Seventh Framework Programme
Theme 6 [SPACE]

Project: 607193 UERRA

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

Deliverable D2.2

Report of observations and datasets assembled for the ensemble-based variational assimilation

WP no:	2
WP leader:	MO
Lead beneficiary for deliverable :	MO
Name of ~~author~~/contributors:	Richard Renshaw, Peter Jermey, Amy Doherty, Jemma Davie, Chris Burrows
Nature:	Report
Dissemination level:	PU
Deliverable month:	24
Submission date: 5 April, 2016	Version nr: 1

Report of observations and datasets assembled for the ensemble-based variational assimilation

Richard Renshaw, Peter Jermey, Amy Doherty, Jemma Davie and Chris Burrows

April 5, 2016

1 Introduction

The Met Office will be running a regional European reanalysis from 1978 to the present day, as part of the EU FP7 funded UERRA project (Uncertainties in Ensembles of Regional Reanalyses, [Unden et al., 2014]).

The reanalysis aims to make extensive use of available observations (Section 2). Most of the observations will come from ECMWF archives, assembled for ERA reanalyses and, from 2003 onwards, for operational NWP. These observations will require careful quality control and selective use in order to avoid degrading the quality of the reanalysis. They will be assimilated into a regional model which has limited vertical extent, covering much of the stratosphere but not extending to the mesosphere. This imposes restrictions on the use of the highest level observations (Section 3).

The regional reanalysis will rely on ERA global reanalyses to provide lateral boundary conditions (Section 4). Surface boundary conditions over the ocean will come from global datasets of daily SST (Sea Surface Temperature) and sea-ice cover (Section 5).

These comprise the major inputs to the reanalysis. This report documents these inputs and their use in the Met Office regional reanalysis.

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], [Jermey and Renshaw, 2016]), 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 reprocessed ground global positioning (GPS) observations, as detailed in Tables 1 and 2.

Reprocessed observations of zenith delay from ground GPS stations (1996-2014) are being provided by Rosa Pacione of ASI/CGS (Agenzia Spaziale Italiana/Centro di Geodesia Spaziale), Italy. This data has been processed in a consistent way over the time period. This avoids the sudden changes in quality and bias that can affect operational data streams. Operational data will be used for 2015 onwards, with bias corrections calculated and applied every month. This is a data source not yet used at ECMWF and so operational data will be taken from Met Office archives.

The TOVS instruments are particularly important in extending the satellite record back to 1978. The ability to assimilate TOVS radiances in Met Office systems has not been maintained and so work was required to add back this capability. Level 1 electronic counts are read from ECMWF MARS (Meteorological Archive and Retrieval System), then processed to radiances and the MSU radiances mapped to the HIRS grid using the NWP-SAF AAPP package http://nwpsaf.eu/site/software/aapp/. These are then assimilated in the same way as other satellite radiometers. The SSU (Stratospheric Sounding Unit) component of TOVS will not be used. Its use is technically complicated and its three channels are sensitive to temperatures high above the top layer of the Met Office regional model.

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.

3 Observation selection

Operationally at the Met Office, observation data is quality controlled before being assimilated. For the EURO4M reanalysis, a static reject list of observation sites was used for the two year period. A fixed reject list is less appropriate for a forty year period and so a new flexible observation automatic monitoring system has been developed for surface, sonde and aircraft observations. This runs monthly. It compares observations against reanalysis model fields (6-hour forecasts). Where observation and model differences are consistently large, those sites are rejected. The system also calculates bias corrections for surface pressure and for aircraft and sonde temperature.

For satellite data, usage will be guided by the experience of the ERA reanalyses. As part of ERA-40 and ERA-Interim, the ECMWF reanalysis team have constructed lists of dates when individual instruments, and even individual channels, are and are not reliable. The Met Office

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	1996-2014 2015-2017	ASI/CGS 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	ECMWF
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	MSU and HIRS radiances	1979-2002	ECMWF

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

reanalysis will use these in its own selection. It will also follow ECMWF in using VarBC (Variational Bias Correction) to apply bias correction to satellite radiances [Dee and Uppala, 2009]. VarBC analyses bias corrections as part of the assimilation process. In this way the biases change with time so as to fit drifts in instrument bias. The method is well tested for global models but has yet to be used in a regional model. It will be seen whether the scheme will allow for slow change in instrument bias without responding to the fast change in model biases expected in regional models on synoptic timescales.

