A statistical approach to image denoising based on kriging and PODs with comments on statistical vs. physical models

Nicolas Durrande[1], Rodolphe Le Riche[1,2] , Stephane Avril[1,2]
[1] Ecole des Mines de St-Etienne
[2] CNRS UMR 5146

Physical versus statistical models (1/3)

Physical models

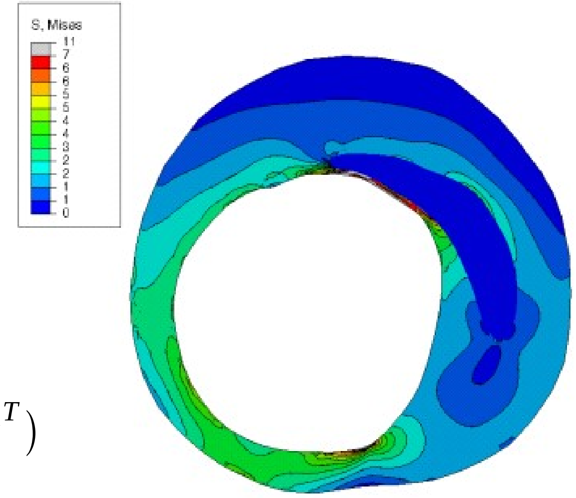
are based on a priori physical information. Typically conservation laws, kinematic descriptions, constitutive laws, boundary conditions, i.e., partial differential equations which are then numerically integrated with solutions in the volume and/or at the boundary.

They aim at being predictive with little experimental data.

Ex. Newtown's 2nd law, $\nabla . \sigma + F = \rho \ddot{u}$

Strain-displacement eq., $\varepsilon = \frac{1}{2}(\nabla u + \nabla u^{T})$

Constitutive eq., $\sigma = C : \varepsilon$



von Mises stresses in a carotid artery
(from Alexandre Franquet)

Physical versus statistical models (2/3)

Statistical models

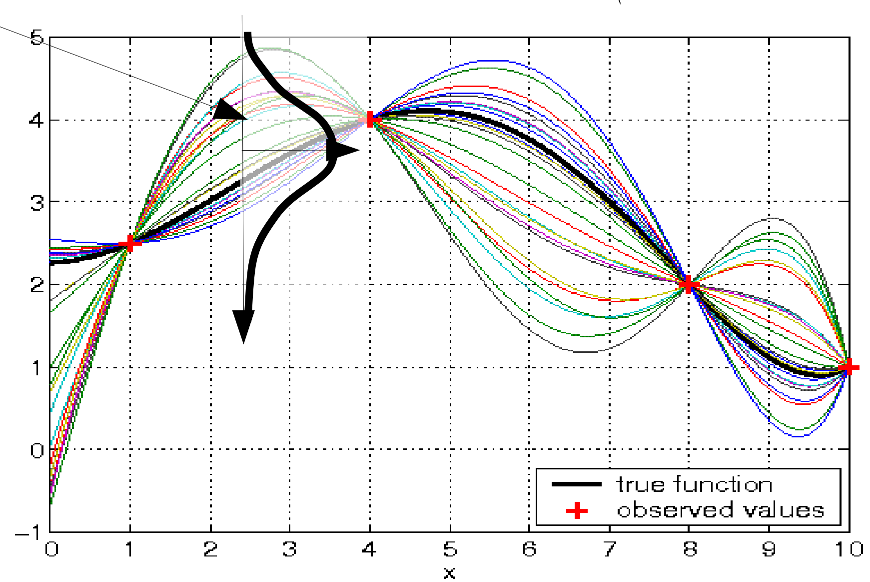
are built by tuning a statistical model from a set of inputs and outputs (red crosses hereafter).
• Do not (need to) calculate entire fields.
• Can usually handle data noise.
• Need data !

$$[\,F(x)\mid f(x^1),\ldots,f(x^t)\,] \sim N\left(m_{OK}(x)\,,\,s^2_{OK}(x)\right)$$

Ex. : ordinary kriging

$$F(x) = \mu + \varepsilon(x)$$

where $\varepsilon(x)$ is a Gaussian process.

