International Journal of Environmental & Agriculture Research (IJOEAR)
ISSN: [2454-1850]
[Vol-12, Issue-10, October- 2026]
Website: www.ijoear.com
Journal DOI: 10.25125/agriculture-journal

Ensemble Kalman Filter Optimization for Convection-Permitting Forecasts in Complex Terrain

Dr. Sachin S. Chinchorkar

Associate Professor, Polytechnic in Agricultural Engineering, Anand Agricultural University, Dahod, Gujarat, India

Received:- 18 September 2026/ Revised:- 29 September 2026/ Accepted:- 11 October 2026/ Published: 31-10-2026
Copyright @ 2026 International Journal of Environmental and Agriculture Research
This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (https://creativecommons.org/licenses/by-nc/4.0) which permits unrestricted Non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract— *Weather forecasting in complex terrain remains challenging due to mountainous topography, which generates high spatial variability, local atmospheric circulations, and rapidly evolving convective systems. This paper focuses on the optimization of the Ensemble Kalman Filter (EnKF) to enhance convection-permitting numerical weather prediction in complex terrain. A high-resolution ensemble forecasting framework is used to analyze the effects of key EnKF parameters, including ensemble size, covariance localization, and inflation. Controlled experiments are used to compare baseline and optimized configurations with respect to meteorological parameters, including temperature, wind, humidity, and precipitation. The analysis shows that ensemble size affects the stability of forecast-error covariance estimation, whereas covariance localization is required to control unrealistic spatial correlations caused by finite ensemble sizes. Ensemble inflation helps ensure sufficient ensemble dispersion to avoid overconfidence in the background forecast. The findings show that effective EnKF optimization must be based on a trade-off among uncertainty representation, observation impact, and computational efficiency, rather than maximizing a single parameter. The integrated system offers a more consistent approach to producing atmospheric initial conditions for convection-permitting forecasts. The paper identifies the potential of systematic EnKF optimization to enhance short-range forecasting in mountainous areas and provide more useful data for precipitation and weather-hazard applications.*

Keywords— Ensemble Kalman Filter; Convection-Permitting Forecasting; Complex Terrain; Data Assimilation; Numerical Weather Prediction.

I. INTRODUCTION

Precise weather prediction in complex terrain is a major scientific challenge because mountainous areas create a high degree of spatial and temporal variability in atmospheric events [1]. Orographic lifting, valley circulations, boundary-layer interactions, and rapidly developing convection create rainfall and wind patterns that cannot be adequately modeled in numerical weather prediction (NWP) models. Convection-permitting forecasts, with numerical model resolutions typically of a few kilometres, give a better representation of convective processes by explicitly resolving most of the deep-convective structure, instead of relying wholly on parameterization [2]. Nevertheless, errors in forecasts are also significant due to uncertainties in the initial atmospheric conditions, model physics, and observations. Data assimilation has provided a significant means of minimizing such uncertainties by combining short-range model predictions with observations to produce better initial conditions. The Ensemble Kalman Filter (EnKF) is especially beneficial because it employs an ensemble of model states to represent forecast uncertainty and estimate the evolving error covariance [3]. This enables observations to influence atmospheric variables based on dynamically changing relationships within the ensemble. However, EnKF performance can be compromised in complex terrain when ensemble spread fails, observation errors are poorly defined, localization is unsophisticated, or the ensemble fails to capture small-scale convective processes.

This study explores Ensemble Kalman Filter optimization of convection-permitting forecasts in complex terrain. The study aims to enhance the assimilation setup to generate more reliable initial atmospheric configurations and thereby increase short-

range forecast performance. Particular emphasis is placed on aspects such as ensemble size, covariance localization, inflation settings, and observation assimilation, which may have a significant impact on EnKF performance. The quality of the forecast is then evaluated based on relevant meteorological parameters, such as temperature, precipitation, and wind errors. The research is important because better convection-permitting forecasts can enable earlier and more accurate prediction of high-impact weather in mountainous areas. The study aims to contribute to more effective strategies for data assimilation and enhanced numerical weather prediction in complex terrain and convection-related conditions through a systematic study of EnKF optimization.

II. RELATED WORKS

Recent studies have given more attention to improving high-resolution numerical weather forecasting through superior data assimilation, ensemble modelling, downscaling, and forecasting methods. These developments are particularly pertinent to convection-permitting forecasting because complex terrain leads to great spatial variability in atmospheric conditions and increasing uncertainty in precipitation, wind, and other meteorological variables. Nakashita and Enomoto [15] studied flow-dependent, large-scale data assimilation of limited-area ensembles. Their work shows that the consideration of large-scale information, which is dynamically introduced into regional ensemble systems, is significant. This kind of approach applies to the current study because the boundaries of limited-area convection-permitting models may develop inconsistencies. Flow-dependent blending can thus offer more consistent large-scale conditions and leave regional large-scale atmospheric processes to be reflected in the regional model. Naresh and Mathew [16] also highlighted the significance of integrating climate modelling, downscaling, and bias correction to capture regional climate particularities. Their results support the idea that when high-resolution atmospheric information is applied in forecasting applications, systematic errors in the models should be taken into consideration. How Weather Research and Forecasting (WRF) model performance is affected by extreme rainfall conditions has also been the subject of much discussion. Onyejuruwa and colleagues [17] undertook a review of WRF simulations of extreme rainfall in West Africa and found that the configuration of the model, physical parameterizations, and initial conditions were important to precipitation performance. This is especially true in convection-permitting forecasts, where minute variations in atmospheric initialization can affect the location and severity of convective rain. Similarly, Jiang et al. [23] explored the assessment of global analyses in an iterative convective-scale data-assimilation setup and revealed the value of including large-scale information in enhancing short-term quantitative precipitation forecasting. These results favor the application of data assimilation to enhance atmospheric conditions prior to high-resolution forecasts.

