A combined orthogonal decomposition and polynomial chaos methodology for data-based analysis and prediction of coastal dynamics

R-S. MOURADI [a,b,*], C. GOEURY[a], O. THUAL[b], F. ZAOUI[a] and P. TASSI[a]

From [a]LNHE, EDF R&D and [b]CECI, CERFACS, CNRS

I. Siltation of nearshore channels due to coastal morphodynamics

Figure 1: Silting of Bray Harbor, Irish Sea (Muir Éireann). *Source: Afloat Magazine.*

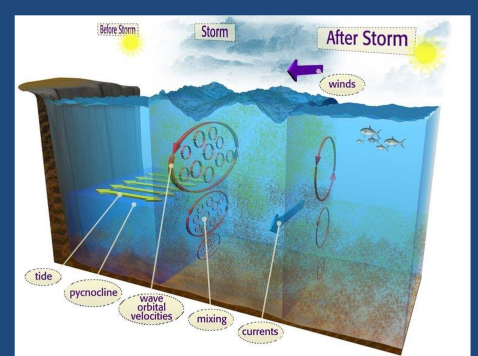


Figure 2: Sediment Resuspension. *Source: T. Miles, University of New Jersey (https://www.travisnmiles.com/)*

Sedimentation and siltation in nearshore channels is a well-known issue in harbors for example (e.g. Figure 1). They imply frequent dredging interventions, with high operational costs, often hindered by tight scheduling.

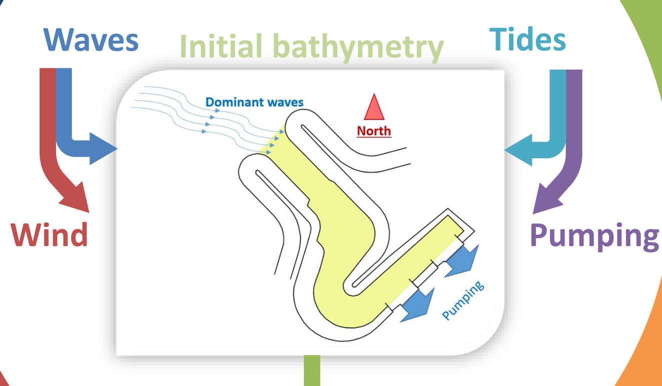
Power plant intakes are submitted to the same constraints, in addition to acting as sinks for sediments because of the water pumping, which attracts the sediments inside.

Outside the intake, many physical forcings influence the sediments dynamics and drive them to the channel.

Inside, a number of industrial forcings (pumping, dredging, etc.) impact the settling of sediments by their action on the flow. Therefore, a new bathymetry is obtained.

The context

The application:
A water intake in a coastal area, forced by many phenomena



II. Proper Orthogonal Decomposition on field measurements

The POD consists of writing an approximation of the bathymetry field $Z(x,t)$ as a finite sum of a separate variables functions product, at a given order $d \in \mathbb{N}^*$. This would be written as $Z(x,t) \approx \sum_{k=1}^{d} a_k(t)\boldsymbol{\varphi}_k(x)$. The functions $\boldsymbol{\varphi}_k(x)$ and $a_k(t)$ are resp. called spatial modes and temporal coefficients. They are orthogonal and are selected so that the order $d \in \mathbb{N}^*$ is minimum.

