Stochastic Response Surface Methodology in Medicine with censored/uncensored data analysis

Teresa Oliveira[1], Conceição Leal[2], Amílcar Oliveira[3]

Abstract— Stochastic Response Surface Methodology (SRSM) is very well known as a tool for modeling uncertainty in simulation models, mainly in the context of Risk Analysis. Developments on SRSM were remarkable over the past few years, mainly due to the technological advances and new powerful tools in what concerns hardware and software. It has also been possible to extend and explore this methodology in many areas and to state its importance in studies connected with Health problems, more specifically in Medicine revealing to be crucial on helping specialists to attain better and accurate Diagnosis and Prognosis. In recent papers, see [4], [5] and [34], [35], the authors paid attention to the application and behavior of the methodology when applied to a sample with censored data. In this work the authors present an extension of previous research involving the analysis of the whole available data contained in Wisconsin Prognostic Breast Cancer database. Uncensored data and right censored together data will be studied in order to compare both censored/uncensored results and to attain more realistic results. The aim is to provide more appropriate models to prediction and to risk analysis, since the models obtained with uncensored data suffer from a right bias. Computational results and graphics were implemented using R software.

Keywords— Breast Cancer Prognosis, Polynomial Chaos, Risk Analysis, Stochastic Response Surface Methodology, Uncertainty.

I. Introduction

In this paper we review Stochastic Response Surface Methodology as a tool for modeling uncertainty in the context of Risk Analysis, specifically in healthcare. An application in the survival analysis in the breast cancer context, with noncensored and censored data, is implemented with R software.

Uncertaity quantification on Risk Analysis has a crucial relevance, particularly in complex dynamical systems. Stochastic process variability and epistemic uncertainty state uncertainty quantification as a prerequisite on probabilistic risk assessment, either in experimental data or in model parameters. Thanks to the recent technological advances, simulation techniques became current to estimate models allowing to predict systems' behaviors, with respect to the probability of occurrence of a specific event and the consequences of this occurrence.

Probabilistic uncertainty quantification demand numerical simulation whose computational costs are often very high, thus the use of metamodels arises as a pressing necessity. The Response Surface Methodology (RSM) is known to be a suitable tool for the quantification of uncertainty. Stochastic Response Surface Methodology (SRSM) was developed as a conceptual extension of the traditional RSM, to approximate model inputs and outputs in terms of random variables, such as standard normal variables, by a polynomial chaos expansion.

The objective of the methodology is to reduce the number of model simulations required for adequate estimation of uncertainty, as compared to conventional Monte Carlo methods. The use of metamodel in conjunction with Monte Carlo method allows a faster estimation of the response system CDF that characterize the uncertainty propagation on the model. The importance of such methodology is very well known, mainly in the survival analysis context, so we will review it in the next section.

II. The Stochastic Response Surface Methodology – Expansion into Polynomial Chaos

Monte Carlo method is the most popular one to probabilistic quantification of covariate uncertainty propagation in the system response and to estimate its probability distribution. However, this may have very high computational costs and it is necessary to rely on methodologies that converge more quickly to the solution.

The SRSM allows the generation of a metamodel, computationally less demanding and statistically equivalent to the complete numerical model, which can be used to implement sensitivity analysis. It is needed a limited number of simulations of the complete model to get the response of the system for the model coefficients estimation. Reference [32] introduce this methodology and present two case studies. The basic idea of the methodology is to represent the response of a model to changes in covariate, using a response surface defined with a polynomial basis that is orthogonal with respect to a probability measure on the space of parameters. The SRSM lies on the assumption that the random variables, whose probability density functions are square integrable, can be approximated by the expansion in stochastic series of random variables or direct transformation of these, see [26] for details.

In the last years SRSM was subject to important developments that release of assumptions that are not always realistic. The arbitrary and data-driven approaches of the methodology are important examples of these developments. In the classic version, a vector of random variables $\xi=(\xi_i)$, $i = 1,..,n$, is selected, under $N(0,1)$ distribution, representing

ISBN: 978-1-61804-253-8

uncertain variables of a model in such way that $x_i = h(\xi_i)$. After this selection, response variables are represented as a function of the same vector of random variables: $Y = f(c, \xi)$, being c a vector of coefficients to estimate. The response of the system model to the various achievements of ξ, allows the estimation of model coefficients. The coefficients c_i quantify the dependence of the response Y on the input vector ξ, for each realization of x.