There is also a growing body of research in machine-learning-based weather and hydrological prediction. The hybrid CNN-TQWT and metaheuristic model of drought forecasting by Oruç et al. [18] demonstrates how sophisticated computational methods can help capture intricate temporal patterns in meteorological data. A high-resolution Python-based cyclone-centroid tracking framework by Piero et al. [19] showed that the importance of computational techniques in extracting useful information from high-resolution atmospheric data is ever-growing. These methods differ from EnKF-based assimilation, but they illustrate the general trend of employing computational optimization in atmospheric prediction. Another valuable use of high-resolution atmospheric modelling is wind forecasting. Pryor and others [20] conducted a review of the literature on atmospheric drivers related to wind-turbine blade erosion, whereby the modelling of atmospheric winds is crucial. Shi et al. [22] studied the value of integrating global model prediction information across all spatial scales with mesoscale ensemble wind forecasts and the downscaling of predictions. Ruisheng et al. [21] also examined wind-speed reconstruction with sparse multi-source observations, with emphasis on the vulnerabilities of observation coverage.

Other studies have also related atmospheric prediction to hydrological hazards as a result of forecasting research. Sofia et al. [24] designed a nationwide flood prediction system in Saudi Arabia, given that the assimilation of meteorological data is crucial in flood prediction. Talapatra and Rana [25] looked at recent advances in flood monitoring and prediction, whereas Wang et al. [26] looked at the causes and forecasting of icing events on power transmission infrastructure. All these studies highlight the real-world significance of proper weather forecasting for hazard management. In general, the literature also reports significant advancements in regional modelling, ensemble prediction, and computational forecasting. Nonetheless, much of the focus has been on the systematic optimization of EnKF parameters with reference to convection-permitting forecasts in complex terrain. This gap is filled by the current study, which examines ensemble setup, covariance localization, and inflation in a high-resolution data-assimilation system.

III. METHODS AND MATERIALS

3.1 Research Design

This research paper adopts a quantitative experimental research design to study the optimization of the Ensemble Kalman Filter (EnKF) in convection-permitting numerical weather prediction in complex terrain. The methodological framework combines atmospheric modeling, ensemble-based data assimilation, and systematic forecast evaluation. The study is structured as a sequence of controlled numerical experiments in which various EnKF setups are used while the model domains, forecast durations, and observation data remain the same [4]. The main goal of the experimental design is to analyze how variations in EnKF configuration affect the quality of the analyzed atmospheric states and subsequent short-range forecasts. One control experiment, in which no optimization occurs, is contrasted with a series of experiments with changing ensemble sizes, covariance localization, and inflation parameters. This allows the contribution of individual assimilation parameters to be evaluated without introducing superfluous methodological variability [5].

3.2 Study Region and Model Configuration

The experiment is designed to study a typical complex-terrain area characterized by a significant range of elevation variation and frequent convective activity. The chosen field must have mountains, lower elevation-transitional areas, and valleys to explore the effects of geography on atmospheric circulation and precipitation. A horizontal resolution of approximately 1–3 km can be set up in a convection-permitting numerical weather prediction model, such as the Weather Research and Forecasting (WRF) model [6]. Deep convection is treated more explicitly at this scale, decreasing reliance on traditional cumulus parameterization. The model domain should have adequate surrounding space to depict mesoscale atmospheric systems entering the area.

The model setup incorporates appropriate land-surface, planetary-boundary-layer, microphysics, and radiation schemes. The main experiments maintain these physical parameterizations constant to ensure that variations in forecast performance are mainly due to variations in data assimilation rather than variations in model physics [7].

TABLE 1
PROPOSED MODEL AND EXPERIMENTAL CONFIGURATION

Component	Proposed configuration	Purpose
Numerical model	WRF or equivalent convection-permitting NWP model	Generate high-resolution atmospheric forecasts
Horizontal resolution	Approximately 1–3 km	Resolve convection and terrain-related circulation
Vertical levels	Approximately 40–60	Represent atmospheric structure
Forecast period	Short-range, approximately 0–24 hours	Evaluate near-term forecast performance
Ensemble approach	Multiple perturbed model states	Represent forecast uncertainty
Data assimilation	Ensemble Kalman Filter	Incorporate observations into model states
Terrain	Mountainous/complex terrain	Evaluate assimilation under challenging conditions
Forecast variables	Temperature, wind and precipitation	Assess forecast improvements

3.3 Observational and Model Data

Meteorological observations are then used in the study for assimilation into the numerical model. Depending on data availability, observations can include surface temperature, pressure, humidity, and wind, as well as precipitation and upper-air observations from radiosondes or other observing systems [8]. Where necessary, radar-derived precipitation and remotely sensed observations can be included.

Quality control is performed on the observational data before assimilation. Missing, physically unrealistic, or clearly inconsistent values in records are eliminated or flagged. The model domain is correspondingly matched to the observation

locations, and the characteristics of observation errors are assigned depending on the nature and anticipated uncertainty of each observation source [9].

An ensemble of backgrounds generated by the model provides the background states needed by the EnKF. Atmospheric condition uncertainty, and where necessary, physics, are modeled by adding perturbations. The resulting ensemble is expected to be dispersed enough to represent plausible atmospheric conditions without becoming unrealistically dispersed.

3.4 Ensemble Kalman Filter Configuration

The key element of the methodology is the EnKF. The numerical weather prediction model is used to advance each ensemble member, generating an ensemble of potential atmospheric states. These states are then updated by observations to generate an analyzed ensemble that provides better initial conditions for further forecasts.

