PNNL-22405

US Army Corps of Engineers®

Prepared for the U.S. Army Corps of Engineers, Portland District,
Under an Interagency Agreement with the U.S. Department of Energy
Contract DE-AC05-76RL01830

Synthesis and Evaluation of Research, Monitoring, and Evaluation and Restoration Effectiveness Data in the Lower Columbia River and Estuary, 2012

DRAFT ANNUAL REPORT

AM Coleman
GE Johnson
AB Borde
HL Diefenderfer
NK Sather
TE Seiple
JA Serkowski

April 2013

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.
(8/2010)

PNNL-22405

Synthesis and Evaluation of Research, Monitoring, and Evaluation and Restoration Effectiveness Data in the Lower Columbia River and Estuary, 2012

DRAFT ANNUAL REPORT

AM Coleman
GE Johnson
AB Borde
HL Diefenderfer
NK Sather
TE Seiple
JA Serkowski

April 2013

Prepared for
U.S. Army Corps of Engineers, Portland District
under an Interagency Agreement with the U.S. Department of Energy

Pacific Northwest National Laboratory
Richland, Washington 99352

Preface

Pacific Northwest National Laboratory (PNNL) conducted this project for the U.S. Army Corps of Engineers, Portland District (Corps). The work was coordinated regionally under the Anadromous Fish Evaluation Program, study code EST-P-12-01. The PNNL project managers were Gary Johnson and Andre Coleman. The Corps technical lead was Cynthia Studebaker. The purpose of the project is to develop a geospatial, web-accessible database (called *Oncor*) for action effectiveness and related data from monitoring and research efforts for the Columbia Estuary Ecosystem Restoration Program (CEERP). The intent is for the *Oncor* database to enable synthesis and evaluation, the results of which can then be applied in subsequent CEERP decision-making. This is the first annual report in what is expected to be a 3- to 4-year project, which commenced on February 14, 2012. This report covers project activities conducted from February 2012 through February 2013.

There are many ongoing database efforts in the Columbia River basin, some of which are coordinated through the Pacific Northwest Aquatic Monitoring Partnership (PNAMP). To avoid duplication of effort, we strived to adopt existing PNAMP structures and terminology for use in *Oncor*. A substantial effort was undertaken to coordinate the *Oncor* database for CEERP data with databases supporting data collected elsewhere in the Columbia River basin. A glossary is included in the front material of this report to aid communication.

The 2012 project included three general activities: coordination among *Oncor* stakeholders, database and user interface development, and CEERP support. A significant work product for the second activity is a document called "Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program." This document will be released under separate cover at a later date. Work products for the third activity, the 2013 CEERP Strategy Report and Action Plan, were delivered separately to the Corps and are not included in this report. For copies of these documents, please contact C. Studebaker (503 808 4788).

Suggested citation for this report:

Coleman AM, GE Johnson, AB Borde, HL Diefenderfer, NK Sather, TE Seiple, and JA Serkowski. 2013. *Synthesis and Evaluation of Research, Monitoring and Evaluation and Restoration Effectiveness Data in the Lower Columbia River and Estuary, 2012*. PNNL-22405, draft annual report submitted to the U.S. Army Corps of Engineers, Portland District, Portland, Oregon, by Pacific Northwest National Laboratory, Richland, Washington.

Summary

The study reported herein was conducted by researchers at the Pacific Northwest National Laboratory (PNNL) for the U.S. Army Corps of Engineers, Portland District (Corps). The purpose of the project is to develop a geospatial, web-accessible database (called *Oncor*) for action effectiveness and related data from monitoring and research efforts for the Columbia Estuary Ecosystem Restoration Program (CEERP). The intent is for the *Oncor* database to enable synthesis and evaluation, the results of which can then be applied in subsequent CEERP decision-making. This is the first annual report in what is expected to be a 3- to 4-year project, which commenced on February 14, 2012. This report covers project activities conducted from February 2012 through February 2013.

The specific objectives and tasks for the 2012 study year were as enumerated here. 1) Coordination − Coordinate with CEERP funding agencies and regional stakeholders to finalize key analysis questions and database needs for research, monitoring, and evaluation (RME) and ecosystem restoration in the lower Columbia River and estuary (LCRE) within CEERP's adaptive management framework. 2) Database Development − Develop and demonstrate a web-based proof-of-concept geospatial database management system and analysis system for CEERP. 3) Analysis and Synthesis[1] − Draft the CEERP 2012 Strategy Report, 2012 Action Plan, and 2013 Synthesis and Evaluation Memorandum, and provide other analytical and programmatic support. Work products for Objective 3 were reported separately.

During 2012, the first year of the *Oncor* database development effort, several important accomplishments advanced upon meeting the objectives, summarized as follows by Objectives 1 and 2:

<u>Objective 1 Coordination</u>

- The extensive coordination effort was fruitful in promoting understanding, constructive input, and buy-in from the user community.
- Substantive coordination occurred with the database development efforts for the Pacific Northwest Aquatic Monitoring Partnership and Bonneville Power Administration's restoration project tracking system, called "cbfish.org."
- Communication and coordination was undertaken with the Lower Columbia Estuary Partnership and its data responsibilities.

<u>Objective 2 Database Development</u>

- An *Oncor* data model was conceived and developed for RME and other data from CEERP efforts in the LCRE.
- An *Oncor* database with user interface was built and launched successfully, establishing proof of concept.
- Data reduction procedures, including Data Exchange Templates, were established for water-surface elevation and sediment accretion.

[1] The 2012 work products from the Analysis and Synthesis objective were delivered separately to the Corps and are not included in this report. For copies of the documents, contact C Studebaker (503 808 4788).

- Example data from a variety of data categories were normalized and uploaded into *Oncor*.
- The *Oncor* data model and database were tested using example data addressing particular analysis questions and use cases.

Much work remains to be done to fully develop and apply *Oncor*. We recommend the following activities during 2013 and beyond:

- Continue coordination efforts (Avenues A, B, and C—three approaches to database coordination established in fiscal year 2012 involving funding agencies and regional stakeholders to coordinate database needs for RME, and ecosystem restoration in the LCRE within the CEERP's adaptive management framework.).
- Disseminate draft procedures for water-surface elevation and sediment accretion and solicit feedback from the user community
- Continue work on new data reduction procedures for other data categories, such as fish and vegetation; develop data entry tools.
- Establish routine data uploads from the Estuary Partnership's Ecosystem Monitoring project, the Washington Department of Fish and Wildlife's restoration Memorandum of Agreement projects, and other studies.
- Populate *Oncor* with regionally available, normalized data to prepare to support estuary-wide meta-analysis of effectiveness data.
- Develop tools for automatic data processing of particular data sets (to be determined) within *Oncor* after data have been uploaded.
- Solicit feedback on *Oncor*'s user interface and update accordingly.
- Post data and links to other databases on the *Oncor* web portal to make them accessible to CEERP stakeholders.

When the primary *Oncor* development effort is complete and the database is ready to be transferred to an entity (to be determined) for long-term operation and maintenance, LCRE regional stakeholders will have a functioning geospatial, web-accessible database for action effectiveness and related data from monitoring and research efforts that can be synthesized and evaluated to support CEERP and its ecosystem restoration mission in the LCRE.

Acknowledgments

We thank the following colleagues for their help with this project during 2012:

- Russell Scranton and Ben Zelinsky of the Bonneville Power Administration
- Catherine Corbett, Keith Marcoe, and Mathew Schwartz of the Lower Columbia Estuary Partnership
- Sarah Apsens, Susan Ennor, and Joe Lettrick of PNNL
- Mike Langeslay and Cynthia Studebaker of U.S. Army Corps of Engineers, Portland District.

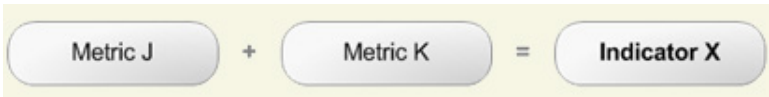
Acronyms and Abbreviations

AEMR	action effectiveness monitoring and research
BiOp	Biological Opinion
BPA	Bonneville Power Administration
CEERP	Columbia Estuary Ecosystem Restoration Program
CHaMP	Coordinated Habitat Monitoring Program
CMOP	Coastal Margin Observation and Prediction
Corps	U.S. Army Corps of Engineers, Portland District
DART	Data Access in Real Time
DET	Data Exchange Template
ELHD	early life history diversity
EP	Estuary Partnership (refers to the Lower Columbia River Estuary Partnership)
EPA	U.S. Environmental Protection Agency
ERTG	Expert Regional Technical Group
F&WP	Fish and Wildlife Program
FCRPS	Federal Columbia River Power System
FPC	Fish Passage Center
FY	fiscal year
GIS	geographic information system
LCRE	lower Columbia River and estuary
LOBO	Land/Ocean Biochemical Observatory
MOA	Memorandum of Agreement
NA	not applicable
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
OHSU	Oregon Health Sciences University
PITAGIS	PIT (passive integrated transponder) Tag Information System
PSMFC	Pacific States Marine Fisheries Commission
PNAMP	Pacific Northwest Aquatic Monitoring Partnership
PNNL	Pacific Northwest National Laboratory
QA/QC	quality assurance/quality control
RDMS	Relational Database Management System
RME	research, monitoring, and evaluation
RPA	reasonable and prudent alternative
SEHI	Salmon Ecosystem Habitat Index
SEV	sum exceedance value
USGS	U.S. Geological Survey

VBA	Visual Basic for Applications
WDFW	Washington Department of Fish and Wildlife

Glossary

action	Restoration activity or project, e.g., dike breach.
analysis question	A query of the database addressing a particular research or management need.
attribute	A data value, numeric or character-based, that identifies one geographic entity from another.
basemap data sets	Mapping data that contains basic reference data, such as roads, cities, prominent landscape features, etc. to orient the user.
category*	A classification rank used for summarizing and reporting that is below subject, above subcategory. For example, Fish or Water Quality
convention	Adopted standards for measurements, metrics, units, etc. included in the Data Exchange Template.
custom format	An alternative format to the Data Exchange Template developed for large data sets that predate *Oncor* and are deemed important to include within the database.
data dictionary	A worksheet in a Data Exchange Template that defines all fields included in that Data Exchange Template (i.e., field names, data type, description, etc.).
data event	The lowest grouping of data in the data model that includes a unique combination of measurement type, place, and time. A data event is an organizing principle of the data model.
Data Exchange Template	An Excel workbook containing the data dictionary, one or more templates for data entry, and one or more data examples. A static "expression" derived from a data reduction workbook intended for upload to *Oncor*.
data generator	Any individual or entity that provides data to *Oncor*.
data layer	An individual geographic information system (GIS) data file representing a theme or parts of a theme such as land cover, elevation, or hydrography.
data model	A conceptual database design process structure that includes a lifecycle of end-user needs assessment, data type definitions, linkages of data, design review, implementation, and testing of design.
data provider	See data generator.
data reduction	The process of transforming raw data by statistical or mathematical functions into a more usable format.
data reduction space	Includes a set of data processing procedures, including quality assurance/quality control, and a Data Exchange Template.
data steward	An individual user or entity that maintains control over a data set.
data table	Numerical and/or textual information structured into rows and columns and may or may not be linked to spatial features.
data theme	A categorization of GIS data, sometimes synonymous with data layer, into groupings of geographic objects that share a common purpose, function, or type such as vegetation type, soil texture, dike, and tide gate locations.

data type	The attribute of a variable, field, or column in a table that determines the kind of data it can store. Common data types include character, integer, decimal, single, double, and string.
database	A collection of structured, interrelated information stored as a series of tables in a commonly accessible information system and can include both spatial and non-spatial data.
data set*	A collection of data, usually presented in tabular form. Each column represents a particular variable. Each row corresponds to a given member of the data set in question. It lists values for each of the variables, such as the height and weight of an object. Each value is known as a datum. The data set may comprise data for one or more members, corresponding to the number of rows. Nontabular data sets can take the form of marked up strings of characters, such as an XML file.
derived data	Using a base set of data, either tabular or spatial, multiple variables and/or mathematical functions are used to convert data to another form, revealing additional metrics or information.
domain table	A table within a database defining unique allowable values for a given column of data to aid in reduction of data errors, e.g., a user can only include one of the following values: 1, 6, 12, 18, 24 for column *X* in Table *Y*.
end-user	The individual, organization, or entity that is using a developed product.
estuary-wide scale	Historical floodplain from Bonneville Lock and Dam to the mouth of the Columbia River.
feature class layer	A collection of geographic features that share a common feature geometry (i.e., point, line, polygon).
feature geometry	Spatial representation of geographic objects within a data theme that are represented by a point, line, polygon, or distributed grid/mesh.
field description	The background information for a column of data in a data table that corresponds to the field name. It may include what the data in the column represent and what the units are for the data.
field name	The given identifier for a column of data in a data table, i.e., the header.
foreign key	Within data tables in a database, an attribute or set of attributes in one table that match the primary key attributes in another table with the intent of joining one or more data tables together. Also see Primary Key.
indicator*	Value resulting from the data reduction of metrics across sites and temporal periods based on applying the procedures in the inference design. A reported value used to indicate the status, condition, or trend of a resource or ecological process; intended to answer questions posed by the objectives of the protocol. Contrast with metric. 
keyword	A single word or short phrase that describes the context and content of a given data set.

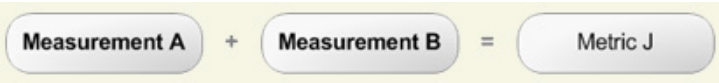
landscape scale — A spatial and ecological scale that makes use of site scale and regionally available information within a larger system to consider process and function in a more system-based and holistic manner.

legacy data — Historic data and/or data collected and structured using an older protocol.

local site — An individual reference, control, or restoration site.

managed data — Spatial and tabular data held in the local *Oncor* database located on a server managed by the *Oncor* development team. These data may come from multiple entities who have agreed to store data locally within the *Oncor* database. Also see Unmanaged Data.

measurement* — A value resulting from a data collection event at a specific site and temporal unit. Measurements can be used to produce metrics using a response design.



metadata* — Literally, "data about data;" provides information about aspects such as the "who, what, where, and when" of data and can be considered from the perspective of both the data producer and the data consumer.

metric* — A value resulting from the reduction or processing of measurements taken at a site and temporal unit at one or more times during the study period based on the procedures defined by the response design. Metrics can be used to estimate an indicator using an inference design. Note that a variety of metrics can be derived from original measurements.

monitored indicators — Values resulting from data reduction of metrics sourced from a time-series of field-collected data around specific themes of 1) water-surface elevation, 2) water temperature, 3) channel cross-section surveys, 4) sediment accretion, 5) vegetation, and 6) fish. Also see Indicators.

non-spatial data — Information structured without reference to a geographic object, these types of information would typically be stored in Data Tables.

normalized data — Data that has undergone processing to conform to a set of rules about the data, such as strictly adhering to the defined Data Type, eliminating or standardizing null data values, and/or checking for anomalous data against the established Domain Tables. This processing is completed to prepare for loading into a database.

pedigree — The recorded source and history of a given data set for purpose of understanding integrity of data and appropriate use and application of the data.

primary key — Within data tables in a database, an attribute or set of attributes in a database that uniquely identifies each record with the intent of joining one or more data tables together. A primary key allows no duplicate values and cannot be null.

published services — A means of making available, publicly or privately, data that can be accessed and transferred over the internet using web-enabled applications in a seamless, behind-the-scenes manner using one of a number of established

	protocols such as REST, WSDL, or JSON; also commonly known as "web services.". See Web Services.
raster data	A type of GIS file format that represents a data theme for a geographic area in a continuous manner using an equal size, cell-based, row-column structure (i.e., a matrix). Examples of data in this format include imagery and digital elevation models. These types of data can be layered into "bands" that represent different phenomena, for example, different ranges in the electromagnetic spectrum, as is found in multi-spectral satellite imagery.
reach	A common hydrogeomorphic area typically using a number of criteria including floodplain boundary, landforms and geology, presence and location of tributaries, gradient, and in the case of estuary systems, salinity and tidal influence.
regional data	Spatial data consistently representing an area with similar physical characteristics or a system or component of a system. In general, regional data are often represented at a coarser spatial scale, but cover a broader geographic area.
relationship	In the context of databases, data from two or more Data Tables are joined through a common Data Field, referred to as a Primary Key or Foreign Key. The linkages to other tables can be set as one-to-one or one-to-many.
SDE	A software technology from Environmental Systems Research Institute (ESRI) for managing spatial data in a Relational Database Management System (RDMS) allowing for enterprise use (large multi-user environment) of geographic data. The technology makes accessing spatial data from the RDMS seamless to the end-user. Also referred to as ArcSDE and Spatial Database Engine.
Site*	The spatial area where one or more measurements are taken at sampling locations and metrics are derived. A more generic term for this is spatial unit.
site-specific data	Field data collected at a specific restoration, control, or reference site
source data	The origin of a particular set of information, whether it is tabular or spatial.
spatial data	Representation of information in a geographic context stored using either one Feature Geometry type or in simple X/Y or longitude/latitude in a Data Table, thus data may or may not be in a standard GIS file format.
standard *Oncor* metrics indicators	a set of metrics and indicators that have been established for each and monitored indicator and are included as a part of the *Oncor* Data Reduction Procedures and Data Exchange Template.
study area	A conglomeration of sites. Also see Site.
subcategory*	A classification rank used for summarizing and reporting that is below category. For example, Fish Abundance or Turbidity.
survey data set	A collection of information sourced from a survey instrument such as a Total Station, theodolite, or global positioning siystem. For research, monitoring, and evaluation work, this type of data is usually collected for cross sections,

	transects, surface or feature elevation points, instrument calibration, or boundary definitions.
tag	See Keyword.
temporal data	Any spatial or tabular data consistently and repeatedly collected over a regular or irregular time interval. This form of data will have date/time stamps associated with the observation value.
unmanaged data	Data owned and maintained by others through a special means of live data access over the internet referred to as "web services." These data are formally referred to as "unmanaged" because the data are not stored within the *Oncor* database nor does the *Oncor* team have control over the data.
use case	Container for analysis questions. Typical application of the database, e.g., Expert Regional Technical Group project template where individuals would have a specific use for the database. A means to organize the analysis questions.
user interface	The aspects of a computer system or program with which a software user can interact, and the commands and mechanisms used to control its operation and input data. In the case of *Oncor*, the user interface is a web-based interface.
vector data	Spatial data taking the form of points, lines, or polygons and stored as a single coordinate pair (in the case of a point) or an ordered list of coordinate pairs representing the vertices of a geographic feature (in the case line or polygon). Compare to Raster Data.
web services	A means of communicating and transferring data over the internet using web-enabled applications in a more seamless, behind-the-scenes manner using one of a number of established protocols such as REST, WSDL, or JSON.
widget	An interactive graphic component of a user interface (such as a button, scroll bar, or menu bar), its controlling program, or the combination of both the component and program. Also see User Interface.

* Definitions obtained from https://www.monitoringresources.org/Resources/Glossary/Index.

Contents

Preface iii
Summary v
Acknowledgments vii
Acronyms and Abbreviations ix
Glossary xi
1.0 Introduction 1.1
1.1 Study Objectives 1.1
1.2 Background 1.1
1.3 Approach and Philosophies 1.4
1.4 Report Contents and Organization 1.5
2.0 Coordination 2.1
2.1 Conceptual Model Updates 2.1
2.2 Regional Coordination 2.3
2.3 Inventory of Existing Databases 2.5
2.4 Data Access and Use Guidelines 2.7
3.0 Database Development 3.1
3.1 Analysis Questions and Example Use Cases 3.1
3.2 Data Reduction Procedures 3.4
3.3 *Oncor* Data Model 3.11
3.3.1 Background 3.12
3.3.2 Data Model Structure 3.13
3.4 Prototype *Oncor* Database and User Interface 3.16
3.4.1 *Oncor* System Architecture 3.16
3.4.2 Interactive Map-Based Interface 3.17
3.4.3 Managed and Unmanaged Data 3.18
3.4.4 Data Loading 3.22
4.0 Accomplishments and Recommendations 4.1
5.0 References 5.1
Appendix A – Outline for Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program A.1
Appendix B – Data Reduction Procedure for Water Surface Elevation B.1

Figures

1.1 CEERP adaptive management process 1.4
1.2 Conceptual model for the *Oncor* database project 1.5
2.1 Conceptual model for SEHI 2.2
2.2 Schematic of the hierarchy of roles and responsibilities from the database administrator to the end-user 2.8
3.1 Schematic of data flow from data generator to *Oncor* 3.6
3.2 PNAMP data classification tree used to guide the *Oncor* data model into a common conceptual structure 3.13
3.3 Simplified *Oncor* data model 3.14
3.4 Web-based geographic information system tools/geoprocessing functions to directly interact with the data 3.18
3.5 Retrieval of "unmanaged" data time-series defined by an interactive map-based box selection and user-based query filter to narrow date range and attributes of interest 3.19
3.6 Plotting of "unmanaged" data retrieved through interactive area-of-interest selection on map interface 3.20
3.7 Interactive selection, access, and retrieval of instrument data and time-series plots for user-defined time periods 3.20
3.8 Workflow to bring data processed through the Data Reduction Procedure into the *Oncor* database 3.23

Tables

2.1 Regional database coordination avenues 2.3
2.2 Discussion points of regional coordination meetings/calls for *Oncor* during 2012 ..... 2.4
2.3 Inventory of databases especially relevant to CEERP and relationship to *Oncor* 2.6
3.1 CEERP objectives with corresponding analysis questions and *Oncor* data 3.1
3.2 Example use cases 3.2
3.3 Example data dictionary for sediment accretion or erosion rate 3.7
3.4 Example metadata and column override options for sediment accretion and erosion rate 3.8
3.5 Example sediment accretion or erosion rate data collected in 2 years at the same location 3.9
3.6 Example sediment accretion or erosion rate means calculated for 2 years in the same location, from data in Table 3.5 ..... 3.10
3.7 Example sediment accretion rate calculated from July 14, 2005 through July 9, 2009 at one location 3.10
3.8 *Oncor* –Example of data tracking process for managed data 3.21
4.1 General objectives for 2012, 2013, and 2014 for the Synthesis and Evaluation project 4.2

1.0 Introduction

The study reported herein was conducted by researchers at the Pacific Northwest National Laboratory (PNNL) for the U.S. Army Corps of Engineers, Portland District (Corps). The goal of the study is to develop an estuary-wide, geospatial data-management system (called *Oncor*) for research, monitoring, and evaluation (RME) studies and restoration project development under the federal Columbia Estuary Ecosystem Restoration Program (CEERP) in the floodplain study area of the lower Columbia River and estuary (LCRE).

