

RESEARCH ARTICLE

Impact of INSAT-3D land surface temperature assimilation via simplified extended Kalman filter-based land data assimilation system on forecasting of surface fields over India

Abhishek Lodh^{1,2,3}  | Ashish Routray¹ | Devajyoti Dutta¹ | Vivek Singh⁴ | John. P. George¹

¹National Centre for Medium Range Weather Forecasting (NCMRWF), Ministry of Earth Sciences (MoES), Noida, India

²Swedish Meteorological and Hydrological Institute, Norrköping, Sweden

³Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

⁴Indian Institute of Tropical Meteorology (New Delhi Branch), Ministry of Earth Sciences, New Delhi, India

Correspondence

Abhishek Lodh, Researcher, Swedish Meteorological and Hydrological Institute, Folkborgsvägen 17, Norrköping, Sweden; Visiting Fellow, Department of Physical Geography and Ecosystem Science, Lund University, Sölvegatan 12, 223 62 Lund, Sweden.

Email: abhishek.lodh@nateko.lu.se; abhishek.lodh@smhi.se

Ashish Routray, National Centre for Medium Range Weather Forecasting (NCMRWF), Ministry of Earth Sciences (MoES), A-50, Sector-62, Noida, UP-201309, India.

Email: ashishroutray.iitd@gmail.com; ashish.routray@gov.in

Abstract

The land surface temperature (LT) is a crucial variable that governs the energy and radiation budget of the earth's atmosphere and influences land-atmosphere interactions. The LT plays a crucial role mainly in the short-range forecast of a numerical weather prediction (NWP) model. The primary research goal in this research work undertaken is to assess the impact of assimilation of LT data from the Indian satellite (INSAT-3D) into the NCMRWF global NWP model (NCUM) through a simplified Extended Kalman Filter (sEKF) land data assimilation system (LDAS), particularly important as there are few screen-level observations over the region. A dedicated stand-alone pre-processing system has been designed to prepare LT observations in a compatible format for the land surface assimilation system. The approach for LT data assimilation from the INSAT-3D satellite reduces the uncertainty associated with the initial state of LT analysis while simultaneously improving the accuracy of forecasts of near surface atmospheric variables. An observing system experiment (OSE) was carried out during both the summer (May) and winter (February) months by assimilating the INSAT-3D LT data in a coupled land-atmosphere analysis-forecast system. The results obtained from the OSE demonstrate that the use of INSAT-3D LT data improves the forecast skill of both maximum and minimum temperature over India, particularly in areas characterized by higher LT variability. The improvement is pronounced in forecasts of maximum (minimum) temperature during “Boreal” summer (“Boreal”

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Meteorological Applications* published by John Wiley & Sons Ltd on behalf of Royal Meteorological Society.

winter) season. The verification scores also indicate that the incorporation of INSAT LT data substantially improves the NCUM model's forecast performance. By assimilating LT, the mean error of maximum and minimum temperature forecasts in India was decreased, accompanied by enhanced forecast accuracy within a time frame of approximately 24 h. The scores for the verification measures, specifically the Probability of Detection (POD), demonstrate a ~15% improvement in both the forecasts for maximum and minimum temperatures. This improves the temperature prediction as well as the ability to forecast intense weather episodes like cold spells and heat waves.

KEYWORDS

INSAT-3D, land surface temperature, land-atmosphere coupling, observing system experiment, simplified extended Kalman filter

1 | INTRODUCTION

According to the 21st Conference of the Parties (COP21) of the United Nations Framework Convention on Climate Change (UNFCCC), climate change is primarily responsible for increased temperatures, melting of sea ice and ice sheets, changes in precipitation, increase in the intensity, and frequency of extreme weather events such as droughts, heat waves, tropical cyclones, and floods etc. Among the South Asia countries, India is on the forefront of adverse effects of extreme weather events due to its vast geography. Among all these extreme weather events, the focus on accurate forecast of intensity and duration of heat waves has gained momentum in context of global warming scenario (Pandey et al., 2022). India generally experiences severe and long-lasting heat waves during the months of April to early July. The months of April to July also see the highest temperatures over India. During the months of April to June, ridge or anticyclonic circulation forms in the lower-middle tropospheric levels over north-west India and surrounding areas during heat wave events. As a result, there is subsidence of the air leading to rise in surface air temperatures, which advects heat eastwards. Hence, accurate and advance forecasts of heat wave events are necessary for developing effective mitigation strategies. The generation of accurate analysis and forecasts of land surface temperature (LT) is necessary for this task. LT has been considered as one of the high-priority parameters of the International Geosphere-Biosphere Programme (Sharifnezhadazizi et al., 2019). The LT is a crucial land surface variable that has been officially recognized as a major variable from the 55 crucial climatic parameters by the global climatic observing system (GCOS) and the World Meteorological Organisation (WMO), which is important in numerous fields, including hydrology, agriculture, identification of urban

heat islands, research on climate change, ecological and biogeochemical studies, and more. The accessibility of surface variables like LT, soil moisture (SM), soil temperature (ST), and surface heat fluxes from in-situ observations is difficult, costly, and is very limited across the globe particularly over Indian region (Candy et al., 2017; Kumar et al., 2017; Lu et al., 2017). Consequently, significant efforts have been made in recent years to assimilate the land surface data obtained from satellite products, which offers extensive spacio-temporal coverage (Kumar et al., 2022; Sgoff et al., 2020; Singh et al., 2016; Trigo et al., 2008). The geostationary satellites measure with high temporal resolution and are thus able to capture the full diurnal variability of the LT, which is not possible in the case of polar orbit satellites (Li et al., 2012; Singh et al., 2016).

Several studies have demonstrated that assimilating land surface observations (such as 2 m temperature, 2 m humidity, satellite SM, LT, snow depth etc.) into weather models has a positive impact on weather forecasts owing to the influence of land surface processes on the partitioning of energy, mass, and momentum fluxes between the land surface and the atmosphere (Boussetta et al., 2015; Ek & Holtslag, 2004; Halder et al., 2018; Koster et al., 2010; Lodh, 2020; Reichle et al., 2010; Seneviratne et al., 2013). SM, ST, and LT are important land surface variables over Indian subcontinent for the partitioning of the total energy between latent and sensible heat fluxes, which can influence, trigger precipitation (extreme temperatures and heat wave events) over weak synoptic forcings (Lodh et al., 2019, 2022; Routray et al., 2023). Thus, accurate analyses and realistic initialization of land surface variables (SM, snow cover, and LT) hold significance for short- to medium-range weather forecasts of land surface and atmospheric variables (Bakketun et al., 2023; Benjamin et al., 2022; de Rosnay

et al., 2011, 2014; Lodh et al., 2022; Sgoff et al., 2020). The LT's initialization has important implications for critical processes like turbulent mixing in planetary boundary layer (PBL), where the frictional forces generated from vegetation and topography produces eddies. As a result, LT data assimilation producing lower boundary analyses, which is important for the numerical model's ability to simulate the dynamics between the surface parameters viz. SM, LT, and PBL, resulting in more accurate forecasts near surface variables and weather patterns (Koster & Suarez, 2003).

The difference between the LT (referred to as surface temperature which is temperature at top few molecules of the Earth's surface in contact with the soil/ground) and the near surface temperature (typically measured at a height of 2m from surface) is of paramount importance in regulating the transfer of moisture, heat, and momentum fluxes between the land and the atmospheric boundary layer (Stull, 2017). The LT is mainly a product of albedo, vegetation cover, and SM, whereas ST is the temperature of the first few centimeters of the soil approximately 0–7 cm, where surface is zero centimeter. In a land surface model, ST is set at the middle of each layer, and heat transfer is calculated at the interfaces between them. Moreover, skin temperature represents the temperature of the uppermost infinitesimally thin surface layer (land or water) at radiative equilibrium, which has no heat capacity and can respond instantaneously to changes in surface fluxes. It is a theoretical temperature that is required to satisfy the surface energy balance and is calculated differently over land and sea (Muñoz-Sabater et al., 2021). In a land surface model, 2m temperature (temperature of air at 2m above surface of land, sea or in-land waters) is calculated by interpolating between the lowest model level and the Earth's surface, taking account of the atmospheric conditions (ECMWF, 2024). Despite the complex interaction between the top of the surface layer and bottom of the atmospheric boundary layer where turbulent mixing is important, the processes between these two layers are intricately intertwined. Thus, LT plays an important role in modulating the exchange of energy and moisture flux within the atmospheric boundary layer (Aires et al., 2004; Sgoff et al., 2020; Singh et al., 2016). Hence, the primary sources of LT are factors that are in close proximity to the earth's surface, such as the communication among SM, ST, atmospheric temperature above the surface, incoming radiation, and vegetation cycle principally leaf area index and surface albedo. The data assimilation of these land surface variables viz. SM, ST, leaf area index, and surface albedo also exhibit a positive influence on forecast of boundary layer humidity, temperature, and convective scale precipitation (Candy, 2014; Capecchi &

Brocca, 2014; Dharssi et al., 2011, 2015; Lodh et al., 2022). The interplay between these energy and moisture fluxes and land surface variables exerts a substantial impact on atmospheric circulation patterns, subsequently leading to significant repercussions on weather and climate conditions at both regional and global scales. Therefore, it is imperative to accurately prescribe the initial and boundary conditions, specifically LT and SM, within a numerical weather prediction (NWP) model. In recent years, there have been several endeavours to assimilate LT satellite retrievals into a land surface scheme coupled with an NWP model. The improvement of forecast skill and enhancement of the model's forecasts can be achieved through the utilization of ensemble or simplified Extended Kalman Filter (sEKF)-based Land Data Assimilation (LDA) methods (Aires et al., 2004; Han et al., 2013; Liu et al., 2013; Singh et al., 2016). Studies like Huang et al., 2008; Sgoff et al., 2020, assimilated leaf area index (LAI), and synthetic satellite LT retrievals into a land-atmosphere NWP model, which resulted in an enhanced representation of ST profiles. All the aforementioned studies provide convincing evidence that the incorporation of LT is of utmost importance in improving the atmospheric boundary layer processes and has a positive influence on forecasting of surface variables like SM, ST, LT, and 2 m temperature (Candy et al., 2017; Ghent et al., 2010; Han et al., 2013; Reichle et al., 2010). Simultaneously, it exerts an influence on atmospheric variables via transfer mechanisms by enhancing the soil conditions in the NWP model. The initialization of atmospheric NWP models with accurate land surface state coupled with land surface/vegetation models is a key step towards developing a coupled LDA system capable of exploiting current and future satellite observations.

The LT retrievals from the Indian satellite, INSAT-3D imager are continuously derived over the Indian subcontinent. However, there have been a few studies conducted to evaluate the assimilation of INSAT-3D derived LT (L2) data over the Indian region using a sophisticated Land Data Assimilation System (LDAS). Singh et al. (2016) conducted a study in which they included the LT data obtained from INSAT-3D into the Weather Research and Forecasting (WRFv3.4) mesoscale model. This was achieved by utilizing a 3D variational (3DVAR) technique. The major limitation of the cited study is that the LT observation assimilated as an atmospheric variable (2 m surface air temperature) through 3DVAR assimilation technique, which is not an appropriate methodology to assimilate the LT observation without using of LDAS. This is because we need a land surface model to properly represent the exchange of fluxes between the surface top and near surface atmosphere that influences the

atmospheric processes within the boundary layer. The efficacy of the land surface model, in conjunction with the LDAS, in accurately representing land surface processes and facilitating land-atmosphere feedback widely acknowledged in the studies (Balsamo et al., 2007; Duerinckx et al., 2015; Hamdi et al., 2014; Lodh et al., 2022; Parada & Liang, 2008; Routray et al., 2023). Nevertheless, despite the limitation, the findings of Balsamo et al., 2007, Singh et al., 2016 suggest that incorporating LT observations into the modelling system yields a favourable effect on the weather forecast, regardless of whether the LT observation is utilized as a substitute for air temperature in the 3DVAR technique.

Land data assimilation is a cross-disciplinary science to bring together computer simulations and real-world surface data in a synergistic manner. The process of land DA involves integrating a priori knowledge of the state, often derived from the short forecast of model, with land surface observations and their corresponding error characterizations, to ascertain the most accurate estimation of the lower boundary condition for the model state. In recent times, there has been a growing adaptation of Kalman filter-based assimilation systems in LDAS coupled with NWP models (Fisher et al., 2005). Kalman filters are used in LDA for their ability to optimally combine model background state with real-time observations (even when measurements are noisy and unavailable), handle uncertainties, non-linear processes, and sequentially update the state of the land model. Thus, Kalman filter-based optimal and sequential LDAS approach aims to incorporate surface observations and atmospheric forcings leading to significant improvements in model analyses and forecasts (Candy et al., 2017; Lodh et al., 2022; Routray et al., 2023; Sgoff et al., 2020). Hence, it is imperative to incorporate the LT data obtained from INSAT-3D into a LDAS that utilizes advanced techniques to enhance the predictive accuracy or forecast performance of the model in simulating weather phenomenon.

The main objective of the present study is to explore and evaluate the impact of assimilation of LT retrievals obtained from INSAT-3D in the sEKF-based LDAS on weather forecasts (related to maximum and minimum temperature) over Indian region. The present research work is a first study to assess the impact of assimilation of LT retrievals from INSAT-3D through sEKF-based LDAS on simulation of extreme weather events over Indian region, where few screen level observations are available. National Centre for Medium Range Weather Forecasting (NCMRWF), India, has adopted the sEKF-based LDAS, from the UK Met Office for SM analysis (Lodh et al., 2016, 2022; Routray et al., 2023). This LDAS is now considered for assimilating INSAT-3D LT satellite observations along with ASCAT SM data into the

global (referred to as NCUM) forecast model. Accurate LT initialization plays a major role in the predictive capacity of any NWP system due to its substantial impact on the exchanges of sensible and latent heat flux through the land surface model (Candy et al., 2017; Lodh et al., 2019). It is important to mention here that LT has a very short “memory” or limited heat capacity in most land surface models, so it has a very minimal impact on NWP forecasts. However, in this research work, the LT observations are assimilated along with the being used to update the states of the SM, ST, and skin temperature. Hence, the forecast changes caused by assimilating LT will come from the analysis updates to the SM and ST, not from the LT itself. This forms an important motivation for the study. This research paper will outline our attempts to assimilate the LT data obtained from INSAT-3D satellite and assess its efficacy in predicting maximum and minimum near surface temperatures across India. To fulfil this main objective, we have two sub-objectives in this study. Initially, a pre-processing procedure was developed to prepare the INSAT-3D LT observations (quality controlled) in a gridded format for assimilation. Secondly, this study aims to examine the impact of integrating pre-processed INSAT-3D LT measurements into the global model using a sEKF-based LDAS for a long period by studying various weather patterns over Indian region. The paper is arranged as follows. Section 2 provides a comprehensive account of the INSAT-3D LT and its associated pre-processing procedure. Section 3 documents the sEKF-based LDA method. Section 4 presents the results and discussion part, and lastly, Section 5 states the final conclusions of the present research along with potential avenues for further investigation.

2 | INSAT-3D LAND SURFACE TEMPERATURE

The INSAT-3D satellite was deployed in 2013 from French Guyana, with an aim to monitor the earth's surface and measure vertical profiles of atmospheric temperature and humidity (Mitra, 2014). The INSAT-3D satellite (orbital position is situated at 82° E) has three payloads: Meteorological – Imager and Sounder, Data relay Transponder, and satellite-aided search and Rescue. The meteorological payload, INSAT-3D, imager provides imaging capability of the earth's disc at six channels in 0.65, 1.7, 3.9, 6.8, 10.8, and 11.9 μm bands (Singh et al., 2016) with a telescope aperture of 310 mm diameter. The imager's 10.8 μm (TIR1) and 11.9 μm channels (TIR2) are situated in the infrared window, having a ground resolution of 4 km x 4 km at the sub-satellite point. Here, TIR stands for Thermal Infrared Retrievals.

The INSAT-3D satellite is equipped with a very precise radiometer, improved spatial resolution of 4 km (Singh et al., 2016), and an increased number of spectral channels in the split-window mode compared with its predecessors, namely INSAT-3A and Kalpana-1. The INSAT-3D imager offers enhanced spatial resolution for meteorological observations at 30-minute intervals and covers six different wavelengths (Singh et al., 2016). The imager monitors land and ocean conditions for weather forecasting and disaster warning systems from a geostationary altitude. Additionally, there is a solitary channel in the visible region. The retrieval of LT measurements over Indian region is done by utilizing the INSAT-3D Imager TIR satellite images and employing a split-window algorithm (Pandya, Shah, Trivedi, & Panigrahy, 2011; Pandya, Shah, Trivedi, Panigrahy, Parihar, & Kirankumar, 2011) utilizing the 10.8 and 11.9 μm imager window channels for the retrieval process. Satellite-based TIR data is directly linked to the LT through the radiative transfer equation (Li et al., 2013). The superiority of TIR retrievals over microwave retrievals for LT is acknowledged in the literature (Singh et al., 2016). A physics-based modified split-window algorithm incorporating the view zenith angle dependency is developed to retrieve LT at Space Applications Centre, ISRO, Ahmadabad, India. The split-window algorithm uses sensor-specific split-window coefficients simulated by a radiative transfer model-MODTRAN.

Although surface air temperature data are regularly collected by national meteorological centers around the world, dense in-situ observations of LT are not widely available over India. Therefore, it is not feasible to initialize high-resolution NWP models with sparsely distributed in-situ LT data, with simple interpolation approaches to the model grid point. Remote sensing techniques provide potential for land surface measurements. The LT observations are useful input to various models related to soil hydrology, thermodynamics, agrometeorology, ecology, and boundary layer atmospheric processes. Thus, the LT observation from INSAT-3D satellite is crucial for research in NWP and land surface processes over India. Furthermore, enhancing the prediction capability of the NWP models relies on better physics along with better observations for both land surface and atmospheric analysis. Hence, in order to reduce the inaccuracies in LT analysis and forecasts, it is essential to incorporate the assimilation of LT observations from satellites over India. These satellite data can be readily accessed from the Meteorological and Oceanographic Satellite Data Archival Centre (MOSDAC), which is an online data repository managed by the Space Application Centre, Ahmadabad, Indian Space Research Organisation (ISRO), India. The INSAT-3D LT data can be downloaded from: [https://](https://www.mosdac.gov.in/)

www.mosdac.gov.in/. Thus, the scarcity of LT and ST observations for assimilation over India is solved by the satellite-retrieved INSAT-3D LT product from ISRO. Following the sEKF process of LT LDA, the NWP model will adjust its ST with respect to LT through Jacobians, resulting in a reduction of their differences with the observed surface temperature.

