

GLOBAL ESTIMATION AND ANALYSIS OF IONOSPHERIC DRIVERS  
WITH A DATA ASSIMILATION ALGORITHM

BY  
AURORA LÓPEZ RUBIO

Submitted in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy in Mechanical and Aerospace Engineering  
in the Graduate College of the  
Illinois Institute of Technology

Approved *Subhrajit Datta*  
Adviser

Chicago, Illinois  
May 2022



## ACKNOWLEDGEMENT

I would first like to thank my Ph.D. adviser Dr. Seebany Datta-Barua for her guidance, help and expertise, invaluable in formulating the research questions and methodology. Thank you!

I would also like to thank my IIT Ph.D. committee members: Dr. Boris Pervan, Dr. Hassan Nagib and Dr. Shuwang Li for their time and their insightful questioning. Thank you to the Illinois Institute of Technology, Armour Collage of Engineering and Graduate College.

I would also like to acknowledge my past and present labmates from the IIT Space Weather Lab for their help and support thought this process.

I am very grateful to my family and friends. Thanks to my parents and the rest of my family, Santi, Nata, Javi, Elodie and Marina for their understanding and for always supporting me with their love. Finally, a special thanks to Alberto, for all his encouragement and for always being there for me.

This work was funded by the following awards: NSF AGS 1352602 and NASA 80NSSC19KOO86 and the IIT Graduate Research Assistantship.

## AUTHORSHIP STATEMENT

I, Aurora López Rubio, attest that the work in this thesis is substantially my own.

In accordance with the disciplinary norm of Mechanical and Aerospace Engineering Major (see IIT Faculty Handbook, Appendix S), the following collaborations occurred in the thesis:

Dr. Alex Chartier of Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA and Dr. Sarah McDonald of Naval Research Laboratory, Washington, DC, USA, contributed to the idea and the discussion of the second contribution presented in this dissertation. Outputs of SAMI3 model used in some sections of this thesis, were also provided by them.

Dr. Gary Bust of Johns Hopkins University Applied Physics Laboratory, Laurel, MD, USA collaborated in the idea and discussion of the direct estimation of the ionospheric drivers in the data assimilation algorithm EMPIRE. Outputs of IDA4D used in Chapters 3 and 6 were provided by Dr. Gary Bust as well.

Dr. Seebany Datta-Barua of Illinois Institute of Technology, Chicago, IL, USA provided guidance and collaborated in the idea and discussion of all the contributions presented in this dissertation.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENT . . . . .	iii
AUTHORSHIP STATEMENT . . . . .	iv
LIST OF TABLES . . . . .	vii
LIST OF FIGURES . . . . .	viii
LIST OF SYMBOLS . . . . .	xiv
ABSTRACT . . . . .	xxv
CHAPTER	
1. INTRODUCTION . . . . .	1
1.1. Contributions . . . . .	4
2. TECHNICAL BACKGROUND . . . . .	9
2.1. Overview . . . . .	9
2.2. Thermosphere . . . . .	10
2.3. Ionosphere . . . . .	14
2.4. Magnetic Field . . . . .	19
2.5. IT interaction . . . . .	23
2.6. Algorithms . . . . .	29
3. VECTOR SPHERICAL HARMONICS FOR EMPIRE NEUTRAL WIND ESTIMATION . . . . .	42
3.1. Vector spherical harmonics derivation . . . . .	42
3.2. Direct estimation of ionospheric drivers . . . . .	46
3.3. Application of VSH to study geomagnetic storm and compar- ison to previous method . . . . .	52
4. QUANTIFICATION OF REPRESENTATION ERROR IN EM- PIRE ESTIMATION . . . . .	70
4.1. Error discussion . . . . .	71
4.2. Method . . . . .	72
4.3. Results . . . . .	78

4.4. Conclusions . . . . .	91
5. INVESTIGATION OF NIGHTTIME IONOSPHERIC LOCALIZED ENHANCEMENT . . . . .	94
5.1. Ingestion of FPI using VSH . . . . .	96
5.2. Storm November 2003 . . . . .	97
5.3. Storm August 2018 . . . . .	116
5.4. Conclusions . . . . .	129
6. ADDITIONAL INGESTION OF ICON NEUTRAL WINDS MEASUREMENTS . . . . .	133
6.1. Ingestion of ICON neutral winds . . . . .	134
6.2. Method . . . . .	137
6.3. Results . . . . .	145
6.4. Conclusions . . . . .	153
7. CONCLUSIONS . . . . .	155
7.1. Summary and discussions . . . . .	155
7.2. Future work . . . . .	157
APPENDIX . . . . .	158
A. MAPPING MATRICES $F_{EXB}$ AND $H_{EXB}$ FOR ELECTRIC POTENTIAL SCALAR SPHERICAL HARMONIC EXPANSION . . . . .	159
B. MAPPING MATRIX $H_U$ FOR VECTOR SPHERICAL HARMONIC DERIVATION . . . . .	163
C. MAPPING MATRICES $F_{UN}$ AND $F_{UE}$ FOR VECTOR SPHERICAL HARMONIC DERIVATION . . . . .	172
D. MAPPING MATRIX $H_{FPI}$ VECTOR SPHERICAL HARMONIC DERIVATION . . . . .	176
BIBLIOGRAPHY . . . . .	178

## LIST OF TABLES

Table		Page
2.1	Inputs to EMPIRE, climate models and measurements, utilized in this dissertation organized by chapters. . . . .	41
3.1	RMS error in m/s between EMPIRE estimate for case VSH, case PS and model HWM14 and FPI measurements at Pisgah, Cariri and Nasca. . . . .	66
3.2	Normalized RMS error with FPI measurements between EMPIRE estimate for case VSH, case PS and model HWM14 and FPI measurements at Pisgah, Cariri and Nasca. . . . .	68
6.1	Variables downloaded from ICON [1]. . . . .	134
6.2	RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Arecibo. . . . .	152
6.3	RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Urbana. . . . .	152
6.4	RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Millstone. . . . .	153

## LIST OF FIGURES

Figure	Page
2.1 Layers of Earth’s Atmosphere. Adapted from [2]. . . . .	11
2.2 Typical midlatitude daytime and nighttime electron density profiles for solar maximum (solid lines) and minimum (dashed lines) [3]. . . . .	15
2.3 Single charged particle trajectory when embedded in a magnetic field according to the force equation. . . . .	17
2.4 $\vec{E}$ cross $\vec{B}$ drift of ions. . . . .	18
2.5 Schematic drawing of the Earth’s magnetic field with the geomagnetic north pole, spin axis and field lines labeled, based on [4]. . . . .	20
2.6 Schematic drawing of the Earth’s magnetic field based on a dipole model and its corresponding L-shell values. . . . .	21
2.7 Schematic drawing of the inclination angle with the local horizon, the magnetic coordinates ( $r$ , $\theta$ and $\phi$ ) and the unit vector parallel to the magnetic field $\hat{b}$ labeled. . . . .	21
2.8 Schematic drawing of the spherical coordinate system ( $\hat{r}$ , $\hat{\theta}$ , $\hat{\phi}$ ) and the magnetic field coordinate system ( $\hat{b}$ , $\hat{\perp}_N$ , $\hat{\perp}_E$ ). . . . .	22
2.9 Schematic drawing of the spherical coordinate system ( $\hat{r}$ , $\hat{\theta}$ , $\hat{\phi}$ ) and the geographic coordinate system ( $\hat{e}$ , $\hat{n}$ , $\hat{r}$ ). . . . .	23
2.10 Simplified EMPIRE model and data processing flow chart. . . . .	38
2.11 Simplified EMPIRE model and data processing flow chart. . . . .	41
3.1 RMS (Root-Mean-Square) difference in $m/s$ between estimated $u_N$ , in green, and $u_E$ , in magenta, and HWM14 model over order $l_{\max}$ over all grid points at 0 UT on 25 October 2011. Black dashed line represents the values at $l_{\max} = 6$ . . . . .	46
3.2 New EMPIRE model with direct estimation of IT drivers described with data processing flow chart. . . . .	51
3.3 (a) Dst index for October 2011 in nT. Studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line. (b) TEC maps over time. Terminator line is indicated with a black solid line and local noon with dashed red line. . . . .	53

3.4	Diagram of the EMPIRE configuration and of the calculation method of storm neutral winds for each of the studied cases. . . . .	56
3.5	Mean (top) and standard deviation (bottom) over all grid points of the error of the estimated neutral winds and the model HWM14 data over order $l_{\max}$ of the VSH expansion at 0 UT on October 25. Zonal component is shown in magenta and meridional component in green. Black dashed line represents the values at $l_{\max} = 6$ . . . .	57
3.6	Neutral wind estimation at 00:10 UT and 3 deg magnetic latitude over magnetic longitude. Top shows meridional direction and bottom zonal direction of neutral speed. Red represents case PS estimate and in blue case VSH is shown. Black dashed line indicates geomagnetic longitude boundary $180 = -180$ deg. . . . .	60
3.7	Neutral winds plotted over time at Pisgah location to the (a) south, (b) west, (c) north and (d) east. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value. . . . .	61
3.8	Neutral winds plotted over time at Cariri location to the (a) west-northwest and (b) north-northwest. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value. . . . .	63
3.9	Neutral winds plotted over time at Nasca location to the (a) west-northwest, (b) north-northeast and (c) north-west. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value. . . .	64
4.1	Ap index for August 2018 in nT. Studied time limits are indicated with black dashed lines and the beginning of geomagnetic storm time is indicated with green dashed-dot line. . . . .	74
4.2	Block Diagram of SAMI3 and EMPIRE. . . . .	75

4.3	Standard deviation in (a) velocity-space and (b) density rate-space over all grid points of the error between the estimated drivers and the SAMI3 data over order $l_{\max}$ at 0:10 UT on August 25, 2018. In (a) the neutral wind zonal component is shown in magenta, the neutral wind meridional component in green, the ion drift magnetic up in orange, magnetic south in purple and magnetic east in cyan. In (b) the transport term corresponding to neutral winds is shown in blue and to the ion drifts is shown in red. The black dashed line represents the values at $l_{\max} = 6$ . . . . .	79
4.4	Zonally averaged error between ion drifts EMPIRE estimation and SAMI3 values in the perpendicular (a) to meridional and to (b) zonal direction plotted over time. . . . .	82
4.5	Zonally averaged error between neutral winds EMPIRE estimation and SAMI3 (HWM14) values in the (a) geographical meridional and (b) zonal direction plotted over time. . . . .	84
4.6	(a) Mean and (b) standard deviation of residual between EMPIRE estimation and SAMI3 values averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude. . . . .	86
4.7	(a) Mean and (b) standard deviation of relative residual between EMPIRE estimation and SAMI3 values averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude. . . . .	89
4.8	Estimated variance in velocity space averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude. . . . .	90
5.1	Simplified EMPIRE model and data processing flow chart. . . . .	98
5.2	Dst index for November 2003 in nT. Studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line. . . . .	99

5.3	TEC maps over time for 20-21 November 2003. Terminator line is indicated with a black solid line and local noon with dashed line. Red star indicates location near magnetic equator, blue where NILE is observed and green Millstone Hill ISR location. . . . .	100
5.4	Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes. Red dashed line represents the mean value of the case. . . . .	104
5.5	Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes. Red dashed line represents the mean value of the case. . . . .	105
5.6	Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes in blue. Red dashed line represents the mean value of the case. . . .	106
5.7	Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes in blue. Red dashed line represents the mean value of the case. . . .	107
5.8	Ion drift geographical meridional and zonal direction at height 250 km estimated with EMPIRE represented in red arrows and TEC plotted in color. Terminator line is represented with a black dashed line. Each subfigure represents a different time: (a) 20/17, (b) 20/19, (c) 20/21, (d) 20/23, (e) 21/1 and (f) 21/3 (dd/hh) November 2003 UT. Red star indicates location near magnetic equator, blue where NILE is observed and green Millstone Hill ISR location. . . . .	109
5.9	Ion drift vertical component (m/s) over time. Solid line represents value estimated with EMPIRE and dashed line the climate model values. In red results at 72 deg W, 5 deg N are plotted (near magnetic equator) and in blue results at 72 deg W, 25 deg N location (over NILE). Black dashed line separates upward (positive) or downward (negative) velocity. . . . .	111
5.10	Ion drift vertical component (m/s) over time. Solid line represents value estimated with EMPIRE and dashed line the climate model values. In green results at 72 deg W, 55 deg N (ISR location) are plotted. Black dashed line separates upward (positive) or downward (negative) velocity. Black solid line is the ISR measurements at Millstone Hill location. . . . .	113

5.11	Neutral winds geographical meridional and zonal direction at height 250 km estimated with EMPIRE represented in magenta arrows and TEC plotted in color. Each subfigure represents a different time: 20/17, 20/19, 20/21, 20/23, 21/1 and 21/3 (dd/hh) November 2003 UT. . . . .	115
5.12	Dst index for August 2018 in nT. The studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line. . . . .	116
5.13	TEC maps over time for 25-26 August 2018. Terminator line is indicated with a black solid line. Magenta markers indicate FPI sites where neutral wind measurements were available: Arecibo (star), Millstone Hill (triangle) and Urbana (square). . . . .	118
5.14	Diagram of the EMPIRE configuration for each of the studied cases.	121
5.15	Ion drift geographical meridional and zonal direction at height 250km estimated with EMPIRE represented in red arrows and TEC plotted in color. Terminator line is represented with a black dashed line. Each subfigure represents a different time: 25/23, 26/00, 26/01, 26/02, 26/03 and 26/04 (dd/hh) August 2018 UT. . . . .	123
5.16	Neutral winds geographical meridional and zonal direction at height 250km estimated with EMPIRE represented in magenta arrows and TEC plotted in color. Each subfigure represents a different time: 25/23, 26/00, 26/01, 26/02, 26/03 and 26/04 (dd/hh) August 2018 UT. . . . .	124
5.17	Neutral winds plotted over time at Urbana location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value. . . . .	126
5.18	Neutral winds plotted over time at Millstone Hill location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value. . . . .	128

5.19	Neutral winds plotted over time at Arecibo location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value. . . . .	130
6.1	Simplified EMPIRE model and data processing flow chart. . . . .	137
6.2	Dst index for 1-5 January 2020 in nT. The studied time limits are indicated with black dashed line. . . . .	138
6.3	Coverage of ICON instrument for altitudes 200-400 km during January 4 2020 (a) 0UT-0:20UT, (b) 0:20UT-0:40UT, (c) 0:40UT-1UT, (d) 1UT-1:20UT, (e) 1:20UT-1:40UT and (f) 1:40UT-2UT. . . . .	140
6.4	Diagram of the EMPIRE configuration for each of the studied cases.	142
6.5	Neutral wind ICON measurements between 1:40 UT and 2:00 UT on January 4, 2020 at (a) points. Meridional component and error (b) in green and Zonal component and error (c) are observed over time. . . . .	144
6.6	Neutral winds estimation with ingestion in blue, with no ingestion in red and ICON measurements in green over a world map for January 4 2020 at times (a) 2:30 UT (b) 2:50 UT (c) 3:10 UT and (d) 3:30 UT . . . . .	146
6.7	Neutral winds plotted over time at Arecibo location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion. . .	147
6.8	Neutral winds plotted over time at Urbana location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion. . .	150
6.9	Neutral winds plotted over time at Millstone location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion.	151

## LIST OF SYMBOLS

Symbol	Definition
$\vec{B}$	Magnetic field
$\vec{B}_0$	Climate magnetic field
$B_0$	Norm of climate magnetic field
$D$	Declination angle
$\vec{E}$	Electric field
$\vec{F}$	Lorentz force
$\mathbf{F}_{\text{exb}}$	Ion drift mapping matrix
$\mathbf{F}_{\mathbf{u}}$	Neutral winds mapping matrix
$\mathbf{F}_{\mathbf{uE}}$	Neutral wind zonal component mapping matrix
$\mathbf{F}_{\mathbf{uN}}$	Neutral wind meridional component mapping matrix
$I$	Inclination angle
$\mathbf{I}$	Identity matrix
$\mathbf{H}$	Mapping matrix
$\bar{\mathbf{H}}$	Normalized mapping matrix
$\mathbf{H}_{\text{exb}}$	Ion drifts transport contribution mapping matrix
$\mathbf{H}_{\text{FPI}}$	FPI neutral winds ingestion mapping matrix
$\mathbf{H}_{\text{ICON}}$	ICON neutral winds ingestion mapping matrix
$\mathbf{H}_{\mathbf{N}}$	Ion continuity equation mapping matrix
$\mathbf{H}_{\mathbf{u}}$	Neutral winds transport contribution mapping matrix
$L$	L-shell
$\mathbf{L}$	Regularization matrix

$N$	Electron density
$N_l^m$	Legendre normalization constant
$P_l^m$	Legendre polynomial function
$P_{l,(\cdot)}^m$	Derivative of legendre polynomial function with respect to $(\cdot)$
$\mathbf{P}_{\mathbf{t}/\mathbf{t}}$	Covariance matrix of the state
$\tilde{\mathbf{P}}_{\mathbf{t}/\mathbf{t}}$	Covariance matrix of the state of correction estimation
$\mathbf{P}_{\mathbf{t}+1/\mathbf{t}}$	Forecast covariance matrix of the state
$\tilde{\mathbf{P}}_{\mathbf{t}+1/\mathbf{t}}$	Forecast covariance matrix of the state of correction estimation
$\mathbf{P}_{\mathbf{0},\mathbf{t}/\mathbf{t}}$	Background covariance matrix of the state
$\tilde{\mathbf{P}}_{\mathbf{0},\mathbf{t}/\mathbf{t}}$	Background covariance matrix of the state of correction estimation
$R_e$	Earth radius
$\mathbf{R}$	Observation covariance matrix
$\tilde{\mathbf{R}}$	Observation covariance matrix of correction estimation
$\mathbf{R}_{\mathbf{a}_0}$	Observation covariance matrix from not estimated background model transport terms
$\tilde{\mathbf{R}}_{\mathbf{a}_0}$	Observation covariance matrix from all background model transport terms
$\tilde{\mathbf{R}}_{\mathbf{FPI}}$	Observation covariance matrix from FPI neutral wind ingestion
$\mathbf{R}_{\mathbf{N}}$	Observation covariance matrix from ion continuity equation
$\tilde{\mathbf{R}}_{\mathbf{N}}$	Observation covariance matrix from correction ion continuity equation
$V$	Potential field
$V_0$	Climate model potential field

$\mathbf{W}_t$	Process noise covariance matrix
$\tilde{\mathbf{W}}_t$	Process noise covariance matrix of correction estimation
$Y_{lm}$	Scalar spherical harmonics
$a_{exb}$	Ion drift transport
$a_{u_{\parallel}}$	Neutral winds parallel transport
$a_{FPI}$	LOS neutral wind
$a_{0,u_{\parallel}}$	Background model neutral winds parallel transport
$a_{0,dfsn}$	Diffusion parallel transport model
$a_{0,exb}$	Background model ion drift transport
$a_{0,g}$	Gravity parallel transport model
$a_{0,loss}$	Loss rate model
$a_{0,prod}$	Production rate model
$a_{0,N}$	Sum of background terms not being corrected at one grid point
$\tilde{a}_{0,N}$	Sum of all background terms at one grid point
$\tilde{a}_{0,FPI}$	Background LOS neutral winds at one grid point
$az$	Azimuth angle
$\mathbf{a}_{ICON}$	ICON neutral wind estimation
$\mathbf{a}_{0,N}$	Sums of background terms not being corrected
$\tilde{\mathbf{a}}_{0,N}$	Sums of all background terms
$\hat{b}$	Unit vector parallel to magnetic field line
$\frac{\partial}{\partial t}$	Partial derivative with respect to time
$dt$	Time step used in filter

$\hat{e}$	Geographic east (zonal) direction unit vector
$el$	Elevation angle
$\mathbf{f}_{\text{exb}}$	Ion drift mapping matrix at one grid point
$\mathbf{f}_{\text{LOS}}$	Line of sight neutral winds mapping matrix at one grid point
$\mathbf{f}_{\mathbf{u}}$	Neutral winds mapping matrix at one grid point
$\mathbf{f}_{\mathbf{uE}}$	Neutral wind zonal component mapping matrix at one grid point
$\mathbf{f}_{\mathbf{uN}}$	Neutral wind meridional component mapping matrix at one grid point
$\mathbf{h}_{\text{exb}}$	Ion drifts transport contribution mapping matrix at one grid point
$\mathbf{h}_{\text{FPI}}$	FPI neutral winds ingestion mapping matrix at one grid point
$\mathbf{h}_{\text{N}}$	Ion continuity equation mapping matrix at one grid point
$\mathbf{h}_{\text{E,ICON}}$	ICON zonal neutral wind ingestion mapping matrix at one grid point
$\mathbf{h}_{\text{N,ICON}}$	ICON meridional neutral wind ingestion mapping matrix at one grid point
$\mathbf{h}_{\mathbf{u}}$	Neutral winds transport contribution mapping matrix at one grid point
$k$	Power series degree for radial direction
$k_{\text{exb}}$	Number of coefficients of electric potential expansion
$k_{\mathbf{u}}$	Number of coefficients of neutral winds expansion
$l$	Legendre polynomial order or power series degree for colatitudinal direction
$l_{\text{max}}$	Maximum Legendre polynomial order
$n_{\phi}$	Number of magnetic longitudes used in EMPIRE

$n_{qtimes}$	Number of time steps during quiet time used in EMPIRE
$\hat{n}$	Geographic north (meridional) direction unit vector
$m$	Legendre polynomial degree
$m_q$	Mass of charged particle
$p$	Power series degree for longitudinal direction
$q$	Charge
$r$	Earth centered radius
$\hat{r}$	Vertical magnetic and geographic direction unit vector
$\vec{r}$	Vertical magnetic and geographic vector
$s_{prod}$	Production rate
$s_{loss}$	Loss rate
$t$	Time
$\vec{u}$	Neutral winds
$\vec{u}_0$	Background model neutral wind vector
$\vec{u}_{\parallel}$	Field-parallel neutral winds
$u_E$	Geographic zonal neutral wind component
$u_{E,ICON}$	ICON neutral wind zonal component
$u_N$	Geographic meridional neutral wind component
$u_{N,ICON}$	ICON neutral wind meridional component
$u_{0,E}$	Background model neutral wind zonal component
$u_{0,N}$	Background model neutral wind meridional component
$\tilde{u}_{0,FPI}$	Background model neutral winds projected in the LOS direction at a grid point

$u_\theta$	Magnetic south neutral wind component
$u_\phi$	Magnetic east neutral wind component
$v_{SAMI3}$	Velocity from SAMI3 climate model
$\hat{v}_{EMPIRE}$	Velocity estimated with EMPIRE
$v_{exb,r}$	Ion drift up component
$v_{exb,\theta}$	Ion drift magnetic south component
$v_{exb,\phi}$	Ion drift magnetic east component
$\vec{v}$	Ion velocity
$\vec{v}_{g,0}$	Parallel velocity due to gravity
$\vec{v}_{exb}$	Ion drift
$\vec{v}_{dfs,0}$	Field-parallel velocity due to diffusion
$\vec{v}_\perp$	Perpendicular ion velocity
$\vec{v}_\parallel$	Field-parallel ion velocity
$\omega_i$	Gyro frequency
$\mathbf{w}_{t/t}$	Process noise
$\tilde{\mathbf{w}}_{t/t}$	Process noise of correction estimation
$\tilde{x}_c^{lm}$	Unknown coefficients corresponding to the cosine expansion of neutral winds correction
$\tilde{x}_s^{lm}$	Unknown coefficients corresponding to the sine expansion of neutral winds correction
$\hat{x}_{fitting}$	Estimated state vector fitting driver expansion to SAMI3
$\mathbf{x}$	State vector
$\bar{\mathbf{x}}$	Normalized state vector

$\tilde{\mathbf{x}}$	State vector of correction estimation
$\mathbf{x}_{\text{exb}}$	State vector of ion drift expansion
$\tilde{\mathbf{x}}_{\text{exb}}$	State vector of ion drift correction expansion
$\tilde{\mathbf{x}}_{\text{E}}$	State vector of zonal neutral winds correction expansion
$\tilde{\mathbf{x}}_{\text{N}}$	State vector of meridional neutral winds correction expansion
$\hat{\mathbf{x}}_{t/t}$	Estimated state
$\tilde{\hat{\mathbf{x}}}_{t/t}$	Estimated state of correction estimation
$\hat{\mathbf{x}}_{t+1/t}$	Forecast state
$\tilde{\hat{\mathbf{x}}}_{t+1/t}$	Forecast state of correction estimation
$\mathbf{x}_{\mathbf{u}}$	State vector of neutral winds expansion
$\tilde{\mathbf{x}}_{\mathbf{u}}$	State vector of neutral winds correction expansion
$\mathbf{x}_{0,\text{exb}}$	Background ion drift state
$\mathbf{x}_{0,t/t}$	Background state at time $t$
$\tilde{\mathbf{x}}_{0,t/t}$	Background state of correction estimation at time $t$
$\mathbf{x}_{0,t+1/t}$	Background state at time $t + 1$
$\tilde{\mathbf{x}}_{0,t+1/t}$	Background state of correction estimation at time $t + 1$
$\mathbf{x}_{0,\text{exb}}$	Background neutral winds state
$y_{\text{FPI}}$	FPI neutral wind measurement at one grid point
$y_{\text{N}}$	Electron density measurement at one grid point
$y_{\text{N,ICON}}$	ICON meridional neutral wind measurement at one grid point
$y_{\text{E,ICON}}$	Zonal meridional neutral wind measurement at one grid point
$\mathbf{y}_{\text{FPI}}$	FPI neutral winds measurements

$\mathbf{y}_{\text{ICON}}$	ICON neutral winds measurements
$\mathbf{y}_{\text{N}}$	Electron density measurements
$z_{\text{FPI}}$	Observation from FPI neutral winds ingestion at one grid point
$\tilde{z}_{\text{FPI}}$	Observation from correction FPI neutral winds ingestion at one grid point
$\tilde{\mathbf{z}}$	Observation of correction estimation
$\mathbf{z}$	Observation
$\mathbf{z}_{\text{FPI}}$	Observations from FPI neutral winds ingestion
$\tilde{\mathbf{z}}_{\text{FPI}}$	Observations from correction FPI neutral winds ingestion
$\mathbf{z}_{\text{ICON}}$	Observations from ICON neutral winds ingestion
$z_{\text{N}}$	Observation from ion continuity equation at one grid point
$\tilde{z}_{\text{N}}$	Observation from correction ion continuity equation at one grid point
$\mathbf{z}_{\text{N}}$	Observations from ion continuity equation
$\tilde{\mathbf{z}}_{\text{N}}$	Observations from correction ion continuity equation
$\delta a_{\text{exb}}$	Ion drift transport correction
$\delta a_{u_{\parallel}}$	Neutral winds parallel transport correction
$\delta a_u$	Neutral winds transport correction
$\delta u$	Neutral wind vector correction
$\delta u_E$	Geographic zonal neutral wind component correction
$\delta u_{\text{FPI}}$	Neutral wind projected into FPI LOS direction correction
$\delta u_N$	Geographic meridional neutral wind component correction
$\delta a_{\text{FPI}}$	LOS neutral wind component correction

$\delta V$	Potential field correction
$\epsilon$	Representation error in the observation space
$\epsilon_N$	Error of IDA4D electron densities
$\epsilon_{rel,IDA4D}$	Percentage of the relative error of IDA4D electron densities
$\bar{\epsilon}_{u_E}$	Zonal averaged representation error of the zonal neutral wind
$\bar{\epsilon}_{u_N}$	Meridional averaged representation error of the zonal neutral wind
$\epsilon_v$	Representation error in the velocity space
$\bar{\epsilon}_{v_{exb,\perp m}}$	Zonal averaged representation error of the perpendicular meridional ion drift velocity
$\bar{\epsilon}_{v_{exb,\perp z}}$	Zonal averaged representation error of the perpendicular zonal ion drift velocity
$\ell$	Normalized L-shell
$\mu_\epsilon$	Mean of representation error during quiet time
$\mu_\sigma$	Mean of difference between representation error and estimated error
$\nu_{in}$	Ion-neutral collision frequency
$\nu$	Observation error from ion continuity equation
$\tilde{\nu}$	Observation error from correction ion continuity equation
$\nu_{FPI}$	Observation error from FPI neutral winds ingestion
$\Phi_l^m$	Harmonic unction
$\Phi_{l,(\cdot)}^m$	Derivative of harmonic function with respect to $(\cdot)$
$\phi$	Magnetic longitude
$\hat{\phi}$	Magnetic east direction unit vector
$\Phi$	Gauss Markov process function

$\rho_+$	Ratio of ion-neutral collision to the gyro frequency of the ions
$\tau$	Kalman filter time constant
$\theta$	Magnetic colatitude
$\theta_0$	Reference magnetic colatitude
$\hat{\theta}$	Magnetic south direction unit vector
$\hat{\perp}_E$	Magnetic field perpendicular east direction unit vector
$\hat{\perp}_N$	Magnetic field perpendicular north direction unit vector
$\sigma_{climate}$	Standard deviation of climate background drivers
$\sigma_{fitting}$	Standard deviation of the drivers due to the fitting of basis functions expansion
$\sigma_0$	Standard deviation of background drivers in the drivers space
$\sigma_{x,0}$	Standard deviation of background state in the x-space
$\sigma_\epsilon$	Standard deviation of representation error during quiet time
Dst	Disturbance Storm Time index
EIA	Equatorial Ionization Anomaly
EMPIRE	Estimating Model Parameters from Ionospheric Reverse Engineering
FPI	Fabry-Perot Interferometer
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
HWM	Horizontal Wind Model
ICON	Ionospheric Connection Explorer
IDA4D	Ionospheric Data Assimilation Four Dimensional

IGRF	International Geomagnetic Reference Field
IRI	International Reference Ionosphere
ISR	Incoherent Scatter Radar
IT	Ionosphere-Thermosphere
LOS	Line Of Sight
MIGHTI	Michelson Interferometer for Global High-Resolution Thermospheric Imaging
NILE	Nighttime Ionospheric Localized Enhancement
NRL-MSISE	Naval Research Laboratory Mass Spectrometer Incoherent Scatter model
PPEF	Prompt Penetration Electric Field
PRE	Pre-Reversal Enhancement
RMS	Root-Mean-Square
SAMI3	Sami3 is Also a Model of the Ionosphere
SNR	Signal-to-Noise Ratio
TAD	Traveling Atmospheric disturbance
TEC	Total electron content
TECu	Total electron content units
UT	Universal Time
VSH	Vector Spherical Harmonics
WGS84	World Geodetic System-1984

## ABSTRACT

This dissertation studies a data assimilation algorithm that estimates the drivers of the ionosphere-thermosphere (IT) region of the Earth. The algorithm, EMPIRE (Estimating Model Parameters from Ionospheric Reverse Engineering) can estimate 2 main drivers of the ionospheric behavior: neutral winds and electric potential by ingesting mainly ionospheric densities obtained through Global Satellite System (GNSS) measurements. Additionally, the algorithm can ingest FPI (Fabry-Perot interferometer) neutral wind measurements. The contributions include 1) Vector spherical harmonic basis function for neutral wind estimation, 2) Quantification of the representation error of the estimations of the algorithm EMPIRE, 3) Analysis of Nighttime Ionospheric Localized density Enhancement (NILE) events and 4) Ingestion of global ICON (Ionospheric Connection Explorer) neutral winds measurements.

The IT region in the atmosphere is characterized by having a large concentration of free ions and electrons, electromagnetic radiation and Earth's magnetic field. The behavior of the region is dominated by the solar activity, that ionizes the free electrons of the region, forming ionospheric plasma and determining its density. Unusual solar activity or any atmospheric disturbance affects the distribution of the ionospheric plasma and the behavior of the IT region. The redistribution of the ionospheric density impacts technology widely used such as telecommunication or satellite navigation, so it is increasingly important to study the IT system response.

The IT behavior can be characterized by what drives its changes. Two drivers that play a key role, the ones we focus on this dissertation, are electric potential, that directly affects the charged ions in the system, and neutral winds, that refers to the velocity of the neutral particles that form the thermosphere. To quantify these drivers, measurements and climate models are available. Measurements are limited as the IT region is vast and covers the entire globe. Climate models can provide information in

all the region, but they are usually not as reliable during the unusual solar activity conditions or disturbances. In this dissertation we use a data assimilation algorithm, EMPIRE, that combines both sources of data, measurements and models, to estimate the IT drivers, neutral winds and electric potential. EMPIRE ingests measurements of the plasma density rate and models the physics of the region with the ion continuity equation. The drivers are represented with basis functions and their coefficients are estimated by fitting the expansions with a Kalman filter.

In previous work and use of the algorithm, the neutral winds were expanded using power series basis function for each of the components of the vector. The first contribution of the dissertation is to use a vector spherical harmonic expansion to describe the winds, allowing a continuous expansion around the globe and self-consistent components of the vector. Before, EMPIRE estimated the correction of the drivers with respect climate model values. In this work, EMPIRE is also modified to directly estimate the drivers. Then, a study of the representation error, which is the discrepancy between the true physics and the discrete model that represents the physics of EMPIRE and its quantification is done. Next, EMPIRE is used to analyze two NILE events, using the global estimation of both winds, from the first contribution, and the electric potential, derived in previous work. Finally, global estimation of winds allows us to implement the ingestion of ICON global winds in EMPIRE, in addition to the plasma density rate measurements.

## CHAPTER 1

### INTRODUCTION

The space environment can affect and has practical repercussions on technology that operates in space. Particularly, between 80 km to 1000 km above Earth's surface, there is a region in the atmosphere called the ionosphere-thermosphere (IT), characterized by having a large concentration of free ions and electrons, electromagnetic radiation and Earth's magnetic field. In this region technology must be engineered to withstand space weather and the changes that this region undergoes. Besides, the reliance on this type of technology is increasing over time [5], as traveling signals such as communication signals can suffer distortions or even complete disruption when abrupt changes in the region happen. Some example of these systems are the International Space Station [6], Global Navigation Satellite Systems (GNSS) or low-Earth orbit satellites, so the knowledge of the IT is crucial for the design criteria of these systems.

The space weather dominated by the Sun is going to determine the IT behavior. The free electrons in the IT region are ionized by solar radiation, giving rise to the ionospheric plasma that forms the ionosphere. The solar radiation is a large amount of energy radiated from the Sun to the Earth's space environment carried by the solar wind. This solar wind, a continuous flow of protons and electrons originating from the Sun's corona, interacts with the Earth's magnetic field, and abrupt changes in the solar wind can consequently cause abrupt variations in the Earth's space environment. Unusual solar activity or any atmospheric disturbance can affect the distribution of the ionization and the behavior of the IT system. Magnetic disturbance around the globe, lasting one or more days, due to variations in the solar activity is denominated "geomagnetic storm-time." The ionospheric plasma is redistributed in the IT region during these geomagnetic storms, and it is increasingly important to study the

system response due to the impact it has on widely used technology, such as telecommunications or satellite navigation. For example, [7] has studied the error in aircraft navigation services during a geomagnetic storm.

The objective of this dissertation is to understand better how the IT behaves and what drives the changes we can observe there, to prepare the way for reducing its effect on technologies affected by this behavior. The study of the plasma drivers will help us understand better the IT region during these events. One of the main drivers is that the electrons are in an electric and magnetic field. The charged particle is going to move in response to the fields and redistribute the plasma. This contribution to the ion velocity is called ion drift. It is described with more detail in Section 2.3. Thermospheric winds, or neutral winds, also play a key role in the ionospheric dynamics, as the neutral particles are coupled to the ions, which will be explained with more detail in Section 2.2.

The IT region is large and it covers the entire globe. This makes difficult to study the system with physical measurements, as it is difficult to have enough available measures of the whole region. Remote sensing methods are usually the most used to measure this region [8], [9], [10] as the alternative, in situ measurements, would require a large coverage. Also, it is difficult to obtain measurements at high altitudes of the IT region. Historically there has been a lack of data of neutral winds with sufficient altitudinal and horizontal resolution [11]. Due to the lack of data, the neutral winds were not as studied as the electron density and its variations [12]. The recent Ionospheric Connection Explorer (ICON) mission [13] has served to fill this gap, as it measures the winds globally.

Other options to study the IT system, such as ionospheric and thermospheric models, also present a challenge during storm times. They are not able to calculate the IT system behavior as good as during quiet time because physical mechanisms

are inadequately described [14]. There are 2 types of models: first-principle models, where the governing dynamics equations are defined and solved through numerical methods. First-principle models have a high computational cost, as it is solved in every location in each time step studied. The second type of model are empirical models, created by using basis functions to fit the ionosphere drivers to a large amount of data over time. These basis functions can be chosen, and they can represent to a greater or lesser extent the phenomena they are describing (for example using a first order polynomial expansion to fit a linear model would be a good representation). The main advantage is that they are not solved in every location, because coefficients to a function which covers all the locations is the unknown in this fitting problem. However, this type of model is not a good representation of the ionosphere if the measurements it is fitted to are not good enough, or in this case, they are too sparse over the region.

Another method to study the IT drivers globally is using data assimilation, in which regional measurements and global climate model data are combined. This technique has been used before in other works to study the density of the ionosphere ingesting GNSS global measurements of TEC (Total Electron Content) (number of electrons present along a specific path) [15, 16, 17, 18, 19, 20]. This method can also be used to update climate models. For example, [21] and [22] have used the global TEC measurements to update the model IRI. [23] developed IRTAM, an algorithm that updates IRI in real time ingesting Global Ionospheric Radio Observatory (GIRO) data. The previous algorithms are able to estimate the ionosphere behavior, but not the plasma drivers.

There are also algorithms that improve the IT drivers using data assimilation: [24] ingests thermospheric temperatures and uses this data to improve the model TIEGCM. Also, [25] improves the outputs of the system WACCM-X + DART ingest-

ing ICON measurements of the neutral winds while [26] does it by ingesting GOLD thermospheric temperatures. These models need measurements of the plasma drivers themselves, which are less plentiful than, for example, global TEC measurements.

There have been efforts to estimate the ionospheric drivers using global TEC measurements. GAIM-FP (Full Physics) model [27] is able to estimate the drivers. [28] have developed the Thermospheric Wind Assimilation Model (TWAM), that estimates the neutral winds correcting the winds output by GAIM-FP that uses observations of seasonal maps of F region ionosphere peak parameters. They estimate the neutral winds at low and mid latitudes. In this dissertation we develop a method for estimating neutral winds globally using the algorithm EMPIRE (Estimating Model Parameters from Ionospheric Reverse Engineering) [29] in Section 2.6.3.

In use for over a decade, EMPIRE was the first to separate the estimation of the drivers as a separate quantity of interest (and second step) from the estimation of plasma density [30]. EMPIRE estimates the physical ionospheric drivers, neutral winds and ion drift, by ingesting global electron density, which are derived from another assimilation algorithm: IDA4D (Ionospheric Data Assimilation Four-Dimensional [18]), which will be described with more detail in Section 2.6.1. IDA4D ingests primarily GNSS TEC measurements. Then, EMPIRE is used to estimate the drivers by using IT climate models in combination with these electron density output of IDA4D. This allows us to have observations distributed over the whole region.

## 1.1 Contributions

This dissertation describes the continued development of the EMPIRE algorithm for improved geomagnetic storm-time state estimation of the physical drivers of the IT system. In Chapter 2 a summary of the IT behavior and the model and algorithms used to study the region are described, including EMPIRE. The contri-

butions to the data assimilation algorithm EMPIRE are described in the following Chapters.

### **1.1.1 Vector spherical harmonics for EMPIRE neutral wind estimation.**

In Chapter 3 the first contribution is described. It consists on estimating global neutral winds with EMPIRE using a vector basis. In previous work, a power series basis expansion [30] was used to estimate the regional neutral winds. Each vector component was fitted separately so large-scale patterns present in the motion not orthogonal to the velocity wouldn't be captured [31]. Also, expanding a vector field directly into two scalar fields, as done previously in [32], included a dependence on the coordinate system used, while a direct representation of a vector basis allows us to capture the properties of the field independent of the coordinate system used [33]. This update in the method will allow us to study ionospheric dynamics during storms and enforces each component of the neutral winds to be consistent with each other.

The EMPIRE system estimates the correction of the drivers with respect to climate model values [30]. In [32], they updated the algorithm and added a Kalman filter solver. This introduced the initialization of the variances of the ingested and climate data. However, since the unknowns in the system were the corrections of the drivers, this step was complex. To simplify the implementation we directly estimate the drivers instead of the correction to a model, and we define an analysis to set up the variances of the new global neutral wind estimation.

### **1.1.2 Quantification of representation error in EMPIRE estimation.**

Space weather forecasting, which is useful as it can alert of upcoming geomagnetic storms and can give the time to react to minimize the impact of these events, also relies on data assimilation techniques. Understanding and separating the different uncertainties associated with the data assimilation algorithms is necessary for the optimal use of the method [34]. Data assimilation algorithms combine regional measurements

with global climate models, to improve estimation and forecasting of state and can be used for space weather analysis.

There are different sources of error in these systems, such as the assumption of unbiased errors, the inaccuracy of the background climate models used, or the errors of the observations. Different studies to assess how well data assimilation algorithms work have been conducted. [35] analyzes the impacts of model, background and observation error produced by a data assimilation algorithm using a variation error estimation method [36]. [37] and the references within analyzed a method to determine the systemic model bias in ensemble Kalman filters.

Analysis of the errors of methods focused on studying the IT has been done before. [19] have studied the model error of the GAIM-FP model. [38] and [39] study the model systematic bias of the TIEGCM algorithm and of a four-dimensional local ensemble transform Kalman filter (4D-LETKF) respectively. Calibration of the background climate models was analyzed by [40]. A method to estimate and account for the model error was also developed by [41].

One of the main challenges in data assimilation is to characterize the errors from discrepancies between observations and the physical model used in the data assimilation algorithm to describe them [42]. This error includes the representation error, which describes the error of using a discrete physical model that is not able to represent all the scales [43] and all the physical processes and dynamics of what is being observed [44]. [44] describes 3 main sources for the representation error: due to different scales represented in the observations and the model, due to the observation operator used to convert measurements into the state space, and quality control or pre-processing of observations.

The goal of this contribution is to give a methodology of the quantification of

the representation error of a data assimilation algorithm to study the fidelity of the algorithm’s model. We focus on the algorithm EMPIRE. In previous work [29, 7, 32, 45], the drivers have been estimated and studied during different geomagnetic events, but they haven’t focused on an assessment of the error of the estimations.

**1.1.3 Investigation of Nighttime Ionospheric Localized Enhancement.** These modifications to EMPIRE algorithm can be applied to study geomagnetic storms more accurately. When these geomagnetic storms are really strong they are called extreme geomagnetic storms. A co-rotating mid-latitude “nighttime ionospheric localized density enhancement” (NILE) has been observed during some geomagnetic storms [46, 47]. The event is characterized by an increase in TEC during local night that originates above the Caribbean and extends into continental USA. This event has been previously studied in [48] for two different geomagnetic storms. They were studied through an electron density analysis, using IDA4D (Ionospheric Data Assimilation four Dimensional [18]) outputs using algorithm SAMI3 [49] as an input, as the use of a physics model like SAMI3, rather than usually used empirical models, will provide for better resolution of the nighttime effects. However, information about the plasma drivers, which EMPIRE can provide, will help us understand and characterize them better. The objective of this contribution is to study the event and analyze the relative contributions of both neutral winds and electric fields with EMPIRE. This is described in Chapter 5. Also, the ingested data will be the IDA4D coupled to SAMI3 algorithm, that gives more accurate outputs during night time.

The estimation of the IT drivers with EMPIRE can be improved, as shown in previous work [32], by assimilating measurements of neutral winds by Fabry-Perot interferometers (FPIs) in addition to TEC IDA4D measurements. The ingestion of the neutral velocities can significantly change the drivers estimation. Another goal of this contribution is to study the NILE event ingesting neutral wind measurements,

but using the global vector spherical harmonics expansion. I will examine if the new vector spherical harmonics expansion changes the drivers estimation.

**1.1.4 Additional ingestion of ICON neutral winds measurements.** Measurements of neutral winds can also be ingested to increase the performance of the algorithm EMPIRE [32]. In previous work, the local measurements of the winds FPI instruments were ingested to improve the EMPIRE estimation performance. However, the measurements were limited to the number of instruments available and to the location they were. As mentioned before, the recent ICON mission [13] measures winds globally, as the MIGHTI (Michelson Interferometer for Global High-Resolution Thermospheric Imaging) instrument that measures the winds is located in the ICON satellite that orbits around the Earth. The first contribution allows EMPIRE to estimate global winds and consequently to ingest global winds measurements. The last contribution described in Chapter 6 modifies EMPIRE to make it able to ingest ICON neutral winds measurements. The estimated IT drivers, both ingesting and not ingesting the winds, for a small geomagnetic storm are analyzed and compared to each other.

## CHAPTER 2

### TECHNICAL BACKGROUND

An overview of the ionosphere-thermosphere region is given in Section 2.1. Section 2.2 introduces the thermosphere and gives a description of the models and instruments that can give information of this region. Similarly, the ionosphere and its possible sources of information are explained in Section 2.3. The magnetic field of the Earth, that plays a role in the dynamic of the ionosphere region, is described in Section 2.4. In this section, the different coordinate systems used in this dissertation are also introduced. In Section 2.5, the interaction between the ionosphere and thermosphere is explained, as well as the geomagnetic storms and general ionospheric features. Finally, Section 2.6 explains the different algorithms and models used throughout this dissertation.

#### 2.1 Overview

The atmosphere is divided into different layers according to their properties and behavior. From lowest to highest, the major layers are the troposphere (0km-12km), stratosphere (12km-50km), mesosphere (50km-80km), thermosphere (80-600km) and exosphere (700km to the outer space around 10,000km). As mentioned before, this work focuses on the thermosphere region. As its own name implies, the thermosphere is characterized by a large increase in temperature. Gradients in temperature and density can cause movement of the gas in this region, gas formed more than 99% by neutral particles. Molecular oxygen ('O') is the dominant species in this region.

The ionosphere consists of ions and free electrons within the thermosphere. The ionosphere is overall electrically neutral with equal number of ions and electrons. Radiation from the Sun brings enough energy to this region to cause ionization of

the neutral atoms and molecules giving rise to the ionospheric plasma that forms the ionosphere. In practice, the lower altitude limit is around 50 to 70 km and there is no true upper limit although it transitions to the plasmasphere. The ionosphere and thermosphere regions overlap, and the behavior of the ions of the ionosphere can affect the dynamics of the neutrals, and on the other hand, neutrals moving due to temperature or pressure gradients, waves and tides can affect the behavior of the ions. The different layers of the atmosphere and the ionosphere are represented in Figure 2.1. In the IT region is where the aurora borealis and aurora australis happen. The ISS (International Space Station) and some low altitude satellites are also located in this region and they are affected by the IT behavior and distribution. For example, the heating and consequent expansion of the atmosphere exert drag on low altitude satellites, changing their orbit characteristics. Radiation in the region can damage equipment and poses a health hazard to the astronauts inside the ISS. Electromagnetic signals transmitted between communication satellites and the ground or from GNSS (Global Navigation Satellite Systems) to ground, air, or sea vehicles also travel through the IT region. The ions in the IT region can delay the signals as they refract them when they travel through the region, which can cause communications outages or loss of positioning accuracy.

## 2.2 Thermosphere

Analyzing individual particles of gas, the motion of the thermosphere is dominated by collisions between those particles. At a large scale, we can study a set of particles as a whole using bulk properties with Navier-Stokes fluid equations, as in the lower to middle thermosphere the density and scale sizes are enough to make this assumption. The distance between collisions is much shorter than the scale size of interest.

We refer to the neutral winds as the velocity of the neutral particles. This

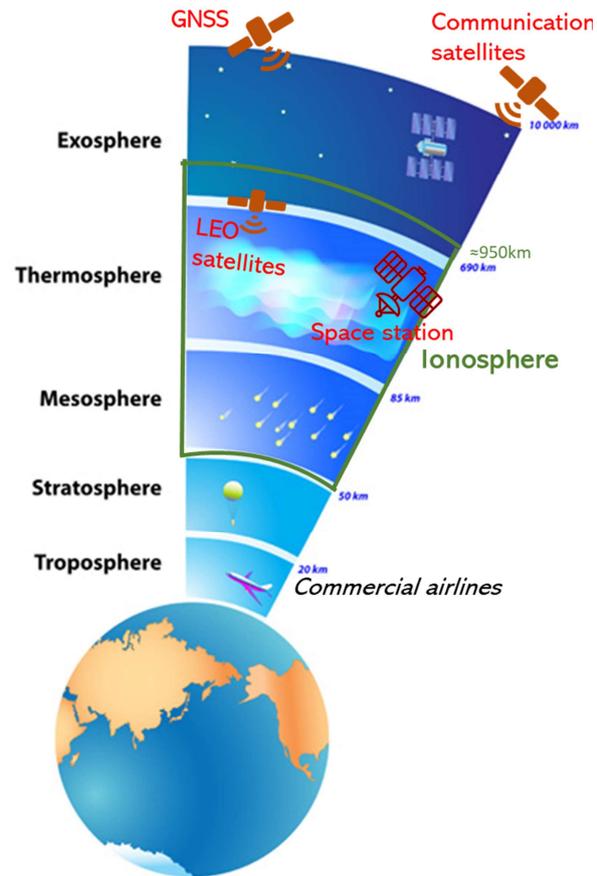


Figure 2.1. Layers of Earth's Atmosphere. Adapted from [2].

motion is controlled by pressure gradients produced in response to differential heating in the thermosphere. When warmer air rises, the neutral particles are displaced by descending cool air. This flow of the neutral particles is mostly horizontal, with some smaller vertical movement.

In the thermosphere, there are different heat sources and losses that can affect the motion of the neutral particles. The main source of heat is the absorption of solar radiation and the principal cooling process is downward heat conduction. The solar heat that the thermosphere absorbs has a diurnal and seasonal variation, as the electron concentration variations described before do. There are also longer-term variations, for example the solar radiation depends on the 27 day rotation period of the Sun and the solar cycle mentioned before.

For altitudes above 200 km, in the F-region of the ionosphere, the thermospheric winds are usually eastward during the night (16:00-04:00 local time (LT)) and switch to westward during day (04-16 LT) [50]. This zonal (east/west) behavior is true for all latitudes on Earth. The meridional component (north/southward) depends on the magnetic latitude of the location studied. The winds in the northern and southern hemispheres will blow equatorward during the local night and poleward during daytime [51].

At high latitudes, during geomagnetic storms there can also be another heat source. Charged particles precipitate, transferring energy to the neutral particles. Also, during these storms, changes in the electric field increase the transfer of energy to the thermosphere. This translates into a coupling between the thermosphere and the ionosphere due to the ion-neutral collision. If studying the ion motion, neutral winds can accelerate or decelerate the ions (ion drag). Therefore, the electrodynamic of the ionosphere influence the dynamics of the thermosphere and on the other hand, the neutral winds can have a significant effect on the ionosphere behavior.

There are different methods to study the neutral winds, climate models or measurements, as described in the introduction. Two instruments that measure the winds are described in following sections, as they will be used in this work both for validation of results and for ingesting in the data assimilation algorithm EMPIRE.

**2.2.1 Models.** Empirical models of thermospheric winds depend almost entirely on available data. The Horizontal Wind Model 2014 (HWM14) [50] is a model that generates winds by fitting ground-based FPI (Fabry-Perot Interferometer), incoherent scatter radar (ISR), satellites, and rocket measurements. This model will be used throughout this work as the climate model for the neutral winds.

**2.2.2 Instruments.** The FPI (Fabry-Perot interferometer) is a ground-based instrument that measures the neutral winds. It uses a passive radio technique that measures the Doppler shift of the recombination electromagnetic nightglow emissions [52]. Nightglow is light emitted by neutral particles when they gain energy by either chemical reactions or collisions with electrons in the ionosphere to release their extra energy. Different emissions come from different species, and because they are moving, the emissions received at the FPI are Doppler-shifted. Depending on the emission line chosen to use in the instrument, different altitudes can be studied as the species are not mixed in the thermosphere. FPI instruments measure line-of-sight (LOS) of the instrument winds only.

The measurements can be used to study the neutral winds. However, the trustworthiness of the measurements need to be considered. Contamination of the measurements can occur when there are large airglow gradients [53], which happen in the storm time mid-latitudes, but not as much at low latitudes. The brightness of the samples can help determine the quality of the measurements. A bright region of the sky reduces the error of the FPI measurements. During storms at mid latitudes, the brightest region is poleward so the northern LOS will have better measurements of the

winds. At equatorial latitudes, the nightglow intensity gradient is not purely zonal or meridional, and it varies significantly from season to season and night to night [54]. Because of the nature of the instruments, cloudy regions can change the emissions received at the FPI location, consequently changing the measurements. There are cloud sensors available at the same location for some FPIs and their information can also be studied to determine the quality of the measurements.

ICON (Ionospheric Connection Explorer) [1] is a NASA mission that focuses on studying the low latitude ionosphere properties. It is positioned in a low-Earth orbit to observe the ionosphere around the equator and its instruments aim to sense at altitudes from 90 kilometers, around the lowest boundary of space, up to 580 kilometers. The period of the orbit is 97 minutes. ICON carries 4 different instruments that help understand how the ionosphere behaves: MIGHTI (Michelson Interferometer for Global High-resolution Thermospheric Imaging), IVM (Ion Velocity Meter), EUV (Extreme Ultra-Violet) and FUV (Far Ultra-Violet) instrument. The MIGHTI instrument measures the neutral wind at low latitudes covering altitudes between 90 and 300 km.

MIGHTI measures the horizontal wind speed and direction (the wind vector). It is formed by 2 MIGHTI units and their viewing angle of the same site separated by 90 deg allowing for estimation of the wind vector as each unit measures the wind along its LOS [55]. It is based on an interferometric technique, similar to the FPI technique described before. The LOS winds are also calculated from the Doppler shift of emission lines in the Earth's upper atmosphere.

### 2.3 Ionosphere

The different layers of the ionosphere are shown in Figure 2.2: D, E and F layer, which splits into F1 and F2 layers during daytime. The ionosphere's vertical

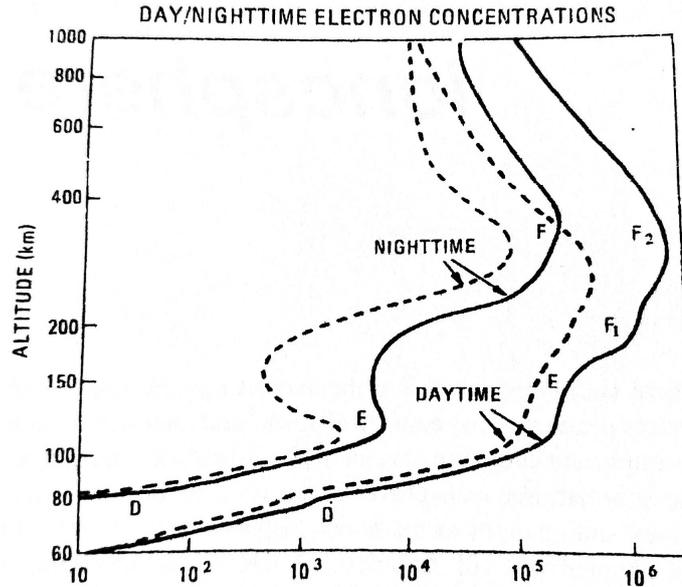


Figure 2.2. Typical midlatitude daytime and nighttime electron density profiles for solar maximum (solid lines) and minimum (dashed lines) [3].

plasma density profile is also shown. It can vary from day to night, with the seasons, with latitude. We can see in Figure 2.2 that in a typical daytime the electron density increases because the photoionization increases due to the solar radiation. At nighttime, the electron concentration decreases as recombination start to happen. Some of the layers, F<sub>1</sub> and D-layers even disappear after sunset. It is also observed that during the sunspot maximum, indicated with a solid line, where there is more solar activity and more energy reaches the ionosphere from the Sun, the electron density is higher that in the solar minimum in all of the layers. This solar cycle peaks and weakens every 11 years. Without these layers, the solar radiation that reaches the ground would be greater.

The solar cycle refers to the Sun's periodic variation in energy and activity, caused by the hot plasma moving in the Sun and the magnetic field it carries [56] and has a period of 11 years. The magnetic field changes consequently cause changes in the solar activity. The beginning of a solar cycle is called solar or sunspot minimum.

The middle of the solar cycle is the solar maximum when the number of sunspots visible on the Sun peaks. As the cycle ends, it fades back to the solar minimum and then a new cycle begins. Each cycle can be more or less active. We are currently in solar cycle 25, since January 2020. Solar cycle 23 goes from approximately August 1996 to December 2008 and solar cycle 24 starts in December 2008 until December 2019 [57]. Solar cycle 24, with its maximum around April 2014 was weaker than solar cycle 23, whose maximum occurred around November 2001 [57].

Ionospheric behavior in terms of electron density movement is influenced by the movement of the plasma, which varies due to two main causes. The first cause for movement is that the charges are in an electric and magnetic field. The other is due to the neutral particles moving in the thermosphere region, which drag the ions and the electrons and causes change in the plasma velocity, described in the previous section.

To understand the plasma velocity due to electric and magnetic forces, first we explain how a single charged particle of that plasma behaves due to magnetic field forces. When this particle moves only in a magnetic field, a force acts on it:

$$\vec{F} = q\vec{v} \times \vec{B}_0 \quad (2.1)$$

where  $q$  is the charge intensity,  $\vec{v}$  is the charged particle's velocity and  $\vec{B}_0$  is the magnetic field. This expression show us that if a particle is moving in a plane and is in presence of a magnetic field perpendicular to this same plane, this particle will start to follow the trajectory of a circle in a counterclockwise or clockwise direction, depending on the charge of the particle. This can easily be seen using the right-hand rule. Figure 2.3 illustrates this motion.

