Improving the communication and accessibility of stock assessment using interactive visualization tools

Journal:	*ICES Journal of Marine Science*
Manuscript ID	Draft
Manuscript Types:	Quo Vadimus
Date Submitted by the Author:	n/a
Complete List of Authors:	Regular, Paul; Fisheries and Oceans Canada Newfoundland and Labrador Region, Robertson, Gregory; Fisheries and Oceans Canada Newfoundland and Labrador Region Rogers, Robert; Fisheries and Oceans Canada Newfoundland and Labrador Region Lewis, Keith; Fisheries and Oceans Canada Newfoundland and Labrador Region
Keyword:	dashboard, data exploration, dynamic documents, integrated analysis, model validation, open stock assessment

Improving the communication and accessibility of stock assessment using interactive visualization tools

Paul M. Regular, Gregory J. Robertson, Robert Rogers, Keith P. Lewis

Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, 80 East White Hills, St. John's, Newfoundland and Labrador, A1C 5X1, Canada

Running title: Interactive stock assessment

Conflict of Interest Statement: The authors have no conflicts of interest to declare.

Data Availability Statement: Data sharing is not applicable to this article as no data were generated and no analysis to draw scientific inference was conducted in this study.

Corresponding author: Paul M. Regular, Fisheries and Oceans Canada, Northwest Atlantic Fisheries Center, 80 East White Hills, St. John's, Newfoundland and Labrador, A1C 5X1, Canada, E-mail: Paul.Regular@dfo-mpo.gc.ca, Phone: 1-709-772-2067

Abstract: Scientists across many fields are faced with the challenge of synthesizing and communicating information from large and complex data sets. The field of stock assessment is no exception as the volume and variety of the data has grown alongside the computational methods used to integrate them. While this growth in data and model complexity has improved the assessments of many stocks, the process of communicating the results to colleagues, stakeholders and fisheries managers in a meaningful way has become more daunting. The traditional approach of presenting information across a series of static slides often fails to convey the richness of information available and, as such, important patterns and details are easily overlooked. Here we contend that this problem can be mediated through the effective use of new open source tools for building interactive visualizations. These tools allow a broader audience to conduct detailed explorations of the results, leading to a deeper and collective understanding of both the data and models used to inform stock assessments. As a consequence, the peer review process is more open and accessible and the resulting science advice is improved and widely supported.

Keywords: dashboard, data exploration, dynamic documents, integrated analysis, model validation, open stock assessment

The greatest value of a picture is when it forces us to notice what we never expected to see.

- Tukey (1977)

Quantitative stock assessment plays a central role in modern fisheries management (Hilborn and Walters, 1992). Over time, and as new methods are developed, there is an increasing amount of data available to inform stock assessments. These data might be richer information on stock structure using an increasing array of markers and biomarkers, and/or the continued lengthening of population status and catch time series. Concurrently, analytical methods are now able to integrate many data sources into one stock assessment model (Maunder and Punt, 2013). For well-monitored stocks, the challenge has shifted from having sufficient data and information for providing sound advice on stock status, to presenting large quantities of data and output from increasingly complex statistical models in a meaningful way. Traditional formats (e.g. slides) for presenting information at stakeholder meetings are often insufficient to convey the richness of information available and the static and sequential nature of these formats can stifle meaningful discussions. A solution to this problem is the effective use of interactive visualization tools (Keena *et al.*, 2016). These are common tools we use every day on a range of web sites, but their use is no longer restricted to web site developers as these tools are being integrated into software commonly used by the research community (Perkel, 2018). As a result, an increasing number of scientists across a wide range of disciplines are starting to apply interactive visualization tools to explore and communicate their results (e.g. Jones *et al.*, 2016; Letcher *et al.*, 2018; Yeatman *et al.*, 2018). This trend includes fisheries research and many of these tools, and the code used to produce them, are available online (see Appendix A for examples).

Here we aim to demonstrate how interactive visualization tools provide an efficient and effective means of exploring and communicating the ever-expanding array of data inputs and model outputs. We especially highlight the use of interactive visuals inside dynamic documents called "dashboards". The concept of using a dashboard was borrowed from the business community where they are frequently used to group a series of interactive visuals and tables to provide at-a-glance views of key performance indicators. We first demonstrate the application of these tools to data that are commonly presented in stock assessments to show how they can simplify the detailed exploration of data from long-term monitoring programs. Second, we focus on the

58 modeling aspect of stock assessment and demonstrate how interactive dashboards can be used to
59 explore, diagnose and communicate results from an integrated assessment model. This structure
60 corresponds to two important steps in the stock assessment process: 1) data exploration, and 2)
61 data modeling. In each case, the dashboards were constructed within the framework of the R
62 programming language (R Core Team, 2017) and the RStudio IDE (RStudio Team, 2015) using
63 three or more of the packages listed in Table (Table 1).

64 Table 1: Names and descriptions of the key packages used to build the interactive applications
65 presented in this paper. Descriptions are directly from the `DESCRIPTION` file of each package.

