A new class of Ansatz

M.Tabone, Y.Zhang, S.Mangiarotti

Centre d'Etude Spatiale de la BIOsphère
CESBIO-OMP UMR UT3-CNES-IRD-CNRS-INRAE

Environmental problematics

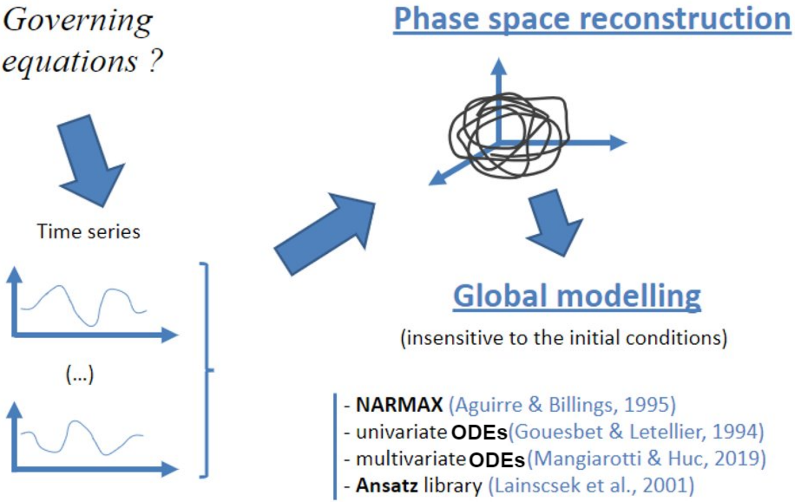
Eco-bio-physical system:

T, h, p, S, etc.

Dynamic system:

$$\dot{\vec{X}} = f(\vec{X})$$

Environmental modeling



Results obtained so far?

	uniODEs	**multiODEs**	**Ansatz**
Theoretical cases	Lorenz-1963 Rössler-1976 Gouesbet & Letellier 1994	Quadratic-cubic Systems (3D-5D), etc. Mangiarotti & Huc 2019	Lorenz-1963 Rössler-1976 Lainscsek et al. 2001, 2003 Malasoma & Boiron, 2003
Experimental systems	Electrodissolution Letellier et al. 1995 Mixing Reactor Letellier et al. 1997	-	-
Environmental observations	Lynx population Maquet et al. 2007 Cereal crops Mangiarotti et al. 2012 Karstic springs Mangiarotti et al. 2019	Bombay plague Mangiarotti 2015 Ebola in West Africa Mangiarotti 2016 Earthworms cycles Mangiarotti 2021	-

Global Modeling

- Original formulation

$(x_1, x_2, \ldots, x_n)$

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, \ldots, x_n) \\ \dot{x}_2 = f_2(x_1, x_2, \ldots, x_n) \\ \ldots \\ \dot{x}_n = f_n(x_1, x_2, \ldots, x_n) \end{cases}$$

- Univariate reformulation

$(x_1, \dot{x}_1, \ddot{x}_1, \ldots)$

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \ldots \\ \dot{X}_n = F(x_i, X_2, \ldots, X_n) \end{cases}$$

Sophus Lie

Lie derivatives + inversion

$$\mathcal{L}_f f_i(\boldsymbol{x}) = \frac{\partial f_i(\boldsymbol{x})}{\partial \boldsymbol{x}} \boldsymbol{f}(\boldsymbol{x}) = \sum_{k=1}^{m} \frac{\partial f_i(\boldsymbol{x})}{\partial \boldsymbol{x}} f_k$$

$F(x_1, x_2, \ldots, x_n)$ → $F(x_1, X_2, X_3, \ldots)$

Global Modeling

- Original formulation

$$(x_1, x_2, \dots, x_n)$$

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, \dots, x_n) \\ \dot{x}_2 = f_2(x_1, x_2, \dots, x_n) \\ \dots \\ \dot{x}_n = f_n(x_1, x_2, \dots, x_n) \end{cases}$$

- Univariate reformulation

$$(x_1, \dot{x}_1, \ddot{x}_1, \dots)$$

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \dots \\ \dot{X}_n = F(x_i, X_2, \dots, X_n) \end{cases}$$

