Statistical Estimation of Aircraft Infrared Signature Dispersion

Sidonie Lefebvre, Antoine Roblin, Suzanne Varet, Gérard Durand
ONERA DOTA, Chemin de la Hunière, 91761 Palaiseau Cedex France (sidonie.lefebvre@onera.fr)

Abstract — Existing computer simulations of aircraft InfraRed Signature (IRS) do not account for the dispersion induced by uncertainty on input data such as aircraft aspect angles and meteorological conditions. As a result, they are of little use to estimate the detection performance of optronic systems: in that case, the scenario encompasses a lot of possible situations that must indeed be addressed, but can not be singly simulated. In this paper, a three-step methodological approach for predicting simulated IRS dispersion of poorly known aircraft is proposed. The first step is a sensitivity analysis. The second step consists in a Quasi-Monte Carlo survey of the code output dispersion. In the last step, a metamodel of the IRS simulation code is constructed. This method is illustrated in a typical scenario, namely an air-to-ground full-frontal attack by a generic combat aircraft, and gives a satisfactory estimation of the infrared signature dispersion.

1. **Introduction**

Knowledge of aircraft InfraRed Signature is indispensable for assessing their detection probability, and thus their survivability in an hostile environment. By signature, we mean all the quantities for predicting the signal that would be observed by an optronic sensor when the aircraft is in its surroundings. For many reasons, the experimental approach is generally not feasible to evaluate the IRS: aircrafts may not be available, and the IRS is also needed in configurations which cannot be reached easily due to safety reasons. Computer programs, which enable to evaluate the IRS of aircraft and backgrounds, are therefore extremly valuable tools. Existing computer simulations of aircraft IRS do not account for the dispersion induced by uncertainty on input data, such as aircraft aspect angles and meteorological conditions. As a result, they are of little use to estimate the detection performance of IR optronic systems: in that case, the scenario encompasses a lot of possible situations that must indeed be addressed, but can not be singly simulated. Hence, the simulated result is no longer a single IRS value, but an interval of possible IRS, which should include the IRS measured at a given instant.

We focus in this paper on a scalar response: the sensor differential irradiance between target and background. The performance criterion associated to an optronic sensor, in this case, simply consists in the probability that the sensor irradiance produced by the aircraft is below some given threshold α, which represents the background clutter. In order to simplify the analysis, only the 3-5 μm spectrally integrated target intensity is considered. ONERA has developed for thirty years a simulation of combat aircraft IRS, CRIRA, initiated by Gauffre (1981). We aim at defining a general methodology for predicting, using CRIRA, simulated IRS dispersion of poorly known military aircraft and non-detection probabilities for typical thresholds. This methodology will be helpful to size surveillance IR sensors and to evaluate their performances.

A black box representation is associated to the IRS computer simulation code f : $Y=f(X_1,\ldots,X_n)$ where the n X_i denote the uncertain input factors of the code, and Y is the output of the simulation. The non-detection probability P_α associated to the threshold α is given by:

$$P\left(\left|f\left(X_1,\ldots,X_n\right)\right|<\alpha\right)=\int_{[0,1)^n} I_{|f(X_1,\ldots,X_n)|<\alpha}\left(\vec{t}\right)p_{X_1,\ldots,X_n}\left(\vec{t}\right)d\vec{t}$$

where I stands for the indicator function and $p_{X_1,\ldots,X_n}$ is the joint probability density function of the X_i. A well-known tool to estimate such probabilities is the Monte Carlo stochastic sampling. The main drawback of this method is the slow convergence, scaling asymptotically with the inverse square root of the number of samples and starting from what is often a large initial error. It is

therefore not uncommon to need more than one million samples to guarantee an accuracy better than one percent. The computational cost can thus be prohibitive. Alternative approaches have been developed, in order to speed up the convergence, such as Quasi-Monte Carlo method. The convergence rate depends on the problem's dimension *n*, which can be decreased by focusing exclusively on input factors that are really significant, according to their impact on the code output dispersion. We thus propose a three-step methodological approach for predicting simulated IRS dispersion of poorly known aircraft. The first step is a sensitivity analysis, which identifies inputs that have negligible influence on the IRS and can be set at a constant value. The second step consists in a Quasi-Monte Carlo survey of the code output dispersion. In the last step, a neural network metamodel of the IRS simulation code is constructed. This method is illustrated in a typical scenario, namely a daylight air-to-ground full-frontal attack by a generic combat aircraft flying at low altitude.
The sensitivity analysis is described in Section 2, the estimation of IRS dispersion in Section 3, and the metamodel construction in Section 4. The results are discussed in Section 5.

