ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

Ensemble prediction with OpenIFS

Pirkka Ollinaho
Academy of Finland Postdoctoral Researcher

With thanks to:
Glenn Carver (ECMWF),
Simon Lang (ECMWF),
Lauri Tuppi (UoH),
Heikki Järvinen (UoH)

6/19/19

Why?

Why?

- Ensemble forecast are useful

Why?

- Ensemble forecast are useful
- They provide...
  - an estimation of the forecast uncertainty
  - efficient means to study impact of initial state
  - a way to gather knowledge about model physics and about closure parameters therein

What?

What?

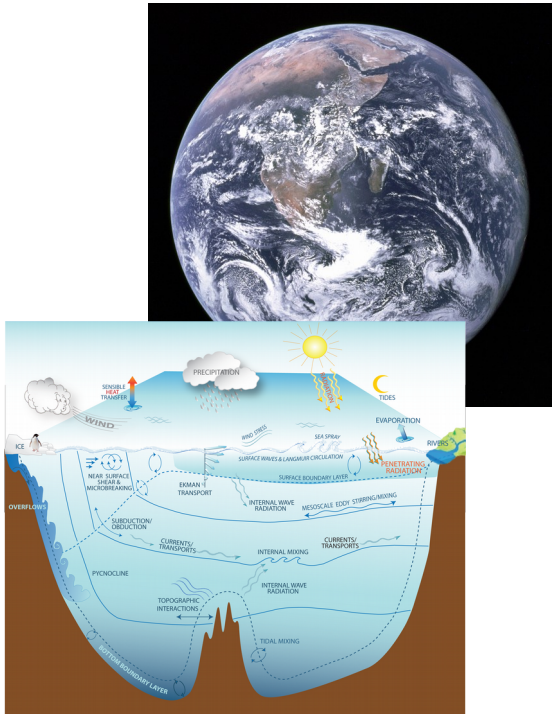
- Future state of the atmosphere is defined by

What?

- Future state of the atmosphere is defined by
  - Its current state

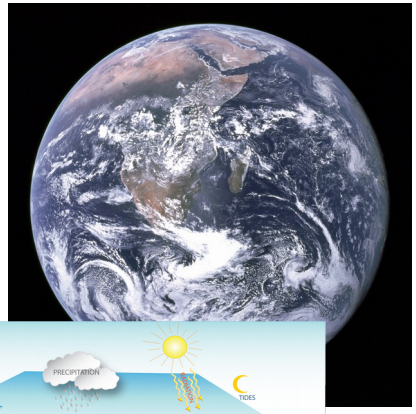
What?

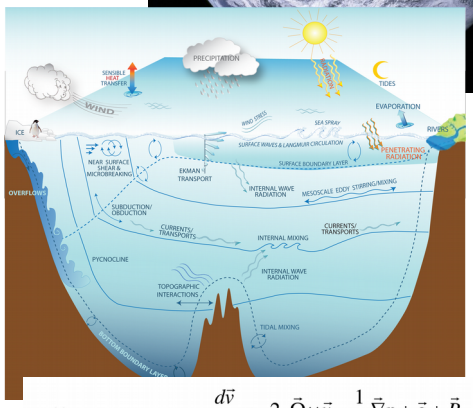
- Future state of the atmosphere is defined by
  - Its current state
  - Boundary interactions with the ocean, land surface, etc.



What?

- Future state of the atmosphere is defined by
  - Its current state
  - Boundary interactions with the ocean, land surface, etc.
  - Governing/primitive equations





Momentum conservation $\frac{d\vec{v}}{dt} = -2\cdot\vec{\Omega}\times\vec{v} - \frac{1}{\rho}\vec{\nabla}p + \vec{g} + \vec{P}_v$

Energy conservation $\frac{dT}{dt} = \frac{R\cdot T\cdot\omega}{c_p p_s \sigma} + P_T$

Water vapour conservation $\frac{dq}{dt} = P_q$

Mass conservation $\frac{dp_s}{dt} = p_s\cdot\left(\vec{\nabla}\cdot\vec{v} + \frac{d}{d\sigma}\frac{d\sigma}{dt}\right)$

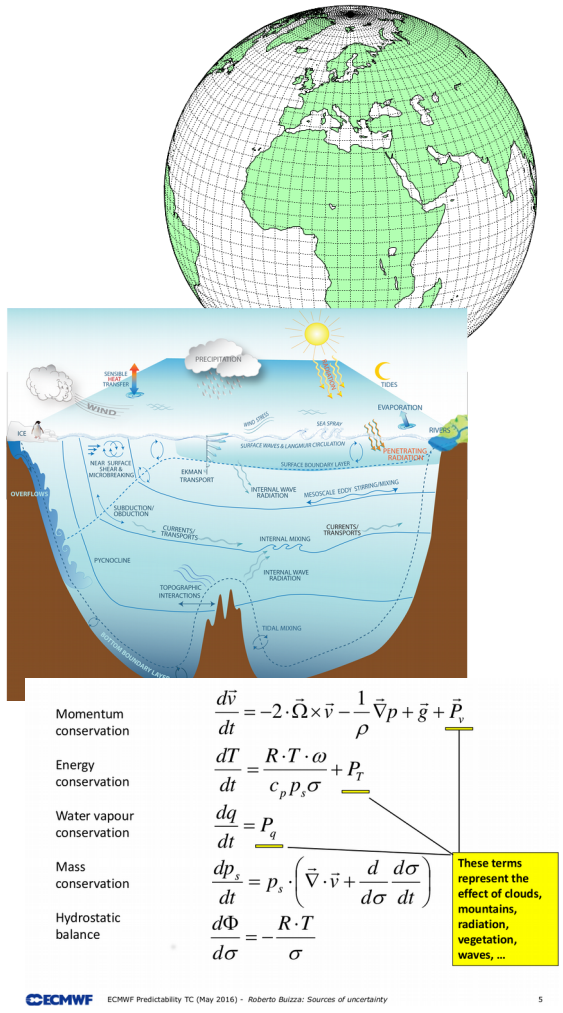
Hydrostatic balance $\frac{d\Phi}{d\sigma} = -\frac{R\cdot T}{\sigma}$

These terms represent the effect of clouds, mountains, radiation, vegetation, waves, …

What?

- Future state of the atmosphere is defined by
  - Its current state

    Not know accurately enough
  - Boundary interactions with the ocean, land surface, etc.
  - Governing/primitive equations

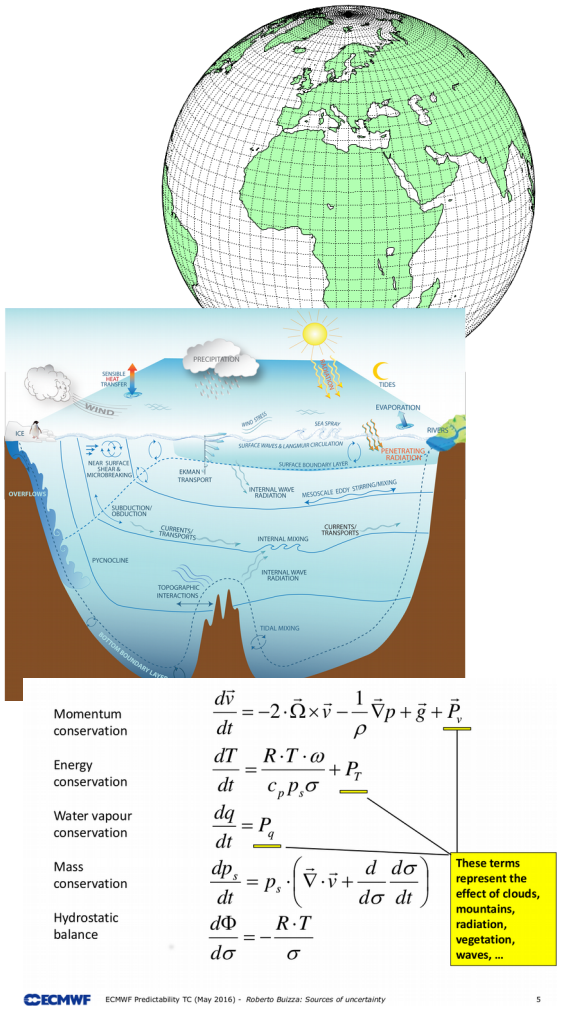


What?

- Future state of the atmosphere is defined by
  - Its current state

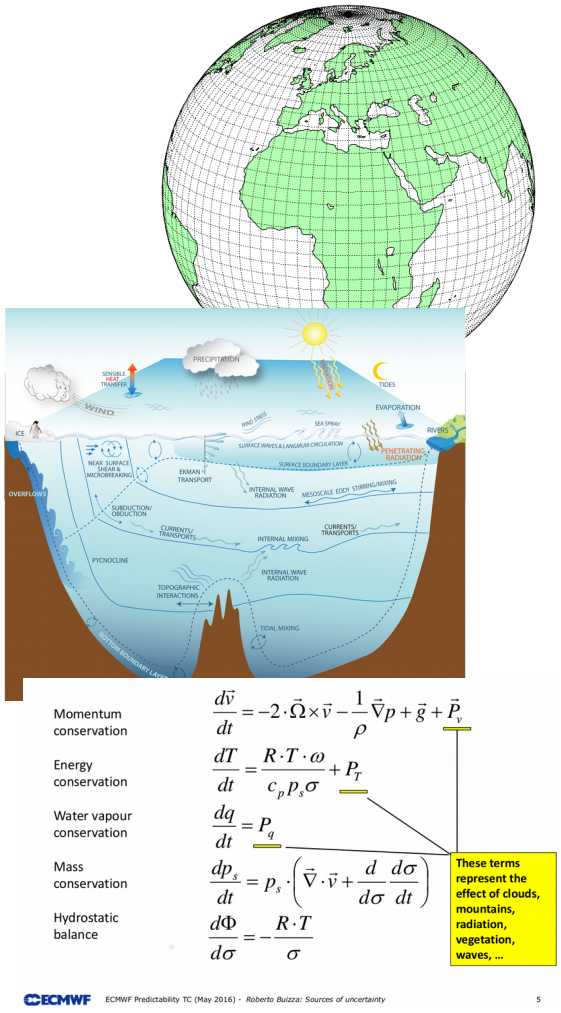
    Not know accurately enough
  - Boundary interactions with the ocean, land surface, etc.

    Lacking understanding to model precisely
  - Governing/primitive equations



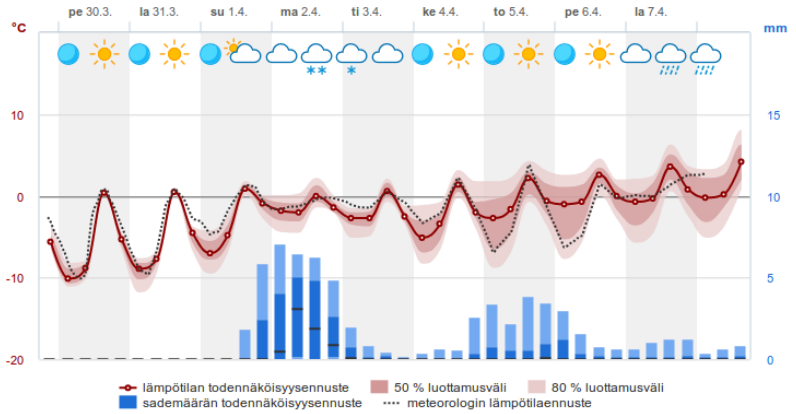
What?

- Future state of the atmosphere is defined by
  - Its current state
    - Not know accurately enough
  - Boundary interactions with the ocean, land surface, etc.
    - Lacking understanding to model precisely
  - Governing/primitive equations
    - No analytical solution –> need to discretize
    - Due to limited computing resources, the discretized series cannot be continued into infinity
    - Models can't thus directly describe the smallest scale phenomena –> parametrizations



What?

