Regime-dependent statistical post-processing: Application to wind speed forecasts

SAM ALLEN

FRANK KWASNIOK

CHRIS FERRO

GAVIN EVANS

PIERS BUCHANAN

Study Overview

- Several studies have recognised that the performance of operational weather forecasting systems depends on the prevailing atmospheric circulation.
- Therefore, forecasters often adjust their predictions depending on the synoptic-scale behaviour of the atmosphere.
  - A more objective approach would be to incorporate the circulation directly into the statistical post-processing model.
- To do this, we propose an analogue approach based on atmospheric regimes.
- The approach can be expressed more generally as a mixture-model forecast, which can incorporate uncertainty regarding the prevailing regime.
- This is applied to wind speeds from a quasigeostrophic model, and reforecast data.
- The full study is available in Allen et al. (2020)

Key Results

- Incorporating regime information can yield significant improvements upon conventional post-processing methods if the climatological wind speed varies between the different weather regimes.
  - The conventional post-processing method is not calibrated conditional on the regimes.
- If the wind speeds do not depend on the regimes then the regime-dependent approach reverts back to the original post-processing method.
  - It should always perform at least as well as the original post-processing, provided sufficient data is available.
- Improvements are largest at longer lead times, when the raw ensemble is less informative, but are only available if the future regime can be accurately predicted.
- Forecasts improve most when the prevailing regime is associated with wind speeds that differ most from climatology.
  - This suggests predictions of extreme weather events could benefit from regime information.

Statistical Post-Processing

- Post-Processing exploits the relationship between the Numerical Weather Prediction Model (NWP) model and the atmosphere in the training data to address systematic errors in ensembles

- What if the relationship changes under different circumstances?

- If these circumstances could be identified then they could be incorporated into post-processing models

Grouped Statistical Post-Processing

- Forecast-observation pairs $(y, \boldsymbol{f})$ in the training data can be assigned to a group

$(y_1, \boldsymbol{f}_1)$	A
$(y_2, \boldsymbol{f}_2)$	A
$(y_3, \boldsymbol{f}_3)$	B
$(y_4, \boldsymbol{f}_4)$	A
…	
$(y_{N-1}, \boldsymbol{f}_{N-1})$	B
$(y_N, \boldsymbol{f}_N)$	B

$$Y \mid \boldsymbol{f}, A \sim g_A$$

$$Y \mid \boldsymbol{f}, B \sim g_B$$

- *Separate post-processing models can then be applied to forecasts in each group*
- *g_A and g_B could be predictive distributions, for example,*
  - *With different post-processing coefficients*
  - *Or even distinct underlying parametric families*

- Groups should be chosen such that different model errors are expected in each group

Weather Regimes

- Atmospheric circulation is the movement of air in the atmosphere

- Regimes are patterns in the circulation that exhibit:
    - Persistence (relative to individual weather events)
    - Recurrence
    - At fixed geographical locations

- The atmosphere can be understood as a flow driven from one metastable equilibrium to another (Charney and Devore, 1979)

- Therefore, separate post-processing models can be applied to forecasts depending on the prevailing weather regime

Motivation

- Weather regimes have a large impact on local weather systems

- The forecasting ability of the NWP model changes when the atmosphere resides in different regimes (Ferranti et al., 2015)

- Weather regimes implicitly incorporate information regarding spatial and multivariate relationships

- "certain weather impacts (such as coastal flooding, extreme heat and poor air quality) are more likely to occur during the occurrence and persistence of a few specific weather patterns" (Met Office website, 2016; Neal et al, 2016)

Ensemble Model Output Statistics (EMOS)

- For exchangeable ensemble members f_j ($j = 1, \ldots, M$) with ensemble mean $\bar{f}$ and ensemble variance s^2, wind speed y can be modelled using a truncated Normal distribution:

$$y|f_1, f_2, \ldots, f_M \sim N_0(\alpha + \beta\bar{f}, \gamma + \delta s^2)$$

where $\alpha, \beta, \gamma, \delta$ are parameters to be estimated

- Parameters are estimated here using maximum likelihood estimation over a training data set of historical forecast-observation pairs

Thorarinsdottir and Gneiting. (2010)

Regime-dependent EMOS

- For exchangeable ensemble members f_j $(j = 1, \ldots, M)$ with ensemble mean $\bar{f}$ and ensemble variance s^2, wind speed y can be modelled using a truncated Normal distribution that depends on the weather regime:

$$y|f_1, f_2, \ldots, f_M, r \sim N_0(\alpha_r + \beta_r \bar{f}, \gamma_r + \delta_r s^2)$$

where r is the prevailing atmospheric regime

- We now have a set of parameters for each regime $(\alpha_r, \beta_r, \gamma_r, \delta_r \; for \; r = 1, \ldots, R)$
- Parameters $\alpha_r, \beta_r, \gamma_r, \delta_r$ are estimated using maximum likelihood over all forecast-observation pairs in the training data that are assigned to regime r
- This can be thought of as a regime-based analogue approach (Barnes et al. 2019)

Allen et al. (2019)