Consider Ψ_i polynomials which form a base of orthogonal polynomials to a given probability measure. The form of the function f result of the polynomials chaos expansion is expressed by:

$Y = f(c,\xi =$

$$= c_0\Psi_0 + \Sigma_{i1=1}^{\infty} c_{i1}\Psi_1(\xi_{i1}) + \Sigma_{i1=1}^{\infty}\Sigma_{i2=1}^{i1} c_{i1i2}\Psi_2(\xi_{i1},\xi_{i2}) + \ldots$$

This series is approximate by a truncated polynomial in n-th power. The roots of the polynomial of $n+1$ degree can be used in the simulation model in order to get the data to estimate the model coefficients.

The methodology is implemented sequentially as follows: (i) representation of uncertain input variables; (ii) representation of the response variable; (iii) estimation of the model parameters; (iv) calculation of the response statistical properties; (v) evaluation of the model answers approximation.

The following algorithm illustrates the application of the methodology.

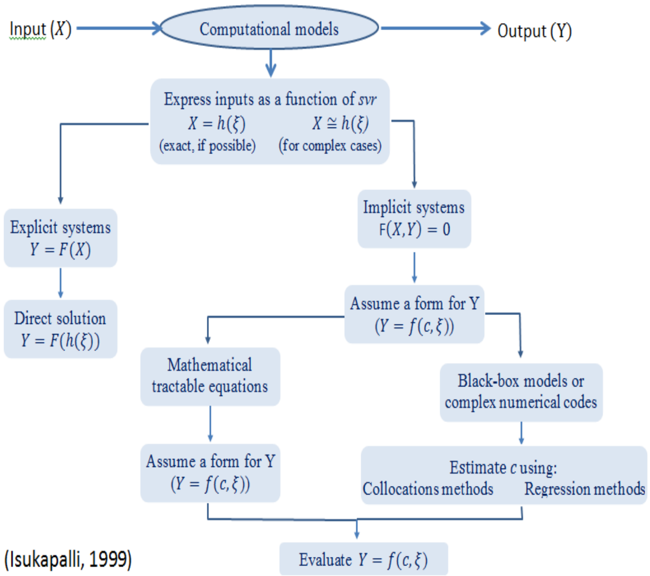


Fig. 1: Schematic depiction of the Stochastic Response Surface Method

In the classical approach the measure is the Gaussian and the polynomials are the Hermite polynomials, which form is as follow (see [19] and [24]).

$$PCE = c_0 + \Sigma_{i=1}^{n} c_i\xi_i + \Sigma_{ii=1}^{n} c_{ii}\left(\xi_i^2 - 1\right) + \Sigma_{i=1}^{n-1}\Sigma_{j>i}^{n} c_{ij}\xi_i\xi_j$$

Reference [10]–[13] showed that it is possible to obtain a better approximation of the response variables using non-Gaussian expansions in polynomials chaos. In this case, the Hermite polynomials are replaced by orthogonal polynomials with respect to the probability measure of input variables, [10]. This approach was designated generalized polynomial chaos expansion. Reference [20] presented conditions on the probability measures involving the mean square convergence of the generalized polynomials chaos expansion.

Reference [30] proposed a new generalization of the methodology, called arbitrary polynomial chaos expansion or data-driven chaos expansion. In this new approach, the probability distributions and the probability measures are arbitrary. Statistical moments are the only source of information that is propagated in the stochastic model. Probability distributions may be discrete, continuous, or continuous discretized, may be specified by analytical way (PDF or through CFD), numerically using a histogram or by using the raw data. In this approach, all distributions are admissible for the input variables of a given model, as long as they have in common a finite number of moments. Thus, in the case of considering truncated polynomial, there is just the need to know a finite number of moments, with no need for complete knowledge of the probability density function or even its existence, which frees the researcher from the need of assumptions that may not always be supported by existing data. According to the literature it is known that this expansion converges exponentially and faster than the classical expansion.

