Environmental Science and Policy 164 (2025) 104018

Contents lists available at ScienceDirect

Environmental Science and Policy

journal homepage: www.elsevier.com/locate/envsci

Environmental management using a digital twin

Jennifer M. Durden

National Oceanography Centre, European Way, Southampton SO14 3ZH, UK

ARTICLE INFO

Keywords:
Evidence-based
Conservation
Protected areas
Environmental impact assessment
Big data
Biodiversity

ABSTRACT

Robust environmental management is based on evidence of ecosystem health and anthropogenic harms gleaned from successful environmental monitoring. Successful monitoring involves the synthesis of observations from a variety of sources to represent a site in its current and past states, the anticipation of future conditions, and communicate the findings to decision-makers for environmental management and other stakeholders; a lack of such synthesis and communication has been identified as a shortcoming in Environmental Impact Assessment. However, a suitable digital platform for this synthesis and communication has not yet been developed. Digital twins, an approach from engineering, may offer a solution with advantages over other approaches traditionally employed in ecosystem monitoring. Here a process and considerations for conducting the use case analysis of a digital twin for environmental monitoring is presented, including identifying users, establishing their requirements, refining use cases based on data practicalities, planning analyses and data/model integrations, and developing the user interface. The process is demonstrated using a case study, developing use cases for an ecological digital twin of a UK Marine Protected Area, which could be generalised as use cases for a digital twin for ecosystem monitoring of a conservation area. Considerations for constructing a digital twin based on these use cases are discussed, including the practicalities of using remotely-sensed biological data; gaps in the scientific, technological and data management capabilities; the role of expertise in adding value beyond simple data collation data; and federation of digital twins. Finally, challenges and benefits to using a digital twin approach to informing conservation management are summarised.

1. Introduction

Environmental and ecological monitoring are critical for understanding the health of our planet and anthropogenic harms to facilitate robust evidence-based management; they are also a legal requirement for conducting industrial activities in many jurisdictions and necessary for meeting UN Sustainable Development Goals (United Nations Environment Programme, 2019). Such monitoring involves the synthesis of environmental observations (sometimes in combination with social and cultural data) from a variety of sources to represent a site in its current and past states, and anticipate future conditions. These syntheses provide the evidence base for decision making, important for directing conservation efforts (Sutherland et al., 2004); improving monitoring survey effectiveness and efficiency (Lindenmayer and Likens, 2010; McDonald, 2003); planning industrial, commercial, or agricultural activities as part of Environmental Impact Assessment (Durden et al., 2017b); and anticipating impacts of climate scenarios (Callaghan et al., 2021). Thus, successful environmental monitoring involves both conducting the data syntheses and communicating the findings to decision-makers. Indeed, regulators have identified these aspects as lacking in evaluating Environmental Impact Assessments (Clark et al., 2020). While substantial effort has been invested in developing standard metrics for evaluating ecosystem health (e.g., Essential Biodiversity Variables; Pereira et al., 2013), a suitable digital platform for (a) synthesising the variety of data across space, time and other perspectives in comparison to thresholds, (b) evaluating future scenarios, (c) optimising monitoring design, and (d) communicating findings to decision-makers has not been developed.

A ’digital twin’ may provide such a platform. It is a concept that originated in engineering projects, where a digital representation of a physical object is used to simulate and optimise design options, by incorporating observational data along with modelling and projections related to scenarios. The idea of applying the digital twin approach to environmental evaluation and decision-making has been mooted conceptually (Blair, 2021; Klippel et al., 2021) and at broad scales, related to Earth system and oceanographic applications (Bauer et al., 2021; Chen et al., 2023), agriculture and land degradation (Purcell et al., 2023; Pylianidis et al., 2021), and related to ocean sustainability

E-mail address: Jennifer.durden@noc.ac.uk.

https://doi.org/10.1016/j.envsci.2025.104018
Received 21 June 2024; Received in revised form 30 January 2025; Accepted 1 February 2025

(Tzachor et al., 2023, 2022). The digital twin approach is viewed positively for its facility to integrate varied environmental data and AI techniques, the potential for dynamic updating with real-time data (de Koning et al., 2023), the potential to address anthropogenic impacts including climate change and pollution (Purcell et al., 2023; Tzachor et al., 2023), and their power to improve corporate social responsibility (Xu et al., 2023), but detail on the application of digital twins to environmental and ecosystem monitoring, particularly in scoping them at a local scale, is lacking.

The digital twin approach has advantages over other approaches traditionally employed in environmental and ecosystem monitoring. Spatial management often uses geographic information system-based collation and representation of data; while this presentation perspective is intuitive given the geospatial nature of much environmental monitoring data, the approach can miss out important syntheses (particularly those of a non-spatial nature), often focuses on observational data and lacks functionality for scenario testing (except state-and-transition simulation models; Daniel et al., 2016), may not present univariate / multivariate / qualitative data well, may not combine or connect different types of data or projections (e.g., interdisciplinary observations, modelling-based outputs), and is data-focussed rather than focused on the key metrics for decision-making. Ecological models of many types can be used as a premise for ecological monitoring (Getz et al., 2018; Jørgensen, 2008), but tend to be specialist, focus on a single functional aspect (e.g., stocks and flows) or one scale (e.g., population, community or ecosystem), lack a user interface to deliver metrics to non-specialist decision-makers, and lack the flexibility to facilitate investigation of the underlying survey data in an accessible view. A well-designed digital twin would combine the strengths of both of these approaches, potentially combining observational data with functional modelling (potentially from multiple types of ecological model) and scenario testing, along with an accessible user interface for non-specialist stakeholders and decision-makers (Tzachor et al., 2023). Another key aspect of digital twins is the opportunity for 'federation', or connecting interoperable digital twins, which could also allow for ecological monitoring at a wider scale or meta analyses by connecting digital twins of similar types (e.g., over countries, across major geographic features, or across sites from a particular industry), or allow more robust forecasting by connecting digital twins modelling or generating future scenarios for environmental parameters with others focused on monitoring sites (e.g., examining impacts of future climate scenarios or management actions). The key to achieving a useful digital twin for informing monitoring decisions is scoping it appropriately, a process known as 'developing use cases', but this process has not yet been defined for this application of digital twins. Similar to scoping any model, developing use cases for digital twins for environmental and ecological monitoring involves understanding the wide variety of potential users and their needs, and consideration for the expert input required, data needs, analyses and synthesis, and presentation necessary to address users' needs. While some uses cases have been identified in discussing the application of digital twin approach to environmental and ecological monitoring, no process for conducting a use case analysis has been presented.

