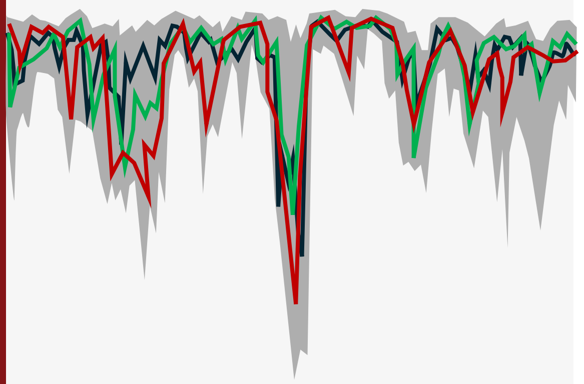
‘Forecast busts’ over Europe in ERA5 reforecasts: Characteristics and predictions using neural networks



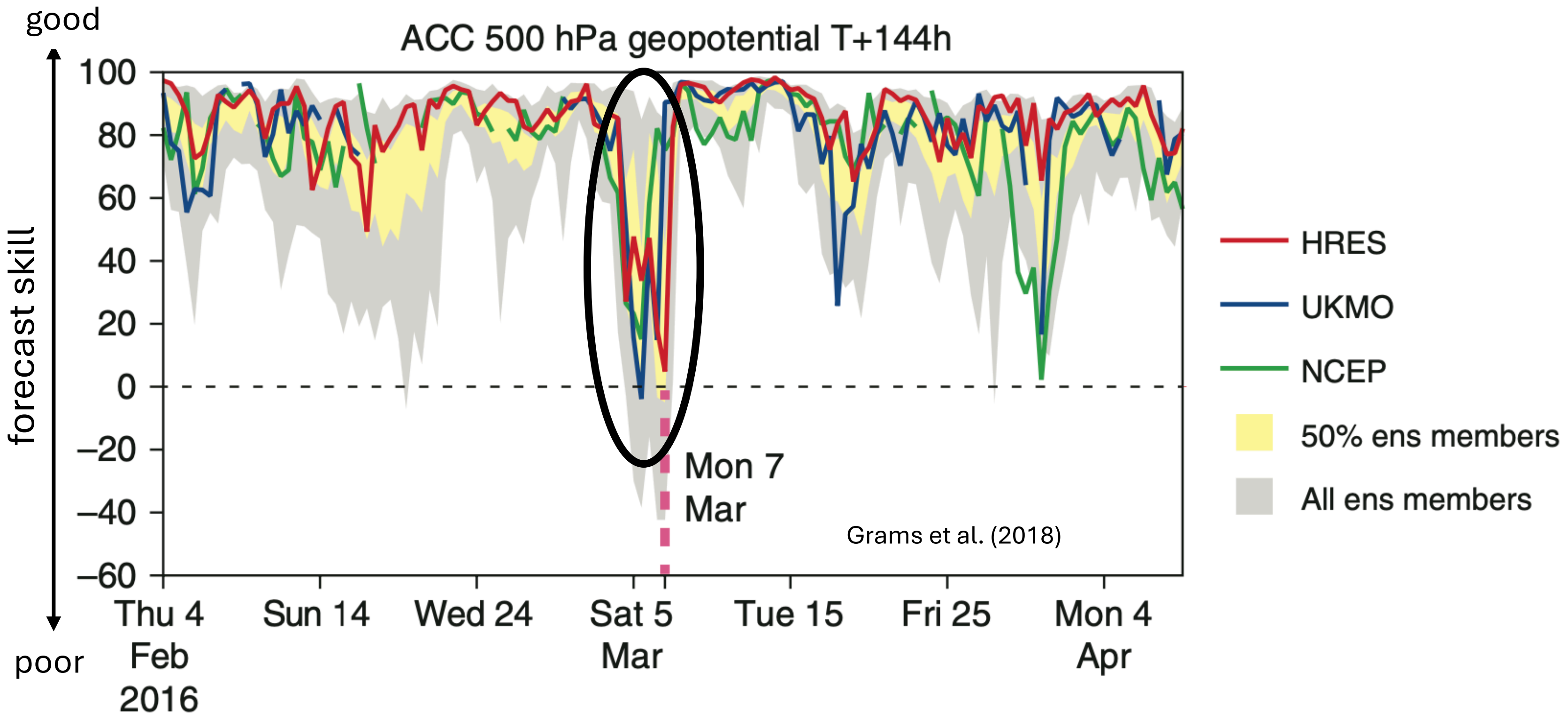
Seraphine Hauser[1,2], David Parsons[1], Steven Cavallo[1], Jonathan Martin[2]

[1] School of Meteorology, University of Oklahoma, Norman, Oklahoma, USA
[2] Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, Wisconsin, USA

seraphine.hauser-1@ou.edu

Workshop on Diagnostics for Global Weather Prediction
(ECMWF, Reading, 9-12 September 2024)

What is a "forecast bust"?



"Forecast busts"

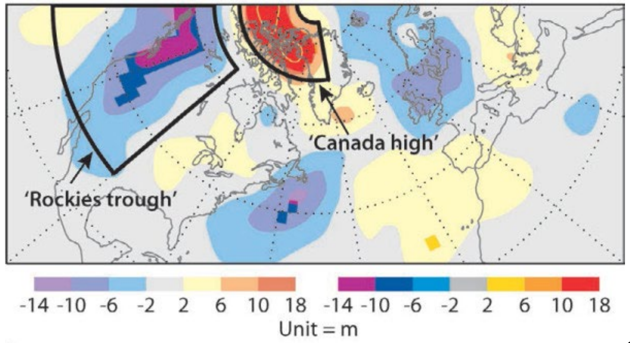
When forecast skill suddenly drops producing a very poor forecast within short lead times
(Rodwell et al., 2013)

Research on ”forecast busts” over Europe

The UNIVERSITY of OKLAHOMA

Definition of forecast busts

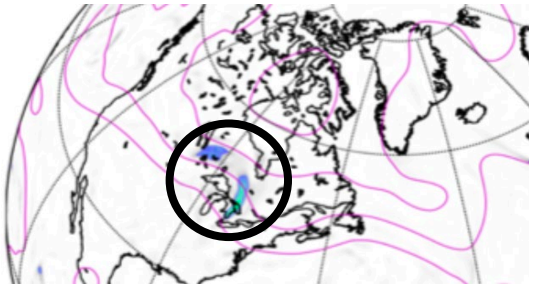
First (small) systematic analysis based on ERA-Interim
(Rodwell et al., 2013)



Link of initial errors to strong diabatic heating

Mesoscale convective systems (Parsons et al., 2019)

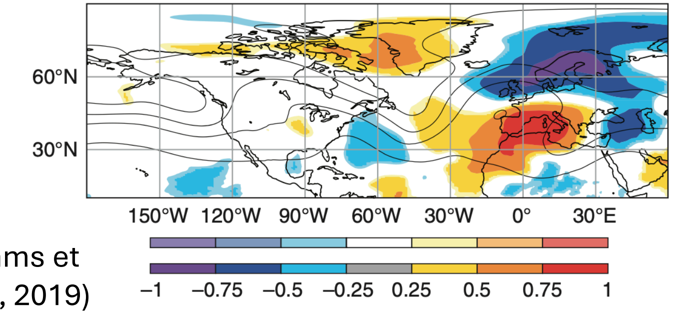
or sub synoptic-scale moist processes linked to warm conveyor belts
(Grams et al., 2018, Pickl et al., 2023)



