University of Reading

The effects of the RAW filter on the climatology and forecast skill of the SPEEDY model

Article

Published Version

Amezcua, J., Kalnay, E. and Williams, P. D. ORCID: https://orcid.org/0000-0002-9713-9820 (2011) The effects of the RAW filter on the climatology and forecast skill of the SPEEDY model. Monthly Weather Review, 139 (2). pp. 608-619. ISSN 0027-0644 doi: https://doi.org/10.1175/2010MWR3530.1 Available at https://centaur.reading.ac.uk/19622/

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.
Published version at: http://dx.doi.org/10.1175/2010MWR3530.1
To link to this article DOI: http://dx.doi.org/10.1175/2010MWR3530.1

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.

www.reading.ac.uk/centaur

University of Reading

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

The Effects of the RAW Filter on the Climatology and Forecast Skill of the SPEEDY Model

JAVIER AMEZCUA AND EUGENIA KALNAY

Department of Atmospheric and Oceanic Sciences, University of Maryland, College Park, College Park, Maryland

PAUL D. WILLIAMS

Department of Meteorology, University of Reading, Reading, United Kingdom

(Manuscript received 4 June 2010, in final form 2 September 2010)

ABSTRACT

In a recent study, Williams introduced a simple modification to the widely used Robert–Asselin (RA) filter for numerical integration. The main purpose of the Robert–Asselin–Williams (RAW) filter is to avoid the undesired numerical damping of the RA filter and to increase the accuracy. In the present paper, the effects of the modification are comprehensively evaluated in the Simplified Parameterizations, Primitive Equation Dynamics (SPEEDY) atmospheric general circulation model. First, the authors search for significant changes in the monthly climatology due to the introduction of the new filter. After testing both at the local level and at the field level, no significant changes are found, which is advantageous in the sense that the new scheme does not require a retuning of the parameterized model physics. Second, the authors examine whether the new filter improves the skill of short- and medium-term forecasts. January 1982 data from the NCEP–NCAR reanalysis are used to evaluate the forecast skill. Improvements are found in all the model variables (except the relative humidity, which is hardly changed). The improvements increase with lead time and are especially evident in medium-range forecasts (96–144 h). For example, in tropical surface pressure predictions, 5-day forecasts made using the RAW filter have approximately the same skill as 4-day forecasts made using the RA filter. The results of this work are encouraging for the implementation of the RAW filter in other models currently using the RA filter.

1. Introduction

There are several time-stepping schemes for the numerical integration of the differential equations representing the evolution of a dynamical system (e.g., Durran 1991). The particular scheme chosen for any given integration will depend upon a compromise between the desired accuracy, stability, computational efficiency, ease of implementation, and run-time memory requirements. While it is always hoped that simulations will be insensitive to time-stepping choices, the evidence suggests that this hope may be forlorn (e.g., Pfeffer et al. 1992; Williamson and Olson 2003; Zhao and Zhong 2009). Therefore, the following question naturally arises: Which of the many possible time-stepping schemes offers the most realistic simulations for the least computational expense?

A centered time-stepping scheme known as the leapfrog—and specifically the Robert–Asselin (RA) filtered version—is a widely used option in contemporary models of the atmosphere and ocean. This popularity is mainly due to three factors: the ease of implementation, the low computational expense (only one evaluation of the model's tendency is needed per time step), and the low run-time storage requirements. The most serious problem associated with the leapfrog scheme is the ''time splitting'' instability associated with the creation of a spurious computational mode. The RA filter provides a considerable amelioration of this problem. The application of this filter, however, while damping the computational mode, can also have the undesired effect of significantly damping the physical mode of the solution, hence degrading its accuracy. In recent work, Williams (2009) introduced a simple modification to the RA filter, with the objective of improving its performance while avoiding its associated

Corresponding author address: Javier Amezcua, Department of Atmospheric and Oceanic Sciences, University of Maryland, College Park, College Park, MD 20742-2425.
E-mail: jamezcua@atmos.umd.edu

DOI: 10.1175/2010MWR3530.1

© 2011 American Meteorological Society

problems; the modification will hereafter be referred to as the Robert–Asselin–Williams (RAW) filter.

To date, the effects of the RAW filter have been tested only in a simple linear model representing inertial oscillations of the simple harmonic type (Williams 2009). In the present paper, the filter will be implemented and tested in the Simplified Parameterizations, Primitive Equation Dynamics (SPEEDY) model (Molteni 2003), a relatively simple nonlinear atmospheric general circulation model (AGCM). Our objectives are to examine whether the use of the RAW filter changes either the climatology or the skill of weather forecasts, or both. For the first objective, we will calculate the local and field significance, following Livezey and Chen (1983). For the second objective, we will calculate the anomaly correlation coefficient (ACC) and the root-mean-square error (RMSE), using base data from the National Centers for Environmental Prediction (NCEP) reanalysis. For both the climatology and the forecasts, we will assess whether the upgrade from RA filter to RAW filter causes any significant changes.

This paper is organized as follows. Section 2 contains a short description of the RAW filter. Section 3 gives a brief description of the SPEEDY model in which we are testing the filter. Section 4 studies the effects of the RAW filter on the climatology of the model; it is divided into two subsections, assessing the local significance and the field significance of the variables, respectively. In section 5, we look for improvements in the skill of short- and medium-term weather forecasts due to the introduction of the RAW filter. Section 6 concludes the paper with a summary and discussion.

2. The RAW filter

The centered discretization scheme known as the leapfrog is implemented as follows:

$$\frac{\partial x}{\partial t} = F(x) \Rightarrow x_{n+1} = x_{n-1} + 2\Delta t F(x_n). \tag{1}$$

The leapfrog scheme is a widely used numerical integration method, in particular for hyperbolic equations and complex models. There are two main reasons for this. First, being a centered scheme, it is reasonably accurate and has an error of order $O(\Delta t)^2$. Second, it requires only one computation of the time derivative per time step, and is therefore reasonably computationally efficient. The leapfrog scheme, however, introduces into the solution of the equation a spurious computational mode besides the actual physical mode (e.g., Kalnay 2003). This undesired mode manifests itself in nonlinear integrations as a spurious, growing oscillation between even and odd time steps.

Several approaches have been proposed to combat the growth of the computational mode; the most widely used is the RA filter. This filter was introduced by Robert (1966) and was shown by Asselin (1972) to suppress the computational mode while leaving the physical mode untouched for low frequencies with long periods compared to the time step Δt. The RA filter is implemented in leapfrog integrations as follows:

$$\begin{aligned} x_{n+1} &= \overline{x}_{n-1} + 2\Delta t F(x_n) \\ \overline{x}_n &= x_n + \frac{\nu}{2}[x_{n+1} - 2x_n + \overline{x}_{n-1}]. \end{aligned} \tag{2}$$

The smoothing parameter ν in Eq. (2) is usually chosen to be $O(0.01 - 0.2)$. The choice of this parameter is important: if its value is too small it can hardly manage to dampen the computational mode, but if it is too large it can lead to loss of accuracy in the solution. In his original analysis, Asselin (1972) studied values up to $\nu = 0.2$. For atmospheric models, Durran (1991) notes that values of $\nu = 0.12$ are typically used in the National Center for Atmospheric Research (NCAR) community (Williamson 1983); Déqué and Cariolle (1986) consider values as high as $\nu = 0.2$ and so does the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model (MOM). For oceanic models Kantha and Clayson (2000) recommend values between $\nu = 0.1$ and $\nu = 0.3$.

