



**HAL**  
open science

# Physics Informed Model Error for Data Assimilation

Jules Guillot, Emmanuel Frénod, Pierre Ailliot

► **To cite this version:**

Jules Guillot, Emmanuel Frénod, Pierre Ailliot. Physics Informed Model Error for Data Assimilation. Discrete and Continuous Dynamical Systems - Series S, In press, 10.3934/dcdss.2022059 . hal-03615129

**HAL Id: hal-03615129**

**<https://hal.science/hal-03615129>**

Submitted on 25 Mar 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Physics Informed Model Error for Data Assimilation

Jules Guillot <sup>\*a</sup>, Emmanuel Frénod <sup>†a1</sup>, and Pierre Ailliot <sup>‡b</sup>

<sup>a</sup>Univ Bretagne - Sud, CNRS UMR 6205, LMBA, F-56000 Vannes,  
France

<sup>b</sup>Univ Brest, CNRS UMR 6205, Laboratoire de Mathématiques de  
Bretagne Atlantique

<sup>1</sup>See-d, Vannes

## Abstract

Data assimilation consists in combining a dynamical model with noisy observations to estimate the latent true state of a system. The dynamical model is generally misspecified and this generates a model error which is usually treated using a random noise. The aim of this paper is to suggest a new treatment for the model error that further takes into account the physics of the system: the physics informed model error. This model error treatment is a noisy stationary solution of the true dynamical model. It is embedded in the ensemble Kalman filter (EnKF), which is a usual method for data assimilation. The proposed strategy is then applied to study the heat diffusion in a bar when the external heat source is unknown. It is compared to usual methods to quantify the model error. The numerical results show that our method is more accurate, in particular when the observations are available at a low temporal resolution.

**Keywords:** Data Assimilation, EnKF, Heat Equation, Model Error, Uncertainty Quantification

---

\*jules.guillot@univ-ubs.fr

†emmanuel.frenod@univ-ubs.fr

‡pierre.ailliot@univ-brest.fr

# 1 Introduction

This paper is part of the framework of data assimilation. In meteorology and oceanography, see Ghil and Malanotte-Rizzoli [8], or for factory issues as explained in Ailliot *et al.*[1], the goal of data assimilation is to reconstruct the real state of a system, for instance the sea surface temperature, taking into account observation data. It consists in combining a dynamical model, which describes the temporal evolution of the system, with observation data that contain measurement errors due to the sensors. The ensemble Kalman filter (EnKF) developed in Evensen [6] is a usual data assimilation method. The principle of this algorithm is to compute an ensemble of possible trajectories for the system, called the ensemble members, which are then corrected using the observation data. However the dynamical model embedded in this method generates an error, with respect to the reality, complicated to quantify. This model error is usually treated using a stochastic term with zero mean and a variance to specify. The main contribution of the paper is to propose an original treatment for the model error which further takes into account the physics of the system. This novel method is applied here to a simple case studied: the heat equation. The model error is treated by using a particular solution of this equation that includes a stochastic part. The efficiency of this method is then assessed by comparing its error to the ones of usual methods.

In Section 2, the model error is defined and different methodologies to mitigate it are given. The EnKF is then detailed in Section 3. The Section 4 describes the studied case and the new model error treatment. Numerical results are then discussed in Section 5. Finally, some remarks and perspectives are given in Section 6.

## 2 The model error in data assimilation

Two steps are involved in the sequential data assimilation process : the forecast and the analysis. For the forecast, the dynamical model is often used with an additive Gaussian white noise, to compensate the lack of knowledge due to unknown physical phenomena, or to correct a wrong parameterization of the dynamical model. The latter can be determined by two different approaches. If physical processes of the system are known, parameterized partial differential equations (PDEs) or ordinary differential equations (ODEs) are solved: this is the model-driven approach. Whereas when the underlying phenomena are unknown, a machine learning tool based on the available observation data may be used: this is the data-driven approach, as illustrated in Lguensat *et al.*[9]. For the analysis, the new observation is used to modify the forecast. The real state is in general partially observed by the sensors, so the observation corresponds to the observed parts of the real state. The observation is estimated by applying an observation operator to the forecast. This observation operator returns the forecast for the observed parts of the real state. The difference between the observation and its estimate is called the innovation. The Kalman gain is then

applied to this innovation to correct the forecast. It is the optimal operator in the sense that when applied to the innovation it is supposed to minimize the mean square error with respect to the real state.

The additive noise in the forecast step is used to probe the difference between the real state and the output of the dynamical model, which is the solution of the dynamical model applied to the previous state estimate. This difference is the consequence of the model error. Estimating the covariance matrix of the model error treatment allows to quantify the model uncertainty, different techniques based on the method of moments or the maximum likelihood of the innovation are shown in Tandeo *et al.*[11].