The estimation of model parameters depends on the model complexity, see [33]**.** In case the model is invertible, the parameters can be obtained directly from the input random variables $(\xi_i)_{i=1}^{n}$. If the model equations are mathematically manipulated, in spite of nonlinearities, then the model coefficients can be obtained afterwards, by an appropriate norm minimization of residuals, replacing the input random variables by the respective transformations in terms of Gaussian variables $N(0,1)$, known as Galerkin method, for details see [33]. When the model equations are difficult to manipulate, the coefficients can be estimated by the collocation points method. Each set of points is chosen such that the model estimates are accurate at these points, given by the set of the N resulting linear equations.

Reference [33] present some methods for parameter estimation, all based on the collocation points methods: Probabilistic Collocation Method, Efficient Collocation Method and Regression Based Method and these authors discuss advantages and disadvantages for each method.

The expansion in polynomials chaos is a simple but powerful tool for stochastic modeling. Probability density function, probability distribution functions or other statistics of interest can be estimated and quickly evaluated via Monte Carlo simulation (MCS), once the evaluation of a polynomial function is faster than the original equations model evaluation.

ISBN: 978-1-61804-253-8

In the case of risk analysis, to use arbitrary expansion, one can directly use a set of large-sized data or probability density function of maximum or relative minimum entropy, since, in this case, the relevant moments of the polynomial chaos expansion are compatible with those of the input variables. The bootstrap resampling method may be used to obtain more precise estimates of the moments from a reduced set of data available, providing a more accurate estimation of the risk assessment model. Reference [31] propose such an application on calibration models to history matching for CO_2 storage in underground reservoirs.

III. LITERATURE APPLICATIONS

The RSM in its various approaches has an important role in generating reduced models to replace the simulator complex processes requiring a very high number of simulations. Applications are diverse and many of them relate to the Stochastic Response Surface Methodology for quantification of uncertainty, in stochastic processes.

Reference [32] applied the SRSM to four case studies cover a range of applications, both from the perspective of the model application (biology, air quality and groundwater) and its complexity.

Reference [6] applied the SRSM to study the role of various hydro geologic parameters in the uncertainty assessment of the chemical contaminant concentration in groundwater, resulting from the nuclear industry, so that one can designing the waste disposal facilities and remedial action plans. This study provides a program of environmental monitoring in the nuclear industry.

Reference [8] applied the SRSM to analyze the reliability of stochastic stability of rock slopes involving correlated non-normal variables.

Reference [29] applied the SRSM, based on the expansion of arbitrary polynomial chaos, to a problem of contaminant transport in 3D heterogeneous aquifer and the risk to human health caused by exposure of the population.

References [27]– [31] have applied the SRSM, with different approaches and combined with other methodologies, in various problems relating to the storage of CO_2 in underground geological formations and the associated risks.

Reference [1] combined subset simulation method with the SRSM to perform probabilistic analysis of a shallow strip footing. As well as other aspects, they analyzing uncertainty propagation of the soil shear strength parameters on the ultimate bearing capacity. This problem, that involves the computation of the ultimate bearing capacity of a strip footing, is presented to demonstrate the efficiency of the proposed procedure by comparing the results given by MCS methodology applied on Polynomial Chaos Expansion (PCE) models and on original deterministic model.

Reference [14] applied the SRSM in a model of gas injection into an incompressible porous media, and showed its effectiveness in uncertainty and sensitivity analysis of complex numerical models.

Reference [2] applied the generalized SRSM to assess leakage detectability at geologic carbon storage sites under parameter uncertainty. They demonstrated how it can be used to construct probability maps for assessing the detection of anomalies in the coverage of underground geological storage formations, in space and time.

Reference [7] applied the SRSM to estimate the propagation of uncertainties in the parameters of the function retention of strontium in the human body.

Reference [9] applied the SRSM to analyze the reliability of an underground cavern, associated with deterministic finite element methods. More specifically, the SRSM was used to perform the probabilistic analysis serviceability performance of the cavern.

SRSM was not yet conveniently explored in applications on the field of medicine, despite uncertainty confound the understanding of the essential medical facts and how they are integrated ([25]. The uncertainty parameters, the heterogeneity of the patient and the stochastic uncertainty results are increasingly important concepts in models of medical decision. [25] and [3] show various methods for analyzing uncertainty and the heterogeneity of patients in decision models.

