Natural Language AI:
Forecasting Ionosphere
by Historical Analogy

Ivan Galkin[1], Paul Song[1], and Dieter Bilitza[2,3]

(1) Space Science Laboratory, UMass Lowell, USA

(2) Department of Physics and Astronomy, George Mason University, Fairfax, VA, USA

(3) Space Physics Data Facility, NASA, GSFC, Greenbelt, MD 20771, USA

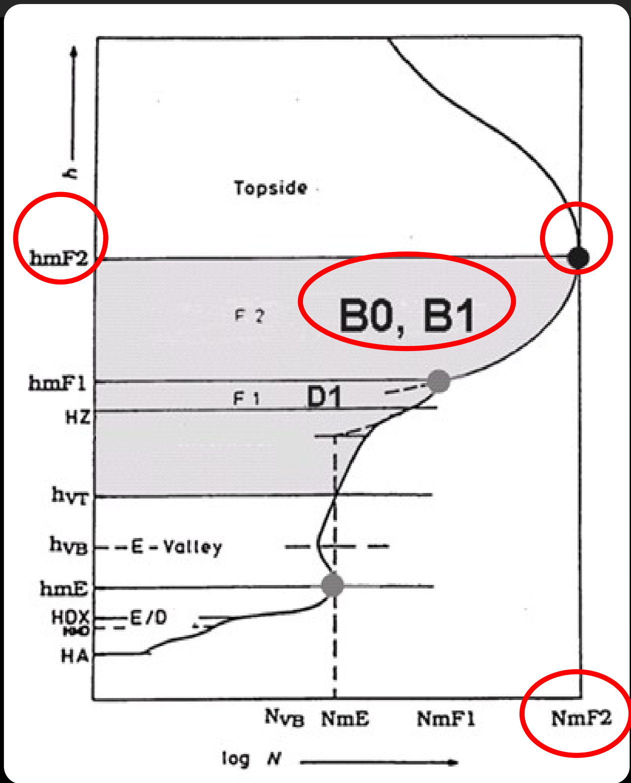
International Workshop on Machine Learning for Space Weather:

ICTP

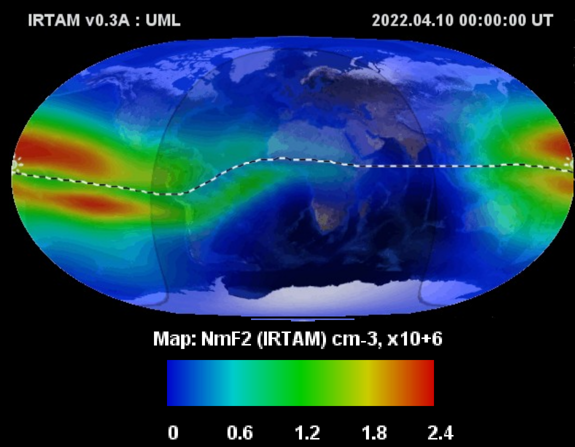
Buenos Aires, Argentina
7-11 November 2022

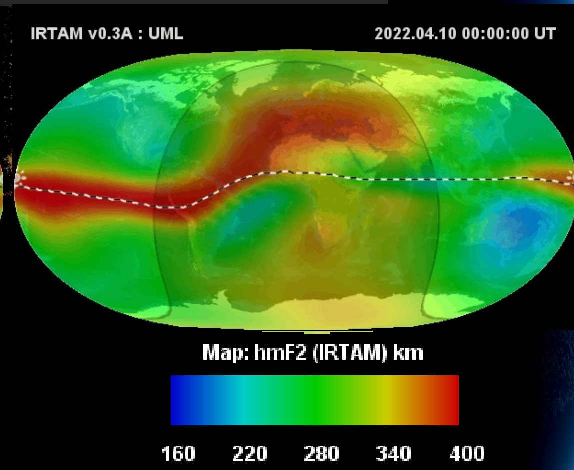
Ionosphere: Galloping introduction

3D volume of plasma and 2D maps of its properties



1D vertical profile of plasma density





N_mF2 h_mF2

One day of ionospheric dynamics

"Ionosphere is a major operational nuisance" © USAF

Layout

- Natural Language AI: Forecasting Ionosphere by Historical Analogy

Deep Learning:
Panacea?
or a Snake Oil…

GNSS PPP / RTK

- **Affected systems: autonomous vehicles and machinery**
- **TID as a *Silent Accuracy Kill***

- **Worse than the loss of lock and**

HF Geo

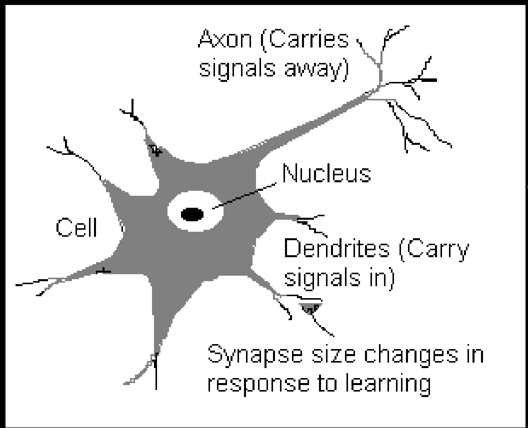
- **Geolocation of uncooperative HF transmitter**
- **Tens of km positioning errors**
- **Short-range catastrophe during TID passage**

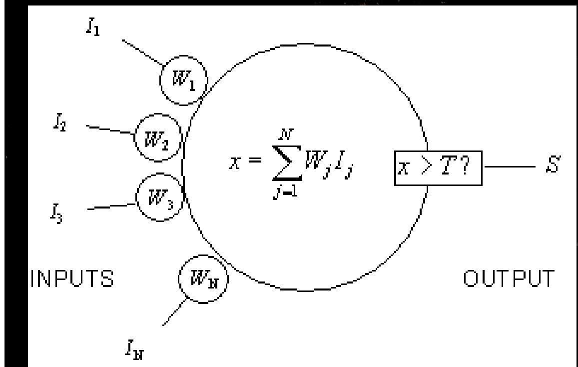
[NRL, 2013]

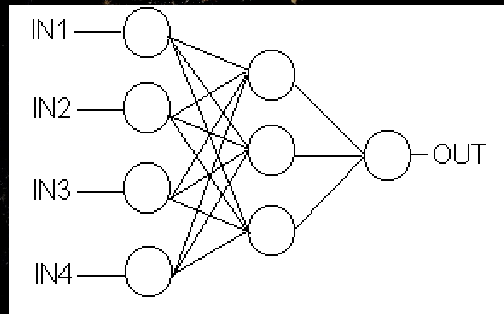
Also, HF comms for civil aviation, spacewalk planning on ISS, etc.