The inflation method established by Anderson and Anderson [2] is usually used to correct the bad forecasting especially due to the treatment of the model error. It consists in increasing the coefficients of the covariance matrix of the forecasts, multiplying this matrix by a time-dependent scalar determined by the equations of Desroziers *et al.*[5]. A temporal smoothing could be necessary if this inflation factor is too noisy.

Another approach is detailed in Brajard *et al.*[4] and Farchi *et al.*[7], where the model error is treated using the output of a neural network.

Stochastic parametrization is also frequently used for numerical weather prediction. Instead of adding a treatment for the model error to the output of the dynamical model, the unknown physical process is directly replaced inside the dynamical model by a stochastic term. The PDEs involved in the dynamical model become Stochastic-PDEs (SPDEs) but are solved with high computational cost to obtain the forecasts. Different strategies of stochastic parametrizations are described in Palmer *et al.*[10] and Berner *et al.*[3].

The new approach detailed in this paper is slightly different from the others: the model error is treated using a randomized solution of the stationary PDE involved in the dynamical model.

### 3 The ensemble Kalman filter

Sequential data assimilation is based on the following state-space model

$$\begin{cases} X_k = M[X_{k-1}] + \eta_k & (1a) \\ Y_k = H[X_k] + \epsilon_k & (1b) \end{cases}$$

where  $X_k \in \mathbb{R}^n$  is the real state at the discrete time  $t_k = (k - 1)dt$ , with  $k \in [1, \dots, K_{final}]$  and  $dt$  is the time step.  $Y_k \in \mathbb{R}^p$  is the observation.  $M$  is the dynamical model applied to the previous state and  $H$  is the observation operator giving the observed components of  $X_k$ .  $\eta_k \sim \mathcal{N}(0, Q)$  and  $\epsilon_k \sim \mathcal{N}(0, R)$  are the uncertainties with Gaussian assumptions, respectively the treatments for the model error and the observation error. Indeed the observation contains an error, due to the sensor, which is often treated using a Gaussian white noise with a covariance matrix  $R$  to specify, but this is not the purpose of the paper. The goal is to estimate the real state given the dynamical model and the observation data.

The EnKF provides an ensemble of  $N$  estimates for the real state at each time  $t_k$ . During the forecast step, the EnKF determines the  $N$  forecasts  $X_k^{f,i}$  for  $i \in [1, \dots, N]$ , thanks to the dynamical model and the treatment for the model error. Then knowing the new observation  $Y_k$ , those forecasts are corrected using the innovation  $d_k^i$  to obtain the  $N$  analyses  $X_k^{a,i}$ .  $X_1^{a,i}$  corresponds to the initial condition  $X_0$  of the dynamical model, plus  $\eta_1^i \sim \mathcal{N}(0, Q)$ , to generate the ensemble members for the first iteration.

Forecast step:

$$X_k^{f,i} = M[X_{k-1}^{a,i}] + \eta_k^i \quad \eta_k^i \sim \mathcal{N}(0, Q) \quad (2)$$

$$X_k^f = \frac{1}{N} \sum_{i=1}^N X_k^{f,i} \quad (3)$$

$$P_k^f = \frac{1}{N-1} \sum_{i=1}^N (X_k^{f,i} - X_k^f)(X_k^{f,i} - X_k^f)^T \quad (4)$$

$P_k^f$  is the estimated covariance matrix of the forecasts and notably depends on  $Q$ .

Analysis step:

$$K_k = P_k^f H^T (H P_k^f H^T + R)^{-1} \quad (5)$$

$$d_k^i = Y_k + \varepsilon_k^i - H X_k^{f,i} \quad \varepsilon_k^i \sim \mathcal{N}(0, R) \quad (6)$$

$$X_k^{a,i} = X_k^{f,i} + K_k d_k^i \quad (7)$$

$$X_k^a = \frac{1}{N} \sum_{i=1}^N X_k^{a,i} \quad (8)$$

The estimation of  $Q$  and  $R$  is a key point in data assimilation as explained in Tandeo *et al.*[11]. Most of the time, the model error treatments are uncorrelated in space. A strategy to further take into account the physics and mitigate the model error is discussed hereafter.