NOT ALWAYS POSSIBLE

Observability problems
Letellier & Aguirre (2001)

L. Aguirre

C. Letellier

Global Modeling

- Original formulation

$(x_1, x_2, \dots, x_n)$

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, \dots, x_n) \\ \dot{x}_2 = f_2(x_1, x_2, \dots, x_n) \\ \dots \\ \dot{x}_n = f_n(x_1, x_2, \dots, x_n) \end{cases}$$

- Univariate reformulation

$(x_1, \dot{x}_1, \ddot{x}_1, \dots)$

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \dots \\ \dot{X}_n = F(x_i, X_2, \dots, X_n) \end{cases}$$

Global Modeling

- Original formulation

$(x_1, x_2, \ldots, x_n)$

$$\begin{cases} \dot{x}_1 = f_1(x_1, x_2, .., x_n) \\ \dot{x}_2 = f_2(x_1, x_2, .., x_n) \\ \ldots \\ \dot{x}_n = f_n(x_1, x_2, .., x_n) \end{cases}$$

- Univariate reformulation

$(x_1, \dot{x}_1, \ddot{x}_1, \ldots)$

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \ldots \\ \dot{X}_n = F(x_i, X_2, .., X_n) \end{cases}$$

Ansatz library: A library of identities between these two formulations

C.Lainscsek Ansatzs library

$$A_1 \equiv \begin{cases} \dot{x} = a_0 + a_1 x + \boxed{a_2 y} + a_4 x^2 \\ \dot{y} = b_0 + b_1 x + b_2 y + b_4 x^2 + b_5 xy + \boxed{b_6 xz} + b_7 y^2 \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

$$A_2 \equiv \begin{cases} \dot{x} = a_0 + a_1 x + \boxed{a_2 y} + a_4 x^2 \\ \dot{y} = b_0 + b_1 x + b_2 y + \boxed{b_3 z} + b_4 x^2 + b_5 xy + b_7 y^2 \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

$$A_3 \equiv \begin{cases} \dot{x} = a_0 + a_1 x + a_4 x^2 + \boxed{a_5 xy} \\ \dot{y} = b_0 + b_1 x + b_2 y + b_4 x^2 + b_5 xy + \boxed{b_6 xz} + b_7 y^2 \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

$$A_4 \equiv \begin{cases} \dot{x} = a_0 + a_1 x + a_4 x^2 + \boxed{a_5 xy} \\ \dot{y} = b_0 + b_1 x + b_2 y + \boxed{b_3 z} + b_4 x^2 + b_5 xy + b_7 y^2 \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

$$A_5 \equiv \begin{cases} \dot{x} = \boxed{a_2 y} \\ \dot{y} = b_0 + b_1 x + b_2 y + b_4 x^2 + b_5 xy + b_7 y^2 + \boxed{b_8 yz} \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

$$A_6 \equiv \begin{cases} \dot{x} = \boxed{a_5 xy} \\ \dot{y} = b_0 + b_1 x + b_2 y + b_4 x^2 + b_5 xy + b_7 y^2 + \boxed{b_8 yz} \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

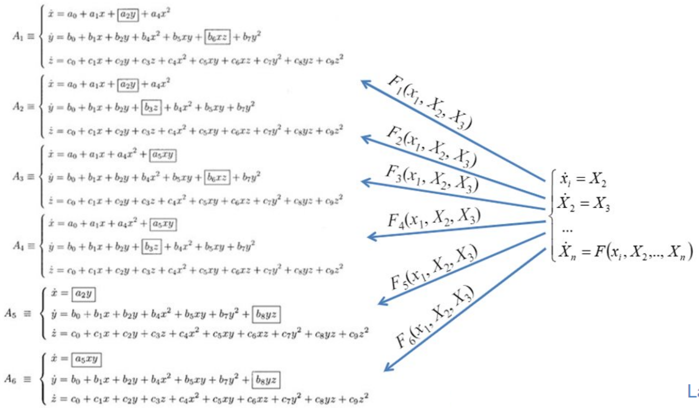
$F_1(x_1, X_2, X_3)$, $F_2(x_1, X_2, X_3)$, $F_3(x_1, X_2, X_3)$, $F_4(x_1, X_2, X_3)$, $F_5(x_1, X_2, X_3)$, $F_6(x_1, X_2, X_3)$ →

