Proceedings of Asset Integrity Management - Ageing and Life Extension Conference
AIM-ALE 2021
November 30-December 2, 2021, Stavanger, Norway

AIM-ALE2021-76217

DIGITAL TRANSFORMATION BY THE IMPLEMENTATION OF THE TRUE DIGITAL TWIN CONCEPT AND BIG DATA TECHNOLOGY FOR STRUCTURAL INTEGRITY MANAGEMENT

U. T. Tygesen*
Department of Energy, Ramboll
Esbjerg, Denmark
Email: utt@ramboll.com

E. J. Cross
Dynamics Research
Department of Mechanical Engineering
University of Sheffield, UK

P. Gardner
Dynamics Research
Department of Mechanical Engineering
University of Sheffield, UK

K. Worden
Dynamics Research
Department of Mechanical Engineering
University of Sheffield, UK

B.A. Qadri
Department of Energy, Ramboll
Esbjerg, Denmark
Email: baq@ramboll.com

T.J. Rogers
Dynamics Research
Department of Mechanical Engineering
University of Sheffield, UK

ABSTRACT

For re-assessment of 'traditional' industry asset management methods the key technology is Structural Health Monitoring (SHM) combined with the recent development within novel Big Data technologies. The technologies support the digital transformation of the industry with the purpose of cost reduction and increase of structural safety level. Today's State-of-the-Art methods encompass novel advanced analysis methods ranging from linear and nonlinear system identification, virtual sensing, Bayesian FEM updating, load calibration, quantification and propagation of uncertainties and predictive maintenance. Challenges approachable with the new methods cover structural re-assessment analysis, Risk- and Reliability-Based Inspection Planning (RBI), and new ground-breaking methods for damage detection; many of which exploit recent advances in Machine Learning and AI and the concept of the 'True' Digital Twin. In this paper, a selection of the new disruptive technologies is presented along with a summary of the limitations of current approaches, leading to suggestions as to where tomorrows' new methods will emerge. New frameworks are suggested for the way forward for future R&D activities based on an Ontological Approach, founded on a shared communication purpose and the systemising/standardisation of the methods for performing SHM. The Ontology Approach can be embedded in, or made compatible with, organising (and decision-supporting) frameworks based on Population-based SHM methods and extended Probabilistic Risk Analysis. The new ideas also offer the potential benefit of gaining information/learning from a large pool of structures (the population) over time and by transfer learning, transferring missing information to individual structures where less (or no) specific data are available.

Keywords: True Digital Twin (TDT), Structural Health Monitoring (SHM), Linear and Nonlinear System Identification, FEM Updating, Uncertainty Quantification, Risk- and Reliability-Based Inspection Planning (RBI), Predictive Maintenance, Damage Detection, Big Data Analytics, Machine Learning, Artificial Intelligence (AI), Ontological Approach, Population-based SHM (PBSHM), Probabilistic Risk Analysis (PRA).

*Address all correspondence to this author.

INTRODUCTION

There is an often-stated industry need to implement novel digitalisation and automation methods for structural integrity management of offshore structures; in fact, for infrastructure generally.

In the near future, serious decisions will be needed for an extensive number of existing oil and gas platforms; i.e., whether to decommission or extend lifetime. Large investments will be required to preserve a safe oil and gas production over the years to come. Furthermore, it is known that offshore structures, over many years in operation, may undergo changes; these could range from variations in topside weights and platform extensions to changes from subsidence, degradation of grouted connections or the impact of un-expected extreme wave events. Such changes may undermine the safety of the structures and increase their susceptibility to failures. A digital transformation is required to support the future needs for data-based decision making, for optimising investment and for securing the required safety levels.

One prevalent process to assess the structural integrity and (potentially) extend structural lifetime while still maintaining or increasing safety is to merge well-known Structural Health Monitoring Systems (SHMS) with the True Digital Twin (TDT) concept, as presented in [1]. Here the term 'True' is added to avoid some confusion in the industry about what a digital twin covers. In this presentation a True Digital Twin is defined as a virtual representation of a system that spans its lifecycle, is updated from real-time data representing the real physics, and uses simulation and reasoning to help decision-making.

SHMS and TDT both offer capabilities for long-term integrity assurance and minimisation of maintenance costs, simultaneously reducing uncertainties and increasing safety levels. The proposed TDT extends other approaches, by incorporating advanced methods for model updating, quantifying model uncertainties and adding a direct link to Risk- and Reliability-Based Inspection Planning (RBI) methods, e.g. [2]. The method briefly presented in the section "Current Framework" has been applied for more than 15 years on a large number of offshore structures and has a proven record in enhancing information for structural integrity assessment. A detailed overview of the concept presented at five levels (L1 to L5) of application is provided by [3]. The current framework can be applied for most of the structural integrity assessments commonly required today. However, there will be requirements for implementing even more advanced technologies for system assessments in the future, where behaviours may be outside the validation range of some of today's technologies. For example, most common applied methods for system identification currently assume linear and stationary system behaviour. New methods for increasing the validation range of the technology are used will be presented in the "Tomorrow's Methods" section. Examples could be systems with nonlinear and/or nonstationary system behaviour as may be observed for, e.g., damaged structures [4–6], and other types of structures e.g. like bridge supports with non-linear friction. For such cases, special considerations and technologies are required.

Apart from extending the validation range of today's methods, the new methods also offer opportunities for new feasibilities. In addition, to facilitate more integrated solutions than that presented in the 5 level (L1 to L5) concept for creating a TDT [3], more powerful methods for quantification of uncertainties are presented in [4]. The extended ideas for quantification and propagation of uncertainties offer direct potential for improving today's Risk- and Reliability-Based Inspection Planning methods. In section "Tomorrow's Methods", new technologies for more advanced damage detection methods based on state-of-the-art machine learning are also presented (see also [5]).

The possibilities are briefly described in section "Tomorrow's Methods" and further explored in section "The Future – New Frameworks". In the first of those sections, several newly emerging technologies for load prediction, system identification and damage identification are discussed. The methods solve some of the challenges of today's methods. Although these ideas are often founded in advanced machine learning and data analytics, the optimal strategy for problem solving will, in most cases, depend on the proper use of priori known physics; the combination of data-based and physics-based predictive models is the objective of the new field of physics-informed machine learning, which includes the construction of grey-box models [4–7]. The section also discusses transfer learning, a principled approach to making inferences about systems or structures for which very little data are measured; the power of the method lies in improving these inferences using richer sources of data from other (but similar) structures [8–10]. Transfer learning is a key technology within the new framework of Population-Based SHM (PBSHM) [11–14]; this is a very recent approach to asset management, which leverages data from all the members of a population of structures to strengthen diagnostic or prognostic capability for individual members/structures. PBSHM is further discussed in the section "The Future - New Frameworks". Also presented here is the concept of an Ontological Approach [15], as the organising principle for a more systematic approach to communication and sharing of knowledge, and possibly the standardisation of future implementations, for digitalisation and automation of processes for structural integrity assessment. Finally, a new framework, exploiting an extended probabilistic approach for risk- and reliability-based inspection planning is presented [16].

In summary, the purpose of this study is to introduce new technologies which enable more detailed insight into structural integrity by unlocking the potential of data-driven information by applying new and advanced methods of Big Data Analytics and exploiting information gained/learned from a larger pool of structures over time with the purpose of transferring missing information from a large measured population of structures to specific individuals. The new framework is also a proposal for a systematisation/standardisation approach for the future common

R&D on the topic.

New technologies will be discussed, which not only cover offshore structures, but in time can be transferred to all other engineering disciplines, facilitating more efficiently, resilient, and sustainable design, enabling maintenance without compromising safety and functionality. The present paper presents technology facilitating offshore oil and gas production, which society still needs for the global transition to renewable energies, thus supporting sustainability; safety for both personnel and the environment. The technology is general and today transfer to disciplines such as wind turbines [17], bridges, high rise buildings, etc. is ongoing.

CURRENT FRAMEWORK

The TDT is an updated structural analysis model that captures the real-life behaviour of offshore structures in real-time. The TDT provides an advanced state-of-the-art methodology to facilitate coupling between the actual physical structure (the physical twin), it's environment and its digital twin.

An overview of the current Ramboll method is shown in Fig. 1, [3]. The process can be split into two primary stages. The first stage is the calibration of a Finite Element (FE) model against measured data from the structure of interest. The second stage is where decisions as to the maintenance and operation of the structure are made.

