Section 1

Assimilation of atmospheric and land observations.
Data impact and sensitivity studies.
Methodological advances.

Optimal interpolation of inhomogeneous fields using neural networks

Ph.L.Bykov

Hydrometeorological Research Center of Russian Federation, e-mail: bphilipp@inbox.ru.

Introduction. In this note we are concerned with the problem of interpolation of a random field $f(\vec{x}), \vec{x} \in \Omega$ from a scattered set of SYNOP stations $\vec{x}_j, j = 1...m$ onto another set of points $\vec{y}$ in a domain Ω on Earth surface. We will be using the moments of the field $f(\vec{x})$: the mean $\mu(\vec{x})$, the spread $\sigma(\vec{x})$, and the correlation function (CF) $K(\vec{x}, \vec{y})$. It is also natural to assume that the interpolation result $\hat{f}$ depends linearly on the interpolated data:

$$f(\vec{y}) \approx \hat{f}(\vec{y}) = w_x(\vec{y})^T [f_x - \mu_x] + \mu(\vec{y}), \qquad (1)$$

where $w_x(\vec{y})$ is the (unknown) column of the interpolation weights we are looking to compute, f_x is column of known values of f. We can state the optimization problem as a problem of minimizing the mean interpolation loss L:

$$L(\mu, \sigma, K) = \sum_{j=1}^{m} e\left(f(\vec{x}_j), \hat{f}_j(\vec{x}_j)\right) \to \min_{\mu, \sigma, K}, \qquad (2)$$

where e is the loss function and $\hat{f}_j$ is the interpolating operator computed from the same set of data without taking into account the value $f(\vec{x}_j)$. The classical optimal interpolation method [1] uses the mean squared error as loss function and computes the interpolation weights from the equation:

$$K_{xx}\left[\sigma_x \circ w_x(\vec{y})\right] = K_{xy} \circ \sigma(\vec{y}),$$

where K_{xx} – the positive definite $m \times m$ correlation matrix, $\circ$ stands for the Hadamard product.

The problem of estimating the correlation function K is easy to solve for homogeneous and isotropic fields f, i.e. in the case $K(\vec{x}, \vec{y}) = K(|\vec{x} - \vec{y}|)$. The case of inhomogeneous fields that we address here is more complicated. It follows from Mercer's theorem [2], [3] that the feature mapping $g : \Omega \to H$ exists for any field f, where H is Hilbert space. Furthermore, the feature mapping g turns inhomogeneous anisotropic field f into a homogeneous isotropic field in a Hilbert space H:

$$K(\vec{x}, \vec{y}) = K_g\left(\|g(\vec{x}) - g(\vec{y})\|_H\right).$$

Several papers [4], [5], [6] suggest methods for approximation of the feature mapping g and maximization of the logarithm of the likehood in Gaussian covariance model:

$$L_{Gauss}(g, K_g) = -\frac{1}{2}\left[\bar{f}_x^T K_{xx}^{-1} \bar{f}_x + \ln \det K_{xx}\right] \to \max_{g, K_g}, \qquad (3)$$

where $\bar{f} = (f - \mu)/\sigma$ is the normalized field. However, this choice is questionable, since the interpolation function (1) and the functional (2) depend only on the local oscillations of the means $\mu(\vec{x}) - \mu(\vec{y})$ and the spreads $\sigma(\vec{x})/\sigma(\vec{y})$, while the functional (3) depends on their absolute values.

Approach. Let $\Theta(\vec{x})$ be a set of predictors of inhomogeneity of the interpolated field f. For minimizing the interpolation loss (2) we shall apply the backpropagation method [7] to the following functions, described by the neural networks and parameters:

1. The feature mapping g to the 4-dimensional space H as a graphic of the neural network $\tilde{g}$:

$$g(\vec{x}) = (\vec{x}, \tilde{g}(\Theta));$$

2. The mean $\mu(\vec{x}) = \mu(\Theta(\vec{x}))$ and the spread $\sigma(\vec{x}) = \sigma(\Theta(\vec{x}))$ as the neural networks;

3. The parameters $\vec{\upsilon} = (\varepsilon, \beta, R_1, R_2)$ of family of CFs:

$$K_g(r, \vec{\upsilon}) = \varepsilon I_0(r) + (1 - \varepsilon)\left[(1 - \beta)(1 + r/R_1)\exp(-r/R_1) + \beta \exp(-r^2/2R_2^2)\right],$$

where $r = |g(\vec{x}) - g(\vec{y})| = \sqrt{|\vec{x} - \vec{y}|^2 + |\tilde{g}(\vec{x}) - \tilde{g}(\vec{y})|^2}$ – the distance in extended space H.

All neural networks we consider are two layers perceptrons with the ReLU activation function and 32 neurons at hidden layer. The Huber loss function [8] is chosen.

Examples. We shall demonstrate our approach for two fields of particular interest: 1) the 2-meter temperature *T2m* model biases; 2) the snow depth *SD*. We consider the following set of predictors of inhomogeneity $\vec{\Theta}$:

A. The first guess field (for *T2m* only; the COSMO-Ru model [9] forecast with lead time 6 hours);
B. The Earth surface altitude;
C. The sine and cosine of the Julian day $2\pi t / T_y$, where t is time, $T_y = 1$ year;

D. The elevation angle of the Sun.

The feature mapping g has to be injection and should take into account all predictors of inhomogeneity Θ.

Results. Adding additional dimensions to the feature mapping allows us to explain most part of the variance of the fields *T2m* and *SD*: the $\lim_{\vec{y} \to \vec{x}} K(\vec{x}, \vec{y})$ is significantly larger (see figure). Moreover, the presented method gives

more accurate and more detailed fields: since due to our choice of the feature mapping the CFs decrease much faster. Finally, the backpropagation for interpolation losses allows us to avoid extra assumptions on the field *f*.
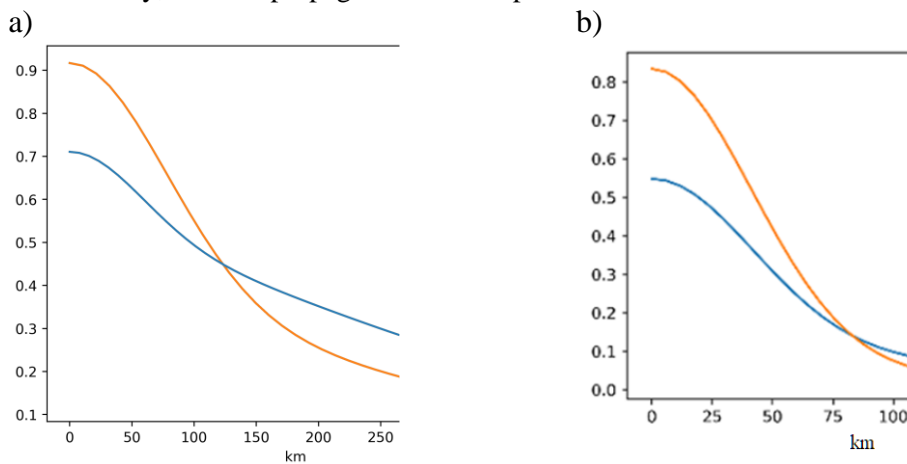
a)   b)

Figure. The CFs for a) *T2m* model biases and b) snow depth *SD* as the functions of the distance (x-axis). Orange curves correspond to the extended 4-dimensional space *H* and the blue curves correspond to the real space with extra assumptions of homogeneity and isotropy.

References
[1] Gandin L.S., Objective Analysis of Meteorological Fields, Jerusalem, 1965, p. 896.
[2] Mercer J., "Functions of positive and negative type and their connection with the theory of integral equations," *Philosophical Transactions of the Royal Society A,* vol. 209, no. 441-458, pp. 415-446, 1909.
[3] Aronszajn N., "Theory of reproducing kernels," *Trans. Amer. Math. Soc.,* vol. 68, no. 3, pp. 337-404, 1950.
[4] Sampson P., Guttorp P., "Nonparametric estimation of nonstationary spatial covariance structure," *J. Amer. Statist. Assoc.,* vol. 87, p. 108–119, 1992.
[5] Bornn L., Shaddick G., Zidek J.V., "Modeling Nonstationary Processes Through Dimension Expansion," *Journal of the American Statistical Association,* vol. 107, no. 497, pp. 281-289, 2012.
[6] Zammit-Mangion A., Tin L. J. Ng, Vu Q., Filippone M., "Deep Compositional Spatial Models," 2019. https://arxiv.org/pdf/1906.02840.pdf.
[7] Rumelhart D.E., Hinton G.E., Williams R.J., "Learning Internal Representations by Error Propagation. In: Parallel Distributed Processing," vol. 1, Cambridge, MIT Press, 1986, pp. 318-362.
[8] Huber P.J., "Robust Estimation of a Location Parameter," *Ann. Math. Statist.,* vol. 35, no. 1, pp. 73-101, 1964.
[9] Rivin G.S., Rozinkina I.A., Vil'fand R.M., Alferov D.Yu., Astakhova E.D., Blinov D.V., Bundel' A.Yu., Kazakova E.V., Kirsanov A.A., Nikitin M.A., Perov V.L., Surkova G.V., Revokatova A.P., Shatunova M.V., Chumakov M.M., "The COSMO-Ru system of nonhydrostatic mesoscale short-range weather forecasting of the Hydrometcenter of Russia: The second stage of implementation and development" *Russ. Meteorol. Hydrol.,* vol. 40, no. 6, pp. 400-410, 2015.

Replacing OPAC with MERRA2 Aerosols for the UFS

Anning Cheng[1] and Fanglin Yang[2]
[1.]IMSG@NOAA/NCEP/EMC [2.]NOAA/NCEP/EMC

1. Motivation

Aerosols play an important role in the energy budget of the Earth-atmosphere system. They directly scatter and absorb electromagnetic radiation (aerosol direct effect) and indirectly interact with cloud macro- and micro-physics (aerosol indirect effect). Non-absorbing aerosols, such as sulfate and organic carbon, scatter solar radiation back to space producing a cooling effect on the climate system. Aerosols also contain absorbing material, such as black carbon, that absorbs solar radiation leading to a warming effect that partly offsets the aerosol scattering effect.

NOAA is currently developing the next generation coupled Unified Forecast System (UFS) for weather, sub-seasonal and seasonal predictions. The aerosols prescribed in the UFS are still adopted from the Optical Properties of Aerosols and Clouds (OPAC) dataset (Hess *et al.*, 1998), which only represents aerosol distributions in 1990's. It also has a very coarse 5x5-deg horizontal resolution. Aerosol concentration is a mixture of different predefined components. An exponential profile is used to define the distribution of aerosol particles with height.

To improve the representation of aerosols in the UFS, the Modern-Era Retrospective Analysis for Research and Applications Version 2 (MERRA2) will be used to replace OPAC. MERRA2 was produced using the three-dimensional variational data analysis (3DVAR) Gridpoint Statistical Interpolation (GSI) meteorological analysis scheme and the GEOS-5 atmospheric model. The MERRA-2 meteorological observing system includes Moderate Resolution Imaging Spectroradiometer (MODIS), the Advanced Very High Resolution Radiometer (AVHRR) instruments, Multi-angle Imaging SpectroRadiometer (MISR), ground-based Aerosol Robotic Network (AERONET), and other bias-corrected observational data as well. MERRA2 has a horizontal resolution of 0.5° latitude by 0.625° longitude, and has 72 levels in the vertical extending from the surface to one pascal. So it has a much higher spatial resolution than OPAC. MERRA2 aerosols consist of 15 modes, five bins of dust and sea salts, two bins of organic carbon and black carbon, and one bin of sulfate, respectively. Described in the following is an evaluation of medium-range weather forecast skill in the UFS atmospheric model, the Global Forecast System version 16 (GFS.v16), with OPAC aerosols being repealed by MERRA2.

2. Results

GFS.v16 has a horizontal resolution of 13 km and has 127 layers in the vertical extending from the surface to the mesopause. To assess the impact of replacing OPAC with MERRA2 aerosols on GFS forecast skill, two control experiments with OPAC aerosols and two sensitivity experiments with MERRA2 aerosols were conducted for one summer and one winter case, respectively. The winter case covers the period from 1 December 2019 through 1 March 2020, and the summer case for the period from 1 June 2019 through 1 September 2019. Each experiment is initialized with GFS.v16 initial conditions at the 00Z cycle, once for every 5 days in each winter and summer time period.

