Data Development Workflow

1. The National Hydrographic Dataset (v2) Data (NHD+)

Overview

NHD+ is a suite of geospatial products that are derived from static snapshots of the National Hydrography Dataset (NHD) stream network (1:100,000-scale), Watershed Boundary Dataset (WBD) hydrologic units (12-digit), and National Elevation Dataset (NED) topography (30m) through a sophisticated data integration process. The NHD, WBD and NED are all maintained through USGS National Geospatial Program (NGP) led stewardship programs involving states and federal agencies. The derived components of NHDPlus are updated using the latest versions of the NHD, WBD and NED on a periodic basis, as determined by programmatic priorities and resources. For example, the production of NHDPlus Version 2 was largely driven by USGS Water Program requirements for improved regional water quality modeling capabilities. USGS Water and the EPA Office of Water worked together to update the NHD in preparation for NHDPlus Version 2 and, also, shared the cost for producing it.   
(Source: <http://www.horizon-systems.com/nhdplus/NHDPlusV2_home.php>)

The DMT Habitat Prioritization Project relies on the NHD+ (version 2) to provide both the spatial units of habitat prioritization (**NHD+ catchments**, eventually up-scaled to the **12-digit hydrologic units**) as well as numerous in-stream and up-slope attributes associated with these catchments. The project also uses a number of raster datasets included with NHD+ to derive additional instream and upslope catchment attributes. Table 1 lists all the datasets obtained from NHD+ used in habitat prioritization.

This section outlines the procedures used to obtain and prepare NHD+ data for use in developing the habitat models used to determine uplift potential.

1. Obtaining the data

Horizon Systems (<http://www.horizon-systems.com/NHDPlus/index.php>) hosts all NHD+ data and related documentation on its servers and allows data to be downloaded free of charge. Most data are bulk-downloaded for regional sections corresponding to major drainage areas; see <http://www.horizon-systems.com/NHDPlus/NHDPlusV2_data.php>. North Carolina touches three of these regions: The South Atlantic North (03N), the Tennessee (06), and the Ohio (05).

1. Uncompressing the files and the NHD+ directory format

We downloaded all the raster, vector, and tabular datasets listed in Table 1 for each of the three regions intersecting North Carolina. Some datasets for regions 03 and 05 were additionally parsed into sub-regional divisions (e.g. 03a and 03b). All subdivisions for each region were downloaded.

The datasets, which are obtained as compressed zip files, were decompressed into the native NHD+ directory format (Figure 1). At the end of this step, all required vector, raster, and tabular NHD+ data could be accessed on a local desktop. Uploading and merging the NHD+ data to the Nicholas School’s GIS Server

1. Importing data to the Nicholas School GIS Server

To facilitate sharing data across multiple machines (and avoid duplication among team workers), we uploaded the NHD+ datasets to the Nicholas School’s GIS Server (“NS-GIS.WIN.DUKE.EDU”) - a Windows 2008 Server running MS SQL Server 2008 (R2) integrated with ESRI’s ArcGIS Server (v 10.2). To do this, we created an ArcGIS Enterprise GeoDatabase on the server (“NHDPlusV2”) using Geodatabase administrative tools in ArcGIS desktop.

All but a few (exceptions listed below) of the NHD+ datasets for a single region (region 03) were imported directly into the server geodatabase using ArcMap’s *Feature Class to Geodatabase*, *Raster to Geodatabase*, and *Table to Geodatabase* tools. Once region 03 datasets were imported into the server geodatabase, corresponding datasets from the regions 05 and 06 were merged directly into the existing server datasets using the ArcGIS *Mosaic* and *Append* tools for raster and vector/tabular datasets, respectively.

**Projections**

All spatial datasets maintained their original coordinate systems. Raster datasets are referenced to the Albers Equal Area (NAD 83) projection. Vector datasets are referenced to the NAD 1983 geographic coordinate system.

**Metadata**

FGDC Metadata from the original datasets downloaded from the Horizon Systems servers were copied over to the server datasets using ESRI’s Import Metadata tool.

**Special case #1: NHDFlowlines**

The NHDFlowlines feature class was handled slightly differently to allow a network dataset to be constructed from it. As ArcMap network datasets can only be created within feature datasets, we first manually created a feature dataset (“NHDFlowlines”) in the server geodatabase, using the coordinate system and XY tolerance information of the NHDFlowlines feature class as a template. We then imported the NHD Flowline dataset for NHD region 03 into this feature dataset and then appended the Flowline datasets from regions 05 and 06 to it (using ArcCatalog’s Load Dataset… function)

**Special case #2: Mean annual and monthly precipitation, temperature, and runoff tables**

The NHD+ provides the tabular data for precipitation, temperature, and runoff as both mean annual averages and as individual monthly averages. Prior to uploading these datasets into the server geodatabase, we combined the annual and monthly means into a single table for each parameter. We did this in two steps, with a Python script written to execute each one.

The first Python script (“NHD\_MergeVPUAttributes.py”) merged the individual parameter tables for the different NHD+ regions into a single table covering all regions. For example, the mean monthly runoff tables for the month of January (ROMM01001.txt) for regions 05 and 06 were each appended to that of region 03. The appended tables were stored in a local file geodatabase as inputs for the second step. A list of the tables is shown in in Table 2.

