Predictive Ecology: a Re-imagined Foundation and Toolkit for Ecological Models

Eliot J. B. McIntire[§1*,2,3] eliot.mcintire@canada.ca ORCID 0000-0002-6914-8316

Alex M. Chubaty[1,3,4*] achubaty@for-cast.ca ORCID 0000-0001-7146-8135

Steven G. Cumming[3*] stevec@sbf.ulaval.ca ORCID0000-0002-7862-2913

Dave Andison[2,5*], andison@bandaloop.ca

Ceres Barros[2] cbarros@mail.ubc.ca ORCID 0000-0003-4036-977X

Celine Boisvenue[1*,2] celine.boisvenue@canada.ca ORCID 0000-0002-6031-7961

Samuel Hache[6], samuel.hache@canada.ca

Yong Luo[1,7*], Yong.Luo@gov.bc.ca ORCID 0000-0002-3748-9773

Tatiane Micheletti[2] tati.micheletti@triade.org.br ORCID 0000-0003-4838-8342

Frances E. C. Stewart[1,8,9] fstewart@wlu.ca ORCID 0000-0001-9344-8346

* Current address/primary affiliation

§ Author of correspondence

1. Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, 506 Burnside Rd W, Victoria, BC, V8Z 1M5, Canada
2. Faculty of Forestry, Forest Resources Management, The University of British Columbia, 2045-2424 Main Mall, Vancouver, BC V6T 1Z4, Canada
3. Département des sciences du bois et de la forêt, Pavillon Abitibi-Price, 2405, rue de la Terrasse, Université Laval, Québec, QC, G1V 0A6, Canada
4. FOR-CAST Research & Analytics, PO BOX 96026 West Springs, Calgary, AB, T3H 0L3, Canada
5. Bandaloop Landscape-Ecosystem Services Ltd., 6552 Littlewood Road, Nelson, BC, Canada, V1L 6S1

6. Canadian Wildlife Service, Environment and Climate Change Canada, Yellowknife, Northwest Territories, Canada
7. Forest Analysis and Inventory Branch, BC Ministry of Forests, Lands, Natural Resource Operations and Rural Development, PO Box 9512 Stn Prov Govt, Victoria, BC, V8W 9C2, Canada
8. University of Victoria, School of Environmental Studies, 3600 Finnerty Rd, Victoria, BC V8W 2Y2, Canada
9. Wilfrid Laurier University, 75 University Ave., Waterloo, Ontario, N2L 3C5, Canada

Statement of authorship

EM, AC, SC conceived of the idea; EM wrote the first draft; DA and SH tested and iterated through the PERFICT approach for management application; EM and AC are lead developers of the SpaDES toolkit; YL, CBa, TM all contributed code to the toolkit; all authors contributed substantially to the ideas; EM, AC, CBa, CBo, YL, TM, FS contributed substantially to revisions. Author order after SC is alphabetical.

Data accessibility statement:

No data were used in this study. All models and code mentioned are freely available and hosted on GitHub.com or on the Comprehensive R Archive Network, with their links provided in the text.

Running title:

Predictive Ecology: A foundation and toolkit

Type of Article:
Methods

Word Counts:
Abstract: 136
Main Text: 4610
Box 1: 170
Box 2: 591

Number of References: 62
Number of Figures: 3
Number of Tables: 1
Number of Boxes: 2

Author of correspondence:
Eliot McIntire, Pacific Forestry Centre, Canadian Forest Service, Natural Resources Canada, 506 Burnside Rd W, Victoria, BC, V8Z 1M5, Canada, Fax: 1-250-363-6004, Phone: 1-250-298-2374 (or temporarily during COVID-19 closures: 1-250-634-3897 -- not a permanent phone number to be published)

"*All models are wrong; some are useful*" George Box, 1951

"*Entities should not be multiplied without necessity*" a version of Occam's razor

"*Predictive Ecology marks a step in science... towards the humble questions that can be answered*" RH Peters 1992, Critique for Ecology

Abstract

Prediction from models and data in Ecology has a long history and can be made from many types of statistical, simulation, and other classes of models. To date, our ability to use the predictive approach as a tool for developing, validating, updating, integrating and applying models across scientific disciplines and to influence management decisions, policies and the public has been hampered by disparate perspectives on prediction and inadequate tools. We present a coherent perspective that follows a **P**redictive **E**cology approach based on 5 principles: **R**eusable, **F**reely available and **I**nteroperable models, built around a **C**ontinuous workflow, which are **T**ested automatically (PERFICT). We describe the SpaDES toolkit that helps implement these principles. We outline some benefits for society of working with these principles, including 1) speeding up scientific advances; 2) data science advances; and 3) improving science-policy integration.

Introduction

All ecologists hope that their models will prove useful. As a result, it is easy to read George Box's famous quote and feel complacent if one's models are not being particularly useful since all *other* models can be collectively criticized as they too are "wrong". It is also widely accepted that ecologists should rely on Occam 's razor which reflects the concept of parsimony to drive our science. Both concepts, however, pivot on a deep subjectivity. Occam's razor is simultaneously an excellent guideline for scientific model development and an ironic jest to the inherent subjectivity of the scientific method. It is easy to allow subjectivity to creep into our perception of data, model choice and structure (Berryman 1992), and decisions based on those. Ecologists continue to build models that are largely under-tested beyond the original study. If ecologists can reconcile these subjective delineations between useful/not useful (or with and "without necessity"), perhaps their relationships with models would also achieve reconciliation. In the case of models informing management and policy -- as one indicator of usefulness -- applied scientific disciplines have long histories (Boutin & Hebert 2002; O'Neill 2002; Fahey *et al.* 2010) with COVID-19 being the most recent and glaring example of how variable the success has been (see Figure 8 in Rijs & Fenter 2020). Within the context of ecological forecasting in the 21st century (Clark 2001), a subjective assessment of a model's usefulness or necessary complexity is no longer a matter of statistical debate: ecologists need the best forecasts for the pressing problems of today (Dietze 2017).

To address these deep struggles along the objectivity-subjectivity continuum during model development, scientists have advocated for a *predictive* approach to *valuing* ecological models (Peters 1977, 1991; McGill *et al.* 2007; Houlahan *et al.* 2015; Mouquet *et al.* 2015; Travers *et al.* 2019). Under this general framework, models that successfully predict the state of a system are inherently more useful than those that do not (Peters 1977). Early calls to this included a critique of *mechanistic* ecological models that allowed for better "understanding" but had predictions that were difficult or impossible to test (Peters 1986). Recent "Ecological Forecasting" allows for a balance between prediction without concern for underlying causes ("empirical" or "phenomenological") and prediction from mechanistic models that may enable forecasting outside of historical conditions (Peters 1991; McGill *et al.* 2007; Evans 2012; Evans *et al.* 2012; Houlahan *et al.* 2015; Mouquet *et al.* 2015; Radchuk *et al.* 2019; Travers *et al.* 2019), echoing "hybrid" models (e.g., Kimmins *et al.* 1999; Boulangeat *et al.* 2014). In parallel, a call for near-term forecasting and a "continuous" perspective on the adaptive modeling following the "predict-test-update-predict" cycle has been made (Dietze *et al.* 2018; White *et al.* 2019). This would dramatically speed up ecological model improvements, be they theoretical, applied, phenomenological, simple or complex. Indeed, economists and weather forecasters update forecasts at regular intervals. Similarly, finding the *right* level of model complexity is a constant challenge (Anderson *et al.* 2000; Horne & Garton 2006; Aho *et al.* 2014). When model selection is between models of vastly different types -- e.g., a simple mathematical model of population dynamics or a complex, spatially-explicit, many-parameter population simulation model -- how do scientists decide what to recommend to assist managers in making decisions about conserving a declining species? We suggest that we still lack (i) a combination of a foundational framework for predictive ecology (PE) as the basis for delineating the thresholds that Box and Occam's razor identified, and (ii) a sufficient toolkit to support this framework.

