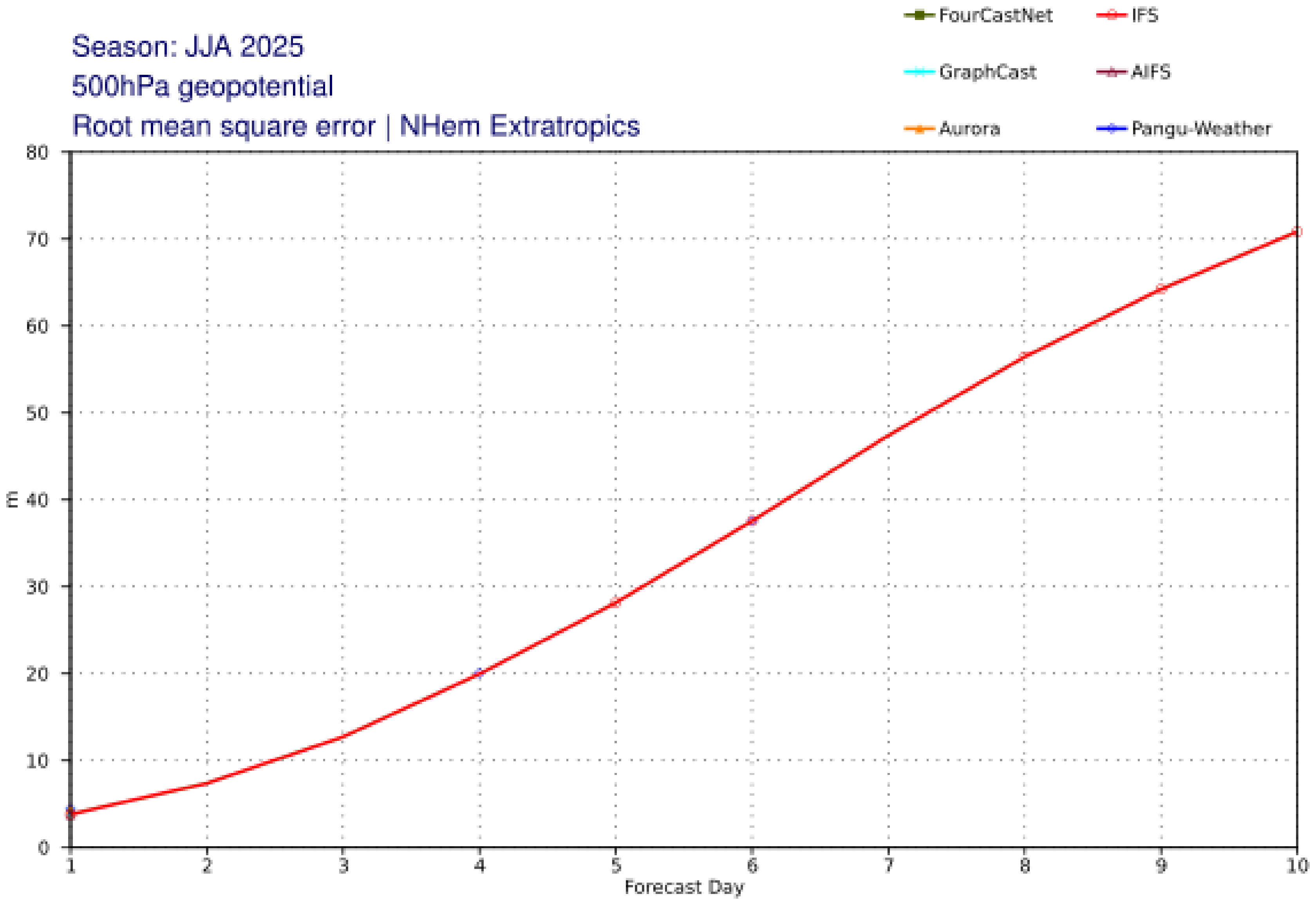
Machine Learning in Weather and Climate

Overview and Motivation

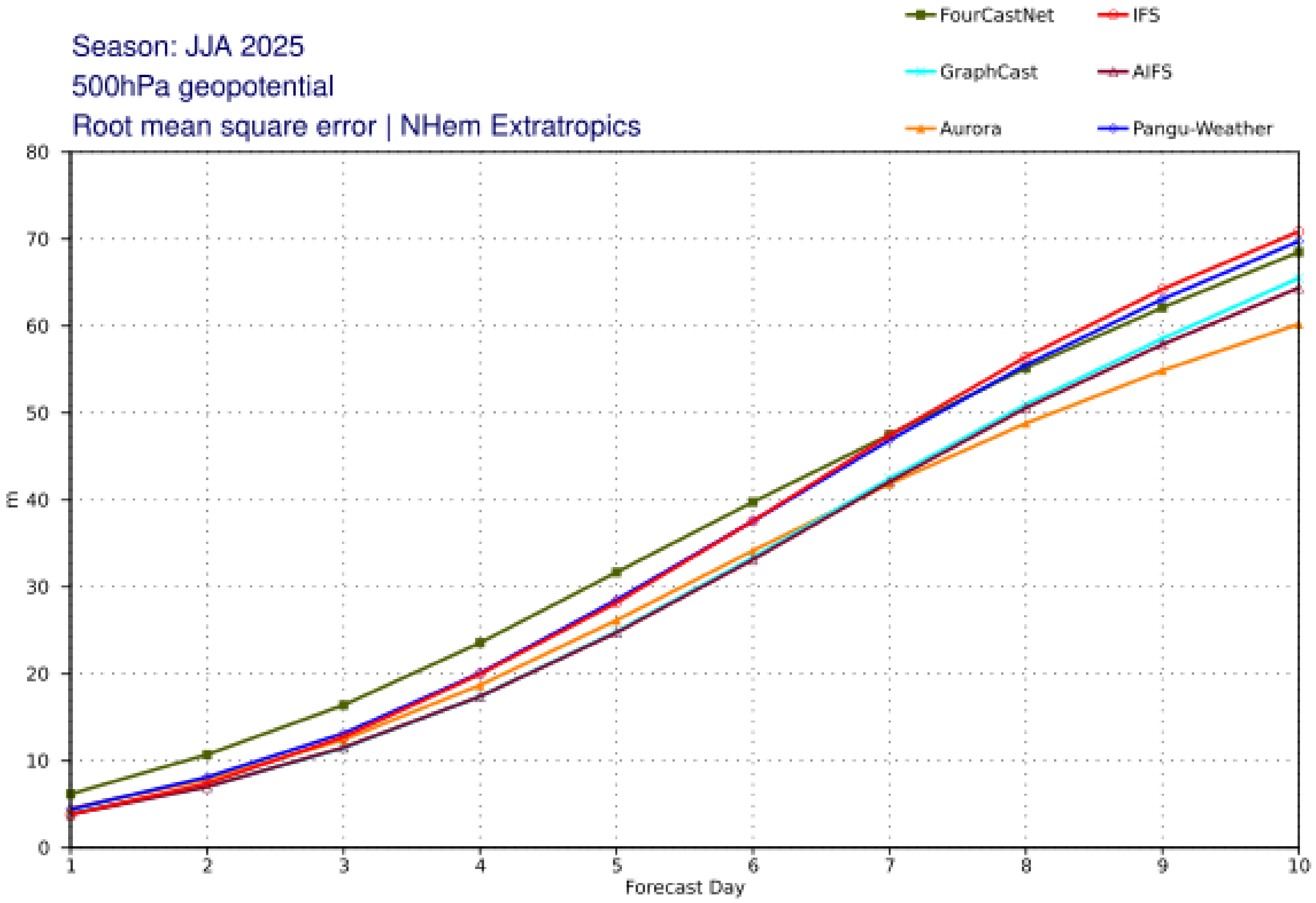
Christian Lessig
christian.lessig@ecmwf.int

Motivation



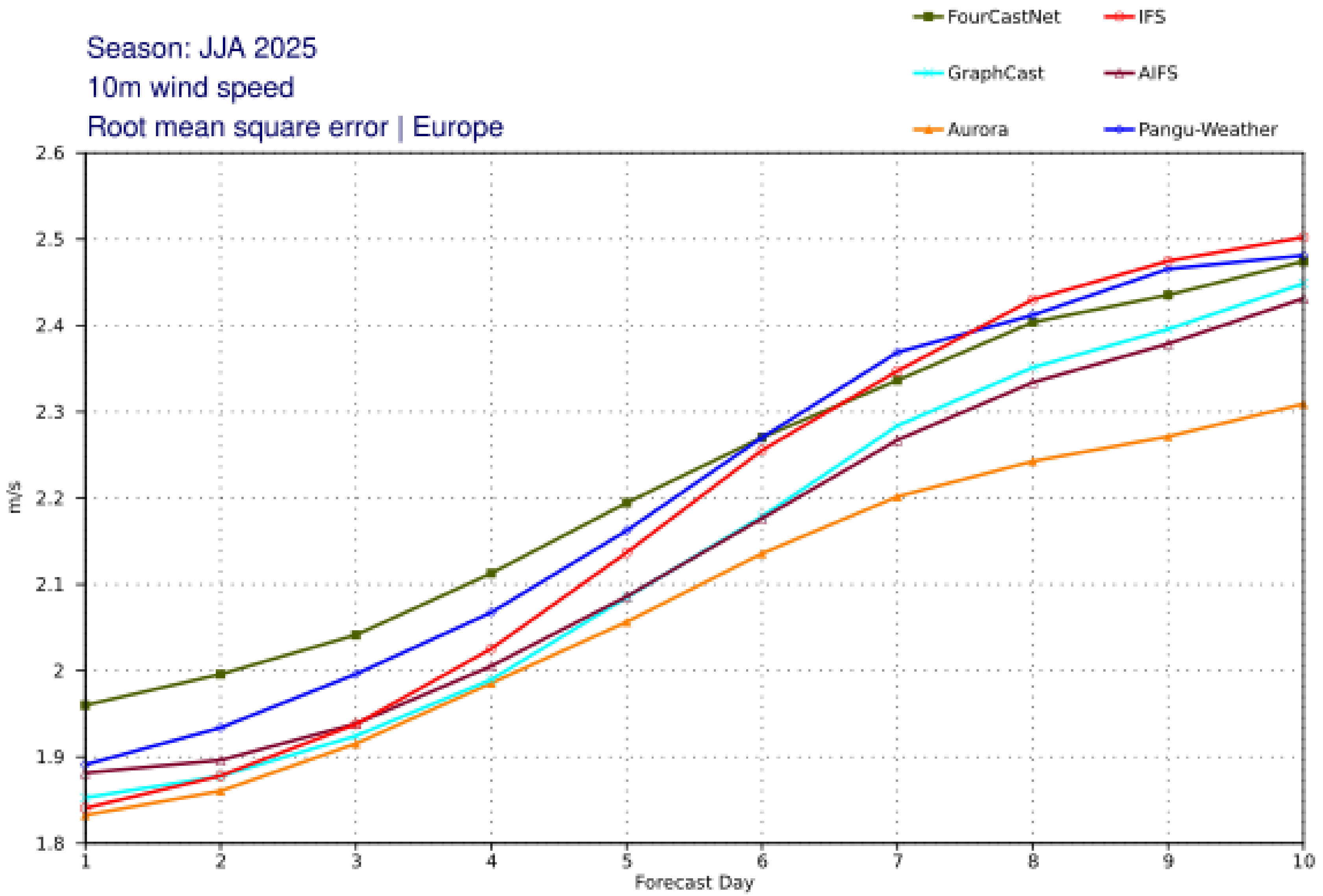
Season: JJA 2025
500hPa geopotential
Root mean square error | NHem Extratropics
FourCastNet
IFS
GraphCast
AIFS
Aurora
Pangu-Weather
m
0
10
20
30
40
50
60
70
80
1
2
3
4
5
6
7
8
9
10
Forecast Day

Motivation



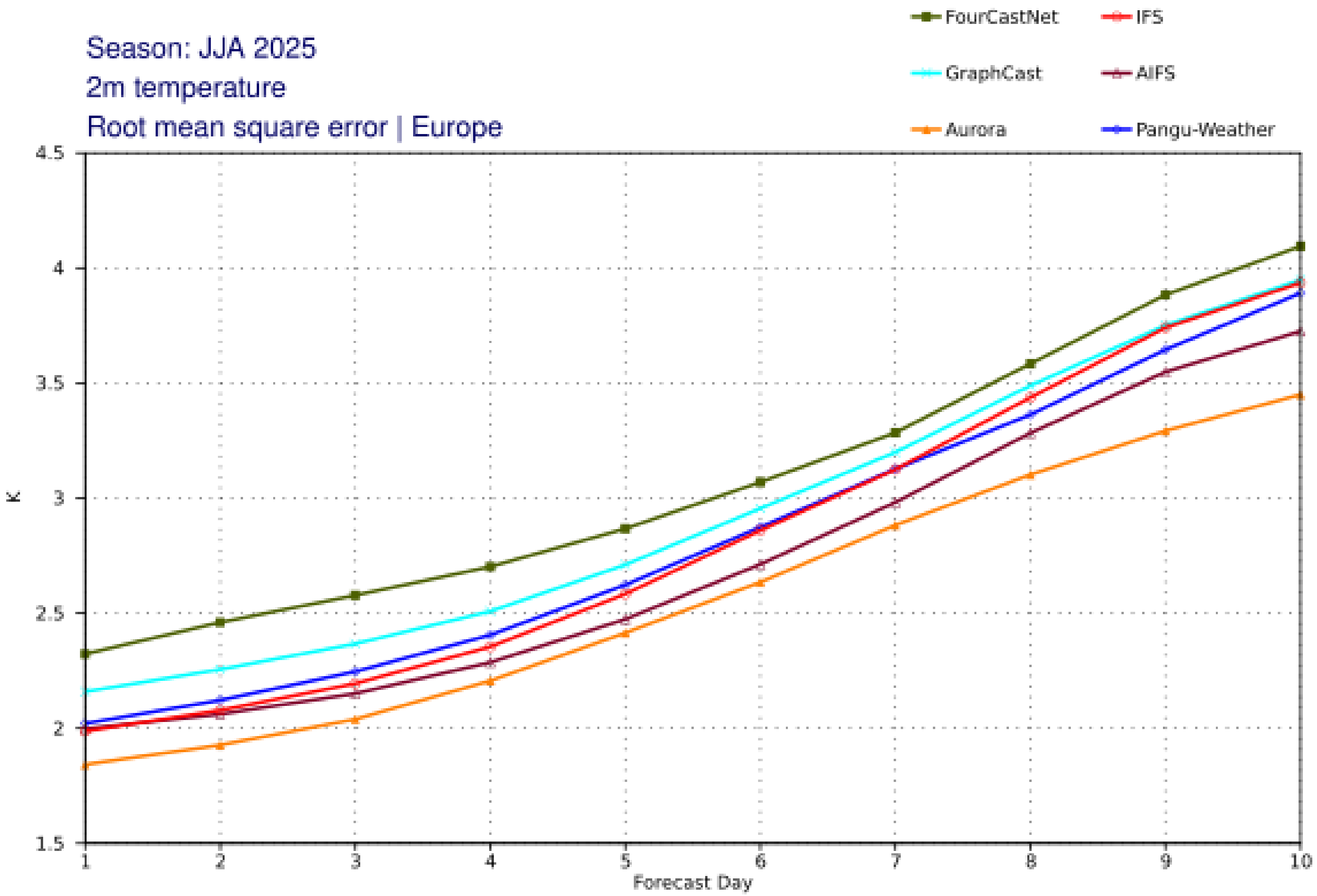
Season: JJA 2025
500hPa geopotential
Root mean square error | NHem Extratropics
FourCastNet
IFS
GraphCast
AIFS
Aurora
Pangu-Weather
m
0
10
20
30
40
50
60
70
80
1
2
3
4
5
6
7
8
9
10
Forecast Day

Motivation



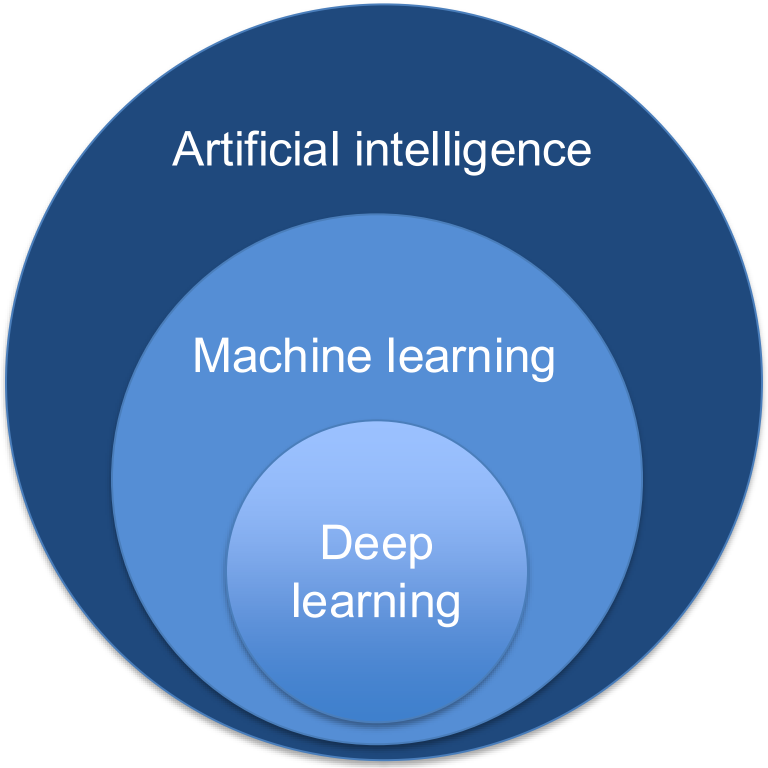
Season: JJA 2025
10m wind speed
Root mean square error | Europe
FourCastNet
IFS
GraphCast
AIFS
Aurora
Pangu-Weather
m/s
2.6
2.5
2.4
2.3
2.2
2.1
2
1.9
1.8
1
2
3
4
5
6
7
8
9
10
Forecast Day

Motivation



ECMWF

What is machine learning?



Artificial intelligence (AI) *is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans.*[1]

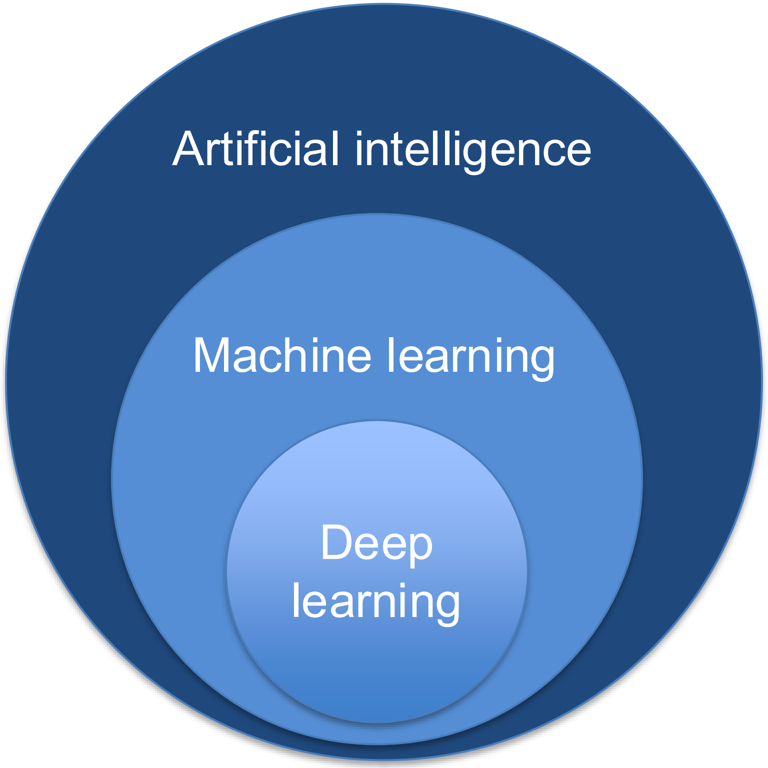
Machine learning (ML) *encompasses algorithms and statistical models primarily derived from data that are used to perform tasks without using explicit instructions.*[1]

Deep learning *are machine learning methods using large/deep artificial neural networks.*[1]

---

[1] Wikipedia

What is machine learning?



Artificial intelligence (AI) *is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans.*[1]

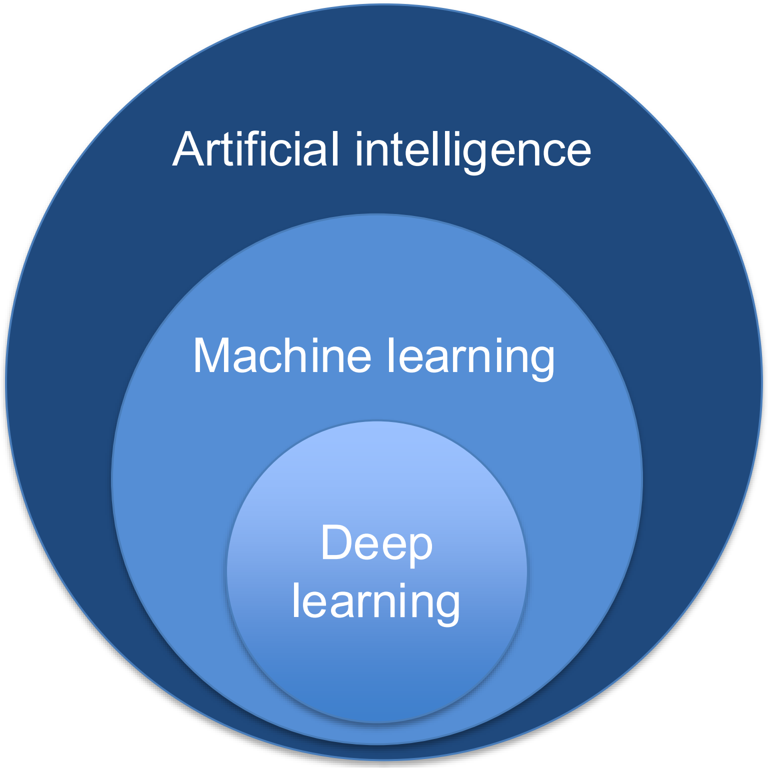
Machine learning (ML) *encompasses algorithms and statistical models primarily derived from data that are used to perform tasks without using explicit instructions.*[1]

Deep learning *are machine learning methods using large/deep artificial neural networks.*[1]

qualitatively different behaviour than small ML methods

[1] Wikipedia

What is machine learning?



Artificial intelligence (AI) *is intelligence demonstrated by machines, in contrast to the natural intelligence displayed by humans.*[1]

Machine learning (ML) *encompasses algorithms and statistical models primarily derived from data that are used to perform tasks without using explicit instructions.*[1]

Deep learning *are machine learning methods using large/deep artificial neural networks.*[1]

qualitatively different behaviour than small ML methods

[1] Wikipedia

What is machine learning?

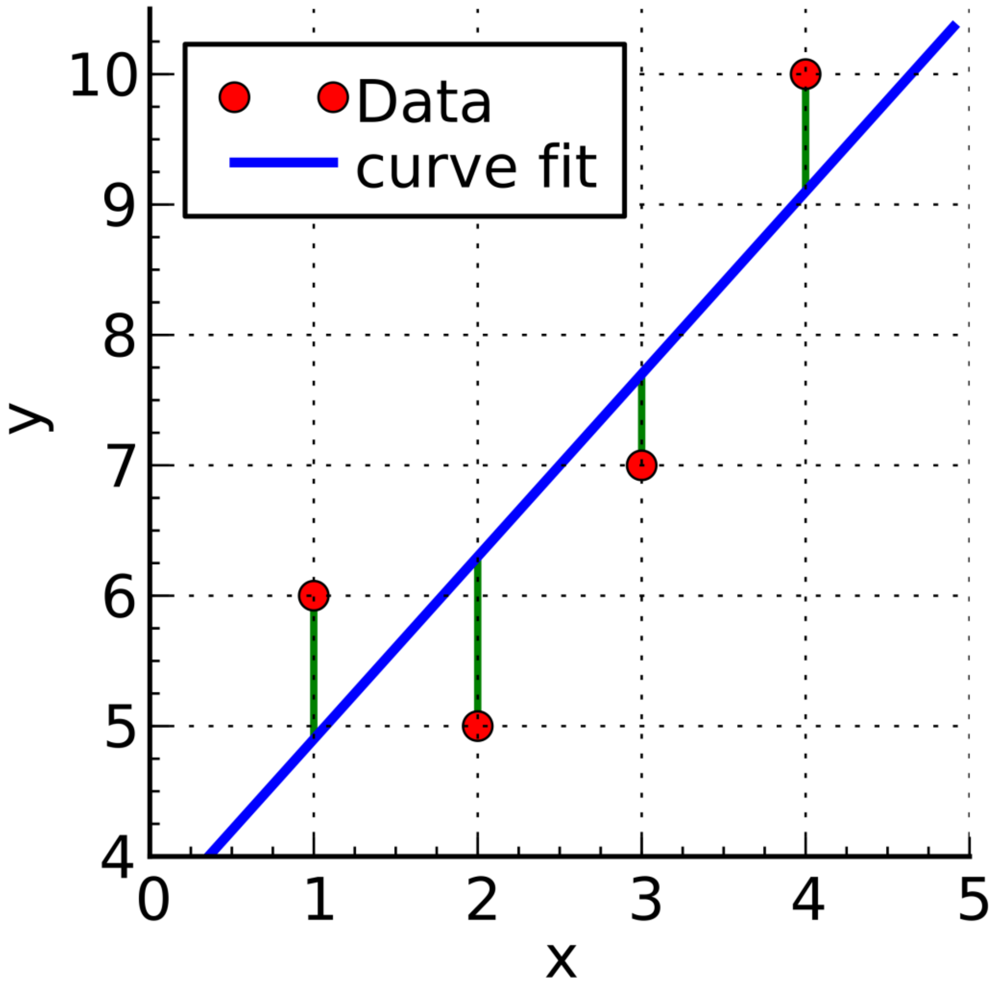
You all know machine learning algorithms:

$$x$$

What is machine learning?

You all know machine learning algorithms:

- Linear regression
  - Training data: point observations (x_i, y_i)
  - Model: linear function $y = f(x) = m\,x + n$
  - Training process: least squares fit / solution of overcomplete linear system
  - Inference: evaluate linear function with new x



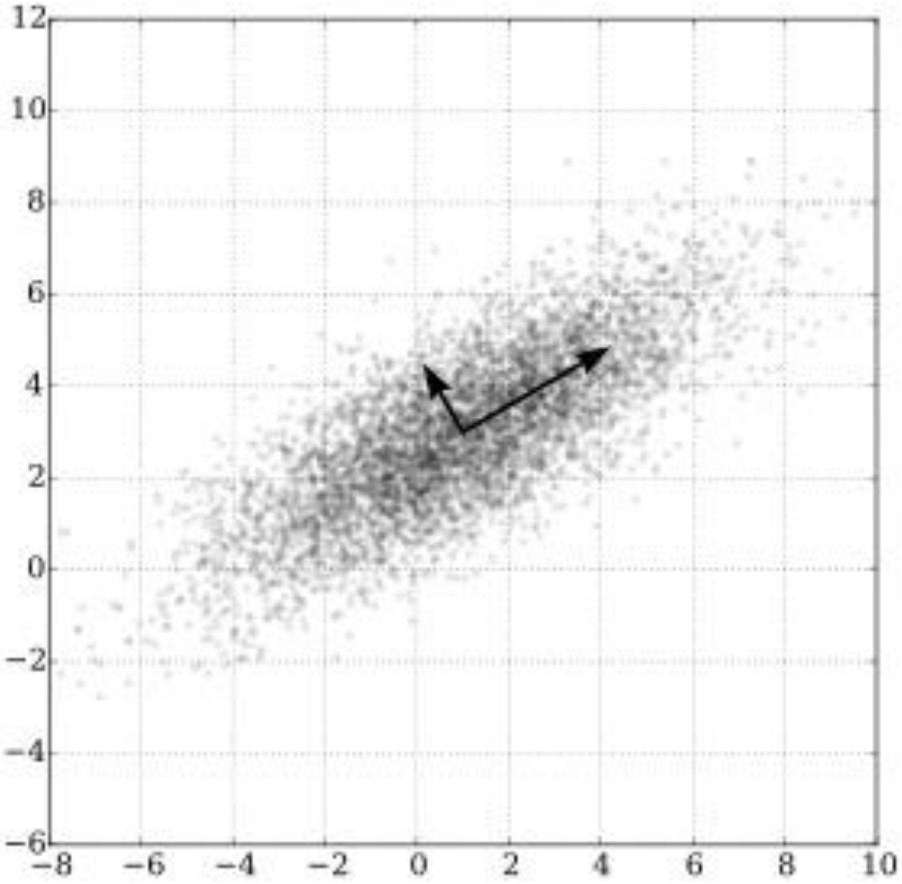
from Wikipedia

What is machine learning?