Although the RA filter is widely used in operational and research models of the atmosphere and ocean (Williams 2009), it has two related problems. First, besides damping the computational mode, the filter also weakly damps the physical mode, especially at high frequencies. This damping may become important for long integrations. Second, the RA filter degrades the accuracy of the unadulterated leapfrog scheme, since, by being uncentered in time, the RA-filtered leapfrog is only first-order accurate.

To ameliorate the negative effects that the RA filter has on the physical solution of the model, Williams (2009) introduced a modification that we refer to as the RAW filter. The original RA filter reduces, by a factor of $(1 - \nu)$, the magnitude of the temporal curvature of the state, and it is this smoothing effect that damps the computational mode. However, the filtering also changes the mean value of the state, averaged over the three time levels:

$$M = \frac{x_{n+1} + x_n + \overline{x}_{n-1}}{3} \neq \frac{x_{n+1} + \overline{x}_n + \overline{x}_{n-1}}{3}. \tag{3}$$

Williams (2009) showed that, when used with the leapfrog scheme, it is this nonmean-conserving feature of the filter that degrades the numerical accuracy. In the same work, the author tackled this problem by introducing an extra step in the filtering process, in order to include the possibility of conserving the mean value. The resulting

RAW filter is implemented in leapfrog integrations as follows:

$$\begin{aligned} x_{n+1} &= \overline{\overline{x}}_{n-1} + 2\Delta t F(\overline{x}_n) \\ \overline{\overline{x}}_n &= \overline{x}_n + \frac{\nu\alpha}{2}[x_{n+1} - 2\overline{x}_n + \overline{\overline{x}}_{n-1}] \\ \overline{x}_{n+1} &= x_{n+1} - \frac{\nu(1-\alpha)}{2}[x_{n+1} - 2\overline{x}_n + \overline{\overline{x}}_{n-1}]. \end{aligned} \quad (4)$$

The RAW filter introduces an extra operation that is simple and does not represent a considerable computational expense with respect to the RA filter. It also introduces a new parameter, $\alpha \in [0, 1]$. Taking an undamped oscillation equation $dF/dt = i\omega F$, Williams (2009) found the amplification relationship for the RAW filter to be

$$A_{\pm}(\alpha, \nu, \omega\Delta t) = \frac{\nu}{2} + \left[1 - \frac{\nu(1-\alpha)}{2}\right] i\omega\Delta t \pm \sqrt{\left(1 - \frac{\nu}{2}\right)^2 - \left[1 - \frac{\nu(1-\alpha)}{2}\right]^2 (\omega\Delta t)^2 + \nu\left(1 - \frac{\nu}{2}\right)(1-\alpha) i\omega\Delta t}. \quad (5)$$

In the amplification relation in Eq. (5), Δt corresponds to the time step of the numerical solution of the equation. Figure 1 (taken from Williams 2009) illustrates the behavior of Eq. (5)—for a fixed value of ν (taken to be 0.2)—with respect to $\omega\Delta t$. Each curve illustrates the effect of a different value of α on the numerical amplification (or numerical dissipation) of a free wave oscillation, which is physically unforced and undamped in the time-continuous differential equation. A value of $\alpha = 1$ corresponds to the traditional RA filter. From this figure we can see that, for a value of $\alpha = 0.53$ one can minimize the spurious, numerical impacts on the physical solution and obtain the closest match to the exact solution over a broad frequency range.

In Williams (2009), the RAW filter was tested in a simple linear system representing harmonic inertial oscillations. For this model, an explicit analytical solution exists and therefore it is easy to visualize and compare the effects of both the RA filter and the RAW filter in the numerical solution of the model. The purpose of the present work is to implement and test the RAW filter in a more realistic atmospheric model, which is described in the next section.

3. The SPEEDY model

In the present paper, we implement and test the RAW filter in a model that is more representative of those used in operational numerical weather predictions and climate simulations. In particular, we choose to use an AGCM known as SPEEDY (Molteni 2003). This model has a spectral primitive equation dynamic core and a set of simplified physical parameterization schemes. The model is chosen because it achieves computational efficiency while maintaining realistic simulations similar to those of state-of-the-art AGCMs with complex physics.

Miyoshi (2005) adapted SPEEDY for use in data assimilation, with output every 6 h. This implementation has a resolution of T30L7, with horizontal spectral truncation at 30 wavenumbers and 7 vertical levels. Data are output on a horizontal grid of 96 longitudinal and 48 latitudinal points. The model includes basic physical parameterizations, the description of which can be found in the appendix of Molteni (2003). The SPEEDY model is formulated in σ coordinates and calculates five field variables: zonal wind u, meridional wind v, temperature T, relative humidity q, and surface pressure ps. The geopotential height z for different pressure levels may be obtained by interpolation. We will also consider the variable precipitation, which the SPEEDY model is able to diagnose. These seven variables are used in the analysis of our results.

The SPEEDY model is based on a spectral dynamical core developed at the GFDL (Molteni 2003). It is a hydrostatic, σ-coordinate, spectral-transform model in the vorticity-divergence form described by Bourke (1974). The time stepping uses a leapfrog scheme, with the RA filter used to suppress the computational mode. In the integration, gravity waves are treated semi-implicitly. This last feature is important to note since some other schemes [such as the Adams–Bashforth third-order method described in Durran (1991)], which could otherwise be more efficient and accurate, become unstable under the semi-implicit scheme, and hence are not suited for this model. The leapfrog scheme does not present this problem. Moreover, an analysis of the favorable performance of the RAW filter under semi-implicit integrations can be found in Williams (2011).

For the RA filter, a value of $\nu = 0.1$ is applied to each prognostic variable. In previous works with the SPEEDY model (e.g., Miyoshi 2005), this value has been found to be adequate for this model, and it is in accordance with the suggested values described in the previous section. For the present study, we have implemented the RAW filter in the model, which required the addition of a single new line of code [corresponding to the change from Eq. (2) to Eq. (4)] and that did not noticeably affect the

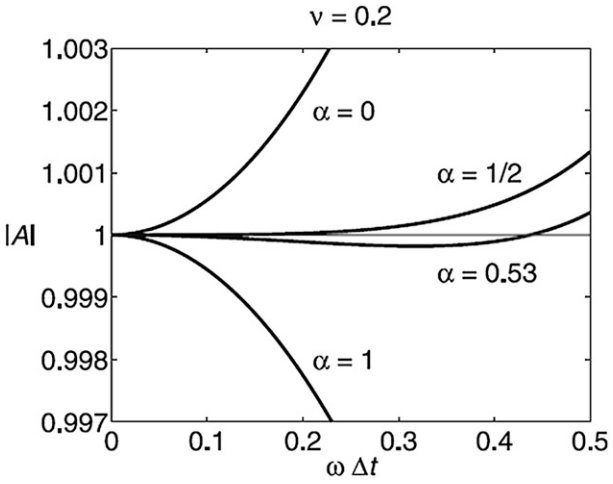

FIG. 1. The impacts of different values of the parameter α of the RAW filter on the numerical amplification of an unforced, undamped wave; taken from Williams (2009). The value of $\alpha = 1$ corresponds to the original RA filter. The value of $\alpha = 0.53$ is a preferred choice, because it keeps the amplification close to its exact value (unity) over a broad frequency range.