1.1 Study Objectives

The overall (2012 through 2015) objectives of the study are as follows:

1. Coordination − Coordinate with CEERP funding agencies and regional stakeholders to establish key analysis questions and database needs for RME and ecosystem restoration within CEERP's adaptive management framework.
2. Database Development − Develop a geospatial data model and prototype LCRE database.[1]
3. Analysis and Synthesis – Analyze and synthesize data to answer key analysis questions addressing CEERP objectives, and provide analytical support at the program level within the CEERP adaptive management process.

The specific objectives and tasks for the 2012 study year were as follows:

1. Coordination − Coordinate with CEERP funding agencies and regional stakeholders to finalize key analysis questions and database needs for RME and ecosystem restoration in the LCRE within CEERP's adaptive management framework.
2. Database Development − Develop and demonstrate a web-based proof-of-concept geospatial database management system and analysis system for CEERP.
3. Analysis and Synthesis[2] − Draft the CEERP 2012 Strategy Report, 2012 Action Plan, and 2013 Synthesis and Evaluation Memorandum, and provide other analytical and programmatic support.

1.2 Background

In January 2011, the Independent Scientific Review Panel expressed concern that RME and project development in the LCRE did not appear to be well-coordinated or well-organized. This situation is cause for concern especially as it pertains to comprehensive reporting requirements of Federal Columbia River Power System (FCRPS) Biological Opinion (BiOp; NMFS 2008). The proposed study is intended to provide an organizational system (a geodatabase) to store past and future data, facilitate data sharing among research and restoration practitioners, and be used as the basis for synthesis and evaluation of data

[1] As used herein, the "data model" is the blueprint or design for structuring a database and the "database" is the structure implemented with real data. We call the LCRE database *Oncor*.
[2] The 2012 work products from the Analysis and Synthesis objective were delivered separately to the Corps and are not included in this report. For copies of the documents, contact C Studebaker (503 808 4788).

in the LCRE. The database is being developed in form and function to relate to other relevant regional data systems (e.g., cbfish.org) and will provide a web-based, publicly accessible, interactive map-centered interface for current and future comprehensive analysis in the LCRE. The subject database will ultimately allow regional managers and stakeholders, such as the Bonneville Power Administration (BPA), the Corps, Lower Columbia River Estuary Partnership (EP), National Marine Fisheries Service (NMFS), Northwest Power and Conservation Council, Oregon Department of Fish and Wildlife, and Washington Department of Fish and Wildlife (WDFW), to adaptively manage and collaborate on RME and habitat restoration project development. In fact, LCRE restoration sponsors committed to uploading data to *Oncor* in their project proposals submitted in February 2013 to BPA as part of the Northwest Power and Conservation Council's geographic review (for more information, see http://www.nwcouncil.org/fw/reviews/ 2014/geographic-review/).

Numerous efforts are under way to increase the survival of Endangered Species Act-listed salmonid stocks in the Columbia River basin. The Reasonable and Prudent Alternatives (RPAs) of the 2008 FCRPS BiOp specifically identify habitat restoration and associated RME in the LCRE as actions that can strongly support this cause. Accordingly, the BPA, Corps, and NMFS have designed and are conducting an extensive RME program related to habitat restoration in the LCRE. The BiOp RME Workgroup Recommendations report (May 2010, available at http://www.salmonrecovery.gov/ Evaluation/Reports.aspx) identified a number of gaps in coverage of the 2008 FCRPS BiOp, including the need for increased action effectiveness research and comprehensive summaries (roll-ups) and evaluations of estuary RME to inform adaptive management of the habitat restoration effort. Our project is helping to address these gaps by developing a geospatial database and instituting it with the existing collaborative, adaptive management process in the CEERP. This process functions to develop, evaluate, adapt, and implement tools to assess and integrate the action effectiveness monitoring and research (AEMR) of LCRE habitat restoration projects.

The LCRE is a 235-km region of the main-stem Columbia River and its floodplain, below Bonneville Lock and Dam and above the entrance to the river at the Pacific Ocean, which does not include the major associated tributary habitats outside of the floodplain. Diking and a more than 40% reduction in flow during the spring freshet (May–July) has reduced the shallow water habitat area available to juvenile salmonids in the LCRE by approximately 62% according to modeled estimates (Kukulka and Jay 2003a, b). Thus, the reconnection of lateral floodplain habitats with the main-stem river by breaching dikes and removing/replacing culverts and tide gates is an important element of landscape-scale restoration programs currently under way on the river (Johnson et al. 2008).

Like most large river floodplain landscapes or "riverscapes" (Wiens 2002), the LCRE is an exceedingly complex region to evaluate by any single measure, and particularly so because of the oceanic influence, which has variable effects depending on the season and on the lateral and longitudinal location of a given site. Regarding habitat types, for instance, the positions of four general tidal wetland vegetation habitat classes—Sitka spruce swamps, riparian woodlands, shrub-scrub, and emergent marsh—vary with changing hydrogeomorphic conditions from Bonneville Dam to the river's mouth (Borde et al. 2011, 2012). At the site scale, plant communities whether in reference condition or during restoration also vary with microtopography (Diefenderfer et al. 2008). The influence of controlling factors such as large woody debris on pool habitat development ranges from considerable in Sitka spruce swamps to perhaps nonexistent in some other plant community types or restoring areas (Diefenderfer and Montgomery 2009).

Anadromous fishes in the LCRE present a complicated situation for managers, because all species, all stocks, and all Evolutionarily Significant Units must pass through the estuary and their estuarine habitat use varies with both biological and environmental factors (Bottom et al. 2005). Furthermore, the survival and physiological condition of juvenile and adult fish collected in the estuary are affected by environmental and anthropogenic factors from the entire life cycle including conditions in the tributaries, main stem, and ocean, confounding attempts at direct cause-and-effect assessment of estuarine habitat influence and necessitating alternative assessment methods (Diefenderfer et al. 2011). In 2012, the Expert Regional Technical Group (ERTG; established under RPA 37) modified the survival benefit estimator described in the FCRPS BiOp (NMFS 2008) to include consideration of ecological relationships for the purpose of restoration project prioritization, but the estimator remains limited by the lack of information about fundamental ecological processes in the LCRE and their effects on salmonid survival (R. Thom, ERTG member, personal communication). Nevertheless, numerous studies on the West Coast of United States and Canada have shown the importance estuarine habitats play in the life histories of some salmonid stocks. Research on salmon distribution patterns in the LCRE, as well as other West Coast estuarine systems, indicates that diverse stocks of subyearling and yearling salmonids use tidal freshwater floodplain and estuarine shallow water habitats (e.g., Reimers and Loeffel 1967; Healey 1980; Levy and Northcote 1982; Shreffler et al. 1990, 1992; Levings et al. 1991; Levings 1994; Sommer et al. 2001; Tanner et al. 2002; Bottom et al. 2008). The FCRPS 2008 and 2010 BiOps call for an extensive habitat restoration program in the LCRE that is currently under way and sponsored in large part by the Corps and BPA.

Habitat restoration and associated RME in the LCRE are being carried out by multiple agencies and entities. However, standard habitat restoration monitoring protocols did not begin to be adopted in the LCRE until 2009 (Roegner et al. 2009), although many individuals working in the region had collaborated on the development of those protocols over the four preceding years and incorporated some elements of them into project-level monitoring. The Roegner Protocols provide a means of reducing barriers to Action Agencies (BPA and Corps) and managers' informed decision-making about restoration actions by coordinating and systematizing future monitoring efforts. Furthermore, data from the AEMR and monitoring are not currently structured using a standard nor are they housed in centrally accessible databases; in some cases, data are not currently available in electronic form. Thus, data integration, assessment, evaluation, and synthesis for BiOps 2013 and 2016 comprehensive reporting pose a significant scientific and organizational challenge, which also must be met if findings are to be applied in adaptive management and restoration prioritization at a programmatic level.

In this study, state-of-the-art geospatial database technology is being applied and implemented within the CEERP adaptive management process (Figure 1.1). The *Oncor* database will support annual CEERP reports, which in turn are used to inform comprehensive BiOp reporting in 2013 and 2016, and support ongoing adaptive management in the LCRE, in accordance with the CEERP process.

Figure 1.1. CEERP adaptive management process. Brown and blue boxes signify adaptive management phases and deliverables, respectively. (From BPA/Corps 2012.)

1.3 Approach and Philosophies

This is a tools-development project to support the Action Agencies' implementation of LCRE ecosystem restoration under CEERP as called for in the 2008 FCRPS BiOp. The study was originally planned as a finite, 3-year effort, assuming receipt of full annual funding. The first year (fiscal year 2012 [FY12]) entailed extensive coordination, database schema/*Oncor* data model and prototype database development, and synthesis and evaluation via the CEERP 2012 Strategy Report and Action Plan. The second and third years will involve deployment and refinement of the *Oncor* data model and database, and development and testing of database tools, analytical tools, and the user interface. The close-out process will involve training and hand-off of the database and analytical tool set to a designated regional entity for long-term support and continued data development. The intent is to create a "living" database and transfer this technology to a regional entity for long-term stewardship. Throughout the life of this project, there will be involvement, coordination, and input from stakeholders, practitioners, database experts, spatial analysts, environmental and fisheries scientists, and end-users. A strength of our approach is the integration of LCRE restoration ecology and science with state-of-the-art geospatial database engineering.

During the inaugural project year (2012), we developed "project philosophies" to guide and define the *Oncor* development effort, including the following:

- Be responsive to the CEERP objectives; i.e., Program Objectives → Analysis Questions → Data Model → Database (Figure 1.2).
- Be responsive to specific use cases; i.e., Use Cases → Analysis Questions → Data Model → Database.
- Develop *Oncor* with a circular workflow allowing multiple modes of entry into the database.

- Be responsive to the data needs of the LCRE regional stakeholders.
- Shift the paradigm from project-specific data to category-specific data.
- Build the most adaptable and scalable data model possible; i.e., make it robust and flexible enough to allow for additional and diverse types and forms of data.
- Allow for the LCRE RME data to be accessed and applied in a meaningful, productive way by various types of users.
- Disseminate data as much as reasonably possible, while protecting data integrity, data pedigree, and ensuring proper use.
- Provide an organized, documented archiving system for users to download data as appropriate.
- Develop methods and tools to allow easy integration into *Oncor* and minimal impact on data generators.
- Coordinate and collaborate with other regional database efforts, e.g., Pacific Northwest Aquatic Monitoring Partnership (PNAMP).
- Produce incremental rollouts of *Oncor* to show progress and solicit feedback.
- Plan for technology transfer to a regional entity at closure of the PNNL-development effort.

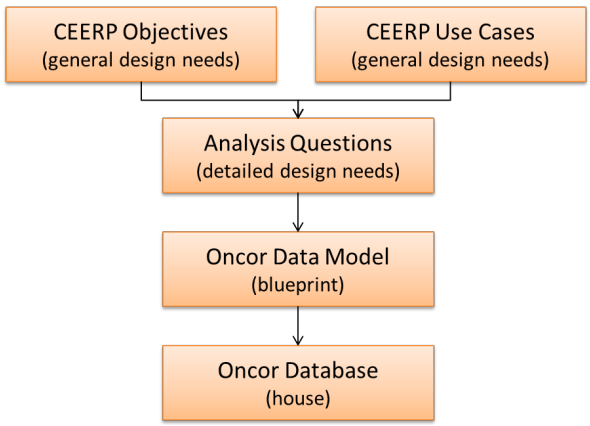


Figure 1.2. Conceptual diagram of the *Oncor* database development effort, using a house as an analogy.

1.4 Report Contents and Organization

This report covers coordination (Objective 1) and database development (Objective 2). CEERP-related work (Objective 3) was delivered to the Corps separately. In the ensuing sections, we present coordination (Section 2.0) and database development (Section 3.0). Material in Sections 2.0 and 3.0 is organized by the tasks for each objective. The report closes with accomplishments and recommendations in Section 4.0 and references in Section 5.0. The appendices contain a detailed outline of "Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program" (Appendix A) and a draft of the data reduction procedure for water-surface elevation (Appendix B).

2.0 Coordination

Coordination among CEERP funding agencies and regional stakeholders to establish key analysis questions and database needs for RME and ecosystem restoration within CEERP's adaptive management framework involved four tasks: 1) conceptual model updates, 2) regional coordination, 3) inventory of existing databases, and 4) development of data access and use guidelines.

2.1 Conceptual Model Updates

The purpose of this task was to prepare for possible future work (2013) to update underlying data, information, and citations in the LCRE conceptual ecosystem model, especially the portion dealing with the Salmon Ecosystem Habitat Index (SEHI) effort under the Salmonid Benefits project (EST-P-09-01) (Diefenderfer et al. 2012).

In one easily navigated electronic tool, the web-based version of the LCRE conceptual ecosystem model that Thom et al. (2004) developed brings together the information provided by existing models of subcomponents of the estuary and the state of the science and knowledge of general estuarine controlling factors, stressors, structures, processes, and functions *as of 2004*. It provides a basis for and structure in which knowledge about the LCRE can be incorporated through updates to a spreadsheet, as the knowledge becomes available, such as new information from Bottom et al. (2008), Diefenderfer et al. (2008), Diefenderfer and Montgomery (2009), Roegner et al. (2008, 2010), Sather et al. (2011), Storch (2011), Storch and Sather (2011), and Thom et al. (2013). The existing conceptual ecosystem model of the LCRE is useful, but needs to be reviewed structurally and updated with new empirical data. For the LCRE conceptual model to maintain currency as a tool, a data model and database system need to be implemented (see Objective 2) to allow the conceptual model to be easily updated whenever relevant data are published in the future. Finally, with the framework in place, the design principles should be refined to accommodate 1) relationship "discovery," and 2) display feature updates indicating linkages and at what level(s) linkages are accessible. The LCRE conceptual ecosystem model (current version) was used to inform the *Oncor* data model for the geospatial database system described below. The 2012 effort for this task involved consultation with scientists working on the Salmonid Benefits project and preparing to update the LCRE conceptual model by identifying 1) measurements and metrics for the SEHI, and 2) improvements and steps to take to make a new version of the LCRE conceptual model.

The SEHI, a subset of the broader LCRE ecosystem conceptual model, includes the measurements and metrics depicted in Figure 2.1, which are included in the *Oncor* data model. The Corps intends to apply SEHI during the planning and prioritization process for restoration actions. Accordingly, monitoring practitioners will collect, reduce, and upload the appropriate data to *Oncor* for eventual application in SEHI. The LCRE ecosystem conceptual model will be updated to include all SEHI measurements and metrics.

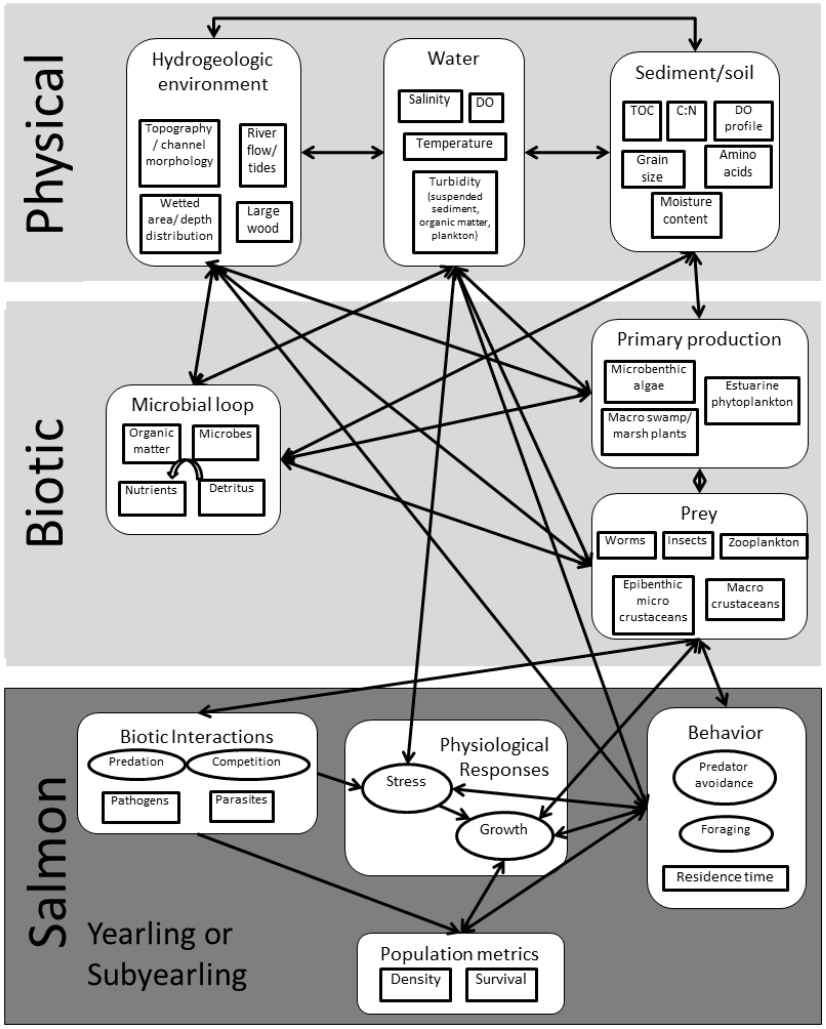


Figure 2.1. Conceptual model for SEHI. (Obtained from Diefenderfer et al. 2012.)

Based on experience with the LCRE and SEHI conceptual models, we offer the following list of possible improvements concerning the LCRE conceptual ecosystem model by Thom et al. (2004). First, use improved mapping data of the estuary; make interactive maps available for users searching for information from specific areas/reaches; and use the new ecosystem classification system for the LCRE (Simenstad et al. 2011). Second, update and populate the model with new data and information that have become available since 2004, including ecological relationships such as the following from Diefenderfer et al. (2012):

- juvenile salmon presence in restoring wetlands as a function of water temperature
- opportunity for juvenile salmon to access a restoring wetland as a function of physical wetted area in combination with biological migration timing
- mass of particulate organic matter produced at a restoring site as a function of the size of the restoring site, the amount of biomass lost over winter per unit area, and an empirically derived constant

- proportion of the mass of particulate organic matter produced at a restoring site that reaches the main-stem river/estuary as a function of distance in kilometers from the site to the main stem
- plant community composition as a function of salinity, land elevation, and inundation
- rate of change of the land elevation at tidal wetland sites as a function of sediment accretion rate
- cross-sectional area at the outlet and edge length of tidal wetland channels as a linear function of catchment area
- number of pools in tidal swamps as a function of the number of log jams.

2.2 Regional Coordination

Regional coordination was a major focus of the *Oncor* project during FY12. Three coordination "avenues" were established in FY12 (Table 2.1) involving funding agencies and regional stakeholders to coordinate database needs for RME and ecosystem restoration in the LCRE within the CEERP's adaptive management framework. The EP Science Work Group provided overall coordination (Avenue A). In managing the coordination task the Work Group reaches out to database and web technologists responsible for technical coordination and logistics and working on similar efforts elsewhere in the Columbia River basin (Avenue B), many of which are funded by BPA on projects conducted under federal BiOp RME and salmon recovery. The third coordination avenue (C), involving the Corps, BPA, and EP, provides for programmatic coordination, which was critical because of each party's respective database work for the LCRE.

Table 2.1. Regional database coordination avenues.

Avenue	Composition	Frequency	Purpose
A	Regional, general: EP Science Work Group	Quarterly	Awareness, feedback
B	Regional, data technologists	Quarterly	Technical coordination, logistics
C	Corps/BPA/EP	Monthly	Programmatic coordination

During CY 2012, nine regional coordination meetings or calls were conducted (Table 2.2). During two Avenue A meetings, we introduced the intent of the *Oncor* database project, presented the *Oncor* data model, and demonstrated a preliminary prototype of the web-based interface accessing the *Oncor* database. We solicited and received feedback on database content and functionality. Two Avenue B meetings started the coordination and communication process among data technologists for BPA, EP, and PNNL. The most active coordination avenue was Avenue C, which involved five calls and meetings during 2012. This avenue is important because BPA is funding multiple database development efforts in support of the Northwest Power and Conservation Council's Columbia Basin Fish and Wildlife Program. Avenue C was a primary coordination mechanism between BPA- and Corps-funded RME database efforts.

