CSMA 2019
14ème Colloque National en Calcul des Structures
13-17 Mai 2019, Presqu'île de Giens (Var)

Virtual, Digital and Hybrid Twins: A new paradigm in data-based engineering and engineered data

[1] *ENSAM ParisTech, 151 rue de l'Hôpital. 75013 Paris, France, Francisco.Chinesta@ensam.eu*
[2] *I3A, Universidad de Zaragoza, Maria de Luna s/n, 50018 Zaragoza, Spain, ecueto@unizar.es*
[3] *ESI GROUP Chair. ECN, 1 rue de la Noe, F-44300 Nantes, France, Emmanuelle.Abisset-chavanne@ec-nantes.fr*
[4] *ESI Group, 3bis rue Saarinen, 94528 Rungis CEDEX, France, {Jean-Louis.Duval;Fouad.El-Khaldi}@esi-group.com*

Résumé — Engineering is evolving in the same way than society is doing. Nowadays, data is acquiring a prominence never imagined. In the past, in the domain of materials, processes and structures, testing machines allowed extract data that served in turn to calibrate state-of-the-art models. Some calibration procedures were even integrated within these testing machines. Thus, once the model had been calibrated, computer simulation takes place. However, data can offer much more than a simple state-of-the-art model calibration, and not only from its simple statistical analysis, but from the modeling and simulation viewpoints. This gives rise to the the family of so-called *twins* : the virtual, the digital and the hybrid twins. Moreover, as discussed in the present paper, not only data serve to enrich physically-based models. These could allow us to perform a tremendous leap forward, by replacing big-data-based habits by the incipient smart-data paradigm.
Mots clés — Virtual twin, Digital Twin, Hybrid Twin, Data-Driven Engineering Science, Machine Learning, Big-Data, Smart-Data.

1 Introduction

As models involved in science and engineering become too complex, their analytical solution is often compromised. On the other hand, computers are able to perform very efficiently only elementary operations. Consequently, it is necessary to transform complex mathematical objects (derivatives, ...) into simpler objects, i.e., elementary operations. At the same time, it is necessary to reduce the number of points and time instants at which the solution of the model is evaluated, by replacing the continuum by a discrete system, treatable by digital computers. Such a procedure is known as numerical simulation and constitutes one of the three pillars of 20th century engineering—modeling and experiments being the other two pillars—. This age has been coined as the *third paradigm* of science.

In the previous (third) industrial revolution, "virtual twins" (emulating a physical system by one, or more, mathematical models to describe its complex behavior) were major protagonists [1]. Nowadays, numerical simulation is present in most scientific fields and engineering domains, making accurate designs and virtual evaluation of systems responses possible—drastically cutting the number of experimental tests.

The usual numerical model in engineering practice (which we will denote here as virtual twin) is something static. This kind of (finite element, finite volume, finite difference) models is nowadays ubiquitous in the design of complex engineering systems and their components. We say that they are static because they are not expected to be continuously fed by data so as to assimilate them. This would be what is today understood as a Dynamic Data-Driven Application System (DDDAS). The characteristic time of standard engineering simulation strategies can not accommodate the stringent real-time constraints posed by DDDAS, specially for control purposes. Real-time simulation for control is typically ensured by techniques based on the use of *ad hoc*—or black box—models of the system (in the sense that they relate some inputs to some outputs, encapsulated into a transfer function). This adapted representation of the system allows proceeding in real-time. However, it becomes too coarse when compared with rich,

1. There seems to be no consensus on the definition of the concepts of virtual, digital and hybrid twins. In this paper we suggest one possible distinction, that seems feasible, attending to their respective characteristics. It is not the sole possibility, of course, nor do we pretend to create any controversy on it.

high fidelity simulations, such as the ones performed using, for example, Finite Element techniques.