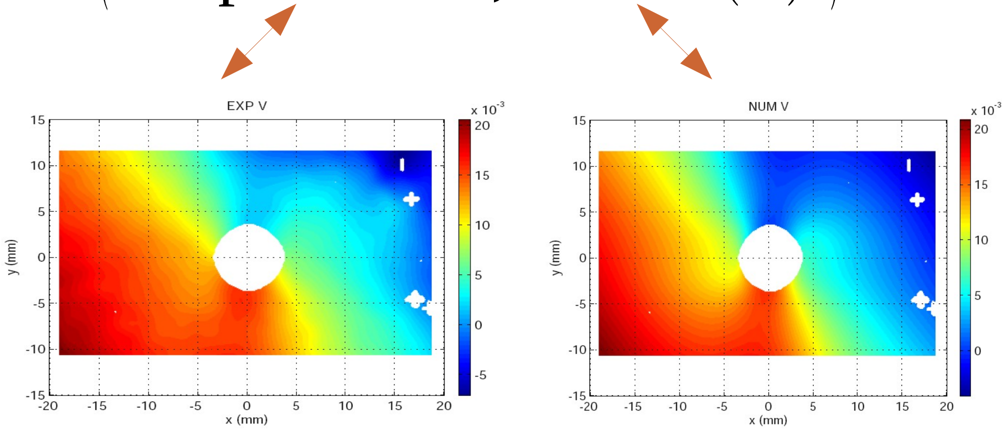


Physical versus statistical models (3/3)

The separation between physical and statistical models is not so clear.

Physical models are tuned from noisy data (inverse approach, identification), just like statistical models.

$$min_{\theta \in S} \text{distance}\left(\text{Experiment} \ , \text{Model}(\theta) \right)$$

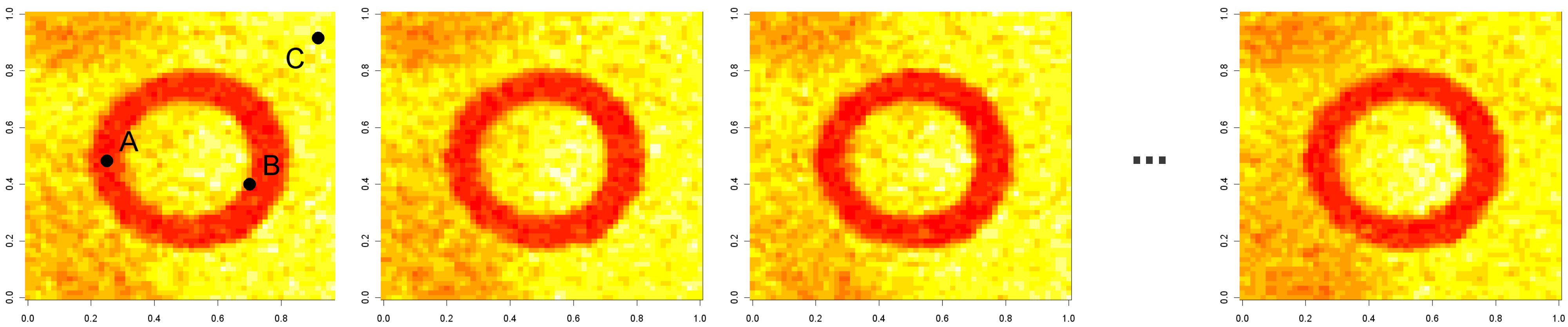


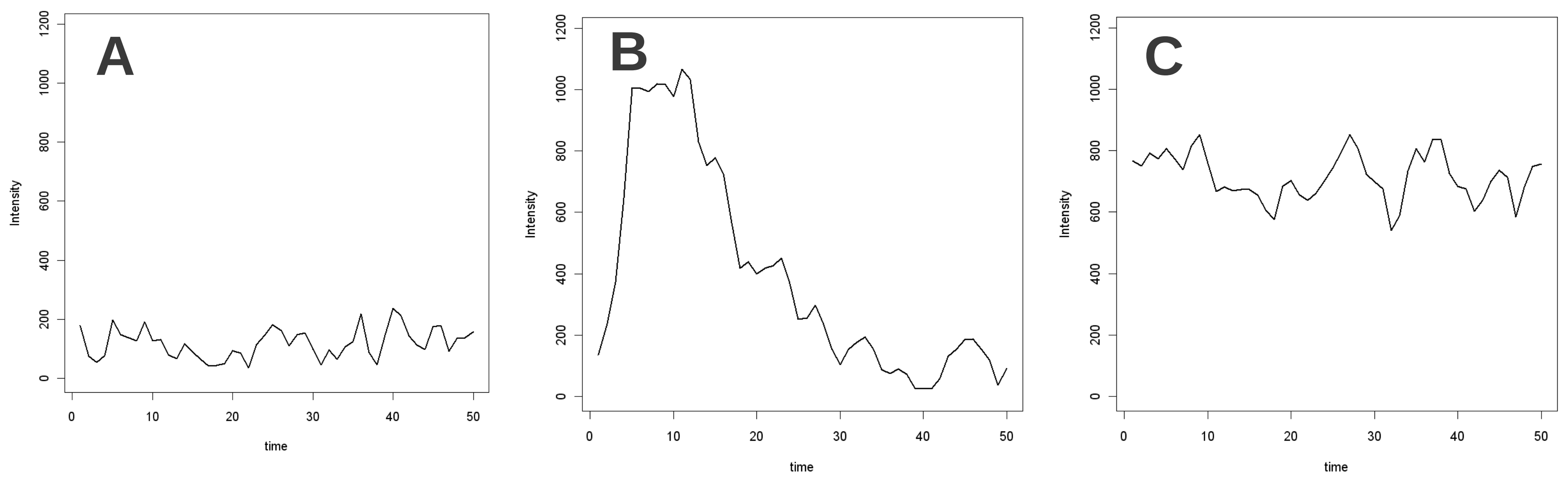
Statistical models need to incorporate some physics in order to spare data and represent high dimensional inputs-outputs relations.