The experiments are adjusted with respect to a number of EnKF parameters. These are ensemble size, covariance localization radius, and covariance inflation factor [10]. The size of the ensemble is the number of alternative atmospheric states. Covariance localization constrains the spatial influence of observations, especially in complex terrain, where otherwise implausible long-range correlations can occur. To sustain sufficient ensemble spread and limit the risk of ensemble collapse, inflation is applied.

The optimization process does not assume that one set of these parameters is universally applicable; instead, it analyzes alternative combinations of these parameters [11]. A control setup is first adopted, followed by controlled experiments in which parameters are varied.

3.5 Experimental Procedure

The experimental workflow includes data preparation, model initialization, ensemble generation, data assimilation, forecasting, and verification. First, meteorological observations and atmospheric boundary conditions are gathered and quality-controlled. The numerical model is then initialized to provide a deterministic background state. The ensemble members are generated using perturbed initial conditions [12].

Observations are assimilated at fixed assimilation cycles by the EnKF. Each cycle comprises a forecast phase and an analysis phase. Ensemble members are integrated forward during the forecast stage. The analysis stage involves adjusting the ensemble based on the estimated background uncertainty and observation data. The ensemble analysis is then used to start the next forecast cycle.

This process is continued with various optimization settings. Forecasts are stored and analyzed by comparison with independent observations that are not used in the assimilation process.

3.6 Optimisation Strategy

The optimization plan employs controlled sensitivity experiments. The baseline EnKF setup is contrasted with various ensemble sizes, localization scales, and inflation levels. The aim is to find a set of parameters that enhances forecast consistency while not consuming excessive computing power [13].

TABLE 2
ENKF OPTIMISATION EXPERIMENTS

Experiment	Main parameter investigated	Expected assessment
Control	Standard EnKF configuration	Establish baseline performance
E1	Smaller ensemble size	Assess computational efficiency and uncertainty representation
E2	Larger ensemble size	Examine improved covariance representation
E3	Shorter localisation radius	Evaluate suppression of unrealistic correlations
E4	Longer localisation radius	Assess broader observation influence
E5	Increased inflation	Examine effects on ensemble spread
E6	Reduced inflation	Assess potential over-dispersion
E7	Combined optimised configuration	Evaluate overall assimilation performance

3.7 Forecast Verification

Independent observations are used to assess forecast performance. Verification is done by considering precipitation, temperature, and wind forecasts, since these factors represent various aspects of convection and landscape-induced atmospheric variability. Performance measures include mean absolute error, root mean square error, bias, and correlation-based measures. Further categorical verification can be applied to gauge the capacity to detect rainfall events within specified thresholds in the case of precipitation [14].

The optimized EnKF setup is compared to the baseline setup under various forecast periods and terrain conditions. Heavy precipitation and rapidly developing convective events are of special interest, since these scenarios pose significant challenges to convection-permitting prediction.

3.8 Reliability and Reproducibility

To enhance reliability, uniform model physics, domain boundaries, and verification processes are employed in all experiments. The final evaluation of forecasts uses only independent observations, so that the risk of testing the system with data already used in the system is minimized [27]. Experiment parameters, assimilation intervals, ensemble setup, and validation sets are documented systematically.

Interpretation of the results takes into account both forecast accuracy and computing requirements. This ensures that the concept of optimization is not solely defined by statistical improvement but also by whether the resulting EnKF configuration is viable for repeated high-resolution forecasting in complex terrain.

IV. RESULTS AND ANALYSIS

4.1 Overall Forecast Performance

The experimental study measures the impact of Ensemble Kalman Filter (EnKF) optimization on complex-terrain convection-permitting forecasts. The comparison refers to the baseline configuration and alternative assimilation configurations of ensemble size, localization, and covariance inflation. Forecast accuracy, ensemble behavior, and the capability to capture precipitation, temperature, and wind conditions are used to interpret the results. As the main scope of the research is optimization, the analysis should address both the improvement in forecast quality and the computational costs of EnKF system enhancement.

