

# The YUIMA Project: a Computational Framework for Simulation and Inference of Stochastic Differential Equations

Stefano M. Iacus

Department of Economics, Management and Quantitative Methods  
University of Milan, Italy

## Abstract

The Yuima Project is an open source and collaborative effort aimed at developing the R package named `yuima` for simulation and inference of stochastic differential equations. In the `yuima` package, stochastic differential equations can be of very abstract type, multidimensional, driven by Wiener process or fractional Brownian motion with general Hurst parameter, with or without jumps specified as Lévy noise.

The `yuima` package is intended to offer the basic infrastructure on which complex models and inference procedures can be built on. The computational framework implemented allow for the estimation of high frequency data and also offer the ability to perform Monte Carlo analysis using cluster infrastructure whenever available in a transparent way to the user.

Some real examples of model implementation and data estimation will be considered.

**Key Words** Stochastic differential equations, high frequency data, inference for stochastic processes

## 1 Introduction

The plan of the YUIMA Project is to construct the bases for a complete environment for simulation and inference for stochastic processes via an R package called `yuima`.

The `yuima` package supplies various functions to carry out simulation and statistical analysis. Theory of statistical inference for stochastic differential equations already has a fairly long history, more than three decades, but it

is still developing quickly new methodologies and expanding the area. The `yuima` package delivers up-to-date methods as a package onto the desk of the user working with simulation and/or statistics for stochastic differential equations. In the `yuima` package, stochastic differential equations can be of very abstract type, multidimensional, driven by Wiener process or fractional Brownian motion with general Hurst parameter, with or without jumps specified as Lévy noise.

### The YUIMA Project Team

The YUIMA Project team is composed by the following researchers: N. Yoshida, Y. Koike and R. Nomura, University of Tokyo (Japan); A. Brouste, University of Le Mans (France); M. Fukasawa, T. Ogihara, Y. Shimuzu, K. Kamatani and M. Uchida, Osaka University (Japan), Masaaki Fukasawa Osaka University; S.M. Iacus, University of Milan (Italy).

### The `yuima` Package

The `yuima` package is hosted on R-Forge and can be installed from R-Forge using `install.packages("yuima", repos="http://R-Forge.R-project.org")`. If, the R-Forge system does not provide binary builds of the `yuima` package, the user can also try `install.packages("yuima", repos="http://R-Forge.R-project.org", type="source")`.

### The Main Object and Classes

There are different classes of objects defined in the `yuima` package and the main class is called the `yuima-class`. Figure 1 represents the different classes and their slots.

The general idea of the `yuima` package is to separate into different sub-class objects the statistical model, the data and the statistical methods. The mathematical description of the statistical model is specified using `setModel` which prepares a `yuima.model` object by filling the appropriate slots. If the aim is the simulation of the solution of the stochastic differential equation specified in the `yuima.model` object then, using the method `simulate`, it is possible to obtain one trajectory of the process. As an output, a `yuima` object is created which contains the original model specified in the `yuima.model` object in the slot named `model` and two additional slots named `data`, for the simulated data, and `sampling` which contains the description of the simulation scheme used as well as other informations. The method `setSampling` allows to specify the type of sampling scheme to be

