First of all, thank you very much for the helpful comments and suggestions you have made about our manuscript. This will definitely help the improvement of the research to a greater extent.

Major comments:

1. One major concern pertains to the classification of three weather regimes using a machine learning (ML) method and the correlation of each identified regime with stability or instability, determined by averaged local Lyapunov exponents. While Charney and Devore's (1979) model is traditionally used for studying non-chaotic weather regimes, Lorenz's (1963a) model is renowned for illustrating chaotic features. It's crucial to note that Lorenz applied a similar approach in the early 1960s to study chaotic and nonlinear oscillatory solutions in a two-layer quasi-geostrophic system (e.g., Lorenz 1962, 1963b; Shen, Pielke Sr., and Zeng, 2023, https://www.mdpi.com/2073-4433/14/8/1279). A brief review of Lorenz's contributions is necessary, and a comparison of the QG-based models should address the variation in the number of Fourier modes concerning chaotic features.

Thank you very much for the suggestion. It is indeed important to discuss Lorenz's contributions and more precisely the comparison between QG - based models in the present context. More elaborated discussion is added in the introduction section of the manuscript from line 45 to 75 as follows

"The pursuit of simplified models for atmospheric phenomena has a long history, dating back to Lorenz's seminal work in the early 1960s (Lorenz, 1960, 1962, 1963a, b). This approach recognizes the value of sacrificing some detail in exchange for a deeper grasp of fundamental physical processes. Lorenz demonstrated the power of this strategy by leveraging Fourier series to distill the barotropic vorticity equation into three ordinary differential equations (Lorenz, 1960). These equations, while omitting smaller scales of motion, yielded valuable insights into atmospheric scenarios such as flow interactions and current stability. Subsequently, he developed a simplified geostrophic model using truncated Fourier-Bessel series (Lorenz, 1962). This eight-equation model captured baroclinic instability, a critical process in atmospheric dynamics, while maintaining key energy relationships. Notably, the model successfully reproduced observed flow regimes and transitions in rotating fluids, suggesting its effectiveness in studying large-scale atmospheric behavior. Lorenz's 1963 research yielded significant advancements in our understanding of atmospheric dynamics through two key publications (Lorenz, 1963a). The first introduced a now-iconic system of three differential equations, derived from a further simplified model for fluid flow. This groundbreaking work unveiled the concept of sensitive dependence on initial conditions, a cornerstone of chaos theory.

In the same year, Lorenz explored a separate avenue by investigating a simplified model for symmetrically heated rotating viscous fluids (Lorenz, 1963b). This work resulted in a system of fourteen ordinary differential equations governed by two external parameters: the thermal Rossby number and the Taylor number. Analytical solutions revealed the existence of purely zonal flow and superimposed steady waves, while numerical integration unveiled a richer tapestry of flow behaviors. Oscillatory solutions with periodic shape changes and irregular non-periodic flow emerged. Interestingly, increasing the Taylor number generally led to greater flow complexity, except at very high values where the model's truncations became unrealistic.

Perhaps most intriguing was the coexistence of unstable purely zonal, steady-wave, and oscillatory solutions. This suggests intricate flow dynamics, with transitions between symmetric and unsymmetric vacillation occurring independently of instability. These findings highlight the ability of simplified models to unveil complex and nuanced behaviors in atmospheric dynamics (Shen et al., 2023).

Lorenz's pioneering work in the early 1960s demonstrated the power of simplified models for understanding atmospheric dynamics. By strategically neglecting certain complexities, he was able to capture key phenomena like baroclinic instability and chaos. However, for large-scale atmospheric simulations, computational efficiency becomes paramount. This is where QG (quasi-geostrophic) models come in. QG models prioritize large-scale features by making specific approximations, allowing for rapid simulations and analyses of broad atmospheric circulation patterns. While they may not capture the intricate details explored by Lorenz's models, QG models remain a workhorse for studying large-scale atmospheric phenomena."

Several issues need attention:

1. Gaussian Mixture Clustering (GMC) was employed for classification, assuming each cluster has a Gaussian component. The suitability of this assumption for chaotic regimes should be addressed, especially considering the regular spatial patterns that appear in the classified regimes. Associating different weather regimes with components of the leading Lyapunov vector, particularly in the

presence of multiple positive Lyapunov exponents, should be addressed. The challenge is heightened by the time-varying components of each Lyapunov vector along the solution orbit, making it difficult to link specific components to zonal or blocking regimes.