The most notable effects are changes in radiative warming/cooling at the surface and the top of the atmosphere (TOA) induced by aerosol reflection and extinction. MERRA2 has less aerosol loading than OPAC over most of the continents due to a reduction in anthropological emissions, and much less sea salt in the Southern and Northern Hemisphere storm track regions. On the other hand, MERRA2 dust aerosol loading from the Sahara desert is much higher than OPAC over northwest Africa. Over East Asia, MERRA2 organic carbon and sulfate aerosol loading are also higher. Normally large aerosol number concentrations are only found in the lower troposphere. Shortwave radiation reflected back by aerosols warms the atmosphere above the aerosol layer. The extinction of shortwave radiation by the aerosol layer also leads to reduced

shortwave (SW) flux at the surface. The reduction can reach as large as 30 W/m^2 over northwest Africa. The absorption of longwave radiation by black carbon and large particles of dust, etc. can warm the aerosol layer and thus increase the outgoing longwave radiation (OLR) at the top of the atmosphere and the downward longwave (LW) flux at the surface. Overall, the magnitude of aerosol longwave radiation forcing is much smaller than that of aerosol shortwave radiation forcing.

It was found from this study that Asian monsoon was also more realistically depicted with MERRA2 aerosols. Aerosols over the Arabian peninsula and Indian subcontinent cause lower troposphere cooling and upper troposphere warming. With MERRA2 aerosols, the sea level pressure increased over the continent and decreased over the ocean in the monsoon region. The Monsoon circulation was stronger and compared better with observations.

The evaluation of medium-range weather forecast skill shows that 500-hPa height anomaly correlations are improved for the Northern Hemisphere and the Pacific and North America (PNA) regions in both winter and summer (Figure 1), with some of the scores passing the 95% confidence Student-t test.
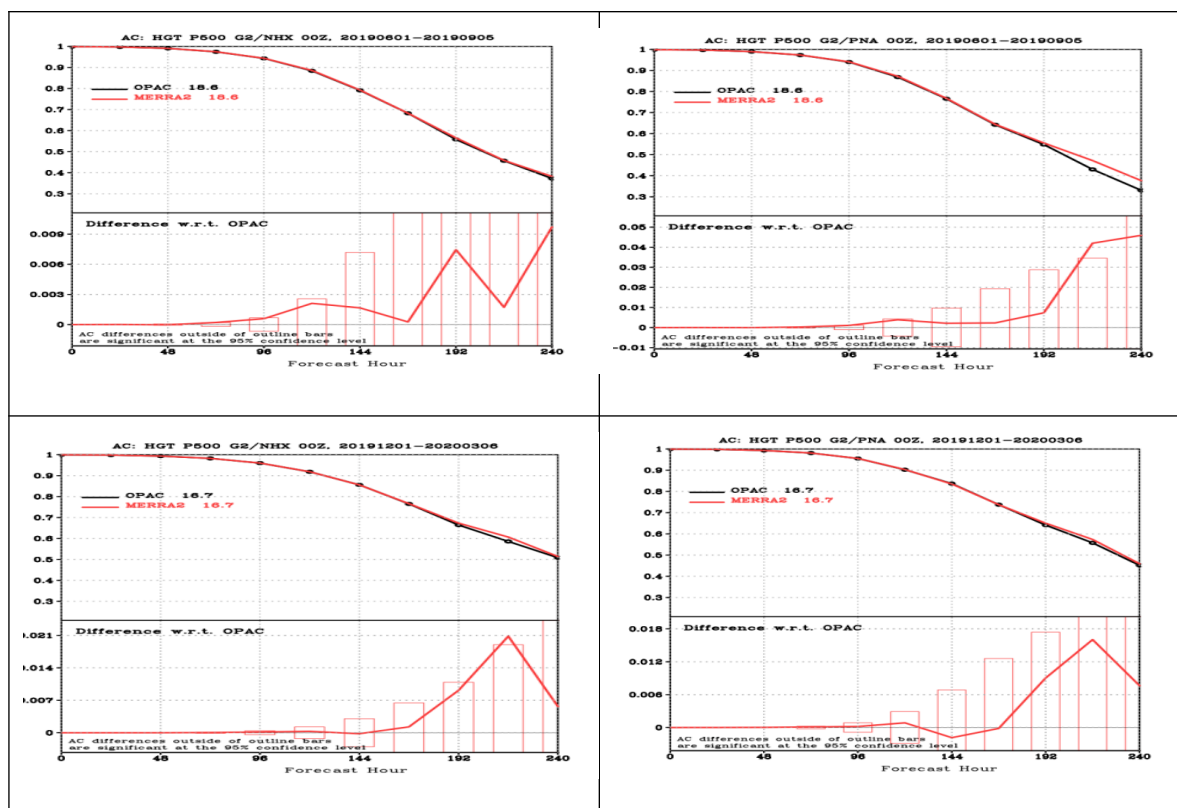

Figure 1. 500-hPa height anomaly correlations for the Northern Hemisphere (left panels) and the Pacific North America (right panels) in summer (upper row) and winter (lower rows), respectively.

References

Buchard,V., C. A. Randles, A. M. da Silva, and et al. 2017, The MERRA-2 aerosol reanalysis, 1980 onward. Part II: Evaluation and case studies. *J. Climate*, **30**,6851–6872.

Hess M., P. Koepke, and I. Schult, 1998, Optical Properties of Aerosols and Clouds: The Software Package OPAC, *Bull. Am. Meteor. Soc.*, **79**, 831-844.

Randles, C. A., A. M. da Silva, V. Buchard, and et al. 2017, The MERRA-2 aerosol reanalysis, 1980 onward. Part I: System description and data assimilation evaluation. *J. Climate*, **30**,6823–6850.

Identifying Atmospheric River Reconnaissance Targets Using Ensemble Forecasts

Travis J. Elless[1*], Xingren Wu[1], and Vijay Tallapragada[2]
[1]I.M. Systems Group at NOAA/NWS/NCEP/EMC [2]NOAA/NWS/NCEP/EMC
*travis.j.elless@noaa.gov

1. Introduction

Model forecasts exhibiting large variability are a detriment to a forecaster's confidence. This can be extremely problematic for large-impact weather events, such as precipitation forecasts in the drought prone Western United States. The Atmospheric River (AR) Reconnaissance field campaign (Ralph *et al*. 2020) provides additional observations in upstream regions/features exhibiting high ensemble forecast variability, with the goal of reducing short-range West Coast precipitation forecast uncertainty. Due to time constraints, mission planning generally uses 00 UTC forecasts to assess variability in the 60-84-h precipitation forecasts and subsequently identify sensitive areas (targets) in model fields at 48 h. Since 2019, the Environmental Modeling Center (EMC) has developed tools to provide potential mission targets to the field campaign, based on ensemble sensitives calculated from a combined 21-member (upgraded to 31-member in 2021) Global Ensemble Forecast System and 21-member Canadian Ensemble Prediction System super ensemble. This report highlights the methodology used for this tool, and provides example outputs that forecasters utilize.

2. EMC ensemble sensitivity tool

The EMC ensemble sensitivity tool utilizes a composite difference approach to identify upstream sensitive regions that are linked to West Coast precipitation forecast variability. Composite differences have been previously used to identify processes and features associated with large ensemble variability in forecast tropical cyclone (TC) position (Torn *et al*. 2015), TC intensity (Rios-Berrios *et al.* 2016), midlatitude cyclone intensity (Lamberson *et al*. 2016), and 500-hPa height errors (Magnusson 2017). Utilizing composite differences for this particular task first requires the identification of subsets of ensemble members based on the 60-84-h West Coast precipitation forecasts. For a given forecast, the standard deviation (SD) in accumulated precipitation is calculated for each grid point and the location with the maximum value is identified. All adjacent grid points within 66% of this maximum SD value are then identified, and precipitation amounts are averaged over this area. Figure 1 provides an example of this identification for a forecast initialized 00 UTC 20 Feb 2021. This averaging allows for sensitivity impacts to be representative of a larger geographical area, instead of focusing on individual grid point values. Ranking the area-averaged precipitation allows the creation of two subsets consisting of the 10 most/least precipitating ensemble members.

For this project, differences between these two subsets at 48 h are evaluated at each individual grid point via

$$\Delta x_i = \frac{\bar{x}_i^{Highest} - \bar{x}_i^{Lowest}}{\sigma_{x_i}}, \tag{1}$$

where Δx_i represents the composite difference for a given state variable (this field project primarily utilizes integrated vapor transport (IVT), 850-hPa equivalent potential temperature (θ_e), and 500-hPa potential vorticity), which is used as a measure of ensemble sensitivity, $x_i^{Highest}$ (x_i^{Lowest}) is the subset mean for the given state variable, and σ_{xi} is the SD of the state variable utilizing all ensemble members. Normalizing the difference by the SD allows for a direct comparison to be made between different model variables and levels to determine fields that are the most sensitive. Differences are found to be statistically significant at the 95% confidence interval from a 1000 iteration bootstrap resampling without replacement process; statistical significance was calculated similarly to Rios-Berrios *et al.* (2016). Figure 2 shows an example of the 48-h composite difference evaluation of 850-hPa θ_e for the 00 UTC 20 Feb 2021 initialization. Differences are primarily found on the periphery of a thermal ridge, suggesting the precipitation variability found in Washington is sensitive to the position of this ridge, such that a more meridionally oriented ridge is associated with increased precipitation in Washington; this feature was subsequently deemed a viable target for additional observations.
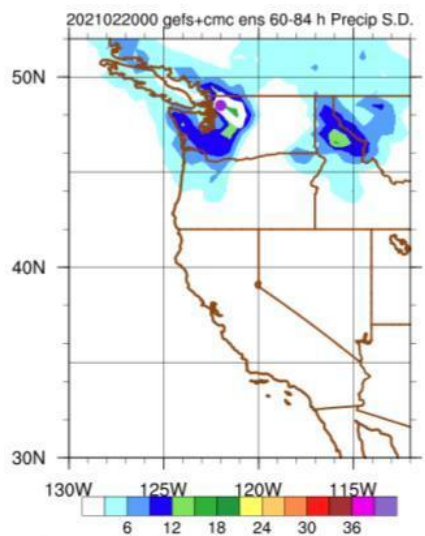
Figure 1. 60-84-h accumulated precipitation standard deviation (color fill; units mm), grid point of maximum standard deviation (purple dot), and area bound by 66% of maximum standard deviation (white line)
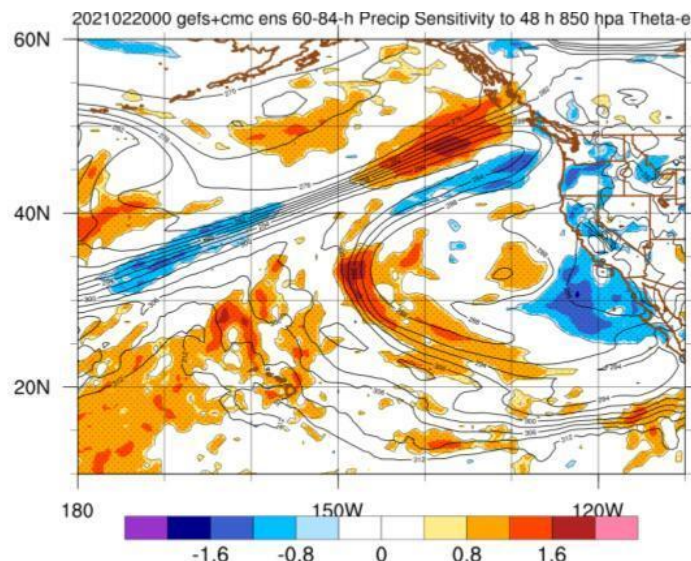

Figure 2. Ensemble mean 48-h 850-hPa equivalent potential temperature (contours) and composite difference between high and low subsets defined for the region in Fig. 1 (color fill; units standard deviation). All differences shown are statistically significant at the 95% confidence interval

3. Other applications

This methodology can be adapted to test the sensitivity of any given metric to any model field, and is a vital tool in understanding processes leading to ensemble variability. For example, this project has also identified targets using subsets created from IVT and mean sea level pressure variability. Future work includes identifying subsets from principal components of precipitation empirical orthogonal function (EOF) analysis, and utilizing this tool to identify sensitive areas associated with operational dropout events.

References

Lamberson, W. S., R. D. Torn, L. F. Bosart, and L. Magnusson, 2016: Diagnosis of the source and evolution of medium-range forecast errors for extratropical cyclone Joachim. *Wea. Forecasting,* **31,** 1197-1214.

Magnusson, L., 2017: Diagnostic methods for understanding the origin of forecast errors. *Quart. J. Roy. Meteor. Soc.,* **143,** 2129-2142.

Ralph, F. M., and Coauthors, 2020: West Coast forecast challenges and development of atmospheric river reconnaissance. *Bull. Amer. Meteor. Soc.,* **101,** E1357-E1377.

Rios-Berrios, R., R. D. Torn, and C. A. Davis, 2016: An ensemble approach to investigate tropical cyclone intensification in sheared environments. Part I: Katia (2011). *J. Atmos. Sci.,* **73,** 71-93.

