Probabilistic Energy Forecasting: Methodologies, Implementations, and Applications

IEEE Power & Energy Society General Meeting
Sunday 17 July 2022
Denver, Colorado

Instructors: Jethro Browell, Tao Hong, Yi Wang, and Bri-Mathias Hodge

Probabilistic Energy Forecasting: Methodologies, Implementations, and Applications

Part 1: Introduction to probabilistic energy forecasting

Instructor: Dr Jethro Browell

@jethrobrowell

Learning objectives

By the end of this course, students should be able to:

Part 1

- Identify, list, and describe the different types of probabilistic forecast

Part 2

- Use a range of metrics and visualizations to evaluate probabilistic forecasts

Part 3

- Describe the processes and data required to produce probabilistic wind and solar power forecasts, electricity price forecasts, and electric load forecasts

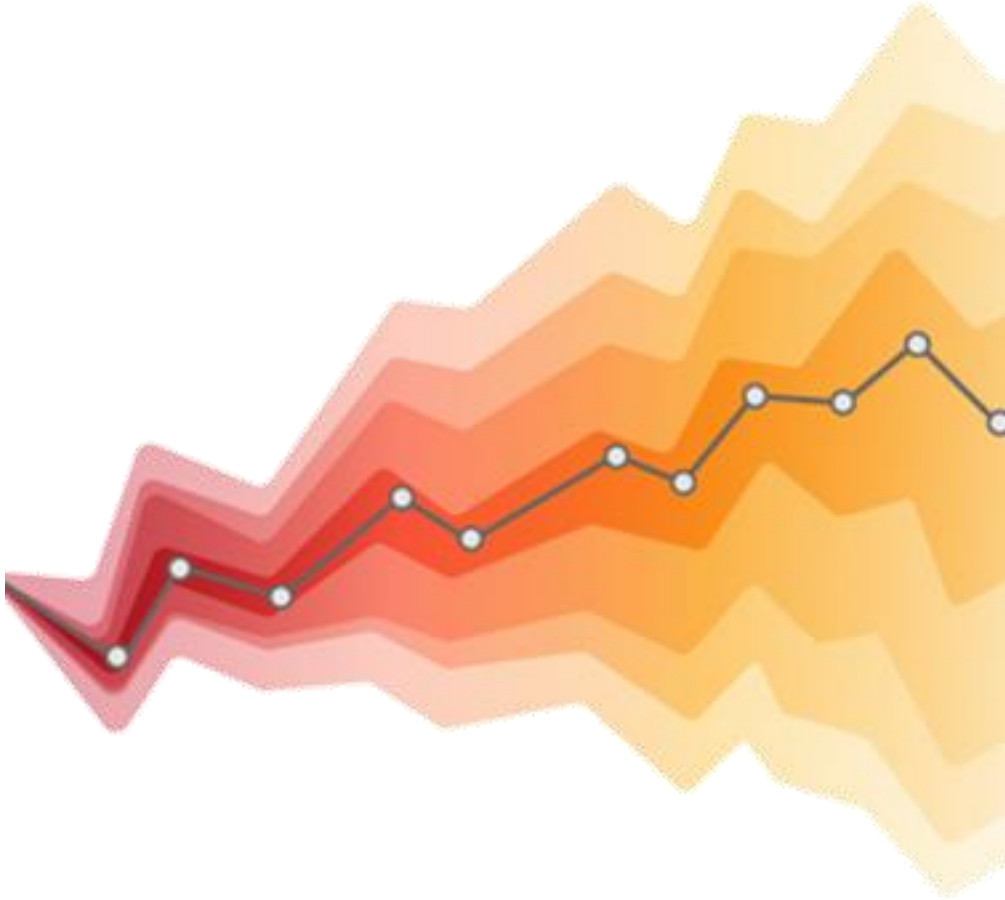
Part 4

- List contemporary statistical and machine learning methods used in probabilistic forecasting, their characteristics and available software implementations
- Describe how probabilistic energy forecasting is used in power systems operation, control, and market participation

Part 1 Contents

Introduction to probabilistic energy forecasting

- Motivation
- Types of probabilistic forecast
- Principles of energy forecasting
  - Timescales
  - Physical and fundamental methods
  - Data-driven methods
- Forecast quality and evaluation
- Decision-making game



Motivation

Probabilistic energy forecasting from minutes to weeks-ahead

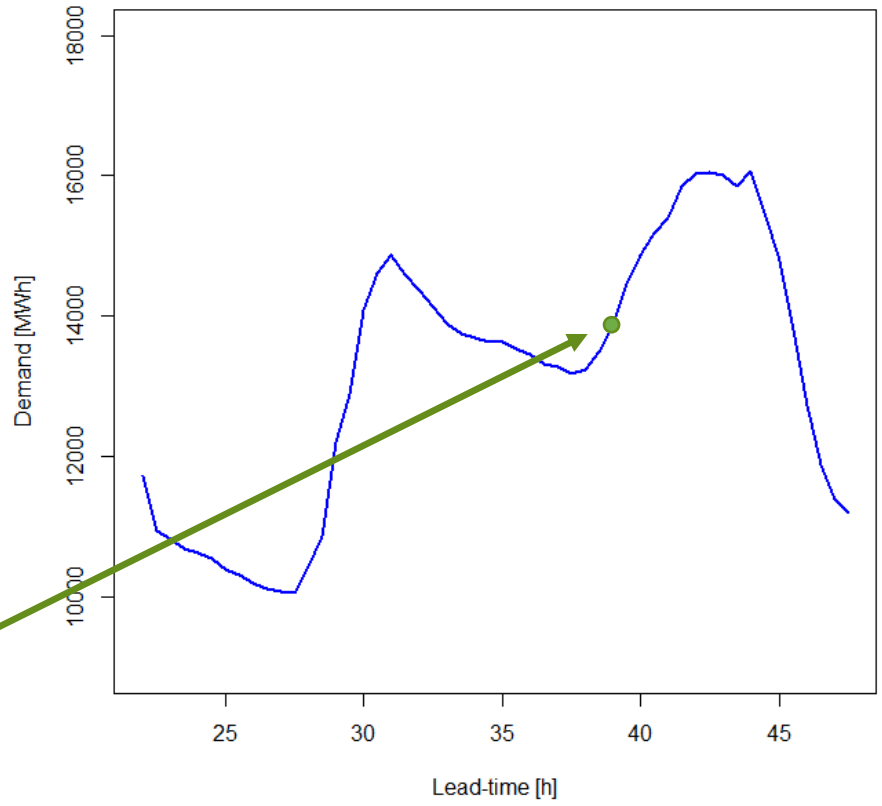
Example: Load

- The objective of electricity markets and system operators is to ensure electricity supply and demand balance
  - This requires forward planning based on forecasts of electricity demand
  - Forecasts are imperfect
  - Some flexibility must be included in schedules to manage forecast errors
  - This comes at a cost, so how much flexibility (often called “reserve”) should be scheduled?

Example: Load

Deterministic forecast

Consider this day-ahead forecast of load on the GB transmission system:



Forecast at each time point is a **single number**

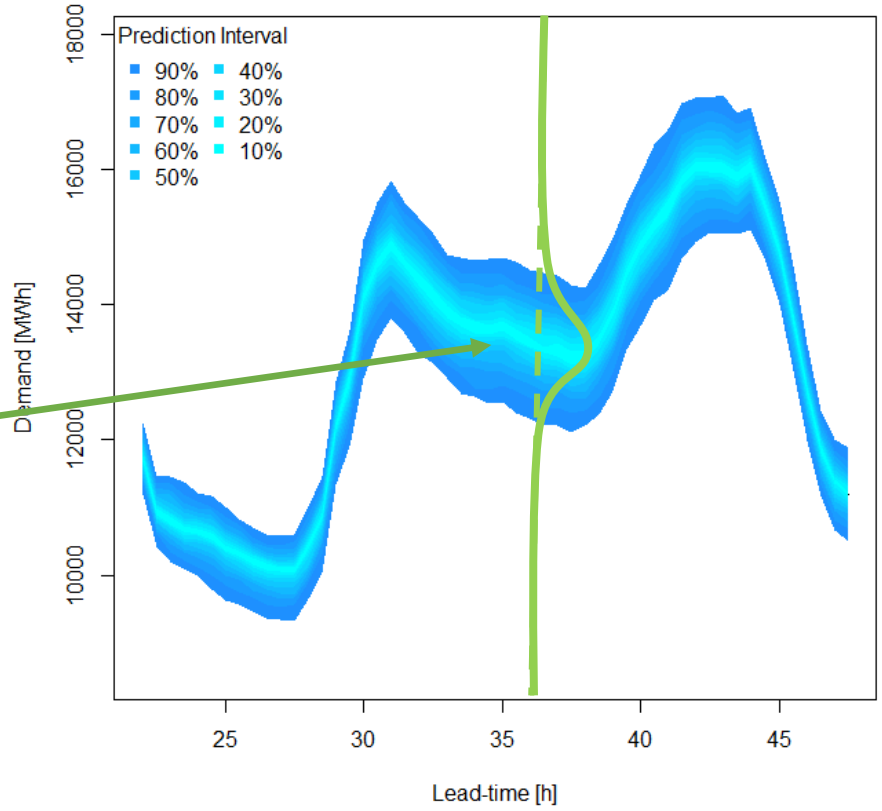
Suitable for decision making if:

1. Cost of over/under predicting is symmetric
2. User is *risk neural*

Example: Load

Probabilistic forecast

Forecast at each time point is a **probability distribution**



Can make decision based on:

1. **Asymmetric costs** of over/under predicting
2. **Risk** indices/metrics

Use case 1: Risk management

When do you take an umbrella?

Risk appetite drives many decisions:

- When there is a chance of rain, at what probability do you change your behavior?
- Power system operators hold “reserve” to meet unexpected short-falls in supply. How much?
- Energy traders must balance risk and reward, what kind of risk is acceptable?

Use case 2: Cost-loss

"Optimal" decisions under uncertainty

Example: Consider the occurrence of an uncertain adverse event, E. If E happens, we incur loss L, if not, there is no loss. We have to decide if we should take precautionary action, with cost C, to avoided the potential loss.

	Adverse events occur	Adverse event does not occur
Precautionary action taken	C	C
Precautionary action not taken	L	0

Given the probability of occurrence $P(E)$, when should with take precautionary action?

Use case 1: Cost-loss

"Optimal" decisions under uncertainty

Options:

1. If we take precautionary action, the expected cost is C
2. If we do not, the expected cost is $\mathrm{P}(E) \times L$

Therefore, if $\boldsymbol{P(E) > \frac{C}{L}}$, the expected cost is minimised by taking precautionary action.

	Adverse events occur	Adverse event does not occur
Precautionary action taken	C	C
Precautionary action not taken	L	0

- Many decision-making problems can be expressed in a form like this
- We may have some control over $P(E)$, C, and/or L
- Calculating $P(E)$ requires a **probabilistic forecast**
- Describing the cost and loss can sometimes be more challenging than forecasting

Example: Wind and Price

Optimal trading

- Traders of renewable energy rely heavily on forecasts
- Penalties for over- or under-delivering against a contract are asymmetric
- The optimal volume to trade may be derived from production and price forecasts
- Aim: achieve equal exposure to imbalance costs for over an under production, i.e., $P(E) = \frac{C}{L}$

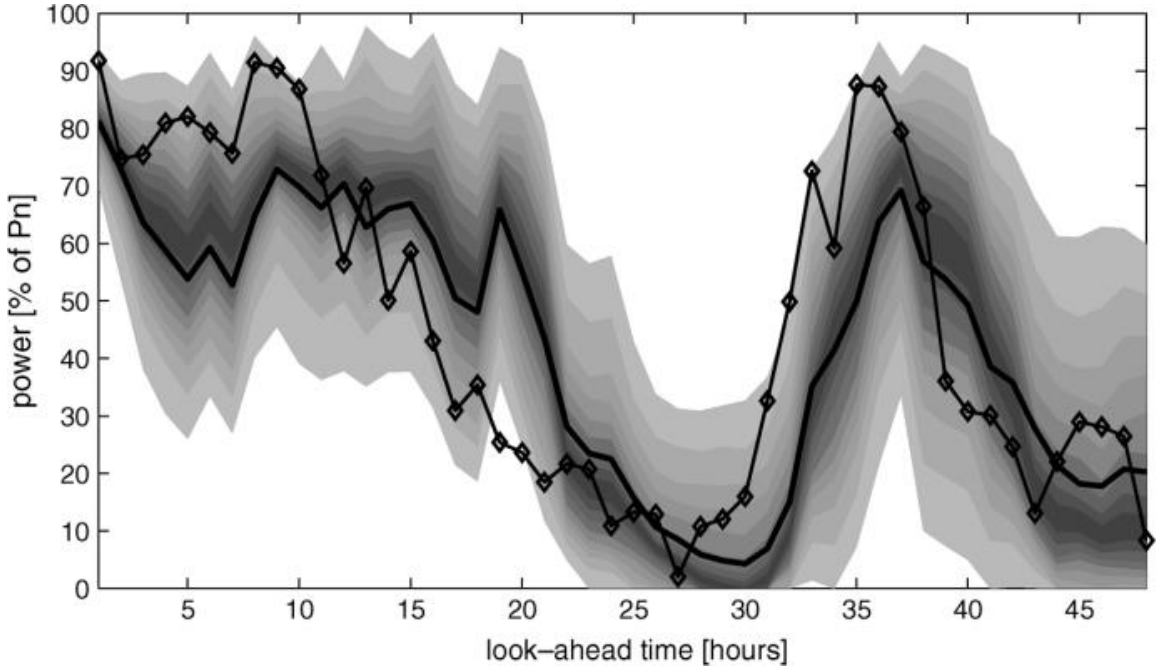


Figure from classic trading example: P. Pinson, C. Chevallier and G. N. Kariniotakis, "Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power," in IEEE Transactions on Power Systems, 2007, doi: 10.1109/TPWRS.2007.901117

Types of probabilistic forecast

Probability and predictions

Desirable properties:

1. Calibration/reliability: Events should occur with the predicted frequency
2. Sharpness: Uncertainty should be minimized/forecast should be as confident as possible
3. Resolution and Discrimination: Forecasts differ from one another, and are different for different outcomes

Note: #1 is a necessary property, #2 & #3 are to be *"optimized"*.

Binary and categorical forecasts

- Predicted probability of a future event: 0-100%, e.g.:
  - Plant failure
  - Generation margin falling below some threshold
  - Loss-of-load (LOLP)
- Probability of multiple (mutually exclusive) categories: 0-100%, summing to 100%, e.g.:
  - Probability of being in a given "*weather regime*"

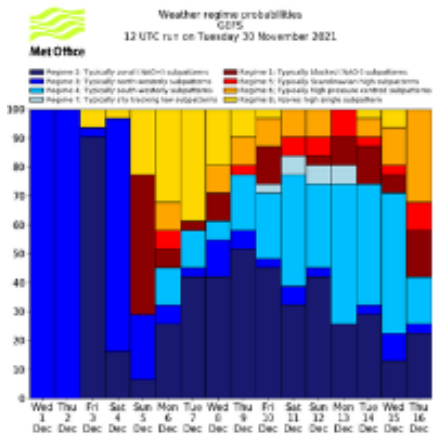
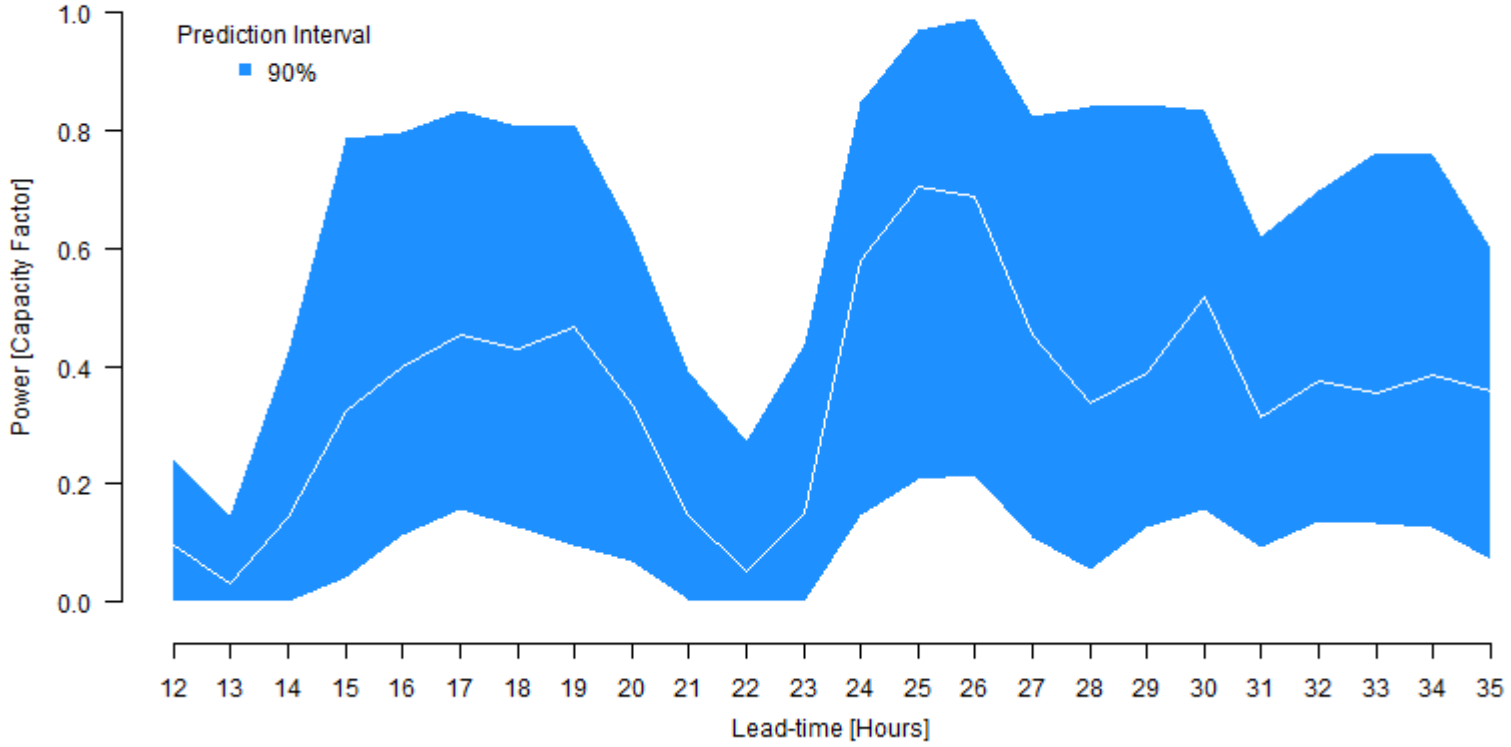


Image source: https://www.metoffice.gov.uk/services/business-industry/energy/decider