2. Sensitivity analysis with a fractional factorial design

Many inputs of CRIRA are uncertain: from twenty up to sixty. Some inputs are set at a constant value by the scenario: they can not induce any uncertainty in IRS and are not taken into account in this statistical study. For our scenario, 28 input data are uncertain:

- 9 describe IR optical properties of the various aircraft surfaces, their symbol is E_...
- 7 are related to flight conditions: altitude, Mach number, engine's power setting, and to aspect angles of the aircraft: cap, yaw, roll and pitch angles,
- 12 are related to atmospheric conditions: visibility (vis), relative humidity (hr), ground air's temperature (ta), atmospheric model, model of aerosol (ihaze), clouds presence, thickness and altitude of base (hbase) of the cloud layer, ground's albedo (salb), hour to compute solar position, deviation from the mean day of the season (iday), deviation of ground temperature from the mean temperature of the season (deltat).

Three factors: model, ihaze and clouds are qualitative, the others are quantitative. The correlations between two or more factors concern the eight first factors related to atmospheric conditions. A single run of our simulation requires about three minutes, we thus keep the number of simulation runs below 4000 for the sensitivity analysis step.
Several approaches enable to carry out the sensitivity analysis of a computer simulation, among which stand out Sobol' indices estimation, described in Saltelli et al. (2000), and use of Design Of Experiments (DOE), reviewed in Myers and Montgomery (1995). We assume that interactions among two or three factors can be significant, but that interactions involving more than three factors are negligible. We want to properly estimate factor effects and interactions between two factors. The evaluation of Sobol' and total indices associated to each factor would require too much CPU time, we thus make use of DOE, and favor fractional factorial designs over screening designs, due to the importance of interactions. Each factor is then described by two levels, chosen thanks to knowledge on operational conditions, meteorological databases, and in order to minimize factors correlations. The factors are normalized, so as to vary between -1 (low point) and 1 (high point). A fractional factorial design does not contain all the 2^{28} possible combinations of factors levels. Hence, it enables to estimate not all the interactions among input variables, but groups of interactions . The size of the interactions forming a group, or aliase, depends on a design property: the resolution. We make use of a resolution VI design, as it insures that factor effects and interactions between two factors are aliased with interactions involving at least four factors, which are assumed to be negligible. For 28 factors, this design decomposes in 2048 runs. We compute the aircraft IRS associated to the design of experiments, and analyze them through a second-degree polynomial model, under the assumption that the residuals are Gaussian and that the factors are independent:

$$Y = c_0 + \sum_i c_i . X_i + \sum_{i<j} c_{ij} . X_i . X_j + \varepsilon_r$$

with Y the size 2048 vector of outputs, X_i the size 2048 vector of values of i^{th} factor, c_0 the outputs mean, c_i the i^{th} factor effect, c_{ij} the interaction effect between the i^{th} and j^{th} factors, and ε_r the size 2048 vector of residuals.

The outputs dispersion is very large, about four orders of magnitude, we thus perform the variance analysis on the natural logarithm of the IRS. The model coefficients are then fitted using least squares method, and a Student's test ascertains their significance. The Pareto plot of Figure 1 depicts the percentages of the 28 $|c_i| \Big/ \sum_{i=1}^{28} |c_i|$, in descending order. The dotted line represents the cumulative part of IRS variance explained by the i most influential factors.