Initial errors (Parsons et al., 2019)

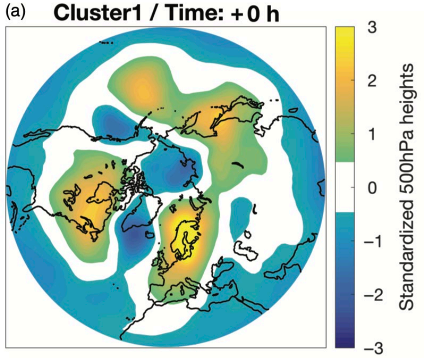
Diagnostics for error development understanding

Error-tracking methods applied to forecast bust cases
(Magnusson, 2017; Grams et al., 2018: Parsons et al., 2019)



Patterns and dynamics of forecast busts

Different flavors of pattern development over the North Atlantic suggesting regime transitions
(Lillo and Parsons, 2017)

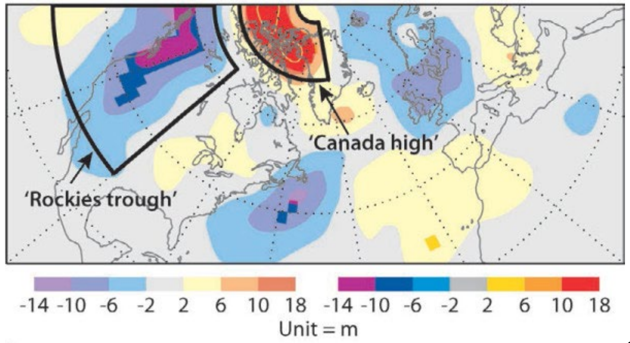


Research on ”forecast busts” over Europe

Open questions addressed

Definition of forecast busts

First (small) systematic analysis based on ERA-Interim
(Rodwell et al., 2013)

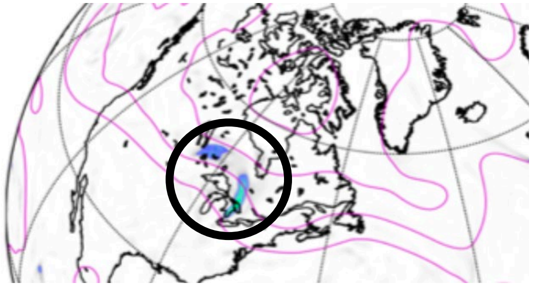


Changes to ERA5? Extension for exceptionally good forecasts?

Link of initial errors to strong diabatic heating

Mesoscale convective systems (Parsons et al., 2019)

or sub synoptic-scale moist processes linked to warm conveyor belts
(Grams et al., 2018, Pickl et al., 2023)

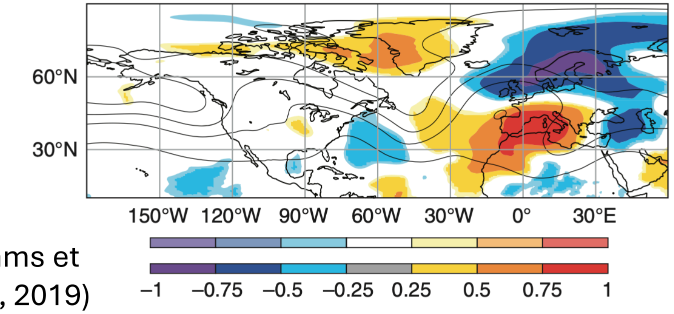


Initial errors (Parsons et al., 2019)

Signs of MCS over US for poor forecasts? How about for good forecasts?

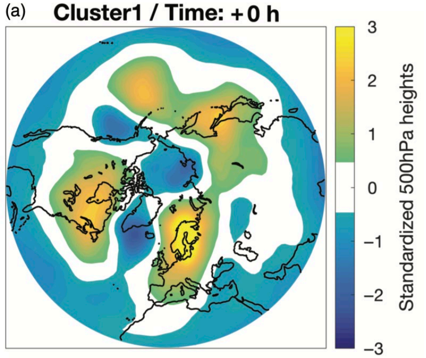
Diagnostics for error development understanding

Error-tracking methods applied to forecast bust cases
(Magnusson, 2017; Grams et al., 2018: Parsons et al., 2019)



Patterns and dynamics of forecast busts

Different flavors of pattern development over the North Atlantic suggesting regime transitions
(Lillo and Parsons, 2017)

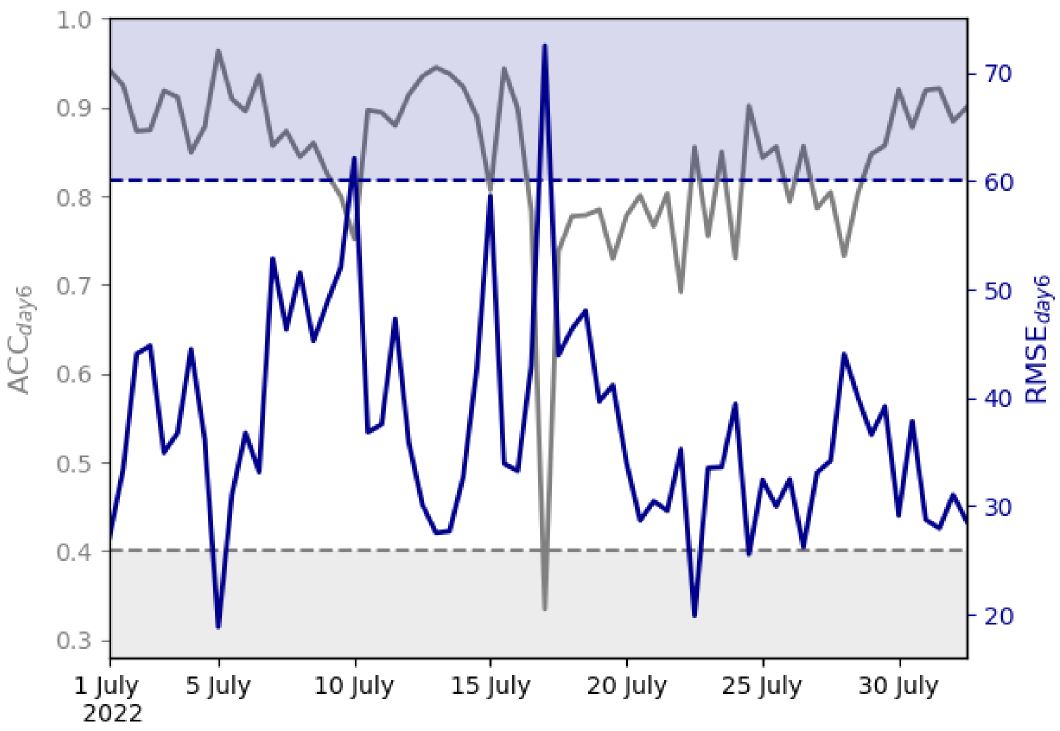


What happens between day 0 and day 6 in the “reanalysis world”?