The Met Office regional reanalysis also differs from global models in having a lower model top. The topmost model level will be at 40km height. Some satellite data are sensitive to atmospheric temperature above 40km, for instance high-peaking radiometer channels and GPS radio occultation data. The implications for observation use have been examined in two studies, [Burrows, 2015] for GPS-RO and Doherty (*pers. comm.*) for satellite sounding.

The reanalysis assimilates bending angle from GPS radio occultation data. Bending angle is a function of refractivity, calculated along the signal path through the Abel transform. The path ranges in height from the top of the atmosphere to the point at which it is nearest Earth (the 'impact height'). Burrows calculated bending angle from model vertical profiles for a range of model top levels. Values of refractivity used above the model top were taken by extrapolating from the topmost level. The bending angles were compared against values calculated from full height model profiles with no truncation. Repeating this for a large set of model profiles gave an estimate of the size of the error in the bending angle from a truncated model top. Figure 1 shows the truncation error for a model top of 40km as a function of impact height. The error increases exponentially with impact height. At 32km the size of the error is of the same magnitude as typical errors in the observed bending angle. For that reason GPS data will only be assimilated in the reanalysis at 30km and below.

Acceptable tolerance	Highest usable impact height
0.01	39km
0.001	32km
0.0001	24km

Table 3: Table showing the highest acceptable impact heights given a range of fractional truncation error tolerances for a lid height of 40km (cf. Figure 1).

It is possible to perform a similar investigation for the use of high-peaking satellite sounding channels. Radiances were calculated using RTTOV [Saunders et al., 1999] from a set of reference profiles. The top part of the profiles were then replaced by extrapolated values and radiances calculated from these. Figure 2 shows the root mean square differences between the two sets of radiances for HIRS. For the lower-peaking and window channels, differences are small. For the stratospheric channels, HIRS channel 1 in particular, differences are large, greater than the assumed instrument error. On this basis, HIRS channels 1, 2 and 3 are rejected. High-peaking

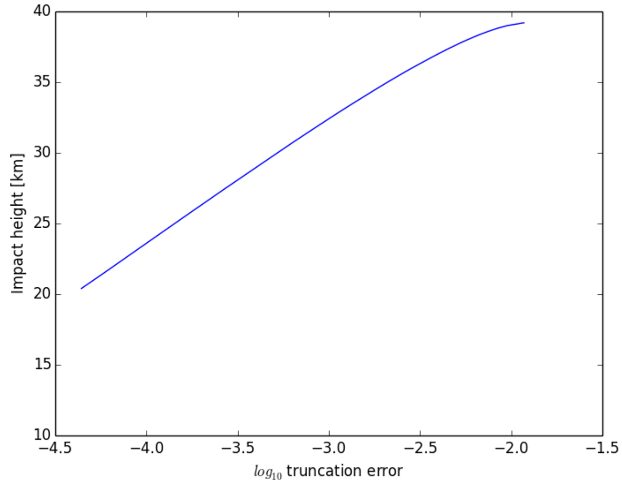

Figure 1: Logarithm (base 10) of the fractional bending angle truncation error due to a 40km lid compared to an 80km lid for a range of impact heights.