- Uncertainty in the predicted state of the atmosphere arises from
  - Its current state
    - Not know accurately enough
  - Boundary interactions with the ocean, land surface, etc.
    - Lacking understanding to model precisely
  - Governing/primitive equations
    - No analytical solution –> need to discretize
    - Due to limited computing resources, the discretized series cannot be continued into infinity
    - Models can't thus directly describe the smallest scale phenomena –> parametrizations



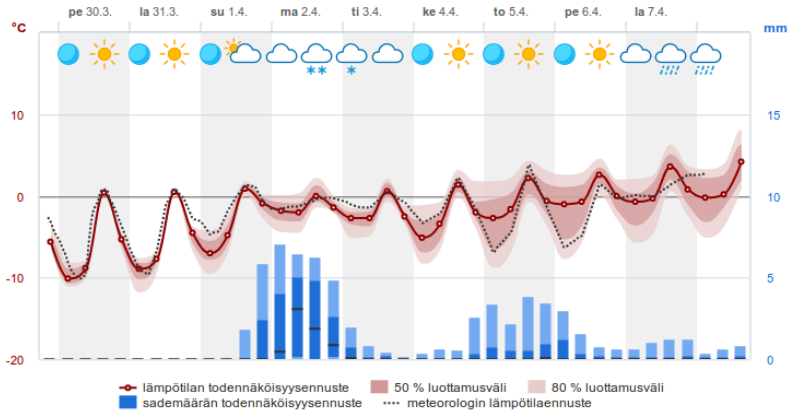
What?

- Uncertainty in the predicted state of the atmosphere arises from
  - Its current state

    Initial state uncertainty
  - Boundary interactions with the ocean, land surface, etc.

    Boundary uncertainty
  - Governing/primitive equations

    Model uncertainty



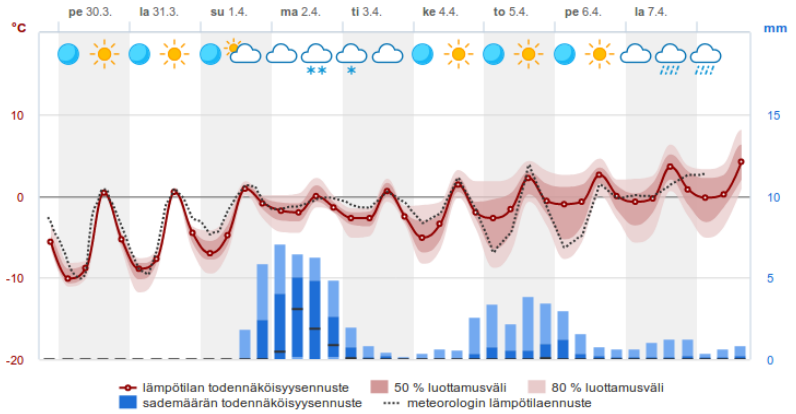
What?

- Uncertainty in the predicted state of the atmosphere arises from
  - Initial state uncertainty

    Initialize model from different atmospheric states
  - Boundary uncertainty

    Add (smart) perturbations into boundary interactions, or directly into ocean or land surface models
  - Model uncertainty

    Use different forecast models, physical parametrizations, or add (smart) perturbations into model equations



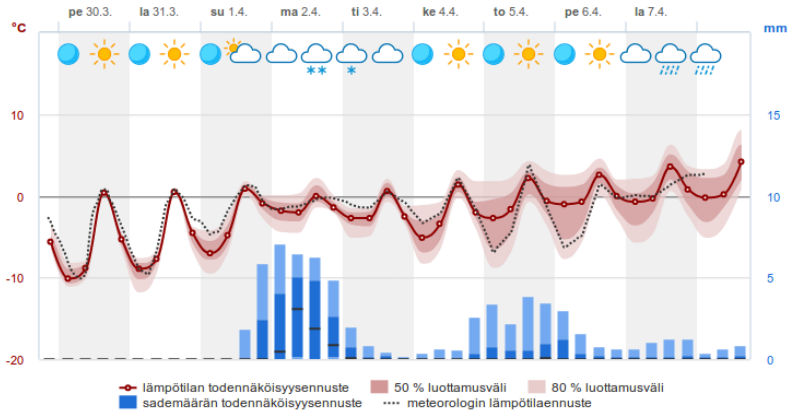
What?

- Uncertainty in the predicted state of the atmosphere arises from
  - Initial state uncertainty

    Initialize model from different atmospheric states
  - Boundary uncertainty

    Add (smart) perturbations into boundary interacations, or directly into ocean or land surface models
  - Model uncertainty

    Use different forecast models, physical parametrizations, or add (smart) perturbations into model equations



COMPLEMENT DETERMINISTIC MODELING!

How?

How?

- Not too many ways of generating ensembles available if you’re not working in an operational forecasting center

How?

- Not too many ways of generating ensembles available if you're not working in an operational forecasting center
- Multi-models?

How?

- Not too many ways of generating ensembles available if you're not working in an operational forecasting center
- Multi-models?
- Multi-physics?

How?

- Not too many ways of generating ensembles available if you're not working in an operational forecasting center
- Multi-models?
- Multi-physics?
- Stochastic physics?

How?

- Not too many ways of generating ensembles available if you're not working in an operational forecasting center
- Multi-models?
- Multi-physics?
- Stochastic physics?
- Stochastic closure parameters?

How?

- Not too many ways of generating ensembles available if you're not working in an operational forecasting center
- Multi-models?
- Multi-physics?
- Stochastic physics?
- Stochastic closure parameters?
- How about initial state perturbations from which ensemble predictions could be started?

Initial state perturbations (for OIFS)

- Replicate what is done operationally at ECMWF
  - IFS CY43R3
  - 50+1 initial states
  - Singular Vectors from the operational ensemble
  - Ensemble of Data Assimilations from the operational ensemble
  - TL159/TL399/TL639 (~120km/~50km/~32km)
  - Dec 2016 – Nov 2017, 00/12UTC

Ensemble workflow manager

- OpenEPS (github.com/pirkkao/OpenEPS)
  - Workflow manager written without any external software (in order to avoid extra hassle)
  - Bash + GNU make
  - Works similarly in HPC, linux cluster and laptop environments
  - NOT the point of this talk, any workflow manager (that gets the job done well) is suitable for using the initial perturbations

Benchmarking the system

- How good is the probabilistic skill of the system
- The experiment design:

Reso	Name (ens size)				
TL159	EDA+SV (N50)	EDA (N50)	SV (N50)	EDA*1.2 (N8)	SV*1.2 (N8)
TL399	EDA+SV (N50)	EDA (N20)	SV (N20)		
TL639	EDA+SV (N20)				

Benchmarking the system

- How good is the probabilistic skill of the system
- The experiment design:

Reso	Name (ens size)				
TL159	EDA+SV (N50)	EDA (N50)	SV (N50)	EDA*1.2 (N8)	SV*1.2 (N8)
TL399	EDA+SV (N50)	EDA (N20)	SV (N20)		
TL639	EDA+SV (N20)				

Benchmarking the system

- What skill score to use?
  - Continuous Ranked Probability Score (CRPS) widely used in benchmarking research experiments

Benchmarking the system

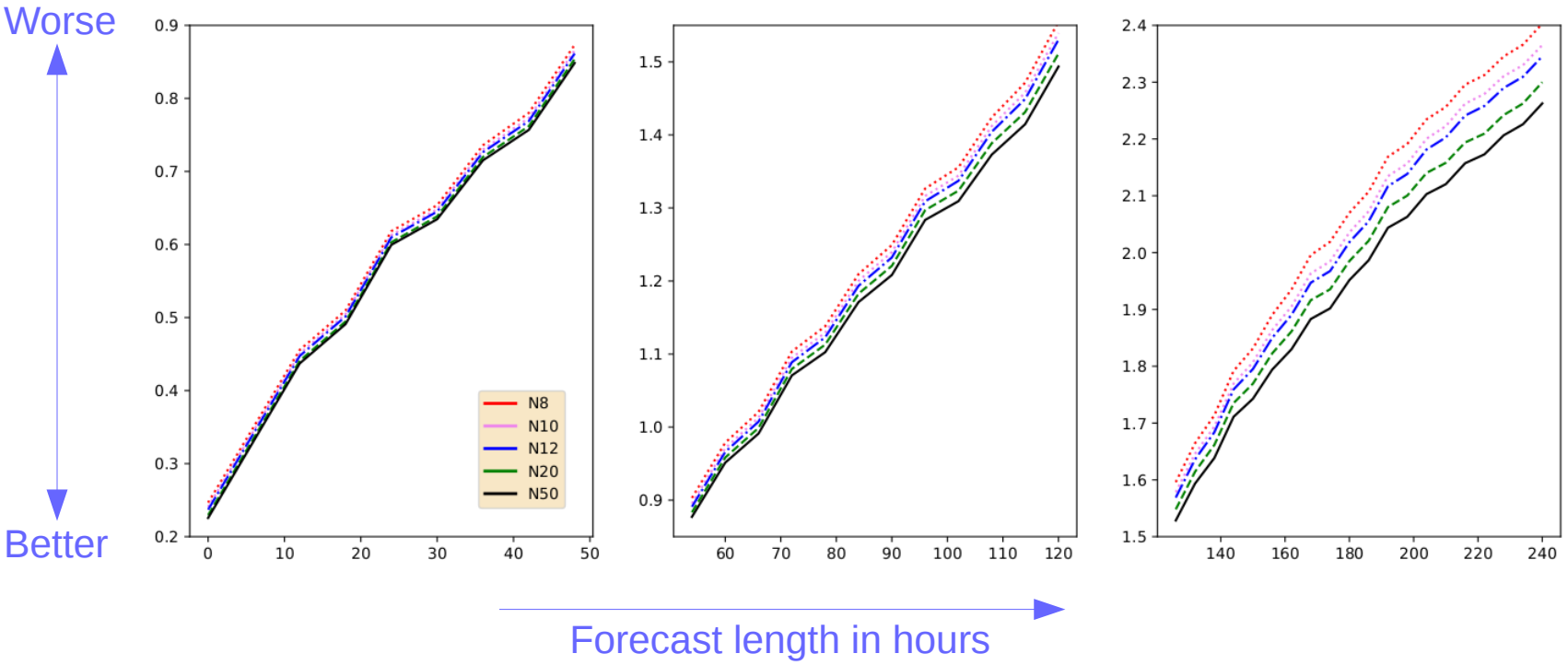
- What skill score to use?
  - Continuous Ranked Probability Score (CRPS) widely used in benchmarking research experiments
  - Dependent on the size of the ensemble
  - Limitations on ensemble size due to computational resource requirements

Benchmarking the system

- What skill score to use?
  - Continuous Ranked Probability Score (CRPS) widely used in benchmarking research experiments
  - Dependent on the size of the ensemble
  - Limitations on ensemble size due to computational resource requirements
- Leutbecher (2018) explored the topic of how a small ensemble size is suitable for finding out the probabilistic skill of a system
  - Fair-CRPS (FAIR) is a skill score independent of ensemble size, i.e. it treats different size ensembles the same way