→ putting more physics into statistics and vice versa is a promising research direction. Example through MRI image denoising.

Presentation of the data

We have 50 images of an artery obtained by MRI.





→ **The images are noisy**
→ **The 50 time steps correspond to one cardiac cycle**

Talk outline

Goal : smooth a sequence (time) of images (space)
Approach : statistical (kriging and POD)

In order to reduce the complexity of the space and time statistical description of our data (artery), two steps

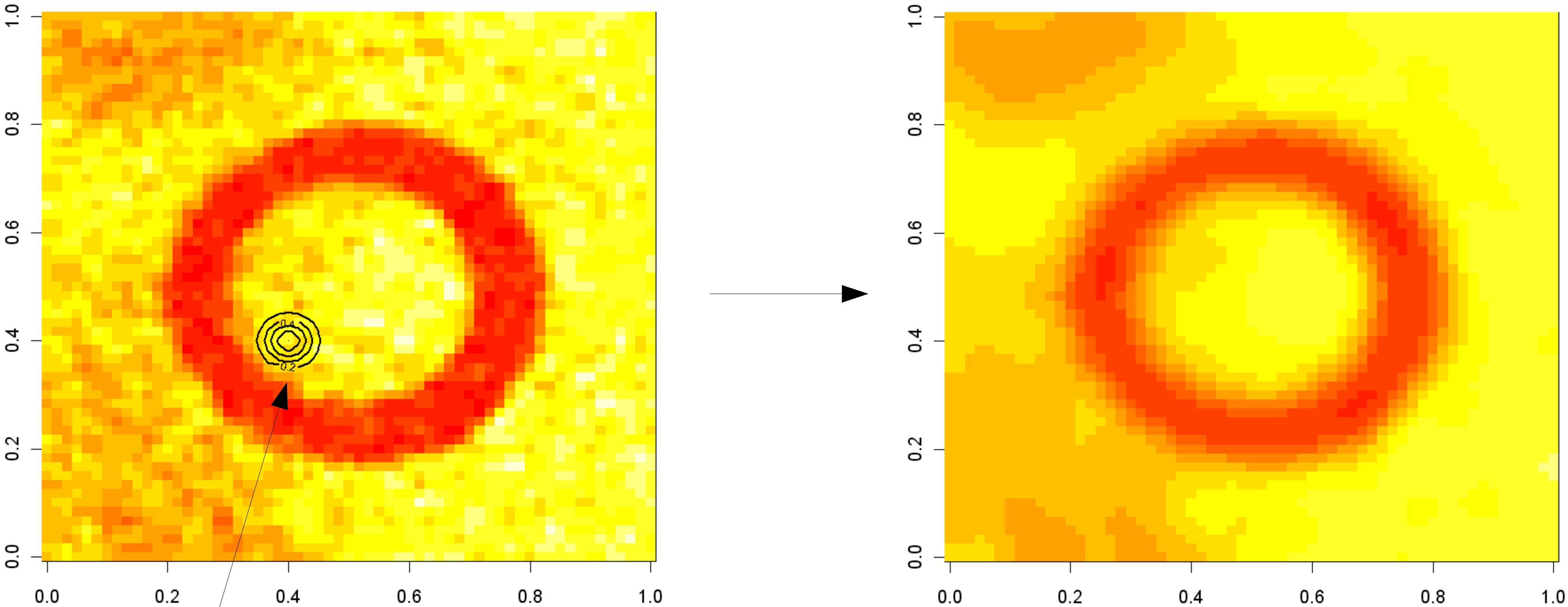
1. Spatial smoothing
2. Time smoothing

and

3. Conclusions

Spatial smoothing (1/6)

A basic approch is to smooth the data with a Gaussian Smoother
→ local average at each pixel