Table 2.2. Discussion points of regional coordination meetings/calls for *Oncor* during 2012. (Contact G. Johnson [PNNL] for copies of meeting notes.)

Meeting	Date	Avenue	Participating Agencies
1	**May 5**	**C**	**BPA/Corps/PNNL**
– Introduced the goal and objectives of the Corps' *Oncor* project to BPA. – Explained BPA's intent for the estuary component of cbfish.org – site-level information on restoration actions in the LCRE. – Understood that *Oncor* and cbfish-estuary development are just beginning; it is a good time for coordination. – Agreed that Avenue C meetings will be useful; e.g., all need a clear understanding of who's providing web services for what.			
2	**Jun 25**	**C**	**BPA/Corps/EP/PNNL**
– Described database efforts for the BPA/NPCC Fish and Wildlife Program (F&WP); e.g., StreamNet provides data for synthesis and evaluation by others. – Explained that BPA is developing terms and conditions, general user agreements, and other data access and use vehicles. BPA to share with *Oncor* when the materials are completed. – Discussed the need for standard data entry templates.			
3	**Jun 26**	**A**	**BPA/Corps/EP/PNNL/Others**
– Presented *Oncor* at a regular monthly meeting of the EP's Science Work Group. – Solicited initial feedback at this introductory session to inform regional stakeholders about plans for *Oncor*. – Noted important programs and associated databases for water quality (U.S. Geological Survey) and toxics (U.S. Environmental Protection Agency and others).			
4	**Aug 15**	**C**	**BPA/Corps/EP/PNNL**
– Continued coordination of development work for cbfish-estuary and *Oncor*. – Emphasized the need to understand the range of use cases and users, recognizing that development will be iterative and a work-in-progress and noted the importance of metadata. – Recognized the need to clarify and establish the respective roles of cbfish-estuary and *Oncor*; need to be clear about the system of record for various data sets. – Stated that cbfish-estuary will be the system of record for restoration actions (projects) for CEERP; i.e., BPA and Corps actions in the LCRE. – Discussed data sharing and access, but more needs to be done here.			
5	**Sep 5**	**B**	**BPA/Corps/EP/PNNL/Sitka**
– Reviewed broadsheets for cbfish-estuary prepared by Sitka. – Discussed systems of record for ERTG documents; possibilities include *Oncor*, EP database, or cbfish-estuary. – Noted the single sign-on technology, called Keystone, Sitka has developed for cbfish.org.			
6	**Sep 7**	**B**	**EP/PNNL/Sitka**
– Explained that BPA's Pisces software program collects project data and Taurus, now called cbfish, is a way to report out the data in Pisces. The system of record for BPA projects, funding, etc. is cbfish.org. – Decided to identify areas to integrate and collaborate between *Oncor* and cbfish. First area will be the Keystone single sign-on technology.			

Table 2.2. (contd)

Meeting	Date	Avenue	Participating Agencies
7	**Sep 19**	**C**	**BPA/Corps/EP/PNNL**
– Discussed need for coordination with PNAMP's various database efforts. Sitka is providing technical support to many of these efforts. – Set a goal of having data entry (collection), reduction, analysis, and a Data Exchange Template (DET) in place for the 2013 spring field season for the Corps' Multi-Scale Action Effectiveness project and the EP's Ecosystem Monitoring project. – Prepared for the presentation on *Oncor* to the EP's Science Work Group on September 25, 2012. – Discussed virtual data from databases providing web services. Will need to identify some priority databases to include as virtual data sets in *Oncor*. – Noted the need for defining terms (lexicon). (See the glossary in this report.) – Discussed use cases and analysis questions, including those for Ecosystem Monitoring, Multi-Scale Action Effectiveness, restoration planning and design, and ERTG project templates. – Reaffirmed the philosophy of making as much data publicly available as reasonably possible, while maintaining data security, integrity, and pedigree. – Had a long debate about definition of a site; more work needed here.			
8	**Sep 25**	**A**	**BPA/Corps/EP/PNNL/Sitka/Others**
– Rolled out the *Oncor* data model and the first prototype of the live *Oncor* database. The purpose of the meeting was to obtain feedback from a broad regional assemblage as represented by participants in the EP's Science Work Group. – Presented data flow—data entry, quality assurance/quality control, reduction, normalization, exchange, upload—and examples of use cases. – Solicited feedback on design, types of data, fields, uses, and potential integrations and links with other databases. We asked: "How would you use this database?" – Identified need to have one place for a user to find data for a given topic, including derived data. Multiple databases with redundant materials will be confusing and counterproductive.			
9	**Nov 21**	**C**	**BPA/Corps/EP/PNNL**
– Discussed use cases. Noted need for a "program manager" use case that would focus on reporting ecological accomplishments to the region and federal court. This would need to be coordinated with the EP's ecosystem indicator initiative. The intent would be to report LCRE ecosystem status to program managers. – Noted that BPA is drafting data-sharing agreements and data-management plans; and will make these available to *Oncor*. – Explained the "Monitoring Explorer" and the "Sample Manager Design" tools that Sitka is developing for PNAMP.			

In particular, we coordinated with the database efforts funded by BPA under the PNAMP umbrella. We worked to adopt PNAMP structures and terminology (such as https://www.monitoringresources.org/Resources/Glossary/). Draft guidelines for data management and sharing (see Section 2.4) were informed by similar material from PNAMP (R. Scranton, personal communication 2/13/14). We linked to monitoringmethods.org, as appropriate. We identified and studied existing PNAMP databases; e.g., Coordinated Habitat Monitoring Program (CHaMP; see Section 2.3). Lastly, we used the example for a Data Exchange Template (DET) set by PNAMP for natural origin spawner abundance.

2.3 Inventory of Existing Databases

In this section we provide an inventory of existing, publicly available databases and determine which might be included in the *Oncor* database as virtual (remotely accessible) or managed data (internal to

Oncor database), or conveniently linked to inform the interested user. The purpose of the inventory is to identify databases that will be useful to *Oncor*'s mission to provide RME data relevant to the CEERP while avoiding duplication of effort.

We used known databases and input from the coordination groups (Avenues A and C) to identify existing databases applicable to *Oncor*. These databases are or will be incorporated into *Oncor* as virtual data (e.g., stream flow data), managed data (e.g., Columbia River Estuary Ecosystem Classification), or as links to data to inform the interested user (e.g., Data Access in Real Time) (Table 2.3).

Table 2.3. Inventory of databases especially relevant to CEERP and relationship to *Oncor*. "Link to inform" means there is a link to leave the *Oncor* website for the user to go access a given database. "Unmanaged" means *Oncor* is retrieving data from a remote site through a web service. "Managed" means the data set is stored locally in the *Oncor* database.

Database	Steward	Web Link	Description	Relationship to *Oncor*
cbfish	BPA/Sitka	www.cbfish.org	Project and portfolio data for the F&WP	Link to inform/ Unmanaged
CHaMP	BPA/Sitka	www.champmonitoring.org	Data from CHaMP monitoring in Intensively Monitored Watersheds	Link to inform
Col R Estuary Ecosystem Classification	USGS/EP	www.estuarypartnership.org/columbia-river-estuary-ecosystem-classification	Col R Estuary Ecosystem Classification, Level 1-5. Links to USGS geographic information system data	Managed (pending permission)
Col R Estuary Land Cover	EP	www.estuarypartnership.org/lower-columbia-river-land-cover	2010 land cover data for the LCRE	Managed (w/ permission)
CO-OPS ODIN	NOAA	tidesandcurrents.noaa.gov	Tide data	Unmanaged
DART	University of Washington	www.cbr.washington.edu/dart	Umbrella website for data from various sources concerning Columbia Basin salmonid, environmental, operational, riverine, ocean and climatic measurements.	Link to inform
Diking layer	EP	NA	Diking maps	Managed (pending permission)
Elevation	Corps	NA	Merged topography and bathymetry data (2009 release)	Managed
FPC	FPC	www.fpc.org	Data on smolt monitoring, adult returns, hatchery releases, spawning, etc.	Link to inform
H- and T-sheets	Univ. of Washington - Wetland Ecosystem Team	https://catalyst.uw.edu/workspace/wet/14965/82924	Historical hydrographic and topographic maps	Managed (pending permission)

Table 2.3. (contd)

Database	Steward	Web Link	Description	Relationship to *Oncor*
Historical imagery	Corps/USGS/ Others	Various http://www.nwp.usace.army.mil/Library/Aerialphotos,realestateandeGIS.aspx https://fortress.wa.gov/ecy/coastalatlas/tools/ShorePhotos.aspx	Historical images and aerial photo's from LCRE	Managed
LOBO Viz	OHSU	http://columbia.loboviz.com/	Land Ocean Biogeochemical Observatory; real-time water quality data	Link to inform/ Unmanaged
Mon. Methods	PNAMP	www.monitoringmethods.org	Monitoring methods and protocols	Link to inform
NWIS	USGS	http://waterdata.usgs.gov/nwis	National Water Information System; streamflow, stage and temperature data	Unmanaged
PITAGIS	PSMFC	www.ptagis.org	Data from monitoring fish tagged with electronic passive integrated transponder tags.	Link to inform
River-scape Analysis Project	University of Montana	http://rap.ntsg.umt.edu	Habitat ranking; historical and projected temperature/flow data; high-res imagery	Link to inform
SATURN	CMOP	http://www.stccmop.org/datamart/observation_network	Various	Link to inform
StreamNet	PSMFC	www.streamnet.org	Fisheries and aquatic habitat data from across the Col. R. Basin.	Link to inform
Toxics	EPA	To be determined	Toxics database	Link to inform
Virtual Col. R.	CMOP	http://www.stccmop.org/datamart/virtualcolumbiariver	Salinity/Temperature forecasts; climate atlas	Link to Inform/ Unmanaged
Washington MOA	WDFW	TBD	WDFW restoration project monitoring data	

2.4 Data Access and Use Guidelines

The purpose of this section is to present draft data use and access guidelines for *Oncor*. This material does not cover "data sharing," which we define as the step where data are transferred from a data generator (researcher) to *Oncor*. Data-sharing agreements are the responsibility of the funding agencies and will be pursued in future years as *Oncor* continues to be developed. There is a hierarchy of roles and responsibilities from data provider or generator to database administrator to end-user (Figure 2.2). Once data are deposited in *Oncor*, the following guidelines could govern data access and data use.

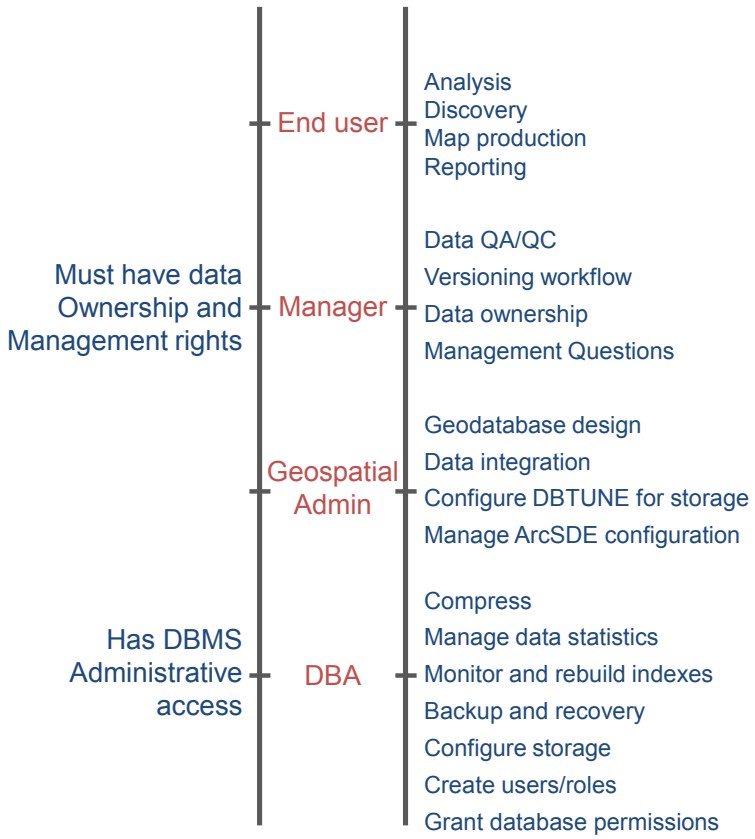


Figure 2.2. Schematic of the hierarchy of roles and responsibilities from the database administrator (DBA) to the end-user.

- A primary philosophy behind *Oncor* is to provide open, web-based access to CEERP RME and other data from the LCRE to interested users, while protecting the rights of researchers providing data. As mentioned above under philosophies, the intent is to disseminate as much data as reasonably possible while maintaining data integrity, pedigree, and appropriate data use.
- Four data access levels have been defined: Level 1 is for *Oncor* database administrators responsible for modifications and additions to the system; Level 2 is for data generators to use *Oncor* in their analyses, but the data will not be available to other users; Level 3 is for users who want to download data; and Level 4 is for users who simply want to discover and view data.
- Users for Levels 1–3 will have to create an account and log in. We are coordinating with the Sitka Technology Group to apply its Keystone single sign-on technology, thereby allowing a user to use a common username/password for the suite of websites supported under cbfish.org, champmonitoring.org, monitoringmethods.org, etc. By creating an account, the user agrees to the terms and conditions.
- Some data may be "embargoed" before they can be downloaded. Embargoing refers to restricting access to downloaded data until the data generator has published the data, but information about the data will be available along with a point of contact.

- A simple form will be provided for users to describe their purpose for downloading data. The user will need to acknowledge the Terms and Conditions.
- Data may not be published without consent of the data generator and the funding agency.
- Data downloaded from *Oncor* and used in a publication of any kind must be credited to the original data generator and access via *Oncor* acknowledged, such as, "Original data provided by [name of individual and agency] via [*Oncor link*] [date of download]."

Data-sharing, access, and use policies for *Oncor* will be determined during 2013 work efforts. The following material, modified from original draft language prepared by BPA (electronic mail from R. Scranton to G. Johnson, February 14, 2013) for its database efforts, will be considered for application to *Oncor*.

- Availability date: Federally funded data shall be made available within 1 year of data collection after quality assurance/quality control (QA/QC) processing is complete. Special cases may limit access if data are associated with particular research studies.
- Data QA/QC processing: The data provider will acknowledge the Terms and Conditions and indicate how to access the metadata records to complete the targeted data collection. The data repository will provide a link to the protocol information and indicate whom to contact to access the data. To ensure proper use of data by end-users, publication links and/or electronic copies should be provided to the database for reference and public comment. Two states of data exist: 1) data that have been through QA/QC processing and are accessible upon request, and 2) data for which QA/QC processing has not been completed so that the data are only accessible to the data provider and funding sources.
- Notification to data provider of download requests: Yes, send email with consumer contact information to data provider; No, notification not required.
- Provider-determined access: The provider grants free access to data; or the provider grants limited access to the funder and people who agree to sharing agreements.
- Data download requirements and user access settings for end-user: Terms and Conditions for open access to information and data will be determined.
- Non-compete and sharing clause: After downloading the data, the data user agrees not to share the data with others or to seek publication based on analysis without consent of the data provider or funder.
- Special cases for limited access to raw may be granted on a case-by-case basis: For example, data need to remain restricted if they protect species from potential poachers, or if landowner agreements preclude sharing raw data or culturally sensitive information.

3.0 Database Development

Elements for *Oncor* database development include analysis questions and use cases, data reduction procedures, the DET, a data model, and a prototype database and user interface. Analysis questions are relevant to the CEERP objectives and consist of key questions that *Oncor* could be used to address during data analysis. Use cases are examples of the application of *Oncor* data for specific purposes. Use cases can also include specific analysis questions asked of the database. Data reduction procedures refer to the steps the data provider takes from raw, field-collected data to data ready for upload to *Oncor*. The DET is the format and information required for uploading data to *Oncor*. This template is the culmination of data reduction procedures. The data model is the architecture for the database. The prototype database and user interface are the hardware and software that make up *Oncor*.

3.1 Analysis Questions and Example Use Cases

The *Oncor* database development effort began with the identification of analysis questions derived directly from the three main CEERP objectives (Table 3.1). The intent was to keep the data development efforts focused, and thus have CEERP program and management needs drive development of the database. To further ensure the relevance and integrity of *Oncor*, we identified analysis questions for three typical use cases (Table 3.2). Additional work on use cases is ongoing. In response to the analysis questions and use cases, *Oncor* data sets and monitored indicators were identified and built into the data model (Section 3.4) and prototype database (Section 3.5). The analysis questions and example use cases are used to test *Oncor*, thus verifying the data model is structured appropriately to address core needs.

Table 3.1. CEERP objectives with corresponding analysis questions and *Oncor* data.

	Obj. 1 Habitat Capacity	Obj. 2 Habitat Access	Obj. 3 Ecosystem Function
Analysis Questions	Increase the capacity and quality of estuarine and tidal-fluvial ecosystems	Increase the opportunity for access by aquatic organisms to and for export of materials from shallow water habitats	Improve realized ecosystem functions
At the site scale, is a restoration activity in the LCRE resulting in increased *[MONITORED INDICATOR*] compared to a control site, or a positive trajectory of [*MONITORED INDICATOR*] toward that at a reference site?	[*MONITORED INDICATOR*] = reduction in mean water temperature, dissolved oxygen, channel cross section, sediment accretion, prey production, macro-detritus export, nutrient export, percent plant cover	[*MONITORED INDICATOR*] = salmon presence/absence, salmon density, fish community structure—richness, fish community structure—species diversity	[*MONITORED INDICATOR*] = growth rate, fish condition, total realized habitat use

Table 3.1. (contd)

	Obj. 1 Habitat Capacity	Obj. 2 Habitat Access	Obj. 3 Ecosystem Function
At estuary-wide or landscape scales, are cumulative restoration activities in the LCRE resulting in increased [MONITORED INDICATOR] compared to 2000 levels?	[MONITORED INDICATOR] = percentage of cover for native plant species, net ecosystem improvement	[MONITORED INDICATOR] = habitat connectivity, total floodplain wetted area, total physical habitat opportunity, total realized habitat opportunity	[MONITORED INDICATOR] = mean survival rate, early life history diversity, genetic stock diversity, mean growth rate, total realized habitat utilization
Action effectiveness questions (examples)	Is planted vegetation increasing in similarity index relative to reference sites? Is water temperature cooler, prey production higher, etc. after restoration than before? Has reed canary grass been eliminated? Is channel cross-section self-maintaining?	What is the pattern in water-surface elevation compared to the reference condition? How muted are the tides or water exchange? Are juvenile salmon able to access the site? Compared to the reference conditions, when are they present, which species, length distribution, densities, or catch rates?	Compared to the reference conditions, what are fish eating at the restored site? Are the fish growing? What are the flux rates of organic matter and prey to the main stem? Has life history diversity increased?

Table 3.2. Example use cases.

Use Case 1 – ERTG Project Templates

Application – Development of ERTG project templates

Anticipated Oncor users – Corps, non-governmental organizations (NGOs), tribes, architect and engineer (A&E) firms

Data sensitivities – none known at this time

Data steward/owner – ??

