Future Trends in Computer-Aided Engineering

J. Marczyk Ph.D., jacek@ontonix.com
Ontonix

Introduction

The paper provides a critical review of certain current practices in Computer Aided Engineering (CAE) as well as a glimpse into its possible future. It is argued that one of the main reasons why CAE is struggling lies in the use of surrogate models such as response surfaces. This appears to be unjustifiable in an era of cheap and fast computers. Surrogate models filter out a significant amount of physics and their use is justified exclusively on the basis of reduced CPU consumption. However, Finite Element (FE) models, which are used to fuel surrogate models via DOE schemes, continue to grow in size. This, in turn, leads to eternal bottlenecks. A further fundamental shortcoming of contemporary CAE is its inability to provide measures of credibility of FE models. In effect, the future of CAE hinges on credible computer models which, as we all know, are supposed to substantially reduce expensive physical testing. However, if the degree of trust in a given FE model (or any other virtual model) is unknown, using it to replace a physical test is clearly a risky undertaking.

The proposed future scenario for CAE is centered on the FE model, not a surrogate. It is argued that such models are capable of pushing CAE to the next level but only in a stochastic simulation perspective. Contemporary CAE is populated by a plethora of tools and methods which often force the engineer to see his problem from the perspective of numerical analysis or algorithms and not physics. There is the pressing need to use simpler and more “natural techniques” that help analysts focus on recognizable patterns, not repeatable details. The use of stochastic techniques will not only allow engineers to include uncertainty in FE models, making them significantly more realistic than surrogates, it will also help to naturally modulate and limit the number of finite elements, reducing IT-specific bottlenecks. Finally, the usage of stochastic techniques shall provide a gateway to a new methodology: complexity-based design. It is shown how the engineering of increasingly sophisticated products cannot proceed without taking complexity into account.

The Present

The fact that CAE is not delivering on its ambitious promises is evident. There are many facts that point to this quite clearly:

- The so-called up-front simulation, advocated for a long time, has not taken off.
- Models have not replaced physical tests to the extents that one would have expected a decade ago.
- Important bottlenecks in the CAE process continue to exist.
- When faced with extremely complex development programs, CAE is unable to cope and deliver solutions on time. In the case of large and ambitious programs, delays of sometimes years are not uncommon.

But why is this happening? After over twenty five years of CAE practice in diverse fields, we believe that the main reasons reside in the following:

- High Performance Computing has not helped turn stochastic simulation into an industrial solution. Most analyses are still deterministic.
- Even though the first multi-physics codes have appeared already fifteen years ago, the majority of the analyses are still single-physics. More than that, most analyses are still linear.
- The majority of today’s CAD models are not parametric. The generation of design options is still in most cases manual.
- The issue of model credibility is not taken seriously. Engineers are not concerned with quantifying the degree of trust of a model even though technology to do so already exists. Sheer model size is seen as guarantee of model validity. A large computer does not beatify a refined mesh.
- Response Surfaces – which are poor surrogates of reality – are used instead of full FE models, mainly to enable and justify the use of (multi-disciplinary) optimization.
- Optimization is used to pursuit the highest performance – this invariably leads to fragile designs.

- Many CAE practitioners forget the fundamental fact that models are only models. When in the right hands models can be extremely useful but, in general, they do not fully reflect the physics they should.

The alarming trend one witnesses is that models grow in size with the excuse (or desire) to gain fidelity and precision. This elevates the magnitude of the old “pre” and “post” bottlenecks without really adding substance. However, there is a natural limit to precision in all spheres of life. This includes engineering. In fact, not everything can be done with arbitrary precision. As Aristotle stated: “An educated mind is distinguished by the fact that it is content with that degree of accuracy which the nature of things permits, and by the fact that it does not seek exactness where only approximation is possible”. Along the same lines Bertrand Russell said that “You only realize the degree of vagueness of things until you try to make them precise”. It is therefore not a matter of gigaflops or gigahertz. There is a limit to precision and a finer mesh will not brake the barrier but simply throw the model below its natural noise-floor. Sophisticated multi-CPU hairsplitting will not extract more information than is naturally contained in a given phenomenon.

