Validation of modeled ionospheric properties during geomagnetic storms: NmF2/foF2, hmF2, and TEC

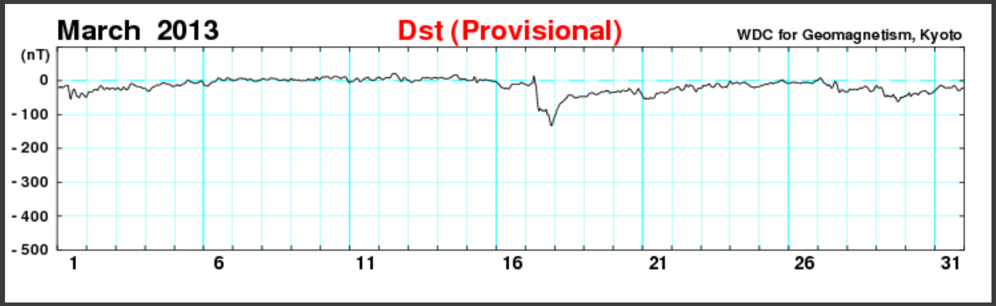
I. Tsagouri

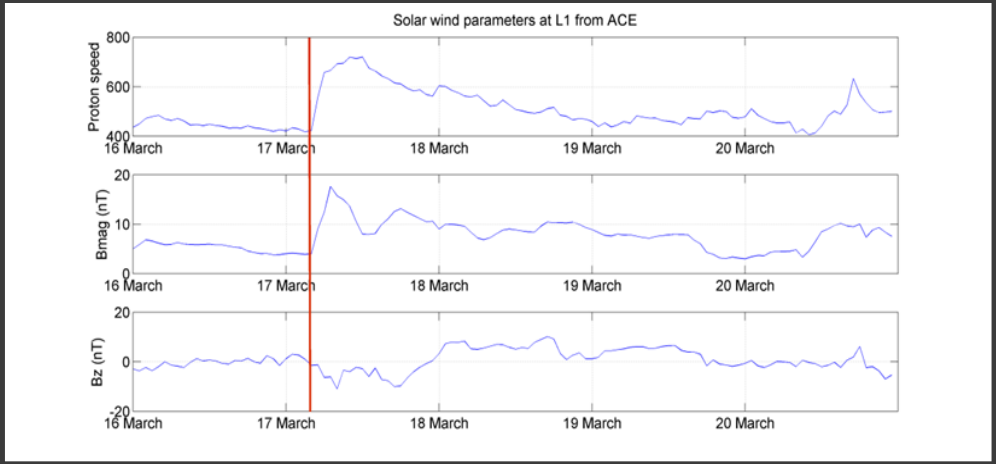
National Observatory of Athens, Greece

CEDAR-GEM Modeling Challenge Session at 2017 CEDAR Workshop,
19 June 2017

Assessment of modeling capabilities in capturing the storm impact (foF2)

Test event: Storm interval 16 – 20 March 2013





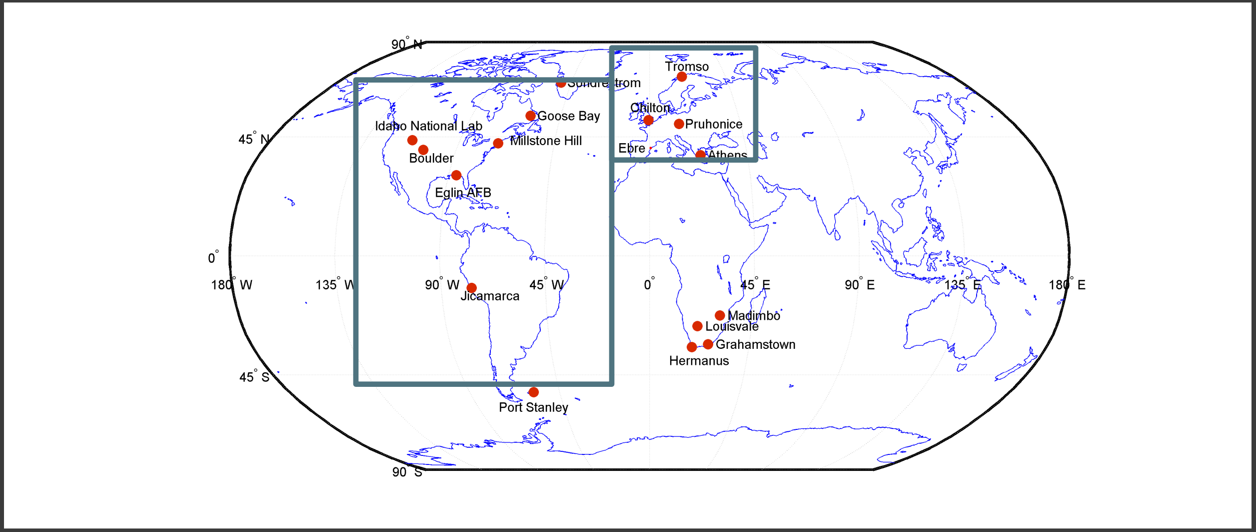
Interplanetary conditions

CME related storm event according to

- ICME list available at http://www.srl.caltech.edu/ACE/ASC/DATA/level3/icmetable2.htm (storm sudden commencement: 17/3/2013, 05:59 UT).

Data presentation

Spatial distribution of the test locations



Data sources:

- Autoscaled values of foF2 from ground-based ionosondes derived from GIRO (http://giro.uml.edu/).