Torn, R. D., J. S. Whitaker, P. Pegion, T. M. Hamill, and G. J. Hakim, 2015: Diagnosis of the source of GFS medium-range track errors in Hurricane Sandy (2012). *Mon. Wea. Rev.,* **143,** 132-152.

Operational Use of GOES-16 Atmospheric Motion Vector (AMV) and ScatSat-1/OSCAT Data in JMA's Global NWP System

NONAKA Kenichi[1] and KOYAMATSU Shin[2]
[1]Numerical Prediction Development Center, Japan Meteorological Agency
[2]Meteorological Satellite Center, Japan Meteorological Agency
E-mail: k-nonaka@met.kishou.go.jp

1. Introduction

Atmospheric Motion Vectors (AMVs) are derived by tracking clouds and water vapor patterns from sequential satellite images, and provide information on tropospheric wind. Scatterometer wind data, such as Advanced SCATterometer (ASCAT) wind data, provide ocean surface wind vectors retrieved from microwave backscatter irradiating toward the sea surface. AMVs and scatterometer wind data are especially important for ocean areas, where in-situ observations are sparse, and are used for data assimilation in operational NWP system worldwide.

GOES-16 AMV and ScatSat-1/OSCAT data were adopted in data assimilation for JMA's global NWP system (GSM) at 00 UTC on 29 July 2020. This report outlines the results of verification experiments using these data in the GSM.

2. GOES-16 AMV

The new-generation GOES-16 geostationary meteorological satellite launched in November 2016 operates in the GOES-East role at 75.2°W, covering the area from the eastern Pacific to the western Atlantic. The satellite is equipped with the Advanced Baseline Imager (ABI) featuring upgraded observation functionality over the previous satellite GOES-13 imager, especially in terms of scanning frequency, spatial resolution and multispectral bands. GOES-16 AMVs are improved, with upgraded imagery and a new derivation algorithm developed for cutting-edge satellites (Daniels et al. 2019).

3. OSCAT Wind

ScatSat-1 equipped with OSCAT was launched in September 2016 and began providing OSCAT wind data in September 2017. Metop series ASCAT wind data are already utilized in the GSM. OSCAT involves the use of a different microwave frequency and scanning technique from ASCAT, and can scan wide regions but is more affected by raindrop attenuation. Based on research using OSCAT wind data and first-guess statistics, ASCAT preprocessing is applied to data usage in assimilation.

4. Assimilation Experiments

Observing-system experiments assimilating both GOES-16 AMV and OSCAT wind data in the GSM were performed to verify effects on analysis and forecast fields. The experiments were performed in an environment equivalent to the operational JMA system in December 2019, and were validated over periods of several months in summer (10 Jun. – 11 Oct. 2019) and winter (10 Nov. 2019 – 11 Mar. 2020). The control experiments (CNTL) had the same configuration as the operational set-up, and the test experiments (TEST) were performed using GOES-16 AMV and OSCAT wind data. Figure 1 shows AMV coverage and the scatterometer wind data utilized in the global analysis for 00 UTC on January 1, 2020. In the figures below, magenta and deep green indicate the GOES-16 AMV and OSCAT wind data used in the analysis, respectively.

5. Verification Results

Figure 2 shows normalized changes in standard deviation against the first guess in the rawinsonde u/v-component wind data applied in the summer and winter experiments. The first-guess wind field for the troposphere changes neutrally or improves slightly to match the rawinsonde observation. Figure 3 is as per Figure 2, but for microwave sounding observation. Here too, the first-guess field is modified to be more consistent with observation from the humidity sounding channels, which are sensitive to upper air (MHS Ch. 3 and ATMS Ch. 21/22), and temperature sounding channels sensitive to lower air (AMSU-A Ch. 5 and ATMS Ch. 6). The results suggest that the upper-troposphere circulation field and lower-air convergence/divergence positions in the first guess are improved by assimilation of GOES-16 AMV and OSCAT wind data.

Figure 4 shows verification of effects on tropical cyclone track prediction in summer 2019 for the Atlantic and eastern Pacific regions using best-track data provided by NOAA (B-decks) for reference. The predicted position errors are seen to be reduced within approximately 72 forecast

hours based on GOES-16 AMV and OSCAT wind data assimilation.

6. Summary

The research reported here verified that analysis field consistency with microwave sounders is improved and the accuracy of tracking prediction for tropical cyclones is slightly modified by the use of GOES-16 AMV and OSCAT wind data in the global NWP system. Both data have been used operationally in GSM data assimilation since 00 UTC on 29 July 2020. OSCAT wind data are also being considered for use in JMA's meso and local scale NWP systems.

References

Daniels, J., et al., (2019). GOES-R Advanced Baseline Imager (ABI) Algorithm Theoretical Basis Document For Derived Motion Winds, Version 3.1.
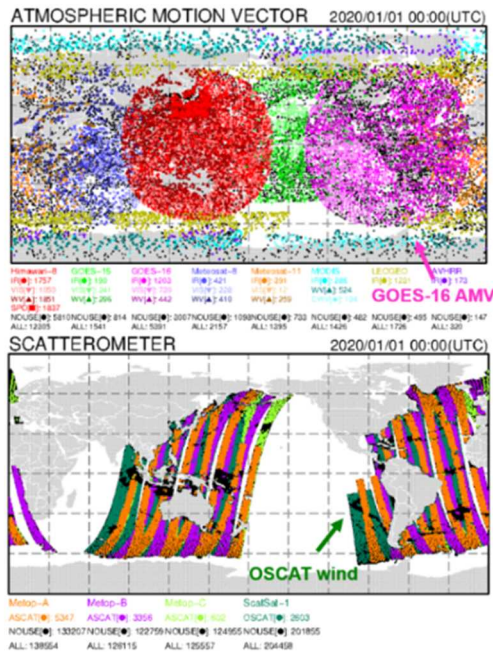
Figure 1. Data coverages of AMV (top) and scatterometer wind data (bottom) used in global analysis at 00 UTC on January 1, 2020. GOES-16 AMVs and OSCAT wind data are plotted in magenta and deep green, respectively.
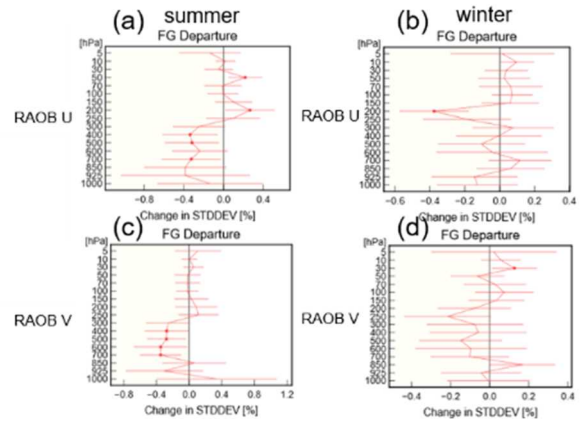

Figure 2. Normalized changes in standard deviation (STDDEV [%]) for first-guess departures in u/v components of rawinsonde (RAOB U/V). a, c and b, d are validated for summer (10 Jun. – 11 Oct. 2019) and winter (10 Nov. 2019 – 11 Mar. 2020), respectively. Error bars represent a 95% confidence interval, and dots represent statistical significance.
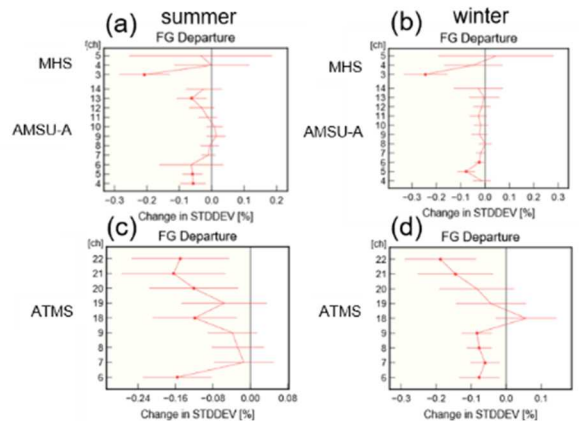

Figure 3. Normalized changes in standard deviation (STDDEV [%]) for first-guess departures in microwave sounding data for individual channels [ch]. a, c and b, d are validated for summer (10 Jun. – 11 Oct. 2019) and winter (10 Nov. 2019 – 11 Mar. 2020), respectively. Error bars represent a 95% confidence interval, and dots represent statistical significance.
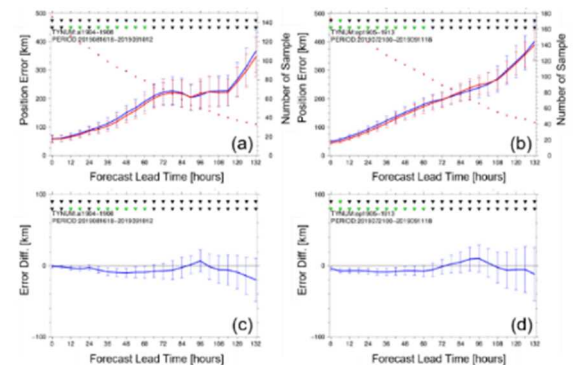

Figure 4. Average track forecast error of tropical cyclones in summer 2019 for (a) the Atlantic and (b) the eastern Pacific. Red and blue lines indicate positional errors in TEST and CNTL, respectively, and red dots indicate the number of samples. The bottom figures show position error differences between TEST and CNTL (TEST-CNTL) for (c) the Atlantic and (d) the eastern Pacific. Error bars represent a 95% confidence interval, and triangles at the top indicate statistical significance (green: significant; black: not significant).

Operational use of Himawari-8 CSR data of band 9 and 10 in JMA's local NWP system

OKABE Izumi
Meteorological Research Institute, Japan Meteorological Agency
E-mail: i-okabe@mri-jma.go.jp

1. Introduction

The Japan Meteorological Agency (JMA) began to use Himawari-8 surface-sensitive band 9 and 10 (6.9 and 7.3 µm) clear-sky radiance (CSR) data in its local NWP system on July 29, 2020 in addition to Band 8 (6.2 µm) CSR data (Ikuta 2017). Experiments indicated that this assimilation with a new radiative transfer (RT) calculation method for JMA's global NWP system (Okabe 2019) had positive impacts on water vapor (WV) field first-guess (FG) and precipitation forecast scores for heavy rain in the local NWP system. The results are reported here.

2. Methodology

The new RT calculation method is the same as that of the global NWP system (Okabe 2019). The land surface emissivity atlas of Wisconsin University (Borbas and Ruston 2010) and retrieved land surface temperatures from window band 13 (10.8 µm) CSR observation data are used in the calculation.

3. Assimilation Experiment

The control experiment performed (referred to here as CNTL) had the same configuration as the operational JMA local NWP system as of March 2020. The test experiment (TEST) was as per CNTL, but surface-sensitive CSRs from Himawari-8 (band 9 and 10) were additionally assimilated. The experiment periods were from June 27 to July 8, 2018 (referred to as summer) and from January 16 to 27, 2018 (referred to as winter).

4. Impacts on the NWP System

Figure 1 shows normalized changes in the standard deviation (STDDEV) of the FG departure (difference between observations and FG) for microwave sounder, microwave imager data and radiosonde observations, which contain information on WV and temperature in the troposphere. The reductions seen indicate improved fitting between FG and other observations, implying improved FG field accuracy.

Figure 2 shows bias scores, threat scores, false alarm ratios and undetected error rates for three-hour cumulative precipitation forecasts in the summer experiment. Although there were only slight differences between TEST and CNTL, false alarm ratio and undetected error rate values were reduced for heavy-rain thresholds (25 mm or more), and slight improvements were seen in threat scores.

5. Summary

JMA began to assimilate surface-sensitive CSRs from Himawari-8 (bands 9 and 10) in the local NWP system on July 29, 2020, and the new RT calculation method used in JMA's global NWP system was applied. Positive impacts from these CSRs on WV and temperature field accuracies for the first guess in the local NWP system were shown in the assimilation experiment, which also revealed improved precipitation forecasting scores for heavy rain.

References

Borbas, E. E. and Ruston, B. C. (2010). The RTTOV UWiremis IR land surface emissivity module, AS Mission Report NWPSAF-MOVS-042, EUMETSAT Numerical Weather Prediction Satellite Applications Facility, 24pp.

Okabe, I., 2019: Operational use of surface-sensitive Clear-Sky Radiance (CSR) data in JMA's Global NWP System. CAS/JSC WGNE Research Activities in Atmospheric and Oceanic Modelling, Rep. 49, 1.15-1.16.

