Dear Editor-in-Chief, Associate Editor, and Reviewers,

We have completed the revision of the paper entitled "Ensemble Kalman–Guided Model Predictive Path Integral Control for Spatially Localized Suppression of Extremes in Chaotic Geophysical Flows" (Manuscript ID: [10.5194/egusphere-2026-419]) and would like to submit the revised version. The paper has been revised according to the comments from the reviewers.

Please find below our detailed answers to the comments of the reviewers. We sincerely thank the Editor-in-Chief, Associate Editor, and Reviewers for their careful reading and constructive suggestions. We are grateful for all comments, which helped us improve the manuscript substantially.

Yours sincerely,

H. Kuroki, K. Hashimoto, Y. Uehara, Y. Sawada, D. Le, and M. Minamide

Reviewer 1's Comments

This manuscript presents a novel and elegant hybrid control strategy, termed EKG-MPPI, aimed at spatially localized interventions in chaotic geophysical flows. By combining the statistical data assimilation capabilities of Ensemble Kalman Control (EnKC) with the nonlinear, derivative-free optimization of Model Predictive Path Integral (MPPI) control, the authors provide a compelling framework for "Control Simulation Experiments". From a systems and control engineering perspective, using the EnKC to generate a sparse, physics-informed prior to guide the sampling distribution of the MPPI is a brilliant approach to solving the sample inefficiency inherent to high-dimensional state spaces. The method is entirely matrix-free and avoids the computation of adjoint models or gradients, which is highly desirable for chaotic, non-linear atmospheric models. However, while the theoretical hybridization is sound, the current implementation simplifies the temporal complexity of the control problem to a single-step actuation. This reduces the framework to a static spatial optimization problem rather than true multi-step trajectory optimization. Furthermore, the comparison with baseline methods lacks rigorous hyperparameter tuning, which somewhat weakens the performance claims. The manuscript is a strong proof-of-concept, but addressing the specific points below would significantly improve its rigor and impact.

Answer: We thank the Referee very much for the careful reading of our paper and provide positive evaluations. We have revised the manuscript in accordance with the reviewer's comments to further improve its clarity and presentation. Detailed responses to the specific comments and questions are provided below.

Q1: *Single-Step Actuation vs. Sequential Trajectory Optimization: The authors explicitly state that the control perturbation is applied only at the first rollout step, with no further input provided during the remaining prediction horizon (T_{MPPI}). Because the problem is reduced to optimizing a single initial impulse, the use of MPPI, a method designed for sequential, multi-step trajectory optimization, seems algorithmically oversized. The authors should discuss why MPPI was chosen over standard derivative-free black-box optimization algorithms (e.g., CMA-ES or Bayesian Optimization) which are highly efficient for static, single-step parameter search. If the intention is to lay the groundwork for future multi-step actuation, this should be stated more explicitly as the primary justification for using the MPPI formalism.*

Answer: We thank the Referee for this important point. We agree that the current implementation applies the perturbation only at the first rollout step, with zero input thereafter. Thus, at each control decision, the optimization reduces to choosing a single localized impulse. Indeed, this design was motivated by the current CSE setting: in chaotic flows, even a small localized perturbation can produce a meaningful downstream effect over the prediction horizon, while restricting the actuation to a single step reduces the effective control dimension and improves sampling efficiency.

Nevertheless, we agree that a clearer justification for using the MPPI formalism is necessary. In the revised manuscript, we emphasized that MPPI still offers several advantages in this setting:

- Parallel sampling: MPPI evaluates a batch of rollouts whose costs and weights can be computed in parallel. This is advantageous when forward simulations are the dominant cost and can be parallelized on multi-core CPUs or GPUs.
- Avoiding sequential acquisition loops: Bayesian optimization typically involves a sequential propose-evaluate loop, which can be less straightforward for real-time implementation.
- Extensions to multi-step control: As the Referee points out, one motivation for adopting MPPI is that it can be generalized to multi-step control sequences and multiple actuators within the same algorithmic pipeline. The single-step setting in this manuscript is a proof of concept toward that goal.

