

Open Space for the Snohomish River

Jacob Kovacs, Richard McGovern, and Kristina Rose
Geography 482 - GIS Data Management
December 8th, 2013

Abstract

The health of the Snohomish River directly affects the health of surrounding ecosystems, and the health of humans who inhabit those ecosystems. In order to protect the health of the Snohomish River, our project team created a geodatabase to support (1) the identification of key open and impervious spaces in Snohomish County, and (2) the selection of appropriate runoff reduction strategies for these spaces in accordance with Best Management Practices. This project is an instance of the green infrastructure approach called for in the 2010 Regional Open Space Strategy (ROSS), and, as such, represents a contribution to the stated mission of the ROSS Project: to reduce runoff and the spread of pollutants using ecologically friendly means. Our geodatabase provides a platform that could be used by a skilled GIS analyst to answer our research question, and thereby inform the planning efforts of authorities in Snohomish County.

Introduction

Purpose of this project

The overarching goal of this project was the creation of a geodatabase that would allow users to answer a research question relating to “green infrastructure” (defined below). We had broad discretion in selecting a specific focus and audience for our project. After some exploration of the literature, we developed a research question that focused on the contribution of impervious surfaces to stormwater runoff, and the potential use of open spaces as a type of green infrastructure for stormwater management. Our research question further focused on managing stormwater in relation to the health of the Snohomish River. Finally, we chose planning authorities in Snohomish County as the target users for our geodatabase, assuming some technical proficiency with GIS analysis on their part.

Project background

To some degree, stormwater runoff occurs whenever rain falls. In an urban context—with the existence of so many impervious surfaces, like parking lots, roads, and rooftops—the phenomenon is exacerbated, and does significant damage to nearby waterways (Schueler, 1994). It is the content and volume of water that does this damage: water carries urban pollutants, while high-volume flows cause erosion, channel instability, and flooding. The harsh ecological impact of urban runoff makes stormwater management necessary, and the major role of impervious surfaces in producing runoff means that management of impervious surfaces is a crucial task for urban water quality improvement.

The traditional “gray” infrastructure approach to urban runoff management relies on a system of pipes to drain water from urban areas into waterways, perhaps filtering the water in treatment facilities at some point along the way. This approach reduces the negative impact of rainfall on urban areas largely by *shifting* that impact to waterways. “Green” infrastructure, on the other hand, aims to reduce the ecological impact of urban runoff by “weaving natural processes into the built environment” (EPA, n.d.).

This interweaving harnesses the impressive filtering and absorptive powers of plants and soil, to reduce the volume of runoff and slow the transport of pollutants. Some examples of this “interweaving” are rain gardens, vegetated swales, and permeable pavements, each of which use plants or soil to absorb water, removing or delaying its entry to the gray infrastructure system and nearby waterways.

One instance of green infrastructure for runoff management is “green” or “open space”, a common term referring to protected/conserved land, areas without human-built structures, and other public spaces that function as runoff mitigation areas (ROSS, 2010). Alongside the cultivation of open space, other strategies for managing impervious surfaces include various types of “disconnection” and the installation of permeable pavement (CAP, 2012). Planners’ choice of management strategy should be guided by land use and site characteristics. For example, impervious surfaces in residential areas with Hydrological Soil Groups (HSG) Type A are best dealt with through “simple disconnection”, which entails routing runoff to nearby pervious areas for absorption (CWP, 2012).

Research and analysis questions

Inside this broad problem space of urban runoff management and green infrastructure solutions, we chose the following research question: in order to protect the health of the Snohomish River, what impervious areas should be targeted by different runoff reduction strategies? Our project focuses on the Snohomish County Water Resource Inventory Area (WRIA 7), and our main concern is the health of the Snohomish River, including its outlet at Port Gardner Bay. Within this scope we seek to identify impervious areas that are good candidates for different runoff reduction methods, including the runoff reduction method of cultivating open spaces.

To answer this research question, our database must support multiple inquiries. We decomposed our primary question into smaller questions, to be answered through spatial analysis (see Table 1). This decomposition helped us determine more precisely what data should be included in our geodatabase, and also helped us judge the feasibility of our initial research question.