Ikuta, Y., 2017: Assimilation of Satellite Soil Moisture Contents and Clear-sky Radiance in Operational Local NWP System at JMA. JpGU-AGU Joint Meeting 2017.
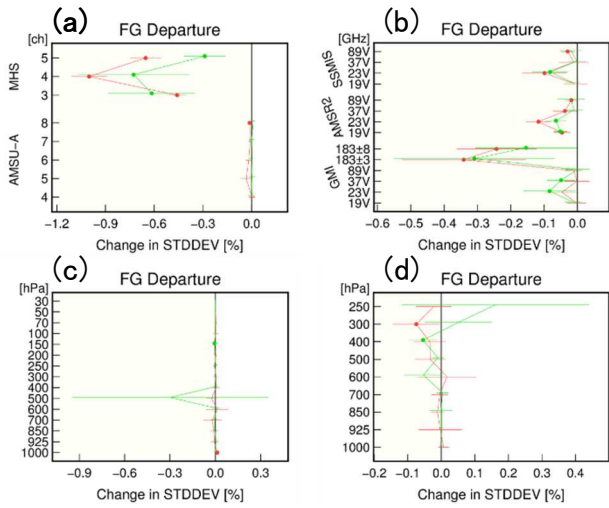
Figure 1. Normalized changes in standard deviation (STDDEV) for first-guess departures in microwave sounding data for individual channels [ch] (a), microwave imager data for individual channels [ch] (b), radiosonde observations indicating temperature (c) and relative humidity (d) for individual pressure heights [hPa]. The validation periods are from June 27 to July 8 2018 (red dots) and from January 16 to 27, 2018 (green dots).
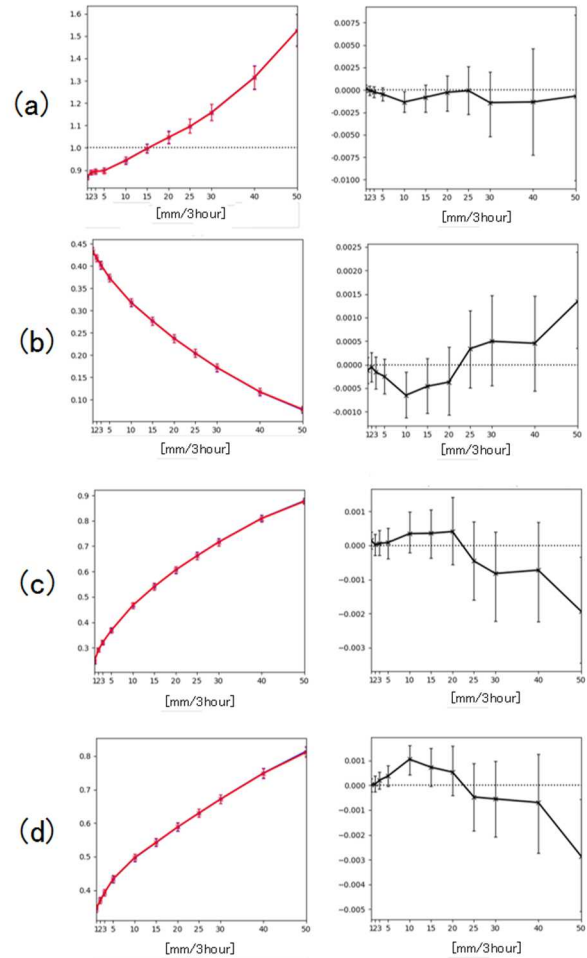

Figure 2. Bias scores (a), threat scores (b), false alarm ratios (c) and undetected error rates (d) for three-hour cumulative precipitation forecasts with different thresholds (x axis). Red lines are scores for TEST and black lines are for CNTL (left panels). The panels on the right show differences between TEST and CNTL scores (TEST - CNTL). The validation periods are from June 27 to July 8, 2018.

Multigrid Beta Filter Scheme for Modeling Background Error Covariance in 3DRTMA

R. James Purser, Miodrag Rancic and Manuel Pondeca

IMSG at NOAA/NCEP/EMC

College Park, Maryland, USA

Email: miodrag.rancic@noaa.gov

1. Parallel multigrid beta filter

This report describes a novel technique for modeling of the background error covariance that is under development at EMC.

Until now, the data assimilation system at EMC was using recursive filters (RF) (e.g., Wu *et al*., 2002; de Pondeca *et al.*, 2011) as a quasi-Gaussian approximation of covariances. Despite some nice properties, RF are essentially sequential operators, exceedingly difficult to successfully parallelize, preventing them to scale well with an increasing number of processing elements.

The Three-Dimensional (3D) Real-Time Mesoscale Analysis (RTMA) system is a project with the goal to provide analyses at remarkably high horizontal resolutions (~2.5 km) at frequent time intervals (~15 min). For the success of such computationally demanding enterprise, the key prerequisite is a vastly improved efficiency, in which scenario the RF represents one of the main hurdles.

Our solution to this problem is the development of a new filter based on the beta distribution, incorporated within a parallel multigrid structure (MGBF).

The new technique has ability to better describe covariances across various scales, to include cross-correlations and to provide negative side lobes, which realistic covariances may possess. Most importantly, the beta filter has a finite support and is therefore more readily parallelizable, which should yield much better scaling.
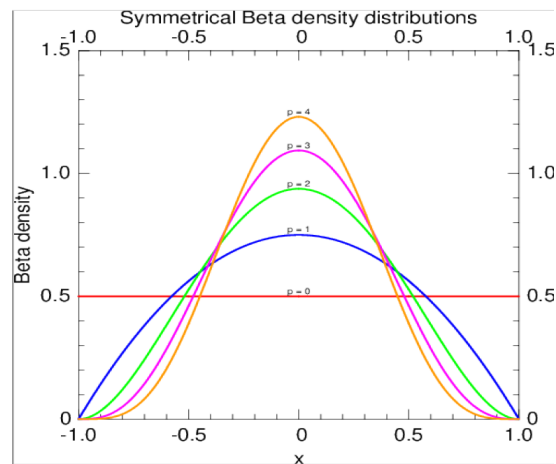

Fig. 1: Examples of beta filter (From Purser *et al.* 2021).

The multigrid method is usually used for solving of elliptic problems based on the idea that the solutions derived at lower and higher resolutions can be combined in an overall more efficient numerical scheme.

In our case, we apply the multigrid by simultaneously calculating quasi-Gaussian approximations by the beta filter over various spatial scales (grid generations), and then combining them together by giving appropriate scale weights to contributions from various grid generations.

An efficient version of the code is developed that allows integration over a wide range of domain decompositions among processing elements.
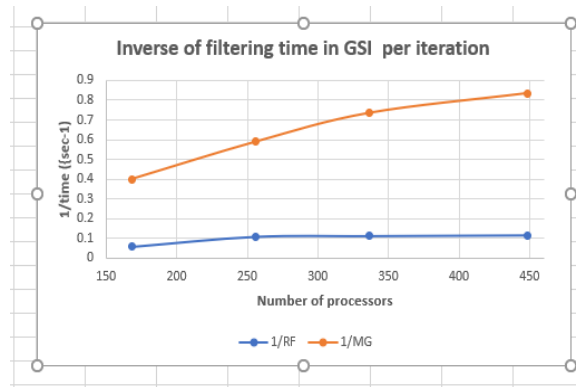
Fig. 2: Scaling derived in single observation tests with
RF and MGBF in GSI (Rancic *et al.* 2021).

2. Preliminary test results

Preliminary scaling derived in a single observation test within GSI (Grid-point Statistical Interpolation), the present data assimilation system at EMC, are shown in Fig. 2. With the increase of number processors, MGBF keeps improving generally better performance, which RF cannot match any longer.

3. Concluding remarks

The MGBF is a multifaceted project, with many attractive features, such as:

- Generation of negative side lobes of covariances through application of a Helmholtz operator
- Introduction of a multivariate option
- Formulation of an efficient, line version of beta filter
- The code is generalized so that it can perform across a wide range of processor decompositions

A preliminary application in GSI confirms a potentially exceptionally large speed up in comparison to RF.

Plans for further development include:

- An updated definition of aspect tensor
- Formulation of a novel, computationally efficient technique for normalization of covariances
- Inclusion of a line version of beta filter
- Further exploration of the multivariate option
- Implementation of MGBF into the JEDI, a new data assimilation system, and formulation of a generalized version for the cubed sphere.

In addition, the line beta filter has a four-dimensional version which opens interesting possibilities for applications in nowcasting. A strategy for application of machine learning to judicially define MGBF scale weights is also formulated, in an effort to further polish estimate of the background error covariance.

References

De Pondeca, M. S., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's Na-tional Centers for Environmental Prediction: Current status and development. Wea. Forecasting, 26, 593–612.

Purser, R. J., M. Rancic, and M. S. De Pondeca, 2021: The Multigrid Beta Function Approach for Modeling of Background Error Covariance in the Real Time Mesoscale Analysis (RTMA). Mon. Wea. Rev. (under review).

Rancic, M., R. J. Purser and M. Pondeca, 2021: Multigrid beta filter scheme for modeling background error covariance in 3DRTMA: Generalization, Helmholtz operator, and line filters. 25IOAS. 101thAnnual Meeting of American Meteorological Society, 9-15 January 2021.

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Wea. Rev., 130, 2905–2914.

Plans to Estimate Adaptive Covariance Parameters Using a Neural Network

Miodrag Rancic, R. James Purser, and Manuel Pondeca

IMSG at NOAA/NCEP/EMC

College Park, MD, USA

Email: jim.purser@noaa.gov

There are plans at EMC to replace the covariances used for background error within the Real-Time Mesoscale Analysis (RTMA, de Pondeca *et al*., 2011) with a dynamically adaptive formulation based on compact-support beta distribution filters ("beta filters") embedded in a computationally efficient distributed multigrid algorithm (Purser *et al*. 2021). In the present operational two-dimensional formulation, the covariance is static, although possessing anisotropic and spatially inhomogeneous features tied to orography. But it uses recursive filters (Wu *et al*., 2002), which have proven to be difficult to parallelize efficiently. The inherently more efficient and more versatile multigrid beta filter (MGBF) covariances that we are extending into three dimensions offer us the opportunity to make these covariances adapt not only to fixed terrain, as is presently done, but also to the ambient flow conditions, and to the evident variability and uncertainty that we can deduce by exploiting the availability of ensembles of short forecasts. In order to make the connection between suitable diagnostics of the background field and ensemble on the one hand, and the covariance amplitude and anisotropy parameters on the other hand, we plan to employ the machine learning techniques of artificial neural networks (NN, Krasnopolsky 2013).

One problem we encounter is that the spatial shape of the covariance response is characterized not by independent scalar parameters, but by a tensor (an "aspect tensor") whose symmetry and positivity must be preserved. The ensemble-averaged outer-product of gradients of the fields formed by the departures of the ensemble members about their collective mean form a tensor which, when suitably mixed with a regularizing horizontally isotropic tensor and the result inverted, provides a tensor of the desired character, stretching the covariance response in the direction indicated by the ensemble. It is then the mixing weights (positive scalars) involved in this process that we can ask the NN to supply, as "outputs" in response to local atmospheric diagnostics, such as windiness, static stability, and ensemble variance, as well as fixed terrain diagnostics, that can be gathered as "inputs" to the NN. The covariance amplitude, i.e., the background error variance, for each analysis variable, is another set of parameters that we can ask a trained NN to provide in response to diagnostics from the background field and ensemble.

Another problem we must address is the "training" of the NN, which entails establishing a formal criterion that corresponds to an objective measure of the quality of the covariance estimate that the NN implies, over numerous archived cases. This needs to be set up in such a way that we can iteratively search for the combination of the very numerous internal weights of the NN that appear to optimize the choices of covariance parameters conditional on the characteristics of ambient flow and of the diagnosed variability within the ensemble. Fortunately, the products of nearby pairs of the observation innovations constitute unbiased (though obviously very noisy and sporadic) "measurements" of the covariance of these innovations. We can therefore use a weighted sum of the squares of the differences between a thinned subset of the products of innovation pairs, and the corresponding modeled covariance estimates, since these latter estimates only differ from the background error covariances by the addition of the small, and reasonably well known, diagonal covariance of the observation errors. It is the fact that the modeled covariances in the multigrid scheme are made up of additive quasi-Gaussian contributions, the beta filters, that will allow us to approximate these contributions by the true analytic Gaussians during the training phase of establishing the NN weights. It is important that the covariances be analytically differentiable with respect to their parameters, since the training process requires that the derivatives of the quality criterion be expressible explicitly with respect to the internal weights of the NN (through the "back-propagation" application of the chain rule) in order that these weights can be efficiently optimized. Once the NN has been successfully trained on a varied and representative archive of past cases, we believe we shall be able to implement a dynamically adaptive analysis for the RTMA using the multigrid beta filter approach, guided by the diagnostics we have available from the ensemble of short forecasts, from the background fields itself, and from the local topography.

References

De Pondeca, M. S., and Coauthors, 2011: The Real-Time Mesoscale Analysis at NOAA's National Centers for Environmental Prediction: Current status and development. Wea. Forecasting, 26, 593–612.

