Advancements in Space Weather Forecasting with Machine Learning: The ACCRUE and Proboost Methods

Enrico Camporeale

CIRES / CU Boulder & NOAA Space Weather Prediction Center

Thanks to: Andong Hu, Brian Swiger, Thomas Berger, Howard Singer, Kent Tobiska

We acknowledge NASA grants 80NSSC20K1580, 80NSSC21K1555, 80NSSC20K1275

Slides available on https://github.com/ecamporeale/talks

University of Colorado Boulder

CIRES

Agenda

- The belated St.Patrick’s day storm (2023)
  - How actionable was SWPC forecast?
- Validation study of SWPC Geospace
- The myth of the analogy between SWx and meteorology
- ML will become the standard way of SWx forecasting by the end of the decade
  - ACCRUE and Proboost methods
- Possible future scenarios for operational Space Weather

“around St.Patrick’s day” storms

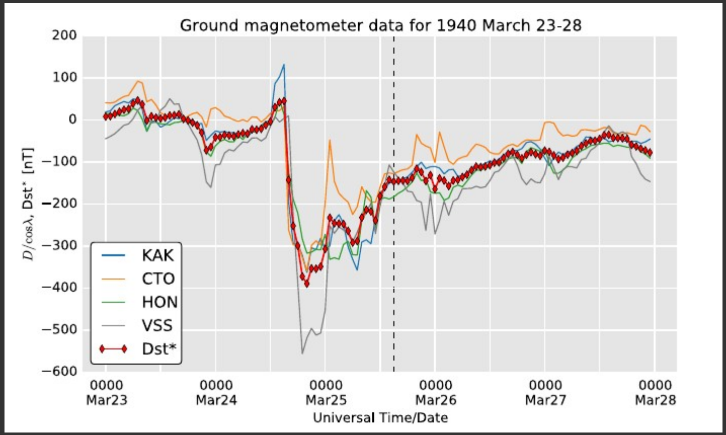
Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **517**, 1709–1723 (2022)
Advance Access publication 2021 December 13

https://doi.org/10.1093/mnras/stab3615

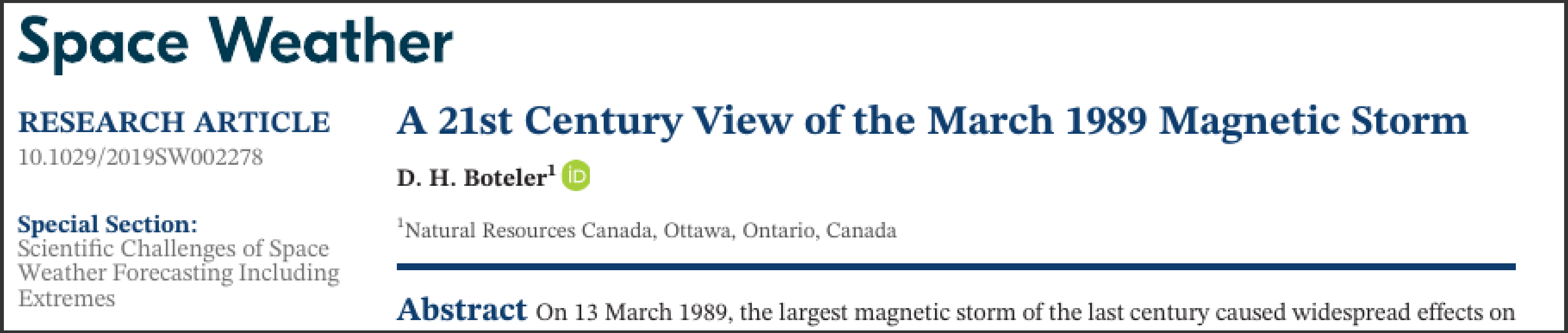
The extreme solar and geomagnetic storms on 1940 March 20–25

Hisashi Hayakawa[1,2,3,4★], Denny M. Oliveira,[5,6] Margaret A. Shea,[7] Don F. Smart,[7] Seán P. Blake,[5,8] Kentaro Hattori[9], Ankush T. Bhaskar,[10] Juan J. Curto,[11] Daniel R. Franco[12] and Yusuke Ebihara[13]



min Dst = -389 nT

“around St.Patrick’s day” storms


Space Weather

RESEARCH ARTICLE
10.1029/2019SW002278

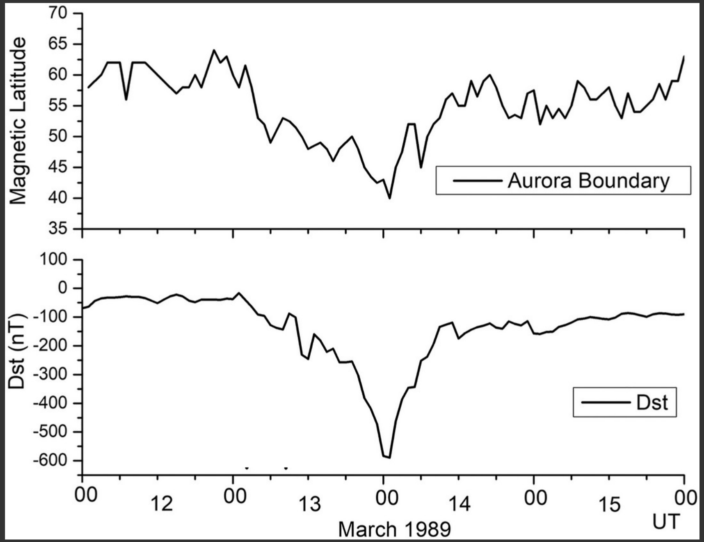
Special Section:
Scientific Challenges of Space Weather Forecasting Including Extremes

A 21st Century View of the March 1989 Magnetic Storm

D. H. Boteler[1]

[1]Natural Resources Canada, Ottawa, Ontario, Canada

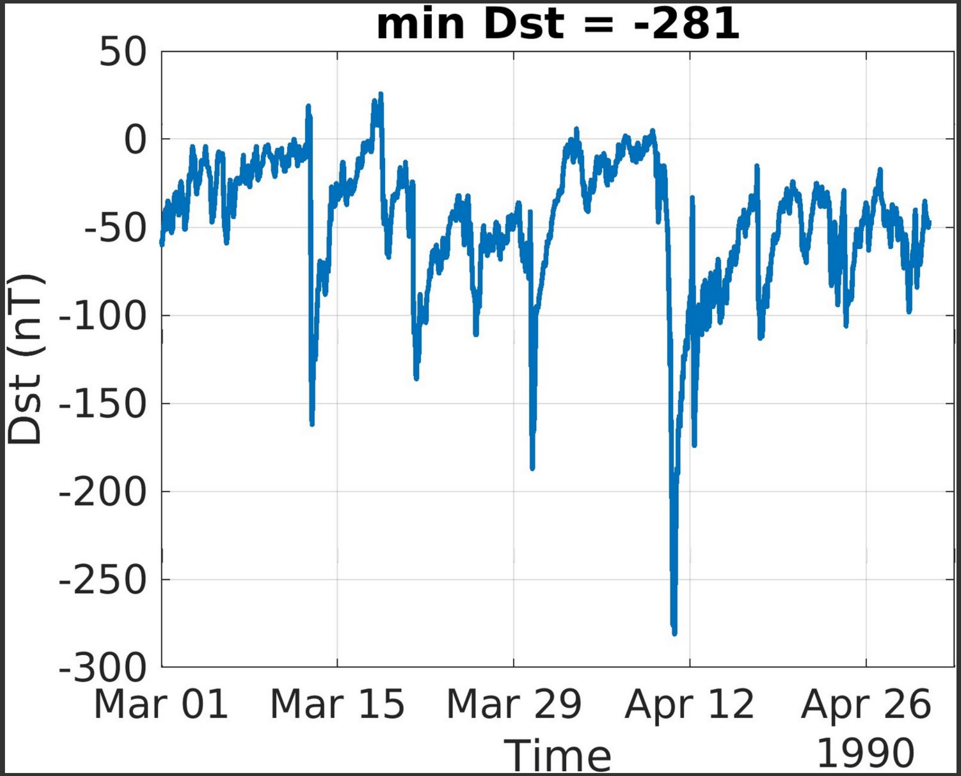
Abstract On 13 March 1989, the largest magnetic storm of the last century caused widespread effects on



min Dst = -589 nT

(strongest storm in recorded history)

“around St.Patrick’s day” storms



min Dst = -281 nT

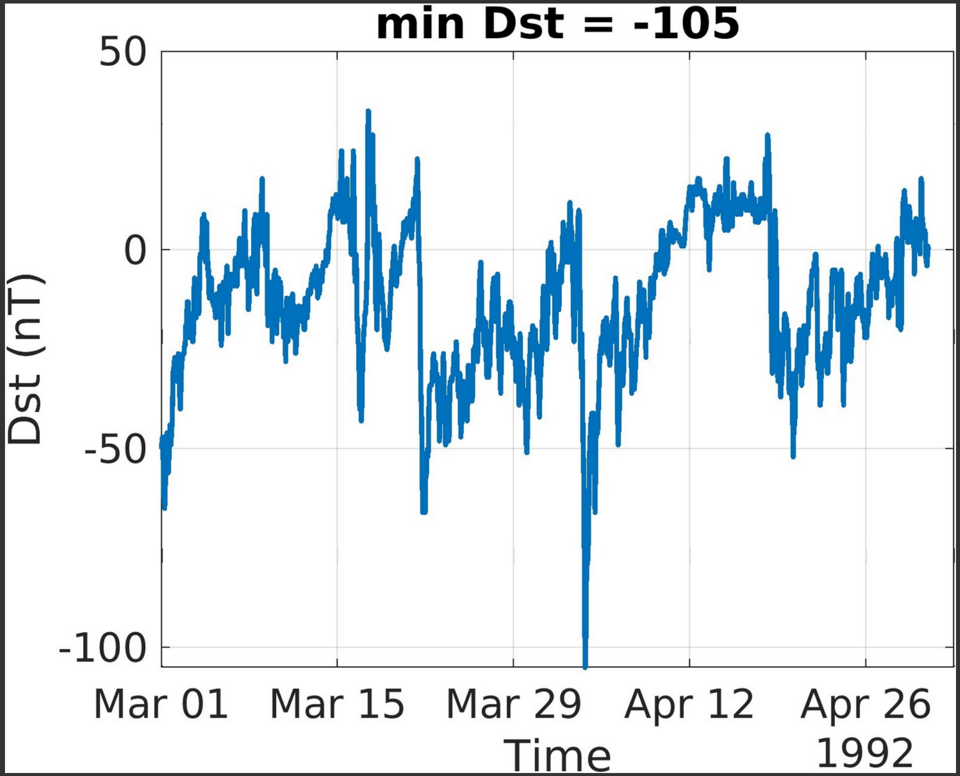
(#12 in recorded history)

“around St.Patrick’s day” storms

min Dst = -298 nT

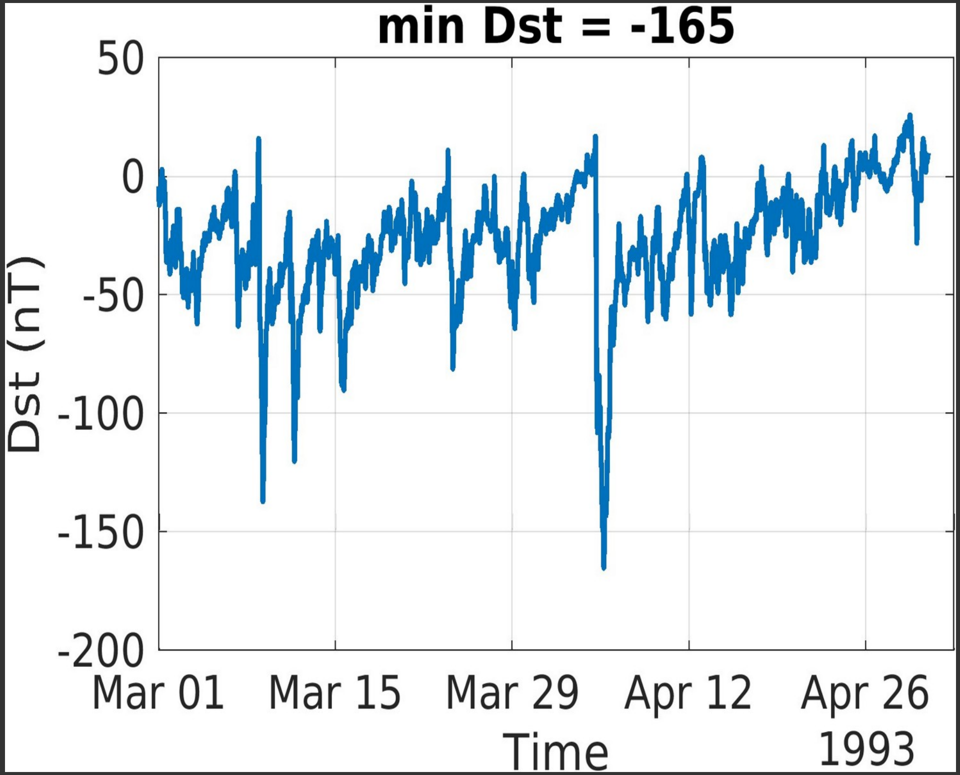
(#8 in recorded history)

"around St.Patrick's day" storms



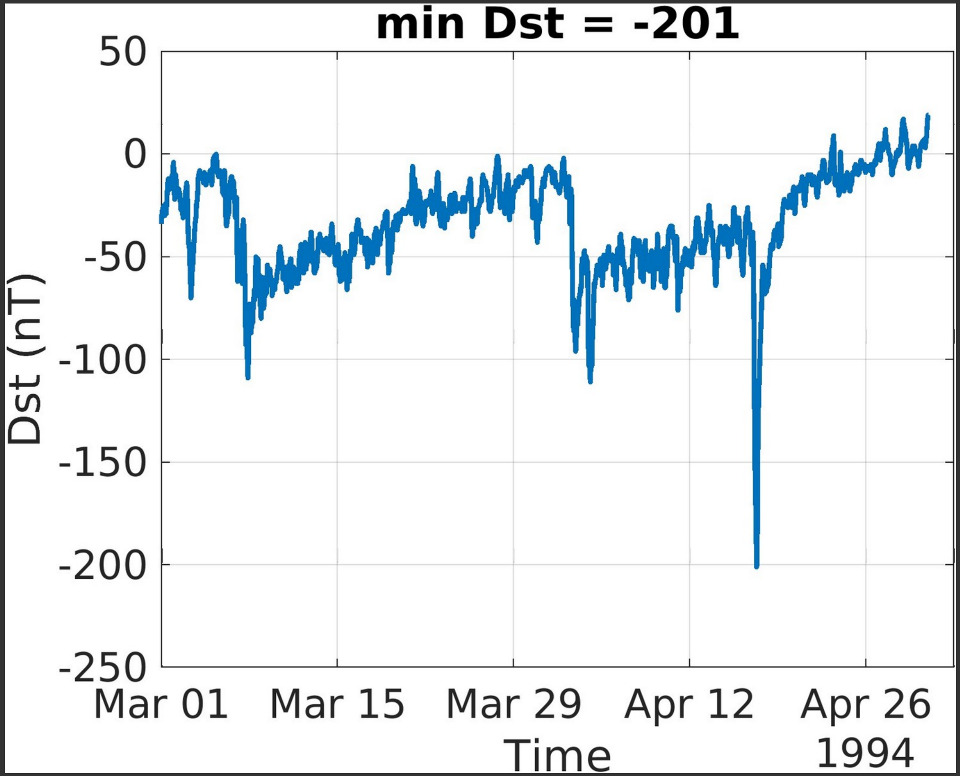
min Dst = -105 nT

"around St.Patrick's day" storms



min Dst = -165 nT

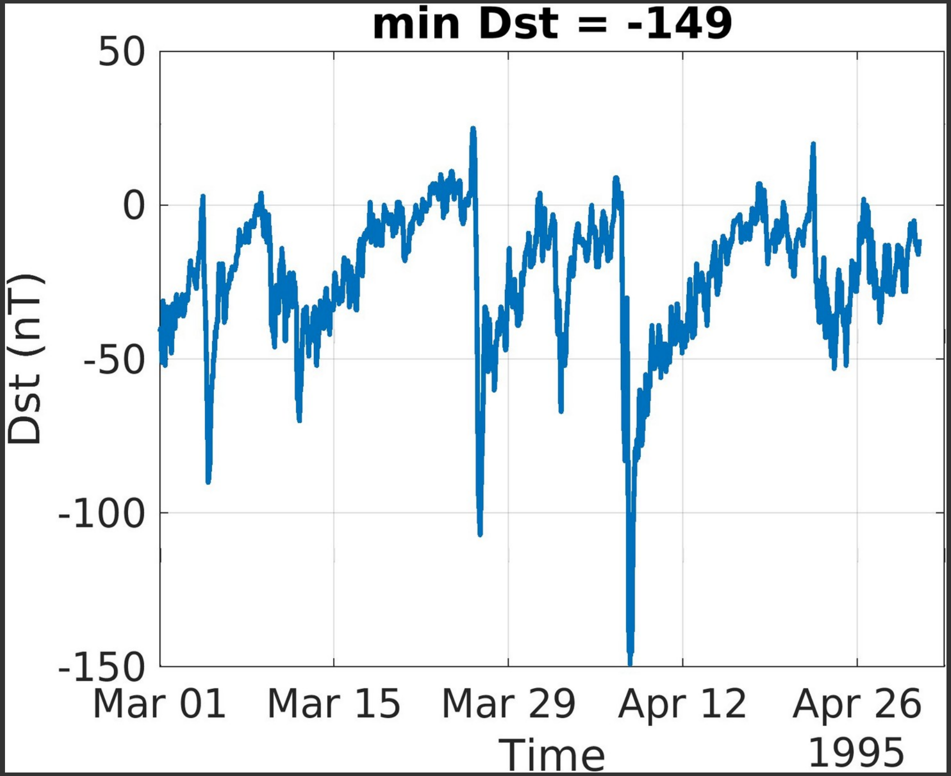
“around St.Patrick’s day” storms



min Dst = -201 nT

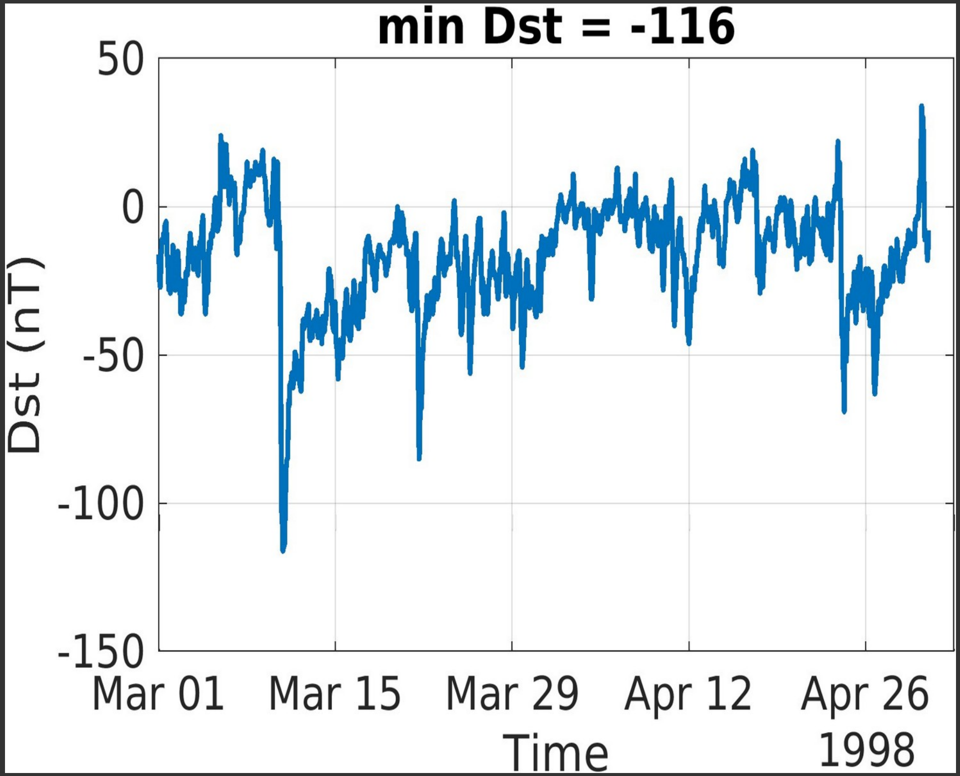
(#27 in recorded history)

“around St.Patrick’s day” storms



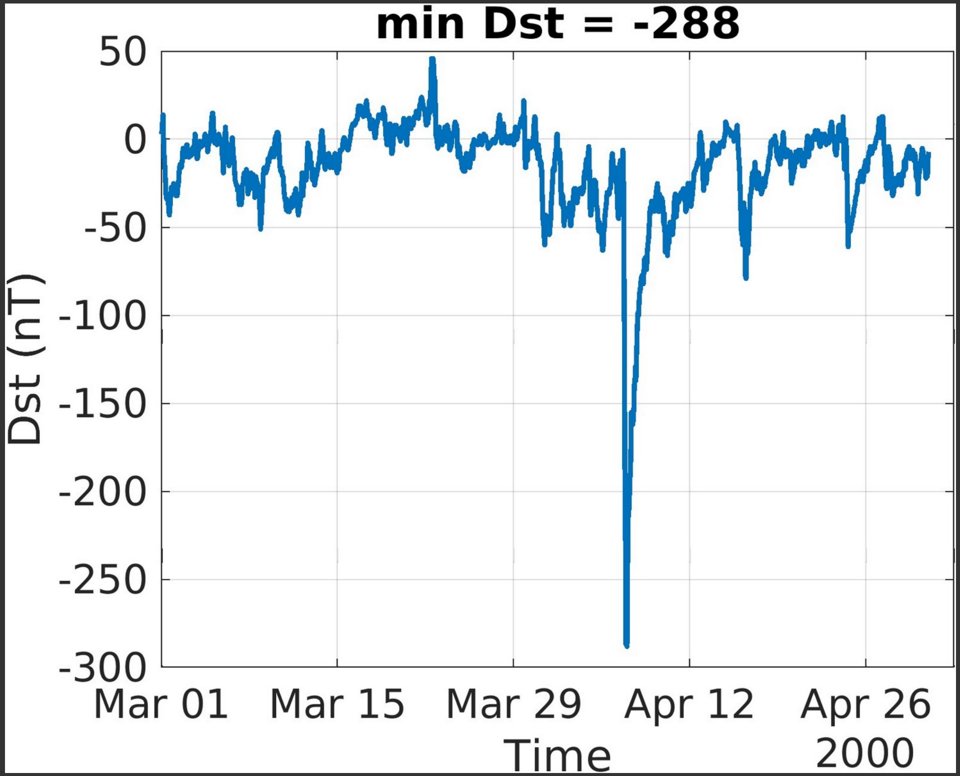
min Dst = -149 nT

“around St.Patrick’s day” storms



min Dst = -116 nT

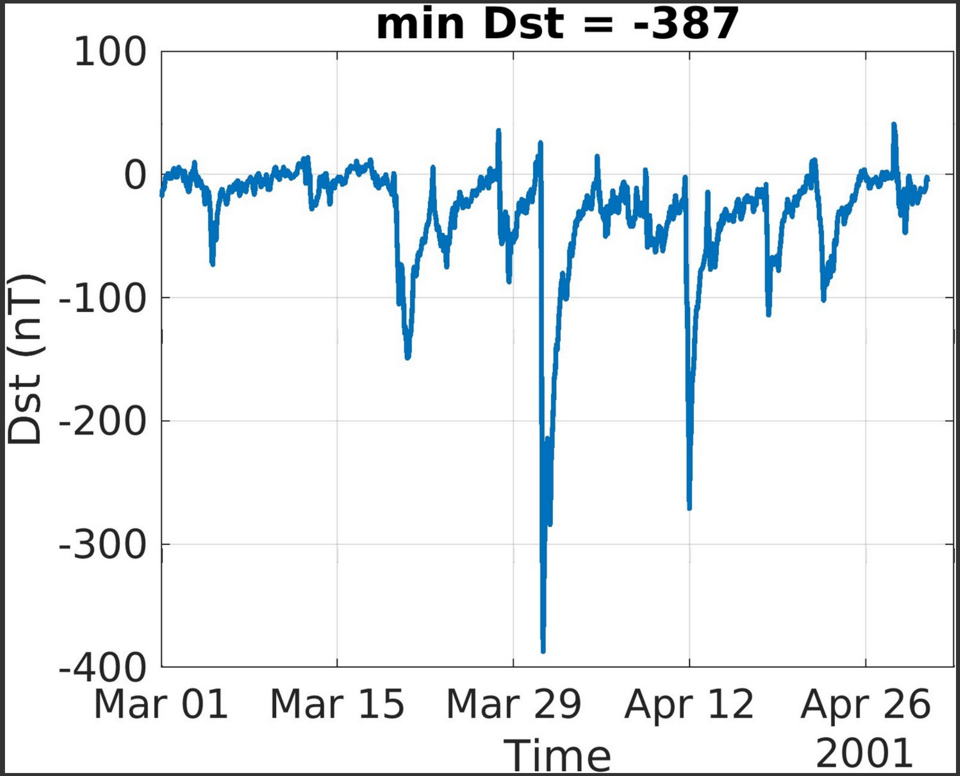
“around St.Patrick’s day” storms



min Dst = -288 nT

(#11 in recorded history)