Ecologists' abilities to build on, reuse and objectively evaluate all but the simplest (e.g., mathematical) models especially in new contexts is rudimentary. This is true for models that vary in complexity from simple statistical models (<5 parameters) to complex simulation models (with >100s of parameters). Attempting to reuse a model built for one place and time to another is, at a minimum, challenging (Yates *et al.* 2018). This can be fraught with challenges arising from data dependencies, software updates, hardware obsolescence, fixed or difficult to get parameters, and continual changes in collaborations with associated loss of expertise (pers. obs. by authors). Some online model banks (e.g., NetLogo model directory https://ccl.northwestern.edu/netlogo/models/) exist but there are no explicit tools for reusing these models (with new data, new study area, or new parameters), apart from becoming equally as proficient with the model code as was the original developer. Alternatively, some models have been built as part of simulation platforms (LANDIS-II: Scheller *et al.* 2007; BIOME-BGC: Bond-Lamberty *et al.* 2015), some with enormous success (Wilensky 1999; Akçakaya & Root 2005; Scheller & Miranda 2015; Schumaker & Brookes 2018). Nevertheless, simulation models are often detached from data and thus require user-driven, and sometimes very extensive, parameterization done independently from running the model. Furthermore, many of the ecological challenges facing the world today cross disciplinary boundaries (Schmolke *et al.* 2010). These platforms have generally been developed within particular disciplines, and often in programming languages that are little known to ecologists, thereby limiting reuse (outside of the original field) and rapid, iterative model development. Therefore, ecologists need a framework for models that can be used widely, that can be deeply rooted in data, and that can transcend disciplinary boundaries, that will combine the rich ideas from the long history of PE (Peters 1982), with near-term, iterative forecasting studies (Dietze *et al.* 2018; White *et al.* 2019) and principles of FAIR (Stall *et al.* 2019) and ARTful (Bodner *et al.* 2020) data.

Here, we present a new framework outlining a reimagined foundation for predictive Ecology, using a PERFICT approach. We then present SpaDES, a toolkit built in R that facilitates an implementation of this framework. We demonstrate numerous benefits that emerge from this framework and show how it is benefiting an applied ecological example currently underway.

The PERFICT approach

The PERFICT approach is an approach to **P**redictive **E**cology that is based on 5 principles: **R**eusable, **F**reely available and **I**nteroperable models, built around a **C**ontinuous workflow, which are **T**ested automatically (Supp. Mat. A for brief version).

Reusable

We define reusability as the ability to take the *algorithms, methods and results* of previous studies “off the shelf” and use them in the same or a new context with little to no changes. Reusability is not a binary: i.e., there are are *levels of reusability* with the benefits of reusability increasing with higher reusability (see Fig. 1), e.g., reusable on the same computer, reusable for a different species or system, reusable for a different study area etc. *Reusability* has the following characteristics about an analysis: it must 1) be scripted, 2) produce the same answer with the same inputs, 3) produce a different, but equivalent, answer with different inputs, 4) have defined inputs and outputs, and 5) work from any system configuration. We consider *reproducibility* as (Borregaard & Hart 2016) a special case of *reusability* that addresses the characteristics 1 and 2 (Begley & Ellis 2012; Klein *et al.* 2014; Nature Editorial 2014; Munafò *et al.* 2017). In our experiences, when inheriting code from a previous project, it has been difficult to arrive at equivalent results. This has been due to several challenges, e.g., final versions of the code and data were difficult to identify (Vines *et al.* 2014), imprecise pseudo-code, manual interventions, missing steps, broken code and unavailable datasets. In some cases, reusability can be achieved by creating functions as they have argument values that can be changed. Often these can be put in a package and hosted in online repositories (e.g., CRAN -- https://cran.r-project.org/). However, functions alone are not sufficient to achieve a robust level of reusability that we present here (see Fig. 2). To leapfrog from one published result to a new project, ecologists need access to actual analyses and model calls, not just the functions themselves.

Freely available

There are many contributions discussing how open science and free, available, interoperable and reusable (FAIR) data helps with the pace of innovation, with transparency, and with the accountability of model predictions (Reichman *et al.* 2011; Stall *et al.* 2019). Having open code and documents allows other scientists to evaluate the implementation of the science, not just the description of the science within a publication. This makes them readily usable and thus testable by others. When model development is open, discussions and challenges can be transparent (e.g., on GitHub issues) and solutions can more rapidly propagate through the community. While a critical component of PERFICT, we do not elaborate on this further here as the benefits for *models* echo those for data as elaborated elsewhere (Stall *et al.* 2019). We note that until all data are FAIR, a hybrid approach to data handling will be required (Fig. 1 light and dark blue module will both be part of many studies).

Interoperable

Interoperability is created by using both modularity *and* standards. Modularity is a description of a component with explicit inputs and outputs with no hidden elements (See Fig. 2). Modular systems have components that can be deconstructed and recombined. The two most important modular design criteria are that modules should 1) be able to run either independently (i.e., decomposable) or as a subcomponent of a larger model (i.e., composable) and 2) communicate with other modules via their defined inputs and outputs (Reynolds & Acock 1997; Voinov et al. 2004). A modularity framework facilitates model comparison and hypothesis testing, while promoting utility, flexibility, adaptability and scientific longevity (Reynolds & Acock 1997). To ensure these modular pieces are interoperable, they must follow standards that define how modules can communicate. This often raises objections, the most common being, “whose standard?” In the case of ecological sciences, this concern is alleviated, in part, as a large community already uses a common language for analyses: the R language. Further alleviating concerns, if the standard is easy to achieve, it is more likely to be used.

Continuous Workflow

Recently, several authors have brought some of the tools and approaches used in the data sciences to ecology and advocated for such workflows to allow for near-term forecasting (Dietze *et al.* 2018; Anderson *et al.* 2019; White *et al.* 2019). A common way for ecologists to *implement* a continuous workflow is through coding all steps, such as data loading, compiling, simulating and reporting, into a single script file. This is generally insufficient due to several challenges. For instance, long computational steps are common and are often manually skipped by a researcher. This manual intervention often masks broken steps as users can make ad hoc adjustments. Further, linear scripts become increasingly challenging to update and understand as they become longer. Building continuous workflows with modular code pieces circumvents many of these challenges (Fig. 1). To support the building of continuous workflows, shared version control tools such as git and cloud hubs such as https://github.com are invaluable. We emphasize that continuous workflows are important in many contexts other than near-term forecasting for specific applied questions. Longer term, strategic planning in many applied fields, e.g., forestry, requires repeated updating of short term goals with long term planning (Paradis *et al.* 2013).

