The Future of Gas Turbine Technology
9th International Gas Turbine Conference
11-12 October 2018, Brussels, Belgium

Paper ID Number: 70-IGTC18

OVERVIEW OF DIGITAL ASSET MANAGEMENT FOR INDUSTRIAL GAS TURBINE APPLICATIONS

Dr. Richard Green[1], Mr. Cody Allen[1], Mr. Brian Fox[1], Dr. Chad Holcomb[1], Mr. Marco Leon[1]

[1] Solar Turbines Incorporated., 2200, Pacific Highway, San Diego, CA 92101.

ABSTRACT

Industrial gas turbines require operational flexibility and availability to be successful in the current business environment. Traditionally, operational risk is managed, in part, through scheduled maintenance (based on operation hours), which is typically formulated through a combination of experience and engineering assumptions. These assumptions are usually inherently conservative and therefore limiting to operational flexibility. Customers are often limited to finite operational hours, starts, firing temperatures and/or rotor speeds. Recent advances in Physics-based Modelling (PBM) and Data Analytics, combined with secure machine data acquisition technologies have provided a platform for the development of Digital Assets. These Digital Assets are created to mirror actual physical assets operating in the field. Digital Assets are a key technology in the implementation of the industrial internet of things (iIOT), providing the flexibility to respond to changes in operation or identify opportunities to optimize the asset's performance during operation, while also maximizing availability.

This paper presents an overview of the critical technologies and approaches needed to successfully build and deploy a functional, efficient Digital Asset, which accomplishes the above goals, in addition to optimizing the life cycle cost of industrial gas turbines. The paper considers a series of key contributing factors, starting with the efficient and secure acquisition of machine data to the development and application of physics-based models, statistical models, and data analytics which utilize the machine data and generate value through a framework of People, Process and Technology.

INTRODUCTION

Industrial gas turbine operators are dependent on the availability of their engines for successful operation of their business. Unplanned down time is costly, disruptive, and represents a significant risk. Equipment reliability is therefore of utmost importance. From an OEM perspective, reliability is critical in satisfying customer needs and expectations.

Historically these considerations have been addressed, with some conservatism, at the component and system design phase. However, as operators expect more capability and value from their industrial units, it has become imperative that OEMs improve their technological and analytical modelling capabilities and leverage them in order to maximize value for their customers.

Currently, a plethora of research is being done on turbine monitoring (for both aero and industrial applications) as seen by the theme of the ASME 2018 Turbomachinery Conference. Gas turbine performance prediction has been broadly classified into two approaches (Sekhon, R., Bassily, H., Wagner, J., 2008). The model-based and model-free approaches each have their merit. Model-free approaches can lead to interesting conclusions, some inconsistent with the physical reality of a gas turbine system.

The newest paragon of architecture for such monitoring and analysis in the industrial turbine setting is an incorporation of an on-site and a remote based system (Simon and Rinehart 2014). The on-site logic continuously monitors control sensors and actuator positions. As part of the architecture, automated data acquisition and transmission systems on-site work to transfer the gathered data to the remote monitoring locations. From these remote monitoring stations, the sampled analog engine measurement data can be further analyzed by algorithms too intricate or memory/processor intensive for on-site feasibility and by system experts. As is true with any modeling endeavor, the key to the performance of the technique is having a model that accurately reflects the nominal operating performance of the actual engine (Simon and Rinehart 2014). Physics-based models are generally very complex, include many variables and require expert knowledge to understand. Furthermore, these models typically require defined inputs and yield defined outputs. Simplified and more flexible approaches abound; indeed, there are numerous investigations for both real-time, model based analysis and post process analysis (Das, S., Sarkar, S., et. al., 2013), (Merrington, G., Kwon, O., et. al., 1991), (Kerr, L., Nemec, T., Gallops, G., 1992).

Furthermore, There are methods that mix both model based approaches; a hybrid modeling approach has been applied where neural networks were trained to engine models and then operated on the underlying engine models enabling a self-tuning mechanism akin to Kalman filter. ((Venturini, M., Puggina N., 2012). In any case, the resultant models and capabilities can be defined as a Digital Asset, which will be further define addressed in the next section.

NOMENCLATURE

PBM	Physics-Based Modeling
iIot	Industrial Internet of Things
N_i	Life to Crack Initiation
ROM	Reduced Order Model
FEA	Finite Element Analysis
EHM	Equipment Health Management
CFD	Computational Fluid Dynamics
RUL	Remaining Useful Life
OEM	Original Equipment Manufacturer
PDF	Probability Density Function

DEFINING THE DIGITAL ASSET

A *digital asset* is a virtual representation of a physical asset, which for this paper, is an industrial gas turbine along with the underlying systems and components. The term is ubiquitous to any industrial equipment which is significantly important, or key, to the successful execution of an industrial process.

A digital asset is a combination of physics-based and data driven models, such as analytics and statistical models, which can be used to define the basic functionality of the asset. The nature of the digital asset is specific to the application of the physical asset and can therefore vary significantly from customer to customer and application to application. The ultimate value of the digital asset is to help define a continuous and accurate risk profile based on machine data, from the available sensors and indicators associated with the equipment. This results in optimal flexibility and utilization of the asset, which can be tailored to individual customers or fleets in (or near) real time.

Typically, complex industrial assets such as gas turbines, are designed with levels of conservatism. This conservatism is a product of assumptions driven from experience and design knowledge as well as expectations to how the equipment will operate. As digital assets are a relatively new development in the industrial gas turbine industry, (GE Research, 2015), specific customers engine operation has not factored directly into legacy designs. Historically engines have been designed around a general application, where the acceptability of the design is governed by the extreme bounds of the desired functionality. This approach ensures a high probability that the asset will function under all potential conditions and is considered safe design. Using experience and design knowledge, it is possible to define a typical range of operations, within which the asset is required to function. The operational range, from mild to extreme conditions can be considered as a probabilistic distribution, see Figure **1**. The mean or average probability would coincide with the typical operation, and the upper or lower bounds would coincide with the most extreme conditions.