Thank you very much for pointing it out. The distribution of the data points on the attractor is not presumably a Gaussian distribution. But the algorithm itself approximates the distribution as Gaussian and identifies clusters based on that approximation. This is pointed out in the manuscript from line 442 to 445 as

"While the distribution of data points on the attractor may not align with a Gaussian distribution, the algorithm proceeds by approximating the distribution as Gaussian and subsequently identifies clusters based on this approximation."

The methodology depicted in this paper is such as computing the local Lyapunov exponents on each datapoint on the attractor, clustering the attractor using GMC and then calculating the average of largest local Lyapunov exponents based on the identified cluster. Even though the attractor has multiple positive Lyapunov exponents, we only investigated the first one since the dynamics of the error will ultimately follow this exponent. To clarify this we added in the manuscript

"Although the attractor exhibits multiple positive Lyapunov exponents, our investigation focused solely on the first exponent, as it governs the dynamics of the error and is therefore considered the most influential."

2. The use of a fixed cluster number (3) in GMC for models with different numbers of positive Lyapunov exponents (LEs) (e.g., 3 positive LEs in Figure 7 and 12 positive LEs in Figure 12) raises concerns. Various clustering values should be explored to illustrate the relationship between the number of weather regimes and the number of positive LEs. For example, can we observe similar weather regimes in the 30 and 165 variable systems? If this is the case, does it imply a consistent number of multiple regimes or equilibrium points across both systems?

Thank you for your input. Evaluating various cluster numbers in the clustering process is crucial for obtaining meaningful outcomes. Our experimentation, spanning from 2 to 6 clusters, revealed distinct flow regimes for 2 and 3 clusters, each exhibiting notable differences. However, as the cluster number increased to 4, we observed overlap between two clusters, resulting in identical structures and flow patterns. This trend persisted with further increases in cluster number. Consequently, we concluded that the attractor achieved optimal clustering with clear and nearly equal data point distribution when utilizing 3 clusters. This is added in the manuscript in the lines 225 to 235 as

"Through experimentation encompassing 2 to 6 clusters, discernible flow regimes emerged for 2 and 3 clusters, showcasing significant distinctions. However, as the cluster count reached 4, we noted convergence between two clusters, leading to identical structures and flow characteristics. This trend persisted with additional cluster increments. Hence, we inferred that the attractor attained optimal clustering with evident and nearly uniform data point distribution when employing 3 clusters."

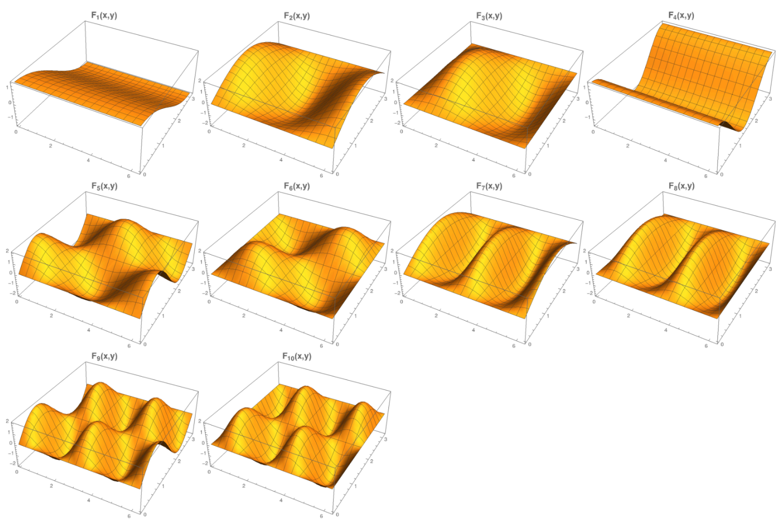
3. Is it possible to compute Lyapunov exponents for individual weather regimes? Could the time evolution of each weather regime be graphed for comparative analysis?

