A Machine Learning Based Approach For Uncertainty Quantification, Application To Launch Vehicle Design

Jorge L. Valderrama

Internship supervisors: M. Balesdent, L. Brevault

August 31, 2021

Introduction

The coupling of uncertainty quantification methodologies with multidisciplinary optimization tools for the early design phase of launch vehicles is computationally intensive. This is mainly due to the strategies for multidisciplinary coupling satisfaction and the required optimal control methods for the trajectory discipline. The early design phase is characterized by a high number of input uncertain variables (*e.g.,* specific impulse, drag coefficient) that render uncertain the output fields (*e.g.,* the optimal speed profile as function of time and the optimal pressure distribution on aerodynamic surfaces). The output fields are comprised of a high number of correlated aleatory variables that make even more daunting the uncertainty quantification task. This work presents an Active Learning (AL) methodology for field variable quantile estimation relying on a surrogate model to reduce the computational cost. These quantiles are useful to characterize flight envelops of launch vehicles. An example case is demonstrated for the quantile estimation of the resulting state variables from the Multidisciplinary Design Analysis and Optimization (MDAO) of a Two-Stage-To-Orbit (TSTO) launch vehicle. The methodology improves the accuracy of the predicted quantiles and outperforms an aleatory enrichment strategy.

Literature review

Three main approaches can be distinguished to deal with uncertainty propagation in the multidisciplinary launch vehicle design problem involving optimal control as presented in [1]. The first one [2] uses Monte Carlo Simulation (MCS) on top of the MDAO problem, generating samples of the input uncertain variables (*e.g.,* engine specific impulse, mass flow rate, drag coefficient) and solving the MDAO problem for each sample. This leads to exact state variable realizations (*e.g.,* velocity, height) and quantities of interest (*e.g.,* dynamic pressure, heat flux, *etc.*). Such a nested approach can easily become prohibitive due to the associated computational cost. The second approach [3] consists on the decomposition of the optimal control problem using spectral methods, modifying the original MDAO problem to create an extended design space and solve a unique but complex optimization problem. The final approach [1] consists on the creation of a surrogate model based on model order reduction and spectral methods that can be exploited at a reduced computational cost. Nevertheless, the predictions issued with the surrogate model are approximations whose error is difficult to assess.

Gaussian processes are a non-parametric machine learning approach whose predicted outputs are random normal variables with known parameters instead of deterministic values. The output variance of the Gaussian process can be interpreted as an error model and has been used in active learning approaches to estimate quantiles of scalar variables [4]. To the best of my knowledge, such an active learning methodology has not been proposed in the literature for the refinement of quantiles of scalar random field variables (*e.g.,* a launcher trajectory).

Proposed methods

An active learning methodology for the estimation of scalar random field variable quantiles is proposed. It relies on the Karhunen-Loève expansion and Gaussian processes. The strategy consists of 5 steps.

Step 1:
A limited-size Monte-Carlo simulation is used to create the input training sample $\mathcal{U}_M$ and solve the exact MDAO function for each realization, this results in an output stochastic process sample $\boldsymbol{X}^*_M$.

Step 2:
The Karhunen-Loève (KL) decomposition of $\boldsymbol{X}^*_M$ is performed to reduce the dimension of the problem and obtain the most significant scaled KL modes and the uncertain variables $\xi_k(\boldsymbol{u})$.

Step 3:
Univariate Gaussian Processes (GP) are trained to mimic the mapping $\boldsymbol{u} \rightarrow \xi_k(\boldsymbol{u})$ and perform predictions on new inputs $\boldsymbol{u}$.

Step 4:
Predictions on a set of new inputs $\mathcal{U}_R$ are issued with

the GP and the KL scaled modes to obtain the estimated stochastic process sample $\hat{\boldsymbol{X}}_R^*$. A quantile of $\hat{\boldsymbol{X}}_R^*$ and its uncertainty are estimated using one of two proposed methods. The first method (called CI Area - A), is based on the aleatory generation of GP trajectories and the second one (called CI Area - B) is based on the analytic propagation of the GP error through the KL expansion. Both methods output an area corresponding to the confidence interval of the estimated quantile (quantile misknowledge).

Step 5:

The surrogate strategy is used to identify the new input realization that contributes the most to the reduction of the misknowledge area. This is achieved by updating the GP models inside an optimization loop. Then, the new input realization is evaluated with the exact function and added to the training sample. This is called an enrichment iteration.

