

# Temporal and Spatial Interpolation Errors of High-Resolution Modeled Atmospheric Fields

N. C. PRIVÉ AND R. M. ERRICO

*Goddard Earth Sciences Technology and Research, Morgan State University, Greenbelt, Maryland*

(Manuscript received 24 June 2015, in final form 2 December 2015)

## ABSTRACT

General circulation models can now be run at very high spatial resolutions to capture finescale features, but saving the full-spatial-resolution output at every model time step is usually not practical because of storage limitations. To reduce storage requirements, the model output may be produced at reduced temporal and/or spatial resolutions. When this reduced-resolution output is then used in situations where spatiotemporal interpolation is required, such as the generation of synthetic observations for observing system simulation experiments, interpolation errors can significantly affect the quality and usefulness of the reduced-resolution model output. Although it is common in practice to record model output at the highest possible spatial resolution with relatively infrequent temporal output, this may not be the best option to minimize interpolation errors. In this study, two examples using a high-resolution global run of the Goddard Earth Observing System Model, version 5 (GEOS-5), are presented to illustrate cases in which the optimal output dataset configurations for interpolation have high temporal frequency but reduced spatial resolutions. Interpolation errors of tropospheric temperature, specific humidity, and wind fields are investigated. The relationship between spatial and temporal output resolutions and interpolation errors is also characterized for the example model.

## 1. Introduction

Modern numerical weather prediction models and general circulation models are integrated with internal time steps on the order of a few minutes or even seconds for physical and dynamical processes. The cost of storage and input/output processing prohibits saving the full temporal and spatial fields generated by the models. Although output is frequently saved at the full spatial resolution, it is very rarely saved at full temporal resolution. In practice, the output interval may be several hours for global models.

While infrequent model output is sufficient for many applications, there are some instances in which more frequent model output is desired. This includes postprocessing of model data, such as for trajectory calculations, the calculation of lateral boundary conditions for limited area models, or the generation of synthetic observations for observing system simulation

experiments (OSSEs). In these cases, temporal interpolation between model states may be used to reproduce the fields at intermediate times. However, this results in interpolation errors in the reproduced fields.

For phenomena that have small-scale features, sharp gradients, or rapid changes or translation of systems, temporal interpolation can smear out details and generate unrealistic atmospheric structures. Gorman (2009) showed examples for several methods of temporal interpolation of a tropical cyclone, where linear interpolation or interpolation using empirical orthogonal functions resulted in a deformed cyclone with two eyes. They found that interpolation using Fourier transforms gave a more realistic rendering of tropical cyclone structures but resulted in large errors over land regions and for larger, complex fields encompassing multiple synoptic features.

The effects of spatial and temporal resolution of fields on trajectory calculations have been investigated in several previous studies. Using a mesoscale model, Kuo et al. (1985) found that the temporal and spatial resolutions of the wind field need to be balanced, rather than one or the other having much higher resolution, in order

---

*Corresponding author address:* Nikki C. Privé, GESTAR, Morgan State University, Code 610.1, NASA/GSFC, Greenbelt, MD 20771.

E-mail: nikki.prive@morgan.edu

to reduce errors in the trajectory calculations. For their available dataset with 400-km horizontal resolution and 12-h temporal frequency, it was found that increasing the frequency of the data in time yielded the most improvement in trajectory error. Kahl and Samson (1986) performed a study using observations from a field campaign at 6-h temporal frequency and spacing on the order of 100–200 km to estimate the role of interpolation errors on trajectory calculations, and they found that spatial resolution had greater impact on interpolation errors than temporal resolution. Stohl et al. (1995) performed a study using the then-operational ECMWF model (0.5° horizontal resolution and 3-h temporal resolution as the validating model output), focusing on trajectory calculations. They found that the largest component of interpolation error for meridional wind was due to temporal interpolation, with a smaller error resulting from spatial interpolation.

When a planned model run is expensive in terms of computation and data storage, such as a nature run used for OSSEs, there is a trade-off between loss of information and the expense of storing and processing the model output. The main goal of this work is to illustrate that recording full-spatial-resolution model output at reduced temporal intervals may not be the optimal choice. An example of the interpolation errors that can result from reduced spatial and/or temporal model output is demonstrated using the Global Earth Observing System, version 5 (GEOS-5; Rienecker et al. 2008), forecast model developed at the National Aeronautics and Space Administration's Global Modeling and Assimilation Office (NASA GMAO). The character of the spatial and temporal interpolation errors for a range of reduced-resolution output is evaluated. A comparison of the interpolation errors that result from different configurations of combined reductions in spatial and temporal resolution is also performed to illustrate the process of selecting the output resolution of a large dataset with storage limitations.

