Hierarchical metamodeling: cross validation and predictive uncertainty

Bianca M. Colosimo[a], Luca Pagani[b], and Matteo Strano[c]

Dipartimento di Meccanica - Politecnico di Milano, 20156 Milano, Italy

[a]biancamaria.colosimo@polimi.it, [b]luca.pagani@polimi.it, [c]matteo.strano@polimi.it

Keywords: Metamodeling, Metal Foams, Hierarchical Fusion Model

Abstract. At Esaform 2013 a hierarchical metamodeling approach had been presented, able to combine the results of numerical simulations and physical experiments into a unique response surface, which is a "fusion" of both data sets. The method had been presented with respect to the structural optimization of a steel tube, filled with an aluminium foam, intended as an anti-intrusion bar. The prediction yielded by a conventional way of metamodeling the results of FEM simulations can be considered trustworthy only if the accuracy of numerical models have been thoroughly tested and the simulation parameters have been sufficiently calibrated. On the contrary, the main advantage of a hierarchical metamodel is to yield a reliable prediction of a response variable to be optimized, even in the presence of non-completely calibrated or accurate FEM models. In order to demonstrate these statements, in this paper the authors wish to compare the prediction ability of a "fusion" metamodel based on under-calibrated simulations, with a conventional approach based on calibrated FEM results. Both metamodels will be cross validated with a "leave-one-out" technique, i.e. by excluding one experimental observation at a time and assessing the predictive ability of the model. Furthermore, the paper will demonstrate how the hierarchical metamodel is able to provide not only an average estimated value for each excluded experimental observation, but also an estimation of uncertainty of the prediction of the average value.

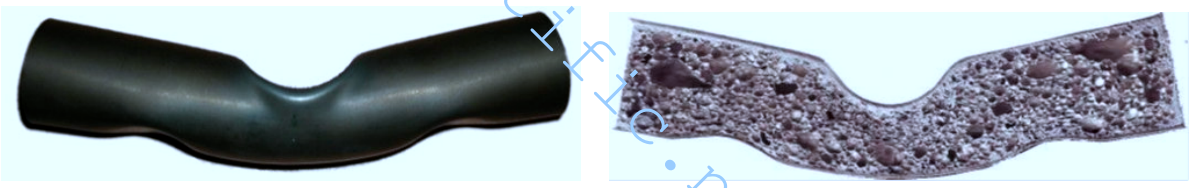
Introduction

The concept of metamodeling for computer simulations of manufacturing problems is known since more than two decades [1, 2]. At the beginning of years 2000, metal forming or plasticity numerical problems, which are usually very time consuming due to the complexity of physical phenomena involved at large deformations, have been increasingly metamodeled, for purposes like uncertainty assessment [3] or design optimization [4]. The typical approach is to use a small set of experimental results in order to either tune or verify a Finite Element Method (FEM) model. Then, a metamodel can be built solely on the base of accurate, calibrated, so-called "high-fidelity" simulations [5]. Alternatively, hierarchical metamodels have been proposed, based on a combination of high fidelity (Hi-Fi) and low fidelity (Lo-Fi) data sets [6]. In the available literature, high vs. low fidelity FEM models typically implies the use of finer vs. coarser meshes [7]. Alternatively, numerical models with different formulations have been used for modeling the same process: as an example in the optimization of sheet metal forming, outcome of a one-step solver (Lo-Fi) can be combined with data from an incremental solver (Hi-Fi) [8]. Another example can be found in the field of CFD (Computational Fluid Dynamics), where a multistage metamodeling technique that links data coming from two different numerical sources (finite volumes and finite differences calculations) has been implemented [9]. These "fusion" metamodels are used because they allow to obtain accurate predictions with less computational effort than metamodels based only on Hi-Fi results, provided that Lo-Fi simulations are faster, thanks to a simpler formulation or to a coarser mesh. These models are often called hierarchical, because there is a hierarchy between different data sets. Clearly, the fidelity of numerical results can only be assessed with reference to real, experimental data of the physical process under investigation. In metal forming and plasticity problems, authors that

deal with FEM simulations must spend a lot of time and effort in order to calibrate the most significant model parameters (coefficients of friction, flow stress and hardening parameters, anisotropy values, damage or failure thresholds, etc.) in order to match the numerical results. As an example, in [10], where a Kriging metamodel is used for optimizing the strength of a clinched joint, an entire Section of the paper is devoted to the validation of the FEM model. In this paper we wish to compare the prediction ability of a "fusion" metamodel based on under-calibrated simulations, with a conventional approach based on calibrated FEM results. Both metamodels will be cross validated with a "leave-one-out" technique, i.e. by excluding one experimental observation at a time and assessing the predictive ability of the model. Furthermore, the paper will demonstrate how the hierarchical metamodel is able to provide not only an average estimated value for each excluded experimental observation, but also an estimation of uncertainty of the prediction of the average value.