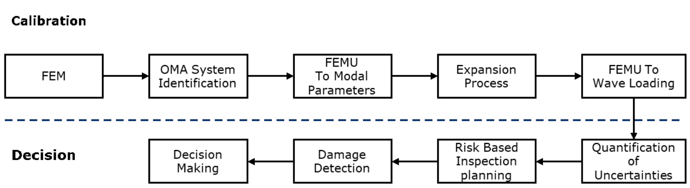

FIGURE 1. Flow diagram giving a top-level view of the current Ramboll method; the process of data-to-decision can be clearly seen.

In the creation of a TDT, a cost-benefit analysis is essential; a five level (L1 to L5) approach is suggested, see Fig. 2. Each level contributes/creates additional value. Following each level L1 to L4, a Decision Gate (DG) is applied to assess whether proceeding to the next level will create value to the operator or not. It is not possible to define generally what the criteria of those gates should be since the cost-benefit for a given system is inherently tied to the business aims of the operator. The definition of the decision gates must be carried out in the scoping stage of the project. It will quantify both the expected costs of advancing through the levels of the TDT and consider the expected benefit of that advancement in terms of the safety and the profitability of the asset.

The levels are only described here in principle. For more detailed information, references are stated in each section.

Five Levels (L1 – L5)	Description
L1 (DG-1)	Screening and diagnostics
L2 (DG-2)	FE model updating
L3 (DG-3)	Load calibration
L4 (DG-4)	Quantification of uncertainties & RBI
L5 (-)	Damage detection & fatigue counter

FIGURE 2. The five levels L1 to L5 for creating a TDT.

Level 1 - Screening and diagnostics (DG-1)

Before deciding on the strategy for digital twin updates and improvements, it is essential to evaluate the performance of the existing digital twin by quantifying its ability to predict the actual structural behaviour. This knowledge enables a diagnosis and creates the basis for deciding whether or not to continue to the next level, i.e. a decision gate must be passed. Linear system identification - also called Operational Modal Analysis (OMA) [18] - is performed to generate a "fingerprint" of the structural characteristics in terms of the modal parameters of the structure. One method of system identification is Stochastic Sub-space Identification (SSI method). The modal parameters are natural frequencies (Fig. 3), mode shapes (Fig. 4), and the associated damping parameters. Typically, the modal parameters are identified based on accelerometer measurements. The fingerprint is evaluated against the predicted modal parameters of the FE model; if measured and predicted modal parameters differ, it is an indication that the FE model needs to be updated. If natural frequencies and/or mode shapes do not match, the correct force/stress distribution in a structure cannot be used for predicting neither nominal stresses, nor hotspot stresses. In this case a decision at DG-1 whether to continue the analysis at Level 2 must be taken. Updating of the modal parameters in Level 2 in case of a poor match is of outmost importance as the updating of the loading in Level 3 is based on updating predicted nominal stresses in the FE model against measured nominal stresses, i.e. the stress distribution in the FE model must be correct before any load calibration is performed in Level 3.

Level 2 - FE Model Updating (DG-2)

The objective of FE model updating is generally to correct masses and stiffnesses (among other parameters) to minimise the discrepancies in the natural frequencies and modes shapes between predicted and measured values. The FE model is updated based on a sensitivity analysis (see Fig. 5), and prior information

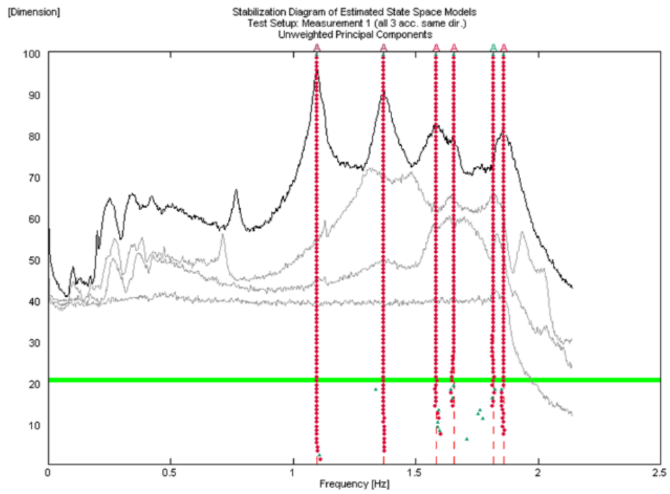

FIGURE 3. Stabilisation diagram from Stochastic Sub-space Identification (SSI) with identification of stable natural frequencies (red).

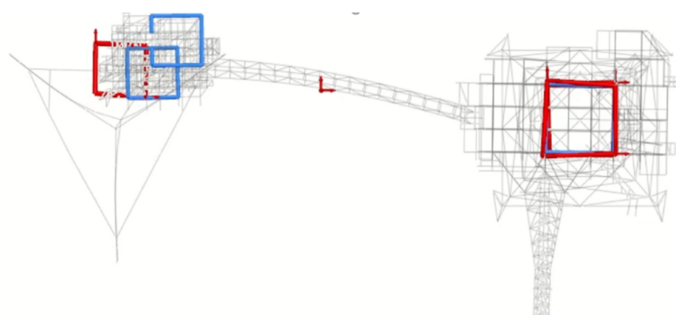

FIGURE 4. Example of comparison of measured and predicted mode shapes (red: predicted, blue:measured). In this case there is a poor match and consequently the analysis must proceed to Level 2 for updating the FE model.

in the form of, for example, a Bayesian parameter estimation approach. Discrepancies in mode shapes are quantified in terms of the Modal Assurance Criterion (MAC) [19], where a MAC of 1.0 indicates a perfect match between a measured and predicted mode shape. Updating of the modal parameters ensures that the static, quasi-static and dynamic characteristics of the prediction model correctly represent the forces and stress flows in the real structure for a given loading. The value created by updated the FE model lies in the improvement in the safety level of the structure compared to the errant FE model that may misrepresent forces and stress in the real structure. However, this update does not ensure that the wave loading is correctly modelled. Typically, load modelling according to applicable code and standards will result in a conservative design. If the conservative design based on the updated model at Level 2 meets the customer need, then the decision can be taken at decision gate DG-2 to stop the analysis at Level 2. In case there is a need for further value creation for example in terms of lifetime extension the decision can be made to continue the analysis in Level 3 for updating also the loading. Level 1 and Level 2 can also support root cause analysis.

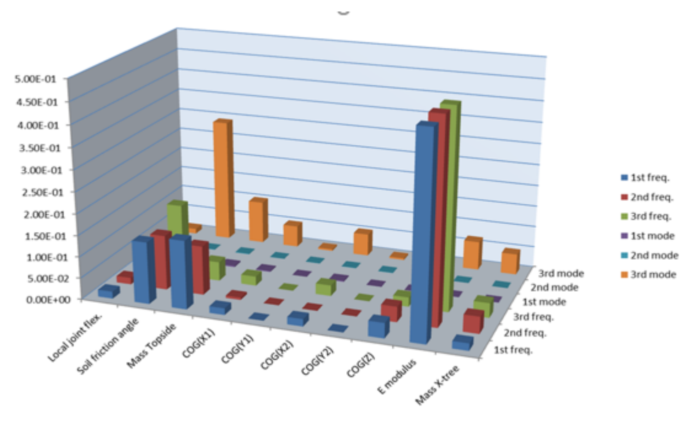

FIGURE 5. Results from an initial sensitivity analysis as basis for a following Bayesian parameter estimator FE model updating.

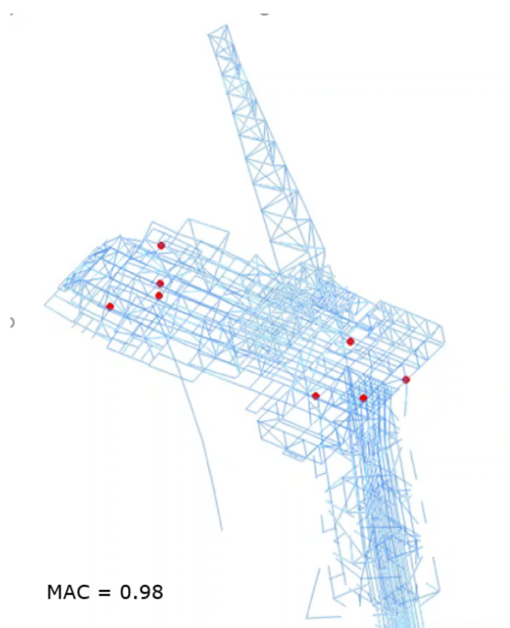

FIGURE 6. Updated mode shape with True Digital Twin (Blue: Mode shape from updated FE model) and the real-world counterpart i.e., the "Physical Twin" (Red: Measured mode shape).

