$$vTEC = \frac{dR \cdot f^2}{40.3}$$

where  $dR$  is the excess ground range and  $f$  is the signal frequency (13.575 GHz). The resulting TEC is then filtered to remove outliers and ground scatter. While the overall precision of the JASON TEC is  $\sim 4$  TECU, it is largely unbiased and provides a crucial constraint over the oceans, where no other dataset has adequate coverage. Jason data is polled at the top of every hour. The Jason altimeter NetCDF files are read using the Matlab provided function ‘nread’.



*Figure 4. Example of JASON-3 vTEC data within the E-CHAIM domain for a two-hour orbit file between 7:17 UT and 9:28 UT on June 5<sup>th</sup>, 2016.*

## 2.4 Satellite-borne GNSS Data

Finally, the NRT system also makes use of COSMIC podTEC line of sight TEC observations. We here choose to assimilate the direct COSMIC TEC observations instead of their traditional inverted radio occultation electron density product, as there are many assumptions made in the RO inversion process that result in 1) performance errors at high latitudes and 2) a loss of overall information. By assimilating the podTEC directly, we preserve the highest amount of ionospheric information possible. All processing of podTEC is completed by CDAAC prior to acquisition by the A-CHAIM system.

COSMIC podTEC data is polled at the top of every hour. The COSMIC podTEC NetCDF files are read using the Matlab provided function 'ncread'.

## 3 Assimilation System

---

### 3.1 Introduction

In order to describe or model a physical system, it is necessary to construct a representation of that system. Data assimilation describes a family of techniques by which physical and empirical models are combined with real measurements in order to produce the most accurate representation of the system given the information available. In order to perform data assimilation there are three procedures which must be completed:

1. **Parameterization of the system:** A set of parameters  $x$  which allow the system to be characterized. These parameters should ideally be able to capture the full range of physical behaviours of the system. These parameterizations are in general not unique for a given system, and so the choice of parameterization can be informed by the objective of the assimilation. The entire space spanned by  $x$  is called the state space, a  $n$  dimensional space where  $n$  corresponds to the number of parameters.
2. **Forward Modelling:** Given a set of model parameters  $x$ , the forward model allows us to make predictions. We can represent this with the forward model operator  $J(x)$  such that, for a set of measurements of the system  $z$

$$z = J(x) \tag{1}$$

The form of  $x$  is determined by parameterization, and the type of measurement.  $J(x)$  is in general a nonlinear operator.

3. **Inverse Modelling:** Given a set of measured observables  $z$ , inverse modelling allows us to reconstruct the state variables  $x$  which best describe the system. This state should then allow us make predictions using the forward model that are more accurate. As the dimension  $n$  of the state space increases, any inversion inevitably becomes poorly conditioned.

Specification of these three components defines an assimilation system.

### 3.2 Parameterization of the System

In atmospheric modelling, it is traditional to parameterize the system using a 2D or 3D geographical grid of points, and to define the state as the set of atmospheric characteristics, such as electron density, at each of those points. While this does allow for a very high spatial resolutions, it does require a very large number of state variables. Not only does this increase computational costs, but it also effectively guarantees that the solution to the inverse model is poorly conditioned. It can also be challenging to ensure that the state is physically realistic, even if the solution is otherwise well conditioned. Finally, grid-based parameterizations also make it non-trivial to update the state in regions with little or no available measurements. Given the sparse and uneven distribution of ionospheric measurements at high latitudes, these challenges demand a different approach.

In A-CHAIM, we use a semi-Epstein layer to prescribe the vertical electron density above each geographic point in the high-latitude region. The shape of this profile is governed by a set of parameters:

$$x_{profile} = [NmF2, hmF2, hmF1, hmE, H_{Bot}, H_{Top}, H_{HF1}, H_E] \quad (2)$$

Electron density for a height  $h$  is given by:

$$N_e(h, x_{profile}) = NmF2 \cdot \text{sech}^2\left(\frac{h - hmF2}{H(h, x)}\right) \quad (3)$$

where

$$H(h, x) = \begin{cases} 2H_{Top} \cdot \left(1 + \frac{rg(h - hmF2)}{rH_{Top} + h(h - hmF2)}\right) & h \geq hmF2 \\ H_B(h, x) \cdot \left(\frac{1}{1 + \exp\left(\frac{hmE - 15 - h}{2.5}\right)}\right) & h < hmF2 \end{cases} \quad (4)$$

where

$$H_B(h) = H_{Bot} + H_{F1} \cdot \text{sech}^2\left(\frac{h - hmF1}{0.5(hmF2 - hmF1)}\right) + H_E \cdot \text{sech}^2\left(\frac{h - hmE}{25}\right) \quad (5)$$

This vertical parameterization is identical to the parameterization used by the E-CHAIM model. It has several advantages for this assimilation application, in particular as it prescribes the entire vertical structure at any geographic position with only 8 parameters. Furthermore, it guarantees that the output of the assimilation is smooth and differentiable. As E-CHAIM serves as the basis of A-CHAIM, this parameterization has the added benefit of making it very simple to transfer information from our background model to the assimilation.

In order to address the challenges that grid-based parameterizations create, we instead build our horizontal variations using a set of basis functions. These functions are a modified version of the spherical harmonics in geographic coordinates:

$$f(\theta, \lambda) = \sum_{l=0}^{\infty} \sum_{m=0}^l \bar{P}_{lm}(\cos\theta) [\bar{C}_{lm} \cos(m\lambda) + \bar{S}_{lm} \sin(m\lambda)] \quad (6)$$

The basis is transformed such that the polar axis of the basis set is aligned with the geomagnetic pole; thus, the colatitude  $\theta$  is defined as a geographic angular distance from the geomagnetic pole, and the longitude  $\lambda$  measures a geographic angular distance around the geomagnetic axis. In addition, the colatitude measure is scaled so that the colatitude at  $180^\circ$  has been moved from the southern geomagnetic pole has been moved up to the  $39^\circ$  parallel. This provides a much more efficient basis set than the unmodified spherical harmonics in the region of interest, while still

being relatively inexpensive to calculate. For any practical purpose thus far, the maximum number of spherical harmonics needed has been 225, corresponding to Legendre polynomials up to 14th order. Therefore, using the 8 profile shape parameters, a maximum of 1800 parameters are needed to define the electron density in the entire 3D volume above 45° magnetic latitude. In practice only 1325 are used, corresponding to polynomials up to 12th orders. The E-CHAIM model can output maps of each of the profile parameters at a given time, and so we can populate the state easily using an inverse spherical harmonic transformation.

While it may be necessary to store 1325 variables in order to describe the entire ionosphere, we can further reduce the number of parameters needed in the state. Many of the profile parameters have a very limited effect on the measurements that are available, and so are unlikely to be improved by assimilation. We can therefore leave these variables unchanged from the values given by E-CHAIM, in order to focus on variables that have a greater effect on the overall profile. Specifically, we are interested in the F2 peak density  $NmF2$ , the F2 peak height  $hmF2$ , the topside scale thickness  $H_{Top}$ , and the bottomside scale thickness  $H_{Bot}$

$$x_{profile} = [NmF2, hmF2, H_{Bot}, H_{Top}] \quad (7)$$

We can further reduce the number of variables in the state space by modifying how we parameterize the state. While it may be required to use a full 12 orders of spherical harmonics to span the data space to the required precision, we are unlikely to be able to resolve variations at those fine scales which correspond to the highest orders. What we do in practice is store the full 12th order expansion for every profile parameter, which we generate at every timestep of the assimilation. These parameters are kept as static, they are not updated as part of the assimilation. We then describe any changes to this base model as a series of perturbations, which are themselves built from a spherical harmonic series. However, we can choose to use fewer orders of spherical harmonics for these perturbations than we would need to span the entire data space with an absolute measurement. As a simple example, consider updating the peak density  $NmF2$  using only the 0th order harmonic. While we still require the 12 orders to accurately capture the finer structures. As a simple example, consider a situation where the E-CHAIM is underestimating the peak density  $NmF2$  globally, but otherwise has the correct general form of the ionosphere. While we still need the 12th order spherical harmonics to describe the finer structures in  $NmF2$ , we can correct the global result using only a single parameter, the 0th order harmonic. By building our state space with these perturbations, we can reduce our state space to one- to two-hundred variables. We can also choose to modify our effective resolution for different profile parameters, in order to take advantage of the varying sensitivity of our measurements to each profile parameter.

**Table 2. An example of how the assimilation state could be defined, with different number of state variables available for each profile parameter of interest.**

Profile Parameter	Basis Order	# of Parameters
$NmF2$	6 <sup>th</sup> order	49
$hmF2$	2 <sup>nd</sup> order	9
$H_{Top}$	3 <sup>rd</sup> order	16
$H_{Bot}$	2 <sup>nd</sup> order	9

With this approach, we are able to use all of the available information from our background model. As the E-CHAIM model is smooth, and our perturbations are themselves smooth, the output of the assimilation will also be smooth. We are able to fine tune the available spatial resolution for different elements of the vertical profile, in order to make the best use of the measurement information. It is also very simple to perform forward modelling, as calculating the electron density at any point only requires calculating the modified colatitude and longitude, and performing the inverse basis transformation for each profile parameter.

### 3.3 Inverse Modelling

Now that we have parameterized our state, it is possible to describe an assimilation scheme. A number of techniques have been developed which could be used in this situation. We use a Sequential Monte Carlo Maximum Likelihood Estimation approach, as it has proven to be more stable than an Extended or Ensemble Kalman Filter, while still being able to deal with the highly nonlinear measurement operator  $H(x)$ . The setup is characterized as follows. The state of the ionosphere at some time  $t_n$  is described by some unknown set of parameters  $x_n$ , an element of our state space. The probability of any given state is dependent on all of the states that came before it  $\{x_{1:n-1}\}$  and all of the measurements up to and including  $t_n$ . Given a set of measurements  $z_n$ , the probability of a given element of the state space is:

$$p(x_n|z_{1:n}) = \frac{p(z_{1:n}|x_{1:n})p(x_{1:n})}{p(z_{1:n})} = \frac{p(z_{1:n}|x_{1:n})p(x_{1:n})}{\int p(z_{1:n}|x_{1:n})p(x_{1:n})dx} \quad (8)$$

Defining and evaluating these probability distributions is impossible for the problem at hand. We can estimate the true value of the state by finding the element of state space which best reproduces the observations at a given time. This requires us to find the set of parameters which maximize the likelihood function

$$l(x_n|z_n) = p(z_n|x_n) \quad (9)$$

While we cannot find the maximum of the likelihood function analytically, we can approximate the maximum value using a Monte Carlo approach. By choosing a random sample  $\mathbf{X}_n^i$  drawn from a distribution with the same support as  $l(x_n|z_n)$ , we can evaluate an empirical measure of the log likelihood as

$$\log(\hat{l}(x_n|z_n)) = \sum_{i=1}^N -0.5(\mathbf{z}_n - H(\mathbf{X}_n^i))^T \mathbf{R}_n(\mathbf{z}_n - H(\mathbf{X}_n^i))\delta x_n^i \quad (10)$$

One possible estimate of the maximum log likelihood is to select the sample that had the maximum log likelihood, but in practice this can often be a poor estimator. Due to the noisy and sparse nature of our observations, it is not uncommon to find a sample that does a good job approximating the specific set of data available at a given time, but small perturbations of this sample fail to be a good estimate. We would expect all samples near the true value of the state to be better than the overall population, even if none of them individually have the best log likelihood.