Following the Referee's suggestion, we additionally performed a BO-based comparison for the present single-step optimization problem using a standalone BO baseline. In our BO implementation, the Expected Improvement (EI) acquisition function was adopted, and seven BO iterations were carried out at each control decision. Although seven iterations represent a relatively modest BO budget, we chose this setting because the resulting per-decision computation time was already substantially larger than that of the proposed method.

The results show that EKG-MPPI outperforms BO in the key metrics:

- Number of exceedances ($X > 12$): 7622 ± 146 (EKG-MPPI) vs. 8299 ± 255 (BO)
- Mean input magnitude: 0.166 ± 0.002 (EKG-MPPI) vs. 0.340 ± 0.004 (BO)
- Mean computation time: $(3.87 \pm 1.69) \times 10^{-4}\,\mathrm{s}$ (EKG-MPPI) vs. $(6.64 \pm 0.82) \times 10^{-2}\,\mathrm{s}$ (BO)

These results indicate that, under the present single-step setting and computational budget, the MPPI-based approach is both more effective and more efficient than the standalone BO baseline. For completeness, we also report in Appendix A the corresponding comparison between BO and vanilla MPPI, which is included only as a supplementary reference for single-step control (following another referee's suggestion).

Changes in the manuscript: In Section 3.4, we added a statement emphasizing why we use MPPI rather than other optimization methods, such as Bayesian optimization. Furthermore, Section 4.1.2 provides a comparison of the intervention effect and intervention magnitude between the proposed EKG-MPPI method and intervention optimization based on Bayesian optimization, while Table 3 in Section 4.1.5 summarizes the computational time. Furthermore, in Appendix A, we present the experimental results comparing BO and vanilla MPPI.

Q2: *Fairness in Baseline Comparison (Hyperparameter Optimization): In Section 4.1.1, the proposed EKG-MPPI is compared against a "vanilla MPPI" baseline. However, the setup for the vanilla MPPI utilizes arbitrarily fixed hyperparameters without any apparent tuning. Specifically, the authors set the temperature parameter to* $\lambda = 0.7$ *for the vanilla MPPI, whereas EKG-MPPI operates with* $\lambda = 0.1$. *Furthermore, the vanilla MPPI relies on a completely static sampling distribution (variance of 0.5 for magnitude and 20.0 for location). Because EKG-MPPI inherently adapts both the mean and variance of its sampling distribution using the EnKC output, it naturally possesses a massive advantage. To ensure a rigorous comparative analysis, the baseline MPPI should ideally undergo a degree of hyperparameter tuning or at least a sensitivity analysis similar to the one performed for EKG-MPPI in Figure 3 to firmly demonstrate that its poorer performance is due to the lack of EnKC guidance rather than sub-optimal static parameters.*

Answer: We thank the Referee for this important comment. We agree that it is important to compare the methods under same values of the hyperparameter λ, and also to compare the proposed method with a baseline that adaptively updates MPPI parameters such as the sampling mean. Specifically, we construct an additional BO-informed MPPI baseline, denoted BO-MPPI, in which Bayesian optimization is used to estimate prior information for the intervention location and magnitude, and these estimates are then supplied to MPPI. For BO-MPPI, the BO objective balances the predicted threshold exceedance over the rollout horizon and the intervention magnitude, where the parameter alpha controls the trade-off between control effectiveness and actuation cost. Here, alpha is the BO-side trade-off parameter in the objective function, whereas λ is the MPPI temperature parameter. These two parameters play different roles and are varied separately in our experiments.

To ensure a fair comparison with EKG-MPPI, BO-MPPI uses the same MPPI temperature λ *and the same sampling variances for intervention location and magnitude as EKG-MPPI.* For each fixed value of λ, we run BO-MPPI with $\alpha = 100, 10, 1, 0.1, 0.01$. The BO budget is fixed at 7 iterations at each control decision, since even with this modest budget the BO-based baseline already requires a substantially longer computation time than the proposed method. We then evaluate the resulting Lorenz-96 simulations in terms of the number of threshold exceedances and the average intervention magnitude.