Systematic analysis: 45 years of ERA5 reforecasts

Data set | ERA5 re-forecasts (1979-2023) Hersbach et al. (2020)

- 45 years of 10-day forecasts based on a constant model version (IFS Cy41r2)
- Northern Hemisphere, 1° spatial resolution, twice-daily initial times (00 and 12 UTC)
- ”True”: ERA5 reanalysis



Definition of forecast busts (day 6, Europe)

Rodwell et al. (2013)

1. Anomaly correlation coefficient (ACC)

   “forecast bust”: $ACC_{day6} < 0.4$ (phase error)

2. Root mean square error (RMSE)

   “forecast bust”: $RMSE_{day6} > 60m$ (amplitude error)

Part 1

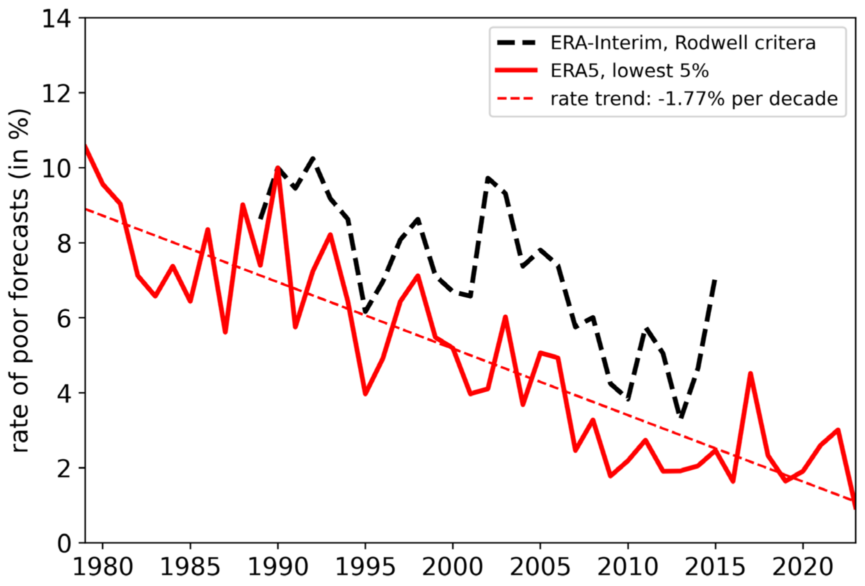
Characteristics of exceptionally poor and good forecast in ERA5 reforecasts

Results | Rate and seasonality of poor/good forecasts

Trend of exceptionally poor and good 6-day forecasts over Europe

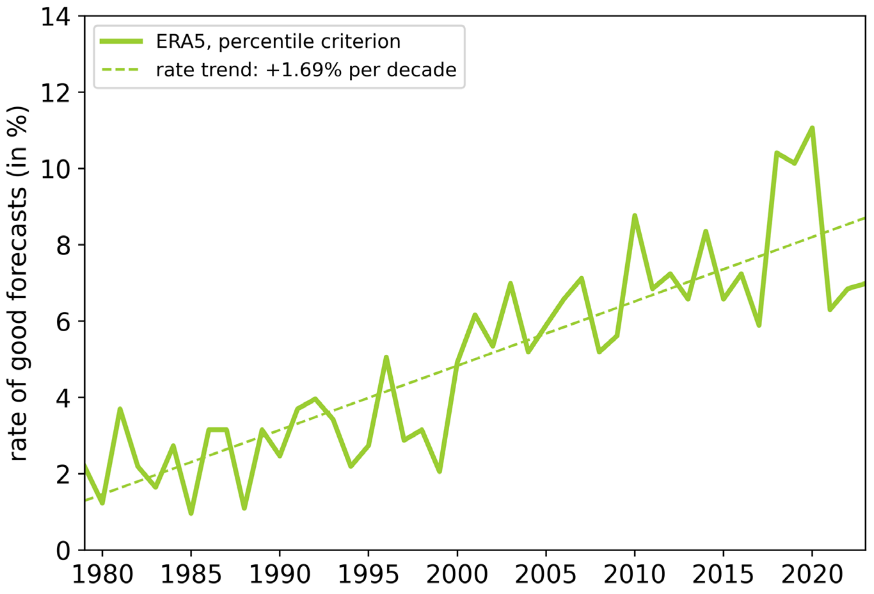
Definition based on ACC only and based on percentiles

Lowest 5% in ACC_{day6}



- Decrease in the number of “forecast busts” with time
  → Good agreement with ERA Interim definition

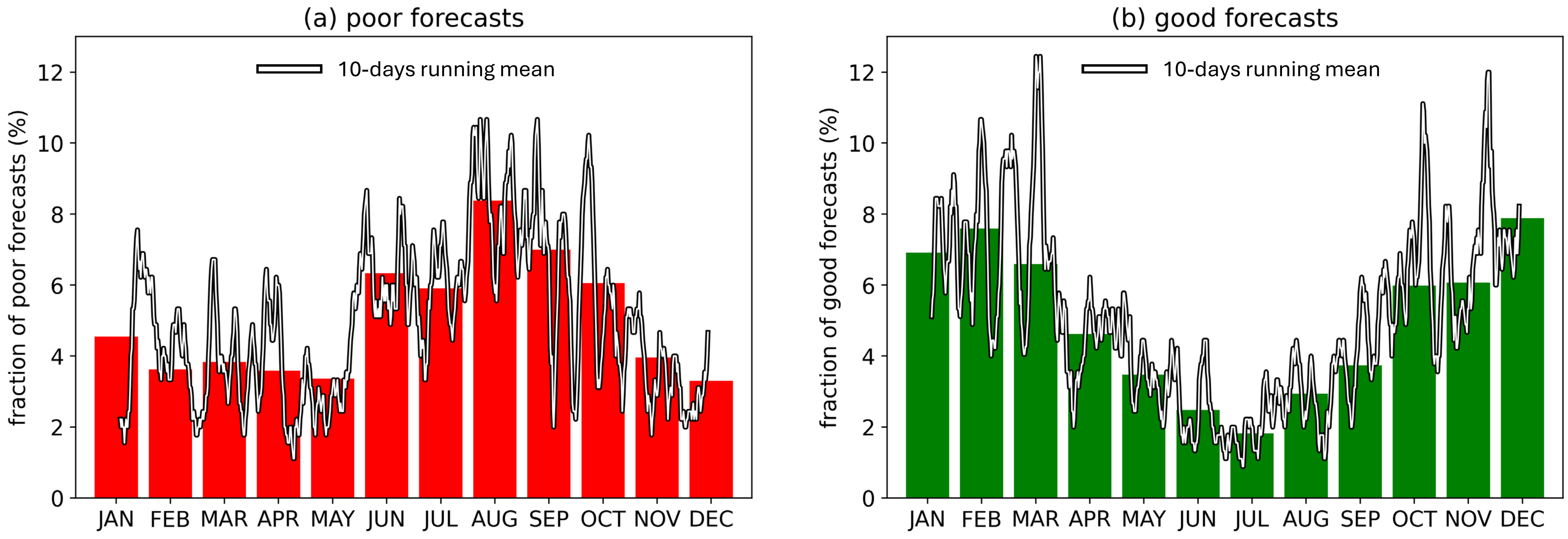
Largest 5% in ACC_{day6}



- Upward trend in exceptionally good forecasts
  → Number and quality of observations improved

