HAL
archives-ouvertes.fr

OpenTURNS, an Open Source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach

Anne Dutfoy, Ivan Dutka-Malen, Alberto Pasanisi, Régis Lebrun, Fabien Mangeant, Jayant Sen Gupta, Maurice Pendola, Thierry Yalamas

▶ **To cite this version:**

Anne Dutfoy, Ivan Dutka-Malen, Alberto Pasanisi, Régis Lebrun, Fabien Mangeant, et al.. OpenTURNS, an Open Source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach. 41èmes Journées de Statistique, SFdS, Bordeaux, 2009, Bordeaux, France, France. inria-00386719

HAL Id: inria-00386719
https://hal.inria.fr/inria-00386719

Submitted on 22 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

OpenTURNS, an Open Source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach

A. Dutfoy, I. Dutka-Malen, A. Pasanisi
EDF R&D, Clamart, France

R. Lebrun, F. Mangeant, J. Sen Gupta
EADS Innovation Works, Suresnes & Toulouse, France

M. Pendola, T. Yalamas
PhiMECA, Clermont-Ferrand & Paris, France

Résumé :

Le besoin d'assurer des performances robustes pour des systèmes complexes ont considérablement accru l'intérêt de modéliser et propager les incertitudes à travers des environnements de simulation physique complexes numériquement. Les enjeux industriels impose que la méthodologie et les méthodes numériques du traitement d'incertitudes soient validées ouvertement et enrichies par la communité scientifique et les autorités de certification. Grâce à un effort commun de groupes industriels et d'institutions académiques, une méthodologie générique a émergé et une liste de méthodes numériques robustes ont été sélectionnées pour l'appuyer. EDF R&D, EADS Innovation Works et PhiMECA ont développé une plateforme open source dédiée au traitement des incertitudes dans le cadre probabiliste ; cette plateforme s'appelle OpenTURNS pour Open source Treatment of Uncertainties, Risks 'N Statistics

Abstract:

The need to assess robust performance for complex systems have led to a considerable rise of industrial interest in the simulation challenges to treat and propagate uncertainty through complex physical and numerical simulation frameworks. The industrial stakes require that both the methodology and the numerical methods for uncertainty treatment have to be openly validated and enriched both by the academic world and the certification authorities. A general methodology has emerged from the joint effort of both industrial companies and academic institutions, and a list of well-established numerical methods have been selected to support this methodology. EDF R&D, EADS Innovation Works and PhiMECA have developed an Open Source software platform dedicated to uncertainty treatment by probabilistic methods, named OpenTURNS for Open source Treatment of Uncertainty, Risk 'N Statistics.

Mots-Clés : Ingénierie – industrie, Statistique mathématique

1. Industrial context and rationale for an Open Source initiative

Uncertainty, along with sensitivity analysis, is the subject of a long-standing concern in the academic community and of substantial on-going research [1, 2, 3]. Although large-scale industrial uncertainty studies have been developed already in the 1990s [4, 5], there has been a considerable rise of interest in many industries in the recent decade as evidenced by the number of international conferences. Industry needs generally to answer tighter regulatory processes (security, safety, environmental control, health impacts...) gradually requiring explicit uncertainty treatment, and to better optimize the design, operation and maintenance of industrial processes or products including uncertainty.
Many attempts to treat uncertainty in large industrial applications have involved domain-specific approaches or standards: metrology [6], reliability [7], differential-based approaches [8], variance

decomposition [9].
Facing the questioning of their certification authorities in an increasing number of different domains or businesses, the large industrial companies to which the authors belong have felt that domain-specific approaches are no more appropriate: in spite of the diversity of terminologies, most of these methods do share in fact many common algorithms, and notably a mixture of estimation of probabilistic quantities involving also the computation of often high-CPU deterministic physical-numerical models. Beyond the leverage effects that can be achieved through the merging of those approaches, the peculiar industrial challenges attached to the recent uncertainty concerns are:

- **transparency**: open consensus that can be understood by outside authorities and experts;
- **genericity**: multi-domain issue that involves various actors along the supply chain;
- **industrial computing capabilities**: to secure the challenging number of simulations generated by uncertainty treatment;

The authors felt that no commercial software fully answered the challenges mentioned here above and launched the OpenTURNS initiative, with the following key features:

- **open source initiative** to secure the transparency of the approach, and its openness to on-going R&D development and expert challenging;
- **generic to the physical or industrial domains** to treat of multi-physical problems;
- **including the largest variety of qualified algorithms**: there is no single ever-relevant methodology, it is preferable to rely on a port-folio of algorithms, up-to-date with R&D;
- **structured in a «practitioner-guidance» methodological approach**: the question in an industrial study is not firstly « how » to compute the uncertainty treatment; it is rather « what » the specific goal to be answered by the uncertainty study is and which best chain of models, given the data practically available on uncertainty, should be selected. ;
- **with advanced industrial computing capabilities**, enabling the use of massive distribution & high performance computing, various engineering environments, large data models etc.

