atmosphere

MDPI

Article

On Applying Large-Scale Correction to Limited-Area Numerical Weather Prediction Models

Anurag Dipankar [1,2,*], Xiang-Yu Huang [1,3] and Peter Heng [1]

[1] Centre for Climate Research Singapore, Meteorological Service Singapore, Singapore 537054, Singapore; xyhuang@ium.cn (X.-Y.H.); peter_heng@nea.gov.sg (P.H.)
[2] Institute for Atmospheric and Climate Sciences, ETH Zürich, 8006 Zürich, Switzerland
[3] Institute of Urban Meteorology, CMA, Beijing 100089, China
* Correspondence: anurag.dipankar@c2sm.ethz.ch

Abstract: This paper presents a new blending approach to applying large-scale correction to the initial condition in a limited-area numerical weather prediction (NWP) model. The new approach combines the implementation benefits of the known approaches and shows significant improvement in the forecast quality when implemented in a tropical NWP model. Sensitivity studies indicate that many improvements come from blending the horizontal winds alone. Adding temperature and specific humidity to the horizontal winds result in forecast quality degradation in the early hours of the simulated tropical environment.

Keywords: numerical weather prediction; data assimilation; tropical weather

Citation: Dipankar, A.; Huang, X.-Y.; Heng, P. On Applying Large-Scale Correction to Limited-Area Numerical Weather Prediction Models. *Atmosphere* **2022**, *13*, 1142. https://doi.org/10.3390/atmos13071142

Academic Editor: Eric A. Hendricks

Received: 13 June 2022
Accepted: 13 July 2022
Published: 18 July 2022

1. Introduction

Limited-area modeling (LAM) of the atmosphere is subject to biases due to the imperfect treatment of the lateral boundaries [1] that provide large-scale information to the LAM. These biases are more apparent for longer time integrations, causing the LAM to drift away from the driving model as in the regional climate simulations [2]. Limited-area numerical weather prediction models also face the same issue, but the biases are assumed to be small because of the shorter integration length. However, imperfect large-scale information affects limited-area NWP models during their initialization step, which is the focus of the current study.

Regional or limited-area NWP models reinforce large-scale information to a greater degree by continuously assimilating observations via data assimilation (DA) techniques within the computational domain, but it is not complete. Being limited in area, these NWP models lack observations from outside the computational domain, making the large-scale representation in the analysis questionable [3]. Misrepresentation of large-scale information in the analysis, together with imperfect treatment of the lateral boundaries, is often found to result in poorer forecast accuracy compared with the global model forecasts [4,5].

NWP systems often perform partial cycling by restarting the DA cycle from global analysis (e.g., [6]) to overcome the issue of large-scale misrepresentation. Restarting from global analysis results in the loss of small-scale features generated by the fine-grid LAM and therefore requires additional simulation time for spin-up. One solution is to combine the large-scale information in global analysis with the small-scale features from the LAM. The early approaches in this direction, known as digital filter-blending, applied a digital filter [7] on both the LAM guess and global analysis on a low-resolution grid and applied the differences to the LAM guess [8]. Recently, Ref. [9] proposed a warm start method to reduce the initial spin-up in a cycling regional model, where they used the global model forecast instead of the analysis to obtain the correct large-scale information.

Another approach is to add an extra term to the cost function of the variational data assimilation to measure the departure of the LAM analysis from the global model analysis [10,11].

Except for [9], all other approaches rely on the global analysis for the implied large-scale correction, which we find restrictive for possible delays in the arrival of the global analysis in the absence of a good network. Moreover, the technique employed in [10,11] required changes in the model, which we wanted to avoid. Our development and the method proposed in [9], which fulfills our requirements, started around the same time but with different aims. Ref. [9] focused on LAM without DA, whereas our aim is to retain the LAM DA.

This approach is commonly known as blending. It uses a scale-selective filter in the physical space to replace the questionable large-scale information in the LAM analysis with presumably correct large-scale information from the global driving model. There are two known ways to implement blending in the DA framework: (a) background blending, wherein the large-scale information from the driving model forecast is blended into the forecast from the LAM to yield an updated background for the DA cycle [5,12]; and (b) analysis blending, wherein the large-scale information from the driving model analysis is blended into the LAM analysis [13,14].

The advantage of background blending over analysis blending is that it can be applied to all the cycles as long as the driving model forecast (instead of analysis) is available, making it more suitable for operational use. For this reason, background blending has become more popular than analysis blending. Recent work includes [15], proposing a dynamic method to select the blending scale (i.e., the scale at which the forecast from the driving model and LAM gradually blend). The Met Office in the U.K. also has plans to implement background blending in their regional NWP model [16]. The advantage of analysis blending is that it is easier to implement, and it does not interact with the DA system directly.

The current study had two aims, the first of which was to propose an alternate blending technique that inherits the benefits of both existing techniques, that is, a technique that is easier to implement while relying on driving model forecasts instead of analysis. The second aim was to perform a systematic study to highlight the sensitivity of blending to the blended variables, which was missing in the previous studies.

The following section presents a description of the different blending approaches, followed by an overview of the simulation details in Section 3. The impact of blending is qualitatively and quantitatively discussed in Section 4. The manuscript ends with conclusions in Section 5.

2. Blending Approaches

The primary objective of any blending approach is to correct the model's large-scale biases to improve the forecast quality. Here, we describe the known blending approaches within the framework of variational data assimilation (VAR) [5] using the notations in [17]. A full list of notations used is listed in Table 1 for quick reference.

For any state vector $\mathbf{x}$ (bold indicates a vector quantity) in the LAM, its blended equivalent is defined as

$$\widetilde{\mathbf{x}} = \mathbf{x} + \overline{(\mathbf{x_G} - \mathbf{x})} \tag{1}$$

Here, overbar (-) represents the low-pass filter, subscript $\boldsymbol{G}$ ($_G$) represents the driving global model variable, and tilde (˜) represents the blended variable. Using $\boldsymbol{\delta x^a}$ as the analysis increment from LAM DA, the analysis $\mathbf{x^a}$ can be written as

$$\mathbf{x^a} = \mathbf{x^b} + \boldsymbol{\delta} \mathbf{x^a} \tag{2}$$

where $\mathbf{x}^{\mathrm{b}}$ is the background state. When large-scale correction is applied through the blending technique, the aim is to obtain a new analysis $\mathbf{x^a_{LC}}$ that is "large-scale-corrected".

In the following section, we first describe the existing blending techniques and propose our new approach at the end.

Table 1. List of variables with their notations.