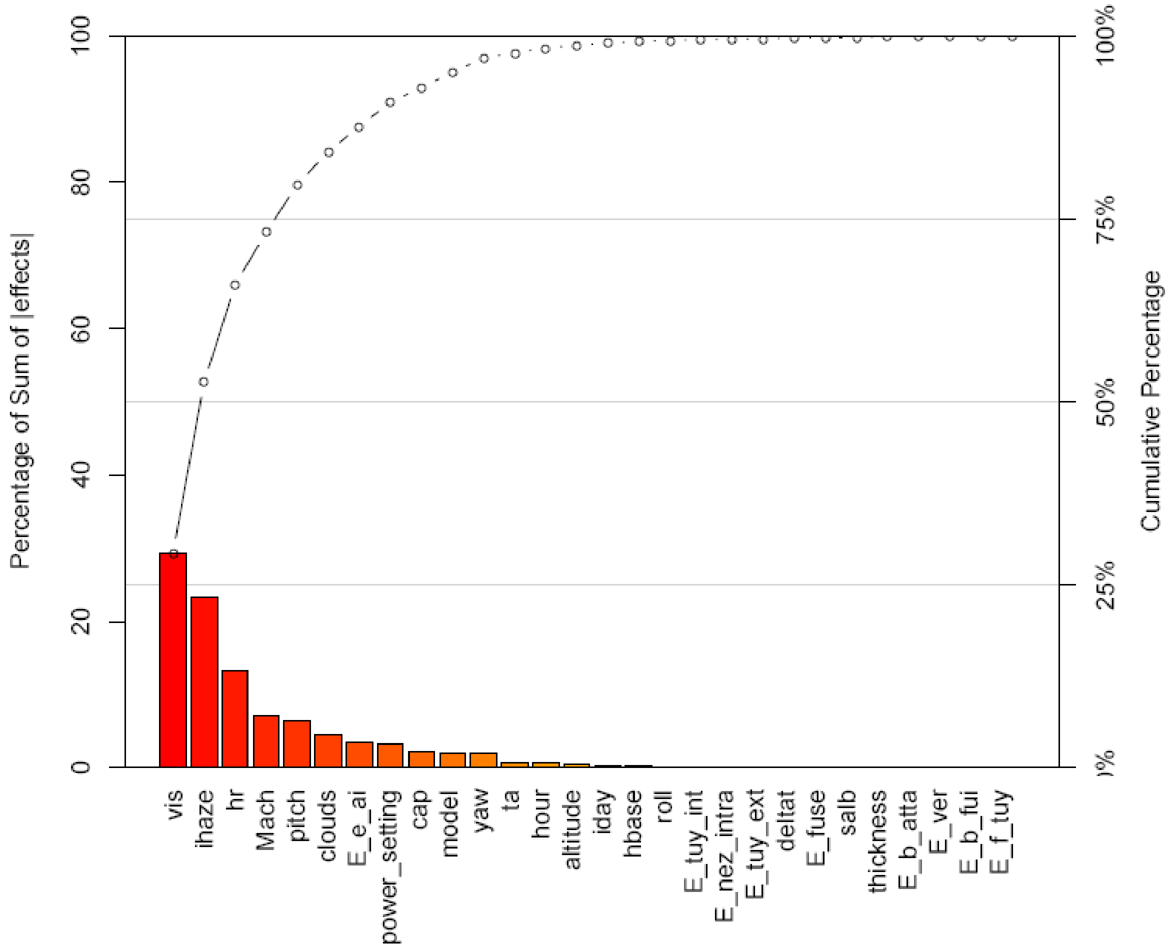

Figure 1. Pareto plot of the 28 factors effects

The analysis of variance shows that 80 % of the IRS variance is explained with only five factors. We do not want to leave off potentially important factors, so we keep the ten most significant variables before going on to the next stage. These variables explain 95 % of the IRS variance: five are atmosphere related factors: vis, ihaze, hr, clouds, model, four are flight conditions variables: mach, pitch, power setting, cap, and one is related to IR optical properties of aircraft surfaces: E_e_ai. The other factors are set at a constant value for the next steps. We have also checked that all the variables involved in significant interactions between two factors were retained. Only one factor associated to aircraft characteristics is important in this scenario, because most of the aircraft properties have been supposed to be perfectly known, and thus do not appear in the sensitivity analysis. Other conclusion could be raised, if the aircraft was considered less known.