Separately, we repeat this comparison for $\lambda = 0.1, 1, 10$. Across all tested values of λ, the EKG-MPPI result lies below the empirical lower envelope (Pareto frontier-like trade-off curve) of the BO-MPPI results obtained by varying α. These additional results show that the advantage of EKG-MPPI is robust to the choice of λ and is not attributable to a single hyperparameter setting.

Changes in the manuscript: A comparison experiment between BO-MPPI described above and the proposed method, EKG-MPPI, has been added to Section 4.1.3.

Q3: *Computational Scalability and Profiling: The results in Table 2 show that EKG-MPPI is the most computationally expensive method tested. While the authors correctly point out that MPPI rollouts are "embarrassingly parallel" and benefit directly from GPU acceleration, a brief computational profiling of the EKG-MPPI execution time would be highly valuable. Specifically, providing a breakdown of the time spent computing the EnKC prior versus the time spent evaluating the MPPI rollouts would clarify where the true bottleneck lies. Given the authors' stated intention in the conclusion to scale this approach to multi-actuator and multi-step scenarios, this profiling would help readers assess the operational feasibility of scaling the framework.*

Answer: We thank the Referee for this useful suggestion and agree that such a runtime breakdown is important for identifying the main computational bottlenecks and assessing scalability. We therefore measured the computation time separately for the EnKC-prior stage and the MPPI rollout/weight stage, and we have added this information to the revised manuscript.

Specifically, in the Lorenz-96 experiment with $K_{MPPI} = 40$, the mean runtime per control decision was $(6.30 \pm 5.23) \times 10^{-5}$ s for the EnKC-prior computation and $(3.24 \pm 1.17) \times 10^{-4}$ s for the MPPI rollout/weight computation. The total mean runtime was therefore 3.87×10^{-4} s per decision step, with approximately 16% of the cost attributable to the EnKC-prior stage and 84% to the rollout and weight stage. This indicates that the main computational bottleneck is the rollout evaluation rather than the EnKC prior itself. Moreover, the total runtime remains well below the sampling interval used in the experiment, supporting the computational feasibility of the proposed method in the current online-control setting.

Changes in the manuscript: We report in Section 4.1.5 a decomposition of the computational time of EKG-MPPI into the EnKC component and the MPPI component.

Q4: *Equation notations: In Section 3.2, Step 3, the embedding functions f_{loc}, g_{loc}, f_{mag}, and g_{mag} are introduced. It would be helpful to the reader to add a brief sentence clarifying their exact mathematical definitions or references earlier in the text, before they are explicitly defined for the experiments in Equations (39) and (50).*

Answer: We thank the Referee for this helpful comment. In the revised manuscript, we clarified the roles of these embedding functions before their experiment-specific definitions are introduced. Specifically, f_{loc} and f_{mag} map the EnKC output to the mean parameters of the sampling distribution for the control location and magnitude, respectively, while g_{loc} and g_{mag} specify the corresponding variance/covariance structure used in the MPPI sampling procedure. Their explicit forms for the present experiments are then provided in Eqs. (39) and (52) (Eq.(50) in the original manuscript).

Changes in the manuscript: We added a brief explanatory sentence for f_{loc}, f_{mag}, g_{loc}, and g_{mag} immediately after Eq. (27).

Reviewer 2's Comments

This manuscript introduces EKG-MPPI, a hybrid control approach for localized suppression of extremes in chaotic geophysical flows. The method combines EnKF for state estimation, EnKC for constructing candidate perturbations, and MPPI to improve on them through nonlinear refinement. The approach is tested on the Lorenz–96 and SQG models, where the authors report reductions in extreme-event, together with comparable or smaller control inputs relative to EnKC. Overall, the paper is well motivated, the methodological idea is interesting, and the combination of EnKC and MPPI is intuitive and potentially useful for this class of problems. The paper shows clearly how the two methods are coupled in practice. The numerical experiments on Lorenz–96 and SQG support the claim that the method works and can improve on EnKC in the examples considered. However, the empirical advantage remains somewhat difficult to assess quantitatively, because the reported performance results in both numerical sections are small and presented without uncertainty estimates. This does not reduce the interest of the proposed method, but it does make the strength of the performance claim harder to evaluate in the present version.