The breast cancer prognosis is markedly heterogeneous, and research has often focused on the effect of prognostic factors related to the disease, such as the expression of estrogen receptors, tumor size and others. Still remains an open question of modeling the data recurrence time, the complexity of the shape of the hazard function over follow-up period, and the identification of factors that may affect it, using a fully parametric approach ([16].

In the next section we present a study in the prognostic breast with uncensored and uncensored data using SRSM.

IV. APPLICATION IN HUMAN HEALTH: CENSORED/UNCENSORED DATA

SRSM was applied to the Wisconsin Breast Cancer Prognostic (WPBC) database which is constituted by data of 253 patients with breast cancer, who have undergone surgical excision of invasive cancer. First, we used only 69 of these cases that correspond to the patients that had remission till the end of the study. Second, we used all cases, including 184 that are right censored because they withdrew from the study or we know only the time of the last examination. The database is very well described by [36], who are responsible for its design and construction.

We estimated, by regression on sample data, a first and a second order polynomial chaos expansion with Hermite

ISBN: 978-1-61804-253-8

polynomials, in which the response variable is the time to recur (TTR) and the uncertainty parameters are three of the ten available covariates, related with morphological characteristics of the nucleus of malignant cells: AREA, TEXTURE and SIZE. The first two variables result from the average of the measurements in the three nuclei with higher values from all nuclei retained in the image. Since there are more than 30 cases the normality of covariates was assumed, as well as independence once there is a low correlation. Different models were simulated with respect to the degree of the polynomial and to the number and nature of the included covariates. The second order model was used to estimate the distribution of the response variable and the resulting survival and risk functions.

The proposed study was implemented by using the R program language with *survival*, *fitdistrplus*, *Hmisc* and *EQL* packages.

The distribution which best fits to TTR data with remission was gamma distribution (lower values of statistics, AIC and BIC criteria obtained with *gofstat* function in PDF than what we got with *fitdist* function, booth from *fitdistrplus* package).

A data-frame was constructed containing the variables processed in accordance with appropriate transformation to a normal distribution: $x_i = \mu_i + \sigma_i\, \xi_i$.

In the first order model, it was observed that variable SIZE did not give a significant contribution to the model. Adjusting the second order model, it was observed that the intersect and the terms $AREA$, $AREA^2 - 1$ and interaction WAREA$\times SIZE$ are significant. A Cox model was fitted and it was observed again that SIZE covariate did not reveal statistical significance. The assumption test to proportional hazards reveals no statistical evidence that these are not proportional to each variable in the overall model.

A similar study was carried out with the censored data.

The distribution that best fits to TTR with censored data was Lognormal distribution (lower values of AIC and BIC criteria obtained with *fitdistcens* function from f*itdistrplus* package).

The Cox model reveal that AREA and SIZE were statistically significant and the second order model shown that the $AREA$, SIZE, $AREA^2 - 1$, $SIZE^2 - 1$ and AREA$\times TEXTURE$ interaction are statistically significant.

Using the second order models the response variable PDFs were simulated by Monte Carlo sampling (Fig.2 –data only with remission and censored data) and it can be compared with empirical PDF (Fig.3).

```
> Nsim<-10^4
> t<-0
> X1<-0
> for (i in 1:Nsim) {
+ u1=rnorm(1)
+ u2=rnorm(1)
+ u3=rnorm(1)
+ TTR<-function(a1,a2,a3) {
+ PCE<- 30.2256 -13.5806  *a1 -4.3718  *a2+ 2.4267*a3+
+ 3.9472 *(a1^2-1)-0.9645*(a2^2-1)+4.7846*(a3^2-1)-2.4691
*(a1*a2)+2.2082   *a1*a3-0.8439 *(a2*a3) }
+    t<-TTR(u1,u2,u3)
+    X1[i]<-t  }
>     X1
```

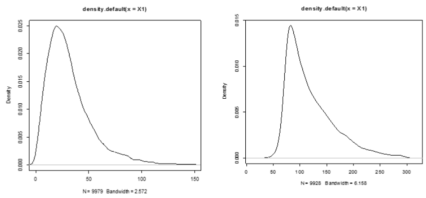


Fig. 2. Simulated density functions of TTR

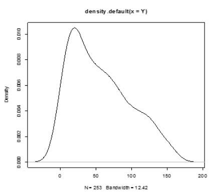
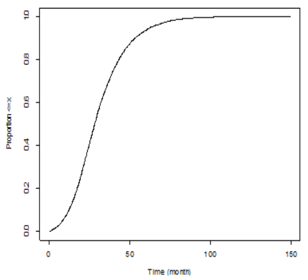


Fig. 3. Empirical density function of TTR

The survival functions (disease-free survival time) can be estimated by: $S(t) = 1 - F(t)$, where F is the distribution function of TTR (Fig.4 and Fig. 6) (with *Ecdf* function from *Hmisc* package) and this function can be compared with the Cox model (Fig. 5 and Fig. 7).



