PARTIAL DIFFERENTIAL EQUATIONS FOR OCEANIC ARTIFICIAL INTELLIGENCE

Jules Guillot[1], Guillaume Koenig[2], Hadi Minbashian[3], Emmanuel Frénod[4], Hélène Flourent[5] and Julien Brajard[6]

Abstract. The Sea Surface Temperature (SST) plays a significant role in analyzing and assessing the dynamics of weather and also biological systems. It has various applications such as weather forecasting or planning of coastal activities. On the one hand, standard physical methods for forecasting SST use coupled ocean- atmosphere prediction systems, based on the Navier-Stokes equations. These models rely on multiple physical hypotheses and do not optimally exploit the information available in the data. On the other hand, despite the availability of large amounts of data, direct applications of machine learning methods do not always lead to competitive state of the art results. Another approach is to combine these two methods: this is data-model coupling. The aim of this paper is to use a model in another domain. This model is based on a data-model coupling approach to simulate and predict SST. We first introduce the original model. Then, the modified model is described, to finish with some numerical results.

Keywords: Data Assimilation, Data-Model Coupling, Artificial Intelligence, SST, PDE

1. Introduction

The general task of this assignment could have been guessed from the title "Partial differential equations for oceanic artificial intelligence". The idea was to use artificial intelligence tools to determine the SST of a given oceanic zone.

This had already been done in Bezenac *et al.* [1]. What they exactly did is using a Convolutional Neural Network (CNN) and a transport equation to predict the evolution of the field of surface temperature for a few days. The CNN was used to identify and predict the velocity field used in the transport equation. They did not consider the boundary conditions for the transport equation nor the diffusion coefficient.

However, other approaches could have been followed. We could have a more physical representation of the problem by adding some terms. Machine learning has been used in this context by Zhang and Lin [7] and could identify the non-linear terms present in the output of an hydrodynamic model. In this case, they did not use a Neural Network, but sophisticated parameters identification tools such as stochastic gradient descent and LASSO objective function.

[1] LMBA/UMR 6205, University of South Brittany

[2] MIO, University of Marseille

[3] Numerical Analysis and Scientific Computing, Department of Mathematics, Technical University of Darmstadt

[4] LMBA/UMR 6205, University of South Brittany and See-d, Vannes

[5] LMBA/UMR 6205, University of South Brittany

[6] LOCEAN, IPSL

Another interesting approach was the one of Chen *et al.* [6] and Ruthotto and Haber [5]. Their idea was to mimick the behavior of a Neural Network by differential equations, ordinary in the first paper and partial in the second. The interest of this approach would be to use the computationnally efficient tools that exists for identifying parameters in differential equations to replace the back propagation in Neural Networks.

In a data-model coupling context, learning PDEs (related to unknown phenomena) from data, had been done in Long *et al.* [3]. More recently, a new Neural Network architecture based on PDEs has been made by Pannekoucke and Fablet [4]. Another work is currently led on the link between discretized PDEs and the layers of a Deep Neural Network.

Finally, we decided to follow a different approach with all this in mind. We decided to use a model with an advection-diffusion problem and some ODEs that was already developed in Flourent *et al.* [2] and to try to fit this to our SST evolution problem. The idea was that the PDE was close enough to the physics of our problem to be efficient, and that the ODEs could be fitted in order to represent the phenomenon not taken into account in the PDE.

2. The model

The model used is a modified version of the model of Flourent *et al.* [2]. The idea is to use an advection-diffusion equation for a given variable and have it passed through several transformations given by ODEs to account for other forcings. In our case, we are interested in the evolution of the SST in time and our forcing variable is the solar radiation at the top of the atmosphere. To simplify our problem, we reduced it to the determination of the average of the temperature on a given zone for an entire year.