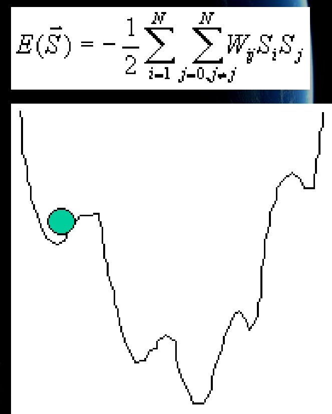
- Pessimist: "when you saw an ionospheric storm, you have seen one storm"
  - a.k.a. "No storms are alike"
- Realist:
  - Yes, the prior ionospheric state does not inform the future state
    - (No inertia… quick reply to drivers)
    - Need the Sun-Earth activity context
  - But individual weather processes admit their

Neural Doctrine: Galloping Introduction





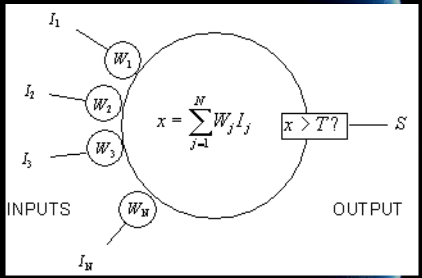


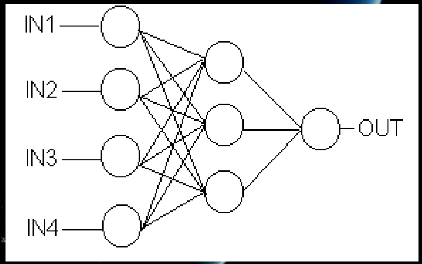


The roaring 1990s: an outburst of algorithmic NNs to replicate human intelligence

Historical Average © Feed-forward NN

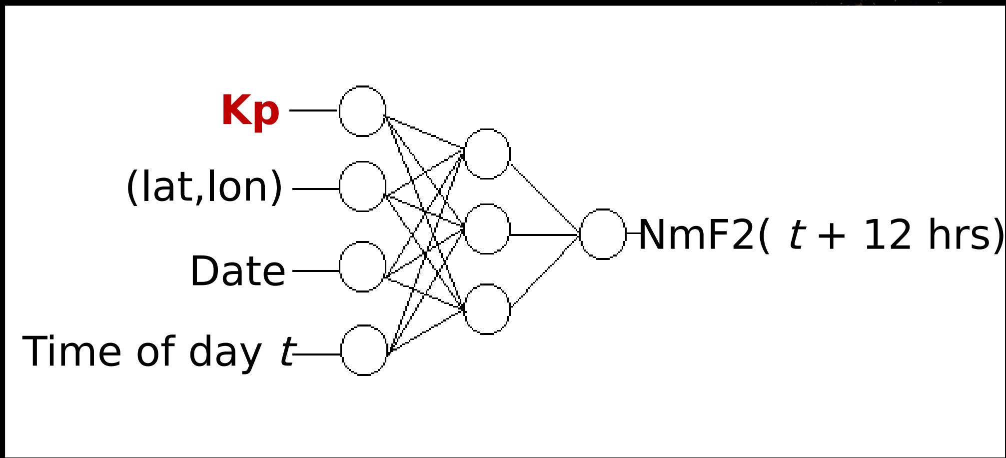
- Training phase
  - present NN with known examples (input and output) for training
    - determining the W_{ij} weights – back-propagation method
- Execution phase
  - WHAT-IF: present trained NN with previously unknown inputs to obtain a predicted output
  - Superior inductive bias of NNs: the capability of gleaning the nature of the system in order to do good WHAT-IFs.
  - Superior but little understood
    - Black-Box: No clue how and why it works well
    - Caused a severe AI Winter in the 2000s
      - NSF would not fund NN projects
      - Physics journals would not publish NN model results
- All feed-forward NN architectures are in "historical analogies" category
  - Subject to *AI Winter*



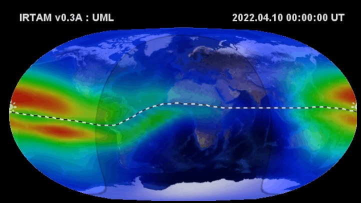


Feed-forward NN for forecasting ionosphere

FORECASTING NN MODEL FOR 12 HOURS AHEAD



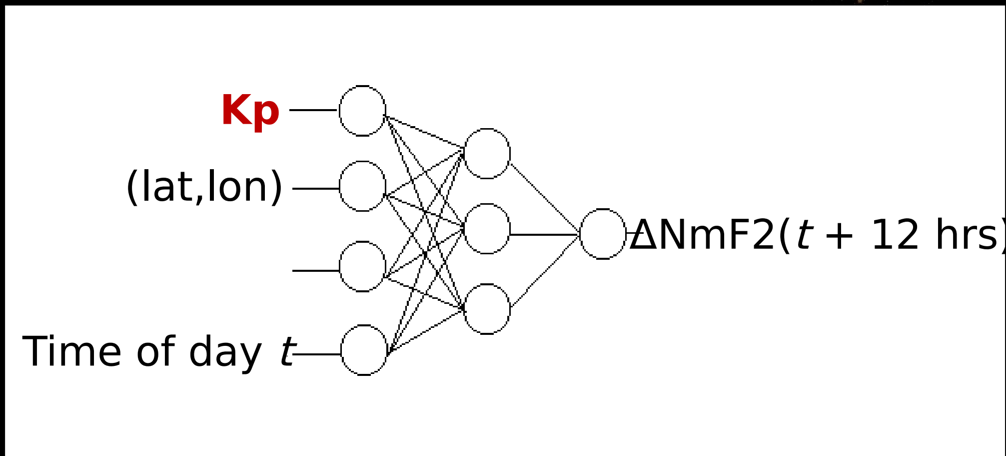
Space Physics community: no-no, this is a SNAKE OIL



- Train a NN to predict peak density in the ionosphere N_mF2 12 hours ahead, as a function of:
  - Time of day
  - Date (year, day of year)
  - Location (lat, lon)
  - Geomag index Kp
- WHAT-IF: run for different Kp values, dates, times, and locations