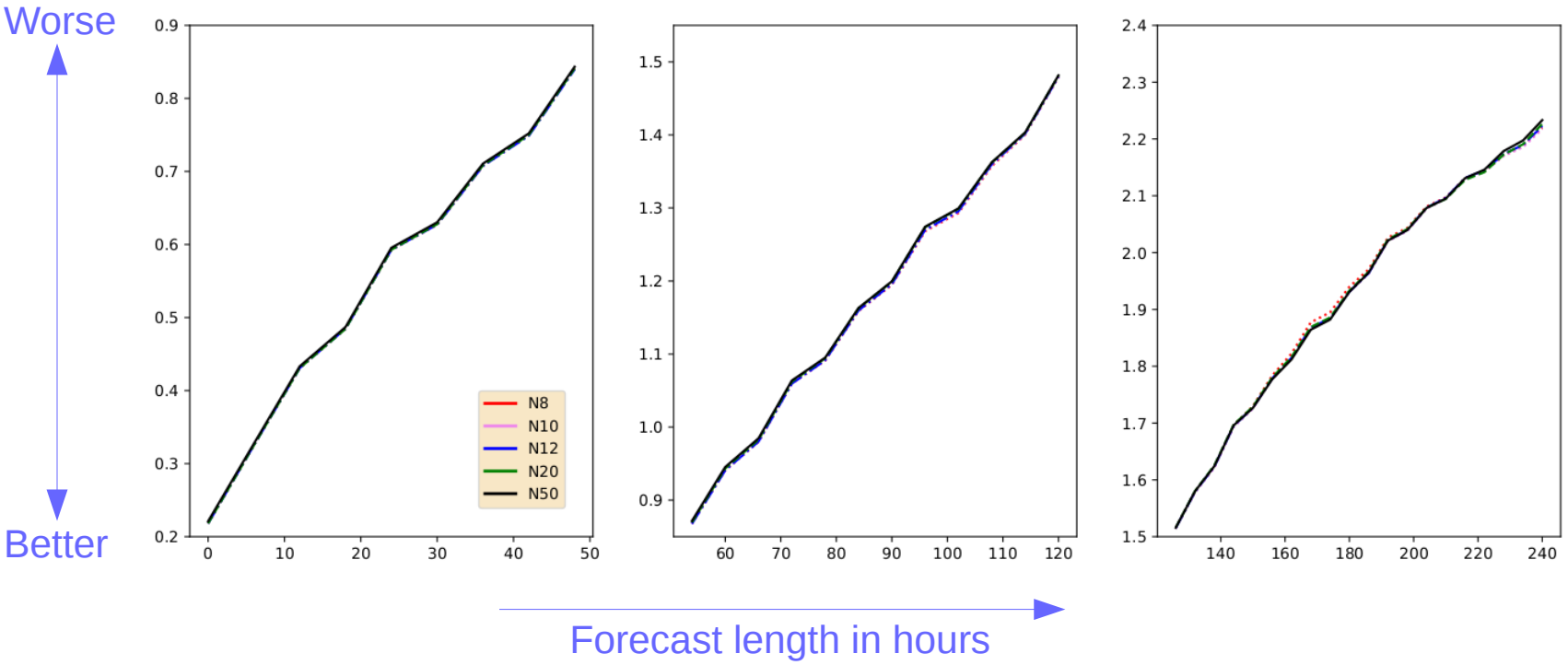
Benchmarking the system

CRPS and different ensemble sizes, TL159. CRPS of temperature at 850hPa in the Northern extra-tropics. Mean over 46 start dates.



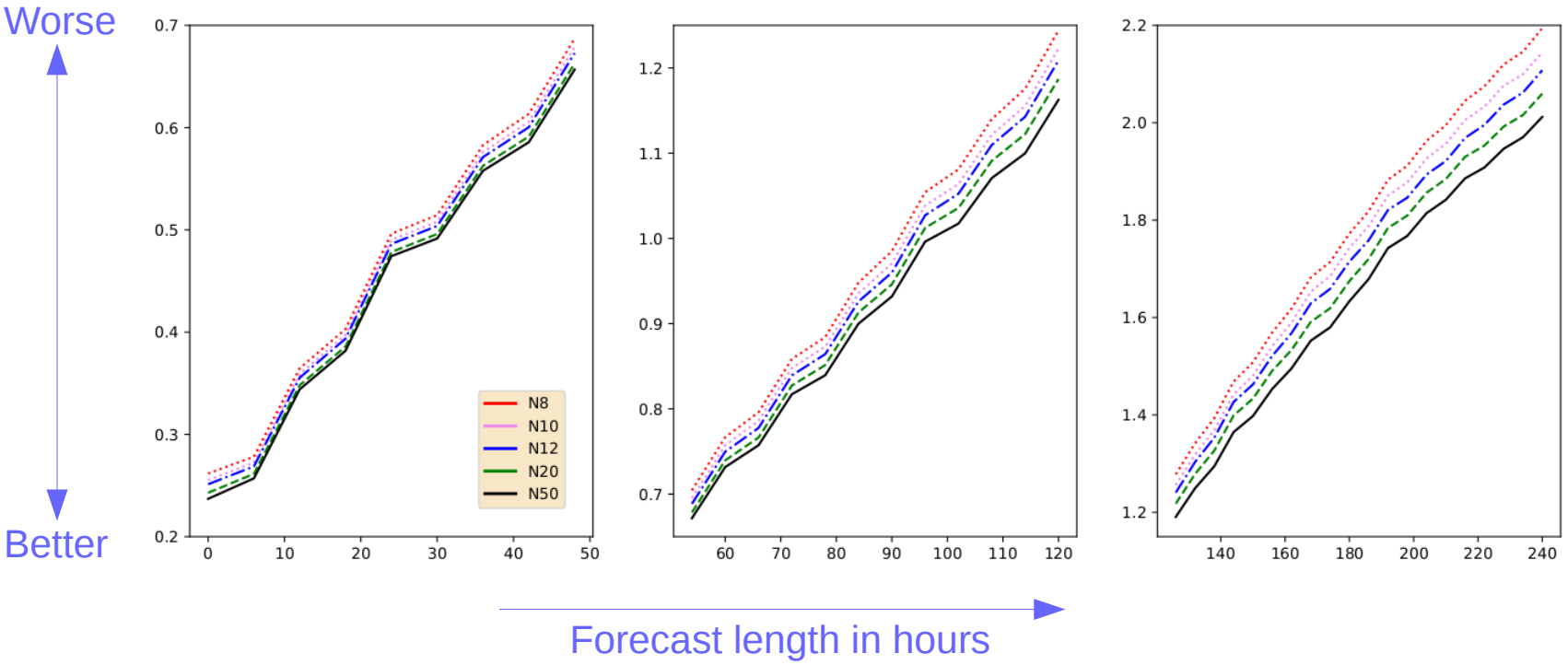
Benchmarking the system

FAIR and different ensemble sizes, TL159. FAIR of temperature at 850hPa in the Northern extra-tropics. Mean over 46 start dates.



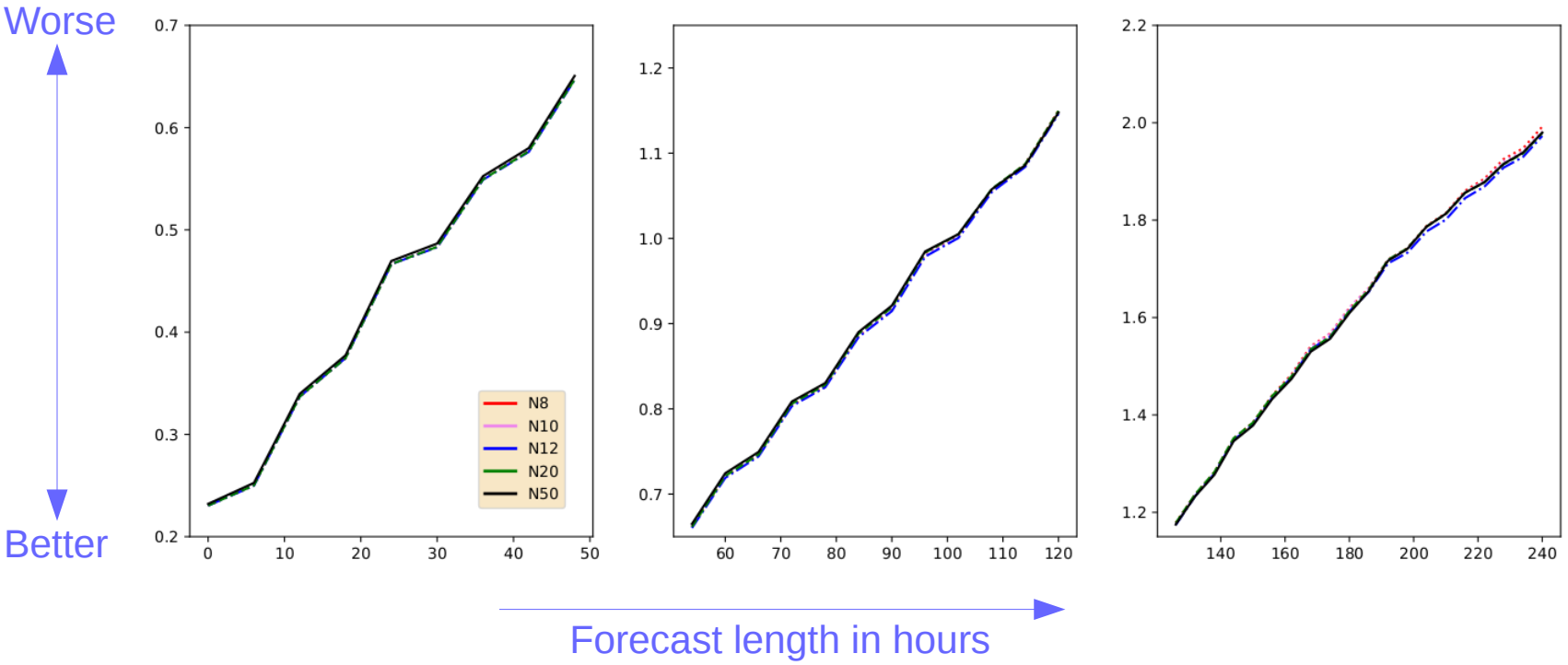
Benchmarking the system

CRPS and different ensemble sizes, TL399. CRPS of temperature at 850hPa in the Northern extra-tropics. Mean over 46 start dates.



Benchmarking the system

FAIR and different ensemble sizes, TL399. FAIR of temperature at 850hPa in the Northern extra-tropics. Mean over 46 start dates.



Benchmarking the system

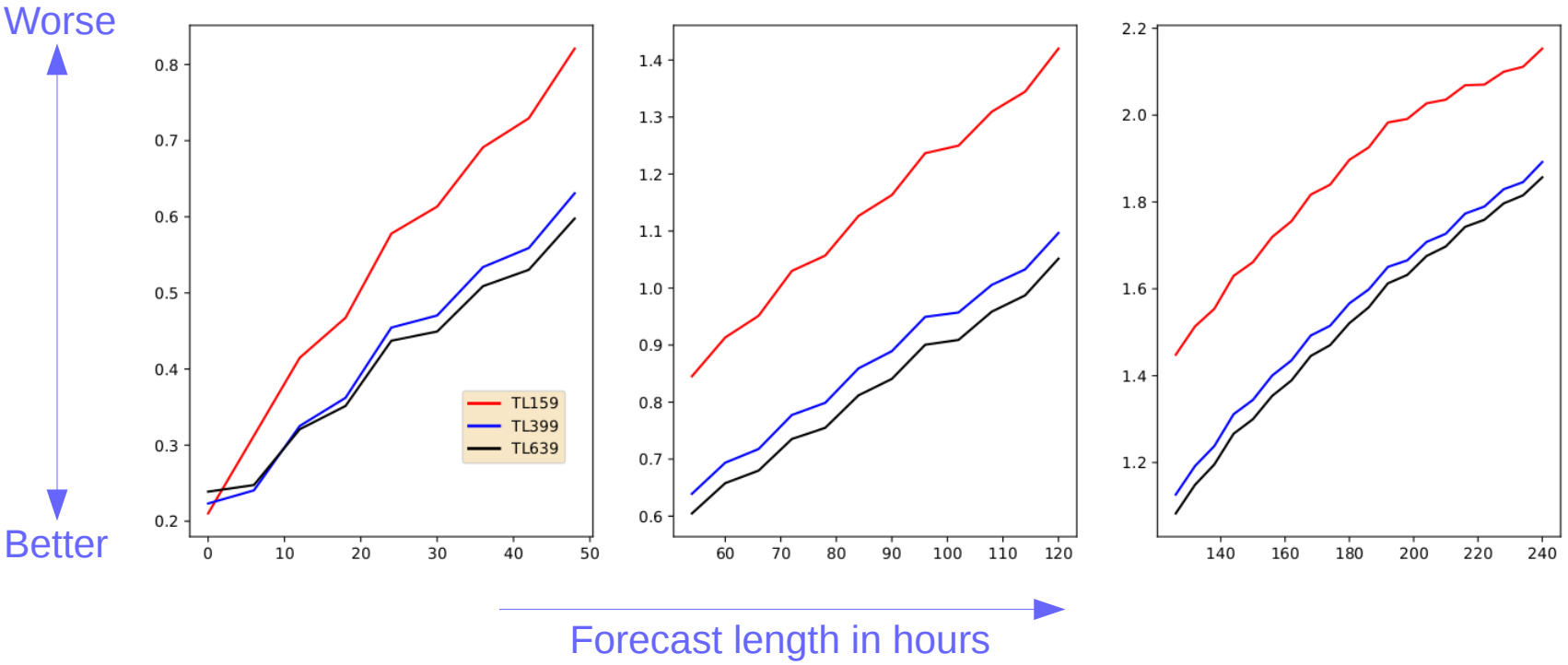
- FAIR seems to work as it should!
- So, lets answer a few questions:
  - How does model resolution affect the probabilistic skill of the system?
  - What are the contributions of SV and EDA perturbations on the skill?