## 4 The physics informed model error

The new approach for the treatment of the model error is applied to the assimilation of temperature data of a metallic bar heated on its center. The real state is the solution of the following inhomogeneous heat equation

$$\frac{\partial X}{\partial t}(x, t) - \alpha \frac{\partial^2 X}{\partial x^2}(x, t) = r(t) \quad \text{for } x \in [0, 1] \text{ and } t > 0 \quad (9)$$

with the initial and boundary conditions

$$X(x, 0) = \sin(\pi x) \quad \text{and} \quad X(0, t) = X(1, t) = 0. \quad (10)$$

$X(x, t)$  is the real temperature at the position  $x$  on the bar at time  $t$ . The goal is to estimate  $X(x, t)$  without knowing the external heat source  $r(t)$  on the right-hand side of equation (9), using the EnKF with the following dynamical model and model error treatment.

The dynamical model:

The misspecified dynamical model used is the homogeneous heat equation:

$$\frac{\partial X}{\partial t}(x, t) - \alpha \frac{\partial^2 X}{\partial x^2}(x, t) = 0 \quad (11)$$

with the same initial and boundary conditions as (10).

The space is then discretized such as  $0 = x_1 < \dots < x_j < \dots < x_n = 1$ , with  $dx$  the space step. The semi-discretization in space of (11) is led by using the centered finite difference scheme

$$\frac{\partial X}{\partial t}(x_j, t) = \alpha \frac{X(x_{j+1}, t) - 2X(x_j, t) + X(x_{j-1}, t)}{(dx)^2} \quad \text{for } j \in [2, \dots, n-1] \quad (12)$$

with  $X(x_j, 0) = \sin(\pi x_j)$  and  $X(x_1, t) = X(x_n, t) = 0$ .

Then this ODE is discretized in time and solved using a RK4-5 scheme with the R-function *ode*. The latter allows to give the solution of the dynamical model for  $t = dt$  knowing the initial condition  $X_0$ .

The model error treatment (physics informed model error):

As the external heat source  $r(t)$  is unknown, the use of a relevant treatment for the model error is necessary to compensate this lack of knowledge. The solution of the stationary heat equation with noisy right-hand side is chosen to treat the model error. To determine it we solve the related stationary equation with  $r(t) = r$  unknown constant

$$\begin{aligned} -\alpha \frac{\partial^2 X}{\partial x^2}(x, t) &= r && \text{since } \frac{\partial X}{\partial t}(x, t) = 0 \\ \iff \frac{\partial X}{\partial x}(x, t) &= -\frac{r}{\alpha}x + c_1 \\ \iff X(x, t) &= -\frac{r}{2\alpha}x^2 + c_1x + c_2. \end{aligned}$$

The constants  $c_1$  and  $c_2$  are determined thanks to the boundary conditions

$$\begin{aligned} X(0, t) &= c_2 = 0 \\ X(1, t) &= -\frac{r}{2\alpha} + c_1 = 0 \Rightarrow c_1 = \frac{r}{2\alpha}. \end{aligned}$$

So the solution of the stationary equation is  $w(x) = \frac{r}{2\alpha}(-x^2 + x)$ . This solution is made noisy by replacing  $r$  by a one-dimensional noise  $r_k^i \sim \mathcal{N}(0, \sigma^2)$ , which refers to the ensemble member  $i$  and the time  $t_k$  for the forecast step. The parameter  $\sigma$  is to specify by practice. The resulting model error treatment is

$$w_k^i = \begin{pmatrix} w_k^i(x_1) \\ \vdots \\ w_k^i(x_j) \\ \vdots \\ w_k^i(x_n) \end{pmatrix} \quad \text{with } w_k^i(x_j) = \frac{r_k^i}{2\alpha}(-x_j^2 + x_j). \quad (13)$$

The algorithm used to estimate the temperature  $X(x, t)$  is detailed in Table 1. In the forecast step, the term  $M[X_{k-1}^{a,i}]$  corresponds to the solution at  $t = dt$  of the dynamical model with the initial condition  $X_0$  replaced by  $X_{k-1}^{a,i}$ . In the analysis step,  $p$  points on the bar among  $n$  are observed, these observation data are contained in  $Y_k$ .

Table 1: EnKF with the physics informed model error.