`shiny`	Makes it incredibly easy to build interactive web applications with R. Automatic "reactive" binding between inputs and outputs and extensive prebuilt widgets make it possible to build beautiful, responsive, and powerful applications with minimal effort (Chang *et al.*, 2018)
`shinydashboard`	Create dashboards with 'Shiny'. This package provides a theme on top of 'Shiny', making it easy to create attractive dashboards (Chang and Borges Ribeiro, 2018).
`rmarkdown`	Convert R Markdown documents into a variety of formats (Allaire *et al.*, 2018).
`flexdashboard`	Format for converting an R Markdown document to a grid oriented dashboard. The dashboard flexibly adapts the size of its components to the containing web page (Allaire, 2017).
`plotly`	Easily translate 'ggplot2' graphs to an interactive web-based version and/or create custom web-based visualizations directly from R (Sievert, 2018).
`leaflet`	Create and customize interactive maps using the 'Leaflet' JavaScript library and the 'htmlwidgets' package. These maps can be used directly from the R console, from 'RStudio', in Shiny applications and R (Cheng *et al.*, 2018).
`crosstalk`	Provides building blocks for allowing HTML widgets to communicate with each other, with Shiny or without (i.e. static .html files) (Cheng, 2016).

66 Data exploration

67 The data sets used in stock assessments are constantly growing due to the continuation of long-
68 term monitoring efforts, the addition of new monitoring programs, or both. As such, stock

assessment biologists need to manage large volumes of data from a variety of sources. Time series of reported landings and catch-at-age are often analyzed in conjunction with data from "fishery-independent" surveys that track changes in abundance and, in many cases, also monitor trends in biological factors such as age composition, growth rates, sex ratios and maturation stages. For some data-rich stocks, mark and recapture studies are also carried out to estimate movement, migration, growth rate, natural mortality, and discard mortality. All of the above-mentioned data sets are complex and as the volume and variety of these data increases, it becomes more difficult to be aware of the details of each data set and discover key patterns within each. It has long been recognized that this challenge can be mediated, to a degree, by the application of interactive visualization tools as they allow detailed exploration of the data behind a plot (Tukey and Tukey, 1985; Fisherkeller *et al.*, 1988). For instance, the ability to zoom in on features or areas of interest, turn off layers and hover over specific points to reveal more information creates an interactive user-driven experience that expedites explorations of data. These abilities are illustrated by two recently developed interactive tools: 1) a tool designed to quickly examine fishery-independent survey data; and 2) an interactive mapping tool developed for the exploration of a long-term tagging study.

Survey data

Fisheries and Oceans Canada (DFO) has been conducting a multi-species stratified-random survey across the Newfoundland and Labrador shelf since the 1970s (Rideout and Ings, 2018). These data are analyzed using a standard stratified analysis via a locally developed R package called RStrap (for details on methodology see Smith and Somerton, 1981). Both the inputs and outputs from RStrap analyses are very large and, depending on the species, the time-series may include > 40 years and span the majority of the Newfoundland and Labrador shelf. In order for people to quickly and reliably explore these data without iteratively modifying R scripts one species at a time, an application (hereafter called RStrap Explorer) was built using a combination of shiny, crosstalk, flexdashboard, plotly and R markdown (Table 1). RStrap Explorer started as a way to visualize estimates of biomass and abundance trends of specific species by supplying RStrap output to a flexdashboard file designed to organize and visualize the results. The shiny package was latter applied as it allows RStrap to be run in the background and therefore allows the user to dynamically explore abundance and biomass estimates from multiple species in one session. Both crosstalk and plotly were incorporated to allow the user to interact with the data.

RStrap Explorer contains four primary tabs: 1) "Survey Indices" contains stock level estimates of biomass and abundance (Figure 1), 2) "Age & Length Distributions" contains length and age frequency plots (Figure 2), 3) "Recruitment" displays recruitment indices, and 4) "Help" provides additional context to the survey data and analysis. The input interface for RStrap Explorer is divided into two parts: basic and advanced inputs. Basic inputs are those like species, season, and survey year(s). Advanced inputs allow the user to define the length or age of recruitment, select whether they would like analyses conducted by length and/or age, and so on. This application was built for data exploration and allow the user to obtain quick and easy visualization of survey data, before the thorough the data checking and analysis required for formal stock assessment begins.