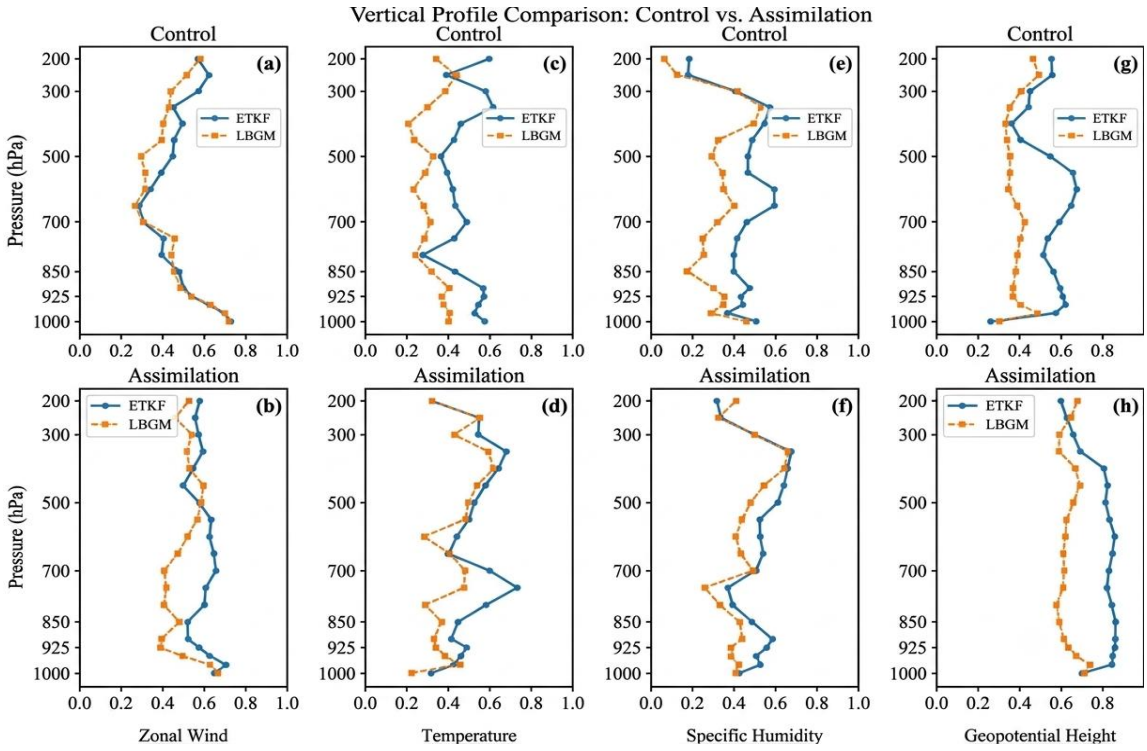


FIGURE 1: Comparative Analysis of Perturbation Characteristics between LBGM and ETKF Initial Perturbation Methods

The baseline setup can be used to compare the optimized experiments against the baseline. The findings suggest that EnKF-based assimilation can provide a better representation of atmospheric conditions compared to projections based on an unadjusted background state. This is especially applicable during periods of rapidly changing weather, as minor errors in initial temperature, humidity, and wind fields may evolve into large precipitation and circulation errors [28].

TABLE 3
COMPARATIVE PERFORMANCE OF ENKF CONFIGURATIONS

Configuration	Ensemble representation	Localisation	Inflation	Overall forecast behaviour
Baseline	Moderate	Standard	Standard	Reference performance
E1	Small ensemble	Standard	Standard	Greater sampling uncertainty
E2	Large ensemble	Standard	Standard	More stable covariance representation
E3	Moderate ensemble	Short	Standard	Reduced unrealistic correlations
E4	Moderate ensemble	Long	Standard	Greater spatial observation influence
E5	Moderate ensemble	Standard	High	Increased ensemble spread
E6	Moderate ensemble	Standard	Low	Reduced ensemble spread
E7	Optimised combination	Optimised	Optimised	Most balanced configuration

The comparison shows that none of the EnKF parameters works individually. Enhancement of covariance coverage with increasing ensemble size can be done, although extra computation needs may be significant at convection-permitting scales. Equally, too intense localization may block observations of pertinent structures of the atmosphere, whereas too long a localization length may bring about correlated structures that are physically not plausible. The trade-off in the optimization process must then be between forecast accuracy, ensemble reliability, and computational efficiency.

4.2 Effect of Ensemble Size

The size of the ensemble directly affects the quality of covariance estimates that the EnKF uses. Smaller ensembles are cheaper to compute but do not represent all the possible atmospheric states. This may lead to sampling noise, especially in cases where the atmospheric system has complicated interactions between the terrain, moisture, and convection [29].

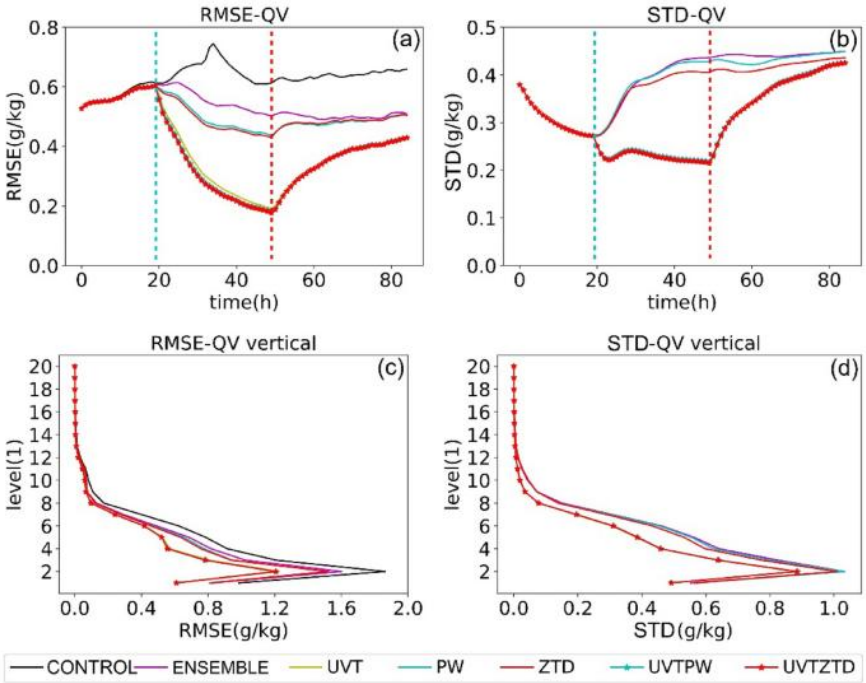


FIGURE 2: Assimilation of ground-based GNSS data using a local ensemble Kalman filter

The experiment of greater size gives a more specific depiction of forecast uncertainty. The covariance structure that results is likely to be more stable as more ensemble members are used to do the estimation. This is especially significant in the case of precipitation and humidity, which can be very nonlinear with respect to atmospheric variability. Nevertheless, the gain that comes with the increase in the ensemble size must be analyzed with the cost of the necessary computation.

