A new temporally flow-dependent EDA estimating background errors in the new Copernicus European Regional Re-Analysis (CERRA)

Adam El-Said[1], Pierre Brousseau[2], Martin Ridal[3], and Roger Randriamampianina[4]

[1]Meteo France
[2]METEO-FRANCE
[3]SMHI
[4]Met Norway

November 23, 2022

Abstract

A new augmented Ensemble of Data Assimilations (EDA) technique to estimate background error covariances (B-matrix) has been developed for the Copernicus European Regional Re-Analysis (CERRA). The B-matrix is modelled on a bi-Fourier limited area model. Background errors are assumed isotropic, homogeneous and non-separable. Linearised geostrophic and hydrostatic balances are incorporated as multivariate relationships, coupling vorticity and geopotential extended to mass-wind and specific humidity fields via the f-plane approximation. The B-matrix is estimated by a new 10-member CERRA-EDA system, temporally tethered to real-time meteorological situations. The EDA forecast differences comprise two main pools: seasonal and daily. The seasonal component is pre-prepared at reanalysis-resolution (5.5km). The new augmentation governs real-time mixture of winter and summer differences. The daily component is an 11km moving 2.5 day average. B-matrix re-estimation occurs every 2 days, with a fixed split of 80-20\% seasonal-daily. We consider a case study to illustrate the potential of CERRA-EDA to estimate weather regime change. The most influential factors are temporal evolution of spatial observation coverage, and varying the seasonal-daily split. Background error statistics, improvements in analysis and forecast skill scores and overall assimilation system performance are shown and discussed.

A new temporally flow-dependent EDA estimating background errors in the new Copernicus European Regional Re-Analysis (CERRA)

A. El-Said[1], P. Brousseau[1], M. Ridal[2], R. Randriamampianina[3]

[1]Metéo-France, CNRM-GMAP, Avenue Gaspard Coriolis, Toulouse Cedex, France
[2]Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
[3]Norwegian Meteorological Institute, Oslo, Norway

Key Points:

- Newly developed augmented temporally flow-dependent EDA system for estimating B matrix for European reanalysis
- EDA successfully estimates weather regime change but observation network and proportions of EDA sourcing are vital
- Demonstrated improvements in statistical profiles and forecast scores and overall data assimilation system performance of CERRA EDA system

Corresponding author: A. El-Said, `adam.el-said@meteo.fr`

Abstract

A new augmented Ensemble of Data Assimilations (EDA) technique to estimate background error covariances (B-matrix) has been developed for the Copernicus European Regional Re-Analysis (CERRA). The B-matrix is modelled on a bi-Fourier limited area model. Background errors are assumed isotropic, homogeneous and non-separable. Linearised geostrophic and hydrostatic balances are incorporated as multivariate relationships, coupling vorticity and geopotential extended to mass-wind and specific humidity fields via the f-plane approximation.

The B-matrix is estimated by a new 10-member CERRA-EDA system, temporally tethered to real-time meteorological situations. The EDA forecast differences comprise two main pools: seasonal and daily. The seasonal component is pre-prepared at reanalysis-resolution (5.5km). The new augmentation governs real-time mixture of winter and summer differences. The daily component is an 11km moving 2.5 day average. B-matrix re-estimation occurs every 2 days, with a fixed split of 80-20% seasonal-daily.

We consider a case study to illustrate the potential of CERRA-EDA to estimate weather regime change. The most influential factors are temporal evolution of spatial observation coverage, and varying the seasonal-daily split. Background error statistics, improvements in analysis and forecast skill scores and overall assimilation system performance are shown and discussed.

Plain Language Summary

The Copernicus European Regional Re-Analysis (CERRA) is a new retrospective analysis currently in production between 1984-2021. It is produced by a statistical analysis package that assimilates data from state-of-the-art observation systems into a regional weather model as well as a land-surface model. This is a cycling process, taking place every 3 hours.

The quality of the retrospective analysis hinges on correctly specifying observation and model background errors. Background error covariances are too numerous to explicitly compute and store in memory. We model the prevalent features of the errors and estimate them. The background errors are modeled to reasonably represent meteorological balances with respect to the Earth's rotation and vertical motion, and assume its distribution in 3D space is equal in all directions. The background errors are estimated by an Ensemble of Data Assimilations (EDA) performed cyclicly.

We demonstrate the capability of our augmented system to capture changes in weather regime. We illustrate this through a case study on two periods with different weather regimes. Improvements background error statistics are shown. Dependencies on the observation network and the mixture of different parts of the EDA are the main factors. We also show our EDA system was optimal for our intended use.

1 Introduction

A new Copernicus European Regional Re-Analysis (CERRA) provided by the Swedish Meteorological and Hydrological Institute (SMHI), Météo-France and Met Norway has been developed. CERRA uses Hirlam Aladin Research Mesoscale Operational Nwp In Europe (HARMONIE), a full NWP system, with the Aire Limitée Adaptation dynamique Développement InterNational (ALADIN) model for physics and dynamics. The reanalysis period spans 1984-2021 and the domain is Europe and North Africa. HARMONIE is currently in operational use by the HI-Res Limited Area Modelling (HIRLAM) consortium across 26 countries in Europe and northern Africa for short-range mesoscale NWP.

Retrospective analyses (re-analyses) provide temporally continuous and spatially coherent descriptions of atmospheric model states over a long time-frame. In effect, it is a re-writing of meteorological and climatological history, with better systems and another retrospective examination of the data. This atmospheric synopsis is achieved by combining a regional weather model with ever-improving observations in statistical optimality using data assimilation. Reanalysis practice is now commonplace for large weather centres since their advent 40 years ago for the First Global atmospheric research program Global Experiment (FGGE), (Bengtsson et al., 1982). The temporal span of global reanalysis ranges between 10 years (ERA-15 for example) to nearly 120 years (ERA20C, (Poli et al., 2016)). Reanalyses from world leading weather centres; Europe (CERA and ERA series), US (NCEP/NCARNCEP-DoE, CFSRR, MERRA), Japan (JRA55) and recently the Chinese Meteorological Administration (CRAI) demonstrate the increasing utility and routine production of reanalyses. It is also a valuable global-scale peer-validation exercise involving the most advanced weather centres in the world. The research community's use of reanalyses continues to grow worldwide. At the time of writing this paper Google scholar quotes in excess of 30,000 citations for NCEP/NCAR's 40-year reanalysis, (Bromwich & Wang, 2005), and 19,000 for ECMWF's outgoing ERA-Interim, (Dee et al., 2011).

Reanalyses for specialist purposes and more spatially confined domains come with the added benefit of increasing clarity for their intended context. It adds clarity to regions of varying orographies or specialist applications requiring different physics, with scope to trade smaller domains for higher resolution. For example there is; the Arctic Reanalysis System using NCAR's WRF-DA, (Bromwich et al., 2016), DWD's COSMO reanalysis of Europe, (Bollmeyer et al., 2015), a North American Regional Reanalysis (NARR), (Mesinger et al., 2006), and an Indian subcontinent reanalysis (IMDAA), (Mahmood et al., 2018). Upcoming specialist and regional reanalyses in Europe provided under the Copernicus Climate Change Services (C3S) framework include Copernicus Arctic Regional Reanalysis (CARRA) and CERRA-Land. Regional reanalyses benefit from higher model spatial resolutions and more observations per unit area. Regional or specialist reanalysis are also coupled with their global-counterparts via the use of global Lateral Boundary Conditions (LBCs). Most of the reanalysis efforts aforementioned use variational techniques for their DA systems with a variety of B-matrix estimation techniques. NCEP and NCAR, Europe (ECMWF ERA series) and Japan (JRA) mostly subscribe to variational methods for fulfilling their respective reanalyses with some minor exceptions (NOAA's 20C reanalysis, for example, uses the Kalman filter). These reason for these exceptions are usually a combination of; application context, computational resources and time-frame. The choices for B-matrix estimation vary more widely, for example, JRA55 estimated B using the NMC-method, while DWD's reanalysis, using their COSMO model, employs an *ensemble nudging* DA method with 21 members and a cycled Kalman filter analysis scheme.

The background error covariance matrix is, undeniably, a vital component of any DA system. It has a direct influence on the quality of the analysis, because it spreads information presented by the observations over the analysis increment. It has been illustrated that on a global scale, for any one analysis increment, the influence of information coming from; assimilated observations, and background, is split 15%, 85% respectively, (Cardinali et al., 2004). Two issues of practicality known about the B-matrix are: it is ill-conditioned and it is too large to represent explicitly. Ill-conditioning increases the number of iterations for convergence, if convergence even remains a possibility, and can seriously degrade analysis quality, (El-Said, 2015). The B-matrix is also too large to store even on the most advanced memory-abundant HPCs in the world since the size of the model state vector is of the order $\mathcal{O}10^{12}$ for most large-scale operational weather models. These two problems are simultaneously and conveniently dealt with by a change of variable, known as the control variable (or vector) transform (CVT). CVT allows implicit B-matrix specification, avoiding B inversion, explicit memory storage. A surrogate control vector is thus introduced, which conveniently preconditions, de-correlates and greatly reduces it to a more manageable size.

Specification of a surrogate control vector requires knowledge of its structure (modelling B), and commensurate estimation of the parameters governing its structure (estimating B). The large body of research behind modelling the structure of B in the last two decades strongly suggests that B is; *anisotropic*, *heterogeneous*, *multivariate* and *flow-dependent*, (Bannister, 2008). The HARMONIE NWP system represents the meteorological fields as spherical harmonics with triangular spectral truncation, inline with the ALADIN spectral model. The spatial correlation structure follows a spectral convolution which renders spatial structures as isotropic and homogeneous. Spectral convolutions are also vertically *non-separable*, meaning that broad horizontal correlations are deep, and less expansive horizontal correlations are isometrically shallow in 3D space. Non-separability facilitates correct, and necessary, specification of mass-wind correlations, (Bartello & Mitchell, 1992; Phillips, 1986). Balance is incorporated by imposing locally linearised geostrophic and hydrostatic balances. Vorticity and geopotential are related via the f-plane assumption, which is appropriate for the ALADIN domain. The Coriolis and Laplacian operator ratio in the f-plane equation is estimated by linear regression, (Berre, 2000). Multiple linear regressions are then used to estimate divergence, temperature and surface pressure, and specific humidity. This entire procedure comprises the balance operator. The extension to specific humidity was added to ALADIN, after adapting the original NCEP method, (Parrish et al., 1997; Derber & Bouttier, 1999). Further augmentations were also successfully implemented for the balance operator. Introducing a non-linear balance equation with the non-linear terms linearised around the background, and a quasi-geostrophic omega equation to aid in the enforcement of balance in regions with strong acceleration of curvature. Finally, modelling the structure of B is dependent on the context of DA system. For example, there exists a diffusion operator more suited to ocean modelling, (Mirouze & Weaver, 2010). For atmospheric data assimilation recursive filters allowed for inhomogeneous and anistropic background error structures in 2D space for 3DVAR initially, (Purser et al., 2003). The wavelet formulation also addressed heterogeneity and anisotropy allowing for better modelling of dynamical flow-dependent features such as the actual location of the tropopause for 3DVAR and 4DVAR, (Fisher, 2003).

As important as modelling B is, its quality depends on a commensurate estimation technique. Three popular methods for estimating the B-matrix are: the observation-background disentanglement technique, (Hollingsworth & Lönnberg, 1986), the NMC method, (Parrish & Derber, 1992), and statistics obtained from forecast differences originating from an Ensemble of Data Assimilation (EDA) analyses, (Fisher, 2003). The latter method enables time-varying background error if updated with sufficient temporal frequency. Theoretical grounding for EDA systems to represent background errors is demonstrated in (Zagar et al., 2005). The authors show that if a general forecast system is assumed to be weakly non-linear, doubling its covariance results in a doubling of the covariance of a 'true' unperturbed forecast produced using unperturbed initial conditions. EDAs can also provide insight into model uncertainties, (Palmer, 2001). The EDA technique performs several analyses, usually at lower resolution, where each member is obtained by perturbing observations and SST fields according to a Gaussian distribution with zero mean and a prescribed observation covariance matrix. This is what distinguishes each member in the ensemble. These analyses are then used to produce short forecasts and differences between these forecasts are computed to subsequently provide the statistical material needed for the B-matrix; the empirical data for the multivariate balance relationships and variances and length-scales required for the modelled covariance structure functions aforementioned.

