Metamodeling the water temperature

Michaël Baudin [1] Fabien Souillé [1]
with contributions by Chu Mai

[1]EDF R&D. 6, quai Watier, 78401, Chatou Cedex - France, michael.baudin@edf.fr

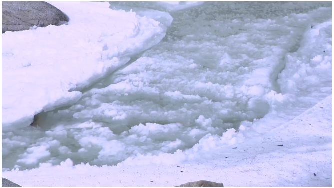
June 2021

Contents

Goals

Industrial topic:

- Frazil ice are small crystal disks.
- They may appear when the water temperature falls below the fusion point.
- The growth can be fast.
- They may accumulate on water intake racks of nuclear power plants [Souillé et al., 2020]: we focus on the Blayais power plant.



Frazil ice in Yosemite Creek, [Commons, 2020]

Blayais CNPE, © OpenStreetMap contributors

Goals

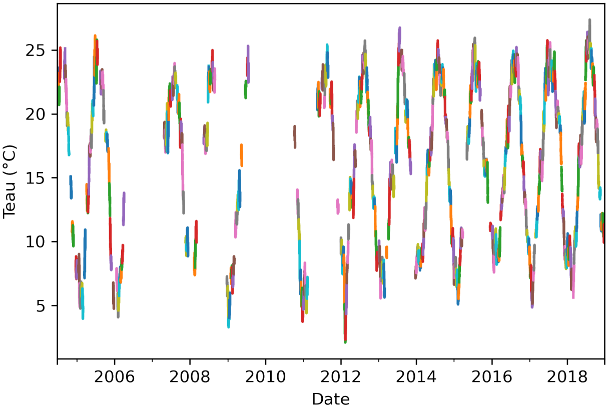
Goals of the study :

- We have measures of water temperature, air temperature, water level and flowrate of the Gironde river in the estuary.
- We want to predict the frazil near the Blayais nuclear power plant.
- We make the hypothesis that the frazil appears when the water temperature of the Gironde falls below 0° C.
- The simulation horizon is 7 days: we make the hypothesis that the water temperature is known from days 1 to 6 and want to predict the water temperature on day 7 (we may want to update this time interval).

Data

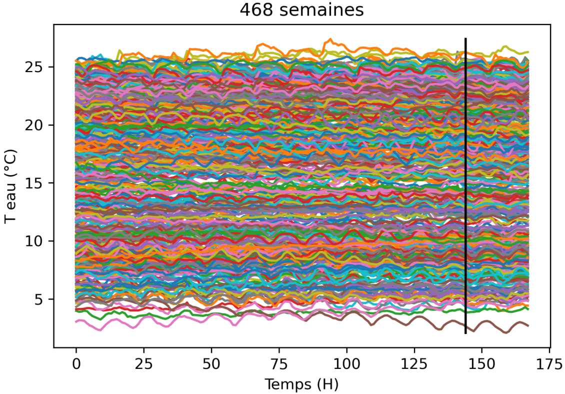
- We have water observations from the Magest station from 2005 to 2016 [Schmidt, 2014].
- We also have observations from the power plant water temperatures sensors from the Blayais power plant for 2015, 2016 and 2017.
- Some observations are lacking for short or longer times (from minutes to months).

Data



Observations of the water temperature from the Magest station and Blayais power plan, merged from 2005 to 2019 at hour time step : 468 continuous weeks.

Data



Observations of the water temperature from the Magest station and Blayais power plan, merged from 2005 to 2019 at hour time step : 468 continuous weeks. The vertical black line represents the cut between the days 1 to 6 (input) and the 7th day (output to be predicted).

Mathematical problem

Below are the goals and constraints:

1. Predict $T_{\text{future water}}$ depending on $T_{\text{past water}}$.
2. Using directly the temporal grid in input and output leads to a too large number of input ($24 \times 6 = 144$ time steps) and output variables (24 time steps).
3. We would like to change easily the time grid.
4. The link between input and output is complicated and not necessarily linear.
5. Possible to take into account other physical variables such as air temperature or water level (not necessarily discretised on the same grid).
6. The observations are noisy and the sample set is large (hundreds).
7. We would like to estimate a confidence interval of the future temperature.