Interval and density forecasts



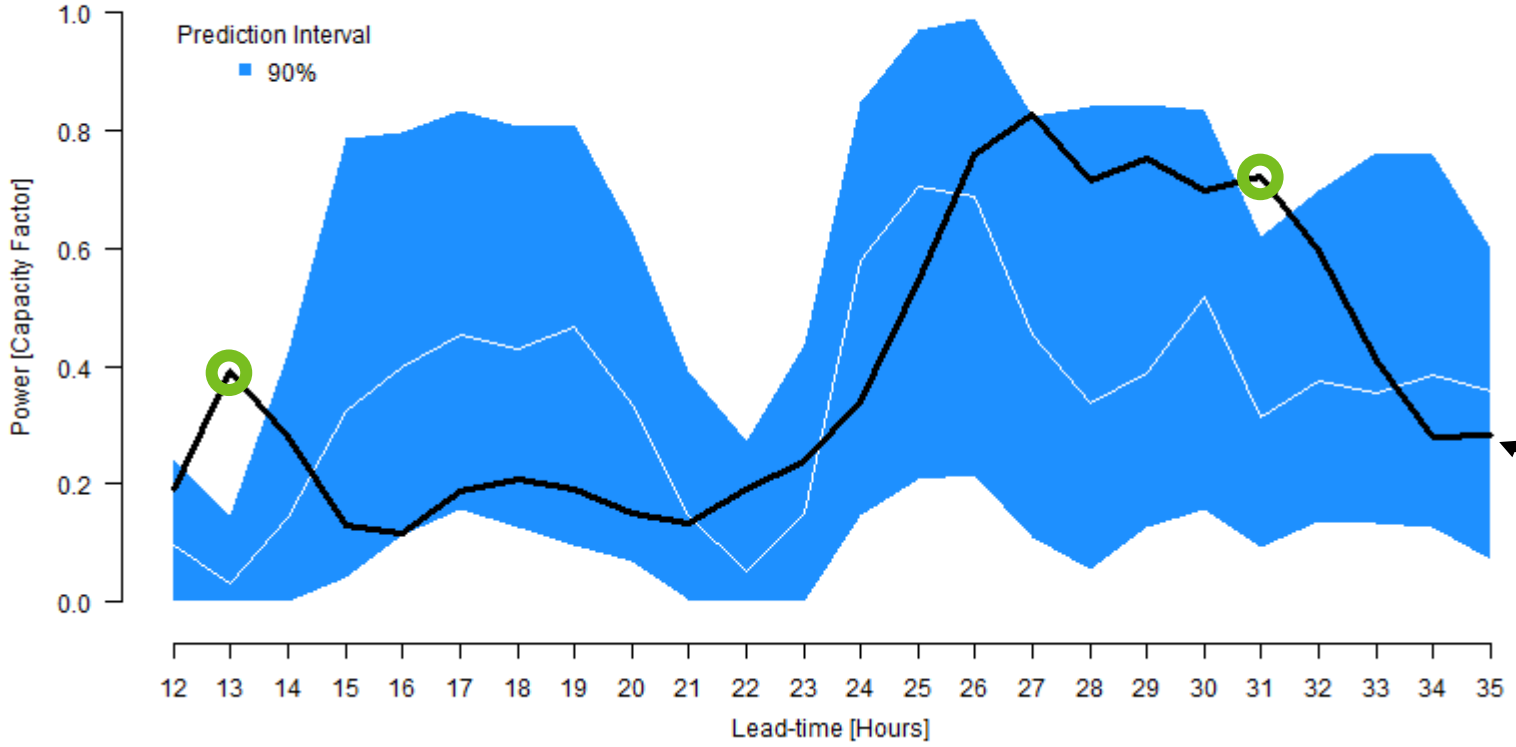
Example: Wind Power Forecast

- 90% prediction interval and deterministic forecast (white)

Notice:

- Variability of interval width
- Asymmetry around deterministic forecast
- How often should we expect the observation to be outside the interval?

Interval and density forecasts



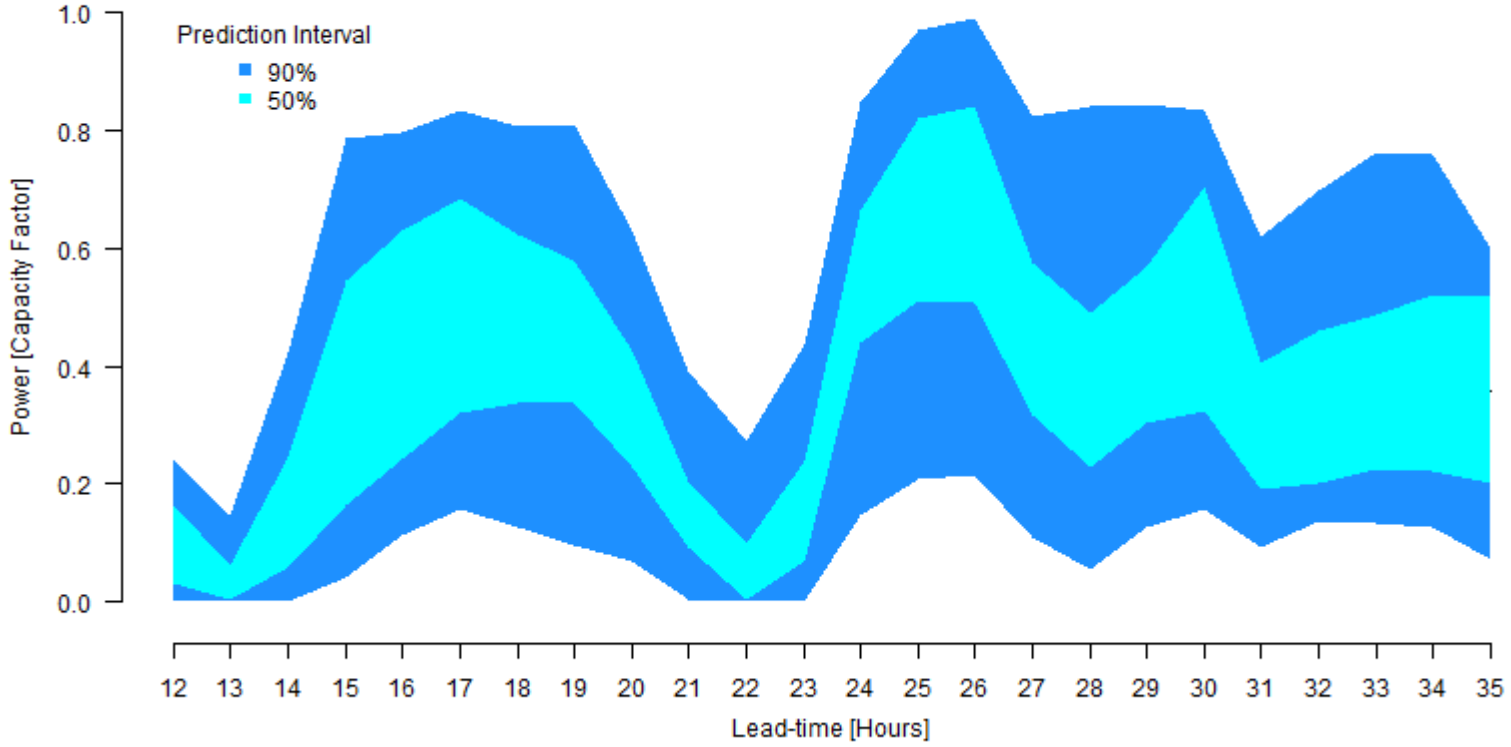
Example: Wind Power Forecast

- 90% prediction interval and deterministic forecast (white)

Notice:

- Variability of interval width
- Asymmetry around deterministic forecast
- How often should we expect the observation to be outside the interval?

Interval and density forecasts



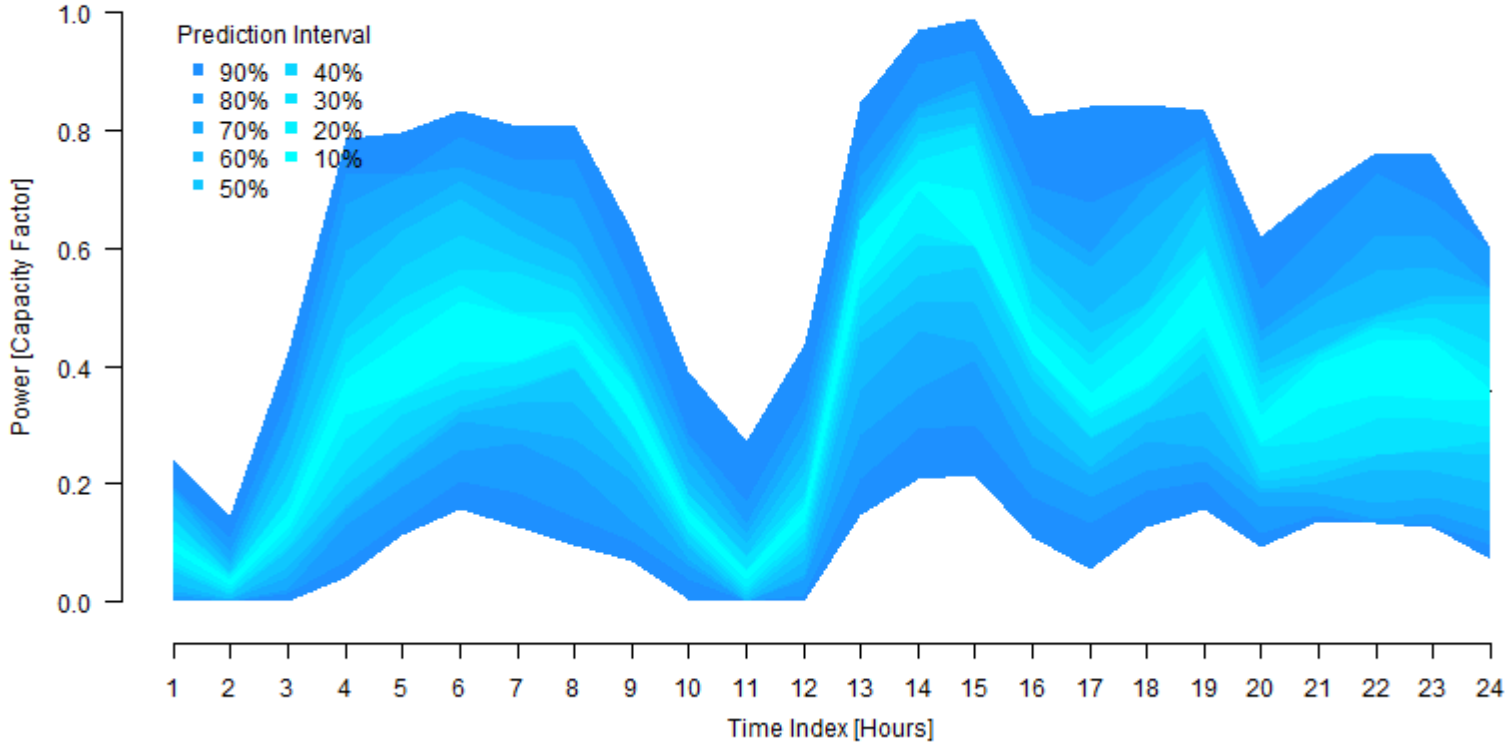
Example: Wind Power Forecast

- 90% and 50% prediction interval

Deterministic forecast not shown:

- Can be misleading: it is not "the most likely trajectory"
- Usually too smooth compared to actual values

Interval and density forecasts



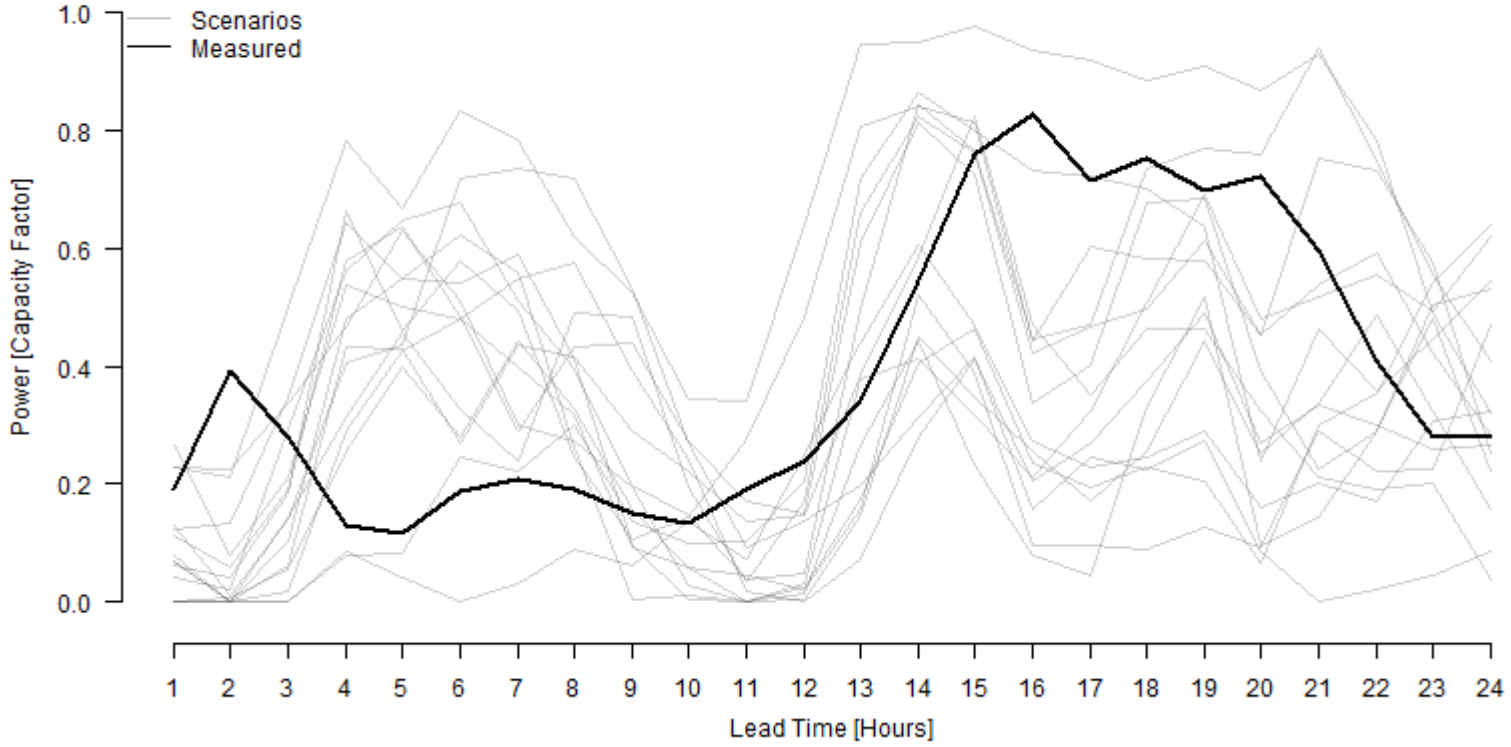
Example: Wind Power Forecast

- 10% to 90% prediction intervals
- We can visualise multiple intervals for more detail
- A density forecast is a smooth function spanning all probability levels

However…

- This forecast doesn't tell us about correlation in time

Scenario/trajectory forecasts

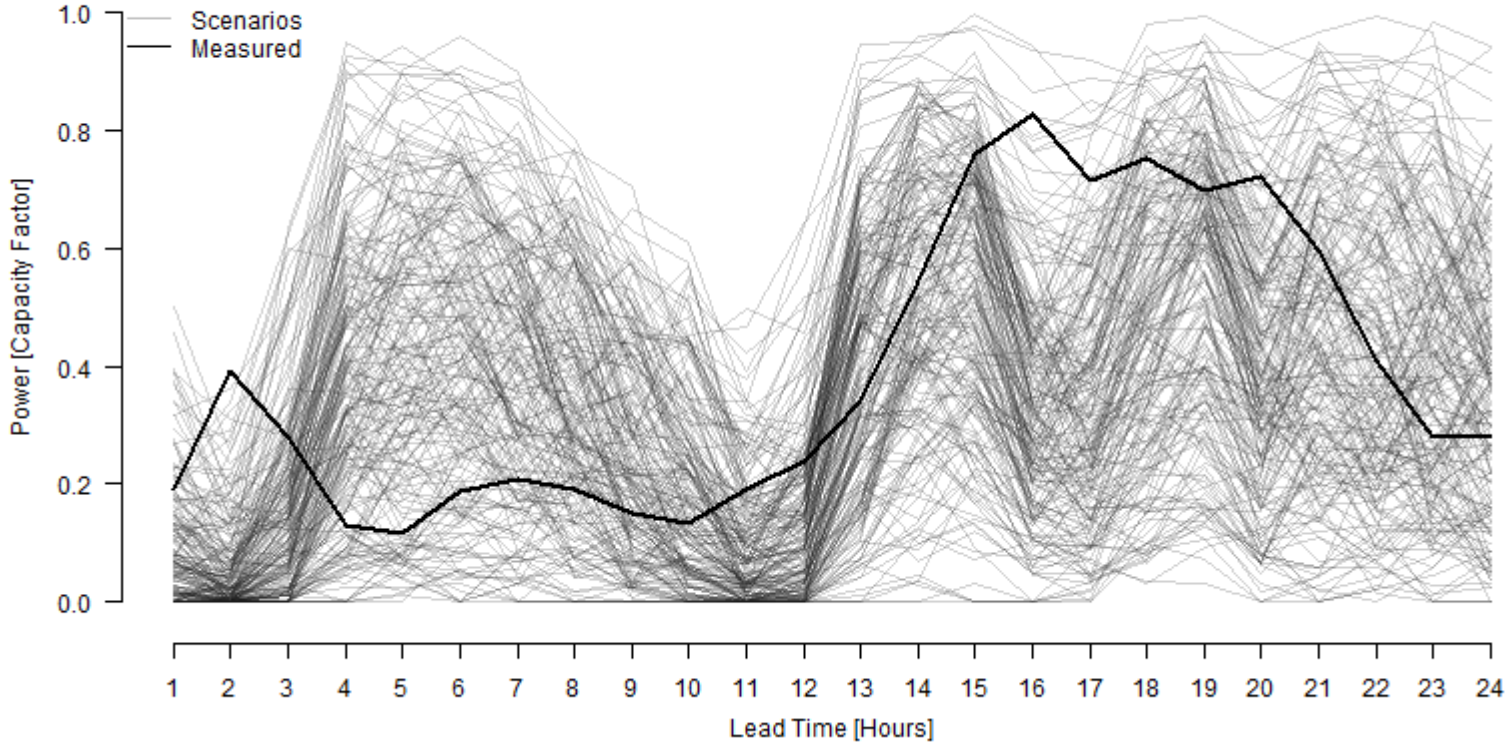


Example: Wind Power Forecast

- 15 scenarios or trajectories
- Temporal correlation is represented…
- …but we've lost the clarity provided by prediction intervals.

Perhaps having more scenarios will help…

Scenario/trajectory forecasts



Example: Wind Power Forecast

- 15 scenarios or trajectories
- Temporal correlation is represented…
- …but we've lost the clarity provided by prediction intervals.

Perhaps having more scenarios will help… if you're a computer !