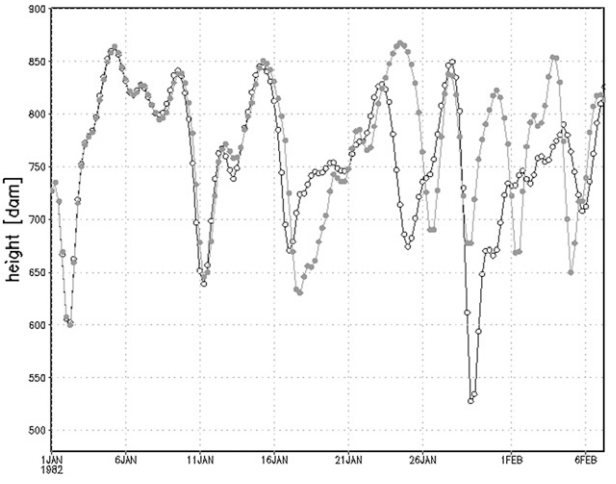

FIG. 2. The evolution of the 500-hPa geopotential height at 38°N, 75°W over 1 month. The line with open circles was obtained with the RA filter. The line with closed circles was obtained with the RAW filter. Each circle denotes a 6-h mean value. The initial conditions were identical in both integrations.

integration speed. We will compare and contrast a control integration, achieved using the original RA filter, with a second integration, achieved using the new RAW filter with $\nu = 0.1$ and $\alpha = 0.53$.

Given the chaotic nature of the SPEEDY model (or any AGCM), the change in the filtering scheme for the numerical integration will produce different solutions beyond a week or so, even when starting from the same initial conditions. Figure 2 illustrates this behavior for the 500-hPa geopotential height at a given location (Maryland at 38°N, 75°W). The temporal evolutions for this variable start to show visible changes after around 8 days, and by 17 days the solutions are completely different. This result agrees well with the limit of predictability for the atmosphere of two weeks estimated originally by Lorenz (1963). Because of this behavior, it is difficult to assess the effects of the RAW filter from single runs [as it was done in Williams (2009) with the simple oscillations model]; instead we will have to look at the statistics for multiple runs.

We seek to answer the following two questions in the remainder of the paper. First, does the new time integration scheme affect the model climatology? This question is of interest because any statistically significant changes in the climatology may require a retuning of the physical parameterizations. We note in passing that such changes to the simulations need not necessarily be improvements, even if a better numerical scheme is implemented, because of the possible problem of compensating numerical and physical biases. And, second, does the more accurate filter improve the short- and medium-term (1–6 day) forecasts of the model?

4. Effects of the RAW filter on the climatology of the SPEEDY model

Since we are interested in possible changes to the climatology of the model, in this section we would like to consider relatively long time averages for our variables. To strike a balance between retaining long averages and avoiding the effects of seasonality, we choose to focus on monthly averages. We will take the variables separately at each of the seven pressure levels. For example, we will consider the mean 510-hPa geopotential height for March, denoted $z_{510_{\text{mar}}}$, and the mean 200-hPa temperature for September, denoted $T_{200_{\text{sep}}}$. Surface pressure and precipitation are two-dimensional fields without vertical dependence. Taking into consideration the previous specifications, we will have 37 variables for each month of the year, giving 444 variables in total.

For each one of the 444 variables, we will look for differences between the climatology generated by the RAW filter and the climatology generated by the RA filter. Hence, we can write our null hypothesis as $\overline{x_{\text{month,RA}}} = \overline{x_{\text{month,RAW}}}$ and our alternative hypothesis as $\overline{x_{\text{month,RA}}} \neq \overline{x_{\text{month,RAW}}}$, where the second subscript indicates the time-stepping method by which the variable was generated.

To generate our climatology, we run the model for $N_{\text{years}} = 8$ yr, and for each filter scheme separately. For each year we compute the monthly means. Since the value of the temporal autocorrelation of the monthly

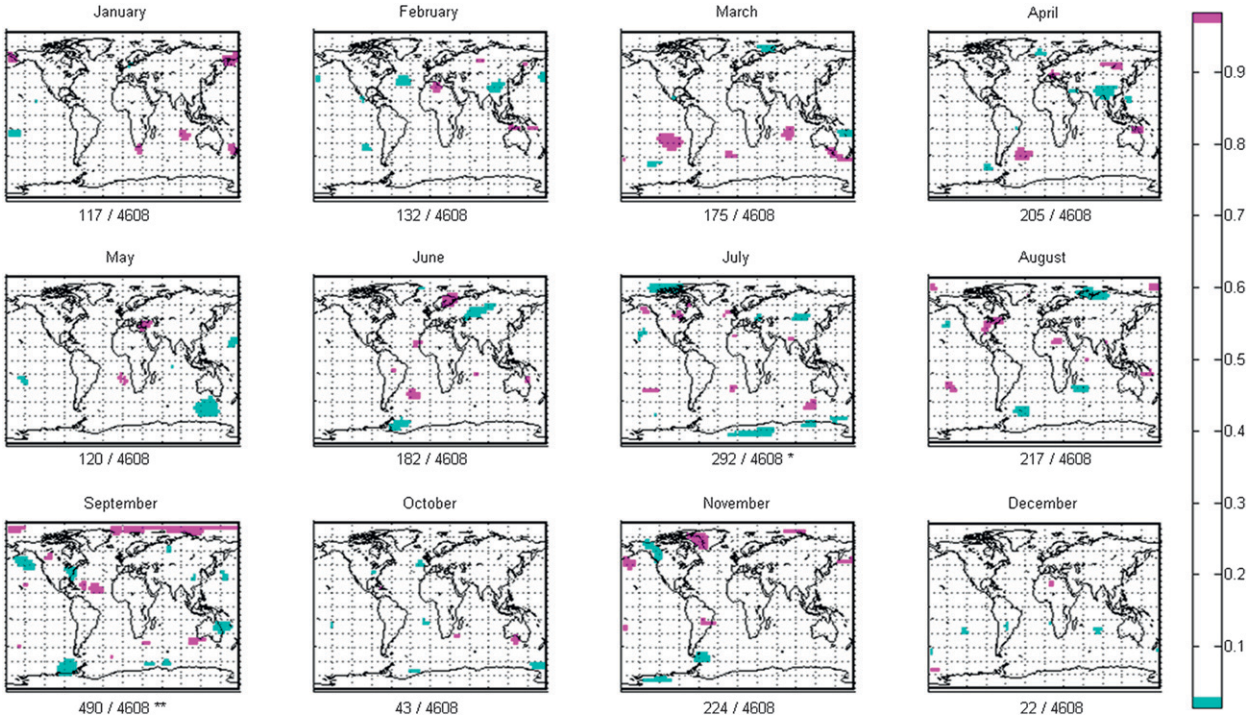

FIG. 3. Results of applying the t test for difference of means in the variables z_{510} for each month with a local significance $\alpha_{\text{local}} = 0.05$. Under each map we indicate the number of grid points that resulted locally significant out of the 96×48 grid. One asterisk denotes that the variable is field significant ($\alpha_{\text{field}} = 0.05$) considering finite sample size, and two asterisks denote that it is field significant considering both finite sample size and spatial correlation. Only the month of September is field significant.

means from one year to the next is very low, it is acceptable to neglect it when computing the statistics. If our variables were daily values instead of monthly averages, then we would surely need to consider this temporal autocorrelation and use a more suitable method, such as the moving blocks bootstrap proposed by Elmore et al. (2006).

a. *Local significance*

It is important to distinguish between local variations and field variations. For the former case, we test the null hypothesis for each variable at each vertical level and at each point on the 96 by 48 grid. The result for each grid point represents the *local significance* (Livezey and Chen 1983). For the latter case, the way in which we take into consideration the set of results for all the grid points of a variable determines the *field significance* (Livezey and Chen 1983), as described in section 4b.