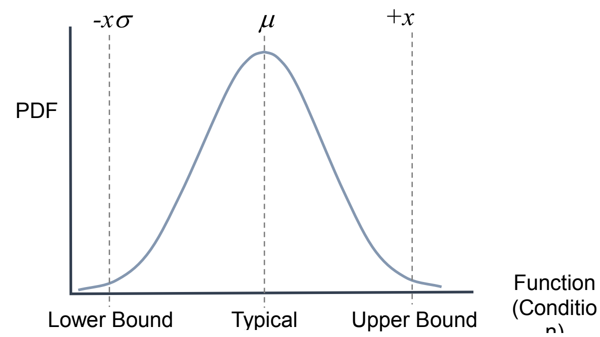


Figure 1: Illustrates the aspects of the safe design approach as a probability distribution.

The confidence required to ensure reliability of the equipment is also the source of the conservatism. Designing the asset to function under extreme conditions that have a relatively low probability of occurring, (i.e. <2.5%) means there is a high probability, (i.e.>97.5%) that the asset is underutilized, (with regard to the functionality in question). For example, in the case of critical hot section components, such as turbine disks, the safe design approach requires the disk is capable of withstanding the most extreme operational profiles, dealing with the highest temperatures and stresses over the greatest number of cycles and dwell periods. The result is a disk that can withstand the extremes of operation, but is probably underutilized for low load applications with respect to the life of the disk.

Unfortunately, this is a necessary byproduct of the safe design approach. It is typically unacceptable for OEMs and operators alike to design, build and operate industrial equipment with a high probability of failure. Therefore, without a condition based approach to asset management there are few options available.

The only way to challenge these assumptions is with machine data and most importantly, a digital asset model capable of processing the data in such a way as to provide value. This becomes increasingly important if the application of the equipment is changing over time with customer needs, as this may invalidate the initial design assumptions. As operational profiles change the need to respond quickly and effectively becomes paramount for optimal management of the asset. If the actual operational profile is known, there is no need to make assumptions and the conservatism can be avoided. thereby, optimizing the full utilization of the asset. It is worth noting, that this does not eliminate all uncertainties, there may be additional system variables which need to be taken into account. This example is only applicable to operation, which is a major source of uncertainty in the design.

It is the customers' needs which drive the functionality of the digital asset. Therefore, functional aspects of the digital asset should be developed incrementally over time in partnership with the operators in order to ensure the digital asset is adding value. This approach affords the developers and operators the opportunity to build capability into the digital asset which is driven by the customers actual needs and is associated with value added.

The recent advent of digital assets can be attributed to advancements in several key areas;

- Computational capability from new software architectures (Zaharia, M., Chowdhury, M., et. al., 2010)
- Data security (Huawei IoT Security, 2017)
- New internet platforms designed to support large scale data processing for industrial applications (iIot) (AWS, 2018) (Google BigQuery, 2018)
- Computationally efficient physics-based and data analytics models (van Paridon, A., Barnes, C. et al) (van Paridon, A., Bacic, M. & Ireland, P.).

However, this industry is still evolving and requires a continued close collaboration between all disciplines to be successful. The future of digital asset development lies in the effective application of machine intelligence. Although a popular term, often indicative of the iIOT zeitgeist, machine intelligence is not synonymous with automation. In fact, autonomous technologies are still early in development, evident by the lack of substantive value to industrial businesses. Rather, the immediate value in machine intelligence is in the ability to effectively communicate large quantities of data quickly and efficiently. Bringing value through a more robust collaboration between the operators and the machine. Effective communication provides real, actionable intelligence from data processing. Examples, such as forecasting the durability of an engine for a given operational profile to minimize downtime and optimize scheduled maintenance are presented later in the paper.

This type of intelligent communication is dependent on several fundamental elements defining the digital asset. In order of importance;

- Machine Data, which can be further categorized into;
  - Ongoing availability of data
  - Secure data acquisition
  - Quality of data
  - Fidelity of data
- Model Definition, which can also be further classified into;
  - Physics-based
  - Analytics
  - Statistical
  - Hybrid

The following sections discuss the requirements for each of the fundamental elements in more detail, followed by examples of how these models interact to communicate actionable intelligence and provide value.

MACHINE DATA

Data is the most important aspect of any digital asset. It is the key resource which determines the effectivity of the asset and the success of the intended application. It is the first consideration when defining the capability of the asset relative to the customer's needs. Before developing models, or defining specifications or capabilities for the digital asset, evaluate and determine the status of the data. The following considerations should be addressed;

Ongoing Availability: This is a fundamental challenge that can be insurmountable. Lack of data is probably the most common challenge when attempting to develop a digital asset. Physics-based models need data for validation, conversely, analytics or statistical models needs data for construction and validation. Availability or even accessibility of data introduces a range of challenges. Even if a physical asset is generating data, is the data accessible and does the accessibility enable scalable applications? For example, a machine may be gathering and recording data, but is the data physically onsite with the machine or can it be remotely accessed? In the case of remote access, there may be physical challenges accessing the data. Often industrial gas turbines reside in remote locations, where internet connectivity is limited. There may be insufficient bandwidth to allow transfer of the data. If there is limited connection capability, is edge computing a solution and if so is that solution scalable? Is a local installation of the software architecture needed to run the digital asset an option? Is there sufficient space, both physical or digital to store and execute the digital asset in the event of a local installation? How will the digital asset communicate with other assets, or operators in order to ensure effective application of machine intelligence.

All these considerations and more are driven by availability of data. Without a scalable solution to access and manipulate data, the intended capabilities of the digital asset can be severely compromised. Remote connectivity is by far the most robust and convenient method of accessing data, however, cyber security becomes a consideration.

Secure Data Acquisition: This follows availability and represents a significant challenge. Cyber security for industrial applications is of paramount importance. Without the ability to securely access the available data, there is no likelihood that the digital asset will provide value. Given the nature of industrial gas turbine applications, security is a key consideration for all customers. Having a secure network which is defensible to cyber-attacks is a critical element of successfully executing their business. If accessing and manipulating data compromises the security of a network it is simply unacceptable. A tremendous amount of effort, preparation and experience in securely handling data is needed to effectively deploy digital assets.

