Incorporating the North Atlantic Oscillation into the post-processing of MOGREPS-G wind speed forecasts

DOI: https://doi.org/10.1002/qj.3983

SAM ALLEN

GAVIN EVANS

PIERS BUCHANAN

FRANK KWASNIOK

UNIVERSITY OF EXETER

Met Office

Introduction

- Statistical post-processing methods have become a key component of operational weather forecasting over the past two decades

- Post-processing removes systematic errors that manifest in ensemble prediction systems (EPSs)…

- … by exploiting the relationship between the EPS and the atmosphere, as identified from past observations

- This relationship may change under different circumstances
  - Techniques exist to account for temporal and spatial structures in the forecast biases
  - But errors may also depend on the prevailing weather regime

Weather Regimes

- Weather regimes are patterns in the atmosphere's circulation that exhibit:
  - Persistence (relative to individual weather events)
  - Recurrence
  - At fixed geographical locations
  - e.g. the North Atlantic Oscillation, the Arctic Oscillation, atmospheric blocking

- These regimes are known to have a large impact on local weather systems (Hannachi et al., 2017)

- Just as separate post-processing models are often fit to individual seasons or locations to account for temporal and spatial biases…

- … 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
- But NWP models are often unable to simulate the observed regime behaviour (Dawson and Palmer, 2015)
- Hence, the forecasting ability of the NWP model changes when the atmosphere resides in different regimes (Ferranti et al., 2015)
- These regimes are also often linked to the occurrence of more extreme weather events
    - “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

Regime-dependent post-processing

- Let Y denote the weather variable we are trying to predict, let x be the output from an ensemble prediction system and let r denote the prevailing weather regime
- Statistical post-processing models typically seek to identify the conditional distribution of the target variable given the output from the EPS

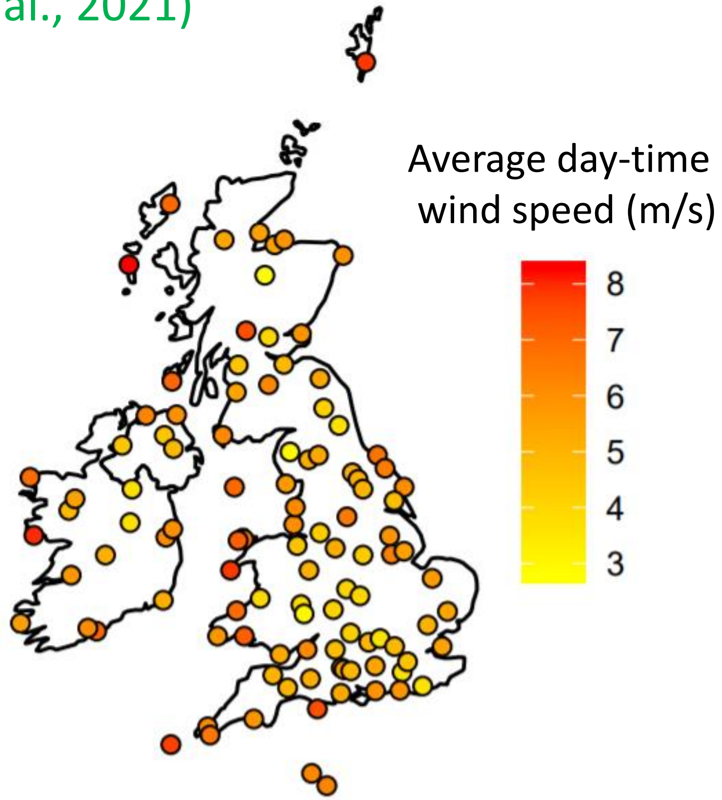
Conventional post-processing	Regime-dependent post-processing
$Y \mid x \sim p(y \mid x)$	$Y \mid x, r \sim p(y \mid x, r)$

- Regime-dependent post-processing extends this to model the conditional distribution given the EPS output and the prevailing weather regime

- This serves as a meteorologically-driven way of adding information to post-processing methods