You all know machine learning algorithms:

- Linear regression
  - Training data: point observations (x_i, y_i)
  - Model: linear function $y = f(x) = m\,x + n$
  - Training process: least squares fit / solution of overcomplete linear system
  - Inference: evaluate linear function with new x

- Principal Component Analysis (PCA)
  - Proper orthogonal decomposition, Karhunen-Loeve transform, Hoteling transform …
  - Training data: data points $x_i \in \mathbb{R}^n$
  - Model: reduced basis representation
  - Training processes: computation of singular value decomposition
  - Inference: project data into reduced basis



from Wikipedia

What is machine learning?

The deep learning revolution:

- 1950s: Artificial intelligence starts as scientific field to match or surpass human intellectual abilities
    - Development with a lot of over-promises and disappointments over the decades

AI: "the science of making machines do things that would require intelligence if done by humans”

John McCarthy

”[the perceptron] may eventually be able to learn, make decisions, and translate languages”

Frank Rosenblatt

"Artificial intelligence is the new electricity”

Andrew Ng

What is machine learning?

The deep learning revolution:

- 1950s: Artificial intelligence starts as scientific field to match or surpass human intellectual abilities
    - Development with a lot of over-promises and disappointments over the decades
- 1950s and 1960s: Perceptron and neural network introduced
    - Often motivated by and for neuroscience research
    - Training larger neural networks remained essentially impossible until late 2000s
    - Key advance was availability of GPUs for general purpose computations (CUDA etc)

AI: "the science of making machines do things that would require intelligence if done by humans"

John McCarthy

"[the perceptron] may eventually be able to learn, make decisions, and translate languages"

Frank Rosenblatt

"Artificial intelligence is the new electricity"

Andrew Ng

What is machine learning?

The deep learning revolution:

- 1950s: Artificial intelligence starts as scientific field to match or surpass human intellectual abilities
  - Development with a lot of over-promises and disappointments over the decades
- 1950s and 1960s: Perceptron and neural network introduced
  - Often motivated by and for neuroscience research
  - Training larger neural networks remained essentially impossible until late 2000s
  - Key advance was availability of GPUs for general purpose computations (CUDA etc)
- 2010-2020: Rapid take over of machine learning by deep neural networks and algorithmic advancements
  - Neural networks quickly morphed from academic niche to practical deployment

AI: "the science of making machines do things that would require intelligence if done by humans”

John McCarthy

”[the perceptron] may eventually be able to learn, make decisions, and translate languages”

Frank Rosenblatt

"Artificial intelligence is the new electricity”

Andrew Ng

What is machine learning?

The deep learning revolution:

- 1950s: Artificial intelligence starts as scientific field to match or surpass human intellectual abilities
  - Development with a lot of over-promises and disappointments over the decades
- 1950s and 1960s: Perceptron and neural network introduced
  - Often motivated by and for neuroscience research
  - Training larger neural networks remained essentially impossible until late 2000s
  - Key advance was availability of GPUs for general purpose computations (CUDA etc)
- 2010-2020: Rapid take over of machine learning by deep neural networks and algorithmic advancements
  - Neural networks quickly morphed from academic niche to practical deployment
- 2020+: scaling of neural networks to trillions of trainable parameters

AI: "the science of making machines do things that would require intelligence if done by humans”

John McCarthy

”[the perceptron] may eventually be able to learn, make decisions, and translate languages”

Frank Rosenblatt

"Artificial intelligence is the new electricity”

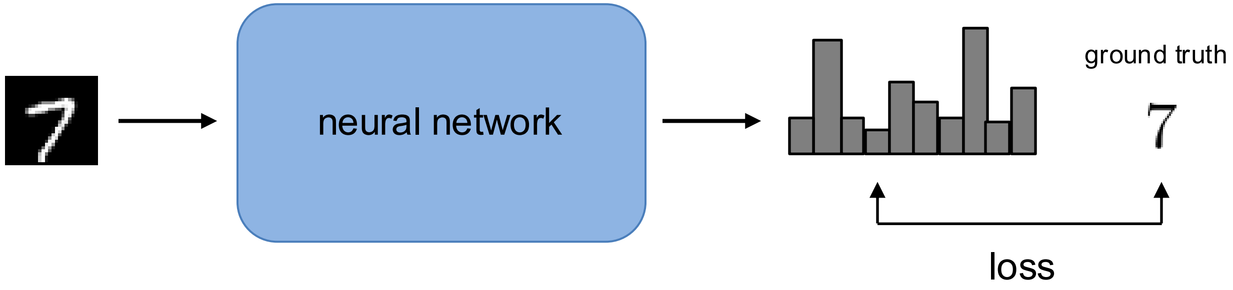
Andrew Ng

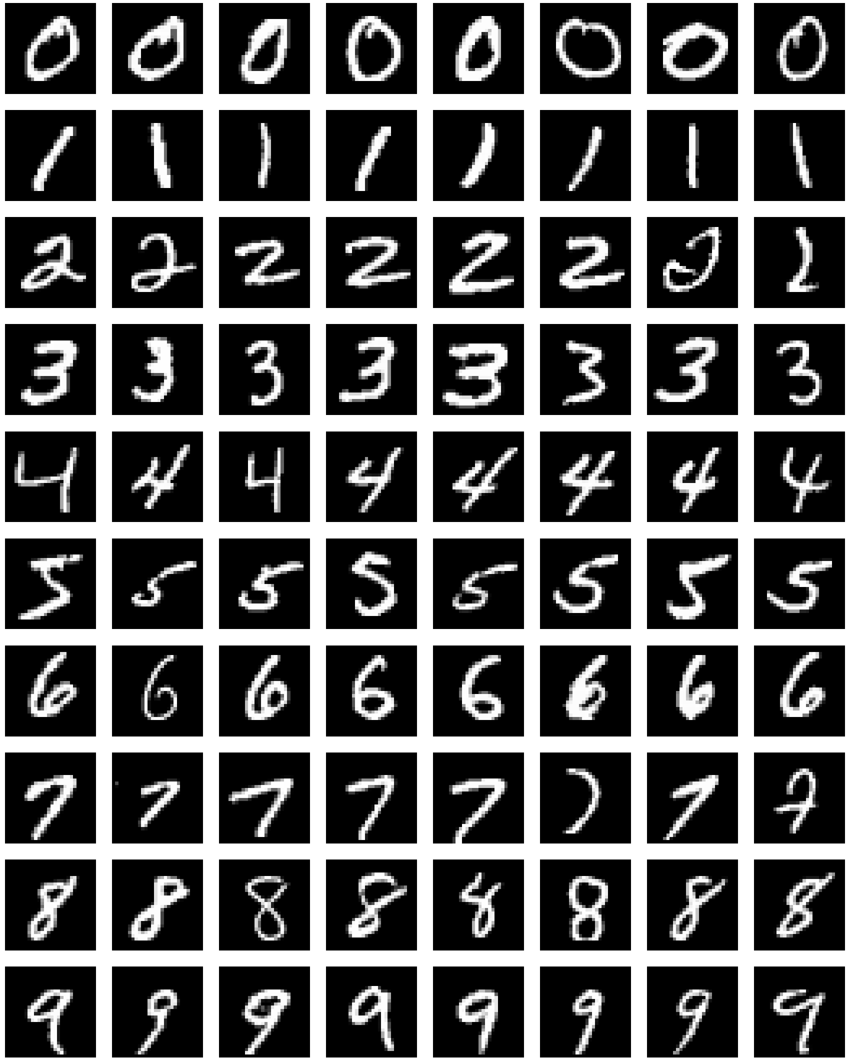
What is machine learning?

The classical MNIST problem:

- Classification of digits in postal codes
- Challenge:
  - Handwritings are very varied and ambiguous
  - Classical (human-crafted) algorithms are ineffective to deal with this variability

- Training data: real-world scanned postal codes that have been manually labelled
  - Labelling is likely not fully accurate
- Model: neural network
- Training process: optimize free parameters in the neural net



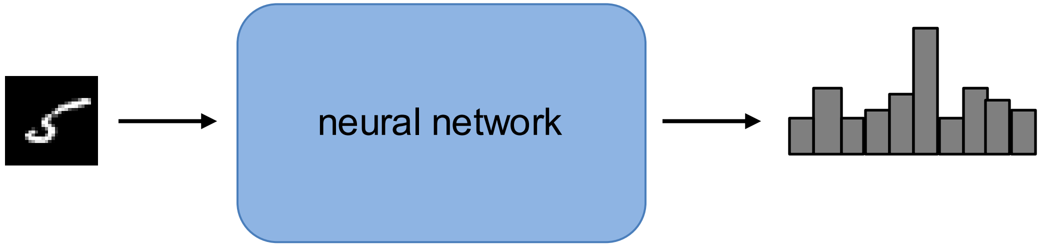


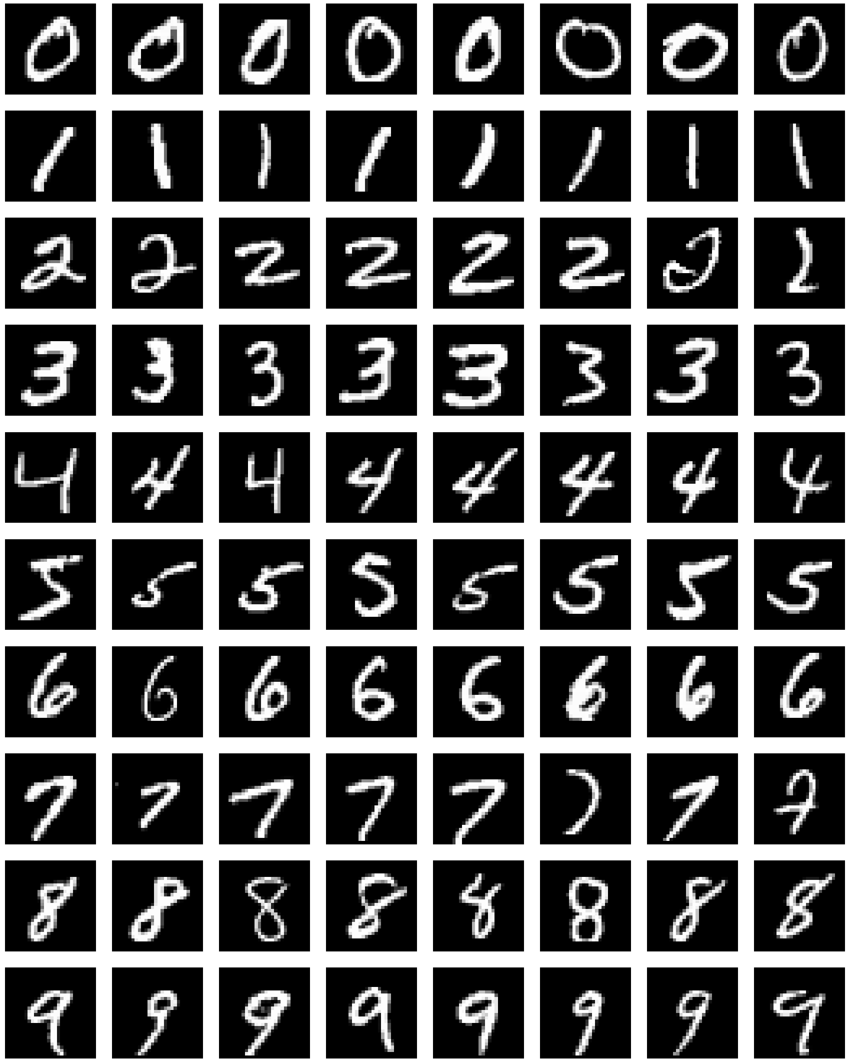
What is machine learning?

The classical MNIST problem:

- Classification of digits in postal codes
- Challenge:
    - Handwritings are very varied and ambiguous
    - Classical (human-crafted) algorithms are ineffective to deal with this variability

- Training data: real-world scanned postal codes that have been manually labelled
    - Labelling is likely not fully accurate
- Model: neural network
- Training process: optimize free parameters in the neural net
- Inference: predict prob over digits given an input image



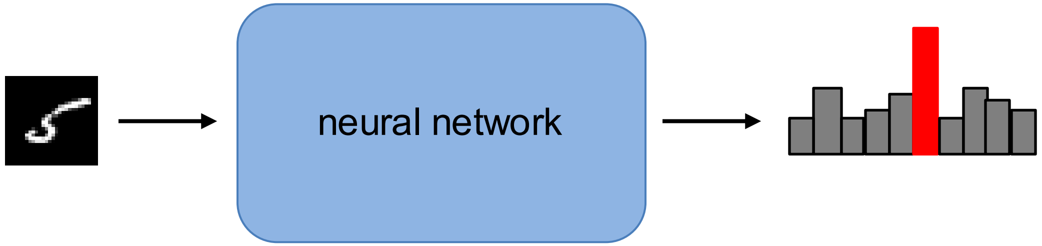


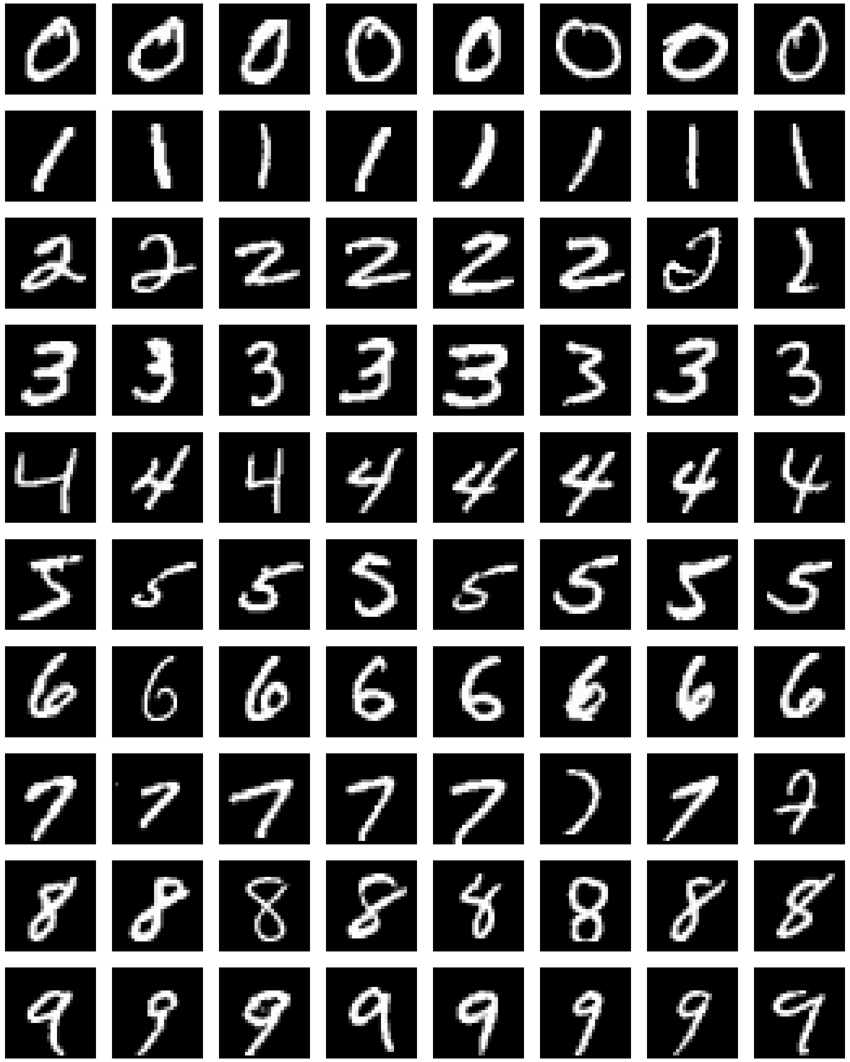
What is machine learning?

The classical MNIST problem:

- Classification of digits in postal codes
- Challenge:
    - Handwritings are very varied and ambiguous
    - Classical (human-crafted) algorithms are ineffective to deal with this variability

- Training data: real-world scanned postal codes that have been manually labelled
    - Labelling is likely not fully accurate
- Model: neural network
- Training process: optimize free parameters in the neural net
- Inference: predict prob over digits given an input image





What is machine learning?

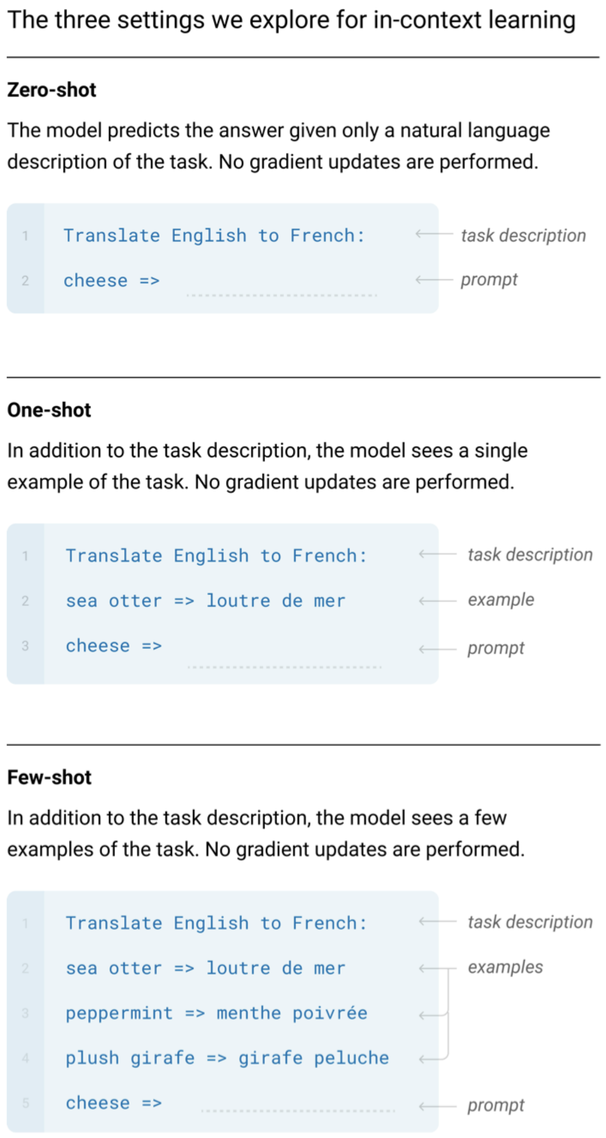
Large Language Models (LLMs):

- Machine translation and understanding was from the outset one of the motivation for artificial intelligence
    - Rule-based based expert/logical systems failed to effectively represent and process the versatility and ambiguity in human language

What is machine learning?

Large Language Models (LLMs):

- Machine translation and understanding was from the outset one of the motivation for artificial intelligence
  - Rule-based based expert/logical systems failed to effectively represent and process the versatility and ambiguity in human language
- GPT-2/3 showed a recipe to improve the skill of language models (around 2020):
  - Self-supervised learning to make use of large amounts of unlabeled text data
  - Sophisticated fine-tuning protocol to adapt LLMs to specific applications
  - Networks have emergent zero-shot abilities at scale, i.e. they can solve tasks they were not specifically trained for
  - Scaling laws: skill increases when data volume, compute and network size are scaled in tandem

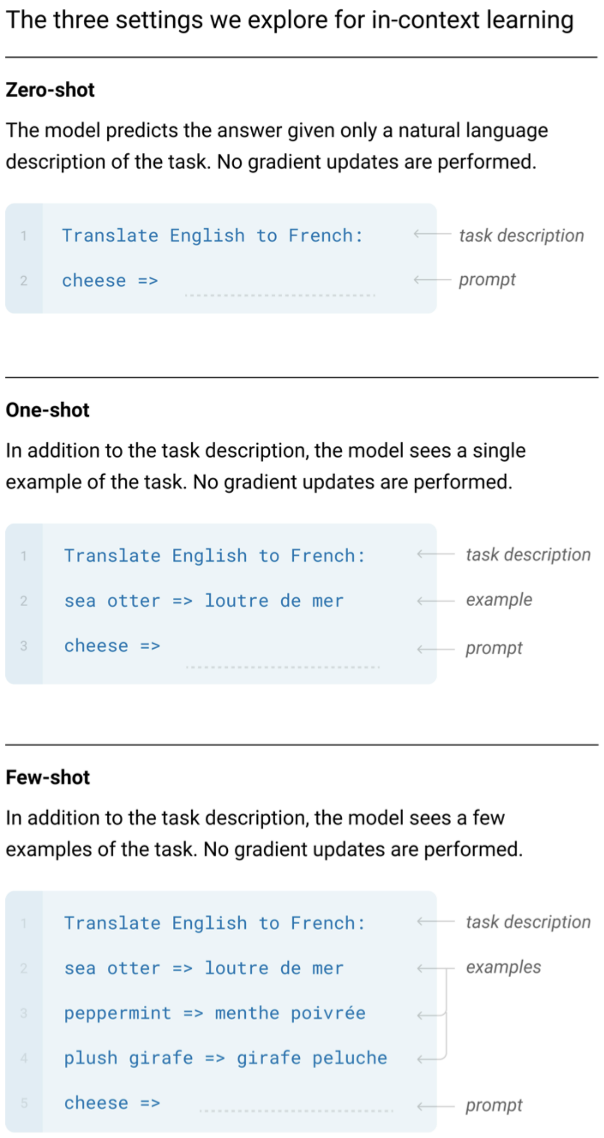


from https://arxiv.org/pdf/2005.14165

What is machine learning?

Large Language Models (LLMs):

- Machine translation and understanding was from the outset one of the motivation for artificial intelligence
    - Rule-based based expert/logical systems failed to effectively represent and process the versatility and ambiguity in human language
- GPT-2/3 showed a recipe to improve the skill of language models (around 2020):
    - Self-supervised learning to make use of large amounts of unlabeled text data
    - Sophisticated fine-tuning protocol to adapt LLMs to specific applications
    - Networks have emergent zero-shot abilities at scale, i.e. they can solve tasks they were not specifically trained for
    - **Scaling laws: skill increases when data volume, compute and network size are scaled in tandem**



from https://arxiv.org/pdf/2005.14165

What is machine learning?

The bitter lessons of machine learning (Rich Sutton):[1]

This is a big lesson. As a field, we still have not thoroughly learned it, as we are continuing to make the same kind of mistakes. To see this, and to effectively resist it, we have to understand the appeal of these mistakes. We have to learn the bitter lesson that building in how we think we think does not work in the long run. The bitter lesson is based on the historical observations that 1) AI researchers have often tried to build knowledge into their agents, 2) this always helps in the short term, and is personally satisfying to the researcher, but 3) in the long run it plateaus and even inhibits further progress, and 4) breakthrough progress eventually arrives by an opposing approach based on scaling computation by search and learning. The eventual success is tinged with bitterness, and often incompletely digested, because it is success over a favored, human-centric approach.

[1] http://www.incompleteideas.net/IncIdeas/BitterLesson.html

What is machine learning?

The bitter lessons of machine learning (Rich Sutton):[1]

This is a big lesson. As a field, we still have not thoroughly lear
same kind of mistakes. To see this, and to effectively resist it, we
mistakes. We have to learn the bitter lesson that building in how
long run. The bitter lesson is based on the historical observation
build knowledge into their agents, 2) this always helps in the sh
the researcher, but 3) in the long run it plateaus and even inhibi
progress eventually arrives by an opposing approach based on
learning. The eventual success is tinged with bitterness, and oft
success over a favored, human-centric approach.

ake the
of these
rk in the
en tried to
fying to
through
nd
se it is

Discussion

Our work reinforces the bitter lesson. The most important factors determining the performance of a sensibly designed model are the compute and data available for training[5] (Tolstikhin et al., 2021). Although the success of ViTs in computer vision is extremely impressive, in our view there is no strong evidence to suggest that pre-trained ViTs outperform pre-trained ConvNets when evaluated fairly. We note however that ViTs may have practical advantages in specific contexts, such as the ability to use similar model components across multiple modalities (Bavishi et al., 2023).

Acknowledgements

We thank Lucas Beyer and Olivier Henaff for feedback on an earlier draft of this note. We also thank Lucas Beyer for providing training speed estimates for ViT models on TPU-v4 devices.

References

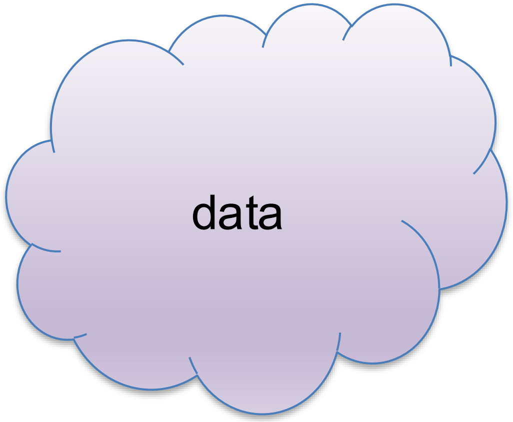
I. Alabdulmohsin, X. Zhai, A. Kolesnikov, and L. Beyer. Getting vit in shape: Scaling laws for compute-optimal model design. *arXiv preprint arXiv:2305.13035*, 2023.

[5] By sensibly designed, we mean models that are sufficiently expressive and have stable gradient propagation.