3 | REGIONAL LDAS AND UTILIZATION OF LTs

3.1 | Overview of simplified Extended Kalman Filter (sEKF)-based LDAS

For the initialization of the forecast model (NCUM), NCMRWF has employed real-time atmospheric analyses (6 hourly intervals) based on the hybrid 4DVAR technique and SM analyses based on sEKF algorithm-based LDAS (Lodh et al., 2016, 2022; Routray et al., 2023). Several prominent NWP research centers, such as ECMWF, Meteo France, and UK Met Office, have employed similar sEKF approach-based LDAS to assimilate surface synoptic observations (SYNOP) of screen level temperature and humidity along with satellite-retrieved observations of SM (ASCAT) to generate a precise initial state or analysis of land surface variables (Albergel et al., 2017; Candy et al., 2017; de Rosnay et al., 2013, 2014, 2022; Drusch et al., 2009; Duerinckx et al., 2015, 2017; Fairbairn et al., 2017; Gómez et al., 2020; Lodh et al., 2022; Mahfouf et al., 2009). The revised surface Hydrology Tiled ECMWF Scheme for Surface Exchanges over Land (H-TESSEL) land surface scheme implemented at ECMWF incorporates soil hydrology and land surface dynamics (Balsamo et al., 2009). The land component of the Met Office NWP system is based on the JULES land surface model (Best et al., 2011; Clark et al., 2011). It divides land use into four non-vegetated types (urban areas, inland water, bare soil, and land ice) and five vegetation types (broadleaf trees, needle leaf trees, temperate (C3) grass, tropical (C4) grass, and shrubs). The International Geosphere-Biosphere Programme (IGBP) provides a global land use data set from which the fraction of each of these kinds at each land grid point is mapped in NCUM. In a similar fashion to UK Met Office at NCMRWF, the offline LDAS based on sEKF and Joint UK Land Environment Simulator (JULES) land surface model has been implemented over the Indian region (Lodh et al., 2016, 2019, 2022; Routray et al., 2023) for the NCUM model for generation of SM and LT analysis. The sEKF-based SM or LT analysis generated is a point-wise one-dimensional data assimilation scheme (Albergel et al., 2017; Duerinckx et al., 2015; Fairbairn et al., 2017;

Gómez et al., 2020; Mahfouf et al., 2009; Parrens et al., 2014). Initial values of the SM background error in sEKF-based LDAS shows similar sized analysis increments at the top-most soil level and larger increments at deeper layers, when compared to the nudging scheme and/or optimal interpolation (Lodh et al., 2022). The strategy implemented in this study for the generation of LT analysis is a one-dimensional system that functions independently at each soil column of the JULES land surface model, utilizing a weakly coupled sEKF-based data assimilation technique between the land and atmosphere (Gómez et al., 2020). The soil column inside the JULES land surface model is partitioned into four layers, each characterized by thicknesses of 0.1, 0.25, 0.65, and 2 metres, respectively, starting from the surface and extending to the bottom (i.e. 3 m depth in total), and SM is calculated based on Richard's equation. The updated equation of SM content in each layer is given by water balance equation. The vertical fluxes of water within the soil layers follows Darcy's law, which relates the water flux to the hydraulic conductivity and gradient of water content with depth or soil water suction (Best et al., 2011). The soil processes that control the hydraulic linkages between soil conductivity and soil suction are explained by the Van Genuchten equations (Van Genuchten, 1980; Walters et al., 2019). Soil texture map is from three separate sources, Harmonised World Soil Database, State soil geographic database and point observations of soil sand, silt and clay fractions (Dharssi et al., 2011). The soil-thermodynamics is illustrated through the exchange of heat through diffusion among the soil layers, as well as the advection of heat between these layers via moisture fluxes (Dharssi et al., 2011, 2015). The analysis increment in LT or SM is applied (Bouttier & Courtier, 2002) to each layer of the soil by using the following mathematical equations:

$$x^a = x^b + \mathbf{K} (y - \mathbf{H}(x^b)), \quad (1)$$

$$\Delta x = \mathbf{K} (y - \mathbf{H}(x^b)).$$

The equation (1) says that the analysis (LT or SM) is obtained by adding to the first-guess (background) the innovation multiplied by $\mathbf{K}$, Kalman gain matrix (Kalnay, 2002). The symbols used in equation (1) denotes: x^a as analysis vector that contains four layers of SM, ST at each layer, and LT. Further, y represents the observation vector (for screen level humidity, screen level temperature, SM observation, and LT observation). The x^b denotes the background fields (model surface temperature, ST, SM) are derived from the short forecast of the previous cycle of NCUM model valid at analysis time.

The atmospheric forcing data for the offline JULES land surface model ($\mathbf{M}$) used (precipitation, surface longwave wave and short-wave radiation, air temperature, humidity, wind speed, and surface pressure) are obtained from short-range forecasts of the atmospheric model with the assimilation window. Atmospheric forcing of air temperature and humidity are applied at a height higher than the screen level (10 m). This allows the sEKF LDAS to also assimilate observations of screen level temperature and humidity. $\mathbf{H}$ represents the non-linear observation operator which performs the conversion from model space to observation space and vice versa. $\mathbf{H}$ does not evolve in time and is ideally a diagnostic block of the land surface model, whereas $\mathbf{H}(x^b)$ represents model equivalent in observation space and is output from the unperturbed JULES forecast (Dharssi, 2010). The $\mathbf{K}$ Kalman gain matrix (Equation 2) specifies the contribution of each observation to update the model parameters and is computed from the observation and background error covariances and the Jacobians (partial differentials or sensitivities).

$$\mathbf{K} = \mathbf{B}\mathbf{H}^T (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}, \quad (2)$$

where $\mathbf{B}$ and $\mathbf{R}$ denote the background and observation error covariance matrices characterizing the forecast and observation vectors. The observation operator in the sEKF system, denoted as $\mathbf{H}$ (and its transpose $\mathbf{H}^T$), is the Jacobian matrix: linearized version of the observation operator H . $\mathbf{H}$ is obtained using a finite difference approach, individually perturbing each component x_j of the initial model state x at time t_0 by a small amount δx_j (perturbation). The Jacobian ($\mathbf{H}$) is expressed as the difference of the prognostic states between perturbed and control run at the end of the assimilation interval at time t_i , divided by the initial perturbation:

$$\mathbf{H} = \frac{H(x + \delta x_j) - H(x)}{\delta x_j}. \quad (3)$$

Thus, linearizing the observation operator is accomplished through linearizing the model around the non-linear trajectory between two consecutive time steps and is obtained by performing a Taylor series expansion of H . The calculation of $\mathbf{H}$ involves estimating of its Jacobians ($\mathbf{H}$) by perturbation forecasts of an offline JULES model run that transforms the model states to the observation space. Therefore, in simple terms, the $\mathbf{H}$ matrix establishes the connection between the observations (in our study INSAT-3D LT) and the analysis vector. In other words, $\mathbf{H}$ is an important parameter that provides a

linkage between the model and observed variables (e.g. SM, LT) that we want to estimate within the soil layers. Thus, the Jacobians are first order partial derivatives that describe the link between the observations and the analyses variables and are computed via finite differences of JULES perturbation runs. Jacobians of ST (computed via Jules perturbation run) link the observations and model analysis vector of ST (all four layers), skin temperature with the assimilation window. The LT data assimilation using the sEKF is conducted within every analysis cycle to ensure that the observations update the background (i.e. prevailing land surface condition) at every grid point. In an ideal Extended Kalman Filter (EKF)-based LDAS, the system incorporates and advances the land model's background (**B**) and forecast error covariances through perturbation runs, by utilizing the offline land surface scheme. The equation for evolving **B** background error covariance matrix at the initial state is given by:

$$\mathbf{B} = \mathbf{M} \cdot \mathbf{A} \cdot \mathbf{M}^T + \mathbf{Q}, \quad (4)$$

where **B** is matrix of error covariances at initial state (flow-dependent in ideal EKF), **M** is the model, non-linear operator which evolves the state vector in time. This is generally the prognostic block of the land surface model. The Jacobians of the tangent linear of JULES (**M**) is calculated using finite difference.

$$\mathbf{M} = \frac{\mathbf{M}(\mathbf{x} + \delta \mathbf{x}_j) - \mathbf{M}(\mathbf{x})}{\delta \mathbf{x}_j}, \quad (5)$$

$$\mathbf{A} = (\mathbf{I} - \mathbf{K} \cdot \mathbf{H}) \cdot \mathbf{B}, \quad (6)$$

where **M**(**x**) is a JULES unperturbed forecast using initial state **x** and **M**(**x** + $\delta \mathbf{x}_j$) is a JULES forecast from the relevant perturbed run. **A** is the matrix of the analysis error covariances, **I** is unit matrix, and **Q** is the matrix of the internal model error covariances. In contrast to the EKF, the static **B** and the **R** matrixes are a distinguishing characteristic of the sEKF, as calculation of **B** matrix is unstable (Kourzeneva, 2014). In sEKF, the diagonal components of **B** matrix consist of error covariances of the initial state and **R** matrix consist of observational error variances (de Rosnay et al., 2013, 2014; Mahfouf et al., 2009; Mahfouf et al., 2009). Also, it is not clear how to specify the model error matrix **Q** (see equation 4) in the sEKF and thus not easy to cycle the background error covariance matrix **B**. The utilization of a static **B** matrix commonly refers to simplified EKF, that is, sEKF algorithm-based LDAS. But the cost of simplification is improper representation of physics. There is almost no

response of $T_{2m}(t)$ to $T_s(0)$ due to short memory of temperature, but strong response of $T_{2m}(t)$ to $T_s(t)$ [refer to Figure 1a]. The JULES model temperature at 2 m is calculated using similarity functions (Walters et al., 2019). For each grid-point analysis and typical assimilation cycle in the CTL and INSAT experiments, the analysis vector ($\mathbf{x}^a$) contains the following soil field:

$$\mathbf{x}^a = [\text{SM from level1 to level4, ST from level1 to level4, LT}].$$

So, the size of analysis vector is nine. The observations screen level temperature, humidity, ASCAT SM, and INSAT LT are utilized in the assimilation cycle. Thus, the dimension of observation vector is four. It is commonly assumed that observations are uncorrelated in error, specifically at every grid point. As a result, the **R** matrix, which represents the covariance structure of the observations, is considered to be diagonal. The **B** matrix is constructed from background errors. The advantages of the sEKF over other LDA methods like optimum interpolation are the dynamically calculated gain coefficients, which makes it easier to include new observation types. In our study, the perturbations for the sEKF are created by offline JULESv5.1 model runs. The number of perturbed model runs is equal to the number of analysis variables. In the sEKF, analysis of SM and ST using the JULES model at 4 soil layers requires a total of 9 ensemble integrations of the offline JULES model (perturbed runs) along with an unperturbed run to calculate the Jacobians of the observation operator. Each element within the set **H** holds the sensitivity of the observation with respect to the analysis vector (Gómez et al., 2020). The atmospheric forcing data necessary for conducting the offline JULES model simulation in the LDAS are acquired from the NCUM short-term forecast. It is important to mention that the sEKF algorithm works on independent land columns of the JULES model in the vertical with uncorrelated errors (Duerinckx et al., 2017). The studies conducted by Lodh et al. (2022) and Routray et al. (2023) provide a comprehensive description of the individual components of the Jacobian matrix and the assimilation process of the SM data obtained from ASCAT. Table 1 presents the observation and background errors as well as the perturbation values employed in the sEKF. Those parameters are utilized to create **B** and **R** that are supposed to be diagonal matrices.

3.2 | Inclusion of INSAT-3D LTs

LT is an important variable for land surface processes and weather prediction. The pre-processing of INSAT-3D LT data is a crucial step in its use within the sEKF-based

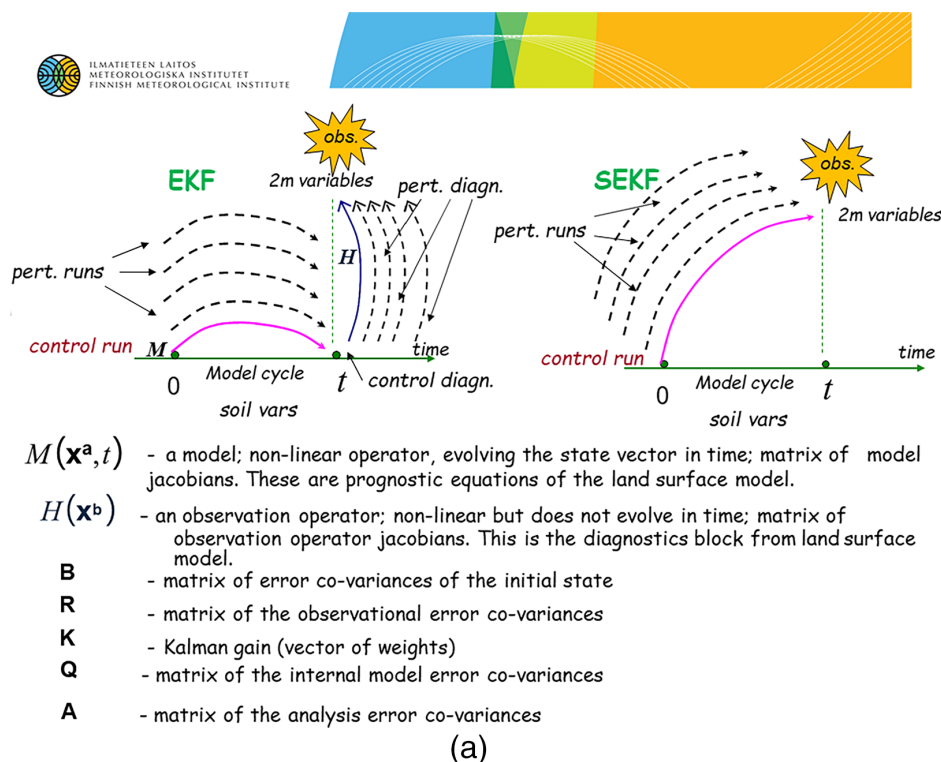
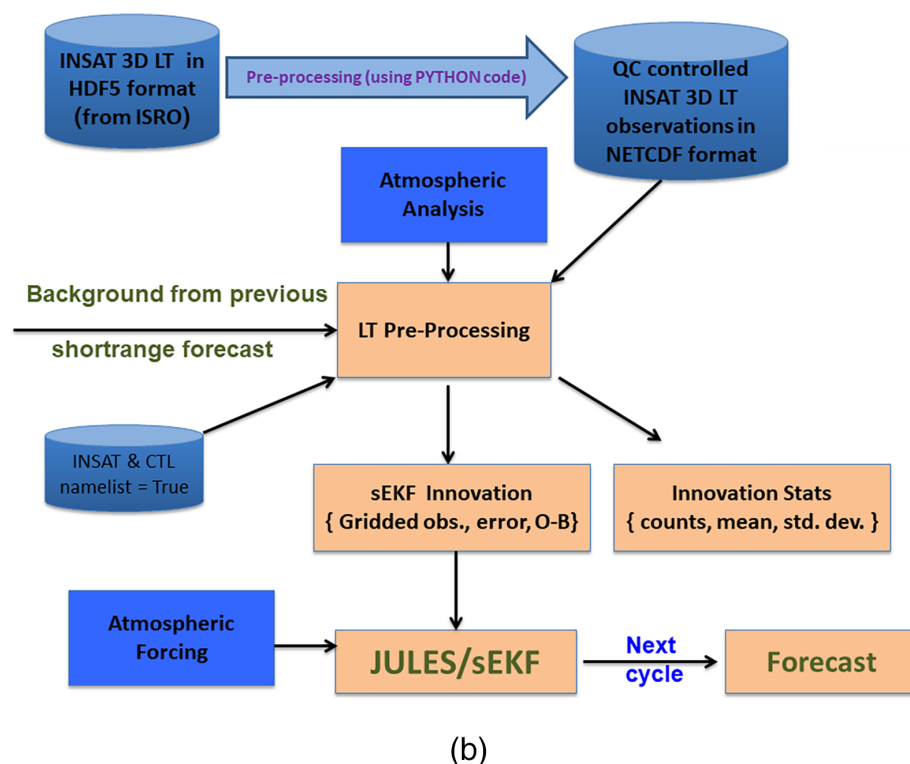


FIGURE 1 (a) Diagram showing difference between Extended Kalman Filter (EKF) and simplified Extended Kalman Filter (sEKF). (b) Flowchart of the INSAT LT data assimilation after pre-processing.



LDAS framework, as discussed in Lodh et al. (2019). The efficacy of LT in LDAS research is constrained by the short memory, often on the range of very few minutes, attributable to its relatively low heat retention capacity (Houser et al., 2010). Level-2 LT observations from INSAT-3D are available at 30-min intervals from

the ISRO's Meteorological and Oceanographic Satellite Data Archival Centre (MOSDAC) repository at <https://www.mosdac.gov.in> (Singh et al., 2016). These data are accessible in the HDF5 (Hierarchical Data Format) and gif format. To facilitate compatibility with the surface processing system (SURF) programmes of the LDAS, a

TABLE 1 Observation Error (OE), Background Error (BE), and Perturbations used in sEKF LDAS for INSAT-3D LT assimilation.