By selecting appropriate normalized weighting factors  $W_n^i$ , we can estimate the expectation value

$$\hat{x}_n = \sum_{i=1}^N W_n^i \delta x_n^i(x_n) \quad (11)$$

These weights are determined by analyzing the log likelihood of each type of observation separately. Each observation type  $k$  gives us a subweight  $w_{nk}^i(\lambda_k)$ , that is scaled by a factor  $\lambda_k$ , which we select for each measurement type.

$$w_{nk}^i = \exp \left[ \frac{-0.5 \lambda_k \hat{l}(\mathbf{X}_n^i | z_{nk})}{\frac{1}{N} \sum_{s=1}^N |\hat{l}(\mathbf{X}_n^s | z_{nk})|} \right] \quad (12)$$

The optimal distribution from which to draw our sample  $\mathbf{X}_n^i$  is in fact the posterior distribution of the previous time step  $\mathbf{X}_{n-1}^i$ . We can then calculate appropriate weights for each particle by comparing their predicted measurements to the real measurements. The normalized weights  $W_n^i$  are calculated from the unnormalized weights  $w_n^i$ , where  $\mathbf{R}$  is the error covariance matrix of the measurements  $\mathbf{z}$

$$W_n^i = \frac{w_n^i(\mathbf{X}_{1:n})}{\sum_{i=1}^N w_n^i(\mathbf{X}_{1:n})} \quad (13)$$

where

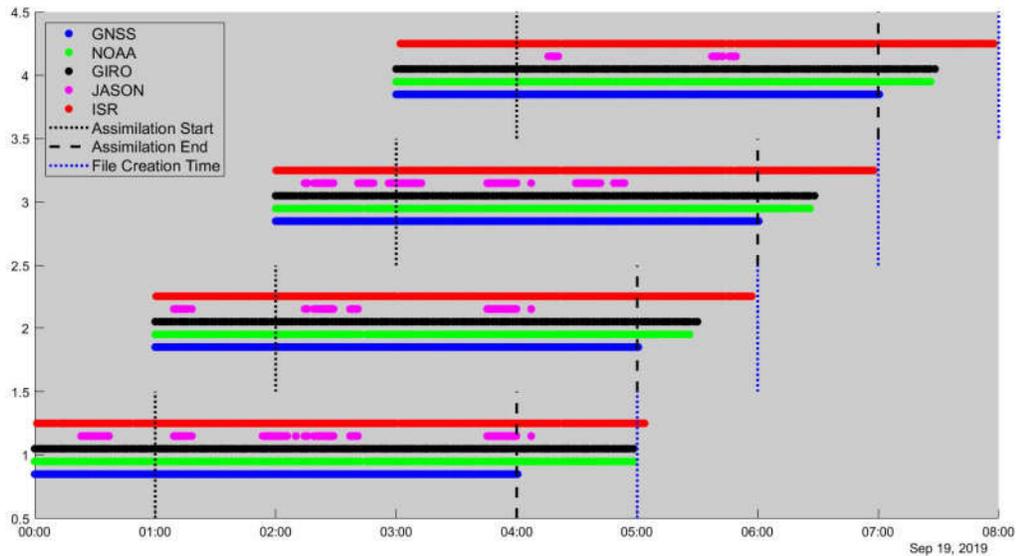
$$w_n^i = w_{n-1}^i p(z_n | x_n) = w_{n-1}^i \left( z_n - H(x_n^i) \right)^T \mathbf{R}_n \left( z_n - H(x_n^i) \right) \quad (14)$$

As we step forward through time, the weights assigned to each low-probability particle become exponentially small and can quickly become numerically unstable. As a result, it is occasionally necessary to perform a resampling procedure. A-CHAIM uses Systematic Resampling, but all resampling techniques have a similar aim, to re-generate a new distribution of particles with more optimal weights. Particles with very low weights do not contribute meaningfully to the estimation, and so resampling aims to replace low weight particles with high weight ones in a way which preserves the statistical validity of the ensemble. After resampling, all weights  $w_n^i(x_{1:n})$  are set to  $\frac{1}{N}$ . While resampling is essential to preserve the usefulness of the ensemble, resampling also increases the variance of the ensemble. Resampling is only performed when the Effective Sample Size (ESS) drops below  $\frac{N}{2}$

$$ESS = \left( \sum_{i=1}^N (W_n^i)^2 \right)^{-1} \quad (15)$$

### 3.4 Assimilation Workflow

As illustrated in Figure 1, at the top of each hour, a .mat file is assembled containing all datasets that were available within the previous four hours. Following the creation of this file (say, at time A), the assimilation system is run, starting at a time four hours prior to the .mat file creation time (say, at time B). The results of the previous assimilation run at time B are used to initialize this new assimilation run. If no previous run was available, the system starts from a fresh state. This assimilation run conducts data assimilation on 15-minute intervals until the end of the data file is reached (say, at time C, e.g. three hours after time B). An illustration of the assimilation timing is provided in Figure 5. In this figure, time A is illustrated with a dotted blue line, time B is illustrated with a dotted black line, and time C is marked with a black dashed line.



**Figure 5. Example of the assimilation workflow on September 19<sup>th</sup>, 2019. The dots reflect periods where various datasets are available, where blue represented GNSS data, green represents ionosonde data retrieved from NOAA, black represents ionosonde data retrieved from GIRO, purple represents JASON altimeter data, and red represents ISR data. Vertical lines are used to represent the time the assimilation program was called (blue dotted line), the time the assimilation run began from (black dotted line), and the time that the assimilation was run until (black dashed line).**

This schedule produces output assimilation files every 15 minutes, with two hours of overlap between subsequent runs. Periods that overlap with previous runs are re-assimilated to reflect the fact that data is being acquired continuously and thus previous assimilations may be improved upon using newly available data. This is an important consideration, as the prior state is used in subsequent assimilation timesteps; thus, any improvement to the previous state will result in improved pre-conditioning the next assimilation step.

### 3.5 User Interfacing with Output Data

Output files for the most recent A-CHAIM assimilation run are provided to the user at <https://chain-new.chain-project.net/index.php/chaim/a-chaim/a-chaim-output-model-files>. Previous run output files are stored for 60 days on the CHAIN server and may be made available upon request. These output files store the E-CHAIM parameter set that was used for the assimilation run, as well as perturbation parameters generated by the assimilation system. No account is necessary in order to access these output files.

To extract electron density from these files, we have provided codes in Matlab and C, available at <https://chain-new.chain-project.net/index.php/chaim/a-chaim/a-chaim-software-downloads>. Users are requested to create an account in order to access this software. Account are used only to ensure that the CHAIM group can contact users with important maintenance or update information. Accounts are free and automatically authorized, so there is no barrier to use. Instructions and examples on the use of the A-CHAIM output file interface software are provided in the software package readme.txt file.

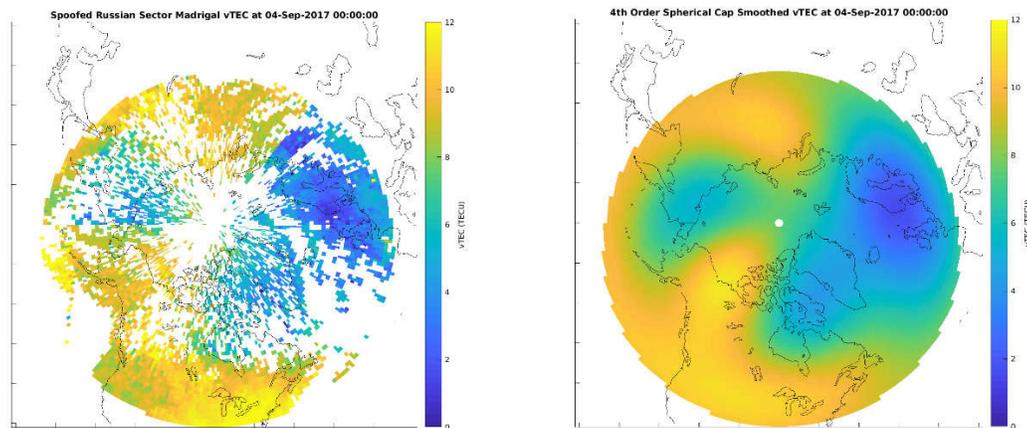
In addition to providing output files and associated interpretation software, we also provide a website interface, functionally identical to what is already provided for E-CHAIM, at <https://chain-new.chain-project.net/index.php/chaim/a-chaim/a-chaim-web-application>. This web application can be used to generate 2D contours of A-CHAIM-derived ionospheric parameters, as well as ASCII files of requested parameters.

## 4 Performance

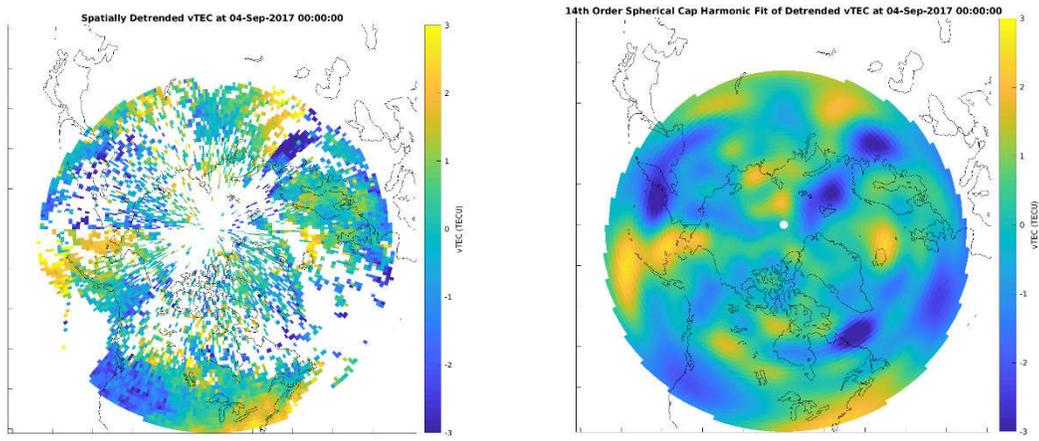
---

### 4.1 Simulations

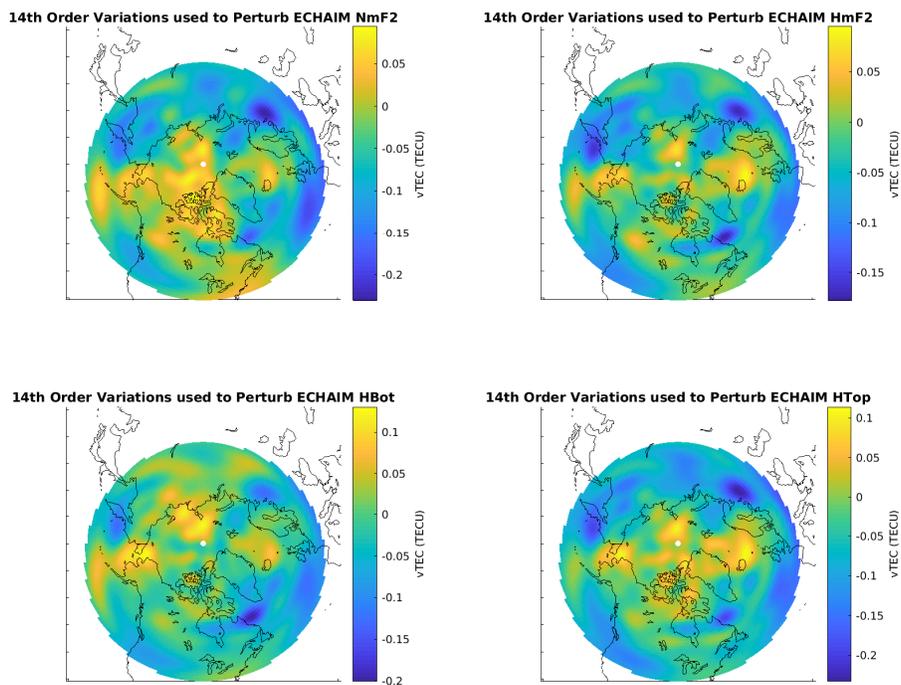
In order to arrive at the above choice of assimilation framework we conducted a series of simulations using a realistic synthetic ionosphere. This ionosphere is generated by superimposing small-scale spatial and temporal variability on an E-CHAIM background. This variability is simulated using Madrigal TEC maps that have been artificially filled in to account for large data gaps. First, we remove a 4<sup>th</sup> order spherical cap harmonic fit of the TEC maps from the data. This fit is illustrated in Figure 6. The remaining small-scale variability is then fit to a set of 14<sup>th</sup> order spherical cap harmonics, illustrated in Figure 7. The remaining distribution of spherical harmonics are randomly perturbed, illustrated in Figure 8, and added the background E-CHAIM state. The resulting state is used as the experimental truth for the following assimilation simulations.