Solar Turbines Incorporated. has developed and deployed over 2000 industrial asset connections that are secure and employ purpose-built connectivity architecture to deliver a full suite of read-only data acquisition from the turbomachinery controller. The connectivity solution includes multiple layers of security including physical and virtual firewalls, and network security protocols. Figure 2 illustrates Solar Turbines Incorporated approach to secure data acquisition.

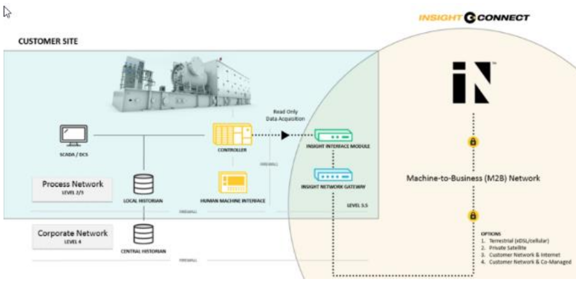


Figure 2: InSight Platform™ Secure Data Acquisition Diagram (Solar Turbines Incorporated, InSight Connect™).

Data Quality: This aspect of data acquisition is intrinsically linked to the model definition and ultimately to the customer's needs. This introduces an iterative element to the development of the digital asset. Without clear model definition prior to the development of the digital asset, there is little to no indication that the data quality is sufficient to achieve the principal functionality of the intended digital asset. Low quality can encompass a variety of events, ranging from minor issues, which can be easily addressed and filtered from the data set (such as a malfunctioning sensor that is part of a larger sensor array), to more severe data quality issues (like missing or corrupted data, that cannot be cross referenced and can only be omitted from the data set). Each data quality issue can be characterized under severity and occurrence. In this way, the overall data set can be managed and evaluated. This framework prevents minor issues being overlooked in the event that the occurrence is high, which ultimately leads to lower quality data.

Data quality can be managed through various tools and processes to ensure repeatability and reproducibility. Approaches to managing data quality issues could include, but are not limited to, probabilistic and statistical methods. Where data quality issues can be quantified, they should be included in a probabilistic output. Figure 3 illustrates an example of how a statistical approach can compensate for data quality.

The challenge with this approach is quantifying the impact of low quality data on the outcome of the digital asset. This is usually predetermined and requires a thorough evaluation prior to the event occurring.

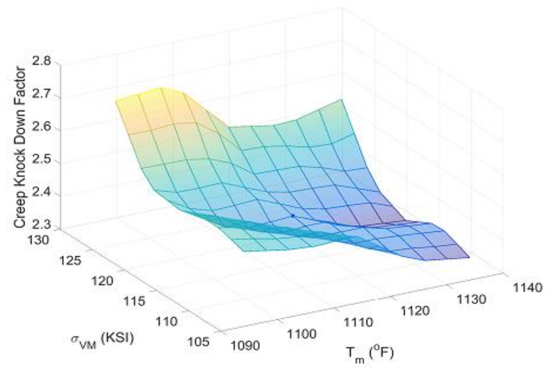
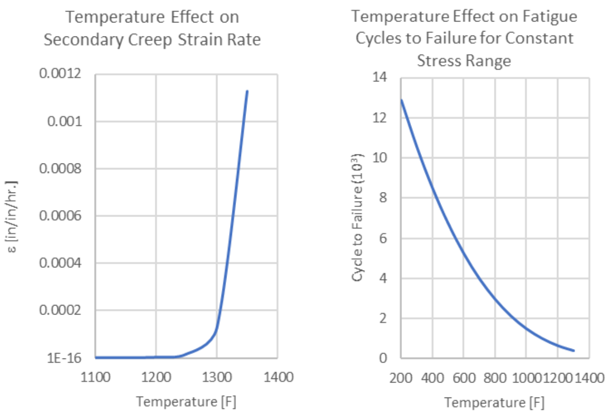


Figure 3: Probabilistic response surface generated for predicting creep strain rate uncertainty over a range of stresses and temperatures for a typical gas turbine Superalloy.

Other approaches which could be utilized include data repair techniques, employed to address known data issues, or data filtering which can clean noisy data. In either case it is important to ensure each correction or repair event is also tracked and quantified to ensure the overall data quality is adequate to perform the intended function with the desired level of confidence. It is important to ensure that the digital asset is operating on real, quality data.

Fidelity: This aspect of the data acquisition is similar to quality in that, the required fidelity is typically driven by the model, which in turn is a function of customer needs. Fidelity can be defined as the combination of the measurement accuracy and the sampling rate. There is the potential for sophisticated models to be employed without the sufficient fidelity in the data to provide any substantive value.

An example of the mismatch between model and data fidelity can be observed when assessing the digital asset for fatigue damage. Most sample rates are a function of time in that a sensor records a reading at a specific time increment. Fatigue damage is calculated using either stress or strain ranges and is a function of frequency. In order for the digital asset to calculate stress or strain ranges, a model is required to translate the machine data that the sensors are recording, such as temperature or speed, into stress or strain at a particular location within the engine. If the data sample rate is insufficient to capture the peak values during a transient event of the machine, the subsequent stress range calculation will be inaccurate, leading to an under prediction of fatigue damage. Therefore, if the purpose of the digital asset is to predict durability and the principal damage mechanism is fatigue, the sample rate needs to be high enough to ensure recording of the peak values during the aforementioned transient, such as, engine starts and stops or large load changes. Fidelity is also important when considering the sensor technology; is the sensor capable of continuously providing the level of accuracy during the life of the digital asset? If the sensor is not robust, or the application is operating in a relatively harsh environment, data may be compromised, leading to the quality issues mentioned above. Is the data sufficiently accurate to achieve the intended function? For example, if temperature sensors are being used to calculate creep damage, is the fidelity of the measurement sufficient to manage the sensitivity in the damage prediction. At typical industrial gas turbine operating temperatures, creep damage can increase or decrease by a factor of 2 for relatively small changes in operating temperatures, such as 20F for commonly used Superalloys. If the repeatability of the measurement is low, then the resultant creep damage could vary significantly, resulting in inaccuracies in the prediction and a loss of capability and value for the digital asset. Figure 4 illustrates the challenge with fidelity for



two common damage mechanisms associated with industrial gas turbines.