Level 3 - Load Calibration (DG-3)

The performance of the TDT is not solely related to its ability to represent the structural static, quasi-static and dynamic characteristics accurately; the load modelling part is of equal concern when evaluating the performance of a TDT for fatigue prediction. To achieve a TDT for fatigue re-assessment purposes or lifetime extension, it is vital that the wave load model accurately represents the real physical conditions. Wave load calibration requires long-term measurement data - typically from wave radars - to represent the loading part, and in some cases strain gauges to supplement measured global platform displacements generated from accelerometer data. In order to ensure a proper load calibration, it is essential that the load calibration is performed in a way consistent with the method adopted for the fatigue re-assessment analysis which follows. Wave load calibration can be performed by calibrating the wave load model parameters e.g., the Cd and Cm values in Morison's equation [20]. Typically, load modelling parameters from codes and standards yield conservative fatigue predictions. The value created in Level 3 Load calibration is typically potential lifetime extension extension and at the same time an increase in the safety of the structure. At Level 3, decision gate DG-3, it can be decided to stop any further analysis dependent on the operators needs. A decision to continue to Level 4 for further value creation in terms of quantifying the real safety level of the structure and/or optimization of inspection plan requires that probabilistic methods are adopted.

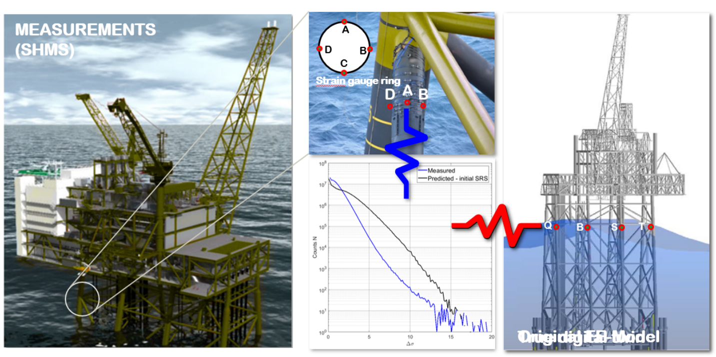

FIGURE 7. Wave load calibration based on calibration to measured and predicted stress range history curves.

Level 4 - Quantification of Uncertainties & RBI (DG-4)

At Level 4, the question of how much the prediction models have improved can be answered in terms of quantifying the uncertainties of the updated model performance against physical measurements. The benefit of continuing to Level 4 is typically only relevant for operators which have needs for reduction of cost for inspection-planning activities and/or just a need for quantifying the real safety of the structure. If the inspection planning is based on Risk- and Reliability-Based Inspection Planning (RBI) methods, the model uncertainties can be quantified in terms of Bias (typically calibrated to Bias=1.0) and the associated Coefficient of Variation (CoV) values. The assessment of the uncertainties should be consistent with the particular RBI approach adopted by the operator e.g. [2]. A number of uncertainties need to be quantified from data and models, ranging from measurement uncertainties and model uncertainties to environmental variations like year-to-year sea state variations, etc. The value creation from a reduction of the uncertainties in terms of reduced CoV values is a reduction in the number of hot spots to be inspected and an increase in the time between inspections, i.e. a reduction in the number of surveys to be performed in the lifetime of the structure. A reduction in the uncertainties typically results in a significant reduction of the inspection costs, and, in some cases, structures have been verified to be inspection free in the remaining lifetime of the structure, i.e. a considerable reduction of OPEX costs. This is exemplified in Fig. 8, where the wave load model of the TDT has been calibrated against measurements. The value creation in Level 4 is in the significant reduction of the uncertainties; as the uncertainties are closely related to the safety, the safety level is improved and now verified by measurements. Based on the vast experience from many projects, a benefit already at the design phase of projects can be offered by the methods, resulting in a reduction in the CAPEX costs. For methods on how to harvest the benefits from a SHM system even at the design phase for a new structure, reference is made to [17]. The method opens for design of new structures adapting a set of reduced partial coefficients. The result is cost reductions for steel material, while still maintaining the same level of safety or an even increase in the safety level. A decision at DG-4 can be taken whether to proceed to Level 5 for continuing the measurements by a permanent SHM system with the potential of further value creation in terms of higher safety level and for providing information supporting critical decision making in case of for example occurrence of un-expected wave events.

Level 5 - Continuous Measurement - Damage Detection

Several methods exist for detecting changes. Some may be useful for specific applications for damage detection, but no approach for robust and general damage detection methods exists as of today. Recently a major R&D project on advanced damage detection based on Big Data Analytics, including Machine Learning/AI technologies in a probabilistic framework, has been funded [21]. The project is a spin-off of years of research performed by many parts on the topics related to the observation of abnormal waves in the Danish North Sea [22].

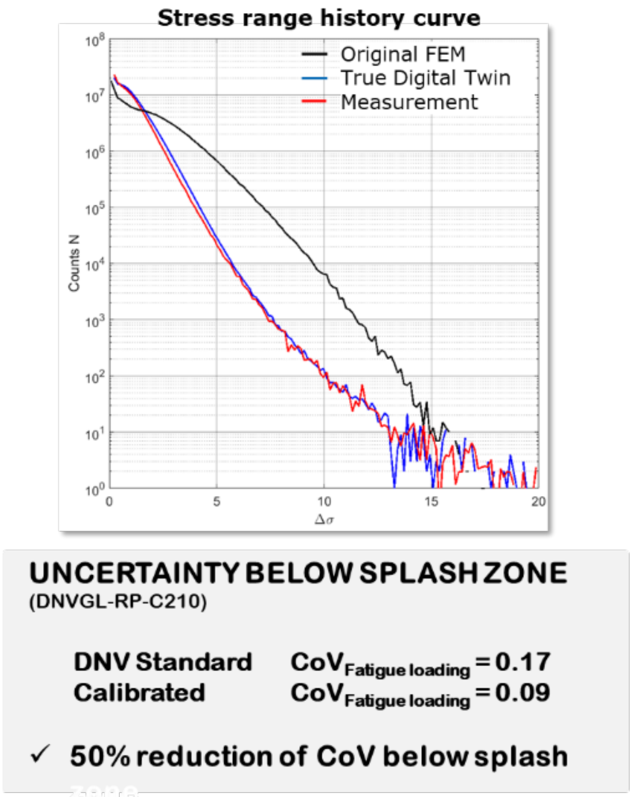

FIGURE 8. Quantification of uncertainties for input to RBI.

Realising the need for advanced damage detection can be exemplified by imagining a storm occurring during the night in which an extreme wave event has occurred, and the structural integrity of the platform might be compromised. In the morning the operator needs sufficient reliable information to support critical decision making on whether it is safe to continue operation, or the platform must be shut down and evacuated due to critical damage.

The main principles are illustrated in Fig. 9, [23]. Starting from the top left with the physical structure in the real world, which in traditional engineering is modelled by a FE model. The new idea is to introduce not only one probabilistic model predicting the present real-world behaviour, but to generate a large number of competing probabilistic models, which constantly will compete in representing not only the present system behavior, but also any future system behavior which could occur including any future damaged conditions.

Monte Carlo simulations for all competing systems are performed for generating scenarios to be stored in a Big Data database. The data from the simulated scenarios are fed into the Big Data Analytics for forming cluster representations as basis for the later quick and reliable automatic identification of damages that may occur in the physical measured data. At the same time, continuing from the top left in Fig. 9, data from observations are stored in another database consisting of real-world measured data. Based on the measured data, the competing models are continuously updated to TDT's, as required. In addition, the measured data is fed into the Big Data Analytics. By combining information from observations with the cluster representations, probabilities can be assigned to each of the competing models for representing the measured data. In the event of an un-expected extreme wave occurring, damage detection can quickly be performed by ranking the competing models by their probabilities to determine the most likely damaged condition.