Let us take a quick look at what an established and popular CAE process known as (multi-disciplinary) optimization looks like and what it implies. For the purpose, let us start with the PDE Euler-Bernoulli equation of a vibrating beam and let us recall some of the assumptions that one must make in order to formulate it:

- The beam is long and slender
- The constraints are perfect
- Loads are applied far from the constraints
- The displacements are small
- Sections remain plain
- The material is linear and elastic
- The material is homogeneous
- Shear is neglected

At this point the equation is discretized via some finite element or finite difference scheme. How much physics has been lost at this point? 5%? 10%? Suppose now that one wishes to optimize the performance of the beam in some of its aspects. One typically proceeds along the following lines:

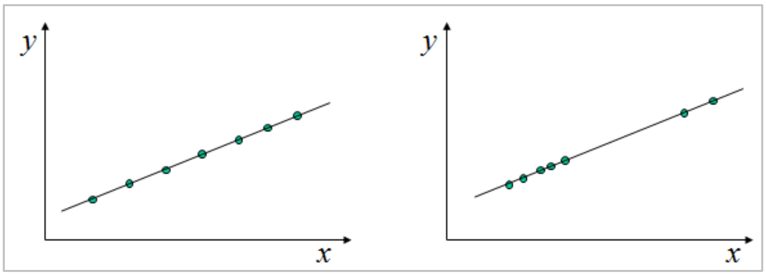
- Some design parameters are selected.
- A DOE (Design Of Experiments) scheme is used to sample the design space. This produces a table of numbers that has nothing to do with the physics of the underlying problem. In fact, this table can be applied to any problem, regardless if it is a CFD problem or a structural mechanics one. Sampling, in fact, is performed prior to having seen the response.
- Based on the results of the DOE sampling and on some regression scheme, a multi-variable Response Surface is built. The degree of the surface is established arbitrarily by the analyst. It is known that when compared to the original FE models Response Surfaces introduce approximations that can range from 5 to 10%.
- At this stage, the usually expensive FE (Finite Element) model is set aside and analysis proceeds based on the Response Surface. The result is that potentially pathological behavior is not spotted.
- An optimization technique (there are hundreds) is selected to find the optimum (global or local). It is known that different optimization techniques yield different results. An “optimum” solution is finally determined.
- A so-called “robustness analysis” is performed using Monte Carlo schemes, by sampling the Response Surface around the optimal design. The design variables are sampled randomly in certain ranges so as to incorporate tolerances into the process. A fundamental flaw that often goes unobserved in the process is related to the fact that input variables are often correlated and cannot be sampled independently. Neglecting this fact grossly overestimates the reliability of the real product.

What is wrong with this process? Actually quite a lot:

- Optimal products, optimal decisions or processes are known to be fragile. Engineers should seek acceptable solutions – this is the art of being an engineer, to quickly spot a solution that works – not pursue perfection. Optimality leaves no margin of error: *corruptio optimi pessima* as the Romans said. One reason behind the economic crisis is that the economy has been over-optimized: everything has been stretched to the limit. This can never be healthy.

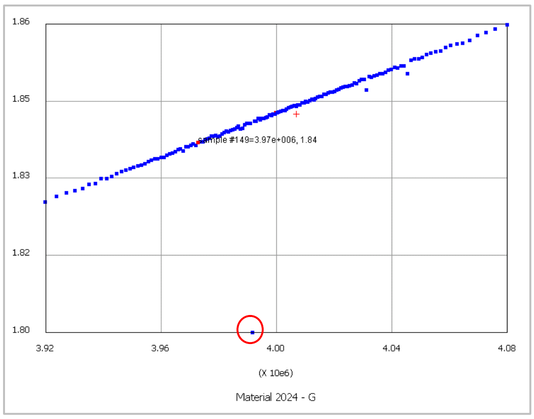
- A (FEM) model is only a model. A response surface, or any other surrogate, is a model of a model. A great industrial FEM model sometimes captures 80-90% of the physics of a real phenomenon. Such cases, however, are rare. But let's stick with the 90%. Let's suppose that a model indeed embraces 90% of reality. How credible is a 5% improvement in performance obtained with a digital model that misses 10% of reality?