$$\frac{\partial \Phi_f}{\partial t}(t,x) + \omega \frac{\partial \Phi_f}{\partial x}(t,x) - c \frac{\partial \left[\chi \frac{\partial \Phi_f}{\partial x} \right]}{\partial x}(t,x) = Q(t,x) - fF(x)\Phi_f(t,x) \tag{1}$$

$$\frac{\partial \Psi}{\partial t}(t,x) = fF(x)\Phi_f(t,x) - u\Psi(t,x) \tag{2}$$

$$\frac{\partial \Xi}{\partial t}(t,x) = -u\Psi(t,x)\Big(\Xi(t,x) - sup\Big) - P\Big(\Xi(t,x) - inf\Big) \tag{3}$$

$$s(t) = \int_\Omega \Xi(t,x)\,dx \tag{4}$$

Here the parameters are the following: Φ_f is the unknown advected quantity, ω is the advection velocity, c and χ are the scaling of the diffusion and the field of the coefficient of diffusion. Q corresponds to the source term and is here given by a scaled version of the solar radiation at the top of the atmosphere. f and F are terms transfering quantities from the advection-diffusion equation towards the ODEs. This was inspired from a biological model where this transfer consists in the extraction of nutrients from the flow by biological agents, see [2].

The parameter Ψ is the first "hidden quantity" related to the non-represented phenomenon. We see that it increases by extracting quantity from the PDE and decrease with the term $u\Psi$ which transfers some of its content to the next ODE.

The quantity Ξ corresponds to another hidden variable. Its behavior is slightly different from Ψ because its equation is not linear and contains saturation terms. The first term links the evolution of Ξ to what is extracted from the quantity Ψ.

Finally we have the quantity $s(t)$ which basically contains the total of the quantity Ξ at each time t.

In our case, we thought that the PDE was able to correctly reproduce the heat absorbed by the water when the solar radiation arrives. And we thought that the ODEs would be able to reproduce efficiently other phenomenona affecting the link between incoming solar radiation and water surface temperature, such as the cloud coverage, the difference in opacity of water or the wind chill for example. But if we wanted those phenomenona to be correctly reproduced, we needed to change some of the parameters.

3. Application to our problem

3.1. General view of the problem

As said earlier, we tried to identify a working relationship between solar radiation at the top of the atmosphere and mean value of the SST. To do so, we tried to fix three parameters in our system : sup, inf and u. Those are not related with the PDE but with the ODEs. We have used data taken from some part of the North Atlantic and solar radiation that we collected from the CERES project.
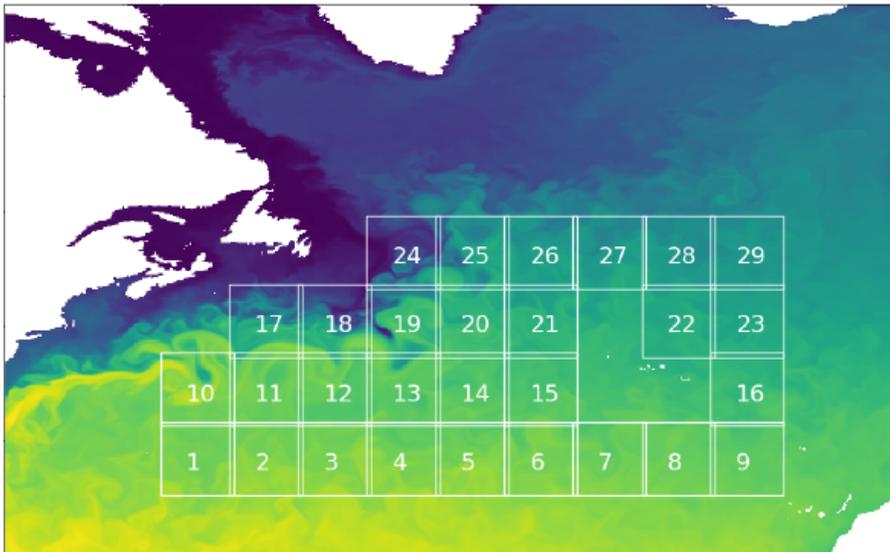

FIGURE 1. Localisation of our dataset. It is present in the rectangle 20.

Finally, we got interested in the spatial average of those data. Also, the variations over the years were of interest because they would give us a minimum error to reach with our model. We then did the spatial averaging which is presented in the figure 2.

