Medium- and Extended-Range Ensemble Weather Forecasting

David Richardson

Abstract The chapter provides an overview of ensemble weather forecasting for the medium- and extended-range (days to weeks ahead). It reviews the methods used to account for uncertainties in the initial conditions and in the forecast models themselves. The chapter explores the challenges of making useful forecasts for the sub-seasonal timescale, beyond the typical limit for skilful day-to-day forecasts, and considers some of the sources of predictability such as the Madden-Julian oscillation (MJO) that make this possible. It then introduces some of the ensemble-based forecast products and concludes with a case study for a European heat wave that demonstrates how ensemble weather forecasts can be used to guide decision making for weather-dependent activities.

Keywords Ensemble • Weather forecast • Uncertainty • Predictability • Medium-range • Extended-range • Sub-seasonal • Madden-Julian oscillation

D. Richardson (✉)
European Centre for Medium-Range Weather Forecasts (ECMWF),
Reading, UK

PREAMBLE

Operational weather forecasts for the medium- and extended-range (days to weeks ahead) are generally based on the output from global Numerical Weather Prediction (NWP) or General Circulation Model (GCM) ensemble forecasts.

INITIAL CONDITION UNCERTAINTIES

The main aim of the ensemble approach is to account for uncertainty in the initial atmospheric conditions. Typically this is done by adding small perturbations to a single "best-estimate" analysis of the current state of the atmosphere. The analysis is generated through assimilation of observations into the NWP model. The size of the perturbations is constrained to be consistent with the known analysis errors. A number of different methodologies are used to generate perturbations that are physically realistic and that will grow to represent the range of possible future states consistent with the initial uncertainty.

The error-breeding method (Toth and Kalnay 1993, 1997) uses a cycling approach where the differences between short-range forecasts are re-scaled to form the initial perturbations for the next forecast. A number of generalisations of the original breeding method have been developed, designed to improve the representation of the analysis uncertainty at each initial time. These include the Ensemble Transform Kalman Filter (ETKF, Bishop et al. 2001), used at the Met Office (Bowler et al. 2008, 2009), and the Ensemble Transform with Rescaling (ETR, Wei et al. 2008) used at the United States National Centers for Environmental Prediction (NCEP).

The Singular Vector (SV) method (Buizza and Palmer 1995) computes new perturbations at each analysis time. The method identifies the fastest growing perturbations over a given time period (e.g. 48 hours). A linear combination of these SVs, scaled to have amplitudes consistent with the analysis error, is added to the best-estimate analysis to make the starting conditions for each ensemble member. The SV method is used operationally at the European Centre for Medium-Range Weather Forecasts (ECMWF) (Leutbecher and Palmer 2008) and the Japan Meteorological Agency (JMA) (Yamaguchi and Majumdar 2010).

Other methods more directly address the observation uncertainty by perturbing the observed values themselves. At ECMWF, perturbations are

also provided from an Ensemble of Data Assimilations (EDA) (Buizza et al. 2008). Each EDA member is an independent data assimilation, using the same set of observations, but introducing perturbations to these observations consistent with the known observation errors. The Meteorological Service of Canada (MSC) uses perturbed observations and an ensemble approach, the ensemble Kalman filter (EnKF, Houtekamer and Mitchell 2005; Houtekamer et al. 2009, 2014), to provide an ensemble of initial conditions. It should be noted that for both the EDA and EnKF, it is necessary to take account of model uncertainties (see below) as well as the observation uncertainties to generate appropriate initial perturbations.

Model Uncertainties

Global NWP ensemble forecasts typically run with a grid spacing of a few tens of kilometres. Many important physical processes (that affect, e.g., clouds and precipitation) work on much smaller spatial scales than can be resolved directly. These processes are represented in the NWP models by "parametrization schemes" that describe the aggregate effect of the smaller-scale unresolved processes on the larger resolved scales.

The finite resolution of the NWP model and the approximations made in the parametrisation schemes are sources of model uncertainty. Most global NWP ensembles also now include a representation of these uncertainties in the model formulation. A range of methods has been developed and sometimes a combination of methods is used in a single ensemble system to account for different aspects of model uncertainty.