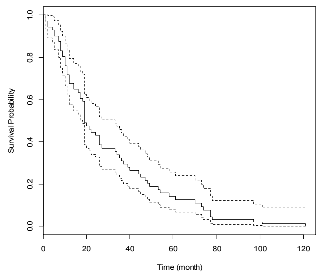
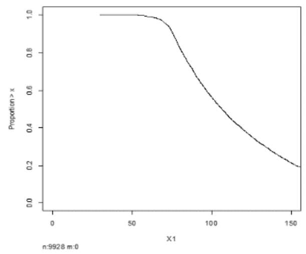


Fig. 4. Simulated survival function only with remission data



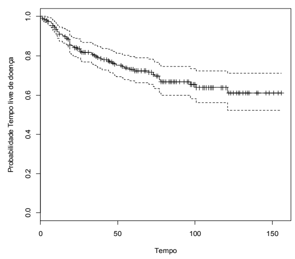


Fig. 6. Simulated survival function with censored data

To estimate the hazard function, which is defined by $H(t) = \frac{f(t)}{S(t)}$, $f(t)$ being the density function of TTR, we need to get the theoretical distribution that best fits the survival functions $S(t)$ (Fig.7 and Fig. 9). These were Gamma and Lognormal functions, respectively. Once identified the distributions that best fit the survival functions, it is possible to estimate H and obtain its graphical representations (Fig. 10 and Fig. 11).

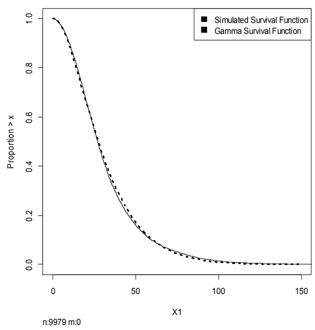


Fig. 8: Survival function adjusted with simulated only with remission data (obtained from the Gamma PDF fitted to simulated data.)

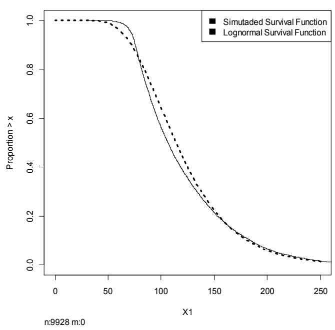


Fig.9: Survival function adjusted with simulated censored data (obtained from the Lognormal PDF fitted to simulated data.)

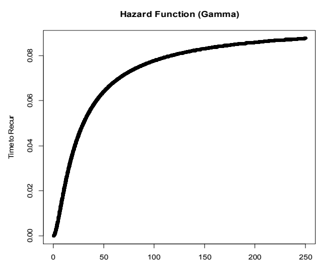


Fig. 10. Hazard function (obtained with Gamma PDF and survival function fitted only with simulated remission data)

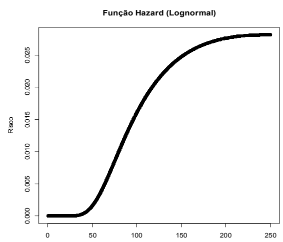


Fig.11. Hazard function (obtained with Lognormal PDF and survival function fitted with simulated censored data)

V. Considerations and Conclusion

The application of SRSM in the computer modelling of complex systems is crucial, since the full model simulation usually requires high computational costs. The methodology has been applied mainly to the analysis of environmental risk problems in particle transport and fluid, and on the analysis of structural reliability. In this work its application to a problem of survival analysis, in connection with morphological features of cell nuclei of breast tumour data, was considered.