- Response surfaces are constructed predominantly based on statistical methods. These ignore the fundamental fact that in cases such as the one indicated below, the two conditions are, from a physical standpoint, totally different.



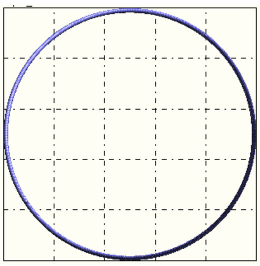
In fact, in both of the above cases, the correlation coefficient between x and y is 100%. In the case on the right, the regression line should not be used in the void between the two groups of points.

- Because Response Surfaces are built on a-priori established DOE tables, they are unable to embrace bifurcations (abundant in crash-type phenomena) or any other pathologies (of which Nature is full) such as outliers. Below is an eloquent case: linear eigenvalue analysis with a COTS FE tool – see the circled outlier.



What should be done with outliers like the one in the above figure? As Francis bacon sustained: "The best way to understand Nature is to understand her anomalies". But a statistician would throw the outlier away so as to enable the use of regression. This is called bending reality to suit one's tools.

- Correlation techniques suffer another fundamental flaw. If one has data points arranged, for example, on a circle, correlation is 0. However, a relationship like the one depicted below is certainly a significant one and surely involves plenty of physics. In other words, correlations are not a good way of learning from data – they can miss very important patterns.



Response surfaces are poor caricatures of reality. Response surfaces are, however, used extensively in other fields, such as economics. In fact, we believe that the use of response surface-based models is one of the reasons why rating agencies and risk managers got is so terribly wrong when evaluating companies and risk exposures, leading, ultimately to the crisis. However, the most surprising thing is that the usage of response surfaces in CAE implies first the synthesis of multi-million degree-of-freedom FE models, only to later project them onto over-simplified surrogates. How wasteful! Just like the Response Surface Method has contributed to the economic crisis, it will continue to damage CAE.

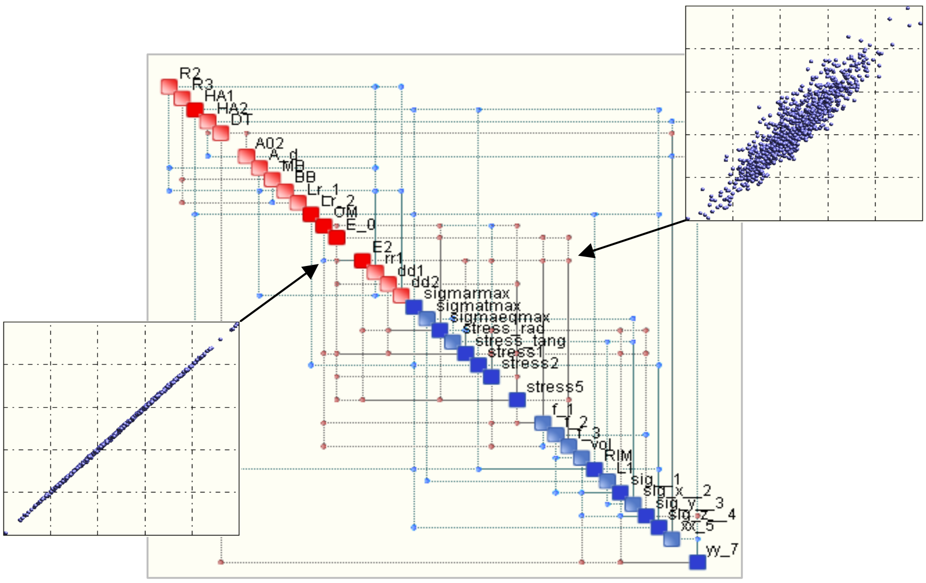
The Future

It is not easy to make predictions. However, there are trends in CAE which are opening new avenues and creating new opportunities for those who are willing to innovate. The recently developed complexity quantification techniques are making it possible for the first time to rationally and quantitatively approach:

- Knowledge capture: how to transform data to knowledge. Especially in a multi-discipline setting.
- Model validation: how to actually measure the degree of credibility of a computer model.
- “Automatic CAD”: how to come up with the simplest solution to a design problem.

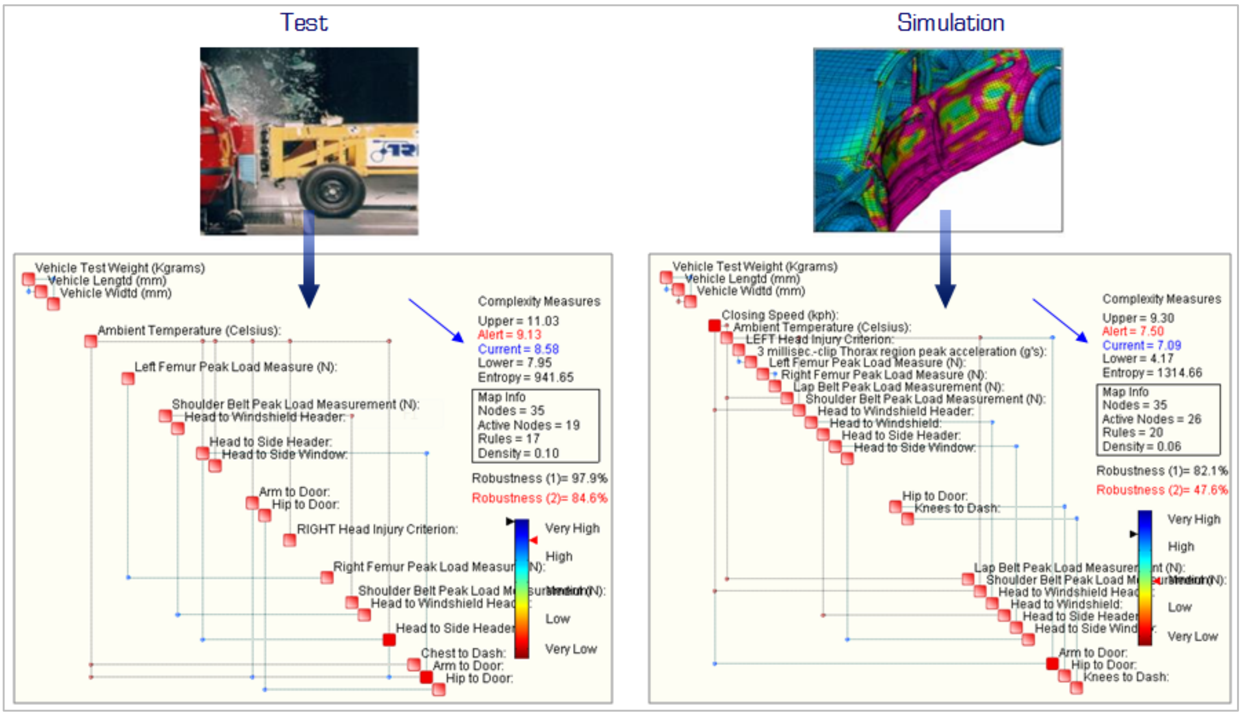
We believe the first two are vital towards the credibility and future of CAE. Complexity quantifies the “degree of sophistication and uncertainty” of a system. In information theory language we could say that complexity measures the “amount of structured information”. Complexity is a quantity that is intimately linked to and establishes an elegant bridge between the concepts of entropy and structure. In a Universe in which there is an overwhelming tendency to decay, how can structure possibly exist? The answer to this question, which has troubled science for decades, lies in complexity. However, in the present paper we will limit ourselves to the two definitions outlined above, as they are sufficient to carry across the necessary ideas.