Real test case

In this Section we briefly introduce the real test case, that is already been analyzed in Colosimo et al. [11]. They have analyzed the design of an anti-intrusion side bar for vehicles, made of an outer tubular steel case and a filling reinforcement made of aluminum foam [12]. The filling of cases made of thin metal sheets or tubes with a reinforcement made of cellular metals (or metal foams), as for the structure shown in Figure 1, allows for the production of lightweight, high performance components, particularly suited for flexural resistance in terms of amount of energy absorbed for a given maximum load.



(a) Stainless steel tube after 3 point bending, filled by an aluminium foam

(b) Longitudinal cross section

Fig. 1: Steel tube

The have considered a closed section with a composite (bi-material) structural beam with initial length L, vertical average dimension H, horizontal average size W, initial average cross section area $S = H \cdot W$, total occupied volume $V = S \cdot L$, mass M, apparent density $\rho = \frac{V}{M}$. In lateral impact, the structure will undergo a flexural state of stress-strain. Given a load P $[kN]$ - deflection δl $[mm]$ diagram in bending of a foam filled bar, up to any value of deflection δl, the load curve profile will exhibit a maximum load value P_{max}, an average load P_{avg} and an amount of absorbed energy per volume $E_{abs} = P_{avg} \cdot \delta l$. The crash force efficiency can be written as the ratio between the mean load and the maximum load of a Force-Displacement curve [13]: $\eta = \frac{P_{avg}}{P_{max}}$. A body with high efficiency will have a large energy absorption, while limiting the maximum load P_{max} (and the corresponding acceleration) transmitted to the vehicle. A Specific Energy Absorption (SEA) can be also defined, as the ratio between the absorbed energy and the total mass: $SEA = \frac{E_{abs}}{M}$. In a lateral crash, for any given tubular composite structure and a given amount of incoming energy, it is important to:

- increase or maximize the energy absorption E_{abs}, given a maximum deflection δl_{max}, while limiting the total mass M (this is equal to maximize the SEA of the structure);
- increase or maximize the crash force efficiency $\eta(\delta l_{max})$;

- minimize the intrusion into the vehicle δl_{max}.

A synthetic objective function y can be built as a factorial combination of the above mentioned indicators:

$$y(\delta l_{max}) = \frac{SEA(\delta l_{max}) \cdot \eta(\delta l_{max})}{\delta l_{max}} = \frac{P^2_{\delta l_{max}}}{P_{max}(\delta l_{max}) \cdot M} \tag{1}$$

and measured in $[\frac{kN}{kg}]$. In the present case, the maximum admissible intrusion has been taken equal to $\delta l_{max} = 48\ mm$ for reasons related to design criteria of cars.

A design vector $\boldsymbol{x}$ has been selected for the proposed example, with two design variables: x_1 is a toughness indicator, related to the plastic material properties of the outer steel skin; the other variable x_2 is a shape factor, related to its geometry.

$$\begin{aligned} x_1 =& \frac{n^{n+1} \cdot K}{n+1} \\ x_2 =& \frac{J}{W} \end{aligned} \tag{2}$$

where K and n are respectively the hardening coefficient and exponent of the flow stress power law, J is the moment of inertia of the tube cross section and W is the depth of the specimen, in the direction of the movement of the punch, *i.e.* the lateral encumbrance of the structure. x_1 $[MPa]$ measures the fracture toughness of the skin material, because it represents the area under the flow stress curve. In x_2 $[mm^3]$, J is divided by W because the encumbrance could be limited by the space available inside the car door.

The optimization problem is to find the $\boldsymbol{x}$-values that maximize y, within the range of the investigated values of $\boldsymbol{x}$. The fusion metamodel and the optimization method described in [11] has been applied in order to find the optimal design in the framework of two different scenarios.

Description of the metamodeling approach

Suppose we have made a deterministic simulation that depends on q parameters at n_l design points, where the sub-index l is used to indicate that simulation provides the Lo-Fi data. A *first-stage* model can be built over the simulations results, by modeling the relationship between the response y_l and the design parameters. This function has usually a complex shape, so we use Gaussian Process model (GP) to describe its behavior. Gaussian Processes, known as Kriging in spatial statistics, are widely used to describe response of a computer code [14]. We assume that the response value y_l in a generic design point $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iq})' \in \mathbb{R}^q$ can be described by the relation:

$$y_l(\boldsymbol{x}_i) = \boldsymbol{f}_l'(\boldsymbol{x}_i)\boldsymbol{\beta} + \eta(\boldsymbol{x}_i), \quad i = 1, \ldots, n_l \tag{3}$$