<p><i>Initialization:</i> for <math>i = 1, \dots, N</math>  generate <math>w_1^i</math>  <math>X_1^{a,i} = X_0 + w_1^i</math>  For <math>k \geq 2</math>  <i>Forecast:</i> for <math>i = 1, \dots, N</math>  generate <math>w_k^i</math>  <math>X_k^{f,i} = M[X_{k-1}^{a,i}] + w_k^i</math>  <math>X_k^f = \frac{1}{N} \sum_{i=1}^N X_k^{f,i}</math>  <math>P_k^f = \frac{1}{N-1} \sum_{i=1}^N (X_k^{f,i} - X_k^f)(X_k^{f,i} - X_k^f)^T</math>  <i>Analysis:</i> for <math>i = 1, \dots, N</math>  generate <math>\varepsilon_k^i \sim \mathcal{N}(0, R)</math>  <math>K_k = P_k^f H^T (H P_k^f H^T + R)^{-1}</math>  <math>d_k^i = Y_k + \varepsilon_k^i - H X_k^{f,i}</math>  <math>X_k^{a,i} = X_k^{f,i} + K_k d_k^i</math>  <math>X_k^a = \frac{1}{N} \sum_{i=1}^N X_k^{a,i}</math></p>
---

## 5 Numerical results

The efficiency of the treatment of the model error previously detailed is assessed by comparing this method to reference methods used to mitigate the model error in the EnKF. The following algorithms used for the comparison embed the same

dynamical model for the forecast step, which corresponds to the homogeneous heat equation (11) where the diffusion coefficient  $\alpha$  is known. However the treatment for the model error in the forecast step is different for each algorithm. The analysis step is the same for all the algorithms, with the same observations simulated by following (1b).

The algorithm that treats the model error by using the method in Section 4 is called the Physics Informed Model Error algorithm (PIME) and the associated parameter  $\sigma$  is denoted as  $\sigma_{PIME}$ . The other algorithms used for the comparison treat the model error using a Gaussian white noise  $\eta_k^i$ , see (2), but with a different covariance matrix  $Q$ . One uses  $Q = \sigma_{QD}^2 I_n$  ( $I_n$  is the  $n$  by  $n$  identity matrix): it is called the  $Q$ -Diagonal algorithm ( $QD$ ). Whereas the other puts  $Q(i, j) = \sigma_{QSS}^2 e^{-\lambda|x_i - x_j|}$  for  $1 \leq i, j \leq n$ : it is called the  $Q$ -Spatial Structure algorithm ( $QSS$ ) because it takes into account the position on the bar. The values of the diagonal of this covariance matrix are equal to  $\sigma_{QSS}^2$  and the more the elements of the matrix are away from the diagonal, the more the values of these elements are close to zero, with a speed that depends on the value of the spatial scale parameter  $\lambda$ . The parameters  $\sigma_{PIME}$ ,  $\sigma_{QD}$ ,  $\sigma_{QSS}$  and  $\lambda$  are to optimize by practice.

As said in Section 4, the real state is the solution of equation (9). The goal of this study is to compare the efficiency of each algorithm to retrieve the real state, without knowing the external heat source  $r(t)$ .

For the real state we put  $r(t) = 0.1 \sin(t)$ , to have an external heat source that will alternatively reheat or refresh the bar over time.

The parameters values for the three algorithms are given in Table 2. Among the  $n$  points of interest on the bar, only one in two is observed with low temporal frequency (every time period  $dt$ ) and with an important observation noise.

Parameter	Value
$n$	100
$p$	50
$N$	30
$R$	$0.01 I_n$
$K_{final}$	30
$\alpha$	0.05
$dt$	1

Table 2: Parameters values.

Each algorithm is assessed using the global Root Mean Square Error

$$\text{Global RMSE} = \frac{1}{K_{final}} \sum_{k=1}^{K_{final}} \sqrt{\frac{1}{nN} \sum_{i=1}^N (X_k^{a,i} - X_k)^T (X_k^{a,i} - X_k)}. \quad (14)$$

For  $QSS$ , the global RMSE is computed for different values of  $\sigma_{QSS}$  and  $\lambda$ . The global RMSE of  $QSS$  is more sensitive to the value of  $\sigma_{QSS}$  than to the

value of  $\lambda$ . To minimize the global RMSE of QSS with respect to  $\lambda$ , we put  $\lambda = 0.01$ .

The parameters  $\sigma_{PIME}$ ,  $\sigma_{QD}$  and  $\sigma_{QSS}$  are optimized by minimizing the global RMSE for each algorithm. To this end, we put  $\sigma_{PIME} = \sigma_{QD} = \sigma_{QSS} = \sigma$  and vary  $\sigma$  using the logarithmic scale  $\{10^{-5}, 10^{-4.9}, 10^{-4.8}, \dots, 10^{-0.1}, 1\}$ . The global RMSE of each algorithm for each value of  $\sigma$  is plotted in Figure 1.