Testing automatically

Testing ecological models is challenging (Oreskes *et al.* 1994). We distinguish two facets of ecological model testing from computer science and ecological science. From a computer science perspective, the testing is to evaluate current and future errors in the *implementation* of the algorithms, such as the software design and translation of mathematics to code. A robust approach to doing each of these comes from software development fields and relies on model creators using abundant code “assertions” (Rosenblum 1995); with assertions active, each time a model runs it is also being tested. Further, developers can create “unit tests” on individual components (e.g., functions). It is only a few extra steps to attach this code to automated continuous integration (CI) systems e.g., such as online services such as GitHub Actions (e.g., via https://github.com). On the other hand, an ecological model can always be improved through calibrating from training data and validating using additional data as our understanding, models, and data are improved. When automated tests are evaluating fit to data (e.g., validations), models can be iteratively improved (Dietze 2017; Parrott 2017) through identifying breakages and validation failures.

Implementing the PERFICT approach: SpaDES

Implementing these principles requires the development of a powerful yet flexible and extensible toolbox. That toolbox would need to have a number of characteristics, especially broad exposure to as many ecological scientists as possible. Therefore, we developed a number of R packages within the SpaDES meta-package (Chubaty & McIntire 2019). If a user uses good coding practices in, say R, version control and continuous integration, they have a strong foundation for creating *some aspects* of the PERFICT approach*.* The SpaDES collection of packages further promotes *freely available* and elevates the *reusability, interoperability, continuous workflow and* automated *testing* required for an operational, yet achievable, implementation of these principles (see Box 1 -- Best Practices).

SpaDES currently comprises five open source packages on CRAN: `SpaDES.core`, `SpaDES.tools`, `SpaDES.addins`, `reproducible`, `quickPlot`, plus several others not on CRAN (but available on https://github.com) including `SpaDES.shiny`, `SpaDES.experiment`, `SpaDES.project`, `petools`, and `pemisc` (see https://spades.predictiveecology.org for links to all packages). The `SpaDES.core` package is built around the concepts of modules and events, similar to other discrete event simulators (e.g., Banks *et al.* 2005). Indeed, some aspects of SpaDES were derived from existing tools with which we had experience (e.g., SELES: Fall & Fall 2001). This foundation provides a generic platform for scheduling arbitrary sequences of modules, enabling pipelining of simple to complex (e.g., linear, cyclic, conditional) sequences. It is indifferent to the data or modeling paradigm (individual/agent based modeling, population modeling, landscape modeling, GIS/raster-based models, statistical models, etc.).

SpaDES facilitates the integration of many model components within and among disciplines as connections are made through shared data. Furthermore, SpaDES provides the infrastructure to build scenarios, experiments, replicates, and ensemble runs, taking advantage of R's parallelism and high performance computing (HPC) capacities. SpaDES packages are *freely* available and open source and every user-facing function is documented and has examples of use as per CRAN policy. The `quickPlot` package allows for visualizations created by *interoperable* modules (e.g., attempts by one module to plot do not interact with another module's plots). The more detailed descriptions of all of SpaDES functionalities are beyond the scope of this manuscript: several vignettes are available on CRAN (https://cran.r-project.org/package=SpaDES) and also can be found in both GitHub (https://github.com/PredictiveEcology) and PredictiveEcology websites (https://spades.predictiveecology.org). While R can interact with code from other languages (e.g., C++, Python, java, julia), the SpaDES framework is currently written in R, requiring that code written in other languages be wrapped for use in R.

To create *interoperability* and enhance *reusability*, SpaDES operates with modules and events as the basic structures of organizing code. A module is a code chunk that represents a coherent idea or concept that is fairly distinct and stand-alone, such as a "statistical analysis of wildlife collar data", "fire simulator" or "GIS analysis of area within 500m of a road network" The SpaDES system defines a standard for metadata (see Supp. Mat. B for example) that identifies algorithmically what a module does. Foremost, the metadata identify a module's *expected* inputs and *created* outputs, making it different from a simple code chunk or function. By building on *expected* inputs, rather than the inputs themselves, a SpaDES module is indifferent to where those inputs come from, enabling *interoperability* and *reuse*. With such an algorithmic representation of its expected inputs and created outputs, a SpaDES module becomes *reusable* because the source of the data is not specified. The content of a SpaDES module is arbitrary and can therefore perform any tasks including data processing, simulation modeling, statistical modeling, visualization, summarizing, validation, data retrieving, and report building.

Within a module, events are arbitrarily complex steps. Each module can have one (e.g., “get tree data from source, clean, and munge it”) or more events (e.g., in a tree dynamics module, events such as “growth”, “calculate competition”, “mortality”, “dispersal”). Events can be scheduled at any time, for any time; they can be run once (e.g., a data preparation module) or many times; at regular intervals or irregular steps; conditional on any arbitrary states or deterministically run at particular times. The SpaDES formulation encourages events to be scheduled within the module itself, so they are not dependent on externalities, formalizing *reusability*. By formalizing modules with events, arbitrarily complex, modular code chunks can be built. By collecting these code chunks into events and modules, rather than linear chunks (e.g., R markdown chunks), they can be included or excluded from a particular project, i.e., *reused* in new ways and shared among researchers.

One of the greatest challenges in creating a *continuous* workflow is that code must be constantly run from start to finish and automatically *tested* or breakages sneak in without being aware of them. This is prohibitive when there are long computational steps.The `reproducible` package has two principle functions to facilitate this. The generic `Cache` that can be used in any context, nested at any arbitrary function depth, and, unlike other R versions of caching, is aware of non-standard R objects that are stored on disk, such as GIS data files. The more specific `prepInputs` is a tool to bring arbitrary local or remotely located data into R that uses checksumming, caching, and a wide array of GIS operations to harmonize spatial data.

Extensible

Any toolkit developed for implementing the PERFICT approach should be able to handle a large array of problems and can be extended by anybody at any time. An important feature of R/SpaDES that distinguishes it from other low-level languages (e.g., Java or C++) or purpose-built modeling platforms (e.g., NetLogo), for example, is that the algorithms can be extended *by the modeler*, i.e., it does not require one of the *software* developers to create, implement and deploy the new algorithm. For example, we built a custom language, NetLogoR (Bauduin *et al.* 2019), which is a reimplementation of the NetLogo language (Wilensky 1999), allowing us to address our GIS needs within individual-based models. Ongoing SpaDES module development with our collaborators and co-authors of this paper includes modules built in Java, Python, and C++. The toolkit must be able to develop rich downstream tools, such visualization, web interfaces (e.g., Supp. Mat. C and D for example web interfaces), analyses of complex simulation experiments, or validation of models. Creating new tools that work with a generic SpaDES module allows researchers to reuse rich components in new contexts. As tools continue to develop, Application Programming Interfaces will likely emerge in other languages (e.g., Python).

Benefits of the PERFICT approach

Based on our experiences leading to the development of the SpaDES package, we have identified a wide range of valuable outcomes coming from the implementation of the PERFICT approach: 1) speeding up scientific advances; 2) data science advances; and 3) improving science-policy integration (Table 1).

Speeding up scientific advances. One of the primary objectives for model estimation as a process that advances science is to avoid overfitting (e.g., via approaches such as AIC).