The high-resolution model and interpolation methods are described in section 2. Characterization of interpolation errors is detailed in section 3, and examples of the selection of a suitable dataset resolution are illustrated in section 4. The results are discussed in section 5.

## 2. Method

A high-resolution global model integration is first generated over a test period. In this test, GEOS-5 was run for a 75-h forecast beginning at 2100 UTC 14 June 2012, with approximately 1.5-km horizontal resolution

(cube–sphere grid with 5760 points along each edge of the cube) and 72 hybrid vertical levels, with internal model time steps of 120 s for the model physics and 1.25 s for the dynamics. Because of the large size of the model output, only selected fields are recorded; for example, state variables are only output at a few pressure levels but on the full-resolution cube–sphere grid. The model run was not generated specifically for this study, and output fields were available only at 10-min intervals. The first 27 h of the model run are excluded from calculations to allow for spinup to the 1.5-km resolution with an initial state from a replay of the Modern-Era Retrospective Analysis for Research and Applications (MERRA; Rienecker et al. 2011) reanalysis at 12.5-km resolution.

To determine the error due to temporal interpolation, a simple method is adopted, similar to the techniques used by Kuo et al. (1985) and Stohl et al. (1995). The full set of available model output is considered as “truth” and called the BASE set (Putman et al. 2016). Degraded output sets are then considered where only the model output at time intervals  $m\Delta t$  from the BASE set is used, where  $m$  is a whole number and  $\Delta t$  is 10 min. The “missing” output times are then recreated by linearly interpolating in time between the fields available in the degraded output set. The error of the interpolated fields is then directly calculated by comparison against the BASE fields. This procedure is performed for  $m$  of 2, 3, 4, 6, 9, 12, and 18 corresponding to time intervals of 20, 30, 40, 60, 90, 120, and 180 min.

A similar method is used to determine the error due to spatial interpolation. The BASE output fields are degraded by using only the model output at intervals  $n\Delta x$ , where  $n$  is a whole number and  $\Delta x$  is a single grid width of approximately 1.5 km. A bilinear interpolation of the degraded output is then used to fill in the fields at the full resolution of the BASE output and compared to calculate the error. The degradation and interpolation are performed for  $n$  of 2, 3, 4, 6, 9, 12, and 18, corresponding to spatial intervals of 3, 4.5, 6, 9, 13.5, 18, and 27 km.

The root-mean-square error (RMSE) of the interpolated field  $x$  in comparison to the truth  $t$  is calculated as

$$\text{RMSE} = \sqrt{\frac{1}{kl} \sum_{j=1}^l \sum_{i=1}^k a_i (x_{i,j} - t_{i,j})^2}, \quad (1)$$

where there are  $l$  times during the period of interest,  $a_i$  is an area-weighting term corresponding to the fractional size of the  $i$ th grid box compared to the total grid area, and there are  $k$  grid points. To avoid difficulties in

interpolation along the edges of the cube–sphere grid, areas within 20 grid points of an edge or 200 grid points of a corner of a face are omitted from all calculations. Calculations are performed over the period from 2350 UTC 15 June 2012 to 2350 UTC 17 June 2012. The RMSE over the 2-day period is calculated separately for each of three regions: 30°–60°N [Northern Hemisphere (NHEX)], 30°–60°S [Southern Hemisphere (SHEX)], and 30°S–30°N (tropics). Any pressure surface points that were below ground were excluded from computations, including the areal weighting term.

Linear temporal interpolations and bilinear spatial interpolations are chosen here because these are the types of interpolation commonly used when generating synthetic observations for OSSEs. Kahl and Samson (1986) found that linear temporal interpolation resulted in smaller temporal interpolation errors than cubic spline or variable acceleration interpolation methods. More sophisticated interpolation methods, such as spectral methods and empirical orthogonal functions, may result in smaller interpolation errors but at a computational expense that becomes onerous when working with large datasets. The crude method of reducing the spatial resolution of the fields used here is chosen because it is the closest analog in space to infrequent output of temporal states.

### 3. Characterizing interpolation errors

The character of interpolation errors for this model configuration is first investigated by modifying the temporal resolution of the output while the spatial resolution is maintained at the full 1.5-km grid. The spatial resolution of the output is then similarly varied while keeping the 10-min temporal resolution. The time mean RMSE calculated over the 2-day period is shown in Fig. 1 for all three regions as a function of resolution, where temporal and spatial interpolation errors for each variable are displayed with the same vertical scale for comparison. Both spatial and temporal interpolation errors have RMSEs that appear to asymptote as the resolution degrades, although the asymptotic value is not yet reached at the lowest resolutions tested.