Principles

Probabilistic energy forecasting from minutes to weeks-ahead

Timescales

	Very Short-term	Short-term	Medium-Term	Long-term
Timescale	Minutes to Hours	Hours to Days	Weeks to Seasons	Years
Applications/ Users	Balancing/TSO Markets/Traders	Markets/Traders Planning/TSO O&M/Operators	O&M/Operators Markets/Traders Planning/TSO	Planning and investment/ TSO, policy makers, developers
Methodology	Time series methods: statistical, machine learning	Numerical Weather Prediction, NWP post-processing, data-driven and physical/fundamental models		Climate projections, Technological forecasting

Timescales

	Very Short-term	Short-term	Medium-Term	Long-term
Timescale	Minutes to Hours	Hours to Days	Weeks to Seasons	Years
Applications/ Users	Balancing/TSO Markets/Traders	Markets/Traders Planning/TSO O&M/Operators	O&M/Operators Markets/Traders Planning/TSO	Planning and investment/ TSO, policy makers, developers
Methodology	Time series methods: statistical, machine learning	Numerical Weather Prediction, NWP post-processing, data-driven and physical/fundamental models		Climate projections, Technological forecasting

Timescales

	Very Short-term	Short-term	Medium-Term	Long-term
Timescale	Minutes to Hours	Hours to Days	Weeks to Seasons	Years
Applications/ Users	Balancing/TSO Markets/Traders	Markets/Traders Planning/TSO O&M/Operators	O&M/Operators Markets/Traders Planning/TSO	Planning and investment/ TSO, policy makers, developers
Methodology	Time series methods: statistical, machine learning	Numerical Weather Prediction, NWP post-processing, data-driven and physical/fundamental models		Climate projections, Technological forecasting

Timescales

	Very Short-term	Short-term	Medium-Term	Long-term
Timescale	Minutes to Hours	Hours to Days	Weeks to Seasons	Years
Applications/ Users	Balancing/TSO Markets/Traders	Markets/Traders Planning/TSO O&M/Operators	O&M/Operators Markets/Traders Planning/TSO	Planning and investment/ TSO, policy makers, developers
Methodology	Time series methods: statistical, machine learning	Numerical Weather Prediction, NWP post-processing, data-driven and physical/fundamental models		Climate projections, Technological forecasting

Principles: Physical and fundamental models

Sometimes the best way to produce a forecast is to model the underlying physical or economic relationships…

- **Weather forecasting:** applying knowledge of fluid dynamics and atmospheric physics to compute future weather
- **Electricity price forecasting:** fundamentals, e.g. price of fuels, plant efficiencies, load, (+ efficient market hypothesis)
- **Wind and solar forecasting:** physical models of energy conversion, e.g. wind turbine power curves and PV panel efficiency

Principles: Physical models

Numerical Weather Prediction

The NWP process:

1. State of the atmosphere ***estimated*** based on observations

2. Equations describing atmospheric physics solved for each cell and its neighbours to estimate future state:

Navier-Stokes Equations

- Set of partial differential equations that describe motion of a viscous fluid.
- Conservation of: Mass, Momentum & Energy
- Field of computational fluid dynamics
- Some processes are “parametrised” if can’t be solved explicitly

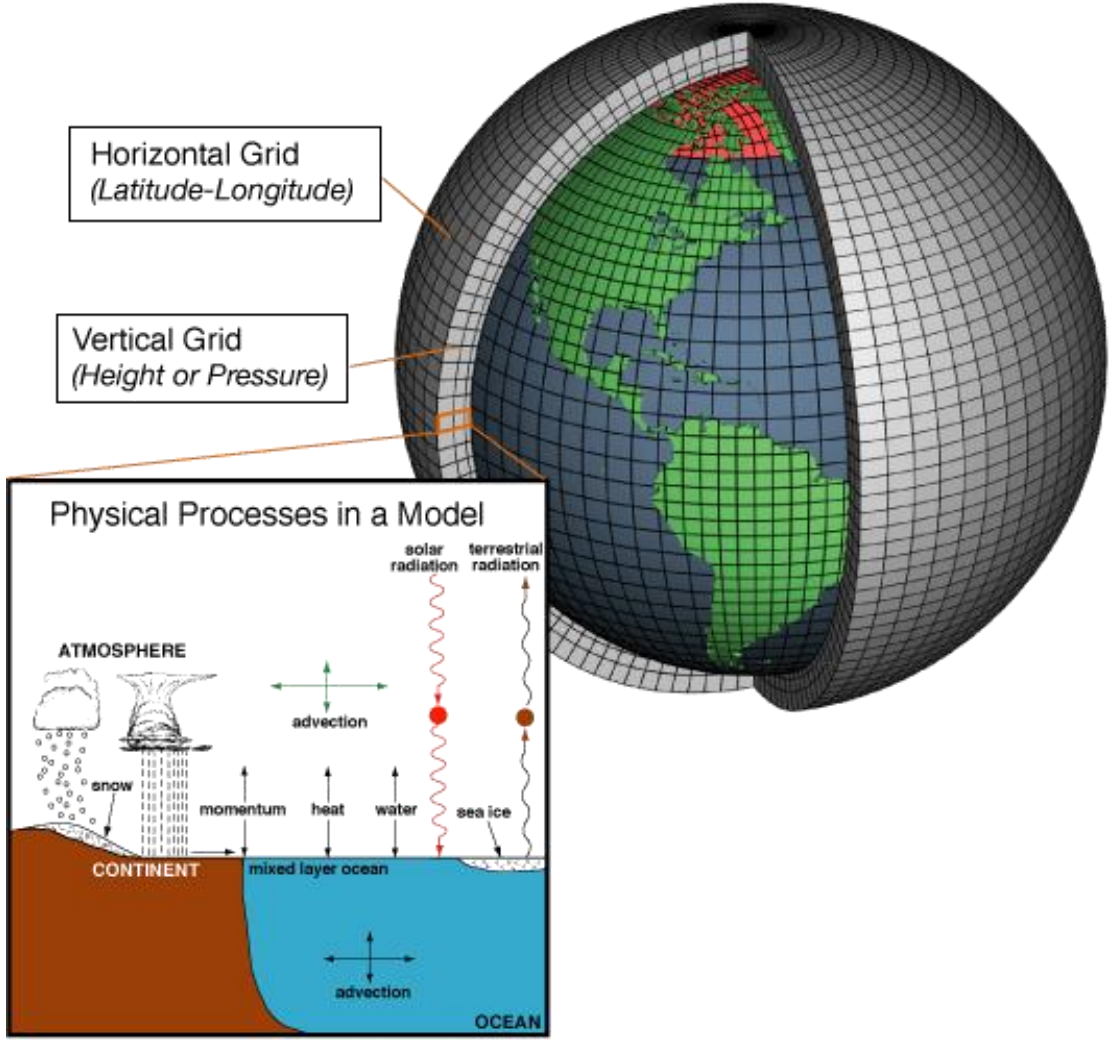


Fig: NOAA.gov

Principles: Physical models

Numerical Weather Prediction

Producing NWP is computationally demanding, placing limitations on frequency and resolution of forecasts:

Temporal Resolution

- Typically, new forecasts issued 4 times per-day
- Sometimes more, e.g. hourly updates for short horizons
- Forecast output available in 1 or 3 hour time steps

Spatial Resolution

- Global models range from 1° to 0.1^o lat x lon (100km-10km). Small (and fast) scale features and processes not resolved
- Local models may be 1-10km resolution, complex terrain and convection may be resolved

Number of ensemble members…

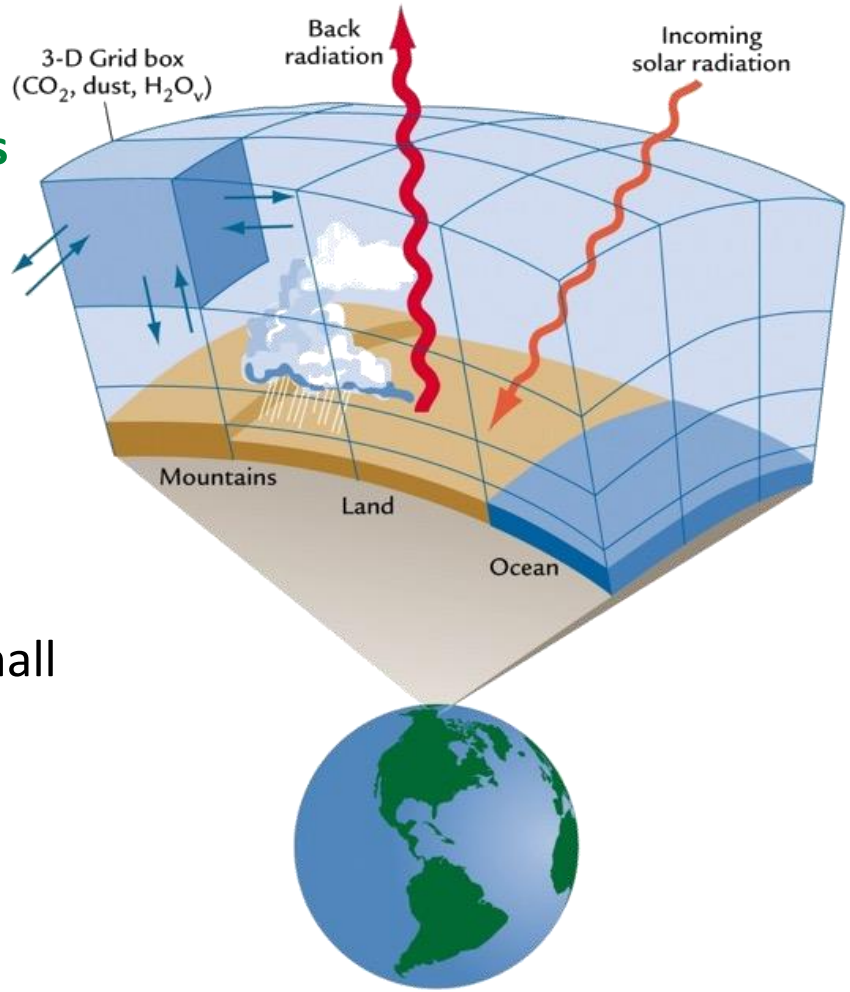


Fig: NOAA.gov

Principles: Physical models

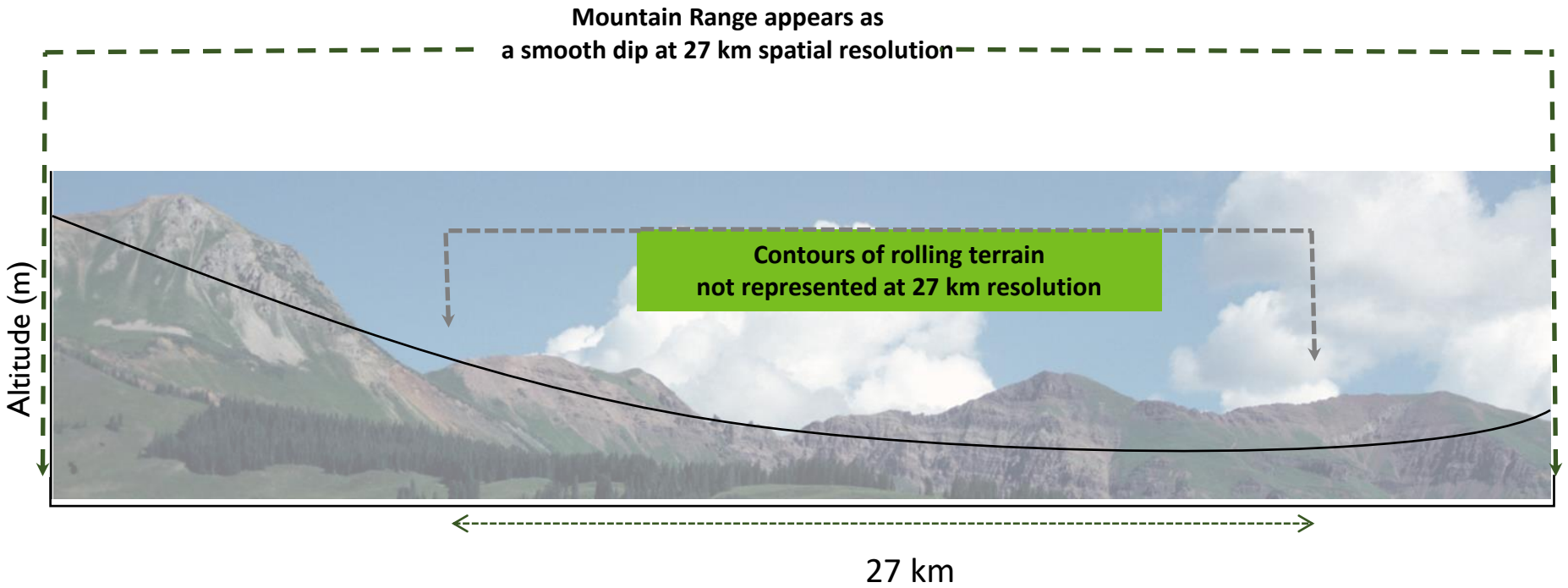
Numerical Weather Prediction

Example: Impact of Terrain and Spatial Resolution of Model

Illustrative example adapted from Justin Sharp, Sharply Focused

Principles: Physical models

Numerical Weather Prediction



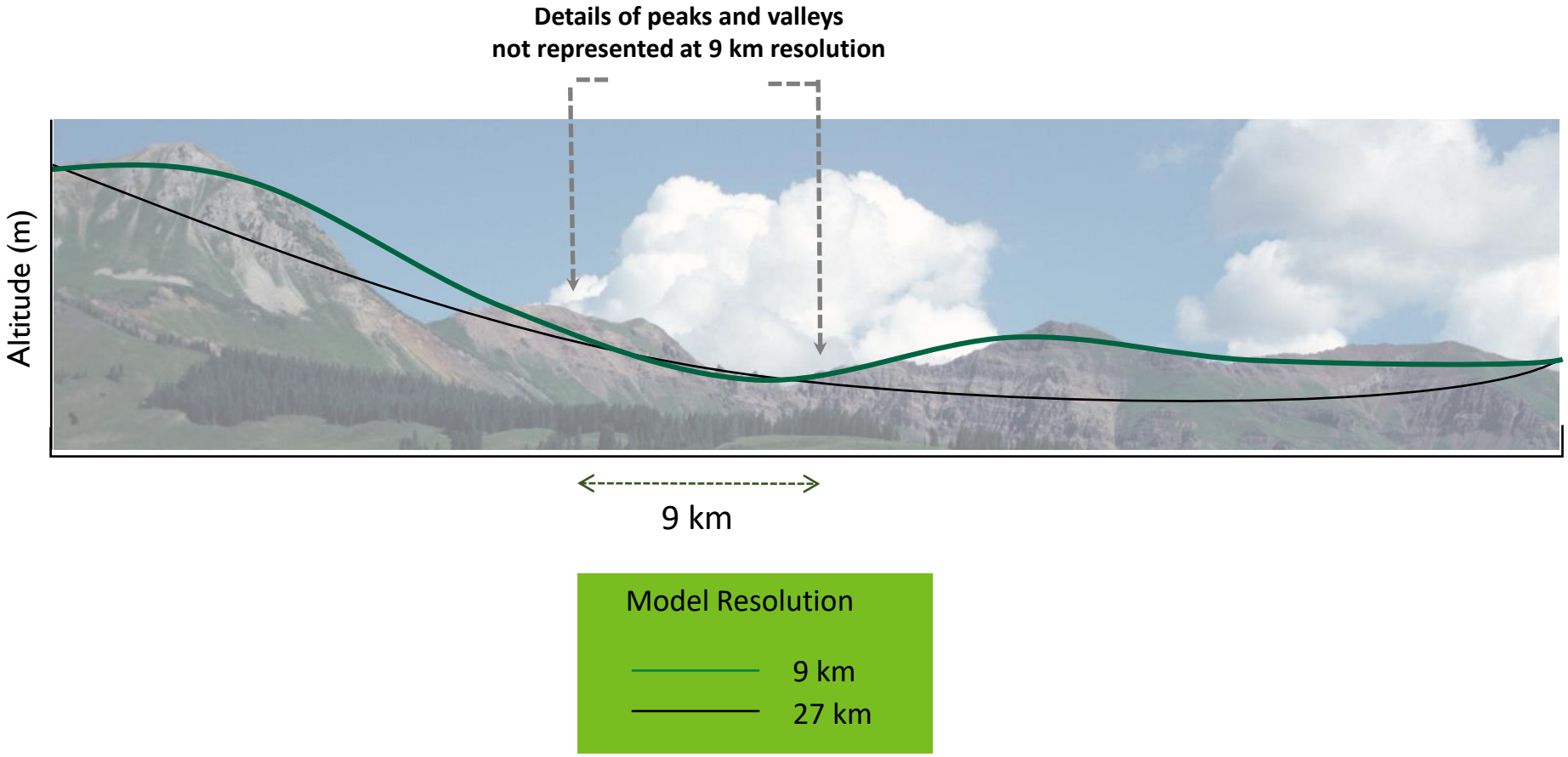
Model Resolution

27 km

Illustrative example adapted from Justin Sharp, Sharply Focused

Principles: Physical models

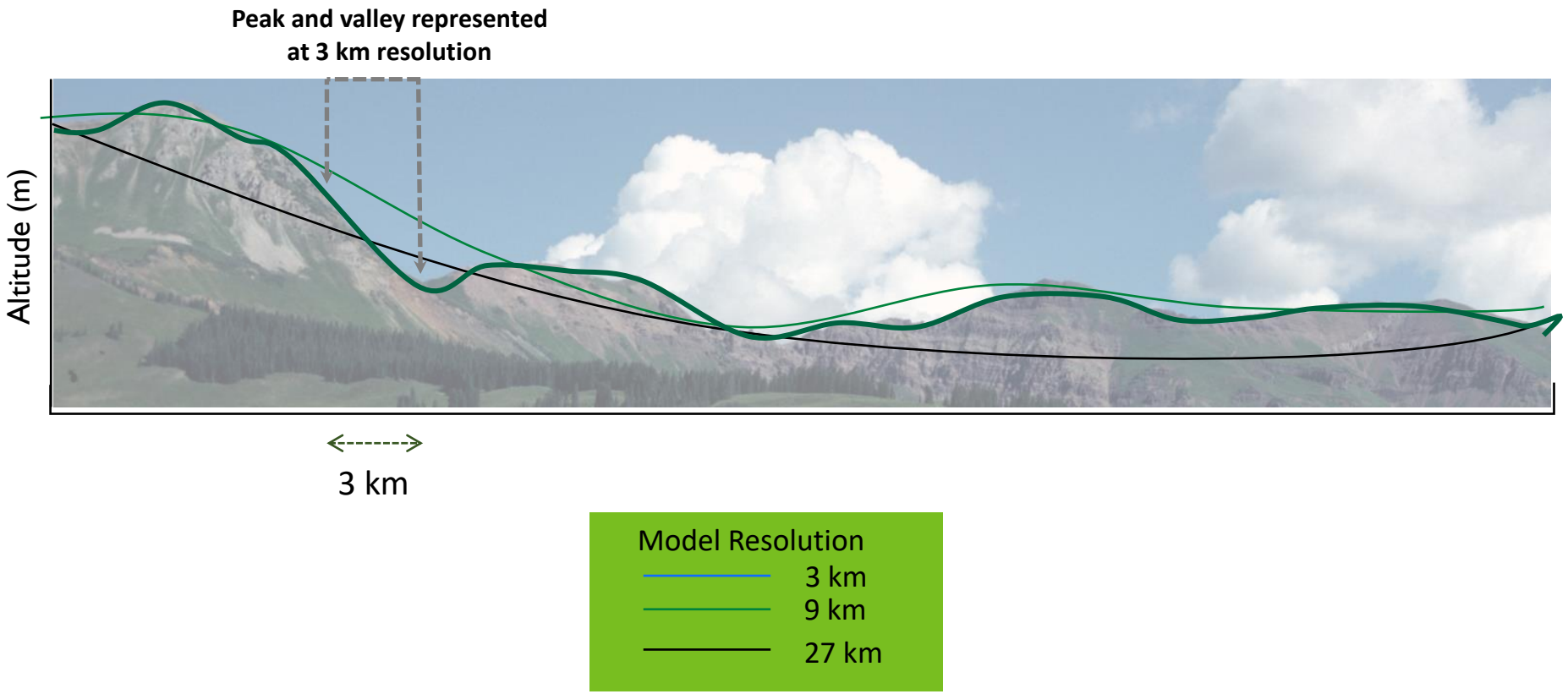
Numerical Weather Prediction



Illustrative example adapted from Justin Sharp, Sharply Focused

Principles: Physical models

Numerical Weather Prediction



Illustrative example adapted from Justin Sharp, Sharply Focused

Principles: Physical models

Ensemble Numerical Weather Prediction

Quantification of uncertainty cause by imperfect initial estimate of atmosphere and model physics