1st challenge: Quantification of the quiet-time ionospheric variation

Available Options

- Average over 5-quietest days within a month
- Average over 5-quietest days within 30-days prior to an event
- Monthly median
- 30-day running median or 30-day median prior to an event – Suitable for "real time" applications

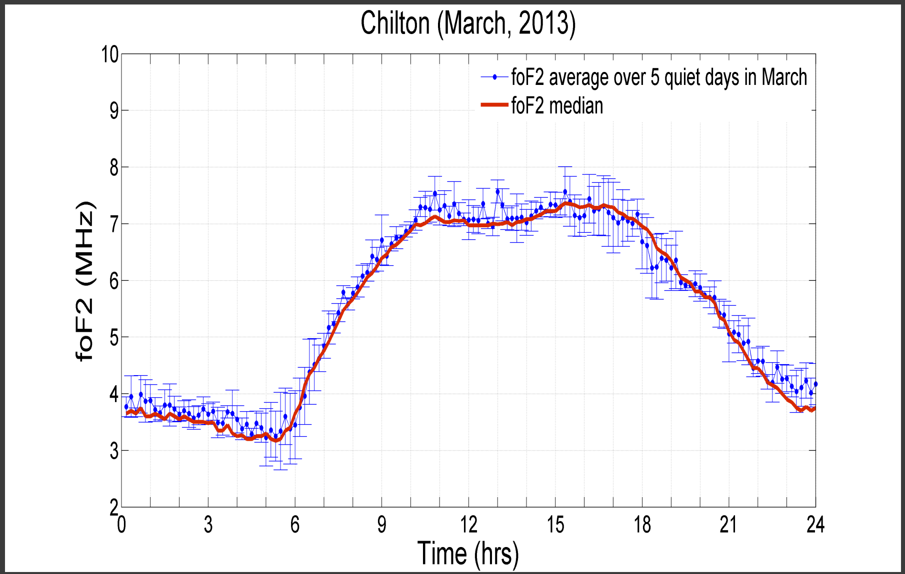
For the selection of the five days we use the following criteria:

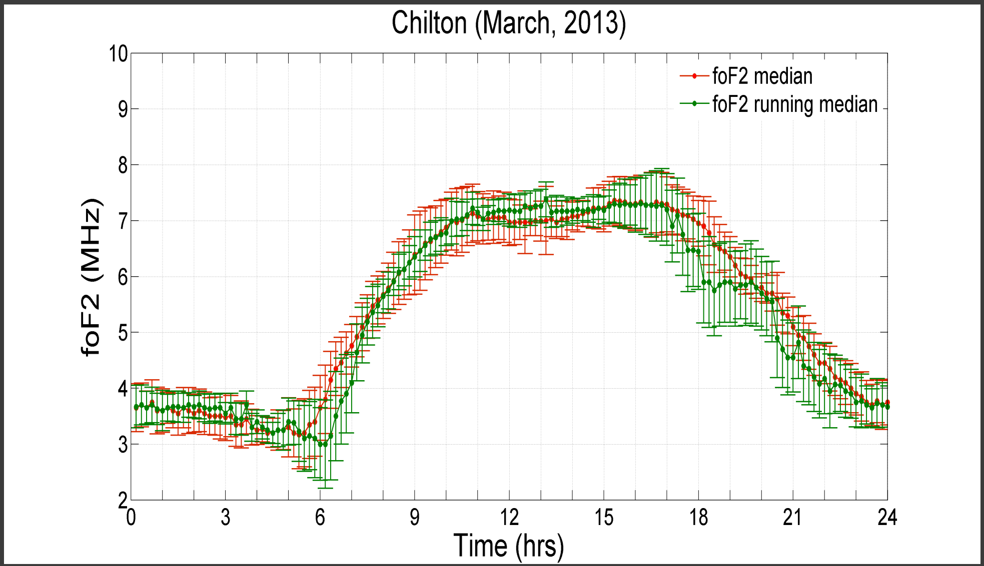
- Min Dst index ≥ -30 nT for the day and the previous one
- Max AE index ≤ 250 nT for the day and the previous one

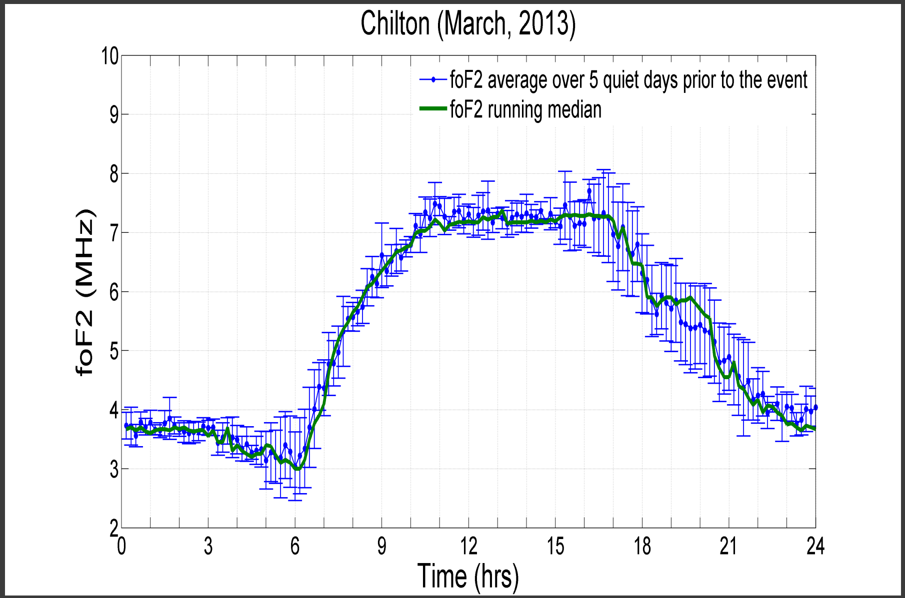
For the test event analyzed here:

Storm time interval	5 Quietest days within the month	5 Quietest days prior to the event
16-20 March 2013	6/3, 7/3, 8/3, 25/3, 26/3	25/2, 27/2, 6/3, 7/3, 8/3

Quiet-time ionospheric variation over Chilton (foF2) – March 2013







Error bars: STDs (Uncertainties mainly due to ionogram autoscaling errors and quiet time variability)

Mean STD monthly medians: **0.4 MHz**
Mean STD 5 quiet days in the month: **0.3 MHz**

Mean STD running medians: **0.4 MHz**
Mean STD 5 quiet days before the event: **0.4 MHz**

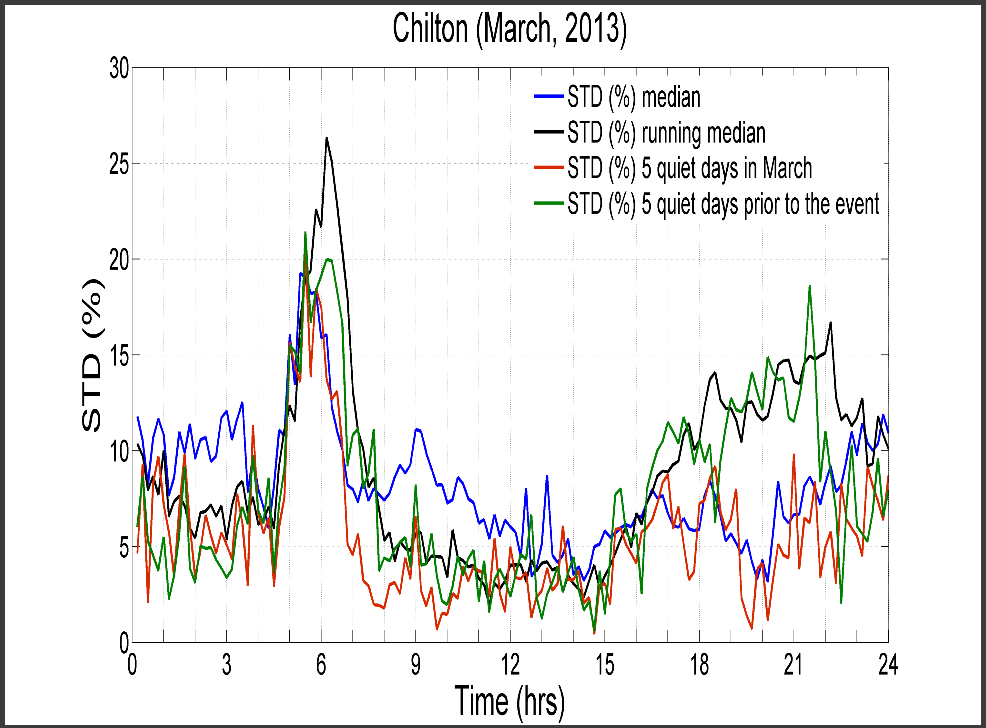
Quiet-time ionospheric variation over Chilton (foF2) – March 2013

STD (%) = (STD_foF2x / foF2x)*100

x: median, running median, average over 5 quiet days

Mean STD (%) monthly medians: **8 %**
Mean STD (%) 5 quiet days in the month: **6 %**
Mean STD (%) running medians: **9%**
Mean STD (%) 5 quiet days before the event: **8%**