Figure 4 illustrates the effects of temperature on creep and fatigue for a standard Superalloy at typical metal temperatures, for industrial gas turbines applications

Conversely, performance degradation of an industrial gas turbine typically takes much longer period of time (weeks, or even months depending on several key operating factors) and as such, lower sample rates can be utilized and still provide the necessary fidelity to perform the intended function of the digital asset.

Therefore, it is critically important to understand the fundamental functions of the digital asset in order to determine if the data being collected is adequate. Again, data presents the biggest challenge to the application of digital assets and should be evaluated prior to the development of any digital asset. It is also worth noting that data acquisition systems require investment and as such it is often important to leverage what is currently available in order to demonstrate value before committing to further investment. Herein lies the dilemma in the use of digital asset for asset management. The potential benefits digital assets can provide in active risk management and asset optimization ensure continued evaluation of this technology but the value of the digital asset in each aspect must be estimated and vetted before significant investment of capital is made to implement it.

MODELS AND ANALYTICAL FRAMEWORKS

As discussed in the previous section, data is the key element of the digital asset. Without an adequate data set the models defining the digital asset cannot adequately function. Therefore, the data along with the customers' needs, define the models needed to construct the digital asset.

Digital asset specifications should start with the customers' needs. What functionality is required to add value to the customers operation? In the case of the industrial gas turbine, basic functionality can be simplified into several categories, i.e.

- Availability
- Durability
- Performance (including emissions)

For the purpose of optimizing the asset from an operational standpoint, models can be developed to predict elements of all the above characteristics. However, before specifying content it is important to evaluate the machine data as discussed in the previous section. Between the limitations defined by the machine data and the requirements defined by the customers' needs, the best models can then be identified to perform the desired function. There are several classifications of models and techniques which can be employed;

- Physics based models (Bound)
- Data driven models (Unbound)
- Hybrid models

Each model type has specific advantages and disadvantages which should be considered for the purpose of optimizing the specific functionality of the overall digital asset.

Physics-Based Models: can be defined as models governed by the laws of physics and are inherently bound. Physics models incorporate the physical characteristics of whatever functionally is being modeled. The models are typically derived from physical laws and have been validated under a range of conditions. The formulation of the model does not change with application, rather, the specific application under consideration is prescribed via the variables defined in the model. For further reference consider any physics or engineering text.

The disadvantage is that physics based models can vary in complexity and can require significant amounts of data for validation of the model. Every aspect of interest of the physical situation being modeled should be captured by the physics model and validated with appropriate data. For example, physics models predicting a material response will typically require material data, as well as operational data, which then requires further translation into variables that can be coupled to the material data, such as strain tensors. This aspect of the physics-based model creates a significant challenge when considering computational capabilities and efficiencies. Often physics-based models rely on advanced numerical techniques to solve, as in the example above, calculating strain tensors for a complex geometry, typically requires a finite element analysis (FEA). Similarly, calculation of velocity and pressure distributions of internal flows with complex geometry often require computational fluid dynamics (CFD), both of which can require significant computational resources and more importantly, time. These types of approaches are not suited to dealing with large operational data sets. To address this, reduced order or surrogate models are typically employed which reduce the computational effort and subsequently time. The challenge here is ensuring that the reduced order models, (ROMs) are capable of capturing the accuracy needed to adequately perform the intended function.

The advantage is that physics models are bound. The models cannot only be used to safely interpolate, but also extrapolate with a high degree of confidence that the predictions are correct. It is for this reason that Physics-based models are the preferred approach for dealing with nonlinear extrapolations. Highly nonlinear stress-strain systems, such as those found in gas turbines, can be very difficult to extrapolate and require domain expertise to understand the relationships which govern the system.

Considering the benefits and limitations, it is important to understand when to employ this approach. Physics-based models are best suited to applications where the resultant outcome has no prior indication of the event occurring and the occurrence of the event is intolerable. For example, a suitable application for physics-based modeling would be predicting the life of an industrial gas turbine disk. Often there is no measurable indication that a part will develop a crack before the event occurs, (since stress and strain is not directly measurable in a gas turbine) and the presence of a crack in a critical rotating component maybe intolerable to the continued operation, resulting in an unplanned shutdown or potential catastrophic failure. Therefore, a physics-based model would be a suitable candidate to predict this event using the available machine data. In this case, the sensor data must be translated into parameters, such as stress and strain which are usually required to predict durability.

Data Driven Models: are unbound and are based solely on the input data and model selection. Represented by different forms of mathematical models (Murphy, 2012), ranging from simple ordinary regression models to highly complex neural networks with many hidden layers. Data driven models coupled with statistical techniques can be employed to find relationships and correlations between any given data sets.

This approach does not require a physical framework and as such can be used with any combinations of data sets. However, the applicability of the output is very much dependent on experience as it may have no grounding in physical relationships that must hold in reality and should be factored into the capability of the digital asset. This provides significant flexibility when considering which data driven model to employ. Data driven models provides more choices and tend to be more computationally economical than physics-based models.

The biggest drawback is the unbound nature of the models, extrapolations and potential correlations maybe nonsensical. This can introduce significant risk when considering the complexity of physical systems, such as industrial gas turbines. At best, the system is highly complex and requires experience and domain knowledge to extract value from the machine data and resulting output of the data driven models. At worst, the interactions are too complex, resulting in spurious correlations and misleading data driven models which generate unwarranted concerns that cannot be substantiated. Such an example can occur in equipment health management, (EHM). Consider the blind application of a data driven model to performance degradation, where the data set acquired is from the site's winter months only. The data driven model is fit to the data and produces low errors upon validation. Then data from summer months is acquired where the ambient temperature is much higher, and upon evaluating the performance of the gas turbine with the data driven model, it appears the gas turbine has degraded substantially, when in actuality the decrease in power comes from the decrease in density of air due to increased temperatures and the resulting effect this has on compression and the combustion process. If the data driven model's conclusion is taken without the domain expert's agreement, a recommendation for shutdown and compressor wash will be suggested to the customer, or worse, a shutdown and full inspection of equipment both resulting in unnecessary downtime.

