Use of Digital Twins for Process Safety Management

Arvind Keprate and Nikhil Bagalkot

Department of Mechanical, Electronics and Chemical Engineering, Oslo Metropolitan University, Norway

Contents

Abstract

The chapter discusses the framework for developing a Digital Twin (DT) for the process safety management (PSM) of small-bore piping (SBP) on a typical offshore platform. One of the important problems during the PSM of SBP is that due to significantly large number of SBPs on a process facility it is very difficult to place sensors at small bore connection (SBC) for stress estimation. In absence of the stress values, it is difficult to estimate the remaining fatigue life (RFL) of SBC which further impedes the inspection planning. Thus, in this chapter, a methodology comprising of CFD, FEA and Machine Learning is sued to obtain a virtual sensor for stress estimation at the SBC. The input to the virtual sensor is process parameters such as pressure and flow velocity while output is the maximum Von-Mises stress at the SBC. Thereafter, probabilistic crack growth law coupled with Bayesian Network is used to develop a DT for RFL estimation of SBP, which in turn is used to obtain reliability curves and inspection plans. Online deployment of the developed DT will give an up-to-date RFL estimates and inspection plans which can be used then be used for PSM of the SBP.

Keywords

Process safety; Digital Twin; Physics Based Modeling, Surrogate Modeling; Bayesian Network

1. Introduction to Digital Twin

In the past two decades, the concept of Digital Twin (DT) has found numerous applications across different industries such as aerospace, manufacturing, automotive, healthcare, process etc. The concept was first introduced in the field of aircraft and aerospace as an information mirroring model for spacecraft state simulation and to obtain accurate data for decision-making assistance (Glaessgen & Stargel, 2012). The first definition of the DT was postulated by Grieves as a "*Physical product in real space, virtual product in virtual space and the connection of data and information that ties the two spaces together*." Ever since then, numerous definitions of DT have existed in the literature

This is a preprint-version of the following journal article/book chapter; DOI: https://doi.org/10.1016/bs.mcps.2022.05.001

arising particularly due to the usage of DTs in different industries. In context to defense and aerospace industries DT is defined as "*an integrated multiphysics, multiscale, probabilistic simulation of an as-built system, enabled by Digital Thread, that uses the best available models, sensor information, and input data to mirror and predict activities/performance over the life of its corresponding physical twin.*" (West and Blacburn 2017). The main focus in the aforementioned definition was on core concepts of structural mechanics, material science, and estimating the long term behavior of the aircraft. However, for the manufacturing industry, besides system performance prediction, the DT also ensures continuous information transmission amongst various phases of product lifecycle, based on computer-aided simulations (Haag and Anderl 2018). The main goal of using DT in the process industry is for online operational monitoring, predictive maintenance, leakage detections and dynamic process optimization. Thus, in context to process industry DT can be defined as "*as a virtual representation of a physical asset enabled through data and simulators for real-time prediction, optimization, monitoring, controlling, and improved decision making*" (Rasheed et al., 2020). Irrespective of which industry DTs are used, the three main pillars of the DT are, physical system, virtual representation of the physical system (this could be 3D CAD model, ontologies or Simulink/Simscape/Modellica model) and a digital thread (continuous stream of data transfer between physical and virtual system). A common architecture of the DT is shown in Fig 1.

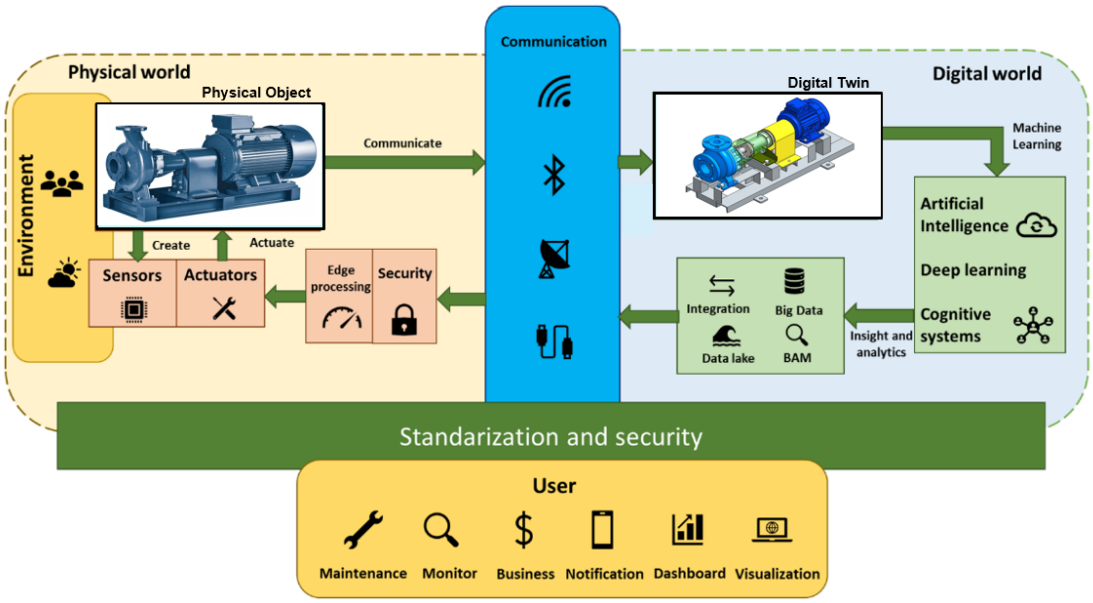

Fig. 1: Common Architecture of a Digital Twin (Adapted from Ferreira et al., 2019)

The concept of DT can be classified based upon its application area and also on how the dataflow occurs between physical system and its virtual representation. Based on the application area the different types of DTs are (Grieves and Vickers, 2017):

1. Digital Twin Prototype (DTP): A DTP consists of the informational sets (such as 3D models, Bills of Materials, Bills of Processes etc.) which is imperative for describing and producing a physical version that emulates the virtual version.
2. Digital Twin Instance (DTI): A DT that represents its physical counterpart throughout its life cycle and is linked to a particular physical asset with the help of continuous digital thread (historical data, live data from sensors, tests and inspection).

3. Digital Twin Aggregate (DTA): An aggregation of all the DTIs, of a particular system or system of systems.
4. Digital Twin Environment: It refers to fully integrated, multi-disciplinary physics application space for operating a DT for predictive (to estimate the future behavior) or interrogative (to understand the historical or current state) purposes.

Based upon how the dataflow occurs between physical system and its virtual representation, distinction can be made between digital model, digital generator, digital shadow and a digital twin, as highlighted in. Fig. 2. For the digital model, the data integration between the physical asset and its virtual representation is not automatic which implies any change either in physical asset or in a virtual model must be executed manually by analyst/engineer. A bit more advanced concept in terms of data integration is a digital shadow which allows unidirectional automatic information flow from the physical world to the virtual world. For e.g. a system where sensors measure information from the physical model and transfer signals to the virtual model is an example of the Digital Shadow. Irrespective of whether information flows is continuous or discretized, as long as it is automatic, the integration level can be determined as a digital shadow (Sanabria, et al., 2022). As can be seen from Fig. 2, for a DT, interaction and convergence are its two key aspects, meaning that autonomous bi-directional data exchange between physical system and virtual system is a must. For a DT, the information flowing from the virtual world to the physical asset will instruct actuators to perform certain operation which will change the physical asset in some way. On the other hand, the data from the sensors on the physical asset generally changes the virtual twin automatically in such a way that it accurately represents the current state and the future evolution of its physical counterpart. Thus, A DT is something alive that changes, improves and evolves while maintaining the comparison between physical and virtual space (Sanabria, et al., 2022).

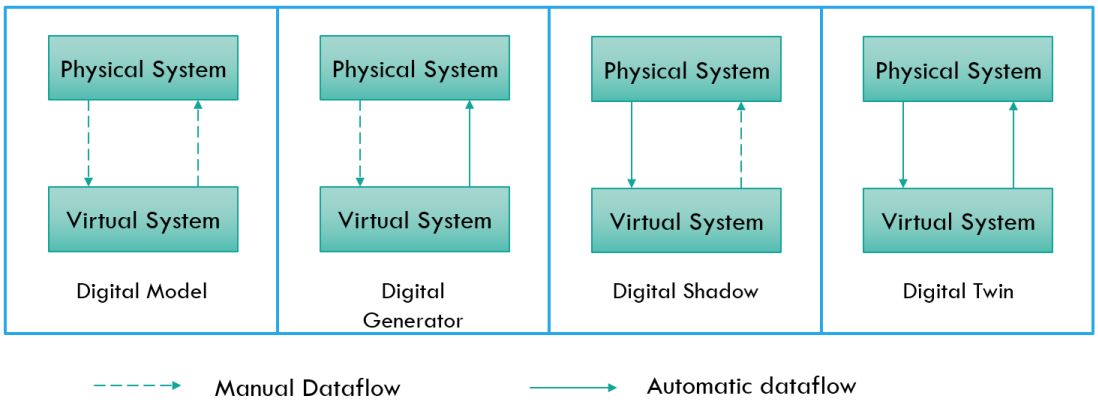

Fig. 2 Difference Between Concepts Related to Digital Twin (Adapted from Tekinerdogan and Verdouw, 2020)

In order to avoid confusion between DT and several other related concepts such as model, simulations, Cyber physical systems (CPSs) and Internet of things (IoT) it is important to create a clear distinction between these concepts. Engineering models (such as mathematical models, statistical models, 3D CAD, data driven etc.) generally consists of many idealizations and approximations and are used to understand the underlying physical phenomenon/system. For example, Paris law is an empirical crack growth model which is used in different industries in order

to calculate remaining fatigue life of the component. Simulations on the other hand primarily involve computer-based models which are used during the design phase to perform what-if analysis for the underlying physical system. Simulation technology has changed considerably from single