In scoping a digital twin, use cases must be feasible and address a decision-making need. One challenge to use case feasibility lies in the data logistics. Environmental data are high in variety (Blair, 2021; Durden et al., 2017a), involving a mixture of data types and scales, including biological data (e.g., organismal, sequencing, observational), environmental parameters (e.g., habitat conditions, physical / chemical / geological / atmospheric / oceanographic parameters), functional or mechanistic linkages, socioeconomic data, and regulatory information; these data occur in multiple dimensions, which presents scientific and technical challenges for integration (e.g., Wicquart et al., 2022). Some data types and historical monitoring data may not be in a digital format, precluding their use in digital twins, though efforts to digitise or rescue valuable historical monitoring datasets are underway (Bledsoe et al., 2022). The uptake of remote sensing for ecosystem monitoring increases the volume of observational data suitable for integration into a digital twin, particularly by increasing the spatial reach and frequency of observations in a digital format, and present opportunities for data generation using AI (e.g., Christin et al., 2019; Høye et al., 2021). This presents challenges for use in a digital twin which are common to Big Data in ecology (Fan et al., 2014; Hampton et al., 2013; Howe et al., 2008; Soranno and Schimel, 2014), such as the lack of data standards which enable data to be Findable, Accessible, Interoperable or Repeatable (FAIR data standards; Wilkinson et al., 2016), though efforts to develop the necessary standards are ongoing (e.g., Durden et al., 2024; Fegraus et al., 2005; Horton et al., 2021; Schoening et al., 2022). Similar challenges were found in developing digital twins for global ocean observations (Snowden et al., 2019) or using Earth observations and physical parameters (Li et al., 2023), suggesting some potential ways to overcome them (Trantas et al., 2023), albeit with less data variety and less variety in desired output metrics for the use cases. Regardless of the anticipated solutions, the challenge of data logistics is an important consideration in developing the use cases for digital twins for environmental and ecological monitoring.

Here a process and considerations for conducting the use case analysis (or scoping) a digital twin for environmental and ecological monitoring and management is presented, including identifying users, establishing their requirements, and refining use cases. The process is demonstrated using a case study, where it is applied to developing use cases for a pilot digital twin of a UK Marine Protected Area. The use cases developed in this case study present a set of generalised use cases for a digital twin for ecosystem monitoring of a conservation area. The practicalities of using remotely-sensed biological data to address the use cases are assessed, and gaps in the scientific, technological and data management capabilities required to integrate such data into a digital twin are identified. The role of expertise in delivering use cases that add value beyond simply collating data, and federation of environmental digital twins are discussed. Finally, benefits to using a digital twin approach to informing conservation management are summarised.

2. Method – A process for use case development for digital twins for environmental management

2.1. Define users and aims

The first step (Fig. 1) is to define the users or target audience for the digital twin – who will use it and what are their main aims? This could simply be defined as whomever is commissioning the digital twin, for example an environmental regulator monitoring the impact of conservation measures or an industrial company reviewing their environmental performance. Although the decision makers are the most obvious users, other users could be included for collaborative decision making and/or communication, for example including all stakeholders involved in the consultation for an environmental impact assessment or conservation area management planning process. A broad set of potential users could include regulatory bodies, government agencies, scientists, companies, industry bodies, environmental consultancies, non-governmental organisations, local community organisations, the public, and users of connected or related digital twins. The main aims of these users (and their detailed needs; see below) may be determined through active engagement (e.g., meetings, workshops, surveys, etc) with them, or from consulting regulatory, legal or governance documents produced or provided by them. The users' aims will imply constraints on the location, spatial / temporal extent and resolution of the digital twin. For example, confined to a conservation area / region or proposed industrial site, and assessing ecological change over a particular period or from a specified activity.

Monitoring of conservation areas typically aims to answer the following questions, sometimes in consultation with stakeholders:

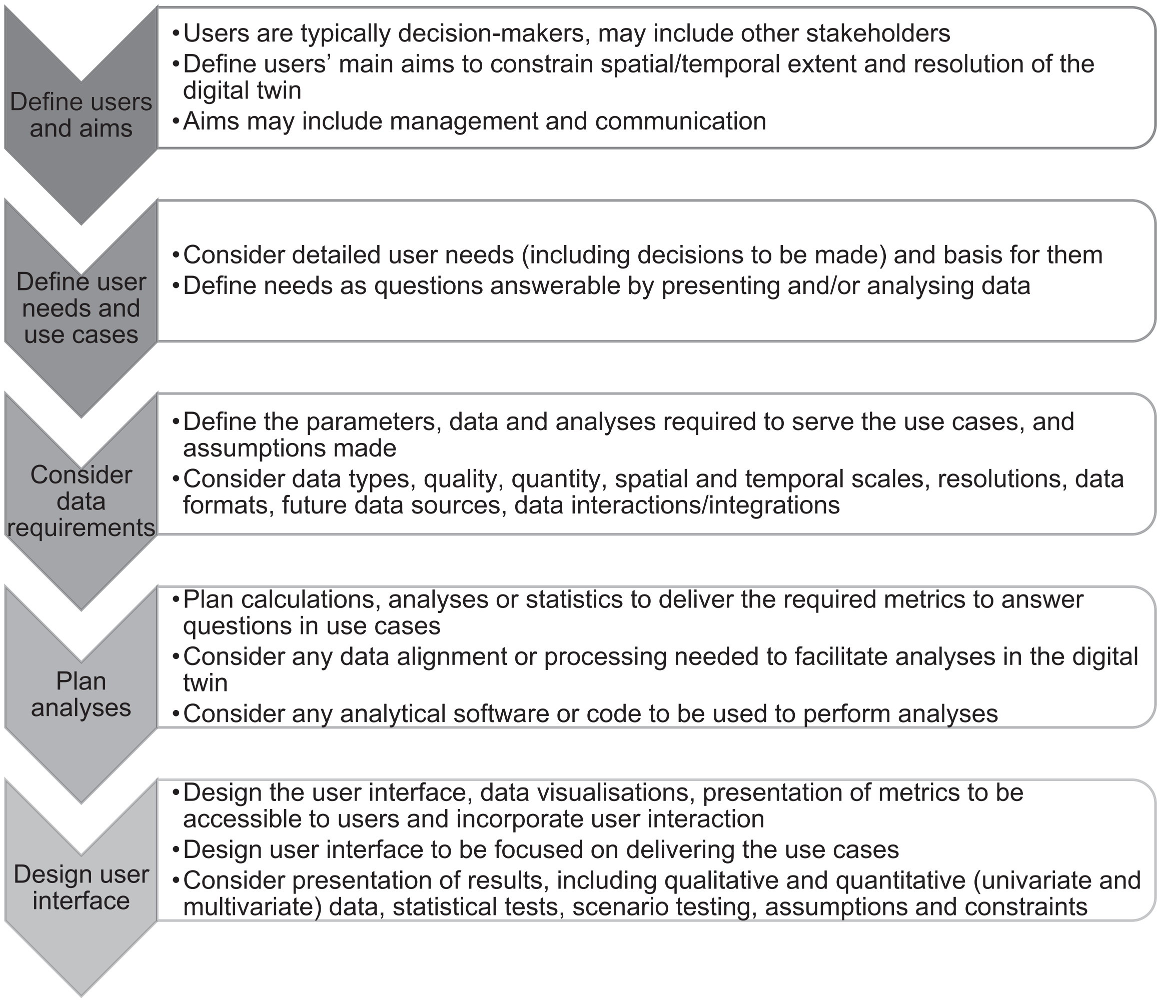

Fig. 1. Steps to defining the use cases (scoping) a digital twin for environmental management and monitoring.

1) What is the condition of the site or area? Is it changing, and if so, how?
2) Are the conservation measures working?
3) What might the future condition(s) of the site be?
4) How might alterations to the conservation measures impact the condition of the site in the future?
5) How can monitoring be conducted more (cost) efficiently, given the existing knowledge of the site?
6) How might stakeholder concerns be addressed, and what impact could those measures have to the site condition?

Preparation for an environmental impact assessment of a proposed industrial activity generally seeks to answer the following questions, and to communicate the findings with stakeholders:

1) What is the condition of the site, and what are the types and magnitudes of natural variability?
2) How might proposed industrial activity impact site conditions? Are the magnitudes and types of impact significantly different to natural variability?
3) How could proposed mitigations or management actions change those impacts to the site?
4) How could environmental impacts of the proposed industrial activity be effectively and efficiently monitored during and after the activity?
5) How might stakeholder concerns be addressed, and what impact could those measures have to the site condition?

2.2. *Define user needs and use cases*

The next step is to define the specific needs of the users and any associated drivers (Fig. 1). Consider the specifics of decisions to be made, the basis or requirements for those decisions, such as applicable regulations, thresholds, metrics, criteria or tests, and details of scenarios that might be tested. The basis for these specific use cases may be regulatory requirements or derived from an ecosystem-based management approach (Tallis et al., 2010). Connections to other digital twins or interoperability with models should be considered; for example, with digital twins to assess future climate scenarios (Voosen, 2020), with models of contaminant plume spread, with ecosystem models, or with digital twins of monitoring equipment such as autonomous vehicles. The specific user needs should be defined in terms of questions that the digital twin could address through the presentation and/or analysis of data. These desired use cases may be revisited, with target users and aims adjusted, to arrive at a set of achievable use cases.

2.3. *Consider data requirements and practicalities*

This step defines what parameters, data and analyses are required to serve the use cases, and the assumptions and expertise required to produce useful metrics and syntheses (Fig. 1). Consider which data may be important for addressing user needs, such as observational and empirical data, and model outputs; parameters representing physical, chemical, biological, geological, geographic, atmospheric, oceanographic, socioeconomic, and/or regulatory aspects (both natural and anthropogenic); and metadata such as locations, dates, times and quality markers. Consider what data are available in a digital format, whether the data meet FAIR data standards, and whether they require any constraint based on expert knowledge. For example, a global dataset of observed presence of organisms may need to be constrained to only those taxa relevant to the site or conservation aims; integration of similar data collected with different equipment or in different units may require expertise to reconcile. The scales of datasets, whether spatial, temporal or other scale, and the ability to integrate across those scales should be considered both within and across parameters. The quantity and quality of data should be sufficiently robust to satisfy any statistical requirements of the use cases. Datasets may have a variety of sources, including publicly-available repositories, institutional repositories, commercial or industry sources, and local indigenous sources; data may be subject to licensing or access constraints. In considering the data availability and practicalities, the use cases may need to be revisited and refined, potentially removing or reframing those that are currently impractical. Finally, consider how future data to be collected will be integrated, what implications this may have to the use cases and technical implementation, and whether there will be feedback between the environmental conditions, monitoring equipment and the digital twin.

2.4. Plan analyses to deliver use cases

Use cases may require calculations, analyses or statistics to deliver the required metrics, and/or the integration of data with models, and these analyses should be planned as part of the use case development (Fig. 1). Analyses may draw on existing or bespoke analytical programs, packages or code, applied to the available data. Expertise may be required to appropriately constrain the data or ensure that analyses are applied where statistical assumptions are met to deliver meaningful outputs rather than allowing spurious ones; for example, data collection methods and units may need to be aligned, data as stored in repositories may need to be aggregated into sample units of appropriate size for statistically-robust analyses, or organism identifications may need to be aligned (e.g., taxonomically or functionally) across datasets. This necessarily requires consideration for the level of processing or analyses required prior to data input into the digital twin, that is, whether the base data are raw or summarised, and whether calculations, analyses, or data extraction (e.g., using AI) occur within the digital twin. Any user inputs or interaction with the digital twin should be planned, along with identifying issues of potential scale mismatches between datasets which may preclude analyses to address the use cases.

2.5. Design presentation of the use cases in a user interface

The use cases should be considered in terms of their presentation in the digital twin, including visualisation and the user interface that does not require expertise (Fig. 1). Pertinent advice comes from a recent study examining digital twins for the brownfield sector, which suggested that the presentation of the digital twin should be visual, intuitive and interactive, while communicating costs, risk and uncertainties (Hammond et al., 2023). As the primary drivers behind the digital twin, the use cases should be in a format or interface that is accessible to the users, and the user interface for the digital twin should be centred on them (rather than on the data). As some ecosystem metrics or indicators are geospatial, the digital twin interface may have a base interface with a geographic information system feel, but in order to provide the meaningful information for decision makers, the structure of the digital twin interface should make the use cases obvious. The analyses or computations performed on the data may have qualitative and/or quantitative results, univariate or multivariate results, may involve statistical tests, may be spatial or temporal in nature, may identify patterns, and may be compared with criteria or metrics as part of the use case; thus, they are likely to involve different presentations to be effective. Consider the presentation of scenario testing and any user interaction, along with users' level of environmental knowledge, management expertise and computer literacy / programming knowledge. The interface design should also include ways for users to examine the underlying data and models; to understand any assumptions, decisions, or constraints imposed on the data; and the magnitudes of confidence and uncertainties in delivering the use cases.