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \dots \\ \dot{X}_n = F(x_i, X_2, \dots, X_n) \end{cases}$$

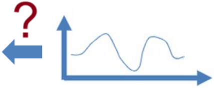
C. Lainscsek

Lainscsek et al. 2001, 2003

C.Lainscsek Ansatzs library



C. Lainscsek



Lainscsek et al. 2001, 2003

C.Lainscsek Ansatzs library

Ansatz identification is presently limited to:

- **3-dimensional dynamics**
- **original formulation of degree 2**

We decided to construct another library extended to **non autonomous equations**.

Construction of a new Ansatz library

Case study: famous non autonomous systems: the **Duffing system** and the **Van der Pol system** which are **cubic sytems**.
$\Longrightarrow$ Include two supplementary terms in the **general original formulation**:

$$(I) = \begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy + \alpha u \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy + \Delta x^3 + \Gamma x^2 y + \beta v \end{cases}$$

with u(t) or v(t) a known **external forcing of unknown dimension**.

Construction of a new Ansatz library

$$(I) = \begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy + \alpha u \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy + \Delta x^3 + \Gamma x^2 y + \beta v \end{cases}$$

Inversion problems
& Restriction to polynomial forms

- *Original structure if x is observed*:

$$(I1) = \begin{cases} \dot{x} = a_0 + a_1 x + y + a_3 x^2 + \alpha u \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy + \Delta x^3 \\ +\Gamma x^2 y + \beta v \end{cases}$$

- *Original structure if y is observed*:

$$(I2) = \begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy + \alpha u \\ \dot{y} = b_0 + x + b_2 y + b_4 y^2 + \beta v \end{cases}$$

Construction of a new Ansatz library

$$(I) = \begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy + \alpha u \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy + \Delta x^3 + \Gamma x^2 y + \beta v \end{cases}$$

- *Original structure if x is observed:*

$$(I1) = \begin{cases} \dot{x} = a_0 + a_1 x + y + a_3 x^2 + \alpha u \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 x^2 + b_4 y^2 + b_5 xy + \Delta x^3 \\ +\Gamma x^2 y + \beta v \end{cases}$$

- *Original structure if y is observed:*

$$(I2) = \begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy + \alpha u \\ \dot{y} = b_0 + x + b_2 y + b_4 y^2 + \beta v \end{cases}$$

The new library:

Observed variables: x, u

$$(A_1) \equiv \begin{cases} \dot{x} = X_2 \\ \dot{X}_2 = \theta_0 + ... + \theta_9 u + \theta_{10} u^2 \\ +\theta_{11} ux + \theta_{12} ux^2 + \theta_{13} uX_2 + \theta_{14} \dot{u} \end{cases}$$

Observed variables: x, v

$$(A_2) \equiv \begin{cases} \dot{x} = X_2 \\ \dot{X}_2 = \theta_0 + ... + \theta_{15} v \end{cases}$$

Observed variables: x, u=v

$$(A_3) \equiv \begin{cases} \dot{x} = X_2 \\ \dot{X}_2 = \theta_0 + ... + (\theta_9 + \theta_{15}) u + \\ \theta_{10} u^2 + \theta_{11} ux + \theta_{12} ux^2 + \theta_{13} uX_2 \\ +\theta_{14} \dot{u} \end{cases}$$

Observed variables: y, u

$$(A_4) \equiv \begin{cases} \dot{y} = Y_2 \\ \dot{Y}_2 = \theta_0 + ... + \theta_9 u \end{cases}$$