The literature uses the term 'flow-dependence' to mean different things. So for the sake of clarity, it is important to distinguish between 'temporal flow-dependence' and 'dynamical flow-dependence'. They are not mutually exclusive but the literature has referred to both interchangeably as 'flow-dependence'. For example, temporal flow-dependence is what was meant in (Brousseau et al., 2012a, 2012b) and dynamical flow-dependence was the emphasis in (Bonavita et al., 2012) while (Isaksen et al., 2007) utilised both the temporal and dynamical while only referring to it as flow dependent. 'Flow-dependence' has also been used when referring to the ability of 4DVAR to propagate in the time dimension, and even when referring to non-linear balances, (Dee et al. (2011), section 2.1.3). Temporal flow-dependence, as we've termed it in this paper, simply means allowing for a degree of real-time tethering to meteorological situations as they arise. This is achieved via a 'live' or continuously cycling EDA running along side the assimilation system. Dynamical flow is the ability of the B-matrix to model dynamical features such as the tropopause position. An advantage of incorporating temporal flow-dependence is alleviating the need to re-configure the B-matrix during re-analysis production. Temporal flow-dependence is a relevant consideration for reanalyses, because of the time-span involved. Allowing background errors to vary in time provides much needed freedom in constraining either more or less-strongly to the observations. This tethering has been shown to provide considerable improvement in the ECMWF IFS, albeit with dynamical flow-dependence included, (Isaksen et al., 2007), and more relevantly to our context in the analogous Meteo-France AROME system, (Brousseau et al., 2012a). ALADIN and AROME are essentially the same package, with slightly different physics for increasingly higher-resolutions available in AROME.

Another reanalysis consideration is observation network coverage, which are susceptible to a plethora of significant changes over decadal time frames. Spatio-temporal correlations of observed quantities related to modelled prognostic variables change significantly, as does spatio-temporal observation density and distribution. If a static B-matrix was used, re-configuring the B-matrix during the re-analysis would become increasingly necessary. Real-time EDA systems enable synergy of temporal flow-dependent meteorological information, alleviating the need for B-matrix re-tuning (a daunting task in operational NWP), while also increasing analysis quality. The EDA-Interim report attests to the absence of the design consideration of changing observation networks, which they stated is needed in a reanalysis context, (Dee et al., 2011), for example. So having a real-time EDA which incorporates temporal flow to account for, observation networks, seasonal and daily variations, and meteorological phenomena is necessary in a reanalysis context. While this is all good news, one cannot over-look increased 'moving parts' introduced by an EDA system, increasing overall system run-time fragility. The EDA itself is not a complex, however, ensuring timely and seamless integration of all its moving parts with its partner DA system can be challenging.

A trade-off in overall system complexity and implementation is inevitable. Further design considerations include, computing parallelisation resource allocation, job queueing and timing, data retrieval and storage for forecast difference synthesis, all while maintaining close human-guided attentiveness to the smooth running of reanalysis production. EDA-specific considerations would therefore be; DA scheme, number of members, resolution, B-matrix update frequency, number of differences used for each update and the ratio of 'online' and 'offline' differences used. We have opted to use 3DVAR coupled with a fixed ratio of 80-20% seasonal-daily mixture of forecast differences to compute the B-matrix every 2 days. Further details of B-matrix design follows in section 3.

In this paper we investigate the effect of various combinations of EDA forecast difference mixing to compute the B-matrix. We do this in the context of a reanalysis system designed for a domain over Europe and North Africa with a total time-span of approximately 40 years. We then view the potential variability in: observation networks over time, resolution, temporal frequencies; seasonal and daily, and finally times of year where climate and weather phenomena are susceptible to different scales of meteorological variability, characterised by weather regimes. The first 4 potential variabilities have been investigated in non-reanalysis contexts and have been shown to produce sufficient variation to warrant B-matrix design consideration, (Brousseau et al., 2012a). In this paper we sufficiently demonstrate that capturing weather regime variability is a real possibility.

In section 2 we detail the HARMONIE-ALADIN system used to produce CERRA, with details on the observations in 2.2, and the data assimilation technique and its minimisation in section 2.1. Section 3 details the structure of the B-matrix, the EDA used to estimate it in section 3.1. The potential of background error statistical variability relevant to our regional reanalysis application is shown in 3.2, and the new forecast difference selection mechanisms introduced to address the change in weather regime, is discussed in section 3.3. Section 4 details our line of enquiry and briefly explains the weather regime paradigm. This sets the scene for the experiment design to demonstrate sufficient weather regime variability capture, which follows in section 4.1. The highlights of our investigation and illustrative examples are discussed in 4.2. Finally, we conclude our main findings in section 5.

2 The CERRA system

CERRA consists of 2 streams. The principal stream, which we call CERRA-DET, has 5.5km resolution and 4-minute time-step. It is accompanied by the second 10-member EDA stream, CERRA-EDA, running at 11km resolution and 10-minute time-step. The CERRA project uses the HARMONIE NWP system (code cycle version: cy40h1), with ALADIN physics, a 3DVAR system for the upper atmosphere, and Optimal Interpolation (OI) for the surface. The model has 106 vertical levels for both streams. The lateral boundary conditions (LBCs) come from ERA5. The ERA5 reanalysis has 31km resolution, 20-minute time-step, and an accompanying 63km 10-member EDA system with a 12-minute time-step, (Hersbach et al., 2020). Each of these streams were split into several temporally segmented streams taking advantage of the ECMWFs parallel HPC architecture. CERRA-DET's domain covers Europe, Northern Africa and South-Eastern parts of Greenland (The full domain can be seen pictorially (Wang & Randriamampianina, 2021), Figure 1). LBC forcing for CERRA-DET comes from ERA5 (31km). CERRA-EDA is forced by LBCs from ERA5's EDA stream, which has a resolution of at 63km at the equator.

ALADIN is coupled to the MESCAN-SURFEX surface analysis system, similar to the one used the predecessor project to CERRA, Uncertainties in Ensembles of Regional ReAnalyses (UERRA). It computes surface prognostic variables (surface and radiative temperatures, roughness length, albedo and emissivity) and fluxes (evaporation, sensible and latent heat fluxes and wind stress), while accounting for the proportion of different surface types (land, ocean, lakes and towns) that are projected in each model gridbox, (Masson et al., 2013). The MESCAN-SURFEX system, adopts it name from 'SURFace EXternalisée' (SURFEX) and the merger between SMHI's 'MEsoscale Surface analysis' (MESan), (Häggmark et al., 2000), and Metéo-France's 'Code d'Analyse Necessaire à ARPEGE pour ses Rejets et son Initialisation' (CANARI) systems. MESCAN uses covariance matrices to characterise surface correlation structures as a function of 2-metre temperature and relative humidity, along with distance functions between each grid-point, both vertically and horizontally.

2.1 Minimisation

The 3DVAR cost-function is solved using an algorithm based on a known quasi-Newton limited-memory technique, (Nocedal, 1980), called M1Q3. M1QN3 is known to be inherently more forgiving of non-strictly quadratic cost-functions compared to the conjugate-gradient alternative CONGRAD. M1QN3 minimises the cost-function by calculating an inner-product of the cost-function gradient and approximated Hessian, without storing

Table 1. Mean number of observations (in thousands) for March[a] in each respective year as used in CERRA-DET.

Year	1985	1991	1998	2003	2012	2018
Total	11.9	12.4	14.5	34.2	79.1	78.9
Satellite	1.0	0.8	0.7	10.6	40.6	38.3

[a]Whole month.

either. Instead, a pair of vectors are stored which can reconstruct the Hessian or gradient as necessary. The Hessian approximation is obtained by an inverse BFGS algorithm, and the step-size computed from a line-search procedure. The number of minimisation iterates of M1QN3 is fixed at 50 with no other iterative halting criterion. Each analysis is performed every 3 hours in a continuous assimilation cycle. The fields analysed are vorticity, divergence, temperature and surface pressure and specific humidity. Other prognostic variables are simply copied from the background.

2.2 Observations

This section details all observations used for CERRA-DET and CERRA-EDA. Most of the observations are stored and retrieved from two main sources; the Meteorological Archival and Retrieval System (MARS) and European Centre File Storage system (ECFS) at the ECMWF. These observations include the so-called conventional observations, like data from SYNOP stations, ships, drifting buoys, radiosonde and aircraft (AIREP, ACAR, AMDAR), satellite radiance from different instruments such as Advanced Microwave Sounding Units (AMSU-A and AMSU-B), Microwave Humidity Sounder (MHS) and Infrared Atmospheric Sounding Inferometer (IASI), and satellite based atmospheric motion vector (AMV) wind data. Further, ground-based observations of zenith total delay (ZTD) from the Global Navigation Satellite System (ground-based GNSS) are also used. The ground-based observations from GNSS are part of a network of reprocessed ZTD's provided by the European Reference Permanent Network (EUREF-EPN, Bruyninx et al. (2019)). Furthermore, GNSS-RO (Radio Occultation) observations provided by the Global Navigation satellite system receiver for Atmospheric Sounding, from EUMETSAT's large network of Satellite Application Facilities (GRAS-SAF) are also reprocessed to increase the signal to noise ratio. Finally, scatterometer observations are fetched directly from the EUMETSAT dissemination centre EUMETCAST. In addition, we have excluded data from NOAA's older Microwave Sounding Unit (MSU) data, which had temporal coverage from 1978 to 2005, due to unresolved technical issues.

The bulk of the observations both prior and post-processing, are increasingly dominated by satellite radiance observations as the years progress up to 2018 as can be seen in Table 1. ATOVS and IASI data are quite dominant in 2018 having the potential to make up 90% of the total observations at a single assimilation time. The number of observations used prior to any processing or thinning can be of the order 10^7, which are then reduced down to an order of 10^4. Further details of the observation specifics are covered in Wang and Randriamampianina (2021). It is important to keep in mind the temporal evolution of observation densities in the context of reanalysis production because of its potential impact on analysis quality in general, but particularly background error covariances, and therefore the quality of the reanalysis, (Brousseau et al., 2012a).

$$\mathbf{C}^{q_u} = \begin{bmatrix} \begin{pmatrix} h^1 & & \\ & \ddots & \\ & & h^L \end{pmatrix}_{(1)} \begin{pmatrix} v^1v^1 & v^1v^2 & \dots & v^1v^L \\ v^2v^1 & v^2v^2 & & \vdots \\ \vdots & & \ddots & \\ v^Lv^1 & v^Lv^2 & \dots & v^Lv^L \end{pmatrix}_{(1)} & & \\ & \ddots & \\ & & \begin{pmatrix} h^1 & & \\ & \ddots & \\ & & h^L \end{pmatrix}_{(N)} \begin{pmatrix} v^1v^1 & v^1v^2 & \dots & v^1v^L \\ v^2v^1 & v^2v^2 & & \vdots \\ \vdots & & \ddots & \\ v^Lv^1 & v^Lv^2 & \dots & v^Lv^L \end{pmatrix}_{(N)} \end{bmatrix} \tag{2}$$

3 B-matrix design

The B-matrix is not stored explicitly. However, we can elucidate each of its constituents to gain insight into its role in CERRA's 3DVAR. The spatial correlation matrix, containing one block for each total wave-number of the spectral model, for the correlations between the 4 prognostic variables: wind (vorticity (ξ), divergence (η)), the mass field (temperature (T) and surface pressure (P_s)) and specific humidity (q) is such that:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}^{\xi} & & & \\ & \mathbf{C}^{\eta_u} & & \\ & & \mathbf{C}^{(T,P_s)_u} & \\ & & & \mathbf{C}^{q_u} \end{pmatrix}. \tag{1}$$

Each block contains vertical and horizontal correlations, for all total wave-numbers $n = 1, ..., N$ and for all model levels $l = 1, ..., L$, $L = 106$. Therefore, $\mathbf{C} \in \mathbb{R}^{4NL\times 4NL}$. Note that vorticity is considered balanced.