3.2. More details on the data

Looking at data, we see a very clear seasonal trend associated with some high frequency variations. This seasonal trend is common to every year while the high frequency variations differ from one year to another. At first we will try to identify this seasonal trend, because it is directly linked to the sun radiation arriving on the surface. To see this, we can have a look at the figure 3.
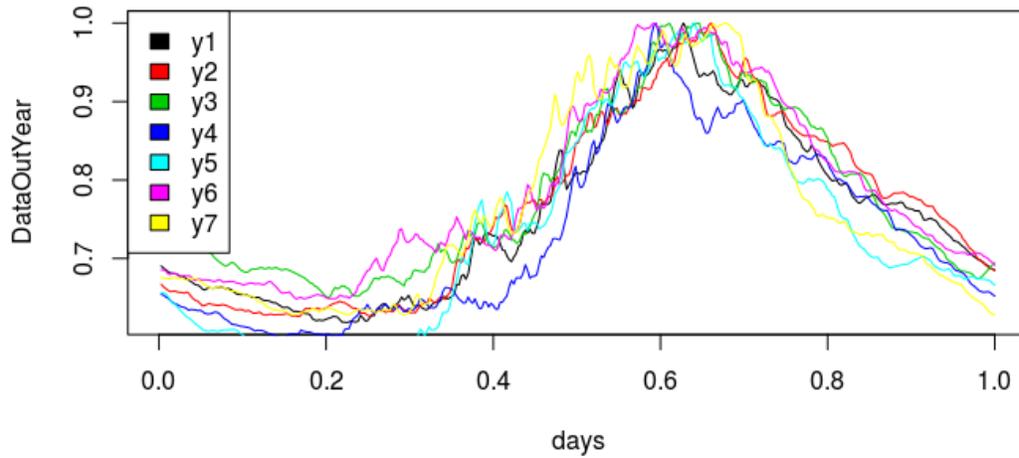
FIGURE 2. Variations of the spatial average of the temperature over different years. This has been normalized to be compatible with our model.
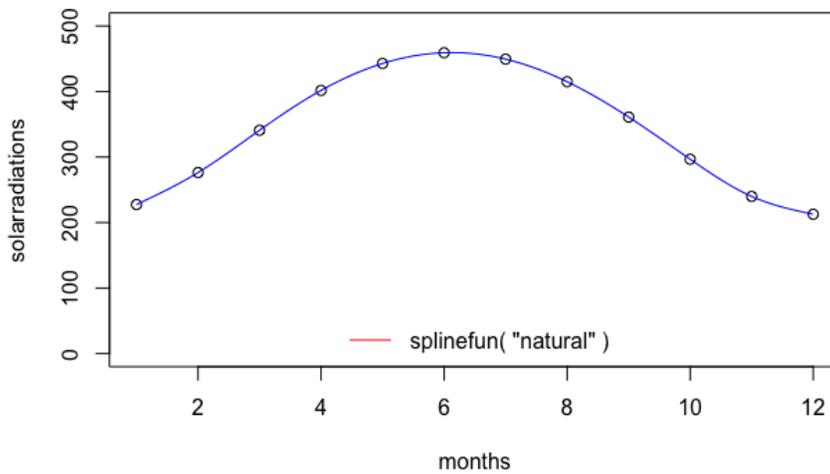

FIGURE 3. Sun radiation arriving at the top of the atmosphere for a standard year in $W.m^{-2}$. The variations between different years are very weak and therefore this year is taken as representative.

3.3. **Physics behind the model**

Though we see a global connection in the shape of the curves of SST and solar radiation, we also see discrepancies. The two large discrepancies are the delay between the moment of highest irradiance (in June) and the moment of highest temperature (in July and August) and the high frequency signal in temperature, which is absent from the solar radiation signal. Finally, the important interannual variations in temperature that are not present in solar radiation are also intriguing.

It is interesting to consider the origin of those data before trying to explain the discrepancies between the curves. First of all, the temperature data come from a model of oceanic circulation (NEMO) and are taken from the upper region of the oceanic model. This upper region is taken as the highest region in the vertical discretization of the model. This highest region is then a cell that represents a few meters of water. Both physically and numerically, that region is affected by complex phenomenona that we will describe further in this paper. However, here, it is interesting also to note that this model is corrected by a variational algorithm that incorporates satellite data for the SST. However, those temperatures correspond to the skin temperature of the water, which is a layer of a few centimeters between the atmosphere and the upper ocean. It is often considered that the temperature in this layer is correlated with the temperature of the first meters of water.