Benchmarking the system

- FAIR seems to work as it should!
- So, lets answer a few questions:
  - How does model resolution affect the probabilistic skill of the system?
  - What are the contributions of SV and EDA perturbations on the skill?

Benchmarking the system

FAIR of temperature at 850hPa in the Northern extra-tropics. Mean over 27 start dates.

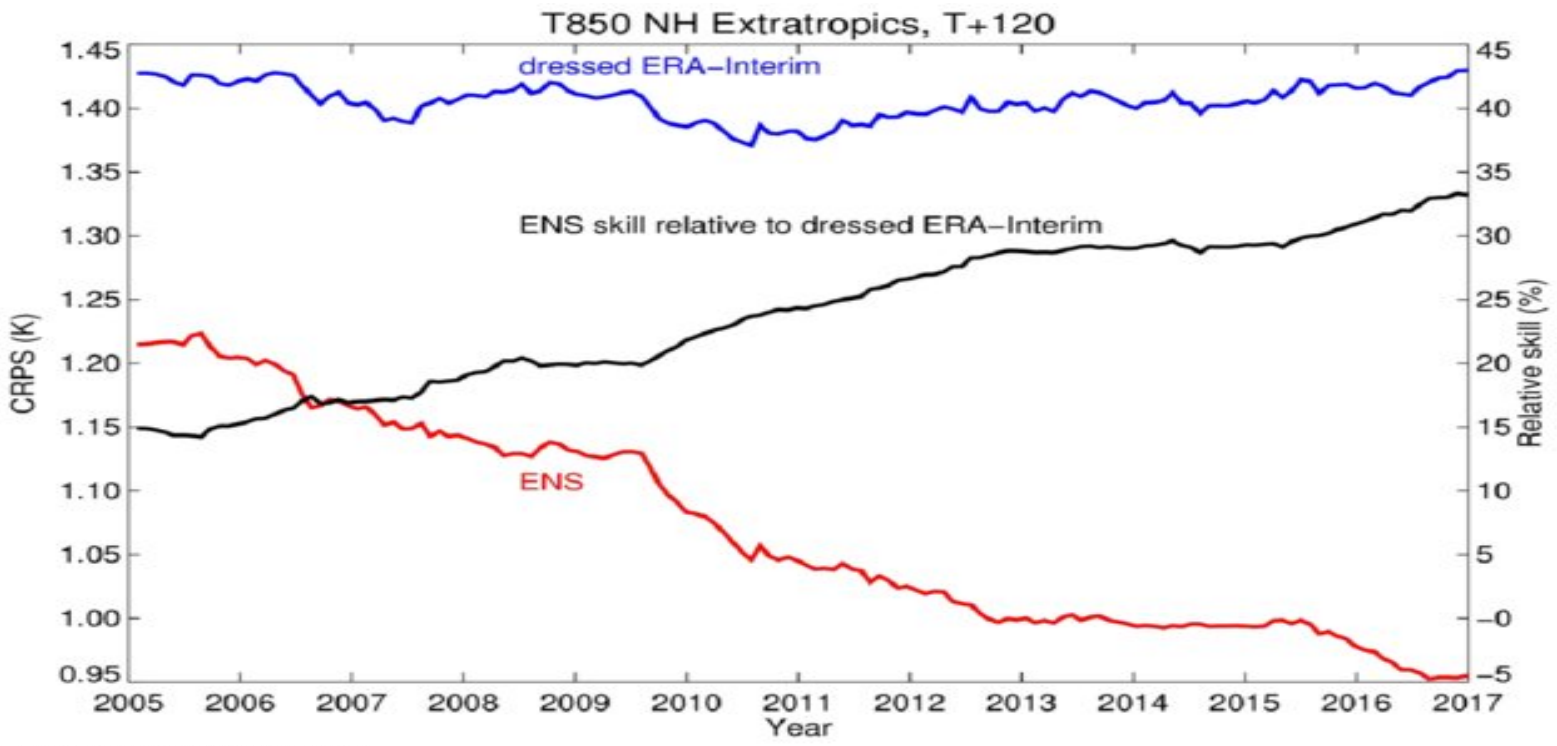


Benchmarking the system

Quick check against ECMWF operational ensemble. CRPS of T850 over the years.

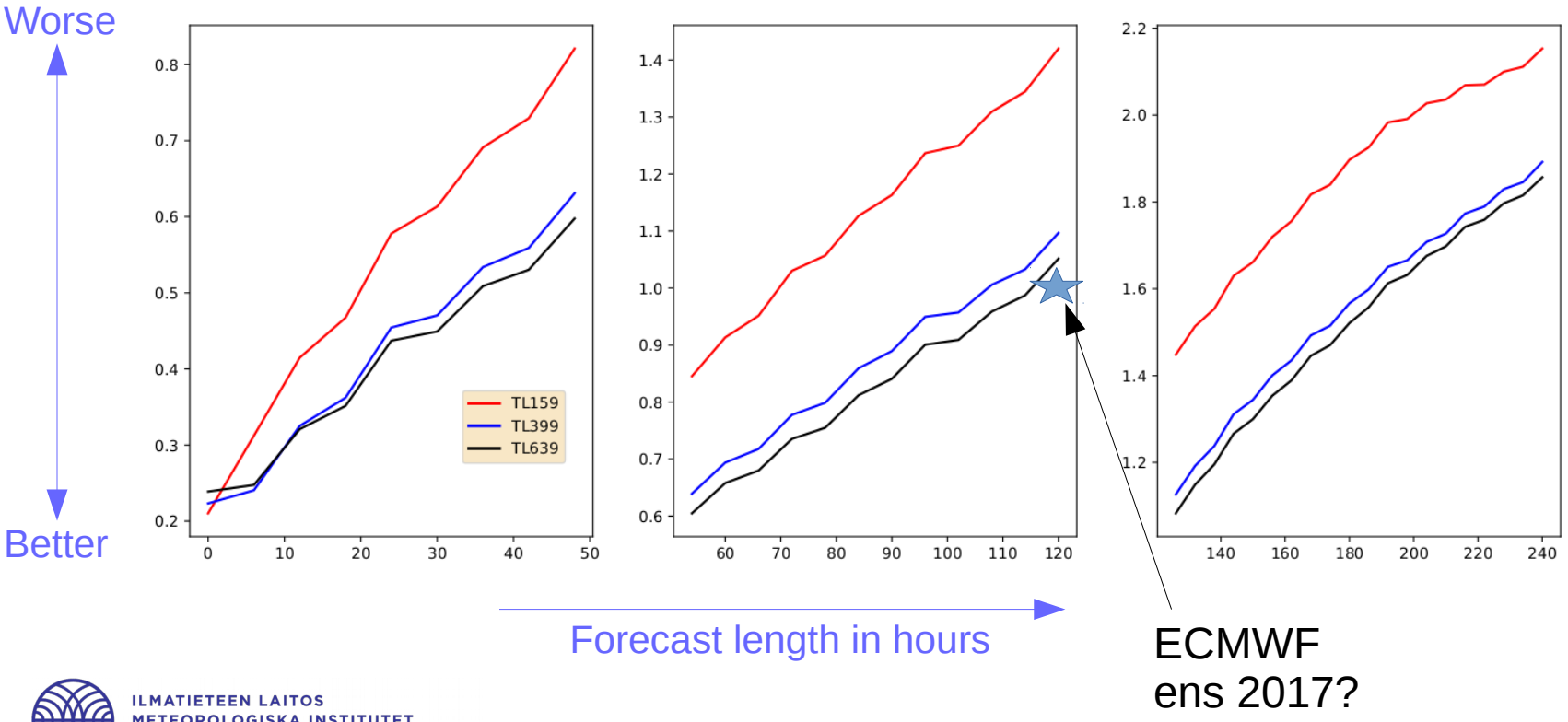
Worse

Better



Benchmarking the system

FAIR of temperature at 850hPa in the Northern extra-tropics. Mean over 27 start dates.

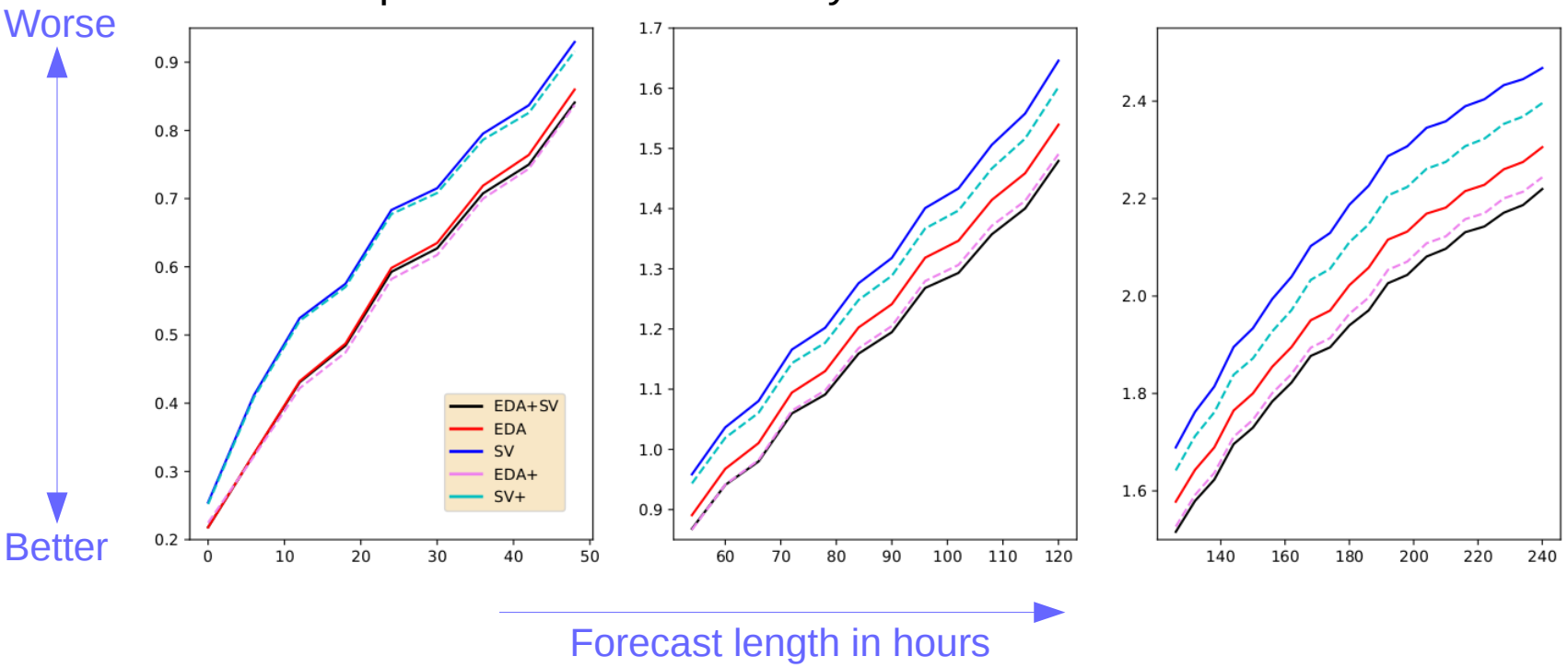


Benchmarking the system

- FAIR seems to work as it should!
- So, lets answer a few questions:
  - How does model resolution affect the probabilistic skill of the system?
  - What are the contributions of SV and EDA perturbations on the skill?

