Hybrid assimilation on a parameter-calibrated model to improve the prediction of heavy rainfall events during the Indian summer monsoon

Sandeep Chinta1,*, V. S. Prasad2 and C. Balaji1,3,4

1Department of Mechanical Engineering, Indian Institute of Technology Madras, Chennai 600 036, India
2National Centre for Medium Range Weather Forecasting, A-50, Sector 62, Noida 201 309, India
3Centre of Excellence in Atmospheric and Climate Sciences, Indian Institute of Technology Madras, Chennai 600 036, India
4Divecha Centre for Climate Change, Indian Institute of Science, Bengaluru 560 012, India

Heavy rainfall events during the Indian summer monsoon cause landslides and flash floods resulting in a significant loss of life and property every year. The exactness of the model physics representation and initial conditions is critical for accurately predicting these events using a numerical weather model. The values of parameters in the physics schemes influence the accuracy of model prediction; hence, these parameters are calibrated with respect to observation data. The present study examines the influence of hybrid data assimilation on a parameter-calibrated WRF model. Twelve events during the period 2018–2020 were simulated in this study. Hybrid assimilation on the WRF model significantly reduced the model prediction error of the variables: rainfall (18.04%), surface air temperature (7.91%), surface air pressure (5.90%) and wind speed at 10 m (27.65%) compared to simulations with default parameters without assimilation.

Keywords: Heavy rainfall events, hybrid assimilation, numerical weather model, parameter calibration, summer monsoon.

The Indian summer monsoon (ISM) is among the oldest global monsoon phenomena occurring with striking regularity every year. Heavy rainfall events during ISM cause landslides and flash floods resulting in a significant loss of life and property each year1. The number of low and moderate rainfall events averaged over the entire Indian region has substantially reduced during ISM, whereas heavy rainfall events have increased over the years2. Also, there has been a noticeable increase in both the average frequency of heavy rainfall events and the percentage of seasonal rainfall contributed by these events3. So, accurately simulating the heavy rainfall events during the ISM is crucial.

The accuracy of a numerical weather prediction (NWP) model depends upon both its ability to accurately represent the physics of the atmosphere and the precision of the initial conditions provided to the model4. In the NWP model, sub-grid-scale processes are parameterized based on reasonable physical or statistical representations. The parameterization schemes require information from the forecast variables about the process to be parameterized using a set of assumptions5. Multiple studies have been conducted to identify the optimal set of parameterization schemes for various regions and types of simulation events6–12.

However, each parameterization scheme contains multiple parameters on which the scheme is formulated. Typically, the default values of these parameters are determined through theoretical or experimental studies by the developers of the scheme13. By calibrating the values of these parameters to observations, the accuracy of the prediction can be improved by increasing the ability of the model to accurately represent the physics of the atmosphere. Parameter calibration based on tuning to an objective can be classified into two categories14. The first category involves optimizing an objective function that evaluates the difference between the model simulation and a corresponding set of observations. To accelerate the optimization process, a model emulator or surrogate model of the actual physical model is constructed. Several studies have used this approach15,16. The second category quantifies critical uncertainty sources in the problem by employing a Bayesian approach. Some methods in this category use the actual model, while others use a statistical emulator to eliminate regions of the parameter space that are not physically possible and yield not-ruled-out-yet parameter space. Although the second category is robust in identifying the viable parameter space, it requires huge computational power compared to the first category. Many studies have used this approach for climate models17,18. As trial-and-error methods typically concentrate on tuning a limited number of parameters (usually one or two) at a time, the above two categories are more beneficial for calibrating a considerable number of parameters. Overtuning is a crucial factor to consider during parameter calibration. It refers to calibrating the values of the
parameters to specific metrics resulting in the model performing well for those metrics, while the performance deteriorates for other processes or metrics. Therefore, conducting validation experiments for different metrics and simulating events not used for calibration can help check whether the parameter calibration resulted in over tuning.

It is computationally expensive to calibrate all the parameters. So, a sensitivity study is necessary to determine the parameters that profoundly impact model prediction. Several studies have used various sensitivity analysis techniques to identify the sensitive parameters for different regions and types of events simulated. Studies have also calibrated the sensitive parameters with respect to observations using advanced optimization techniques to improve model prediction. The model parameters have been optimized to obtain a better forecast for various models such as watershed model, atmospheric general circulation model, and climate model to reduce the prediction error.