**Figure 6.** *Right: An example of the filled in Madrigal TEC map. Left: 4<sup>th</sup> Order spherical cap harmonic fit of the example in the map on the left.*

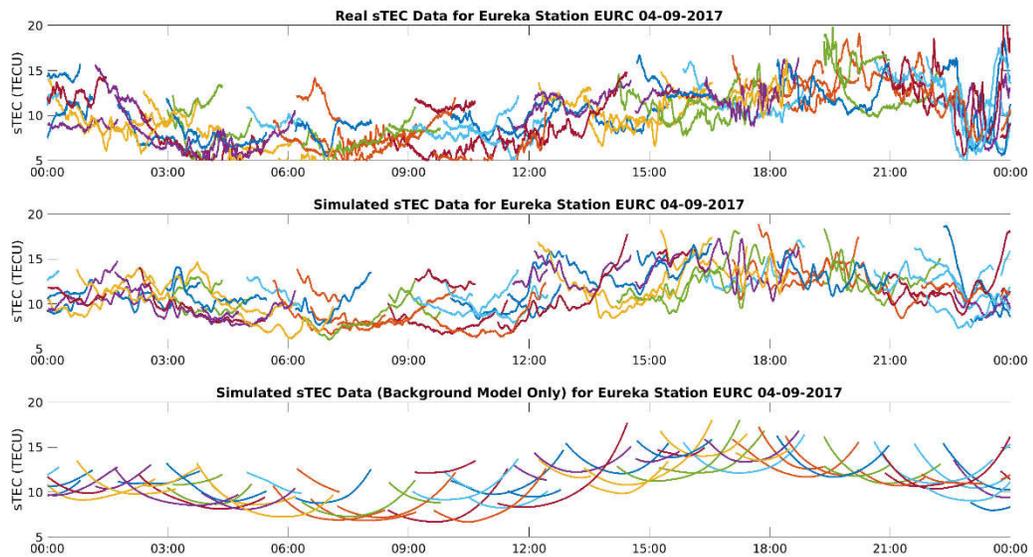


**Figure 7.** Right: The small scale TEC variability remaining after removing Figure 6 (right) from Figure 6 (left) . Left: 14<sup>th</sup> Order spherical cap harmonic fit of the map on the left.



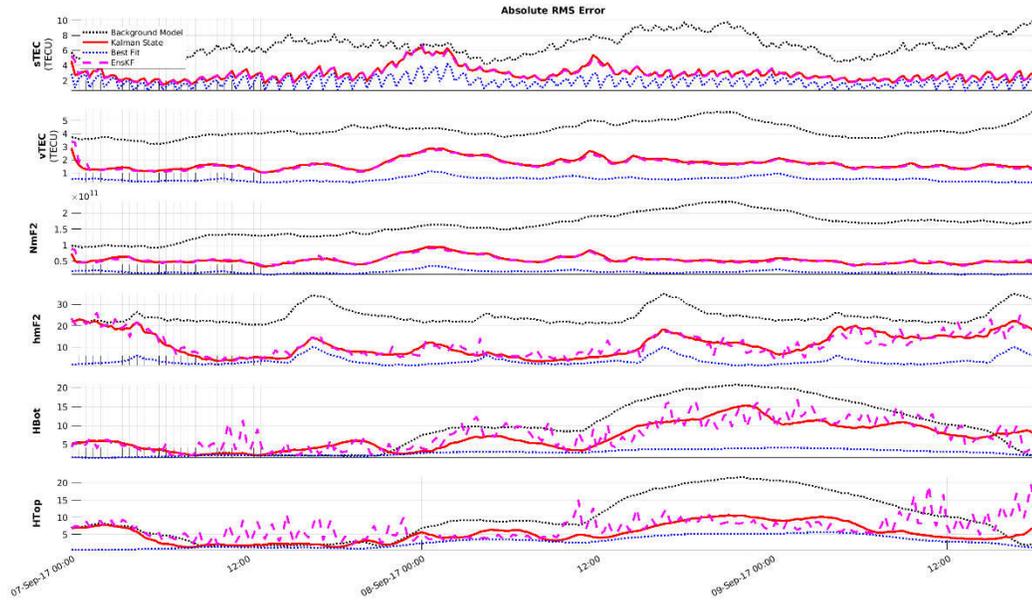
**Figure 8.** Small scale variability superimposed on the background E-CHAIM state for each A-CHAIM assimilation parameter.

First, to test the stability of the method against the possible situation of the background E-CHAIM model performing poorly, we use a quiet-time day of unperturbed E-CHAIM data as the background for the data assimilation simulations. We then generate synthetic data from the simulated truth state. This synthetic data is then perturbed using a random sampling of measurement errors. In the case of ionosonde data, the distribution of these errors is generated using comparisons between ARTIST-autoscaled data and corresponding manually scaled data in order to get a realistic measure of the errors inherent in the data the real assimilation system will face. For the GPS TEC data, random noise is added to reflect errors in the carrier phase observable. An example of the simulated sTEC at Eureka is presented in Figure 9.



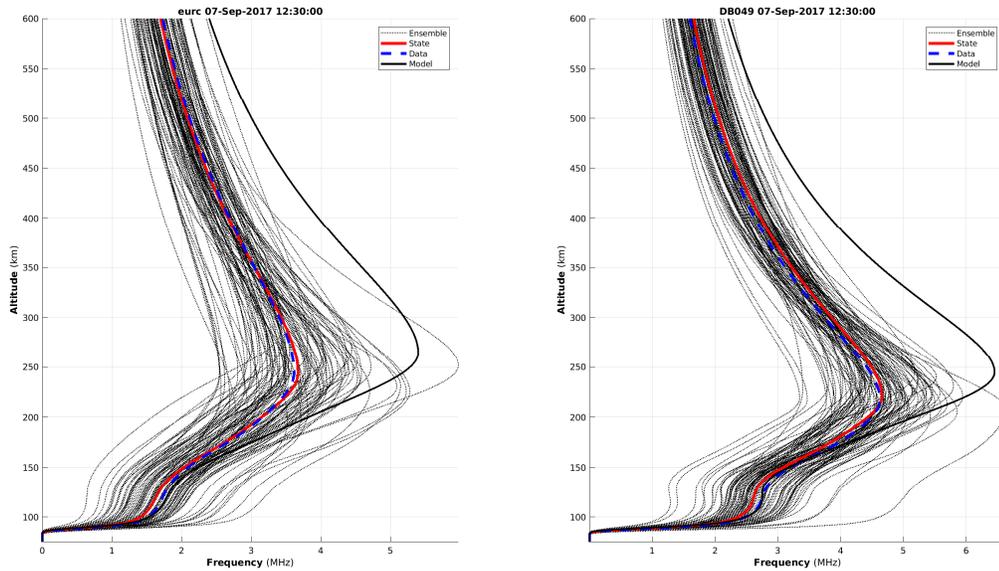
***Figure 9. Real measured sTEC (top) at Eureka on September 4<sup>th</sup>, 2017, with the corresponding simulated measurements in the middle plot and the E-CHAIM unperturbed sTEC in the bottom plot.***

With a perturbed known truth, assumed poor performance of the background model, and physically realistic measurement errors, we can now proceed with the assimilation simulations. For brevity, we will only present one of these simulations here. In this simulation, the truth is generated for the period between September 7<sup>th</sup> and 10<sup>th</sup>, 2017, corresponding to a severely disturbed ionospheric state associated with a geomagnetic super storm ( $K_p = 8+$ ). The background model used for this simulation is a quiet day from one year prior to the event to ensure that the background is sufficiently different from the truth state.



**Figure 10. RMS errors of the particle filter assimilation scheme (red), an EnKF scheme (pink), the background model (black), and the theoretical best fit state (blue) for the period between September 7<sup>th</sup> and September 10<sup>th</sup>, 2017. The top plot corresponds to the performance in representing the measured  $sTEC$ , the second plot corresponds to the errors in representing  $vTEC$ , with respect to the truth state, the third plot is that for  $NmF2$  (in  $e/m^3$ ), the fourth is for  $hmF2$  (in km), the fifth is for  $H_{Bot}$  (in km), and the bottom plot is the corresponding RMS errors in  $H_{Top}$  (in km).**

In Figure 10, we present the time series of the absolute RMS errors for  $NmF2$ ,  $hmF2$ ,  $H_{Top}$ , and  $H_{Bot}$  for the particle filter assimilation system (red), the background model (black), the best possible assimilation result given our chosen basis set (blue), and an Ensemble Kalman Filter (EnKF) implementation (pink). In general, the system is clearly capable of significantly improving upon the background model and approaches the theoretical best fit on several occasions. Even when available ionosonde observations severely diminish during the geomagnetic storm, due to absorption, the assimilation system maintains its stability with only minor losses in  $hmF2$  performance. The choice to use the particle filter over other approaches is highlighted by the increased stability of the red curve (particle filter) with respect to the pink curve (EnKF) in the inversion of the bottomside and topside thicknesses. An illustration of the vertical electron density profile at Eureka (no data assimilated at this location) and Dourbes (foF2 and  $hmF2$  assimilated at this location) is presented in Figure 11.



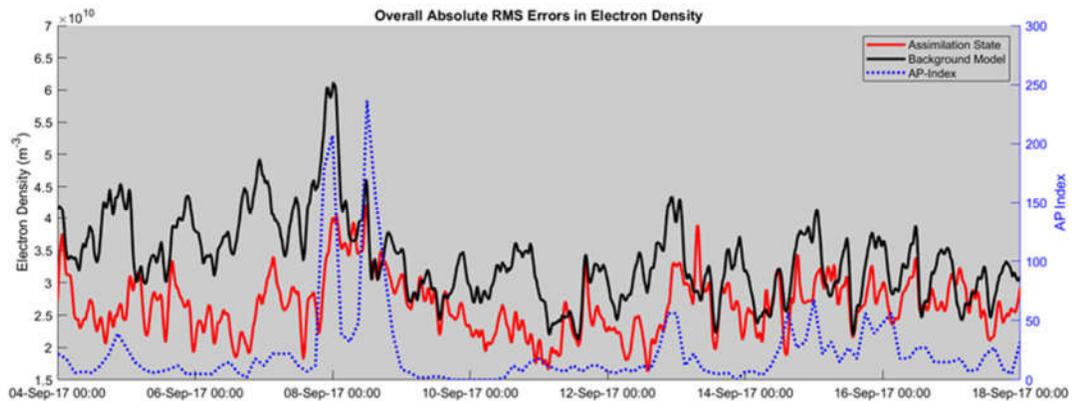
**Figure 11. Example ionospheric electron density profiles at Eureka (left) and Dourbes (right) during the disturbed periods leading up to the September 8<sup>th</sup> geomagnetic storm. The black dotted lines correspond to a random sample of the particle filter ensemble members, the solid black lines correspond to the background model profile, the blue dashed lines correspond to the true ionospheric state, and the red lines correspond to the particle filter inverted state.**

As you will note, contrary to many other ensemble ionospheric assimilation systems, which typically are only sensitive to NmF2 variability, the ensemble generated by the particle filter is clearly capable of spanning an array of hmF2's and layer thicknesses, ensuring that the system has the flexibility to fit these parameters, if sufficient data is available. These simulations prove to be promising, where we appear to be able to constrain components of the vertical electron density profile, not possible with other assimilation approaches, which notoriously have challenges capturing hmF2 and layer thicknesses. Based on these simulations we are confident that the particle filter approach is the optimal approach for the high latitude ionospheric assimilation problem.