Although science was preeminently data-based at the early years (think of Tycho Brahe, for instance), it was at the end of the 20th century that data irrupted massively in most scientific fields, and in particular in the one we are specially interested in : engineering. For many years data have been widely incorporated to usual practice in many disciplines where models were more scarce or less reliable—with respect to engineering sciences—. Thus, massive data were classified, visualized (despite its frequent multi-dimensional nature), curated, analyzed, ... thanks to the powerful techniques recently developed in the wide areas of artificial intelligence and machine learning. When correlations are removed from data, a certain simplicity emerges from the apparent complexity, as proved by advanced nonlinear dimensionality reduction techniques based on manifold learning. Moreover, a number of techniques were proposed to establish the relations between outputs of interest and certain inputs, assumed to be sufficient to explain and infer the outputs. These are the so-called "model learners", based on the use of linear and nonlinear regressions, decision trees, random forests, neural networks—inevitably linked to deep-learning techniques—, among many others.

The solution of physically-based models, very well established and largely validated in the last century, was partially or totally replaced by these data-based models, due to their computational complexity. This is especially true for applications requiring real-time feedback. Thus, massively collected and adequately curated data, as just discussed, provided interpretation keys to advise on an imminent fortuitous event. This makes possible improved data-based predictive maintenance, efficient inspection and control, ... that is, allows for real-time decision making. The price to pay is an as rich as possible learning stage. This takes considerable time and efforts, as the establishment of validated models took in the previous engineering revolution.

Important success was reported, many possibilities imagined, ... justifying the exponential increase in popularity of these "digital twins". There has been a fast development of data-driven models for representing a system with all its richness while ensuring real-time enquiries to its governing model. However, replacing the rich history of engineering sciences (that proved their potential during more than a century with spectacular successes) led to feelings of bitterness and of waste of acquired knowledge.

This new incipient engineering consists of "virtual twins", that operate offline in the design stage, and their digital counterparts, based on data, taking over in online operations. However, the domain of applicability of the last, even if they are superior in what concerns their rate of response, continues to be narrower. A combination of both, the "virtual" and "digital" twins seems to be the most appealing solution. However, prior to combining both, a major difficulty must be solved : the real-time solution of physically-based models.

All the just introduced problems can not be overcome by simply employing more powerful computers—in other words, by employing modern supercomputing facilities—. Even when this is a valuable route, it strongly limits the accessibility to the appropriate simulation infrastructure. This is true also in what concerns to its integration in deployed platforms. Recent history has proved that this is a prohibitive factor for small and medium-sized companies. An effort must be paid towards the to democratization of simulation.

Again, it was at the end of the past century and the beginning of the 21st century, that major scientific accomplishments in theoretical and applied mathematics, applied mechanics, and computer sciences contributed to new modeling and simulation procedures. Model Order Reduction (MOR) techniques were one of these major achievements [1]. These techniques do not proceed by simplifying the model, models continue to be well established and validated descriptions of the physics at hand. Instead, they rely on an adequate approximation of the solution that allows simplifying the solution procedure without any sacrifice on the model solution accuracy, in view of accommodating real-time constraints.

A feasible alternative within the MOR framework consists of extracting offline the most significant modes involved in the model solution, and then projecting the solution of similar problems in that reduced space. Consequently, a discrete problem of very small size must be solved at each iteration or time step. Thus, MOR-based discretization techniques provide significant savings in computing-time. Another MOR-based route consists of computing offline, using all the needed computational resources and computing time, a parametric solution that contains the solution of all possible scenarios. This parametric solution can then be online particularized to any scenario using deployed computational facilities, including tablets or even smartphones. It allows then to perform efficient simulation, optimization, inverse analysis, uncertainty propagation and simulation-based control, all under real-time constraints. Such a solution has been demonstrated on many applications where the Proper Generalized Decomposition (PGD) method is used [2, 3].

The next generation of twins was born, the so-called “hybrid twin$^{\text{TM}}$", that combines physically-based models within a MOR framework (for accommodating real-time feedback) and data-science.

