


SUPPLEMENTARY MATERIAL

# Supplementary Material for "Chaotic systems learning with hybrid Echo State Network/Proper Orthogonal Decomposition based model"

M. Lesjak<sup>1</sup> and N. A. K. Doan<sup>2</sup> \*

<sup>1</sup>Department of Mechanical Engineering, Technical University of Munich, Boltzmannstr. 15, 85747 Garching, Germany

<sup>2</sup>Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 1, 2629 HS Delft, Netherlands

\*Corresponding author. E-mail: [n.a.k.doan@tudelft.nl](mailto:n.a.k.doan@tudelft.nl)

Received: August 17, 2021

## 1. Additional details on the parameters of the CDV system

The Charney-DeVore (CDV) system is a chaotic low-order atmospheric model inspired by (Crommelin et al., 2004) that models the barotropic flow in a  $\beta$ -plane with orography. A truncated version of the CDV system is considered here and it is governed by the following set of ODEs (Doan et al., 2020; Crommelin et al., 2004; Crommelin and Majda, 2004):

$$\begin{aligned} \dot{u}_1 &= \gamma_1^* u_3 - C(u_1 - u_1^*) \\ \dot{u}_2 &= -(\alpha_1 u_1 - \beta_1) u_3 - C u_2 - \delta_1 u_4 u_6 \\ \dot{u}_3 &= (\alpha_1 u_1 - \beta_1) u_2 - \gamma_1 u_1 - C u_3 + \delta_1 u_4 u_5 \\ \dot{u}_4 &= \gamma_2^* u_6 - C(u_4 - u_4^*) + \epsilon(u_2 u_6 - u_3 u_5) \\ \dot{u}_5 &= -(\alpha_2 u_1 - \beta_2) u_6 - C u_5 - \delta_2 u_4 u_3 \\ \dot{u}_6 &= (\alpha_2 u_1 - \beta_2) u_5 - \gamma_2 u_4 - C u_6 + \delta_2 u_4 u_2 \end{aligned} \quad (1)$$

The model coefficients in the CDV systems were chosen in this work to ensure a chaotic and intermittent behaviour and their exact definitions and values are provided hereunder.

The model coefficients are given by:

$$\begin{aligned} \alpha_m &= \frac{8\sqrt{2}m^2(b^2 + m^2 - 1)}{\pi(4m^2 - 1)(b^2 + m^2)}, & \beta_m &= \frac{\beta b^2}{b^2 + m^2} \\ \delta_m &= \frac{64\sqrt{2}}{15\pi} \frac{b^2 - m^2 + 1}{b^2 + m^2}, & \gamma_m^* &= \gamma \frac{4\sqrt{2}mb}{\pi(4m^2 - 1)} \\ \epsilon &= \frac{16\sqrt{2}}{5\pi}, & \gamma_m &= \gamma \frac{4\sqrt{2}m^3 b}{\pi(4m^2 - 1)(b^2 + m^2)} \end{aligned} \quad (2)$$

for  $m = 1, 2$ . In this work, we set the parameters as in (Wan et al., 2018),  $(u_1^*, u_4^*, C, \beta, \gamma, b) = (0.95, -0.76095, 0.1, 1.25, 0.2, 0.5)$ , which ensures a chaotic and intermittent behaviour.

## 2. Additional details on the hyperparameters search for the ESN and values of the hyperparameters used

### 2.1. Details on the hyperparameters search

To determine a set of optimal hyperparameters, the following approach is used in our work:

1. The reservoir size is fixed to 500 neurons throughout the hyperparameter search.
2. Following (Lukoševičius, 2012), we optimize the ESN parameters in the following order: (i) input scaling,  $\sigma_{in}$ , (ii) spectral radius,  $\rho$ , (iii) degree of connectivity,  $\langle d \rangle$ , (iv) leaking rate,  $\alpha$ , and (v) Tikhonov regularization factor,  $\gamma$ .
3. To obtain a "first guess" of appropriate values for these parameters, for each hyperparameter successively, we vary it in a pre-specified range and compute the prediction horizon of the resulting ESN. The resulting best performing hyperparameters are retained.
4. To refine the values of the hyperparameters, starting from the values found at the previous stage, we perform successively for each hyperparameter (in the order specified in step 2.) a line search with a given step  $h$  (i.e. we estimate the accuracy of the ESN when the considered hyperparameter is varied by  $h$  - both when it is increased and decreased by  $h$ ). If the modified hyperparameter does not yield improved accuracy, then the step size  $h$  is decreased. Otherwise, the hyperparameter value is updated to this newfound value. Once a maximum number of optimization step has been reached, we move on to the next hyperparameter (in the order specified in step 2.).

