Digital twin of an industrial cluster: a proof of concept on the Humber Estuary

Regulator’s Pioneer Fund project

December 2023

Contents

Executive summary

The climate is changing, and impacts are already being felt. Recent heatwaves and droughts in Europe have led to nuclear and gas power generators needing to reduce capacity or breach environmental permits (Qvist, 2019), with adverse impacts on power supply, the economy and the natural environment. With droughts and heatwaves predicted to increase (Climate Change Commitee, 2021), the UK's Climate Change Risk Assessment already identifies this as a challenge to sectors requiring plentiful, clean water, such as power generation or energy-intensive industries. It also poses a risk to the deployment and operation of new technologies, including carbon capture and storage and hydrogen production. Understanding local environmental capacity and the cumulative impacts of decarbonising industrial clusters is essential to protect people and the environment. Critically, it will also ensure that industrial clusters and the technologies they will rely on are sustainable over the long term, including economically viable as pressure on natural resources increases from both human activity and climate change.

Using a grant from the Regulators' Pioneer Fund, this project has produced a proof of concept (PoC) digital twin model to begin to assess the cumulative impacts of an industrial cluster on the water environment. The PoC digital twin model uses both real and notional environmental and industrial data, combined with climate change projections, to look at the potential impact of the emerging industry over time and how limited natural resources may affect deployment. The PoC focuses on an industrial sub-cluster along the River Trent and River Ouse leading into the Humber Estuary.

The PoC had several key challenges to overcome during the design phase including the handling of data from multiple sources with varying temporal and spatial detail (including incomplete data sets); ensuring the software tools interacted correctly to develop an interactive model framework and visualisations; and determining how complex hydrological models could best be represented in a simplified form for the purpose of proving the concept.

The PoC looks at river flow and water temperature under climate change scenarios against potential industry cooling water requirements and assesses any impact on the aquatic environment using environmental thresholds. The model allows the user to configure the industrial activities, adjusting the abstraction and discharge volumes as well as the temperature of the thermal discharge. This allows the user to repeatedly configure and run simulations in the digital twin to potentially create a favourable scenario where industry could take the necessary resources whilst ensuring environmental protection. The PoC digital twin outputs can be visualised both in graph form, allowing interrogation of the data, as well as providing a 3D video simulation showing impacts over time.

Through the creation of the PoC and a series of test simulations this project has determined that digital twin technology could assess cumulative impacts and environmental capacity in low carbon industrial clusters, and potentially provide significant benefits for short- and long-term environmental planning.

At this stage, the PoC digital twin cannot be used to make regulatory decisions, nor can it be used by the industry to design projects. To do this, additional data sources would be needed, and the environmental interactions built within the model need refining to improve accuracy. However, the PoC provides an important foundation to demonstrate how digital twins could indeed be used to inform regulation, identify environmental risk and potential future environmental or planning issues. With collaboration with other stakeholders, including other regulatory bodies and industry, additional data and further refinement, digital twin technology has the potential to be a useful tool in future strategic and regulatory decision making.

Project background

Low carbon technologies such as carbon capture and hydrogen production play an important role to meet UK Net Zero 2050, if they can be deployed sustainably. Powering Up Britain sets out the government's commitment to deliver four low carbon industrial clusters by 2030 and at least one net zero cluster by 2040.

The Environment Agency's 'Environmental Capacity in Industrial Clusters Humber Pathfinder' project (funded by the Department of Business Energy and Industrial Strategy in 2021-22) explored the water needs and constraints of deploying carbon capture and storage and hydrogen production in the Humber Estuary, part of the East Coast Cluster. The pathfinder concluded that current and future water resources in the Humber region are already under pressure, exacerbated by growing demand and the impacts of climate change, with additional abstractions in some locations not being viable. As such, current plans for low and zero carbon technologies are not consistent with current and future challenges to water in the Humber. This could severely hinder the operation of the technology and the ability for the Humber Cluster to reach net zero by 2040 and UK Net Zero by 2050.

Ultimately, planning our net zero future must include what a sustainable and climate resilient industrial cluster looks like (Environment Agency, 2022). Any new development, including low-carbon development, must be able to operate with the context of local environmental capacity. The ability to predict and forecast the availability of natural resources, and how different (and often competing) industrial sectors will use those resources, is critical to the successful deployment and long-term resource efficiency of low carbon technologies.

In November 2022, the Department of Business, Energy and Industrial Strategy awarded the Environment Agency a grant of £271,975 via the Regulators' Pioneer Fund to deliver the Humber Estuary digital twin. The Regulators' Pioneer Fund is a grant-based fund to enable UK regulators and local authorities to help create a UK regulatory environment that encourages business innovation and investment. The current £12m round is being delivered by the Department of Science, Innovation and Technology.

The project proposed an innovative research and development approach to create a proof of concept digital twin that provides a visual representation of the water environment in the Humber industrial cluster, future potential environmental capacity, and environmental pressures of carbon capture and hydrogen technologies. The model is informed by current environmental data, climate projections and proposed infrastructure projects. It is not an exact replica of the Humber estuary environment or proposed industry as there is currently insufficient detailed data to support this level of accuracy, as well as time constraints to translate the data into a working digital twin.

Digital twin

A digital twin is a virtual replica of an object, system or environment which can be altered to explore different scenarios and determine the impact on the real-world object or environment, the effect is then visualised by the digital twin. Digital twins are used to remotely monitor, replicate, and improve the efficiency of complex industrial processes, or to virtually play out scenarios historically or in the future, to establish potential impacts. The technology is now beginning to be explored for environmental sectors to assess and manage current and future environmental risk and aid short- and long-term strategic decision making (Government Office for Science, 2023). The potential benefit of a digital twin over a digital model is often realised in the connection to and similarity of its physical twin. It adds value by incorporating historical, real time or current data along with long-term future predictions to demonstrate the impacts on the physical (real) twin and inform short and, or long-term strategic decision making. The effectiveness of a digital twin is reliant on how accurately it recreates the physical twin and is heavily dependent on aspects like data quality and accuracy, modelling assumptions and understandability of visual output (Bolton A, Enzer M, Schooling J et al, 2018).

Project aim

The project aimed to create and test a digital representation of a low carbon cluster within its environmental setting and identify environmental and operational limits by simulating the operation of multiple low carbon technologies, now and in the future. It aimed to provide opportunities to innovate, simulate and mitigate environmental impacts and understand limitations of data and digital twin technology.

To facilitate the creation of the digital twin model several key questions were posed:

1. Can a digital twin approach be used to determine if there will be cumulative impacts on the environment from the deployment of a low carbon industry?
2. Will a threshold be reached where the deployment of low carbon technology is constrained by the environment (including in the future) through reduced flows and increased water temperatures?
3. Can digital twin technology be used to engage and inform stakeholders, environmental policy, and regulation?

