Data Collection Guideline for Water-Energy-Food Nexus Digital Twin

Saeed, N.[1*], Ardakanian, A.[1], Trout, A.[2], Choe Peng, L.[3], and Hui Weng, G[4]

*e-mail: nagham.saeed@uwl.ac.uk

1 School of Computing and Engineering, University of West London, UK 2 Slingshot Simulations, UK 3 School of Chemical Engineering, Universiti Sains Malaysia, Malaysia 4 River Engineering and Drainage Research Centre (REDAC), Universiti Sains Malaysia, Malaysia

14-19 April 2024 – Vienna, Austria

Introduction

Achieving food security for the growing human population without widespread environmental degradation is one of the biggest challenges of the twenty-first century. Constructed wetlands (CWs) are an excellent example of how the integrated management of water-energy-food nexus in a single system can effectively reverse environmental degradation. CWs are artificial wetlands engineered to provide multiple ecosystem services to semi-arid areas; by treating wastewater and storm-water runoff through microorganisms (as much as 90%), while also providing habitat for wildlife and producing food, through low-tech methods that contribute to significant cost savings. Although these benefits have been locally demonstrated on a small scale, the quantitative benefits across all three resources have not been widely realised.

Aim

In this project, and for the first time, Digital Twins (DT) will represent the virtual environment of a CW in Malaysia. The prototype will be developed based on a successfully operating CW in a remote village in Cyprus. Consecutively, this will be tested and validated by implementing it in a similar CW in SA and Malaysia. This, will enable us to simulate how different processes and equipment configurations affect energy consumption, water quality and the crops that are being yielded from the treated water. The use of the DT will break the silos around the water-energy-food sectors and will add to the understanding of the nexus concept, minimise the risk of costly failures, and allow for more rapid innovation and development through integration. Additionally, by providing a shared, interactive platform for understanding and exploring complex systems, DT can help to build a more informed and engaged leadership and community around water-energy-food nexus projects in the global south. DT can revolutionise how we approach projects involving the water-energy-food nexus allowing more efficient and effective decision-making, by optimising the interlinkages and interdependencies of these three resources/sectors.

Site Visit to Constructed Wetland in Malaysia Frangipani Langkawi Resort and Spa

Figure 1 Location of The Frangipani Langkawi Resort and Spa

Our Malaysian collaborators visited the constructed wetland depicted in Figure 1 to explore water quality and harvesting (Figure 2a), energy production (Figure 2b), and food production (Figure 2c). During the visit, eight water samples were collected, along with data on soil composition, solar dehydrator usage, fruit and vegetable cultivation, and animal husbandry.

a) **Water:** Rainwater Harvesting (RWH) System

b) **Energy:** Solar Photovoltaic cells (on roof) and solar dehydrator

c) **Food:** Organic Farming

Figure 2 The Frangipani Langkawi Resort and Spa

Methodology

Phase I: Digital Twin Feature (Based on DT Periodic table)

Phase II: Qualitative Data Collection (Pre-constructed Wetland)