Apart from the calibration of parameters, accuracy of the model prediction also depends on the exactness of the initial conditions. Data assimilation is used to improve the initial conditions utilizing observational data. Data assimilation algorithms require the background state of the atmosphere obtained from the short-range forecast of the previous cycle. The background state, as it is a forecast, contains some uncertainty. Different assimilation algorithms handle this uncertainty differently in the form of an error covariance matrix. The assimilation algorithm corrects the background state using data from the observations and considering the error covariances of the background state and the observations. The background error covariance (BEC) matrix plays a critical role in data assimilation, particularly for weather systems such as heavy rainfall events that are intermittent and transient.

Three-dimensional variational (3DVar) assimilation method utilizes a static BEC and does not contain flow-dependent spatial covariance, which means that the errors pertaining to the flow of the day are assumed to be invariant. Ensemble Kalman filter data assimilation generates the ensembles using the Monte Carlo method, and flow-dependent BEC is obtained from the ensemble of forecasts. However, sampling error in the ensemble BEC is significant because of the relatively small sample size. In hybrid assimilation, the weighted average of the static and flow-dependent BEC is used within a 3DVar framework. Several studies have broadly concluded that hybrid data assimilation performs better than the variational assimilation techniques and is sometimes equivalent or better than pure ensemble assimilation methods.

This study builds up on two previous studies. In the first study, sensitivity analysis was performed and the parameters that significantly influenced heavy rainfall events prediction were identified using the Morris one-at-a-time method. In the second study, these sensitive parameters were calibrated with respect to observations using a multi-objective adaptive surrogate model-based optimization method. These two previous studies only addressed one aspect of improving the NWP model prediction: increasing the accuracy with which the model represents the physics of the atmosphere. However, these studies did not consider the second aspect of improving the model prediction: increasing the accuracy of the initial conditions provided to the model. The present study addresses both these aspects by implementing hybrid assimilation on a parameter-calibrated model to improve prediction.

**Data assimilation methodology**

Two data assimilation algorithms, namely three-dimensional variational (3DVar) assimilation and three-dimensional ensemble variational (3DEnVar) hybrid assimilation have been used in this study. These algorithms are implemented using the Weather Research and Forecasting (WRF) model data assimilation system (WRFDA). In the 3D-Var algorithm, the analysis field \( X_A \) is estimated by minimizing a cost function (that calculates the distance from the background \( X_b \) and from the observations \( Y_o \) as formulated as

\[
J(X) = (X - X_b)^T B^{-1} (X - X_b) + [Y_o - H(X)]^T R^{-1} [Y_o - H(X)],
\]

where \( J(X) \) is the cost function that needs to be minimized, \( B \) the BEC matrix estimated as the error (averaged over many forecasts) between two short-range forecasts valid at a particular time, \( H \) the forward operator that maps variables from the state space to the observation space and \( R \) the observation error covariance matrix that contains the instrument error. In the National Meteorological Center method, BEC is formulated as

\[
B_{3DVar} \approx \alpha E[X_f(24 \text{ h}) - X_f(12 \text{ h})]
\]

\[
\times [X_f(24 \text{ h}) - X_f(12 \text{ h})]^T,
\]

where \( X_f(24 \text{ h}) \) is the 24th hour and \( X_f(12 \text{ h}) \) is the 12th hour regional forecast valid at a particular time. Different control variable (cv) transform options are available in WRFDA to evaluate \( B_{3DVar} \) where each cv transform focuses on different control variables. Some studies have observed that cv6 option performed better compared to other options, such as cv3, cv5 and cv7 in WRFDA. The control variables used in cv6 are stream function \( \psi \), unbalanced temperature \( T_u \), unbalanced velocity potential \( \zeta_u \), unbalanced surface pressure \( P_{su} \) and unbalanced pseudo relative humidity \( RH_u \). As the assimilation is performed on a parameter-calibrated model, the value of \( B_{3DVar} \) is evaluated with the cv6 option using simulations from a parameter-calibrated model. Simulations were done for July–August.
2016, and the 24th hour and 12th hour forecasts were evaluated at 12-h intervals for these two months. These forecasts were utilized to calculate $B_{3DVar}$ using eq. (2).