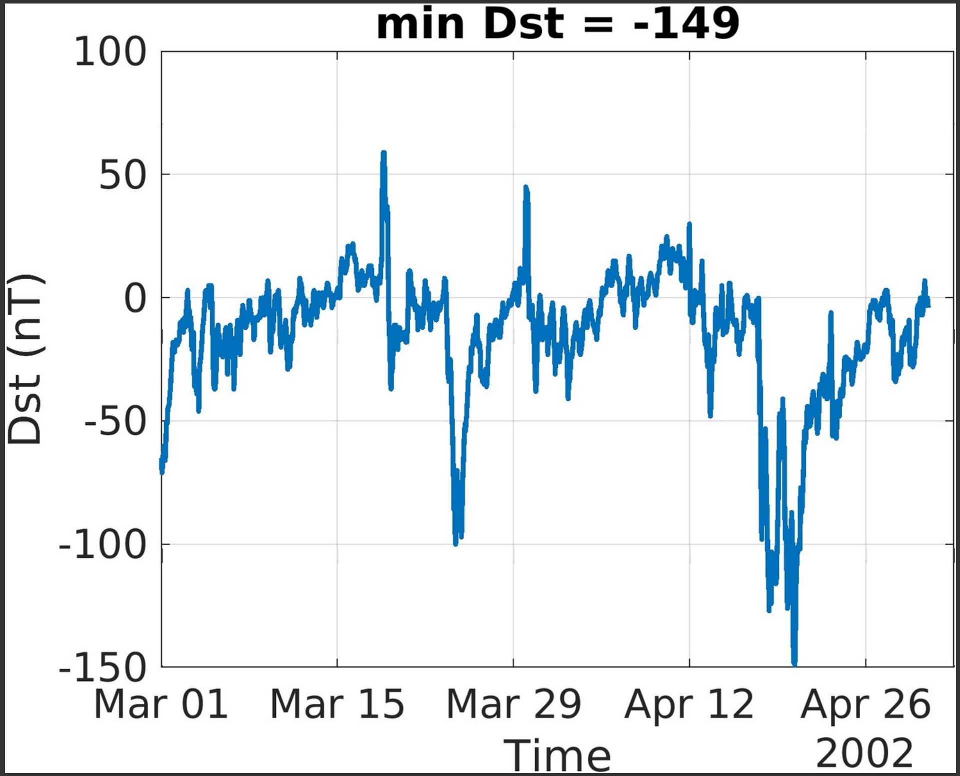
“around St.Patrick’s day” storms



min Dst = -387 nT

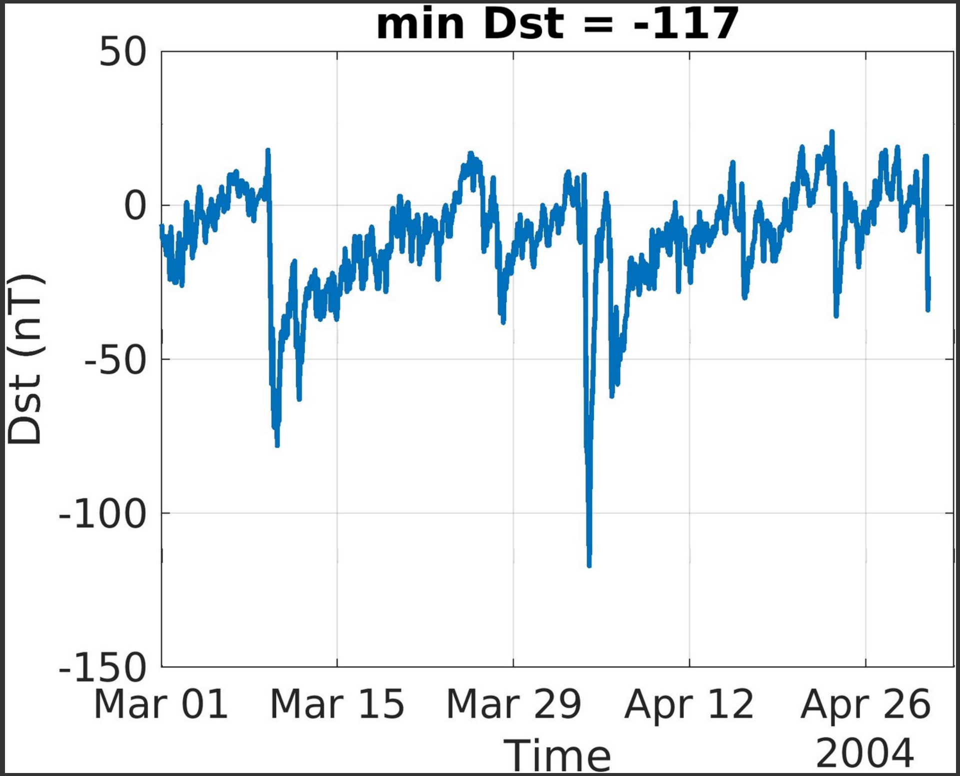
(#3 in recorded history)

“around St.Patrick’s day” storms



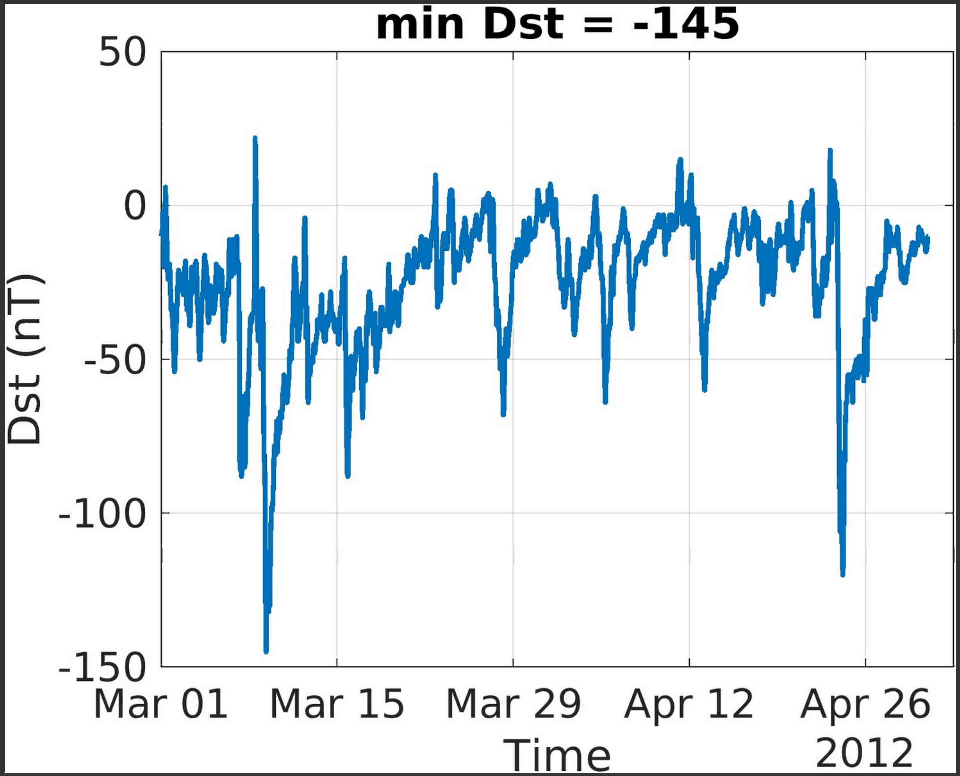
min Dst = -149 nT

“around St.Patrick’s day” storms



min Dst = -117 nT

“around St.Patrick’s day” storms

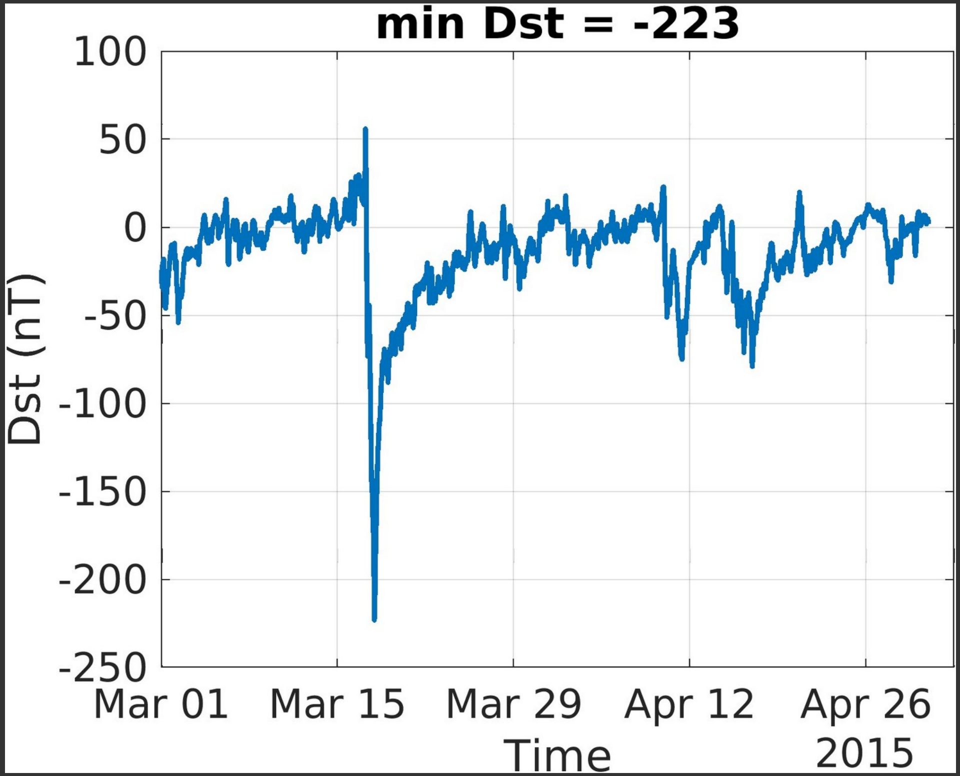


min Dst = -145 nT

“around St.Patrick’s day” storms

min Dst = -132 nT

“around St.Patrick’s day” storms



The St.Patrick’s day storm!

min Dst = -223 nT

(#21 in recorded history)

"around St.Patrick's day" storms

J. Space Weather Space Clim., 6, A9 (2016)
DOI: 10.1051/swsc/2016004

SWSC

RESEARCH ARTICLE — OPEN ACCESS

Overview of the 2015 St. Patrick's day storm and its consequences for RTK and PPP positioning in Norway

Knut Stanley Jacobsen* and Yngvild Linnea Andalsvik

Norwegian Mapping Authority, PO 600 Sentrum, 3507 Hønefoss, Norway

Space Weather Message Code: WARK04
Serial Number: 2484
Issue Time: 2015 Mar 17 1844 UTC

EXTENDED WARNING: Geomagnetic K-index of 4 expected
Extension to Serial Number: 2483
Valid From: 2015 Mar 17 0430 UTC
Now Valid Until: 2015 Mar 18 1000 UTC
Warning Condition: Persistence
www.swpc.noaa.gov/noaa-scales-explanation
Potential Impacts: Area of impact primarily poleward of 65 degrees Geomagnetic Latitude.
Induced Currents - Weak power grid fluctuations can occur.
Aurora - Aurora may be visible at high latitudes such as Canada and Alaska.

Space Weather Message Code: WATA30
Serial Number: 119
Issue Time: 2015 Mar 17 1808 UTC

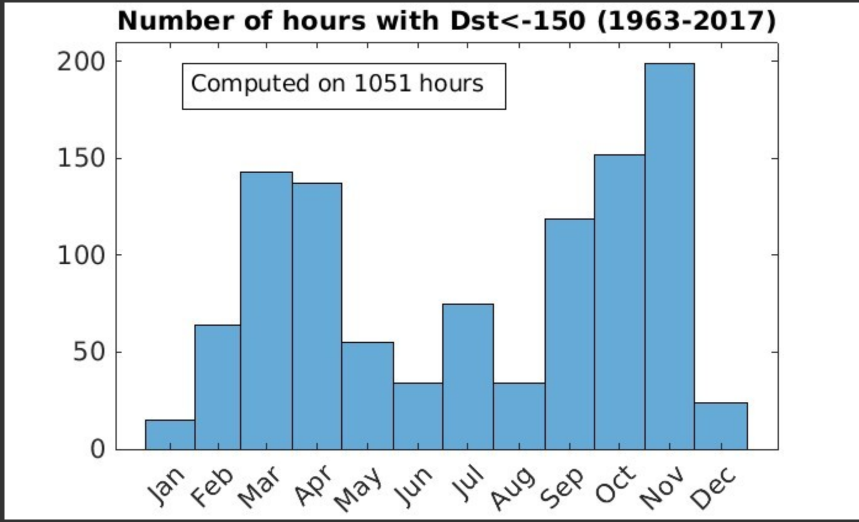
WATCH: Geomagnetic Storm Category G2 Predicted
Highest Storm Level Predicted by Day:
Mar 18: G2 (Moderate) Mar 19: None (Below G1) Mar 20: None (Below G1)
THIS SUPERSEDES ANY/ALL PRIOR WATCHES IN EFFECT
www.swpc.noaa.gov/noaa-scales-explanation
Potential Impacts: Area of impact primarily poleward of 55 degrees Geomagnetic Latitude.
Induced Currents - Power grid fluctuations can occur. High-latitude power systems may experi
Spacecraft - Satellite orientation irregularities may occur; increased drag on low Earth-orbit sat
Radio - HF (high frequency) radio propagation can fade at higher latitudes.
Aurora - Aurora may be seen as low as New York to Wisconsin to Washington state.

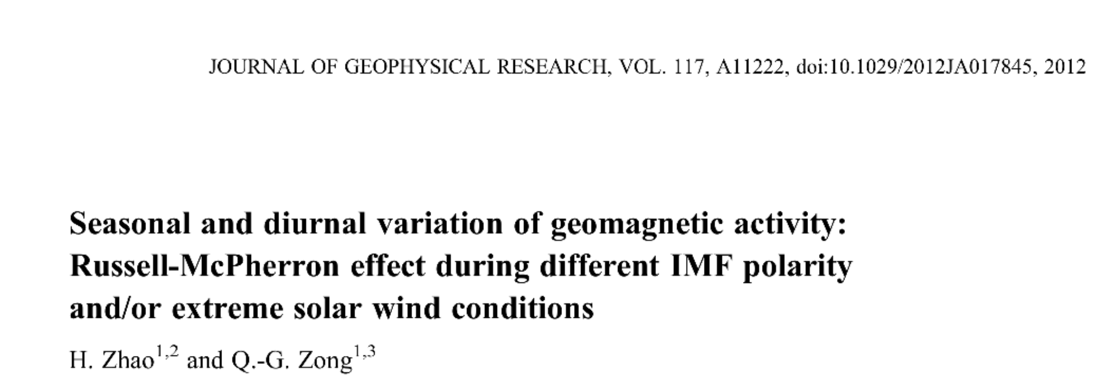
1. Introduction

On 17–18 March 2015, the first storm of solar cycle 24 to reach the G4 level on the NOAA scale (Poppe 2000) occurred. As March 17th is St. Patrick's day, we will refer to the storm as the St. Patrick's day storm. The storm was notable for two reasons: the first that it was at that point the strongest storm of the solar cycle, the second that space weather agencies around the world failed to predict it. Geomagnetic storm warnings had been issued, but only for a minor storm, which would not be a concern to most users. As an example, this is an extract of the weekly report by the space weather prediction centre of NOAA.[1]

> Space weather outlook 16 March–11 April, 2015
>
> Solar activity is expected to continue at moderate levels until 19 March when Region 2297 transits off the visible disk. · · ·⟨snip⟩· · · Geomagnetic field activity is expected to be at unsettled to active levels with minor storm periods likely on 18 March due to a combination of CH HSS effects as well as the arrival of the 15 March CME by mid to late on 17 March.

“around St.Patrick’s day” storms



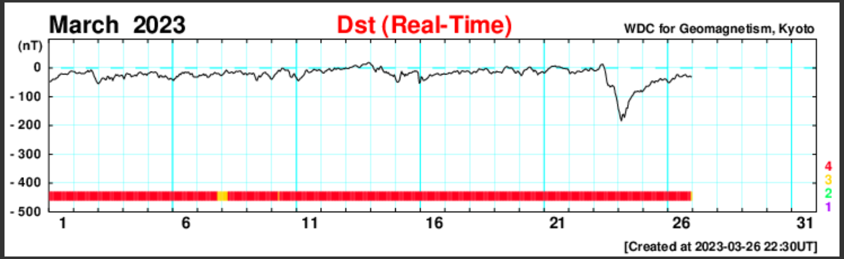

JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 117, A11222, doi:10.1029/2012JA017845, 2012

Seasonal and diurnal variation of geomagnetic activity: Russell-McPherron effect during different IMF polarity and/or extreme solar wind conditions

H. Zhao[1,2] and Q.-G. Zong[1,3]

Out of the 50 strongest storms in recorded history:
- 11 happened in March/April
- 15 happened in September/October

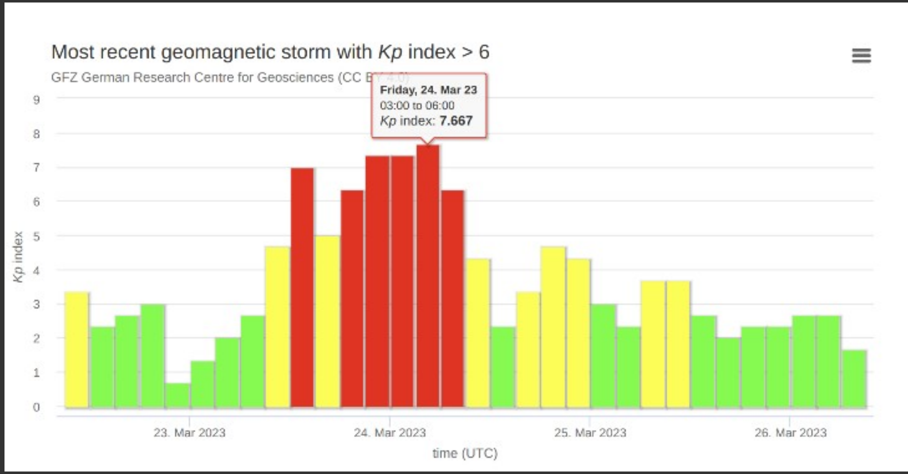
The belated St.Patrick's day storm (2023)
(National Chocolate Covered Raisins Day)



Reached (quick-look) **Dst = -184**
on 03/24 at 3:00UTC

3-DAY GEOMAGNETIC FORECAST

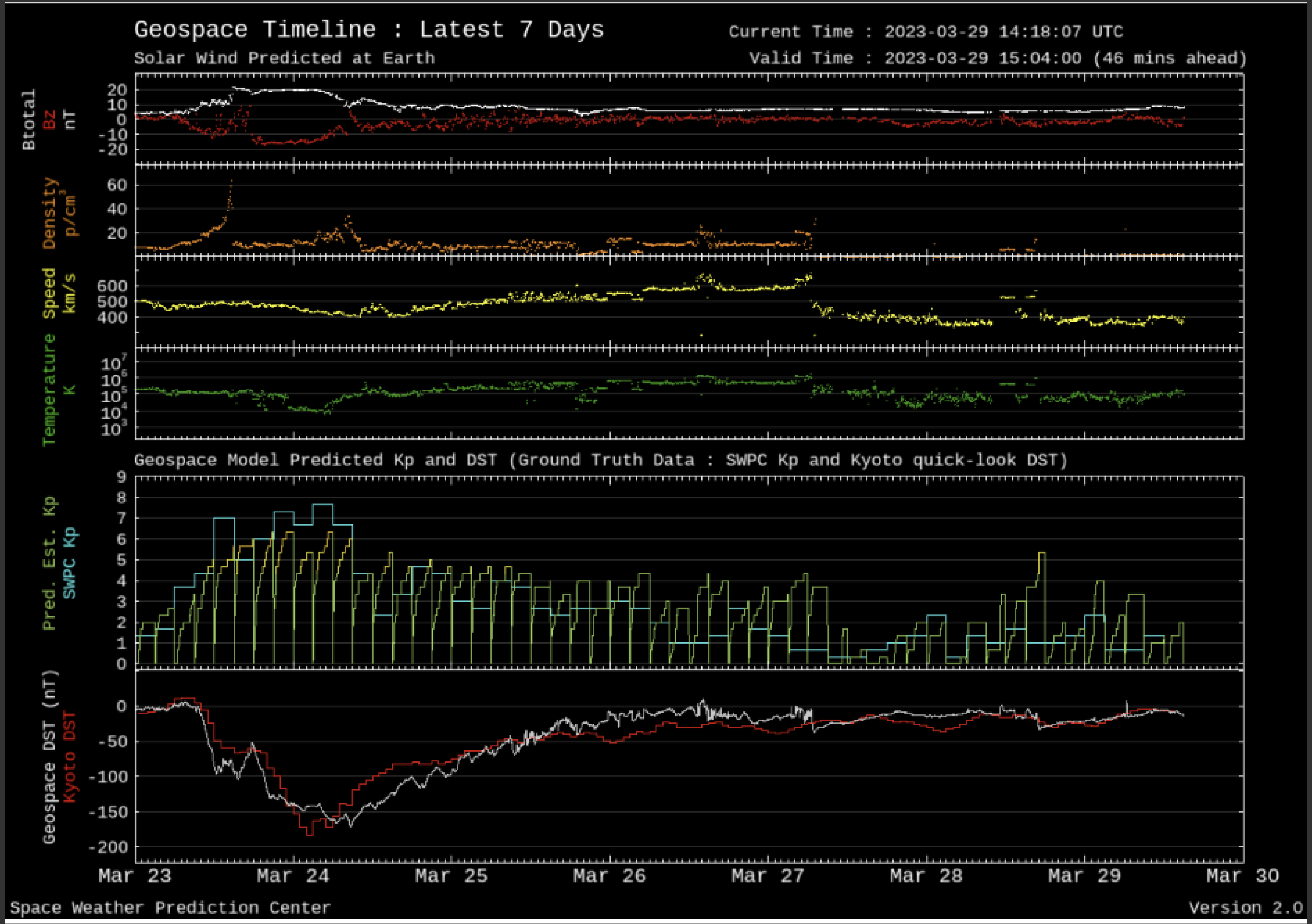
```
:Product: Geomagnetic Forecast
:Issued: 2023 Mar 23 2205 UTC
# Prepared by the U.S. Dept. of Commerce, NOAA, Space
#
NOAA Ap Index Forecast
Observed Ap 22 Mar 017
Estimated Ap 23 Mar 044
Predicted Ap 24 Mar-26 Mar 035-030-020

NOAA Geomagnetic Activity Probabilities 24 Mar-26 Mar
Active                  01/20/25
Minor storm             40/35/30
Moderate storm          45/20/15
Strong-Extreme storm    15/10/05

NOAA Kp index forecast 24 Mar - 26 Mar
             Mar 24    Mar 25    Mar 26
00-03UT       4.67      3.67      2.67
03-06UT       5.67      4.67      3.33
06-09UT       5.67      4.33      4.67
09-12UT       4.33      4.00      3.67
12-15UT       4.00      4.00      3.67
15-18UT       3.67      4.00      3.33
18-21UT       2.67      4.33      3.00
21-00UT       2.67      4.33      3.00
```



max Kp = 7.66 (G4)

Ranked #33 among all storms!!