Contour lines of the Gaussian smoother @ px (25,25)

The nature of the data is not respected :

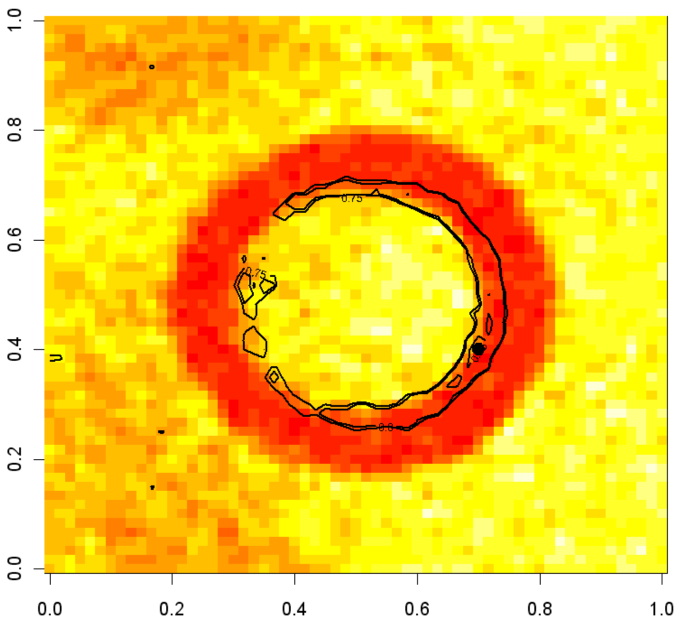
The smoother has to be adapted

Spatial smoothing (2/6)

The basic idea is to adapt the neighbourhood of each pixel to the problem at hand

→ we consider neighborhoods based on the empirical covariance matrix

The intensity at each pixel is seen as a random variable and each time step gives us a realization of this RV.



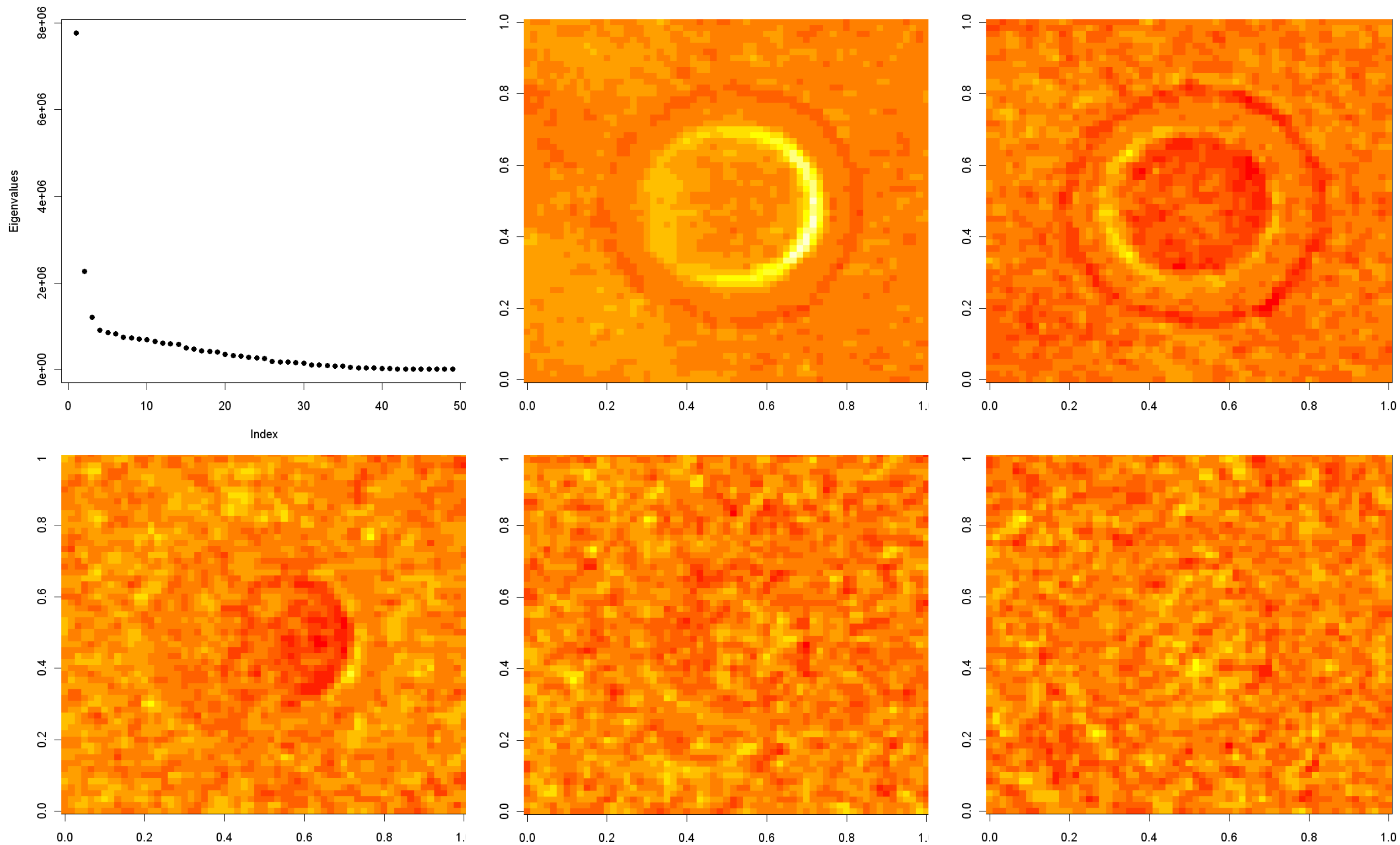
$$C = \frac{1}{N}\sum_{i=1}^{N}\left(X^{(i)} - \bar{X}\right)\left(X^{(i)} - \bar{X}\right)^{T}$$