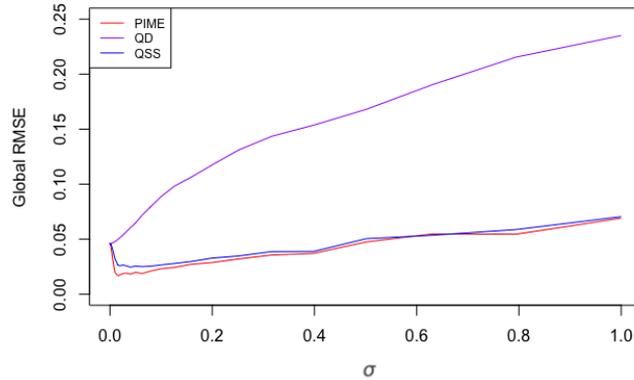


Figure 1: Optimization of  $\sigma_{PIME}$ ,  $\sigma_{QD}$  and  $\sigma_{QSS}$ .

For each algorithm, the value of  $\sigma$  that minimizes the global RMSE is shown in Table 3. The numerical results reported hereafter are obtained using these values.

$\sigma_{PIME}$	0.016
$\sigma_{QD}$	0.001
$\sigma_{QSS}$	0.050

Table 3: Optimal values for  $\sigma_{PIME}$ ,  $\sigma_{QD}$  and  $\sigma_{QSS}$

The temporal evolution of the heat diffusion in the bar is shown in Figure 2. The average of the analysis ensemble  $X_k^a$  is used to obtain the image of the diffusion for each algorithm. Contrary to QD and QSS, PIME partially retrieves the reheating of the bar (at  $t_9$ ,  $t_{15}$  and  $t_{28}$ ) due to the unknown external heat source.

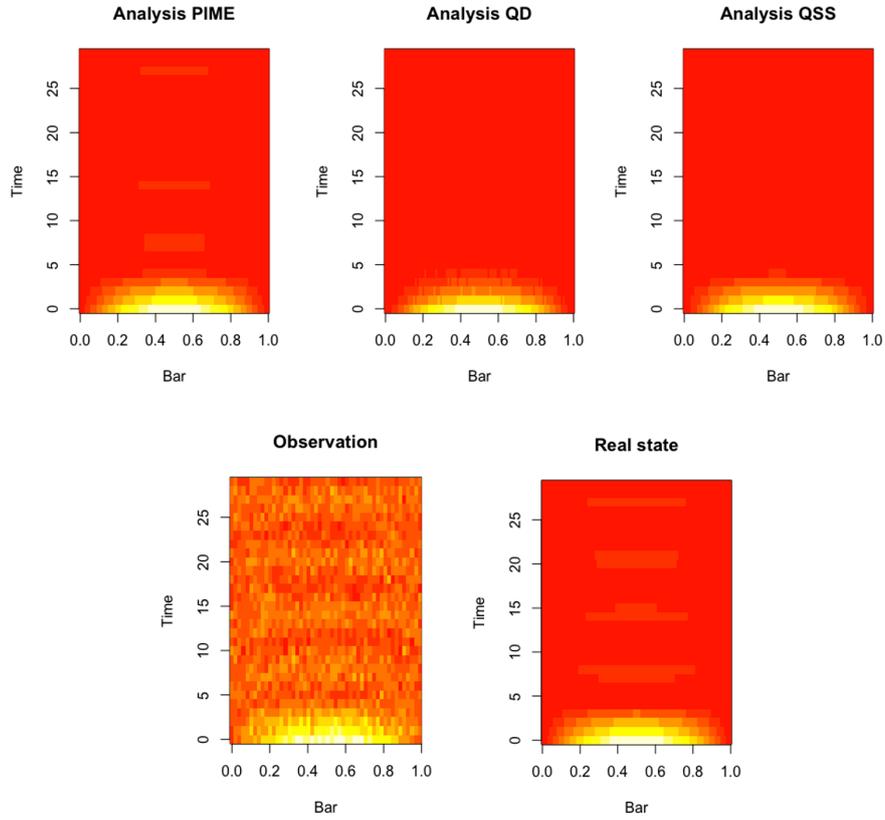


Figure 2: Comparison of the evolution of the heat diffusion for the analysis of each algorithm (the more the color is red, the more the temperature is close to zero).

The global RMSE is computed for each algorithm and given in Table 4. PIME is the most accurate, followed by QSS and QD.

Algorithm	Global RMSE
PIME	0.017
QD	0.048
QSS	0.025

Table 4: Global RMSE of each algorithm.

Focusing on time  $t_4$ ,  $X_4^f$  and  $X_4^a$  are plotted for each algorithm with the real state  $X_4$  and the observation  $Y_4$  in Figure 3.