One approach is to use a number of different parametrisation schemes within the ensemble. For example, there are various ways to parametrise convective processes and an ensemble can be generated by running some members using one convection parametrisation scheme, while other members use a different convection scheme (Charron et al. 2010). An alternative is to use a single parametrisation, but to perturb some of the key parameters in the scheme (Bowler et al. 2008). Other schemes represent the uncertainty from the sub-grid scale by stochastically perturbing the tendencies from the parametrisation schemes, as, for example, in the Stochastically Perturbed Parametrisation Tendency scheme (SPPT, Leutbecher et al. 2017; Buizza et al. 1999). Backscatter schemes are designed to simulate the transfer of energy from the unresolved sub-grid scales to the larger scales that are resolved by the model (Shutts 2005; Berner et al. 2009).

OPERATIONAL GLOBAL MEDIUM-RANGE ENSEMBLES

A number of meteorological centres produce operational medium-range ensemble forecasts. Initial condition and model uncertainties are represented using many of the above methods, with different centres adopting different approaches. The TIGGE (The International Grand Global Ensemble) project provides access to regular global ensemble predictions from ten of the leading global NWP centres to support research and has facilitated a comprehensive evaluation of the global ensembles produced by different NWP centres (Swinbank et al. 2016; Bougeault et al. 2010; and references therein).

Comparison of the forecasts from the TIGGE centres confirms that while the different centres each have their strengths and weaknesses, the different perturbation methodologies all have merit. It is more important that an NWP system produces an ensemble that accounts for both initial condition and model uncertainties than the precise methodology used to produce the perturbations. However, it is also important to carry out proper and comprehensive evaluation to ensure that the perturbations are consistent with the uncertainties of the system. Figure 8.1 shows an example of the skill of operational ensemble forecasts from five global centres in predicting the large-scale weather patterns over the extra-tropical northern hemisphere up to two weeks ahead (the temperature at 850 hPa is a good indicator of whether a location is under the influence of a warm or cool circulation pattern). Skill is measured using the Continuous Ranked Probability Skill Score (CRPSS), a standard measure for assessing the usefulness of probabilistic forecasts, which can also be interpreted as an indication of the potential economic value of the forecast systems (Palmer and Richardson 2014). CRPSS ranges from a maximum value of 1 (perfect knowledge of what the weather will be) to zero (only the climatological information is known). This evaluation shows that all the forecasting systems have positive skill in forecasting day-to-day changes in the weather for up to two weeks ahead.

EXTENDED-RANGE ENSEMBLES

The medium-range ensembles described above typically produce forecasts for one to two weeks ahead. This is usually considered the limit for day-to-day predictability, as the influence of the atmospheric initial conditions is much reduced at longer range.
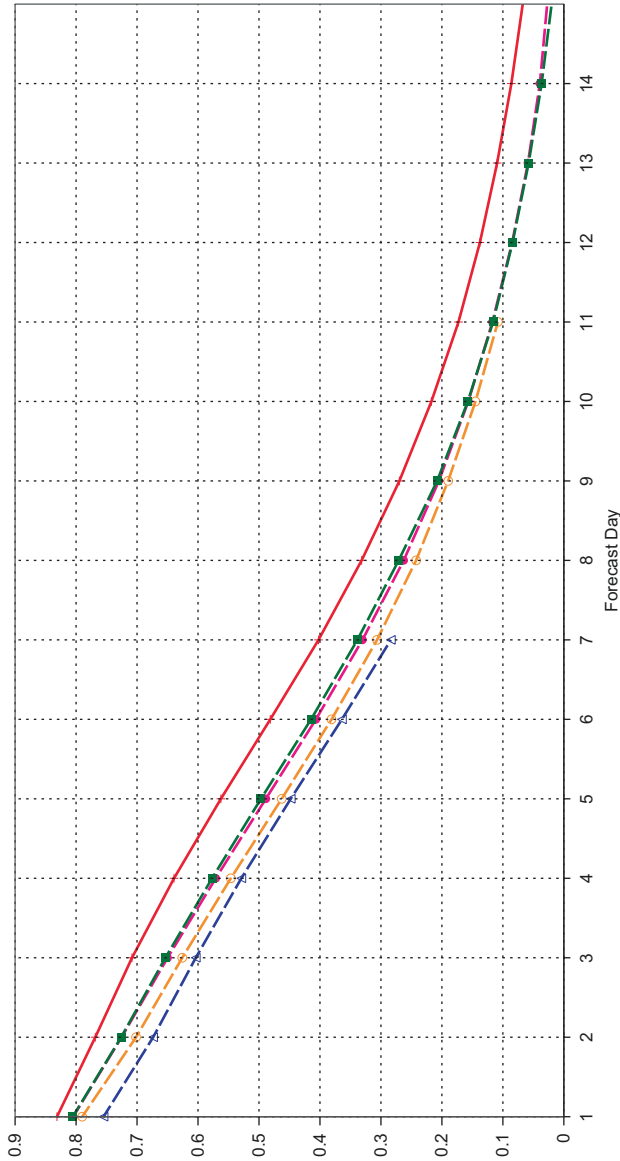
Fig. 8.1 Skill of ensemble forecasts for temperature at 850 hPa in the northern hemisphere extra-tropics for 2016. The verification is performed against each centre's own analysis, with the forecast and analysis data taken from the TIGGE archive. CMC = Canadian Meteorological Centre, JMA = Japan Meteorological Agency, UKMO = United Kingdom Met Office; NCEP = The National Centres for Environmental Prediction (USA)