To test the null hypothesis in the local context, we perform the Satterthwaite–Welch (SW) version of the t test. This test requires the data to come from normal distributions, allows small samples, and permits the two groups compared to have different variances. The test statistic is

$$t_{\text{month}} = \frac{\overline{x}_{\text{RA,month}} - \overline{x}_{\text{RAW,month}}}{\sqrt{\dfrac{s^2_{\text{RA,month}}}{N_{\text{years}}} + \dfrac{s^2_{\text{RAW,month}}}{N_{\text{years}}}}}. \tag{6}$$

In Eq. (6), $\overline{x}$ represents the interannual mean and s^2 represents the interannual variance. The statistic has a t distribution with f degrees of freedom, where f is calculated as indicated in Eq. (7). This expression for the "effective" number of degrees of freedom is the main difference of the SW t test from the standard t test:

$$f = \frac{\left(\dfrac{s^2_{\text{RA,month}}}{N_{\text{years}}} + \dfrac{s^2_{\text{RAW,month}}}{N_{\text{years}}}\right)^2}{\dfrac{\left(\dfrac{s^2_{\text{RA,month}}}{N_{\text{years}}}\right)^2}{N_{\text{years}} - 1} + \dfrac{\left(\dfrac{s^2_{\text{RAW,month}}}{N_{\text{years}}}\right)^2}{N_{\text{years}} - 1}}. \tag{7}$$

We perform the two-tailed version of the SW t test on all our variables, using a significance level of $\alpha_{\text{local}} = 0.05$. [Note our use of a subscript here, to distinguish this variable name from the unsubscripted α used in Eq. (4) for the RAW filter.] Figure 3 presents the results for the

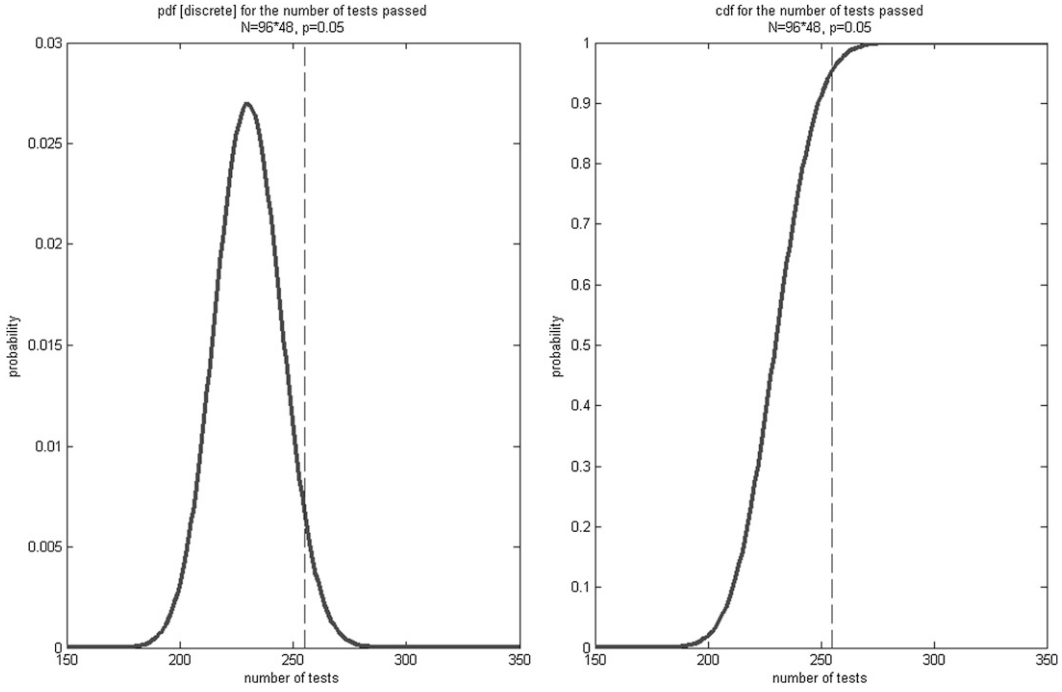

FIG. 4. (left) Probability mass function and (right) cumulative probability function for the binomial distribution representing the total number of local significance tests passed (assuming independence). For a total of 96×48 tests of local significance $\alpha_{\text{local}} = 0.05$, at least 255 must be passed in order to achieve a field significance $\alpha_{\text{field}} = 0.05$.

variable z_{510} (i.e., the 510-hPa geopotential height) for every month of the year. For the maps shown in this figure, we color in blue the points with $p \leq 0.025$ and in pink those with $p \geq 0.975$. Hence, the pink regions are grid points at which the climatology generated by the RAW filter has significantly smaller values than the climatology generated by the RA filter, while the blue regions are grid points at which the climatology generated by the RAW filter has significantly larger values than the climatology generated by the RA filter.

In Fig. 3 we see no preferred regions for the significant points, but they are instead scattered around the globe without coherency from one month to the next. This is true not only for this variable but for the others too (not shown). Moreover, since we are performing the same test in each grid point, some of the tests can be passed just by chance. This is called the ''multiplicity problem'' by Wilks (2005) and can lead to erroneous conclusions. One has to ask the following question (Livezey and Chen 1983): what is the minimum number of tests (out of the 96×48) that must be passed in order to achieve some desired field significance α_{field}?

b. Field significance

As indicated above, one must look at the results together in a ''field'' sense. To obtain this field significance, two effects must be taken into consideration (Livezey and Chen 1983). The first is finite sample size. We are performing the significance test at each of the $M = 96 \times 48$ grid points of the model. Each test may be regarded as a Bernoulli trial with a probability of success equal to the significance of the local t test, $\alpha_{\text{local}} = 0.05$. For the moment, let us assume that each of the M trials is independent from each of the others. Then we can regard the total number of tests passed as a random variable from a binomial distribution with a total of $M = 96 \times 48$ trials and an individual probability of success of $\alpha_{\text{local}} = 0.05$.

The mass probability function and the cumulative probability density function for this discrete binomial distribution are shown in Fig. 4. The distribution is centered on 5% of 96×48 tests (i.e., 230.4 tests). To have a field significance of $\alpha_{\text{local}} = 0.05$, the minimum number of tests that must be passed corresponds to the $100 \times (1 - \alpha_{\text{field}})$th percentile of this binomial distribution. Therefore, if we choose the field significance to be $\alpha_{\text{field}} = 0.05$, then at least $m_0 = 255$ tests must be passed.