The circular motion is governed by a frequency, that is denominated gyro frequency  $\omega_i$ . This angular frequency depends on the charge  $q$  and mass  $m_q$  of the

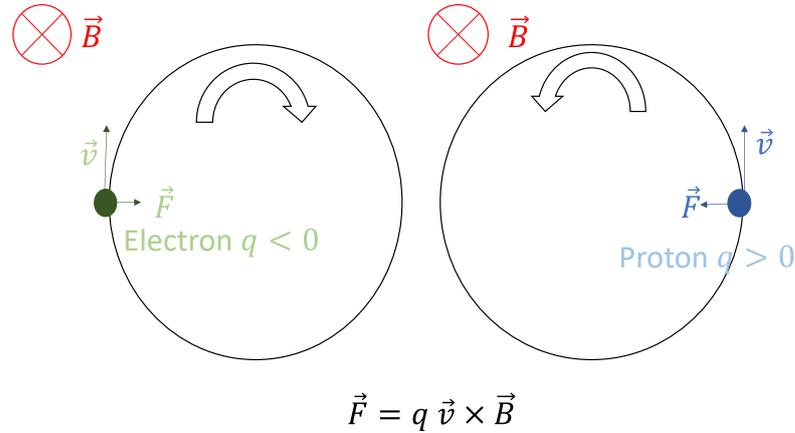


Figure 2.3. Single charged particle trajectory when embedded in a magnetic field according to the force equation.

charged particle as well as the magnetic field magnitude  $B_0$ .

$$\omega_i = \frac{qB_0}{m_q} \quad (2.2)$$

However, in the ionosphere the charges are also affected by the presence of electric fields,  $\vec{E}$ . When an external force also acts on a charged particle, besides the magnetic field, these particles drift through the magnetic field in addition to the basic circle motion mentioned before. Also, because the electric force  $\vec{E}$  depends on the charge  $q$  of the particle, this drift is charge-independent, so the drift direction and speed is the same for both electrons and protons. Summarizing, the ions in the ionosphere drift at rate  $\vec{v}$  given by the following the equation (derived from Equation 2.1):

$$\vec{v} = \frac{\vec{E} \times \vec{B}_0}{B_0^2} \quad (2.3)$$

where  $B_0$  is the norm of the magnetic field vector  $\vec{B}_0$ . Figure 2.4 shows the trajectories of an electron and a proton for a given magnetic field and electric field direction. Both

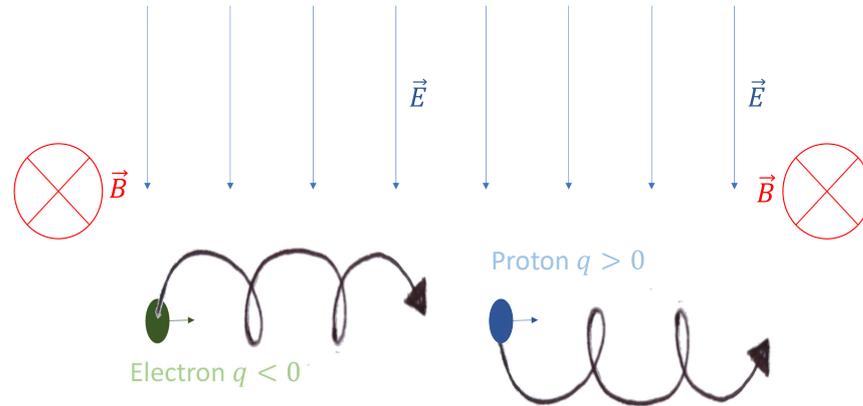


Figure 2.4.  $\vec{E}$  cross  $\vec{B}$  drift of ions.

the electron and proton in this example drift to the right.

**2.3.1 Models.** To study the ion drifts, there are different climate models or measurements available to do so. The Weimer model [58] is used in this work in Chapters 3, 5 and 6 to describe the high latitude electric field  $\vec{E}$ . We also use the ion drifts derived in the model SAMI3 [49] for analyses in Chapter 4. The magnetic field  $\vec{B}_0$  is obtained from climate model International Geomagnetic Reference Field (IGRF-11) [59].

**2.3.2 Measurements.** The ISR (incoherent scatter radar) is a ground-based instrument that measures the ion drifts. It uses a technique based on the transmission of high-power radio waves with high frequencies. The radar wave scatters incoherently, randomly in direction, from the electrons in the ionosphere [60]. A higher electron concentration translates into more reflected electromagnetic waves and more power returning to the radar. The amount of energy reflected by one electron is known, so using the total received intensity measured, the electron density in the studied region

can be calculated. If the ions in the plasma are moving, then the received signal will be Doppler shifted, so the data will have different frequencies. This allows the ion drift at the scattering site to be estimated via ISR [61].

The TEC (Total Electron Content) is a measurement of the electron density  $N$ . TEC is defined as the integral of electron density present along a path between a radio transmitter, such as GPS (Global Positioning Systems) or GNSS (Global Navigation Satellite System) satellites and a ground receiver [62]. It is calculated as the integral of the electron density  $N$  along the path between the radio transmitter and the receiver. As density  $N$  is measured in number of electrons per cubic meter, TEC is electrons per square meter. By convention, 1 TEC Unit (TECu) is defined as  $10^{16}e^-/m^2$ . The order of magnitude of the vertical TEC in Earth's ionosphere is around  $\approx 1 - 100$  TECu. Measurements of slant TEC through dual-frequency GNSS indicates how the plasma in the ionosphere is redistributed when geomagnetic storms occur. However, the information of what is driving the TEC changes is given by IT drivers.

## 2.4 Magnetic Field

The Earth's magnetic field  $\vec{B}_0$ , which plays a role in the ion drift dynamic, can be simplified to a dipole model with its axis offset by approximately 11 degrees from Earth's spin axis. A schematic diagram of this model is shown in Figure 2.5, with the spin and geomagnetic axes are indicated.

To characterize the field lines in the dipole model the L-shell  $L$  parameter can be defined. L-shell describes the set of magnetic field lines which cross the Earth's magnetic equator at a number of Earth-radius ( $R_e$ ) equal to the L value.

$$L = \frac{r}{R_e \sin^2 \theta} \quad (2.4)$$

Where  $\theta$  is the magnetic colatitude. A representation of this parameter is shown

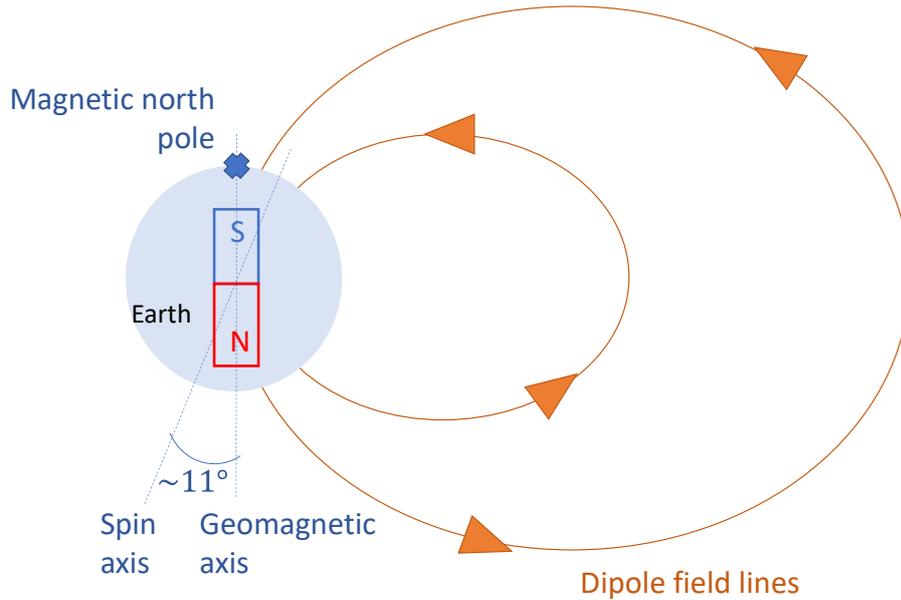


Figure 2.5. Schematic drawing of the Earth's magnetic field with the geomagnetic north pole, spin axis and field lines labeled, based on [4].

in Figure 2.6, where magnetic field lines from the dipole model from the previous example in Figure 2.5 corresponds to  $L = 3.5$  and  $L = 6$ .

The angle between the local horizon and the magnetic field lines is the inclination angle ( $I$ ). At the magnetic equator this angle will be 0 degrees and at the north pole it is 90 degrees, as convention states that north of the magnetic equator the inclination angle is positive and south it is negative. In Figure 2.7 the inclination angle  $I$  is shown for a specific location (radius  $r$ , magnetic colatitude  $\theta$  and magnetic longitude  $\phi$ ) as well as the local horizontal represented with a dashed pink line. The unit vector  $\hat{b}$  parallel to the magnetic field line direction is defined and show in the schematic drawing in orange.

It is useful later in the work to use a magnetic field coordinate system, which is defined by the unit vector  $\hat{b}$  parallel to the magnetic field line direction, instead of the spherical magnetic coordinate system given by  $(\hat{r}, \hat{\theta}, \hat{\phi})$ , represented in Figure

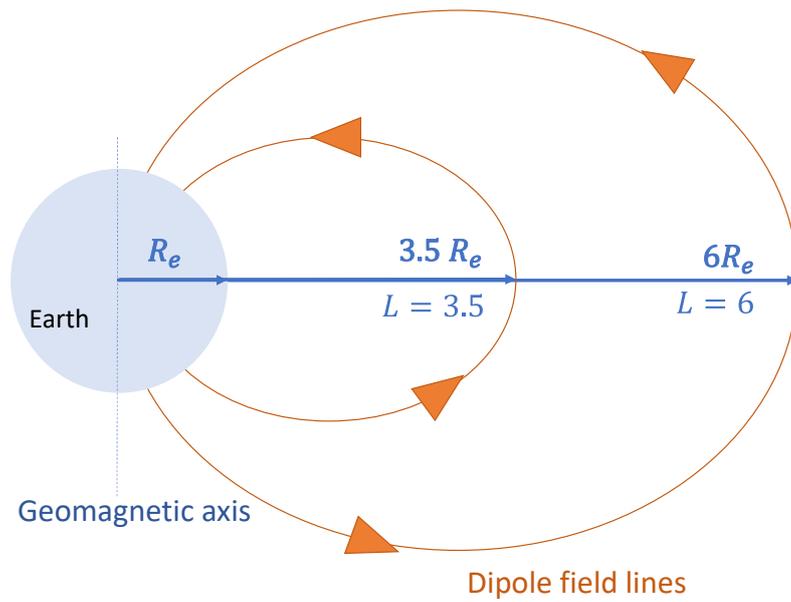


Figure 2.6. Schematic drawing of the Earth's magnetic field based on a dipole model and its corresponding L-shell values.

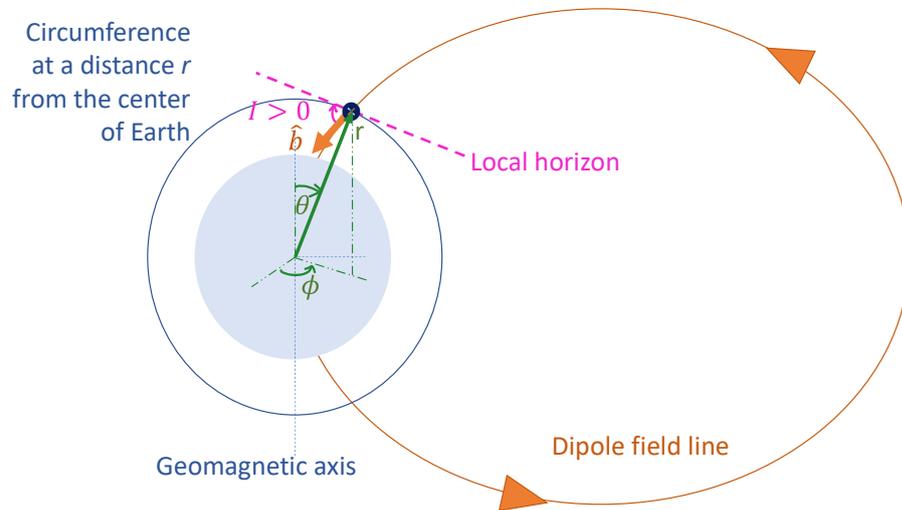


Figure 2.7. Schematic drawing of the inclination angle with the local horizon, the magnetic coordinates ( $r$ ,  $\theta$  and  $\phi$ ) and the unit vector parallel to the magnetic field  $\hat{b}$  labeled.

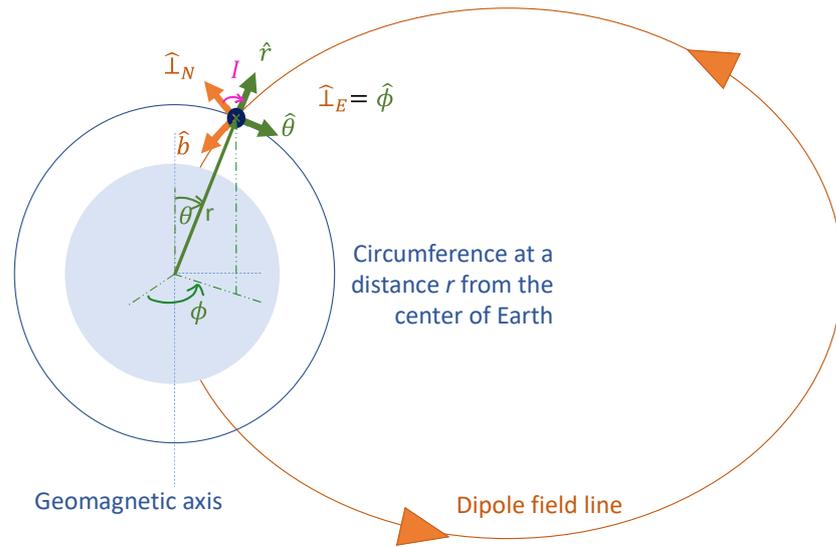


Figure 2.8. Schematic drawing of the spherical coordinate system  $(\hat{r}, \hat{\theta}, \hat{\phi})$  and the magnetic field coordinate system  $(\hat{b}, \hat{\perp}_N, \hat{\perp}_E)$ .

2.8 with green arrows. The magnetic  $\hat{\theta}$  is parallel to the local horizon. The magnetic field coordinate system is defined by the orthogonal basis:  $(\hat{b}, \hat{\perp}_N, \hat{\perp}_E)$ , where the perpendicular north  $\hat{\perp}_N$  and perpendicular east  $\hat{\perp}_E$  are unit vectors perpendicular to the magnetic field line given by unit vector  $\hat{b}$ . The direction  $\hat{\perp}_N$  is in the same plane as  $\hat{b}$  and the local horizon given by  $\hat{\theta}$ , and  $\hat{\perp}_E$  coincides with the magnetic zonal direction  $\hat{\phi}$ . This coordinate system is shown in Figure 2.8 with orange arrows. To transform these coordinates to geographical coordinates we use the inclination angle.

On the other hand, the neutral winds are usually given and studied in a geographic coordinate system. The geographic coordinate system is defined by the basis vectors  $(\hat{e}, \hat{n}, \hat{r})$ , where  $\hat{e}$  is geographic east or zonal,  $\hat{n}$  is geographic north or meridional and  $\hat{r}$  is the vertical direction. The system is characterized by the declination angle  $D$ , which is the angle between the magnetic north and true north, where the geographic North Pole is located. The angle is positive when magnetic north is east of true north and varies depending on the location on the Earth's surface and over time.

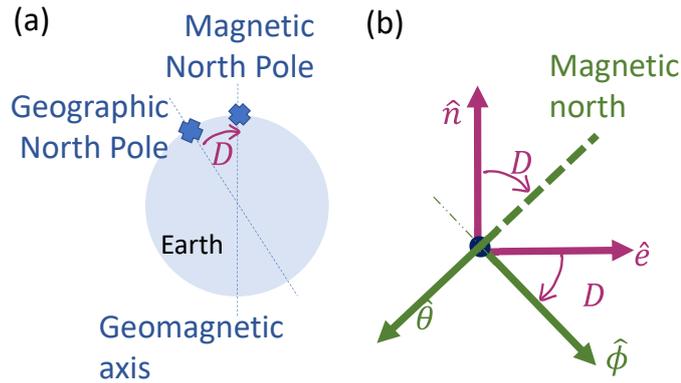


Figure 2.9. Schematic drawing of the spherical coordinate system  $(\hat{r}, \hat{\theta}, \hat{\phi})$  and the geographic coordinate system  $(\hat{e}, \hat{n}, \hat{r})$ .

The angle is indicated in Figure 2.9a. In Figure 2.9b the change of coordinates from magnetic spherical coordinates, shown in green, to geographical coordinates system, indicated in purple arrows, is represented.

## 2.5 IT interaction

The thermosphere is a fluid medium governed by frequent collisions between individual particles of gas, so the use of governing equations such as the Navier-Stokes equations, continuity and energy, to characterize its properties is possible. However, the thermosphere layer overlaps with the ionosphere, so the use of these equations is valid until the fluid approximation is not possible. This limit is given when the collisions is short compared to the scale sizes, around 600 km.

The dynamics of this coupled region is driven mainly by neutral winds and ion drifts. To understand their relationship, first we review the ions motion due to neutral winds considering as if the electric field was absent. We consider the case where the ions and neutral atmosphere are present only in a magnetic field [63]. The motion would be governed by the ratio  $\rho_+$  of ion-neutral collisions  $\nu_i n$  to the gyro

frequency of the ions,  $\omega_i$ , defined in the Section 2.3. The velocity of the ions,  $\vec{v}$ , would be:

$$\vec{v} = \frac{1}{1+\rho_+^2} [\rho_+^2 \vec{u} + \rho_+^2 \vec{u} \times \hat{b} + (\vec{u} \cdot \hat{b})\hat{b}] \quad (2.5)$$

with

$$\rho_+ = \frac{2\pi\nu_i n}{\omega_i} \quad (2.6)$$

The neutral wind is represented with  $\vec{u}$  and the magnetic field direction as  $\hat{b}$ . At high altitudes (F-region, above an altitude of approximately 150 kilometers), the ion-neutral collision frequency  $\nu_{in}$  is very small:  $\rho_+ \ll 1$ , so the above equation can be simplified to:

$$\vec{v} = (\vec{u} \cdot \hat{b})\hat{b} \quad (2.7)$$

The hypothesis of collision free means that the motion of the ions due to neutral winds only contributes in the direction of the magnetic field line  $\hat{b}$ . On the other hand, from Equation 2.3, we observe that the ion drift is perpendicular to the magnetic field line  $\hat{b}$ , which is the direction in which the magnetic field vector  $\vec{B}_0$  is oriented. Because of this, it is common to use a coordinate system oriented with the magnetic field direction, described in Figure 2.8.

**2.5.1 Geomagnetic storm.** The study of these drivers and their effect on the redistribution of the plasma in the IT region is important during geomagnetic storms. As described previously, a geomagnetic storm is a significant disturbance in the Earth's magnetic field [64] caused by specific conditions and variations in the solar wind that reaches the Earth. The largest storms are associated with solar coronal mass ejections (CMEs), where solar plasma embedded with its magnetic field carried with the solar wind arrives at the Earth [65].

There are different geomagnetic indices such as Dst, AE and Kp, which summarize global geomagnetic variability and help detect geomagnetic storms [66]. [64] gives a threshold of the Dst (Disturbance Storm Time) index for which we can con-

sider an event a geomagnetic storm or not. The solar wind can provoke major changes in the currents, the plasma distribution and the fields in the Earth's magnetosphere, such as a ring of westward current around Earth. The Dst index measures the intensity of this ring current and acts as an indicator of geomagnetic activity and the size of the storms [67]. Geomagnetic storms are characterized by three phases: initial, main and recovery [68]. The main phase occurs when the energy input from the solar wind intensifies and consequently increases the intensity of the ring current. The ring current produces a magnetic field that inside the ring is opposite of Earth's internally generated field. Ground magnetometers measure a decrease in magnetic field strength during storms (main phase) and then weakening ring current produces a magnetic field increase to baseline (recovery phase) that the magnetometers also indicate.

Other effects the storm has are explained below. The changes in the solar wind arriving at Earth and interacting with the Earth magnetic field produce particle precipitation followed by a sudden heating in the high latitudes and altitudes of the atmosphere. Additionally, Joule heating of the thermosphere also occurs at high latitude. Joule heating arises from the currents flowing in the ionosphere, that causes friction between the neutral particles and the ions [69]. This sudden high energy has different effects, such as thermospheric circulation or currents, TADs (traveling atmospheric disturbances) or changes in the electrodynamics processes. At mid latitudes, effects of the storm are also visible. The TADs can propagate neutral winds disturbances from auroral to lower latitudes, and because they are coupled with the ions behavior, they modify the plasma behavior at these latitudes [70]. Prompt penetration electric field (PPEF) at high latitudes can also cause disturbances in the ions dynamics at mid-latitudes [71]. PPEF refers to the interplanetary electric fields carried by the solar winds appearing immediately in the Earth's ionosphere after the solar wind reaches the Earth's magnetic field [72], causing disturbances in the plasma dynamics.

**2.5.2 General Ionosphere Features.** There are several features characteristic of the IT region, which are not linked to a geomagnetic storm as they also occur during quiet time, which will be used in this work. For example, the South Atlantic anomaly, which describes a region where the magnetic field strength is weaker and there is an increase of ions in this area. It is caused by the geomagnetic axis not passing through the center of the Earth. This feature is always present [73].

The equatorial fountain describes the effect of the plasma moving upwards and rising several hundred kilometers at the equator [74] and this feature is also not characteristic to geomagnetic storms. It is caused by the north geomagnetic field (at the equator the vector parallel to the magnetic field lines  $\hat{b}$  is oriented in the magnetic north direction  $\hat{b} = -\hat{\theta}$ ) combined with a daytime east-west ionospheric electric field. Using Equation 2.3, the resulting ion drift, that is the cross multiplication of the electric field and the magnetic field, is upwards. During daytime, the electric field is typically eastward [75], which creates an upward ion drift rising the plasma to higher altitudes. During local sunset, also called evening terminator, the eastward electric field strengthens before it reverts to westward during local night time. This phenomena is called Pre-reversal enhancement (PRE) and it causes an increase of the vertical drifts from the equatorial fountain during sunset [76]. At the terminator line there is also a gradient of the electron concentration, as there are more ions on the day side due to the ionization from solar UV radiation [75].

After the ions move vertically to higher altitudes, pressure and gravitational forces will cause the plasma to diffuse poleward along magnetic field lines to higher magnetic latitudes, usually to  $\pm 10$  deg. At higher altitudes, the recombination rate is smaller [77] so the electron concentration will increase. This phenomenon, where the ionization increases around the magnetic equator, is denominated equatorial ionization anomaly (EIA) [78]. This phenomenon is visible everyday during daytime, when

the eastward electric field and the northward magnetic field at the equator generate the upward ion motion.

However, the storm time fountain effect can be enhanced and transformed to a “super” fountain during geomagnetic storms. The eastward electric fields at low latitudes can be enhanced during geomagnetic storms due to the Joule heating that consequently alters the generation of electric fields [79], this effect is called ionospheric disturbance dynamo, or due to the PPEF that can occur during geomagnetic storms. This also enhanced the EIA, resulting in ionization peaks at higher magnetic latitudes [48].

The EIA can also be distributed asymmetrically (more TEC enhancement in one hemisphere than in another) due to the behavior of the neutral winds. There are two theories that can explain the asymmetries in the EIA [80]. Intra-hemisphere transport (within one hemisphere) says that southern winds drive plasma up in the northern hemisphere and consequently creates a higher crest of enhanced TEC within the Northern hemisphere than in the Southern hemisphere. This form of EIA asymmetry occurs when the equatorial fountain effect is stronger than the neutral winds. Southern (northern) winds move plasma along the field lines but a strong fountain effect at the equator restricts the plasma motion in the windward direction and consequently the ions move upward and poleward, creating a TEC enhancement in the Northern (Southern) hemisphere. The trans-equatorial transport occurs when the fountain effect is weaker than the winds. In the trans-equatorial case southern (northern) neutral winds create an asymmetry of the EIA with a TEC enhancement in the Southern (Northern) hemisphere. The winds will move the plasma along the field lines to higher altitudes in the opposite hemisphere since the winds overcome the weak fountain effect, forming the asymmetry in the EIA. [80] found that strong winds (Intra-hemisphere transport) explained most of the cases observed and modeled in

that work.

**2.5.3 NILE.** A feature that has been observed by analyzing the TEC behavior for some geomagnetic storms is a NILE (nighttime ionospheric localized density enhancement). It has been observed so far at mid-latitudes at higher altitudes [47], only around one spot in the world (near the Gulf of Mexico region) and at nighttime. It is characterized by an increase in the total electron content. There are a several cases where this phenomena has been observed, for example, during solar cycle 23: October 2003 [47], April 6-7 2000 and July 15-16 2000 [81]. Previous work [7] also detected some cases where NILE seemed to appear on the same location but not during a extreme storm. This case belongs to the solar cycle 24 and it occurred on October 25th 2011. NILE for 2 geomagnetic storm has been previously analyzed in [48] through an electron density analysis, but information about the plasma drivers would help us understand and characterize them better. Possible causes for this NILE hypothesized in the literature include an electrodynamic effect tied to the South Atlantic Anomaly or an equatorial super-fountain [82] associated with a stormtime prompt penetration electric field (PPEF).

In recent studies [48], two NILE events were studied through a TEC analysis and they show that the NILE appears to originate from the enhancement of the equatorial ionization anomaly (EIA) in the recovery phase of the geomagnetic storm. The EIA for the cases studied in [48] was asymmetric as the northern crest, where the NILE appears to originate, was more enhanced than the southern crest. In [82] they hypothesize that the South Atlantic anomaly creates a preferred configuration of longitude and time where the enhanced TEC of the NILE builds up and that during geomagnetic storms the enhanced electric fields during strong storms at the terminator redistribute the plasma from the EIA crests to magnetically-conjugate regions.

## 2.6 Algorithms

First we describe the IDA4D algorithm that outputs global electron density  $N$ . Then, a small background of model SAMI3 is given. Data assimilation algorithm EMPIRE is described is also described.

**2.6.1 IDA4D.** The algorithm IDA4D (Ionospheric Data Assimilation Four - Dimensional [18]) is a data assimilation algorithm that outputs global electron density values  $N$  and ingests a variety of data sources linearly and non linearly dependent on the electron density  $N$ , like slant Global Navigation Satellite Systems (GNSS) total electron content (TEC), GNSS radio TEC, electron density altitude profiles from ISR (incoherent scatter radar), ionosonde electron density profiles (in-situ electron density such as DMSP, CHAMP, SWARM, COSMIC-2), space based observations of electron density [83] (LEO satellite beacon TEC) and topside GNSS TEC (looking upwards from LEO satellites to constellation). In general, the most complete data set is the GNSS TEC, so it is the main ingested quantity of this algorithm. It uses all these measurements to correct a climate background model, both empirical or physical.

In this dissertation, the IDA4D algorithm outputs are used for contributions 1, 3 and 4 (in Chapters 3, 5 and 6 respectively). Two different sources of density  $N$  are used as background model for IDA4D, depending on the cases we are studying. The first one is the IRI (International Reference Ionosphere) [84] background model (Chapters 3 and 6) and the second one is SAMI3 [49] (Chapter 5). The use of a physics model, such as SAMI3 rather than empirical model, will provide for better resolution of the nighttime effects like NILE.

For example, in Chapter 6, the inputs to IDA4D are: Global Positioning System (GPS) slant TEC from 592 stations (each of the stations measured between  $\approx 30000, 40000$  measurements during one day), DMSP (Defense Meteorological Satellite

Program) electron density profiles from 4 satellites (each of the satellites measured  $\approx 80000$  measurements during one day) and GPS radio occultation data of electron density profiles, advantageous for improving the situation of large data gaps over the oceans, from 6 receivers (each of the receivers measured  $\approx 230000$  measurements during one day).

**2.6.2 SAMI3.** The SAMI3 algorithm is a physics-based ionosphere model that calculates the plasma and chemical evolution of seven ion species ( $H^+$ ,  $He^+$ ,  $N^+$ ,  $O^+$ ,  $NO^+$ ,  $N_2^+$  and  $O_2^+$ ). It studies the behavior of the plasma along the Earth's dipole and is based on the two-dimensional model SAMI2 [49]. For neutral driving, the neutral composition and temperature are provided by the Naval Research Laboratory Mass Spectrometer Incoherent Scatter (NRLMSISE-00) model [85] and the neutral winds are obtained from the HWM models [50]. It also uses the [86] model for the ion production and loss and for the gravity and diffusion effect on the parallel direction of the ion velocity, similarly to EMPIRE. The magnetic field is provided by the International Geomagnetic Reference Field (IGRF-11) model [59] and the electric field is solved with a potential solver derived from current conservation [87], [88]. The ion drift is calculated self-consistently with the solved perpendicular electric field using Equation 2.3 in the low to mid-latitude region of the ionosphere. At high latitudes, SAMI3 uses the Weimer [58] model.

**2.6.3 EMPIRE.** In this section, we give a summary of the EMPIRE algorithm, as developed in previous work. It estimates the ionosphere drivers over a region during a specified period of time by combining information from different background models and from measurements. This is useful during unusual ionospheric conditions, because the background models are usually a mean over a long period of time, so they do not predict unusual behavior of the ionosphere. Real measurements do capture geomagnetic storms and unusual behavior, but they are usually scarce, as the

ionosphere region is large.

EMPIRE ingests electron density measurements  $N$  and estimates the IT drivers using the ion continuity on a gridded region over time. At each grid point, characterized by a geomagnetic radius, colatitude and longitude  $(r, \theta, \phi)$ , the ion continuity equation is imposed:

$$\frac{\partial N}{\partial t} = s_{prod} + s_{loss} - \vec{\nabla} \cdot (N\vec{v}) \quad (2.8)$$

where  $s_{prod}$  and  $s_{loss}$  are source terms (production and loss rate, respectively),  $\vec{v}$  is the ion velocity and  $\frac{\partial N}{\partial t}$  represents the electron density rate. Although the electron density  $N$  also varies in space,  $(r, \theta, \phi)$ , Equation 2.8 is imposed at each grid point, so there is only variation with time. The assumption of electron density  $N$  equal to the ion density is made when solving the ion continuity equation, as the plasma is quasi-neutral. The change in ions over time in a grid point is going to depend on the production and loss of ions and on the transport of ions in that region, represented by the divergence term in Equation 2.8.

The electron density rate  $\frac{\partial N}{\partial t}$  is finite differenced from the electron density  $N$ , obtained from the data assimilation algorithm IDA4D (Ionospheric Data Assimilation Four-Dimensional [18]).

To solve the system in Equation 2.8, EMPIRE divides the ion velocity,  $\vec{v}$ , into two directions according to the magnetic coordinate system:  $\parallel$  and  $\perp$  directions, aligned with and perpendicular to the magnetic field line, respectively.

$$\frac{\partial N}{\partial t} = \underbrace{s_{prod}}_{a_{0,prod}} + \underbrace{s_{loss}}_{a_{0,loss}} - \underbrace{\vec{\nabla} \cdot (N\vec{v}_{\perp})}_{a_{exb}} - \underbrace{\vec{\nabla} \cdot (N\vec{v}_{\parallel})}_{a_{\parallel}} \quad (2.9)$$

The notation  $a$  is used to refer the different terms in the ion continuity equation that are not measurements in the system. The subscript “0” denotes that the quantity is assumed to be known based on a background model. The models used to calculate the production and loss terms  $a_{0,prod}, a_{0,loss}$  can be found in [86].

The ion drift, represented with the subscript “*exb*”, is assumed to be the dominant velocity in the perpendicular direction  $\vec{v}_\perp \equiv \vec{v}_{exb}$ , as described in a previous section (Section 2.5). Below the ion drift  $\vec{v}_{exb}$ , defined in Equation 2.3, is repeated but substituting the electric field  $\vec{E}$  for the gradient of an electric potential field  $V$ .

$$\vec{v}_{exb} = \frac{-\nabla V \times \vec{B}_0}{B_0^2} \quad (2.10)$$

This is a reasonable assumption at F region altitudes, described in the previous chapter 2. The ion drift results from crossing the electric field  $\vec{E}$  associated with ion motion with the Earth’s magnetic field  $\vec{B}_0$ . The electric field  $\vec{E}$  is calculated as the gradient of an electric potential  $V$ . The potential  $V$  is one of the IT drivers we actually estimate. The magnetic field  $\vec{B}_0$  is obtained from the International Geomagnetic Reference Field (IGRF-11) model [59].

In the parallel direction we consider the other effects that can contribute to the ion velocity  $\vec{v}_\parallel$ : neutral winds, gravity and diffusion. The last two effects are calculated following the model described in [86]. This model is simplified as explained in [30] to be:

$$\vec{v}_\parallel = \vec{u}_\parallel + \vec{v}_{g,0} + \vec{v}_{dfs,0} \quad (2.11)$$

where  $\vec{u}_\parallel$  represents the neutral wind in the parallel to the magnetic field direction. Neutral winds  $\vec{u}$  is the second driver we are going to estimate with EMPIRE, by projecting the geographic meridional and zonal components,  $u_N$  and  $u_E$  respectively, onto the field-aligned direction  $\hat{b}$  as shown in [45]. The assumption of negligible vertical winds in the vertical direction is made. The velocity due to gravity is represented by the  $\vec{v}_{g,0}$  term and lastly  $\vec{v}_{dfs,0}$  is the diffusion contribution. To calculate these we need to have values of the ion and electron temperatures, obtained from the International Reference Ionosphere (IRI 2007) [84], and of the neutral density and temperature which are obtained from the Naval Research Laboratory Mass Spectrometer Incoherent Scatter (NRL-MSISE00) model [85].

Once the background climate values in EMPIRE are computed, we express Equation 2.9 as:

$$\frac{\partial N}{\partial t} = a_{0,prod} + a_{0,loss} + a_{exb} + \underbrace{a_{u_{\parallel}} + a_{0,dfsn} + a_{0,g}}_{a_{\parallel}} \quad (2.12)$$

Because of the hypothesis of ion drifts being the dominant term in the perpendicular direction, we only consider the neutral wind projection in the parallel direction as contributing to the electron density rate of change:  $a_u = a_{u_{\parallel}}$ .

In general, EMPIRE is configured to compute the neutral winds and the electric potential contributions,  $a_u$  and  $a_{exb}$ , because we currently assume that during storm time these are the main contributions to a change in electron density. This thesis implements a new method to solve the system and to estimate these contributions. It will be described in the Chapter 3.

In previous work, to solve the system, EMPIRE estimated the correction of these contributions. To do that, each ionospheric driver,  $v_i$ , was divided into a unknown correction term and a background model term, where subscript  $i$  represents the ionospheric drivers: neutral wind geographical zonal and meridional  $u_E$  and  $u_N$  and potential  $V$  giving rise to  $\vec{v}_{exb}$ . They can be expressed as:

$$v_i = \delta v_i + v_{i,0} \quad (2.13)$$

Where  $\delta$  indicates correction term, although it does not necessarily needs to be small. Neutral wind background value  $\vec{u}_0$  is obtained from HWM14 climate model [50] and potential field background model term is obtained using Weimer model [58]. Substituting these expressions for each driver in the transport term in Equation 2.12, we can express the ion continuity equation as:

$$\underbrace{\frac{\partial N}{\partial t}}_{y_N} = a_{0,prod} + a_{0,loss} + \underbrace{a_{0,exb} + \delta a_{0,exb}}_{a_{exb}} + \underbrace{a_{0,u_{\parallel}} + \delta a_{0,u_{\parallel}}}_{a_{u_{\parallel}}} + a_{0,dfsn} + a_{0,g} \quad (2.14)$$

The new method allows for EMPIRE to directly estimate the contributions  $a_u$  and  $a_{exb}$  instead of the correction to a background model term  $\delta a_u$  and  $\delta a_{exb}$ . The whole derivation will be shown in Section 3.2.

To estimate the correction terms, each correction of ionospheric driver term,  $\delta V$ ,  $\delta u_N$  and  $\delta u_E$ , is expanded as a multiplication of basis functions  $\mathbf{f}_i$  and a set of unknown coefficients  $\tilde{\mathbf{x}}_i$ , where subscript  $i = exb, u_N, u_E$ . The potential field  $V$  at each grid point  $j$  can be expanded as:

$$\delta V_j = \mathbf{f}_{\mathbf{exb},j} \tilde{\mathbf{x}}_{\mathbf{exb}} \quad (2.15)$$

Where bold symbols indicate an array. Each  $j$ th grid point is given by geomagnetic coordinates altitude, colatitude and longitude  $(r, \theta, \phi)$ . From now on, a tilde notation “ $\tilde{\phantom{x}}$ ” is introduced to refer to terms that change within the old and new method of solving EMPIRE. The mapping matrix  $\mathbf{f}_{exb}$  is obtained expanding the potential field  $\delta V_j$  using a spherical harmonic basis function [45]. Note that this mapping matrix is not going to change in the new method, as the expansion is also done. The expansion is defined as:

$$\delta V_j = \sum_{l=1}^{l_{\max}} \sum_{m=1}^l N_l^m P_l^m(\ell(r, \theta)) \Phi_l^m(\phi, \tilde{\mathbf{x}}_{\mathbf{exb}}) \quad (2.16)$$

where  $N_l^m P_l^m$  is the fully normalized associated Legendre polynomial as described by [89] and  $\ell$  is the normalized L-shell parameter (defined in Equation 2.4), between  $[-1, 1]$ . By assuming constant electric field along the dipole field lines, this allows for the spatial potential field to be a function of only two independent variables  $(L, \phi)$ . The term  $\Phi_l^m$  contains the harmonic term and the unknown coefficients  $\tilde{\mathbf{x}}_{\mathbf{exb}}$ . The final form of mapping matrix  $\mathbf{f}_{\mathbf{exb}}$  is shown in Appendix A.

The neutral winds zonal and meridional component can be expanded as:

$$[\delta u_E \ \delta u_N]^T = \mathbf{f}_{\mathbf{u},j} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (2.17)$$

The meridional direction is expanded using a power series basis at each  $j$  grid point:

$$\delta u_{Nj} = \sum_{k=0}^{k_{\max}} \sum_{l=0}^{l_{\max}} \sum_{p=1}^{p_{\max}} \tilde{x}_N^{klp} \left(\frac{R}{R_e}\right)^k (\theta - \theta_0)^l (\phi)^p \quad (2.18)$$

$$\tilde{\mathbf{x}}_{\mathbf{N}} = \left[ \tilde{x}_N^{000} \quad \tilde{x}_N^{001} \quad \dots \quad \tilde{x}_N^{klp} \quad \dots \quad \tilde{x}_N^{k_{\max}, l_{\max}, p_{\max}} \right]^T \quad (2.19)$$

Where  $\theta_0$  and  $R_e$  are reference colatitude and altitude respectively. A similar expansion to Equation 2.18 is used to define the zonal  $\delta u_{Ej}$  and  $\tilde{x}_E$ . Substituting these expansions into Equation 2.17 and considering negligible vertical winds, mapping matrix  $\mathbf{f}_{\mathbf{u}}$  can be obtained. State  $\tilde{\mathbf{x}}_{\mathbf{u}}$  is obtained by concatenating both directions:

$$\tilde{\mathbf{x}}_{\mathbf{u}} = [\tilde{\mathbf{x}}_E^T \tilde{\mathbf{x}}_{\mathbf{N}}^T]^T. \quad (2.20)$$

One of the contributions is to update this expansion to vector spherical harmonics basis function. This will change the initial form of the expansion and consequently the final values of the mapping matrix  $\mathbf{f}_{\mathbf{u},j}$ . The new derivation will be described in Section 3.1 and Equations 2.18 to 2.20 will be substituted by a new basis function.

Substituting expansions (2.15) and (2.17) in the corresponding correction transport terms  $\delta a_i$ , defined in Equation (2.14), we can express the  $\delta a_i$  terms as a multiplication of basis functions  $\mathbf{h}_i$  and the same set of coefficients  $\tilde{\mathbf{x}}_{\mathbf{i}}$ :

$$\delta a_{exb,j} = \mathbf{h}_{\text{exb},j} \tilde{\mathbf{x}}_{\text{exb}} \quad (2.21)$$

$$\delta a_{u,j} = \mathbf{h}_{\mathbf{u},j} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (2.22)$$

Concatenating each of these terms of the different drivers, we arrive at the mapping matrix  $\mathbf{h}_{\mathbf{N}j}$  and the state of the system  $\tilde{\mathbf{x}}$ , where subscript  $N$  indicates that the equation comes from the ion continuity equation:

$$\underbrace{\tilde{z}_{N,j}}_{y_N - \tilde{a}_{0N}} = \mathbf{h}_{\mathbf{N}j} \tilde{\mathbf{x}} + \tilde{\nu}_j \quad (2.23)$$

where  $\tilde{a}_{0N}$  is the sum of all the background terms and the term  $\tilde{\nu}_j$  is the observation error, that is assumed to be a white sequence with known observation covariance  $\tilde{\mathbf{R}}_N$ . The state  $\tilde{\mathbf{x}}$  is defined as a column vector:

$$\tilde{\mathbf{x}} = [\tilde{\mathbf{x}}_{\text{exb}}^T \tilde{\mathbf{x}}_{\text{u}}^T]^T \quad (2.24)$$

and its size will depend on the number of coefficients of each driver expansion.

Equation 2.23 is imposed over all the  $j$  points in the region at each time step. Stacking them the system can be transformed to:

$$\underbrace{\tilde{\mathbf{z}}_N}_{\mathbf{y}_N - \tilde{\mathbf{a}}_{0N}} = \mathbf{H}_N \tilde{\mathbf{x}} + \tilde{\nu} \quad (2.25)$$

Column vector  $\tilde{\mathbf{z}}$  is obtained by stacking the difference of the electron density rate  $y_N$  and the background terms,  $\tilde{a}_{0N}$  of each grid point and depends on the number of grid points our region is divided into. Mapping matrix  $\mathbf{H}_N$  and column vector  $\tilde{\nu}$  are also derived by stacking the mapping row matrix  $\mathbf{h}_{Nj}$  and the observation error  $\nu_j$  of each grid point. Mapping matrix  $\mathbf{H}_N$  will be formed by mapping matrices of each of the estimated drivers:  $\mathbf{H}_N = [\mathbf{H}_{\text{exb}} \mathbf{H}_{\text{u}}]$ . The derivation and the definition of the ion drift mapping matrix  $\mathbf{H}_{\text{exb}}$  can be found in [45] and a summary is shown in Appendix A. The vector spherical harmonics basis function will be introduced for the neutral wind expansion, so  $\mathbf{H}_{\text{u}}$  will be defined in Section 3.1. The state  $\tilde{\mathbf{x}}$  is the same at all points allowing us to solve the system in the whole region at the same time.

Without additional ingestion of measurements, we define the EMPIRE system as:

$$\underbrace{\tilde{\mathbf{z}}}_{\mathbf{z}_N} = \underbrace{\mathbf{H}}_{\mathbf{H}_N} \tilde{\mathbf{x}} + \tilde{\nu} \quad (2.26)$$

where  $\tilde{\nu}$  is characterized by  $\tilde{\mathbf{R}} = \tilde{\mathbf{R}}_N$  EMPIRE solves the linear system defined in equation 2.26, which is overdetermined, with a Kalman filter that was implemented

in previous work [32]. A time update is first applied to the state, assuming a Gauss-Markov state transition:

$$\hat{\tilde{\mathbf{x}}}_{t+1/t} = \tilde{\mathbf{x}}_{0,t+1/t} + \mathbf{\Phi}(\hat{\tilde{\mathbf{x}}}_{t/t} - \tilde{\mathbf{x}}_{0,t/t}) + \tilde{w}_{t/t} \quad (2.27)$$

Where the subscript “ $t/t$ ” indicates values in the time  $t$  given  $t$  and the subscript “ $t + 1/t$ ” describes the time updated value at time  $t + 1$  given  $t$  using information from the previous time step.  $\hat{\tilde{x}}_{t+1/t}$  is the forecast state at time  $t + 1$  and  $\hat{\tilde{x}}_{t/t}$  is the measurement update at previous time  $t$ . The hat  $\hat{\phantom{x}}$  notation indicates estimated state.  $\tilde{x}_{0,t+1/t}$  is the background state at time  $t + 1$  and  $\tilde{x}_{0,t/t}$  is the background state at time  $t$ . It is set, for both times to  $\tilde{\mathbf{x}}_0 = \mathbf{0}$ , which assumes that the initial point for the filter is that there is no correction to the drivers. The process noise  $\tilde{w}_{t/t}$  is assumed to be normally distributed with zero mean and covariance  $\tilde{\mathbf{W}}_t$ . Matrix  $\mathbf{\Phi}$  is defined as:

$$\mathbf{\Phi} = \exp\left(\frac{-dt}{\tau}\right) \mathbf{I} \quad (2.28)$$

where  $\mathbf{I}$  is the identity matrix and  $\tau$  is a time constant that quantifies how rapidly measurements are “forgotten” and  $dt$  is the time step used in the filter. The time update co- variance is given by:

$$\tilde{\mathbf{P}}_{t+1/t} = \mathbf{\Phi} \tilde{\mathbf{P}}_{t/t} \mathbf{\Phi}^T + \underbrace{(1 - \mathbf{\Phi})^2 \tilde{\mathbf{P}}_{0,t/t}}_{\tilde{\mathbf{W}}_t} \quad (2.29)$$

Where  $\tilde{P}_{t+1/t}$  is the forecast covariance matrix of the state at time  $t + 1$ ,  $\tilde{P}_{t/t}$  is the forecast covariance matrix of the state at time  $t$  and  $\tilde{P}_{0,t/t}$  is the background covariance matrix at time  $t$ . The process noise covariance  $\tilde{W}_t$  is defined so that the time updated covariance  $\tilde{P}_{t+1/t}$  reverts to the background model covariance  $\tilde{P}_{0,t/t}$  in the absence of measurements over time. The background error is defined in [32], but in this work we will change it and will be defined in Chapter 3.

Then, the state is updated using the ingested measurements of electron density:

$$\hat{\tilde{\mathbf{x}}}_{t/t} = \hat{\tilde{\mathbf{x}}}_{t/t-1} + (\mathbf{H}^T \tilde{\mathbf{R}}^{-1} \mathbf{H} + \tilde{\mathbf{P}}_{t/t-1}^{-1})^{-1} \mathbf{H}^T \tilde{\mathbf{R}}^{-1} (\tilde{\mathbf{z}} - \mathbf{H} \hat{\tilde{\mathbf{x}}}_{t/t-1}) \quad (2.30)$$

$$\tilde{\mathbf{P}}_{t/t} = \tilde{\mathbf{P}}_{t/t-1} - \tilde{\mathbf{P}}_{t/t-1} \mathbf{H}^T (\mathbf{H} \tilde{\mathbf{P}}_{t/t-1} \mathbf{H}^T + \tilde{\mathbf{R}}^{-1}) \mathbf{H} \tilde{\mathbf{P}}_{t/t-1}^T \quad (2.31)$$

where  $\tilde{\mathbf{R}}$  contains information about the error covariance  $\nu$  of the measurements  $\partial N/\partial t$ , obtained from the propagation of the error of the plasma density  $N$  given by IDA4D algorithm [18] and from the error of the background ionospheric effects  $\tilde{\mathbf{a}}_0$  production, loss, diffusion, gravity, neutral winds and ion drift.

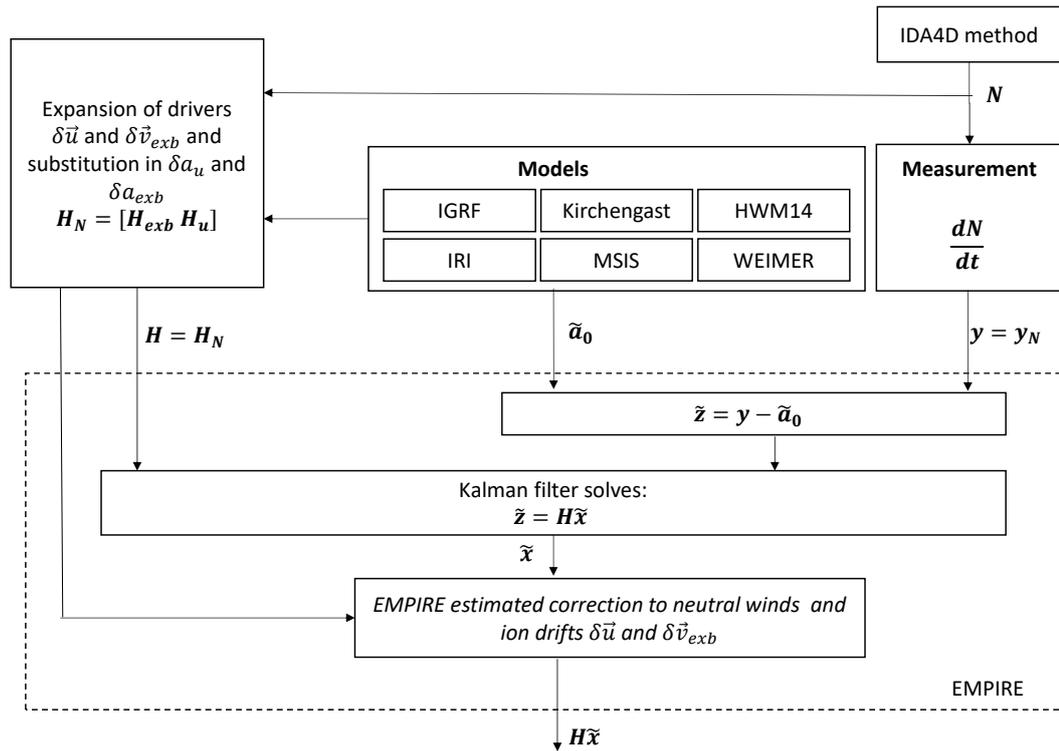


Figure 2.10. Simplified EMPIRE model and data processing flow chart.

A summary of the algorithm is shown in Figure 2.10, with a flow chart describing how all these background models and the electron density measurements enter into EMPIRE. It is based on a diagram shown in [32].

**2.6.3.1 Ingestion of winds.** EMPIRE also ingests measurements of neutral winds from FPI (Fabry-Perot interferometer) instruments [32]. They measure the line-of-sight (LOS) thermospheric wind speeds with associated uncertainties at an assumed

altitude of 250 km, as described in Section 2.2.2.

To include the information in the algorithm, additional equations in every  $i$ th location where a measurement is available are included. These equations are added to the original EMPIRE system defined in Equation 2.26. Note that  $i$  does not need to coincide with the number of grid points  $j$  at which we are imposing the ion continuity equation. They are expressed as:

$$y_{FPI} - \tilde{u}_{0,FPI} = \delta u_{FPI,i} \quad (2.32)$$

Where  $y_{FPI}$  represents the LOS neutral wind measurement from FPI instruments at each  $i$  location and the term  $\tilde{u}_{0,FPI}$  represents the background model neutral winds projected in the LOS direction at each  $i$  location. To calculate it, the neutral wind geographical zonal and meridional components from point quantity vector  $\vec{u}_0 = [u_{0,E} \ u_{0,N} \ 0]^T$  from model HWM14 [50] are projected into the LOS of the measurement. The look direction of each  $i$  LOS is characterized by elevation  $el_i$  and azimuth  $az_i$  angle:

$$\tilde{u}_{0,FPI} = \mathbf{f}_{\mathbf{LOS},i} \vec{u}_0 \quad (2.33)$$

$$\mathbf{f}_{\mathbf{LOS},i} = \begin{bmatrix} \cos(el_i) \cos(az_i) & \cos(el_i) \sin(az_i) & \sin(el_i) \end{bmatrix} \quad (2.34)$$

The correction term  $\delta u_{FPI,i}$  is obtained by transforming the basis function expansion of the neutral winds into the LOS direction. Using the defined expansion in Equation 2.17 we obtain the following expression:

$$\delta u_{FPI,i} = \mathbf{f}_{\mathbf{LOS},i} \begin{bmatrix} \delta u_E & \delta u_N & 0 \end{bmatrix}^T = \mathbf{h}_{\mathbf{FPI},i} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (2.35)$$

Mapping matrix  $\mathbf{h}_{\mathbf{FPI},i}$  is going to depend on the expansion used for the neutral winds. One of the contributions is to update this expansion to a vector spherical harmonics basis function. Definition of mapping matrix  $\mathbf{h}_{\mathbf{FPI},i}$  for this case will be described in Section 5.1.

Using the same notation as used in the EMPIRE system before, in Equation 2.23, we can write Equation 2.32 as the following:

$$\tilde{z}_{FPI,i} = y_{FPI} - \underbrace{\tilde{a}_{0,FPI}}_{\tilde{u}_{0,FPI}} = \underbrace{\delta a_{FPI,i}}_{\delta u_{FPI,i}} + \nu_{FPI,i} \quad (2.36)$$

Where  $\nu_{FPI,i}$  represents the measurement error. Stacking all the  $i$  equations and substituting the correction term defined in Equation 2.35:

$$\tilde{\mathbf{z}}_{\mathbf{FPI}} = \mathbf{H}_{\mathbf{FPI}} \tilde{\mathbf{x}}_{\mathbf{u}} + \nu_{\mathbf{FPI}} \quad (2.37)$$

Column matrix  $\tilde{\mathbf{z}}_{\mathbf{FPI}}$  size is going to be equal to the number of  $i$  measurements ingested. Mapping matrix  $\mathbf{H}_{\mathbf{FPI}}$  is obtained by stacking the  $i$  row matrices  $\mathbf{h}_{\mathbf{FPI},i}$ . Measurement error  $\nu_{\mathbf{FPI}}$  is assumed to have zero mean and the associated covariance matrix  $\mathbf{R}_{\mathbf{FPI}}$  is the error of the neutral winds given with the measurements.

The EMPIRE system, defined in Equation 2.23, adding the ingestion of FPI measurements is now:

$$\underbrace{\begin{bmatrix} \tilde{\mathbf{z}}_{\mathbf{N}} \\ \tilde{\mathbf{z}}_{\mathbf{FPI}} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} \mathbf{y}_{\mathbf{N}} \\ \mathbf{y}_{\mathbf{FPI}} \end{bmatrix}}_{\mathbf{y}} - \underbrace{\begin{bmatrix} \tilde{\mathbf{a}}_{0\mathbf{N}} \\ \tilde{\mathbf{a}}_{0,\mathbf{FPI}} \end{bmatrix}}_{\tilde{\mathbf{a}}_0} = \underbrace{\begin{bmatrix} \mathbf{H}_{\text{exb}} & \mathbf{H}_{\mathbf{u}} \\ \mathbf{0} & \mathbf{H}_{\mathbf{FPI}} \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} \tilde{\mathbf{x}}_{\text{exb}} \\ \tilde{\mathbf{x}}_{\mathbf{u}} \end{bmatrix} + \begin{bmatrix} \tilde{\nu} \\ \nu_{\mathbf{FPI}} \end{bmatrix} \quad (2.38)$$

The system has the same form as with no ingestion, as seen in Equation 2.26, and it is solved with a Kalman filter using Equations 2.27 to 2.31. It needs to be considered that there are new error measurements  $\tilde{\nu}_{\mathbf{FPI}}$  in the definition of covariance matrix of the measurements  $\tilde{\mathbf{R}}$ , that appears in the measurement update of the filter. The EMPIRE system simplified system adding the ingestion of winds is shown in the diagram in Figure 2.11.

**2.6.3.2 EMPIRE inputs.** A summary of the inputs used in this work organized by chapters is shown below.

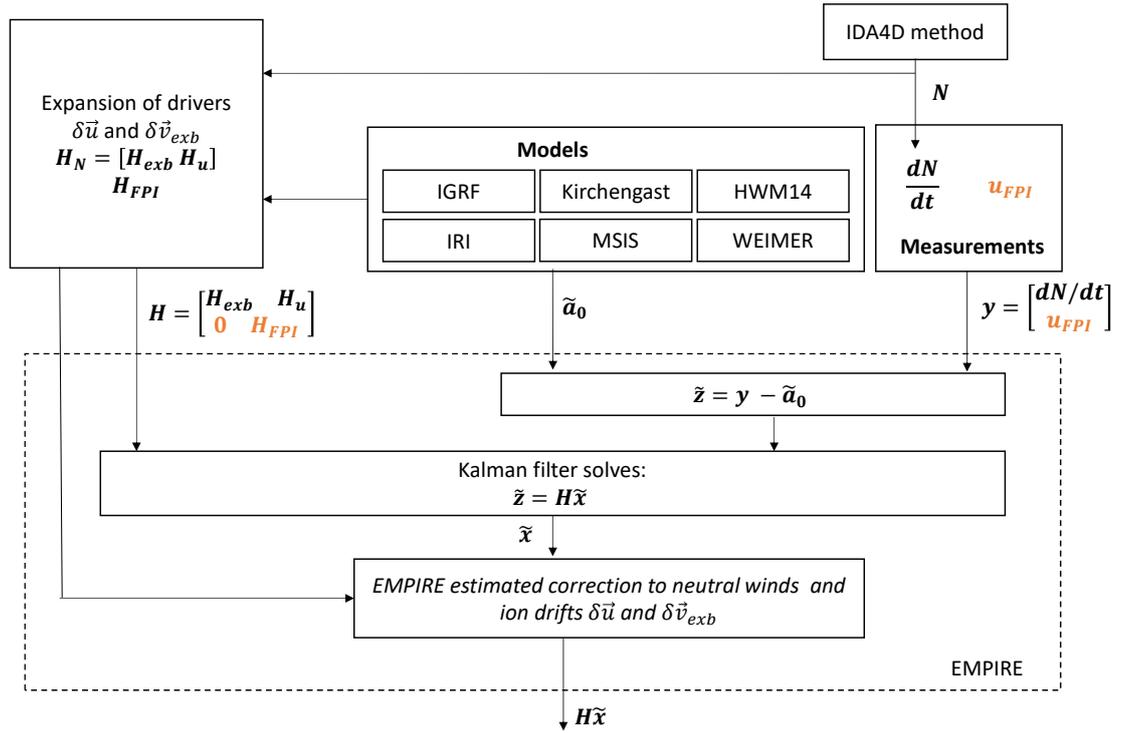


Figure 2.11. Simplified EMPIRE model and data processing flow chart.