Regime-dependent EMOS

- There is typically uncertainty regarding the atmospheric regime at the forecast validation time
- To account for this, model the wind speed using a weighted mixture of predictive distributions:

$$y|f_1, f_2, \dots, f_M, r \sim \sum_{r=1}^{R} w(r) N_0(\alpha_r + \beta_r \bar{f}, \gamma_r + \delta_r s^2)$$

where $w(r)$ specifies the probability of the atmosphere residing in regime r at the validation time

- The weight is a function of the prevailing atmospheric flow, not just a parameter
- The model on the previous slide is a specific case when the regimes are known exactly – the weights in this case are indicator functions
- If the weight is not an indicator function (i.e. regimes are not known with certainty) then all parameters are estimated simultaneously using maximum likelihood over all available training data

Mixture-model weights

- The mixture-model weight $w(r)$ can be thought of as a prediction of the future regime
- We consider three choices of the weight:

1. The regime at the forecast initialisation time
   - i.e. a persistence forecast for the future regime
   - Weight is an indicator function since the regime can be determined from current analyses

2. The proportion of ensemble members predicting each regime at the validation time
   - Weight is not an indicator function so all parameters are estimated simultaneously

3. The regime that actually occurs at the forecast validation time
   - This is not known in practice, but is available when working with historical data
   - It provides an upper bound on the improvements gained from incorporating regimes
   - Weight is an indicator function since the regime can be determined from observations

Outline

- We apply these approaches to wind speed forecasts in two scenarios:
  - Data from a quasigeostrophic model of the Northern Hemisphere
  - Data from the National Oceanic and Atmospheric Administration's (NOAA) Reforecasting project

- Forecasts are assessed using the continuous ranked probability score (CRPS)
  - And the associated skill-score (CRPSS), using the original truncated Normal (TN) approach as a reference forecast

- Regime-dependent truncated Normal (RDTN) approaches use a mixture-model with the:
  - Regime at the initialisation time (-init)
  - Proportion of ensemble members predicting each regime at validation time (-ens)
  - True regime at validation time (-true)

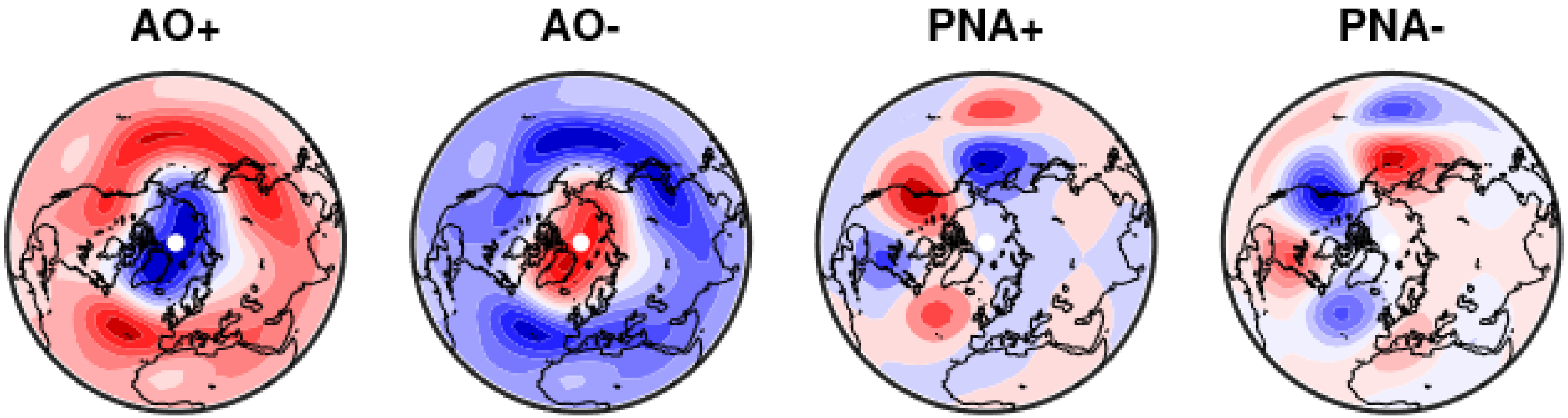
  as regime weights

Quasigeostrophic (QG) model

- Use a three-layer quasigeostrophic model truncated at wavenumber 21 (Kwasniok, 2019)
  - Complex enough to generate atmospheric patterns that appear in climate reanalyses
  - Simple enough to allow a large amount of data to be simulated
- The same QG model truncated at wavenumber 19 is used to generate forecasts
- The training and test data both consist of 15 years worth of daily forecast-observation pairs
- Post-processing is performed locally at 1024 grid points in the Northern Hemisphere