Observed variables: y, v

$$(A_5) \equiv \begin{cases} \dot{y} = Y_2 \\ \dot{Y}_2 = \theta_0 + .. + \theta_{10} \dot{v} + \theta_{11} v + \theta_{12} v^2 \\ +\theta_{13} yv + \theta_{14} y^2 v + \theta_{15} Y_2 v \end{cases}$$

Observed variables: y, u=v

$$(A_6) \equiv \begin{cases} \dot{y} = Y_2 \\ \dot{Y}_2 = \theta_0 + ... + \theta_{10} \dot{v} + (\theta_{11} + \theta_9) v \\ +\theta_{12} v^2 + \theta_{13} yv + \theta_{14} y^2 v + \theta_{15} Y_2 v \end{cases}$$

Two categories of coefficients

Example for A_2:

- $\theta_0 = b_4 a_0^2 + b_0 - a_0 b_2$
- $\theta_1 = b_1 - a_1 b_2 + 2b_4 a_0 a_1 - a_0 b_5$
- $\theta_2 = b_3 - a_3 b_2 + b_4 a_1^2 + 2b_4 a_0 a_3 - a_1 b_5 - \Gamma a_0$
- $\theta_3 = \Delta + 2b_4 a_1 a_3 - a_3 b_5 - a_1 \Gamma$
- $\theta_4 = b_4 a_3^2 - a_3 \Gamma$
- $\theta_5 = b_2 a_1 - 2b_4 a_0$
- $\theta_6 = b_4$
- $\theta_7 = -2b_4 a_1 + 2a_3 + b_5$
- $\theta_8 = \Gamma - 2b_4 a_3$
- $\theta_{15} = \beta$

Global coefficients θ and original coefficients a, b, α, β, Γ and Δ.

GPoM Package

The Comprehensive R Archive Network

- Sylvain Mangiarotti & Mireille Huc
- Global Modelling Tool
- Obtain equations close to the dynamics of the observed system

Retrieving the Original system

(I) Ansatz identification is limited to **low-dimensional dynamics**

(II) Global modelling (with **GPoM** algorithm) could be used to get low-dimensional approximaton of high-dimensional dynamics but:

Its identification strategy cannot be used straightfoward

$$\begin{cases} \dot{x}_i = X_2 \\ \dot{X}_2 = X_3 \\ \dots \\ \dot{X}_n = F(x_i, X_2, \dots, X_n) \end{cases}$$

GPoM

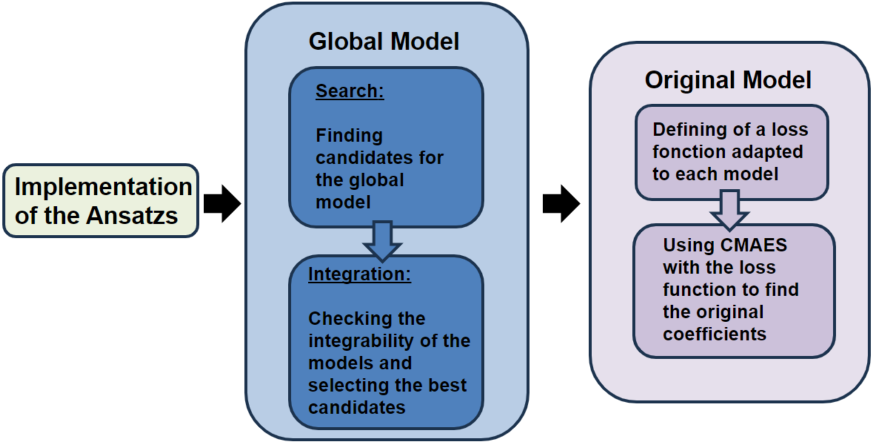
$$\sim P(x_1, X_2, X_3) = \theta_0 + \theta_1 x_1 + \theta_2 x_1 X_2 + \dots$$

?

