# Technical Appendix B: GIS Methodology

1. **Overlaying the Impact Segments with Available GIS Data Sets**
2. **Calculating Flood Depth in the Skagit Basin**
3. **Determining Which Flooding Scenario Contributed to the Greatest Depth for a Given Flood Year**

## Overlaying the Impact Segments with Available GIS Data Sets

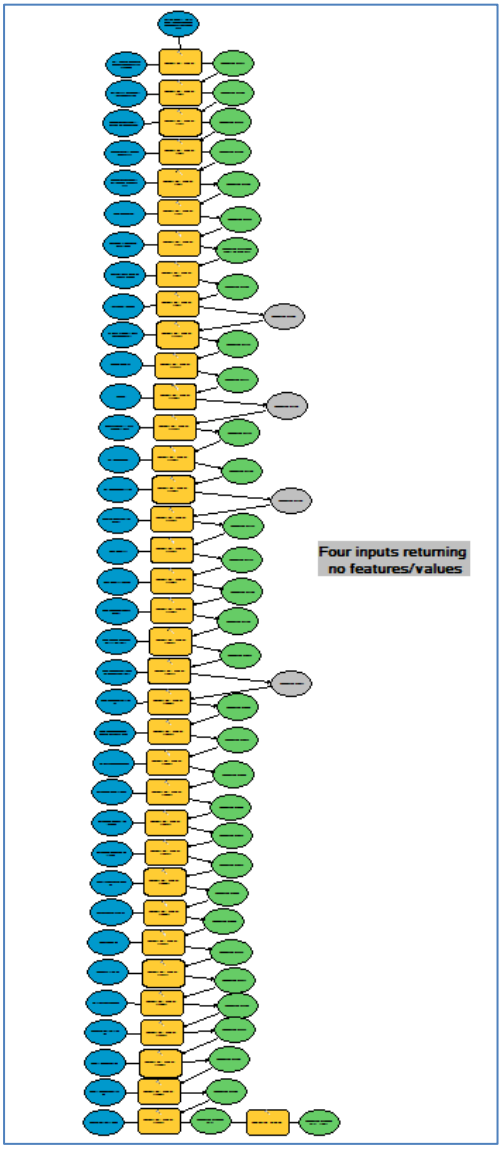
Elizabeth Lanzer, November, 2014

The process described below created tables that were simplified and interpreted by Mark Maurer to create the Segment characterization narratives provided in Appendix X (Profiles) of this report.

The project team identified datasets that would impact their evaluation of adaptation needs and options. They asked for some specific WSDOT datasets, and also asked for other data by general topic. To meet these needs, I used my 15 years of experience supporting WSDOT's environmental assessment work with GIS data and analysis, to pull over 35 data themes from the agency's 700+ dataset GIS library. These 35+ datasets were clipped to western Skagit County and saved to a local file geodatabase to provide for efficient geoprocessing.

The impact segments were defined using the further extent of impacts across all flooding scenarios as interpreted by Simon Page. These road centerlines were buffered by 200 feet (with flat ends to avoid intersections). 200 feet was used based on accuracies of the datasets being overlaid (1:12,000 – 1:100,000) and the desire to include rather than exclude nearby conditions.

**Figure 1 Primary overlay geoprocessing model (not intended to be legible) showing how characteristics from 30+ layers were compiled onto the impact segments.**




The primary overlay consisted of about 35 cascaded Spatial Join operations (see graphic that is not intended to be legible but presented to illustrate the how the data was compiled into one feature class). Fields were dropped in the Spatial Join to leave only the key characteristic(s) that informed adaptation evaluation in the output. Merge rules were used to summarize multiple value returns. The feature attribute table was exported to Excel. In Excel, two tabs were prepared – one with the Impact segments as columns, and one with the Impact segments as rows.