Notations	Meaning	Notations	Meaning
$\mathbf{x}$	State vector	$\delta\mathbf{x}$	Analysis increment
$\overline{\mathbf{x}}$	Filtered (low-pass) state vector	$\mathbf{x^a_{LC}}$	Large-scale corrected analysis
$\mathbf{x_G}$	Driving model state vector	$\Delta\mathbf{x_{LC}}$	Large-scale correction
$\tilde{\mathbf{x}}$	Blended state vector	$\mathbf{K}$	Kalman gain
$\mathbf{x^a}$	Analysis vector	$\mathbf{y}^o$	Observations
$\mathbf{x^b}$	Background state vector	H	Observation operator
Δ	Grid size	$\mathbf{B}$	Background error covariance matrix
L_c	Cut-off length scale for low-pass filter	ε	Filtering parameter (see Equation (13))

2.1. Analysis Blending

As depicted in Figure 1, analysis blending is applied after the DA step and requires global analysis. The resultant analysis is a blend of large-scale information from global analysis and small-scale information from LAM analysis, i.e.,

$$\mathbf{x^a_{LC}} = \tilde{\mathbf{x}}^{\mathbf{a}} = \mathbf{x^a} + \overline{\mathbf{x^a_G} - \mathbf{x^a}} = \mathbf{x^b} + \overline{\mathbf{x^b_G} - \mathbf{x^b}} + \delta\mathbf{x^a} + \overline{\delta\mathbf{x^a_G} - \delta\mathbf{x^a}} \quad (3)$$

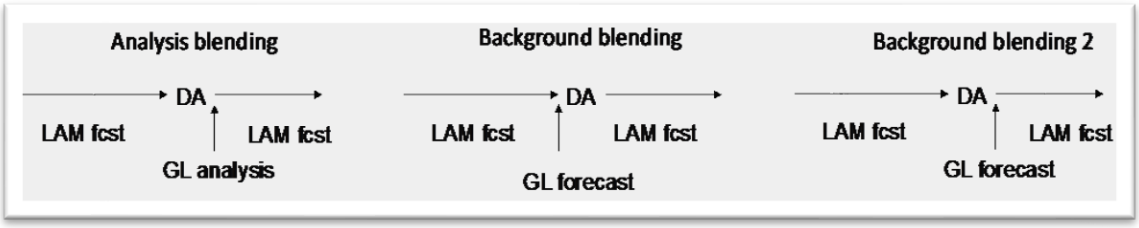

Figure 1. Schematic showing different blending approaches. From left: Analysis blending corrects the large-scale information in the LAM analysis using global (GL) analysis; background blending corrects the large-scale information in the LAM background using global forecast (fcst), which is then used in the LAM data assimilation (DA); background blending 2 corrects the large-scale information in the LAM background as in background blending but after DA.

Another way to interpret this change to the old analysis ($\mathbf{x^a}$) is to view it as an equivalent large-scale correction ($\Delta\mathbf{x_{LC}}$) of the following form

$$\Delta\mathbf{x_{LC}} = \mathbf{x^a_{LC}} - \mathbf{x^a} = \overline{\mathbf{x^b_G} - \mathbf{x^b}} + \overline{\delta\mathbf{x^a_G} - \delta\mathbf{x^a}} \quad (4)$$

Equation (4) implies that the large-scale correction is applied to both the background and the analysis increment. This approach has a clear advantage in using the best estimate of large-scale correction from the global analysis. Such an approach makes sense in NWP centers running both global and regional models because the delay in receiving global analysis is minimal, but not for other centers.

2.2. Background Blending

Realizing the practical limitations of analysis blending, [5] proposed blending recent global model forecasts with the LAM forecasts to create a large-scale-corrected background.

This approach requires the blending to be implemented before the DA step (see Figure 1), modifying the background state as

$$\tilde{\mathbf{x}}^{\mathbf{b}} = \mathbf{x}^{\mathbf{b}} + \overline{\mathbf{x_G} - \mathbf{x}^{\mathbf{b}}} \tag{5}$$

which is then used to evaluate the analysis increment

$$\hat{\delta\mathbf{x}} = \mathbf{K}\left[\mathbf{y}^{o} - H\left(\tilde{\mathbf{x}}^{\mathbf{b}}\right)\right] \tag{6}$$

where $\mathbf{K}$ is the Kalman gain, and $\mathbf{y}^{o}$ indicates the observations. Note the use of blended background with the observation operator H. Circumflex () is used with the increment to indicate that the use of $\tilde{\mathbf{x}}^{\mathbf{b}}$ in Equation (6) has appropriately modified the original increment. Finally, the large-scale-corrected analysis can be written as

$$\mathbf{x}^{\mathbf{a}}_{\mathbf{LC}} = \mathbf{x}^{\mathbf{b}} + \overline{\mathbf{x_G} - \mathbf{x}^{\mathbf{b}}} + \hat{\delta\mathbf{x}} \tag{7}$$

giving the equivalent large-scale correction term as:

$$\Delta\mathbf{x_{LC}} = \overline{\mathbf{x_G} - \mathbf{x}^{\mathbf{b}}} + \mathbf{K}\left[H\left(\overline{\mathbf{x}^{\mathbf{b}} - \mathbf{x_G}}\right)\right] \tag{8}$$

Like analysis blending, the large-scale correction in this approach also affects the background and the analysis increment. Note that unlike analysis blending, where the LAM analysis increments are directly corrected through the use of the global analysis increment, the analysis increment in background blending is indirectly corrected through the use of a corrected background that is then fed into the DA algorithm.

While this approach has a clear advantage for real-time use, it suffers from drawbacks. First, it experiences complications due to its dependence on the DA technique. For 3DVAR, Equations (5) and (6) need to be executed only once; for the advanced techniques such as 3DVAR first guess at appropriate time (FGAT), these equations would have to be executed several times within the analysis window. If performed correctly, global forecasts at the matching observation time are needed for Equation (6), which adds a burden in terms of data transfer and creates further complications due to the need to ensure that the global forecasts come from the same cycle; otherwise, Equation (6) may feel an unwanted jump. The second concern is the suitability of Equation (6), which assumes the background state $\mathbf{x}^{\mathbf{b}}$ is from the LAM that was used to derive the background error covariance matrix $\mathbf{B}$ implied in $\mathbf{K}$. The use of a blended background $\tilde{\mathbf{x}}^{\mathbf{b}}$ in Equation (5) therefore makes its suitability questionable.

Despite the inconsistent $\mathbf{x}^{\mathbf{b}}$ and $\mathbf{K}$, background blending has been demonstrated to significantly improve the forecast quality in different parts of the world [5,12,15,16]. This suggests that the errors due to incorrect large-scale information in LAMs are perhaps more significant than the errors due to the aforementioned inconsistency.

2.3. Forecast Blending

The approach adopted in the present study is meant to resolve the inconsistency between $\mathbf{x}^{\mathbf{b}}$ and $\mathbf{K}$ in the background blending approach. It uses the global forecast at the nominal analysis time to estimate large-scale correction by comparing it with the consistent LAM background (see Figure 1). This large-scale correction is then appropriately added to the analysis increment. That is, the large-scale-corrected analysis in this case takes the following form

$$\mathbf{x}^{\mathbf{a}}_{\mathbf{LC}} = \mathbf{x}^{\mathbf{b}} + \delta\mathbf{x}^{\mathbf{a}} + \Delta\mathbf{x_{LC}} \tag{9}$$

using

$$\Delta\mathbf{x_{LC}} = \overline{\mathbf{x_G} - \mathbf{x}^{\mathbf{b}}} \tag{10}$$

Equation (10) is essentially the first term on the right-hand side (RHS) of Equation (8). Adding Equations (9) and (10) results in

$$\mathbf{x}^{\mathbf{a}}_{\mathbf{LC}} = \mathbf{x}^{\mathbf{b}} + \overline{\mathbf{x}_{\mathbf{G}} - \mathbf{x}^{\mathbf{b}}} + \delta\mathbf{x}^{\mathbf{a}} \tag{11}$$

which can be equivalently written as the sum of a blended background and analysis increment.

$$\mathbf{x}^{\mathbf{a}}_{\mathbf{LC}} = \tilde{\mathbf{x}}^{\mathbf{b}} + \delta\mathbf{x}^{\mathbf{a}} \tag{12}$$

Equation (12) implies that the presented approach (indirectly) corrects only the background, leaving the analysis increment untouched. The fact that the DA algorithm does not use the blended background to estimate the new increment helps to resolve the issue of inconsistency in the background blending approach. It also makes the implementation easy, as only one global forecast corresponding to the nominal analysis time is needed instead of several, as in the background blending approach.