[1] http://www.incompleteideas.net/IncIdeas/BitterLesson.html

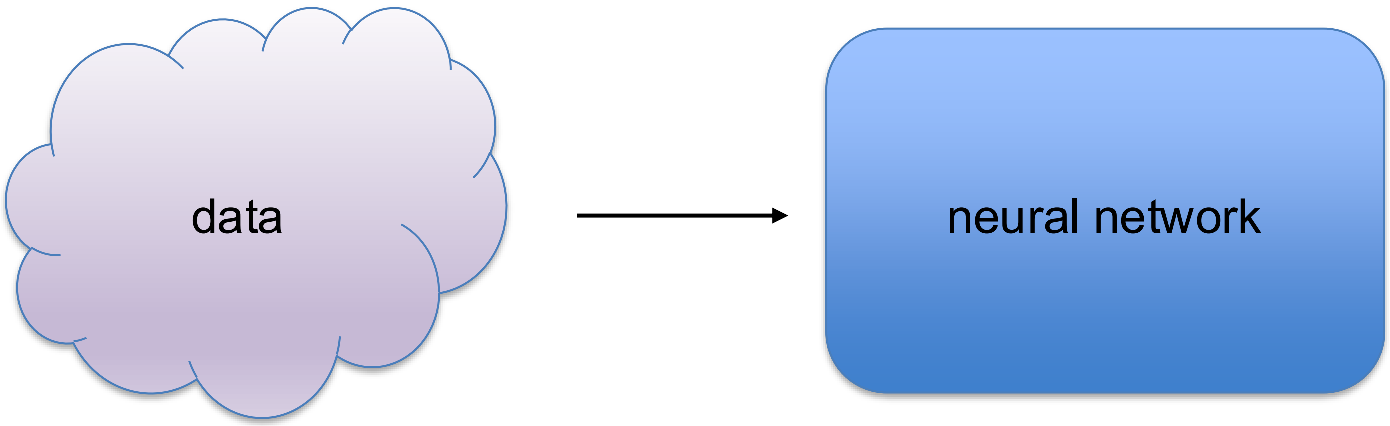
[2] https://arxiv.org/abs/2310.16764

What is machine learning?



the stuff to model
(usually with ambiguity
and large variability)

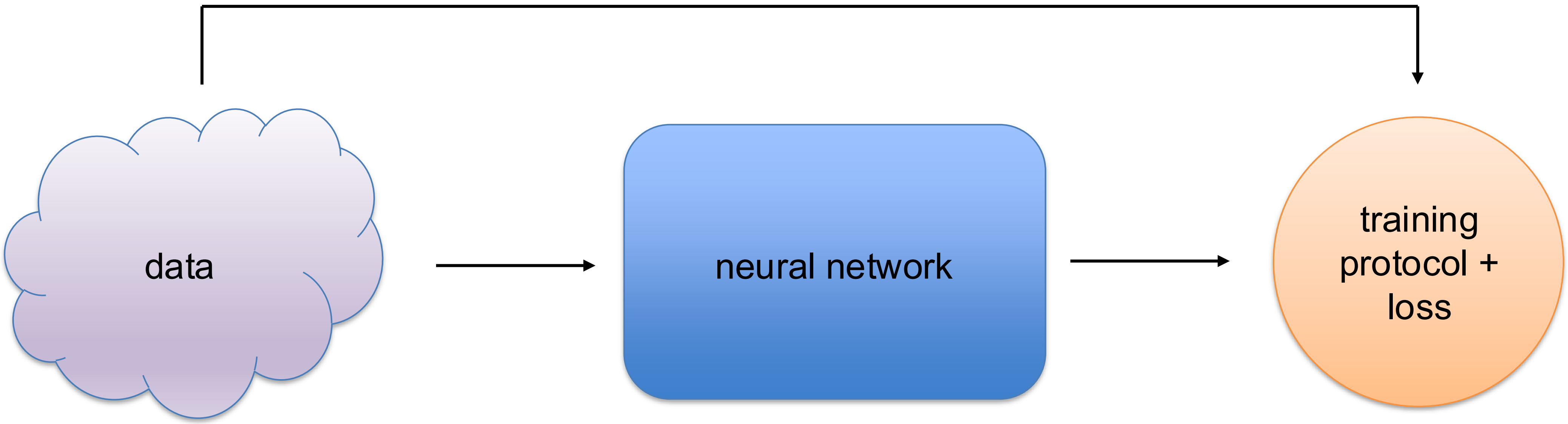
What is machine learning?



the stuff to model
(usually with ambiguity
and large variability)

a lot of free parameters combined
weakly nonlinearly to model
the data

What is machine learning?

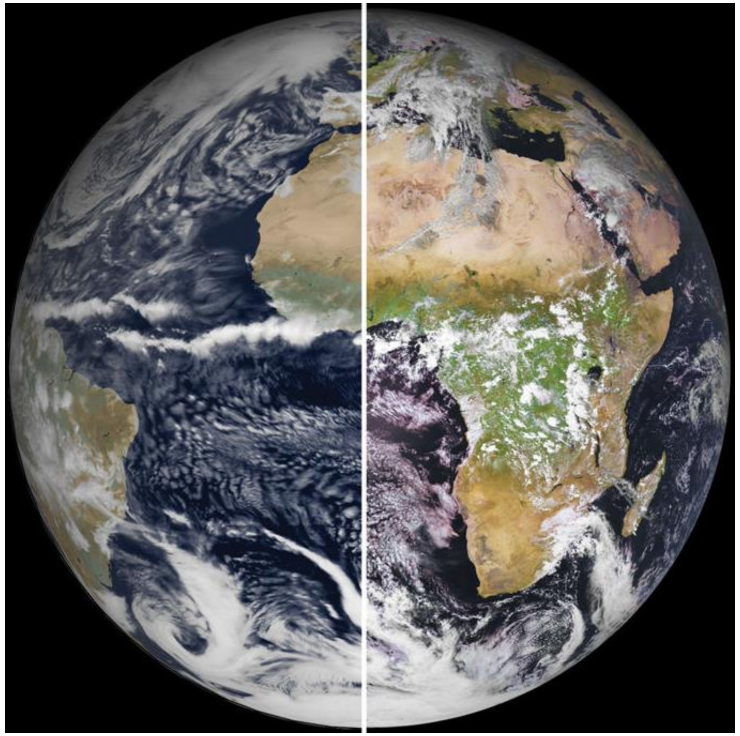


the stuff to model (usually with ambiguity and large variability)

a lot of free parameters combined weakly nonlinearly to model the data

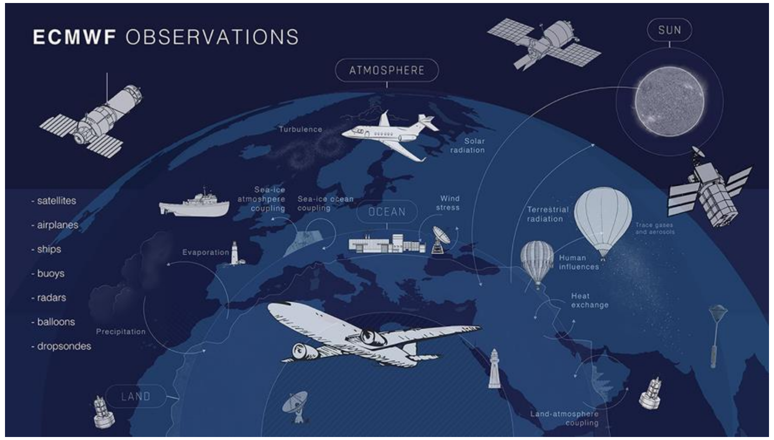
nonlinear optimization that can be computed efficiently

Machine learning for weather forecasting



Predicting the weather is different from classical ML challenges:

- Powerful conventional numerical weather prediction models
    - Physics of most processes in the Earth system is well understood
    - Equation-based models provide highly skilful weather forecasts



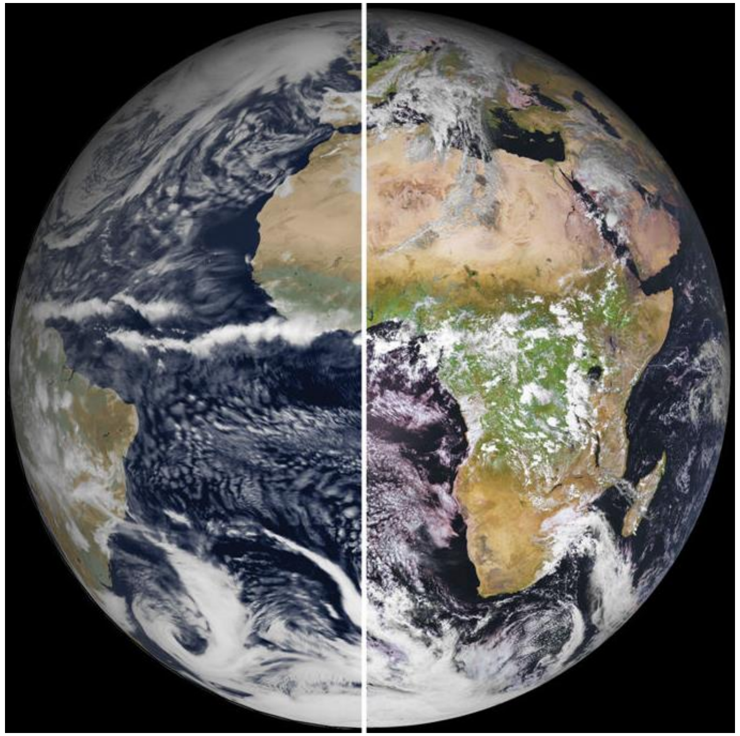
Machine learning for weather forecasting

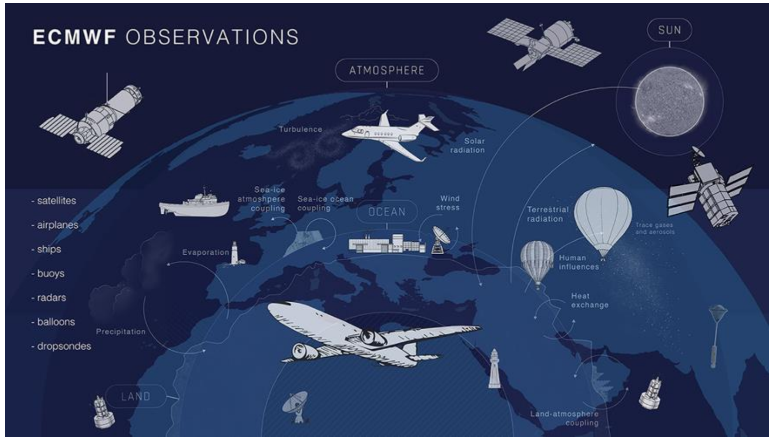
Predicting the weather is different from classical ML challenges:

- Powerful conventional numerical weather prediction models
  - Physics of most processes in the Earth system is well understood
  - Equation-based models provide highly skilful weather forecasts

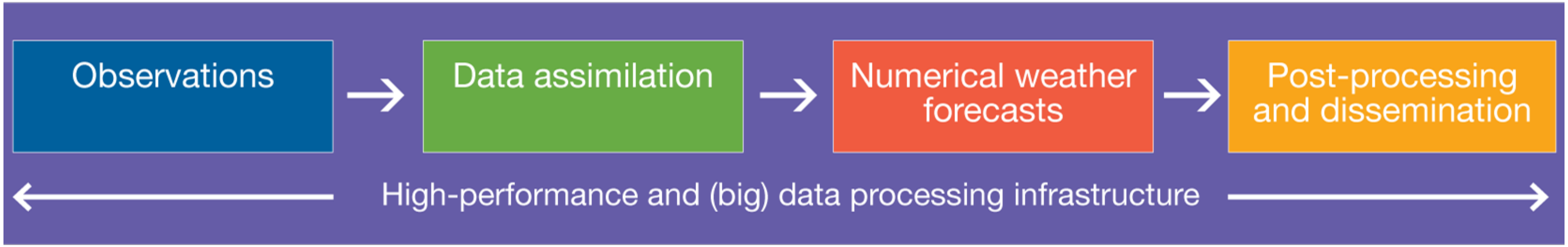
Why is ML for weather prediction an opportunity?

- Weather prediction is inherently uncertain (has ambiguity)
- Modeling of many small-scale processes and their interaction across the entire globe is highly challenging with classical approaches
- Large amounts of data
  - Reanalysis (ERA5 is in total about 6 PB)
  - Forecast archives
  - Observations, e.g. modern satellites records 10s TB of data every day, which are difficult to exploit with conventional approaches

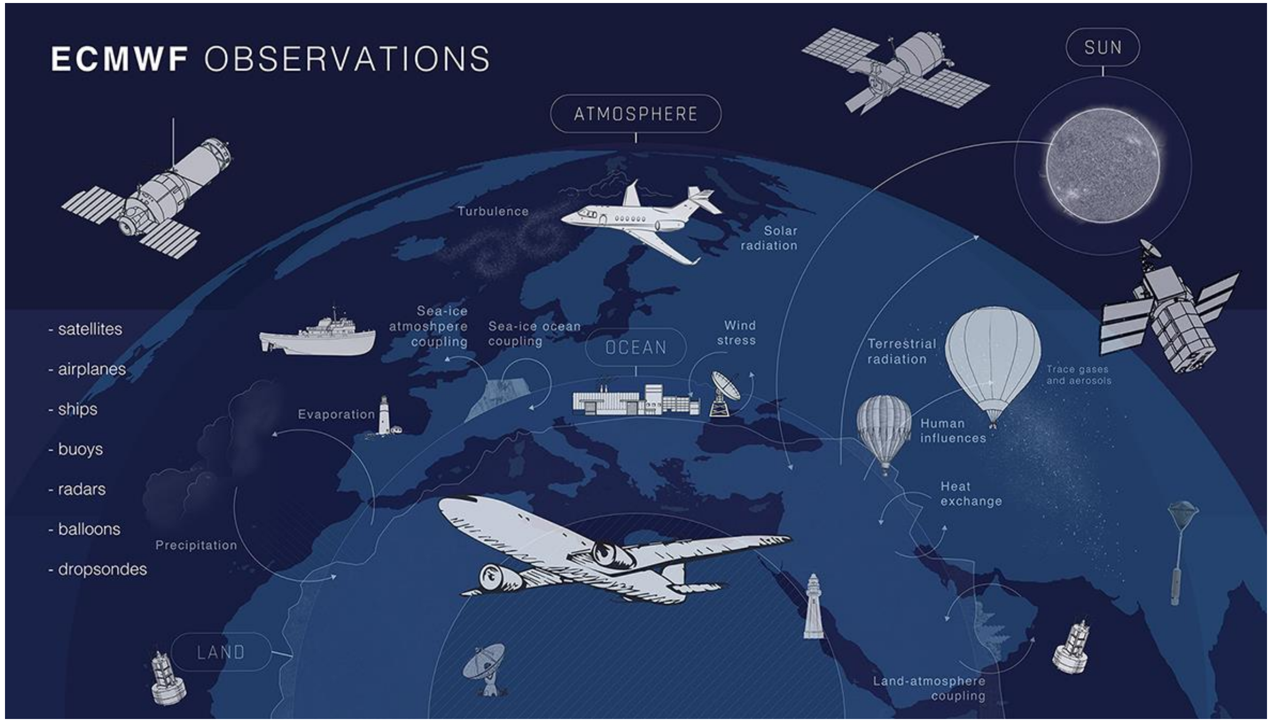




Conventional numerical weather prediction pipeline



Conventional numerical weather prediction pipeline

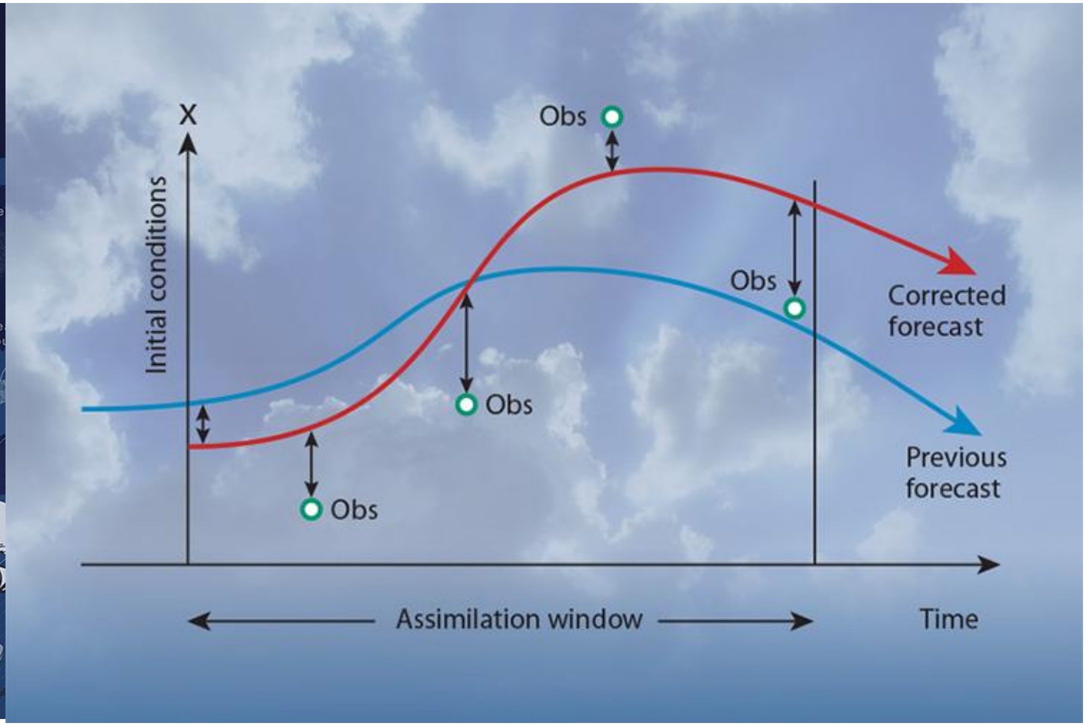


Observations → Data assimilation → Numerical weather forecasts → Post-processing and dissemination

← High-performance and (big) data processing infrastructure →

ECMWF

Conventional numerical weather prediction pipeline

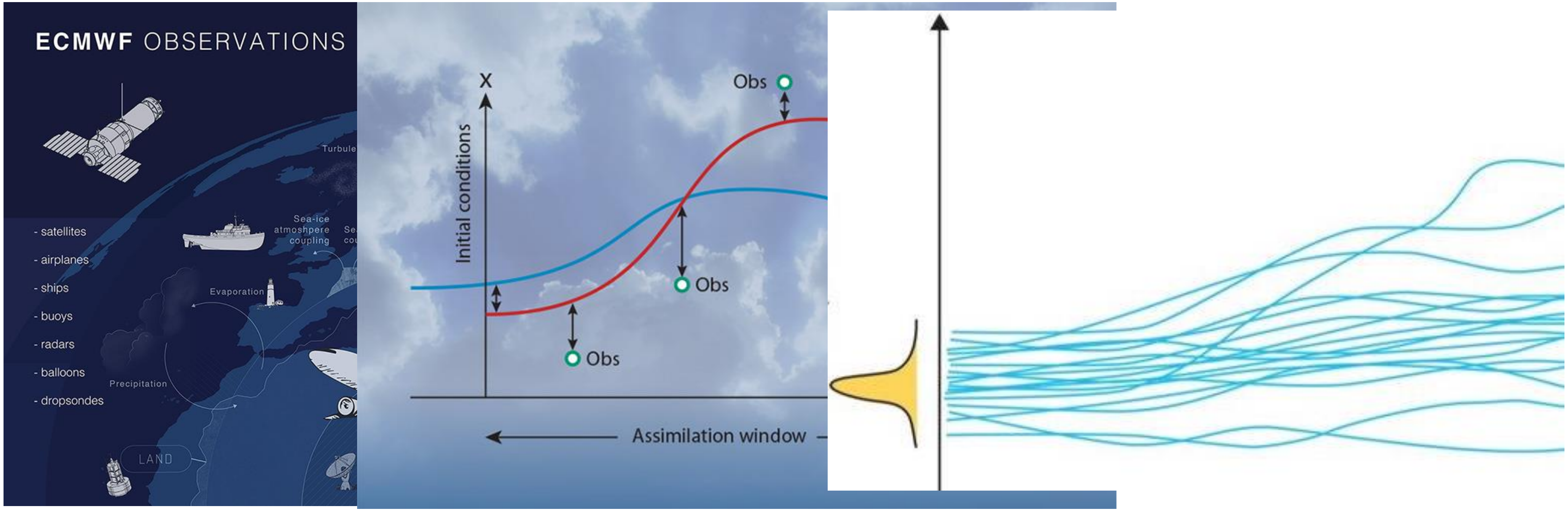


Observations → Data assimilation → Numerical weather forecasts → Post-processing and dissemination

← High-performance and (big) data processing infrastructure →

ECMWF

Conventional numerical weather prediction pipeline



https://www.researchgate.net/profile/Peter-Groenquist/publication/337020034/figure/fig1/AS:821718931214336@1572924468990/ECMWF-Ensemble-Prediction-System.jpg

Observations → Data assimilation → Numerical weather forecasts → Post-processing and dissemination

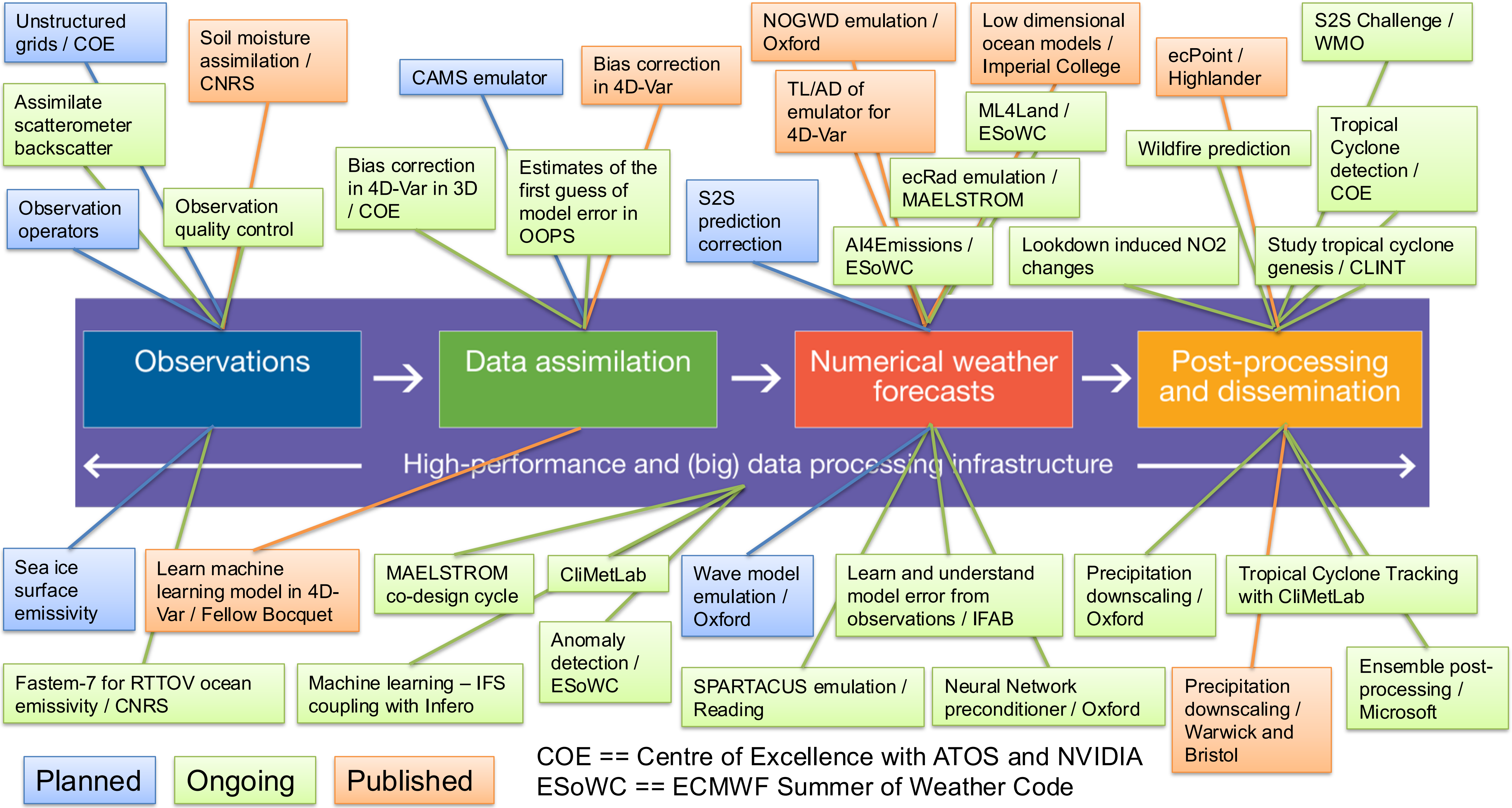
← High-performance and (big) data processing infrastructure →

ECMWF

Machine learning for weather forecasting: a short history

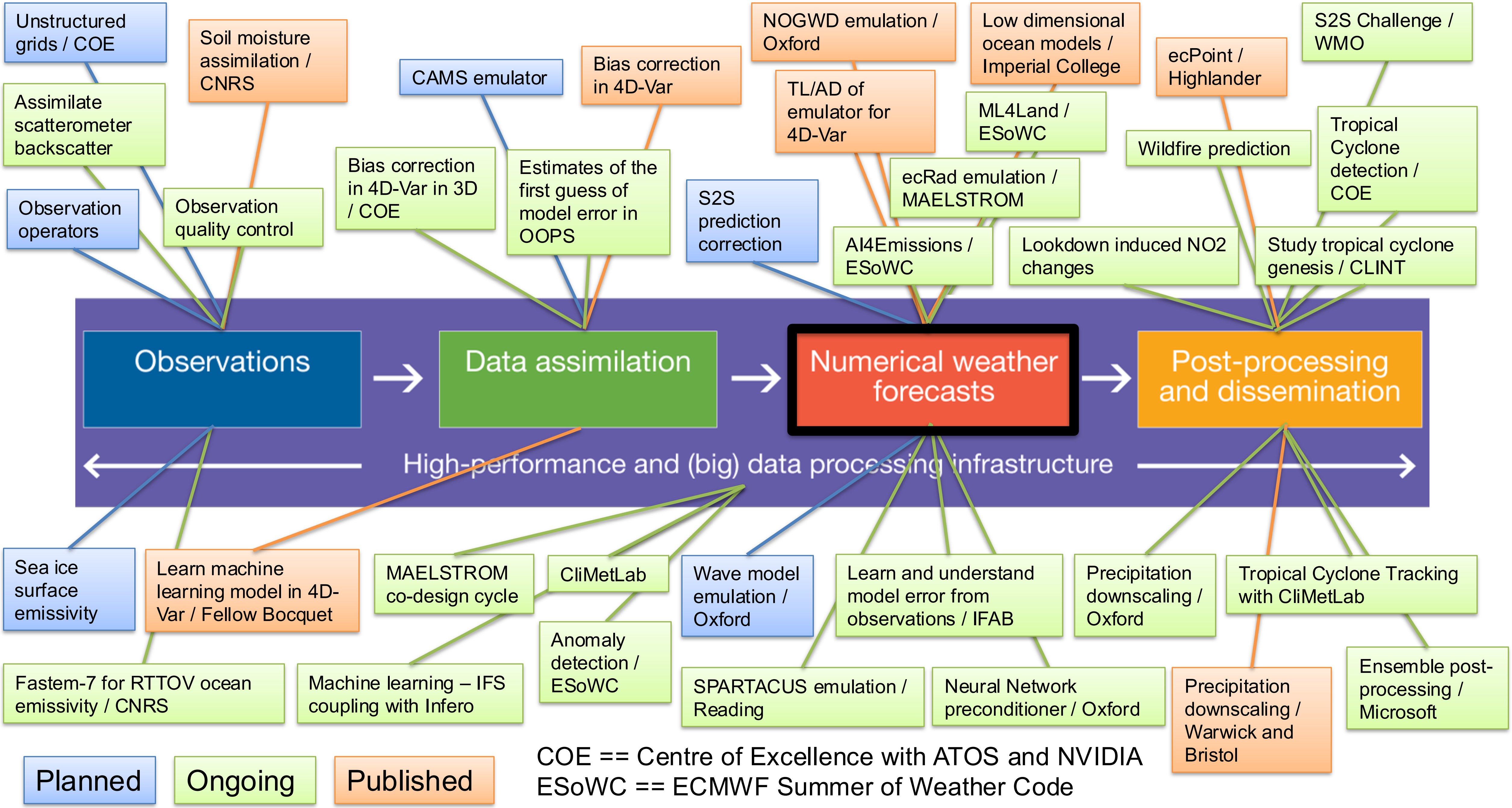
- 1990s-present: Neural networks have been explored and used in Earth system modeling since the 1990s
  - Used typically as data-optimized interpolation method
- 2010-present: Learning parametrizations
  - Philosophy: we understand large scale motion in the atmosphere but struggle with representing
  - Still many efforts in this direction