This aspect of deploying data driven models for predictive analytics introduces one of the most significant challenges. In the case of very expensive industrial equipment, such challenges to use purely data driven models are only overcome by utilization of large, high quality and high-fidelity data sets, which cover most of, if not all of, the domain of operation for a given piece of equipment. Furthermore, domain expertise if required to differentiate low quality and high-quality data sets. It should be clear that such data sets create an artificial bound on the model, in the sense that, if the data set used to train the model spans the entire data space there is no need for a predictive model and the exercise becomes self-defeating. However, the existence of such an expansive data set is rarely found in reality. Indeed, it is both costly and impractical to expect customers to allow the asset managers to perform daily tests in situ on all available machine data for the purpose of building fully expansive data sets. It is much more practical to build a predictive model capable of achieving this outcome. This is a key differentiator in the application of data driven models to large scale IoT domains. Even with the application of data driven models, domain expertise is crucial. The combination of physics models and data driven models is where the digital asset realizes its full potential; such combinations can be referred to as hybrid models.

Hybrid Models: are blends of all the different classes of models to achieve an optimal configuration for the digital asset. This provides the developers considerable flexibility by taking advantage of the aspects of each model approach maximize the advantages and mitigate the disadvantages.

Depending on the complexity of the system being modeled, it is not always possible (or at least feasible) to model each aspect of the system with the most representative model. There are serval reasons why this maybe the case, ranging from inadequate data (as mentioned in the section above) to practical considerations, such as sufficient computational resources. This is characterized by the durability example given earlier. The most appropriate model to represent engine or component durability for a gas turbine engine, that fits the model definition requirements is a physics-based model. It is particularly challenging to translate machine data, represented by a collection of operational parameters such as temperatures, pressures and speeds, into the corresponding transient stresses and strains at a given location on a key component. Further computational effort is required to operate on the stresses and strains needed to calculate damage for the damage mechanisms of interest, (such as fatigue, creep, oxidation, etc.) which is then used to determine remaining useful life (RUL). Then this must be done for multiple locations on multiple key components within the engine. This process flow only represents the computational element of determining durability. To have a more complete assessment of durability, additional data such as physical data, (via component inspection) is required and is discussed in more detail in the next section.

Traditional approaches (full order finite element models) which convert gas path temperatures, pressures and rotor speeds are impractical and as discussed previously, would require excessive computational effort. Therefore, reduced order (or surrogate) models are required to simulate the full order model approach and are not considered physics-based models, but are derived from the physics-based models. These types of hybrid models (van Paridon, A. et al) are typically lower fidelity, but are much more feasible and computationally efficient, allowing the calculations for stress and strain to be completed in practical time frames. Although not strict physics-based models, these types of ROMs do follow a basic physical frame work, represented by the lumped mass approach, but are unbound, in that the coefficients are not limited to physical bounds. The results of the ROMs are key to calculating damage using the material models, which are true physics-based models. Without this hybrid model approach, it would be impractical to process large operational data sets spanning multiple years.

Hybrid models also encapsulate data driven models which are trained on the high fidelity, physics-based models. For example, Solar Turbines Incorporated can use its high fidelity thermo-dynamics simulation physics-based model to generate a range of operational data. Data driven models such as Neural Networks, Support Vector Machines and Ensemble methods can learn the physics-based model parameter space. These models are a type of surrogate model, however, often materialize as black boxes. This black box nature can be one of the drawbacks of such data driven models, however, these models can be specified at various levels of complexity and often do capture steady state dynamics of a particular system provided enough of the parameter space is capture in the training data. It should be noted that accuracy is a function of the input data and should not be confused with the capability of the model, see above section on data integrity.

IDEALIZED DIGITAL ASSET STRATEGY

Value can be defined in a variety of ways, but is typically associated with the ability to accurately forecast behaviors or events which can be used to optimize the asset. Figure 5 illustrates the digital asset concept.

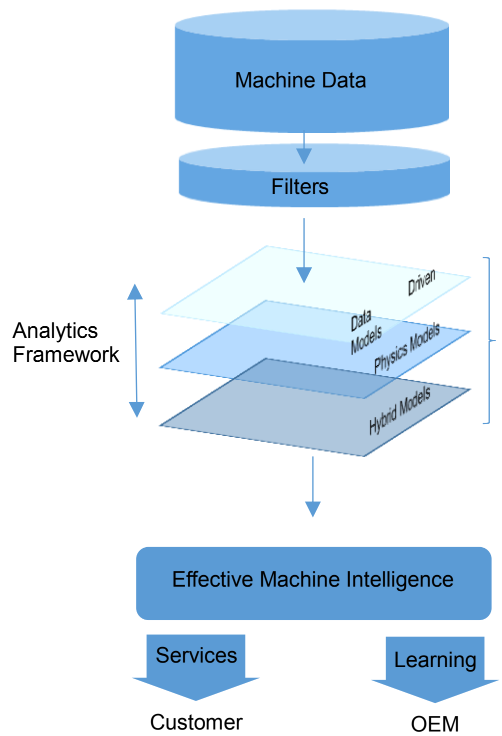


Figure 5: Illustrates the constitutive elements of the digital asset relative to the flow of data and ultimate purpose.

The layering of models, combined with the choice of model types is dictated by the desired functionality, which is driven by customer's needs and available useful machine data, (as discussed in previous sections). The number and definition of the model layers as well as the interaction between layers will be specific to the digital asset. It is typical to have model layers that are completely independent and do not interact with other model layers.

There are also model layers that are critically dependent on other model layers, such as in the hybrid model discussed in the previous section. The level of interdependency is defined by the desired functional output.