Correlation between the pixel (43,25) – black dot – and the other pixels

This covariance matrix catches the physical partition of the space.

Spatial smoothing (3/6)

To denoise, we study the diagonalization $C = P\,D\,P^t$



The first three eigenvectors seem meaningful.
From the 4th on, we consider that they represent noise.

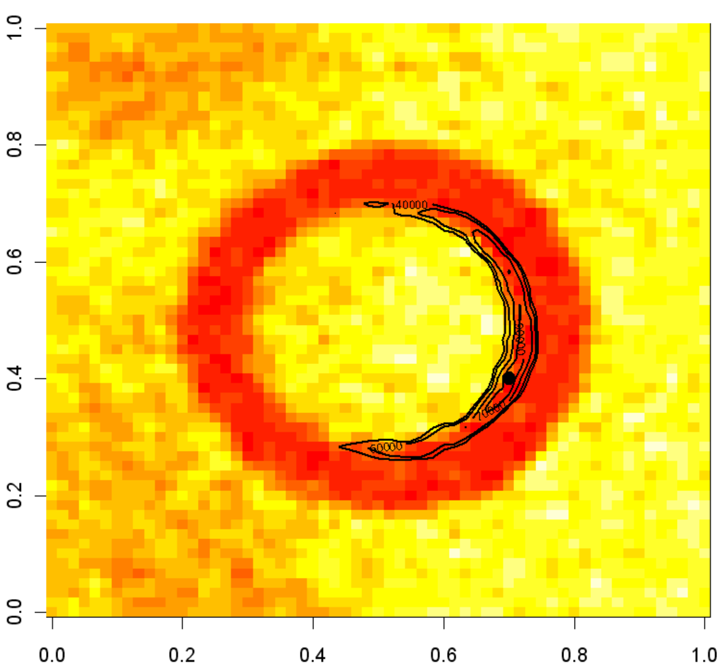
Spatial smoothing (4/6)

We thus split the covariance matrix in two groups

$$C = \begin{pmatrix} P_1 & \dots & P_p \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_p \end{pmatrix} \begin{pmatrix} P_1 & \dots & P_p \end{pmatrix}^T$$

$$= P\left(D_s + D_n\right)P^T = C_s + C_n$$

And we look at the neighborhood given by C_s



The use of covariance for smoothing is called *kriging*.

Here, as the covariance structure is learnt empirically, *kriging* is equivalent to *Proper Orthogonal Decomposition (POD)*.

Spatial smoothing (5/6)

The smoothed image at t is given by the kriging average:

$$I_s(x,t) = c_s(x)^T \left(C_s + C_n\right)^{-1} I(t)$$

where $c_s(x)$ is the covariance between the pixel x and the other pixels hence a line of C_s and the so-called neighborhood.

