Project: 607193 - UERRA

Seventh Framework Programme
Theme 6 [SPACE]

Project: 607193 UERRA

Full project title:
Uncertainties in **E**nsembles of **R**egional **R**e-**A**nalyses

Deliverable D2.1

Development of ensemble variational data capability and report demonstrating ensemble uncertainty products

WP no:	2
WP leader:	MO
Lead beneficiary for deliverable :	MO
Name of author/contributors:	Peter Jermey, Jemma Davie, Sana Mahmood, Richard Renshaw
Nature:	Report
Dissemination level:	PU
Deliverable month:	21
Submission date: September 29, 2015	Version nr: 1

Development of Ensemble Reanalysis System and Products

Peter Jermey, Jemma Davie, Sana Mahmood and Richard Renshaw

November 11, 2015

1 Introduction

The Met Office has completed and performed a technical test on the regional ensemble of reanalyses system that will be used to produce an ensemble of forty year regional reanalyses for the European domain.

The system will be used to produce a number of realisations of forty years of a wide range of atmospheric variables, available hourly. The set of realisations can also be used to derive mean and spread (uncertainty) for each variable.

A deterministic reanalysis will also be produced for the satellite era (1978-2017). This will make use of the ensemble information to improve assimilation of observations into the system. The combined dataset will provide users with a high resolution dataset, along with a lower resolution uncertainty estimate, which will be useful for a wide variety of applications. The deterministic and ensemble reanalyses are the principal contribution of the Met Office to the Uncertainties in Ensembles of Regional Reanalyses (UERRA) project, [Unden et al., 2014].

This document is a report on the technical test of the system which has been used to demonstrate the ensemble products that will be available from the production. This report details observations used in Section 2 and the configuration of the ensemble and deterministic systems are described in Sections 3 and 4, respectively. The test system is described in Section 5, with results from the short technical test in Section 6. Examples of products from the ensemble system are presented in Section 7, and the report is summarised in Section 8.

2 Observations

Following the achievement of the Met Office pilot reanalysis for the European Reanalysis and Observations for Monitoring (EURO4M, [Klein Tank et al., 2014]), a similar set of observations will be assimilated for the UERRA reanalyses. This set of observations comprises those processed for assimilation in the European Centre for Medium Range Weather Forecasting (ECMWF) reanalyses, [Dee et al., 2011], with additional Met Office ground global positioning (GPS) observations, as detailed in Tables 1 and 2.

Development of gridded precipitation data is also in progress. It is hoped that gridded precipitation data from the European Climate Assessment and Dataset (ECA&D, [Klein Tank et al., 2002]) will be assimilated into the deterministic reanalysis, but due to time constraints, it is unlikely this will be included in the ensemble reanalysis. It is expected that precipitation assimilation will have greater positive impact at higher resolutions and so it is expected to be of lower significance in the ensemble reanalysis than the deterministic reanalysis.

Operationally at the Met Office, observation data is quality controlled before being assimilated. For the EURO4M reanalysis, a fixed 'blacklist' of observations was used for the two year period. A fixed blacklist is less appropriate for a forty year period and so a new flexible observation automatic monitoring system will be used for surface, sonde and aircraft observations. This system has been initially developed to process surface station data and other observation types are currently under development.

3 Ensemble System

The regional reanalysis ensemble system is based on the Met Office operational forecast model (Unified Model, UM, [Davies et al., 2005]). The system is configured using the new dynamical core (Even Newer Dynamics for General atmospheric modeling of the environment, ENDGame, [Wood et al., 2014]) which has been operational since July 2014. ENDGame has been developed to improve coupling of the forecast model to physical parameterisations and to improve the handling of Rossby waves.

As in the Met Office operational global forecasting system, the system features six-hour cycling. For each member, a forecast is carried out from a combination of the analysis increment, provided by the data assimilation system, and the previous forecast (background) from T-3 to, at least T+9. The reanalysis fields are output at T+0 (the centre of the six hour assimilation window) and the background for the next cycle at T+3. The system is shown in Figure 1. As shown, the ensemble members are independent of one another and do not require recentring around a control.

The ensemble system uses 4DVAR data assimilation, [Rawlins et al., 2007], to draw the system close to observational data. 4DVAR produces an increment to the background which estimates the optimum state of the atmosphere, given the background and observations across the six hour window T-3 to T+3, taking into account three-dimensional position and time of each observation. As is common practise, the assimilation will take place at half the horizontal resolution of the model.

Observation	Subtypes	Variables	Dates	Source
Land SYNOP	Land synoptic observations (LNDSYN), Meteorological airfield reports (METARS), Mobile synoptic observations (MOBSYN), Australian point estimates of sea-level pressure (PAOBS), American surface bogus observations (BOGUS)	Surface pressure, temperature, humidity, wind	1978-2017	ECMWF
SHIP	Ship synoptic observations (SHPSYN)	Surface pressure, wind, temperature, humidity	1978-2017	ECMWF
Buoy	Buoy	Surface pressure, wind, temperature	1979-2017	ECMWF
Sondes	Radiosondes (TEMP), Wind profilers (WINPRO), Dropsondes (DROPSOND), Wind only sondes (PILOT)	Upper-air wind, temperature, humidity	1978-2017	ECMWF
Aircraft	Aircraft Meteorological Data Relay (AMDARs), Air Report (AIREPs), Tropospheric Airborne Meteorological (TAMDAR), Met Office bogus observations (TCBOGUS), American upper air bogus observations (BOGUS)	Flight-level temperature, wind	1978-2017	ECMWF
AIRS	Advanced Infared Sounder (AIRS)	AIRS	2003-2017	ECMWF
ATOVS	Global Operational Vertical Sounder (ATOVSG)	HIRS/AMSU radiances	1998-2017	ECMWF

Table 1: List of observations (Part 1 of 2).