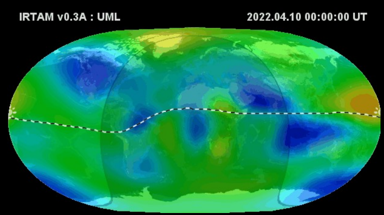
Feed-forward NN for forecasting ionosphere

FORECASTING NN MODEL FOR 12 HOURS AHEAD



- Train a NN to predict *deviation* of N_mF2 from the expected quiet-time behavior 12 hours ahead, as a function of:
  - Time of day
  - Location (lat, lon)
  - Geomag index Kp
- Run it for different Kp values and locations (what-if)
  - Obtain 2D map of ΔNmF2
  - Apply Δ to quiet-time predicted 2D map of NmF2

Space Physics community: still not good



"Storm" option of NmF2 in IRI

FORECASTING MODEL FOR UP TO 24 HOURS AHEAD

$$\{\mathbf{Ap}\} \longrightarrow \Delta\mathrm{NmF2}(\mathrm{lat},\mathrm{lon},t_{fore})$$

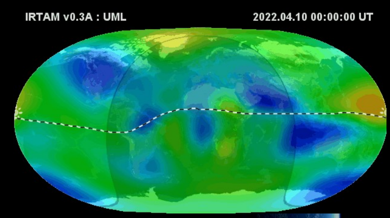
$$t_{fore} = (t,\ t + 24\ \mathrm{hrs})$$

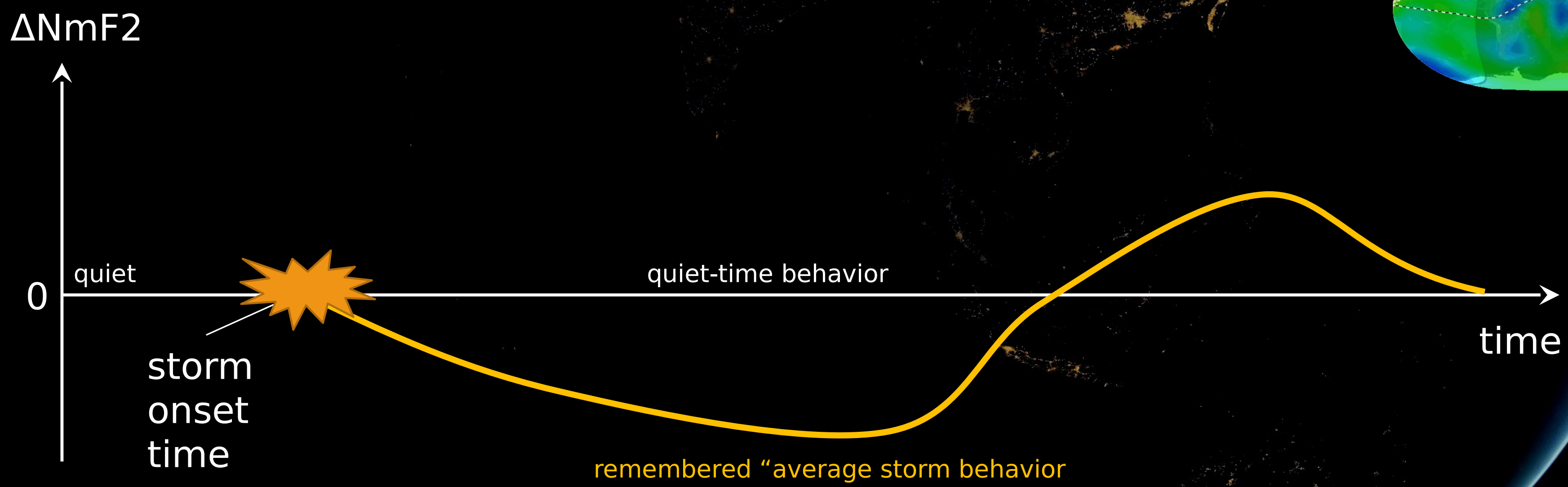
Research funding agencies:
but this is an empirical model… oh well

- "Storm" option for NmF2 forecast in IRI, [Fuller-Rawell et al, 1999]
  - Ap is tested for a threshold value to determine if the day is **quiet** or **disturbed**
  - This is an "average" **storm behavior** of ionosphere on disturbed days
  - The storm behavior is stored as ΔNmF2 for any location and forecast time up to 24 hours ahead
- Other "storm" options are pursued based on this principle
  - Blanch and Altadill [2012]

International Workshop on Machine Learning for Space Weather:

Concept of the “Triggered” storm option



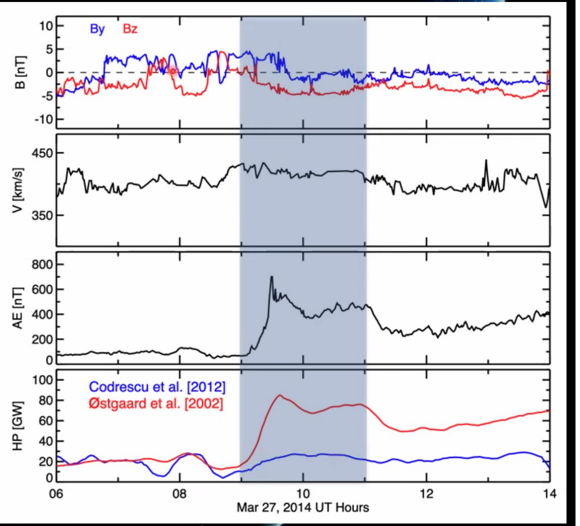


Forecast by analogy to an average “anomaly” timeline

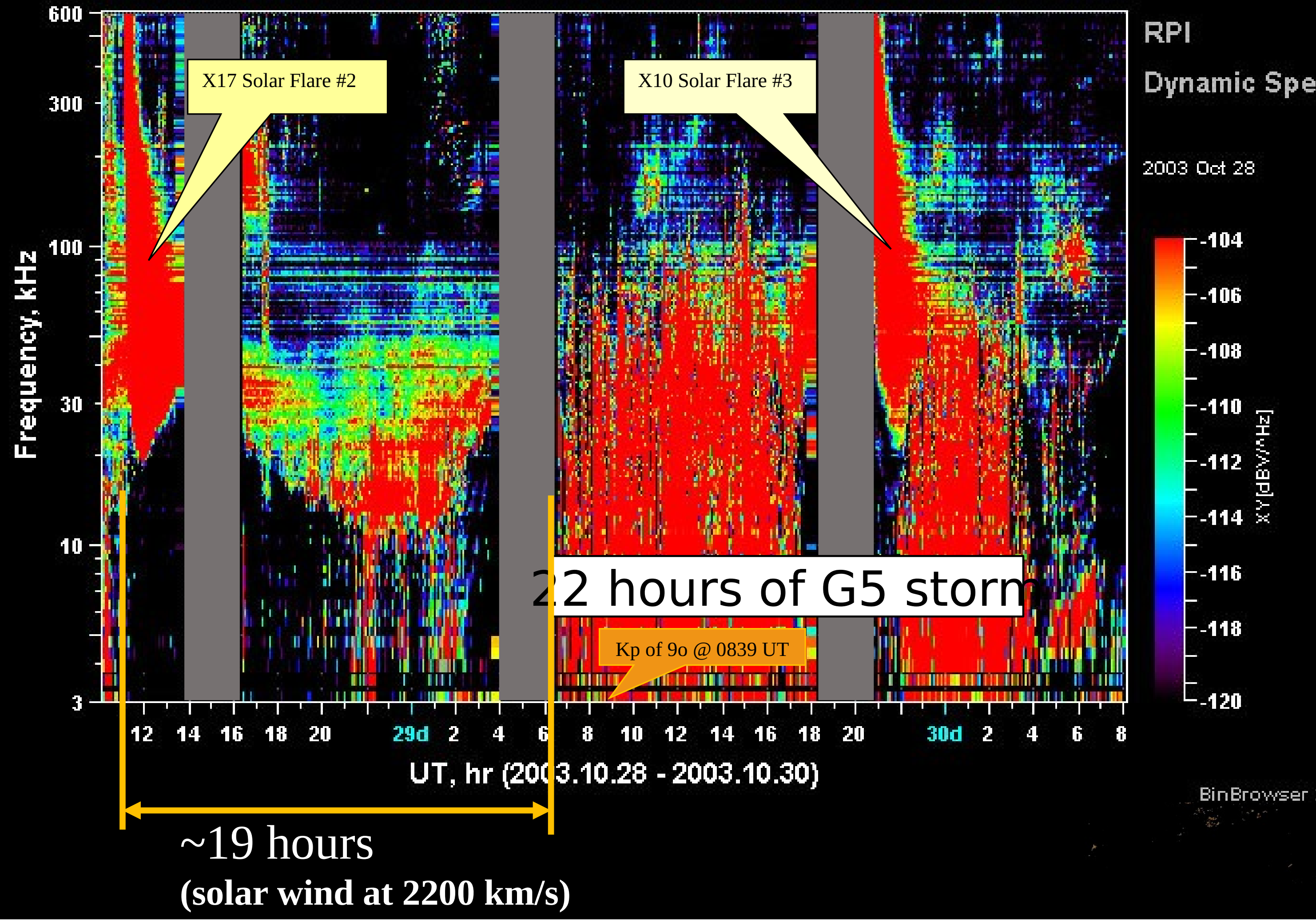
Next step: Library of the storm storylines

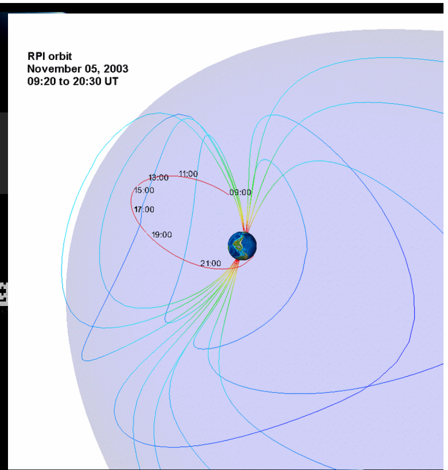
instead of the “average storm timeline”

- Instead of an “average” storm, keep a library of previous storm storylines of ΔNmF2
  - To forecast, just find the most relevant storm in the library
- Each storyline must be remembered in the context of the activity in the Sun-Earth environment
  - i.e., not just replay a storm line using one “trigger”
  - Need to build a grand storyline of driving events in the heliospace and geospace
  - Assumption: if we know all driver stories well, the ionosphere can be explained
    - This is a strong assumption
- Need good ideas for
  - The storm library
  - Search-and-retrieval algorithms
  - *Tweaking the library copy to current conditions*