Results | Rate and seasonality of poor/good forecasts

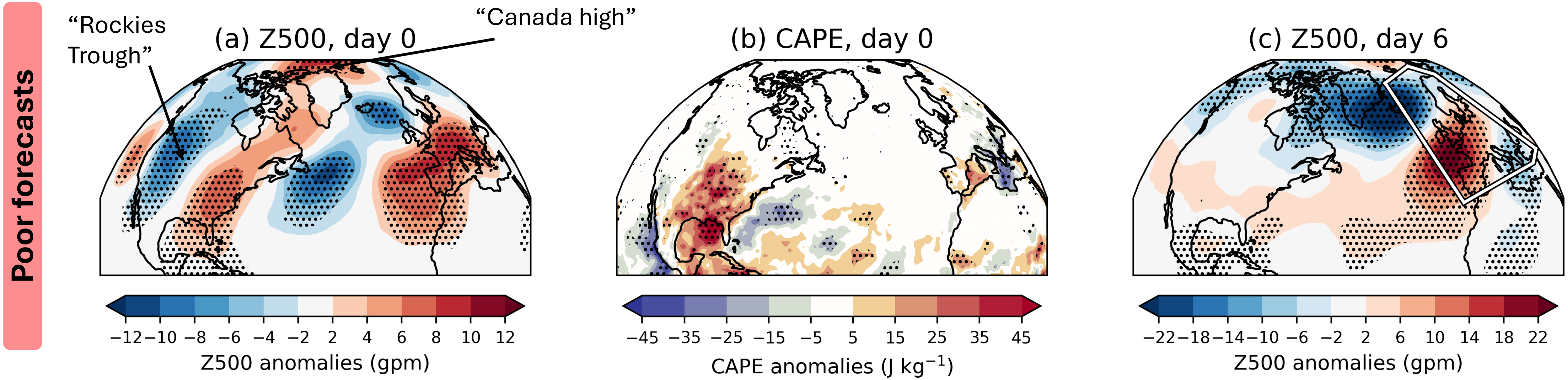
Distinct differences in seasonal occurrence



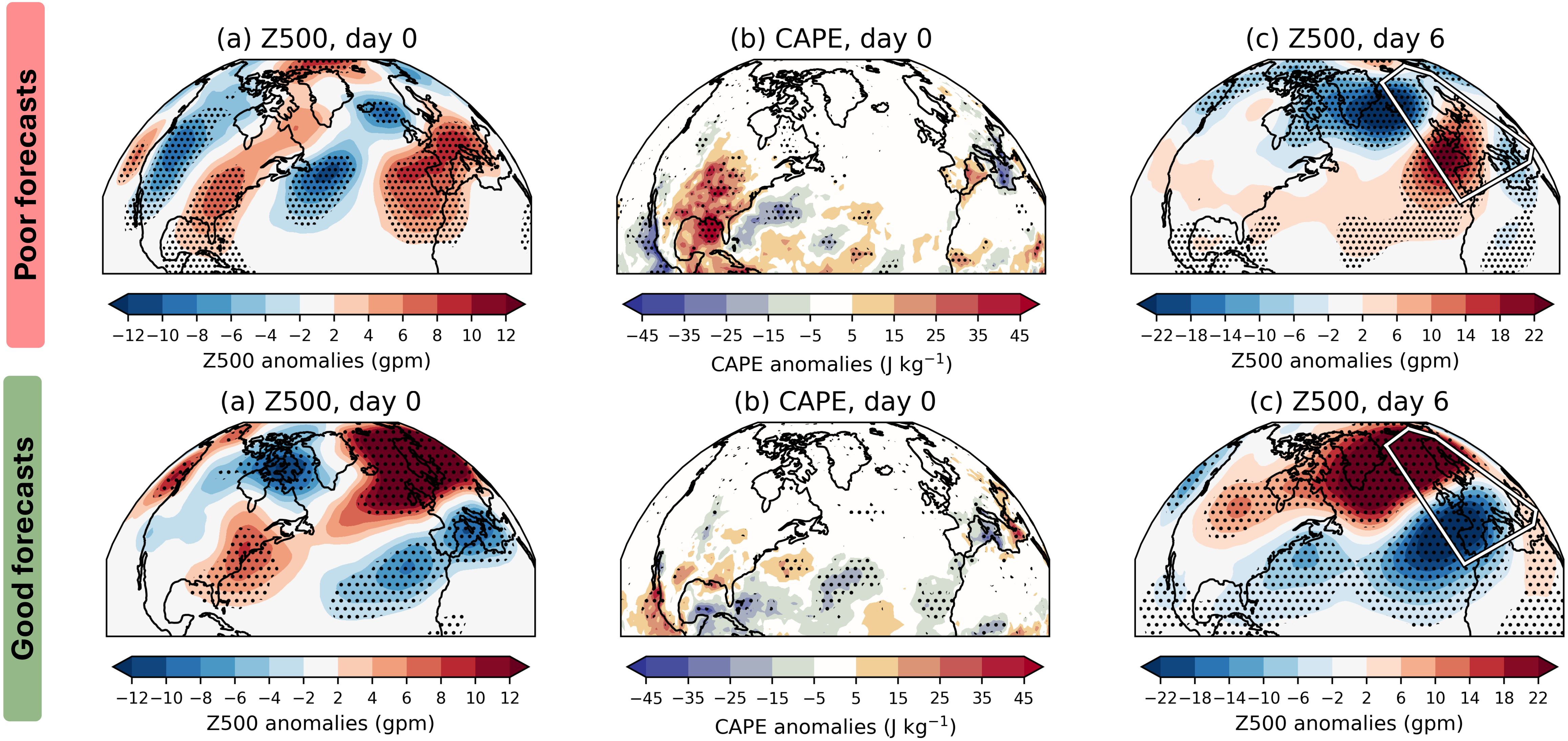
- Peak occurrence from June to October
- Link to summertime convection/hurricane season?

- Peak occurrence in the cold season, minimum in summer
- Increased wintertime skill related to low-frequency modes?

Results | Patterns of exceptionally poor vs. good forecasts



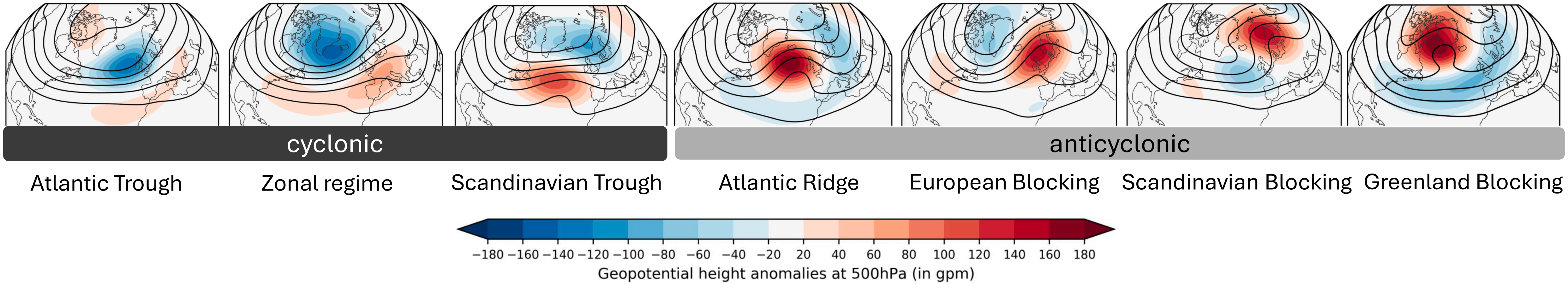
Results | Patterns of exceptionally poor vs. good forecasts



How can we capture what happens within those 6 days **and** investigate variability?