QG model

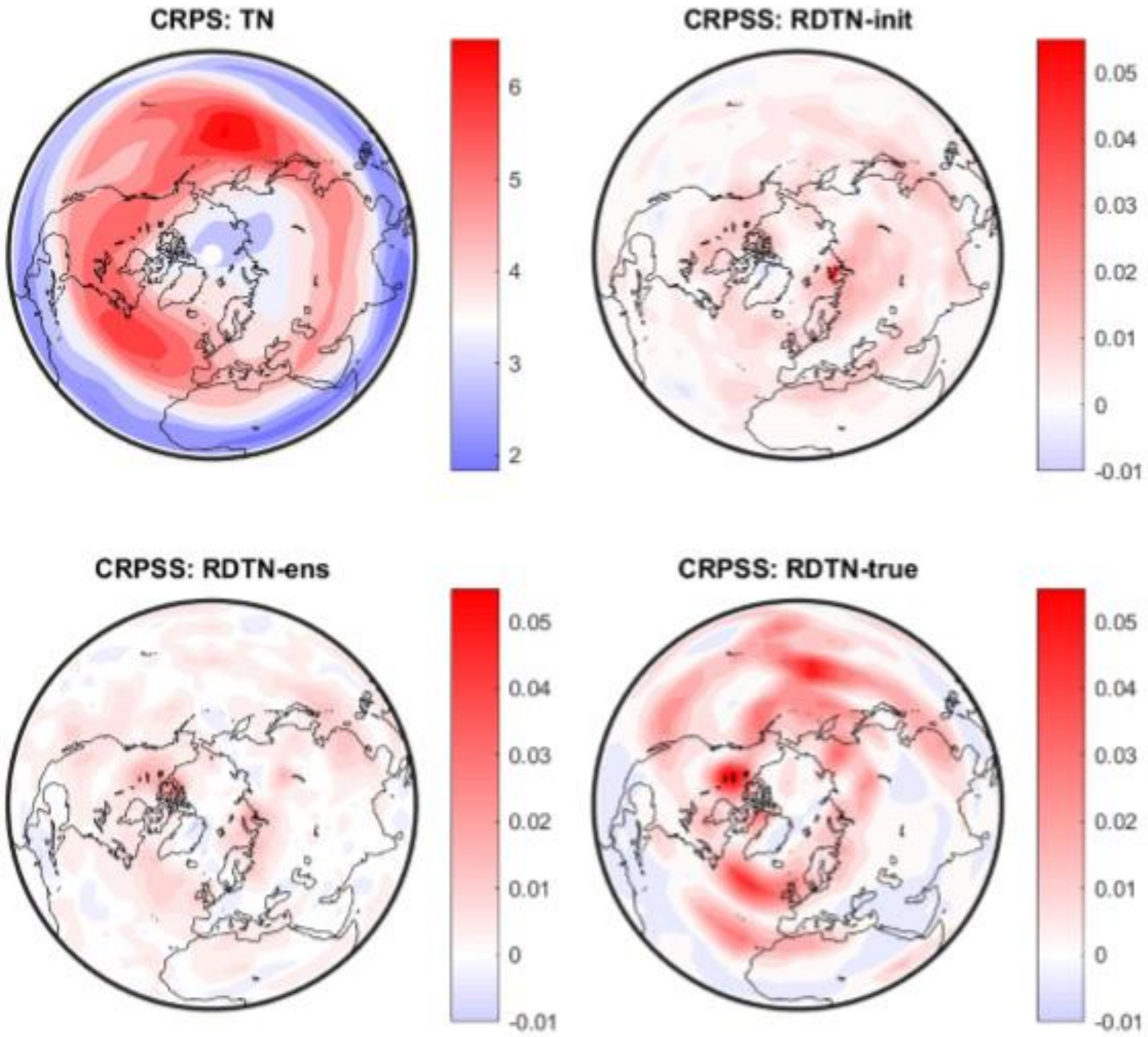
- We identify 4 regimes by fitting a hidden Markov model to 500mb streamfunction anomalies
- Regime centres look similar to the positive and negative phases of the Arctic Oscillation (AO) and Pacific-North America pattern (PNA)



Blue (red) regions represent negative (positive) streamfunction anomalies

QG model

- Wind speeds are least predictable over the Pacific and Atlantic basins
- Skill scores for regime-dependent methods are close to zero at locations where the regimes have little effect on the wind speeds
- Large improvements are available at locations surrounding the centres of the regimes when using the true regime at forecast validation time
- These improvements are much smaller when the future regime is unknown