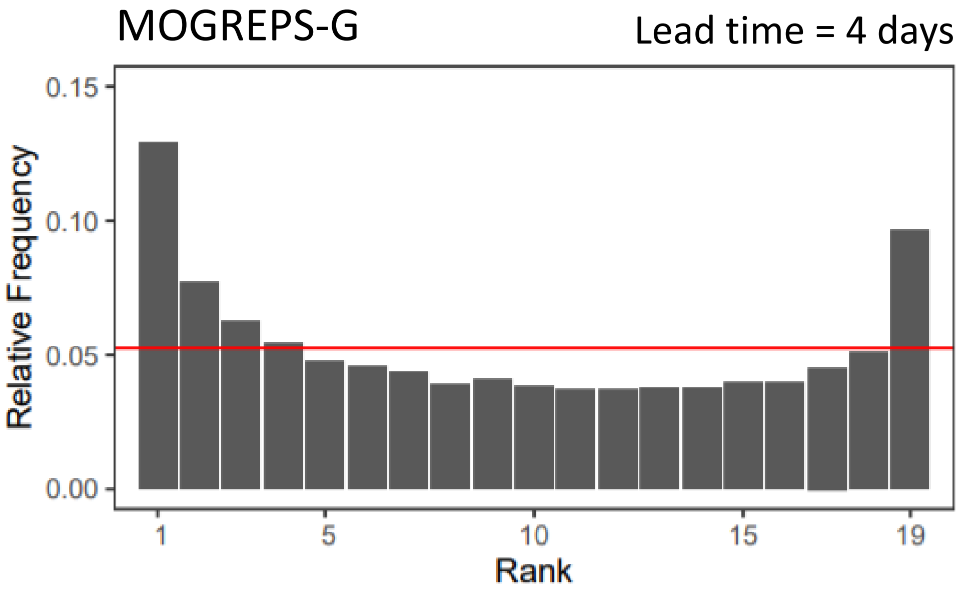
Application

- We apply this regime-dependent post-processing framework to wind speed forecasts issued by the Met Office's global ensemble prediction system MOGREPS-G (Allen et al., 2021)
- MOGREPS-G generates 18 ensemble members at a 20km horizontal resolution
- The gridded MOGREPS-G forecast fields are downscaled to 106 stations over the UK and Ireland
- Data is available for the 2 year period between 2018 and 2019
- Forecasts are considered at lead times at 12 hour intervals up to 6 days in advance



MOGREPS-G

- We can check the calibration of the MOGREPS-G ensemble forecasts using rank histograms
- A calibrated forecast will generate a uniform histogram
- The MOGREPS-G ensemble forecasts are underdispersed, i.e. overconfident



Forecasts aggregated over all days in 2019 and all locations

Statistical Post-Processing

- Assume wind speed at a given time and place follows a logistic distribution truncated below at zero

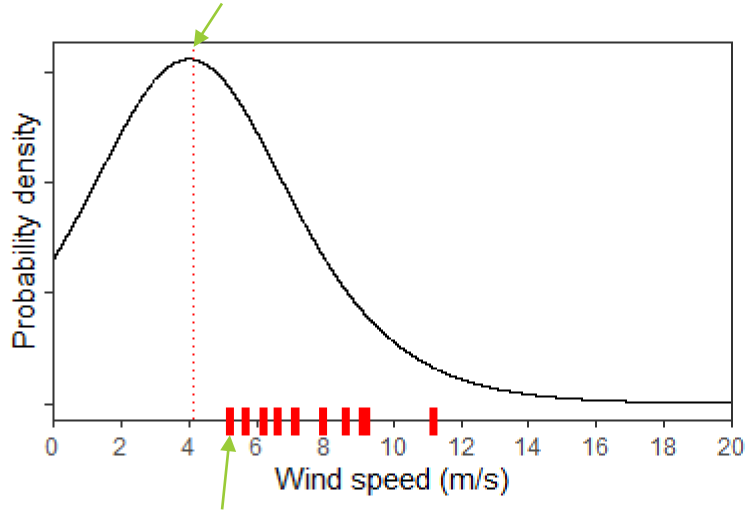
$$Y|x \sim L_0(\mu, \sigma), \qquad \mu = \alpha + \beta \bar{x}, \qquad \sigma^2 = \gamma + \delta s^2$$

- With location μ a linear function of the ensemble mean $\bar{x}$
- and spread σ^2 a linear function of the ensemble variance s^2

(Gneiting et al., 2005; Messner et al., 2014)

- Separate parameters are estimated for each location using forecasts and observations over 2018
- The resulting forecasts are evaluated over 2019