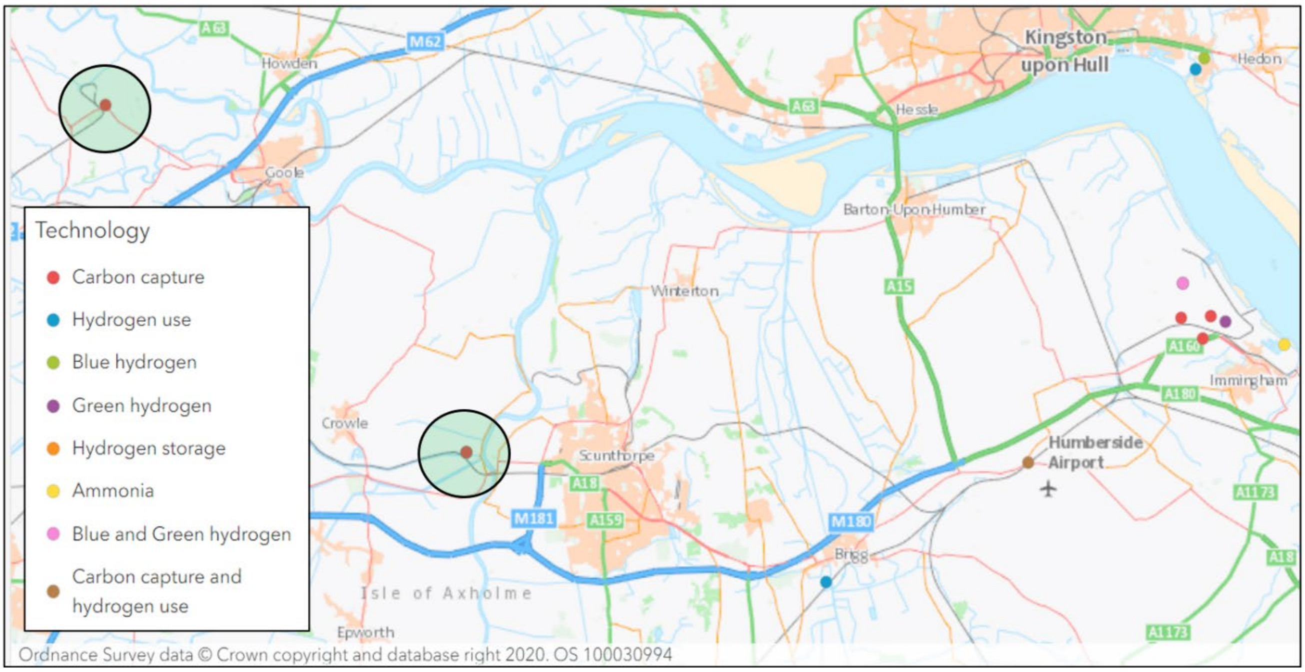
4. Can the capability of the model be expanded to provide a more holistic picture of environmental conditions and changes, and what gaps or barriers may exist?

Study area and scope

The initial project scope was to build a digital twin model of the Humber Estuary. Due to the short timeframe for the project, modelling complexities, geographical scale of the Humber estuary and data limitations a refined scope was agreed with the Regulators' Pioneer Fund to focus on smaller industrial clusters within the Humber area. A sub cluster at Immingham was discounted due to the lack of estuarine data and additional complexities modelling tidal flow. As a result, sub clusters on the River Trent, at Keadby and the River Ouse, near Drax, were selected to demonstrate how the digital twin technology could be used to predict environmental capacity and cumulative impacts from potential low carbon technologies. Both rivers feed into the head of the Humber Estuary (Map 1).

The model focuses on the two key parameters of water availability and water temperature as these can adversely impact the natural environment, as well as the potential efficiencies of the low carbon technologies. There are other parameters that could be incorporated into a digital twin to understand the cumulative impacts in more detail, however due to project time constraints a decision was taken to focus on water availability and temperature for the PoC.

Map 1: Humber Estuary- proposed low carbon cluster technologies. Digital twin PoC sites are highlighted in green



Data

Baseline environment

The baseline environment in the digital twin was created using gauged river flows from the nearest assessment points at North Muskham and Skelton, on the Rivers Trent and Ouse, respectively. Met Office historical air temperature (to present date) and Environment Agency water sample data was also incorporated. Due to limited data sets (some only having monthly or annual figures) the data has been extrapolated creating a notional data set. The limitations of this approach and recommendations are set out later in the report.

Climate projections

Future river flow data was extracted for the assessment points at North Muskham and Skelton from the enhanced future flows and groundwater (eFLaG) database. eFLaG is a set of nationally consistent hydrological projections derived from the UK climate change projections. The digital twin incorporates three river flow models available from eFLaG. For each model the minimum, median and maximum flow predictions were included to allow climatic averages and extreme scenarios to be assessed.

Currently, water temperature predictions do not exist for estuarine waters or rivers. The Environment Agency is in the process of producing riverine water temperature predictions for England, but the results were not available within the timeframe of this project. In the absence of water temperature projections, UKCP18 air temperature projections were used due to the understood direct relationship between the parameters, and a linear regression applied based on a small sample of routine temperature data collected by the Environment Agency. Datasets show the mean and maximum temperature projections for the Humber region at intervals between the present day and 2100.

Thresholds

Hands-off flows for the River Trent and River Ouse, taken from the Environment Agency's abstraction licensing strategies, were used as thresholds to determine the river level the hypothetical low carbon industries would need to cease abstracting at. Hands-off flows are designed to protect the environment and the right of other abstractors. If the flow in the river or groundwater level drops below the local hands-off flows required to protect the environment, abstraction must stop. Therefore, in dry years, restrictions are likely to apply more often, which will affect the reliability of supply, potentially limiting the growth and long-term sustainability of clusters. The hands-off flow is static overtime in the PoC model.

The model uses established environmental temperature thresholds for migratory fish species (non-cyprinids, cyprinids and sea lamprey, a protected species within the Humber Special Area of Conservation) to determine the potential cumulative impact of the thermal discharges. (Environment Agency, 2018) A cooling water intake limit of 23°C was also incorporated to show the potential impact of the thermal discharges and climate change on the sustainability of the river water used for industrial cooling as higher cooling water intake temperatures are associated with reduced efficiency in powerplants (Environment Agency, 2022).

Industrial clusters

At present there is limited data available for the low carbon industries in the Humber Estuary, therefore publicly available local industry cooling water abstraction and discharge permit data was used to demonstrate how a cluster could be added and assessed in the digital twin.

Development of the digital twin

Key challenges

The key challenge faced by the development team was to create a digital platform that enabled data sources of differing levels of detail, timescale, and geographic location, to be incorporated into an interactive and intuitive visual product – in a relatively short time frame. This led to specific challenges, including:

Incorporating data from multiple levels of granularity: hourly, daily, monthly, annually. A way to integrate the different data sources levels was required.

Integrating multiple software tools: for example, analytics, streaming data, future projection simulations, digital twin modelling and digital twin visualisation. Although Microsoft was chosen for this PoC and the Azure suite used, each tool has a different way of working, individually and collectively, with differing strengths and weaknesses. The challenge of connecting these components together in a way that made sense logically as well as technically needed to be resolved.

Product maturity: Some of the components of the Microsoft products used for the PoC are new and immature. digital twinning is a relatively new venture for many software vendors and consequently some of the tooling is in its infancy. This created challenges in terms of data integration as well as the final interface capabilities.

Complex environmental modelling: Some compromises in terms of the complexity of the mathematical models representing the environment was required. Considerable time would have been required to integrate complex models needed to represent the water environment, for example, tidal flow, hydrodynamic effects, riverbank topography, localised effects of specific industrial activities). A modular approach was taken where some mathematical models were kept simple to prove the concept. Further complexity and nuances of interactions and models could be added later.

Visualisation: More than one visualisation tool was chosen to represent the results of the model and meet the requirements of different end user groups. Power BI (statistical data visualisation software) was selected to enable self-service analytical work on the data sets, and Azure digital twin tooling enabled the creation of 3D simulations of the data environment, for users and stakeholders requiring a ‘real world’ representative view and allowing interaction and engagement with the data on a new level.

Addressing the challenges

There were two key technical scenarios we wanted to represent in the PoC.

1. Real time (low latency) sensor data through to the visualisation tools.
2. Climate change projections and industry data through to the final visualisation tools.

Real time data

To simulate real world sensors in the environment and demonstrate how the data would feed through into the digital twin, a high-level architecture was employed (Figure 1). The IOT central hub application simulated real world sensors for a range of water quality parameters (Figure 2). This data feeds into the event hub as if it were a real sensor in the real environment. This framework was designed to allow potential future sensors situated in the real environment to be easily plugged into the architecture. This proof of concept project demonstrated how this could be achieved. The event hub directs data into two separate streams. The first, through Stream Analytics into a Power BI dataset. This enables analysis of the streamed (‘real time’) and historical data. The second, feeds data into the Azure digital twin where it is modelled against digital twin thresholds and represented in the Azure digital twins visualisation layer (Figure 3).