- | |
|--|
| <ol style="list-style-type: none"> 1. What area drains into the Snohomish River? 2. What potential open spaces lie within this identified area? 3. What impervious surfaces lie within this identified area? 4. Which of these identified impervious surfaces are good candidates for different runoff reduction methods? For example: <ol style="list-style-type: none"> a. Which parking lots are government-owned, and therefore easier to mandate their replacement with permeable material? b. What are the landuse categories associated with identified impervious surfaces, and what does this imply for our choice of runoff reduction method? |
|--|

Table 1: *Spatial analysis questions*

Potential beneficiaries of this project

Currently the Snohomish River is monitored by Snohomish County Surface Water Management, the Washington State Department of Ecology (WA DOE), and the Environmental Protection Agency (EPA). The cities of Monroe, Snohomish, Marysville, and Everett are also main

players in the monitoring process. The present focus of all these agencies with regard to water pollution is on point-source pollution (transported through the sewer system) and non-point source pollution, a type typically harder to manage but one which our green infrastructure specifically addresses. Of this list of stakeholders, our geodatabase is likely to be most useful to analysts working for Snohomish County. City governments probably lack the technological infrastructure to make use of GIS, and the EPA and WA DOE are not immediately concerned in the day-to-day management of stormwater at the county level. Beyond our target users, likely beneficiaries of this project are all Snohomish County residents (in a very diffuse way, owing to ecological interconnectedness).

Project report outline

The remainder of this project report will cover our project methods (including data acquisition, database design, and database implementation); different challenges we encountered during database implementation, and how we were able (or not) to deal with these challenges; our project findings and conclusions; and, finally, our recommendations to others who are interested in conducting a similar study. More detailed information referenced in several sections is included in the appendices at the end of the report.

Project methods

Database design and construction

While there is some overlap with Yeung and Hall's (2007) "Systems Development Life Cycle" and "Database Development Life Cycle", in the end our workflow was tailored to suit the demands of the project. For instance, we couldn't do a thorough initial study for the database, which would have required an assessment of user needs. Nor was the database operationalised, which would have entailed user education and training. These modifications arose from the fact that our project lacked real-life users. Table 2 presents an overview of the project workflow as adapted to our constraints.

<i>Project planning</i>	<i>Geodatabase design</i>	<i>Geodatabase implementation</i>
<ul style="list-style-type: none"> ● Define channels of communication ● Establish spaces for collaboration ● Allocate tasks and set deadlines 	<ul style="list-style-type: none"> ● Conceptual ● Logical ● Physical 	<ul style="list-style-type: none"> ● Gather, assess, and document data ● Revise database model as needed ● Build geodatabase

Table 2: *Overview of project workflow*

During the project planning phase, we laid the foundation for subsequent project activities by defining our channels of communication and establishing spaces for collaboration. The bulk of our collaboration occurred through documents on Google Drive. The capability to jointly edit and comment was crucial to creating products that captured everyone's input, when we weren't able to spend many

hours together in person. Even when we were working side by side, in fact, we would often be working in the same Google Drive document. Cloud computing gave us a quick way to collaborate without having to reconcile different versions or stitch together multiple documents. Another factor enabling us to collaborate productively was that we always took time to establish an unambiguous division of labor and clear deadlines---a step that sounds simple, but is easy to overlook or dismiss as ‘too formal’ for short lab assignments.

With our project planning complete, we started into the actual process of database design (Table 3), proceeding in three stages: conceptual, logical, and physical design. Conceptual design relied heavily on our exploration of related literature. Once we adequately surveyed the problem space, we were able to generate a coherent, focused research question, then decompose it into narrower questions that corresponded to specific spatial analytic operations. Our logical design emerged from this conceptual work; we translated important nouns from our questions into the logical vocabulary of ArcGIS software, identifying necessary *feature classes* and grouping them into thematically or spatially related *datasets*. To these feature classes and datasets, we added attributes, rasters, topologies, and networks as needed. Finally, the physical design stage of geodatabase design was fairly straightforward. Difficulties with the server led us to implement our geodatabase as a Personal GDB (rather than through ArcSDE), and ArcCatalog shields users from having to consider the intricacies of the physical aspect of their geodatabase¹. We did, however, give some thought to the overall size of the database, deciding to clip data that exceeded the geographic extent of our project. We also backed up the finished geodatabase on a jumpdrive.