For spatial interpolation errors (left panels), a higher RMSE indicates that the field of interest has more noise and/or features with smaller length scales. A lower RMSE for spatial interpolation errors indicates a BASE field that is spatially smooth with structures having longer length scales. Similarly, a high RMSE for temporal interpolation errors (right panels) indicates a rapidly changing field or one with large temporal variance, while a low RMSE indicates a field that changes slowly or has a small temporal variance.

The RMSEs for zonal wind demonstrate that the same variable can have different behaviors at different levels of the atmosphere. At 850 hPa, the RMSE of wind for spatial interpolation is very similar for all three regions, while at 250 hPa there is significant spread in RMSE between the different regions. At 250 hPa, the greatest spatial interpolation RMSEs occur in the tropics, with the lowest RMSEs in the SHEX. This may be due to convective outflow at upper levels in the tropics and to a lesser extent in the summer hemisphere extratropics, while the winter hemisphere winds are relatively smooth at upper levels. The RMSEs for spatial interpolation of wind at 850 hPa are of greater magnitude than the RMSEs at 250 hPa, even though wind speeds are greater in the upper troposphere and the RMSE is not normalized.

Temporal interpolation errors of wind do not show an asymptotic relationship between RMSE and resolution at upper levels, with a more clearly asymptotic relationship in the lower troposphere. In particular, the SHEX RMSE of 250-hPa wind has the most linear relationship with resolution, indicating that the wind field in that region changes on longer time scales. At both upper and lower levels, the RMSEs for temporal interpolation of wind are very similar for the tropical and NHEX regions. This similarity may be due to the influence of moist convection on the wind field, given the summer conditions in the NHEX.

As with wind, the magnitude of RMSE for temperature interpolation is considerably larger at 850 than at 250 hPa. At 850 hPa, the spatial interpolation RMSE is very similar for the NHEX and the SHEX, which is unexpected due to the seasonal differences between the two hemispheres as well as the geographical/orographical differences. Interpolation errors of specific humidity at 850 hPa show clearly asymptotic behavior for both spatial and temporal interpolations, with the largest RMSE in the moist tropics and the smallest RMSE in the SHEX region, where humidity is low in the winter season.

### 4. Selecting dataset resolutions

When generating a long model run of a high-resolution global model, storage requirements for the output dataset are an important concern. For example, the recently generated GEOS-5 nature run (Gelaro et al. 2015) used a 7-km horizontal resolution, 72-level version of the GEOS-5 with 30-min output over a 2-yr forecast integration, and required approximately 4 PB of storage. A similar nature run generated with the 1.5-km, 72-level, 10-min resolution used in the BASE dataset of this study and output on the native cube–sphere grid would require approximately 144 PB of storage,