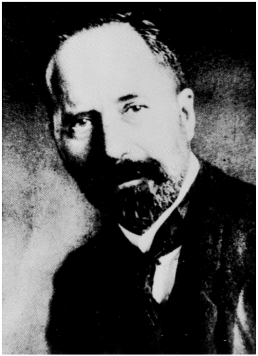
$$\begin{cases} \dot{x} = a_0 + a_1 x + a_2 y + a_3 z + a_4 x^2 + a_5 xy \\ \quad + a_6 xz + a_7 y^2 + a_8 yz + a_9 z^2 \\ \dot{y} = b_0 + b_1 x + b_2 y + b_3 z + b_4 x^2 + b_5 xy \\ \quad + b_6 xz + b_7 y^2 + b_8 yz + b_9 z^2 \\ \dot{z} = c_0 + c_1 x + c_2 y + c_3 z + c_4 x^2 + c_5 xy \\ \quad + c_6 xz + c_7 y^2 + c_8 yz + c_9 z^2 \end{cases}$$

Original coefficients (a_i, b_i, c_i) unavailable

Two step process



The case study

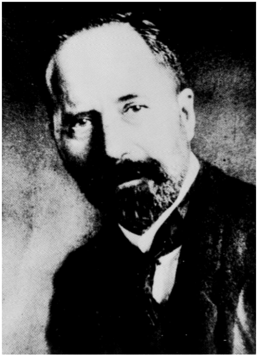


Georg Duffing
(1861-1944)

Duffing system:

$$\begin{cases} \dot{x} = y \\ \dot{y} = -x^3 - ay + u \\ \dot{u} = v \\ \dot{v} = -\omega^2 u \end{cases}$$

The case study



Georg Duffing
(1861-1944)

Modified Duffing system :

$$\begin{cases} \dot{x} = y + 8u \\ \dot{y} = -0.02y - x - 5x^3 \\ \dot{u} = v \\ \dot{v} = -0.25u \end{cases}$$

Search

- Candidates for the global model:
  Approximation of the θ_i:

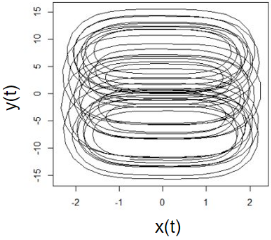
```
[1] 1
dX1/dt = 1 X2


dX2/dt = -0.00204296  + 7.95036759 X3  -0.01973319 X2  + 0.002279
71 X2^2  -1.182612 X1  + 0.00018522 X1 X2  -0.0024935 X1^2  -0.00
029093 X1^2 X2  -4.83310754 X1^3  + 4.92e-06 X1^4


dX3/dt = +


dX4/dt = +


[1] 2
dX1/dt = 1 X2


dX2/dt = 7.950436 X3  -0.01975 X2  + 0.002139 X2^2  -1.182801 X1
+ 0.000186 X1 X2  -0.005197 X1^2  -0.00028 X1^2 X2  -4.833024 X1^
3  + 0.000861 X1^4


dX3/dt = +


dX4/dt = +
```

- Outputs:

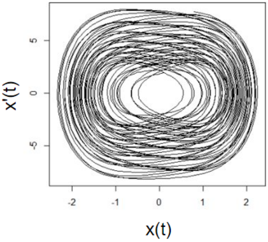
SuperMat	list [37]
SuperStruct	list [2022]
SuperAns	integer [2022]
SupernbFinal	double [37]

We obtained 37 candidates for the global model.