A model with Hermite polynomials was fitted to estimate the time to recurrence of breast cancer after tumor excision, according to three characteristics of the malignant cells nuclei: extreme area, extreme texture and tumor size. Two studies were performed: the first one only with recurrence data, and therefore uncensored data, and second, adding to these the censored data, data from patients who had no recurrence until the end of the study or those that left the study before it has finished.

In the first study it was expected that results would suffer from a bias. As indicated by [37], most observed recurrences were at relatively short time (a mean of 24 months) and, therefore, the simulation on the model obtained with a regression method using just these data should result in low prediction of recurrence time, confirming the bias of this particular data set (mean of 31 months). However, it should be noted that [22] refer that it is expected a peaks of recurrence about 18 months after surgery.

In the censored data, the higher probability of recurrence occurs not very far from 60 months and mean of TTR is about 119 months. The peak at 60th month is often referred in the literature, see for example [18] and [17]. In the survival function it can be observed that it is about the sixtieth month that the probability of recurrence begins to decrease significantly, coinciding with that peak of occurrence.

Acknowledgment

This work was partially sponsored by national founds through the Fundação Nacional para a Ciência e Tecnologia, Portugal - FCT under the project (PEst-OE/MAT/UI000 6/2014).

References

[1] A. Ahmed, A. H Soubra, “Extension of subset simulation approach for uncertainty propagation and global sensitivity analysis.” *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards* 6.3: pp. 162-176, 2012.

[2] A. Y.Sun, M. Zeidouni, J. P. Nicot, Z. Lu, D. Zhang, “Assessing leakage detectability at geologic CO< sub> 2</sub> sequestration sites using the probabilistic collocation method”, Advances in Water Resources, 56, pp. 49-60, 2013.

[3] B. G.Koerkamp, M. C. Weinstein, T. Stijnen, M. H. Heijenbrok-Kal, M. M. Hunink, “Uncertainty and patient heterogeneity in medical decision models”, Medical Decision Making, 30(2), pp.194-205, 2010.