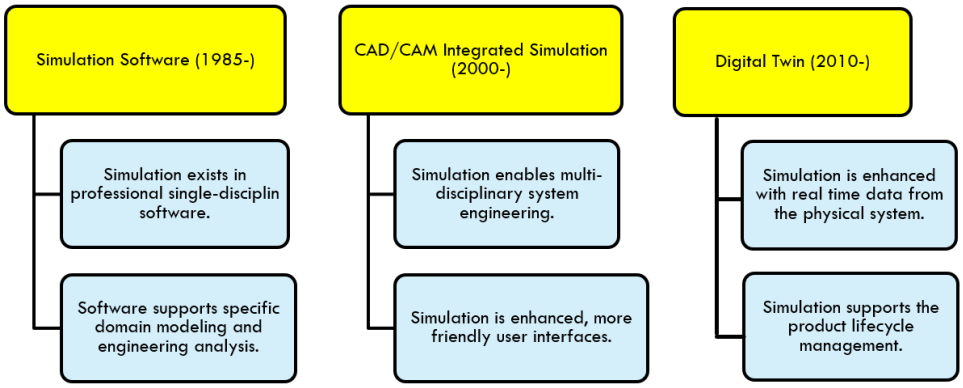

Fig. 3. Difference Between Simulation and a Digital Twin

discipline software (1985) to CAD/CAM integrated multi-disciplinary software (2000) as shown in Fig 3. Nevertheless, simulation results depend upon the input given by the analyst and hence it is a static virtual representation of the physical system. DTs overcome this gap by coupling simulation modeling with IoT platform, and thus the input to the DT is the real-time data captured with the help of sensors on the physical system and thus enabling the analyst to virtually observe the operation of the physical system. Thus, simulation, models, IoT platform are combined to form the DT in order to enhance the value across the entire life cycle of the product. It must be mentioned there that many real assets (such as Wind Turbines) are cyber-physical systems, meaning that they consist of both the physical/hardware elements (such as Blades, tower, Nacelle, generator) and cyber/software elements (such as control system etc.). While building a DT of a CPS, the physical elements are represented by a virtual model (which are approximations of reality as discussed above), while the cyber elements, can be included directly in the DT, which can then be used for modeling, simulating, and optimizing the CPS. The DTs can be used during entire life cycle of the system, for example during design phase they can be used for virtual testing, while during manufacturing they can simulate the production process. From the process safety management point of view, application of DT during the operational phase, offers the maximum benefit as they can be utilized to continuously reflect the past, present and future health of the system/asset, which in turn can be used to formulate inspection and maintenance plans. Besides this, DTs also allow analyst to perform complex and safety critical simulations, in order to calculate the dynamic risk, before such an event happens in real life. The operators can then take necessary decisions in order to mitigate the dynamic risk estimated by the DT, and thus maintain process safety at the industry regulated levels.

2. Process Safety Management of Small-Bore Piping

On a typical process facility, the process piping can be categorized into mainline piping and Small Bore Piping (SBP), as shown in Fig.4. The SBP is a process piping having diameter of two inches or less which typically branches off the main process piping and is often used as relief lines, nozzles, instrumentation ports and drains. The connection of SBP with the mainline piping is called as Small-

Bore Connection (SBC). In the recent years the refining industry has had a history of failures occurring in Small-bore Connections (SBCs) (Baggett 2019). Likewise, both the petroleum and

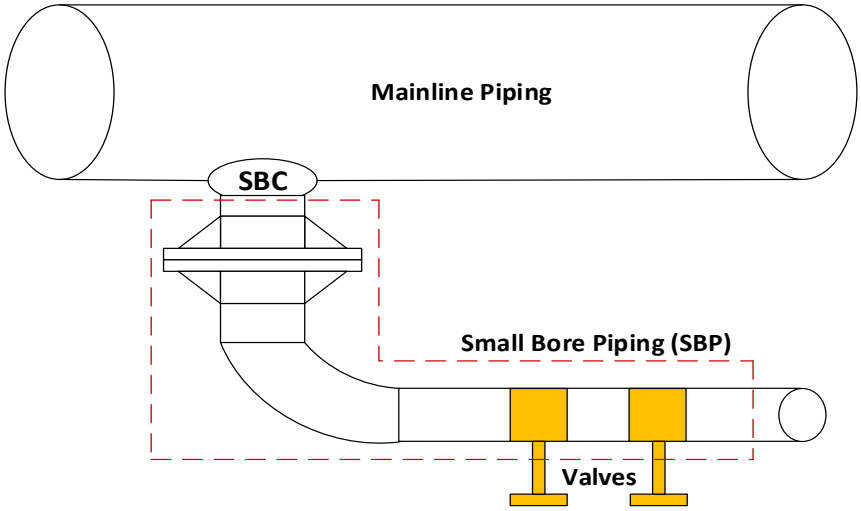

Fig. 4 Schematic of typical process piping

maritime industry also report significant SBP failures leading to Hydrocarbon Release (HCR). Although all HCRs emanating from failed pipework are of concern, the main danger comes from gaseous releases, as the gas cloud produced has the potential to quickly spread across the OOG platform. For instance, a gas release from a 6-mm diameter hole in a piping system operating at about 150 bar (i.e. 15 MPa) has the potential to turn into a major HCR (i.e. greater than 300 kg) in approximately eight minutes (EI, 2013). If ignited, HCR can escalate into a major accident, resulting in significant economic losses and environmental damage, whilst causing a serious threat to the lives of personnel. Different degradation mechanisms (such as corrosion, erosion, stress corrosion cracking, corrosion under insulation and fatigue, etc.) are responsible for diminishing the health state of SBP, nevertheless fatigue (emanating due to vibrations) is attributed as the most prominent deterioration phenomenon causing failure of SBP as depicted in Fig. 5.

In a process piping system, cyclic loads emanate due to the excitation of the mainline piping carrying the hydrocarbons. This consequently leads to the magnified vibrations in small bore piping and ultimately causing the fatigue failure of the SBCs leading to HCR. Hence, from process safety perspective it is necessary to comprehend the chain of actions causing the fatigue failure of the SBCs. Fig. 6 depicts the failure chain due to the Vibration Induced Fatigue (VIF). It can be seen from the Fig. 6, that the excitation forces responsible for vibrating the mainline piping are classified as steady state or transient. The former occurs during the normal operating conditions and emanates due to phenomenon such as flow-induced turbulence, flow-induced pulsations, pulsations from pumps/compressors, etc. (EI Guidelines, 2007). On the contrary the latter originate during events such as changing operating conditions, additional units coming online, normal start-up and shutdown, emergency shutdown, valve operation, etc. (HSE-OTR-028, 2002). In order to mitigate the likelihood of the HCR from the process piping on the offshore platforms, it is necessary to deeply understand the excitation mechanisms depicted in Fig. 6. The taxonomy of the fatigue originating in process piping is illustrated in Fig. 7. Process piping especially near rotary equipment is subjected to High Cycle Fatigue (HCF) and VIF is the major reason for it's failure. Although piping vibration

is seldomly detected visually, nevertheless, expert knowledge (from standards such as EI Guidelines and from practicing engineers) coupled with condition monitoring data can help to frame effective inspection and maintenance (I&M) strategies for preventing unwanted piping breakdown due to fatigue. In order to cater for better understanding of the fatigue loading of SBP, EI guidelines suggest a more detailed analysis of SBP using advanced simulation techniques, such as Computational Fluid Dynamics (CFD), Finite Element Analysis (FEA), as these techniques allow, analyst to perform the dynamic simulation of pressure, temperature and flow characteristics in process facilities, which in turn are used to compute risk and reliability of the SBP (EI, 2007).

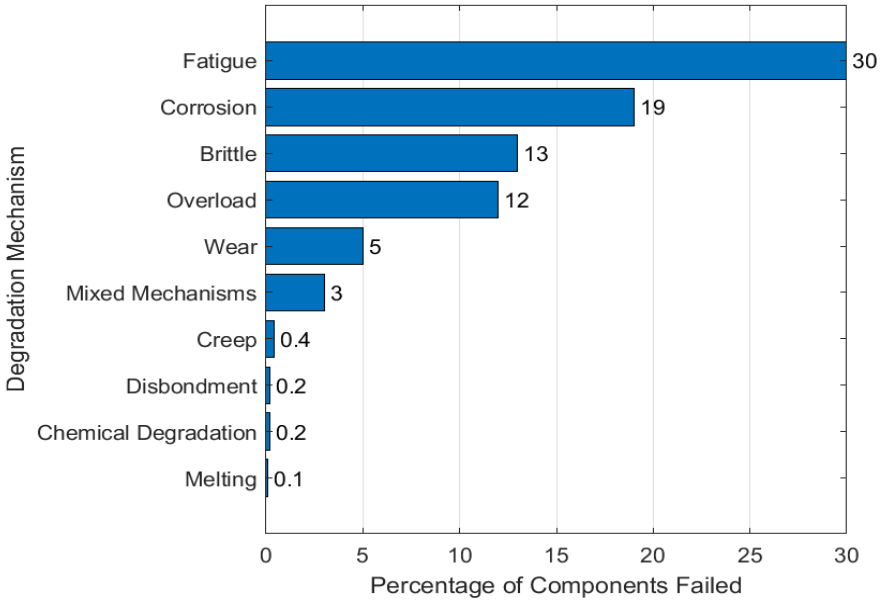

Fig. 5 Most Dominant degradation Mechanisms (adopted from DNV, 2020).

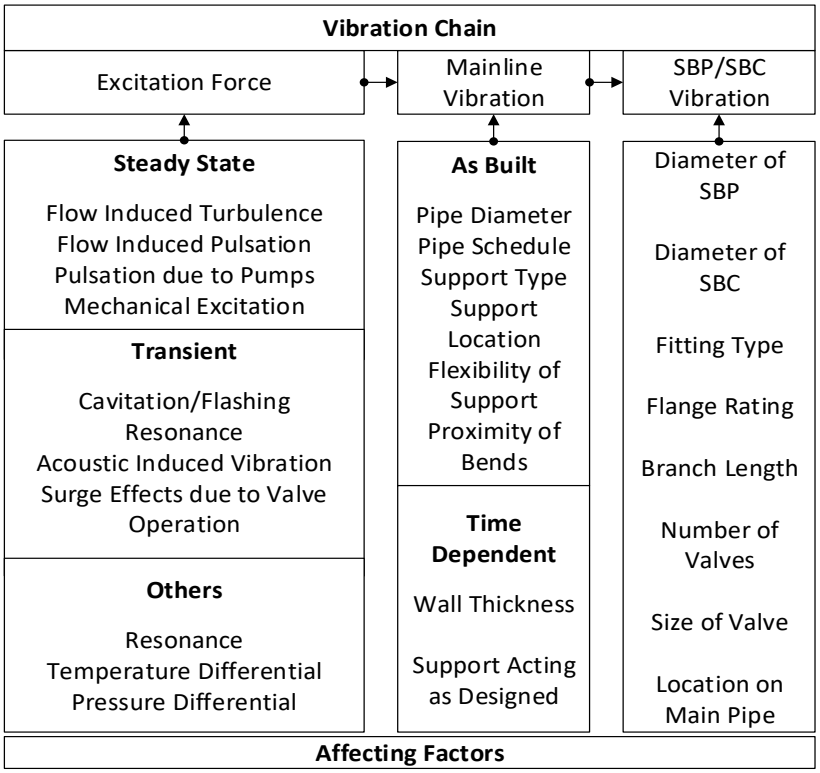

Fig. 6. Vibration Induced Fatigue Failure Chain (adapted from Swindell, 2003).

