USDA
United States
Department of
Agriculture

Natural Resources Conservation Service

SMARTER
FLOODPLAIN
MANAGEMENT

USER GUIDE

Restoration Enhancement Modeling Tool

Beta Version | December 2018

WILLAMETTE
PARTNERSHIP

The
Freshwater
Trust

Prepared for Willamette Partnership by The Freshwater Trust

Table of Contents

1 Introduction

In 2014, Willamette Partnership received a Conservation Innovation Grant (CIG) from the Natural Resources Conservation Service (NRCS) to integrate the science, policy, and tools to support floodplain markets and incentives. The Freshwater Trust (TFT) is supporting this effort by developing and adapting tools that quantify the environmental benefits from floodplains. The tools developed as part of this effort will help communities quantify the water quality, hydraulic, and hydrologic benefits of floodplain restoration actions and communicate the range of ecosystem services provided by floodplains.

Physical characteristics, including stream channel complexity, lateral connectivity, flood storage, and the interaction of water with the floodplain, are the primary drivers of water quality changes in aquatic systems. These physical characteristics alter the movement of water, allowing high flows to inundate the floodplain and potentially lead to improved water quality. Floodplain inundation can increase infiltration and nutrient cycling, allowing for burial and uptake of phosphorus and the complete removal of nitrogen through denitrification. During periods of inundation, floodplains can act as a sediment sink, as sediment suspended in the water column can settle and be deposited on the floodplain. Complex physical characteristics of the channel, such as meander bends or variable geometry, can increase hyporheic exchange flows. These hyporheic exchange flows can buffer surface water temperatures and produce cold-water patches that support a natural thermal regime.

Multiple models and tools currently exist and provide a means to evaluate different floodplain features and functions. Many of these models can be used together to generate a more holistic understanding of the ecosystem services provided by intact floodplains. At this time, however, none of the available tools are able to quantify all of the water quality benefits that may result from floodplain restoration or management actions. The tool that TFT is currently developing captures and quantifies these environmental benefits, specifically the water quality, hydraulic, and hydrologic benefits of floodplain restoration and management actions. In addition, it provides information to help communicate the range of ecosystem services provided by floodplains.

It is important to note that multiple site-level models that rely on extensive input data to predict changes in water quality and quantity already exist, such as HEC-RAS, Heat Source, CE-QUAL-W2, and AQUATOX. When the requisite data are available, these models can be used to accurately quantify the water quality and quantity benefits of a change in floodplain management. This quantification tool, however, should not be seen as a replacement for these sophisticated models. Rather, TFT designed this quantification tool to function well in situations when extensive data inputs are *not* available to the user. The goal of this effort was not to recreate these complex models, but rather to develop a predictive tool that can be used to assess the benefits of a floodplain restoration project or a change in floodplain management in the absence of extensive data. The intent of this tool is to allow floodplain managers and project developers to evaluate the impact of different management alternatives without engaging in extensive modeling.

The quantification tool does not require that the user complete sophisticated site-level modeling or have a statistical background. The statistical components are a part of the back-end of the tool and will be updated as part of tool maintenance. TFT has designed the quantification tool to allow for a range of users and varying levels of input data. TFT has worked with Willamette Partnership to understand the needs of the project's stakeholders to ensure that the tool remains flexible and functional for a diverse group of users.

In July 2016, TFT completed an initial version of the quantification tool built using the R software environment. In the initial version, the quantification tool included two R scripts that required the user to interface with the R

software directly using a desktop application. In this recent version, TFT made substantial improvements to the user interface of the tool. The primary improvement was changing from desktop-based script to a web-based tool with a graphical user interface. The graphical user interface provides the user with additional information and guidance on parameter selection.

This user guide provides information on how to access, set up, and run the quantification tool. This guide also provides information on how to process and interpret the tool output. The final version of the user guide will also include the procedure to update and maintain the quantification tool.

1.1 SYSTEM REQUIREMENTS

The quantification tool was built using the R software environment. R is a free statistical computing and graphics software environment that includes functionality for data manipulation, calculation, and graphical display (Venables et al., 2016). The R environment includes a large number of statistical procedures, including: linear and generalized linear models, nonlinear regression models, time series analysis, classical parametric and nonparametric tests, clustering, and smoothing. In addition to the statistical procedures, TFT utilized an open-source R package called Shiny, which provides a framework for building interactive web applications within R. Using Shiny, once a tool has been developed it can be uploaded to a server hosted by RStudio, then shared as a webpage that serves as the user interface to the tool. One advantage of this approach is that the only system requirement for the user is the ability to access the web. Moreover, Shiny allows the tool to be built with a responsive design that adapts to the size of the screen or web browser being used. The result is that the tool can be easily viewed on any desktop or laptop computer, tablet, or smartphone.