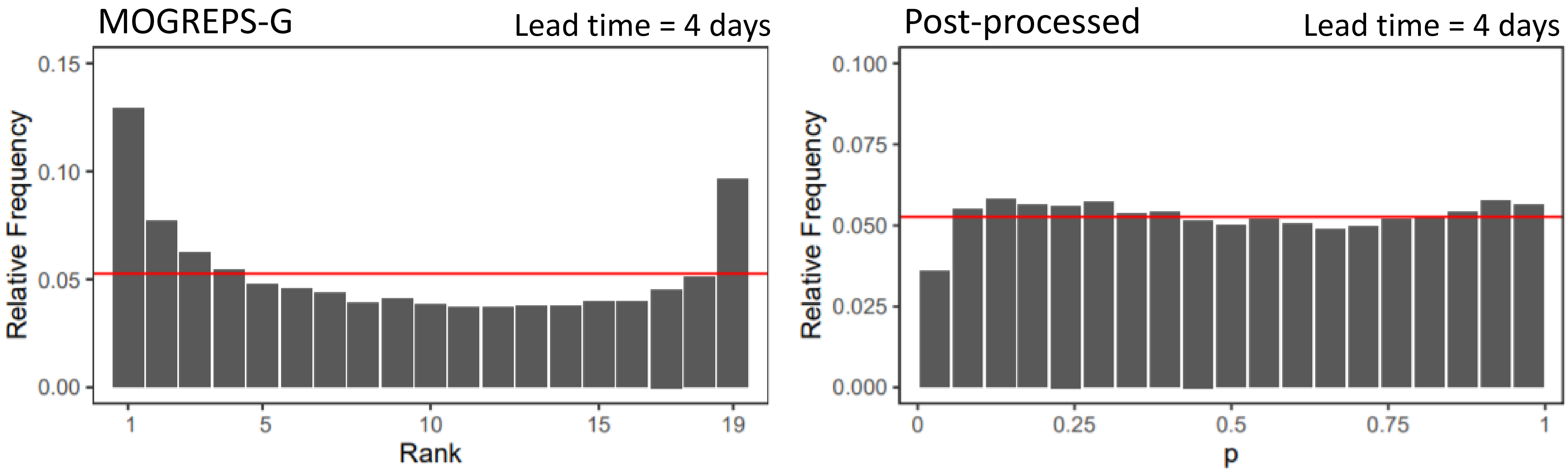
Adjusted ensemble mean



MOGREPS-G ensemble members

Statistical Post-Processing

- The post-processed forecasts are considerably better calibrated than the raw MOGREPS-G output



- But the post-processed forecast may still be subject to conditional biases
- Can regime-dependent post-processing improve upon this established post-processing approach?

Regime-dependent Post-Processing

- Condition post-processing on the prevailing regime, r

$$Y|x, r \sim L_0(\mu_r, \sigma_r), \qquad \mu_r = \alpha_r + \beta_r \bar{x}, \qquad {\sigma^2}_r = \gamma_r + \delta_r s^2$$

- A separate set of parameters is now estimated for each regime
- Parameters are again estimated individually for each location

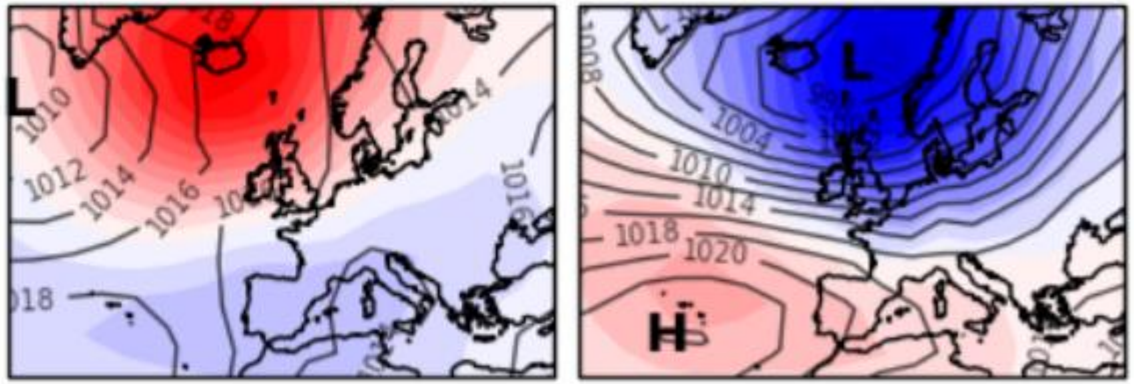
using forecasts and observations over 2018 associated with the relevant regime

- This constitutes a regime-based analogue approach to post-processing
- We want to utilise regimes that have the largest effect on the errors of MOGREPS-G ensembles

The North Atlantic Oscillation

- The North Atlantic Oscillation (NAO) has a strong influence on European weather
  - The positive phase is associated with high wind speeds and precipitation over northern Europe
  - The negative phase is associated with calmer, more stable weather conditions
- The NAO is as defined in the Met Office's Decider tool (Neal et al., 2016)
- We condition post-processing on three regimes:
  - The negative phase of the NAO (NAO-)
  - The positive phase of the NAO (NAO+)
  - When neither NAO regime occurs (Other)

