Peer-Review of
"Data Driven Regional Weather Forecasting"

November 30, 2022

1 Summary of the Content

The objective of this paper is to provide a data-driven forecast method to replace the conventional model forecast. The authors use the time-delayed observations to reduce the uncertainty of data-driven forecast, and then set up a parameterization for the forecast model. The parametrization framework consists of a polynomial term and a analog-type of method. Once the parameterization is set, the parameters are estimated using standard methods. The authors tested this method using a shallow water equation on a 10×10 grid. The authors considered several different observed regions. The overall numerical performance is quite good.

2 Overall Feedback

The general idea is easy to understand and the numerical results look good and clear. However, certain aspects still need to be clarified. I recommend major revision for this manuscript.

3 Comments

1) Lines 580-584, 622-628

In lines 580-584, the authors claim that the polynomial term is used to represent the polynomial terms in the original dynamics. SWE has nonlinear terms, so the polynomials should have degree larger than 1. However, in lines 622-628, the authors seem suggest that the choice of linear polynomial is due to their insight about the dynamics of SWE. This contradicts to their reasoning in lines 580-584. Please explicitly explain why degree 1 polynomials are good choices for SWE, and what is your insight into the dynamics of SWE.

2) I recommend the authors provide their realization of their algorithm with the paper. This would make introspection/adaption/expansion of their results by the community much easier and faster.

3) Please add some figures to visualize your choice of "centers", i.e. the data at which of the time steps are used in the RBFs?

4) The scientific story of this manuscript needs to be justified with more reasoning and tests.

I have another interpretation of the methodology presented in this manuscript. The use of "centers" in this method essentially makes use of the ergodic nature of the SWE with forcing. The "centers" are sample points on the attractor of the system. The Gaussian RBFs serve as a weight functionals and the coefficients $w_{\alpha q}$ are approximately the true dynamics corresponding to the state $\mathbf{TD}^c(q)$. In this case, the polynomial term could serve as a residual term. Can the authors quantify the contributions from the polynomial term and the RBF term, and discuss about the results?

5) Throughout the manuscript the authors use $\mathcal{R}$ to denote the set of real numbers. Conventionally, people use $\mathbb{R}$ instead of $\mathcal{R}$.

6) Some of the notation/index in the equations are confusing.

For instance, In Eq.(11), you probably need a subscript a in $\mathbf{S}(n)$ and $\mathbf{S}(n+1)$. At the end of Eq.(17), what does $\mathbf{O}_j(n)$ mean?