Overlays for the soils groups and types were overlaid separately to provide a better characterization of soils types in a segment by % of a soils group like Hydrologic D or Hydric. The Soil Survey Geographic Database (SSURGO) soils data for Skagit County was clipped to the impact segment buffers, and then Statistics were run to summarize the area of each soils group within each impact segment.

To address "outlier" values in the primary overlay result, "Select by Attribute" then "Select by Location" tools were run against the highway characteristics datasets (Traffic, Federal Functional Class & Freight Class). The features on the impact segment highways were selected out by State Route attributes, then, re-selected for the individual segment's spatial extent to create tables showing the range of values falling within the segment buffer, but only on the segment highways. This was needed due to Segments 2 and 10 where the buffer overlay captured values from I-5 when those segments intersected I-5.

### GIS Data Input Screening

To help put the impact segments in context with the USACE Study and other local planning:

- Damage Reaches from the USACE Geotechnical Investigation (GI)
- County Comprehensive Plan Zoning

To characterize the Highway's core transportation functions:

- Average Annual Daily Traffic
- Truck percentage
- Design High Volume
- Federal Functional Classification

- Local Transit Routes
- Park & Rides
- Structures (bridges and under crossings)
- WSDOT Facilities (Sites, Buildings, Fuel Stations)

To characterize the drainage issues and infrastructure currently in place:

- Culverts & Culvert Ends (can't tell which ends go to the same pipes)
- Stormwater Discharge Points
- Approved stormwater treatments including
    - Ponds
    - Ditches
    - Slopes
    - Flow Restrictors
    - Energy Dissipaters
- FEMA Flood Zone / Floodway
- Chronic Environmental Deficiencies (locations requiring repeated maintenance due to recurring location based conditions that are typically drainage related – fixed or yet to be fixed permanently)
- Fish Passage Site Inventory (barriers, fixed structures, and non-barrier river/road crossings)

To characterize soils and geological conditions, other known hazards:

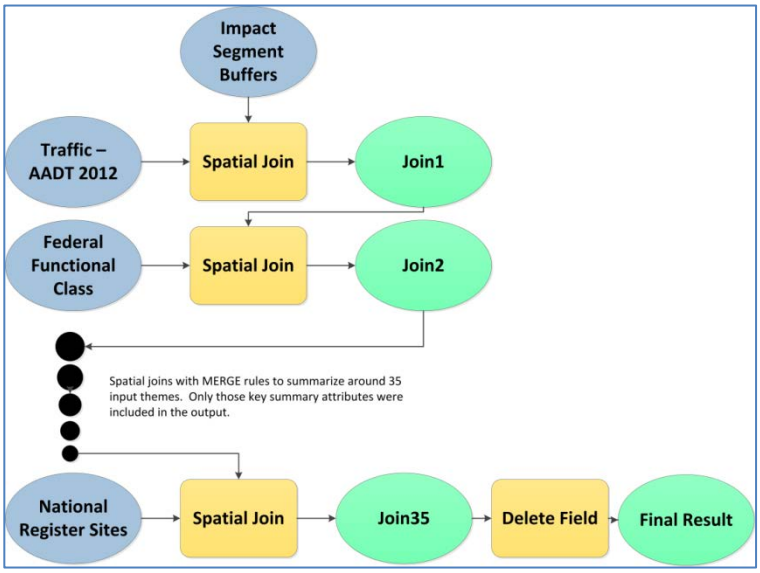
- NRCS Soils (SSURGO)
    - Hydrologic Groups
    - Hydric
- Landslides
- Unstable Slopes Along State Highways
- Liquefaction Susceptibility
- Tsunami Inundation / Evacuation Zones
- WSDOT Climate Impact Vulnerability Assessment ratings

To identify resources or conditions that should be protected or avoided:*

- Hospitals
- Firehouses
- Cemeteries
- Known National Register Sites (including barns)
- Confirmed or Suspected Contaminated Sites

*We wanted to get life-line routes, but weren't able to get to the right source in time for this analysis.