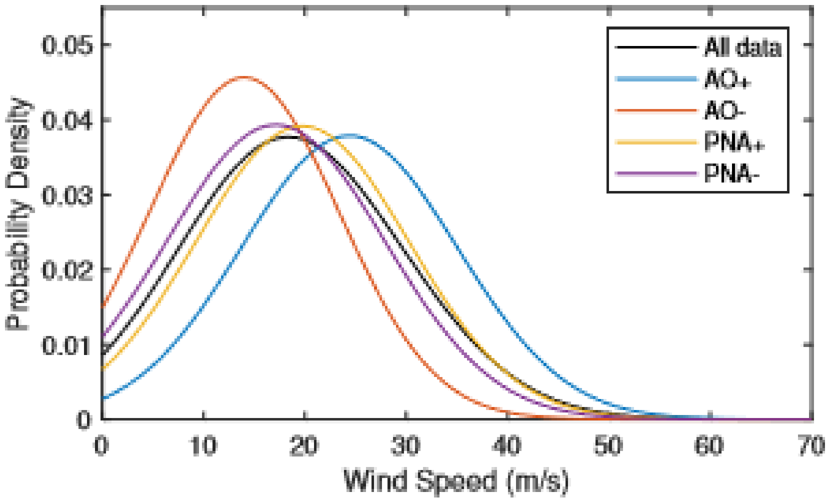


Lead time: 6 days

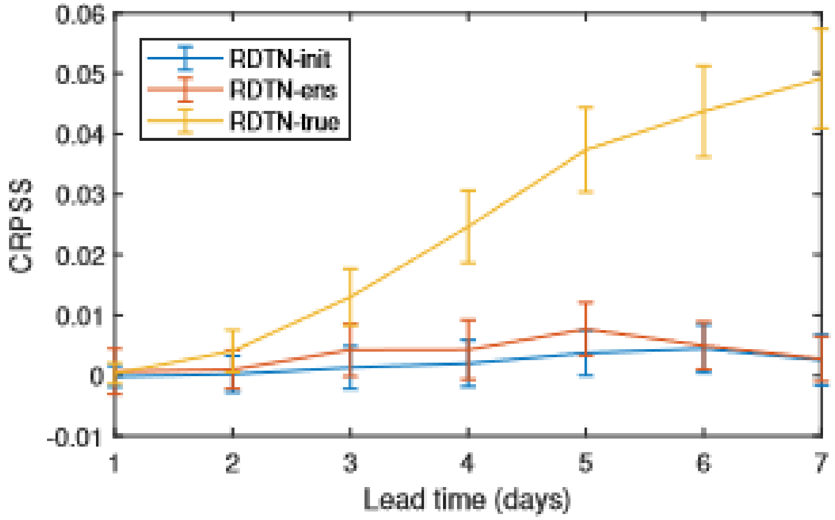
QG model

- Consider forecasts at one location in the Atlantic Ocean where the wind speed varies considerably between the AO regimes.
- Improvements increase with lead time, but only when the true regime is known.
  - RD methods can improve forecasts by almost 5% upon conventional post-processing

Climatological wind speed distributions



CRPSS against lead time

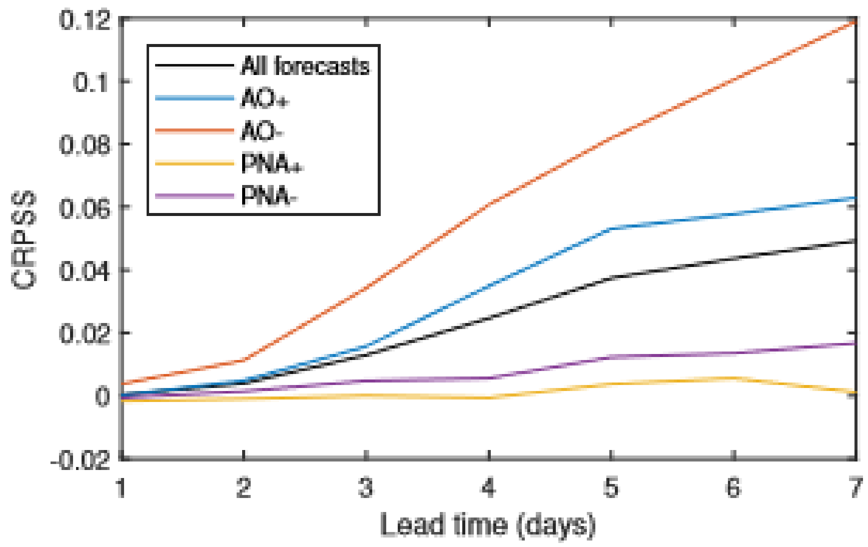


QG model

- Improvements are largest in regimes for which the wind speed differs most from climatology
  - Up to 12% improvements for forecasts when the AO- regime occurs at validation time
  - Up to 6% improvements for forecasts when the AO+ regime occurs at validation time

- The AO+ regime is synonymous with high wind speeds at this location
  - Regime-dependent methods could produce more accurate forecasts of more extreme weather events