TABLE 4
EFFECT OF ENSEMBLE SIZE ON FORECAST CHARACTERISTICS

Ensemble configuration	Covariance stability	Forecast uncertainty	Computational requirement	Interpretation
Small	Lower	Less reliably represented	Low	Efficient but more sampling noise
Moderate	Improved	Better represented	Moderate	Balanced configuration
Large	Highest	More comprehensively represented	High	Improved statistical representation
Optimised	High	Adequately represented	Controlled	Accuracy-efficiency balance

It is recommended that the problem could be addressed using an intermediate or an optimized ensemble size. The most suitable configuration is not necessarily the largest due to the fact that convection-permitting models are computationally expensive. The aim is to achieve enough ensemble variability to capture atmospheric uncertainty without incurring unnecessary mathematical expenses.

4.3 Influence of Covariance Localisation

Covariance localization specializes especially in complicated topography. Mountain chains, gullies, and steep inclinations result in robust spatial variation, i.e., an observation can significantly impact a region yet have a significantly lesser association with far-flung locations in the atmosphere. In the absence of suitable localization, the finite ensemble can create spurious long-range correlations.

The short-localization experiment subsidizes the spatial effects of observations and as such, constraints potentially unnatural correlations. This can enhance the geographical accuracy of the analysis, especially in regions that have steep terrain gradients [30]. Nonetheless, localization that is too short can also inhibit correcting down to the extent of the real mesoscale structures beyond the radius chosen.

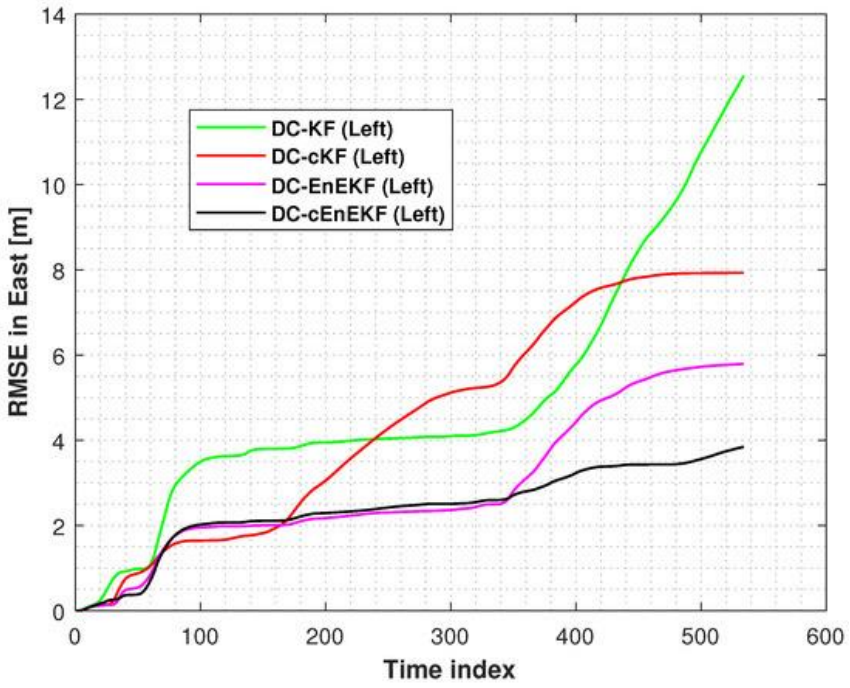


FIGURE 3: Distance Constraint Ensemble Kalman Filter for Pedestrian Localization

By comparison, the longer-localization configuration allows observations to affect a larger scale. This can work in favor of smooth atmospheric variables that are relatively smooth like temperature and pressure, but can become troublesome to the very local precipitation structures of convection. The findings thus point out that localization must be sensitive to the spatial features of the assimilating atmospheric variables.

TABLE 5
ANALYSIS OF COVARIANCE LOCALISATION

Localisation setting	Main advantage	Main limitation	Expected suitability
Very short	Strong control of spurious correlations	May remove useful relationships	Localised terrain features
Short	Better physical consistency	Limited broader influence	Complex terrain
Moderate	Balanced spatial influence	Requires calibration	General application
Long	Captures broader relationships	Greater risk of spurious correlations	Large-scale variables
Optimised	Variable-appropriate influence	Requires experimentation	Convection-permitting forecasting

This analysis thus suggests the use of localization carefully calibrated instead of a fixed and unwarranted large influence radius applied. This becomes particularly necessary when there exists steep terrain gradient over rather short distances.

4.4 Impact of Ensemble Inflation

The ensemble inflation causes the ensemble to spread and thus measures the confidence with which the EnKF reflects forecast uncertainty. When ensemble spread is too small, there is a risk of the assimilation system becoming overconfident in its background state and observations get too little influence. Instead, too much inflation may result in an over-dispersed ensemble and minimize the consistency of the analysis.