What is knowledge? What is a body of knowledge? When related to a system or phenomenon, knowledge can be defined as a dynamic set of inter-related fuzzy rules of the type “if A then B”. The fact the rules are inter-related (for example “if it rains then drive slowly” and “if it rains then start windscreen wipers” and “if it rains then increase the distance to the car in front”) means that together they form structure. Rules can be more or less strong, or crisp. The strength of a rule may be measured in different ways. Most importantly, rules may be arranged in groups or maps, such as the one illustrated below. The case in questions relates input (red nodes) variables to performance descriptors (blue nodes) of a mechanical system. Two rules – indicated by means of red and blue connectors - are also illustrated in the figure below.



The rule on the left resembles a straight line and is, evidently, more crisp. The other rule is cigar-shape and is more “fuzzy” in that for each value of “A” the resulting “B” is not clearly defined. Nevertheless, the rule does reflect an evident structure in the nature of the relationship between the two parameters (Young’s modulus and third natural frequency). Maps such as the one above are created automatically by postprocessing the results of Monte Carlo Simulation or, in the case of time-domain simulations, of a single computer run or physical experiment. The value of such maps (known as Process or Knowledge Maps) is phenomenal: not only do they portray the structure of information flow within a given system; they also indicate critical parameters (known as hubs), help measure the degree of redundancy or fail-safe of a design and, most importantly, allow one to measure the complexity of the system in question. Realizing which parameters influence which ones and in what measure is, when seen in a global context, equivalent to having knowledge of that particular system or phenomenon.

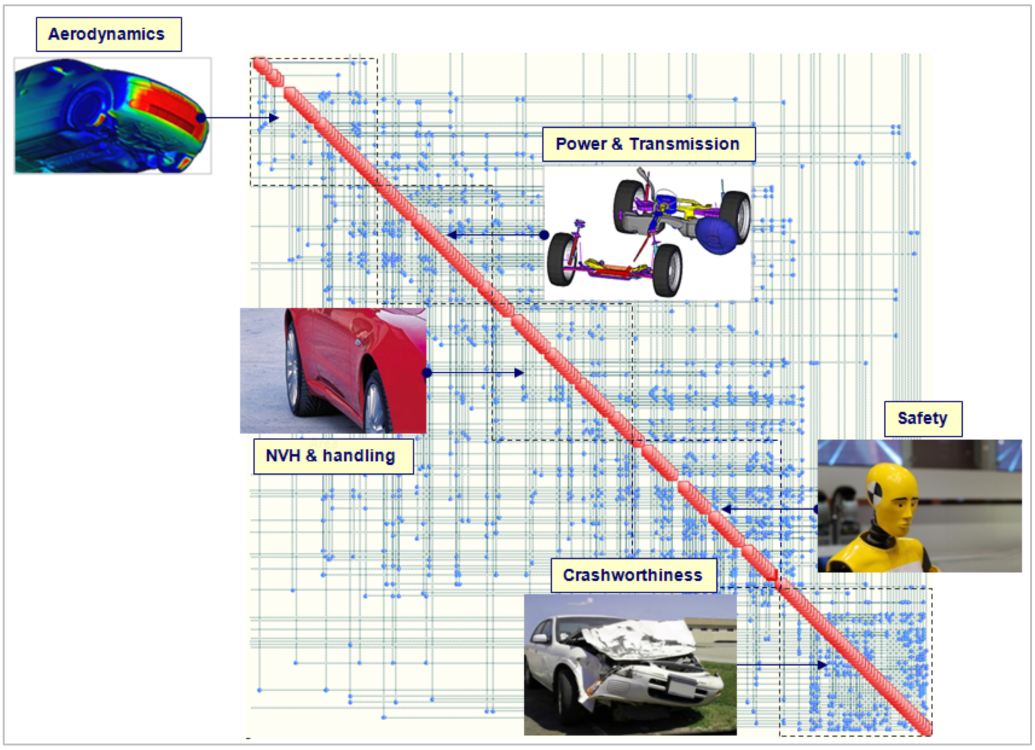
Obtaining knowledge is simple. It is sufficient to organize existing data, of which CAE produces phenomenal amounts on a daily basis, and to process it using tools such as OntoSpace™. The process is illustrated below both for physical testing as well as computer analysis.