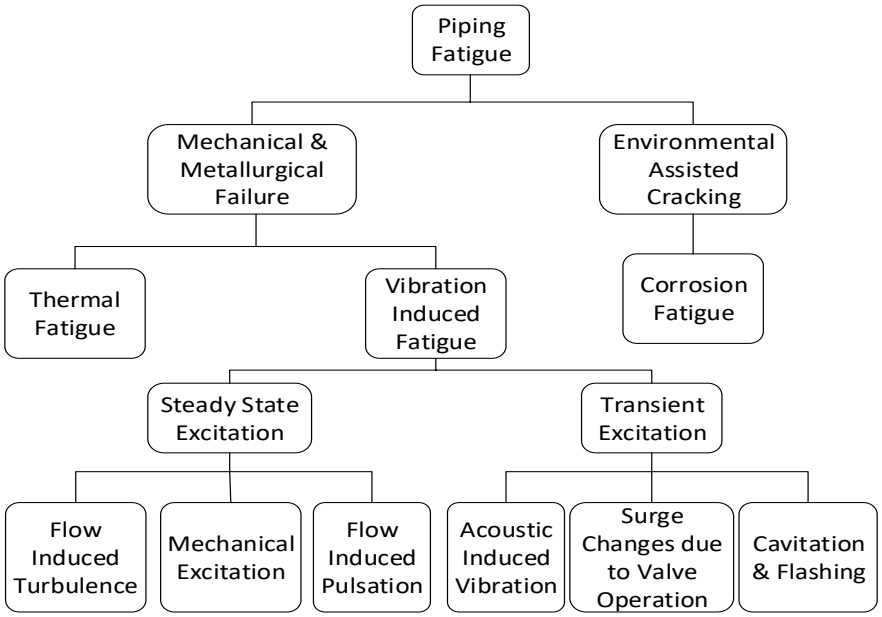

Fig. 7. Taxonomy of the piping fatigue.

3. Digital Twin for Small Bore Piping

As discussed, a DT is the most realistic representation of a physical asset, which incorporates models and all necessary available information (operational, organizational, and technical) of the asset. Some of the areas where a DT can be used effectively in the process industry is for the early anomaly detection of production process parameters, prognostics and health management (PHM) of assets, support for decision-making on process safety management (PSM) (EL Saddik, 2018; Tao et al., 2019). Lately, in the process industry the concept of DT is being applied for predictive maintenance of the various assets such as pumps, heat exchangers, etc. (Min, et al, 2019). However, when it comes to usage of DT technology for SBP, very little work has been done in this direction although SBP is the most commonly failed equipment. The primary challenge for the implementation of the DT for PHM of SBP is due to lack of sensor data and the current industry practice of selecting fatigue critical inspection location based on ad-hoc basis. The lack of sensor data stems from the fact that on a typical offshore platform there could be more than 100,000 SBPs, hence placing a physical sensor for all these SBPs would be practically impossible. Thus, if a DT needs to be built for the SBP, then it is vital to forge a connection between the real-world observations (for example from experiments, inspections) and the information obtained from the virtual sensors (data driven models) on the virtual model (CAD model or Simulink model). An analogy that fits this scenario is the one given by Allen and Tildesley (1987) in the book Computer Simulations of Liquids, where they demonstrate how coupling experiments (and experience), with theory and computer simulations, can enhance the up-to-date knowledge of a real system (See Fig. 8). Similar connections between real-world observations and data-driven models are equally relevant while making a DT for the SBP, so that an analyst /inspection engineer can get an up-to date health state of the asset which could then be used to estimate risk and frame I&M strategies. Hence, while building a DT for the SBP, it is necessary to combine the expert knowledge (either given in the standards or available from the practicing engineers), the data (coming from inspection, experiments) and the virtual models.

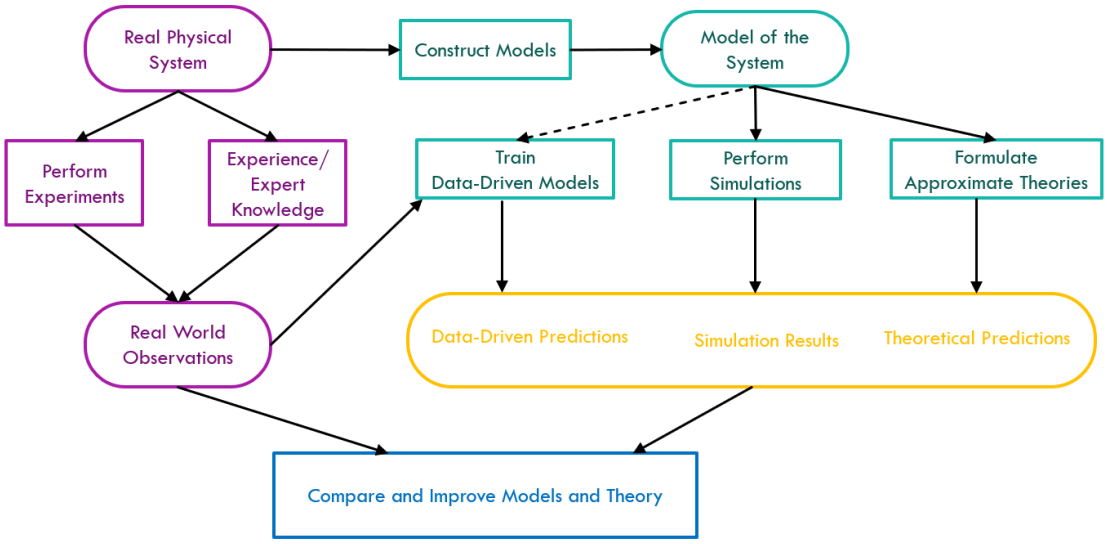

Fig. 8. Coupling real world observations with information from virtual model (Adapted from Allen and Tidesley, 1987)

Modeling methods can either be physics-based (they try to capture the actual physics of the process e.g., CFD, FEA) or data-driven models (they try to find trend based upon the data e.g., machine learning, deep learning). It is a common practice in process industry to perform risk analysis of engineering assets by leveraging the results from CFD and FEA techniques owing to its ability to simulate the physical phenomenon with great accuracy. For example., using CFD dispersion modelling, Fiates and Vianna (2016) performed risk assessment of gas leakage on a typical offshore topside plant. Kumaresh et al. (2016) calculated transient ignition probability of gas leaks in enclosure using CFD, Li et al. (2019) presented a CFD based approach to describe the behavior of underwater gas release and dispersion from subsea gas pipelines leak. However, as argued by Keprate et al. (2019) the higher accuracy of the advanced dynamic simulation techniques is achieved at the expense of the greater time required for simulations to run, which makes these techniques computationally expensive and time-consuming. In order to overcome the aforementioned shortcomings of CFD and FEA, recently engineering community has resorted to surrogate models (SMs) as they are faster than the main simulation code (i.e. FEA, CFD), while still being usefully accurate (Keane, Forrester et al. 2008). Many researchers such as Forrester et al. (2006), Moore et al. (2016), Wilson et al. (2017), Loy et al. (2018), Jeon et al. (2019), and Keprate et al. (2017), Bagalkot et al. (2021) have used SMs to replace CFD simulations for various engineering applications and demonstrated the accuracy and reduced computational expense of SMs. Thus, while building a DT for the SBP, a hybrid method which consists of physical modeling (using CFD and FEA) and data driven modeling (consisting of machine learning models) can be used. Furthermore, when the DT is to be used for risk assessment of safety critical equipment, it is necessary to account for uncertainty in the physical phenomenon and thus combine the above two modelling methods with the probabilistic graphical models such as Bayesian Networks (Hafver et al., 2017). Fig. 9 demonstrates the components used for building DT for estimating the RFL for the SBP. As can be seen the process flow parameters such as pressure, flow velocity serves as an input to the DT model. In reality the value of these parameters can be continuously communicated to the DT using principles of IoT, however, for the sake of illustration, in the case study, authors shall use

discrete values of these parameters. The next step is to build a virtual sensor for stress estimation using CFD and FEA coupled with Machine learning models. The input to the virtual sensor is pressure and flow velocity and the output is stress value at the SBC. Thereafter, the calculated stress acts as one of the inputs (along with initial crack size, critical crack size and material parameters) to the empirical crack growth model (Paris law) which is then coupled with a Bayesian Network model to estimate the probability density function (PDF) of the RFL of the SBP. Finally, the Cumulative Density Function (CDF) is used to establish the reliability curve which in turn is used to frame inspection intervals for SBP.

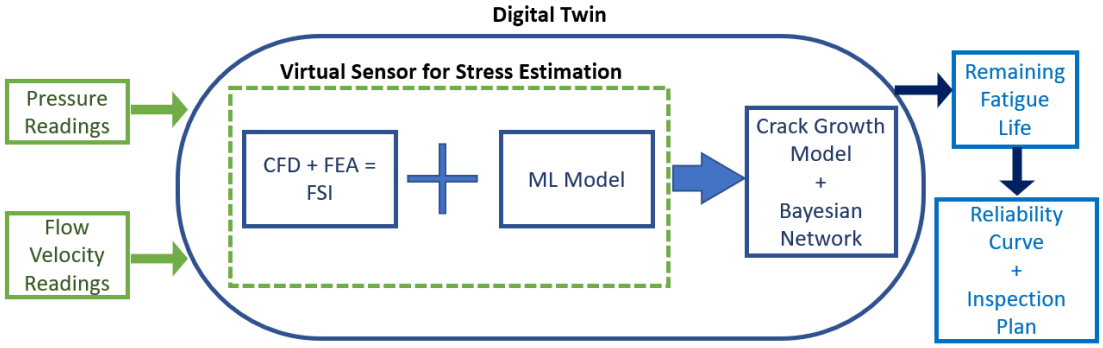

Fig. 9. Digital Twin Framework for Estimating RFL of SBP

4. Illustrative Case Study

4.1 General

For this case study, a DT for estimating the Remaining Fatigue Life (RFL) of SBP shall be developed using hybrid approach which includes physics based modeling (captured by CFD/FEA and analytical crack growth model), the surrogate models (using machine learning models) and logical models (Bayesian Newtork). AISI Type 304 Stainless Steel (SUS 304) having yield strength of 215MPa, is chosen as the material of the process piping. A semi-elliptical surface crack (having crack depth of a, and crack length of $2c$) is assumed to be initiated at the weld toe of the SBC, which shall propagate under the influence of the stress acting at the weld toe interface. A total of 4 small bore sizes have been used for a fixed size of main pipe. For a 10-inch (Nominal 80S) size main pipe, four small bore size 0.5-inch, 1-inch, 1.5-inch, and 2-inch (all Nominal 80S) have been used as shown in Table 1.