The structure of one of these blocks, for specific humidity for example, has the form shown in (2), where $\mathbf{C}^{q_u} = Diag\left(\mathbf{h}^{q_u}_{(1)}\mathbf{V}^{q}_{(1)}, ..., \mathbf{h}^{q_u}_{(N)}\mathbf{V}^{q}_{(N)}\right)$, for $\mathbf{h}_{(n)}, \mathbf{V}_{(n)} \in \mathbb{R}^{L\times L}$, and $\mathbf{C}^{q_u} \in \mathbb{R}^{NL\times NL}$. The subscripts (n) denote the wave-number and the super-scripted values inside each constituent block, h^l, v^l, denote the horizontal and vertical correlation values at level l, respectively. The balance operator contains the multivariate balance relationships as originally devised by Derber and Bouttier (1999) and later applied to ALADIN in Berre (2000), such that:

$$\mathbf{L} = \begin{pmatrix} \mathbf{I} & & & \\ \mathbf{MH} & \mathbf{I} & & \\ \mathbf{NH} & \mathbf{P} & \mathbf{I} & \\ \mathbf{QH} & \mathbf{R} & \mathbf{S} & \mathbf{I} \end{pmatrix}. \tag{3}$$

This $\mathbf{L}$ component accounts for the mass-wind and specific humidity balances, separate from the correlations dealt with in $\mathbf{C}$. The theoretical idea is that the constituents of $\mathbf{L}$, address the hydrostatic and geostrophic balances in the vertical and horizontal. The vertical balances, for mass-wind are taken into account by $\mathbf{M}$, $\mathbf{N}$ and $\mathbf{P}$ and these are related to specific humidity via $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$. The horizontal balances are applied with the horizontal balance operator $\mathbf{H}$, which is a diagonal matrix taking spectral vorticity coefficients and obtaining balanced geopotential by multiplication of the assumed linear regression coefficients. Balanced geopotential is the balanced part of the linearised mass variable deduced from (T, P_s) via the linearised hydrostatic relationship (Parrish et al., 1997). Finally, $\mathbf{\Sigma}$ is a diagonal matrix of σ_b along the diagonal, representing the background error for the respective wave-number, level and variable in spectral space. The B-matrix can therefore be written as sparse matrices:

$$\mathbf{B} = \mathbf{L}^T \mathbf{\Sigma}^T \mathbf{C} \mathbf{\Sigma} \mathbf{L}, \tag{4}$$

which is identical in its formulation to (Berre, 2000).

The regression coefficients and standard deviations are updated every time the forecast differences are harvested from the EDA, which we now describe.

3.1 B-matrix estimation: CERRA-EDA

CERRA-EDA is used to estimate the B-matrix and for uncertainty quantification for CERRA-DET. CERRA-EDA is a 10-member ensemble where each member has its observations perturbed using a diagonal observation error covariance matrix. The observation error covariance matrix is assumed to have zero-mean and follow a Gaussian distribution. Once each observation is perturbed a 3DVAR assimilation is performed and an analysis follows. This assimilation of 3DVAR is performed continuously, cycled every 6 hours. The analysis is then used to produce a 6-hour forecast for each ensemble member. These forecasts are forced by perturbed LBCs from ERA5-EDA, which have 63km resolution at the equator. This EDA implementation permits the consideration of observation error and background error, including LBC error, in the data assimilation process. Model error is not taken into account and thus none of the conventional techniques such as Stochastic Perturbation of Physical Tendencies (SPPT) Stochastic Kinetic Energy Backscatter (SKEB) has beeb included.

Each forecast is 6-hours long. The differences of these 6-hour forecasts are computed between each adjacent ensemble member. The forecast differences are computed such that

$$\mathbf{d}_{i,i+1}^{(i)} = \mathbf{x}_i^f - \mathbf{x}_{i+1}^f, \tag{5}$$

for $i = 0, ..., 8$ and the difference between last and first member $\mathbf{d}_{9,0}^{(10)} = \mathbf{x}_9 - \mathbf{x}_0$. These differences are then used to compute the correlations and balance relationships that comprise the B-matrix.

There are two distinct EDAs performed in this way by CERRA-EDA to estimate the B-matrix used in CERRA-DET. A high-resolution 5.5km EDA performed just once, and a lower-resolution 11km EDA following CERRA-DET in parallel. The purpose of the high-resolution EDA is mainly to capture seasonal variability, and to estimate background error covariances for CERRA-DET's high-resolution horizontal scales from 11km to 5.5km. The purpose of the lower-resolution EDA is to have live tethering to the current meteorological situation, as is realised on a daily basis, but only for horizontal scales higher than 11km. The daily forecast differences are used to update the B-matrix in CERRA-DET every 2 days.

The choice of 3DVAR, 10-members, 11km resolution and 10-minute time-step for the EDA were a sufficient compromise to meet constraints on computational expense and implementation. Reducing the resolution from 5.5 to 11km yields half the number of grid points in both directions of the bi-harmonics, and allows for an increase in time-step. So this is why CERRA-DET has a 4-minute time-step in comparison to CERRA-EDA's 10-minute time-step. The numerical cost of the 10-member 11km EDA is twice the cost of CERRA-DET.

We have chosen this route of using an EDA with 3DVAR mainly as an optimal compromise of operational deadline and resource constraints such as running cost, complexity (maintenance), and running time (speed). Another reason is that in the context of the reanalysis, the reasonable measure of uncertainties that can be obtained from an EDA is an attractive prospect. Finally, the potential of the EDA to alleviate the need for B-matrix re-tuning served as another good reason to have an EDA-estimated flow-dependent B-matrix.

3.2 Potential statistical variability

Prior to development of CERRA-EDA, a preliminary system was created and successfully tested on Metéo-France's VORTEX/OLIVE-ALADIN system. This acted as the blueprint for the ensuing HARMONIE-ALADIN CERRA-EDA system. It is this preliminary system that was used to identify potential variabilities requiring consideration and decisions with regards to developing CERRA-EDA. This section discusses plots produced using this preliminary system. The OLIVE-ALADIN EDA system had a 5.5km resolution, 5-members and forced by 42km LBCs from AEARP (Metéo-France's global EDA system for ARPEGE), cycled 6 hourly with differences taken from 6-hour forecasts. The only differences between this preliminary system and the CERRA-EDA system are; the number of members and the resolution of the LBC's. The two periods used for the results shown in this section were $1^{st} - 18^{th}$ July 2017 and $1^{st} - 18^{th}$ December 2018.

The potential areas of statistical variability relevant to our study here are: the model and its spatial resolution, time-scales (daily, weekly, monthly and seasonal), observation systems over decadal time-scales and weather regimes. These areas of variability have also been echoed in (Brousseau et al., 2012a). Weather regime change, if occurring on time-scales of a few days or weeks say, could potentially be categorised under daily or weekly variability. An illustration of the potential variability that can be exhibited day-to-day is shown in Figures 1 and 2.

Figures 1 and 2 both simultaneously show potential seasonal and daily variabilities. Figure 1 shows total σ_b values in the vertical, where between 1000hPa and 800hPa there is potential of a 50% change in the base σ_b value for specific humidity from summer to winter for example. Similarly, larger potential variability can be seen in the horizontal scales, where nearly an entire order of magnitude of change in specific humidity (top-left plot, Figure 2), is possible, for the largest scales ($\sim 10^3$) and the smaller scales ($>10^2$).

In Figure 3(a) the total σ_b^{Ps} value over time has potential to change significantly from the base value that would otherwise be predicted by a seasonally static set of statistics arising from a climatological B-matrix (dotted-lines).

An important parameter that we derive from the general theme exhibited by Figure 3, more specifically Figure 3(d) for vorticity, is that the steepest gradient in σ_b occurred from days 10 to $12\frac{1}{2}$ (approximately). This illustrates that the maximum change shown in this preliminary experiment for a prognostic variable is $2\frac{1}{2}$ days. This contributed to our choice of using an EDA moving average of $2\frac{1}{2}$ days, for CERRA-EDA and updating the B-matrix every 2 days in CERRA-DET.

3.3 Forecast difference mixing

There are 2 pools of forecast differences used to produce the B matrix. The high-resolution (5.5km) climatological part, $D(t)^H_{clim}$ is a mixture of summer and winter differences, D^H_e and D^H_h respectively, where superscript 'H' denotes high-resolution and the subscripts denote e, (ete - summer) and h, (hiver - winter) respectively. The forecast differences for D^H_e come from 1^{st}-10^{th} July-2017. The forecast differences for D^H_h come from 1^{st}-10^{th} January-2018. The forecast differences for these periods have been done previously (offline) and are stored. A time-dependent function, which governs the proportion of summer and winter differences used based on the time of year, then synergises an appropriate number of forecast differences from each season to make up the $D(t)^H_{clim}$ component. Figure 4 shows the weather regimes of each respective period. The second part

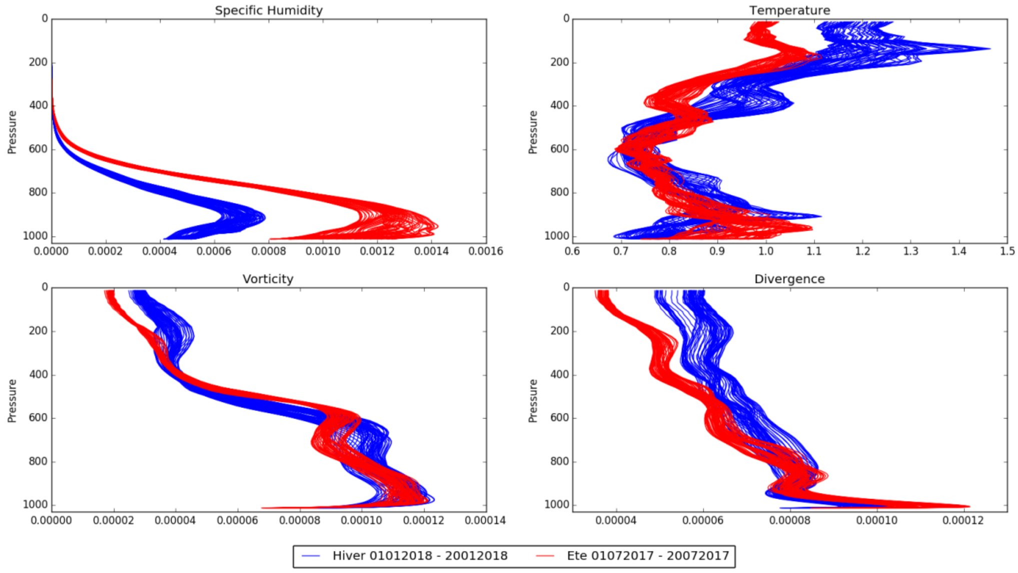

Figure 1. Vertical daily and seasonal variabilities using OLIVE-ALADIN 5-member EDA system. Vertical profiles of horizontally averaged σ_b vs pressure (hPa) for; specific humidity (top-left), temperature (top-right), vorticity (bottom-left) and divergence (bottom-right). Each profile represents σ_b for one assimilation time; summer is represented by the period 01-07-2017 (red-lines) winter is represented by the period 01-01-2018 (blue-lines).

