Marine Renewable Energy:
Data Transferability and Collection Consistency

Andrea E. Copping
Alicia M. Gorton
Mikaela C. Freeman

Pacific Northwest National Laboratory

January 2018

Table of Contents

Abbreviations

EMEC	European Marine Energy Test Centre
EMF	electromagnetic field
ICOE	International Conference on Ocean Energy
MRE	marine renewable energy
MREPA	marine renewable energy archetype
PNNL	Pacific Northwest National Laboratory

Introduction

Background

As the marine renewable energy (MRE) industry progresses in US and waters worldwide, the increasing demand for data and information about how MRE technologies (wave and tidal devices) may interact with the marine environment continues. Our understanding of the potential environmental effects of MRE development is slowly increasing, informed by monitoring data collected around devices in several nations and a growing body of research studies. Information derived from monitoring and research is published in scientific journals and technical reports, which may not be readily accessible or available to regulators and other stakeholders.

Regulators at the federal and state level in the US, and analogously in other nations, must satisfy legal and regulatory mandates in order to grant permission to deploy and operate MRE devices. Inherent in these laws and regulations is a concept of balancing risk to the environment and human uses of public resources against economic development and human well-being. Research efforts related to the potential effects of MRE development are focused on this concept of risk, and the interactions between devices and the environment most likely to cause harm, or those for which the greatest uncertainty exists, are garnering the most attention (Copping et al. 2016). The components of risk—probability of occurrence and consequence of occurrence—are fundamental to the process by which regulators evaluate project compliance with environmental statutes. The concept of risk also provides an excellent context for discussing research outcomes and assisting regulators in learning more about potential effects.

The MRE industry is struggling with the high costs of baseline assessments and post-installation monitoring, as well as long timelines for obtaining permits, which lead to uncertainty and risk related to project financing. Regulators require assessment and monitoring information to allow them to carry out the necessary analyses to describe, permit, and manage the environmental risks associated with new MRE technologies and new uses of ocean space. One way to reduce risks to the industry and the environment, and to allow for acceleration of this new form of low carbon energy, could be the ability to transfer learning, analyses, and data sets from one country to another, among projects, and across jurisdictional boundaries.

As the MRE industry matures, the ability to readily transfer research and monitoring results, data, study designs, data collection methods, and best practices from project to project will lead to cost reductions for baseline environmental studies and post-installation monitoring. Regulators and stakeholders currently lack access to synthesized and contextualized data emerging from early-stage projects and there are no mechanisms by which to apply data and information across geographically distinct projects. This leads to each individual project bearing the full burden of information requirements on a site-by-site basis. In addition, data are collected around early-stage MRE devices using many different methods, instruments, and measurement scales. If similar parameters and accessible methods of collection were used for baseline and post-installation monitoring data around all early-stage devices, the results would be more readily comparable. This comparability would lead to a decrease in scientific uncertainty and support a common understanding of the risk of MRE devices to the marine environment. This in turn would facilitate more efficient and shorter permitting processes, which would decrease financial risk for MRE project development.

Purpose of the White Paper

This white paper defines the challenge of data transferability and data collection consistency by applying the state of the knowledge of environmental research, as well as analogous research from other marine industries.

Specifically, this white paper seeks to accomplish the following:

1. Determine methods, criteria, and guidance for allowing the use of MRE environmental effects data collected in one location or jurisdiction to be applied to consenting/permitting processes in another location or jurisdiction.
2. Outline a process for creating best practices for transferring data from one location or project to another.
3. Explore a pathway to developing best practices for data collection to encourage the collection of consistent data types to address each major MRE effect.

State of Knowledge of MRE Environmental Effects

The 2016 Annex IV State of the Science (Copping et al. 2016) report provides the best assessment of the state of knowledge of MRE environmental effects worldwide. The State of Science report was developed using published research, monitoring studies, and the best scientific judgment available at the time; additional papers and reports published since January 2016 have been examined to augment the original assessment. Based on this state of knowledge, it is clear that considerable progress has been made in understanding specific interactions between MRE devices and marine animals, habitats, and ecosystem processes. It is also clear that considerable work is yet to be accomplished: certain interactions need to be discounted or “retired” in order to simplify siting and permitting processes; other interactions will likely require mitigation in order to reduce potential harm to the marine environment.

Through discussions with regulators in the US and abroad, and based on the experience of early-stage MRE developers, it is not clear that the state of knowledge has been clearly communicated and understood by many regulators. It appears that regulators in many jurisdictions are not eager to rely on data sets, information, and outcomes generated from other locations to make permitting decisions.

Needs of the Regulatory Community

The Annex IV team at Pacific Northwest National Laboratory (PNNL) has engaged with the regulatory community in the US (federal and state regulators, resource managers, and advisers) through a webinar held in March 2017, a survey of regulators’ knowledge of and preferences related to permitting MRE development, a second webinar in November 2017, and informal interactions. Through these interactions, it has become clear that there is still a need and an appetite for additional outreach and engagement to ensure that existing information is well known. It is also clear from the survey results and subsequent discussions that data transferability is of interest to the regulators, but most have no clear understanding of how this might work.

Based on the interactions with the regulatory community, progress can be made through three distinct pathways:

1. Information Dissemination – There is a need for wide dissemination of what is known about MRE interactions with the marine environment, and that knowledge needs to be put into context to ensure that regulators and other members of the MRE community have a common understanding of the risks.
2. Data Transfer – A case should be made with regulators that data can be transferred from one MRE project to another, and a set of best practices for data transfer data collection consistency should be developed and promulgated.
3. New Research – Outstanding questions remaining about interactions of MRE developments and the marine environment will require new research. These questions will be collated throughout the process of regulator engagement and the workshop and made available to funding sources.

The Need for Consistency in Data Collection

Inherent in the effort to enable the transfer of monitoring data about MRE devices and their applications from one jurisdiction to another is the need to understand how similar the data might be. Ensuring that the data used from one (origin) location are compatible with the needs of another (target) location, and that multiple data sets from one or more locations can be pooled or aggregated, requires an evaluation of the degree to which the data are consistent. To date, few efforts have prescribed or compared collection methods, instrumentation, or analyses. A key example of this is shown in data collected to evaluate acoustic output from wave devices to evaluate the potential deleterious effects the noise might have on marine animals (Table 1; Copping et al. 2016).

Table 1. Field measurements of acoustic data from Copping et al. (2016) to illustrate the variety of measurements used when collecting environmental effects data.