The belated St.Patrick's day storm (2023)

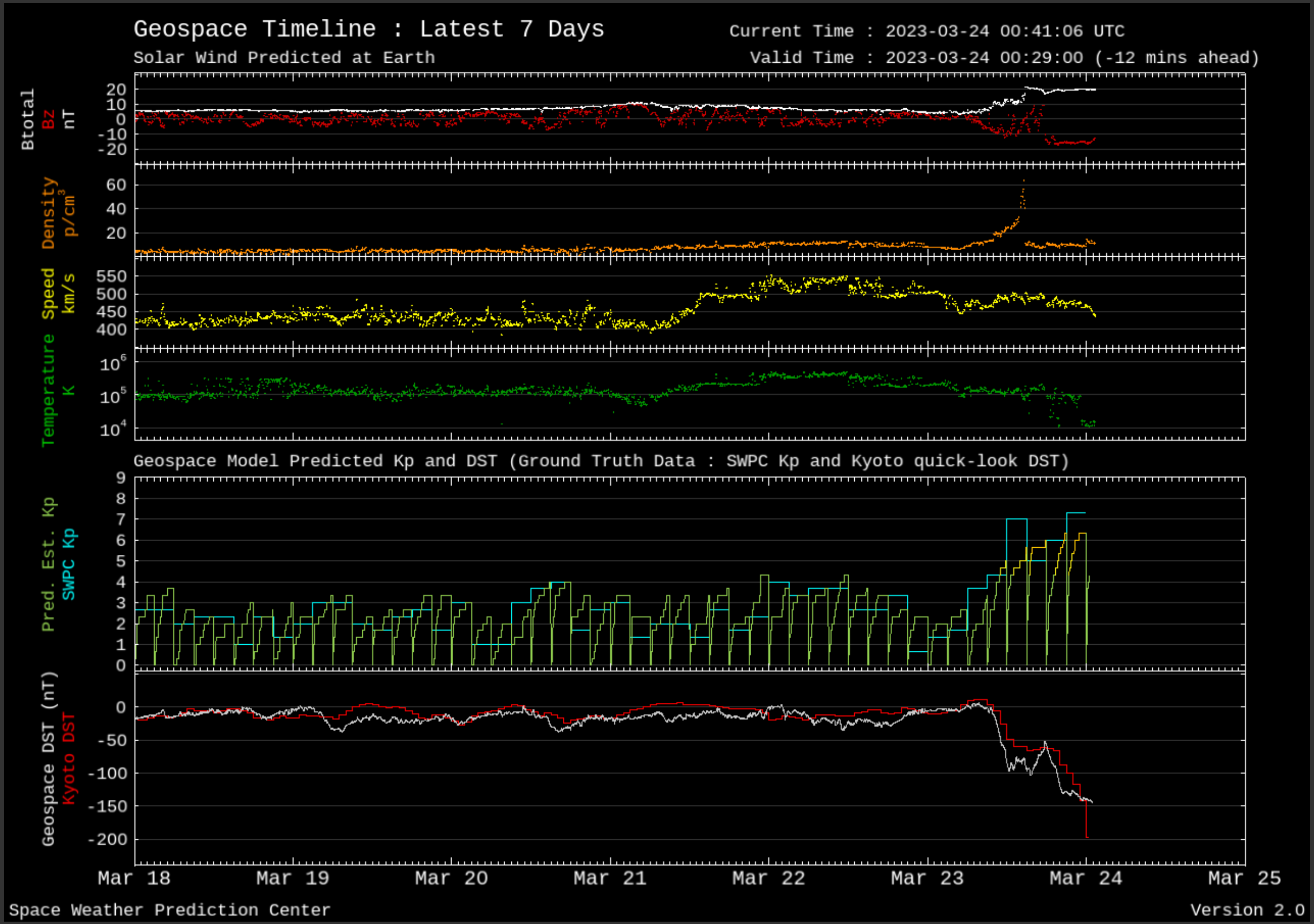


Note for future testbed:

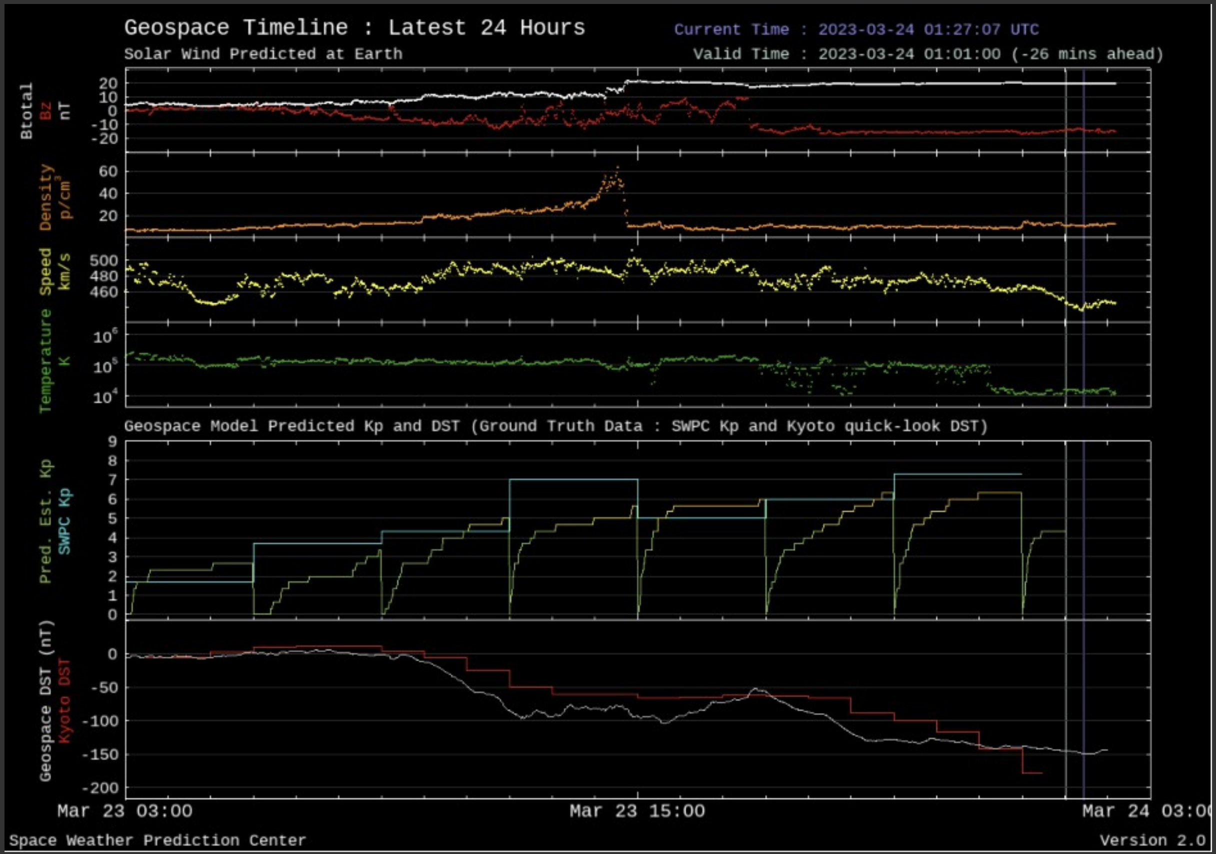
We need to be able to analyze what was predicted at a given time.

That info is not readily available from SWPC

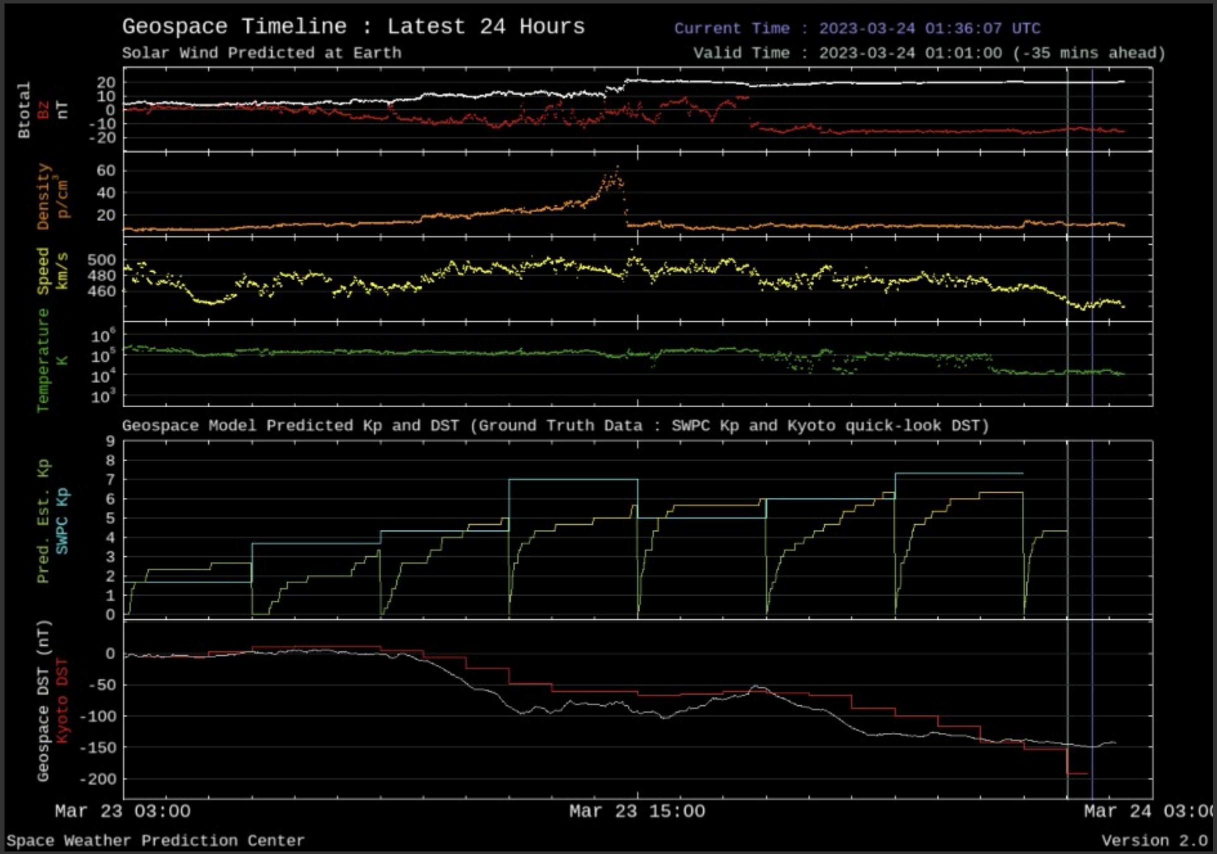
The belated St.Patrick's day storm (2023)



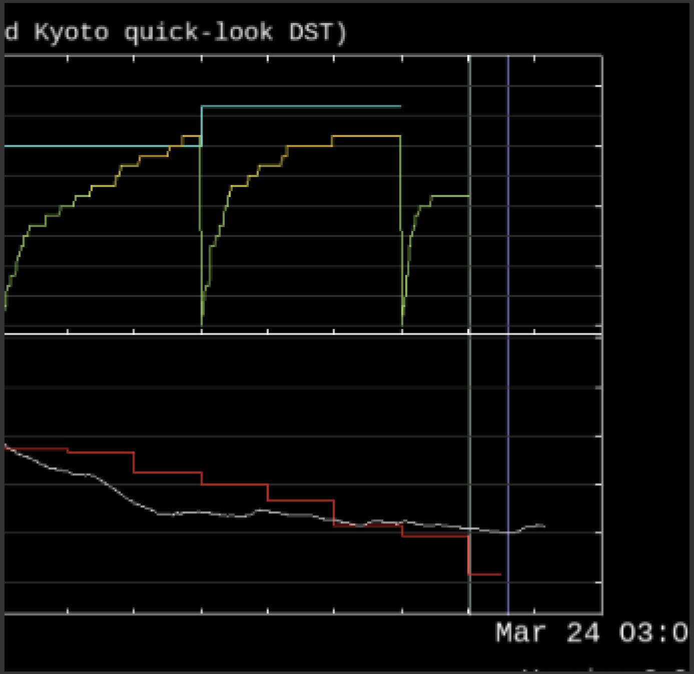
The belated St.Patrick's day storm (2023)

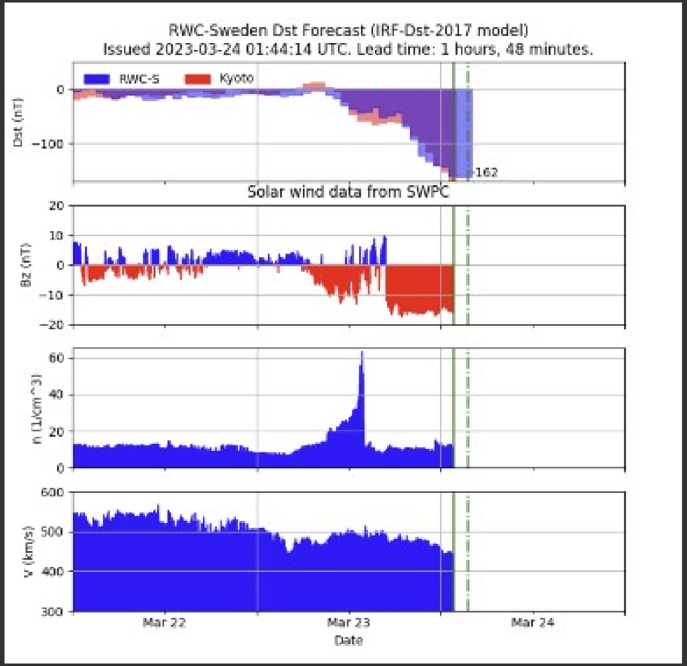


The belated St.Patrick's day storm (2023)



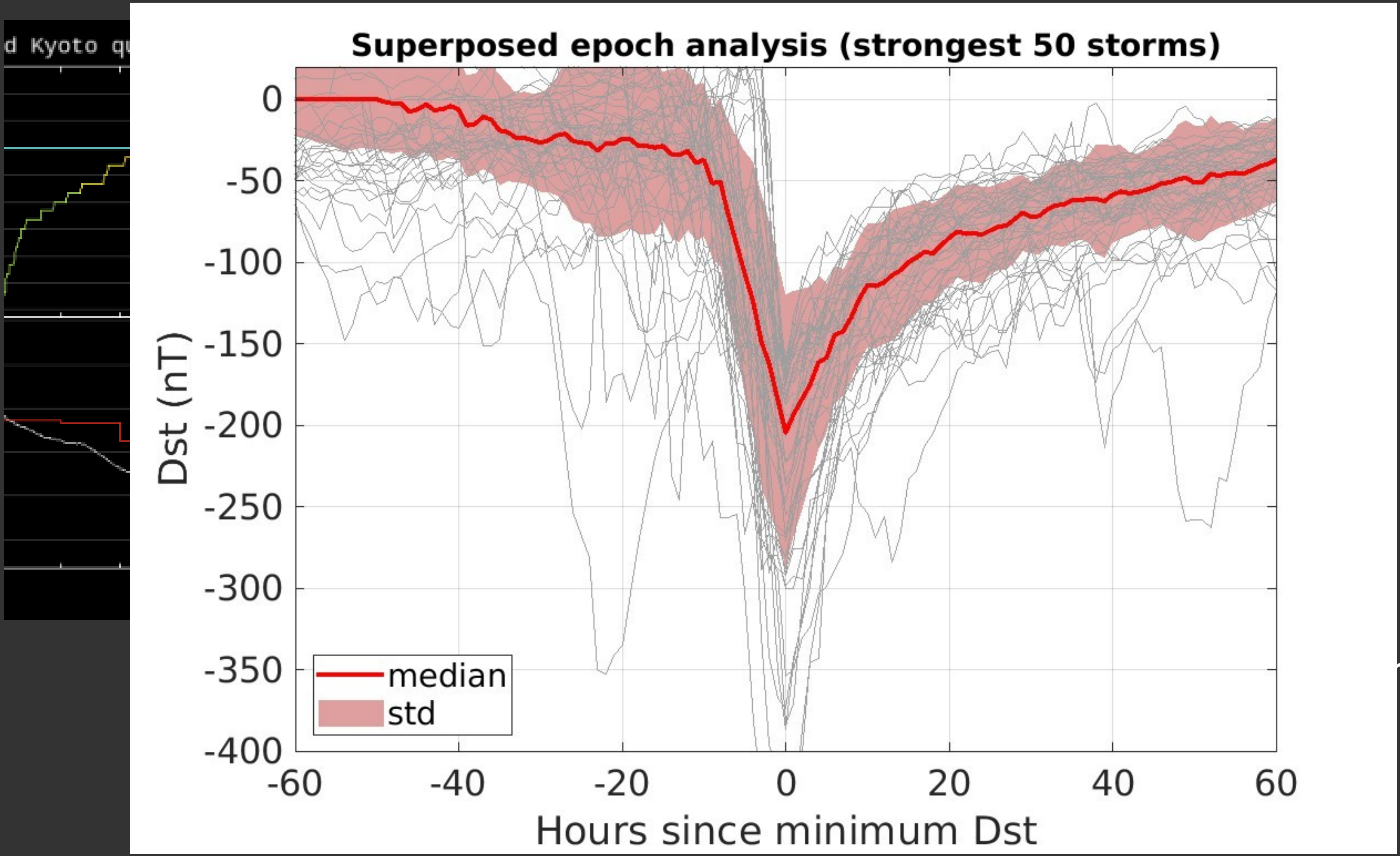
The belated St.Patrick's day storm (2023)





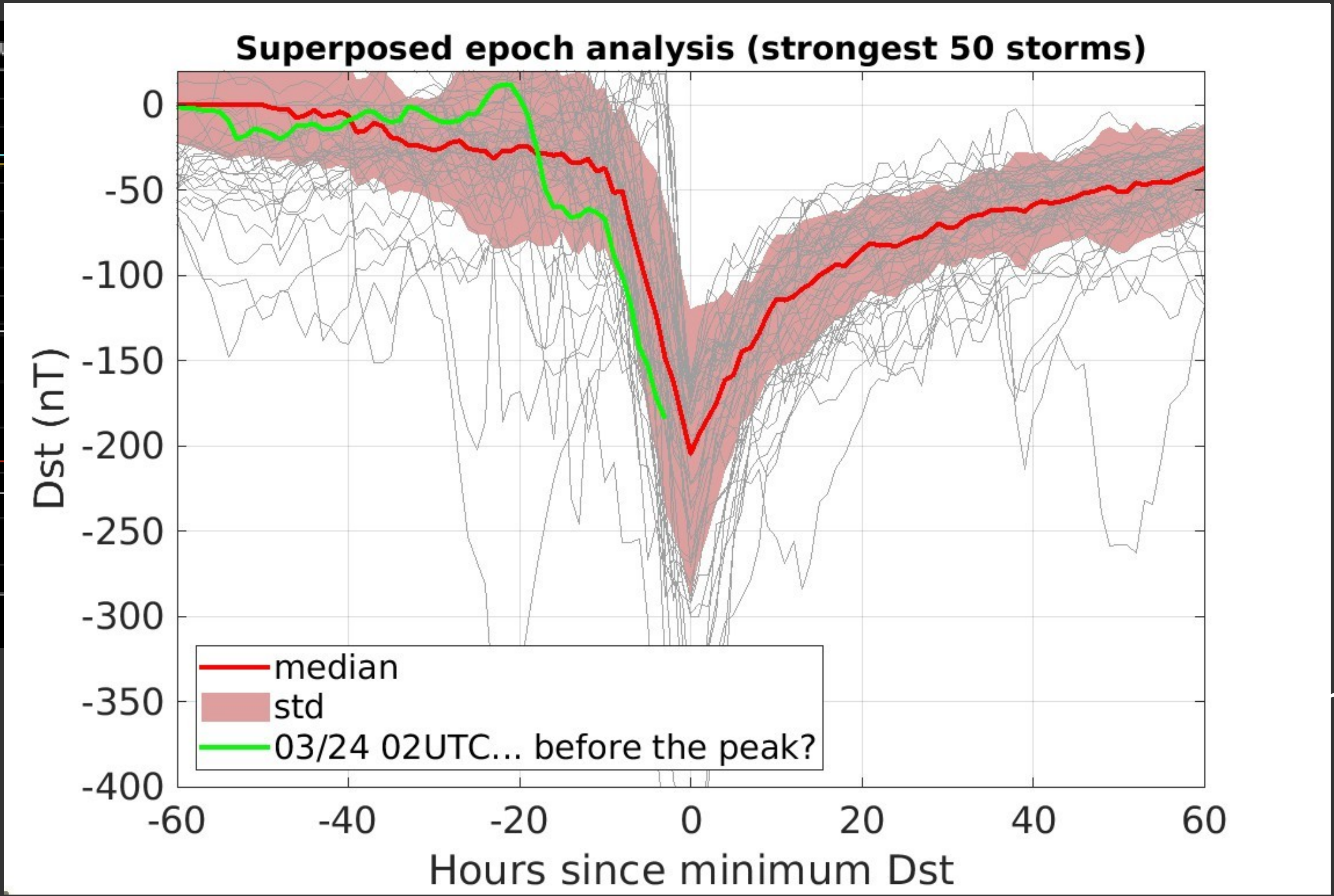
At 01:44UTC ESA predicts Dst = -162 for 03:32UTC.
Observed Dst = -163 at 03:00UTC
Dst = -161 at 04:00UTC