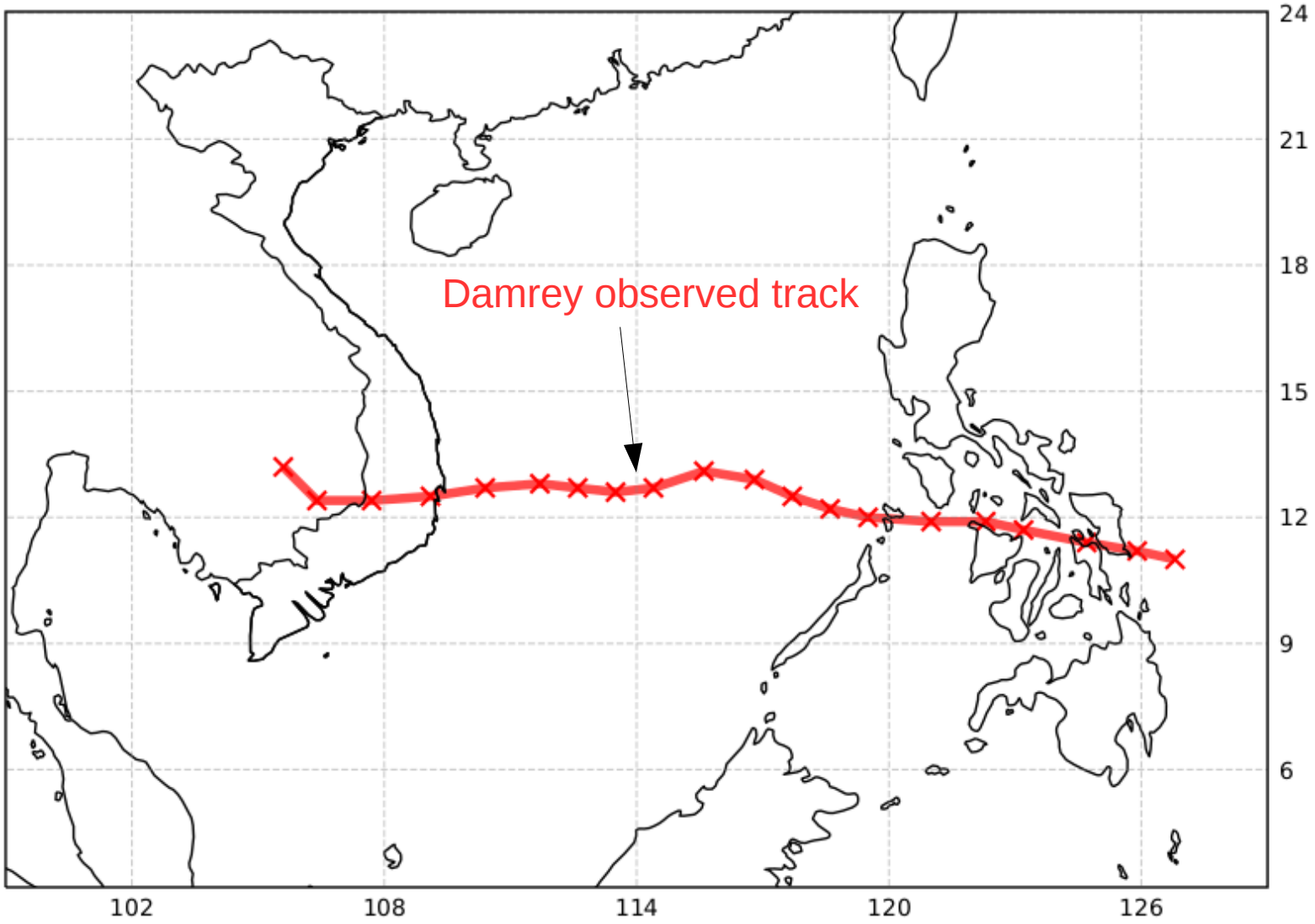
Benchmarking the system

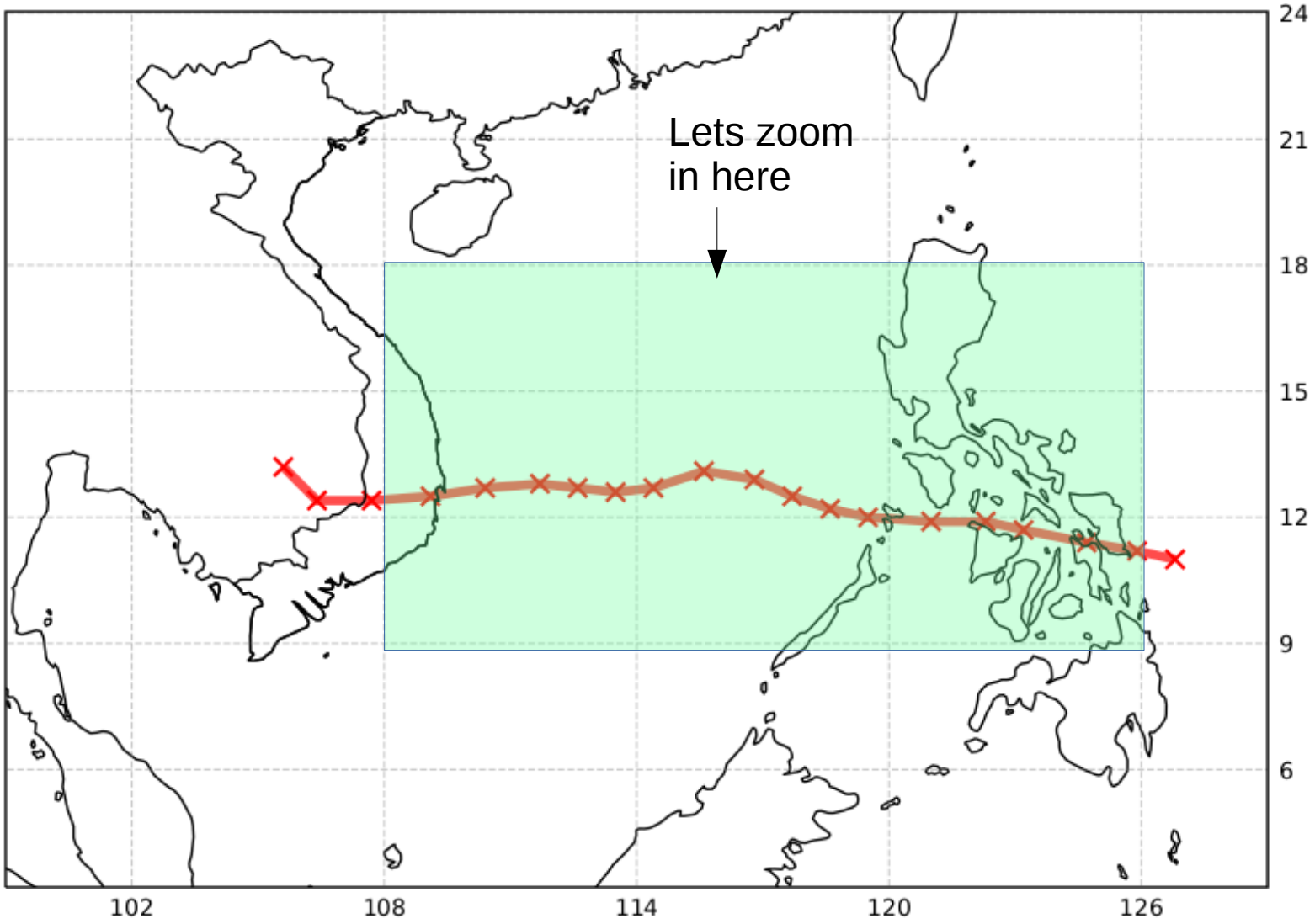
FAIR of temperature at 850hPa in the Northern extra-tropics. Mean over 47 start dates.
EDA+: EDA perturbations inflated by 1.2.
SV+: SV perturbations inflated by 1.2.