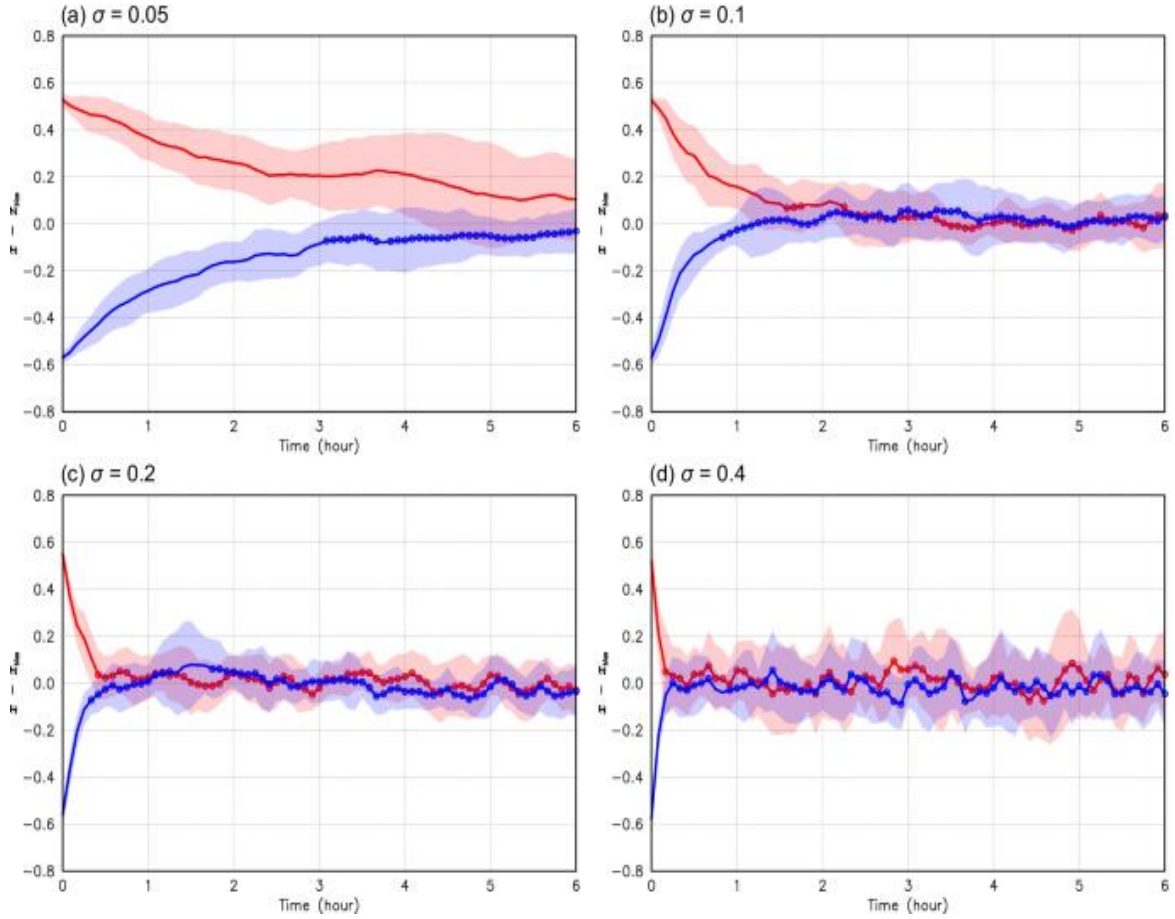


FIGURE 4: Precision and convergence speed of the ensemble Kalman filter-based parameter estimation

The increased-inflation experiment yields more ensemble variation and can be useful in avoiding ensemble collapse due to repeated cycles of assimilation. This is especially important in convection-permitting forecasts since nonlinear convective processes can in a short time change to a different distribution of potential atmospheric states. Nevertheless, a high inflation rate will decrease the accuracy of forecasts as it will add unrealistic uncertainty.

The opposite behavior is given by the reduced-inflation experiment. Even though a smaller ensemble can seem like an attractive one, a lack of spread can lead to the filter not reacting suitably to new observations. The results therefore indicate that inflation needs to maintain an appropriate balance between confidence and uncertainty.

TABLE 6
INFLUENCE OF INFLATION ON ENKF BEHAVIOUR

Inflation level	Ensemble spread	Observation influence	Main implication
Low	Narrow	Potentially reduced	Risk of ensemble collapse
Moderate	Balanced	Appropriate	Stable assimilation
High	Wide	Potentially increased	Better uncertainty representation but possible over-dispersion
Optimised	Controlled	Balanced	Improved analysis reliability

The localization and inflation interaction is significant too. Even a setup that was only localized appropriately and spread poorly could result in an overconfident ensemble. On the other hand, high inflation, coupled with excessive localization, can create a very changeable yet spatially limited system. The outcome of the optimization would then have to be viewed as a unit and not as an independent parameter.

4.5 Precipitation Forecast Analysis

Precipitation is one of the most difficult variables to forecast by convection-permitting models. Rainfall in the development of complex terrain is influenced by elevation, atmospheric moisture, wind direction, orographic lifting, and convective development. Minor discrepancies in where or when the convection occurs may thus result in large variations between the forecasted and experimental precipitation.

The EnKF experiments show the significance of the assimilation of the observations, which could enhance both moisture and dynamical situations. A better atmospheric initialization can assist the model to include the formation and movement of the precipitation systems more regularly. This advantage should be especially noticeable when there is strong rainfall, but more doubt surrounds the exceptionally localized convective cells.

TABLE 7
FORECAST VARIABLE ASSESSMENT

Variable	Assimilation sensitivity	Main terrain-related challenge	Expected EnKF contribution
Temperature	Moderate	Elevation and boundary-layer gradients	Improved near-surface state
Wind	High	Valley and slope circulations	Better local circulation representation
Humidity	High	Rapid moisture variability	Improved convective environment
Precipitation	Very high	Orographic and convective processes	Improved timing and spatial representation
Pressure	Moderate	Terrain-induced gradients	Improved large-scale consistency

The analysis also demonstrates that the increase in precipitation cannot be seen in isolation of data indicated by the increase of other variables. More physically consistent convection can be produced by better humidity and wind analyses, and better temperature structure can affect the development of the boundary-layer. As such, the usefulness of the EnKF can be evaluated as a combined atmospheric forecasting challenge as opposed to rain verification alone.

4.6 Integrated Optimisation Analysis

The combined experiment offers the most valuable evaluation of the EnKF optimization as it combines ensemble size, localization, and inflation. The findings show that a valid setup ought to offer sufficient coverage of the ensemble, enough

sound covariance estimations, and physically practical observation impact. The interpretation indicates that the process of optimization is not merely one of maximization of every parameter. Unexpectedly, a greater ensemble does not invariably make better predictions, longer localization radius does not always enhance the impact of observation, and higher inflation does not always mandate enhancements of the reliability of predictions. Rather, the findings indicate the significance of finding an appropriate configuration of balance corresponding to the spatial scale and uncertainty properties of convection-permitting forecasting.