Observation	Subtypes	Variables	Dates	Source
GPSRO	GPSRO	Bending Angle	2006-2017	ECMWF
Ground GPS	Integrated Water vapour (GPSIWV)	Zenith Delay	1999-2017	MO
IASI	Global (IASIG) Local (IASIL)	IASI	2007-2017	ECMWF
satwinds/AMVs	ESA Cloud Motion Winds (ESACMW), ESA High Resolution Wave mode (ESAHRWVW), Geostationary Operational Environmental (GOESBUFR), Imaging Spectroradiometer (MODIS), Meteosat 2nd Generation satellite winds (MSGWINDS), Satellite Observations (SATOB)	wind wind wind wind wind	1988-2017	MO
scatwinds	SeaWinds, WindSat, Advanced Scatterometer (ASCAT), High Resolution Advanced Scatterometer (ASCATHR), ESA High Resolution Wave mode (ESAHRWVW), ESA Scatterometer (ESAUWI)	wind	1992-2017	ECMWF
SEVIRIclear	Meteosat 2nd Generation clear sky radiances (MSGCSR), Meteosat 2nd Generation radiances (MSGRAD)	clear sky	1982-2017	ECMWF
SSMIS	Special Sensor Microwave	water vapour, ocean wind speed	2009-2017	ECMWF
TOVS	Operational Vertical Sounder	radiances	1979-2002	ECMWF

Table 2: List of observations (Part 2 of2).

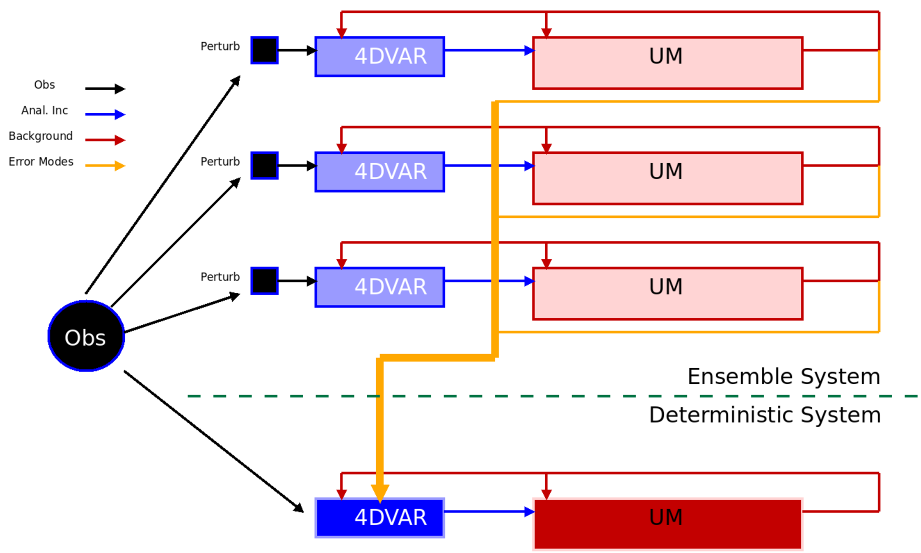

Figure 1: System design - ensemble and deterministic system. Only three members of the ensemble system are shown.

4DVAR makes use of error covariances to allow for differences in uncertainty in the observations and background. These are dependent on the background error covariance matrix (**B**), as well as the observation error covariance matrix. For the ensemble system **B** is a static smoothed parameterisation, trained by a long time-series of regional atmospheric states. This parameterisation is referred to as a static background error covariance matrix, since it does not vary with each cycle.

So that the ensemble of reanalyses represents the uncertainty in the reanalysis system, uncertain aspects of the system are perturbed so that each member receives a different realisation of them, within the uncertainty estimation.

Each assimilated observation is quality controlled and bias corrected. The observations are then individually perturbed. A value for each realisation of an observation is randomly chosen from within an estimated confidence interval, assuming a Gaussian distribution.

It can be shown that 4DVAR increments are drawn from the same distribution as UM model error, assuming that reanalyses are drawn from the same distribution as the truth. Therefore different realisations of the model are produced by perturbing the model with a randomly selected archived increment in each six hour window within the forecast model. This follows a system for global ensembles used at the Met Office for research into model error [Piccolo and Cullen, 2015].

The upper boundary of the model is a rigid lid, which is not a source of uncertainty and, therefore, features no perturbation. For each member a separate surface analysis defines the lower land boundary using the Met Office surface analysis scheme (SURF, [Candy, 2014]). Uncer-

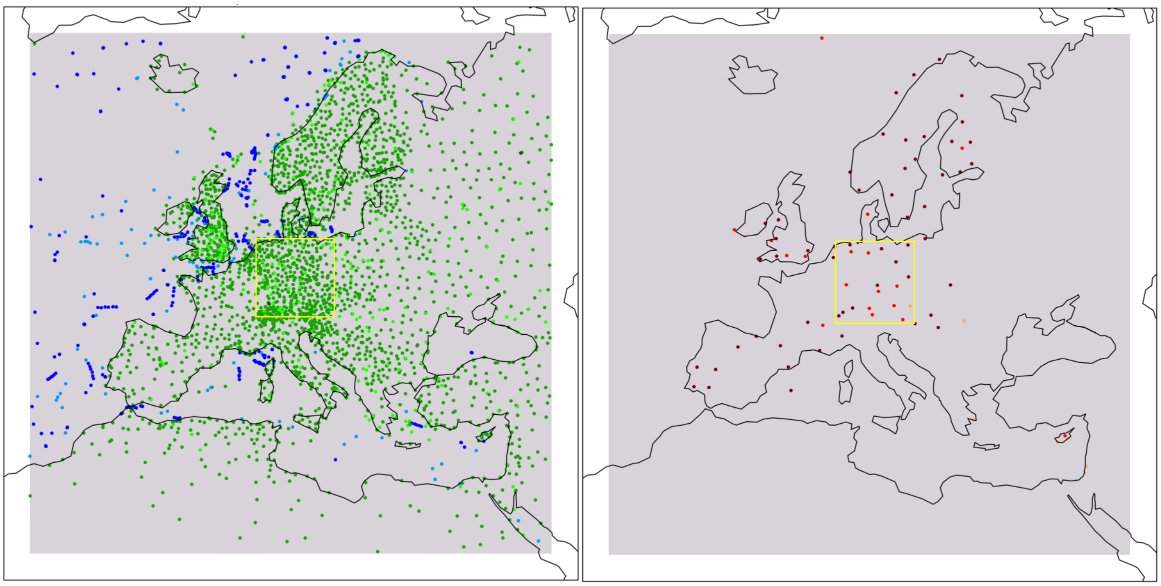

Figure 2: UERRA domain (grey) with German domain for verification (yellow box). The left and right hand plots show positions of surface and sonde observations, respectively within the assimilation window for 06Z on the 9th March 2010. Light and dark green indicate SFERICS and SYNOP observations, respectively. Light and dark blue indicate ships and buoys, respectively. Orange, light red and dark red indicate PILOT, TEMP and WINPRO sondes.