Table 2.1. Inputs to EMPIRE, climate models and measurements, utilized in this dissertation organized by chapters.

	Chapter 3	Chapter 4	Chapter 5	Chapter 6
N observation	IDA4D	SAMI3	IDA4D	IDA4D
N background	IRI	SAMI3	SAMI3	IRI
FPI observation	-	-	FPI	ICON
Ion and electron temperature			IRI	
Neutral temperature and density			NRL-MSISE00	
Neutral winds background			HWM14	
Potential field background	Weimer	SAMI3	Weimer	Weimer
Magnetic field			IGRF-11	

## CHAPTER 3

### VECTOR SPHERICAL HARMONICS FOR EMPIRE NEUTRAL WIND ESTIMATION

EMPIRE estimates each of the ionosphere drivers by expanding them using basis functions. Previously, power series basis functions were used to represent both meridional and zonal neutral wind components  $u_N$  and  $u_E$  [29], allowing the study of regional cases. The use of expansions that are continuous on a spherical domain allows EMPIRE to be used to study global cases. A vector spherical harmonics (from now on abbreviated as “VSH”) expansion will be introduced in this contribution in Section 3.1 to decompose the velocity, as defined in [90]. In addition, a direct estimation of the drivers will be derived in Section 3.2, in order to define the errors of the Kalman filter in a more methodical way.

Then, the VSH derivation will be applied to study a geomagnetic storm and the results will be compared to the previous method used in EMPIRE. The storm and the EMPIRE configuration will be described and results and conclusions will be shown in Section 3.3.

#### 3.1 Vector spherical harmonics derivation

In a power series basis expansion, there were two states for the neutral winds, one for each direction, as they were calculated independently of each other. However, for a VSH decomposition, the state  $\tilde{\mathbf{x}}_{\mathbf{u}}$  is the same for both directions, so they will be consistent with each other. This derivation has two parts: the definition of the vector  $\vec{u}$  and the mapping from velocity space to the  $a$  (density rate) space.

The vector spherical harmonic form is the following:

$$\vec{u} = \sum_{l=0}^{l_{max}} \sum_{m=0}^l Y_{lm} \hat{r} + r \vec{\nabla} Y_{lm} + \vec{r} \times \vec{\nabla} Y_{lm} \quad (3.1)$$

where  $\vec{r} = r\hat{r}$  is the radial magnitude and direction [91]. The expansion is defined by

the degree  $l$  and by the order  $m$ , that goes to a maximum degree  $l_{max}$  that we choose.

The function  $Y_{lm}$  is the scalar spherical harmonics defined as:

$$Y_{lm} = N_l^m P_l^m(\cos \theta) \Phi_l^m(\phi) \quad (3.2)$$

with  $\theta$  being the geomagnetic colatitude and  $\phi$  the geomagnetic longitude. We expand the neutral wind using magnetic coordinates to be consistent with the EMPIRE grid. The function  $N_l^m P_l^m$  describes the fully normalized associated Legendre polynomial, described in [89] as:

$$N_l^m = \sqrt{\frac{2l+1}{2} \frac{(l-m)!}{(l+m)!}} \quad (3.3)$$

$$P_l^m(\cos \theta) = \frac{(-1)^m}{2^l l!} \sqrt{(1-\cos^2 \theta)^m \frac{d^{l+m}}{d(\cos \theta)^{l+m}} (\cos^2 \theta - 1)^l} \quad (3.4)$$

The harmonic part can be found in the  $\Phi_l^m(\phi)$  function:

$$\Phi_l^m(\phi) = \tilde{x}_s^{lm} \sin(m\phi) + \tilde{x}_c^{lm} \cos(m\phi) \quad (3.5)$$

where  $\tilde{x}_c^{lm}$  and  $\tilde{x}_s^{lm}$  are the unknown coefficients that form the state  $\tilde{\mathbf{x}}_{\mathbf{u}}$  for each combination of  $l, m$ . The size of the state will depend on the maximum order  $l_{max}$  we select for the expansion. The state  $\tilde{\mathbf{x}}_{\mathbf{u}}$  can be expressed as:

$$\tilde{\mathbf{x}}_{\mathbf{u}} = [\tilde{x}_c^{00} \tilde{x}_c^{10} \dots \tilde{x}_c^{l_{max} l_{max}} \tilde{x}_s^{00} \tilde{x}_s^{10} \dots \tilde{x}_s^{l_{max} l_{max}}]^T \quad (3.6)$$

$$= [\tilde{\mathbf{x}}_c^T \tilde{\mathbf{x}}_s^T]^T \quad (3.7)$$

Equations 2.18, 2.19 and 2.20 are now substituted by the new expansion given by Equations 3.1, 3.2, 3.3, 3.4, 3.5, 3.6 and 3.7.

The second part of the derivation is to map the neutral winds to the electron density rate “ $a$ ” space. This is done by taking the divergence of the neutral wind

vector projection in the magnetic field direction multiplied by the electron density, as described in the following equation 3.8:

$$\delta a_{u,j} = -\vec{\nabla} \cdot (N \underbrace{(\delta \vec{u} \cdot \hat{b}) \hat{b}}_{\delta \vec{u}_{\parallel}}) = \mathbf{h}_{\mathbf{u},j} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (3.8)$$

Stacking the row matrix  $\mathbf{h}_{\mathbf{u},j}$  at each  $j$ th grid point, as explained in Section 2.6.3, we obtain the mapping matrix  $\mathbf{H}_{\mathbf{u}}$ , that is defined in Appendix B:

$$\delta \mathbf{a}_{\mathbf{u}} = \mathbf{H}_{\mathbf{u}} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (3.9)$$

The correction of the horizontal wind vector  $\delta \vec{u} = [u_{E,j} \ u_{N,j} \ 0]$  at each grid point  $j$ th grid point can be expressed in the geographical coordinate system  $(\hat{e}, \hat{n}, \hat{u})$ , with  $\hat{e}$  representing the geographic east unit vector and  $\hat{n}$  the geographic north, as defined in Equation 2.17. It can be divided into:

$$\delta \vec{u}_j = \begin{bmatrix} \mathbf{f}_{\mathbf{uE},j} \\ \mathbf{f}_{\mathbf{uN},j} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (3.10)$$

Stacking the row matrices  $\mathbf{f}_{\mathbf{uE},j}$  and  $\mathbf{f}_{\mathbf{uN},j}$  at each  $j$ th grid point, we can obtain the neutral wind vector of the whole studied region:

$$\delta \mathbf{u}_{\mathbf{E}} = \mathbf{F}_{\mathbf{uE}} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (3.11)$$

$$\delta \mathbf{u}_{\mathbf{N}} = \mathbf{F}_{\mathbf{uN}} \tilde{\mathbf{x}}_{\mathbf{u}} \quad (3.12)$$

The mapping matrices  $\mathbf{F}_{\mathbf{uN}}$  and  $\mathbf{F}_{\mathbf{uE}}$  are defined in Appendix C.

**3.1.1 Selection of  $l_{\max}$ .** The number of terms of the decomposition is limited by the maximum order selected  $l_{\max}$ , which appears in 3.1. The number of coefficients of the expansion  $k_u$  is equal to:

$$k_u = (2 \sum_{l=0}^{l=l_{\max}} (l+1)) \quad (3.13)$$

Where the 2 appears because the state is divided into coefficients for the cosine terms and for the sine terms. The  $l, m$  terms that multiply “0” mapping matrix coefficients will be removed to avoid singularities when inverting the mapping matrix. Using Equations B.39 and B.43, we need to take into account that the cosine coefficients multiply the harmonic cosine term ( $\cos m\phi$ ) and its derivative ( $-m \sin m\phi$ ) while the sine coefficients multiply the harmonic sine term ( $\sin m\phi$ ) and its derivative ( $m \cos m\phi$ ). For the sine terms for all the  $m = 0$  coefficients the corresponding expansion value is also 0, so they are not going to be estimated. For the cosine coefficients, for  $m = 0$  only the derivative part that they multiply in the expansion is 0. However, when the order is also  $l = 0$ , the Legendre polynomial and its derivatives is 0, so the corresponding coefficient is not estimated for  $l = m = 0$ . Number of coefficients is reduced from  $k_u$  to  $(k_u - 1 - \sum_{l=0}^{l=l_{\max}} l)$ .

A study of the error of the VSH fitting to HWM14 winds over the maximum order is done to select the  $l_{\max}$ . We use Equations 3.11 and 3.12 at a specific time to do the fitting. The time selected is the first time step of the studied period, 25 October of 2011 at 0 UT. State  $\tilde{x}_u$  is estimated with a Moore-Penrose pseudoinverse [92]. Figure 3.1 shows the analysis of the error of the VSH fitting.

Figure 3.1 shows the RMS (Root-Mean-Square) of the residual between the fitted  $u_N$  and  $u_E$  terms and the model HWM14 over all the grid points at 0UT on October 25. They are plotted over  $l_{\max}$ , to analyze which order to chose for the VSH expansion. The RMS difference tends to go to an asymptotic value of  $115m/s$  for the zonal component in magenta and a value of  $80m/s$  for the meridional term in magenta from approximately  $l_{\max} = 14$  and  $l_{\max} = 10$  respectively. The number of the unknown coefficients increases with  $l_{\max}$ , so there is a compromise between the order of the VSH expansion and the computational time needed. We select for the fitting a value of  $l_{\max} = 6$ , indicated with a black vertical dashed line. At this value,

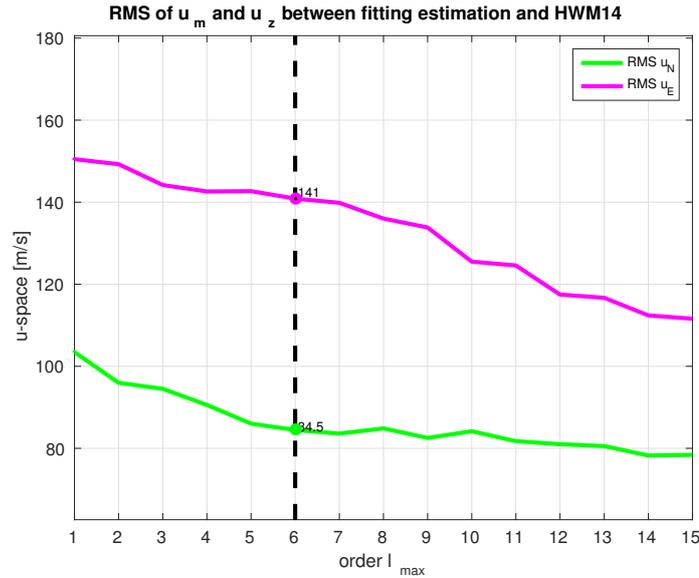


Figure 3.1. RMS (Root-Mean-Square) difference in  $m/s$  between estimated  $u_N$ , in green, and  $u_E$ , in magenta, and HWM14 model over order  $l_{\max}$  over all grid points at 0 UT on 25 October 2011. Black dashed line represents the values at  $l_{\max} = 6$ .

a decrease in the slope of the RMS curve is visible for the meridional RMS in green and the RMS value starts to stabilize around the asymptote.

**3.2 Direct estimation of ionospheric drivers** In previous work, the state  $\tilde{\mathbf{x}}$  defined in Equation 2.24 express the correction of the drivers with respect to a background model. This implied that the background covariance matrix  $\tilde{\mathbf{P}}_{0,t/t}$ , defined in the Kalman filter in Equation 2.29, contained the errors of the correction to the background model, not of the ionospheric drivers themselves. These errors are harder to understand physically and to set, so we have modified EMPIRE to directly estimate the density rate  $a_i$  due to a given driver, where  $i$  represents neutral winds or electric potential, instead of the correction  $\delta a_i$ .

First, the drivers itself,  $v_i$ , are expanded using a basis function. Equations 2.15 and 2.17 are now:

$$V_j = \mathbf{f}_{\text{exb},j} \mathbf{X}_{\text{exb}} \quad (3.14)$$

$$\vec{u}_j = \mathbf{f}_{\mathbf{u},j} \mathbf{X}_{\mathbf{u}} \quad (3.15)$$

Where the mapping matrices  $\mathbf{f}_{\text{exb},j}$  and  $\mathbf{f}_{\mathbf{u},j}$  are the same as the ones defined for the previous method. Substituting these expansions in their corresponding transport terms  $a_i$ , defined in Equation 2.12, we obtain:

$$a_{\text{exb},j} = \mathbf{h}_{\text{exb},j} \mathbf{x}_{\text{exb}} \quad (3.16)$$

$$a_{\mathbf{u},j} = \mathbf{h}_{\mathbf{u},j} \mathbf{x}_{\mathbf{u}} \quad (3.17)$$

Concatenating each of these terms, as done for the previous method, we arrive at the mapping matrix  $\mathbf{h}_j$ , that is the same as the old method, and the new state of the system  $\mathbf{x}$ :

$$\underbrace{z_{Nj}}_{\frac{\partial N}{\partial t} - a_{0N}} = \mathbf{h}_{Nj} \mathbf{x} + \nu_j \quad (3.18)$$

where  $a_0$  is the sum of all the background terms not corrected (production, loss, diffusion and gravity) and  $\nu_j$  is the observation error of  $z_{Nj}$ . The state  $\mathbf{x}$  is defined as a column vector:

$$\tilde{\mathbf{x}} = [\mathbf{x}_{\text{exb}}^T \ \mathbf{x}_{\mathbf{u}}^T]^T \quad (3.19)$$

and its size will depend on the number of coefficients of each driver expansion.

Again, we impose Equation 3.18 over all the  $j$  points in the region at each time step and they are stacked obtaining the following system:

$$\underbrace{\mathbf{z}_N}_{\frac{\partial N}{\partial t} - \mathbf{a}_{0N}} = \mathbf{H}_N \mathbf{x} + \nu \quad (3.20)$$

Mapping matrix  $\mathbf{H}_N$  is the same as the old method and observation error vector  $\nu$  is assumed to have zero mean and covariance matrix  $\mathbf{R}_N$ , given by the measurement errors of the terms forming  $\mathbf{z}_N$ . Without additional ingestion of measurements, we can define the EMPIRE system as:

$$\underbrace{\mathbf{z}}_{\mathbf{z}_N} = \underbrace{\mathbf{H}}_{\mathbf{H}_N} \mathbf{x} + \nu \quad (3.21)$$

The state  $\mathbf{x}$  is solved with a Kalman filter. Equations 2.27 to 2.31 are used, but removing the  $\sim$  notation. They are repeated below:

$$\hat{\mathbf{x}}_{t+1/t} = \mathbf{x}_{0,t+1/t} + \Phi(\hat{\mathbf{x}}_{t/t} - \mathbf{x}_{0,t/t}) + \mathbf{w}_{t/t} \quad (3.22)$$

$$\mathbf{P}_{t+1/t} = \Phi \mathbf{P}_{t/t} \Phi^T + \underbrace{(1 - \Phi)^2 \mathbf{P}_{0,t/t}}_{\mathbf{w}_t} \quad (3.23)$$

The background state at time  $t+1$  and  $t$  given  $t$ ,  $\mathbf{x}_{0,t/t}$  and  $\mathbf{x}_{0,t+1/t}$ , in the new implementation is not equal to 0, as it was in the previous implementation, but it can be calculated using climate models. This new implementation also allows for the direct calculation of the corresponding background covariance matrix  $\mathbf{P}_{0,t/t}$  in terms of the climate data and their errors.

We calculate  $\mathbf{x}_0$ , in time  $t+1$  or  $t$ , using the definition of the state from Equation 3.19. The state is formed by a contribution by the potential field and another by the neutral wind vector. Using the described expansions of the drivers, Equations 3.14 and 3.15, we can calculate the background state for each driver. Stacking the expansions for all the  $j$  grid points:

$$\mathbf{V} = \mathbf{F}_{\text{exb}} \mathbf{x}_{\text{exb}} \quad (3.24)$$

$$\vec{\mathbf{u}} = \mathbf{F}_{\mathbf{u}} \mathbf{x}_{\mathbf{u}} \quad (3.25)$$

Applying these definitions to the background model case, where the drivers are known information obtained from climate models, we obtain the following:

$$\mathbf{x}_{0,exb} = (\mathbf{F}_{\text{exb}})^\dagger \mathbf{V}_0 \quad (3.26)$$

$$\mathbf{x}_{0,u} = (\mathbf{F}_u)^\dagger \vec{\mathbf{u}}_0 \quad (3.27)$$

The matrices equations are solved with a Moore-Penrose pseudoinverse [92], represented with a  $\dagger$ . The climate data for the neutral winds  $\vec{\mathbf{u}}_0$  is obtained from HWM14

model (Horizontal Wind Model) [50] and for the electric potential  $\mathbf{V}_0$  the Weimer model is used [58].

The background covariance matrix  $\mathbf{P}_{0,t/t}$  is defined as:

$$\mathbf{P}_{0,t/t} = \text{diag}(\sigma_{x_{j,0}}^2) \quad (3.28)$$

where  $j$  represents the drivers we are estimating  $j = \{V, u\}$ . Variance  $\sigma_{x_{j,0}}^2$  represents a matrix of the variances in the  $j$  “state space”, e.g., having units of m/s for neutral wind, transformed to state space. They can be expressed as a function of the error in the “velocity” space using the drivers expansions from Equations 3.24 and 3.25. We can obtain a mapping matrix for each  $i$  velocity component we estimate, where  $i$  represents the velocity components of the ion drifts in magnetic coordinates and of the neutral winds in geographical coordinates  $i = \{v_{exb,r}, v_{exb,\theta}, v_{exb,\phi}, u_E, u_N\}$ :

$$\mathbf{v}_i = \mathbf{F}_i \mathbf{x}_i \quad (3.29)$$

where the mapping matrices  $\mathbf{F}_i$  of the ion drift components are defined in [45], and a summary is given in Appendix A, and the matrices of the neutral winds can be found in Appendix C. Then, the errors of each  $j$  studied term in state space  $\sigma_{x_{j,0}}^2$  are defined as:

$$\sigma_{x_{j,0}}^2 = \mathbf{F}_j^\dagger \text{diag}(\sigma_{j,0}^2) (\mathbf{F}_j^\dagger)^T \quad (3.30)$$

where  $\sigma_{j,0}^2$  represent the variances of the background terms studied in the velocity space. Each of the studied terms  $j$  is composed by 3 (ion drift case) or 2 (neutral wind case)  $i$  velocity components.  $\mathbf{F}_j$  is obtained by stacking the corresponding  $\mathbf{F}_i$  matrices, from Equation 3.29, and  $\text{diag}(\sigma_{j,0}^2)$  is a diagonal matrix in which the diagonal is formed by the corresponding ( $\sigma_{i,0}^2$ ) background errors of each velocity component. The division of  $j$  and  $i$  terms is needed because there is only one set of coefficients for each  $j$  term studied, but the  $\mathbf{F}_i$  mapping matrices are defined for each velocity

component  $i$ . The background variance of each velocity component  $i$  is defined as:

$$\sigma_{i,0}^2 = (\sigma_{i,climate}^2 + \sigma_{i,fitting}^2) \quad (3.31)$$

where both the error of the climate model used  $\sigma_{i,climate}$  and the error of the fitting used to expand the drivers  $\sigma_{i,fitting}$  for each of the drivers  $i$  are taken into account and are assumed to be independent. Parameter  $c$  is a scaling factor.

Them, the time update step of the Kalman filter with the new implementation is:

$$\hat{\mathbf{x}}_{t/t} = \hat{\mathbf{x}}_{t/t-1} + (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H} + \mathbf{P}_{t/t-1}^{-1})^{-1} \mathbf{H}^T \mathbf{R}^{-1} (\mathbf{z} - \mathbf{H} \hat{\mathbf{x}}_{t/t-1}) \quad (3.32)$$

$$\mathbf{P}_{t/t} = \mathbf{P}_{t/t-1} - \mathbf{P}_{t/t-1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{t/t-1} \mathbf{H}^T + \mathbf{R}^{-1})^{-1} \mathbf{H} \mathbf{P}_{t/t-1} \quad (3.33)$$

Another difference with the previous case is that matrix  $\mathbf{R}$  defined in 3.33 contains information about the error covariance of the measurements  $\partial N / \partial t$ , obtained from the propagation of the error of the plasma density  $N$  given by IDA4D algorithm [18] but also about the error of the ionospheric effects we are not estimating: the production, loss, diffusion and gravity ( $\mathbf{a}_0$ ). We hypothesize the covariance matrix  $\mathbf{R}_{\mathbf{a}_0}$  of those background model terms that are not modeled and the error covariance matrix  $\mathbf{R}_{\mathbf{N}}$  of the measurements  $N$  to be independent, so that:  $\mathbf{R} = \mathbf{R}_{\mathbf{N}} + \mathbf{R}_{\mathbf{a}_0}$ .

Figure 3.2 shows the new EMPIRE configuration simplified in a data processing flow chart.

In comparison to the old method, shown in diagram from Figure 2.10, the drivers are directly estimated with the Kalman filter. Models HWM14 and Weimer, used before to calculate the background transport terms of the IT drivers, neutral winds  $a_{0,u}$  and ion drifts  $a_{0,exb}$  respectively, are no longer needed to calculate the  $\mathbf{z}$  array.

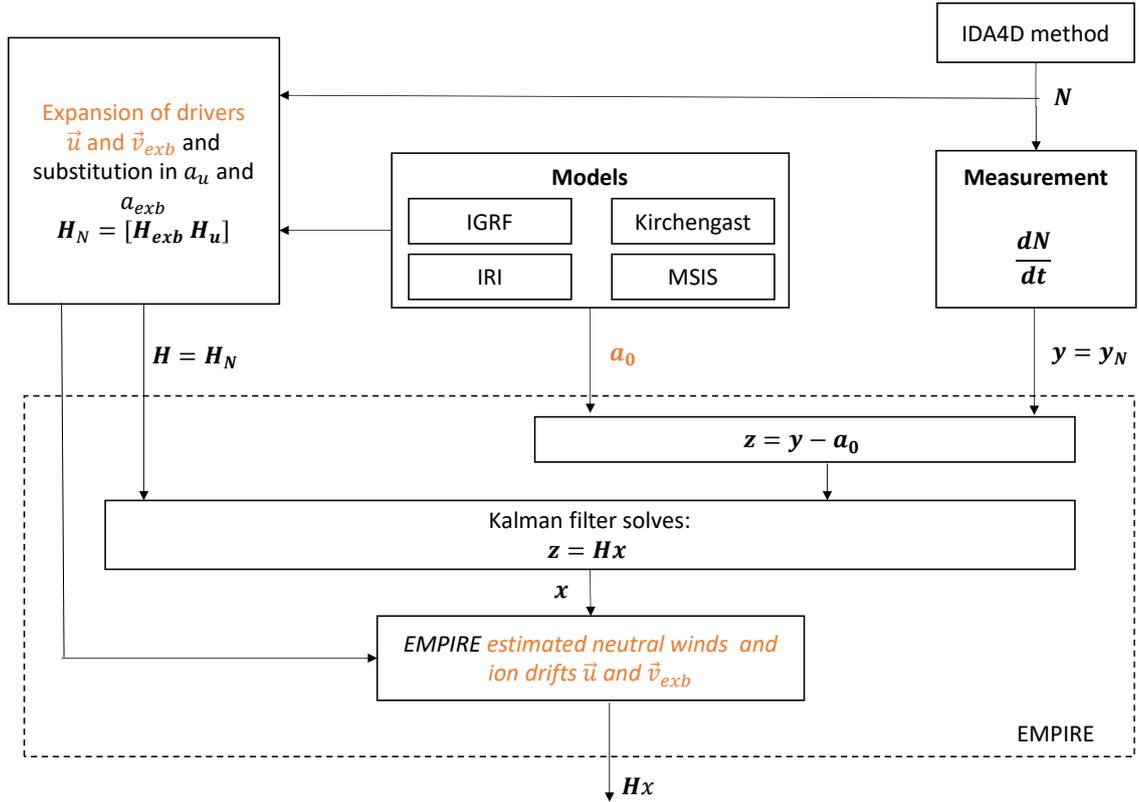


Figure 3.2. New EMPIRE model with direct estimation of IT drivers described with data processing flow chart.

In addition, to solve the system, we normalize the state by doing a non-dimensionalization on Equation 3.21. A regularization matrix  $\mathbf{L}$  is defined to do this transformation. It is a square matrix with size equal to the length of the state vector ( $k_{exb} + k_u$ ). The number of coefficients of each of the expansions,  $k_{exb}$  for the potential field  $V$  and  $k_u$  for the neutral winds  $\vec{u}$ , are defined by the maximum order chosen in the drivers expansion. This step is done to assure full rank of the mapping matrix  $\mathbf{H}$  and avoid singularities when estimating the state  $\mathbf{x}$ :

$$\mathbf{z} = \underbrace{[\mathbf{HL}]}_{\tilde{\mathbf{H}}} \underbrace{[\mathbf{L}^{-1}\mathbf{x}]}_{\tilde{\mathbf{x}}} + \nu \quad (3.34)$$

$$\mathbf{L} = \text{diag}(l_1^{exb}, \dots, l_{k_{exb}}^{exb}, l_1^u, \dots, l_{k_u}^u) \quad (3.35)$$

where  $\bar{\mathbf{H}}$  is the normalized mapping matrix and  $\bar{\mathbf{x}}$  is the normalized state. Non-normalized  $\mathbf{H}$  and  $\mathbf{x}$  are substituted by  $\bar{\mathbf{H}}$  and  $\bar{\mathbf{x}}$  throughout Equations 3.21 to 3.33. For the neutral winds the order of magnitude selected to normalize is  $l^u = 10m/s$  for all the  $k_u$  coefficients and for the potential field is  $l^{exb} = 1kV$  for all the  $k_{exb}$  coefficients. After solving the system, the not normalized state  $\hat{\mathbf{x}}_{t/t}$  and its corresponding variance  $\mathbf{P}_{t/t}$  are calculated and saved, as they are used to calculate the IT drivers using Equation 3.29.

### 3.3 Application of VSH to study geomagnetic storm and comparison to previous method

The purpose of the test described here is to examine how the vector basis improves the neutral wind estimation with respect the scalar basis used previously in EMPIRE. We examine one storm period and estimate the neutral winds globally using the previously-existing power series and using VSH. We compare the estimated winds for both cases over longitude to check the continuity over the globe. Then, we validate the results comparing the estimate to the measurements made at three sites within the estimation region. To compare the results quantitatively we also calculate the root-mean-square of the difference between the estimation and the measurements.

**3.3.1 Method.** In this section the storm analyzed and the EMPIRE configuration for both cases are described.

**3.3.1.1 Storm.** The geomagnetic storm on 25 October 2011 is analyzed. One of the reasons it was chosen is because a previous analysis using the power series was already done in [32]. Also, there were three available sites over the globe with neutral winds measurements to compare the estimates to.

The Dst index, that is an indicator of the magnetic activity, is shown in Figure 3.3 (a) over time. Following the classification given by [64], we consider Dst index

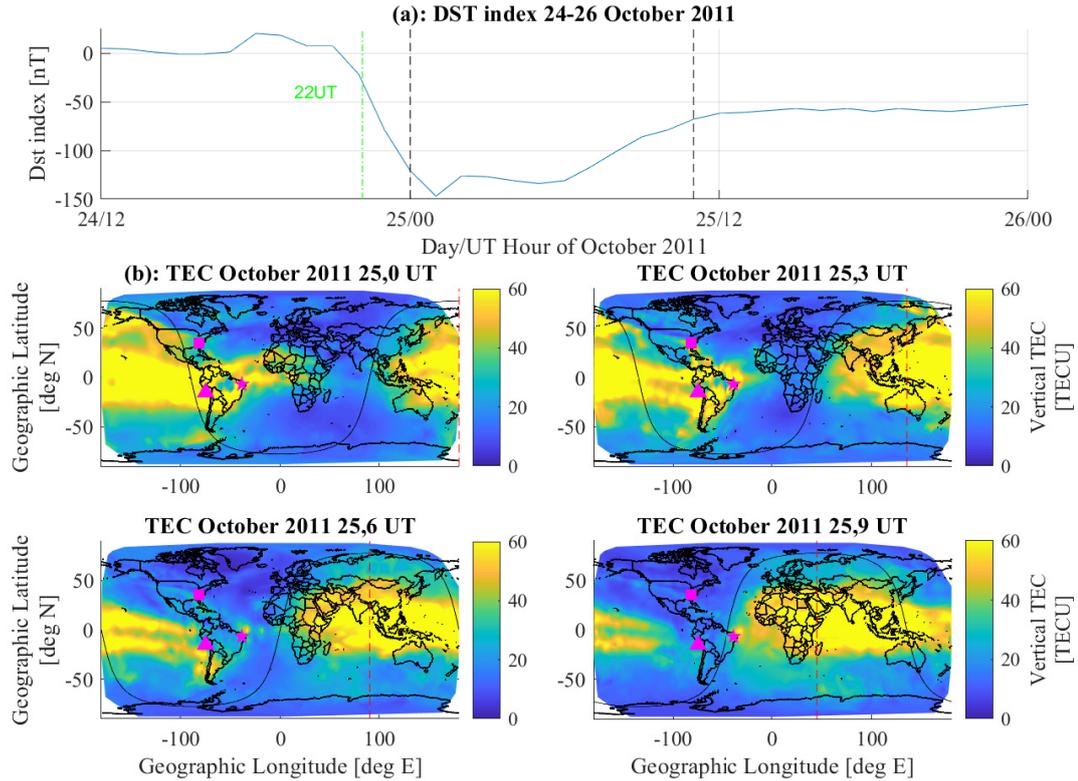


Figure 3.3. (a) Dst index for October 2011 in nT. Studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line. (b) TEC maps over time. Terminator line is indicated with a black solid line and local noon with dashed red line.

below  $-30$  nT as a storm event. At the end of the previous day, 24th of October, around 22:07 UT, the index crosses this Dst value and continues decreasing, indicating the rise of geomagnetic activity.

To observe the ionosphere conditions over this day, the TEC values at 0, 3, 6, and 9 UT are also shown in Figure 3.3 (b). The day/night terminator is indicated with a black solid line and it specifies local night, that is when neutral wind measurements are available.

During this day, there were 4 FPIs (Fabry-Perot interferometers) available that measured the neutral winds: Pisgah, Cariri, Cajazeiras and Nazca FPI. Cariri and Cajazeiras sites ( $7.38\text{S}$ ,  $36.53\text{W}$  and  $6.88\text{S}$ ,  $38.56\text{W}$ ) during this event are paired

and were observing the same volume. Combined they can give estimations of neutral winds at that volume, but both have data for different directions the instruments were looking to. Both locations are close and given the EMPIRE grid resolution, the system is not going to distinguish between these two locations, so only one the FPI's data will be used. The data from the Cajazeiras site is less reliable than the Cariri site, as it was more cloudy as explained in the technical background, so in this study only the measurements from the Cariri site are used. The TEC at this location decreases with time.

The Nazca FPI (14.97S,74.89W) and the Pisgah FPI (35.2N, 82.85W) are also used. At the Nazca site the TEC also decreases over the local night. The Nazca FPI belongs to a FPI network in Perú, but for this event only the Nazca one is reliable [93]. At the Pisgah site the TEC shows a higher peak around 3UT and then decreases. These 3 sites, Cariri, Pisgah and Nazca, are marked in Figure 3.3 with a magenta star, square and triangle respectively. These measurements will be used as validation data and they will be compared to the EMPIRE estimations.

The FPI measurements are given in different LOS (line of sight) directions of the instrument. The estimated geographic neutral winds will be projected to this LOS direction for the comparison. For the Pisgah FPI there are measurements available in the 4 geographic directions (north, south, east, west). At the Cariri site we have data for only 2 LOS directions, as it was operating in a common-volume mode with the Cajazeiras site (not used here due to cloud cover). The Nazca FPI was also operating in the common volume mode and there are 3 LOS available. The pierce point of the LOS is not the same as the location of the sites. The separation between the site location and the observation pierce point at 250 km is  $<2.5$  degrees, which is less than the resolution of 6 degrees of our estimate. We will use the site location as the point where we have the neutral winds data, as the grid will not distinguish between

this point and the actual region where the measurements are available.

**3.3.1.2 EMPIRE configuration.** The electron density rate will be derived from the electron density  $N$  from the IDA4D algorithm [18]. Examples of the vertically integrated densities of the density data that are input to EMPIRE are shown in Fig. 3.3(b). The [86] model is used for the ion production  $a_{0,prod}$  and loss  $a_{0,loss}$  and for the gravity  $a_{0,g}$  and diffusion  $a_{0,difsn}$  effect on the parallel direction of the ion velocity. To calculate these last two terms we use the neutrals properties characterized by the NRL-MSISE00 model [85] and the electron and ion temperature that are provided by the IRI model [84]. Finally, the magnetic field  $\vec{B}_0$  is provided by the IGRF model [59].

A global analysis is done by using a grid of magnetic colatitude  $\theta$  between 3 and 177 degrees with a resolution of 6 degrees, magnetic longitude  $\phi$  between -180 and 180 degrees with a step of 6 degrees and altitude  $h$  between 200 *km* and 500 *km* with a step of 50 *km*, following the altitude limits set in [32]. The grid has 12810 points at each time step. The analysis period is of 11 hours starting at 0 UT on the 25th of October 2011 with 20-minute increments, because during this time there are FPI measurements available with which to validate the estimations.

**3.3.1.3 Cases.** Two different cases are run in EMPIRE. In the first one, we use the VSH derivation to estimate the neutral winds in the meridional and zonal directions, from now on “Case VSH.” In this case, the state will be calculated with the direct estimation described in Section 3.2. For the second case, from now on “Case PS,” the power series expansion will be used to describe the meridional neutral wind. The zonal direction will not be estimated using EMPIRE, as there is little observability in this direction using the power series expansion [45]. With the vector basis function we can expand the neutral wind projection into the magnetic field into geographic north and east directions, while with the power expansion in the geographical directions is

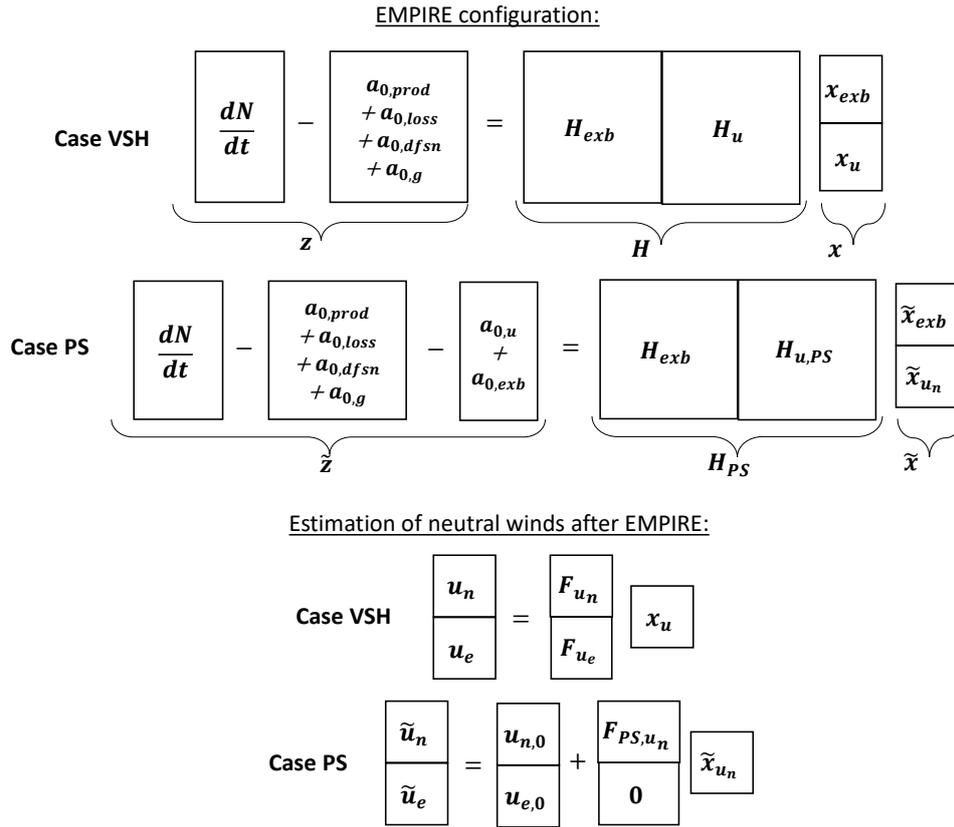


Figure 3.4. Diagram of the EMPIRE configuration and of the calculation method of storm neutral winds for each of the studied cases.

not possible. The estimation of the zonal component with the scalar expansion will be noisy at the points with no observability, so we use as the zonal component in the Case PS the model HWM14 data. Another difference between the two cases is that the state will represent a correction to the model in Case PS, as implemented in previous work [45].

In Figure 3.4 the EMPIRE configuration is shown for each of the cases. Each block represents the vector or matrix indicated inside. After EMPIRE is run, the estimation of the neutral winds is done. The *tilde* notation represents the case PS, where the correction to climate model values is calculated and the power series expansion [32] is used to determine the mapping matrix  $\tilde{H}_{u_n}$ .

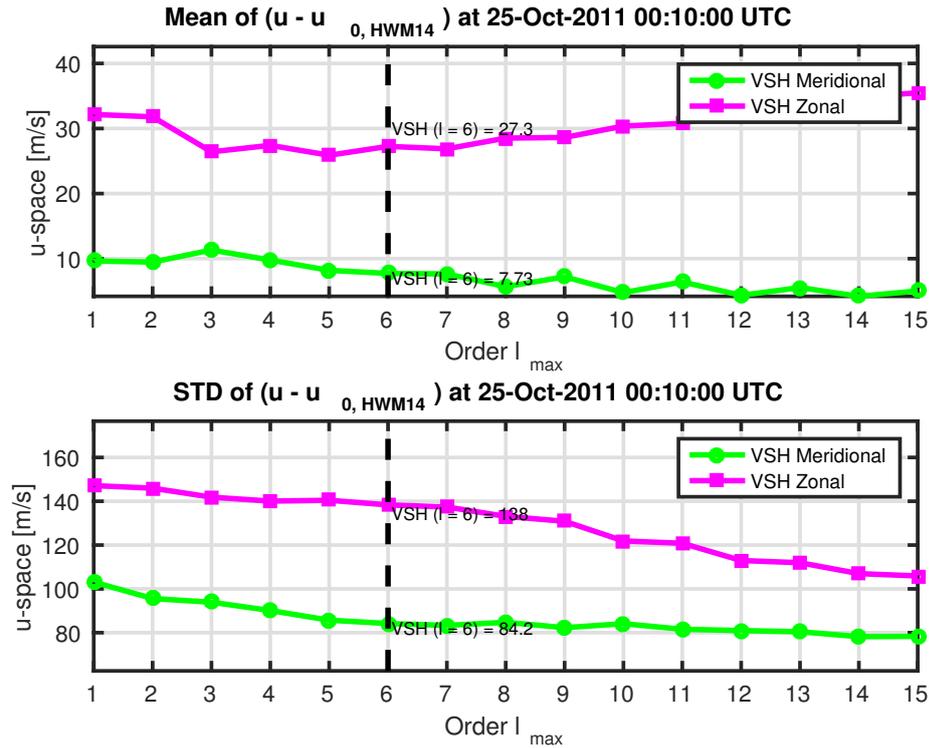


Figure 3.5. Mean (top) and standard deviation (bottom) over all grid points of the error of the estimated neutral winds and the model HWM14 data over order  $l_{\text{max}}$  of the VSH expansion at 0 UT on October 25. Zonal component is shown in magenta and meridional component in green. Black dashed line represents the values at  $l_{\text{max}} = 6$ .

For Case VSH the vector spherical harmonics decomposition is used to estimate the global neutral winds  $u_E$  and  $u_N$ . The  $l_{\text{max}}$  order selected is  $l_{\text{max}} = 6$ , as shown before in Figure 3.1. To select the error of the fitting needed for the Kalman filter, defined in Equation 3.31 as  $\sigma_{u_N, \text{fitting}}^2$  and  $\sigma_{u_E, \text{fitting}}^2$ , we analyze the results of the fitting of the expansion to HWM14 values, similar to results shown in Figure 3.1. In Figure 3.5 the mean and the standard deviation of the residual of the fitting for both neutral wind directions are shown over maximum degree  $l_{\text{max}}$  chosen. Top plot of 3.5 represents the mean over all the grid points of the error between the neutral winds model and the estimated values with the VSH fitting at 0 UT on October 25. In magenta the meridional values are shown and in green the zonal direction. Both are plotted over  $l_{\text{max}}$  order. Black dashed line represents the values at  $l_{\text{max}} = 6$ . The

bottom plot represents the standard deviation of the error between the fitting and the model HWM14. It is observed that the mean of the neutral wind residual is really stable for any order of the expansion. For the zonal direction is around  $55m/s$  and for the meridional component around  $-15m/s$ .

In the bottom plot of Figure 3.5, the mean of the residual neutral winds is shown over maximum order of the expansion. We can see that the values decreases with the order  $l_{max}$ . Choosing a value of  $l_{max} = 6$  is a good compromise between the number of coefficients to estimate and the standard deviation of the residual error. This error of the fitting is used to set up the background covariance  $\sigma_{u_N,0}^2$  and  $\sigma_{u_E,0}^2$  defined in Equation 3.31, for the meridional direction we use  $\sigma_{u_N,fitting}^2 = (100m/s)^2$  and for the zonal direction  $\sigma_{u_E,fitting}^2 = (130m/s)^2$ . The scale factor  $c_u$  of this error is set as 5, based on engineering judgment, to take into account the error in the mapping matrix  $H_u$  itself, as it depends on the measurement  $N$ .

For the power series decomposition in Case PS, we use an order of  $(l_{max}, k_{max}, p_{max}) = (3, 3, 3)$  as chosen in [32]. The configuration of the variances used for this case are also described in [32]. For the scalar spherical harmonics fitting for the potential field estimation  $V$ , the maximum degree is also  $l_{max} = 6$ , as described in [45]. The error of the fitting to the model data was calculated for every direction and the standard deviation of the residual error was:  $\sigma_{v_{exb,r},fitting}^2 = (50m/s)^2$ ,  $\sigma_{v_{exb,\theta},fitting}^2 = (110m/s)^2$  and  $\sigma_{v_{exb,\phi},fitting}^2 = (190m/s)^2$ . Again, the scale factor  $c_{exb}$  for this driver is set to 5 to take into account the propagated error of  $H_{exb}$  due to the error of the electron density ingested  $N$ .

Once all the orders of the expansions have been defined, the state vector size can be determined. For case VSH it is  $[90 \times 1]$  and  $[108 \times 1]$  for case PS: 42 coefficients for  $V$  and for  $\vec{u}$  there are 48 terms in the Case VSH and 64 terms in the Case PS.

The regularization matrix  $\mathbf{L}$ , defined in Section 3.2, depends on the state chosen and its size. For the neutral winds the order of magnitude selected to normalize is  $100m/s$  and for the potential field is  $1kV$ . The Kalman filter that solves the EMPIRE system also needs to be configured. In each time step, we need a first guess of the estimated drivers. The neutral winds  $\vec{u}_0$ , used to calculate  $\mathbf{x}_{u,0}$  in Equation 3.27, are provided by the Horizontal Wind Model 2014 (HWM14) [50] and the climate ion drift  $\vec{v}_{exb,0}$ , used to calculate  $\mathbf{x}_{exb,0}$  in Equation 3.26, is given by the [58] model. The error of the neutral winds climate model,  $\sigma_{u,climate}^2$  defined in Equation 3.31, is set as:  $\sigma_{u_n,climate}^2 = (75m/s)^2$  and  $\sigma_{u_e,climate}^2 = (75m/s)^2$ , as defined in [32]. For the ion drift, we select a  $\sigma_{v_{exb},climate}^2 = (20m/s)^2$  at each of the directions.

The measurement covariance matrix  $\mathbf{R}$ , used in the measurement update of the Kalman filter in Equation 3.32, is derived from the IDA4D densities error, considering that the measurement in our system is the finite difference electron density, and from the errors of  $\mathbf{a}_0$ . The errors of  $\mathbf{a}_0$  are calculated from different climate models, so there is no error available associated. We make the hypothesis of an error of 1% of the actual value of  $\mathbf{a}_0$  as during storm time we are assuming that the main changes in the ion drift are due to ion drift and neutral winds, and not due to production, loss, gravity and diffusion, effected embedded in density rate term  $\mathbf{a}_0$ . Results with 10% are comparable to results with 1%. The time constant  $\tau$  needed for the Gauss-Markov transition in the matrix defined in Equation 2.28 is set to  $\tau = 7200s$ , similarly to previous work. The filter at any time step will “forgot” the measurements ingested from 2 hours before.

**3.3.2 Results.** Neutral wind estimation for both cases over longitude is shown in Figure 3.6 at 0:10 UT on October 25 at geomagnetic latitude 3deg. Top plot in Figure 3.6 represents the northward components of the winds. In red the PS estimation is represented and in blue the VSH results. The black dashed line indicates

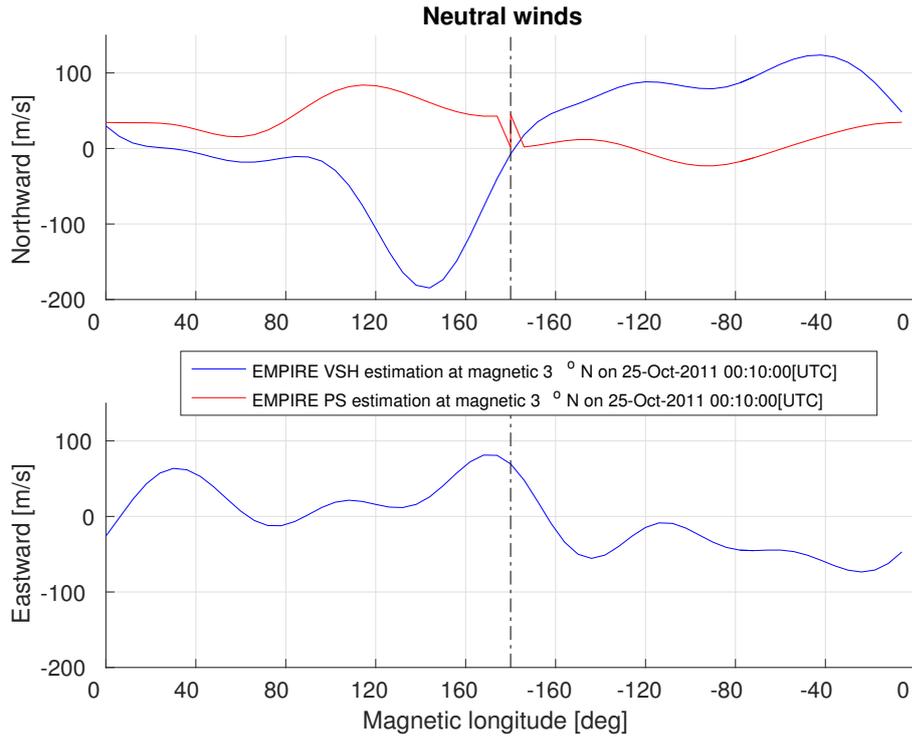


Figure 3.6. Neutral wind estimation at 00:10 UT and 3 deg magnetic latitude over magnetic longitude. Top shows meridional direction and bottom zonal direction of neutral speed. Red represents case PS estimate and in blue case VSH is shown. Black dashed line indicates geomagnetic longitude boundary  $180 = -180$  deg.

the geomagnetic longitude boundary  $180 = -180$  deg. We can see that the VSH in blue is continuous over this boundary while the PS estimation in red has a gap at this longitude. The bottom plot of Figure 3.6 represents the geographical component of the winds. It is observed that there is no PS estimate for this direction, as we are not correcting this direction for this case. The blue VSH winds are continuous over the sphere in the  $180 = -180$  deg boundary as expected.

VSH case is continuous over the globe as imposed with the vector spherical harmonics expansion. To analyze how the actual values of the winds changes by choosing the VSH case instead of the PS case, the results are compared to measurements at the 3 FPI sites described in the previous section. The neutral winds estimation for each of the two cases calculated by EMPIRE at the Pisgah FPI location is shown in

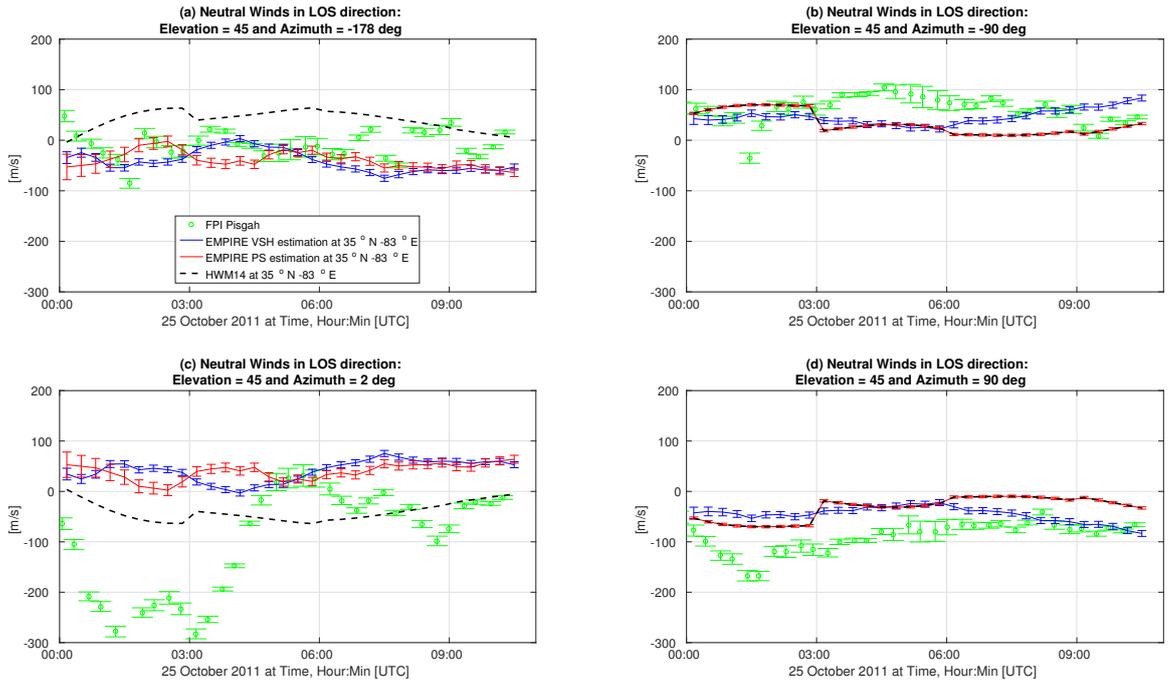


Figure 3.7. Neutral winds plotted over time at Pisgah location to the (a) south, (b) west, (c) north and (d) east. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value.

Figure 3.7. The y-axis indicates the speed along the LOS direction, positive away from the FPI. The x-axis represents the studied period of time. Each of the subplots represents the LOS direction of the FPI instrument, given by the azimuth angle, and (a) is south, (b) west, (c) north and (d) east. In blue the results for Case VSH are shown, in red for Case PS and in green the FPI measurements that are used for validation. The black dashed line is the model HWM14 values, that are corrected by EMPIRE using the electron density rate measurements for Case PS and used as an initial guess for Case VSH. As expected, in cases (b) and (d) the estimated speed for Case PS in red coincides with the model values in black because the zonal component of the wind is not being corrected for this case.

In the south LOS (a) the Case VSH estimation in blue follows the trend of the measurements and it is within most of the error bars until approximately 7 UT, when

the model HWM14 in black is closer to the FPI values. The Case PS estimation in red is also close to the measurements, but doesn't capture some of the motion of the measurements, such as the shift to positive values around 3 UT. From 6 UT onward both cases show similar results.

On the other hand, in the northward direction (c), we can see that both PS and VSH cases estimate similar results. Both cases indicate a northward wind instead of a southward wind, as HWM14 says, and the measurements say until 4 UT. From 4 UT, the measurements and both estimates are closer to each other, until 6 UT, when the estimate for both cases diverge from the measurement values.

In the West LOS (b), it is observed that the PS Case results coincides with the HWM14 model. This is expected as we are only correcting the meridional component of the wind for this case. However, in the VSH Case, the zonal direction can be estimated thanks to the vector basis, so that the projection of the results into the West LOS doesn't coincide with the model values. Until 3 UT the VSH estimate follows the model trend but with a 20 m/s difference. From 3 UT to 6 UT the estimation is really close to the model and differs from the measurements. However, around 7 UT the VSH estimation is able to pick up the stronger westwards winds that the measurements are indicating. From 9 UT the measurements weaken and the VSH estimate is not able to capture it.

In the East LOS (d), the behavior of the estimate is similar to the West LOS comparison. But at the end of the studied period, around 7 UT, the estimate VSH is within the error bars of the measurements.

In Figure 3.8, the results at Cariri are shown. For this day, there were only measurements from the FPI in two LOS directions. The first LOS (a) is looking in a west-northwestward direction, given by the FPI azimuth angle of  $-76^\circ$ . Because the

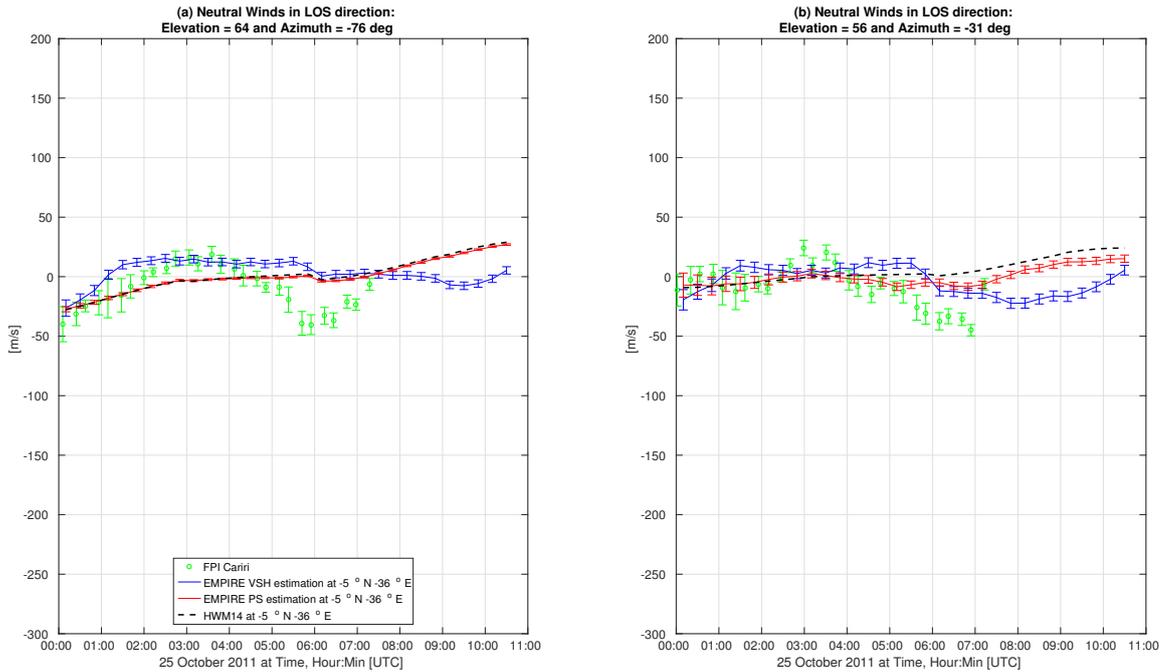


Figure 3.8. Neutral winds plotted over time at Cariri location to the (a) west-northwest and (b) north-northwest. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value.

zonal component is the same as the background model for the PS case, we expect the results projected in this direction to be close to the model projection. However, the VSH can give an estimate in the zonal direction and it is observed that it follows a similar trend to the measurements. The first 4 hours it is able to capture the change in the motion from negative to positive and follows closely the measurements. From 4 UT the FPI data show a decrease in the winds that any of the cases is able to capture.

In the second direction (b), the LOS is north-northwestward. The PS results in red, projected into the LOS, now differ more from the model values than in the previous LOS as expected. It follows the FPI data closely until 5 UT. The estimate for Case VSH has a similar trend to the FPI measurements after 3 UT, although again looks delayed with respect the FPI data. For example, we can also see a decreasing

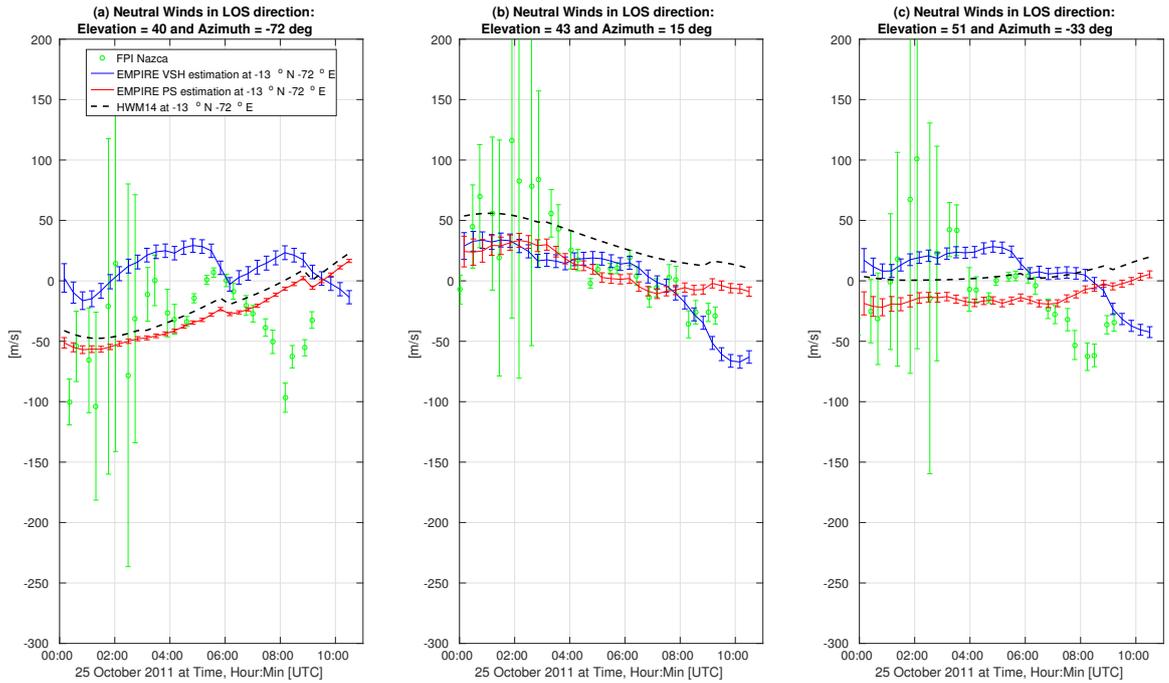


Figure 3.9. Neutral winds plotted over time at Nasca location to the (a) west-northwest, (b) north-northeast and (c) north-west. Green circles indicate the measurement from the FPI, blue indicates the value at the closest grid point to the FPI location for Case VSH and the red for Case PS. Black dashed line indicates the HWM14 value.

wind in this direction from 5 UT with shift from positive to negative.