Figure 1: Real time architecture

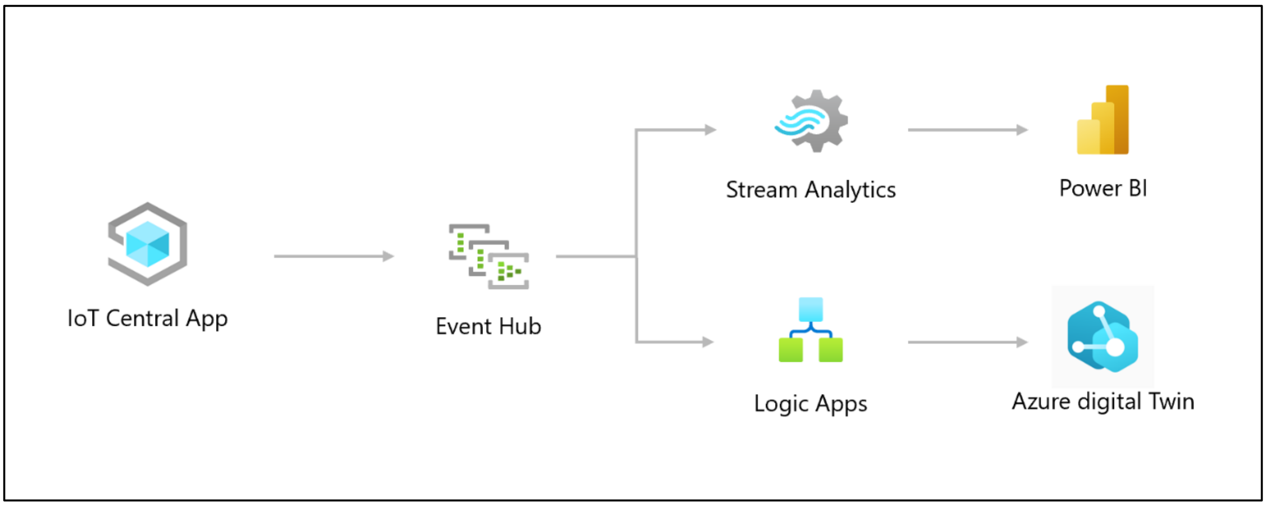


Figure 2: IOT central app showing one of the simulated water sensors

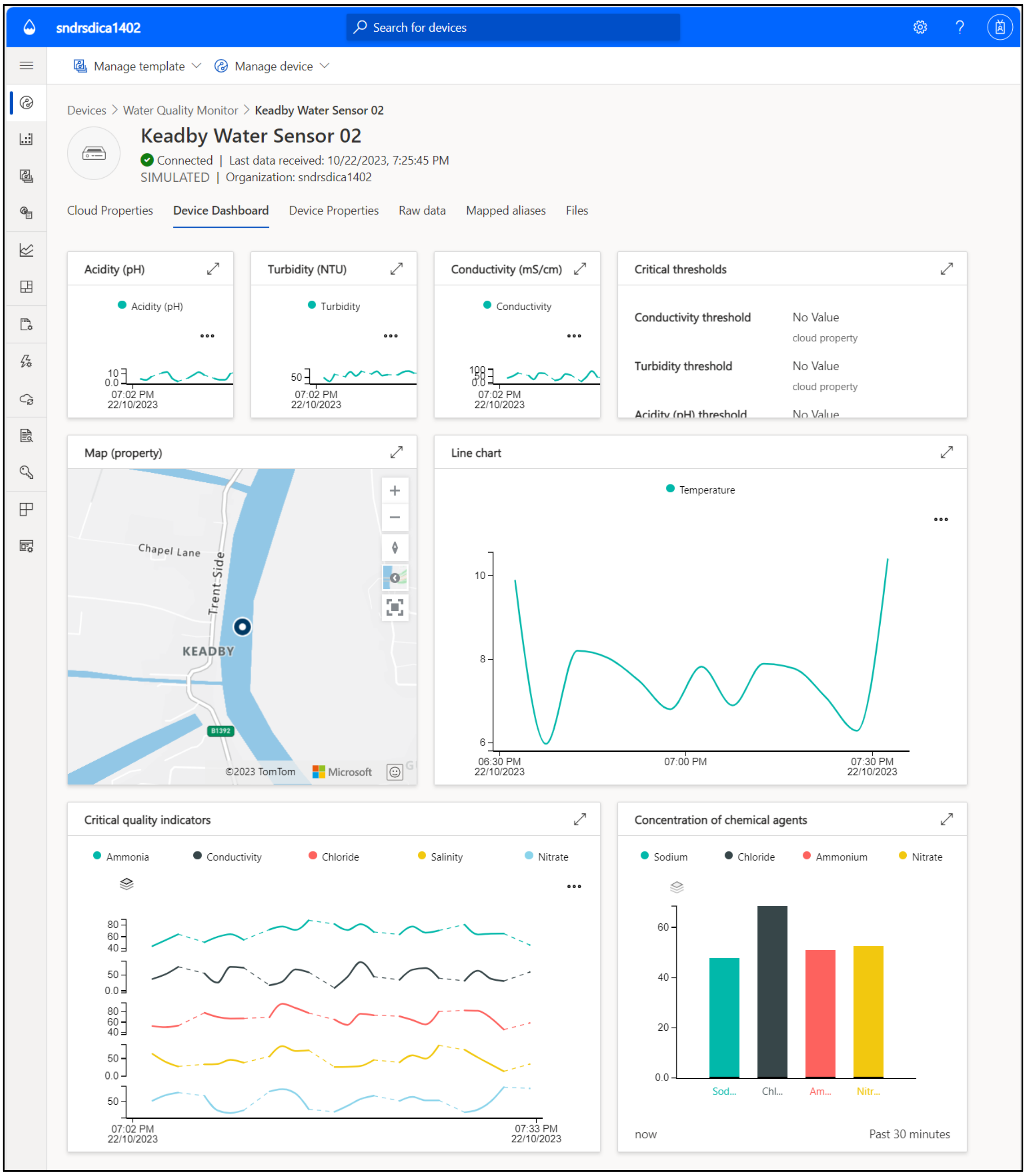


Figure 3: Azure digital twin 3D visualisations

Climate change projections and industry data

To incorporate climate change projections and industry data a similar architecture to the real time sensor feed was employed, with a couple of key additions to translate batch data to a flow data and allow the digital twin user to configure the model and set thresholds to assess potential future industrial activity and environmental impacts (Figure 4).

The data was incorporated and processed using Power Query and loaded into Power BI datasets. Power BI was selected to perform this function as it enabled flexible data modelling, maximum customisation, and simpler ingestion of new future data sets. It also enabled further interaction points for other downstream analytics and provided the ability to create rich analytical views against the data, aiding user analysis (Figure 5).