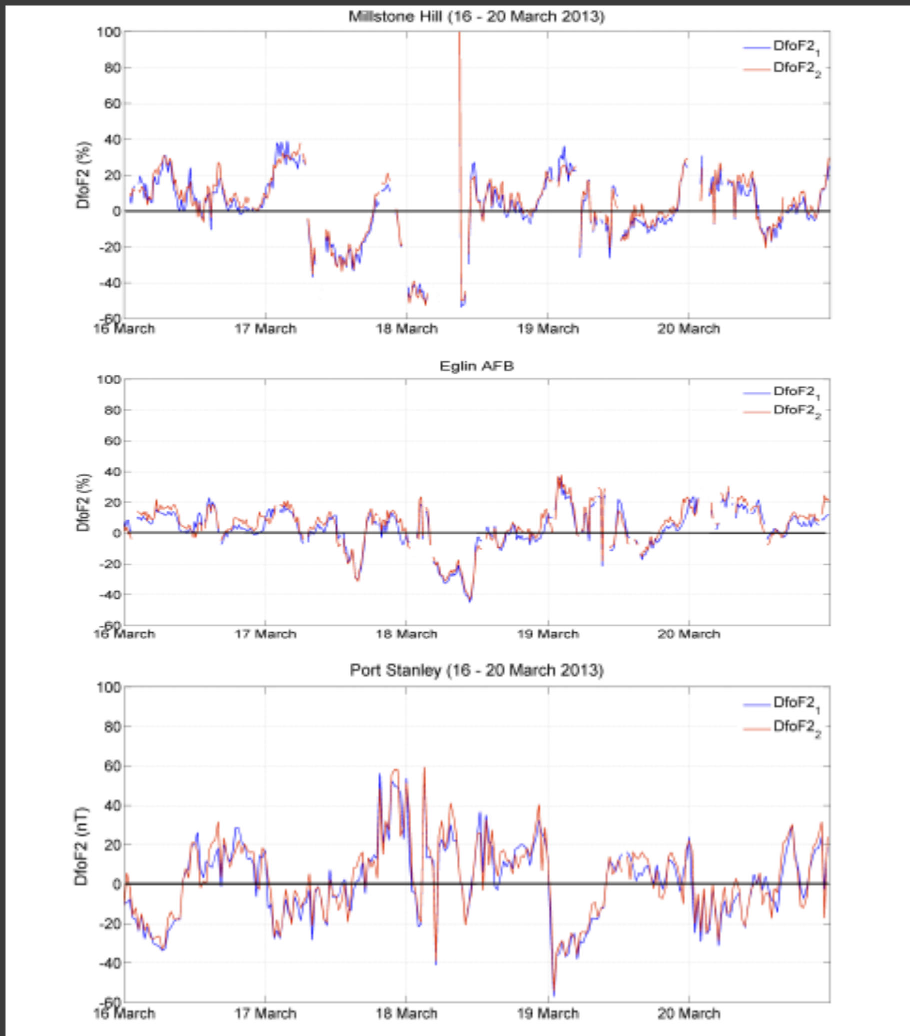
STD (%) is estimated for each observation time of the day

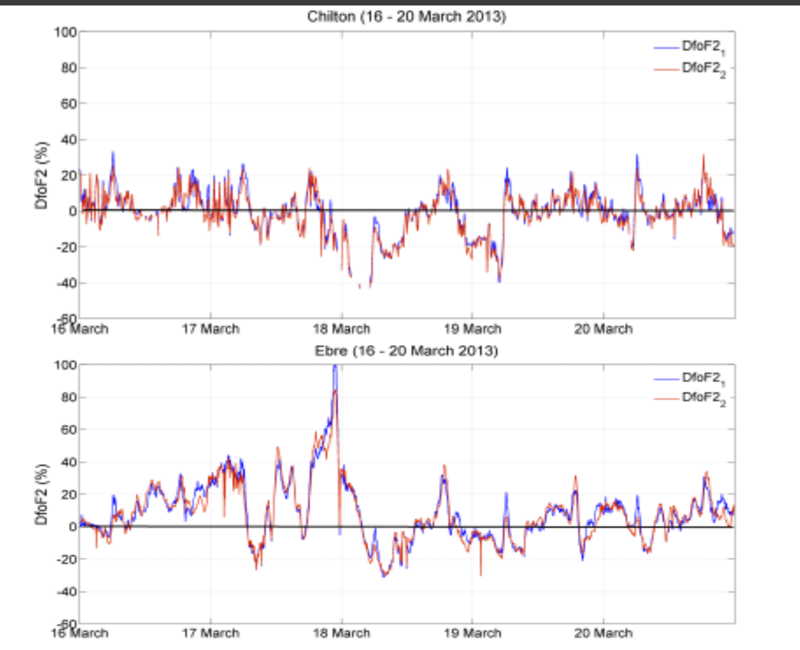


Information that may be extracted

- Local time dependence of the uncertainties: e.g., for the case under study here the uncertainties are significantly larger in dawn sector in all terms (for Chilton UT=LT)
- Monthly medians are comparable to the average of 5 quiet days within the month, while running medians are comparable to the average of 5 quiet days prior to the storm event. On average, all approaches may be considered comparable.
- On average, ionospheric variations of about 10% wrt quiet conditions may be ignored in any case.

Quantification of the storm impact





The foF2 storm-time response is equivalently determined by using two different options for the background conditions in agreement with the previous results.

$$dfoF2_1=[(foF2-foF2_{median\_30\,days\,prior})/foF2_{median\_30\,days\,prior}]*100$$

$$dfoF2_2=[(foF2-foF2_{5\,quiet days})/foF2_{5\,quiet days}]*100$$

Ionospheric storm characteristics (with respect to 30-days median)

Significant disturbances > ± 20% (2STDs in the variability of medians)

Station	dfoF2 > 20%						dfoF2 < -20%					
	Start Time		Max	t_max		Duration	Start time		Min	t_min		Duration
	UT	LT	(%)	UT	LT	(hrs)	UT	LT	(%)	UT	LT	(hrs)
Europe												
Chilton							22.2	22.2	-43%	5.2	5.2	14.0
Pruhonice	11.0	12.0	47%	11.75	12.75	10.5	22.5	23.5	-46%	4.0	5.0	14.0
Ebre	11.5	12.5	101%	22.6	0.6	16.3	3.8	4.8	-31%	7.7	8.7	9.5
Athens	10.5	12.5	83%	22.75	0.75	17.0						
North America												
Idaho Nat. Lab							6.25 5.5	23.25 22.5	-47% -41%	9.5 12.0	2.5 5.0	26.5 8.0
Millstone Hill							7.25 22.5	2.25 17.5	-48% -53%	9.25 9.25	4.25 4.25	11.0 12.25
Boulder							7.0	0.0	-45%	10.0	3.0	19.25
Eglin AFB							4.0	22.0	-45%	10.75	4.75	8.25
South America												
Jicamarca												
Port Stanley	15.5	12.5	58%	3.0	0.5	17.5						

European sector: storm time response is mainly detected in the middle latitudes, and characterized by a initial positive phase that was followed by a negative phase.