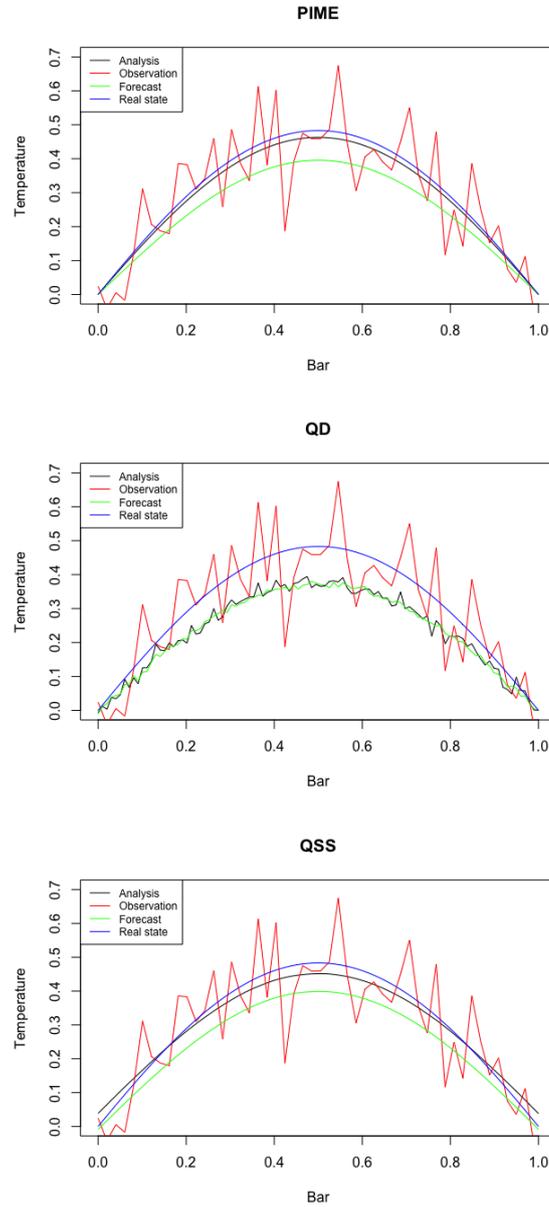


Figure 3: Comparison of the values of  $X_4^f$  and  $X_4^a$  for the different algorithms.

In Figure 3, for the three algorithms the analysis is better than the forecast as expected. PIME gives the best analysis. The one of QSS is slightly less

accurate, mainly at the boundaries of the bar.  $QD$  gives the worst analysis and the temperature fields are not smooth in space: this is a consequence of using a diagonal matrix for  $Q$  and thus no spatial correlation for the model error treatment.

The efficiency of each algorithm is now studied for the middle of the bar, which is the point where the error is the most important. To achieve this, the value of the 50th component of  $X_k^a$  is plotted as a function of  $t_k$  for each algorithm. In Figure 4, PIME and  $QSS$  can retrieve the oscillations of the temperature due to  $r(t)$ , with a better accuracy for PIME. Whereas  $QD$  is unable to reconstruct these oscillations.

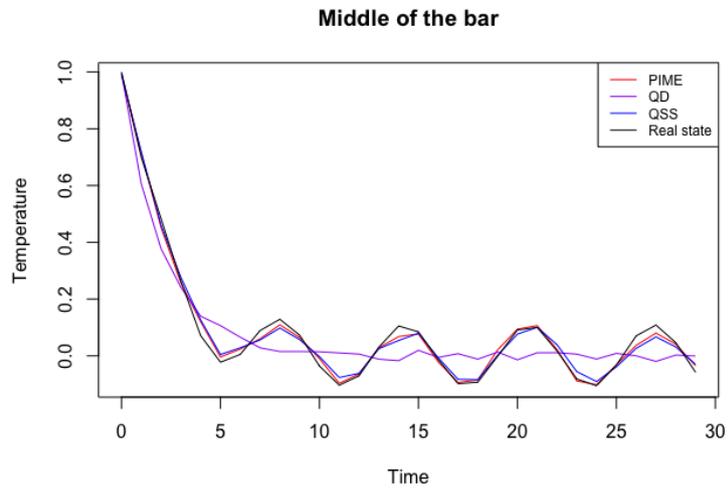


Figure 4: Temporal evolution of the estimated temperature of the middle point for each method.

The same study is done when the time period between two consecutive observations is increased to  $dt = 1.5$ . As there are less observations that contain informations about  $r(t)$ , the model error is more important and its treatments are to adapt by optimizing  $\sigma_{PIME}$ ,  $\sigma_{QD}$  and  $\sigma_{QSS}$  with the method related to Figure 1.

PIME again gives the best estimate in Figure 5 and widens the gap with the other methods.