Results | Weather regime perspective

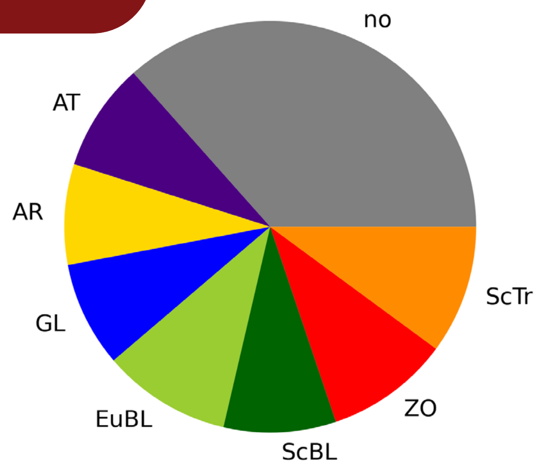
Year-round definition of North Atlantic-European weather regimes (Grams et al., 2017)

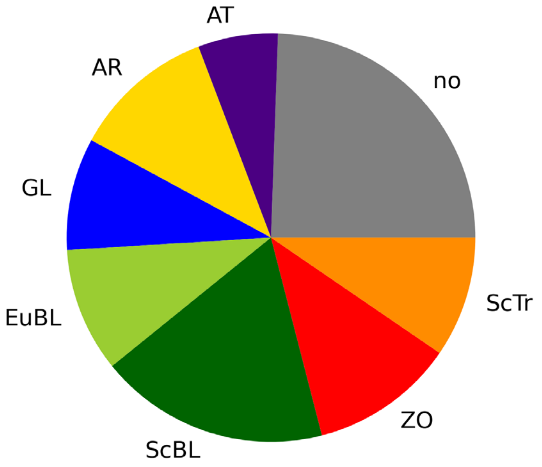


Which regime is active at day 0?

Poor forecasts

Large-scale flow is often assigned to none of the seven regimes





Good forecasts

Mostly regimes are active, dominance of **Scandinavian Blocking**

Results | Regime development day 0 – day 6

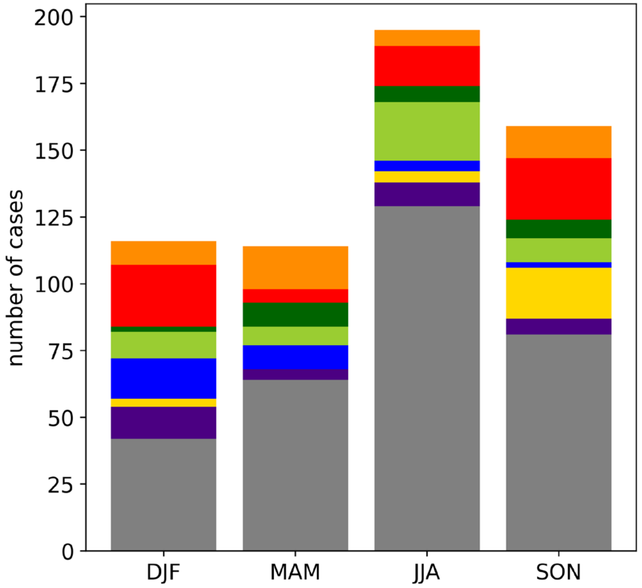
day 0 day 6

Case 1

$\text{regime}_{day0} = \text{regime}_{day6}$

Poor forecasts

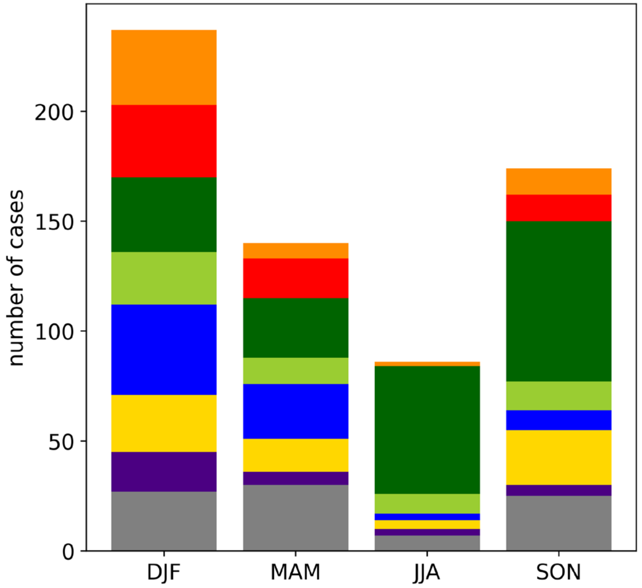
36%



- Poor forecasts often in periods of no regime
- Good forecast predominantly predict the continuity of a pattern