On one hand, real-time solution of physically-based models allows us to assimilate data collected from physical sensors, to calibrate them. Therefore, it also exhibits predictive capabilities to anticipate actions. Thus, simulation-based control was made possible, and successfully implemented in many applications, often by using deployed computing devices (e.g., Programmable Logic Controllers). Despite an initial euphoric and jubilant period in which high-fidelity models were exploited in almost real-time by using standard computing platforms, unexpected difficulties appeared as soon as they were integrated into data-driven application systems.

Significant deviations between the predicted and observed responses have been detected, nevertheless, by following this approach. The origin of these deviations between predictions and measurements can be attributed to inaccuracy in the employed models, in parameters or in their time evolution. These often continue to be crude descriptions of the actual systems. Attacking this ignorance can done by developing on-the-fly data-driven models that could eventually correct this deviation between data and model predictions.

Indeed, a DDDAS consists of three main ingredients : (i) a simulation core able to solve complex mathematical problems representing physical models under real-time constraints ; (ii) advanced strategies able to proceed with data-assimilation, data-curation and data-driven modeling ; and (iii) a mechanism to online adapt the model to evolving environments (control). The Hybrid Twin$^{\text{TM}}$[4] embraces these three functionalities into a new paradigm within simulation-based engineering sciences (SBES).

2 From Virtual to Hybrid Twins

A given physical system is characterized by a number of continuous or discrete variables. In general, to manipulate these variables in a computer, continuous variables are discretized, i.e., more than looking for those variable at any point, it is assumed that variables at any point can be expressed from the ones existing in some particular locations (the nodes, if we employ the finite element terminology) by using adequate interpolations.

In what follows the discrete form of the variables defining the system state at time t is denoted by $\mathbf{X}(t)$. As just indicated, they could include, depending on the considered physics, nodal temperatures, velocities, displacements, stresses, etc.

The system evolution is described by its state $\mathbf{X}(t)$, evolving from its initial state at the initial time $t = t_0 = 0$, denoted by $\mathbf{X}_0$. Numerical models based on well established physics allow making this prediction of the system state at time t from the knowledge of it at the initial time t_0, by integrating its rate of change (coming from the physical laws adequately discretized) given by $\dot{\mathbf{X}}(\tau)$ at $\tau \in (0,t]$.

This contribution will be expressed by $\dot{\mathbf{X}}(t;\mu) = \mathbf{A}(\mathbf{X},t;\mu)$—we emphasize its parametric form—, where μ represents the set of involved parameters that should be identified offline or online.

Remark 1. In the previous expression the semicolon $(\cdot;\cdot)$ makes a distinction between the coordinates before the semicolon—in this case, time—and the model parameters after it—here, μ.

Thus, if we assume a model to accurately represent the subjacent physics involved in the system, predictions are easily performed by integrating $\mathbf{A}(\mathbf{X},t;\mu)$. Here, if real-time feedbacks are needed, standard integration (based on the use of well experienced numerical techniques like finite elements, finite differences, finite volumes, spectral methods, meshless (or meshfree) techniques, ...) of the dynamical system expressed by $\mathbf{A}(\mathbf{X},t;\mu)$, turns out to be unsatisfactory. As previously discussed, the employ of model reduction techniques opened new routes in this sense. In particular, the Proper Generalized Decomposition—PGD—precomputes (offline) the parametric solution, thus making possible to accommodate real-time constraints.

When model calibration is performed online, model parameters μ are calculated by enforcing that the associated model prediction fits as much as possible to the experimental measurements, at least at the measurement points. In the context of process or system control, external actions can be applied to drive the model towards the given target. Thus, the state rate of change (if we neglect noise for the moment) is composed by two terms,

$$\dot{\mathbf{X}}(t;\mu) = \mathbf{A}(\mathbf{X},t;\mu) + \mathbf{C}(t), \tag{1}$$

that expresses the physical and forced (external goal-oriented actions) contributions, **A** and **C**, respectively.

Remark 2. In general, control actions, here represented by the term **C** could depend on measures and/or on the inferred model parameters, but here, and without loss of generality, we only indicate explicitly its dependence on time.