Parameters	Value
BE: Soil moisture level 1	0.04 m ³ m ⁻³
BE: Soil moisture levels 2–4	0.026 m ³ m ⁻³
BE: Soil temperature	2.0 K
BE: Skin temperature	2.0 K
OE: Screen humidity	0.0009 g/g
OE: Screen temperature	2.0 K
OE: ASCAT soil moisture	0.035 m ³ m ⁻³
OE: Near surface temperature (INSAT LT)	2.0 K
OE: Screen relative humidity	8%
Perturbation: Soil moisture	−0.005 m ³ m ⁻³
Perturbation: Soil temperature	−0.05 K
Perturbation: Skin temperature	−0.05 K

stand-alone proprietary Python software has been developed at NCMRWF to convert the LT data from HDF5 to NetCDF format. This conversion step ensures compatibility between the satellite observations and the processing system. The spatial dimensions of the INSAT-3D LT data are represented by the GeoX and GeoY values. The GeoX values span from 1000 to 1800, whereas the GeoY values vary from 530 to 1350. The pixel (GeoX) and scan (GeoY) directions for the TIR₁ and TIR₂ channels of the L2B products are determined by these values. A Python utility program is written to process the INSAT-3D LT data, including extraction of the GeoX and GeoY values mentioned above. The detail about the pre-processing technique applied for raw LT data obtained from the MOSDAC, ISRO, is described in the technical report published by Lodh et al. (2019). The aim of the LT pre-processing is to create a quality-controlled (QC) gridded data set covering the Indian landmass. It is essential to read the LT observations available in the assimilation window (± 30 min centered on analysis time) as well as subsequent LT from model forecast and cloud cover (i.e. available 1hrly frequency). This is because LT has a very short memory. The LT observations need to pass through various types of QC checks (Figure S1) and be discarded (Candy, 2013; Candy et al., 2017) if any of the following conditions are true:

- Data are placed over ocean, sea, and mountain.
- Reject observations if model cloud fraction is more than 0.2.
- Snow cover stacked at observation location. The threshold value for snow not exceeding 0.05 kg m⁻².
- Maximum allowed orography is 2000 m, and LT values are within the range of 250–350 K.

- Gross error of observation excess of 50% by comparing the model forecast LT using the Bayesian theory (Lorenc & Hammon, 1988) that test removes the cloud contaminated observations.

Figure 1b illustrates the workflow of the sEKF-based LDAS, demonstrating the assimilation of INSAT-3D LT within the global NCUM modelling system. Computing the Jacobians provides the link between the INSAT-3D LT observations and the soil levels of the JULES model. After this, the “sEKF pre-processor” will generate the quality-controlled gridded LT (O-B) innovations on the model grid (See equation 1). To compute the innovation, the background LT field consistent to the observation time is used (the time interval is 30 minutes within the assimilation cycle). After pre-processing of LT, gridded observation, innovation, and other statistics are produced, which is an input to the next step for sEKF-based LDAS. In Figure 2a,b, show the typical distribution of INSAT-3D LT observations, valid at 12UTC on May 2, 2018, and February 2, 2019, respectively, which are used in the LDAS assimilation cycle. Figure 2c,d shows the zoomed LT observations over the Indian domain. The INSAT-3D LT observations are extracted and assimilated within the LDAS along with the model's background fields, within the assimilation window (± 30 min). Subsequently, QC tests are applied to the satellite LT observations and converted into a gridded format compatible with the model grid, capturing the average LT value per grid point, LT innovation, and LT-associated uncertainty. The LT uncertainty for INSAT-3D is fixed at 3.5 K. The observation error is 2 K, used in Bayesian calculation to run the LT pre-processing algorithm (Lodh et al., 2019). The gridded LT observations, obtained after undergoing pre-processing and QC checks are used as surface initial condition for ST for each soil layer and skin temperature. Figure 3a,b represents the mean LT analysis and analysis increment (Analysis-Background, i.e. A-B) in a gridded format after LDAS from a typical DA cycle 12UTC, 20180501. Figure S2a,b shows the mean of ASCAT SM observations and SM innovations in gridded format after pre-processing of ASCAT observations, for a typical data assimilation cycle of INSAT experiment. Figure S3a,b shows the mean of INSAT-3D surface temperature observations and LT innovations in gridded format over the global domain, after pre-processing of INSAT-3D observations, for a typical data assimilation cycle of INSAT experiment. Figure 3c shows the total number of observations assimilated before and after QC in the pre-processing of observations (± 30 min assimilation window) during the period 00UTC 20180501 to 18UTC 20180507. It is essential to mention here, in CTL experiment, the gridded SM observations (from ASCAT) are used in the sEKF-based LDAS to obtain SM initial condition.

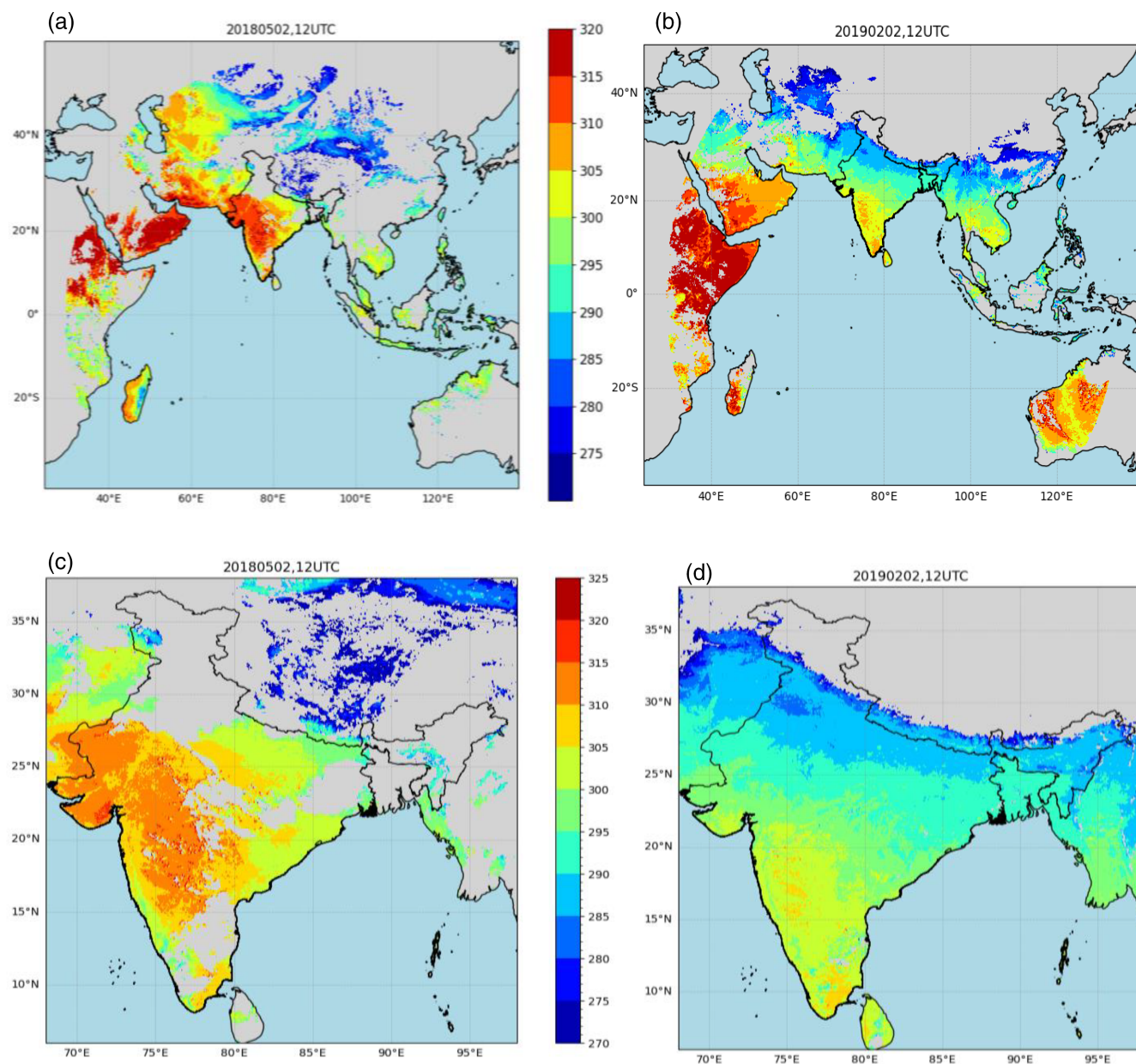


FIGURE 2 (a) LT (K) observations from INSAT-3D and (c) mean observation zoomed over the Indian domain used in the data assimilation (DA) cycle for a typical date valid at 12 UTC 20180502. (b–d) same as (a–c) but valid at 12 UTC, 20190202.

For the INSAT experiment, in addition to INSAT-3D LT, the LDAS also incorporates the ASCAT (MetOp-A and B) SM product along with surface synoptic observations (SYNOP) of screen level temperature and humidity. The spatial distribution of in-situ surface synoptic observations exhibits non-uniformity across India. Therefore, a simplified variational analysis scheme, that is, 3DVAR is utilized to generate a two-dimensional spatial dataset of surface-level observations at every grid point, prior to its utilization in the sEKF. It is crucial to note that the land surface analysis operations conducted within LDAS (surface processing system is acronymed “SURF” at NCMRWF) are carried out independently of the atmospheric upper-air

analysis using Hybrid-4DVAR, constituting the weakly coupled land-atmospheric data assimilation system.

4 | NCM ANALYSIS-FORECAST SYSTEM AND NUMERICAL EXPERIMENTS

The unified model is an NWP and climate modelling system developed by the UK Met Office (Brown et al., 2012; Hurrell et al., 2009). Since 2012, NCMRWF has adopted and operationalized the unified model (NCUM) from the UK Met Office (Rajagopal et al., 2012). The NCUM model framework provides a flexible configuration that can be

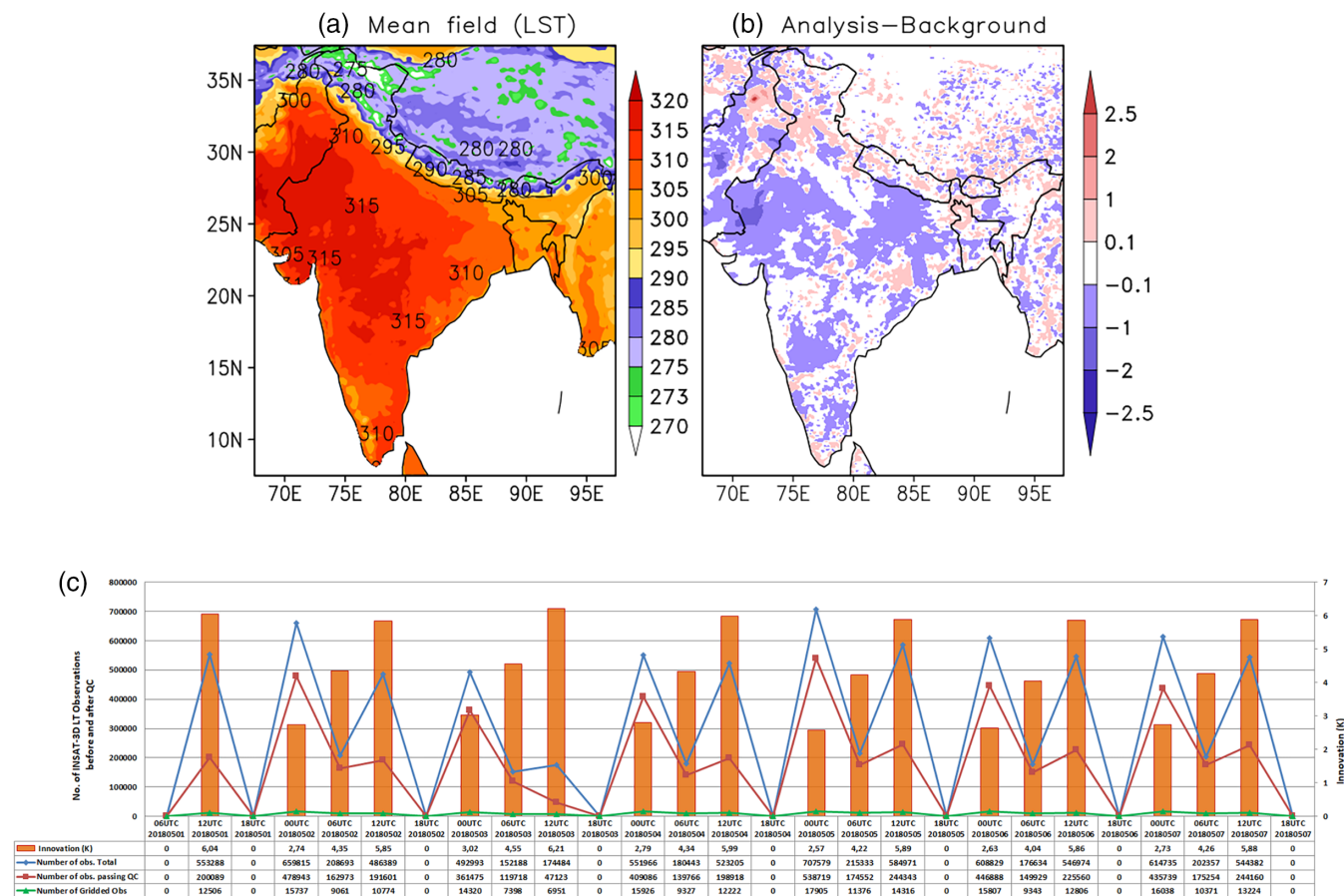


FIGURE 3 (a) Mean of LT (model units: K) analysis over Indian domain after QC and (b) “Analysis – background” after SURF pre-processing of observations, for a typical DA cycle 12UTC, 20180501. (c) LT innovations (i.e. observation – background) and total number of observations assimilated before and after QC in the SURF pre-processing of observations during the period 00UTC 20180501 to 18UTC 20180507 (± 30 minutes assimilation window).

used for various purposes, including global and regional weather prediction, seasonal forecasts, and climate studies. The model has configurations that can be run at different resolutions and for various domains, depending on the forecast requirements (Routray et al., 2019). The Even Newer Dynamics for General Atmospheric Modelling of the Environment (ENDGame) functions as the dynamic core of the NCUM modelling system. The NCUM model runs globally, with ~ 12 km horizontal resolution (N1024) and 70 vertical levels extending up to a height of 80 km. The global model (NCUM) utilizes a hybrid height coordinate system that follows the terrain and is non-hydrostatic in nature for the vertical dimension, which is used to generate medium-range forecasts lasting up to 10 days. Additionally, it employs an Arakawa-C grid staggering method for the horizontal dimension. The investigation of dynamics of the deep atmosphere is conducted through the utilization of a semi-implicit, semi-Lagrangian, predictor-corrector numerical technique (Cullen et al. in 1997 and further developed by Davies et al. in 2005). The land surface scheme employed in this study is JULES (Best et al., 2011; Clark et al., 2011), in conjunction with

various physical parameterization methods for cloud microphysics, radiation, and boundary layer physics etc. The details about NCUM forecast modelling system is discussed in the works of Davies et al. (2005) and Routray et al. (2019). The Hybrid-4DVAR DA technique is used to prepare the atmospheric analysis, whereas surface boundary conditions mainly SM and ST analyses prepared through the sEKF-based LDAS algorithm (Lodh et al., 2016, 2022; Routray et al., 2023). To assess and comprehend the effects of using INSAT-3D LT data through sEKF-based LDAS on numerical weather forecast, two sets of assimilation-forecast experiments are conducted here. Here, two different sets of experiments are conducted with NCUM model (N1024), which corresponds to approximately 12 km.

- CTL experiment: The assimilation of satellite-derived SM observations from ASCAT, along with screen level observations of 2 m temperature and 2 m humidity, is performed using a sEKF-based LDAS.
- INSAT experiment: Assimilation of INSAT-3D LT observations along with the other land surface

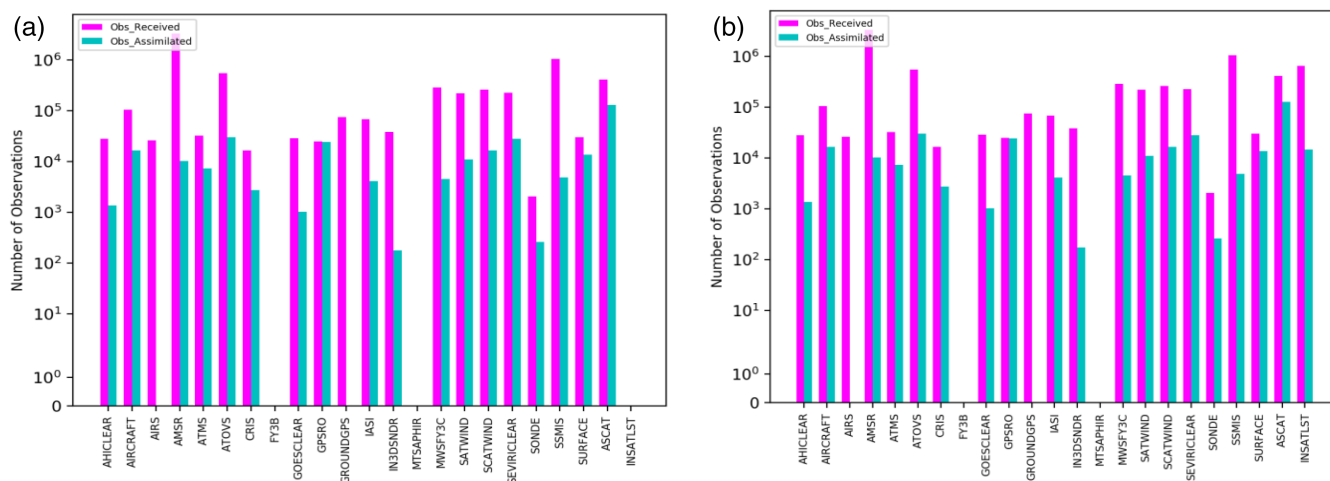


FIGURE 4 Bar plot of the various observational data included in (a) CTL and (b) INSAT LT assimilation for a typical DA cycle 12UTC 20180502.