Table 1: *Details of the size of the main pipe and small bore*

DN (inch)	DN (mm)	OD (mm)	ID (mmm)	Thickness(mm)
0.5 (small-bore)	15	21.3	13.84	3.73
1 (small-bore)	25	33.4	24.3	4.55
1.5 (small-bore)	40	48.3	38.14	5.08
2 (small-bore)	50	60.3	49.22	5.54
10 (main pipe)	250	273.1	247.7	12.7

The first step in creation of the DT is to create a virtual representation of the piping system and use this virtual model for FEA and CFD analysis to estimate the stress acting on the SBC. Since CFD and FEA simulations are computationally expensive, so a surrogate model is used as a virtual model, for estimating stress. The detailed methodology for stress prediction using CFD, FEA, and ML approach as shown in the Figure 3 is discussed. ANSYS software is used to perform Fluid Structure Interaction (FSI) by coupling CFD and FEA analysis. The details of each step in the methodology shall be presented in the upcoming sections. As shown in Fig 10, first the CFD analysis shall be carried out, the results of this analysis are then transferred to a FE analysis tool, which will estimate the stress distribution on the weld toe of the SBP. Thereafter nineteen different Machine Learning (ML) algorithm are trained, tested and validated using the data generated from the CFD/FEA analysis. The most accurate ML algorithm (Optimizable Gaussian Process Regression in our case) is then utilized to predict value of the stress for 10000 samples (of pressure and velocity variables) generated using Latin Hypercubic Sampling (LHS). Thereafter, different distributions are fitted on the predicted maximum stress value and the Akaike Information Criterion (AIC) value is calculated to determine the best distribution fit for the stress acting on SBP.

4.2 Physics Based Modeling

4.2.1 Creation of the Model:

A single small-bore attached to the main pipe configuration is used to carry out the analysis in the current study. The dimensions of the main pipe and small-bore is adopted from Energy Institute (EI, 2008). Fig. 11 shows the 3D CAD model with a small-bore and main pipe used to carry out both CFD and finite element (FE) analysis. The 3D CAD model is segmented into a fluid region representing the fluid flow and a solid region representing the pipe body.

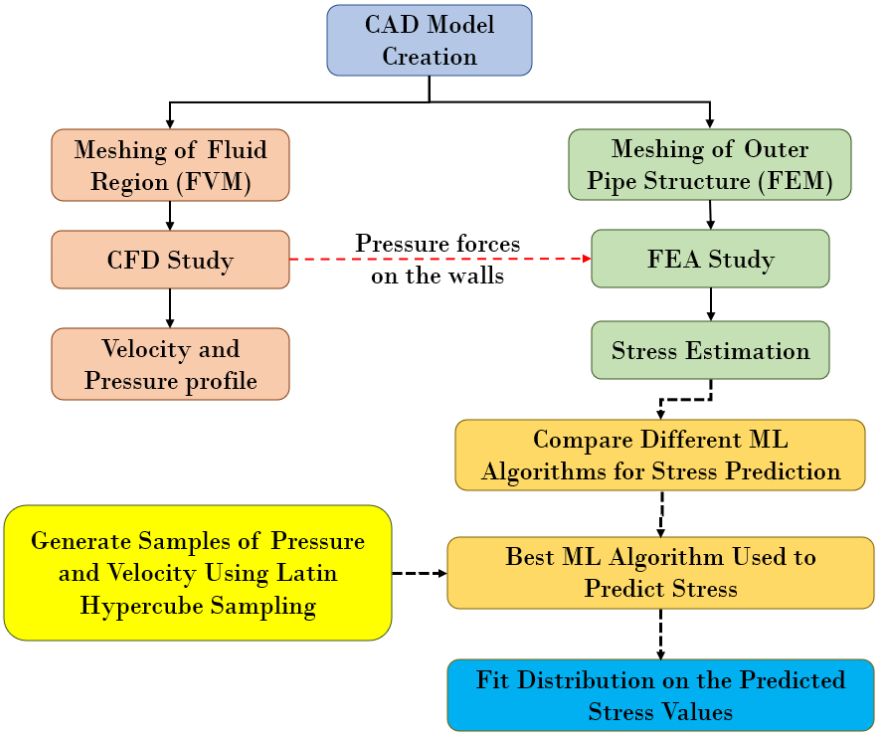

Fig 10. Methodology used for stress estimation (Using Virtual Sensor)

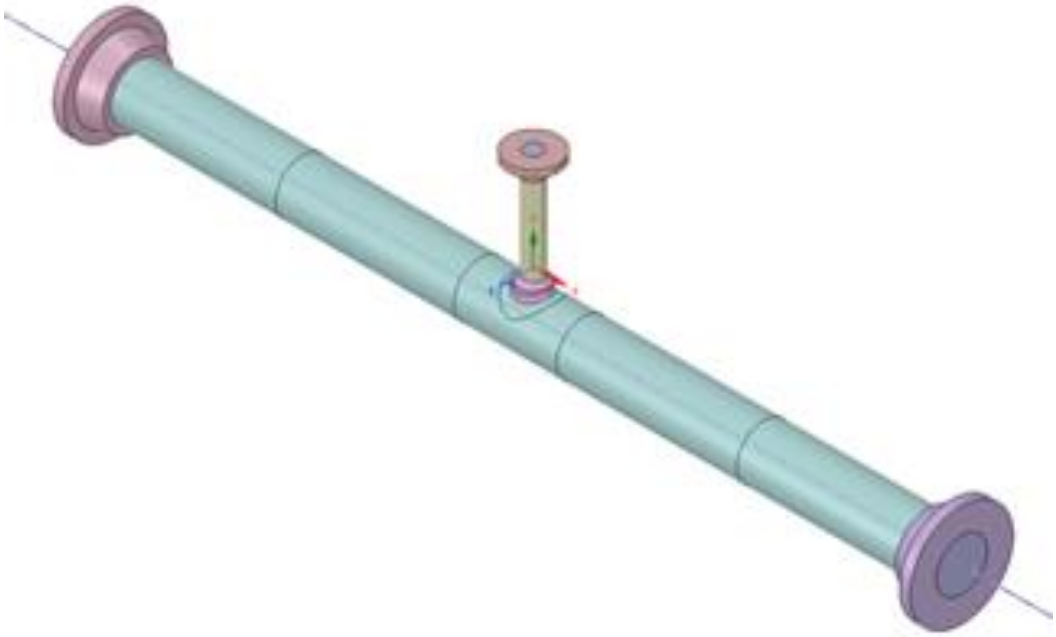

Fig 11. CAD model of mainline piping and SBP

4.2.2 Model Meshing

For creating high fidelity physics based model using CFD/FEA, meshing plays a very important role. Too coarse mesh can lead to lowering the accuracy of the model, while too fine mesh can significantly increase the computation time. Thus a mesh sensitivity study was performed for this work. The 3D model shown in the Fig 11 was meshed using ANSYS Meshing Tool. Since there are two modes of simulation CFD and FEA, there will be two types of meshing finite volume and finite element. For the CFD analysis a finite volume mesh was created on the fluid flow region. Fig 12b shows the mesh distribution at the junction of small bore and main pipe, the distribution of inflation layers as shown in Fig 12c, along with mesh details. The mesh size was selected after a comprehensive mesh sensitivity study. The size of the first layer of the inflation was based on the $Y+ < 1$ criteria, a total of 12 layers of inflation with growth rate of 1.15 was used for CFD meshing. The minimum face size of the mesh was selected to be $8*10^{-5}$ m and max face size was restricted to $2*10^{-3}$ m, the increment in the size is based on the growth rate of 1.15. For the region represented by the structure of the main pipe and small bore, a finite element mesh is implemented. The nodes of the mesh on the pipes will be able to transfer the pressure loads arising from fluid flow and impact and transfer is as stresses within the pipe structure. The element distribution at the main pipe and small bore is shown in Fig. 13. Similar to CFD mesh the mesh size for stress analysis was selected after a comprehensive mesh sensitivity analysis. The supports were provided at the either ends of the main pipe.

4.2.3 CFD and FEA Analysis:

After the model is meshed, ANSYS FLUENT 2021 R2 is used to carry out a comprehensive CFD analysis. For each size of bore the CFD analysis is carried out range of operational pressure and velocity. Due to limited time availability for illustration purposes, the analysis was carried out for four operating pressures 5, 15, 25, and 35 bar, for each of these operating pressures, the analysis

was carried out for 7 operating flow rates, the flow rates are represented as velocity (keeping the inlet density and area constant) in the range of 1.5 m/s to 8.4 m/s. From the boundary condition perspective, a velocity inlet with inlet gauge pressure is used as inlet boundary condition, and pressure outlet is implemented at the exit to the main pipe. The small bore acts a dead-leg and has no flow exiting it. Compressible air is used for the simulation with Peng Robinson model. A two-equation k-epsilon, with scalable wall function turbulence model option was implemented. The pressure force applied by the fluid on the internal walls of the pipe calculated by the FLUENT software is transferred to the FEA software, to estimate the mechanical stress on the structure arising due to the pressure loads. The software is setup to estimate the Equivalent Von-Mises stress at the junction of main pipe and small bore. The stress on the internal and external wall of the SBP is estimated. The boundary condition includes fixed support at the end of the main pipe, and gravity effect, the influence of the weight of equipment at the end of small bore like check valves, pressure or temperature sensors, was neglected.

All the results of the simulation are of similar nature except from change in magnitude, therefore as a representative case the FSI (CFD and FEA) results for the case study of 1.0- inch small bore at fixed pressure of 5 bar and at three different flow rates will be presented in detail. For the rest of the cases the output FSI result (Von-Mises stress) will be reported, which will later be used to build the surrogate model for stress prediction.

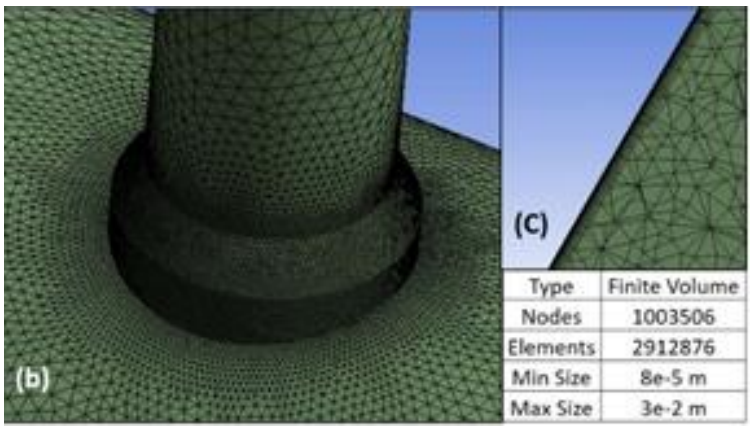

Fig. 12 Mesh distribution and mesh details for CFD analysis.

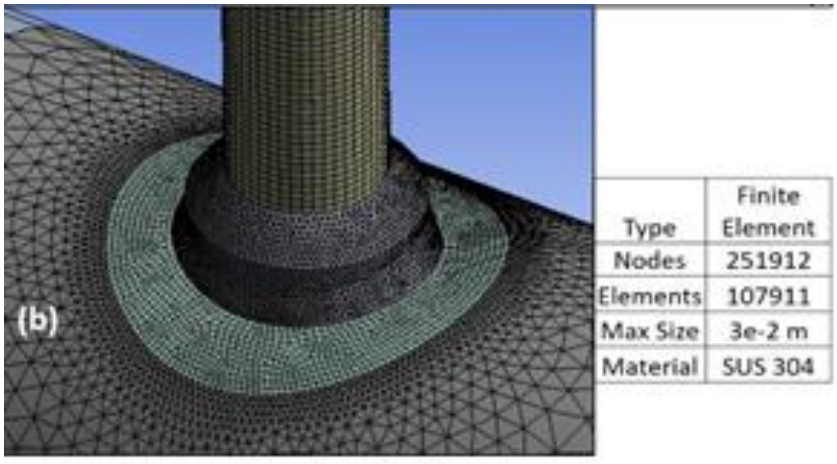

Fig. 13 Mesh distribution and mesh details for FE analysis.