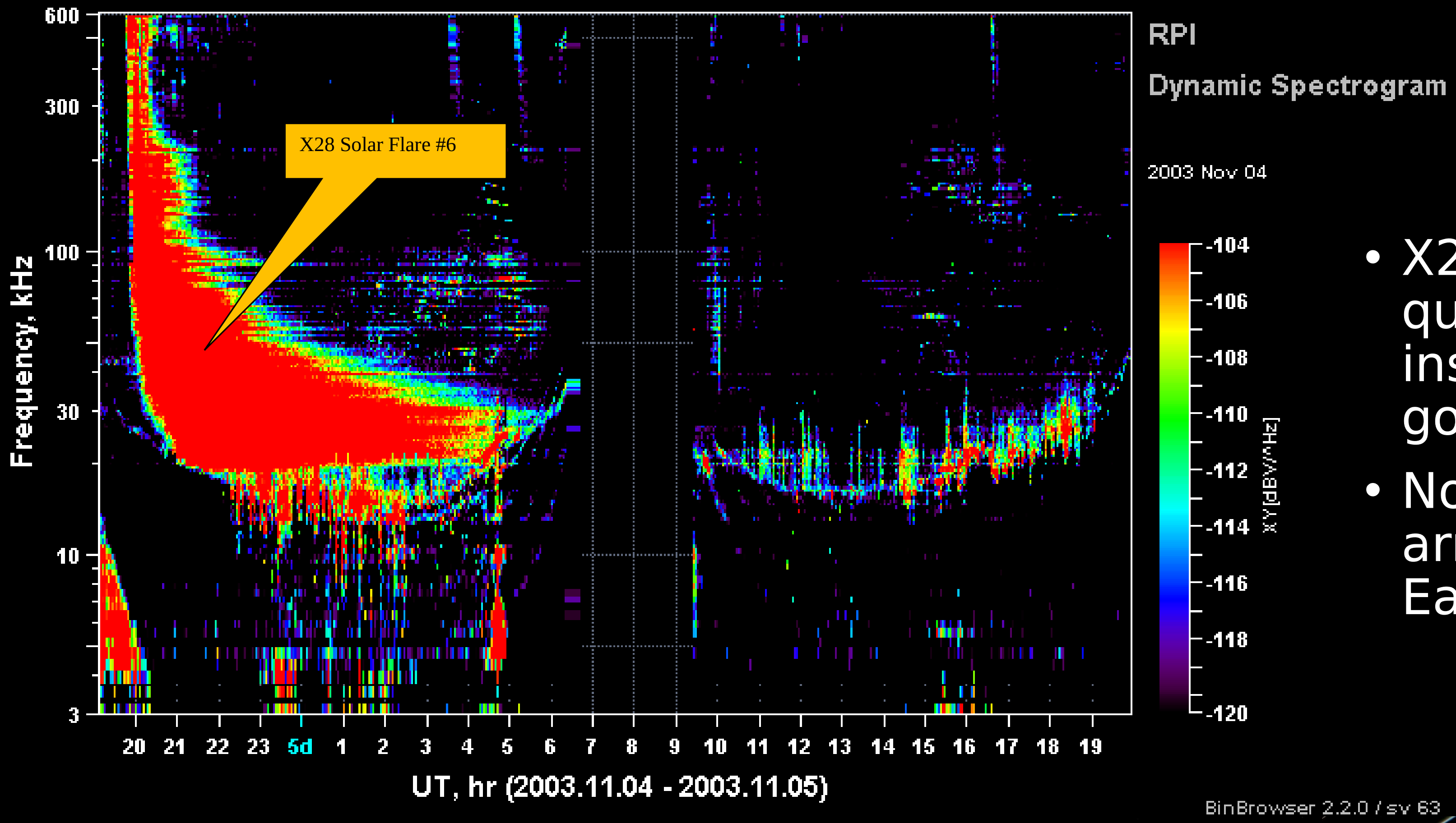
Example: Halloween storm as seen by RPI





- Solar wind velocity can vary
- Timelines will vary as well
- But… natural language AI will come to the rescue!
    - varying speed of words

Unrelated example: mega-flare 6 days later

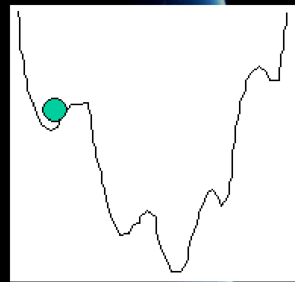


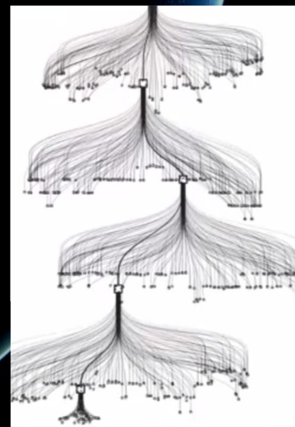
- X28 ranking is questioned: instruments got blinded!!
- Nothing arrived to the Earth, though

International Workshop on Machine Learning for Space Weather:

Alexa, play Yesterday by The Beatles

- "Yesterday" is interpreted in the context of "play"
  - Not a reference to one day earlier
  - A title to be fetched from the database of song titles
- DEEP LEARNING: multi-layered recurrent (feed-back) network topologies
  - Support interpretation of subelements in the context of other cues
    - Starting position of NN (the green ball) is determined from the context
    - Network evolves into the closest stable condition (remembered state)
    - That state propagates to the next layer of the network
  - Appears matching to the idea of interpreting ionospheric dynamics in the context of the external forces acting on it
    - Context: reports of ongoing Sun-Earth activity
    - Output: ionospheric dynamics fetched from the historical record database
    - What is different? Deep Learning the interplay of helio- and geo-activity markers





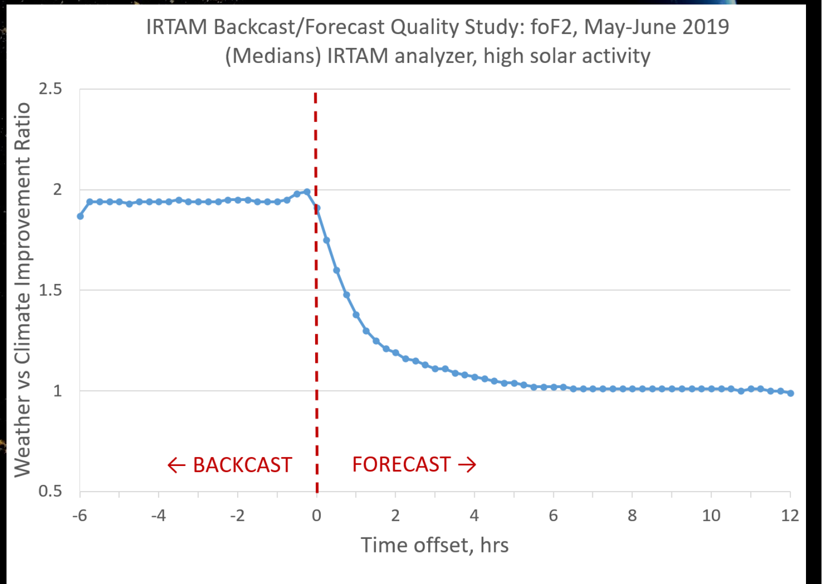
Natural Language AI for space weather?

- Detect “Alexa!”
  - Recognition of the storm onset
    - Solar flare?.. signature of CME?.. Solar wind pressure?.. lots of ideas!
    - Maybe all of the markers must be used to determine reference time
- Then, somehow, interpret the available “Play Yesterday by The Beatles”
  - Extract **context cues** to retrieve the best-matching storyline in the Storm Library
  - Context of the sentence == Context of the relevant system driver storylines
  - Combination of context cues and their relative strength = Deep Learning
- Retrieve and *process* the closest storm storyline from the library
  - Process? Encoding is needed to avoid varying timing of the processes
    - [to support varying speed of word pronunciation]
- Apply the processed storyline to forecast the upcoming departure of the stormy weather from the quiet-time model
- REPEAT

International Workshop on Machine Learning for Space Weather:

Capturing Context of Ionospheric Dynamics

- Ionosphere: immediate response to external forcing
  - Thus its current conditions do not inform future states
- Need to use storylines of all external drivers as context
  - Cannot be just one instant “triggering” driver (e.g., Kp=6)
  - Driver dynamics is matched (paired) to the ionospheric storm dynamics
    - Across the complete forecast storyline from onset to end
  - Important: which driver is relevant out of the set? (Deep Learning helps; *inductive bias*)

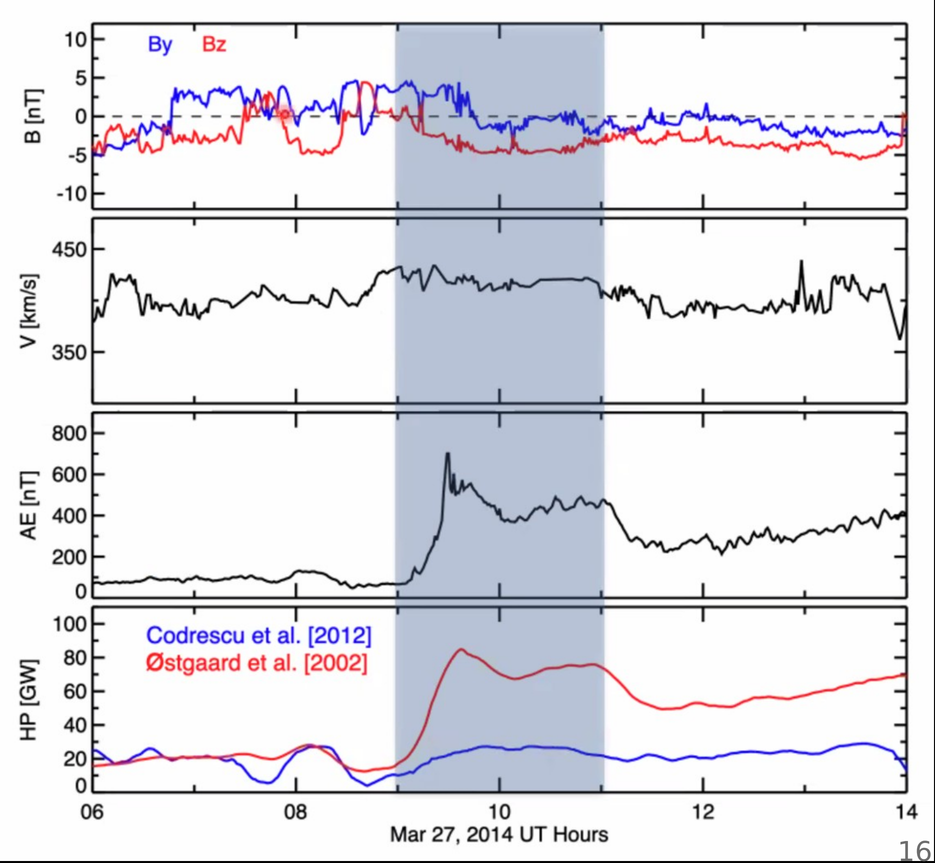


Sensor data @ 2 hour latency are useless

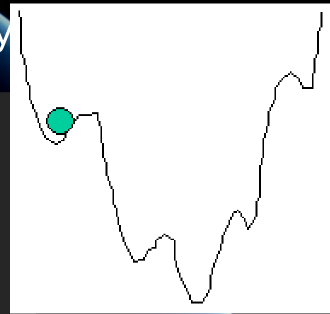
high solar activity, mostly quiet time

Why “REPEAT” step is needed?

- Ionospheric response to the storm-time impacts is not just a “triggered option for a disturbed plasma day”
  - Context of the disturbed ionosphere dynamics is a continuous function of *t*
    - Driver storylines need to be complete to retrieve the best matching storm in the library
    - But in the forecast scenario, only an initial fragment of the storylines may be available
- Forecast shall be repeated as time progresses and larger fragments of the storylines become available

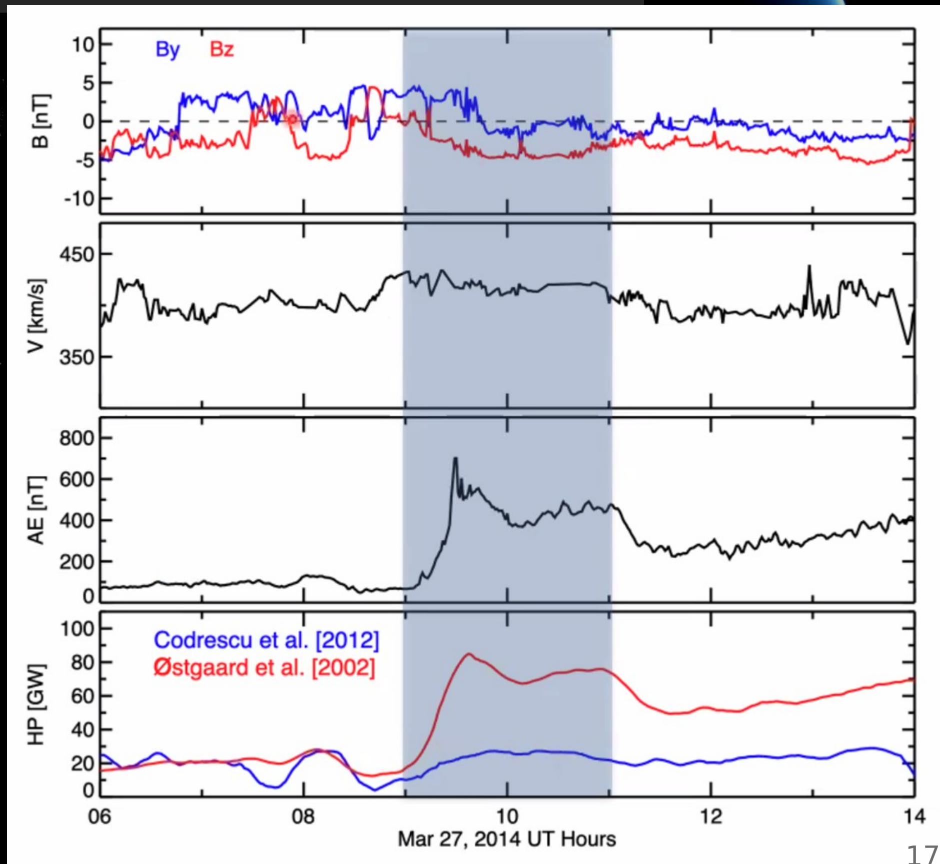


Associative memory



How to get storyline from a fragment

- Note: not the storyline of the storm, but of the storm *drivers*
  - Simpler task… *divide and conquer*
  - It is the interplay of drivers that matters
- Associative Memory is one possibility
  - Used in recognition of handwriting
  - Also for recalling stored data from their noisy and incomplete realizations
- Recursive, feed-back NN architecture
  - Hopfield networks
  - No input layer, no output layer
  - Neurons are clipped to available data and evolve into the nearest local minimum of E