- Multiple forecasts produced with different initial conditions
- *Spread* of ensemble members indicates uncertainty in forecast
- Most valuable in “medium-range” and beyond (3+ days-ahead)
- As we look further ahead, ensemble approaches “climatology”, i.e. the range of conditions typical of the time of year

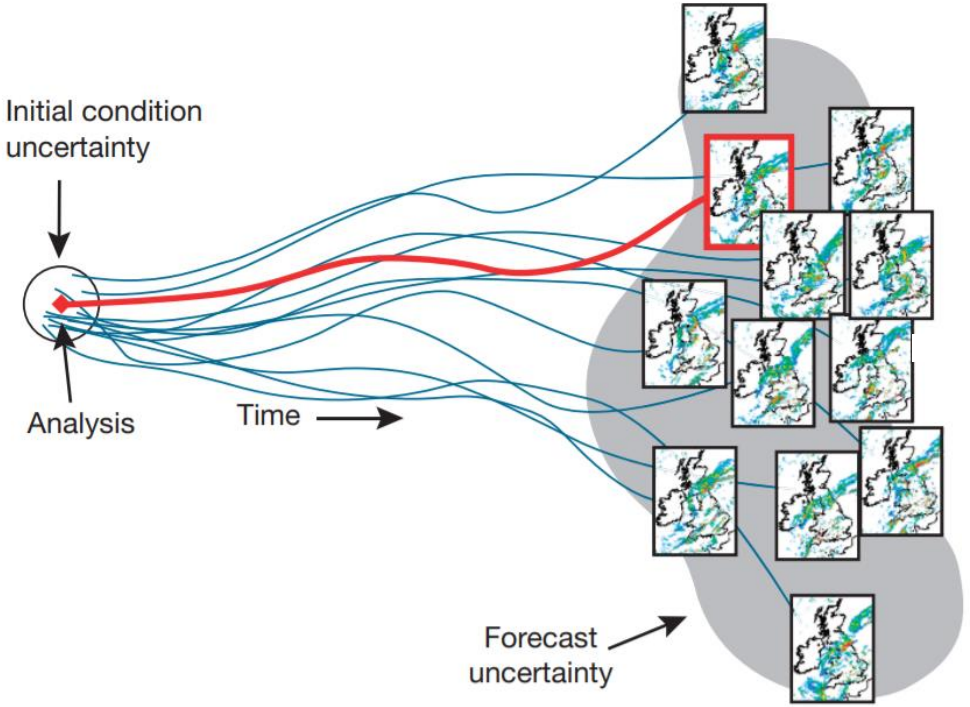


Figure source: Bauer, P., Thorpe, A. & Brunet, G. The quiet revolution of numerical weather prediction. *Nature* **525,** 47–55 (2015). https://doi.org/10.1038/nature14956

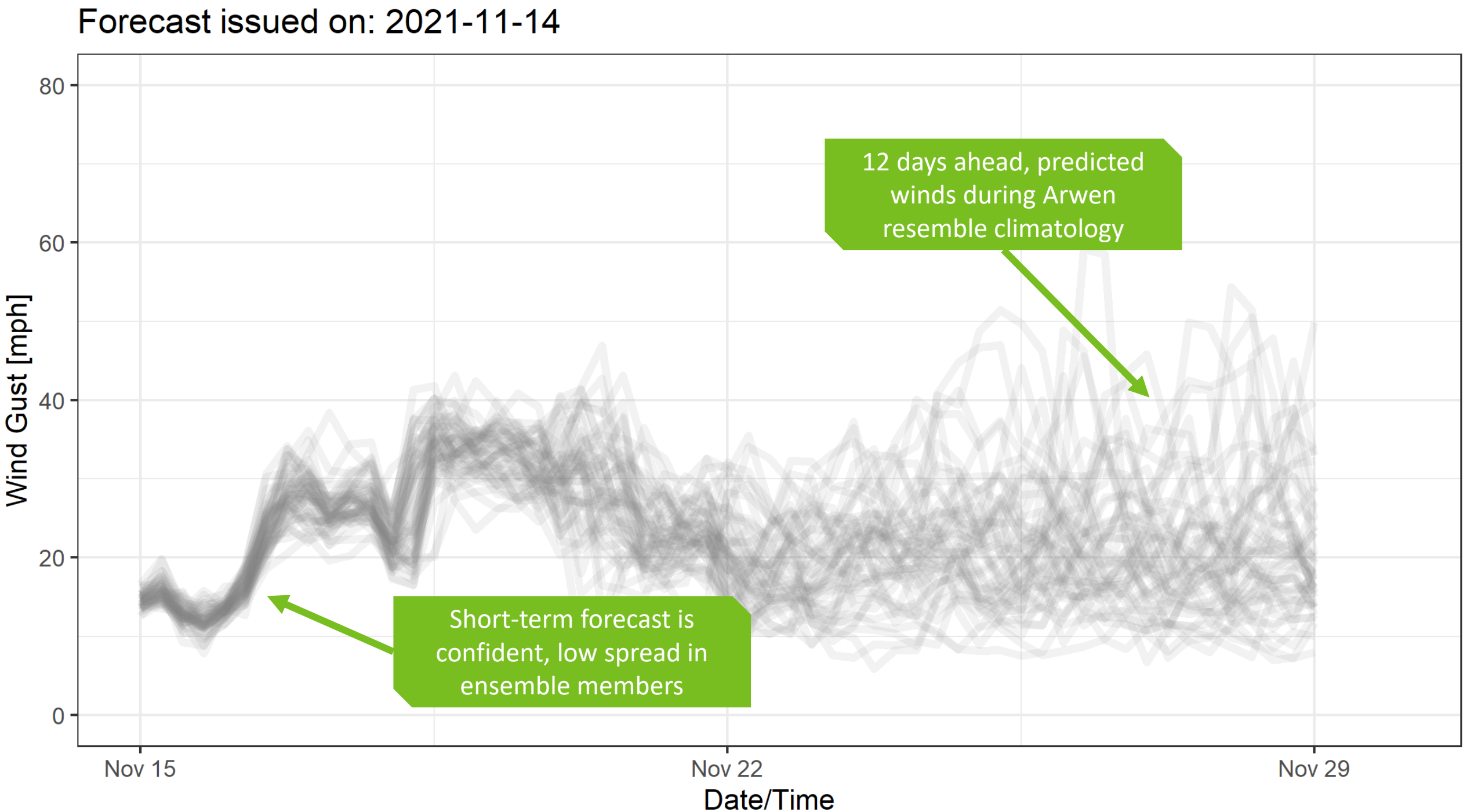
Principles: Physical models

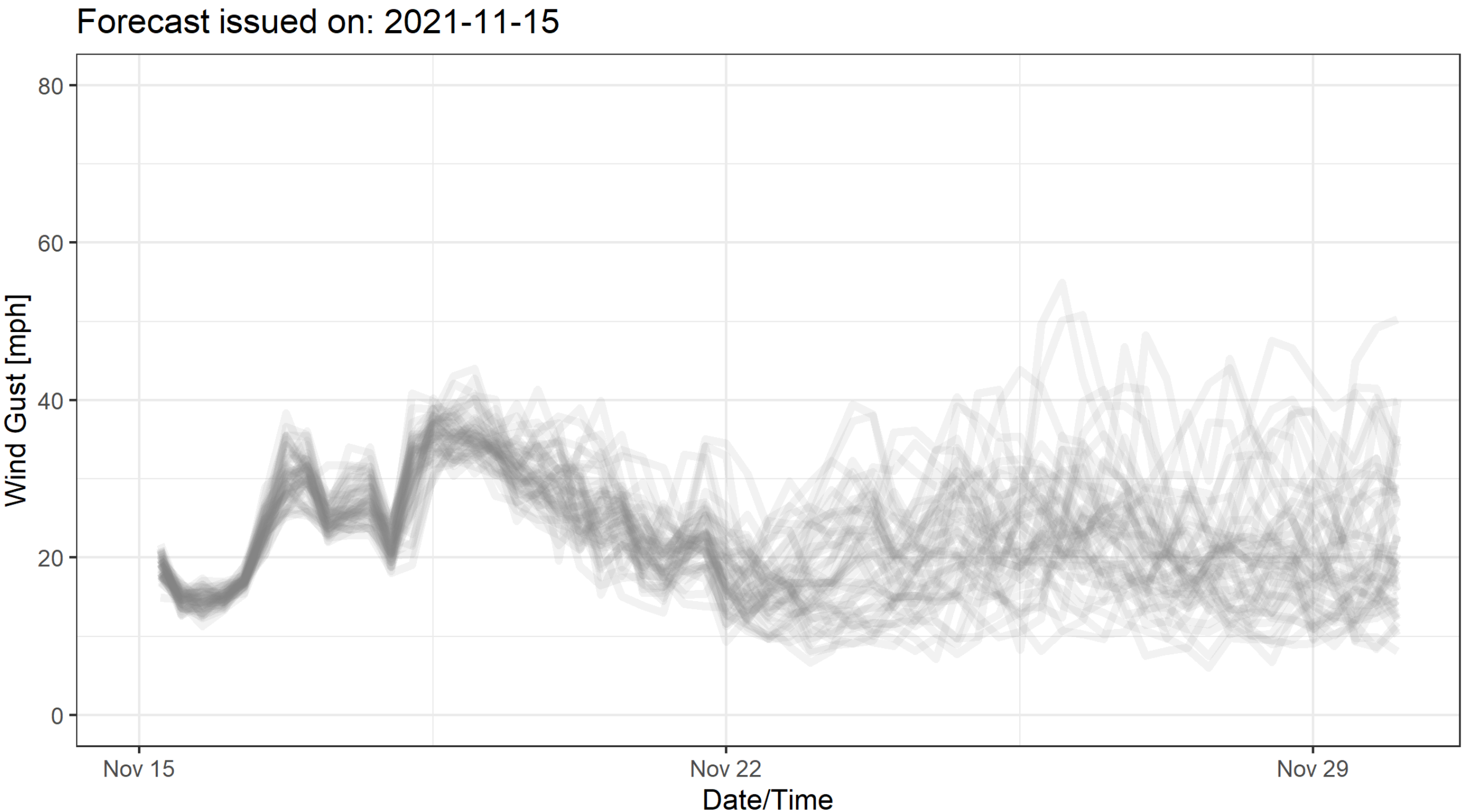
Ensemble Numerical Weather Prediction

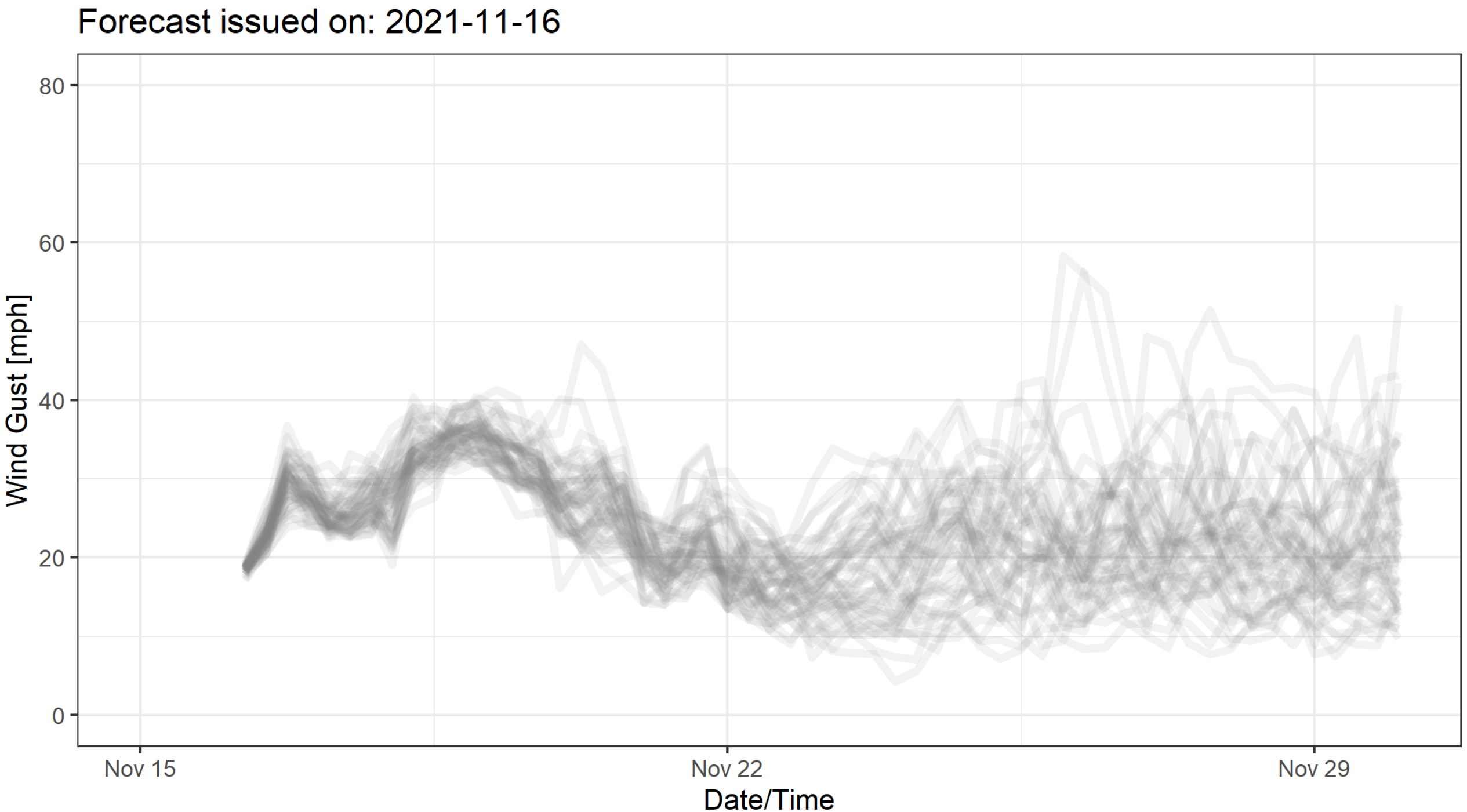
Example: **Storm Arwen**

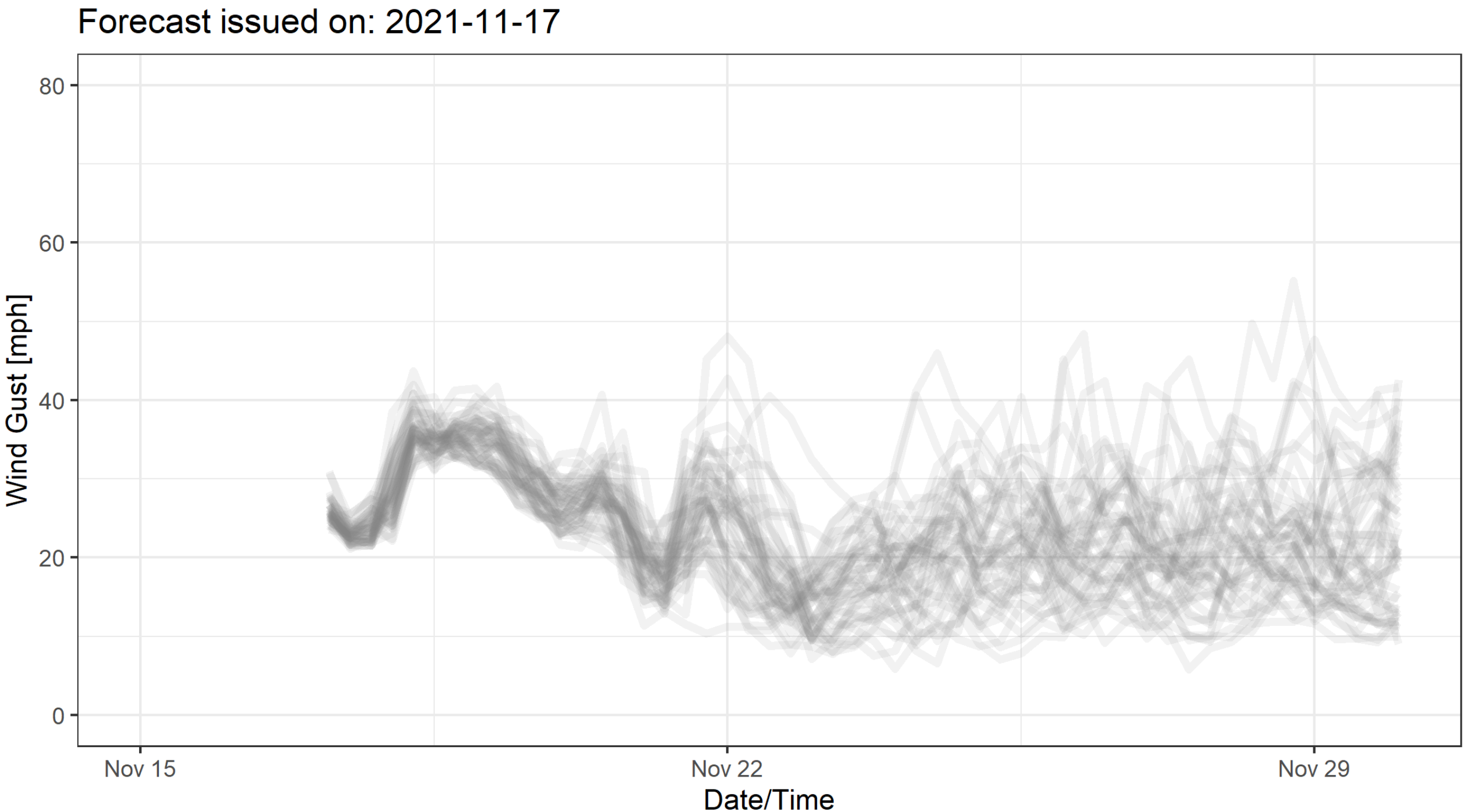
- Major storm in UK and Europe
- 26-28 November 2021
- High winds and snow
- Over 1 million customers affected
- 10,000 still without power a week later