Fig 14 shows the velocity vector at the junction of the main pipe and small-bore for 5 bar (1inch small bore, operational velocity 8.5 m/s). Fig 14a is along a central plane parallel to the flow direction (positive x-axis) and Fig 14b is along the plane perpendicular to the flow direction. The velocity vector gives a clear picture of where the stress might get induced in the pipe due to the flow. As can be seen a major turbulence region is created in the vicinity of junction, this turbulence gives rise to region of pressure fluctuations or uneven pressures as shown in Fig 15

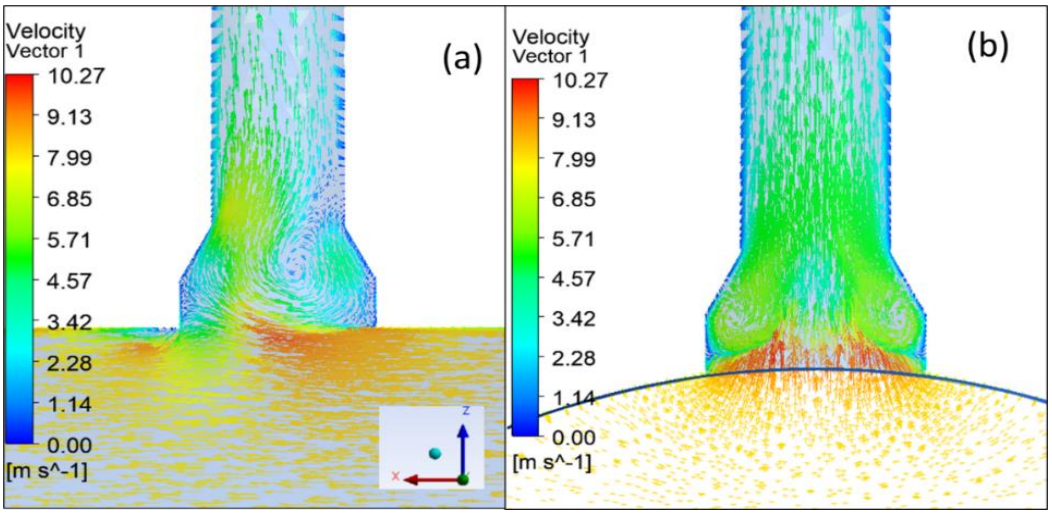

Fig. 14 Velocity vector at the SBC

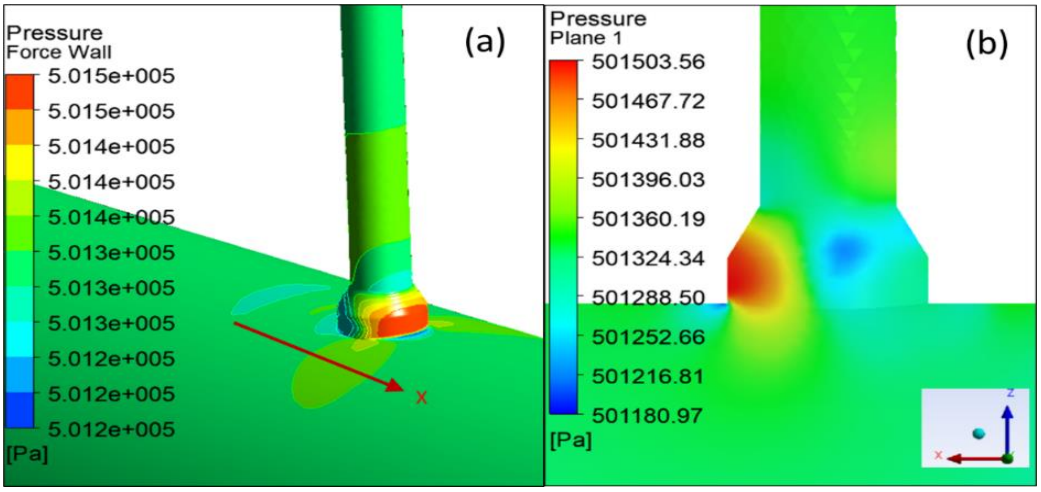

Fig. 15 Pressure Concentration at the SBC

Fig 16 shows the Von-Mises stress estimated by the ANSYS Mechanical software using the pressure loads from the ANSYS FLUENT software for flow condition mentioned in description of above CFD results. Fig 16a and Fig 16b show the stress distribution at the external and internal surface of the junction respectively and flow direction is in the positive X-direction. It may be observed that the maximum stress concentration, as well as stress distribution (regions of low and high stresses) stress is at SBC. The observations made in Fig 16 is supported from the flow and force results shown in Fig 14 and Fig 15. Furthermore, the stress is slightly higher on the inner diameter of the pipe

compared to outer diameter. Similarly, for a fixed size, the larger the flow rate, the greater is the stress at the junction of small bore and main pipe as can be seen in Fig. 16.

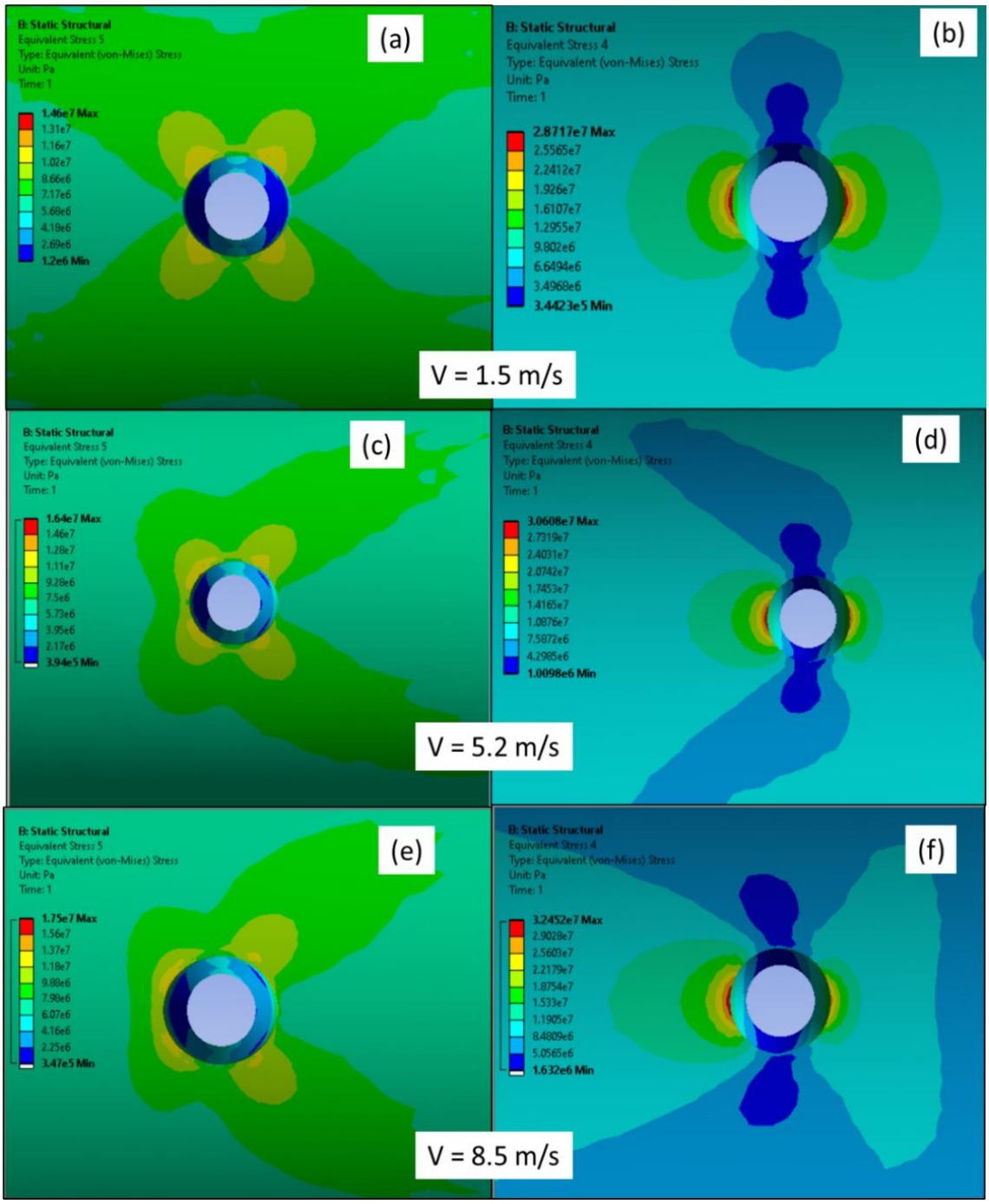

Fig. 16 Stress distribution at SBC at different velocity (1-inch small bore, 5 bar)

Although FEA/CFD is a good approach for emulating the physical phenomenon, but is can be seen that it is practically impossible to estimate stress at SBC for different values of parameters (pressure and volume flow rate). Thus, we shall employ surrogate model which once trained and tested can act as a replacement to the computationally expensive and/or time-consuming CFD/FEA simulations, without compromising the accuracy of the output.

4.3 Surrogate Modelling

The most commonly used SMs in the engineering domain are linear regression, Tree based methods, Ensemble methods, Boosting methods, Support Vector Machine and GPR. The mathematical background and theory of the various SMs used in this manuscript are discussed briefly (Keprate et al. 2017). The data generated from CFD/FSI simulations was used for training, testing and validation of these SMs. Eighty-Four data points are used to train the aforementioned SMs, while Twenty-Eight data points are used as for the validation purpose. Nineteen different SMs (consisting of linear regression, Tree based methods, Ensemble methods, Boosting methods, Support Vector Machine and GPR) are evaluated using 4 metrics, namely, Root Mean Square Error (RMSE), Maximum Absolute Error (MAE), Coefficient of Determination (R^2) and training time. Thereafter, using the flowchart shown in Fig. 17, different SMs are constructed to predict the value of the Maximum Stress. K-fold (10 folds in our case) cross validation is used during the training of the surrogate models. Afterwards, the most accurate SM is further optimized using Bayesian Optimization to further improve its accuracy. Finally, the optimized SM is trained on validation dataset and the predicted values of maximum stress are compared to the values obtained from CFD and FEA.