Alternative graphic for primary overlay:



## Calculating Flood Depth in the Skagit Basin

Stu Smith, October, 2014

This process describes how the maximum flood water depth, by flood year, for each raster cell was calculated.

Input data consisted of two sources:

1. Arc Grids from the Corps of Engineers that depict the flood water surface elevation (WSE) on a 400' X 400' cell size, based on a series of flood models based on two factors:
   - 10-, 25-, 50-, 100-, and 500-year floods
   - The location of where levees would breach ("scenario")
   - The grids were named in the form XXwYY, where XX = flood year (e.g., 25), and YY = scenario (e.g., 464, 1239)

2. A lidar-based Grid terrain model of the earth's surface on a 6' X 6' cell size: skagitmosaic

Step 1: Select – by flood year – the maximum water *surface elevation* at each 400' X 400' cell from among the various scenarios:

- The "Cell Statistics" Tool was used. Input grids were the several Corps of Engineers scenarios for that flood year (as described in #1 above), with the Overlay statistic = "maximum" and the "Ignore NoData" option checked on.
- The output was the maximum flood water surface elevation grid for each flood year (10 year was not calculated since there was only one scenario for that year), stored in max_surface_elevation_**XXX**_year_flood, where XXX = that flood year, e.g., 10, 25, 50 100, and 500.

Step 2: Calculate – by flood year – the water *depth* using the 6' X 6' cell size of the terrain model in #2 above.

This was calculated with the Raster Calculator tool. The algebraic expression is:

- (max_surface_elevation_**XXX**_year_flood – skagitmosaic)
- The output raster is: max_depth_**XXX**_year_flood
- Environment settings for the Raster Calculator tool were set:
- Raster Analysis > Cell Size > same as layer skagitmosaic
  - This insures that the output grid will have the 6' x 6' cell size from skagitmosaic, rather than the 400' x 400' cell size from the max_surface_elevation_**XXX**_year_flood grids.

These results were posted to a file geodatabase depth.gdb containing ten rasters:

Five rasters created by Step 1 depict the maximum water depth, at a 6' X 6' cell size, that occurred among the various scenarios for each flood period. Depth was calculated by subtracting the earth's elevation (contained in the raster skagitmosaic) from the maximum flood surface elevation among the several flooding scenarios (the 10-year flood had only one scenario) for each flood period.

max_depth_10_year_flood
max_depth_25_year_flood
max_depth_50_year_flood
max_depth_100_year_flood
max_depth_500_year_flood

Five 400' X 400' flood surface elevation rasters created by Step 2 are also located in the geodatabase as:

max_surface_elevation_10_year_flood
max_surface_elevation_25_year_flood
max_surface_elevation_50_year_flood
max_surface_elevation_100_year_flood
max_surface_elevation_500_year_flood

Example, using the 50-year flood:

max_depth_50_year_flood = (max_surface_elevation_50_year_flood – skagitmosaic)

## Determining Which Flooding Scenario Contributed to the Greatest Depth for a Given Flood Year

Stu Smith, October, 2014

The process described below is a follow-up to the analysis that calculated the greatest flood depth for each flood year.

DOT scientists were not interested in 500-year floods, so analysis for that flood year was not attempted. Analysis was done for the other flood years: 10, 25, 50 and 100.

Tools used:

- Raster Calculator
- Cell Statistics
- Create Mosaic Dataset
- Add Raster to Mosaic Dataset
- Extract by Attributes
- Build Raster Attribute Table
- Set Null
- Highest Position
- Append

Prior analysis had determined the greatest flood depth that would occur at each cell, by flood year. However, DOT scientists were also interested in which flood scenario contributed to a given cell's greatest depth. Highest Position is the appropriate tool for determining which scenario contributed to the maximum depth. Unfortunately, simply running Highest Position is not sufficient because it will only output a value for those cells where ALL scenarios overlap (think Venn diagram). Where one or more scenarios do not overlap, Highest Position outputs nodata, which is unacceptable. This document describes the process for determining which scenario was responsible for the greatest flood depth for all overlap combinations.

