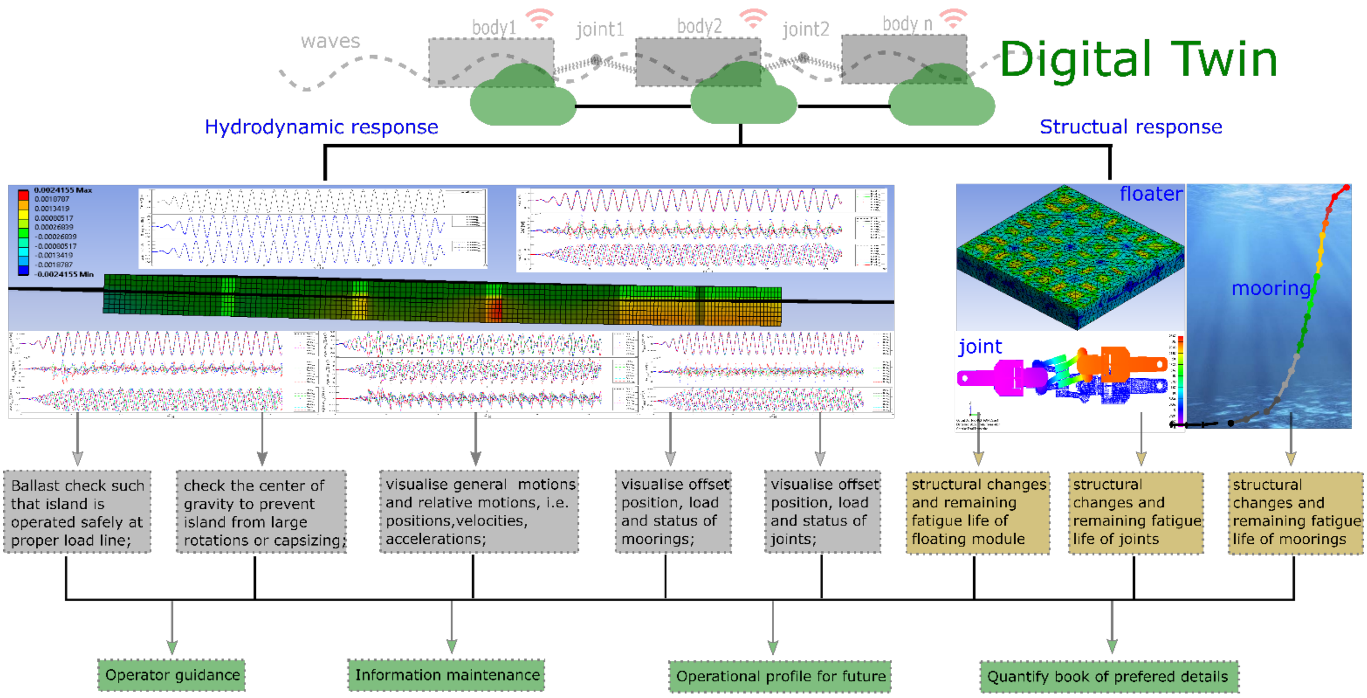

Figure 54 Proposed digital twin output

By using the digital twin, the following views are able to be provided

- The island configuration with basic alarms, alerts etc.
- Environmental conditions wind, sea, current, temperature etc.
- the global island motion state with respect to
  - horizontal movement in the mooring arrangement
  - loading overview of the mooring arrangements
  - global overview of relative motions and linking joint loads
- individual floater status popups with
  - ballast and load line status
  - generic alarm status (bilge alarms, fire, smoke, gas)
  - floater local motions
  - floaters own local sensors

Firstly, the ballast configuration and management system can be checked for each module, such that the island is operated at proper load lines. Meanwhile, the centre of gravity (CoG) is also monitored, since it dominates the floater's hydrodynamic performance, and a good CoG can prevent the island from large rotational motions. For an operator, the motions, relative motions (offset positions, velocities and accelerations), and loads for each module can be visualised. For instance, this is quite helpful when adding or removing modules, adding or adjusting moorings etc. The offset position, loads, and status of mooring system can also be real-time provided. Similar functionality also applicable to joints, as well as other connectors. The surrounding environment in terms of weather, waves, currents, ship, and aircraft traffic are collected through all the sensors, and data is postposed and visualised. This way, the impact of the island on the environment can be assessed, and the agreement of observed behaviour with expectations can be checked.