In Figure 3.9, the results at the Nazca site are shown. There are 3 LOS direction measurements available for this FPI. The measurements error bars for this site are larger than those of the other two sites, although from 3UT the errors tend to be smaller than at the beginning of the studied period. In the first LOS (a), the direction of the measurements is west-northwestward. The PS case is close to the model, as expected, because the zonal component is the model HWM14 value. The VSH estimation in this look direction differs more from the measurements than the PS and model. However, it is observed that it follows a similar trend, but delayed in time around 2 hours and 30 m/s above what the measurements indicate. We can see this relation by comparing the negative peak of the measurements at 4 UT with the negative peak of the VSH estimate around 6 UT.

In the second LOS (b), the direction of the winds is mostly northward. The VSH estimate is close to the model until 6 UT, when it starts to decrease to negative winds. The FPI winds show until 3 UT high positive speeds, but then they decrease following the trend of the VSH estimate. The PS estimate is also close to the FPI values, but at the end of the period, around 8 UT, is not able to capture the more negative speed of the neutrals.

In the last LOS direction (c) the direction of the FPI data is north-westward. VSH in blue indicates a positive motion until 6 UT while the PS estimate in red stays negative for almost all the period. Until 4UT the FPI values indicate a positive velocity that the VSH also indicates. It is observed that then the FPI measurements decrease to 0 m/s for 2 hours from 4 UT to 6 UT followed by a decrease to negative winds. The VSH estimate has a similar behavior but delayed in time approximately 2 hours.

To compare both cases quantitatively, the root-mean-square (RMS) difference between the EMPIRE estimates and the FPI measurements is calculated. EMPIRE results are linearly time-interpolated to the FPI measurement times to compute the residuals. Table 3.3.2 shows the RMS of the residuals for each of the LOS directions at each of the sites for the Case VSH, Case PS and HWM14 data. Boldtype highlights the smallest RMS and normalized RMS for each of LOS, among the 2 cases and the model.

It is observed in Table 3.3.2 that in the Pisgah location, the RMS for the VSH estimation is better than the model and than the PS estimation except in the north direction (LOS3 direction represented in Figure 3.7 (c)). In the south, east, and west directions the RMS is reduced by 20, 7, and 7 % compared to the background model. We can also see that the RMS in the PS case in the west and east direction (LOS2 and LOS4 directions represented in Figure 3.7 (b) and (d)) is the same as the model

Table 3.1. RMS error in m/s between EMPIRE estimate for case VSH, case PS and model HWM14 and FPI measurements at Pisgah, Cariri and Nasca.

RMS ( <i>m/s</i> )	Pisgah S	Pisgah W	Pisgah N	Pisgah E	Cariri WNW	Cariri NNW	Nasca WNW	Nasca NNE	Nasca NW
VSH	<b>48</b>	<b>42</b>	177	<b>56</b>	<b>18</b>	17	56	30	<b>38</b>
PS	48	49	174	60	19	<b>16</b>	<b>36</b>	<b>29</b>	39
HWM	61	49	<b>124</b>	60	20	21	38	32	40

RMS, because we are not correcting the zonal direction for this case. Using VSH allows EMPIRE to have an estimate of the zonal component, that in the case of Pisgah, improves the RMS of the neutral winds compared to the FPI measurements. In the north direction (LOS3 direction represented in Figure 3.7 (c)) the RMS of the model is better than any of the EMPIRE estimations, but all of them exceed 100 m/s.

The analysis in this work has been done assuming the FPI measurements as our truth reference, but the measurements have been shown to contain an atmospheric scattering component, as explained by [54], in which they say errors up to 400 m/s can be reached during storms in the neutral wind measurements. Looking at the brightness of the samples, not shown here, the North LOS at Pisgah looks the most trustworthy and the South the least trustworthy. EMPIRE is not able to capture the southward surge seen in the data in Figure 3.7 (a) from 0-5 UT. The strong southern winds are likely a TAD (traveling atmospheric disturbances) from sudden heating in the storm [70]. EMPIRE might not be able to capture it because the time resolution used, 20 minutes, is too large compared with the disturbance.

At Cariri site, there is a 10 – 20% reduction in the RMS using EMPIRE over the model. In both directions, the RMS for Case PS and Case VSH are similar. This indicates that using the VSH gives us similar results to that of PS, but allows the

system to be global and to have zonal estimations.

Finally, for the Nasca site, we can observe that the estimation of Case VSH is closer to the FPI measurements than the model and the Case PS for only the last LOS northwest, represented in Figure 3.9 (c). However, the RMS residual for the second LOS, plotted in Figure 3.9 (b), is really similar to the PS case. Finally, in the direction west-northwest, shown in Figure 3.9 (a), the VSH estimate is worse than both the model HWM14 and the PS results. However, Nasca is the site whose FPI measurements have one-sigma errors of about 75-100 m/s until 3UT, which encompasses the VSH estimates much of the time. These high errors are also coupled by a weaker intensity of the samples, not shown in this study, so they are less reliable. It may be originated by a penetration electric field during this storm, that lifts the plasma increasing the F-region peak. At higher altitudes the recombination rate decreases which reduces the brightness of the sample and the SNR (Signal to Noise Ratio) of the FPI data. Consequently, it is followed by an inflation of the uncertainties over this site. Later in the night, changes in the speed is generally attributed to the disturbance dynamo, which is driven by enhanced energy deposition into the high latitude ionosphere, as explained by [94]. This effect has a bigger time scale allowing the FPI sample to be clearer and smaller errorbars.

The normalized RMS with respect to the FPI measurements is also calculated, to help interpret the results. The normalized RMS errors are shown in Table 3.3.2.

It is observed that the trend of most of the cases is similar to the not normalized RMS previously analyzed. At Pisgah, the West LOS estimate in this analysis is better for the PS case and the climate model HWM14 case than the VSH case. However, in [54] they explain that the east LOS is more reliable than the west looking direction and in the East LOS the normalized RMS still indicates a better agreement between the measurements and the VSH case.

Table 3.2. Normalized RMS error with FPI measurements between EMPIRE estimate for case VSH, case PS and model HWM14 and FPI measurements at Pisgah, Cariri and Nasca.

Normalized RMS	Pisgah S	Pisgah W	Pisgah N	Pisgah E	Cariri WNW	Cariri NNW	Nasca WNW	Nasca NNE	Nasca NW
VSH	<b>5.4</b>	1.4	6.7	<b>0.5</b>	2.2	53.9	<b>12.6</b>	3.7	478.2
PS	8.6	<b>0.9</b>	5.2	0.7	<b>1.1</b>	<b>3.7</b>	26.6	<b>2.8</b>	273.8
HWM	9.5	0.9	<b>4.2</b>	0.7	1.2	26.4	21.0	5.8	<b>51.7</b>

At Cariri, in both LOSs the normalized RMS indicates that the PS case is in better agreement with the measurements than the VSH case or the climate model. In the first LOS, represented in Figure 3.8a, from approximately 0 to 2 UT, the PS estimate agrees with the measurements better and from 2 to 4 UT the VSH case is observed to follow the trend better. There is negative motion from 5 to 7 UT that neither of the estimates or the climate model is able to capture, but the PS estimate is closer to the values than the VSH estimate, improving the total calculated RMS.

At Nasca, the normalized RMS at the first LOS, WNW represented in Figure 3.9a, indicates a better agreement of the VSH estimates out of all the cases with the measurements than the non normalized RMS. However, at the last LOS, represented in Figure 3.9c, the best agreement given by the normalized RMS with the measurements is the HWM output. The normalized RMS is really large for both EMPIRE estimates in this direction. It may be because the measurements at this LOS, which are the values used to normalize the residual between the estimations and the measurements, were close to 0  $m/s$  between 4 and 6 UT, increasing the normalized RMS errors to high values.

**3.3.3 Conclusions.** Neutral winds have been estimated with EMPIRE using a vector spherical harmonics expansion and a geomagnetic storm was studied globally during the 25th of October 2011. The use of the vector spherical harmonic expansion

allows the winds to be global and consistent in every direction. The vector basis also permits estimation in all directions and to be independent of the coordinate system used in the algorithm. Validation of the results was done by comparing them to FPI measurements at three different locations. A study of the RMS of the difference of the estimations and the measurements was also done.

As seen in the results and in the RMS values, in the majority of the LOS of the 3 sites the global vector spherical harmonics estimation is slightly better than the previous method and than the model when compared to the FPI values. In the Cariri site in both LOS studied (west-northwest and north-northwest) and in Nasca, specially in the north-west and north-northeast directions, the VSH results are able to follow the trend of the FPI measurements better than the model and the PS results. The northward direction of the FPI in Pisgah is the only LOS in where the model HWM14 behaves better than the EMPIRE estimations and this disagreement might be due to a bigger time scale used in EMPIRE than the ionospheric dynamics scale.

Doing a VSH estimation with EMPIRE gives similar results to the previous method utilized, although in some regions is better. However it allows EMPIRE to be a global system and be continuous over the whole region. A global basis also allows the ingestion of global measurements of neutral winds, like ICON measurements. The vector basis also has the advantage of allowing an estimation in any direction while the scalar basis functions was limited to the meridional direction. In the Pisgah comparison of the east and west LOS we can observe this limitation of the PS method. Vector basis self-consistency will also allow a better estimation of neutral winds in a region when ingesting FPI measurements in a different point, as the vector basis enables a relationship between all the points. Investigation of storms ingesting FPI and ICON measurements will be shown in Chapters 5 and 6.

## CHAPTER 4

### QUANTIFICATION OF REPRESENTATION ERROR IN EMPIRE ESTIMATION

The governing equation in EMPIRE is the ion continuity equation of the ionized atomic oxygen ( $O^+$ ) as this ionized species primarily dominates the F layer ionosphere. The EMPIRE observation equation is a simplification of the behavior that doesn't include chemistry or momentum and energy physics of the plasma. This reduction of the physics to the ion continuity equation will introduce a representation error in the algorithm.

The objective is to analyze how well EMPIRE is able to predict the global ionospheric drivers by reducing the observation operator to the ion continuity equation. A perfect representation of the ionosphere to use as comparison to study the representation error and to be considered as our "truth" does not exist. In this contribution, we use a self-consistent source that feeds both climate models and measurements data that EMPIRE needs to the algorithm and we will use a climate model as if it were, and it will be considered as our "truth". Comparing the results to the chosen self-consistent model, we see how well the EMPIRE simplification of considering only the ion continuity equation captures the dynamics. This comparison analysis and representation error characterization help us understand the strengths and limitations of the data algorithm and is used in future runs.

The self-consistent physics model that we use is the SAMI3 (Sami3 is Also a Model of the Ionosphere) algorithm [49], developed at the Naval Research Laboratory (NRL). SAMI3 models the ionosphere by solving the continuity and momentum equations of seven different ion species ( $H^+$ ,  $He^+$ ,  $N^+$ ,  $O^+$ ,  $NO^+$ ,  $N_2^+$  and  $O_2^+$ ) and the energy equations of three of them ( $H^+$ ,  $He^+$  and  $O^+$ ).

Section 4.1 discusses the different error sources considered for this study. Sec-

tion 4.2 describes the data used to run the EMPIRE algorithm as well as the configuration parameters for EMPIRE. EMPIRE analysis of the results are shown in Section 4.3. Section 4.4 presents a brief summary of the main conclusions.

#### 4.1 Error discussion

The representation error, which we want to characterize for the data assimilation algorithm EMPIRE, comes from the discrepancy between the true observation  $\mathbf{H}\mathbf{x}$  and what the data assimilation algorithm estimates  $\mathbf{H}\hat{\mathbf{x}}$ :

$$\epsilon = \mathbf{H}\hat{\mathbf{x}} - \mathbf{H}\mathbf{x} \quad (4.1)$$

$\epsilon$  is defined in the observation space, electron density rate-space, but can be translated to velocity space:

$$\epsilon_{\mathbf{v}} = \hat{v}_{EMPIRE} - v_{SAM3} \quad (4.2)$$

Using a self-consistent source for both the model background values and the measurements of electron density that EMPIRE needs allows us to study and characterize the representation error of choosing the EMPIRE simplified model to represent the ionosphere. Our “truth” is known and the representation error can be characterized by comparing the estimates to the original ingested self-consistent source.

Another error that comes into play in the EMPIRE formulation is the observation  $\mathbf{R}$  covariance error that characterizes the observation error  $\nu$  from Equation 3.20. It takes into account both  $\frac{\partial N}{\partial t}$  and  $\mathbf{a}_0$  terms forming the observation vector  $\mathbf{z}$ . We hypothesize the covariance matrix  $\mathbf{R}_{\mathbf{a}_0}$  of those background model terms that are not modeled (e.g. production, loss, gravity and diffusion) and the error covariance matrix  $\mathbf{R}_N$  of the measurements  $N$  to be independent, so that:  $\mathbf{R} = \mathbf{R}_N + \mathbf{R}_{\mathbf{a}_0}$ . The covariance matrix of the background transport terms,  $\mathbf{R}_{\mathbf{a}_0}$ , is assumed to be a diagonal matrix with the square error of the  $a_0$  term at each grid point. The covariance

matrix of the electron density rate  $\partial N/\partial t$ , defined as  $\mathbf{R}_N$ , is obtained by propagating the error of the  $N$  electron density observations as defined in [32].

On the other hand, the dynamic system used in EMPIRE has an associated process noise  $\mathbf{w}_{\mathbf{t}/\mathbf{t}}$ , defined in Equation 3.22 and characterized by the process covariance error  $\mathbf{W}_{\mathbf{t}}$ . The covariance  $\mathbf{W}_{\mathbf{t}}$  depends on the variances of the background velocity components  $\sigma_{i,0}^2$ , as shown throughout Equations 3.28 to 3.30, which are defined in Equation 3.31 as a function of a fitting error  $\sigma_{i,fitting}$ , a climate model error  $\sigma_{i,climate}$  and a scale factor  $c_i$ .

## 4.2 Method

The goal of this study is to quantify the distribution means  $\mu_{i,fitting}$  and standard deviations  $\sigma_{i,fitting}$  of Equation 3.31 and to quantify the representation error in the velocity space  $\epsilon_v$  in Equation 4.2.

The fitting error is needed to calculate the process noise  $\mathbf{w}_{\mathbf{t}/\mathbf{t}}$  from Equation 2.27. The climate model used as background in EMPIRE is considered as our “truth” in this work, as we want to be able to characterize the representation error of algorithm. Thus, the error associated with the climate model is set as  $\sigma_{i,climate} = 0$  for all of the drivers.

Then, we determine the fitting error of each of the drivers  $\sigma_{i,fitting}^2$  from Equation 3.31 by fitting the drivers to the background model values: SAMI3 for the ion drifts and HWM14 for the neutral winds. Equation 3.29 is solved for all  $i$  drivers for the maximum expansion order chosen  $l_{\max}$  at a specific time:

$$\hat{x}_{i,fitting} = \mathbf{F}_i^\dagger v_{i,SAMI3} \quad (4.3)$$

Then the estimated fitted drivers are calculated using Equation 3.29 again and they are compared to the background model values to define the fitting variance for each

driver:

$$\sigma_{i,fitting} = std(\mathbf{F}_i \hat{x}_{i,fitting} - v_{i,SAMI3}) \quad (4.4)$$

The self-consistent source used to estimate the representation error  $\epsilon_v$  of EMPIRE in this study is SAMI3. SAMI3 uses the [86] model for the ion production and loss and for the gravity and diffusion effect on the parallel direction of the ion velocity. The magnetic field is provided by the IGRF model and the electric field is solved with a potential solver derived from current conservation [87], [88]. The ion drift is calculated self-consistently with the solved perpendicular electric field in the low to mid-latitude region of the ionosphere. At high latitudes, SAMI3 uses the [58] model. The neutral winds are obtained from the HWM14 (Horizontal Wind Model) [50]. The neutral properties are characterized by the NRL-MSISE00 model [85]. The continuity and momentum equations of seven different ion species (H+, He+, N+, O+, NO+, N2+ and O2+) and the energy equations of three of them (H+, He+ and O+) are used to represent the ionosphere.

A quiet day case is selected for this study because the electron density rate  $\frac{\partial N}{\partial t}$  derived from the model SAMI3 electron density values  $N$  is more accurate than during storm time [95]. EMPIRE inputs need to have the smallest possible error so the disagreement between its estimations and SAMI3 is mainly due to the EMPIRE model itself. The day we study is August 25, 2018. The Ap index, which is an input to SAMI3, is shown in Figure 4.1. Given the classification by [96], a quiet-minor event corresponds to an Ap index between 0 – 20 nT while indices from 30 – 50 nT corresponds to a moderate storm. The Ap index reaches values higher than 20 nT between 17 UT and 18 UT, and after the index starts to increase, indicating that the geomagnetic storm starts to develop. From 21 UT, the index reaches the 30 nT threshold and on the 26th at 0 UT increases significantly. We consider as storm time after 18 UT, on the supposition that SAMI3 will not provide inputs as representative

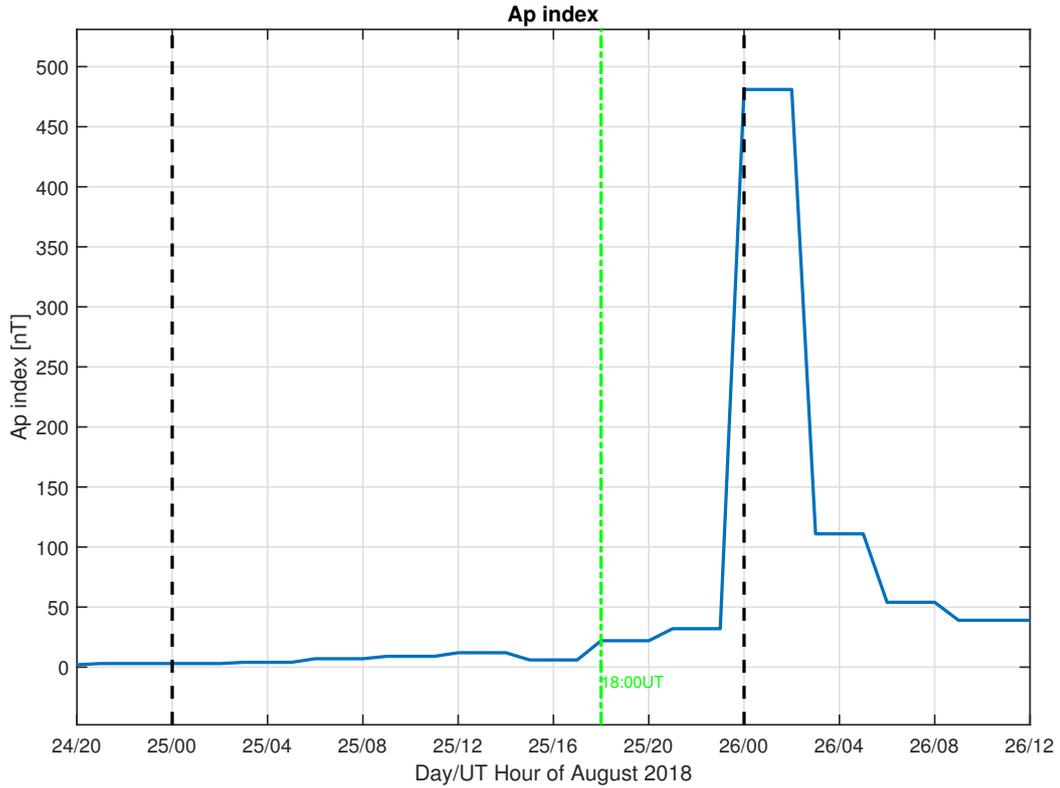


Figure 4.1. Ap index for August 2018 in nT. Studied time limits are indicated with black dashed lines and the beginning of geomagnetic storm time is indicated with green dashed-dot line.

of reality to EMPIRE when geomagnetic activity is present. However, our study period will go from 0 UT on the 25th to 0 UT on the 26th, to check if the estimation differs more from the model SAMI3 during storm time, as might be expected.

We run EMPIRE to compare the driver estimations to SAMI3 true values. SAMI3 values of the electron density  $N$  will be used as the pseudo-measurements in the data assimilation algorithm. The electron density  $N$  from SAMI3 is then finite differenced to obtain the electron density rate  $\frac{\partial N}{\partial t}$ , that EMPIRE ingests. Also, in EMPIRE we use the same background models that SAMI3 uses to make all the sources that EMPIRE ingests consistent with SAMI3. [86] model is used to specify the production, loss, gravity and diffusion effects ( $a_{loss,0}$ ,  $a_{prod,0}$ ,  $\vec{v}_{g,0}$  and  $\vec{v}_{dfs,0}$ ). IGRF model is used to describe the magnetic field  $\vec{B}_0$ . The neutral properties will be given

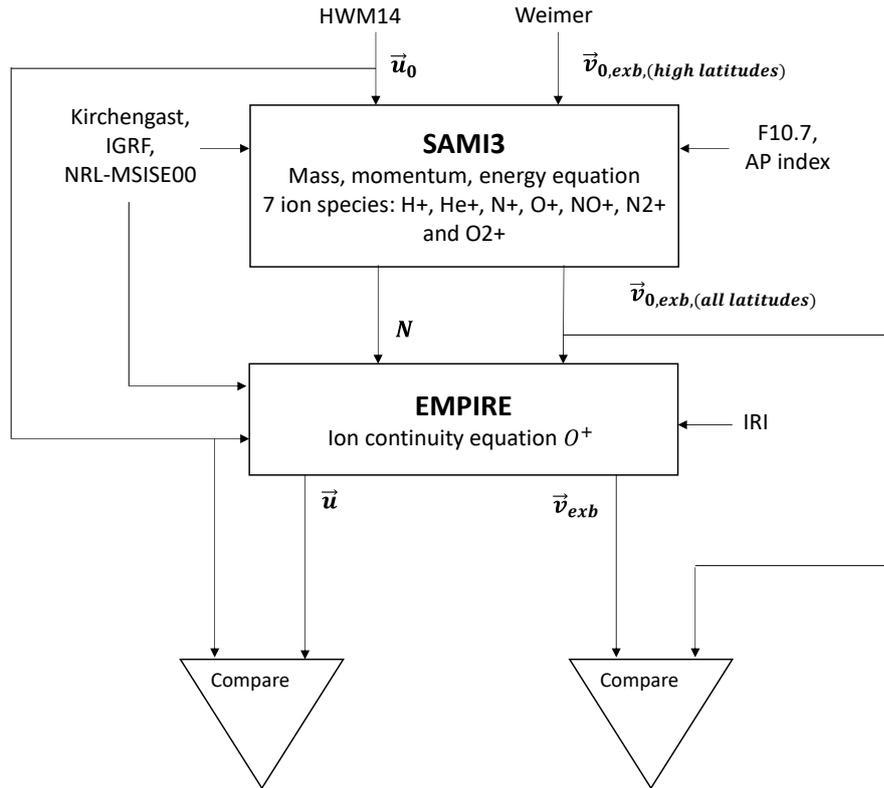


Figure 4.2. Block Diagram of SAMI3 and EMPIRE.

by the NRL-MSISE00 model as in SAMI3. However, the electron and ion temperature will be provided by the IRI model, instead of by the SAMI3 output values, as they were not available for the chosen case. It has been checked (not shown) that the EMPIRE system is not really sensitive to a change in these temperatures, so using IRI instead of SAMI3 will not change the EMPIRE estimation. A block diagram of how SAMI3 works and how it is coupled to EMPIRE is shown in Figure 4.2. The results of the estimated drivers with EMPIRE, ion drift  $\vec{v}_{exb}$  and neutral winds  $\vec{u}$  will be compared to the SAMI3 values to assess mean and standard deviation of the representation error of EMPIRE.

A global analysis is done by using a grid of magnetic colatitude  $\theta$  between 3 and 177 degrees with a resolution of 6 degrees, magnetic longitude  $\phi$  between -180

and 180 degrees with a step of 6 degrees and altitude  $h$  between 200 *km* and 650 *km* with a step of 50 *km*. The grid has 12810 points at each time step. The analysis period is of 24 hours starting at 0 UT on 25 August 2018 at 20-minute increments.

To estimate the global neutral winds  $u_E$  and  $u_N$ , a vector spherical harmonics decomposition is used. The number of terms of the decomposition is limited by the maximum degree selected  $l_{max} = 6$ . For the scalar spherical harmonics fitting for the potential field estimation  $V$  the maximum degree is also  $l_{max} = 6$ , as described in [45]. The state vector size for each of the cases is  $[90 \times 1]$ :  $k_u = 48$  coefficients for  $\vec{u}$  and another  $k_{exb} = 42$  terms for  $V$ .

In each time step, we need a first guess of the estimated drivers  $\mathbf{x}_{0,t+1/t}$  and  $\mathbf{x}_{0,t/t}$ , defined in Equation 2.27. SAMI3 values of the ion drift will be used as the background drift used in EMPIRE  $\vec{v}_{exb,0}$ . The neutral winds background values  $\vec{u}_0$  will be given by the HWM14 model, just as SAMI3 utilizes. This choice makes all the inputs of EMPIRE essentially self-consistent with each other. The corresponding background state is obtained by using Equation 3.29. The time constant value of the Kalman filter from Equation 2.29 is chosen to be  $\tau = 2$  hours.

The observation covariance of the background transport terms  $\mathbf{R}_{a_0}$  is obtained by making the hypothesis of an error of 1% of  $a_0$  to calculate the error of  $a_0$ . This small percentage is chosen because the model is considered as the truth to EMPIRE, so a small error in the inputs allows us to study how well EMPIRE itself is modeling the ionosphere. The observation covariance matrix  $\mathbf{R}_N$  of the electron density rate is obtained by propagating the density  $N$  error. The electron density  $N$  is given by the model SAMI3, used in this work to assess the representation error. [97] did a study case of the uncertainties of the SAMI3 electron densities compared to measurements and established that they are less than 25 %. Based on this, a hypothesis of an error of 20% of  $N$  is made without considering any correlation between individual errors.

They can be correlated in space, but their quantitative estimates are difficult and will increase the computational time of the filter. The measurement errors are distributed diagonally in the error covariance matrix of the electron density.

Once the filter is configured as described, the state is estimated with EMPIRE at each time step and the estimated neutral winds and ion drifts are obtained at every grid point for the desired period of time. To study the representation error of the ion drifts, we will compare our results to the SAMI3 drift output. The perpendicular-to-the-magnetic-field zonal and meridional components are going to be studied. The zonal averaged difference between the estimation and SAMI3 values of the 2 components, perpendicular meridional  $\bar{\epsilon}_{v_{exb,\perp m}}$  and zonal  $\bar{\epsilon}_{v_{exb,\perp z}}$  to the magnetic field line direction, will be shown over time and magnetic latitude at an altitude of 250 km. This height is chosen to be consistent with previous analysis using EMPIRE, shown in previous Chapter 3, in which neutral winds results are compared and validated to FPI (Fabry-Perot interferometer) values at this altitude.

To study the neutral winds representation error we will compare our results to the HWM14 model, as it is the neutral wind input that SAMI3 uses. The geographic zonal and meridional directions are analyzed using the zonal mean difference between the estimation and SAMI3 values (that uses HWM14), of the 2 components  $\bar{\epsilon}_{u_E}$  and  $\bar{\epsilon}_{u_N}$ . They will also be calculated at the same altitude of 250 km. The zonally averaged difference for all of the terms will be calculated at each time and magnetic colatitude as:

$$\bar{\epsilon}_v(\theta; t) = \frac{1}{n_\phi} \sum_{\phi=-180}^{\phi=180} \epsilon_v(\theta, \phi; t) \quad (4.5)$$

where  $\epsilon_v$  is the representation error in the velocity space defined in Equation 4.2 and the magnetic longitude step is 6 degrees and  $n_\phi$  is the number of magnetic longitudes used in the EMPIRE grid,  $n_\phi = 61$ .

The mean and standard deviation of the difference between the estimation of each of the components of the drivers and the model SAMI3 will also be calculated to assess the representation error of EMPIRE. They will be calculated over magnetic latitude by taking the mean  $\mu_\epsilon(\theta)$  and the standard deviation  $\sigma_\epsilon(\theta)$  over longitude and during quiet time:

$$\mu_\epsilon(\theta) = \frac{1}{n_\phi + n_{qtimes}} \sum_{t=0UT}^{t=18UT} \sum_{\phi=-180}^{\phi=180} \epsilon_v(\theta, \phi; t) \quad (4.6)$$

$$\sigma_\epsilon(\theta) = \left( \frac{1}{n_\phi + n_{qtimes}} \sum_{t=0UT}^{t=18UT} \sum_{\phi=-180}^{\phi=180} [\epsilon_v(\theta, \phi; t) - \mu_\epsilon]^2 \right)^{1/2} \quad (4.7)$$

where  $n_{qtimes}$  is the number of step times during quiet time  $n_{qtimes} = 55$  because the step of the time summation is 20 minutes.

Finally, we will compare the representation error  $\epsilon_v$  to the estimated error  $\hat{\sigma}_x$ , contained in the estimated covariance matrix  $\hat{\sigma}_x^2 = \text{diag}(\mathbf{P}_{\mathbf{t}/\mathbf{t}})$ . The error is transformed to the corresponding velocity space  $\hat{\sigma}_i^2$  using the mapping matrices from Equation 3.29. The mean  $\mu_\sigma$  of the error is calculated over time and longitude, similarly to the methodology in Equation 4.6:

$$\mu_\sigma(\theta) = \frac{1}{n_\phi + n_{qtimes}} \sum_{t=0UT}^{t=18UT} \sum_{\phi=-180}^{\phi=180} \hat{\sigma}_i \quad (4.8)$$

### 4.3 Results

To characterize the process noise  $\mathbf{w}_t$ , the fitting variance  $\sigma_{i,fitting}$  is analyzed as described in the previous section. The standard deviation value of the residual error between SAMI3 and the fitted drivers averaged over all grid points  $j$  is plotted over  $l_{\max}$  for all the  $i$  drivers in Figure 4.3a.

Figure 4.3a shows the standard deviations for each of the  $i$  drivers over the maximum degree  $l_{\max}$  chosen for the expansion:  $u_N$  in green,  $u_E$  in magenta,  $v_{exb,r}$  in

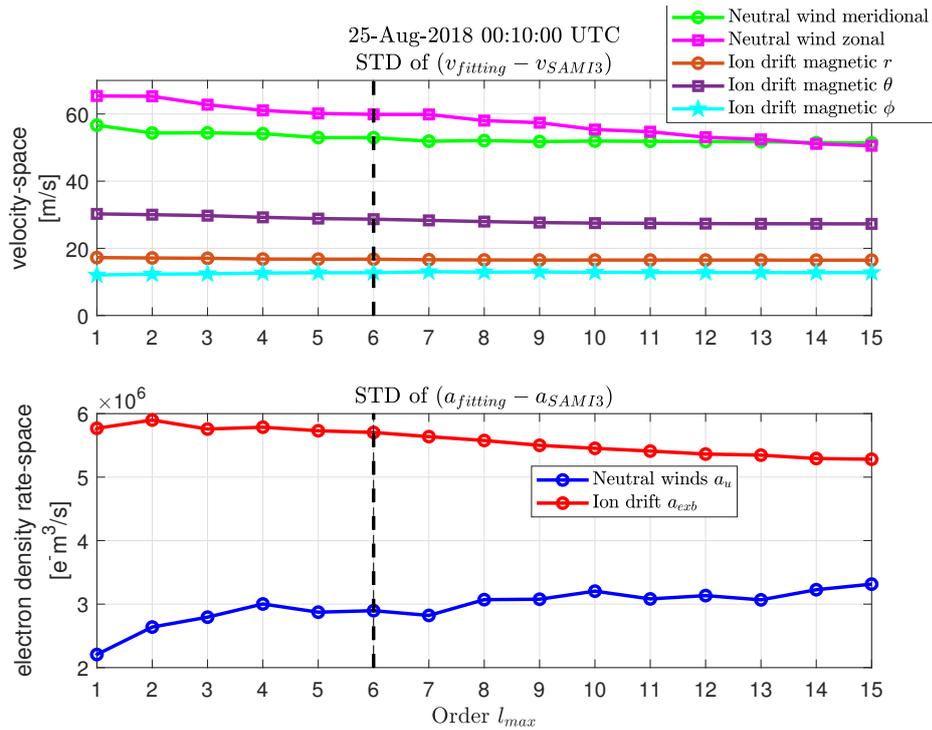


Figure 4.3. Standard deviation in (a) velocity-space and (b) density rate-space over all grid points of the error between the estimated drivers and the SAMI3 data over order  $l_{max}$  at 0:10 UT on August 25, 2018. In (a) the neutral wind zonal component is shown in magenta, the neutral wind meridional component in green, the ion drift magnetic up in orange, magnetic south in purple and magnetic east in cyan. In (b) the transport term corresponding to neutral winds is shown in blue and to the ion drifts is shown in red. The black dashed line represents the values at  $l_{max} = 6$ .

orange,  $v_{exb,\theta}$  in purple and  $v_{exb,\phi}$  in cyan. It is observed that the standard deviation of the fitted drivers decreases when the  $l_{\max}$  increases as expected for a fitting for the neutral wind components. For the ion drift components, the decrease of the residual with  $l_{\max}$  is not that clear. This may be because the selected time for the fitting is a quiet period, as seen in Figure 4.1. The ion drift absolute values are not that high and the order of the fitting does not affect the fitting. However, the expected tendency of a decrease of the residual with order for these drivers is observed for other studied times. In Chapter 3 during a storm time, this tendency was observed, although it is not shown in this dissertation. For  $l_{\max} = 6$ , as selected previously in [32], the following standard deviations between the fitting estimations and the background model values are obtained for the ion drift:  $\sigma_{v_{exb,r},fitting}^2 = (20m/s)^2$ ,  $\sigma_{v_{exb,\theta},fitting}^2 = (30m/s)^2$  and  $\sigma_{v_{exb,\phi},fitting}^2 = (10m/s)^2$ , where each one represent one of the components of the ion drift vector. For the neutral winds and the same maximum  $l_{\max} = 6$ , as selected in previous Chapter 3, the standard deviations between the estimated fitting wind and the HWM14 model are  $\sigma_{u_N,fitting}^2 = (60m/s)^2$  and  $\sigma_{u_E,fitting}^2 = (50m/s)^2$ .

The mean of the fitting error was also studied for different order  $l_{\max}$ , although it is now shown here. For the fitting of the ion drifts to SAMI3 drifts, the mean of the residual of the fitting for  $l_{\max} = 6$  were approximately  $8m/s$ ,  $0m/s$  and  $4m/s$  for the magnetic up  $\hat{r}$ , magnetic south  $\hat{\theta}$  and magnetic east  $\hat{\phi}$  respectively. In the neutral winds study, the mean of the residuals for  $l_{\max} = 6$  was around  $8m/s$  for both the zonal and the meridional direction. A non-zero mean error is denominated systematic error and to consider it in the Kalman filter solver a bias parameter would need to be introduced in the state vector. However, we make the assumption of zero-mean in the process noise  $w_t$ , so we do not consider these mean values of the fitting error.

EMPIRE estimates the drivers in electron density rate space, or observation space, as described in Equation 2.9. The corresponding  $a_{i,fitting}$  value for each of

the  $i$  drivers can be calculated after doing a fitting to SAMI3 values using Equation 2.9. Figure 4.3b shows the standard deviation in the density rate space for each of the  $i$  drivers over the maximum degree  $l_{\max}$  chosen. It is calculated by comparing the transport term  $a_{i,fitting}$  of each of the estimated drivers with its corresponding SAMI3 value  $a_{i,SAMI3}$ . Results for  $a_u$  are indicated in blue and  $a_{exb}$  in red. In the electron density rate-space, the standard deviation of the fitting does not monotonically decrease with the increase in the number of coefficients given by  $l_{\max}$ .

To characterize the representation error  $\epsilon$ , we study the difference between EMPIRE and SAMI3 outputs over time over the whole global region and we also study the dependence of the results with latitude, as the dependence with longitude has been shown to be less relevant in analysis of the results not shown here.

Figure 4.4 shows the zonally averaged ion drift difference  $\bar{\epsilon}_v$  in the perpendicular meridional direction (a) and in the perpendicular zonal direction (b) between EMPIRE results and SAMI3 “true” values, defined in Equation 4.5. The perpendicular zonal direction coincides with the magnetic  $\hat{\phi}$  direction. Positive values, in yellow, indicate that the estimation is higher than the model values. On the other hand, negative values in blue indicate an underestimate estimation compared to the SAMI3 values. The parallel ion velocity is not shown, as we have made the hypothesis that the ion drift is only seen in the perpendicular to the magnetic field direction. They are plotted over time for the whole day on 25 August 2018 and we consider quiet time until 18 UT, when the Ap index indicates geomagnetic activity.

The estimation in the perpendicular meridional direction, shown in the Figure 4.4a, is close to the model SAMI3 values, as the error is close to 0  $m/s$  in almost all region. We can see a higher disagreement close to the magnetic poles and also close to the magnetic equator over time. From 0 UT to 12 UT, in the northern hemisphere EMPIRE tends to overestimate the perpendicular meridional component and in the

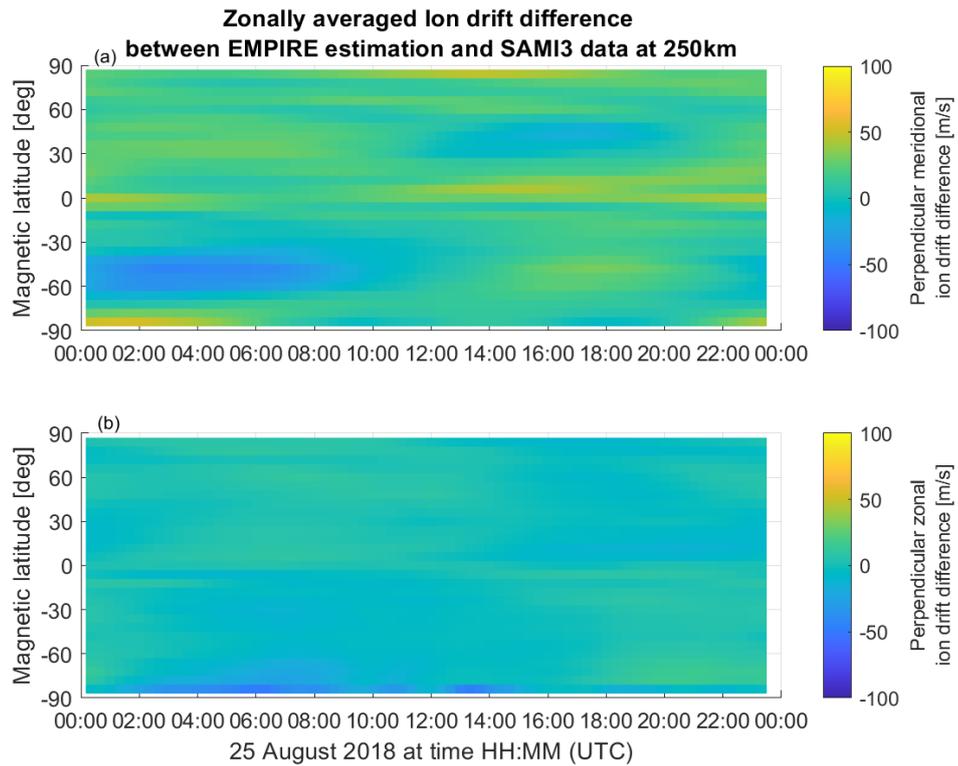


Figure 4.4. Zonally averaged error between ion drifts EMPIRE estimation and SAMI3 values in the perpendicular (a) to meridional and to (b) zonal direction plotted over time.

southern hemisphere it tends to underestimate it. From that time, the contrary is observed. In the perpendicular zonal direction, Figure 4.4b, the error is also small in almost all latitudes and times. We can see a higher disagreement at the high latitudes, especially in the southern hemisphere. The small order of magnitude of the errors may also be because the ion drifts during quiet times at lower latitudes are not that high.

EMPIRE estimates the neutral winds contribution to the electron density change simultaneously with the ion drifts. Figure 4.5 shows the zonally averaged error of the neutral wind in comparison with the model SAMI3 values over time and over magnetic latitude. Figure 4.5a shows the geographic meridional direction and Figure 4.5b the geographic zonal component of the neutral speed. The vertical component of the vector is assumed to be zero in EMPIRE, so it is not shown.

For both directions, meridional and zonal in Figure 4.5a and 4.5b respectively, at mid and high latitudes the EMPIRE estimations differ from SAMI3 more. In the meridional direction, Figure 4.5a, Around 18UT at all latitudes the error starts to increase and a step is visible in the plot. In both hemispheres the high error starts to span to lower latitudes. This may correspond to the developing of the geomagnetic storm around 18 UT, as we showed in the Ap index in Figure 3.3b, that is an input to SAMI3. Because SAMI3 is a model, the electron density output of SAMI3 can be underestimated [95] during geomagnetic storms while it has been analyzed that SAMI3 performs quite well in in comparison to daily averaged global measurements [98]. In the zonal direction, shown in Figure 4.5b, from 18 UT an increase of the error at low and mid latitudes that may correspond with the beginning of the geomagnetic storm is also visible.

We also observe that at some times when one of the neutral wind components is overestimated, the other component is underestimated. An example of this is

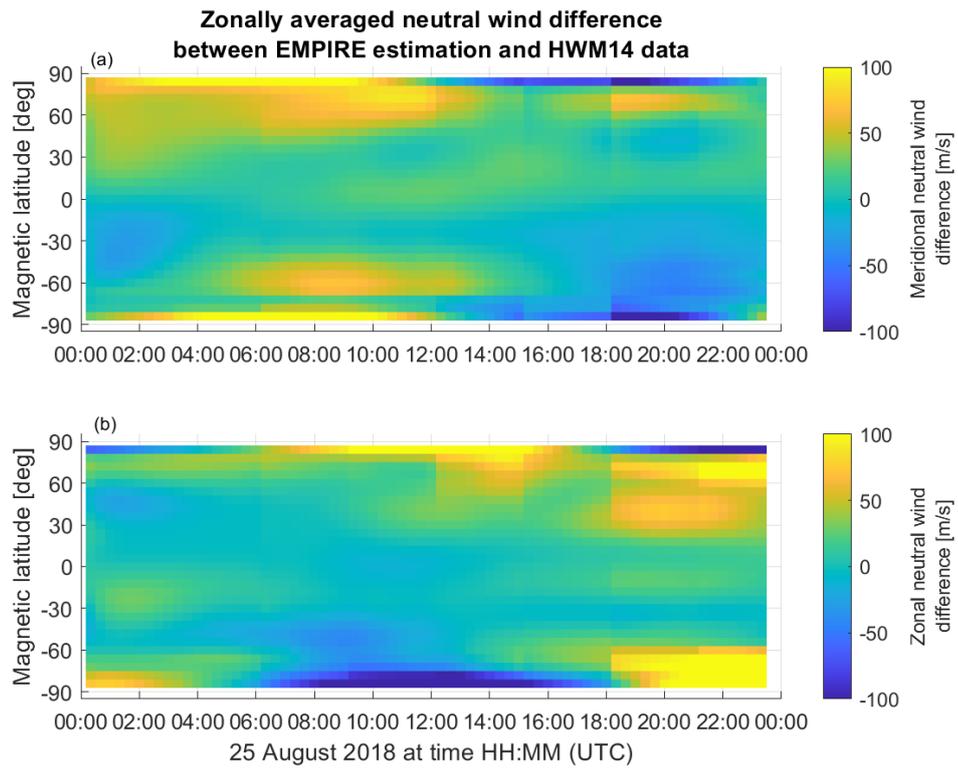


Figure 4.5. Zonally averaged error between neutral winds EMPIRE estimation and SAMI3 (HWM14) values in the (a) geographical meridional and (b) zonal direction plotted over time.

observed close to the poles between 6 UT to 10 UT around  $-60$  deg latitude, where the zonal error is negative (blue) while the meridional error is positive (yellow). This might be influenced by the estimation of the expansion in the density rate space. The divergence of the neutral wind vector is what EMPIRE estimates, so some accuracy might be lost when processing the results from the density rate space to the neutral winds space.

To obtain an order of magnitude of the errors by latitude, the mean and the standard deviation of the error averaged over longitude and over the quiet time (from 0 UT to approximately 18 UT) is computed. Figure 4.6a shows the results of the mean error over latitude for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction and Figure 4.6b shows the standard deviation.

The mean  $\mu_\epsilon$  of the residual  $\epsilon$ , in Figure 4.6a, is close to 0  $m/s$  at low and mid-latitudes for all of the estimated drivers. We can see that at high latitudes, from about 60 deg and  $-60$  deg, the mean of the neutral winds components, zonal in magenta and meridional in green, increases up to 50 – 60  $m/s$ . The mean of the error the ion drift estimations, perpendicular meridional in purple and perpendicular zonal in cyan, also increase at higher latitudes but only up to 20 – 25  $m/s$ . Close to the magnetic equator the ion drift perpendicular meridional component estimation also increases, as observed also in Figure 4.4.

In Figure 4.6b, the standard deviations of the residuals are shown. The perpendicular meridional direction, in purple, is about 20 – 25  $m/s$  for low and mid magnetic latitudes, approximately up to 70 degrees. For perpendicular zonal the standard deviation in cyan reaches values of 10  $m/s$  at these low and mid latitudes. The fitting error imposed in the Kalman filter to these drifts were set in Section 4.2 as  $\sigma_{v_{exb,r},fitting}^2 = (20m/s)^2$ ,  $\sigma_{v_{exb,\theta},fitting} = (30m/s)^2$  and  $\sigma_{v_{exb,\phi},fitting} = (10m/s)^2$ .

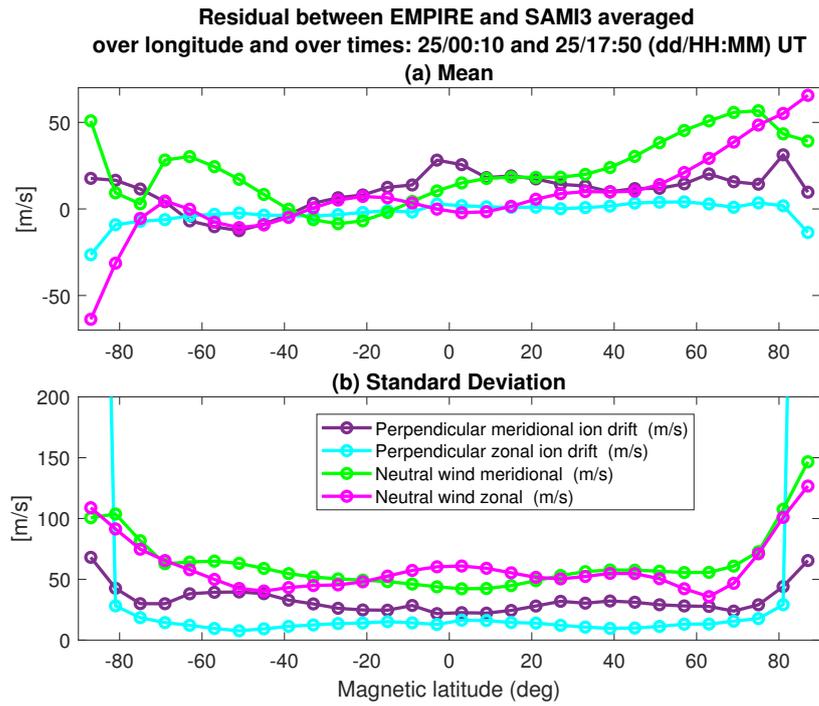


Figure 4.6. (a) Mean and (b) standard deviation of residual between EMPIRE estimation and SAMI3 values averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude.

The perpendicular zonal direction coincides with the magnetic  $\hat{\phi}$  direction. The standard deviation of the residual in this direction has the same order as the fitting error. The perpendicular meridional direction depends on the inclination angle  $I$ , that describes the angle between the geomagnetic field and the Earth surface. The error will be a combination between the magnetic  $\hat{r}$  and  $\hat{\theta}$  direction. Thus, the standard deviation of error in this direction is similar to the variance of the fitting imposed in the Kalman. This indicates us that the representation error of this component is mainly due to a fitting error from using the mapping matrix.

However, at higher latitudes, both the meridional and zonal standard deviation increase to higher order of magnitudes, about 1000  $m/s$  for the perpendicular zonal and 70  $m/s$  for the perpendicular meridional direction. The expansion used to estimate the potential field, described in [45], models the electric potential as constant along a dipole magnetic field line. This hypothesis becomes inaccurate around geomagnetic latitudes of 60 – 70 degrees, when L-shell reaches values of 10 [45]. Also, the magnetic field lines at high latitudes are “open”, not represented in the dipole model. The increasing error in the EMPIRE estimations with respect to the model SAMI3 around these magnetic latitudes may be due to this hypothesis not being true in these locations.

The neutral wind standard deviation of both meridional and zonal direction residuals, in Figure 4.6b in green and magenta respectively, at high latitudes increase with respect to the rest of the latitudes, to approximately 100 – 150  $m/s$ . At low latitudes, up to 20 deg, the meridional standard deviation in green is smaller than in the zonal direction, about 40  $m/s$ . At mid latitudes, it is about 60  $m/s$ . The zonal deviation, in magenta, at low and mid latitudes is around 50 – 60  $m/s$ . It is also seen that close to the magnetic equator the meridional error decreases while the zonal one increases a similar value, accentuating the fact that the estimation of

both components are connected. The variance imposed in the Kalman filter in the process noise due to the fitting for the neutral winds were  $\sigma_{u_N,fitting}^2 = (60m/s)^2$  and  $\sigma_{u_E,fitting}^2 = (50m/s)^2$ . The meridional and zonal residual at low latitudes decreases with respect to the fitting error, and at mid latitudes is about the same. This means that at low and mid latitudes the representation error of the neutral winds estimation of the data assimilation algorithm is mainly due to the use of the mapping matrix. The higher error of the neutral winds at high latitudes might be affected by the high error of the ion drifts at these latitudes. The assumption used for the calculation of the ion drifts of electric potential being constant along the dipole magnetic field becomes inaccurate from latitudes around 60 – 70 deg and poleward. If EMPIRE is attributing a high change in the electron density rate due to the ion drifts, but what is really ingested is smaller, then to compensate it EMPIRE attributes a high contribution due to the neutral winds. Also, some error at high latitudes might be due to not considering the auroral production as a source of electron density in the ion continuity equation in Equation 2.9. Another additional source of error may be due to the observation error not being exactly zero, as the error of the ingested SAMI3 electron density  $N$  was defined to be 1% of the value.

To have a sense of the order of magnitude of the residual, the mean and standard deviation are normalized with the climate model values. Figure 4.7a shows the results of the mean relative error over latitude for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction and Figure 4.7b shows the standard deviation.

It is observed that the mean of the relative residual in Figure 4.7a follows a similar trend as the one described for the non normalized results in Figure 4.6a. For the normalized standard deviation results, shown in Figure 4.7b, the normalized

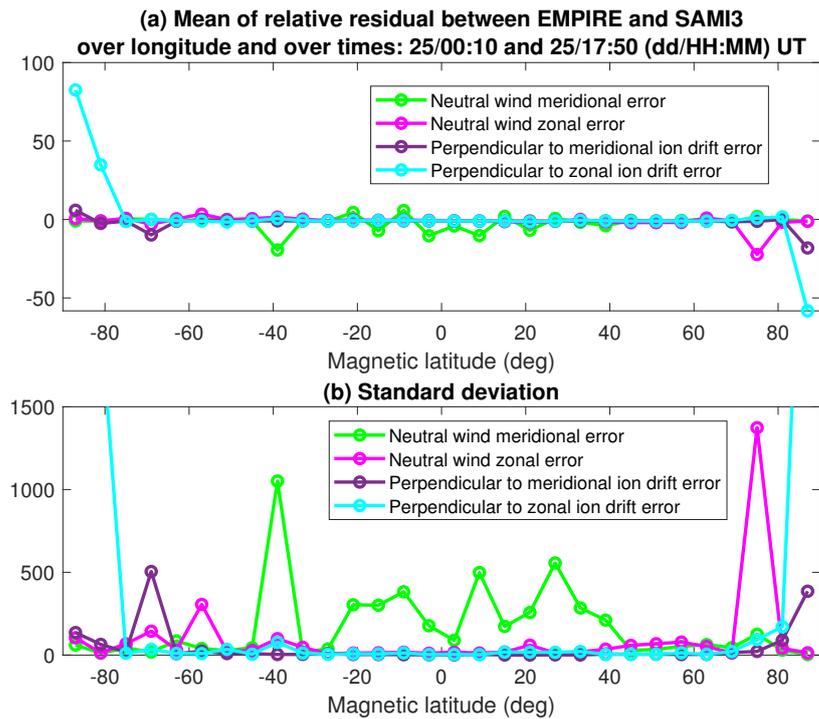


Figure 4.7. (a) Mean and (b) standard deviation of relative residual between EMPIRE estimation and SAMI3 values averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude.

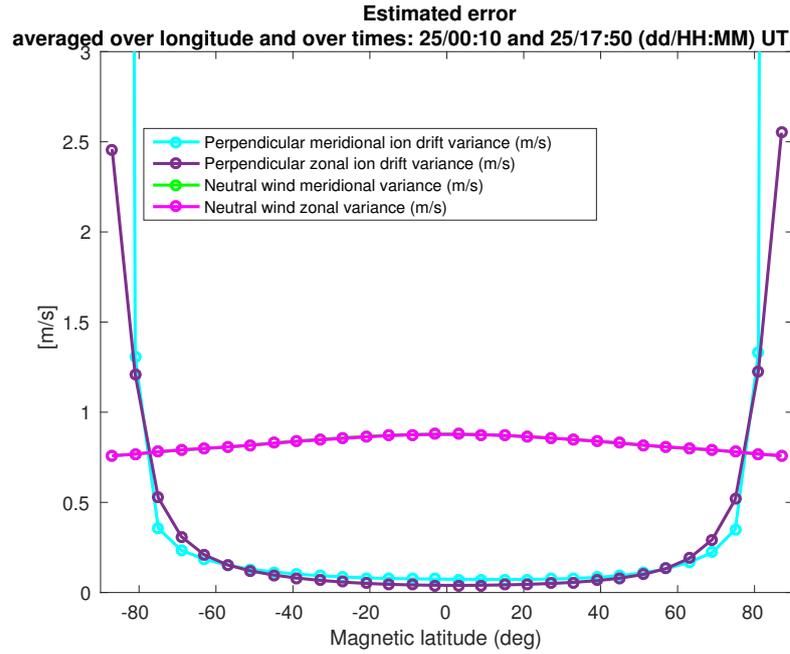


Figure 4.8. Estimated variance in velocity space averaged over longitude and quiet time for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction plotted over magnetic latitude.

standard deviation of the meridional wind error is higher than any of the others components at low and mid latitudes, indicating a higher disagreement between the EMPIRE estimates and the SAMI3 values. This may be because the meridional winds SAMI3 values are small and are dividing the residual, increasing the relative error. At high magnetic latitudes, the standard deviation of the drivers increases too.

For all the estimated velocity components, the covariance error matrix  $\mathbf{P}_{t/t}$ , from Equation 3.33, is analyzed by studying the mean of the estimated error over longitude and time  $\mu_\sigma$ , from Equation 4.8. Results are shown in Figure 4.8 for the ion drifts in the perpendicular meridional (purple) and zonal (cyan) direction and for the neutral winds in the geographical meridional (green) and zonal (magenta) direction.

The variance of the estimated velocity components, for geographic zonal and

meridional neutral winds and perpendicular meridional and zonal ion drifts, is small for all geomagnetic latitudes. Neutral meridional and zonal variance are almost the same, so the green meridional is not seen in Figure 4.8. From Equation 3.33, it is observed that the estimated covariance matrix depends on the forecast covariance matrix  $P_{t/t-1}$  and on a second term that contains the observation covariance matrix  $\mathbf{R}$ . The forecast covariance matrix, defined in Equation 3.23, is small. It depends on the background variance,  $\sigma_{i,0}^2 = \sigma_{i,fitting}^2 + \sigma_{i,climate}^2$ , and we have assumed that the climate contribution is negligible. On the other hand, the observation covariance matrix  $\mathbf{R}$ , that is calculated using the errors of the electron density  $N$  and the electron density terms not estimated  $a_0$  (production, loss, diffusion and gravity), is also small. We are assuming that SAMI3 is the “truth” so that the errors of its outputs are hypothesized to be small. Both contributions have the same order so the resultant time updated covariance matrix  $P_{t/t-1}$  contains small variances too and consequently so does the error of the estimated velocities at each grid point over time. The estimated variance is smaller than the representation error for all the velocity terms studied, but they do not need to be the same in the corresponding space [44], that is why it is important to study it separately.