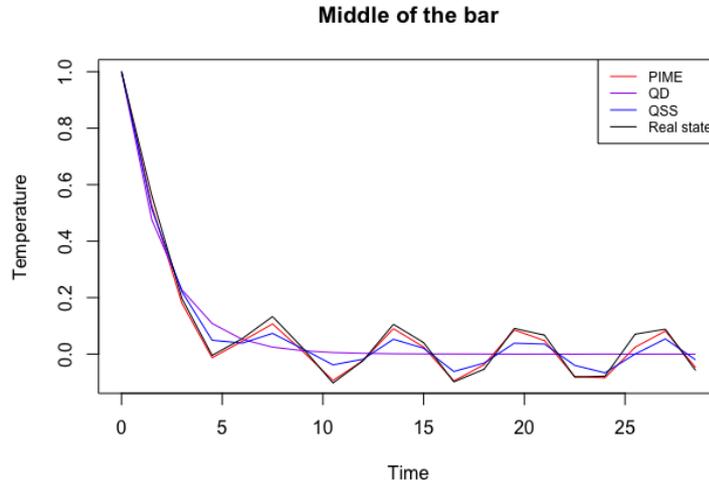


Figure 5: Temporal evolution of the estimated temperature of the middle point for each method with  $dt = 1.5$ .

The efficiency of PIME, in comparison with the other algorithms, is more obvious when  $dt$  is important. This is also shown in Figure 6, where the global RMSE for each value of  $dt$  is plotted for each algorithm. The parameters  $\sigma_{PIME}$ ,  $\sigma_{QD}$  and  $\sigma_{QSS}$  are optimized for each value of  $dt$ .

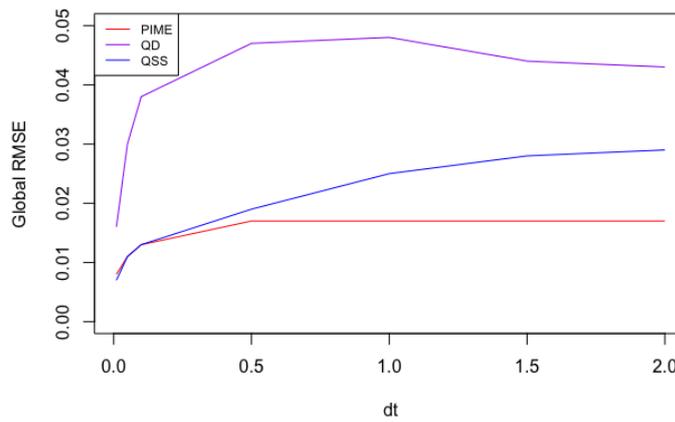


Figure 6: Global RMSE of each algorithm according to the value of  $dt$ .

When  $dt$  is smaller, PIME and QSS are more accurate and have nearly the same error. There are more observations that give informations on the real

states, that is why the error of each algorithm is less important. The added value of PIME cannot be seen easily when  $dt$  is too small, because the system does not evolve enough to let appear a sort of stationary state which is approached by PIME thanks to its physics informed model error.

Lastly, to check if the results discussed above for PIME are robust versus random perturbations in the observation noise and in the noise of the model error treatment, a confidence interval for the temporal evolution of the RMSE of PIME is built. To this end, the experience of the heat diffusion in the bar is repeated 100 times for PIME with the parameters values of Table 2 and  $\sigma_{PIME} = 0.016$ . The real state, dynamical model and model error remain the same, whereas the values of the observations and of the model error treatment vary from one experience to another because of the random noise. At each experience, the temporal evolution of the RMSE of PIME is determined by computing the RMSE for each time  $t_k$

$$\text{RMSE}(t_k) = \sqrt{\frac{1}{nN} \sum_{i=1}^N (X_k^{a,i} - X_k)^T (X_k^{a,i} - X_k)}. \quad (15)$$

The average of these temporal evolutions of RMSE is plotted in red in Figure 7 and the 95% confidence interval in grey is computed for each time  $t_k$ .

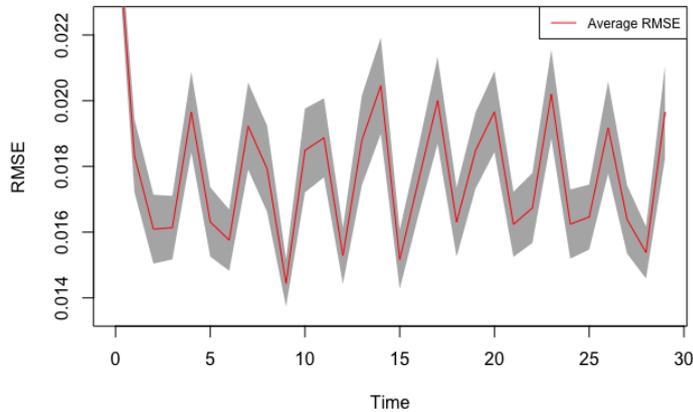


Figure 7: Confidence interval for the temporal evolution of the RMSE of PIME.