By including structural factors in a digital twin, the remaining fatigue life of floating module, joints and mooring system can be regularly updated in an automated fashion within the digital twin. Structural repairs, changes to operations, and refits, each influence the stress conditions of a floating structure, which affect how fatigue damage accumulates. If a floating structure's operational history is well documented, the digital twin can assess how changes in each of these aspects modify structural degradation. This approach allows a quantification of the influence of combined structural and operational changes on a structure's fatigue life. Without well-defined operational history, such comparisons can only show relative changes as the baseline structural condition is not well known.

Another possible contribution from digital twin is that it can provide operator guidance related to reducing fatigue damage accumulation. This guidance will involve identifying island configurations and relative incident wave angles with acceptable motions and limited fatigue damage accumulation rates using spectral fatigue analysis. In the future, software developed for on board use should consider additional factors, such as fuel usage or the nature of the current operations. By monitoring the stress state of an island throughout its life and keeping good maintenance records, local structural performance can be quantified. The number of cycles at each level of applied stress can be compared with the number of detected cracks. This quantification of fatigue performance could inform structural detail design guides. Designers could then understand whether the proposed structural detail had been assessed in stress conditions similar to those expected in a new design.

5.2 State estimation / dimensionality reduction

In most complex, real-word systems there will be features that cannot be observed – or, more precisely, quantities that are prohibitively difficult to measure, due to practical, technical, and/or financial limitations. A common way to improve the measurability (observability) of such system is to use mathematical models that incorporate knowledge of the dynamical properties of the system, i.e. incorporating the properties of the whole island from the measured properties of limited individual floaters for this project. These are often referred to as state observers or state estimators, depending on whether or not they facilitate predictions, respectively.

There exist many different observer and estimator designs, ranging from simple Luenberger observers to more advanced Kalman filters - estimator designs that also treat signal and measurement noise as well as future predictions to reduce phase legs and increase robustness. There also exist observer designs that are tailored for specific systems, such as the Nonlinear Passive Observer (NPO), which aims to filter out cyclic wave-induced vessel motions that cannot be compensated for in the vessel's dynamical position control system. For a linear system, the dynamics of a linear system is commonly expressed in a state-space equation. The state-space model describes the dynamical behaviour of the system using the state variables of the system. State variable are the minimal set of variable capable of capturing the whole response of the prescribed system. In the context of rigid body dynamics and structural dynamics, displacement, velocity at all degree of freedom (DOF) are often the state variables. For this digital twin, the discrete-time state-space can be applied, sine we are dealing with real-world dynamics problems. The discrete-time space is compatible with the sampling methods, which are inevitably discrete. Equation 5.1 is the general discrete-time state-space equation, and Equation 5.2.2 is the corresponding measurement equation. There two equations are fundamental in the framework of Kalman filtering.

$$\boldsymbol{x_{k+1}} = \boldsymbol{A}\boldsymbol{x_k} + \boldsymbol{B}\boldsymbol{p_k} \tag{5.2.1}$$

$$\boldsymbol{y_k} = \boldsymbol{C}\boldsymbol{x_k} + \boldsymbol{D}\boldsymbol{p_k} \tag{5.2.2}$$

Where $\boldsymbol{x_k} \in \boldsymbol{R}^{N_x}$ is the state vector; $\boldsymbol{y_k} \in \boldsymbol{R}^{N_y}$, the measurement vector; $\boldsymbol{p_k} \in \boldsymbol{R}^{N_p}$, the input vector; $\boldsymbol{A} \in \boldsymbol{R}^{N_x \times N_y}$, the system matrix; $\boldsymbol{B} \in \boldsymbol{R}^{N_x \times N_p}$, the input matrix; $\boldsymbol{C} \in \boldsymbol{R}^{N_y \times N_x}$, the output matrix; $\boldsymbol{D} \in \boldsymbol{R}^{N_y \times N_p}$, the feedthrough matrix. The Kalman filter (KF) is a recursive Bayesian solution for state estimation. The filter utilises two pieces of information, including 1) a linear mathematical model for the system and 2) partial measurements related to the state variables directly from the physical system, to estimate the full state. To apply the KF, additive Gaussian white noises are assumed in both the state and measurement equation.

$$\boldsymbol{x_{k+1}} = \boldsymbol{A}\boldsymbol{x_k} + \boldsymbol{B}\boldsymbol{p_k} + \boldsymbol{w_k} \tag{5.2.3}$$

$$\boldsymbol{y_k} = \boldsymbol{C}\boldsymbol{x_k} + \boldsymbol{D}\boldsymbol{p_k} + \boldsymbol{v_k} \tag{5.2.4}$$

Where $\boldsymbol{w_k} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{Q}) \in \boldsymbol{R}^{N_x}$ is the process noise vector, and $\boldsymbol{v_k} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{R}) \in \boldsymbol{R}^{N_y}$ is the measurement noise vector. Transforming these two equations into stochastic equations allows us to take into account the modelling errors, stemming from the imperfect mathematical modelling of the real system, and the measurement errors, resulting from the imperfect sensing apparatuses, respectively.