During the hyperparameter search performed above, the accuracy of the ESN is estimated by computing the prediction horizon when the ESN (or hybrid-ESN) is evolving in an autonomous manner for 50 different initial conditions.

### 2.2. Hyperparameters of the ESN for the prediction of the CDV system

The values of the hyperparameters of the different ESNs for the CDV system are provided in Table 1 for the data-only ESN and in Table 2 for the ESN in the hybrid approaches.

**Table 1.** Hyperparameters of the data-only ESN for the prediction of the CDV system.

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
3.5	0.8778	0.64999	0.7	$10^{-5}$

**Table 2.** Hyperparameters of the ESN for (left) hybrid-ESN-A and (right) hybrid-ESN-B for the prediction of the CDV system.

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$	$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
3.5	0.8778	0.8	0.55	$5 \times 10^{-5}$	2.5	0.8099	0.40499	0.65	$10^{-5}$

### 2.3. Hyperparameters of the ESN for the prediction of the Kuramoto-Sivashinsky system

The values of the hyperparameters of the different ESNs for the prediction of the Kuramoto-Sivashinsky equation are provided in Table 3 for the data-only ESN, in Table 4 for the ESN in the hybrid-ESN-A and in Table 5 for the ESN in the hybrid-ESN-B.

**Table 3.** Hyperparameters of the ESN for the prediction of the KS system

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
3.5	1.0	0.1	0.3	$10^{-5}$

**Table 4.** Hyperparameters of the hybrid-ESN-A for the prediction of the KS system (left: with 19 modes; right: with 29 modes)

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
3.0	1.0	0.378	0.4	$5 \times 10^{-5}$

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
2.5	0.8499	0.5	0.35	$10^{-4}$

**Table 5.** Hyperparameters of the hybrid-ESN-B for the prediction of the KS system (left: with 19 modes; right: with 29 modes)

$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
2.5	0.95	0.4	0.3	$5 \times 10^{-5}$

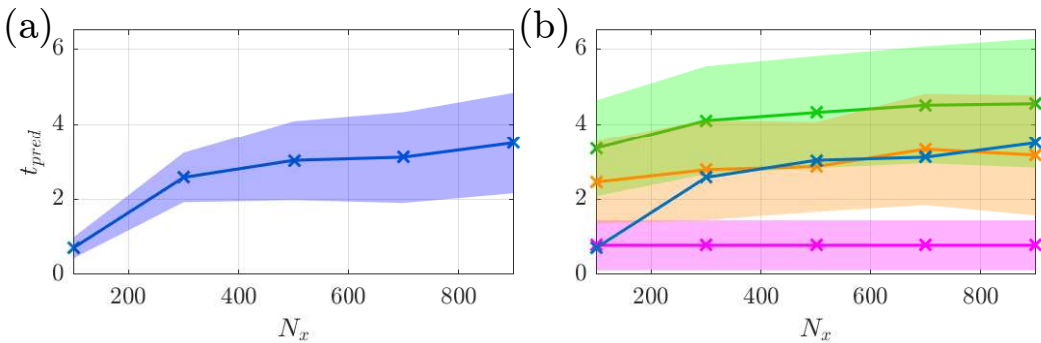
$\langle d \rangle$	$\alpha$	$\sigma_{in}$	$\rho$	$\gamma$
3.0	0.95	0.045	0.3	$7.2 \times 10^{-5}$

### 3. Prediction horizon from different ESN realizations

In this section, we present the prediction horizon of the data-only ESN, hybrid-ESN-A and hybrid-ESN-B when it is computed using 10 different realizations of the ESN given that the performance of the ESN may be affected by the random seed used to generate the matrices  $W_{in}$  and  $W$  (Haluszczynski and Räth, 2019). We assess here whether our results are affected by this. The calculation of the prediction horizon for each realization is made for 50 different initial conditions.