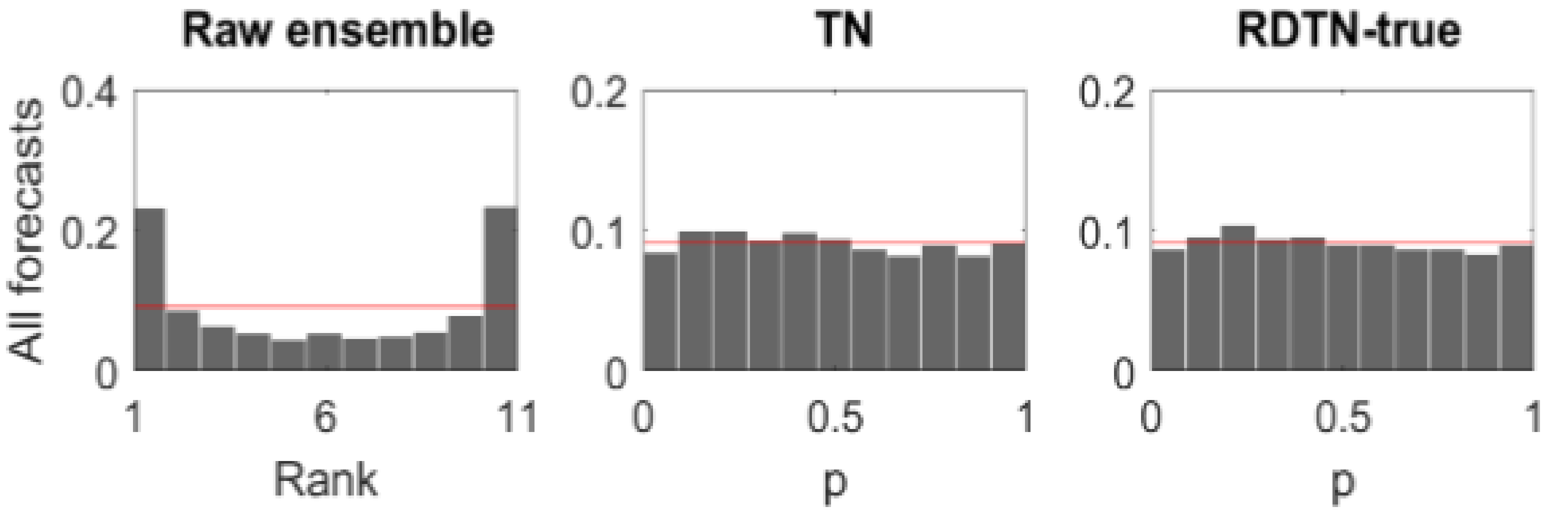
- The regime-dependent approach does not perform worse than conventional post-processing even for regimes that have little effect on the wind speeds



CRPSS against lead time for RDTN-true in each regime

QG model

- Rank and PIT histograms graphically assess probabilistic forecasts
  - Uniform histograms (bars lying close to red line) indicate forecasts are calibrated
- U shaped histogram shows the raw ensemble forecast is underdispersed
- All post-processing methods produce forecasts that appear calibrated

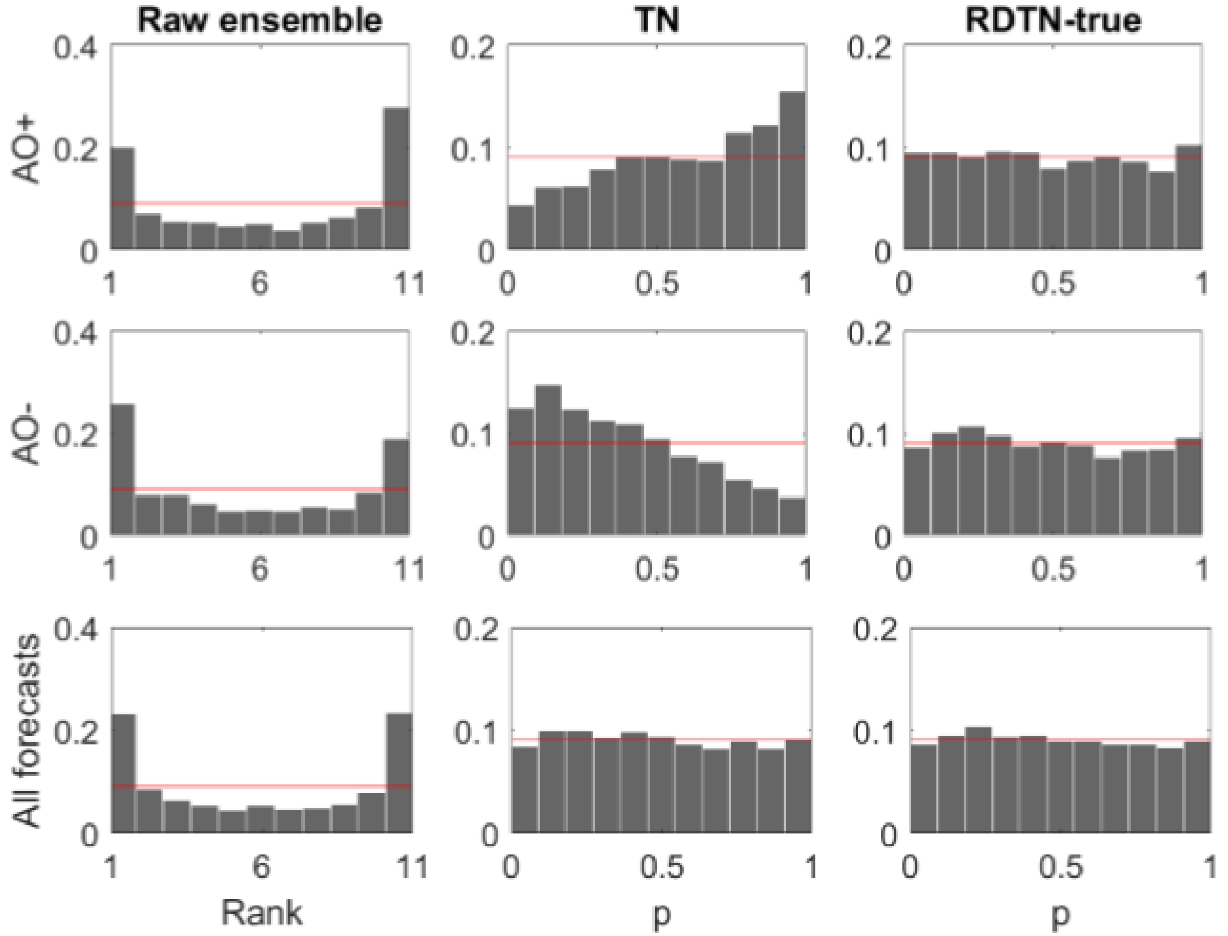


QG model

- But the TN approach is oppositely biased in the AO- and AO+ regimes
  - The conventional post-processing is not calibrated with respect to the regimes