However, there are important sources of predictability that do give the potential to make useful forecasts for the sub-seasonal scale (one or two months ahead). One important phenomenon is the Madden-Julian oscillation (MJO), a feature of the tropical atmosphere that also impacts on the weather in the extra-tropics (Lin et al. 2009; Cassou 2008). The MJO evolves over a period of 40–60 days, and so is a potential source of forecast skill for several weeks ahead. Initial conditions in the stratosphere can affect the circulation in the troposphere over the following month, providing another source for sub-seasonal predictability (Baldwin and Dunkerton 2001). Other potential sources of predictability at this timescale include the land surface (Koster et al. 2010) and snow cover (Jeong et al. 2013) conditions at the start of the forecast.

There have been significant improvements in sub-seasonal forecasts in recent years, with large improvements in skill for predicting the MJO (Fig. 8.2), as well as its influence on other regions, including Europe (Vitart 2014). There is now a growing interest in developing applications to exploit these forecasts as well as to improve the forecasts themselves.

Following the success of TIGGE, a new sub-seasonal to seasonal prediction project (S2S) has been initiated by the World Weather Research Programme (WWRP) and World Climate Research Programme (WCRP). The main goal of this five-year project is to improve forecast skill and understanding of the sub-seasonal to seasonal timescale and to promote its uptake by operational centres and its exploitation by the applications community (Vitart et al. 2012).

The S2S database includes near real-time ensemble forecasts for up to 60 days ahead, from 11 forecasting centres: Australia's Bureau of Meteorology (BOM); the China Meteorological Administration (CMA); ECMWF; Environment and Climate Change Canada (ECCC); Italy's Institute of Atmospheric Sciences and Climate (CNR-ISAC); the Hydrometeorological Centre of Russia (HMCR); the Japan Meteorological Agency (JMA); the Korea Meteorological Administration (KMA); Météo-France; the US National Centers for Environmental Prediction (NCEP); and the UK Met Office (Vitart et al. 2017).