The solution is to run Highest Position multiple times, once for each overlap combination. Here's how to identify each unique overlap combination, using the four scenarios for the 25-year flood:

| Scenario | Value |
|---|---|
| 464 | 1 |
| 1239 | 2 |
| 1379 | 4 |
| 2159 | 8 |

Using the 1, 2, 4, 8… sequence assures that the scenarios will be uniquely represented in subsequent analysis. Adding the "Values" for each scenario combination results in a unique "combination value". Here are the possible combinations for the 25-year flood (for example, 464 by 1379 results in 1 + 4 = 5):

| 2-way | Combination value |
|---|---|
| 464 by 1239 | 3 |
| 464 by 1379 | 5 |
| 464 by 2159 | 9 |
| 1239 by 1379 | 6 |
| 1239 by 2159 | 10 |
| 1379 by 2159 | 12 |

| 3-way | Combination value |
|---|---|
| 464 by 1239 by 1379 | 7 |
| 464 by 1239 by 2159 | 11 |
| 464 by 1379 by 2159 | 13 |
| 1239 by 1379 by 2159 | 14 |

| 4-way | Combination value |
|---|---|
| 464 by 1239 by 1379 by 1259 | 15 |

- (Note that the 4-way overlap creates the same result as using the default Highest Position, described above.)

Thus, every 1-, 2-, 3-, and 4-way overlap combination is uniquely represented by the numbers 1 through 15.

Flood years (such as 50) with just two scenarios, will use only the values 1 and 2, with 3 being the only possible combination value. Flood years with more scenarios will extend the 1, 2, 4, 8 sequence described above. For example, the 100-year flood has seven scenarios, assigned values of 1, 2, 4, 8, 16, 32, 64, with combination values ranging up to 127 (1 + 2 + 4 + 8 + 16 + 32 + 64 = 127). It is important that the values be consistently assigned in ascending scenario number order among all flood years; that is, the scenario with the lowest number, say 464, is assigned value 1, and the second-lowest number is assigned value 2, and so forth.

The analysis process starts by assigning the "values" to each scenario flood depth raster (e.g., depth_XXw_YY_int, where XX = flood year, such as 25, and YY = scenario, such as 464) that had been created in the preceding analysis, outputting the result to a new raster, named depth_XXw_YY_int_ZZ, where ZZ = scenario value:

(depth_25w_464_int * 0) + 1 = depth_25w_464_int_1

(depth_25w_1239_int * 0) + 1 = depth_25w_1239_int_2

(depth_25w_1379_int * 0) + 1 = depth_25w_1379_int_4

(depth_25w_2159_int * 0) + 1 = depth_25w_2159_int_8

- Note that multiplying the input raster "resets" it to zero, and the ensuing addition calculates every cell throughout the raster to that integer value.

Now, by adding the rasters, a sum of combination values can be created for each cell.

Use the Cell Statistic tool to generate a single output raster for each flood year.

- The output rasters from above are the inputs
- Overlay statistic is "sum"
- Ignore nodata is checked on.
- Environments > Processing extent = "union of inputs"
- Output name = depth_XXw_int_cellstats, where XX = flood year (for example, the 25-year flood name would be depth_25w_int_cellstats)

A polygon view of the overlap combinations for each flood year can be generated by 1) creating a Mosaic Dataset and 2) adding all the scenario depth rasters to the mosaic dataset. Then look at the polygon footprints to see which scenario combinations exist, along with the total number of scenario combinations. This two-step process starts with the Create Mosaic Dataset tool, used once per flood year:

- Output location = …\depth_scenario.gdb
- Mosaic Dataset Name = depth_XXw_mosaic_dataset, where XX = flood year
- Coordinate system = NAD83, State Plane North, FIPS 4601, feet