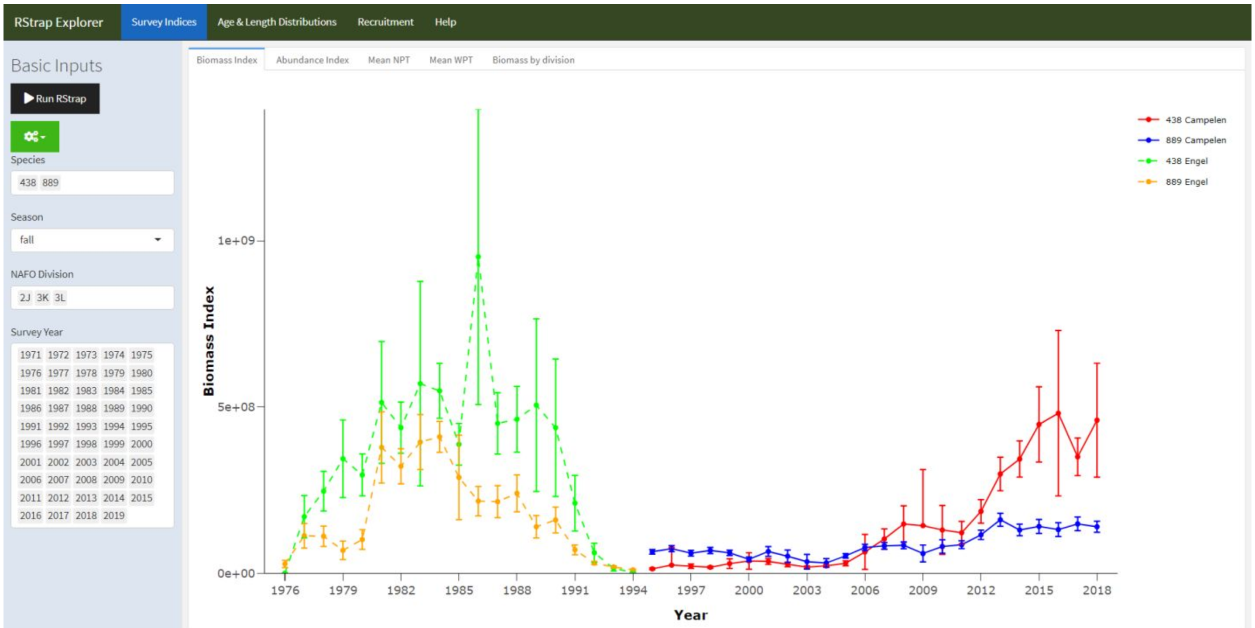

Figure 1: - Screen shot of the “Survey Indices” page of the RStrap Explorer application. The “Basic Inputs” sidebar is fixed across all pages and allows the user to specify the species, NAFO Divisions, season, and survey years of interest. The “Biomass Index” sub-tab is illustrated here. This sub-tab shows a time-series of biomass data with 95% confidence intervals, with a break in the lines when the survey gear changed from an Engle to Campelen trawl, for both cod (*Gadus morhua*; species code 438) and American Plaice (*Hippoglossoides platessoides*; species code 889). All trend lines can be toggled off using plotly’s dynamic user interface. The user also has the ability to hover over the points to compare estimates and export the figure as a png file.

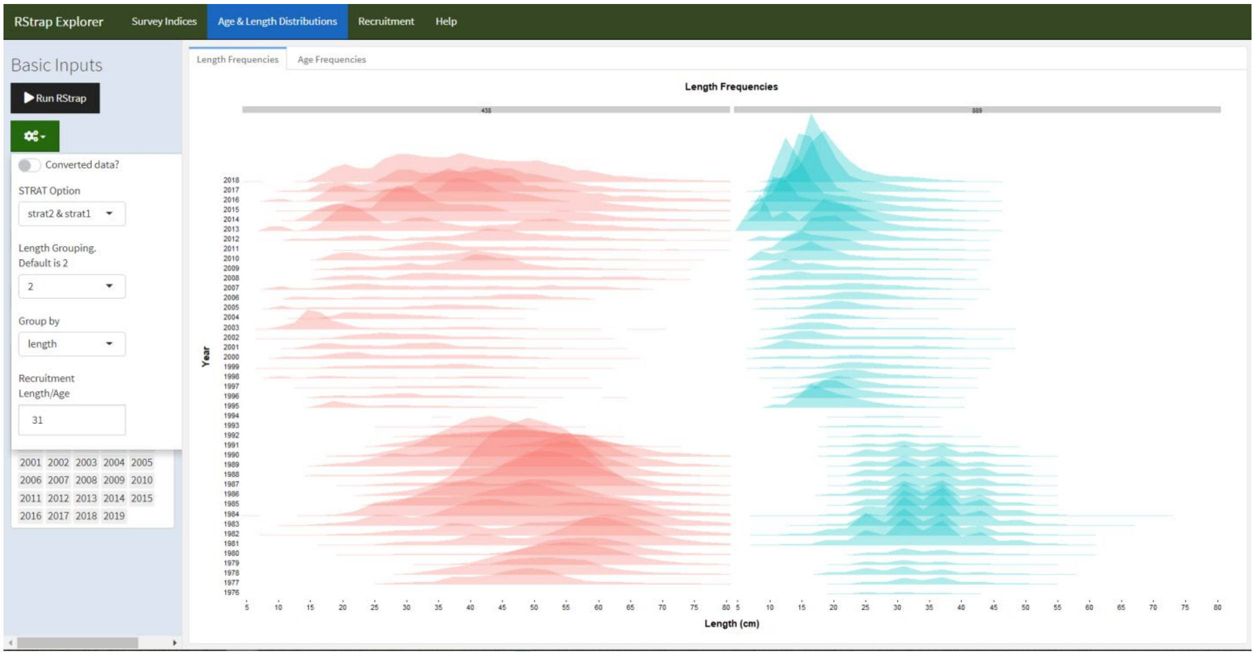

Figure 2: - "Age and Length Frequencies" tab of the RStrap Explorer tool. This tab illustrates the length frequency distributions of both cod (*Gadus morhua*; species code 438) and American plaice (*Hippoglossoides platessoides*; species code 889) binned into 2 cm length bins. The binning is easily modified using the "Length Group" field in the "Advanced Options" drop down menu (shown on left side).