Analysis questions and information needs:

- What are the physical conditions at a proposed restoration site or location(s) within a 5-km radius?
  - Mean, minimum, and maximum tidal range
  - Ordinary-high-water tide elevation
  - Extreme-high-water elevation
  - Two-year flood elevation
  - Mean, minimum, and maximum salinity
  - Mean, minimum, and maximum water temperature
- What are the vegetation conditions at a proposed restoration site or location(s) within a 5-km radius?
  - Species composition by percent cover
  - Percentage invasive species by percent cover
  - Map
- What are the fish characteristics at a proposed restoration site or location(s) within a 5-km radius?
  - Species composition, including Endangered Species Act-listed species
  - Salmonid densities, including species and stocks from upriver
  - Salmon populations in watersheds nearby

Metrics and datasets supporting the use case − xxx

Canned data reduction and analysis pieces − xxx

Table 3.2. (contd)

Use Case 2 – Monitoring and Research for Vegetation and Vegetation-Elevation

Application –Restoration project design and engineering; ERTG project templates; research ecology; AEMR
Anticipated Oncor users – Corps, federal and state agencies, NGOs, tribes, A&E firms
Data sensitivities – none known at this time
Data steward/owner – currently PNNL
Analysis questions

- Herbaceous Plant Species and Elevation
  - What is the average percent cover of species s, or all species, or species {1,2,3…s} at site n? At all sites {1,2,3….n}? At site n in year(s) y? At all sites {1,2,3….n}in year(s) y? At sites with attribute(s) a? [e.g. cover type]. In zone(s) z? Plot bar graph of results (applicable to any of the above with multiple x or multiple y or both).
  - How many native plant species have been observed at site n? At all sites {1,2,3….n}? In zone(s) z? At river kilometer k? What are their Latin names? What are their common names? Plot bar graph of results (applicable to any of the above with multiple x or multiple y or both).
  - How many non-native plant species have been observed at site n? At all sites {1,2,3….n}? In zone(s) z? At river kilometer k? What are their Latin names? What are their common names? Plot bar graph of results (applicable to any of the above with multiple x or multiple y or both).
  - What are the minimum and maximum elevations at which species s occurs at site n? At all sites {1,2,3….n}? At sites with attribute(s) a [e.g., cover type]? In zone(s) z? At river kilometer k? Plot bar graph of results with species on the x-axis and selected y-axis (elevation, average percent cover, or both).
  - What is the average elevation of the vegetation survey area at site n? At all sites {1,2,3….n}? At sites with attribute(s) a [e.g., cover type]? In zone(s) z?
  - List the plant species that have been observed at river kilometer k. At site n? At all sites {1,2,3….n}? At sites with attribute(s) a [e.g., cover type]? In zone(s) z? In the LCRE?
  - What is the average percent cover of the x most dominant herbaceous plant species at restoration site n in year(s) y?
  - What are the Relative Sorensen Similarity Indexes for herbaceous vegetation species in year(s) y at the set of sites {1,2,3…n}? Plot matrix with labels for year and site attribute reference or restoration.
  - How have the Relative Sorensen Similarity Indexes for herbaceous vegetation species changed in the y years since the restoration action was implemented at site n?
  - Plot the species-area curves and Bray-Curtis (Sorensen) distance curves for the set of quadrats collected at site n or the set of sites {1,2,3…n}.
- Shrubs and Elevation
  - What is the density (stems per hectare) of species s at site n or at the set of sites {1,2,3…n}?
  - What is the relative density percent of species s at site n or at the set of sites {1,2,3…n}?
  - What is the frequency (number of plots) of species s at site n or at the set of sites {1,2,3…n}?
  - What is the relative frequency percent of species s at site n or at the set of sites {1,2,3…n}?
  - What is the mean elevation of species s at site n or at the set of sites {1,2,3…n}?
- Trees and Elevation
  - What is the density (stems per hectare) of species s at site n or at the set of sites {1,2,3…n}?
  - What is the relative density percent of species s at site n or at the set of sites {1,2,3…n}?
  - What is the frequency (number of plots) of species s at site n or at the set of sites {1,2,3…n}?
  - What is the relative frequency percent of species s at site n or at the set of sites {1,2,3…n}?
  - What is the dominance of species s at site n or at the set of sites {1,2,3…n}?
  - What is the relative dominance percent of species s at site n or at the set of sites {1,2,3…n}?
  - What re the minimum, maximum, mean, and median diameter at breast height of species s at site n or at the set of sites {1,2,3…n}?
  - What is the mean elevation of species s at site n or at the set of sites {1,2,3…n}?

Metrics and datasets supporting the use case − xxx
Canned data reduction and analysis pieces − xxx

Table 3.2. (contd)

Use Case 3 – Monitoring and Research for Fish
Application – Research ecology; ERTG project templates; AEMR
Anticipated Oncor users – Corps, federal and state agencies, NGOs, tribes, A&E firms
Data sensitivities – none known at this time
Data steward/owner – currently PNNL
Analysis questions

- Physical Conditions
  - What are the monthly, site-specific water temperature and dissolved oxygen levels in a given study area?
  - What is the monthly mean water temperature across all sites?
- Fish Community and Density
  - What is the percent of total catch for fish species over a given time period and study area and/or site?
  - What are the seasonal mean densities for salmon, non-native, and other native (excluding salmon) taxa in a given study area over a given time period? What is the standard error of the means?
  - What species composed greater than 1% of the total catch in a given study area (or site) over a given time period?
  - What is the monthly mean density of the most abundant species (as determined from the proceeding question)? What are the standard errors associated with the means?
  - What is the site specific catch per unit effort (CPUE; # of fish/beach seine haul) over a given time period?
  - What are the site-specific mean densities of selected fish species over a given year, month, and/or season?
- Data Summaries
  - What are the summary statistics associated with sizes of selected fish species over a given time period in a particular study area?
  - What is the monthly mean length for selected species within a given study area? What are the standard errors of the means?
  - What are the seasonal mean densities for selected salmon species in a given study area?
  - What are the monthly mean densities of selected salmon species by year in a given study area? What is the standard error of the means?
  - What are the monthly mean fork lengths of selected salmon species by year in a given study area? What is the standard error of the means?
  - What are the seasonal length frequency distributions for a particular fish species at a given site or study area?
- Genetics
  - What is the estimated percent contribution of genetic stock group composition of unmarked Chinook salmon sampled within a given study area?
  - What is the estimated percent contribution of genetic stock group composition of marked Chinook salmon sampled within a given study area?
  - What is the monthly genetic stock proportion of unmarked Chinook salmon sampled in a given study area?
- Early Life History Diversity
  - What is the annual ELHD index within a given study area?
  - What is the seasonal ELHD index within a given study area?
  - What is the genetic stock ELHD index within a given study area?
  - What is the seasonal genetic stock ELHD index within a given study area?

Metrics and datasets supporting the use case − xxx
Canned data reduction and analysis pieces – xxx

3.2 Data Reduction Procedures

In research and monitoring, the measurements made by scientists or technicians in the field or laboratory are colloquially referred to as "raw data." This section describes the effort by the *Oncor* development team to provide detailed and efficient methods for those collecting raw data for the CEERP ("data generators") to transform data as needed to meet typical reporting requirements and load data into

Oncor. Once data are in the *Oncor* database, they will be available for the data generators and CEERP managers to compare across data collection sites, times, and monitoring programs in the LCRE.

Data reduction is defined as the process of transforming raw data by statistical or mathematical functions into a more useable form. Transformation can be as simple as changing the unit—a common example being survey feet to meters. Typically, various transformations need to be performed on restoration monitoring data, with quality control checks at each step to ensure that the original values are correctly presented. Final data used for reporting or loaded into *Oncor* should be traceable to raw data through documentation of any transformations that were made. Ultimately, data reduction should produce documentation sufficient to permit an independent data auditor to determine whether data are accurate, complete, traceable, and meet specifications, although most data collection in the CEERP is not subjected to such auditing processes. These types of procedures also help to ensure that data are not lost when staff members change at a data-generating organization, by making it easy for any staff member to open a data package and know what to expect and how it can and cannot be used.

In the LCRE, many people who monitor restoration and reference sites use the methods described by Roegner et al. (2009), a document developed by the Corps' Cumulative Effects study (study code EST-P-04-04): *Protocols for Monitoring Habitat Restoration Projects in the Lower Columbia River and Estuary*. For example, the current geographic review proposals by five "umbrella projects" in the CEERP—Columbia Land Trust, Cowlitz Indian Tribe, Columbia River Estuary Study Taskforce, EP, and Washington Estuary Memorandum of Agreement (see www.cbfish.org)—state that Roegner et al. (2009) will be the basis of monitoring in 2014. The Roegner et al. (2009) document was developed to support the CEERP and is freely available at the following URL:
http://www.nwfsc.noaa.gov/publications/displayinclude.cfm?incfile=technicalmemorandum2009.inc.
The individual methods provided by Roegner et al. (2009)—e.g., hydrology, vegetation—are also available for selection at www.monitoringmethods.org, where according to the PNAMP terminology they are termed "methods" not "protocols" (*Oncor* adopts the PNAMP terminology). However, the methods described by Roegner et al. (2009) focus on field data collection with few references to the procedures required for data reduction, analysis, and reporting, and questions have arisen about how to ensure data quality procedures and standardized, comparable results in the CEERP (see for example Borde et al. 2012).

Because of the widespread adoption of the Roegner et al. (2009) methods, the *Oncor* development team is building detailed instructions and demonstration examples to help users efficiently transform raw data collected with these methods into measurements, metrics, and indicators[1] in formats that will automatically load into *Oncor* (Figure 3.1). We are calling the activities that occur after data collection

[1] *Oncor* adopts the relevant PNAMP terminology, which is available from
https://www.monitoringmethods.org/Glossary/Definition/:
Measurement: A value resulting from a data collection event at a specific site and temporal unit.
Metric: A value resulting from the reduction or processing of measurements taken at a site and temporal unit at one or more times during the study period based on the procedures defined by the response design. Metrics can be used to estimate an indicator using an inference design. Note that a variety of metrics can be derived from original measurements.
Indicator: A value resulting from the data reduction of metrics across sites and temporal periods based on applying the procedures in the inference design. A reported value used to indicate the status, condition, or trend of a resource or ecological process; intended to answer questions posed by the objectives of the protocol. Contrast with metric. Per the inference design, metrics are combined or reduced to produce indicators.

and before *Oncor* data loading "data reduction space." Data reduction space includes data processing procedures (e.g., mathematical transformations and quality control checks), as well as several key elements that can be stored in an Excel data reduction workbook: 1) a data dictionary defining each field (i.e., column header or "attribute") of the data(Table 3.3); 2) data and associated metadata (e.g., data owner, contact, intended use, etc.) (Table 3.4); 3) example data (Table 3.5); 4) calculated data (Table 3.6 and Table 3.7); and 5) spatial data associated with the data (e.g., shape file of the location of a water surface level gauge or vegetation plot).

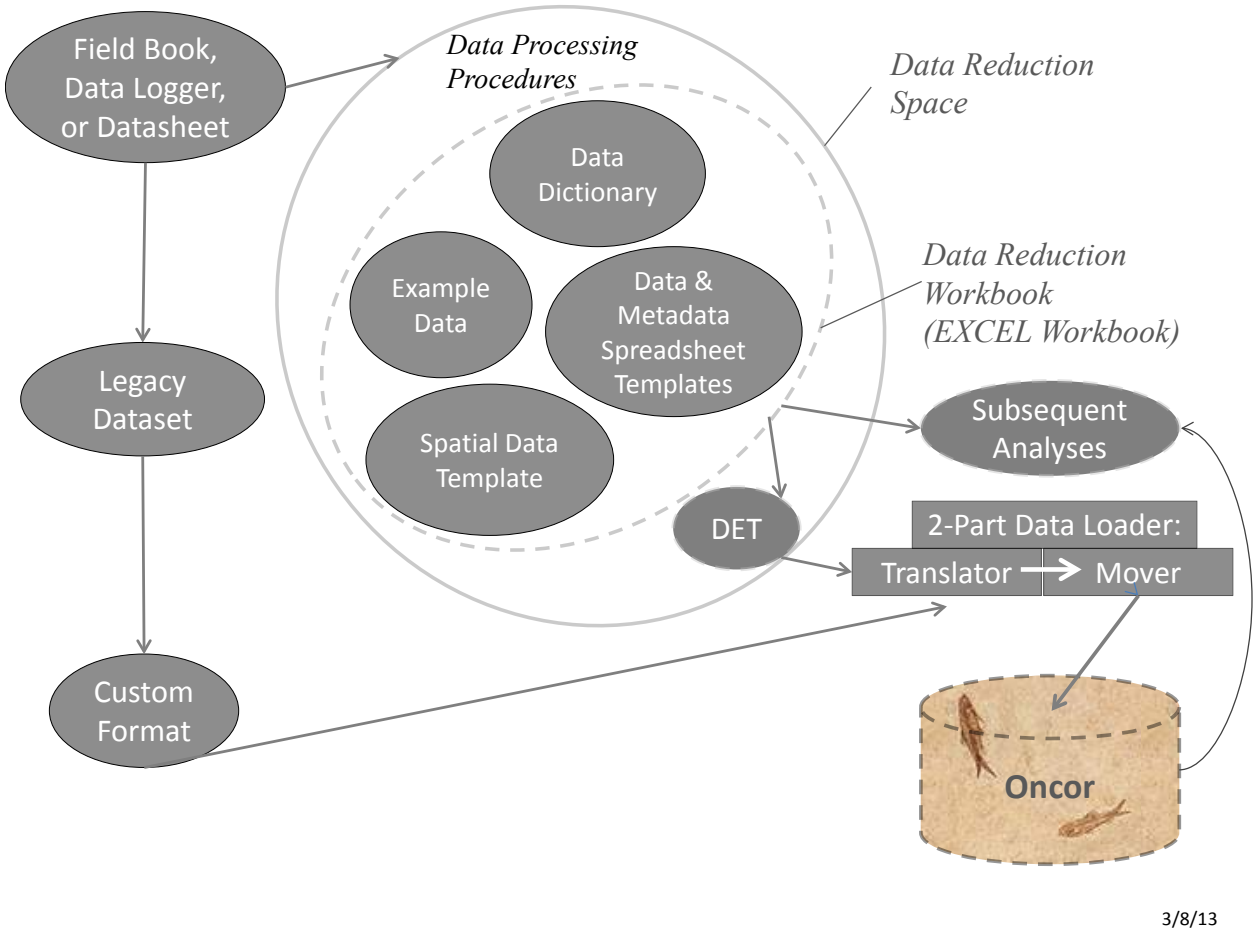


Figure 3.1. Schematic of data flow from data generator to *Oncor*. Data flow includes data reduction and entry into a Data Exchange Template (DET) by the data generator, prior to automated loading to the *Oncor* database.

Table 3.3. Example data dictionary for sediment accretion or erosion rate.

Field Name	Data Type	Req.	Validation	Description	Codes/ Conventions	Metadata/ Measurement/ Metric/ Indicator 0/1/2/3	Calculation (Y/N)	Algorithm
Site_ID	TEXT	Y	Lookup Table	Location where the sampling area is and the measurements are made		0	N	
SAMPLING_AREA_ID	TEXT	N	Lookup Table	Area within a site where measurement was made, if a large site has been divided into subareas.		0	N	
SURVEY_DAY	DATE	N		Day and month of measurement		0	N	
SURVEY_YEAR	DATE	Y		Year of measurement		0	N	
RECORDER	TEXT	N		The initials of the person who recorded the data.			N	
SS_ID	TEXT	Y		Identification code for the set of stakes, developed by the data generator.		0	N	
SAMPLE_PT_X	DOUBLE	N		The X-coordinate for the sediment accretion stakes. As an alternative, may provide the shape file.		0	N	
SAMPLE_PT_Y	DOUBLE	N		The Y-coordinate for the sediment accretion stakes. As an alternative, may provide the shape file.		0	N	
PROJECTION	TEXT	N		The projection for the shapefile containing the sediment stake point(s). An alternative to providing X-Y coordinates.		0	N	
SHAPE_SOURCE	TEXT	N		The filename of the point shapefile containing the sediment stake point(s). An alternative to providing X-Y coordinates.		0	N	
HOR_DIST	SHORT INTEGER	Y	< 100 cm		Unit is cm.	1	N	
VERT_DIST	DOUBLE	Y	< 100 cm		Unit is cm.	1	N	

Table 3.3. (contd)

Field Name	Data Type	Req.	Validation	Description	Codes/ Conventions	Metadata/ Measurement/ Metric/ Indicator 0/1/2/3	Calculation (Y/N)	Algorithm
CDIRECTION	TEXT	N	Lookup Table	A notation at the zero stake to indicate the cardinal direction of the stake relative to the other stake.	N, S, E, W, NE, NW, SE, SW	0	N	
MEAN EL	DOUBLE	N		Average of 11 points from one site at one survey date/year.		2	Y	AVERAGE (X1:X11)
SD EL	DOUBLE	N		Standard deviation of average of 11 points from one site at one survey date/year.		2	Y	STDEV (X1:X11)
SED RATE	DOUBLE	N		Rate per year, calculated as the difference between mean elevations over time measured in daily increments.	May be a positive number (=accretion) or a negative number (=erosion).	3	Y	((AVERAGE EARLIER YR-AVERAGE SUBSEQUENT YR)/(DATE SUBSEQUENT YEAR - DATE EARLIER YEAR))*365

Table 3.4. Example metadata and column override options for sediment accretion and erosion rate.

DATACATEGORY	DATASETOWNER	DATASETPROVIDER	CONTACT	CONTACTEMAIL	METHOD	PROJECTION	SHAPE_SOURCE
TEXT	TEXT	TEXT	TEXT	TEXT	TEXT	TEXT	TEXT
SedimentRate	USACE	PNNL	HL Diefenderfer	Heida.Diefenderfer@pnnl.gov	Roegner et al. (2009)	UTM10, NAVD88	CE_Sedstakes.shp

Table 3.5. Example sediment accretion or erosion rate data collected in 2 years at the same location.

Site ID	SUBAREA ID	SURVEYDATE	SURVEYYEAR	RECORDER	SS_ID	HOR_DIST	VERT_DIST	CDIRECTION
TEXT		DATE	DATE		TEXT	SHORT INTEGER	DOUBLE	TEXT
KF	GRD	7/14/2005	2005	RMT	KF-GRD	0	68.8	N
KF	GRD	7/14/2005	2005	RMT	KF-GRD	10	69.0	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	20	68.0	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	30	68.4	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	40	68.2	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	50	66.5	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	60	65.6	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	70	65.4	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	80	64.0	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	90	64.6	
KF	GRD	7/14/2005	2005	RMT	KF-GRD	100	62.0	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	0	59.4	N
KF	GRD	7/9/2009	2009	HLD	KF-GRD	10	60	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	20	60.8	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	30	58.9	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	40	59.5	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	50	57.9	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	60	59.2	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	70	58.6	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	80	58.9	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	90	56.4	
KF	GRD	7/9/2009	2009	HLD	KF-GRD	100	56.2	

Table 3.6. Example sediment accretion or erosion rate means calculated for 2 years in the same location, from data in Table 3.5.

SITE_ID	SUBAREA_ID	DATE	MEAN_EL	SD_EL
TEXT	TEXT	DATE	DOUBLE	DOUBLE
KF	GRD	7/14/2005	66.41	2.29
KF	GRD	7/9/2009	58.71	1.41

Table 3.7. Example sediment accretion rate calculated from July 14, 2005 through July 9, 2009 at one location.

SITE_ID	SUBAREA_ID	STARTDATE	ENDDATE	SED_RATE
TEXT	TEXT	DATE	DATE	DOUBLE
KF	GRD	7/14/2005	7/9/2009	1.93

The *Oncor* development team is providing a document with the working title of "Data Reduction for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program" (see outline in Appendix A). The scope of this stand-alone document is the data reduction space between field data collection and loading to *Oncor* (Figure 3.1). We expect this document to provide CEERP programmatic guidance to support quality management programs at individual organizations that collect restoration research and monitoring data in the LCRE. The two general routes by which a prospective data generator may submit data to the *Oncor* data loader are 1) a DET and 2) a custom format. Custom format loaders will be developed by PNNL only for large, legacy data sets deemed critical for inclusion in *Oncor*, and are not included in the CEERP Data Reduction Procedures document (Figure 3.1). A strength of this document will be that its authors also participated in development of the Roegner et al. (2009) data collection methods document. In fact, we plan to solicit peer review of the Data Reduction Procedures document by Dr. Curtis Roegner (NMFS).

The *Oncor* development team is also building example data reduction workbooks in Excel that correspond to each key data category described by Roegner et al. (2009): hydrology, water quality, elevation, landscape features, plant species composition and cover, and fish community. The workbooks will be available to data generators online and may be adopted at will, but are not required, because many data generators have preferred data reduction workbooks already in use. However, the DET, a simplified spreadsheet including only the final data loaded into *Oncor*, will be required for all data generated in the future. The DETs are flexible and ultimately will identify a full set of fields representing all known data collection in the LCRE that are relevant to the CEERP. DETs will include fields for each measurement, metric, and indicator that can be entered into the *Oncor* database. Data generators likely will not collect data corresponding to all fields, so DETs submitted to *Oncor* will generally contain fewer than the total available number of fields. A subset of fields for each indicator will, however, be required for data entry into *Oncor*; e.g., a plant species percent cover or beach seine haul cannot be entered without a corresponding location and date.

The data reduction procedures will be released under separate cover in a document called "Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program." A detailed outline of this document is presented in Appendix A. A draft of the data reduction procedure for water-surface elevation is presented in Appendix B.

3.3 *Oncor* Data Model

The *Oncor* system is founded on the principle of a data model, which fundamentally begins an information system lifecycle and drives the system in the types of data to be included, how the data interact with one another, how a user interacts with the data, and what kinds of questions can be asked. We developed a flexible and scalable draft data model for the LCRE to support past, current, and future data development by multiple entities and integrated connections with other regional data-management systems. This approach feeds directly into a core project philosophy, "build the most adaptable and scalable data model possible; i.e., robust and flexible enough to allow for additional and diverse types and forms of data."

A data model can be thought of as simply an architectural blueprint; it is the design for what is to be built. As is true for building a physical structure, building a model involves a lifecycle that includes needs assessment, proper planning, design review, implementation, and testing with feedback loops

integrated into the entire process. The proper design of a data model will affect the function and usability of an information system, such as *Oncor*, at all levels of its implementation, now and into the future. In addition, a well-designed data model has the specific advantages of improving unforeseen analytical capabilities, addressing multi-dimensionality and spatiotemporal system dynamics, and putting forth a template that can be shared among other entities performing similar work in other areas.