Project Location	Device Type	Developer, Project/ Device Name	Project Phase	Project Scope	Sound Levels and Pressure Spectral Densities	Organism Type	Results
Strangford Lough, Northern Ireland	Tidal; two 16 m open-bladed rotors, attached to a pile in the seabed in 26.2 m of water	MCT (Marine Current Turbines) SeaGen™	Ambient	Used hydrophones to measure ambient noise	Range of 115 to 125 dB re 1 µPa	NA	High frequencies (200 Hz – 70 kHz) attributed to sound of tidal flow.
			Construction	Measure noise levels of construction activities and marine mammal response to construction noise	• Driving pin-piles: 136 dB 1 µPa at 28 m; 110 dB 1 µPa at 2130 m • Drilling: 20-100 Hz. Equiv. to background noise at 464 m	Harbor porpoise	Temporary displacement of harbor porpoises during construction. Baseline abundances resumed following completion of construction.
			Construction	Calculate the perceived noise levels by marine animals during drilling	• Harbor seal: 59 dB_{ht} at 28 m and 30 dB_{ht} at 2130 m • Herring: 62 dB_{ht} at 28 m and 25 dB_{ht} at 2130 m	Harbor seals, harbor porpoise, herring, dab, trout	Perceived levels of sound from pin-pile driller were generally lower than ambient levels of sound in the narrows. Calculations of perceived noise suggest marine animals in Strangford Lough were unlikely to be disturbed at distances more than 115 m from drilling.
			Operation	Determine harbor seal behavior in area of operating device	Ambient plus device signature	Harbor seals	No significant displacement of seals or porpoises. Marine mammals swam freely in the Lough during operation. Noted evasion at channel center during turbine operation.
Cobscook Bay, Maine, USA	Tidal; a single, barge-mounted, cross-axis turbine generator unit in 26m of water	Ocean Renewable Power Company, Cobscook Bay Tidal Energy Project	Operation	Measure noise levels of the barge-mounted turbine	Less than 100 dB re $\mu Pa^2/Hz$ at 10m	NA	At 200 to 500 m from the turbine, sound was not detectable above ambient noise within the bay.
East River, New York, USA	Tidal; six three-bladed unducted turbines bottom-mounted in 10 m of water	Verdant Power, Roosevelt Island Tidal Energy Project	Operation	Measure noise levels around the array of tidal turbines	Up to 145 dB re 1µPa @ 1m from the array	14 fish species in the area	During the study, blades on one turbine were broken and another turbine was failing, resulting in more noise generation than would be expected. Conclude sound at damaged turbine array did not reach levels known to cause injury for 13 species of fish examined.
Puget Sound, Washington, USA	Wave; 1/7th-scale wave buoy	Columbia Power Technologies, SeaRay™	Ambient and Operation	Measure sound signature of the wave device and surrounding area	• Ambient: 116-132 dB re 1µPa in frequency of 20 Hz to 20 kHz when ships were nearby. • Device: 126 dB re 1µPa	NA	Ambient noise levels masked the wave device sound. Sound from the SeaRay was closely correlated to the wave period.

Overall Roadmap for Data Transferability

We have examined recent permits and licenses in the US and abroad, and conducted a literature review, allowing us to examine and learn about elements of data transferability and collection consistency. We have used this background to develop a plan that will further the community's ability to use MRE environmental data collected from one project location at another.

The data transferability plan consists of the following steps:

1. Develop a framework for examining MRE monitoring data, based on what has been learned from the literature and our familiarity with available MRE data;
2. Gather examples of data for the major interactions of marine animals and habitats with MRE devices, based on State of the Science priorities;
3. Engage with the US regulatory community to determine their impressions of the adequacy and applicability of data collected to date; we expect the reviewers to examine these data within the context of their needs for reviewing and accepting permit applications in their areas of responsibility. We hope to understand their assumptions and concerns about specific aspects of the proposed data transferability and collection consistency, and to understand their appetites for risk, in order to predict their willingness to extend the use of datasets from one location to another. We also intend to introduce the framework to the regulators to garner their impressions of its usefulness.
4. Adjust the context and details of the framework to take into account the knowledge base, impressions, and risk appetite of the regulators.
5. Prepare background material on data transferability, data collection consistency, and the framework for participants in a workshop to be held in conjunction with ICOE in June 2018.
6. Engage workshop participants, with the intent of incorporating their input, to develop a draft best practices document, targeted for September 2018.

Implementation of the plan will be undertaken after sufficient review and acceptance among participants and Annex IV analysts. It is expected that the interactions with US regulators will encourage other Annex IV analysts to carry out similar interactions with their regulators.

Literature Review

Data Transferability

A literature review was conducted to understand how challenges related to data transferability and data collection consistency have been addressed in other industries. The literature review was conducted by reviewing articles found via Google Scholar and Web of Science. Search terms used for the literature review included "data transferability," "environmental data transferability," "data transferability framework," "transferability framework," "data consistency," "data management," "environmental data management," and "data model transferability." This review allowed us to investigate potential data transferability frameworks, models, and approaches, and to determine the limits of data collection consistency in supporting data transferability. The literature that proved to be most pertinent came from a wide range of fields, including economics, transportation, ecology, and land system science. Summaries of the seminal papers are provided in Appendix A.

Several of the reviewed studies focused on data needs and best practices related to data transferability. For example, Briassoulis (2001) presents a policy-oriented analysis of data needs for integrated land-use change. The evaluation concluded that policy-oriented analysis of land-use change requires the following:

- Data must be spatially and temporally compatible, consistent, reliable, easily and inexpensively available and georeferenced.
- Systematic, compatible, consistent, and reliable definitions must be used.
- Compatible, consistent, reliable, easy, and inexpensive data collection procedures must be followed.

A report prepared by the Volpe National Transportation Systems Center (2005) summarizes the results of a peer exchange on data transferability organized and sponsored by the US Federal Highway Administration. The exchange brought together representatives of state and local departments of transportation, metropolitan planning organizations, academics, and transportation consultants. Significant discussions focused on the following topics: developing transferability guidelines to encourage proper data transfer; determining whether certain variables were more transferable than others; and developing requirements for testing data comparability. Drummond et al. (2009) summarized the conclusions and recommendations of the task force that was assembled to investigate the transferability of economic data. The summary recommends several good research practices related to transferability, including recommendations for statistical analyses and modeling, along with guidance when considering the appropriateness of data derived from different jurisdictions.

In addition to data needs and best practices, much of the reviewed literature evaluated statistical methods, models, and frameworks related to data transferability. Vanreusel et al. (2007) investigated the transferability of habitat-based predictive distribution models for two regionally threatened butterflies in northern Belgium. One conclusion of the study was that models depending on area-specific conditions (e.g., landscape structure, microclimate, soil type) may be over fitted to the local conditions, which could limit their transferability. The authors hypothesize that models based on combined data could possibly have greater potential for generalization, leading to a higher potential for transferability. Wenger and Olden (2012) proposed a method for evaluating ecological model transferability through the application of trout species distribution modeling. The authors concluded that traditional linear models have greater transferability, while machine-learning techniques such as random forests and artificial neural networks can produce models with excellent in-sample performance but poor transferability (unless complexity is constrained). Heikkinen et al. (2012) investigated 10 modeling techniques related to (1) species distributions of birds, butterflies, and plants, and (2) climate and land cover in Finland to determine whether good model interpolative prediction accuracy comes at the expense of transferability. The results showed that the machine-learning techniques (MAXENT) and the generalized boosting method (GBM), along with generalized additive models (GAM; a regression-based method), had a desirable combination of good prediction accuracy and good transferability. The authors noted that the challenge of model transferability is due to the need to include all relevant environmental variables without having the model become too complex or over fitted. Rashidi et al. (2013) evaluated the effectiveness of Bayesian updating to synthesize travel demand data as a means of reliably transferring distribution models to areas where data collection is too costly or unfeasible. Of particular interest and relevance, Václavík et al. (2016) investigated the transferability potential of research from 12 regional projects that focused on issues of sustainable land management across four continents. The study used a previously developed concept of land system archetypes (Václavík et al*.* 2013) to estimate the transferability potential of project research by calculating the statistical similarity of locations across the world to the project archetype, assuming a higher degree of transferability in locations that had similar land system characteristics. The results showed that areas of high transferability potential are often clustered around project locations; however, high transferability potentials can be found in geographically distant locations, especially when the values of the considered variables are close to the global mean or when the project archetype is driven by large-scale conditions (e.g., environmental, socioeconomic). The proposed transferability framework presented by Václavík et al*.* (2016) provides a blueprint for research programs that are interested in investigating the transferability potential of place-based studies to other geographic areas, while also assessing possible gaps in research efforts.