The factors independence and the gaussian residuals assumptions are not fully borne out. We thus have proposed in Varet et al. (2009) to use two-levels fractional factorial designs to get an estimation of the Sobol' indices. This approach enables to estimate the variance part explained by factor effects and interaction effects between two factors assuming solely the independence of eleven base factors. For the chosen scenario, this method leads to very same input selection.

3. Quasi-Monte Carlo estimation of the IRS dispersion

The Quasi-Monte Carlo method makes use of a slightly different kind of sampling than the Monte Carlo one, as described in Tuffin (2007): the pseudo-random numbers are replaced with uniformly distributed determinist sequences, the low discrepancy sequences, to improve the accuracy of approximations for a fixed number N of simulation runs. The discrepancy D_N^* is a measure of the uniformity of the points dispersion. A low discrepancy sequence is characterized by a $O\left(\frac{\log(N)^n}{N}\right)$ discrepancy, where n is the problem's dimension, that is to say the number of significant factors in this paper. Koksma-Hlawka theorem gives an upper bound of the convergence rate, but it is quite difficult to estimate for practical purposes. However, several authors, as Caflisch, Morokoff and Owen (1997) report a better convergence rate than Monte Carlo one's even for large n. Construction algorithms of low discrepancy sequences lead to correlations among points coordinates. Hence in high dimension, most of low discrepancy sequences present distribution irregularities on low order projections, in particular for consecutive dimensions, as stated in Morokoff and Caflisch (1994) and Schmid (2001). If the interaction between the associated variables is significant, these irregularities are quite impeding. Moreover, the determinist nature of these sequences is a major drawback for confidence intervals estimation. Randomization methodologies have thus been developed, like the scrambling ones, initiated by Owen (1995): they preserve the low discrepancy, add randomness and decrease projections irregularities.

In this study, we use a scrambled Faure's sequence, with a Faure-Tezuka scrambling (2003). N data set $(X_{i1},X_{i2},...,X_{iN})$, i in $\{1,..10\}$, are generated for $X_1,...,X_{10}$ the ten significant factors, and the N outputs $(y_1,y_2,...,y_N)$ computed enable to estimate:

- the empirical cumulative distribution function of the IRS for our scenario by:
-

$$F_N(x) = \frac{1}{N}\sum_{i=1}^{N} I(y_i \le x)$$

- the non-detection probability P_α by:
-

$$P_{\alpha N} = \frac{1}{N}\sum_{i=1}^{N} I\left(\left|IRS\left(X_{1i},...,X_{10i}\right)\right| < \alpha\right)$$

Given the computation time, we limit N to 10000. The non significant input variables are fixed at a constant value. Among the ten most significant variables, only five (vis-ihaze-hr-nuages-model) can not be described by a uniform law. These factors are related to atmospheric conditions, and are dependent. We can find their values in meteorological databases, but we have not found enough data yet to estimate a joint probability density function. We thus perform random sampling with replacement from the database to obtain a combination of real values, instead of using scrambled Faure's sequences for these factors. Figure 2 depicts the empirical cumulative distribution function of the IRS. For confidentiality reasons, all the IRS values are scaled by an arbitrary constant.