Tagging data

Northern cod (NAFO Divisions 2J3KL) has a rich history of tagging, starting in in 1954 (Taggart *et al.*, 1995) and continuing to this day. The tagging and recovery data are captured in a standardized database, with fields typical of most tagging programs. This data base has over 600,000 records as of early 2019, with 2,000-10,000 tags deployed annually in recent decades. The tagging and capture data are used in the current assessment model for this stock (Cadigan, 2016), but tools to explore this extensive data set were limited, especially from a spatial perspective. To begin to explore and understand this large data set, we built a simple application using shiny and leaflet (Table 1) that allows a user to quickly and dynamically subset the data (e.g. ranges of release and capture years, specific geographic locations). As the application developed, it was incorporated into shinydashboard, to take advantage of the number of layout options available in shinydashboard (Table 1). Given the large number of tags released, often at nearby sites, visualizing the data with static mapping was particularly challenging. The markercluster (https://github.com/Leaflet/Leaflet.markercluster) function available in leaflet was particularly useful as a means to dynamically scale the level of pooling of spatial points (Figure 3). This basic mapping tool allowed us to quickly become familiar with the data, identify outliers and incorrect data entries, as well as explore options on how to spatially pool the data for subsequent demographic analysis. Further tabs were added to provide basic summaries of the selected data (Figure 4).

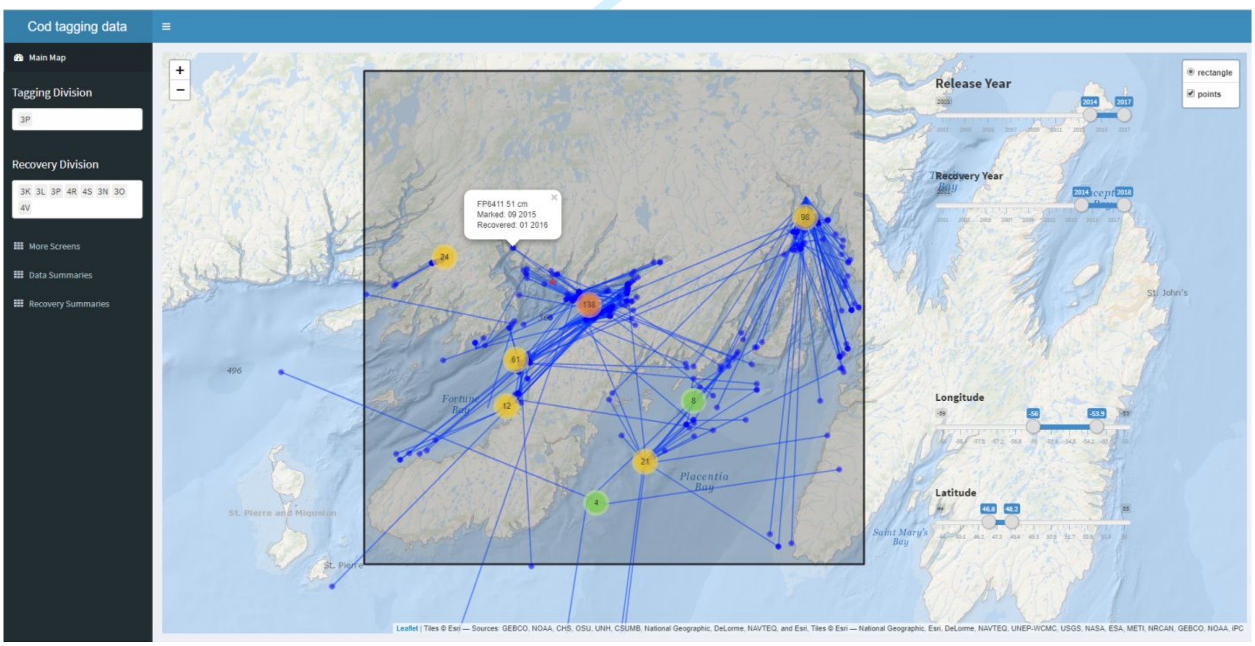

Figure 3: Screen shot of cod tag mapping tool using shinydashboard and leaflet. The markercluster function dynamically splits or pools tagging locations (red, orange, green or yellow points) depending on zoom level, the recoveries positions (blue) are much fewer, and are

left to be plotted individually at all scales. Options to include pop up labels are included, so specific information on each point can be retrieved with a mouse click (in this case: tag number, fish length, date released, and date captured), which is particularly useful when error checking.

Figure 4: Basic summaries of the recovery data from the tags selected within the shiny dashboard cod tag mapping tool. In this case, histograms and summary statistics of the recovery positions are returned, along with a simple map of kernels showing the 2D spatial distribution of the selected tag recoveries.