Data Collection Consistency

Many of the papers reviewed for data transferability stressed the need for data collection consistency. Biassoulis (2001) explained that different data collection procedures that produce a variety of data, or in this case were collected using different measurements, can greatly affect the transferability of data. Transferability can also be affected by the spatial scale, temporal scale, definition, and context of the data collected (Briassoulis 2001; Volpe National Transportation Systems Center 2005). While admitting that it is not realistic to expect that the same instruments and measurements will be used in the wide array of studies and environmental monitoring, the Volpe National Transportation Systems Center (2005) pointed out that developing common standards for data collection can aid in the comparability of findings and data transferability.

Briassoulis (2001) recommended compatible, consistent, reliable, easy, and inexpensive data collection procedures be followed, but also noted that adopting standardized and uniform procedures is often not realistic unless it is coordinated internationally or by a single agency. In order for collection consistency to be possible for an industry, researchers and developers need to work together to develop best practices for measurements and procedures, at the same time communicating with policy-making bodies or agencies to ensure data collection procedures and measurements produce policy-relevant data that are compatible for use in permitting and consenting (Briassoulis 2001). Volpe National Transportation Systems Center (2005) provided a research plan for "*Identifying Needs and Approaches for Standardization of Travel Model Input Data*" that offers a valuable model for assessing the need for and benefits of collection consistency, associated costs, and practical implementation. The model might be applicable to environmental effects data collected by the MRE industry.

Framework for Data Transferability

From examining the literature, and listening to regulator concerns, it appears that a framework is needed to guide how data generated in one location can be transferred for regulatory use in another location. Such a framework will bring together data sets in an organized fashion, compare the applicability of each data set for use in other locations, and guide the process for comparison and data transfer. The framework proposed here can be used to accomplish the following:

- Develop a common understanding of data types and parameters that are most useful in determining and addressing the potential effects of MRE development.
- Create a set of best practices that will harmonize the consistent collection of data that address key interactions between MRE installations and the marine environment.
- Engage regulators in testing the framework and soliciting input to test the limits of their appetites to embrace data transfer.
- Set limits and considerations for how best practices can be applied to assist with effective and efficient siting, permitting, post-installation monitoring, and mitigation.

Choosing Variables and Data Sets for Transfer

The choice of variables and data sets that might be considered for transfer from one location to another must be driven by regulatory requirements; studies and analyses to date have concentrated broadly on applicable regulations and permit guidelines (Copping et al. 2016). From these studies and analyses, it is clear that a common and consistent set of key interactions can be identified in almost all countries (Table 2); this set of variables and interactions will guide the development of the data transferability framework.

Table 2. Interactions and variables that act as stressors derived from MRE devices and applicable MRE technology types.

Interaction or Variable (Stressor)	Applicable MRE Technology
Risk of marine animals colliding with turbine blades	Tidal
Effects of acoustic output from devices on marine animal behavior	Wave and tidal
Effects of electromagnetic fields from cables and devices on marine animals	Wave and tidal
Changes in nearfield habitat, including reefing of marine animals because of the presence and operation of devices	Wave and tidal
Changes in flow fields, sediment transport, and effects on farfield habitats because of the presence and operation of devices	Wave and tidal

Drivers for Developing a Data Transferability Framework

Examination of the scientific literature about data transferability, discussions with regulators, and examination of recent permits and licenses in the US and abroad clearly indicate that criteria need to be developed for use in transferring data between locations. These criteria will guide the choice of data sets for transfer, form the basis for developing best practices, and help give regulators confidence that their needs are being met. These criteria must include the following:

- comparability of parameters and methods for how the data were collected;
- sufficient description of the physical, chemical, and biological environment to determine comparability among sites;
- assessment of the similarity of MRE technology devices and balance of station; and
- description of the application of the data set for siting and regulatory purposes, at the location of origin.

Framework Outline

The proposed framework will consist of

- a method for describing the environment and evaluating the comparability of data sets (MRE project archetypes);
- a series of steps that will describe the applicability of the framework to MRE technologies; and
- a method for describing the application of a data set from one site to another, to support regulatory processes.

MRE Archetypes

The most promising transferability methodology and framework that might be applied to MRE permitting is gleaned from the literature presented by Václavík et al. (2016) for sustainable land management purposes. The authors' concept of defining a project "archetype" based on a variety of indicators can be applied to other place-based studies, including MRE studies. By adopting the concept of an "MRE project archetype" (MREPA), a combination of stressors, site conditions, MRE technologies, and receptors can be applied to help meet MRE regulatory needs. The comparability between archetypes at the location of origin of the data set and the location to which data will be transferred must be evaluated.

Defining MRE Archetypes

The key premise of the MREPA concept is that MRE projects with like MREPAs will have the highest potential for data transferability. Four variables that define each project MREPA are the stressors, site conditions, MRE technology types, and receptor groups.

Stressors

Portions of MRE devices or other system components affect environmental receptors, such as marine mammals and habitats (Copping et al. 2016). These stressors include

- collision risk;
- effects of underwater noise;
- effects of electromagnetic fields (EMFs);
- changes in nearfield habitats and reefing patterns; and
- changes in physical systems, sediment transport, and farfield environmental effects.

Site Conditions

Information about site conditions at the site of origin and the target site to which data will be transferred is pertinent when determining the data transferability potential. Site conditions have been defined as follows for each of the stressors listed above:

- For collision risk (tidal turbine), the site can match one of four conditions: (1) shallow and narrow channel, (2) deep and wide channel, (3) shallow and wide channel, and (4) deep and narrow channel.
- For effects of underwater noise, site conditions include (1) insolated/quiet environments and (2) noisy environments (acoustic peaks above US regulatory standards for thresholds, broadband).
- For effects of EMFs, site conditions include (1) buried cables, (2) unburied cables laid on the seafloor, (3) shielded cables, and (4) unshielded cables.
- For changes in nearfield habitat and reefing patterns, site conditions include (1) hard bottom habitats, (2) soft-bottom habitats, and (3) in the water column.
- For changes in physical systems, sediment transport, and farfield habitat changes, site conditions include (1) enclosed basins and (2) open coastlines.

MRE Technologies

Each stressor listed can be related to specific MRE technology types, as follows:

- Collision risk can be related to a tidal device that is (1) bottom-mounted or (2) suspended in the water column (floating).
- Effects of acoustic noise can be related to either (1) tidal devices or (2) wave devices.
- Effects of EMFs can be related to (1) seafloor cables or (2) draped cables.
- Changes in nearfield habitats and reefing patterns can be related to (1) foundations/anchors or (2) floats/mooring lines.
- Changes in physical systems can be related to (1) tidal devices or (2) wave devices.

Receptors

Each stressor and MRE technology type has the potential to have an effect on a particular group of environmental receptors:

- For collision risk, receptors include (1) marine mammals, (2) fish, and (3) diving birds.
- For effects of acoustic noise, receptors include (1) marine mammals and (2) fish.
- For effects of EMFs, receptors include (1) elasmobranchs and (2) mobile/sedentary invertebrates.
- For changes in nearfield habitat and reefing patterns, receptors include (1) benthic invertebrates, (2) demersal fish, and (3) shoaling fish.
- For changes in physical systems, receptors include (1) sediment transport and (2) water quality/food web.