However, it does suffer from a potential double-counting of large-scale information: one from the large-scale increment (the second term on the RHS of Equation (11)) and another from the LAM analysis increment (the last term on the RHS of Equation (11)). This double-counting is more apparent if the global analysis is used instead of the forecast in Equation (11), with a perfect LAM DA that can incorporate the information of observations outside the LAM domain. In this case, the large-scale analysis increment, i.e., $\overline{\delta\mathbf{x}^{\mathbf{a}}}$, is likely to have an overlap with the contributions due to blending the increment $\left(\overline{\mathbf{x}_{\mathbf{G}} - \mathbf{x}^{\mathbf{b}}}\right)$. However, because the current approach uses a global forecast instead of analysis and perfect LAM DA is not possible, we think that the effect of this double-counting, if any, would be negligible.

3. Simulation Details

The model used in this study was SINGV [18,19], which is based on the Met Office's Unified Model [20]. It uses 3DVAR FGAT and is cycled every 3 h. The analysis increments were added through incremental analysis update (IAU). The simulation domain, centered over Singapore, covers the Malay Peninsula and the Sumatran island of Indonesia with 1092 × 1026 points in longitude and latitude, respectively. Eighty levels were used in the vertical that reached up to 38.5 km. The grid resolution was 1.5 km in the horizontal direction and was vertically stretched. Further details of the model's physics and the DA technique can be found in [19,21].

Four sets of simulations were performed for June 2019, cycling eight times per day. Data from two cycles, 03Z and 15Z, were used for analyses as they were of interest to the forecasters for operational use. The forecast length for these two cycles was 42 h. The four sets of simulations were: (a) CTRL, which did not blend any variable; (b) EXPT_uv, which blended horizontal winds u and v; (c) EXPT_uvT, which blended potential temperature in addition to the winds; and (d) EXPT_uvTQ, which blended specific humidity on top of EXPT_uvT. For the purpose of blending, the 6th-order low-pass tangent filter proposed in [22] was used. The filter is implicit and was applied in the physical space, having the following amplitude response

$$F(k) = \left[1 + \varepsilon \, \tan^6\left(\frac{k\Delta}{2}\right)\right]^{-1} \tag{13}$$

for wavenumber k and grid size Δ. The filtering parameter ε is defined as

$$\varepsilon^{-1} = \tan^6\left(\frac{k_c\Delta}{2}\right) \tag{14}$$

such that the energy at the cut-off wavenumber k_c is half the energy at $k = 0$. The cut-off length scale L_c ($=\frac{2\pi}{k_c}$) was set at 900 km for the simulations performed here, which was slightly larger than half the simulation domain (~750 km). The scale selective nature of the

filter is demonstrated in Figure 2, where the horizontally averaged kinetic energy spectra at z = 5 km are shown for the unblended SINGV, driving model (ECMWF), and the blended SINGV for one of the cycles. It should be noted that a shorter cut-off length scale of 500 km is used in this figure for better illustration.

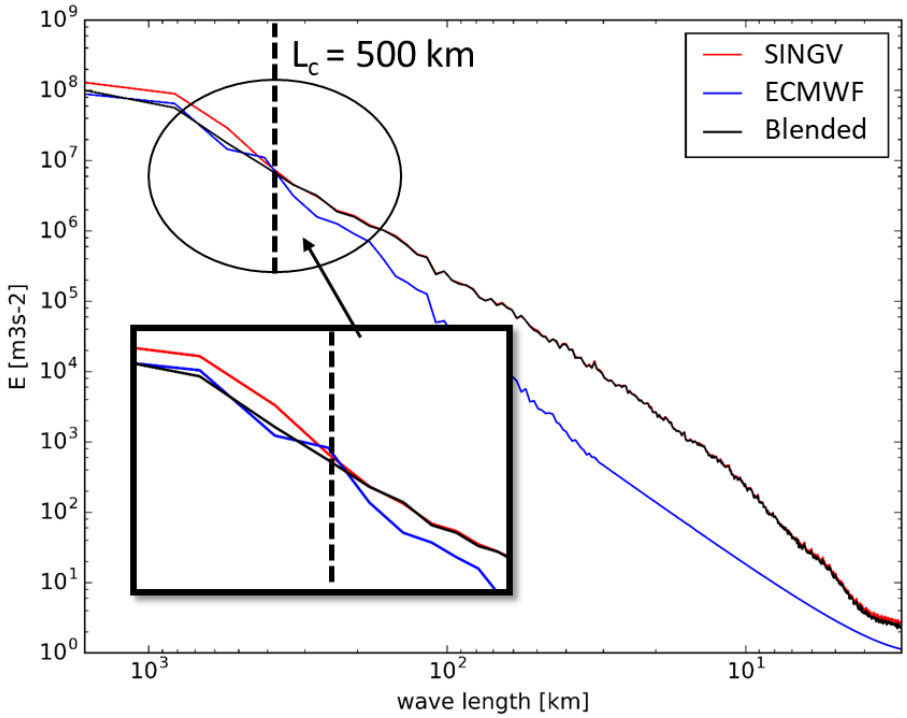

Figure 2. Comparing the horizontally averaged kinetic energy spectra at z = 5 km from unblended SINGV, ECMWF, and blended SINGV forecasts just after initialization to demonstrate the scale-selective nature of the used algorithm. Scales in wavelengths larger than Lc (=500 km) in unblended SINGV (in red) are replaced with that from the ECMWF forecast (in blue), resulting in the blended spectrum (in black).

4. Results and Discussion

The simulation results are discussed in this section. Blended and unblended simulations are first compared against ECMWF forecasts and analysis in Section 4.1 to assess the effectiveness of the blending algorithm. The added value in the blended simulations is then evaluated against observation in Section 4.2.

4.1. Blending Effectiveness

The aim of any blending algorithm is to align the large scales of the otherwise drifting LAM to the driving model. Given that the presented algorithm uses a global model forecast to provide the desired large-scale correction instead of the analysis, it is helpful to check how well the blended forecasts compare against the global model analysis. Figure 3 compares the domain-averaged vertical profile of the root mean squared error (RMSE) against EMCWF analysis, with temperature and eastward wind being used as representative thermodynamic and kinetic variables. The rest of the variables showed similar behavior and are therefore not shown here.