NAO- **NAO+**



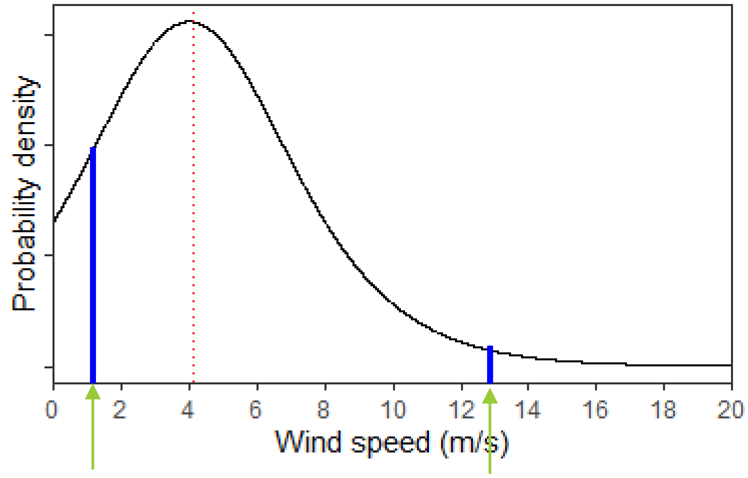
Above average mean sea level pressure

Below average mean sea level pressure

Results

- The conventional post-processing approach generates forecasts that are overdispersed when the more stable NAO- occurs…
- …and underdispersed when the NAO+ occurs.

The observed wind speed should fall between the 0.05 and 0.95 quantiles of the predictive distribution 90% of the time



0.05-quantile

0.95-quantile

Coverage of 90% prediction intervals for forecasts associated with each regime

Method	NAO-	NAO+	Other	Overall
Conventional	92.03	87.69	91.56	91.01

Forecasts aggregated over all days in 2019 and all locations

Lead time = 1 day

Results

- The conventional post-processing approach generates forecasts that are overdispersed when the more stable NAO- occurs…
- …and underdispersed when the NAO+ occurs.
- Regime-dependent post-processing addresses this

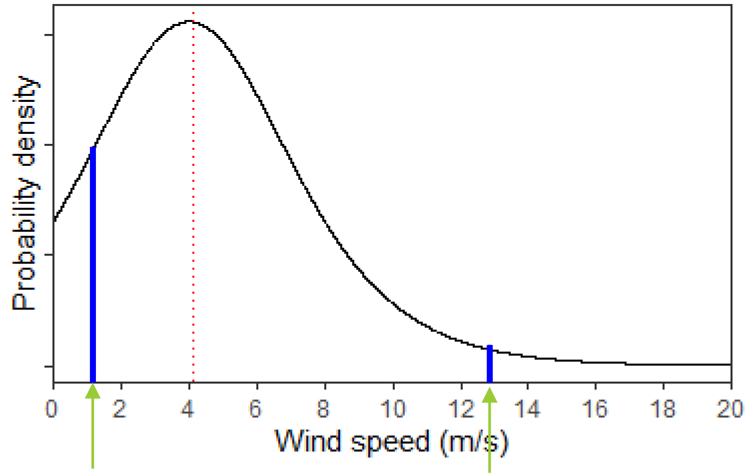
Coverage of 90% prediction intervals for forecasts associated with each regime

Method	NAO-	NAO+	Other	Overall
Conventional	92.03	87.69	91.56	91.01
Regime-dependent	90.26	89.86	90.17	90.14

Forecasts aggregated over all days in 2019 and all locations

Lead time = 1 day

The observed wind speed should fall between the 0.05 and 0.95 quantiles of the predictive distribution 90% of the time



0.05-quantile

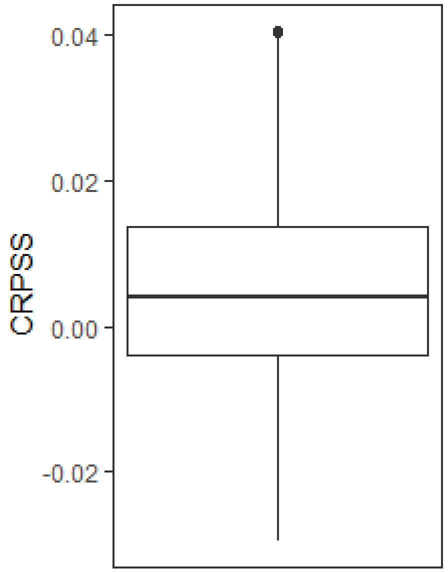
0.95-quantile

Results

- Assess the accuracy of forecasts using the Continuous Ranked Probability Score (CRPS)
- and its skill score (CRPSS; using the conventional post-processing approach as the reference)
- e.g. CRPSS of 0.04 corresponds to a 4% improvement gained by regime-dependent post-processing relative to the conventional framework

- Regime-dependent post-processing is beneficial at roughly 66% of locations

- At which locations is the approach not beneficial?