Each iteration of the KF includes two updating steps: 1) time update and 2) measurement update. The first step updates the mean and covariance of our estimation on the next state solely based on the mathematical model. The second step then incorporates the measurements in the estimation of the mean and covariance. The Kalman gain matrix regulates the proportion of the information fusion process. At time step *k+1*, assuming we have obtained the state estimate $\hat{x}_{k|k}$ and the estimation error covariance matrix $\hat{\sum}_{k|k}$ from the previous step using a set of measurements up to time step *k*, the new estimate $\hat{\boldsymbol{x}}_{k+1|k+1}$ and the error covariance matrix $\hat{\sum}_{k+1|k+1}$ are updated via the following two steps:

i) Time update

$$\hat{\boldsymbol{x}}_{k+1|k} = \boldsymbol{A}\hat{\boldsymbol{x}}_{k|k} + \boldsymbol{B}\boldsymbol{p}_k \tag{5.2.5}$$

$$\boldsymbol{P}_{k+1|k} = \boldsymbol{A}\boldsymbol{P}_{k|k}\boldsymbol{A}^T + \boldsymbol{Q} \tag{5.2.6}$$

ii) Measurement update

$$\boldsymbol{K}_{k+1} = \boldsymbol{P}_{k+1|k}\boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{P}_{k+1|k}\boldsymbol{C}^T + \boldsymbol{R})^{-1} \tag{5.2.7}$$

$$\hat{\boldsymbol{x}}_{k+1|k+1} = \hat{\boldsymbol{x}}_{k+1|k} + \boldsymbol{K}_{k+1}(\boldsymbol{y}_{k+1} - \boldsymbol{C}\hat{\boldsymbol{x}}_{k+1|k} - \boldsymbol{D}\boldsymbol{p}_k) \tag{5.2.8}$$

$$\boldsymbol{P}_{k+1|k+1} = \boldsymbol{P}_{k+1|k} - \boldsymbol{K}_{k+1}\boldsymbol{C}\boldsymbol{P}_{k+1|k} \tag{5.2.9}$$

Hence, with an initial guess on the state $\hat{\boldsymbol{x}}_{0|0}$ and error covariance $\boldsymbol{P}_{0|0}$, the Kalman filter algorithm can run efficiently to estimate the full state at every time step. Then this approach can be adopted to estimate the state of rigid body motions, as well as the structural dynamics. For further information, please go to Lau (2018) [Ref. 5].

5.3 Signal quality assessments / anomaly detection

Vast volumes of data are continuously generated, collected, and stored. The collection of the data is a process that places a high importance on the accuracy of the data collected. Data acquisition and pre-processing deal with issues such as; sampling, de-noising, removing outliers, compression, identifying missing data, and time synchronisation. The application of the proper storage principles is paramount to the quality of the data. The details are described in Chapter 3.

The signal quality assessment is the problem of finding patters in data that do not conform to a priori expected behaviour. This is related to the problem in which some samples are distinct, in terms of given metric, from the rest of the dataset, where these anomalous samples are indicated as outliers. This is important for decision makers to be able to detect them in order to take appropriate actions. The correct detection of such types of unusual information empowers the decision maker with capacity to act on the system in order to correctly avoid, correct or react to the situations associated with them. Within the S@S project, the environmental profile, operational profile, mooring systems, and et al., are intensively monitored by quite an amount of sensors, and each that send measurements with high frequency for the digital twin. The early detection of behaviours that could potentially lead to the floating structure or the attached mooring systems/joints failures in the application context of interest. One additional characteristic of this problem is that these floating modules have different operational profiles. Therefore, in order to correctly detect future anomalies, it is essential to segment the dataset available in order to automatically discover the operational regime of the floater in the recent past. This segmentation algorithm would allow one to discriminate the changes of the operational profile from anomalies and faults, as manual changes are not logged, and sometimes, those modifications take place without human supervision.