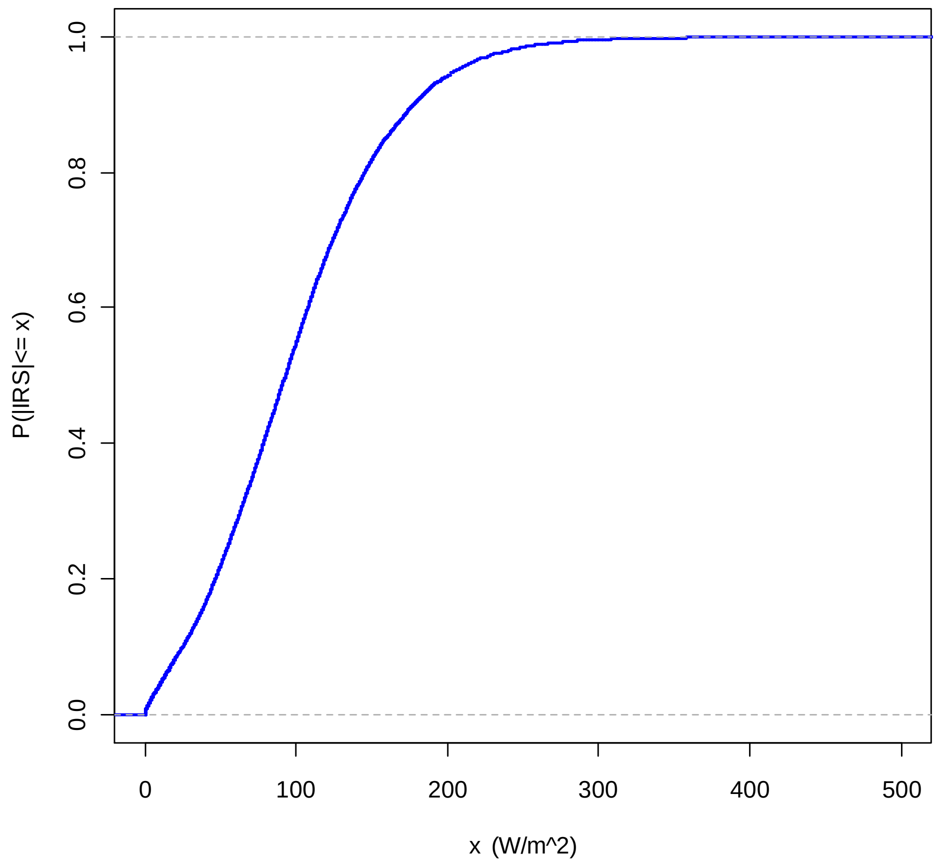

Figure 2. Empirical cumulative distribution function of the IRS

We have checked that, if we perform 50 bootstrap draws of two-thirds of the 10000 IRS values, the corresponding gap among empirical cumulative distribution functions is negligible, about 0.5 %. Realistic thresholds for non-detection probabilities depend a lot on the optronic sensor we want to size. We therefore test our methodology by estimating three quantiles β: 1 %, 5 % and 25 %, which correspond to typical non-detection probabilities. We make use of the empirical estimator: $\inf\{y, F_N(y) > \beta\} = y_{\lceil \beta N \rceil}$ after reordering. Table 1 gives 95% confidence level bootstrap estimation based on 5000 draws among the 10000 IRS values, for different sample sizes. A good evaluation of 5 % and 25 % quantiles is obtained with 2000 values.

	1 %	5 %	25 %
250	[1.63, 1.71]	[11.36, 11.56]	[54.8, 55.2]
500	[1.57, 1.62]	[11.44, 11.6]	55
1000	[1.34, 1.38]	[11.12, 11.24]	55
2000	[1.23, 1.26]	[11, 11.08]	55
5000	[1.18, 1.2]	[10.94, 10.98]	55
10000	[1.16, 1.17]	10.92	55

Table 1. Bootstrap estimations of three quantiles: 1 % - 5 % - 25 %

4. A neural network metamodel

In order to check whether some optronic sensor meets required specifications, we do not need accurate estimations of extreme IRS values. Very small IRS values are indeed neither detected, whereas large IRS are consistently detected. It is therefore of great interest to build a metamodel of the IRS simulation code, suited to a chosen scenario, and to save the use of the much more expensive simulation for IRS close to typical detection thresholds. It would allow carrying through computationally demanding tasks, such as optimization of optronic sensor properties. Our computer simulation of aircraft IRS is complex, and linear regression models give poor predictions, we thus preferred nonlinear modeling, more precisely neural network metamodels described in Hastie et al. (2001).

A neural network can be described as an oriented graph built up from a set of neurons, which are nonlinear parametric functions organized in successive layers. All the neurons of a layer work in parallel. The most used network is the multi hidden layer back-propagation network, or multi-layer perceptron (MLP): the unknown parameters, or weights, are associated to the neurons input. Hornik (1989) and Barron (1993) theorems state that a single hidden layer MLP set up of a finite number of neurons, with the same nonlinear activation function, like the hyperbolic tangent, and a linear output neuron, is a parsimonious universal approximator. The weights are estimated by supervised training.