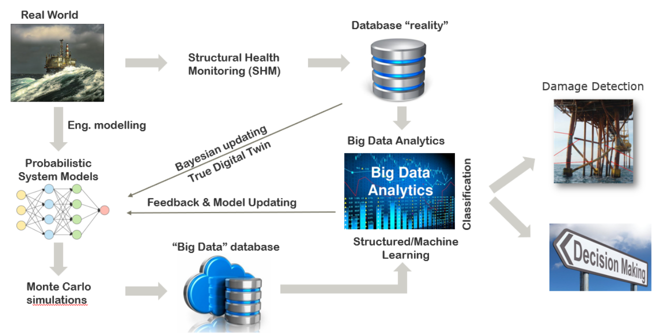

FIGURE 9. Main principle of ongoing R&D for more advanced damage detection based on a combination of utilisation of simulated data from competing models for millions of years with continuous measured data.

TOMORROW'S METHODS

The purpose of this section is to highlight technologies that have been developed in recent years by the authors, in the context of other State-of-The-Art work, which has the potential to extend and improve aspects of the current framework, with a view to increased autonomy, insight and robustness. Each of these developments is built upon increasingly data-centric ideas; in other words, they consider how to make use of most effectively, and learn from, data collected from a structure. The methodologies span the related areas of statistical modelling, machine learning and artificial intelligence. The core aim of integrating new technology into an asset management setting is to increase knowledge, in the sense that the operator can make more informed decisions about the usage and maintenance of a structure.

Considering how one might realise an increased recovery of information, three possible routes are considered, which each show promise in different areas of the current framework. The three points of discussion are: the use of grey-box modelling to combine physical insight with machine learning tools, the application of Bayesian techniques to embed prior engineering knowledge; and the potential of transfer learning for Population-Based SHM. With these methodologies, two different tasks relevant to asset management are investigated. The first is to improve the estimation of the loading experienced by a given structure; if successful, this is a route to reducing uncertainty in the fatigue damage calculations. The second task is that of online damage detection. A framework for damage detection will follow the principles as presented in Rytter's Hierarchy [24, 25].

First, a grey-box modelling paradigm is introduced. A grey-box is a combination of white- and black-box methods. A white-box is constructed based entirely on physical laws governing the system, e.g., Morison equation or an FE model. A white-box approach retain high levels of interpretability, high user confidence and the ability to extrapolate to new conditions (subject to the validated space of the model). However, white-boxes are limited by the expressivity of the physical model and have no capacity to learn behaviour that is not explicitly included (e.g. knowledge gaps). A black-box philosophy sets aside the desire for the model to reflect the physics in a meaningful way, in favour of a more flexible set of modelling methods with the capacity to learn from observed data. A black-box sacrifices the benefits of interpretability in return for improved predictive performance by discovering functional structures from data.

With respect to the grey-box approach: by the combination of physical components from the white-box and data-based components from the black-box, it is hypothesised that the benefits of each could be realised while limiting the respective drawbacks of each approach. In other words, there is a potential of realising the "best of both worlds".

Another important concept used in several recent developments is that of a Bayesian framework for modelling of uncertainties [26]. As engineers seek to construct statistical models of structures of interest, it is essential to consider how meaningful information about uncertainty might be included. There are several of approaches to this quantification, both probabilistic and imprecise. In the name of brevity, the complete array of methods will not all be discussed [27], instead, some reasoning for adopting a Bayesian approach, in a number of the examples, is offered. Two major strengths of the Bayesian methodology motivate its use in this work. Firstly, it is concerned with determining the distributions over variables of interest, by virtue of recovering these distributions it works well as a component within larger analyses where one may wish to propagate uncertainty. Secondly, it provides a rigorous mathematical framework for the incorporation of engineering intuition via description of this knowledge as probability distributions in the prior. These descriptions are flexible and powerful enough to express confidence (e.g. in a time domain) in a manner that might aid algorithms in quantifying a level of belief.

One such example of a Bayesian method which is used in this work is that of the Gaussian process (GP) [28, 29]. The GP is a tool which allows an engineer to learn a distribution over functions [29]. An example pertinent to this paper is to contrast a deterministic viewpoint of learning the loading time-history with the Bayesian alternative of determining the distribution of possible loads a structure has experienced. To do so, the GP expresses some prior belief about the form of the function via the kernel which is then combined with observations of the inputs and outputs of the function of interest to build a posterior distribution over the functions. The mathematical tools to do this are application of Bayes Rule and identities for multivariate Gaussian distributions, for details see [29].

Load Estimation: The loading experienced by a given structure is intrinsically linked to the fatigue life of that structure. As such, by quantifying and reducing uncertainty regarding loading, confidence in the fatigue performance of assets can be improved. Considering the load estimation problem, two different solutions are considered in this work: the first is a direct modelling of the nonlinear dynamics of the waves using a grey-box model – a Gaussian process NARX model; the second a joint input-state-parameter model of the dynamics of the system and the loads themselves.

In Pitchforth et al. [7] an approach for grey-box modelling of wave loading is proposed. The form of the grey-box considered is that of a combination of Morison's equation [20] and the GP-NARX as a nonparametric learner. The NARX formulation is a particular version of a GP model specifically for time series, where lagged versions of the inputs and outputs, i.e. past realisations, are used as additional inputs to the model. This form is a nonlinear version of the well-known Auto-Regressive (AR) model with eXogeneous inputs. Results compare a fully black-box approach (only the GP-NARX model), and a fully white-box approach (only the Morison equation [20]) with a combination of the two. Importantly, this grey-box combination provides the benefits of the flexible GP-NARX machine learner in terms of predictive accuracy provides the robustness of the physics-inspired model of Morison's equation when data are sparsely available. The work demonstrates the power of grey-box approaches in terms of the model's ability to extrapolate, that is to make meaningful and trustworthy predictions outside the domain where data have been previously observed.

An alternative approach towards load estimation was shown in [4, 30], where a Gaussian Process Latent Force Model (GP-LFM) is used to perform joint input-state-parameter inference over a dynamic system. The problem of joint input-state inference is well covered in the literature [31, 32], and efforts have been made towards combined input-state-parameter estimation [33, 34]. In these closely linked problems, dynamicists attempt to infer the internal states of a system (its displacements and velocities) in combination with its inputs – i.e. loads – and in some cases parameters, such as modal properties. Two notable features are present in [4, 30]: first, the unmeasured load is modelled as a GP which can be inferred through a linear state-space model; second that the distributions over unknown parameters of the system are also recovered. In the context of the current paper, the approach of [4, 30] provides valuable information regarding the load that a structure has experienced – intrinsically linked to fatigue damage accrual – and combines this with the process of OMA. In future, this idea may provide a more holistic approach to structural identification for offshore structures in operation, as learning the properties of that system is now coupled with identifying the actual loads that the structure has experienced. In Fig. 10, the example of recovered force from the method on a dynamic system is shown. It can be seen that the approach exhibits excellent recovery of the true force (shown in red) when compared to the mean of the estimate (shown in dark blue); note that these lines overlap, obscuring the blue line because of the very close fit). The shaded area in the figure indicates that the model has recovered these forces with high certainty. It shall be noted that not only the loading is recovered in the process, also the other FE model parameters are updated at the same time.

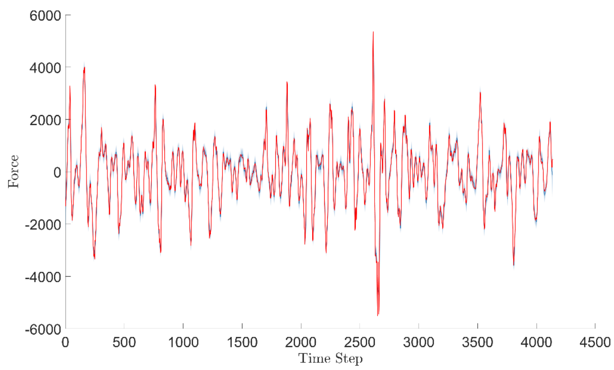

FIGURE 10. Load recovery using the Gaussian process latent force model [4].

This subsection has discussed two possible technologies for estimation of loads on offshore structures; each adopts a grey-box approach combining physical insight with flexible machine learning tools. Estimation of the loading time-history would be a key component in a digital twin system for structures, such as those offshore, which operate in harsh and variable environments with and end of life commonly dictated by the fatigue life of the system. The feature of a more advanced load estimation via a True Digital Twin [1] is a mechanism for reducing uncertainty about the condition of the structure and enabling predictive maintenance. The methods presented here avoid the shortcomings of some of todays' system identification methods in that both non-linear and nonstationary system now can be accounted for. In addition, it shall be noted that the presented formulation allows for performing the levels L1 to L4 as one integrated analysis, where system identification, Bayesian FEM updating, load calibration and quantification of uncertainties are all performed at the same time in one integrated analysis. It facilitates for much more detailed information for detection of for example knowledge gaps (e.g. outside of validation range issue) and uncertainty propagation and hence increased knowledge about the safety of the structure.