3. Results – A case study of use case development for monitoring a Marine Protected Area

3.1. Users and aims

The main aim was to create a pilot digital twin for the Greater Haig Fras Marine Conservation Zone (jncc.gov.uk/our-work/greater-haig-fras-mpa/), a Marine Protected Area (MPA) that protects approximately 2000 km^2 of seabed in the Celtic Sea west of England. It includes the Haig Fras rocky reef in the Celtic Sea that was designated as a Special Area of Conservation in 2008 (jncc.gov.uk/our-work/haig-fras-mpa/), and the surrounding seabed with substantial areas of mixed rock-sediment habitat. Conservation of the site is monitored by the Joint Nature Conservation Committee on behalf of the UK Government.

The main users of the digital twin were the regulator and researchers that have collected habitat and ecological monitoring data at the site. Representatives from the Joint Nature Conservation Committee were interviewed, and documents produced by them detailing the monitoring aims of the site were consulted (see web links above); their aim is to monitor conditions at the site. Government representatives and the public were identified as users whose aims were to view the data for the MPA and the basic functionality of a digital twin for environmental monitoring. Other minor users included developers of other digital twins, whose aim is to explore the federation between environmental digital twins, and software developers interested in technical aspects of the backend underpinning the outputs of the pilot digital twin (these latter technical aspects are not discussed here).

3.2. Specific use cases

The specific use cases were mainly structured around the conservation aims of the Greater Haig Fras MPA in a document by the UK Government (UK Government Department of Environment and Rural Affairs, 2016). These aims focus on four types of sedimentary habitat that are protected features, one habitat feature of conservation importance, and the Haig Fras rocky reef. The conservation aims require the monitoring of habitat extents and integrity; biological community structure (specifically abundance, diversity), function, and composition in each habitat; and changes over time, including discerning changes from natural versus anthropogenic processes. Each of these aspects is addressed in the specific use cases, which are presented in Table 1. Two main sources of anthropogenic inputs have been identified by the Joint Nature Conservation Committee: fishing and subsea cables. These industrial inputs could be considered as part of assessment of the effectiveness of conservation measures to date, and in terms of future scenarios. These two industrial users were out of scope in terms of developing specific use cases for their needs, but were considered potential users of use cases already established. A use case to monitor species of interest was added to understand changes to the seabed community. Species of interest were identified as either (a) being species of conservation interest listed in the conservation aims (i.e., sea pens) or statistically identified as being characteristic of particular seabed habitats from the initial survey (Benoist et al., 2019), or (b) invasive species.

Users from the Joint Nature Conservation Committee identified that a use case to aid in designing efficient but statistically robust surveys would be valuable, with decision-making focused on the number of seabed images required to characterise the benthic communities found in the different sedimentary habitats on site. A potential future use case could involve connection to a digital twin of the autonomous vehicle

Table 1
Specific use cases for the digital twin case study of the Haig Fras Marine Protected Area, a conservation area in the Celtic Sea where benthic habitats and sea pens are the focus of monitoring. * indicates use case excluded following data logistics analysis.

Use case	Users	Detail / desired metric	Data required	Analyses	Display
Habitat/biotope identification, condition, change	Regulator/ agency, scientist	Determination of habitats/ biotopes (how many, what type, where are they, are they the protected features)	Bathymetry / seabed photograph surveys / habitat maps with habitats identified classified	Automated identification and classification of habitats in images Integration of geological / chemical data	Show locations / extents of habitats Show metrics Show example seabed photos of each habitat type
		Habitat size / distribution (location, extent)		Habitat distribution modelling	Show changes over time Show comparison of data with regulated habitat types and thresholds
		Change in number, types, distribution over time		Calculation of change between time periods (trends, variance)	
Status of and changes to benthic community structure	Regulator/ agency, scientist	Density, diversity of benthic community by habitat/ biotope per survey	Seabed image datasets from across all habitats with known image areas / locations Human-based annotations for individual-based biodiversity Counts and identifications of organisms in other sample types (e.g., sediment cores)	Automated annotations of organisms in images Calculations of biodiversity metrics (density, species richness, etc.) based on sample units derived from aggregated imagery and associated annotations for each time period Employ 'vegan' ecological statistics package in R or similar	Show metrics for each survey and habitat type Show comparisons between habitats Show change(s) over time Show comparison of data with regulated thresholds
		Change over time		Calculation of change between surveys	
Identification and monitoring of species of interest (species of conservation interest and invasive species)	Regulator/ agency, public	Density and distribution of species of conservation interest in different habitats	Seabed image locations in which species of conservation interest found Observations of species of conservation interest in other datasets	Population size, density in each habitat, species distribution models from each survey for each species of conservation interest	Show metrics for each species of interest Show change in metrics or presence over time Show distribution maps in relation to extent of regulated conservation area(s)
		Change over time		Statistical comparisons of above metrics between time points yielding trends, variance	
		Potential detection and location of invasive species	Observations (location, number, date) of invasive species	Challenge: to establish which species may be invasive, with few monitoring surveys	
Optimisation of survey design based on existing site knowledge	Regulator/ agency, scientist	Number of images (or seabed area coverage by images) needed per sample unit	Density and diversity of organisms in previous surveys per image, seabed areas of images	Accumulation curves of density / diversity	User interaction to alter desired statistical power of comparisons Display of impact of user inputs on number of sample units needed
		Number of sample units required	Variation between sample units from previous surveys (from use case above) User selection of statistical power of comparisons	Statistical power calculation (see Durden et al., 2016)	
Testing effectiveness of conservation measures*	Regulator/ agency, scientist	Assess changes to the conditions of habitats and benthic biota in relation to the imposition of conservation measures	Data from the habitat and benthic community assessments above Locations, dates and natures of conservation measures applied Locations and details of fishing activities prior to and since conservation measures applied; locations and dates of cable laying	Statistical analysis of changes to habitat and benthic community between before/after conservation measure implementation (with location considered, as appropriate)	Show change in parameters over time, with date of conservation measures marked Show locations of fishing activities, cable laying and synoptic observations of species of interest, and/or benthic community metrics in habitats/areas in ecologically-relevant proximity to industrial activities
Testing the outcomes of climate scenarios*	Users of other digital twins, regulatory/ agency, scientist	Produce future scenarios of community structure and function at the seabed, including distribution of indicator species, based on climate scenarios	Results of community analysis (above) Climate data (e.g., sea surface temperature) for period of past site monitoring and future scenarios, gridded at a meaningful spatial scale for the site Correlative model or mechanistic understanding of relationship between sea surface temperature and community structure and function at this site	Calculated changes in benthic community over time Challenge: so few time points in monitoring data resulting in no understanding or model of relationship between sea surface temperature and benthic community (either correlative or mechanistic); mismatch in spatial scales between climate model outputs and site observations	Overlay sea surface temperature scenarios Show results of future scenarios of benthic community structure and function
Exploration of available data	Regulator/ agency,	Location / date / extent of environmental surveys and observational datasets	Imagery, bathymetry and observational data with capture metadata	Calculation of perimeter latitude and longitudes, earliest and latest dates per survey	Display datasets as "layers" similar to GIS presentation for exploration