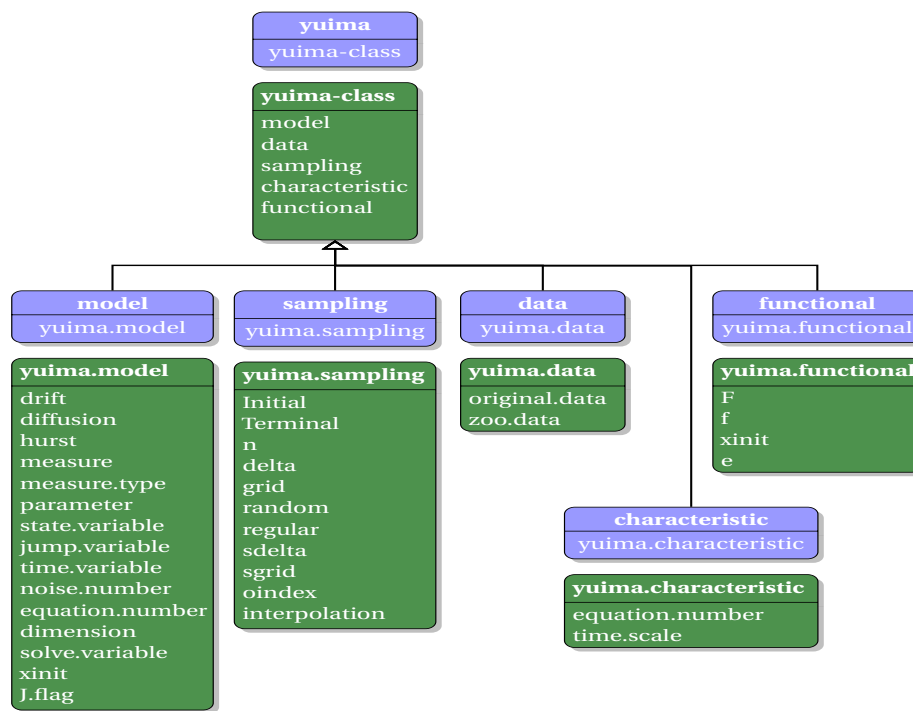


Figure 1: The main classes in the `yuima` package.

used by the `simulate` method. But `yuima` object may contain the slot `data` non only as the outcome of `simulate` but also because the user decides to analyse its own data. In this case the method `setData` is used to transform most types of R time series object into a a proper `yuima.data` object. When the slots `data` and `model` are available, many other methods can be used to perform statistical analysis on these SDE's (stochastic differential equation) models. Further, functionals of stochastic differential equations can be defined using the `setFunctional` method and evaluated using asymptotic expansion methods.

### The `yuima.model` Class

At present, in `yuima` three main classes of stochastic differential equations can be easily specified. In `yuima` one can describe three main families of stochastic processes at present. These models can be one or multidimensional and eventually described as parametric models. Let  $X_0 = x_0$  be the initial value of the process, then, the main classes are as follows: *i*) **diffu-**

**sion models** described as  $dX_t = a(t, X_t, \theta)dt + b(t, X_t, \theta)dW_t$ , where  $W_t$  is a standard Brownian motion; *ii*) **SDE's driven by fractional Gaussian noise**, with  $H$  the Hurst parameter, described as  $dX_t = a(t, X_t, \theta)dt + b(t, X_t, \theta)dW_t^H$ ; *iii*) **diffusion process with jumps and Lévy processes** solution to

$$dX_t = a(t, X_t, \theta)dt + b(t, X_t, \theta)dW_t + \int_{|z|>1} c(X_{t-}, z)\mu(dt, dz) + \int_{0<|z|\leq 1} c(X_{t-}, z)\{\mu(dt, dz) - \nu(dz)dt\}$$

The functions  $a(\cdot)$ ,  $b(\cdot)$  and  $c(\cdot)$  may have a different number of arguments. The `yuima` package is able to accept any user-specified notation for the state variable  $x$  (for  $X_t$ ) and the time variable  $t$ .

### One Dimensional Diffusion Processes

Assume that we want to describe the following stochastic differential equation  $dX_t = -3X_tdt + \frac{1}{1+X_t^2}dW_t$ . This model can be described in `yuima` by specifying the drift and diffusion coefficients as plain R expressions passed as strings `mod1<-setModel(drift="-3*x",diffusion="1/(1+x^2)")`. Now, with `mod1` in hands, it is extremely easy to simulate a trajectory by Euler-Maruyama scheme of the process with `simulate(mod1)`. The `simulate` function fills in addition the two slots `data` and `sampling` of the `yuima` object.

### Specification of Parametric Models

Assume now that we want to specify a parametric model like this  $dX_t = -\theta X_tdt + \frac{1}{1+X_t^\gamma}dW_t$ , where  $a(x, \theta) = -\theta x$  and  $b(x, \gamma) = 1/(1 + x^\gamma)$ . Then, `yuima` attempts to distinguish the parameters' names from the ones of the state and time variables `mod2<-setModel(drift="-theta*x",diffusion="1/(1+x^gamma)")`. so, in this case, `theta` and `gamma`, which are different from `x` and `t`, are assumed to be parameters. In order to simulate the parametric model it is necessary to specify the values of the parameters  $\theta$  and  $\gamma$  in this way `simulate(mod2,true.param=list(theta=1,gamma=3))`.