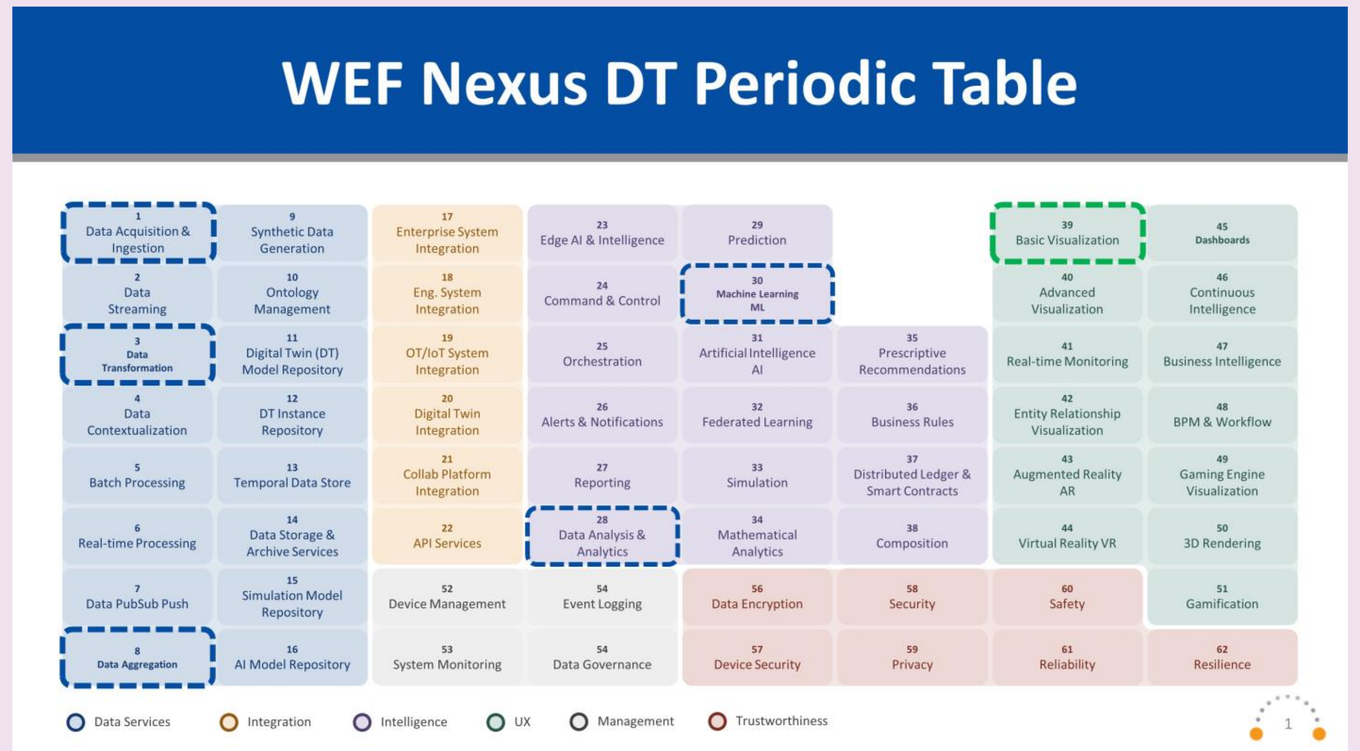
Phase III: Quantitative Data Collection from Existing Constructed Wetland Manual + WEFE Site Analyst

Phase IV: WEF Nexus Model Development

Phase V: Evaluation and Testing

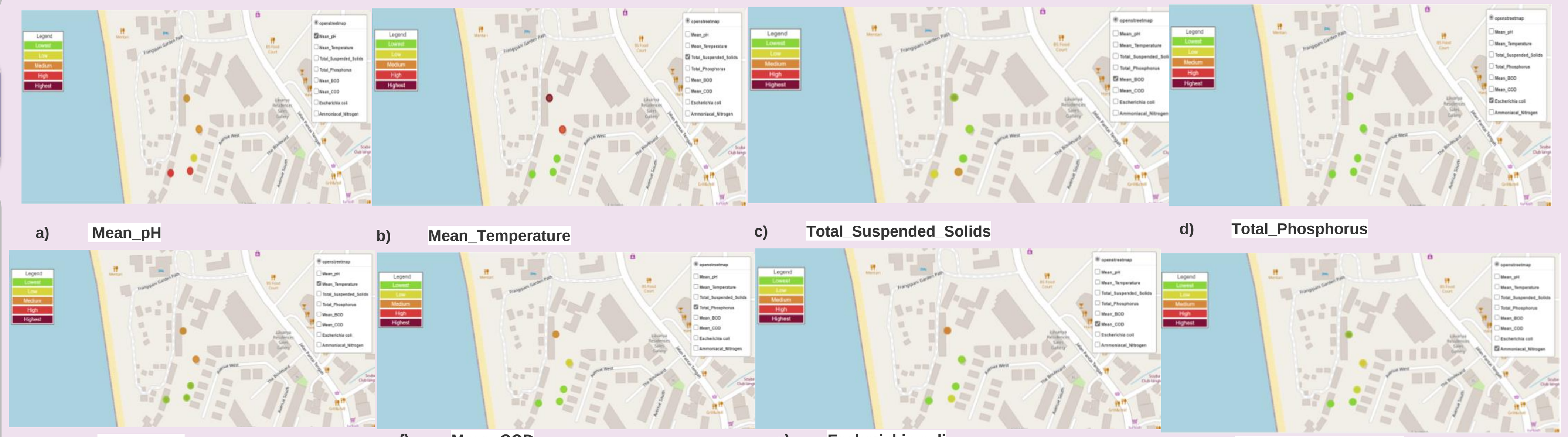
Figure 3 Methodology for the WEF Nexus Digital Twin

