

A new R bundle for design and analysis of computer experiments

Céline Helbert^{1,*}, Delphine Dupuy¹ and Yves Deville²

1. École des Mines de Saint-Étienne (Dpt. G2I/3MI) 158, Cours Fauriel - 42023 Saint-Étienne Cedex 2 (France)

2. Statistical Consultant - 37, rue Lamartine 73000 Chambéry (France) (www.alpestat.com)

* Contact author: helbert@emse.fr

Keywords: computer experiments, space-filling designs, metamodeling, kriging, global optimization.

The DICE consortium* is a partnership between academic and industrial laboratories from various fields (e.g. petroleum, automotive, nuclear etc.). DICE aims at developing and applying statistical methods to solve problems related to computer experiments: uncertainty propagation through a flow simulator, global optimization of a car crash simulator or calculus of failure probabilities in a nuclear process.

Each methods used through the consortium applications are implemented and integrated in a numerical toolbox that allows to fully process a case study. This toolbox is actually a R bundle comprising the following 4 packages:

- `DiceDesign` for space-filling design of experiments;
- `DiceEval` for evaluation and comparison of metamodels;
- `DiceKriging` for the estimation of a response surface via Gaussian Processes;
- `DiceOptim` for global optimization.

Each package addresses a specific task where classical and new methods are implemented. Among the new methods, `DiceDesign` proposes the construction of numerical designs of experiments based on stochastic processes and the computation of discrepancy and distance criteria. `DiceEval` is especially dedicated to fit usual models. A validation procedure (containing numerical criteria, graphical plots and cross-validation methods) is also proposed to validate a fitted model. `DiceKriging` implements a large panel of kriging models used for uncertainty propagation or for global optimization. The component `DiceOptim` performs different versions of EGO algorithm (Jones et al., 1998).

The different case studies proposed by the industrial partners of the DICE consortium have been processed by the bundle. One of these case studies will served as a running example to illustrate its main functionalities. In this application, the space-filling design (inputs of the simulations) is done by `DiceDesign`. `DiceEval` is then used to modelize the output of the simulator and allows to fit classical models: linear models, `additive{gam}`, `PolyMARS{polspline}` and `kriging{DiceKriging}` models. Moreover, `DiceEval` provides graphical tools in order to compare the quality of the fitted models on learning, validation and test sets.

Acknowledgements.

This work was conducted within the frame of the DICE consortium between ARMINES, Renault, EDF, IRSN, ONERA and Total S.A. The authors wish to thank J. Franco, O. Roustant, D. Ginsbourger and L. Carraro for their contributions.

References

- Hastie T., Tibshirani R. and Friedman J. (2001). **The Elements of Statistical Learning: Data Mining, Inference and Prediction**, Springer.
- Jones D.R., Schonlau M. and Welch W.J. (1998). Efficient Global Optimization of Expensive Black-Box Functions. *Journal of Global optimization*, 13, 455–492.
- Welch W.J., Buck R.J., Sacks J., Wynn H.P., Mitchell T.J. and Morris M. (1992). Screening, Predicting, and Computer Experiment. *Technometrics*, 34 (1), 15–25.

*DICE stands for *Deep Inside Computer Experiments* (<http://dice-consortium.fr/>).