Thank you very much for this suggestion. In our current analysis, we're computing local Lyapunov exponents for every data point on the attractor. These exponents help us understand the stability of trajectories in our system. Once we've classified the data points using Gaussian mixture clustering, we determine the average of the largest local Lyapunov exponent within each cluster. This process allows us to identify the predictability horizon associated with each cluster. Furthermore, we analyze the flow regimes by examining the average geopotential height at 500 hPa within each cluster. This helps us characterize the different atmospheric circulation patterns that emerge from the data. By combining these approaches, we gain insights into both the predictability and the underlying dynamics of the system, enabling a deeper understanding of its behavior.

4. Concerning predictability in typical dynamical systems, characterized by systems of ordinary differential equations (ODEs), a positive Lyapunov exponent (LE) usually signifies temporal chaos (distinct from spatial-temporal chaos). In estimating predictability horizons under different conditions, a higher positive LE, on average, implies a greater average growth rate, indicating faster error growth and thus diminished predictability horizons. However, when applying this concept to assess predictability in spatial-temporal systems, it becomes imperative to

account for errors related to spatial movement. The analogy of whether more intense hurricanes (with higher growth rates) are less predictable encourages authors to contemplate the influence of spatial movement on error predictions. Therefore, in contrast to a zonal flow, although a blocking regime is linked to instability manifested by a larger LE, the consideration of spatial movement is essential when comparing errors in zonal and blocking cases in order to compare their predictability horizons.

The QGS land atmosphere coupled model is a spectral model where we project ODEs on several basis functions that represent different spatial patterns. Precisely in the



manuscript we used 10 different basis functions which represent the spatial pattern shown in the figure above. So computing the perturbations of the spatial mode or computation of Lyapunov exponents in the spectral space is indeed considering both spatial and temporal chaos. In this context, spatial amplification of the modes will not be calculated. As you mentioned, the spatial evolution of error is of considerable interest, its significance becoming more pronounced in higher resolution runs, which we intend to explore in a subsequent investigation.

Specific Comments.

Page 5: Please indicate whether and how Eqs. (1) and (2) are coupled with Eqs. (3) and (4).

Thank you for the comment. Discussion about the coupling is included in the revised manuscript from line 177 to 185 as

“The hydrostatic relation in pressure coordinates $\partial\Phi = -1/\rho a$ where the geopotential height $\Phi_i = f_0\psi^i_a$ and the ideal gas relation $p = \rho_a RT_a$ allow one to write the spatially dependent atmospheric temperature anomaly $\delta T_a = 2f_0\theta_a/R$,with θ_a as the baroclinic stream function. This can be used to eliminate the vertical velocity ω. This changes the independent dynamical field to the stream function field ψ_a and the spatially dependent temperatures δT_a and δT_g.”

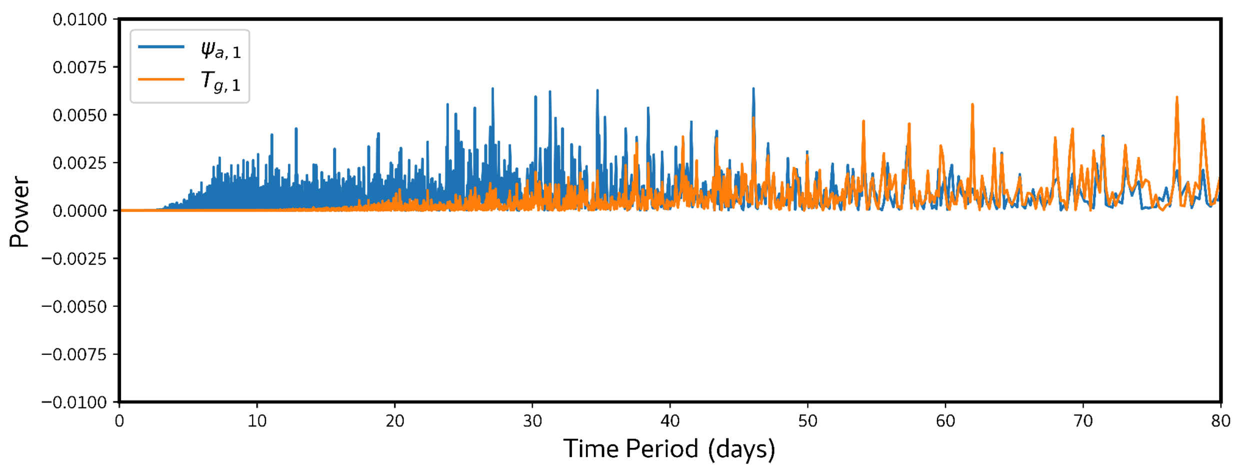
Page 7: Please indicate the equation(s) that contains Cg.