Figure 4: Climate change projections and industry data architecture

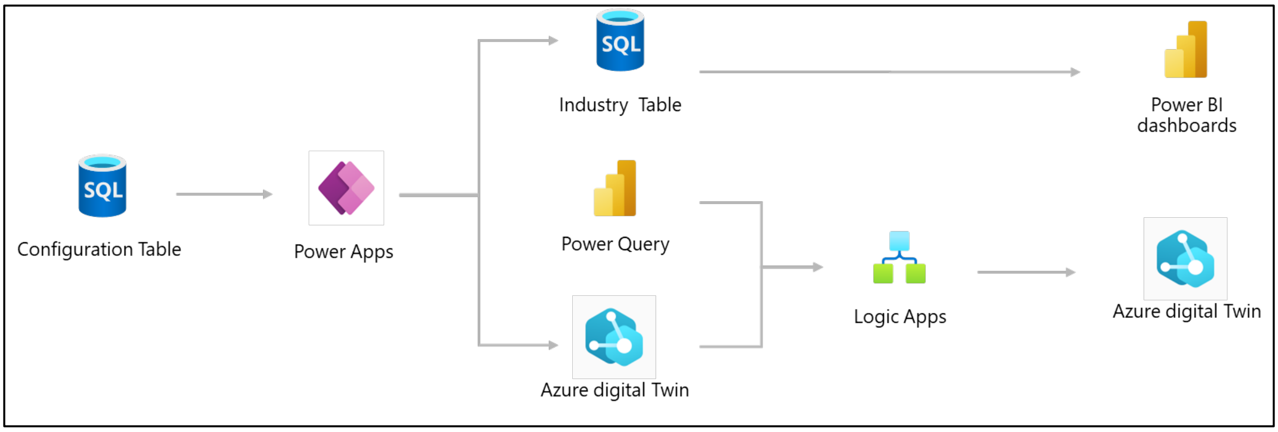
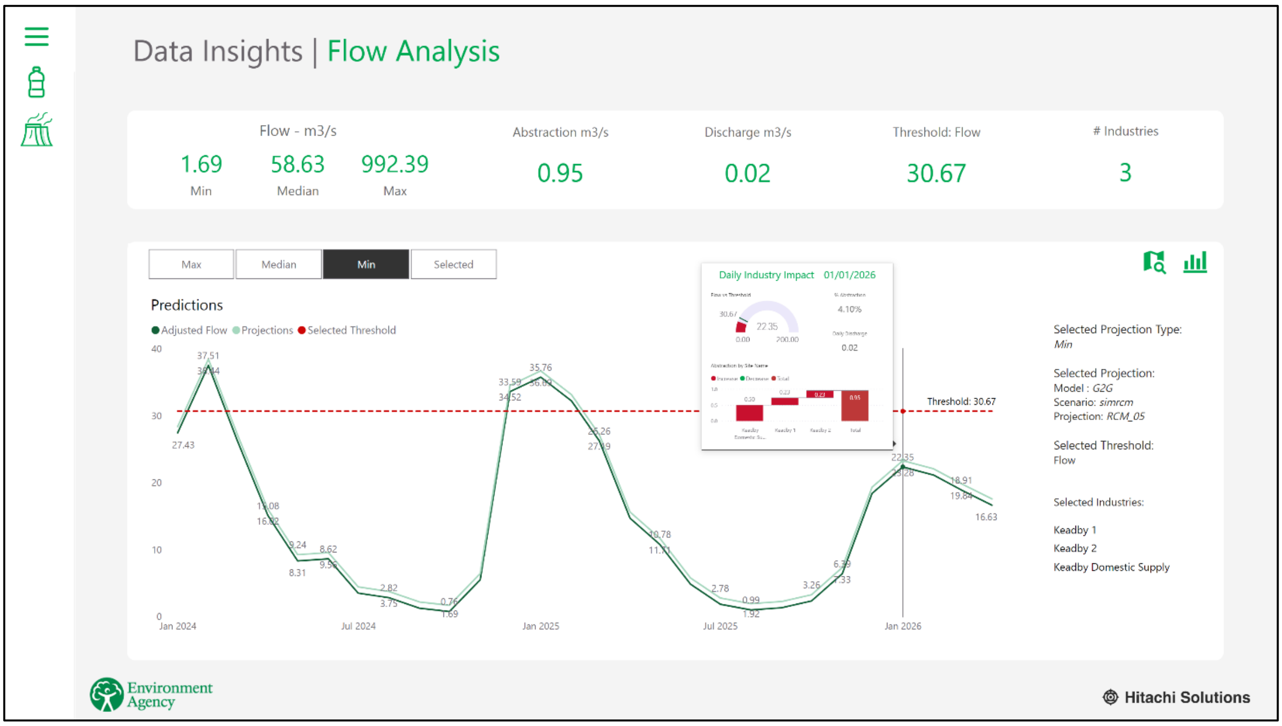


Figure 5: Power BI output showing projected river flows and industry abstraction



Using the digital twin

To use the digital twin the user configures the model to choose the desired geographic location, data set, environmental or industrial thresholds, and industry units (Figure 6). The user confirms the configuration and initiates the model run. At this point, the back end of the process gathers data from various sources and then initiates the simulation over time passing the data in the form of a playback stream to the Power BI the data visualisation tool and the Azure digital twin 3D model (Figure 7). The model determines the cumulative impact of the industry units (abstraction volumes versus discharges, or thermal discharges) and uses the environmental thresholds to determine if industrial activity would need to cease. The model calculates the number of threshold exceedances over time and demonstrates the potential impact the industry units would have on the water environment.

Figure 6: Power app user configuration page

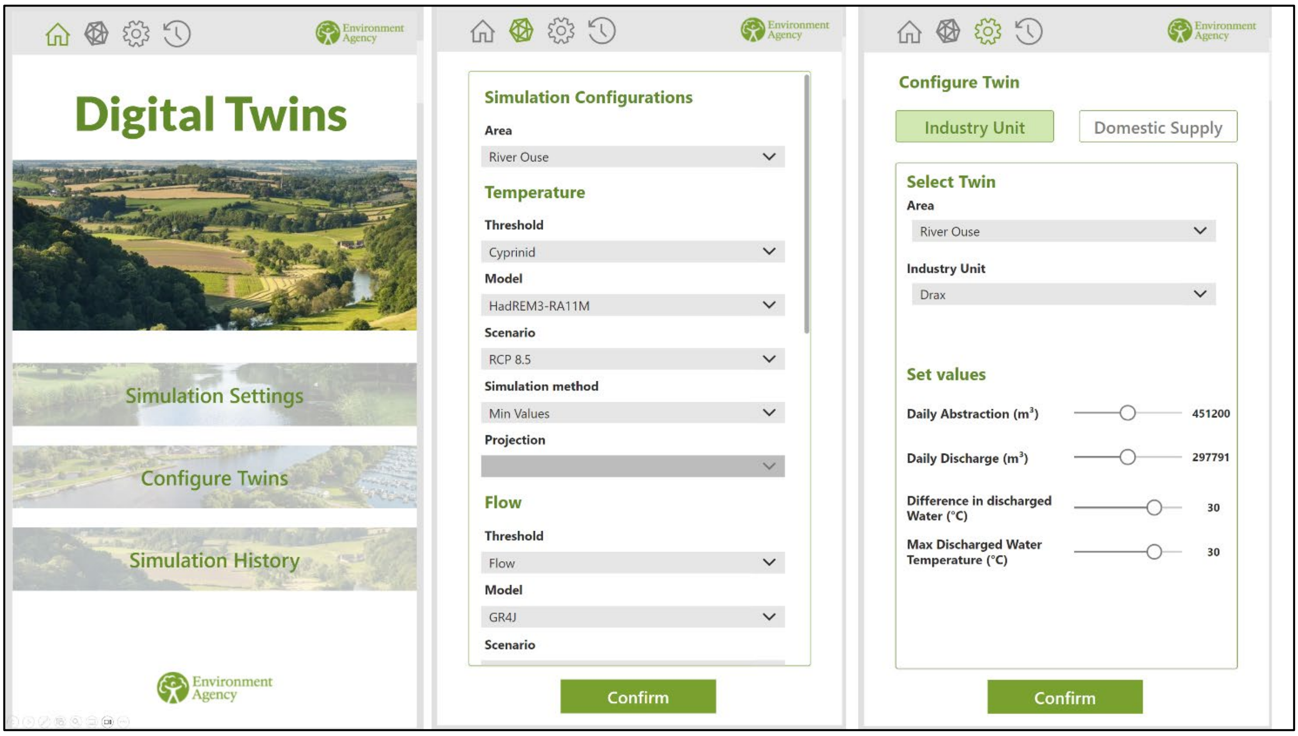
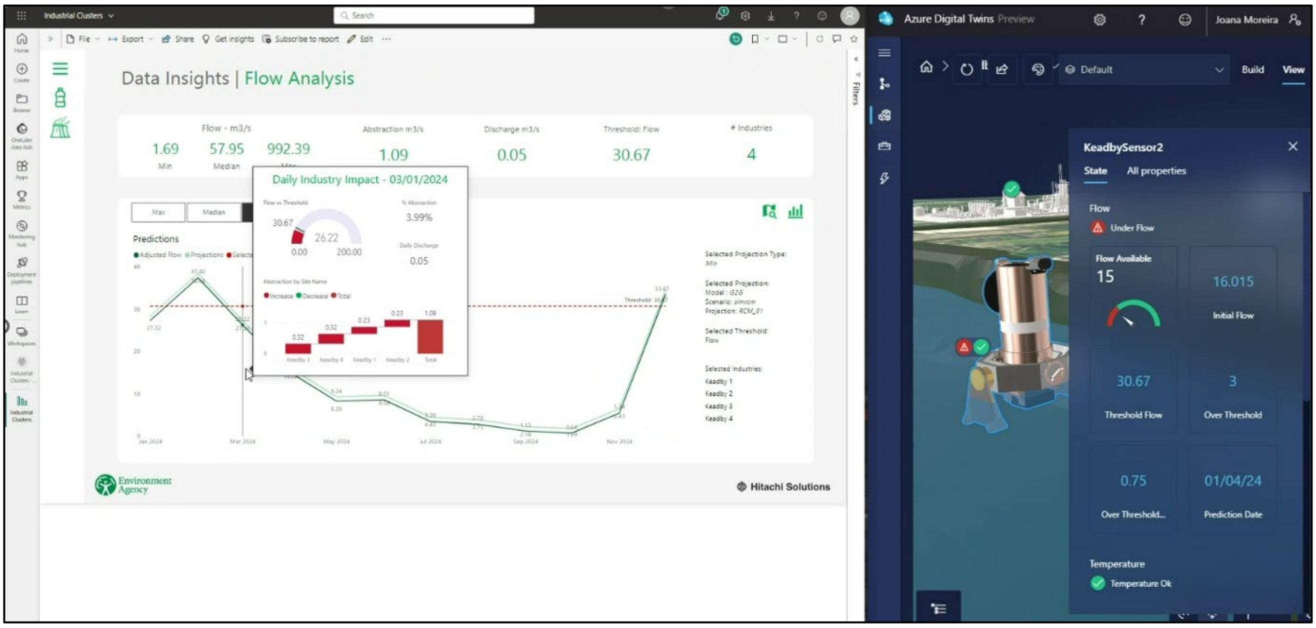


