How to calibrate a climate model with neural network based physics?

Blanka Balogh, David Saint-Martin, Aurélien Ribes

CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
blanka.balogh@meteo.fr

22/05/2022

This presentation participates in OSPP

EGU

Outstanding Student & PhD candidate Presentation contest

Calibration of parameterizations

- Physical parameterizations are both tuned *offline* and *online* (e.g., to calibrate the mean climate).
- Current neural network (NN) based parameterizations can only be tuned *offline* : there is no guarantee that they will be accurate once they have been plugged into the model.

How to calibrate a dynamical system with NN-based physics?

Calibration of parameterizations

- Physical parameterizations are both tuned *offline* and *online* (e.g., to calibrate the mean climate).
- Current neural network (NN) based parameterizations can only be tuned *offline* : there is no guarantee that they will be accurate once they have been plugged into the model.

How to calibrate a dynamical system with NN-based physics?

Calibration of NN parameterizations

- Step 1 : generate a new learning sample $((x, \theta), y)$.
- Step 2 : train an NN f to approximate y as a function of (x, θ).
- Step 3 : implement the NN parameterization into the model & calibrate the value of θ, noted θ^*. Use θ^* in further simulations.

Proof-of-concept experiment : L63

The Lorenz'63 model

The Lorenz'63 (hereafter, L63, Lorenz, 1963) is defined with a set of 3 equations.

$$\dot{x}_1 = \sigma(x_2 - x_1), \tag{1}$$
$$\dot{x}_2 = x_1(\rho - x_3) - x_2, \tag{2}$$
$$\dot{x}_3 = x_1 x_2 - \beta x_3. \tag{3}$$

The NN $\widehat{f}$ learns to approximate Eq.3 :

$$\widehat{\dot{x}_3} = \widehat{f}(x_1, x_2, x_3; \beta) \tag{4}$$

The learning sample is obtained by generating several L63 orbits, with $\beta \in [1.6, 3]$. Finally, $\dot{x}_3$ is replaced with the NN model.

Choice of the metric m. In our case, the metric of interest is long-term climate (mean value) and standard deviation.

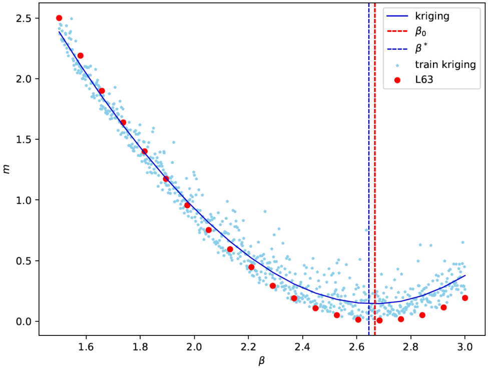
Let $[\mathbf{x}]^{\widehat{f}}(\beta)$ be a long orbit generated using the NN component, the mean value of which is noted $(\widehat{\mu}_1(\beta), \widehat{\mu}_2(\beta), \widehat{\mu}_3(\beta))$ and of standard deviation $(\widehat{\sigma}_1(\beta), \widehat{\sigma}_2(\beta), \widehat{\sigma}_3(\beta))$. If the reference orbit has a mean state $(\mu_1^0, \mu_2^0, \mu_3^0)$ and standard deviation $(\sigma_1^0, \sigma_2^0, \sigma_3^0)$, the estimate value of the metric m we have used can be expressed as :

$$\widehat{m}(\beta) = \sum_{i=1}^{3} (\widehat{\mu}_i(\beta) - \mu_i^0) + (\widehat{\sigma}_i(\beta) - \sigma_i^0). \quad (5)$$

Please note that m do not depend on the initial condition of the orbits: they are both considered long enough not to depend on it.

Proof-of-concept experiment : L63

Calibration of β is achieved by minimizing a metric m, penalizing long-term model bias. Optimization is done on a kriging metamodel, that smoothes the $\widehat{m}(\beta)$.



Value of metric m (MSE on mean climate) as a function of the tunable NN input parameter, β.

- Reference value : $\beta_0 = 2.667$
- Optimal value : $\beta^* = 2.664$.