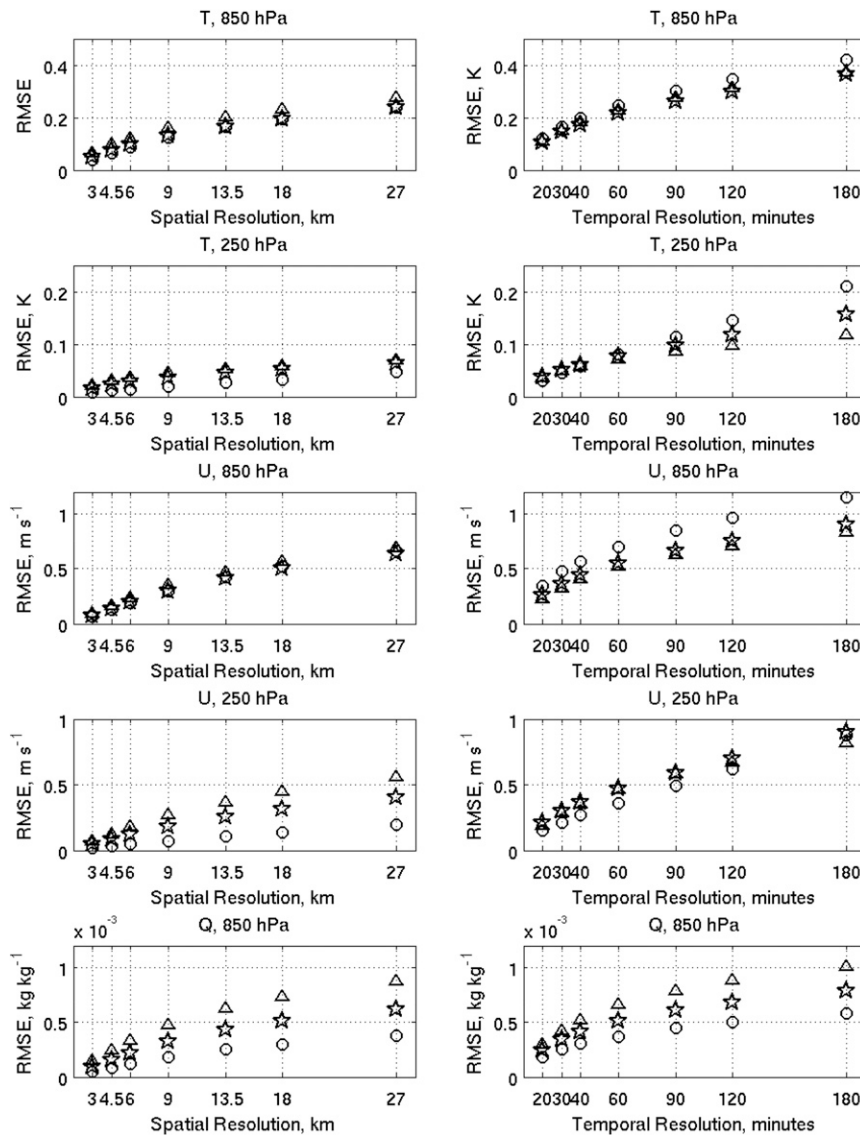


FIG. 1. RMSE for temporal and spatial interpolations calculated for NHEX (star), SHEX (circle), and tropics (triangle) at various resolutions (abscissa) for temperature (K), zonal wind ( $\text{m s}^{-1}$ ), and specific humidity ( $\text{kg kg}^{-1}$ ) at 250 and 850 hPa. (left) Spatial interpolation error and (right) temporal interpolation error.

which is not reasonable given currently available resources.

If storage is expected to be a limitation when generating such a large dataset, then a simple test can be used to determine the best temporal and spatial resolutions of the output dataset. First, the selected model is run for a short period with output at full spatial resolution and at a temporal resolution higher than a reasonable frequency for the full dataset. Then different configurations of output resolution can be compared by the method of artificially reducing the resolution and interpolating to reproduce the full fields as described in section 2. The

tested configurations are chosen to result in a full dataset that meets the storage limitations.

Two examples are illustrated in this section to demonstrate the procedure. In the first example, it is assumed that the file output size must be reduced by a factor,  $S$ , of at least 16 compared to the 1.5-km horizontal resolution, 10-min output dataset. Four combinations of spatial and temporal resolutions are compared: 1.5-km, 180-min output; 3-km, 40-min output; 4.5-km, 20-min output; and 6-km, 10-min output. For these simple illustrations, the RMSE interpolation errors are calculated for only two fields: 850-hPa temperature and 250-hPa winds. In

TABLE 1. RMSEs for spatial and temporal interpolation of datasets with the resolution reduced to decrease storage requirements by a factor of  $S \geq 16$  compared to BASE. Boldface values indicate the lowest interpolation error for that variable and region for the set of four comparison cases.

Output resolution			850-hPa $T$			250-hPa $U$		
(km)	(min)	$S$	NHEX	SHEX	Tropics	NHEX	SHEX	Tropics
1.5	180	18	0.38	0.44	0.39	0.95	0.96	0.83
3.0	40	18	0.17	0.19	0.18	0.36	0.25	0.35
4.5	20	18	0.12	0.12	0.13	0.22	0.14	0.21
6.0	10	16	<b>0.10</b>	<b>0.09</b>	<b>0.12</b>	<b>0.12</b>	<b>0.05</b>	<b>0.18</b>

practice, a collection of the most important variables for the study of interest should be tested.

Table 1 shows the RMSE for interpolation errors calculated over the same time period and regions described in section 2. The lowest interpolation RMSEs are found for the case with 6-km horizontal resolution and 10-min output. The largest RMSEs are found for the case with 1.5-km horizontal resolution and 180-min output, which is expected given the results seen in Fig. 1.

For this first example, the choice of output resolution might be influenced by the primary region and fields of interest. While the 6-km, 10-min output resolution gives the lowest interpolation errors overall, the 4.5-km, 20-min output configuration has slightly smaller storage requirements ( $S$  of 18 compared to 16). Comparing these two cases, the 850-hPa temperature interpolation errors are of the same magnitude, but the 250-hPa zonal wind interpolation errors in the midlatitudes are much lower in the case with 6-km, 10-min resolution.

In the second example, the same procedure is demonstrated but for a more limited storage capability requiring a reduction in output dataset size by a factor of at least 32. In this example, four cases are compared: 3-km, 90-min output; 4.5-km, 40-min output; 6-km, 20-min output; and 9-km, 10-min output. An additional case with 1.5-km, 40-min output ( $S = 4$ ) is also included for comparison as discussed in greater detail later in this section. The interpolation RMSEs are displayed in Table 2.