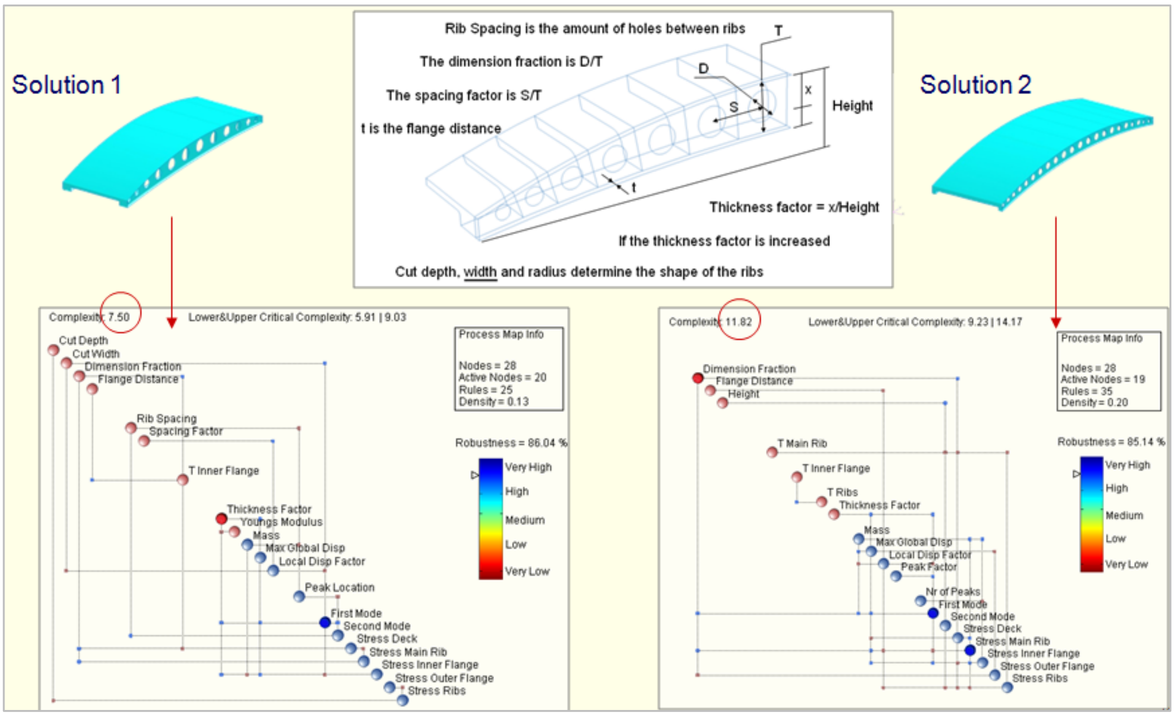
The above image prompts the idea to use Process Maps to measure the "distance between a computer model and the experiment" it is supposed to emulate. The concept is simple: if the model correctly captures the physics of the experiment, then the complexities of both maps (this is the weak condition) as well as the topologies of the maps (this is the strong condition) should be identical (or very similar). The degree to which this holds can easily be measured – we finally have a physics-based metric of model credibility. Today, the degree of "model-test correlation" is measured via a Euclidean norm of a vector whose entries are the differences of measured and computed values (e.g. frequencies or displacements). The first applications of this new technology are taking place in the automotive industry.