Krasnopolsky, V. M., 2013: The Application of Neural Networks in the Earth Sciences. Springer, 204 pp.

Purser, R. J., M. Rancic, and M. S. De Pondeca, 2021: The Multigrid Beta Function Approach for Modeling of Background Error Covariance in the Real Time Mesoscale Analysis (RTMA). Mon. Wea. Rev. (under review).

Wu, W.-S., R. J. Purser, and D. F. Parrish, 2002: Three-dimensional variational analysis with spatially inhomogeneous covariances. Mon. Wea. Rev., 130, 2905–2914.

Assimilation of satellite derived surface temperature over land in AROME-France model

Zied Sassi, Nadia Fourrié, Camille Birman, Vincent Guidard

CNRM, Météo-France and CNRS, 42 av Coriolis, 31057 Toulouse, France

The surface temperature is a key parameter in Earth radiative balance simulation and numerical weather prediction. However, its high spatial and temporal variability makes its observation and simulation complex over land. Nowadays, the meteorological satellites observe the Earth with high coverage and several algorithms enable the retrieval of the surface temperature from satellite observations (Li et al, 2013). Although it is used for the assimilation of satellite radiances in AROME model (Seity et al, 2011), the satellite derived land surface temperature (LST) itself is not currently assimilated. The aim of this work is to evaluate the benefit of assimilating the SEVIRI (Aminou, 2002) LST in AROME and to study its impact on the assimilation of different observations and also on the forecasts.

Since this project represents the first work of LST assimilation in AROME, the first step consisted in implementing the assimilation of SEVIRI LST in the surface analysis system as described on figure 1. AROME model is coupled to SURFEX platform (Masson et al, 2013) for surface modelisation. The operational version of the surface analysis system uses a 2D Optimal Interpolation to assimilate as a first step the 2 meter temperature and relative humidity observations, then a 1D Optimal Interpolation based on 2 meter temperature and relative humidity increments is used to analyze the soil temperature and water content in the first and second soil layers. As described in figure 1, the assimilation of SEVIRI LST has been implemented and is used in addition to the 2 meter temperature to analyze the soil temperature.
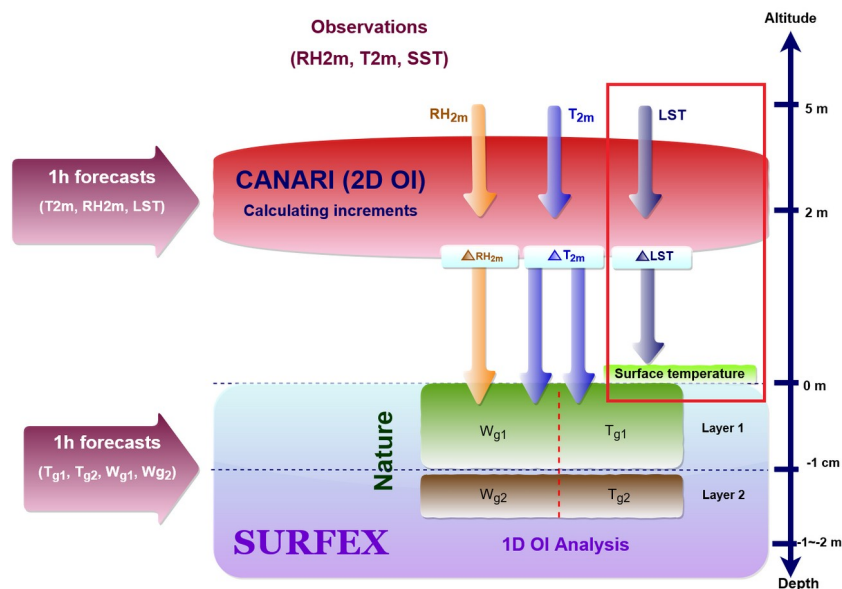

Figure 1: Implementation of the LST assimilation in the surface analysis of AROME model

Based on the results of previous work showing a better synergy between infrared sensors derived LST during nighttime (Sassi et al, 2020) and on the diagnostics of LST observations and model errors showing lower uncertainties by nighttime, the assimilation of SEVIRI LST has been first implemented at nighttime analysis times (00 and 03 h). The results obtained in this experiment were compared to the ones from a reference experiment without the assimilation of SEVIRI LST.

The second step of this work consisted in evaluating the impact of SEVIRI LST assimilation over a two-month period on the assimilation of surface and atmospheric observations. A slight decrease of the difference between the 2 meter temperature observation and the model first guess has been observed during the first assimilation analysis times of the day (00-06 h) as shown in figure 2. A consistent decrease in difference between the first guess and the 2 meter relative humidity has been also observed during nighttime assimilation.
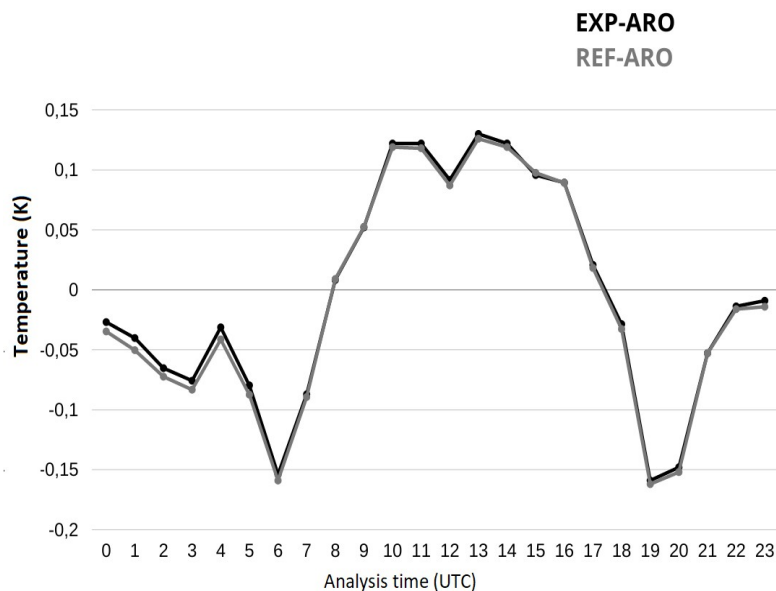
Figure 2: Mean differences between 2 m temperature (K) observations and first guess for July and August 2019. Black line represents the first guess departure in the experiment assimilating LST and grey line the results for the reference.

Moreover, the evaluation of satellite channels in the atmospheric assimilation has shown a decrease of difference between first guess and observation for microwave sensors and an improvement of the background for most analysis times.

Concerning the impact on forecasts, when compared to surface stations, a slight but significant and consistent improvement has been observed on the temperature and relative humidity at 2 meters during nighttime forecast ranges. Compared to radiosondes observations, a decrease of the RMSE of the temperature and relative humidity parameters has been observed, mainly between 1000 and 700~hPa levels, for the first forecast ranges in most cases, up to 24h.

To further evaluate the benefit of the satellite derived LST assimilation, initially implemented for nighttime assimilation, it would be interesting to evaluate the impact of the LST assimilation at all analysis times, including the daytime assimilation times. Further steps will also include the extension of the methodology to other instruments and the use of LST for soil moisture analysis.

References:

Aminou, D. [2002], 'MSG's SEVIRI Instrument', ESA bulletin (111), 15–17.

Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I. F. and Sobrino, J. A. [2013], 'Satellite-derived land surface temperature : Current status and perspectives', Remote Sensing of Environment 131, 14 – 37.

Masson, V., Le Moigne, P., Martin, E., Faroux, S., Alias, A., Alkama, R., Belamari, S., Barbu, A., Boone, A., Bouyssel, F., Brousseau, P., Brun, E., Calvet, J.-C., Carrer, D., Decharme, B., Delire, C., Donier, S., Essaouini, K., Gibelin, A.-L., Giordani, H., Habets, F., Jidane, M., Kerdraon, G., Kourzeneva, E., Lafaysse, M., Lafont, S., Lebeaupin Brossier, C., Lemonsu, A., Mahfouf, J.-F., Marguinaud, P., Mokhtari, M., Morin, S., Pigeon, G., Salgado, R., Seity, Y., Taillefer, F., Tanguy, G., Tulet, P., Vincendon, B., Vionnet, V. and Voldoire, A. [2013], 'The SURFEXv7.2 land and ocean surface platform for coupled or offline simulation of earth surface variables and fluxes', Geoscientific Model Development 6(4), 929–960.

Sassi, M. Z., Fourrié, N., Guidard, V. and Birman, C. [2019], 'Use of infrared satellite observations for the surface temperature retrieval over land in a NWP context', Remote Sensing 11(20), 2371.

Seity, Y., Brousseau, P.and Malardel, S., Hello, G.and Bénard, P., Bouttier, F., Lac, C. and Masson, V. [2011], 'The AROME-France Convective-Scale Operational Model', Monthly Weather Review 139, 976–991.

Effect of supersaturation constraint in a variational data assimilation system

SAWADA Ken[1] and HONDA Yuki[2]

[1] *Meteorological Research Institute, Japan Meteorological Agency, Tsukuba, Japan*
[2] *Japan Meteorological Agency, Tokyo, Japan*
(e-mail: ksawada@mri-jma.go.jp)

1 Introduction

The reproducibility of precipitation in the early stages of forecasts (spin-down or spin-up problem) has been a significant issue in numerical weather prediction. This issue is thought to be caused by moisture imbalance in the initial data given by data assimilation. In the case of the Japan Meteorological Agency (JMA) mesoscale data assimilation system JNoVA, we found that the imbalance stems from the existence of unrealistic supersaturated states in the minimal solution of the cost function in JNoVA. We implemented a penalty function method for the mixing ratio within JNoVA to suppress unrealistic supersaturated states and investigated the effects on the reproducibility of precipitation.

2 Method

The fundamental solution for handling the moisture imbalance problem are to construct proper control variables and set proper background error covariance (see, [1]). However, they are quite difficult, thus we apply an exterior penalty function method to avoid generating unrealistic moisture states and to obtain appropriate moisture balance, without entering into the problem of reconfiguring the control variables and error covariances.

An exterior penalty function method is one of the numerous algorithms to solve a constrained optimization problem [3, 2]:

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \quad \text{subject to } g(\mathbf{x}) \leq 0, \quad (1)$$

where $f(\mathbf{x})$ is the objective function (or cost function), $g(\mathbf{x})$ is the (nonlinear) constraint function, and X is the control space. In this method, the original constrained problem (1) is converted to an unconstrained problem

$$\min_{\mathbf{x} \in X} f_2(\mathbf{x}), \quad (2)$$

by introducing the auxiliary function defined by

$$f_2(\mathbf{x}) = f(\mathbf{x}) + \lambda \max\{0, g(\mathbf{x})\}^\alpha, \quad (3)$$

where $\lambda > 0$ and $\alpha \geq 1$ are the penalty parameters. The second term of equation (3) is called the "penalty term" or "penalty function". If there is more than one constraint, one additional penalty term is added for each constraint.

The implementation of the exterior penalty function method in the variational assimilation system is very simple, it is done by just adding penalty term

$$J_{\mathrm{qv}}(\mathbf{x}) = \lambda \sum_i \left(\max\{0, g_{1i}(\mathbf{x}), g_{2i}(\mathbf{x})\} \right)^\alpha \quad (4)$$

to the original cost function, where

$$\begin{cases} g_{1i}(\mathbf{x}) &= \mathrm{qv}_i(\mathbf{x}) - \mathrm{qvs}_i(\mathbf{x}) \\ g_{2i}(\mathbf{x}) &= -\mathrm{qv}_i(\mathbf{x}). \end{cases} \quad (5)$$

Here, qv_i and qvs_i are the water vapor mixing ratio and the saturation mixing ratio, respectively, at the i-th grid point. Here, $g_{1i} \leq 0$ is the constraint for supersaturation at the i-th grid point, and $g_{2i} \leq 0$ is that for the negative mixing ratio. Since the mixing ratio has large values in the lower troposphere, this construction of the penalty function (5) is intended for the effective modification of the atmospheric fields in the lower troposphere, which is closely related to the initiation and development of deep convection and the generation of precipitation.

3 Results and Summary

To investigate the impact of the penalty function method on the analysis and forecast, we conducted twin data assimilation cycle experiments from June 28th to July 8th, 2018. In the following, the experiment that employs the original JNoVA system is called "Ctrl," and those utilizing the new JNoVA system equipped with the penalty term J_{qv} are called "Tests." In the Tests, we set $\alpha = 1$ and $\lambda = 100, 200, 500,$ and 1000, which are labeled "L100," "L200," and so on.

Figure 1 shows the impact of the penalty function on the modification for the violation defined by $\max\{0, g_{1i}(\mathbf{x}), g_{2i}(\mathbf{x})\}$. The violations are substantially reduced by the penalty function method as the value λ becomes large.