Indicated in the revised manuscript as

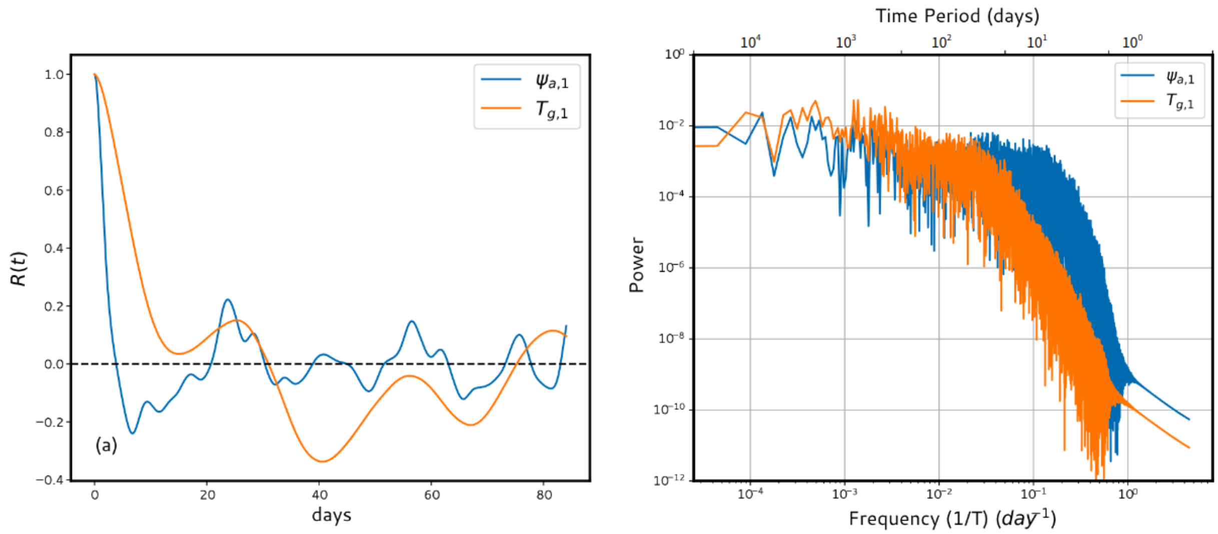
“The dimensional meridional differential shortwave solar radiation absorbed by the land and the atmosphere are given by $\delta R_g = \sqrt{2}C_g\cos(y/L)$ and $\delta R_a = \sqrt{2}C_a\cos(y/L)$ respectively. Hence we decide to provide $C_a = 0.4C_g$. The variable C_g is a dimensional parameter, which is an indicator of the meridional difference in solar heating absorbed by the land between the walls, and it is the crucial parameter in our land atmosphere coupled model.”

Page 7, Figure 1c with ACC. Is it possible to identify a time scale of approximately 36 days within the timeframe spanning 20 to 80 days?