American sector: only negative storm effects in the North American sector,
no significant disturbances in Jicamarca.
only positive effects are observed in the South Hemisphere (Port Stanley).

2nd challenge: Modeling the local time dependence of the storm-time response

First test on models' assessment

Models (available at CCMC)

1. IRI2012
2. SAMI3
3. CTIPe
4. TIE-GCM

3rd challenge: Scores for metrics

- ME (Mean Error: modeled to obs) and STDs of the ME, MAE (Mean Absolute Error) as indicators of bias and error magnitudes
- The RMSE and MRE (mean relative error) as indicators of the prediction accuracy.

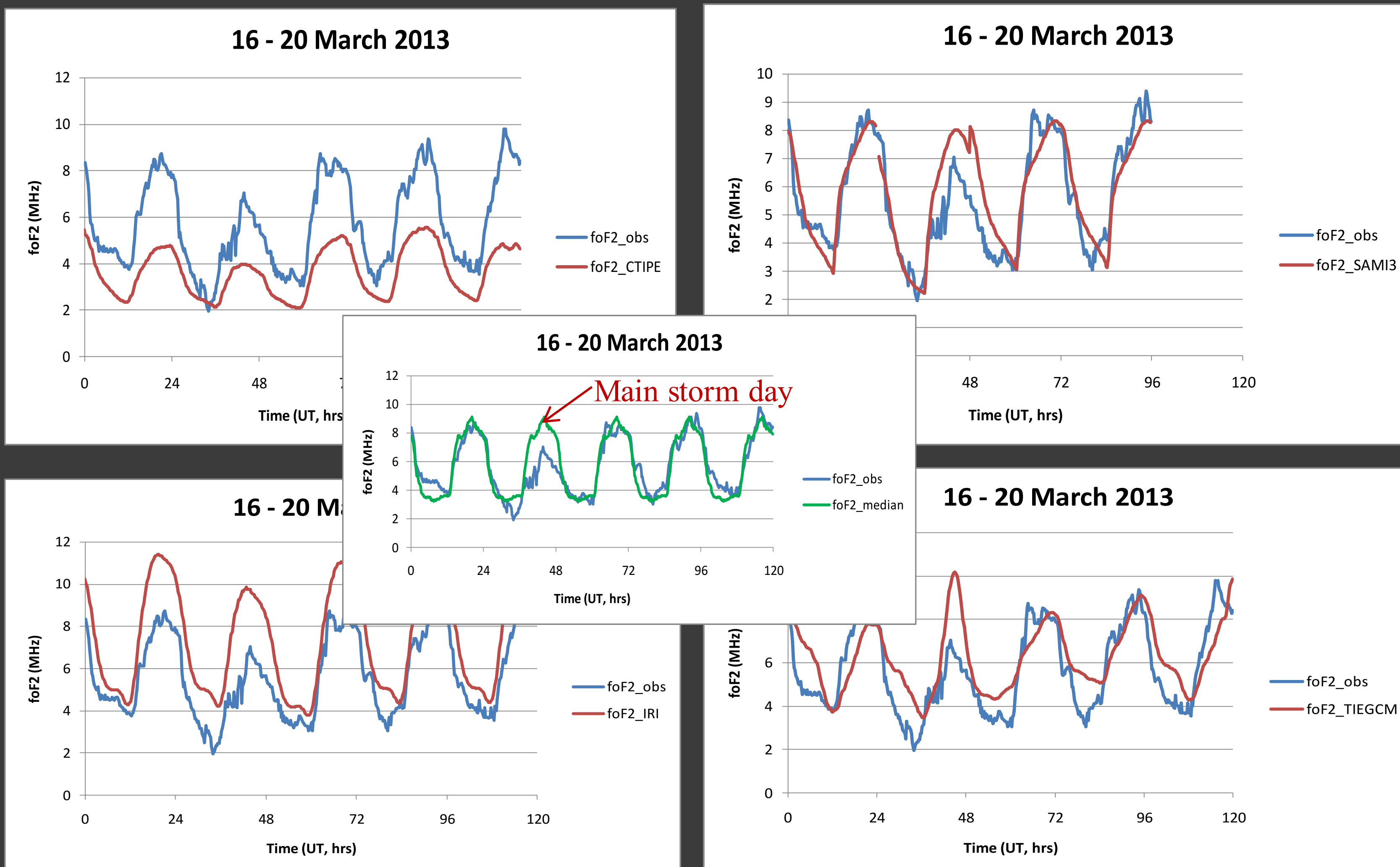
$$MRE = 1/N \sum(|foF2obs - foF2mod|)/foF2obs$$

- Relative improvement over climatology (i.e. median values) to test improvements over standard prediction approaches

$$\% \text{ improvement} = \frac{\text{RMSE}_Y - \text{RMSE}_X}{\text{RMSE}_Y} \times 100,$$