where $\boldsymbol{f}_l'(\boldsymbol{x}_i)$ is the transpose of $\boldsymbol{f}_l(\boldsymbol{x}_i) \in \mathbb{R}^r$ that is a function of the design variable $\boldsymbol{x}_i$, $\boldsymbol{\beta} \in \mathbb{R}^r$ is a vector of unknown parameters, $\eta(\boldsymbol{x}_i) \sim \mathcal{GP}(0, \sigma^2_\eta, \boldsymbol{\vartheta}_l)$ is a GP with zero mean and variance-covariance matrix (called kernel) defined through the parameter σ^2_η and the vector of scale parameters $\boldsymbol{\vartheta}_l \in \mathbb{R}^q$. The core of the GP is the variance-covariance matrix; we use the "matern" covariance function defined as:

$$\begin{aligned} \mathbb{C}ov[\eta(\boldsymbol{x}_i), \eta(\boldsymbol{x}_j)] =& \sigma^2_\eta r_\eta(\boldsymbol{x}_i, \boldsymbol{x}_j) \\ =& \sigma^2_\eta \prod_{k=1}^{r} \frac{1}{\Gamma(\nu) 2^{\nu-1}} \left(\frac{2\sqrt{\nu}\,|x_{ik} - x_{jk}|}{\vartheta_k} \right)^{\nu} K_\nu \left(\frac{2\sqrt{\nu}\,|x_{ik} - x_{jk}|}{\vartheta_k} \right) \end{aligned} \tag{4}$$

where $\Gamma_\nu(\cdot)$ is the modified Bessel function of order ν and $\boldsymbol{\vartheta}_l$ the vector of smoothness parameters. In this report we set the parameter ν to $\frac{5}{2}$, so we have a twice differentiable process and

with this covariance function the correspondent correlation function is not ill conditioned [15]. If we set the parameter ν there are two unknown parameters to be estimated (ϑ_1 and ϑ_2).

Let $\boldsymbol{y}_l = \left(y_l(\boldsymbol{x}_1), y_l(\boldsymbol{x}_2), \ldots, y_l(\boldsymbol{x}_{n_l})\right)'$ represent the observed values and $\boldsymbol{F}_l = \begin{bmatrix} \boldsymbol{f}_l'(\boldsymbol{x}_1) \\ \boldsymbol{f}_l'(\boldsymbol{x}_2) \\ \vdots \\ \boldsymbol{f}_l'(\boldsymbol{x}_{n_l}) \end{bmatrix} \in \mathbb{R}^{n_l \times r}$ represent the model matrix. The estimation of all unknown GP parameters can be done by maximizing the logarithm of the restricted likelihood [16], given by:

$$l_{\hat{\boldsymbol{\beta}}} = -\frac{n_l - r}{2}\log 2\pi + \frac{1}{2}\log\left(\boldsymbol{F}_l'\boldsymbol{F}_l\right) - \frac{n_l - r}{2}\widehat{\sigma}_\eta^2 - \frac{1}{2}\log \mid \boldsymbol{R}_\eta \mid + \\ - \frac{1}{2}\log \mid \boldsymbol{F}_l'\boldsymbol{R}_\eta^{-1}\boldsymbol{F}_l \mid - \frac{n_l - r}{2} \tag{5}$$

where $\boldsymbol{R}_\eta = \{r_\eta(\boldsymbol{x}_i, \boldsymbol{x}_j)\}$ is the correlation matrix of the Lo-Fi points, $\widehat{\sigma}_\eta^2 = \frac{1}{n_l - r}\left(\boldsymbol{y}_l - \boldsymbol{F}_l\widehat{\boldsymbol{\beta}}\right)' \boldsymbol{R}_\eta^{-1} \left(\boldsymbol{y}_l - \boldsymbol{F}_l\widehat{\boldsymbol{\beta}}\right)$ is the restricted maximum likelihood estimator of σ_η^2 and $\hat{\boldsymbol{\beta}} = \left(\boldsymbol{F}_l'\boldsymbol{R}_\eta^{-1}\boldsymbol{F}_l\right)^{-1}\boldsymbol{F}_l'\boldsymbol{R}_\eta^{-1}\boldsymbol{y}_l$ is the maximum likelihood estimator of $\boldsymbol{\beta}$. We use a *quasi-Newton* algorithm to optimize the function (5), using the Matlab `fmincon` function.