2.1 Model updating

When models represent the associated physics poorly, a non negligible deviation between their predictions and the actual evolution of the system, acquired from collected data, is expected. This deviation is expected to be biased, because it represents the modeler's ignorance on the subjacent physics. The unbiased deviation contribution is associated to modeling or measurement noise and is easily addressed by using adequate filters. However, biased deviations express hidden physics and required a particular treatment, that is, their online modeling by assimilating collected data.

Indeed, the deviation (gap between the model prediction $\mathbf{X}(t;\mu)$ and measurements $\mathbf{X}^{\exp}(t)$) when considering the optimal choice of the model parameters μ, and, more precisely its time derivative, should be used for the online construction (under the already mentioned severe real-time constraints) of the so-called data-based correction model. This correction, also referred as deviation model, is here denoted by $\mathbf{B}(\mathbf{X},t)$. Even if, in what follows, the presence of unbiased noise is ignored, its inclusion is straightforward.

Thus, the fundamental equation governing a hybrid twins writes

$$\dot{\mathbf{X}}(t;\mu) = \mathbf{A}(\mathbf{X},t;\mu) + \mathbf{B}(\mathbf{X},t) + \mathbf{C}(t) + \mathbf{R}(t), \tag{2}$$

expressing that the rate of change of the system state at time t contains four main contributions :

1. the model contribution, whose rate of change related to the model parameters μ reads $\mathbf{A}(\mathbf{X},t;\mu)$. MOR is crucial at this point to ensure real-time feedback ;
2. a data-based model $\mathbf{B}(\mathbf{X},t)$ describing the gap between prediction and measurement ;
3. external actions $\mathbf{C}(t)$ introduced into the system dynamics in order to drive the model solution towards the desired target. It also includes any other kind of decision based on the collected and analyzed data ;
4. the unbiased noise $\mathbf{R}(t)$, that has been traditionally addressed using appropriate filters. This term also includes external actions for which there is no possible prediction. Typically, human intervention on the system.

Here we have omitted a very important distinction, the necessity of collecting appropriate data with different aims : (i) to calibrate the considered physically-based model, assumed to represent the first-order contribution to predictions and for explicative purposes ; (ii) to construct on-the-fly the data-driven model update ; and (iii) for decision-making proposes (control) by using data-analytics on the collected data.

It is also worth noting that the better locations and frequency of acquisition for collecting data could differ, given the volume of data to treat and data-assimilation rate, depending on the purpose : calibration, modeling and control. In the present framework, the model could help to infer the *smartest* data to acquire, and when and where they should be collected. Thus, Big Data could be replaced by Smart Data in the framework of a new multi-scale data science and theory of information, bridging the gap between data (microscopic), information (mesoscopic) and knowledge (macroscopic).

The construction of the data-based model deserves some additional comments :

1. Deviations inform us about the model possibly becoming inaccurate. In our approach, model updating is based on the deviation model, and then, it is added to the first-order model when it exists. Other authors suggested to update the model itself within a stochastic framework.
2. In some cases the first order, physically-based model, **A**, does not exist, or simply it is ignored as was the case in most digital twins, motivated by difficulties related to its real-time solution, to its accuracy, etc ... In this case, when constructed from scratch, many data are required to reach a sufficient accuracy. However, when the data-driven model is only expected to fill the gap between the first-order model predictions and the measurements, the higher is the model accuracy, the smaller the data-driven contribution, implying that the required volume of data significantly decreases. It is worth mentioning that collecting data and processing them is expensive, and could compromise the real-time constraints.
3. The recent exponential growing of machine learning techniques (data-mining, deep-learning, manifold learning, linear and nonlinear regression techniques ... to cite but a few) makes it possible to construct on-the-fly such a data-based deviation model ;
4. Another possibility consists of expressing the deviation within a parametric form within the PGD framework. To that end, a sparse-PGD is constructed—here viewed as an advanced powerful nonlinear regression technique—, operating on the deviations. These deviations are the difference between the physically-based prediction and the measurements. The main advantage of this procedure is that the parametric expression of the deviation can be added to the expression of the model based on the known physics, **A**, that was already expressed using the same format (parametric PGD separated representation).