Minimizing overfitting by using a single dataset will have limited success (Reunanen 2003); using “independent” data for validation (e.g., predictive validation Power 1993; Wenger & Olden 2012) is usually recommended (Reunanen 2003). Nevertheless, the widespread use of independent data for fitting/validating is limited, because, we believe, most models and projects are not ready when independent data become available, particularly if the models have complex data requirements. Using modular, reusable and interoperable models can greatly contribute to speeding up scientific advances by enabling faster and iterative re-evaluation and updating of these models (and model fit) -- by the original model creators or others -- when new data become available for validation and/or prediction. As the number of such models grows, models can routinely become part of meta-model comparisons and the appropriate level of complexity can be determined. This will help overcome the “dinosaur problem of simulation models,” where models get “bigger, bigger, bigger, useless” (H. Kimmins, pers. comm.) because there is always another process that seems critical to include. Over time, forecast success from models will improve and the forecast horizon will extend outwards (Petchey et al. 2015). Furthermore, scientists will have access to complete model objects (e.g., *sensu* R language), to which statistical and graphical “methods” (e.g., R functions like predict, AIC, drop) can be applied -- instead of tables of coefficients -- from published work. Using fully functioning models from other researchers, we gain more power and flexibility for forecasting, for iterative improvements (e.g., because they contain the variance-covariance structures), for meta-modelling and testing alternative hypotheses (e.g. ensemble or consensus forecasts; Marmion et al. 2009), and even near-automatic meta-analyses across studies and systems (Hedges et al. 1999, Koricheva et al. 2013). Finally, rewriting widely used models, while labour intensive, can be profitable for the broader community through increased interoperability and reusability (Thiele & Grimm 2015). A community of contributors accelerates the implementation of new insights (e.g., data inadequacies, ecological processes) and helps discover and fix bugs (Boisvenue et al. in prep; Barros et al. in prep) and with internal modularity building and adding new components will be easier in the future. With the PERFICT approach, every project can be its own collection of modules; modules can be added, removed, adjusted depending on their (automatic) testing against data. This modular complexity is the practical implementation of Occam's razor.

Data Science advances. One of the reasons to adopt the PERFICT approach is to build formalization and thus gain powerful tools from neighbouring data science field – i.e. pipelines (Beaulieu-Jones and Greene 2017), online databases and repositories (using or building), online services (e.g., Google Earth Engine, Moore and Hansen 2011), online data visualization (e.g., leaflet; Crickard III 2014) and web applications (e.g., shiny; Chang et al. 2019) – which can be algorithmically linked throughout a project. These links can be made by a data-savvy scientist, built into functions, packages and modules, and then used more broadly. For example, user access control (UAC) is a reality for many datasets: not all datasets are yet FAIR (Stall *et al.* 2019). Building on top of UAC tools (e.g., Google Authentication), users of our SpaDES modules will be allowed to give credentials when required, without breaking the continuous workflow. When a new user downloads a module, the module automatically downloads the data it needs, assisting the user with advanced tools such as checkpointing and spatial cropping, projecting, masking and data integrity checking. By maintaining the connection to the original data sources, a user can get updates as needed. These links between data and models also enable a quicker re-parameterization and re-validation against new data or when using the model in a new study area. Following the PERFICT approach, parameter estimation modules and validation modules can be developed and included as a part of a project to link both calibration data and validation data (See Fig. 1; Barros et al in prep.), and continuous parameterization and validation can be realized. Furthermore, the PERFICT approach, e.g., via predictive validation, creates a formal and rapid way to let the data tell us which data are better for a particular question. This is particularly important in Ecology, where various data sources whose quality and quantity range widely, e.g. high quality but modestly sized field data vs. very large remote sensing datasets (of varying quality).

Improving science-policy integration. The approach used by the Intergovernmental Panel on Climate Change (IPCC) provides a template for science-policy integration. The IPCC brings together more scientists than a typical research project, runs many different models, and integrates model outputs to test hypotheses, to understand model uncertainty and divergent or common outputs, to build iterative forecasts of the future, and to compare data as the future represents a forecasting-based hypothesis test (https://www.ipcc.ch/assessment-report/ar6/). Replicating this approach for every applied ecological problem will require major improvements in how ecologists integrate across scientific disciplines and models, utilize large and novel data, and repeat this process. The PERFICT approach outlines a way to replicate the process of the IPCC, but with vastly fewer resources. It encourages a nimble approach to applied decision making (Box 2) that allows for both changeable process complexity (e.g., a simple fire model or a complex fire model) and management complexity (e.g., manage fire risk in isolation, or within the context of forest management, species-at-risk, climate change and pest management). Scientifically, the easier testing of alternative models and hypotheses using the PERFICT approach offers an objective ground to resolve contradictions from models. From a management perspective, competing land management goals such as carbon sequestration and species at risk conservation can be evaluated, crossing traditional scientific disciplines and synergies can be identified. With new and more data, the predictions from potential models are checked against data, reported clearly and rapidly, and repeated regularly. This translates directly to policy spheres that have a regular reporting requirement (e.g., Kurz *et al.* 2009). It also brings decision making into a continuous improvement process, allows for the creation of nimble decisions support systems and builds confidence in science-informed decision-making. Finally, PERFICT improves science-policy integration by increasing model interoperability horizontally (e.g. integrating across disciplines) and vertically (Fig. 3). While the literature has a track record of this sort of science-policy integration happening (e.g., Schmolke *et al.* 2010), the PERFICT approach will allow this to become ubiquitous and speed the transfer of vertical information. This expands the reach of ecological models beyond ecologists and promotes co-production by enabling the direct participation and feedback of non-experts, like policy and decision makers.

Conclusion

The future of applied ecology requires solutions that cross disciplines and transcend scientific, statistical, computational, and human cultural paradigms. Historically, there have been successes for “wicked” problems (Parrott 2017), but they are too few and too infrequent (Travers et al. 2019). Ecologists must embrace the current data revolution and the unprecedented computational power to be at the table of every decision and policy that affects ecosystems worldwide. Often these situations demand rapid answers that cannot wait for a grant cycle and multi-year projects to complete: we have to be nimble and ready to give data-driven solutions to new problems or meta-problems. The PERFICT approach and the toolkit we introduce here provide a path forward. This approach can bring the language of ecologists to the language of policy makers and land managers through dynamic (e.g., automatically generated) reporting or web interfaces and through rapid answers to complex trade-offs in ecosystem management. It provides a readily achievable and objective solution to the inherent subjectivity of Occam’s razor and George Box’s quote about model utility. The tools identified here focus around R and the SpaDES ecosystem of packages and modules. These tools are likely only the first iteration of such a foundation; many meta-tools are emerging on top of the SpaDES standard. Training the next generation of ecologists to think in the PERFICT approach will be a challenge; but tools are now available with minimal training, especially for those who know R already. This reimagining of Predictive Ecology from the empirical version presented by Robert H. Peters in the late 1970s (Peters 1977) echoes his utilitarian and objectivity goals, but expands into the data and computational revolutions of the 21st century. Being at the management, policy making, and political tables with timely answers to challenging ecological questions is the ultimate goal.

Acknowledgements

Funding was provided by fRI Research (Healthy Landscapes Program and Mountain Pine Beetle Ecology Program), Environment and Climate Change Canada (through the Canadian Wildlife Service), Natural Resources Canada (through Assistant Deputy Minister's Innovation Fund), and Natural Sciences and Engineering Research Council of Canada (NSERC Strategic Funds [to Cumming] and Discovery Grant [to McIntire]). Mitacs Accelerate Award no. IT10289 and Mitacs Elevate Fellowship no. IT12391, in partnership with the Foothills Research Institute (fRI), both awarded to C Barros.

References Cited

Aho, K., Derryberry, D. & Peterson, T. (2014). Model selection for ecologists: the worldviews of AIC and BIC. *Ecology*, 95, 631–636.