(*continued on next page*)

Table 1 (*continued*)

Use case	Users	Detail / desired metric	Data required	Analyses	Display
	scientist, industry, public	Location / type / extent of modelled data (if any) Location of seabed features for conservation Location of MPA boundaries	Modelled data with location and extent Coordinates from regulator / agency Coordinates from regulator / agency	Calculation of perimeter latitude and longitudes, dates of model	Display perimeter of extent of survey Display season / year of observations / samples
Communication of overview of the MPA	Regulator/ agency, scientist, industry, public	Key locations and metrics about the site Seabed photograph examples	From use cases above	From use cases above Selection of photos of habitats, species of interest	Show locations, photos and metrics in a compelling way for non-specialist audience May need text with site conservation aims / description

conducting a survey, in which the survey design parameters calculated here could be used in vehicle mission control.

Users desiring to federate this digital twin with other digital twins were interested in the potential for assessing the role of climate change in shaping benthic communities at the site. This use case was suggested to involve accepting the output of a digital twin for climate change in Europe under development (Voosen, 2020), as future scenario inputs to this digital twin from which the impacts to benthic communities could potentially detected, should the relationship between surface ocean climate parameters and benthic conditions and organisms be established.

Finally, a use case was defined for communicating important information about the site to the public. Many of these specific use cases are generalisable to digital twins of other conservation areas.

3.3. *Data practicalities*

The use cases for the digital twin require a variety of observational data types, including some monitoring data bespoke to the site and some publicly-available data from external websites. The primary data types required for this digital twin are remotely-sensed observational data from monitoring surveys: bathymetry (Zelada Leon et al., 2020), and seabed photographs (Benoist et al., 2023; Bernardi et al., 2022). Biodiversity and habitat data were extracted manually in these images prior to input to the digital twin, but future digital twins could incorporate artificial intelligence to generate these data (Høye et al., 2021),

As this is a recently-designated site, there are few monitoring surveys; seabed photographs were captured approximately every 3 years over a decade using an autonomous underwater vehicle (Benoist et al., 2019). Biological data (counts, identifications) has been extracted from the images manually by experts with consistent methods and identifications. However, these repeated surveys covered only a small portion of the site (1.9 ha) concentrated in one area away from the rocky reef. Another, unrepeated, photographic survey (JNCC Offshore Benthic Images CEND0513) captured sparsely-spaced images across the remainder of the site, but used entirely different photographic methods, metadata and habitat and organism naming standards. The integration of this other survey data was predicated on aligning these metadata (particularly taxonomic identifications) and reducing the method bias between photographic methods. Secondary data sources to address the use cases included biological observations from a publicly-available dataset (nbn.org.uk/the-national-biodiversity-network/archive-information/nbn-gateway/), constrained to the location and to marine organisms, and publicly-available habitat maps. Although required to address the use cases, suitable data on local fishing pressure or on installation / maintenance of seabed cables were not available, so these use cases were discounted (denoted * in Table 1).

3.4. *Analyses to deliver use cases (Table 1)*

Expert input was required to constrain the publicly-available datasets, to aggregate the observational data into sample units of statistically-relevant size, and to implement the calculation of ecological metrics based on those sample units (based on Durden et al., 2016).

3.5. *Visualisation and user interface*

The use cases suggested that the interface would require a variety of visualisations (Table 1), including map-based graphics with the ability to show numerical, qualitative and statistical metrics shown in scatter or line plots, and descriptions or graphics of assumptions, limitations and methodologies. The use cases required two types of user engagement with the interface: one to allow the user to select the time periods of interest in determining change, and another to facilitate the user selecting statistical power required in assessing the optimal survey design. The users desired a "point-and-click" style interface, rather than one requiring text-based entries or coding.

4. Discussion

The development of use cases for the case study identified some practical challenges for applying digital twins to environmental monitoring. Four major challenges have implications for generating meaningful outputs to address users' needs, and need to be tackled to facilitate the wider adoption of the digital twin approach to ecosystem science, in addition to technical challenges identified by Trantas et al. (2023):

1) Lack of data in some dimensions at appropriate scales, along with a lack of monitoring and data standards

   The issue of the lack of suitable available data has been identified in relation to environmental and ecological digital twins (Blair, 2021; de Koning et al., 2023; Tzachor et al., 2023), particularly as it relates to lack of data meeting FAIR data principles. Lack of digitalisation of data is an issue that is common to environmental restoration, where it also hampers the application of digital twins (Hammond et al., 2023). Solutions to this issue for digital data involve the development of data standards along with a shift in data culture (Durden et al., 2017a). However, the lack of suitable data, along with ensuring data are of sufficient quality, is a problem for monitoring and evidence-based management of the environment more fundamentally, and not just for incorporation into a digital twin; this is not just a lack of FAIR principles, but a lack of data collection. Insufficient suitable data precludes the ability to represent a whole ecosystem, or conduct robust statistics to understand change, and to separate natural variation from anthropogenic change; these problems can be solved by implementing an ecosystem-based monitoring approach (Danovaro et al., 2020), along with well-designed surveys and monitoring plans (Lindenmayer and Likens, 2010) and suitable quality standards (Ferretti, 2011). International monitoring initiatives are developing data standards including quality standards (e.g., Miloslavich et al., 2018; Pereira et al., 2013) as are industries with highly regulated environmental monitoring. Increases to the quantity, spatial reach, timing and frequency of ecological data collection are occurring with the increasing use of remote sensing in monitoring (e.g., Latifi et al., 2023; Willis, 2015), and initiatives to improve best practises and standards related to its uptake (e.g., Hill and Wilkinson, 2004; Hitchin et al., 2015; Obura et al., 2019). Regardless, these limitations of data quantity and quality on the evidence-base must be made clear to decision-makers by acknowledging them during the use case development and implementation, and by presenting this information in the user interface of the digital twin.