Good forecasts

39%



ScTr ZO ScBL EuBL GL AR AT no

Results | Regime development day 0 – day 6

day 0

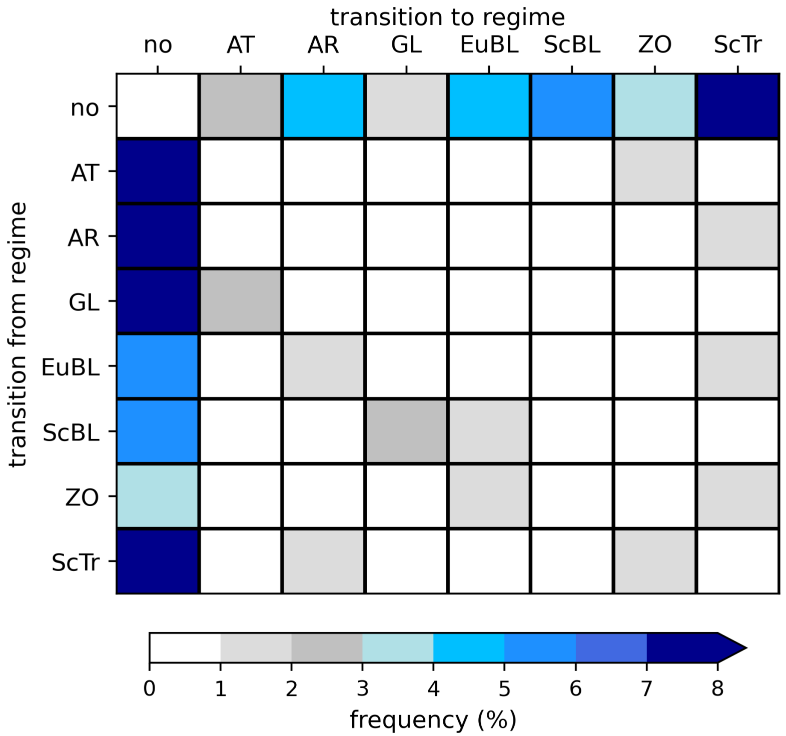
day 6

Case 2

$\text{regime}_{\text{day0}} \neq \text{regime}_{\text{day6}}$

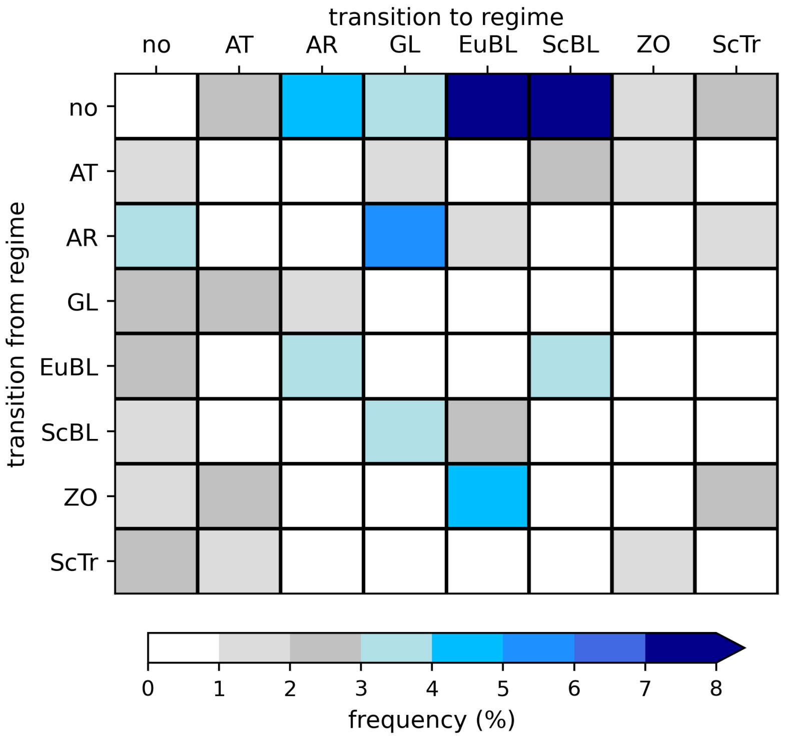
Poor forecasts

64%



Good forecasts

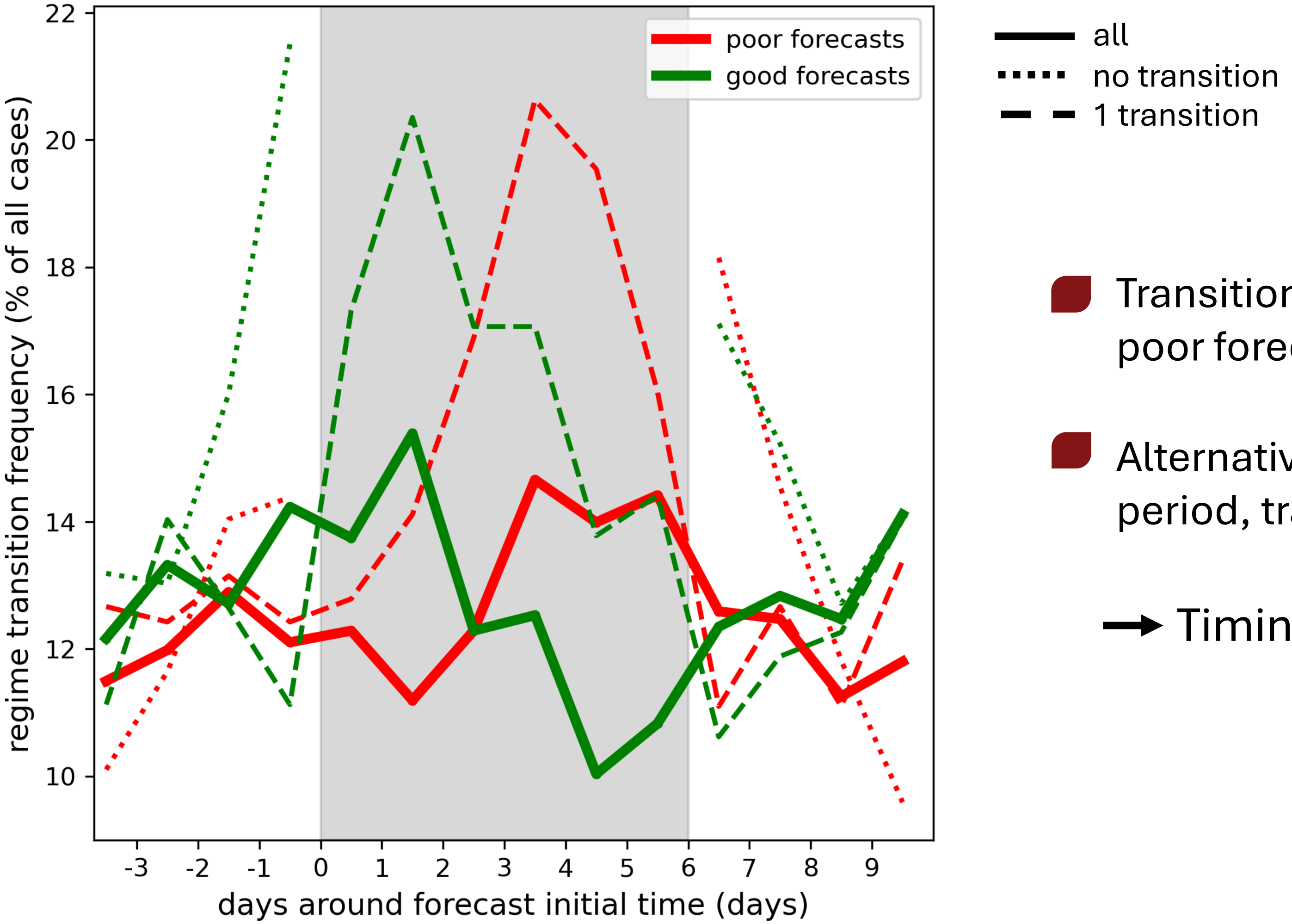
61%



- Transitions from or into the ‘no regime’ of poor forecasts
- Good forecasts perform well with onsets of blocked regimes or transitions between blocked regimes

Results | The timing of the transition matters!

When do regime transitions occur?



- Transitions occur earlier in good forecasts and later in poor forecasts
- Alternatively, if no transition occurs within the forecast period, transitions occur immediately beforehand

→ Timing matters!

Summary | Part 1

What did we do?

- Redefined the definition of large-scale forecast busts and extended it to exceptionally good forecasts in a 45-years period of ERA5 re-forecasts

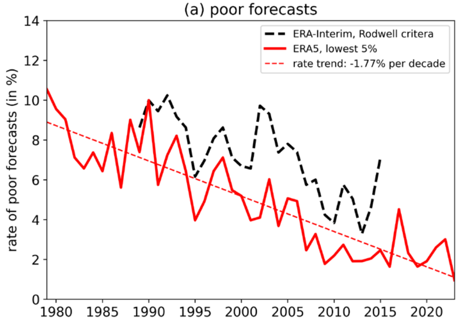
Our four key findings

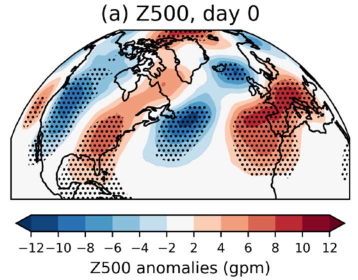
1. Decreasing (increasing) number of poor (good) forecasts with time despite no changes in model physics probably due to improved observations

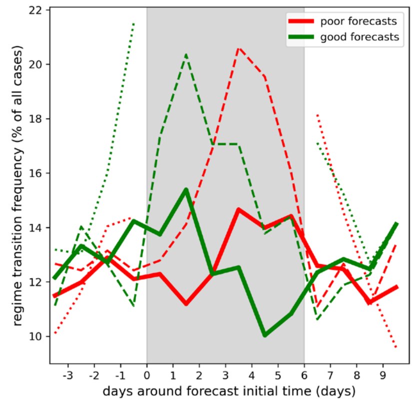
2. High agreement in the large-scale patterns at initial time of forecast busts between ERA-Interim (Rodwell et al., 2013) and ERA5 busts, indicating struggles independent of the model version