Data modeling

Synthesizing data from multiple sources presents a key challenge to stock assessment. Analyses of different data sources were traditionally carried out independently and the summaries or parameters from these analyses were used in the assessment model. This approach, however, is less than ideal because information may be lost and uncertainty may be unaccounted for when

161 we "do statistics on the statistics" (Link, 1999; Maunder and Punt, 2013). Such issues have
162 largely been curtailed in contemporary stock assessments thanks to advances in statistical
163 computing that have facilitated the analysis of all available data, in as raw a form as appropriate,
164 in a single integrated analysis (Maunder and Punt, 2013). Specifically, statistical modeling tools
165 such as JAGS (Plummer, 2003), AD Model Builder (Fournier *et al.*, 2012) and Template Model
166 Builder (Kristensen *et al.*, 2016) allow the construction of a joint likelihood for an array of
167 observations to, in theory, extract as much information as possible about the biological and
168 fishery processes. From a computational perspective, analyses of a variety of large data sets has
169 never been easier. However, from a human perspective, contemporary stock assessment
170 biologists are faced with the challenge of understanding and integrating data from multiple
171 sources into a single model and communicating the methods and results to stakeholders and
172 fisheries managers. This challenge was palpable for the data-rich case of Northern cod.

173 The Northern cod stock off southern Labrador and eastern Newfoundland is one of the most well
174 studied stocks in eastern Canada. As such, there are multiple monitoring programs that help
175 inform the status of the stock and data from most of these programs have been integrated into a
176 state-space stock assessment model, called NCAM (Cadigan, 2016). The model includes
177 information from research vessel autumn trawl surveys (1983-present), Sentinel fishery surveys
178 (1995-present), inshore acoustic surveys (1995-2009), fishery catch-at-age compositions and
179 partial fishery landings (1983-present), and tagging (1983-present). Using a series of observation
180 equations, this TMB based model reduces thousands of historical data points into quantities such
181 as recruitment, spawning stock biomass, fishing mortality and natural mortality. Once the model
182 is fit to the data, the next step is to produce visual representations of the data and model output.
183 The traditional approach would involve producing static presentations and documents with a

series of figures and tables. However, with large amounts of model inputs and outputs, this approach quickly becomes overwhelming for both the analyst and the stakeholders involved for at least two reasons. First, it is no longer feasible for the analyst to include and describe every figure and table produced in a single document. Second, it is difficult for stakeholders to efficiently digest the information that has been compressed into a series of static slides or pages. Interactive documents provide a potential solution to this problem as they allow much more information to be contained and accessible on a single screen.

In the pursuit of an easier and more efficient way to communicate results from NCAM, an interactive and self-contained "dashboard", called NCAM explorer, was developed for the 2018 assessment of Northern cod (Dwyer *et al.*, 2019). We used R-based packages (Table 1) to construct a tool for exploring the input and output of NCAM, specifically the flexdashboard package to group interactive plotly-based visuals into a dynamic document. We also used the crosstalk package to link the data displayed across multiple plots. Via R Markdown, the dashboard is rendered into a self-contained html file that is reproducible, interactive, and easy to update following modifications to the model or the addition of new data.

The NCAM dashboard (Supplement 2) contains a series of pages, the first of which provides terse point-form background on the model (page named "Background"). Subsequent pages provide a series of diagnostic plots for assessing model fits to catch ("Catch"), survey ("RV survey", "SN survey" and "SS survey"; accessed from the "Surveys" drop-down menu), and tagging ("Tagging") data. For instance, the "RV survey" page includes plots of observed and predicted values of mean numbers per tow captured in the research vessel survey (Figure 5). The dashboard also includes pages focused on model estimates such as catchability and selectivity ("Catchability"), stock size and vital rates ("Trends"; Figure 6), and stock productivity

207 ("Productivity"). Finally, some results from a retrospective analyses are included under the "Retro" page, trends from different models are compared under the "Comps" page, details on the projections are accessed from the "Projections" drop-down menu ("Assumptions", "Past projections", "Retro projections", and "Results" pages), and key inputs and outputs are accessed from "Tables" drop-down menu ("Inputs", "Settings", "Outputs" pages). The plots and tables included in the dashboard are similar to those typically presented at assessment meetings and in research documents, however, there are two key benefits of this approach over the *modus operandi* of producing static documents and slides: 1) interactive plots nested in a dashboard permit relatively easy and efficient access to the details as it replaces scrolling through tens, if not hundreds, of pages or slides with mouse-clicks across pages holding data-rich illustrations (i.e. both broader patterns and finer details in the data are accessible via zooming and tooltips); and, 2) the automated nature of the dashboard circumvents the monotonous, time-consuming and error-prone task of copying and pasting figures, tables and values into documents and slides. Both benefits expedite the process of exploring a range of model configurations as the automated output facilitate quick views of standard diagnostics and the interactive plots facilitate detailed explorations and comparisons of models with different configurations. The dashboard also makes it easy to share the results with colleagues and stakeholders as it is rendered into a self-contained html file. This allows others to independently scrutinize details of both data and the model that are typically only accessible to the analyst. Such access improves the transparency of the stock assessment model which, in turn, leads to richer discussion and scrutiny of the biological and statistical rigor of the model. For instance, visualizations depicting the model assumptions, process errors and confidence intervals around the projections raised important questions on the impact of the assumptions on the projections from the model. These questions were raised during

the first assessment meeting in which this tool was used and the ensuing discussion helped the assessment biologists, managers and fisheries representatives in the room gain a deeper understanding of the uncertainties and the patterns in the risk tables.