Figure 7: Digital twin 3D visualisation



Results

To demonstrate the capabilities of the PoC digital twin model four key questions were posed at the start of the project. To answer the first two questions, scenarios were simulated on the PoC digital twin.

1. Can a digital twin approach be used to determine if there will be cumulative impacts on the environment from the deployment of a low carbon industry?
2. Will a threshold be reached where the deployment of low carbon technology is constrained by the environment (including in the future) through reduced flows and increased water temperatures?

Scenario 1: Does the low carbon cluster have an impact on water availability for cooling water?

For this scenario simulations were run for the year 2024 using the median flow model projection. The PoC digital twin shows there is an increase in the number of times the hands-off flow would be reached if more than one industry is added to the modelled area on the River Trent. The results of the simulation are summarised in Table 1. Figures 8 and 9 show the digital twin model visual outputs.

Table 1: Exceedances of the hands-off flow 2024 with increased low carbon industrial activity

Number of industries	G2G Model (median)
0	3
1	3
2	4
3	4
4	4

Figure 8: Statistical model output showing impact of water consumption from four industries in 2024

Key: Adjusted flow - abstraction impact of the selected industries; Projections- climate model selected for the simulation; Threshold- environmental/ industrial temperature threshold.

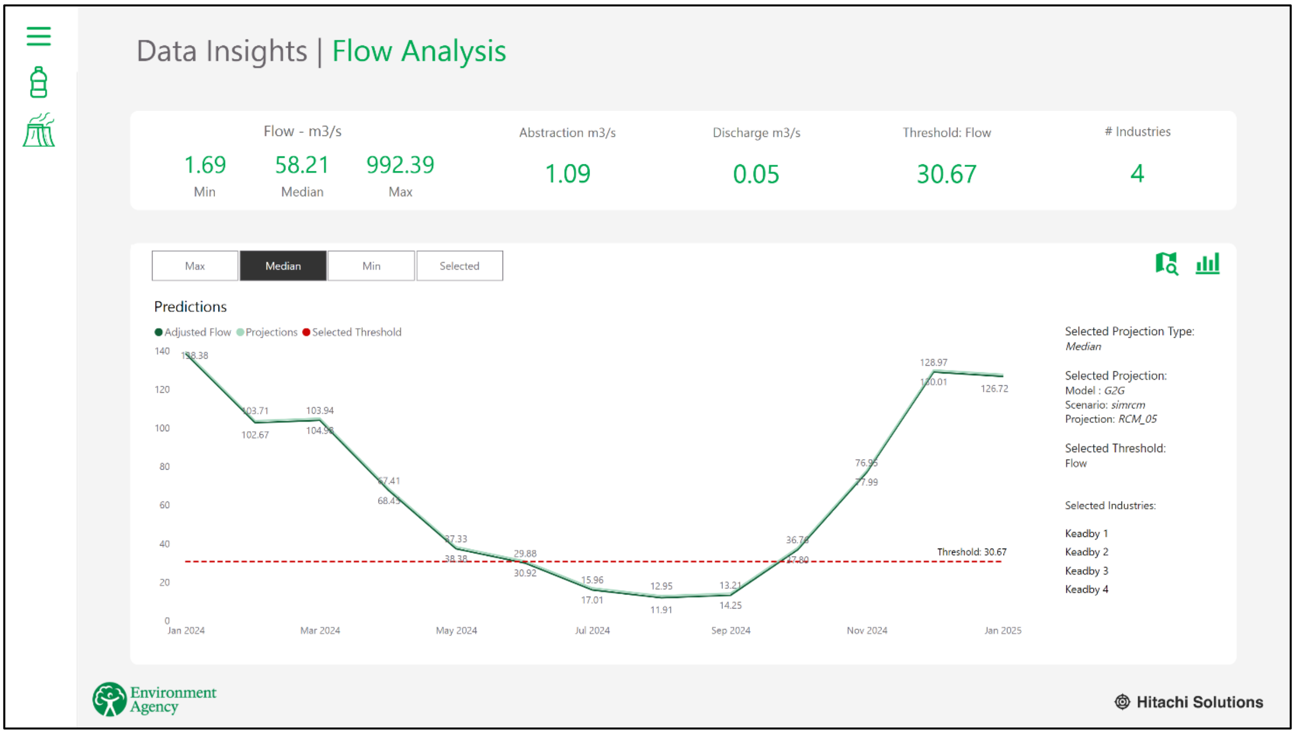
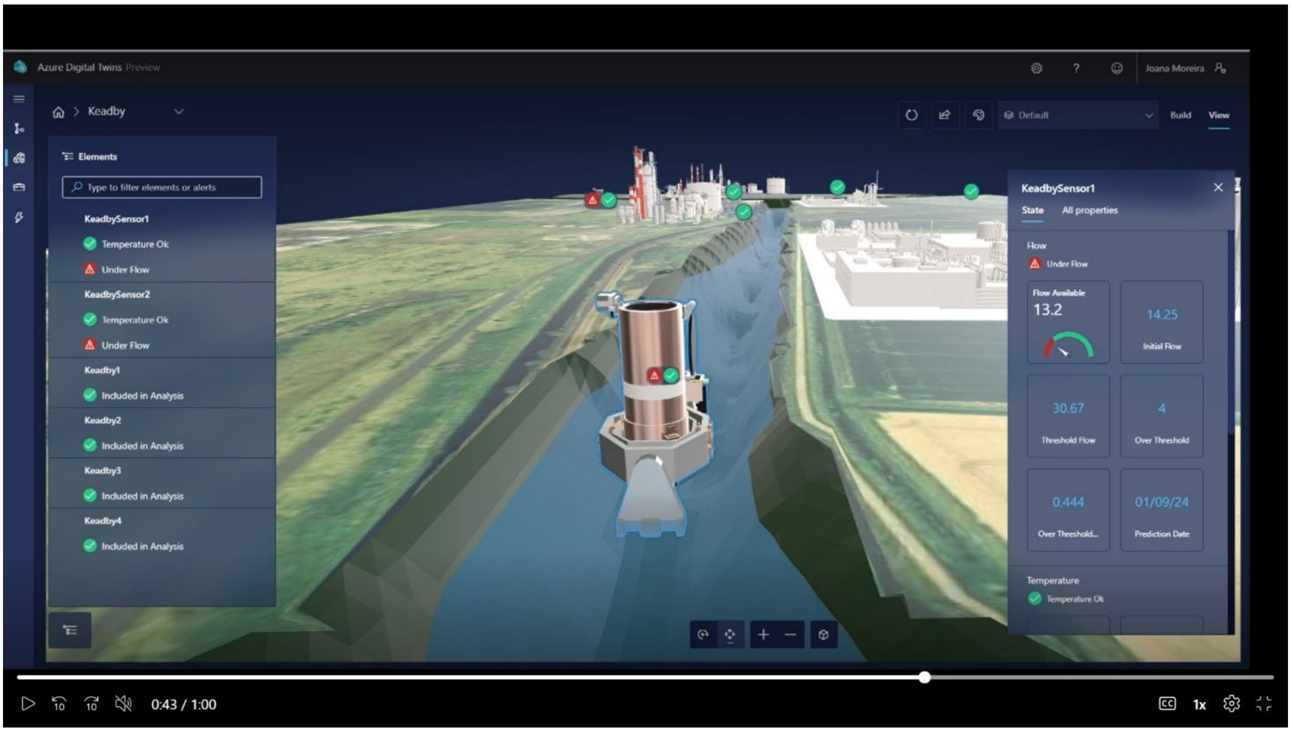