Figure 2 shows the three-hour accumulated precipitation in the forecasts from the initial data at 09 UTC on June 28th, which is the result of the first cycle. Strong rainfall along the baiu front in the sea northwest of Kyushu, in the radar/raingauge analyzed precipitation data (RA, treated as the obser-

vations), is not adequately reproduced in the forecast of Ctrl. In the forecast of the Tests, the reproducibility (including the location, distribution, and amount) of precipitation is improved. The distributions of precipitation in the Tests (L200, L500, and L1000) are similar to each other, and the maximum amounts of the precipitation in L500 and L1000 are almost comparable to that of the RA.

We also performed verifications of the 12-hour precipitation forecasts of Ctrl and the Tests every 12 hours, during the cycle period (the initial times are 00UTC and 12UTC). The fractions skill scores (FSSs) with a 10 km verification grid and thresholds of 1.0 mm/h and 10.0 mm/h are shown in Figure 3. We can see clear improvements in the FSSs of the Tests at both thresholds in the early stages of the forecast. For other cases with different verification grid sizes and thresholds, we confirmed that the Tests are superior to Ctrl in general (not shown). But it is difficult to determine what value of λ gives the best improvement in the FSS, since the scores vary with the threshold and the atmospheric state even though L1000 seems to be better among the Tests from Fig. 3. These indicate that the improvement is robust for the values of λ. One possibility is that the value $\lambda = 100$ is sufficient in the case of cycle assimilation because the differences of the atmospheric fields among the Tests are small compared to the differences between Ctrl and the Tests (not shown).

From these results, we conclude that the new moisture balance introduced by the penalty function method has a positive impact on the reproducibility of precipitation in the early stages of forecasts.

References

[1] R. N. Bannister, H. G. Chipilski, and O. Martinez - Alvarado. Techniques and challenges in the assimilation of atmospheric water observations for numerical weather prediction towards convective scales. *Quart. J. Roy. Meteor. Soc.*, 146:1–48, 2019.

[2] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear programming: Theory and Algorithms*. John Wiley and Sons, 3 edition, 2006.

[3] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Method*. Academic Press, 1982.
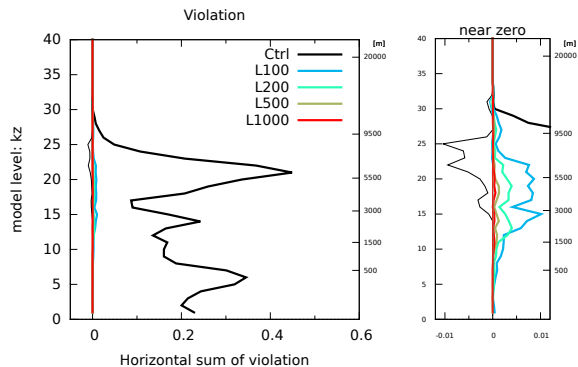
Figure 1: Vertical profile of horizontal summation of the violations with an enlarged view shown in the panel to the right.
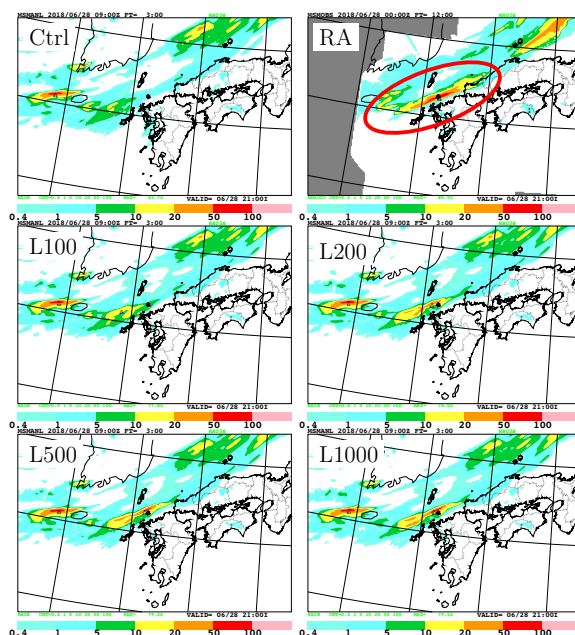

Figure 2: Three-hour accumulated precipitation ([mm/3h]; shaded as in the color bar) at 12UTC on June 28th, 2018. The different panels show the radar/raingauge analyzed precipitation (RA) and the forecasts of Ctrl and the Tests (L100, L200, L500, and L1000).
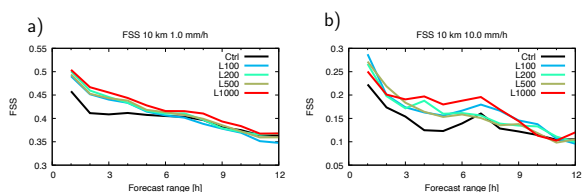

Figure 3: Fractions Skill Scores (FSSs) of the forecasts up to 12 hours. The verification grid is 10 km, and the precipitation thresholds are a) 1.0 [mm/h] and b) 10.0 [mm/h].

A New Atmospheric River Quantification Metric

Keqin Wu[1], Xingren Wu[2], and Vijay Tallapragada[3]
[1]NGI at EMC/NCEP/NOAA, College Park, MD 20740, [2]IMSG at EMC/NCEP/NOAA, College Park, MD 20740,
[3]EMC/NCEP/NOAA, College Park, MD 20740

Email: Keqin.Wu@noaa.gov

1. Introduction

Atmospheric rivers (ARs) are long, narrow bands of atmospheric moisture that are responsible for most of the horizontal transport of water vapor outside the tropics. They play a critical role in creating extreme precipitation, flooding, drought, etc. Despite the fact that ARs are attracting increasing interest from various communities, no consensus has yet been reached on a metric to track and measure AR related quantities for an AR forecast. A preferred metric should help scientists, forecasters, and the public easily interpret and understand the characteristics of ARs.

2. AR Quantification Metric

Our new AR quantification metric is inspired by the widely used automated tropical cyclone forecasting (ATCF) hurricane system track. We define AR tracks as the path of the selected AR centers in consecutive time steps. A few other quantities would also be calculated and recorded along an AR track. The AR scale proposed by CW3E [1] categorizes the impacts of ARs into 5 scales, based on duration and magnitude of the Integrated Water Vapor Transport (IVT). It does not fully reflect the locations and time-varying characteristics of an AR. Our new metric covers both the intensity and tracks of ARs to provide a more comprehensive perspective on location related AR characteristics, impacts, and landfalls. The time series of a few AR quantities recorded along the path would also provide a simplified data set for AR studies, to reduce requirements for data archive and transfer.
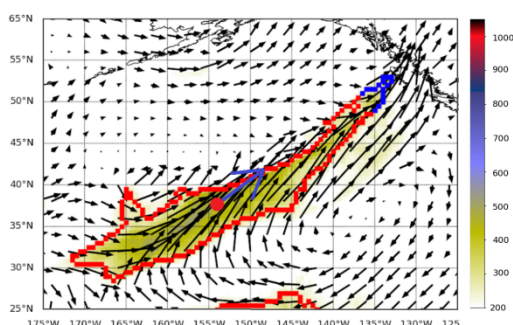

Figure 1. AR quantification by AR Center (red dot), AR Intensity (maximum IVT), AR size (area measurement of the AR region), AR direction (blue arrow), and AR front (blue curve).

First, ARs are extracted from regions with *IVT > 200 kg m^{-1} s^{-1}*, region length > *2000 km*, length/width ratio > *2*, and other conditions based on Guan and Waliser's AR extraction method [2]. From an extracted AR (Figure 1), we calculate *AR Center* as the mean center of the AR region weighted by IVT values, *AR intensity* as the maximum IVT within the AR region, *AR Size* as the region area, *AR direction* as the direction of the mean IVT vector within the AR region, and *AR front* (front part of AR boundary along the AR direction). The AR size and maximum IVT are good indicators of an AR's impact. AR centers in consecutive time steps form an *AR track*, which, along with AR fronts, project an AR's paths and landfalls.

3. Applications

The new metric is applied to study AR forecast uncertainty using the NCEP global ensemble forecast system version 12 (GEFSv12) reanalysis and reforecast data.

Figure 2 shows tracks from different ensemble members and also spaghetti plots showing the contours of AR boundaries. The purple shades show the degree of overlap among the AR regions predicted by different ensemble members. The overlaid AR tracks indicate the uncertainty of AR paths among ensemble members.

Figure 3 shows composite tracks from a few consecutive GEFSv12 cycles. The tracks are generated from the ensemble mean of each cycle. The background color map shows the swath of maximum IVT. Plotting these AR tracks in a single map provides a quick overview of the variance among the same ARs predicted at

different initial times. For example, even when the ensemble means are used, different AR predictions can still occur with different cycles' forecast data.
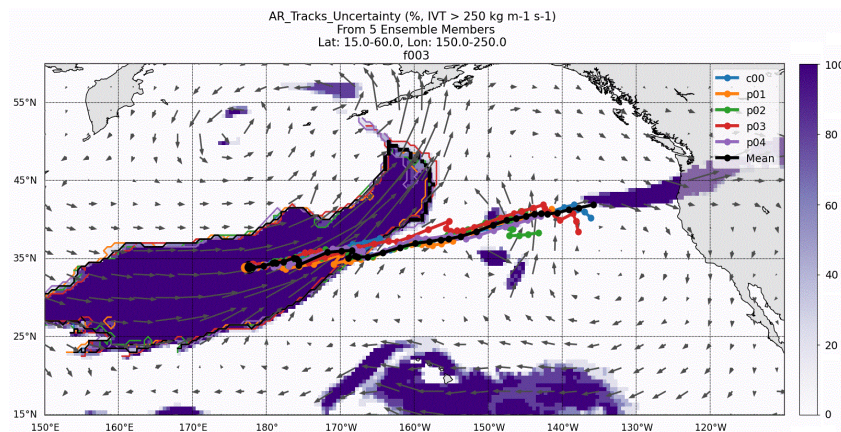

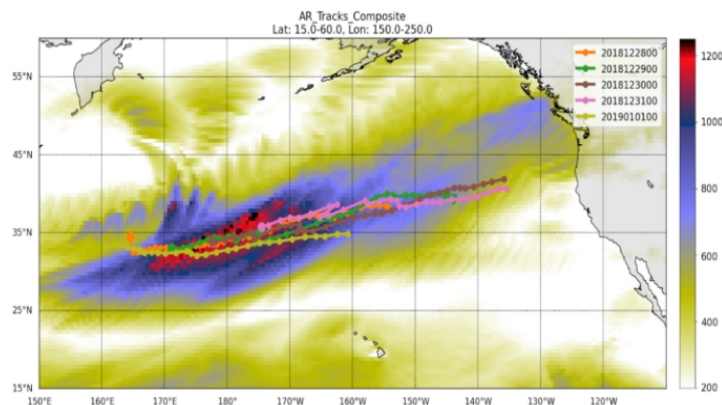Figure 2. AR tracks of GEFSv12 ensemble in East Pacific Ocean on 12/30/2018 (5 members).

Figure 3. Composite AR tracks of GEFSv12 in East Pacific Ocean (cycle 2018122800 - 2019010100).

4. Conclusion and Future Work

We propose a new AR quantification metric based on AR path extraction. The example applications demonstrate that the new metric provides an effective tool to study ARs. Future work may include refining the calculation of AR centers, so the AR tracks may reflect times and locations of the ARs' landfalls and the maximum impacts more precisely, collectively selecting the set of AR related quantities along the tracks with the community, and including more applications to verify the usefulness of those AR quantities.

References

[1] F. M. Ralph, J. J. Rutz, J. M. Cordeira, M. Dettinger, M. Anderson, D. Reynolds, L. J. Schick, and C. Smallcomb, A Scale to Characterize the Strength and Impacts of Atmospheric Rivers, Bulletin of the American Meteorological Society. (2019) 100 (2): 269–289, doi:https://doi.org/10.1175/BAMS-D-18-0023.1

[2] B. Guan and D. E. Waliser (2015), Detection of atmospheric rivers: Evaluation and application of an algorithm for global studies, Journal of Geophysical Research: Atmospheres, 120, 12514–12535, doi:10.1002/2015JD024257

Impact of Atmospheric River Reconnaissance Dropsonde Data on NCEP GFS Forecast: A Case Study

Xingren Wu[1], Vijay Tallapragada[2], Stephen Lord[3], Keqin Wu[4], and Marty Ralph[5]

[1]IMSG at EMC/NCEP/NOAA, College Park, MD 20740, [2]EMC/NCEP/NOAA, College Park, MD 20740, [3]UCAR at EMC/NCEP/NOAA, College Park, MD 20740, [4]NGI at EMC/NCEP/NOAA, College Park, MD 20740, [5]CW3E, Scripps Institution of Oceanography, UC San Diego, CA 92093

Email: Xingren.Wu@noaa.gov

1. Introduction

Atmospheric rivers (ARs) are long narrow corridors of water vapor transport that serve as the primary mechanism to advect moisture into mid-latitude continental regions, including the U.S. West Coast. They are responsible for most of the horizontal water vapor flux outside of the tropics and are an important source of precipitation. Although the advances in satellite data assimilation have greatly improved global model forecast skill, including the NCEP global forecast system (GFS), forecasting AR features and the corresponding precipitation remains a challenge due in part to their formation and propagation over the ocean, where in-situ and ground-based observations are extremely limited (Zheng *et al*. 2021). The AR Reconnaissance (AR Recon) Campaigns (Ralph *et al*. 2020) that took place during the winters of 2016 and 2018-2021 provide additional data by supplementing conventional data assimilation with dropsonde observations of the full atmospheric profile of water vapor, temperature, and winds within the ARs.