PITHIA-NRF and T-FORS: European SW

- Real-time data for forecasting by historic analogy are not easy to come about
  - Need a consortium of real-time data providers
- PITHIA-NRF is an emerging space physics data infrastructure in Europe
    - Look it up! www.pithia-nrf.eu
    - HORIZON 2020 project
    - Based on EGI Foundation mega-facility of computing resources
      - Public funding = better prospects of longevity
    - And a Network of Research Facilities (BRF)
      - some facilities have decades of uninterrupted operation
  - T-FORS is the pilot project to leverage PITHIA-NRF collections
    - TID Forecasting System
    - HORIZON 2020 project
    - Listen to Elvira Astafyeva talk later this morning (TID)

The EU Framework Programme for Research and Innovation
HORIZON 2020

PITHIA-NRF Research Infrastructure

egi
www.egi.eu

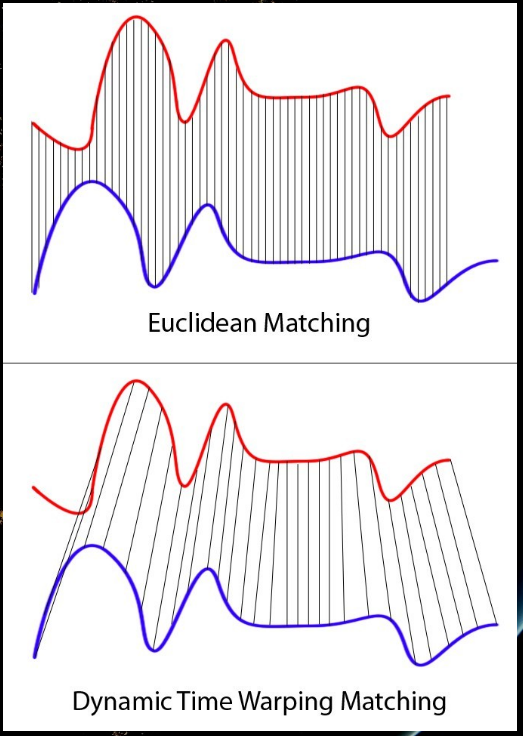
UMASS LOWELL

Forecast Architecture

REPEAT: forecast quality improves with each iteration as longer storylines are coming in

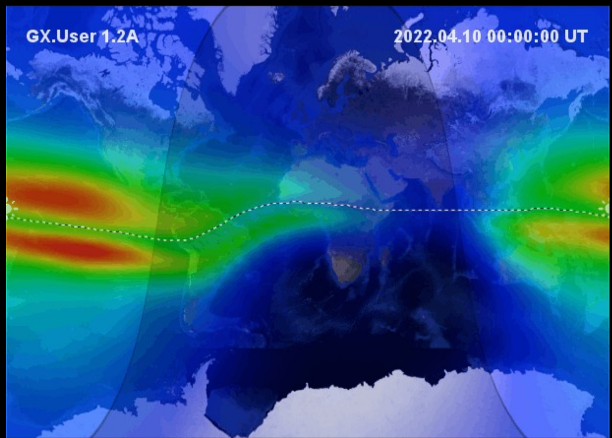
Dynamic Time Warping (DTW)

- Warp library-provided storm storyline
  - DTW finds similarity between 2 storylines
  - Driver storylines may be indicative of *how different* the actual storm timing is from the Library copy
  - Corresponding time warping shall be applied to the correction ΔNmF2