Figure 9: Digital twin simulation showing impact of water consumption from four industries in 2024



Scenario 2: Is a threshold reached where the aquatic environment is adversely impacted from increased river temperature from thermal discharges?

The PoC digital twin indicates the non-cyprinid fisheries threshold would be breached twice after all four thermal discharges from the low carbon industries are added over a 12-month period in 2024. The results of the simulation are summarised in Table 2. Figures 10 and 11 show the digital twin model visual outputs. River temperature data is limited so for this example the notional maximum temperature projection data model was used.

Table 2: Exceedances of the non-cyprinid fisheries threshold for the year 2024 with increasing low carbon industrial activity

Number of industries	HadREM3-RA11M Model (maximum)
0	0
1	0
2	0
3	0
4	2

Figure 10: Statistical model output showing the impact of thermal discharges from four industries in 2024

Key: Adjusted temperature - thermal discharge impact of the selected industries; Projections - climate model selected for the simulation; Threshold - environmental/industrial temperature threshold.

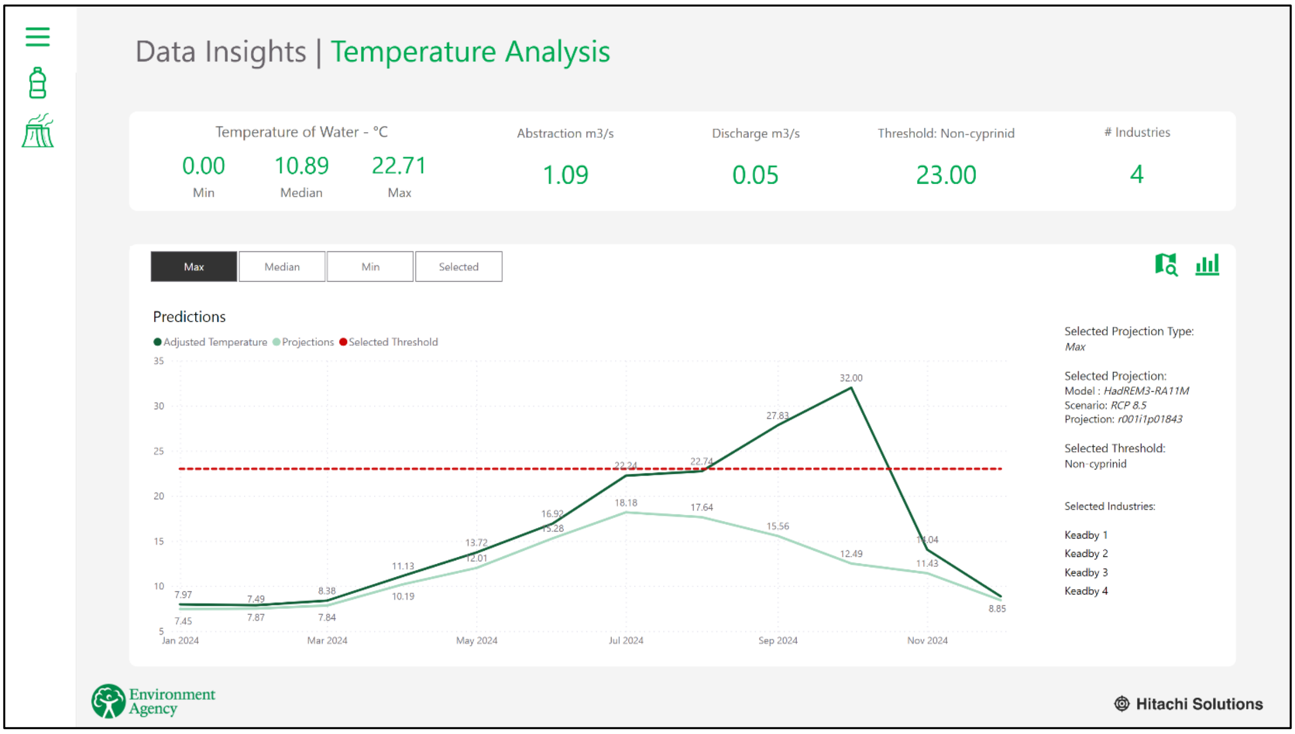


Figure 11: Digital twin simulation showing the impact of thermal discharges from four industries in 2024

Scenario 3: Does climate change constrain the deployment of low carbon technologies?

The PoC digital twin model shows an increase in the number of times the hands-off flow threshold is reached as well as the fisheries and cooling water thresholds overtime, up to 2080, when minimum flows and maximum water temperatures were used. Simulations were also run using the median flow which also showed an increase in the number of times the hands-off flow threshold is reached. These results indicate the environment (available flow and thermal tolerance) would become more restrictive over the next few decades. Unless action could be taken to mitigate the impact, operational activity would have to cease until river flows and thermal capacity increased to acceptable levels. The results of the simulation are summarised in Table 3. Figures 12 and 13 show the digital twin visual outputs.

Table 3: Exceedances of environmental thresholds for flow and thermal discharges for 5-year time periods between 2030 and 2080

Temperature simulations were based on the min flow model. Temperature projection data was not available for the period 2040-2060.

Time period	Industry units	Hands-off flow threshold - minimum	Hands-off flow threshold - median	Lamprey threshold (21°C)	Cooling water intake threshold (23°C)
2030-35	4	58	28	21	17
2035-40	4	61	30	25	19
2065-70	4	60	32	25	26
2070-75	4	57	35	31	29
2075-80	4	60	34	31	30

Figure 12: Statistical model output showing thermal impacts between 2075-2080

Key: Adjusted temperature- thermal discharge impact of the selected industries; Projections- climate model selected for the simulation; Threshold- environmental/ industrial temperature threshold.