The belated St.Patrick’s day storm (2023)



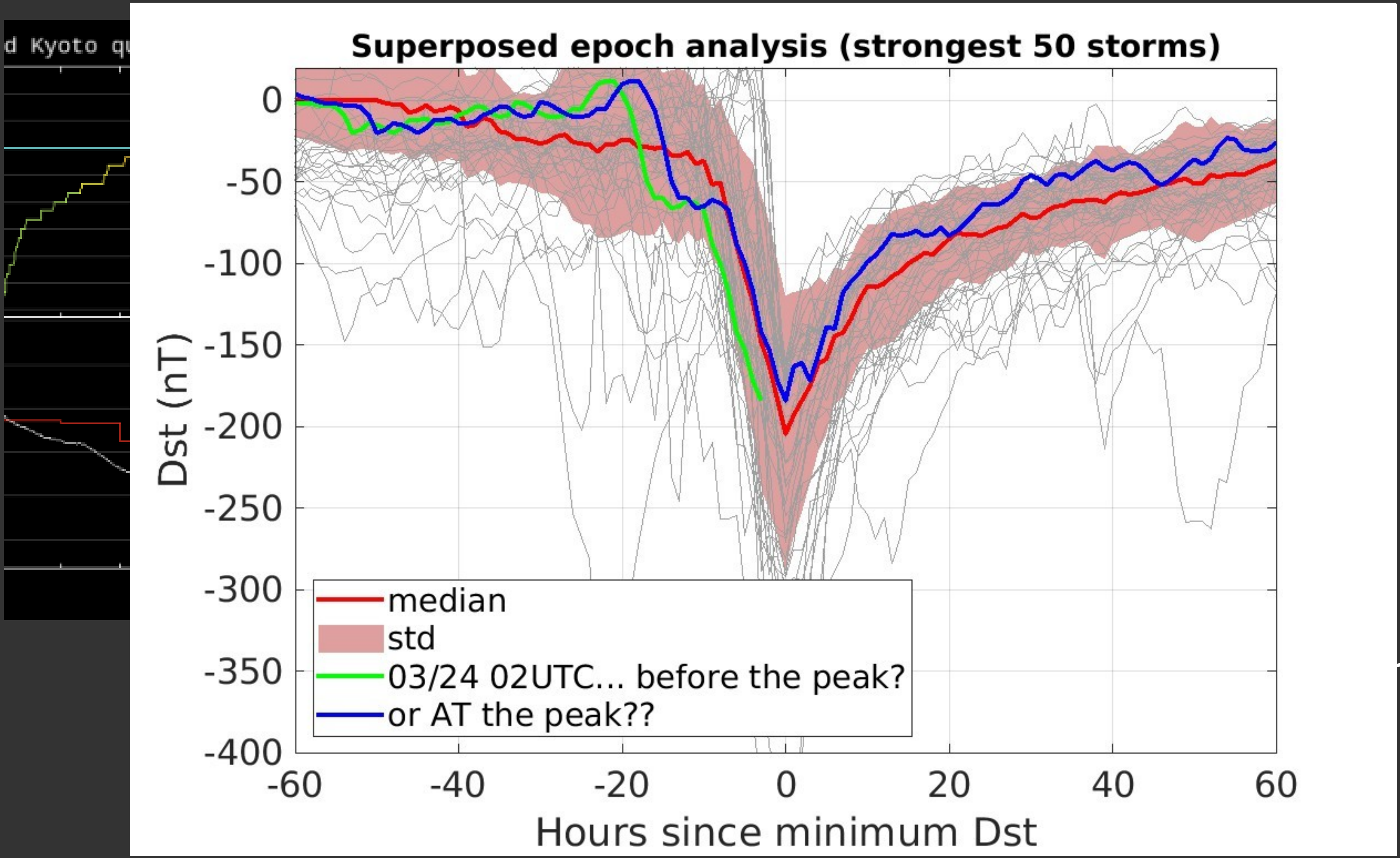
03:32UTC.

The belated St.Patrick's day storm (2023)



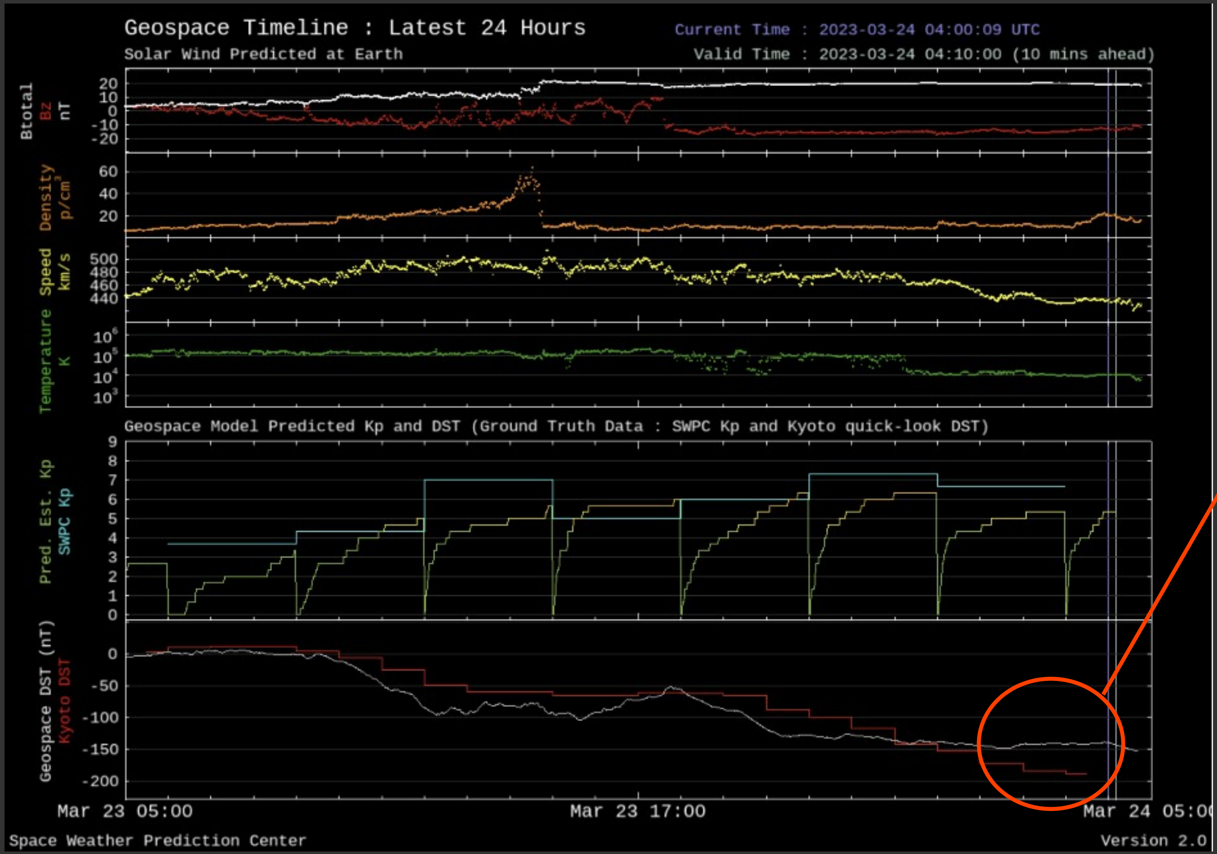
03:32UTC.

The belated St.Patrick’s day storm (2023)



03:32UTC.

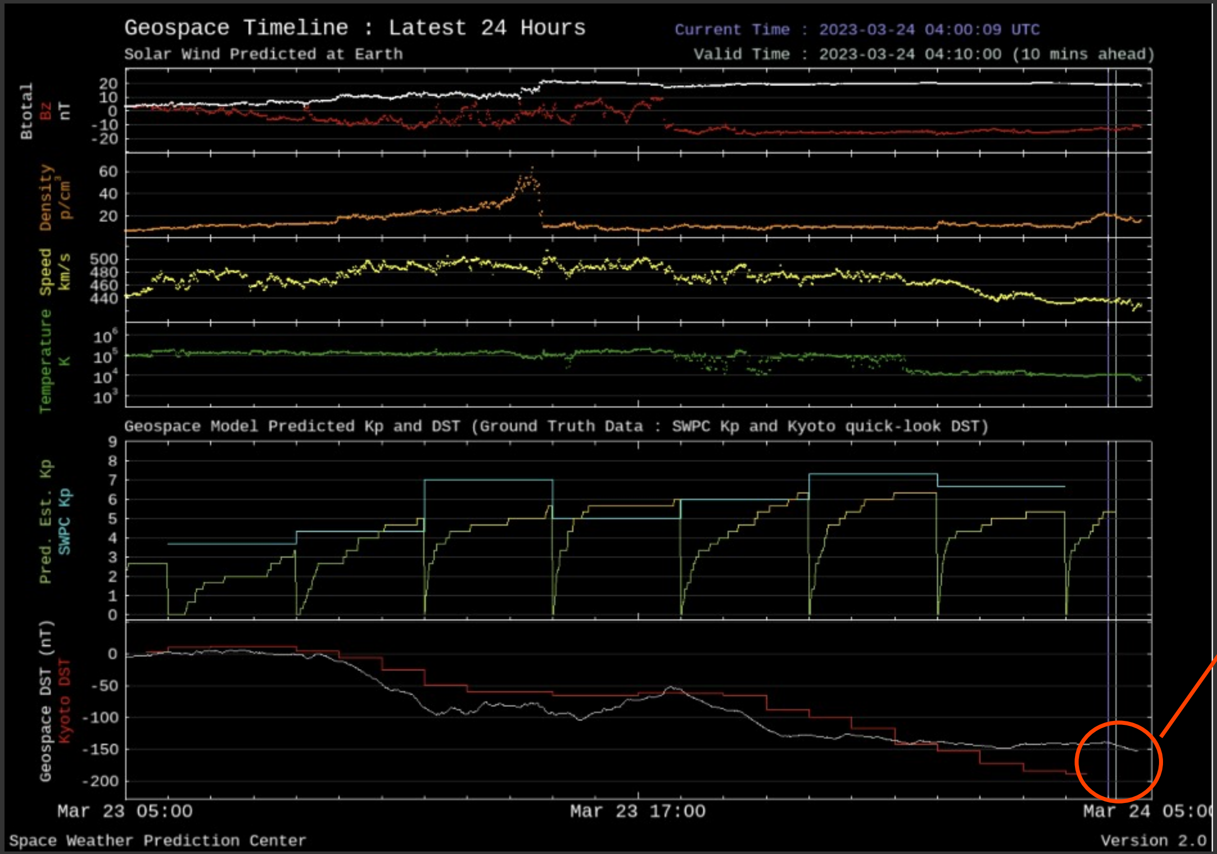
The belated St.Patrick's day storm (2023)



What makes SWPC prediction **non actionable**?

- Geospace predicts a slow recovery of the storm while observed Dst continues to decrease

The belated St.Patrick’s day storm (2023)



What makes SWPC prediction **non actionable**?

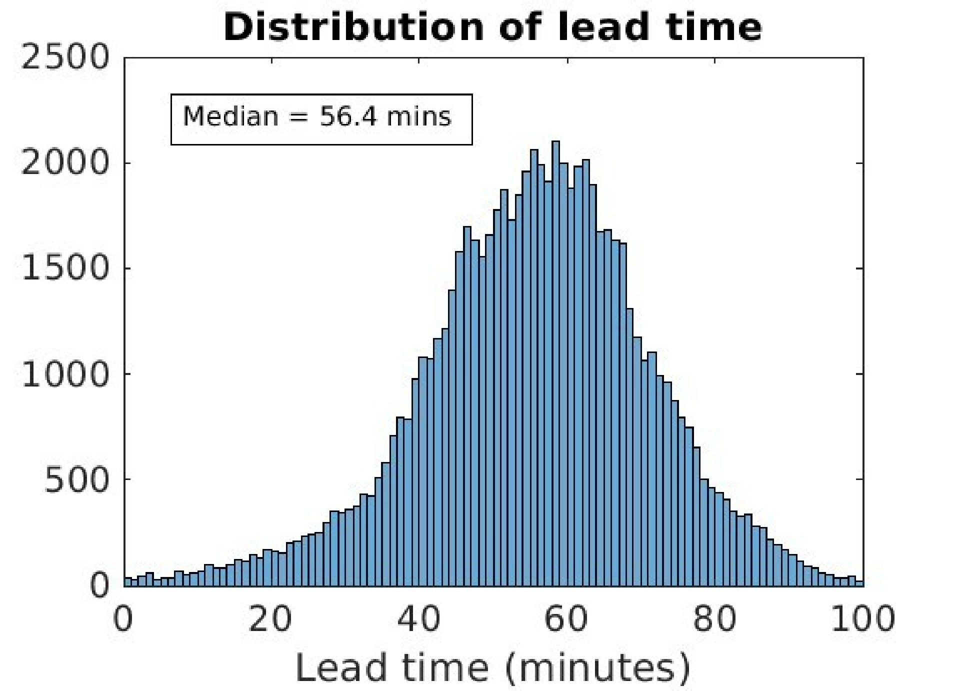
- Geospace predicts a slow recovery of the storm while observed Dst continues to decrease
- At the time of observed min Dst, Geospace predicts a decrease of Dst

  Is it just closing the gap, or should it be interpreted as the storm not being at its peak?

The belated St.Patrick's day storm (2023)

What makes SWPC prediction **non actionable**?

- Very short lead-time (a problem with the CONOPS more than the model itself)

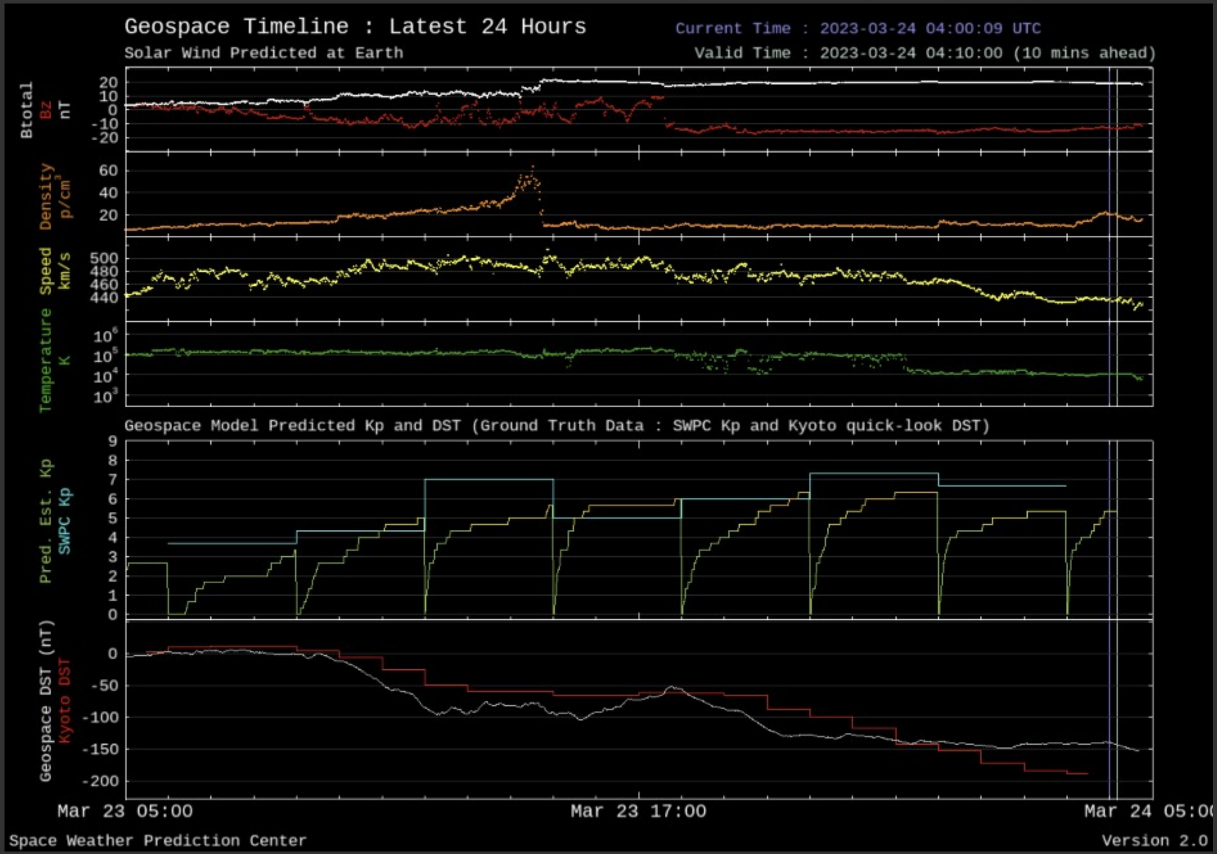


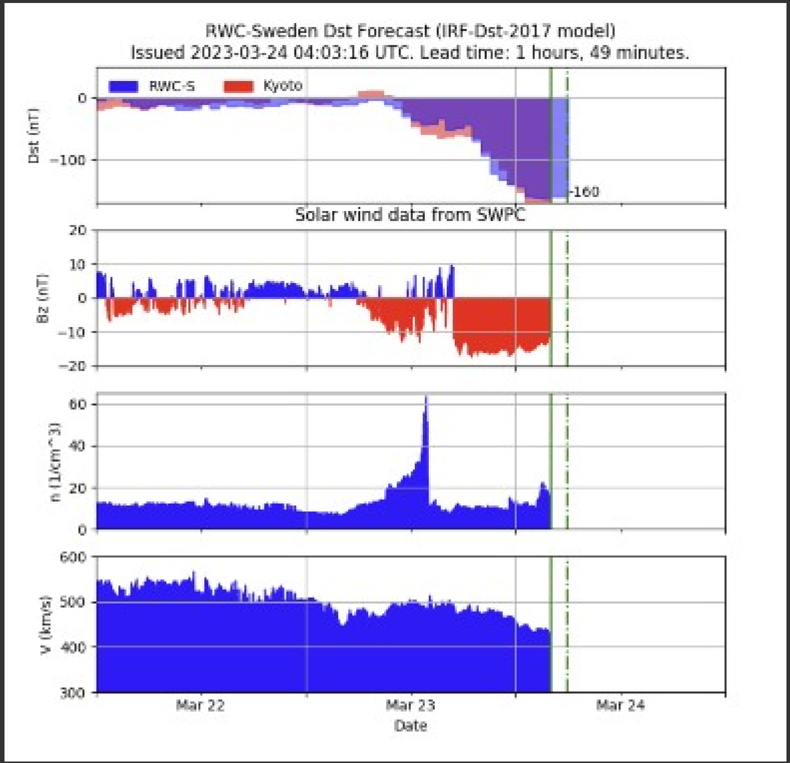
The belated St.Patrick’s day storm (2023)

What makes SWPC prediction **non actionable**?

- Very short lead-time (a problem with the CONOPS more than the model itself)
- No quantification of uncertainty (error bars)

The belated St.Patrick's day storm (2023)





At 04:03UTC ESA predicts Dst = -160
for 05:52UTC
Observed Dst = -156 at 06:00UTC

Statistical validation of SWPC Geospace

Space Weather

RESEARCH ARTICLE
10.1029/2022SW003049

Key Points:

- Space Weather Modeling Framework (SWMF) simulations were carried out for 122 geomagnetic storms from 2010 to 2019
- SWMF simulations of ground magnetic disturbances provide predictive results with a median Heidke Skill Score (HSS) of 0.45 for magnetometers in all regions
- Simulation performance for high-latitude magnetic perturbations have lower HSS with a median of 0.32

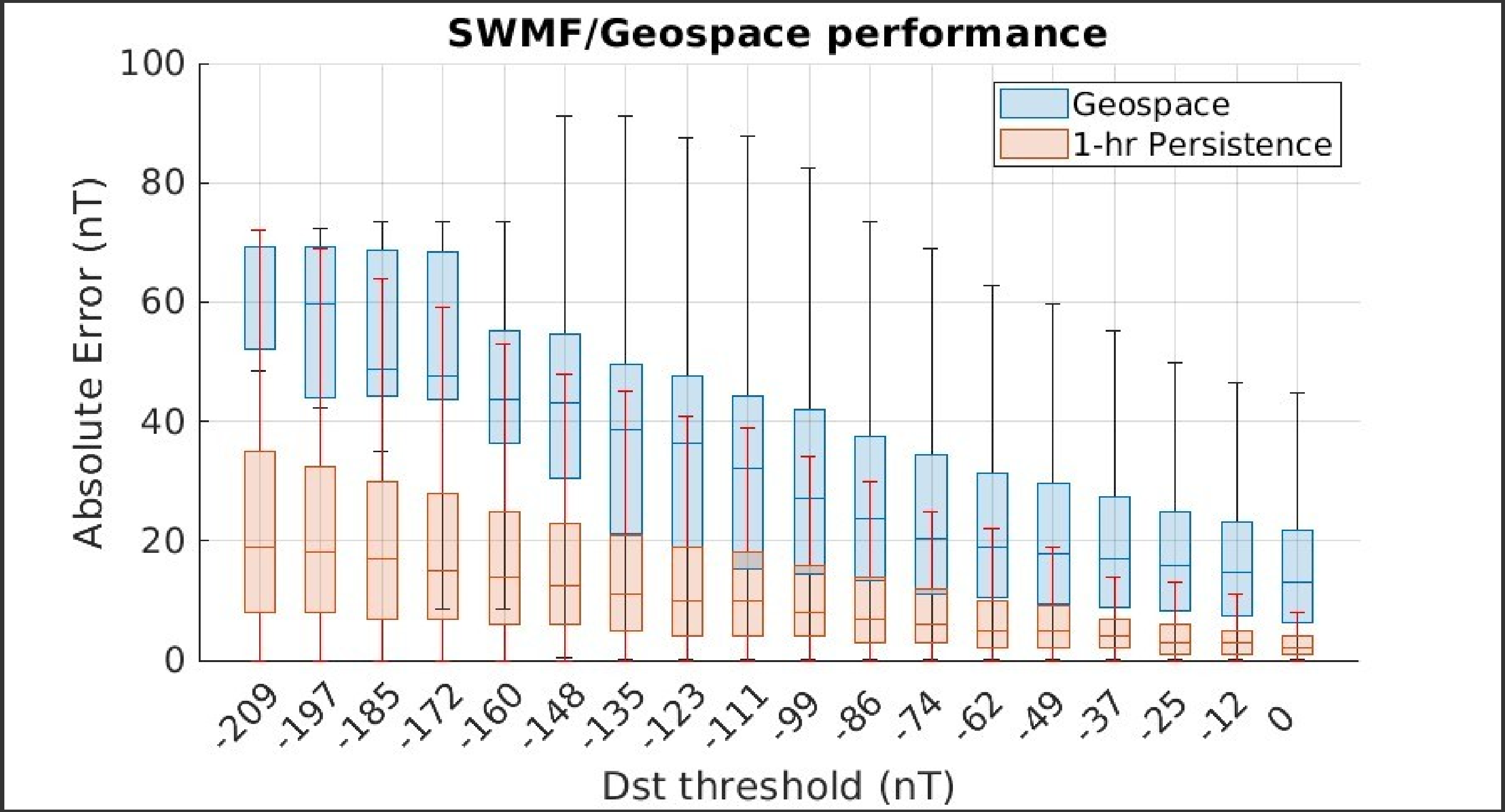
A Large Simulation Set of Geomagnetic Storms—Can Simulations Predict Ground Magnetometer Station Observations of Magnetic Field Perturbations?