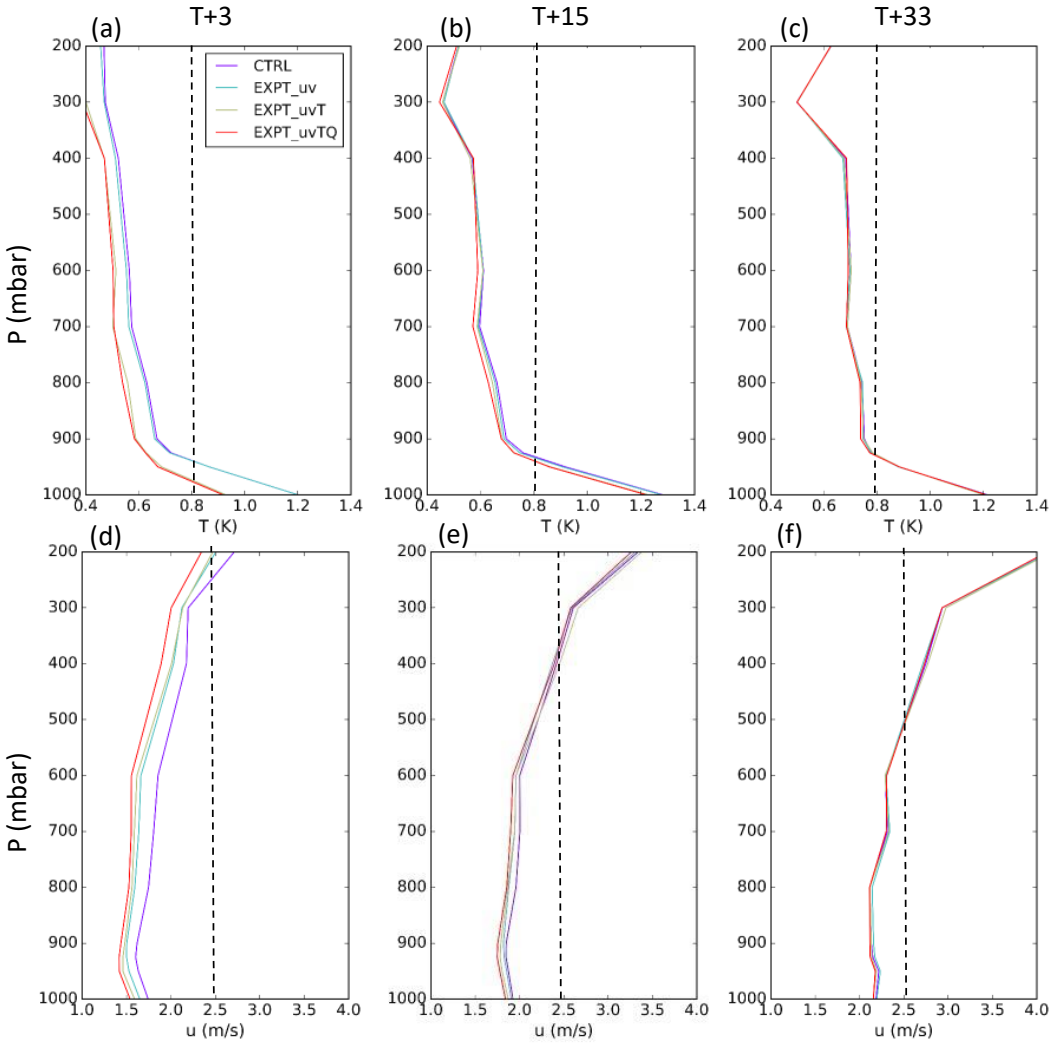

Figure 3. Vertical profiles of root mean squared error (RMSE) of temperature (T), in (**a**–**c**), and eastward wind (u), in (**d**–**f**), with respect to ECMWF analysis for CTRL and experiments at the indicated forecast lengths in hours. The vertical dashed line is at 0.8 K and 2.5 m/s in the top and bottom panels to assist in visualizing the increase in RMSE with the forecast.

Figure 3 shows that RMSE for all the experiments and variables increased with forecast length. For both variables, RMSE had the trend EXPT_uvTQ < EXPT_uvT < EXPT_uv < CTRL in the early forecasts, indicating that the blending algorithm worked as expected despite using the global model forecast. These differences became relatively small toward the end of the simulation, as evident in Figure 3c,f. RMSE reached a minimum when both thermodynamic and kinematic variables were blended, suggesting that blending all the variables is advantageous, as performed in previous studies. However, only blending the winds also helped in reducing the RMSE for temperature (and specific humidity, not shown).

One of the advantages of the current approach is that the large-scale correction $\Delta \mathbf{x_{LC}}$ (see Equation (10)) is combined with the analysis increment to be incrementally applied through the IAU operator. This ensures that the model does not feel unwanted jumps during initialization. This is demonstrated in Figure 4, which compares the time evolution of the absolute mean sea level pressure tendency $\left(\left|\frac{dP_{msl}}{dt}\right|\right)$, horizontally averaged over the model domain, between CTRL and experiments through different cycles. As expected, the pressure tendency in experiments and the CTRL were relatively similar at the nominal analysis time (03Z and 06Z) and a few hours later in the forecast except for in EXPT_uvTQ,

where the differences in tendency between successive cycles were slightly larger than in the CTRL.

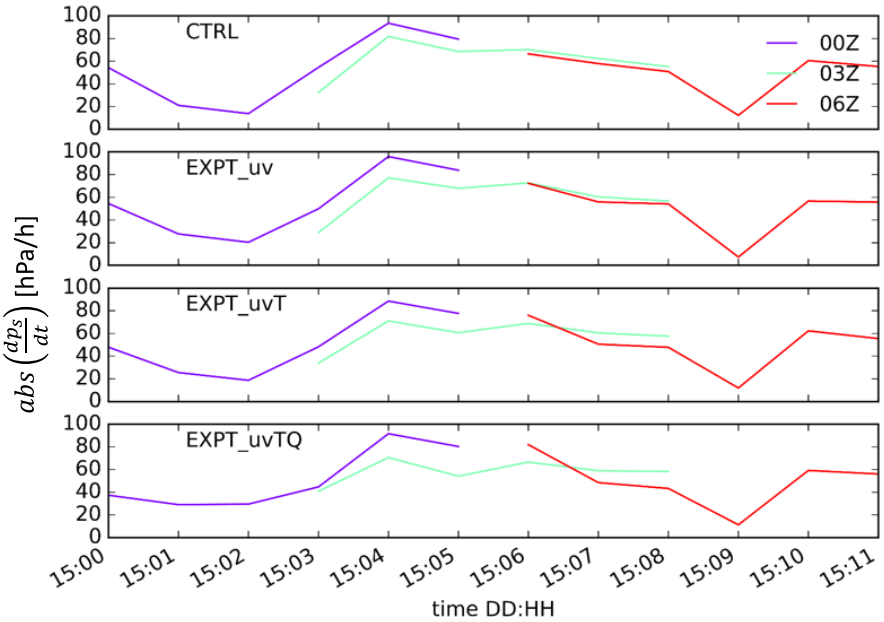

Figure 4. Time evolution of absolute mean sea level pressure tendency $\left(\left|\frac{dP_{msl}}{dt}\right|\right)$, averaged horizontally over the simulation domain, on a randomly selected day.

Figures 3 and 4 suggest that the blending algorithm works as expected by smoothly adjusting the prognostic variables in LAM forecasts (in a large-scale sense) toward a state that is most accurate at the time (ECMWF analysis in present case). To assess how it influences precipitation forecast, which is of high interest in this region, the monthly mean rainfall forecasts from the experiments and CTRL are compared against the ECMWF forecast. It is shown in Figure 5 for the entire simulation domain. Overall, all the experiments and the CTRL reproduced the ECMWF rainfall pattern. The western Sumatran coast (see box 1 in the ECMWF panel) has the heaviest rainfall due to the pronounced diurnal cycle in that region [23,24]. On and off of the Sumatran coast, the rainfall intensity was largest in EXPT_uvTQ and weakest in CTRL. The rainfall pattern in box 1 was best captured in CTRL even in the absence of blending. Box 2, close to the northern boundary, shows light rainfall in ECMWF, whereas the experiments showed signs of heavy rainfall. Box 3 over the South China Sea east of Singapore is of particular interest, where EXPT_uvTQ showed no rainfall, but ECMWF showed signs of light rain.