After creation, the Add Rasters to Mosaic Dataset tool is used, with:

- Raster Type = Raster Dataset
- Input Data = Dataset
- Source = the flood depth rasters for that flood year (e.g., depth_XXw_YY_int, where XX = flood year and YY = scenario)
- Update Cell Size Ranges and Update Boundary boxes are checked on

Continuing with the 25-year flood example, there are four polygon footprints with the combination values 1, 3, 12, 15. Note that not every possible combination will necessarily exist within each flood year; in this 25-year flood example only 4 out of a possible 15 combinations exist.

Next, the separate scenario combinations need to be "exploded" into separate rasters, for later use as masks. No "explode" tool exists, so the process must be done individually using the Extract by Attributes tool:

- Input is the Cell Statistics raster from above (e.g. depth_25w_int_cellstats)
- Where clause selects the appropriate combination value, e.g. "Value = 3"

- The output raster name is in the form of depth_XXw_gNN, where XX = the flood year, g stands for "grid", and NN= combination value. For example, depth_25w_g3 is the raster for the 25-year flood where scenarios 464 and 1239 overlap (hence 1 + 2 = 3).
- This is repeated for each combination value. In the 25-year flood, here are the outputs:
  depth_25w_g1
  depth_25w_g3
  depth_25w_g12
  depth_25w_g15

Each of these output grids, therefore, simply delineates the rectangular grid extent covered by that particular scenario combination. However, within that extent there will likely be cells where all input layers are nodata. Such areas are of no interest since there is no possibility that any of the input layers would be responsible for the highest flood level. The following describes the process for determining what cells are all nodata:

The process ingests the depth rasters appropriate to a given flood year / scenario and sums the depth values. Previously, the nodata values had been converted to -100,000,000, so cells with large negative numbers were nodata. The size of the negative values indicated the number of overlapping depth rasters. For example, -200,000,000 indicated that two rasters were involved (for example, scenarios 464 and 1239). -300,000,000 would indicate three rasters were involved, and so on.

Note that some raster depth values were actually negative (below sea level???); however, none of them ever approached the magnitude of the -100,000,000 value.

The summation was done with the Raster Calculator tool. Here's an example statement:

depth_25w_464_int + depth_25w_1239_int

The output raster is named sum_XXw_gNN, where XX is the flood year and Y is the scenario combination. Non-overlap cells are then converted to nodata with the Set Null tool with these parameters:

- Input conditional raster = sum_XXw_gNN (the output raster from the sum, directly above)
- Expression = the summed negative value according to the number of overlapping grids (for example, -200,000,000).
- Input false raster or constant value = Input false raster or constant value
- Output raster = maskXXw_gNN, where XX is the flood year and NN is the scenario combination

These above two steps (summation and set null) are contained in models hard-coded for each flood year / scenario combination. These models are named using the form summation_XX_gNN, where XX = flood year and NN = combination value. Each model has the following environment setting:

Model>Model Properties>Environments>Raster Analysis>Mask = depth_XXw_gNN. This mask raster, created above, is unique to each model and therefore is appropriately hard-coded for each.

The Set Null tool should, by default, create an attribute table for the output raster maskXXw_gNN. However, there were times when this did not occur. Therefore, at this point, inspect each of the output rasters and use the Build Raster Attribute Table tool if necessary.

The Highest Position tool is run next, once for each scenario combination. Manually enter the depth layers appropriate to a given scenario combination into the tool in a specific order. The entry order is critical, and must be consistent across all scenario combinations. Depth layers were entered according to their scenario number, in numerical order, with the smallest number first. For example, using the 25-year flood, 464 X 1239 combination, the depth layers were entered into the tool in this order:

depth_25w_464_int
depth_25w_1239_int

Since 464 is a smaller number than 1239, it was entered first.