Akçakaya, H.R. & Root, W. (2005). RAMAS GIS: linking spatial data with population viability analysis. *Appl. Biomath. Setauket N. Y.*

Anderson, B., Chamberlain, S., Krystalli, A., Mullen, L., Ram, K., Ross, N., *et al.* (2019). *rOpenSci Packages: Development, Maintenance, and Peer Review*.

Anderson, D.R., Burnham, K.P. & Thompson, W.L. (2000). Null hypothesis testing: Problems, prevalence, and an alternative. *J. Wildl. Manag.*, 64, 912–923.

Banks, J., Carson, J., Nelson, B.L. & Nicol, D. (2005). *Discrete-Event System Simulation*. 4th ed. Prentice Hall.

Begley, C.G. & Ellis, L.M. (2012). Raise standards for preclinical cancer research. *Nature*, 483, 531–533.

Berryman, A.A. (1992). On Choosing Models for Describing and Analyzing Ecological Time Series. *Ecology*, 73, 694–698.

Bodner, K., Fortin, M. & Molnár, P.K. (2020). Making predictive modelling ART: accurate, reliable, and transparent. *Ecosphere*, 11.

Bond-Lamberty, B., Fisk, J.P., Holm, J., Bailey, V.L., Bohrer, G. & Gough, C. (2015). Moderate forest disturbance as a stringent test for gap and big-leaf models. *Biogeosciences Online*, 12.

Borregaard, M.K. & Hart, E.M. (2016). Towards a more reproducible ecology. *Ecography*, 39, 349–353.

Boulangeat, I., Georges, D. & Thuiller, W. (2014). FATE-HD: A spatially and temporally explicit integrated model for predicting vegetation structure and diversity at regional scale. *Glob. Change Biol.*, 20, 2368–2378.

Boutin, S. & Hebert, D. (2002). Landscape ecology and forest management: developing an effective partnership. *Ecol. Appl.*, 12, 390–397.

Chubaty, A.M. & McIntire, E.J.B. (2019). *SpaDES: Develop and Run Spatially Explicit Discrete Event Simulation Models. An R meta package*.

Clark, J.S. (2001). Ecological Forecasts: An Emerging Imperative. *Science*, 293, 657–660.

Dietze, M.C. (2017). *Ecological forecasting*. Princeton University Press, Princeton, NJ.

Dietze, M.C., Fox, A., Beck-Johnson, L.M., Betancourt, J.L., Hooten, M.B., Jarnevich, C.S., *et al.* (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proc. Natl. Acad. Sci.*, 201710231.

Environment and Climate Change Canada. (2018). *Pan-canadian approach to transforming species at risk conservation in Canada* (No. ISBN: 978-0-660-27223-8). Gatineau, QC, Canada.

Evans, M.R. (2012). Modelling ecological systems in a changing world. *Philos. Trans. R. Soc. B Biol. Sci.*, 367, 181–190.

Evans, M.R., Norris, K.J. & Benton, T.G. (2012). Predictive ecology: systems approaches. *Philos. Trans. R. Soc. B Biol. Sci.*, 367, 163–169.

Fahey, T.J., Woodbury, P.B., Battles, J.J., Goodale, C.L., Hamburg, S.P., Ollinger, S.V., *et al.* (2010). Forest carbon storage: ecology, management, and policy. *Front. Ecol. Environ.*, 8, 245–252.

Fall, A. & Fall, J. (2001). A domain-specific language for models of landscape dynamics. *Ecol. Model.*, 141, 1–18.

Grimm, V. & Railsback, S.F. (2012). Pattern-oriented modelling: a “multi-scope” for predictive systems ecology. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, 367, 298–310.

Horne, J.S. & Garton, E.O. (2006). Likelihood cross-validation versus least squares cross-validation for choosing the smoothing parameter in kernel home-range analysis. *J. Wildl. Manag.*, 70, 641–648.

Houlahan, J.E., McKinney, S.T. & Rochette, R. (2015). On theory in ecology: Another perspective. *BioScience*, 65, 341–342.

Kimmins, J.P., Mailly, D. & Seely, B. (1999). Modelling forest ecosystem net primary production: the hybrid simulation approach used in forecast. *Ecol. Model.*, 122, 195–224.

Klein, R.A., Ratliff, K.A., Vianello, M., Adams, R.B., Bahník, Š., Bernstein, M.J., *et al.* (2014). Investigating Variation in Replicability: A “Many Labs” Replication Project. *Soc. Psychol.*, 45, 142–152.

Kurz, W.A., Dymond, C.C., White, T.M., Stinson, G., Shaw, C.H., Rampley, G.J., *et al.* (2009). CBM-CFS3: a model of carbon-dynamics in forestry and land-use change implementing IPCC standards. *Ecol. Model.*, 220, 480–504.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., *et al.* (2007). Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecol. Lett.*, 10, 995–1015.

Micheletti, T., Stewart, F.E.C., Cumming, S., Haché, S., Stralberg, D., Tremblay, J.A., *et al.* (in review). Teasing apart the pathways of climate change effects on boreal landbird distributions in Northwestern Canada using SpaDES. *Front. Ecol. Evol.*

Mouquet, N., Lagadeuc, Y., Devictor, V., Doyen, L., Duputié, A., Eveillard, D., *et al.* (2015). Predictive Ecology In A Changing World. *J. Appl. Ecol.*

Munafò, M.R., Nosek, B.A., Bishop, D.V.M., Button, K.S., Chambers, C.D., Percie du Sert, N., *et al.* (2017). A manifesto for reproducible science. *Nat. Hum. Behav.*, 1.

Nature Editorial. (2014). Journals unite for reproducibility. *Nat. News*, 515, 7.

Northrup, J.M. & Gerber, B.D. (2018). A comment on priors for Bayesian occupancy models. *PLOS ONE*, 13, e0192819.

O'Neill, J. (2002). *Ecology, policy and politics: Human well-being and the natural world*. Routledge.

Oreskes, N., Shrader-Frechette, K. & Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263, 641–646.

Paradis, G., LeBel, L., D'Amours, S. & Bouchard, M. (2013). On the risk of systematic drift under incoherent hierarchical forest management planning. *Can. J. For. Res.*, 43, 480–492.

Parrott, L. (2017). The modelling spiral for solving 'wicked' environmental problems: guidance for stakeholder involvement and collaborative model development. *Methods Ecol. Evol.*, 8, 1005–1011.

Peters, R.H. (1977). The unpredictable problems of tropho-dynamics. *Environ. Biol. Fishes*, 2, 97–101.

Peters, R.H. (1982). Useful Concepts for Predictive Ecology. In: *Conceptual Issues in Ecology* (ed. Saarinen, E.). Springer Netherlands, Dordrecht, pp. 215–227.

Peters, R.H. (1986). The role of prediction in limnology. *Limnol. Oceanogr.*, 31, 1143–1159.

Peters, R.H. (1991). *A critique for ecology*. Cambridge University Press, Cambridge.

Power, M. (1993). The predictive validation of ecological and environmental models. *Ecol. Model.*, Theoretical Modelling Aspects, 68, 33–50.

Radchuk, V., Kramer-Schadt, S. & Grimm, V. (2019). Transferability of mechanistic ecological models is about emergence. *Trends Ecol. Evol.*, 34, 487–488.

Reichman, O.J., Jones, M.B. & Schildhauer, M.P. (2011). Challenges and Opportunities of Open Data in Ecology. *Science*, 331, 703–705.