tainty in the lower ocean boundary and the lateral boundary conditions is represented by use of external ensemble datasets, version 2 of the Hadley Centre Sea Ice & Sea Surface Temperature data set (HadISST2, successor to HadISST, [Rayner et al., 2003]) and ECMWF's Reanalysis 5 (ERA5, [Dee, 2014]), respectively. Both these external ensembles have only ten members and so some regional reanalysis members will share boundary condition realisations. If every member of a regional ensemble has a different boundary condition realisation then the ensemble spread will be strongly influenced by uncertainty from the boundaries. However, if a subset of the ensemble shares boundary condition realisations then their influence on the ensemble spread will be decreased in favour of other sources of uncertainty, i.e. observations and model. Therefore using fewer boundary condition realisations than there are ensemble members is not expected to be detrimental to the ensemble.

The ensemble reanalysis is planned to be produced for the satellite era over the EUR-11 domain of the Coordinated Regional climate Downscaling Experiment (CORDEX, [Jacob et al., 2014]), as shown in Figure 2, with twenty members at half horizontal resolution, i.e. 0.22 degree grid boxes (approximately 24km), on 63 levels to approximately 40km.

4 Deterministic System

The deterministic system will be very similar to the ensemble system, except that it will feature a single member and no perturbations to represent uncertainty. Instead of 4DVAR, hybrid-4DVAR will be used, which couples the deterministic system to the ensemble system.

In hybrid-4DVAR the background error covariance will be a linear combination of the static matrix and a matrix calculated from the ensemble of (re)forecasts from the same cycle. In this way the ensemble will provide the deterministic system with error characteristics dependent on the synoptic situation, which improves the accuracy of the reanalysis. The coupling of the two systems is shown in Figure 1. Hybrid-4DVAR has been used operationally for global data assimilation at the Met Office since July 2011, [Clayton et al., 2013]. The regional variant will be used for the first time with the deterministic reanalysis.

As with the ensemble reanalysis, deterministic reanalysis will be run for the satellite era over the CORDEX EU-11 domain, but at full horizontal resolution, i.e. 0.11 degree grid boxes (approximately 12km), again on 63 levels to approximately 40km.

5 Technical Test System

A short period technical test (15th February 2010 to 15th March 2010) has been carried out using a test version of the ensemble reanalysis system. The test system is close to the intended production system, but features some differences, which are detailed here.

Perturbations of the model and observations follow those of the intended production system, but a regional variant of the surface analysis system was not available for the test. Instead reconfigured global operational fields from 2010 were used as the lower land boundary. On each cycle the difference in these fields from the previous cycle was calculated. These different fields were randomly perturbed and added to the original fields to produce a different realisation of the lower land boundary for each member. A bug in this scheme has been identified, but is not expected to have significantly impacted the test results.

Also at the time of testing, ERA5 data was not available for the lateral boundary conditions. Instead the ERA5 team kindly gave access to early test data, which also defined the period chosen for testing.

Since the technical test was the first time the system was run, no data was available to train the static background covariance matrix. In its place the EURO4M background covariance was used. This is appropriate for the domain, since the EURO4M domain is similar to CORDEX-11, but was developed for a different dynamical core ('New Dynamics').

Component	EURO4M	Test	Ensemble	Deterministic
Domain	EURO4M	EU-11	EU-11	EU-11
Levels	70 to 80km	70 to 80km	63 to 40km	63 to 40km
Forecast Model	UM (New Dynamics)	UM (ENDGame)	UM (ENDGame)	UM (ENDGame)
Resolution	0.11deg, 12km	0.22deg, 24km	0.22deg, 24km	0.11deg, 12km
Assimilation	4DVAR	4DVAR	4DVAR	hybrid 4DVAR
Resolution	0.22deg, 24km	0.44deg, 48km	0.44deg, 48km	0.22deg, 24km
B	Static (EURO4M)	Static (EURO4M)	Static (UERRA)	Static (UERRA) & Ensemble
Members	1	10	20	1
Land boundary	Reconfigured Global	Reconfigured Global	Regional SURF	Regional SURF
LBCs	ERA-Interim	ERA5 (test)	ERA5	ERA5

Table 3: Comparison of reanalysis systems.

At the time of testing, porting to ECWMF's machine was not yet complete and so the test was carried out at the Met Office, processing all observations from the Met Office archive, instead of mostly using ECMWF's observations for reanalysis. To improve efficiency for the test, a ten member ensemble was run. The test was carried out at the resolution planned for production, but on 70 levels to 80km. The differences between the ensemble production, deterministic production and ensemble test systems are shown in Table 3.

6 Validation

The results presented here are from the system used for technical testing only. It is expected that actual products will improve on this due to the differences between production and test systems, described in Table 3.

An ensemble is well structured if

1. each ensemble member is equally likely,
2. the ensemble mean has greater accuracy than a deterministic control at the same resolution
3. the ensemble spread matches the error in the mean of the ensemble and,
4. the probability of an event in the ensembles matches that in the observed data [Gofa, 2010].

In this section, the test system and a control (featuring no perturbations, one member and at the same resolution as the ensemble members) are assessed against surface observations and radiosondes. To ensure independence of the observation data from the reanalyses, six hour forecasts are assessed.

Two variables are assessed - 2m temperature (T2) and 500hPa (500H) geopotential height. It is problematic to compare error of the ensemble mean with ensemble spread using point observation data since the data will include potentially large representivity differences. The error of the ensemble mean is calculated by comparing point information with grid box averages, whereas the spread is calculated only by using grid box averages. To ameliorate this effect global ensembles are often assessed using the relatively large-scale variable 500hPa geopotential height. However large-scale variables are known to not be as well represented by regional systems as global systems, [Jones and Macpherson, 2010]. Therefore 2m temperature is also assessed here. The systems are compared across the whole domain and a smaller German domain which features a dense observation network, as shown in Figure 2.