For 850-hPa temperatures in the extratropics, the two configurations with the lowest RMSE were the 6-km, 20-min output case and the 9-km, 10-min output case. The 6-km, 20-min case had the lowest RMSE for both temperature and wind in the tropics. For the midlatitudes, the lowest wind interpolation errors occurred for the 9-km horizontal resolution, 10-min output case. For midlatitude 850-hPa temperatures, the RMSE was nearly the same for both the 6-km, 10-min case and the 6-km, 20-min case. The 9-km, 10-min resolution configuration has a greater savings in storage requirements than the 6-km, 20-min resolution configuration, so the choice of the best

TABLE 2. RMSEs for spatial and temporal interpolation of datasets with the resolution reduced to decrease storage requirements by a factor of  $S \geq 32$  compared to BASE. Boldface values indicate the lowest interpolation error for that variable and region for the set of four comparison cases.

Output resolution			850-hPa $T$			250-hPa $U$		
(km)	(min)	$S$	NHEX	SHEX	Tropics	NHEX	SHEX	Tropics
3.0	90	36	0.26	0.30	0.27	0.58	0.47	0.58
4.5	40	36	0.17	0.19	0.18	0.36	0.25	0.35
6.0	20	32	<b>0.13</b>	<b>0.13</b>	<b>0.14</b>	0.22	0.14	<b>0.24</b>
9.0	10	36	<b>0.13</b>	<b>0.13</b>	0.16	<b>0.18</b>	<b>0.07</b>	0.27
1.5	40	4	0.18	0.20	0.19	0.38	0.31	0.34

output resolution would depend on the region of greatest interest and the availability of computing resources.

It is instructive to compare the RMSE of the interpolated fields from the cases with a combination of both spatial and temporal resolutions with the RMSE from the previously calculated cases with only a single aspect of degraded resolution. For example, the cases with 4.5-km, 40-min output and 3-km, 40-min output can be compared to the case with 1.5-km, 40-min output, listed at the bottom of Table 2. Surprisingly, the RMSEs of 250-hPa winds are larger for the 1.5-km case than the lower-spatial-resolution cases, as well as for the midlatitude 850-hPa temperature.

This result can be explained by examining the spatial variances and covariances of the interpolated fields and the verifying BASE fields. The spatial variance of a horizontal field  $t$  averaged over  $k$  grid points is

$$\text{Var}(t) = \frac{1}{k} \sum_{i=1}^k a_i (t_i - \mu)^2, \tag{2}$$

where  $\mu$  is the area-weighted mean of all  $t$ . The mean-square error (MSE) of interpolated field  $x$ , which is an estimate of  $t$  over the same region at a particular time, can be defined as

$$\text{MSE} = \frac{1}{k} \sum_{i=1}^k a_i (x_i - t_i)^2, \tag{3}$$

which can be related to the variances and covariance of  $t$  and  $x$  by assuming that there is no bias between  $x$  and  $t$ , so that  $\mu$  is the mean of both fields:

$$\begin{aligned} \text{MSE} &= \frac{1}{k} \sum_{i=1}^k a_i [(x_i - \mu) - (t_i - \mu)]^2, \\ &= \frac{1}{k} \sum_{i=1}^k a_i [(x_i - \mu)^2 + (t_i - \mu)^2 - 2(x_i - \mu)(t_i - \mu)], \\ &= \text{Var}(x) + \text{Var}(t) - 2 \text{Cov}(x, t). \end{aligned} \tag{4}$$