In hybrid assimilation, BEC was estimated using a weighted average of the variational and ensemble BEC, as already mentioned. The background error covariance $B_{hyb}$ was evaluated using eq. (3).

$$B_{hyb} = (1 - \beta)B_{Ens} + \beta B_{3DVar},$$

where $\beta$ is the weighting factor and $B_{Ens}$ is the ensemble BEC. Figure 1 presents a flowchart of the hybrid data assimilation methodology adopted in this study. The ensemble BEC was evaluated using ensemble forecasts. Global Ensemble Forecast System (GEFS) data were used as the driving data for providing initial and boundary conditions for the ensembles. As GEFS data contain 21 ensemble members, the hybrid assimilation also contains 21 ensemble members ($k = 21$). A value of $\beta = 0.25$ was used to evaluate $B_{hyb}$ (ref. 44), i.e. 25% of $B_{3DVar}$ and 75% of $B_{Ens}$. After the processing, quality control and thinning of observation data in WRFDA, the background forecast was updated using data observations and $B_{hyb}$. The updated background forecast is called analysis and contains the improved initial conditions. This was used as the background for the following forecast cycle. Ensemble members were also updated using the observation data. After ensemble assimilation, the ensemble background forecasts were updated to ensemble analysis for the next forecast cycle. This procedure was repeated continuously for several cycles. Single-resolution assimilation was performed in this study, where both the deterministic and ensemble forecasts were simulated on a similar grid resolution. The ensemble analyses are recentred in some studies using the deterministic analysis for the next forecast cycle. However, recentring was not used in this study due to its small impact on the ensemble analyses.

**Design of experiments**

**Domain and model configuration**

The domain has a resolution of 12 km in the horizontal direction (Figure 2). WRF model version 4.0 was used in this study. The domain covers the Indian subcontinent and its...
Table 1. List of parameters and their default and calibrated values

<table>
<thead>
<tr>
<th>Scheme</th>
<th>Parameter</th>
<th>Default value</th>
<th>Calibrated value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cumulus</td>
<td>pd</td>
<td>1</td>
<td>2</td>
<td>Multiplier for the downdraft mass flux rate</td>
</tr>
<tr>
<td></td>
<td>pe</td>
<td>1</td>
<td>0.5693</td>
<td>Multiplier for the entrainment mass flux rate</td>
</tr>
<tr>
<td></td>
<td>phusl</td>
<td>150</td>
<td>349.9996</td>
<td>Starting height of downdraft over USL (hPa)</td>
</tr>
<tr>
<td></td>
<td>timec</td>
<td>2,700</td>
<td>2,770.13</td>
<td>Mean consumption time of CAPE (s)</td>
</tr>
<tr>
<td>Microphysics</td>
<td>ice stokes</td>
<td>14,900</td>
<td>18,147.36</td>
<td>Scaling factor applied to ice-fall velocity (s⁻¹)</td>
</tr>
<tr>
<td>Shortwave</td>
<td>cssca fac</td>
<td>1e-5</td>
<td>5.73e-6</td>
<td>Scattering tuning parameter (m² kg⁻¹)</td>
</tr>
<tr>
<td>Land surface</td>
<td>porsl</td>
<td>1</td>
<td>0.5</td>
<td>Multiplier for saturated soil water content</td>
</tr>
<tr>
<td></td>
<td>bsw</td>
<td>1</td>
<td>1.168</td>
<td>Multiplier for Clapp and Hornberger $b$ parameter</td>
</tr>
<tr>
<td>Planetary boundary layer</td>
<td>Brs cr sb</td>
<td>0.25</td>
<td>0.4423</td>
<td>Critical Richardson number for the boundary layer of land</td>
</tr>
</tbody>
</table>

Figure 3. Daily regional average rainfall in the MCR during Indian summer monsoon for the period 2018–2020. Solid line boxes show the events that are simulated.

surrounding regions. It comprises 470 points in the zonal direction and 460 points in the meridional direction. The central point is 19°N, 80°E. The time-step used is 40 s. The model is discretized into 40 sigma ($\sigma$) layers vertically with the top layer at 50 hPa level in the atmosphere. The initial and lateral boundary conditions are obtained from the National Centers for Environmental Prediction Global Forecast System model six-hourly data at 0.5° × 0.5° resolution.