3. Very clear differences in the seasonality and patterns (good vs. poor forecasts)

4. Similar characteristics in the number of regime transitions during the 6-day period but some differences in certain transitions and in particular the timing of transitions







Part 2

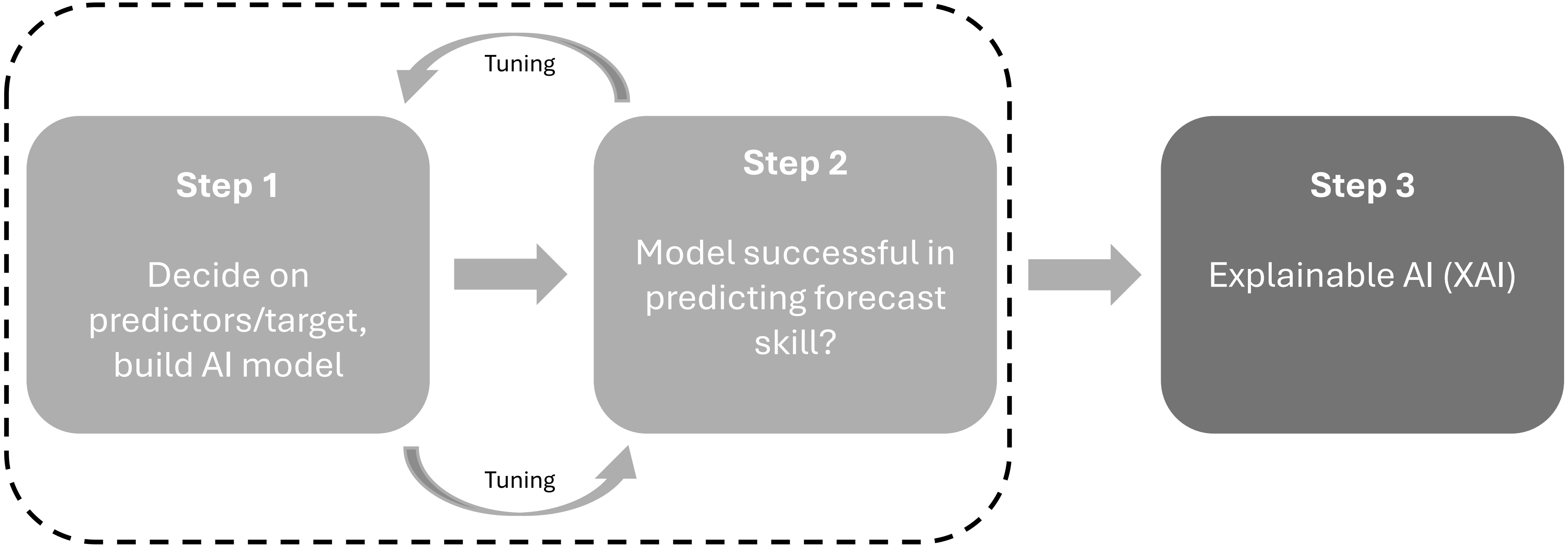
Prediction of exceptionally poor and good forecast

(preliminary results, work in progress!)

Predicting forecast skill over Europe

- **Recap Part 1:** composites and variability of patterns linked to exceptionally poor and good large-scale forecasts

  ⟶ Do we miss certain precursors of exceptionally poor/good forecasts?
  ⟶ Can we use artificial intelligence (AI) to identify which variables and regions are important to look at to predict the occurrence of good/poor forecasts?



Predicting forecast skill over Europe

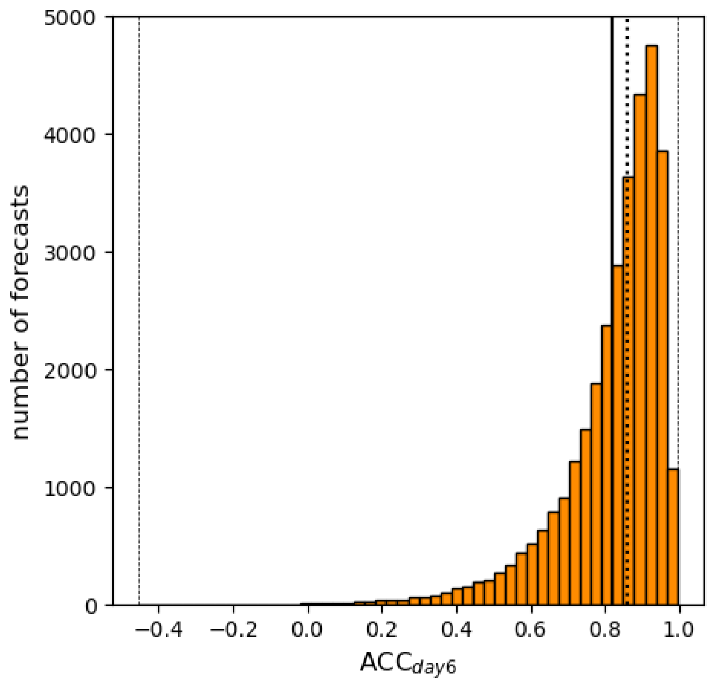
Convolutional neural network (CNN)

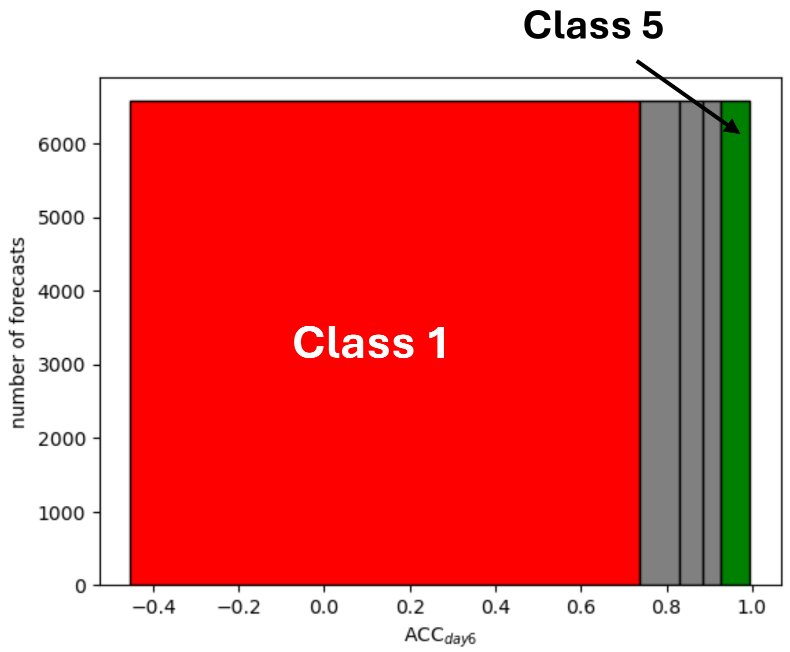
(1) Target

ACC over Europe at day 6

↓

Classification CNN instead of ACC value prediction (5 classes in total based on ACC percentile thresholds)