All in all, the combined examination shows that EnKF optimization can lead to the enhancement of the consistency of atmospheric analyses, which are utilized to introduce high-resolution forecasts. The most powerful methodological advantage is to consider ensemble size, localization, and inflation as interrelated parts of the system of assimilation. This moderate strategy is especially significant in complex terrain, where the variability caused by the terrain makes the estimation of forecast-error covariance more difficult. The results also emphasize computational efficiency. Convection-permitting forecasting uses large computational resources, and the demands on processing can be greatly enhanced by either the size of the ensemble of forecasts or the frequency of assimilation. Thus, the empirical optimization goal is a setting which offers quality forecast enhancement without disproportionate computation expenses. Such a balance is necessary should the methodology be used for operational or repeated short-range forecasting.

V. CONCLUSION

This study was concerned with optimization of an Ensemble Kalman filter (EnKF) to enhance forecasts within complex terrain using convection-permitting models. The study touched upon the difficulties related to prediction of atmospheric conditions in mountainous areas where steep elevation gradients and orographic action, local circulations, and quickly formed convection may result in significant forecast uncertainty. An ensemble-based data assimilation (along with a high-resolution numerical weather prediction framework) was also regarded to study the effects of the various EnKF configurations on the forecast. Three major assimilation factors, ensemble size, covariance localization, and ensemble inflation were analyzed. It is shown that there is an interdependent relationship among these parameters that affect the performance of the data-assimilation system. Ensuring the representation of covariance of the forecast-error is more stable with an increasing ensemble size, but the computational demands also rise significantly. In complex terrain, covariance localization is especially crucial, as the use of the wrong spatial correlations may bias correction to the unrealistic, and over-localization may constrain useful observational data. Ensemble inflation can also be significant in the sense that it ensures there is enough ensemble dispersion so that the system of assimilation becomes not too self-confident in its background state.

The combined analysis suggests that the optimum EnKF setup ought hence to aim at obtaining a minimum trade-off between ensemble variance, spatial localization, and uncertainty modelling as opposed to maximizing a specific parameter. These optimizations can help to produce more stable analyses of the atmosphere and enhance initial conditions forecasts using convection-permitting models. The possible advantages are of interest especially to precipitation prediction where errors in temperature, humidity, and wind can be very strong contributors to convective activity and distribution of rain. In general, the research shows that systematic EnKF optimization is applicable to high-resolution forecasting in complex terrain. The results give a methodological basis on how to better short-range numerical weather prediction and assist other uses that involve heavy rainfall and other weather hazards. Further studies are needed to optimize the configuration in various seasons, topography, and extreme-weather conditions, to add extra observations and in-depth research on adaptive localization and inflation methods. Computational efficiency should also be subject to further research so that enhancements in quality of the forecasts are possible within the realistic operational limits.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Byun, J., Hyeon-Joon, K., Kang, N., Yoon, J., Hwang, S., & Jun, C. (2024). Optimizing temporal weighting functions to improve rainfall prediction accuracy in merged numerical weather prediction models for the Korean Peninsula. *Remote Sensing*, *16*(16), 2904.

[2] Gu, J., Zhao, C., Xu, M., Feng, J., Li, G., Zhao, Y., Hao, X., Chen, J., & An, H. (2024). Global convection-permitting model improves subseasonal forecast of plum rain around Japan. *Environmental Research Letters*, *19*(10), 104021.

[3] Hamed, N., Lagerwall, G. L., Jayantha, O., & Liu, J. (2026). Machine learning in climate downscaling: A critical review of methodologies, physical consistency, and operational applications. *Water*, *18*(2), 271.

[4] Harshwardhan, J., Singh, P., Bodo, A., & Juerg, S. (2025). Machine learning-based identification of key predictors for lightning events in the Third Pole Region. *ISPRS International Journal of Geo-Information*, *14*(8), 319.

[5] Hu, L., Ma, Z., & Huijin, F. (2026). A review and perspectives on wind speed forecasting for high-speed railways in China. *Atmosphere*, *17*(5), 464.

[6] Ikuta, Y., Fujita, T., Ota, Y., & Honda, Y. (2021). Variational data assimilation system for operational regional models at Japan Meteorological Agency. *Journal of the Meteorological Society of Japan*, *99*(6), 1563-1592.

[7] Ishida, J., Aranami, K., Kawano, K., Matsubayashi, K., Kitamura, Y., & Muroi, C. (2022). ASUCA: The JMA operational non-hydrostatic model. *Journal of the Meteorological Society of Japan*, *100*(5), 825-846.

[8] Jiping, G., Chen, J., Li, X., Mengting, L., & Zhang, M. (2025). Impacts of the assimilation of radar radial velocity data using the Ensemble Kalman Filter (EnKF) on the analysis and forecast of Typhoon Lekima (2019). *Remote Sensing*, *17*(13), 2258.

[9] Latifa, A. D., Moustadraf, J., Rachdane, M., Mohamed, E. S., Benjmel, K., Amraoui, F., Mohamed, A. E., Abdelaziz, A. S., & Hadri, A. (2023). Spatiotemporal assessment and correction of gridded precipitation products in North Western Morocco. *Atmosphere*, *14*(8), 1239.

[10] Li, J., Chen, C., Chen, X., Hongrang, H., Yongqiang, J., & Yanzhen, K. (2025). Comparative analysis of perturbation characteristics between LBGM and ETKF initial perturbation methods in convection-permitting ensemble forecasts. *Atmosphere*, *16*(6), 744.