<i>Conceptual</i>	<i>Logical</i>	<i>Physical</i>
<ul style="list-style-type: none"> ● Become familiar with problem space ● Generate research question ● Decompose research question into several questions 	<ul style="list-style-type: none"> ● Identify desired datasets and feature classes ● Identify desired attributes, rasters, topologies, and networks 	<ul style="list-style-type: none"> ● Personal GDB---didn't use ArcSDE (issues with server) ● Clipped to reduce required storage space ● Copied finished GDB onto jumpdrive

Table 3: *Overview of geodatabase design stages*

Despite the careful thought that went into our three phases of design, implementation turned out to be an iterative process, with unanticipated difficulties leading to substantial departures from our logical geodatabase design (summarized in Appendix 1). For example, we removed some feature classes that weren't available as such, but could be produced later from existing feature classes through spatial analysis. We also removed feature classes that turned up as attributes inside another feature class. Finally, we added feature classes that seemed potentially relevant to our research or analysis questions, even though we had not identified them during our conceptual design stage.

¹ We regretted the need to abandon working in a server environment, but reasoned that our target audience would likely be working from a Personal GDB as well, not from a server. Thus, we think that this technical difficulty didn't undermine the purpose of our project.

We built the geodatabase by gathering, assessing, and documenting data, storing it on one group member's H drive in the geography lab. When all data was located, evaluated, and recorded, we created feature datasets in ArcCatalog and then imported our feature classes. Finally, rasters were added, and several topologies were defined.

Data sources

We acquired most of our data---including zoning, rainfall, soil characteristics, and elevation orthophotos---from the US Department of Agriculture through their data gateway site. We also found additional datasets specific to Snohomish County from different pages on their website (Information Services Directory, and Surface Water Management). We also gathered some auxiliary datasets from WAGDA, Environmental Response Management Application (ERMA), and the Washington State Department of Ecology. (Refer to Appendix 2 for a complete list of data sources.)

Challenges in the geodatabase implementation process

Lack of metadata was a serious challenge for this project. Without metadata, we had to download and open data to assess its quality and fit; this was very time consuming. Even after opening a file to investigate it, we were often unable to reach a definite judgment about its quality (particularly with respect to its completeness and recency). Lack of metadata was a problem even though our data was sourced from reputable providers, as noted above. Other challenges also related to data fit. Some data was available at too fine a resolution for our purposes (like water quality measurements from hundreds of monitoring stations, taken on an hour basis). Other data was available at a national scale, which we dealt with by simply clipping it with our WRIA 7 boundary.

Additional difficulties arose in the data acquisition process. We had numerous issues finding the data we wanted. A shapefile from one source was completely inaccessible, corrupt despite multiple download attempts. Other data (mainly orthophotos) was available in chunks that would have needed to be tediously stitched together to cover our WRIA. Some government agencies would post tantalizing data online in an interactive map format or as a pdf, but would not make the shapefile or table available for download. Lastly, some data was potentially available but barriers of time, money, or hardware would need to be surmounted. The scope of our project meant that we couldn't access data that was available on a fee basis, or available through ordering (with a wait time), or available on a CD through the library.

One final puzzle we never resolved was a loss of spatial referencing for certain feature classes and rasters during geodatabase construction. When originally checked in ArcMap or ArcCatalog the data was spatially referenced, but once the geodatabase was built and opened in ArcMap to attempt some visualizations, we learned that some files² had lost their spatial referencing and couldn't be accurately visualized. Perhaps relatedly, our DEM files were spatially referenced, but visualized in the