- Calibrating forecasts in each regime separately alleviates these errors

- If the regime at the validation time is not predicted well (as for RDTN-init and RDTN-ens here) then biases are similar to those for TN



GEFS Reforecasts

- Forecasts from the National Centers for Environmental Prediction's (NCEP) Global Ensemble Forecasting System (GEFS) (Hamill et al. 2013)

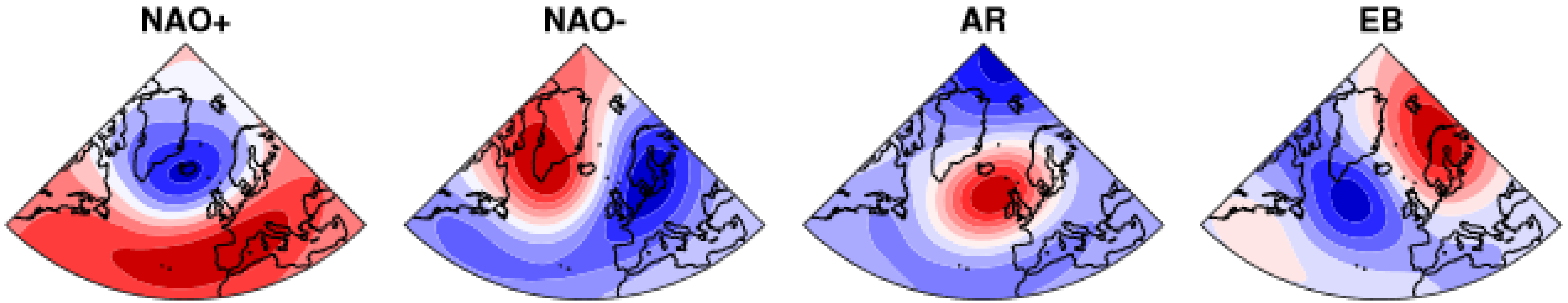
- Post-processing is performed locally at 1353 grid points in the Euro-Atlantic region
  - A subset of the domain on which regimes are identified

- Training data is 15 winter seasons (Nov – Mar) between 1985 and 1999

- Test data is 10 winter seasons between 2000 and 2009

GEFS Reforecasts

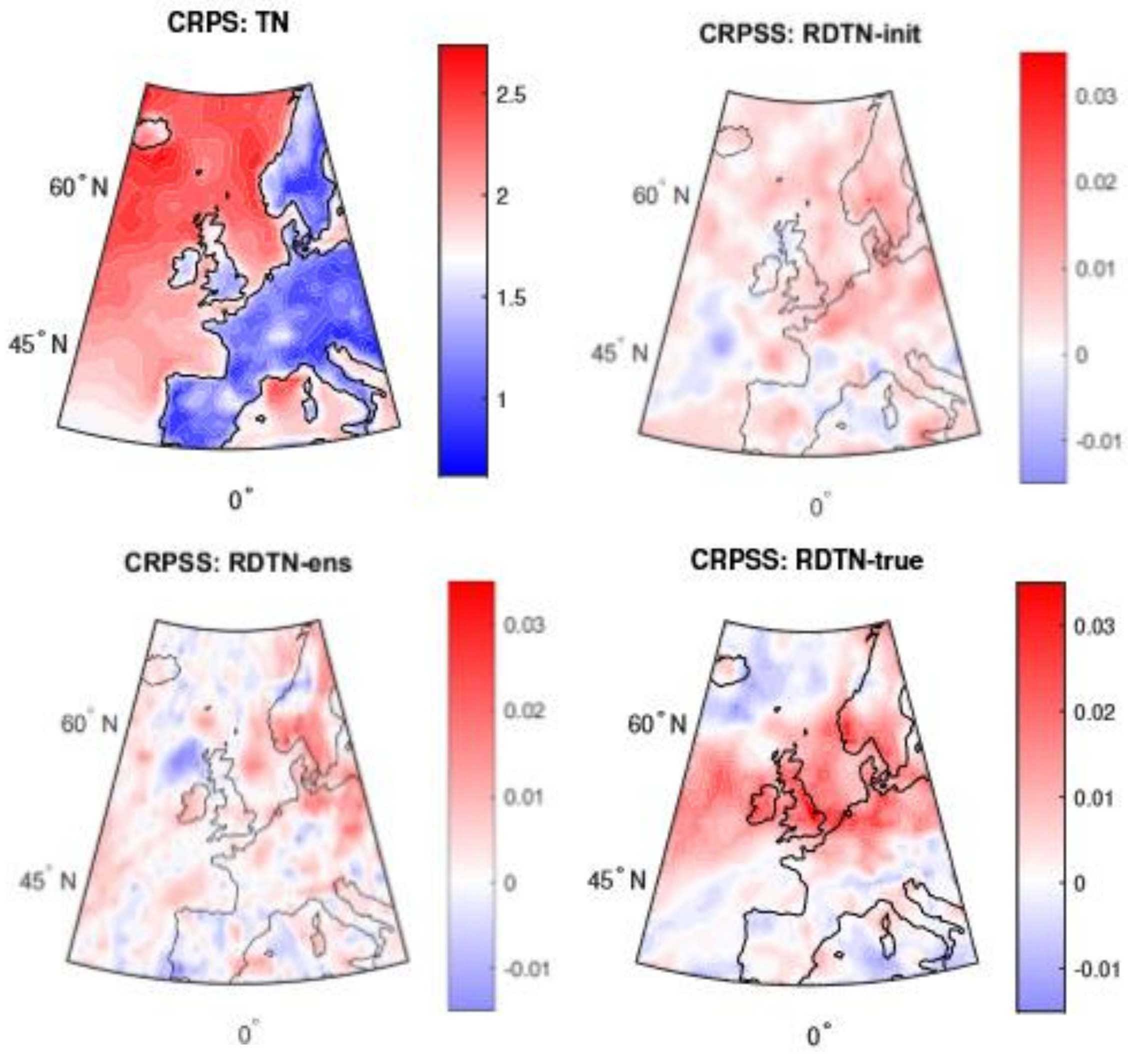
- We identify 4 regimes by applying k-means clustering to 500mb geopotential height anomalies
- Regime centres look similar to the positive and negative phases of the North Atlantic Oscillation (NAO), an Atlantic Ridge (AR) and European Blocking (EB; or a Scandinavian High)