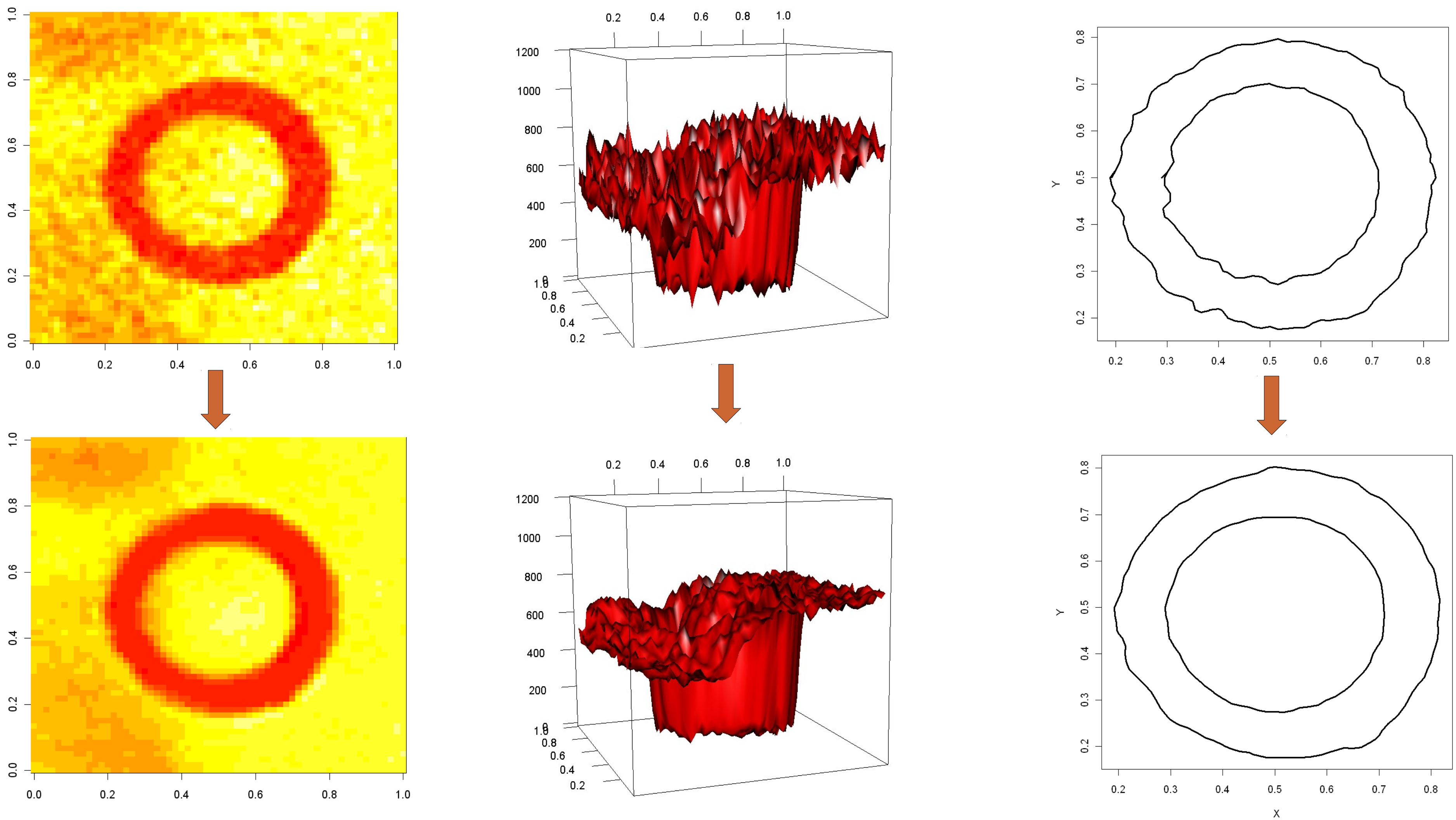
From kriging to POD :

$$I_s(t) = C_s \left(C_s + C_n\right)^{-1} I(t) = \left(P D_s P^T\right)\left(P D^{-1} P^T\right) I(t) = P D_s D^{-1} P^T I(t)$$

$$= P \begin{pmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix} P^T I(t)$$