   Thus, the resulting solution contains some modes coming from the discretization of the equations representing the known physics, while the remaining ones are associated to the detected deviations. In any case, both together represent the actual system that contains hidden physical mechanisms, more complex that the ones retained in the first-order model, pragmatically captured even when ignoring its real nature.

   If the real solution evolves in a manifold, its projection on the manifold defined by the physical model, **A**, allows computing the best choice for the involved parameters, i.e., μ (calibration). The orthogonal complement represents the deviation model. All of them, the real, the physical and the deviation models can be cast into a parametric PGD separated representation form.

Remark 3. In the previous expression, Eq. (2), the data-based contribution $\mathbf{B}(\mathbf{X},t)$ justifies the “hybrid" appellation, because the model is composed of two contributions, one coming from well established and validated physics, the other based on data. This double nature makes the difference between usual digital twins and their hybrid counterparts.

Remark 4. When enriching a dynamical system with a data-based contribution, before reaching a sufficient accuracy, a stable system can become unstable, thus compromising long-time predictions. In this case the control term could encompass a numerical stabilization to ensure that the enriched dynamical systems remains stable.

Remark 5. Deep learning, based on the use of deep neural networks, allows impressive accomplishments. However, it generates nowadays a certain frustration in a scientific community that for centuries tried to explain reality through models. That aim is almost lost when using deep learning. Even if many efforts are being paid with the purpose of explaining these machine learning techniques, today their impressive performance is not fully understood. However, within the hybrid twin rationale, things become less uncomfortable, since these techniques, whose predictions are difficult to explain, are being used to model a physics that escapes to our understanding, what we have called ignorance.

3 Perspectives

The most complete member of the twin family involves many different methodologies, in particular :

1. Real-time simulation based on Model Order Reduction ;
2. Real-time calibration ;
3. Real-time data-assimilation and data completion ;
4. Real-time data-analytics ;
5. Real-time data-driven modeling.

The previous requirements were deeply addressed in the recent review [4]. The hybrid twin, that perfectly encompasses the functionalities of its two predecessors, the so-called virtual and digital twins, consists of :

1. the pre-assumed physical contribution, efficiently addressed by using Model Order Reduction techniques ;
2. a data-based modeling of the gap between predictions and measurements ;
3. external actions to drive the model solution towards the desired target (control and decision making) ;
4. the unbiased noise filtering ;

where sufficient data is required with three main aims : (i) to calibrate the physical model ; (ii) to construct the data-based model ; and (iii) to make decisions to keep the system under control and progressing to the wished target.

Control and decision making is efficiently performed by using artificial intelligence and machine learning techniques, as soon as the learning state is successfully accomplished. On the other hand, the data-based model construction can be performed :

— from the use of machine learning techniques (data-mining, regression, deep-learning, manifold learning, ... as previously described) ;
— by expressing the deviation in a parametric form within the PGD framework by using the regression PGD —rPGD— discussed before. In this framework, data-science could be used offline to define the smartest data so be considered, and in particular, what data, and when and where they should be collected, defining the new smart-data paradigm.

It is important to note that in some circumstances the physical model is almost unattainable. Thus, the only possible contribution concerns the data-based model that is constructed from scratch by using any of the available techniques discussed in the present paper, but requiring a larger amount of “smart" data.

From the discussion addressed in the present work, some actions seem urgent to us :

1. In what concerns model order reduction, one of the main challenges is that of constructing consistent interpolations of pre-computed solutions (non-intrusive PGD) on the solution manifold so as to be able to proceed even when solutions exhibit localization. The parametric solutions of models exhibiting bifurcations is another major issue.