Machine learning at ECMWF



COE == Centre of Excellence with ATOS and NVIDIA
ESoWC == ECMWF Summer of Weather Code

Machine learning at ECMWF



COE == Centre of Excellence with ATOS and NVIDIA
ESoWC == ECMWF Summer of Weather Code

Machine learning for weather forecasting: a short history

- 1990s-present: Neural networks have been explored and used in Earth system modeling since the 1990s
  - Used typically as data-optimized interpolation method
- 2010-present: Learning parametrizations
  - Philosophy: we understand large scale motion in the atmosphere but struggle with representing
  - Still many efforts in this direction

- 2018: Dueben and Bauer: First attempt for pure neural network-based forecasting
- 2020/21: Weyn, Duran et al.: Skilful ensemble forecast close to IFS skill
- 2022: Keisler: graph neural network approach with skill close to IFS

Machine learning for weather forecasting: a short history

- 1990s-present: Neural networks have been explored and used in Earth system modeling since the 1990s
    - Used typically as data-optimized interpolation method
- 2010-: Learning parametrizations
    - Philosophy: we understand large scale motion in the atmosphere but struggle with representing
    - Still many efforts in this direction

- 2018: Dueben and Bauer: First attempt for pure neural network-based forecasting
- 2020/21: Weyn, Duran et al.: Skilful ensemble forecast close to IFS skill
- 2022: Keisler: graph neural network approach with skill close to IFS

inflection point

Machine learning for weather forecasting: a short history

- 1990s-present: Neural networks have been explored and used in Earth system modeling since the 1990s
  - Used typically as data-optimized interpolation method
- 2010-: Learning parametrizations
  - Philosophy: we understand large scale motion in the atmosphere but struggle with representing
  - Still many efforts in this direction

- 2018: Dueben and Bauer: First attempt for pure neural network-based forecasting
- 2020/21: Weyn, Duran et al.: Skilful ensemble forecast close to IFS skill
- 2022: Keisler: graph neural network approach with skill close to IFS

inflection point

"To do [medium range weather forecasting] with the same level of skill using AI would likely require an exceptional (and hence unrealistic) amount of training data. […]", Tim Palmer (https://arxiv.org/abs/2007.04830)

Machine learning for weather forecasting: a short history

- 1990s-present: Neural networks have been explored and used in Earth system modeling since the 1990s
  - Used typically as data-optimized interpolation method
- 2010-: Learning parametrizations
  - Philosophy: we understand large scale motion in the atmosphere but struggle with representing
  - Still many efforts in this direction

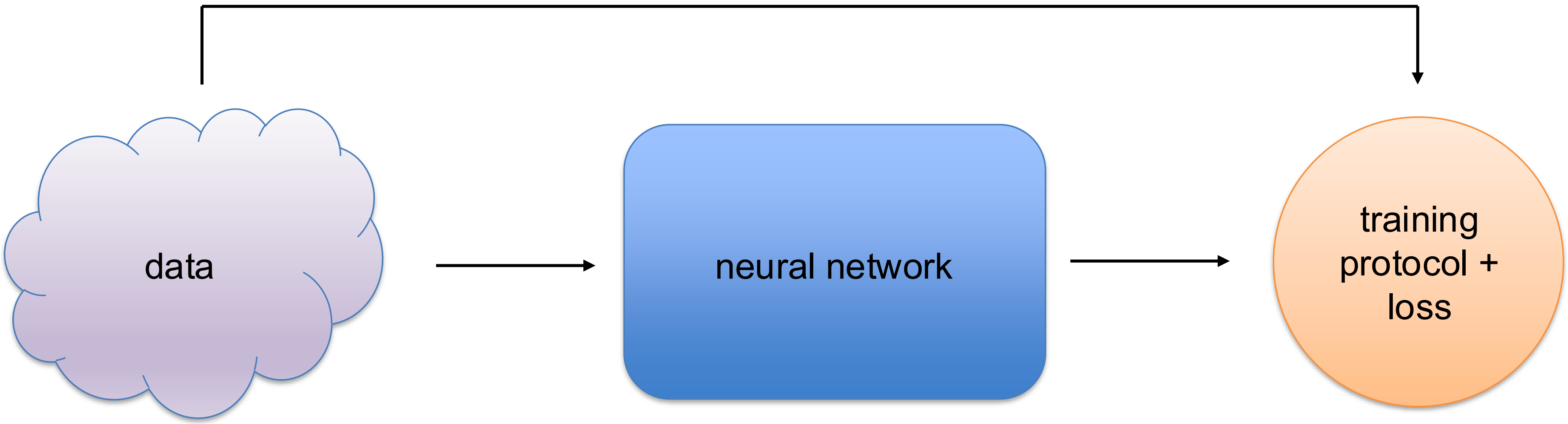
- 2018: Dueben and Bauer: First attempt for pure neural network-based forecasting
- 2020/21: Weyn, Duran et al.: Skilful ensemble forecast close to IFS skill
- 2022: Keisler: graph neural network approach with skill close to IFS

inflection point

- 2023: PanguWeather, GraphCast: First models whose skill surpasses operational IFS in headline scores
- 2024: NeuralGCM: hybrid model surpassing IFS
- 2025: AIFS: first operational weather forecasting model
- 2024/2025: GenCast and AIFS-Ensemble: AI-based ensemble methods surpassing IFS-ENS

Machine Learning-based Weather Forecasting

Machine learning-based weather forecasting

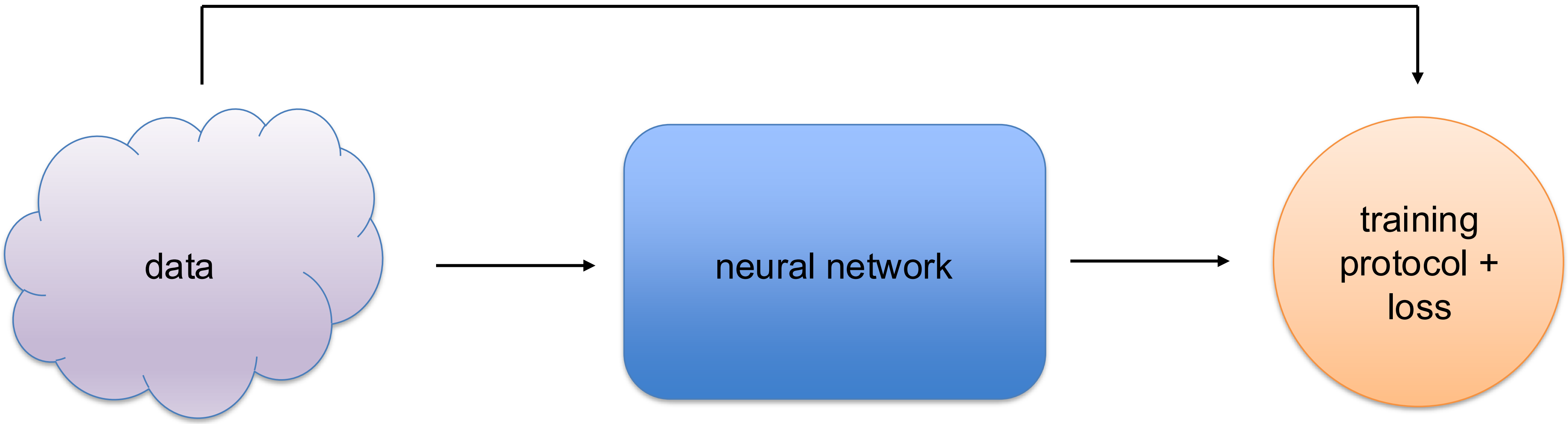


the stuff to model (usually with ambiguity and large variability)

a lot of free parameters combined weakly nonlinearly to model the data

nonlinear optimization that can be computed efficiently

Machine learning-based weather forecasting

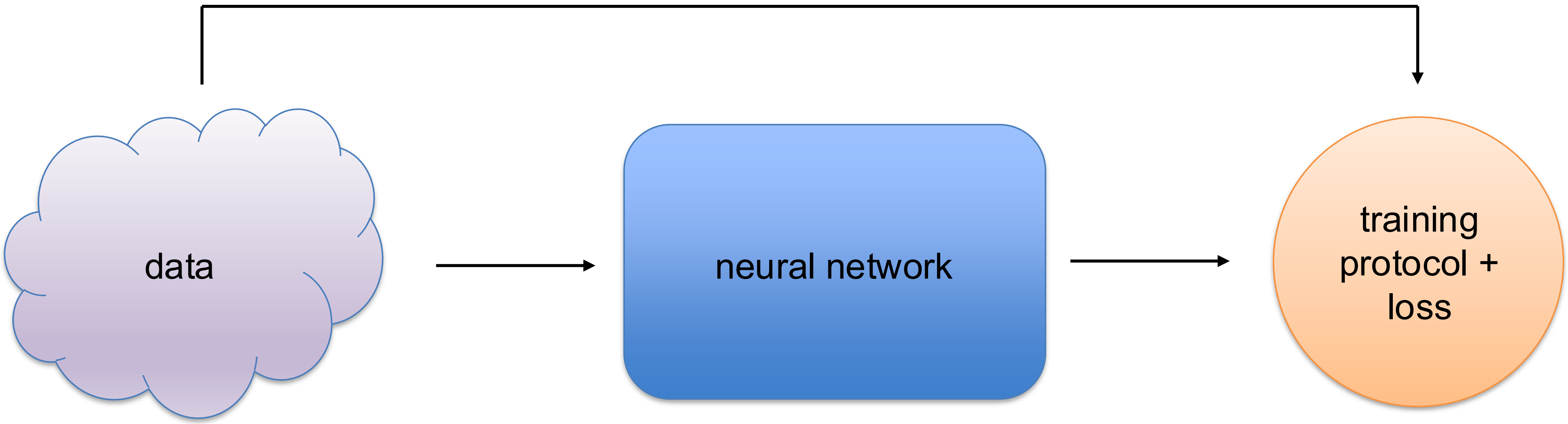


reanalysis / analysis
(e.g. ERA5)

a lot of free parameters combined
weakly nonlinearly to model
the data

nonlinear optimization
that can be computed
efficiently

Machine learning-based weather forecasting

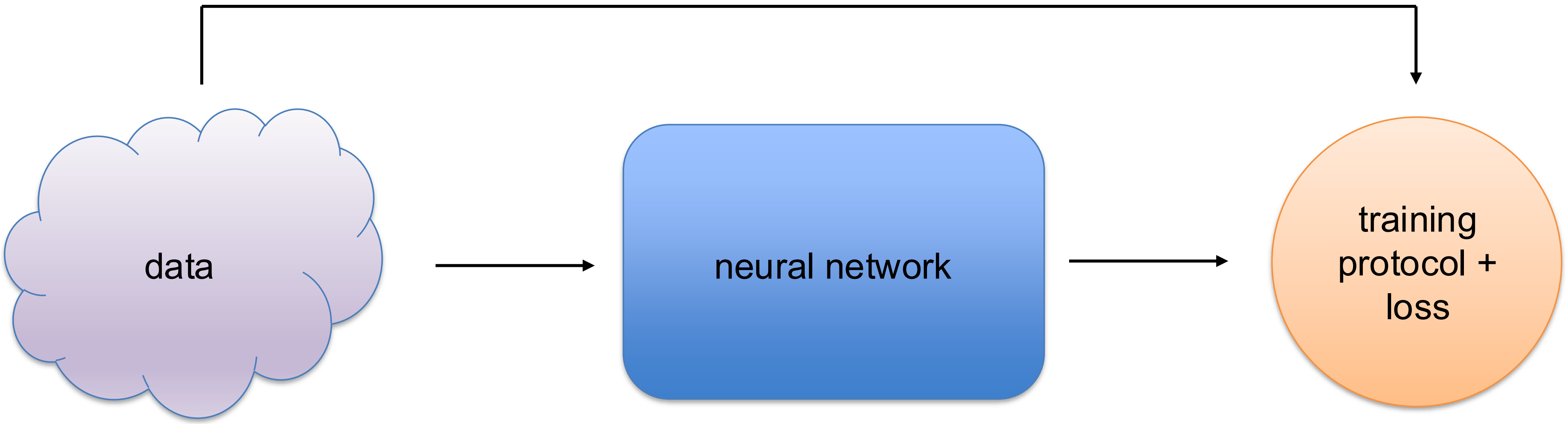


reanalysis / analysis
(e.g. ERA5)

50-200 million trainable parameters
(transformers, graph neural nets, conv-nets, …)

nonlinear optimization that can be computed efficiently

Machine learning-based weather forecasting

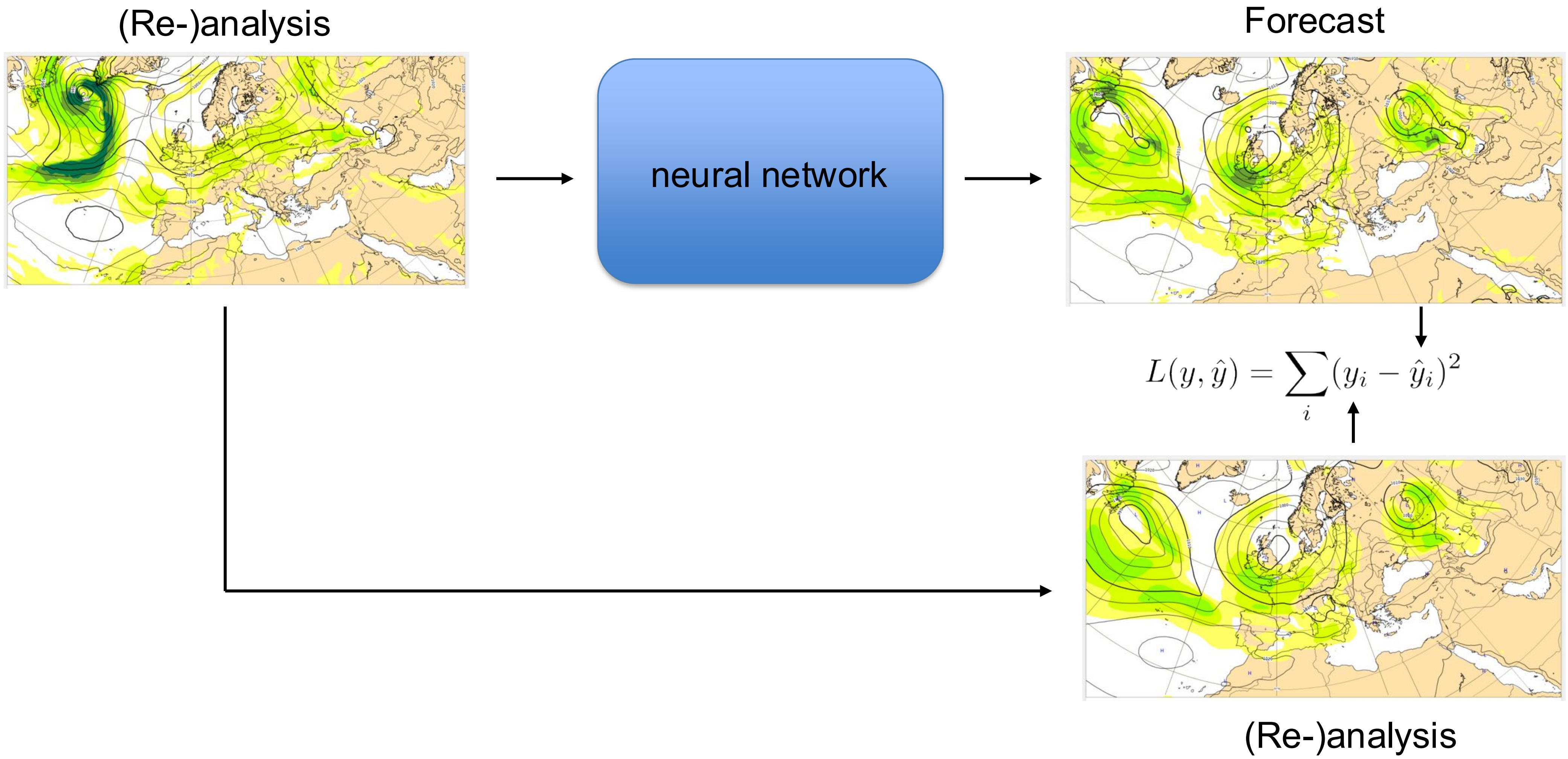


reanalysis / analysis
(e.g. ERA5)

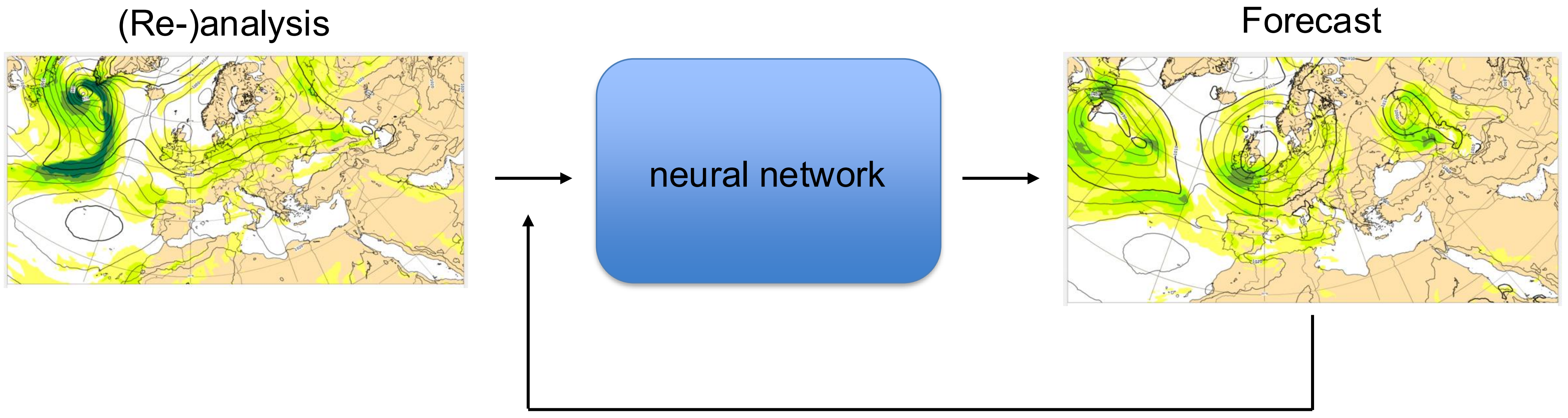
50-200 million trainable parameters
(transformers, graph neural nets,
conv-nets, …)

minimize MSE
between prediction
and training data

Machine learning-based weather forecasting

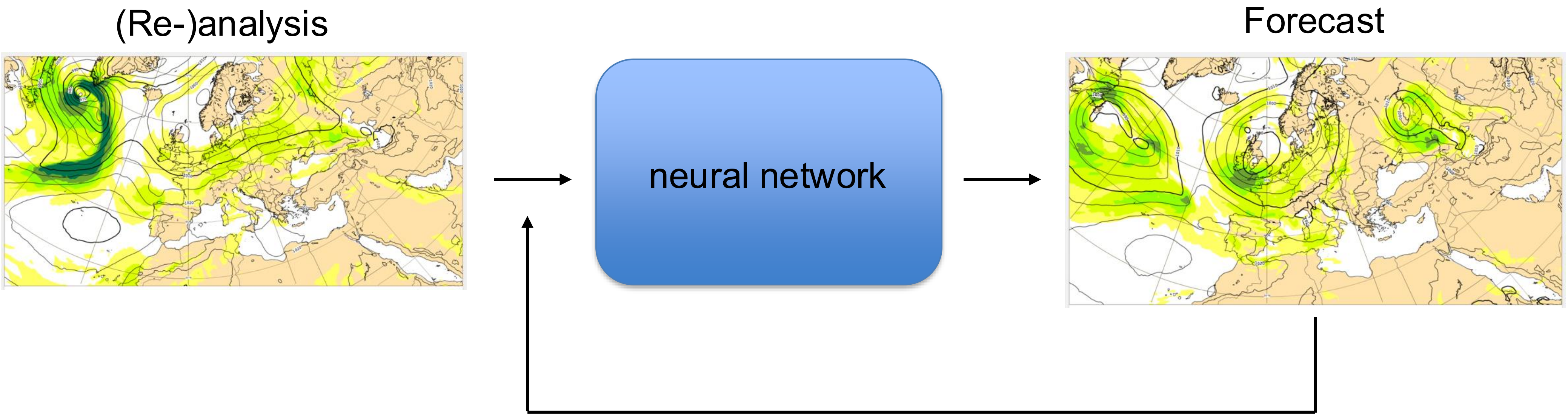


Machine learning-based weather forecasting



time stepping similar to existing models:
forecast is used as input for neural net for
next time step

Machine learning-based weather forecasting

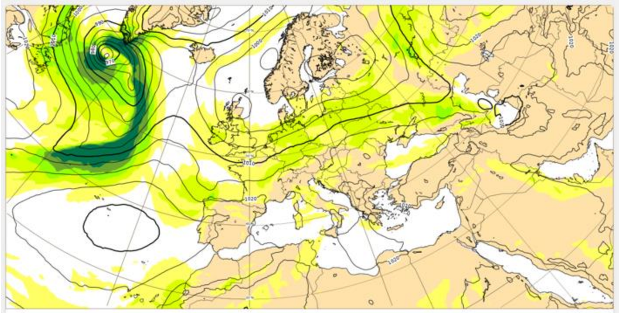


time stepping similar to existing models:
forecast is used as input for neural net for
next time step

Time step: 6h or longer
(compared to O(min) for conventional models

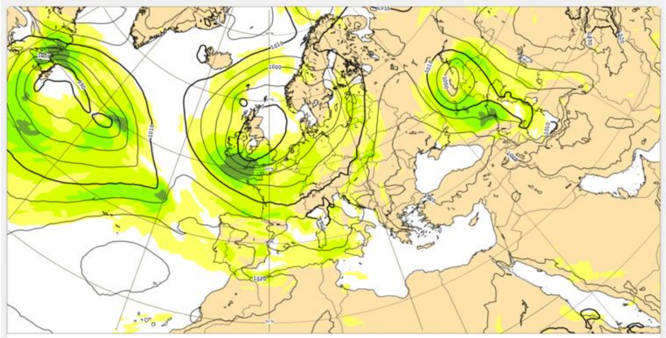
Machine learning-based weather forecasting

(Re-)analysis



neural network

Forecast



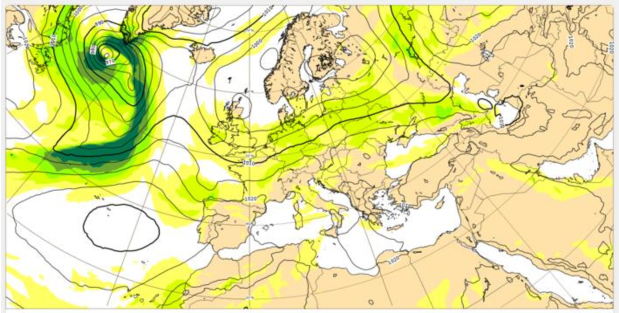
Fine-tuning on operational analysis to avoid mismatch between training and inference time

time stepping similar to existing models: forecast is used as input for neural net for next time step

Time step: 6h or longer
(compared to O(min) for conventional models

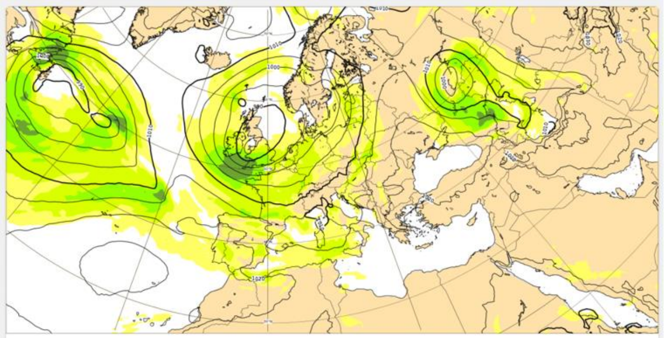
Machine learning-based weather forecasting

(Re-)analysis



neural network

Forecast



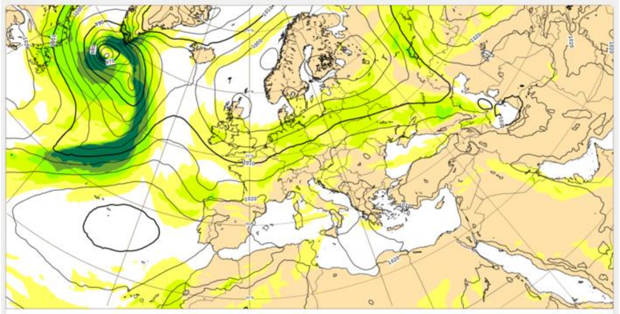
Fine-tuning on operational analysis to avoid mismatch between training and inference time

time stepping similar to existing models: forecast is used as input for neural net for next time step

Time step: 6h or longer
(compared to O(min) for conventional models

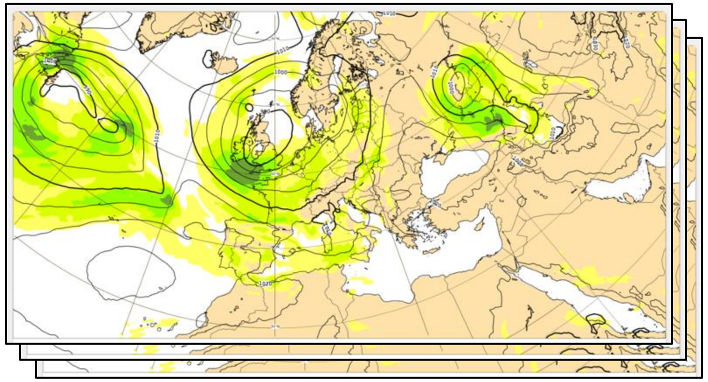
Machine learning-based weather forecasting