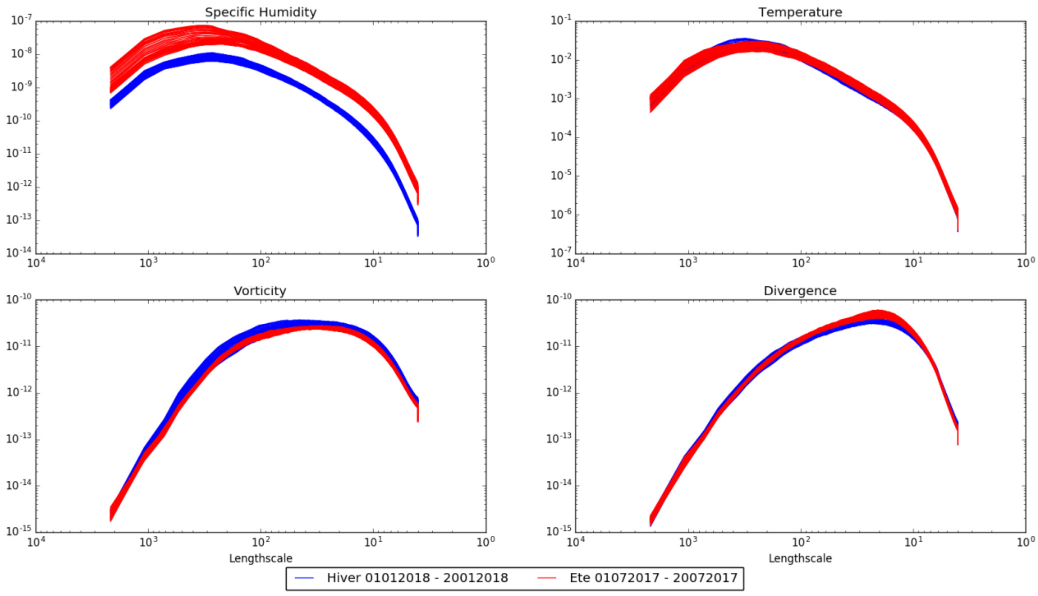

Figure 2. Horizontal daily and seasonal variability using OLIVE-ALADIN 5-member EDA system. Horizontal values of σ_b (vertically averaged) vs length-scale (km) for; specific humidity (top-left), temperature (top-right), vorticity (bottom-left) and divergence (bottom-right). Each profile represents σ_b for one assimilation time; summer is represented by the period 01-07-2017 (red-lines) winter is represented by the period 01-01-2018 (blue-lines).

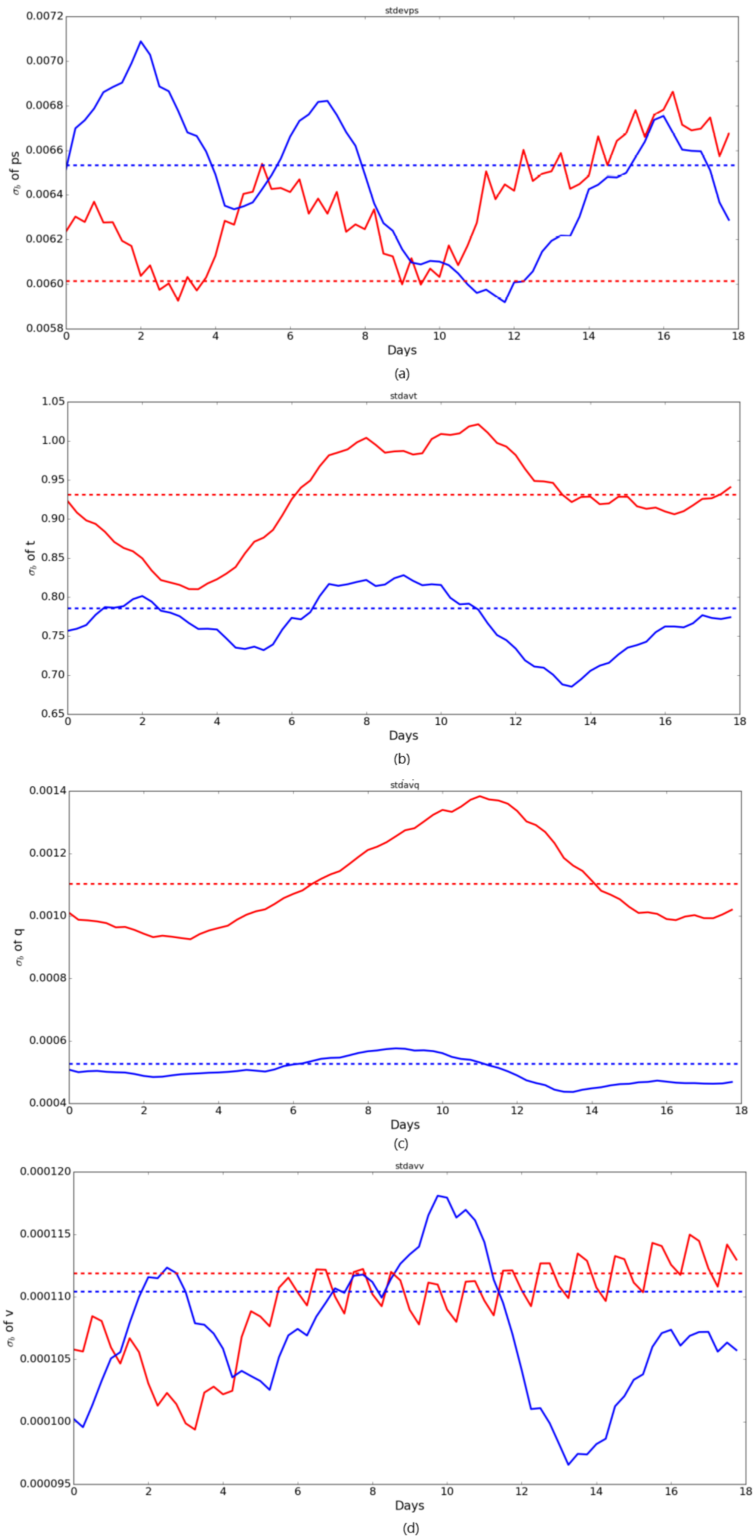

Figure 3. Potential daily variability over 1 month using OLIVE-ALADIN 5-member EDA system. Total σ_b value for (a) surface pressure, (b) temperature, (c) specific humidity and (d) vorticity. Each line represents the total σ_b value for each day in the respective month; first 18 days of December-2018 (blue-line) and first 18 days of July-2017 (red-line). The dotted lines are climatological σ_b value taken from static climatological B-matrix; summer seasonal average (red-dotted line) winter seasonal average (blue-dotted line).

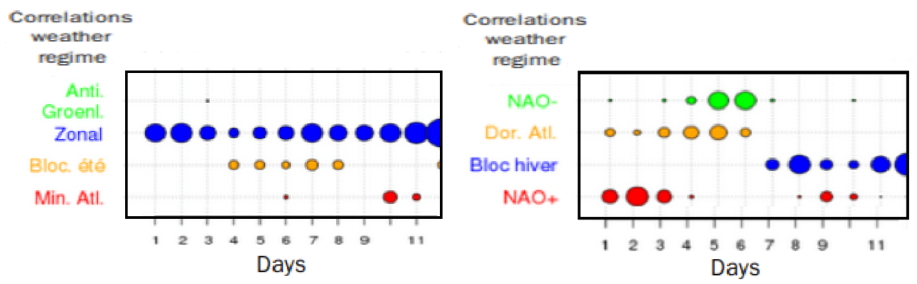

Figure 4. Weather regimes of periods used for $D(t)^H_{clim}$ climatological differences. The summer period represented by the chosen days in July-2017 is dominated by a zonal regime, whereas the winter period represented by January-2018 shows a near-half between NAO+/NAO- followed by a less-imposing winter blocking regime. Regimes: Bloc. été/hiver - summer/winter blocking, Dor. Atl. - Atlantic Ridge, Min. Atl. - Atlantic Minimum

comprises low-resolution (11km) daily differences, D^L_{jour}, coming from the 'online' CERRA-EDA running in parallel to CERRA-DET.

The function governing the proportion of forecast differences mixed is such that:

$$D = \underbrace{\left[D^H_e(1 - t(d)) + D^H_h t(d)\right]}_{D^H_c} \alpha + (1 - \alpha) D^L_j, \tag{6}$$

where D is the total pool of forecast differences used to create the B matrix, α is the forecast differencing mix ratio and:

$$\begin{aligned} r(d) &= \begin{cases} -d \mod h & \text{for } d < h \\ d \mod h & \text{for } d > h \end{cases}, \\ t(d) &= \frac{r(d)}{h}, \end{aligned} \tag{7}$$

is the seasonal linear weighting function. h is half the number of days in the year and $d \in [1, 364]$ is the day within the current year. As an illustrative if that time of year is 1^{st} January, then $t(d)$ would dictate that *all* of the differences would come from D^H_h and none from D^H_e. Preliminary studies for choice of α indicated that $\alpha = 0.8$ for CERRA-DET is an optimal compromise. This allowed us to achieve the most desirable analysis, as illustrated in the following sections. The forecast differences, D^L_j, constitute a $2\frac{1}{2}$ day moving average from CERRA-EDA. This is to account for the maximum gradient shown to occur within a $\sim 2\frac{1}{2}$-day period, as discussed in our preliminary experiment in section 3.2. This would also allow for abrupt changes arising for example from a weather regime change, which may be instigated by sub-grid-scale processes.

4 Case study: Can the EDA capture a change in weather regime?

Identifying a sudden change in large-scale weather phenomena such as a weather regime would improve the analysis, since the ensuing state estimate would be more accurate at any given time. Subsequent improvement of forecast quality should also ensue, however this is isn't a given. So we begin by asking the question: Can a flow-dependent B-matrix be dynamic enough to detect a sudden change in weather regime? If so are there any caveats? And what impact does this have on analysis and forecast quality? Before answering these questions we briefly explain how a weather regime is characterised.

While there are numerous methods which are used to identify weather regimes, the data we use utilises the methodology detailed in (Vautard, 1990). The way the authors

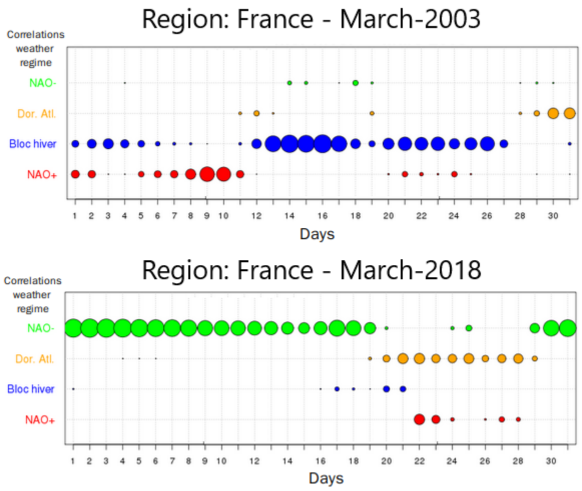

Figure 5. The correlation of the actual weather regimes exhibited in periods; March-2003 (top plot) and March-2018 (bottom plot). Region: France. Plot and data courtesy of Météo-France: http://seasonal.meteo.fr/content/suivi-clim-regimes-quot.

compute the weather regime is by using 24-hour centered finite difference of vectors containing principal components of 'large-scale tendencies' to compute 'instantaneous tendencies'. These large-scale tendencies depend on: themselves, small-scale components and other factors. A composite tendency function, which gives a weighted-average of, small-scale, 'other factors' and large-scale tendencies, is minimised using a least-squares approach, The solution of this least-squares function becomes the large-scale tendency, or weather regime. In this way, 4 weather regimes categorically emerge: a European blocking dipole, enhanced zonal flow, a positive anomaly over Greenland and a ridge over the eastern Atlantic Ocean. Weather regimes in Europe are generally identified by quasi-stationary centres of pressure ratios at ∼500hPa roughly over the Azores Islands, west Portugal and Iceland. The weather regime paradigm is used to characterise large-scale circulation patterns over regional domains. It is also used as a proxy to predict significant short-to-mid-term changes in the statistical probability of hot and cold extremes and precipitation occurrences across Europe. We will term categorisations of weather regimes in Europe, as mentioned previously, as: North-Atlantic Oscillation (NAO $+/-$), Atlantic Ridge, blocking and 'zonal'.