Here's another example using the 25-year flood, 464 X 1379 X 2159 scenario combination (combination value= 13). The depth rasters were entered in this order:
depth_25w_464_int
depth_25w_1379_int
depth_25w_2159_int

The Highest Position tool's output raster is named:
highest_scenario_XX_year_floods_gNN, where XX is the flood year and NN is the scenario combination

The Highest Position tool's Environment Settings are modified, each time the tool is run, as follows:
Environments>Raster Analysis>Mask>mask_XXw_gNN. This is the raster mask from the previous step, and its presence in the Highest Position tool's environment ensures that only non-nodata cells are output.

Then build a .vat with the Build Raster Attribute Table tool.

Note that if all scenarios have the same depth for a given cell, the Highest Position tool selects the first scenario in the input list by default. This was not a problem for the scientists when this work was done in 2014.

The final step is to merge all of the highest_scenario_XX_year_floods_gNN rasters, from the previous step, into one overall raster for that flood year. The output is named in the form:
highest_scenario_XX_year_floods, where XX = a given flood year

Ideally, the Mosaic to New Raster tool should work, but it fails, even with extensive troubleshooting review. Sadly, the Mosaic tool fails as well.

Append is the only tool that worked, but it required some caution:

1. Append combines input rasters into an existing target raster. For consistency, the lowest numbered highest_scenario_XX_year_floods_gNN rasters (from the prior step) were appended into the highest numbered raster. For example, for the 25-year flood combinations, these rasters:
   - highest_scenario_25_year_floods_g1
   - highest_scenario_25_year_floods_g3
   - highest_scenario_25_year_floods_g12

   were appended to:
   - highest_scenario_25_year_floods_g15

2. Since the Append tool overwrites an existing raster (in the above example highest_scenario_25_year_floods_g15 is overwritten) its original version needs to be preserved. This was accomplished by copying it to a remote .gdb prior to the append. After the append has completed, the overwritten raster was renamed and then the original raster was moved back to the working .gdb. Here's an example workflow:
   a. Copy highest_scenario_25_year_floods_g15 to a remote .gdb
   b. Append rasters to highest_scenario_25_year_floods_g15
   c. Rename highest_scenario_25_year_floods_g15 to highest_scenario_25_year_floods
   d. Move highest_scenario_25_year_floods_g15 from the remote .gdb to the working .gdb.

After the append process is complete, the text field "scenario" was added to the highest_scenario_XX_year_floods raster's .vat and populated manually, using the field calculator, with the text values for each scenario. For example, here's the .vat contents for the raster highest_scenario_25_year_floods:

| Objectid | Value | Count | scenario |
|---|---|---|---|
| 1 | 1 | 1497864 | Depth_25w_0464_int |
| 2 | 2 | 9457265 | Depth_25w_1239_int |
| 3 | 3 | 9022531 | Depth_25w_1379_int |
| 4 | 4 | 3092864 | Depth_25w_2159_int |

The scenario values' purpose is to identify and/or label each cell with the scenario that contributed to the highest depth for a given flood year.

The highest_scenario_XX_year_floods rasters are the final output product, and are intended for use by the scientists. They depict which scenario is responsible for the maximum depth at each cell, by flood year

The resulting rasters that identify which scenario within a flood year (10, 25, 50 and 100) contributed to the highest water level, per 6' X 6' raster cell across the Skagit basin:

highest_scenario_10_year_floods
highest_scenario_25_year_floods
highest_scenario_50_year_floods
highest_scenario_100_year_floods

These grids were overlaid on the maximum water level data grids. That enabled determining both the maximum water height and the scenario that contributed to it, by flood year.

Two considerations with this data:

1. There was only one scenario for the 10-year flood, thus it is the only contributor for that flood year.
2. If two scenarios contributed equally to a cell's maximum water level within a flood year, the scenario with the smallest "scenario number" is recorded as the highest. For example, in the 25-year flood, scenario 464 would be given preference over scenario 1379.

This page intentionally left blank