From Figure 5, we inferred that the effectiveness of blending, as seen in the prognostic variables, does not necessarily translate into the derived quantities such as rainfall. This is not surprising given that rainfall results from the coupled and nonlinear interaction of those prognostic variables. At the same time, however, the poor performance (of the experiments) against the ECMWF forecast does not imply poor forecast quality unless verified against the observation. This is described in the next section.

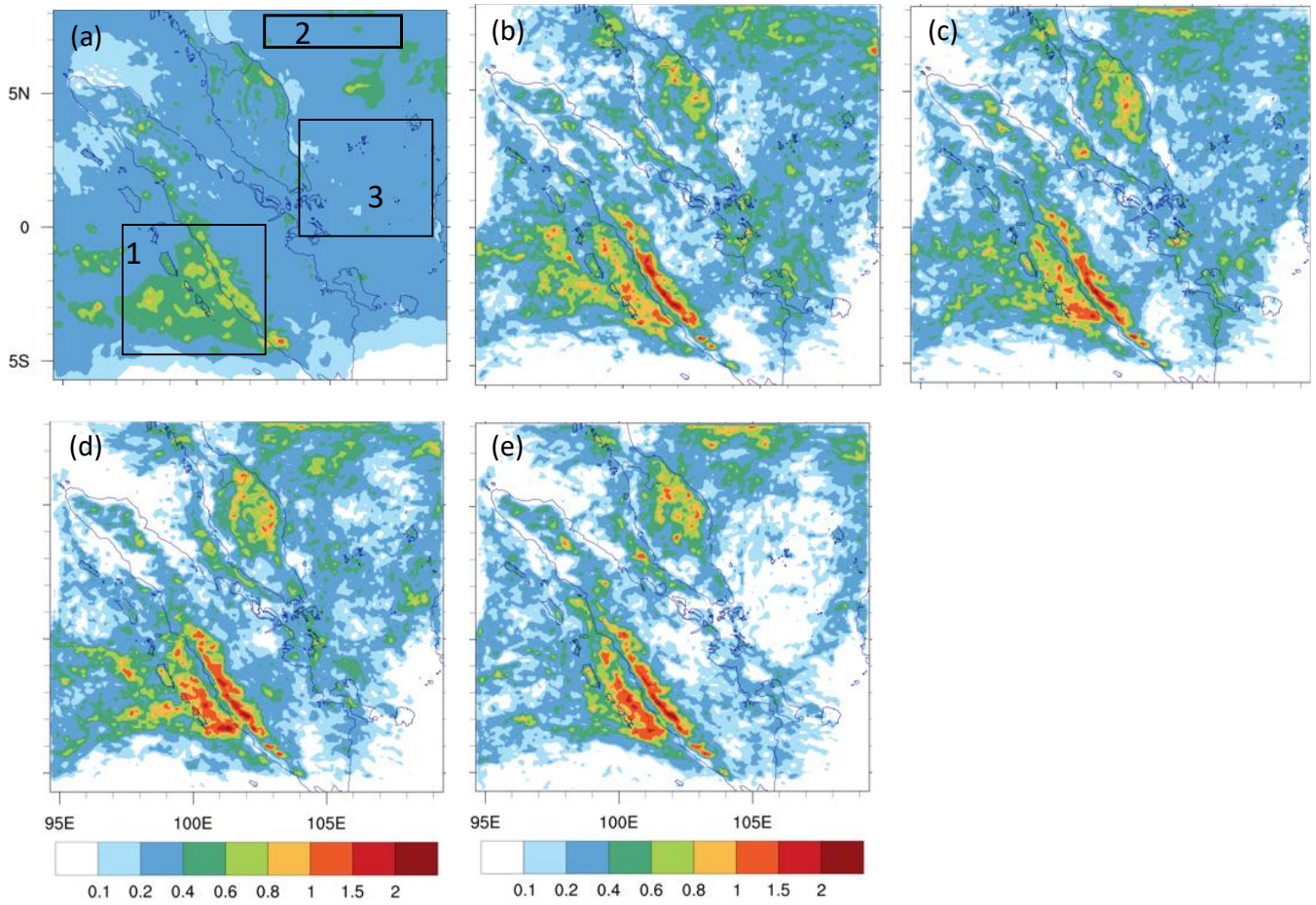

Figure 5. Monthly (June 2019) mean rainfall in millimeters per hour from: (**a**) ECMWF, (**b**) CTRL, (**c**) EXPT_uv, (**d**) EXPT_uvT, and (**e**) EXPT_uvTQ. CTRL and experiments are regrided to ECMWF resolution. The boxes in (**a**) indicate the regions where experiments show marked differences compared to ECMWF forecast.

4.2. Forecast Verification

The advantage of blending is presented in this section, in terms of improvements in forecast verification scores. Standard atmospheric variables u, v, T, and RH for the entire month from the cycles at 03Z and 15Z were verified against the data from 19 radiosonde stations in the whole simulation domain. Error statistics (bias and RMSE) were calculated for selected vertical levels for lead times T+09 and T+33, as a representative for early and late forecast hours, respectively. Attention is given to identifying a systematic of blending on the experiments, across vertical levels or in time or bot).

With respect to CTRL, the relative humidity (RH) error statistics in Figure 6a do not show a systematic impact of blending horizontal winds (EXPT_uv). Adding thermodynamic variables, temperature, and specific humidity in addition to the winds (i.e., EXPT_uvT and EXPT_uvTQ), however, resulted in a systematic reduction in RMSE at 925 and 500 hPa at all lead times. The mean bias, on the other hand, was found to have increased at 925 and 500 hPa for the short (T+09) lead time. The temperature error statistics in Figure 6b show a large spread. There was a systematic reduction in mean bias when temperature is added to the winds (EXPT_uvT) at upper levels at all the lead times. The RMSE was also found to decrease in all the blending experiments at 925 hPa at all lead times. As for the horizontal winds, there was no significant impact of blending, except for a systematic reduction in RMSE in the meridional wind at 925 and 700 hPa at T+09 in all blending experiments (see Figure 6c).