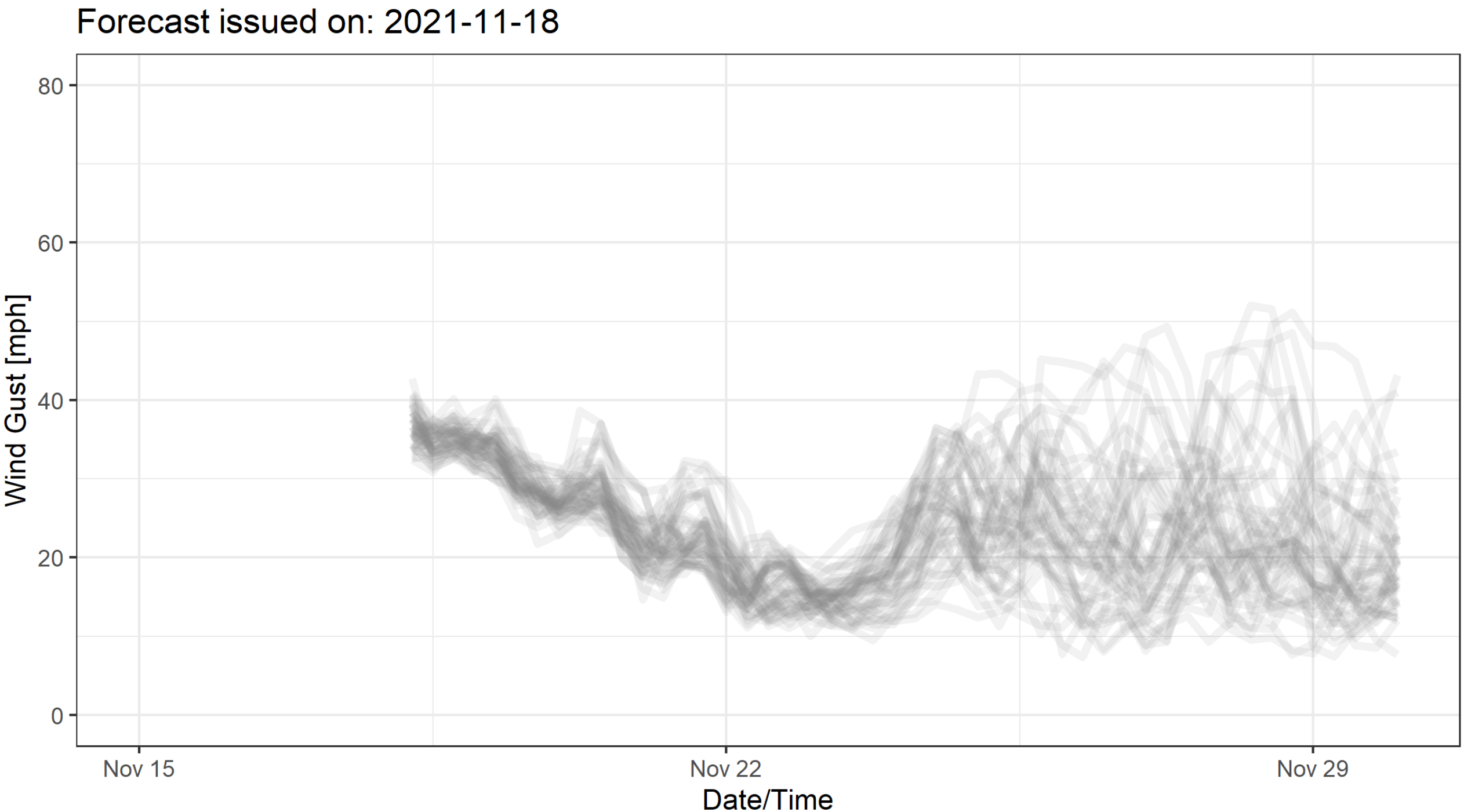
Lets see how the ensemble weather forecast captured this event…

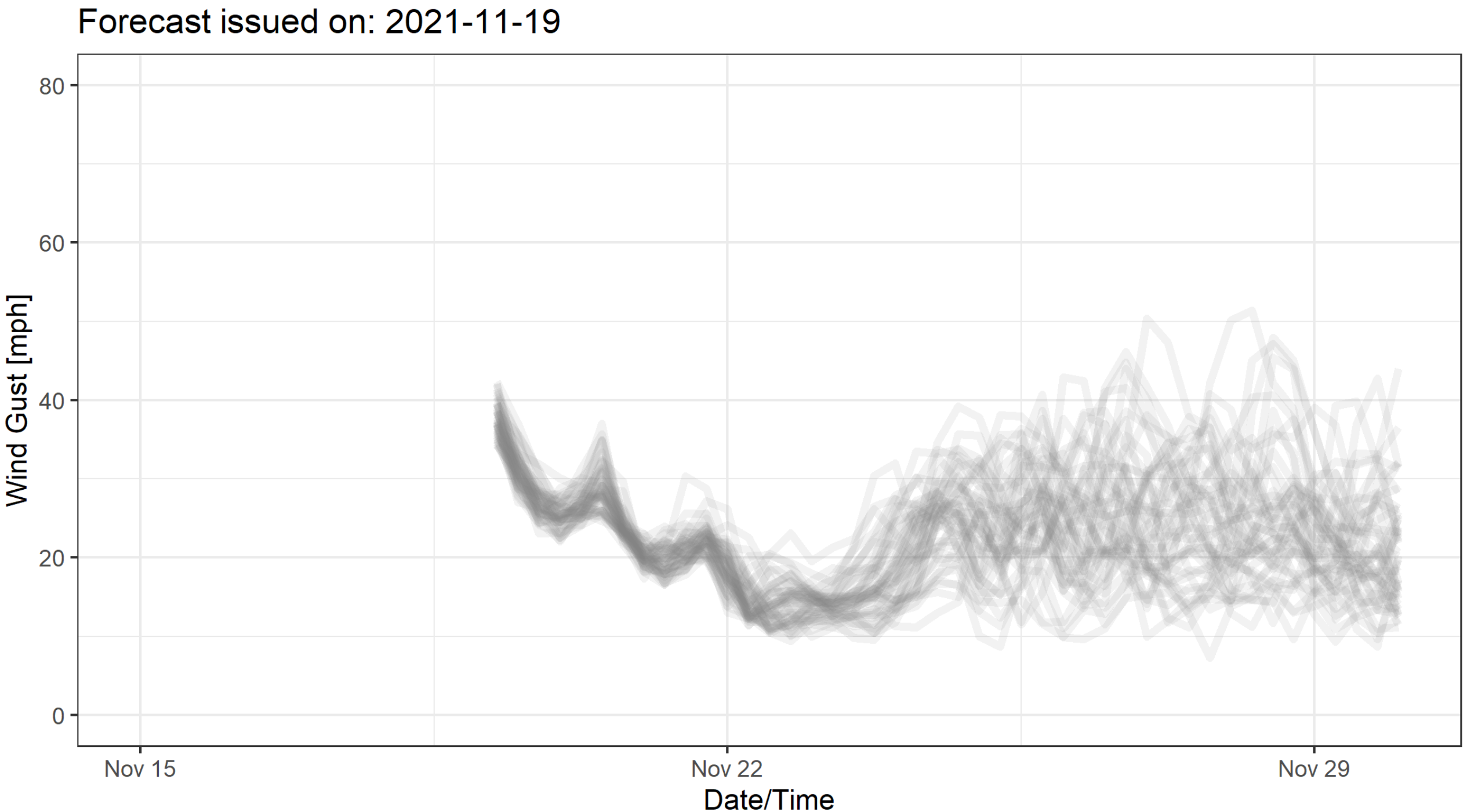

Forecast issued on: 2021-11-14
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time
12 days ahead, predicted winds during Arwen resemble climatology
Short-term forecast is confident, low spread in ensemble members

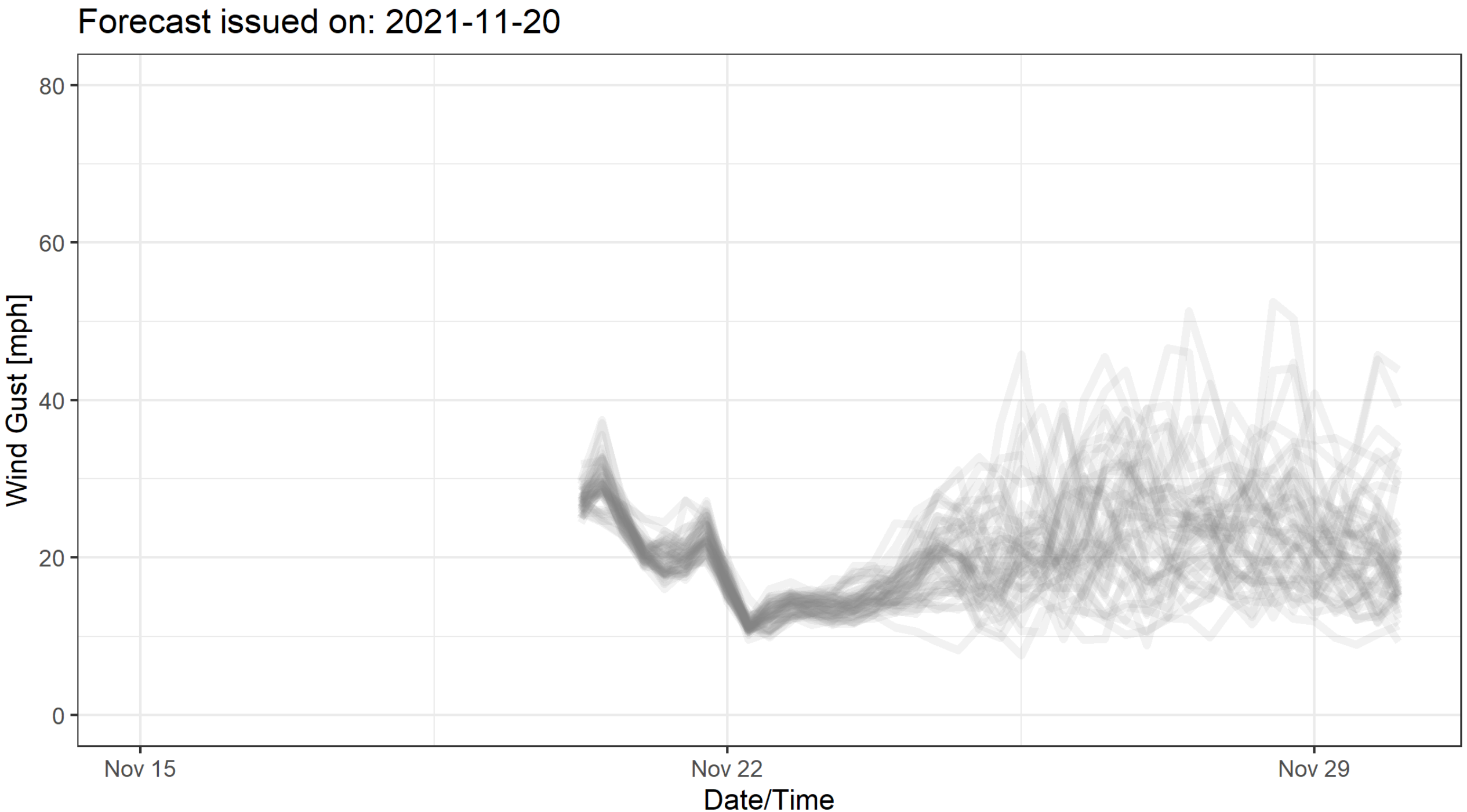

Forecast issued on: 2021-11-15
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

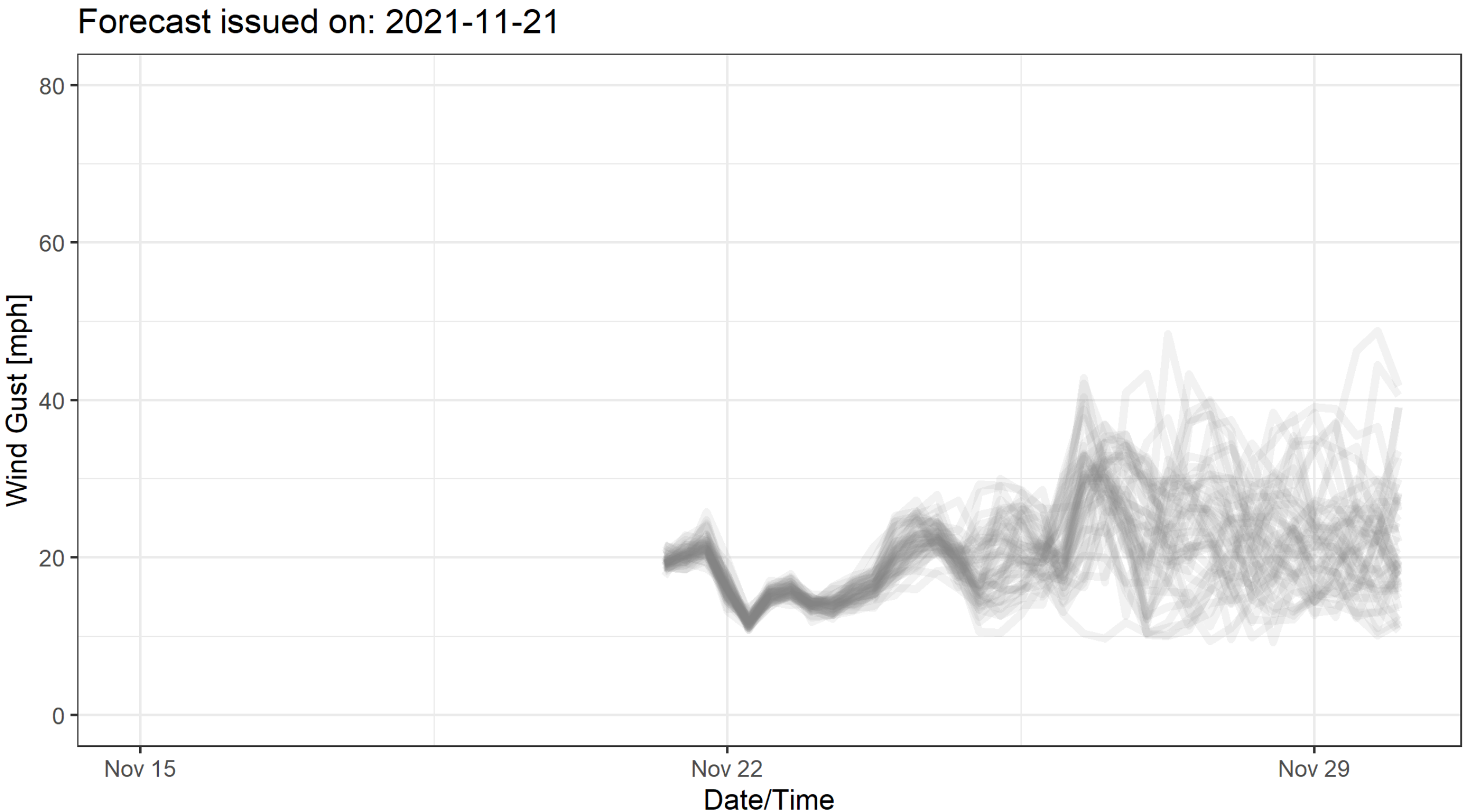

Forecast issued on: 2021-11-16
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

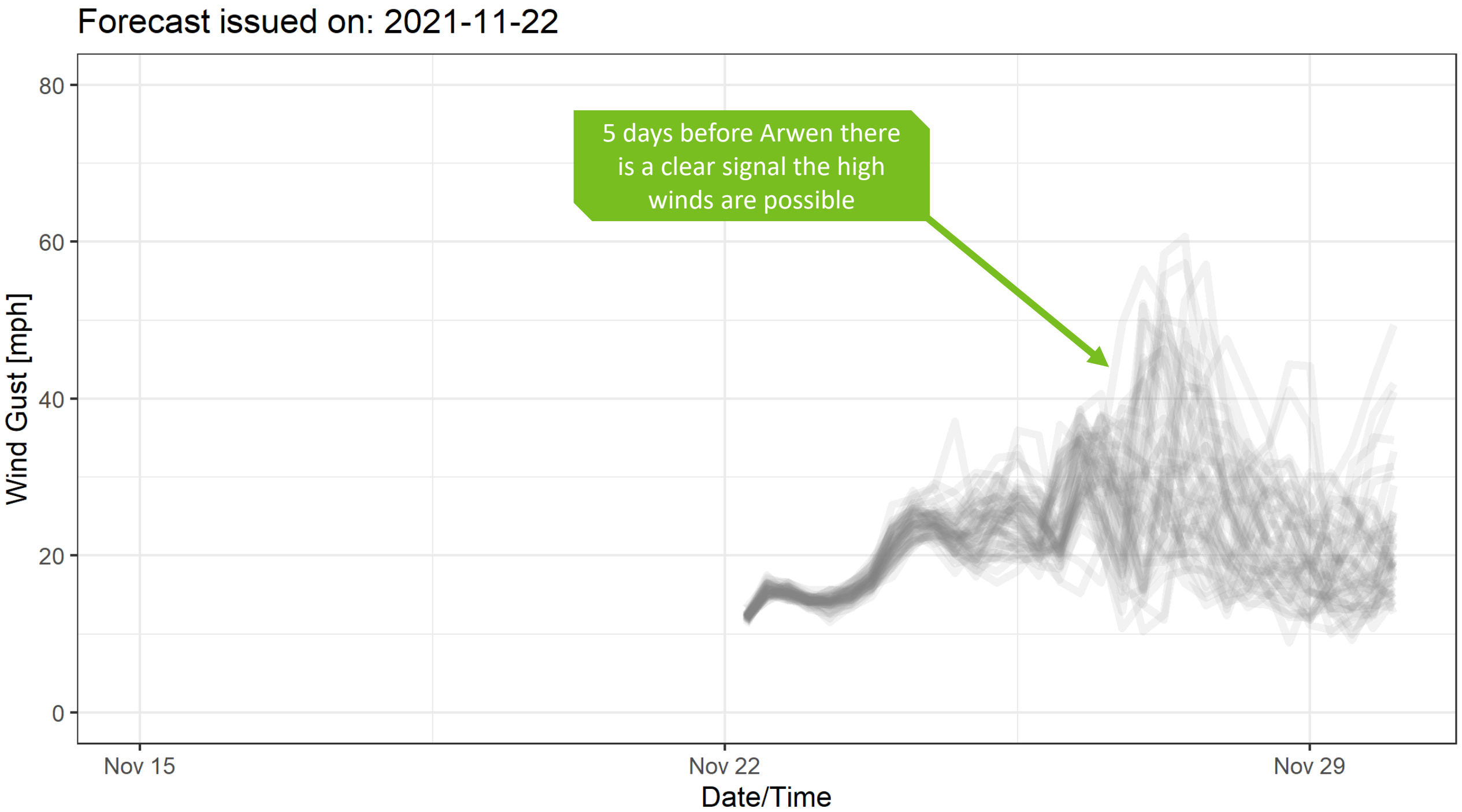

Forecast issued on: 2021-11-17
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