[4] C. Leal, T. A. Oliveira and A. Oliveira, "Stochastic Response Surface Methodology: A study on polynomial chaos expansion", SMTDA Proceedings, Lisbon. In Press, 2014.
[5] C. Leal, T. A. Oliveira and A. Oliveira, "Stochastic Response Surface Methodology: A Study in the Human Health Area", in T. E. Simos, G. Psihoyios, Ch. Tsitouras and Z. Anastassi (eds.), Numerical Analysis and Applied Mathematics ICNAAM 2014, AIP Conference Proceedings, American Institute of Physics, to be published, 2014.
[6] D. Datta, H.S. Kushwaha, "Uncertainty Quantification Using Stochastic Response Surface Method Case Study-Transport of Chemical Contaminants through Groundwater", International Journal of Energy, Information & Communications, 2(3), 20011.
[7] D. Datta, "Uncertainty modeling of retention function in biokinetic model using polynomial chaos theory-development of computational algorithm", International Journal of Mathematical Archive (IJMA) ISSN, pp.2229-5046, 4(4), 2013.
[8] D. Li, Y. Chen, , W. Lu, C. Zhou, "Stochastic response surface method for reliability analysis of rock slopes involving correlated non-normal variables", Computers and Geotechnics, 2011, 38(1), pp. 58-68.
[9] D. Q. Li, S. H. Jiang, Y. F. Chen, C. B. Zhou, "Reliability analysis of serviceability performance for an underground cavern using a non-intrusive stochastic method", Environmental Earth Sciences, 71(3), pp. 1169-1182, 2014.
[10] D. Xiu, G. E. Karniadakis, "Modeling uncertainty in steady state diffusion problems via generalized polynomial chaos", Computer Methods in Applied Mechanics and Engineering, 191(43), pp. 4927-4948, 2002a.
[11] D. Xiu, G. E. Karniadakis, "The Wiener--Askey polynomial chaos for stochastic differential equations", SIAM Journal on Scientific Computing, 24(2), 619-644, 2002b.
[12] D. Xiu, G. E. Karniadakis, "Modeling uncertainty in flow simulations via generalized polynomial chaos", Journal of Computational Physics, 187(1), pp 137-167, 2003a.
[13] D. Xiu, G. E. Karniadakis, "A new stochastic approach to transient heat conduction modeling with uncertainty", International Journal of Heat and Mass Transfer, 46(24), pp.4681-4693, 2003b.
[14] E. Bastug, A. Menafoglio, T. Okhulkova, "Polynomial Chaos Expansion for an Efficient Uncertainty and Sensitivity Analysis of Complex Numerical Models", Conference Paper. ESREL 2013, Amsterdam, Netherlands, 2013.
[15] G. Blatman, "Adaptive sparse polynomial chaos expansions for uncertainty propagation and sensitivity analysis", Ph.D Thesis, Université Blaise Pascal, Clermont-Ferrand ,2009.
[16] I. Ardoino, E. M. Biganzoli, C. Bajdik, P.J. Lisboa, P. Boracchi, F. Ambrogi, Flexible parametric modelling of the hazard function in breast cancer studies. Journal of Applied Statistics, 39(7), 1409-1421, 2012.
[17] I. Jatoi, A. Tsimelzon, H. Weiss, G. M. Clark, S. G. Hilsenbeck, "Hazard rates of recurrence following diagnosis of primary breast cancer", Breast cancer research and treatment, 89(2), 173-178, (2005.
[18] M. W. Retsky, R. Demicheli, D. E. Swartzendruber, P. D. Bame, R. H. Wardwell, G. Bonadonna,, ... P. Valagussa, "Computer simulation of a breast cancer metastasis model", Breast cancer research and treatment,45(2), 193-202, 1997.
[19] N. Wiener, "The homogeneous chaos", American Journal of Mathematics, 60(4), pp. 897-936, 1938.
[20] O. G. Ernst, A. Mugler, H. J. Starkloff, E. Ullmann, "On the convergence of generalized polynomial chaos expansions", ESAIM Math. Model. Numer. Anal, 46(2), pp.317-339, 2012.
[21] O. L. Mangasarian, W. N. Street, W. H. Wolberg, "Breast cancer diagnosis and prognosis via linear programming", Operations Research, 43(4), pp. 570-577, 1995.
[22] R. Demicheli, A. Abbattista, R. Miceli, P.Valagussa, G. Bonadonna, "Time distribution of the recurrence risk for breast cancer patients undergoing mastectomy: further support about the concept of tumor dormancy", Breast cancer research and treatment, 41(2), pp. 177-185,1996.
[23] R Development Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0, 2012.
[24] R. G. Ghanem, P. D. Spanos, "Stochastic finite elements: a spectral approach (Vol. 41). New York: Springer-Verlag, 1991.
[25] R. S. Dittus, S.D. Roberts, J. R. Wilson, "Quantifying uncertainty in medical decisions", Journal of the American College of Cardiology, 14(3), A23-A28, 1989.
[26] S.Balakrishnan, A. Roy, M. G. Ierapetritou, G. P. Flach, P.G. Georgopoulos, "Uncertainty reduction and characterization for complex environmental fate and transport models: An empirical Bayesian framework incorporating the stochastic response surface method", Water Resources Research, 39(12) , 2003.
[27] S. Oladyshkin, H. Class, R. Helmig, W. Nowak. "Highly Efficient Tool for Probabilistic Risk Assessment of CCS Joint with Injection Design", Computational Geosciences, 13, 451-467, 2009.
[28] S. Oladyshkin, W. Nowak, "Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion", Reliability Engineering and System Safety,106, pp.179-190, 2010.
[29] S. Oladyshkin, H. Class, R. Helmig, W. Nowak, "A concept for data-driven uncertainty quantification and its application to carbon dioxide storage in geological formations", Advances in Water Resources, 34(11), pp.1508-1518, 2011.
[30] S. Oladyshkin, W. Nowak, "Data-driven uncertainty quantification using the arbitrary polynomial chaos expansion", Reliability Engineering & System Safety, 106, pp.179-190, 2012.
[31] S. Oladyshkin, H. Class, W. Nowak, "Bayesian updating via bootstrap filtering combined with data-driven polynomial chaos expansions: methodology and application to history matching for carbon dioxide storage in geological formations", Computational Geosciences, 1-17, 2013.
[32] S .S. Isukapalli, S. S., Roy, A., Georgopoulos, P. G.: Stochastic response surface methods (SRSMs) for uncertainty propagation: application to environmental and biological systems. Risk analysis, 18(3), pp. 351-363, 1998.
[33] S.S. Isukapalli, "An uncertainty analysis of transport transformation models", Ph.D. Thesis, New Brunswick, New Jersey: The State University of New Jersey,1999.
[34] T.A. Oliveira, C. Leal, A. Oliveira, "Stochastic Response Surface Methodology: A Study in the Human Health Area". ICNAAM Proceedings, Rhodes-Greece. In Press, 2014.
[35] T.A Oliveira, C. Leal and A. Oliveira, " Response Surface Methodology: a review of applications to risk assessment", in Kitsos, C., Oliveira, T., Rigas, A. and Gulati, S. (eds.), Chapter XXIX in Theory and Practice of Risk Assessment, Springer Proceedings in Mathematics and Statistics, to be published (2015).
[36] W. H. Wolberg, W. N. Street, O. L. Mangasarian, "Importance of nuclear morphology in breast cancer prognosis", Clinical Cancer Research, 5(11), pp.3542-3548, 1999.
[37] W. N. Street, O. L., Mangasarian, W. H. Wolberg, "An inductive learning approach to prognostic prediction" In ICML (pp. 522-530) ,1995.
[38] Wisconsin Breast Cancer Prognosis Dataset. http://pages.cs.wisc.edu/~olvi/uwmp/cancer.html#prog