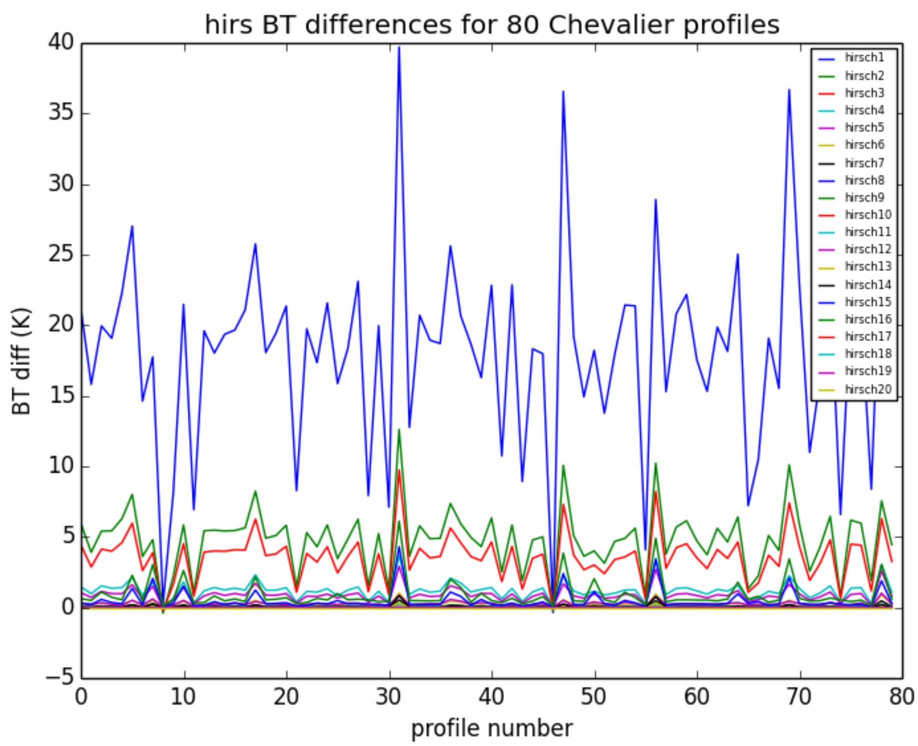

Figure 2: Differences between radiances (Kelvin) calculated from full and truncated/extrapolated vertical profiles for HIRS channels.

channels for other instruments have been rejected on a similar basis.

4 Lateral Boundary Conditions

Regional models are driven by global models around the boundaries of the model domain - lateral boundary conditions (LBCs). For EURO4M, the Met Office reanalysis made use of ERA-Interim. Global fields were available 6-hourly at a resolution of approximately 80km and on 60 vertical levels. For UERRA, the Met Office reanalysis will be an ensemble system. The spread in the ensemble comes from a spread in the various inputs (observations, surface fields, LBCs) as well as in adding terms to represent model error. If the spread in the various input values is a good representation of the uncertainty in those variables, then it is to be hoped that the spread in the reanalysis outputs will be a good representation of reanalysis errors.

ERA-Interim is soon to be superceded by ERA5, 31km resolution and 137 levels. ERA-Interim was purely a deterministic reanalysis and so provides no spread in LBCs. ERA5 includes an ensemble of 10 members which is used in 4DVar to provide an improved representation of model background errors, compared to the usual climatological estimate. ERA5 is an attractive alternative to ERA-Interim as a source of LBCs. ERA5 production has started (January 2016) and is expected to continue until the end of 2017. UERRA milestone MS5 requires the Met Office reanalysis to be complete by end of June 2017. This would seem to exclude ERA5 as an option for the whole period 1978-present. It is not known how much difference the choice will make. This is to be tested on a short period.

5 Sea Surface Temperature and Sea Ice

The lower boundary over water is described by sea ice fraction (SI) and sea surface temperature (SST). For the majority of the period this data will be provided by the second version of the Hadley Centre Sea Ice and Sea Surface Temperature data set (HadISST2.1), [Kennedy, 2015]. The dataset includes a ten member ensemble of daily SST realisations derived from satellite and in-situ data gridded at 0.25°. These realisations are consistent with a single realisation of SI at 1° derived from satellite data and historical ice charts.

HadISST2.1 data is available from 1899 to 2010. For 2010 - present the lower boundary over water will be based on the National Centre for Ocean Forecasting (NCOF) Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) dataset, [Donlon et al., 2012]. This includes a single realisation of daily SST derived from satellite and in-situ data gridded at 0.05°. The dataset also includes the European Organisation for the Exploitation of Meterological Satellites (EUMETSAT) Ocean and Sea Ice Satellite Applications Facility (OSI-SAF, http://www.osi-saf.org/),