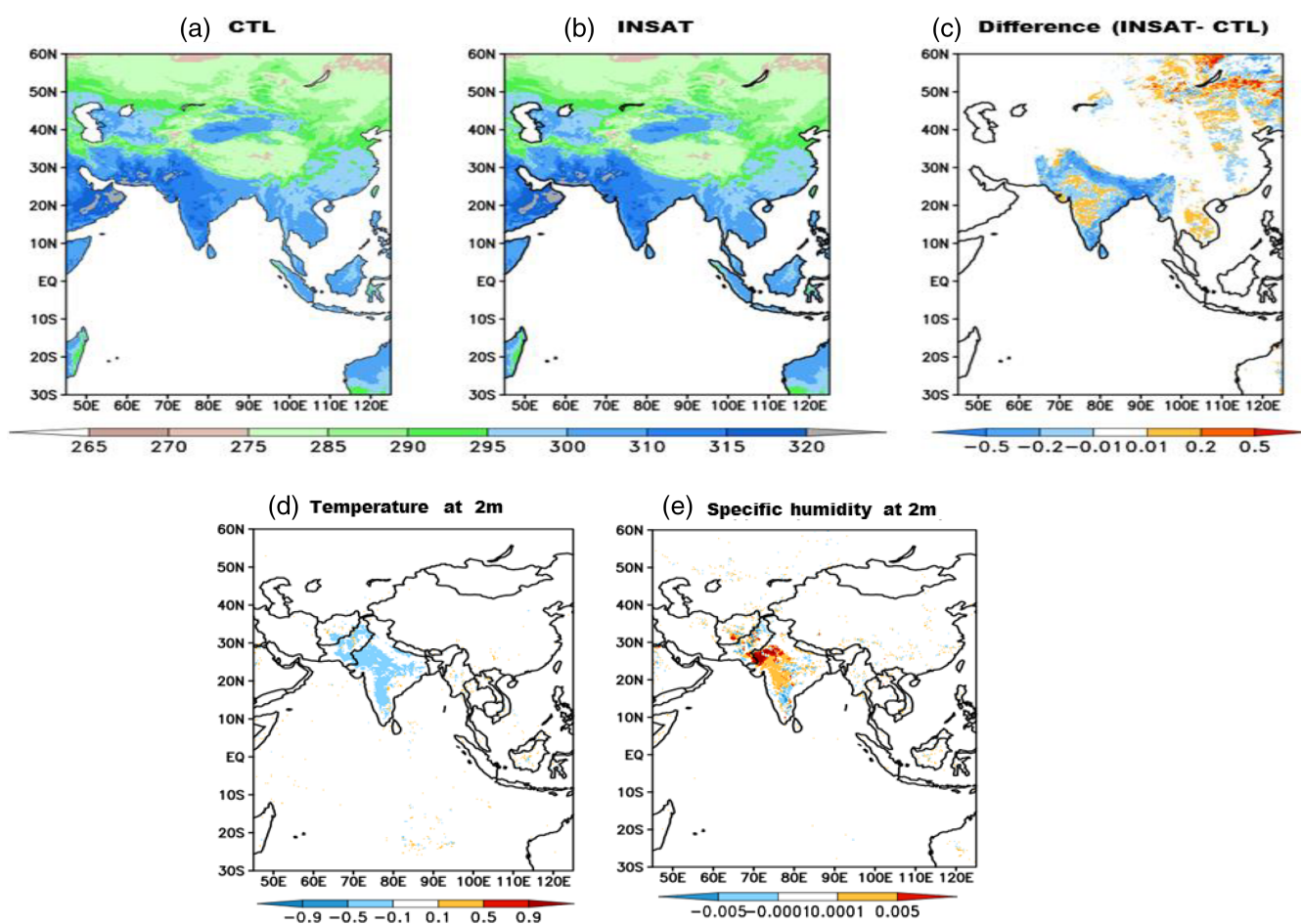


FIGURE 5 Comparison of soil temperature analysis (model units: K) produced at NCMRWF from (a) CTL (b) INSAT experiment*. Difference (INSAT - CTL) of (c) soil temperature (layer 1), (d) temperature at 2 m, and (e) specific humidity at 2 m, for a typical data-assimilation cycle valid at 06UTC 20180502. (* Kindly note in this experiment INSAT-3D LT assimilated over Indian region only).

observations utilized in the CTL experiment. This forms the observing system experiment (OSE) experiment.

The atmospheric initial conditions for both numerical experiments are generated through Hybrid-4DVAR DA method by utilizing the available conventional and

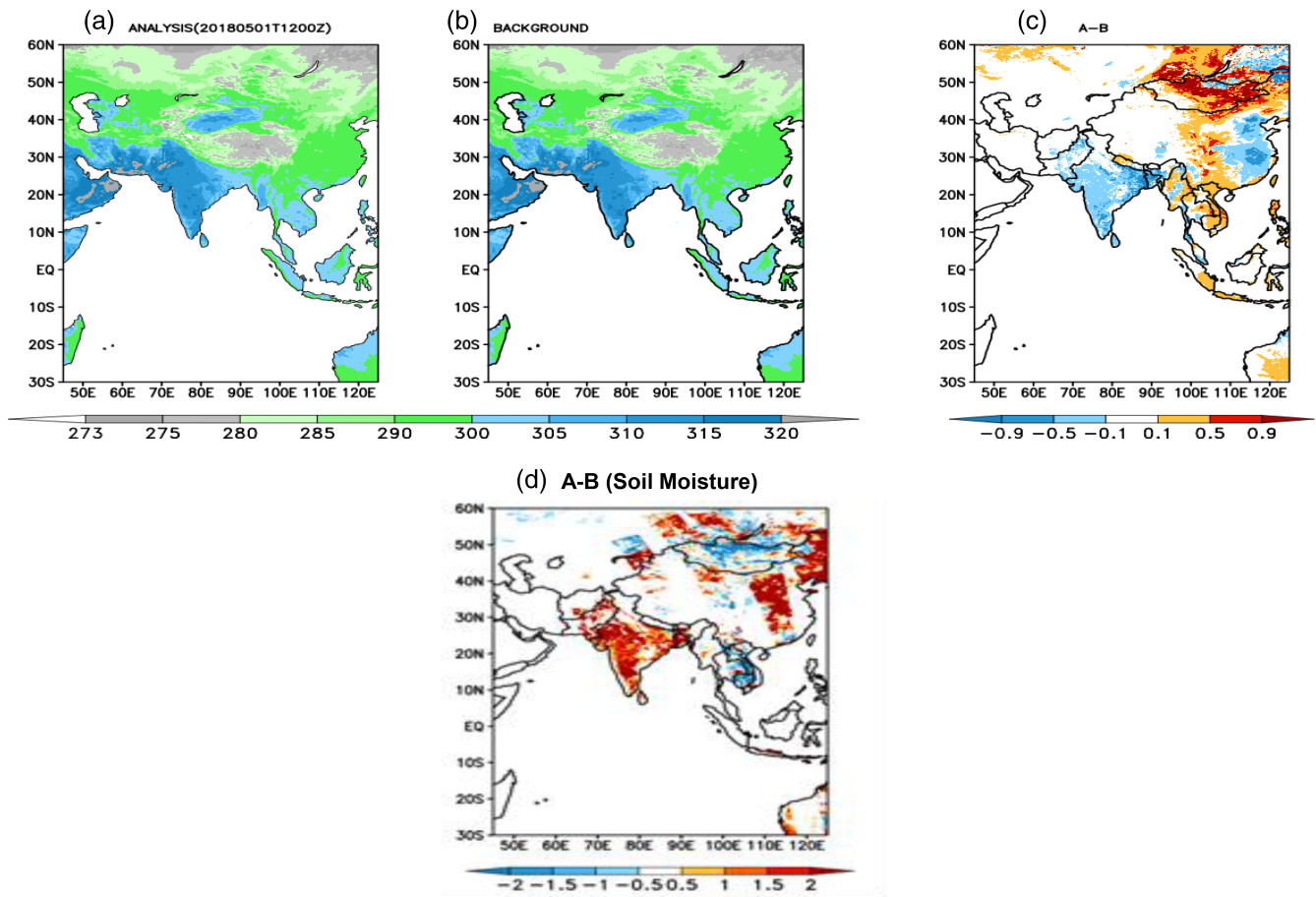


FIGURE 6 Mean surface soil temperature (model units: K) for (a) analysis from INSAT experiment (A); (b) model background (B); and (c) difference (Analysis – Background) on a typical time of 12 UTC 20180502. (d) Same as (c) but for soil moisture difference (Units: kg/m²).

non-conventional atmospheric observations. It is worth mentioning that the atmospheric analyses are produced separately from the land surface analyses (SM and LT). The land and atmosphere coupling happen when the forecast model initiates the forecast and makes use of both land surface initial condition from NCUM LDAS and atmospheric initial condition from Hybrid-4DVAR. The integration of the JULES model within the LDAS and the NCUM global forecasting system effectively incorporates the feedback mechanism. The essential element of the sEKF-based LDAS for NCUM entails the calculation of the Jacobians of the observation operator **H**. These Jacobians create the relationship between the satellite-derived INSAT-3D LT observations and the land surface variables by utilizing perturbation forecasts from the offline JULESv5.1 model. Computing the Jacobians provide the physical link INSAT LT gridded observations and the soil levels of the JULES model. After this the “INSAT LT to sEKF pre-processor” will generate the quality-controlled gridded LT (O-B) innovations on the model grid. For INSAT LT data pre-processing, to compute the innovation, the nearest

background LT field consistent to the observation time is used (the current time interval is 30 min within the assimilation cycle).

The LT observations acquired from the INSAT-3D via ISRO and processed by the MOSDAC are gathered within the geographical boundary of India (5–40° N; 65–100° E). Following a QC, the high-quality LT observations are chosen for ingesting into the NCUM's LDAS. A pre-processing system incorporating several QC methods, calibration techniques, and consistency checks is implemented before INSAT-3D LT assimilation. Details can be found in Lodh et al. (2019). It is essential to note that INSAT-3D LT observations are accessible at 00, 06, and 12 UTC, but not at 18 UTC (Lodh et al., 2019). As a result, the LT observations are not utilized in the 18 UTC of the DA cycle. It is crucial to note that the assimilation window for the INSAT 3D LT is ± 30 min relative to the analysis time for every assimilation cycle. Figure S4 shows the spatial extent of the mean of gridded INSAT-3D LT observations over Indian domain after QC for a typical LDA cycle 12UTC 20190202.

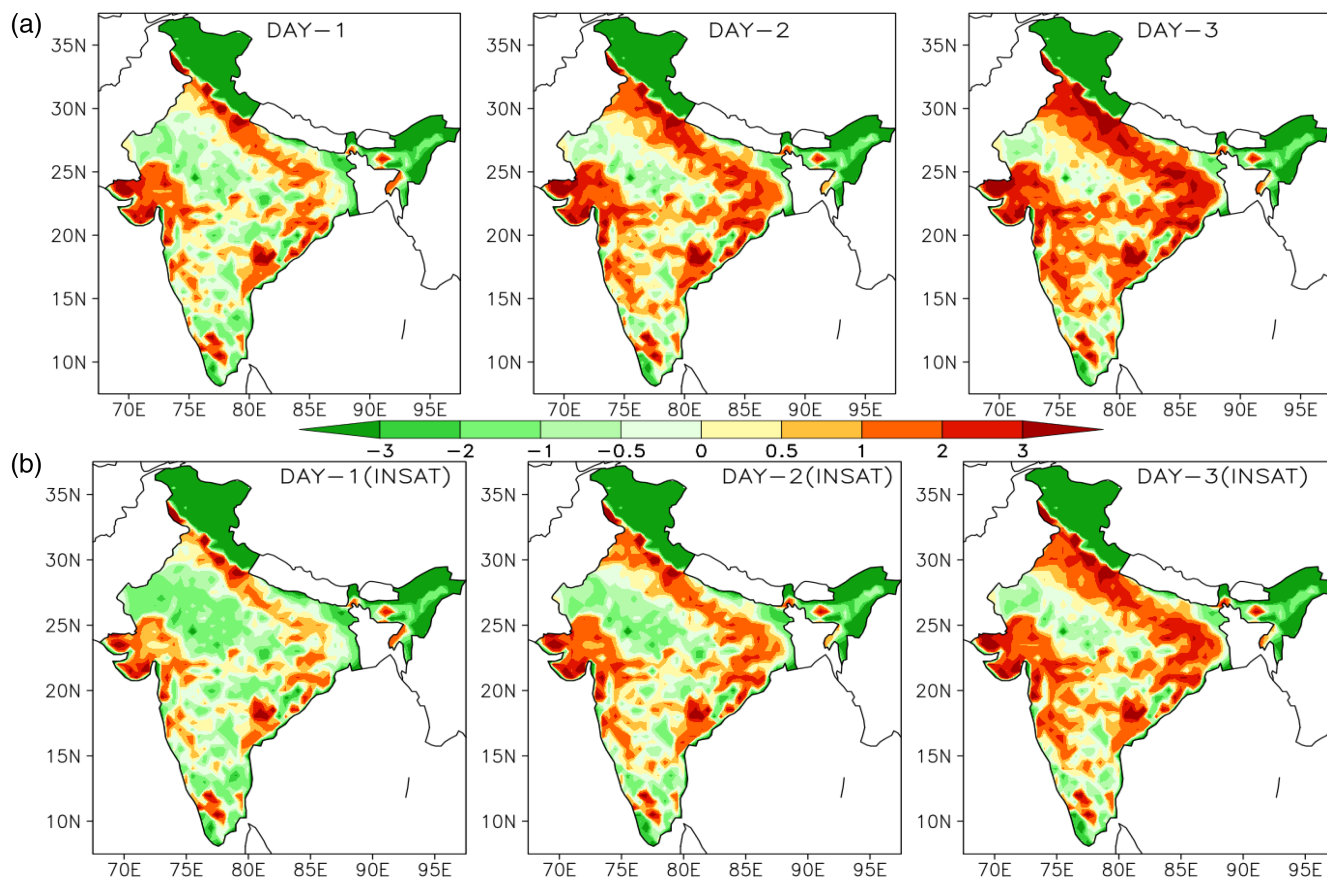


FIGURE 7 Spatial distribution of mean error (ME) of maximum temperature ($T_{\max}$; K) for summer period (a) CTL (upper panel) and (b) INSAT (lower panel) experiments for the period of 31 days from May 8 to June 7, 2018.

The observation sensitivity experiments were conducted by selecting a summer month (May 2018) and a winter month (February 2019). Furthermore, the research undertaken is the first study to calculate the Jacobians pertaining to surface-level observations and satellite measurements of skin temperature or LT derived from INSAT-3D over India, employing the JULESv5.1 model. The LT temperature perturbations utilized in the forecasts has a magnitude of -0.05 K, as indicated in Table 1. The SM and LT analyses generated internally at the NCMRWF are utilized to set up the initial surface boundary conditions (BCs) for the NCUM long-forecast simulations. For our study, the NCUM model in conjunction with the JULES land surface scheme generates atmospheric forecasts up to 7 days. In the conducted tests, seven-day forecasts were prepared for all days in May 2018 and February 2019, utilizing the Initial Condition (IC) obtained at 00 UTCs. The CTL and INSAT experiments are carried out during the summer (winter) months, where the model ran for the period 00UTC 20180501 to 18UTC 20180610 (00UTC 20190201 to 18UTC 20190228). Figure 4a,b depict various observations (both atmospheric and land

observations) acquired from distinct platforms and assimilated in the CTL and INSAT experiments for a typical assimilation cycle valid at 12 UTC on May 2, 2018, respectively. It is noteworthy to emphasize that except for ASCAT SM and INSAT-3D LT, all other observations assimilated in the upper-air atmospheric data assimilation system.