Boxplot of the relative improvement gained by regime-dependent post-processing at each location

Lead time = 5 days

n = number of wind speed observations in the training data set
n_r = number of forecasts associated with regime r in the training data set
$\bar{y}$ = mean wind speed in the training data set
$\bar{y}_r$ = mean wind speed during regime r in the training data set

Results

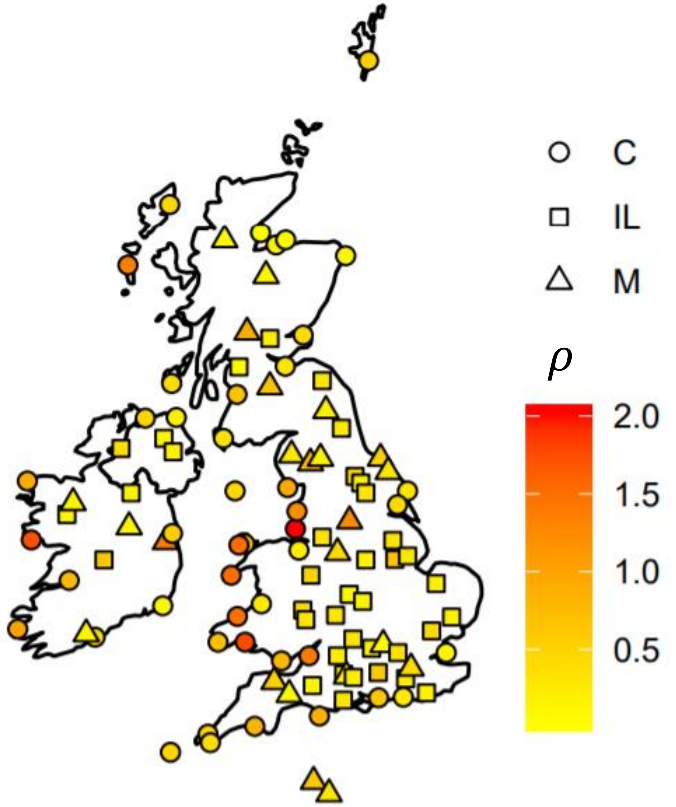
- We expect improvements when the regimes affect the local weather

- Define a measure of regime-dependency as the amount of variation in the wind speeds that is explained by changes in the regime

$$\rho = \sum_{r=1}^{3} \frac{n_r}{n} (\bar{y}_r - \bar{y})^2$$

- We can calculate this measure individually for each location

- Coastal locations (C) are affected more by the phase of the NAO than inland (IL) and mountainous (M) stations

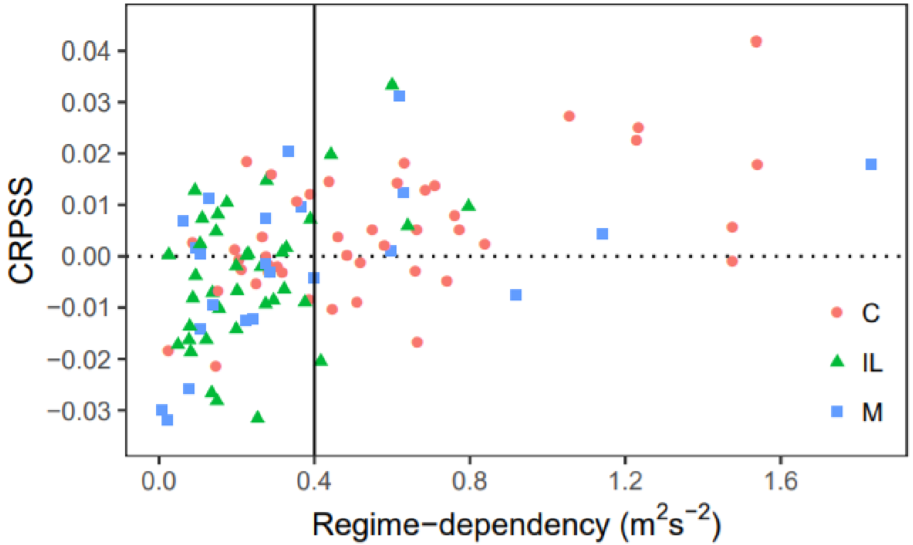


Results

- There is a clear association between our measure of regime-dependency and the improvements gained by regime-dependent post-processing
- Larger improvements tend to occur at coastal locations
- Forecasts become worse when the regime-dependency is low

Lead time = 5 days

Relative improvement



The relative improvement gained by regime-dependent post-processing at each location as a function of the regime-dependency