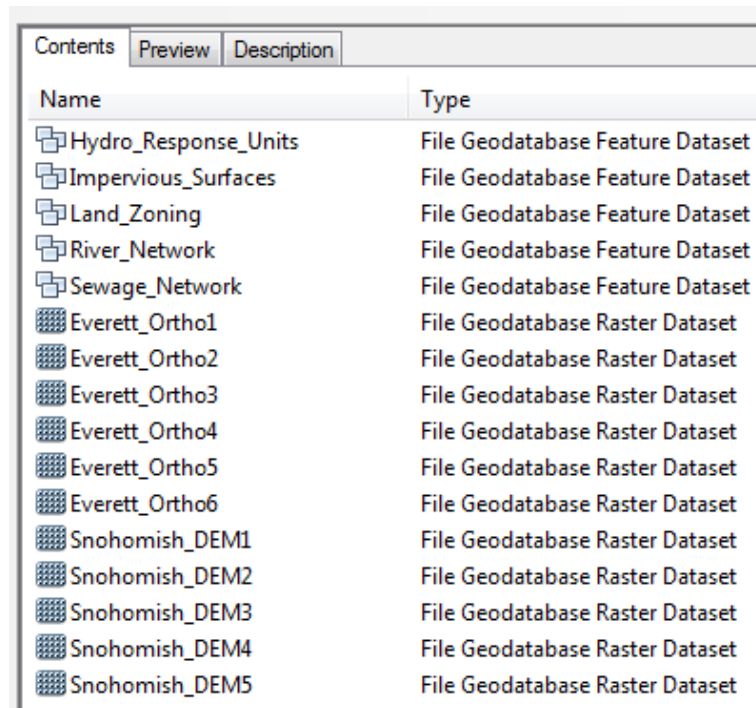
² Quite a few, namely Soil_Characteristics; Average_Precipitation; Interstate_Roads; Residential_Roads; WRIA7_Streams; River_Network_Topology; Sewage_Network; Sewage_Network_Topology; and all orthophotos.

wrong location (vis a vis our Land_Zoning basemap). This was most likely an issue with projections.

Findings

Description of results

As depicted in Figure 1, our datasets were compiled together with rasters (DEMs and orthophotos). The DEMs and orthophotos contain elevation data and surfaces that could---after some additional work, in the form of digitization---be combined with River_Network, Sewage_Network, Hydro_Response_Units and Impervious_Surfaces to answer our first three analysis questions³. The schema in Figure 2 shows the geodatabase in more detail, revealing the feature classes inside the feature datasets. Hydro_Response_Units contains soil characteristics and average precipitation, for example, to help inform planners' understanding of how surface water might flow over different areas. The Land_Zoning dataset contains data about ownership of parcels and information about landuse. This information is crucial to answering our fourth analysis question⁴. It is this information that makes our database most valuable. We envision that this data could be combined with BMP selection criteria described by the Center for Watershed Protection (2012) and thereby make a larger contribution to supporting the decision process of Snohomish county planners.



Contents	
Name	Type
Hydro_Response_Units	File Geodatabase Feature Dataset
Impervious_Surfaces	File Geodatabase Feature Dataset
Land_Zoning	File Geodatabase Feature Dataset
River_Network	File Geodatabase Feature Dataset
Sewage_Network	File Geodatabase Feature Dataset
Everett_Ortho1	File Geodatabase Raster Dataset
Everett_Ortho2	File Geodatabase Raster Dataset
Everett_Ortho3	File Geodatabase Raster Dataset
Everett_Ortho4	File Geodatabase Raster Dataset
Everett_Ortho5	File Geodatabase Raster Dataset
Everett_Ortho6	File Geodatabase Raster Dataset
Snohomish_DEM1	File Geodatabase Raster Dataset
Snohomish_DEM2	File Geodatabase Raster Dataset
Snohomish_DEM3	File Geodatabase Raster Dataset
Snohomish_DEM4	File Geodatabase Raster Dataset
Snohomish_DEM5	File Geodatabase Raster Dataset

Figure 1: *Overview of geodatabase contents in ArcCatalog*

³ What area drains into the Snohomish River? What potential open spaces lie within this identified area? What impervious surfaces lie within this identified area?

⁴ Which of these identified impervious surfaces are good candidates for different runoff reduction methods? For example: Which parking lots are government-owned, and therefore easier to mandate their replacement with permeable material? What are the landuse categories associated with identified impervious surfaces, and what does this imply for our choice of runoff reduction method?