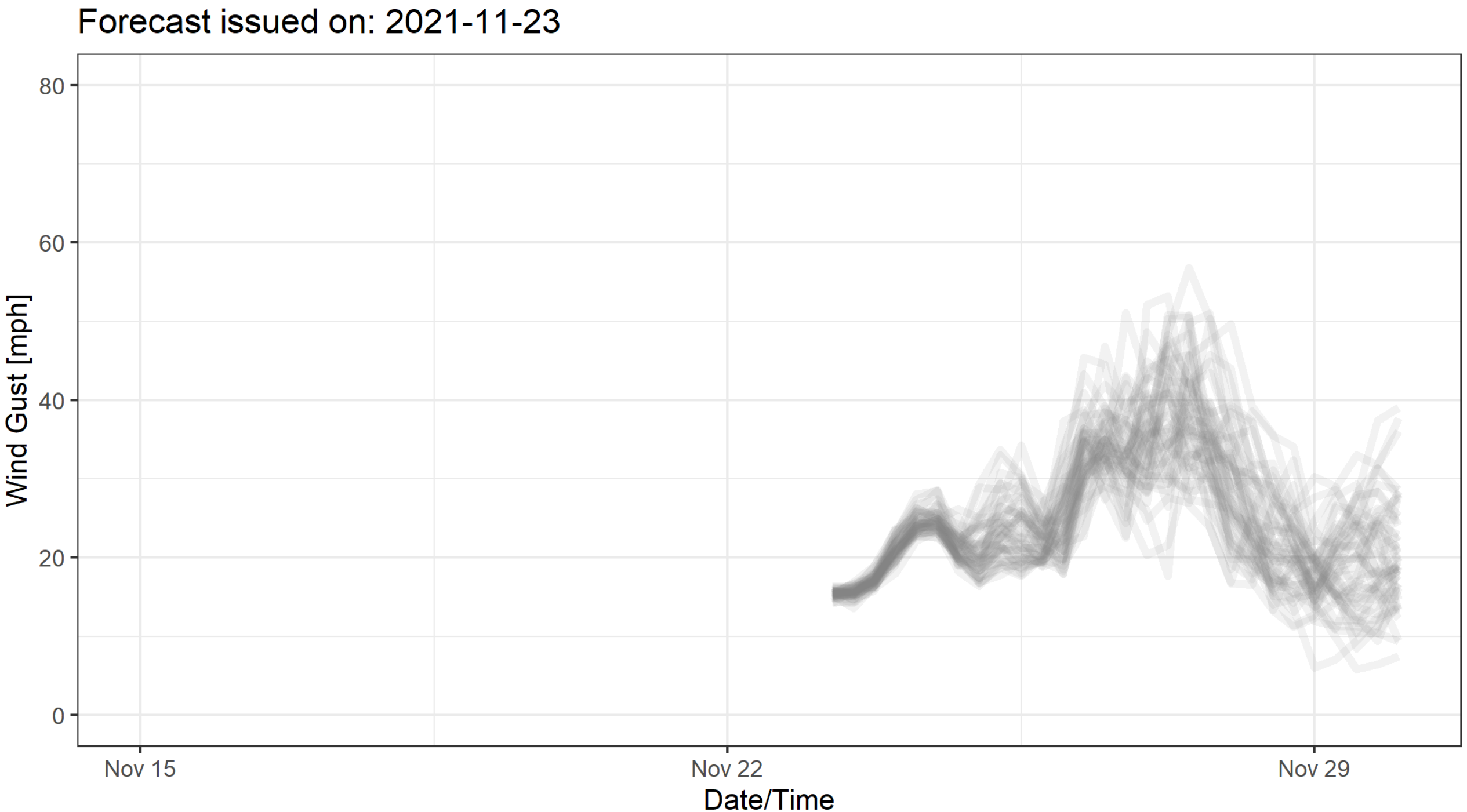

Forecast issued on: 2021-11-18
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

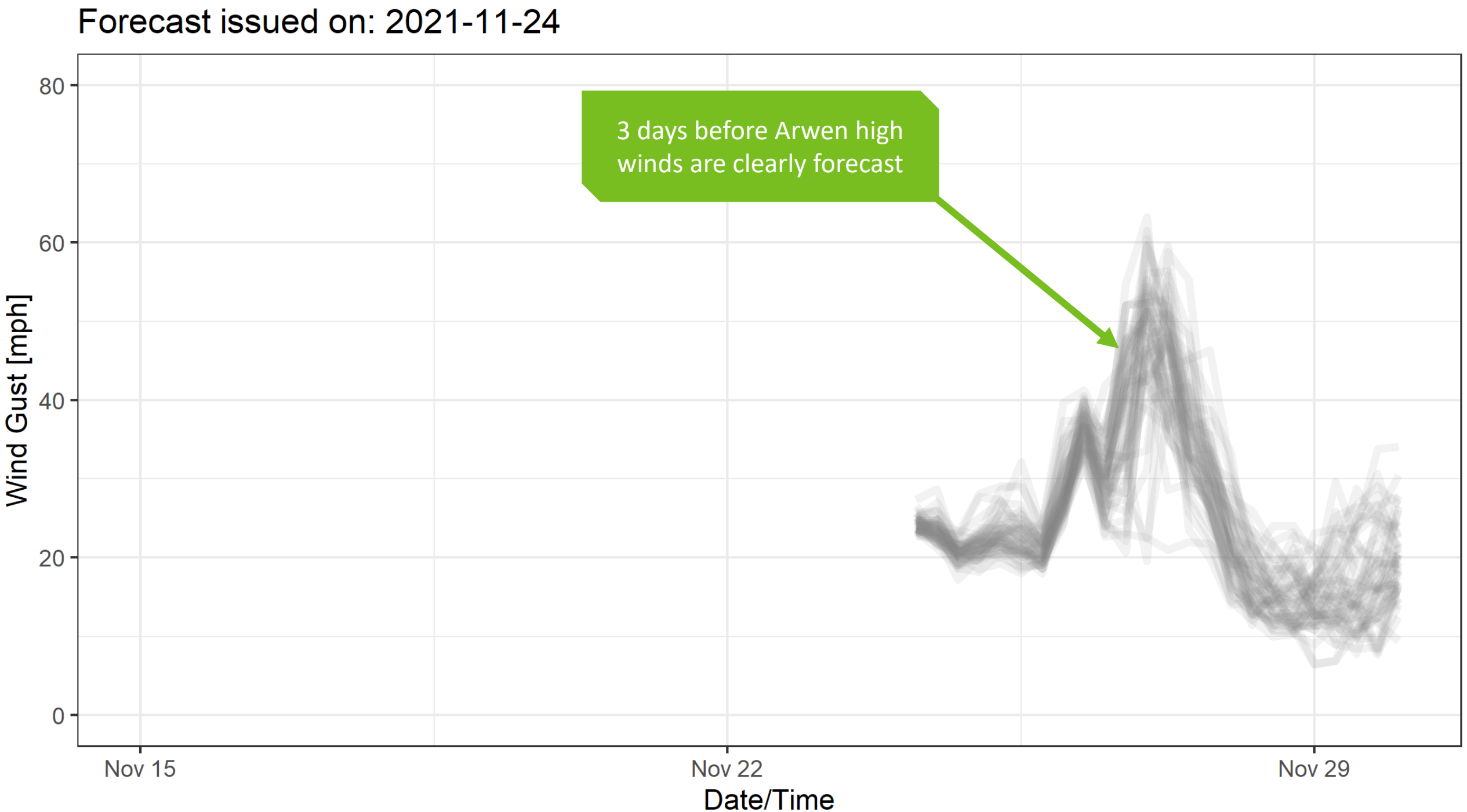

Forecast issued on: 2021-11-19
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

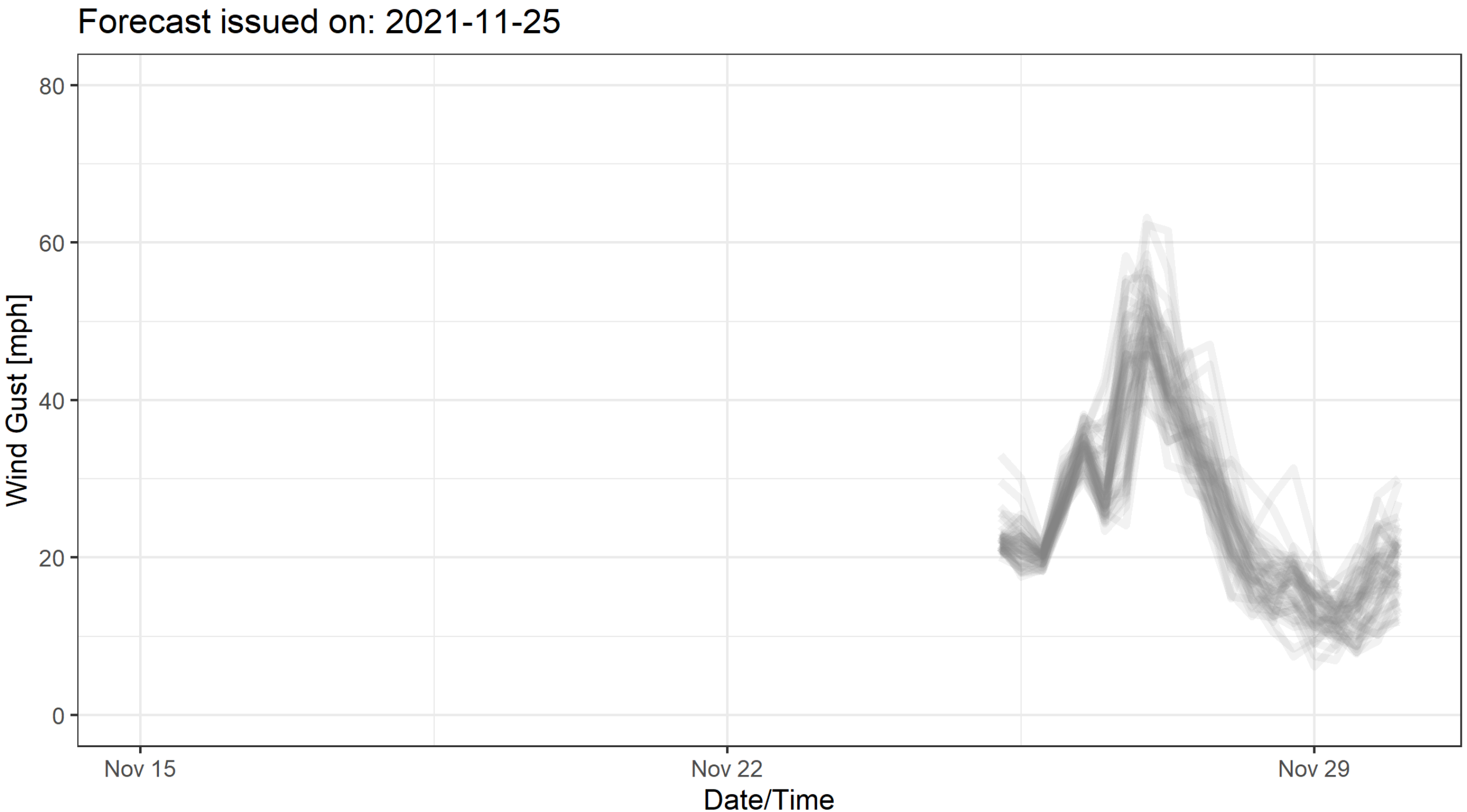

Forecast issued on: 2021-11-20
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time


Forecast issued on: 2021-11-21
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

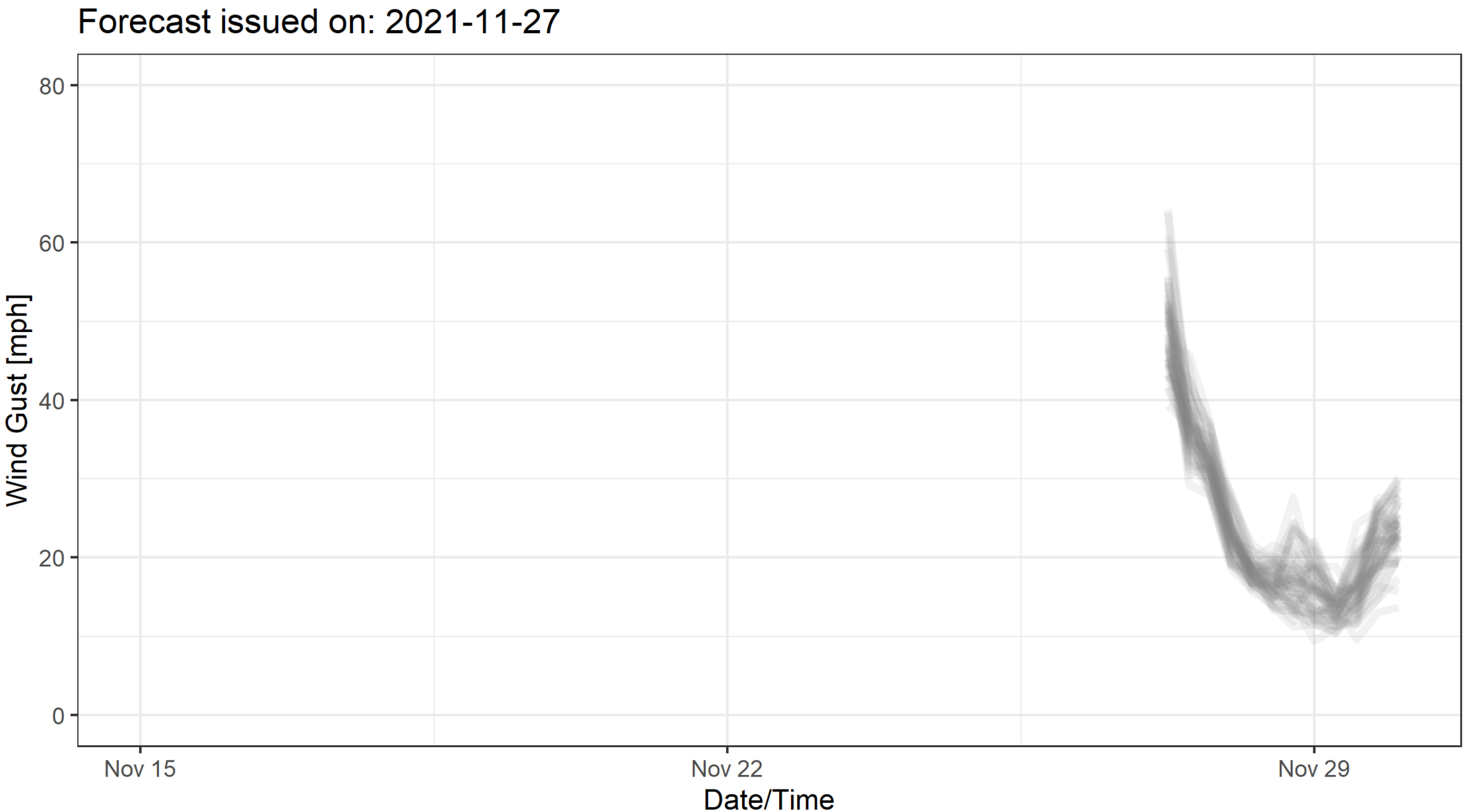

Forecast issued on: 2021-11-22
5 days before Arwen there is a clear signal the high winds are possible
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time


Forecast issued on: 2021-11-23
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time


Forecast issued on: 2021-11-24
3 days before Arwen high winds are clearly forecast
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

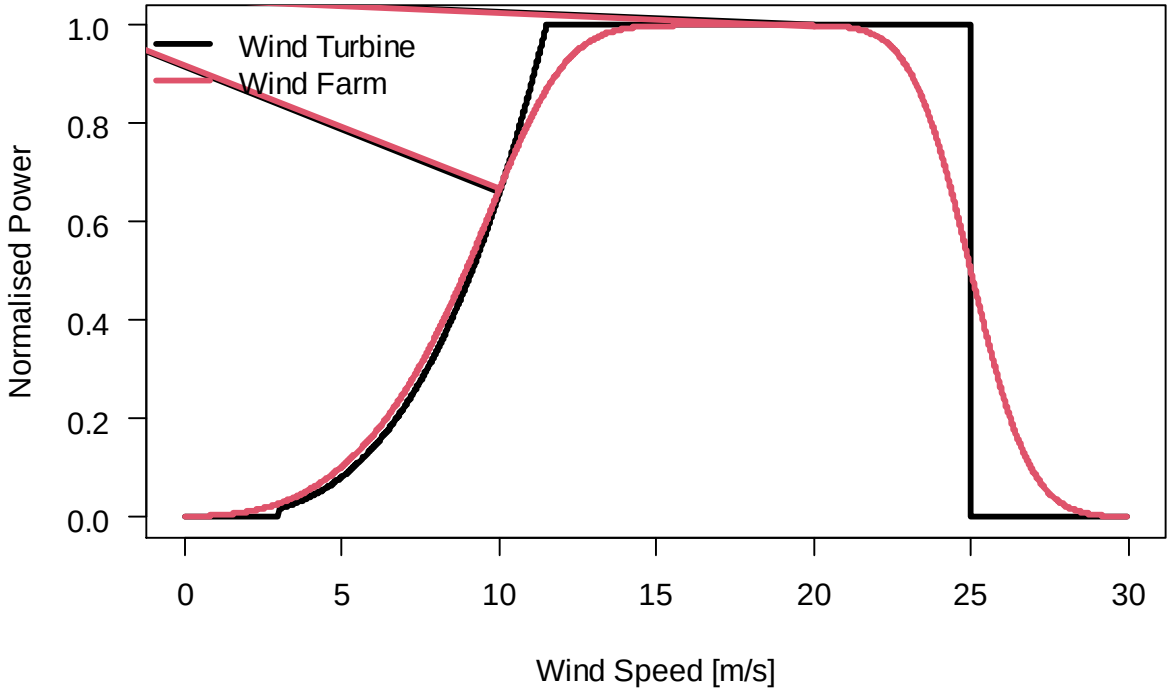

Forecast issued on: 2021-11-25
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

Forecast issued on: 2021-11-26
1 day before peak wind speeds, extreme wind is predicted
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time


Forecast issued on: 2021-11-27
Wind Gust [mph]
80
60
40
20
0
Nov 15
Nov 22
Nov 29
Date/Time

Principles: Physical and fundamental models

Sometimes the best, or only, way to produce a forecast is to model the underlying physical or economic relationships…

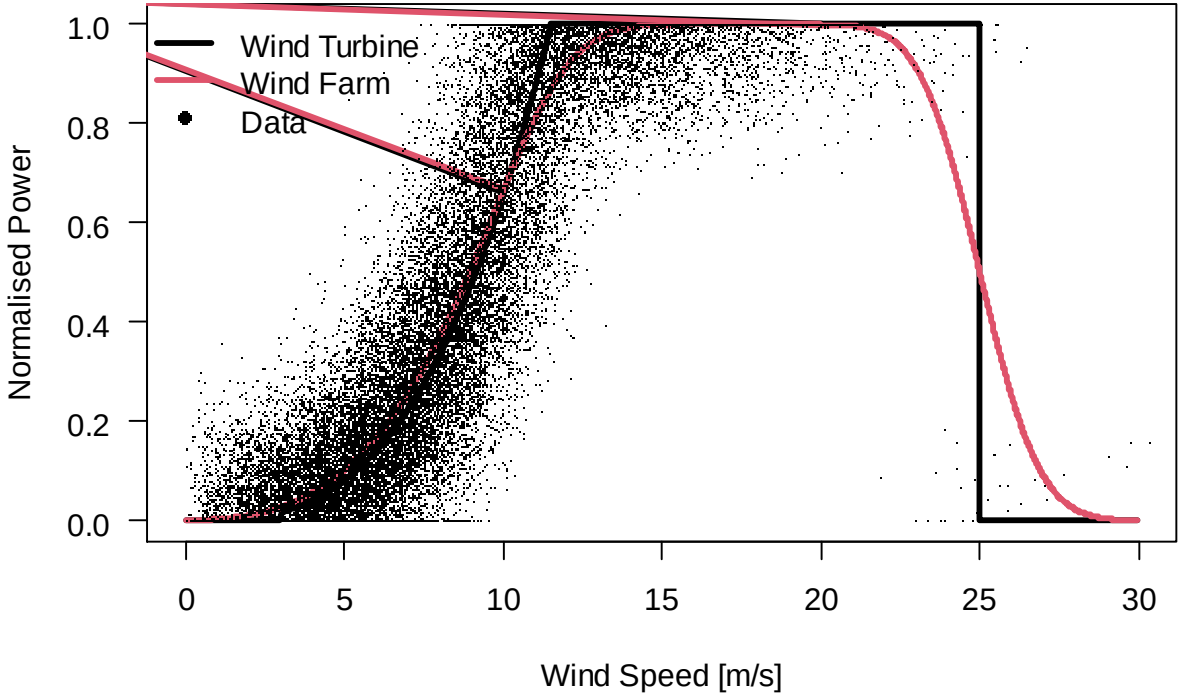


Consider a new wind farm:

- The *power curve* for individual turbines is known
- The smoothing effect of combining these in a farm may be estimated (imperfectly)
- We can convert wind speed forecasts into wind power forecasts!