Furthermore, this S@S case is also facing a Big Data problem, since each floater has many sensors that submit information to the data hub every few seconds. A floating island will be comprised by a number of such floaters. This characteristic imposes extra requirements on the low computational complexities of algorithms and methods to be applied, as well as on the supporting computational engine. In order to deal with such a noisy data, time series segmentation is identified as a necessary technique to be used as a pre-processing step for time series analysis. This step must be able to identify blocks of homogeneous data that can be analysed in a separate fashion. Thus, the massive amount of data to be processed in an on-line fashion pose a challenge to current time series segmentation methods.

Anomaly detection techniques have been proposed in the literature, based on distribution, distance, density, clustering and classification. Their application very depending on the user, the problem domains and even the dataset. In many cases, the anomaly detection is related to outlier detection. In statistics, outliers are data distance that deviate from a given sample in which they occur. Anomaly detection techniques can be summarised by grouping them into sets of similar approaches:

- Distribution-based approaches: a given statistical distribution is used to model the data points. Then, points that deviate from the model are flagged as anomalies or outliers. These approaches are unsuitable for moderately high-dimensional datasets and require prior knowledge of the data distribution. They are also called parametric and non-parametric statistical modelling.
- Depth-based approaches: this computes the different layers of convex hulls and flag objects in the outer layer as anomalies or outliers. It avoids the requirement of fitting a distribution to the data, but has a high computational complexity.
- Clustering approaches: many clustering approaches can detect anomalies or outlier as elements that do not belong, or that are near, to any cluster.
- Distance-based approaches: distance-based anomalies or outlier detection marks how distance an element is from a subset of the elements close to it. It has been pointed out that these methods cannot cope with datasets having both dense and sparse regions, an issue denominated the multi-density problem.
- Density-based approaches: density-based anomalies or outlier detection have been proposed to overcome the multi-density problem by means of the local outlier factor (LOF). LOF measures the degree of outlier for each dataset element and depends on the local density of its neighbourhood. The local correction integral (LOCI) method, and its outlier matric, the multi-granularity deviation factor (MDEF), were proposed with the purpose of correctly dealing with multi-density and multi-granularity.
- Spectral decomposition: spectral decomposition is used to embed the data in low dimensional subspace in which the data distance can be discriminated easily. Many techniques based on principle component analysis (PCA) have emerged. Some of them decompose space into normal, anomaly and noise subspaces. The anomalies then can be detected in the anomaly subspace.
- Classification approach: in this case, the problem is posed as the identification of which categories to which an observation belongs. It operates in two phases: 1) it learns a model based on subset observation; and 2) it infers a class for new observations based on the learned model. This method operates under the assumption that a classifier distinguishes between normal and anomalous classes that can be learned in the given feature space. Based on the labels available for the training phase, anomaly detection techniques based on classification can be grouped into two broad categories: multi-class and one-class anomaly detection techniques.

For the S@S project, the novel segmentation algorithm, proposed in Marti et al. (2015) [Ref. 6], that is able to correctly identify those bocks of data at a viable computational cost can be adopted. The novelty of this approach is the combination of a fast and high-quality segmentation method with a one-class support vector machine (SVM) approach for efficient anomaly detection. The one-class SVM learns a region that contains the training data instances (a boundary). Kernels, such as radial basis functions (RBF) (Buhmann, 2003) [Ref. 7], linear Fourier, etc. (Rüping, 2001) [Ref. 8], can be used to learn complex region. If a test instance falls within the learned region, it is declared as normal; else, it is declared as anomalous. This technique with a time series segmentation method is combined to prune noisy, unreliable and inconsistent data. For further information, please see Marti et al. (2015) [Ref. 6].

Besides the signal-based fault detection described above, a model-based approach can also be applicable to the present project. The model-based approaches to failure detection, isolation and identification (FDI) is based on analytical redundancy or functional redundancy, meaning dissimilar signals are compared and evaluated to identify the existing faults in the system or its components. This comparison is between the measured signal and the estimated values generated by the mathematical model of the system. Figure 55 shows a general structure of model-based approaches.