5 | RESULTS AND DISCUSSIONS

Figure 5a–c shows the ST (K) from both the analyses (CTL and INSAT experiments) along with their difference (INSAT-CTL) over the Indian region for a typical DA cycle valid at 06 UTC May 2, 2018. Figure 5c illustrates the effect of assimilating INSAT-3D satellite LT data on ST and humidity over India, whereas Figure 5d,e displays the 2 m temperature and humidity difference between INSAT-CTL, respectively, over the Indian region. The impact on LDAS on 2m-specific humidity is of opposite sign to temperature. Also, it is noticed that the assimilation of LT observations updates the screen temperature over the Indian region, especially over the

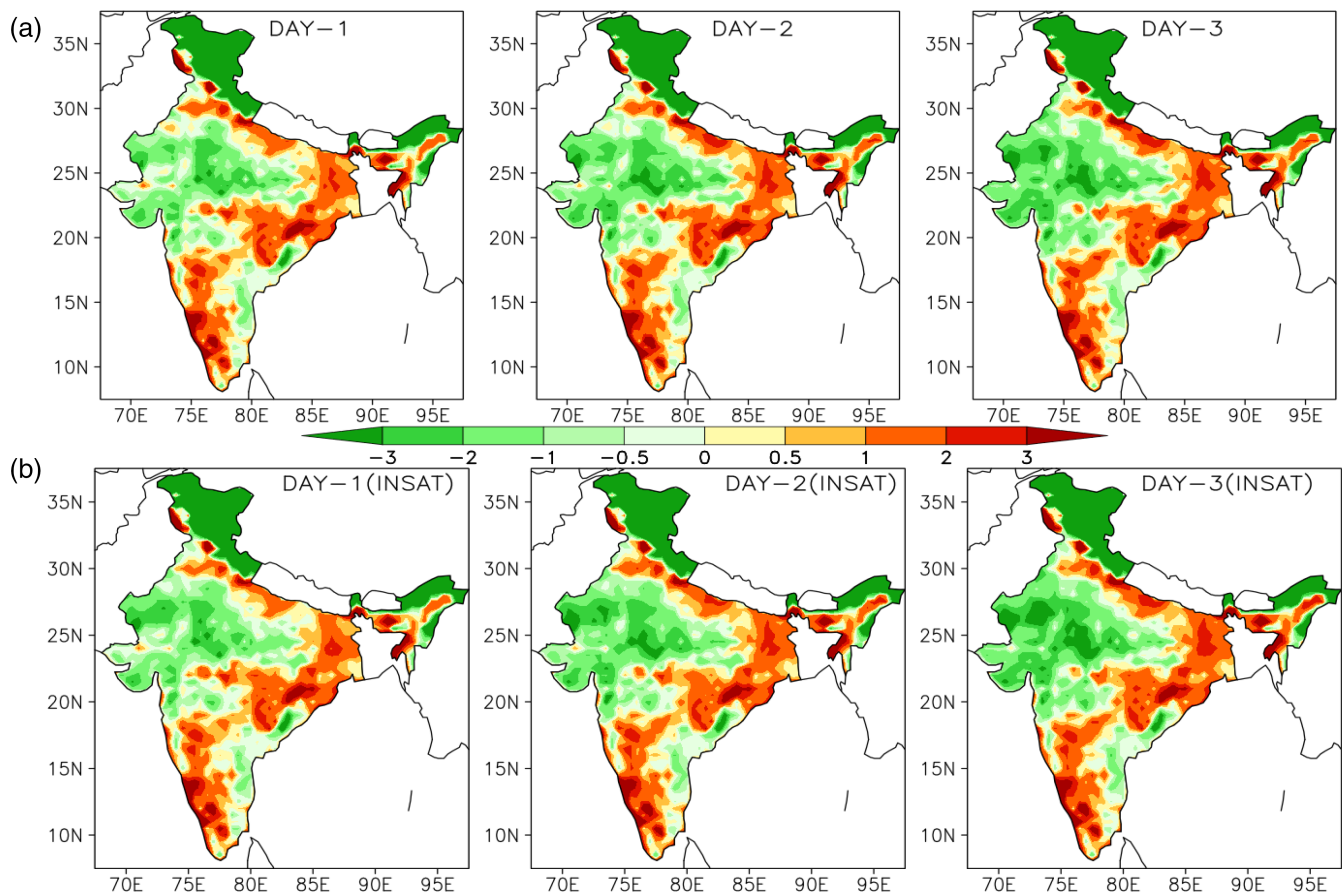


FIGURE 8 Spatial distribution of mean error (ME) of minimum temperature ($T_{\min}$; K) for summer period (a) CTL (upper panel) and (b) INSAT (lower panel) experiments for the period 31 days from May 8 to June 7, 2018.

arid and semi-arid regions, where there are few LT observations. Figure S5a–c displays the soil temperature analysis from CTL, INSAT experiment, and difference in soil temperature analysis between the INSAT and CTL experiment (valid at 06 UTC May 2, 2018), showing the geographical coverage and impact of assimilating INSAT-3D satellite LT data on ST and humidity over its whole coverage area. Figure 6a–d depicts the ST analysis from INSAT experiment (A) and model background (B) along with the differences (A–B) in ST and SM, respectively, valid at 12 UTC May 2, 2018. It is clearly seen from the Figure 6c,d that the ST and SM is modified over the regions, where the INSAT-3D LT data are assimilated, with reduction in the mean error. With respect to the model first-guess, the LT analysis is reduced over the India region by up to -0.9 K. It is noteworthy to mention that the LT observations are only considered over the Indian region in the DA cycles. Also, over the ocean and snow, there are no LT observations assimilated in the model because these are discarded through QC procedures. Hence, the innovations from the A–B (Figure 6c,d) can be noticed only over the land only (not oceans and terrain regions). Therefore, the maximum change in ST

due to INSAT-3D LT assimilation is seen over the Indian regions. It is noticed that there are changes in the A – B values of LT over other parts of the globe, which can be attributed to ASCAT assimilation. Over the Indian domain, the adjustment in the LT analysis after assimilation of INSAT 3D LT observations can be attributed to the better representation of skin and ST, land-atmosphere interactions, and vegetative feedback via the JULES model in the LDAS, demonstrating the potential positive impact of INSAT-3D LT data for enhancing the surface temperature analysis over India.

Figures 7a,b and 8a,b, respectively, depicts the spatial distribution of mean error of maximum and minimum near surface temperature for the summer (May 8 to June 7, 2018) and winter (February 2019) months, as determined by the CTL and INSAT experiments. The mean error is calculated with respect to the observed maximum and minimum temperature ($\sim 0.25 \times 0.25^\circ$ resolution) values provided by India Meteorological Department (IMD; Srivastava et al., 2009). According to the IMD, the maximum temperature is usually measured in the afternoon, around 2:30 PM local time (i.e. 14:30 IST), whereas the minimum temperature is typically recorded in the

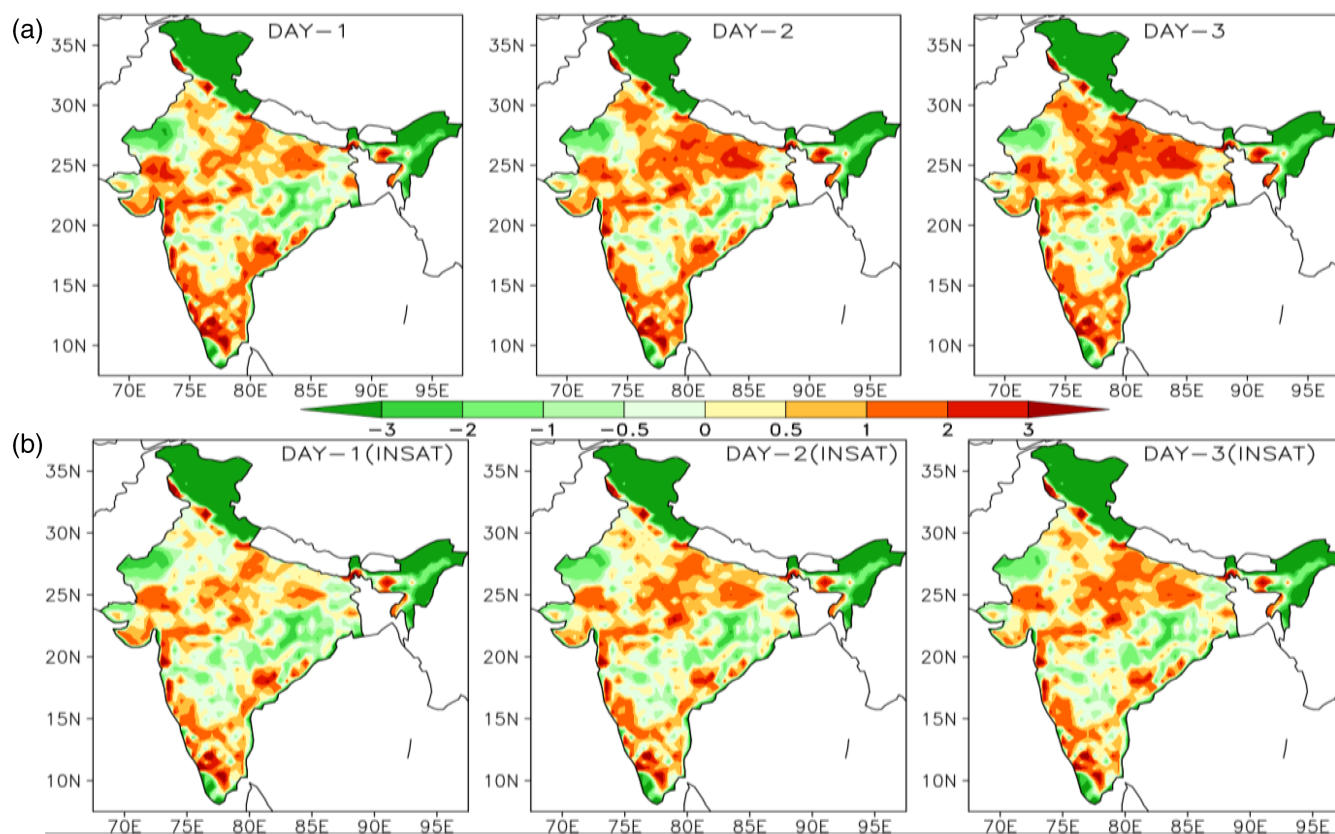


FIGURE 9 Spatial distribution of mean error (ME) of maximum temperature ($T_{\max}$; K) for winter period (a) CTL (1st row) and (b) INSAT (2nd row) experiments for the period of 28 days from February 1 to 28, 2019.

early morning, around 5:30 AM local time (05:30 IST), when the earth's surface has lost the most heat during the night, and temperatures tend to be the lowest. From the Figures 7 and 8, it is confirmed that the mean error in maximum/minimum temperature forecasts over the Indian domain is reduced for all the forecast days by 0.5°K to 1°K during the summer month by the assimilation of INSAT-3D LT in the INSAT run, in comparison with the CTL simulation. The day 3 forecast (Figure 7) of INSAT simulation shows that the mean error in the day time maximum temperature is varying between $\pm 0.5^{\circ}\text{K}$ over the central part of India. From Figure 8, it is shown that during the summer month of May 2018, the mean error in forecasts of minimum temperature is reduced over the north-west and central parts of India as well as over the Ganges Plains. Comparing Figures 5c, 6c, 7 and 8, it is observed that there is decrease in the ST increments all over north-west India, Gangetic Plains, peninsular India at the initial time of the model run (06UTC 20,180,502). Whereas, the mean errors in $T_{\max}$ and $T_{\min}$ are reduced over central India, the Ganges Plains, Chotanagpur plateau, and peninsular India region.

Similarly, the spatial distribution of mean error of maximum and minimum temperature for the winter

month of February 2019 from the CTL and INSAT simulations is shown in Figures 9a,b and 10a,b, respectively. The mean error of maximum/minimum temperature is once again reduced around 1°K to 2°K in the winter month (February) for all the forecast days, particularly over the Indo-Gangetic Plains and north-west of India. However, the mean error in minimum temperature (Figure 10) is reduced over peninsular India, Rajasthan, Gujarat, West Bengal, and Bihar by -3°K to -0.5°K . The improvement in the forecast errors in maximum and minimum temperature in the INSAT experiments can be attributed to the fact that the sEKF-based LDAS assimilated the LT observations to update the land surface state variables after every 6-h cycle, which resulted in better analysis and forecasts of surface temperature in the INSAT experiment.

The NCUM model simulated LT is also evaluated with the location-specific global Telecommunication System (GTS) $T_{\max}$ and $T_{\min}$ observations (Source: IMD) during the summer and winter months. Figure 11a-j shows the time series of model simulated (day 1–5) maximum and minimum temperature forecasts from CTL and INSAT experiments compared with the station observed data at the Delhi location

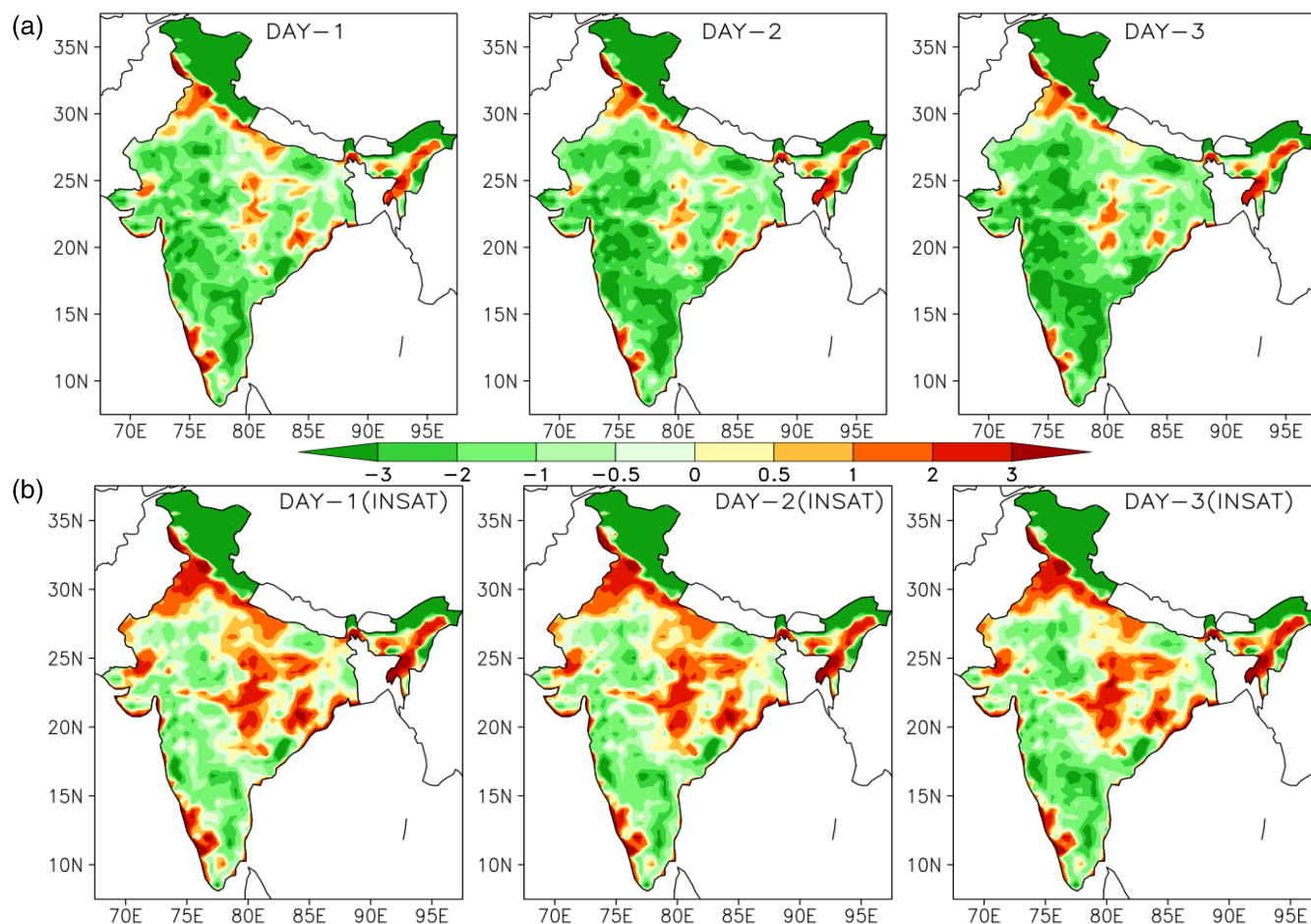
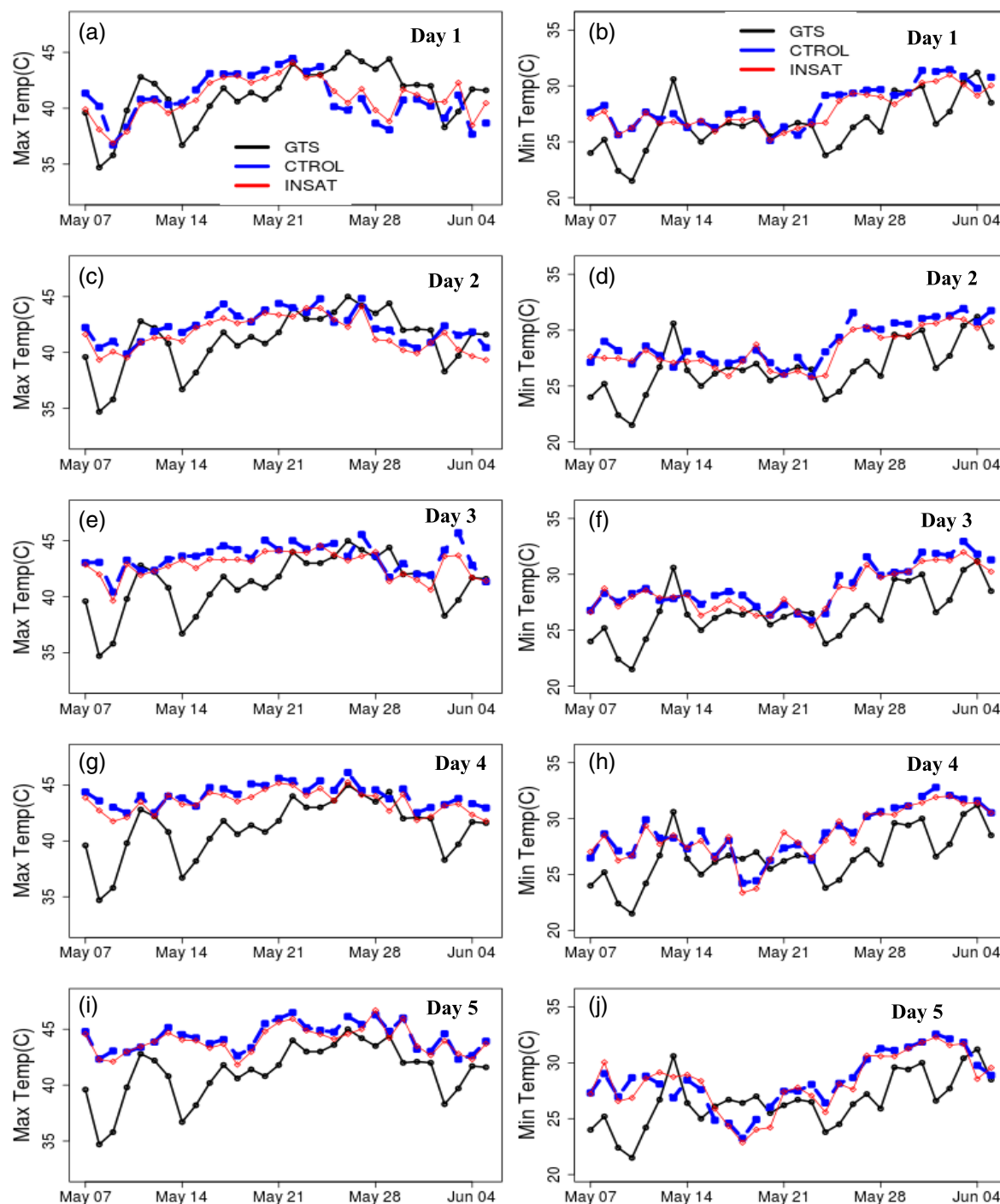


FIGURE 10 Spatial distribution of mean error (ME) of maximum temperature ($T_{\min}$; K) for winter period (a) CTL (1st row) and (b) INSAT (2nd row) experiments for the period of 28 days from February 1 to 28, 2019.