Considerations for a data model design begin with the application domain, where the core requirements of the information system are established. Defining what the system will do provides a systems analysis perspective and leads to an initial design, or a conceptual computational model, that evolves iteratively throughout the development lifecycle as requirements become more focused and/or additional functionality/needs are identified. Bringing the conceptual computational model into reality, or systems implementation, particularly in the case of *Oncor* where state-of-the-art database methods are being implemented into spatially aware databases, requires incremental roll-outs and repeated testing on a number of levels, including testing of different use cases (Section 3.1), local system performance, web-based system performance, security, and the perspectives of a range of users. A progression of complex steps is involved in the development of a data model, which will be continuously refined into FY13 to ensure that users (the public, stakeholders, managers, scientists, and analysts) can efficiently use available data through multiple means of discovery and use.

3.3.1 Background

The *Oncor* data model was developed in response to the need to bring together various types of data that are largely geospatial and spatiotemporal in nature. In cases where data were collected without the idea of bringing them into a spatial framework, the data can typically be adapted to represented locations. Upon searching for existing data models to meet the needs of habitat restoration and associated RME efforts, it was initially believed the well-designed ArcMarine (http://dusk.geo.orst.edu/djl/arcgis/) spatial data model, developed by Oregon State University, could be adapted to meet the needs in the LCRE; however, differences in requirements for the estuary data, in addition to newer spatial database technologies, ultimately led to the development of the *Oncor* data model.

Early system needs and data resources were initially defined under the Corps' *Cumulative Ecosystem Response to Restoration Projects in the Lower Columbia River and Estuary* (Johnson et al. 2012). At the onset of the current project, it became clear, and was defined as one of the project philosophies, that a transition from project-specific data to category-specific data was necessary to avoid "stove-piping" available data; similar data from multiple projects needed to be integrated to make the best use of the collected data. This notion evolved to one of bringing data together under categories, subcategories, indicators, and metrics as defined by the PNAMP data classification (see Figure 3.2), while also maintaining the metadata (data about the data) including data pedigree and intended use (also defined by a project philosophy).

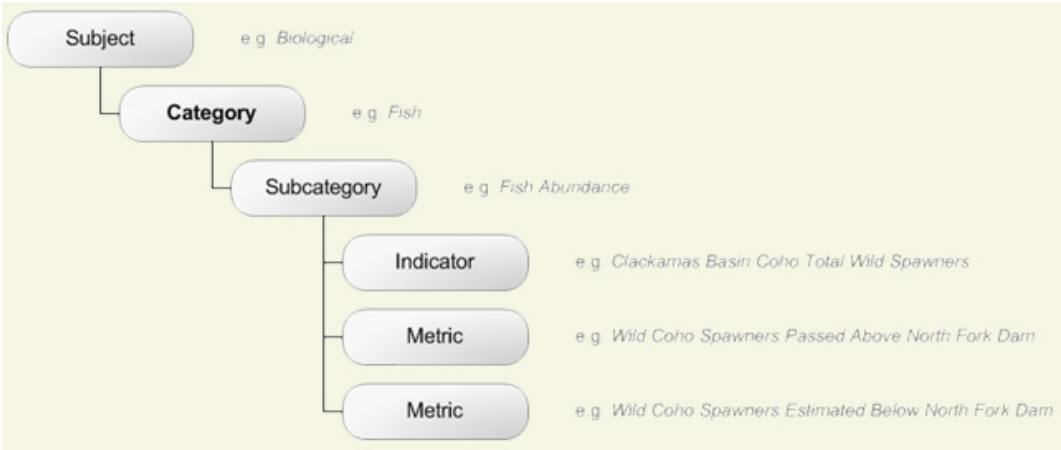


Figure 3.2. PNAMP data classification tree used to guide the *Oncor* data model into a common conceptual structure (http://monitoringmethods.org).

3.3.2 Data Model Structure

A somewhat unique aspect of the *Oncor* system is that the capture and assimilation of data are being layered through a Geospatial Relational Database Management System rather than a traditional Relational Database Management System (RDMS). The distinction is that an additional and highly functional entry point for accessing the data within the data model structure comes through the use of geography. The data model ties feature geometry (a spatial representation of a feature, measurement location, or site of interest represented as a line, point, polygon; see Figure 3.3) to a data event or series of data events at that location or area. For example, a water-surface elevation sensor can be represented in the data model as a single point with an associated X and Y (longitude, latitude) coordinate, but a series of data is tied to that single point representing water-surface elevation collected every hour for the past 5 years. An additional benefit for developing the data model around geographic principles is that spatial relationships can be established among the feature geometries included in *Oncor*. This allows a user to filter data not only by standard RDMS queries (i.e., retrieve data for a given site collected in a given time period), but also by proximal geography (i.e., retrieve data for a given time period for sites within 5 km of a restoration site, 1 km from the main-stem river, and between river miles 15 to 65). The notion of a spatial hierarchy is established within the *Oncor* data model that allows inheritance of data from multiple spatial scales. The spatial hierarchy is defined as follows, from the fine to the coarse spatial scale:

- Level 0 – Sampling Location: aspecific place where a measurement was taken
- Level 1 – Sampling Area: acollection of sampling locations or transects
- Level 2 – Site: an area defining where measurements are taken from Level 0-1
- Level 3 – Study Area: an area defining a conglomeration of sites from Level 0-3
- Level 4 – Consists of several larger spatial domains including:
  - Reach: adefined hydrogeomorphic reach within which the study area is contained
  - County: the county within which the study area is contained
  - State: the state within which the study area is contained.

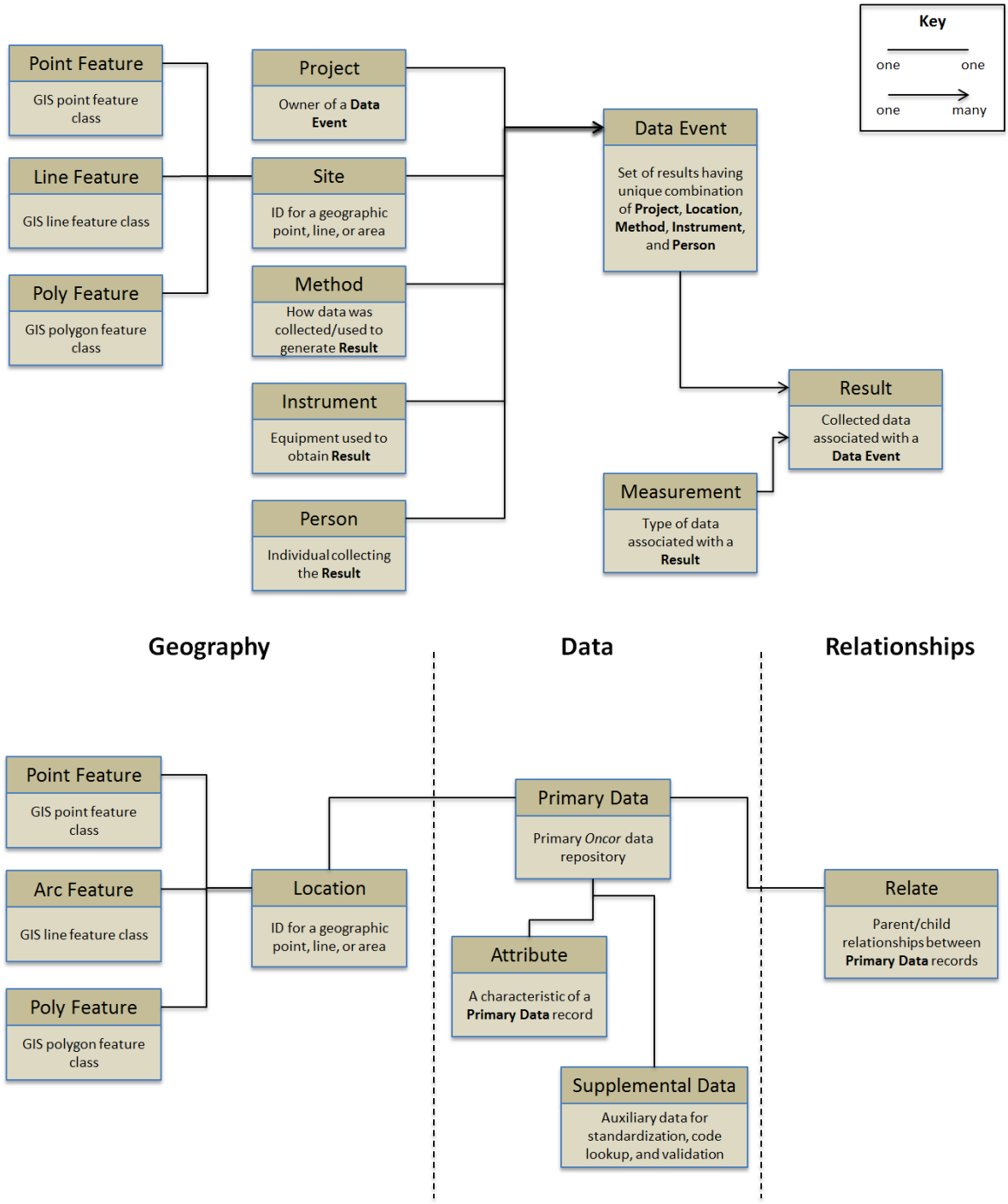


Figure 3.3. Simplified *Oncor* data model. [method box above – change "was" to "were"]

It should be noted that the progression of data from Level 0 to Level 4 does not need to be continuous, but at an absolute minimum, Level 0 needs to be established for any data set. For example, a U.S. Geological Survey (USGS) stage gage is represented as a single point that satisfies a Level 0 – Sampling Location; however, this station is not a member of a Level 1 – Sampling Area, Level 2 – Site, Level 3 – Study Area, but it can be assigned to Level 4. The spatial hierarchy provides the opportunity

for users to insert as much spatial detail as they choose to best represent the data that are being collected and the field sampling plan that has been established, at least from a spatial perspective. The users define their spatial hierarchies in the DET.

Using the established data reduction procedures ultimately leads the user to a DET for a defined category of data. A major purpose of developing the DET is to offer a standardized approach, blending the needs of individual data collectors with database requirements, for loading and making data available in *Oncor*. As is noted elsewhere in this report, the data from the DET are put through a data transformation and data loading (see Section 3.4.4) process that puts the data into the structure of the data model and consequently data are loaded into the *Oncor* database. In considering future needs and capabilities of the system, providing a standard mechanism for delivering data (via the DET) allows for data uploads by data collectors/stewards directly through the *Oncor* website—benefiting individual users by providing a standard procedure and *Oncor* administrators by allowing efficiency and operation of the system. The key to planning for this capability is the establishment of a flexible and scalable data model with ability to handle many types and forms of data.

Figure 3.3 represents a simplified version of the *Oncor* data model. The model may be divided into three parts: geography, data, and data groupings. The "geography" structure consists of geographic information system (GIS) "features," which are spatial representations of points, arcs (or lines), and polygons (arbitrary two-dimensional shapes). All data in *Oncor* are associated with one or more spatial features. An inherent characteristic of a GIS database is that relationships between locations are maintained automatically. For example, a water-surface elevation measurement is typically collected at a point location. This point is automatically associated with any polygon region that encloses it. So, if the point is located inside a study-site boundary polygon, it is automatically linked to that study site. This is one method by which *Oncor* users can compile data collected by multiple projects.

The "data" component of the *Oncor* model houses the primary scientific measurements collected in the field. To accommodate a wide variety of data categories, the *Oncor* data model must be extremely flexible. Unlike many databases, which store specific categories of data in separate tables with fixed attribute fields, *Oncor* uses one table structure to store virtually all data. This is accomplished through the concept of an "attribute." Every primary data record in *Oncor* consists of an attribute and a value. The attribute defines the kind of data that are associated with the value. For example, a water-surface elevation measurement may have attributes such as measurement date, having a date value; water-surface elevation, having a number value; and instrument name, having a string value. This system is flexible because the addition of a new attribute to store in the database does not require restructuring of the data model; the new attribute must only be added to a list of valid *Oncor* attributes. Moreover, attributes can define a virtually unlimited number of metadata records that may be very specific to a particular project or data generator. To provide structure to the data records, attributes can be grouped into logical categories, such as those suggested by PNAMP (Figure 3.3). In this manner, users can request groups of attributes that apply to a particular category or subcategory.

In addition to primary measurement data, *Oncor* includes a supplemental data repository to manage data common to all data categories. For example, all data categories include attributes that involve the specification of people, instruments used, sampling methods, etc. To ensure standardization across the database, values for these attributes are maintained in the supplemental data structure and given unique codes.

With all primary data stored in a single structure, *Oncor* uses a unique grouping paradigm to organize relationships among records. Every primary-data record is associated through a parent-child relationship to at least one other primary-data record. To continue the example of water-surface elevation data, the records for measurement date and elevation might be "children" to a "parent" record that defines the data-logger deployment event. In turn, the deployment event may be the child of a logger-network parent record. Hierarchical organization permits measurement data and metadata storage at many levels, depending on the scope. Likewise, users retrieving data can "drill down" through the hierarchy to the desired level of detail. This scheme dovetails with the PNAMP concept of measurements combining to create metrics, and metrics combining to produce indicators.

3.4 Prototype *Oncor* Database and User Interface

A primary objective of this project is to develop an estuary-wide data-management and information discovery/retrieval system for RME studies and restoration project development using a web-accessible geospatial database. The integration of analysis questions and use cases, and development of the data reduction procedures, DETs, and data model directly contribute to development of the web-based end-user interface allowing access, interaction, exploration, and analysis with the data. A prototype database and a live user interface have been developed and are discussed herein. The majority of the established project philosophies directly apply to the prototype database and user interface, and in many cases are the terminus of a project philosophy. The project philosophies described in Section 1.3 apply to the database and user interface.

A fundamental guiding philosophy applied to the database and user interface, which also stems from the data model, is that "*Oncor* is developed with a circular workflow allowing multiple modes of entry into the database." The design of *Oncor* is explicit in providing several entry points from which to access the same sets of data, thus providing a flexible mechanism for a variety of users (e.g., public, managers, restoration practitioners, etc.) who may think about data in different ways. One such way is to use *Oncor* as one would any standard database. For example, a query can be made to retrieve all fish catch records for the estuary between March and June 2008 and the user will receive a table of the fish catch records according to table definitions provided in the data dictionary of the fish monitored indicator DET. Taking this a step further, as described in Section 3.3.2, questions can be asked of the database that involve spatial rules such that a user can present a query that retrieves fish catch records between March and June 2008 for areas between river kilometers 100 through125.

3.4.1 *Oncor* System Architecture

The *Oncor* information system is based on a standard three-tier system architecture, which includes layers for client presentation, business logic, and data access. These layers are often referred to by their respective software component names as web, application, and database servers. Implementing a multi-tier system architecture is more costly and complex, but offers the best performance, security, and scalability. It is also a very common enterprise architecture, which makes it possible to migrate one or more *Oncor* components to another organization in the future.

Using an application server (middle tier) can enhance system security by eliminating any direct communication between the web server (clients) and database server. Encapsulating business logic in

a separate middle tier also protects *Oncor* from the volatility of ever-changing web technologies that often require client code upgrades or even migration.

The current version of *Oncor* implements the following commercial-based technology stack

- IIS7 (web server)
- ArcGIS Server, ArcSDE, and custom WCF SOAP services (application servers)
- Structured Query Language Server 2008 R2 (database server).

Using commercial technologies offers many immediate benefits, such as predictable integration and operation, which enables rapid deployment. However, in the future, we will be testing and migrating components to open source technologies, thus keeping long-term maintenance costs more reasonable.

3.4.2 Interactive Map-Based Interface

A major element of the *Oncor* user experience involves an interactive map-based interface. This interface provides an intuitive means to explore, access, retrieve, and analyze data (i.e., compare the use of Google Maps to find a gas station or particular type of restaurant in comparison to submitting a Structured Query Language statement to a database to find the same data). An example of using this kind of interface would be retrieving data by simply defining a box or polygon on the map for the area of interest, then filtering the data down using any combination of subject, category, subcategory, metric, indicator, and time range. In addition, simple geoprocessing functions that are found in GIS software are or are planned to be available in the user interface. For example, a profiling tool that allows the user to draw a simple or complex line and retrieve results back from an underlying data layer such as elevation (e.g., 2009 Corps Light Detection and Ranging/bathymetry data), land cover, ecosystem classification, etc. (see Figure 3.4). A user can also run similar functions by drawing a polygon and retrieving statistics about the area such as elevation range, percentage land cover, or other statistics by GIS layer. In terms of data retrieval, design plans have been made to allow a user to select an area of interest in the map interface that would allow clipping and downloading of the spatial data of interest. By offering an intuitive map-based user environment, the interface enables a user to more intuitively discover and interact with data while using the power of a relational database.

Expanding further upon the circular workflow philosophy, additional means by which data can be accessed combine the RDMS, Geospatial Relational Database Management System, and geospatial queries. To allow for ease of use, yet exercise the capabilities that *Oncor* provides, we are developing the user interface to look up, run, and export results from example use cases (see Table 3.2) as well as those for the CEERP objectives (see Table 3.1). In both cases, it will also be possible to only retrieve and download data around specific analysis questions while applying rules around current geospatial queries. These capabilities address several of the *Oncor* project philosophies. The incremental roll-outs of the *Oncor* database and user interface also allow for stakeholder feedback so we ensure the system can meet the needs of the user community.

The current externally published web-interface for *Oncor* can be viewed at http://gisx.pnl.gov/lcre/ and the next release of the *Oncor* web-interface will be published at http://oncor.pnnl.gov.

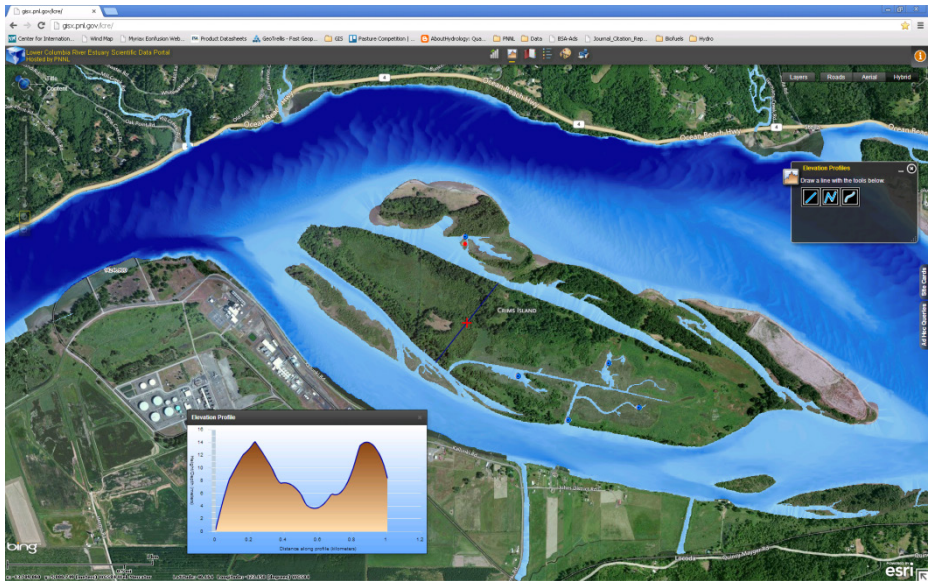


Figure 3.4. Web-based geographic information system tools/geoprocessing functions to directly interact with the data, such as querying Light Detection and Ranging and bathymetry data to produce cross-section profiles.

3.4.3 Managed and Unmanaged Data

An important concept to understand with regard to the *Oncor* database and user interface is "managed" and "unmanaged" data. As a guiding philosophy in the development of *Oncor,* the project team is to "coordinate and collaborate with other regional database efforts." This happens on two levels: first, as was described in Section 2.2, meetings with various regional entities address a variety of purposes; and second, *Oncor* uses data owned and maintained by others through a special means of live data access over the internet referred to as "web services." These data are formally referred to as "unmanaged" because the data are not stored within the *Oncor* database nor does the *Oncor* team control the data. An organization or entity that provides data as a web service can allow other servers on the internet to directly access the data, according to any security rules that have been implemented, as if the data were located on the local server. Significant benefits of offering data as a web service are found largely in content management allowing data to be controlled and revised by a single source. In other words, if an update to a data set is required, the data steward can make the required updates and push the fresh data to the web service making it instantly available to all software clients that make use of the service—there is no other formal dissemination of data required, i.e., post ZIP files to a web or file transfer protocol server and wait for users to download the data and manually bring them into their current architecture.

Another important feature regarding unmanaged data is the idea that it is possible for organizations to consume and share web services with each other as a mechanism for collaboration. For example, BPA/NPCC F&WP – Projects and Portfolios (http://www.cbfish.org) can make available a web service for its Project Action Location sites (polygon areas) along with their associated ERTG scores. The cbfish.org site owns and maintains the data, but the data can be used along with RME or other data holdings from *Oncor*. Conversely, data holdings in *Oncor's* managed data, or data held locally in the *Oncor* database, can be made available to others as a web service for those who need certain data sets for

their workflows. Data held locally in *Oncor* must go through a "Data Loading" process described in more detail below. In the prototype development of *Oncor*, a well-known web service was built into the user interface to enable and test the capability. The USGS offers a very comprehensive web service (see http://waterservices.usgs.gov/) to retrieve water-related data from the National Water Information System. This web service and the interactive spatial queries of the *Oncor* user interface allow a user to select an area on a map and select the attributes of interest (i.e., flow, stage, temperature, date range, etc.) for USGS gage stations found within the selection box, and data are queried and retrieved for these sites (see Figure 3.5 and Figure 3.6). Locally-held managed data (see Figure 3.7) interfaces with unmanaged data in the same manner, as if all data are held on the server locally; this comes about through good design and development of the web-based user interface.