Damage Detection: Two possible solutions are considered with respect to the damage detection challenges presented by Rytter's hierarchy [24, 25]. The challenges being the detection, localisation, classification, quantification and prognosis of damage on a system [35]. Both solutions address an essential challenge in damage detection which is especially important for offshore systems. The challenge being the lack of availability of data relating to the structure while in a damaged state. Therefore, the proposed approaches circumvent the requirement for extensive damage-state data at the start of operation by either learning adaptively online in a semi-supervised manner [5] or by transferring knowledge between structures [8–10]. The combination of these two methodologies may well be possible and beneficial and is the subject of ongoing research.

Within machine learning, it is possible to characterise learning methods in terms of the available information which might be exploited by that learner. For supervised learning the training is performed based on both observed inputs and observed outputs. The learning consists of labelling and recognising data in classes or groups related to different normal operational and damaged conditions. The disadvantage of supervised learning can be a long training period and the requirement that only conditions that have been part of the training can be recognised. Unsupervised learning is defined as a situation in which data is available without any labels. The methods are dominated by two-class classification tasks (either undamaged or damaged) tasks based on outlier analysis. The advantage of the unsupervised learning methods is that it requires only a limited learning period/data and that there is no requirement for labelling of data by manual intervention. Another advantage of unsupervised learning is that the method can detect damaged conditions that have not been part of the training. The obvious thing then is to combine the two methods in semi-supervised learning. The advantage of a semi-supervised approach being that information from a few labels can improve performance with a very large dataset. In practice when a learner has been in operation, it is expected that all normal occurring operational conditions have been labelled and from here onwards any unforeseen changes of the structure can be efficiently detected and identified. In general, semi-supervised methods will outperform supervised and unsupervised methods, see [36] for greater detail. In [5], the weakest possible semi-supervision is employed, whereby unsupervised clusters are assigned meaningful labels following inspection of a data point in that cluster; this may be referred to as label propagation; there are however many other approaches that could be taken [37].

In [5], a promising solution for damage detection where damage data are not available is presented. The proposed model is based upon a Dirichlet Process Mixture of Gaussians (DPGMM). Intuitively, the model tries to learn a set of multivariate Gaussian (Normal) distributions that capture the distribution of the extracted features. It is a clustering approach that seeks to group similar data together that are in some way "close" in the feature space. Since this model is probabilistic, the data are grouped by considering their membership of one of a set of probability distributions, in this case multivariate Gaussians. Formally, the model attempts to model the density of all the data as a joint mixture density; that is, a weighted set of Gaussians. To learn such a model, one must determine two things, first the parameters (mean and covariances) of the individual components, second the mixing proportions – what fraction of the data does each component represent. The interesting feature of the DPGMM is that the mixture is comprised of an infinite set of Gaussians and thus can represent any arbitrary density. In a damage detection setting, one assigns physical meaning to each of these different components – which are called the clusters – such that data can be grouped into interpretable sets. In a very simple example, one might have two clusters one relating to the normal condition and one relating to a damage condition. The contribution of [5] was to present an online application of the DPGMM in a damage detection setting where new clusters could be initiated online; practically, this meant that data from structural conditions which had not been observed before could be grouped into new clusters.

The DPGMM provides a powerful approach for a damage detection system which might learn and adapt online, removing the need for extensive data collection from a range of conditions before operation. A further improvement of the DPGMM method could be to combine it with transfer learning in Population Based SHM (PBSHM). PBSHM is a new branch of SHM which seeks to make use of information between structures to improve the efficiency of SHM tasks - for a thorough introduction see [11–14]. One of the key tools in the PBSHM arsenal is transfer learning, specifically domain adaptation [8–10]. The aim of domain adaptation is to map data from two dissimilar datasets onto a common space, such that a single classifier might be used with improved performance. It is especially powerful when a user is data rich in one dataset (the source domain) and lacks information in another (the target domain). Consider an example in the offshore context, a company has monitored a structure for a number of years and has collected extensive data from a range of operating and damage conditions. Now a new structure has had a monitoring system installed for only one month and little is known about how damage may affect the features of interest. In transfer learning, even if the structures are different (e.g. one is monopile and one is a jacket) the two datasets might be combined to allow better predictions on the new structure. [8] shows that this may even be possible between data from an FE model and the physical structure, see Fig. 11. The domain adaptation approach is therefore, a powerful tool for life-cycle management offshore. The reason is that it allows users to leverage more value from monitoring data by transferring the knowledge contained in that data between structures. This increased value will be an important component of a cost-effective damage detection programme, especially for operators with many structures, which may be expensive to instrument. An even larger benefit from PBSM could be achieved in case operators/owners across industries could agree on sharing data for the benefit of all.

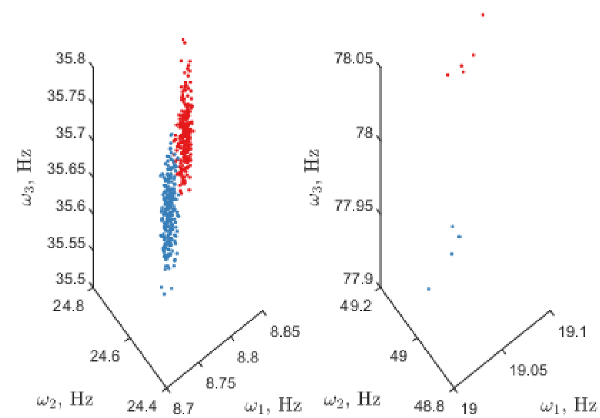

FIGURE 11. To illustrate the power of transfer learning: the left figure shows samples of the first three natural frequencies from a noise-contaminated FE model of a laboratory three-storey structures, in an undamaged and a damaged condition and the right shows experimental measurements from the structure with sparse information. Note the large mismatch in scales between left and the right figure; the FE model has been deliberately created with the wrong geometry and materials in order to test the algorithm. The example show how experience/learning detection of damage from other type of structure in principle can be transferred to a different type of structure with limited measured data.

To briefly conclude this section, an overview of several new developments related to monitoring and management of offshore structures has been given. Two main tasks have been highlighted as challenges which could be addressed in the near future; these were to estimate the unmeasured loading time-history a structure has experienced, with a view to improving fatigue life predictions and to provide methods for real-time damage detection which learn and react online, providing information about the current health state of the structure. Two different grey-box models for load estimation were shown, one where a machine learner (GP-NARX) is combined with the well-known Morison's equation and one where the dynamics of the structure are coupled with a machine learning estimate of the forcing. The presented meth-

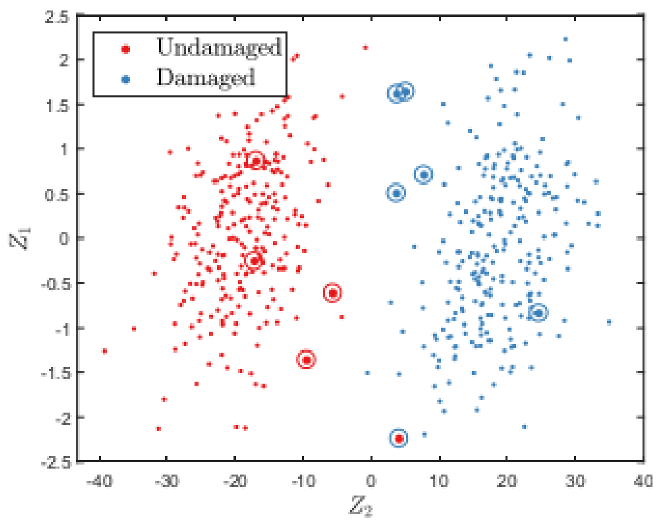

FIGURE 12. shows the frequency data mapped into a common feature space using domain adaptation [8]; in this domain, a classifier trained on the model data is 90% accurate on the measured data.

ods overcome some of the issues with today's methods, which on occasion could be used outside of their validation range without evidence that this is permissible or safe. For damage detection, lack of availability of damage-state data at "Day 0" was highlighted as a challenge for current methodologies. The DPGMM method was proposed to address this via a semi-supervised approach with algorithms that are self-training, recognising and labelling conditions with no human intervention (automatically). The idea is that during a training period, the algorithm has identified and clustered most normally-occurring conditions. Any changes not covering normal operational conditions will from then on be detected and identified.

