

Python package for model STructure ANalysis (pySTAN)

S. Van Hoey, J. van der Kwast, I. Nopens, P. Seuntjens

11/06/2013

Abstract

The selection and identification of a suitable hydrological model structure is more than fitting parameters of a model structure to reproduce a measured hydrograph. The procedure is highly dependent on various criteria, i.e. the modelling objective, the characteristics and the scale of the system under investigation as well as the available data. Rigorous analysis of the candidate model structures is needed to support and objectify the selection of the most appropriate structure for a specific case (or eventually justify the use of a proposed ensemble of structures). This holds both in the situation of choosing between a limited set of different structures as well as in the framework of flexible model structures with interchangeable components.

Many different methods to evaluate and analyse model structures exist. This leads to a sprawl of available methods, all characterized by different assumptions, changing conditions of application and various code implementations. Methods typically focus on optimization, sensitivity analysis or uncertainty analysis, with backgrounds from optimization, machine-learning or statistics amongst others. These methods also need an evaluation metric (objective function) to compare the model outcome with some observed data.

However, for current methods described in literature, implementations are not always transparent and reproducible (if available at all). No standard procedures exist to share code and the popularity (and amount of applications) of the methods is sometimes more dependent on the availability than the merits of the method. Moreover, new implementations of existing methods are difficult to verify and the different theoretical backgrounds make it difficult for environmental scientists to decide about the usefulness of a specific method. A common and open framework with a large set of methods can support users in deciding about the most appropriate method. Hence, it enables to simultaneously apply and compare different methods on a fair basis.

We developed and present pySTAN (python framework for STructure Analysis), a python package containing a set of functions for model structure evaluation to provide the analysis of (hydrological) model structures. A selected set of algorithms for optimization, uncertainty and sensitivity analysis is currently available, together with a set of evaluation (objective) functions and input distributions to sample from. The methods are implemented model-independent and the python language provides the wrapper functions to apply administer external model codes. Different objective functions can be considered simultaneously with both statistical metrics and more hydrology specific metrics.

By using so-called *reStructuredText* (*sphinx documentation generator*) and *Python documentation strings* (*docstrings*), the generation of manual pages is semi-automated and a specific environment is available to enhance both the readability and transparency of the code. It thereby enables a larger group of users to apply and compare these methods and to extend the functionalities.

Keywords: python, model analysis, open-source

Postal Address:

ir. Stijn Van Hoey, BIOMATH, Department of Applied Mathematics, Biometrics and Process Control, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

dr. Johannes van der Kwast, UNESCO-IHE Institute for Water Education, 2601 Delft, The Netherlands

Prof. dr. ir. Ingmar Nopens, BIOMATH, Department of Mathematical Modelling, Statistics and Bioinformatics, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

Prof. dr. ir. Piet Seuntjens, VITO, Boeretang 200, 2400 Mol, Belgium

Email:

stvhoey.vanhoey@ugent.be

J.vanderKwast@unesco-ihe.org

Ingmar.Nopens@biomath.ugent.be

piet.seuntjens@vito.be

Phone +32 (0)9 264 5935