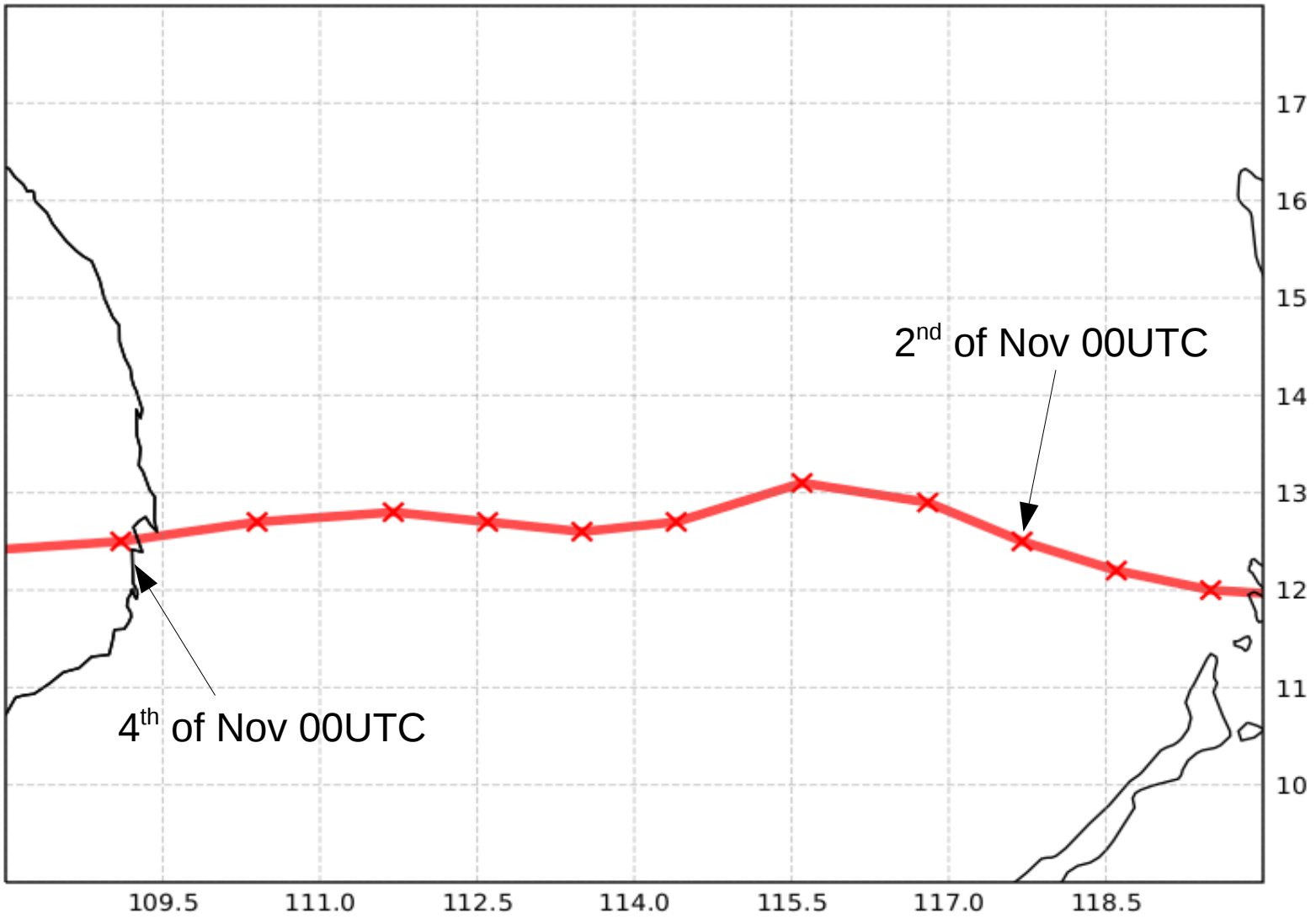


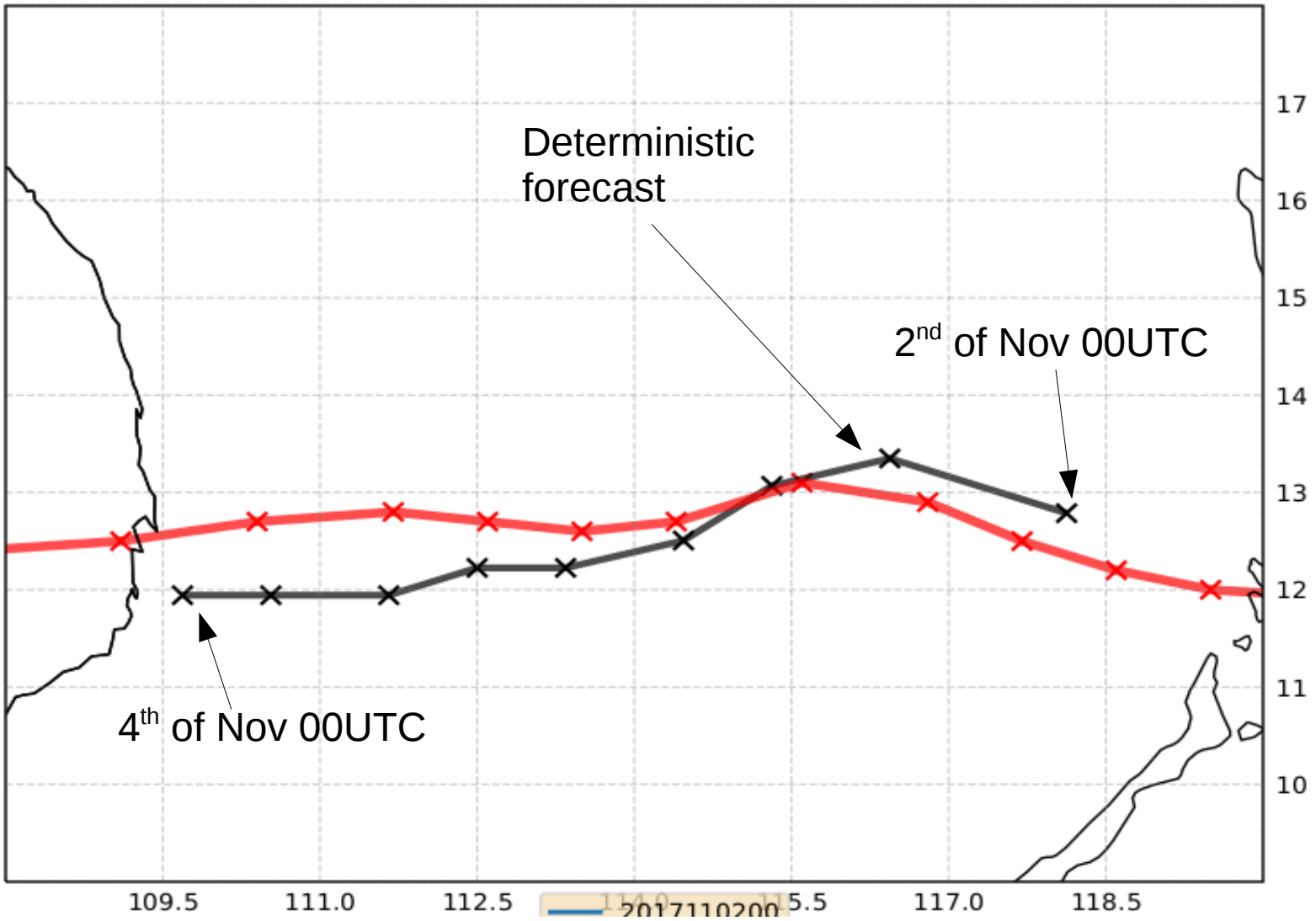
Use in case studies: Tropical Cyclone example

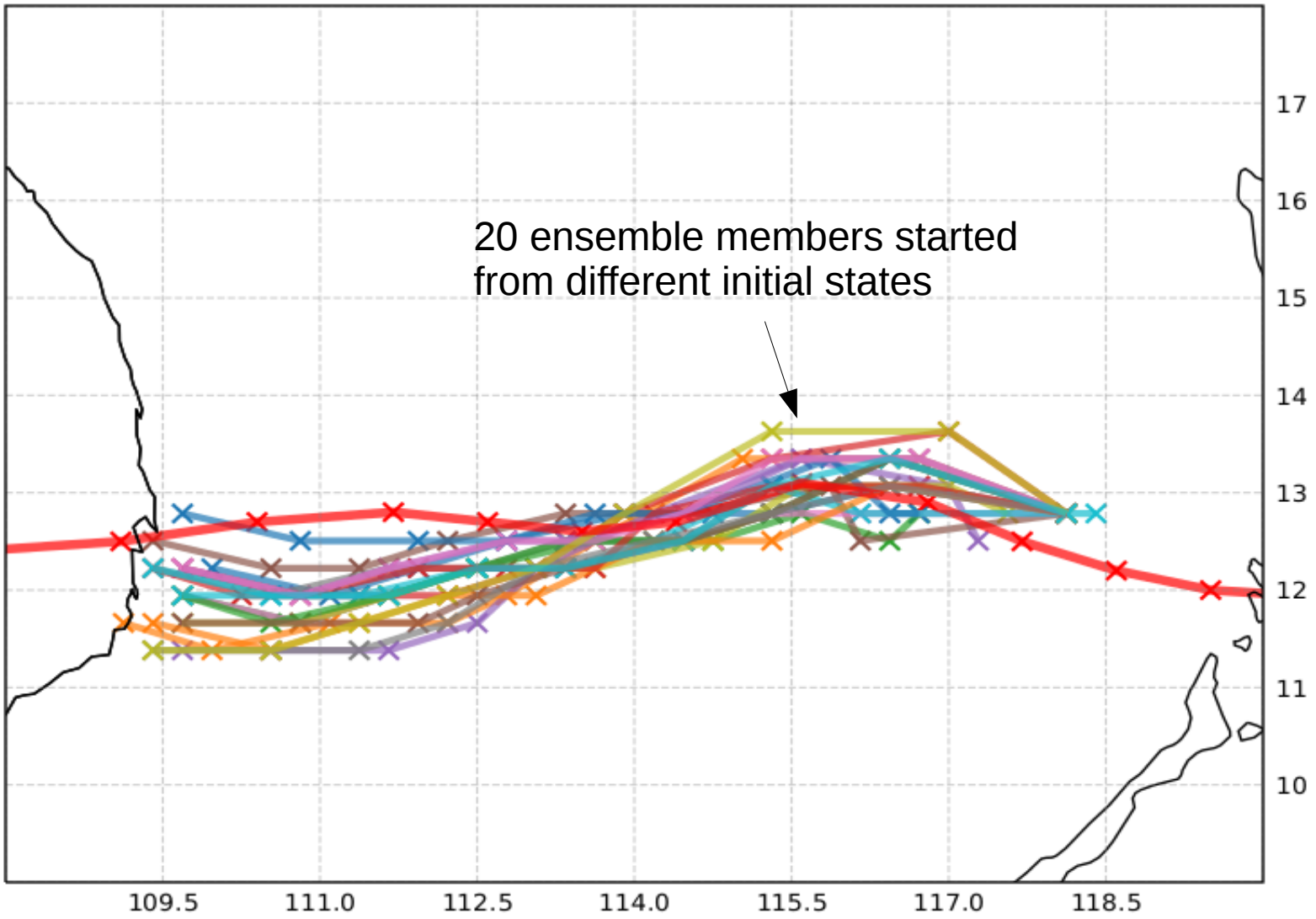
- Damrey typhoon
  - Made landfall in Vietnam 00UTC 4th of November 2017
- Lets simulate the case with OIFS
  - CY40R1
  - TL639
  - 20 ensemble members
  - SV and EDA perturbations
- Tracking the cyclone
  - Simply find the MSLP minimum


Damrey observed track
24
21
18
15
12
9
6
102
108
114
120
126


Lets zoom
in here
24
21
18
15
12
9
6
102
108
114
120
126


17
16
15
14
13
12
11
10
109.5
111.0
112.5
114.0
115.5
117.0
118.5
2nd of Nov 00UTC
4th of Nov 00UTC


Deterministic forecast
2nd of Nov 00UTC
4th of Nov 00UTC
17
16
15
14
13
12
11
10
109.5
111.0
112.5
117.0
118.5


20 ensemble members started
from different initial states
17
16
15
14
13
12
11
10
109.5
111.0
112.5
114.0
115.5
117.0
118.5

Use in case studies: Tropical Cyclone example

- Only using deterministic forecast does not provide enough information (wrong, too southerly track etc.)
- An ensemble started with initial state perturbations starts to find correct solutions of the track

Use in case studies: Tropical Cyclone example

- Only using deterministic forecast does not provide enough information (wrong, too southerly track etc.)
- An ensemble started with initial state perturbations starts to find correct solutions of the track
- What is more crucial, better representation of the TC initial structure, or better representation of the prevailing flow situation?
- Is it enough to get some aspects of the initial state better represented (e.g. winds)?

Discussion

- The dataset of ensemble initial states will be made public once it has been documented
  - FTP server
  - 22TB of data in tarz-format
- I’ve only discussed initial state perturbations here
  - CY40R1 includes a model uncertainty representation as well (SPPT)
  - CY43R3 includes an alternative model uncertainty representation currenly under development at ECMWF (SPP)

Summary

- Why?
  - Ensembles open up a lot of new ways to tackle scientific problems
- How?
  - Get a workflow manager
  - Download OIFS ensemble initial states once available
  - Include a model uncertainty representation