Suggested solutions

Suggested solutions :

1. Reduce the dimension using the Karhunen-Loève decomposition (e.g. [Sullivan, 2015], chap.11).
2. Use sparse generalized polynomial chaos metamodel [Ghanem and Spanos, 2003, Xiu and Karniadakis, 2002, Maître and O.M.Knio, 2010, Blatman and Sudret, 2011].
3. The method suggested in [Dutfoy and Lebrun, 2017] combine both previous methods (and discusses sensitivity analysis).

Karhunen-Loève decomposition

The Karhunen-Loève decomposition of a stochastic process $X(t)$ is:

$$X(t) = \sum_{k \geq 0} \sqrt{\lambda_k} \xi_k \varphi_k(t)$$

where $\xi_k \in \mathbb{R}$ is the k-th coefficient in the decomposition:

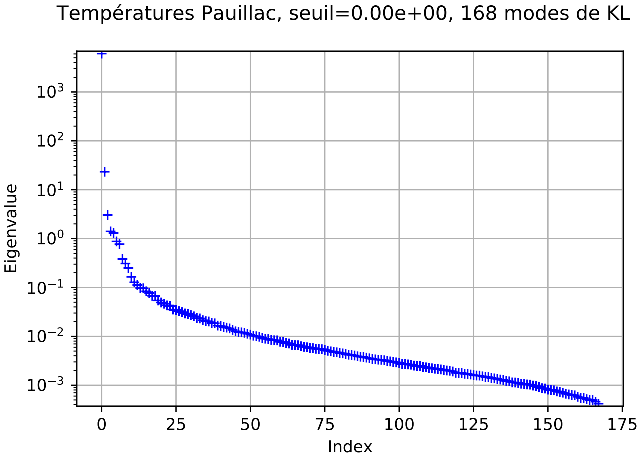
$$\xi_k = \frac{1}{\sqrt{\lambda_k}} \int_{\mathcal{D}} X(t) \varphi_k(t) dt$$

for $k \geq 0$ and

- $(\lambda_k)_{k \geq 0}$ are the eigenvalues,
- $(\varphi_k)_{k \geq 0}$ are the eigenvectors,

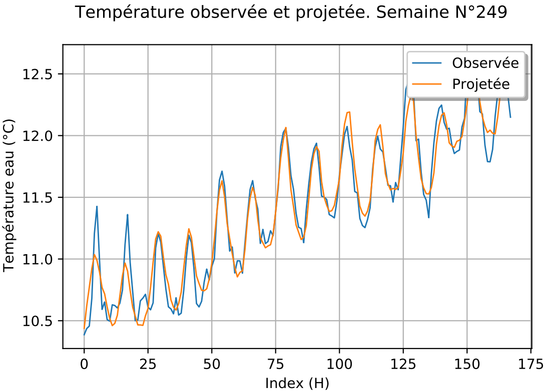
of the Fredholm problem.

Karhunen-Loève decomposition



Decreasing sequence of eigenvalues of the Karhunen-Loève decomposition of the Gironde water temperature time series.

Karhunen-Loève decomposition



Comparison of the 249th trajectory of the Gironde water temperature and its projection on the truncated Karhunen-Loève decomposition with 9 modes.

Field metamodel

$$
\begin{tikzcd}
X \arrow[r, "g"] \arrow[d, "\pi"'] & Y \\
(\xi_k)_{k \in \mathbb{N}} \arrow[r, "g_\xi"] & (\xi_k')_{k \in \mathbb{N}} \arrow[u, "(\pi')^{-1}"']
\end{tikzcd}
$$

Dimension reduction and metamodel of the stochastic process for water temperature of the Gironde.

- The function g is the model from the input (past) water temperature X to the output (future) water temperature Y.
- The projection of X on the K.L. basis leads to the coefficients $(\xi_k)_{k\geq 0}$ and the projection of Y leads to the coefficients $(\xi_k')_{k\geq 0}$.
- The model g_ξ maps the K.L. coefficients.
- The metamodel $\tilde{g}_\xi$ is a surrogate for g_ξ.