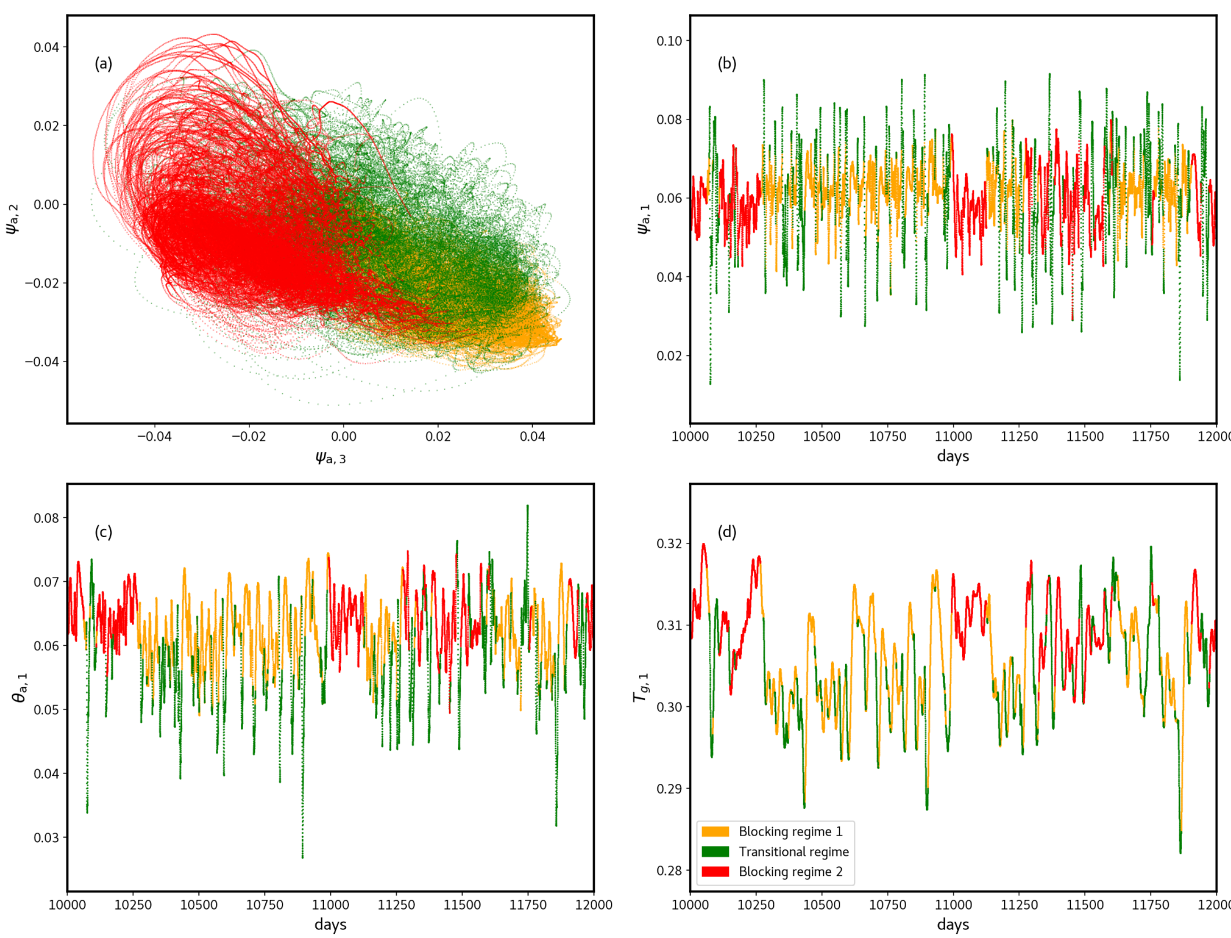
Yes, Through a comprehensive examination of the power spectrum, we found it inconclusive to find oscillations every 36 days since we are observing several peaks



during the course of 80 days.



Page 7: Can a figure analogous to Figure 1 be generated for each of the categorized regimes?