The parameterization schemes were the same as those used in parameter calibration⁴⁹. The Kain–Fritsch eta scheme was used for cumulus parameterization⁵⁶, the WSM 6 single-class scheme for microphysics parameterization⁵⁷, the Dudhia scheme for shortwave radiation⁵⁸, the RRTM scheme for longwave radiation⁵⁹, the MM5 Monin–Obukhov scheme for the surface layer⁶⁰, the Yonsei University scheme for the planetary boundary layer⁶¹, and the Noah scheme for land surface parameterization⁶². As mentioned earlier, each scheme has multiple parameters that can be calibrated. Table 1 lists nine sensitive parameters that influence the prediction⁴⁷. The table also presents the default and calibrated values of these sensitive parameters. The calibrated parameter values minimized the prediction errors of daily accumulated rainfall (RAIN), surface air temperature (SAT), surface air pressure (SAP) and wind speed at 10 m (WS10)⁶⁹.

Events simulated

The monsoon core region (MCR) is a critical zone where the variation of ISM rainfall all over the country is similar
Figure 4. Distribution of conventional observations at 00 UTC on 7 July 2019 assimilated for event F.

Table 2. Overview of the observations assimilated

<table>
<thead>
<tr>
<th>NCEP prepbufr observations</th>
<th>Satellite winds: GEOAMV</th>
<th>Land surface: SYNOP, METAR, SONDE SFC</th>
<th>Marine surface: BUOY, SHIPS</th>
<th>Upper air: SOUND</th>
</tr>
</thead>
</table>

Satellite radiance observations

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Satellite</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMSU-A</td>
<td>NOAA 15,16,18,19; EOS-Aqua; METOP-A</td>
</tr>
<tr>
<td>AIRS</td>
<td>EOS-Aqua</td>
</tr>
<tr>
<td>HIRS-4</td>
<td>NOAA 18</td>
</tr>
<tr>
<td>IASI</td>
<td>METOP-A</td>
</tr>
<tr>
<td>MHS</td>
<td>METOP-A; NOAA 18,19</td>
</tr>
<tr>
<td>SSMI</td>
<td>DMSP 16</td>
</tr>
</tbody>
</table>

Table 3. Details of numerical experiments

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEF_NA</td>
<td>Simulations with default parameters without data assimilation</td>
</tr>
<tr>
<td>CAL_NA</td>
<td>Simulations with calibrated parameters without data assimilation</td>
</tr>
<tr>
<td>DEF_3DV</td>
<td>Simulations with default parameters and three-dimensional variational data assimilation</td>
</tr>
<tr>
<td>DEF_HYB</td>
<td>Simulations with default parameters and three-dimensional ensemble variational hybrid data assimilation</td>
</tr>
<tr>
<td>CAL_3DV</td>
<td>Simulations with calibrated parameters and three-dimensional variational data assimilation</td>
</tr>
<tr>
<td>CAL_HYB</td>
<td>Simulations with calibrated parameters and three-dimensional ensemble variational hybrid data assimilation</td>
</tr>
</tbody>
</table>

The observation data used for assimilation are summarized in Table 2, and include both conventional and satellite radiance data. Observations were assimilated at six-hourly intervals. Conventional observations comprise global surface and upper-air data collected by NCEP and are provided in PREPBUFR format. The data contain land surface, marine surface, radiosonde, pibal and aircraft reports from the Global Telecommunications System, profiler and satellite wind data. Figure 4 presents the conventional data from various sources valid at 00 UTC on 7 July 2019, used in the assimilation for event F. Satellite radiance data (BUFR format) from various instruments on-board different satellites were used for assimilation. The instruments include Advanced Microwave Sounding Unit-A, Atmospheric Infrared Sounder (AIRS), High-resolution Infrared Sounder-4, Infrared Atmospheric Sounding Interferometer, Microwave Humidity Sounder and Special Sensor Microwave/Imager.

The WRF model output variables were compared with the verification data to evaluate the accuracy of the simulations. RAIN was validated using IMD daily accumulated gridded rainfall data at 0.25° × 0.25° resolution. SAT, SAP and WS10 were verified with the Indian monsoon data assimilation and analysis regional reanalysis data at 0.12° × 0.12° resolution.

Experimental set-up

This study assesses the influence of hybrid assimilation on a parameter-calibrated WRF model on predicting heavy
Table 4. RMSE values of the variables rainfall (RAIN), surface air temperature (SAT), wind speed at 10 m (WS10), and surface air pressure (SAP), averaged over 48 days (12 four-day events) for different experiments.