Q. Al Shidi[1], T. Pulkkinen[1], G. Toth[1], A. Brenner[1], S. Zou[1], and J. Gjerloev[2]

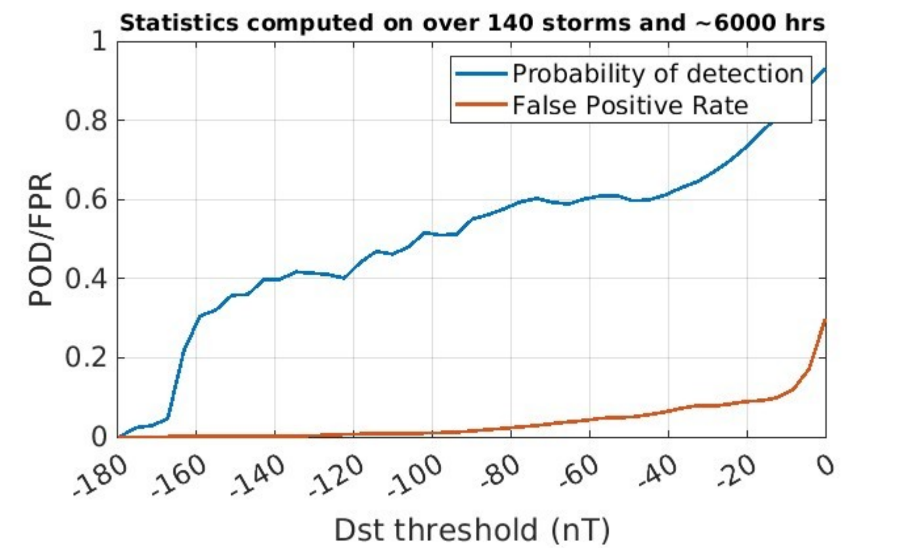
[1]Climate and Space Sciences and Engineering, University of Michigan, Ann Arbor, MI, USA, [2]Applied Physics Laboratory, John Hopkins University, Baltimore, MD, USA

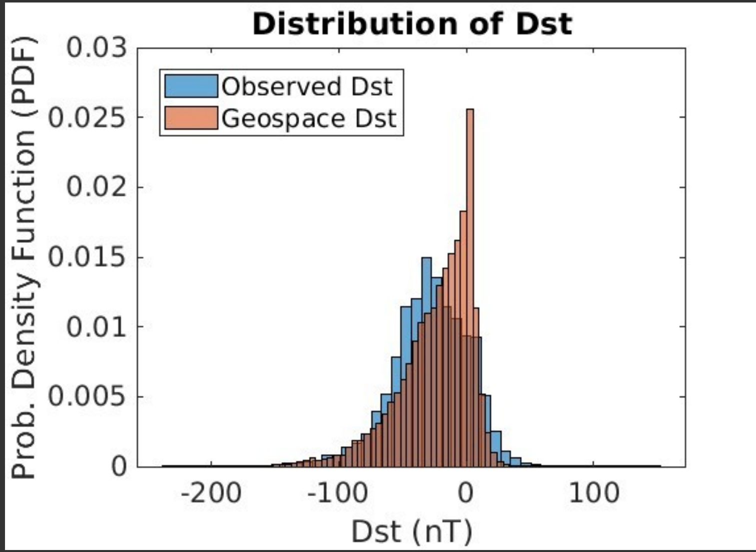
Abstract We use the Space Weather Modeling Framework Geospace configuration to simulate a total of 122 storms from the period 2010–2019. With the focus on the storm main phase, each storm period was run for 54 hr starting from 6 hr prior to the start of the Dst depression. The simulation output of ground magnetic

Statistical validation of SWPC Geospace

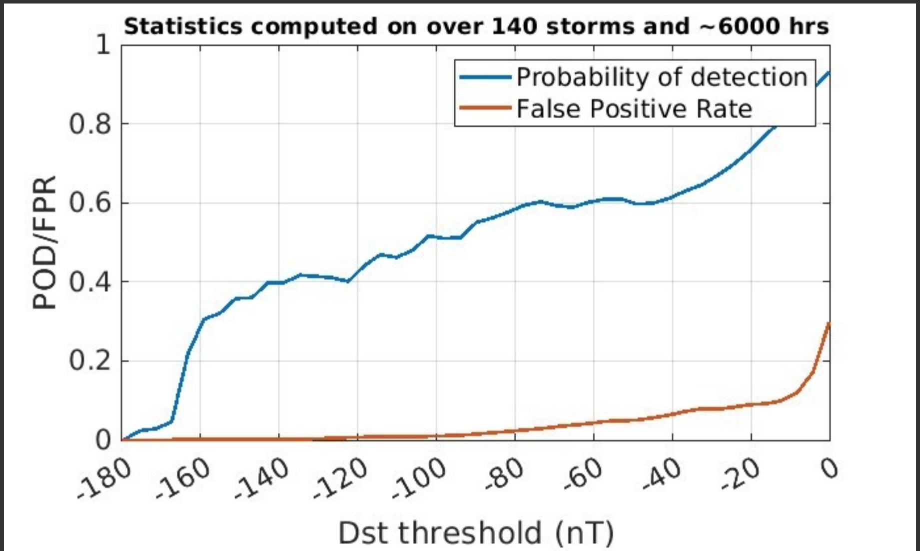


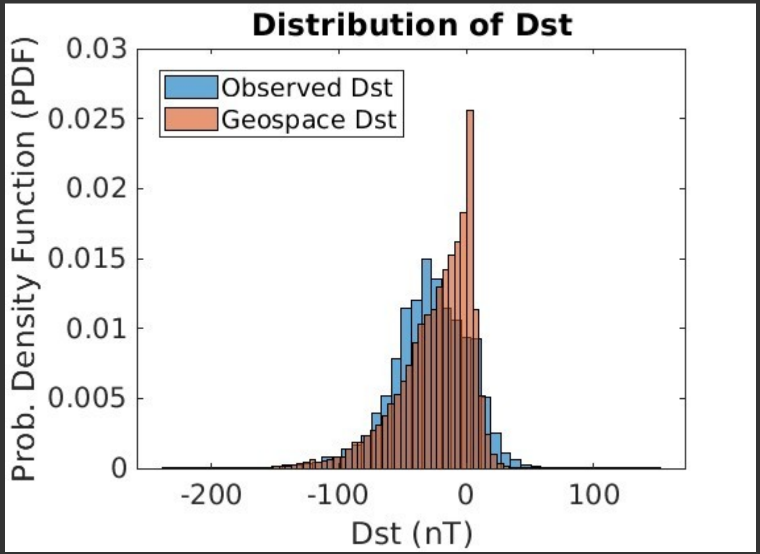
Statistical validation of SWPC Geospace





Statistical validation of SWPC Geospace

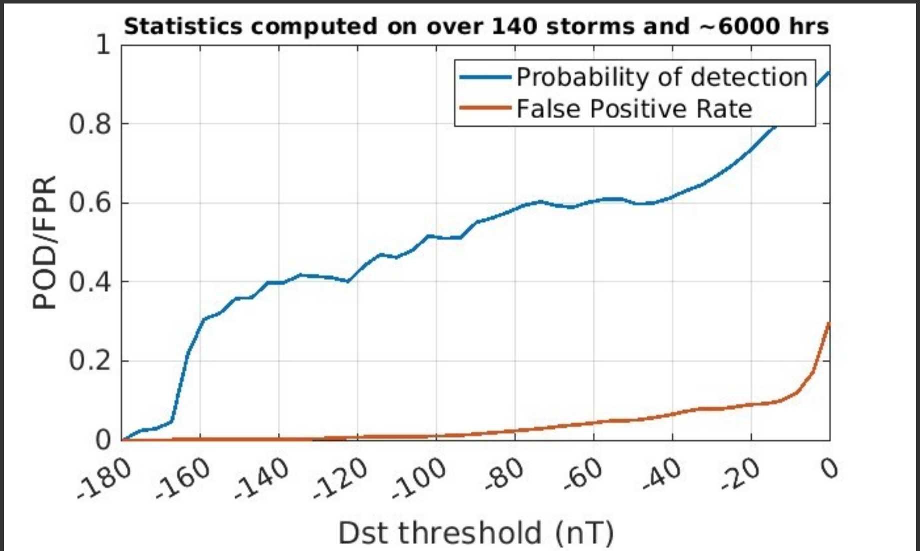


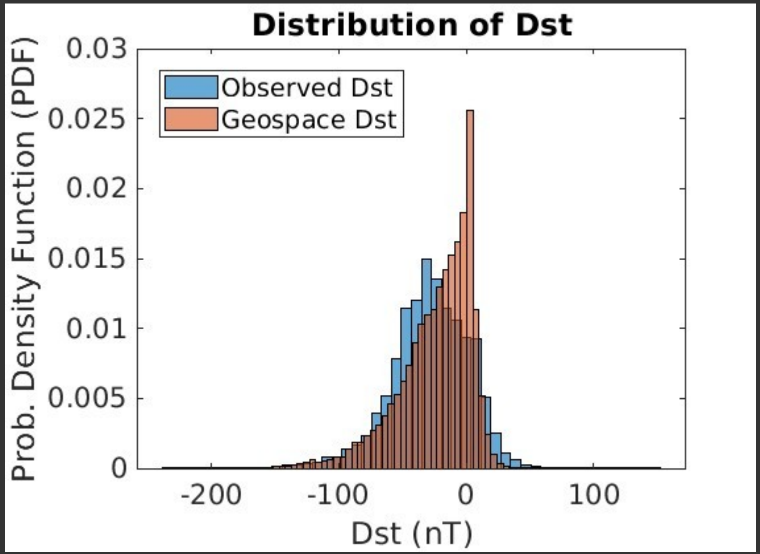


From SWAG recommendations (DRAFT)
https://www.weather.gov/media/nws/Final-SWAG-Slides_18-20-Jan-2023-r.pdf

"AI/ML methods have an important role in future SWx models,
but they can not replace physics based models for extreme events"

Statistical validation of SWPC Geospace

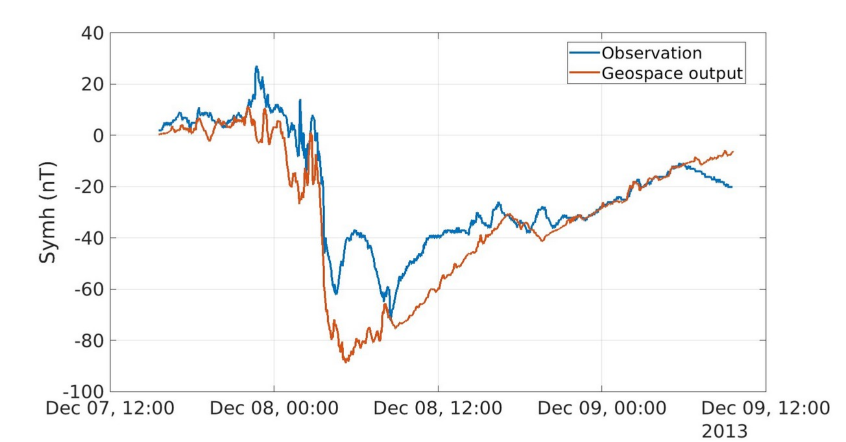


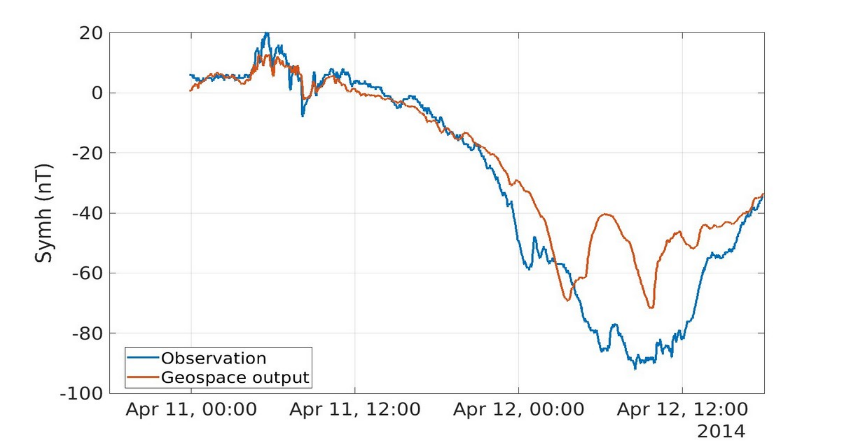


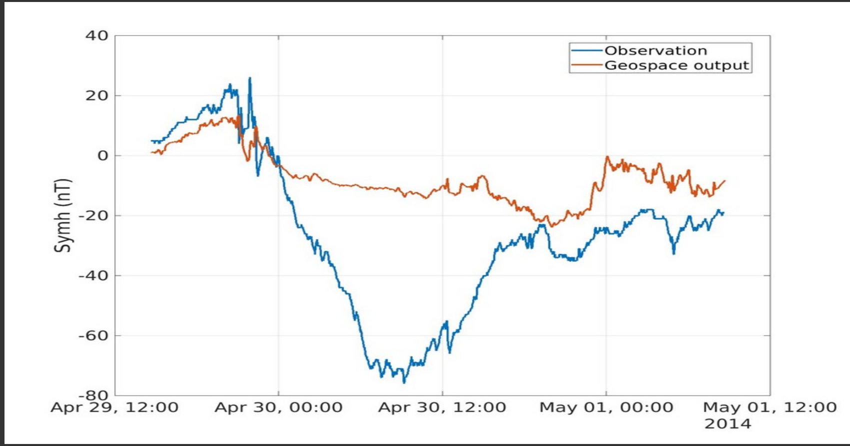
From SWAG recommendations (DRAFT)
https://www.weather.gov/media/nws/Final-SWAG-Slides_18-20-Jan-2023-r.pdf

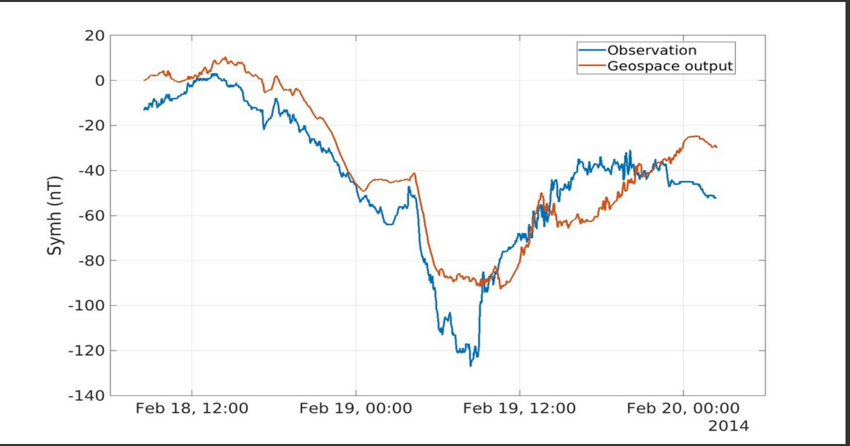
"AI/ML methods have an important role in future SWx models,
but they can not replace physics based models for extreme events"

Statistical validation of SWPC Geospace

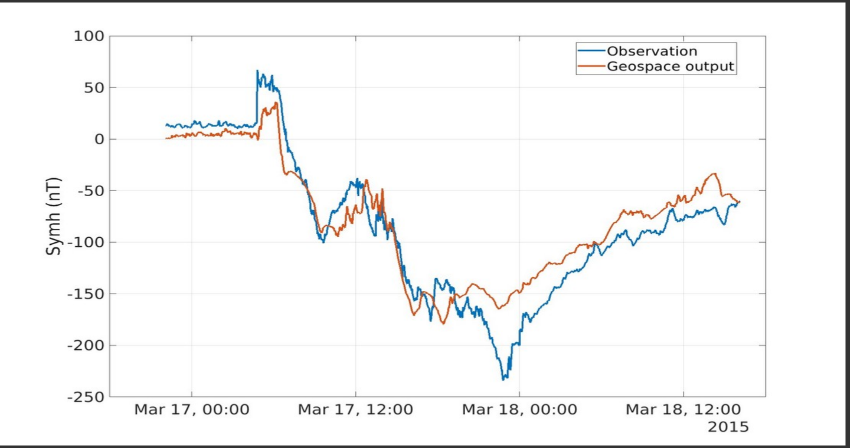


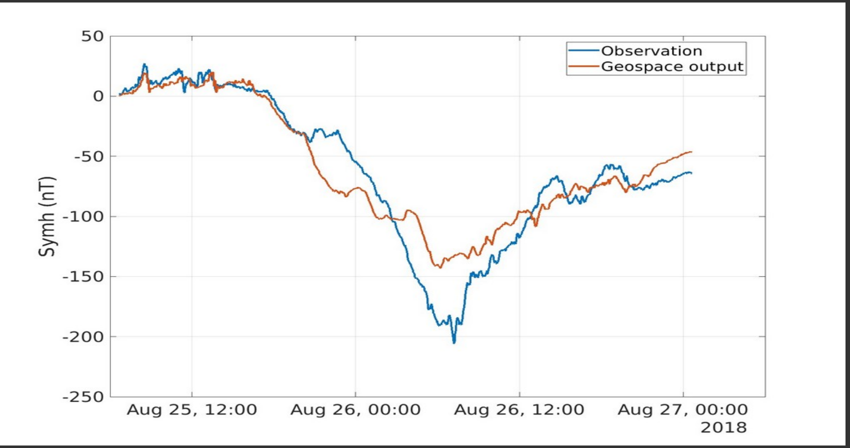


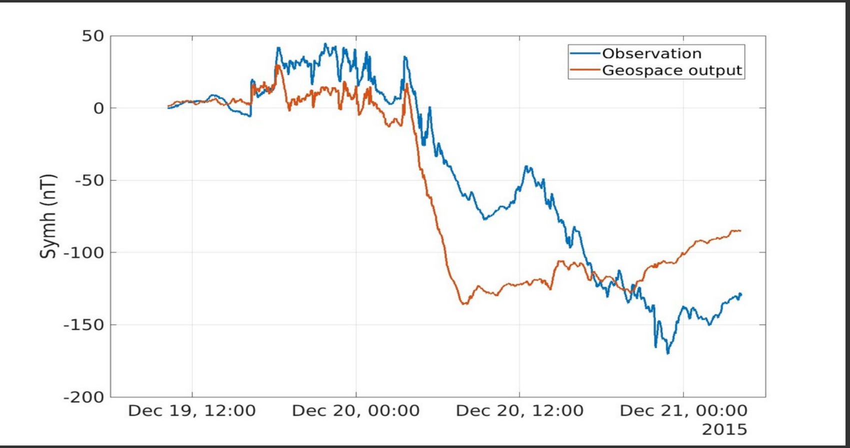


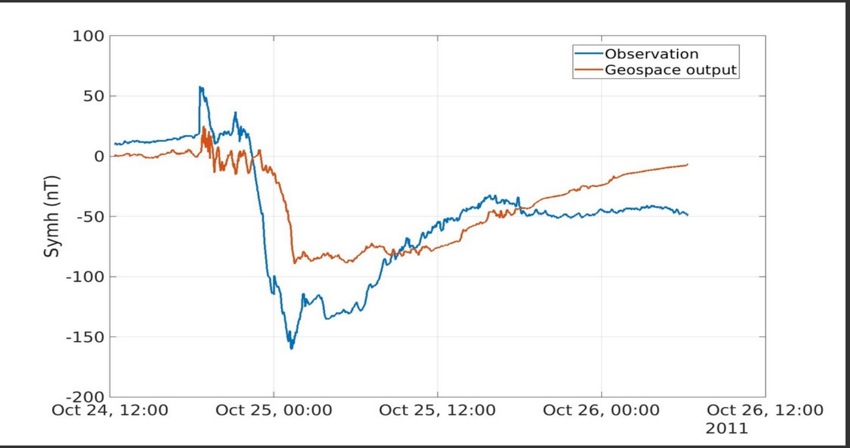


Statistical validation of SWPC Geospace



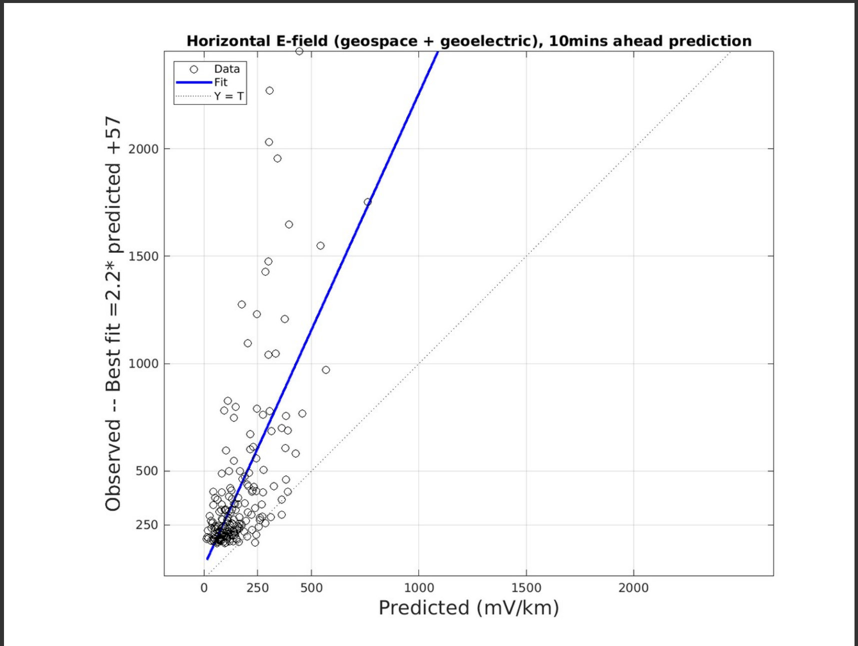


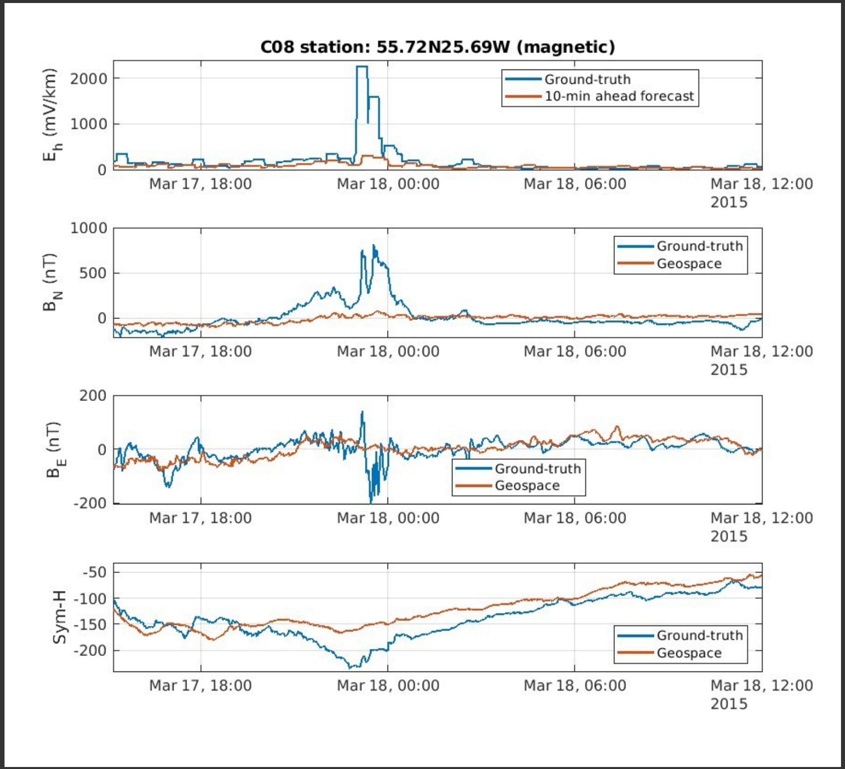




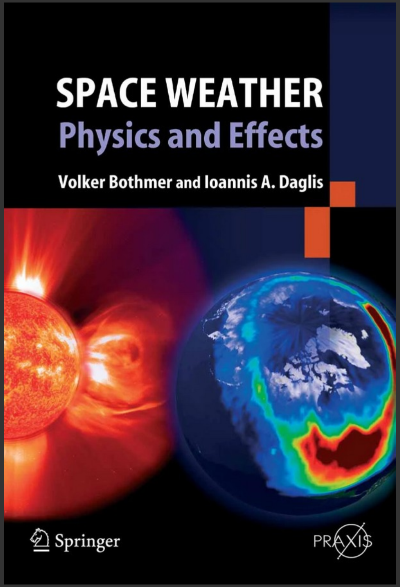
Statistical validation of SWPC Geospace

Prototype of Geospace driving Geoelectric model
(10 minutes ahead predictions)