2. Experiments

In this study the NCEP GFS version 15 (GFSv15) was used to examine the impact of AR supplemental dropsonde observation data on the GFS forecast. GFSv15 was implemented into operations in June 2019 at NCEP. It was developed with the finite volume cubed-sphere dynamical core and microphysics from GFDL, and 4D-Hybrid En-Var data assimilation (DA). The dropsonde data used were from the AR Recon 2020 campaigns, including 17 intensive observation periods (IOPs) from January 24 to March 11. Global control (CTRL) and denial (DENY) experiments were conducted by using or denying the dropsonde data in the GFSv15 in the period from January 24 to March 18 for both DA and model forecasts; the additional 7 days of runs are needed to verify the forecasts from the last IOP of March 11.

3. Results

The standard NCEP *vsdb* verification system [1] was used to evaluate the CTRL and DENY experiments. Overall the global verification metrics were very similar between CTRL and DENY, with slightly better forecast skill noted over the Pacific North American (PNA, 180-320E, 20-75N) region when the supplemental dropsonde data were used (Figure 1). For cases where the prediction skill is relatively low (i.e., the prediction is challenging), the data collected from the dropsondes helped to improve AR related precipitation forecasts and increase the 5-day anomaly correlation, including geopotential height, temperature, and wind (Figure 1). Precipitation observations are critical for verifying AR impact. The impact of the dropsonde data over the CONUS was small and insignificant. However, precipitation prediction over the U.S. West Coast improved significantly in the CTRL when dropsonde data are used (Figure. 2). It is also associated with improvement in the water vapor transport forecasts in the CTRL (not shown).
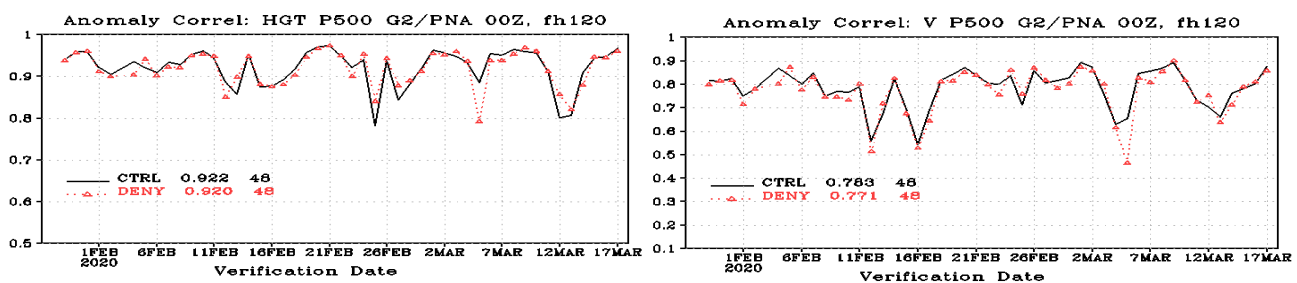

Figure 1. Anomaly correlation of day-5 forecast for geopotential height (HGT) and meridional wind (V) at 500 hPa over PNA.
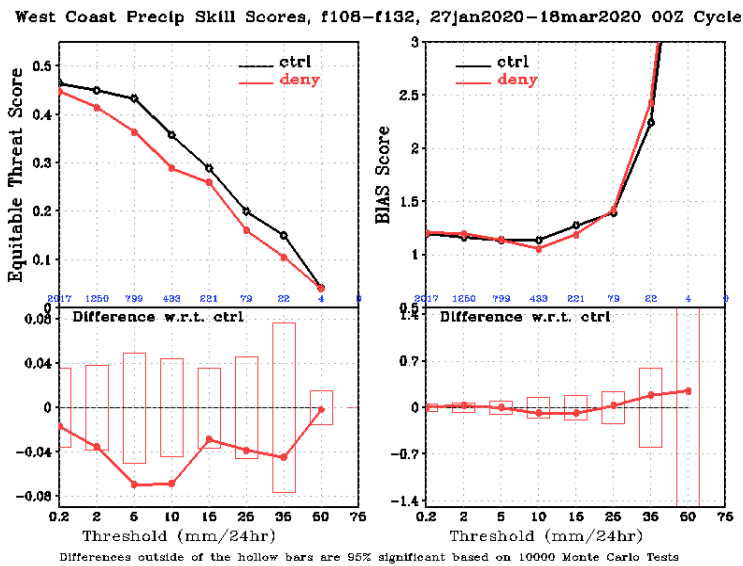
West Coast Precip Skill Scores, f108–f132, 27jan2020–18mar2020 00Z Cycle

Figure 2. 24-h precipitation Equitable Threat Scores of day-5 forecast for CTRL and DENY over the U.S West Coast.

4. Summary

This study indicates that there was a small overall positive impact on the GFS forecast skill for the PNA region when dropsonde data were available and used. Data impact was greater when dropsonde observations from consecutive IOPs were available and used. The AR supplemental observations helped fill the data gaps needed for DA to provide better model initial conditions (Zheng *et al*. 2021). There is a systematic improvement in the precipitation prediction over the U.S. West Coast when the dropsonde data are used. This is associated with improvement in the water vapor transport forecast.

References

Ralph F.M., and Coauthors, 2020: West Coast forecast challenges and development of Atmospheric River Reconnaissance. *Bull. Amer. Meteor. Soc.*, **101**, E1357–E1377, https://doi.org/10.1175/BAMS-D-19-0183.1.

Zheng M., L.D. Monache, X. Wu, F.M. Ralph, B. Cornuelle, V. Tallapragada, J.S. Haase, A.M. Wilson, M. Mazloff, A.C. Subramanian, F. Cannon, 2021: Data Gaps within Atmospheric Rivers over the Northern Pacific. *Bulletin of the American Meteorological Society,* **102**, E492–E524, https://doi.org/10.1175/BAMS-D-19-0287.1. 2021.

[1] https://github.com/yangfanglin/gfs_verif/tree/vsdb_grib2

Preliminary Comparison and Evaluation of Soil Moisture Produced in the NCEP Global Land Data Assimilation System (GLDAS)

Youlong Xia[1], Helin Wei[1], Mollie D. Gaines[2], Jesse Meng[1], George Gayno[1], Hang Lei[1], Rongqian Yang[1], Yihua Wu[1], Fanglin Yang[3], Michael Barlage[3], Daryl T. Kleist[3], and Vijay Tallapragada[3]

[1]IMSG at EMC/NCEP/NOAA, College Park, Maryland, USA; [2]Center for Geospatial Analytics, North Carolina State University, Raleigh, North Carolina, USA; [3]EMC/NCEP/NOAA, College Park, Maryland, USA
Youlong.Xia@noaa.gov

Background of the NCEP GLDAS

The offline land modeling system GLDAS was operationally implemented at NCEP in 2011 (*Meng et al., 2012*). Its main purpose is to provide land surface states such as soil moisture and temperature to the Climate Forecast System (CFS), including the CFS Reanalysis and Reforecast project and the CFS version 2, in support of seasonal analysis and forecasting at NCEP. As in traditional LDAS systems, GLDAS uses the CPC (Climate Prediction Center) Merged Analysis of Precipitation and NCEP's Global Data Assimilation System (GDAS) surface meteorological forcing to drive the Noah land model to produce soil moisture and soil temperature. In 2020, GLDAS was updated and extended to support the development of the Global Forecast System version 16 (GFSv16). Unlike the GLDAS used in the CFS, the updated GLDAS is forced by the CPC's 1/8th degree global gauge-based daily precipitation and GDAS surface meteorological forcing, and uses an updated MODIS (Moderate Resolution Imaging Spectroradiometer) IGBP (International Geosphere-Biosphere Programme) vegetation type and STATSGO (State Soil Geographic) soil type. The system was updated to a newer version of the NOAH Land Surface Model (LSM). It is run once a day to provide soil moisture and soil temperature initial conditions for GFS forecasts. Figure 1 shows how GLDAS was used in GDAS and GFSv16. It should be noted that GLDAS is not a data assimilation system, as neither in-situ observations nor satellite retrievals are assimilated into the NOAH LSM.
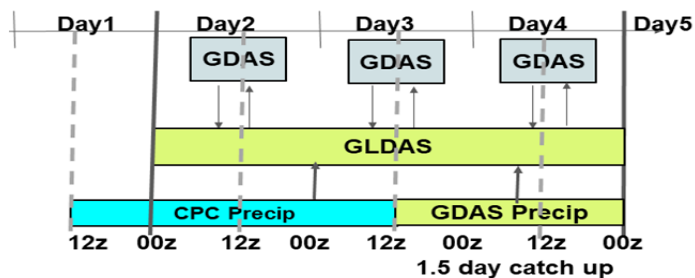

Figure 1. A diagram of the daily GLDAS run when the CPC precipitation and GDAS surface meteorological forcing are used. CPC daily precipitation observations are accumulated for a 12Z to 12Z cycle. GLDAS has a 1.5 day catch up as the CPC precipitation lags 1.5 days behind real time.

Reference Datasets and Evaluation Method

GLDAS soil moisture was evaluated by using multiple reference datasets including the operational GFSv15 model product, North American LDAS reanalysis product (NLDAS, *Xia et al., 2012*), Soil Moisture Operational Products System (SMOPS) satellite retrievals, and in-situ observations from the International Soil Moisture Network (ISMN). The NLDAS daily soil moisture simulated in the three land models has been comprehensively evaluated against in situ observations (*Xia et al., 2015*). In this study we compared and evaluated soil moisture for summertime (1 July - 31 August 2019) and wintertime (1 December 2019 – 31 January 2020) against our reference soil moisture products. We used time series comparisons of regional averages and station measurements, where stations were selected using the nearest neighbor method. A few examples are shown in the Results and Summary section below.

Results and Summary

Regionally averaged top 10 cm soil moisture was compared with the operational GFSv15 and NLDAS (Fig.2a), and SMOPS (Fig.2b). GLDAS is closer to NLDAS Noah and Mosaic than the GFSv15. GLDAS uses the same CPC gauge-based precipitation, while NLDAS uses the Regional Climate Data Assimilation System (RCDAS) surface meteorological forcing. The NLDAS evaluation results showed that VIC (Variable Infiltration Capacity) overestimates and Mosaic underestimates the observed soil moisture, while Noah is closer to the observations (*Xia et al., 2015*). Overall, GLDAS soil moisture in dry regions may tend to be even drier. The same result can be found when SMOS soil moisture was compared (Fig.2b). GLDAS soil

moisture is outside of one-sigma standard deviation and tends to overestimate dryness when compared with GFSv15. In spite of these magnitude differences, GLDAS reasonably captures the daily variability of NLDAS and SMOPS soil moisture. GLDAS soil moisture was also compared with in-situ observations in Austin, TX (Fig.3a) and at Cochora Ranch, CA (Fig.3b). The results show that in the summer GLDAS soil moisture estimates are drier than the GFSv15 and in-situ observations, and in the winter GLDAS and GFSv15 estimates are closer to each other than to the observed soil moisture. There is a very large difference between the models and observations. The major reason for the dry soil moisture estimates may be due to lower precipitation amounts generated in GFSv16 and less observed precipitation when compared with precipitation generated in the operational GFS. In the summer, low precipitation results in less infiltration into soil, so the soil will become drier and drier until the wilting point is reached. However, the actual cause remains unclear and needs further investigation in the future. The large difference between the GLDAS and in-situ observations may come from (1) spatial scale mismatch, (2) soil type and related soil and hydrologic parameter differences, (3) vegetation type differences leading to different ET, (4) missing/misrepresenting physical processes, and (5) surface meteorological forcing errors. Nevertheless, comparison and evaluation of multiple references have shown that GLDAS reasonably captures the observed daily variability. Therefore, GLDAS was included in GFSv16 for an operational implementation at NCEP in March 2021.
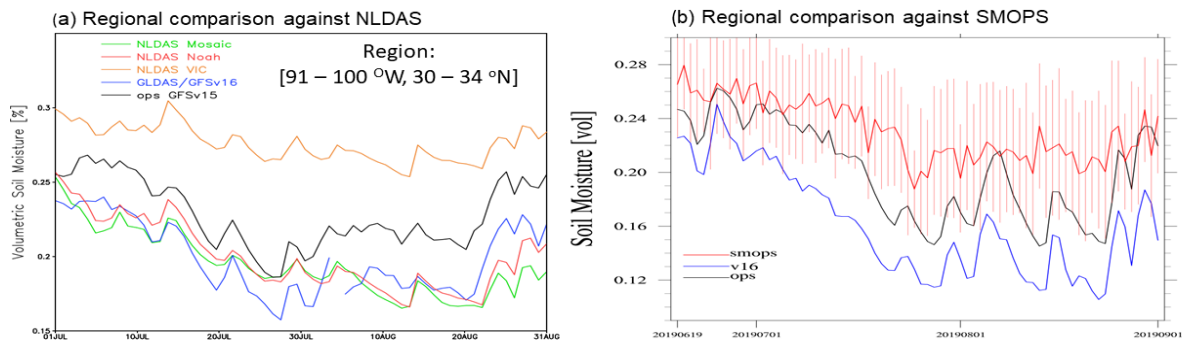