These models are generally different from the NWP models used to produce medium-range forecasts at the same centres. Most are coupled to an ocean model, as it is important to take account of the evolution of the sea-surface temperature and its interaction with the atmosphere over the longer time periods of these forecasts. For the same reason, some systems also include an active sea ice model. Some centres, such as ECMWF, that
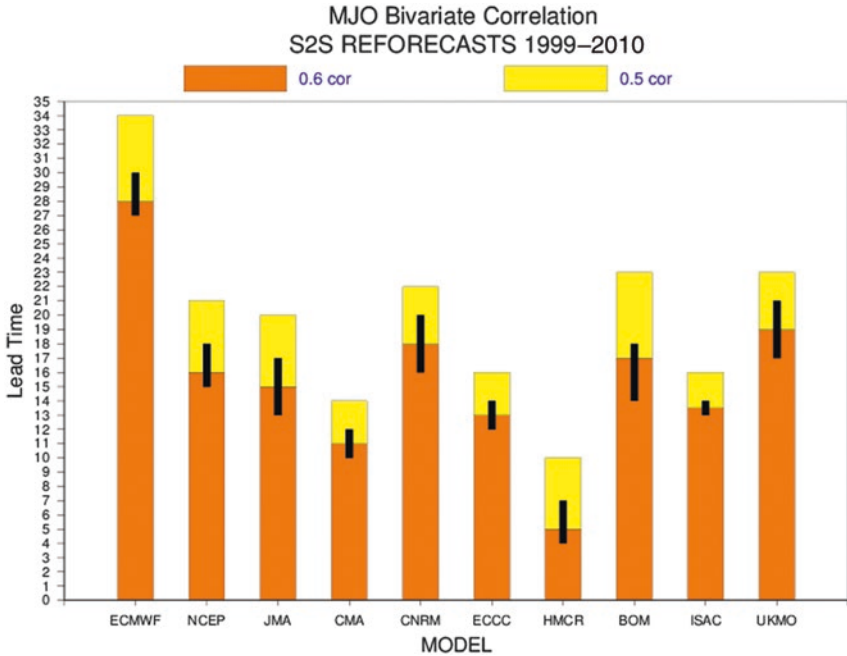
Fig. 8.2 Forecast lead-time (in days) when a correlation-based measure of accuracy of the prediction of the Madden-Julian Oscillation (MJO) reaches 0.6 correlation (orange bars) and 0.5 correlation (yellow bars) (1.0 would indicate a perfect forecast). The black lines indicate the 95% confidence interval of the time when the 0.6 correlation is reached. Results are based on the re-forecast from 1999 to 2010 from all the models, verified against ERA-Interim analyses. Correlations of 0.5 and 0.6 are often used as indication of useful forecast skill (Vitart 2014)

already include an ocean model in the medium-range ensemble, use the same ensemble system for both the medium-range and sub-seasonal forecasts, hence providing a seamless set of predictions covering the timescales from days to months ahead. While the skill in predicting the day-to-day weather may be limited to around two weeks in general (as seen in Fig. 8.1), many of the S2S models demonstrate substantial skill out to three or even four weeks ahead for the MJO (Fig. 8.2); this can lead to enhanced predictability, for example, over Europe, in certain situations.

Another important aspect of the sub-seasonal forecasts is the need to account for model errors. Systematic model errors (biases) can accumulate

during the forecast, and while they are often small enough to be neglected for medium-range forecasts, they become too large to be ignored at the longer sub-seasonal range. Hence, an additional set of historical ensemble integrations is generated by making forecasts from start dates covering the last 15–20 years. These re-forecasts (or hindcasts) are used to estimate the model climate. This can then be used to remove the model biases from the real-time forecasts in a statistical post-processing step (Vitart 2014).

Ensemble Weather Forecast Products

Ensembles are designed to take account of the uncertainties in the initial conditions and in the NWP model used to make the forecast. The set of ensemble forecasts provides a direct quantitative indication of the range of possible future weather scenarios that may occur. Most ensembles are constructed so that each member of the ensemble is equally likely. The proportion of ensemble members forecasting a specific weather event gives an indication of the probability for it to occur. Grouping the ensemble into a small number of clusters can be valuable for those cases when there are distinct alternative scenarios within the ensemble (Ferranti et al. 2015).

The forecast values of weather variables (temperature, wind, rainfall, etc.) are typically generated on a 10–50-kilometre spatial grid for medium- and extended-range ensembles. These values are not directly comparable to the measurements recorded at specific locations since they represent the average for the area covered by a grid-box. Statistical post-processing can substantially improve the forecasts by tailoring (down-scaling) the grid-box forecasts to smaller areas or individual sites, and also accounting for the finite ensemble size (Hemri et al. 2014). Re-forecasts are useful for calibrating medium-range forecasts as well as for the sub-seasonal range, and re-forecast datasets are increasingly becoming part of the medium-range forecast configurations. This can be especially important for severe weather forecasting by providing information about how the model performed for severe events in the re-forecast period.