Answer: We thank the Referee very much for the careful reading of our paper and provide positive evaluations. We have revised the manuscript in accordance with the reviewer's comments to further improve its clarity and presentation. Detailed responses to the specific comments and questions are provided below.

Q5: *In both the Lorenz-96 and SQG sections, the conclusions are reported without uncertainty estimates. For Lorenz-96, Table 1 reports extreme event suppressions; for SQG, the conclusions are drawn from raw peak wind and mean input comparisons. But in both cases no confidence intervals or variability estimates are provided for these performance diagnostics. Since for both metrics the improvement is moderate (less than 10%), it is then difficult to assess how significant the improvement of EKG-MPPI over EnKC is.*

Answer: We thank the Referee for pointing out the lack of uncertainty quantification in the original manuscript. In the revised manuscript, we addressed this concern by performing repeated experiments with different random seeds and reporting the mean and standard deviation of the relevant performance metrics.

For the Lorenz–96 experiments, we repeated the simulations 10 times with different random seeds, using the temperature parameter $\lambda = 0.25$ for EKG-MPPI. We evaluated both the number of threshold exceedances (i.e., the extreme-event count) and the mean control-input magnitude. The results are as follows:

- Number of exceedances ($X > 12$): 7622 ± 146 (EKG-MPPI) vs. 7820 ± 142 (EnKC)
- Mean input magnitude: 0.166 ± 0.002 (EKG-MPPI) vs. 0.191 ± 0.001 (EnKC)

These results indicate that EKG-MPPI consistently reduces the number of extreme events relative to EnKC while also requiring a smaller control input. Moreover, the differences in the mean values are larger than the corresponding standard deviations, suggesting that the observed improvements are not solely attributable to stochastic variability.

For the SQG experiments, we conducted five independent runs with different random seeds for each scenario, and computed the mean and standard deviation of the maximum wind speed and the control-input magnitude. For brevity, we present here only the results for Scenario 5; the results for all scenarios are provided in the revised manuscript:

- Maximum wind speed: 3.341 ± 0.050 (EKG-MPPI) vs. 3.403 ± 0.005 (EnKC)
- Control input magnitude: 0.0135 ± 0.0038 (EKG-MPPI) vs. 0.0137 ± 0.0015 (EnKC)

In this case as well, EKG-MPPI yields a lower mean peak wind speed than EnKC while maintaining a comparable level of control effort. Although the variability in the control input is somewhat larger for EKG-MPPI, the reduction in peak wind speed is consistently observed across runs.

Overall, these results suggest that the improvements of EKG-MPPI over EnKC are robust across random seeds and are unlikely to be attributable solely to stochastic variability.

Changes in the manuscript: We added uncertainty estimates (mean $\pm$ standard deviation over repeated runs with different random seeds) for all comparison experiments; see Tables 1–3 and Fig. 4.

Q6: *In the Lorenz-96 experiments, EKG-MPPI uses $\lambda = 0.1$ and an EnKC-informed proposal distribution, while vanilla MPPI uses $\lambda = 0.7$ and uninformed Gaussian sampling for magnitude and location. As a result, the reported difference does not isolate the effect of the proposed guidance alone. I suggest clarifying that vanilla MPPI is included as a simple reference only, and that the principal comparison is with EnKC.*

Answer: We thank the Referee for this important comment. We agree that the original comparison with vanilla MPPI did not isolate the effect of the proposed guidance alone, since vanilla MPPI differs from EKG-MPPI not only in the absence of EnKC-based guidance but also in the temperature parameter λ and the proposal distribution.