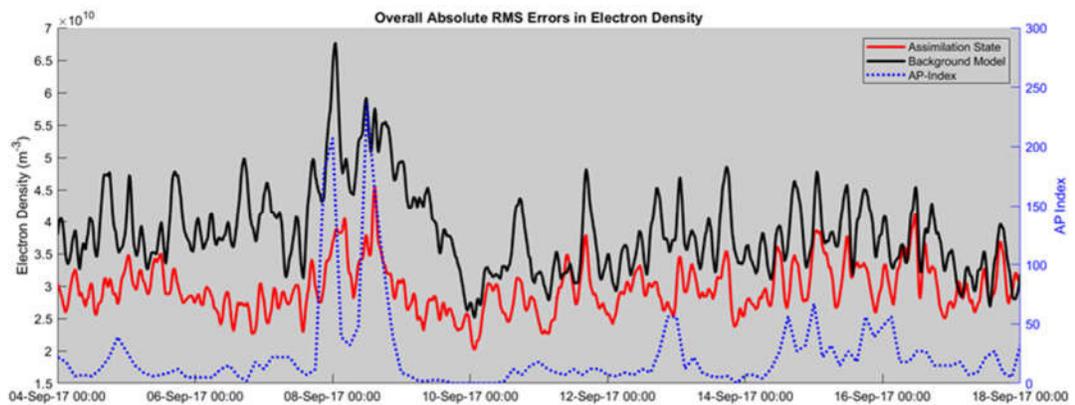
## 4.2 Further Simulations

To assess the stability and robustness of the current approach, we have conducted a further series of simulations using increasingly larger background-to-truth errors. To this end, we have conducted identical simulations using an E-CHAIM (Figure 12) and IRI (Figure 13) background from a year before the simulated truth dataset. Neither background here reflects a realistic background situation, but rather, they provide an assessment of the worse-case-scenario where 1) for the E-CHAIM background: the background model has some similar features to the truth but has large absolute errors, 2) for the IRI background: the background model is highly smoothed, not representing many of the features in the truth, and has large absolute errors. These simulations

take place over a week spanning a large geomagnetic storm (September 8<sup>th</sup>, 2017). The results of these simulation runs demonstrate that the system is stable regardless of the level of complexity in the background or of degree of discrepancy from the truth. Due to the IRI background being worse than that of E-CHAIM in this situation there is more improvement available to be gained through assimilating data. This results in a more striking improvement over the background than when using the E-CHAIM background, but also results in larger overall errors in the final assimilated product. While the E-CHAIM background performs better than the IRI, the results from both background choices are stable.

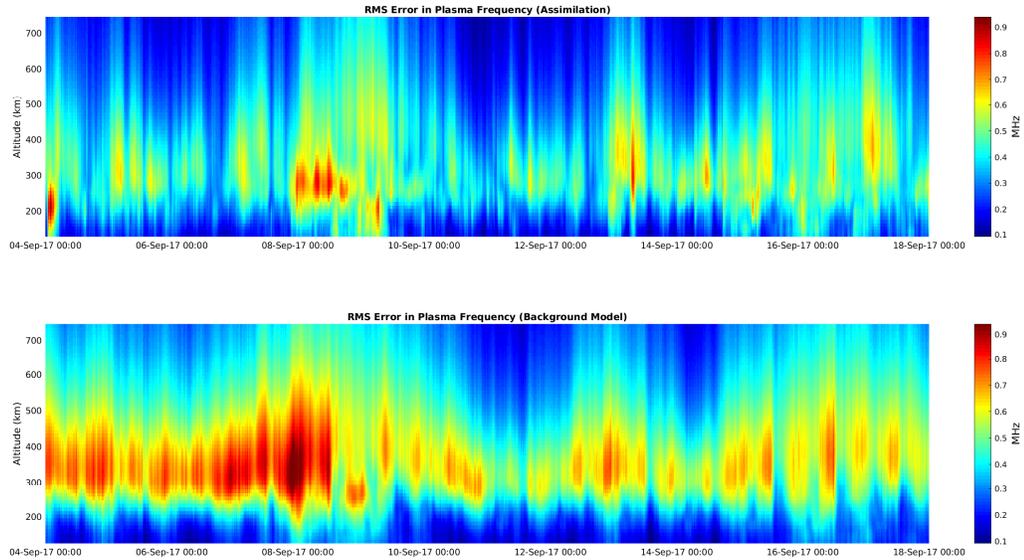


**Figure 12.** Overall RMS errors in the electron density at 10 altitudes spanning from 100km to 600km in 50km steps from the assimilation (red) and background (black) using E-CHAIM as the background model in the assimilation. The Ap index for the period is presented in blue.



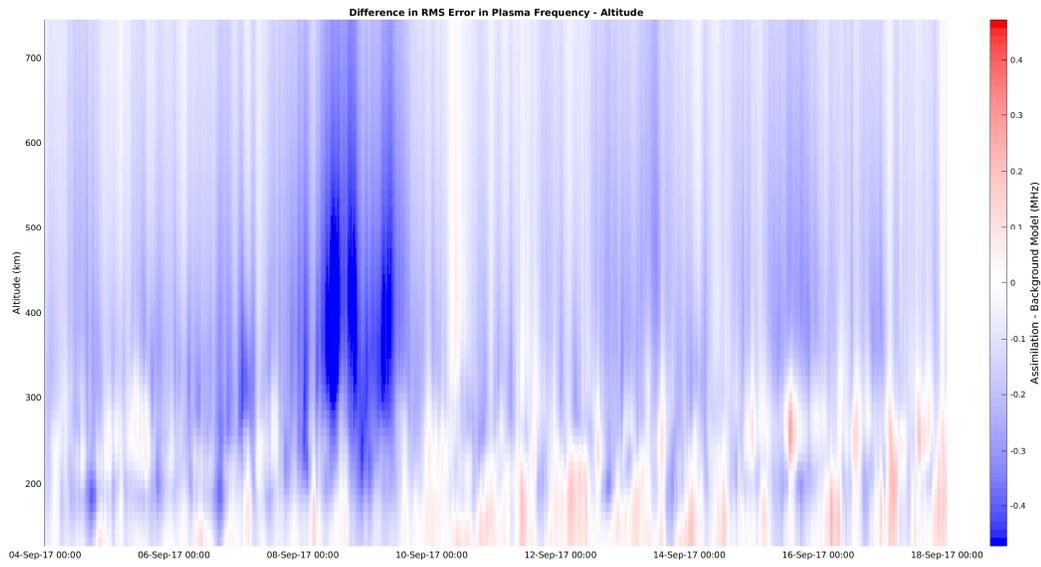
**Figure 13.** Overall RMS errors in the electron density at 10 altitudes spanning from 100km to 600km in 50km steps from the assimilation (red) and background (black) using the IRI as the background model in the assimilation. The Ap index for the period is presented in blue.

Vertical profiles of the errors in plasma frequency resulting from the assimilation are compared to those of the background model in Figure 14. Note the significant improvement in electron density produced by the assimilation in the F-Region.



**Figure 14. Vertical profiles of RMS errors in plasma frequency over the course of the simulation for the assimilation system (top) and the background model (bottom).**

Overall, performance is generally improved in all regions where there are measurements; however, performance degrades significantly in regions without data. For example, in terms of altitude, GNSS station baselines of less than 300km are necessary in order to constrain bottomside electron density. This geometry is only available in the United States and some regions of Europe, so overall errors with respect to altitude are only modestly improved with respect to the background in the lower bottomside globally. This is represented in Figure 15 where we present altitude-time contours of the relative performance of the assimilation with respect to the background in terms of plasma frequency.

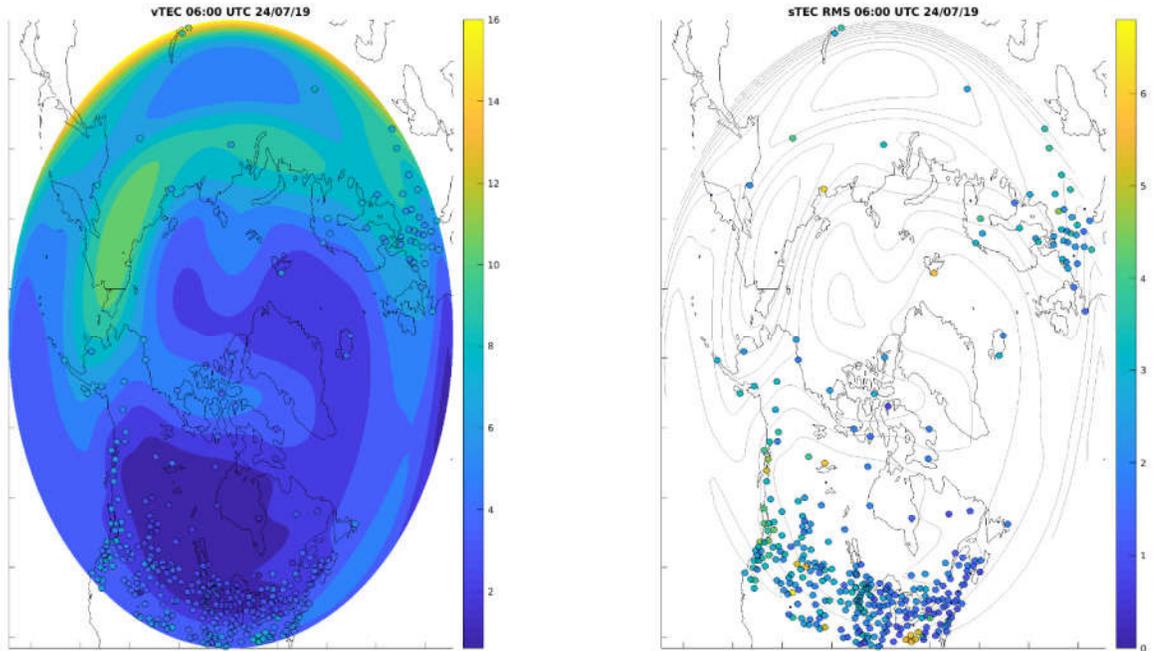


*Figure 15. Difference between background and assimilation errors in plasma frequency with respect to altitude (blue implies improvement by the assimilation).*

Horizontally, a similar challenge is seen over regions like the Pacific Ocean, which have no data available.

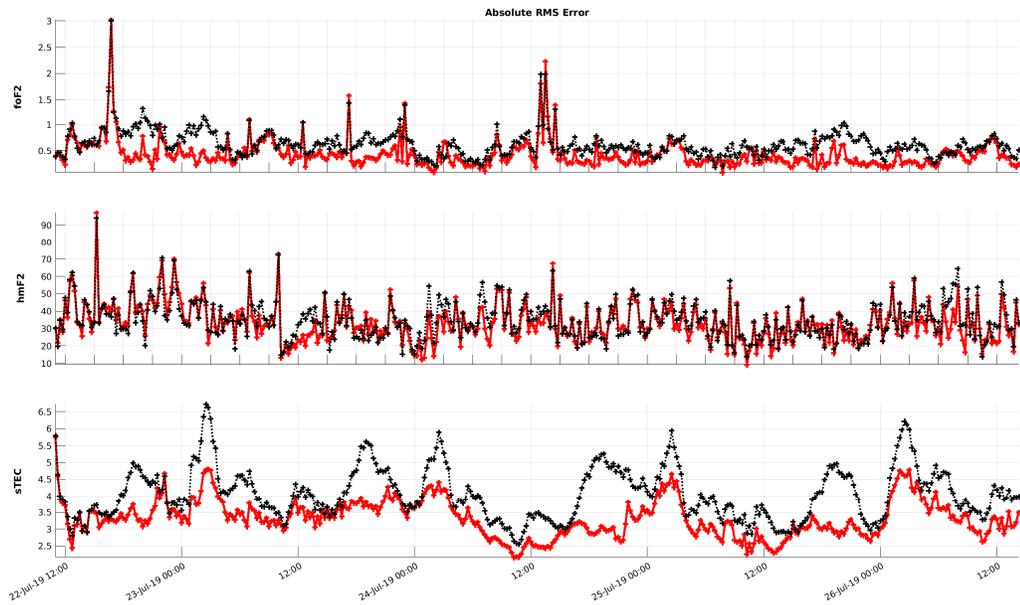
### 4.3 Preliminary Assessment with Real Data

Following the completion of the assimilation simulations, which assure us that the assimilation system is viable, a limited validation with real, near-real-time data was conducted as part of Phase IIb. Unfortunately, due to a formatting issue (now fixed), a significant portion of the GNSS and ionosonde data described in Section 2 were not included in the following assimilation trial. To illustrate the distribution of GNSS data that were included in the trial, we present an example of the GNSS station distribution superimposed on a map of assimilation-derived vertical TEC (vTEC) with corresponding station RMS errors in slant TEC (sTEC) in Figure 16.