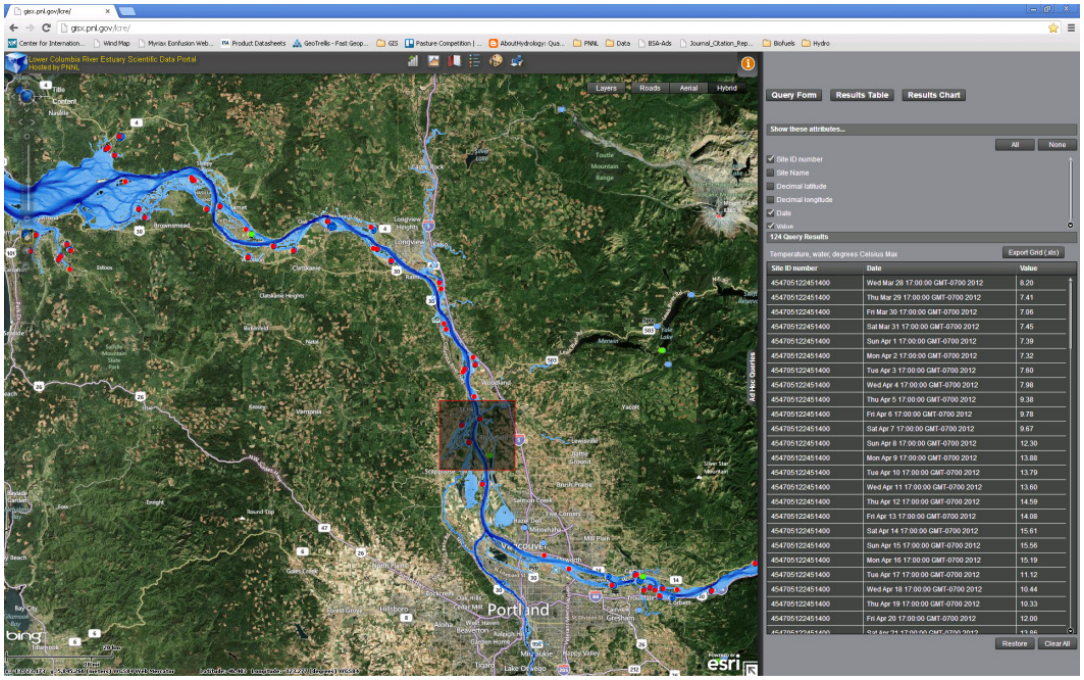


Figure 3.5. Retrieval of "unmanaged" data (via web services) time-series defined by an interactive map-based box selection and user-based query filter to narrow date range and attributes of interest (e.g., water temperature, stage, etc.).

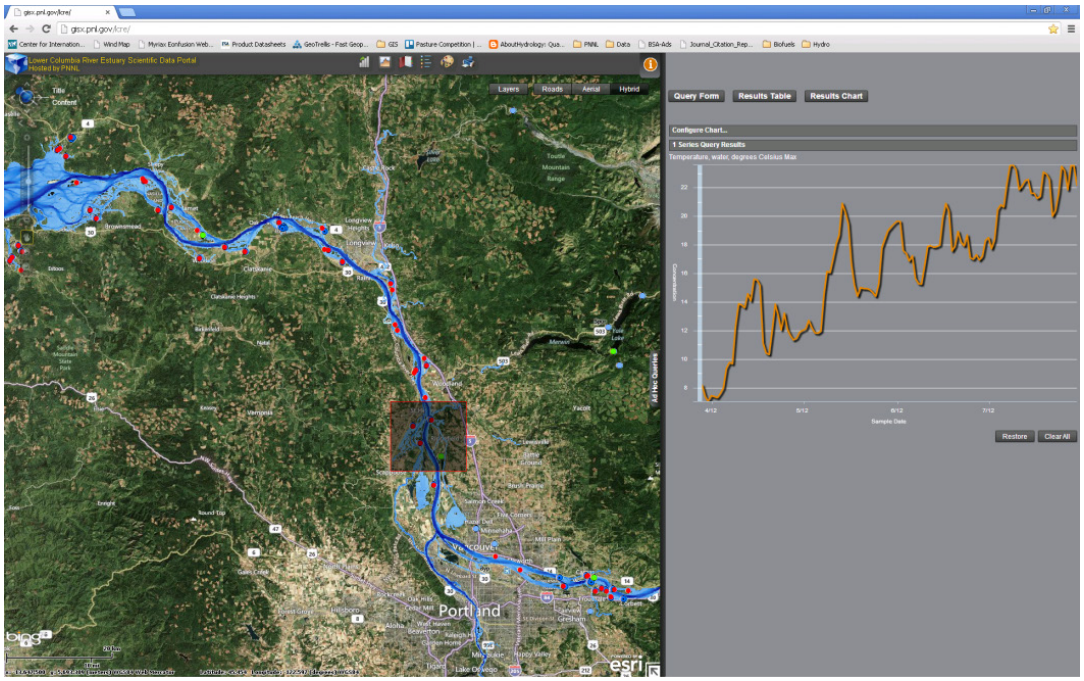


Figure 3.6. Plotting of "unmanaged" data (via web services) retrieved through interactive area-of-interest selection on the map interface.

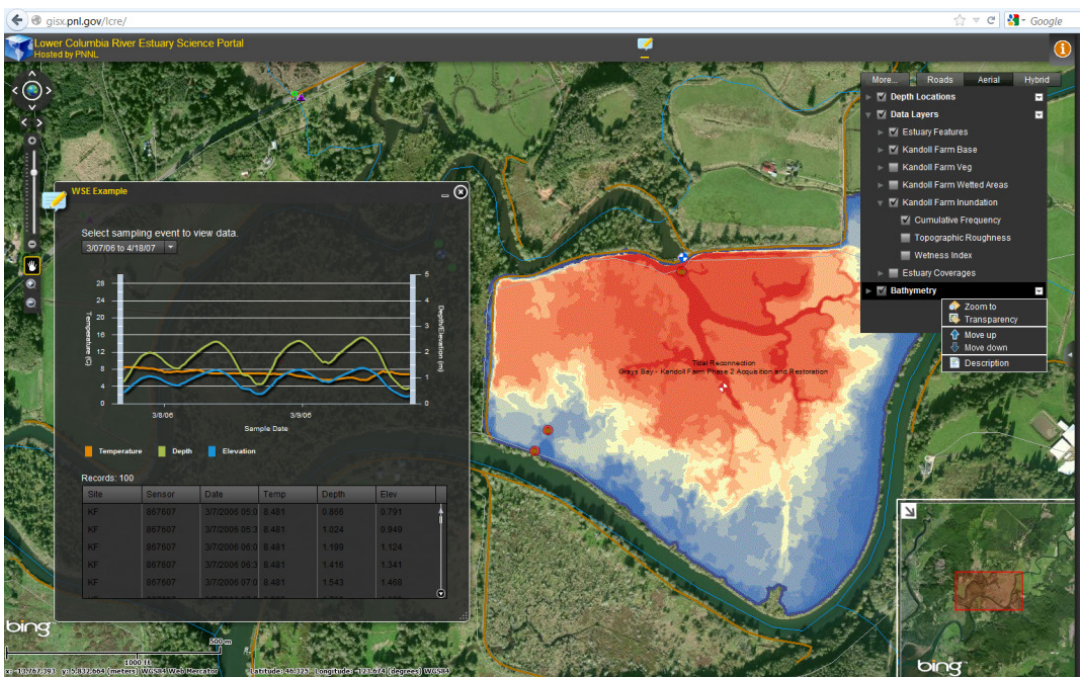


Figure 3.7. Interactive selection, access, and retrieval of instrument data and time-series plots for user-defined time periods.

Data that are eventually included within the *Oncor* system (managed) can be owned by anyone who grants permission to include the data in the system. As indicated in the metadata portion of the DETs, full credit and citation are read into the database, stored, and provided to users who access each and every

data set included in *Oncor*. A workflow is followed to prepare data for inclusion in the system. An example of *Oncor* managed data categories and subcategories is included in Table 3.8 and provides a means by which to track data from inclusion in the data model (if not included in the base set of subcategories), to DET generation and loading, and to explicitly ensuring the data set is included in the user interface and is available for use. Elements of this workflow are discussed under the *Oncor* Data Model (Section 3.3), Data Reduction Procedures (Section 3.2), and Data Exchange Template (Section 3.2). Section 3.4.4 describes the process of transforming and loading data.

Table 3.8. *Oncor* – Example of data tracking process for managed data. Black cells indicate completion; gray cells indicate in-progress.

Category	Subcategory	Built Into Data Model	Reduced & Normalized	QA/QC	DET Ready Uploading	Transform / Data	User-Interface Developed	Comment
Survey	Cross Sections							
	Sediment Stake							
	Photo Locations							
Water	Properties							
	Quality							
Hydrology	Water Surface Elevation							
	Inundation Model							
Vegetation	Veg Herb							
	Veg Tree							
	Veg Shrub							
	Veg Transect							
	Biomass							
Fish	Catch							
	Salmon Size							
	Non-Salmon Size							
	Stomach Content							
	Condition							
	Genetics							
Invertebrates	Benthos							
	Neuston							
	Insect Fallout							
Virtual	Streamflow							
	Water Quality							
	Stage							
Raster	Topo/Bathym.							
	Eco Classification							
	Land Cover							
	Historical Imagery							

Table 3.8. (contd)

Category	Subcategory	Built Into Data Model	Reduced & Normalized	QA/QC	DET Ready Uploading	Transform / Data	User-Interface Developed	Comment
	Historical Topographic Maps	■	■	■	■			
	Diking Layer	■						
	H- and T-sheets	■						
ERTG	Work Products							
	SBU Reports							
	Meeting Notes							
Library	Reports	■	■	■				
	Journal Articles	■	■	■				

3.4.4 Data Loading

The most critical component of *Oncor* is the data it holds. As such, a primary goal of *Oncor* is to provide data generators with a convenient mechanism by which to share their data. For this goal to be realized, the preparation of data for loading into the *Oncor* database must be made as easy as possible, while maintaining high data quality standards. The DET (see Section 3.2) was developed to establish a standard format that data generators can use to deliver data to *Oncor*. However, the *Oncor* development team recognizes the presence of voluminous quantities of valuable legacy data that also require a loading mechanism, but time and resources are often not available to transform and load data into a format that is suitable for the DETs. Figure 3.8 summarizes the general workflow for uploading both DET-based and legacy data to *Oncor*.

In the context of *Oncor*, legacy data are considered large data sets that are desirable for inclusion in *Oncor* that the data owner cannot, or will not, convert into DET format. This generally includes data in external databases, or data not loaded in any database. Legacy data can be imported into *Oncor* by first converting them to a DET format using a custom-developed translation program, then following the normal DET-based loading procedure, described below.

Ideally, the majority of data arriving for input into *Oncor* will be in DET format. These data will first pass through a translation and validation step. Here the data are checked for conformance to the DET standard. Any format discrepancies are flagged and the entire DET is rejected and returned to the data generator with details of the problem, which likely means data are missing from required fields. If the data format is correct, selected fields are then validated against a set of rules defined in the DET Data Dictionary. These rules only identify blatant errors or physically impossible values and are not meant to replace a full QA validation, which could come about for a number of reasons, including faulty instrument, incorrect transcription of data, incorrect units reported, etc. Rule violations result in a rejection of the DET and the data generator is notified of the offending records. Finally, data are compared with existing records in *Oncor* to verify that duplicates will not be created and that the necessary prerequisite information is present. Certain values, such as sampling methods, instrument

names, or people records, must be present in the database prior to loading data that references them. Again, if violations occur, the DET is returned to the data generator. When all validation checks succeed, the data are translated into a structure defined by the *Oncor* data model.

The last step of data loading consists of actually performing the transactions to migrate the new data into the production database. This step includes the assignment of unique identifiers to the new records and documentation of successful loading. Finally, as the DETs are being established, the data loading workflow is being developed and tested.

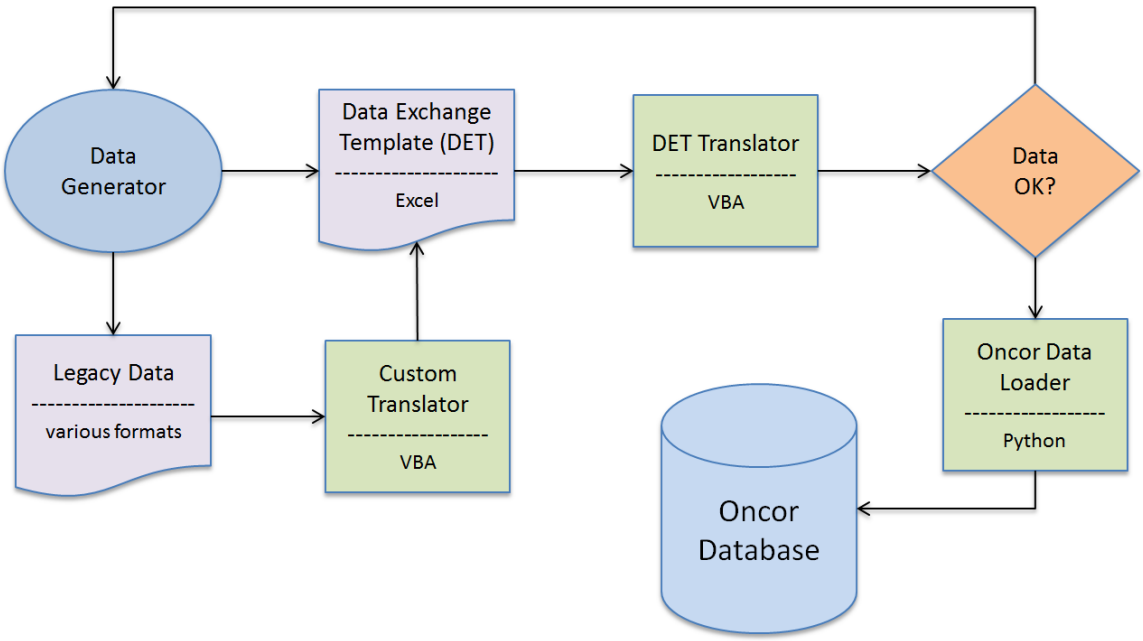


Figure 3.8. Workflow to bring data processed through the Data Reduction Procedure into the *Oncor* database. (VBA is Visual Basic for Applications.)

4.0 Accomplishments and Recommendations

Accomplishments during the first year (FY 2012) of the *Oncor* database development effort can be summarized as follows:

- The extensive coordination effort was fruitful in promoting understanding, constructive input, and buy-in from the user community.
- Substantive coordination occurred with the database development efforts for PNAMP and BPA's restoration project tracking system, called "cbfish.org."
- Communication and coordination was undertaken with the EP and its data responsibilities.
- An *Oncor* data model was conceived and developed for RME and other data from CEERP efforts in the LCRE.
- An *Oncor* database with user interface was built and launched successfully, thereby establishing proof-of-concept.
- Data reduction procedures, including DETs, were established for water-surface elevation and sediment accretion.
- Example data from a variety of data categories were normalized and uploaded to *Oncor*.
- The *Oncor* data model and database were tested using example data addressing particular analysis questions and use cases.

Much work remains to fully develop and apply *Oncor*. Generally, objectives for FYs 2013 and 2014 involve continuing *Oncor* development and transferring *Oncor* technology (Table 4.1). We recommend the following activities during 2013 and beyond:

- Continue coordination efforts (Avenues A, B, and C).
- Disseminate draft procedures for water-surface elevation and sediment accretion and solicit feedback from the user community.
- Continue work on new data reduction procedures for other data categories, such as fish and vegetation.
- Develop data entry tools.
- Establish routine data uploads from the EP's Ecosystem Monitoring project and WDFW's restoration Memorandum of Agreement projects, and other studies.
- Populate *Oncor* with regionally available, normalized data to prepare to support estuary-wide meta-analysis of effectiveness data.
- Develop tools for automatic data processing of particular data sets (to be determined) within *Oncor* after data have been uploaded.
- Solicit feedback on *Oncor*'s user interface and update accordingly.
- Post data and links to other databases on the *Oncor* web portal to make them accessible to CEERP stakeholders.

Table 4.1. General objectives for 2012, 2013, and 2014 for the Synthesis and Evaluation project, which includes regional coordination, *Oncor* database development, CEERP analysis and synthesis, and information dissemination. Note that annual scopes of work depend on available funding levels.

	2012 Coordinate Regionally and Establish Prototype Database	2013 Refine Database Application	2014 Transfer to Regional Entity
Obj. 1 Coordination	Obtain stakeholder input and guidance during database development	Obtain stakeholder review and feedback on prototype database	Obtain stakeholder coordination for eventual transfer of technology
Obj. 2 Database Development	Develop estuary data model and prototype database (example using PNNL-collected data); confirm that outputs support key analysis questions; identify and coordinate with other compatible regional data systems	Continue development of data reduction protocols; refine the Estuary Data Model and continue to build the LCRE Database and populate or link it with regionally available, normalized data to prepare to support estuary-wide meta-analysis of effectiveness data	Finalize geospatial database management and analysis system; document database structure; train new users
Obj. 3 Analysis and Synthesis	Draft and finalize the CEERP 2012 Strategy Report and Action Plan; draft the 2013 Strategy Report and Action Plan	Analyze and synthesize *pilot* data to address program objectives; apply results to science-based decisions in development of the 2013 Synthesis Memorandum; 2014 Strategy Report and 2014 Action Plan	Analyze and synthesize data to address program objectives; apply results to science-based decisions in development of the 2014 Synthesis Memorandum; 2015 Strategy Report and 2015 Action Plan; perform LOE evaluation
Obj. 4 Dissemination	---	Continue web-interface development; post *pilot* data and links to other databases on the database web portal to make it accessible to stakeholders	Continue to post data on the database web portal to make it accessible to stakeholders; continued updates to web portal as needed.

5.0 References

Borde AB, SA Zimmerman, RM Kaufmann, HL Diefenderfer, NK Sather, RM Thom. 2011. *Lower Columbia River and Estuary Restoration Reference Site Study: 2010 Final Report and Site Summaries*. PNWD-4262, prepared for the Lower Columbia River Estuary Partnership by Pacific Northwest National Laboratory, Marine Sciences Laboratory, Sequim, Washington.

Borde AB, VI Cullinan, HL Diefenderfer, RM Thom, RM Kaufmann, J Sagar, and C Corbett. 2012. *Lower Columbia River and Estuary Ecosystem Restoration Program Reference Site Study: 2011 Restoration Analysis*. PNNL-21433, final report prepared for the Lower Columbia River Estuary Partnership by Pacific Northwest National Laboratory, Richland, Washington.

Bottom DL, CA Simenstad, J Burke, AM Baptista, DA Jay, KK Jones, E Casillas, and M Schiewe. 2005. *Salmon at River's End: The Role of the Estuary in the Decline and Recovery of Columbia River Salmon*. NOAA Technical Memorandum NMFS-NWFSC-68. NOAA Fisheries, Northwest Fisheries Science Center, Seattle, Washington.

Bottom DL, G Anderson, A Baptista, J Burke, M Burla, M Bhuthimethee, L Campbell, E Casillas, S Hinton, K Jacobson, D Jay, R McNatt, P Moran, GC Roegner, CA Simenstad, V Stamatiou, D Teel, and JE Zamon. 2008. *Salmon Life Histories, Habitat, and Food Webs in the Columbia River Estuary: An Overview of Research Results, 2002–2006*. Report prepared for the U.S. Army Corps of Engineers, Portland District, Northwest Division, Portland, Oregon, and Bonneville Power Administration, Environment, Fish, and Wildlife Division, Portland, Oregon, by the Northwest Fisheries Science Center, National Marine Fisheries Service, Seattle, Washington.

BPA/Corps (Bonneville Power Administration and U.S. Army Corps of Engineers). 2012. *Columbia Estuary Ecosystem Restoration Program: 2012 Strategy Report*. Final report, Portland, Oregon.

Diefenderfer HL, A Coleman, AB Borde, and IA Sinks. 2008. "Hydraulic Geometry and Microtopography of Tidal Freshwater Forested Wetlands and Implications for Restoration, Columbia River, U.S.A." *Ecohydrology & Hydrobiology* 8:339–361. doi:10.2478/v10104-009-0027-7. PNNL-SA-59093.

Diefenderfer HL and DR Montgomery. 2009. "Pool spacing, channel morphology, and the restoration of tidal forested wetlands of the Columbia River, U.S.A." *Restoration Ecology* 17(1):158–168. doi:10.1111/j.1526-100X.2008.00449.x. PNNL-SA-55592.

Diefenderfer HL, RM Thom, GE Johnson, JR Skalski, KA Vogt, BD Ebberts, GC Roegner, and EM Dawley. 2011. "A levels-of-evidence approach for assessing cumulative ecosystem response to estuary and river restoration programs." *Ecological Restoration* 29:111–132.