In the revised manuscript, we therefore clarified that vanilla MPPI is included only as a simple reference, and that the principal comparison in this paper is between EnKC and the proposed EKG-MPPI method, since EKG-MPPI is designed as an EnKC-based nonlinear refinement. To avoid overinterpreting the vanilla MPPI results, we removed them from the main performance table and revised Table 1 to focus exclusively on the comparison between EnKC and EKG-MPPI.

Moreover, to provide more informative baseline comparisons, we additionally introduced two baselines in the revised manuscript: a standalone Bayesian optimization (BO) method and a BO-informed MPPI method. The standalone BO baseline directly optimizes the location and magnitude of the single-step perturbation under the same setting, and therefore serves as a relevant derivative-free comparator for the present low-dimensional problem. The BO-informed MPPI baseline, in turn, uses BO to provide adaptive prior information for MPPI, thereby offering a stronger adaptive baseline than vanilla MPPI with a fixed uninformed proposal. These additional baselines allow us to distinguish more clearly between the benefit of adaptive proposal guidance and that of the optimizer itself. We also revised the corresponding discussion to make the role of each baseline explicit and to avoid overstating the significance of the vanilla MPPI comparison.

Changes in the manuscript: We clarified that vanilla MPPI is included only as a simple reference, revised Table 1 to focus on the comparison between EnKC and EKG-MPPI, and added BO and BO-informed MPPI baselines in Section 4.1.3.

Q7: *Figure 4's caption states that the evaluation metric is the percentage difference ((EKG - EnKC)/EnKC) × 100, but the figure itself shows raw maximum wind speed and raw mean input magnitude by scenario. Either this text should be removed and replaced with useful information about the figure, or the results about the percentage difference should be added.*

Answer: We thank the Referee for pointing out this inconsistency. The original caption was misleading. In the revised manuscript, we correct the caption of Figure 4 so that it directly describes the quantities actually shown in the figure, i.e., the raw maximum wind speed and the raw mean input magnitude for each scenario. We remove the incorrect statement about the percentage difference.

Changes in the manuscript: We have clarified in the caption of Figure 4 that the evaluation metrics are the raw values of wind speed and intervention magnitude.

Q8: *The figures could be made more informative. In Figure 2, 3 and 5, the axes lack clear units or precise information on the quantities shown, and some legends remain too limited to guide the reader which needs to infer from the text outside of the figure. E.g. "X variables" is not defined, and most colour bars come without text. I would encourage the authors to improve the figure labels, legends, and captions so that each figure can be understood more directly on its own.*

Answer: We agree that these figures should be more self-contained. In the revised manuscript, we have improved not only the captions but also the axis labels, legends, and colour-bar annotations. For Figure 2, we removed the percentile-based visualization and instead present the comparison of the methods in terms of the extreme-event count and the input magnitude in a table. For Figure 3, we added explicit labels for the quantities shown on the axes, namely the temperature parameter and the sample number, and descriptive colour-bar titles for all panels: "Number of Extreme Values", "Number of Interventions", and "Mean Intervention Magnitude". We also revised the caption so that the meanings of panels (a)-(c) can be understood without referring to the main text. For Figure 5, we added spatial axis labels with units ("Distance [km]") and a labelled colour bar ("Wind Speed [m/s]"). In addition, the caption now explicitly states that the plotted field is the wind-speed magnitude and that the white rectangle denotes the target region.

We believe that these revisions make the figures substantially more informative and easier to interpret independently of the main text.

Changes in the manuscript: We revised the axis labels and color bars in Figures 4 and 5. In addition, we removed Figure 2 and instead present the performance comparison in Table 1.