The data for solar radiations come from satellite observations given at the CERES. They correspond to the solar radiation reaching the top of the atmosphere, and not to solar radiation reaching the surface of the Earth. Therefore, it is not affected by meteorological phenomenona such as clouds or the concentration of chemical species in the atmosphere. That explains the small interannual variability. Therefore, the major part of their variations is due to the tilt of the rotation axis of the Earth and to the variation of distance Earth-Sun during a year. This distance and the tilt of the rotation axis of Earth change during each year, but very slowly. That explains the low interannual variability and the absence of high frequency signal.

Now we can try to point out some phenomenona that explain interannual variability and the high frequency signal in the SST. As stated above, the state of the atmosphere, such as cloud coverage or the concentration of some chemical species in the atmosphere can have a strong effect on the radiation reaching the surface and therefore on the energy available to heat the ocean. Purely oceanic parameters, such as the state of the sea surface, the opacity of water or the depth of the surface layer may also affect the way the radiation heat up the upper layer. Finally, external phenomenona such as the temperature of the water when it enters the zone, the temperature of the atmosphere or the heat transfer with the lower part of the ocean may also play a role. All those phenomenona may affect both the high frequency variations and the delay observed, as well as the interannual variability of the SST.

Another effect that plays a role in the delay is the high heat capacity of water. Therefore, it takes some time before the sun radiation heats the water up and we end up with a maximum of temperature that is delayed compared with the maximum of solar radiation.

3.4. **Link with the model**

In our case, we use the solar radiation forcing directly into the advection-diffusion equation. We therefore count on it to transfer the part directly linked to the solar radiation. However, to model all the other phenomenona responsible notably for the delay or the high frequencies, we rely on the ODEs. For this reason, we decided to focus our efforts on the determination of the parameters of the ODEs, for those represent the less understood part of the problem.

4. NUMERICAL RESULTS

We began by testing the model of Flourent *et al.* [2] directly with our forcings and our SST data and try to see if it was able to reproduce correct behavior. We began by a learning phase where we tried to learn the different *sup*, *inf* and *u* parameters, for an average of the different years for the temperature. This is available on the figure 4. Those were obtained with the input shown on figure 5. The parameters of the model are presented on the table 1.
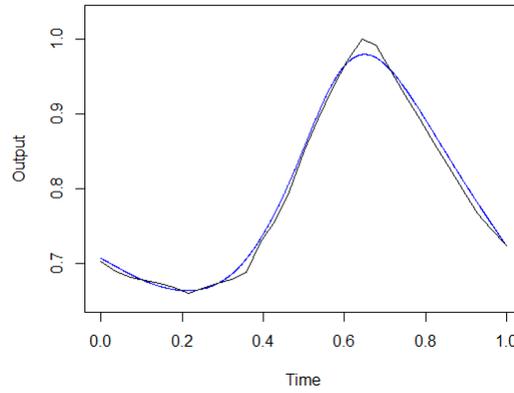
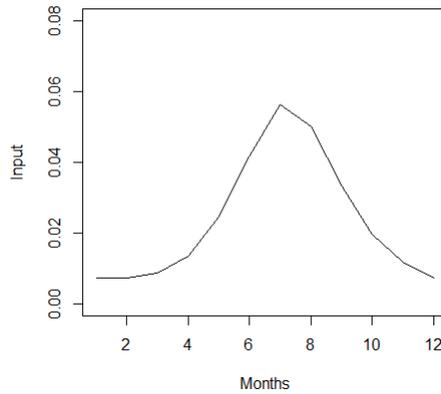FIGURE 4. Comparison between model outputs (in blue) and data for an average year (in black).

FIGURE 5. Input function used for the model of Flourent *et al.* [2].

TABLE 1. Values of the parameters.