As the graphic depicts, analytical frameworks are designed to integrate multiple layers and provide predictive capabilities. An example framework may utilize physics-based, system specific models for a given engine system in one layer while utilizing fleet data driven or fleet statistical models in another layer. As an example of fleet driven statistics, consider an OEM monitoring the health of a remote asset. The fleet statistics layer may compare descriptive statistics of various comparable engines running on the same operational envelope to classify nominal vs off design behavior. For example, in the figures 6a and 6b, a specific engine's first and second bearing vibration displacement (peak to peak) is shown. In one case, the engine, colored green, is well within the expected range of the fleet. The other, can be seen to be the 5th highest by average, in the fleet. However, note that the data range as described by the typical boxplot parameters, appear perfectly reasonable if compared only with itself in the same time frame, which is also seen in the time series plot.

This example illustrates the power of the hybrid model approach, where both physics-based and data driven models amalgamate in the analytics framework to deliver maximum value to the customer.

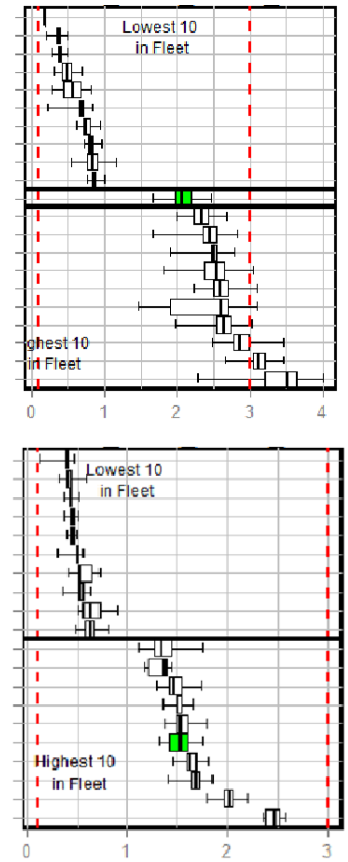


Figure 6(a): Example vibration [mil p-p] plot for a given engine, two different bearings, compared with the top 10 and bottom 10 engines in the comparable fleet. Note that the top bearing plot is well within the fleet margin whereas the bottom plot shows the vibration to be among the largest in the comparable fleet.

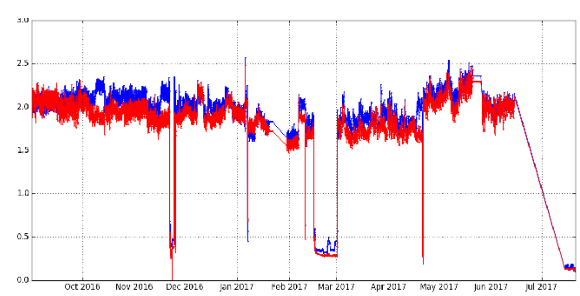


Figure 6(b) Time series data of same engine compared only with itself. The data has noise, but doesn't appear to be significantly different in time.

The analytics framework, regardless of model layers, must be validated properly before the digital asset can deliver reliable value. Sources of validation should be as abundant as possible and should include many independent sets of data such as

- *Fleet wide historical data*

- *Acceptance Test data*
- *Disassembly data*
- *In situ inspection data*

From experience, it is clear that adequate data sets are necessary but not sufficient to create a successful digital asset. Domain experts with an understanding of system design and intent are required for the successful implementation of an analytics framework, and construction of corresponding digital assets. Care should be taken when acting on insights from data alone.

APPLICATIONS

To emphasize the concepts discussed in the previous sections, two examples of digital asset functionality are presented which demonstrate how a blended approach to digital asset development can provide an effective solution to address traditionally difficult applications.

Application I: Predicting the Remaining Useful Life (RUL) of Turbine Disks with physics-based models: Along with other key components, turbine disks are critically important components in the determination of overall engine durability which, as discussed in the previous section, is a principal element of asset optimization. For this example, RUL, is represented by life to crack initiation, N_i.

As discussed previously, traditional life prediction approaches would require complex full order models (FEA) to translate the machine data, such as gas path temperatures and engine speeds into a representative stress state at a given life limited location. These models are impractical for this application, requiring too much computational effort to predict the actual RUL. Therefore, hybrid models, in the form of ROMs, should be used for this application. A ROM for each life limiting location is required as there may be several different life limiting locations on the disk. Continuous condition based assessments require the digital asset to assess multiple damage mechanisms simultaneously, identifying the most damaging location based on varying operating profiles. Therefore, multiple ROMs of key locations, (such as disk or blade firtree lobes and cooling holes), are created to represent the turbine disk and the Analytical Framework must be able to handle all relevant damage mechanisms and their interactions. Each ROM creates a specific stress and strain hysteresis response for each engine cycle. Typically, the material response is nonlinear and path dependent. Therefore, the complete stress strain hysteresis should be representative of the entire loading history. Figure 6 illustrates the process flow and subsequent material response for a given operational profile.

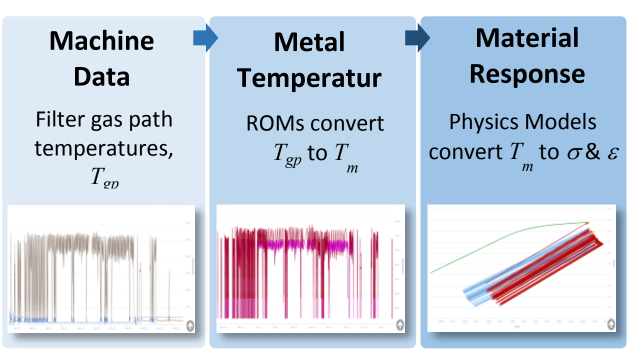


Figure 7: Hybrid (ROM) approach for predicting material response for a complex operational profile

This is critically important and is perhaps the most valuable aspect of a condition based life prediction approach. Evaluating the stress strain hysteresis as a continuous history identifies interactions between various damage mechanisms caused by variations in the loading cycles. The interactions can result in significantly more damage, compared with evaluations performed on identical (stabilized) cycles or when damage mechanisms are considered independently, as in the case of an equivalent hours approach (Green, R. et al). Failure to account for these types of interactions between major damage mechanisms may result in severely under predicting damage which could result in unintended consequence, such as catastrophic failure and unplanned downtime.