Let us see how many variables fulfill the requirement to be field significant at the level $\alpha_{\text{field}} = 0.05$. For each of the $2 + 5 \times 7 = 37$ variables and each of the 12 months, Table 1 shows the number of points at which the t test was passed (i.e., the number of points that were locally significant). Considering the finite sample size effect, 119 out of the 444 variables are field significant. In the table,

TABLE 1. For each one of the variables, this table presents the number of grid points (out of 96 × 48) that resulted locally significant $\alpha_{\text{local}} = 0.05$ after applying the *t* test. The variables that result field significant with a value of $\alpha_{\text{field}} = 0.05$ are set bold if they are field significant considering only finite sample size (and independence), and they are also presented in italics (in addition to the boldface) if they are field significant considering both finite sample size and spatial correlation. The minimum number of locally significant points to achieve field significance is $m_0 = 255$ under independence and $m_0 \sim 390$ considering spatial correlation.

m_0			u						v						T						q						z					
255	***390***	ps	Prec	950	835	685	510	340	200	80	950	835	685	510	340	200	80	950	835	685	510	340	200	80	950	835	685	510	340	200	80	950	835	685	510	340	200	80
Jan		57	118	96	98	112	120	190	190	**305**	138	164	198	221	**268**	**307**	**351**	138	164	198	221	**268**	**307**	**351**	202	160	135	126	114	130	133	121	103	84	117	153	146	181
Feb		38	195	140	157	203	219	245	226	198	163	173	191	186	199	**261**	**310**	163	173	191	186	199	**261**	**310**	227	**257**	248	226	180	230	207	139	95	99	132	119	88	49
Mar		40	185	**269**	229	243	**265**	223	186	127	239	**256**	**274**	**256**	247	213	226	239	**256**	**274**	**256**	247	213	226	221	188	216	205	**292**	232	201	188	180	167	175	166	130	67
Apr		63	**263**	**306**	**319**	**294**	**268**	231	210	198	162	188	152	203	213	228	**290**	162	188	152	203	213	228	**290**	158	149	145	177	216	198	204	231	233	215	205	153	125	191
May		65	84	210	213	217	183	164	146	113	240	239	248	**276**	**280**	241	**264**	240	239	248	**276**	**280**	241	**264**	**300**	252	**274**	**339**	209	124	119	91	96	102	120	159	154	100
Jun		95	80	227	230	231	**263**	**256**	**267**	232	**264**	**258**	**290**	**301**	**293**	253	242	**264**	**258**	**290**	**301**	**293**	253	242	**288**	**285**	**315**	221	243	179	176	84	93	130	182	194	174	143
Jul		62	198	**350**	**348**	**349**	***403***	***396***	***400***	**342**	**255**	**259**	246	245	230	146	182	**255**	**259**	246	245	230	146	182	247	225	**277**	**255**	164	170	158	175	174	179	**292**	**308**	163	135
Aug		67	246	**328**	**316**	248	**264**	251	153	116	254	**262**	245	208	176	181	214	254	**262**	245	208	176	181	214	**327**	**289**	**318**	251	165	164	176	241	241	211	217	**380**	**348**	135
Sep		45	**349**	**346**	**383**	**376**	***393***	**337**	**337**	**297**	248	**270**	**278**	**301**	**276**	243	234	248	**270**	**278**	**301**	**276**	243	234	**308**	**310**	**323**	174	246	182	202	**361**	**376**	***430***	***490***	***425***	***408***	**304**
Oct		45	45	120	103	113	91	115	116	65	161	192	180	141	108	83	71	161	192	180	141	108	83	71	154	136	135	140	176	215	204	60	40	42	43	78	100	88
Nov		46	**261**	**276**	247	**255**	**266**	239	196	154	**305**	**345**	**314**	**294**	216	185	180	**305**	**345**	**314**	**294**	216	185	180	227	209	172	125	181	194	187	**269**	**280**	**282**	224	148	112	69
Dec		65	4	122	106	106	93	119	105	61	166	174	175	143	113	74	55	166	174	175	143	113	74	55	136	103	102	94	175	227	157	2	3	2	22	45	37	61

these variables are bolded. The month with the most field significant variables is September, with 26 out of 37 variables. There is apparently no preferred pressure level or variable for the field significance to appear.

Given only the above analysis, we would lean toward concluding that the RAW filter is indeed changing the climatology of the SPEEDY model for a considerable number of variables. However, a second effect must be taken into consideration: spatial correlation. When considering the total number of tests locally passed as a binomial distribution, we had to assume that the tests were independent from each other. That is, we considered that the result of a *t* test in a given grid point would not affect the result of the test in the surrounding grid points. We now improve this analysis by replacing the binomial distribution with a null empiric distribution in which the spatial correlation is embedded. A way to construct this distribution is Monte Carlo simulation. Elmore et al. (2006) describes how to generate the distribution by correlating random numbers with the data for each one of the variables for a number of trials. We selected this number of trials to be 1000.

Figure 5 shows the results of generating these empirical distributions for $z_{510_{\text{mar}}}$, $T_{200_{\text{sep}}}$, $v_{835_{\text{jul}}}$, and $u_{950_{\text{ago}}}$. (We generated the empirical distributions only for those variables that had resulted field significant.) One can immediately notice that these empirical distributions are substantially broader than the corresponding binomial distribution, having considerably heavier tails. For each variable, the shape of the distribution will be unique, since it contains the particular information of the spatial correlation for that variable. They are all, however, expected to present a qualitative similarity, since there is a common pattern of spatial interdependence for all the variables. For our purposes (evaluating the field significance of the individual *t* tests), we will be particularly interested in the upper tail of each of the distributions.

As one can see from Fig. 5—and as previously noted by Livezey and Chen (1983) and Elmore et al. (2006)—spatial correlation makes it more difficult to achieve the same level of field significance. The minimum number of local tests required to be passed is larger than with the binomial distribution. With the field significance level we had selected, $\alpha_{\text{field}} = 0.05$, the minimum number of tests that must be passed under independence is 255, whereas for the empirical distributions, the minimum numbers of tests are considerably larger. Table 2 shows these numbers for the four variables under consideration. Considering more than just the 4 variables in the table, the new minimum number of tests required to be passed ranges from around 380 to 530.