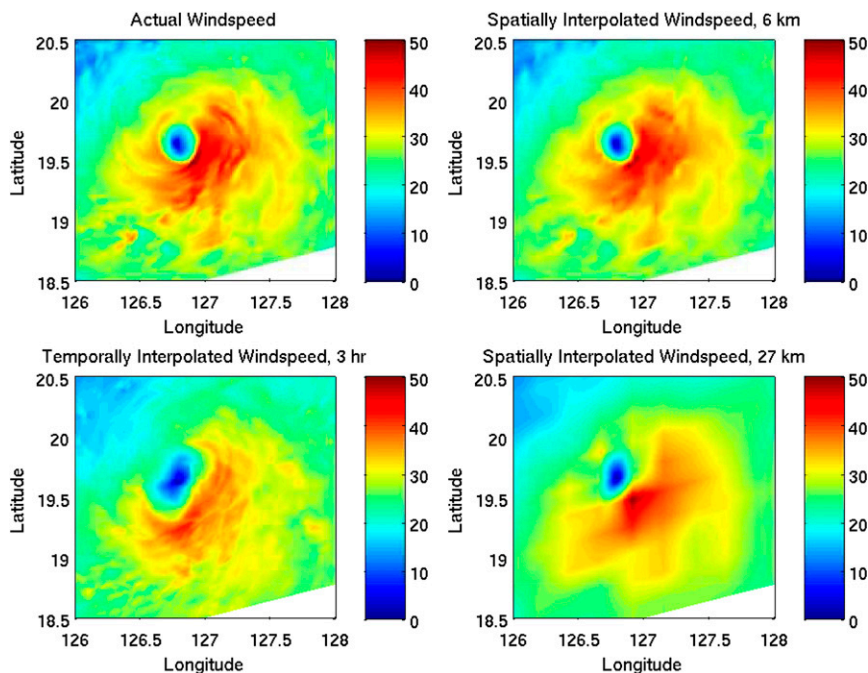


FIG. 2. Snapshot of 850-hPa horizontal wind speed at 2120 UTC 17 Jun 2012 for Typhoon Guchol: (top left) actual wind speed from BASE, (top right) spatially interpolated wind speed degraded to 6-km resolution, (bottom left) temporally interpolated wind speed with 3-h degraded resolution (center time shown), and (bottom right) spatially interpolated wind speed degraded to 27-km resolution. All units in  $\text{m s}^{-1}$ ; all panels have the same color scale.

At times that fall exactly on the 40-min intervals where the interpolation occurs in space only, the 1.5-km configuration is expected to have smaller RMSE than the 3-km or 4.5-km cases, following the results of the previous section. For intermediate times when both spatial and temporal interpolations occur however, the MSE of the 1.5-km case may be larger. Calculating the variance and covariance terms in (4) at intermediate times (not shown) reveals that the  $\text{Var}(x)$  term can be smaller in the 3- and 4.5-km cases than the 1.5-km case, while the  $2\text{Cov}(x, t)$  term does not have as large a decrease.

In these cases, small-scale features are generally associated with short time scales, even if the dynamical mechanism is simply advection. When the spatial resolution is degraded, these small-scale features are removed, thus predominantly eliminating elements with high temporal frequency. These high-temporal-frequency features are a source of error when interpolating in time, so the fields where resolution is degraded in both space and time can have a smaller RMSE than fields that have degraded temporal frequency but high spatial resolution.

## 5. Discussion

The push for ever-finer model resolution is often lead by the desire to capture smaller and smaller spatial scales, with

the internal time step of the model dictated by the Courant–Friedrichs–Lewy condition (Courant et al. 1928) and/or demands of the model physics. Many important atmospheric processes occur at finescales, and the finer model resolution ideally allows less dependence on parameterizations while producing more realistic results. It can therefore be compelling to try to capture the finest spatial resolution available in the model output, but this will necessarily come at the expense of the temporal frequency of the output given limited resources. If the required fields will be primarily at intermediate times to those available in the recorded dataset, the temporal interpolation errors may overwhelm any benefit of the high spatial resolution. A compromise between the desire for fine spatial features, the expense of high-temporal-resolution output, and the limitations of storage and processing capabilities has been sought here.

Degrading the spatial resolution of the fields results in a loss of small-scale information, but interpolation of degraded temporal output can cause damage across a range of spatial scales. As illustrated by Gorman (2009), atmospheric structures can be deformed in unrealistic and nonphysical ways due to temporal interpolation. Even large, coherent features can be significantly deformed by temporal interpolation if the rate of change of