For the purpose of being able to clearly identify if the B-matrix is capable of detecting a change in weather regime, we deliberated in selecting two periods where the winter regimes differ: March-2003 and March-2018, as shown in Figure 5.

4.1 Experiment design

The B-matrix used for CERRA-EDA itself is a pre-populated 5.5km static climatological B-matrix, not to be confused with the temporally updated CERRA-DET B-matrix computed from CERRA-EDA. This same static climatological B-matrix acts as the referential comparator in the experiments that follow. This static climatological B is composed of differences from $D(t)^{H}_{clim}$, as described earlier in section 3.3.

We seek to establish whether the ability of the B-matrix to recognise a change in weather regime is contingent on:

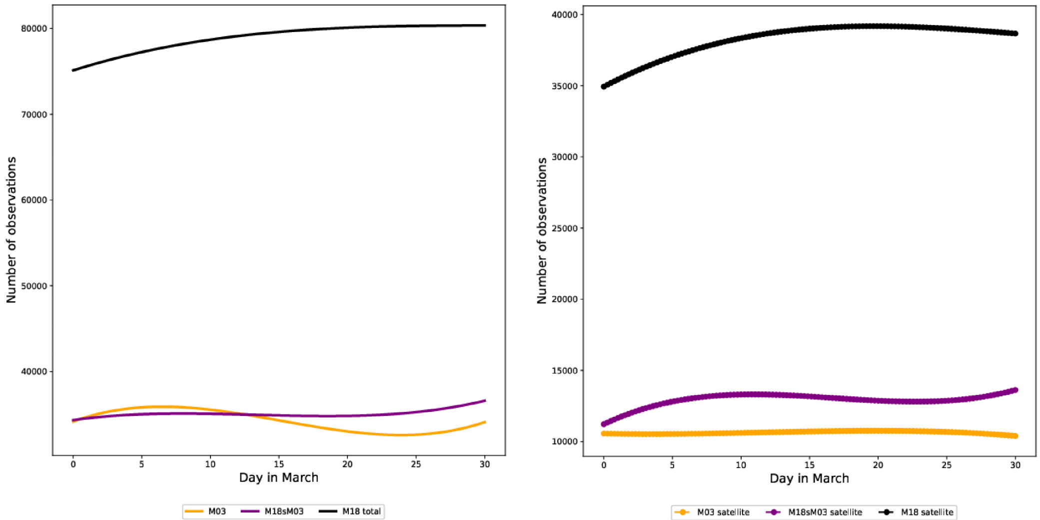

Figure 6. Total numbers of observations (left plot) and satellite observations (right plot), by day in March (x-axis), for: March-2003 (orange line), March-2018 simulating March-2003 observations and March-2018 with default settings (black line). The observation data was averaged using a 4^{th} degree polynomial.

1. Varying climatology and daily forecast difference influence via α.
2. The observation network used in both respective periods.

To investigate varying the influence of climatology and daily differences, we compute the following B-matrices:

BS Climatology only. No daily. $\alpha = 1$. $t(d) = 0.5$ fixed.
B8020 Climatology-dominant with daily. $\alpha = 0.8$. $t(d)$ varying.
B5050 Climatology and daily equal. $\alpha = 0.5$. $t(d)$ varying.
B2080 Daily-dominant with climatology. $\alpha = 0.2$. $t(d)$ varying.

And to investigate the contribution of the observation network to the statistics we run use the B-matrices above to produce analyses for the following periods:

(i) March-2018 with default observation settings. (M18)
(ii) March-2018 with observations settings closely mimicking March-2003. (M18sM03)
(iii) March-2003 with default observation settings. (M03)

The abbreviations are what we used to refer to each period later in the plots in this section. M18sM03 allows us to isolate general contribution of the observation network to the ensuing results. To allow the observation network of M18 to mimic M03 as closely as possible (M18sM03), we removed the following observations: SATOB (polar and geostrophic), Atmospheric Motion Vectors (AMV), Aircraft observations (AMDAR, AIREP, ACARS), and IASI. Figure 6 shows the total and satellite observation numbers for (i,ii,iii). An example of the aircraft observation coverage is shown in Figure 7. Figure 7 in 2018 shows significant aircraft coverage of regions close to pressure centres and flow-corridors to detect weather regime change. For changes in weather regime to be detected this is significant. It is important to note that these aircraft observations were removed for M18sM03. Finally, it was shown in Wang and Randriamampianina (2021) that the aircraft obser-

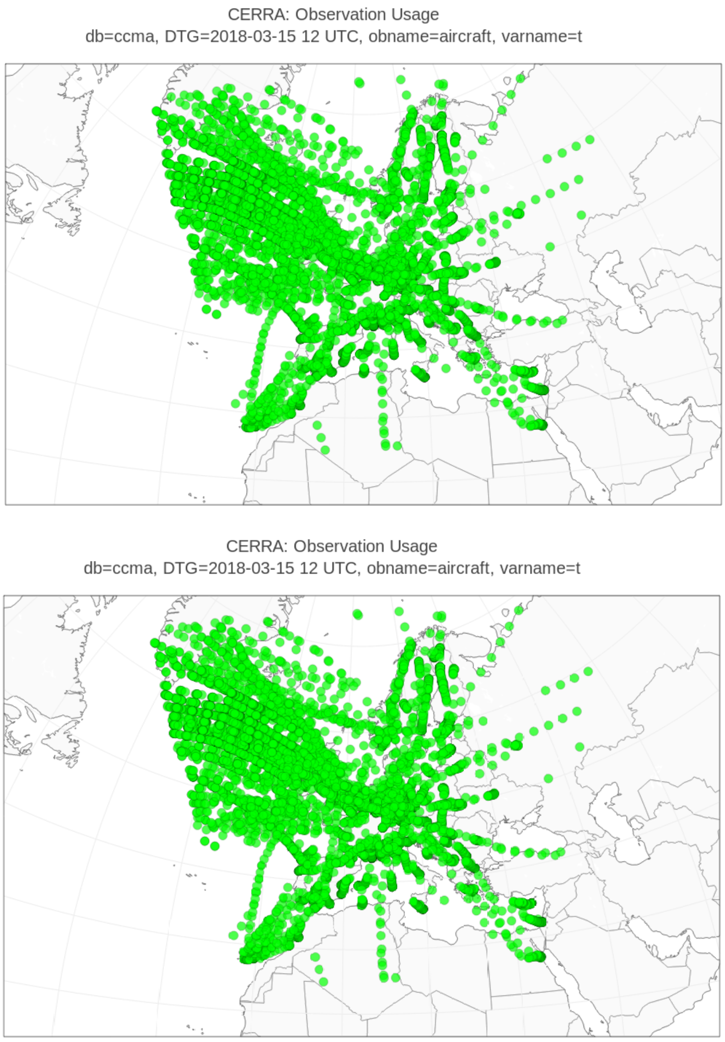

Figure 7. Aircraft observation coverage at all vertical levels.

vations had the most impact on CERRA analyses, which is relevant to us here since these form a part of the observations we removed for M18sM03.

4.2 Results

In this section, we describe our results in three parts; the temporal evolution of the computed σ_b value, the horizontal and vertical background error variance spectra, and their impact on the quality of the analysis and forecast.

4.2.1 Temporal evolution of σ_b

The time series of σ_b^q indicates two interesting aspects. The first aspect is the contrasting range of σ_b values for both periods. The mean range of M18 (red-line) is approximately $\sigma_b = 5$, showing a strong tendency to remain near the static σ_b value, except in times of WR change (days 21-28 for M18, Figure 5). M03 has a very wide σ_b range of nearly $\sigma_b = 4$, remaining virtually flat even in times of WR change, keeping in mind that the M03 WR change is between days 6-11. Since M18sM03 also has a wide range of σ_b values, it is fair to say that the large σ_b range is attributed to the difference in observation network principally, since M18 would naturally have the improved observation network.

The second aspect is the flexibility of σ_b. M18 with its original observation network is the only period showing any significant change at times of WR change. This is connected to α because we only see these changes for $\alpha = 0.2, 0.5$. M18 σ_b values exhibit more readiness to change, for example between days 7-9, while remaining close to the static value, $\sigma_b = 4.8$. Another strong fluctuation of σ_b between days 23-27 further illustrates the potential for rapid change, given a more weighty influence on the daily EDA statistics, ie. for $\alpha = 0.2$. This is mimicked in a more dilute sense for $\alpha = 0.5$. M03 and even M18sM03 show near-to-no change in comparison to M18, except between days 26-28, and even then it is very small.

It is therefore fair to conclude from Figure 5, that the coverage of the improved observations of M18 over M03, (Table 1) aid in; the flexibility of σ_b at WR change time (days 21-28), and the range of the σ_b values of each respective period. Aiding this conclusion is both M03 and M18sM03 have nearly identical σ_b value-ranges and σ_b temporal stasis. This is due to having almost identical observation networks. In the experiment with the most improved observation network (relatively), the α value facilitates the breadth of change σ_b can exhibit. This illustrates both the ability of the EDA to adjust to sudden changes in the meteorological situation, and the need to estimate α correctly.

4.2.2 Variance spectra and profiles

The horizontal spectra in Figure 10 illustrates the primacy of having an updated and denser observation network. Comparing M18 (blue-lines) with M18sM03 (green-lines), the blue line consistently has a higher variance profile across all wavelengths above 11km. This is explained both; by interpolated daily forecast differences constituting half of the differences ($\alpha = 0.5$), which are interpolated from the EDA resolution of 11km to the reanalysis production resolution of 5.5km. M18sM03 otherwise behaves almost identically to the bulk of the M03 spectra (red-lines), where the few spectra lines where M03 has increased variance in the lower wavelength range vorticity and divergence, is due to the different weather regimes present. M03 is mainly blocking (Figure 4), where the 3 red profiles that are visibly higher (bottom two plots of Figure 10), almost reaching the same values as M18 (blue lines), represent the 1 day where there is a slight increase in geopotential. M18 is mainly NAO-, with higher geopotential values, with better observations, which explains the consistently higher vorticity and divergence profiles in the bottom two plots of Figure 10.

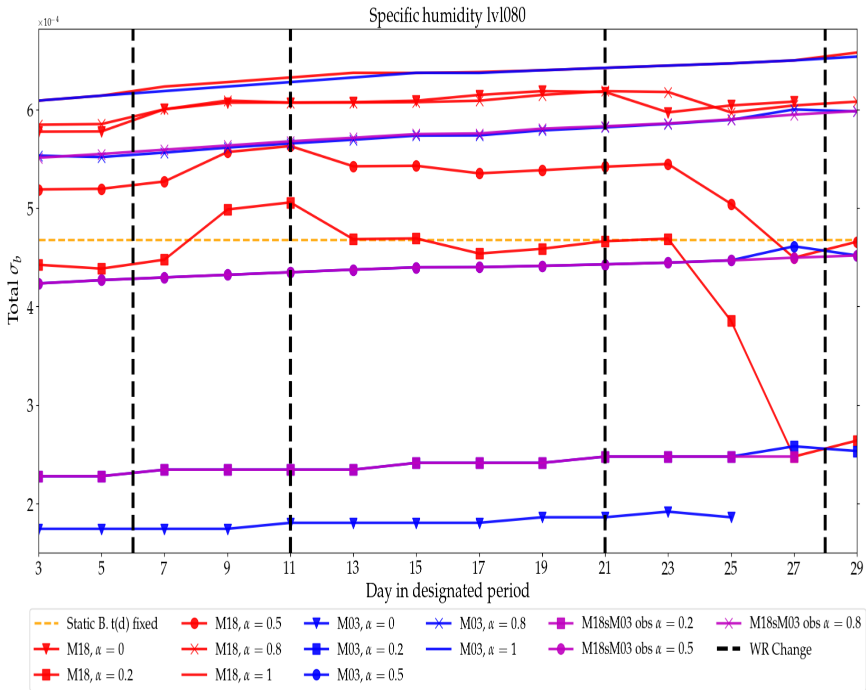

Figure 8. Total σ_b value for specific humidity at ~800hPa from 3^{rd}-29^{th} March. Periods: M18 (red-lines), M03 (blue-lines) and M18sM03 (purple-lines). The B-matrices are as follows: BS (yellow-dotted-line), B8020 $\alpha = 0.8$ (crosses), B5050 $\alpha = 0.5$ (dots), B2080 $\alpha = 0.2$. The vertical black lines show points of weather regime change from both periods seen in Figure 5.