1.2 QUANTIFICATION OVERVIEW

TFT is using Bayesian networks to represent the relationships and the probabilities associated with the floodplain characteristics that promote changes in water quality. A Bayesian network is a probabilistic graphical model representing the key factors of a system and their conditional dependencies (Stewart-Koster et al., 2010). Bayesian networks can represent the relationship between a set of input variables and a response variable, and are widely used for knowledge representation and reasoning under uncertainty in natural resource management (Kragt, 2009). As probabilistic models, Bayesian networks are inherently good at representing relationships between variables if the relationships involve some amount of uncertainty. They can be a useful modelling tool in situations where different types of variables and knowledge from various sources need to be integrated within a single framework (Kragt, 2009). The Bayesian network used in the tool developed by TFT incorporates the full range of published, quantified outcomes from floodplain conservation or restoration to be considered when determining potential project benefits.

The ability of a Bayesian network to predict the probability of a given outcome is a function of the volume of data that was used to build the conditional probability tables for that network. As such, the focus of the development of this Bayesian network was two-fold: 1) to identify the major drivers of sediment, phosphorus, nitrogen, temperature, and flood storage changes in floodplains in order to create a list of potential nodes for the Bayesian network; and 2) to identify for which of these drivers there is a substantial amount of data available in the scientific literature from which to build conditional probability tables for the Bayesian network.

To develop the tool, TFT used the findings of the literature review completed in Task 1 to determine the key correlations or other influences affecting each constituent of interest. Once these preliminary relationships were identified, a supplemental literature search was performed to gather additional evidence to build the

conditional probability tables. This information was then used to build a graphical structure representing the expected relationships and causal influences, which served as the initial Bayesian network. Due in part to the wide range of different explanatory variables that were investigated in the studies reviewed, the initial network included a large number of nodes, many of which included a limited number of data points to inform the relationship to the outcome variable. Therefore, for the draft tool the network was simplified to include only the nodes whose general relationship to the constituent of interest was clear and supported by a variety of observations from multiple sites or studies within the scientific literature (Figure 1). Overall, 2,729 observations were included in the dataset used to build the first draft of the Bayesian network.
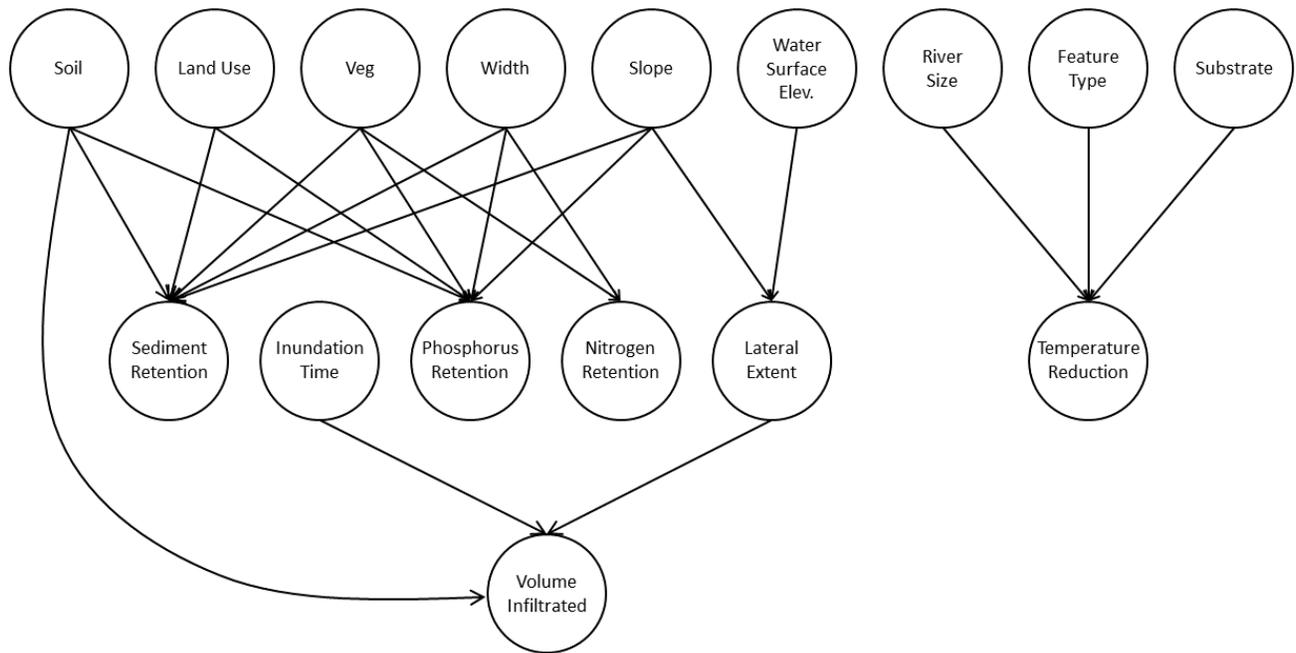