In addition to this, transfer learning - specifically domain adaptation - was discussed as a technology to move information between structures and – via the framework of PBSHM – maximise value across multiple (possibly disparate) systems.

THE FUTURE - NEW FRAMEWORKS

As the last section has clearly shown, a number of new predictive modelling tools have emerged recently, which offer the promise of revolutionising diagnostic and prognostic capability across many engineering sectors, not least offshore. In most cases, to optimally exploit their full power, the algorithms will need to be embedded in some overall decision framework capable of transforming data into asset management decisions. In order to automate the data-to-decision process as far as possible, holistic systems with minimal human intervention are the goal. As in the case of 'methods', the recent past has seen the emergence of candidate 'frameworks' which could be implemented independently or carefully combined. Three such possibilities will be discussed here: Ontologies, Probabilistic Risk Assessment and Population-Based SHM (PBSHM).

Ontologies: A key requirement of any SHM framework will be the acquisition and storage of data and its refinement into knowledge. Part of that process of refinement will involve the 'methods' discussed earlier. Acquisition and storage of data is the domain of database technology; however, acquisition and storage of knowledge is another matter. Doing this 'by hand' given the huge expanse of data and information in the literature and distributed in the World Wide Web, is simply not feasible. In order to automate this process, ontologies can be used. Ontologies do not have a universally-accepted definition; a simple definition, capturing the properties important for this paper is that they are 'systems developed to organise knowledge . . . coupled with a view on the "world" (or domain of interest) that has motivated it' [38]. In general, ontologies are used to store, share, describe, process, and reuse domain knowledge in specific application domains. They can take a variety of forms, but as specified in [39], they will necessarily '. . . include a vocabulary of terms, and some specification of their meaning. This includes definitions and an indication of how concepts are inter-related which collectively impose a structure on the domain and constrain the possible interpretations of terms.' It is the vocabulary here, and the broad concept of inter-relation, which extends the idea of the ontology well beyond a 'simple' database. Ontologies are usually encoded using dedicated languages e.g. OWL [40].

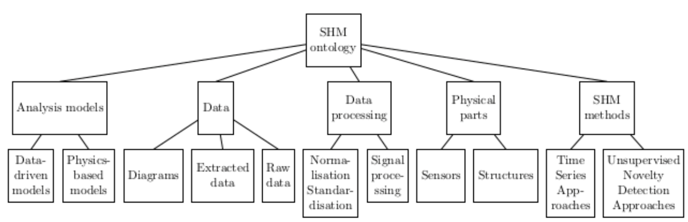

FIGURE 13. The five SHM ontology superclasses and some of their subclasses.

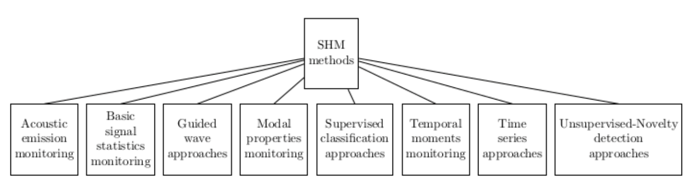

FIGURE 14. SHM "methods" subclass [15].

Regardless of the language that they are expressed in, all ontologies are comprised of the following components:

1. *Individuals* (instances): represent objects in the domain of interest.
2. *Classes* (concepts): sets of individuals (instances). These can be defined using formal descriptions stating precisely the requirements for class membership. Classes may be organised into a superclass-subclass hierarchy, which is also known as a taxonomy.
3. *Properties*: are characteristics that individuals in the domain of interest can have.
4. *Relations*: describe the various ways that individuals and classes are related.
5. *Axioms*: are statements that are asserted to be true in the domain described and comprise the overall theory of the particular application.

No comprehensive ontologies for the purposes of SHM currently exist (although an ontology associated with physics-based modelling [41] has been proposed and following the principle of reusability central to the ontology field, this could be incorporated into more comprehensive structures, like those proposed here, as part of the model-based SHM component of an SHM ontology for example.) The barest outline of how an SHM ontology might be constructed is given in [15]; even so, the main schematic is too complex to include here. Instead, the main classes are given in Fig. 13 for illustration.

Population of the ontology is a particular problem. Of course, this may be carried out by hand, using OWL (or the ontology language of choice); however, this is very restricted. It is more efficient and effective if the ontology can be populated automatically. Fortunately, there has been a body of research based on automatic extraction of knowledge for ontologies from the internet [42]. An initial ontology can be populated by hand to help focus the search.

Population-Based SHM (PBSHM). PBSHM has already appeared in this paper in the context of the 'methods' section. This fact is because PBSHM depends critically on transfer learning algorithms which allow one to improve health inferences on structures with little or no data by using data from other structures within some population. However, PBSHM itself is a 'framework' in which algorithms operate and is much more than a playground for one algorithm. The basic papers on the subject are [11–14], although the subject is in its infancy and is developing currently at a substantial rate.

Critical issues in PBSHM concern when it is appropriate to attempt transfer learning at all; this is important because of a phenomenon known as negative transfer, in which an attempt to transfer between two problems which are not sufficiently similar will cause a degradation in predictive capability on the target structure. This problem is at the root of the basic formulation of PBSHM, and is addressed as follows. The structures of interest, in the population of interest, are converted to 'points' in an abstract space of structures. The key point is that the space of structures should be a metric space; i.e., it should be possible to define a 'distance' between two given structures. This distance is then used as a measure of similarity between structures; such that, if two structures are 'close' in the metric, then they are structurally similar and transfer may be attempted.

The two main steps in finding which 'point' in the abstract space corresponds to a given structure are, first to create an irreducible element (IE) model, then to convert this to an attributed graph (AG). The IE models are quite simple conceptually; one represents the structure of interest by substructures which are structurally or dynamically basic, like beams, plates or shells. This representation is then converted into a graph, where each IE is assigned a vertex (node) in the graph; two nodes are then joined by an edge if their corresponding elements are joined physically in the real structure. At this level the graph only represents the topology of the structure; i.e., its connectedness. However, one can assign attributes to each node – the geometrical and material properties of the element. Suppose an element had the type [plate], then the geometrical attributes associated with the element would be the length, width and thickness; the material properties might be Young's modulus, Poisson's ratio etc. The attributes are assembled into a vector, which is then associated with the node in the graph; the properties of joints between elements are stored as attribute vectors associated with edges. This is all rather abstract, and many details are omitted (see [12]); however, it provides a usable space of structures – the space of attributed graphs is a metric space as desired – and thus a means of assessing the similarity of structures. If this sounds implausible, the reader can consult [43], in which the similarities of several real bridges are assessed; in this case, bridges of similar types are identified effectively; a suspension bridge is paired with another suspension bridge etc (see Fig. 15). A very recent paper – in an aerospace rather than civil context – shows that the AG models of two aircraft wings of different types provide guidance on which data should be used for optimal transfer [44]; in fact, an SHM classifier trained on one wing is shown (via transfer learning) to give 100% accuracy on the other wing.

The wing problem raises another interesting problem in PBSHM, which is that, although similarity assessment is based on the structures themselves, transfer is conducted on SHM feature data measured on the structures. This observation means that the AG/IE models of the structures have to be stored in some database, in which their associated feature data also live. Within this large framework, some 'methods' will compare the similarity of structures; if transfer is indicated, this will be accomplished by operating on the structures' feature data using other methods. As one might imagine, such a database/framework will need to be quite sophisticated; first steps towards construction of such an object are discussed in [45, 46].

FIGURE 15. Illustration of how the PBSHM metric has allowed similarity measures between real bridge structures. The matrix shows the similarity scores between IE/AG models of a set of eight bridges. The score is normalised to unity for a perfect match (the diagonal); note the high scores for the pairs of beam and slab bridges, suspension bridges and truss bridges.

Ultimately, the PBSHM database/framework is intended to incorporate an ontology to store knowledge associated with the SHM context, in order to support the transfer of inferences between structures. Even at this level of complexity, an important ingredient is missing, a PBSHM-based asset management should provide decision support. The necessary ingredients are intended to come from a third 'framework' – probabilistic risk analysis.

Probabilistic Risk Analysis (PRA): Originated decades ago in the nuclear energy industry. It allows risk-informed decisions on design and operation of safety-critical or high-value assets. Preliminary work on a risk-based decision framework has already combined elements of PRA with the SHM paradigm [16], using probabilistic graphical models. Optimal decisions are found by maximising expected utility. The use of utility is an important aspect of the framework; rather than simply assessing the probability that the structure is in some damage-state, the utility weights that probability by a consequence or cost measure. This action is clearly what is needed by an optimal decision process; one should not conduct expensive inspections or maintenance if the system has confidently detected inconsequential damage.