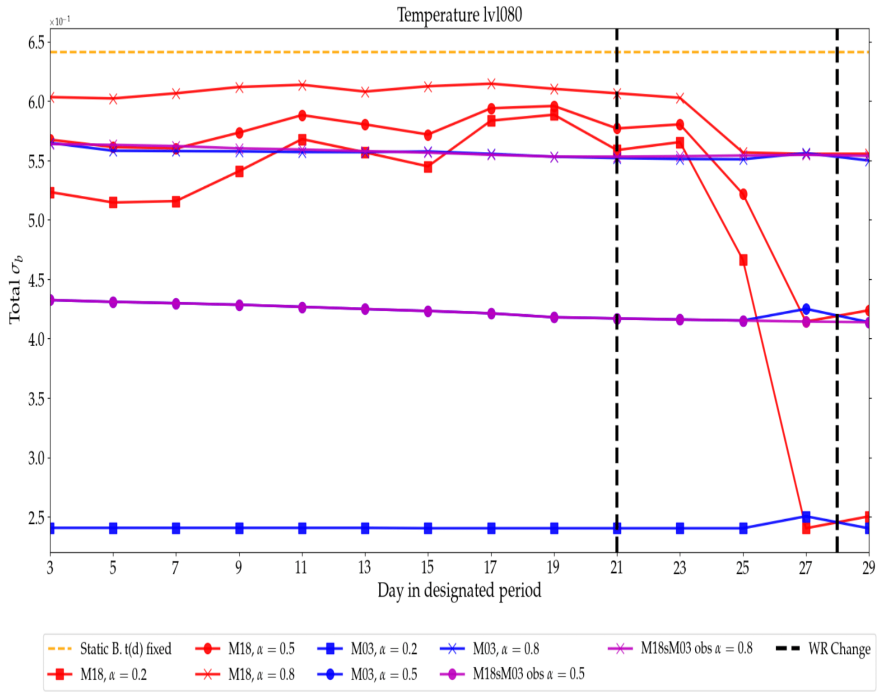

Figure 9. Same as Figure 8 but for the Temperature variable.

The vertical profiles in Figure 11 clearly show that increasing α, namely increasing daily forecast difference, increases all of the standard deviation values for temperature throughout the vertical, regardless of period. Conversely, having the minimum value of $\alpha = 0.2$ as seen in Figure 11, the standard deviation of temperature is reduced by up to 1/3 of its original value.

Finally, comparing M18 to M18sM03 (left-most and middle plots, Figure 11), we observe that the range of potential values of temperature standard deviation is much wider for M18. This is mainly due to the improved general observation coverage of M18 over M03 (Figure 6). This is most visible for example for $\alpha = 0.2$, where for M18M03 (Figure 11, middle plot, red-lines) , σ_b^T does not exceed ~ 0.4, whereas for M18 σ_b^T can reach 0.75, similar to the static B-matrix standard deviation, which has no daily influence. This illustrates the dynamic potential of the B-matrix with more daily-EDA influence than climatological, which also shows in the total σ_b value, Figure 8.

4.2.3 Analysis impact

In this section, we discuss results obtained by using B-matrices mentioned in section 4.1 in the same system used for CERRA. We contrast the results from Table 2 with Figure 8 for specific humidity and Figure 9 for Temperature. The behaviour of the wind variable is identical to temperature (Figure 9), and is therefore not shown. Anytime σ_b is mentioned, Figures 8, 9 are being referred to for brevity. Diagnostics of relative difference of analysis and background departure RMS's normalised by BS values are shown in Table 2. These are for spatio-temporally regularised observations: aircraft measurements of temperature and wind, and ground-based GNSS for specific humidity. Relative differences of RMS's for background and analysis departures are referred to as RDRMS(O-B) and RDRMS(O-A) respectively.

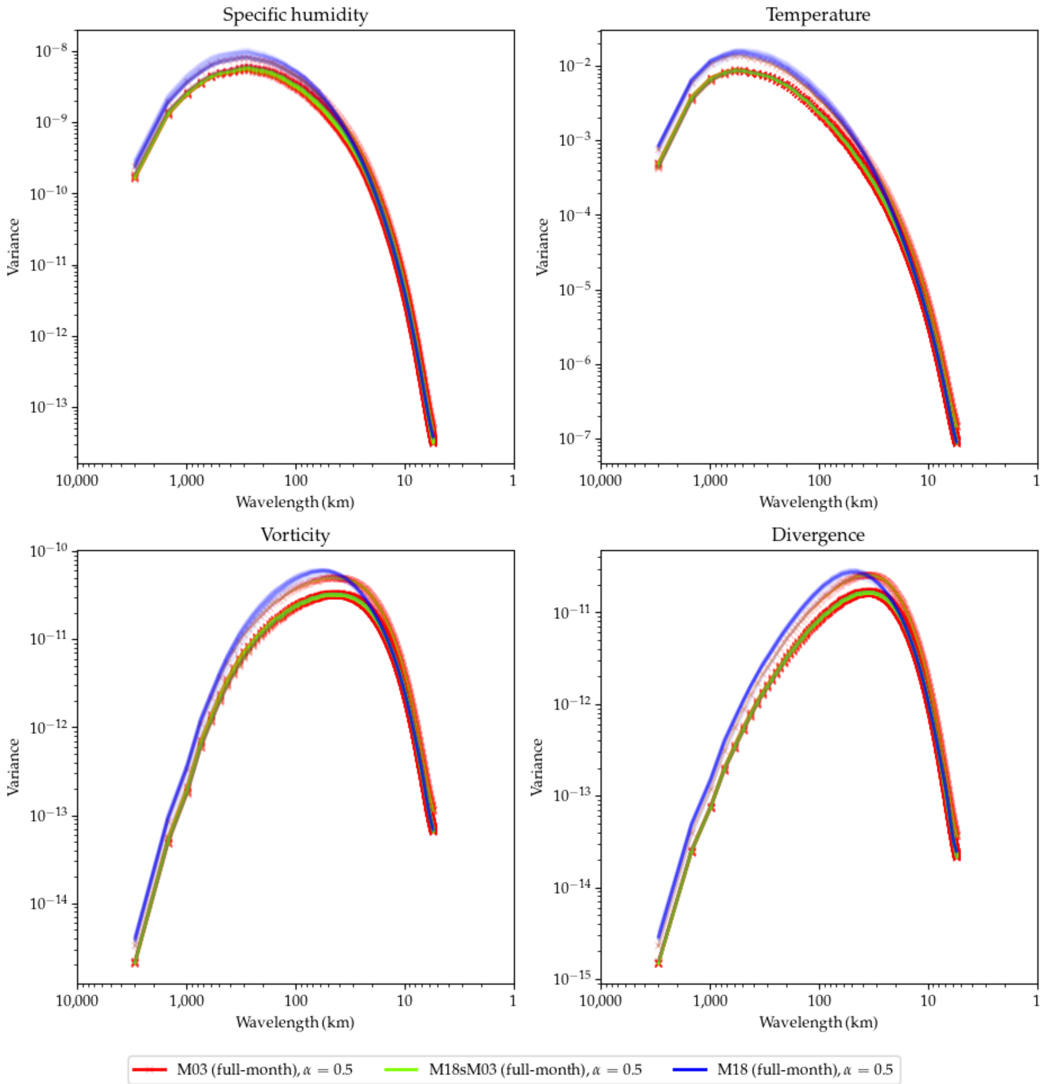

Figure 10. Horizontal variance spectra at 1000hPa of; specific humidity (top-left), temperature (top-right), vorticity (bottom-left) and divergence (bottom-right) of forecast errors for B-matrix B5050 during periods; M03 (red-lines), M18sM03 (teal-lines), M18 (blue-lines).

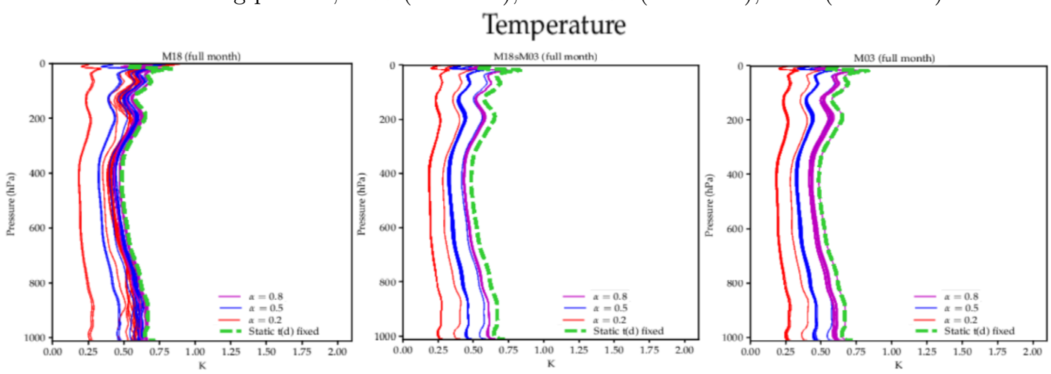

Figure 11. Vertical standard deviation profiles of Temperature for periods (left to right); M18, M18sM03 and M03. Standard deviations of; BS (green-dotted-line), B8020 $\alpha = 0.8$ (purple lines), B5050 $\alpha = 0.5$ (blue-lines) and B2080 $\alpha = 0.2$ (red-lines).

In short, RDRMS(O-B) for B8020 for example, measures how far the background-state produced by B8020 actually is from the observations, in comparison to the BS-equivalent, averaged over the period indicated. The same is analogously true for analysis departures, RDRMS(O-A). Positive values show how much the respective dynamic matrix has caused the analysis to move away from the observations in comparison to static.

σ_b of temperature (and wind) is smaller for all dynamic matrices implying more trust in the background compared to BS. This causes the positive RDRMS(O-A) values of temperature and wind (first column of Wind and Temperature, Table 2), indicating an analysis farther from the observations. Conversely, for specific humidity, σ_b values are less than BS equivalent, which show that the dynamic matrices are causing less trust in the background. As a result the analysis fits the observations more closely (first column of Specific Humidity, Table 2).

Examining the (O-A) values, on average for Mar-18, σ_b for temperature (Figure 9) and wind (not shown) are lower in the dynamic matrices compared to BStatic. It follows that the background is more trusted during the 3DVAR minimisation resulting in higher RMS(O-A) values, (as showed by positive RDRMS(O-A) values in Table 2). Conversely, specific humidity σ_b values are higher for the dynamic B-matrices compared to BStatic. It follows that the background is less trusted during the 3DVAR minimisation and RMS(O-A) is smaller.

The O-B values show that changes seen in wind and specific humidity benefit the subsequent background in the assimilation cycle. Therefore RMS(O-B) values are slightly reduced (the lowest reduction being -0.3% for specific humidity in B8020 and the highest being 0.7% for wind in B5050). However, these RDRMS(O-B) values are statistically significant. The opposite is true for temperature where the RMS(O-B) values increase slightly, but the changes are not statistically significant.