MRE Archetype Matrices

A series of matrices have been developed for each stressor that identify the potential site conditions, MRE technology types, and receptors that can be applied to an MRE project at the origin site and at the target site (Table 3–Table 7). From each matrix, an MREPA can be identified for a particular project or set of data that might be useful for transfer. For example, projects related to collision risk have the potential to be classified as one of 22 possible MREPAs based on the project site conditions, MRE technology types, and receptors. Defining the project MREPA is the first step in determining the transferability potential of data from a project, as discussed in the following section.

Table 3. Marine Renewable Energy Project Archetype (MREPA) Matrix for Collision Risk.

Site Condition [(a)]	Technology	Receptors
Shallow and Narrow Channels	Tidal Device, Bottom-Mounted	Marine Mammals
		Fish
		Diving Birds
	Tidal Device in the Water Column	Marine Mammals
		Fish
		Diving Birds
Shallow and Wide Channels	Tidal Device, Bottom-Mounted	Marine Mammals
		Fish
		Diving Birds
	Tidal Device in the Water Column	Marine Mammals
		Fish
		Diving Birds
Deep and Wide Channels	Tidal Device, Bottom-Mounted	Marine Mammals
		Fish
	Tidal Device in the Water Column	Marine Mammals
		Fish
		Diving Birds
Deep and Narrow Channels	Tidal Device, Bottom-Mounted	Marine Mammals
		Fish
	Tidal Device in the Water Column	Marine Mammals
		Fish
		Diving Birds

(a) Shallow channels are defined as having a depth less than 40 m. Deep channels are defined as having a depth greater than 40 m. Narrow channels are defined as having a width of less than 2 km. Wide channels are defined as having a width greater than 2 km.

Table 4. MREPA Matrix for Effects of EMFs.

Site Condition	Technology	Receptors
Buried Cables	Seafloor Cables	Elasmobranchs
		Mobile /Sedentary Invertebrates
Cables Laid on Seafloor	Seafloor Cables	Elasmobranchs
		Mobile/Sedentary Invertebrates
Shielded Cables	Seafloor Cables	Elasmobranchs
		Mobile/Sedentary Invertebrates
Unshielded Cables	Seafloor Cables	Elasmobranchs
		Mobile/Sedentary Invertebrates
	Draped cables	Elasmobranchs
		Mobile/Sedentary Invertebrates

Table 5. MREPA Matrix for Effects of Acoustic Noise.

Site Condition	Technology [(a)]	Receptors
Isolated/Quiet Environment	Tidal Device	Marine Mammals
		Fish
	Wave Device	Marine Mammals
		Fish
Noisy Environment	Tidal Device	Marine Mammals
		Fish
	Wave Device	Marine Mammals
		Fish

(a) Sound levels generally caused by specific portions of each technology: tidal device sound from blade and rotor rotation, as well as power take offs; wave device sound from power take offs. In addition, some lower levels of sound may be generated by mooring systems and interactions between the device and the surface waters, but these sounds were considered to be of less amplitude and unlikely to be of concern for marine mammals (Copping et al. 2016). Isolated/Quite Environments are those with noise measuring less than 80 db. Noisy Environments are those with noise measuring greater than 80 db,

Table 6. MREPA Matrix for Nearshore Changes to Habitat and Reefing Patterns.

Site Condition	Technology	Receptors
Hard Bottom Habitat	Foundation/Anchors	Benthic Invertebrates
		Demersal Fish
		Shoaling Fish
Soft-Bottom Habitat	Foundation/Anchors	Benthic Invertebrates
		Demersal Fish
		Shoaling Fish
Water Column	Floats/Mooring Lines	Marine Mammals and Sea Turtles
		Demersal Fish
		Shoaling Fish

Table 7. MREPA Matrix for Changes to Physical Systems and Farfield Habitat Changes.

Site Condition	Technology	Receptors
Enclosed Basin	Tidal Device	Sediment Transport
		Water Quality/Food Web
Open Coast	Wave Device	Sediment Transport
		Water Quality/Food Web

Applying the Framework

The preferred outcome of applying the data transferability framework is characterization of the level of risk associated with each key MRE technology interaction with the marine environment, simplification of the questions associated with these key interactions, and hence decreased need for extensive onsite data collection or ancillary research studies to elucidate the level of risk. By implementing the data transferability framework, the siting and permitting processes for installation of single MRE devices and arrays could be shortened and scarce funding resources could be directed toward the interactions that remain most uncertain.

Characterize Origin Project

The first step in determining the transferability of data sets from an MRE project is identifying the MREPA for the origin project by examining the stressors, site conditions, MRE technology types, and receptors, as defined above. Figure 1 provides an example of characterizing data from an origin project that investigates collision risk for marine mammals, specifically harbor seals. By following the matrix provided in Table 6, the project is characterized by stressor, site conditions, MRE technology, and receptor.

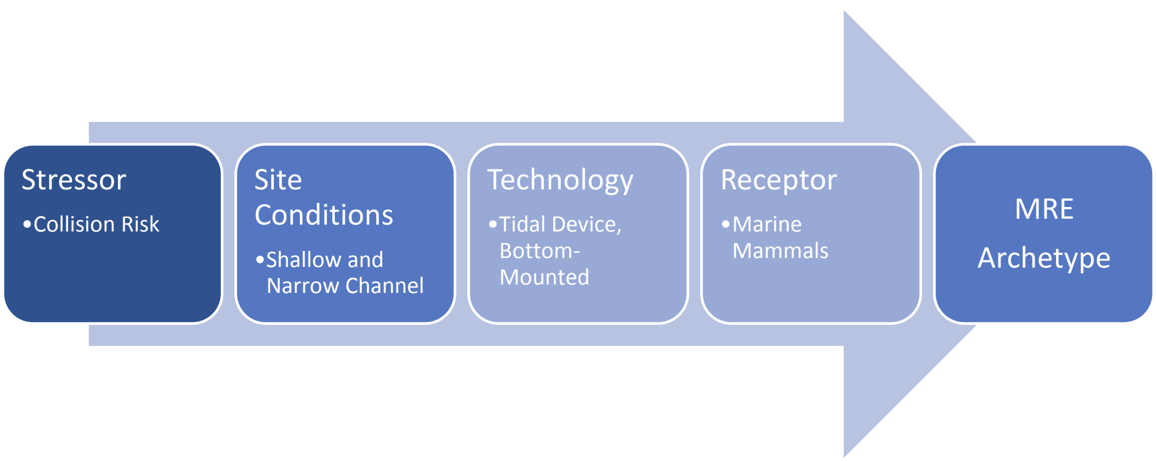

Figure 1.Example of an MREPA for a project site of origin.

Characterize Target Project

Once the MREPA of the origin project site is characterized, the MREPA for the target site also needs to be evaluated. As an example, the potential target project site might require investigation of collision risk to marine mammals, such as endangered killer whales (orca). This comparison assures that the MREPA for the target site is identical to the MREPA of the origin project shown in Figure 1.

Transferability Potential

Rules have been developed to evaluate the potential transferability of data between an origin project site and a target project site (Figure 2). In order for data transferability to be considered, the origin and target project site *must share the same MREPA*, thereby ensuring that the two locations share the same stressors, site conditions, MRE technology types, and receptors. Sharing an MREPA means there is potential for transferring data from the origin project to the target project. Next, the degree of transferability should be evaluated by examining the receptor species, specific technology types, wave or tidal resource, and geographical proximity of the projects to one another. The more variables the origin and target project sites have in common, the more transferable data will be from the origin site to the target project site.