**Figure 16.** *Left) An example vTEC contour map generated from the assimilation system with station measured vTEC indicated by the filled circles. Right) RMS errors in STEC for each station. The thin lines reflect the vTEC contours of the figure on the left.*

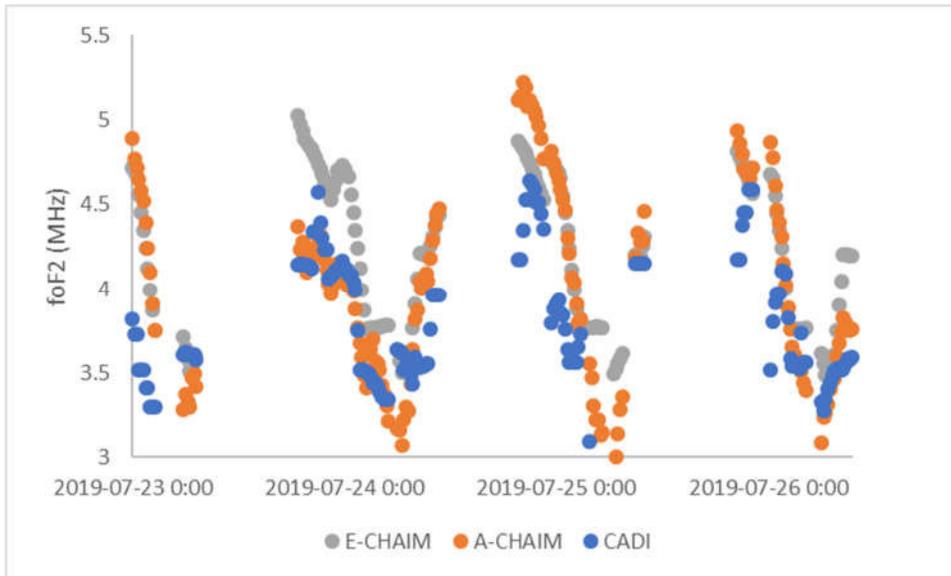
One will note that a number of stations were not included from Europe and the United Kingdom. Despite the missing data, fitting performance was acceptable for the real-data run. The performance of the assimilation system against the assimilated data is provided in Figure 17.



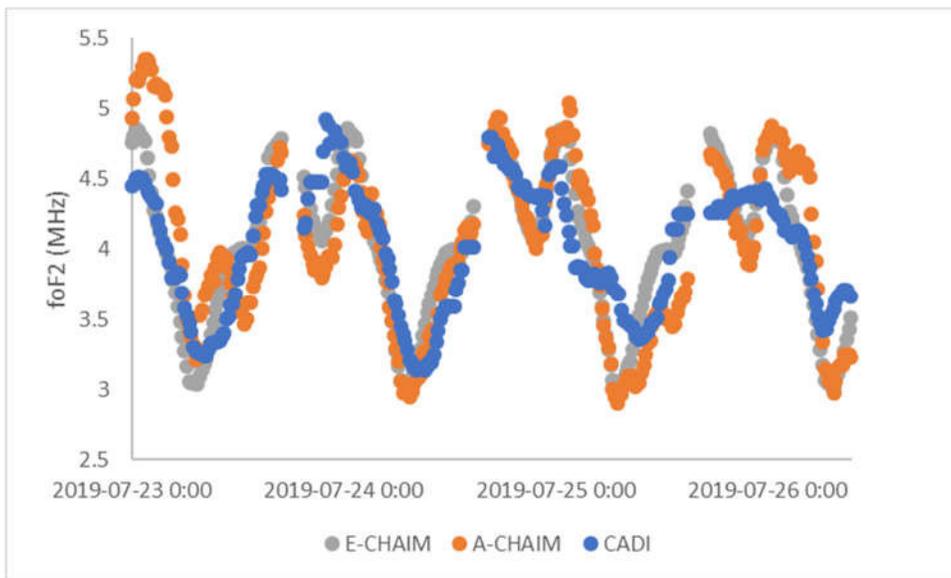
**Figure 17. RMS errors in foF2 (top), hmF2 (middle), and sTEC (bottom) for the background E-CHAIM model (black) and the assimilation system (red).**

Despite the lack of data, one will note from Figure 17 that systematic improvements in sTEC and foF2 were achieved, where foF2 errors were generally maintained below the 0.5 MHz level and sTEC errors were reduced by as much as 2 TEC units (TECU). Minor improvements were seen in hmF2; however, the full GNSS and ionosonde dataset would be necessary to fully assess the system's capacity for determining hmF2, as hmF2 is highly sensitive to station density. Limited hmF2 improvement also stems from the limited reliability of the auto-scaled hmF2 data, which already have measurement errors on the order of 25km. These errors imply that the ionosonde hmF2 data does not provide much value unless E-CHAIM is exhibiting hmF2 errors greater than 25km. This highlights the potential benefit if improved ionosonde scaling software were available. One should also note that the outliers in foF2 performance in Figure 17 reflect errors in ionosonde-derived foF2 autoscaling and not spontaneous excursions by the model.

A limited set of CHAIN ionosonde data was also processed to assess the performance of the above assimilation run in regions with little to no data assimilated in order to verify the stability of the system.



**Figure 18.** Example foF2 at Qikiqtarjuaq from CADI (blue), E-CHAIM (grey), and A-CHAIM (orange).



**Figure 19.** Same as Figure 18 but for Cambridge Bay.

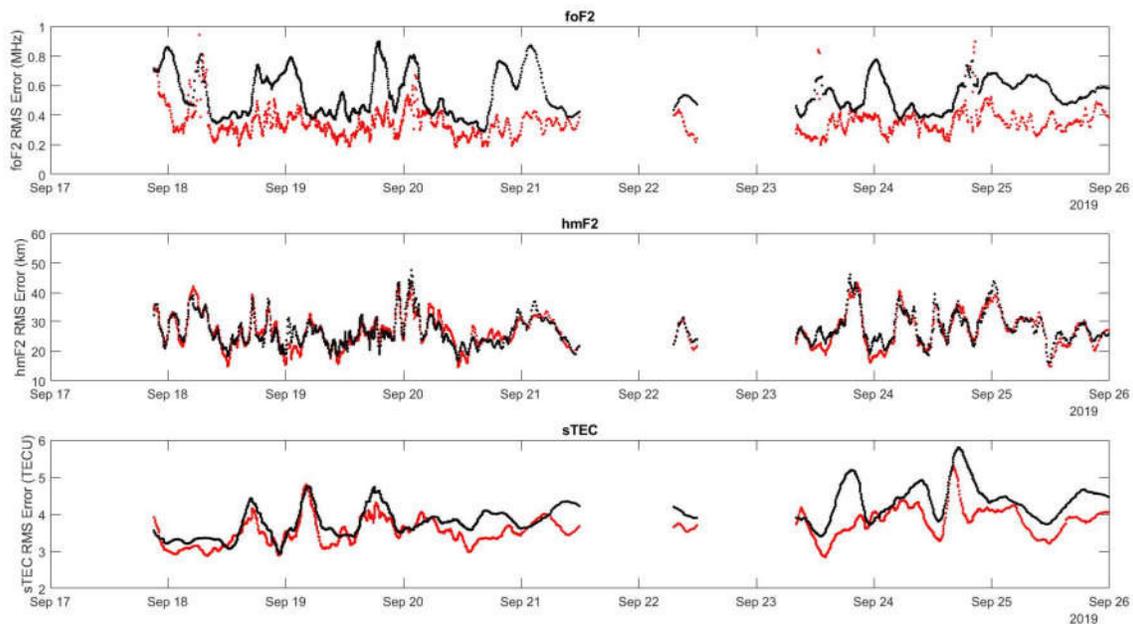
From the above figures, we note that, even in regions with a severe absence data, the system remains stable and converges to the background. When background errors are sufficiently large, such as during the daytime on August 23<sup>rd</sup> at Qikiqtarjuaq, the assimilation system rightly diverges from the background toward improvement. As it is currently summer, when E-CHAIM performs at its best, only minor errors are available for improvement by the assimilation system.

## 4.4 Further Testing with Real Data

As E-CHAIM errors become somewhat larger into the fall equinox, the performance of the assimilation system can be better assessed; as such, validations were conducted as we approach the October 1<sup>st</sup>, 2019, delivery deadline. As the system was continuously under refinement and development as we approached the October 1<sup>st</sup> deadline, we will here focus on a period between September 18<sup>th</sup> and September 28<sup>th</sup>, 2019, closest to the delivery date of the assimilation system. As part of this testing, we will explore two main questions: what is the overall performance of the assimilation system and what is the impact of assimilation delay on performance?

### 4.4.1 Overall Performance of fitting dataset

To assess the overall performance of the A-CHAIM assimilation system, we have restarted the assimilation system on September 18<sup>th</sup>, 2019, and run the system for almost ten days. To begin this assessment, we shall first present the performance of the model with respect to its fitting dataset. In Figure 20, we present A-CHAIM's overall RMS errors in foF2, hmF2, and TEC with respect to those of E-CHAIM for several days. For this figure, E-CHAIM and A-CHAIM foF2, hmF2, and slant TEC (sTEC) are directly compared to that measured by the various instruments contributing to the assimilation.

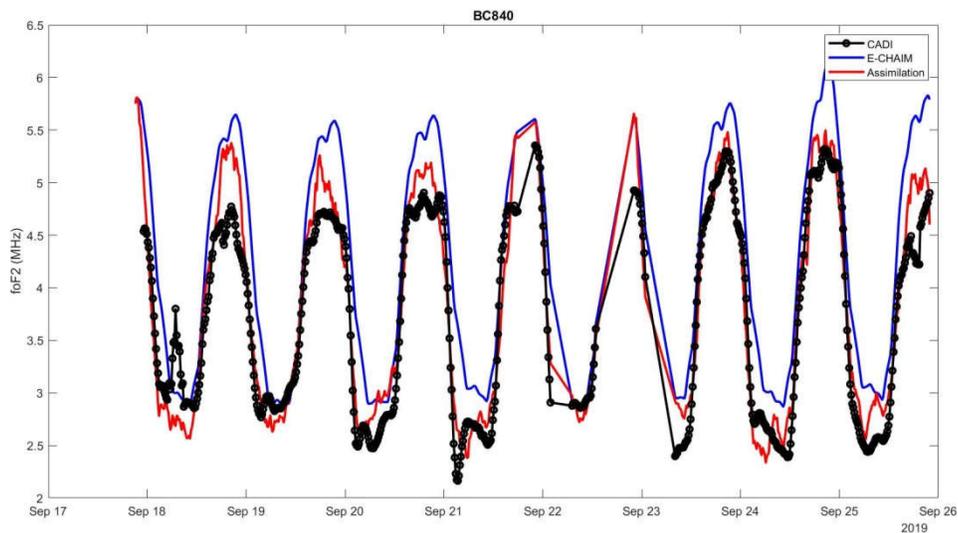


**Figure 20.** RMS errors in foF2 (top), hmF2 (middle), and sTEC (bottom) using E-CHAIM (black) and A-CHAIM (red), calculated against the A-CHAIM assimilation dataset.