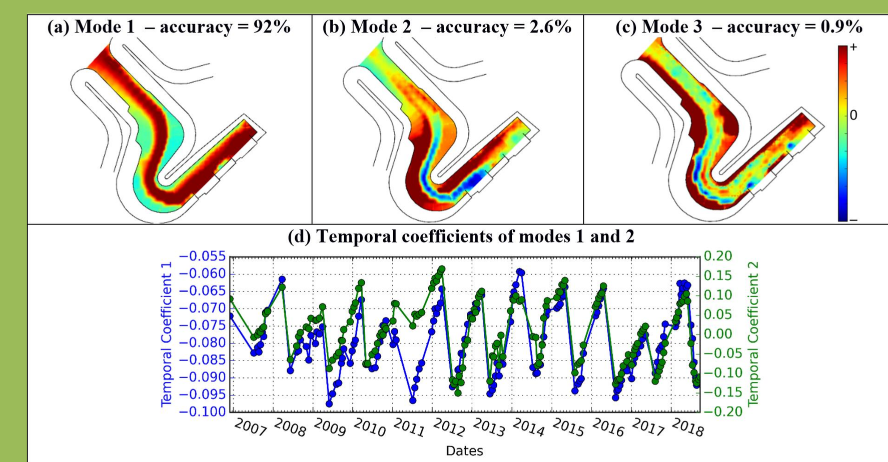


Figure 3: The first three elements of the POD basis applied to the bathymetry measurements and the first two associated temporal coefficients.

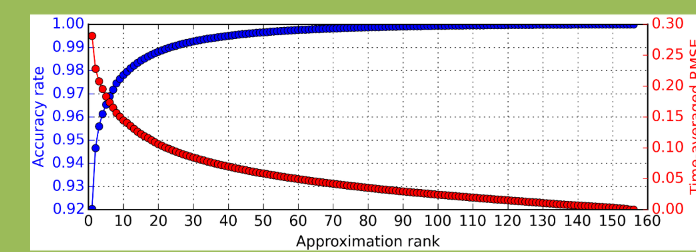


Figure 4: POD on bathymetry - Accuracy rate and time-averaged RMSE as functions of the approximation rank.

POD (Proper Orthogonal Decomposition)

Each spatial mode represents a dynamical pattern. The temporal variations of the associated temporal coefficients represent a certain percentage of the variance of $Z(x,t)$. The modes $a_k(t)\boldsymbol{\varphi}_k(x)$ are therefore associated to a represented variance percentage, here called ***"accuracy rate"*** (see Figure 3). When increasing the order $d \in \mathbb{N}^*$, these variance percentages are added, giving an increasing accuracy rate (see Figure 4).

III. Data-based learning with Polynomial Chaos Expansion

Let $(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_V)$ be a set of forcing parameters. We can construct a dynamical model $\boldsymbol{\mathcal{H}}_k$ for each a_k as: $a_k(t_2) = \boldsymbol{\mathcal{H}}_k(a_k(t_1), t_2 - t_1, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_V)$

If we consider that $(\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_V)$ live in the space of real random variables with finite second order moments, $\boldsymbol{\mathcal{H}}_k$ can be constructed with PCE. The latter allows a polynomial approximation of a random variable Y as:

$$Y(\mathbf{X}_1, \dots, \mathbf{X}_V) = \boldsymbol{\mathcal{M}}_0 + \sum_{i=1}^{V} \boldsymbol{\mathcal{M}}_i(\mathbf{X}_i) + \sum_{1 \le i < j \le V} \boldsymbol{\mathcal{M}}_{i,j}(\mathbf{X}_i, \mathbf{X}_j) + \dots + \boldsymbol{\mathcal{M}}_{1,\dots,V}(\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_V),$$

PCE (Polynomial Chaos Expansion)

where $\boldsymbol{\mathcal{M}}_0$ is the mean of Y and $\boldsymbol{\mathcal{M}}_{I \subseteq \{1,\dots,V\}}$ represents the common contribution of the variables $I \subseteq \{1,\dots,V\}$ on the variation of Y, in a polynomial form:

$$X_1^{\alpha_1} X_2^{\alpha_2} \dots X_V^{\alpha_V}$$

A "training set" is used to learn the PCE model, and a "prediction set" to evaluate it on real scenarios. As shown in Figure 5, the fitting works best when the signal shows some consistency. For a more chaotic function, as the temporal coefficient 3, the PCE fitting is poor, yet it approaches the order of magnitude and seems to capture some peaks in the dynamics.