The second Python script (“NHD\_JoinVPUAttributes.py”) joined the 12 monthly mean values to the annual mean tables for cumulative total precipitation, cumulative total temperature, incremental precipitation, incremental temperature, and runoff, respectively. Records were joined using the *FeatureID* attribute, and the value fields were renamed to include the month of record (e.g. “PRECIPVC\_01”). The result here is a single table for each parameter (precipitation, temperature, and runoff) that include both annual and monthly mean values.

We wrote an additional Python script to calculate the minimum and of the monthly mean values for each catchment. This script is titles “NHD\_CalculateVPUSummaries.py”.

**Special case #3: Incremental and Cumulative NLCD tables**

The NHD+ also provides incremental and upstream cumulative area of NLCD land cover classes for each catchment via its VPU Attribute Extension (<http://www.horizon-systems.com/NHDPlus/V2NLCD2011.php>). As in the case with the precipitation, temperature, and runoff tables, we merged NLCD tables for each region into a single table of all regions and uploaded into the ArcGIS Server database. We used the same Python script as above ((“NHD\_MergeVPUAttributes.py”) to automate this.

Additionally, however, we also combined several land cover attributes to produce more generalized land cover classes for analysis. Table 3 lists the regroupings of the original NLCD land cover classes.

**Table 1**. NHD+ (v2) Datasets obtained from Horizon Systems for NC Regions (03N, 05, & 06).

|  |  |  |
| --- | --- | --- |
| Format | Name | Description |
| Raster | elev\_cm | Elevation, in cm |
| Raster | CAT | NHD+ catchments |
| Raster | fac | Flow accumulation |
| Raster | shdreleif | Shaded relief |
| Raster | fdr | Flow direction |
| Raster | fdrnull | Flow direction; streams set to null values |
| Vector | NHDflowline | NHD+ flowline features |
| Vector | CatchmentFeatures | NHD+ catchment features |
| Vector | WBD\_Subwatershed | National watershed boundary database (HUC12 features) |
| Vector | WBDHU2 | 2-digit hydrologic unit polygons |
| Vector | WBDHU4 | 4-digit hydrologic unit polygons |
| Vector | WBDHU6 | 6-digit hydrologic unit polygons |
| Vector | WBDHU8 | 8-digit hydrologic unit polygons |
| Table (dbf) | CumulativeArea | Cumulative area upstream of an NHDFlowline feature |
| Table (dbf) | ElevSlope | Elevation and slope derived for NHDFlowline features |
| Table (dbf) | PlusFlowlineVAA | NHD+ “Value Added Attributes” for NHDFlowline features |
| Table (csv) | CumTotNLCD2011 |  |
| Table (csv) | IncrTotNLCD2011 |  |
| Table (csv) | CumTotPrecipMA | Mean annual rainfall across the area upstream of a catchment |
| Table (csv) | CumTotTempMA | Mean annual temperature across the area upstream of a catchment |
| Table (csv) | IncrPrecipMA | Mean annual rainfall over the area of each NHD+ catchment |
| Table (csv) | IncrTempMA | Mean annual temperature over the area of each NHD+ catchment |
| Table (dbf) | EROM\_MA0001 | Extended unit runoff method (EROM) annual flow estimates |
| Table (dbf) | EROM\_*mm*0001 | Extended unit runoff method (EROM) monthly flow estimates |
| Table (csv) | ROMA0001 | Mean annual runoff recorded within each NHDFlowline feature |

**Table 3.** Reclassification of NLCD classes into more generalized land cover classes

|  |  |
| --- | --- |
| **Original class** | **Generalized class** |
| 11 – Open water | 11 – Open Water |
| 12 – Perennial Ice/Snow | *- Omitted -* |
| 21 – Developed, Open Space | 21 – Developed, Open Space |
| 22 – Developed, Low intensity | 20 - Developed |
| 23 – Developed, Medium Intensity | 20 - Developed |
| 24 – Developed, High intensity | 20 - Developed |
| 31 – Barren | 31 – Barren |
| 41 – Deciduous forest | 41 – Deciduous forest |
| 42 – Evergreen forest | 42 – Evergreen forest |
| 43 – Mixed forest | 43 – Mixed forest |
| 51 – Dwarf shrub | *- Omitted -* |
| 52 – Scrub shrub | 50 – Shrub |
| 71 – Grassland/herbaceous | 71 – Grassland/herbaceous |
| 72 – Sedge/herbaceous | *- Omitted -* |
| 73 – Lichens | *- Omitted -* |
| 74 – Moss | *- Omitted -* |
| 81 – Pasture/hay | 81 – Pasture/hay |
| 82 – Cultivated crops | 82 – Cultivated crops |
| 90 – Woody wetlands | 90 – Woody wetlands |
| 91 – Emergent herbaceous wetlands | 91 – Emergent herbaceous wetlands |

**Figure 1.** Directory format created when NHD+ downloaded zip files are decompressed.