Let us conservatively consider one of the smallest of these numbers (e.g., 390) as our minimum number of

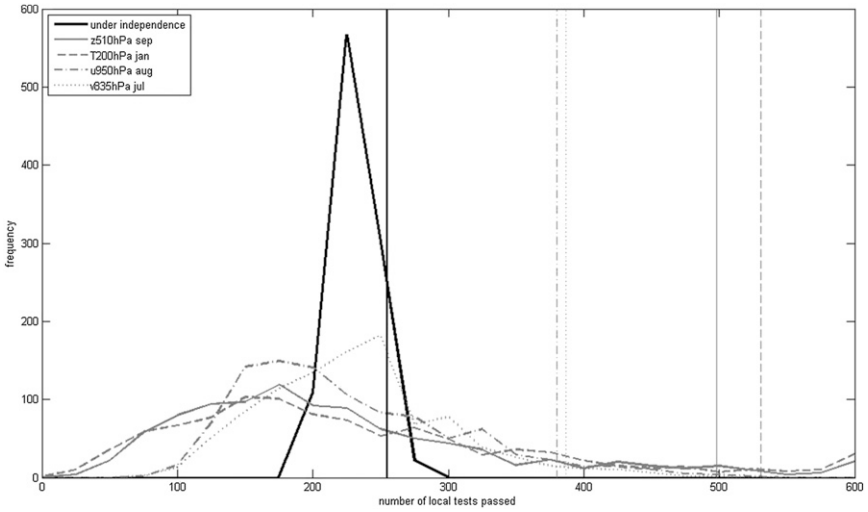

FIG. 5. Distribution of the total number of local significance tests passed. The binomial distribution (black line) corresponds to the assumption of field independence. The empirical distributions (gray lines), which consider the spatial correlation, are shown for four variables. These distributions were constructed via Monte Carlo simulation with 1000 iterations, as described in the text. The vertical lines indicate the 95th percentile for each distribution. It is noticeable that these values are substantially higher than the value of 255 (associated with the binomial distribution) appropriate for the spatially correlated variables.

local tests required to be passed in order to achieve the field significance $\alpha_{\text{field}} = 0.05$, and let us reconsider the results of Table 1. After considering the effect of spatial correlation, only 8 out of the 444 variables are field significant at $\alpha_{\text{field}} = 0.05$. That is, only 1.8% of the variables suffered a significant change. These variables are identified in Table 1 with italics (in addition to the boldface). This is clearly a huge reduction from the 119 field significant variables we had obtained under the assumption of independence.

Hence, considering spatial correlation, we conclude that there is no evidence to support the hypothesis that the climatology of the SPEEDY model generated by integrating with the RAW filter is different from that generated by integrating with the RA filter. This is an advantageous finding, in the sense that the new scheme does not require a retuning of the parameterized physics.

5. Effects of the RAW filter on the skill of short-term and medium-term forecasts

Since the climatology of the SPEEDY model is unchanged by the introduction of the new filter, we can now proceed to answer the question of accuracy: are solutions obtained with the RAW filter more accurate than solutions obtained with the RA filter?

To assess any possible accuracy improvement, we use the anomaly correlation coefficient (ACC) for h-hour forecasts. The ACC is a measure of the agreement between the spatial variations in the forecast and the analysis, each with respect to the climatology. The ACC is calculated using

$$\text{ACC} = \frac{\sum_{i=1}^{N}[(f_i - \text{cs}_i)(a_i - \text{cr}_i)\cos\varphi_i]}{\sqrt{\sum_{i=1}^{N}[(f_i - \text{cs}_i)^2\cos\varphi_i]\sum_{i=1}^{N}[(a_i - \text{cr}_i)^2\cos\varphi_i]}}, \tag{8}$$

where f_i is the forecast, a_i is the analysis, cr_i is the climatology of the reanalysis, cs_i is the climatology of the SPEEDY model, φ_i is the latitude, and N is the total number of grid points for the variable. Note that we use the SPEEDY model's own climatology rather than the

TABLE 2. Minimum number of tests (out of 96×48) to be passed with a local significance $\alpha_{\text{local}} = 0.05$ to achieve a field significance $\alpha_{\text{field}} = 0.05$.

Variable	Min no. of locally significant points
Any variable under spatial independence	255
$z_{510_{\text{mar}}}$	498
$T_{200_{\text{sep}}}$	531
$v_{835_{\text{jul}}}$	387
$u_{950_{\text{aug}}}$	380

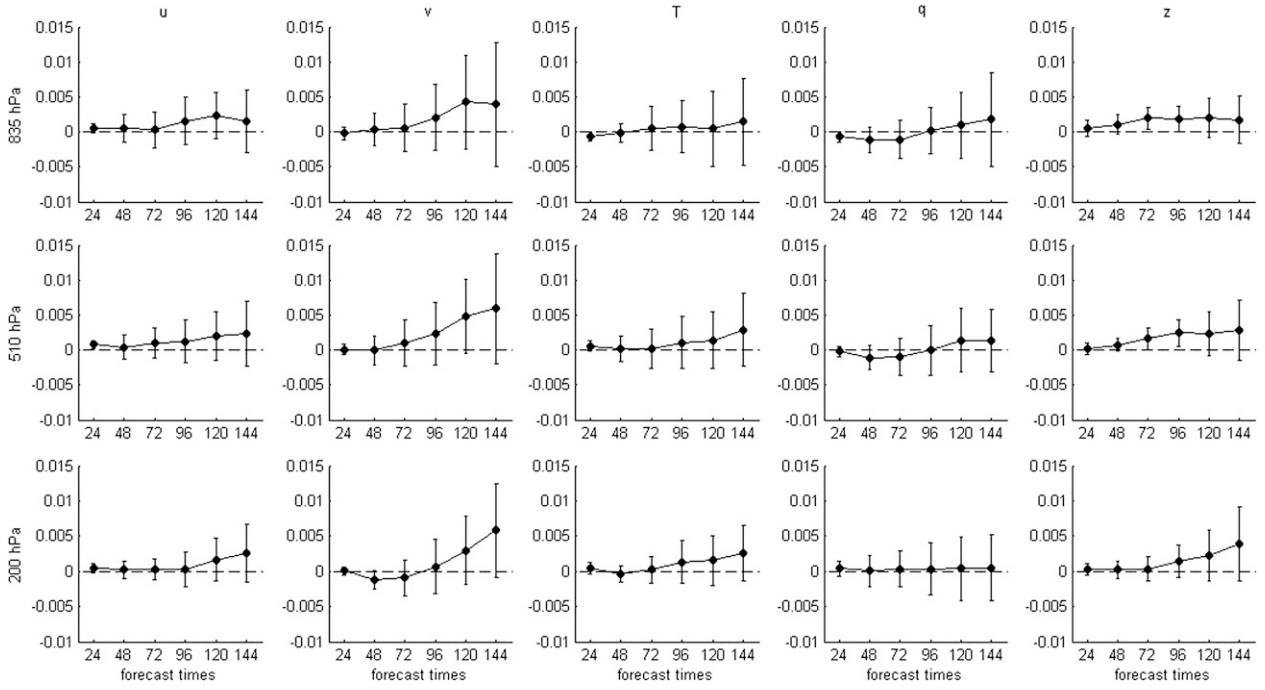

FIG. 6. Increase in anomaly correlation coefficient ($\overline{\mathrm{ACC}_{\mathrm{RAW}}} - \overline{\mathrm{ACC}_{\mathrm{RA}}}$) for six different forecast times. The values were computed globally for three different pressure levels and for each of the five variables. The most benefited variables are the meridional wind and the geopotential height. The bars denote one standard deviation of the difference.

reanalysis climatology to define forecast anomalies, because the SPEEDY model has resolution much lower than operational forecast models, and hence larger climatological errors. The subscript i labels the points on the grid.