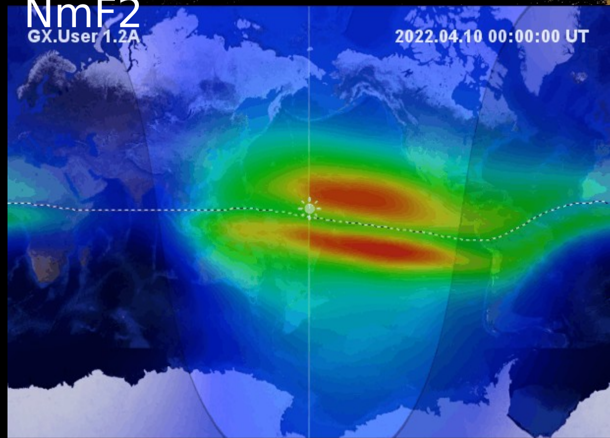


Library of storms in longitude-neutral form

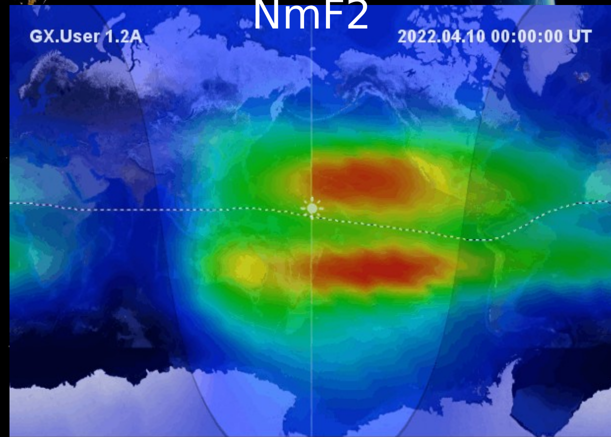
“Classic” view of NmF2

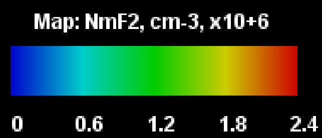


Local Time Noon view of NmF2

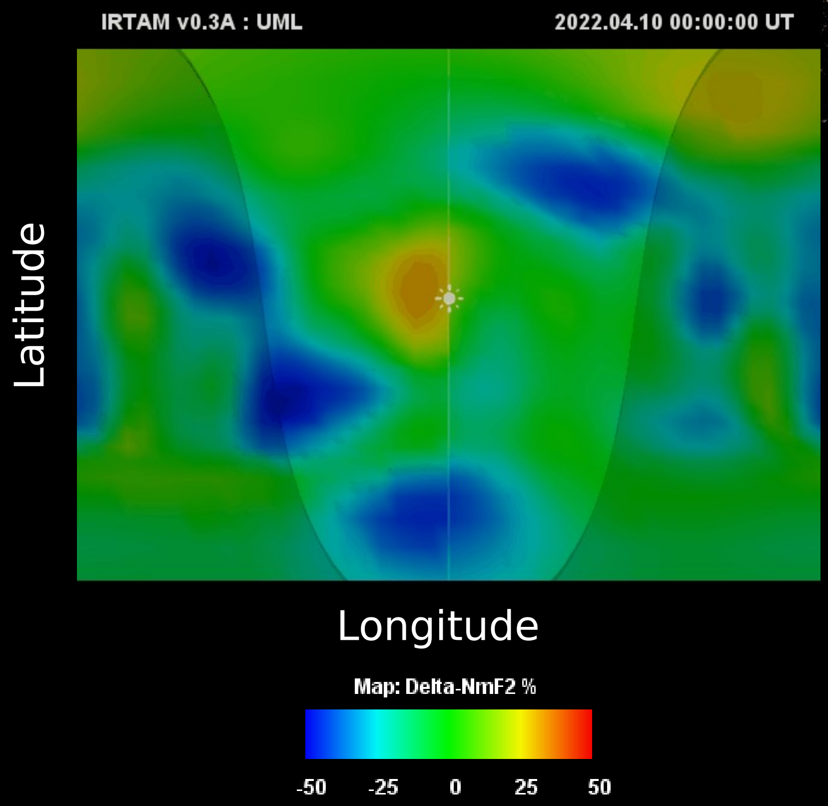


De-magnetized view of NmF2

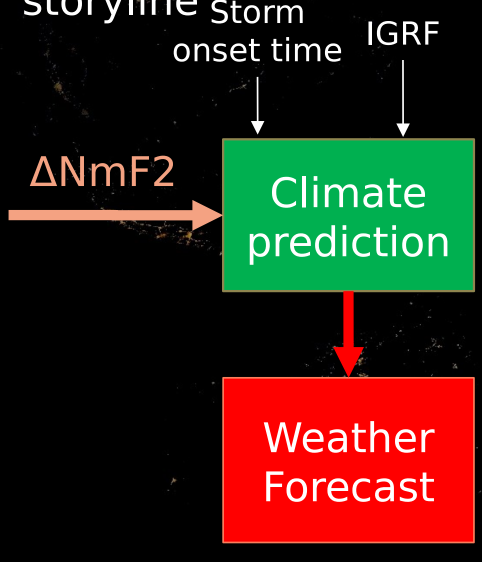




Encoding library storylines of ΔNmF2



- Total 1024 coefficients to store 24 hour global animated 2D maps
- Can be expanded to 2048 coefficients to store 2-day storyline



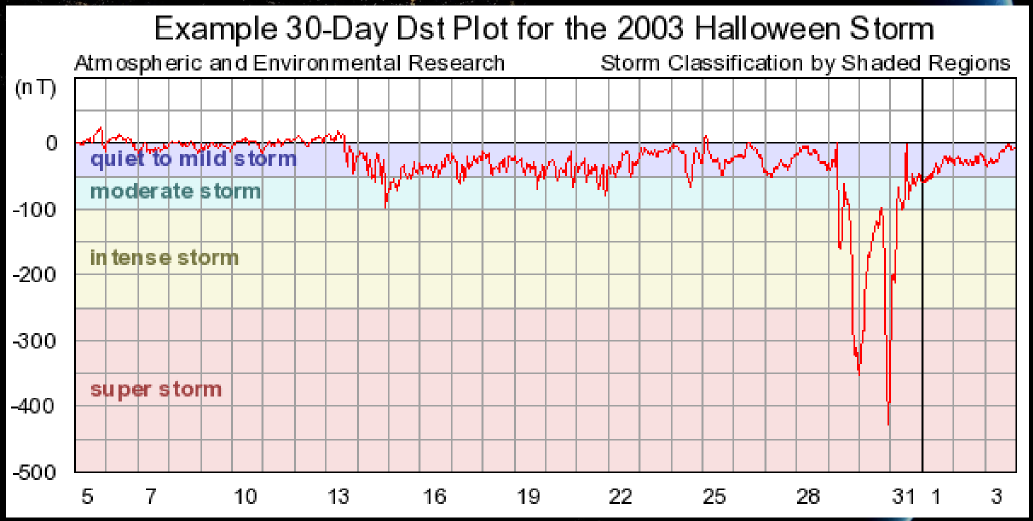
The IGRF model of the Earth magnetic field is needed to re-magnetize the retrieved library

Forecasting by context-driven memory

- NOT to build a least-square regression on 1024 unknowns
- NOT to build a back-prop feed-forward NN with 1024 outputs
- Just memorize them, *cleverly*
  - Associate the timeline of ionospheric dynamics with timelines of ionospheric state drivers
    - Deep Learning: placing the storm vocabulary into the context of a “sentence” of ongoing geospace activity
    - Rely on NN superior inductive bias to build the context
    - Plus other tricks:
      - Dynamic Time Warping (DTW)
      - Associative memory (AM)
        - Restore a driver’s full storm timeline from its initial observed fragment

Natural Language Processing as DTW example

- Analogous to Sound/Syllable recognition
- Custom Language to describe storm progression



Summary

- Deep Learning "Ice-Break" is ongoing in NN-based forecasting
  - DL learns the system from its previous behavior
- A concept study of DL-based forecast of the ionospheric storm storylines:
  - Forecast deviation timeline of the disturbed ionosphere
    - Deviation from the quiet-time LT-centered/demagnetized ionosphere
  - Sync the deviation timeline to the actual/definitive storm onset time (Alexa!)
  - Use Dynamic Time Warping to maintain a smaller vocabulary of the storm behavior
  - Deep Learning to describe ionosphere timeline in the context of key storm driver timelines
  - For each activity driver, use associative memories to retrieve a full-length storyline from the initially observed fragment
- Procedure:
  - Detect storm onset, obtain full-length driver storylines
  - Take 30-day median current ionosphere, LT-center, de-magnetize,
  - Retrieve deviation storyline from the storm library, time-warp to current activity
  - Apply deviation to the median, position at reference LT, re-magnetize.