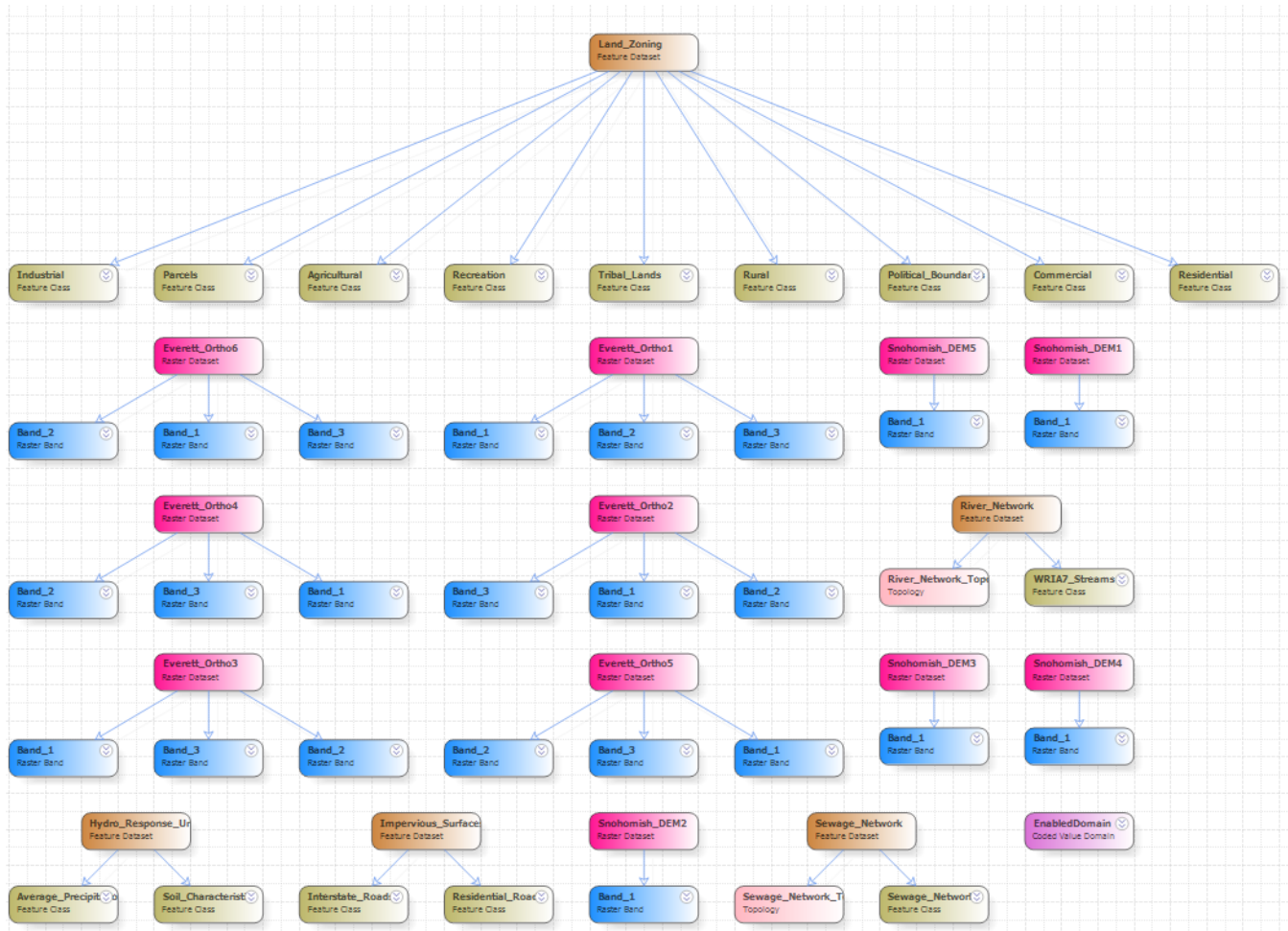


Figure 2: Overview of geodatabase contents in ArcDiagrammer

Discussion of results

As mentioned earlier, the implemented geodatabase design was different from what was planned (see Appendix 1). Nonetheless, the results were mostly as expected. By marshalling diverse data on parcel ownership, landuse, infrastructure and surface water runoff, we have assembled a database with potential to be useful for anyone involved in stormwater management in Snohomish County, people working to develop green infrastructure, as well as others invested in the environmental health of WRIA 7. We say ‘potential’ because there is significant work that remains to be done in the realm of analysis, before the data can be used to answer our research question. For example, major impervious surfaces would need to be digitized from orthophotos. As well, we hoped to find data on flow direction to help understand the surface which stormwater and other liquid pollutants flow over, but since this is generally data that is derived, it was not publicly available. Future projects should use ArcHydro tools to derive their own flow direction data for incorporation into network datasets.