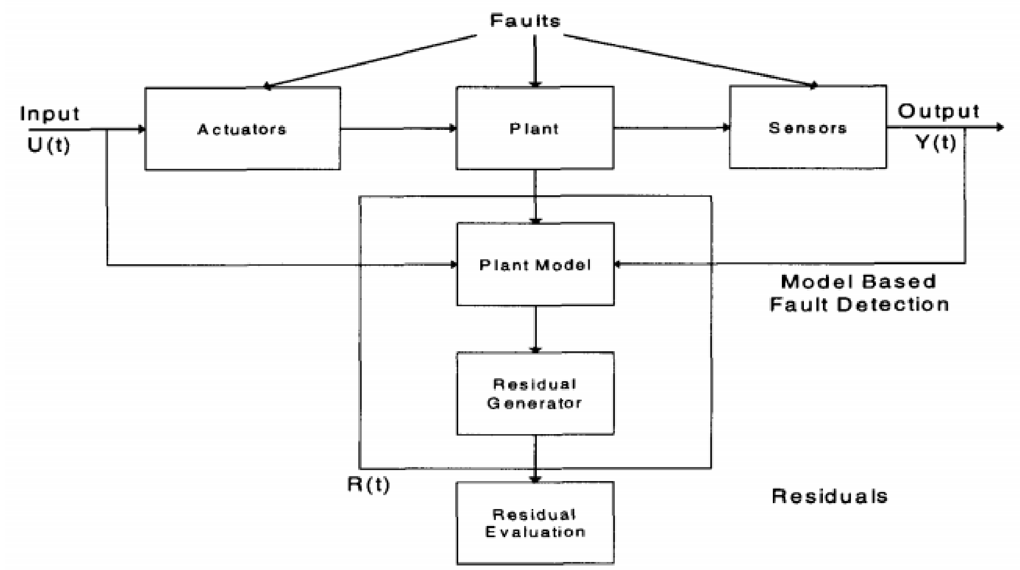

Figure 55 Model based approach [1]

Residual generations is the heart of the model-based approach. However, the techniques involved in model-based diagnosis differ in the generation and definition of a residual, for instance in some cases it is the discrepancy of output (from the system) estimation and in some cases, it is the error in parameter (of the system's model) estimation itself. It is imperative that the generated residual be dependent only on the faults in the system and not on its operating state. Several techniques have been proposed in the literature for this residual generation are a modification or improvement of the following three principles:

- Observer-based approaches;
- Parameter estimation technique;
- Parity space approach.

Observer-based approaches reply on estimating the outputs from either Luenberger observers or Kalman filters. The approach is centred on the idea that the state estimation error is zero in a fault free environment and it is not so otherwise. Dedicated observer, fault detection filters, and output observers are the three important subcategories that fall under this approach. The basic idea behind the parameter estimation techniques is that the faults affects the outputs through the system's parameters. Hence this approach is centred on generating online estimation of the parameters and analysing the changes in the estimates. In the equation error methods which analyse the parameters directly, least-squares estimation is quite often used; in the output error methods which compute the output numerical optimisation techniques are often used. The principle of parity space relations is to check for parity of the measurements from the process, generating a residual by comparing the model and the process behaviour. This approach has been shown to be in close correlation with the observer-based techniques. As stated before the model-based FDI approaches are based on identifying (constructing) models that mimic the system. However, an ideal model is never obtained because of nuances of the pragmatic world such as noise, etc., and to be effective, the model-based FDI should learn to differentiate between these uncertainties and the changes due to failures. Another difficulty is to identify not only the existing faults, but the incipient faults which may not (yet) significant affect the system.

6 Demonstrator data collection system WP10

It will be some time before a true large scale modular structure will be built. A model scale setup was thus used to verify the merits of the considerations discussed in previous paragraphs. The following aspects were listed in particular:

- Distributed acquisition of local parameters across multiple modules
- Distributed storage of collected data
- Time stamping using synchronised clock references
- Data filtering suppressing outliers and measurement noise
- Data combination and aggregation into meaningful quantities

Space@Sea WP10 scheduled extensive demonstration model tests to validate the design concept and the numerical calculation models. The tested full island configuration as tested is shown in Figure 56.