The PRA methodology has been designed using a multi-level structural representation - structure, sub-structure, component - with faults occurring at the component level. The extension to PBSHM will move to whole populations. In fact, population-based PRA generalises the concept of SHM of 'systems-of-systems'. Consider SHM of an offshore wind farm; this entails conducting SHM of an individual turbine, then sub-structures within the turbine, like the tower or blade. If one considers the drivetrain as a structure, one might wish to detect faults at the bearing or gearbox: i.e., component level. It is immediately clear that asset management of a company's entire inventory of farms will require a 'systems-of-systems-of-systems ... of-systems' approach or SN.

Although the three 'frameworks' in this section have been presented separately, there is no bar to combining the features of all three into an overarching system. One can build a PBSHM core, in which an ontology is fused to the underlying database/schema, in such a way that knowledge in the ontology can help match structures in order to maximise the possibility of successful transfer. In this system, a decision-support interface based on PRA and fueled by the diagnostics/prognostics from the PBSHM core can be added. In terms of creating an industry standard, an ontology could be created to share knowledge between users in different companies and across platforms; this is what ontologies are designed for. In fact, multiple users could also contribute their data to the PBSHM core of a shared system; although this may seem implausible from the point of view of data security and company confidentiality, it is probably not; new ideas like differential privacy [47], offer possibilities. A differential privacy system could allow users to improve inferences on their structures, using PBSHM, while keeping the actual data from other structures invisible.

CONCLUSION

In the present paper, the framework for today's, tomorrow's and for the future implementation of novel state-of-the-art technologies for digitalisation and automation for the purposes of structural integrity management has been presented.

The framework of today based on the creation of a True Digital Twin has been presented in terms of a five-level approach (L1- L5), in which value is created at each level as a differentiated service dependent on an owner's or operator's special needs. The State-of-The-Art methods of today encompass novel advanced analysis methods ranging over linear and nonlinear system identification, virtual sensing, Bayesian FEM updating, load calibration, quantification and propagation of uncertainties, Risk- and Reliability-Based Inspection Planning (RBI) and damage detection; many of which exploit recent development in machine learning and the concept of the True Digital Twin.

The value created ranges from verification of prediction model against measurement for improving the safety level of facilities to root-cause analysis, lifetime extension, predictive maintenance, reduction of OPEX costs in terms of reduction of costs for inspections, reduction of CAPEX costs already at the design phase of a new structure, to continuously monitor of changes (damage detection) and monitoring of accumulated fatigue (fatigue counter).

Today's methods have some shortcomings when it comes to nonlinear and/or nonstationary system behaviours, which can be relevant for some structural systems. For such systems the behaviour may be outside of the validation range of parts of the current technologies. In “Tomorrow’s methods” a range of additional machine learning and AI-based technologies are introduced for extending the validation range of current methods among other benefits. Four initiatives are introduced.

The first advance is the use of grey-box modelling; this combines the best of the two worlds of the white-box (maintaining physical insight) and the black-box with the ability to learn/extend the validation range and improve prediction where the physical model fails. A grey-box formulation is presented as tomorrow’s technology for performing OMA for both nonlinear and nonstationary systems.

The second advance is the application of Bayesian techniques with the benefit of embedding prior engineering knowledge for e.g., FEM updating purposes. Engineering knowledge including the incorporation of uncertainties, can be embedded as probability distributions. Introducing a Bayesian method based on a Gaussian Process (GP) formulation, combined with a NARX strategy (GP-NARX), allows for performing all of the Levels L1 to L4 in one integrated analysis. The presented method, in one integrated analysis, performs nonlinear/nonstationary system identification (OMA), as well as Bayesian FEM updating, load calibration and by the introduction of the Gaussian Process formulation also quantifies the uncertainties in an even more detailed format than the today’s structural reliability methods.

The third advance is the introduction of two novel methods for damage detection. First a method for advanced damage detection based on Big Data Analytics is presented. The main idea is based on the concept of introduction of competing models, which constantly compete in representing the real world data from physical measurements both for un-damaged and damaged conditions. Machine learning on million of years of simulated data in a virtual world representation by the TDT concept is combined with physical measurements for quick damage detection in case of occurrence of un-expected extreme wave events. Second a classification/clustering approach for advanced damage detection based on a Dirichlet Process Mixture of Gaussians (DPGMM) is presented. Both the advantages and disadvantages of unsupervised and supervised learning are described, resulting in a final proposal for a semisupervised non-parametric clustering approach. The advantage of a semi-supervised approach is that the algorithms are selftraining, recognising and labelling conditions with no human intervention (automatically). The idea is that, the algorithm has identified and clustered most normal-occurring conditions during a training period. Any changes not covering normal operational conditions will from then on be detected.

The fourth final advance is presented as a Population-based SHM (PBSHM) approach with the potential benefit of acquiring and then transferring information within a large pool of measured structures and transferring knowledge and experience from years of measured data from many other structures to individual structures for which there are limited measured data.

An example presented in the section “FUTURE – NEW FRAMEWORKS”, shows the potential for a huge benefit in combining the DPGMM approach with the PBSHM approach. Adapting information from the measurement of a large pool of structures and combining with measurements from a new structure with limited data, shows that even with limited information one will then be able to make early diagnostics even with access to limited data from an individual structure.

In order to gain optimal benefit from PBSHM a common framework across operators, organisations and engineers etc., is proposed. An ontological approach is suggested and described as a basis for a more systematic, guided, standardised process for the community to collectively enhance structural performance; a common way in the future to gather information and drive the green transition to renewables and societal sustainability. New practices are essential, not only for oil & gas but for all disciplines.

ACKNOWLEDGMENT

Given the bread of the illustrations presented here, the paper could have had many co-authors; instead the authors would like to thank a number of colleagues, who made essential contributions to some of the work used for illustration here; in particular, Julian Gosliga, George Tsialiamanis , Aidan Hughes and Dan Pitchforth of University of Sheffield, and David Hester and Andrew Bunce of Queen’s University Belfast

REFERENCES

[1] Tygesen, U. T., Worden, K., Rogers, T., Manson, G., and Cross, E. J., 2019. “State-of-the-art and future directions for predictive modelling of offshore structure dynamics using machine learning”. In Dynamics of Civil Structures, Volume 2, S. Pakzad, ed., Springer International Publishing, pp. 223–233.

[2] DNV-GL-RP, 2019 edition. “Probabilistic methods for planning of inspection for fatigue cracks in offshore structures”. *DNVGL-RP-C210. Sept. 1*(C).

[3] Tygesen, U., Jepsen, M., Vestermark, J., Dollerup, N., and Pedersen, A., 2018. “The true digital twin concept for fatigue re-assessment of marine structures”.

[4] Rogers, T., Worden, K., and Cross, E., 2020. “On the application of gaussian process latent force models for joint input-state-parameter estimation: With a view to bayesian operational identification”. *Mechanical Systems and Signal Processing,* ***140***, p. 106580.

[5] Rogers, T., Worden, K., Fuentes, R., Dervilis, N., Tygesen, U., and Cross, E., 2019. "A bayesian non-parametric clustering approach for semi-supervised structural health monitoring". *Mechanical Systems and Signal Processing,* ***119***, pp. 100–119.

[6] Rogers, T. J., Worden, K., and Cross, E. J., 2020. "Bayesian joint input-state estimation for nonlinear systems". *Vibration,* ***3***(3), pp. 281–303.

[7] Pitchforth, D., Rogers, T., Tygesen, U., and Cross, E., 2021. "Grey-box models for wave loading prediction". *Mechanical Systems and Signal Processing,* ***159***, p. 107741.

[8] Gardner, P., Liu, X., and Worden, K., 2020. "On the application of domain adaptation in structural health monitoring". *Mechanical Systems and Signal Processing,* ***138***, p. 106550.

[9] Bull, L., Gardner, P., Dervilis, N., Papatheou, E., Haywood-Alexander, M., Mills, R., and Worden, K., 2021. "On the transfer of damage detectors between structures: An experimental case study". *Journal of Sound and Vibration,* ***501***, p. 116072.