How did we get where we are?
The myth of the analogy with meteorology



Published in 2007

Space weather forecasting historically viewed through the lens of meteorology

George Siscoe

The history of progress in the effectiveness of meteorological forecasting can be divided into ten stages: (1) recognition of societal need; (2) development of rules for forecasts based on visual observations; (3) quantification of storm parameters through instrument observations; (4) development of retrospective synoptic weather maps; (5) institution of forecast centers after the technological means of forecasting (the telegraph and instrument-based weather maps) came into being; (6) development of models of storm structure; (7) subjective analysis based on weather chart analysis; (8) objective analysis based on empirical formulas; (9) numerical predictions based on integrating the equations of atmospheric motion; and (10) storm tracking by radar and satellites. A parallel division of the history of space weather forecasting is here recounted. Whereas the effectiveness of meteorological forecasting dramatically increased with the advent of the numerical forecasting (stage 9), space weather forecasting is presently making progress through massively expanding its repertoire of objective forecast algorithms (stage 8). The advent of physics-based numerical space weather predictions (the stage of dramatic improvement in forecast effectiveness in meteorology, stage 9) is still in the future for space weather, although codes to achieve such predictions are under development. The crucial role that teaching forecasting in core meteorology courses has played in producing researchers motivated to improve forecasting effectiveness (and its absence in space weather curricula) is emphasized.

2.1 SIBLING SCIENCES

How did we get where we are?
The myth of the analogy with meteorology

REVIEW

doi:10.1038/nature14956

The quiet revolution of numerical weather prediction

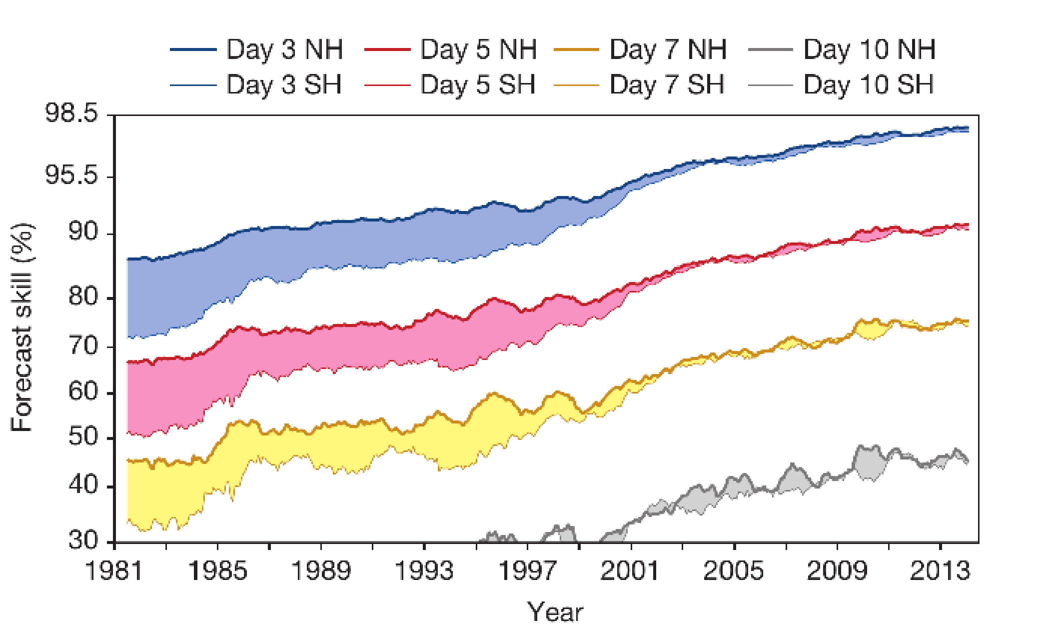
Peter Bauer[1], Alan Thorpe[1] & Gilbert Brunet[2]

Advances in numerical weather prediction represent a quiet revolution because they have resulted from a steady accumulation of scientific knowledge and technological advances over many years that, with only a few exceptions, have not been associated with the aura of fundamental physics breakthroughs. Nonetheless, the impact of numerical weather prediction is among the greatest of any area of physical science. As a computational problem, global weather prediction is comparable to the simulation of the human brain and of the evolution of the early Universe, and it is performed every day at major operational centres across the world.

This Review explains the fundamental scientific basis of numerical weather prediction (NWP) before highlighting three areas from which the largest benefit in predictive skill has been obtained in the past—physical process representation, ensemble forecasting and model initialization.

Bauer, P., Thorpe, A., & Brunet, G. (2015). The quiet revolution of numerical weather prediction. Nature, 525(7567), 47-55.

How did we get where we are?
The myth of the analogy with meteorology



A measure of forecast skill at three-, five-, seven- and ten-day ranges, computed over the extra-tropical northern and southern hemispheres.

How did we get where we are?
The myth of the analogy with meteorology

- For the last 20 years the accepted narrative in the space weather community has been that SWx is a ‘younger sibling’ of terrestrial weather.
- As such, the hope has always been that SWx will soon achieve the same level of extraordinary success witnessed in numerical weather predictions.
  - As a consequence the majority of effort (and funding) has been spent in developing physics-based models.

How did we get where we are?
The myth of the analogy with meteorology

- It is now time for revisiting that narrative and to recognize that physics-based models have not delivered what was expected ~20 years ago, and probably will never do.

How did we get where we are?
The myth of the analogy with meteorology

- It is now time for revisiting that narrative and to recognize that physics-based models have not delivered what was expected ~20 years ago, and probably will never do.
- The fundamental reason is that in the Sun-Earth system, we will never reach the required level of accuracy in:
  - Physical processes representation (i.e. separation of scales between kinetic and fluid physics)
  - Ensemble modeling (lack of in-situ observations of space weather drivers)
  - Model initialization/data assimilation (lack of observations to be assimilated – in NWP on the order of 10M observations)

State-of-the-art operational physics-based models

	Lead-time	Actionable (lead-time + Uncertainty Quantification)	Accuracy	Cost
WSA/ENLIL	1 – 3 days ahead prediction	No uncertainty	No better than zero-cost models	~ **50,000** CPU-hours/year Approximately **$2,000/year** (excluding storage, RAM, etc)
SWMF (Geospace)	20 – 60 mins ahead	No uncertainty	Good at timing but no better than persistence in accuracy	~ **420,000** CPU-hours/year Approx. **$15,000/year** (excluding storage, RAM, etc)
WAM/IPE	2 days ahead prediction	No uncertainty	Not sure about validation	~**320,000** CPU-hours/year Approx. **$10,000/year** (excluding storage, RAM, etc)

	Lead-time	**Actionable (lead-time + Uncertainty Quantification)**	**Accuracy**	**Cost**
Machine Learning	1 hour – 3 days	Ensemble or built-in UQ	Typically better than physics-based equivalent	Example: LiveDst (a forecast every 15 minutes) ~ **250** CPU-hours/year costs **<$10/year** (including everything: storage, I/O time, RAM, etc.)

	Lead-time	Actionable (lead-time + Uncertainty Quantification)	Accuracy	Cost
Machine Learning	1 hour – 3 days	Ensemble or built-in UQ	Typically better than physics-based equivalent	Example: LiveDst (a forecast every 15 minutes) ~ **250** CPU-hours/year costs **<$10/year** (including everything: storage, I/O time, RAM, etc.)

Take home message:
ML will become the standard way of SWx forecasting by the end of the decade

Why?

😁

	Lead-time	**Actionable (lead-time + Uncertainty Quantification)**	**Accuracy**	**Cost**
Machine Learning	1 hour – 3 days	Ensemble or built-in UQ	Typically better than physics-based equivalent	Example: LiveDst (a forecast every 15 minutes) ~ **250** CPU-hours/year costs **<$10/year** (including everything: storage, I/O time, RAM, etc.)

Take home message:
ML will become the standard way of SWx forecasting by the end of the decade

Why?

1) Physics based model have not delivered (and probably will never do)

😁

	Lead-time	Actionable (lead-time + Uncertainty Quantification)	Accuracy	Cost
Machine Learning	1 hour – 3 days	Ensemble or built-in UQ	Typically better than physics-based equivalent	Example: LiveDst (a forecast every 15 minutes) ~ **250** CPU-hours/year costs **<$10/year** (including everything: storage, I/O time, RAM, etc.)

Take home message:
ML will become the standard way of SWx forecasting by the end of the decade

Why?

1) Physics based model have not delivered (and probably will never do)

2) ML models have shown to outperform physics-based in lead-time, accuracy, actionability, and cost

😁

	Lead-time	Actionable (lead-time + Uncertainty Quantification)	Accuracy	Cost
Machine Learning	1 hour – 3 days	Ensemble or built-in UQ	Typically better than physics-based equivalent	Example: LiveDst (a forecast every 15 minutes) ~ **250** CPU-hours/year costs **<$10/year** (including everything: storage, I/O time, RAM, etc.)

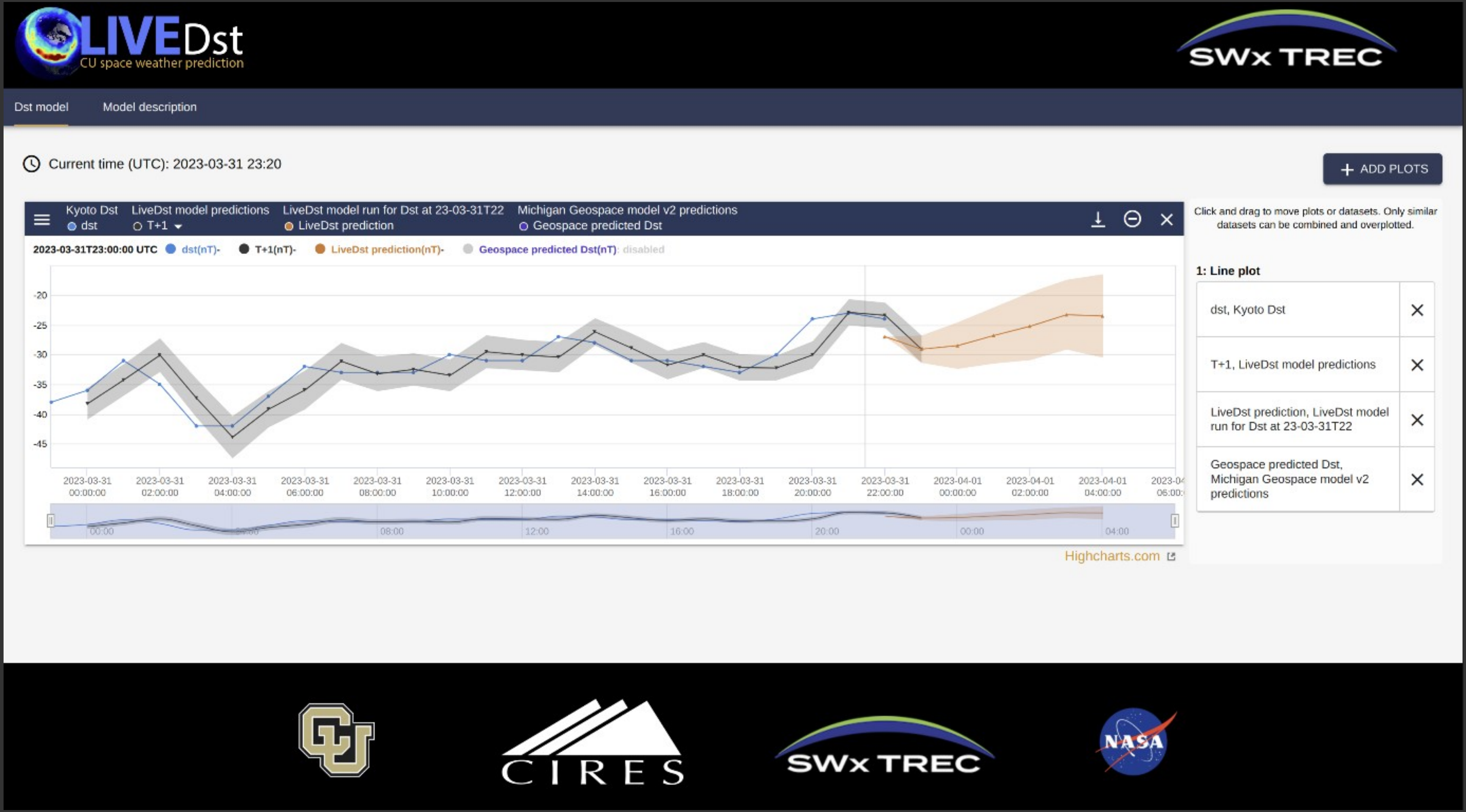
Take home message:
ML will become the standard way of SWx forecasting by the end of the decade

Why?

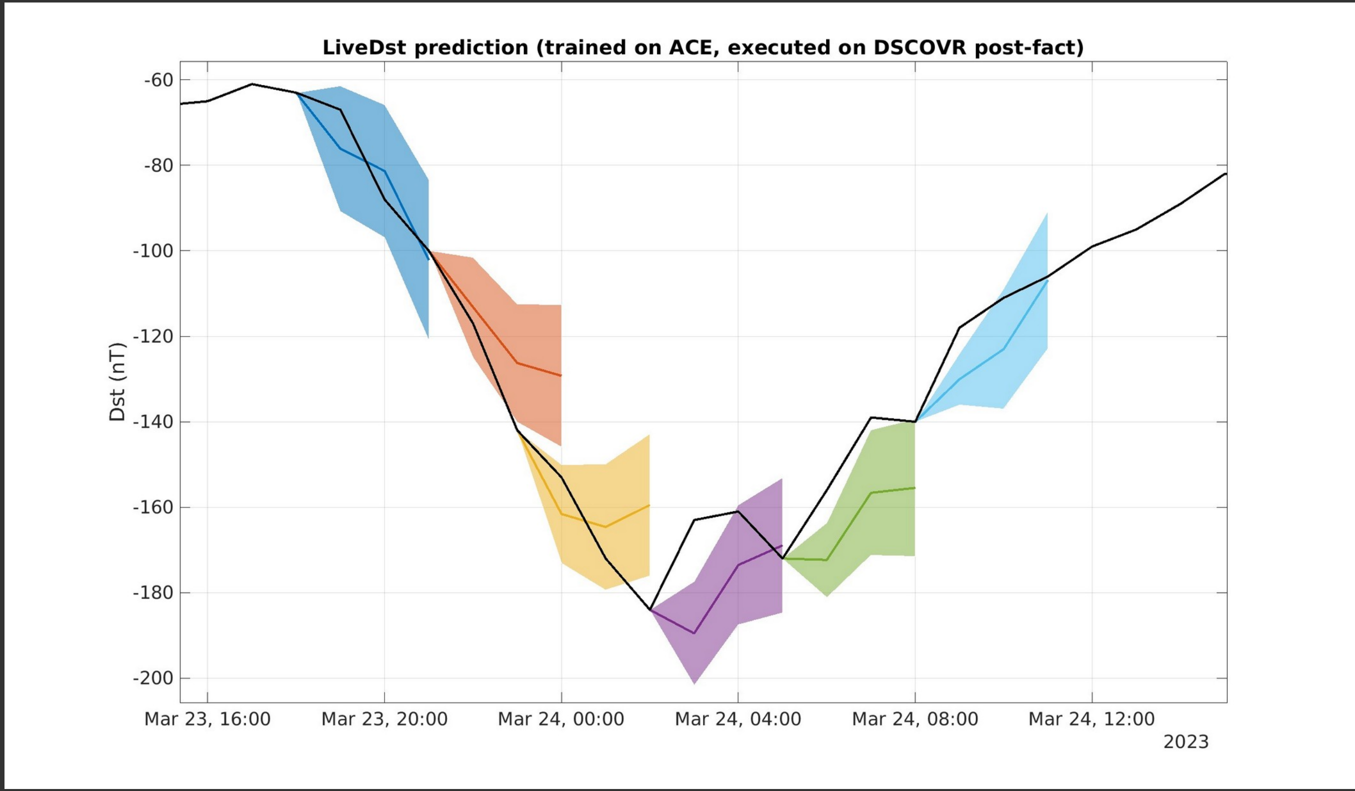
1) Physics based model have not delivered (and probably will never do)

2) ML models have shown to outperform physics-based in lead-time, accuracy, actionability, and cost

3) It’s going to happen no matter what!



https://swx-trec.com/dst/



Post-hoc LiveDst forecast of 03/24/23 storm

Credit: Andong Hu (CIRES)

BIG caveat: LiveDst was not running live during the 03/24/23 event, because ACE didn't return any data.
At this time we don't have a back-up system that uses DSCOVR when ACE is not available (it's being worked on and will be ready by Space Weather Workshop!)