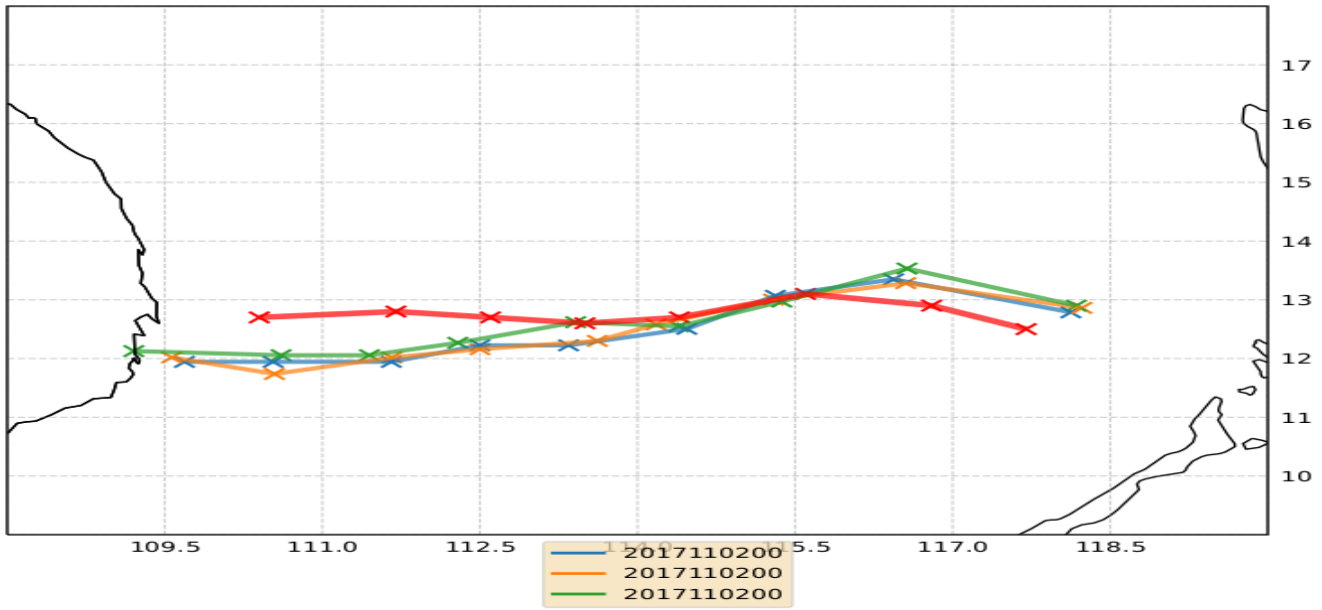
ILMATIETEEN LAITOS
METEOROLOGISKA INSTITUTET
FINNISH METEOROLOGICAL INSTITUTE

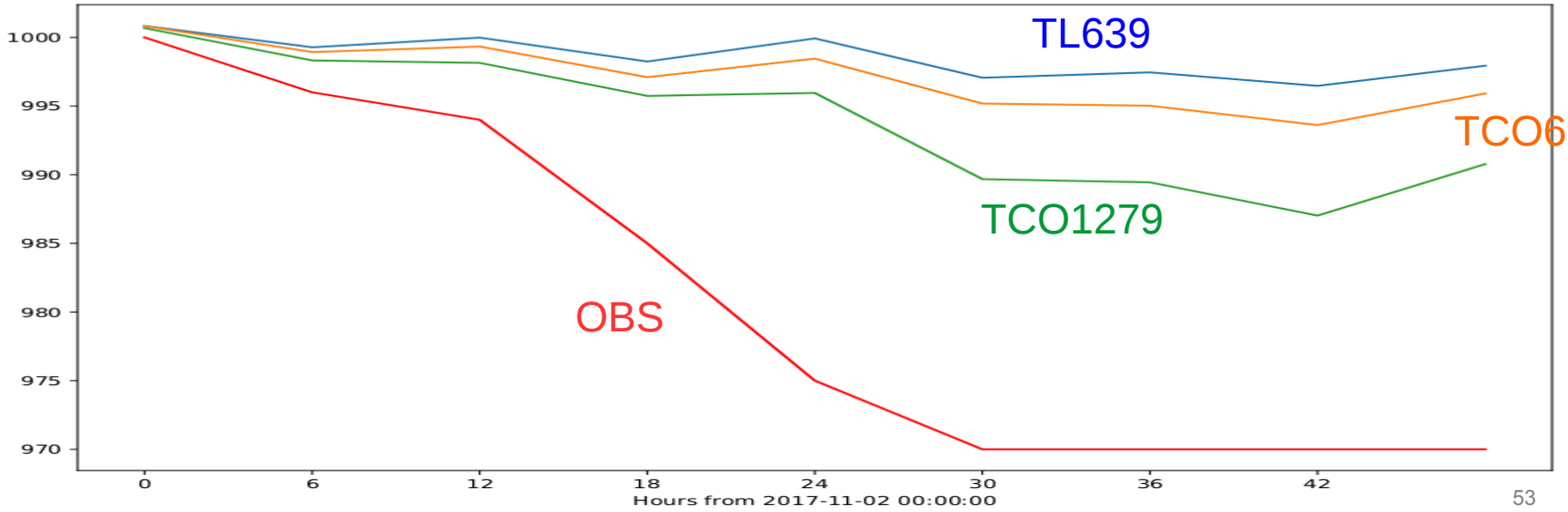
Acknowledgements

ECMWF Special Project: “Parameter estimation (EPPES) in HarmonEPS “

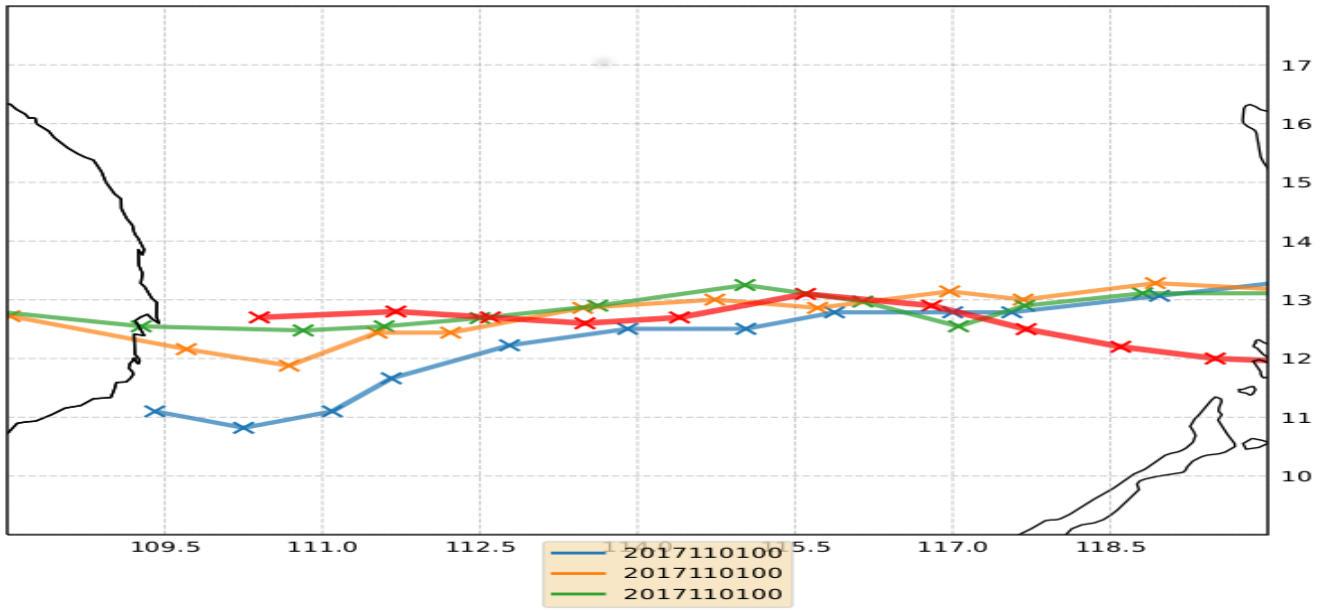
Academy of Finland, Postdoctoral Researcher, 316939: “Quantifying model-specific uncertainties in climate simulations”