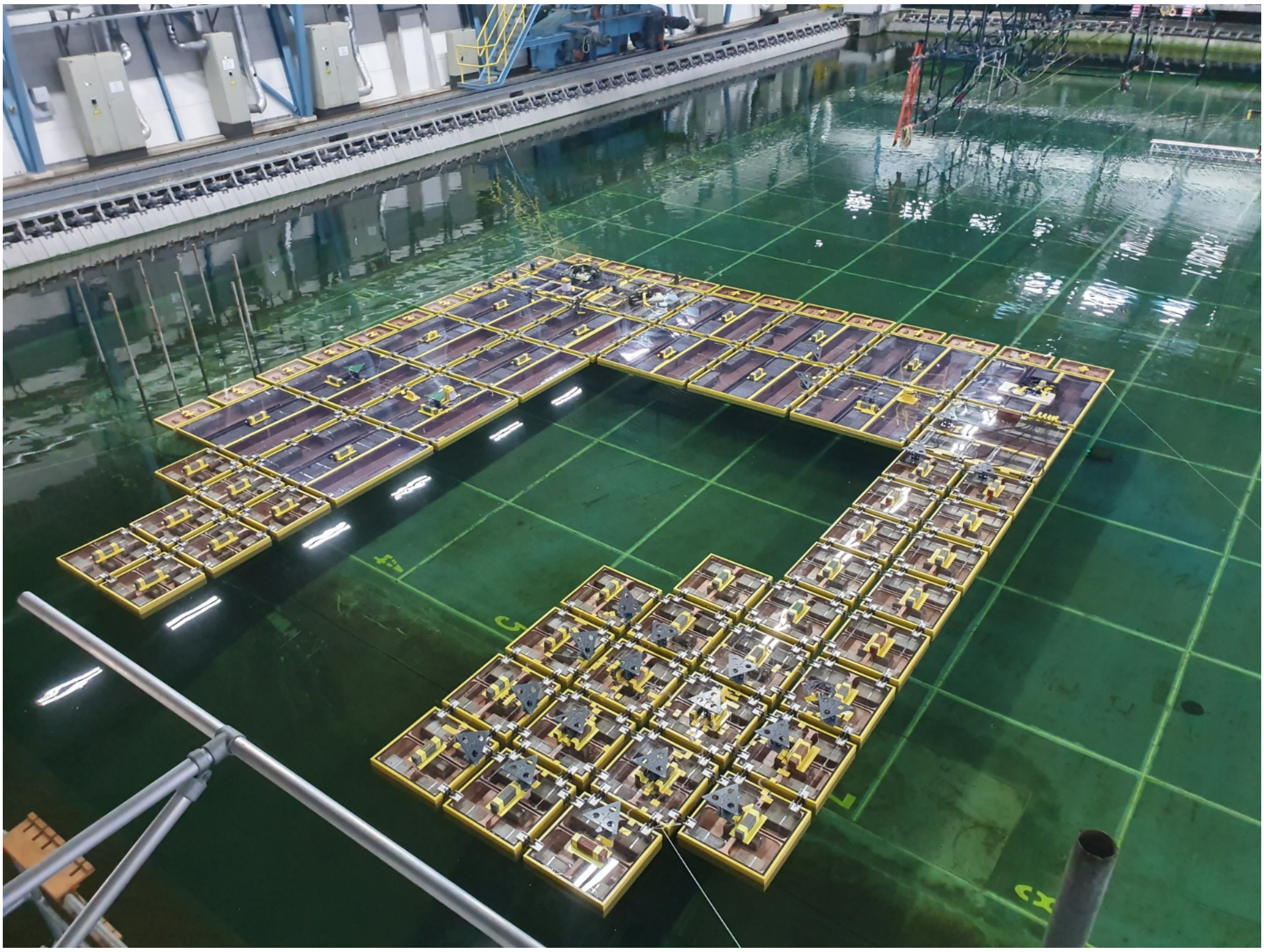

Figure 56 Full island configuration WP10 model tests

A simplified test setup using the same floater modules was used to verify the distributed data logging system as specified for O&M remote monitoring. The outline of the tested configuration is shown in Figure 57