Rijs, C. & Fenter, F. (2020). The academic response to COVID-19. *Front. Public Health*, 8.

Rosenblum, D.S. (1995). A practical approach to programming with assertions. *IEEE Trans. Softw. Eng.*, 21, 19–31.

Scheller, R.M., Domingo, J.B., Sturtevant, B.R., Williams, J.S., Rudy, A., Gustafson, E.J., *et al.* (2007). Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecol. Model.*, 201, 409–419.

Scheller, R.M. & Miranda, B. (2015). *LANDIS-II Biomass Succession v3.2 Extension User Guide* (User Guide).

Schmolke, A., Thorbek, P., DeAngelis, D.L. & Grimm, V. (2010). Ecological models supporting environmental decision making: A strategy for the future. *Trends Ecol. Evol.*, 25, 479–486.

Schumaker, N.H. & Brookes, A. (2018). HexSim: a modeling environment for ecology and conservation. *Landsc. Ecol.*, 33, 197–211.

Stall, S., Yarmey, L., Cutcher-Gershenfeld, J., Hanson, B., Lehnert, K., Nosek, B., *et al.* (2019). Make scientific data FAIR. *Nature*, 570, 27–29.

Thiele, J.C. & Grimm, V. (2015). Replicating and breaking models: good for you and good for ecology. *Oikos*, 124, 691–696.

Travers, H., Selinske, M., Nuno, A., Serban, A., Mancini, F., Barychka, T., *et al.* (2019). A manifesto for predictive conservation. *Biol. Conserv.*, 237, 12–18.

Vines, T.H., Albert, A.Y.K., Andrew, R.L., Débarre, F., Bock, D.G., Franklin, M.T., *et al.* (2014). The Availability of Research Data Declines Rapidly with Article Age. *Curr. Biol.*, 24, 94–97.

Wenger, S.J. & Olden, J.D. (2012). Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods Ecol. Evol.*, 3, 260–267.

White, E.P., Yenni, G.M., Taylor, S.D., Christensen, E.M., Bledsoe, E.K., Simonis, J.L., *et al.* (2019). Developing an automated iterative near-term forecasting system for an ecological study. *Methods Ecol. Evol.*, 10, 332–344.

Wilensky, U. (1999). NetLogo.

Wilson, G., Aruliah, D.A., Brown, C.T., Chue Hong, N.P., Davis, M., Guy, R.T., *et al.* (2014). Best practices for scientific computing. *PLoS Biol.*, 12, e1001745.

Wilson, G., Bryan, J., Cranston, K., Kitzes, J., Nederbragt, L. & Teal, T.K. (2017). Good enough practices in scientific computing. *PLOS Comput. Biol.*, 13, e1005510.

Yates, K.L., Bouchet, P.J., Caley, M.J., Mengersen, K., Randin, C.F., Parnell, S., *et al.* (2018). Outstanding Challenges in the Transferability of Ecological Models. *Trends Ecol. Evol.*, xx, 1–13.

Benefit	Example	PERFICT approach enables the benefit by:
Speeding up scientific advances	Occam’s razor	Allowing for an objective evaluation of how much complexity is right for a given project as models of arbitrary complexity can be readily compared
	Informative priors	Easing the process of moving from a previous study’s Bayesian posteriors to a new study’s priors, lessening the problems with specifying uninformative priors (Northrup & Gerber 2018)
	Forecast horizon	Repeatedly iterating a forecasting model with regularly updated data and model, expanding the forecast horizon
	Community of contributors	Allowing manageable projects with 100s of contributors to quickly update our understanding of a system
	Predictive validation	Facilitating use of predictive validation. Using truly out-of-sample data to test models becomes easier with reusable, interoperable modules
	Rewriting models	Encouraging rewriting in a widely known language (R/SpaDES) allowing many experts to see and understand code.
	Many eyes	Establishing a modeling standard that is understandable by many scientists with sufficient capacity. Bug fixes and improvements are identified and implemented very quickly.
Data Science advances	Building on data science tools	Facilitating the use of sophisticated cloud repositories, user access control, data caching, cloud services etc. for researchers who do not have the capacity or time to learn and develop them.
	Data quality and quantity	Building a complete data-model-validation pipeline from reusable components allowing for assessment of different data sources
	Linking models to data	Keeping the linkage between canonical, original data sources and models live at all times. This allows for rapid reparameterization and updating with continual testing
	Cross disciplinarity	Lessening the technological, data and cultural barriers that make cross disciplinary work challenging

Improving science-policy integration	Regular reporting	Lowering the effort required to produce regular updates for policy reporting
	IPCC-like process	Allowing lower budget projects to achieve IPCC-like integration with its benefits such as regular updating, ensemble modeling, and direct policy making
	Different users	Creating a complete framework that allows for all types of expertise -- from land managers, rights holders and the public, to scientists and computer programmers -- to interact (Fig. 3)
	Web and decision support applications	Allowing for the development of generic web and decision support tools that can be reused widely
	Coping with contradictions	Opening the science informed decision-making and policy-making process to shed light on cases where models contradict one another and offering an objective way to resolve those contradictions

Table 1: Benefits and examples of the PERFICT approach and how these benefits can be realized. See text for details.

Figure Legends

Fig. 1. Example abstraction of a PERFICT approach to a study. In this view, the study can be nimble as many components are reused, the entire workflow is continuous, changing data availability is accommodated, alternative modules for the same process are explicit, each sub-project has high modularity (few arrows cross sub-project boundaries), and there are very few data sources that are external to each module demonstrating that the cross-disciplinary connections are minimal. For a given sub-project, a ubiquitous workflow is to have 2 generic modules (one for parameter estimation -- hexagon -- and one for forecasting or predicting -- square), with zero or more idiosyncratic modules. In cases of maximum reusability, ρ is the only idiosyncratic dataset that must be supplied. In our experience, each project begins with many idiosyncratic datasets and non-reusable modules to deal with those idiosyncratic datasets, in part because we do not yet know what is reusable. But as we identify the components that are reusable, over time and use, more and more elements move from idiosyncratic modules to generic modules (e.g., elements in G are moved to H). Similarly, as all the data in a project become freely available, the idiosyncratic modules may be dropped, simplifying the project, and maximizing reusability. We include an alternative collection ζ that represents the same ecological process as γ; the two together can inform "consensus" forecasts, be treated as alternative hypotheses, help to estimate model uncertainty etc. Modules J and K can be built to provide feedbacks into any arbitrary modules. Where there is a need to use heuristic optimization (e.g., pattern oriented modeling: Grimm & Railsback 2012), a single objective function can be developed to update arbitrary parameters (not shown). A common, traditional Ecology study would include h and G, i.e., closed data and low reusability models. Squares (forecasting, prediction) and hexagons (statistical or parameter estimation) represent modules; circles represent data. Arrows represent data-module connections with freely (solid lines), and not freely (dashed) available data. Greek letters indicate sub-projects which are collections of

modules that create a coherent data-module workflow for a single idea (e.g., “fire forecasting”), which could be from a scientific publication. Alphabet letters are arbitrary labels for data (lowercase) or modules (uppercase) showing that data input expectations will generally (though not necessarily) be unique for a given module. Within the data types, green is proprietary or truly idiosyncratic data, yellow is freely available and open data, beige is data inputs and (possibly) outputs of modules, orange is outputs that are not inputs (e.g., for visualization, reporting etc.), and purple is shared data. We show data in different colors to emphasize their different roles; within a project, they are simply arbitrary data objects. Within the module types, the darkness of the coloration indicates how generic it is, therefore how reusable it is in different contexts. We differentiate estimation modules from forecasting modules by shape to identify their different roles; the modules have no structural differences. ρ is the study area for a project.