Another issue equally crucial to the fate of CAE is the linking of disjoint disciplines and providing engineers and designers with a holistic system-like view of things. When complexity of a design reaches certain levels, a systems approach is a must: highly complex systems often fail because of un-thought of interactions between apparently "distant" components or disciplines. Understanding a system from a global standpoint will be a driving trend not only in engineering but also in economics and medicine. Science has suffered in the past few decades precisely because of a deterministic obsession with perfection and a maniac pursuit of details, while losing the global picture. An example, again from the auto industry, is illustrated below.

The structure of the global Process Map below reveals not only the interactions between the various aspects of a sophisticated product, it also measures the degree of couplings and allows the team to approach any desired design changes is a more natural and collaborative fashion. In other words, reaching compromises is easier and these are always necessary. Collaborative design doesn't mean just connecting dispersed teams using the internet.



This brings us to the last point: CAD. Should parametric models ever become mainstream, complexity technology could allow us to envision a new way of doing CAD. The concept is illustrated below. You generate multiple solutions to a given CAD problem via random perturbations of a parametric CAD model – this is done using COTS Monte Carlo Simulation tools. Out of the multiple solutions the ones that fulfill objectives and constraints are selected. Out of these the one with the lowest complexity is chosen as the final design: the best solution is the simplest that works. Two compatible solutions to a CAD problem are shown below: the one on the left has complexity 7.5, the other around 12. Both have the same performance.



Conclusions

The manufacturing industry is facing tremendous problems when facing huge development programs and highly sophisticated products. Failures and delays in such cases cause immediate losses. High product complexity inevitably leads to liability costs (recalls, warranty, law suits). This, in effect, is logical because conventional design neglects complexity as a design attribute. Complexity will have to become a central design attribute of new sophisticated products and multi-disciplinary product development processes. This can already happen today at the CAD level, providing that parametric models are used.

CAE has been fairly fast at adopting High Performance Computing technology and notoriously creates a huge amount of data. However, this data is processed using traditional techniques, more focused on producing images than on extracting knowledge. The highest degree of “compression” and synthesis is when one passes from data to knowledge. But data compression *per se* is of course not the objective – the goal is to better understand highly interconnected systems which embrace multiple disciplines so as to enable faster design and to avoid failures. Data contains hidden structure which can be easily extracted providing that statistical methods or other conventional methods are not used for the purpose.

A fundamental necessary step forward that CAE needs to take is to abandon the usage of surrogates. In the age of cheap parallel computers, their usage is unjustified. Models have to be realistic, not precise but precision, unfortunately, is still the focus. This is evident in the pursuit of optimal solutions. It can be shown on analytical grounds - even though it is also intuitively obvious – that optimal solutions are fragile. When highly sophisticated product development is faced, one needs to incorporate complexity, not optimality, into the design loop. Highly complex systems simply cannot be optimized but there needs to be a limit to their complexity. This is because excessive complexity is a prelude to fragility – think of the global financial system. So, the future design paradigm needs to be focused not on maximizing performance or efficiency but on limiting complexity. This will be good enough.

Unfortunately, contemporary CAE practitioners are convinced that the “pure CAE” issue is (almost) solved and have focused their attention on such aspects as PLM, simulation data management or collaborative development. This is all necessary technology but where CAE has to do most of its homework is at solver level and at guaranteeing that a digital model, which is supposed to replace a physical test, is credible and that its level of credibility is known. With its current structure, CAE is not equipped to face the challenges of the future.

References

[1] “Principles of Simulation-based CAE”, J. Marczyk, FIM Publications, Madrid, Spain, 1999.

[2] “Beyond Optimization in CAE”, J. Marczyk, CIMNE, Barcelona, Spain, 2002.

[3] “Complexity Management: New Perspectives and Challenges for CAE in the 21-st Century”, J. Marczyk, BenchMark Magazine, NAFEMS, July 2008.

[4] “Practical Complexity Management”, J. Marczyk, Editrice Uni-Service, Trento, Italy, 2009.