Q9: *Here are some very small corrections that would make the manuscript easier to read and more complete:*

1. *In Algorithm 1, the Require line lists H but not H_c, although line 5 calls EnKC with H_c.*
2. *In Eq. (1), the model equation uses q_{t-1}, while the text below defines q_t as the model error.*
3. *In Eq. (10), the text just above introduces an ensemble $\{x_t^{\alpha(i)}\}_{i=1}^m$, but Eq. (10) then uses $x^a_{t+T_{\mathrm{EnKC}}}$ as if it were a single, already defined state. It is unclear from a quick read whether this denotes one ensemble member, the ensemble mean, or another quantity.*
4. *The switch from the generic EnKC horizon T_{EnKC} to T_c in the Lorenz–96 section could be made more explicit.*
5. *The Gaussian sampling notation should be clarified. In Sect. 3, the second Gaussian parameter is introduced as a variance/covariance term, but in the numerical sections the paper sets $g_{\mathrm{mag}}(u_t^{\mathrm{EnKC}}) = \|u_t^{\mathrm{EnKC}}\|_1/2$ in Lorenz–96 and $g_{\mathrm{mag}}(u_t^{\mathrm{EnKC}}) = \|u_t^{\mathrm{EnKC}}\|_1$ in SQG, which seems like a natural choice for a standard deviation. It would help to clarify whether the authors mean $m \sim \mathcal{N}(\mu, g_{\mathrm{mag}})$ or $m \sim \mathcal{N}(\mu, g^2_{\mathrm{mag}})$.*
6. *In the paragraph introducing Fig. 5, the order "no-control, EKG-MPPI, and EnKC" should be made consistent with Fig. 5, where the order is "no-control, EnKC, and EKG-MPPI."*
7. *The introduction refers to Henderson as an early 4D-Var-based study, but the corresponding reference is missing from the bibliography. The citation should be added explicitly: Henderson, J. M., Hoffman, R. N., Leidner, S. M., Nehrkorn, T., and*

Grassotti, C. (2005), "A 4D-Var study on the potential of weather control and exigent weather forecasting," Quarterly Journal of the Royal Meteorological Society, 131(612), 3037–3051, doi:10.1256/qj.05.72.

8. *For the first CSE citation on line 34, I think the original CSE paper from 2022 is better suited: Miyoshi, T. and Sun, Q. (2022), "Control simulation experiment with Lorenz's butterfly attractor," Nonlinear Processes in Geophysics, 29, 133–145.*

9. *The manuscript currently cites Sawada (2024a) as a preprint; the authors may also update that entry to the published journal version: Sawada, Y. (2024), "Ensemble Kalman Filter Meets Model Predictive Control in Chaotic Systems," SOLA, 20, 400–407, doi:10.2151/sola.2024-053.*

Answer: We thank the Referee for these careful and helpful comments, and we have addressed each of the points raised by implementing the corresponding revisions individually.

Changes in the manuscript:

1. In Algorithm 1, we add H_c to the Require line and explicitly define it as the operator that maps the state variables to the control criteria.

2. In Eq. (1) and the text below it, we correct the notation so that the model-error term is written consistently as q_{t-1}.

3. After introducing the analysis ensemble in Eqs. (8) and (9), we now state explicitly that $\bar{x}_t^a$ denotes the analysis ensemble mean at time t. We also clarify the notation around Eq. (10) by specifying that the state at $t + T_{\mathrm{EnKC}}$ refers to the ensemble-mean forecast obtained by propagating the analysis ensemble to that lead time.

4. In the Lorenz–96 section, we replace T_c with T_{EnKC} in order to unify the notation with the generic EnKC formulation.

5. We clarify that the second parameter of the Gaussian distribution denotes a variance/covariance. Accordingly, we revise the definitions of g_{mag} so that they are given on the variance scale, and we have added this clarification to Step 3 in Section 3.2. Specifically, we now define $g_{\mathrm{mag}}(u_t^{\mathrm{EnKC}}) = \|u_t^{\mathrm{EnKC}}\|_1^2$ in the SQG experiments.

6. We correct the ordering in the paragraph introducing Fig. 5 so that it matches the figure itself, namely: no control, EnKC, and EKG-MPPI.

7. We add the missing Henderson reference to the bibliography: Henderson et al. (2005), "A 4D-Var study on the potential of weather control and exigent weather forecasting."

8. We replace the first CSE citation with the original paper by Miyoshi and Sun (2022), as suggested.
9. We update the Sawada (2024a) reference from the preprint entry to the published journal version in *SOLA*.