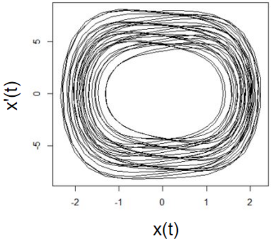
Phase portraits



(a) Modified Duffing system in (x, y) projection

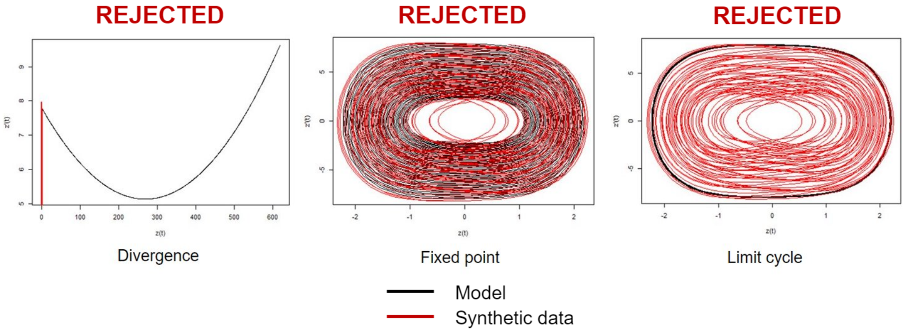


(b) Modified Duffing system in $(x, \dot{x})$ projection



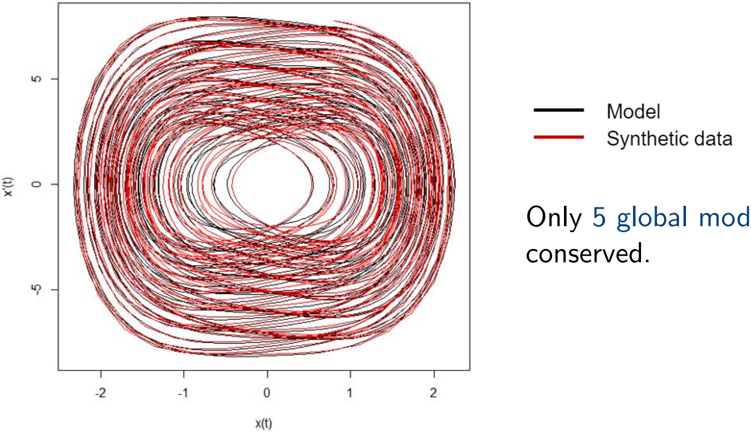
(c) Equivalent Ansatz of the modified Duffing in $(x, \dot{x})$ proj.

Numerical Integration



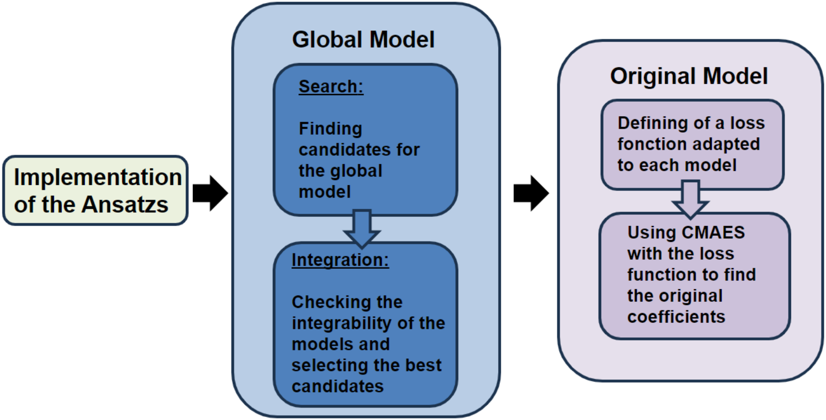
Numerical Integration

CONSERVED



Only 5 global models are conserved.

Two step process



Reminder: Two categories of coefficients

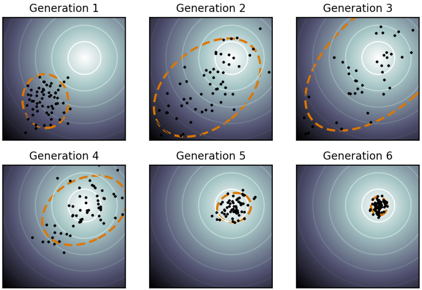
Example for A_2:

- $\theta_0 = b_4 a_0^2 + b_0 - a_0 b_2$
- $\theta_1 = b_1 - a_1 b_2 + 2b_4 a_0 a_1 - a_0 b_5$
- $\theta_2 = b_3 - a_3 b_2 + b_4 a_1^2 + 2b_4 a_0 a_3 - a_1 b_5 - \Gamma a_0$
- $\theta_3 = \Delta + 2b_4 a_1 a_3 - a_3 b_5 - a_1 \Gamma$
- $\theta_4 = b_4 a_3^2 - a_3 \Gamma$
- $\theta_5 = b_2 a_1 - 2b_4 a_0$
- $\theta_6 = b_4$
- $\theta_7 = -2b_4 a_1 + 2a_3 + b_5$
- $\theta_8 = \Gamma - 2b_4 a_3$
- $\theta_{15} = \beta$

Global coefficients θ and original coefficients a, b, α, β, Γ and Δ.