Once all the GP parameters have been estimated, a prediction at any new design point $\boldsymbol{x}_0$ can be computed using the BLUP (*Best Linear Unbiased Predictor*) estimator (see for example [14] or [17]), defined as:

$$\widehat{y}_l(\boldsymbol{x}_0) = \boldsymbol{f}_l'(\boldsymbol{x}_0)\widehat{\boldsymbol{\beta}} + \widehat{\boldsymbol{r}}_\eta'\widehat{\boldsymbol{R}}_\eta^{-1}\left(\boldsymbol{y}_l - \boldsymbol{F}_l\widehat{\boldsymbol{\beta}}\right) \tag{6}$$

where $\boldsymbol{r}_\eta = \left(R(\boldsymbol{x}_1, \boldsymbol{x}_0),\ R(\boldsymbol{x}_2, \boldsymbol{x}_0), \ldots, R(\boldsymbol{x}_{n_l}, \boldsymbol{x}_0)\right)'$ is the correlation vector between the new design point $\boldsymbol{x}_0$ and all the other n_l points where the response was observed (i.e., simulated).

Predictions obtained via equation (6) are based on the Lo-Fi simulations only. However, in most practical cases experimental results are available, too. We will refer to these results as Hi-Fi y_h using the subscript h to distinguish them from the Lo-Fi y_l ones, since they are usually more accurate with respect to the simulations but available in limited number, so we can write $n_h \ll n_l$.

The main objective of the fusion metamodeling is to combine the Lo-Fi and Hi-Fi data in order to improve predictions achievable by using the Lo-Fi or the Hi-Fi data sets alone. The core of the data fusion model is a *linkage* or *second-stage* model, which represents the connection between Lo-Fi and Hi-Fi data and can be expressed as [18, 9]:

$$y_h(\boldsymbol{x}_i) = \rho(\boldsymbol{x}_i)\hat{y}_l(\boldsymbol{x}_i) + \delta_0 + \delta(\boldsymbol{x}_i) + \varepsilon_h(\boldsymbol{x}_i) \quad i = 1, \ldots, n_h \tag{7}$$

where the aim is to correct the Lo-Fi predictions $\hat{y}_l(\boldsymbol{x}_i)$ (i.e., predictions done using the Lo-Fi simulation only) using a "scale" and a "shift" effects, represented by $\rho(\boldsymbol{x}_i)$ and $\delta_0 + \delta(\boldsymbol{x}_i)$, respectively. We assume that the term $\varepsilon_h(\boldsymbol{x}_i)$ is the random independent error of Hi-Fi points, known also as the nugget, with $\mathbb{C}ov[\varepsilon_h(\boldsymbol{x}_i), \varepsilon_h(\boldsymbol{x}_j)] = 0 \quad \forall\ i \neq j$ and $\varepsilon_h(\boldsymbol{x}_i) \sim \mathcal{N}(0, \sigma_{\varepsilon_h}^2)$. We must add this random "nugget" effect because the response of a single experiment is not deterministic.

Following [9], the following scale effect is assumed

$$\rho(\boldsymbol{x}_i) = \boldsymbol{f}_h'(\boldsymbol{x}_i)\boldsymbol{\rho}, \tag{8}$$

where $\boldsymbol{f}_h(\boldsymbol{x}_i)$ is a function of $\boldsymbol{x}_i$ and $\boldsymbol{\rho}$ is a vector of unknown parameters to be estimated. Note that the subscript h is used to show that this function can be different from (3) used to

to model the Lo-Fi results. According to [9], a linear function is usually enough to represent the scale effect.

The shift effect is represented by $\delta_0 + \delta(\boldsymbol{x}_i)$, where δ_0 is a constant and $\delta(\boldsymbol{x}_i)$ is a GP:

$$\delta(\boldsymbol{x}_i) \sim \mathcal{GP}(0, \sigma_\delta^2, \boldsymbol{\vartheta}_h). \tag{9}$$

It is possible to prove [19] that only δ_0 can be estimated analytically.

According to the assumed combination of the linkage (or "second-stage") (7) and "first-stage" (3) models, the metamodel obtained via data fusion allows to predict a process realization at each new location $\boldsymbol{x}_0$ as:

$$\widehat{y}_h(\boldsymbol{x}_0) = \widehat{\rho}(\boldsymbol{x}_0)\widehat{y}_l(\boldsymbol{x}_0) + \widehat{\delta}_0 + \left(\widehat{\sigma}_\delta^2 \widehat{\boldsymbol{r}}_\delta + \widehat{\boldsymbol{\sigma}}_0\right)' \widehat{\boldsymbol{\Sigma}}_h^{-1} \left(\boldsymbol{y}_h - \widehat{\boldsymbol{P}}\widehat{\boldsymbol{y}}_l - \widehat{\delta}_0 \mathbf{1}_{n_h}\right) \tag{10}$$