Figure 1. Conceptual diagram of the draft Bayesian network.

2 Instructions

TFT developed the quantification tool using the R software environment. The following sections provide instructions on how to use the quantification tool, the necessary data inputs, and how to process and interpret the results. The quantification tool calls upon the Bayesian model described in the previous section.

The web-based tool can be accessed at the following URL: https://thefreshwatertrust.shinyapps.io/REM_2016

2.1   INPUT DATA

The first user input required by the quantification tool is the categorization of the output probabilities (i.e., model results) into High, Medium, and Low categories. This is done using the sliders within the Goal Setting tab of the tool (Figure 2). This feature allows users to incorporate any threshold values or other values of interest into the tool outputs in order to make the results more relevant to their specific site goals. If the user chooses not to make any adjustments, the model will be run using the default threshold values. When the sliders have been adjusted to their desired positions, the user can click on the "Click to submit settings" button at the bottom left side of the page. This will prompt the tool to run the back-end script and create the conditional probability tables for the Bayesian network using the output categories defined by the user. This process will take a few minutes.

Figure 2. The Goal Setting tab at the top of the page can be used to modify the output categories.

Next, the user can navigate to any of the three tool tabs at the top of the page (REM-Sediment and Nutrients, REM-Temperature, and REM-Infiltration) to adjust site characteristic inputs. Each tab contains a drop down menu for each input parameter that is an explanatory variable for the constituent in question (Figure 3). After the desired changes have been made for each of the input variables the user can run the model by clicking on the "Click to apply changes" button in the bottom left corner. If the updated input settings change the resulting output probabilities, the bar graphs at the center of the page will change accordingly.
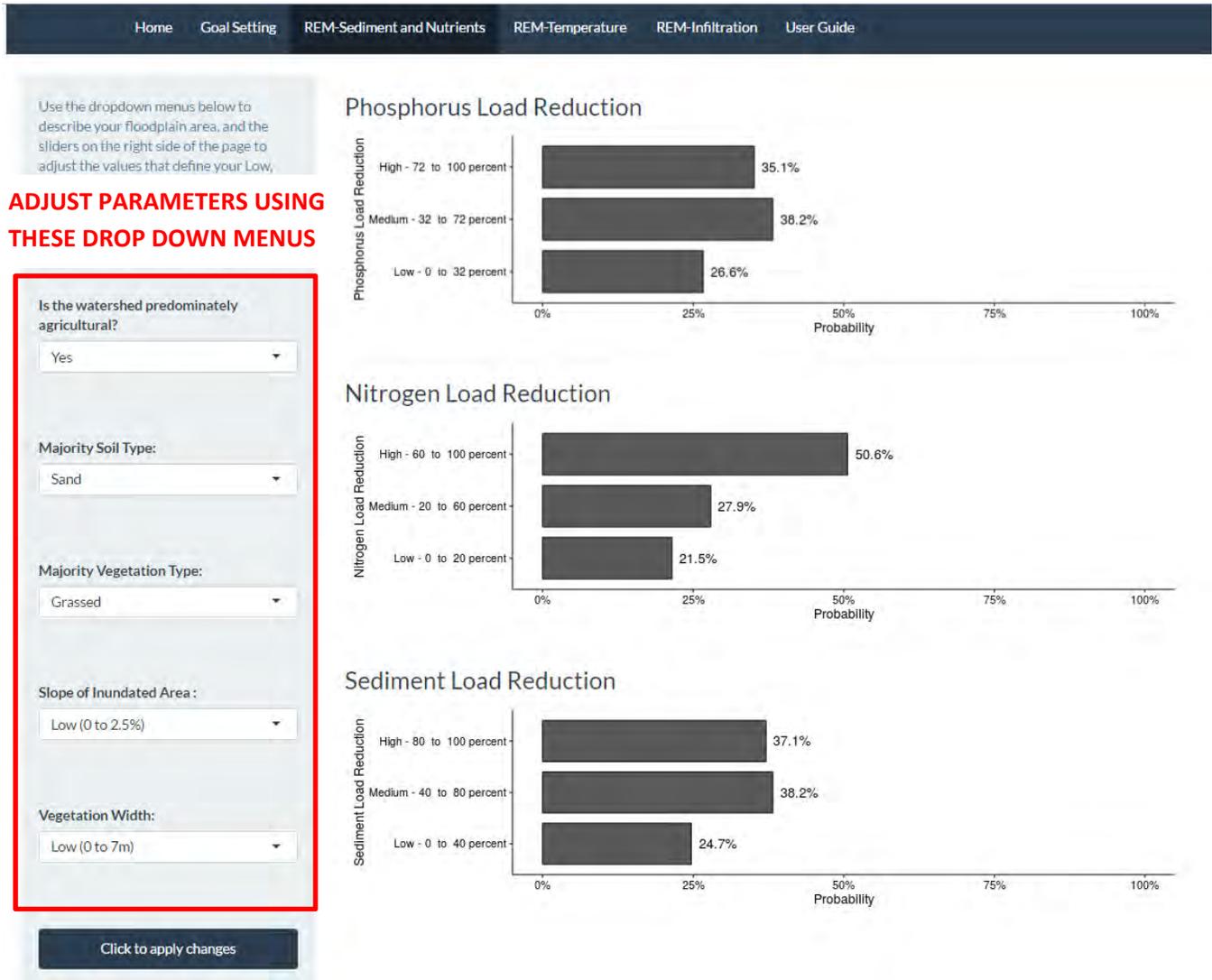
Figure 3. Example of the drop down menus used to adjust input parameters for the REM-Sediment and Nutrient