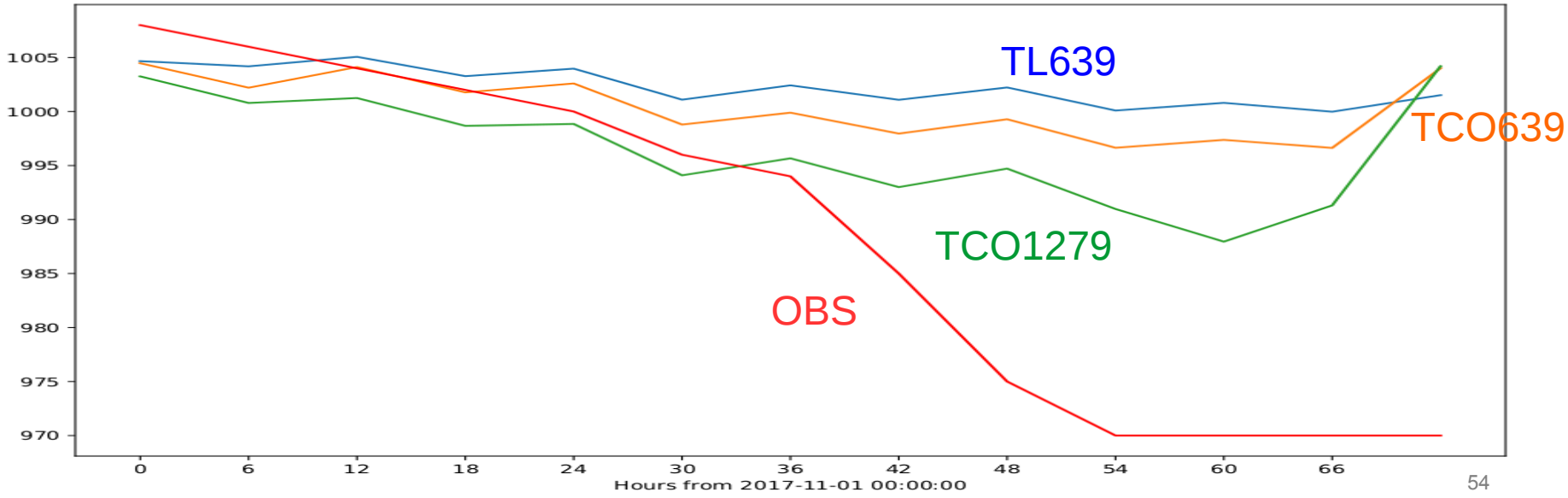
2nd of Nov, 00UTC. +48h



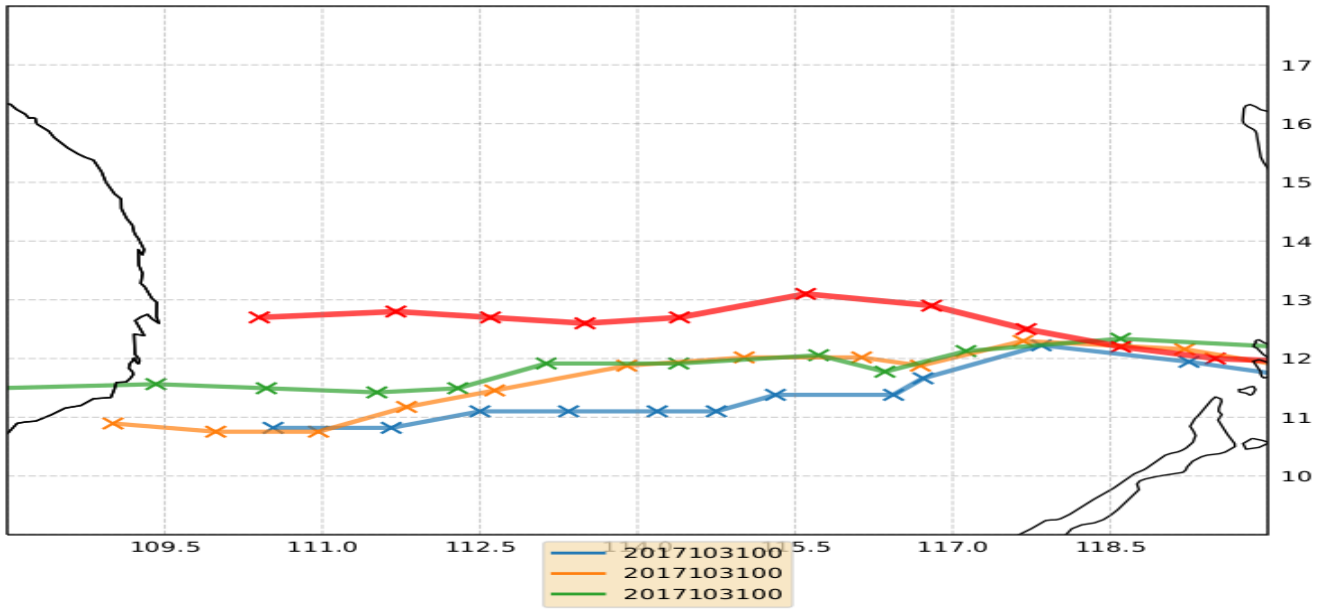


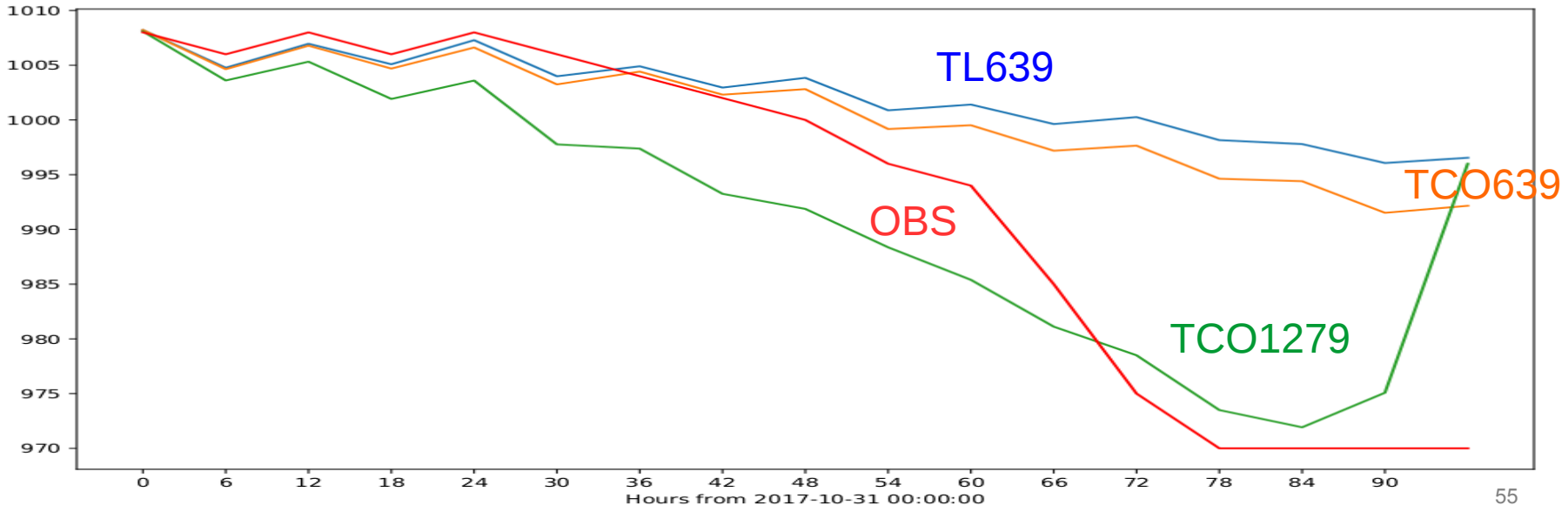
1st of Nov, 00UTC. +72h



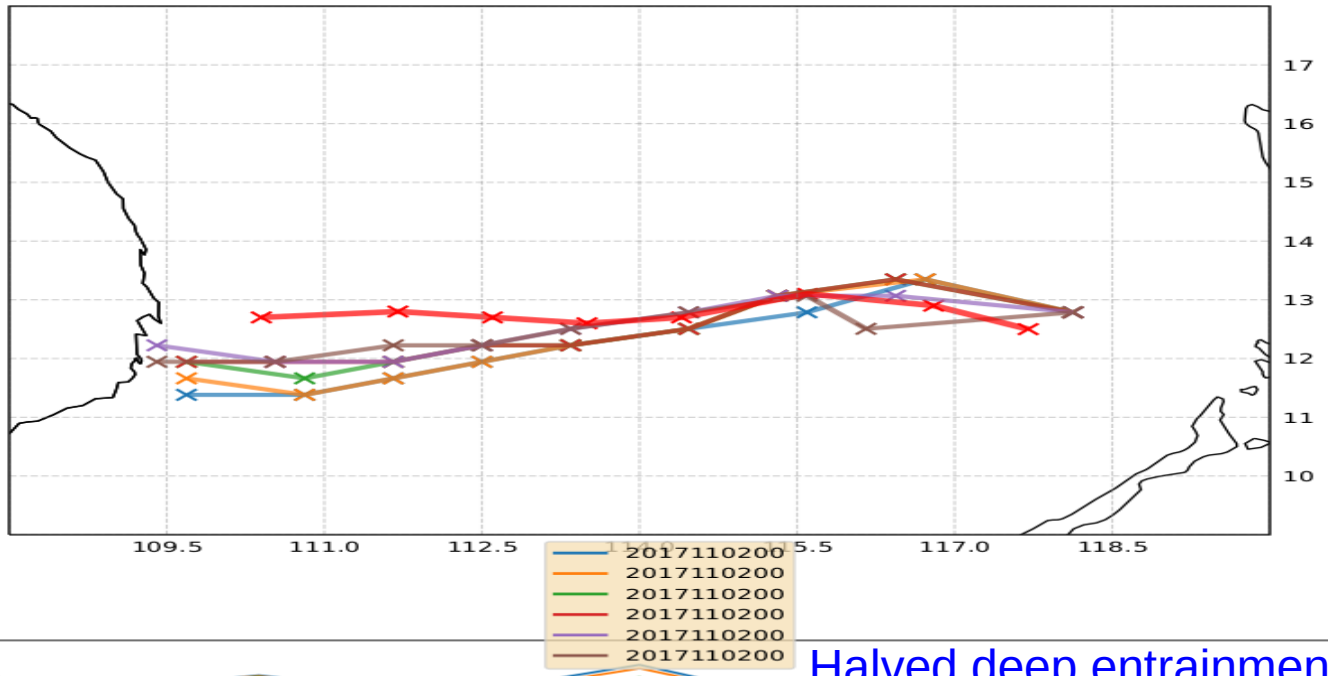


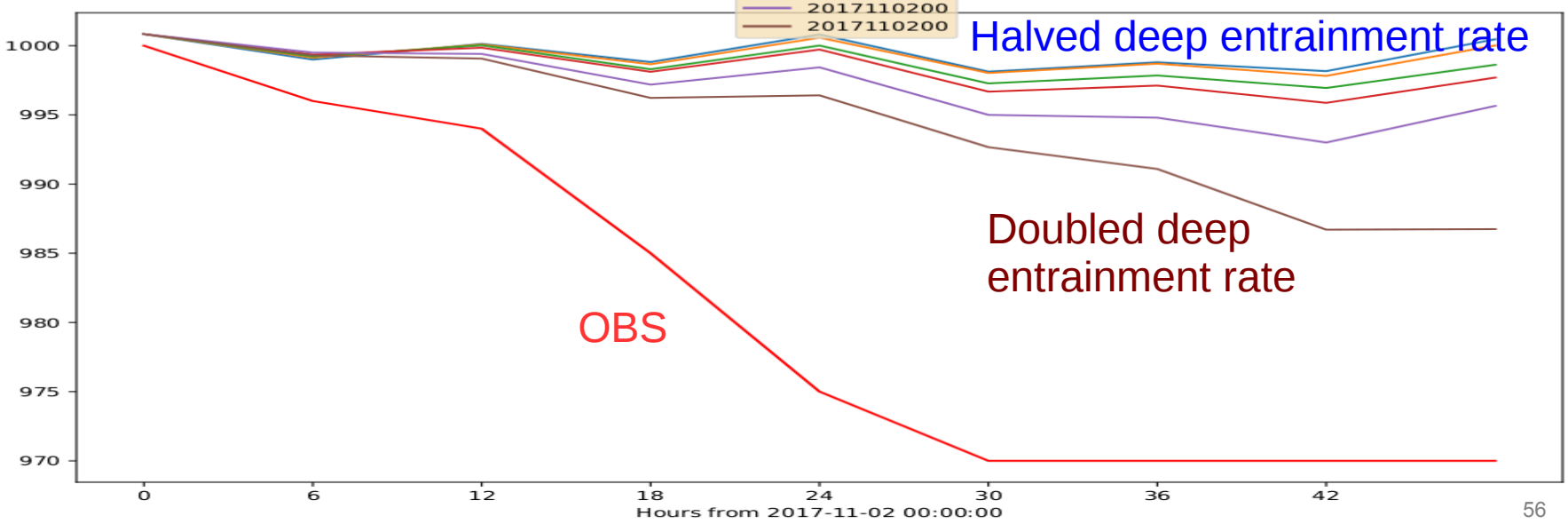
31st of Oct, 00UTC. +96h



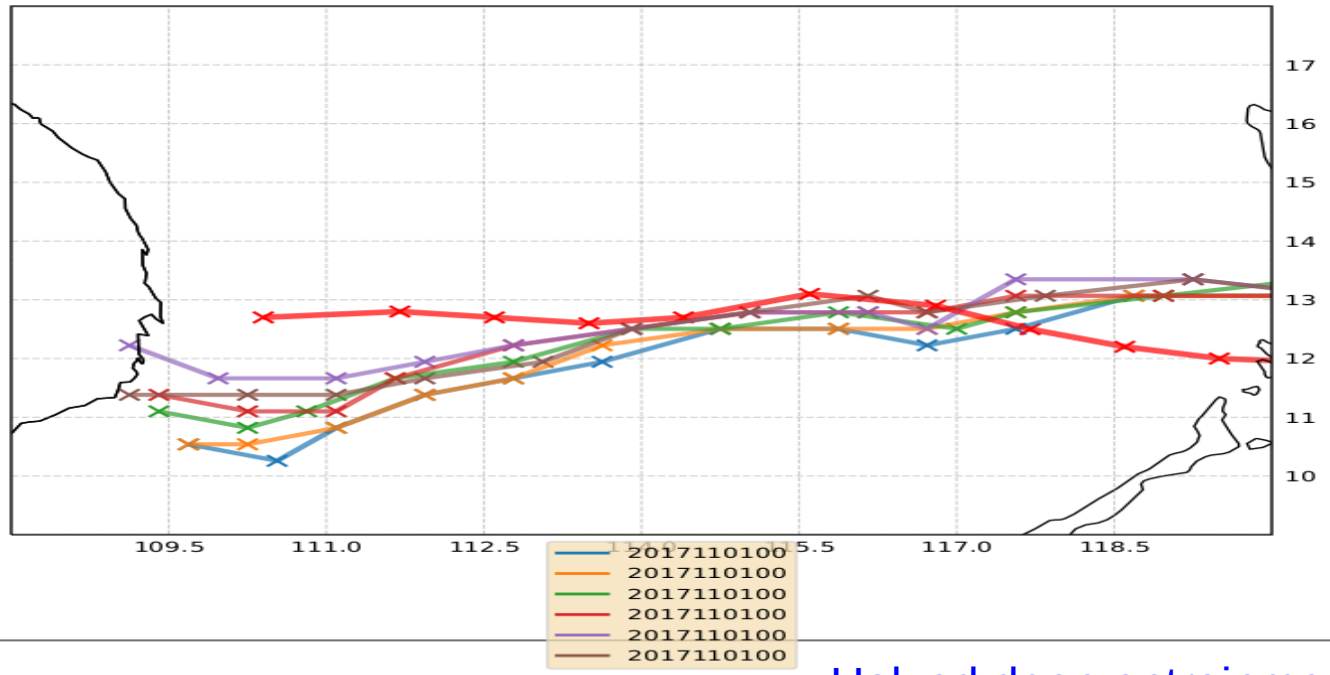


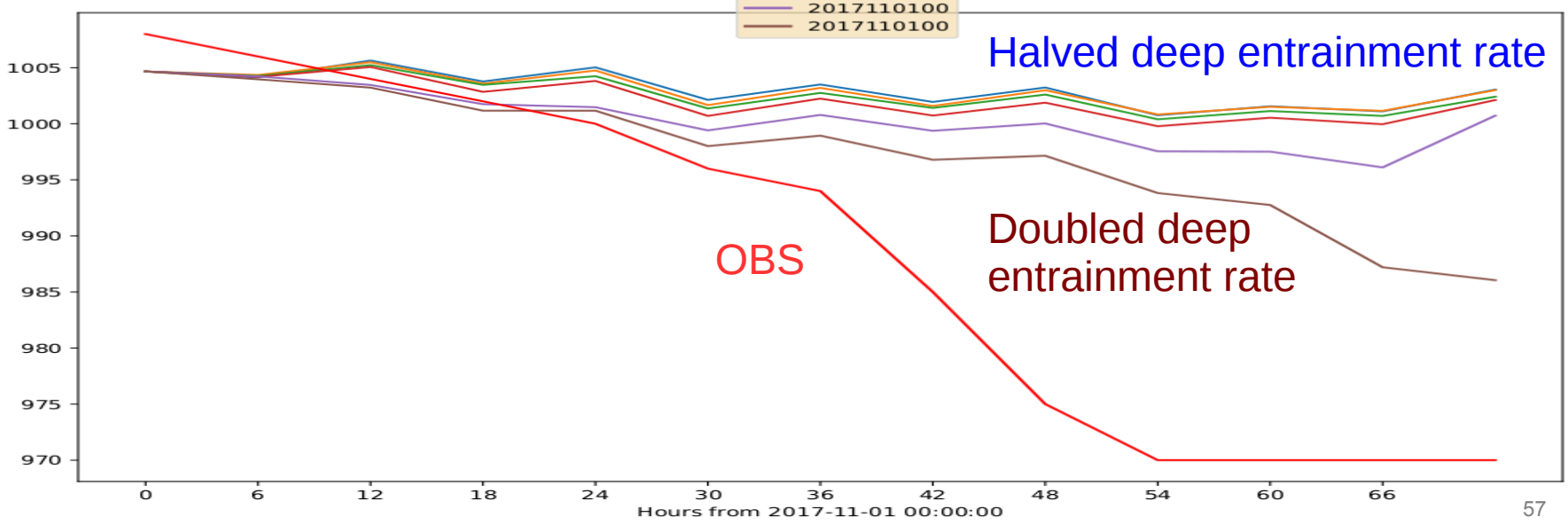
2nd of Nov, 00UTC. +48h





2nd of Nov, 00UTC. +72h





2nd of Nov, 00UTC. +96h