(2) Predictors

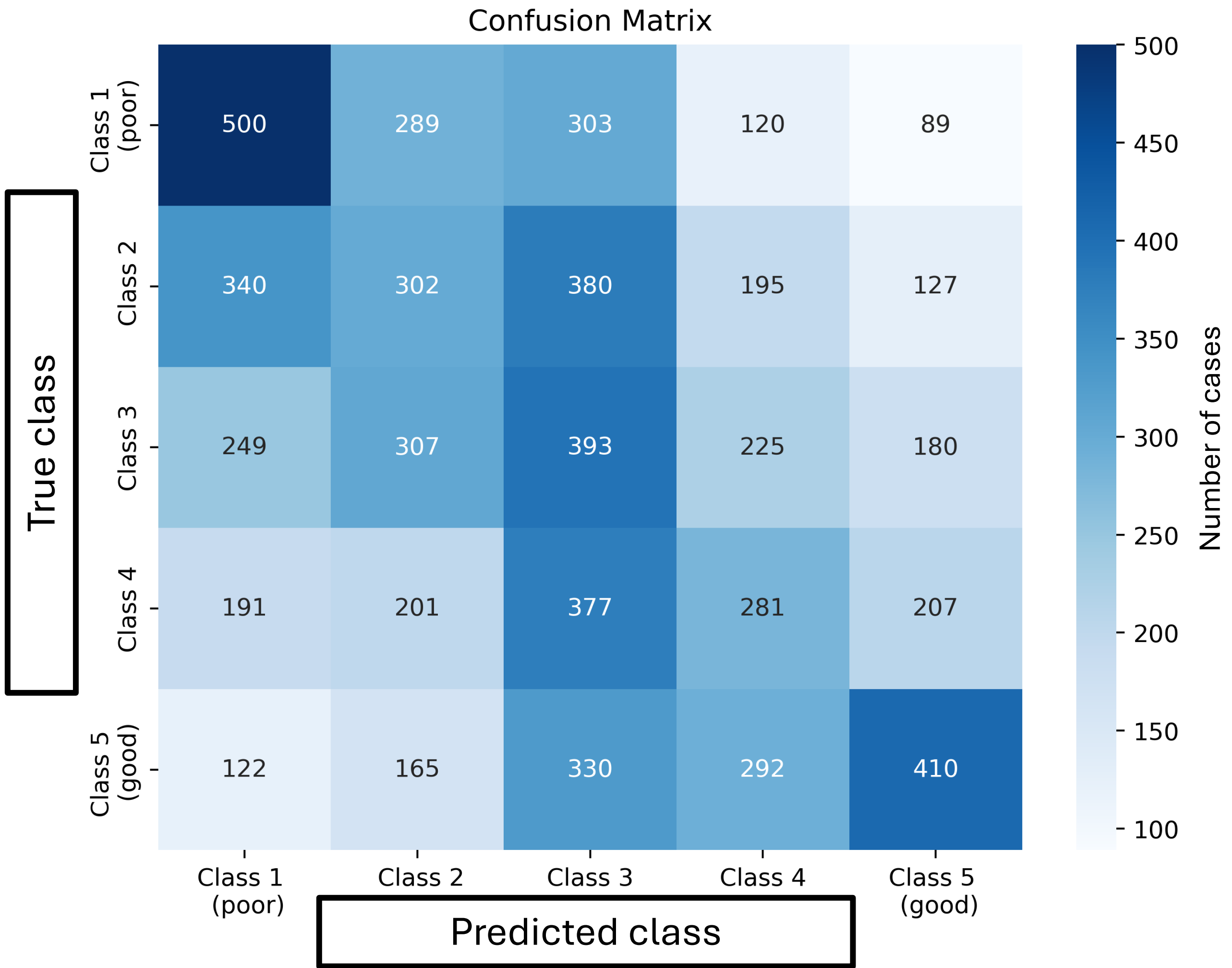
Z500 | MSLP | CAPE | T@850hPa | Divergent wind@250hPa | TCW | Day of the year

2D lat-lon fields at initial time (day 0), full Northern Hemisphere

Predicting forecast skill over Europe

CNN prediction of forecast accuracy at day 6

Testing on 6575 forecast initial times



1. Model is not 'perfect': correct predictions in only 29% of all test cases
2. If predicting the neighboring class is considered as 'correct prediction': 65%
3. Very poor and very good forecasts (20^{th}, 80^{th} percentile) show highest hit rate
4. Very poor skill of the CNN for the "medium skill classes"

Summary | Part 2

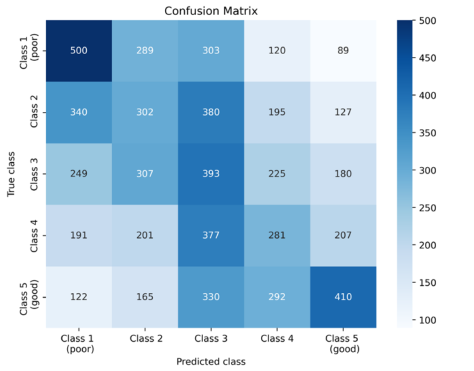
The UNIVERSITY of OKLAHOMA

What did we do?

- Built up a classification CNN to predict the forecast skill class over Europe at day 6 just based on fields at forecast initial time (no ensemble information!)

First finding

- CNN is partly successful, better in predicting skill of very good/poor forecasts



Next steps

- Further improve model (Redefine classes? Check/change predictors? Change CNN architecture? ...)

- Once model performance is acceptable: Which predictor is important for decision and are there specific regions the model focusses on?

Happy for feedback and ideas!

seraphine.hauser-1@ou.edu

References

Grams, C., Beerli, R., Pfenninger, S. *et al. (2017):* Balancing Europe's wind-power output through spatial deployment informed by weather regimes. *Nature Clim Change* 7, 557–562. https://doi.org/10.1038/nclimate3338

Grams, C. M., L. Magnusson, and E. Madonna (2018): An atmospheric dynamics perspective on the amplification and propagation of forecast error in numerical weather prediction models: A case study. *Quart. J. Roy. Meteor. Soc.,* 144, 2577–2591, doi:10.1002/qj.3353.

Hersbach H, Bell B, Berrisford P, et al. (2020): The ERA5 global reanalysis. *Q J R Meteorol Soc.* , 146: 1999–2049. https://doi.org/10.1002/qj.3803

Lillo, S.P. and Parsons, D.B. (2017): Investigating the dynamics of error growth in ECMWF medium-range forecast busts. Q.J.R. Meteorol. Soc., 143: 1211-1226. https://doi.org/10.1002/qj.2938

Magnusson, L. (2017): Diagnostic methods for understanding the origin of forecast errors. Q.J.R. Meteorol. Soc, 143: 2129-2142. https://doi.org/10.1002/qj.3072

Parsons, D.B.; Lillo, S.P.; Rattray, C.P.; Bechtold, P.; Rodwell, M.J.; Bruce, C.M. (2019): The Role of Continental Mesoscale Convective Systems in Forecast Busts within Global Weather Prediction Systems. *Atmosphere* 2019, *10*, 681. https://doi.org/10.3390/atmos10110681

Pickl, M., Quinting, J.F. & Grams, C.M. (2023): Warm conveyor belts as amplifiers of forecast uncertainty. *Quarterly Journal of the Royal Meteorological Society*, 149(756), 3064–3085. Available from: https://doi.org/10.1002/qj.4546

Rodwell, M. J., and Coauthors (2013): Characteristics of Occasional Poor Medium-Range Weather Forecasts for Europe. *Bull. Amer. Meteor. Soc.*, **94**, 1393–1405, https://doi.org/10.1175/BAMS-D-12-00099.1.