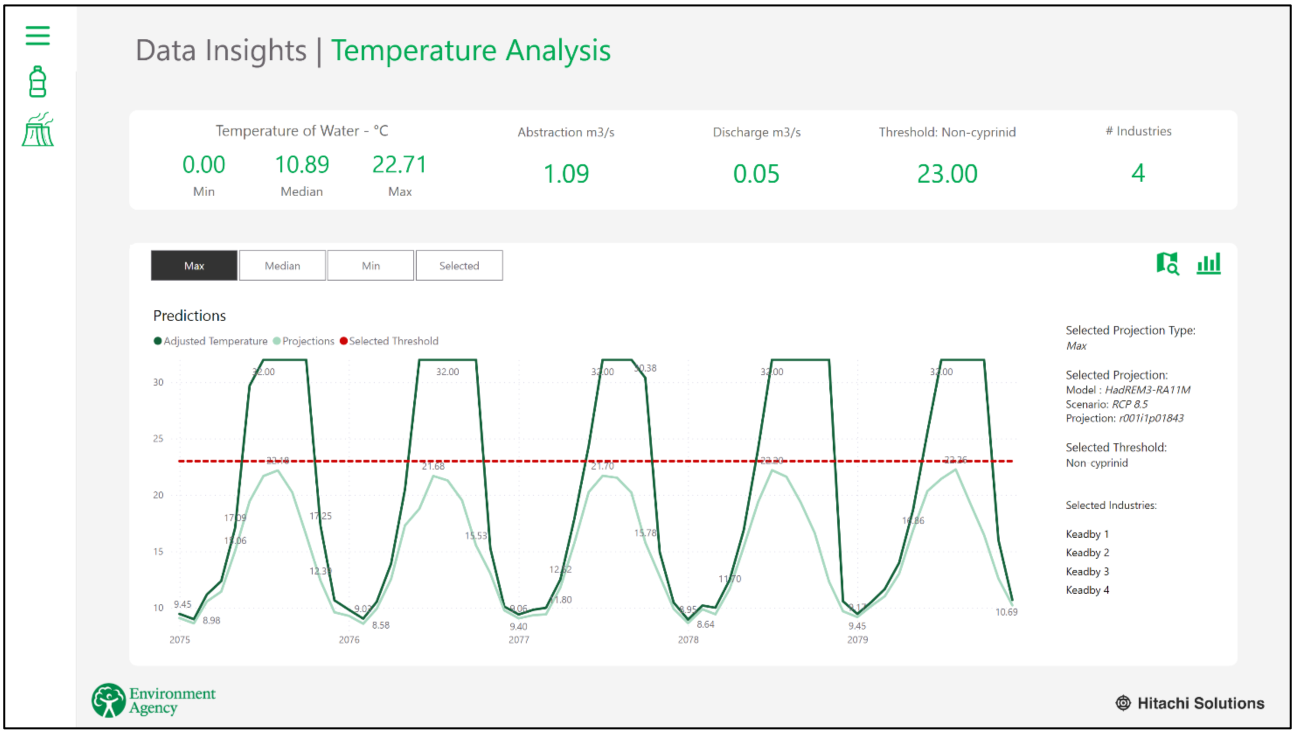
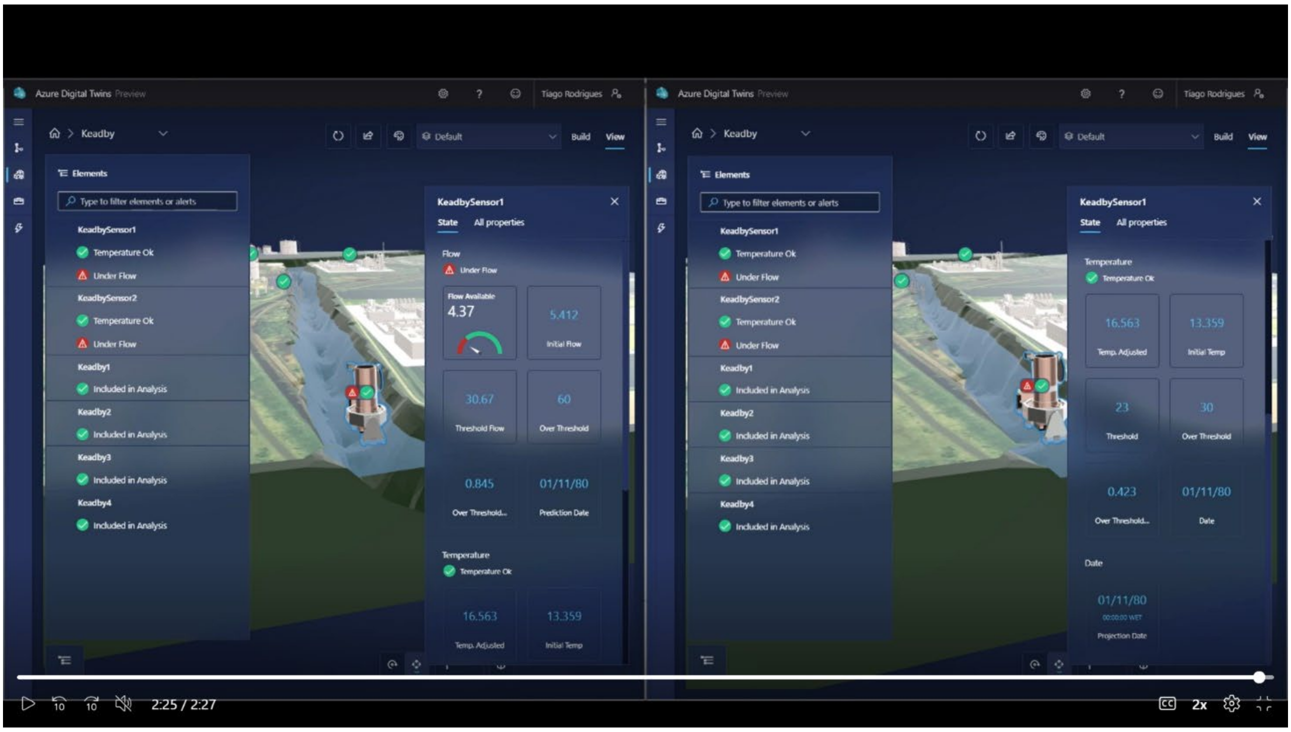


Figure 13: Digital twin simulation showing flow and thermal impacts between 2075-2080



Discussion

At present environmental risk from permitted industrial activities is often considered on site-by-site basis rather than in-combination or at a spatial, cluster-scale. Alone, the impact of industrial activities may not significantly impact the environment, or the impacts can be mitigated. However, where industry is clustered into a small spatial area, the impacts may be amplified or there may be long term unintended consequences. Understanding the cumulative impacts of decarbonising industrial clusters is essential to protect people and the environment, whilst ensuring the technologies are environmentally and economically sustainable over the long term, accounting for increased resource pressures and the impacts of climate change.

The ability to build an industrial cluster within an accurately replicated digital environment would allow policy makers, regulators, local planners, and industry to design clusters that would be compatible with the natural environment, now and in the future, and avoiding unintended consequences.

The PoC has shown how the deployment of a low carbon industry could be constrained by the environment (including in the future) through reduced flows and increased water temperatures. The PoC can be configured to add industrial activities, environmental data, specified times, and thresholds. The ability to include real time sensor data and for the user to be able configure elements adds value to the model by allowing the user to manage water consumption, thermal discharges, or industrial growth, potentially identifying mitigation or alternative solutions where resources are limited. Conversely, the model may show that a favourable outcome cannot be created, indicating where environmental capacity is too constrained to accommodate the proposed activities. The digital twin outputs can be visualised in graph form, allowing further interrogation of the data, and on the Azure digital twin, providing a 3D simulation demonstrating how the industrial activity interacts with the environment over time.

Engaging and informing stakeholders, environmental policy, and regulation

With further development, the PoC has shown that digital twin technology has potential for application by a broad range of stakeholders. There is an opportunity for industry to use digital twin technology to collectively or individually test and design proposals, developing solutions to reduce environmental risk and ensure economic and environmental sustainability. Similarly, this approach could be used by regulators and/or land use planners to test conditions and environmental thresholds in environmental permitting and land use planning decisions. The inclusion of a user-friendly interface to create the clusters, as well as statistical outputs and 3D visualisations means the digital twin can be used to engage and inform a range of stakeholders on proposals, including local public engagement.