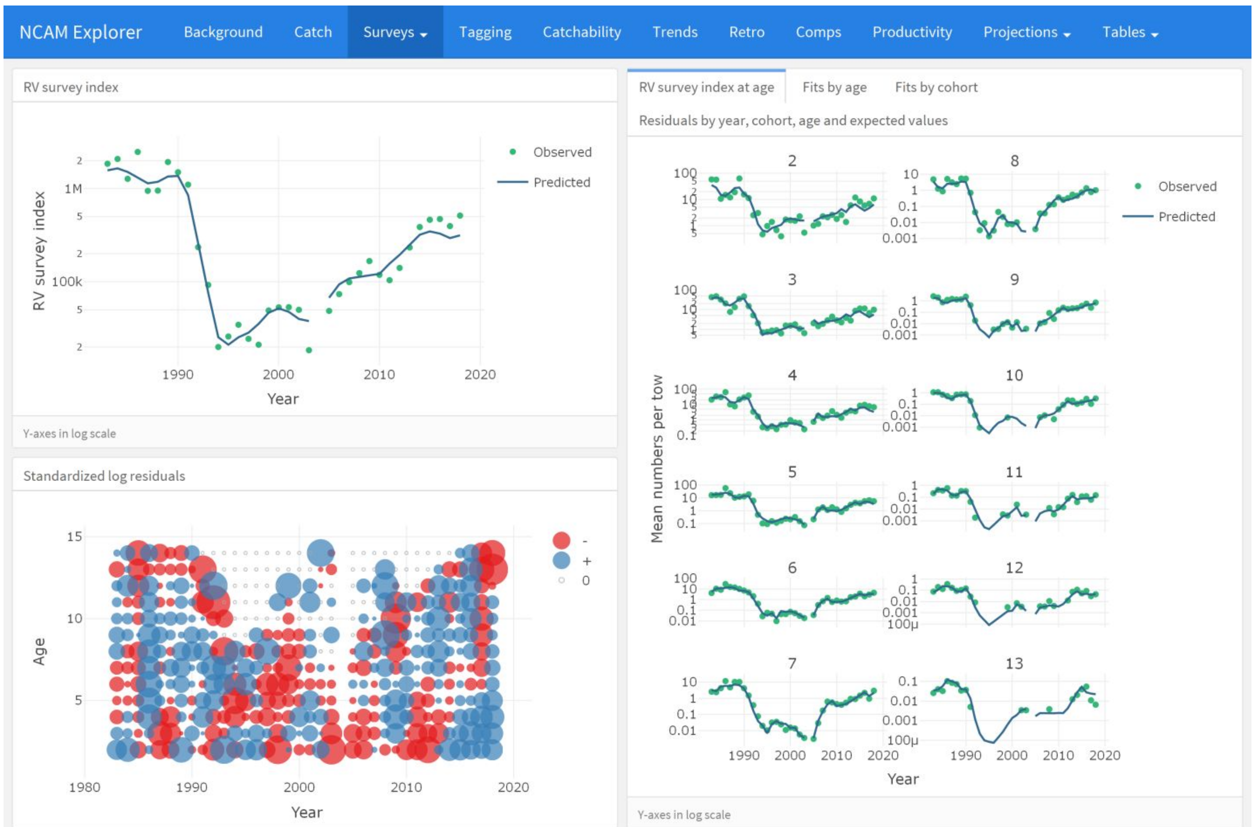

Figure 5: Screenshot of the “RV survey” tab from the NCAM dashboard where total observed (dots) and NCAM model predicted values (lines) for the DFO RV survey index are shown in the upper left panel and scaled matrix plot of age-disaggregated standardized log residuals are shown in the lower left panel (blue = positive, red = negative, symbols scaled by size; grey = index values of zero). Age-disaggregated observed (dots) and predicted (lines) values from the DFO RV survey are shown in the right panel.

Figure 6: Screenshot of the “Trends” tab which displays estimates of recruitment, stock size, and stock size relative to B_{lim} (left panels) and mortality rates (F, M, Z, right panels).

Towards open stock-assessment

The amount of data available to scientists has grown by orders of magnitude in recent decades as has the complexities of data management, exploratory data analysis, formal analyses and associated diagnostics (Lewis *et al.*, 2018). The majority of this sequence of events, sometimes called “the data pipeline” (Leek and Peng, 2015a), have not traditionally been part of the peer-review process which sees only the end products of an analysis. However, decisions made along the data pipeline increasingly influence the outcome of the study. Gelman and Loken (2014) coined the term “garden of forking paths” to illustrate that different conclusions can be arrived at depending on what decisions are made along during different stages of the analysis. Due to a number of limitations, such as available pages in journals, much of the data pipeline is not transparent nor is it reproducible. A number of authors have recently advocated for a culture of open science and reproducible research, i.e., a change in the transparency and reproducibility of science (Hampton *et al.*, 2013, 2015; Leek and Peng, 2015a, 2015b). Proponents of open science and reproducible research highlight a number of benefits including a more productive and responsible scientific culture, an ability to address larger and more complex questions, as well as a more efficient workflow and ability to reproduce one's own work (Fomel and Claerbout, 2009; Lowndes *et al.*, 2017).