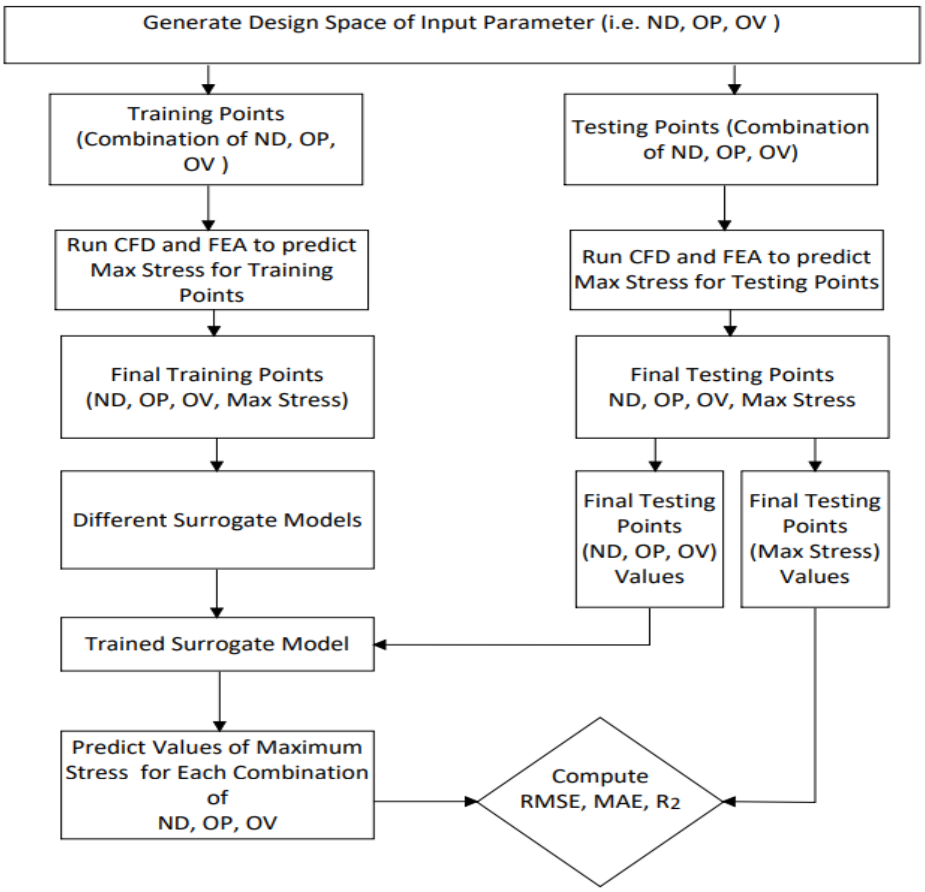

Fig. 17 Flowchart to build Surrogate models for Maximum Stress prediction

In order to compare the performance of the SMs, four metrics, namely, Root Mean Square Error (RMSE), Maximum Absolute Error (MAE), coefficient-of-determination (R^2) and training time are used. Mathematically, these are written as:

$$RMSE = \left(\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\right)^{0.5}$$

$$MAE = \max(|y_1 - \hat{y}_1|, |y_2 - \hat{y}_2|, .., |y_n - \hat{y}_n|) \quad (1)$$

$$R^2 = 1 - \frac{SS_{Error}}{SS_{Total}}$$

Generally, the most efficient algorithm should have minimum value of RMSE, MAE and training time, and the value of R^2 should be closer to 1. The comparison of different SMs is presented in Table 2 and it can be seen that Ensemble Bagged trees is the worst performing algorithm while, Gaussian Process Regression (GPR) having a squared exponential kernel function is the best performing algorithm.

Table 2. Comparison of different Surrogate Models

Surrogate Model	***RMSE (MPa)***	***MAE (MPa)***	**R^2**	**Training Time (s)**
Linear Regression	7.91	5.87	0.9	2.41
Linear Regression with Interaction	3.99	5.71	0.97	0.50
Robust Linear Regression	7.97	7.87	0.89	0.58
Stepwise Linear Regression	3.99	3.03	0.97	2.06
Fine Tree Algorithm	13.99	11.26	0.68	0.69
Medium Tree Algorithm	16.81	12.86	0.53	0.17
Coarse Tree Algorithm	19.01	14.85	0.4	0.17
Ensemble Boosted Trees	7.54	4.9	0.91	2.05
Ensemble Bagged Trees	24.41	19.76	0.01	1.65
Linear SVM	8.00	5.63	0.89	0.84
Quadratic SVM	3.08	2.46	0.98	0.30
Cubic SVM	2.21	1.83	0.99	0.24
Fine Gaussian SVM	18.17	13.86	0.45	0.29
Medium Gaussian SVM	5.51	3.65	0.95	0.21
Coarse Gaussian SVM	7.67	5.03	0.90	0.28
Rational Quadratic GPR	2.02	1.40	0.99	0.98
Squared Exponential GPR	2.01	1.39	0.99	0.46
Matern 5/2 GPR	2.04	1.42	0.99	0.53
Exponential GPR	3.31	2.12	0.98	0.46

It is further possible to enhance the accuracy of GPR, by performing Bayesian optimization. In total 30 iterations are performed across the entire hyperparameter space (such as Basis Function, kernel function, kernel scale etc.) of the GPR to arrive at the optimized parameters (which are constant basis function, non-isotropic matern 5/2 kernel function, 34.89 kernel scale). The value of evaluation metrics for optimized GPR are RMSE of 1.81MPa, MAE of 1.28MPa and R^2 of 0.99. However, the training time increased to 102.99 seconds, which is logical as 30 iterations were

performed with different values of hyperparameters, to arrive at the optimized parameter values. Furthermore, the plot comparing the stress values predicted by CFD/FEA and GPR are shown in Fig. 18, in which true response represents the former stress values while the predicted response represents the stress value estimated by optimized GPR. It is seen that for most of observations GPR predicts accurately the value of stress. However, it is vital to check the optimized GPR performance on the validation dataset. As can be seen from Fig. 19 and also inferred from the value of RMSE which is 3.012MPa, and R^2 which is 0.974, thus performance of optimized GPR is good even on validation dataset. The, trained optimized GPR shall now serve as a virtual/soft sensor which can be used to predict maximum stress for SBP for the given flow conditions.

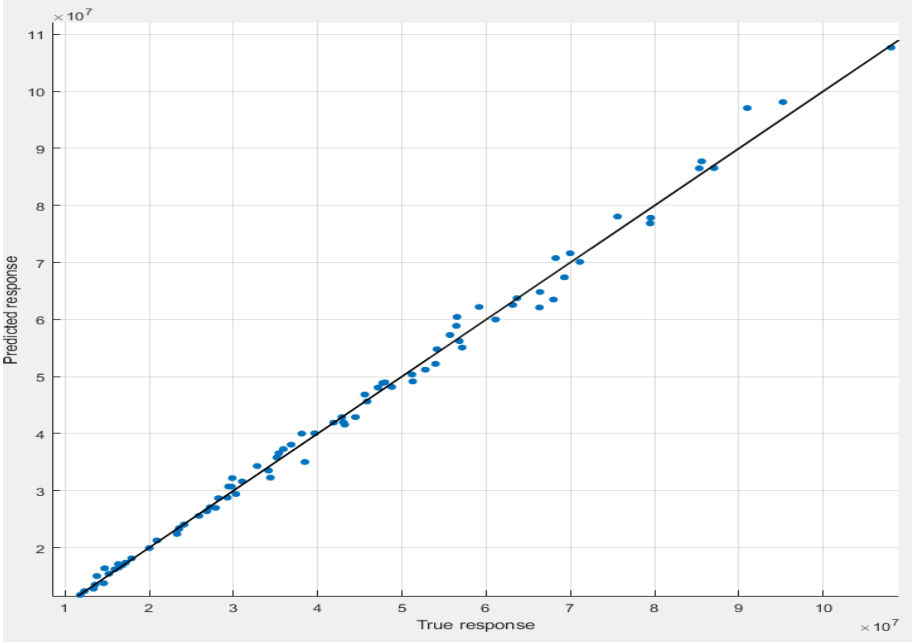

Fig. 18 True vs. predicted stress values for training dataset

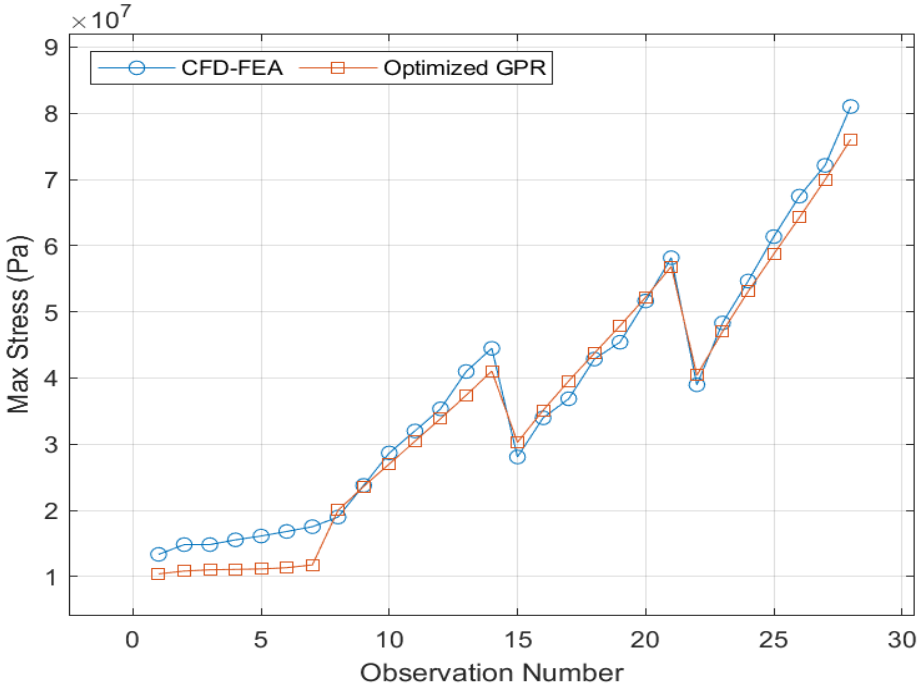

Fig. 19 Maximum stress values predicted for validation dataset

If the flow data and pressure readings are obtained in real time from the process facility, then they can serve as in out data to the developed soft sensor. However, if it is not possible then we can utilize various sampling techniques to generate the input data. For illustrative purpose, Latin Hypercubic Sampling (LHS) is utilized to obtain 10000 samples of pressure variable (between 5bar and 35 bar), and of velocity variable between 1.5m/s and 8.7m/s, which act as input to the virtual sensor, which in turn estimates the value of max stress for these 10000 samples for a 2inch (50mm Nominal Diameter) SBP. Thereafter, different distributions are fitted on predicted maximum stress value. The two best fitted distributions are lognormal (log location = 3.37 and log scale = 0.388) and 2-parameter Weibull distribution (scale = 35.23, shape r = 2.86) as can be seen in Fig.20. The Akaike information criterion (AIC) value was calculated to determine of the quality of the distribution fit. Generally, the distribution with the low AIC value is usually the preferred distribution. Amongst the two distributions, lognormal distribution was the best fit for the given, as its AIC value was lower than Weibull distribution (AIC_LN = 76939, while AIC_Wbl = 77113).