2) Lack of mechanistic understanding of causative relationships between parameters

   The lack of mechanistic understanding of causative relationships, particularly between biological or ecological parameters is a problem common to other ecological models (de Koning et al., 2023), and monitoring and scenario evaluation more generally. This lack of mechanistic understanding poses challenges for integrating observational and modelled data, and may obscure or preclude meaningful outputs. Therefore, it is important to develop use cases with this limitation in mind, particularly to understand where scenario testing can provide either simply correlation or a causative understanding that addresses a user need, and to quantify uncertainties and confidence levels where relationships are based on models.

3) Substantial software development requirements

   The construction of a digital twin that implements the use cases identified in the case study involves significant software development infrastructure investment, which is likely beyond the resources of many monitoring projects. The scale and nature of software development required is related to the use cases, data and design selected, but is likely to require technical skills in scientific data analytics, model development, data management, user interface (graphic) design and/or web application development, and a timeline on the order of several months to develop. One solution to reduce this barrier could be to develop standard digital twin structures and interfaces are needed to serve a standard set of use cases for common applications, for example to support Environmental Impact Assessments for a particular jurisdiction or industry, or to manage types or groups of conservation areas. Such standard structures, in combination with data and monitoring standards, would reduce the bespoke software development required for a project and so could facilitate uptake, particularly with government, agencies, or ecologists with limited software development expertise, and facilitate federation (for meta-analyses across the jurisdiction or industry). The set of use cases for ecosystem monitoring for conservation presented here could be generalised for use in building the needed standard structures.

4) Substantial maintenance requirements

Once constructed, the digital twin would require substantial effort to maintain and update it, even in relation to the use cases. The most obvious updates involve incorporating new observational data, particularly from new surveys, or updated model inputs, but maintenance could also include aspects that impact the aims and use cases of the digital twin and their visualisation. This type of maintenance requires expert oversight to ensure that the digital twin continues to provide meaningful outputs; for example, adjusting analyses or metrics to account for changes to data collection methodology or aligning standards; adjusting the integration of modelled and observational data as a result of new mechanistic understanding; altering constraints to externally-sourced datasets; realigning use cases to updated monitoring requirements, legislation or best practises; or introducing interoperability with other digital twins. There is also desire for digital twins to incorporate dynamic data inputs (e.g., de Koning et al., 2023), which would require frequent maintenance to ensure that dynamic inputs are resulting in sensible outputs. Maintenance or updating could be triggered as part of monitoring review, for example updating data inputs and synthetic outputs following periodic surveys or as part of a regulatory reporting program, updating use cases when regulations change, or updating software aspects when monitoring methods and instrumentation alter data types and formats.

5. Conclusions

Digital twins become popular tools across a variety of applications, including some in ecosystem science, but for these tools to be adopted widely, the focus should be on the meaning they can provide, not on the data within them. To centre that focus on meaningful outputs, this paper presents a process and major considerations for developing use cases for a digital twin for environmental monitoring. The process is generalised so that it could be employed in developing use cases for a variety of applications, including conservation area management, environmental assessments and management of industrial developments, or long-term ecological research. The application of this process to developing use cases for a digital twin for environmental monitoring of a Marine Protected Area provides a framework for a standard set of use cases for conservation area management, which could be built into a standard digital twin structure for this application.

In conducting this use case analysis, several benefits to applying digital twins to environmental monitoring became apparent:

1) It offers an approach to integrate and synthesise data beyond disciplinary siloes to create a concept of the site that facilitates interpretation and meaning.
2) It offers a tool for decision-makers to access and interact with environmental data.
3) It is a flexible approach that can accommodate the analysis and presentation of qualitative information, univariate and multivariate analyses along with geospatial assessments, and temporal comparison, in one platform.
4) It can incorporate existing and developing code-based ecological and data science analytical tools (e.g., R packages)
5) It offers a tool for communicating issues of ecological importance to the public, government and other non-scientist stakeholders, including interactive components that convey the potential outcomes of different scenarios.
6) It connects environmental ecological data and analyses to others nationally/internationally and to other data types, accelerating the bridging of gaps.
7) It facilitates connections with other areas of science / expertise that could bring important improvements to ecological data handling and modelling, in turn facilitating making more ecological data FAIR.
8) It can be adapted to incorporate or connect to data pipelines that incorporate AI to generate biodiversity data, reducing the time from data collection to the generation and presentation of evidence for decision-making.

CRediT authorship contribution statement

Jennifer M. Durden: Conceptualization, Methodology, Investigation, Writing – Original draft preparation, Writing – Reviewing and Editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This project was supported by the UK Natural Environment Research

Council (NERC) through the Piloting an Information Management Framework for Environmental Digital Twins project (NE/X016765/1) and AtlantiS Project (NE/Y005589/1). Thanks to Veerle A.I. Huvenne, Daniel O.B. Jones, Fionnuala McBreen, James Albrecht, Ross Bullimore and Anna Downie for helpful comments on the approach.

Data availability

Data are cited in the text.

References

Bauer, P., Stevens, B., Hazeleger, W., 2021. A digital twin of Earth for the green transition. Nat. Clim. Change 11, 80–83.

Benoist, N., Bett, B.J., Morris, K., Ruhl, H., 2023. Greater Haig Fras autonomous underwater vehicle seafloor survey - mosaicked image tiles used to assess benthic assemblages and seabed types (2012), NERC EDS British Oceanographic Data Centre.

Benoist, N.M.A., Morris, K.J., Bett, B.J., Durden, J.M., Huvenne, V.A.I., Le Bas, T.P., Wynn, R.B., Ware, S.J., Ruhl, H.A., 2019. Monitoring mosaic biotopes in a marine conservation zone by autonomous underwater vehicle. Conserv Biol. 33, 1174–1186.

Bernardi, M., Hosking, B., Petrioli, C., Bett, B.J., Jones, D., Huvenne, V.A., Marlow, R., Furlong, M., McPhail, S., Munafò, A., 2022. AURORA, a multi-sensor dataset for robotic ocean exploration. Int. J. Robot. Res. 41, 461–469.

Blair, G.S., 2021. Digital twins of the natural environment. Patterns (N. Y) 2, 100359.