Table 1 WEF Nexus Digital Twin feature



The methodology employed for developing the Digital Twin for the Frangipani Langkawi Resort and Spa (CW) is outlined in Figure 3. It commenced with Phase I, where we identified the principal DT features to encapsulate a WEF Nexus CW, referencing the DT Periodic Table [1] as illustrated in Table 1. Key features selected encompassed Data Acquisition & Ingestion, Data Transformation, Data Aggregation, Data Analysis & Analytics, Machine Learning and Basic Visualization. Phase II involved pinpointing the CW alongside the three-element WEF, and the associated metrics, which were successfully delineated at the "Frangipani Langkawi Resort and Spa" in Malaysia, with measures established for data collection. The subsequent stage, Phase III, encompassed manual data collection and utilization of the WEFE Site Analyst platform [2]. Phase IV entailed model development, followed by Phase V for evaluation and testing.

Results



a) Mean_pH b) Mean_Temperature c) Total_Suspended_Solids d) Total_Phosphorus e) Mean_BOD f) Mean_COD g) Escherichia coli h) Ammoniacal_Nitrogen

Figure 4 Some of the WEF DT Model results

As depicted in Figure 4, the model effectively plotted each metric according to location, color-coding their values from green for the lowest to dark maroon for the highest. Moreover, the DT model provides visualization of all metrics grouped by month, or a breakdown for each metric by month. Mean values were computed for certain metrics like pH, temperature, BOD, and COD, while total values were calculated for others such as Suspended Solids and Phosphorus.

Model Evaluation

Table 2 OLS Regression Results for Total Precipitation against Air Temperature

OLS Regression Results

Dep. Variable:	Total_Precipitation	R-squared (uncentered):	0.868
Model:	OLS	Adj. R-squared (uncentered):	0.856
Method:	Least Squares	F-statistic:	72.20
Date:	Sat, 30 Mar 2024	Prob (F-statistic):	3.67e-06
Time:	20:58:52	Log-Likelihood:	9.6433
No. Observations:	12	AIC:	-17.29
Df Residuals:	11	BIC:	-16.80
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Air_Temperature	0.0100	0.001	8.497	0.000	0.007	0.013