<table>
<thead>
<tr>
<th>Experiments</th>
<th>RAIN (mm/day)</th>
<th>SAT (K)</th>
<th>WS10 (m/s)</th>
<th>SAP (hPa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEF-NA</td>
<td>28.93</td>
<td>1.72</td>
<td>2.69</td>
<td>874.21</td>
</tr>
<tr>
<td>CAL-NA</td>
<td>26.54 (8.24%)</td>
<td>1.42 (17.54%)</td>
<td>2.21 (17.77%)</td>
<td>825.93 (5.52%)</td>
</tr>
<tr>
<td>DEF-3DV</td>
<td>24.74 (14.47%)</td>
<td>1.86 (-8.60%)</td>
<td>2.16 (19.80%)</td>
<td>827.13 (5.39%)</td>
</tr>
<tr>
<td>DEF-HYB</td>
<td>24.82 (14.21%)</td>
<td>1.68 (2.71%)</td>
<td>2.07 (23.15%)</td>
<td>823.63 (5.79%)</td>
</tr>
<tr>
<td>CAL-3DV</td>
<td>24.09 (16.74%)</td>
<td>1.69 (1.88%)</td>
<td>1.99 (25.99%)</td>
<td>821.85 (5.99%)</td>
</tr>
<tr>
<td>CAL-HYB</td>
<td>23.71 (18.04%)</td>
<td>1.59 (7.91%)</td>
<td>1.95 (27.65%)</td>
<td>822.59 (5.90%)</td>
</tr>
</tbody>
</table>

Values within brackets represent the reduction in value of RMSE for the corresponding experiment with respect to the default experiment.

Figure 5. Comparison of RMSE values of the WRF model variables for DEF-NA and DEF-HYB experiments: a, rainfall; b, surface air temperature; c, wind speed at 10 m; d, surface air pressure. Values above the bars indicate reduction in RMSE of DEF-HYB compared to DEF-NA.

rainfall events during ISM. Six experiments (for each of the 12 events) were performed. Table 3 summarizes the details. The first experiment was performed with default parameters without data assimilation (DEF-NA) to set a benchmark for comparison. The second experiment was performed with calibrated parameters without data assimilation (CAL-NA) to evaluate the influence of calibration on the model output. The third and fourth experiments were performed with
default parameters using 3DVar (DEF_3DV) and 3DEnVar assimilation (DEF_HYB) respectively, to evaluate the influence of assimilation on the default model parameters. The fifth and sixth experiments were performed with calibrated parameters using 3DVar (CAL_3DV) and 3DEnVar assimilation (CAL_HYB) respectively, to evaluate the influence of assimilation on the calibrated model parameters.

For experiments without data assimilation (DEF_NA and CAL_NA), the simulations were started at 00 UTC on the first day of each event and run continuously for 96 h. For experiments with data assimilation, simulations were started 24 h before the beginning of each event. A 6 h spin-up followed by four cycles of assimilation at six-hourly intervals was performed. After four cycles of assimilation, the model was continuously run for 96 h. Root mean square error (RMSE) as defined in eq. (4) was used for evaluating the simulations. It was also used to calibrate the parameters in the earlier study, as the objective was to determine the parameter values that minimize RMSE.

\[
\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (\text{sim}_i^t - \text{obs}_i^t)^2}{N \times T}},
\]

where \(\text{obs}_i^t\) and \(\text{sim}_i^t\) are the observed and simulated values at grid point \(i\) and time \(t\) respectively, \(N\) the total number of grid points in MCR and \(T\) is the number of simulation days.

**Results and discussion**

Table 4 summarizes the results obtained from the experiments. RMSE values of the variables RAIN, SAT, WS10

---

**Figure 6.** Comparison of RMSE values of the WRF model variables for CAL_NA and CAL_HYB experiments: 
- **a.** rainfall;  
- **b.** surface air temperature;  
- **c.** wind speed at 10 m;  
- **d.** surface air pressure. Values above the bars indicate reduction in RMSE of CAL_HYB compared to CAL_NA.
and SAP were evaluated for each experiment and averaged across 12 heavy rainfall events. The percentage values within brackets represent the reduction in RMSE for the corresponding experiment with respect to the DEF_NA. Calibrated parameters performed better than default parameters by reducing RMSE for all variables, viz. RAIN (8.24%), SAT (17.54%), WS10 (17.77%) and SAP (5.52%) without assimilation. Using assimilation on default parameters improved the overall prediction corresponding to both algorithms, with HYB performing similar to or better than 3DV for all variables compared to DEF_NA. The results were similar for DEF_3DV and DEF_HYB for RAIN (≈14%) and SAP (≈5%). However, DEF_HYB performed better than DEF_3DV for WS10 and SAT. Even for calibrated parameters, HYB performed similarly to or better than 3DV for all variables. CAL_HYB performed better than CAL_NA for all variables except SAT.