Bledsoe, E.K., Burant, J.B., Higino, G.T., Roche, D.G., Binning, S.A., Finlay, K., Pither, J., Pollock, L.S., Sunday, J.M., Srivastava, D.S., 2022. Data rescue: saving environmental data from extinction. Proc. R. Soc. Lond. Ser. B-Biol. Sci. 289.

Callaghan, M., Schleussner, C.-F., Nath, S., Lejeune, Q., Knutson, T.R., Reichstein, M., Hansen, G., Theokritoff, E., Andrijevic, M., Brecha, R.J., Hegarty, M., Jones, C., Lee, K., Lucas, A., van Maanen, N., Menke, I., Pfleiderer, P., Yesil, B., Minx, J.C., 2021. Machine-learning-based evidence and attribution mapping of 100,000 climate impact studies. Nat. Clim. Change 11, 966–972.

Chen, G., Yang, J., Huang, B., Ma, C., Tian, F., Ge, L., Xia, L., Li, J., 2023. Toward digital twin of the ocean: from digitalization to cloning. Intell. Mar. Technol. Syst. 1.

Christin, S., Hervet, É., Lecomte, N., Ye, H., 2019. Applications for deep learning in ecology. Meth Ecol. Evol. 10, 1632–1644.

Clark, M.R., Durden, J.M., Christiansen, S., 2020. Environmental Impact Assessments for deep-sea mining: Can we improve their future effectiveness? Mar. Policy 114.

Daniel, C.J., Frid, L., Sleeter, B.M., Fortin, M.J., Kriticos, D., 2016. State-and-transition simulation models: a framework for forecasting landscape change. Meth Ecol. Evol. 7, 1413–1423.

Danovaro, R., Fanelli, E., Aguzzi, J., Billett, D., Carugati, L., Corinaldesi, C., Dell'Anno, A., Gjerde, K., Jamieson, A.J., Kark, S., McClain, C., Levin, L., Levin, N., Ramirez-Llodra, E., Ruhl, H., Smith, C.R., Snelgrove, P.V.R., Thomsen, L., Van Dover, C.L., Yasuhara, M., 2020. Ecological variables for developing a global deep-ocean monitoring and conservation strategy. Nat. Ecol. Evol. 4, 181–192.

Durden, J.M., Schoening, T., Althaus, F., Friedman, A., Garcia, R., Glover, A., Greniert, J., Jacobsen Stout, N., Jones, D.O.B., Jordt-Sedlazeck, A., Kaeli, J.W., Koser, K., Kuhnz, L., Lindsay, D., Morris, K.J., Nattkemper, T.W., Osterloff, J., Ruhl, H.A., Singh, H., Tran, M., Bett, B.J., 2016. Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. In: Hughes, R.N., Hughes, D.J., Smith, I.P., Dale, A.C. (Eds.), Oceanogr Mar Biol: Ann Rev. CRC Press, pp. 1–72.

Durden, J.M., Murphy, K., Jaeckel, A., Van Dover, C.L., Christiansen, S., Ortega, A., Jones, D.O.B., 2017b. A procedural framework for robust environmental management of deep-sea mining projects using a conceptual model. Mar. Policy 84, 193–201.

Durden, J.M., Luo, J.Y., Alexander, H., Flanagan, A.M., Grossmann, L., 2017a. Integrating "big data" into aquatic ecology: challenges and opportunities. Limnol. Oceano.: Bull. 26, 101–108.

Durden, J.M., Schoening, T., Curtis, E.J., Downie, A., Gates, A.R., Jones, D.O.B., Kokkinaki, A., Simon-Lledó, E., Wright, D., Bett, B.J., 2024. Defining the target population to make marine image-based biological data FAIR. Ecol. Inform. 80.

Fan, J., Han, F., Liu, H., 2014. Challenges of Big Data Analysis. Natl. Sci. Rev. 1, 293–314.

Fegraus, E.H., Andelman, S., Jones, M.B., Schildhauer, M., 2005. Maximizing the value of ecological data with structured metadata: an introduction to ecological metadata language (EML) and principles for metadata creation. Bull. Ecol. Soc. Am. 86, 158–168.

Ferretti, M., 2011. Quality assurance: a vital need in ecological monitoring. CABI Reviews.

Getz, W.M., Marshall, C.R., Carlson, C.J., Giuggioli, L., Ryan, S.J., Romanach, S.S., Boettiger, C., Chamberlain, S.D., Larsen, L., D'Odorico, P., O'Sullivan, D., 2018. Making ecological models adequate. Ecol. Lett. 21, 153–166.

Hammond, E.B., Coulon, F., Hallett, S.H., Thomas, R., Hardy, D., Beriro, D.J., 2023. Digital tools for brownfield redevelopment: stakeholder perspectives and opportunities. J. Environ. Manag. 325, 116393.

Hampton, S.E., Strasser, C.A., Tewksbury, J.J., Gram, W.K., Budden, A.E., Batcheller, A. L., Duke, C.S., Porter, J.H., 2013. Big data and the future of ecology. Front Ecol. Environ. 11, 156–162.

Hill, J., Wilkinson, C., 2004. Methods for Ecological Monitoring of Coral Reefs: A Resource for Managers. Australian Institute for Marine Science.

Hitchin, R., Turner, J.A., Verling, E., 2015. NMBAQC/JNCC epibiota remote monitoring from digital imagery: operational guidelines. Natl. Mar. Biol. Anal. Qual. Control Scheme Jt. Nat. Conserv. Comm. 25.

Horton, T., Marsh, L., Bett, B.J., Gates, A.R., Jones, D.O.B., Benoist, N.M.A., Pfeifer, S., Simon-Lledó, E., Durden, J.M., Vandepitte, L., Appeltans, W., 2021. Recommendations for the standardisation of open taxonomic nomenclature for image-based identifications. Front Mar. Sci. 8, 620702.

Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., Hill, D.P., Kania, R., Schaeffer, M., St Pierre, S., Twigger, S., White, O., Yon Rhee, S., 2008. Big data: the future of biocuration. Nature 455, 47–50.

Høye, T.T., Arje, J., Bjerge, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C., Raitoharju, J., 2021. Deep learning and computer vision will transform entomology. Proc. Natl. Acad. Sci. 118.

Jørgensen, S.E., 2008. Overview of the model types available for development of ecological models. Ecol. Model. 215, 3–9.