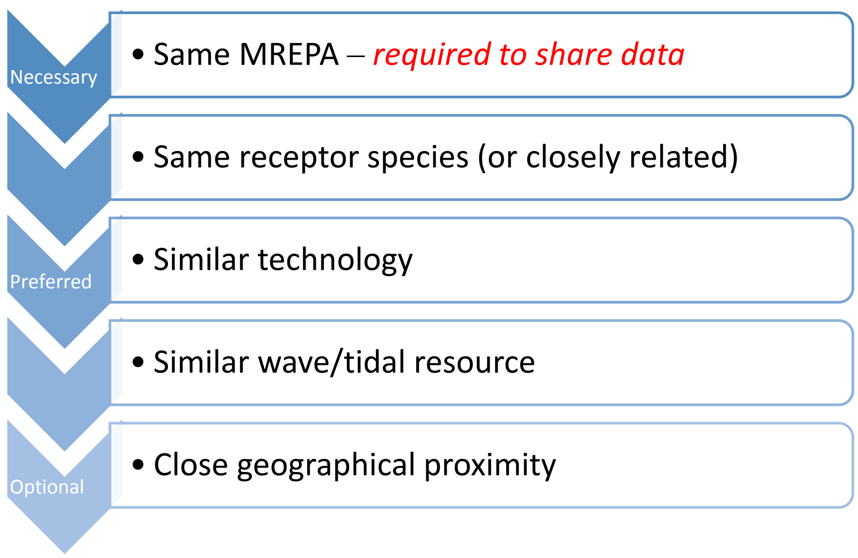

Figure 2. Rules of transferability to evaluate transferability potential.

Interacting with Regulators

The framework is developed to provide a background against which discussions with regulators can proceed to understand the limits of transferability, based on the confidence individual regulators have to accept data and information collected in one location for information analyses of applications for MREs in her/his jurisdiction. The framework will also help to understand where additional data collection, analysis, and interpretation can help increase the degree of data transferability. The following plan lays out the steps and processes for achieving these goals.

Plan for working directly with regulators

The active outreach and engagement plan described here is organized around regulator focus groups. The purposes of the regulator focus groups are as follows:

- Understand regulators' real-world challenges for interpreting data and analyses for MRE projects (or analogous industry projects in the absence of significant experience with MRE applications).
- Share existing data sets with regulators and obtain their feedback on their perceived limitations for accepting data generated in other locations and other jurisdictions for their own regulatory analyses.
- Develop methods for transferring data sets from one project to another.
- Integrate lessons learned from the variety of federal and state regulators who are constrained by differing legal and regulatory regimes for permitting activities in a variety of waterbodies and geographic regions.

Preparation of Data Sets and Example Projects

In preparation for meeting with regulators, synthetic data sets will be acquired and prepared for sharing with the regulators. These data sets will simulate the types of data and information that could be available from the locations of early MRE deployments, representing the origin projects, as described in the framework. The purpose of developing and sharing the data sets is to elicit impressions from the regulators, to understand which aspects of the data they might be comfortable with including in their regulatory analyses, and to understand which aspects continue to concern them.

The MREPA for from which the data are acquired (the origin project) will be constructed and an example project, drawn from the jurisdiction of the participating regulators (the target project), will be constructed for use in demonstrating the framework.

Synthetic data sets will be collected for each of the key stressors (collision risk, effects of underwater noise, EMF effects, nearfield benthic habitat changes, and physical changes), and specific data sets will be targeted for each stressor (Table 8). We will work with Annex IV partners, member nation analysts, and other collaborators to acquire the data.

Table 8. Synthetic data sets to be acquired for interaction with regulators on data transferability.

Stressor	Data Set or Information Source	Comments
Collision risk	Video clips taken around turbines at the European Marine Energy Test Centre (EMEC), other locations in the UK, and Kvichak River.	Clips of both fish and marine mammals
Effects of underwater noise	Sound outputs from wave energy converters at the Wave Energy Test Site and turbines at EMEC; compare to regulatory thresholds.	Marine mammal and fish thresholds

Effects of EMF	Results of Bureau of Ocean Energy Management field experiments and PNNL lab experiments.
Changes in nearfield habitat,	Video clips of outer continental shelf to show consistent soft-bottom habitat from the Pacific Marine Energy Center; video from Admiralty Inlet to show rocky/cobble habitat.
Changes in physical systems	Numerical models of tidal areas and WEC wave energy converter deployment locations.

Regulator Focus Groups

Regulator focus groups will be made up of US state and federal regulators, drawn from existing contacts and those who engaged during the 2017 outreach and survey process.

The goal of the regulator focus groups is to understand regulator acceptance of and concerns about data transferability, and to articulate the real-world challenges regulators have about applying data from origin projects to projects in their jurisdictions. Regulator focus groups may be held in person or online, by region, or by the major concentrations of regulators' requirements (i.e., marine animals, water quality, habitat, etc.).

Each regulator focus group will be conducted to provide information and seek feedback as follows:

- The concept and background information about data transferability will be introduced, as they apply to the current status of the MRE industry.
- Regulators will be asked to articulate their field of regulatory focus, in terms of MRE development.
- The synthetic data sets pertinent to the particular regulators will be presented; the regulators will be asked to identify what they regard as being applicable to their jurisdiction, and what they would still be lacking after viewing the data.
- The concept of the MREPAs will be introduced, along with the rules for acceptance of data sets and information from other projects, and the regulators will be asked to react to the use of the framework.
- Discussions of the need for data collection consistency will be held to ensure that the regulators understand the need to encourage consistent data collection for pre- and post-installation monitoring within a single project and among projects, and so they can provide their input.
- Additional recommendations will be sought from regulators about how we might accomplish the task of data transferability.

Applying Regulator Focus Group Learning and Next Steps

After the regulator focus groups, the knowledge gained from the groups will be brought together and a process for sharing the information at an ICOE workshop developed. The goal of the workshop will be to provide input to a set of best practices for data transferability and consistent data collection, and to provide a semi-quantitative output of the process. Attendance at the workshop will be solicited through direct invitation to key researchers and

MRE developers, as well as Annex IV and Ocean Energy Systems Executive Committee members. ICOE will also advertise the workshop and invite interested registrants.

Logistical preparation for the workshop at ICOE will proceed after the focus groups have concluded. The workshop participants will be provided with a summary of the input from the regulator focus groups, including:

- impressions from US regulators about when and where data sets and other learning can be transferred from one location or project to another;
- challenges noted by the regulators to carrying out that transfer of data; and
- examples where US regulators believe that transferability of data would assist with their regulatory analyses.

The workshop will share the MREPA concept for classifying projects and project sites. Feedback will be sought on improvement of the matrices and rules of use. A semi-quantitative process for evaluating the transferability potential for specific data types will be attempted. Discussions will be held around the need for consistent data collection and particular parameters that are deemed most important to be collected for each key stressor. The primary outcome will be the capture of workshop results that will serve as the foundation for continued work toward development of best practices for transferring data between and among projects.

Next Steps

After the workshop at ICOE, a report will be prepared that summarizes the outcomes of the literature review, preparation and examination of synthetic data sets, regulator focus groups, and the workshop discussions. The report will provide an outline for best practices for data transferability and data collection consistency. Regulators who participated in the process and workshop participants will have the opportunity to review the draft report and their input will be considered during development of the final report in late 2018.