where $\boldsymbol{r}_\delta = \mathbb{C}orr(\boldsymbol{y}_h, y_h(\boldsymbol{x}_0))$, $\boldsymbol{\sigma}_0$ is a vector with entries $\sigma_{0i} = \rho(\boldsymbol{x}_i) \cdot \rho(\boldsymbol{x}_0) \cdot \mathbb{C}ov(\widehat{y}_l(\boldsymbol{x}_i), \widehat{y}_l(\boldsymbol{x}_0))$ $\forall\, i = 1, \ldots, n_l$, $\boldsymbol{y}_h = (y_h(\boldsymbol{x}_1), y_h(\boldsymbol{x}_2), \ldots, y_h(\boldsymbol{x}_{n_h}))'$ are the observed Hi-Fi experimental values, $\boldsymbol{P}$ is a diagonal matrix with entries $\rho(\boldsymbol{x}_i), \quad i = 1, \ldots, n_h$ given in equation (8), $\boldsymbol{\Sigma}_h$ is the variance-covariance matrix of the model and $\mathbf{1}_{n_h} \in \mathbb{R}^{n_h}$ is a vector of ones.

Model comparison

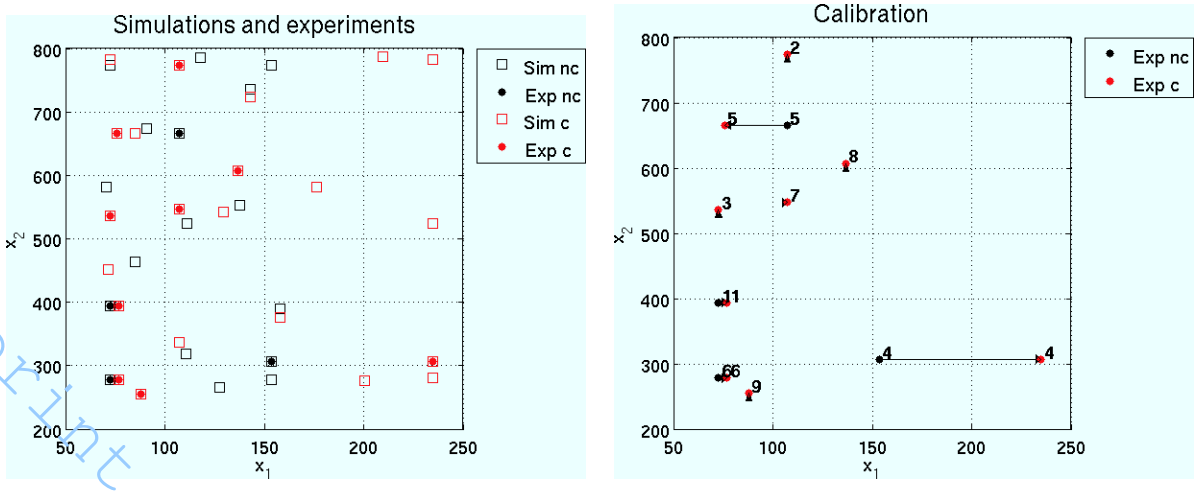
The purpose of this paper is to compare the predicting ability of the fusion model in different scenarios: in one case (scenario c, calibrated), the FEM simulations have been calibrated by tuning the material properties (K and n) of the tubular materials. In the other scenario (nc, non-calibrated), the K and n values have been calculated using the nominal properties of the materials. In Figure 2a all the available simulations (empty squares) and experiments (dots) are reported. Since 9 experimental design points are available, they have been replicated numerically in order to apply the fusion metamodels. 13 additional design points have been added according to a random, space filling design of computer simulations. The location of the design points along the x_1 axis of the simulations in the two scenarios (c, calibrated; nc, non-calibrated) are different because the calibration changes its values. Besides, the space-filling design has been planned independently for each scenario.

In Figure 2b the points of the experiment before (black) and after (red) the calibration are shown, with an associated number which is the running order of the experiments. It is possible to see that the size of the design of the non-calibrated scenario (black points) is almost half in the x_1 direction with respect to the design of the calibrated scenario (red points). As a consequence, the resulting metamodeled surfaces, built for each scenario, could be considerably different in shape from each other. Nevertheless, it is easy to compare the results provided by each metamodel at the 9 locations of the hi-fi (experimental) points and verify their accuracy. A leave-one-out cross validation method has been used: when predicting the y-value at any given location $\boldsymbol{x}_0$, the results of the experiments at that location are excluded for the estimation the metamodel parameters.