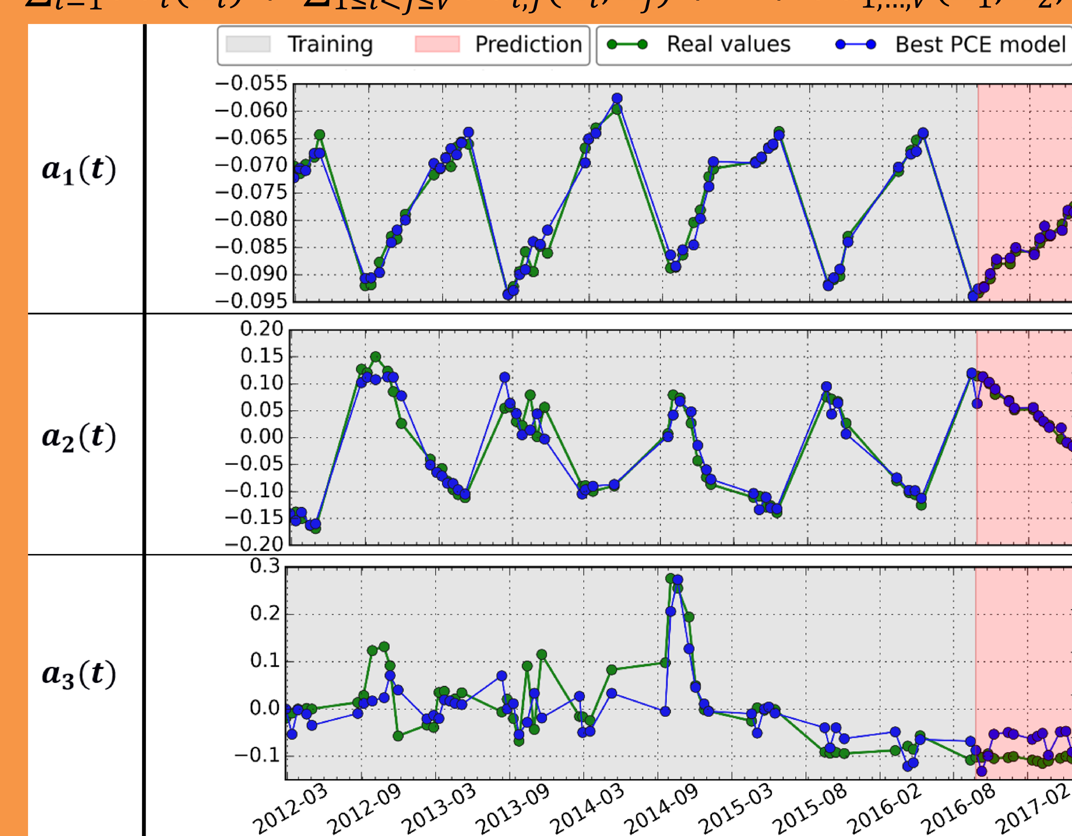


Figure 5: PCE fits for the first three POD temporal coefficients using a training set of 50 members. . The "best model" designation corresponds to a chosen polynomial degree with) minimal training RMSE (Root Mean Squared Error).

IV. POD-PCE coupling as a data driven predictor

Prediction

We attempt to project the estimation of future temporal coefficients $a_k(t_2)$ in the bathymetry POD basis, in order to construct a full prediction field as:

$$Z(x,t_2) \approx \sum_{k=1}^{d} a_k(t_2)\boldsymbol{\varphi}_k(x)$$
$$\approx \sum_{k=1}^{d} \boldsymbol{\mathcal{H}}_k(a_k(t_1), t_2 - t_1, \boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_V)\boldsymbol{\varphi}_k(x)$$

To identify the errors of all the algorithm's steps on the final prediction, we plot the associated averaged residuals in time, for each geographical point of the channel, as shown in Figure 6.