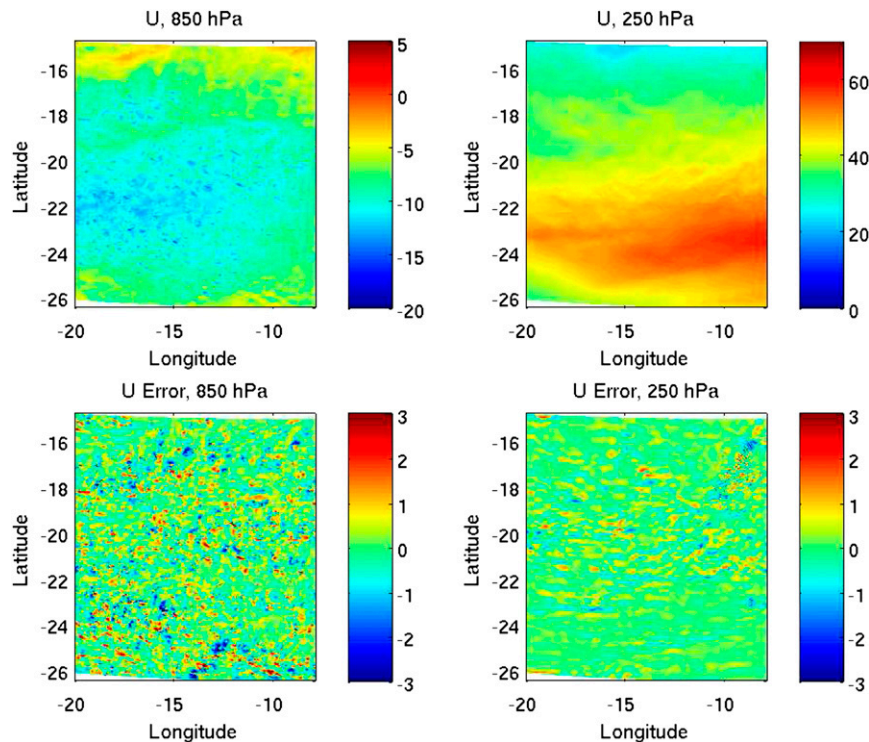


FIG. 3. Snapshot of (top) zonal wind field and (bottom) zonal wind interpolation errors for a resolution of 27 km for 2350 UTC 15 Jun 2012. Wind and error at (left) 850 and (right) 250 hPa. All units in  $\text{m s}^{-1}$ ; note the different color scale for the top two panels.

the feature is rapid. For example, temporal interpolation of a tropical cyclone can result in a cyclone with two smeared “eyes,” and temporal interpolation of a moving jet results in a pair of weaker, blurred jets.

An illustration of this effect from the model run considered here comparing the spatial and temporal interpolations of Typhoon Guchol in the western Pacific is shown in Fig. 2. The true model output wind speed at 850 hPa near the center of the tropical cyclone is shown in the top-left panel for BASE. The lower-right panel shows the effect of reducing the spatial resolution of the wind speed to 27 km and then interpolating to the full resolution. While many small-scale features are lost, the strongest wind speeds are comparable to the full-resolution field, there are some hints of banding structures, and the eye has a smooth ellipsoid shape. In the bottom-left panel, the temporal resolution of the wind speed is reduced to 3-hourly output and interpolated to the original time (with the displayed time as the halfway between the output times for greatest illustration). For the temporal interpolation, the wind speeds are weaker and the shape of the eye is considerably deformed, as is the wind field in the eyewall and inner cyclone. While the temporally interpolated field appears to contain many smaller features, these features do not correspond to the features in the corresponding BASE

field. As a comparison, the top-right panel shows the wind speed spatially interpolated to 6-km resolution, the resolution from the first example in section 4 with the lowest RMSE. While some of the finescale features are lost at 6-km resolution, many structures are still resolved, and the overall depiction of the tropical cyclone is realistic.

The behavior of spatial and temporal interpolation errors has been characterized over a range of different output resolutions given a particular baseline integration of the GEOS-5 forecast model. For most of the variables tested, the interpolation errors showed asymptotic behavior with degraded spatial and temporal resolution. Saturation of the interpolation RMSE at an asymptotic state is an indication that the information content of the full-resolution state is lost at these levels of reduced resolution. This behavior is not expected for the range of resolutions tested here but would be anticipated as the resolution was much more drastically degraded. In contrast, an RMSE that continues to increase with reduced resolution indicates that some information from the full-resolution state is retained at these more coarse resolutions. Therefore, those fields for which the RMSE begins to asymptote quickly as the resolution is degraded are the fields of greatest concern. None of the examined fields are close to saturation at the

coarsest resolutions tested here; however, the RMSEs for most fields show some signs of asymptotic behavior.

One unexpected finding was the larger magnitude of errors in the lower troposphere compared to the upper troposphere for both wind and temperature. As the RMSE was not normalized, the upper troposphere might be anticipated to yield greater wind interpolation errors as the spatial mean and variance of wind are much larger than in the lower troposphere. The results of [Stohl et al. \(1995\)](#) also found that the interpolation errors increased with larger wind speed, especially for temporal interpolations. On the other hand, [Kahl and Samson \(1986\)](#) found that spatial interpolation errors were greater with higher wind speed, but temporal interpolation errors were not dependent on wind speed.

The cause of larger RMSEs at 850 hPa is illustrated in [Fig. 3](#) for a region over the southern Atlantic Ocean, showing a snapshot of the wind fields at 1.5-km resolution from a randomly chosen time (2350 UTC 15 June) and the corresponding errors due to spatial interpolation from the same field reduced to 27-km resolution. The upper-level wind field has a strong jet streak with high wind speeds, but the wind field is relatively smooth, with the greatest interpolation errors occurring not near the jet maximum but instead flanking the jet to the north. However, at lower levels, while wind speeds are relatively low compared to upper levels, the wind field has a significant amount of variance at small spatial scales over the entire region shown. Thus, in (4), at lower levels the covariance of the true field and the interpolated field is small relative to the variance of each field, while at upper levels the covariance remains large relative to the variance of each field. These small-scale features in the wind field are expected to have a short time scale for variance, which would similarly lead to the higher temporal interpolation RMSEs in the lower troposphere compared to the upper troposphere.