[10] Gardner, P., Bull, L., Dervilis, N., and Worden, K., 2021. "Overcoming the problem of repair in structural health monitoring: Metric-informed transfer learning". *Journal of Sound and Vibration,* ***510***, p. 116245.

[11] Bull, L., Gardner, P., Gosliga, J., Rogers, T., Dervilis, N., Cross, E., Papatheou, E., Maguire, A., Campos, C., and Worden, K., 2021. "Foundations of population-based shm, part i: Homogeneous populations and forms". *Mechanical Systems and Signal Processing,* ***148***, p. 107141.

[12] Gosliga, J., Gardner, P., Bull, L., Dervilis, N., and Worden, K., 2021. "Foundations of population-based shm, part ii: Heterogeneous populations – graphs, networks, and communities". *Mechanical Systems and Signal Processing,* ***148***, 02, p. 107144.

[13] Gardner, P., Bull, L., Gosliga, J., Dervilis, N., and Worden, K., 2021. "Foundations of population-based shm, part iii: Heterogeneous populations – mapping and transfer". *Mechanical Systems and Signal Processing,* ***149***, 02, p. 107142.

[14] Tsialiamanis, G., Mylonas, C., Chatzi, E., Dervilis, N., Wagg, D., and Worden, K., 2021. "Foundations of population-based shm, part iv: The geometry of spaces of structures and their feature spaces". *Mechanical Systems and Signal Processing,* ***157***, 08, p. 107692.

[15] Tsialiamanis, G., Wagg, D., Antoniadou, I., and Worden, K., 2020. "An ontological approach to structural health monitoring".

[16] Hughes, A., Barthorpe, R., Dervilis, N., Farrar, C., and Worden, K., 2021. "A probabilistic risk-based decision framework for structural health monitoring". *Mechanical Systems and Signal Processing,* ***150***, p. 107339.

[17] Augustyn, D., Tygesen, U., Ulriksen, M., and Sørensen, J., 2019. "Data-driven design and operation of offshore wind structures".

[18] Reynders, E., 2012. "System identification methods for (operational) modal analysis: Review and comparison". *Archives of Computational Methods in Engineering,* ***19***, 03, pp. 51–124.

[19] Friswell, M., and Mottershead, J., 1995. *Finite Element Model Updating in Structural Dynamics*. SpringerLink.

[20] Morison, J., Johnson, J., and Schaaf, S., 1950. "The Force Exerted by Surface Waves on Piles". *Journal of Petroleum Technology,* ***2***(05), 05, pp. 149–154.

[21] Brincker, R., Tygesen, U., Faber, M., Katsanos, E., Tychsen, M., and Stang, H., 2020. "Innovative structural health monitoring and risk-informed structural integrity management". *Innovation Fund Denmark – Grand Solutions. Case number: 0175-00028A, InnoSHM, Denmark*.

[22] Tychsen, J., and Dixen, M., 2016. "Summary of the impact on structural reliability of the findings of the tyra field extreme wave study 2013-2015". *The 3rd Offshore Structural Reliability Conference. OSCR2016*.

[23] Faber, M., Glavind, S., Astudillo, J., and Quin, J., 2019. "Big data techniques for probabilistic system representation". *International Forum on Engineering Decision Making IFED 2019*.

[24] Rytter, A., 1993. "Vibrational based inspection of civil engineering structures". PhD thesis, Denmark. Ph.D.-Thesis defended publicly at the University of Aalborg, April 20, 1993 PDF for print: 206 pp.

[25] Worden, K., and Dulieu-Barton, J., 2004. "An overview of intelligent fault detection in systems and structures". *Structural Health Monitoring-an International Journal - STRUCT HEALTH MONIT,* ***3***, 03, pp. 85–98.

[26] Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B., 2004. *Bayesian Data Analysis*, 2nd ed. ed. Chapman and Hall/CRC.

[27] Klir, G., and Smith, R., 2001. "On measuring uncertainty and uncertainty-based information: Recent developments". *Ann. Math. Artif. Intell.,* ***32***, 08, pp. 5–33.

[28] O'Hagan, A., 1978. "Curve fitting and optimal design for prediction". *Journal of the Royal Statistical Society: Series B (Methodological),* ***40***(1), pp. 1–24.

[29] Rasmussen, C. E., and Williams, C. K. I., 2005. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.

[30] Rogers, T., Worden, K., Manson, G., Tygesen, U., and Cross, E., 2018. "A bayesian filtering approach to operational modal analysis with recovery of forcing signals". pp. 223–233.

[31] Lourens, E.-M., Papadimitriou, C., Gillijns, S., Reynders, E., De Roeck, G., and Lombaert, G., 2012. "Joint input-response estimation for structural systems based on reduced-order models and vibration data from a limited

number of sensors". *Mechanical Systems and Signal Processing,* ***29***, 05, pp. 310–327.

[32] Nayek, R., Chakraborty, S., and Narasimhan, S., 2019. "A gaussian process latent force model for joint input-state estimation in linear structural systems".

[33] Dertimanis, V. K., Chatzi, E. N., Eftekhar Azam, S., and Papadimitriou, C., 2019. "Input-state-parameter estimation of structural systems from limited output information". *Mechanical Systems and Signal Processing,* ***126***, July, pp. 711–746.

[34] Maes, K., Theland, F., and Lombaert, G., 2019. "Tracking of inputs, states and parameters of linear structural dynamic systems". *Mechanical Systems and Signal Processing,* ***130***, 09, pp. 755–775.

[35] Farrar., C., and Worden, K., 2013. *Structural health monitoring A machine learning perspective*, 1. ed. ed. John Wiley & Sons, Ltd.

[36] Bull, L., Worden, K., and Dervilis, N., 2020. "Towards semi-supervised and probabilistic classification in structural health monitoring". *Mechanical Systems and Signal Processing,* ***140***, p. 106653.

[37] Rogers, T. J., Dervilis, N., Worden, K., Cross, E., and Bull, L., 2021. "Managing system inspections for health monitoring: A probability of query approach". *Proceedings of 39th the International Conference on Modal Analysis*.

[38] Borgo, S., 1970. *How Formal Ontology can help Civil Engineers*, Vol. 61. 01, pp. 37–45.

[39] Uschold, M., King, M., Moralee, S., and Zorgios, Y., 1998. "The enterprise ontology". *Knowl. Eng. Rev.,* ***13***, pp. 31–89.

[40] Heflin, J. "An introduction to the owl web ontology language. lecture notes, lehigh university".

[41] GROSSE, I. R., MILTON–BENOIT, J. M., and WILEDEN, J. C., 2005. "Ontologies for supporting engineering analysis models". *Artificial Intelligence for Engineering Design, Analysis and Manufacturing,* ***19***(1), p. 1–18.

[42] Mitchell, T. M., Betteridge, J., Carlson, A., Hruschka, E., and Wang, R., 2009. "Populating the semantic web by macro-reading internet text". In The Semantic Web - ISWC 2009, A. Bernstein, D. R. Karger, T. Heath, L. Feigenbaum, D. Maynard, E. Motta, and K. Thirunarayan, eds., Springer Berlin Heidelberg, pp. 998–1002.

[43] Gosliga, J., Hester, D., Bull, L., Worden, K., and Bunce, A., 2021. "On population-based structural health monitoring for bridges". *Submitted to Mechanical Systems and Signal Processing,* ***148***.

[44] Gardner, P., Bull, L., Gosliga, J., Poole, J., Gowdridge, T., Dervilis, N., and Worden, K., 2021. "A population-based methodology for transferring knowledge between heterogeneous structures in structural health monitoring: damage localisation on a population of aircraft wings". *Submitted to Journal of Sound and Vibration.*

[45] Wickramarachchi, C., Brennan, D., Lin, W., Harvey, D., Cross, E., and Worden, K., 2021. "Towards a population-based structural health monitoring: Part v, networks and databases". *Topics in Modal Analysis & Testing. Proceedings of the 39th IMAC, A Conference and Exposition on Structural Dynamics 2021*, October.

[46] Brennan, D., Wickramarachchi, C., Cross, E., and Worden, K., 2021. "Implementation of an organic database structure for population-based structural health monitoring". *Topics in Modal Analysis & Testing. Proceedings of the 39th IMAC, A Conference and Exposition on Structural Dynamics 2021*, October.

[47] Dwork, C., 2008. "Differential privacy: A survey of results". In Theory and Applications of Models of Computation, M. Agrawal, D. Du, Z. Duan, and A. Li, eds., Springer Berlin Heidelberg, pp. 1–19.