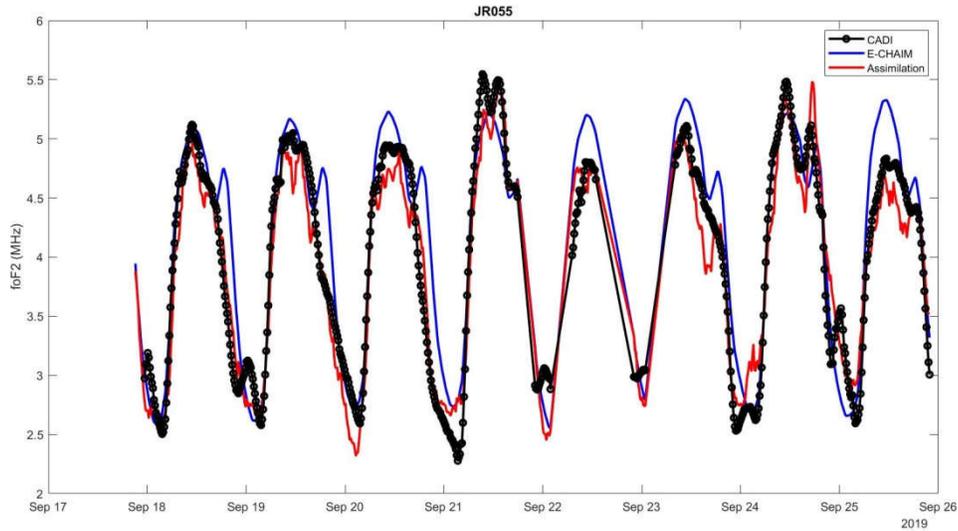
In this figure, one will note that foF2 is maintained at a performance level better than 0.5MHz for the majority of the assimilation, whereas E-CHAIM RMS errors exhibited an diurnal variation with maximum errors of approximately 0.9 MHz. One may also note that hmF2 performance is largely unchanged by the assimilation. In this specific assimilation test case, observational hmF2 errors have been inflated based on assumed ionosonde autoscaling error behaviour; however, this inflation is still in need of adjustment, as such errors should remain below 25km. This means that the assimilation should perform at the 25km accuracy level if the chosen hmF2 basis functions are sufficiently robust to capture the required spatial hmF2 variability. This component of the assimilation system is, thus, still being refined; however, given the lack of high latitude ionosondes contributing to the assimilation dataset and the limited constraint GNSS measurements provide with respect to hmF2, performance gains in hmF2 will likely be limited to sub-auroral latitudes.

Finally, one will note from Figure 20 that sTEC is systematically improved through assimilation, with errors being maintained largely at the 4 TECU level. Given the lack of a significant storm during this period and E-CHAIM's already-strong performance, the systematic improvement here is notable.

To better illustrate the behaviour of foF2 at the assimilated station locations, we present the foF2 timeseries for this period of the Boulder (BC840) and Juliusruh (JR055) ionosondes in Figure 21 and Figure 22, respectively. In these figures, we have plotted A-CHAIM (red), E-CHAIM (blue), and measured (black) foF2 for the entire run period.



**Figure 21. foF2 at Boulder from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period.**

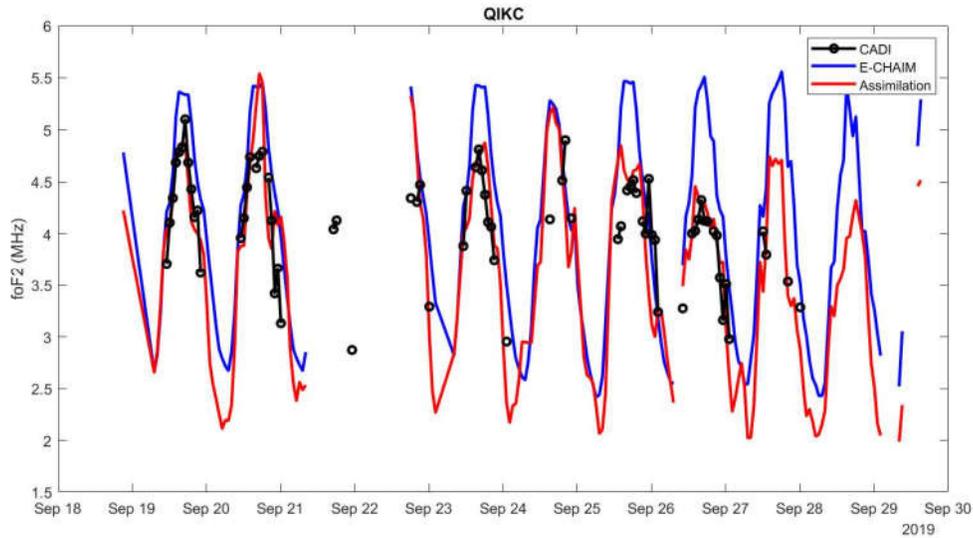


*Figure 22. foF2 at Juliusruh from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period.*

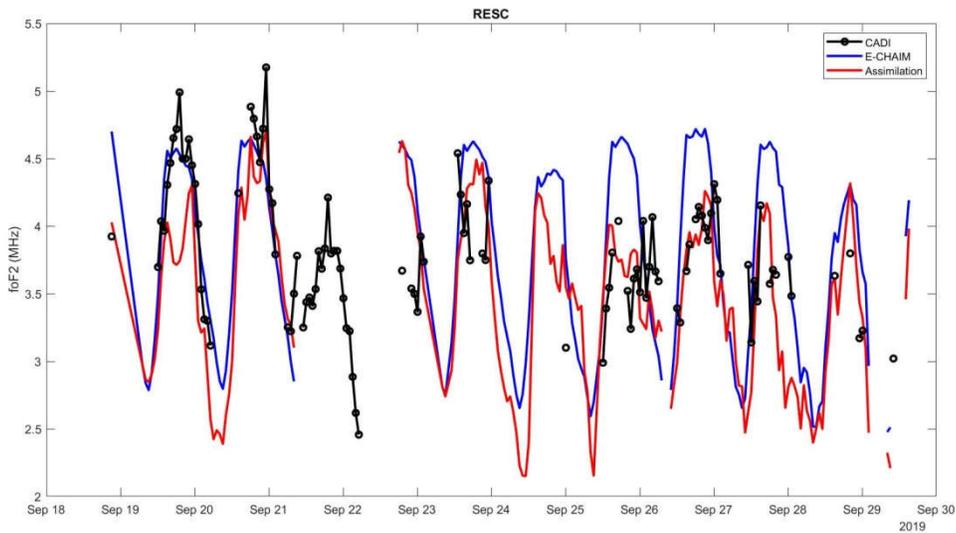
From Figure 21 and Figure 22, one will note that the assimilation system appears capable to representing the observed temporal variation in foF2 at these locations. The overall RMS error of A-CHAIM is 0.29 MHz at Boulder and 0.26 MHz at Juliusruh, while E-CHAIM performed with RMS errors of 0.63 MHz at Boulder and 0.38 MHz at Juliusruh.

#### 4.4.2 Performance against CHAIN assets

As described in Section 2.2 processed CHAIN ionosonde data is not currently available for NRT assimilation due to a lack of automated scaling software; as such, the CHAIN stations were not assimilated in these tests and serve as a good test of the performance of the assimilation system in regions devoid of contributing data. For this purpose, we present two examples of assimilation performance over the full period at Qikiqtarjuaq and Resolute in Figure 23 and Figure 24, respectively. Overall errors for each of these runs are respectively 0.44 MHz and 0.46 MHz for the assimilation, while the E-CHAIM background performed with overall RMS errors of 0.69 MHz and 0.55 MHz, respectively. Both locations exhibit an improvement with respect to the background model.



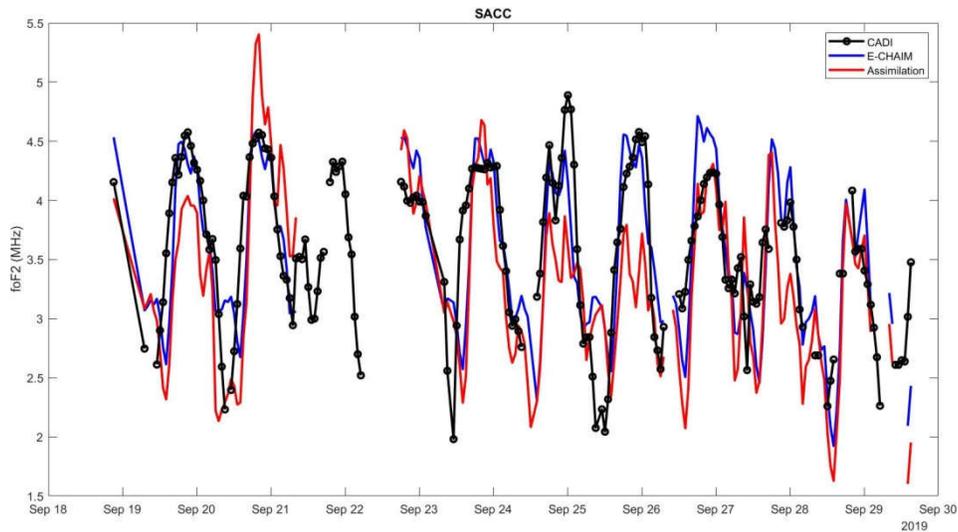
**Figure 23.** *foF2 at Qikiqtarjuaq from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period.*



**Figure 24.** *foF2 at Resolute from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period.*

As one tends closer to the pole, assimilated data becomes highly scarce; thus, the assimilation performs at a similar level to the background E-CHAIM model in these regions. Furthermore, due to the scarcity of data in these regions, assimilation resolution is limited; as such, features at smaller spatial scales cannot be resolved. To illustrate this point, we also present results from the

Sachs Harbour ionosonde in Figure 25. The overall foF2 performance at Sachs Harbour during this period was 0.55 MHz for the assimilation system and 0.46 MHz for the background model.



**Figure 25. foF2 at Sachs Harbour from ionosonde (black), E-CHAIM (blue), and A-CHAIM (red) for the sample test period.**

The largest source of error at Sachs Harbour appears to result from the underestimation of foF2 between September 25<sup>th</sup> and 26<sup>th</sup>. This underestimation is a direct result of the limited resolution available in this region with respect to significant spatial features in electron density, as foF2 at both Resolute and Qikiqtarjuaq saw significant decreases during this period, not observed at Sachs Harbour. The assimilation’s attempt to correct the foF2 for the Eastern Polar Cap region has spilled over into the Western Polar Cap, where no such feature is observed. The poorer performance of the assimilation on the 19<sup>th</sup> can largely be attributed to the assimilation spin up, as the assimilation system was restarted without a prior state on the 18<sup>th</sup>; furthermore, additional ionosonde data from Russia was acquired and incorporated into the assimilation system on the 26<sup>th</sup>.

#### 4.4.3 Preliminary performance against assimilation delay

In this section we briefly test the performance of the assimilation system against how near the assimilation was conducted with respect to the time the run was started (e.g. the blue dotted line in Figure 5). This test is to assess the impact of progressive data loss as one approaches the run start time. To this end, we present overall RMS errors at the Qikiqtarjuaq and Resolute test ionosondes and the Boulder and Juliusruh assimilated ionosondes. location against the delay with respect to the current time in Table 3.

**Table 3. Overall foF2 RMS errors for the period for the September test period using E-CHAIM and A-CHAIM with various assimilation delay times.**

	RESC	QIKC	BC840	JR055
1-hour Delay	0.4739	0.3722	0.2814	0.2683
2-hour Delay	0.4738	0.4429	0.285	0.2658
3-hour Delay	0.4634	0.445	0.2917	0.2578
E-CHAIM	0.5453	0.6925	0.6268	0.3772

Based on these preliminary tests, it becomes apparent that delay, and thereby the progressive loss of GNSS data as one approaches the current time, has little effect on the assimilation outcome at virtually all sites. In fact, the largest differences are observed at Qikiqtarjuaq, where we see RMS errors actually decrease as one approaches the current time. This suggests that the increased abundance of GNSS data may negatively affect the assimilation in that region. This result would suggest that the inclusion of additional GNSS data is reducing the weight of the ionosonde-derived foF2 observations, which appear to be more useful in constraining foF2 at Qikiqtarjuaq, or the errors in the assimilated sTEC are large enough to contaminate the resulting assimilation. Further investigation of this effect, over a longer test dataset, is necessary in order to fully assess the impact of delay on assimilation outcomes. It remains to be seen whether performance can be maintained with sub-hourly delays, e.g. in the absence of GNSS data; however, this should be tested.