#### 4.4 Conclusions

We have conducted a study to quantify the representation error in EMPIRE estimation of ionospheric drivers during quiet time. Neutral winds and ion drifts have been estimated with EMPIRE ingesting all inputs, usually measurements and background model data, from the same self-consistent source. SAMI3 is the background model used for this analysis as well as the observation source, and truth comparison. The goal has been to analyze the uncertainties of the EMPIRE model: study how precise the model is in estimating the ionosphere with just the ion continuity equation from O+, compared to a more complete ionospheric model, SAMI3.

We investigated the ionosphere behavior during a quiet day on 25 August 2018, so that all the inputs in EMPIRE are self-consistent with SAMI3. The results from EMPIRE have been compared to SAMI3 model values, as the estimated results were expected to be close to the SAMI3 results. The representation error of each of the drivers is analyzed. We show that EMPIRE is able to estimate the neutral wind and ion drifts with smaller errors at low and mid geomagnetic latitudes. We also show that the representation error of EMPIRE has a similar order of magnitude to the fitting error from the mapping matrices used. The simplification of only using the ion continuity equation is good enough as the main component of the representation error is due to the fitting done to model the neutral winds and potential field with basis functions. The estimated variance being small agrees with the Kalman filter formulation and the hypotheses done for this study. It does not contain the representation error, that is why it is important to do this study and analyze it separately. We conclude that during storm time this error analysis works less well as the estimation values for both drivers differs from the model SAMI3 outputs. We were also able to characterize the error of these estimations, that can be used for future runs on EMPIRE as the representation error of the algorithm by transforming the errors from the drivers space to the observation state.

This analysis helps us understand the strengths and weaknesses of EMPIRE. However, these representations errors may not be the best case scenario achievable by EMPIRE, as during a quiet day the electron density rate, which is the main input of EMPIRE, is not as large a signal as during geomagnetic storms. But if we have chosen a geomagnetic storm for the analysis, outputs from background models are generally less reliable and they might not be a self-consistent source of “true ionosphere” for EMPIRE. We would not expect our estimations to be close to the model results. For future work, one way of calculating the “truth” could be through an error analysis with EMPIRE changing the time step size until it converges to the same value and

use that as the “truth.” Also, in the future we expect to use EMPIRE to study geomagnetic storms understanding how much we can trust the results: reasonably in the low and mid latitudes with appropriate weights chosen based on this analysis.

## CHAPTER 5

### INVESTIGATION OF NIGHTTIME IONOSPHERIC LOCALIZED ENHANCEMENT

In the previous chapter the global estimation and the assessment of the error of the IT drivers were studied. These analyses allow us to study global geomagnetic storms with EMPIRE globally and to understand how well the drivers are going to be estimated.

A Nighttime Ionospheric Localized Enhancement (or NILE) of the electron density has been observed at northern mid-latitudes during the recovery phase of some geomagnetic storms in the ionosphere. In all cases observed to date, the NILE appears to originate above the Caribbean and sometimes extends into the continental USA. The TEC maps during this phenomena indicates that the plasma extends northwestward, and it is believed to be due to an enhanced eastward electrojet, that increases the effects of the equatorial ionization anomaly [99]. However, geomagnetic storms' high energy injection can provoke drastic changes that the NILE may originate from something different, like neutral winds.

In recent studies [48], the event was studied and they show that the NILE appears to be originated from the enhancement of the equatorial ionization anomaly (EIA) in the recovery phase of the geomagnetic storm. This “anomaly”, which happens always in quiet and storm time, occurs during daytime and dawn, when the sunlight hits the Earth heating the atmosphere and consequently causing a thermal expansion. Near the magnetic equator, the ions are displaced to higher altitudes, followed by an increase in the electron density. The ions are also displaced poleward, at quiet times to latitudes up to 10 deg.

In [48] some questions were raised. They concluded that the NILE appears to be driven by ion drift rather than neutral winds. They [48] also question why

only the northern part of the EIA, from which the NILE appears to be originated, persists into the night and why the southern part does not. Another objective of the contribution is to try to answer this question by studying the estimated global neutral winds with EMPIRE, specifically the meridional direction of the winds. As summarized in Section 2.5.3, [80] explains that the EIA asymmetry is related to the meridional direction of the winds and the strength of the fountain effect, if there is one.

The enhancement of the EIA can be caused by different effects during geomagnetic storms like the enhancement of the equatorial fountain effect due to PPEF (prompt penetration of the electric field) at high latitudes or to ionosphere disturbance dynamo (caused by Joule heating), as explained in the Section 2.5.3. Understanding these physical effects would help to model and simulate the NILE event to be taken into account in the future.

The objective of this contribution is to investigate the NILE event and the EIA asymmetry, so the meridional winds and the fountain effect, and consequently the ion drifts, are going to be studied and estimated with EMPIRE. Two different possible NILE events are going to be studied, the same ones previously studied in [48].

In addition, for one of those storms, additional FPI neutral winds measurements are going to be ingested with EMPIRE, using the new VSH expansion for the neutral winds. Besides investigating the NILE event during this storm, we also want to analyze if the ingestion of neutral winds with the VSH derivation from Chapter 3 improves the estimation of the winds, as it did with the old implementation as shown in [32].

In the first Section 5.1 the ingestion of FPI measurements to EMPIRE using

the VSH expansion will be explained. Then, Section 5.2 describes the first storm studied, the EMPIRE algorithm configuration used and the results. In Section 5.3 the second storm is introduced and the EMPIRE estimates are analyzed for this case. Finally, in Section 5.4 we will conclude the NILE analysis.

### 5.1 Ingestion of FPI using VSH

EMPIRE is modified to ingest FPI measurements of neutral winds using the VSH expansion and the direct estimation of the drivers explained in Chapter 3. The ingestion of measurements of neutral winds from FPI was already implemented in previous work [32] and summarized in 2.2.2. An additional  $i$  equations are augmented to the EMPIRE system defined in Equation 3.20 to assimilate wind measurements. Similarly to the neutral wind measures described in Equation 2.32, we impose:

$$y_{FPI} = u_{FPI,i} \quad (5.1)$$

Where  $y_{FPI}$  represents the LOS neutral wind measurement from FPI instruments at each  $i$  location. The neutral winds  $\vec{u} = [u_E \ u_N \ 0]^T$  VSH expansion defined in 3.10 is projected to the LOS direction of each of the measurements:

$$u_{FPI,i} = \mathbf{f}_{LOS,i} \vec{u} = \mathbf{h}_{FPI,i} \mathbf{x}_u \quad (5.2)$$

Matrix  $\mathbf{f}_{LOS,i}$  is defined in Equation 2.34 and mapping matrix  $\mathbf{h}_{FPI,i}$  is obtained by substituting the VSH expansion. The whole derivation of the matrix can be found in Appendix D.

To add the new ingested measurements to EMPIRE, we change notation of Equation 5.1 to:

$$z_{FPI} = y_{FPI} - \underbrace{a_{0,FPI}}_0 = \underbrace{a_{FPI,i}}_{u_{FPI,i}} \quad (5.3)$$

Stacking all the  $i$  equations and substituting the expansion of  $u_{FPI,i}$  using Equation

5.2, the following equation is obtained:

$$\mathbf{z}_{\text{FPI}} = \mathbf{H}_{\text{FPI}}\mathbf{x}_{\text{u}} \quad (5.4)$$

where the mapping matrix  $\mathbf{H}_{\text{FPI}}$  is obtained by stacking the  $i$  row matrices  $\mathbf{h}_{\text{FPI},i}$  of Equation 5.2. The EMPIRE system, defined in Equation 3.20, is now augmented.

$$\underbrace{\begin{bmatrix} \mathbf{z}_{\text{N}} \\ \mathbf{z}_{\text{FPI}} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} \mathbf{y}_{\text{N}} \\ \mathbf{y}_{\text{FPI}} \end{bmatrix}}_{\mathbf{y}} - \underbrace{\begin{bmatrix} \mathbf{a}_{0\text{N}} \\ \mathbf{a}_{0,\text{FPI}} \end{bmatrix}}_{\mathbf{a}_0} = \underbrace{\begin{bmatrix} \mathbf{H}_{\text{exb}} & \mathbf{H}_{\text{u}} \\ \mathbf{0} & \mathbf{H}_{\text{FPI}} \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} \mathbf{x}_{\text{exb}} \\ \mathbf{x}_{\text{u}} \end{bmatrix} \quad (5.5)$$

State  $\mathbf{x}$  is solved with a Kalman filter, as described in Equations 3.22 to 3.33. It needs to be considered that there are new measurements ingested in the definition of the covariance matrix of the measurements  $\mathbf{R}$ , that appears in the measurement update of the filter. EMPIRE data processing flow chart for this case is shown in Figure 5.1

## 5.2 Storm November 2003

In this section we analyze the neutral winds and ion drifts estimated with EMPIRE for the geomagnetic storm on November 20-21, 2003, when a NILE event has been observed in [48] and analyzed by the plasma distribution.

The Dst index, which is an indicator of magnetic activity, is shown in Figure 5.2 over time for this case. The Dst index crosses the -30 nT boundary approximately at 6:30 UT and continues decreasing, indicating the beginning of the magnetic storm following the classification given by [64]. The NILE for this storm was observed during local nighttime above the Caribbean, from 23 UT to around 4 UT. The Dst index in these times is increasing after its minimum, indicating the recovery phase of the storm.

In Figure 5.3 the TEC maps for different times are plotted. The TEC values

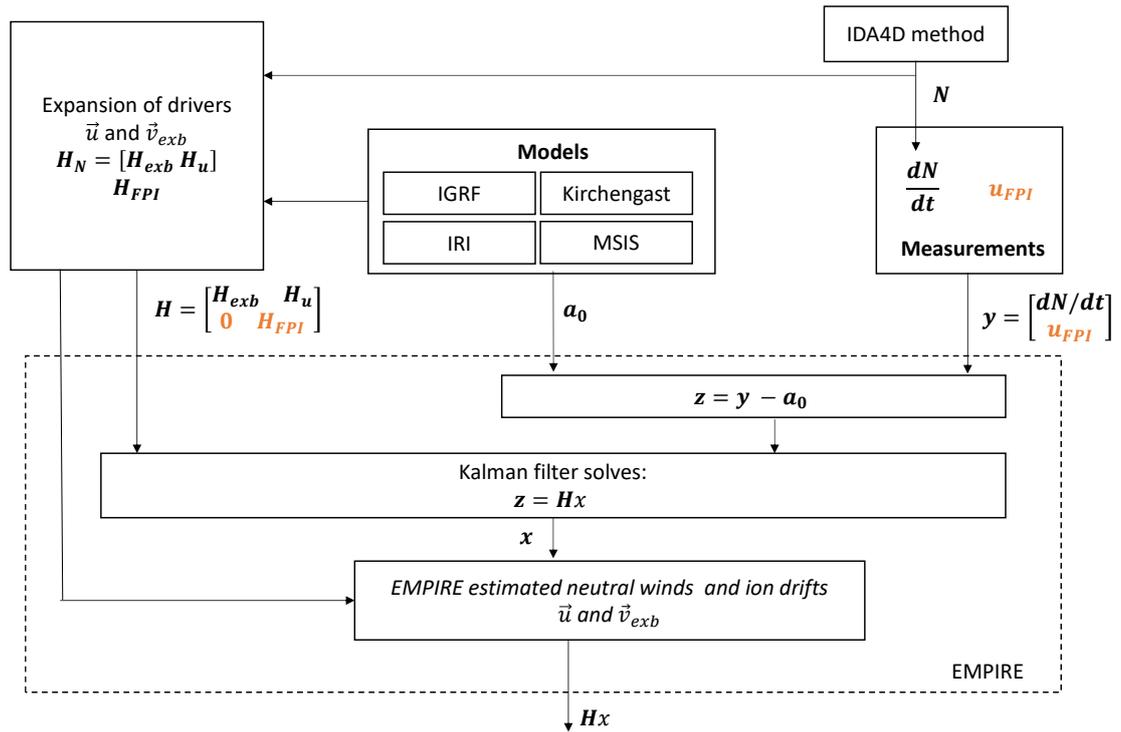


Figure 5.1. Simplified EMPIRE model and data processing flow chart.

are the data ingested in IDA4D algorithm, that outputs the global electron density values that EMPIRE uses as measurements. The TEC maps will also show the NILE event and the conditions of the ionosphere during this time. The TEC values are shown from 20/17 (dd/hh) UT to 21/03 UT (a,b,c,d,e and f) every 2 hours. The day/night terminator is indicated with a black solid line and it specifies local night, that is when the NILE event happens. Local noon is represented with a dashed red line.

The EIA is seen in the TEC maps for this storm. Around the magnetic equator close to the noon, indicated with the vertical red dashed line, at 17 UT (Figure 5.3a) an enhancement of the TEC is visible. We can see that it reaches latitudes of 30 deg in the northern hemisphere. Over this region is local day, that is when the EIA happens.

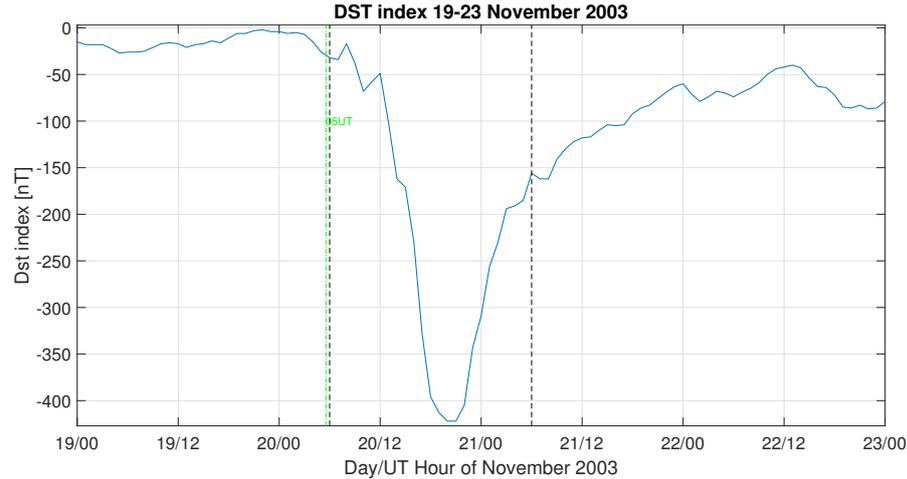


Figure 5.2. Dst index for November 2003 in nT. Studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line.

The EIA is enhanced, as in quiet days it only reaches latitudes up to 10 deg, and it peaks around 21 UT (Figure 5.3c). Then, the EIA enhancement decreases and goes back to normal EIA over time. However, the northern crest of the EIA persists into the night, east of the terminator line in black, as observed clearly at 21 and 23 UT (Figure 5.3c and Figure 5.3d respectively). It remains visible hours after the peak of the enhanced EIA at 21 UT. The NILE appears to originate from this northern TEC crest and it is first observed around 1 UT on the 21st (Figure 5.3e) above the Carribean (approximately at -75 to -60 longitude and 30 degrees latitude). The enhancement occurs during local night time and seems to rotate with Earth, as it extends to continental USA, over Florida, as seen at 3UT on the 21st (Figure 5.3 f). The NILE is observed until approximately 4 UT on the same day. We can also see in Figure 5.3f, when the effects of the enhanced EIA are smaller, the PRE effect. The TEC increases close to the terminator line, represented with a black line.

As mentioned before, the EIA causes an increase in the TEC because the anomaly originates from an upward motion of the ions near the equator (fountain effect). At higher altitudes the recombination rate of the ions is smaller, so following

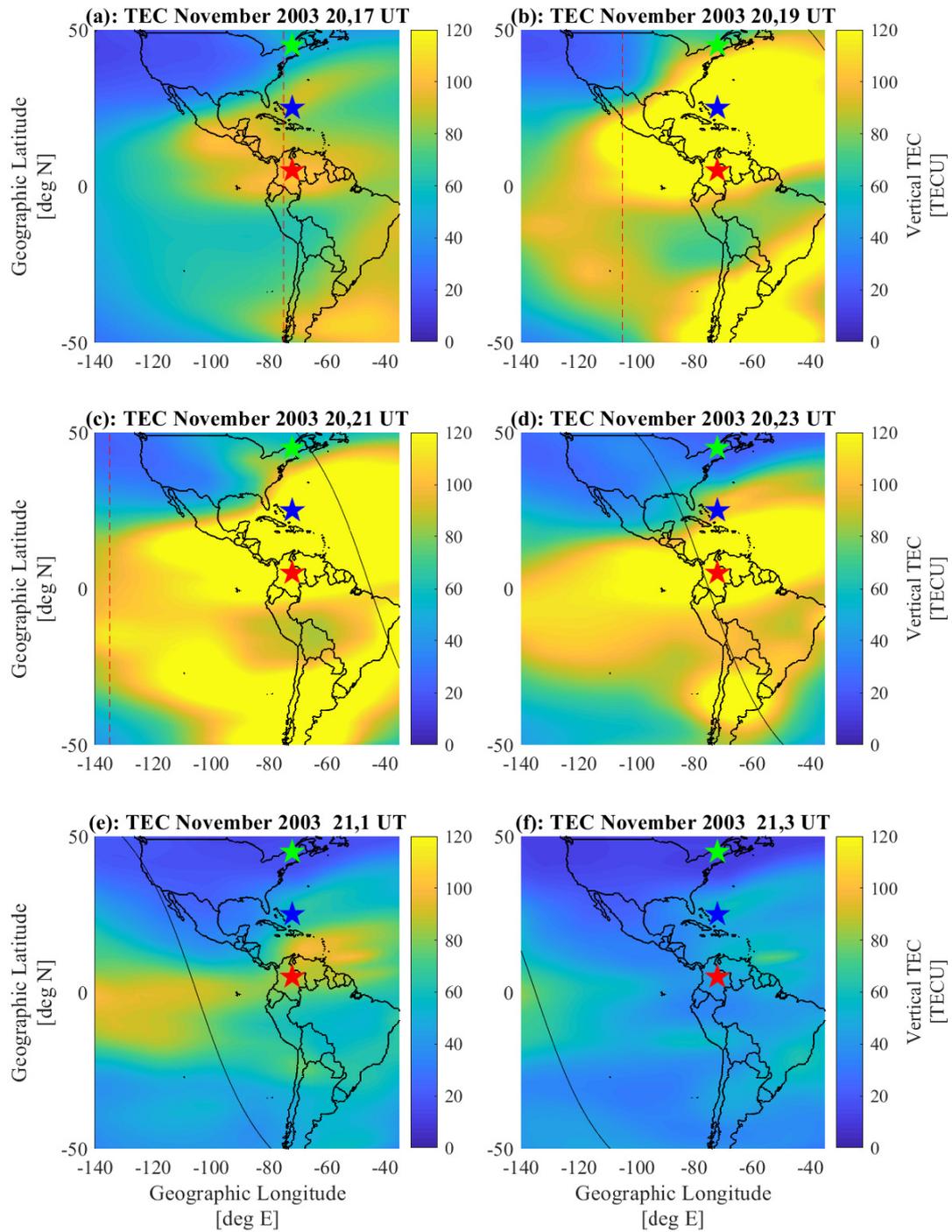


Figure 5.3. TEC maps over time for 20-21 November 2003. Terminator line is indicated with a black solid line and local noon with dashed line. Red star indicates location near magnetic equator, blue where NILE is observed and green Millstone Hill ISR location.

the ion continuity equation, the electron density and the TEC increases. Near the equator, this upward displacement is followed by a poleward motion of the ions, which gives an increased total electron content not only at the equator but at higher latitudes too.

To check the behavior, we are going to analyze in following sections the vertical ion velocity at the 3 different locations marked in Figure 5.3: one near the magnetic equator in the northern EIA crest (red star), one where the NILE is observed (blue star), and at the Millstone Hill ISR location (green star). At the last location measurements of the vertical drift are available so they will be used for comparison with the estimated drifts. EMPIRE is used to estimate the drivers for this storm and its configuration is described below.

**5.2.1 EMPIRE configuration.** In this section I describe how EMPIRE is configured to estimate the IT drivers during the geomagnetic storms selected to study the NILE event. The electron density rate will be derived from the electron density  $N$  from the IDA4D algorithm [100] coupled with the SAMI3 algorithm [49]. IDA4D will correct the data from SAMI3 ingesting mainly measurements of GNSS TEC. Examples of this density data ingested by IDA4D is shown in Figure 5.3. The [86] model is used to calculate the background ion production and loss source terms  $a_{0,prod}$  and  $a_{0,loss}$ . The same model is used to model the gravity and diffusion transport terms effect on the parallel direction of the ion velocity  $a_{0,g}$  and  $a_{0,dfsn}$ . To calculate these terms we use the neutrals properties characterized by the NRL-MSISE00 model [85] and the electron and ion temperature will be provided from IRI model [84]. Finally, the magnetic field  $\vec{B}_0$  is provided by the IGRF model [59].

A global analysis is done, as in the last 2 chapters, by using a grid of magnetic colatitude  $\theta$  between 3 and 177 degrees, magnetic longitude  $\phi$  between -180 and 180 degrees with the same 6 degrees resolution and altitude  $h$  between 200  $km$  and

500 km, with a step of 50 km, following the altitude limits set in [32]. The grid will have 12810 points at each time step. The period starts at 6 UT on the 20th, when the storm starts, with 20-minute increments. The end of the period is a little after the NILE stops to be observed. The analysis period is of 24 hours.

The neutral winds  $\vec{u}$  are estimated using the vector spherical harmonics derivation from Chapter 3 and the potential field is expanded with a spherical harmonics basis function [45]. The order of both expansions is set to  $l_{\max} = 6$ , as determined in previous analysis in Section 3.3.1.2. The state vector size will be  $[90 \times 1]$ , where 48 terms come from the neutral winds expansion and the other 42 from the ion drift derivation.

The background covariance and state are needed for the time update in the Kalman filter in Equations 3.22 and 3.23. The background state for the ion drifts  $\mathbf{x}_{0,\text{exb}}$  is calculated with Equation 3.26 using ion drifts from the SAMI3 algorithm [49] and the background state for the neutral winds  $\mathbf{x}_{0,\mathbf{u}}$  is calculated using Equation 3.27 with neutral winds data from HWM14 model [50]. To calculate the background covariance matrix  $\mathbf{P}_{0,t/t}$  with Equation 3.28, we need to define the error of the fitting  $\sigma_{i,\text{fitting}}$ , the error of the model  $\sigma_{i,\text{model}}$  and a scale factor  $c$  for each of the estimated drivers, as seen in Equation 3.28. The error of the ion drift model is set to  $\sigma_{\text{exb},\text{model}}^2 = (20\text{m/s})^2$  at each of the directions of the vector velocity, as set in previous chapters. Similarly, the HWM14 error is set to  $\sigma_{u_{n,e},\text{model}}^2 = (75\text{m/s})^2$  for both geographical zonal and meridional directions.

For the error of the fitting of the ion drifts we set  $\sigma_{\text{exb},r,\text{fitting}} = (25\text{m/s})^2$ ,  $\sigma_{\text{exb},\theta,\text{fitting}} = (150\text{m/s})^2$  and  $\sigma_{\text{exb},\phi,\text{fitting}} = (140\text{m/s})^2$  in every direction. For the neutral winds, we set:  $\sigma_{u_n,\text{fitting}} = (100\text{m/s})^2$  and  $\sigma_{u_e,\text{fitting}} = (80\text{m/s})^2$ ,  $c_{\text{exb}} = c_{\mathbf{u}} = 2$ . They are selected by doing a fitting analysis with  $l_{\max} = 6$ , similarly to the analysis described in Section 3.3.1.3.

**5.2.1.1 Error of ingested measurements.** The measurement covariance matrix  $\mathbf{R}$  used in the Kalman filter in Equation 3.32 will have size  $[12810 \times 12810]$  and is derived from IDA4D densities  $N$  error and from the errors of  $\mathbf{a}_0$ . The errors of  $\mathbf{a}_0$ , stored in covariance matrix  $\mathbf{R}_{\mathbf{a}_0}$ , are calculated from different climate models, so there is no error available associated. We make the hypothesis of an error of 1% of the actual value, as we are assuming that during storm time the major contributor to electron density change is the ion drifts and the neutral winds.

The error in densities  $N$  from IDA4D correcting SAMI3 are not given for this case. The hypothesis of the error being a percentage of the actual densities observations  $N$  is made. To calculate this percentage we analyze other storms where the error of the densities  $\epsilon_N$  from only IDA4D was given with the data set. The analysis is shown in Figure 5.4, Figure 5.5, Figure 5.6 and Figure 5.7. They show the percentage of the relative error,  $\epsilon_{rel,IDA4D}$  of the electron density  $N$  averaged over different (a) times, (b) colatitudes, (c) longitudes and (d) altitudes:

$$\epsilon_{rel,IDA4D}(r, \theta, \phi, t) = \frac{\epsilon_N}{N} \cdot 100\% \quad (5.6)$$

In Figure 5.4 the results are shown for the data on 25 October 2011. The error can depend for example on the number of sources used in IDA4D or on the intensity of the storm. We will not focus on analyzing the sources or a tendency, we will only obtain an order of magnitude of the IDA4D error to use in EMPIRE when this error is not given. It is seen that the mean error for all the averaged cases indicated with a red line, represented in a dashed red line, is around 40%.

For a second case, on March 17 2015 in Figure 5.5, the mean error averaged for the different times is around 30%, Figure 5.5a, while in the other subplots, for specific latitudes, longitudes and altitudes in 5.5b, 5.5c and 5.5d respectively. The relative error of the electron density is not over 20%.

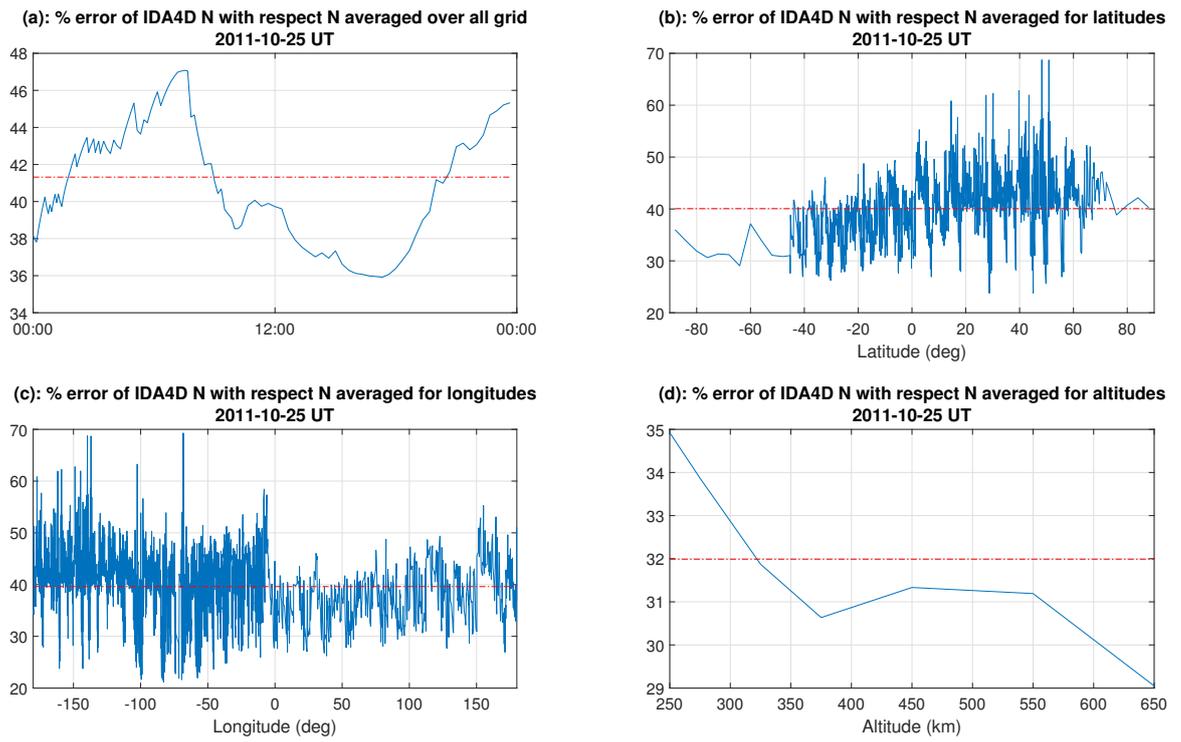


Figure 5.4. Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes. Red dashed line represents the mean value of the case.

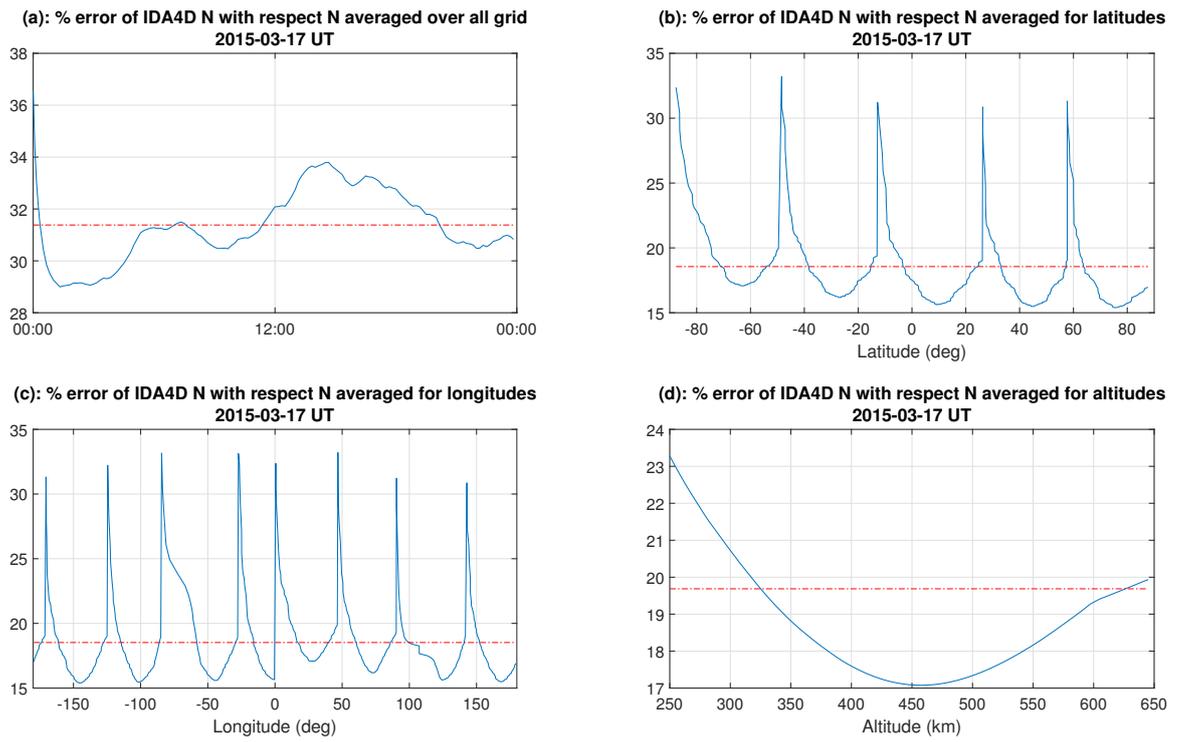


Figure 5.5. Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes. Red dashed line represents the mean value of the case.

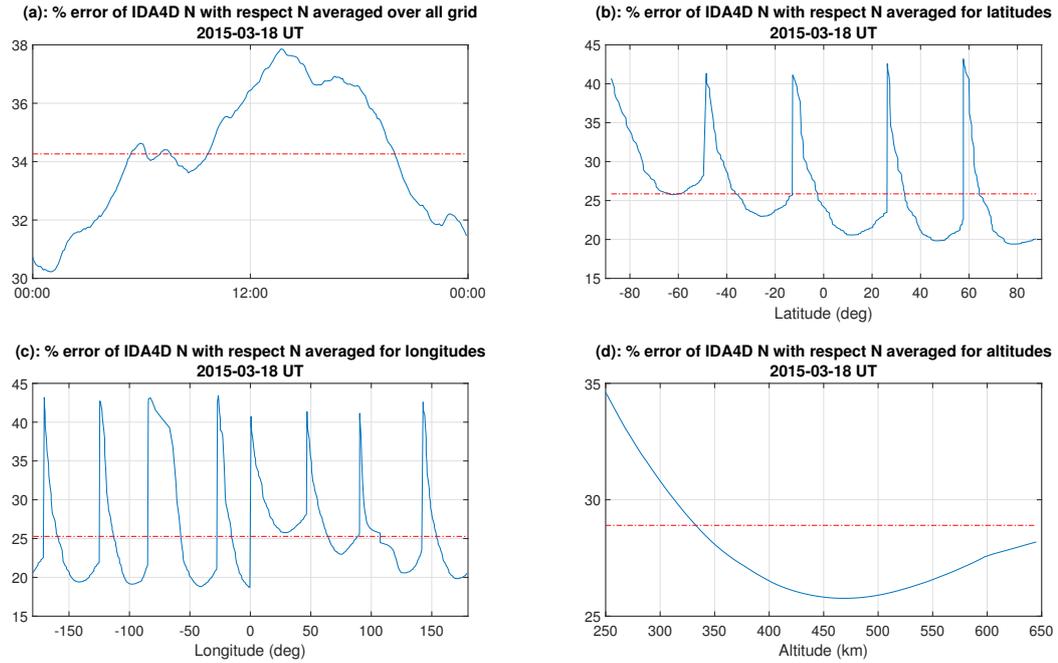


Figure 5.6. Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes in blue. Red dashed line represents the mean value of the case.

IDA4D data for March 18 2015 is also available and the same analysis is done in Figure 5.6. The mean relative error of the electron density is around 30%.

Finally, a fourth case on 4th January 2020 is shown in Figure 5.7. The relative error of the electron density for this day is around 20 – 25%.

Taking into account the 4 cases we have available data of the IDA4D electron densities error, we set an error for the electron density from IDA4D correcting SAMI3 of 35%, which allows us to build the covariance error matrix  $R_N$ .

Once the state is calculated using the described configuration, the estimated neutral winds and ion drifts are calculated in every grid point for the desired period of time. We will show the horizontal neutral winds and ion drifts results over time over the region where the NILE is observed. We will also show the vertical ion drift at the 3 marked locations. First, in the locations indicated in Figure 5.3 near the magnetic

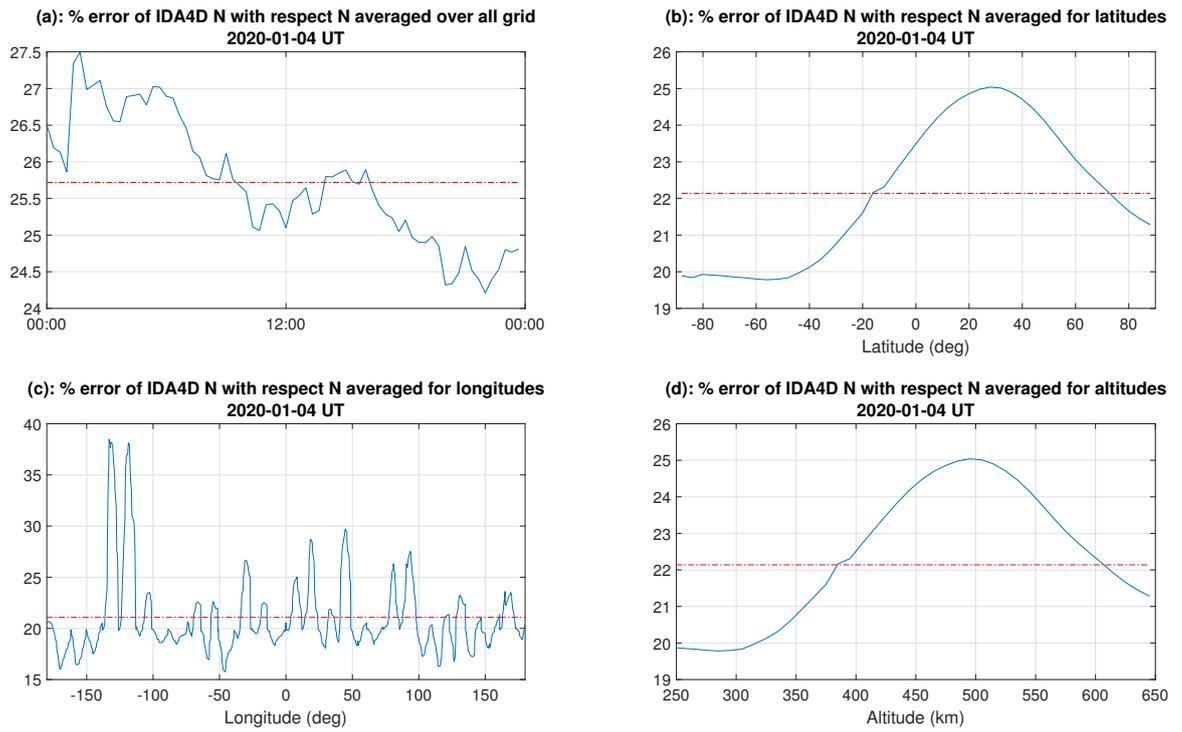


Figure 5.7. Averaged relative error of electron density in percentages plotted for different (a) times, (b) latitudes, (c) longitudes and (d) altitudes in blue. Red dashed line represents the mean value of the case.

equator and near where the NILE is observed. Then, to validate the EMPIRE vertical ion drift, estimates in the ISR Millstone Hill location will be plotted with the actual measurements.

**5.2.2 Results.** In this section the results of the estimated ion drifts and neutral winds are shown. The vertical component of the ion drift will also be plotted to see if the uplift of ions near the magnetic equator cause by the enhanced EIA is visible and to validate the results at ISR Millstone Hill.

**5.2.2.1 Ion drifts.** To analyze the fountain effect and its strength, that affects the formation of the asymmetry of the EIA, the ion drifts are studied. Figure 5.8 shows the estimated ion drifts at 250 km over a map that includes the region where the NILE was seen, above the Caribbean, and the magnetic equator. Each of the maps represents a different time: 17 UT, 19 UT, 21 UT and 23 UT on the 20th and 1 UT and 3 UT on the 21st. The black line represents the terminator line and the red dashed line the local noon. Color on the figures represent the vertical TEC values. Locations at which we are going to study the vertical component of the drifts are marked with a star: green for Millstone Hill, where measurement drifts from ISR are available, blue for a grid point close to where the NILE is observed, and red for a grid point close to the geomagnetic equator.

Around the longitudes where the NILE is observed, between  $-65^\circ$  and  $-70^\circ$ , a poleward ion drift motion is seen from 17 UT (Figure 5.8a) until 23 UT (Figure 5.8d), consistent with the enhanced equatorial ionization anomaly seen at low and mid latitudes at these times. Between 23 UT and 1 UT on the following day (Figure 5.8e) the ion drifts shift to equatorward motion and become much weaker in speed. The enhanced TEC persisting during nighttime at mid latitudes also starts to go to lower latitudes, but the NILE originates from this northern crest and is visible at this time eastwards of Florida, just above the blue star. We can see an increase in

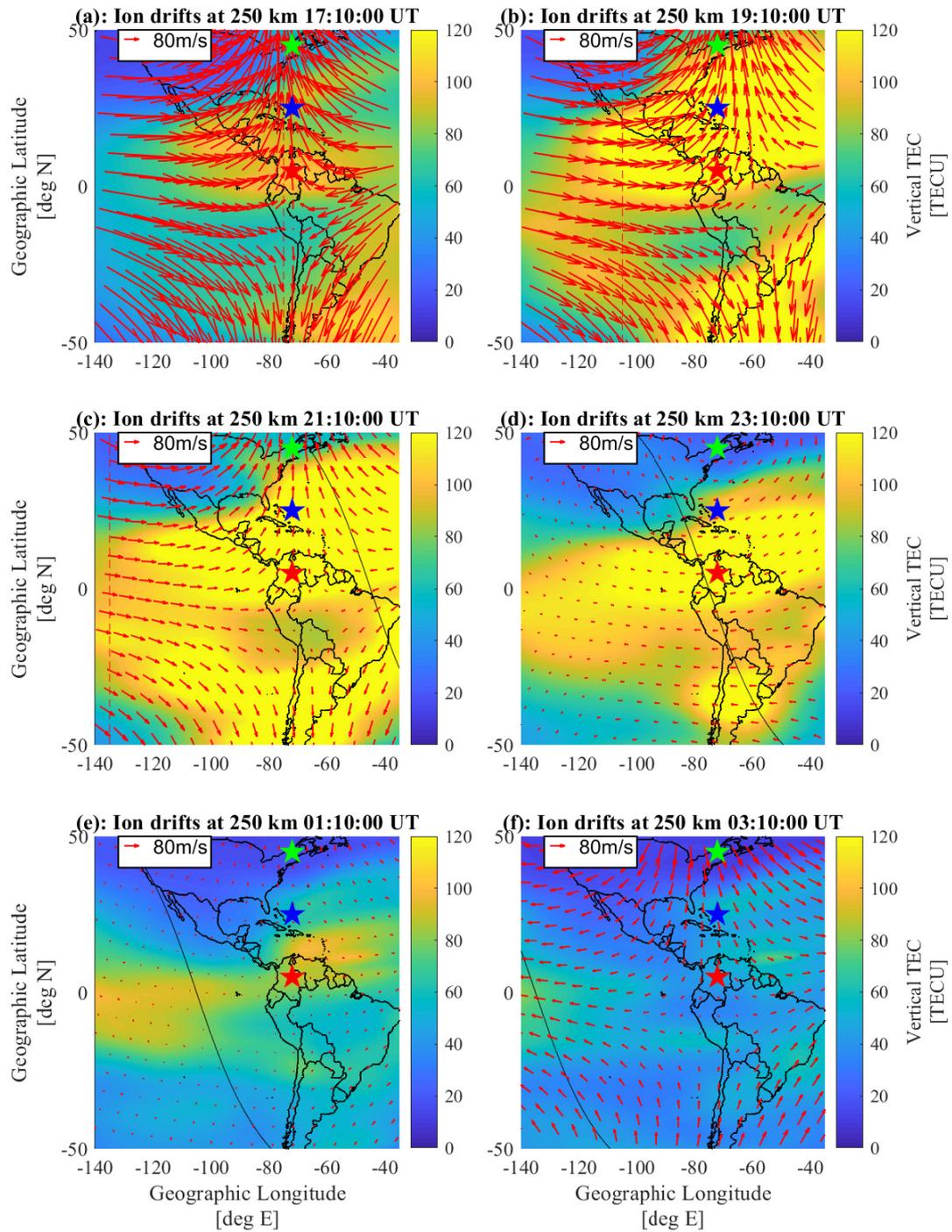


Figure 5.8. Ion drift geographical meridional and zonal direction at height 250 km estimated with EMPIRE represented in red arrows and TEC plotted in color. Terminator line is represented with a black dashed line. Each subfigure represents a different time: (a) 20/17, (b) 20/19, (c) 20/21, (d) 20/23, (e) 21/1 and (f) 21/3 (dd/hh) November 2003 UT. Red star indicates location near magnetic equator, blue where NILE is observed and green Millstone Hill ISR location.

the velocity from 1 UT to 3 UT (Figure 5.8e), as the red arrows are larger. Later, not shown here, ion drifts start to weaken in magnitude, the NILE dissipates and the TEC decreases too, indicating a correlation between ion drift activity and TEC enhancement.

**5.2.2.2 Ion drifts validation.** We also analyze the ion drift vertical component at the locations marked in the maps in 5.8. Results are shown in Figure 5.10 at the following locations: longitude  $72^\circ$  W and latitude  $5^\circ$  N in red near the magnetic equator, latitude  $25^\circ$  N in blue near the zone where the NILE event is observed. In Figure 5.9 results at longitude  $72^\circ$  W latitude  $45^\circ$  N are illustrated in green, which is the location where an ISR instrument with vertical ion drifts measurements for this day is located.

The results of the vertical ion drift are plotted in Figure 5.10 over local time. In red the location near the magnetic equator is shown and in blue the results at the location over NILE. The dashed lines represent the model SAMI3 values and the solid lines the EMPIRE estimation. The black dashed line separates the upward and downward velocity, upward being positive velocities and downward negative ones. It is observed for the 2 locations a high upward velocity from 5 LT on the 20th until 19 LT, with a peak in the positive velocity around 10:30 LT with drifts of  $250 - 300\text{m/s}$ . This is consistent with the EIA physics explained in Section 2.5.3 and with the theory of the superfountain effect, where the F region is lifted close to the equator, consequently causing a poleward motion and a poleward uplift later in time. From that time the velocity is downwards, which is consistent with the decrease of TEC in that location. Compared to the model (the dashed lines) it is observed that EMPIRE is able to estimate the uplift of the plasma layer due to the enhanced EIA. Between 21 and 23 LT, we can see a change in the slope of the velocity, almost becoming a positive slope, that agrees with the more visible NILE in the TEC maps in Figure 5.8. Local sunset

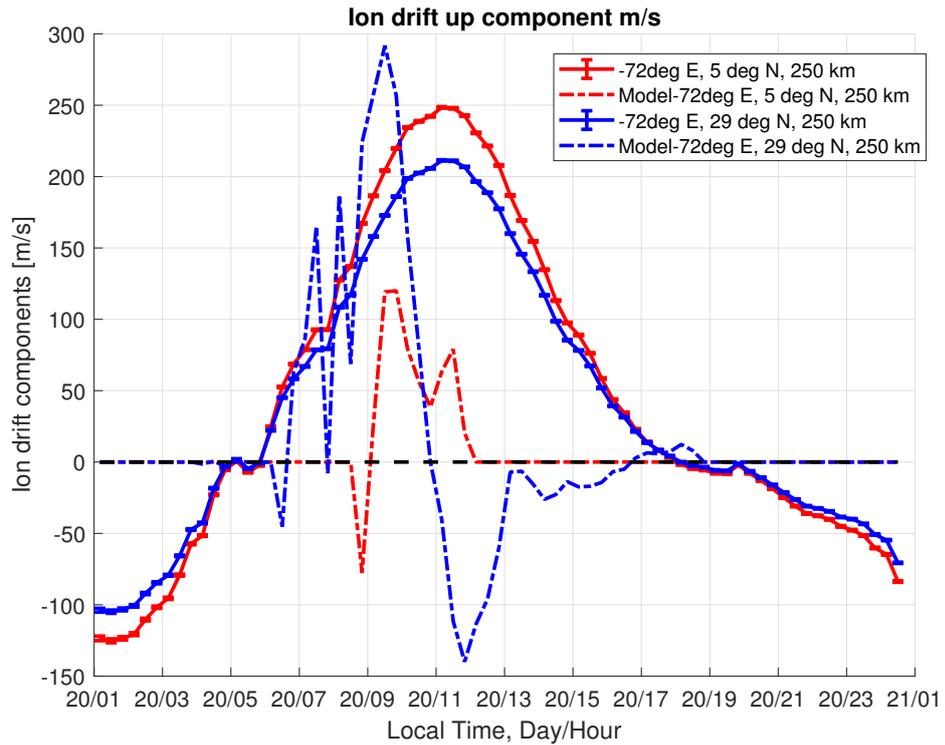


Figure 5.9. Ion drift vertical component (m/s) over time. Solid line represents value estimated with EMPIRE and dashed line the climate model values. In red results at 72 deg W, 5 deg N are plotted (near magnetic equator) and in blue results at 72 deg W, 25 deg N location (over NILE). Black dashed line separates upward (positive) or downward (negative) velocity.

at the 2 locations is around 16:20 LT. Looking at the model values, the dashed lines, we observe an uplift of the layer after that time, more pronounced for the red location (near magnetic equator). This uplift is consistent with the PRE and the polarization terminator effects described in [82]. The estimates are also positive at those times and the slope of the vertical velocity becomes more horizontal.

The error bars of the 2 locations are small compared to the actual values of the ion drift. This may happen because the error imposed for the measurements or the ion drift background values are too high. When this happens, the filter thinks that the estimated values of the ion drifts are within both sources of error, observations and background model, so the estimate error bars are smaller.

Figure 5.9 shows the vertical ion drift at the Millstone Hill ISR location, marked in green in the maps from Figure 5.8.

The EMPIRE results at the location where the Millstone Hill ISR measurements are available are plotted in green in Figure 5.9 with a solid line. The model HWM14 values at the same point are shown for comparison with a green dashed line. The solid black line represents the measurements at Millstone Hill ISR. The behavior of the estimated vertical drift at this higher latitude is similar to the ones described for the other 2 locations (near magnetic equator and near NILE). In comparison to the Millstone Hill ISR measurements in black, we observe that the estimate is able to capture the increase in vertical velocity when the model cannot.

**5.2.2.3 Neutral winds.** Neutral winds are also estimated with EMPIRE and results are plotted in Figure 5.11. They are plotted at altitude of 250 km in a geographical map that includes the region where the NILE was seen, above the Caribbean. Each of the maps represents a different time: 17 UT, 19 UT, 21 UT and 23 UT on the 20th and 1 UT and 3 UT on the 21st. The black dashed line represents the terminator

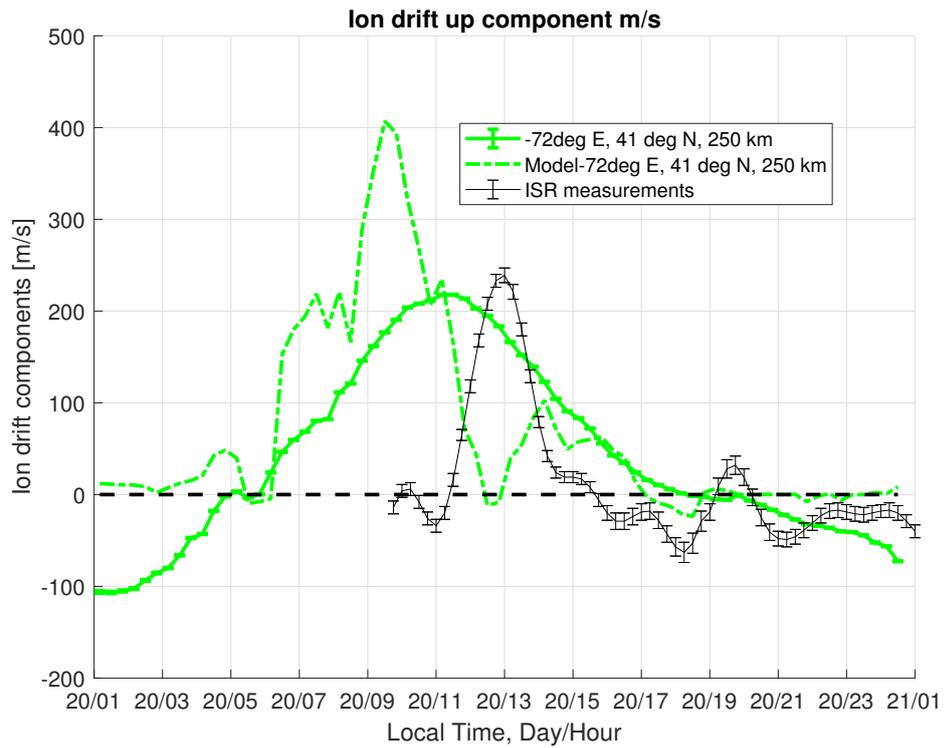


Figure 5.10. Ion drift vertical component (m/s) over time. Solid line represents value estimated with EMPIRE and dashed line the climate model values. In green results at 72 deg W, 55 deg N (ISR location) are plotted. Black dashed line separates upward (positive) or downward (negative) velocity. Black solid line is the ISR measurements at Millstone Hill location.

line. Color on the figures represent the vertical TEC values.

The order of magnitude of the winds is higher than the ion drift and at some points even reaches values of 500 m/s. However, these speeds were observed before in previous analysis of this storm in [101], where they estimated also speeds of 600 *m/s*. It is observed in the southern hemisphere in the first time step, at 17 UT on the 20th in Figure 5.11a, that the neutral velocity in the local morning longitudes of the map (around  $-140$  to  $-100^\circ$ ) is mostly equatorward. Then, when the enhanced EIA is more visible over longitudes between  $-100$  and  $-80^\circ$  at midday, the winds are southward in both hemispheres. Finally, for post-noon longitudes in the map, from approximately  $-80$  to  $-40^\circ$  the winds in the northern hemisphere are eastward while in the southern hemisphere the southwards winds remain. This behavior of the winds is similar until 21 UT, Figure 5.11c, which is the time when the EIA maximum enhancement was found.

Then, from 23 UT in Figure 5.11d, winds weaken. In the longitudes where the NILE is observed ( $-60$  to  $-80^\circ$ ), winds are northward in the northern hemisphere and in the southern hemisphere. Looking at the TEC at this time we can observe a more intense northern crest of the enhanced EIA. We can also see the South Atlantic Anomaly effect on the TEC at latitudes around  $-30^\circ$  and longitudes around  $-70^\circ$ . The asymmetry in the Northern hemisphere produced by mostly southward winds in previous times agrees with the intra-hemisphere transport described in Section 2.5.3, when the fountain effect is stronger than the neutral winds themselves. In the estimated drifts, analyzed in Figure 5.8, we can see strong poleward drifts corresponding to a storm fountain effect, agreeing with the intra-hemisphere transport theory. It is explained by the neutral winds' ability to move ions and electrons along the magnetic field lines and consequently changing the electron density because of the height dependency. At 1 UT in Figure 5.11e winds in these longitudes are still northward.

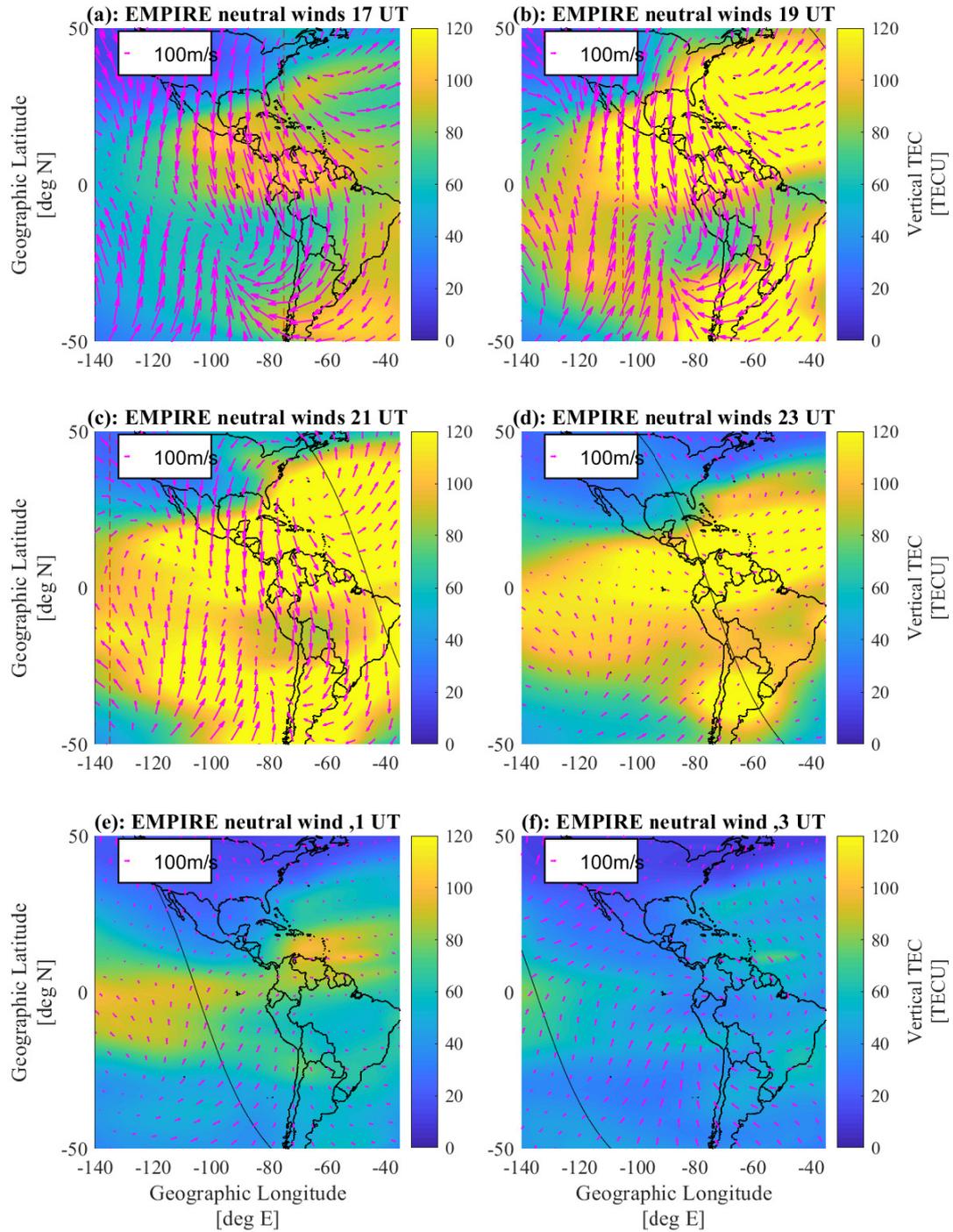


Figure 5.11. Neutral winds geographical meridional and zonal direction at height 250 km estimated with EMPIRE represented in magenta arrows and TEC plotted in color. Each subfigure represents a different time: 20/17, 20/19, 20/21, 20/23, 21/1 and 21/3 (dd/hh) November 2003 UT.