The input variables for each of the three tabs are shown in Table 1 below. Five of the input variables are inherently categorical, and each of the six remaining numeric variables has been converted to a categorical variable—Low, Medium, or High—using the Jenks method of finding natural breaks in the distribution of values for each variable.

Table 1. User inputs and categorical options for each parameter.

Input Variables: REM – Sediment and Nutrients	Majority Soil Type	Is the watershed predominantly agricultural?	Major Vegetation Type	Vegetation Width	Slope of Inundated Area
Categorical options	Silt	Yes	Grassed	Low (0 to 7m)	Low (0 to 2.5%)
	Sand	No	Herbaceous	Med (7.1 to 21.3m)	Med (2.6 to 6%)
	Clay		Forested	High (>21.4m)	High (> 6.1%)
	Loam				

Input Variables: REM-Temperature	River Size	Dominant Substrate			
Categorical options	Small (5th order or smaller)	Gravel			
	Large (larger than 5th order)	Cobble			
		Fine			
		Sand			

Input Variables: REM-Infiltration*	Inundation Time (Days)	Water Surface Elevation (m)	Slope of Inundated Area	Majority Soil Type	
Categorical options	Low (0 to 1.14 days)	Low (0 to 0.216m)	Low (0 to 2.5%)	Silt	
	Med (1.15 to 3.75 days)	Med (0.217 to 0.384m)	Med (2.6 to 6.0%)	Sand	
	High (>3.75 days)	High (> 0.384m)	High (> 6.0%)	Clay	
				Loam	

* Buffer slope and soil type are also input variables for flood storage

This script and the accompanying database will be updated regularly as part of the tool maintenance process when additional data become available.