Klippel, A., Sajjadi, P., Zhao, J., Wallgrün, J.O., Huang, J., Bagher, M.M., 2021. Embodied digital twins for environmental applications. ISPRS Ann. Photogramm., Remote Sens. Spat. Inf. Sci. V-4-2021, 193–200.

de Koning, K., Broekhuijsen, J., Kühn, I., Ovaskainen, O., Taubert, F., Endresen, D., Schigel, D., Grimm, V., 2023. Digital twins: dynamic model-data fusion for ecology. Trends Ecol. Evol. 38, 916–926.

Latifi, H., Valbuena, R., Silva, C.S., 2023. Towards complex applications of active remote sensing for ecology and conservation. Meth Ecol. Evol. 14, 1578–1586.

Li, X., Feng, M., Ran, Y., Su, Y., Liu, F., Huang, C., Shen, H., Xiao, Q., Su, J., Yuan, S., Guo, H., 2023. Big Data in Earth system science and progress towards a digital twin. Nat. Rev. Earth Environ. 4, 319–332.

Lindenmayer, D.B., Likens, G.E., 2010. The science and application of ecological monitoring. Biol. Conserv 143, 1317–1328.

McDonald, T., 2003. Review of environmental monitoring methods: survey designs. Environ. Monit. Assess. 85, 277–292.

Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley Jr., D.M., Chiba, S., Duffy, J.E., Dunn, D.C., Fischer, A., Gunn, J., Kudela, R., Marsac, F., Muller-Karger, F.E., Obura, D., Shin, Y.J., 2018. Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Chang Biol. 24, 2416–2433.

Obura, D.O., Aeby, G., Amornthammarong, N., Appeltans, W., Bax, N., Bishop, J., Brainard, R.E., Chan, S., Fletcher, P., Gordon, T.A.C., Gramer, L., Gudka, M., Halas, J., Hendee, J., Hodgson, G., Huang, D., Jankulak, M., Jones, A., Kimura, T., Levy, J., Miloslavich, P., Chou, L.M., Muller-Karger, F., Osuka, K., Samoilys, M., Simpson, S.D., Tun, K., Wongbusarakum, S., 2019. Coral reef monitoring, reef assessment technologies, and ecosystem-based management. Front Mar. Sci. 6.

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H.G., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H.M., Cardoso, A.C., Coops, N.C., Dulloo, E., Faith, D.P., Freyhof, J., Gregory, R.D., Heip, C., Höft, R., Hurtt, G., Jetz, W., Karp, D.S., McGeoch, M.A., Obura, D., Onoda, Y., Pettorelli, N., Reyers, B., Sayre, R., Scharlemann, J.P.W., Stuart, S.N., Turak, E., Walpole, M., Wegmann, M., 2013. Essential biodiversity variables. Science 339, 277–278.

Purcell, W., Neubauer, T., Mallinger, K., 2023. Digital Twins in agriculture: challenges and opportunities for environmental sustainability. Curr. Opin. Environ. Sustain. 61.

Pylianidis, C., Osinga, S., Athanasiadis, I.N., 2021. Introducing digital twins to agriculture. Comput. Electron. Agric. 184, 105942.

Schoening, T., Durden, J.M., Faber, C., Felden, J., Heger, K., Hoving, H.-J., Kiko, R., Köser, K., Karämmer, C., Kwasnitschka, T., Möller, K.O., Nakath, D., Naß, A., Nattkemper, T.W., Purser, A., Zurowietz, M., 2022. Making marine image data FAIR. Sci. Data 9.

Snowden, D., Tsontos, V.M., Handegard, N.O., Zarate, M., O' Brien, K., Casey, K.S., Smith, N., Sagen, H., Bailey, K., Lewis, M.N., Arms, S.C., 2019. Data Interoperability Between Elements of the Global Ocean Observing System. Front Mar Sci 6..

Soranno, P.A., Schimel, D.S., 2014. Macrosystems ecology: big data, big ecology. Front Ecol. Env 12, 3. -3.

Sutherland, W.J., Pullin, A.S., Dolman, P.M., Knight, T.M., 2004. The need for evidence-based conservation. Trends Ecol. Evol. 19, 305–308.

Tallis, H., Levin, P.S., Ruckelshaus, M., Lester, S.E., McLeod, K.L., Fluharty, D.L., Halpern, B.S., 2010. The many faces of ecosystem-based management: making the process work today in real places. Mar. Policy 34, 340–348.

Trantas, A., Plug, R., Pileggi, P., Lazovik, E., 2023. Digital twin challenges in biodiversity modelling. Ecol. Inform. 78.

Tzachor, A., Sabri, S., Richards, C.E., Rajabifard, A., Acuto, M., 2022. Potential and limitations of digital twins to achieve the Sustainable Development Goals. Nat. Sustain. 5, 822–829.

Tzachor, A., Hendel, O., Richards, C.E., 2023. Digital twins: a stepping stone to achieve ocean sustainability? npj Ocean Sustain. 2.

UK Government Department of Environment and Rural Affairs, 2016. No. 9 Wildlife Environmental Protection Marine Management: The Greater Haig Fras Marine Conservation Zone Designation Order 2016, ⟨https://www.legislation.gov.uk/ukmo/2016/9/pdfs/ukmo_20160009_en.pdf⟩.

United Nations Environment Programme, 2019. Measuring Progress: Towards Achieving the Environmental Dimension of the SDGs. United Nations Environmentl Programme.

Voosen, P., 2020. Europe builds 'digital twin' of Earth to hone climate forecasts. Science 370, 16–17.

Wicquart, J., Gudka, M., Obura, D., Logan, M., Staub, F., Souter, D., Planes, S., 2022. A workflow to integrate ecological monitoring data from different sources. Ecol. Inform. 68.

Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J., Groth, P., Goble, C., Grethe, J.S., Heringa, J., t Hoen, P.A., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., van der Lei, J., van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B., 2016. The FAIR Guiding Principles for scientific data management and stewardship. Sci. Data 3, 160018.

Willis, K.S., 2015. Remote sensing change detection for ecological monitoring in United States protected areas. Biol. Conserv 182, 233–242.

Xu, Y., Wang, L., Xiong, Y., Wang, M., Xie, X., 2023. Does digital transformation foster corporate social responsibility? Evidence from Chinese mining industry. J. Environ. Manag. 344, 118646.

Zelada Leon, A., Huvenne, V.A.I., Benoist, N.M.A., Ferguson, M., Bett, B.J., Wynn, R.B., 2020. Assessing the repeatability of automated seafloor classification algorithms, with application in marine protected area monitoring. Remote Sens. 12.