LiveDst is trained on ACE data and this plot shows the prediction that could have been issued if executed on DSCOVR data

Space Weather

RESEARCH ARTICLE
10.1029/2022SW003286

Special Section:
Machine Learning in Heliophysics

Key Points:

- A new multi-hour ahead Dst prediction model developed from solar wind observations using Gated Recurrent Unit networks is proposed
- The uncertainty of the proposed Dst model is estimated by applying the Accurate and Reliable Uncertainty Estimate method
- A multi-fidelity method is developed to boost the performance of the model

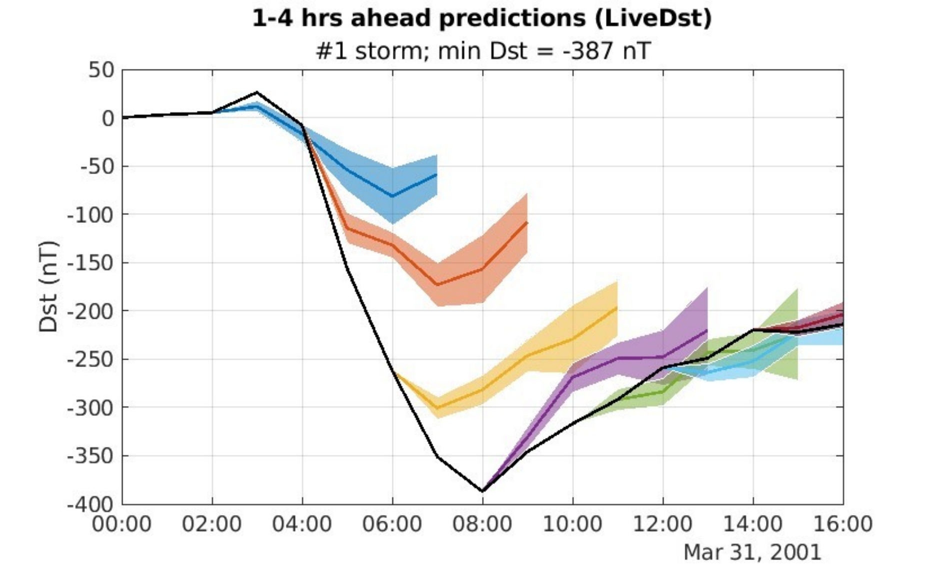
Correspondence to:

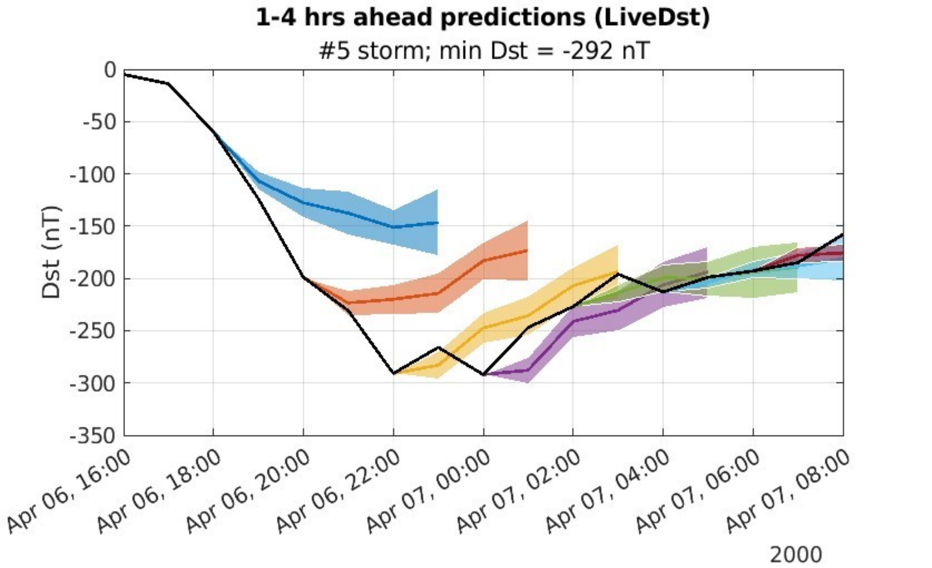
Multi-Hour-Ahead Dst Index Prediction Using Multi-Fidelity Boosted Neural Networks

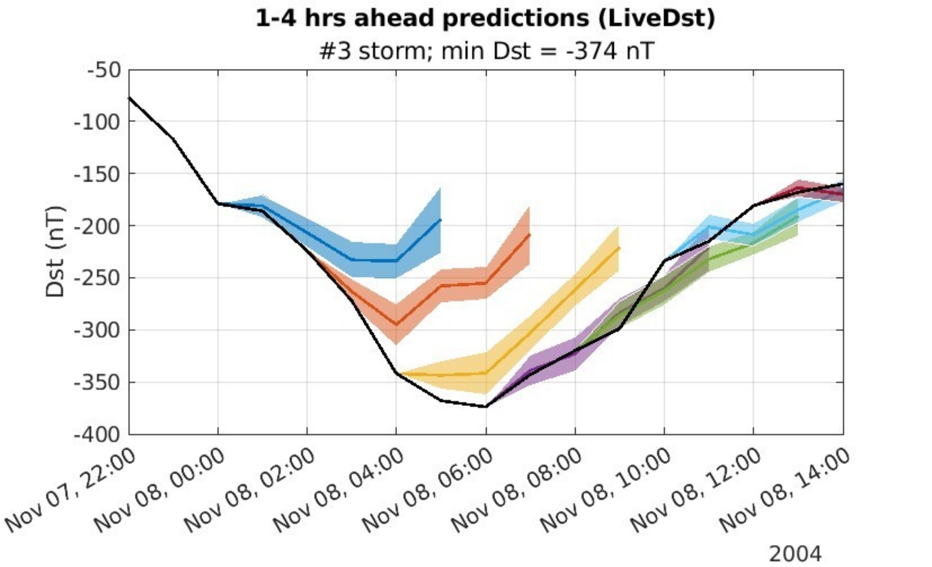
A. Hu[1], E. Camporeale[1,2], and B. Swiger[1,2]

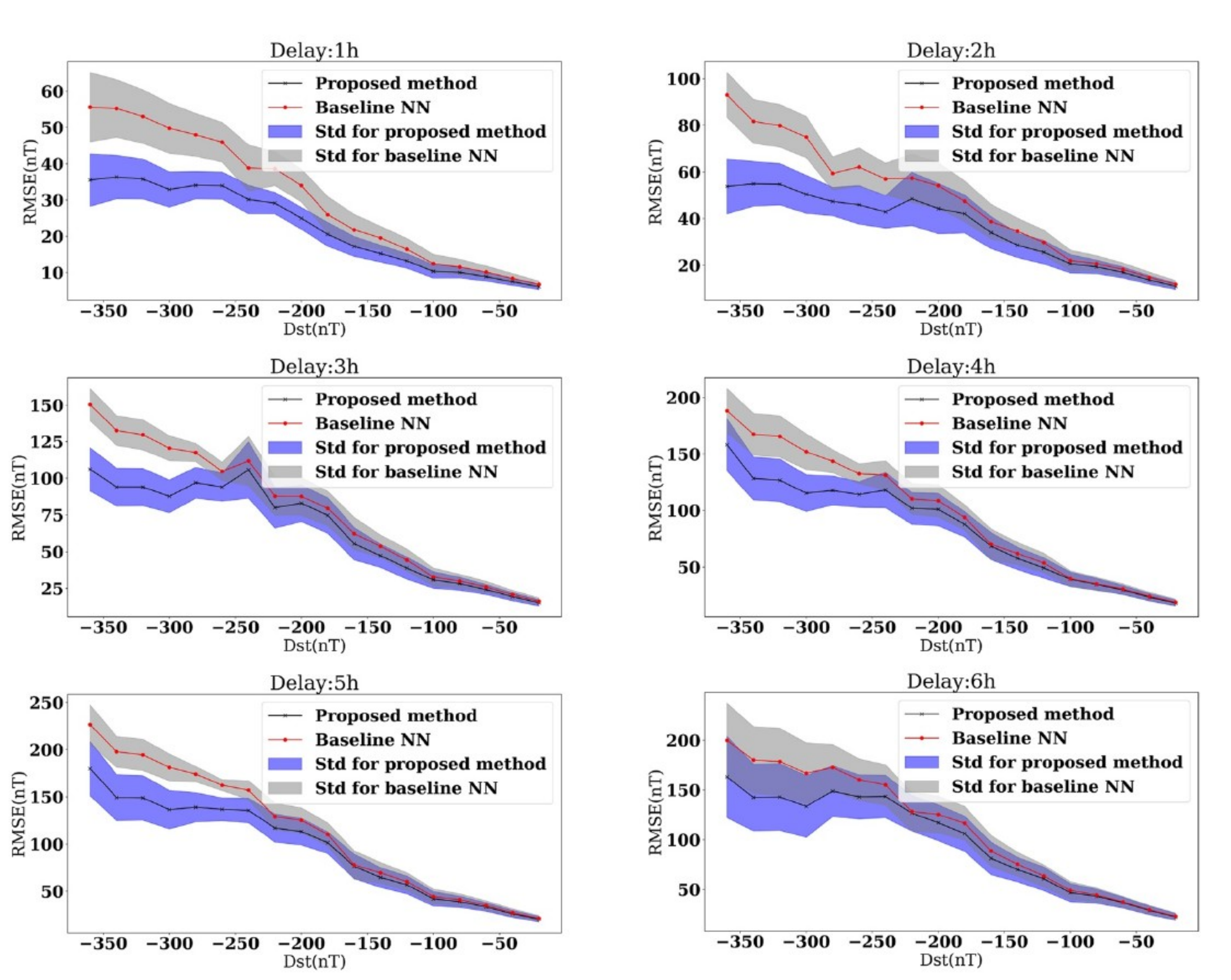
[1]CIRES, University of Colorado, Boulder, CO, USA, [2]NOAA Space Weather Prediction Center, Boulder, CO, USA

Abstract The Disturbance storm time (Dst) index has been widely used as a proxy for the ring current intensity, and therefore as a measure of geomagnetic activity. It is derived by measurements from four ground magnetometers in the geomagnetic equatorial region. We present a new model for predicting Dst with a lead time between 1 and 6 hr. The model is first developed using a Gated Recurrent Unit (GRU) network that is trained using solar wind parameters. The uncertainty of the Dst model is then estimated by using the Accurate and Reliable Uncertainty Estimate method (Camporeale & Carè, 2021, https://doi.org/10.1615/int.j.uncertaintyquantification.2021034623). Finally, a multi-fidelity boosting method is developed in order to enhance the accuracy of the model and reduce its associated uncertainty. It is shown that the developed model can predict Dst 6 hr ahead with a root-mean-square-error of 13.54 nT. This is significantly better than a persistence model or a single GRU model.


1-4 hrs ahead predictions (LiveDst)
#1 storm; min Dst = -387 nT
Dst (nT)
50
0
-50
-100
-150
-200
-250
-300
-350
-400
00:00
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
Mar 31, 2001


1-4 hrs ahead predictions (LiveDst)
#5 storm; min Dst = -292 nT
Dst (nT)
0
-50
-100
-150
-200
-250
-300
-350
Apr 06, 16:00
Apr 06, 18:00
Apr 06, 20:00
Apr 06, 22:00
Apr 07, 00:00
Apr 07, 02:00
Apr 07, 04:00
Apr 07, 06:00
Apr 07, 08:00
2000


1-4 hrs ahead predictions (LiveDst)
#3 storm; min Dst = -374 nT
Dst (nT)
-50
-100
-150
-200
-250
-300
-350
-400
Nov 07, 22:00
Nov 08, 00:00
Nov 08, 02:00
Nov 08, 04:00
Nov 08, 06:00
Nov 08, 08:00
Nov 08, 10:00
Nov 08, 12:00
Nov 08, 14:00
2004


Delay:1h
Proposed method
Baseline NN
Std for proposed method
Std for baseline NN
RMSE(nT)
Dst(nT)
Delay:2h
Delay:3h
Delay:4h
Delay:5h
Delay:6h

ACCRUE
(ACCurate and Reliable Uncertainty Estimate)

The ACCRUE method:

- Estimates the uncertainties associated with single-point outputs generated by a deterministic model, in terms of Gaussian distributions;
- Ensures the optimal trade-off between accuracy and reliability;
- Does not need to run ensembles. It costs as much as training a neural network
- Code available: zenodo.1485608

Space Weather

RESEARCH ARTICLE

10.1029/2018SW002026

On the Generation of Probabilistic Forecasts From Deterministic Models

E. Camporeale[1,2], X. Chu[3], O. V. Agapitov[4], and J. Bortnik[5]

[1]Center for Mathematics and Computer Science (CWI), Amsterdam, The Netherlands, [2]Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, USA, [3]Laboratory for Atmospheric and Space Physics, University of Colorado, Boulder, CO, USA, [4]Space Sciences Laboratory, University of California Berkeley, Berkeley, CA, USA, [5]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, CA, USA

Key Points:
- We introduce a new method to estimate the uncertainties associated with single-point outputs generated by a deterministic model
- The method ensures a trade-off between accuracy and reliability of the generated probabilistic forecasts
- Computationally cheap model:

International Journal for Uncertainty Quantification, 11(4):81–94 (2021)

ACCRUE: ACCURATE AND RELIABLE UNCERTAINTY ESTIMATE IN DETERMINISTIC MODELS

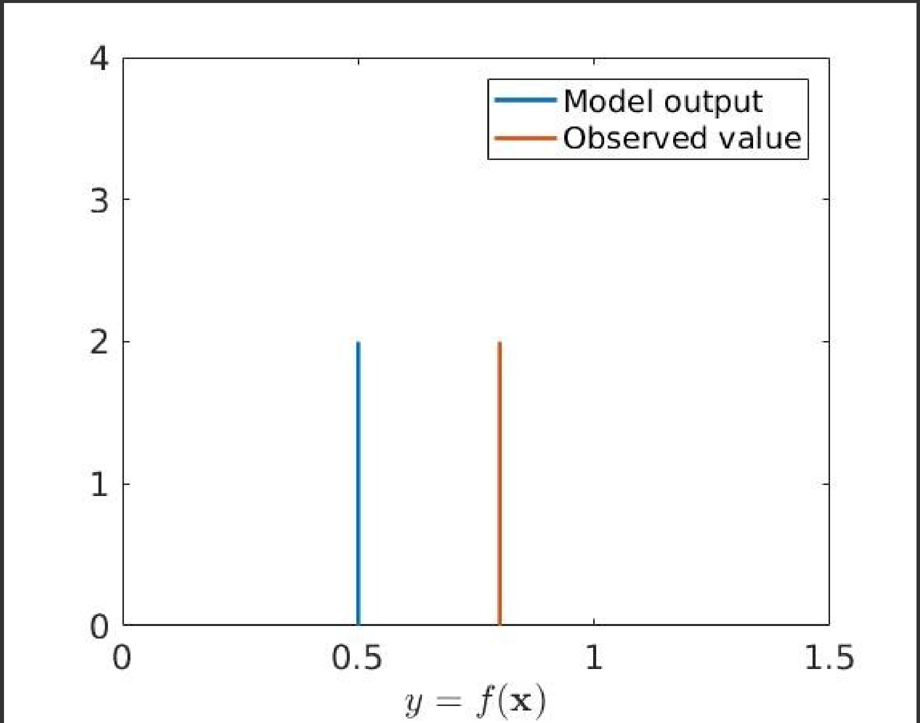
Enrico Camporeale[1,] & Algo Carè[2]*

ACCRUE Approach

Let us assume that for a single (multidimensional) input **x**, our model predicts an output y = f(x).

Blue line → Model output
Red line → Real (observed value)

<u>Working hypothesis:</u>
We want to use the model output as the mean of a Gaussian distribution that is interpreted as a probabilistic forecast.



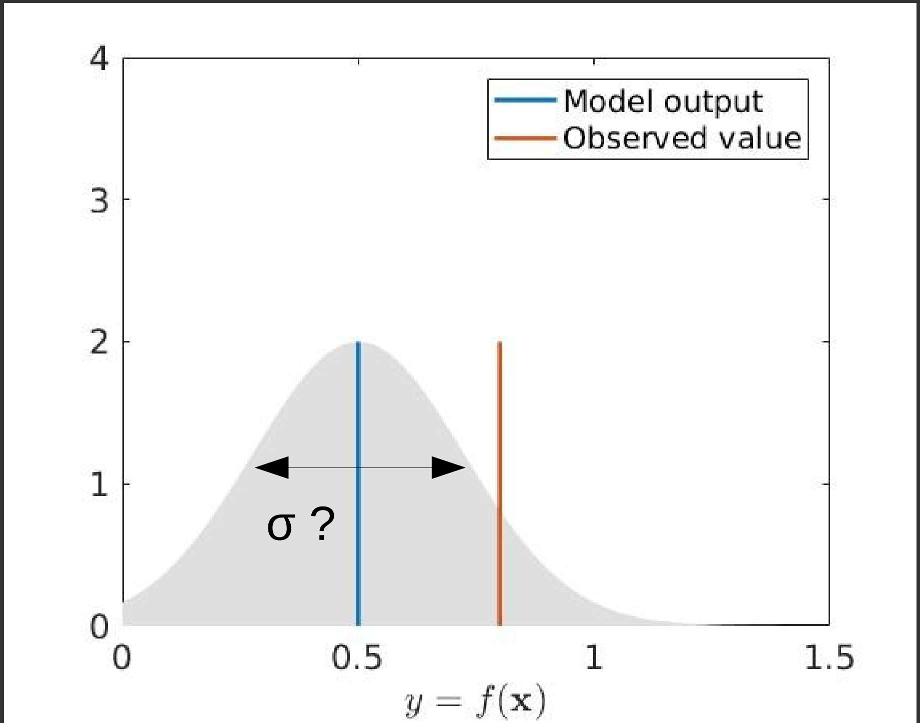
ACCRUE Approach

Let us assume that for a single (multidimensional) input **x**, our model predicts an output y = f(x).

Blue line → Model output
Red line → Real (observed value)

<u>Working hypothesis:</u>
We want to use the model output as the mean of a Gaussian distribution that is interpreted as a probabilistic forecast.

What is the optimal width of a Gaussian forecast?



ACCRUE Recipe

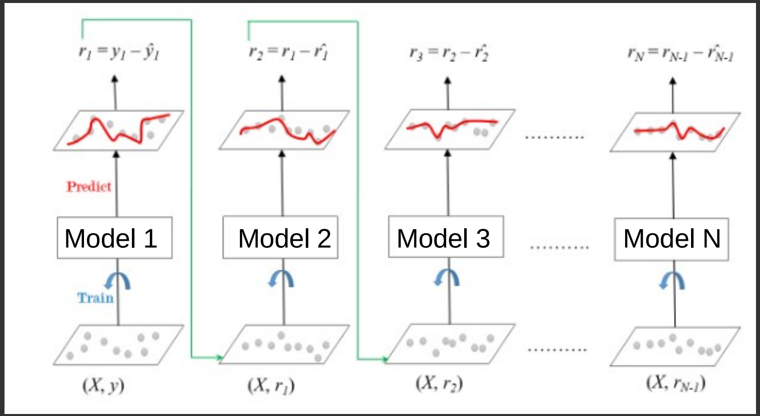
- The optimal Gaussian width (std) is the one that optimizes both accuracy and reliability
- This is a two-objective optimization problem, because reliability and accuracy are competing objectives.
- We define the Accuracy-Reliability (AR) cost function:

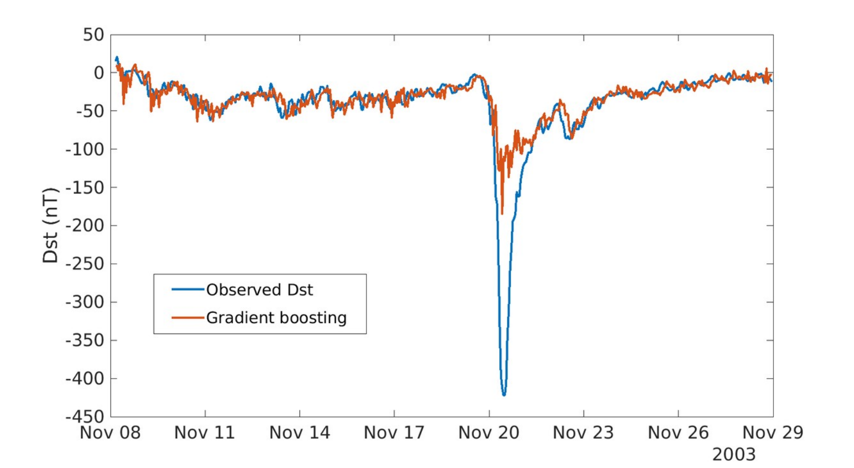
$$AR = \underbrace{CRPS}_{\text{Accuracy}} + \underbrace{RS}_{\text{Reliability}}$$

- Accuracy and Reliability cannot both be minimized simultaneously and we have to find the best trade-off
- The minimization problem is solved using a neural network

Gradient Boosting in one slide

- A hierarchy of models is built
- Each model is trained on the residuals (errors) of the previous one
- The final model is an additive combination of all sub-models
- One of the strongest algorithm
  - But it under-performs on imbalanced datasets





ProBoost in one slide

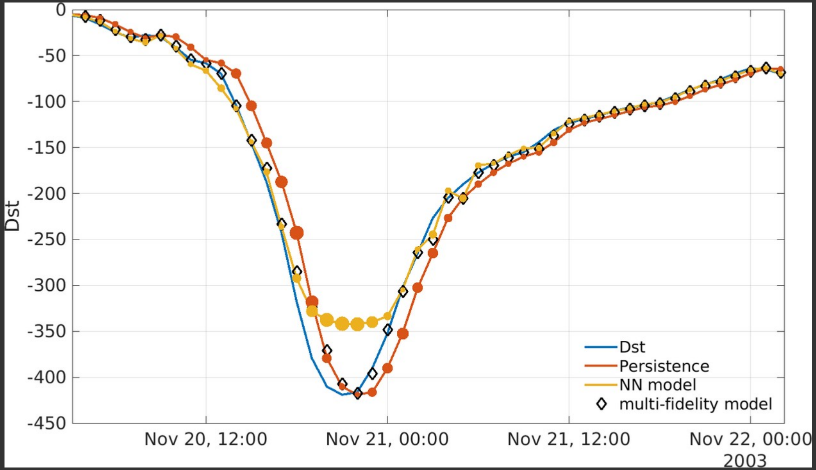
- We can use the power of ACCRUE to identify regions (in feature space) where a predictor works well and where it doesn’t

ProBoost in one slide

- We can use the power of ACCRUE to identify regions (in feature space) where a predictor works well and where it doesn’t
- An ensemble of models is built by sub-sampling the training set
  - The sub-sampling criterion is based on the ACCRUE uncertainty

ProBoost in one slide

- We can use the power of ACCRUE to identify regions (in feature space) where a predictor works well and where it doesn't
- An ensemble of models is built by sub-sampling the training set
  - The sub-sampling criterion is based on the ACCRUE uncertainty
- Sub-models are combined weighted by their precision = $1/\sigma^2$



Extending the lead-time to 1-2 days ahead

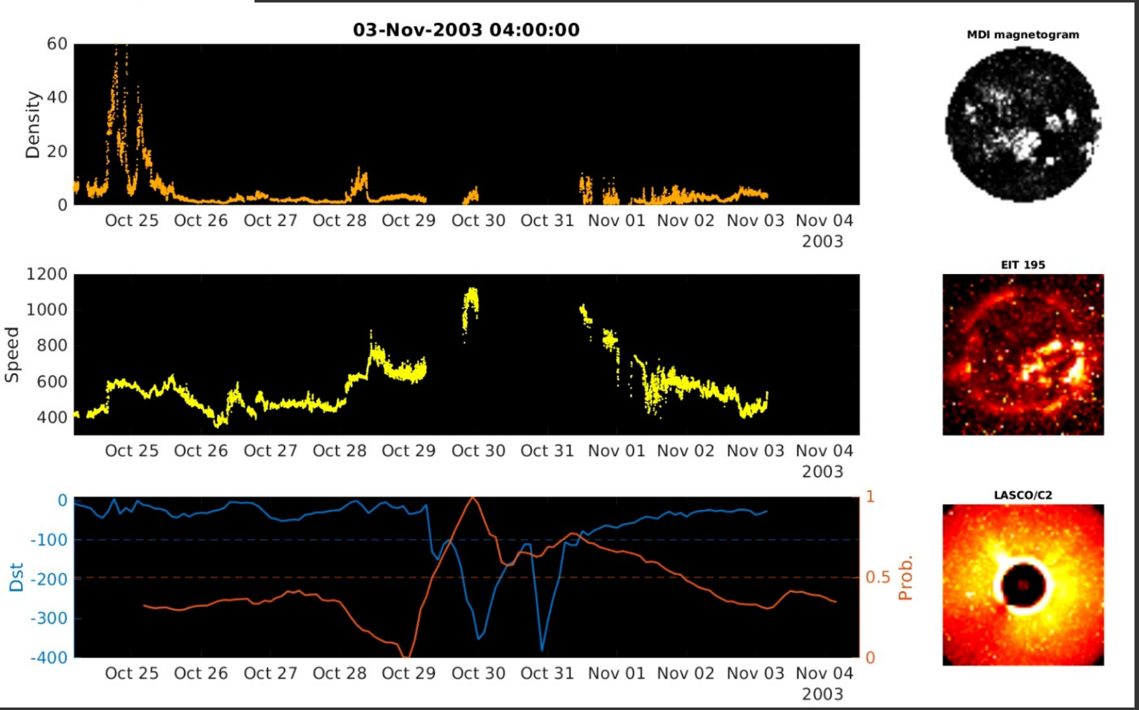
Space Weather®

Research Article | Open Access

Probabilistic Prediction of *Dst* Storms One-Day-Ahead Using Full-Disk SoHO Images

