Weather Clim. Dynam. Discuss., referee comment RC1
https://doi.org/10.5194/wcd-2022-55-RC1, 2022

Comment on wcd-2022-55

Anonymous Referee #1

Referee comment on "Improved extended-range prediction of persistent stratospheric perturbations using machine learning" by Raphaël de Fondeville et al., Weather Clim. Dynam. Discuss., https://doi.org/10.5194/wcd-2022-55-RC1, 2022

General comments

This manuscript quantifies, and aims to improve, the predictability of sudden stratospheric warming events using statistical methods. The contribution is significant and valuable, as the authors manage some impressive improvements in predictability while keeping the machine learning physically constrained and transparent. The data processing pipeline is simple enough to be reproduced, which is not always the case with similar studies. They also exercise care in checking the robustness of their results with cross-validation. I do have some technical questions and concerns about some "hyper-hyper-parameters", in particular the motivation for the SSW definition. Provided these choices are explained, the paper should be suitable to publish after minor revisions.

Specific comments

1. My first and primary concern is with the data-driven index for slow-recovering SSWs. My understanding is that you opted not to use a standard definition, e.g., the reversal of zonal-mean zonal wind at 60N and 10hPa as per the WMO, for two reasons: (1) to have a continuous sliding scale, and (2) to avoid arbitrariness. However, you could easily modify the WMO definition to make it continuous, e.g., by changing the zonal-wind threshold or the duration. Also, standard definitions enjoy the significant advantage of being simple and objective, whereas data-driven methods such as the one you describe (PCA on temperature profiles) seems to require far more arbitrary choices, such as the size of the polar cap region, the pressure levels, and the spatiotemporal resolution. You do back up your choice by citing multiple preceding studies linking the temperature profile with surface impacts, so I'm not asking you to change the procedure, but the standard definition could be dealt with more thoroughly. I would hope that many key conclusions of the study, such as statistically significant precursors, would not change much if you were to swap in the standard definition. However, you do refer to the standard criterion at various later points in the paper; overall, I was unsure to what extent the WMO definition was being used as a parameter or just a point of comparison. Is a large value of your

perturbation index, I, highly correlated with the occurrence of an SSW defined by the WMO? Some comments about this would be welcome.

2. Figure 3 can use some clarification. First of all, I didn't see the amount of variance in the first four principal directions reported anywhere---did I miss it? This seems like relevant information. Second, I found the colors and overlapping curves to be hard to parse. The high- and low-pressure levels are colored quite similarly, and the "fast downward progress of the anomalies" (c. line 176) is not so obvious to me. A heat map might be more clear, with each row corresponding to a pressure level and each column a time sample. That is how I'm used to visualizing downward propagating anomalies, e.g., as in Baldwin & Dunkerton (2001) Fig. 2. Third, it is hard to connect mixtures of these principal directions with a pattern of vertical temperature change.

3. What is the reason to choose sPCA instead of linear regression (perhaps with regularization to promote sparsity, e.g., LASSO)?

Technical corrections

Fig. 4: Some material is cut off at the interface between the left and right panels. For example, the vertical axis should have a "PC2" label, whereas the "Jul 2009" axis tick label is cut off. Further, because the PCs are not standard, it was not obvious (at least not to me) which direction to follow the black curves in the right-hand panel. I trust they all move in the same direction (counterclockwise, as suggested by the text)?

Circa line 175 (time-delay embedding notation): there is some mixing between the subscripts on time (t_1,...,t_N) and the time itself (t=1,...,N-(T-1)). It would be better to stick to one notation for simplicity.