The input data are the ten most significant factors of § 2, they are reduced and centered. The output is ln|IRS|. We have compared the predictions of several single-layer perceptrons, with 2000, 4000 or 6000 training samples chosen by random sampling among the 10000 of § 3, and 5, 7 or 10 hidden neurons. The best results were obtained with a single-layer perceptron with 7 hidden neurons and 4000 training samples. The coefficient of regression R^2, computed on the 6000 remaining test samples, is 0.85. Figure 3 compares the empirical cumulative distribution functions of the IRS, obtained with the computer simulation CRIRA or with the neural network metamodel. The metamodel gives very satisfying IRS predictions and enables to estimate small non-detection probabilities with errors in the order of 1 %. We have also estimated the three quantiles: 1 % - 5 % - 25 % of 2000 neural networks with identical architecture but different training samples, drawned by bootstrap among the 10000 of § 3. Table 2 gives 95% confidence level estimation of these quantiles.

1 %	5 %	25 %
[1.42, 1.44]	[13.5, 13.6]	[53.9, 54]

Table 2. Bootstrap estimations of three quantiles of the metamodel: 1 % - 5 % - 25 %

The metamodel leads to a good approximation of the three quantiles, even if it slightly underestimates the 25 % quantile and overestimates the 1 % and 5 % ones. The neural network model is therefore a very useful tool to focus on special ranges of IRS.