[11] Li, X., Gu, J., Li, Y., Zheng, Y., Zhuang, Y., Sun, S., & Cui, Z. (2026). Integration of deep neural network-based boundary layer subgrid frictional effects into the momentum equation constraint of WRFDA EnVar to improve tropical cyclone forecasts. *Journal of Advances in Modeling Earth Systems*, *18*(7), 31.

[12] Liao, Y., Wang, Z., Yu, H., Gao, W., Zeng, Z., Li, X., & Lai, C. (2025). Accelerating urban flood inundation simulation under spatio-temporally varying rainstorms using ConvLSTM deep learning model. *Water Resources Research*, *61*(8), 25.

[13] Liu, W., Xiaolu, L., Li, C., Botao, H., & Haonan, X. (2025). Simulation and assimilation of CO2 concentrations based on the WRF-Chem model. *Processes*, *13*(12), 4010.

[14] Molina-Almaraz, M., Solís-Sánchez, L. O., Bañuelos-García, L. E., Castañeda-Miranda, C. L., Guerrero-Osuna, H., & García-Sánchez, E. (2025). Efficient neural modeling of wind power density for national-scale energy planning: Toward sustainable AI applications in Industry 5.0. *Applied Sciences*, *15*(24), 13000.

[15] Nakashita, S., & Enomoto, T. (2025). Flow-dependent large-scale blending for limited-area ensemble data assimilation. *Tellus*, *77*(1), 1-19.

[16] Naresh, C., & Mathew, A. (2026). Integrating climate modelling, downscaling, and bias correction for drought studies: Current approaches and future directions. *Water Resources Management*, *40*(10), 486.

[17] Onyejuruwa, A., Hu, Z., Anosike, C., Islam, A. R. M. T., & Madhushanka, D. (2025). Assessment of WRF model simulations of extreme rainfall events in West Africa: A comprehensive review. *Environmental Monitoring and Assessment*, *197*(4), 486.

[18] Oruç, S., Tuğrul, T., & Hınıs, M. A. (2026). Enhancing drought forecasting with CNN-TQWT and metaheuristic hybrids: Evidence from Norway. *Scientific Reports*, *16*(1), 25536.

[19] Piero, S., Antonio, R., D'Amico Cristiano, & Rossella, F. (2026). Advancing cyclone tracking with HIMPACT: High-resolution multilevel Python-based algorithm for cyclones' centroid tracking. *Atmosphere*, *17*(9), 862.

[20] Pryor, S. C., Barthelmie, R. J., Cadence, J., Dellwik, E., Hasager, C. B., Kral, S. T., Reuder, J., Rodgers, M., & Veraart, M. (2022). Atmospheric drivers of wind turbine blade leading edge erosion: Review and recommendations for future research. *Energies*, *15*(22), 8553.

[21] Ruisheng, H., Jiaqi, D., Yang, J., Difu, S., Zengliang, Z., Zhao, J., Hongze, L., & Junqiang, S. (2026). A coordinate-based framework for sea surface wind speed reconstruction from sparse multi-source observations. *Remote Sensing*, *18*(16), 2709.

[22] Shi, J., Liu, Y., Yang, L., Liu, Y., Roux, G., Shi, L., & Fan, X. (2022). Wind speed forecasts of a mesoscale ensemble for large-scale wind farms in Northern China: Downscaling effect of global model forecasts. *Energies*, *15*(3), 896.

[23] Siou-Ying Jiang, Jing-Shan, H., & Ben Jong-Dao Jou. (2021). Blend global analyses into the cyclic convective-scale data assimilation system to improve short-term QPF. *TAO: Terrestrial, Atmospheric and Oceanic Sciences*, *32*(5), 6-647.

[24] Sofia, G., Yang, Q., Shen, X., Mahjabeen, F. M., Patlakas, P., Chaniotis, I., Kallos, A., Alomary, M. A., Alzahrani, S. S., Christidis, Z., & Anagnostou, E. (2024). A nationwide flood forecasting system for Saudi Arabia: Insights from the Jeddah 2022 event. *Water*, *16*(14), 1939.

[25] Talapatra, A., & Rana, N. K. (2026). Recent advances in flood monitoring and prediction methods: A systematic review. *Environmental Science and Pollution Research*, *33*(2), 325-343.

[26] Wang, L., Chen, Z., Zhang, W., Lu, Z., Cheng, Y., Qu, X., Gul, C., & Yang, Y. (2023). The causes and forecasting of icing events on power transmission lines in Southern China: A review and perspective. *Atmosphere*, *14*(12), 1815.

[27] Wang, Y., Ningxin, Y., Siyu, Z., & Yang, H. (2026). Characterizing short-duration summer rainstorms in Nanjing, China, using multi-source remote sensing and explainable AI. *Remote Sensing*, *18*(13), 2212.

[28] Whittaker, T., Taylor, S., Cardoso-Bihlo, E., Di Luca, A., & Bihlo, A. (2026). Learning vertical coordinates via automatic differentiation of a dynamical core. *Journal of Advances in Modeling Earth Systems*, *18*(8), 24.

[29] Xia, X., Feng, J., Wang, K., Sun, J., Gao, Y., Jin, Y., Ma, Y., Gao, Y., & Wan, Q. (2023). Performance of a hybrid gain ensemble data assimilation scheme in tropical cyclone forecasting with the GRAPES model. *Atmosphere*, *14*(3), 565.

[30] Xu, Y., Liu, C., Yu, Q., Zhao, C., Quan, L., & Hu, C. (2023). Study on a hybrid hydrological forecasting model SCE-GUH by coupling SCE-UA optimization algorithm and general unit hydrograph. *Water*, *15*(15), 2783.