We perform the ACC computation for the month of January 1982. For the analysis data, we use the NCEP–NCAR reanalysis dataset (Kalnay et al. 1996) interpolated onto the SPEEDY grid. The climatology of SPEEDY is computed from the 8-yr runs for the RA filter and the RAW filter. Following the conclusion from section 3, we compute the climatology as follows:

$$\overline{x_{\mathrm{JAN}}} = (\overline{x_{\mathrm{JAN,RA}}} + \overline{x_{\mathrm{JAN,RAW}}})/2. \qquad (9)$$

We select 3 of the 7 vertical levels of the model, representing roughly the upper atmosphere (200 hPa), the middle atmosphere (510 hPa), and the lower atmosphere (835 hPa). The ACC analysis is performed for the model variables (u, v, T, q, z) in each of the above levels, and it is also computed for the surface variable ps.

The ACC analysis is first performed globally. The results for the five variables (excluding ps) are presented in Fig. 6, which displays the differences $\overline{\mathrm{ACC}_{\mathrm{RAW}}} - \overline{\mathrm{ACC}_{\mathrm{RA}}}$. There is a clear, general improvement due to the use of the RAW filter, and the improvements are around $O(10^{-3})$ in magnitude. The improvement increases with lead time and is more important for medium-term forecasts with lead times of 96, 120, and 144 h. The variables that benefit most from the RAW filter are z and v, while q is the only variable that has no apparent improvement. There are almost no cases where the difference $\overline{\mathrm{ACC}_{\mathrm{RAW}}} - \overline{\mathrm{ACC}_{\mathrm{RA}}}$ is negative.

To examine regional differences, we finally perform the ACC analysis for three latitudinal bands: the tropics (25°S–25°N), the Northern Hemisphere midlatitudes (25°–75°N), and the Southern Hemisphere midlatitudes (75°–25°S). Figure 7 shows the results for the two variables that were globally most benefited by the RAW filter: the geopotential height and meridional wind.

For the geopotential height z, the largest improvements in the ACC occur in the tropics. Moreover, the improvements start to be noticeable in the 72-h forecast, which is earlier than for the other variables. The difference $\overline{\mathrm{ACC}_{\mathrm{RAW}}} - \overline{\mathrm{ACC}_{\mathrm{RA}}}$, which is of the order of +0.02 for medium-range forecasts, is larger for this variable and region than for any other. Results are similar for the surface pressure ps (not shown). These improvements in the skill of medium-range forecasts, which arise directly from the upgrade to the RAW filter, increase the anomaly correlation coefficient for surface pressure (and 500-hPa geopotential height) in the tropics by 10%–20%, as seen in Fig. 8. As a consequence, 5-day forecasts made using the RAW filter have approximately the same skill

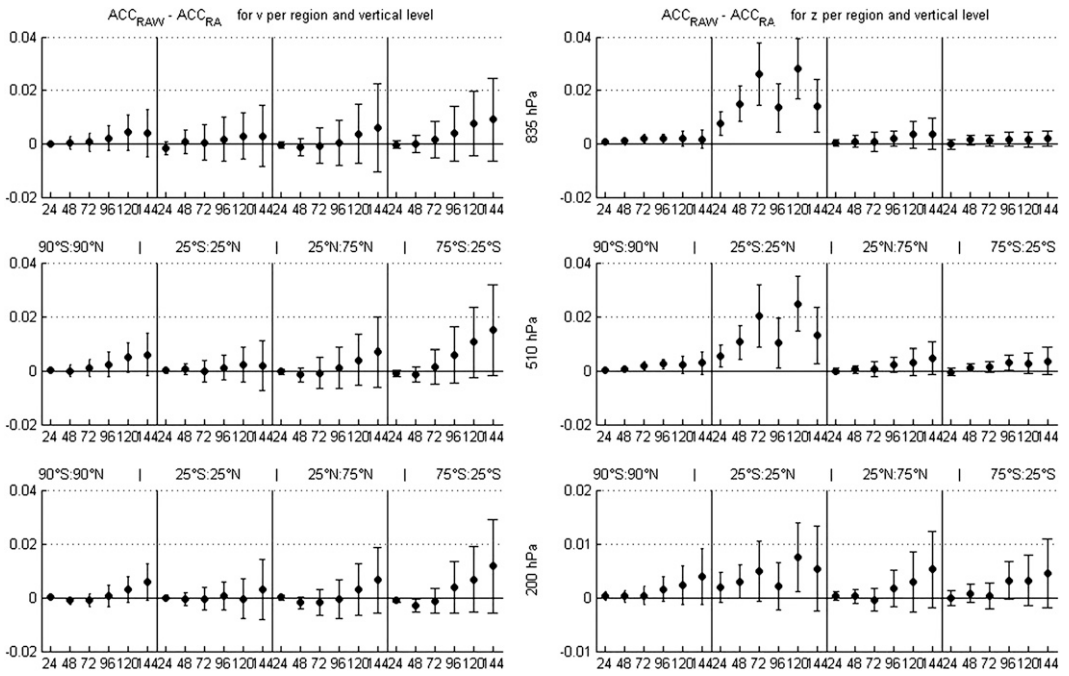

FIG. 7. Increase in anomaly correlation coefficient ($\overline{\mathrm{ACC}_{\mathrm{RAW}}} - \overline{\mathrm{ACC}_{\mathrm{RA}}}$) for six different forecast times for two variables (geopotential height and meridional wind) at three pressure levels and four different latitudinal bands. The bars denote one standard deviation of the difference.

as 4-day forecasts made using the RA filter, and 4-day forecasts made using the RAW filter have approximately the same skill as 3-day forecasts made using the RA filter.

For the meridional wind v, the largest improvements in the ACC occur outside the tropics, in the three levels of the atmosphere, and they are more noticeable as the forecast time increases. The improvements in the temperature T (not shown) are very similar to those for v, with the largest values occurring in the Northern Hemisphere and especially in the middle atmosphere. For the zonal wind u (not shown), there is a moderate improvement for the medium-term forecasts, but it is not as striking as for the previously listed variables, and the improvement never exceeds 0.005. For the relative humidity q (not shown), we consistently get an improvement close to zero.

To complement the ACC analysis, an additional RMSE analysis is performed. This statistic does not involve the climatology; instead it compares directly the forecast (generated by integrating with any of the two filters) with the reanalysis data. We calculated a latitude-weighted RMSE as expressed in Eq. (10):

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f_i - a_i)^2 \cos\varphi_i}. \qquad (10)$$

In this expression f_i is the forecast, a_i is the analysis, φ_i is the latitude, and N is the total number of grid points for the variable. We computed the difference $\overline{\mathrm{RMSE}_{\mathrm{RAW}}} - \overline{\mathrm{RMSE}_{\mathrm{RA}}}$ for each forecast time and for each variable. This difference should be negative for the cases in which the RAW filter is improving the accuracy of the forecasts. This experiment yielded results similar to the ACC analysis; the figures generated are not shown. For

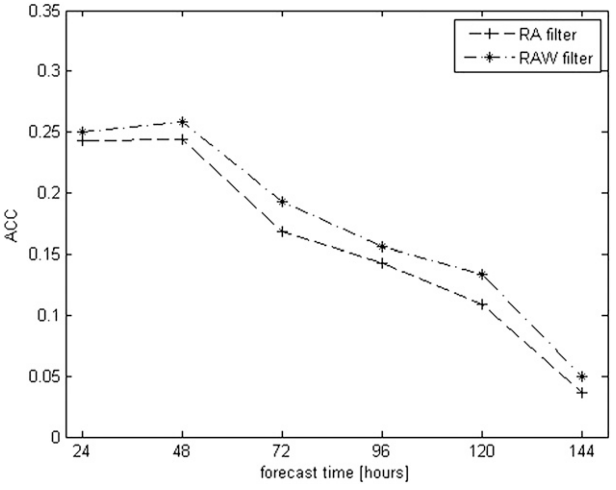

FIG. 8. ACC for forecasts of surface pressure in the tropics. Notice that 96-h forecasts using the RAW filter have approximately the same skill as 72-h forecasts using the RA filter. Also, 120-h forecasts using the RAW filter have approximately the same skill as 96-h forecasts using the RA filter.

the majority of the variables we observed a reduction in the RMSE, particularly for medium-term forecasts. In the RMSE, however, it is more difficult to asses the relative impact of the filter among the different variables, since for each of the variables we have different units, while the ACC is nondimensional.