Principles: Physical and fundamental models

Sometimes the best, or only, way to produce a forecast is to model the underlying physical or economic relationships…



Warning!

There are many complex processes at play in energy conversion (/consumption/pricing) and building them all into a **physical** or **fundamental** model may not be possible.

Data driven methods may be preferred, once there is sufficient data to train them…

Principles: Physical and fundamental models

Sometimes the best way to produce a forecast is to model the underlying physical or economic relationships…

Pros	Cons

• Interpretable
• Effective (sometimes necessary) in new situations, i.e. those not in historical data, such as
  • New market arrangements
  • New power plant
  • New technology
  • Unusual fuel costs
  • Extreme events in general
	• Can be less accurate then statistical/ML methods, especially weather-to-power relationships
	• Can be complex:
	  • Difficult to build and maintain
	  • Uncertainty quantification can be challenging
	  • High computational burden

Principles: Data-driven forecasting

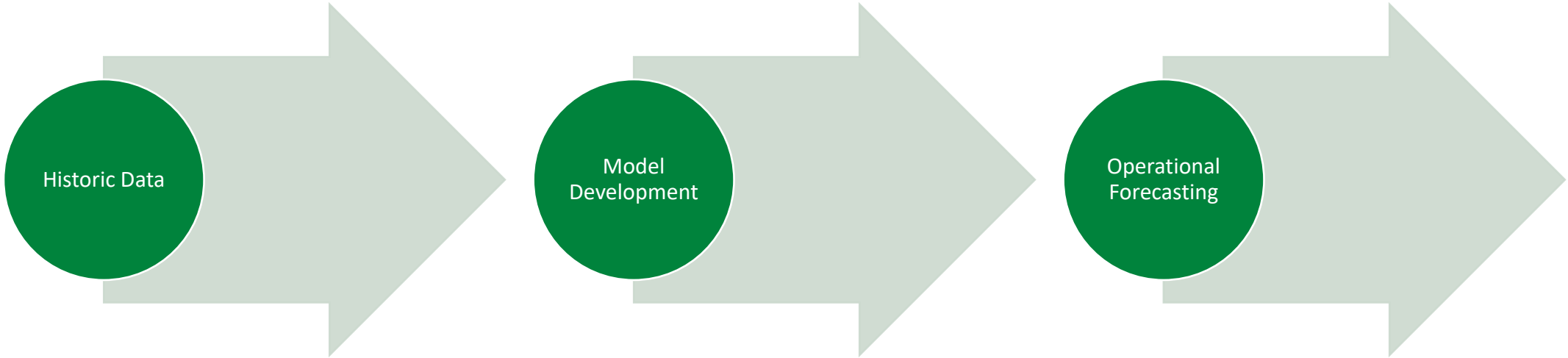
Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Brief introduction, more in parts 2 & 3 of this tutorial…

Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Big idea: Learn the relationship between data available when making a forecast and future values of a quantity of interest from historic data.



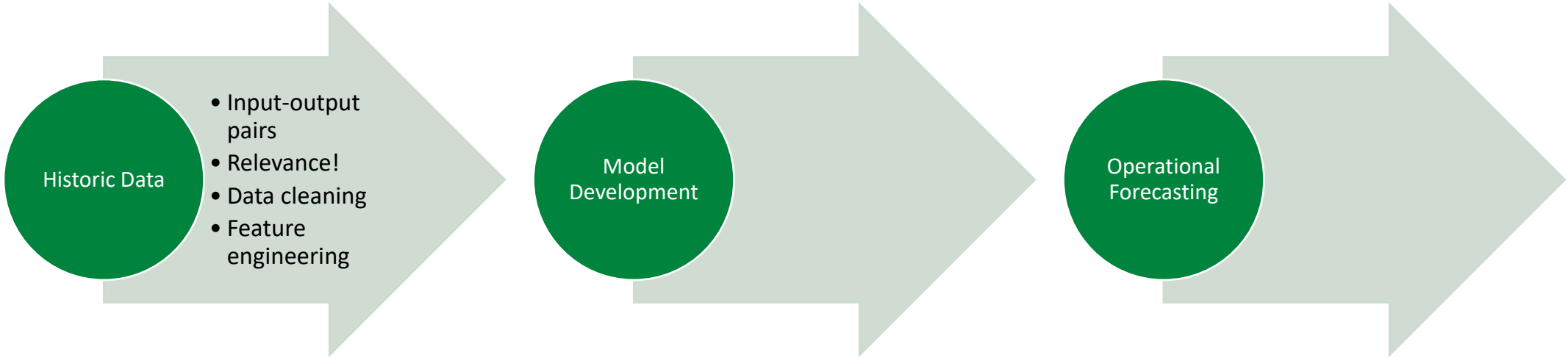
Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Assumptions:
- The relationship between inputs and outputs in the same now as it was in historic data
- A model capable of learning this relationship can be found

Requirements:
- Sufficient *volume* of *relevant* data is available

Other considerations:
- Data quality/cleaning
- Feature engineering



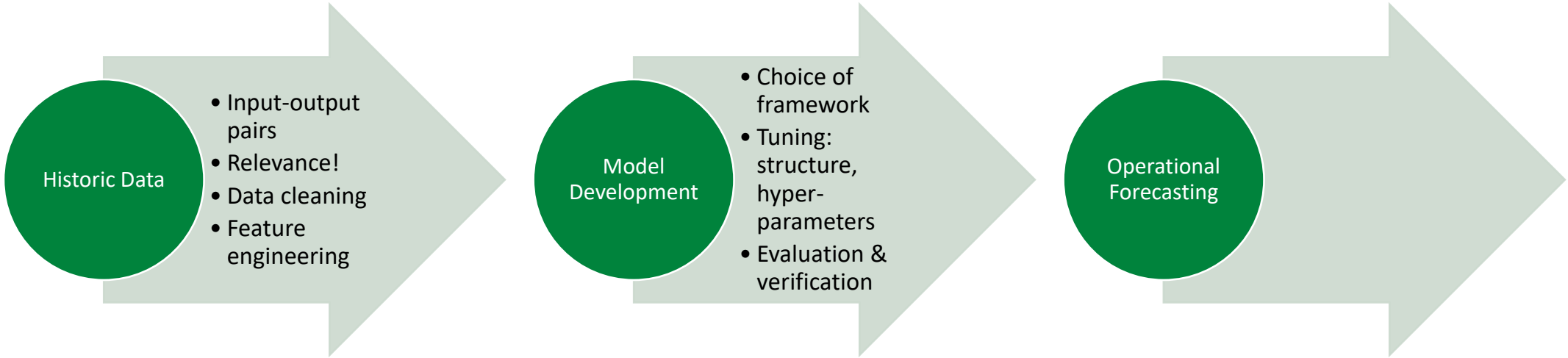
Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Vast array of models/frameworks to choose from!
Any method implemented well will likely outperform a fancy method implemented poorly.

Expert judgement and *hyper-parameters* tuning play a large role in performance
(including featuring engineering!)

Model/forecast evaluation should match end-use
Produce forecast for maximum value!



Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Operational Requirements

- Input data must be available in time to produce the forecast!
- Robustness to missing data
- Continuous/periodic updating of model
- Maintenance: data and underlying processes may change

Historic Data

- Input-output pairs
- Relevance!
- Data cleaning
- Feature engineering

Model Development

- Choice of framework
- Tuning: structure, hyper-parameters
- Evaluation & verification

Operational Forecasting

- Live operation
- Robustness
- Up-dating / online learning
- Maintenance

Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

Pros	Cons
• Can be highly accurate • Huge range of powerful methods available • Several frameworks for uncertainty quantification	• Accuracy depends on data quality, volume and relevance • Poor (and unknown) performance in “unseen” situations • Interpretability

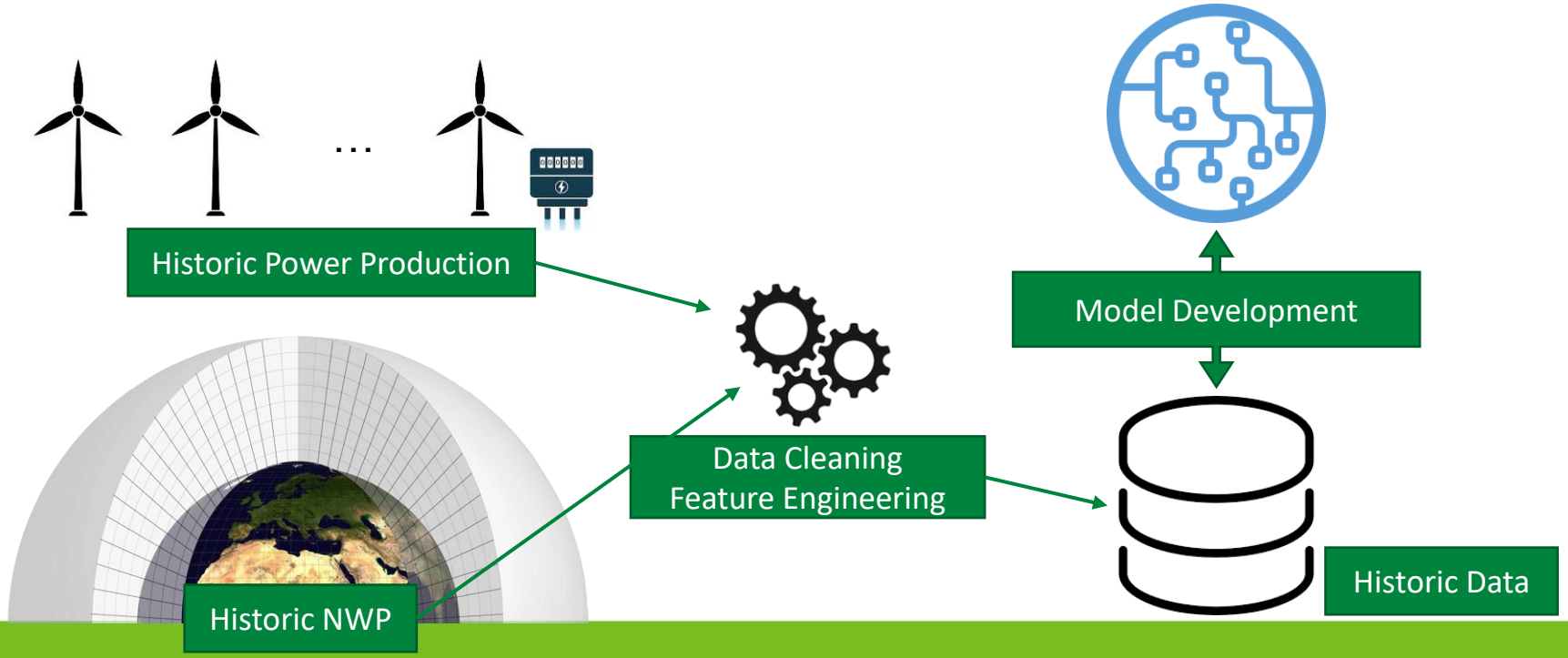
Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

In practice, many best-in-class forecasting systems combine physical, fundamental, and driven method.

Lets take wind power as an example again:

Training Phase



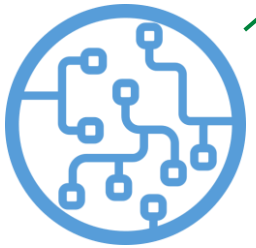
Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

In practice, many best-in-class forecasting systems combine physical, fundamental, and driven method.

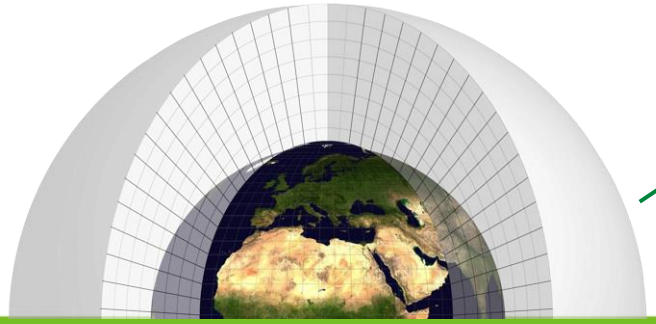
Lets take wind power as an example again:

Power Forecast



Operational Phase

Latest NWP

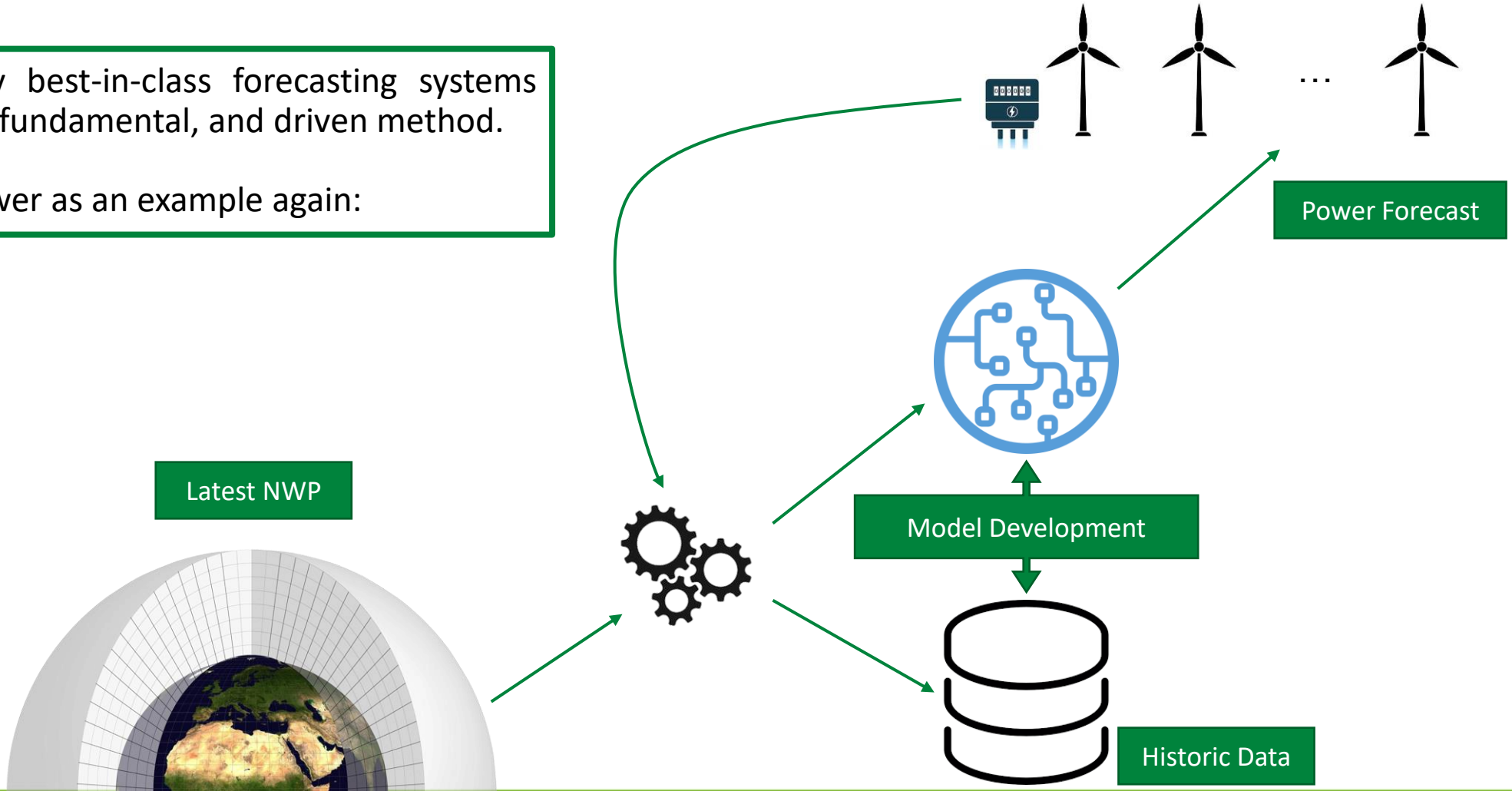


Principles: Data-driven forecasting

Statistics… Machine Learning… Artificial Intelligence… call it what you like!

In practice, many best-in-class forecasting systems combine physical, fundamental, and driven method.

Lets take wind power as an example again:



Operational Phase with Update

Forecast quality and evaluation

What is a *'good'* forecast?

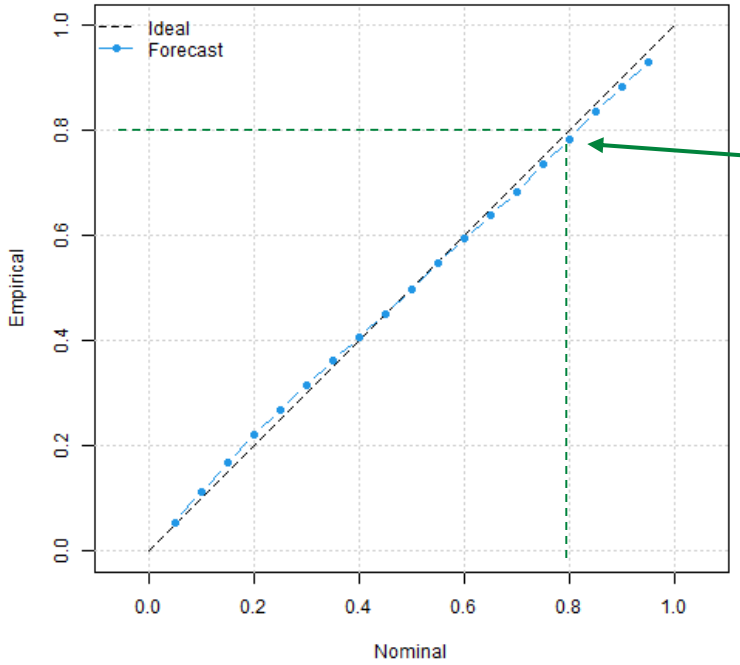
What makes a *'good'* forecast?

Desirable properties:

1. Calibration/reliability
2. Sharpness
3. Resolution and Discrimination
4. **Better decisions!**

Calibration/reliability

Reliability Diagrams

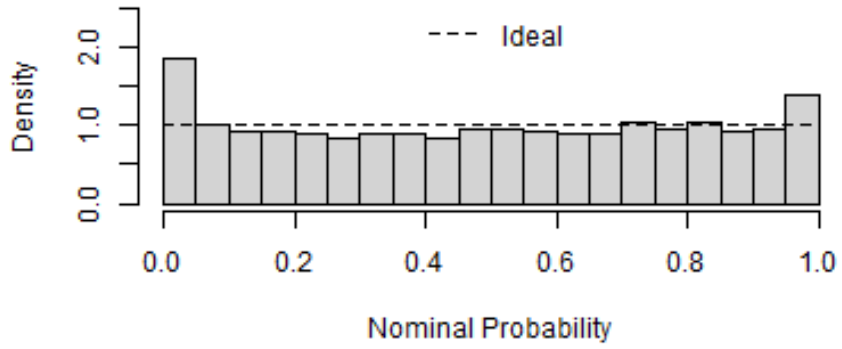


- For univariate probabilistic forecasts
- Frequency of observed events should match their nominal probability
  - I.e., 80% of observations should fall below the 80% quantile of a density forecast
  - With finite data, some deviation from “perfect” is to be expected, can be assessed with confidence intervals

- Recall: calibration/reliability is a necessary condition so should always be verified.