(Re-)analysis



neural network

Forecast



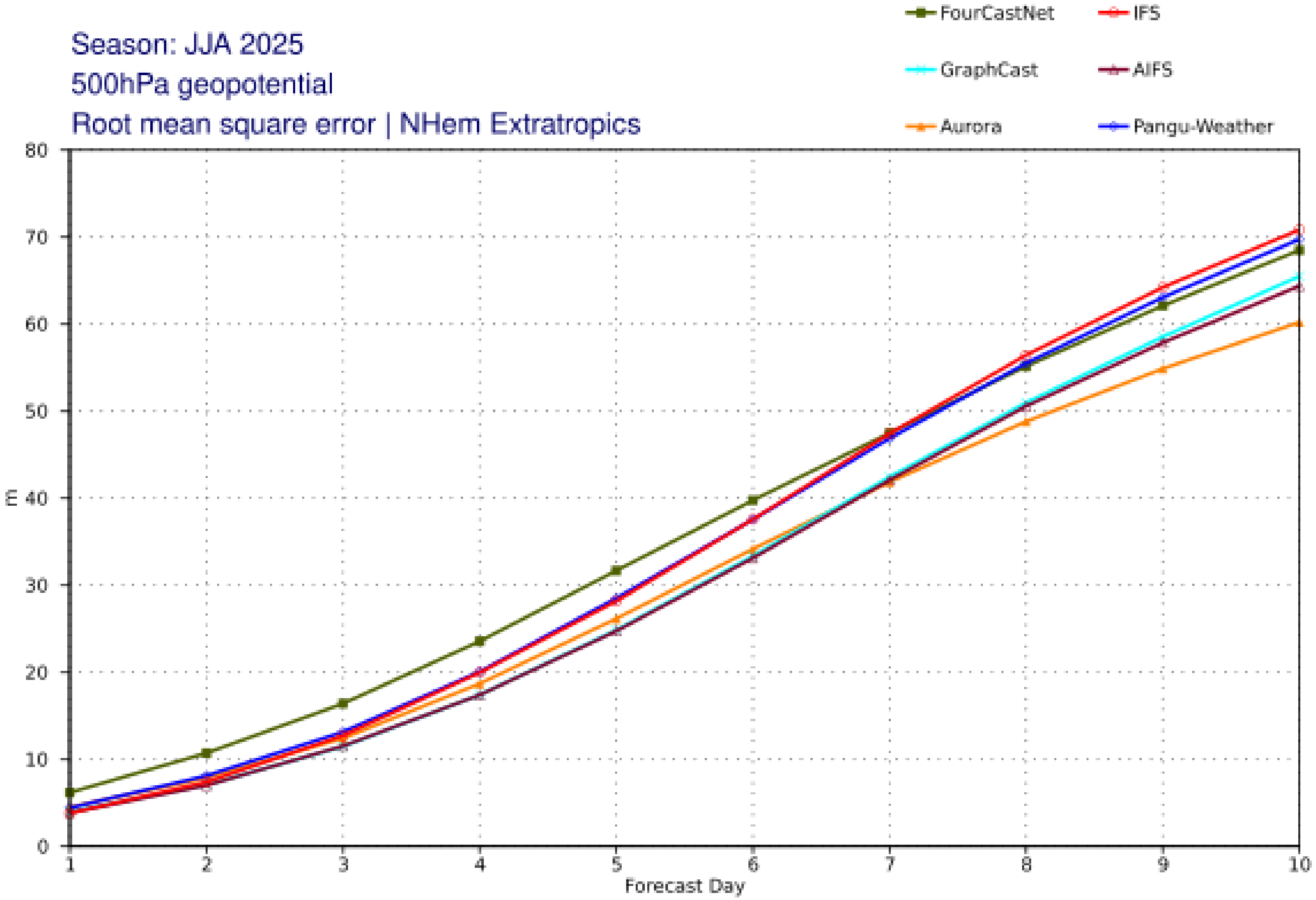
ensemble methods

Fine-tuning on operational analysis to avoid mismatch between training and inference time

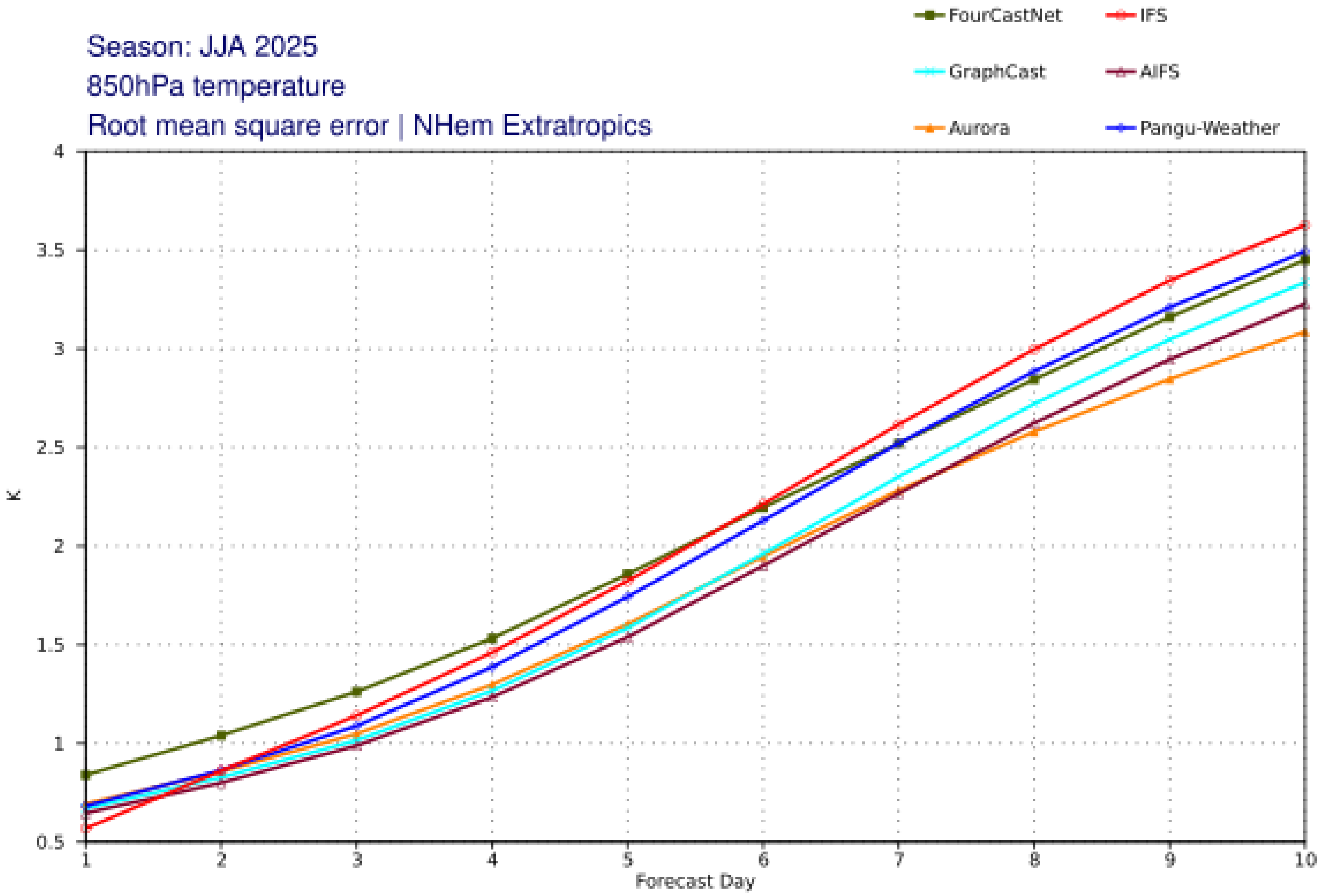
time stepping similar to existing models: forecast is used as input for neural net for next time step

Time step: 6h or longer
(compared to O(min) for conventional models

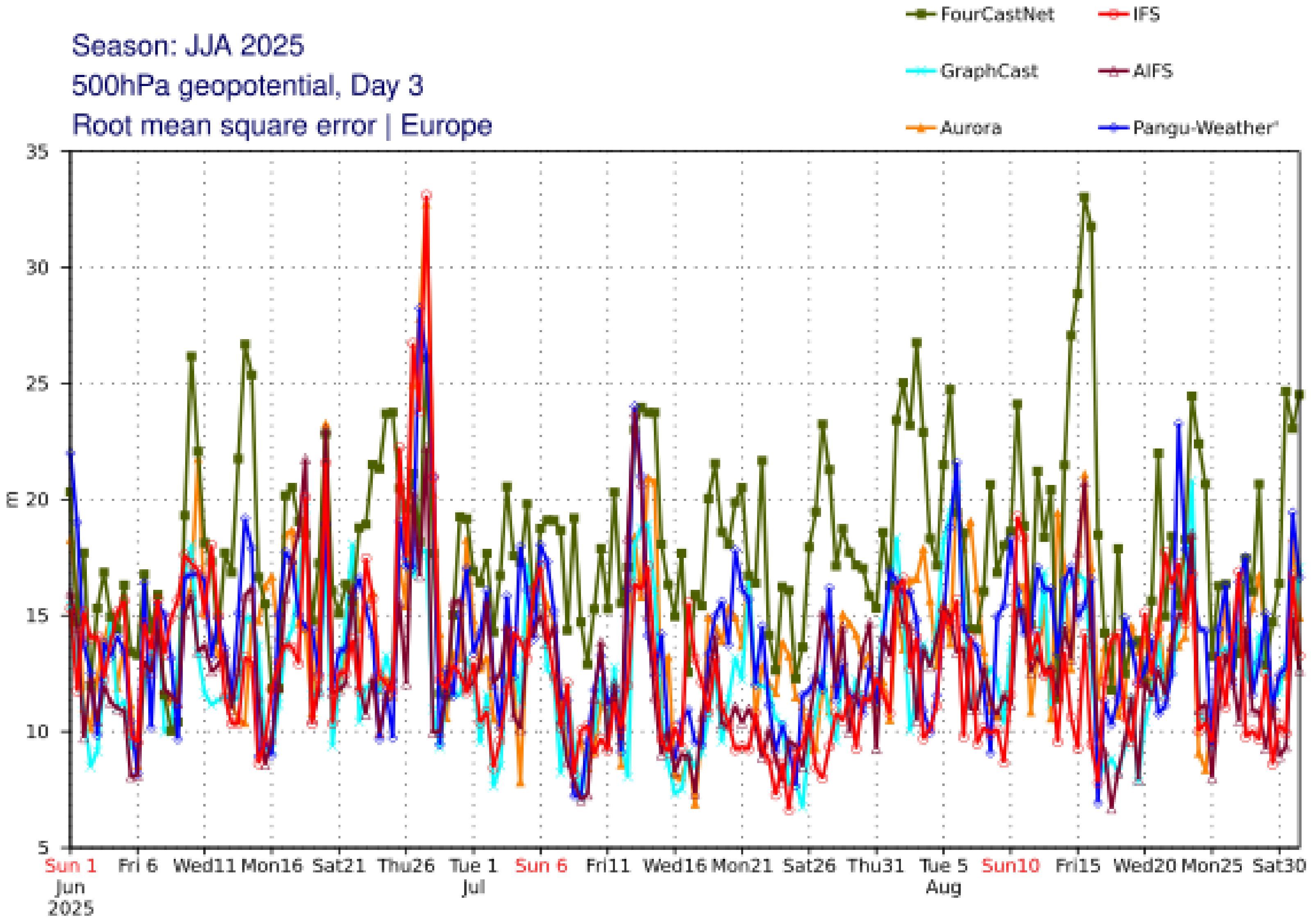
Machine learning-based weather forecasting



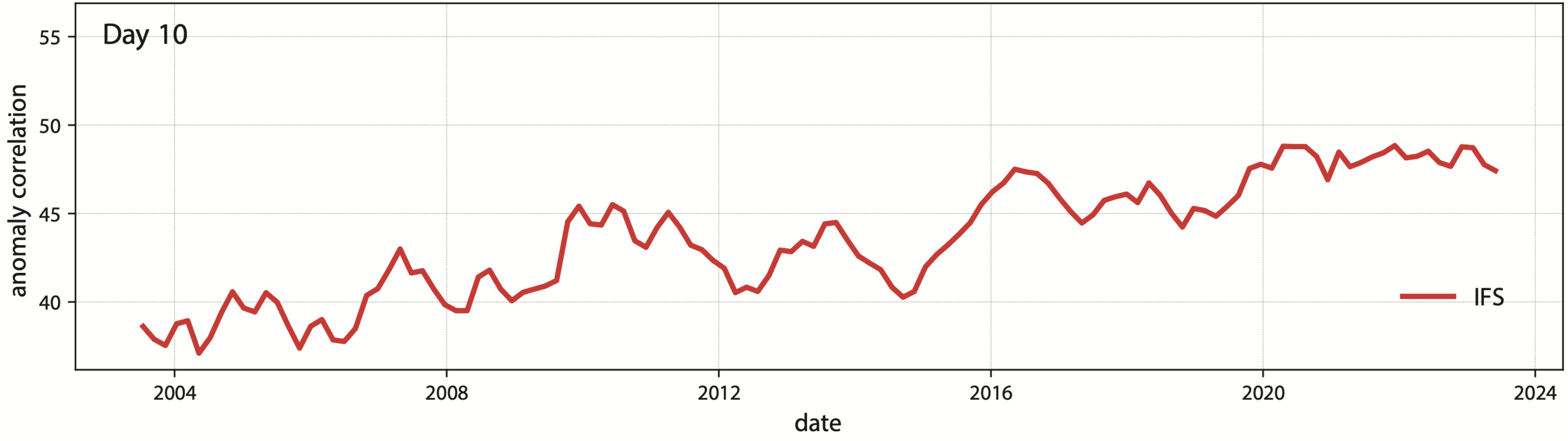
Machine learning-based weather forecasting



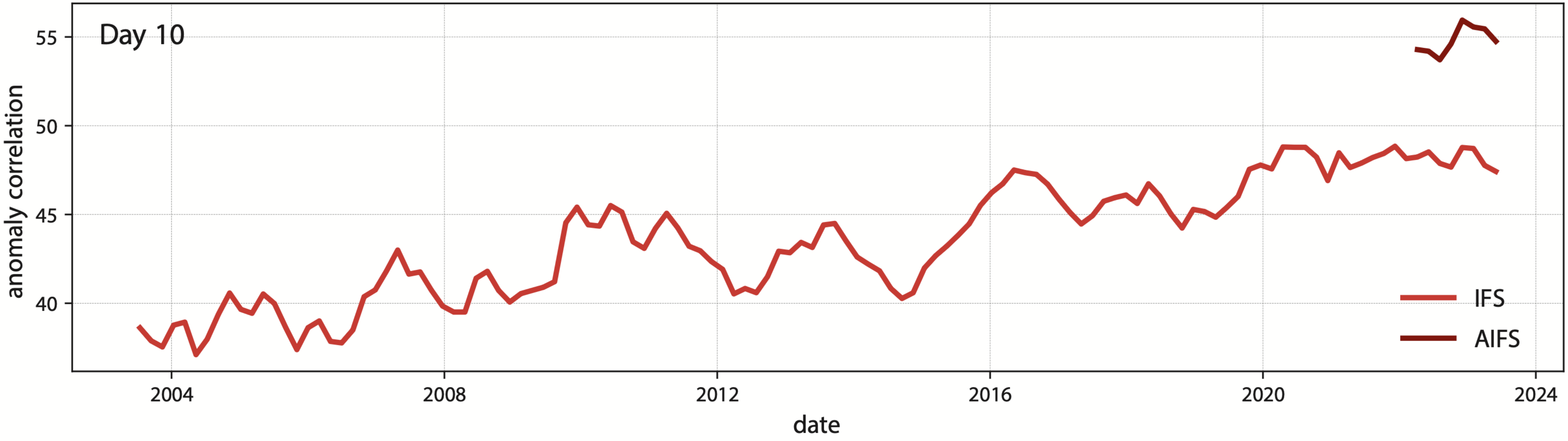
Machine learning-based weather forecasting



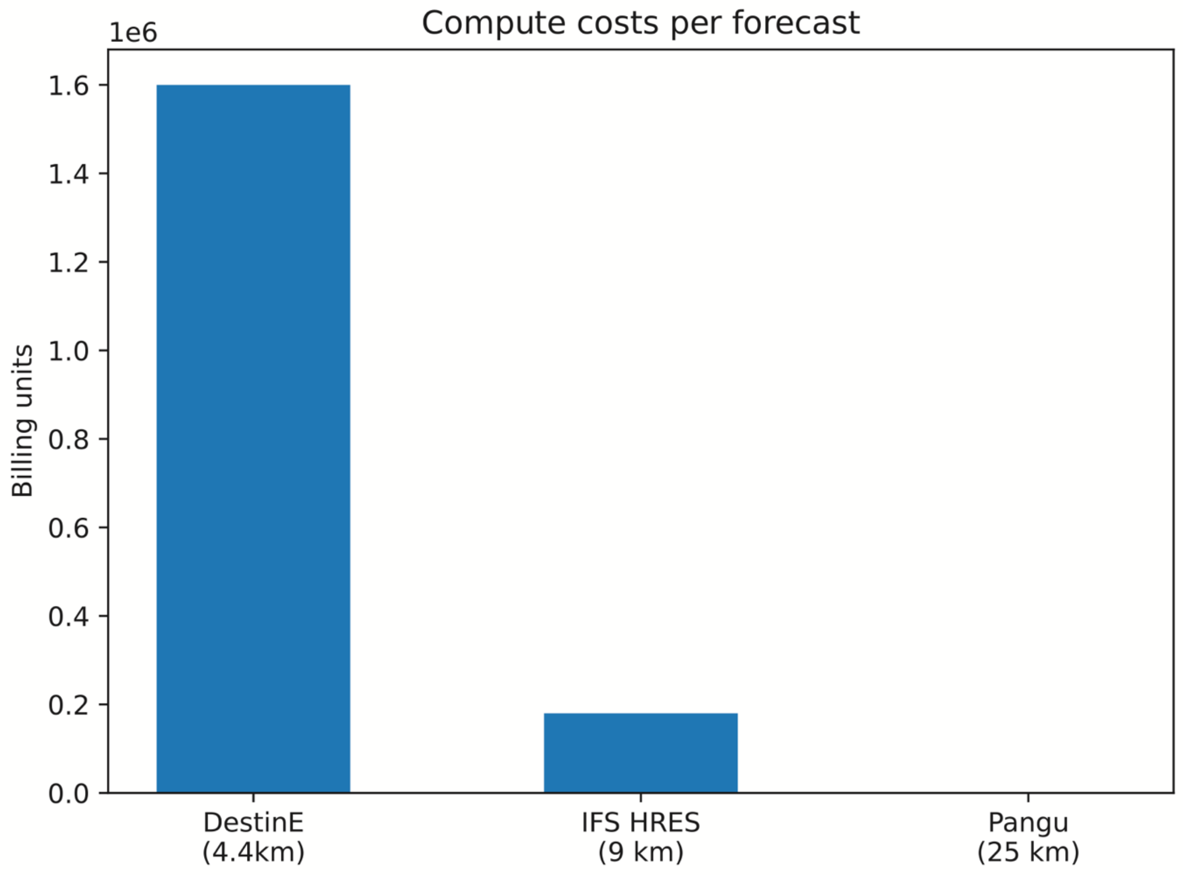
Machine learning-based weather forecasting



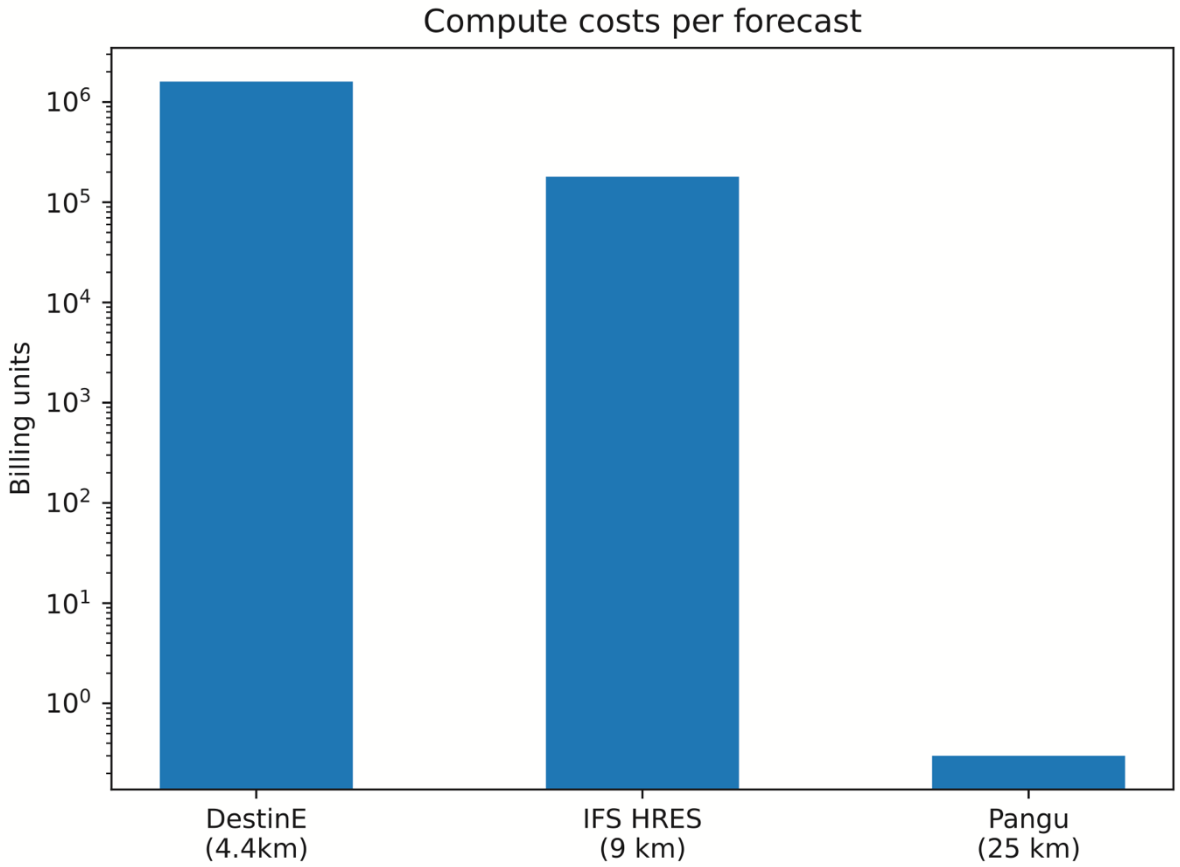
Machine learning-based weather forecasting



Machine learning-based weather forecasting



Machine learning-based weather forecasting



Machine learning-based weather forecasting

- State-of-the machine learning models for global weather forecasting
  - Trained on ERA5 reanalysis, potentially fine-tuned on operational analysis
  - Variety of network architectures
  - Robust forecasts with skill comparable or better to the best conventional models
  - Approximately three orders of magnitude cheaper than conventional models in terms of computational costs
  - Ensemble models
    - Diffusion models or with CRPS(-type) training
    - Skill comparable or better than conventional models

Machine learning-based weather forecasting

- State-of-the machine learning models for global weather forecasting
  - Trained on ERA5 reanalysis, potentially fine-tuned on operational analysis
  - Variety of network architectures
  - Robust forecasts with skill comparable or better to the best conventional models

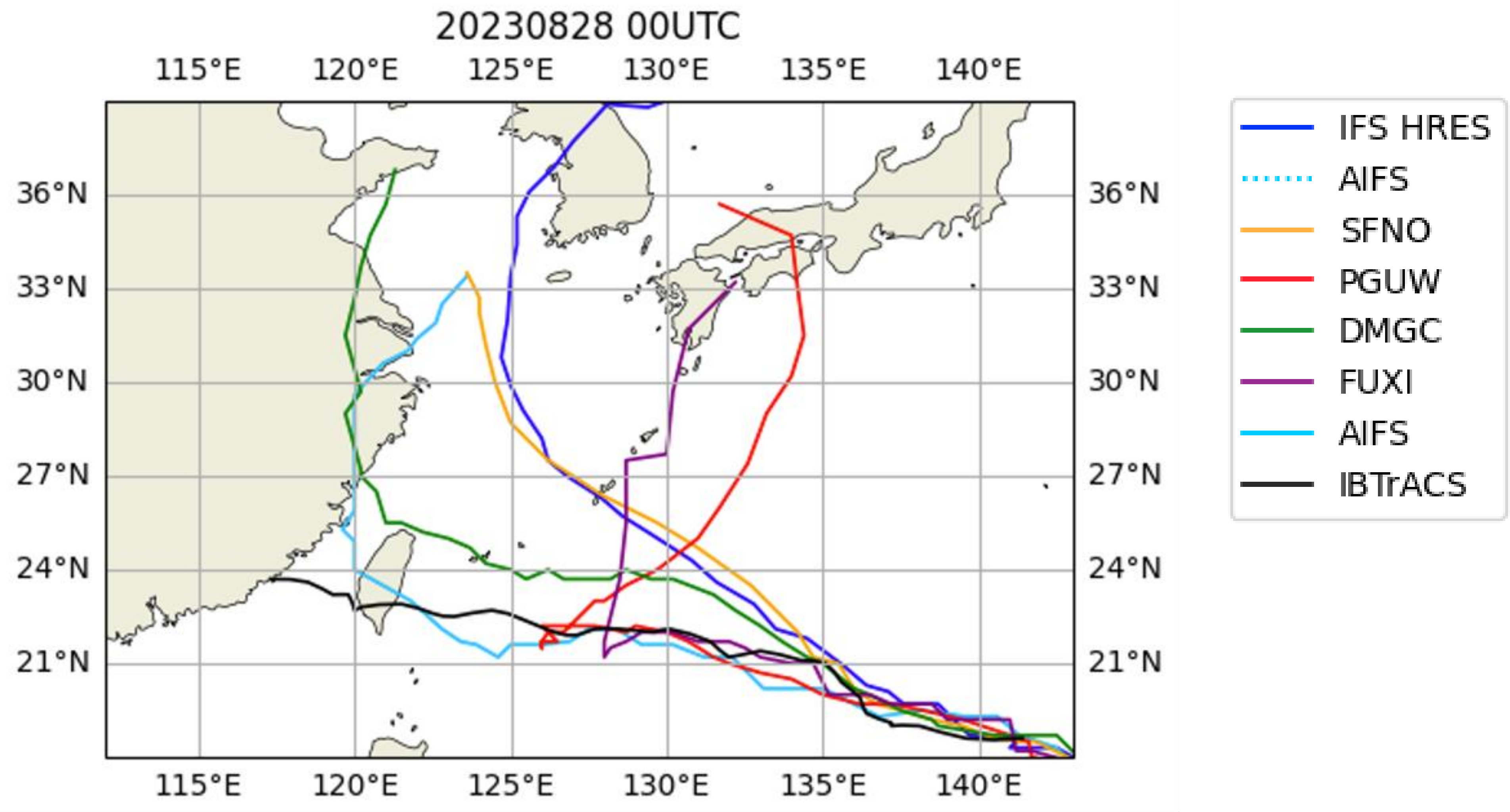
- Ensemble models
  - Diffusion models or with CRPS(-type) training

Machine learning-based weather forecasting

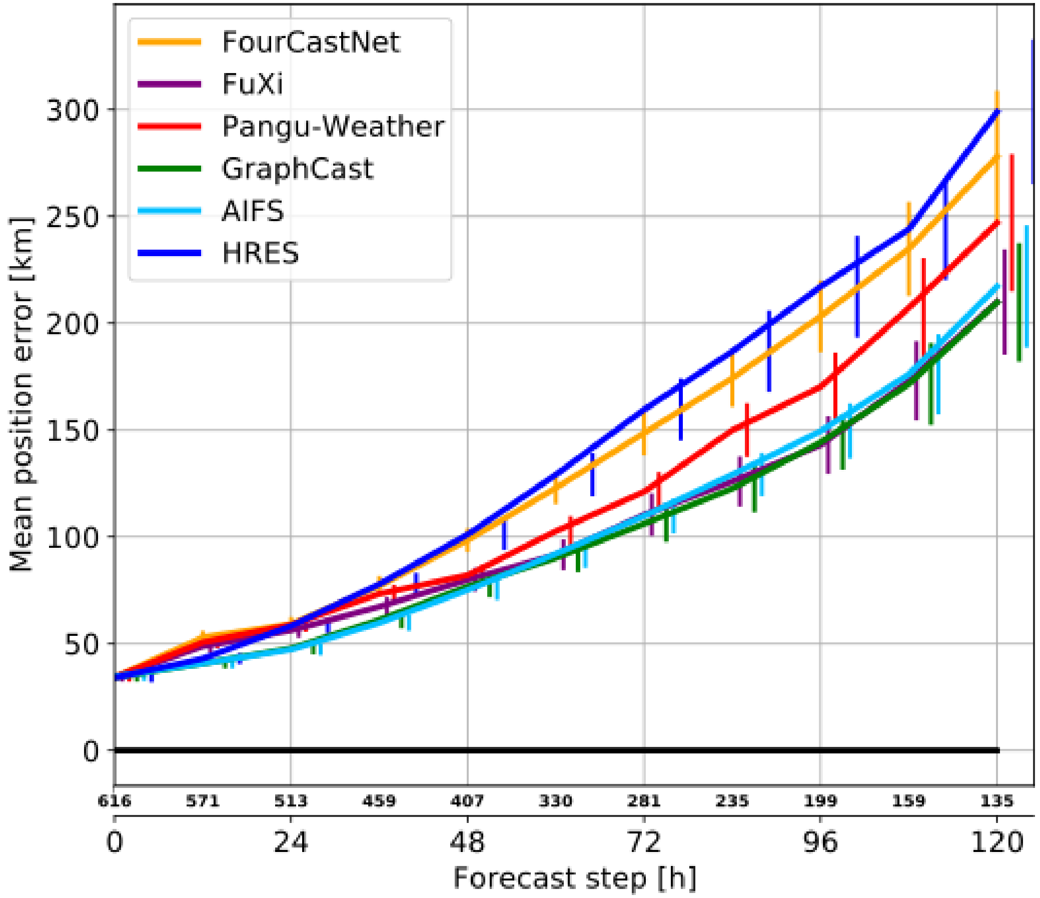
- Beyond global scores
  - How do state-of-the-art perform for extreme events?
  - How physical are the forecasts of machine learning models?