(28.56° N, 77.22° E) during May 8 to June 7, 2018, respectively. The evolution pattern of the temperature in all the forecast days is reasonably well-captured by both the model simulations during the period. However, the temperature obtained from the INSAT experiment is relatively well matching with the IMD observations. It is clearly seen that the bias, correlation coefficient, and root mean square error (RMSE) are considerably improved in the INSAT simulations in all the forecast days compared with the CTL. A decrease in bias suggests that the INSAT runs are closer to the observation over Delhi compared with the CTL. Thus, the INSAT experiment captures the LT conditions more accurately, with smaller systematic errors than that of CTL. Higher correlation also indicates that the LT forecasts from the INSAT experiment better capture the variations and patterns of LT at the Delhi location, leading to improved agreement with the station observed data. RMSE quantifies the average magnitude of the differences between the forecasted and observed values. A lower RMSE

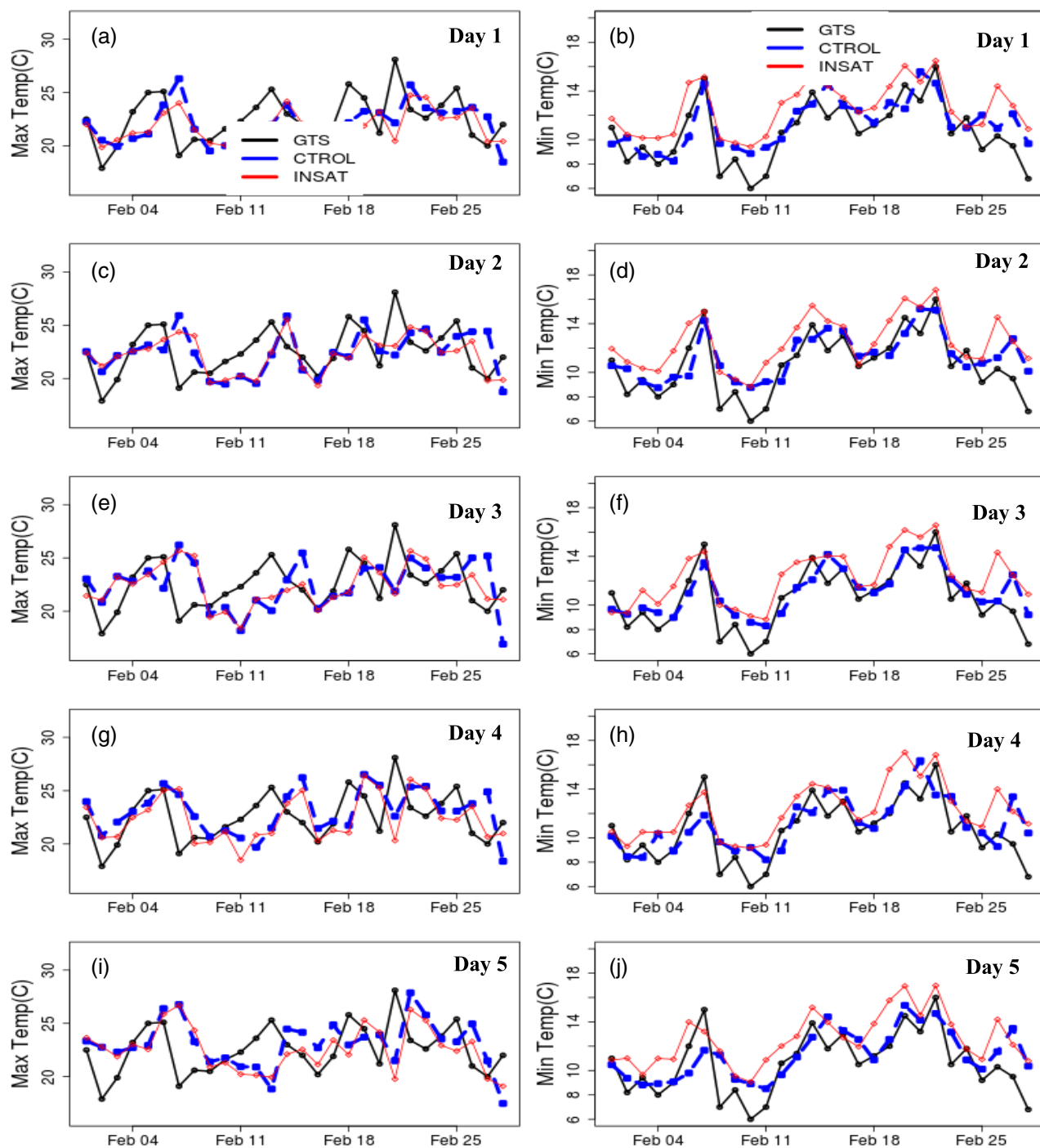
indicates that the INSAT experiment produces LT forecasts that are closer to the actual observations, indicating better accuracy and precision in capturing the variability of LT at the Delhi location. At the same time, it is again observed that the reduced bias, increased correlation, and decreased RMSE in the INSAT runs for the month of February 2019 (Figure 12) indicate improvements in the accuracy, reliability, and skill of the surface temperature predictions after assimilation of INSAT LT observations into the LDAS compared with the CTL experiment. These results highlight the values and effectiveness of incorporating the INSAT data that enhanced the model forecast skill for LT.

Figure 13 shows the hourly forecasted visibility, near surface temperature, and relative humidity (RH) obtained from both the model simulations compared with the available observation over Indira Gandhi International (IGI) airport, Delhi (28.56° N, 77.22° E) during the period February 14 and 16, 2019. The diurnal variation of surface temperature and RH are well



(a) Max. Temp.	DAY1-CTL	DAY1-INSAT	DAY2-CTL	DAY2-INSAT	DAY3-CTL	DAY3-INSAT	DAY4-CTL	DAY4-INSAT	DAY5-CTL	DAY5-INSAT
Bias	0.25	0.25	-1.16	-0.50	-2.31	-1.64	-2.91	-2.31	-3.13	-2.79
Correlation	0.20	0.44	0.39	0.42	0.20	0.28	0.43	0.39	0.51	0.50
RMSE	2.86	2.33	2.59	2.34	3.43	2.92	3.66	3.24	3.78	3.52
(b) Min. Temp.	DAY1-CTL	DAY1-INSAT	DAY2-CTL	DAY2-INSAT	DAY3-CTL	DAY3-INSAT	DAY4-CTL	DAY4-INSAT	DAY5-CTL	DAY5-INSAT
Bias	-1.63	-1.17	-2.22	-1.64	-2.32	-1.91	-2.16	-1.97	-1.82	-1.65
Correlation	0.51	0.52	0.48	0.50	0.56	0.56	0.49	0.50	0.38	0.44
RMSE	2.65	2.33	3.10	2.66	3.08	2.76	3.15	3.00	3.19	3.02

FIGURE 11 Time series (May 8 to June 7, 2018) of model simulated and observed (GTS) a) $T_{\max}$ (K) and (b) $T_{\min}$ (K) over Delhi location (28.56° N, 77.22° E) for day 1 forecast. Similarly, (c,d); (e,f); (g,h); and (i,j) are same as (a,b) but for day 2 to 5, respectively.



(a) Max. Temp.	DAY1-CTL	DAY1-INSAT	DAY2-CTL	DAY2-INSAT	DAY3-CTL	DAY3-INSAT	DAY4-CTL	DAY4-INSAT	DAY5-CTL	DAY5-INSAT
Bias	0.60	0.74	0.14	0.31	0.04	0.18	-0.48	0.17	-0.55	-0.05
Correlation	0.14	0.16	0.13	0.27	0.04	0.08	0.12	0.13	0.07	0.10
RMSE	2.80	2.51	2.80	2.48	3.30	2.64	2.91	2.91	3.32	3.21
(b) Min. Temp.	DAY1-CTL	DAY1-INSAT	DAY2-CTL	DAY2-INSAT	DAY3-CTL	DAY3-INSAT	DAY4-CTL	DAY4-INSAT	DAY5-CTL	DAY5-INSAT
Bias	-1.94	-0.70	-1.91	-0.69	-1.68	-0.51	-1.61	-0.53	-1.96	-0.71
Correlation	0.78	0.90	0.79	0.89	0.83	0.86	0.68	0.85	0.71	0.81
RMSE	2.25	1.73	2.24	1.70	2.13	1.51	2.09	1.96	2.46	1.93

FIGURE 12 Time series (February 1 to 28, 2019) of model simulated and observed (GTS) (a) $T_{\max}$ (K) and (b) $T_{\min}$ (K) over Balasore location (21.52° N, 86.93° E) for day 1 forecast. Similarly, (c,d); (e,f); (g,h); and (i,j) are same as (a,b) but for day 2 to 5, respectively.

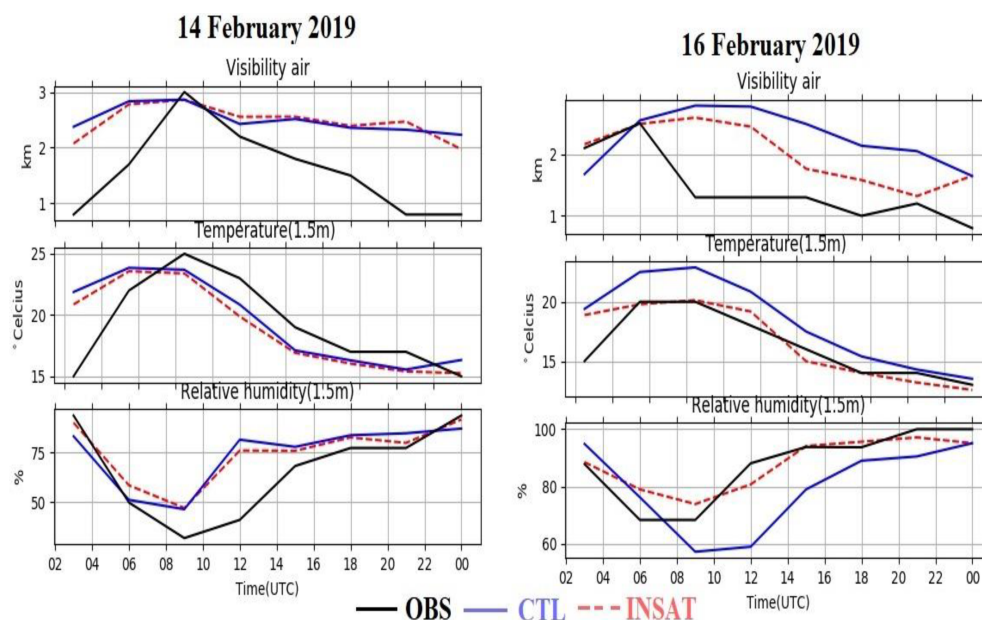


FIGURE 13 Compared hourly time series of model simulated visibility (km), temperature (C) and relative humidity at 1.5 m (%) compared with IGI airport Delhi (28.56° N, 77.22° E) observation during February 14 and 16, 2019.

simulated by both the model simulations. However, the variation pattern and values are relatively well-captured by the INSAT simulation on both the days, when compared to the CTL simulation. In the case of the visibility, the model simulations are far from the observed one during February 14, 2019. By contrast, the visibility is considerably improved in the INSAT simulation on February 16, 2019, whereas more deviation is noticed in the CTL simulation during the period.

Several fundamental weather verification metrics defined by the World Meteorological Organization (WMO) are used to assess the accuracy and skill of temperature forecasts over India. The metrics used in our study are:

- Probability of Detection (POD) or the hit rate is defined as number of hits divided by total number of hits and misses.
- Equitable Threat Score (ETS) measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance.
- Critical Success Index (CSI) or threat score measures the fraction of observed and/or forecast events that were correctly predicted.
- False Alarm Ratio (FAR) is the false alarms divided by total number of hits and false alarms.
- Heidke Skill Score (HSS) measures the fraction of correct forecasts after eliminating those forecasts which would be correct due purely to random chance.
- Symmetric Extremal Dependency Index (SEDI).

For more details about the metrics, refer CAWCR (n.d.), Hogan et al., 2009, Routray et al., 2023, Singh et al., 2017.

These indices provide quantitative measures of forecast accuracy, reliability, and skill. The IMD gridded $T_{\max}$ and $T_{\min}$ observed datasets are used for verification purposes (Singh et al., 2017; Srivastava et al., 2009). The various statistical skill scores are calculated for the boreal summer (May 8 to June 7, 2018) and boreal winter (February 2019) periods. The skill scores calculated for the boreal summer period ($T_{\max}$) and boreal winter period ($T_{\min}$) at different temperature ranges from the CTL and INSAT experiments are depicted in Figures 14a–f and 15a–f, respectively. The verification indices are computed for all the forecast days from day 1 to day 5. It can be observed that the skill scores of maximum temperatures forecasts during the boreal summer obtained from the INSAT experiment of all the forecast days is improved at different temperature bins of 32–34°C, 34–36°, 36–38°C, 38–40°C to 40–42°C with 1 day (24 hours) gain in advanced forecasts. The POD calculated based on day 1 (day 5) forecasts from the INSAT experiment has increased from ~0.35 to ~0.42 (~0.18 to ~0.22) for the prediction of temperature range, 40–42°C. This means more ratio of hits than misses, indicating correctly forecasted the events to the total number of observed events in comparison with the CTL, by approximately 15%. It is deduced that after assimilation of INSAT-3D LT, the skill of the forecast model is improved and also the ability to detect the occurrence of heat wave (extreme temperature) events. The ETS measures the fraction of observed and/or forecast events that were correctly predicted, adjusted for hits associated with random chance. The ETS LT forecasts from the INSAT experiment have a higher proportion of correct forecasts than the CTL (increased to ~0.15 for day 1 forecasts of 40–42°C temperature range), which relative to what would be expected by chance alone. The HSS considers both hits and

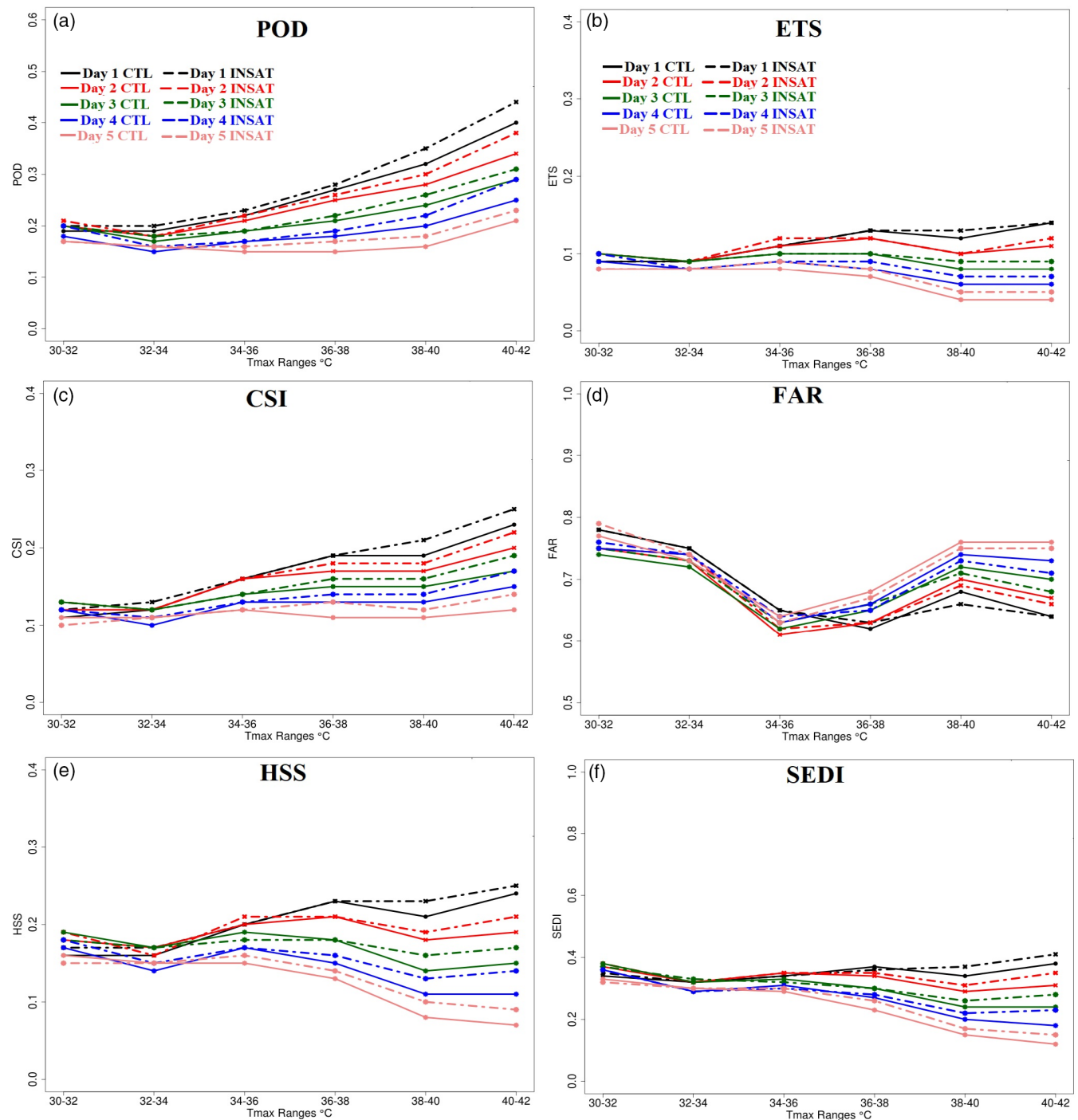


FIGURE 14 Verification metrics of $T_{\max}$ (a) Probability of Detection (POD); (b) Equitable Threat Score (ETS); (c) Critical Success Index (CSI); (d) False Alarm Ratio (FAR); (e) Heidke Skill Score (HSS); and (f) Symmetric Extremal Dependency Index (SEDI) from CTL (straight line) and INSAT (dashed line) at various forecast days (Day 1 to 5) for the summer period (May 8 to June 7, 2018).