2.2   OUTPUT AND RESULTS

The output consists of two parts: (1) the output category ranges as defined by the user on the Goal Setting tab of the tool, for each of the five outputs; and (2) the resulting probability associated with each outcome category range (Figure 4).
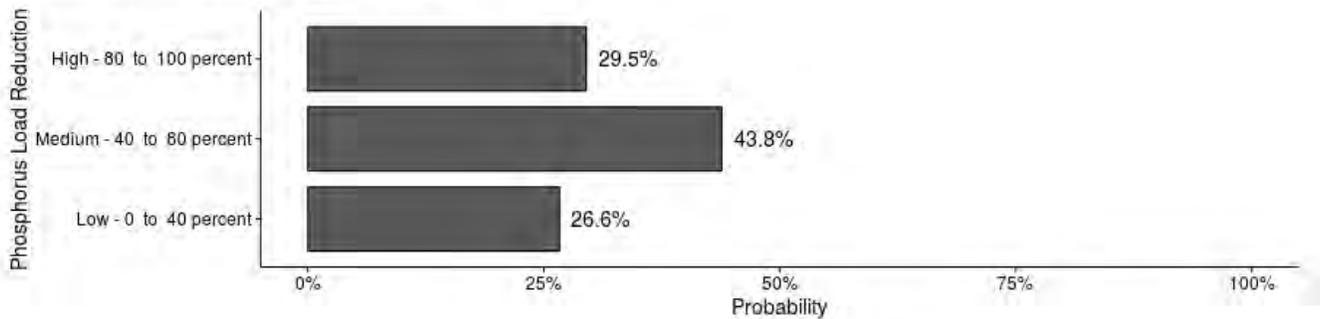
Figure 4. Example of the output produced by the quantification tool. The horizontal bars represent the probability of seeing a phosphorus load reduction within the Low, Medium, and High ranges described on the y-axis.

2.3   PROCESSING AND INTERPRETING THE RESULTS

After running the tool, the user will be presented with a set of results and associated probabilities. These results are based on the user-selected input variables, and include the calculated probability associated with each user-defined benefit category for each parameter. These probabilities represent the expectation of achieving these low, medium, or high benefits based on the reported literature.

2.3.1   Nitrogen, Phosphorus, and Sediment

The nitrogen, phosphorus, and sediment benefits presented are percent reductions. These percent reductions can be converted to a nutrient or sediment load reduction using incoming water quality data. Incoming nitrogen, phosphorus, and sediment water quality data, as well as discharge data are necessary to complete the load calculation. Nutrient and sediment loads can be measured and quantified using pre-project water quality and discharge data from the project area. In cases where insufficient water quality data are available to quantify the incoming load, the load can be estimated based on land use and land cover. Each land use category is associated with a runoff export coefficient that represents the amount of a pollutant loaded to a system from a defined area each year. The incoming nutrient and sediment load can be estimated using regional runoff export coefficients and watershed land use characteristics using the Export Coefficient Model approach described by Reckhow et al. (1980).

Once the incoming nitrogen, phosphorus, or sediment loads have been calculated, the quantification tool results can be applied:

$$\text{Load Reduction} = [\text{Percent Reduction}] \times [\text{Incoming Load}]$$

The load reduction calculation must be completed for the upper and lower value of each benefit range to characterize the range of the water quality benefits.

2.3.2   Temperature

The temperature benefits expressed by the quantification tool are those associated with cold-water microhabitats or thermal refugia. The water temperature benefits predicted by the quantification tool represent the expected difference in water temperature as a result of hyporheic exchange flows through a geomorphic feature. The water temperature difference expressed by the tool represents the difference in temperature between the mainstem water column and an individual thermal refuge.

The literature values used to populate the database are from conditions observed during the summer. As such, the predicted thermal benefits are not expected to occur year-round. Additionally, the temperature difference described is the daily maximum difference in water temperature. That is, these are the thermal refuges that are available to aquatic species when mainstem water temperature peaks.

The abundance and type of geomorphic features present within a reach will drive the availability of thermal refugia. At this time, the thermal benefit of a cold-water refuge cannot be translated into a thermal load reduction without more detailed information regarding hyporheic exchange flows through the feature. The tool outputs, along with information regarding the presence of geomorphic features, can be used to determine the abundance of thermal refugia within a project reach.

2.3.3   Water Infiltrated
Infiltrated water benefit represents the volume of water infiltrated at a single cross-section of the stream, which has been normalized to a one-meter slice of longitudinal extent. The result can be applied to the full stream reach by multiplying by the reach length. The resulting volume represents the infiltrated water quantity associated with the input characteristics.

Based on a particular combination of soil type, water surface elevation, overbank duration, and bank slope, an infiltration volume range is estimated from literature values. Differences in this quantity of retained floodwater as a result of restoration actions can be used to represent the potential for peak discharge reduction, increased baseflow, and improved floodplain connectivity.