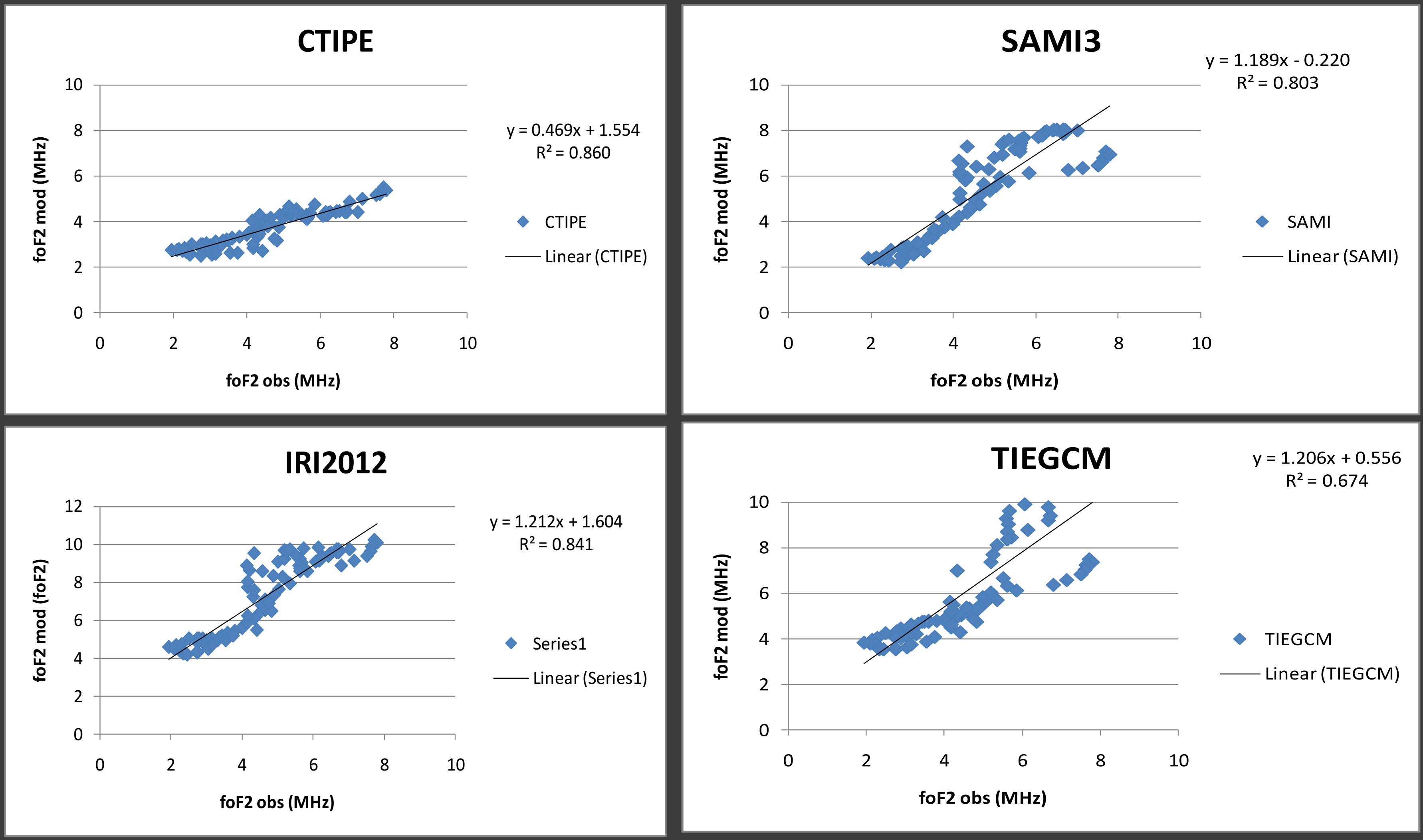
- Correlation coefficient (R) and coefficient of determination (r^2) as indicators of prediction efficiency

Modeling results over Boulder.



The output of IRI2012 (CTIPE) consistently overestimates (underestimates) the observations

Scatter plots of modeled vs observed values over Boulder.



This test demonstrates rather successful performance for IRI2012 and CTIPE

Boulder: Scores calculated over the main storm day based on the observed values (17 March, 2013)

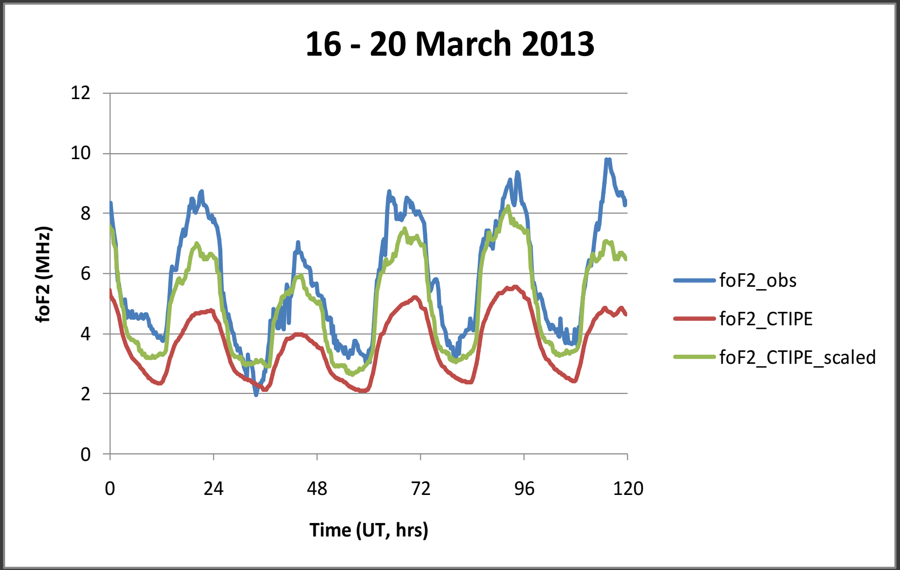
Model	ME (MHz)	MAE (MHz)	STD (MHz)	RMSE (MHz)	MRE (%)	% impr
Median	1.07	1.32	1.26	1.65	29%	-
CTIPe	**-0.91**	**1.02**	**0.88**	**1.26**	**20%**	**24%**
SAMI3	**0.66**	**0.90**	**0.97**	**1.16**	**18%**	**30%**
TIEGCM	1.51	1.59	1.35	2.02	36%	-23%
IRI2012	**2.59**	**2.59**	**0.89**	**2.74**	**60%**	**-20%**