Once the full stress strain hysteresis response is predicted, the damage can be determined using physics models. In this example, the physics models represent material models for predicting primarily creep and fatigue damage. Material damage models will vary depending on the component and subsequent damage mechanisms governing the life of that component. This is where domain knowledge and experience is important, to ensure the correct material models are considered for specific life limited locations.

Knowledge of the design, materials and overall operational limits are critically important to understanding which damage mechanisms should be considered. Lack of representation, or inadequate model selection may lead to underpredicting damage and increased risk. In the case of industrial gas turbine disks, the predominant damage mechanisms are creep and fatigue. It is important to select a material model capable of predicting damage from both mechanisms as well interactions. There are several accepted life prediction approaches capable of modeling this type of behavior, (Manson, S.S., Halford, G.R.). Selecting the most appropriate is dependent on many factors, including experience, available material data and even the business models and preferred approach to asset management. One such approach is Ductility Exhaustion (Ainsworth, R. A., et al.), (Green, R. et al, U.S. Patent No. 9.200,984) which is a physics-based model that can be used to calculate damage from creep and fatigue as well as accounting for interactions between both mechanisms. The subsequent damage predictions for each cycle can then be accumulated for the entire operational history. Figure 7 illustrates the computational process to move from measurable machine data to final durability prediction

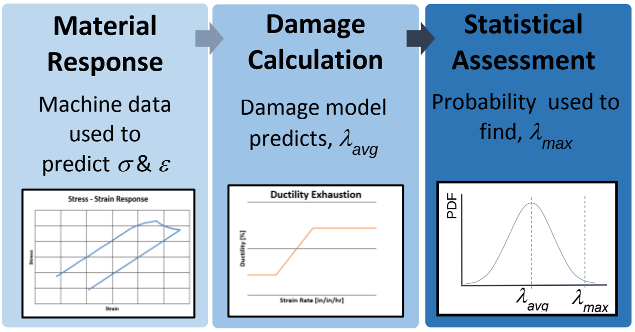


Figure 8: Process flow for predicting durability from machine data

By integrating this damage prediction into a statistical framework, it is then possible to generate probabilistic models which can be used to manage risk. The probabilistic model is used to determine a confidence in the historical damage accumulation but can also be used to provide a framework for forecasting. Forecasting in this example is defined as a method of predicting the future damage state for a given operational profile, or the probability of generating an operational profile based on external conditions. It is this functionality which provides most value to equipment operators or asset managers. Through the ability to forecast the future state of durability, it is possible to adapt operational parameters in order to minimize planned down time and optimize life cycle costs. Therefore, through the effective blending of hybrid, physics and statistical models, within an analytical framework, it is possible to create models which can be used to predict durability based on machine data.

Application II: Recoverable Performance Estimation via Data Driven Models: In the subsequent example, we compare a physics-based model of compressor efficiency degradation against a data driven model of compressor degradation. Consider the performance of a gas turbine axial compressor. A basic adiabatic efficiency calculation computes the ratio of isentropic work to actual work and can elucidate the state of the compressor health (Mattingly, 1996). The question is how these values are calculated. In a first approximation, we can take the ratio of isentropic enthalpy over observed enthalpy, and we may further simplify by assuming an ideal gas. In a second approximation, we may add more complexity by removing the ideal gas assumption. In a high-fidelity simulation, we may utilize CFD outputs of inlet and exit temperatures and pressures, which employ the physics-based model approach. We can in turn, build a pure data driven model by utilizing machine data, (if enough data is available) and we have domain expertise to help identify the correct model and parameters.

For example, knowing that ambient temperatures and pressures effect the compressor discharge lead us to account for both of these parameters in our data driven model. If we then require the data be filtered to only include data from a certain operational envelope, we can build a data driven model to fit the compressor discharge pressure based on some set of independent parameters. The data driven model may be a well-defined, continuous curve or it could be a discontinuous machine learning regression type model such as a decision tree or neural network. If the data used to train the model is from a new compressor, or shortly after a water wash has been performed, the data may capture enough of the operational envelope such that the data driven model's output very well approximate the compressor's output pressure given some set of inputs. Then, the data driven model may be used to create a ratio of actual pressure to predicted pressure, giving a measure of performance or efficiency (Allen, C., Holcomb, et. al, 2018)

Both models were adjusted to begin at 100% efficient. In the physics-based approach (red data), a high-fidelity thermodynamics model was used with the engine's ambient conditions and is the ratio of observed compressor discharge pressure to predicted discharge pressure. The data driven model is a third order polynomial bound between ambient temperature and compressor discharge pressure. The model was fit to the initial data period, where the data (used for training) has been filtered based on the operational envelope.

Figure 8 (a) shows that the data driven model, while not as accurate as the physics-based model, does capture salient features of the compressor efficiency degradation. Figure 8(b) shows the percent error between the data driven model and physics-based model. The data driven model has a maximum error of 2%, depending on the operating condition, however, the results are achieved in near real-time with minimal computation effort in comparison to the full order physics-based model.

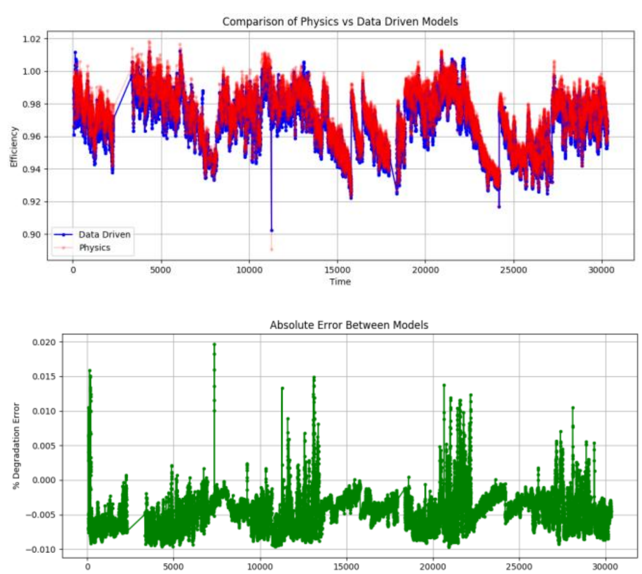


Figure 9: (a) Data driven vs physics-based model comparison of compressor efficiency, (b) relative error plot

of data driven model with physics-based model. In both, time is in hours.