The predictions of metamodels built with calibrated and non-calibrated data are compared (at each of the 9 locations) to the average of the experimental values. The experimental values are averaged because every experiment has been replicated three times. In Figure 3 the absolute values of the prediction errors are plotted, where the prediction error is calculated as in Equation (11).

$$\widehat{\varepsilon}_{i,j} = \overline{y}_{i,exp} - \widehat{y}_{i,j}, \forall\, j \in \{sim_c, fus_c, fus_{nc}\} \tag{11}$$

where $\overline{y}_{i,exp}$ is the mean of the experimental values at location *i-th* and $\widehat{y}_{i,j}$ is the prediction at the location *i-th* with the *j-th* model. sim_c stands for calibrated simulations, fus_c for fusion model based calibrated simulations and fus_{nc} for fusion model with non-calibrated simulations. Needless to say, the FEM simulation process is deterministic.



(a) Design (b) Calibration effect

Fig. 2: Design of calibrated and non calibrated scenario

In Figure 3 the error of the two metamodels. The error of the "nc" metamodel is larger at locations 1 to 5 and smaller at locations 6 to 9. We can note that at location 4 both errors are quite large because that point is isolated from other hi-fi points.

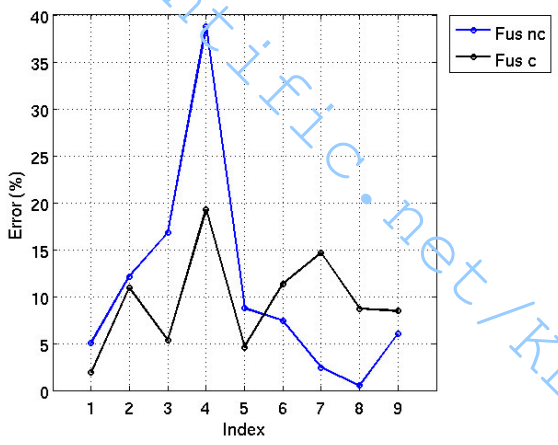
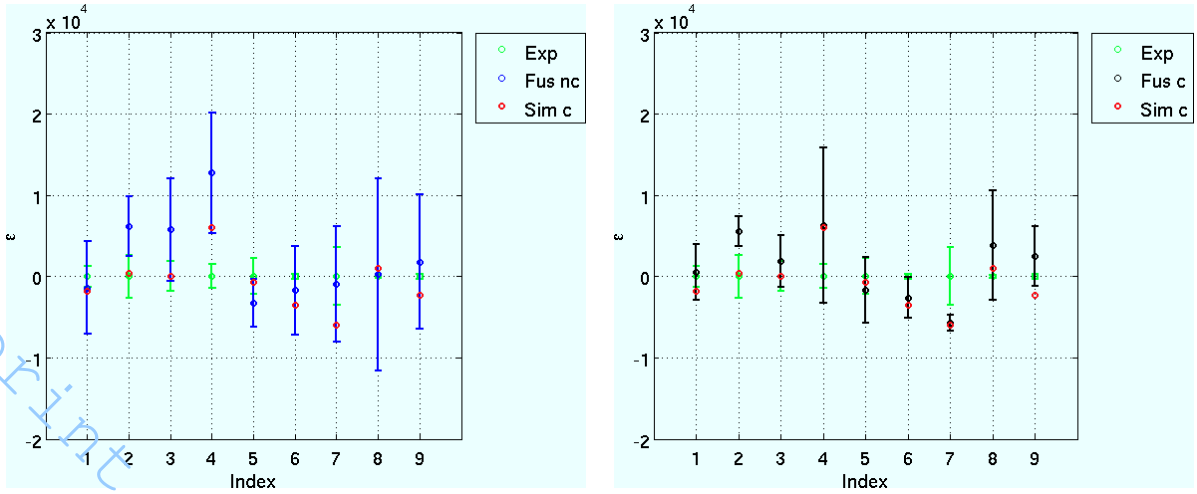


Fig. 3: Absolute value, $\left|\frac{\widehat{\varepsilon}_{i,j}}{\overline{y}_{i,exp}}\right|$

In Figure 4a we show the confidence interval of the experimental error (green) and the confidence interval of the error of the fusion model based on non-calibrated simulations (blue). Similarly, in Figure 4b we draw the confidence interval of the error of the fusion model based on calibrated simulations (black). We use a confidence level of 99% for each interval, and we make the assumption the data are normally distributed. We can appreciate from both figures where the uncertainty of the prediction of the metamodels is big, and consequently where it is possible to add points to improve the prediction. We also note that at location 4, where there is a big prediction error, the confidence interval on the mean value is consequently very large.

In Table 1 we report the p-value of the difference between each model and the experimental results. In order to compute these results we make the assumption that the predictions of the fusion models and the experiments are independent, otherwise it is impossible to compute the variance of the difference between the fusion model and the experiments (the correlation between the random variables is unknown).