6. Summary and discussion

The first question asked in the present paper is: are there any statistically significant changes in the monthly climatology of the SPEEDY model caused by the upgrade in the numerical integration scheme from Robert–Asselin (RA) filter to Robert–Asselin–Williams (RAW) filter? To answer this question, we performed a Satterthwaite–Welch t test for the difference of means for each variable, in order to assess local significance at the 5% level. At some grid points the tests were passed, but these points appeared to be scattered around the globe and showed no particular preference for location. In field significance tests, after considering the effects of both finite sample size and spatial correlation, we found that there is no significant evidence to reject the null hypothesis of identical climatologies. In other words, for each month, the climatology generated by integrating with the RA filter is the same as the one obtained with the RAW filter. Hence, the RAW filter is suitable for use in the SPEEDY model.

The second question asked is: is there a statistically significant improvement in the skill of short- to medium-term (24–144 h) forecasts caused by the upgrade from RA filter to RAW filter? To answer this question, an ACC analysis was performed for 24-, 48-, 72-, 96-, 120-, and 144-h forecasts for the month of January 1982. As analysis data we used the NCEP–NCAR reanalysis dataset interpolated onto the SPEEDY grid. The model climatology was generated by 8-yr integrations of SPEEDY. The ACC analysis was performed on three pressure levels (835, 510, and 200 hPa), both globally and by latitude. A complementary RMSE analysis was performed following the same scheme, and yielding the same conclusions as the ACC analysis.

In general, an improvement of order $O(10^{-3})$ in the ACC can be attributed to the use of the RAW filter, and the improvement is larger for medium-term forecasts with lead times of 72, 120, and 144 h. The geopotential height was strongly benefited in the tropics, with ACC increases as large as 0.02 for a 72-h forecast and 0.025 for a 120-h forecast. As a consequence, 5-day forecasts made using the RAW filter have approximately the same skill as 4-day forecasts made using the RA filter, and 4-day forecasts made using the RAW filter have approximately the same skill as 3-day forecasts made using the RA filter. The meridional wind was strongly benefited in the extratropics. The improvements in surface pressure mimicked those in geopotential height, and the impacts on temperature were very similar to those on meridional velocity. The improvements for the zonal velocity were less noticeable and there were no significant improvements in the relative humidity.

The results of this work are encouraging for the use of the RAW filter in the numerical solution of models based on the widely used RA filter. More generally, we have found that the skill of medium-range weather forecasts is sensitive to the time-stepping method, about as much as could be expected from the use of different physics parameterizations to improve forecast skill. We suggest that, in future work, numerical time schemes be revisited as a potentially important component of model error.

Acknowledgments. The authors thank Dr. Hong Li for providing the NCEP–NCAR reanalysis data already interpolated onto the grid of the SPEEDY model, as well as Dr. Sajal Kar and an anonymous reviewer for their constructive criticism that helped enhance this work. PDW acknowledges funding through a University Research Fellowship from the Royal Society (Reference UF080256). The support of NASA Grants NNX07AM97G and NNX08AD40G, and DOE Grant DEFG0207ER64437 is also gratefully acknowledged.

REFERENCES

Asselin, R., 1972: Frequency filter for time integrations. *Mon. Wea. Rev.,* **100,** 487–490.

Bourke, W., 1974: A multilevel spectral model. I. Formulation and hemispheric integrations. *Mon. Wea. Rev.,* **102,** 687–701.

Déqué, M., and D. Cariolle, 1986: Some destabilizing properties of the Asselin time filter. *Mon. Wea. Rev.,* **114,** 880–884.

Durran, D. R., 1991: The third-order Adams–Bashforth method: An attractive alternative to leapfrog time differencing. *Mon. Wea. Rev.,* **119,** 702–720.

Elmore, K., and Coauthors, 2006: Field significance revisited: Spatial bias errors in forecasts as applied to the eta model. *Mon. Wea. Rev.,* **134,** 519–531.

Kalnay, E., 2003: *Atmospheric Modelling, Data Assimilation, and Predictability*. Cambridge University Press, 364 pp.

——, and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Kantha, L., and C. Clayson, 2000: *Numerical Models of Oceans and Oceanic Processes*. International Geophysical Series, Vol. 66, Academic Press, 750 pp.

Livezey, R., and W. Chen, 1983: Statistical field significance and its determination by Monte Carlo techniques. *Mon. Wea. Rev.,* **111,** 46–59.

Lorenz, E., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.,* **21,** 130–141.

Miyoshi, T., 2005: Ensemble Kalman filter experiments with a primitive-equation global model. Ph.D. thesis, University of Maryland, College Park, College Park, MD, 226 pp.

Molteni, F., 2003: Atmospheric simulations using a GCM with simplified physical parametrizations. I: Model climatology and variability in multi-decadal experiments. *Climate Dyn.,* **20,** 175–191.

Pfeffer, R. L., I. M. Navon, and X. Zou, 1992: A comparison of the impact of two time-differencing schemes on the NASA–GLAS climate model. *Mon. Wea. Rev.,* **120,** 1381–1393.

Robert, A. J., 1966: The integration of a low order spectral form of the primitive meteorological equations. *J. Meteor. Soc. Japan,* **44,** 237–245.

Wilks, D. S., 2005: *Statistical Methods in the Atmospheric Sciences*. 2nd ed. Academic Press, 648 pp.

Williams, P. D., 2009: A proposed modification to the Robert–Asselin time filter. *Mon. Wea. Rev.,* **137,** 2538–2546.

——, 2011: The RAW filter: An improvement to the Robert–Asselin filter in semi-implicit integrations. *Mon. Wea. Rev.,* in press.

Williamson, D. L., 1983: Description of NCAR Community Climate Model (CCMOB). NCAR Tech. Note NCAR/TN-210+STR, 88 pp.

——, and J. G. Olson, 2003: Dependence of aqua-planet simulations on time step. *Quart. J. Roy. Meteor. Soc.,* **129,** 2049–2064.

Zhao, B., and Q. Zhong, 2009: The dynamical and climate tests of an atmospheric general circulation model using the second-order Adams–Bashforth method. *Acta Meteor. Sin.,* **23,** 738–749.