Field metamodel: polynomial chaos decomposition

The method has three steps: project, predict, lift.

1. Project. We decompose the input on $m_X \in \mathbb{N}$ K.-L. modes

$$\xi_k(\omega) = \frac{1}{\sqrt{\lambda_k}} \int_{\mathcal{D}_{past}} X(t,\omega)\varphi_k(t)dt$$

   for $k = 1, \ldots, m_X$, where $\mathcal{D}_{past} = [0, t_{present}]$.

2. Predict. Map from the input to the output coefficients of the basis, based on a vector-to-vector metamodel :

$$\boldsymbol{\xi}'(\omega) = \tilde{g}_{\boldsymbol{\xi}}(\boldsymbol{\xi}(\omega))$$

   where $\boldsymbol{\xi} \in \mathbb{R}^{m_X}$, $\boldsymbol{\xi}' \in \mathbb{R}^{m_Y}$.

3. Lift. Compute the output on $m_Y \in \mathbb{N}$ K.-L. modes :

$$Y(t,\omega) = \sum_{k=0}^{m_Y} \sqrt{\lambda_k^{(Y)}} \xi'_k(\omega)\varphi_k^{(Y)}(t)$$

   for any $t \in \mathcal{D}_{future} = [t_{present}, t_{final}]$.

Field metamodel: KL decomposition of inputs and ouputs

We consider a basis of orthogonal multivariate polynomials ψ_j, that is, for which the weighted $\mathcal{L}^2$ scalar product is zero for different polynomials:

$$(\psi_j, \psi_k) = 0,$$

for $j \neq k$.
The output coefficients in the K.L. decomposition can be decomposed into a linear combination of multivariate polynomials:

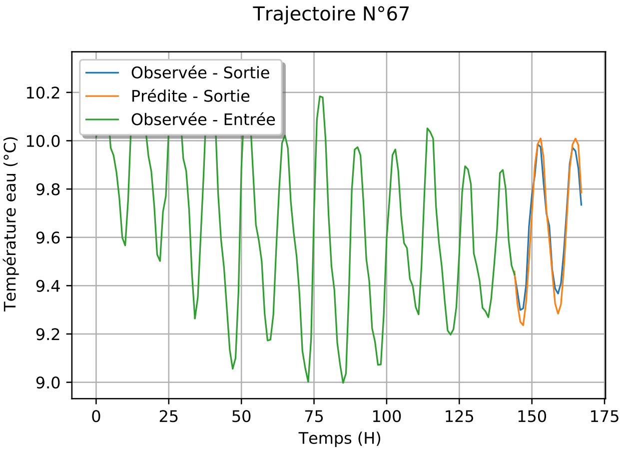
$$\boldsymbol{\xi}' = \sum_{j=0}^{p} \boldsymbol{a}_j \psi_j \left(\mathring{\boldsymbol{\xi}}\right), \qquad \mathring{\boldsymbol{\xi}} = T^{-1}(\boldsymbol{\xi})$$

where $p \in \mathbb{N}$ is the number of coefficients in the truncated polynomial chaos decomposition, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_p \in \mathbb{R}^{m_Y}$ are the coefficients of the decomposition, $m_X \in \mathbb{N}$ is the number of input K.-L. modes, $\mathring{\boldsymbol{\xi}} \in \mathbb{R}^{m_X}$ is a random vector, $m_Y \in \mathbb{N}$ is the number of output K.-L. modes, $\boldsymbol{\xi}' \in \mathbb{R}^{m_Y}$ is the output random vector, T is the isoprobabilistic transformation, $\boldsymbol{\xi} \in \mathbb{R}^{m_X}$ are the input K.L. coefficients.

Field metamodel

- We use $m_X = 8$ K.L. coefficients on input and $m_Y = 4$ K.L. coefficients on output.
- The distribution of the input is modeled by histograms, with independent marginals.
- The multivariate polynomial basis was based on linear enumeration.
- The polynomial was created based on a degree 3 sparse least squares polynomial : 31 coefficients to estimate.
- The cross validation (75% for training) leads to a predictivity coefficient equal to 99% for the first output K.L. mode and between 49% and 62% for modes 2, 3, 4.
- The standard deviation of the metamodel residual is an increasing function of time (remote future predictions are less accurate) and between 0.17 and 0.38 °C.
- Some trajectories, especially with late change of trend, are predicted less accurately.