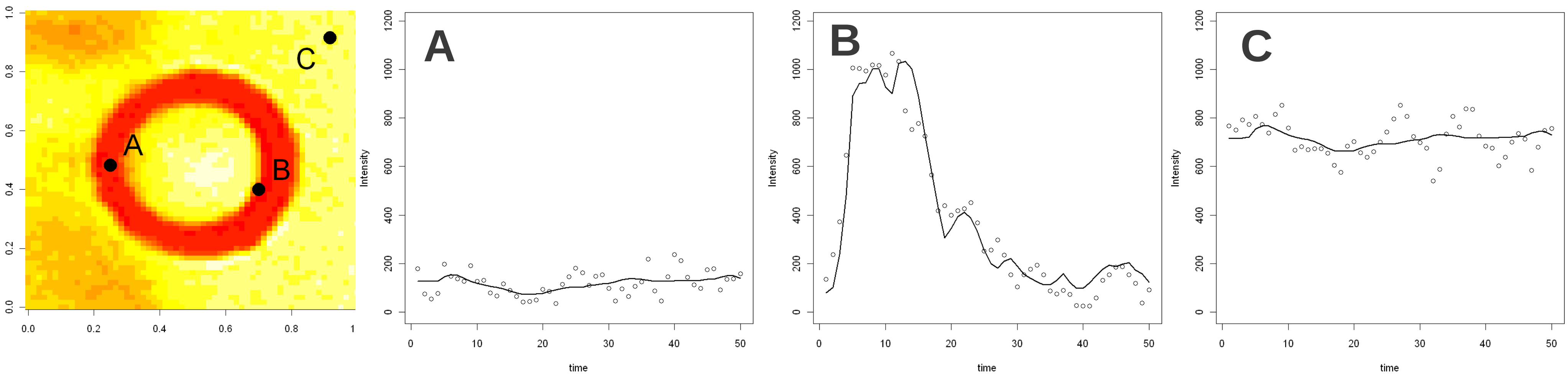
Spatial smoothing (6/6)

Result with effect on contour (right)



Time smoothing

Effect of spatial smoothing (previous slides) on time response :



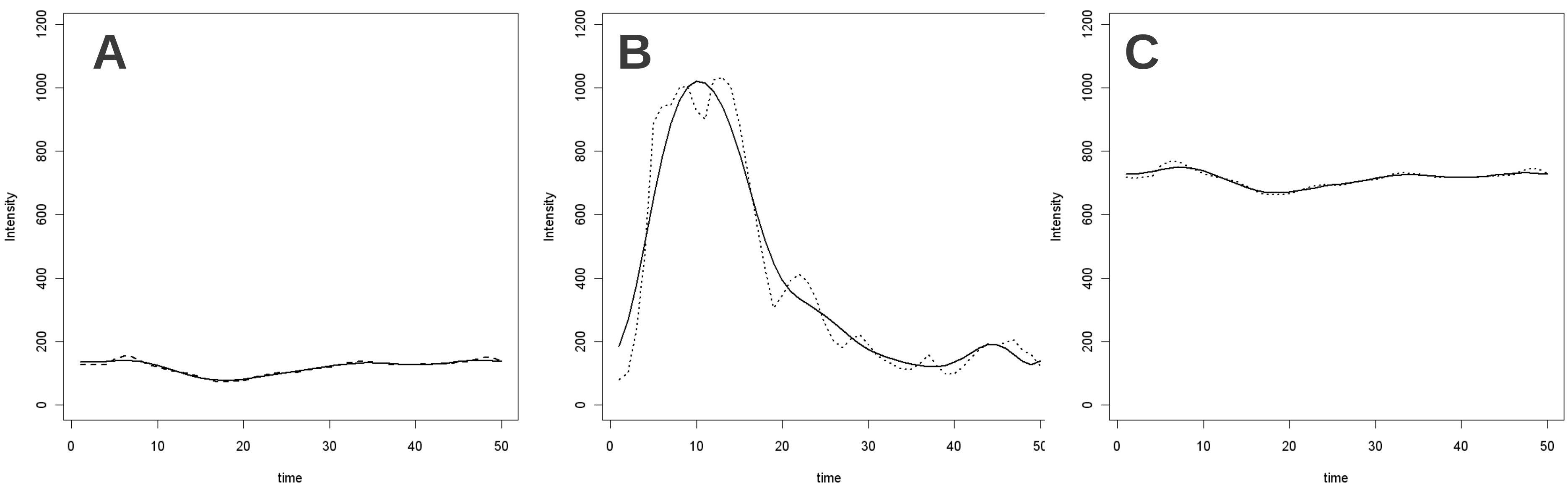
→ **The curves are much less noisy**

→ **The periodicity is not guaranteed (observed here because data is periodic)**

Time smoothing

→ **The artery movements are periodic: we apply a time smoothing (covariance in time) which is periodic.**

$$C_x(t_1, t_2) = Cov(I(x, t_1), I(x, t_2)) = \exp(-\sin^2(t_1 - t_2))$$



Both smoothings are krigings since they are based on covariance.

The shape of the functions I(x,t) after smoothing *has very little functional assumptions* (infinite number of basis functions in the RKHS). But there are assumptions in the covariance structure (i.e., the RKHS).