Fig. 2. Functions are the foundation of many models. Functions are deeply modular, and if collections of functions are wrapped into packages, there are many tools that enable easy dissemination, quality control, continuous integration, documentation, and writing. Functions may have default values for arguments, but they are not intended to do something without the user understanding the function and providing input arguments. We define modules to be similar to functions because they have inputs and convert those inputs into some output. However, modules are collections of one or more functions that have computer readable metadata describing their inputs and outputs. Unlike functions, this module metadata contains the information that describes how modules fit (or do not) together. Modules, as we present them here, are the basic unit of code that enables and facilitates all the elements of the PERFICT approach. In analogy, functions are Lego® pieces, often supplied in a package (similar to collections of functions), and modules are Lego® structures made with those pieces, such as trucks, houses, roads, space shuttles etc. A given structure has inherent value, i.e., a truck can be the end goal of a project and can stand alone. The metadata (implicit in Lego®) describe the

ways these structures interact, i.e., a road can take things with wheels (input); a bus has wheels (output) so can go on a road, but a house does not so cannot. Using a structure by itself or combining multiple structures together makes simple to complex “models”, such as neighbourhoods, villages, cities, space stations, etc. Many modules fit together (a truck and a road); others do not (a truck and a space station). The structures can be used in many new ways, pieces added to structures, and collected into complex meta-structures. If we want to build a Lego® city, we could either start with raw blocks or build a new configuration by reusing some or all pre-existing structures. Furthermore, other toy “brands” -- or computer languages, e.g., Python, C++ -- can be added to the city. Using the PERFICT approach, ecologists can build robust, reusable modules, enabling rapid creation, use, testing and reformulating of models.

Fig. 3. The different users, and their contributions to applied decision making (arrow-heads show increasing importance), who interact with ecological modeling and their forecasts. The PERFICT approach is a modular framework that allows many entry points into a science-policy system. The SpaDES toolkit enhances existing open data and tools, facilitates the implementation of this approach, and improves the ability to engage a wide variety of users.

Box 1 Implementing a PERFICT approach: best practices

To work with the PERFICT approach, best practices begin at project initiation, continue during project development, and end always in a fully reproducible and reusable state that can become part of a continuous adaptive management process. In our experience, the extra effort involved in learning and implementing are relatively modest; the payoffs are large. Below, capital letters in parentheses indicate which elements of the PERFICT approach are addressed. When the SpaDES collection of R packages has explicit tools for this, we indicate. See text for details.

1. Use a modular approach to code development (R, I; SpaDES);
2. Follow coding best practices (R) (e.g., Wilson *et al.* 2014, 2017);
3. Use a standard interface to code (SpaDES);
4. Use data directly from a cloud repository (R, F, I; SpaDES).
5. Use open solutions and approaches unless it is functionally impossible (R, F, SpaDES);
6. Provide worked examples (R, SpaDES);
7. Maintain code to be always functioning (F, I, C; SpaDES);
8. Write, maintain and run tests (R, T; SpaDES);
9. Document code (I, SpaDES);
10. Use version control (R)

BOX 2 -- PERFICT in action

Pushing against parsimony, it is now clear that we often need heterogeneous models that integrate across disciplines so that realistic land management challenges can be addressed (Houlahan *et al.* 2015). For example, land managers in Northwest Territories, Canada, are attempting to manage declining woodland caribou populations and listed bird species-at-risk in the context of protected areas planning and indigenous peoples' rights (Micheletti *et al.* in review in review; S. Haché pers. comm; T. Micheletti pers. comm). This formed the basis for a pilot project for the PERFICT approach. To forecast these values, there were many ecological and land management issues that had to be addressed. For example, wildfires had to be forecasted under changing climate, changing fuels (vegetation), and changing fire suppression practices. Vegetation is shifting due to direct and indirect effects of climate, such as species and biome shifts, permafrost melt, tree species drought-induced mortality, and accelerated forest succession dynamics, to name a few. These landscapes are also currently and historically inhabited and used by Indigenous peoples (https://native-land.ca/maps/territories/sahtu-dene-and-metis/). There are important road networks for mining and other anthropogenic development. There are enormous carbon stores in the frozen peatlands that are melting, releasing these to the atmosphere. The overarching question was how to best manage the Species-at-Risk, alongside all these other values, given changing climate and indigenous peoples rights.

To correctly manage these landscapes, all these elements must be included in forecasting, so decisions can evaluate consequences, synergies and trade-offs across multiple disciplines. It is likely inappropriate to give any of these issues short shrift and have each project treat the issues that are not well studied by the team as "externalities"; yet, building large collaborative projects with models that do not interoperate is extremely onerous, time consuming, and ultimately very costly. We need to focus on management problems, while including *all the best disciplinary models*. This challenge led to the creation of a new class of ecological scientist: the *integrator*. Like a traditional generalist, this person acts within the big picture perspective, yet has just enough knowledge of the modules (and potentially the community of scientists who developed them) to be able to work at the interface between modules.

For the pilot, we assembled a 24-member collaborative team and brought together 19 modules with 7 lead module developers to assist with this problem (see Supp. Mat. E to see module interdependency plot). The team included partners from two levels of government, and scientists from three universities, and the initial pilot was pulled together in four months. With the PERFICT approach, the technical parts of linking the models were a minor component of the whole project. The challenges we faced were not from integration, as we were using SpaDES-compatible modules, but from the immature science that some of the 19 modules addressed.

In the expansion of the pilot -- the “Western Boreal Initiative” -- we are working with over 40 modules, with 10 different sub-projects including endangered species conversation, Indigenous land management, the Pan-Canadian Approach to conservation (Environment and Climate Change Canada 2018), caribou management and carbon management. Some of these projects are addressing whole-system management questions, others are very specific. Model components were either new or were existing; many required improved algorithms or had access to improved datasets as compared to the pilot. In all cases, each benefited from working within the PERFICT approach, allowing for nimble updates at any point, swapping out of previous models, and weaving in new elements including the long term process of Indigenous rights on the land. These are co-produced, works in progress.