The tools presented here are additions to the growing number of interactive stock assessment tools being developed and applied across the globe (see Appendix A). We believe that such tools represent an important a step forward in terms of open science and reproducible research in addition to being a major step forward in how stock assessments are presented and critiqued. Although these tools do not reveal the entire data pipeline, in conjunction with posting code and data on-line, they represent a significant advance over more traditional methods. Similarly, these tools do not fully alleviate pragmatic issues associated with reproducible research. For example, it is unlikely that an individual has the time to fully review and recreate a particular analysis in its entirety when most participants are already overcommitted (Banks, 2011). Further, in the case of NCAM, considerable statistical experience and expertise is required simply to run the model. However, the dashboard approach removes many of these obstacles. The html file contains all model inputs, outputs and diagnostics in an open, and easy to use, interactive document that can be distributed prior to meetings. As such, participants can assess the results and diagnostics at their leisure. Specific experience with the tools used to generate the results are not required to constructively and critically review of the results presented in the dashboard.

Conclusions

We must acknowledge there are costs to adopting the use of interactive tools. The first and obvious one being that staff are required to have both the time and training to effectively develop these tools. In our three examples, the developers (first, second and third authors) were relatively proficient with R and have natural aptitudes for programming, but all had their formal training in field-based population ecology. While some learning time was needed (a few weeks), the learning curve was relatively minor because there was no need to learn a new programming language (e.g. JavaScript, CSS, etc.). Moreover, a growing user base means that most programming issues were readily addressed through simple internet searches. We expect the accessibility and quality of these tools to continue to improve as more people in the fisheries community develop and use interactive documents. It is our experience that the upfront investment in time has already paid dividends in terms of efficiencies of delivering products for subsequent stock assessments, and in the quality of the data exploration and understanding. Finally, scientific exploration is rarely a linear process, and the ability to navigate around a dashboard fits the natural way people pursue ideas.

Acknowledgements

We thank the numerous colleagues and participants of various stakeholder and stock assessment meetings who encouraged us to further develop these interactive visualization tools, and especially those who took the time to make suggestions on how to make them more accessible and useful. We are also grateful for the constructive feedback from Gary Carvalho, Kelli Johnson, Colin Millar and two anonymous reviewers.

Appendix A

Partial list of applications developed to explore and communicate marine and fisheries science:

- Marine Aggregates Application (https://openscience.cefas.co.uk/ma_tool/) from Cefas
- Benthic Non Native Species Tool (https://openscience.cefas.co.uk/invasive_species/) from Cefas
- Species Dashboard (https://shiny.marine.ie/speciesdash/) from The Marine Institute
- The Stock Book (https://shiny.marine.ie/stockbook/) from The Marine Institute
- Data-Limited Fisheries Toolkit demo (https://www.datalimitedtoolkit.org/demo) from the Fisheries Center, University of British Columbia
- Exploring Groundfish in an Ecosystem Context (https://incorporatingecosystemapproach.shinyapps.io/indiapp/) from Fisheries and Oceans Canada, Maritimes Region
- Catchy Data (https://eu.oceana.org/en/catchy-data) from Oceana
- VISA tool (e.g. https://ices-tools-dev.github.io/VISA_tool/hke.27.3a46-8abd.html#ices_advice_2018) from ICES, which was developed to meet an "EU request on dissemination of ICES advice beyond pdf files"

References

Allaire, J. 2017. Flexdashboard: R markdown format for flexible dashboards. https://CRAN.R-project.org/package=flexdashboard.

Allaire, J., Horner, J., Xie, Y., Marti, V., and Porte, N. 2018. Markdown: 'Markdown' rendering for r. https://CRAN.R-project.org/package=markdown.

Banks, D. 2011. Reproducible research: A range of response. Statistics, Politics, and Policy, 2: 4. bepress.

Cadigan, N. G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 73: 296–308. NRC Research Press.

Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. 2018. Shiny: Web application framework for r. https://CRAN.R-project.org/package=shiny.

Chang, W., and Borges Ribeiro, B. 2018. Shinydashboard: Create dashboards with 'shiny'. https://CRAN.R-project.org/package=shinydashboard.

Cheng, J. 2016. Crosstalk: Inter-widget interactivity for html widgets. https://CRAN.R-project.org/package=crosstalk.

Cheng, J., Karambelkar, B., and Xie, Y. 2018. Leaflet: Create interactive web maps with the javascript 'leaflet' library. https://CRAN.R-project.org/package=leaflet.

Dwyer, K., Brattey, J., Cadigan, N., Healey, B., Ings, D., Lee, E., and Maddock Parsons, D. *et al.* 2019. Assessment of the northern cod (*Gadus morhua*) stock in nafo divisions 2J3KL in 2018. DFO Can. Sci. Advis. Sec. Res. Doc., 2019/nnn: v + 126p.