Teresa A. Oliveira

Teresa A. Oliveira is Assistant Professor of Statistics in the Department of Sciences and Technology, at the Universidade Aberta (UAb), Portugal, www.uab.pt, since 2000. She coordinates de Master in Biostatistics and Biometry. She holds a PhD in Statistics and Operations Research - Experimental Statistics and Data Analysis (University of Lisbon-2000), and MS in Statistics and Operations Research (Faculty of Sciences, University of Lisbon-1994). Her research interests include Experimental Design, Statistical Modelling, Data Analysis and Risk Analysis as well as Mathematical and Statistical e-Learning. Co-editor of the volumes "Risk Assessment Challenges: Theory and Practice", by "Springer Proceedings in Mathematics and Statistics" and of the volume "Statistical and Biometrical Challenges: Theory and Applications", she has published several papers in international journals, books and proceedings. She has supervised several PhD and master thesis regarding her interest fields. She co-organized several international conferences on Statistics, Mathematics and Computation (8^{th} Edition - March 2014), as well as Portuguese-Polish Workshops on Biometry (4^{rd} Edition September 2014). She is an integrated member of the Center of Statistics and Applications of the University of Lisbon and she is a collaborator in UIED-Research Unit on Education and Development, Faculty of Sciences and Technology of New University of Lisbon, in IPM- Preventive Medicine

Institute, Faculty of Medicine of University of Lisbon and in LEaD - Laboratory of Distance Learning of Universidade Aberta. She is a member of Portuguese Statistical Society SPE), International Biometric Society (IBS) and International Statistical Institute (ISI). She is the scientific secretary of the Committee on Risk Analysis of the International Statistical Institute (CRA-ISI), see http://www.isi-web.org/44-com/com/126-ra from January 2014.

Conceição Leal

Conceição Leal teaches Mathematics in Escola Secundária de Paços de Ferreira, Portugal, since 1991. She was Department Coordinator and Principal Investigator of the Self-School Evaluation. She holds a degree in Mathematics and a pos-graduation on Statistics, Mathematics and Computation by Universidade Aberta (UAb), Portugal. She is finishing Master thesis on Statistics, Mathematics and Computation and the pos-graduation on Biostatistics and Biometry. Her main research interests include Pure and Applied Mathematics, Experimental Design, Statistical Modelling, Data Analysis and Risk Analysis. She has participated in several international conferences with oral presentations and she has published several papers in Response Surface area.

Amílcar Oliveira

Assistant Professor of Statistics on courses of 1st and 2nd cycle offered by the Department of Sciences and Technology, Universidade Aberta (UAb), Lisbon, Portugal.
He holds a Ph.D. in Mathematics, (Statistical Modelling), Universidade Aberta, an Master degree. in Statistics and Optimization, Faculty of Sciences and Technology, New University of Lisbon.
His research interests include Statistical Modelling. He has published several papers in international journals, books and proceedings regarding these fields. He is integrated member of the Center of Statistics and Applications, University of Lisbon.

ISBN: 978-1-61804-253-8