Figure 5 shows the event-wise RMSE values for DEF_HYB and DEF_NA experiments to assess the impact of data assimilation on the default parameters. These results show a general trend with a substantial reduction in RMSE of all variables for DEF_HYB, except SAT. RMSE values for RAIN were reduced for all events with an average decrease of 14.21% for DEF_HYB compared to DEF_NA. A similar trend was observed for WS10 (23.15%) and SAP (5.79%), where DEF_HYB performed better for almost all the events. However, for SAT, RMSE increased for most of the events. Although the average RMSE (2.71%) had reduced, it was skewed primarily because of event I, where it reduced (43%) from 3.36 K to 1.92 K. Figure 6 shows the event-wise RMSE values for CAL_HYB and CAL_NA experiments to assess the impact of data assimilation on the calibrated parameters. RMSE values of RAIN (10.67%) and WS10 (12.02%) had reduced after assimilation for most of the events. The RMSE of SAP were similar with and without assimilation, with no significant increase or decrease for almost all events after assimilation. However, the RMSE values of SAT increased after assimilation for
most of the events, with an overall increase of −11.68%. Although there was an increase in RMSE for SAT after assimilation, the overall effect of calibration and assimilation had reduced (7.91%) compared to the default parameters without assimilation.

Figure 7 compares the spatial patterns of bias (simulated minus observed) for the DEF NA and CAL HYB experiments for RAIN, SAT, WS10 and SAP, averaged over 48 days (12 four-day events). The bias of RAIN was similar for both DEF NA and CAL HYB in the western region. However, a strong positive bias in the eastern part for DEF NA was replaced by a negative bias for CAL HYB. Although the reduction in overall bias is not evident from this figure, the RMSE values (18.04%) indicate a significant decrease for CAL HYB compared to DEF NA. In the case of SAT, a strong positive bias in the northern region for DEF NA was replaced by a weak negative bias for CAL HYB. Also, a weak positive (negative) bias in the western (eastern) part was replaced by a strong negative bias. Overall, a reduction in RMSE (7.91%) was observed for CAL HYB compared to DEF NA. A strong positive bias in the entire region for DEF NA was replaced by a weak negative or positive bias for CAL HYB, which is consistent with a considerable reduction in RMSE (27.65%) for CAL HYB compared to DEF NA. In the case of SAP, a weak negative bias in DEF NA was replaced by a weak positive bias for CAL HYB in the entire region. The reduction in RMSE was also small (5.90%) for CAL HYB compared to DEF NA. Event-wise spatial comparison, presented in Supplementary Figures 1–12, also shows a general trend of improvement in the spatial pattern of all variables for CAL HYB compared to DEF NA.

Conclusion

The impact of assimilation on the parameter-calibrated WRF model was assessed. Twelve heavy rainfall events were simulated. The calibrated parameters performed better than the default parameters, reducing RMSE for all variables. There was considerable improvement in predicting all variables with assimilation on both the default and calibrated parameters, except for SAT. Also, 3DEnVar expectedly performed better than the 3DVar assimilation method. Overall, hybrid assimilation with calibrated parameters showed a significant improvement for the variables RAIN (18.04%), SAT (7.91%), WS10 (27.65%) and SAP (5.90%) compared to the default parameters without assimilation. A further improvement in prediction could be obtained using different methods. Performing an observing system simulation experiment can help identify observations that do not improve the initial conditions. These observations can be omitted, thereby improving the results from assimilation. Assimilating observations from other satellite instruments and radars could also help improve the predictions. Also, implementing advanced assimilation techniques such as four-dimensional ensemble variational (4DEnVar) hybrid assimilation and particle filter could further improve the model prediction.


ACKNOWLEDGEMENTS. This work was supported by the National Monsoon Mission (Project No. IITM/MM-II/IITM/2018/IND-8) of the Ministry of Earth Sciences (MoES), Government of India. Simulations were performed on the Aaditya high-performance computing system at the Indian Institute of Tropical Meteorology in Pune. Initial and boundary conditions as well as observation data were from the National Centers for Environmental Prediction, USA. We thank the National Centre for Medium Range Weather Forecasting, MoES, GoI for the IMDAA reanalysis.

Received 27 June 2022; revised accepted 24 December 2022

doi: 10.18520/cs/v124/i6/693-703