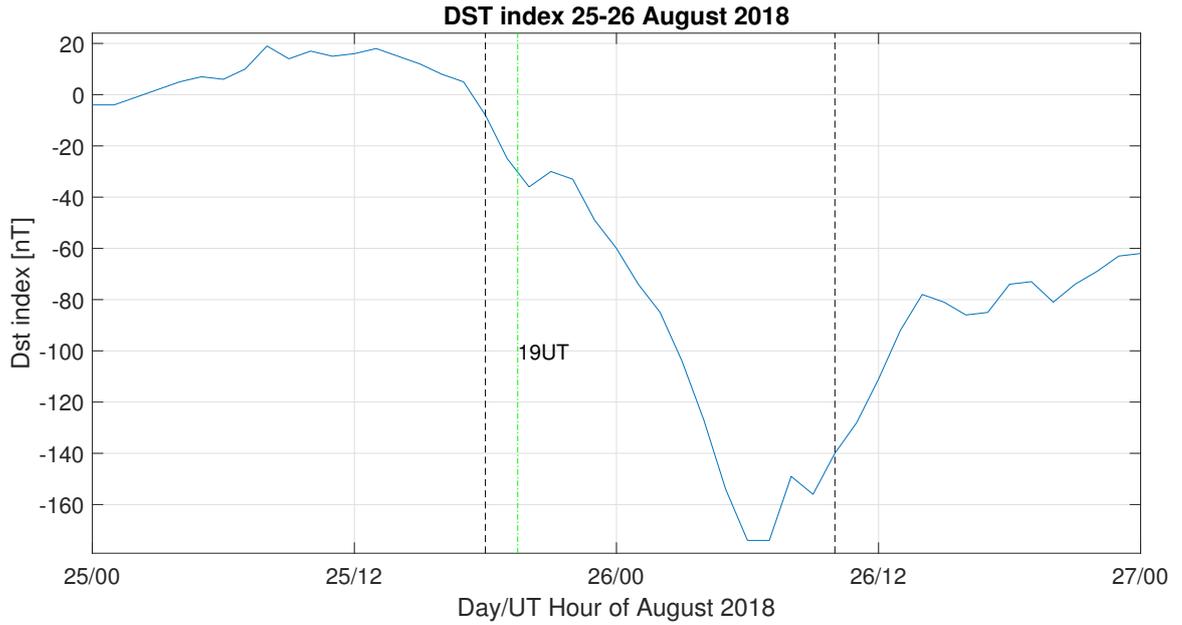


Figure 5.12. Dst index for August 2018 in nT. The studied time limits are indicated with black dashed lines and beginning of geomagnetic storm time is indicated with green dashed-dot line.

Close to the magnetic equator they move to the east. At this time the NILE is visible above the Caribbean. At 3 UT in Figure 5.11f the NILE is still observed and winds over this region are mostly equatorward.

### 5.3 Storm August 2018

The second storm studied where the NILE event was observed is 25-26 August 2018. The Dst index is shown in Figure 5.12 over time.

The minimum order of magnitude of the index is about  $-140$  nT, that is much less extreme than the November 2003 storm index of about  $-400$  nT, on Figure 5.2. This indicates that the August 2018 storm effects are much smaller in magnitude than the November 2003 one. The Dst index crosses the  $-30$  nT boundary approximately at 20 UT, indicating the start of the storm following the classification given in [64]. The storm time also differs with the November 2003, as that one started in the early UT hours. At local nighttime over the zone where the NILE event is observed, from

approximately 23 UT to 4 UT hours, the storm has not reached the minimum Dst index, so the storm is in its main phase.

In Figure 5.13 the TEC maps for different times are plotted showing the evolution of the storm. The TEC values are the data ingested by the IDA4D algorithm, that outputs the electron density values  $N$  correcting the SAMI3 algorithm. The color scale of the TEC is reduced compared to the November storm on Figure 5.3 (TEC goes to 20 TECu vs 80 TECu). The TEC values are shown for the same times, from 25/23 (dd/hh) UT to 26/04 UT (a,b,c,d,e and f). The terminator line is represented with a black solid line.

Similarly to the November 2003 storm, an enhancement of the TEC consequent of the enhanced EIA is visible at 23 UT (Figure 5.13a). The high TEC extends to mid latitudes during the daytime longitudes (before the terminator line). It is also observed that the enhanced TEC from the EIA also persists into the nighttime, duskward of the terminator line. The NILE event can be observed for this storm over the central USA, some degrees in latitude higher and farther west than the feature observed on the November 2003 storm, from 2 UT (Figure 5.13d) to 3 UT (Figure 5.13e). It is seen that it is moving westward, that is the same motion the enhanced EIA has. The NILE feature is smaller and less intense than the one on the November 2003 storm. It is observed that it is originated from the northern EIA crest.

We can see an enhancement of the TEC around the sunset terminator moves with it over time. Besides the terminator line, the enhanced EIA crest is moving parallel to this terminator line, leaving a mid-latitude increase in the TEC over the central USA.

Although the feature on this day is smaller, there is better observational coverage, as there is more GNSS data on 2018 [48] (that is the main input to IDA4D

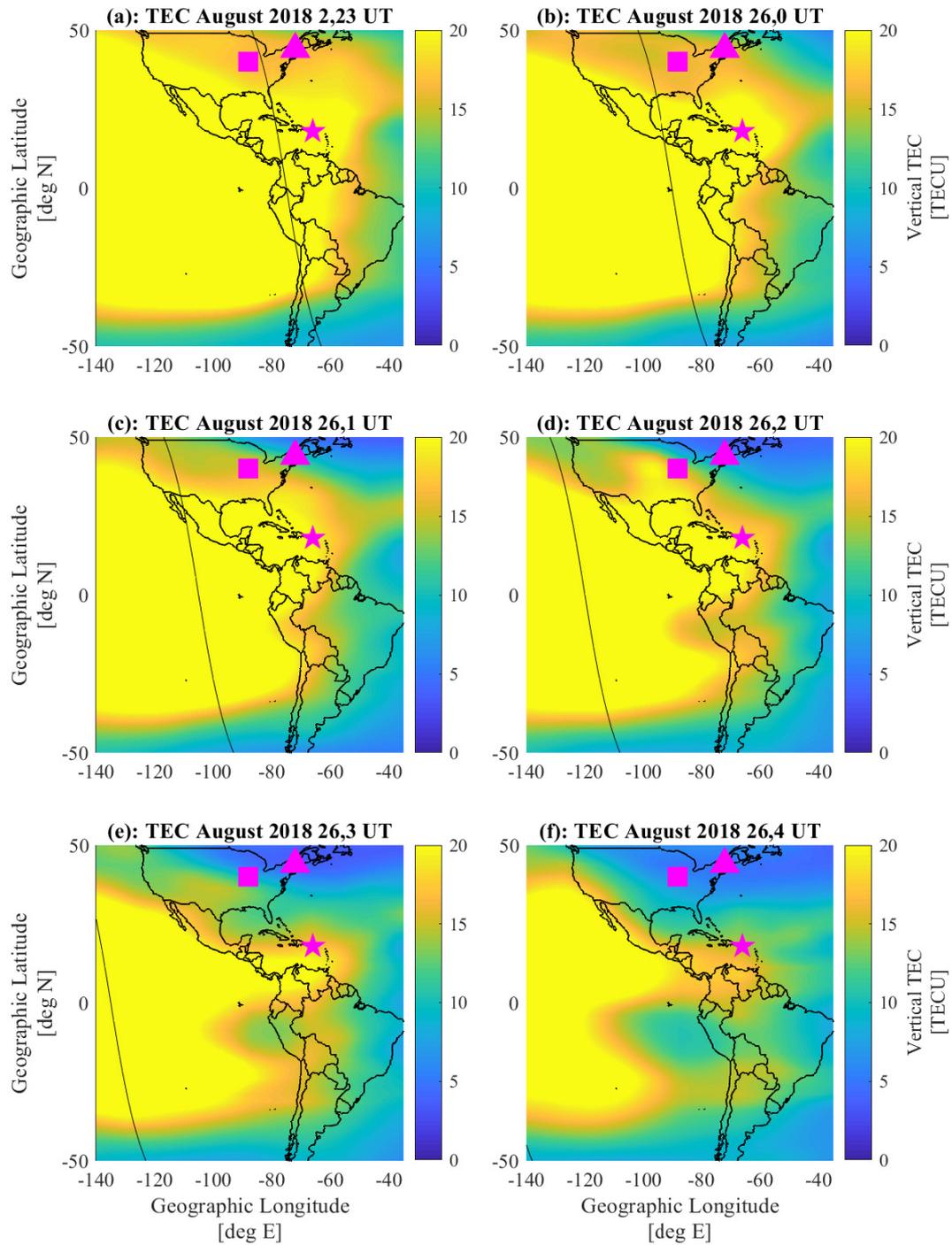


Figure 5.13. TEC maps over time for 25-26 August 2018. Terminator line is indicated with a black solid line. Magenta markers indicate FPI sites where neutral wind measurements were available: Arecibo (star), Millstone Hill (triangle) and Urbana (square).

algorithm, which electron density outputs are ingested by EMPIRE). There were also 3 FPI available that measures neutral winds on this day: Arecibo, Millstone Hill and Urbana. Ingesting these wind measurements additionally in EMPIRE changes the driver estimations significantly [32]. Arecibo ( $18^\circ$  N,  $66^\circ$  W), Millstone ( $44^\circ$  N,  $72^\circ$  W) and Urbana ( $40^\circ$  N,  $88^\circ$  W) site are marked in Figure 5.13 with a magenta star, triangle and square, respectively.

To study the effect of ingesting winds with the new VSH derivation from Chapter 3, I will run EMPIRE for 2 different cases: with additional ingestion of measurements from Urbana FPI, from now on “Case ingestion”, and with no ingestion of additional winds, from now on “Case no ingestion”. We select the Urbana location to ingest winds in the “Case ingestion” as it is the closest location to where the NILE is observed for this day. The other sites’ winds will be used for validation and compared to the EMPIRE estimations of the neutral winds. Selecting to ingest Urbana winds, will also help us understand the effect of the distance between the grid point where additional winds are ingested and the grid point studied on the estimations.

As in previous storms analyzed, the FPI measurements are given in different LOS directions of the instrument. The estimated geographical neutral winds will be projected to this LOS to compare the results to the measurements. For all of the FPI sites there are measurements available in the 4 geographic directions (north, south, east and west). As explained previously, the pierce point of the LOS is not the same as the location of the sites. However, because the separation between the site location and the observation pierce point at 250 km is less than 2.5 degrees, the grid will not distinguish between these two points, as the EMPIRE grid resolution used is 6 degrees. We will use the site location as the point there we have the neutral wind data.

**5.3.1 EMPIRE configuration.** The EMPIRE configuration is similar to the one

described for the November storm on Section 5.2.1 for both cases. The differences with that case are described here.

In the “Case ingestion”, the number of grid points will be augmented to  $12810 + i$ , depending on how many measurements available  $i$  are in each time step. However, number of neutral winds  $i$  is much smaller than the number of measurements of  $N$ , 12810. EMPIRE will ingest northward and eastward measurements from Urbana FPI, as they are more reliable than south and west LOS, as explained in Section 2.2.2. The analysis period for both cases is of 14 hours starting at 18 UT on the 25th with 20-minute increments.

The variances needed to define the Kalman filter, Equations 3.22 to 3.33, are set for both cases as:  $\sigma_{exb,r,fitting} = (20m/s)^2$ ,  $\sigma_{exb,\theta,fitting} = (20m/s)^2$  and  $\sigma_{exb,\phi,fitting} = (10m/s)^2$ . For the neutral winds, we set:  $\sigma_{u_n,fitting} = (75m/s)^2$  and  $\sigma_{u_e,fitting} = (75m/s)^2$ . The scale factor  $c$  for both drivers is set to  $c_u = 2$  for the neutral winds and  $c_{exb} = 6$  for the ion drifts, to take into account the propagated error of  $\mathbf{H}$  due to the error of the electron density ingested  $N$ .

The measurement covariance matrix  $\mathbf{R}_N$  used in the Kalman filter in Equation 3.32 for “Case ingestion” needs to take into account that the system has been augmented to ingest FPI measurements of neutral winds. It is a diagonal matrix with size  $(12810 + i)$  by  $(12810 + i)$ , where  $i$  is the number of north and east measurements available from Urbana FPI at each time step. Number of terms  $i$  can change in each time step. The first part of the matrix will be set up the same as for the November storm. The rest, the  $[i \times i]$  terms left, will be obtained from the error of the FPI measurements.

Once EMPIRE is configured, the state is calculated for both cases, solving the systems shown in Figure 5.14.

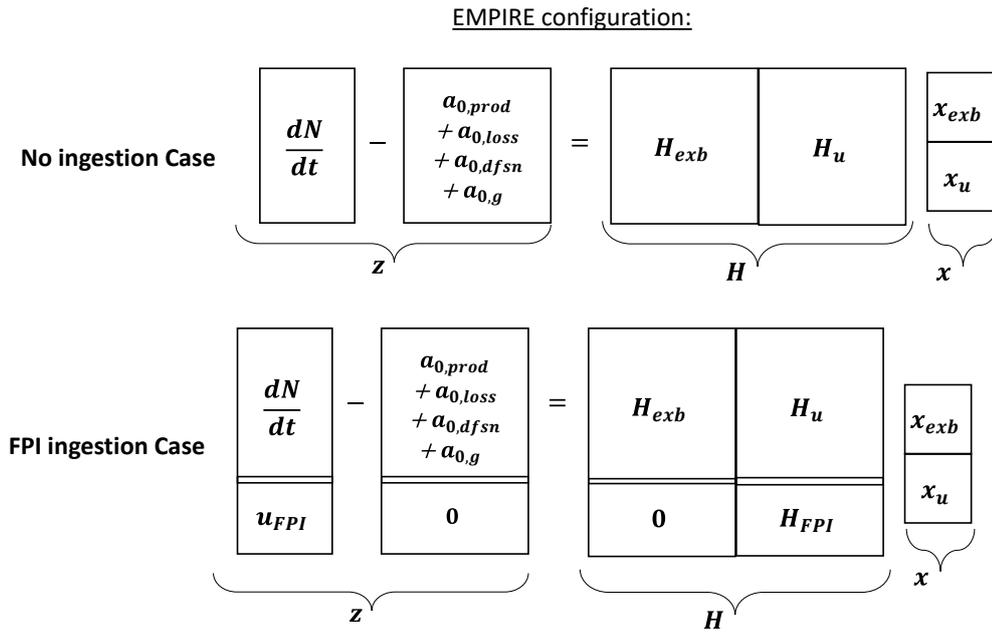


Figure 5.14. Diagram of the EMPIRE configuration for each of the studied cases.

The global ion drifts and neutral winds for the “Case ingestion” will be analyzed to investigate the EIA and its origin. Then, to validate the results and study the impact of ingesting winds, neutral winds from both cases, ingestion and no ingestion, at 3 FPIs, marked in Figure 5.13, will be studied over time: Urbana FPI, from which we are ingesting the measurements in “Case ingestion”, and Millstone Hill FPI and Arecibo FPI whose measurements are not assimilated. The estimates will also be compared to the measurements at these sites.

**5.3.2 Results.** Analysis of the August 25-26 storm is shown in this section. Ion drifts and neutral winds for the “Case ingestion” will be shown and validation of the neutral winds will be done by comparing the results at 3 FPI sites with the FPI measurements and the estimates of the “Case no ingestion”.

**5.3.2.1 Ion drifts.** First, the ion drifts are studied in Figure 5.15. The EMPIRE

estimation of the ion drifts at 250km is plotted with red arrows over a map for different times: 23 UT on the 25th and 0 UT, 1 UT, 2 UT, 3 UT and 4 UT on the 26th. The black line represents the terminator and the color on the figures represents the vertical TEC values. As before, the limits of the vertical TEC values changes with respect the 2003 November case from 120 TECu to 20 TECu.

In this case, the storm starts later in the day, around 19 UT. In the November storm, besides being stronger, the storm starts earlier in the day at 5 UT. We can observe west and equatorward ion drifts at 23 UT on Figure 5.15a over the NILE longitudes (around  $-70^\circ$ ). At  $-140^\circ$  we can observe that the drifts are slightly equatorward. At 0 UT on the following day, Figure 5.15b, the directions of the drifts are similar but with weaker magnitudes. From 1 UT, represented in Figure 5.15c, the drifts over the NILE longitudes become equatorward. At the same time the enhanced EIA at midlatitudes starts to move back to lower latitudes, as we can observe at 2 UT in Figure 5.15d, which is consistent with the poleward motion. At 3 UT, Figure 5.15e, the northern crest of the EIA persists and the NILE appears to originate from it. In this case the NILE is observed over the continental US. The drifts are still equatorward and weaker.

**5.3.2.2 Neutral winds.** The neutral winds are also estimated with EMPIRE and the results can be seen in Figure 5.16. They are also plotted at an altitude of 250km in a geographical map for different times: 23 UT on the 25th and 0 UT, 1 UT, 2 UT, 3 UT and 4 UT on the 26th. Again, the black line represents the terminator and the color represents the vertical TEC values. FPI sites are marked with a magenta star (Arecibo), triangle (Millstone Hill) and square (Urbana Atmospheric). The Urbana measurements of neutral winds were ingested for the EMPIRE estimation and the measurements from Arecibo and Millstone Hill will be used for validation.

We can see that at 23 UT, in Figure 5.16a, winds are poleward in the local

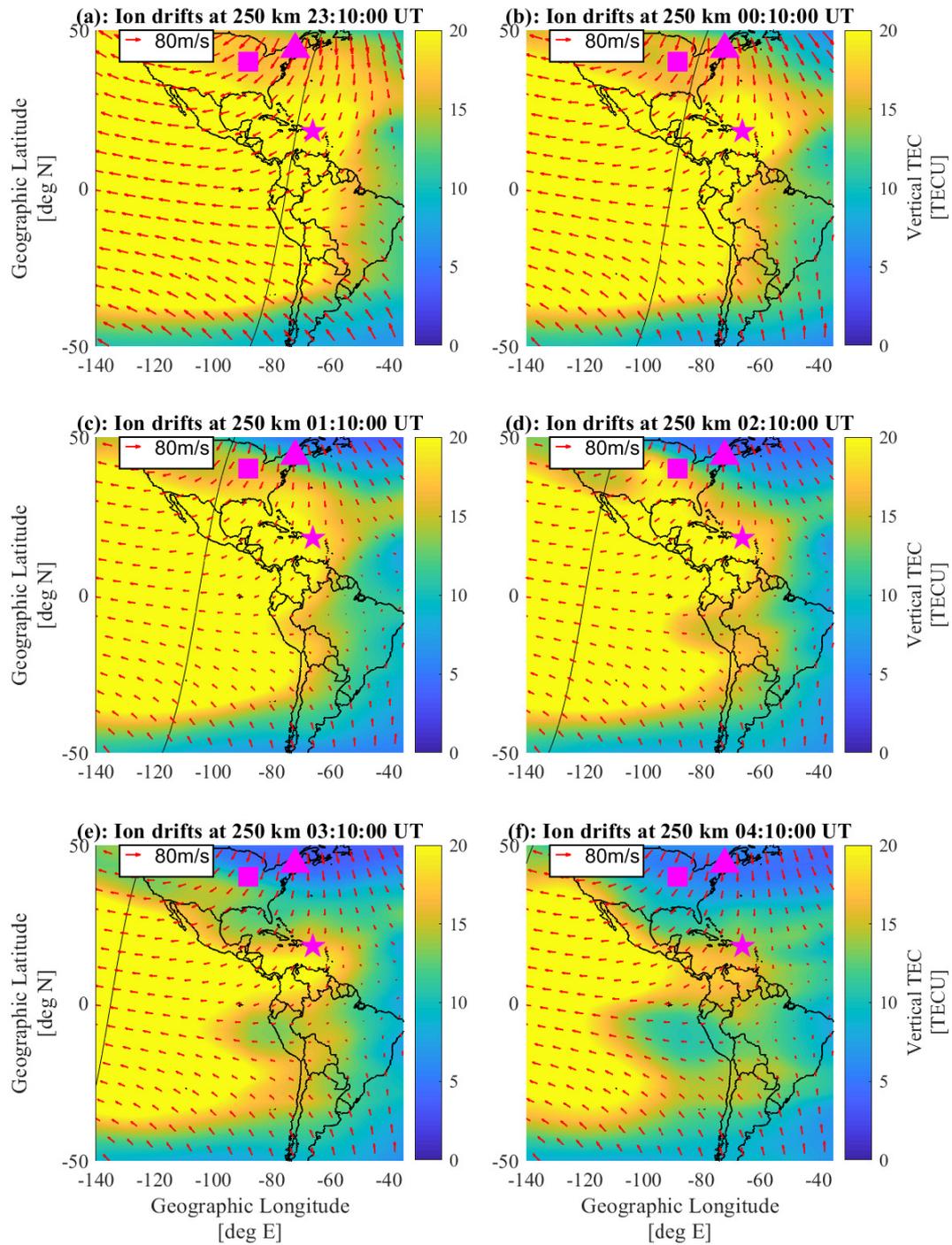


Figure 5.15. Ion drift geographical meridional and zonal direction at height 250km estimated with EMPIRE represented in red arrows and TEC plotted in color. Terminator line is represented with a black dashed line. Each subfigure represents a different time: 25/23, 26/00, 26/01, 26/02, 26/03 and 26/04 (dd/hh) August 2018 UT.

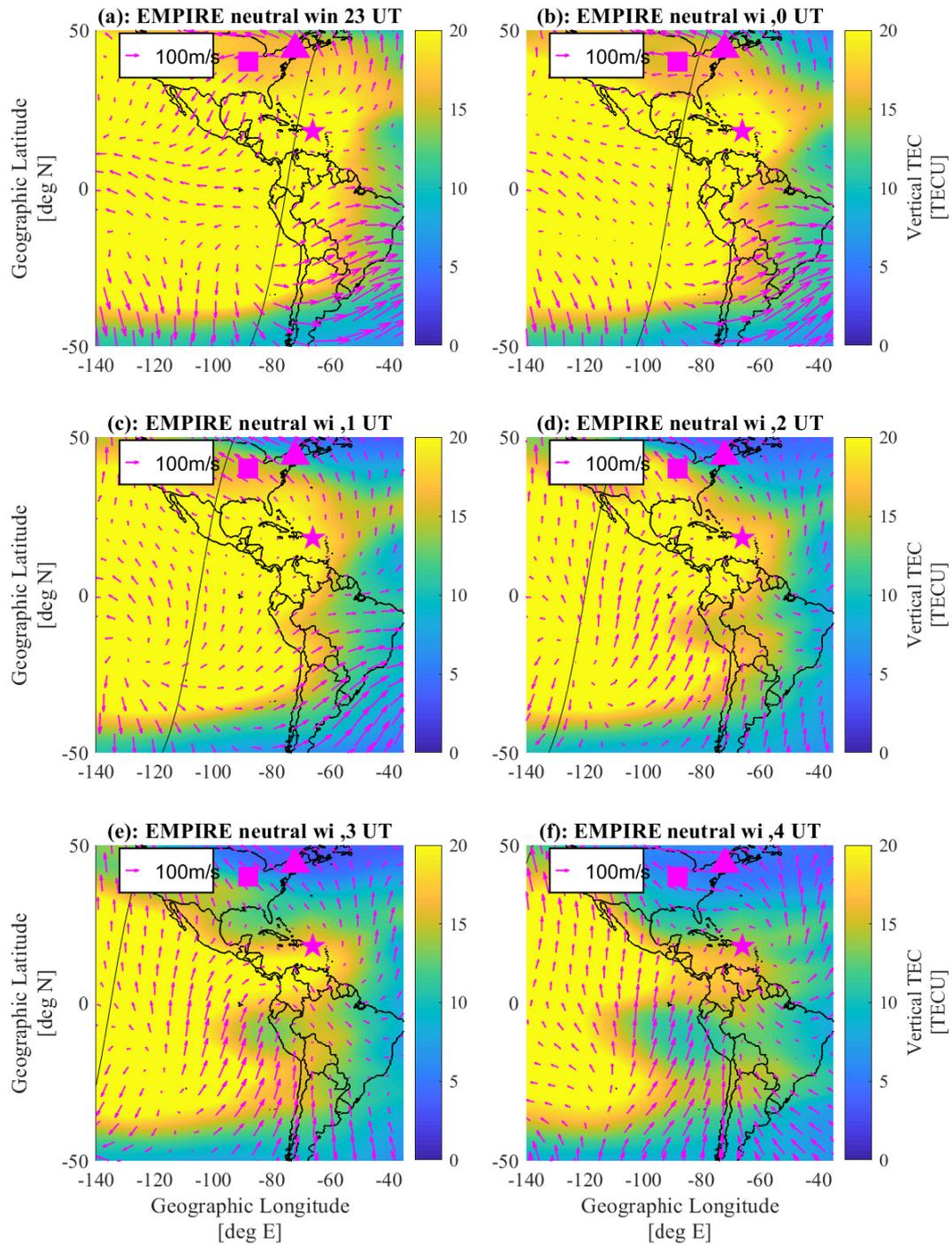


Figure 5.16. Neutral winds geographical meridional and zonal direction at height 250km estimated with EMPIRE represented in magenta arrows and TEC plotted in color. Each subfigure represents a different time: 25/23, 26/00, 26/01, 26/02, 26/03 and 26/04 (dd/hh) August 2018 UT.

day (before the terminator line) in the southern hemisphere and they are west and poleward in the northern hemisphere. On local night longitudes the winds in the northern hemisphere are poleward and in the southern they are mostly eastward and a little bit northward. At 0 UT and 1 UT, Figure 5.16b and Figure 5.16c, neutral wind behavior in both hemispheres is similar, although the poleward direction at the south hemisphere starts to dissipate and a more northern direction in the winds appears over all longitudes in the north hemisphere and at the local night in the south hemisphere.

The enhanced northern crest producing an asymmetry in the enhanced EIA, that has persisted into the night, combined with northern winds agrees with the trans-hemisphere theory described in Section 2.5.3. For that to happen, the fountain effect should be weaker than the winds, so that northern winds, such as we observe over the NILE longitudes from 0 UT, move the plasma to northern latitudes and higher altitudes, increasing the TEC in the north hemisphere. In the estimated drifts in Figure 5.15, the drifts were mostly westward and they do not restrict the effect of the meridional winds, agreeing with the trans-hemisphere transport theory. At 2 UT, Figure 5.16d, the winds in the east longitudes of the map become poleward. The northern crest over continent US reaches higher latitudes, that agrees with the northern winds over this section and the trans-hemisphere transport. Then, at 3 and 4 UT, Figure 5.16e and Figure 5.16f, the winds in the northern hemisphere start to become westward over the NILE longitudes. The NILE is visible over US and it has moved west from the previous time. This may be due to the west neutral winds, as the ion drifts at this time, Figure 5.15e, are mostly equatorward.

**5.3.2.3 Neutral winds validation.** Neutral wind estimations at the FPI locations are also analyzed and compared to the measurements. Figure 5.17 shows the results for the different LOS at the Urbana location (40 deg N and 88 deg W), where the

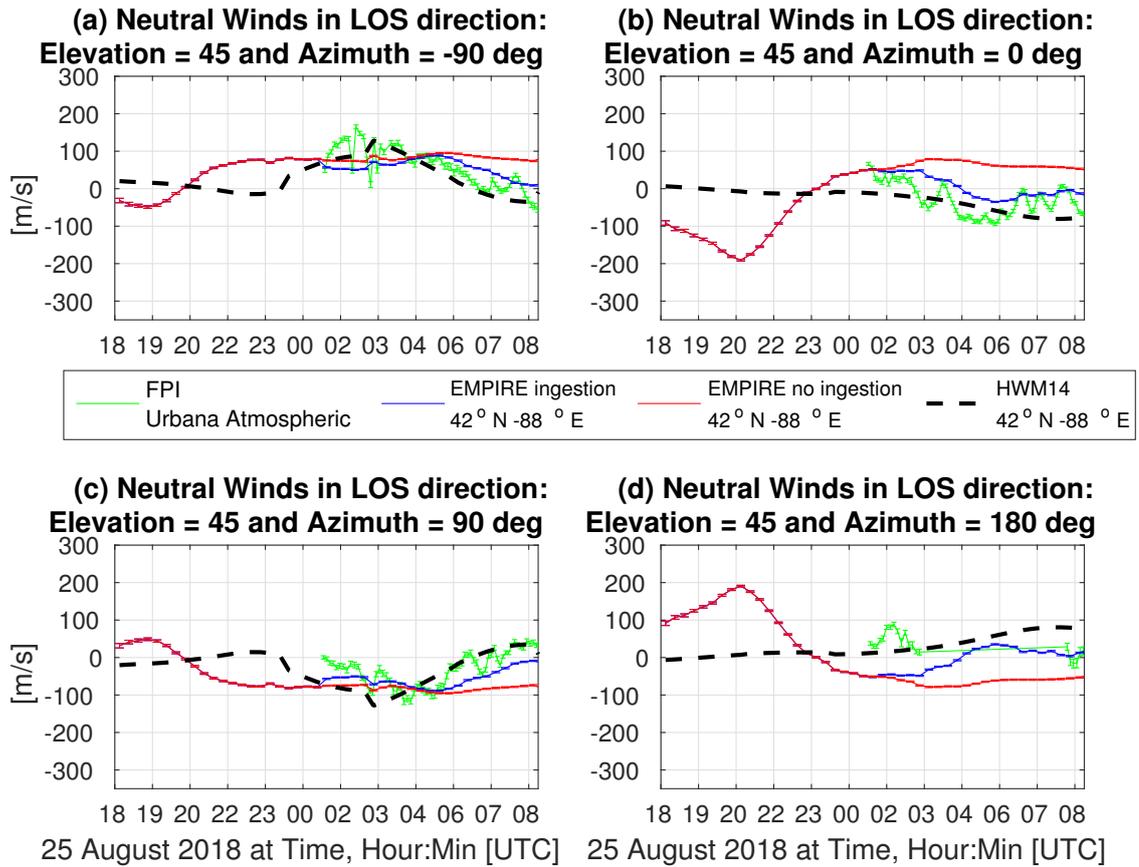


Figure 5.17. Neutral winds plotted over time at Urbana location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value.

neutral winds measurements were ingested in addition to the global electron density. EMPIRE results for “Case ingestion” are shown in blue, HWM14 model values are indicated with a black dashed line and the FPI measurements are observed in green over time. Positive values mean that the winds are going away from the FPI and negative values refer to winds going in the LOS direction to the FPI. The results for the “Case no ingestion” are shown in red.

In Figure 5.17a, the LOS is looking to the west direction, as the azimuth angle is  $-90^\circ$ . We can see that the estimation using the ingestion of winds at this location works properly, as the blue estimation is much closer to the measurements than the

red estimation (no ingestion). It is also visible that until 1 UT, when there are these additional measurements of winds available, the red and blue estimations coincide.

Figure 5.17b represents the north LOS, with an azimuth angle of  $0^\circ$ . This direction is ingested for the blue results. We can see that the blue estimation follows the measurements in green as expected. The red estimation is not close to the measurements for this LOS, and it gives a worse estimation than the model HWM14 in black. However, not correcting this direction with the VSH derivation means that the zonal direction is not corrected either. For the zonal LOS, Figure 5.17a and 5.17c, the red line is closer to the measurements than the model at some times.

Figure 5.17c indicates an eastward LOS with an azimuth angle of  $90^\circ$ . This direction is also ingested in EMPIRE for the blue estimation. It is observed, just like in the west LOS, that the EMPIRE estimation with ingestion of winds, follows the measurements in green. Finally, the Figure 5.17d indicates a southward LOS as the azimuth angle is  $180^\circ$ . For this case, there is a disagreement between the measurements in north and green LOS. As described in Section 2.2.2, south LOS are usually less reliable than the northern ones as the brightness of the samples is smaller.

Next, the results are compared at sites where the winds were not used, so we can use the measurements as validation of the EMPIRE estimation. The Millstone Hill FPI ( $44^\circ\text{N } 72^\circ\text{W}$ ) is the closest to the Urbana FPI, where the winds were ingested. The results at this location over time are shown in Figure 5.18 for the different LOS where measurements were available. Measurements are indicated in green, model HWM14 values in black, EMPIRE results ingesting Urbana winds in blue, and EMPIRE results not ingesting additional winds in red.

Again, Figure 5.18a, b, c and d represents a west, north, east and southward LOS direction respectively, indicated by the azimuth angle. In the west and east

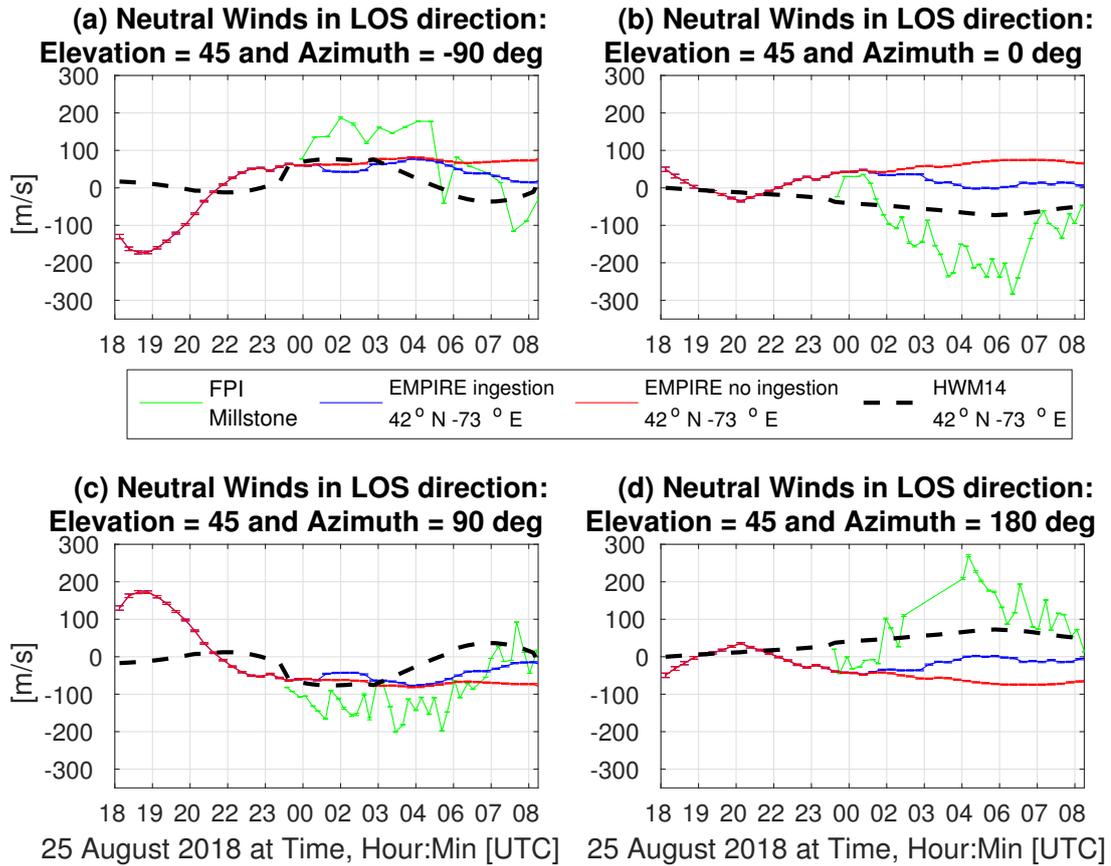


Figure 5.18. Neutral winds plotted over time at Millstone Hill location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value.

LOS, we can see that from 1 to 6 UT, the blue estimation is actually worse than the estimation with no ingestion in red, as the no ingestion case estimates a more eastward drift than the estimation with ingestion. However, in the meridional estimation, in the north and south LOS, the estimation in blue is better than the estimation with no ingestion. Nevertheless, for this case the climate model is better than both estimations. A better study to understand why this is happening at this location needs to be done.

Finally, results at the Arecibo FPI site ( $18^{\circ}\text{N}$  and  $66^{\circ}\text{E}$ ) for the different LOS where measurements were available are shown in Figure 5.19. Measurements are represented with a green line, model HWM14 values with a black dashed line, EMPIRE results ingesting Urbana winds with a blue curve, and EMPIRE results not ingesting additional winds with red.

Figure 5.19a, b, c and d represents a west, north, east and southward LOS direction respectively, indicated by the azimuth angle. First, we can see that the measurements at this location reaches higher values in all of the LOS than in the previous studied FPIs. The meridional ingestion estimation, looking at the north LOS Figure 5.19b, indicates a more southward velocity than the model and the estimation with no ingestion, agreeing better with the measurements from 3 UT. From 22 to 3 UT approximately, both estimations are able to capture a more northward wind than the model represented in black dashed line, just as the measurements indicate. The zonal estimation, shown in Figures 5.19a and 5.19c, is also more westward at the end of the studied period than the model and no ingestion estimation case, like the measurements indicate.

#### 5.4 Conclusions

Neutral winds and ion drifts have been estimated with EMPIRE for two geo-

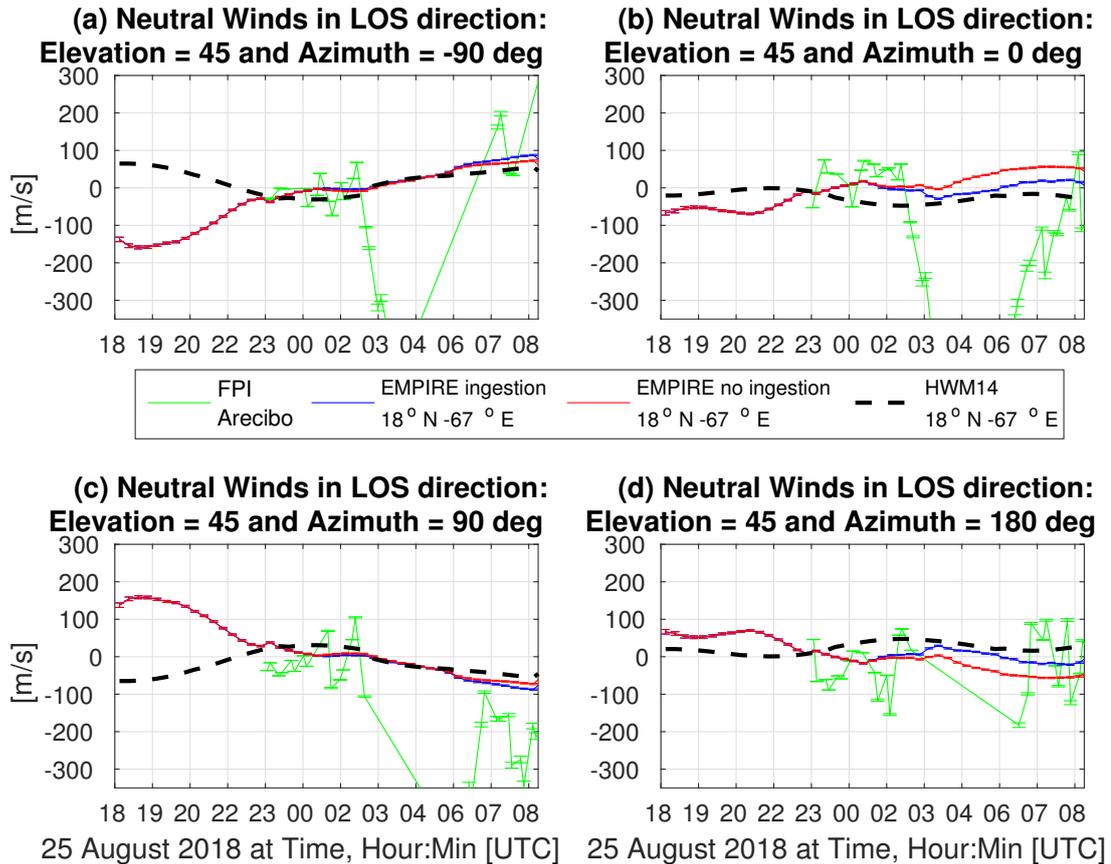


Figure 5.19. Neutral winds plotted over time at Arecibo location to the (a) west, (b) north, (c) east and (d) south. Green circles indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no addition ingestion. Black dashed line indicates the HWM14 value.

magnetic storms, November 2003 and August 2018, where a NILE enhancement was observed. The use of global vector spherical harmonics basis allows to study globally this event and the assessment of EMPIRE error helped to determine which drivers are more trustworthy and where. The ion velocity was studied for both storms, separating it in ion drifts and neutral winds. For the storm on August 2018, Urbana FPI measurements of neutral winds were ingested in addition to normal electron densities inputs.

For the November 2003 storm, the neutral wind meridional behavior and the strong fountain effect observed in the estimated ion drifts are consistent with a stronger EIA crest in the Northern Hemisphere. The EIA asymmetry may be the reason why only the northern crest persists into the night time, so maybe is the plasma source of NILE. The ion drift estimation agrees with a poleward and upward motion of an enhanced EIA driven by a super fountain. The vertical drifts at Millstone Hill ISR were compared to the measurements, and both showed a similar trend.

The global results for the August 2018 storm also indicates that the asymmetry of the enhanced EIA, from where the NILE appears to originate from, is also due to the neutral winds meridional behavior and a weak fountain effect. The fountain effect is not as strong as in the previous case, although it also appears that it is the driver that causes the enhanced TEC at mid latitudes during the first hours of the storm. However, it may also be due to only a PRE effect on sunset stronger because it happened during the magnetic storm.

An analysis to determine if the self-consistent basis function used to expand the neutral winds helps with the estimation in locations near the Urbana FPI site was done. That case is also compared to a case where no additional measurements of winds are ingested to estimate the drivers. Validation of the results of the 2 cases was done by comparing the neutral winds to FPI measurements at 2 additional locations

and to results. We conclude that the ingestion helps the estimation in the location where winds are ingested. Also, in the location farthest from the ingestion point, the estimation ingesting winds agrees better than the model with the measurements almost all the studied time. However, at mid-latitudes, the location closest to the point where measurements are ingested, the estimation differs more from the measurements. For the future, a weighting function depending on the distance to the point at which we are ingesting winds can be determined, as it is possible the effects of ingesting at a location constrain too much a nearby location, increasing the error in the estimation. Also, a Kalman filter was used as EMPIRE is designed to one day work in real time. However, analyses of previous geomagnetic storms could benefit by using other solvers in which the information used is maximized. Applying information or smoothing filters to EMPIRE may optimize the use of the observations ingested, as they add additional backward time updates to the estimate.

## CHAPTER 6

## ADDITIONAL INGESTION OF ICON NEUTRAL WINDS MEASUREMENTS

In this contribution the goal is to ingest global neutral winds measurements from ICON to estimate the neutral winds with EMPIRE. In the previous contribution we have seen that the ingestion of winds help the estimation in the location where we are ingesting, but at locations near the ingested data site, the estimation may agree less with the independent measurements. Global ICON measurements might help with this problem, as it measures neutral winds globally at different sites in a period of time. The global expansion using the vector spherical harmonics derivation from Chapter 3 allows this study. We modify EMPIRE to ingest ICON neutral winds as measurements in the observation system matrix in Section 6.1.

The MIGHTI (Michelson Interferometer for Global High-resolution Thermospheric Imaging) instrument included in the ICON satellite takes measurements at low latitudes, so we expect to observe the ionosphere with EMPIRE better at these latitudes when ingesting the measures. To check if the EMPIRE estimations improve ingesting ICON measurements with respect to only ingesting electron densities, a period of time where ICON measurements are available will be analyzed for these 2 cases (ingesting and not ingesting ICON neutral winds measurements). The EMPIRE configuration and the day selected for this analysis are described in Section 6.2. The results of the neutral winds estimation for these two cases are shown in Section 6.3. Neutral winds will be plotted in a map over the zone where the ICON measurements are available for different times. They will also be plotted at specific locations over time, where FPI measurements are available, for validation of the results. Finally, Section 6.4 concludes and summarizes the results found in this contribution.

Table 6.1. Variables downloaded from ICON [1].

Variable name	Description
Zonal Wind	Zonal component of the horizontal neutral wind. Positive Eastward (m/s)
Meridional Wind	Meridional component of the horizontal neutral wind. Positive Northward (m/s)
Zonal Wind Error	Error in the zonal wind estimate (m/s)
Meridional Wind Error	Error in the meridional wind estimate (m/s)
Wind Quality	Quantification of the quality, from 0 (Bad) to 1 (Good)
Epoch	Sample time (number of ms since Jan 1, 1970) (ms)
Altitude	WGS84 altitude of each wind sample (km)
Longitude	WGS84 longitude of each wind sample (deg)
Latitude	WGS84 latitude of each wind sample (deg)

## 6.1 Ingestion of ICON neutral winds

The ingestion of measurements of neutral winds from FPI for VSH was described in Section 5.1. ICON outputs the meridional and zonal geographic winds for different locations and times with associated uncertainties. EMPIRE is modified to additionally ingest ICON measurements of neutral winds.

**6.1.1 ICON data.** The data is obtained from [icon.ssl.berkeley.edu/Data](http://icon.ssl.berkeley.edu/Data) [1]. The ICON Data Product 2.2 is used for this research. Specifically, the variables that use are described in Table 6.1, where the WGS84 (World Geodetic System-1984) is a standard datum for use in satellite navigation, that is given in geographic coordinates.

A difference from the ground-based FPI measurements is that the FPI winds were given in the LOS direction of the instrument, while for the ICON measurements the meridional and zonal wind products are provided. Before ingesting the data in the algorithm, a few steps to process these data are taken. First, for each time interval, defined by the Epoch variable in Table 6.1, the points outside the EMPIRE domain

are removed using the altitude, longitude and latitude values defined in Table 6.1. Then, from this reduced set of points, measurements that are “not a number” (NaN) are discarded. Finally, measurements that have a quality flag (defined in Table 6.1) equal to 0 are also removed.

**6.1.2 Modification in EMPIRE algorithm.** To include these measurements in the algorithm, similarly to the FPI winds ingestion, additional  $2i$  equations are imposed to the EMPIRE system defined in Equation 3.20, where  $i$  is the number of different locations where ICON measurements are available at a time step. Note that both sources of neutral winds measurements, FPI and ICON, could be ingested at the same time by increasing the EMPIRE system including the corresponding equations. The additional equations are:

$$y_{N,ICON} = u_{N,ICON,i} \quad (6.1)$$

$$y_{E,ICON} = u_{E,ICON,i} \quad (6.2)$$

Where  $y_{N,ICON}$  and  $y_{E,ICON}$  represents the meridional and zonal neutral wind measurement, respectively, from the MIGHTI instrument of ICON at each  $i$ th location. The neutral winds VSH expansion, defined in Equation 3.10, is used to expand the ICON winds  $u_{N,ICON,i}$  and  $u_{E,ICON,i}$ :

$$u_{N,ICON,i} = \underbrace{\mathbf{f}_{N,i}}_{\mathbf{h}_{N,ICON,i}} \mathbf{x}_{\mathbf{u}} \quad (6.3)$$

$$u_{E,ICON,i} = \underbrace{\mathbf{f}_{E,i}}_{\mathbf{h}_{E,ICON,i}} \mathbf{x}_{\mathbf{u}} \quad (6.4)$$

Matrices  $\mathbf{f}_{N,i}$  and  $\mathbf{f}_{E,i}$  are defined in C and mapping matrices  $\mathbf{h}_{N,ICON,i}$  and  $\mathbf{h}_{E,ICON,i}$  transform the coefficients into neutral wind components.

To add the new ingested measurements to EMPIRE, we stack Equations 6.1

and 6.2 at each  $i$  location and change the notation to:

$$\mathbf{z}_{\text{ICON},i} = \underbrace{\begin{bmatrix} y_{N,\text{ICON}} \\ y_{E,\text{ICON}} \end{bmatrix}}_{\mathbf{y}_{\text{ICON}}} - \underbrace{\begin{bmatrix} a_{0N,\text{ICON}} \\ a_{0E,\text{ICON}} \end{bmatrix}}_{\mathbf{a}_{0,\text{ICON}}=0} = \begin{bmatrix} u_{N,\text{ICON},i} \\ u_{E,\text{ICON},i} \end{bmatrix} \quad (6.5)$$

Substituting the expansion of  $u_{N,\text{ICON},i}$  and  $u_{E,\text{ICON},i}$  using Equations 6.3 and 6.4,

$$\mathbf{z}_{\text{ICON},i} = \underbrace{\begin{bmatrix} \mathbf{h}_{N,\text{ICON},i} \\ \mathbf{h}_{E,\text{ICON},i} \end{bmatrix}}_{\mathbf{h}_{\text{ICON},i}} \mathbf{x}_{\mathbf{u}}. \quad (6.6)$$

Stacking all the  $2i$  equations defined in Equation 6.6, the following equation is obtained.

$$\mathbf{z}_{\text{ICON}} = \mathbf{H}_{\text{ICON}} \mathbf{x}_{\mathbf{u}}, \quad (6.7)$$

where the mapping matrix  $\mathbf{H}_{\text{ICON}}$  is obtained by stacking the  $i$  2-row matrices  $\mathbf{h}_{\text{ICON},i}$ . The EMPIRE system, defined in Equation 3.20, is now augmented:

$$\underbrace{\begin{bmatrix} \mathbf{z}_{\text{N}} \\ \mathbf{z}_{\text{ICON}} \end{bmatrix}}_{\mathbf{z}} = \underbrace{\begin{bmatrix} \mathbf{y}_{\text{N}} \\ \mathbf{y}_{\text{ICON}} \end{bmatrix}}_{\mathbf{y}} - \underbrace{\begin{bmatrix} \mathbf{a}_{0\text{N}} \\ \mathbf{a}_{0,\text{ICON}} \end{bmatrix}}_{\mathbf{a}_0} = \underbrace{\begin{bmatrix} \mathbf{H}_{\text{exb}} & \mathbf{H}_{\mathbf{u}} \\ \mathbf{0} & \mathbf{H}_{\text{ICON}} \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} \mathbf{x}_{\text{exb}} \\ \mathbf{x}_{\mathbf{u}} \end{bmatrix} \quad (6.8)$$

The state  $\mathbf{x}$  is solved with a Kalman filter, as described in Equations 3.22 to 3.33. Since there are new measurements ingested the covariance matrix of the measurements  $\mathbf{R}$  is also augmented, which appears in the measurement update of the filter. A diagonal matrix with the errors of the ingested winds  $\mathbf{R}_{\text{FPI}}$ , with size  $i$  by  $i$  is block-added to the measurement covariance matrix of the ion continuity equation observations  $\mathbf{R}_{\text{N}}$ , which size is  $j$  by  $j$ , where  $j$  is the number of grid points in the

EMPIRE domain:

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}_N & \mathbf{0}_{i \times i} \\ \mathbf{0}_{j \times j} & \mathbf{R}_{\text{FPI}} \end{bmatrix} \quad (6.9)$$

EMPIRE data processing flow chart for this case is shown in Figure 6.1.

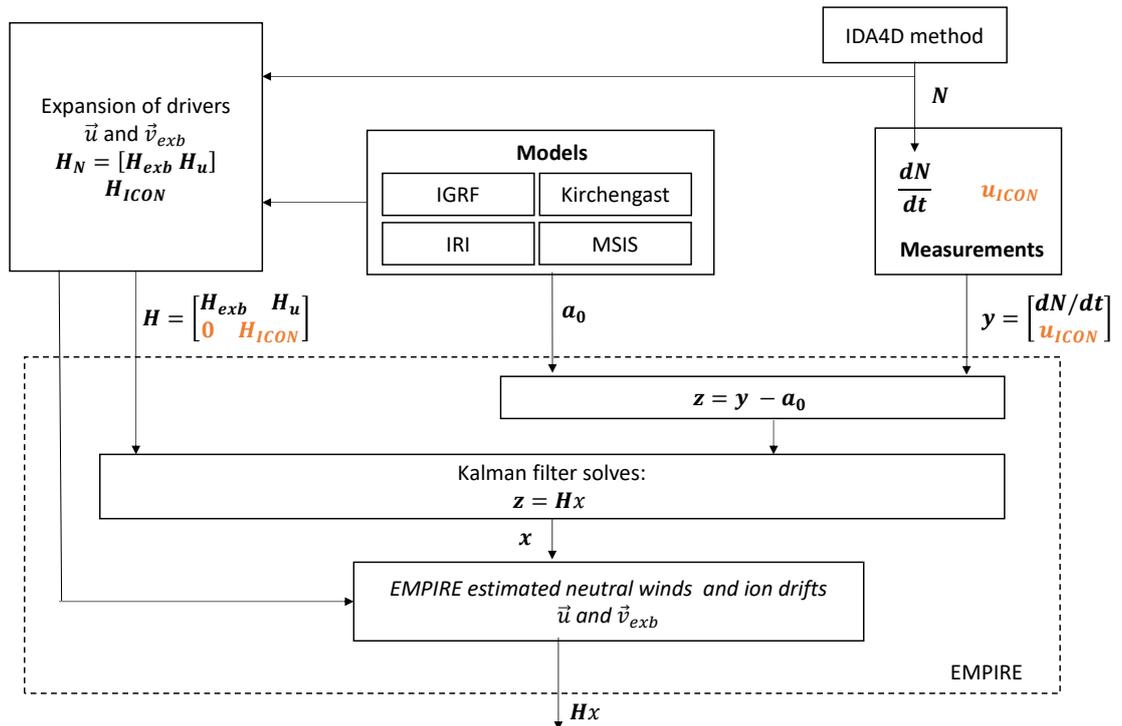


Figure 6.1. Simplified EMPIRE model and data processing flow chart.

## 6.2 Method

In this section we describe the case studied, where ICON data is available to ingest and FPI measurements are available to validate the results. We also shown the data ingested by EMPIRE and the EMPIRE configuration for two different cases.

**6.2.1 January 2020.** The case selected to study the ICON assimilation in EM-

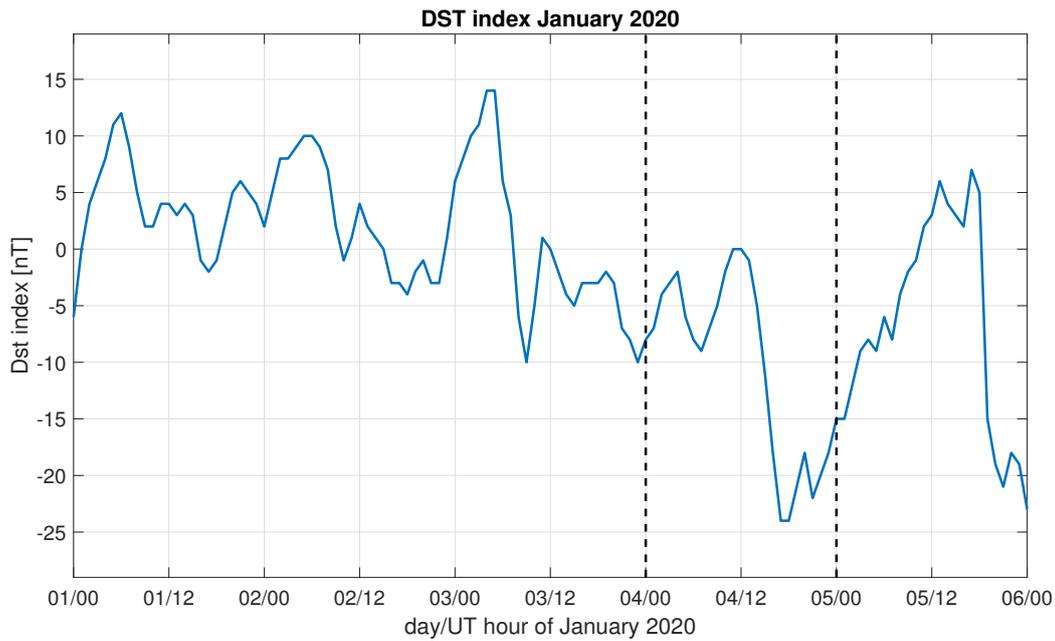


Figure 6.2. Dst index for 1-5 January 2020 in nT. The studied time limits are indicated with black dashed line.

PIRE is the day January 4, 2020. The Dst index, indicator of the geomagnetic activity is shown in Figure 6.2 over January 1 to 5. The Dst index doesn't reach very extreme values, but the minima over this period of time are found on the 4th. ICON measurements are only available from the end of solar cycle 24 to the present, beginning of solar cycle 25. As explained in Section 2.1, these times represent the solar cycle minimum and in addition solar cycle 24 was less active. We chose this date to do the ICON study because IDA4D results and ICON MIGHTI measurements between end of December 2019 to the first few days of January 2020 were both made available by ICON science team member Dr. Gary Bust, were as strong southward neutral winds over the US were detected.

The coverage of the MIGHTI instrument is illustrated in Figure 6.3 with red dots. For the day we study this implementation, January 4 2020, the locations at which measurements were made are shown in Figures 6.3a - 6.3f with red dots. Each of the subfigures represents the coverage of the instrument for 20 minutes, as it is

the time step used in EMPIRE, for a range of altitudes between 200 km and 400 km, where there are measurements of the neutral winds available. We plot 2 hours of coverage because the period of the orbit is of 97 minutes. The measurements at these points will be ingested by EMPIRE and they will also be compared to the EMPIRE estimations.

During this event, there were 3 ground-based FPIs available that measured the neutral winds: Arecibo, Millstone Hill and Urbana Atmospheric Observatory. Arecibo site ( $18^{\circ}\text{N}$ ,  $66^{\circ}\text{W}$ ), Millstone ( $44^{\circ}\text{N}$ ,  $72^{\circ}\text{W}$ ) and Urbana site ( $40^{\circ}\text{N}$ ,  $88^{\circ}\text{W}$ ) are marked in Figure 6.3 with a magenta star, triangle and square respectively. The measurements will be used as validation data and they will be compared to EMPIRE estimations of the neutral winds. As in previous storms analyzed, the FPI measurements are given in different LOS directions of the instrument. The estimated geographical neutral winds will be projected to this LOS direction for the comparison. For all of the FPI sites there are measurements available in the 4 geographic directions (north, south, east and west). The pierce point of the LOS is not the same as the location of the sites. The separation between the site location and the observation pierce point at 250 km is less than 2.5 degrees, which is smaller than the EMPIRE grid resolution of 6 degrees. We will use the site location as the point at which we have the neutral wind data, as the grid does not distinguish between this point and the actual region where the measurements are given.