The period of WR-change shows the most significant change of σ_b values for the dynamic B-matrices compared to BS (Figure 9). These changes are seen in the reduction of RMS(O-B) values (second columns and second rows for wind and specific humidity, Table 2) with a maximum reduction of 2.4% (highlighted in red) for Ground-based GNSS observations in B5050. This significant change between these two periods is not true for B2080 however. B2080 shows RMS values for (O-A) and (O-B) for wind and specific humidity variables (second row of B8020 and second row of B8020, Table) that are close to B2080 values. This indicates that increasing the proportion of daily information from the EDA has no added benefit in this case. A plausible reason for this is the disparity between the resolutions of the daily (11km) and seasonal (5.5km) forecast differences. The additional information from the daily differences is only relevant for model scales above 11km.

We have shown that the use of dynamical B-matrices has the potential to improve the general behaviour of the data assimilation system.

The impact of dynamic B-matrix changes on the quality of the forecast during the data assimilation cycle are further evaluated using precipitation skill scores shown in Figure 12. The 24-hour accumulated precipitation is simulated by the sum of eight 3-hour forecasts from the data assimilation cycle between 6 UTC and 6 UTC the following day, and compared to rain-gauge measurements. There are approximately 4850 rain-gauge measurements available each day. Figure 12 shows the relative difference of Hiedke Skill Score (HSS) for measured precipitation thresholds every 24 hours (0.2, 2, 5 and 10 mm/24h) for B8020, B5050 and B2080 experiments compared to BS. Positive values indicate that the dynamic B-matrix experiment is closer to the observations than BS. Circles on the curves indicate that the differences are statistically significant.

B8020 exhibits slightly better HSS values than BS (positive values, HSS is higher in B8020 than BS) for all the thresholds around the 1% relative difference region. How-

Table 2. Relative differences of Root Mean Square (RMS) values of Observation-Analysis (O-A) and Observation-Background (O-B) in the observation space for aircraft measurements of wind, temperature and ground-based GNSS observations. Each row shows B-matrices: B8020, B5050 and B2080 compared to and normalised by BS during two distinct periods: 1-31st March-2018, and 26-31st March-2018. Positive values in rows for each period indicate that the RMS for the respective B-matrix is larger than BS. Bold number indicates statistically significant differences using a student's t-test with 95% confidence interval.

B-matrix	Period	Wind[1]		Temperature[1]		Specific Humidity[2]	
		O-A	O-B	O-A	O-B	O-A	O-B
B8020	1-31st	**4.1**	**-0.5**	**3.6**	0.3	**-8.9**	**-0.3**
	26-31st	**6.0**	**-1.2**	**5.0**	0.2	**-10.0**	**-1.5**
B5050	1-31st	**6.2**	**-0.7**	**6.1**	**0.5**	**-2.8**	**-0.5**
	26-31st	**12.0**	**-2.1**	**11.0**	0.1	**-1.1**	**-2.4**
B2080	1-31st	**7.0**	-0.3	**7.2**	0.3	0.2	0.2
	26-31st	**7.1**	-0.2	**6.8**	0.2	**-10.1**	**-1.7**

[1]Aircraft observations.
[2]Ground-based GNSS.

ever, this improvement is significant only for 0.2mm and 10mm thresholds. B5050 only shows a significant improvement for the 5mm threshold, while B2080 is significantly better for 0.2, 2 and 5mm thresholds. For other thresholds the HSS differences are not significant.

The diagnostics drawn from Table 2 and Figure 12 confirmed our choice of B8020 B-matrix for CERRA-EDA and consequentially, CERRA production, providing us with the best compromise in cyclic analysis and background-state quality for the entirety of the reanalysis. This configuration also permits dynamic adjustment of covariances arising from horizontal scales above and below 11km, albeit homogeneously and isotropically. Increased weighting on forecast differences from the 5.5km EDA ensures the reduction of polluting numerical noise that could potentially arise from the 11km EDA. It is clear that B2080 allows for this, to the detriment of the cyclic quality of the background and analyses states, as well as the skill scores.

While one of the main purposes of a reanalysis system is to provide optimal estimation of the atmosphere state at a given time, the system can also be used to initialize longer-range forecasts for operational NWP. To examine this potential and the aptitude of the system to provide input data for longer range forecasts, Figures 13 and 14 show normalised RMSE of the analysis, 12-hour and 24-hour forecasts (F00, F12 and F24, respectively) all valid at 00H00 and 12H00 UTC, using radiosonde measurements. Each of these forecasts arise from using the dynamic B-matrices against forecasts using BS as our benchmark comparator. So positive values (to the right of the 0 in the figures) indicate improved forecast quality of the respective dynamic B-matrix over BS. As previously discussed, the smaller σ_b values of dynamic B-matrices for temperature and wind compared to BS result in subsequent analyses which fit the observations less, ie. the background is more trusted. As Table 2 generally shows larger increments to RDRMS(O-A), which explain the subsequent apparently degradation in analysis, F12 and F24 forecasts are not significantly worse. March-2003 shows differences between the dynamic B-matrices and BS which aren't significantly different from zero. Conversely, during March-2018 these forecast ranges exhibit statistically significant improvements (positive normalised RMSE differences) for the 600-300hPa layer, indicating that the changes in the data assimila-

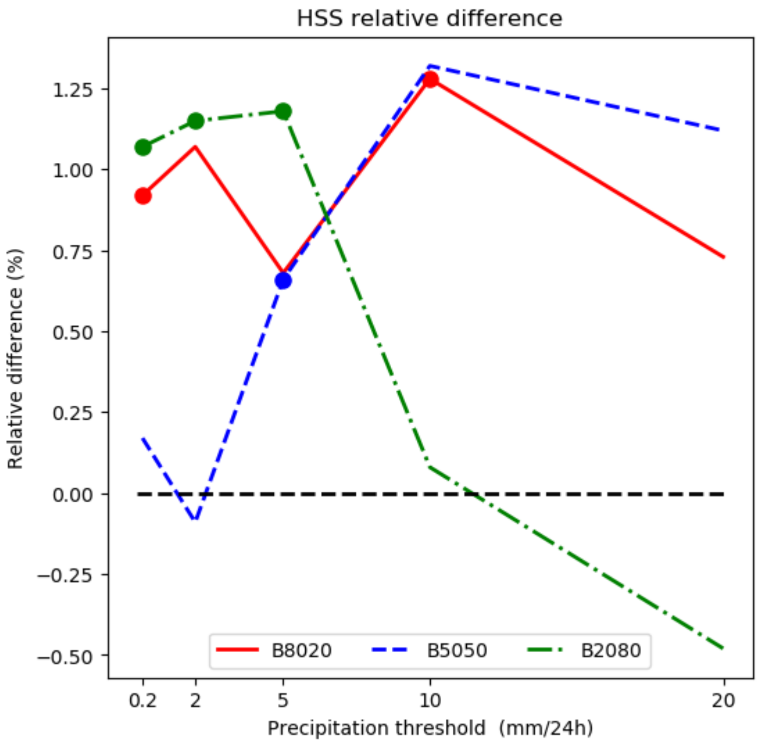

Figure 12. Relative differences of Hiedke Skill Scores (HSS) for 24h accumulated precipitation measurement thresholds (mm). HSS is compared to the chance score. The differences for each threshold are between measurements by rain-gauges (4850 on the geographical domain) and the sum of the 3-hour range forecasts from the data assimilation cycle (background) between 06UTC and 06UTC the next day. This is for the period from 01 to 31 March 2018. Each line shows the respective B-matrices; B8020 (solid red-line), B5050 (dashed blue-line) and B2080 (dash-dot green-line). Each experiment is compared to and normalised by BS for 0.2, 2, 5 and 10mm/24h thresholds. Circles on lines indicate statistically significant differences using a Boot-strap test with 95% confidence interval.

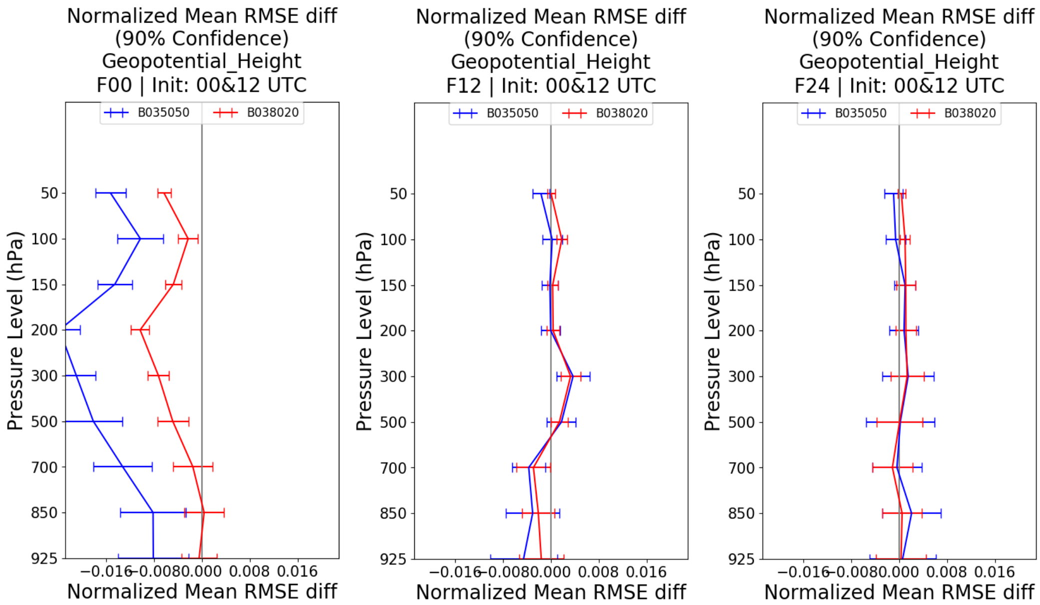

Figure 13. Difference in the mean of the root mean square error (RMSE), normalised by the mean scores for; the analysis (F00), 12-hour (F12) and 24-hour (F24) forecasts of geopotential field against radiosonde observations. Each line represents differences between forecasts with BS and forecasts with: B8020 (red line) and B5050 (blue line). Each line represents the average of forecast differences starting at both 00 UTC and 12 UTC during March-2003. Positive values, to the right side of the vertical line at 0, represent an improvement over BS.

tion system are of benefit to the forecast quality once again. Dynamic B-matrix σ_b values have the potential to increase over BS, providing analyses closer to observations and also better subsequent forecast as a consequence. B8020 exhibits better results in the lower atmospheric levels, at F24 below 850hPa for example, than B5050 or B2080.

5 Conclusions

In this paper we focused on detailing our new temporally flow-dependent augmented EDA system. The CERRA-EDA system was designed specifically for use in a regional 5.5km ∼40-year reanalysis. The system comprises a time-varying selection of seasonal differences at higher-resolution (80%), and a lower-resolution continuously cycled 6-hourly EDA (20%), averaged over 2.5 days. The B-matrix used in CERRA-DET is updated every 2 days.

Our line of scientific enquiry began with investigating if it was possible to capture weather regime change and if so, what caveats does it entail. We also wanted to know what impact this would have on the analysis and forecast. In conclusion it is possible to estimate and statistically realise weather regime change. It depends mainly on the observations, but it also depends on α, ie. the proportion of seasonal-daily forecast differences used for the EDA system. α can greatly impact the amplitude of σ_b change at times of WR change.