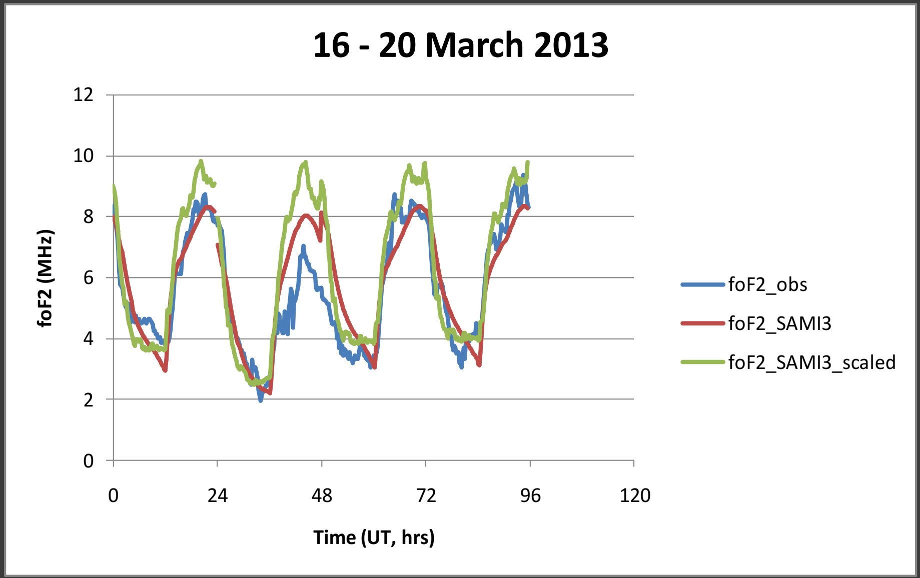
The models' performance over Boulder differ for different scores (green: better red: worse performances).

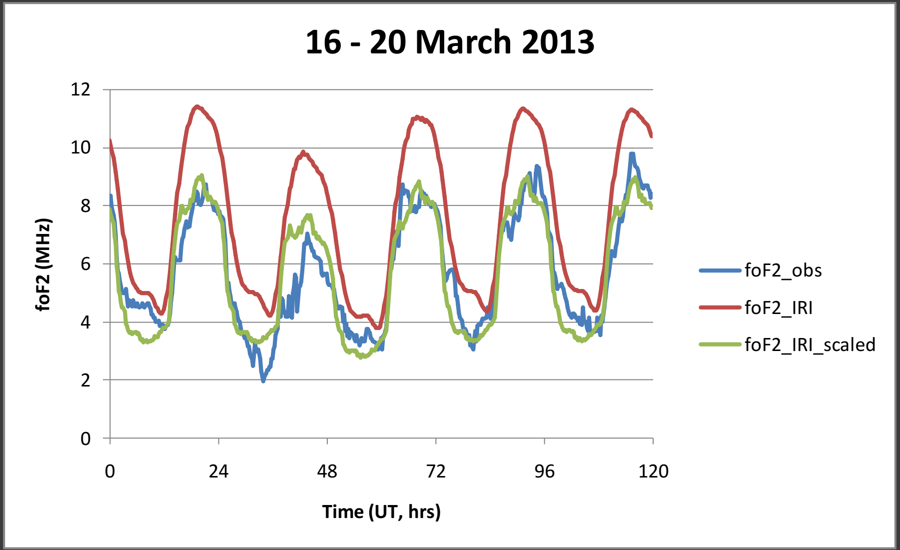
Model	R	R^2
Median	1.12	0.67
CTIPe	**0.47**	**0.86**
SAMI3	**1.19**	0.80
TIEGCM	1.21	0.67
IRI2012	1.21	**0.84**