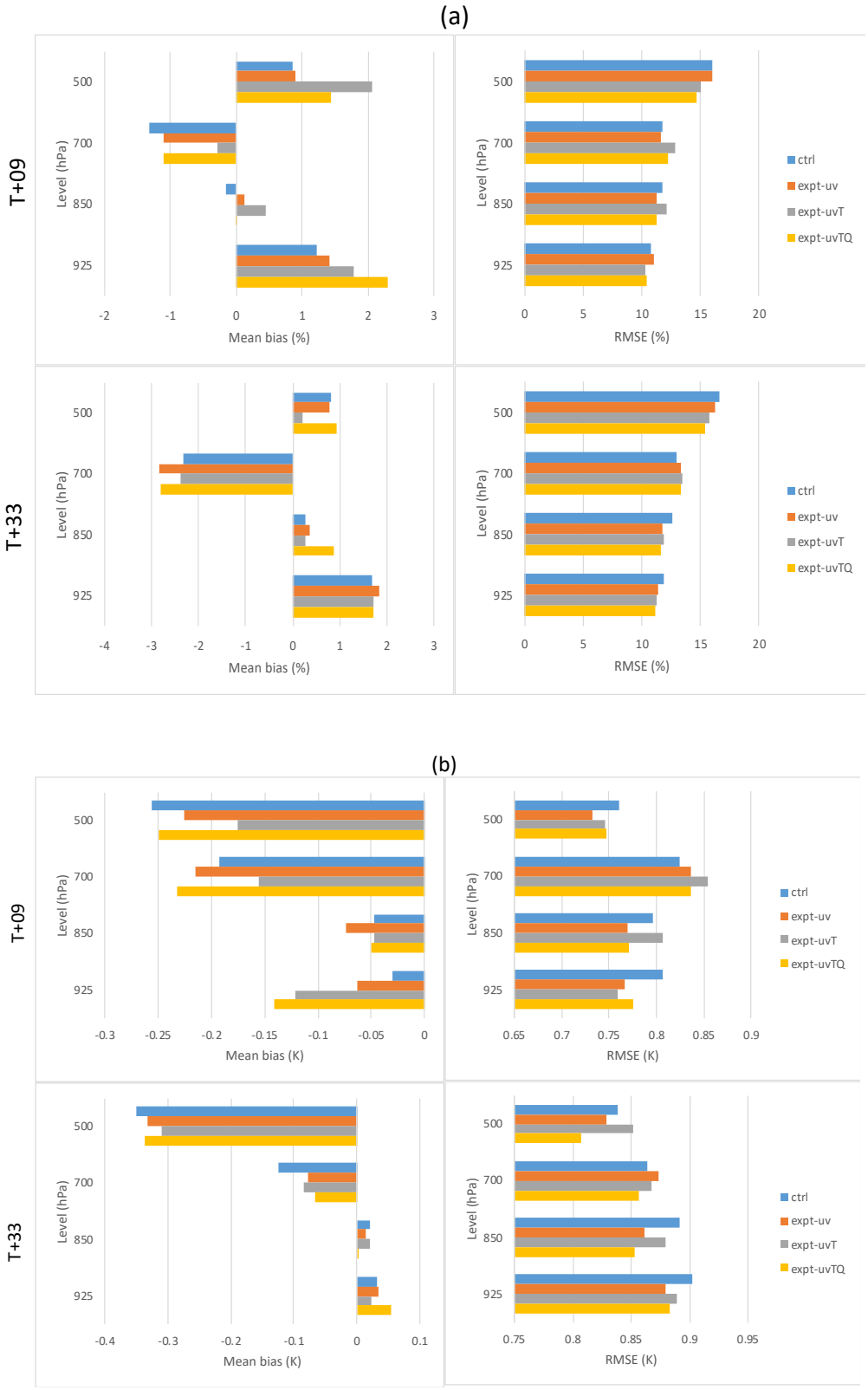

Figure 6. *Cont.*

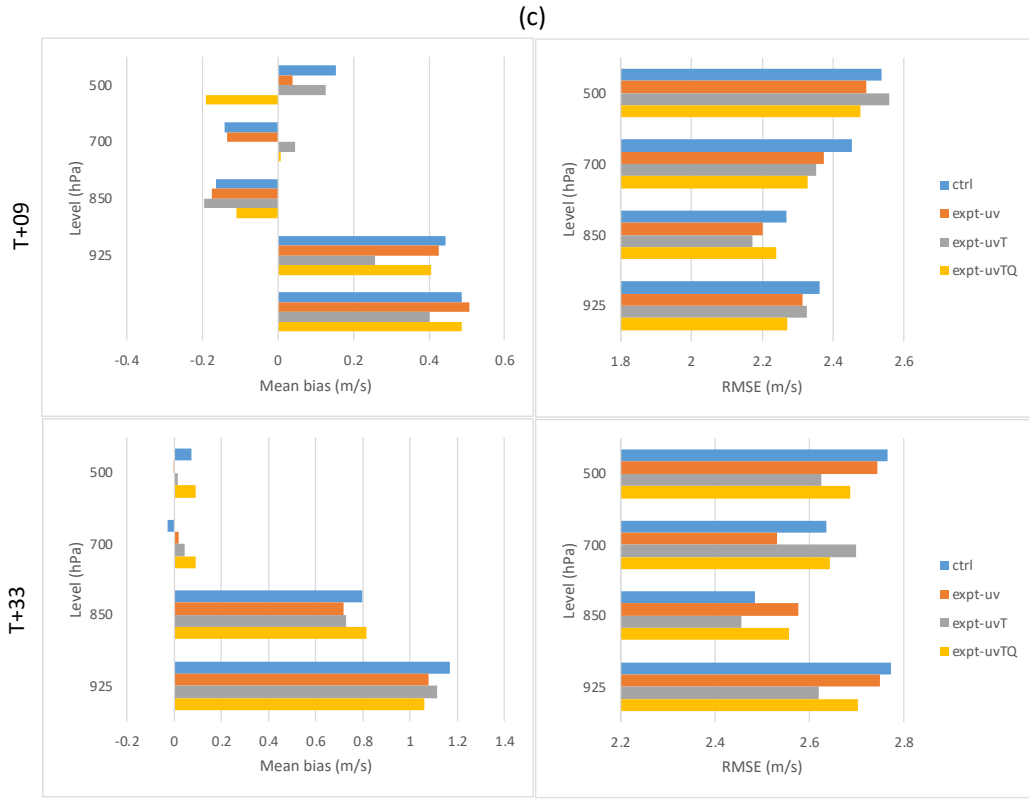

Figure 6. Mean error statistics: (**a**) relative humidity (RH), (**b**) temperature (T), and (**c**) meridional wind (v) estimated against all stations in the domain are shown for 03Z and 15Z cycles at the indicated levels.

Not shown here, but the differences in mean bias and RMSE between the experiments generally fell within the uncertainty (95% confidence) and therefore could not be concluded to be statistically significant.

The precipitation forecast was evaluated against the Global Precipitation Measurement (GPM) [25] precipitation product. Fractions skill score (FSS) diagnostics [26] were used to allow for a scale-selective evaluation of the precipitation forecasts. Nearest neighbors were used to identify the scale of interest. The scores were then calculated for these neighbors of different sizes for different forecast lead times. The scores vary from 0 (worst) to 1 (best) for a given "truth". Here, we used the GPM as the truth and calculated the FSS using three-hourly accumulated rainfall for the entire month, from 03Z and 15Z cycles. Scores were calculated for different length scales but here we summarize the results for 10 km × 10 km grid boxes, equivalent to one GPM grid box. Hinton diagrams [27,28] were used to visualize a matrix of the scores whereby the abscissa was the forecast length and the ordinate was the precipitation threshold (both absolute and relative). The triangle area was the difference in FSSs between experiment and control (e.g., EXPT_uv-CTRL in Figure 7a), with green upward-pointing triangles indicating higher skill in the experiment and purple downward-pointing triangles indicating otherwise. The maximum difference was 0.0731, which dictates the size of the largest triangle. Bold black lines around the triangles indicate statistically significant scores at the 0.05 level, as determined by the Wilcoxon signed-rank test [29].

Contrasting the findings in Figure 5, the verification scores in Figure 7 show marked improvements in the precipitation forecast due to blending. In particular, we found a significant improvement in EXPT_uv (see Figure 7a) in the first 18 h. Adding temperature to the blend (see Figure 7b) made T+3, T+21, and T+24 forecast significantly worse, reducing the benefits of blending the horizontal winds. Adding specific humidity (EXPT_uvTQ) to the blend further degraded forecasts up to T+6, but improvements were found in

the later hours. This result is puzzling as a similar study using SINGV over the same simulation domain [30] found significant improvement in the precipitation scores (see Figures 9 and 10 in Ref. [30]) due to the addition of the thermodynamic variables in the assimilation cycle. Drawing parallels between [30] and the present study is, however, not fair as the moisture (and temperature) source in [30] was a localized radiosonde as opposed to large-scale contributions from the driving model in the present study. However, Ref. [30] highlighted the longer adjustment period after initialization, which is due to the inclusion of thermodynamic variables (see Section 5b in [30]). We think this is the reason for forecast degradation in the early hours in the present study.