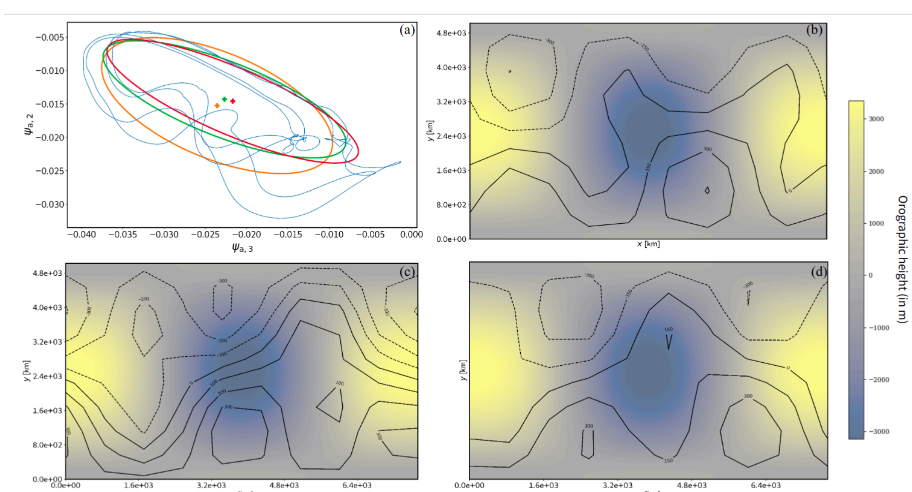


Thank you for the comment. Yes, it could be drawn by giving different colors to the data points belonging to different flow regimes or respective clusters. In the above figure, phase space dynamics of the model projected on the ($\psi_{a,3}$, $\psi_{a,2}$)-plane is shown in panel (a) for k_d = 0.08 . The figure illustrates the Gaussian mixture clustering results, where each data point is colored according to its corresponding cluster or flow regime. Panel (b), (c) and (d) represents the temporal evolution of barotropic ($\psi_{a,1}$) and baroclinic ($\theta_{a,1}$) atmospheric stream function along with the ground temperature ($T_{g,1}$) for C_g= 300 W m^{-2} and k_d = 0.08 with colors corresponding to their respective clusters. We included this figure into the appendix as another way of representing clusters.

Page 9: The discussions on the Oseledec method should be expanded to incorporate insights from Lorenz's contributions. (e.g., Lorenz 1965; please see a review by Shen, Pielke Sr, and Zeng, 2023).

Thank you very much for drawing our attention toward this paper. In Lorenz's 1965 paper, he uses a 28-variable atmospheric model which is developed by extending the equations of a two-level geostrophic model using truncated double-Fourier series. This model accounts for nonlinear interactions among disturbances of varying wavelengths. Numerical integration is employed to find nonperiodic time-dependent solutions. By comparing solutions with slightly different initial conditions, the rate of growth of small initial errors is investigated. Lorenz's error growth estimation is based on using singular value decomposition which is not the direction we wanted to proceed with the current paper as we computed the Lyapunov exponents that are asymptotic properties of the attractor. Hence we will not use that reference in the current context.

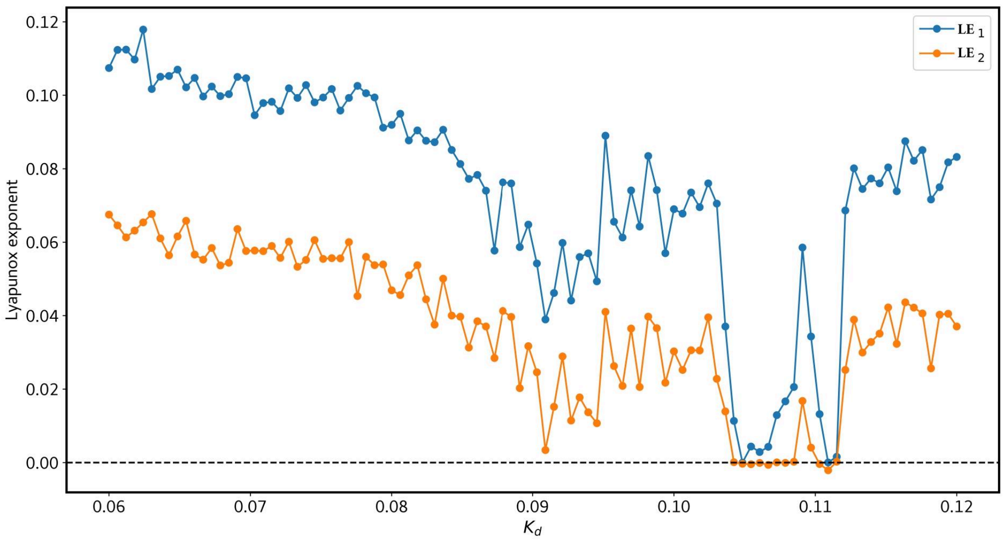
Page 11: Please add Figure B3 to include flows for Cg = 400 or kd = 0.12, for periodic flows.