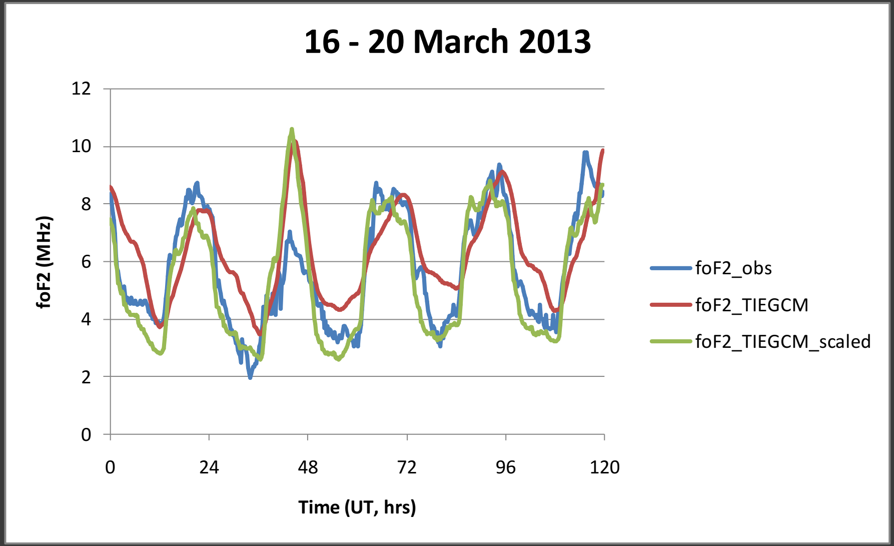
Alternative approach: the assessment of the models' performance in terms of the scaled values

Boulder









Scaled_foF2=foF2mod*(obs_median/mod_median)

Boulder: Scores calculated over the main storm day based on scaled values (17 March, 2013)

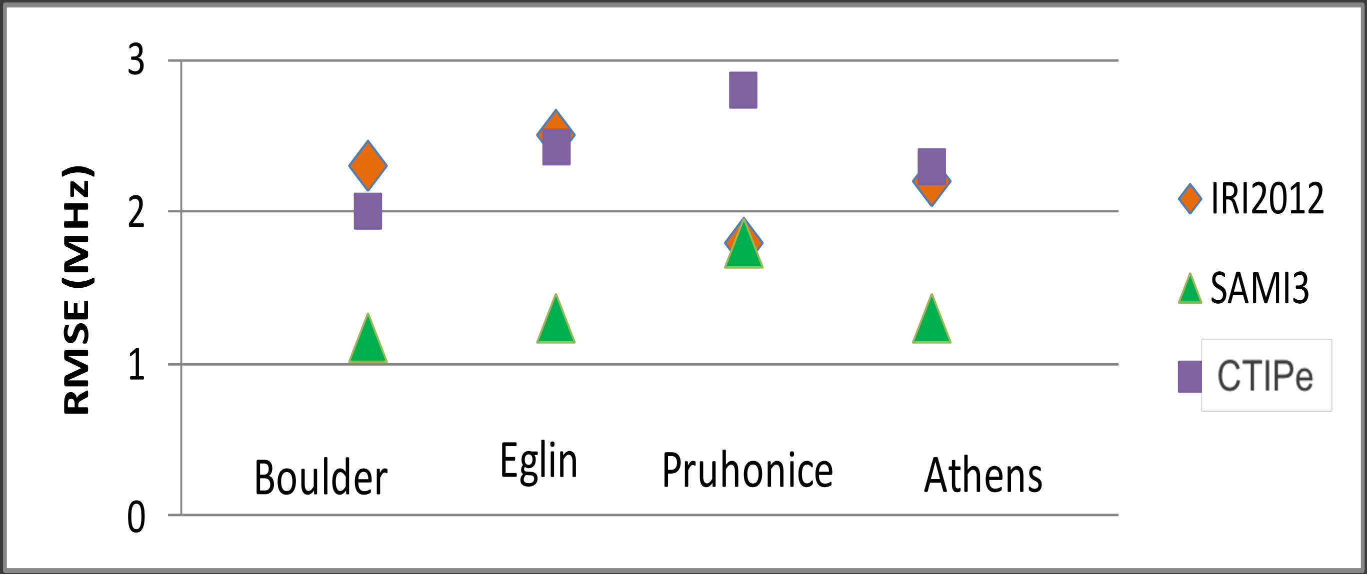
Model	ME (MHz)	MAE (MHz)	STD (MHz)	RMSE (MHz)	MRE (%)	% impr
Median	1.07	1.32	1.26	1.65	29%	-
CTIPe	**0.26**	**0.58**	**0.64**	**0.68**	**13%**	**59%**
SAMI3	**1.07**	**1.45**	**1.59**	**1.90**	**29%**	**-15%**
TIEGCM	**0.73**	**1.24**	**1.99**	**1.65**	**24%**	**0%**
IRI2012	**0.76**	**0.92**	**0.88**	**1.16**	**23%**	**76%**

Model	R	R^2
Median	1.12	0.67
CTIPe	**0.70**	**0.87**
SAMI3	**1.41**	**0.70**
TIEGCM	1.28	0.67
IRI2012	**0.93**	**0.73**

The models' predictive capabilities in capturing the storm impact appear to be improved in most of the cases when the scaled values are considered, but this is not the case for e.g. SAMI3.

The above finding may indicate that some of the modeling limitations may come from shortcomings in modeling the quiet-time (background) variation.

RMSE



The examination of the scores (e.g. RMSE) over different locations indicate that the models' performance may also depend on the location.

Concluding remarks

- The performance of each model depends on the selected scores:
  Fair evaluation of the modeling capabilities requires complementary analysis of a set of scores.
- The fair evaluation of the models' predictive capabilities in capturing the storm impact should include the evaluation of the models' performance in modeling the quiet-time ionospheric variation.
- The performance of each model may depend also on the location:
  Detailed diagnostics require consideration of spatial dependencies (e.g. dependence on the latitude and longitude).

Next challenges:

- Implementation of the proposed methodology to hmF2 and vTEC
- Analysis over a set of storm events and comprehensive assessment of the modeling capabilities.