**6.2.2 EMPIRE configuration.** The ingested electron density measurements  $N$  will be obtained from the IDA4D algorithm [102]. The ion production  $a_{0,prod}$  and  $a_{0,loss}$  are calculated using the [86] model. The gravity  $a_{0,g}$  and diffusion  $a_{0,dfsn}$  effect on the parallel direction of the ion velocity are also obtained by applying the [86] model. To characterize the neutral properties, needed to calculate these transport terms, the NRL-MSISE00 model is utilized [85]. The electron and ion temperatures

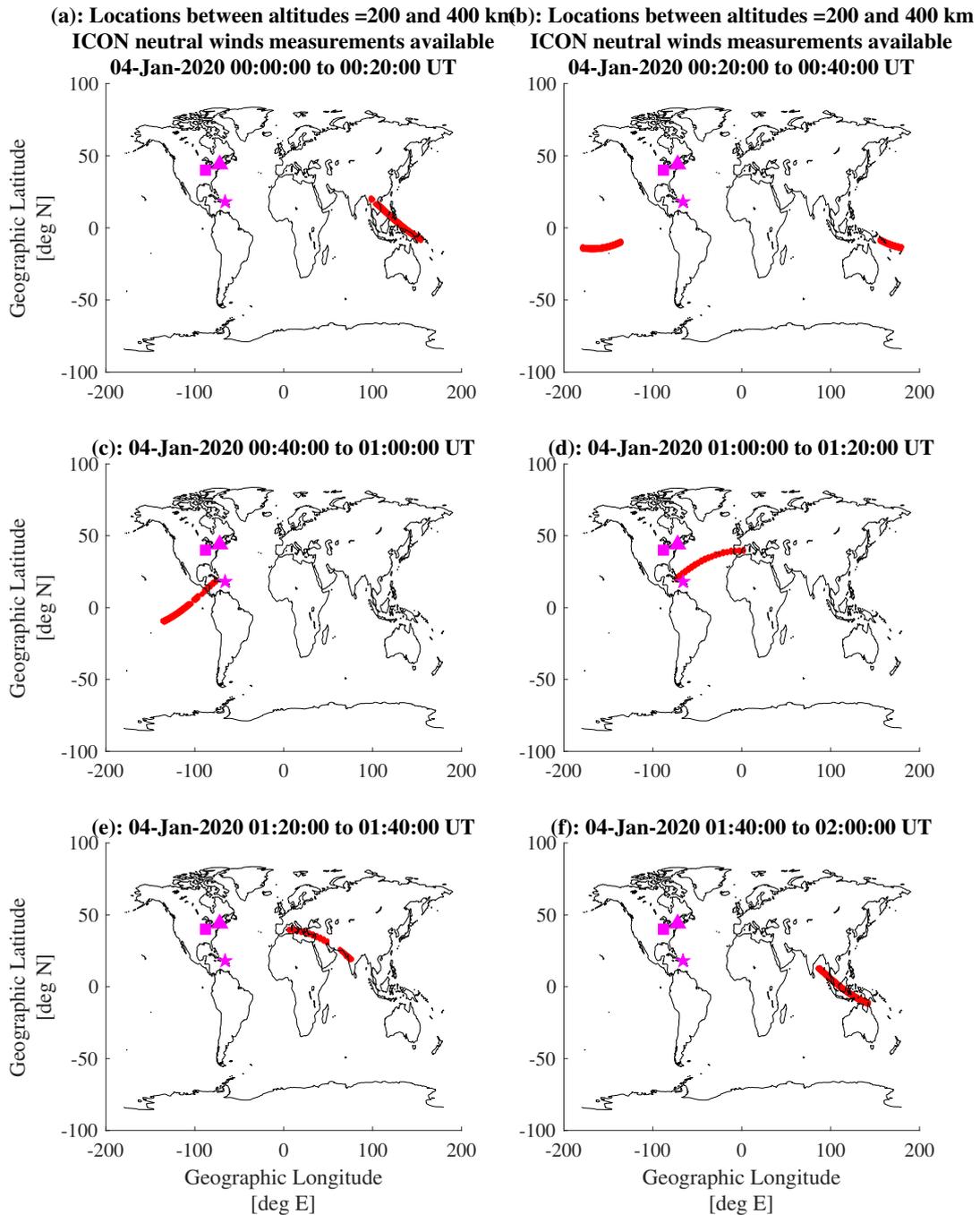


Figure 6.3. Coverage of ICON instrument for altitudes 200-400 km during January 4 2020 (a) 0UT-0:20UT, (b) 0:20UT-0:40UT, (c) 0:40UT-1UT, (d) 1UT-1:20UT, (e) 1:20UT-1:40UT and (f) 1:40UT-2UT.

are provided by the IRI model [84] and the magnetic field  $\vec{B}_0$  is provided by the IGRF model [59].

A global grid is imposed in EMPIRE with magnetic colatitude  $\theta$  between 3 and 177 degrees with a resolution of 6 degrees, magnetic longitude  $\phi$  between -180 and 180 degrees with the same resolution of 6 degrees and altitude  $h$  between 200 km and 500 km with a step of 50 km. The grid has 12810 points at each time step. The analysis period is of 24 hours, from 0 UT on January 4 with 20-minute increments.

**6.2.2.1 Cases.** To analyze the results of assimilating ICON measurements of neutral winds, two different cases are run in EMPIRE. In the first one, we only ingest the electron density  $N$ , from now on “no ingestion Case”. In the second case, from now on “ICON ingestion Case,” besides the electron density  $N$  EMPIRE will also ingest additional measurements of neutral winds obtained from ICON. The augmented system for this case is described in Section 6.1.2. In Figure 6.4 the EMPIRE configuration is shown for each of the 2 cases, where each block represents the vector or matrix indicated inside.

For both cases, to obtain the mapping matrix for each of the drivers,  $\mathbf{H}_u$  and  $\mathbf{H}_{\text{exb}}$ , we use the VSH decomposition described in Section 3 for the neutral winds and a spherical harmonics basis function for the potential field, as shown in [45]. The number of terms of the decomposition is limited by the maximum order selected  $l_{\text{max}}$ . For both basis functions, the order selected is  $l_{\text{max}} = 6$ . The state vector has a size of  $[90 \times 1]$ , where 48 coefficients come from the neutral winds expansion and the other 42 terms from the ion drifts derivation.

The background covariance and state for the Kalman filter, defined in Equations 3.32 and 3.33, are set up the same as in the previous Chapters. The model used to define the background state for the neutral winds  $\mathbf{x}_{0,u}$  is the HWM14 model

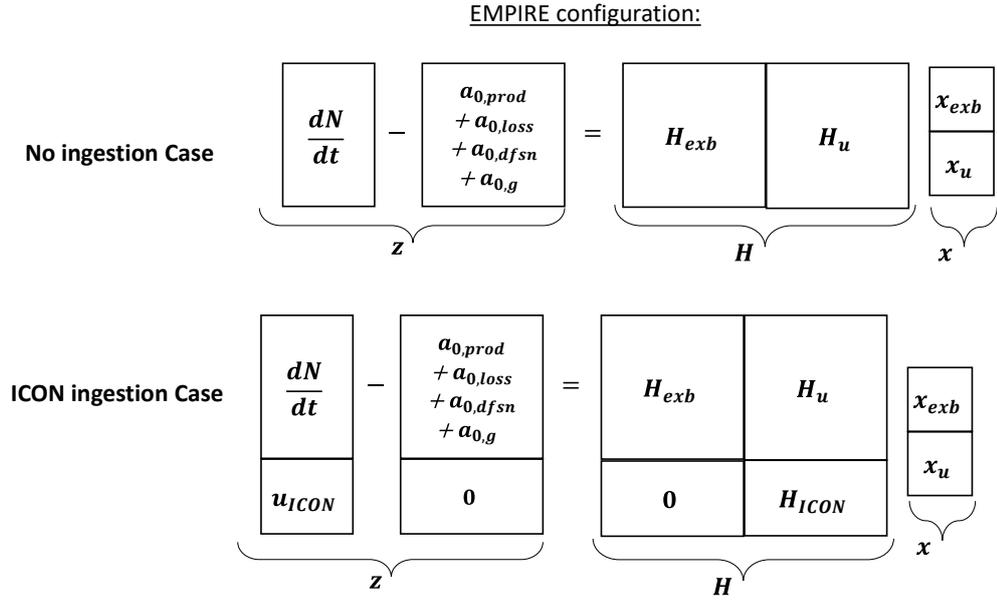


Figure 6.4. Diagram of the EMPIRE configuration for each of the studied cases.

[50]. For the ions drift background state  $\mathbf{x}_{0,exb}$  the Weimer model is used [58]. The background covariance  $\mathbf{P}_{0,t/t}$  is defined setting the following variances:  $\sigma_{u_{E,N,model}}^2 = (20m/s)^2$ ,  $\sigma_{u_{N,fitting}}^2 = (45m/s)^2$ ,  $\sigma_{u_{E,fitting}}^2 = (70m/s)^2$ ,  $\sigma_{exb,model}^2 = (20m/s)^2$  and  $\sigma_{exb,r,fitting}^2 = (20m/s)^2$ ,  $\sigma_{exb,\theta,fitting}^2 = (70m/s)^2$  and  $\sigma_{exb,\phi,fitting}^2 = (70m/s)^2$ . The scale factors  $c_u$  and  $c_{exb}$ , which take into account the propagated error of  $\mathbf{H}$  due to the error of the electron density ingested  $N$ , are set to 1, as in this case the electron density is considered more reliable as it is not a magnetic storm. The state is normalized with the regularization matrix  $\mathbf{L}$ , defined in Equation 3.34. For the neutral winds an order of magnitude of 100 m/s is used to normalize the state and for the potential field a order of magnitude of 1 kV. The time constant  $\tau$  needed for the Gauss-Markov transition in the matrix defined in Equation 2.28 is set to  $\tau = 7200s$ , similarly to previous chapters. The filter at any time step will “forget” the measurements ingested from 2 hours before.

For the “No ingestion” case, the measurement covariance matrix  $\mathbf{R} = \mathbf{R}_N + \mathbf{R}_{\mathbf{a}_0}$  used in the Kalman filter in 3.32 will have the same size as the EMPIRE grid:  $[12810 \times 12810]$ . It is derived from the IDA4D densities  $N$  errors and from the errors of  $\mathbf{a}_0$ . The array vector  $\mathbf{a}_0$  is calculated from different climate models, so there is not an error associated with the values used in EMPIRE. We make the hypothesis of an error of 1% of the actual value, as we are assuming that the main contributor to the electron density change is the ion drifts and neutral winds, to calculate covariance matrix  $\mathbf{R}_{\mathbf{a}_0}$ .

For the ICON “ingestion” case, the measurement covariance matrix  $\mathbf{R}$  will take into account the new ICON measurements ingested. It is defined as a diagonal matrix with size  $[(12810 + i) \times (12810 + i)]$ , where  $i$  is the number of ICON measurements of neutral winds available at each time step. The number of terms  $i$  can change at each time step. The first part of the matrix will be set up the same as the “No ingestion” case. The rest  $[i \times i]$  terms left, will be obtained from the error of ICON measurements, that are given with the measured data. An example of the ingested winds is illustrated in Figure 6.5. The ICON wind measurements, at locations marked in red in Figure 6.5a, are plotted in Figures 6.5b and 6.5c, in the meridional, in green, and zonal, in magenta, directions respectively. The errorbars represent the error of the measurements, that are contained in the covariance matrix  $\mathbf{R}$ .

Once the state is calculated for both cases using the described configuration, the estimated neutral winds are calculated at every grid point for the period of time studied. Results over the ICON coverage are plotted and compared to the measurements for both cases for different times. Estimated neutral winds over time for both cases will also be shown at the FPI sites whose neutral winds measurements are available for validation.

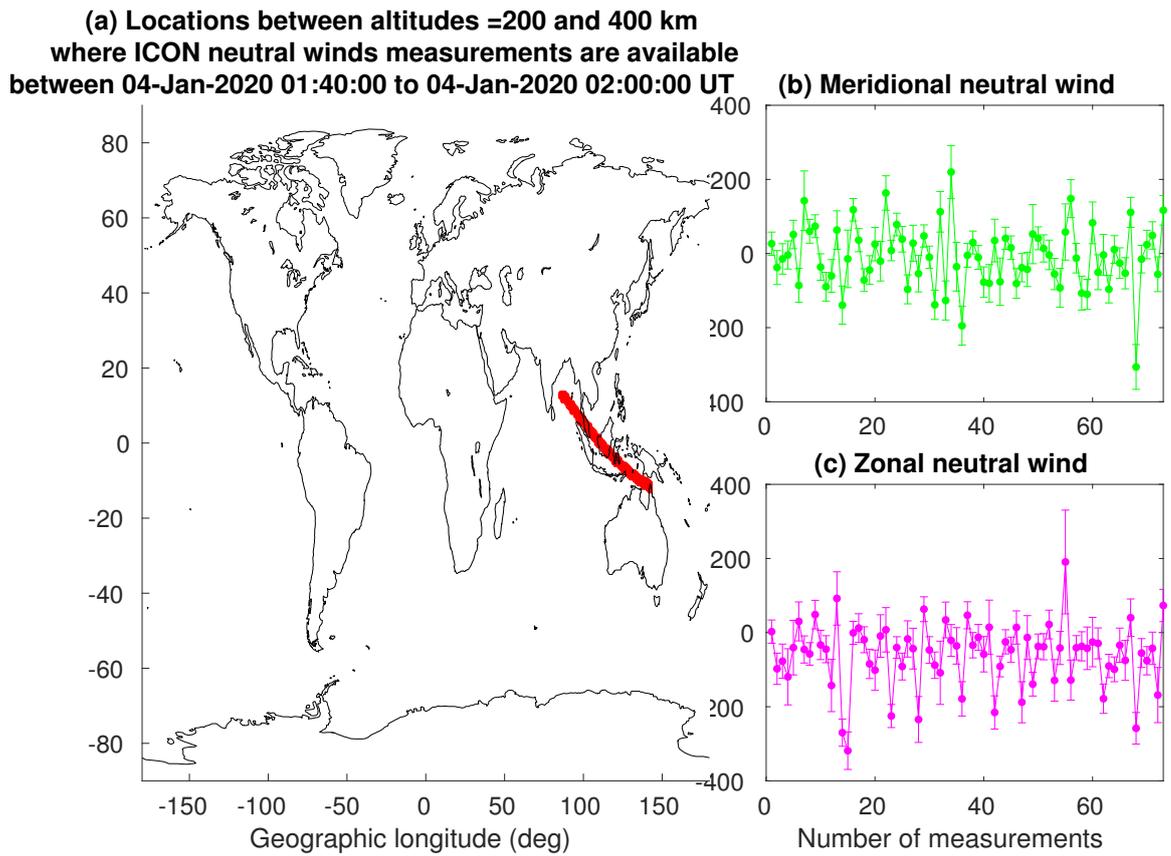


Figure 6.5. Neutral wind ICON measurements between 1:40 UT and 2:00 UT on January 4, 2020 at (a) points. Meridional component and error (b) in green and Zonal component and error (c) are observed over time.

### 6.3 Results

First, we analyze the global estimated winds for different times and we compare it to the ICON measurements. Figure 6.6 shows maps of the winds from ICON in green arrows, from EMPIRE estimation with ingestion in blue and the estimates with no ingestion in red. The different subplots represent different times of the studied period: 2:30 UT, 2:50 UT, 3:10 UT and 3:30 UT. We chose these time limits because they represent one period of the ICON orbit taking measurements.

We can see that at 2:30 UT, in Figure 6.6a, the ICON satellite is measuring over longitudes between  $160^\circ$  and  $80^\circ$  W, indicated with the green arrows. It is measuring eastward winds over central America of more than 100 m/s, a weaker poleward/eastward wind near the equator around  $120^\circ$  W and a equatorward/westward wind around  $150^\circ$  W. We can observe that the estimation without ingesting the measurements, in red, has a poleward component in the longitudes  $140 - 120^\circ$  W, but over central America the zonal component is the opposite direction of what ICON indicates. The winds estimation ingesting the additional ICON measurements, in blue, agrees with the measurements as expected. At the rest of the studied times, the ICON coverage moves eastward until the 97 minutes orbital period finishes. A similar behavior between the winds and the measurements as in the first time analyzed is observed.

The FPI sites whose data are used for validation are: Arecibo, Urbana and Millstone Hill (marked in Figure 6.6 with a magenta star, square and triangle). Results over time at the Arecibo FPI are shown in Figure 6.7 for different LOS of the instrument. Results of EMPIRE estimation with ingestion of ICON measurements are shown in blue, with the no ingestion case in red and the FPI measurements plotted in green.

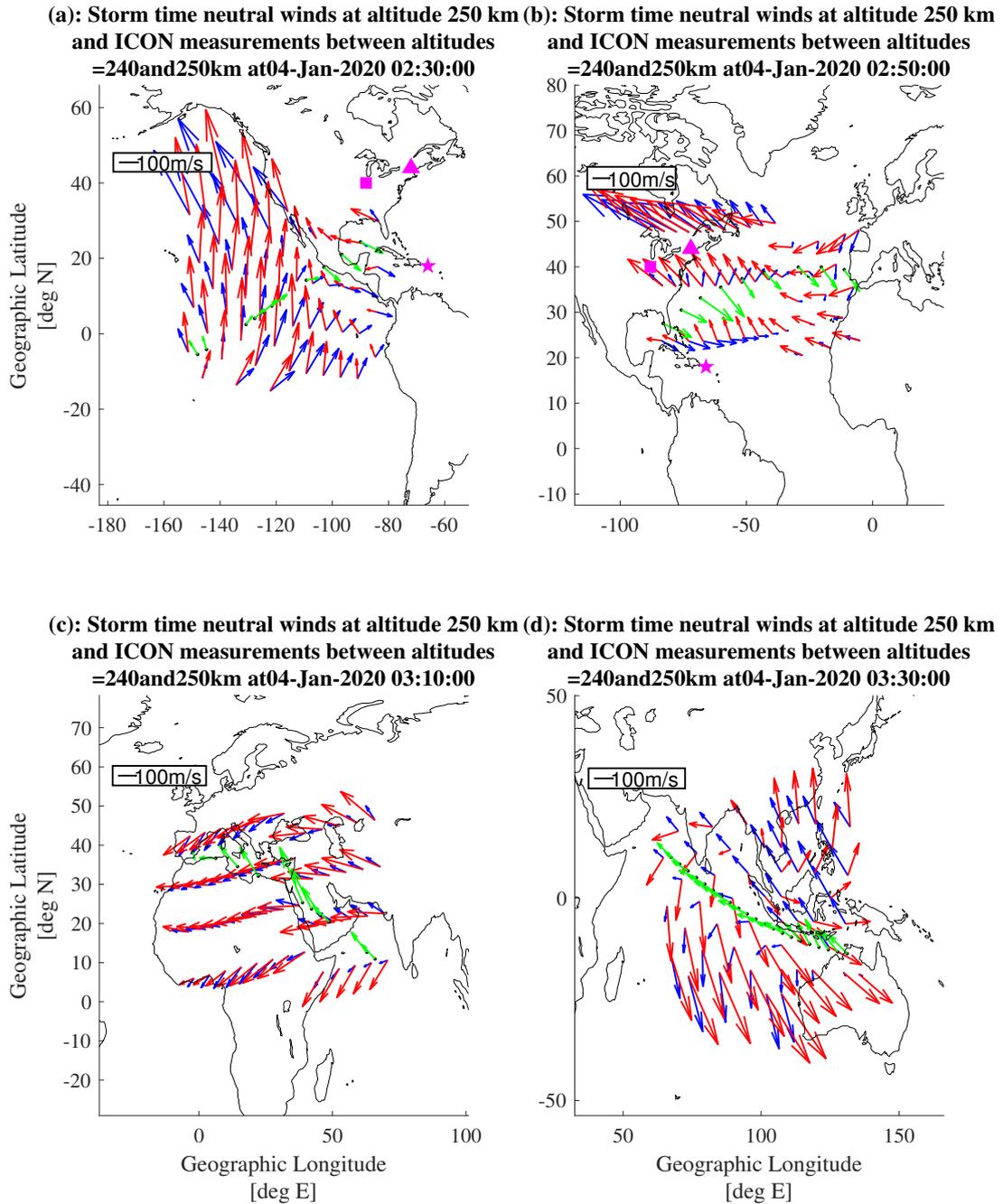


Figure 6.6. Neutral winds estimation with ingestion in blue, with no ingestion in red and ICON measurements in green over a world map for January 4 2020 at times (a) 2:30 UT (b) 2:50 UT (c) 3:10 UT and (d) 3:30 UT .

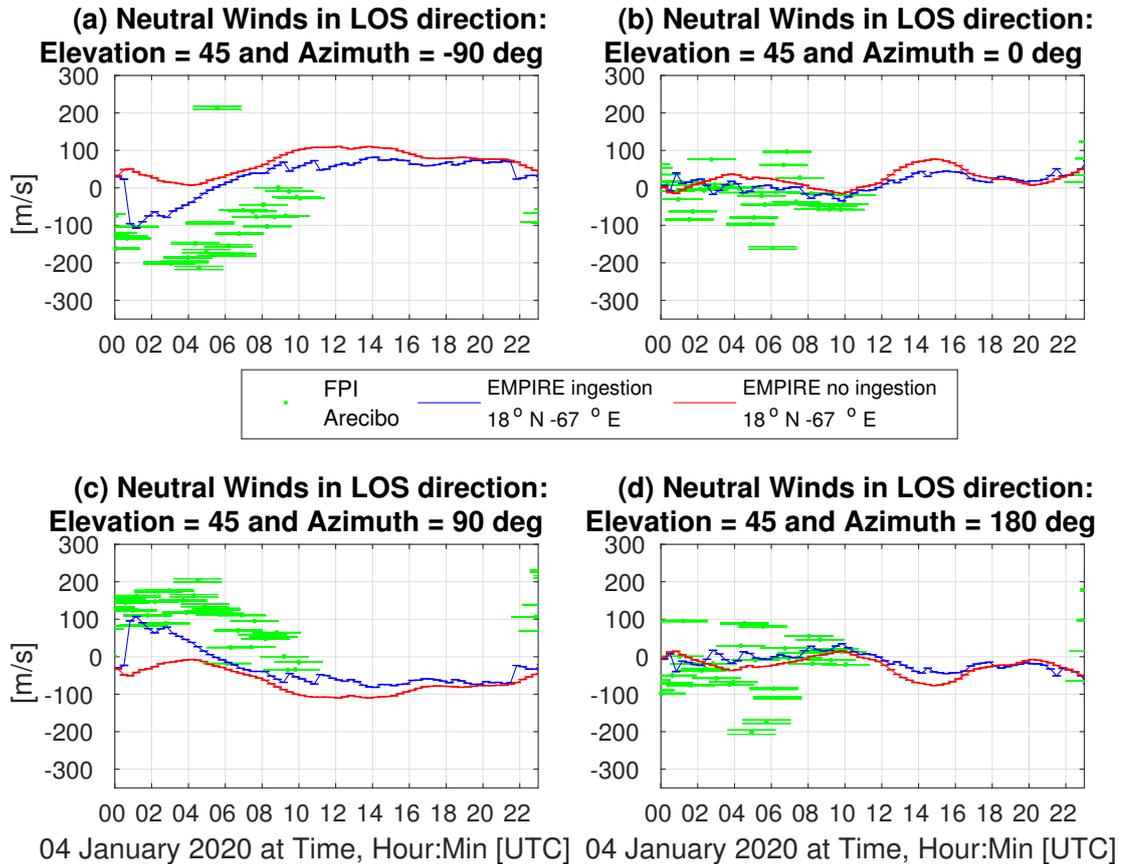


Figure 6.7. Neutral winds plotted over time at Arecibo location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion.

At Arecibo the 4 LOS of measurements available were west (Figure 6.7a), north (Figure 6.7b), east (Figure 6.7c) and south (Figure 6.7d), indicated by the azimuth angle. In the zonal direction, in both eastward and westward LOS from Figures 6.7c and 6.7a, the FPI measurements indicate a strong eastward wind of around 200 m/s between 0 and 4 UT and then they decrease to almost 0 m/s. The estimation with no additional ingestion in red is not able to estimate this motion, while the estimation with the ICON ingestion is able at the beginning to estimate these stronger winds. In the meridional direction, in the north LOS in Figure 6.7b, the measurements in green indicates a mostly southward motion from 0 to 8 UT (with some points indicating north). The EMPIRE estimation with ICON ingestion indicates also a southward motion, although not as strong as the measurements indicate. Around 8 UT, when the measurements are of a smaller order of magnitude, around 50 m/s, the EMPIRE estimation in blue agrees better with these values. In the south LOS, in Figure 6.7d, the measurements indicates a more northward motion during the period of time between 0 and 8 UT. However, southern LOS are usually less reliable than northern as less brightness arrives to the FPI, as explained in Section 2.2.2.

Looking at the coverage of ICON in Figure 6.6, there are ICON measurements available close to the FPI site between 2:30 and 2:50 UT. We can see in the 4 LOS that at that time and at some later specific times that the blue curve, which ingests the ICON measurements, corrects its trajectory. This may be the points where measurements are ingested and EMPIRE is trying to adapt the winds to what the measurements are indicating.

The error bars of both estimates are really small over time at all the LOS studied. This may be because the January 4, 2020 was a quiet day, as observed in Figure 6.2, so the ingested electron density  $N$  from IDA4D and consequently its error is smaller in absolute value. The error of the electron density  $N$  affects the observation

error of the system  $\mathbf{R}$ , so a possibility of the small error bar in the solutions might be the filter assuming the observations are too accurate and estimating the solution there. A detailed study of the observation error and the process error, which also plays a role in the final estimation error, should be done for quiet time analyses.

The results at another FPI site, in Urbana, are shown below in Figure 6.8. Again, the color green represents the FPI measurements, the color blue the EMPIRE estimation of the winds with additional ingestion of ICON winds and the color red the EMPIRE estimation of winds with no additional ingestion. They are plotted over time and each sub figure represents a different LOS in which the FPI took measurements, given by its azimuth angle: (a) west, (b) north, (c) east and (d) south.

The measurements in green in the east LOS direction, in Figure 6.8c, indicate westward winds of almost 100 m/s at 0 UT decreasing to around 0 m/s around 3 UT and then turning eastward with a magnitude of approximately 50 m/s around 6 UT, and finally decreasing again to almost 0 m/s at 8 UT to 12 UT. Both estimates, in blue with ingestion of ICON winds and in red with no ingestion, are able to estimate this westward surge at the beginning, although the order of magnitude of the estimate is stronger than measured by the FPI. Then the estimate in red is not able to capture the almost 0/eastward motion from 2 UT, while the blue estimate agrees better with the measurements. We can see around 16 UT, when close to this site there are no ICON measurements being ingested, that the blue curve tends to go back to the red one. In the westward direction a similar behavior is observed in Figure 6.8a.

For the meridional winds, the north LOS measurements in green from Figure 6.8b indicates a southward motion in the first 2 hours of the studied time and then a small northward wind the rest of the time, with a peak around 6 UT reaching around 50 m/s. The estimates, in blue with ingestion and in red with no ingestion, are not able to capture this initial southward motion, but they do stabilize around 0 m/s

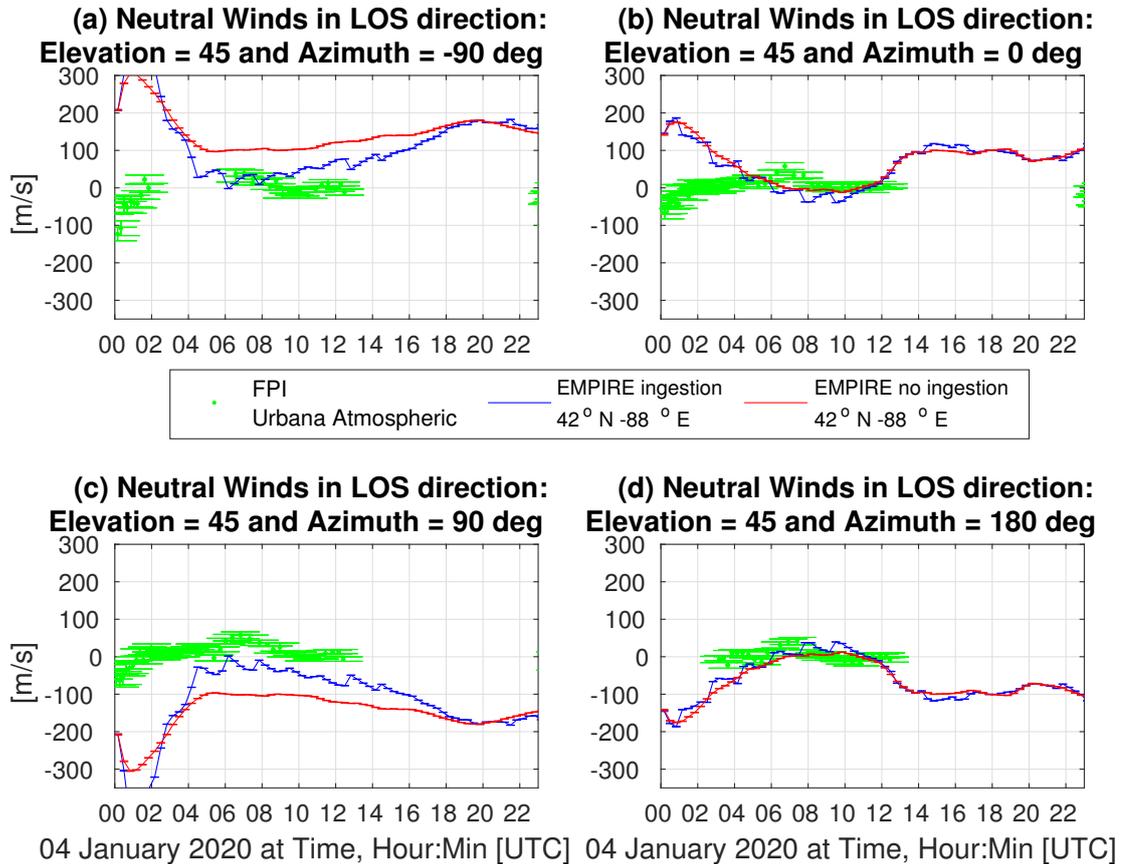


Figure 6.8. Neutral winds plotted over time at Urbana location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion.

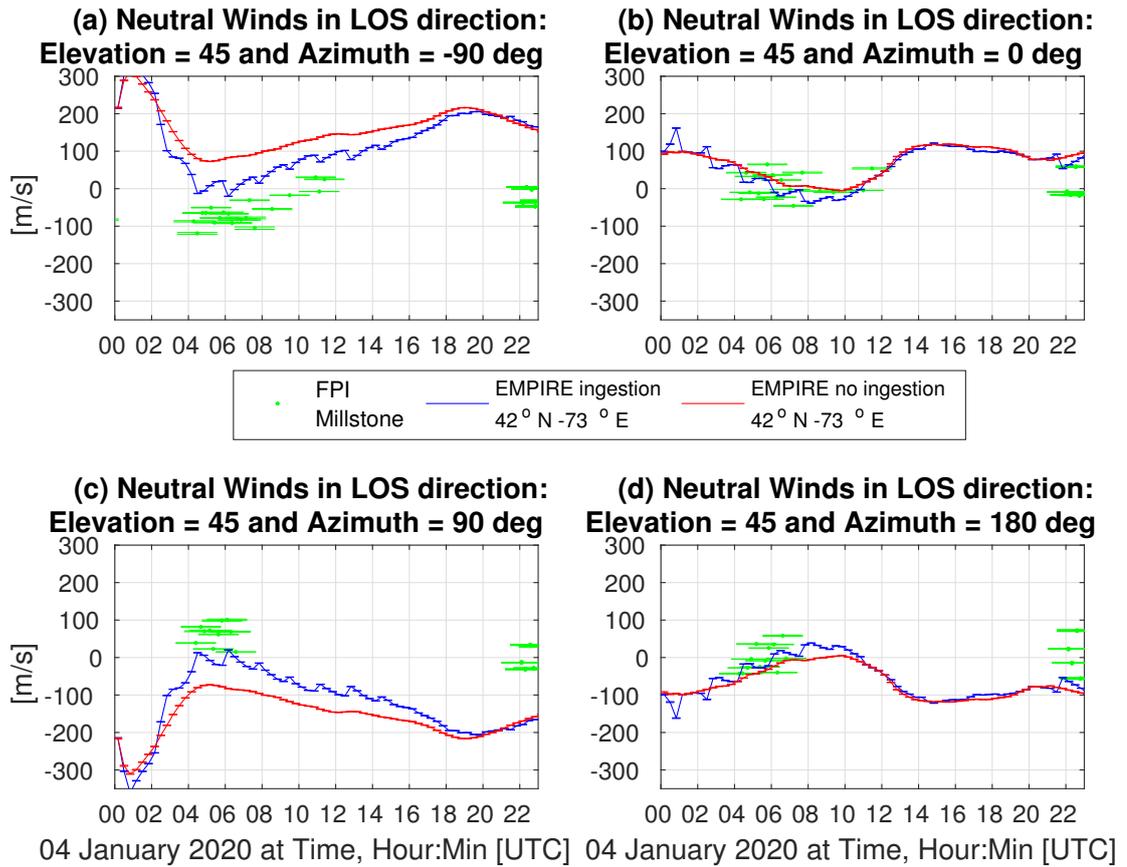


Figure 6.9. Neutral winds plotted over time at Millstone location to the (a) west, (b) north, (c) east and (d) south. Green dots indicate the measurement from the FPI, blue indicates the EMPIRE value at the closest grid point to the FPI location with ingestion of winds, and red indicates the EMPIRE estimation with no ICON ingestion.

around 4 UT, like the measurements do. In the south LOS in Figure 6.8d, there are not measurements of that initial south motion. But both estimations after 2 UT follow the trend of the measurements.

Finally, Figure 6.9 shows the results over time at the Millstone Hill FPI location. Green shows the measurements from the FPI instrument, blue the estimates ingesting ICON winds and red the estimates not ingesting additional winds measurements. Each subfigure represents a different LOS given by its azimuth angle: (a) west, (b) north, (c) east and (d) south.

Table 6.2. RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Arecibo.

RMS (m/s)	Arecibo S	Arecibo W	Arecibo N	Arecibo E
No ingestion	158.83	50.45	144.14	52.62
ICON ingestion	127.41	41.35	114.69	52.45

Table 6.3. RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Urbana.

RMS (m/s)	Urbana S	Urbana W	Urbana N	Urbana E
No ingestion	172.35	90.45	157.13	77.99
ICON ingestion	173.77	90.93	152.0	77.64

We can see in the zonal direction, looking at the eastward LOS from Figure 6.9c, that the estimate ingesting the ICON winds is able to estimate the eastward peak shown by the measurements after 4 UT, while the estimate not using ICON measurements is not able to capture it. In the meridional direction at this location, both EMPIRE wind estimates with and without ingestion similarly agree with the measurements between 4 and 8 UT.

To compare the improvement between the 2 cases, the RMS error between the estimations and the FPI measurements is calculated for each of the cases. Results for the Arecibo location are shown in Table 6.2.

There is an improvement of the RMS in the ICON case with respect the no ingestion case of  $\approx 20\%$  in the south, west and north LOS at the Arecibo location.

The results of the RMS between the estimations and the measurements at Urbana are shown in Table 6.3.

At this site, the RMS in the ICON ingestion case is very similar in all the di-

Table 6.4. RMS error in m/s between EMPIRE estimate for case ICON ingestion and no ingestion case and FPI measurements at Millstone.

RMS (m/s)	Millstone S	Millstone W	Millstone N	Millstone E
No ingestion	207.58	73.47	182.38	84.77
ICON ingestion	188.24	71.18	162.43	82.13

rections to the no ingestion case. The ICON measurements are taken at low-latitudes, closer to Arecibo site than the Urbana FPI (as we can observe in Figure 6.3). That may be the reason why at Arecibo there is a higher improvement of the RMS while at Urbana is much smaller.

Finally, the RMS results at the Millstone site are shown in Table 6.4.

It is observed that at Millstone FPI the RMS between the estimations and the measurements for the ICON ingestion for ICON ingestion case is smaller than for the no ingestion case. There is a RMS improvement of the estimations in the ingestion case of  $\approx 10\%$  in the meridional direction with respect the no ingestion case. In the zonal direction the RMS are similar for both cases.

## 6.4 Conclusions

Ingestion of additional neutral winds measurements from ICON satellite has been implemented to the EMPIRE algorithm. Neutral winds have been estimated with EMPIRE for January 4 2020. The vector spherical harmonics basis implemented in previous contributions allows the ingestion of ICON global measurements. Results were also compared to the results with no ingestion of winds, and validation of the results was done by comparing them to FPI measurements at three different locations.

As seen in the results, the use of ICON winds improves the estimation the points where the winds are ingested and in the surrounding area of these points with

respect to the “no ingestion” case. Comparing the results to the FPI measurements, available only during local nighttime, results ingesting the ICON winds are similar or better than the no ingestion case in the 3 sites we have compared the estimates to the measurements.

Ingesting ICON winds in EMPIRE makes the estimation of the winds closer to the measurements, both the ICON winds ingested and FPI winds not ingested but nearby. Additional studies to check how the ion drift estimation behaves when ingesting neutral winds from ICON could be done.

## CHAPTER 7

### CONCLUSIONS

This dissertation has described the data assimilation algorithm EMPIRE, which estimates ionospheric drivers neutral winds and ion drifts by ingesting primarily plasma density  $N$  and neutral winds measurements from Fabry-Perot interferometer ground instruments and/or ICON satellite winds, and by correcting the driver values of climate models. The main contributions include 1) a vector spherical harmonic derivation to expand the neutral winds for its global estimation, 2) a quantification of the representation error of the estimated drivers of EMPIRE, 3) a study of a nighttime ionospheric localized enhancement (NILE) event that is sometimes observed during some geomagnetic storms, and 4) ICON neutral winds ingestion.

#### 7.1 Summary and discussions

Chapter 3 derived and described a vector spherical harmonics expansion for the neutral wind driver. This allowed global estimation of the winds and consistent values of the vector in every direction. Analysis of a geomagnetic storm during October 25 2011 was done using the global expansion and results were compared to the previous method, in which power series basis was used to decompose each wind component. New results were better than or similar to the old implementation in comparison to FPI measurements.

Chapter 4 quantified the representation error of the drivers estimation of EMPIRE. Representation error describes the discrepancy between the ionosphere that EMPIRE describes and the real ionosphere. To study the representation error, the climate model SAMI3 is chosen as our “truth”. Because SAMI3 outputs the ionospheric drivers and all the data that EMPIRE ingests, we are able to compare the EMPIRE estimation with a “true” value (SAMI3 values). We also discuss the differ-

ent sources of errors of the algorithm and conclude that the representation error is mostly due to the fitting error of the drivers at low and mid altitudes, and that at high latitudes the EMPIRE estimations differ more from what is considered the true values of the drivers.

In Chapter 5, the global representation of the winds allow us to study the NILE (Nighttime Ionospheric Localized Enhancement) event, that has been observed during some geomagnetic storms over continental America. We analyze the event for 2 different cases to try to determine the effect of the meridional neutral wind and the fountain effect on its formation. FPI ingestion of winds is also adapted to the global vector spherical harmonic decomposition used in EMPIRE. For the NILE event on November 2003, we found that the ion drift behavior was consistent with a enhanced EIA (Equatorial Ionization Anomaly) and an enhanced super fountain effect. We also show that the neutral winds meridional behavior could be the cause of the asymmetry of the EIA (north hemisphere electron concentration is higher than in the south hemisphere), which may be why only the northern crest, from which the NILE appears to originate from, persists into the night. For the NILE event observed on August 2018, which was weaker, we also show that the asymmetry of the EIA may be due to the neutral winds meridional behavior combined with a weaker fountain effect found in the ion drift estimation. Only 2 different NILE events were studied. However, to better determine the mechanisms that cause this effect more storms need to be run. We also validate the ingestion of winds by comparing the results to wind measurements at 3 FPI sites. We conclude that the ingestion helps the estimation at the location where winds are ingested and that it affects the results with more strength at closer locations.

Finally, in Chapter 6, ingestion of ICON measurements is implemented in the algorithm. A study of how EMPIRE estimation of neutral winds changes with respect

to the non ingestion of winds is done. We analyze the neutral winds on a quiet day January 4, 2020 and we validate the results with 3 additional FPI measurements of the winds. We show that the ingestion of ICON measurements improves the estimation in the region where winds are being ingested and in the surrounding area. We also show that the estimation ingesting ICON agrees better with the FPI winds than the “no ingestion” case.

## 7.2 Future work

### 7.2.1 Neutral winds not limited to the parallel to the magnetic field line.

Neutral winds are being estimated by only considering the projection of the winds onto the field-parallel direction, as described in Chapter 2. We lose information of the driver in the perpendicular-to-the-magnetic-field direction. In principle, the contribution of these components is much smaller than the ion drifts contribution. However, for the future it would be helpful to estimate the whole neutral wind vector, as it will eliminate some of the observability problems.

**7.2.2 Refining of Kalman filter setup.** Additional Kalman filter setup analysis may help the estimation. For example, we may introduce numerically in the variance’s configuration of EMPIRE the error in the mapping matrix due to the electron density measurement, which was the main contribution of the representation error. The scale factor could also be studied to have a more methodological way to set it up. Finally, to determine the fitting error for any run we need to do a prior analysis according to the maximum order chosen of the decomposition of the drivers. However, it could be useful to implement this analysis in EMPIRE, so that this step could be automatically done when running the algorithm.

Also, implementing an information smoothing filter may improve the estimation. Smoother filters updates the estimation from the end of the time domain to

the beginning, so that the estimations in the first step are updated even using measurements from the last time step. Although these filters would not be useful for a real-time application of the algorithm, they may help in improving the estimation of the IT drivers in storms that have already happened and that research want to characterize, as the smoothing filters maximize the use of the ingested measurements.

Finally, an analysis to determine the effect of the start time and the time constant in the process noise of the time update on the result may help determine the optimal time to start a case with EMPIRE to obtain improved results. The time constant is used to “forget” the estimate from previous time and remove the accumulated error, so an optimal balance between these two factors can be found empirically to improve the estimates.

**7.2.3 Analysis of ion drift estimation when ingesting ICON measurements of neutral winds.** Chapter 6 studies the neutral winds behavior when ingesting ICON neutral winds. However, the ion drifts drivers are also being estimated and considered as a contribution in the electron density rate change in the system. An analysis on how this driver changes when ingesting or not ingesting winds will be helpful to determine and redefine how useful is the ICON ingestion to EMPIRE.

APPENDIX A  
MAPPING MATRICES  $F_{EXB}$  AND  $H_{EXB}$  FOR ELECTRIC POTENTIAL  
SCALAR SPHERICAL HARMONIC EXPANSION

In this appendix the terms that form the mapping matrices  $\mathbf{F}_{\text{exb}}$  and  $\mathbf{H}_{\text{exb}}$  are described. They are obtained by expanding the electric potential field with spherical harmonics basis function. The whole derivation is shown in [45], but a summary is shown below.

Mapping matrix  $\mathbf{f}_{\text{exb},j}$  is described in Equation 2.15 and transforms the state coefficients  $\tilde{\mathbf{x}}_{\text{exb}}$  into the potential field  $V$  space. It is obtained by using the expansion of potential field introduced in Equation 2.16. The expression is repeated below, but implying the sum with the  $lm$  indices for brevity:

$$\delta V_j = N_l^m P_l^m(\ell(r, \theta)) \Phi_l^m(\phi, \tilde{\mathbf{x}}_{\text{exb}}) \quad (\text{A.1})$$

$$\Phi_l^m(\phi, \tilde{\mathbf{x}}_{\text{exb}}) = \tilde{x}_{c,exb}^{lm} \cos \phi + \tilde{x}_{s,exb}^{lm} \sin \phi \quad (\text{A.2})$$

where  $N_l^m P_l^m$  is the fully normalized associated Legendre polynomial as described by [89] and the variable  $\ell(r, \theta)$  describes the normalized magnetic field L-shell between domain  $[-1, 1]$ . Consequently, the spatial potential field is a function of only two independent variables  $(\ell, \phi)$ . This can be done by assuming constant electric potential along the dipole field lines. The term  $\Phi_l^m$  contains the harmonic term and the unknown coefficients  $\tilde{\mathbf{x}}_{\text{exb}}$ . L-shell was defined in Equation 2.4 as:

$$L = \frac{r}{R_e \sin^2 \theta} \quad (\text{A.3})$$

Where  $R_e$  is Earth's mean radius [3]. The normalized L-shell,  $\ell$ , is described as:

$$\ell = -1 + \frac{2}{(L_{\max} - L_{\min})} (L - L_{\min}) \quad (\text{A.4})$$

where  $L_{\max}$  and  $L_{\min}$  are limits on the L-shells of the domain. They can be computed for grid domain limits  $(\theta, r)$  by plugging them into Equation A.3.

The state is defined as:

$$\tilde{\mathbf{x}}_{\text{exb}} = [\tilde{x}_{c,exb}^{11} \tilde{x}_{c,exb}^{21} \dots \tilde{x}_{c,exb}^{l_{\max} l_{\max}} \tilde{x}_{s,exb}^{11} \tilde{x}_{s,exb}^{21} \dots \tilde{x}_{s,exb}^{l_{\max} l_{\max}}]^T \quad (\text{A.5})$$

$$= [\tilde{\mathbf{x}}_{\text{c,exb}}^T \tilde{\mathbf{x}}_{\text{s,exb}}^T]^T. \quad (\text{A.6})$$

The coefficients for the lowest orders  $l = 0$  and  $m = 0$  are truncated because taking the gradient of this potential makes them unobservant.

The final form of mapping matrix  $\mathbf{f}_{\text{exb},j}$  is for the  $j$ th point:

$$\delta V_j = \begin{bmatrix} f_{\text{exb},1}^{11} & f_{\text{exb},1}^{21} & \cdots & f_{\text{exb},1}^{l_{\text{max}}l_{\text{max}}} & f_{\text{exb},2}^{11} & f_{\text{exb},2}^{21} & \cdots & f_{\text{exb},2}^{l_{\text{max}}l_{\text{max}}} \end{bmatrix}_j \tilde{\mathbf{x}}_{\text{exb}} \quad (\text{A.7})$$

where

$$f_{\text{exb},1}^{lm} = N_l^m P_l^m(\ell(r, \theta)) \cos \phi \quad (\text{A.8})$$

and

$$f_{\text{exb},2}^{lm} = N_l^m P_l^m(\ell(r, \theta)) \sin \phi \quad (\text{A.9})$$

Stacking the  $j$  matrices  $\mathbf{f}_{\text{exb},j}$ , the mapping matrix  $\mathbf{F}_{\text{exb}}$  that transform the state to potential field becomes:

$$\mathbf{F}_{\text{exb}} = [\mathbf{f}_{\text{exb},1}^T \mathbf{f}_{\text{exb},2}^T \cdots \mathbf{f}_{\text{exb},j}^T \cdots \mathbf{f}_{\text{exb},j_{\text{end}}}^T]^T \quad (\text{A.10})$$

The  $\mathbf{F}_{\text{exb}}$  matrix maps coefficients to potential. Next, the  $\mathbf{H}_{\text{exb}}$  matrix, which maps coefficients to density rate, is computed. There are 3 major operations to calculate the  $\mathbf{H}_{\text{exb}}$  mapping matrix: taking the gradient of the potential, crossing the potential with the magnetic field to obtain the drifts and then calculating the divergence to obtain the transport term. Substituting the expansion of the potential field described in Equation A.1 in the ion drift definition from Equation 2.10:

$$\delta \vec{v}_{\text{exb}} = \underbrace{\frac{-\nabla \delta V \times \vec{B}_0}{B_0^2}}_{\vec{v}_{0,\text{exb}}} + \underbrace{\frac{-\nabla(N_l^m P_l^m(\ell)\Phi_l^m) \times \vec{B}_0}{B_0^2}}_{\delta \vec{v}_{\text{exb}}} \quad (\text{A.11})$$

Then, the correction term  $\delta \vec{v}_{\text{exb}}$  can be substituted in the corresponding transport term  $\delta a_{\text{exb}}$ , defined in Equation 2.9:

$$\delta a_{\text{exb}} = -\vec{\nabla} \cdot \left( N \frac{-\nabla(N_l^m P_l^m(\ell)\Phi_l^m) \times \vec{B}_0}{B_0^2} \right) \quad (\text{A.12})$$

$$= \underbrace{\begin{bmatrix} h_{exb}^{11} & h_{exb}^{21} & \dots & h_{exb}^{l_{\max}l_{\max}} \end{bmatrix}}_{\mathbf{h}_{exb,j}} \tilde{\mathbf{x}}_{exb} \quad (\text{A.13})$$

where

$$h_{exb,j}^{lm} = \begin{bmatrix} a_1 & a_2 \end{bmatrix}_j \begin{bmatrix} b_{1c}^{11} & b_{1c}^{21} & \dots & b_{1c}^{l_{\max}l_{\max}} & b_{1s}^{11} & b_{1c}^{21} & \dots & b_{1c}^{l_{\max}l_{\max}} \\ b_{2c}^{11} & b_{2c}^{21} & \dots & b_{2c}^{l_{\max}l_{\max}} & b_{2s}^{11} & b_{2c}^{21} & \dots & b_{2c}^{l_{\max}l_{\max}} \end{bmatrix}_j \quad (\text{A.14})$$

and

$$a_1 = -\frac{1}{B_0^2 r \sin \theta} (B_{0\theta} \frac{\partial N}{\partial r} - \frac{B_{0r}}{r} \frac{\partial N}{\partial \theta}) + \frac{6N(1 + \cos^2 \theta)}{\|\vec{B}_0\| r^2 (1 + 3 \cos^2 \theta)^{3/2}} \quad (\text{A.15})$$

$$a_2 = \frac{1}{B_0^2 r \sin \theta} (\frac{2B_{0r}}{\tan \theta} + B_{0\theta}) \frac{\partial N}{\partial \phi} \quad (\text{A.16})$$

$$b_{1c}^{lm} = -m N_l^m P_l^m(\ell) \sin(m\phi) \quad (\text{A.17})$$

$$b_{1s}^{lm} = +m N_l^m P_l^m(\ell) \cos(m\phi) \quad (\text{A.18})$$

$$b_{2c}^{lm} = \frac{2N_l^m P_{l(\ell)}^m \cos(m\phi)}{R_e \sin^2 \theta (L_{\max} - L_{\min})} \quad (\text{A.19})$$

$$b_{2s}^{lm} = \frac{2N_l^m P_{l(\ell)}^m \sin(m\phi)}{R_e \sin^2 \theta (L_{\max} - L_{\min})} \cos(m\phi) \quad (\text{A.20})$$

Where the magnetic field  $\vec{B}_0$  is described as  $\vec{B}_0 = (B_{0r}, B_{0\theta}, B_{0\phi})$  in spherical coordinates and  $P_{l(\ell)}^m$  is the derivative of the Legendre polynomial with respect to  $\ell$ . It was derived in [45] as:

$$P_{l(\ell)}^m(\ell) = \begin{cases} \frac{\sqrt{l(l+1)}}{\sqrt{1-\ell^2}} P_l^1 & m = 0 \\ \frac{m \ell P_l^m - \sqrt{(l+m)(l-m+1)} \sqrt{1-\ell^2} P_l^{m-1}}{1-\ell^2} & m > 0 \end{cases} \quad (\text{A.21})$$

APPENDIX B  
MAPPING MATRIX  $H_U$  FOR VECTOR SPHERICAL HARMONIC  
DERIVATION

In this section we describe the terms that form the matrix  $\mathbf{H}_u$  defined in Equation 3.9 when using a vector spherical harmonics expansion for the neutral winds vector. Starting at Equation 3.9, we substitute for each  $j$ th grid point the vector spherical harmonics expansion  $\vec{u}$ , defined in Eq. 3.1. The expression is repeated below, but implying the sum with the  $lm$  indices for brevity.

$$\vec{u} = Y_{lm}\hat{r} + r\vec{\nabla}Y_{lm} + \vec{r} \times \vec{\nabla}Y_{lm} \quad (\text{B.1})$$

where

$$Y_{lm} = N_l^m P_l^m(\cos\theta)\Phi_l^m(\phi) \quad (\text{B.2})$$

The gradient of the scalar spherical harmonic function  $\vec{\nabla}Y_{lm}$  is calculated below. The coordinates selected used to describe the vector are the magnetic coordinates radial distance, colatitude and longitude:  $r, \theta, \phi$ . We take into consideration that  $Y_{lm} = Y_{lm}(\theta, \phi)$ .

$$\vec{\nabla}Y_{lm} = Y_{lm(r)}\hat{r} + \frac{1}{r}Y_{lm(\theta)}\hat{\theta} + \frac{1}{r\sin(\theta)}Y_{lm(\phi)}\hat{\phi} \quad (\text{B.3})$$

where  $\hat{r}, \hat{\theta}, \hat{\phi}$  are the unit vectors of a spherical coordinate system and a subscript “(x)” indicates the partial derivative with respect to  $x$ . We substitute the gradient in the vector spherical harmonics definition and we impose negligible vertical winds (in the  $\hat{r}$  vertical direction):

$$\vec{u} = \underbrace{\left(Y_{lm(\theta)}\hat{\theta} + \frac{1}{\sin(\theta)}Y_{lm(\phi)}\hat{\phi}\right)}_{r\vec{\nabla}Y_{lm}} + \underbrace{\left(\frac{-1}{\sin(\theta)}Y_{lm(\phi)}\hat{\theta} + Y_{lm(\theta)}\hat{\phi}\right)}_{\vec{r} \times \vec{\nabla}Y_{lm}} \quad (\text{B.4})$$

Rearranging the equation:

$$\vec{u} = \left(Y_{lm(\theta)} - \frac{1}{\sin(\theta)}Y_{lm(\phi)}\right)\hat{\theta} + \left(Y_{lm(\theta)} + \frac{1}{\sin(\theta)}Y_{lm(\phi)}\right)\hat{\phi} \quad (\text{B.5})$$

We can express it as:

$$\vec{u} = u_\theta\hat{\theta} + u_\phi\hat{\phi} \quad (\text{B.6})$$

where

$$u_\theta = (Y_{lm(\theta)} - \frac{1}{\sin(\theta)} Y_{lm(\phi)}) \quad (\text{B.7})$$

$$u_\phi = (Y_{lm(\theta)} + \frac{1}{\sin(\theta)} Y_{lm(\phi)}) \quad (\text{B.8})$$

We define

$$Y_{lm(\theta)} \equiv N_l^m P_l^m \Phi_{l(\theta)}^m \quad (\text{B.9})$$

and

$$Y_{lm(\phi)} \equiv N_l^m P_l^m \Phi_{l(\phi)}^m \quad (\text{B.10})$$

, where the subscript in parentheses indicates the partial derivative, such that:

$$\Phi_{l(\phi)}^m = m[-x_c^{lm} \sin(m\phi) + x_s^{lm} \cos(m\phi)] \quad (\text{B.11})$$

$$P_{l(\theta)}^m(\cos \theta) = \frac{P_l^m(\cos \theta)}{d\theta} = \frac{P_l^m(\cos \theta) \cos \theta}{d \cos \theta} \frac{d \cos \theta}{d\theta} = -\sin(\theta) P_{l(\cos \theta)}^m \quad (\text{B.12})$$

The first derivative of  $P_{lm}(\cos \theta)$  with respect to  $\cos \theta$  is derived in [45] and is described in Equation A.21 with  $\ell = \cos \theta$ .