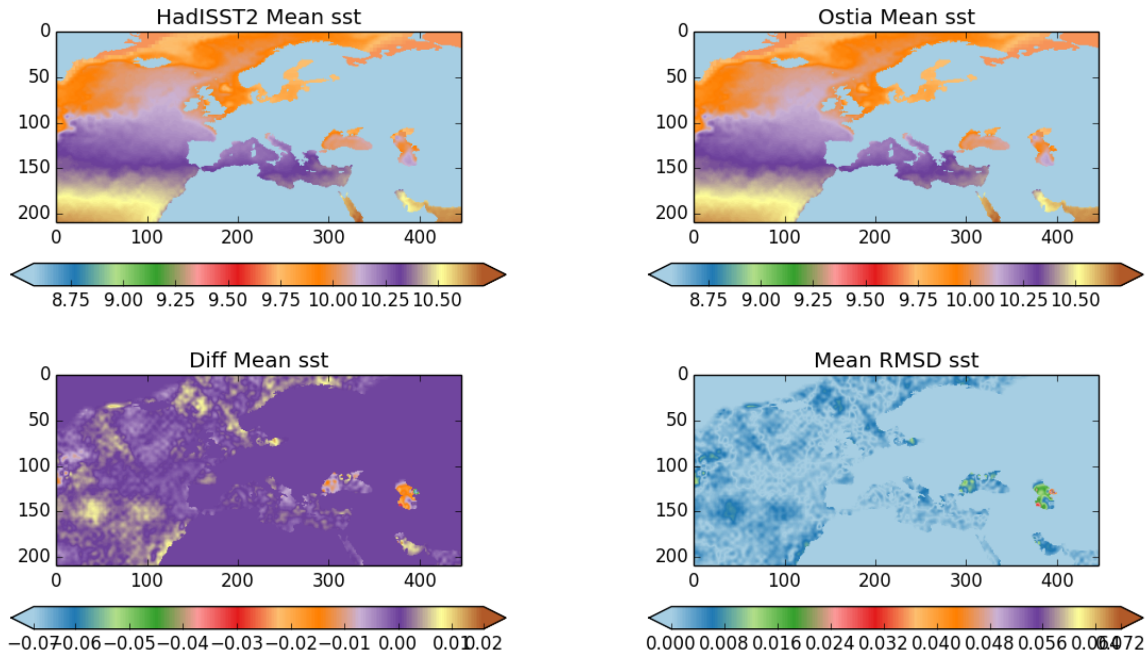

Figure 3: SST comparison of HadISST2.1 with OSTIA for February 2010. Top row - mean daily SST HadISST2.1 (L) and mean daily SST OSTIA (R). Bottom row - difference in monthly mean SST (L) and mean of RMS difference in daily SST (R).

which is based on satellite data at 10km resolution.

The differences between the two datasets are shown for SST and SI in Figures 3 and 4, respectively. As shown in Figure 3, the differences in SST are small: less than 0.1° K throughout the domain. The differences in the number of ice days are not so trivial, as shown in Figure 4. The difference is substantially reduced by mapping the OSI-SAF data onto the HadISST2.1 grid, but some locations have differences of several days in the month.
The lower boundary over water for the reanalysis is given in Table 4. The second period will begin in September 2010, the month with the least SI, to minimise the impact of the change. For the second period the OSTIA and OSI-SAF data will be mapped onto the HadISST2.1 grid. Ten realisations of the OSTIA SST will be obtained by adding a set of randomly selected error perturbations from HadISST2.1. The set of perturbations will be selected so that the month of the perturbations match the month of the OSTIA data to ensure perturbations are appropriate for the time of year. By applying perturbations from the earlier period, the spread of the SST will remain of a similar size throughout the production period.

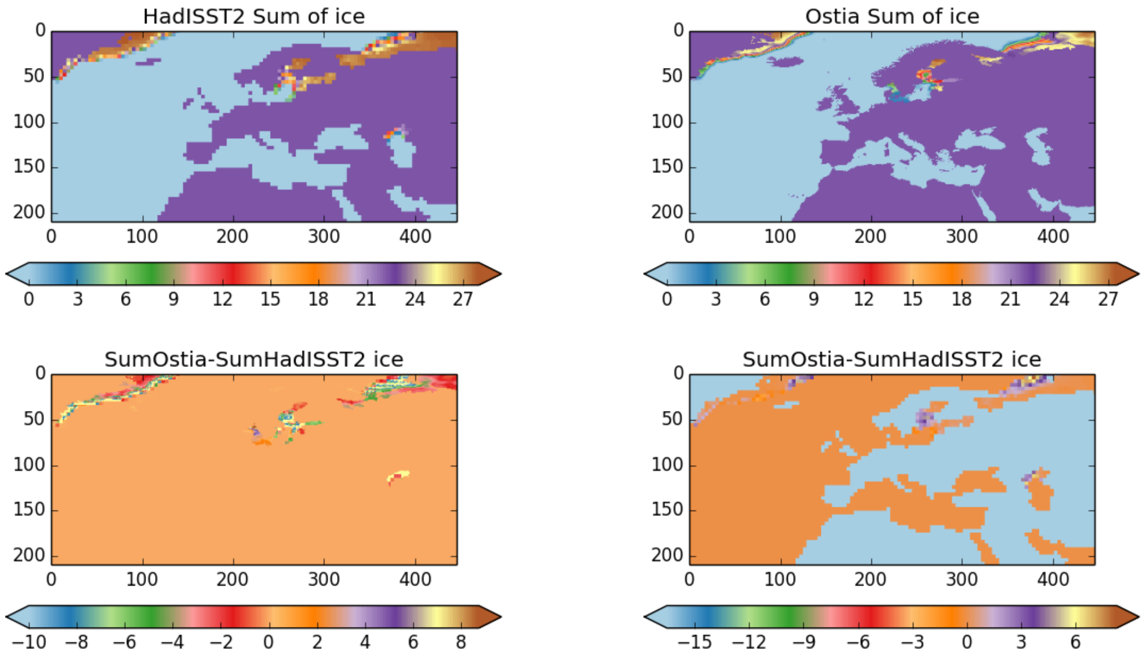

Figure 4: SI comparison of HadISST2.1 with OSI-SAF for February 2010. Top row - number of ice days HadISST2.1 (L) and number of ice days OSI-SAF (R). Bottom row - difference in number of ice days on native grids (L) and difference in number of ice days on HadISST2.1 grid (R).