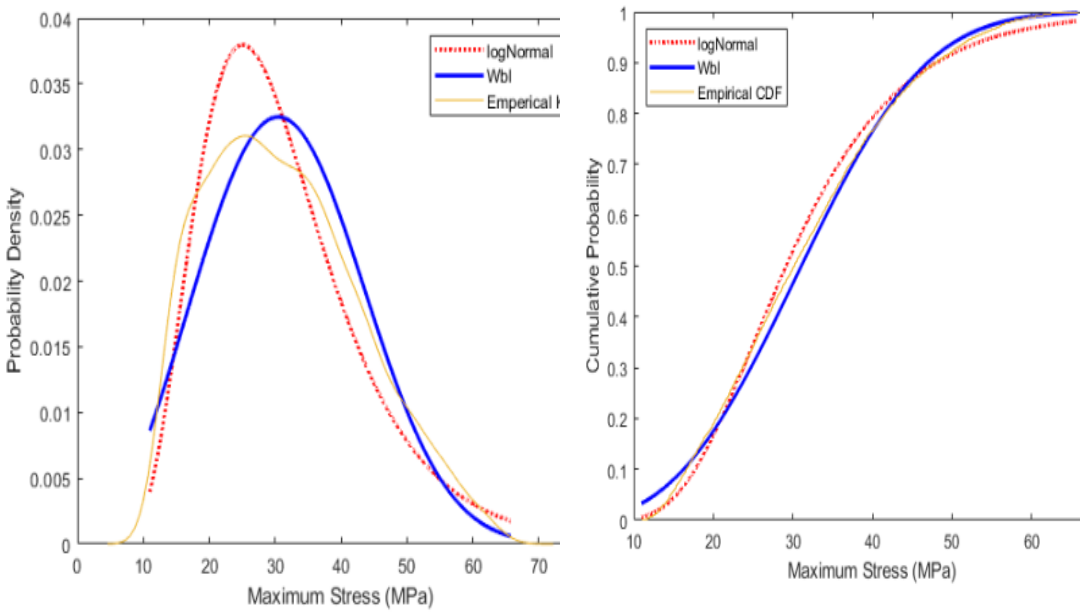

Fig. 20 Distribution fitting on stress values estimated from the Virtual Sensor

4.4 Bayesian Network Modeling for Estimating RFL

The next step is to use probabilistic graphical model i.e. Bayesian Network (BN) to build a DT for estimating the RFL of SBP. The main component of the BN is the empirical Paris crack growth law, which captures the physics behind the crack propagation. Paris law is stated as:

$$\frac{da}{dN} = C(\Delta K)^m \quad for\ \Delta K_{th} \leq \Delta K \leq K_{mat} \tag{1}$$

where a is the crack size in m, N is the number of cycles, $\frac{da}{dN}$ is the increment in the crack growth rate in m/cycle, ΔK is the stress intensity factor (SIF) range in MPa$\sqrt{m}$, ΔK_{th} is the threshold value of the SIF, $K_{mat}$is the material fracture toughness, and m and C are the material parameters. The range of the SIF depends upon the parameters such as the range of the remote stress $(\Delta\sigma)$, crack size (a) and the geometric function (Y). The number of fatigue loading cycles (RFL)

required to grow a crack from the (ICS), a_i, to the final crack size, a_N, can be determined by integrating Eq. 2.

$$RFL = \frac{a_N^{1-\frac{m}{2}} - a_i^{1-\frac{m}{2}}}{C\left(1-\frac{m}{2}\right)(\Delta\sigma\, Y\, \sqrt{\pi})^m} \qquad (2)$$

The various parameters of the Paris model are considered as random variables, and are represented by a suitable probability distribution and associated parameters as shown in Table 3. For our analysis we consider the uncertainty only in three parameters namely, initial crack size (a_i), material parameter (C) and the remote stress range ($\Delta\sigma$), while other parameters (critical crack size and material parameter (m)) are treated as non-random variables. The value of initial crack size is calculated using Equivalent Initial Flaw Size (EIFS) concept given in (Sankararaman et al., 2011), while the value of parameter C is taken from the literature. Furthermore, since Paris model is an empirical law, hence it is also subjected to uncertainty which is represented by a modelling error ε_r whose value in practice is obtained from the experiments. However, in this chapter for the sake of illustration, ε_r is represented by a 5% Gaussian white noise. The value of remote stress range ($\Delta\sigma$) is obtained from the distribution fitting to the data obtained from the virtual sensor developed earlier in the chapter, while the value of geometric function Y is calculated using suitable industrial standards. All the yellow color nodes are input to the BN, while orange color node represents the crack size evolution with time. The BN can finally generate the distribution of RFL (using 10000 samples) which is shown in Fig. 21. It must be mentioned here that for the sale of illustration, authors are manually assigning values to the various parameters of the BN, however, in real world implementation of the DT, the entire data input can be automated.

Table 3. Uncertainty Quantification for RFL estimation

Uncertainty Source	Random Variable	Parameter Value
Initial crack size	a_0 (mm)	LN (0.47, 0.0333)
Material parameter	C	LN (-21.9, 0.3708)
Material parameter	m	3.5
Remote stress range	$\Delta\sigma$ (MPa)	LN (3.37, 0.388)
Critical crack size	a_c(mm)	12.7
Geometric function	Y	0.952

A closer look at the Probability Density Function (PDF) of the RFL tells us that the mean value of the RFL equals to 2.7492e06 cycles, while the 95% confidence interval (CI) is calculated as [6.1557e04 and 1.68e07] cycles. The estimated RFL can be used to estimate the reliability (in the form of reliability curve as shown in Fig 22) of SBP, which in turn can be used to frame the future inspection intervals. In this method the operator sets the target reliability depending upon the limit states and the safety class. For example, if we consider fatigue limit state and high safety class then the target reliability level for the piping system should be 0.9999. From Fig. 23, the corresponding number of cycles for a target reliability of 0.9999 approximately equals to 20000 cycles (as shown by red colored line on the figure). Thus, the first inspection of 2inch SBP must be done after 20000 cycles. The same DT concept can be used to frame inspection intervals of different sizes of SBPs (0.5-inch, 1 inch etc.) and in this way the inspection plan of SBPs for entire process plant can be

generated in a short period of time. Such a proactive strategy would help inspection engineers to timely inspect the SBP and thus help in preventing unwanted breakdown due to fatigue, thus enhancing process safety.

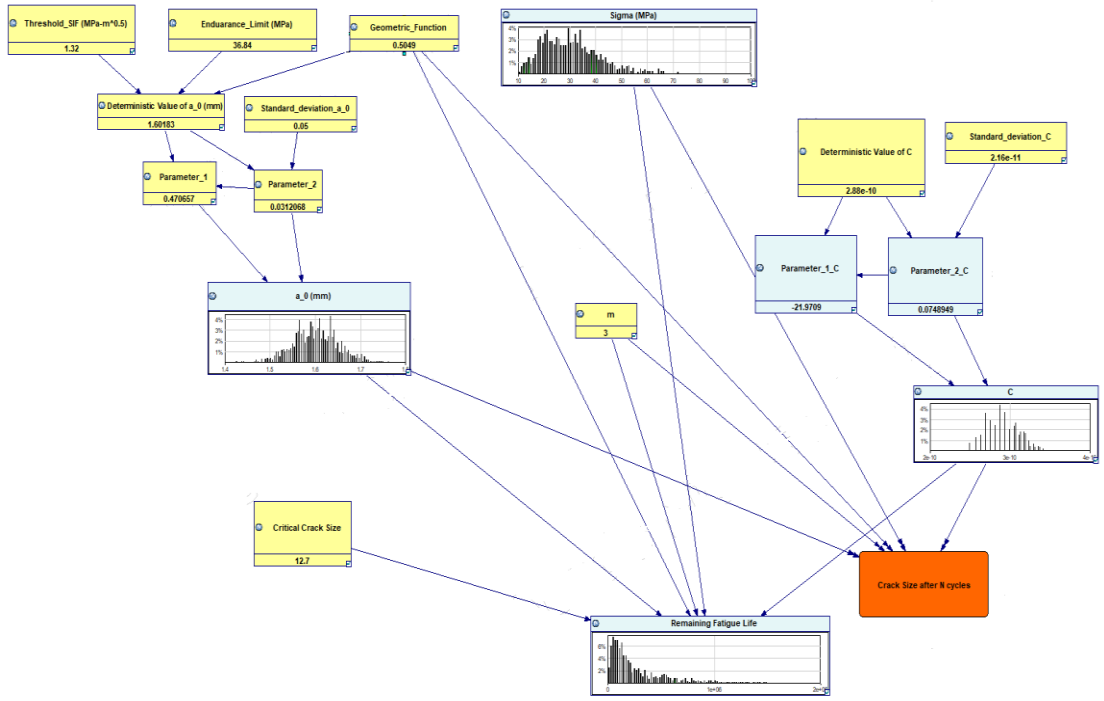

Fig. 21 Bayesian Network Enabled DT for RFL Estimation of SBP

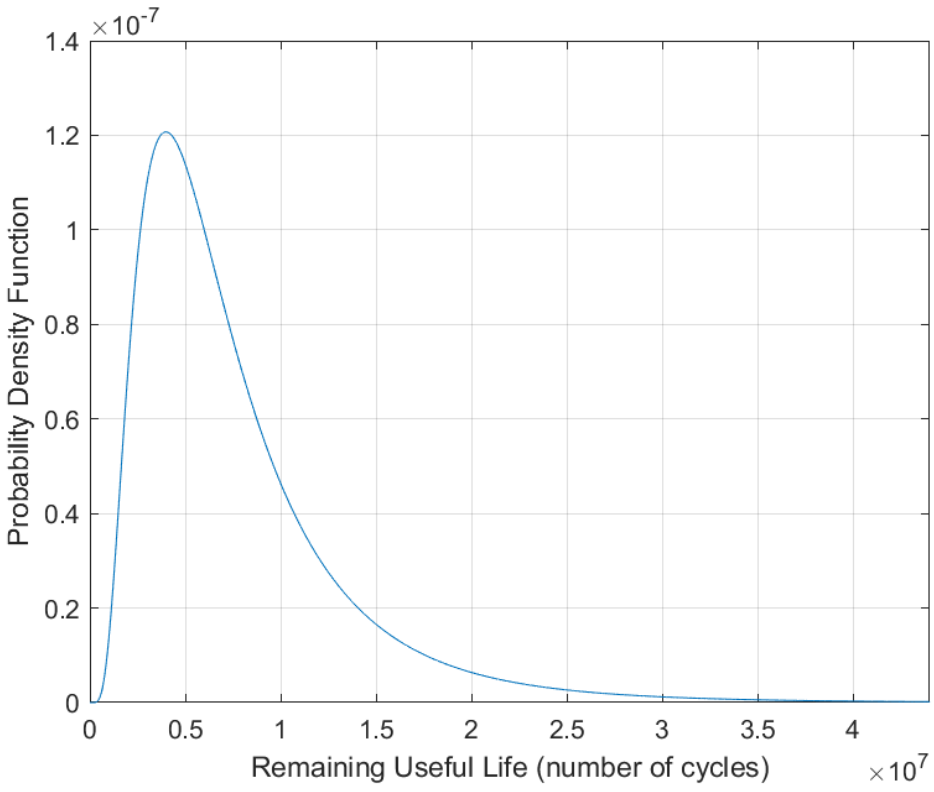

Fig. 22 Probability density function of RFL of SBP (10000 samples)