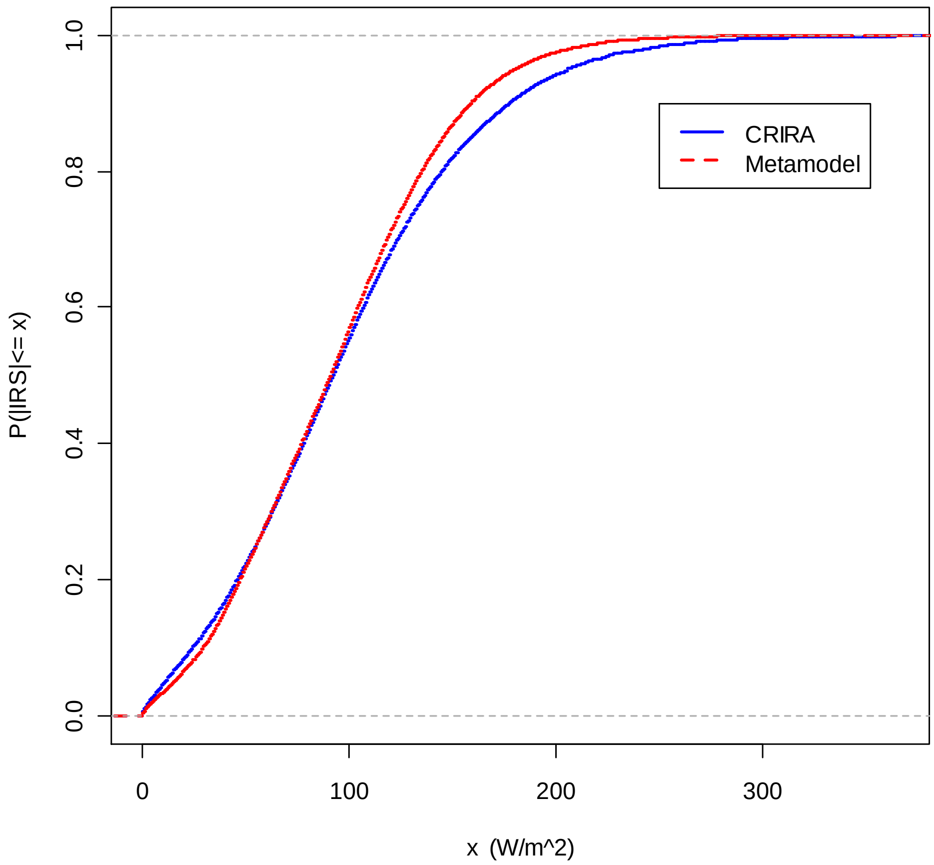

Figure 3. A comparison of empirical cumulative distribution functions of the IRS

5. Conclusion

We have developed a three steps statistical methodology to estimate infrared signature dispersion of a combat aircraft from our existing signature code. In the typical scenario chosen to illustrate this methodology: a daylight air–to-ground full-frontal attack, by a generic but partially known combat aircraft flying at low altitude, the atmosphere related factors are predominant and explain 75 % of the IRS dispersion. The three most significant variables are the aerosol's model, the relative humidity, and the visibility. We have checked that the IRS dispersion was properly estimated with ten thousand samples. The single-layer perceptron metamodel enables to obtain very satisfying estimation of IRS empirical cumulative distribution function and good approximation of quantiles of interest : 1 % - 5 % - 25 %.
Other approaches could have been applied instead of Quasi-Monte Carlo, but variance reduction methods such as importance sampling or controlled stratification were either unsuitable due to the lack of joint probability density function for the meteorological factors, or unprofitable due to the number of training samples (about 2000) needed to insure a good correlation between indicator functions of the code and of its metamodel. Although the methodology presented is general and can be used in different contexts, the results, in particular the significant input data, would most likely be different. We are currently listing all interesting battle scenarios and applying our methodology to study IRS dispersion for each. The case of a spatially resolved aircraft is particularly interesting, as it imply dealing with complete pictures (10x10 pixels) instead of a scalar signature as it was presented in this paper.
The present work constitutes a first attempt to statistically estimate signatures of a poorly known aircraft in statistically defined environmental conditions. As it is quite conclusive, it would be interesting to extend it to other military objects

Acknowledgments

This work has been partially supported by the French Defence Procurement Agency (DGA) grants. Discussions with S. Cohen, M. Sergent and R. Phan-Tan-Luu were very helpful.

Bibliography

Gauffre, G. (1981) Aircraft infrared radiation modeling. *Recherche Aerospatiale*, 4, 245–265.

Saltelli, A., Chan, K. and Scott M. (Eds.) (2000) *Sensitivity Analysis*. Wiley Series in Probability and Statistics. New York: John Wiley and Sons.

Myers, R. H., Montgomery, D. C. (1995) *Response surface methodology: process and product optimization using designed experiments,* NY, John Wiley & Sons.

Varet, S., Lefebvre, S., Durand, G., Roblin, A. and Cohen, S. (2009) Effective Discrepancy and Numerical Experiments, submitted to MCQMC2008 proceedings.

Tuffin, B. (1997) *Simulation accélérée par les méthodes de Monte Carlo et Quasi-Monte Carlo : théorie et application,* PhD in Applied Mathematics, University of Rennes 1.

Caflisch, R. E., Morokoff, W., and Owen, A. B. (1997), Valuation of mortgage backed securities using brownian bridges to reduce effective dimension, *Journal of Computational Finance*, 1, 27—46.

Morokoff, W., and Caflisch, R. E. (1994), Quasi-random sequences and their discrepancies, *SIAM Journal on Scientific Computing*, 15, 1251—1279.

Schmid, W. C. (2001), Projections of digital nets and sequences, *Mathematics and Computers in Simulation*, 55, 239—247.

Owen, A. B. (1995), Randomly permuted (t,m,s)-nets and (t,s) sequences in *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing*, H. Niederreiter and P.J.-.S. Shiue, editors, Springer Verlag, 299—317.

Tezuka, S. and Faure, H. (2003) I-binomial scrambling of digital nets and sequences. *J. of Complexity*, 19,744–757.

Hastie, T., Tibshirani, R., and Friedman, J. (2001) *The Elements of Statistical Learning,* NY, Springer.

Hornik, K., Stinchcombe, M., and White, H. (1989) Multilayer feedforward networks are universal approximators, *Neural Networks*, 2, 359–366.

Barron, A. (1993) Universal approximation bounds for superposition of a sigmoidal function, *IEEE Trans. on Information Theory*, 39, 930–945.