Proof-of-concept experiment : L96

Lorenz'96 (Lorenz, 1996)

K large-scale variables x_k are coupled to J fine-scale variables $y_{k,j}$ each.

$$\frac{\mathrm{d}x_k}{\mathrm{d}t} = -x_{k-1}(x_{k-2} - x_{k+1}) - x_k + F \underbrace{\boxed{-\frac{hc}{b}\sum_{j=1}^{J} y_{k,j}}}_{\text{coupling term } B}, \tag{6}$$

$$\frac{1}{c}\frac{\mathrm{d}y_{k,j}}{\mathrm{d}t} = -by_{k,j+1}(y_{k,j+2} - y_{k,j-1}) - y_{k,j} + \frac{hc}{b}x_k. \tag{7}$$

NN parameterization $\widehat{f}$, with $\theta = c$:

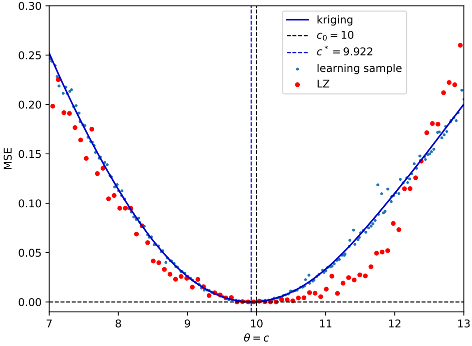
$$\widehat{B} = \widehat{f}(\mathbf{x}; c)$$

And the resulting dynamical model will be :

$$\widehat{\frac{\mathrm{d}x_k}{\mathrm{d}t}} = -x_{k-1}(x_{k-2} - x_{k+1}) - x_k + F + \widehat{B} \tag{8}$$

L96 : perfect model calibration

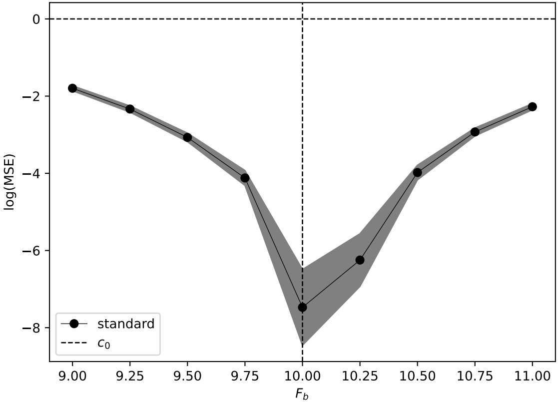
In this case, the 'dynamical' part of L96 (Eq.8) is calibrated using the reference value of the parameters, i.e., $(h, F, b) = (h_0, F_0, b_0)$ (= perfect model calibration). Calibration of c is achieved by minimizing a metric m (analoguous to the metric m used in the L63 case), penalizing long-term model bias.



- Reference value : $c_0 = 10.0$
- Optimal value : $c^* = 9.92$.

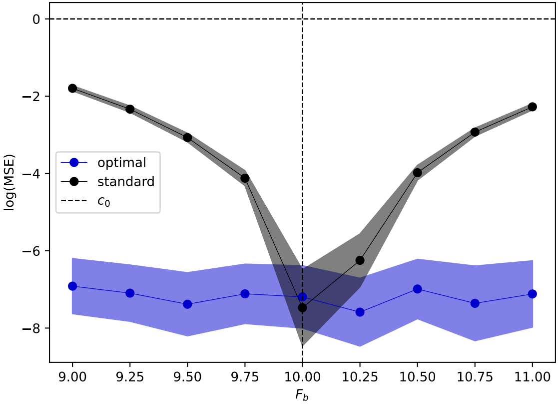
L96 : imperfect model calibration

In the following, we use a biased value $F_b \neq F_0$ in Eq.8. We suppose that F_b is not a tunable parameter.
Using $\widehat{B}(x, c_0)$ leads to high model error. The curve below represents $\log(\widehat{m}(c_0))$ for different values of F_b.



Imperfect model calibration

Using $\widehat{B}(x, c_0)$ leads to increasing model errors, as the bias on F increases.
Using $\widehat{B}(x, c^*)$ leads to low model errors, even with large biases on F.



Conclusions and perspectives

- It is possible to obtain NN-based physics that can be calibrated online, by adding some parameters θ among the input variables of the NN. The value of these additional parameters θ can be tuned to ensure low long-term bias of the resulting model.
- We have illustrate our method using the L63 model. We have successfully recovered the reference value of parameter β using our method and c in the L96 example.
- We have also successfully reduced the value of the metric m when one of the model parameters that cannot be tuned carried biases in the L96 case study.

Next step : implementing our method in a climate model.

You can also check out our paper on Essoar (v1) or in GRL.