Covariance Matrix Adaptation - Evolution Strategy(CMAES)

- Stochastic optimization algorithm
- Developed Heike Trautmann, Olaf Mersmann and David Arnu
- Available on the CRAN
- Used to approximate the original coefficients



Loss fonction: Mean Square Error

The definition of the Mean square error is:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2$$

with:

- n the size of the sample
- y the variable we want to approximate
- f the function to test
- x the explanatory variables

Validation of the approximation of the original model

```
[1] "Nous en sommes à la  241 ème sous structure"
[1] "l'ansatz utilisé est le 2"
[1] "Les variables qui entrent en compte sont"
 [1] "ct"        "X4 "       "X3 "       "X3^2 "     "X2 "       "X2 X3 "    "X2^2 "
 [8] "X1 "       "X1 X3 "    "X1 X2 "    "X1^2 "     "X1^2 X3 "  "X1^2 X2 "  "X1^3 "
[15] "X1^4 "
> ValLoss2[[241]]
[1] 0.0001124674
> struct[[241]]
 [1] 1 1 0 1 0 0 1 0 1 1 0 1 1 1 0 0
> ValLoss3[[241]]
[1] 0.0001124642
> BestPar3[[241]]
[1]  2.582894e-02 -9.068510e-04 -3.453248e-04  7.674675e+00 -1.740495e+00
[6] -1.935096e-02  1.771724e-05  7.726695e-04 -4.614817e+00
> inMod
[1] 17
> Maxit
[1] 1000
>
```

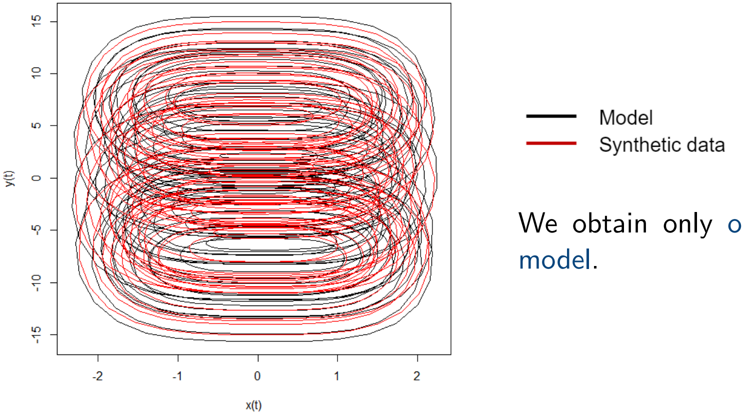
Original Model

$$
\begin{cases}
\dot{x} = y + 8u \\
\dot{y} = -0.02y - x - 5x^3 \\
\dot{u} = v \\
\dot{v} = -0.25u
\end{cases}
$$

Approximation of the Original Model

```
dx/dt = 0.0258289 + 7.6746747 u + 1 y -0.0009069 x  -0.0003453 x^2

dy/dt = -0.019351 y + 1.77e-05 y^2 -1.7404947 x + 0.0007727 x y -4.6148173 x^3
```

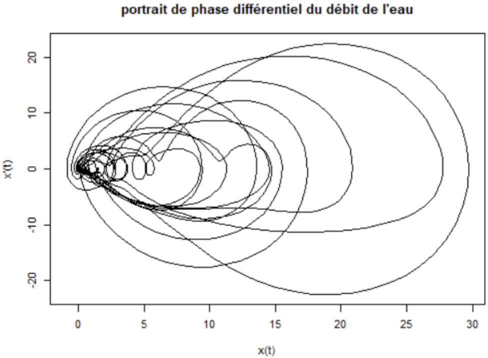
Validation of the approximation of the original model



We obtain only one original model.

Perspectives

- Complete the library with some **other structures**
- Use this method on **observed data**



Thank you!