The evolution of the RMSE depends on the evolution of the external heat source: larger RMSE values are obtained when  $|r(t_k)|$  is larger.

The peak at  $t_1$  is due to  $X_1^{a,i}$  for which the initial condition  $X_0$  is perturbed by  $w_1^i$  to generate the ensemble members (see Table 1), producing first a relatively important RMSE that then decreases by using the dynamical model in

the forecast step at  $t_2$ . This peak is also observed for the other algorithms and for the same reason.

The confidence interval is globally narrow. This shows that PIME is stable with respect to the observation noise and the noise in the model error treatment.

## 6 Conclusion

A new treatment for the model error in the EnKF was introduced, taking further into account the physics of the system. This treatment is a randomized solution of the stationary PDE involved in the dynamical model. This method was applied to the heat diffusion in a bar, where the goal was to estimate the temperature of the bar over time without knowing the external heat source. In this context, our method was compared to reference methods to mitigate the model error. The numerical results showed the efficiency of our physics informed model error, especially when the frequency of the observations was low.

The online estimate of  $\sigma_{PIME}$  by using the state augmentation method may be interesting to further vary in time the physics informed model error and could lead to better results.

The physics informed model error may be extended to more complex physical phenomena such as the estimation of the sea surface temperature (SST).

## References

- [1] Pierre Ailliot, Anne Cuzol, Gilles Durrieu, Hélène Flourent, Emmanuel Frénod, Jules Guillot, Jean-Paul Lucas, and François Septier. Synthèse des questions mathématiques soulevées par la mise en oeuvre de jumeaux numériques pour le suivi et le pilotage de systèmes dynamiques en entreprises. *hal-03167416*, 2021.
- [2] Jeffrey L Anderson and Stephen L Anderson. A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Monthly weather review*, 127(12):2741–2758, 1999.
- [3] Judith Berner, Ulrich Achatz, Lauriane Batte, Lisa Bengtsson, Alvaro De La Camara, Hannah M Christensen, Matteo Colangeli, Danielle RB Coleman, Daan Crommelin, Stamen I Dolaptchiev, et al. Stochastic parameterization: Toward a new view of weather and climate models. *Bulletin of the American Meteorological Society*, 98(3):565–588, 2017.
- [4] Julien Brajard, Alberto Carrassi, Marc Bocquet, and Laurent Bertino. Combining data assimilation and machine learning to infer unresolved scale parametrization. *Philosophical Transactions of the Royal Society A*, 379(2194):20200086, 2021.

- [5] Gérald Desroziers, Loic Berre, Bernard Chapnik, and Paul Poli. Diagnosis of observation, background and analysis-error statistics in observation space. *Quarterly Journal of the Royal Meteorological Society: A journal of the atmospheric sciences, applied meteorology and physical oceanography*, 131(613):3385–3396, 2005.
- [6] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans*, 99(C5):10143–10162, 1994.
- [7] Alban Farchi, Patrick Laloyaux, Massimo Bonavita, and Marc Bocquet. Using machine learning to correct model error in data assimilation and forecast applications. *Quarterly Journal of the Royal Meteorological Society*, 147(739):3067–3084, 2021.
- [8] Michael Ghil and Paola Malanotte-Rizzoli. Data assimilation in meteorology and oceanography. *Advances in geophysics*, 33:141–266, 1991.
- [9] Redouane Lguensat, Pierre Tandeo, Pierre Ailliot, Manuel Pulido, and Ronan Fablet. The analog data assimilation. *Monthly Weather Review*, 145(10):4093–4107, 2017.
- [10] Tim N Palmer, Roberto Buizza, F Doblas-Reyes, Thomas Jung, Martin Leutbecher, Glenn J Shutts, Martin Steinheimer, and Antje Weisheimer. Stochastic parametrization and model uncertainty. *ECMWF Technical Memoranda*, 2009.
- [11] Pierre Tandeo, Pierre Ailliot, Marc Bocquet, Alberto Carrassi, Takemasa Miyoshi, Manuel Pulido, and Yicun Zhen. A review of innovation-based methods to jointly estimate model and observation error covariance matrices in ensemble data assimilation. *Monthly Weather Review*, 148(10):3973–3994, 2020.