Figure 2. Regionally averaged top 10 cm GLDAS soil moisture comparison with ops GFSv15 (black line) and (a) NLDAS (Mosaic, Noah, VIC) and (b) SMOPS (red line). Vertical line is +/- 1 standard deviation representing SMOPS soil moisture variation.
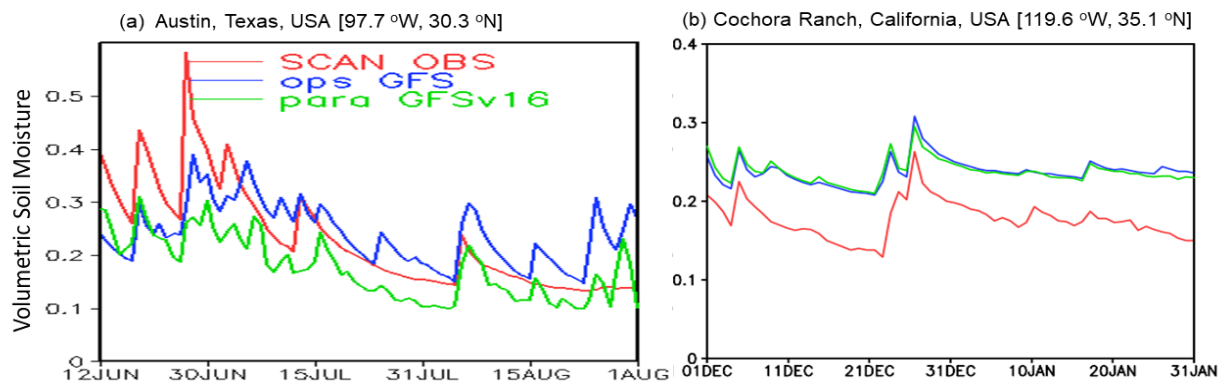

Figure 3. Comparison of top 10 cm GLDAS soil moisture with in-situ observations at Austin and Cochora Ranch. Soil Climate Analysis Network (SCAN) is included in ISMN and data is obtained from ISMN.

References

Meng, J., Yang, R., Wei, H., Ek, M., Gayno, G., Xie, P., and Mitchell, K., 2012: The Land Surface Analysis in the NCEP Climate Forecast System Reanalysis. *J. Hydromet.*, **13**, 1621-1630.

Xia, Y., Mitchell, K. E., Ek, M., Sheffield, J., et al., 2012: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project phase 2 (NLDAS-2): 1. Intercomparison and application of model products. *J. Geophys. Res.*, **117**, D03109, doi:10.1029/2011JD016048.

Xia, Y., Ek, M. B., Wu, Y., Ford, T., and Quiring, S. M., 2015: Comparison of NLDAS-2 Simulated and NASMD Observed Daily Soil Moisture, Part I: Comparison and Analysis. *J. Hydromet.,* **16**, 1962-1980.

Improving ensemble-based background error covariances of the hybrid 4DVar in JMA's global analysis

[1]YOKOTA Sho, [1]KADOWAKI Takashi, [1]ODA Mayuko, [2]OTA Yoichiro
[1]Japan Meteorological Agency, Japan
[2]Ministry of Education, Culture, Sports, Science and Technology, Japan
e-mail: s_yokota@met.kishou.go.jp

1. Introduction

The four-dimensional variational (4DVar) data assimilation system used to determine atmospheric analysis fields for JMA's operational global model (JMA 2019) was updated in December 2019 to a hybrid version with weighted averages for climatological and ensemble-based background error covariances (B_c and B_e, respectively) as initial background error covariances (Kadowaki et al. 2020). The initial ensemble forecast conditions for B_e are determined using a Local Ensemble Transform Kalman Filter (LETKF; Hunt et al. 2007), producing values to represent flow-dependent non-uniform uncertainties not represented by B_c. However, improvement based on B_e is limited due to insufficient ensemble size and hybrid covariance weight. This report outlines more effective use of B_e in hybrid 4DVar with increased ensemble size and hybrid covariance weight, and effects on forecasting.

2. Update Overview

B_e in JMA's 4DVar global analysis is created from three-hour ensemble forecasting initialized using the LETKF with 50 ensemble members (Kadowaki et al 2020). B_c and B_e are blended via the extended control variable method (Lorenc 2003). The weight for hybrid covariance is 0.85 for B_c and 0.15 for B_e below 50 hPa, with values approaching 1 and 0 above 50 hPa, respectively. Localization for B_e involves a Gaussian function application with scales of $1/\sqrt{e}$ set to 800 km (horizontally) and 0.8 scale height (vertically) in 4DVar, and 400 km (horizontally) and 0.4 scale height (vertically) in LETKF. Analysis from 4DVar is used to re-center LETKF analysis. This report details the settings updated as below.

(a) Ensemble size for B_e production is increased from 50 to 100 to suppress sampling errors.
(b) Hybrid covariance weights below 50 hPa are changed from 0.85 for B_c and 0.15 for B_e to 0.50 and 0.50, respectively, to reflect B_e values improved as a result of (a).
(c) Additional revisions:
   1) Horizontal and vertical weights for addition of divergence in initialization of LETKF analysis (Hamrud et al. 2015) are modified for global uniformity based on the global average of horizontal wind ensemble spread.
   2) The vertical localization scale for LETKF is expanded from 0.4 scale height to 0.6 scale height to reduce negative impacts in assimilating observation with vertical integrals (e.g., brightness temperature).
   3) The horizontal localization scale for specific humidity in 4DVar is reduced from 800 to 400 km to suppress specific humidity sampling errors in B_e. As a result, the horizontal scale of B_e between specific humidity and other variables, along with related peak values, is also smaller.

3. Update Effects

To verify the effects of updates (a) − (c), several sensitivity experiments on ensemble size, weights of hybrid covariances and additional revisions were conducted based on the configuration of JMA's operational global NWP system as of December 2019 for the period July 21 – September 11 2018. Figure 1 shows initial cost functions in 4DVar divided by the number of assimilated observations (profiles for brightness temperature), indicating the standard deviation of first-guess departure to assimilated observations normalized by the observation error standard deviation in 4DVar. This indicator was the smallest in the experiment with weight 0.30 for 50 ensemble members and 0.50 for 100 ensemble members. Revisions of (c) also produced slightly smaller values. Such improvements were additionally observed in the boreal winter experiment (not shown).

Comparison between the experiments without (CNTL) and with (TEST) (a) − (c) updates showed larger ensemble spreads in the latter, especially for the stratosphere. Root mean square errors in forecasts of geopotential height, temperature, zonal wind and specific humidity were also smaller, especially in the winter hemisphere (not shown). Position errors for tropical cyclone forecasts were smaller in the eastern North Pacific and the Atlantic where tropical cyclone bogus data were not

assimilated (Figure 2). These improvements were also observed in the boreal winter experiment (not shown).

4. Summary
The updates involving (a) increasing ensemble size from 50 to 100, (b) increasing the weight of hybrid covariances for B_e from 0.15 to 0.50 below 50 hPa, and (c) revising initialization and localization improved JMA's global analysis. These updates were applied in March 2021 (Ujiie et al. 2021).

References
Hamrud, M., M. Bonavita, and L. Isaksen, 2015: EnKF and hybrid gain ensemble data assimilation. Part I: EnKF implementation. *Mon. Wea. Rev.*, **143**, 4847–4864.
Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Physica. D.*, **230**, 112–126.
JMA, 2019: Outline of the operational numerical weather prediction at the Japan Meteorological Agency, Appendix to WMO Technical Progress Report on the Global Data-processing and Forecasting System (GDPFS) and Numerical Weather Prediction (NWP) Research. Japan Meteorological Agency, Tokyo, Japan.
Kadowaki, T., Y. Ota, and S. Yokota, 2020: Introduction of a new hybrid data assimilation system for the JMA Global Spectral Model. *WGNE. Res. Activ. Earth. Sys. Modell*, **50**, 1.9–1.10.
Lorenc, A. C., 2003. The potential of the ensemble Kalman filter for NWP: a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **129**, 3183–3203.
Ujiie, M., M. Higuchi, T. Kadowaki, Y. Kuroki, K. Miyaoka, M. Oda, K. Ochi, R. Sekiguchi, H. Shimizu, S. Yokota, and H. Yonehara, 2021: Upgrade of JMA's Operational Global Model. *WGNE. Res. Activ. Earth. Sys. Modell*, **51**, 6-09--6-10 (current volume).
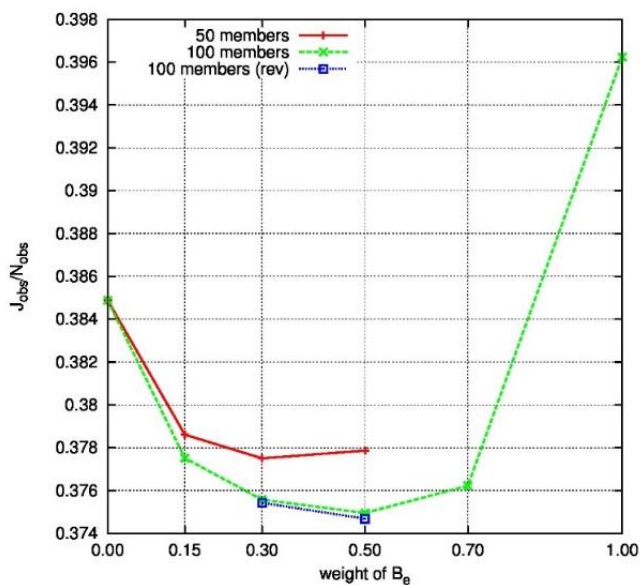
Figure 1. Initial cost functions in 4DVar divided by the number of assimilated observations (profiles for brightness temperature) averaged for August 1 − 31. The horizontal axis represents the weight of hybrid covariances for B_e below 50 hPa (and above 50 hPa for weight 1.00). The red and green lines show results with ensemble sizes of 50 and 100, respectively, without the additional revisions of (c). The blue line shows results with all revisions of (c) for 100 ensemble members. The weight points of 0.15 in the red line and 0.50 in blue line are CNTL and TEST, respectively.
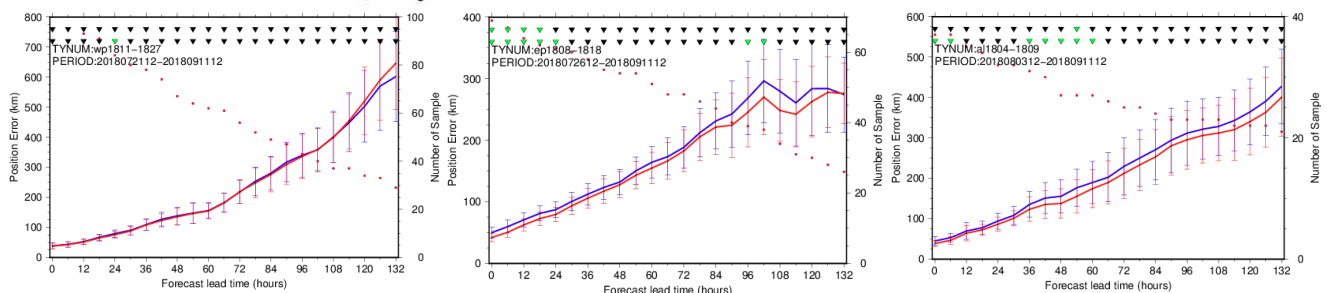

Figure 2. Average position errors (km) for tropical cyclones from July 21 to September 11 2018 in the western North Pacific (left), the eastern North Pacific (center) and the Atlantic (right). The blue and red lines show results from CNTL and TEST, respectively, and the red points are sample numbers. Error bars represent 95% confidence intervals.