Figure 3 shows rank histograms for 2m temperature and 500hPa height. In both cases and both domains, approximately 50% of the observations occur within the spread of the ensemble (a perfect score would be 80%), indicating the system is under spread or biased. For 500hPa height, there is clearly a bias whereby observations are more likely to be above the ensemble range than below it. For 2m temperature across the whole domain, observations are also more likely to be warmer than the ensemble range than cooler. However in the smaller German domain, with a dense observation network, this bias is reduced, indicating the apparent bias in temperature may actually be related to representivity differences between the model and the observations. For temperature, the probability of each observation rank is reasonably constant within the ensemble range, suggesting that each member is equally likely in this variable.

NB due to a bug in the verification process, one of the ensemble members is missing from the rank histograms.

Figure 4 shows time series of RMSE for the ensemble mean and a deterministic control for 2m temperature and 500hPa height over both domains. For 2m temperature, there is a sudden increase in error around the middle of the test period. The cause of this is unclear at the time of writing. The ensemble mean error is only smaller than the control error during this period, indicating that the ensemble perturbations are not sufficiently estimating the uncertainty in the

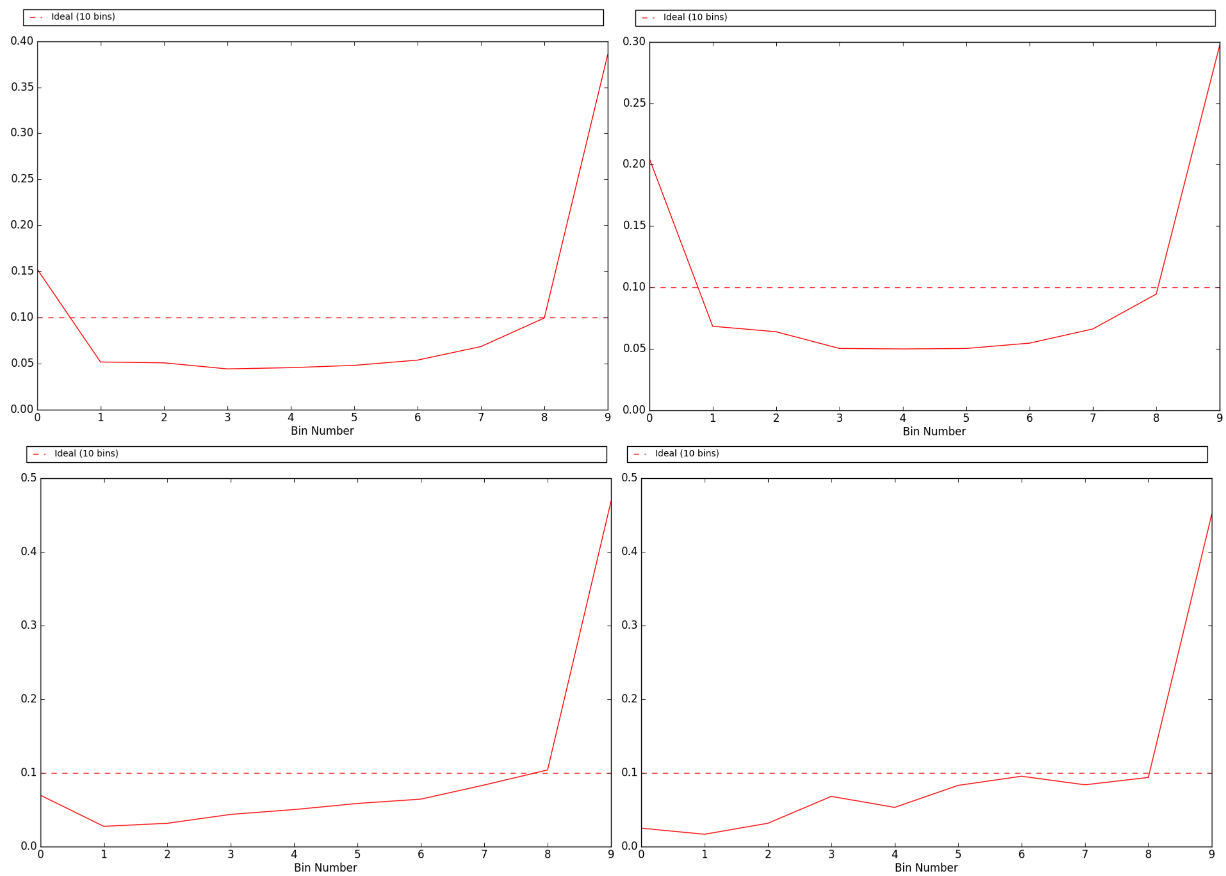

Figure 3: UERRA test ensemble rank histograms at six hours forecast time. The top row shows 2m temperature and the bottom row shows 500hPa height. The left hand and right hand columns show the full domain and the smaller German domain, respectively.