## Multidimensional Processes

Next is an example of a system of two stochastic differential equations for the couple  $(X_{1,t}, X_{2,t})$  driven by three independent Brownian motions  $(W_{1,t}, W_{2,t}, W_{3,t})$

$$\begin{aligned} dX_{1,t} &= -3X_{1,t}dt + dW_{1,t} + X_{2,t}dW_{3,t} \\ dX_{2,t} &= -(X_{1,t} + 2X_{2,t})dt + X_{1,t}dW_{1,t} + 3dW_{2,t} \end{aligned}$$

but this has to be organized into matrix form with a vector of drift expressions and a diffusion matrix. For this system it is now necessary to instruct `yuima` about the state variable on both the left-hand side of the equation and the right-hand side of the equation, i.e. there is the need to specify also the `solve.variable` for the left hand side of the SDE:

```
sol <- c("x1","x2") # variable for numerical solution
a <- c("-3*x1","-x1-2*x2") # drift vector
b <- matrix(c("1","x1","0","3","x2","0"),2,3)
mod3 <- setModel(drift = a, diffusion = b, solve.variable = sol)
```

### 1.1 Fractional Gaussian Noise

The `yuima` allows for the description of stochastic differential equations driven by fractional Brownian motion of the following type  $dX_t = a(X_t)dt + b(X_t)dW_t^H$ , where  $W_t^H$  is a normalized fractional Brownian motion (fBM) with Hurst exponent  $H \in (0, 1)$ . If we want to specify the following model  $dY_t = 3Y_tdt + dW_t^H$ , with  $H = 0.3$ , we proceed as follows `setModel(drift="3*y",diffusion=1,hurst=0.3,solve.var="y")`.

## Lévy Processes

Jump processes can be specified in different ways in mathematics and hence in `yuima` package. Here we briefly present one possibility but the package allows for more general specifications. A simple Ornstein-Uhlenbeck process with Gaussian jumps like this  $dX_t = -\theta X_tdt + \sigma dW_t + dZ_t$  is then specified as follows

```
mod5 <- setModel(drift="-theta*x", diffusion="sigma",
+ jump.coef="1", measure=list(intensity="10",
+ df=list("dnorm(z, 0, 1)")),
+ measure.type="CP", solve.variable="x")
```

Another possibility is to specify the jump component via the Lévy measure  $\nu(\cdot)$ . Without going into too much details, here is an example of specification of a simple Ornstein-Uhlenbeck process with IG (Inverse Gaussian) Lévy measure  $\nu(\cdot)$  and no Poisson component  $dX_t = -xdt + dZ_t$ :

```
mod6 <- setModel(drift="-x", xinit=1, jump.coeff="1",
+   measure.type="code", measure=list(df="rIG(z, 1, 0.1)"))
```

## 1.2 Inference for Stochastic Differential Equations

The `yuima` package allows for the following statistical techniques: Quasi-Maximum Likelihood and Adaptive Bayes Estimation; Asynchronous Covariance Estimation; Change Point Analysis; LASSO-type estimation and model selection.

## 1.3 Other features

The package also allows for automated asymptotic expansion of functionals of diffusion processes with direct applications to asian option pricing for generic models. Random sampling and different simulation algorithm are included for jump processes, fractional Gaussian noise but also diffusion processes.

## Acknowledgements

This work was in part supported by Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research No. 24340015 (Scientific Research), No. 24650148 (Challenging Exploratory Research); the Global COE program “The Research and Training Center for New Development in Mathematics” of the Graduate School of Mathematical Sciences, University of Tokyo; by a Cooperative Research Program of the Institute of Statistical Mathematics and by the project PRIN 2009JW2STY, Ministero dell’Istruzione dell’Università e della Ricerca.

Project YUIMA I was in part supported by JST Basic Research Programs PRESTO. Project YUIMA II is in part supported by NS Solutions Corporation. NS Solutions Corporation and The Graduate School of Mathematical Sciences, University of Tokyo are conducting a joint study (Azzurro project) on “The Application of Advanced Mathematical Statistics Theories in the Financial Industry”.