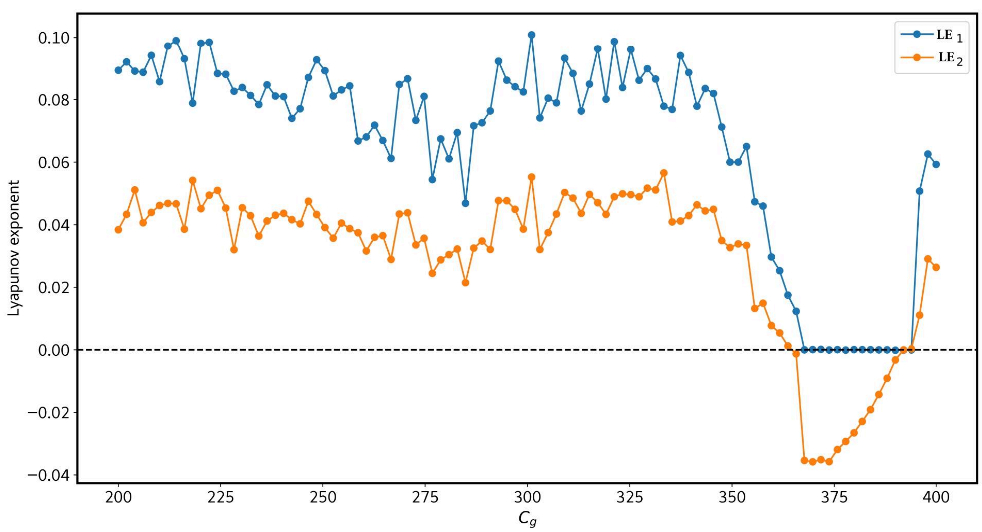
System enters into periodic behavior when k_d = 0.105 to 0.115. Hence we added figure B3 with C_g=300 W/m^2 and k_d= 0.105.

Pages 12 and 13: while $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are used for representing the 1st and 2nd LEs, respectively, the symbol lambda indicates heat exchange. Please consider making changes to reduce confusion.

Thank you very much for pointing this out. $\boldsymbol{\lambda}_1$ and $\boldsymbol{\lambda}_2$ are changed into LE_1 and LE_2 for



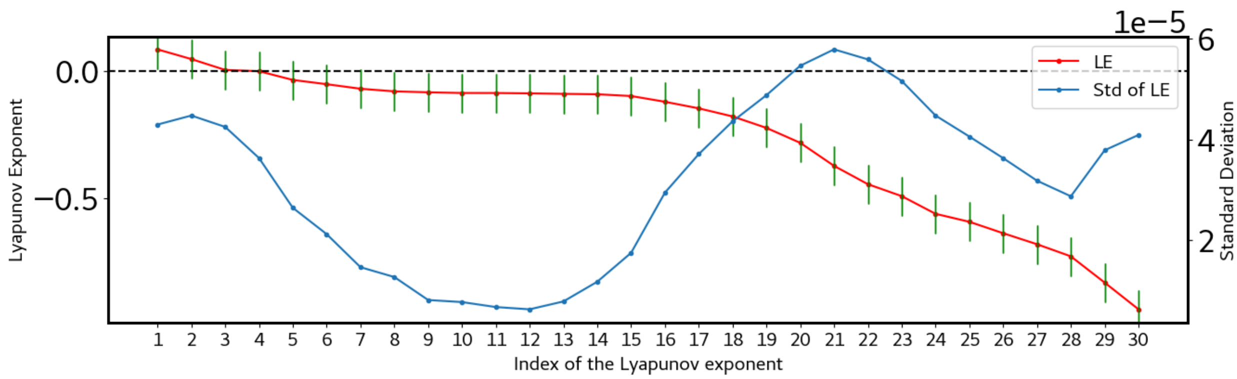
avoiding confusions in figure 6 and 8.



Page 12, line 280. Does the selection of lambda = 0 result in an uncoupled model? How can this be contrasted with the models proposed by Charney and Devore, Lorenz (1962) and/or Lorenz (1963b)?

Selection of lambda = 0 will not result in an uncoupled model. It only ceases the heat exchange between land and the atmosphere. The land and atmosphere components in the model are still interacting via incoming shortwave radiations, outgoing long wave radiations as per the equations. The model proposed by Charney and Devore has similar dynamics where the only difference is the energy balance system. The land atmosphere coupled model has a realistic energy balance system as in Barsugli and Battisti (1998). The coupling of the atmospheric components with the ground is constituted by the surface friction and the radiative flux whereas Charney and Devore are using an energy balancing system based on Newtonian cooling coefficient.

Pages 12 & 17, (in Figures 7 & 12), please offer perspectives on whether the presence of the plateau suggests the existence of singular eigenvalues with higher multiplicity.



After analyzing the uncertainty using bootstrap method, we can now say confidently that the values forming the plateau are very close to each other, suggesting the existence of a potential degeneracy of the eigenvalues. Whether or not it corresponds to a geometrical degeneracy of the Lyapunov vectors (a tangency) as in Vannitsem & Lucarini (J. Phys. A, 2016) remains to be investigated.