Parameter	Value
ω	20
c	0.0002
f	500
P	4.3
u	150
sup	13.5
inf	0.44
$InitSortie$	0.72
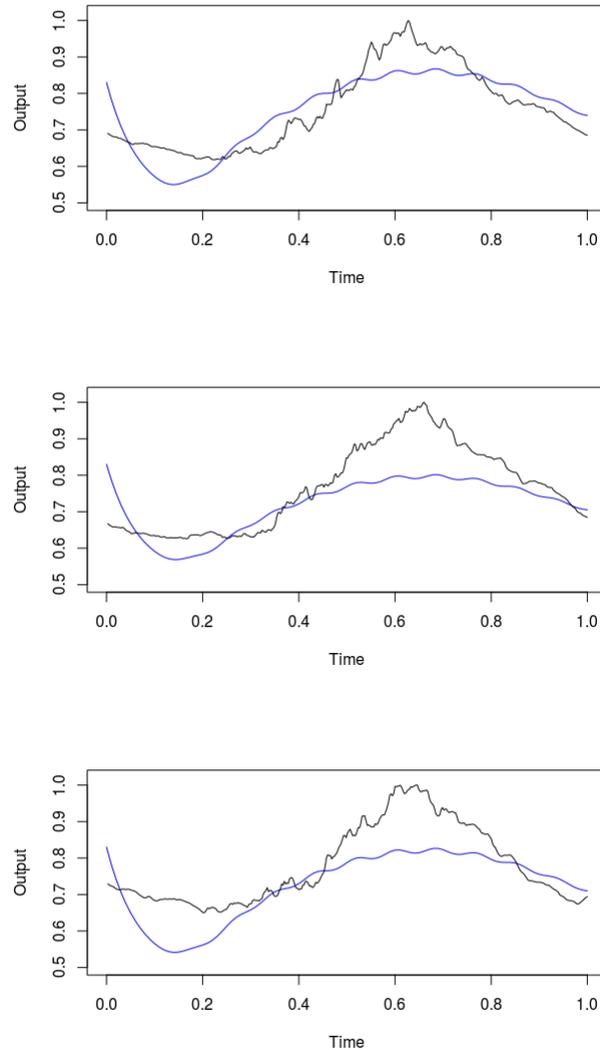
FIGURE 6. Comparison between the model outputs in blue and the data for different years.

4.1. **Interannual variability**

We then tried the same methodology but for the SST of different years. Since the solar radiation varied only slightly, we kept the same values for each year. The idea was to see how the model would adapt to those variations. The data are presented on the figure 6 and the associated parameters on the table 2.

There are some obvious features. We see that there is a correct fit in the timing of the highest temperature, but that the values are quite off. The period of spring though is correctly fitted. Maybe the two biggest discrepancies are the beginning of the year and the fact that there is no high frequency variations. For the high frequency, since we have a low frequency input function and a linear model applied on it, it was expected.

Year	sup	inf	u	objective function
1	56.6872427983148	0.5037952	199.999999999961	0.0488598607471862
2	68.4254272816738	0.4976314	199.999993030828	0.0461400247339915
3	51.7736625511663	0.5224552	199.999999999882	0.0579528378552093
4	62.5269300303108	0.4888889	199.999999999908	0.0554632691635029
5	68.8096166836958	0.4444428	199.999811832358	0.0640272984624972
6	52.5925925925795	0.5375095	199.999999999961	0.0575910849643628
7	54.1394604472804	0.5199167	199.999999999961	0.0733631601644184

TABLE 2. Parameters determined for the different years.

4.2. **2D-extension of the model**

We also tried to use Freefem++ and Fenics to build a 2D representation for the model. Both are finite element modelling software. We used them for solving the PDE and the explicit Euler scheme for the ODEs. However, we were facing an issue: the finite-elements modelling are not good for problem of advection-diffusion. Therefore, we decided to use only the advection problem and to take advantage of the numerical diffusion to play the role of the physical diffusion. This approach was limited since it is really dependant on the scale of the elements in the model and on the time step used. Also, it is quite sensitive to the family of elements used.