Results and discussion

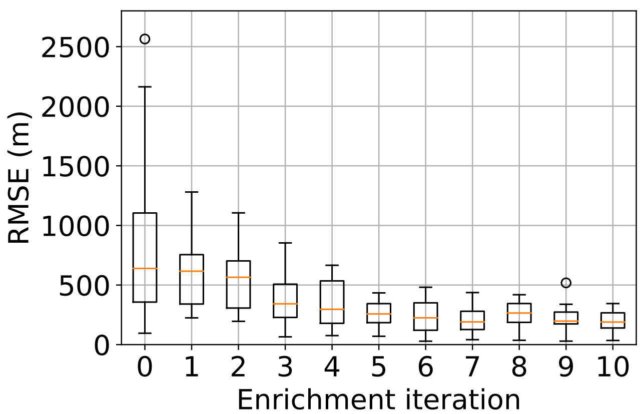

Figure 1: RMSE between estimated quantiles using the AL method and a validation set

7 random input variables, like drag coefficient and residual mass of propellants, are considered in the MDAO of a TSTO vehicle. An initial training sample of size 200 is generated and 10 enrichment iterations using the AL methodology are executed to improve the estimation of the 99% quantile of the altitude profile using the CI Area - A method. The procedure is repeated 10 times using initial aleatory training samples to generalize the analysis. The Root Mean Square Error (RMSE) between the estimated quantile using the AL method and the estimated quantile using a validation set is reduced more than 2 times in average after 10 enrichment iterations (see Fig. 1). A similar result is obtained using the CI Area - B method. No improvement in the RMSE is observed after enriching the original training set with 10 randomly chosen samples. The AL methodology is also executed to refine the 99% quantile of the heat flux profile, adding also 10 samples to the original training set. The added samples in the centered output space can be seen in Fig. (2). It can be noticed that some added sample are close to the 99% quantile in the first portion of the flight (0-18 %) and different ones are close to the same quantile in the second portion of the flight (18-40 %). This highlights the capacity of the AL method to locally refine the surrogate model.

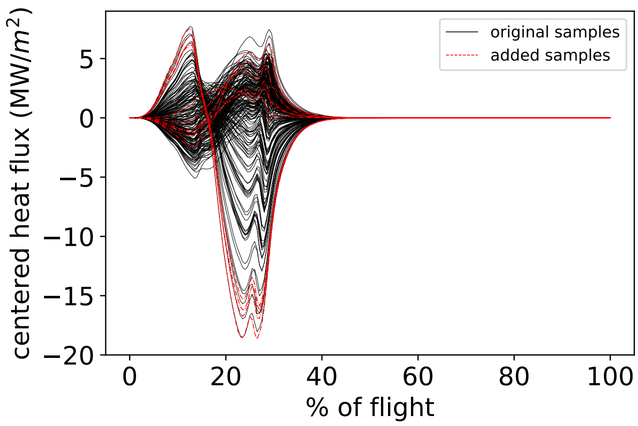

Figure 2: Training samples for the estimation of the 99% quantile of the heat flux profile

Conclusions

An active learning method for the estimation of quantiles of scalar random field variables is proposed within the context of MDAO for launch vehicle design and uncertainty quantification. It reduces the RMSE between the estimated quantile using the surrogate model and the quantile estimated using the validation set while controling the error. The AL methodology refines locally the surrogate model in the regions close to the target quantile.

Selected References

[1] L. Brevault and M. Balesdent, “Uncertainty quantification for multidisciplinary launch vehicle design using model order reduction and spectral methods,” *Acta Astronautica*, vol. 187, pp. 295–314, Oct. 2021.

[2] L. Brevault and M. Balesdent, “Uncertainty-Based Multidisciplinary Design Optimization (UMDO),” in *Springer Optimization and Its Applications*, vol. 156 of *Aerospace System Analysis and Optimization in Uncertainty*, pp. pp.235–292, 2020.

[3] F. Wang, Y. Shuxing, F. Xiong, Q. Lin, and J. Song, “Robust trajectory optimization using polynomial chaos and convex optimization,” *Aerospace Science and Technology*, vol. 92, June 2019.

[4] M. Moustapha, B. Sudret, J.-M. Bourinet, and B. Guillaume, “Quantile-based optimization under uncertainties using adaptive Kriging surrogate models,” *Structural and Multidisciplinary Optimization*, vol. 54, pp. 1403–1421, Dec. 2016.