Blue (red) regions represent negative (positive) height anomalies

GEFS Reforecasts

- CRPS is larger over sea than land, and is particularly large close to Iceland, a mode of North Atlantic storm-track variability
- Significant improvement is only available when regimes affect local wind speeds
- Large improvements are again seen only when the regime is known



Lead time: 7 days

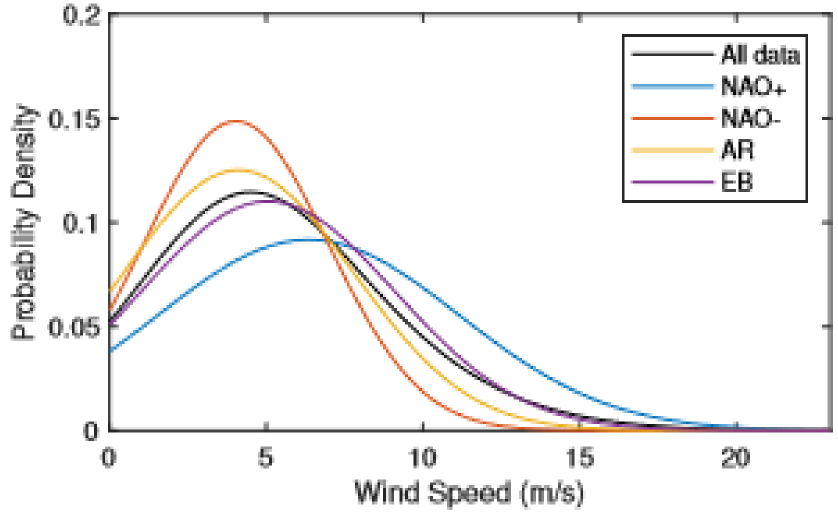
GEFS Reforecasts

- Consider forecasts at one location - close to Bergen, Norway
- High wind speeds typically occur in the positive phase of the NAO, and low wind speeds in the NAO-
- CRPS for the raw ensemble changes between the regimes
  - Highest when the NAO+ occurs and lowest for the AR regime
  - Suggests model biases that differ between the regimes

CRPS	NAO+	NAO-	AR	EB	Total
Raw Ensemble	1.44	1.18	1.06	1.24	1.23

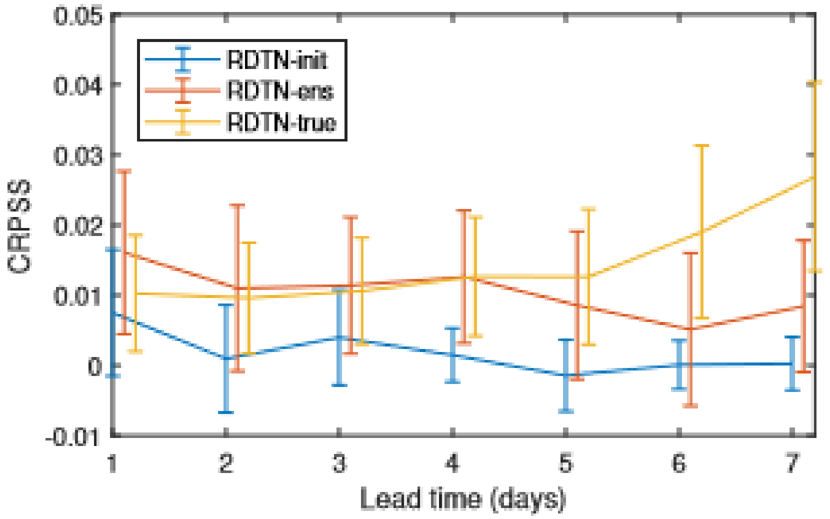
CRPS for raw ensemble forecasts in each regime