References

Briassoulis; H. (2001). Policy-Oriented Integrated Analysis of Land-Use Change: An Analysis of Data Needs. Environmental Management. Vol. 27. No. 1. pp. 1-11.

Copping, A.; Sather, N.; Hanna, L.; Whiting, J.; Zydlewski, G.; Staines, G.; Gill, A.; Hutchison, I.; O'Hagan, A.; Simas, T.; Bald, J.; Sparling, C.; Wood, J.; Masden, E. (2016). Annex IV 2016 State of the Science Report: Environmental Effects of Marine Renewable Energy Development Around the World. Pp. 224.

Drummond, M.; Barbieri, M.; Cook, J.; Glick, H.A.; Lis, J.; Malik, F.; Reed, S.D.; Rutten, F; Sculpher, M; Severens, J. (2009). Transferability of Economic Evaluations Across Jurisdictions: ISPOR Good Practices Task Force Report. Value in Health. Vol. 12. No. 4. pp. 409-418.

Heikkinen, R.K.; Marmion, M.; Luoto, M. (2012). Does the interpolation accuracy of species distribution models come at the expense of transferability? Ecography. Vol. 35. pp. 276-288.

Rashidi, T.H.; Auld, J.; Mohammadaian, A. (2013). Effectiveness of Bayesian Updating Attributes in Data Transferability Applications. Transportation Research Record: Journal of the Transportation Research Board. No. 2344. pp. 1-9.

Václavík, T.; Lautenbach, S.; Kuemmerle, T.; Seppelt, R. (2013). Mapping global land system archetypes. Global Environmental Change. Vol. 23. pp. 1637-1647.

Václavík, T.; Langerwisch, F.; Cotter, M.; Fick, J.; Häuser, I.; Hotes, S.; Kamp, J.; Settele, J.; Spangenberg, J.H.; Seppelt, R. (2016). Investigating potential transferability of place-based research in land system science. Environmental Research Letters. Vol. 11. pp. 1-16.

Vanreusel, W.; Maes, D.; Van Dyck, H. (2007). Transferability of Species Distribution Models: a Functional Habitat Approach for Two Regionally Threatened Butterflies. Conservation Biology. Vol. 21. No. 1. pp. 201-212.

Volpe National Transportation Systems Center. (2005). Summary Report for the Peer Exchange on Data Transferability: Held December 16, 2004. Pp. 44.

Wenger, S.J.; Olden, J.D. (2012). Assessing transferability of ecological models: an underappreciated aspect of statistical validation. Methods in Ecology and Evolution. Vol. 3. pp. 260-267.

Appendix – Literature Review – Summary of Seminal Papers

Transferability of Economic Evaluations Across Jurisdictions: ISPOR Good Research Practices Task Force Report

Michael Drummond, Marco Barbieri, John Cook, Henry A. Glick, Joanna Lis, Farzana Malik, Shelby D. Reed, Frans Rutten, Mark Sculpher, Johan Severens
Value in Health, Volume 12, No. 4, 2009

In 2004, a task force was put together to investigate the transferability of economic data. Their objectives were (1) to review what national guidelines for economic evaluation say about transferability; (2) to discuss which elements of data could potentially vary from setting to setting; and (3) to recommend good research practices for dealing with aspects of transferability (including analytic strategies and guidance for considering the appropriateness of evidence from other countries). The following is a summary of the conclusions and recommendations of the task force:

- Simple descriptive statistics should be used to examine potential differences among jurisdictions before statistical modeling occurs.
- The level of sophistication of subsequent statistical modeling (i.e., fixed effects vs. random effects) should be guided by the following criteria: (1) number of jurisdictions (e.g., countries, clinical centers); (2) exchangeability or nonexchangeability of data; and (3) the availability of covariates (e.g., at center and country level). With more jurisdictions, partial exchangeability of data, and greater availability of covariates, hierarchical modeling is to be preferred.
- Analysts should carefully consider which parameters need to be jurisdiction specific, wherever possible justifying assumptions with empirically derived data.
- Analysts should use scenario analysis (a form of multiway sensitivity analysis) to explore the implication of different assumptions about economic data transferability.
- There should be more investment in data collection for those parameters that are thought to differ most from place to place.

Effectiveness of Bayesian Updating Attributes in Data Transferability Applications

Taha H. Rashidi, Joshua Auld, Abolfazi (Kouros) Mohammadaian
Transportation Research Record: Journal of the Transportation Research Board, No. 2344, 2013

This paper presents methods for applying Bayesian updating to model the household total number of work trips per day. Bayesian updating has been recognized as having great potential for use in the transportation field, and this paper cites many examples. For local areas where comprehensive data collection is too costly and infeasible, Bayesian updating can be used to synthesize travel demand data. The Bayesian updating method - which gives an updated probability distribution of some variable, model parameter, or other element of interest through a combination of a current sample of data about the attribute and some prior knowledge of its distribution - presents an approach for reliable transfer of models in a scientifically valid way. This study shows that, in general, updating small local samples of travel attribute data with prior information from national data sources provides an improved estimate of local travel attributes compared with using the local sample only. However, including all available historical data in the prior distributions does not necessarily improve the quality of the updating results.

Transferability of Species Distribution Models: a Functional Habitat Approach for Two Regionally Threatened Butterflies

Wouter Vanreusel, Dirk Maes, Hans Van Dyck
Conservation Biology, Volume 21, No.1, 2007

This study tested the transferability of habitat-based predictive distribution models for two regionally threatened butterflies within and among three nature reserves in northern Belgium. The study adopted a functional resource-based concept where a species requires a set of specific resources and conditions to survive and reproduce. The authors used resources directly related to ecological functions (host plants, nectar resources, shelter, microclimate) rather than environmental surrogate variables. All models were transferable among the independent areas within the same broad geographical region. The authors argue that habitat models based on essential functional resources could transfer better in space than models that use indirect environmental variables.

Other general conclusions/observations:

- Models based on combined data could possibly have a greater potential for generalization
- Most predictive distribution models for birds, insects, or other species are landscape-scale models that use large-scale and abiotic variables, including topography and climate. Abiotic variables explain animal distributions most indirectly through correlations of the variables with functional ecological resources. When these correlations depend on area-specific conditions (landscape structure, microclimate, soil type, topography), models could be over fitted to the local conditions. This can be one potential explanation of poor transferability.
- The authors hypothesize that models based on combinations of functional relations are likely to have good transferability among areas given that resource use and resource distribution are similar. For example, the good transferability of fish models compared with models for terrestrial organisms might be due to the fact that fish microhabitat variables such as stream characteristics are similar in range of variation among stream and have a more direct functional relationship to the study species than terrestrial abiotic or biotic variables.

Summary Report for the Peer Exchange on Data Transferability: Held December 16, 2004

Prepared by the Volpe National Transportation Systems Center (US Department of Transportation)

This report summarizes the results of a peer exchange on data transferability organized and sponsored by the Federal Highway Administration (FHWA) Travel Model Improvement Program (TMIP) and co-sponsored by the TRB Committees on Urban Transportation Data and Information Systems (ABJ30), Traveler Behavior and Values (ADB10), and National Transportation Data Requirements and Programs (ABJ10). The exchange brought together representatives of state and local departments of transportation (DOTs), metropolitan planning organizations (MPOs), academics, and transportation consultants. It consisted of presentations on data transferability topics, followed by a discussion of data transferability issues structured around a set of questions prepared prior to the meeting. The following sections summarize the relevant peer exchange discussions.