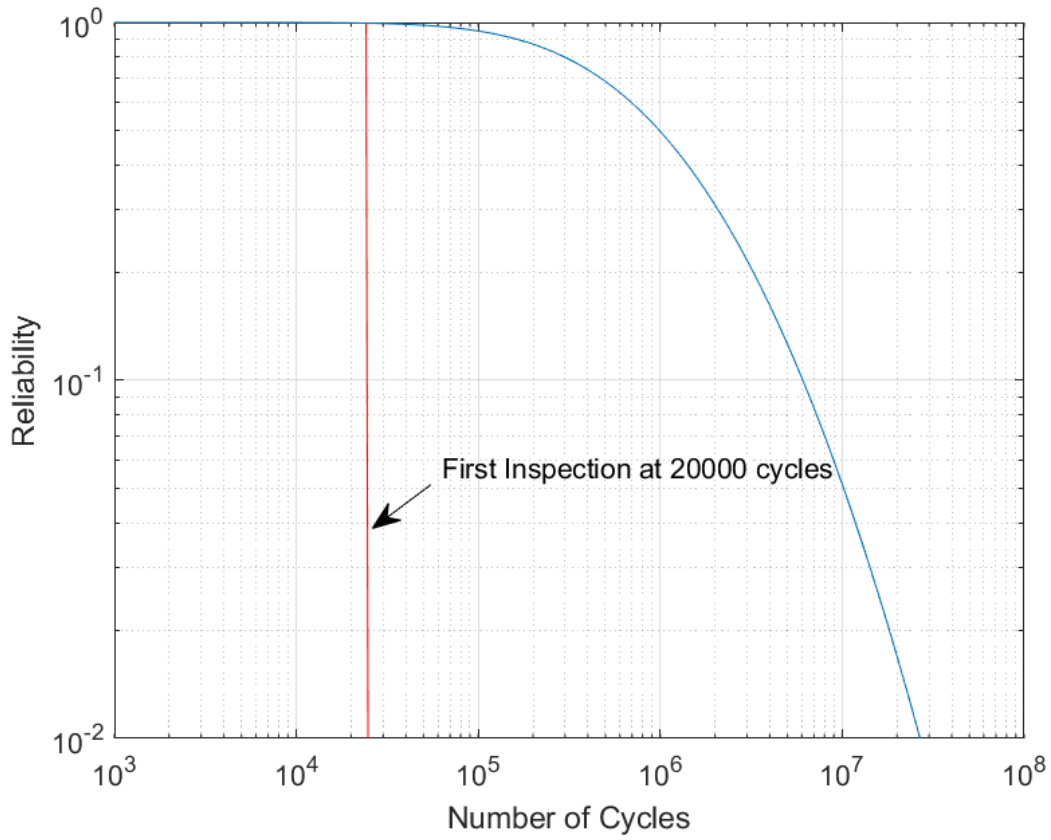

Fig. 23 Reliability Curve for the 2inch SBP (10000 samples)

Conclusion:

In this chapter, a Digital Twin Framework for estimating the Remaining Fatigue Life of Small Bore Piping was developed. Due to large number of SBPs on a typical process facility, the problem of placing a physical sernsor was also discussed. As a solution to the aofrementioned problem, a virtual sensor for utilizing CFD/FEA and Surroagte model was developed. The input to the virtual sensor were process parameters (i.e. presure and flow velocity) while maxium Von-Mises stress at Small Bore Connection (SBC) was the output. The unceratinty in the estimated stress was catered by fitting the most suitable probability disribution (logormal distrubtion for our case) to the stress value. Thereafter, crack propagation in the piping was modeled using Probabilistic Paris Law, which had input parameters such as intial crack size, critical crack size, material parameters and sress. A Bayesian Newtork was developed for representing the probabilistic crack growth phenomenon and estimating the RFL. Finally, reliability curve coupledwith threshold reliability value, was utilized to arrive at the inspection plan of the SBP. In this chapter the input values to the virtual ensors were entered manually for the sake of demonstration, howver, if the DT is deployed on cloud, it is also possible to direlty feed real time process parameters int the DT and estimate RFL. Likewise, more input process parameters such as tempterature, pH etc. can aslo act as input to the DT. Although the DT framework has been deeloped for RFL estimation of SBPonly, it can easily be extended to other assets on a praocess faiclity and also for other degradation mechanisms such as corssion, erosion etc.

Reference

Allen. M.P., and Tildesley. D.J. (1987). Computer Simulation of Liquids. Oxford Scholarship.

Bagalkot, N., Keprate, A., Orderloøkken R. (2021). Combining Computational Fluid Dynamics and Gradient Boosting Regressor for Predicting Force Distribution on Horizontal Axis Wind Turbine. Vibration **4**(1): 248-262.

Baggett, R. T. (2019). Inspect Your Primary Process Small Bore Piping the Right Way, Using Profile Radiography. Inspectioneering Journal **25**(2): 34-37.

Botín-Sanabria, D.M.; Mihaita, A.-S.; Peimbert-García, R.E.; Ramírez-Moreno, M.A.; Ramírez-Mendoza, R.A.; Lozoya-Santos, J.d.J. (2022). Digital Twin Technology Challenges and Applications: A Comprehensive Review. Remote Sens., 14, 1335.

Campos-Ferreira, A.; Lozoya-Santos, J.J.; Vargas-Martínez, A.; Mendoza, R.; Morales-Menéndez, R. (2019). Digital Twin Applications: A review. In Memorias del Congreso Nacional de Control Automático; Asociación de México de Control Automático: Puebla, Mexico,; pp. 606–611.

DNV, (2020). https://www.dnvgl.com/oilgas/laboratories-testsites/article/tubes-and-piping-are-the-most-failureprone-components.html

EI Guidelines, (2007). Guidelines for the avoidance of vibration induced fatigue failure in process pipework. The Energy Institute London, UK.

EI Guideline, (2013). Guidelines for the Design, Installation and Management of Small Bore Tubing Assemblies, The Energy Institute London, UK.

Fiates, J. and S. S. Vianna (2016). Numerical modelling of gas dispersion using OpenFOAM. Process Safety and Environmental Protection **104**: 277-293.

Forrester, A. I., Bresslof, N.W., and Keane. A.J. (2006). Optimization using surrogate models and partially converged computational fluid dynamics simulations. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **462**(2071): 2177-2204.

Grieves, M.W. (2014). Digital Twin: Manufacturing Excellence through Virtual Factory Replication. White Paper.

Grieves, M.W., & Vickers, J.H. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. Transdisciplinary perspectives on complex systems, 85-113.

HSE-OTR 028, (2002). Offshore hydrocarbon releases statistics and analysis. HID Statistics Report by Health Safety Executive, UK.

Haag, S. and R. Anderl (2018). "Digital twin – Proof of concept. Manufacturing Letters **15**: 64-66.

Jeon, K., Yang, S., Kang, D., Na, J., and Lee, W.B. (2019). Development of surrogate model using CFD and deep neural networks to optimize gas detector layout. Korean Journal of Chemical Engineering **36**(3): 325-332.

Keprate, A., Ratnayake, R.M.C., and Sankararaman, S. (2017). "Adaptive Gaussian process regression as an alternative to FEM for prediction of stress intensity factor to assess fatigue degradation in offshore pipeline." International Journal of Pressure Vessels and Piping **153**: 45-58.

Keprate, A., Ratnayake, R.M.C., and Sankararaman, S. (2017). Comparing different metamodelling approaches to predict stress intensity factor of a semi-elliptic crack. International Conference on Offshore Mechanics and Arctic Engineering, American Society of Mechanical Engineers.

Keprate, A., Ratnayake, R.M.C., and Sankararaman, S. (2019). Experimental validation of the adaptive Gaussian process regression model used for prediction of stress intensity factor as an

alternative to finite element method. Journal of Offshore Mechanics and Arctic Engineering **141**(2).
Kumaresh, G. R., Fossan, I., Venkatraman, M.M., Gilarhus, .K.E., and Jensen, S. (2016). CFD-based transient ignition probability modeling of gas leaks in enclosures. International Journal of Computational Methods **13**(02).
Li Xinhong, Chen Guoming, Zhang Renren, Zhu Hongwei, Fu Jianmin, (2018). Simulation and assessment of underwater gas release and dispersion from subsea gas pipelines leak. Process Safety and Environmental Protection, 119, 46-57.
Loy, Y., Rangaiah, G.P., and Lakshminarayan, S. (2018). Surrogate modelling of net radiation flux from pool fires in a hydrocarbon storage facility. Process Safety and Environmental Protection **114**: 296-309.
Min Q., Lu Y., Liu Z., Su C., Wang B. (2019). Machine Learning based Digital Twin Framework for Production Optimization in Petrochemical Industry, International Journal of Information Management, 49, 502-519.
Moore, W. O. W., Jetmarova, H.M., Gebreslassie, M., Tabor, G.R:, Belmnot, M.R., Savic, D.A. (2016). Comparison of multiple surrogates for 3D CFD model in tidal farm optimisation. Procedia Engineering **154**: 1132-1139.
West T, Blackburn M. (2017). Is digital thread/digital twin affordable? A systemic assessment of the cost of DoD's latest Manhattan project. In: Complex adaptive systems, Chicago, USA.
Wilson, B., Wakes, S., Mayo, M. (2017). Surrogate modeling a computational fluid dynamics-based wind turbine wake simulation using machine learning. IEEE Symposium Series on Computational Intelligence (SSCI).
Rasheed. A., San. O. and Kvamsdal. T. (2020). Digital Twin: Values, Challenges and Enablers From a Modeling Perspective. IEEE Access, 8,21980-22012.
Saddik. A. El. (2018). Digital Twins: The Convergence of Multimedia Technologies. *IEEE MultiMedia*, 25(2), 87-92.
Swindell, R., (2003). Vibration fatigue in process pipework: A risk-based assessment methodology. Bureau Veritas, Southampton, UK.
Tao. F., Zhang. M. and Nee. A. Y. C. (2019). Digital Twin Driven PHM. In: Digital Twin Driven Smart Manufacturing, edited by F. Tao, M. Zhang and A. Y. C. Nee Academic Press.