Diefenderfer HL, GE Johnson, RM Thom, AB Borde, CM Woodley, LA Weitkamp, KE Buenau, and RK Kropp. 2012. *An Evidence-Based Assessment of the Cumulative Effects of Tidal Freshwater and Estuarine Ecosystem Restoration: Early Life-Stage Habitat Functions for Endangered Salmonids*. Prepared by Pacific Northwest National Laboratory and NOAA Fisheries for the U.S. Army Corps of Engineers, Portland District, Portland, Oregon.

Endangered Species Act (ESA) of 1973, as amended. 16 USC 1531 et seq.

Healey MC. 1980. "Utilization of the Nanaimo River estuary by juvenile Chinook salmon, *Oncorhynchus tshawytscha*." *Fishery Bulletin* 77:653–668.

Johnson GE, HL Diefenderfer, BD Ebberts, C Tortorici, T Yerxa, JC Leary, and JR Skalski. 2008. *Research, Monitoring, and Evaluation for the Federal Columbia River Estuary Program*. PNNL-17300, Pacific Northwest National Laboratory, Richland, Washington.

Johnson GE, HL Diefenderfer, RM Thom, GC Roegner, BD Ebberts, JR Skalski, AB Borde, EM Dawley, AM Coleman, DX Woodruff, SX Breithaupt, A Cameron, CX Corbett, EM Donley, DA Jay, Y Ke, K Leffler, C McNeil, C Studebaker, and J Tagestad. 2012. *Evaluation of Cumulative Ecosystem Response to Restoration Projects in the Lower Columbia River and Estuary, 2010*. PNNL-20296, prepared for the U.S. Army Corps of Engineers, Portland District, Portland, Oregon, by Pacific Northwest National Laboratory, Richland, Washington.

Kukulka T and DA Jay. 2003a. "Impacts of Columbia River discharge on salmonid habitat I. a non-stationary fluvial tide model." *Journal of Geophysical Research* 108:3293. doi 10.1029/2002JC001382.

Kukulka T and DA Jay. 2003b. "Impacts of Columbia River discharge on salmonid habitat II. Changes in shallow-water habitat." *Journal of Geophysical Research* 108:3294. doi 10.1029/2003JC001829.

Levings CD. 1994. "Feeding behavior of juvenile salmon and significance of habitat during estuary and early sea phase." *Nordic Journal of Freshwater Research* 69:7–16.

Levings CD, K Conlin, and B Raymond. 1991. "Intertidal habitats used by juvenile Chinook salmon (*Oncorhynchus tshawytscha*) rearing in the North Arm of the Fraser River estuary." *Marine Pollution Bulletin* 22:20–26.

Levy DA and TG Northcote. 1982. "Juvenile salmon residency in a marsh area of the Fraser River estuary." *Canadian Journal of Fisheries and Aquatic Sciences* 39:270–276.

NMFS (National Marine Fisheries Service). 2008. *Biological Opinion – Consultation on Remand for Operation of the Federal Columbia River Power System, 11 Bureau of Reclamation Projects in the Columbia Basin and ESA Section 10(a)(1)(A) Permit for Juvenile Fish Transportation Program*. National Marine Fisheries Service (NOAA Fisheries) – Northwest Region, Seattle, Washington. Available at: http://www.salmonrecovery.gov/.

Reimers PE and RE Loeffel. 1967. "The length of residence of residence of juvenile fall Chinook salmon in selected Columbia River tributaries." *Oregon Fish Commission Research Briefs* 13:5–19.

Roegner GC, A Baptista, DL Bottom, J Burke, L Campbell, C Elliot, S Hinton, D Jay, MA Lott, T Lundrigan, R McNatt, P Moran, CA Simenstad, D Teel, E Volk, J Zamon, and E Casillas. 2008. *Estuarine Habitat and Juvenile Salmon − Current and Historical Linkages in the Lower Columbia River and Estuary, 2002-04*. Report prepared for the U.S. Army Corps of Engineers, Portland District, Northwest Division, Portland, Oregon, and Bonneville Power Administration, Environment, Fish, and Wildlife Division, Portland, Oregon, by the Northwest Fisheries Science Center, National Marine Fisheries Service, Seattle, Washington.

Roegner GC, HL Diefenderfer, AB Borde, RM Thom, EM Dawley, AH Whiting, SA Zimmerman, and GE Johnson. 2009. *Protocols for Monitoring Habitat Restoration Projects in the Lower Columbia River and Estuary*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration Technical Memorandum NMFS-NWFSC-97, Northwest Fisheries Science Center, Seattle, Washington.

Roegner GC, EW Dawley, M Russel, A Whiting, and DJ Teel. 2010. "Juvenile salmonid use of reconnected tidal freshwater wetlands in Grays River, Lower Columbia River Basin." *Transactions of the American Fisheries Society* 139:1211–1232.

Sather NK, DJ Teel, AJ Storch, GE Johnson, ES Van Dyke, EM Dawley, DL Kuligowski, TA Jones, A Bryson, and KL Sobocinski. 2011. "Juvenile Salmon and Fish Community Characteristics." In *Ecology of Juvenile Salmon in Shallow Tidal Freshwater Habitats of the Lower Columbia River, 2007-2010*, pp. 2.1–2.38, PNNL-20083, Pacific Northwest National Laboratory, Richland, Washington.

Shreffler DK, CA Simenstad, and RM Thom. 1990. "Temporary Residence by Juvenile Salmon in a Restored Estuarine Wetland." *Canadian Journal of Fisheries and Aquatic Sciences* 47:2079–2084.

Shreffler, DK, CA Simenstad, and RM Thom. 1992. "Foraging by Juvenile Salmon in a Restored Estuarine Wetland." *Estuaries* 15:204–213.

Simenstad CA, JL Burke, JE O'Connor, C Cannon, DW Heatwole, MF Ramirez, IR Waite, D Counihan, and KL Jones. 2011. *Columbia Estuary Ecosystem Classification – Concept and Application*. U.S. Geological Survey Open File Report 2011-1228, Portland, OR.

Sommer TR, ML Nobriga, WC Harrell, W Batham, and WJ Kimmerer. 2001. "Floodplain rearing of juvenile Chinook salmon: Evidence of enhanced growth and survival." *Canadian Journal of Fisheries and Aquatic Sciences* 58:325–333.

Storch AJ. 2011. "Bioenergetics." In *Ecology of Juvenile Salmon in Shallow Tidal Freshwater Habitats of the Lower Columbia River, 2007-2010*, pp. 5.1–5.23, PNNL-20083, Pacific Northwest National Laboratory, Richland, Washington.

Storch AJ and NK Sather. 2011. "Feeding Ecology." In *Ecology of Juvenile Salmon in Shallow Tidal Freshwater Habitats of the Lower Columbia River, 2007-2010*, pp. 4.1–4.20, PNNL-20083, Pacific Northwest National Laboratory, Richland, Washington.

Tanner CD, JR Cordell, J Rubey, and LM Tear. 2002. "Restoration of freshwater intertidal habitat functions at Spencer Is., Everett, Washington." *Restoration Ecology* 10:564–576.

Thom RM, AB Borde, NR Evans, CW May, GE Johnson, and JA Ward. 2004. *A Conceptual Model for the Lower Columbia River Estuary*. PNNL-SA-43444, report to the U.S. Army Corps of Engineers, Portland District, Portland, Oregon, by Pacific Northwest National Laboratory, Richland Washington. Available at: www.nwp.usace.army.mil/pm/lcr/science.asp.

Thom RM, NK Sather, GC Roegner, and DL Bottom. 2013. *Columbia Estuary Ecosystem Restoration Program: 2012 Synthesis Memorandum*. Final report prepared by Pacific Northwest National Laboratory and National Marine Fisheries Service, submitted to the U.S. Army Corps of Engineers, Portland, Oregon.

Wiens JA. 2002. "Riverine landscapes: Taking landscape ecology into the water." *Freshwater Biology* 47:501–515.

Appendix A

Outline for
Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program

Appendix A

Outline for Data Reduction Procedures for the *Oncor* Database of the Columbia Estuary Ecosystem Restoration Program

1. Introduction
    - 1.1. Background
    - 1.2. Purpose
2. General Procedures
    - 2.1. Terminology
    - 2.2. Quality Control and Quality Assurance
    - 2.3. Documentation
    - 2.4. *Oncor* Data Loading Options
3. Indicator-Specific Procedures
    - 3.1. Water Surface Level
        - 3.1.1. Data Processing Procedures
        - 3.1.2. Data Reduction Workbook
        - 3.1.3. Data Exchange Template
    - 3.2. Water Temperature
        - 3.2.1. Data Processing Procedures
        - 3.2.2. Data Reduction Workbook
        - 3.2.3. Data Exchange Template
    - 3.3. Channel Cross-Section Surveys
        - 3.3.1. Data Processing Procedures
        - 3.3.2. Data Reduction Workbook
        - 3.3.3. Data Exchange Template

3.4. Sediment Accretion

3.4.1. Data Processing Procedures

3.4.2. Data Reduction Workbook

3.4.3. Data Exchange Template

3.5. Vegetation

3.5.1. Herbaceous

3.5.1.1. Data Processing Procedures

3.5.1.2. Data Reduction Workbook

3.5.1.3. Data Exchange Template

3.5.2. Shrub

3.5.2.1. Data Processing Procedures

3.5.2.2. Data Reduction Workbook

3.5.2.3. Data Exchange Template

3.5.3. Tree

3.5.3.1. Data Processing Procedures

3.5.3.2. Data Reduction Workbook

3.5.3.3. Data Exchange Template

3.6. Fish

3.6.1. Data Processing Procedures

3.6.2. Data Reduction Workbook

3.6.3. Data Exchange Template

4. Conclusions and Recommendations

5. Literature Cited

Appendix B

Data Reduction Procedure for Water Surface Elevation

Appendix B

Data Reduction Procedure for Water Surface Elevation

Prepared by Amy Borde, Shon Zimmerman, Andre Coleman, and Ron Kaufmann
Pacific Northwest National Laboratory

B.1 Introduction

Roegner et al. (2009) identifies hydrology, and specifically, measures in the variation of water-surface elevation (WSE) as part of the "Core Monitored Metrics" (see Section 2.2.1 – Hydrology). Pressure transducers with automated hourly logging are the recommended approach for capturing the variability in water levels. Section 4.1 – Hydrology (Protocol 1) of Roegner et al. (2009) provides the field deployment and general calculation and analysis methodology. The information contained in this document provides a procedure for processing data collected with a data-logging pressure transducer (hereafter water level logger), methods for calculating WSE, quality control measures, and final data formatting for entry into the *Oncor* database. This document also refers to a number of folders and files that correspond to suggested directory structure for processing data and example spreadsheets that help the user to process data into the final Data Exchange Template format. In the future, these companion tools/examples will be made available on the *Oncor* website.

B.2 The Sensor Log

The sensor log is a way of tracking and organizing the information collected as part of the field deployment and retrieval of water level loggers. The sensor log should contain information such as the following:

- serial number of each sensor for data tracking
- deployment and retrieval date and time
- initials of personnel conducting deployment/retrieval
- water depth above the sensor at the time of deployment and retrieval. This information is used to verify that the sensor is measuring correctly.
- distance from the sensor to the top of deployment post (or some way of marking position of the sensor) at deployment and retrieval. This information is critical to determining the sensor elevation when elevation is surveyed at the top of the deployment post. This distance is also useful for determining whether the sensor moved during the deployment period.
- distance from the sediment to the top of the post. This measurement is not needed for the processing, but indicates whether conditions surrounding the sensor changed during the deployment (e.g., erosion, deposition) or if the post moved.
- elevation of the top of the deployment post and/or elevation of the sensor
- notes about the deployment or retrieval.

B.3 Steps for Determining Water-Surface Elevation

The following sections provide the steps needed to correct the absolute pressure readings measured with the pressure transducer to sensor depth and subsequently to WSE.

An organized file structure is important for keeping track of the various files created throughout the data correction process. A suggested file structure is provided as part of the Data Exchange Template and includes the following folders:

- Step0_Sensor Log
- Step1_Original Files
- Step2_Corrections
  - Step1_AtmCorrection
  - Step2_ElevCorrection
- Step3_QAQC
- Step4_Final.

These folders are referenced in the sections below with the following notation to indicate nested folders: \\Step2_Corrections\Step1_AtmCorrection

Suggested filenames and locations in the file structure are provided in the sections below.

B.4 Atmospheric Pressure Correction

Water level loggers record absolute pressure, which is converted to water depth by processing software. Absolute pressure includes atmospheric pressure and water pressure. Atmospheric pressure is nominally 100 kPa (14.5 psi) at sea level and must be separated from the water pressure to determine water depth. In addition, atmospheric data fluctuates with weather and altitude; left uncompensated, barometric variations could result in errors of 0.6 m (2 ft) or more.

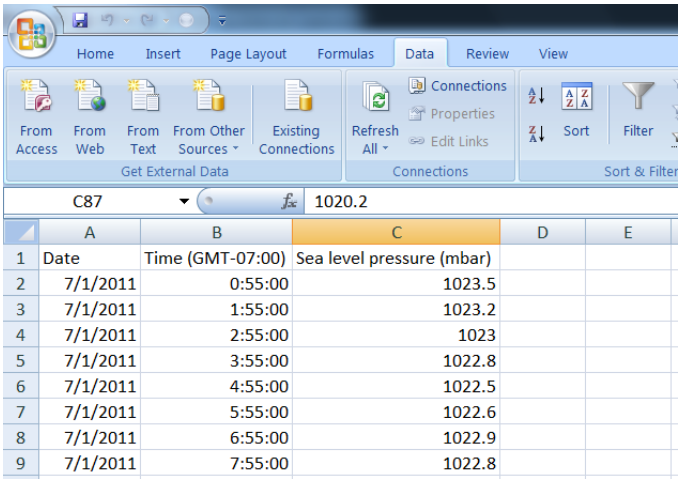
To compensate for barometric pressure changes, hourly atmospheric pressure data need to be acquired from either a local meteorological station (i.e., from the nearest station listed under the National Oceanographic and Atmospheric Administration's National Climatic Data Center [NCDC]) or from another water level logger that was deployed near the site and out of the water, specifically to collect atmospheric pressure. In either case, the time period of atmospheric pressure data needs to span the same time period over which the water level logger was deployed and at the same temporal resolution (i.e., hourly). Some software allows for the direct input of atmospheric data when they are collected with the same kind of data-logger. If this is the case, then simply follow the manufacturer's instructions for the correction process. If atmospheric data are needed then they can be downloaded from the NCDC at: http://cdo.ncdc.noaa.gov/pls/plclimprod/poemain.accessrouter?datasetabbv=DS3505

Follow these steps at the NCDC website:

1. Click "*I Agree to these terms (continue)*".

2. Click "*Continue With SIMPLIFIED Options*".
3. Select the "*Country – United States*" radio button and then click "*Continue*".
4. Select "*Oregon or Washington*" from the dropdown menu and click continue.
5. Select the station of interest with a comparable data range (e.g., Portland Troutdale, WBAN# 24242) and click "*Continue*".
6. Choose your date and time window from the dropdown menus and click "*Continue*". (Note: hourly-only data can be selected; however, this can sometimes result in an incomplete data set. Check the resulting data set to be sure adequate data exist and if not then re-submit request without hourly-only data being selected.)
7. Check the "*Inventory Review*" box, type in your email address and click "*Submit request*". You will receive an email from NCDC with links to your data. Save the data as a text file and proceed to the next step.

Format the data file according to the software's specifications for using an imported atmospheric data file. For example, Onset Computer HOBOware® software requires a tab delimited text file with only three columns: "Date", "Time (in same time zone as data was collected)", and "Sea level pressure (SLP) (mbar)", with the data in the following format: mm/dd/yyyy, hh:mm:ss, and the SLP should have one decimal place visible (see example below). Formatting can be done in Excel or any other spreadsheet.



	A	B	C	D	E
1	Date	Time (GMT-07:00)	Sea level pressure (mbar)		
2	7/1/2011	0:55:00	1023.5		
3	7/1/2011	1:55:00	1023.2		
4	7/1/2011	2:55:00	1023		
5	7/1/2011	3:55:00	1022.8		
6	7/1/2011	4:55:00	1022.5		
7	7/1/2011	5:55:00	1022.6		
8	7/1/2011	6:55:00	1022.9		
9	7/1/2011	7:55:00	1022.8		

Follow the manufacturer's software instructions for converting pressure data to water depth using the atmospheric data file and specifying the water density. Water density is determined based on salinity and/or the water temperature if freshwater. After the water depth is calculated, export the data for further manipulations in a spreadsheet.

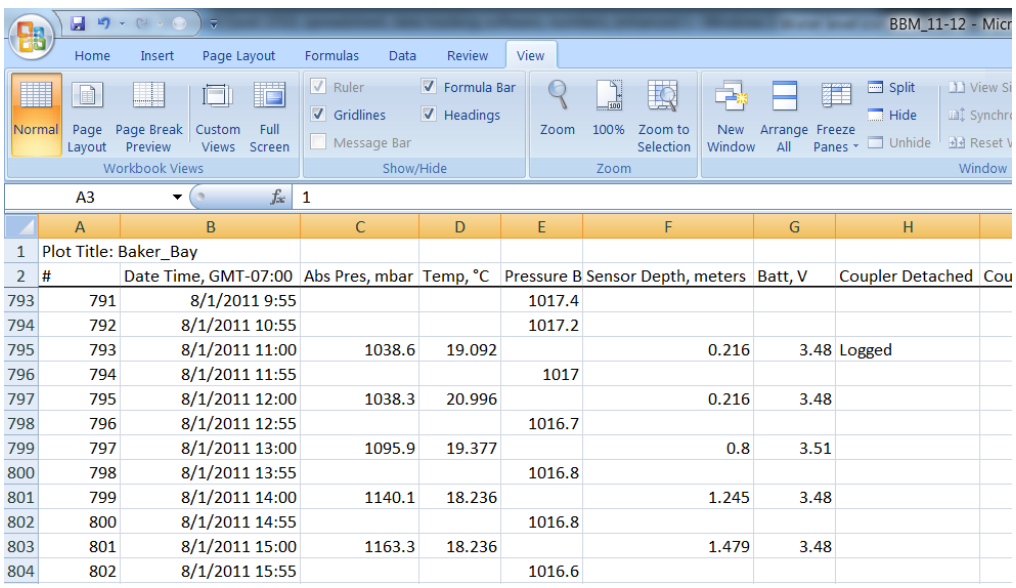
Example filename: SITE_atm_correct_mmddyy.xlsx
Located in: \\Step2_Corrections\Step1_atm_corrected

B.5 Elevation Correction

NOTE: In the future, a script will be written to automate the process in this section and it will be included as part of the Data Exchange Template package.

The WSE is determined by the sum of the sensor depth below the water level surface and the surveyed elevation of the water level logger (see Roegner et al. 2009, 4.1 Hydrology). The WSE will be relative to the datum in which the elevation was collected and can be converted into different data if desired.

The file exported from the water depth processing software will look something like the screen shot below.



	A	B	C	D	E	F	G	H
1	Plot Title: Baker_Bay							
2	#	Date Time, GMT-07:00	Abs Pres, mbar	Temp, °C	Pressure B	Sensor Depth, meters	Batt, V	Coupler Detached
793	791	8/1/2011 9:55			1017.4			
794	792	8/1/2011 10:55			1017.2			
795	793	8/1/2011 11:00	1038.6	19.092		0.216	3.48	Logged
796	794	8/1/2011 11:55			1017			
797	795	8/1/2011 12:00	1038.3	20.996		0.216	3.48	
798	796	8/1/2011 12:55			1016.7			
799	797	8/1/2011 13:00	1095.9	19.377		0.8	3.51	
800	798	8/1/2011 13:55			1016.8			
801	799	8/1/2011 14:00	1140.1	18.236		1.245	3.48	
802	800	8/1/2011 14:55			1016.8			
803	801	8/1/2011 15:00	1163.3	18.236		1.479	3.48	
804	802	8/1/2011 15:55			1016.6			

1. Save the file "atm_correct" as "elev_correct"

   Example filename: SITE_elev_correct_mmddyy.xlsx

   Located in: \\Step2_Corrections\Step2_elev_corrected

   Remove unnecessary data columns. Keep the "Plot Title: Site ID" text at the top of the spreadsheet. The only needed columns are:

   a. "Date time"

   b. "Temp, °C"

   c. "Sensor Depth, meters".

2. If the atmospheric pressure and the absolute pressure do not occur at the same logging interval, blank cells will occur in the Sensor Depth column and need to be removed.