Calibration/reliability

PIT Histograms



- For univariate probabilistic forecasts
- Probability Integral Transformation (PIT): inverse predictive CDF evaluate at corresponding observation.
- PIT transformed data should be uniformly distributed
- Similar information as reliability diagram, but presented slightly differently
  - Example from the same data as previous slide
  - Highlights that the this forecasting system may not be calibrated at high and low probability levels

- (Multivariate extension: “rank histograms” for ranking vectors based on some pre-rank function)

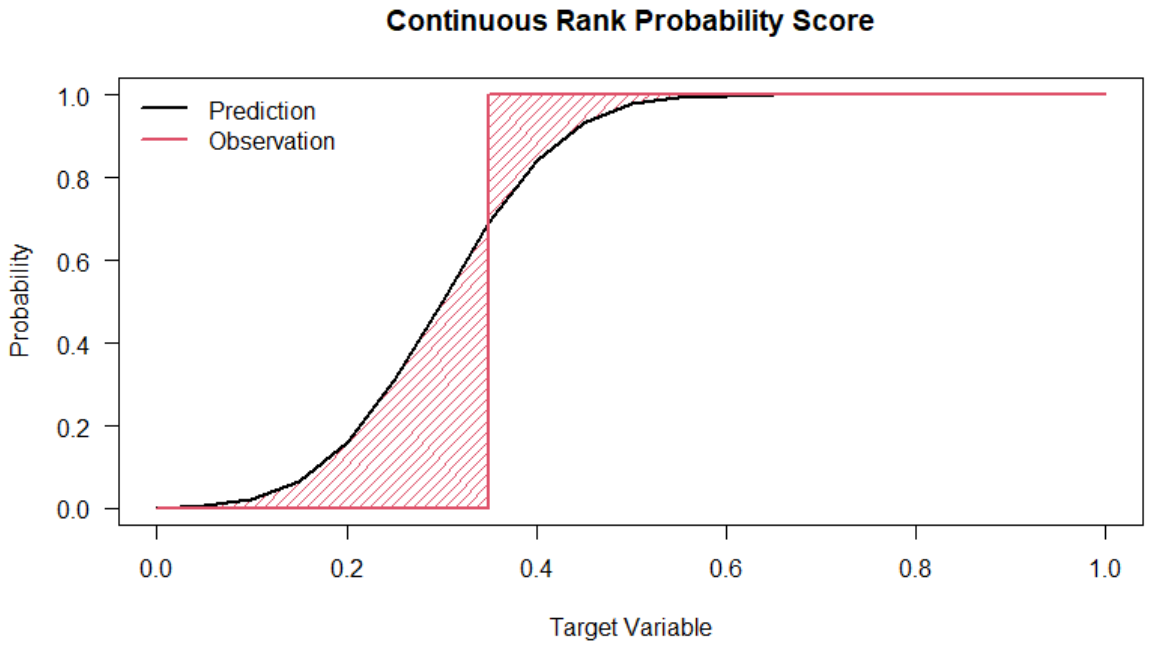
Calibration/reliability

Multi-variate forecasts

- Examples:
  - Temporal trajectories (seen earlier)
  - Spatio-temporal trajectories (multiple locations)
  - Multi-variate spatio-temporal trajectories (multiple variables and locations)
- Verifying calibration can be very difficult!
  - Verify margins (individual locations/lead-times/variables) as if univariate
  - Verify dependency structures (rank histograms, covariance)
  - Verify calibration of relevant “event” forecasts, e.g., ramps

Scores, metrics and skill

Pinball and CRPS



- Continuous Ranked Probability Score evaluates full distribution
- Discrete form, Pinball Score, evaluates individual quantile forecasts
- Reward sharpness as calibration (but calibration should always be verified separately)

$$CRPS = \frac{1}{N}\sum_{t=1}^{N}\left(\underbrace{\hat{F}(x)}_{\text{Density Forecast (CDF)}} - \underbrace{\mathbb{1}(x \leq y_t)}_{\text{Observation}}\right)^2 dx$$

Scores, metrics and skill

Other scores

Univariate

- CRPS/Pinball
- Log Score (Ignorance Score, and extensions Conditional/Censored Likelihood Score)

Multi-variate

- Log Score
- Energy Score (generalization of CRPS)
- Variogram Score
- Dawid-Sebastiani Score

When comparing forecast performance, it may be necessary to test whether apparent differences are "statistically significant". Several methods are at our disposal but are beyond the scope of this course. See *Further reading*.

Scores, metrics and skill

Skill scores

We often want to compare the performance of two forecasting systems, one with metric M and a reference system M_{ref}. Skill scores can have greater discrimination than raw scores alone:

$$Skill = \frac{M_{\mathrm{ref}} - M}{M_{\mathrm{ref}} - M_{\mathrm{perf}}}$$

M_{perf} is the “perfect” score for the metric in question.

Scores, metrics and skill

Economic evaluation

- Forecasts only add value when they lead to better decisions
- Calculating the value ($) of forecast improvement can be difficult or impossible! (Various simulation-based approach have been used, most commonly “shadow trading”)
- The utility of forecasts, e.g. situational awareness or risk reduction, may be more difficult still to quantify

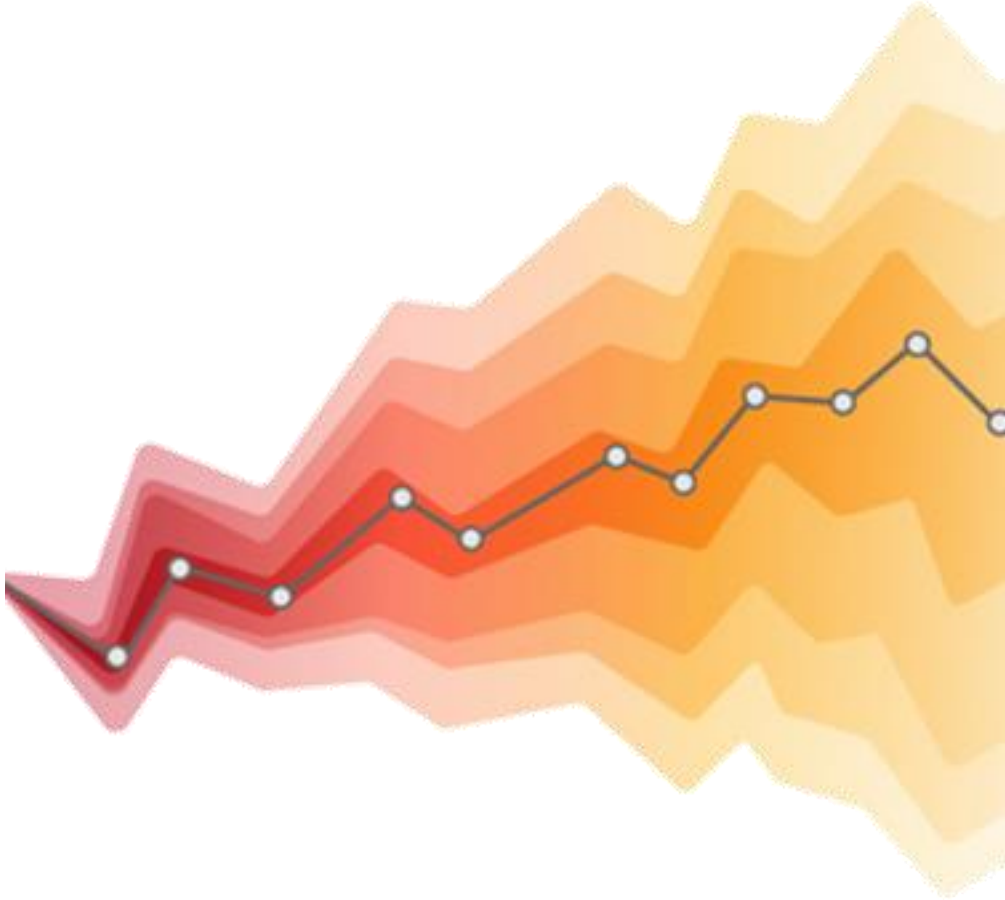
The IEA Recommended Practice for Forecast Selecting Renewable Power Forecasting Solutions suggests a multi-criteria approach:

- Visual Inspection
- Use of multiple metrics tailored to use-case
- Use of histogram or boxplot for evaluation of outliers
- Use of contingency tables for specific event analysis
- Use of skill scores relative to a relevant reference forecast for comparisons

Part 1 Re-cap

Introduction to probabilistic energy forecasting

- Motivation
- Types of probabilistic forecast
- Principles of energy forecasting
  - Timescales
  - Physical and fundamental methods
  - Data-driven methods
- Forecast quality and evaluation
- Decision-making game



Let’s play!

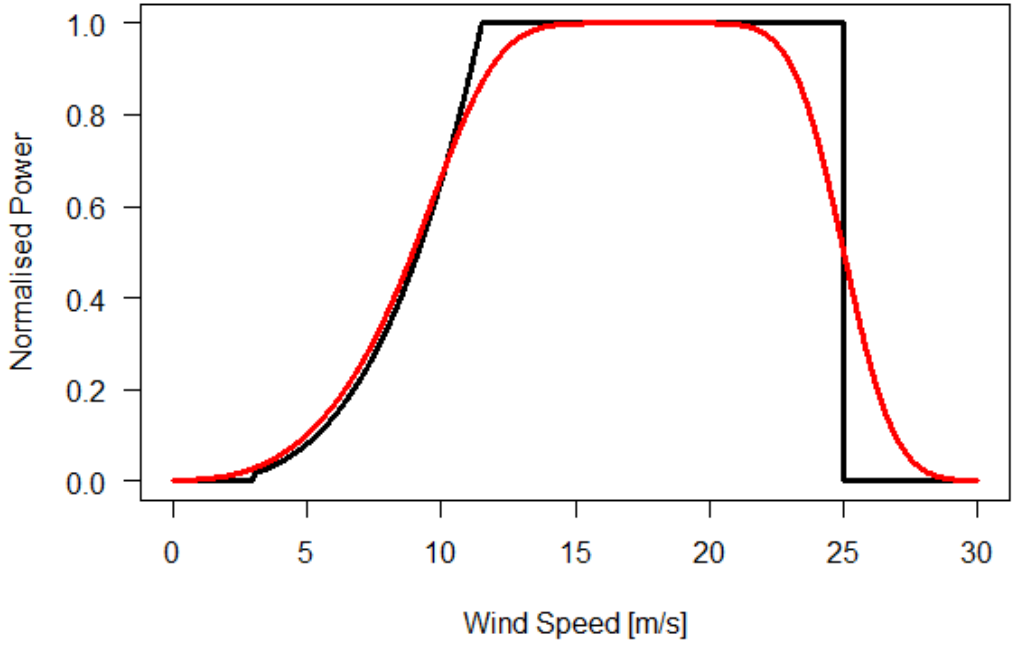
Decision-making under forecast uncertainty

Let’s play!

High-speed shutdown game

High-speed shutdown (HSSD)
A high-speed shutdown events typically occur when wind speed exceeds 21-27m/s. Individual turbines cut-out above 25m/s but use both wind gusts and the mean wind to determine HSSD, and wind farms contain many turbines...

The Game
Decide in whether to trade 50% or 100% of the generation from of an offshore wind farm according to the forecast given the possibility of a high-speed shutdown, where the wind park stops generating due to excessive wind conditions.



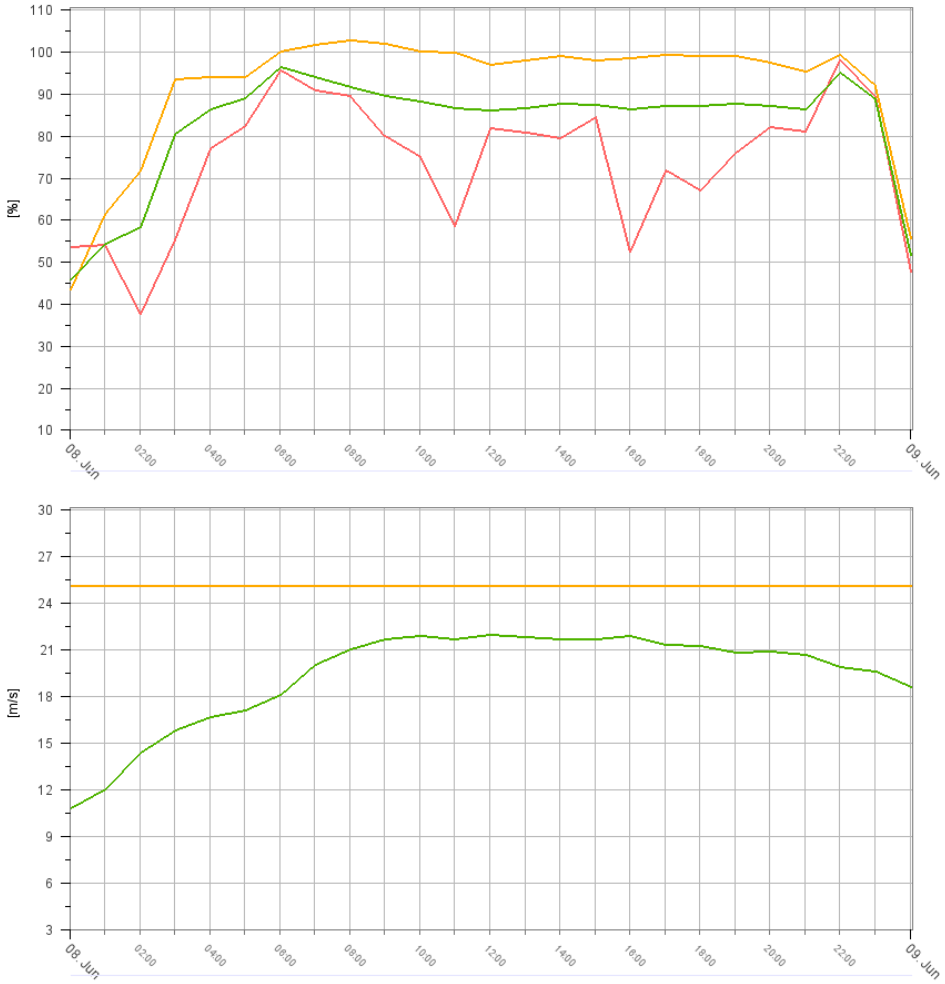
Let’s play!

High-speed shutdown game

First type of forecast

- 3 independent deterministic wind power forecasts in the unit [% of installed capacity] based on 3 different NWP (numerical weather prediction) models ↗

- 1 wind speed forecast in the unit [m/s] and a reference line for the 25m/s threshold reference value for high-speed shutdown. ↘



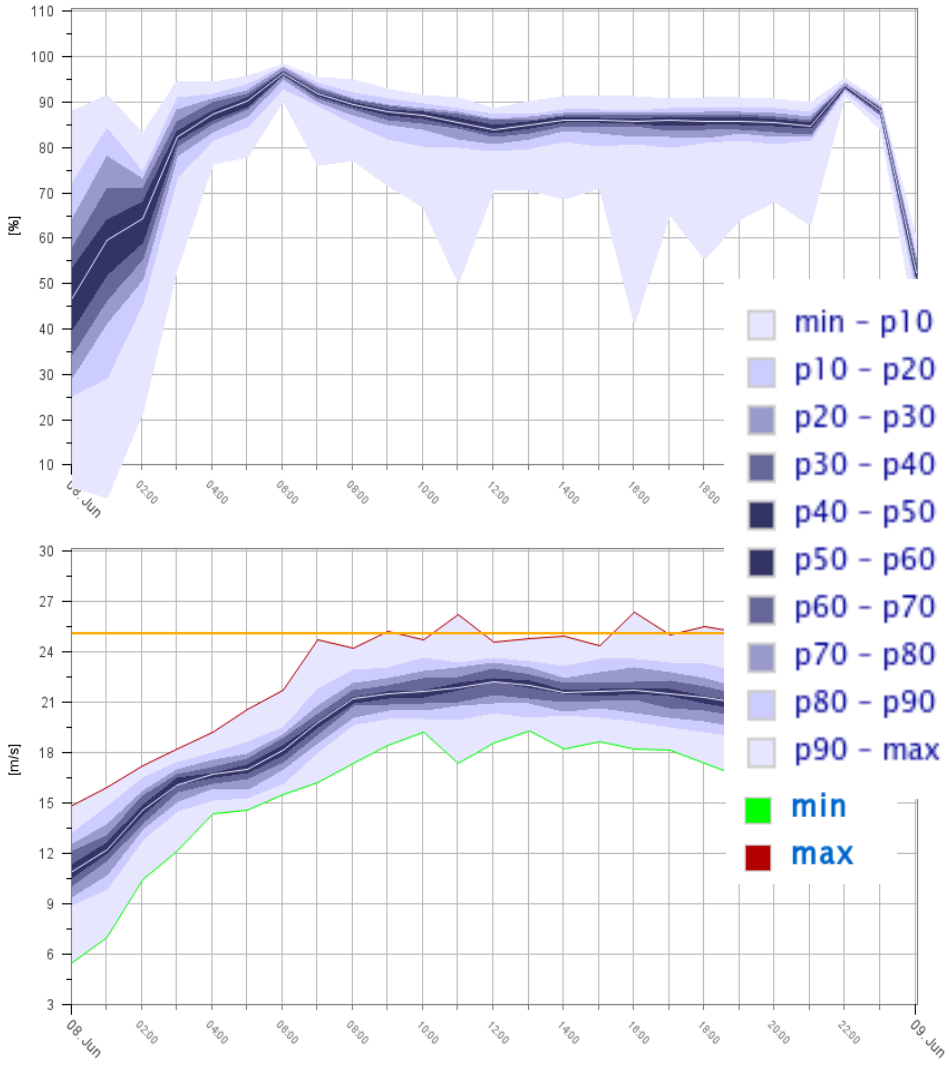
Let’s play!

High-speed shutdown game

Second type of forecast

- 9 wind power percentiles (P10..P90) and a mean (white line) in the unit [% of installed capacity] generated from 75 NWP forecasts of a multi-scheme ensemble prediction system ↗

- 9 wind speed percentiles (P10..P90) and a mean (white line) in m/s] generated from 75 NWP forecasts of a multi-scheme ensemble prediction system ↘



Let’s play!

High-speed shutdown game

To reflect the costs of large and small errors we have defined a simplified cost function for the period, where high-speed shutdown (HSSD) can take place.

Set-up:

- The wind farm capacity is 100MW and the spot market price is 50 EUR/MWh
- Balancing costs are equivalent to spot market prices
- The cost function will only consider your choice for the hours where the actual generation is full load or no generation

		Outcome	
	Trade	HSSD	No HSSD
Decision:	100%	-5,000	5,000
	50%	0	2,500

Note that trading 100% is a risky choice that can both increase your income and loss.
The more conservative 50% trading strategy eliminates the risk of a loss, because balance costs are equal to spot market prices, and you can curtail the wind farm to avoid balance costs.

Let’s play!

High-speed shutdown game

https://meteorology.mpib.dev/wind-power-decisions
Click: “Play the Game”

Game provided by:
Nadine Fleischhut
Max-Planck Institute for Human Development
fleischhut@mpib-berlin.mpg.de

Corinna Möhrlen
WEPROG
com@weprog.com

Further reading

Weather Matters for Energy, Eds. A. Troccoli, L. Dubus, S.E. Haupt, DOI: 10.1007/978-1-4614-9221-4

Weather and Climate Services for the Energy Industry, Eds. A. Troccoli DOI: 10.1007/978-3-319-68418-5

Messner *et al*, Evaluation of wind power forecasts – An up-to-date view, DOI: 10.1002/we.2497

Pinson *et al,* Trading Wind Generation From Short-Term Probabilistic Forecasts of Wind Power, DOI: 10.1109/TPWRS.2007.901117