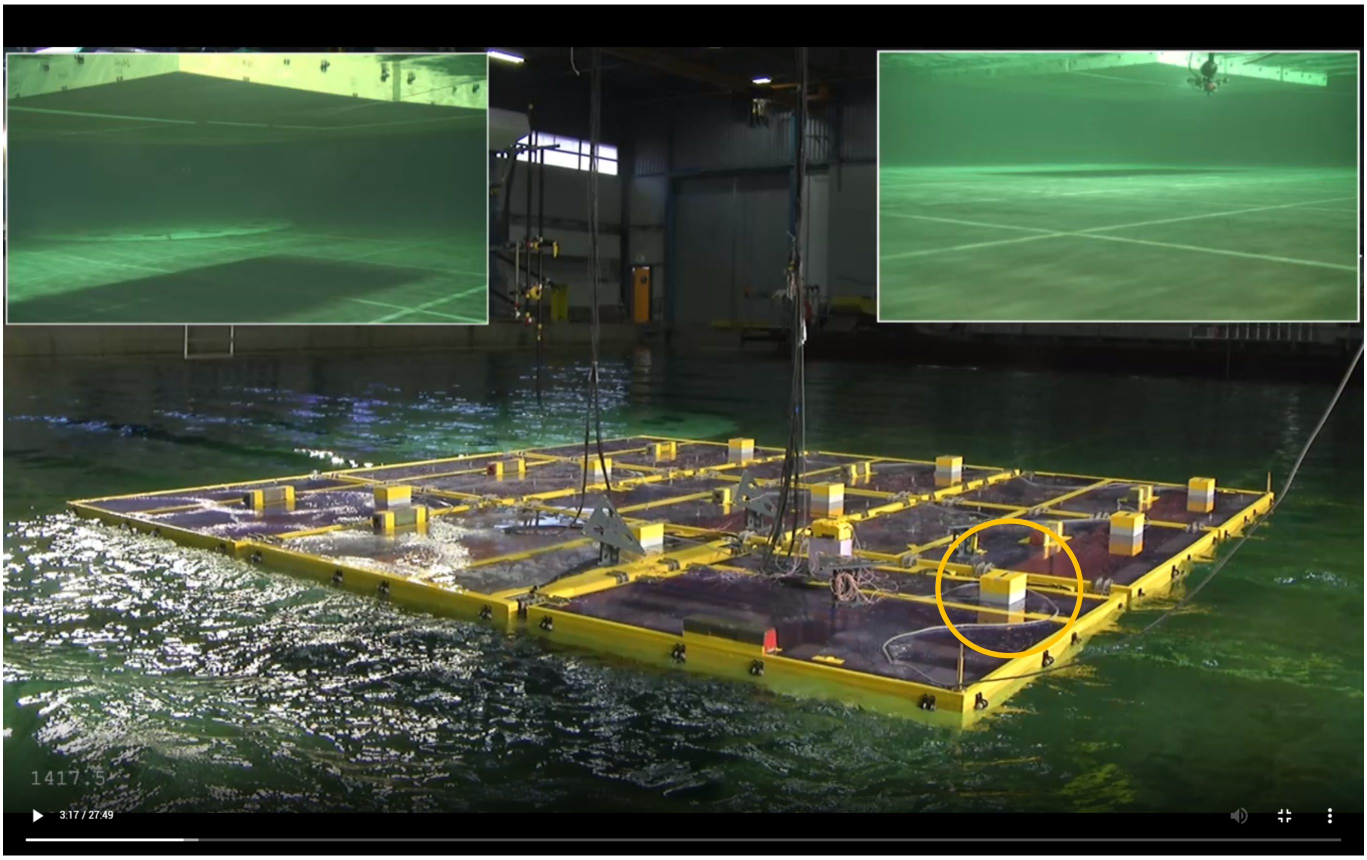

Figure 57 Test configuration model tests with Remote monitoring sensors included

The data logging controllers are fitted inside the white boxes at each of the floaters. The tests were performed on 2020, August 05. The measurements with the standard basin logging system were successful but the local data loggers unfortunately malfunctioned. Real time data was unavailable because of complications with the cabling. Local data storage was found to have not be triggered after the tests completed. The test setup is described nevertheless below.

6.1 Distributed acquisition and storage

The focus for the demonstration case was on collecting the motion response and state of a floating structure in waves. The motion response of the island is in principal defined by the combined motion responses of individual floaters and the forces in between them. The objective of the demonstrator was to reconstruct the island motion behaviour from motion and mooring measurements captured at individual floaters and express this in an intelligible way.

Motion and load measurements were performed by motion measurement units (IMU) and strain gauge load cells, each wired to a custom data acquisition controller per floater. The controller stores the collected output to local SD card storage. The logging system on each floater is a stand-alone unit which is hooked up onto a network along the platform to share power, data and timing in line with the requirements defined for the “multi modular” floating island DCS system.

The sensor setups were assembled from parts that are used regularly at MARIN. For this particular demonstrator the explicit requirement of time base alignment to 10 micro seconds was added. The synchronisation was based on the Inter Range Instrumentation Group timecoding IRIG_B.