3. Often the data are collected in the time zone at the time of deployment (e.g., daylight savings time, which is GMT-7 in the Pacific time zone). If desired, convert the time to a different time zone such as GMT-0 or "local" time, which will switch between daylight savings and standard time on the correct date.

a. If converting to local time, complete the following steps:

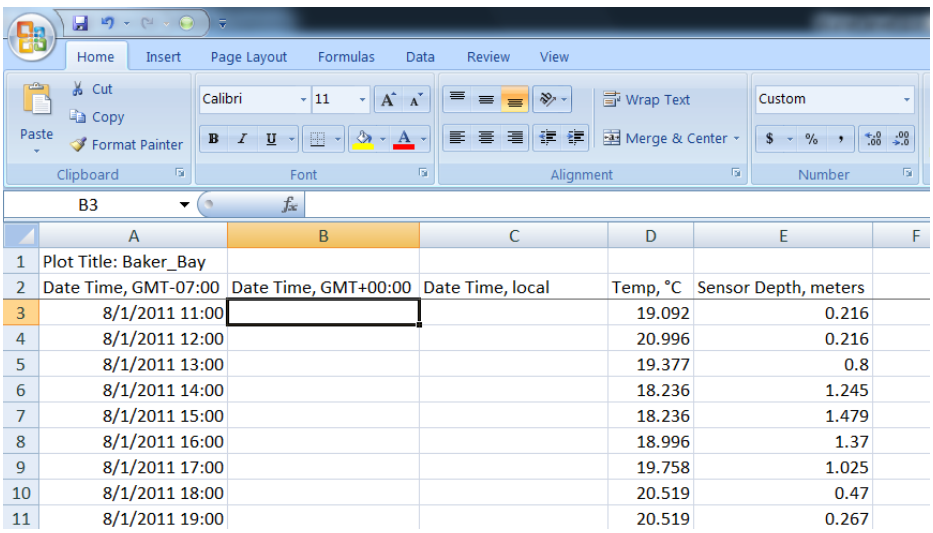
   i. From the Data Exchange Template package open: WL_Localtime_correction_template_v2.xls located in (\\Step2_Corrections\time_template).

   ii. Copy the TimeRef worksheet from its original workbook to the WSE workbook you are currently editing.

   iii. Insert a column for the "Date Time, local."

   iv. Copy the formula found in cell C3 from the example worksheet in the "WL_Localtime_correction_template_v2.xlsx" workbook.

   v. Paste the copied formula into the first blank cell in your new column below the Date Time, local header. Populate the rest of the column with the formula.



	A	B	C	D	E	F
1	Plot Title: Baker_Bay					
2	Date Time, GMT-07:00	Date Time, GMT+00:00	Date Time, local	Temp, °C	Sensor Depth, meters	
3	8/1/2011 11:00			19.092	0.216	
4	8/1/2011 12:00			20.996	0.216	
5	8/1/2011 13:00			19.377	0.8	
6	8/1/2011 14:00			18.236	1.245	
7	8/1/2011 15:00			18.236	1.479	
8	8/1/2011 16:00			18.996	1.37	
9	8/1/2011 17:00			19.758	1.025	
10	8/1/2011 18:00			20.519	0.47	
11	8/1/2011 19:00			20.519	0.267	

b. If converting to GMT+0, follow the steps above except populate the column with the sum of GMT-7 date/time and the value 0.2916666 or the sum of GMT-8 (Pacific Standard Time) and the value 0.33333333.

4. Remove the data with time stamps that are outside the deployment period—i.e., times when the logger is recording, but is not in the water—by referring to the deployment/retrieval times in the sensor log. Be sure times zones are the same in the sensor log and in the data file.

5. In the sensor depth column, if the sensor was out of the water at any time the values will be negative. These negatives need to be changed to 9999 so they are easily identifiable during later processing.

6. Label column F "ElevCorrection". Get the sensor elevation value from the sensor log and place it in the first row under the new heading.

7. Label column G "WLElev (m, NAVD88)". In the first row under this header, sum the values from columns E ("Sensor Depth, meters") and F ("ElevCorrection"). The resultant value is the WSE. Populate the rest of columns F ("ElevCorrection") and G ("WLElev (m, NAVD88)") with the same calculations as are in the first row.

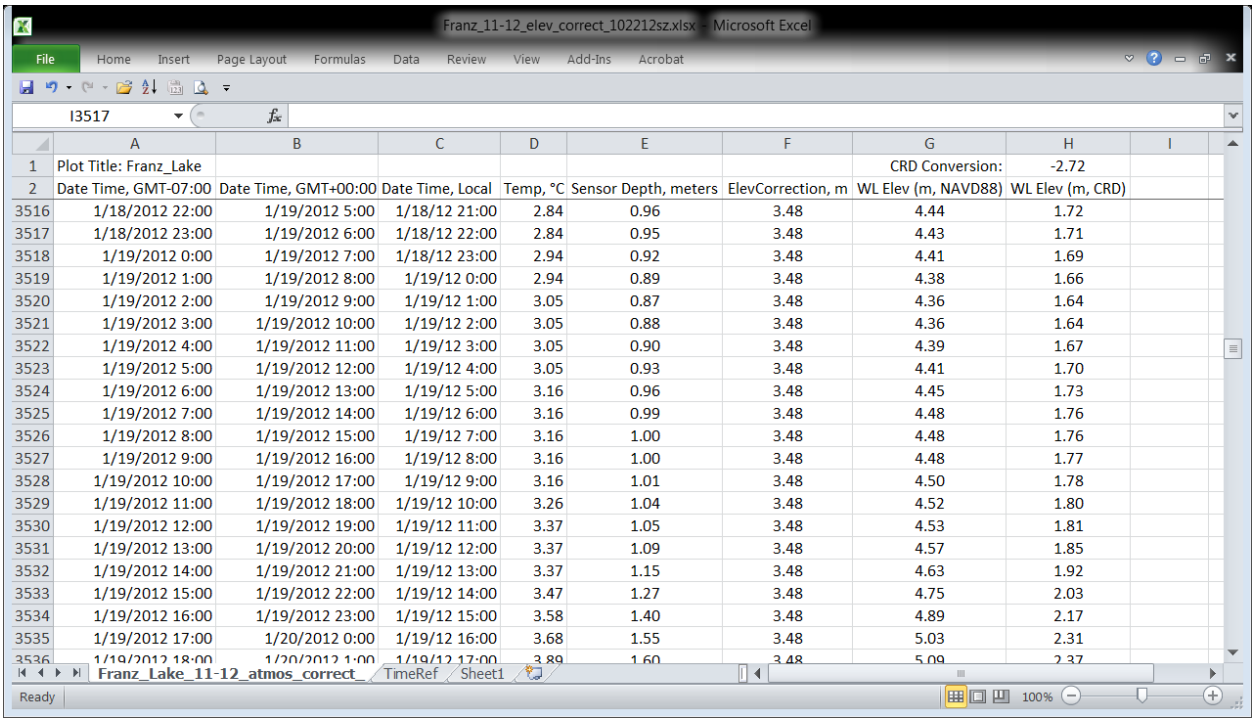
8. To convert to the Columbia River Datum (CRD), insert the heading “CRD Conversion:” in cell G1. Get the CRD conversion value for the site from file: \\Step2_Corrections\Step3_CRD conversions.xlsx

   a. Place the CRD Conversion value in the cell H1.

   b. In column H add the header “WL Elev (m, CRD)”.

   c. Below the header “WL Elev (m, CRD)” enter the equation “=IF(E3<9000,(G3+H1),9999)”. This will produce the final WSE in meters, relative to the Columbia River Datum. Populate the rest of the column with this equation.

See screen shot below for example of the elevation corrected worksheet.



	A	B	C	D	E	F	G	H	I
1	Plot Title: Franz_Lake						CRD Conversion:	-2.72	
2	Date Time, GMT-07:00	Date Time, GMT+00:00	Date Time, Local	Temp, °C	Sensor Depth, meters	ElevCorrection, m	WL Elev (m, NAVD88)	WL Elev (m, CRD)	
3516	1/18/2012 22:00	1/19/2012 5:00	1/18/12 21:00	2.84	0.96	3.48	4.44	1.72	
3517	1/18/2012 23:00	1/19/2012 6:00	1/18/12 22:00	2.84	0.95	3.48	4.43	1.71	
3518	1/19/2012 0:00	1/19/2012 7:00	1/18/12 23:00	2.94	0.92	3.48	4.41	1.69	
3519	1/19/2012 1:00	1/19/2012 8:00	1/19/12 0:00	2.94	0.89	3.48	4.38	1.66	
3520	1/19/2012 2:00	1/19/2012 9:00	1/19/12 1:00	3.05	0.87	3.48	4.36	1.64	
3521	1/19/2012 3:00	1/19/2012 10:00	1/19/12 2:00	3.05	0.88	3.48	4.36	1.64	
3522	1/19/2012 4:00	1/19/2012 11:00	1/19/12 3:00	3.05	0.90	3.48	4.39	1.67	
3523	1/19/2012 5:00	1/19/2012 12:00	1/19/12 4:00	3.05	0.93	3.48	4.41	1.70	
3524	1/19/2012 6:00	1/19/2012 13:00	1/19/12 5:00	3.16	0.96	3.48	4.45	1.73	
3525	1/19/2012 7:00	1/19/2012 14:00	1/19/12 6:00	3.16	0.99	3.48	4.48	1.76	
3526	1/19/2012 8:00	1/19/2012 15:00	1/19/12 7:00	3.16	1.00	3.48	4.48	1.76	
3527	1/19/2012 9:00	1/19/2012 16:00	1/19/12 8:00	3.16	1.00	3.48	4.48	1.77	
3528	1/19/2012 10:00	1/19/2012 17:00	1/19/12 9:00	3.16	1.01	3.48	4.50	1.78	
3529	1/19/2012 11:00	1/19/2012 18:00	1/19/12 10:00	3.26	1.04	3.48	4.52	1.80	
3530	1/19/2012 12:00	1/19/2012 19:00	1/19/12 11:00	3.37	1.05	3.48	4.53	1.81	
3531	1/19/2012 13:00	1/19/2012 20:00	1/19/12 12:00	3.37	1.09	3.48	4.57	1.85	
3532	1/19/2012 14:00	1/19/2012 21:00	1/19/12 13:00	3.37	1.15	3.48	4.63	1.92	
3533	1/19/2012 15:00	1/19/2012 22:00	1/19/12 14:00	3.47	1.27	3.48	4.75	2.03	
3534	1/19/2012 16:00	1/19/2012 23:00	1/19/12 15:00	3.58	1.40	3.48	4.89	2.17	
3535	1/19/2012 17:00	1/20/2012 0:00	1/19/12 16:00	3.68	1.55	3.48	5.03	2.31	
3536	1/19/2012 18:00	1/20/2012 1:00	1/19/12 17:00	3.89	1.60	3.48	5.09	2.37	

B.6 Quality Assurance/Quality Control of Water-Surface Elevation Data

At this stage, the water level calculations and data that went into those calculations need to be “quality assured” to ensure quality of data for any following tasks involving these data.

The quality assurance/quality control (QA/QC) log file, provided as part of the Data Exchange Template package (\\Step3_QAQC\1_QAQC_log.xlsx), provides a working example. All QA/QC steps should be listed in the QA/QC log regardless of the results. Any issues found during the QA process should be marked and commented on in the Excel file so they can be corrected for the final version. Open the QA/QC log and create a new entry for the site you are checking. Follow the format of the example QA/QC log provided.

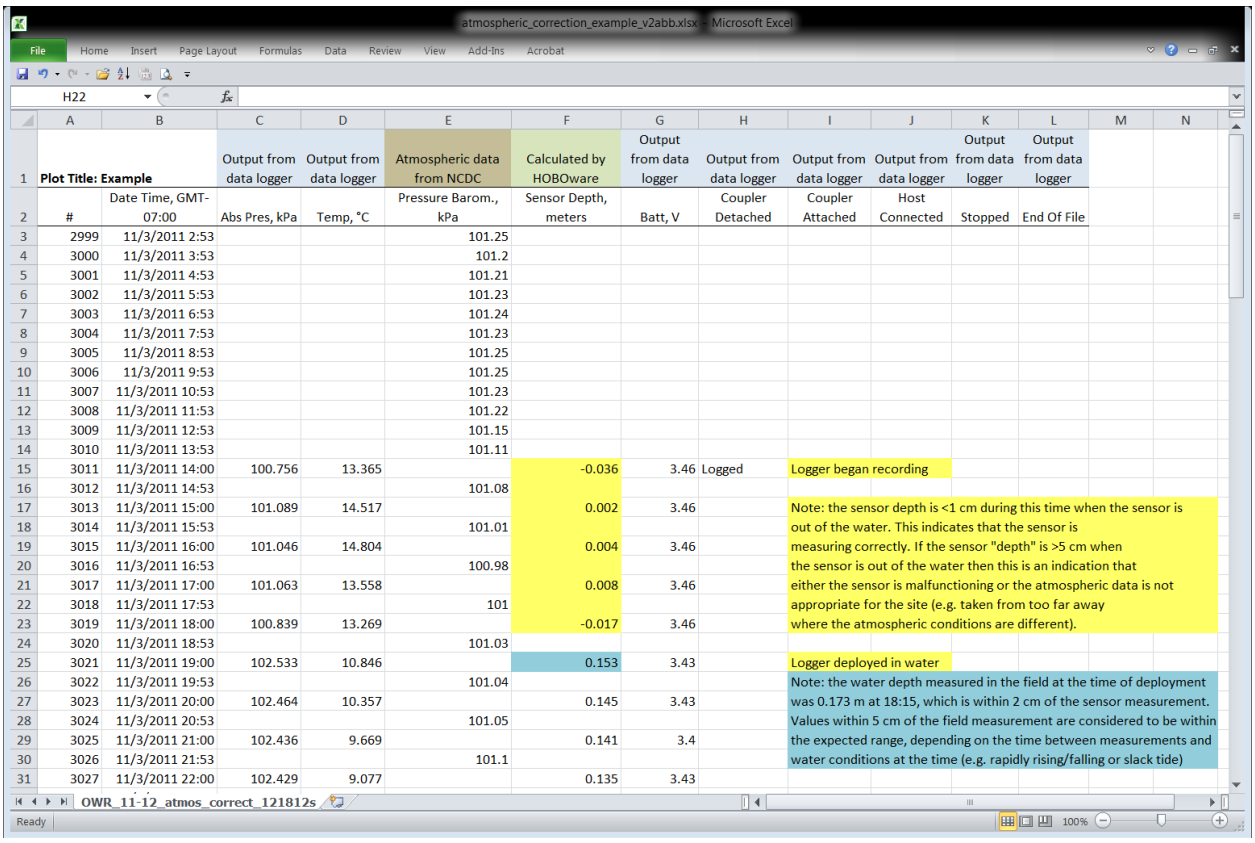
1. Atmospheric data correction check.

   The atmospheric check is made to look for sensor malfunction or variability and to ensure the atmospheric data were appropriate for the site. Pre- and post-deployment of the sensor should record a "depth" value close to 0 and not vary too much.

   Open the atm_corrected.csv file of interest from the location at \\Step2_Corrections\Step2_atm_corrected.

   Using the data in the sensor log, determine the date/times of deployment and retrieval. Check the sensor depth column before and after the dates listed for deployment. In the QA log, record the greatest value within 2 hours prior to deployment or after retrieval.

   See screen shot below and the file "atmospheric_correction_example.xlsx" for examples of what to look for in this file.



	A	B	C	D	E	F	G	H	I	J	K	L
1	Plot Title: Example		Output from data logger	Output from data logger	Atmospheric data from NCDC	Calculated by HOBOware	Output from data logger	Output from data logger	Output from data logger	Output from data logger	Output from data logger	Output from data logger
2	#	Date Time, GMT-07:00	Abs Pres, kPa	Temp, °C	Pressure Barom., kPa	Sensor Depth, meters	Batt, V	Coupler Detached	Coupler Attached	Host Connected	Stopped	End Of File
3	2999	11/3/2011 2:53			101.25							
4	3000	11/3/2011 3:53			101.2							
5	3001	11/3/2011 4:53			101.21							
6	3002	11/3/2011 5:53			101.23							
7	3003	11/3/2011 6:53			101.24							
8	3004	11/3/2011 7:53			101.23							
9	3005	11/3/2011 8:53			101.25							
10	3006	11/3/2011 9:53			101.25							
11	3007	11/3/2011 10:53			101.23							
12	3008	11/3/2011 11:53			101.22							
13	3009	11/3/2011 12:53			101.15							
14	3010	11/3/2011 13:53			101.11							
15	3011	11/3/2011 14:00	100.756	13.365		-0.036	3.46	Logged	Logger began recording			
16	3012	11/3/2011 14:53			101.08							
17	3013	11/3/2011 15:00	101.089	14.517		0.002	3.46		Note: the sensor depth is <1 cm during this time when the sensor is			
18	3014	11/3/2011 15:53			101.01				out of the water. This indicates that the sensor is			
19	3015	11/3/2011 16:00	101.046	14.804		0.004	3.46		measuring correctly. If the sensor "depth" is >5 cm when			
20	3016	11/3/2011 16:53			100.98				the sensor is out of the water then this is an indication that			
21	3017	11/3/2011 17:00	101.063	13.558		0.008	3.46		either the sensor is malfunctioning or the atmospheric data is not			
22	3018	11/3/2011 17:53			101				appropriate for the site (e.g. taken from too far away			
23	3019	11/3/2011 18:00	100.839	13.269		-0.017	3.46		where the atmospheric conditions are different).			
24	3020	11/3/2011 18:53			101.03							
25	3021	11/3/2011 19:00	102.533	10.846		0.153	3.43		Logger deployed in water			
26	3022	11/3/2011 19:53			101.04				Note: the water depth measured in the field at the time of deployment			
27	3023	11/3/2011 20:00	102.464	10.357		0.145	3.43		was 0.173 m at 18:15, which is within 2 cm of the sensor measurement.			
28	3024	11/3/2011 20:53			101.05				Values within 5 cm of the field measurement are considered to be within			
29	3025	11/3/2011 21:00	102.436	9.669		0.141	3.4		the expected range, depending on the time between measurements and			
30	3026	11/3/2011 21:53			101.1				water conditions at the time (e.g. rapidly rising/falling or slack tide)			
31	3027	11/3/2011 22:00	102.429	9.077		0.135	3.43					

2. Sensor depth verification.

   Open the Excel file (*.xls, *.xlsx) of interest found at \\Step2_Corrections\Step2_elev_corrected.

   Find the times in the Excel file closest to the times of deployment and retrieval. Compare the water level measurement in the sensor log with the sensor depth measurement in the water level file (make sure the times and time zones are the same). Note any differences in the QA/QC log (1_QAQC_log.xlsx). See screen shot above and notes highlighted in blue for an example.

   If the sensor log has notes about potential issues with this measurement, copy the note into the QA/QC log.

3. Elevation correction check.

   a. Check the depth sensor elevation value to make sure the correct value was used. Refer to the Depth Sensor Log.

   b. Check the CRD conversion value to make sure the correct value was used. Refer to the CRD Conversion workbook located at \\Step2_Corrections\Step3_CRD_Elevations\CRD conversions.xlsx.

   Also check the equations where these values are used to make sure that they are correctly calculated.

4. Check for sensor movement during deployment.

   Within the Depth Sensor Log, check for differences in the measurement of the sensor to top of post ("Dist to top of post" in Log) between deployment and retrieval. Note any difference in the QA/QC log; differences greater than 10 cm require an investigation of the water level data to see if the time of change can be determined.

5. Check the Sensor Depth column for 9999 values; there can be no negative values in this column or subsequent calculations will be incorrect.

6. Create a hydrograph.

   The last step of the QA/QC is the creation of the hydrograph to check for any anomalies or odd trends in the data.

   a. Create a new tab in the workbook and label it "graph".

   b. Copy the "Time, Local" and "WL Elev (m, CRD)" columns into the new tab. Select the relevant time period to plot.

   c. In the ribbon, go to the Insert tab and select a Scatter with Smooth lines Plot type. Select the time column for the x value and select the wl elev column for the dependent variable.

   Note any anomalies in the QA/QC log.

If there are errors in the water level file make sure they are highlighted and commented then save the file and change the file name from "elev_correct" to QA/QC with date and initials at the end of the filename in the format mmddyyii.

Example filename: SITE_QAQC_mmddyyii.xlsx
Located in: \\Step3_QAQC.

B.7 Finalization of Water-Surface Elevation Processing

This step completes the preparation of the WSE data for use in analysis or by others.

1. If uncorrectable errors are found, the QA/QC version of the file should not be finalized. If there are no errors or any detected errors are corrected, save the file in the Final folder.

   Example filename: SITE_FINAL_mmddyyii.xlsx
   Located in: \\Step4_Final

2. Delete the CRD conversion value and its label in the first row. Delete any of the date/time columns except the desired time zone. Delete columns with any conversion values. Make sure the plot title is still in cell A1. Resave the document.

The time-series WSE is the final product produced for the calculated WSEs. There are derivatives including SEV (sum exceedance value) and Area-Time Inundation Index calculations that use these data, but the time-series data show plainly what the water level is doing at the site over time.

B.8 Reference

Roegner GC, HL Diefenderfer, AB Borde, RM Thom, EM Dawley, AH Whiting, SA Zimmerman, and GE Johnson. 2009. *Protocols for Monitoring Habitat Restoration Projects in the Lower Columbia River and Estuary*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration Technical Memorandum NMFS-NWFSC-97, Northwest Fisheries Science Center, Seattle, Washington.

US Army Corps of Engineers®

Prepared for the U.S. Army Corps of Engineers, Portland District, under a Government Order with the U.S. Department of Energy Contract DE-AC05-76RL01830

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

U.S. DEPARTMENT OF ENERGY

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99352
1-888-375-PNNL (7665)
www.pnl.gov