The Extreme Forecast Index (EFI) was developed at ECMWF to highlight potential anomalous weather events, by comparing the real-time forecast to the re-forecast model climate distribution (Lalaurette 2003).

Beyond a few days ahead it is no longer possible to predict the day-to-day changes in the weather at specific locations. However, by considering the average conditions over a period of time, it is possible to give skilful forecasts for longer lead times (Buizza and Leutbecher 2015). Sub-seasonal forecasts typically predict average conditions for each week of the

coming month. At this longer forecast range, the prediction of changes in large-scale weather patterns is important, for example, to give an early indication of the onset of heat waves or cold spells.

In summer 2015, a heat wave affected large parts of Europe—temperature records were broken in many places, including in Germany, France and Spain. Early indications of widespread warmer than normal conditions during the first week of July can be seen in the ECMWF forecast from 16 June (Fig. 8.3). The signal becomes noticeably stronger in the forecast
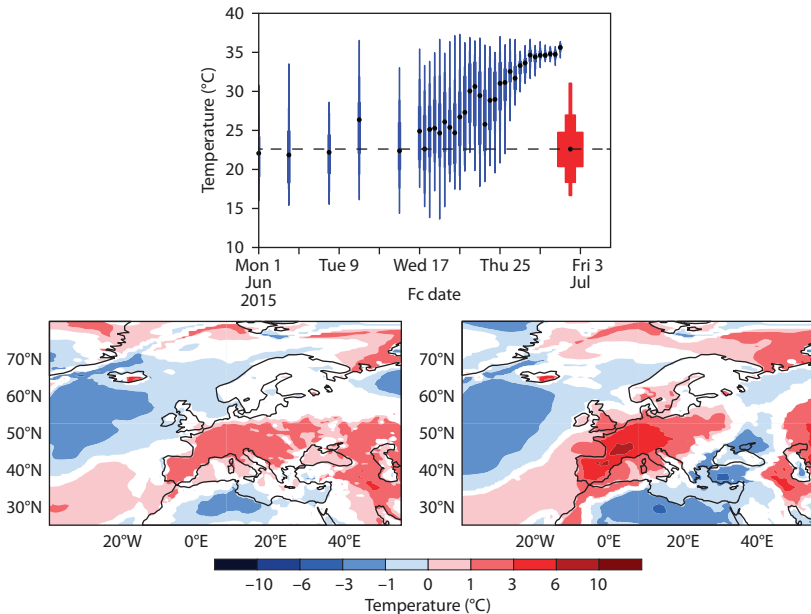

Fig. 8.3 ECMWF forecasts for the heat wave over Europe in July 2015. Lower panel shows the 2-metre temperature anomaly forecasts for the 7-day period 29 June to 5 July initialised on 18 June (left) and 22 June (right). Areas where the forecast distribution is significantly different from climatology are shaded. Upper panel shows the evolution of the ensemble forecasts for the temperature in Paris at 12 UTC on 1 July; the dates on the horizontal axis indicate the start time of each forecast. The box-and-whisker plots show the 1st, 10th, 25th, 75th, 90th and 99th percentile of the forecast, while black dot shows the median of the distribution. The temperature distribution of the model climate (generated from reforecasts for late June and early July for the last 20 years) is shown in red (the dotted line highlights the climate median). Magnusson et al. 2015.

from 22 June. While these successive forecasts give a good indication of the general situation, it is also interesting to consider the forecast for a specific time and location. The top panel of Fig. 8.3 shows a sequence of forecasts for the temperature in Paris at 12 UTC on 1 July (this was the second warmest day on record for Paris, with temperatures reaching to almost 40°C later in the afternoon). Each forecast is represented as a box-and-whisker plot that summarises the information in the ensemble. The forecasts from early June indicate a range of possible outcomes, similar to the model climate distribution (shown in red), with no clear signal for temperature to be warmer or cooler than normal. This is not surprising: as noted in the previous paragraphs, we should not expect skill in predicting the precise temperature at a specific place and time out to a month ahead. By mid-June there is a noticeable shift in the forecasts: while not certain, the ensemble indicates that high temperatures are much more likely to be above normal than below normal, and there is a significant chance that the temperature in Paris could reach more than 30°C. The risk of extreme temperatures increases in consecutive forecasts, and by 26 June the outcome is almost certain (see Magnusson et al. 2015 for more details).