correct rejections as well as assesses the overall forecast skill. The HSS scores calculated from the INSAT simulations (day 5 to day 1) show improved forecasts compared with the CTL forecast (increased to ~ 0.22 for day 1 forecasts of 40–42 $^{\circ}\text{C}$ temperature range). The SEDI score from the INSAT and the CTL forecasts shows that the LT forecasts from the INSAT experiment have better ability to capture

the dependence between observed and forecasted extremes (increased to ~ 0.4 for day 1 forecasts of 40–42 $^{\circ}\text{C}$ temperature range). The FAR score from the INSAT experiment is lower than that from the CTL forecast (decreased to ~ 0.75 to 0.6 for day 1 forecasts of 40–42 $^{\circ}\text{C}$ temperature range), thus indicating that the frequency of false alarms is less after the assimilation of INSAT-3D LT and also indicates

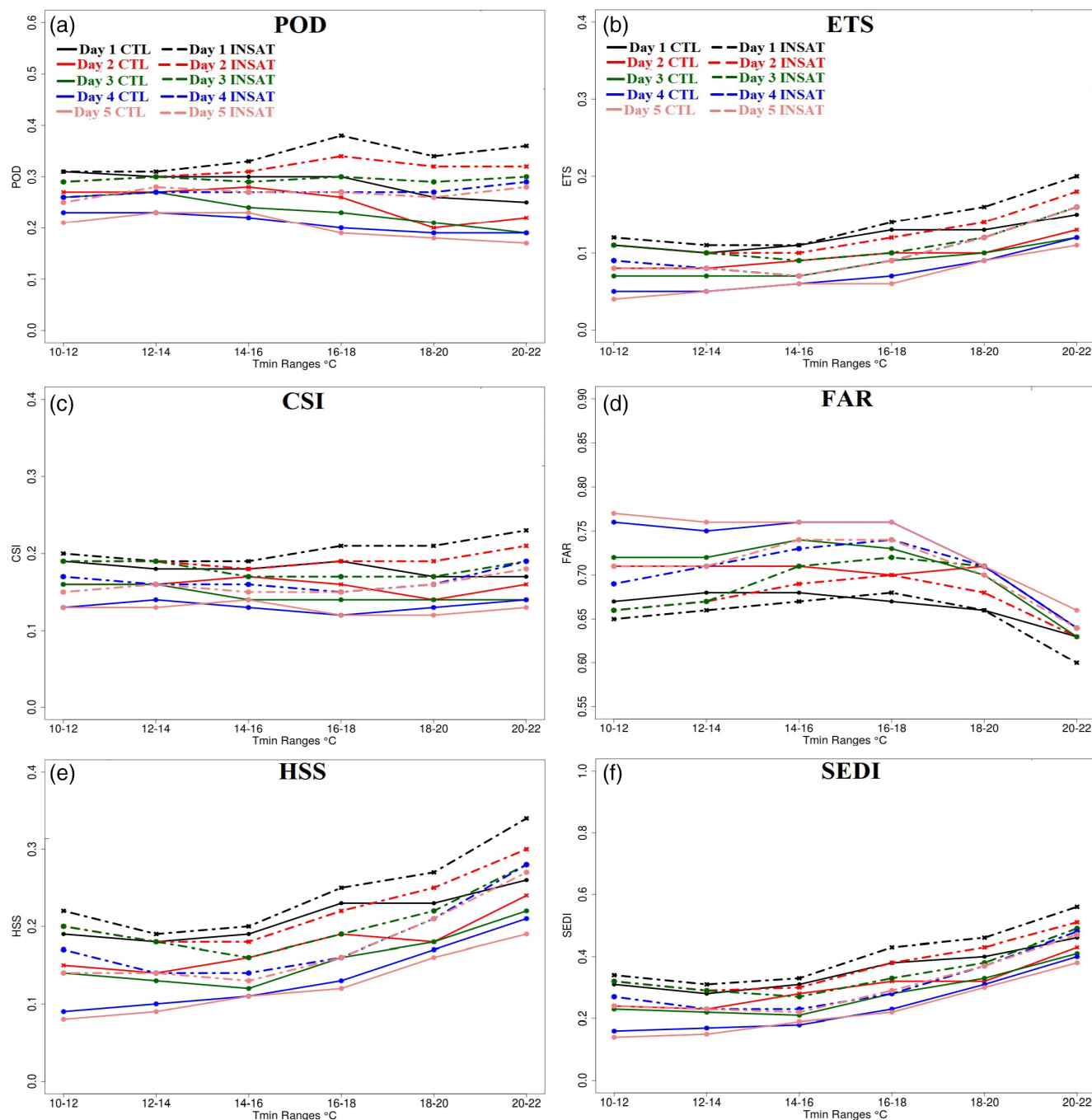


FIGURE 15 Verification metrics of $T_{\min}$ (a) Probability of Detection (POD); (b) Equitable Threat Score (ETS); (c) Critical Success Index (CSI); (d) False Alarm Ratio (FAR); (e) Heidke Skill Score (HSS); and (f) Symmetric Extremal Dependency Index (SEDI) from CTL (straight line) and INSAT (dashed line) at various forecast days (Day 1 to 5) for the winter period (February 1 to 28, 2019).

the higher level of correct forecasting of LT. Hence, the LT forecasts from the INSAT experiment have higher skill in forecasting extreme weather events, such as high temperatures. During the boreal summer month, it is also shown that the forecast for the minimum temperature (figure not proven) gets better in the 22–24 $^{\circ}\text{C}$, 24–26 $^{\circ}\text{C}$, and 26–28 $^{\circ}\text{C}$ temperature ranges with higher verification (especially POD) scores. Similarly, various verification metrics based

on INSAT and CTL simulations are also computed for minimum temperature forecasts during the boreal winter period (February 2019) depicted in Figure 15a–f, which is same as to Figure 14a–f. The HSS scores for minimum temperature forecasts for boreal winter calculated from the INSAT simulations (day 5 to day 1) show improved forecasts compared with the CTL forecast (increased to ~ 0.35 for day 1 forecasts of 20–22 $^{\circ}\text{C}$ temperature range). A similar trend was also

noticed for FAR score. This clearly inferred the skill of both maximum and minimum temperature forecasts during boreal summer and winter months has improved due to the assimilation of INSAT-3D LT in the sEKF-based LDAS. The best impact of LT assimilation on the near surface parameters is achieved with the improved control vector of land variables through sEKF and through the land-atmosphere coupling via JULES land surface scheme. The assimilation of LT can have a positive impact on the near surface (2 m) temperature forecasts even after 24 h (Lodh et al., 2019; Nayak, 2020). Hence, through the application of the sEKF technique for assimilating INSAT-3D LT data, a notable improvement of approximately 24 h in forecasting of maximum and minimum temperatures during boreal summer and winter months is achieved over India.

6 | SUMMARY AND CONCLUSIONS

The current research work is a first study to assimilate the INSAT-3D LT directly into the sEKF-based weakly coupled LDAS within the NCUM modelling system over the Indian region. The NCMRWF LDAS is now capable of generating global analyses of screen level parameters (2 m temperature and relative humidity), sea surface temperature, snow depth, SM, and LT at N1024 (12 km) resolution. Past study like Singh et al., 2016 assimilated the LT into the WRF-3DVAR analysis system by considering the LT observations as an atmospheric variable (2 m surface temperature), which is not an appropriate methodology to assimilate the LT observation without using a LDAS system at 25 km grid resolution using a mesoscale WRFv3.4 model. It is most important to assimilate the LT through LDAS that properly establishes the positive correlation between surface temperature and atmospheric boundary layer through a well-mixed boundary layer (Sgoff et al., 2020). Many past studies suggested the utilization of LT observations in the land assimilation system which is used to provide more accurate information and reduce the uncertainty in the ICs of surface state to the NWP forecast model. By incorporating the satellite LT estimates into the NCUM model, the assimilation process can correct any discrepancies between the model's initial state and the real-world conditions represented by the satellite observations. The assimilation of LT allows us to evaluate the impact of surface state as well as atmospheric forecasts. Here, we have conducted two independent numerical experiments such as CTL (assimilation of SM from ASCAT satellite plus in-situ screen level observations) and INSAT (assimilation of INSAT-3D LT observations along with the

observations utilized in the CTL) for Boreal summer (May 8 to June 7, 2018) and Boreal winter (February 2019) period. The atmospheric analyses are generated through the Hybrid-4DVAR data assimilation technique by utilizing the available conventional and non-conventional atmospheric observations for both experiments.

The model analysis from the INSAT experiment shows the assimilation of LT observations modulates the screen temperature over the Indian region, especially in the arid and semi-arid regions. At the same time, the LT analysis increments are reduced up to 1.0 K over the Indian region, where most of the LT observations are considered for the assimilation cycle. The experiment with LT assimilation was validated against in-situ observations, and overall, the RMSE of soil and near surface temperature of the first-guess (and analysis) are reduced. Further, the impact of assimilating INSAT-3D LT on weather forecasts over India is evaluated by comparing the forecast of maximum and minimum temperature with the observations from IMD stations. The results show that assimilating INSAT-3D LT improves the forecast skill of both maximum and minimum near surface temperature over India, especially in regions with high LT variability. The assimilation also reduces the mean error and RMSEs of temperature forecast over India. The improvement is more pronounced in forecasts of maximum (minimum) temperature during the summer (winter) season. The assimilation of INSAT-3D LT also enhances the representation of the diurnal cycle of LT with adjustments in latent heat flux, and sensible heat flux in the model (Liu et al., 2013). The various statistical skill scores pronounced that the assimilation of INSAT-3D LT has considerably improved the forecast skill of the model. The accurate initialization of lower boundary conditions like SM and LT in a coupled NWP model is vital for capturing the complex land-atmosphere interactions, thereby enhancing the accuracy of seasonal predictions and extreme weather events like heat waves. It can be further improved by using high-resolution LT data from various satellites (geostationary and polar orbiting), improving the quality control algorithm, and optimizing the assimilation parameters. The study is a contribution to the community as it enables the utilization of INSAT-3D LT data for a wide range of applications, contributing to our understanding of LT variations and their implications for weather, climate, and environmental studies. The promising findings derived from this study over India provide impetus for additional efforts aimed at assimilation of wider range of LT data obtained from diverse satellite sources like MODIS.

Currently land surface model parameters have an under-spread in sEKF-based LDAS. Also, in the sEKF, it is not easy to cycle the background error covariance matrix **B**.

Future work will involve research, development on ensemble techniques for performing coupled LDAS experiments (de Rosnay et al., 2022; Kurosawa et al., 2023) as the error covariances will have flow-dependent contributions based on ensemble trajectories (Bojarova et al., 2011). The ensemble techniques also has the capability to include horizontal model error correlations, optimal combination of the flow-dependent and climatological statistics based on trade-off between performance and ensemble size.

AUTHOR CONTRIBUTIONS

Abhishek Lodh: Conceptualization; methodology; software; data curation; investigation; validation; formal analysis; writing – original draft; visualization; writing – review and editing. **Ashish Routray:** Writing – review and editing; validation; investigation; resources; methodology; project administration. **Devajyoti Dutta:** Software; data curation; resources. **Vivek Singh:** Writing – review and editing. **John. P. George:** Supervision; project administration; methodology.

ACKNOWLEDGEMENTS

The authors acknowledge to MOSDAC, ISRO, for providing access to the INSAT-3D LT observations in real time. The authors thank to the IMD for their provision of temperature observations utilized in the present study. We would like to express our appreciation and gratefulness to the scientists, UK Met Office for their valuable technical support. Acknowledgement is also extended to Head, NCMRWF, NOIDA (MoES India), for his encouragement and provided the necessary facilities to perform the study. The first author also expresses gratitude towards SMHI and Lund University, Sweden, for help. Dr. Jelena Bojarova (SMHI, Norrköping, Sweden) and Dr. Ekaterina Kurzenova (FMI, Helsinki, Finland) are acknowledged for academic discussions. The authors sincerely appreciate the insightful comments and suggestions provided by the Editor and anonymous reviewers, which have greatly contributed to improving the quality of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors whose names are listed in the manuscript certify that they have NO affiliations with or involvement in any organization or entity with any financial interest in the subject matter or materials discussed in this manuscript. We hereby declare that the work described in this research article has been entirely carried by us, and it has not been submitted earlier either wholly or partly to any other journals.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Abhishek Lodh  <https://orcid.org/0000-0002-4951-0226>

REFERENCES

- Aires, F., Prigent, C. & Rossow, W.B. (2004) Temporal interpolation of global surface skin temperature diurnal cycle over land under clear and cloudy conditions. *Journal of Geophysical Research: Atmospheres*, 109(D4), 1–18. Available from: <https://doi.org/10.1029/2003JD003527>
- Albergel, C., Munier, S., Leroux, D.J., Dewaele, H., Fairbairn, D., Barbu, A.L. et al. (2017) Sequential assimilation of satellite-derived vegetation and soil moisture products using SURFEX_v8.0: LDAS-monde assessment over the Euro-Mediterranean area. *Geoscientific Model Development*, 10, 3889–3912. Available from: <https://doi.org/10.5194/gmd-10-3889-2017>
- Bakketun, Å., Blyverket, J. & Müller, M. (2023) Using a reanalysis-driven land surface model for initialization of a numerical weather prediction system. *Weather and Forecasting*, 38(11), 2155–2168. Available from: <https://doi.org/10.1175/WAF-D-22-0184.1>
- Balsamo, G., Beljaars, A., Scipal, K., Viterbo, P., van den Hurk, B., Hirschi, M. et al. (2009) A revised hydrology for the ECMWF model: verification from field site to terrestrial water storage and impact in the integrated forecast system. *Journal of Hydrometeorology*, 10(3), 623–643. Available from: <https://doi.org/10.1175/2008JHM1068.1>
- Balsamo, G., Mahfouf, J.-F., Bélair, S. & Deblonde, G. (2007) A land data assimilation system for soil moisture and temperature: an information content study. *Journal of Hydrometeorology*, 8(6), 1225–1242. Available from: <https://doi.org/10.1175/2007JHM819.1>
- Benjamin, S.G., Smirnova, T.G., James, E.P., Lin, L.-F., Hu, M., Turner, D.D. et al. (2022) Land–snow data assimilation including a moderately coupled initialization method applied to NWP. *Journal of Hydrometeorology*, 23(6), 825–845. Available from: <https://doi.org/10.1175/JHM-D-21-0198.1>
- Best, M.J., Pryor, M., Clark, D.B., Rooney, G.G., Essery, R.L.H., Ménard, C.B. et al. (2011) The joint UK land environment simulator (JULES), model description – part 1: energy and water fluxes. *Geoscientific Model Development*, 4, 677–699. Available from: <https://doi.org/10.5194/gmd-4-677-2011>
- Bojarova, J., Gustafsson, N., Johansson, Å. & Vignes, O. (2011) The ETKF rescaling scheme in HIRLAM. *Tellus A: Dynamic Meteorology and Oceanography*, 63(3), 385–401. Available from: <https://doi.org/10.1111/j.1600-0870.2011.00513.x>
- Boussetta, S., Balsamo, G., Dutra, E., Beljaars, A. & Albergel, C. (2015) Assimilation of surface albedo and vegetation states from satellite observations and their impact on numerical weather prediction. *Remote Sensing of Environment*, 163, 75–86. Available from: <https://doi.org/10.1016/j.rse.2015.03.009>
- Bouttier, F. & Courtier, P. (2002) Data assimilation, concepts and methods. Meteorological Training Course Lecture Series ECMWF. Available from: <https://www.ecmwf.int/en/elibrary/79860-data-assimilation-concepts-and-methods> [Accessed 27 November 2024]
- Brown, A., Milton, S., Cullen, M., Golding, B., Mitchell, J. & Shelly, A. (2012) Unified modeling and prediction of weather and climate: a 25-year journey. *Bulletin of the American*