The POD error decreases when increasing the rank d. However, the PCE error increases dramatically from Rank 2 to 3, and becomes stationary for higher ranks. This is due to the fact that higher order temporal coefficients don't vary much. As a consequence, in order to decrease the forecasting error, a better approximation of coefficient 3 is essential.

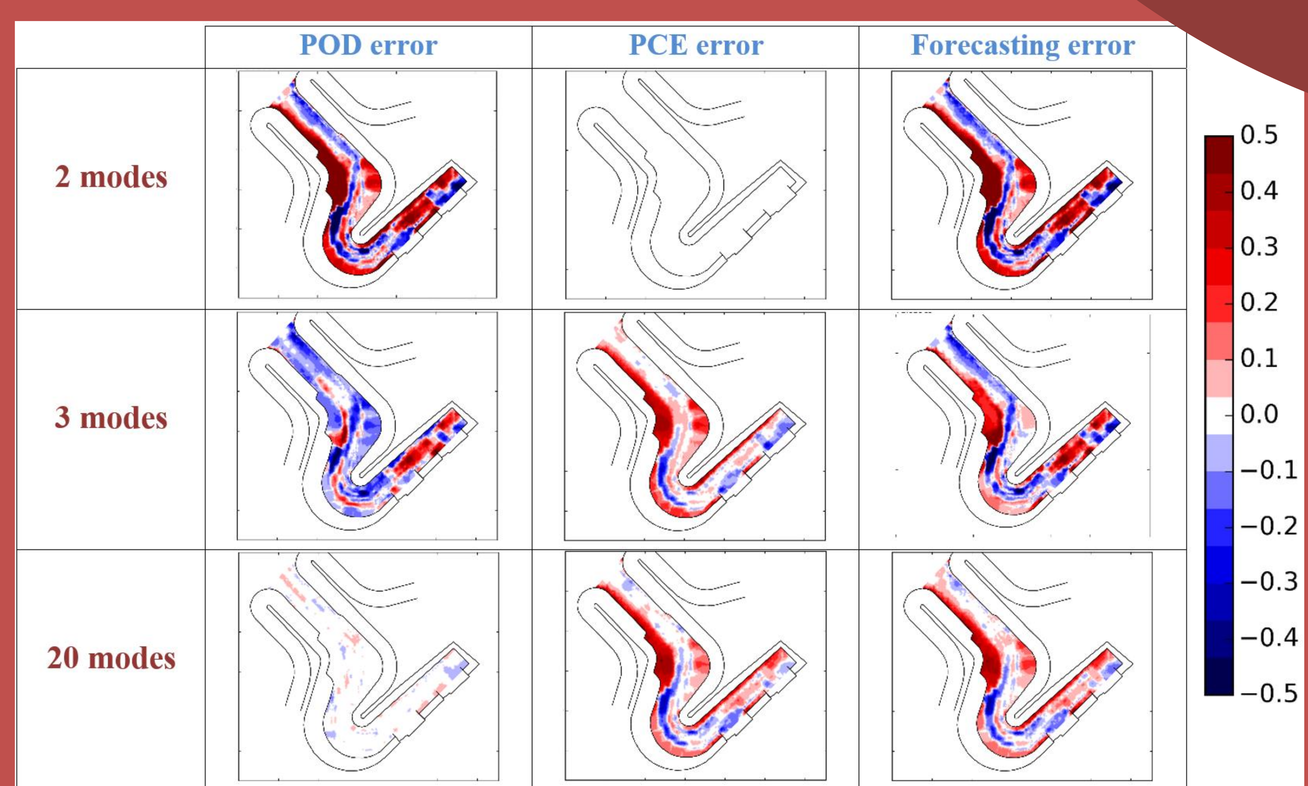


Figure 6: Time-averaged error for each approximation step. Evolution with the POD rank.

Contact: Rem-Sophia MOURADI, PhD student at EDF R&D and CERFACS. E-mail: rem-sophia-r.mouradi@edf.fr

EDF SA au capital de 924 433 331 euros - RCS Paris B 552 081 317 - Copyright EDF Médiathèque