## 5 Conclusions

---

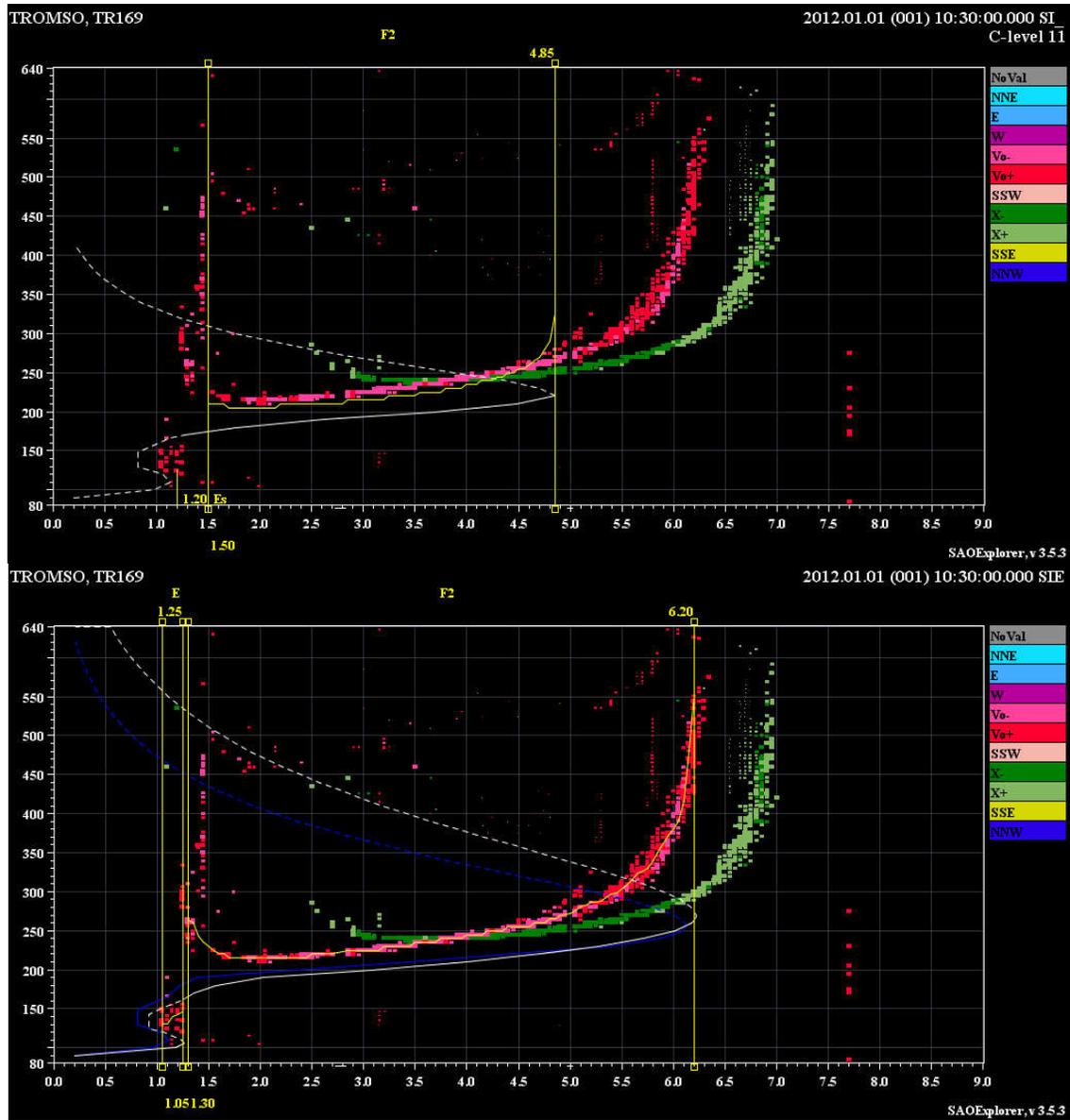
We have here developed robust near-real time data acquisition and processing algorithms. In parallel, we have developed an advanced particle filter assimilation system. This system is a unique approach to the ionospheric data assimilation problem, opting to use non-linear basis sets in both the vertical and horizontal dimensions. This approach was necessitated by the severe paucity of ionospheric data at high latitudes, particularly when considering a near-real time application. We have here tested the assimilation system in realistic simulations of disturbed ionospheric environments. Through these simulations we have refined the assimilation system to be sufficiently assured that it can globally achieve improved performance over the background without becoming unstable or converging too heavily to the background model. Performance evaluations suggest that the system is capable of improving upon E-CHAIM ionospheric TEC and foF2, particularly at sub-auroral latitudes, where data availability is greatest. A-CHAIM has been here demonstrated to outperform E-CHAIM in terms of foF2 in the Canadian Arctic, despite the extreme scarcity of data in this region.

## **6 What to look forward to in Phase III: R-CHAIM**

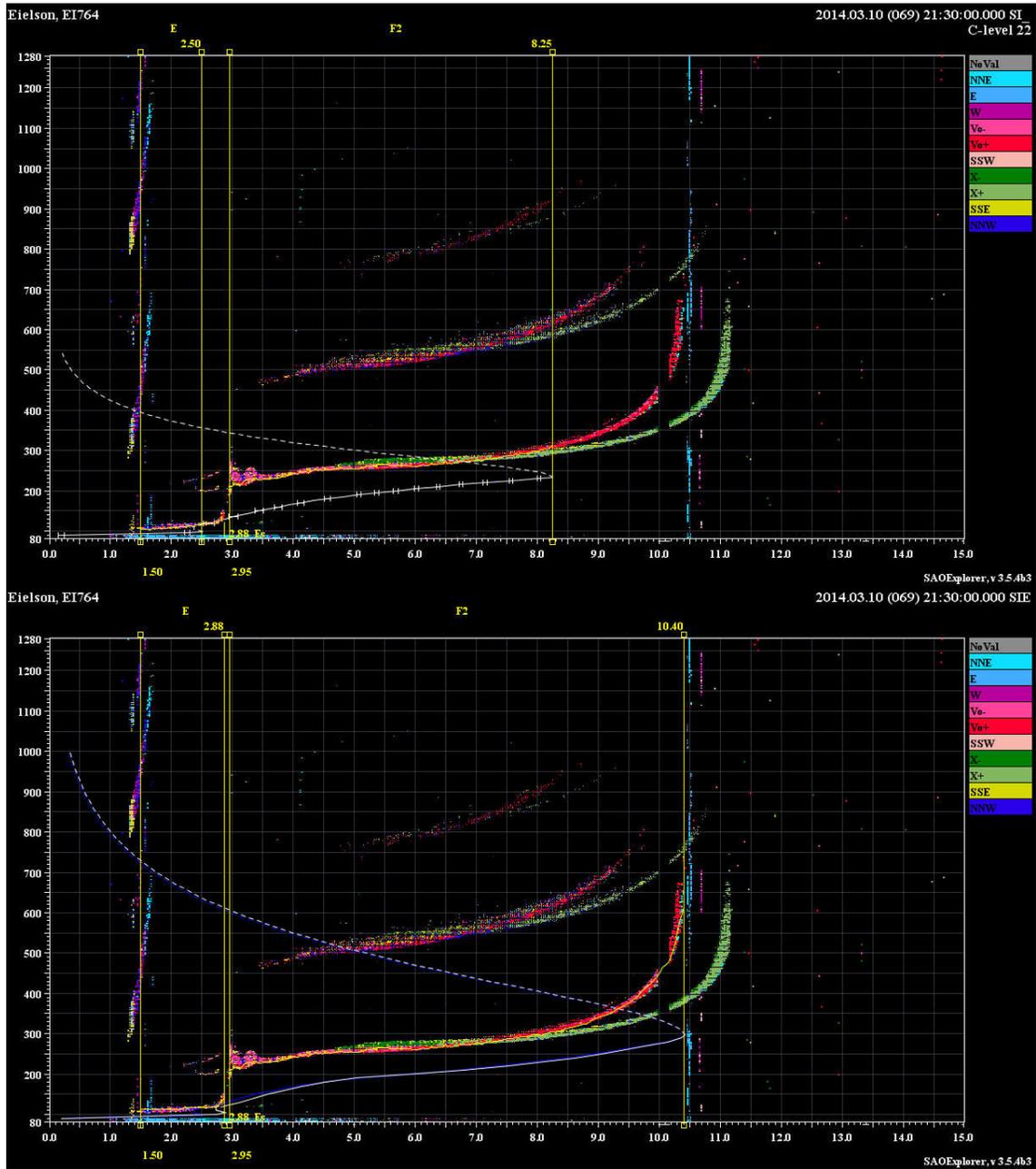
---

Phase 3 will see the development of a reanalysis assimilation system based on A-CHAIM. This system will incorporate other datasets, not available in near-real-time. This reanalysis system will provide electron density at performance levels beyond that currently available with E-CHAIM or A-CHAIM, but only for retrospective time periods (one month latency or greater). The development of this reanalysis will provide DRDC with the future capacity to include other datasets within the A-CHAIM assimilation system, should datasets that are not currently available in near-real-time become available at those latencies in the future. This system will also allow DRDC to test the relative benefit of various datasets for data assimilation and assist DRDC in identifying what ionospheric datasets should be prioritized for further development in the future.

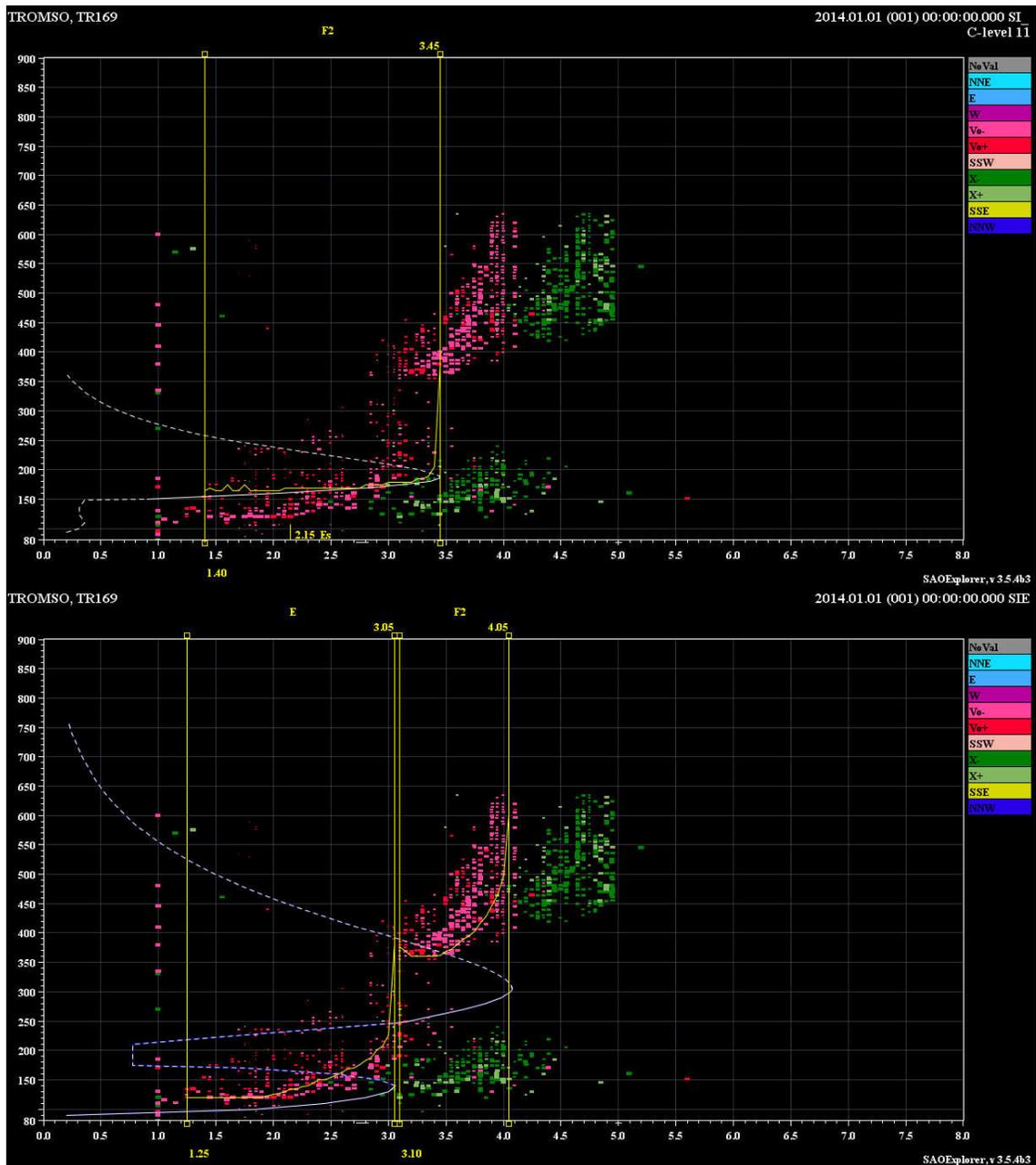
## 7 Appendix



*Figure 26. Top: Ionogram using the ARTIST autoscaling program, demonstrating an early-stopping error with 100% quality score. Bottom: Ionogram with corrected manual scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile.*



*Figure 27. Top: Ionogram using the ARTIST autoscaling program, demonstrating an early-stopping error with 80% quality score and multi-hop traces. Bottom: Ionogram with corrected manual scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile.*



*Figure 28. Top: Ionogram using the ARTIST autoscaling program, erroneously scaling the E-region as the F-region with 100% quality score. Bottom: Ionogram with corrected manual scaling. The scaled trace is in yellow. O-mode echoes are in pink-red. X-mode echoes are in green. White is the corresponding inverted profile. Note: Even with the correct manual scaling, the inverted profile is unphysical as the inversion software's choice of valley is exaggerated. In the case of the manually scaled profile, only the F2-peak parameters would be used by the assimilation system.*

## References/Bibliography

---

Carrano, C, and K. Groves (2009). Ionospheric Data Processing and Analysis. Workshop on Satellite Navigation Science and Technology for Africa. The Abdus Salam ICTP, Trieste, Italy.