- Why can machine learning models outperform sophisticated equation-based models?

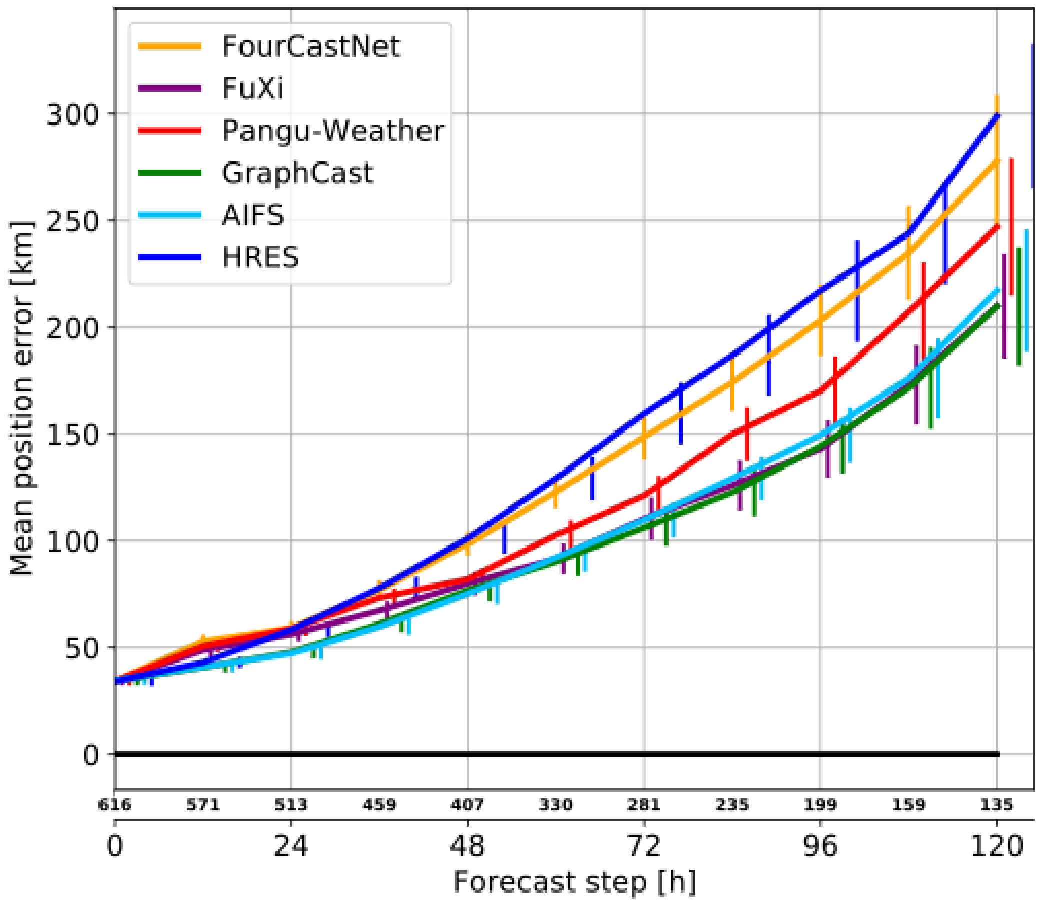
Machine learning-based weather forecasting

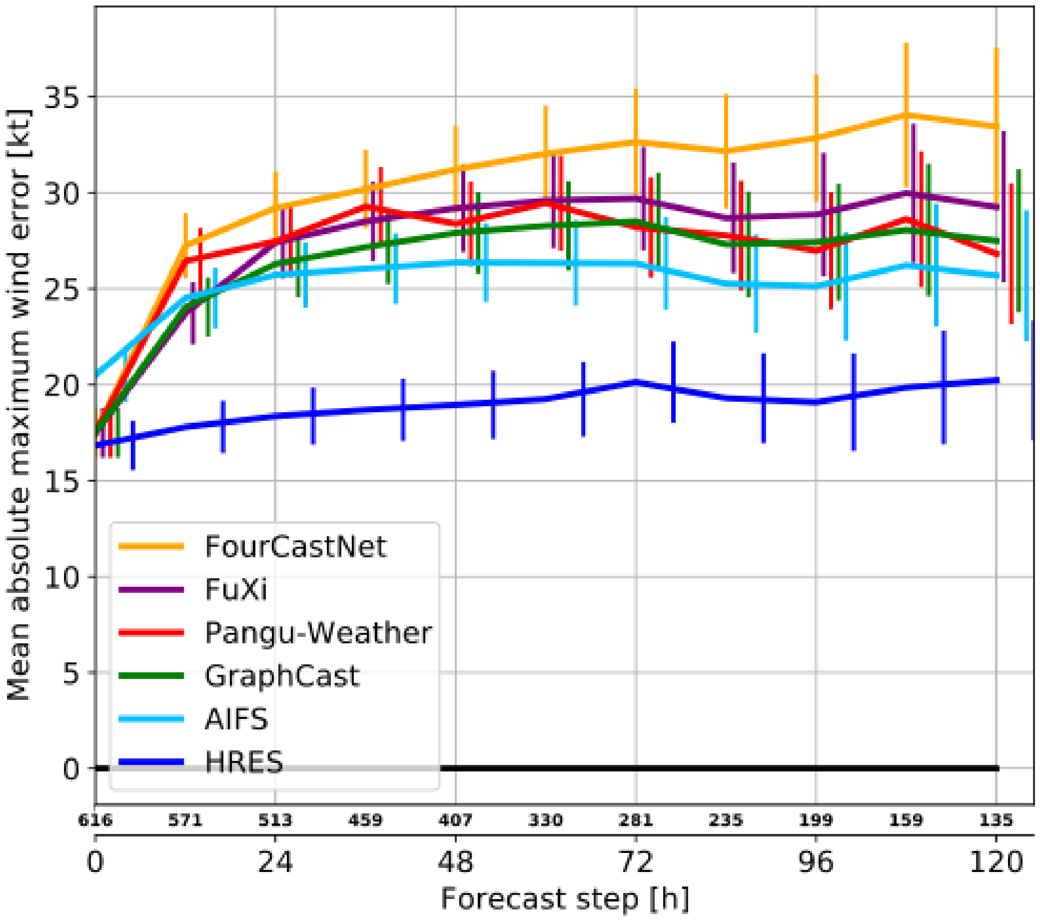


Machine learning-based weather forecasting



Machine learning-based weather forecasting





Machine learning-based weather forecasting

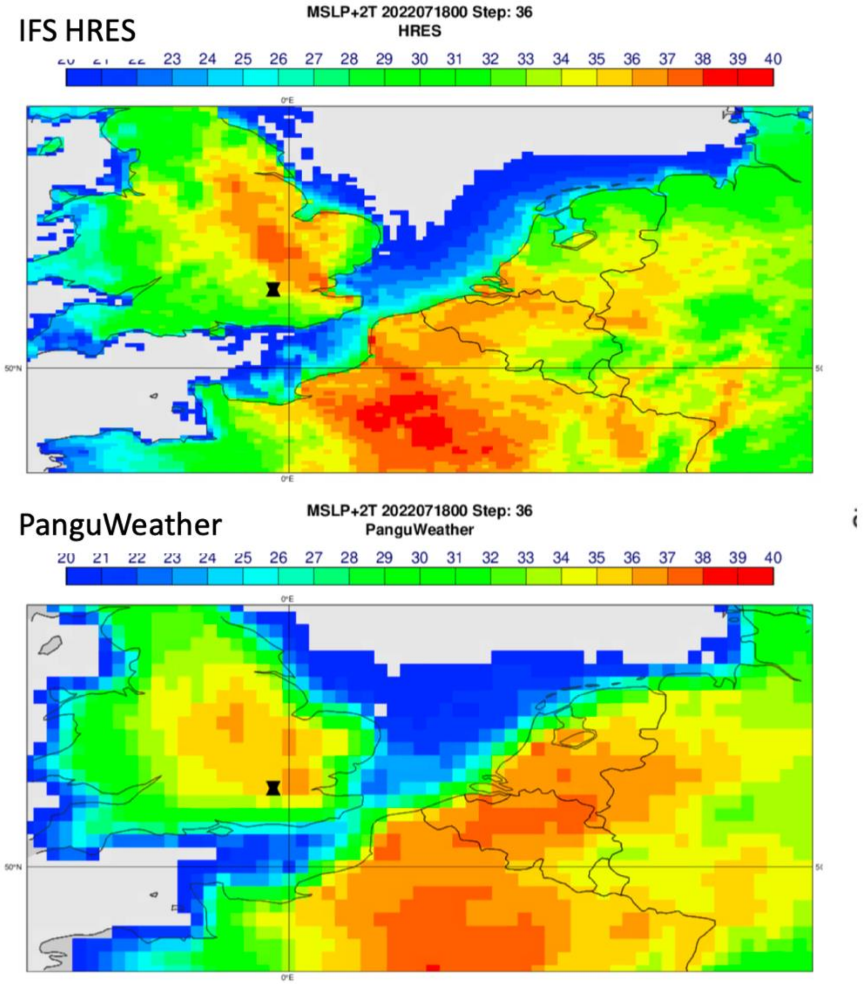


Figure from Zied Ben-Boualleague, ECMWF

Machine learning-based weather forecasting

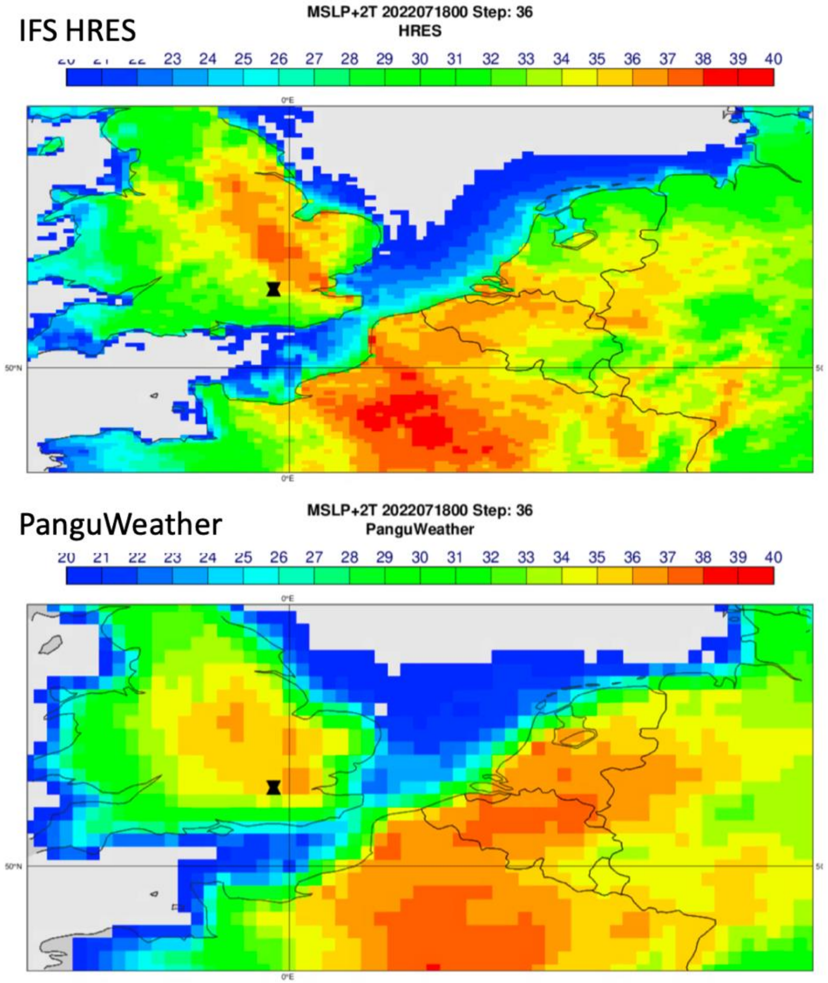
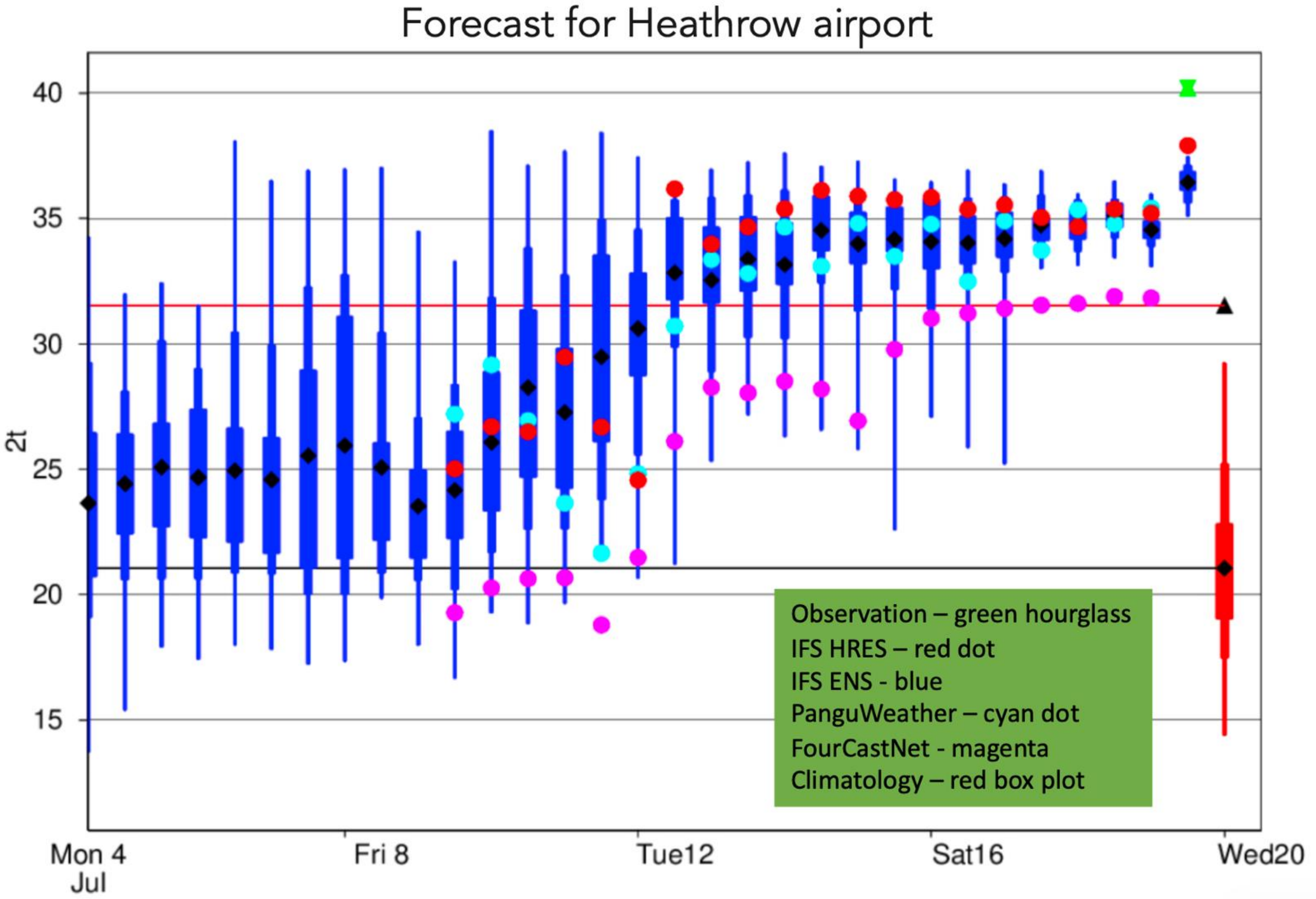


Figure from Zied Ben-Boualleague, ECMWF



Machine learning-based weather forecasting

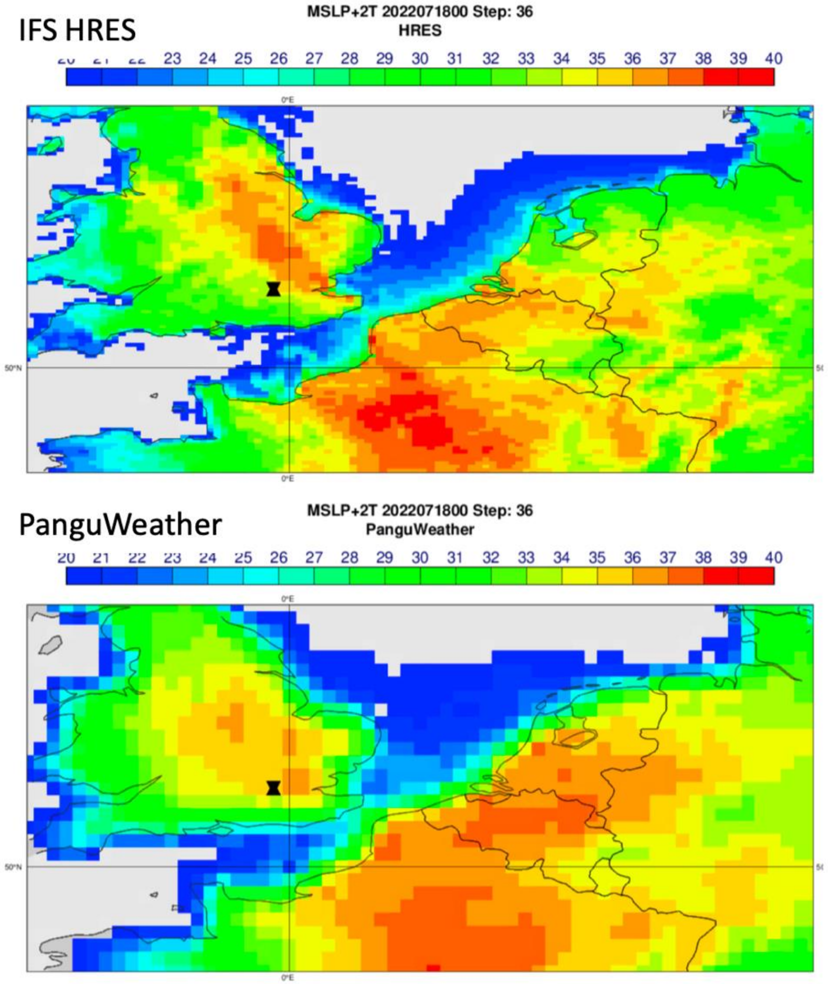
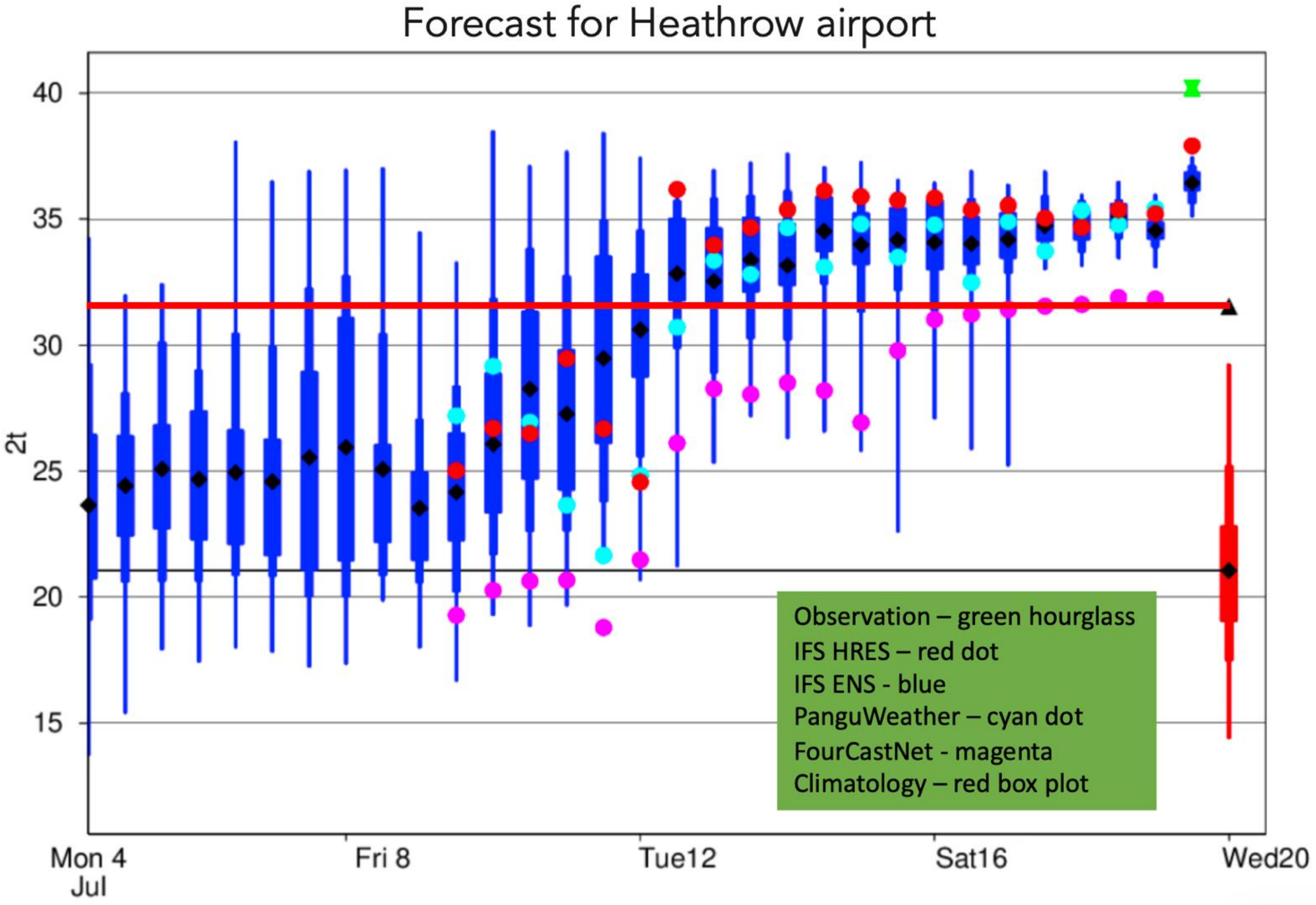
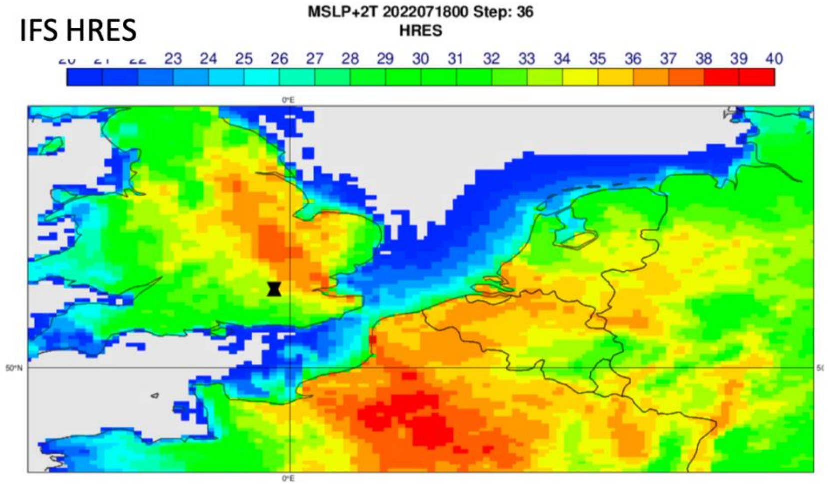


Figure from Zied Ben-Boualleague, ECMWF



Machine learning-based weather forecasting

IFS HRES



PanguWeather

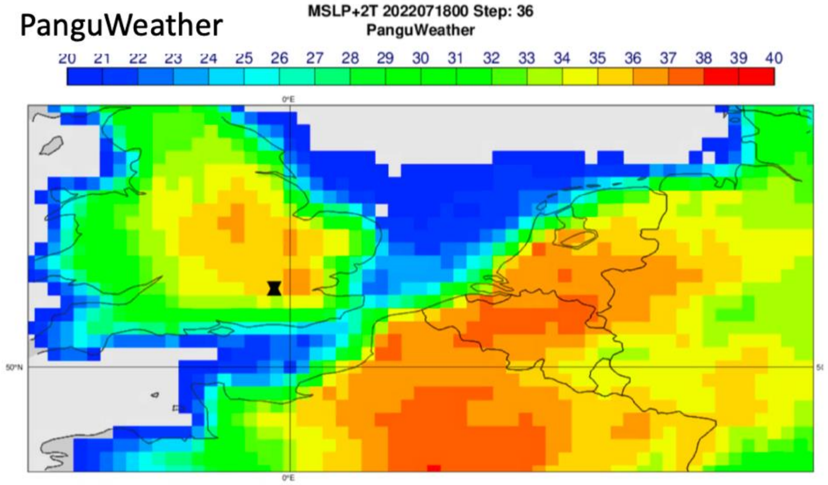
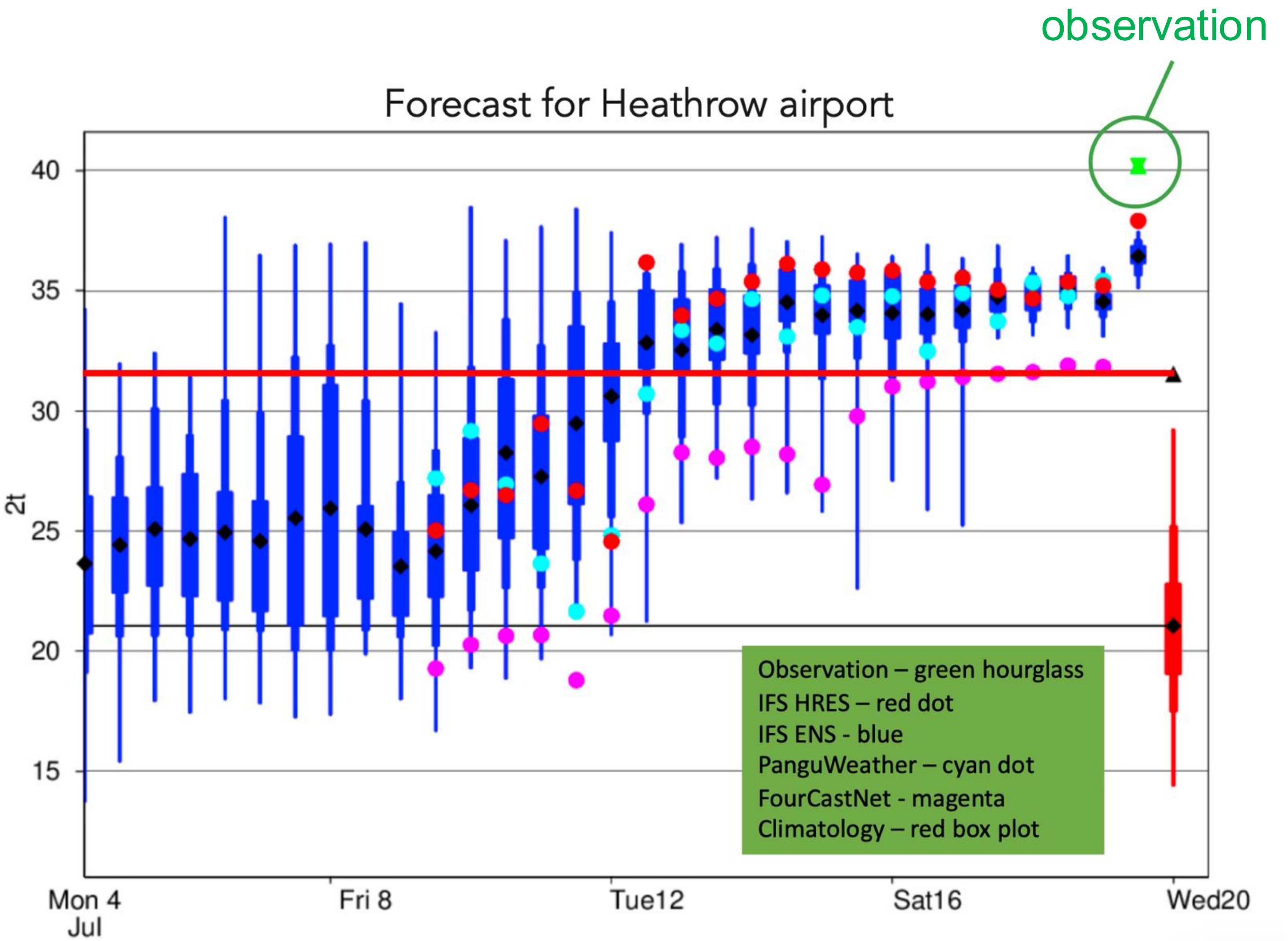