It is an important finding as none of the earlier studies had attempted isolating the impact of blending different variables, which, based on the results shown here, is clearly required.

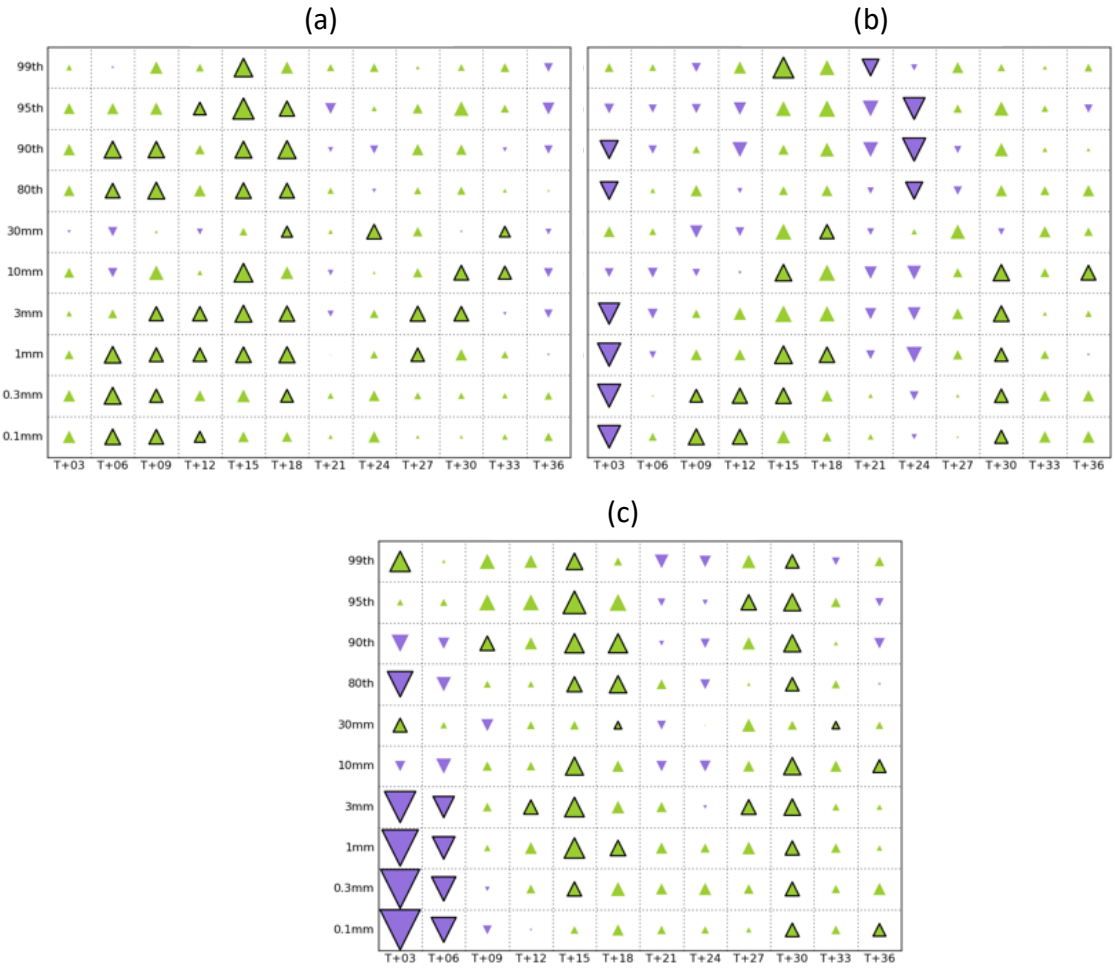

Figure 7. Hinton diagrams showing the differences in fractions skill scores (max = 0.073) for June 2019 for (**a**) EXPT_uv—CTRL, (**b**) EXPT_uvT—CTRL, and (**c**) EXPT_uvTQ—CTRL. Here, green upward-pointing triangles indicate higher skill for experiments, and purple downward-pointing triangles indicate otherwise. A range of absolute and relative rainfall thresholds (ordinate) for a grid box size 10 km × 10 km are used to summarize the impact at increasing forecast lengths along the abscissa.

5. Conclusions

An alternate blending approach was presented that has merits over the existing approaches. In contrast to the most widely used background blending approach of [5], the presented approach is simpler to implement and does not suffer from inconsistencies in the background state ($\mathbf{x^b}$) and the background error covariance matrix ($\mathbf{B}$). The new

approach is similar to analysis blending in its implementation, which is also applied after the DA step. However, because it relies on forecasts instead of the analyses for the implied large-scale correction, this technique is termed forecast blending.

This new approach was implemented in SINGV, which uses 3DVAR-FGAT, and was tested over the tropical domain used operationally at MSS. In addition to the control, which was an unblended simulation, three simulations were performed to assess the benefits of blending dynamic and thermodynamic variables.

In agreement with the earlier findings, the results presented here demonstrate a clear advantage of improved large-scale information in the initial condition of a limited-area NWP model. However, caution is warranted when deciding which variables to use for blending. When comparing the verification scores for CTRL, EXPT_uv, EXPT_uvT, and EXPT_uvTQ, blending the horizontal winds alone (EXPT_uv) seemed to be best, especially over the tropics, where reliable precipitation forecast is of high importance. Significant degradation in forecast quality occurred in the early hours when the thermodynamic variables are added to the blend.

It is also noted that the presented approach was only tested over a specific domain for a limited period, and more tests are needed to demonstrate its usability elsewhere.

Author Contributions: Conceptualization, X.-Y.H.; methodology, A.D. and. X.-Y.H.; software, A.D., validation, A.D., X.-Y.H. and P.H.; formal analysis, A.D. and X.-Y.H.; data curation, A.D. and P.H.; writing original draft preparation, A.D.; writing-review and editing, A.D., X.-Y.H. and P.H.; visualization, A.D. and P.H.; supervision, A.D.; project administration, X.-Y.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Meteorological Services Singapore internal funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Jerry Liu and Joshua Lee from MSS for their continued support and discussion on the topic. Discussions with Adam Clayton and Andrew Lorenc from the U.K. Met Office are also appreciated. The first author would also like to thank the MSS NWP team and CCRS director Dale Barker for bringing this research effort into operation through SINGV 5.1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Warner, T.T.; Peterson, R.A.; Treadon, R.E. A Tutorial on Lateral Boundary Conditions as a Basic and Potentially Serious Limitation to Regional Numerical Weather Prediction. *Bull. Am. Meteorol. Soc.* **1997**, *78*, 2599–2617. [CrossRef]
2. von Storch, H.; Langenberg, H.; Feser, F. A Spectral Nudging Technique for Dynamical Downscaling Purposes. *Mon. Weather Rev.* **2000**, *128*, 3664–3673. [CrossRef]
3. Guidard, V.; Fischer, C. Introducing the coupling information in a limited-area variational assimilation. *Q. J. R. Meteorol. Soc.* **2008**, *134*, 723–735. [CrossRef]
4. Kalnay, E. *Atmospheric Modeling, Data Assimilation and Predictability*; Cambridge University Press (CUP): Cambridge, UK, 2002.
5. Yang, X. Background blending using a spatial filter. *HIRLAM Newsl.* **2005**, *49*, 3–11.
6. Sun, J.; Trier, S.B.; Xiao, Q.; Weisman, M.L.; Wang, H.; Ying, Z.; Xu, M.; Zhang, Y. Sensitivity of 0–12-h Warm-Season Precipitation Forecasts over the Central United States to Model Initialization. *Weather Forecast.* **2012**, *27*, 832–855. [CrossRef]
7. Lynch, P.; Huang, X.-Y. Initialization of the HIRLAM Model Using a Digital Filter. *Mon. Weather Rev.* **1992**, *120*, 1019–1034. [CrossRef]
8. Brožkovà, R.; Klarić, D.; Ivatek-Šahdan, S.; Geleyn, J.-F.; Casse, V.; Širokà, M.; Radnoti, G.; Janoušek, M.; Stadlbacher, K.; Seidl, H. *DFI Blending: An Alternative Tool for Preparation of the Initial Conditions for LAM*; WMO CAS/JSC WGNE Report; Ritchie, H., Ed.; World Meteorological Organization-Publications-Wmo Td: Geneva, Switzerland, 2001; Volume 31, Chapters 1.7–1.8.
9. Short, C.J.; Petch, J. Reducing the spin-up of a regional NWP system without data assimilation. *Q. J. R. Meteorol. Soc.* **2022**, *148*, 1623–1643. [CrossRef]
10. Dahlgren, P.; Gustafsson, N. Assimilating host model information into a limited area model. *Tellus A Dyn. Meteorol. Oceanogr.* **2012**, *64*. [CrossRef]