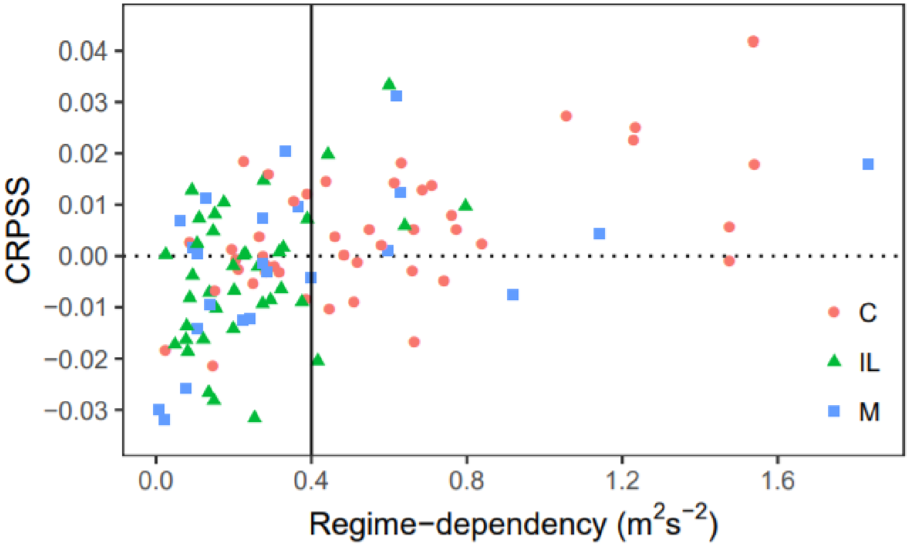
Results

- These results motivate the use of a hybrid post-processing approach that applies regime-dependent post-processing only at locations that are strongly affected by the regimes
  - Defined as locations whose regime-dependency exceeds a chosen threshold

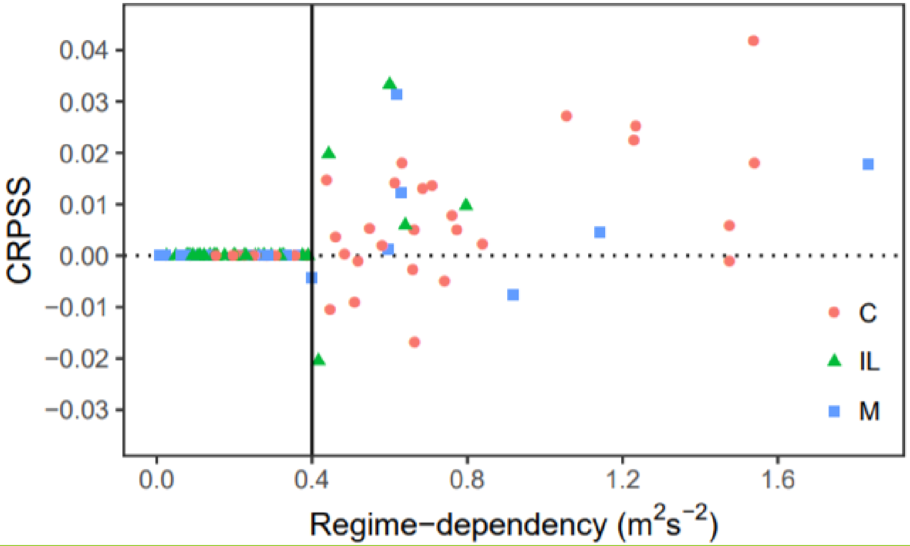
- This removes the negative impact below the chosen threshold, but maintains improvements at locations where the regime-dependency is large