Figure from Zied Ben-Bouallegue, ECMWF



Machine learning-based weather forecasting

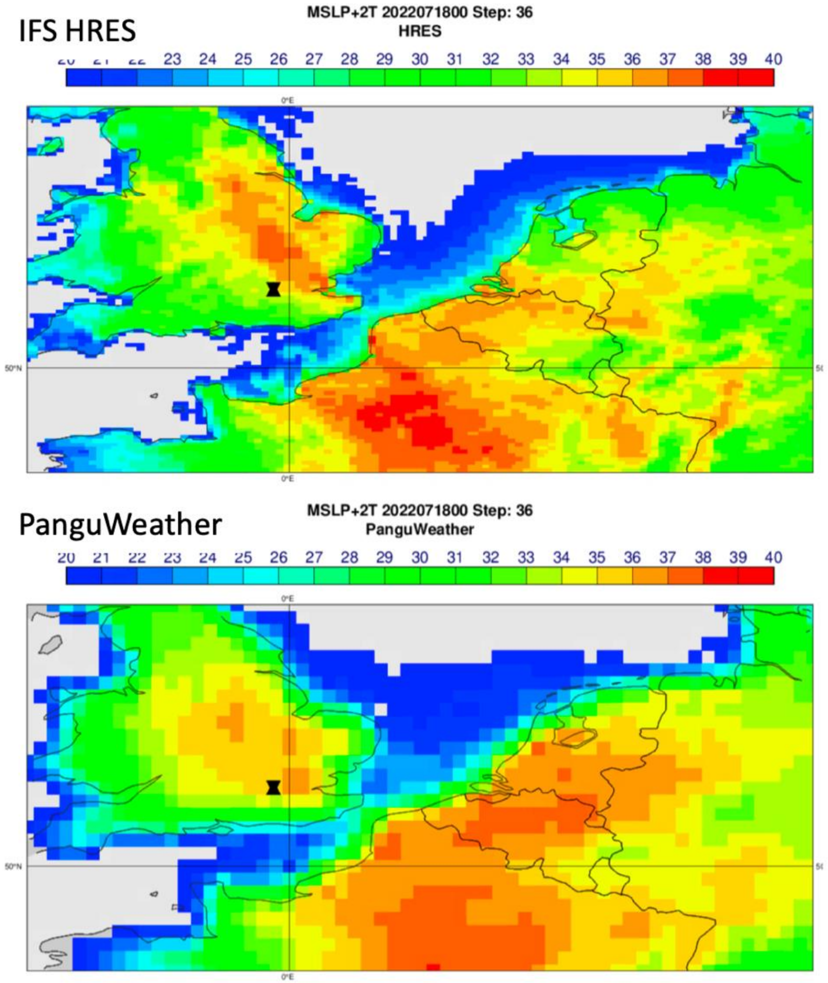
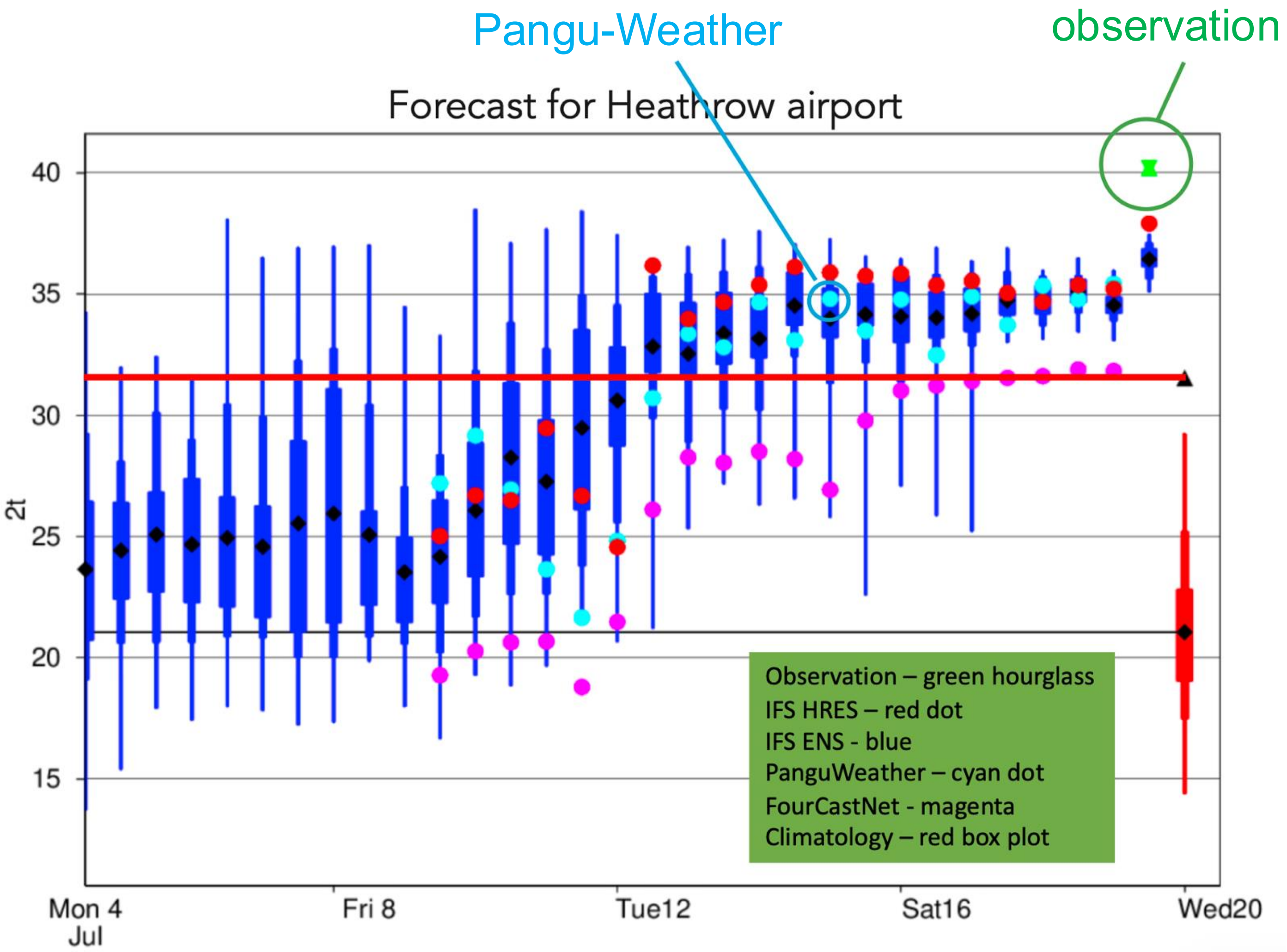
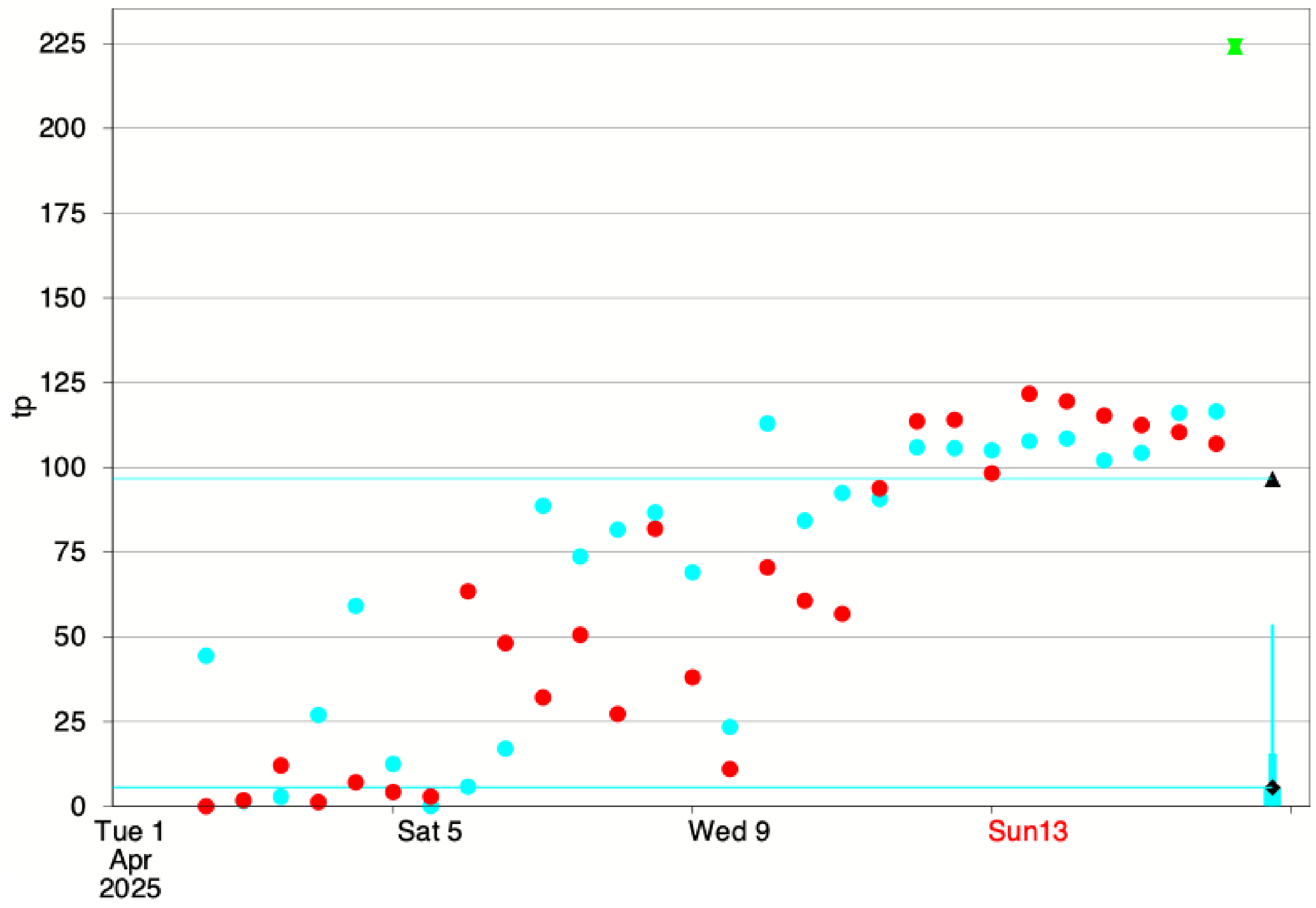


Figure from Zied Ben-Boualleague, ECMWF



Machine learning-based weather forecasting

Forecast Evaluation @ 6 UTC, 17 Apr valid datetime — IFS vs AIFS (lead-time sweep)

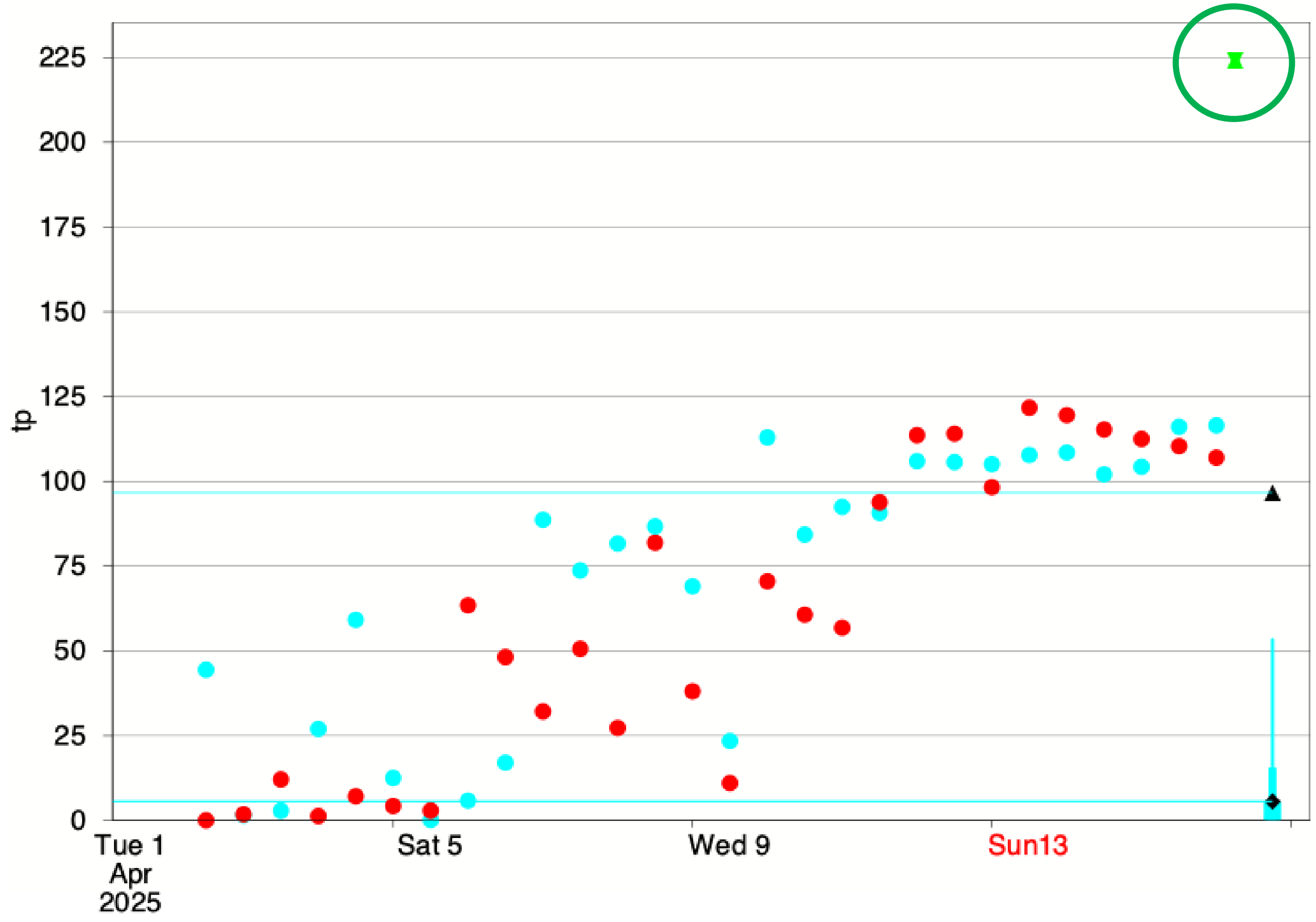


Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Simplon Pass: 0.5° box
16→17 Apr (06–06 UTC).

Machine learning-based weather forecasting

Forecast Evaluation @ 6 UTC, 17 Apr valid datetime — IFS vs AIFS (lead-time sweep)

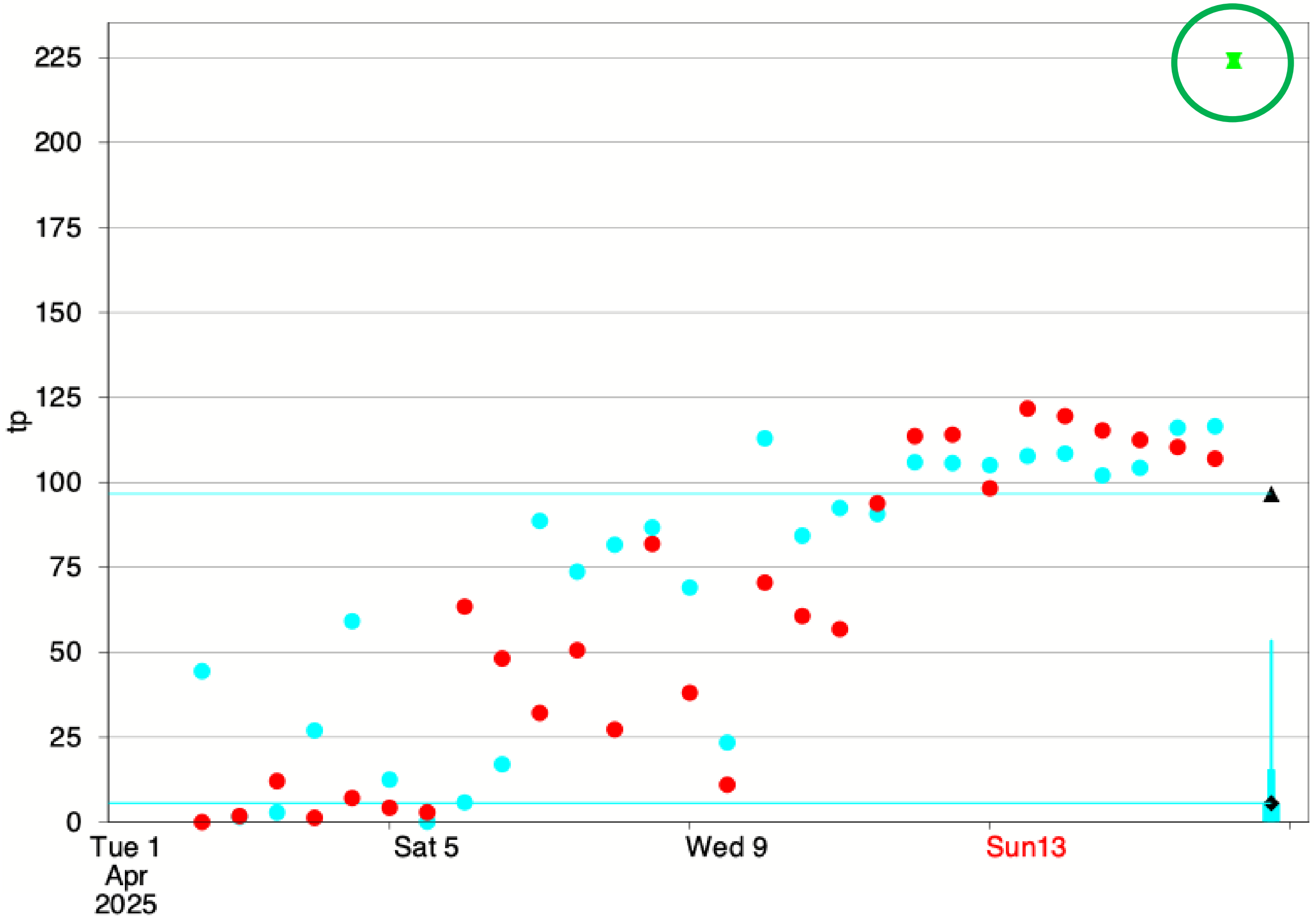


Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Simplon Pass: 0.5° box
16→17 Apr (06–06 UTC).

Machine learning-based weather forecasting

Forecast Evaluation @ 6 UTC, 17 Apr valid datetime — IFS vs AIFS (lead-time sweep)



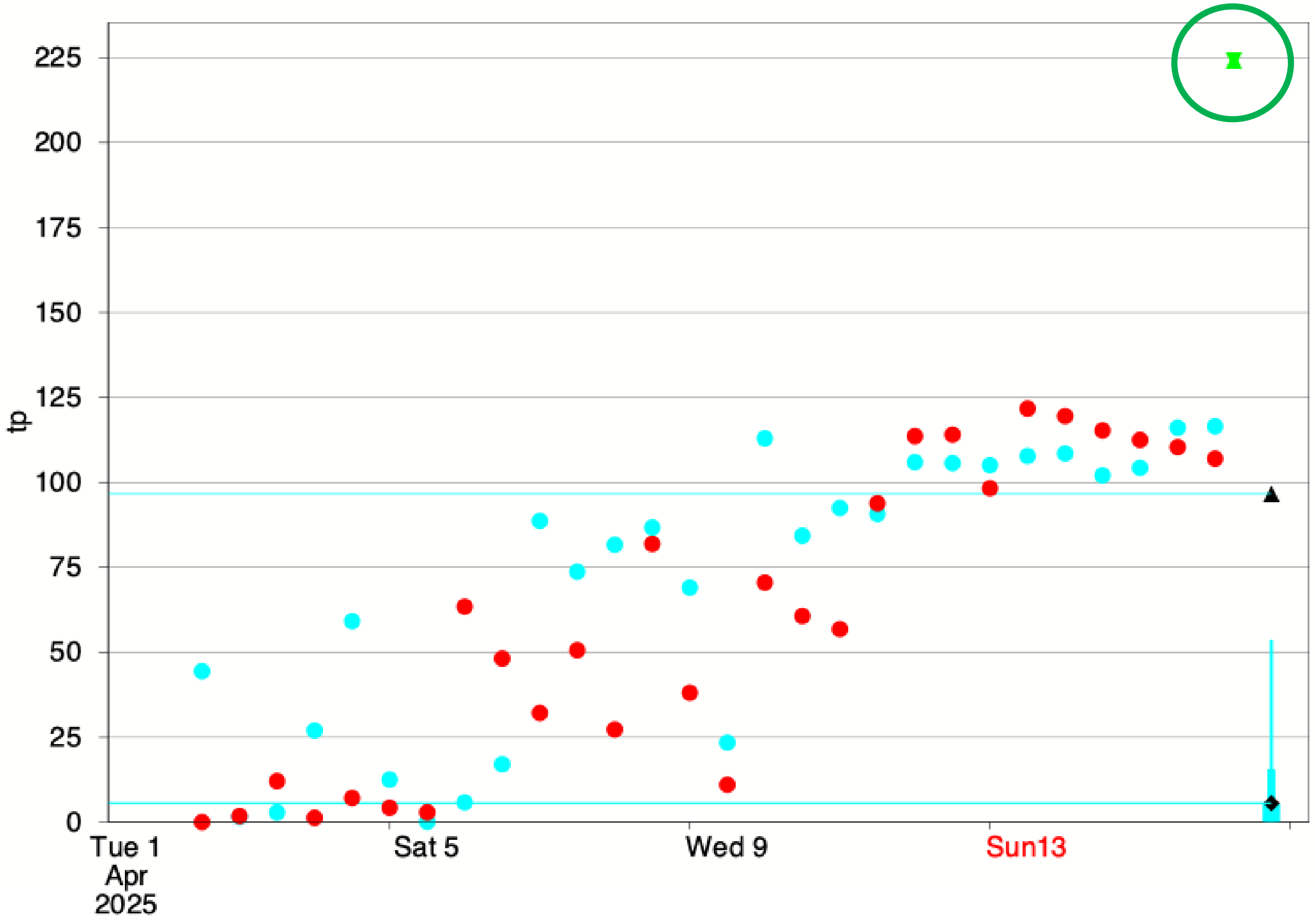
Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Simplon Pass: 0.5° box
16→17 Apr (06–06 UTC).

forecast initialization time

Machine learning-based weather forecasting

Forecast Evaluation @ 6 UTC, 17 Apr valid datetime — IFS vs AIFS (lead-time sweep)



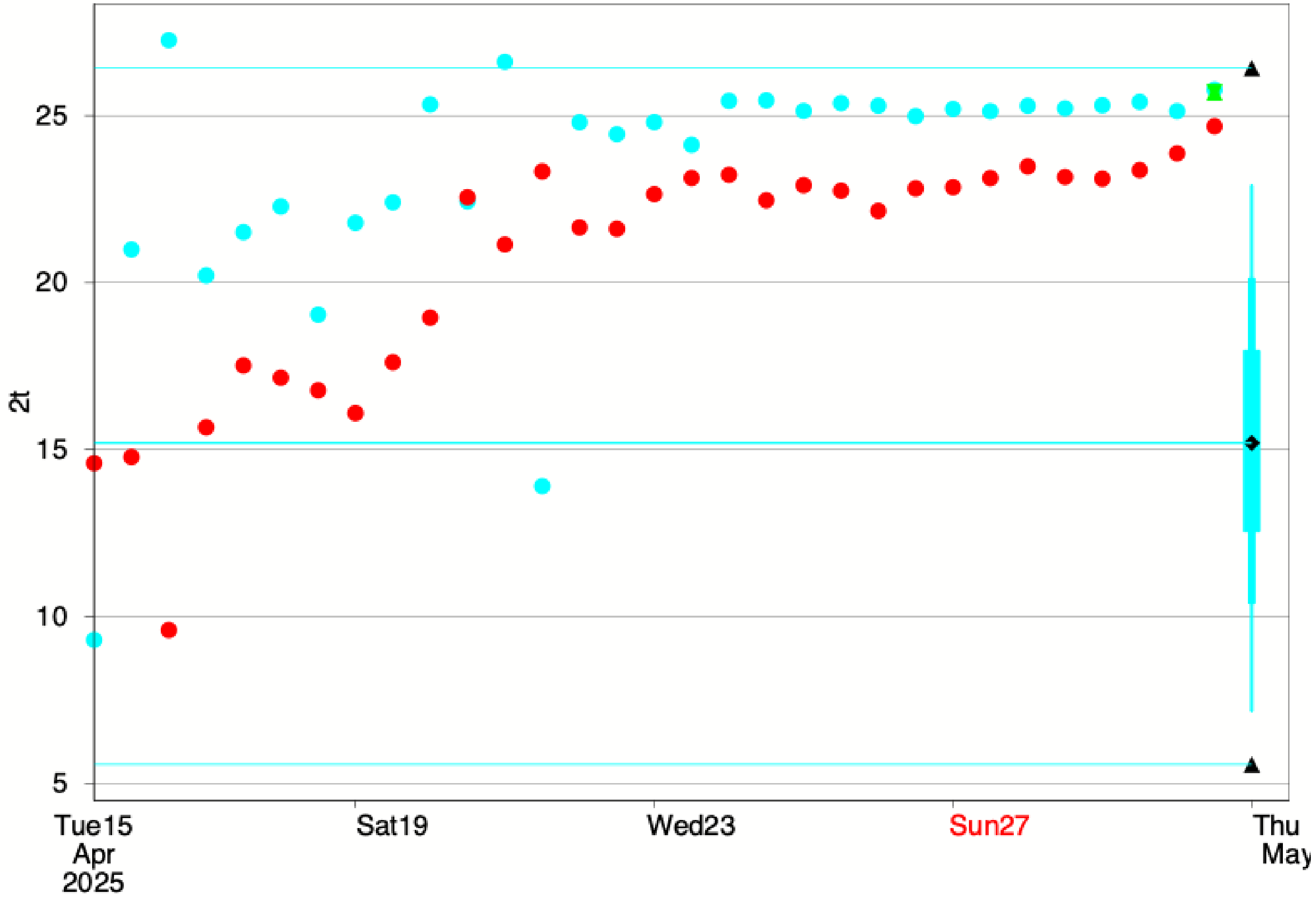
Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Simplon Pass: 0.5° box
16→17 Apr (06–06 UTC).