The value is inherent when managing large fleets, where scalability is paramount for data processing and predictive analytics. The physics-based models are generally too computationally intensive to run in real time or close to real time. As discussed in the previous section, physics-based models usually require a large number of input arguments, some of which are not directly measurable. In many remote monitoring applications, the trend in a parameter may be more informative than the absolute accuracy, which can be realized very efficiently with data driven models.

CONCLUSIONS & FURTHER WORK

A high-level overview has been presented that considers the fundamental aspects and requirements for developing and deploying digital assets, specifically for industrial turbine applications. We have discussed the critical aspects of what defines a digital asset. What type of data and acquisition methods and systems are necessary for a digital asset to function. As well as provide a description of the various model frameworks and approaches, along with an integrated strategy, which is needed to provide sustainable value.

In addition, two examples have been provided demonstrating the practical use of digital assets for both durability and performance predictions. Both examples show the value of the different modeling approaches identified in the modeling framework. It has been proposed that useful deployment of digital assets requires much more than the application of complex mathematical models. Indeed, to successfully provide actionable information through the application of machine intelligence, it is necessary to consider multiple options and approaches within the framework of people, process and technology. Machine intelligence requires a balanced platform from which to drive correlations, otherwise there is a risk of generating nonsensical or misleading data. This outcome can lead to nuisance predictions that have no physical meaning. Moreover, correlations without causation may materialize, potentially leading to inaccurate or incorrect conclusions which ultimately impacts business performance.

It is therefore critically important that digital assets are developed and deployed within the framework of people, process and technology. Thereby leveraging domain expertise within a stable executable platform, which applies the effective technology in order to create sustainable value.

ACKNOWLEDGEMENTS

The authors would like to acknowledge the support of Christopher Meyer[1], Brandon Stein[1], John Mason[1] & Graham Benjamin[1].

REFERENCES

Sekhon, R., Bassily, H., Wagner, J., A Comparison of Two Trending Strategies for Gas Turbine Performance Prediction, Journal of Engineering for Gast Turbines and Power, 130(4), 041601, 2008.

Simon, D., Rinehart, A, A Model Based Anomaly Detection Approach for Analyzing Streaming Aircraft Engine Measurement Data. ASME Turbo Expo 2014: Turbine Technical Conference and Exposition, 2014.

Das, S., Sarkar, S., Ray, A., Srivastava, A., Simon, D.L., Anomaly Detection in Flight Recorder Data: A Dynamic Data-Driven Approach. American Control Conference, 2013

Merrington, G., Kwon, Oh-Kyu, Goodwin, G., Carlsson, B., Fault Detection and Diagnosis in Gas Turbines. Journal of Engineering for Gas Turbines, 1991.

Kerr, L., Nemec, T., Gallops, G., Real-Time Estimation of Gas Turbine Engine Damage Using a Control Based Kalman Filter Algorithm, 114(2), 187-195, 1992.

Venturini, M., Puggina, N., Prediction Reliability of a Statistical Methodology for Gas Turbine Prognostics, Journal of Engineering for Gas Turbines and Power, 134(10), 101601, 2012.

Zaharia, M., Chowdhury, M., Franklin, M., Shenker, S., Stoica, I., Spark: Cluster Computing with Working Sets, HotCloud 2010.

Huawei IoT Security White Paper, Huawei Inc., 2017.

Amazon Web Services (AWS) White Paper, Amazon Inc., 2018.

Google BigQuery Documentation, https://cloud.google.com/products/big-data/, 2018.

van Paridon, A., Bacic, M., Ireland, P., Barnes, C., Lewis, L., Reduced Order Transient Disc Temperature Models For Online Health Monitoring. Proceedings of ASME Turbo Expo 2014: Turbine Technical Conference and Exposition. GT2014-26296, June 16 – 20, 2014, Düsseldorf, Germany

van Paridon, A., Bacic, M., Ireland, P., Kalman Filter Development For Real Time Proper Orthogonal Decomposition Disc Temperature Model. Proceedings of ASME Turbo Expo 2016: Turbomachinery Technical Conference and Exposition. GT2016-56330, June 13 – 17, 2016, Seoul, South Korea

Connectivity Solutions for InSight Platform™, Solar Turbines Incorporated, https://insight.solar.cat.com/web/connect.html, 2018

Murphy, K., Machine Learning: A Probabilistic Perspective. MIT Press, 2012.

Green, R., Douglas, J., Blankenship, E., Impact of Engine Operation on Gas Turbine Component Durability Using Ductility Exhaustion. The Future of Gas Turbine Technology. 8th International Gas Turbine Conference. 12-13 October 2016, Brussels, Belgium

Manson, S.S., Halford, G.R., Fatigue and Durability of Metals at High Temperatures, ISBN: 978-0-87170-718-5, 2009.

Ainsworth, R. A., Hales, R., Budden, P. J., Martin, D. C , (eds), Assessment procedure for the high temperature response of structures. Nuclear Electric Rep. R5, Issue 3, Revision 2, 2014.

Green, R., et al, Condition Based Lifing of Gas Turbine Engine Components, US 9.200,984 B2. Dec 1st 2015.

Allen, C., Holcomb, C., de Oliveira, M., Estimating Recoverable Performance Degradation Rates and Optimizing Maintenance Scheduling. ASME Turbo Expo 2018: Turbine Technical Conference and Exposition, 2018.

Mattingly, J., Elements of Gas Turbine Propulsion, McGraw-Hill, 1996.

GE Global Research, The Digital Power Plant, Marketing Report, 2015.