To start up the Open Source Initiative, a partnership was set by Electricité de France (EDF), European Aerospace Defense and Space Company (EADS) and PhiMECA.

The following paragraphs will detail the key characteristics of OpenTURNS in its initial version (§2.) as well as the wider objectives envisioned by the initiative (§3.).

2. OpenTURNS key characteristics

Technical aspects

The software has been developed since 2004. It is composed of a **scientific C++ library** where the data model and algorithms are programmed. A **python module**, based on the library, allows to access high level operators. With the conviviality of an interpreted language. For the moment, OpenTURNS is a Unix/Linux software but it will be available on Windows OS this year.
The software is **open source** under the LGPL licence. It can be downloaded on http://www.openturns.org. It is also avalaible on Debian and Ubuntu Linux distributions.

Scientific content

The algorithms of OpenTURNS have been selected following many discussions between industrial partners like within the European Safety, Reliabity and Data Association (ESReDA). Beyond the standard methodologies, some are scientifically more recent algorithms. OpenTURNS is consistent with the methodological framework proposed by the ESReDA Uncertainty Project Group [10], which decomposes an uncertainty study into four steps:

- ***Step A:*** **specify the inputs, model, outputs and the decision criterion**. The criterion could be the central dispersion of the interest variable, its distribution or a probability to exceed a threshold. The model can be as well an analytical formula as a complex finite element model.

- ***Step B*****: quantify the sources of uncertainty**. This step consists in modeling the probabilistic density function of the random input vector. If data is avalaible, statistical functionalities are avalaible:
  - density function fitting methods for parametric models and kernel smoothing (with any types of kernel) for non parametric models;
  - quantitative and qualitative tests (X^2, Kolmogorov Smirnov, Anderson Darling, BIC, Henry line, QQ-plots, graphical superposition of the empirical and fitted CDF
  - methods to estimate the dependence between two scalar variables (Pearson, Spearman and X^2 tests, R^2 coefficient of a linear regression, graphical superposition of 2D samples and regression line, Fisher test)

  If no sample is available, there are three ways to model the probabilistic density function:
  - assigning multidimensional gaussian distribution and giving a large choice of 1D distributions as in [11-13]
  - using usual probabilistic density function to each of the 1D input and copulas to describe the dependence structure. The available copulas are the independent copula, the gaussian copula and archimedean ones (Gumbel, Frank, Clayton). The correlation matrix is evaluated by the Spearman rank matrix or the Kendall tau matrix.
  - Defining the PDF as a mixture of usual probabilistic density functions.

- ***Step C*****: propagate uncertainties**. To evaluate the mean and standard deviation of the variable of interest, one can use quadratic combination method or simulation to analyze the central dispersion as in [14]. To evaluate the probability to exceed a threshold, one can use FORM and SORM approximations validated by a Strong Maximum Test or simulation with standard and accelerated algorithms (Monte Carlo, HLS, Directional Simulation, Importance Sampling) with the estimation of a confidence interval.

- ***Step C'*****: analyze sentivity and rank uncertainties**. OpenTURNS proposes the importance factors and sensitivity indices evaluated from the Taylor quadratic variance approximation or the correlation coefficients (Pearson, Spearman, SRC, SRRC, PCC, PRCC). More advanced variance decomposition techniques (such as Fast or Sobol) should come up within later versions. Sobol indices are in the code but are yet to be validated. In the case where we study the probability to exceed a given threshold, OpenTURNS proposes the importance factors evaluated from the FORM and SORM methods.

Moreover, aside these four steps, OpenTURNS gives the possibility to build a response surface used in the analysis instead of the model. OpenTURNS proposes to build either a global information response surface from a second degree polynomial expression, either a local information response surface from a one or second degree Taylor approximation of the model. Polynomial chaos will be available in the next release of the software, which is planned to be released before the conference.

Innovative aspects of OpenTURNS

OpenTURNS works directly in the multidimensional context by placing the probabilitic density function of the random vector in the center of all its algorithms. All methods are implemented in the multidimensional context.
Distributions are classifed (continuous, discrete, elliptic, ...) in order to take the best benefit of their properties in algorithms.
OpenTURNS offers recent statistical methods as kernel smoothing methods and the possibility to define a probabilistic density function as a linear combination of probabilistic density functions.
OpenTURNS uses the copula theory to model the dependence between variables. Gaussian copula and three archimedean copulas are available.
OpenTURNS offers some innovative algorithm as the Strong Maximum Test for FORM, proposed by EDF R&D.
OpenTURNS used the function algebra until the second order, all the functions are defined in the nD context as they go from R^p to R^n, with n and $p \geq 1$.
If the external model evaluates its gradients and its Hessian matrix, OpenTURNS takes them into account.
OpenTURNS is able to compose functions as well as their gradients and Hessian matrix.

OpenTURNS architecture makes it a software platform for easy integration of innovations from the open source community in the field of uncertainty treatment.