Climatological wind speed distributions

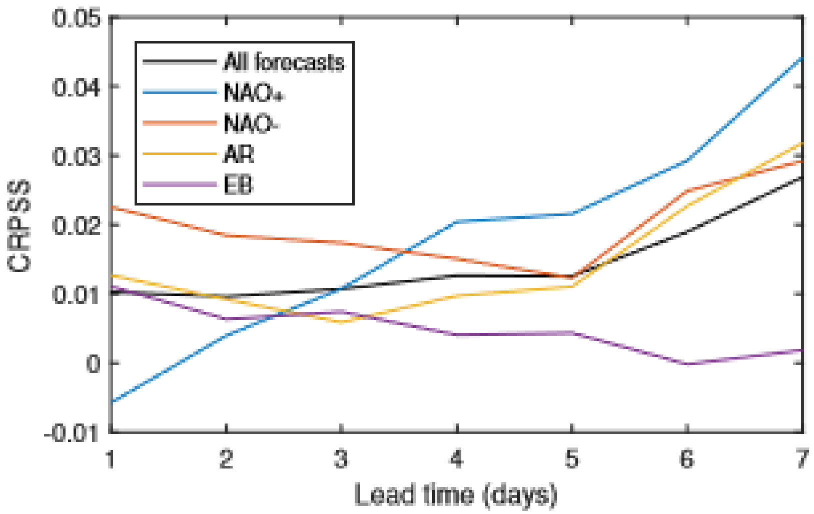


GEFS Reforecasts

- Improvements decrease with lead time for RDTN-init and RDTN-ens, as forecasts of the regime become worse
  - Skill-score increases with lead time when the regime at validation time is known
- Longer-range forecasts benefit most in the NAO+, synonymous with above average wind speeds



CRPSS against lead time



CRPSS against lead time for RDTN-true in each regime

Conclusions & Extensions

- Incorporating atmospheric circulation can improve statistical post-processing methods
- The method here uses a mixture of truncated Normal predictive distributions
  - This is more complex and requires more training data, but adds flexibility to the post-processing model
  - Study using a high-resolution model for which reforecasts are not available is currently ongoing
  - Different predictive distributions could be used in different regimes
- Little improvement is expected when the regimes don’t affect the local wind speeds
- Not sufficient to know the regime at the forecast initialisation time
  - Require a more informative prediction of the future regime
- More improvements available at longer lead times
  - Post-processing should issue the climatological distribution as the raw forecast becomes uninformative
  - Regime-dependent methods issue the climatological distribution within each regime

Conclusions & Extensions

- Forecasts improve most when the prevailing regime corresponds to wind speeds that differ largely from climatology
  - Forecasts of extreme weather events may benefit from including regime information

- The mixture-model approach extends to other ways of grouping the forecasts
  - When would we most expect biases to occur?
  - Optimum choice, and number, of regimes may change for different variables and locations

- Regimes here have the benefit that they are physically meaningful
  - They can account for relationships between different weather variables and spatial locations
  - Sensible for use within multivariate post-processing frameworks

References

- Allen S, Ferro C. A. T., Kwasniok F. 2019: Regime-dependent statistical post-processing of ensemble forecasts. Q. J. R. Meteorol. Soc. 145: 3535–3552.

- Allen S, Ferro C. A. T., Kwasniok F. 2020: Recalibrating wind speed forecasts using regime-dependent Ensemble Model Output Statistics. Q. J. R. Meteorol. Soc. Accepted.

- Barnes C, Chandler R. E., Brierley C. M. 2019: New approaches to postprocessing of multi-model ensemble forecasts. Q. J. R. Meteorol. Soc. 145: 3479–3498.

- Charney J, Devore J. 1979: Multiple Flow Equilibria in the Atmosphere and Blocking. J. Atmos. Sci. 36: 1205-1216.

- Ferranti L, Corti S, Janousek M. 2015: Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. Q. J. R. Meteorol. Soc. 141: 916–924.

References

- Hamill T. M., Bates G. T., Whitaker J. S., Murray D. R., Fiorino M, Galarneau Jr T. J., Zhu, Y., Lapenta, W. 2013: NOAA's second generation global medium-range ensemble reforecast dataset. Bull. Amer. Meteor. Soc, 94:1553–1565.

- Kwasniok F. 2019: Fluctuations of finite-time lyapunov exponents in an intermediate-complexity atmospheric model: a multivariate and large deviation perspective. Nonlinear Processes in Geophysics, 26:195–209.

- Neal R, Fereday D, Crocker R, Cromer R. 2016: A flexible approach to defining weather patterns and their application in weather forecasting over Europe. Met. App. 23: 389-400.

- Thorarinsdottir T. L., Gneiting T. (2010). Probabilistic forecasts of wind speed: ensemble model output statistics by using heteroscedastic censored regression. J. Roy. Stat. Soc., 173:371–388.