A. Hu, C. Shneider, A. Tiwari, E. Camporeale

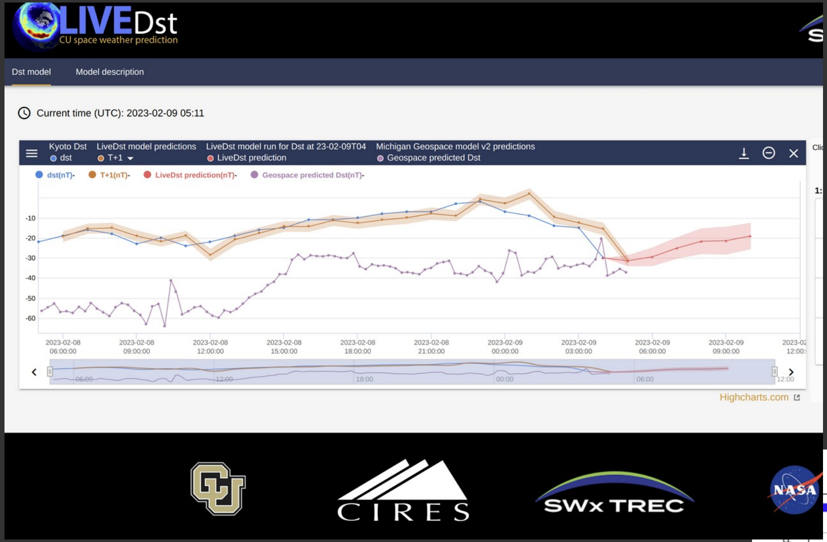
First published: 23 May 2022 | https://doi.org/10.1029/2022SW003064



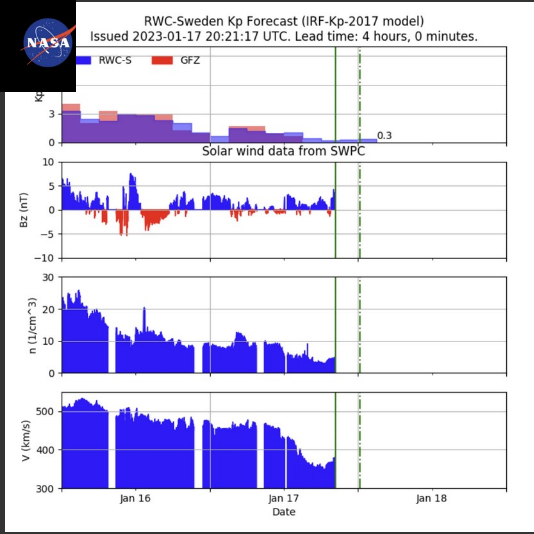
Continuation of this work is supported by NASA O2R grant 80NSSC21K1555

CIRES/SWPC postdoc: Brian Swiger

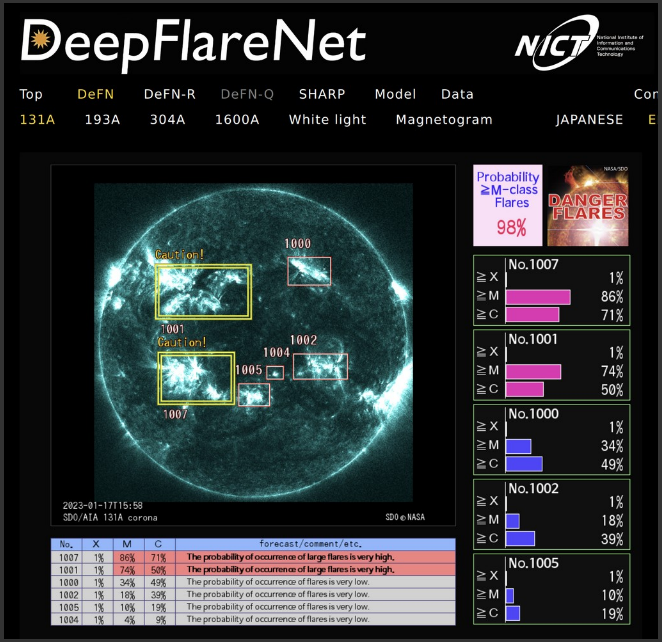
Operational ML models



https://swx-trec.com/dst/



https://swe.ssa.esa.int/irf-federated



https://defn.nict.go.jp/top_eng.html

See also:

https://spaceweather.gfz-potsdam.de/products-data/forecasts/

What’s the future like for SWx?

- ML model will soon supersede traditional methods
  - Physics-based models will become more and more irrelevant for forecasting

What’s the future like for SWx?

- ML model will soon supersede traditional methods
  - Physics-based models will become more and more irrelevant for forecasting

- Small research groups and private companies will have the means of developing, deploying and updating real-time predictive models (with UQ)
  - Training ML models will become (computationally and economically) cheaper and cheaper: no need to train them ‘from scratch’

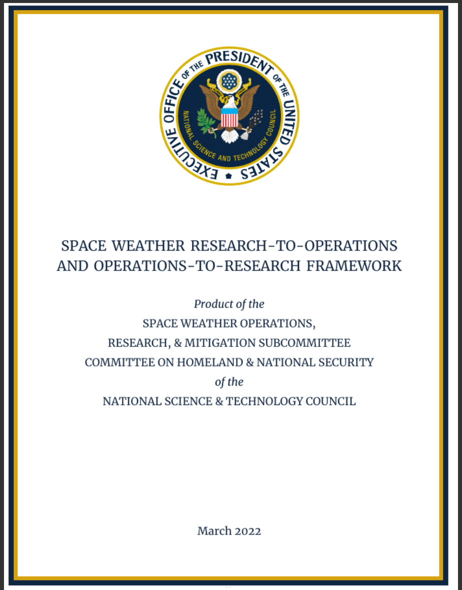
What's the future like for SWx?

- ML model will soon supersede traditional methods
  - Physics-based models will become more and more irrelevant for forecasting
- Small research groups and private companies will have the means of developing, deploying and updating real-time predictive models (with UQ)
  - Training ML models will become (computationally and economically) cheaper and cheaper: no need to train them 'from scratch'
- In contrast, SWPC might take years to push a model from RL4 to RL9 with the current O2R2O framework
  - By the time a model is operational it is most likely outdated by a better model

What's the future like for SWx?

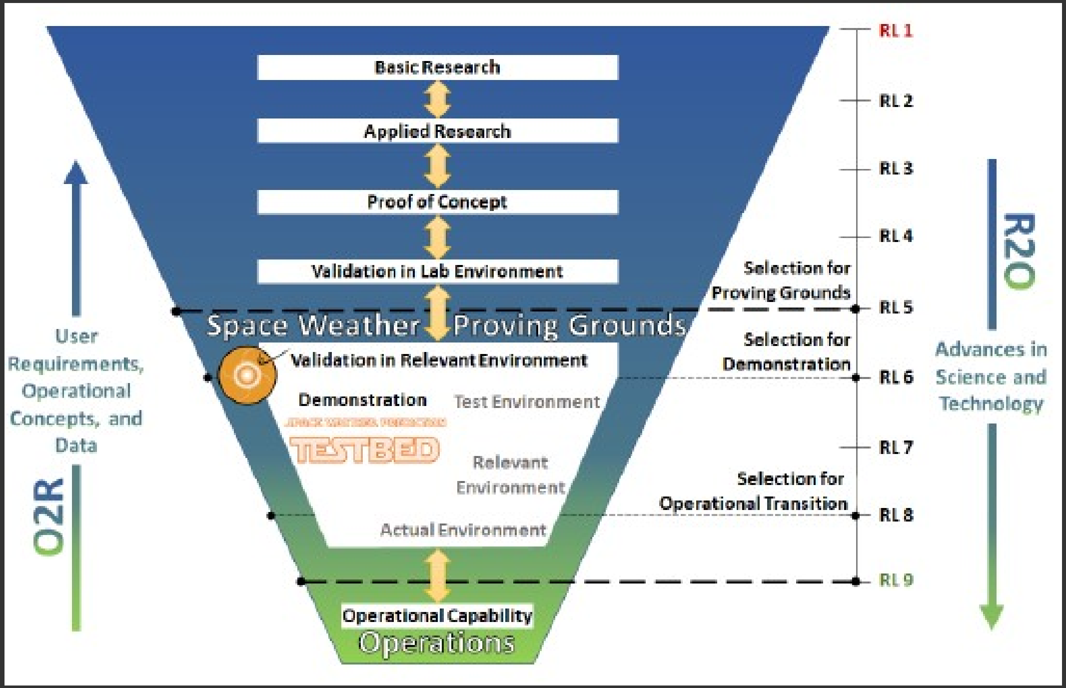
We need to address an urgent question:

How do we operationalize models in view of the rapidly changing landscape ??


SPACE WEATHER RESEARCH-TO-OPERATIONS AND OPERATIONS-TO-RESEARCH FRAMEWORK

Product of the
SPACE WEATHER OPERATIONS, RESEARCH, & MITIGATION SUBCOMMITTEE
COMMITTEE ON HOMELAND & NATIONAL SECURITY
of the
NATIONAL SCIENCE & TECHNOLOGY COUNCIL

March 2022



What's the future like for SWx?

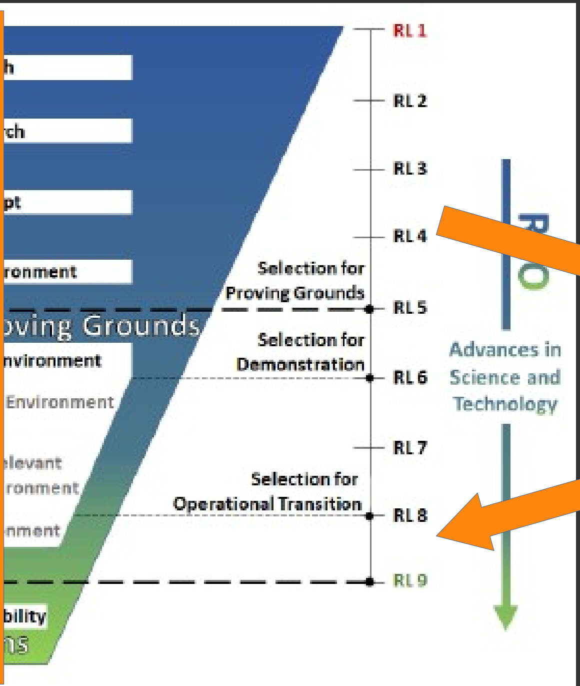
We need to address an urgent question:

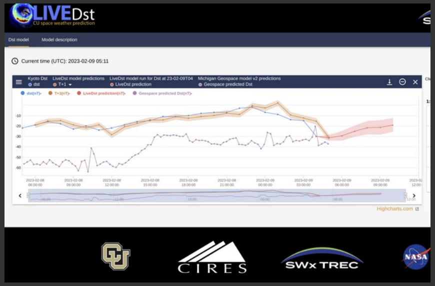
How do we operationalize models in view of the rapidly changing landscape ??

The current O2R2O framework is not "AI-aware"

I think there are (at least)

four different future scenarios





LiveDst runs on the cloud (AWS)

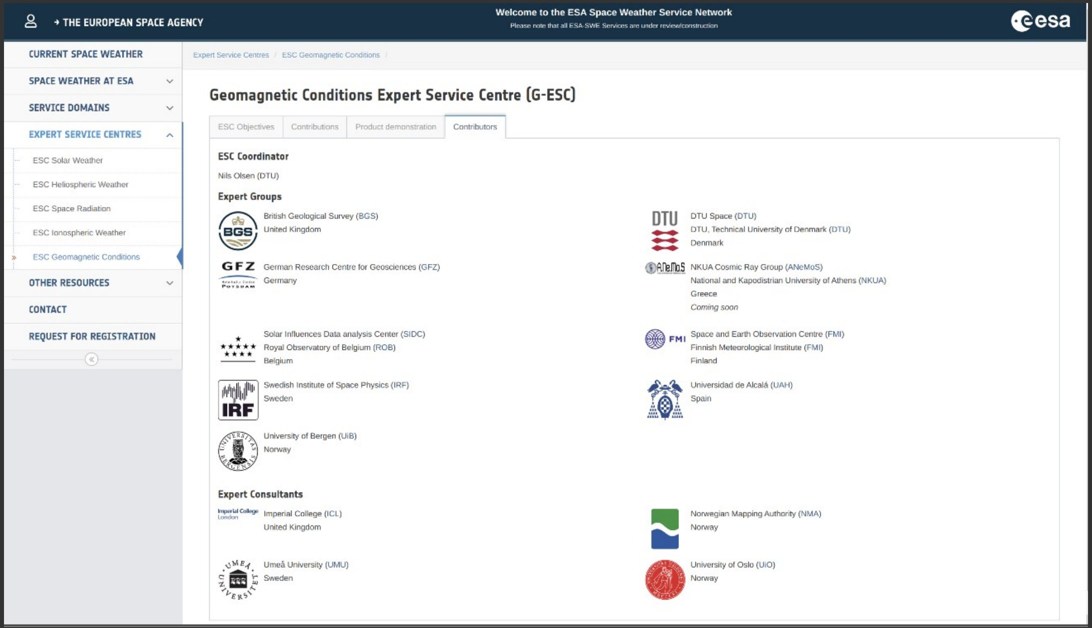
1) The NETFLIX scenario

- Subscription-based model
- SWPC buys **model outputs** from model developers (academia or private companies)
- SWPC scientists and software developers make their own visualizations to show the general public and SWFO
- Model output data IS data!
  - Commercial data program

2) The esa scenario

European Space Agency

- ESA Space Weather data and products are provided through a network of Expert Service Centres. Each of these comprises a distributed set of expert groups contributing particular data, products and/or expertise

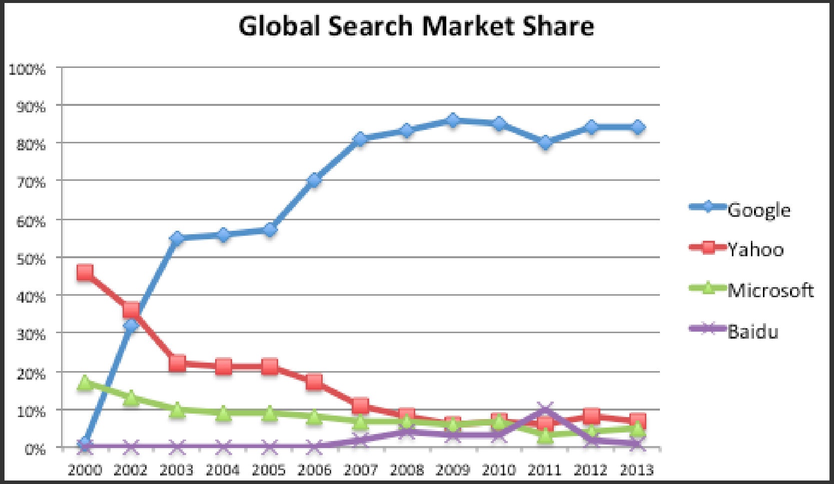


Mostly academic institutes and a few private companies

3) The yahoo! scenario

- Yahoo used to be the best (and only?) way to browse the web
- Today, Yahoo is not even trying to compete with Google
- They just keep doing what they do best (e.g., Yahoo Finance)
- SWPC becomes a data provider
- Forecasting is left to others

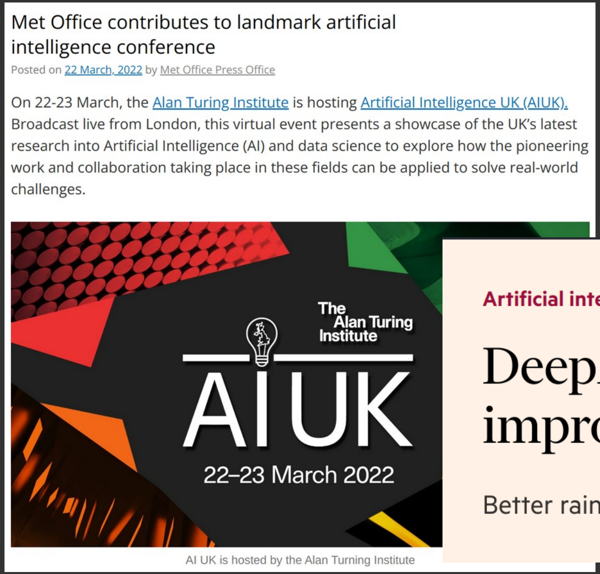
  (private companies, academia, etc.)
- Analogy with AccuWeather,
  The Weather Company, etc.



4) The UK/ South Korea scenario

- Massive Investments in AI from UK Met Office and Korean Space Weather Center
- Roadmap to build and retain in-house AI expertise
- Collaboration with world-leading AI powerhouses

    DeepMind, Alan Turing Institute, etc.


Met Office contributes to landmark artificial intelligence conference

Posted on 22 March, 2022 by Met Office Press Office

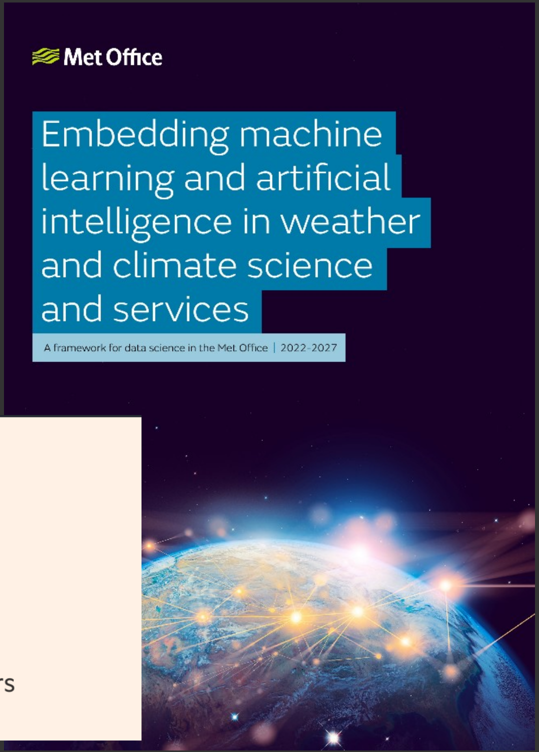
On 22-23 March, the Alan Turing Institute is hosting Artificial Intelligence UK (AIUK). Broadcast live from London, this virtual event presents a showcase of the UK's latest research into Artificial Intelligence (AI) and data science to explore how the pioneering work and collaboration taking place in these fields can be applied to solve real-world challenges.

AI UK is hosted by the Alan Turning Institute

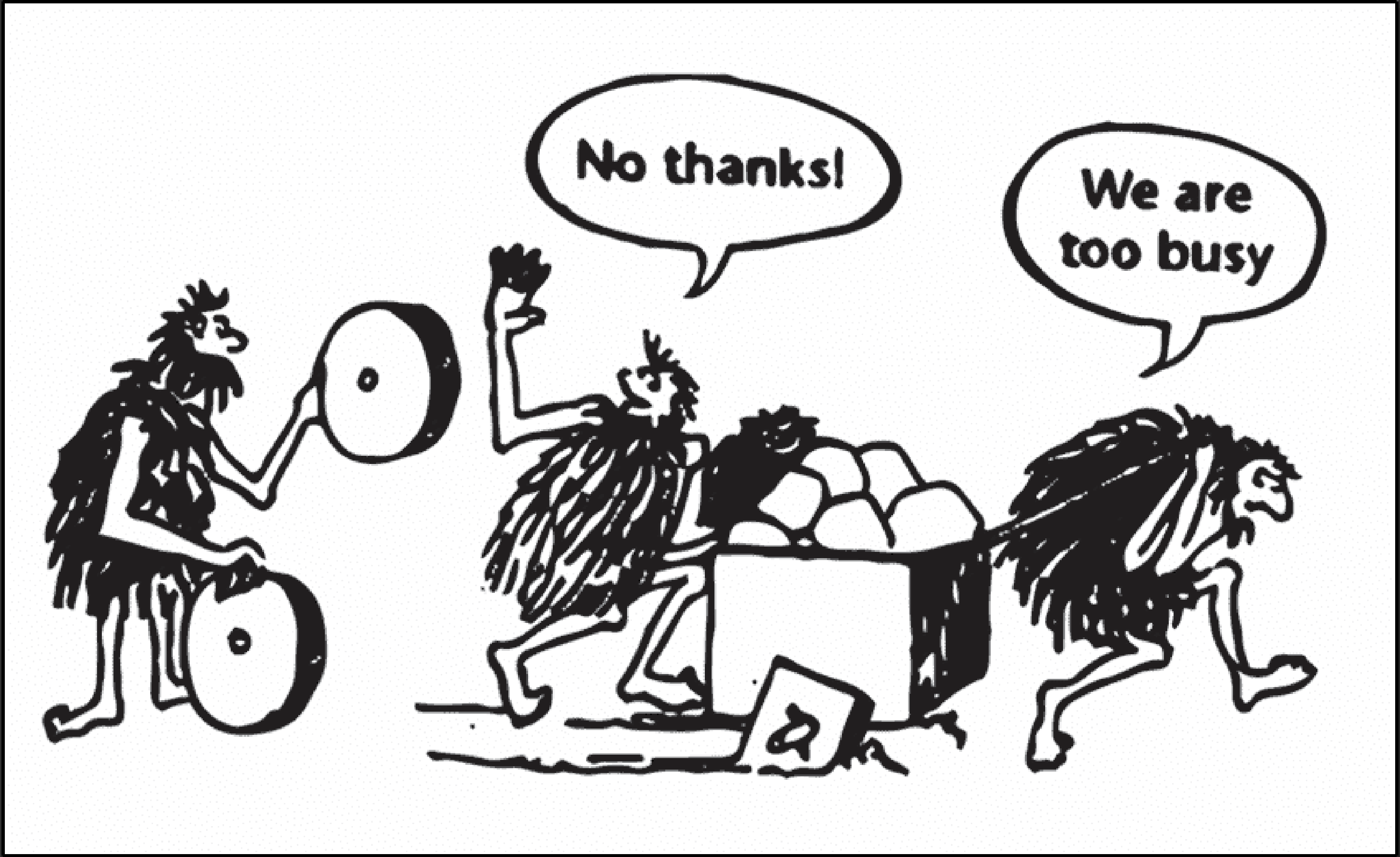

Artificial intelligence + Add to myFT

DeepMind and UK's Met Office use AI to improve weather forecasts

Better rainfall predictions could save lives when floods threaten, say researchers



5) The worst-case scenario



Conclusions

- This is the best time to be in SWx!
  - We are witnessing the dawn of a new era.
  - ML will play the major role in SWx forecasting by the end of the decade
- The analogy between SWx and meteorology that has been narrated in the last 20+ years was based on false premises.
- CU/CIRES + SWx-TREC is at the forefront of ML models for SWx.
  - Next models to be deployed on www.swx-trec.com:
    - Radiation belt electron fluxes (collaboration with UCLA/Bortnik)
    - Regional induced electric fields (work by Andong Hu)
    - Solar wind speed forecast
- The ability of adapting to the changing landscape and of leveraging the AI revolution will determine the future key players in SWx forecasting

Slides on https://github.com/ecamporeale/talks

Enrico.Camporeale@noaa.gov