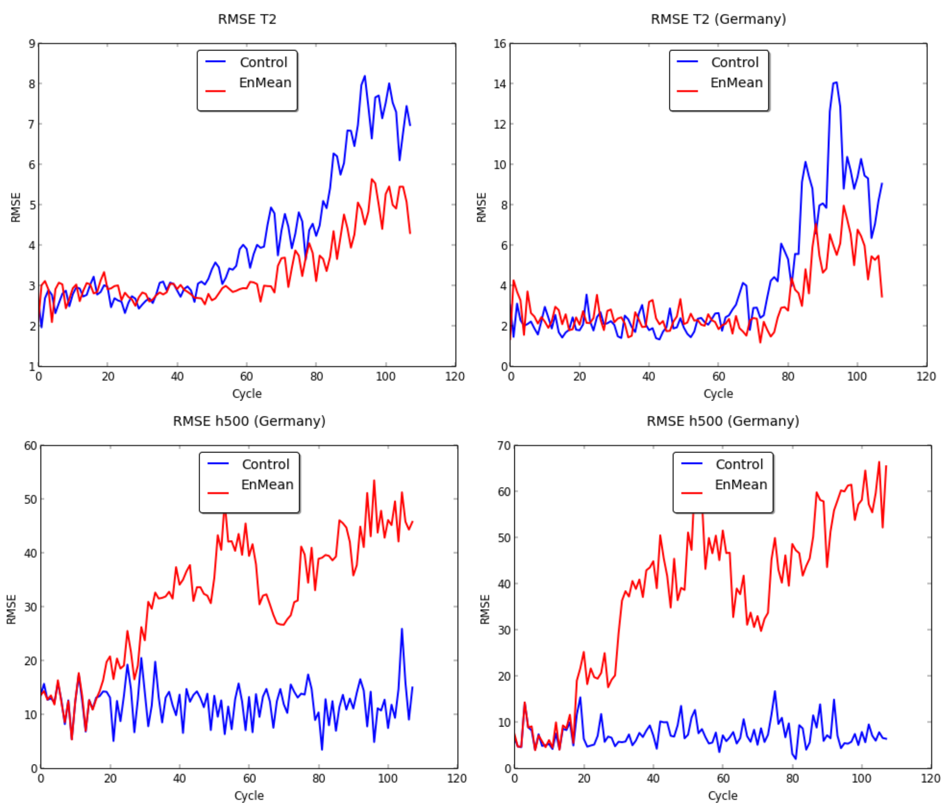

Figure 4: UERRA ensemble mean and deterministic control RMSE at six hours forecast time. The top row shows 2m temperature and the bottom row shows 500hPa height. The left hand and right hand columns show the full domain and the smaller German domain, respectively.

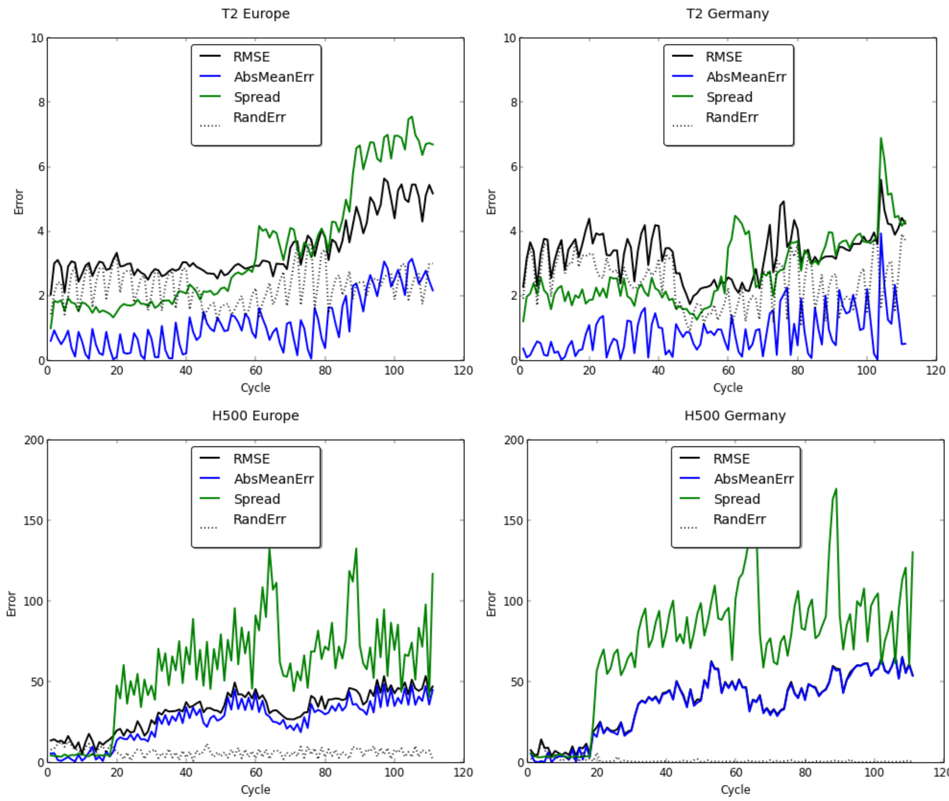

Figure 5: UERRA test spread and mean RMSE, mean bias (AbsMeanError) and random error (RMSE-AbsMeanError) at six hours forecast time. The top row shows 2m temperature and the bottom row shows 500hPa height. The left hand and right hand columns show the full domain and the smaller German domain, respectively.

deterministic system. Results for 500hPa height are similar in that the ensemble mean error is only smaller than that of the control during the period of increased error. For this variable the increased error is only shown in the ensemble mean. This is thought to be caused by a uncertain position of the centre of a low pressure system. The centre varies with ensemble member, creating too large area of low pressure in the ensemble mean, which perhaps indicates the spread in the low pressure centre is too high.

Figure 5 compares the ensemble mean error with the spread for both domains. The spread of the 2m temperature across the whole domain increases during the test period and has a reasonably large difference with the RMSE of the ensemble mean. However the spread in the German domain is a much closer match to the RMSE, indicating that much of the difference calculated in the whole domain may be due to representivity differences between the observations and the model. In both the whole domain and the German domain, the 500hPa spread behaves similarly. After a short period of the spread being much smaller than the mean error, it becomes much larger than the mean error. Throughout the test period the spread retains a relationship to the