The accuracy of the predictions is dependant in the quality and quantity of the data as well as the integrity of the internal models used to represent complex environmental dynamics. At this stage, the PoC model cannot be used to make regulatory decisions or design project proposals. To do this, additional data sources need to be integrated and the environmental interactions within the model refining to give more accurate outputs. However, the PoC provides an important foundation to demonstrate how digital twins could indeed be used to inform regulation, identify environmental risk and potential future environmental or planning issues. The collective deployment of low and zero carbon technologies requires early engagement and the exchange of information between policy makers, regulators, industry, and stakeholders to explore environmental impacts and limits of low and zero carbon clusters. Digital twin technology has the potential to be a key tool in future strategic and regulatory decision making.

Future capability of the model

The PoC digital twin has been designed in a modular form, enabling the geographic coverage to be expanded, new data sources to be added and complex modelling to be integrated to reflect interactions more accurately in the environment. This would require further resource and funding. The limitations of the current model, and potential improvements are discussed below.

Limitations of the proof of concept

The PoC digital twin has not created an exact replica of the environment as accessible data was not available at a spatial and temporal scale sufficient to recreate the environment with the required accuracy and within the time available for this project. As a result, a mix of actual and notional/representative data has been used. The model currently uses simplified logic and assumptions to recreate the complex interactions between the industrial clusters and the water environment. This limits the model, particularly when predicting thermal impacts. Additionally, at present the model includes monthly air temperature predictions which have been adjusted using a small sample of river water temperature data. This limits the model when predicting multiple breaches during warmer months.

At present the PoC digital twin only considers surface waters and does not take the influence of groundwater into account. Hydrodynamic models have not been incorporated into the digital twin due to the short timeframe of the project. However, it is possible in theory to incorporate hydrodynamic models (such as CORMIX) and tidal models into the digital twin in the future. This requires considerable time and resource which was not available during this project. This was a barrier to creating the complexity required to produce a digital twin at a large geographic scale and with the complexity required to accurately reflect the cumulative impacts on the environment. Additionally, for some climate change projections, some data is available annually, other data monthly, limiting the predictive capabilities of the model.

The PoC digital twin was designed to easily include future refined data sources and in a modular form to incrementally improve accuracy, geographic coverage, and complex mathematical models. The digital twin could have incorporated entirely simulated data; but some existing datasets were used to replicate the local environment as closely as possible at this time. The project was an opportunity to determine what information exists and begin to understand the barriers to creating a fully functional digital twin model.

Conclusion

This project has demonstrated that digital twin technology can be used to explore cumulative impacts for the present and into the future. The PoC has proved the viability (and limitations) of the chosen toolset, process, and methodology to assess environmental capacity and cumulative impacts.

With additional spatial and temporal data, and the inclusion of hydrodynamic and water quality models, sections of the Humber Estuary could be replicated more accurately to determine cumulative impacts of new low carbon technologies on the water environment. Some of this data is not yet available for the Humber Estuary, however industry will hold useful baseline data which could be included. The model infrastructure allows new future data sources to be added to improve its accuracy (for example in-situ real time data sensors) as well expand the geographic coverage. However, the larger the geographic area and more complex the data or relationships between the industrial units and the water environment the more resource would be required. Further and wider collaboration is recommended to continue to identify and prioritise sources of data and models for incorporation into the digital twin and develop the framework further. The Environment Agency will incorporate this project and its findings into our on-going communications and engagement work on industrial clusters, and carbon capture and hydrogen deployment. Lessons learnt will also be shared with other digital twin projects being run by the Environment Agency and with the Regulator Pioneer Fund team.

Recommendations

Explore opportunities to:

- Integrate the findings of this project into the Environment Agency's engagement and communications work on industrial clusters and carbon capture and hydrogen deployment.
- Explore potential funding sources to move the digital twin from PoC to a deployed capability that can help inform decision-making.
- Collaborate with industry, the research community and other government organisations (such as Natural England and local planning authorities) who would gain value from this capability.

- Integrate new data sets, such as CS-NOW future flows data (Climate services for a Net Zero resilient world CS-NOW) and English rivers temperature projections (ongoing Environment Agency project).
- Collaborate with and use data held or proposed by industry or submitted to the Environment Agency as part of environmental permit applications. This could be a valuable source of baseline environmental data and operational data for proposed low carbon infrastructure to improve the accuracy of model. This data is not currently collated into a database for use in a digital twin or other modelling purposes.
- Develop a monitoring programme for the Humber to increase the accuracy of the environment being created in the digital twin, as well opportunities to include real time data.
- Incorporate satellite data and potentially install additional sensors to use the model in real time, potentially improving flow and water quality data sets.
- Explore how much of the estuary would need to be built into the digital twin model to accurately determine cumulative impacts from industrial sub clusters and whole cluster. Expand on the model through the modular approach taken in the PoC.
- Develop and incorporate a hydrodynamic model of the Humber Estuary, or relevant areas. A hydrodynamic model of the Tees Estuary is currently being created to understand the impact of industry on the environment.
- Collaborate with and incorporate water company data or hydrodynamic nutrient models required through Price Review 24 (by 2027), including the Humber Estuary.
- Incorporate groundwater models.
- Develop a data storage platform to store historical sensor data (real time data).

References

Bolton A, Enzer M, Schooling J, and others. (2018). The Gemini Principles: Guiding values for the national digital twin and information management framework. Retrieved from Centre for Digital Built Britain.

Climate Change Committee. (2021, June 16). Independent assessment of UK climate risk. Retrieved from Advice to government for the UK's third Climate Change Risk Assessment (CCRA3).

Department for Energy Security and Net Zero. (2023). Powering up Britain.

Environment Agency. (2022). Environmental Capacity in Industrial Clusters Humber Pathfinder. Internal report.

Environment Agency. (2022). Revised guidance for cooling water systems used by the electricity generation industry and associated industries. Draft guidance.

Environment Agency. (2023). Abstraction licensing strategies (CAMS process).

Government Office for Science. (2023). RTA: digital twins.

Humber Industrial Cluster. (2023). Humber industrial cluster plan: background.

Met Office. (2023). UK Climate Projections (UKCP).

Qvist, S. (2019). Nuclear and heat waves: August 2019. Retrieved from Energy for Humanity.

UK Centre for Ecology and Hydrology. (2023). Enhanced future flows and groundwater (eFLaG) portal.

Acknowledgements

The digital twin project team would like to thank Environment Agency colleagues and Hitachi Solutions for their important contributions and support to this proof of concept digital twin project.

Glossary

Azure suite: Used to create digital models of physical environments. Part of the Microsoft package.

Batch data: Large groups of records processed at the same time.

Cyprinid: A family of freshwater fish, including minnow and carp.

Digital twin: A virtual, realistic replica of an object, system or environment which can be modified to understand how changes could happen in the real-world twin.

Environmental threshold: A level which must not be breached to protect people and the environment.

Flow data: Records processed one after the other in a stream of data.

Hands-off flow: River level licensed water abstraction must stop at to protect other abstractors and environment.

Hydrodynamic model: A model showing how the water moves.

IOT central hub: Data platform to hold data and engage with other software. Part of the Microsoft package.

Industrial cluster: A geographic area with a number of industrial sites, of a similar type.

Non-cyprinid: Fish species not covered under the cyprinid classification, such as salmon.

Real time data: Data is delivered immediately from the sensor to the model to be processed. Most real time data is low latency as there is a short delay between the recording of the data and receipt of the data by the model.

Power BI: Interactive data visualisation tool. Part of the Microsoft package.

Power Query: Collect and format data for use in the digital twin model.

Spatial data: Data for a geographic location.

Sub cluster: A small group of industries situated within the industrial cluster.

Temporal data: Data spanning a time period.

Appendix 1: Data sources

Data	Source
Climate change air temperature projections	UK Climate Projections (UKCP) - Met Office
Climate change river flow projections (including model ensembles)	Hydrological projections for the UK, based on UK Climate Projections 2018 (UKCP18) data, from the enhanced future flows and groundwater (eFLaG) project
Gauged river flows	The National River Flow Archive (NRFA)
Industry data	Environmental permits and abstraction licences: public register
River water quality data	Environment Agency's Water Information Management System (WIMS) database