11. Vendrasco, E.P.; Sun, J.; Herdies, D.L.; de Angelis, C.F. Constraining a 3DVAR Radar Data Assimilation System with Large-Scale Analysis to Improve Short-Range Precipitation Forecasts. *J. Appl. Meteorol. Clim.* **2016**, *55*, 673–690. [CrossRef]
12. Wang, H.; Huang, X.-Y.; Xu, D.; Liu, J. A scale-dependent blending scheme for WRFDA: Impact on regional weather forecasting. *Geosci. Model Dev.* **2014**, *7*, 1819–1828. [CrossRef]
13. Yang, X. Analysis blending using a spatial filter in grid-point model coupling. *HIRLAM Newsl.* **2005**, *48*, 49–55.
14. Hsiao, L.-F.; Huang, X.-Y.; Kuo, Y.-H.; Chen, D.-S.; Wang, H.; Tsai, C.-C.; Yeh, T.-C.; Hong, J.-S.; Fong, C.-T.; Lee, C.-S. Blending of Global and Regional Analyses with a Spatial Filter: Application to Typhoon Prediction over the Western North Pacific Ocean. *Weather Forecast.* **2015**, *30*, 754–770. [CrossRef]
15. Feng, J.; Sun, J.; Zhang, Y. A Dynamic Blending Scheme to Mitigate Large-Scale Bias in Regional Models. *J. Adv. Model. Earth Syst.* **2020**, *12*. [CrossRef]
16. Clayton, A. UK Met Office, Personal Communications.
17. Ide, K.; Coutier, P.; Ghil, M.; Lorenc, A.C. Unified notation for data assimilation: Operational, sequential and variational. *J. Meteorol. Soc. Jpn.* **1997**, *75*, 181–189. [CrossRef]
18. Huang, X.Y.; Barker, D.; Webster, S.; Dipankar, A.; Lock, A.; Mittermaier, M.; Sun, X.; North, R.; Darvell, R.; Boyd, D.; et al. SINGV–the Convective-Scale Numerical Weather Prediction System for Singapore. *ASEAN J. Sci. Technol. Dev.* **2019**, *36*, 81–90. [CrossRef]
19. Dipankar, A.; Webster, S.; Sun, X.; Sanchez, C.; North, R.; Furtado, K.; Wilkinson, J.; Lock, A.; Vosper, S.; Huang, X.-Y.; et al. SINGV: A convective-scale weather forecast model for Singapore. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 4131–4146. [CrossRef]
20. Bush, M.; Allen, T.; Bain, C.; Boutle, I.; Edwards, J.; Finnenkoetter, A.; Franklin, C.; Hanley, K.; Lean, H.; Lock, A.; et al. The first Met Office Unified Model–JULES Regional Atmosphere and Land configuration, RAL1. *Geosci. Model Dev.* **2020**, *13*, 1999–2029. [CrossRef]
21. Heng, B.C.P.; Tubbs, R.; Huang, X.-Y.; MacPherson, B.; Barker, D.; Boyd, D.F.A.; Kelly, G.; North, R.; Stewart, L.; Webster, S.; et al. SINGV-DA: A data assimilation system for convective-scale numerical weather prediction over Singapore. *Q. J. R. Meteorol. Soc.* **2020**, *146*, 1923–1938. [CrossRef]
22. Raymond, W.H. High-Order Low-Pass Implicit Tangent Filters for Use in Finite Area Calculations. *Mon. Weather Rev.* **1988**, *116*, 2132–2141. [CrossRef]
23. Mori, S.; Jun-Ichi, H.; Tauhid, Y.I.; Yamanaka, M.D.; Okamoto, N.; Murata, F.; Sakurai, N.; Hashiguchi, H.; Sribimawati, T. Diurnal Land–Sea Rainfall Peak Migration over Sumatera Island, Indonesian Maritime Continent, Observed by TRMM Satellite and Intensive Rawinsonde Soundings. *Mon. Weather Rev.* **2004**, *132*, 2021–2039. [CrossRef]
24. Dipankar, A.; Webster, S.; Huang, X.-Y.; Doan, V.Q. Understanding Biases in Simulating the Diurnal Cycle of Convection over the Western Coast of Sumatra: Comparison with Pre-YMC Observation Campaign. *Mon. Weather Rev.* **2019**, *147*, 1615–1631. [CrossRef]
25. Skofronick-Jackson, G.; Petersen, W.A.; Berg, W.; Kidd, C.; Stocker, E.F.; Kirschbaum, D.B.; Kakar, R.; Braun, S.; Huffman, G.J.; Iguchi, T.; et al. The Global Precipitation Measurement (GPM) Mission for Science and Society. *Bull. Amer. Meteor. Soc.* **2017**, *98*, 1679–1695. [CrossRef] [PubMed]
26. Roberts, N.M.; Lean, H.W. Scale-Selective Verification of Rainfall Accumulations from High-Resolution Forecasts. *Mon. Weather Rev.* **2008**, *136*, 78–97. [CrossRef]
27. Hinton, G.E.; Shallice, T. Lesioning an attractor network: Investigations of acquired dyslexia. *Psychol. Rev.* **1991**, *98*, 74–95. [CrossRef] [PubMed]
28. Bremner, F.J.; Gotts, S.J.; Denham, D.L. Hinton diagrams: Viewing connection strengths in neural networks. *Behav. Res. Methods Instrum. Comput.* **1994**, *26*, 215–218. [CrossRef]
29. Wilks, D.S. *Statistical Methods in the Atmospheric Sciences*, 3rd ed.; Elsevier: Amsterdam, The Netherlands, 2011; pp. 162–166.
30. Lee, J.C.K.; Dipankar, A.; Huang, X.-Y. On the Sensitivity of the Simulated Diurnal Cycle of Precipitation to 3-Hourly Radiosonde Assimilation: A Case Study over the Western Maritime Continent. *Mon. Weather Rev.* **2021**, *149*, 3449–3468. [CrossRef]