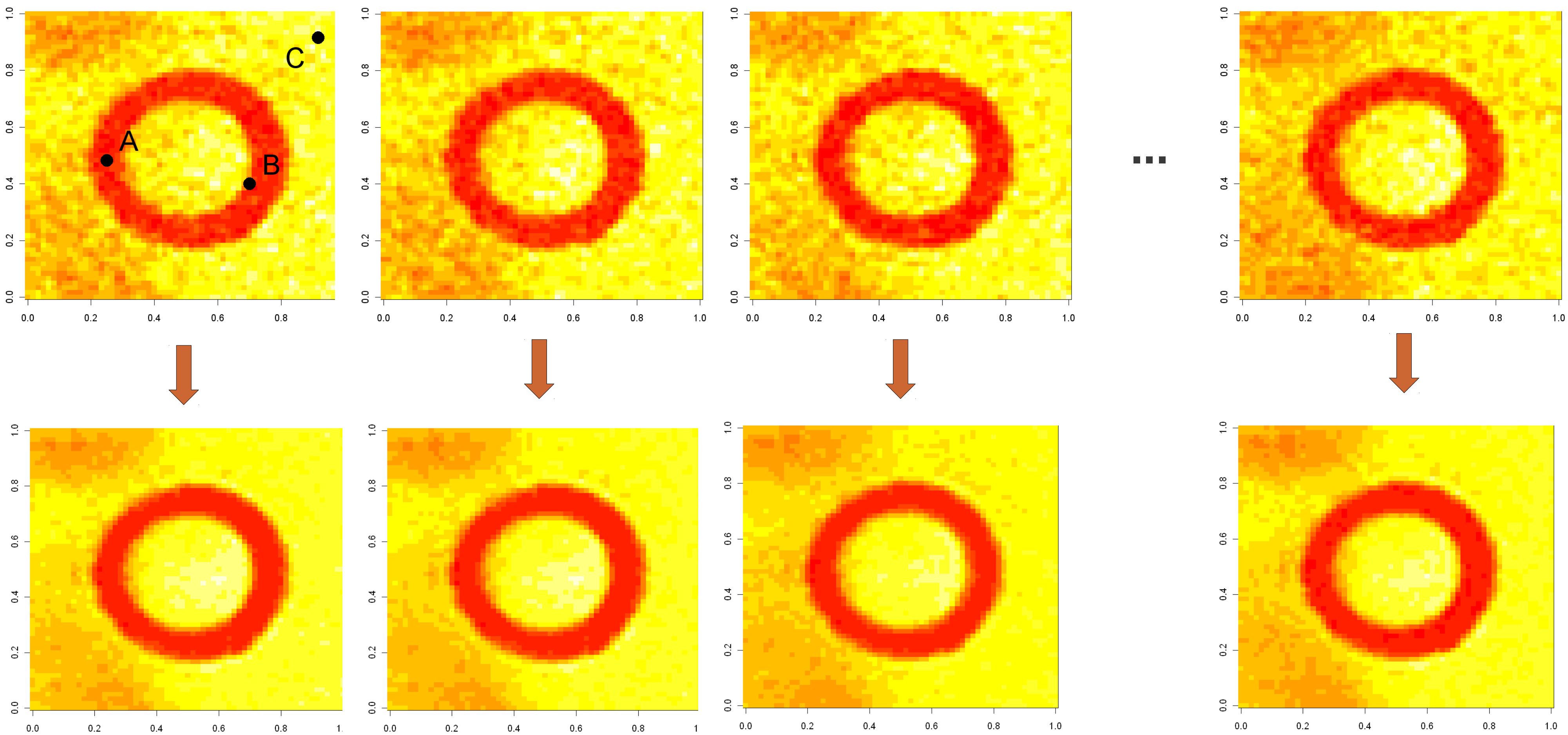
Final results

The sequence (space then time) allows to reduce the complexity

Space smoothing: inversion of a 3721 x 3721 matrix

Time smoothing: inversion of a 50 x 50 matrix

Time and space: inversion of a 186050 x 186050 matrix !!



Conclusions (1/2)

• Space and time have been handled in the following way

1. Space treatment, time taken as a random event
2. Time treatment, independently at each point in space

⇨ never construct the complete time-space covariance matrix (186050^2 here).

• Some physical features are accounted for : spatial neighborhoods based on data and a priori that time neighborhoods are periodical.

Conclusions (2/2)

- Such statistical model is well-adapted to noisy data. The effect of noise on physical models is difficult to control. Here, it is controlled through the choice of the covariances (with effects on likelihood, variances …).

- In other words, statistical models focus on data, physical models on physics. Physical models can predict response for new inputs (e.g., new material parameters), statistical models can simulate new data (with or without noise, for the same input).

- A promising research direction : mixing statistical and physical models. Examples:
  - Use statistical model to denoise (get contours …), then physical identification. This is a current practice.
  - More originally: define covariance based on PDE's.