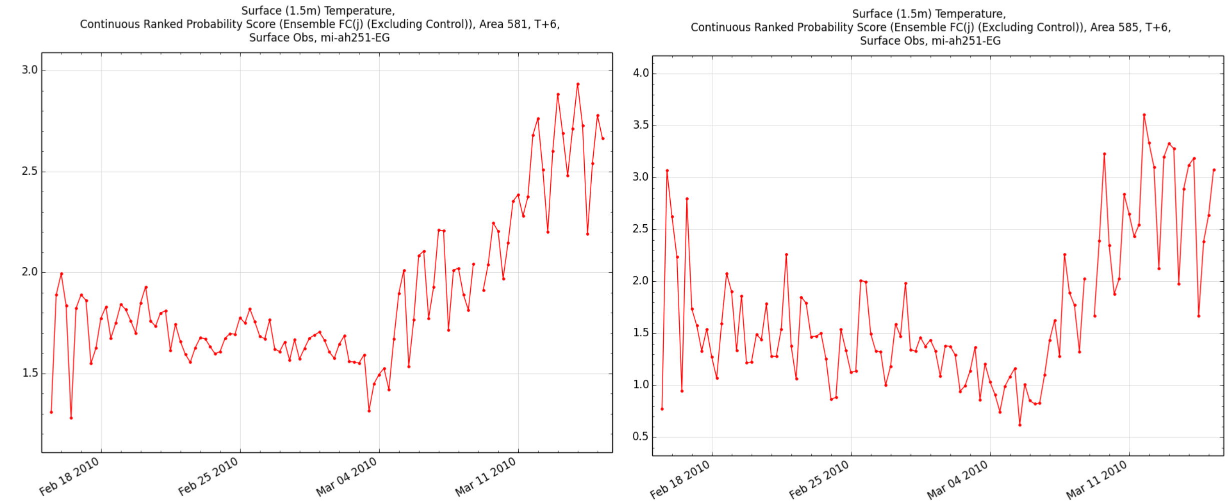

Figure 6: UERRA test CRPS for 2m temperature at six hours forecast time. The left hand and right hand columns show the full domain and the smaller German domain, respectively.

mean error, but both are large and the mean error is dominated by bias. This demonstrates that global reanalysis systems are more appropriate than regional systems for representing large-scale variables, such as pressure and geopotential height. As with the control RMSE, large errors are seen in these metrics which needs further investigation.

Figure 6 shows the continuous rank probability score (CRPS) for 2m temperature for the entire domain and for the smaller German domain. The CRPS is a measure of the difference between the forecast and observation distributions. The plots show reasonable consistency between the ensemble and observation distributions, but, as with the other metrics, shows increasing difference towards the end of the period.

7 Products

The ensemble produces a number of realisations of atmospheric states on an hourly basis. This means that the mean and spread (uncertainty) are available for a wide range of consistent atmospheric fields that are complete across the European domain.

Figure 7 shows the mean of the ensemble mean and the spread at reanalysis times for 2m temperature for the test period. This shows a much smaller spread in the boundary than in the internal domain, indicating that the regional reanalysis system produces a more realistic uncertainty estimate for small scale surface variables than the global parent ensemble. The largest regions of spread are over Eastern Ukraine, Central Europe and the Western Scottish Highlands.

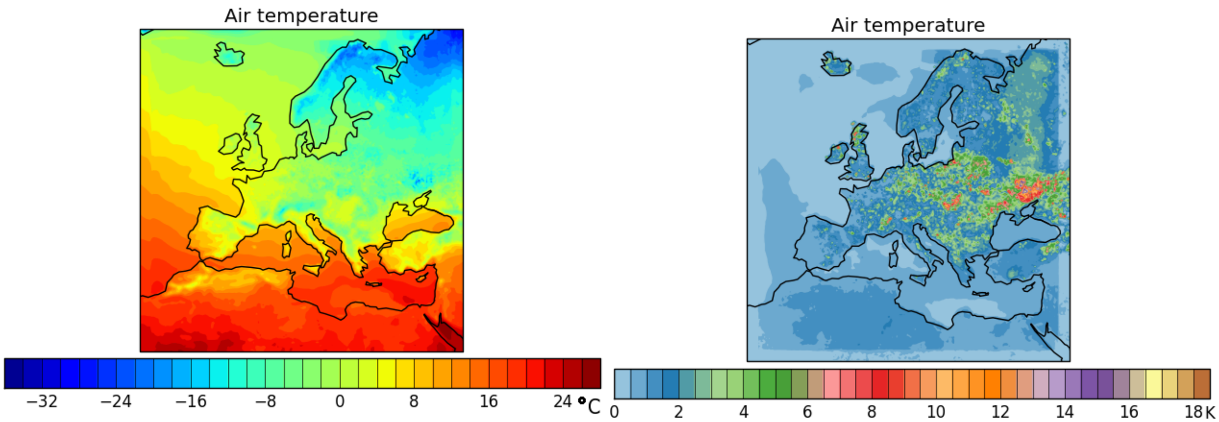

Figure 7: UERRA test mean of 2m temperature ensemble mean (left hand side) and spread (right hand side).

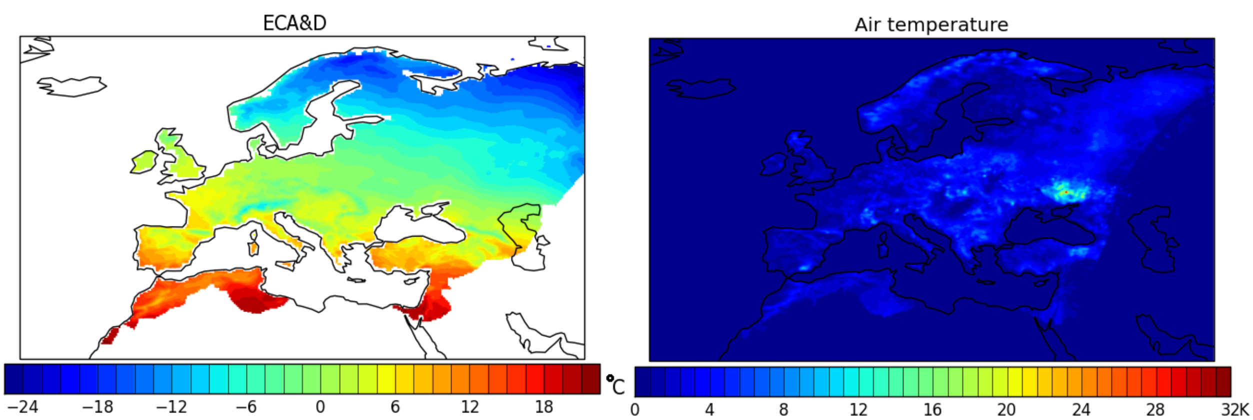

Figure 8: Test mean of 2m temperature ECA&D (left hand side) and absolute difference with UERRA mean (re)forecast (right hand side).