A first example of discretization and solving of the advection problem is available on the figure 7.
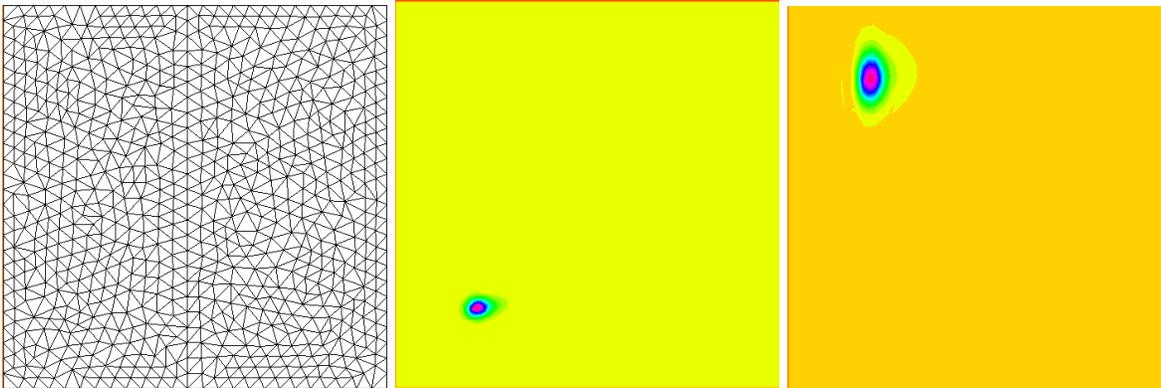

FIGURE 7. Grid, initial condition and result of the advection model after a few time step for the Freefem model with only an advection term and the ODEs.

We also tried it with a finer discretization. We observe less diffusion as expected. We still were using a second order polynomial basis. Results are available on the figure 8. Another tryout was to see the impact of the time step on the diffusion. Results for that are shown on the figure 9. Finally we tried with a discontinuous Galerkin basis to disminish the diffusion. The results are shown on the figure 10.

5. CONCLUSION

We adapted a model, based on a data-model coupling approach, to another framework with the same underlying phenomena: the SST data-assimilation. This model contains a PDE and ODEs with parameters learnt by the available data. To go further, we could use a Neural Network to optimize those parameters instead of the Statistical Learning Tool of the previous model.
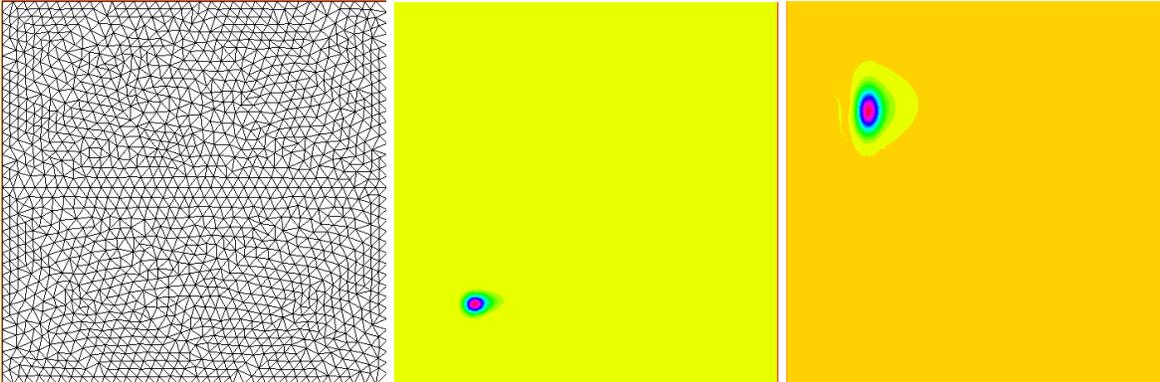
FIGURE 8. Grid, initial condition and result after a few time step for an advection problem and ODEs.
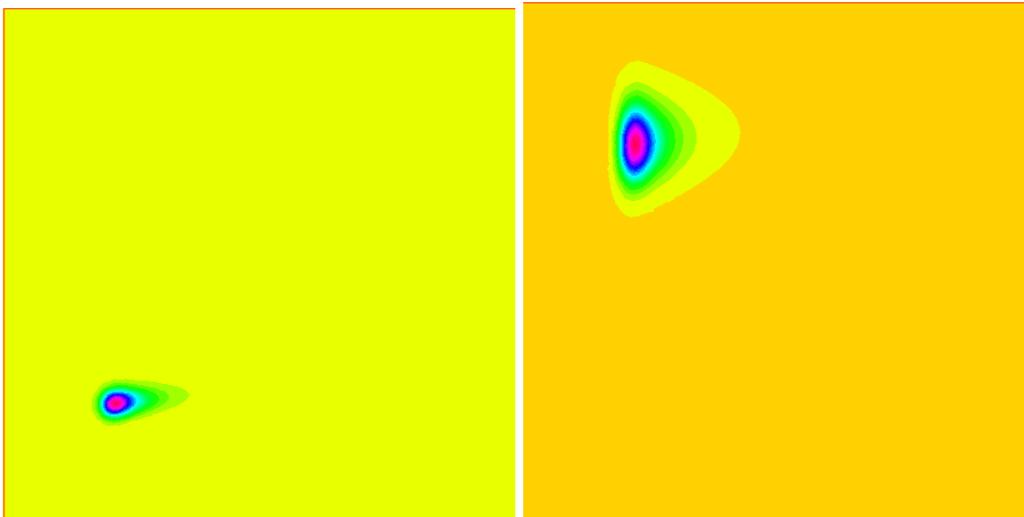