1. *What are your ideas about data transferability for travel demand/activity models? How do you define "transferability spatially and temporally?*

   Participants agreed there are several "layers" of transferability: (1) a conceptual layer which consists of the modeling structure or mechanisms, (2) the parameters layer, and (3) the outcomes layer.

   Participants felt that data transferability guidelines would be helpful for the entire travel demand modeling community for preventing technically invalid data transfers while encouraging proper data transfer. Standards for transferability of data would lay out criteria and guidelines on what data are transferable, define a correct method to conduct data transfer, and provide a method for measuring whether data transfer was performed successfully and correctly (beyond data matching).

   The importance of determining whether certain variables or types of variables are more transferable than others was also discussed.
2. *For which types of applications does data transferability already occur, and how has the transfer been achieved? Was it successful? What applications have data/parameters that are not typically transferred currently, but might be difficult to estimate originally due to future data limitations? What are some new and different applications for which transferability of data/parameters might work?*

   Participants stated that a prerequisite for successful data transfer is that the source data set and the target data set be comparable. To determine if two data sets are comparable, one should combine the data sets and perform usability and reasonability tests, such as testing whether a variable works the same way before and after the data set combination. One common mistake is to overlook scaling of the data. It would be beneficial to have an outline of some basic requirements for testing data comparability.
3. *What types of data/parameters can be transferred or should not be transferred?*

   Some participants felt that, although temporal transferability is currently used regularly, its validity has not been sufficiently studied by modelers. A controlled study of temporal transferability would help the industry learn how to model temporal changes such as increases in trip rates. Trends over time should also be analyzed to determine whether the context of the data is changing.

Panelists all believed that there are probably key or core variables in travel demand modeling that are transferable and that there are context sensitive variables that are less transferable. Some participants suggested that for the time being, until further research is performed, the most stable variables could be transferred with reasonable confidence.

There are currently some elements for which modelers have little understanding and which are difficult to transfer. For example, care should be taken with constants as they represent factors that the data may not explicitly explain.

4. *How are data/parameters transferred in current applications? What are correct methods for data/parameter transfer?*

   Ideally, modelers should have data from both the source and recipient areas to determine the suitability of the data transfer for each specific case. It is very important that there is a basic understanding of the source and recipient circumstances before transferring data, such as understanding of what type of errors are associated with the source data.
   Modelers should also perform a "goodness of fit" test to determine whether data can be transferred. To aid users in identifying transferability, it would be beneficial to come up with a set of supplementary model specification tests for transferability.
   Agencies should be careful when transferring results from models written using different software. The software being used for modeling can affect the resulting value of the coefficients.
   Data are also sometimes transferred as distributions instead of averages.
5. *What are the implications in using transferred data (e.g. need to use same input variables)?*

   For transferability to be successful, modelers must understand the context in which transferred data were gathered and the context in which models and parameters were estimated. The data generation process could be standardized to include the required context so that data from different regions can be pooled and exchanged.

The following research topics and scopes were chosen by the group for advancing the concepts of data transferability:

A. Identifying needs and approaches for standardization of travel model input data
B. Use of standardized metadata in improving the documentation and transferability of Spatial and travel model data
C. Analysis of temporal stability and dynamics in activity-travel behavior
D. Part 1: Regional impact on travel behavior

Part 2: Drivers of travel behavior
Part 3: Facilitation of travel data and model transferability

E. A guidebook that outlines data transferability issues and guides a user step-by-step through evaluating data transferability
F. Simulation of household activities and travel behavior data
G. Employment data and transferability issues in modeling

Assessing transferability of ecological models: an underappreciated aspect of statistical validation

Seth J. Wenger, Julian D. Olden
Methods in Ecology and Evolution, Volume 3, 2012

This study proposes a method for evaluating ecological model transferability based on techniques currently in use in the area of species distribution modeling. The method involves cross-validation in which data are assigned non-randomly to groups that are spatially, temporally or otherwise distinct, thus using heterogeneity in the data set as a surrogate for heterogeneity among data sets. The authors present an example by applying the method to distribution modeling of brook trout and brown trout in western US. The show that machine-learning techniques such as random forests and artificial neural networks can produce models with excellent in-sample performance but poor transferability, unless complexity is constrained. The authors have found that traditional linear models have greater transferability. Other conclusions of the study include:

- Predictor-response relationships that have a sound ecological basis and direct causal linkages are likely to be more transferable than those based on indirect relationships or pure correlation
- In devising a transferability assessment, the researcher must make several key decisions requiring a degree of professional judgment
  - Deciding how many groups into which to divide the data set which is essentially a decision on how conservative a test to run. The authors found that for the example they present, the fewer the groups, the more conservative the assessment. They expect this to be a general rule and to be true regardless of the size of the data set.
  - Deciding how to assign data to the groups
    - All of the fitting data sets should cover a large portion of the range of variability of the predictor variables of interest.
    - The heterogeneity among the groups (in terms of predictor-response relationships) should be in the range of the expected heterogeneity between the full data set and other locations or the data sets for which inferences are of interest.
- With small data sets, where it is possible for a particular grouping to significantly affect the outcome, it may be useful to repeat the transferability assessment multiple times with different group assignments in a form of ensemble prediction.
- If projections and inferences do not extend beyond the conditions represented by the data used to fit the model, transferability is less relevant.
- Dividing the data into subsets provides some inferences into how a model will perform with a new data set (e.g., a different region or time period), but the actual performance could be substantially better or worse.

Does the interpolation accuracy of species distribution models come at the expense of transferability?

Risto K. Heikkinen, Mathieu Marmion, Miska Luoto
Ecography, Volume 35, 2012

This study investigated 10 modeling techniques on both (1) species distribution of birds, butterflies, and plants and (2) climate and land cover in Finland to investigate whether good interpolative prediction accuracy for models comes at the expense of transferability. Results show that extrapolation to new areas is a greater challenge for all included modeling techniques than simple filling of gaps in a well-sampled area, but there are also differences among the techniques in the degree of transferability. Among the machine-learning modeling techniques, MAXENT, generalized boosting methods (GBM), and artificial neural networks (ANN) showed good transferability while the performance of GARP and random forest (RF) decreased notably in extrapolation. Among the regression-based methods, generalized additive models (GAM) and generalized linear models (GLM) showed good transferability. A desirable combination of good prediction accuracy and good transferability was evident for three modeling techniques: MAXENT, GBM, and GAM. However, examination of model sensitivity and specificity revealed that model types may differ in their tendencies to either increased over-prediction of presences or absences in extrapolation, and some of the methods show contrasting changes in sensitivity versus specificity (e.g., ANN and GARP).
The authors note that the challenge of model transferability is related to the general problem of developing species distribution models that include all important environmental variables yet still are not too complex or overfitted. Model complexity may arise from two sources. First, techniques that effectively fit non-linear trends may be susceptible to producing unrealistically complex response functions between species and environmental factors that do not necessarily generalize to other others. Second, model complexity may arise as a result of inclusion of too many predictor variables. Some methods automatically include all predictor variables in the models and may therefore be inherently prone to overfitting. In theoretical terms, the most overfitting-prone techniques might be those that both allow for complex non-linear responses and automatically include all predictor variables in the models, such as some recent machine-learning techniques.