3. Native partners objectives: industrial uses and challenges scientific opportunities

EDF-R&D internal goal is to use OpenTURNS in studies where the uncertainty treatment has already been undertaken for some time, but in a rather domain-specific way using commercial softwares or research applications, particularly in the thermo-hydraulic, mechanical, or environmental field… Furthermore, EDF R&D aims to extend the uncertainty treatment methodology to other fields that do not perform probabilistic studies yet, enjoying OpenTURNS as a user-friendly support for a consistent uncertainty treatment methodology, integrating the multi-disciplinary experience of other fields.
EADS-Innovation Works goal is first to use OpenTURNS for demonstrative studies where uncertainty treatment is forecast: structural mechanics, electromagnetic interferences, tolerancing… The main goal is to spread an uncertainty methodology treatment enabling to change the design rules applicable within the aeronautical business. The OpenTURNS environment will enable to promote such a general approach at different levels for different domains. It also represents a very efficient way to begin a collaborative work with the software industry to introduce these functions in commercial numerical tools.
PhiMECA goal is to promote and capitalize some of the innovative uncertainty methods developed by its R&D team in order to use and propose them in an efficient way for its customers. The proposal may concern personalized services such as uncertainty consultancy, effective deployment of the tool, theoretical and applicative training courses or software developments using OpenTURNS. Applications concern various fields in advanced mechanics for automotive, aeronautics, nuclear, civil engineering or defense industries.

The three native partners welcome the initiative of the open source community to use OpenTURNS for their uncertainty treatment studies or research developments.
In particular, they wish to increase the open community of users both industrials and from universities, in order to facilitate the uses and sharing of innovations in the uncertainty field.
They are also motivated to enrich the OpenTURNS into a new one, including not only more complex

algorithms dedicated to specific needs (like other dependence structures: Archimedean copulas or recent sensitivity algorithms for example), but also new notions (like Bayesian aspects for example).
At last, they want to increase the open source community of contributors to enrich the software with developments of different public.
To give them the possibility to reach these objectives, a Consortium has been created around OpenTURNS, regrouping the three native partners but it is open to others.

4. Conclusion

This paper aimed at presenting the OpenTURNS initiative and the platform in its initial version. It explained why the three partners EDF R&D, EADS Innovation Works and PhiMECA started such a collaboration in 2005: in order to master industrial and environmental risks and uncertainties, with open practices, they launched the development of an open source software dedicated to the treatment of uncertainty, which would be generic (multi-physical / multi-domain issue) and consistent with a global methodology for the treatment of uncertainty.
The wider objectives envisioned by the initiative are to make OpenTURNS a reference software in the field of uncertainty, shared both by the industrial world and the academic one. The platform has already been used for a large number of studies in different divisions of EADS such as Airbus or Astrium, using complex softwares as physical model (Abaqus, code-aster, ASERIS, etc.).

Bibliographie

[1] Granger Morgan M., Henrion M. *(*1990*), A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis,* Cambridge University Press.
[2] Knight, F.H. (1921), *Risk, Uncertainty and Profit.* Hart, Schaffner & Marx.
[3] Madsen H. O., Krenk, S., Lind N. C. (1986), Methods of Structural Safety, Prentice Hall.
[4] Helton, J. *(*1993*),* Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal*, Rel. Eng. & Syst. Saf., 42,* pp. 327-367
[5] Helton, J.C., Oberkampf W.L. (2004)*,* Alternative Representations of Epistemic Uncertainty*, Special Issue of Rel. Eng. & Syst. Saf., vol. 85 n°1-3.*
[6] *GUM, Guide for the Expression of Uncertainty of Measurement,* ISO standard.
[7] Ditlevsen O. & Madsen H.O. (1996)*, Structural reliability Methods,* John Wiley & Sons.
[8] Cacuci, D.G., et al. (1980), Sensitivity Theory for General Systems of Nonlinear Equations, *Nucl. Sc. & Eng. 75*
[9] Saltelli, A., Tarantola, S., Campalongo, F., Ratto, M. *(*2004*), Sensitivity analysis in practice: a guide to assessing scientific models,* Wiley.
[10] De Rocquigny E., Devictor N., Tarantola S. (2008), Uncertainty in Industrial Practice, a guide to quantitative uncertainty management, John Wiley & Sons.
[11] Breitung K. (1984) Asymptotic Approximation for multinormal Integrals, *Journal of Engineering Mechanics, ASCE, 110(3)*, pp. 357-366.
[12] Breitung K. (1989) Asymptotic Approximation for Probability Integrals, *Probabilistic Engineering Mechanics, Vol 4, No 4,*, pp. 187-190.
[13] Cambanis S. , S. Huang, G. Simons (1981) On the theory of elliptical contoured distributions, *University of North Carolina, Journal of multivariate analysis vol 11*, pp. 368-385.
[14] Robert C.P., Casella G. *(*2004). *Monte-Carlo Statistical Methods,* Springer, ISBN 0-387-21239-6, 2nd ed.