Lead time = 5 days

Relative improvement

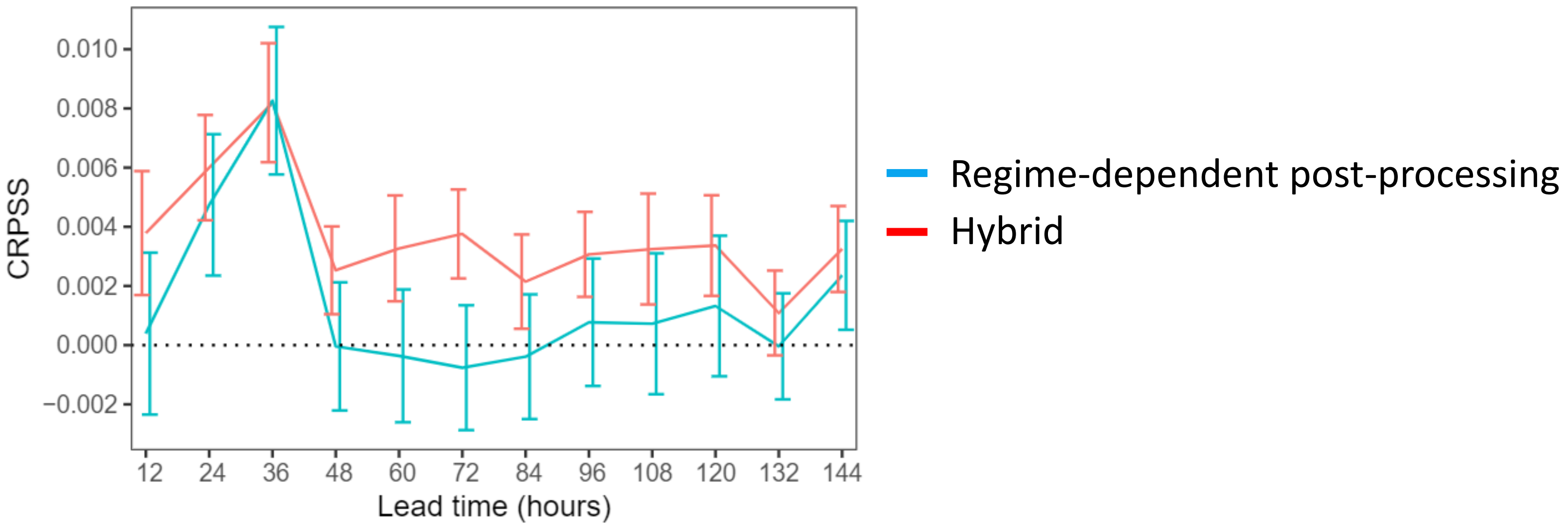


Lead time = 5 days



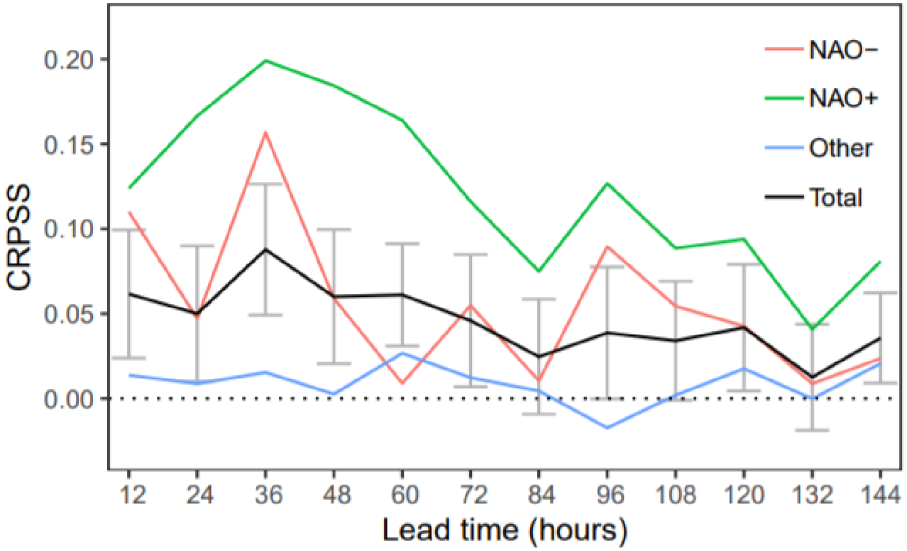
Results

- The regime-based post-processing approaches offer improvements upon conventional methods
- But these improvements are fairly small when averaged over all locations and days
- The benefit is small at longer lead times since we cannot accurately predict the regime far in advance (Allen et al., 2020)



Results

- But at coastal locations, the benefits of regime-dependent post-processing are considerable
- Particularly for forecasts associated with the NAO+ regime