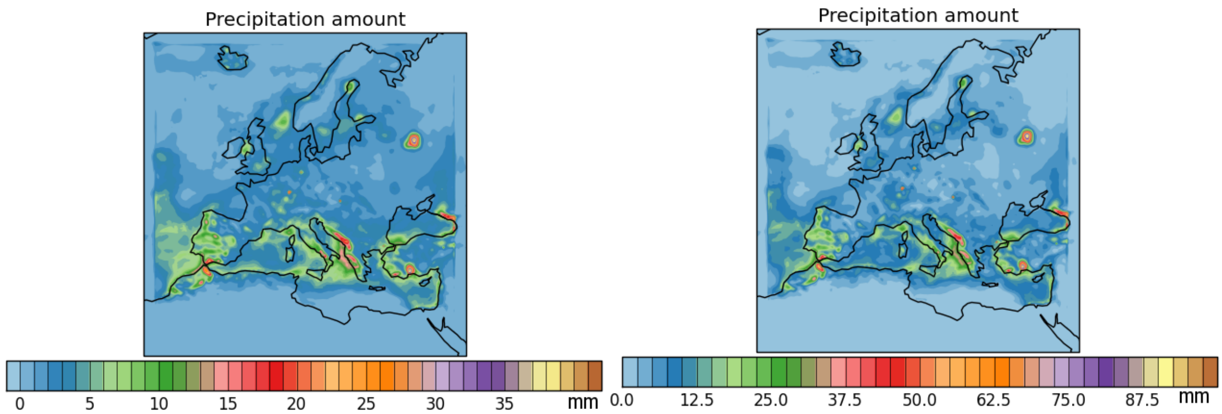

Figure 9: UERRA test mean of 24h precipitation ensemble mean (left hand side) and spread (right hand side).

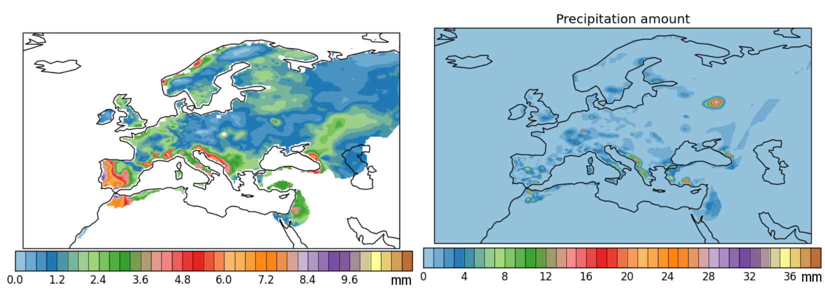

Figure 10: Test mean of 24h precipitation ECA&D (left hand side) and absolute difference with UERRA ensemble mean (right hand side).

Figure 8 shows the mean of 2m temperature from the ECA&D, [Klein Tank et al., 2002], for the test period, and its absolute difference with the ensemble six hour (re)forecast. The six hour (re)forecast is used in place of the reanalysis as a comparison to ensure independence, since ECA&D shares a subset of observations with the reanalysis. Figure 8 shows the largest difference between the ECA&D and the ensemble is over Eastern Ukraine, with relatively large differences over Central Europe and Western Scotland. This demonstrates that the ensemble produces largest spread in areas of greatest uncertainty, as expected.

Figure 9 shows the mean daily precipitation of the ensemble mean and spread. The largest values of mean precipitation are on the Adriatic Coast and in Central Russia. The largest mean spread are in similar positions, indicating the large uncertainties associated with representation of precipitation. The Russian precipitation is mostly due to the record-breaking snowstorm of 20th - 21st February [NOAA, 2010].

Figure 10 shows the mean daily ECA&D precipitation, for the test period, and its difference with that of the ensemble mean. Precipitation observations have not been assimilated into the

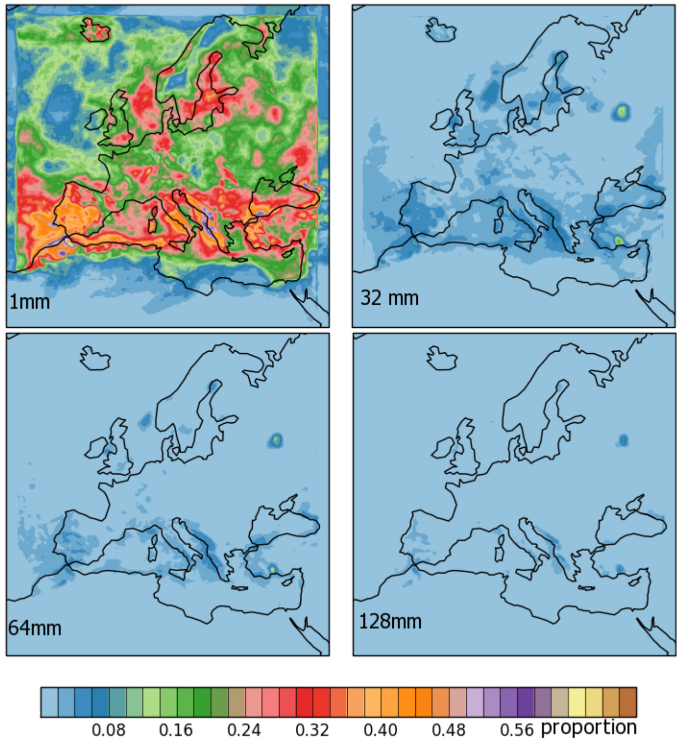

Figure 11: UERRA test probability of 1mm, 32mm, 64mm, and 128mm for 24h in the test period.

ensemble reanalysis and therefore ECA&D is independent. This demonstrates that areas with large (low) spread correspond reasonably well to areas of greatest (least) difference between the ensemble mean and the ECA&D data.

As well as the mean and spread, ensemble members allow user to estimate probability of a chosen event. Figure 11 shows the probability of precipitation events of greater than 1mm, 32mm, 64mm and 128mm in twenty-four hours for the test period. This shows that the probability of extreme precipitation events is greatest over Central Russia and the Adriatic Coast. Figure 12 shows the probability at reanalysis times that wind speed is between 3ms-1 and 15ms-1, the operating range of wind turbines, at 00Z, 06Z, 12Z and 18Z. This highlights the best locations for positioning turbines during the test period.