Field	Ctrl 1978-2010	member $1 \leq i \leq 10$ 1978-2010	member $11 \leq i \leq 20$ 1978-2010	Ctrl 2010-pres.	member $1 \leq i \leq 10$ 2010-pres.	member $11 \leq i \leq 20$ 2010-pres.
SST	$\frac{\sum_{j=1}^{n} h_j^s}{n}$	h_i^s	h_{i-10}^s	$d(s^s)$	$d(s^s) + H_i^s$	$d(s^s) + H_{i-10}^s$
SI	h^i	h^i	h^i	$d(s^i)$	$d(s^i)$	$d(s^i)$

Table 4: Lower boundary over water for Met Office reanalysis where h_j^s and h^i are HadISST2.1 SST and SI, respectively and s^s and s^i are OSTIA SST and OSI-SAF SI, respectively. A mapping from the OSTIA grid to the HadISST2.1 grid is represented by $d()$. For 2010-present the SST fields are perturbed with ensemble perturbations H_i^s selected from a randomly chosen date from HadISST2.1. The month of the ensemble perturbations will match that of the mean field to ensure perturbations are appropriate for the season.

6 Summary

The Met Office reanalysis will use observations from the ECMWF archive, put together for ERA reanalyses and for ECMWF operational NWP. This will be augmented by reprocessed ground GPS station data, courtesy of ASI/CGS. For *in situ* observation types, data selection will be based on monthly reject lists produced by monitoring observation minus background values. For satellite observations, selection is based on ERA usage. Additionally, some satellite data is rejected because of the relatively low model top of the regional model. Satellite radiances will be bias corrected using VarBC. Regional VarBC is untested on long periods and so this will need careful monitoring.

An ensemble of lateral boundary conditions will be available from ERA5 but possibly not on the timescales required for UERRA. Work is planned to test the impact of ERA5 ensemble LBCs versus the single realisation available from ERA-Interim.

Ensemble values of SST and sea-ice will come from HadISST2 for most of the period of the reanalysis, to be replaced by OSTIA data from 2010. Ensemble values for OSTIA will be generated by adding perturbations from HadISST2.

With thanks to Rosa Pacione (Agenzia Spaziale Italiana/Centro di Geodesia Spaziale, Italy) and Gemma Halloran (Met Office) for their work in making available reprocessed ground-based GPS data.

References

[Burrows, 2015] Burrows, C. (2015). Potential use of bending angle data in assimilation systems with low model tops. *Satellite Applications Technical Memo*, 43. available on request from the Met Office.

[Dee and Uppala, 2009] Dee, D. P. and Uppala, S. M. (2009). Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. *Q.J.R. Meteorol. Soc*, 135(644):1830–1841.

[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.

[Donlon et al., 2012] Donlon, C. J., Martin, M., Stark, J., Roberts-Jones, J., Fiedler, E., and Wimmer, W. (2012). The operational sea surface temperature and sea ice analysis (OSTIA) system. *Remote Sensing of Environment*, 116:140–158.

[Jermey and Renshaw, 2016] Jermey, P. M. and Renshaw, R. J. (2016). Precipitation representation over a two-year period in regional reanalysis. *Q.J.R. Meteorol. Soc.* (currently online only).

[Kennedy, 2015] Kennedy, J. (2015). ERA-CLIM2 deliverable D3.15 HadISST2 updates. http://www.era-clim.eu/ERA-CLIM2/Products/ERA-CLIM2_D3.15.pdf.

[Klein Tank et al., 2014] Klein Tank, A. M. G. et al. (2014). European Reanalysis and Observations for Monitoring. http://www.euro4m.eu.

[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.

[Saunders et al., 1999] Saunders, R. W., Matricardi, M., and Brunel, P. (1999). An improved fast radiative transfer model for assimilation of satellite radiance observations. *Q.J.R. Meteorol. Soc*, 125:1407–1425.

[Unden et al., 2014] Unden, P. et al. (2014). Uncertainties in Ensembles of Regional Reanalyses. http://www.uerra.eu.