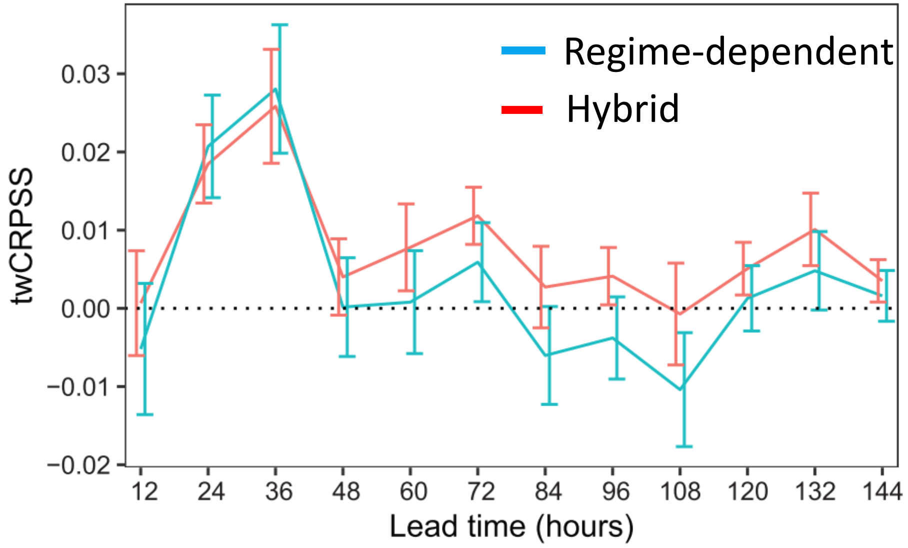


Relative improvement of regime-dependent post-processing at one location on the south-west coast of Wales

- In this example, forecasts improve by almost 20% during the NAO+ at a lead time of 36 hours

Results

- The NAO+ is associated with more extreme wind speeds
- We can emphasise the upper tail of the forecast distribution by evaluating forecasts using the threshold-weighted CRPS (Gneiting and Ranjan, 2011)
  - We employ an indicator weight in the twCRPS that puts emphasis on wind speeds above 8m/s
  - The skill-score is again calculated using the conventional post-processing approach as the reference



- Even when averaged over all locations, the regime-based post-processing methods offer significant improvements at short lead times when assessed using the twCRPS

- Regime-dependent post-processing can improve forecasts made for extreme weather events

Conclusions

- Incorporating the NAO into statistical post-processing methods can improve wind speed forecasts over the UK and Ireland

- Improvements are available when the regimes affect the local climatology
  - Additional parameter uncertainty may outweigh the information provided by the regimes otherwise

- Forecasts improve most when the prevailing regime is associated with wind speeds that differ largely from average (e.g. the positive phase of the NAO)

- Hence, forecasts of extreme weather events can benefit from the inclusion of regime information

Future work

- Apply the regime-dependent framework to other weather variables
  - Precipitation biases, for example, may be more dependent on the weather regime than the season
- Employ within multivariate post-processing methods
  - Weather regimes implicitly incorporate information regarding spatial and multivariate relationships
- Compare with alternative post-processing approaches that add information to the forecast
  - For example, methods based on machine learning
- Develop more accurate forecasts of the future regime in order to benefit at longer lead times

References

- Allen, S., Ferro, C. A. T., & Kwasniok, F. (2019). Regime-dependent statistical post-processing of ensemble forecasts. *Quarterly Journal of the Royal Meteorological Society*, *145*, 3535-3552.

- Allen, S., Ferro, C. A. T., & Kwasniok, F. (2020). Recalibrating wind-speed forecasts using regime-dependent ensemble model output statistics. *Quarterly Journal of the Royal Meteorological Society*, *146*, 2576-2596.

- Allen, S., Evans, G. R., Buchanan, P., and Kwasniok, F. (2021). Incorporating the North Atlantic Oscillation into the post-processing of MOGREPS-G wind speed forecasts. *Quarterly Journal of the Royal Meteorological Society, 147,* 1403-1418

- Dawson, A., and Palmer, T. N. (2015). Simulating weather regimes: Impact of model resolution and stochastic parameterization. *Climate Dynamics, 44,* 2177-2193.

- Ferranti, L., Corti, S., & Janousek, M. (2015). Flow-dependent verification of the ECMWF ensemble over the Euro-Atlantic sector. *Quarterly Journal of the Royal Meteorological Society*, *141*, 916-924.

References

- Gneiting, T., & Ranjan, R. (2011). Comparing density forecasts using threshold-and quantile-weighted scoring rules. *Journal of Business & Economic Statistics*, *29*, 411-422.

- Hannachi, A., Straus, D. M., Franzke, C. L., Corti, S., & Woollings, T. (2017). Low-frequency nonlinearity and regime behavior in the Northern Hemisphere extratropical atmosphere. *Reviews of Geophysics*, *55*, 199-234.

- Messner, J. W., Mayr, G. J., Wilks, D. S., & Zeileis, A. (2014). Extending extended logistic regression: Extended versus separate versus ordered versus censored. *Monthly Weather Review*, *142*, 3003-3014.

- Met Office Website. (2016) Retrieved 26/04/2021 from www.metoffice.gov.uk/research/news/2016/new-weather-patterns-for-uk-and-europe

- Neal, R., Fereday, D., Crocker, R., and Comer, R. E. (2016). A flexible approach to defining weather patterns and their application in weather forecasting over Europe. *Meteorological Applications*, *23,* 389–400