8 Summary & Future Work

A test version of the ensemble reanalysis system has been constructed and tested for a short test period. There is error growth within both the control and ensemble which may be due to

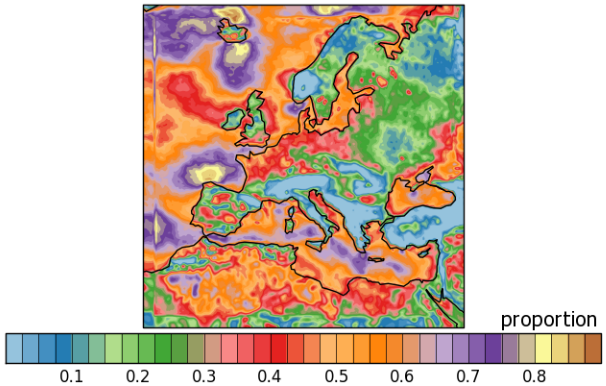

Figure 12: UERRA test probability of wind speed is between 3ms-1 and 15ms-1 in the test period.

non-production quality LBCs, but requires further investigation. These results also suggest that ensemble perturbations may require tuning. Nevertheless, the results suggest that the production ensemble will be well structured.

The production system will use a background error covariance calculated for the appropriate domain and dynamical core, which is expected to reduce the error in the ensemble mean. The production system will also feature a regional surface analysis scheme. This may have a small impact on the accuracy of the ensemble mean, but it will also allow a more appropriate set of realisations of the lower land boundary to be used. Use of production quality ERA5 lateral boundary conditions may also improve the quality of the ensemble reanalysis.

Although it is expected the production will feature 20 members at 24km, it may be that 10 members will be used if the larger ensemble is shown not to significantly improve the quality.

The ensemble reanalysis will produce mean and spread (uncertainty) across a wide range of physically consistent atmospheric variables which will be valuable to a wide range of users. Additionally probabilities of particular weather events may be easily calculated using the ensemble members.

With thanks to Hans Hersbach and ECMWF for early access to ERA5 test data.

References

[Candy, 2014] Candy, B. (2014). Assimilation of satellite data for the land suface. *ECMWF Annual Seminar Proceedings: Use of Satellite Observations in Numerical Weather Prediction.*

[Clayton et al., 2013] Clayton, A. M., Lorenc, A. C., and Barker, D. M. (2013). Operational implementation of a hybrid ensemble/4d-var global data assimilation system at the met office. *Q.J.R. Meteorol. Soc*, 139(675):1445–1461.

[Davies et al., 2005] Davies, T., Cullen, M. J. P., Malcolm, A. J., Mawson, M. H., Staniforth, A., White, A. A., and Wood, N. (2005). A new dynamical core for the met office's global and regional modelling of the atmosphere. *Q.J.R. Meteorol. Soc*, 131(608):1759–1782.

[Dee et al., 2011] Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hlm, E. V., Isaksen, L., Kllberg, P., Khler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thpaut, J.-N., and Vitart, F. (2011). The era-interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc*, 137(656):553–597.

[Dee, 2014] Dee, R. (2014). Use of satellite data in reanalysis. *ECMWF Annual Seminar Proceedings: Use of Satellite Observations in Numerical Weather Prediction.*

[Gofa, 2010] Gofa, H. (2010). Guidelines for verification of ensemble forecasts. *Hellenic National Meteorological Service.*

[Jacob et al., 2014] Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O., Bouwer, L., Braun, A., Colette, A., Dqu, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Krner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P. (2014). Euro-cordex: new high-resolution climate change projections for european impact research. *Regional Environmental Change*, 14(2):563–578.

[Jones and Macpherson, 2010] Jones, C. and Macpherson, B. (2010). The benefit of continuous data assimilation in a regional NWP model. *Met Office Tech Rep*, 566.

[Klein Tank et al., 2014] Klein Tank, A. M. G. et al. (2014). European reanalysis and observations for monitoring.

[Klein Tank et al., 2002] Klein Tank, A. M. G., Wijngaard, J. B., Knnen, G. P., Bhm, R., Demare, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Mller-Westermeier, G., Tzanakou, M., Szalai, S., Plsdttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberfeld, R., van Engelen, A. F. V., Forland, E., Mietus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplova, E., Cegnar, T., Antonio Lpez, J., Dahlstrm, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L. V., and Petrovic, P. (2002). Daily dataset of 20th-century surface air temperature and precipitation series for the european climate assessment. *Int. J. Climatol.*, 22(12):1441–1453.

[NOAA, 2010] NOAA (2010). State of the climate: Global hazards for february 2010. *NOAA National Centers for Environmental Information.* http://www.ncdc.noaa.gov/sotc/hazards/201002.

[Piccolo and Cullen, 2015] Piccolo, C. and Cullen, M. (2015). data assimilation using a unified representation of model error. *Mon. Wea. Review.* accepted.

[Rawlins et al., 2007] Rawlins, F., Ballard, S. P., Bovis, K. J., Clayton, A. M., Li, D., Inverarity, G. W., Lorenc, A. C., and Payne, T. J. (2007). The met office global four-dimensional variational data assimilation scheme. *Q.J.R. Meteorol. Soc*, 133(623):347–362.

[Rayner et al., 2003] Rayner, N. A., Parker, D. E., Horton, E. B., Folland, C. K., Alexander, L. V., Rowell, D. P., Kent, E. C., and Kaplan, A. (2003). Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century. *Journal of Geophysical Research: Atmospheres*, 108(D14).

[Unden et al., 2014] Unden, P. et al. (2014). Uncertainties in ensembles of regional reananlyses.

[Wood et al., 2014] Wood, N., Staniforth, A., White, A., Allen, T., Diamantakis, M., Gross, M., Melvin, T., Smith, C., Vosper, S., Zerroukat, M., and Thuburn, J. (2014). An inherently mass-conserving semi-implicit semi-lagrangian discretization of the deep-atmosphere global non-hydrostatic equations. *Q.J.R. Meteorol. Soc*, 140(682):1505–1520.