Policy-Oriented Integrated Analysis of Land-Use Change: An Analysis of Data Needs

Helen Briassoulis
Environmental Management, Volume 27, No. 1, 2001

This paper offers an analysis of the main data issues for the integrated land-use change in the perspective of their utilization in supporting policy design for sustainable land use. The main dimensions of the data are: (1) system of spatial reference, (2) system of temporal reference, (3) definitions, and (4) data collection procedures. The initial evaluation concluded that policy-oriented integrated analysis of land-use change requires that for the most important variables, at least, data are spatially and temporally compatible, consistent, reliable, easily and inexpensively available and georeferenced; that systematic, compatible, consistent, and reliable definitions are used; and that compatible, consistent, reliable, easy, and inexpensive data collection procedures are followed. The following is a summary of each data dimension.

- Spatial dimension: Systems of spatial reference are rarely compatible in terms of level of spatial resolution, coverage, and spatial definition. Different jurisdictions often employ different systems of spatial reference. Additionally, changes in spatial references over time force the analyst to make assumptions to disaggregate available data, which results in the variable itself as well as its relation to other variables being treated inconsistently.
- Temporal dimension: Systems of temporal reference have similar issues as the spatial dimension, where there may be differences in the unit of temporal aggregation, spacing and number of observations, etc. Systems of temporal reference change over time, and if the transition from one system to another is not planned and indicated, data from different systems should be treated differently.
- Definitions: Definitions pertain to the particular ways concepts are expressed and measured. Definitions may vary between jurisdictions and can change over time, which creates problems in the compatibility of data as they may refer to the same variable but are measured differently. The problems are increased when explicit definitions are not given, when changes in definitions are not indicated, or when data from different sources are combined.
- Data collection: Data collection procedures and rules (even for the same variable) differ between agencies as well as between countries unless they are standardized internationally. Data collection procedures change over time with changes in technology, organizations, etc., and they affect the quality of available data.

The author proposed framework guidelines to address the above challenges and data needs.

- Spatial dimension: A system of spatial reference should be established that is GIS-based and should incorporate clear and transparent aggregation algorithms for consistency in applications and spatial transferability of data. Moreover, GIS may provide finer levels of spatial aggregation and a reasonable degree of easy and inexpensive data retrieval. If standardized systems of spatial reference are used internationally, reliable comparisons among different geographical areas can be performed. Once in place, the common spatial system should be used by all disciplines for data collection and reporting.

- Temporal dimension: Systems of temporal reference used by different disciplines should be harmonized so that different types of variables can be analyzed simultaneously. Common denominators for the temporal scales employed in different disciplines should be found as all as rules for valid aggregations of temporal data. Moreover, rules for the dates of data collection should be adopted so that the use of data collected at various dates does not seriously distort the temporal order of the real events. The standardization of the temporal systems should be common to all countries to facilitate comparison and policy-making for and over different geographical areas at different time periods. Temporal standardization should be done, ideally, in conjunction with spatial standardization to secure spatiotemporal compatibility and consistency.
- Definitions: Standardization of conceptual and operation definitions is an absolute necessity. For past data, "translation" rules have to be devised to assist in their consistent use in analyses.
- Data collection: Standardization of several aspects of data collection should be done in conjunction with the suggested standardization of classification systems. At a minimum, the following must be harmonized for the variables concerned: system of spatial and temporal reference for data collection; operational definitions at each level of detail; dates, method (census, survey), and format of data collection; techniques for data cleaning, coding, recording, and updating; technological infrastructure (computers, GIS); and training personnel. Qualitative data collection, especially for past time periods, requires special attention on issues such as: (1) georeferencing the existing historic information, (2) spatial and temporal aspects of historic data, (3) operational definitions of the data collected, and (4) harmonization of historic and qualitative information with quantitative information for the same variable(s). Effective data collection following the proposed guidelines requires a coordinating data management body with a lattice organizational structure; i.e., it will operate horizontally to cover the diverse types of data needed and vertically from the international to the local level.

Investigating potential transferability of place-based research in land system science

Tomáš Václavík, Fanny Langerwisch, Marc Cotter, Johanna Fick, Inga Häuser, Stefan Hotes, Johannes Kamp, Josef Settele, Joachim H Spangenberg, and Ralf Seppelt
Environmental Research Letters, Volume 11, 2016

This study utilizes a previously developed concept of land system archetypes (LSAs) to investigate potential transferability of research from 12 regional projects implemented in a large joint research framework that focus on issues of sustainable land management across four continents. For each project, the authors characterize its project archetype, i.e. the unique land system based on a synthesis of more than 30 datasets of land-use intensity, environmental conditions, and socioeconomic indicators. They estimate the transferability potential of project research by calculating the statistical similarity of locations across the world to the project archetype, assuming higher transferability potentials in locations with similar land system characteristics. Results show that areas with high transferability potentials are typically clustered around project sites but for some case studies can be found in regions that are geographically distant, especially when values of considered variables are close to the global mean or where the project archetype is driven by large-scale environmental or socioeconomic conditions. Using specific examples from the local case studies, the authors highlight the merit of their approach and discuss the differences between local realities and information captured in global datasets. The proposed method provides a blueprint for large research programs to assess potential transferability of place-based studies to other geographical areas and to indicate possible gaps in research efforts.
Study assumptions, details, and conclusions include:

- The authors assume that similarity of land systems constitutes the potential for transferability (i.e. the more similar two sites are in terms of land use, environmental, and socioeconomic conditions), the higher the probability that methods, results, and conclusions from a project site prove applicable at a similar site.
- The authors estimated the transferability potentials for 12 regional projects by calculating the statistical similarity of all 5 arc-min pixels across the world to the unique land system present in each project study areas. They assumed that if the project study area overlaps with a specific LSA, then its research is potentially relevant for other geographical regions that belong to the same archetype.
- First, they analyzed the conditions in each project as reflected by the considered variables and determined the 'project archetype" (i.e. the unique land system in the study area).
- Second, they calculated statistical similarity of the project archetype (represented by each grid cell within the project) to each global grid cell in the multi-dimensional space defined by considered variables, assuming higher transferability potentials in locations with similar land systems. An 'absolute distance' was used as a measure of similarity.kep
- Third, using the inverse of the absolute distance, they mapped the gradient of transferability potentials for each project in the geographical space.

- The authors used ordinary least square regression analysis to examine the relationship between the total variability of conditions in the study area (calculated as the sum of standard deviations for all variables) and the extent of the 'high' transferability level.
- To illustrate the potential effects that differences in global versus local data may have on the final analysis, they replaced the values of 6 original variables (from datasets with a global extent) with those for the same variables from local distances.
- The mapped levels of transferability potential varied regionally, often exhibiting spatial clustering of highly similar conditions around project sites. This patterns suggests that considered land-use intensity, environmental, and socioeconomic conditions are spatially dependent (i.e. autocorrelated) and that calculated statistical distance partially corresponds to geographical distance.
- In contrast, highly similar conditions were found for a number of projects in locations that are geographically distant from the study sites. This was typical for projects where variable values were close to the global mean or where the project archetype was driven by large-scale environmental or socioeconomic conditions.
- The refined analysis of transferability potentials revealed dependency of the results on the resolution and accuracy of the considered input data. Despite the considerable improvements in global-scale geospatial datasets, the main sources of uncertainty remain in the quality of input data and the availability of socio-cultural information in a globally standardized format.
- This new approach illustrates that rather than offering a way to test local-scale transferability of specific findings *per se*, the authors' approach provides a starting point to identify broad-scale regions with potential transferability of place-based research by calculating envelopes that define the general boundaries of projects' relevance outside of their study areas.