

**State of California**  
**The Resources Agency**  
**DEPARTMENT OF FISH AND WILDLIFE**

**Coastal Salmonid Monitoring Database Management:**

**Crosswalk  
and  
Supplemental  
Attribute Tables**

**Version 2.1 using the Sampling-frame GIS<sup>1</sup> layer  
and the Aquatic Surveys<sup>2</sup> Database  
March 31, 2015**

**By**

David Dixon, Tom Christy, and Douglas Burch

---

<sup>1</sup> This document is an addendum to the Coastal Monitoring Database Management: Metadata and Protocol for the Aquatic Surveys Sampling-frame Attribute Table document (Dixon, Christy, and Burch 2015)

<sup>2</sup> The Aquatic Surveys Database, also referred to as the Aquatic Survey Program (ASP) is the current database system used by the Coastal Monitoring Plan (CMP) effort (Burch et. al. 2015).

A task of the CMP Technical Team's Sampling-frame subgroup was to come up with Sampling-frame metadata attributes. This led us to realize that there is an apparent need for some way of managing attributes that might not be conveniently placed in the Sampling-frame's GIS layer's attribute table, although these "supplemental" attributes, i.e., supplemental to the "location specific" attributes<sup>3</sup>, would likely be needed during Sampling-frame development, e.g., during quality control and, in some cases, mapping certain location specific biological information, such as "focus species by reach" (the primary species of concern within the sampling reach).

This document discusses a possible strategy whereby projects can, at various points during the Sampling-frame development timeline, use different reach location numbers or Sampling-frame draw numbers that are not the eventual Sampling-frame attributes, e.g., this strategy might be used for candidate CMP project reach location codes that are non-GRTS<sup>4</sup> in origin; or, it might be used for pre-draw temporary FrameOrder and DrawOrder numbers used in the development process, but perhaps deprecated without actually being used. In other words, we need the ability to accommodate and manage reach location codes and other attributes used before and during the Sampling-frame transition phase, i.e., for development or data collected before complete adoption of the Sampling-frame methodology, because there can be situations where location codes and other attributes change or are deprecated prior to adoption of the standardized codes/numbers used under the Sampling-frame methodology.

GRTS values, i.e., FrameOrder and DrawOrder attributes, are used as reach numbers and indexed sort values, respectively, managed by the luLocation lookup table in the Aquatic Surveys database and the Sampling-frame attribute table in the GIS layer: they are numeric; each represents a routed linear geographic feature (referred to as reaches) on stream based routed hydrography; and, their order is critical to the GRTS methodology. The task is to come up with a system or protocol whereby these temporary numbers can be updated, or adapted to different scenarios, depending on new or changed circumstances, e.g., addition or removal of reaches, etc., without confounding, or irreversibly losing legacy location attributes. We think the solution is to maintain project level "crosswalk" tables which can be linked to the ASP database, or the Sampling-frame's GIS layer, in order to reference more than one set of reach or draw sort index numbers which can represent design changes over time or during the development life cycle of a specific project's Frame<sup>5</sup>. These crosswalk tables would reside with the specific candidate project, but eventually could be gathered and maintained collectively by the statewide GIS Data Manager.

At the project level, suggested protocol for using a crosswalk table would be that the ASP luLocation lookup table's "LocationCode" field always maintain the current location code or reach number being used by the project, and the crosswalk table be used to access legacy reach attribute values. Regardless of whether the ASP database's LocationCode attribute is unique or not, the "LocationID" field is the primary key field of the luLocation lookup table, and as such is ALWAYS globally unique and never changes. Therefore, the LocationID attribute is our "Rosetta Stone" which we use to reference different sets of location codes (reach numbers) and/or Sampling-frame draw numbers. Because the LocationID attribute is constant after creation, there is no need to maintain different ASP attribute fields in the luLocation

---

<sup>3</sup> The term location specific as it is used here means GIS attribute elements which are descriptive only to the location itself, e.g., length of the stream segment would be a specific description about the location itself while descriptions regarding the degree to which that length of stream provides refuge for Coho salmon are tangential to describing the location. Although not trivial to the sampling-frame's spatial data, all tangential elements are considered "supplemental" and are more efficiently captured by way of related (linkable) attribute tables.

<sup>4</sup> "GRTS" is the acronym and short name for the statistical methodology referred to as "Generalized Random Tessellation Stratified" sampling (McDonald 2003) and implemented as a strategic component of the coastal monitoring effort.

<sup>5</sup> "Frame" implies a subset of the Sampling-frame universe specific to a single project, and this shorthand terminology is most often used when referring to portions of the Sampling-frame that are still under development.

lookup table, or the Sampling-frame’s GIS attribute table, i.e., to manage the various legacy reach codes or in-development Sampling-frame draw numbers.

At the program level, i.e., statewide-CDFW and interagency partners, ideally the GIS Data Manager would maintain copies of the project level reach crosswalk tables, and the Biogeographic Data Branch would be responsible for acquiring these tables and naming them according to a convention that self-documents their origins, perhaps providing a suffix to the table name for clarity, e.g., “ReachCrosswalkProjID101.” A unique table naming convention is necessary, because some legacy reach location codes might be deprecated from the ASP, or never get used, i.e., not referenced by any data. This being the case, the ASP database’s luLocation lookup table, and the Sampling-frame’s GIS layer, would have no record of these “orphaned” reach numbers or in-development sample-frame temporary draw numbers.

The Figure 1 diagram illustrates relationships between the ASP database’s luLocation lookup table, the Sampling-frame’s GIS attribute table, two of the supplemental<sup>6</sup> tables, and a possible crosswalk table, i.e., the “ReachCrosswalkProjID101” table; what follows are field definitions for the example crosswalk table.