2.3.4   Interpreting Benefit Probabilities
The probabilities included with the tool results provide the user with an estimate of the likeliness of the water quality or quantity benefit occurring given the project features. These probabilities are based on the benefits documented in the literature. The probability associated with a benefit range indicates the level of confidence that the site features will result in that range of benefits. Higher probabilities indicate a higher level of confidence in the occurrence of that range of benefits.

For some feature combinations, there may not be sufficient information in the initial network to refine the probability of each benefit category for a parameter. For example, the calculated result may be 33.3% for each benefit category for a single parameter. In this example, there is insufficient information to refine the prediction further and each benefit category is equally as likely to occur. As new information becomes available and is used to update the conditional probability tables, the frequency of these results will be reduced.

3 Future Modifications

The quantification tool is currently in draft form. The focus of the initial tool development effort (completed in July 2016) was to construct the foundational networks and data aggregation that serve as the basis of the tool. TFT continued to enhance the tool by refining the user interface and workflow as part of Task 3. As part of that refinement, TFT simplified the user's experience with the tool by migrating the tool from a desktop application to a web-based tool. The new version of the tool has streamlined the process for setting input conditions, and has incorporated additional flexibility to meet the needs of the user.

The next phase of the project includes piloting the quantification tool with the support of the project stakeholders. Through the piloting phase, TFT intends to identify the aspect(s) of the tool that would benefit from additional refinement. Additionally, through collaboration with the project stakeholders, TFT will work to prioritize the set of tool improvements in order to maximize the utility and usability of the quantification tool. Some of the potential tool improvements that TFT will explore and prioritize with the project stakeholders could include:

1) At this time, the quantification tool requires that the user supply all of the required site-level data. A future tool improvement could include an automated GIS process to assist the user with the aggregation and processing of spatial data.

2) An additional tool improvement could include a modification to how the tool presents and processes the site-level outputs. At this time, the user must manually convert the results to a pollutant load. A subsequent iteration of the tool could provide the user with an additional feature to calculate the incoming pollutant load and the resulting load reduction. The feature could use watershed land use characteristics and the Export Coefficient Model approach described by Reckhow et al. (1980) to determine the incoming nutrient and sediment loads. These watershed characteristics could be combined with the tool results to provide the user with an estimate of the nutrient and sediment load reduction, reducing the need for the user to complete additional calculations.

3) The current version of the tool can quantify the temperature benefit of instream features (riffles, pools, etc.). A future expansion of the tool could include a feature to rapidly quantify the thermal benefits from riparian shade without complex modeling. This expanded temperature feature could allow the user to assess the thermal benefits from riparian shade based on local reach conditions, such as channel width, stream aspect, and riparian vegetation. The results of the rapid assessment would be reported as a thermal load reduction (kilocalories/day).

Once the final form of the quantification tool has been determined with the project stakeholders, TFT will develop a tool maintenance protocol to facilitate the long-term maintenance of the quantification tool. The protocol will outline the steps necessary to update or modify the tool, particularly how to incorporate new or additional data into the tool. One of the benefits of using a probabilistic approach to quantify benefits is that the tool can easily be updated as more information becomes available. In fact, the predictive power of the quantification tool increases with an increase in data. Consequently, TFT has built the quantification tool to allow for future updates to the conditional probability tables based on new data. As part of the finalization of the tool, TFT will document the data update procedure, including how to incorporate additional data and how to update the conditional probability tables. The update procedure will be included in the final User Guide.

After receiving feedback from the project team and stakeholders, TFT will incorporate the feedback into a subsequent iteration of the quantification tool. Any necessary modification will be documented in the updated User Guide and Background documents.

4 References

Kragt, M. E. (2009). *Technical Report No.9: A beginners guide to Bayesian network modelling for integrated catchment*. Landscape Logic Technical Report.

Reckhow, K. H., Beaulac, M. N., & Simpson, J. T. (1980). *Modeling Phosphorous Loading and Lake Response Under Uncertainty: A Manual and Compilation of Export Coefficients* (Vol. 1). EPA-440/5-80-011, U.S. Environmental Protection Agency, Washington, D.C.

Stewart-Koster, B., Bunn, S. E., MacKay, S. J., Poff, N. L., Naiman, R. J., & Lake, P. S. (2010). The use of Bayesian networks to guide investments in flow and catchment restoration for impaired river ecosystems. *Freshwater Biology*, *55*(1), 243–260.

Venables, W. N., Smith, D. M., & R Core Team. (2016). An Introduction to R. Notes on R: A Programming Environment for Data Analysis and Graphics, Version 3.3.0 (2016-05-03).