This example shows how ensemble weather forecasts can be used to guide decision making for weather-dependent activities in the days and weeks ahead. For the coming days, detailed information can be obtained on the weather at a particular place and time. Looking further ahead, these details are less predictable, but the forecasts can give an indication of the likely general weather situation, and what alternative scenarios may be. While the weather is not always predictable out to a month ahead, there are particular situations where the predictability is enhanced—on such occasions the ensemble will show higher confidence and a more limited range of alternatives. In all cases, the ensemble approach provides important information about possible alternative scenarios, and enables users to make appropriate decisions, taking account of the confidence and risks quantified by the ensemble.

References

Baldwin, M. P., & Dunkerton, T. J. (2001). Stratospheric harbingers of anomalous weather regimes. *Science, 294*, 581–584.

Berner, J., Shutts, G. J., Leutbecher, M., & Palmer, T. N. (2009). A spectral stochastic kinetic backscatter scheme and its impact on flow-dependent predictability in the ECMWF ensemble prediction system. *J. Atmos. Sci., 66*, 603–626.

Bishop, C. H., Etherton, B. J., & Majumdar, S. J. (2001). Adaptive sampling with the ensemble transform Kalman filter. Part I: theoretical aspects. *Monthly Weather Review, 129*, 420–436.

Bougeault, P., et al. (2010). The THORPEX Interactive Grand Global Ensemble (TIGGE). *Bulletin of the American Meteorological Society, 91*, 1059–1072.

Bowler, N. E., Arribas, A., Mylne, K. R., Robertson, K. B., & Beare, S. E. (2008). The MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society, 134*, 703–722. https://doi.org/10.1002/qj.234.

Bowler, N. E., Arribas, A., Beare, S. E., Mylne, K. R., & Shutts, G. J. (2009). The local ETKF and SKEB: Upgrades to the MOGREPS short-range ensemble prediction system. *Quarterly Journal of the Royal Meteorological Society, 135*, 767–776. https://doi.org/10.1002/qj.394.

Buizza, R., & Leutbecher, M. (2015). The Forecast skill horizon. *Quarterly Journal of the Royal Meteorological Society.* https://doi.org/10.1002/qj.2619.

Buizza, R., & Palmer, T. N. (1995). The singular vector structure of the atmospheric general circulation. *Journal of the Atmospheric Sciences, 52*, 1434–1456.

Buizza, R., Miller, M., & Palmer, T. N. (1999). Stochastic representation of model uncertainties in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society, 125*, 2887–2908.

Buizza, R., Leutbecher, M., & Isaksen, L. (2008). Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quarterly Journal of the Royal Meteorological Society, 134*, 2051–2066.

Cassou, C. (2008). Intraseasonal interaction between the Madden–Julian Oscillation and the North Atlantic Oscillation. *Nature, 455*, 523–527. https://doi.org/10.1038/nature07286.

Charron, M., Pellerin, G., Spacek, L., Houtekamer, P. L., Gagnon, N., Mitchell, H. L., et al. (2010). Toward random sampling of model error in the Canadian ensemble prediction system. *Monthly Weather Review, 138*, 1877–1901. https://doi.org/10.1175/2009MWR3187.1.

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. https://doi.org/10.1002/qj.2411.

Hemri, S., Scheuerer, M., Pappenberger, F., Bogner, K., & Haiden, T. (2014). Trends in the predictive performance of raw ensemble weather forecasts. *Geophysical Research Letters, 41*, 9197–9205. https://doi.org/10.1002/2014GL062472.

Houtekamer, P. L., & Mitchell, H. L. (2005). Ensemble Kalman filtering. *Quarterly Journal of the Royal Meteorological Society, 131*, 3269–3289. https://doi.org/10.1256/qj.05.135.

Houtekamer, P. L., Mitchell, H. L., & Deng, X. (2009). Model error representation in an operational ensemble Kalman filter. *Monthly Weather Review, 137*, 2126–2143.