A direct practical application of this work regards decision-making when planning a nature run for an OSSE. One of the limitations when generating a nature run is the massive storage needed for high-resolution, long model forecasts with many output fields. The two example cases in [section 4](#) illustrated one method for determining the best output resolution of a large dataset suitable for interpolation that meets specific storage requirements. For both examples, the optimal configuration had relatively high temporal output frequency with some reduction in spatial resolution compared to the full model resolution. It was also found that as temporal output resolution is degraded, very high spatial resolution does not improve interpolation errors

and may actually result in greater interpolation errors compared to reduced spatial resolutions. Large savings in file output size can also be obtained with a relatively small degradation in horizontal resolution due to two-dimensionality.

It is not possible to use these results to extrapolate interpolation errors that might occur for different models or different internal model resolutions. However, the methods demonstrated in this manuscript require only modest computing resources and could be performed before the generation of computationally expensive large datasets. In addition to or instead of the limited variables examined here, the fields of interest for the intended studies could be tested in a similar manner. Some variables, such as cloud fraction, and integrated quantities, such as precipitation, may not be well suited for the type of simple interpolation performed here. Vertical interpolation was not examined here, but it could also be tested if the vertical coordinate were to be transformed between the model and output fields. Thoughtful testing of the dataset output resolutions can save considerable time and expense and may yield better ultimate results.

*Acknowledgments.* Model output fields were kindly provided by William Putman. Support for this project was encouraged by Steven Pawson and provided by NASA GMAO core funding. Comments from four anonymous reviewers led to significant improvements in the manuscript.

## REFERENCES

- Courant, R., K. Friedrichs, and H. Lewy, 1928: Über die partiellen differenzengleichungen der mathematischen physik. *Math. Ann.*, **100**, 32–74, doi:10.1007/BF01448839.
- Gelaro, R., and Coauthors, 2015: Evaluation of the 7-km GEOS-5 nature run. R. D. Koster, Ed., Tech. Rep. Series on Global Modeling and Data Assimilation, Vol. 36, NASA Tech. Memo. NASA/TM2014-104606, 285 pp. [Available online at <http://gmao.gsfc.nasa.gov/pubs/docs/Gelaro736.pdf>.]
- Gorman, R., 2009: Intercomparison of methods for the temporal interpolation of synoptic wind fields. *J. Atmos. Oceanic Technol.*, **26**, 828–837, doi:10.1175/2008JTECH0588.1.
- Kahl, J., and P. Samson, 1986: Uncertainty in trajectory calculations due to low resolution meteorological data. *J. Climate Appl. Meteor.*, **25**, 1816–1831, doi:10.1175/1520-0450(1986)025<1816:UITCDT>2.0.CO;2.
- Kuo, Y.-H., M. Skumanich, P. Haagenson, and J. Chang, 1985: The accuracy of trajectory models as revealed by the observing system simulation experiments. *Mon. Wea. Rev.*, **113**, 1852–1867, doi:10.1175/1520-0493(1985)113<1852:TAOTMA>2.0.CO;2.
- Putman, W. M., M. J. Suarez, and A. Trayanov, 2016: Preparing NASA global modeling and assimilation systems for future HPC applications/architectures. *Second Symp. on High Performance Computing for Weather, Water, and Climate*, New Orleans,



- LA, Amer. Meteor. Soc., 4.2. [Available online at <https://ams.confex.com/ams/96Annual/webprogram/Paper281668.html>.]
- Rienecker, M., and Coauthors, 2008: The GEOS-5 data assimilation system—Documentation of versions 5.0.1, 5.1.0, and 5.2.0. M. J. Suarez, Ed., Tech. Rep. Series on Global Modeling and Data Assimilation, Vol. 27, NASA Tech. Memo. NASA/TM-2008-104606, 101 pp. [Available online at <http://gmao.gsfc.nasa.gov/pubs/docs/Rienecker369.pdf>.]
- , and Coauthors, 2011: MERRA: NASA's Modern-Era Retrospective Analysis for Research and Applications. *J. Climate*, **24**, 3624–3648, doi:10.1175/JCLI-D-11-00015.1.
- Stohl, A., G. Wotawa, P. Siebert, and H. Kromp-Kolb, 1995: Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories. *J. Appl. Meteor.*, **34**, 2149–2165, doi:10.1175/1520-0450(1995)034<2149:IEIWFA>2.0.CO;2.