Conclusions and recommendations

Generalizations emerging from this project

We do feel that the geodatabase that we produced clearly supports the problem statement crafted at the commencement of our data processing activity. We set out looking to manage stormwater runoff into the Snohomish River by identifying open spaces and impervious surfaces where green infrastructure could potentially be implemented. Though this goal is relatively straight forward the actual construction of the geodatabase was not without effort and obstacles. From this data management challenge our project team has come away with a better understanding of the sheer scope of projects such as this.

Recommendations for future work along these lines

In the broadest sense we recommend that those undertaking a similar study be very clear regarding the intended purpose of their geodatabase, so as to filter out unnecessary data. During the early stages of our work it became apparent that it would be extremely easy to collect data that did not lend itself to addressing the research. If one is not continually looking back to the research question and asking if the data is truly crucial, then the database can quickly become cluttered and unmanageable.

Another recommendation emerging from this project is to carefully maintain documentation of sources as they are found. At best, without documentation, a good deal of time would be wasted tracking down the sources after the fact. At worst, losing track of data that makes up the database would mean that the credibility of the database (and any analysis based on it) would be undermined. Our documentation included the original source of data, its planned role in the database, and any comments we had (e.g. about data quality)---see Appendix 2.

A related recommendation is to use a cloud platform to maintain this documentation, if a project is done by a group that does not share physical space like an office or lab. We used Google Docs for this purpose. Because we were constructing this geodatabase as a group, working on a cloud platform was immeasurably helpful, reducing the time and energy expended on the project. A particularly useful feature of Google Docs is the commenting function that runs alongside a document, enabling group members to conduct a conversation on the 'margins' without changing the content of the document until a decision is reached.

We recommend that those undertaking a similar project pay close attention to spatial referencing early in the construction of the geodatabase. Data that had been checked in ArcMap and/or ArcCatalog was found to lack spatial referencing when checked again after the actual construction of the geodatabase. In addressing this issue, groups undertaking a similar task or attempting to move forward with this one would need to be conscious of spatial referencing and projections every step of the way.

Finally, we recommend that those undertaking a similar task pay special attention to the *ownership* of parcels. After analysis, when results need to be put into action, knowledge of the

ownership of pieces of land is essential.

In moving forward this project will require a good deal of both time and funds in order for data to be brought to the level of specificity necessary for comprehensive analysis. Because the scale in which we are working is so particular the data needs to be just as particular. To create a detailed map of open spaces and impervious surfaces for WRIA 7 would be a time-consuming process of piecing together WRIA 7 orthophotos, digitizing new features, and calculating runoff flow. Ultimately the next steps for this project are performing analysis to create new feature classes from the data we gathered (e.g. Hydro_Response_Units, Parking_Lots, etc.); remedying the lack of spatial referencing for some features classes and rasters; and fleshing out substantial gaps in the metadata, so that future users can be more confident and accurate in assessing the quality and fit of the database contents.

References

- Center for Watershed Protection. (2012). *West Virginia stormwater management and design guidance manual*. Charlottesville, VA: West Virginia Department of Environmental Protection and the Center for Watershed Protection.
- Environmental Protection Agency. (n.d.). *What is green infrastructure?* Retrieved from http://water.epa.gov/infrastructure/greeninfrastructure/gi_what.cfm
- Environmental Protection Agency. (n.d.). *A watershed approach*. Retrieved from <http://water.epa.gov/type/watersheds/approach.cfm>
- Regional Open Space Strategy. (2010). *Executive summary: Envisioning Seattle's green future*. Retrieved from <http://depts.washington.edu/open2100/book/executivesummarybook.pdf>
- Schueler, T. (1994). The importance of imperviousness. *Watershed Protection Techniques*, 1(3):100-111.
- Yeung, A. K. W. & Hall, G. B. (2007). *Spatial database systems: Design, implementation, and project management*. Dordrecht, the Netherlands: Springer.

