anySim: Stochastic simulation of processes with any marginal distribution and correlation structure

Tsoukalas I. and Kossieris P.

Department of Water Resources and Environmental Engineering, School of Civil Engineering, National Technical University of Athens, Heroon Polytechneiou 5, 15780 Zographou, Greece

Correspondence: itsoukal@mail.ntua.gr; pkossier@central.ntua.gr

Abstract

`anySim` is an R package (Tsoukalas and Kossieris, 2019) for the stochastic simulation of processes with any marginal distribution and dependence structure. Currently, the package provides models for the simulation of univariate stationary and cyclostationary processes, exhibiting continuous, discrete and mixed-type marginal distributions as well as any valid (i.e., positive definite) short-range or long-range autocorrelation structure (parsimoniously parameterized using theoretical correlation structures; for more details see the references below). Furthermore, it implements a multivariate stationary stochastic model with similar capabilities, preserving also the lag-0 cross-correlation coefficients among the processes. The package can be used for the generation of synthetic time series (e.g., rainfall, runoff, temperature, wind speed, etc.) with the desired marginal and stochastic properties.

Implementation details: Methods & Models

The methodology is based on the concept of Nataf's joint distribution model (Liu and Der Kiureghian, 1986; Nataf, 1962) according to which the joint distribution of random variables with any target arbitrary marginal distributions can be obtained by mapping an auxiliary multivariate standard Gaussian distribution via the inverse cumulative distribution functions (ICDFs). This enables to exploit the link that exists between correlation coefficients in the Gaussian and the target domain, reproducing also the target correlations.

Moving to stochastic process simulation, `anySim` employs a similar concept (for more details see, Kossieris et al., 2019; Tsoukalas, 2019; Tsoukalas et al., 2019, 2018b, 2018a, 2018c, 2017; and references therein) that is based on the mapping (through the ICDF) of an auxiliary Gaussian process (Gp) through the ICDF in order to establish processes with the target marginal distribution and correlation structure. The package comprises the following stochastic simulation models:

- Autoregressive To Anything model of order p - ARTA(p): This model uses an appropriately parameterised univariate AR(p) to simulate an auxiliary Gp to establish the target correlation structure. It is noted that a similar, yet lower-order (i.e., with $p = 1$), model has been proposed by Cario and Nelson (1996).
- Sum of n Autoregressive models of order 1 - nARTA(1): This model uses the sum of n, appropriately parameterised, univariate AR(1) models to simulate an auxiliary Gp to establish the target correlation structure.

- Symmetric Moving Average (neaRly) To Anything - SMARTA(q): This model uses an appropriately parameterised SMA(q) model to simulate an auxiliary Gp to establish the target correlation structure. In the final step, the Gp realisation is mapped to the actual domain through the ICDF of the target distribution (Tsoukalas, 2019; Tsoukalas et al., 2019, 2018b).
- Stochastic Periodic Autoregressive To Anything model of order 1 - SPARTA: This model uses an appropriately parameterised univariate PAR(1) model to simulate a cyclostationary auxiliary Gp to establish the target season-to-season correlation structure (Tsoukalas, 2019; Tsoukalas et al., 2019, 2018a, 2017).

References

Cario, M.C., Nelson, B.L., 1996. Autoregressive to anything: Time-series input processes for simulation. Oper. Res. Lett. 19, 51–58. https://doi.org/10.1016/0167-6377(96)00017-X

Kossieris, P., Tsoukalas, I., Makropoulos, C., Savic, D., 2019. Simulating Marginal and Dependence Behaviour of Water Demand Processes at Any Fine Time Scale. Water 11, 885. https://doi.org/10.3390/w11050885

Liu, P.L., Der Kiureghian, A., 1986. Multivariate distribution models with prescribed marginals and covariances. Probabilistic Eng. Mech. 1, 105–112. https://doi.org/10.1016/0266-8920(86)90033-0

Nataf, A., 1962. Statistique mathematique-determination des distributions de probabilites dont les marges sont donnees. C. R. Acad. Sci. Paris 255, 42–43.

Tsoukalas, I., 2019. Modelling and simulation of non-Gaussian stochastic processes for optimization of water-systems under uncertainty. PhD Thesis, Department of Water Resources and Environmental Engineering, National Technical University of Athens (Defence date: 20 December 2018).

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2019. Building a puzzle to solve a riddle: A multi-scale disaggregation approach for multivariate stochastic processes with any marginal distribution and correlation structure. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2019.05.017

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2018a. Stochastic Periodic Autoregressive to Anything (SPARTA): Modeling and simulation of cyclostationary processes with arbitrary marginal distributions. Water Resour. Res. 54, 161–185. https://doi.org/10.1002/2017WR021394

Tsoukalas, I., Efstratiadis, A., Makropoulos, C., 2017. Stochastic simulation of periodic processes with arbitrary marginal distributions, in: 15th International Conference on Environmental Science and Technology. CEST 2017. Rhodes, Greece.

Tsoukalas, I., Kossieris, P., 2019. anySim: Stochastic simulation of processes with any marginal distribution and correlation structure. R package.

Tsoukalas, I., Makropoulos, C., Koutsoyiannis, D., 2018b. Simulation of stochastic processes exhibiting any-range dependence and arbitrary marginal distributions. Water Resour. Res. https://doi.org/10.1029/2017WR022462

Tsoukalas, I., Papalexiou, S., Efstratiadis, A., Makropoulos, C., 2018c. A Cautionary Note on the Reproduction of Dependencies through Linear Stochastic Models with Non-Gaussian White Noise. Water 10, 771. https://doi.org/10.3390/w10060771