The last step to obtain the mapping matrix is to substitute the expressions of  $\vec{u}$  derived in Equation B.5 into the  $a_u$  term defined at each  $j$ th grid point.

$$a_u = -\vec{\nabla} \cdot [N \underbrace{[(u_\theta \hat{\theta} + u_\phi \hat{\phi}) \cdot \hat{b}] \hat{b}}_{\vec{u}}] \quad (\text{B.13})$$

To transform the magnetic field unit vector  $\hat{b}$  to the geomagnetic coordinate system, the inclination angle  $I$ , defined in Section 2.2, is used. The magnetic zonal direction  $\hat{\phi}$  is perpendicular to  $\hat{b}$ , so component  $u_\phi$  doesn't contribute to the projection in the parallel direction. Vector  $\hat{b}$  can be expressed in magnetic coordinates as  $\hat{b} = -\sin I \hat{r} - \cos I \hat{\theta}$ , so that Equation B.13 can be expressed as:

$$a_u = \vec{\nabla} \cdot (N u_\theta \cos I \hat{b}) = \vec{\nabla} \cdot (N \cos I \hat{b}) u_\theta + N \cos I (\vec{\nabla} u_\theta \cdot \hat{b}) \quad (\text{B.14})$$

We divide the expression in Equation B.14 into terms multiplying  $u_\theta$  and terms multiplying  $\vec{\nabla} u_\theta \cdot \hat{b}$  by expanding the gradient component:

$$a_u = [(\vec{\nabla} \cos I) \cdot \hat{b}N + (\vec{\nabla} N) \cdot \hat{b} \cos I + (\vec{\nabla} \cdot \hat{b})N \cos I]u_\theta + [N \cos I](\vec{\nabla} u_\theta \cdot \hat{b}) \quad (\text{B.15})$$

so that:

$$a_u = a_u|_{u_\theta} u_\theta + a_u|_{\vec{\nabla} u_\theta \cdot \hat{b}} (\vec{\nabla} u_\theta \cdot \hat{b}) \quad (\text{B.16})$$

Equation B.14 gives us the definition of:

$$a_u|_{u_\theta} = (\vec{\nabla} \cos I) \cdot \hat{b}N + (\vec{\nabla} N) \cdot \hat{b} \cos I + (\vec{\nabla} \cdot \hat{b})N \cos I \quad (\text{B.17})$$

and

$$a_u|_{\vec{\nabla} u_\theta \cdot \hat{b}} = N \cos I \quad (\text{B.18})$$

The component  $u_\theta$ , which the transport term  $a_u$  depends on, can also be expanded following Equation B.5:

$$u_\theta = P_{l(\theta)}^m \Phi_l^m + \frac{1}{-\sin \theta} P_l^m \Phi_{l(\phi)}^m \quad (\text{B.19})$$

where we can define:

$$u_\theta = u_\theta|_\Phi \Phi_l^m + u_\theta|_{\Phi_\phi} \Phi_{l(\phi)}^m \quad (\text{B.20})$$

Using Equation B.19, we obtain:

$$u_\theta|_\Phi = P_{l(\theta)}^m \quad (\text{B.21})$$

and

$$u_\theta|_{\Phi_\phi} = \frac{1}{-\sin \theta} P_l^m \quad (\text{B.22})$$

Next, to define all the terms that compose the transport term  $a_u$  in Equation B.16, we calculate the product of the divergence of  $u_\theta$  and  $\hat{b}$ , considering that  $u_\theta = u_\theta(\theta, \phi)$  and that  $\hat{b}$  direction doesn't include the  $\hat{\phi}$ .

$$\vec{\nabla} u_\theta \cdot \hat{b} = \left( \frac{\partial u_\theta}{\partial r} \hat{r} + \frac{1}{r} \frac{\partial u_\theta}{\partial \theta} \hat{\theta} + \frac{1}{r \sin(\theta)} \frac{\partial u_\theta}{\partial \phi} \hat{\phi} \right) \cdot \hat{b} = \frac{-\cos I}{r} \frac{\partial u_\theta}{\partial \theta} \quad (\text{B.23})$$

Since  $u_\theta$  in Equation B.5 contained first derivatives of  $Y_{lm}$ ,  $\frac{\partial u_\theta}{\partial \theta}$  in Eq. B.23 will require second derivatives. Second derivatives of the scalar spherical harmonics  $Y_{lm}$  are introduced in the derivation:  $Y_{lm(\theta\theta)} \equiv P_{l(\theta\theta)}^m \Phi_l^m$  and  $Y_{lm(\theta\phi)} \equiv P_{l(\theta)}^m \Phi_{l(\phi)}^m$ , where:

$$P_{l(\theta\theta)}^m \equiv \frac{\partial^2}{\partial \theta^2} (P_l^m(\cos \theta)) = -\cos \theta P_{l(\cos \theta)}^m + \sin^2 \theta P_{l(\cos \theta \cos \theta)}^m \quad (\text{B.24})$$

and second derivative with respect to  $\cos \theta$ ,  $P_{l(\cos \theta \cos \theta)}^m$ , is obtained by deriving the first derivative, derived in [45] and defined in Equation A.21 with  $\ell = \cos \theta$ :

$$P_{l(\cos \theta \cos \theta)}^m = \begin{cases} \frac{m P_l^m + m \cos \theta (1 - \cos^2 \theta) P_{l(\cos \theta)}^m}{(1 - \cos^2 \theta) \sqrt{1 - \cos^2 \theta}} - \sqrt{(l+m)(l-m+1)} P_{l(\cos \theta)}^{m-1} & m > 0 \\ \sqrt{l(l+1)} \frac{P_{l(\cos \theta)}^1 (1 - \cos^2 \theta) + \cos \theta P_l^1}{(1 - \cos^2 \theta) \sqrt{1 - \cos^2 \theta}} & m = 0 \end{cases} \quad (\text{B.25})$$

So,

$$\begin{aligned} \vec{\nabla} \delta u_\theta \cdot \hat{b} &= \frac{-\cos I}{r} \left[ P_{l(\theta\theta)}^m \Phi_l^m + \frac{1}{-\sin \theta} P_{l(\theta)}^m \Phi_{l(\phi)}^m + \frac{-\cos \theta}{\sin^2 \theta} P_l^m \Phi_{l(\phi)}^m \right] \\ &= \frac{-\cos I}{r} P_{l(\theta\theta)}^m \Phi_l^m + \frac{\cos I}{r} \left[ \frac{1}{\sin \theta} P_{l(\theta)}^m + \frac{\cos \theta}{\sin^2 \theta} P_l^m \right] \Phi_{l(\phi)}^m \end{aligned} \quad (\text{B.26})$$

And similarly to the previous steps, we can define:

$$\vec{\nabla} \delta u_\theta \cdot \hat{b} = \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi} \Phi_l^m + \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi_\phi} \Phi_{l(\phi)}^m \quad (\text{B.27})$$

where, using the expression in Equation B.26:

$$\vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi} = \frac{-\cos I}{r} P_{l(\theta\theta)}^m \quad (\text{B.28})$$

and

$$\vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi_\phi} = \frac{\cos I}{r} \left[ \frac{1}{\sin \theta} P_{l(\theta)}^m + \frac{\cos \theta}{\sin^2 \theta} P_l^m \right] \quad (\text{B.29})$$

Substituting the expressions for divergence term  $\vec{\nabla} u_\theta \cdot \hat{b}$  and the term  $u_\theta$  from Equations B.27 and B.20 respectively in Eq. B.16, the  $a_u$  term can be expressed as:

$$a_u = a_u|_{u_\theta} \underbrace{\left( u_\theta|_{\Phi} \Phi_l^m + u_\theta|_{\Phi_\phi} \Phi_{l(\phi)}^m \right)}_{u_\theta} \quad (\text{B.30})$$

$$+ a_u|_{\vec{\nabla} u_\theta \cdot \hat{b}} \underbrace{\left( \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi} \Phi_l^m + \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi_\phi} \Phi_{l(\phi)}^m \right)}_{(\vec{\nabla} u_\theta \cdot \hat{b})} \quad (\text{B.31})$$

where  $\Phi_l^m$  is defined in Equation 3.5,  $\phi_{l(\phi)}^m$  in Equation B.11 and the rest of the terms have been defined in Equations B.17, B.18, B.21, B.22, B.28 and B.29. Rearranging:

$$a_u = \left( a_u|_{u_\theta} u_\theta|_\Phi + a_u|_{\vec{\nabla}u_\theta \cdot \hat{b}} \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_\Phi \right) \Phi_l^m \quad (\text{B.32})$$

$$+ \left( a_u|_{u_\theta} u_\theta|_{\Phi_\phi} + a_u|_{\vec{\nabla}u_\theta \cdot \hat{b}} \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi_\phi} \right) \Phi_{l(\phi)}^m \quad (\text{B.33})$$

We define:

$$a_u = a_{u,1}^{lm} \Phi_l^m + a_{u,2}^{lm} \Phi_{l(\phi)}^m \quad (\text{B.34})$$

where  $a_{u,1}^{lm}$  and  $a_{u,2}^{lm}$  are obtained using the definition from Equation B.32:

$$a_{u,1}^{lm} = a_u|_{u_\theta} u_\theta|_\Phi + a_u|_{\vec{\nabla}u_\theta \cdot \hat{b}} \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_\Phi \quad (\text{B.35})$$

and

$$a_{u,2}^{lm} = a_u|_{u_\theta} u_\theta|_{\Phi_\phi} + a_u|_{\vec{\nabla}u_\theta \cdot \hat{b}} \vec{\nabla} \delta u_\theta \cdot \hat{b} \Big|_{\Phi_\phi} \quad (\text{B.36})$$

Rearranging the  $a_u$  expression of Equation B.34 as a matrix multiplication at each  $j$ th grid point gives:

$$a_{u,j} = \left[ a_{u_1}^{00} \ a_{u_1}^{10} \ \dots \ a_{u_1}^{l_{\max} l_{\max}} \ a_{u_2}^{00} \ a_{u_2}^{10} \ \dots \ a_{u_2}^{l_{\max} l_{\max}} \right]_j \begin{bmatrix} \Phi_1^m \\ \Phi_{1(\phi)}^m \end{bmatrix}_j \quad (\text{B.37})$$

$$= \underbrace{\left[ \mathbf{a}_{u,1} \ \mathbf{a}_{u,2} \right]_j}_{\mathbf{A}_{u_j}} \begin{bmatrix} \Phi_1^m \\ \Phi_{1(\phi)}^m \end{bmatrix}_j \quad (\text{B.38})$$

where  $\mathbf{A}_{u_j}$  is a row matrix, with length equal to the number of coefficients to estimate  $k_u$ .

We can also express  $\Phi_l^m$  and its derivative  $\Phi_{l(\phi)}^m$  as a matrix multiplication. From Equation 3.5, we can define  $\Phi_l^m$ , which is a column matrix with size equal to

number of coefficients  $k_u$  by 1, where half comes from the cosine coefficients  $\mathbf{x}_c$  and the other half from the sine coefficients  $\mathbf{x}_s$ .

$$\Phi_1^m = \begin{bmatrix} b_{1c}^{00} & 0 & \dots & 0 & b_{1s}^{00} & 0 & \dots & 0 \\ 0 & b_{1c}^{10} & \dots & 0 & 0 & b_{1s}^{10} & \dots & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & b_{1c}^{l_{\max}l_{\max}} & 0 & 0 & \dots & b_{1s}^{l_{\max}l_{\max}} \end{bmatrix}_j \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.39})$$

$$= \begin{bmatrix} \text{diag}(\mathbf{b}_{1c}) & \text{diag}(\mathbf{b}_{1s}) \end{bmatrix}_j \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.40})$$

Where  $\mathbf{b}_{1c}$  and  $\mathbf{b}_{1s}$  are row matrices with length equal to the half of the number of coefficients  $\frac{k_u}{2}$ .

$$b_{1c}^{lm} = \cos(m\phi) \quad (\text{B.41})$$

$$b_{1s}^{lm} = \sin(m\phi) \quad (\text{B.42})$$

The subscript ‘‘c’’ or ‘‘s’’ corresponds to the coefficients  $\mathbf{x}_c$  or  $\mathbf{x}_s$  to which the row array is multiplied to.

On the other hand, the first derivative matrix  $\Phi_{1(\phi)}^m$  using Equation B.11 can be expressed as:

$$\Phi_1^m = \begin{bmatrix} b_{2c}^{00} & 0 & \dots & 0 & b_{2s}^{00} & 0 & \dots & 0 \\ 0 & b_{2c}^{10} & \dots & 0 & 0 & b_{2s}^{10} & \dots & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & \dots & b_{2c}^{l_{\max}l_{\max}} & 0 & 0 & \dots & b_{2s}^{l_{\max}l_{\max}} \end{bmatrix}_j \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.43})$$

$$= \begin{bmatrix} \text{diag}(\mathbf{b}_{2c}) & \text{diag}(\mathbf{b}_{2s}) \end{bmatrix}_j \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.44})$$

where  $\mathbf{b}_{2c}$  and  $\mathbf{b}_{2s}$  are row matrices, with length equal to  $\frac{k_u}{2}$ , defined as:

$$b_{21}^{lm} = -m \sin(m\phi) \quad (\text{B.45})$$

$$b_{22}^{lm} = m \cos(m\phi) \quad (\text{B.46})$$

We define  $\mathbf{B}_{11}$ ,  $\mathbf{B}_{12}$ ,  $\mathbf{B}_{21}$  and  $\mathbf{B}_{22}$  matrices as diagonal matrices whose diagonals are formed by the  $\mathbf{b}_{1c}$ ,  $\mathbf{b}_{1s}$ ,  $\mathbf{b}_{2c}$  and  $\mathbf{b}_{2s}$  respectively, from Equations B.40 and B.44.

$$\mathbf{B}_j = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}_j = \begin{bmatrix} \text{diag}(\mathbf{b}_{11}) & \text{diag}(\mathbf{b}_{12}) \\ \text{diag}(\mathbf{b}_{21}) & \text{diag}(\mathbf{b}_{22}) \end{bmatrix}_j \quad (\text{B.47})$$

This allows to express the harmonics function and its derivative, stacking Equations B.40 and B.44, as:

$$\begin{bmatrix} \Phi_1^m \\ \Phi_{1(\phi)}^m \end{bmatrix}_j = \mathbf{B}_j \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.48})$$

Where  $\mathbf{B}_j$  is a square matrix with size  $k_u$  by  $k_u$ .

Using the mapping forms from Equations B.38 and B.48, we can express the transport term  $a_{u,j}$  from Equation B.34 as:

$$a_{u,j} = \underbrace{\mathbf{A}_{u,j} \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}_j}_{\mathbf{h}_{u,j}} \begin{bmatrix} \mathbf{x}_c \\ \mathbf{x}_s \end{bmatrix} \quad (\text{B.49})$$

where mapping matrix  $\mathbf{h}_{u,j}$  is a row matrix with length equal to the number of coefficients to estimate  $k_u$ .

Finally, the last step is to stack the row matrix  $\mathbf{h}_{u,j} = \mathbf{A}_{u,j}\mathbf{B}_j$  to obtain the mapping matrix in the whole region  $\mathbf{H}_u$ :

$$\mathbf{H}_u = [\mathbf{h}_{u,1}^T \ \mathbf{h}_{u,2}^T \ \dots \ \mathbf{h}_{u,j}^T \ \dots \ \mathbf{h}_{u,j_{\text{end}}}^T]^T \quad (\text{B.50})$$

The mapping matrix  $\mathbf{H}_u$  will have size  $j$  by  $k_u$ , where  $j$  is the number of grid points in the EMPIRE domain and  $k_u$  is the number of coefficients we estimate.

APPENDIX C  
MAPPING MATRICES  $F_{U_N}$  AND  $F_{U_E}$  FOR VECTOR SPHERICAL HARMONIC  
DERIVATION

Matrices  $\mathbf{F}_{\mathbf{u}_N}$  and  $\mathbf{F}_{\mathbf{u}_E}$  transform the coefficients, in  $x$ -space, into neutral winds, in  $u$ -space. They are defined in Eqs. 3.11 and 3.12 as:  $\mathbf{u}_N \equiv \mathbf{F}_{\mathbf{u}_N} \mathbf{x}_u$  and  $\mathbf{u}_E \equiv \mathbf{F}_{\mathbf{u}_E} \mathbf{x}_u$ .

To obtain geographic meridional and zonal  $u_N$  and  $u_E$  at each grid point  $j$ , we transform the neutral wind expansion  $\vec{u}$ , expressed in terms of  $u_\theta$  and  $u_\phi$  in Equation B.6, to the geographic coordinate system using the declination angle  $D$ , defined in Section 2.2. The relationship between both coordinate systems is given by:

$$R_{(\hat{r}, \hat{\theta}, \hat{\phi})} = \begin{bmatrix} 0 & 0 & 1 \\ -\sin D & -\cos D & 0 \\ \cos D & -\sin D & 0 \end{bmatrix} R_{(\hat{e}, \hat{n}, \hat{u})} \quad (\text{C.1})$$

The neutral wind expansion  $\vec{u}$ , expressed in terms of  $u_\theta$  and  $u_\phi$  in Equation B.6, is transformed, using C.1, to:

$$\vec{u} = u_\theta \underbrace{(-\cos D \hat{n} - \sin D \hat{e})}_{\hat{\theta}} + u_\phi \underbrace{(-\sin D \hat{n} + \cos D \hat{e})}_{\hat{\phi}} \quad (\text{C.2})$$

Rearranging the terms into the geographic directions,

$$\vec{u} = (-u_\theta \cos D - u_\phi \sin D) \hat{n} + (-u_\theta \sin D + u_\phi \cos D) \hat{e} \quad (\text{C.3})$$

We define:

$$\vec{u} = u_N^{lm} \hat{n} + u_E^{lm} \hat{e} \quad (\text{C.4})$$

with

$$u_N^{lm} = (-u_\theta \cos D - u_\phi \sin D) \quad (\text{C.5})$$

$$u_E^{lm} = (-u_\theta \sin D + u_\phi \cos D) \quad (\text{C.6})$$

We substitute the values of  $u_\theta$  and  $u_\phi$  from Equations B.7 and B.8 to obtain each of the elements. For the meridional, defined in C.5:

$$u_N^{lm} = - \underbrace{\left( Y_{l(\theta)}^m - \frac{1}{\sin \theta} Y_{l(\phi)}^m \right)}_{u_\theta} \cos D - \underbrace{\left( Y_{l(\theta)}^m + \frac{1}{\sin \theta} Y_{l(\phi)}^m \right)}_{u_\phi} \sin D \quad (\text{C.7})$$

Then, we substitute the values of  $Y_{l(\theta)}^m$  and  $Y_{l(\phi)}^m$  from Equations B.9 and B.10 in Equation C.7 and we group the terms together such that:

$$u_N^{lm} = [-P_{l(\theta)}^m (\cos D + \sin D)] \Phi_l^m + \left[ \frac{P_l^m}{\sin \theta} (\cos D - \sin D) \right] \Phi_{l(\phi)}^m \quad (\text{C.8})$$

where we can define:

$$u_N^{lm} = u_{N,1}^{lm} \Phi_l^m + u_{N,2}^{lm} \Phi_{l(\phi)}^m \quad (\text{C.9})$$

with

$$u_{N,1}^{lm} = -P_{l(\theta)}^m (\cos D + \sin D) \quad (\text{C.10})$$

$$u_{N,2}^{lm} = \frac{P_l^m}{\sin \theta} (\cos D - \sin D) \quad (\text{C.11})$$

Similar steps are done to obtain the zonal component  $u_E^{lm}$ . First, we substitute the values of  $u_\theta$  and  $u_\phi$  from Equations B.7 and B.8 in Equation C.6.

$$u_E^{lm} = - \underbrace{\left( Y_{l(\theta)}^m - \frac{1}{\sin \theta} Y_{l(\phi)}^m \right)}_{u_\theta} \sin D + \underbrace{\left( Y_{l(\theta)}^m + \frac{1}{\sin \theta} Y_{l(\phi)}^m \right)}_{u_\phi} \cos D \quad (\text{C.12})$$

Then, we substitute the values of  $Y_{l(\theta)}^m$  and  $Y_{l(\phi)}^m$  from Equations B.9 and B.10 in Equation C.12 and we group the terms together such that:

$$u_E^{lm} = [P_{l(\theta)}^m (\cos D - \sin D)] \Phi_l^m + \left[ \frac{P_l^m}{\sin \theta} (\cos D + \sin D) \right] \Phi_{l(\phi)}^m \quad (\text{C.13})$$

and we define:

$$u_E^{lm} = u_{E,1}^{lm} \Phi_l^m + u_{E,2}^{lm} \Phi_{l(\phi)}^m \quad (\text{C.14})$$

where,

$$u_{E,1}^{lm} = P_{l(\theta)}^m (\cos D - \sin D) \quad (\text{C.15})$$

$$u_{E,2}^{lm} = \frac{P_l^m}{\sin \theta} (\cos D + \sin D) \quad (\text{C.16})$$

To express the neutral winds components of each  $j$ th grid point as a matrix multiplication and to obtain the row matrices  $\mathbf{f}_{\mathbf{u}_N,j}$  and  $\mathbf{f}_{\mathbf{u}_E,j}$  at each  $j$ th grid point, we use the matrix multiplication definition of the harmonic function and its derivative from Equation B.48. The meridional component from Equation C.9 can be expressed as:

$$u_{N,j} = \begin{bmatrix} u_{N,1}^{00} & u_{N,1}^{10} & \dots & u_{N,1}^{l_{\max}l_{\max}} & u_{N,2}^{00} & u_{N,2}^{10} & \dots & u_{N,2}^{l_{\max}l_{\max}} \end{bmatrix}_j \begin{bmatrix} \Phi_1^m \\ \Phi_{l(\phi)}^m \end{bmatrix}_j \quad (\text{C.17})$$

$$= \underbrace{[\mathbf{u}_{N,1} \ \mathbf{u}_{N,2}]_j}_{\mathbf{f}_{\mathbf{u}_N,j}} \mathbf{B}_j \mathbf{x}_u \quad (\text{C.18})$$

Where  $\mathbf{B}_j$  matrix has been defined in Equation B.47 and  $[\mathbf{u}_{N,1} \ \mathbf{u}_{N,2}]_j$  is a row matrix with length equal to the number of coefficients to estimate  $k_u$ . Similarly, the zonal component from Equation C.14 is expressed as:

$$u_{E,j} = \begin{bmatrix} u_{E,1}^{00} & u_{E,1}^{10} & \dots & u_{E,1}^{l_{\max}l_{\max}} & u_{E,2}^{00} & u_{E,2}^{10} & \dots & u_{E,2}^{l_{\max}l_{\max}} \end{bmatrix}_j \begin{bmatrix} \Phi_1^m \\ \Phi_{l(\phi)}^m \end{bmatrix}_j \quad (\text{C.19})$$

$$= \underbrace{[\mathbf{u}_{E,1} \ \mathbf{u}_{E,2}]_j}_{\mathbf{f}_{\mathbf{u}_E,j}} \mathbf{B}_j \mathbf{x}_u \quad (\text{C.20})$$

Stacking the  $j$  row matrices  $\mathbf{f}_{\mathbf{u}_N,j}$  and separately the  $\mathbf{f}_{\mathbf{u}_E,j}$  we obtain the mapping matrices  $\mathbf{F}_{\mathbf{u}_N}$  and  $\mathbf{F}_{\mathbf{u}_E}$  that transform the state to neutral winds space.

$$\mathbf{F}_{\mathbf{u}_N} = \left[ \mathbf{f}_{\mathbf{u}_N,1}^T \ \mathbf{f}_{\mathbf{u}_N,2}^T \ \dots \ \mathbf{f}_{\mathbf{u}_N,j}^T \ \dots \ \mathbf{f}_{\mathbf{u}_N,j_{\text{end}}}^T \right]^T \quad (\text{C.21})$$

$$\mathbf{F}_{\mathbf{u}_E} = \left[ \mathbf{f}_{\mathbf{u}_E,1}^T \ \mathbf{f}_{\mathbf{u}_E,2}^T \ \dots \ \mathbf{f}_{\mathbf{u}_E,j}^T \ \dots \ \mathbf{f}_{\mathbf{u}_E,j_{\text{end}}}^T \right]^T \quad (\text{C.22})$$

## APPENDIX D

MAPPING MATRIX  $H_{FPI}$  VECTOR SPHERICAL HARMONIC DERIVATION

In this section we describe the terms that form the matrix  $\mathbf{H}_{\text{FPI}}$ . This mapping matrix transforms the coefficients, in x-space, into neutral winds projected into a LOS direction, given by an elevation  $el$  and azimuth  $az$  angle. Starting at Equation 5.2, we substitute the VSH expansion of the neutral wind vector from Equation 3.10, to obtain the mapping matrix  $\mathbf{h}_{\text{FPI},i}$  at each  $i$  point where measurements from an FPI are available.

$$u_{\text{FPI},i} = \underbrace{\begin{bmatrix} \cos(el_i) \cos(az_i) & \cos(el_i) \sin(az_i) & \sin(el_i) \end{bmatrix}}_{\mathbf{f}_{\text{LOS},i}} \begin{bmatrix} \mathbf{f}_{\text{uE},i} \\ \mathbf{f}_{\text{uN},i} \\ \mathbf{0} \end{bmatrix} \mathbf{x}_{\text{u}} \quad (\text{D.1})$$

$$= \begin{bmatrix} \cos(el_i) \cos(az_i) \mathbf{f}_{\text{uE},i} + \cos(el_i) \sin(az_i) \mathbf{f}_{\text{uN},i} \end{bmatrix} \mathbf{x}_{\text{u}} \quad (\text{D.2})$$

$$= \mathbf{h}_{\text{FPI},i} \mathbf{x}_{\text{u}} \quad (\text{D.3})$$

where  $\mathbf{f}_{\text{LOS},i}$  is defined in Equation 2.34 and mapping matrices  $\mathbf{f}_{\text{uN},i}$  and  $\mathbf{f}_{\text{uE},i}$  are described in Appendix C in Equations C.18 and C.20 respectively. Stacking the  $i$  row matrices  $\mathbf{h}_{\text{FPI},i}$  we obtain the mapping matrix  $\mathbf{H}_{\text{FPI}}$  that transform the state to neutral winds in a specific look direction.

$$\mathbf{H}_{\text{FPI}} = [\mathbf{h}_{\text{FPI},1}^T \ \mathbf{h}_{\text{FPI},2}^T \ \dots \ \mathbf{h}_{\text{FPI},i}^T \ \dots \ \mathbf{h}_{\text{FPI},i_{\text{end}}}^T]^T \quad (\text{D.4})$$

## BIBLIOGRAPHY

- [1] T. Immel, S. England, S. Mende, R. Heelis, C. Englert, J. Edelstein, H. Frey, E. Korpela, E. Taylor, W. Craig, S. Harris, M. Bester, G. Bust, G. Crowley, J. Forbes, J.-C. Gérard, J. Harlander, J. Huba, B. Hubert, and M. Sirk, “The ionospheric connection explorer mission: Mission goals and design,” *Space Science Reviews*, vol. 214, 12 2017.
- [2] Center for Educational Technologies Wheeling Jesuit University, “Volcanoes and global climate change,” 2015. Available online at: [http://www.cotf.edu/ete//gcc/?/volcanoes\\_teacherpage/](http://www.cotf.edu/ete//gcc/?/volcanoes_teacherpage/), last accessed on 04-14-2022.
- [3] T. F. Tascione, *Introduction to the Space Environment*. Krieger Publisher Company, 2 ed., 2010.
- [4] D. S. Miladinovich, “Data assimilation for ionosphere-thermosphere storm-time state estimation,” *PhD thesis, Illinois Institute of Technology, Department of Mechanical and Aerospace Engineering*, 2018.
- [5] C. Hively, “Space foundation report reveals global space economy at \$ 329 billion in 2016,” 2017.
- [6] D. Nixon, “International space station: architecture beyond earth,” *Circa Press illustrated ed.*, 2016.
- [7] S. Datta-Barua, T. Walter, G. S. Bust, and W. Wanner, “Effects of solar cycle 24 activity on waas navigation,” *Space Weather*, vol. 12, pp. 46–63, 2014. doi:10.1002/2013SW000982.
- [8] Y. Su, “Kilometer-spaced gnss array for ionospheric irregularity monitoring,” *PhD Thesis*, 2017.
- [9] J. J. Makela, D. J. Fisher, J. W. Meriwether, R. A. Buriti, and A. F. Medeiros, “Near-continual ground-based nighttime observations of thermospheric neutral winds and temperatures over equatorial brazil from 2009 to 2012,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 103, pp. 94–102, oct 2013.
- [10] C. J. Heinselman and M. J. Nicolls, “A bayesian approach to electric field and e-region neutral wind estimation with the poker flat advanced modular incoherent scatter radar,” *Radio Science*, vol. 3, no. 5, p. RS5013, 2008.
- [11] S. E. McDonald, D. P. Drob, J. Emmert, C. Englert, D. Siskind, J. Huba, J. Krall, and S. Basu, “The importance of thermospheric winds for ionospheric modeling,” 2012.
- [12] N. Balan, J. Souza, and G. Bailey, “Recent developments in the understanding of equatorial ionization anomaly: A review,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 171, pp. 3–11, 2018. Vertical Coupling in the Atmosphere-Ionosphere System: Recent Progress.
- [13] T. J. Immel, S. England, S. Mende, R. Heelis, C. Englert, J. Edelstein, H. Frey, E. Korpela, E. Taylor, W. Craig, *et al.*, “The ionospheric connection explorer mission: mission goals and design,” *Space Science Reviews*, vol. 214, no. 1, pp. 1–36, 2018.

- [14] J. Sojka, R. Schunk, T. van Eyken, J. Kelly, C. Heinselman, and M. McCready, "Ionospheric challenges of the international polar year," *Eos, Transactions American Geophysical Union*, vol. 88, no. 15, pp. 171–171, 2007.
- [15] T. Matsuo and E. Araujo-Pradere, "Role of thermosphere-ionosphere coupling in a global ionospheric specification," *Radio Science*, vol. 46, 12 2011.
- [16] N. Ssessanga, M. Yamamoto, and S. Saito, "Assessing the performance of a northeast asia japan-centered 3-d ionosphere specification technique during the 2015 st. patrick's day geomagnetic storm," *Earth, Planets and Space*, vol. 73, no. 1, 2021.
- [17] J. Qiao, Y. Liu, Z. Fan, Q. Tang, X. Li, F. Zhang, Y. Song, F. He, C. Zhou, H. Qing, and Z. Li, "Ionospheric tec data assimilation based on gauss–markov kalman filter," *Advances in Space Research*, vol. 68, no. 10, pp. 4189–4204, 2021.
- [18] G. S. Bust, G. Crowley, T. W. Garner, T. L. G. II, R. W. Meggs, C. N. Mitchell, P. S. J. Spencer, P. Yin, and B. Zapfe, "Four-dimensional gps imaging of space weather storms," *Space Weather*, vol. 5, p. S02003, 2007. doi:10.1029/2006SW000237.
- [19] L. Scherliess, R. W. Schunk, J. J. Sojka, D. C. Thompson, and L. Zhu, "Utah state university global assimilation of ionospheric measurements gauss-markov kalman filter model of the ionosphere: Model description and validation," *Journal of Geophysical Research: Space Physics*, vol. 111, no. A11, 2006.
- [20] L. Mandrake, B. Wilson, G. Hajj, C. Wang, X. Pi, and B. Iijima, "Usc/jpl gain: A real-time global ionospheric data assimilation model," *AGU Fall Meeting Abstracts*, vol. -1, p. 01, 11 2004.
- [21] A. Pignalberi, J. Habarulema, M. Pezzopane, and R. Rizzi, "On the development of a method for updating an empirical climatological ionospheric model by means of assimilated vtec measurements from a gnss receiver network," *Space Weather*, vol. 17, no. 7, pp. 1131–1164, 2019.
- [22] R. Song, K. Hattori, X. Zhang, and C. Yoshino, "The three-dimensional ionospheric electron density imaging in japan using the approximate kalman filter algorithm," *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 219, p. 105628, 2021.
- [23] I. Galkin, B. Reinisch, X. Huang, and D. Bilitza, "Assimilation of giro data into a real-time iri," *Radio Science*, vol. 47, no. 4, 2012.
- [24] J. He, X. Yue, and Z. Ren, "The impact of assimilating ionosphere and thermosphere observations on neutral temperature improvement: Observing system simulation experiments using enfk," *Space Weather*, vol. 19, no. 10, p. e2021SW002844, 2021. e2021SW002844 2021SW002844.
- [25] C.-T. Hsu and N. M. Pedatella, "Assessing the impact of icon/mighti zonal and meridional winds on upper atmosphere weather specification in a whole atmosphere data assimilation system: An observing system simulation experiment," *Journal of Geophysical Research: Space Physics*, vol. 126, no. 9, p. e2021JA029275, 2021. e2021JA029275 2021JA029275.

- [26] F. I. Laskar, N. M. Pedatella, M. V. Codrescu, R. W. Eastes, J. S. Evans, A. G. Burns, and W. McClintock, “Impact of gold retrieved thermospheric temperatures on a whole atmosphere data assimilation model,” *Journal of Geophysical Research: Space Physics*, vol. 126, no. 1, p. e2020JA028646, 2021. e2020JA028646 2020JA028646.
- [27] L. Scherliess, D. C. Thompson, and R. W. Schunk, “Ionospheric dynamics and drivers obtained from a physics-based data assimilation model,” *Radio Science*, vol. 44, no. 1, 2009.
- [28] L. Lomidze and L. Scherliess, “Estimation of thermospheric zonal and meridional winds using a kalman filter technique,” *Space Weather*, vol. 13, no. 11, pp. 747–760, 2015.
- [29] S. Datta-Barua, G. S. Bust, G. Crowley, and N. Curtis, “First storm-time plasma velocity estimates from high-resolution ionospheric data assimilation,” *Journal of Geophysical Research: Space Physics*, vol. 118, no. 11, 2013.
- [30] S. Datta-Barua, G. S. Bust, G. Crowley, and N. Curtis, “Neutral wind estimation from 4-d ionospheric electron density images,” *Journal of Geophysical Research*, vol. 114, 2009.
- [31] Gaia Collaboration, Klioner, S. A., Mignard, F., Lindegren, L., Bastian, U., McMillan, P. J., Hernández, J., Hobbs, D., Ramos-Lerate, M., Biermann, M., Bombrun, A., de Torres, A., Gerlach, E., Geyer, R., Hilger, T., Lammers, U., Steidelmüller, H., Stephenson, C. A., Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Babusiaux, C., Creevey, O. L., Evans, D. W., Eyer, L., Hutton, A., Jansen, F., Jordi, C., Luri, X., Panem, C., Pourbaix, D., Randich, S., Sartoretti, P., Soubiran, C., Walton, N. A., Arenou, F., Bailer-Jones, C. A. L., Cropper, M., Drimmel, R., Katz, D., Lattanzi, M. G., van Leeuwen, F., Bakker, J., Castañeda, J., De Angeli, F., Ducourant, C., Fabricius, C., Fouesneau, M., Frémat, Y., Guerra, R., Guerrier, A., Guiraud, J., Jean-Antoine Piccolo, A., Masana, E., Messineo, R., Mowlavi, N., Nicolas, C., Nienartowicz, K., Pailler, F., Panuzzo, P., Riclet, F., Roux, W., Seabroke, G. M., Sordo, R., Tanga, P., Thévenin, F., Gracia-Abril, G., Portell, J., Teyssier, D., Altmann, M., Andrae, R., Bellas-Velidis, I., Benson, K., Berthier, J., Blomme, R., Brugaletta, E., Burgess, P. W., Busso, G., Carry, B., Cellino, A., Cheek, N., Clementini, G., Damerdji, Y., Davidson, M., Delchambre, L., Dell’Oro, A., Fernández-Hernández, J., Galluccio, L., García-Lario, P., Garcia-Reinaldos, M., González-Núñez, J., Gosset, E., Haigron, R., Halbwachs, J.-L., Hambly, N. C., Harrison, D. L., Hatzidimitriou, D., Heiter, U., Hestroffer, D., Hodgkin, S. T., Holl, B., Janßen, K., Jevardat de Fombelle, G., Jordan, S., Krone-Martins, A., Lanzafame, A. C., Löffler, W., Lorca, A., Manteiga, M., Marchal, O., Marrese, P. M., Moitinho, A., Mora, A., Muinonen, K., Osborne, P., Pancino, E., Pauwels, T., Recio-Blanco, A., Richards, P. J., Riello, M., Rimoldini, L., Robin, A. C., Roegiers, T., Rybizki, J., Sarro, L. M., Siopis, C., Smith, M., Sozzetti, A., Ulla, A., Utrilla, E., van Leeuwen, M., van Reeve, W., Abbas, U., Abreu Aramburu, A., Accart, S., Aerts, C., Aguado, J. J., Ajaj, M., Altavilla, G., Álvarez, M. A., Álvarez Cid-Fuentes, J., Alves, J., Anderson, R. I., Anglada Varela, E., Antoja, T., Audard, M., Baines, D., Baker, S. G., Balaguer-Núñez, L., Balbinot, E., Balog, Z., Barache, C., Barbato, D., Barros, M., Barstow, M. A., Bartolomé, S., Bassilana, J.-L., Bauchet, N., Baudesson-Stella, A., Becciani, U., Bellazzini, M., Bernet, M., Bertone, S., Bianchi, L., Blanco-Cuaresma, S., Boch, T., Bossini, D., Bouquillon, S., Bramante, L., Breedt, E., Bressan, A., Brouillet, N., Bucciarelli, B., Burlacu, A., Busonero, D., Butkevich, A. G.,

Buzzi, R., Caffau, E., Cancelliere, R., Cánovas, H., Cantat-Gaudin, T., Carballo, R., Carlucci, T., Carnerero, M. I., Carrasco, J. M., Casamiquela, L., Castellani, M., Castro-Ginard, A., Castro Sampol, P., Chaoul, L., Charlot, P., Chemin, L., Chiavassa, A., Comoretto, G., Cooper, W. J., Cornez, T., Cowell, S., Crifo, F., Crosta, M., Crowley, C., Dafonte, C., Dapergolas, A., David, M., David, P., de Laverny, P., De Luise, F., De March, R., De Ridder, J., de Souza, R., de Teodoro, P., del Peloso, E. F., del Pozo, E., Delgado, A., Delgado, H. E., Delisle, J.-B., Di Matteo, P., Diakite, S., Diener, C., Distefano, E., Dolding, C., Eappachen, D., Enke, H., Esquej, P., Fabre, C., Fabrizio, M., Faigler, S., Fedorets, G., Fernique, P., Fienga, A., Figueras, F., Fouron, C., Fragkoudi, F., Fraile, E., Franke, F., Gai, M., Garabato, D., Garcia-Gutierrez, A., García-Torres, M., Garofalo, A., Gavras, P., Giacobbe, P., Gilmore, G., Girona, S., Giuffrida, G., Gomez, A., Gonzalez-Santamaria, I., González-Vidal, J. J., Granvik, M., Gutiérrez-Sánchez, R., Guy, L. P., Hauser, M., Haywood, M., Helmi, A., Hidalgo, S. L., Hladczuk, N., Holland, G., Huckle, H. E., Jasniewicz, G., Jonker, P. G., Juaristi Campillo, J., Julbe, F., Karbevská, L., Kervella, P., Khanna, S., Kochoska, A., Kordopatis, G., Korn, A. J., Kostrzewa-Rutkowska, Z., Kruszyńska, K., Lambert, S., Lanza, A. F., Lasne, Y., Le Champion, J.-F., Le Fustec, Y., Lebreton, Y., Lebzelter, T., Leccia, S., Leclerc, N., Lecoœur-Taïbi, I., Liao, S., Licata, E., Lindstrøm, H. E. P., Lister, T. A., Livanou, E., Lobel, A., Madrero Pardo, P., Managau, S., Mann, R. G., Marchant, J. M., Marconi, M., Marcos Santos, M. M. S., Marinoni, S., Marocco, F., Marshall, D. J., Martin Polo, L., Martín-Fleitas, J. M., Masip, A., Massari, D., Mastrobuono-Battisti, A., Mazeh, T., Messina, S., Michalik, D., Millar, N. R., Mints, A., Molina, D., Molinaro, R., Molnár, L., Montegriffo, P., Mor, R., Morbidelli, R., Morel, T., Morris, D., Mulone, A. F., Munoz, D., Muraveva, T., Murphy, C. P., Musella, I., Noval, L., Ordénovic, C., Orrù, G., Osinde, J., Pagani, C., Pagano, I., Palaversa, L., Palicio, P. A., Panahi, A., Pawlak, M., Peñalosa Esteller, X., Penttilä, A., Piersimoni, A. M., Pineau, F.-X., Plachy, E., Plum, G., Poggio, E., Poretti, E., Poujoulet, E., Prsa, A., Pulone, L., Racero, E., Ragaini, S., Rainer, M., Raiteri, C. M., Rambaux, N., Ramos, P., Re Fiorentin, P., Regibo, S., Reylé, C., Ripepi, V., Riva, A., Rixon, G., Robichon, N., Robin, C., Roelens, M., Rohrbasser, L., Romero-Gómez, M., Rowell, N., Royer, F., Rybicki, K. A., Sadowski, G., Sagristà Sellés, A., Sahlmann, J., Salgado, J., Salguero, E., Samaras, N., Sanchez Gimenez, V., Sanna, N., Santoveña, R., Sarasso, M., Schultheis, M., Sciacca, E., Segol, M., Segovia, J. C., Ségransan, D., Semeux, D., Siddiqui, H. I., Siebert, A., Siltala, L., Slezak, E., Smart, R. L., Solano, E., Solitro, F., Souami, D., Souchay, J., Spagna, A., Spoto, F., Steele, I. A., Süveges, M., Szabados, L., Szegedi-Elek, E., Taris, F., Tauran, G., Taylor, M. B., Teixeira, R., Thuillot, W., Tonello, N., Torra, F., Torra, J., Turon, C., Unger, N., Vaillant, M., van Dillen, E., Vanel, O., Vecchiato, A., Viala, Y., Vicente, D., Voutsinas, S., Weiler, M., Wevers, T., Wyrzykowski, L., Yoldas, A., Yvard, P., Zhao, H., Zorec, J., Zucker, S., Zurbach, C., and Zwitter, T., “Gaia early data release 3 - acceleration of the solar system from gaia astrometry,” *A&A*, vol. 649, p. A9, 2021.

- [32] D. S. Miladinovich, S. Datta-Barua, G. S. Bust, and J. J. Makela, “Assimilation of thermospheric measurements for ionosphere-thermosphere state estimation,” *Radio Science*, vol. 51, no. 12, pp. 1818–1837, 2016.
- [33] Mignard, F. and Klioner, S., “Analysis of astrometric catalogues with vector spherical harmonics,” *A&A*, vol. 547, p. A59, 2012.
- [34] P. L. K. C. Niels Bormann, William Bell, “Ecmwf/nwp-saf workshop on the

treatment of random and systematic errors in satellite data assimilation.” Virtual Event: ECMWF/EUMETSAT NWP SAF Workshop, 2020.

- [35] A. Vlasenko, P. Korn, J. Riehme, and U. Naumann, “Estimation of data assimilation error: A shallow-water model study,” *Monthly Weather Review*, vol. 142, no. 7, pp. 2502 – 2520, 2014.
- [36] E. Parmuzin, F.-X. L. Dimet, and V. Shutyaev, “On error analysis in variational data assimilation problem for a nonlinear convection–diffusion model,” vol. 21, no. 2, pp. 169–183, 2006.
- [37] S. Gillijns and B. De Moor, “Model error estimation in ensemble data assimilation,” *Nonlinear Processes in Geophysics*, vol. 14, 01 2007.
- [38] J. Durazo, E. J. Kostelich, and A. Mahalov, “Data assimilation for ionospheric space-weather forecasting in the presence of model bias,” *Frontiers in Applied Mathematics and Statistics*, vol. 7, 2021.
- [39] D. Koshin, K. Sato, K. Miyazaki, and S. Watanabe, “An ensemble kalman filter data assimilation system for the whole neutral atmosphere,” 11 2019.
- [40] P. M. Mehta and R. Linares, “A new transformative framework for data assimilation and calibration of physical ionosphere-thermosphere models,” *Space Weather*, vol. 16, no. 8, pp. 1086–1100, 2018.
- [41] H. L. Mitchell and P. L. Houtekamer, “An adaptive ensemble kalman filter,” *Monthly Weather Review*, vol. 128, no. 2, pp. 416 – 433, 2000.
- [42] P. Laloyaux and M. Bonavita, “Improving the handling of model bias in data assimilation,” 2020.
- [43] A. C. Lorenc, “A global three-dimensional multivariate statistical interpolation scheme,” *Monthly Weather Review*, vol. 109, no. 4, pp. 701 – 721, 1981.
- [44] T. Janjić, N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston, “On the representation error in data assimilation,” *Quarterly Journal of the Royal Meteorological Society*, vol. 144, no. 713, pp. 1257–1278, 2018.
- [45] D. S. Miladinovich, S. Datta-Barua, A. L. Rubio, S. Zhang, and G. S. Bust, “Assimilation of guss measurements for estimation of high-latitude convection processes,” *Space Weather*, vol. 18, aug 2020.
- [46] S. Datta-Barua, “Ionospheric threats to space-based augmentation system development,” 2004.
- [47] S. Datta-Barua, T. W. A. J. Mannucci, and P. Enge, “Altitudinal variation of midlatitude localized tec enhancement from ground- and space-based measurements,” *Space Weather*, vol. S10D06, 2008. doi:10.1029/2008SW000396.
- [48] A. T. Chartier, S. Datta-Barua, S. E. McDonald, J. L. Tate, G. S. Bust, G. Romeo, R. K. Schaefer, and L. P. Goncharenko, “Night-time ionospheric localized enhancements (nile) observed in north america following geomagnetic disturbances,” *Earth and Space Science Open Archive*, p. 42, 2021.

- [49] J. Huba, G. Joyce, and J. Fedder, “Sami2 is another model of the ionosphere (sami2): A new low-latitude ionosphere model,” *Journal of Geophysical Research*, vol. 105, no. 23, p. 35, 2000.
- [50] D. P. Drob, J. T. Emmert, J. W. Meriwether, J. J. Makela, E. Doornbos, M. Conde, G. Hernández, J. Noto, K. A. Zawdie, S. E. McDonald, J. D. Huba, and J. H. Klenzing, “An update to the horizontal wind model (hwm): The quiet time thermosphere,” *Earth and Space Science*, vol. 2, jul 2015.
- [51] M. Dhadly, J. Emmert, D. Drob, M. Conde, E. Doornbos, G. Shepherd, J. Makela, Q. Wu, R. Niciejewski, and A. Ridley, “Seasonal dependence of northern high-latitude upper thermospheric winds: A quiet time climatological study based on ground-based and space-based measurements,” *Journal of Geophysical Research: Space Physics*, vol. 122, no. 2, pp. 2619–2644, 2017.
- [52] J. Makela, D. Fisher, J. Meriwether, R. Buriti, and A. Medeiros, “Near-continual ground-based nighttime observations of thermospheric neutral winds and temperatures over equatorial brazil from 2009 to 2012,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 103, 10 2013.
- [53] B. J. Harding, J. Qin, and J. J. Makela, “Ground-based optical measurements of quiet time thermospheric wind and temperature: Atmospheric scattering corrections,” *Journal of Geophysical Research: Space Physics*, vol. 122, no. 11, pp. 11,624–11,632, 2017.
- [54] B. J. Harding, J. J. Makela, J. Qin, D. J. Fisher, C. R. Martinis, J. Noto, and C. M. Wrasse, “Atmospheric scattering effects on ground-based measurements of thermospheric vertical wind, horizontal wind, and temperature,” *Journal of Geophysical Research: Space Physics*, vol. 122, no. 7, pp. 7654–7669, 2017.
- [55] C. R. Englert, J. M. Harlander, C. M. Brown, K. D. Marr, I. J. Miller, J. E. Stump, J. Hancock, J. Q. Peterson, J. Kumler, W. H. Morrow, and et al., “Michelson interferometer for global high-resolution thermospheric imaging (mighti): Instrument design and calibration,” *Space Science Reviews*, vol. 212, no. 1-2, p. 553–584, 2017.
- [56] K. Erickson, “What is the solar cycle?,” Jul 2021. Available online at : <https://spaceplace.nasa.gov/solar-cycles/en>, last accessed on 04-14-2022.
- [57] SILSO, “Silso: World data center for the production, preservation and dissemination of the international sunspot number,” 2022. SILSO — World Data Center for the production, preservation and dissemination of the international sunspot number, Royal Observatory of Belgium, Brussels.
- [58] D. R. Weimer, “An improved model of ionospheric electric potentials including substorm perturbations and application to the geospace environment modeling november 24, 1996, event,,” *Journal of Geophysical Research*, vol. 106, no. A1, p. 407, 2001.
- [59] C. C. Finlay, S. Maus, C. D. Beggan, T. N. Bondar, A. Chambodut, T. A. Chernova, A. Chulliat, V. P. Golovkov, B. Hamilton, M. Hamoudi, R. Holme, G. Hulot, W. Kuang, B. Langlais, V. Lesur, F. J. Lowes, H. Luhr, S. Macmillan, M. Manda, S. McLean, C. Manoj, M. Menvielle, I. Michaelis, N. Olsen, J. Rauberg, M. Rother, T. J. Sabaka, A. Tangborn, L. Toffner-Clausen, E. Thébaud, A. W. P. Thomson, I. Wardinski, Z. Wei, and T. I. Zvereva, “International Geomagnetic Reference Field: the eleventh generation,” *Geophysical Journal International*, vol. 183, pp. 1216–1230, 12 2010.

- [60] R. A. Marshall and C. M. Cully, “Chapter 7 - atmospheric effects and signatures of high-energy electron precipitation,” in *The Dynamic Loss of Earth’s Radiation Belts* (A. N. Jaynes and M. E. Usanova, eds.), pp. 199–255, Elsevier, 2020.
- [61] W. E. Gordon, “Incoherent scattering of radio waves by free electrons with applications to space exploration by radar,” *Proceedings of the IRE*, vol. 46, no. 11, pp. 1824–1829, 1958.
- [62] A. O. Olabode and E. A. Ariyibi, “Geomagnetic storm main phase effect on the equatorial ionosphere over ile-ife as measured from gps observations,” *Scientific African*, vol. 9, p. e00472, 2020.
- [63] M. A. Macleod, “Sporadic e theory. i. collision-geomagnetic equilibrium,” *Journal of the Atmospheric Sciences*, vol. 23, pp. 96–109, jan 1966.
- [64] W. D. Gonzalez, J. A. Joselyn, Y. Kamide, H. W. Kroehl, G. Rostoker, B. T. Tsurutani, and V. M. Vasyliunas, “What is a geomagnetic storm?,” *Journal of Geophysical Research: Space Physics*, vol. 99, no. A4, pp. 5771–5792, 1994.
- [65] C. E. Navia, M. N. de Oliveira, and C. R. A. Augusto, “The highest geomagnetic storms of the solar cycle observed at ground level,” in *Extreme Weather* (P. J. Sallis, ed.), ch. 3, Rijeka: IntechOpen, 2018.
- [66] C. Haines, M. J. Owens, L. Barnard, M. Lockwood, and A. Ruffenach, “The variation of geomagnetic storm duration with intensity,” *Solar Physics*, vol. 294, no. 11, 2019.
- [67] “Geomagnetic dst index,” *World Data Center for Geomagnetism*, 2015.
- [68] R. de Jesus, Y. Sahai, F. Guarnieri, P. Fagundes, A. de Abreu, J. Bittencourt, T. Nagatsuma, C.-S. Huang, H. Lan, and V. Pillat, “Ionospheric response of equatorial and low latitude f-region during the intense geomagnetic storm on 24–25 august 2005,” *Advances in Space Research*, vol. 49, no. 3, pp. 518–529, 2012.
- [69] K. Cole, “Joule Heating of the Upper Atmosphere,” *Australian Journal of Physics*, vol. 15, no. 2, p. 223, 1962.
- [70] E. Blanch, S. Marsal, A. Segarra, J. Torta, D. Altadill, and J. Curto, “Space weather effects on earth’s environment associated to the 24-25 october 2011 geomagnetic storm,” *Space Weather*, vol. 11, pp. 153–168, 04 2013.
- [71] C.-S. Huang, J. C. Foster, and M. C. Kelley, “Long-duration penetration of the interplanetary electric field to the low-latitude ionosphere during the main phase of magnetic storms,” *Journal of Geophysical Research: Space Physics*, vol. 110, no. A11, 2005.
- [72] B. T. Tsurutani, O. P. Verkhoglyadova, A. J. Mannucci, A. Saito, T. Araki, K. Yumoto, T. Tsuda, M. A. Abdu, J. H. A. Sobral, W. D. Gonzalez, H. McCreddie, G. S. Lakhina, and V. M. Vasyliūnas, “Prompt penetration electric fields (ppefs) and their ionospheric effects during the great magnetic storm of 30–31 october 2003,” *Journal of Geophysical Research: Space Physics*, vol. 113, no. A5, 2008.
- [73] F. J. Pavón-Carrasco and A. De Santis, “The south atlantic anomaly: The key for a possible geomagnetic reversal,” *Frontiers in Earth Science*, vol. 4, 2016.

- [74] N. Balan and G. J. Bailey, “Equatorial plasma fountain and its effects: Possibility of an additional layer,” *Journal of Geophysical Research: Space Physics*, vol. 100, no. A11, p. 21421–21432, 1995.
- [75] L. Scherliess and B. G. Fejer, “Radar and satellite global equatorial f region vertical drift model,” *Journal of Geophysical Research: Space Physics*, vol. 104, no. A4, pp. 6829–6842, 1999.
- [76] P. Ghosh, Y. Otsuka, S. Mani, and H. Shinagawa, “Day-to-day variation of pre-reversal enhancement in the equatorial ionosphere based on gaia model simulations,” *Earth, Planets and Space*, vol. 72, no. 1, 2020.
- [77] J. K. Hargreaves, *The Solar-Terrestrial Environment: An Introduction to Geospace - the Science of the Terrestrial Upper Atmosphere, Ionosphere, and Magnetosphere*. Cambridge Atmospheric and Space Science Series, Cambridge University Press, 1992.
- [78] R. Schunk and A. Nagy, *Ionospheres: Physics, Plasma Physics, and Chemistry*. Cambridge Atmospheric and Space Science Series, Cambridge University Press, 2 ed., 2009.
- [79] M. Blanc and A. Richmond, “The ionospheric disturbance dynamo,” *Journal of Geophysical Research: Space Physics*, vol. 85, no. A4, pp. 1669–1686, 1980.
- [80] S. M. Khadka, C. E. Valladares, R. Sheehan, and A. J. Gerrard, “Effects of electric field and neutral wind on the asymmetry of equatorial ionization anomaly,” *Radio Science*, vol. 53, no. 5, pp. 683–697, 2018.
- [81] S. Datta-Barua, “Ionospheric threats to the integrity of airborne gps users,” *PhD thesis, Stanford University, Department of Aeronautics and Astronautics*, 2007.
- [82] J. Foster and P. Erickson, “Ionospheric superstorms: Polarization terminator effects in the atlantic sector,” *Journal of Atmospheric and Solar-Terrestrial Physics*, vol. 103, pp. 147–156, 2013. Recent Advances in Equatorial, Low- and Mid-latitude Aeronomy.
- [83] G. S. Bust and S. Datta-Barua, “Scientific investigations using ida4d and empire,” in *Modeling the Ionosphere-Thermosphere System* (J. Huba, R. Schunk, and G. Khazanov, eds.), pp. 283–297, John Wiley & Sons, Ltd, 1 ed., 2014.
- [84] D. Bilitza, “The international reference ionosphere: Rawer’s iri and its status today,” *Advances in Radio Science*, vol. 12, pp. 231–236, nov 2014.
- [85] J. M. Picone, A. E. Hedin, D. P. Drob, and A. C. Aikin, “Nrlmsise-00 empirical model of the atmosphere: Statistical comparisons and scientific issues,” *Journal of Geophysical Research: Space Physics*, vol. 107, pp. SIA 15–1–SIA 15–16, dec 2002. doi: <https://doi.org/10.1029/2002JA009430>.
- [86] G. Kirchengast, “The graz ionospheric flux tube simulation model,” *Solar-Terrestrial Energy Program: Handbook of Ionospheric Models (edited by R. W. Schunk)*, vol. 23, pp. 96–109, 1996.
- [87] J. D. Huba, G. Joyce, and J. Krall, “Three-dimensional equatorial spread f modeling,” *Geophysical Research Letters*, vol. 35, no. 10, 2008.

- [88] J. D. Huba, G. Joyce, J. Krall, C. L. Siefring, and P. A. Bernhardt, “Self-consistent modeling of equatorial dawn density depletions with sami3,” *Geophysical Research Letters*, vol. 37, no. 3, 2010.
- [89] M. Abramowitz and I. Stegun, “Handbook of mathematical functions,” *New York: Dover Publications*, 1965. Accessed <http://people.math.sfu.ca/cbm/aands/>.
- [90] R. G. Barrera, G. A. Estevez, and J. Giraldo, “Vector spherical harmonics and their application to magnetostatics,” *European Journal of Physics*, vol. 6, pp. 287–294, oct 1985.
- [91] A. D. Richmond, “Ionospheric electrodynamics using magnetic apex coordinates,” *Journal of geomagnetism and geoelectricity*, vol. 47, no. 2, pp. 191–212, 1995.
- [92] A. Ben-Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications, 2nd ed.* New York: Springer, 2003.
- [93] L. A. Navarro and B. G. Fejer, “Storm-time thermospheric winds over peru,” *Journal of Geophysical Research: Space Physics*, vol. 124, no. 12, pp. 10415–10427, 2019.
- [94] L. A. Navarro and B. G. Fejer, “Storm-time coupling of equatorial nighttime f region neutral winds and plasma drifts,” *Journal of Geophysical Research: Space Physics*, vol. 125, no. 9, p. e2020JA028253, 2020. e2020JA028253 2020JA028253.
- [95] J. L. Lean, S. E. McDonald, J. D. Huba, J. T. Emmert, D. P. Drob, and C. L. Siefring, “Geospace variability during the 2008–2009 whole heliosphere intervals,” *Journal of Geophysical Research: Space Physics*, vol. 119, no. 5, pp. 3755–3776, 2014.
- [96] J. Palacios, A. Guerrero, C. Cid, E. Saiz, and Y. Cerrato, “Defining scale thresholds for geomagnetic storms through statistics,” *Natural Hazards and Earth System Sciences Discussions*, vol. 2018, pp. 1–17, 2018.
- [97] J. Krall, J. Huba, V. Jordanova, R. Denton, T. Carranza, and M. Moldwin, “Measurement and modeling of the refilling plasmasphere during 2001,” *Journal of Geophysical Research: Space Physics*, vol. 121, pp. n/a–n/a, 03 2016.
- [98] S. E. McDonald, J. L. Lean, J. D. Huba, G. Joyce, J. T. Emmert, and D. P. Drob, *Long-Term Simulations of the Ionosphere Using SAMI3*, ch. 11, pp. 119–131. American Geophysical Union (AGU), 2014.
- [99] A. J. Mannucci, B. T. Tsurutani, B. A. Iijima, A. Komjathy, A. Saito, W. D. Gonzalez, F. L. Guarnieri, J. U. Kozyra, and R. Skoug, “Dayside global ionospheric response to the major interplanetary events of october 29–30, 2003 “halloween storms”,” *Geophysical Research Letters*, vol. 32, no. 12, 2005.
- [100] G. S. Bust, T. W. Garner, and T. L. Gaussiran II, “Ionospheric data assimilation three-dimensional (ida3d): A global, multisensor, electron density specification algorithm,” *Journal of Geophysical Research: Space Physics*, vol. 109, no. A11, 2004.

- [101] I. Fernandez-Gomez, M. Fedrizzi, M. V. Codrescu, C. Borries, M. Fillion, and T. J. Fuller-Rowell, “On the difference between real-time and research simulations with ctipe,” *Advances in Space Research*, vol. 64, no. 10, pp. 2077–2087, 2019. Variability and Coupling of the Equatorial, Low- and Mid-Latitude Mesosphere, Thermosphere and Ionosphere: Latest Developments of Monitoring and Modeling Techniques.
- [102] G. S. Bust and C. N. Mitchell, “History, current state, and future directions of ionospheric imaging,” *Reviews of Geophysics*, vol. 46, no. 1, p. RG1003, 2008.