Komjathy, A., (1997). Global Ionospheric Total Electron Content Mapping Using the Global Positioning System. Ph.D. dissertation, Department of Geodesy and Geomatics Engineering Technical Report No. 188, University of New Brunswick, Fredericton, New Brunswick, Canada

Shim, J. S., et al. (2011), CEDAR Electrodynamics Thermosphere Ionosphere (ETI) Challenge for systematic assessment of ionosphere/thermosphere models: NmF2, hmF2, and vertical drift using ground-based observations, *Space Weather*, 9, S12003, doi:10.1029/2011SW000727.

Shim, J. S., Tsagouri, I., Goncharenko, L., Rastaetter, L., Kuznetsova, M., Bilitza, D., et al. (2018). Validation of ionospheric specifications during geomagnetic storms: TEC and foF2 during the 2013 March storm event. *Space Weather*, 16, 1686–1701.  
<https://doi.org/10.1029/2018SW002034>

Themens, D.R., P.T. Jayachandran, I. Galkin, and C. Hall (2017). The Empirical Canadian High Arctic Ionospheric Model (E-CHAIM): NmF2 and hmF2, *J. Geophys. Res. Space Physics*, doi: 10.1002/2017JA024398

Themens, D.R., et al. (2018a). Topside Electron Density Representations for Middle and High Latitudes: A Topside Parameterization for E-CHAIM based on the NeQuick, *J. Geophys. Res. Space Physics*, 123, doi: 10.1002/2017JA024817

Themens, D.R., P.T. Jayachandran, A.M. McCaffrey, and B. Reid (2018b), The Empirical Canadian High Arctic Ionospheric Model (E-CHAIM): Validation and Release. Report submitted to Defence Research and Development Canada, Report Number: DRDC-RDDC-2018-C185 ([http://cradpdf.drdc-rddc.gc.ca/PDFS/unc327/p808067\\_A1b.pdf](http://cradpdf.drdc-rddc.gc.ca/PDFS/unc327/p808067_A1b.pdf))

Themens, D.R., P.T. Jayachandran, A.M. McCaffrey, B. Reid, and R.H. Varney (2019). A bottomside parameterization for the Empirical Canadian High Arctic Ionospheric Model (E-CHAIM), *Radio Sci.*, doi: 10.1029/2018RS006748

This page intentionally left blank.

## List of symbols/abbreviations/acronyms/initialisms

---

DND	Department of National Defence
A-CHAIM	Assimilation Canadian High Arctic Ionospheric Model
E-CHAIM	Empirical Canadian High Arctic Ionospheric Model
IRI	International Reference Ionosphere
NmF2	Peak density of the F2-layer
hmF2	Height of the F2-layer peak
foF2	The F2-layer critical frequency
M3000F2	Ratio of the MUF to foF2
MUF(3000)	Maximum Usable Frequency at 3000km distance
CHAIN	Canadian High Arctic Ionospheric Network
IRTAM	IRI-based Real-Time Assimilative Model
GIRO	Global Ionospheric Radio Observatory
PHaRLAP	Australian Defence Science and Technology Organisation Raytracing Toolkit
DMSP	Defense Meteorological Satellite Program
MIT	Main Ionospheric Trough
DRDC	Defence Research and Development Canada
DSTKIM	Director Science and Technology Knowledge and Information Management

**DOCUMENT CONTROL DATA**

\*Security markings for the title, authors, abstract and keywords must be entered when the document is sensitive

1. ORIGINATOR (Name and address of the organization preparing the document. A DRDC Centre sponsoring a contractor's report, or tasking agency, is entered in Section 8.)  University of New Brunswick 3 Bailey Drive P.O. Box 4400 Fredericton, New Brunswick Canada E3B 5A3		2a. SECURITY MARKING (Overall security marking of the document including special supplemental markings if applicable.)  CAN UNCLASSIFIED
		2b. CONTROLLED GOODS  NON-CONTROLLED GOODS DMC A
3. TITLE (The document title and sub-title as indicated on the title page.)  The Assimilation Canadian High Arctic Ionospheric Model (A-CHAIM): Implementation, Workflow, and Performance		
4. AUTHORS (Last name, followed by initials – ranks, titles, etc., not to be used)  Themens, D. R.; Reid, B.; McCaffrey, A.M.; Jayachandran, P.		
5. DATE OF PUBLICATION (Month and year of publication of document.)  March 2020	6a. NO. OF PAGES (Total pages, including Annexes, excluding DCD, covering and verso pages.)  50	6b. NO. OF REFS (Total references cited.)  8
7. DOCUMENT CATEGORY (e.g., Scientific Report, Contract Report, Scientific Letter.)  Contract Report		
8. SPONSORING CENTRE (The name and address of the department project office or laboratory sponsoring the research and development.)  DRDC – Ottawa Research Centre Defence Research and Development Canada, Shirley's Bay 3701 Carling Avenue Ottawa, Ontario K1A 0Z4 Canada		
9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant number under which the document was written. Please specify whether project or grant.)  99aa	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)  W7714-186507/001/SS	
10a. DRDC PUBLICATION NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document.)  DRDC-RDDC-2020-C045	10b. OTHER DOCUMENT NO(s). (Any other numbers which may be assigned this document either by the originator or by the sponsor.)	
11a. FUTURE DISTRIBUTION WITHIN CANADA (Approval for further dissemination of the document. Security classification must also be considered.)  Public release		
11b. FUTURE DISTRIBUTION OUTSIDE CANADA (Approval for further dissemination of the document. Security classification must also be considered.)		

12. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Use semi-colon as a delimiter.)

OTHR (Over the Horizon Radar); E-CHAIM, IRI, HF; A-CHAIM; Ionosphere; International Reference Ionosphere (IRI); Canadian High Arctic Ionospheric Model

13. ABSTRACT/RÉSUMÉ (When available in the document, the French version of the abstract must be included here.)

The high latitude ionosphere represents a challenging environment for space weather modeling. This region of the ionosphere is very dynamic and undersampled by available observations; thus, models struggle adequately represent this region for operational purposes, such as for use with Over-The-Horizon-Radar (OTHR). It has been shown [e.g. Shim et al., 2011; 2018] that physics-based models of the ionosphere are not capable of matching the performance of empirical climatologic models, such as the International Reference Ionosphere (IRI) and Empirical Canadian High Arctic Ionospheric Model (E-CHAIM) [Themens et al., 2017; 2018a; 2019], regardless of the region tested; however, by design, climatological models cannot adequately capture the short-term variability of the ionosphere and thus are limited in their capacity to represent high latitude regions. While empirical storm adjustments to these models have been implemented to some success, the fact remains that, in order to satisfy the needs of operational users, data assimilation is necessary [Themens et al, 2018b]. To be able to provide the best possible understanding of the current state of the ionosphere, it is necessary to augment traditional empirical models with near-real-time data sources, primarily ionosondes and Global Navigation Satellite System (GNSS) receivers; however, these instruments are both sparsely and unevenly distributed at high latitudes, which creates challenges for traditional assimilation approaches. In this study, we present the first data assimilation system specifically designed to mitigate these limitations for implementation at high latitudes.

Following the development of the Empirical Canadian High Arctic Ionospheric Model (E-CHAIM), we have been tasked with the development of a near-real-time (NRT) data assimilation system based on E-CHAIM. In the following work, we present the methodology, workflow, data composition, and validation of the Assimilation CHAIM (A-CHAIM). This assimilation system follows a sequential maximum likelihood methodology, implemented using a particle filter (e.g. a Monte Carlo-based system), to update E-CHAIM characteristics with NRT measurements from ground-based Global Navigation Satellite System (GNSS) receivers, ionosondes, satellite-borne altimeters, and Incoherent Scatter Radars (ISR).

L'ionosphère supérieure constitue un environnement complexe pour la modélisation de la météorologie spatiale. Cette région de l'ionosphère en est une très dynamique qui a fait l'objet de très peu d'observations; ainsi, les modèles ont de la difficulté à représenter adéquatement cette région à des fins opérationnelles, notamment pour une utilisation avec le radar transhorizon (OTHR). Il a été démontré [p. ex. par Shim et coll., 2011; 2018] que les modèles de l'ionosphère fondés sur la physique ont un rendement inférieur à celui des modèles climatologiques empiriques, comme le Modèle de l'ionosphère internationale de référence (IRI) et le Modèle empirique ionosphérique de l'Extrême-Arctique canadien (E-CHAIM) [Themens et coll., 2017; 2018a; 2019], sans égard à la région échantillonnée. Cependant, de par leur conception, les modèles climatologiques ne peuvent pas détecter convenablement la variabilité à court terme de l'ionosphère, ce qui explique leur capacité limitée à représenter les régions situées en hautes latitudes. Bien que des ajustements empiriques relatifs aux tempêtes ont été mis en œuvre avec un certain succès dans ces modèles, il est nécessaire d'assimiler les données afin de satisfaire les besoins des utilisateurs opérationnels [Themens et coll., 2018b]. Afin de comprendre le mieux possible l'état actuel de l'ionosphère, il faut bonifier les modèles empiriques traditionnels avec des sources de données en temps quasi réel, principalement des ionosondes et des récepteurs du Système mondial de navigation par satellite (GNSS). Toutefois, ces instruments sont répartis de façon éparsée et inégale dans les hautes latitudes, ce qui crée des défis pour les méthodes traditionnelles d'assimilation. Dans la présente étude, nous présentons le premier système d'assimilation de données spécifiquement conçu pour atténuer ces contraintes d'assimilation dans les hautes latitudes.

À la suite de l'élaboration du Modèle empirique ionosphérique de l'Extrême-Arctique canadien (E-CHAIM), nous avons eu pour mandat d'élaborer un système d'assimilation de données en temps quasi réel (TQR) fondé sur le modèle E-CHAIM. Dans la présente étude, nous présentons la méthode, le flux de travail, la composition des données et la validation du modèle CHAIM d'assimilation (A-CHAIM). Ce système d'assimilation repose sur une méthode d'estimation séquentielle de vraisemblance maximale, laquelle est mise en œuvre au moyen d'un filtre à particules (p. ex. un système de type Monte Carlo), afin de mettre à jour les caractéristiques du modèle E-CHAIM avec des mesures en TQR provenant de récepteurs GNSS au sol, d'ionosondes, d'altimètres embarqués sur des satellites et de radars à diffusion incohérente.