forecast initialization time

Machine learning-based weather forecasting

Forecast Evaluation @ 12 UTC, 1 Aug valid datetime — IFS vs AIFS (lead-time sweep)



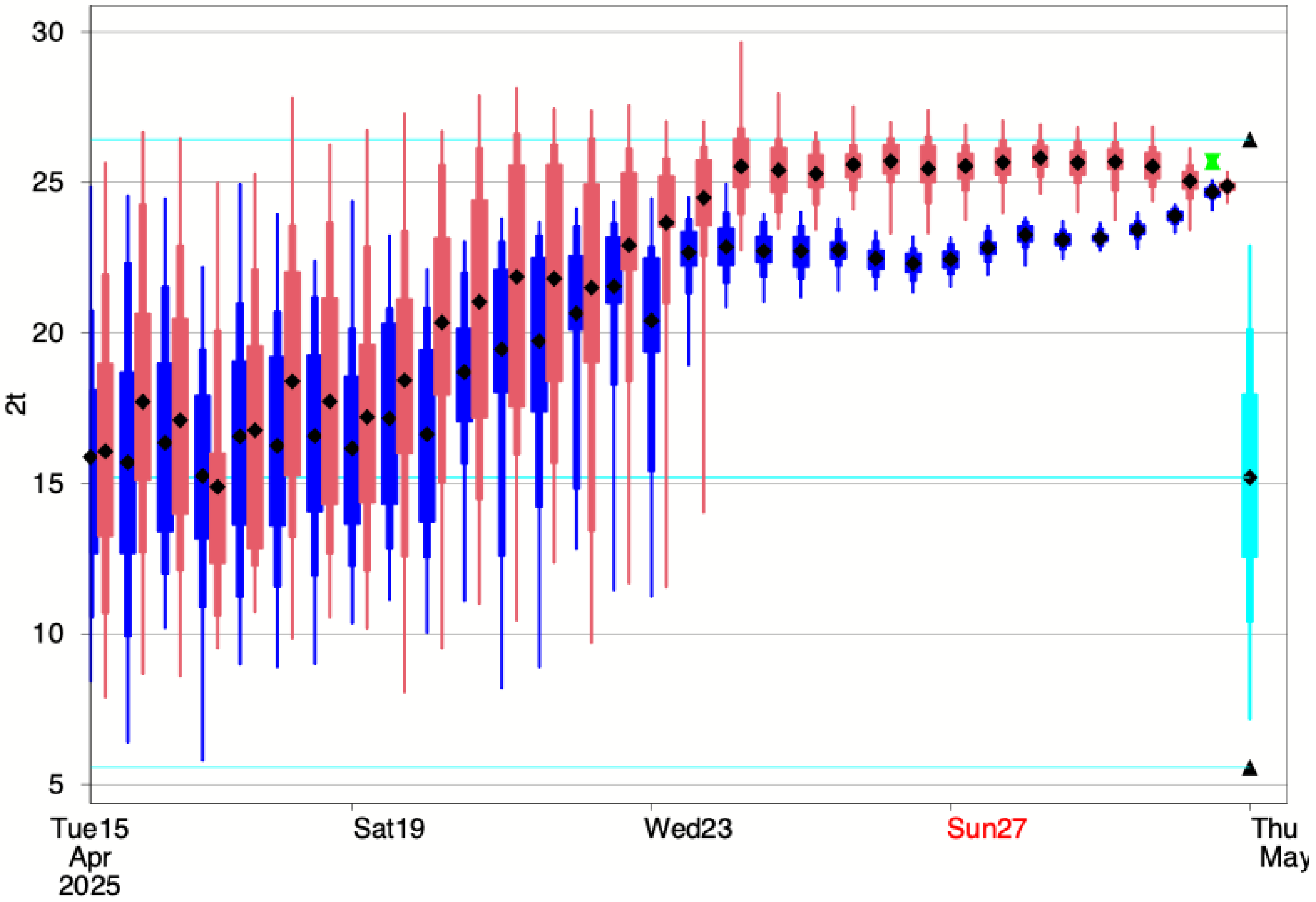
Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Troyes, Fr: 0.25°×0.25° box
12 UTC 30 Apr

forecast initialization time

Machine learning-based weather forecasting

Forecast Evaluation @ 12 UTC, 1 Aug valid datetime — IFS vs AIFS (lead-time sweep)



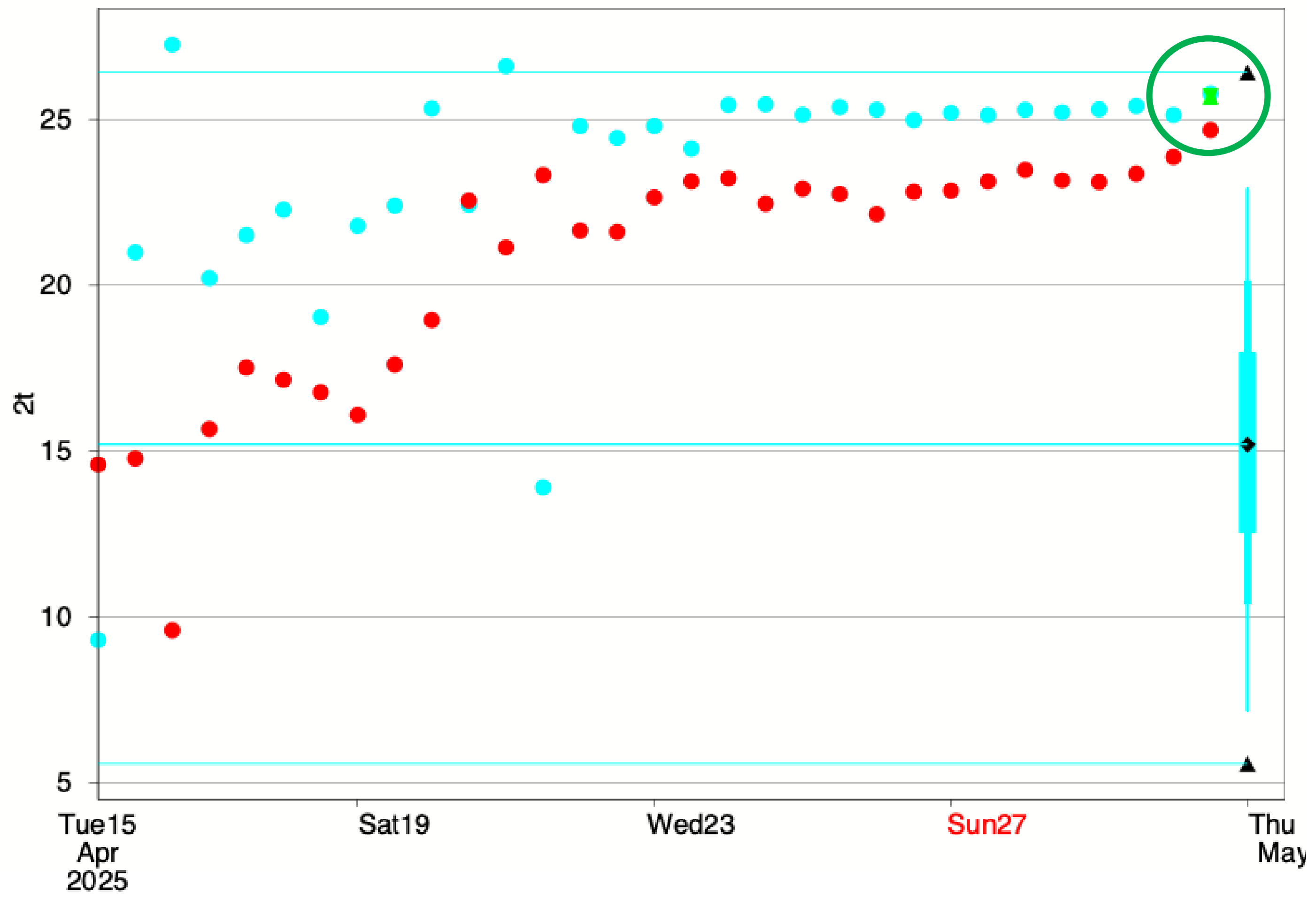
Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Troyes, Fr: 0.25°×0.25° box
12 UTC 30 Apr

forecast initialization time

Machine learning-based weather forecasting

Forecast Evaluation @ 12 UTC, 1 Aug valid datetime — IFS vs AIFS (lead-time sweep)



Observation	
IFS HRES	Red
AIFS Single	Cyan
Climatology	Cyan

Troyes, Fr: 0.25°×0.25° box
12 UTC 30 Apr

forecast initialization time

Machine learning-based weather forecasting

- Beyond global scores
  - How do state-of-the-art perform for extreme events?
  - How physical are the forecasts of machine learning models?

- Why can machine learning models outperform sophisticated equation-based models?

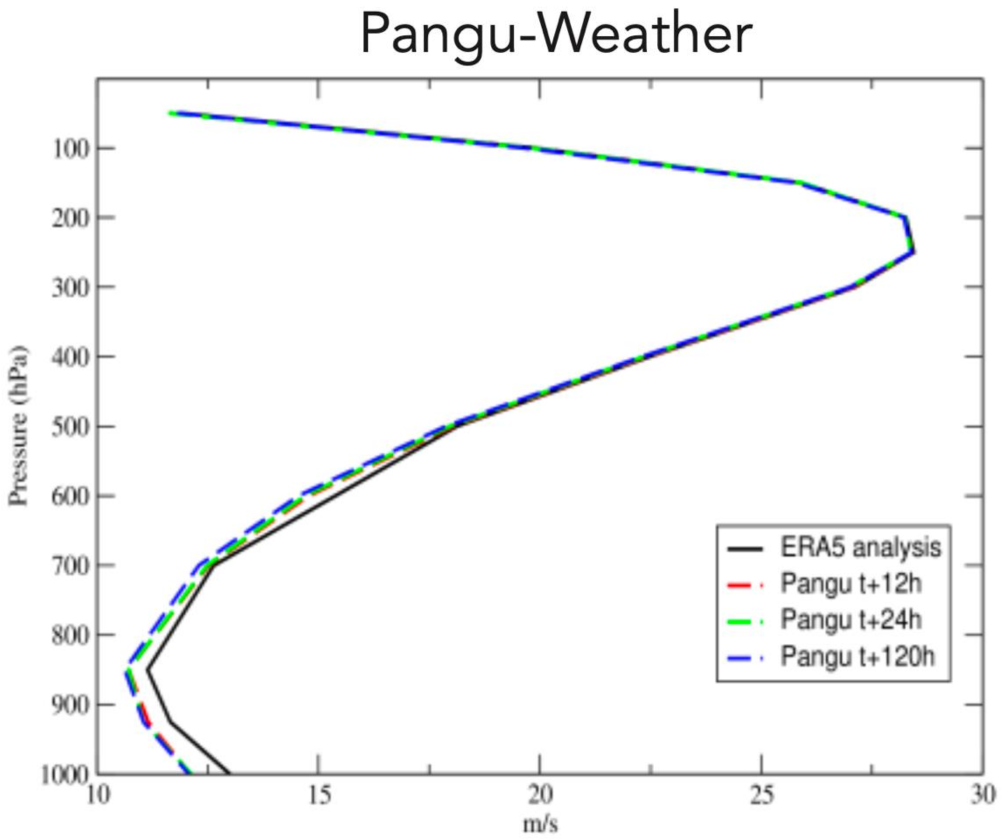
Machine learning-based weather forecasting

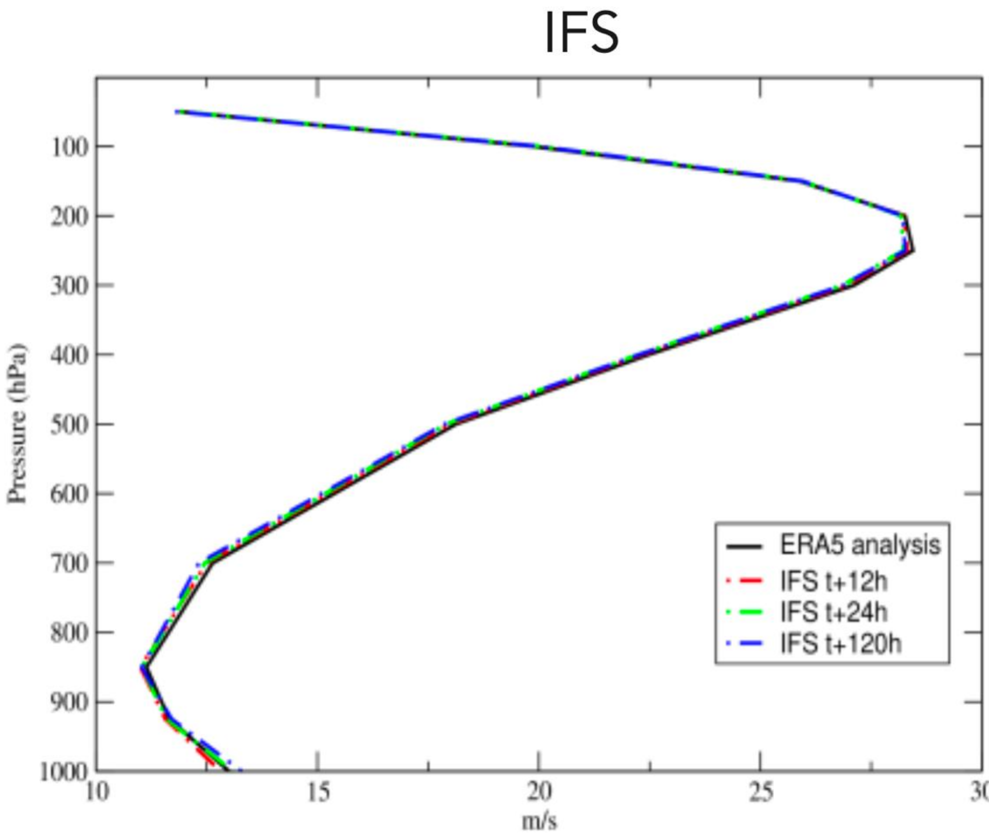
- Beyond global scores
  - How do state-of-the-art perform for extreme events?
  - **How physical are the forecasts of machine learning models?**

- Why can machine learning models outperform sophisticated equation-based models?

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
    - Geostrophic balance

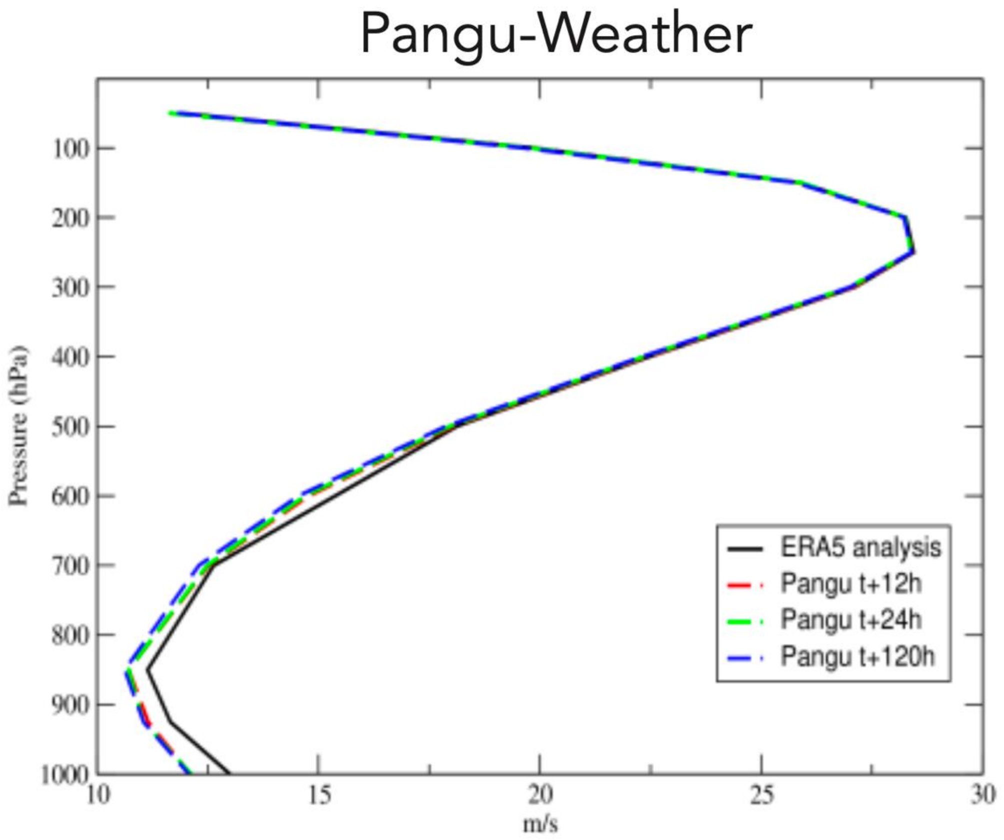


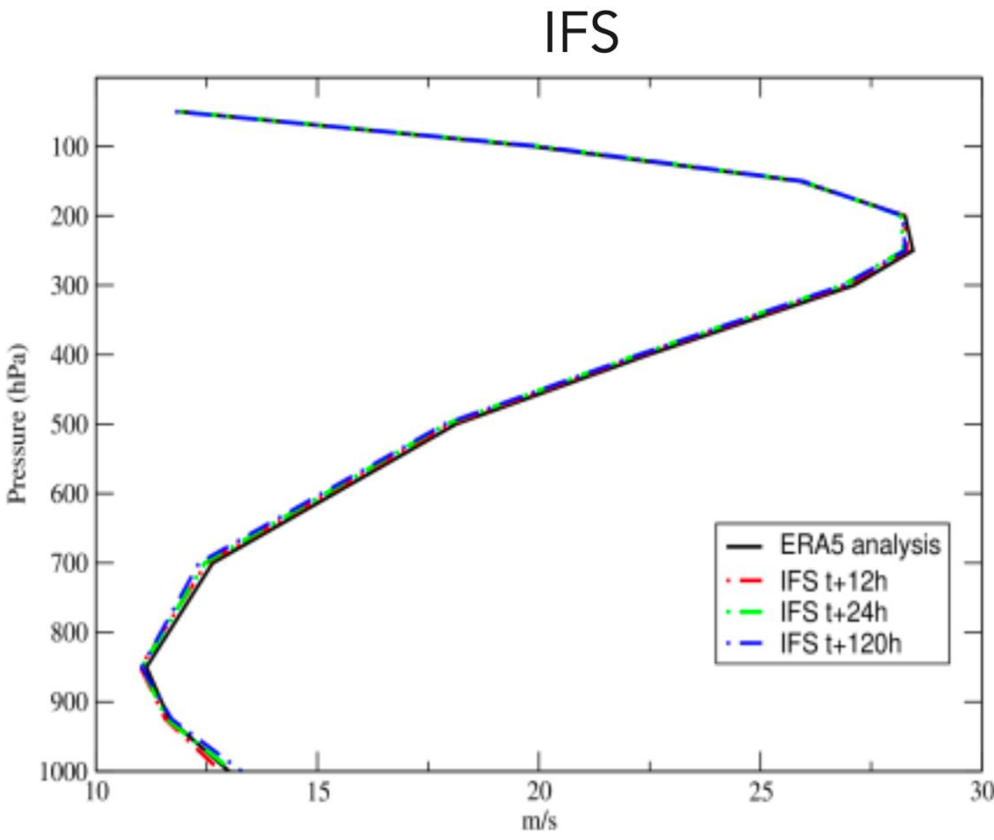


https://arxiv.org/abs/2309.08473

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
  - Geostrophic balance—but is it necessary for a forecast to be physical?





https://arxiv.org/abs/2309.08473

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
  - Geostrophic balance—but is it necessary for a forecast to be physical?

Geometric optics

- Describes the propagation and physical behaviour of light

Maxwell's equations

- Describes the propagation and physical behaviour of light

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
  - Geostrophic balance—but is it necessary for a forecast to be physical?

Geometric optics
- Describes the propagation and physical behaviour of light
- No diffraction
- No caustics
- No interference
- …

Maxwell’s equations
- Describes the propagation and physical behaviour of light

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
  - Geostrophic balance—but is it necessary for a forecast to be physical?

Geometric optics

- Describes the propagation and physical behaviour of light
- No diffraction
- No caustics
- No interference
- …

Yet geometric optics is a completely valid physical theory.

Maxwell's equations

- Describes the propagation and physical behaviour of light

Machine learning-based weather forecasting

- How physical are the forecasts of machine learning models?
  - Geostrophic balance—but is it necessary for a forecast to be physical?

Geometric optics
- Describes the propagation and physical behaviour of light
- No diffraction
- No caustics
- No interference
- …

Yet geometric optics is a completely valid physical theory.

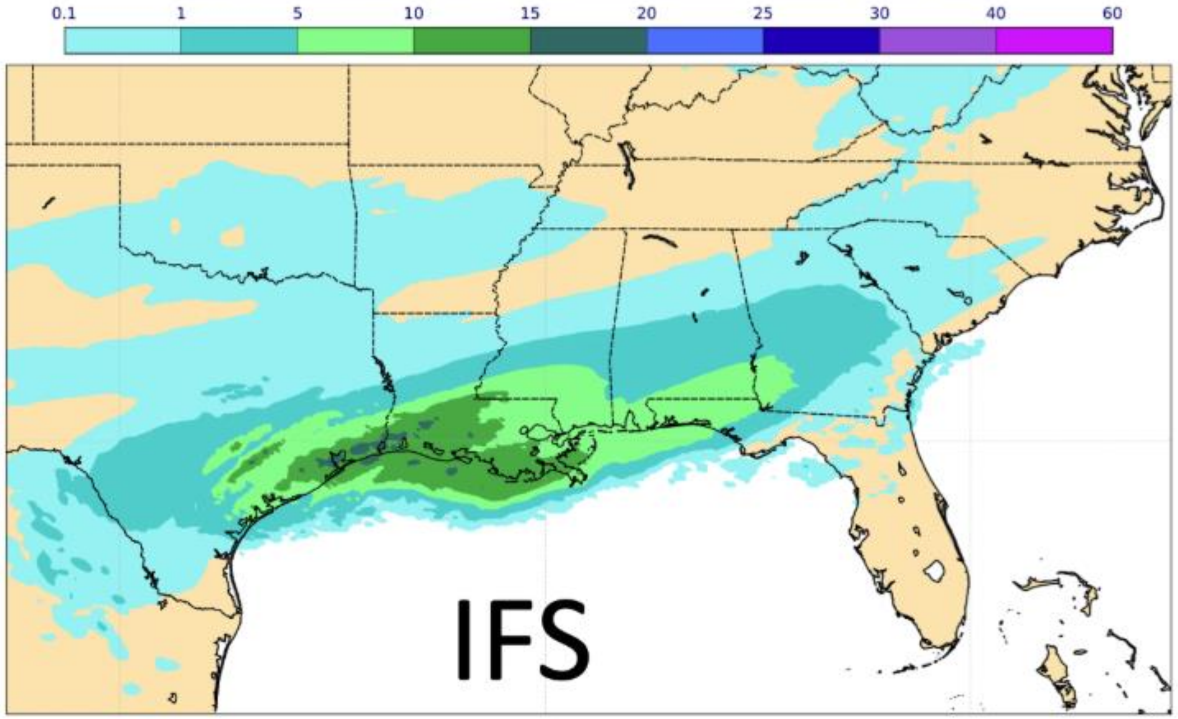
Maxwell's equations
- Describes the propagation and physical behaviour of light

Fun reading on the subject: Asimov, The relativity of wrong

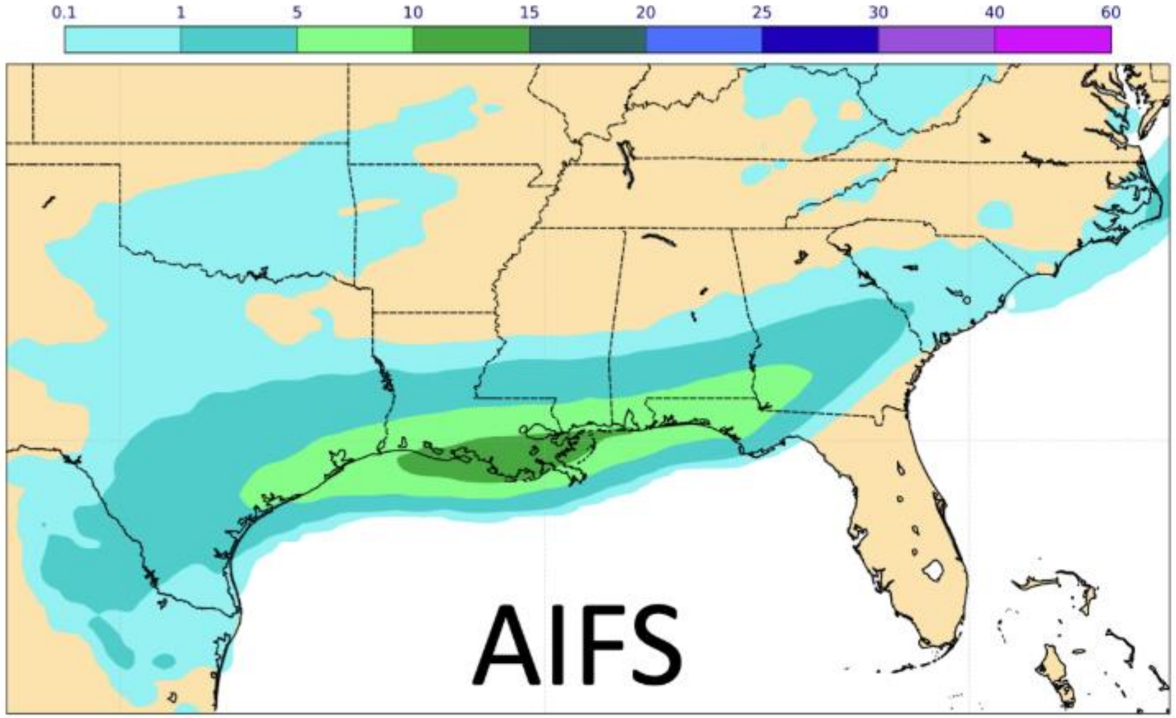
Machine learning-based weather forecasting

2-day forecast of **rare snow** along the Gulf Coast

20 January 2025 00 UTC t+48: 24h snowfall



20 January 2025 00 UTC t+48: 24h snowfall



Conclusion

- Machine learning: derive models from data
- Deep learning
  - Large amounts of data and many free parameters
  - Qualitatively different behaviour than small neural nets

Conclusion

- Machine learning: derive models from data
- Deep learning
  - Large amounts of data and many free parameters
  - Qualitatively different behaviour than small neural nets

- Machine learning is promising for weather forecasting
  - Large amounts of data
  - Various processes and the interaction of many subsystems are difficult to model conventionally
  - Probabilistic problem

Conclusion

- Machine learning: derive models from data
- Deep learning
  - Large amounts of data and many free parameters
  - Qualitatively different behaviour than small neural nets

- Machine learning is promising for weather forecasting
  - Large amounts of data
  - Various processes and the interaction of many subsystems are difficult to model conventionally
  - Probabilistic problem

- Machine learning is currently revolutionizing weather forecasting
  - End-to-end forecasting models outperform conventional models
  - Many untapped opportunities, e.g. use of observations and combining different datasets

What do you expect from the course?

ECMWF | EUROPEAN CENTRE FOR MEDIUM-RANGE WEATHER FORECASTS