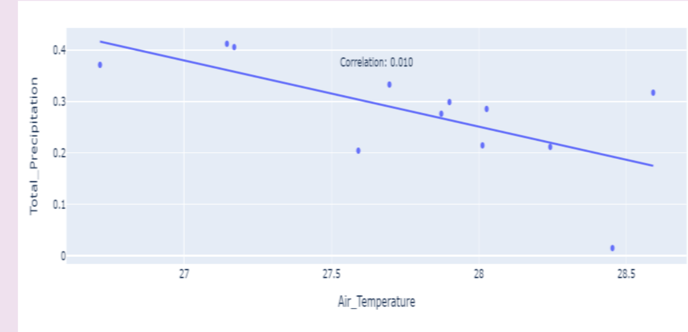
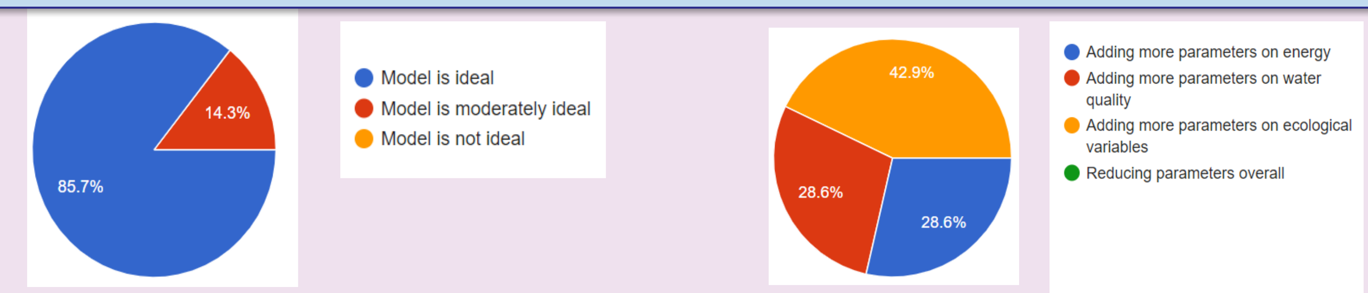


Figure 5 Correlation test between Total Precipitation against Air Temperature

Notes:
[1] R^2 is computed without centering (uncentered) since the model does not contain a constant.
[2] Standard Errors assume that the covariance matrix of the errors is correctly specified.

The Ordinary Least Squares function (OLS), implemented for analysis and evaluation, is a widely used method in linear regression analysis, assessing the variance between individual data points and the predicted best-fit line to quantify error. In our analysis, no constant term was included in the model, and weightings were not applied. The R-squared value, highlighted in Table 2, is a crucial metric indicating the proportion of variance in the independent variable explained by changes in the dependent variables, the R-squared value of 0.868 implies that our model accounts for 86.8% of the variability in the 'Air Temperature' variable. Adjusted R-squared, is valuable for assessing the effectiveness of multiple dependent variables on the model. It adjusts the R-squared metric to account for the number of variables, with a higher adjusted score, 0.856 in the table, indicating the proper contribution of variables to the model's overall explanatory power. The F-statistic, illustrated as 72.2 in the table, evaluates the significance of the linear model against a null model where the variables' effects are set to zero. The Prob (F-Statistic) value assesses the accuracy of the null hypothesis, indicating a 3.67 e-06 chance in this case. Log-likelihood, represented as 9.6433 in the table, quantifies the likelihood that the produced model fits the observed data. It aids in comparing coefficient values for each variable during model creation. Additionally, AIC and BIC, with values of -17.29 and -16.80 respectively in the table, are utilized to compare model efficacy in linear regression, incorporating penalties for multiple variables.

Model Evaluation



a) Model Impression

Figure 7 Model Evaluation results

B) Model future Improvement

During a collaborative workshop between the Royal Academy of Engineering and the Institution of Engineering Malaysia, stakeholders were introduced to the WEF Nexus DT Model for evaluation. The survey results, depicted in Figure 7 (a), were highly positive, indicating a desire to further enhance the model. Stakeholders expressed interest in its expansion to support natural wetlands and suggested incorporating additional metrics pertaining to energy and food, as shown in Figure 7 (b)

References

[1] van Schalkwyk, P. (2022) *Digital Twin Capabilities Periodic Table, Digital Twin Consortium*. Available at: https://www.digitaltwinconsortium.org/wp-content/uploads/sites/3/2022/06/Digital-Twin-Capabilities-Periodic-Table-User-Guide.pdf (Accessed: 28 March 2024).

[2] Fleischmann, J. (2023) *RL-Institut/WEFESITEANALYST: Automated Collection and analysis of open datasets for the site-tailored planning of integrated water-energy-food-environment systems (iwefes)., GitHub*. Available at: https://github.com/rl-institut/WEFESiteAnalyst (Accessed: 30 March 2024).