Our case study showed that the statistics of the B-matrix; the time evolution of σ_b, and horizontal and vertical variance spectra for our prognostic variables, do indeed capture the changes instigated by weather regime changes from NAO- during March-2018. It is clear from our penultimate experiment, simulating March-2003 observations in March-

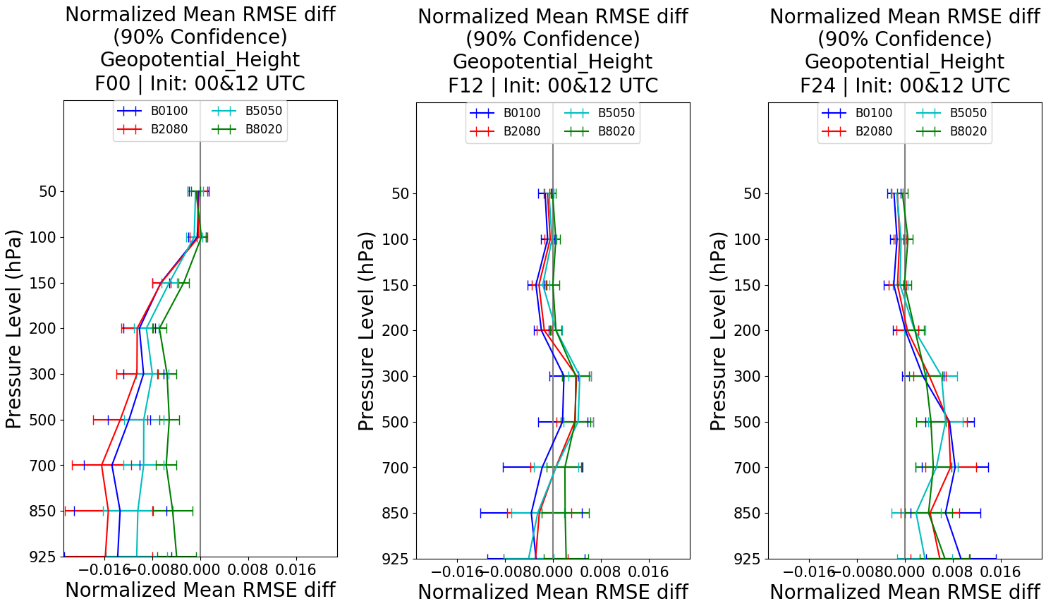

Figure 14. Same as Figure 13 but for March-2018. The additional B-matrix B2080 is represented by the green-dotted-line.

2018, that the observations are the reason. To further support our hypothesis, we illustrated the coverage differences between March-2003 and March-2018 (Figure 7). It was clear that March-2018 had far better aircraft observation coverage of the pressure centres characterising the advent of any weather regime. In addition, March-2018 had just over double the total observations and roughly 3 times as much satellite observations.

We also discussed the impact of varying α. It sufficiently increases the range of potential values that σ_b can take. This does not necessarily translate to better forecast skill scores, as can be seen for example in section 4.2.3, Figure 14. However, the range of σ_b values, while having some form of governance attributed to α, also depends on observation coverage as shown in section 4.2.1, Figure 8. It is shown that $\alpha = 0.2, 0.5$ captures WR far better than the rest, while also performing worse than BS. $\alpha = 0.8$ shows minor σ_b adjustments at times of WR change while having optimal forecast skill performance. The apparent analysis degradation for $\alpha = 0.8$ is due to the assimilation system placing less trust in the background, Table 2, and more in the observations. This provides better performance overall, when viewed in light of the cyclic nature of the assimilation system.

So while varying α is a useful tuning tool we found that it is intrinsically linked to observation coverage. This is known and can be understood analytically by viewing the analysis update equation, which shows the B-matrix acting to 'spread' information depending on the availability of observations at a discrete point in space. It is clear that the additional information, if taken into account without too much weighting beyond 20% in our case, that it the impact is positive overall. Cases where the weighting on the daily component of the EDA was 50-80% allowed for additional noisy information from the 11km scales to pollute the 5.5km scales. This is clearly something to be avoided.

In the context of our reanalysis the overall significance can only be seen in the years over improved observations, approximately 5 years out of the $/sim40$ year reanalysis timeframe. In the general context of EDA B-matrix estimation for DA in NWP however, it is a positive finding. It shows that it is possible for B-matrices to quickly adapt in the face of large-scale and small-scale phenomena if correctly tuned. It is also important to note, that the EDA in a reanalysis context also has the utility of uncertainty quantification. So while a modest improvement in forecast skill is welcomed, we cannot forget the added benefit of uncertainty quantification as tool to improve our efforts in the next reanalysis iteration.

Further improvements to our system would involve the following. Increasing the number of EDA members to reduce sampling error would be a great first step, but this is contingent on computing capability. Partitioning the scales where the EDAs can influence the B-matrix, for example, not allowing the daily 11km EDA to have any input on the 5.5km scales. This would ensure that no irrelevant information is used. This could perhaps be achieved by a scale-dependent function to allow 100% weighting to the 5.5km scales coming from the seasonal component. Finally, EnVar techniques would probably provide scope for further improvement. However, this would require extensive testing in a reanalysis context, to ensure that an improvement over non-EnVar techniques which are currently used.

Acknowledgments

The funding for the CERRA project comes from the Copernicus Climate Change Services (C3S_322 Lot 1) contract. There are no conflicts of interest. The authors would like to thank the entire CERRA team without which this project wouldn't have happened: Eric Bazile, Patrick Le Moigne, Per Unden, Esbjörn Olsson, Semjon Schimanke, Per Dahlgren and Lars Berggren.

Data Availability Statement

The input and output data of the experiments described in the paper are freely available for research purposes from ECMWF and can be requested following the procedures described online (at https://www.ecmwf.int/en/forecasts/datasets).

References

Bannister, R. N. (2008). A review of forecast error covariance statistics in atmospheric variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, *134* (637), 1951–1970.

Bartello, P., & Mitchell, H. L. (1992). A continuous three-dimensional model of short-range forecast error covariances. *Tellus A: Dynamic Meteorology and Oceanography*, *44* (3), 217–235.

Bengtsson, L., Kanamitsu, M., Kéllberg, P., & Uppala, S. (1982). Fgge research activities at ecmwf. *Bulletin of the American Meteorological Society*, *63* (3), 277–303.

Berre, L. (2000). Estimation of synoptic and mesoscale forecast error covariances in a limited-area model. *Monthly weather review*, *128* (3), 644–667.

Bollmeyer, C., Keller, J., Ohlwein, C., Wahl, S., Crewell, S., Friederichs, P., ... others (2015). Towards a high-resolution regional reanalysis for the european cordex domain. *Quarterly Journal of the Royal Meteorological Society*, *141* (686), 1–15.

Bonavita, M., Isaksen, L., & Hólm, E. (2012). On the use of eda background error variances in the ecmwf 4d-var. *Quarterly journal of the royal meteorological society*, *138* (667), 1540–1559.

Bromwich, D. H., & Wang, S.-H. (2005). Evaluation of the ncep–ncar and ecmwf 15-and 40-yr reanalyses using rawinsonde data from two independent arctic field experiments. *Monthly Weather Review*, *133* (12), 3562–3578.

Bromwich, D. H., Wilson, A. B., Bai, L.-S., Moore, G. W., & Bauer, P. (2016). A comparison of the regional arctic system reanalysis and the global era-interim reanalysis for the arctic. *Quarterly Journal of the Royal Meteorological Society*, *142*(695), 644–658.

Brousseau, P., Berre, L., Bouttier, F., & Desroziers, G. (2012a). Flow-dependent background-error covariances for a convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *138*(663), 310–322.

Brousseau, P., Berre, L., Bouttier, F., & Desroziers, G. (2012b). Flow-dependent background-error covariances for a convective-scale data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *138*(663), 310-322. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.920` doi: 10.1002/qj.920

Bruyninx, C., Legrand, J., Fabian, A., & Pottiaux, E. (2019). Gnss metadata and data validation in the euref permanent network. *GPS Solutions*, *23*(4), 1–14.

Cardinali, C., Pezzulli, S., & Andersson, E. (2004). Influence-matrix diagnostic of a data assimilation system. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *130*(603), 2767–2786.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., . . . Vitart, F. (2011). The era-interim reanalysis: configuration and performance of the data assimilation system. *Quarterly Journal of the Royal Meteorological Society*, *137*(656), 553-597. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.828` doi: 10.1002/qj.828

Derber, J., & Bouttier, F. (1999). A reformulation of the background error covariance in the ecmwf global data assimilation system. *Tellus A: Dynamic Meteorology and Oceanography*, *51*(2), 195–221.

El-Said, A. (2015). *Conditioning of the weak-constraint variational assimilation problem for numerical weather prediction* (Unpublished doctoral dissertation). The University of Reading.

Fisher, M. (2003). Background error covariance modelling. In *Seminar on recent development in data assimilation for atmosphere and ocean* (pp. 45–63).

Häggmark, L., Ivarsson, K.-I., Gollvik, S., & Olofsson, P.-O. (2000). Mesan, an operational mesoscale analysis system. *Tellus A: Dynamic Meteorology and Oceanography*, *52*(1), 2–20.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horényi, A., Muñoz-Sabater, J., . . . Thépaut, J.-N. (2020). The era5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.3803` doi: 10.1002/qj.3803

Hollingsworth, A., & Lönnberg, P. (1986). The statistical structure of short-range forecast errors as determined from radiosonde data. part i: The wind field. *Tellus A*, *38*(2), 111–136.

Isaksen, L., Fisher, M., & Berner, J. (2007). Use of analysis ensembles in estimating flow-dependent background error variance. In *Proc. ecmwf workshop on flow-dependent aspects of data assimilation* (pp. 65–86).

Mahmood, S., Davie, J., Jermey, P., Renshaw, R., George, J. P., Rajagopal, E., & Rani, S. I. (2018). Indian monsoon data assimilation and analysis regional reanalysis: Configuration and performance. *Atmospheric Science Letters*, *19*(3), e808.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., . . . others (2013). The surfexv7. 2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes.

Mesinger, F., DiMego, G., Kalnay, E., Mitchell, K., Shafran, P. C., Ebisuzaki, W., . . . others (2006). North american regional reanalysis. *Bulletin of the American Meteorological Society*, *87*(3), 343–360.

Mirouze, I., & Weaver, A. (2010). Representation of correlation functions in variational assimilation using an implicit diffusion operator. *Quarterly Journal of the Royal Meteorological Society*, *136*(651), 1421–1443.

Nocedal, J. (1980). Updating quasi-newton matrices with limited storage. *Mathematics of computation*, *35*(151), 773–782.

Palmer, T. N. (2001). A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal of the Royal Meteorological Society*, *127*(572), 279–304.

Parrish, D. F., & Derber, J. C. (1992). The national meteorological center's spectral statistical-interpolation analysis system. *Monthly Weather Review*, *120*(8), 1747–1763.

Parrish, D. F., Derber, J. C., Purser, R. J., Wu, W.-S., & Pu, Z.-X. (1997). The ncep global analysis system: Recent improvements and future plans (gtspecial issueltdata assimilation in meteology and oceanography: Theory and practice). *Journal of the Meteorological Society of Japan. Ser. II*, *75*(1B), 359–365.

Phillips, N. A. (1986). The spatial statistics of random geostrophic modes and first-guess errors. *Tellus A: Dynamic Meteorology and Oceanography*, *38*(4), 314–332.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., ... others (2016). Era-20c: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083–4097.

Purser, R. J., Wu, W.-S., Parrish, D. F., & Roberts, N. M. (2003). Numerical aspects of the application of recursive filters to variational statistical analysis. part i: Spatially homogeneous and isotropic gaussian covariances. *Monthly Weather Review*, *131*(8), 1524–1535.

Vautard, R. (1990). Multiple weather regimes over the north atlantic: Analysis of precursors and successors. *Monthly weather review*, *118*(10), 2056–2081.

Wang, Z. Q., & Randriamampianina, R. (2021). The impact of assimilating satellite radiance observations in the copernicus european regional reanalysis (cerra). *Remote Sensing*, *13*(3), 426.

Zagar, N., Andersson, E., & Fisher, M. (2005). Balanced tropical data assimilation based on a study of equatorial waves in ecmwf short-range forecast errors. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, *131*(607), 987–1011.