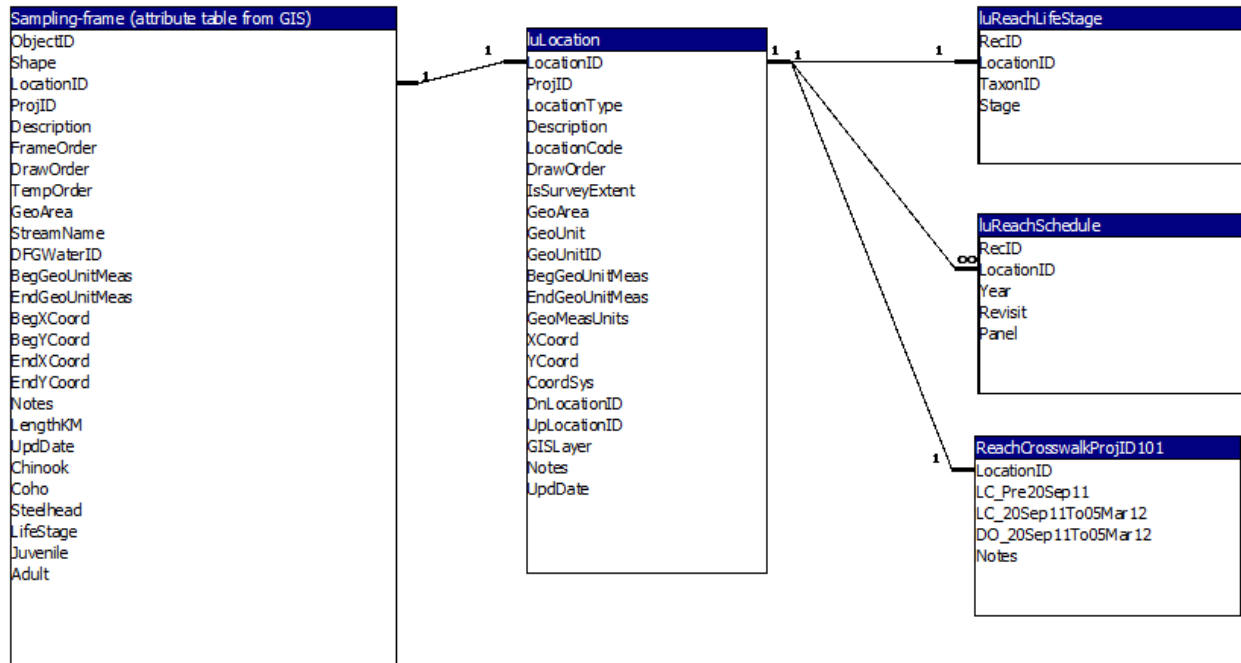


Figure 1. Sampling-frame attributes and relationship to luLocation lookup table and supplemental attributes.

<sup>6</sup> Field definitions for the “luReachLifeStage” and “luReachSchedule” lookup tables can be found in the data dictionary for the ASP database.

**ReachCrosswalk** (table) – Crosswalk table used by a project to manage any legacy reach identifiers, frame order numbers, or draw order numbers that have been created for the project regardless of whether they have ever been used or referenced by any data collected during field operations.

**LocationID** – This attribute has the same definition as that used by the ASP database’s luLocation lookup table, i.e., the same name, data-type, and definition.

**LC\_DDMonthYYToDDMonYY** – The LC refers to LocatinCode and the *DDMonthYY* represents a date formatted as, for example, 04May12, i.e., days and years would always be two characters and the month would always be three characters and in “camel” case, e.g., Jan, Feb, Mar, Apr, etc. So LC values used between September 20, 2011 and March 5, 2012 would be “LC\_20Sep11To05Mar12”, and the field name would be self-documenting -- field data type would be the same as the ASP database’s LocationCode attribute, i.e., a 10 character text field or column in the table. Metadata for this field would be managed by way of the ReachCrosswalk table’s “Notes” field, e.g., notes or log entries.

**LC\_Pre20Sep11** – Original reach location numbers used before a candidate project which began on September 20, 2011.

**LC\_20Sep11To05Mar12** – Reach location codes used while still a candidate project, i.e., between September 20, 2011 and March 5, 2012.

**DO\_20Sep11To05Mar12** – Temporary draw order numbers used while still a candidate project between the time period of September 20, 2011 and March 5, 2012.

**Notes** – Notes provided to document or explain purpose of the attributes, i.e., metadata that the project lead uses to document or explain what the location code or Sampling-frame number was used for, or why it was deprecated.

## References

Dixon, D., Christy, T., Burch, D. Coastal Salmonid Monitoring Database Management: Metadata and Protocol for the Aquatic Surveys Sampling-frame Attribute Table. California Department of Fish and Wildlife; 3/31/2015. [Cited 2015 March 31]. Available from: <https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=100933>

Burch, D., Gallagher, S. Ricker, S., Bellmer, R. Coastal Salmonid Monitoring Database Management: Report and Concept for Success. California Department of Fish and Wildlife; 3/31/2015. [Cited 2015 February 19]. Available from: <https://nrm.dfg.ca.gov/FileHandler.ashx?DocumentID=92766>

McDonald, T. 2003. GRTS for the average Joe: A GRTS sampler for windows. West Inc., Cheyenne, Wyoming. 10pp. Available: [www.west-inc.com/reports/grts.pdf](http://www.west-inc.com/reports/grts.pdf).