- Meteorological Society*, 93(12), 1865–1877. Available at: <https://doi.org/10.1175/BAMS-D-12-00018.1>
- Candy, B. (2013) Implementation of satellite LST into the EKF plan of work. Available from: <https://code.metoffice.gov.uk/doc/surf/surf-2017.07.0/doc/SUDP18sciplan.pdf>
- Candy, B. (2014) The use of satellite information in the land data assimilation to support operational NWP. Available from: <https://code.metoffice.gov.uk/doc/surf/surf-2017.07.0/doc/SSD P05.pdf>
- Candy, B., Saunders, R.W., Ghent, D. & Bulgin, C.E. (2017) The impact of satellite-derived land surface temperatures on numerical weather prediction analyses and forecasts. *Journal of Geophysical Research: Atmospheres*, 122(18), 9783–9802. Available from: <https://doi.org/10.1002/2016JD026417>
- Capecchi, V. & Brocca, L. (2014) A simple assimilation method to ingest satellite soil moisture into a limited-area NWP model. *Meteorologische Zeitschrift*, 23(2), 105–121. Available from: <https://doi.org/10.1127/0941-2948/2014/0585>
- CAWCR. (n.d.) Forecast Verification: A Practitioner's Guide to Atmospheric Science. Available from: https://www.cawcr.gov.au/projects/verification/verif_web_page.html [Accessed 26th September 2024].
- Clark, D.B., Mercado, L.M., Sitch, S., Jones, C.D., Gedney, N., Best, M.J. et al. (2011) The joint UK land environment simulator (JULES), model description – part 2: carbon fluxes and vegetation dynamics. *Geoscientific Model Development*, 4, 701–722. Available from: <https://doi.org/10.5194/gmd-4-701-2011>
- de Rosnay, P., Balsamo, G., Albergel, C., Muñoz-Sabater, J. & Isaksen, L. (2014) Initialisation of land surface variables for numerical weather prediction. *Surveys in Geophysics*, 35(3), 607–621. Available from: <https://doi.org/10.1007/s10712-012-9207-x>
- de Rosnay, P., Browne, P., de Boissésou, E., Fairbairn, D., Hirahara, Y., Ochi, K. et al. (2022) Coupled data assimilation at ECMWF: current status, challenges and future developments. *Quarterly Journal of the Royal Meteorological Society*, 148(747), 2672–2702. Available from: <https://doi.org/10.1002/qj.4330>
- de Rosnay, P., Drusch, M., Balsamo, G., Albergel, C. & Isaksen, L. (2011) Extended Kalman filter soil-moisture analysis in the ifs. ECMWF No. 127, pp. 12–16.
- de Rosnay, P., Drusch, M., Vasiljevic, D., Balsamo, G., Albergel, C. & Isaksen, L. (2013) A simplified extended Kalman filter for the global operational soil moisture analysis at ECMWF. *Quarterly Journal of the Royal Meteorological Society*, 139, 1199–1213. Available from: <https://doi.org/10.1002/qj.2023>
- Dharssi, I. (2010) The extended Kalman filter land data assimilation system: planning paper. Surface Working Paper No. 2. Available from: <https://code.metoffice.gov.uk/doc/surf/surf-2017>
- Dharssi, I., Bovis, K.J., Macpherson, B. & Jones, C.P. (2011) Operational assimilation of ASCAT surface soil wetness at the met Office. *Hydrology and Earth System Sciences*, 15, 2729–2746. Available from: <https://doi.org/10.5194/hess-15-2729-2011>
- Dharssi, I., Candy, B. & Steinle, P. (2015) Analysis of the linearised observation operator in a soil moisture and temperature analysis scheme. *SOIL Discuss.*, 2, pp.505–535. <https://doi.org/10.5194/soild-2-505-2015>
- Draper, C.S., Mahfouf, J.F. & Walker, J.P. (2009) An EKF assimilation of AMSR-E soil moisture into the ISBA land surface scheme. *Journal of Geophysical Research: Atmospheres*, 114(D20104). Available from: <https://doi.org/10.1029/2008JD011650>
- Drusch, M., Scipal, K., de Rosnay, P., Balsamo, G., Andersson, E., Bougeault, P. et al. (2009) Towards a Kalman filter based soil moisture analysis system for the operational ECMWF integrated forecast system. *Geophysical Research Letters*, 36, L10401. Available from: <https://doi.org/10.1029/2009GL037716>
- Duerinckx, A., Hamdi, R., Deckmyn, A., Djebbar, A., Mahfouf, J.-F. & Termonia, P. (2017) Combining an EKF soil analysis with a 3D-Var upper-air assimilation in a limited-area NWP model. *Quarterly Journal of the Royal Meteorological Society*, 143(708), 2999–3013. Available from: <https://doi.org/10.1002/qj.3141>
- Duerinckx, A., Hamdi, R., Mahfouf, J.-F. & Termonia, P. (2015) Study of the Jacobian of an extended Kalman filter for soil analysis in SURFEXv5. *Geoscientific Model Development*, 8(3), 845–863. Available from: <https://doi.org/10.5194/gmd-8-845-2015>
- ECMWF (European Centre for Medium-Range Weather Forecasts), 2024. ERA5-Land: Data Documentation. [online] Available at: <https://confluence.ecmwf.int/display/CKB/ERA5-Land%3A+data+documentation> [Accessed 27 November 2024].
- Ek, M.B. & Holtslag, A.A.M. (2004) Influence of soil moisture on boundary layer cloud development. *Journal of Hydrometeorology*, 5, 86–99. Available from: [https://doi.org/10.1175/1525-7541\(2004\)005<0086>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0086>2.0.CO;2)
- Fairbairn, D., Barbu, A.L., Napoly, A., Albergel, C., Mahfouf, J.-F. & Calvet, J.-C. (2017) The effect of satellite-derived surface soil moisture and leaf area index land data assimilation on streamflow simulations over France. *Hydrology and Earth System Sciences*, 21, 2015–2033. Available from: <https://doi.org/10.5194/hess-21-2015-2017>
- Fisher, M., Leutbecher, M. & Kelly, G.A. (2005) On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. *Quarterly Journal of the Royal Meteorological Society*, 131, 3235–3246. Available from: <https://doi.org/10.1256/qj.04.142>
- Ghent, D., Kaduk, J., Remedios, J., Ardö, J. & Balzter, H. (2010) Assimilation of land surface temperature into the land surface model JULES with an ensemble Kalman filter. *Journal of Geophysical Research: Atmospheres*, 115(D19), 1–16. Available from: <https://doi.org/10.1029/2010JD014392>
- Gómez, B., Charlton-Pérez, C.L., Lewis, H. & Candy, B. (2020) The met Office operational soil moisture analysis system. *Remote Sensing*, 12(22), 3691. Available from: <https://doi.org/10.3390/rs12223691>
- Halder, S., Dirmeyer, P.A., Marx, L. & Kinter, J.L. (2018) Impact of land surface initialization and land-atmosphere coupling on the prediction of the Indian summer monsoon with the CFSv2. *Frontiers in environmental Science*, 5, 1–19. Available from: <https://doi.org/10.3389/fenvs.2017.00092>
- Hamdi, R., Degrauwe, D., Duerinckx, A., Cedilnik, J., Costa, V., Dalkilic, T. et al. (2014) Evaluating the performance of SURFEXv5 as a new land surface scheme for the ALADINcy36 and ALARO-0 models. *Geoscientific Model Development*, 7, 23–39. Available from: <https://doi.org/10.5194/gmd-7-23-2014>
- Han, X., Hendricks Franssen, H.-J., Li, X., Zhang, Y., Montzka, C. & Vereecken, H. (2013) Joint assimilation of surface temperature and L-band microwave brightness temperature in land data assimilation. *Vadose Zone Journal*, 12(3), 1–16. Available from: <https://doi.org/10.2136/vzj2012.0072>

- Hogan, R.J., O'Connor, E.J. & Illingworth, A.J. (2009) Verification of cloud-fraction forecasts. *Quarterly Journal of the Royal Meteorological Society*, 135(643), 1494–1511. Available at: <https://doi.org/10.1002/qj.481>
- Houser, P.R., De Lannoy, G.J.M. & Walker, J.P. (2010) Land surface data assimilation. In: Lahoz, W.A., Khattatov, B. & Ménard, R. (Eds.) *Data assimilation: making sense of observations*. Berlin: Springer, pp. 549–597. Available at: https://doi.org/10.1007/978-3-540-74703-1_21
- Huang, C., Li, X., Lu, L. & Gu, J. (2008) Experiments of one-dimensional soil moisture assimilation system based on ensemble Kalman filter. *Remote Sensing of Environment*, 112(3), 888–900. Available from: <https://doi.org/10.1016/j.rse.2007.06.026>
- Hurrell, J., Meehl, G.A., Bader, D., Delworth, T.L., Kirtman, B. & Wielicki, B. (2009) A unified modeling approach to climate system prediction. *Bulletin of the American Meteorological Society*, 90(12), 1819–1832. Available from: <https://doi.org/10.1175/2009BAMS2752.1>
- Kalnay, E. (2002) *Atmospheric Modeling, data assimilation and predictability*. Cambridge: Cambridge University Press. Available from: <https://doi.org/10.1017/CBO9780511802270>
- Koster, R., Mahanama, S., Yamada, T., Balsamo, G., Boissarie, M., Dirmeyer, P. et al. (2010) The contribution of land surface initialization to sub-seasonal forecast skill: first results from the GLACE-2 project. *Geophysical Research Letters*, 37, L02402. Available from: <https://doi.org/10.1029/2009GL041677>
- Koster, R.D. & Suarez, M.J. (2003) Impact of land surface initialization on seasonal precipitation and temperature prediction. *Journal of Hydrometeorology*, 4, 408–423. Available from: [https://doi.org/10.1175/1525-7541\(2003\)4<408>2.0.CO](https://doi.org/10.1175/1525-7541(2003)4<408>2.0.CO)
- Kourzeneva, E. (2014) Assimilation of lake water surface temperature observations using an extended Kalman filter. *Tellus A: Dynamic Meteorology and Oceanography*, 66(1), 21510. Available from: <https://doi.org/10.3402/tellusa.v66.21510>
- Kumar, J., Hoffman, F.M., Hargrove, W.W. & Collier, N. (2017) Understanding the representativeness of FLUXNET for upscaling carbon flux from eddy covariance measurements. *Earth System Science Data Discuss*, 1–25. <https://doi.org/10.5194/essd-2016-36>
- Kumar, S., Kolassa, J., Reichle, R., Crow, W., de Lannoy, G., de Rosnay, P. et al. (2022) An agenda for land data assimilation priorities: realizing the promise of terrestrial water, energy, and vegetation observations from space. *Journal of Advances in Modeling Earth Systems*, 14, e2022MS003259. Available from: <https://doi.org/10.1029/2022MS003259>
- Kurosawa, K., Kotsuki, S. & Miyoshi, T. (2023) Comparative study of strongly and weakly coupled soil moisture data assimilation with a global coupled land-atmosphere model. *EGU sphere* [Preprint]. <https://doi.org/10.5194/egusphere-2023-887>
- Li, Z., Li, J., Li, Y., Zhang, Y., Schmit, T.J., Zhou, L. et al. (2012) Determining diurnal variations of land surface emissivity from geostationary satellites. *Journal of Geophysical Research: Atmospheres*, 11, 1–14. Available from: <https://doi.org/10.1029/2012JD018279>
- Li, Z.-L., Tang, B.-H., Wu, H., Ren, H., Yan, G., Wan, Z., Trigo, I.F. & Sobrino, J.A., (2013) Satellite-derived land surface temperature: Current status and perspectives. *Remote Sensing of Environment*, 131, 14–37. Available at: <https://doi.org/10.1016/j.rse.2012.12.008>
- Liu, C., Shu, S. & Gao, W., (2013) Improved sensible and latent heat flux estimation of community land model by using ensemble Kalman filter assimilation. *Proc. SPIE 8869, Remote Sensing and Modeling of Ecosystems for Sustainability X*, 886917, <https://doi.org/10.1117/12.2021713>
- Lodh. (2020) Reassessment of land-atmosphere interactions over India during summer monsoon using state-of-the-art regional climate models. *Theoretical and Applied Climatology*, 142(3–4), 1–12. Available from: <https://doi.org/10.1007/s00704-020-03395-x>
- Lodh, A., George, G., Singh, H., George, J.P. & Rajagopal, E.N. (2019) Assimilation of INSAT-3D land surface temperature in EKF based land data assimilation system. Technical report, NMRF/TR/07/2019.
- Lodh, A., George, J.P. & Rajagopal, E.N. (2016) Extended Kalman filter based land data assimilation system for soil moisture analysis at NCMRWF. Technical report, NMRF/TR/06/2016.
- Lodh, A., Routray, A., Dutta, D., George, J.P. & Mitra, A.K. (2022) Improving the prediction of monsoon depressions by assimilating ASCAT soil moisture in NCUM-R modeling system. *Atmospheric Research*, 272, 106130. Available from: <https://doi.org/10.1016/j.atmosres.2022.106130>
- Lorenc, A.C. & Hammon, O. (1988) Objective quality control of observations using Bayesian methods. Theory, and a practical implementation. *Quarterly Journal of the Royal Meteorological Society*, 114, 515–543. Available from: <https://doi.org/10.1002/qj.49711448012>
- Lu, Y., Steele-Dunne, S.C., Farhadi, L. & van de Giesen, N. (2017) Mapping surface heat fluxes by assimilating SMAP soil moisture and GOES land surface temperature data. *Water Resources Research*, 53, 10858–10877. Available from: <https://doi.org/10.1002/2017WR021415>
- Mahfouf, J.F., Bergaoui, K., Draper, C., Bouyssel, F., Taillefer, F. & Taseva, L. (2009) A comparison of two offline soil analysis schemes for assimilation of screen level observations. *Journal of Geophysical Research*, 114, D08105. Available from: <https://doi.org/10.1029/2008JD011077>
- Mitra, A. (2014) INSAT-3D data products Catalog. [online]. Available from: http://satellite.imd.gov.in/dynamic/INSAT3D_Catalog.pdf [Accessed 14th September 2024].
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D.G., Piles, M., Rodríguez-Fernández, N.J., Zsoter, E., Buontempo, C. and Thépaut, J.-N., (2021) ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4371. <https://doi.org/10.5194/essd-13-4349-2021>
- Nayak, S. (2020) Remote sensing for national development: the legacy of Dr. Vikram Sarabhai. *Journal of the Indian Society of Remote Sensing*, 48, 1101–1120. Available from: <https://doi.org/10.1007/s12524-020-01156-x>
- Pandey, P.C., Chauhan, A. & Maurya, N.K. (2022) Evaluation of earth observation datasets for LST trends over India and its implication in global warming. *Ecological Informatics*, 72, 101843. Available from: <https://doi.org/10.1016/j.ecoinf.2022.101843>
- Pandya, M.R., Shah, D.B., Trivedi, H.J. & Panigrahy, S. (2011) Simulation of at-sensor radiance over land for proposed

- thermal channels of imager payload onboard INSAT-3D satellite using MODTRAN model. *Journal of Earth System Science*, 120(1), 19–25. Available from: <https://doi.org/10.1007/s12040-011-0021-9>
- Pandya, M.R., Shah, D.B., Trivedi, H.J., Panigrahy, S., Parihar, J.S. & Kirankumar, A.S. (2011) Evaluation of split-window algorithms for retrieving land surface temperature from INSAT-3D imager observations. *Vayumandal*, 37, 31–37.
- Parada, L.M. & Liang, X. (2008) Impacts of spatial resolutions and data quality on soil moisture data assimilation. *Journal of Geophysical Research*, 113, D10101. Available from: <https://doi.org/10.1029/2007JD009037>
- Parrens, M., Mahfouf, J.-F., Barbu, A.L. & Calvet, J.-C. (2014) Assimilation of surface soil moisture into a multilayer soil model: design and evaluation at local scale. *Hydrology and Earth System Sciences*, 18, 673–689. Available from: <https://doi.org/10.5194/hess-18-673-2014>
- Rajagopal, E.N., Iyengar, G.R., George, J.P., Das Gupta, M., Mohandas, S., Siddharth, R. et al. (2012) Implementation of the UM Model-Based Analysis–Forecast System at NCMRWF. NMRWF/TR/2012.
- Reichle, R.H., Kumar, S.V., Mahanama, S.P.P., Koster, R.D. & Liu, Q. (2010) Assimilation of satellite-derived skin temperature observations into land surface models. *Journal of Hydrometeorology*, 11(5), 1103–1122. Available from: <https://doi.org/10.1175/2010JHM1262.1>
- Routray, A., Lodh, A., Dutta, D., George, J.P. & Mitra, A.K. (2023) Influence of ASCAT soil moisture on prediction of track and intensity of landfall tropical cyclones. *International Journal of Remote Sensing*, 44(1), 341–380. Available from: <https://doi.org/10.1080/01431161.2022.2164232>
- Routray, A., Singh, V., Singh, H., Dutta, D., George, J.P. & Rakhi, R. (2019) Evaluation of different versions of NCUM global model for simulation of track and intensity of tropical cyclones over bay of Bengal. *Dynamics of Atmospheres and Oceans*, 78, 71–88. Available from: <https://doi.org/10.1016/j.dynatmoce.2017.04.001>
- Seneviratne, S.I., Wilhelm, M., Stanelle, T., van den Hurk, B., Hagemann, S., Berg, A. et al. (2013) Impact of soil moisture-climate feedbacks on CMIP5 projections: first results from the GLACE-CMIP5 experiment. *Geophysical Research Letters*, 40(19), 5212–5217. Available from: <https://doi.org/10.1002/grl.50956>
- Sgoff, C., Schomburg, A., Schmidli, J. & Potthast, R. (2020) Assimilating synthetic land surface temperature in a coupled land-atmosphere model. *Quarterly Journal of the Royal Meteorological Society*, 146(733), 3980–3997. Available from: <https://doi.org/10.1002/qj.3883>
- Sharifnezhadazizi, Z., Norouzi, H., Prakash, S., Beale, C. & Khanbilvardi, R. (2019) A global analysis of land surface temperature diurnal cycle using MODIS observations. *Journal of Applied Meteorology and Climatology*, 58(6), 1279–1291. Available from: <https://doi.org/10.1175/JAMC-D-18-0256.1>
- Singh, H., Arora, K., Ashrit, R. & Rajagopal, E.N. (2017) Verification of pre-monsoon temperature forecasts over India during 2016 with a focus on heatwave prediction. *Natural Hazards and Earth System Sciences*, 17, 1469–1485. Available from: <https://doi.org/10.5194/nhess-17-1469-2017>
- Singh, R., Singh, C., Ojha, S.P., Kumar, A.S., Kishtawal, C.M. & Kumar, A.S.K. (2016) Land surface temperature from INSAT-3D imager data: retrieval and assimilation in NWP model. *Journal of Geophysical Research: Atmospheres*, 121, 6909–6926. Available from: <https://doi.org/10.1002/2016JD024752>
- Srivastava, A.K., Rajeevan, M. & Kshirsagar, S.R. (2009) Development of a high-resolution daily gridded temperature data set (1969–2005) for the Indian region. *Atmospheric Science Letters*, 10(4), 249–254. Available from: <https://doi.org/10.1002/asl.232>
- Stull, R. (2017) *Practical meteorology: an algebra-based survey of atmospheric science – version 1.02b*. Vancouver, BC: University of British Columbia. ISBN 978-0-88865-283-6.
- Trigo, I.F., Monteiro, I.T., Olesen, F. & Kabsch, E. (2008) An assessment of remotely sensed land surface temperature. *Journal of Geophysical Research: Atmospheres*, 113, 1–12. Available from: <https://doi.org/10.1029/2008JD010035>
- Van Genuchten, M.T. (1980) A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Science Society of America Journal*, 44, 892–898. Available from: <https://doi.org/10.2136/sssaj1980.03615995004400050002x>
- Walters, D., Baran, A.J., Boutle, I., Brooks, M., Earnshaw, P., Edwards, J. et al. (2019) The met Office unified model global atmosphere 7.0/7.1 and JULES global land 7.0 configurations. *Geoscientific Model Development*, 12, 1909–1963. Available from: <https://doi.org/10.5194/gmd-12-1909-2019>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lodh, A., Routray, A., Dutta, D., Singh, V., & George, J. P. (2024). Impact of INSAT-3D land surface temperature assimilation via simplified extended Kalman filter-based land data assimilation system on forecasting of surface fields over India. *Meteorological Applications*, 31(6), e70019. <https://doi.org/10.1002/met.70019>