(a) Non calibrated simulations (b) Calibrated simulations

Fig. 4: Confidence interval of the prediction error of the fusion models

We use an α-value 1% for each test. Since there are 9 different tests for each scenario, and using the Bonferroni's inequality the first type error for each scenario is less or equal to 9% (9α). At locations 2 and 4, the metamodel with non-calibrated data fails to provide reliable results (i.e. the p-value is smaller than 0.01). At locations 2 and 7, the metamodel with calibrated data fails. Each model fails at two different locations. There is a little but tolerable difference of performance between the two fusion models. For a further comparison, in the fourth column $delta_{fus}$ the p-value of the difference between the two metamodels is printed. Also in this test we make the assumption of independence between the two random variable involved. Looking at this column it is not possible to reject the null hypothesis of equality of the two model at all locations considered (all p-values are larger than 0.01). In other words, the two fusion models yield the same results, from a statistical point of view.

Index	fus_{nc}	fus_c	$delta_{fus}$
1	0.6441	0.7743	0.5728
2	0.0077	0.0006	0.7753
3	0.0853	0.3240	0.2757
4	0.0010	0.1995	0.2987
5	0.0777	0.4665	0.5408
6	0.5400	0.0445	0.7684
7	0.8139	0.0023	0.1894
8	0.9706	0.2702	0.6053
9	0.6751	0.1810	0.8755

Table 1: p-value of the difference between the models and the experiments

To verify that the choice of the variance-covariance matrix does not affect the results, we repeated the analysis with the powered exponential covariance function:

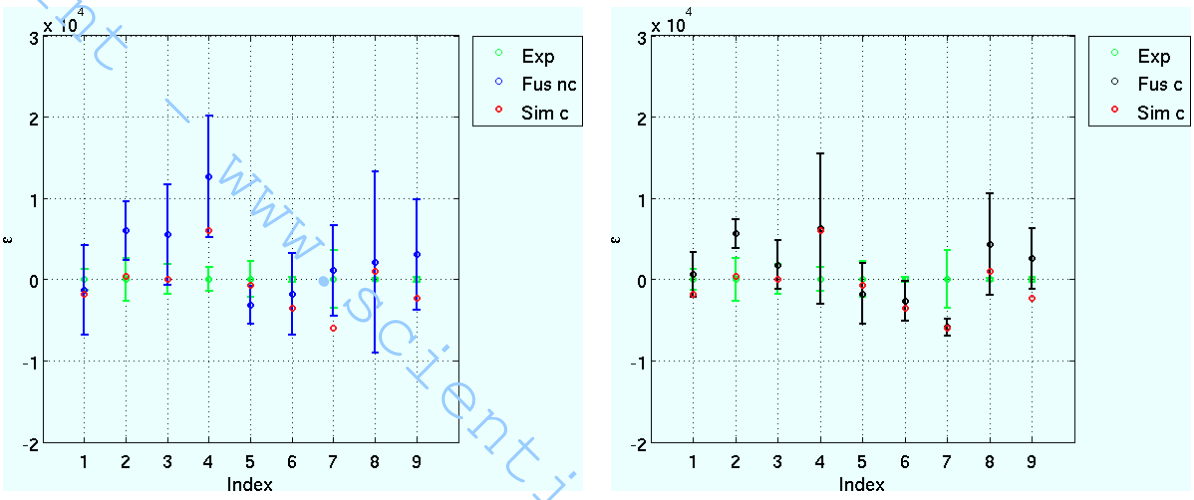
$$\mathbb{C}ov[\eta(\boldsymbol{x}_i), \eta(\boldsymbol{x}_j)] = \sigma_\eta^2 r_\eta(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sigma_\eta^2 \exp\left\{-d(\boldsymbol{x}_i, \boldsymbol{x}_j)\right\}. \tag{12}$$

where $r_\eta(\boldsymbol{x}_i, \boldsymbol{x}_j)$ is a correlation function that it is dependent on the distance between two distinct points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$. The weight distance is defined as:

$$d(\boldsymbol{x}_i, \boldsymbol{x}_j) = \sum_{k=1}^{q} \vartheta_k^l \left(x_{ik} - x_{jk}\right)^{p_k}. \tag{13}$$

The difference between this covariance function and the matern covariance function in Equation (4) is that the second one is twice differentiable, while the degree of smoothness of the powered exponential depend on the coefficient p_k and it is infinite if $p_k = 2$ or zero otherwise. Another difference is that with the powered exponential correlation function we have to estimate four parameters, while with the matern correlation function only two.