Appendix 1: Planned logical design versus implemented design

Planned logical geodatabase design	Implemented logical geodatabase design
<ul style="list-style-type: none"> ★ Hydro_Response_Units <ul style="list-style-type: none"> ○ Soil_Characteristics ○ Rainfall ○ Flow_Direction ★ BMP_Choice <ul style="list-style-type: none"> ○ Parcels ★ WRIA7 <ul style="list-style-type: none"> ○ Political_Boundaries ★ Open_Spaces <ul style="list-style-type: none"> ○ Recreational_Areas ○ Conservation_Areas ○ Rural_Resource_Management ★ River_Network <ul style="list-style-type: none"> ○ Hydro_Network ○ Hydro_Junction ○ HydroEdge_Stream ★ Impervious_Surfaces <ul style="list-style-type: none"> ○ Rooftops ○ Parking_Lots ○ Roads 	<ul style="list-style-type: none"> ★ Hydro_Response_Units <ul style="list-style-type: none"> ○ Soil_Characteristics ○ Average_Precipitation ★ Land_Zoning <ul style="list-style-type: none"> ○ Industrial ○ Parcels ○ Agricultural ○ Recreation ○ Tribal_Lands ○ Rural ○ Political_Boundaries ○ Commercial ○ Residential ★ River_Network <ul style="list-style-type: none"> ○ WRIA7_Streams ○ River_Network_Topology ★ Sewage_Network <ul style="list-style-type: none"> ○ Sewage_Network ○ Sewage_Network_Topology ★ Impervious_Surfaces <ul style="list-style-type: none"> ○ Interstate_Roads ○ Residential_Roads ★ Everett_OrthoX (1 - 6) ★ Snohomish_DEMX (1 - 6)
<hr/> <p><i>Legend</i></p>	
<ul style="list-style-type: none"> ★ Dataset ○ Feature class 	

Appendix 2: Documentation of data sources

Source (URL)	Feature class	Notes
http://www1.co.snohomish.wa.us/Departments/Information_Services/Services/Maps/orthophotos.htm ; http://wagda.lib.washington.edu/data/type/photography/snohomish/	<i>Rooftops</i>	We imagined that users could digitize impervious surfaces from orthophotos --- which exist, at a fine resolution, but must be paid for (from Snohomish) or spatially referenced (from WAGDA)
http://www1.co.snohomish.wa.us/Departments/Information_Services/Services/Maps/orthophotos.htm	<i>Parking_Lots</i>	[as above]
http://datagateway.nrcs.usda.gov/GDGOrder.aspx	<i>Roads</i>	Snohomish county major roads --- looks sparse
http://datagateway.nrcs.usda.gov/GDGOrder.aspx	<i>Roads, con't</i>	Snohomish county roads --- looks comprehensive
http://datagateway.nrcs.usda.gov/GDGOrder.aspx ; http://gis.ess.washington.edu/data/raster/tenmeter/byquad/index.html	<i>Elevation</i>	Data from NRCS corrupt after multiple download attempts; ESS-sourced data needs to be pieced together
http://datagateway.nrcs.usda.gov/GDGOrder.aspx	<i>Soil_Characteristics</i>	“soil_mu_a_wa661” feature has comprehensive soil data
http://datagateway.nrcs.usda.gov/GDGOrder.aspx	<i>Average_Precipitation</i>	(Need better data specifically for Snohomish/WRIA7)
http://www1.co.snohomish.wa.us/Departments/Public_Works/Divisions/SWM/D/DrairageInvGISData.htm	<i>Snohomish_Sewer_System</i>	This is a network dataset for Snohomish’s existing gray infrastructure, containing feature classes for (1) facilities, (2), intersections, (3) drains, (4) catchbasins, (5) pipes connecting all these points.
https://www.erma.unh.edu/northwest/erma.html#/x=-122.31468&y=48.02196&z=10&layers=417+410+1767	<i>HydroEdge_Stream</i>	Flowlines (streams, plus imaginary lines inside waterbodies). Clipped, using <i>Zoning</i> .
http://datagateway.nrcs.usda.gov/GDGOrder.aspx	<i>Parcels</i>	Can’t find metadata to determine what the labels mean for each parcel, but it’s certainly the kind of thing we’re looking for!
https://data.snostat.org/Government/Snohomish-County-GIS-Cities/9wxp-sxyi	<i>Political_Boundaries</i>	Identifies city boundaries
ftp://ftp.snoco.org/Planning_and_Development_Services/Zoning/Current%20Shape%20Files	<i>Zoning</i>	Replaces <i>Recreational_Areas</i> , <i>Conservation_Areas</i> , <i>Rural_Resource_Management</i> ;

		functions as a basemap
--	--	------------------------