#### 3.1. Prediction horizon for the CDV system

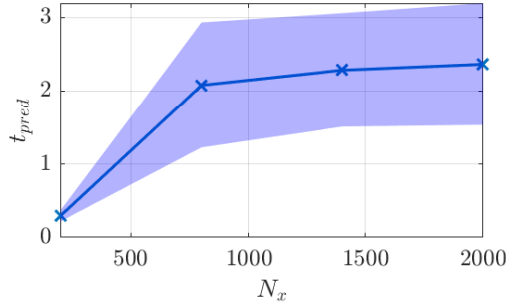
The prediction horizon obtained as described above for the data-only ESN, hybrid-ESN-A and hybrid-ESN-B is shown in Fig. 1 hereunder for the CDV system. It is seen that the trend observed is similar to the one in Fig. 9 of the manuscript. Namely, hybrid-ESN-B shows the longest accuracy while hybrid-ESN-A has a better accuracy than the data-only ESN for small reservoir sizes but is of similar accuracy than the data-only ESN for medium to large reservoir sizes.



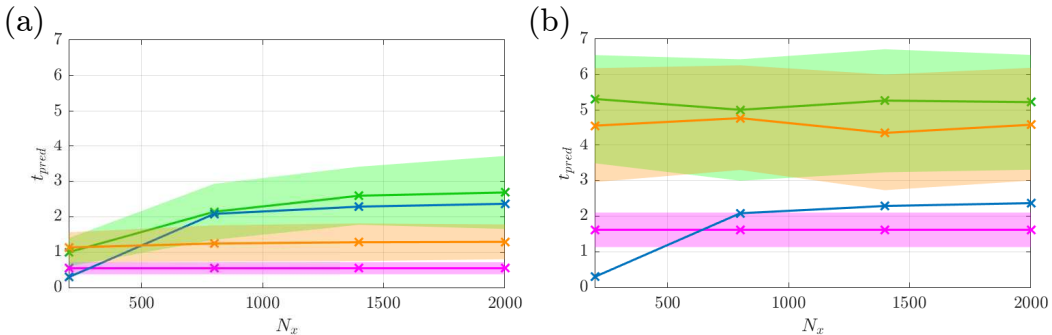
**Figure 1.** Prediction horizon for the CDV system of (a) data-only ESN and (b) the ROM only (magenta line and shaded area), the hybrid-ESN-A (orange dotted line and shaded area), the hybrid-ESN-B (green line and shaded area), of the data-only ESN (blue line) for various reservoir sizes, computed from 10 different ESN realizations (and 50 initial conditions for each realization). Shaded area indicates the standard deviation of the prediction horizon.

### 3.2. Prediction horizon for the KS system

The prediction horizon obtained as described above for the data-only ESN, hybrid-ESN-A and hybrid-ESN-B is shown in Figs. 2 and 3. It is seen that the trend observed is similar to the one in Fig. 18 of the manuscript, indicating that our results are robust with respect to the ESN realization.



**Figure 2.** Prediction horizon of the data-only ESN for various reservoir sizes, computed from 10 different ESN realizations (and 50 initial conditions for each realization). Shaded area indicates the standard deviation of the prediction horizon.



**Figure 3.** (a) With ROM of 19 modes: prediction horizon of the ROM (magenta line and shaded area), of the hybrid-ESN-A (orange line and shaded area), of hybrid-ESN-B (green line and shaded area) for different reservoir sizes and of the data-only ESN (blue line). (b) With ROM of 29 modes: prediction horizon of the ROM (magenta line and shaded area), of the hybrid-ESN-A (orange line and shaded area), of hybrid-ESN-B (green line and shaded area) for different reservoir sizes and of the data-only ESN (blue line). All prediction horizons are computed from 10 different ESN realizations (and 50 initial conditions for each realization). Shaded areas indicate the standard deviation from the average prediction horizon.

## References

- Crommelin, D. T. and Majda, A. J. (2004). Strategies for model reduction: Comparing different optimal bases. *J. Atmos. Sci.*, 61(17):2206–2217.
- Crommelin, D. T., Opsteegh, J. D., and Verhulst, F. (2004). A mechanism for atmospheric regime behavior. *J. Atmos. Sci.*, 61(12):1406–1419.

- Doan, N. A. K., Polifke, W., and Magri, L. (2020). Physics-Informed Echo State Networks. *J. Comput. Sci.*, 47:101237.
- Haluszczyński, A. and R ath, C. (2019). Good and bad predictions: Assessing and improving the replication of chaotic attractors by means of reservoir computing. *Chaos*, 29(10).
- Lukoševičius, M. (2012). A Practical Guide to Applying Echo State Networks. In Montavon, G., Orr, G. B., and Muller, K.-R., editors, *Neural Networks: Tricks of the Trade*. Springer.
- Wan, Z. Y., Vlachas, P., Koumoutsakos, P., and Sapsis, T. P. (2018). Data-assisted reduced-order modeling of extreme events in complex dynamical systems. *PLoS One*, 13(5):1–22.