   Many engineering problems involve trajectories : processes (incremental forming, additive manufacturing, ...), agent trajectories, etc ... The issue of parametrizing a trajectory remains an open issue of major interest nowadays. Finally, reduced models of components should be integrated at the system level, and consequently efficient ROM-interfaces defined.

2. Concerning tests, the issue of unbiased and biased noise must be addressed, as well as its collection at different scales. Inverse techniques must be developed in order to have access to non-measurable variables, because of its nature or accessibility.

   In the same way that a single test is able to offer a rich amount of data (e.g., image correlation) one could imagine replacing the test machine by a computer, and expecting that by solving a problem that activates as many parametric values as possible, one could expect having access to the parametric solution from a single (few) numerical simulation(s).

3. Regarding the incipient smart-data paradigm, efforts must be paid to create a multi-scale theory of data, a sort of data-thermodynamics, that should work at equilibrium and off-equilibrium, to offer a response to four key questions : (i) what data should be collected ? ; (ii) where ? ; (iii) when ? ; and (iv) at which scale(s) ?

4. For model learners and data-driven modeling, different questions arise. One of them concerns the nature of state variables (able to encapsulate all the history-dependent present state) and the way of identifying them from collected data. Another extremely exciting topic concerns the similarities between deep-learning based on neural networks and more physically based model learners as the ones discussed previously. Finally addressing noise and outliers, and differentiate them from multi-scale physically events remains also an open crucial issue.

5. Finally, concerning data and manifold learning (PCA and its nonlinear counterparts and variants), they are most of the times is based on Euclidean distances. It seems that the extraction of uncorrelated parameters from data needs alternative metrics. Looking at two trees, even a child is able to conclude on their similarity (both are recognized as trees in real-time) even if the Euclidean distance among them could be very large. In this regard, TDA (Topology Data Analysis) is attracting interest because of its appealing properties and spectacular capacity of classifying. Topology persistence, persistent homology, mappers, computational geometry, ... are opening a field of unimaginable opportunities.

   Moreover, the use of persistence diagrams allows us to define metrics based on topology (of major interest when addressing shape and topology optimization) and its associated persistent images (eventually combined with sparse sensing) allows defining interpolation, a crucial aspect when addressing reduced order modeling.

   Very often, similarity must be judged and stablished outside a vector space. Imagine establishing similarity between traffic signals or color words (yellow, red, ...). Identifying the similarity of words referring to color requires their transcription to a vector in a given vectorial space that allows for applying standard analysis tools. This transcription can be successfully accomplished using Word2Vect techniques.

It is at this point the dilemma of data versus models totally loses it sense. Both are not concurrent, they should be considered together, one enriching the other and vice-versa. Physics allows determining what observations should be considered when establishing a predictive data-based model while avoiding major risks, as for example, the violation of the frame invariance or thermodynamical consistency (energy conservation and entropy production). On the other hand, data-science could drive physics towards the most pertinent data offering the maximum amount of pertinent information (smart-data versus big-data). The model-data circle is definitively closed.

Références

[1] F. Chinesta, A. Huerta, G. Rozza, K. Willcox. Model Order Reduction. Chapter in the Encyclopedia of Computational Mechanics, Second Edition, Erwin Stein, René de Borst & Tom Hughes Edt., John Wiley & Sons Ltd., 2015.

[2] F. Chinesta,A. Leygue, F. Bordeu, J.V. Aguado, E. Cueto, D. Gonzalez, I. Alfaro, A. Ammar, A. Huerta. PGD-based computational vademecum for efficient design, optimization and control. Arch. Comput. Methods Eng., 20, 31-59, 2013.

[3] F. Chinesta, R. Keunings, A. Leygue. The Proper Generalized Decomposition for advanced numerical simulations. A primer. Springerbriefs, Springer, 2014.

[4] F. Chinesta, E. Cueto, E. Abisset-Chavanne, J.L. Duval, F. El Khaldi. Virtual, Digital and Hybrid Twins : A new paradigm in data-based engineering and engineered data. Arch Computat Methods Eng DOI 10.1007/s11831-018-9301-4