Houtekamer, P. L., Deng, X., Mitchell, H. L., Baek, S.-J., & Gagnon, N. (2014). Higher resolution in an operational ensemble Kalman filter. *Monthly Weather Review, 142*, 1143–1162. https://doi.org/10.1175/MWR-D-13-00138.1.

Jeong, J.-H., Linderholm, H. W., Woo, S.-H., Folland, C. K., Kim, B.-M., Kim, S.-J., et al. (2013). Impacts of snow initialization on subseasonal forecasts of surface air temperature for the cold season. *Journal of Climate, 26*, 1956–1972. https://doi.org/10.1175/JCLI-D-12-00159.1.

Koster, R. D., et al. (2010). Contribution of land surface initialization to subseasonal forecast skill: First results from a multi-model experiment. *Geophysical Research Letters, 37*, L02402. https://doi.org/10.1029/2009GL041677.

Lalaurette, F. (2003). Early detection of abnormal weather conditions using a probabilistic extreme forecast index. *Quarterly Journal of the Royal Meteorological Society, 129*, 3037–3057. https://doi.org/10.1256/qj.02.152.

Leutbecher, M., & Palmer, T. N. (2008). Ensemble forecasting. *Journal of Computational Physics, 227*, 3515–3539.

Leutbecher, M., Lock, S.-J., Ollinaho, P., Lang, S. T. K., Balsamo, G., Bechtold, P., et al. (2017). Stochastic representations of model uncertainties at ECMWF: State of the art and future vision. *Quarterly Journal of the Royal Meteorological Society.* https://doi.org/10.1002/qj.3094.

Lin, H., Brunet, G., & Derome, J. (2009). An observed connection between the North Atlantic oscillation and the Madden–Julian oscillation. *Journal of Climate, 22*, 364–380. https://doi.org/10.1175/2008JCLI2515.1.

Magnusson, L., Thorpe, A., Buizza, R., Rabier, F., & Nicolau, J. (2015). Predicting this year's European heat wave. ECMWF. *Newsletter, 145*, 4–5.

Palmer, T. N., & Richardson, D. S. (2014). Decisions, Decisions...! ECMWF. *Newsletter, 141*, 12–13.

Shutts, G. J. (2005). A kinetic energy backscatter algorithm for use in ensemble prediction systems. *Quarterly Journal of the Royal Meteorological Society, 131*, 3079–3102.

Swinbank, R., Kyouda, M., Buchanan, P., Froude, L., Hamill, T. M., Hewson, T. D., et al. (2016). The TIGGE project and its achievements. *Bulletin of the American Meteorological Society, 97*, 49–67.

Toth, Z., & Kalnay, E. (1993). Ensemble forecasting at NMC: The generation of perturbations. *Bulletin of the American Meteorological Society, 74*, 2317–2330.

Toth, Z., & Kalnay, E. (1997). Ensemble Forecasting at NCEP and the breeding method. *Monthly Weather Review, 125*, 3297–3319.

Vitart, F. (2014). Evolution of ECMWF sub-seasonal forecast skill scores. *Quarterly Journal of the Royal Meteorological Society, 140*, 1889–1899. https://doi.org/10.1002/qj.2256.

Vitart, F., Robertson, A. W., & Anderson, D. L. T. (2012). Subseasonal to seasonal prediction project: Bridging the gap between weather and climate. *WMO Bulletin, 61*, 23–28.

Vitart, F., Ardilouze, C., Bonet, A., Brookshaw, A., Chen, M., Codorean, C., et al. (2017). The Subseasonal to Seasonal (S2S) Prediction Project Database. *Bulletin of the American Meteorological Society, 98*, 163–173. https://doi.org/10.1175/BAMS-D-16-0017.1.

Wei, M., Toth, Z., Wobus, R., & Zhu, Y. (2008). Initial perturbations based on the ensemble transform (ET) technique in the NCEP global operational forecast system. *Tellus A, 60*, 62–79.

Yamaguchi, M., & Majumdar, S. J. (2010). Using TIGGE data to diagnose initial perturbations and their growth for tropical cyclone ensemble forecasts. *Monthly Weather Review, 138*, 3634–3655.