Fig. 1

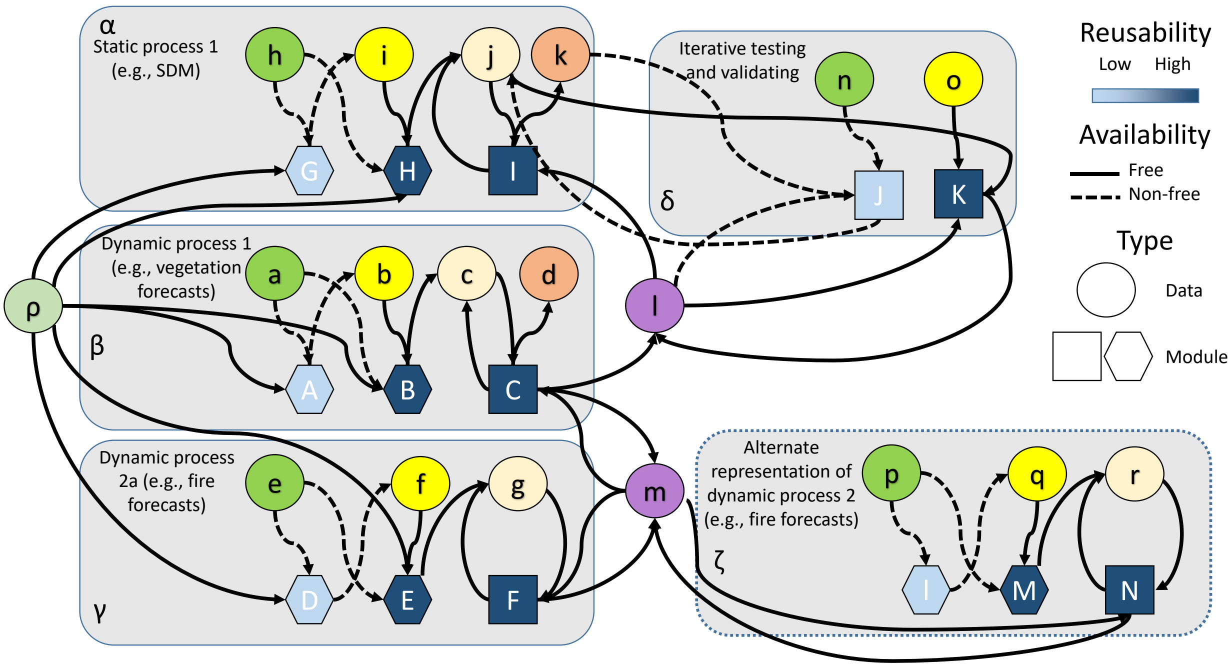
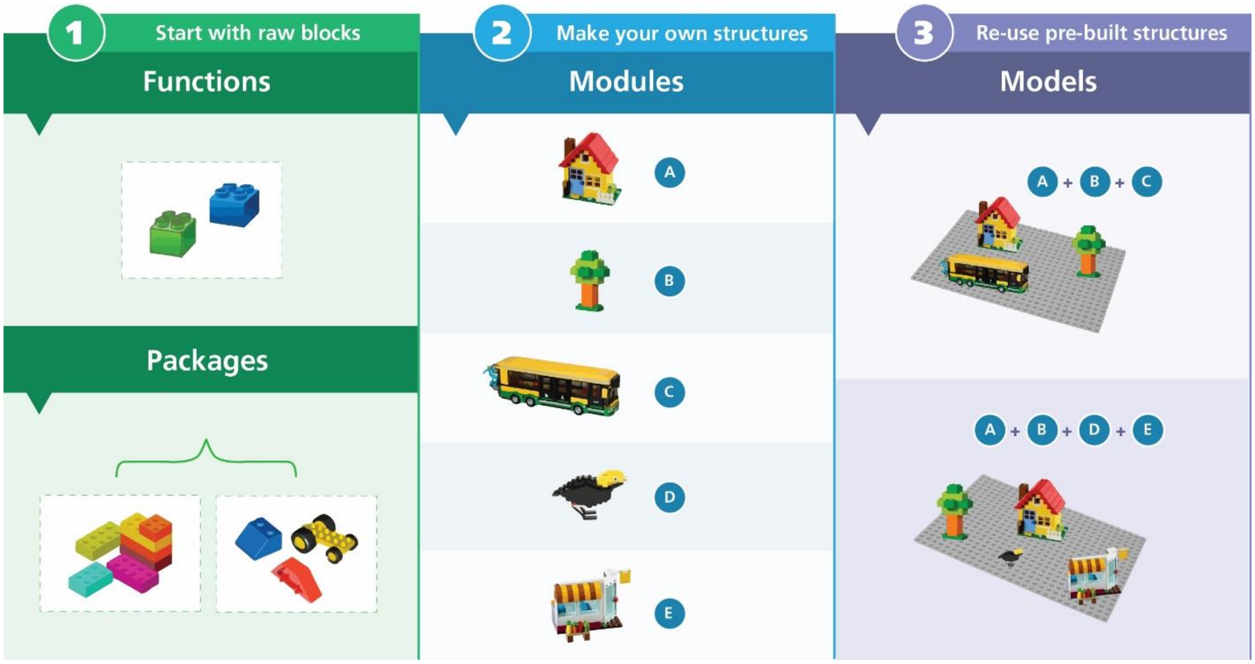
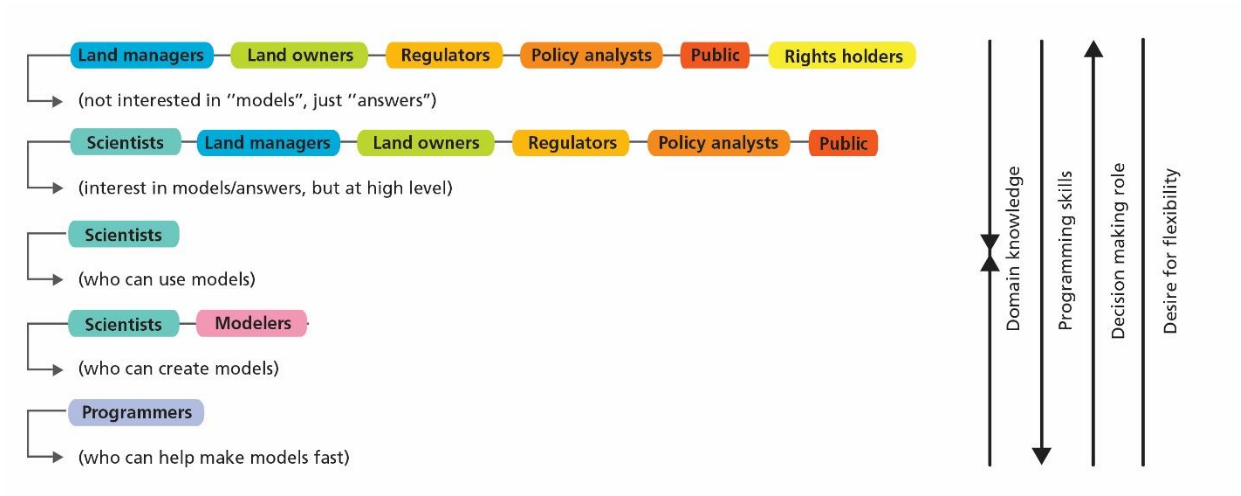

α
Static process 1
(e.g., SDM)
h
i
j
k
G
H
I
Iterative testing
and validating
n
o
δ
J
K
Reusability
Low
High
Availability
Free
Non-free
Type
Data
Module
Dynamic process 1
(e.g., vegetation
forecasts)
a
b
c
d
ρ
β
A
B
C
l
Dynamic process
2a (e.g., fire
forecasts)
e
f
g
m
γ
D
E
F
Alternate
representation of
dynamic process 2
(e.g., fire forecasts)
p
q
r
ζ
L
M
N

Fig. 2


1
Start with raw blocks
Functions
Packages
2
Make your own structures
Modules
A
B
C
D
E
3
Re-use pre-built structures
Models
A + B + C
A + B + D + E

734 Fig. 3



Land managers
Land owners
Regulators
Policy analysts
Public
Rights holders
(not interested in "models", just "answers")
Scientists
Land managers
Land owners
Regulators
Policy analysts
Public
(interest in models/answers, but at high level)
Scientists
(who can use models)
Scientists
Modelers
(who can create models)
Programmers
(who can help make models fast)
Domain knowledge
Programming skills
Decision making role
Desire for flexibility

735