FIGURE 9. Same situation as on the figure 8 but with a different time step.

Acknowledgments

References

[1] Emmanuel De Bézenac, Arthur Pajot, and Patrick Gallinari. Deep learning for physical processes: Incorporating prior scientific knowledge, 2018. https://arxiv.org/abs/1711.07970v2.

[2] Hélène Flourent, Emmanuel Frénod, and Vincent Sincholle. An innovating statistical learning tool based on partial differential equations, intending livestock data assimilation, 2019. https://hal.archives-ouvertes.fr/hal-02079750.

[3] Zichao Long, Yiping Lu, Xianzhong MA, and Bin Dong. Pde-net: Learning pdes from data, 2018. https://arxiv.org/abs/1710.09668.
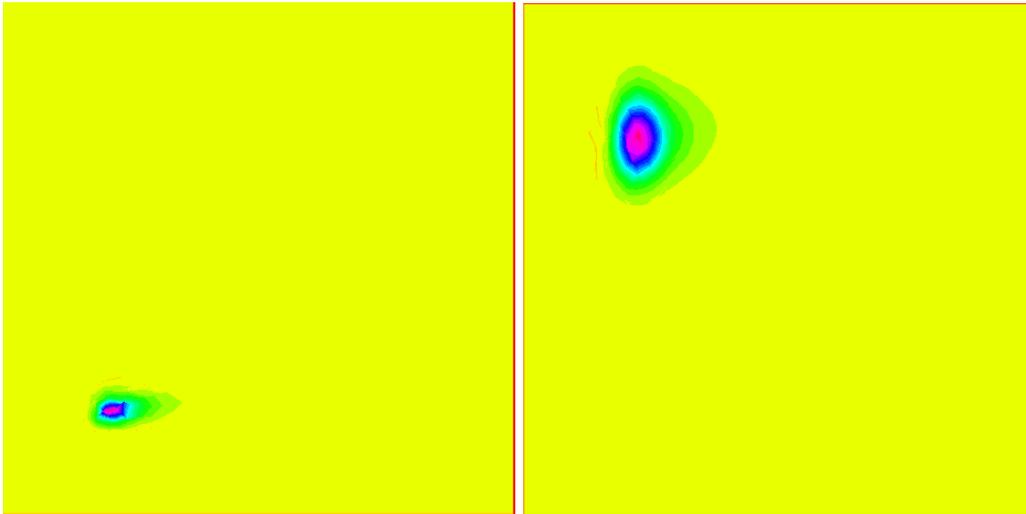
FIGURE 10. Same situation as on the figure 9 but with a discontinuous Galerkin basis.

[4] Olivier Pannekoucke and Ronan Fablet. Pde-netgen 1.0: from symbolic pde representations of physical processes to trainable neural network representations, 2020. `https://arxiv.org/abs/2002.01029`.

[5] Lars Ruthotto1 and Eldad Haber. Deep neural networks motivated by partial differential equations, 2018. `https://arxiv.org/abs/1804.04272v2`.

[6] Ricky T.Q Chen, Yulia Rubanova, Jesse Bettencourt, and David Duvenaud. Neural ordinary differential equations, 2019. `https://arxiv.org/abs/1806.07366v5`.

[7] Sheng Zhang and Guang Lin. Robust data-driven discovery of governing physical laws with error bars, 2018. `https://royalsocietypublishing.org/doi/10.1098/rspa.2018.0305`.