Fisherkeller, M. A., Friedman, J. H., and Tukey, J. W. 1988. PRIM-9: An interactive multidimensional data display and analysis system. Dynamic Graphics for Statistics: 91–109. Wadsworth & Brooks/Cole.

Fomel, S., and Claerbout, J. F. 2009. Reproducible research. Computing in Science & Engineering, 11: 5. IEEE Computer Society.

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M. N., and Nielsen, A. *et al.* 2012. AD model builder: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. Optimization Methods and Software, 27: 233–249. Taylor & Francis.

Gelman, A., and Loken, E. 2014. The statistical crisis in science. American Scientist, 102: 460.

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., and Duke, C. S. *et al.* 2013. Big data and the future of ecology. Frontiers in Ecology and the Environment, 11: 156–162. Wiley Online Library.

Hampton, S. E., Anderson, S. S., Bagby, S. C., Gries, C., Han, X., Hart, E. M., and Jones, M. B. *et al.* 2015. The tao of open science for ecology. Ecosphere, 6: 1–13.

Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment: Choice, dynamics and uncertainty. Chapman; Hall.

Jones, A. S., Horsburgh, J. S., Jackson-Smith, D., Ramírez, M., Flint, C. G., and Caraballo, J. 2016. A web-based, interactive visualization tool for social environmental survey data. Environmental modelling & software, 84: 412–426. Elsevier.

Keena, N., Etman, M. A., Draper, J., Pinheiro, P., and Dyson, A. 2016. Interactive visualization for interdisciplinary research. Electronic Imaging, 2016: 1–7. Society for Imaging Science; Technology.

Kristensen, K., Nielsen, A., Berg, C., Skaug, H., and Bell, B. 2016. TMB: Automatic differentiation and laplace approximation. Journal of Statistical Software, 70: 1–21. https://www.jstatsoft.org/v070/i05.

Leek, J. T., and Peng, R. D. 2015a. Statistics: P values are just the tip of the iceberg. Nature News, 520: 612.

Leek, J. T., and Peng, R. D. 2015b. Reproducible research can still be wrong: Adopting a prevention approach. Proceedings of the National Academy of Sciences, 112: 1645–1646. National Acad Sciences.

Letcher, B. H., Walker, J. D., O'Donnell, M. J., Whiteley, A. R., Nislow, K. H., and Coombs, J. A. 2018. Three visualization approaches for communicating and exploring passive integrated transponder tag data. Fisheries, 43: 241–248. Wiley Online Library.

Lewis, K. P., Vander Wal, E., and Fifield, D. A. 2018. Wildlife biology, big data, and reproducible research. Wildlife Society Bulletin, 42: 172–179. Wiley Online Library.

Link, W. A. 1999. Modeling pattern in collections of parameters. The Journal of Wildlife Management, 63: 1017–1027. [Wiley, Wildlife Society].

Lowndes, J. S. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., and Jiang, N. *et al.* 2017. Our path to better science in less time using open data science tools. Nature ecology & evolution, 1: 0160. Nature Publishing Group.

Maunder, M. N., and Punt, A. E. 2013. A review of integrated analysis in fisheries stock assessment. Fisheries Research, 142: 61–74. Elsevier.

Perkel, J. M. 2018. Data visualization tools drive interactivity and reproducibility in online publishing. Nature, 554: 133–134.

Plummer, M. 2003. JAGS: A program for analysis of bayesian graphical models using gibbs sampling. *In* Proceedings of the 3rd international workshop on distributed statistical computing. Vienna, Austria.

381 R Core Team. 2017. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

383 Rideout, R., and Ings, D. 2018. Research Vessel Bottom Trawl Survey Report (NL Region): A stock-by-stock summary of survey information up to and including the 2017 spring and autumn surveys. Canadian Technical Report of Fisheries; Aquatic Sciences No. 3267.

386 RStudio Team. 2015. RStudio: Integrated development environment for r. RStudio, Inc., Boston, MA. http://www.rstudio.com/.

388 Sievert, C. 2018. Plotly for R. https://plotly-book.cpsievert.me.

389 Smith, S., and Somerton, G. 1981. STRAP: A User-Oriented Computer Analysis System for Groundfish Research Trawl Survey Data. Canadian Technical Report of Fisheries; Aquatic Sciences No. 1030.

392 Taggart, C., Penney, P., Barrowman, N., and George, C. 1995. The 1954-1993 Newfoundland cod tagging database: Statistical summaries and spatial-temporal dynamics. Canadian Technical Report of Fisheries and Aquatic Sciences, 2042.

395 Tukey, J. 1977. Exploratory data analysis. Addison-wesley series in behavioral science. Addison-Wesley Publishing Company. https://books.google.ca/books?id=UT9dAAAAIAAJ.

398 Tukey, J. W., and Tukey, P. A. 1985. Computer graphics and exploratory data analysis: An introduction. *In* Proceedings of the sixth annual conference and exposition: Computer graphics'85, pp. 773–785. National Computer Graphics Association, Fairfax, VA.

Yeatman, J. D., Richie-Halford, A., Smith, J. K., Keshavan, A., and Rokem, A. 2018. A browser-based tool for visualization and analysis of diffusion mri data. Nature communications, 9: 940. Nature Publishing Group.