Field metamodel



For a given trajectory in the validation sub-sample, comparison with the trajectory predicted by the metamodel.

Conclusions and future works

OpenTURNS classes used:

- This work was performed using OpenTURNS v1.16.
- We used the `KarhunenLoeveSVDAlgorithm` for the K.L. decomposition and `FunctionalChaosAlgorithm` for P.C. decomposition.
- The example [OpenTURNS Consortium, 2021] showing how to create a metamodel of a field function was used as a starting point.

OpenTURNS evolutions: development of new classes:

- new field classes: `VertexValuePointToFieldFunction`, `KarhunenLoeveReduction`, `KarhunenLoeveValidation`.
- new methods for K.L. validation: `KarhunenLoeveResult.drawEigenValues()` and `ProcessSample.drawCorrelation()`
- new metamodel classes: `KFoldSplitter`, `LeaveOneOutSplitter`

End

Future work:

- Provide the `FieldMetamodelAlgorithm`, `FieldMetamodelResult` and `FieldMetamodelValidation` classes in OpenTURNS?
- Analysis of a simplified ordinary differential equation which models dynamics of frazil ice concentration.
- Estimate confidence intervals for the temperature mean and of a future observation.

Thank you for your attention!
Questions?

OpenTURNS: www.openturns.org

OpenTURNS

An Open source initiative for the Treatment of Uncertainties, Risks'N Statistics

- Multivariate probabilistic modeling including dependence
- Numerical tools dedicated to the treatment of uncertainties
- Generic coupling to any type of physical model
- Open source, LGPL licensed, C++/Python library

Références I

Baudin, M., Dutfoy, A., Iooss, B., and Popelin, A.-L. (2016).
OpenTURNS: An Industrial Software for Uncertainty Quantification in Simulation, pages 1–38.
Springer International Publishing, Cham.

Baudin, M. and Souillé, F. (2020).
Métamodélisation de la température de l'eau dans l'estuaire de la Gironde.
Technical Report 6125-3119-2020-03392-FR, EDF R&D.

Blatman, G. and Sudret, B. (2011).
Adaptive sparse polynomial chaos expansion based on least angle regression.
Journal of computational Physics, 230(6):2345–2367.

Commons, W. (2020).
File: Frazil ice in Yosemite Creek.png — Wikimedia Commons, the free media repository.
[Online; accessed 1-June-2021].

Références II

Dutfoy, A.-H. and Lebrun, R. (2017).
Méta modélisation et analyse de sensibilité de données fonctionnelles.
Technical Report 6125-1612-2017-01470-FR, EDF R&D, Airbus Group Innovation.

Ghanem, R. G. and Spanos, P. D. (2003).
Stochastic finite elements: a spectral approach.
Courier Corporation.

Maître, O. L. and O.M.Knio (2010).
Spectral methods for uncertainty quantification.
Springer.

OpenTURNS Consortium (2021).
Metamodel of a field function.
http://openturns.github.io/openturns/master/auto_meta_modeling/fields_metamodels/plot_fieldfunction_metamodel.html.

Références III

Schmidt, S. (2014).
Le réseau MAGEST en Gironde : retour d'expérience de 10 ans de suivi haute-fréquence de la qualité des eaux de l'estuaire de la Gironde.
somlit.epoc.u-bordeaux1.fr.

Souillé, F., Taccone, F., and El Mertahi, C. (2020).
A multi-class frazil ice model for shallow water flows.
In *Online proceedings of the papers submitted to the 2020 TELEMAC-MASCARET User Conference October 2020*, pages 122–129.

Sullivan, T. (2015).
Introduction to Uncertainty Quantification.
Springer International Publishing Switzerland.

Xiu, D. and Karniadakis, G. E. (2002).
The wiener–askey polynomial chaos for stochastic differential equations.
SIAM journal on scientific computing, 24(2):619–644.