We report in Figure 5 the 99% confidence interval of the prediction error of the fusion models based on the powered exponential covariance function. It is possible to see that there is not any significant difference with Figure 4, that is the model have the same prediction performances. Similarly, Table 1 and 2 can be compared to verify that both correlation functions of the kriging metamodel are substantially equivalent.



(a) Non calibrated simulations (b) Calibrated simulations

Fig. 5: Confidence interval of the prediction error of the fusion models

Conclusion

In this paper we have shown that, in a specific test case, a hierarchical fusion model based on calibrated simulations yields the same prediction ability of a given response variable of a fusion metamodel built on non-calibrated simulations. In other words, when FEM simulations are run in order to optimize a specific response variable with a metamodeling approach, the need for extensive fine-tuning or calibration of the uncertain simulation parameters is reduced.

The paper has also shown that the uncertainty of prediction can be calculated (as in Fig. 4), and this helps improving the design space, suggesting where new hi-fi design points are required. In the tested case, two different correlation functions of the kriging hierarchical metamodel (matern and powered exponential) have been used, yielding very similar results.

References

[1] L.W. Friedman and I. Pressman. "The Metamodel in Simulation Analysis: Can It Be Trusted ?" In: *The Journal of the Operational Research Society* 39(10) (1988), pp. 939–948.

[2] B. Yu and K. Popplewell. "Metamodels in manufacturing: a review". In: *International Journal of Production Research* 32(4) (1994), pp. 787–796.

[3] L. Baghdasaryan et al. "Model Validation Via Uncertainty Propagation Using Response Surface Models". In: *Detc2002/dac-34140*. 2002, pp. 1–12.

[4] T.T. Do, L. Fourment, and M. Laroussi. “Sensitivity Analysis and Optimization Algorithms for 3D Forging Process Design”. In: *AIP Conference Proceedings 712*. 2004, pp. 2026–2031.
[5] H. Wiebenga, A.H. Boogaard, and G. Klaseboer. “Sequential robust optimization of a V bending process using numerical simulations”. In: *Structural and Multidisciplinary Optimization* 46 (2012), pp. 137–156.
[6] R. Hino, F. Yoshida, and V.V. Toropov. “Optimum blank design for sheet metal forming based on the interaction of high- and low-fidelity FE models”. In: *Archive of Applied Mechanics* 75 (2006), pp. 679–691.
[7] D. Huang et al. “Sequential kriging optimization using multiple-fidelity evaluations”. In: *Structural and Multidisciplinary Optimization* 32(5) (2006), pp. 369–382.
[8] G. Sun et al. “Multi-fidelity optimization for sheet metal forming process”. In: *Structural and Multidisciplinary Optimization* 44(1) (2010), pp. 111–124.
[9] Z. Qian et al. “Building Surrogate Models Based on Detailed and Approximate Simulations”. In: *ASME Journal of Mechanical Design* 128 (2006), pp. 668–677.
[10] E. Roux and P.O. Bouchard. “Kriging metamodel global optimization of clinching joining processes accounting for ductile damage”. In: *Journal of Materials Processing Technology* 213 (2013), pp. 10387–1047.
[11] B.M. Colosimo, L. Pagani, and M. Strano. “Metamodeling Based on the Fusion of FEM Simulations Results and Experimental Data”. In: *Key Engineering Materials*. Vol. 554 - 557. 2013, pp. 2487–2498.
[12] M. Strano, V. Mussi, and M. Menno. “Non-conventional technologies for the manufacturing of anti-intrusion bars”. In: *International Journal of Material Forming* 3 (2010), pp. 1111–1114.
[13] S.C.K. Yuen and G.N. Nurick. “The Energy-Absorbing Characteristics of Tubular Structures With Geometric and Material Modifications: An Overview”. In: *Transactions of the ASME: Applied Mechanics Reviews* 61 (2008), pp. 1–15.
[14] T.J. Santner, B.J. Williams, and W.I. Notz. *The Design and Analysis of Computer Experiments*. Springer Verlag, 2003.
[15] M. L. Stein. *Interpolation of Spatial Data: Some Theory for Kriging*. Springer Series in Statistics, 1999.
[16] D.A. Harville. “Maximum likelihood approaches to variance component estimation and to related problems”. In: *Journal of the American Statistical Association* 72 (1977), pp. 320–338.
[17] O. Shabenberger and C.A. Gotway. *Statistical Methods for Spatial Data Analysis*. Chapman & Hall/CRC, 2005.
[18] M.C. Kennedy and A. O’Hagan. “Predicting the Output from a Complex Computer Code When Fast Apporximations Are Available”. In: *Biometrika* 87 (2000), pp. 1–13.
[19] L. Pagani. “Multisensor Data Fusion for Quality Inspection of Free-Form Surfaces”. MA thesis. Politecnico di Milano, 2011.