Donna Creek Hydrometeorology Data Search

Stephen Bird

2019-10-21

Data Sources

I've assembled the data sources required for our analysis. This includes:

- Daily Gridded Meteorlogical data made available by the Pacific Climate Impacts Consortium
- Daily hydrometric data acquired by the Water Survey of Canada
- Shuttle Radar Topography Mission 30 meter elevation data
- National Hydro Network - NHN - GeoBase Series vector coverage of rivers and lakes

I've downloaded the hydrometric data (see below) but the rest of the data can be downloaded from source once we finalize the analysis. We can calculate a range of basin morphometric parameters from both the SRTM DEM and the vector coverages (e.g. basin area, relief, lake area, etc.). Tony Cheong's Master's thesis would be a good guide for calculation of these features if you think they would be useful. I also have a separate manual for defining and calculating these parameters.

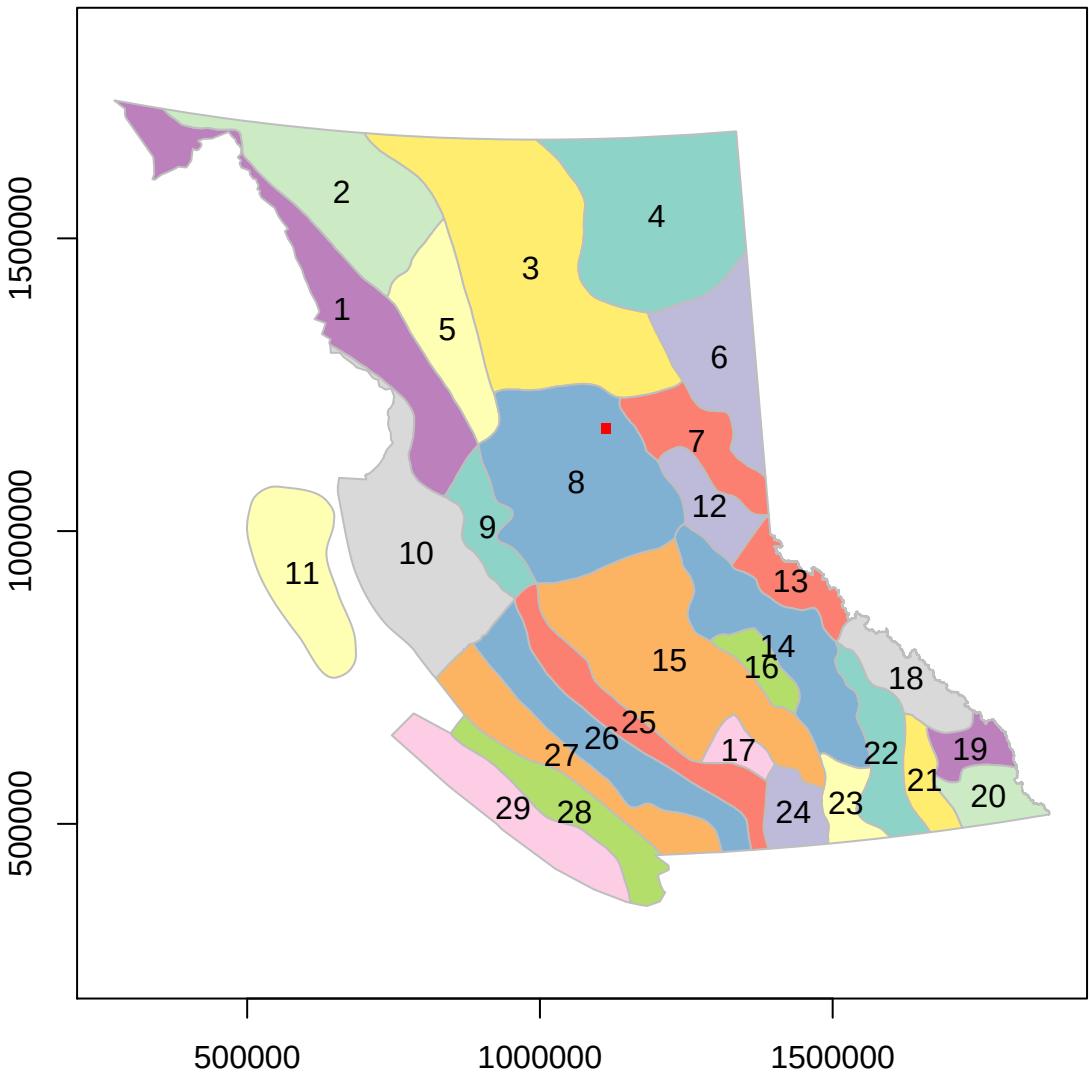
There are a few options for the daily gridded meteorological data to consider. We can choose one of three data sets:

- PNWNAMET (1945-2012): includes precipitation, max. and min. temperature, and wind data
- NRCANMET (1950-2012): includes precipitation, max. and min. temperature, and wind data
- PBCMET (1950-2004): includes precipitation, max. and min. temperature data

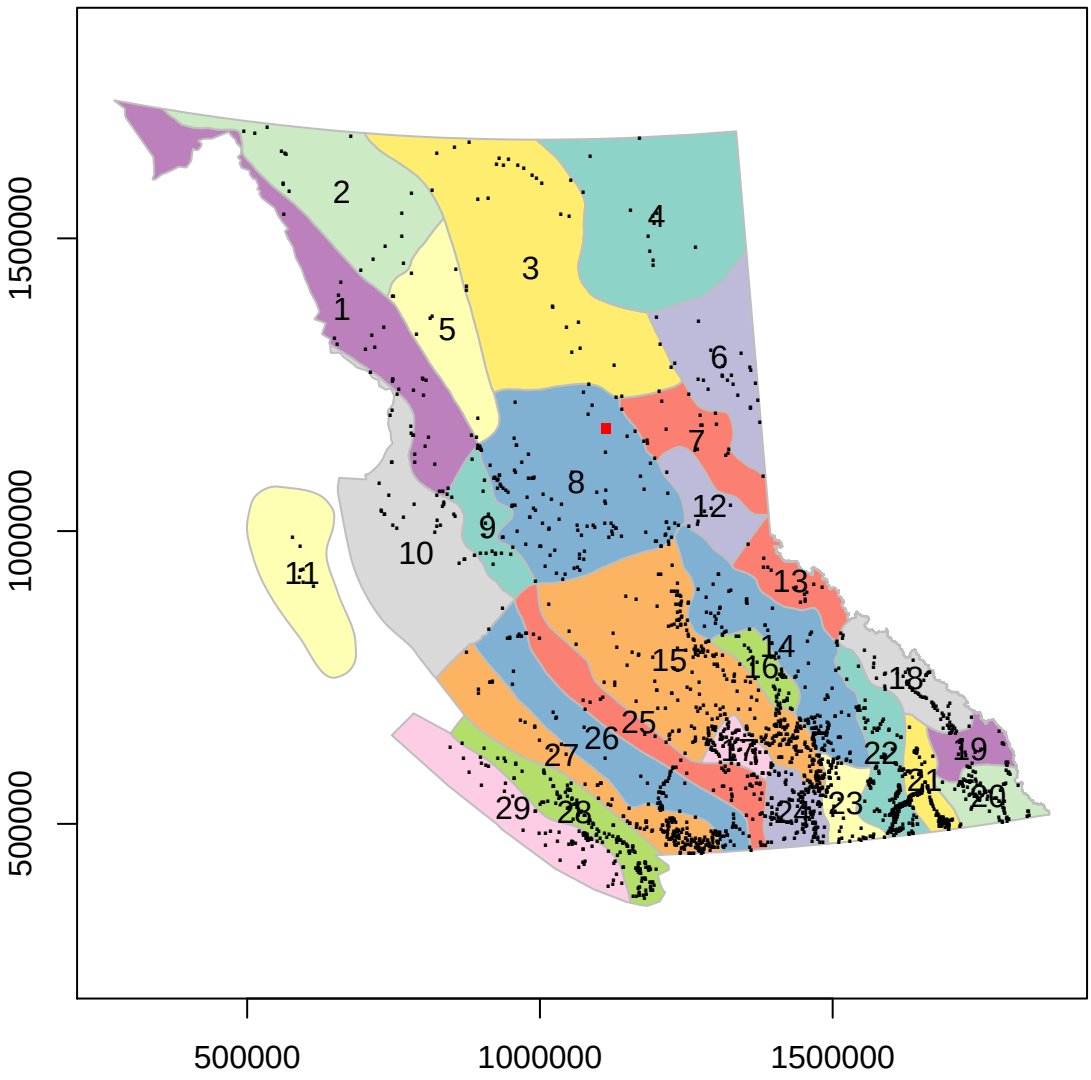
Or, we can obtain raw data for stations in BC and interpolate a grid ourselves. The gridded set named PNWNAMET is the most recent and is the longest record available so I think it is the best choice for us. The data are gridded at 1/16° (about 6 km^2) and most of our basins will be much larger so the spatial resolution of the grid is suitable for our work. However, the raw data source includes snow-pack data and this may be useful to our model (although I assume the recurrent neural network can learn and infer a snow-pack depth from precipitation - temperature - runoff relations). It would require an extra step as we would have to grid our own data. The other advantage of the raw data is that we can create grids up to 2018 and we would eventually like to predict stream-flow in Donna Creek for years spanning 1971 - 2018. We can think about this.

Hydrometric Station Short-list

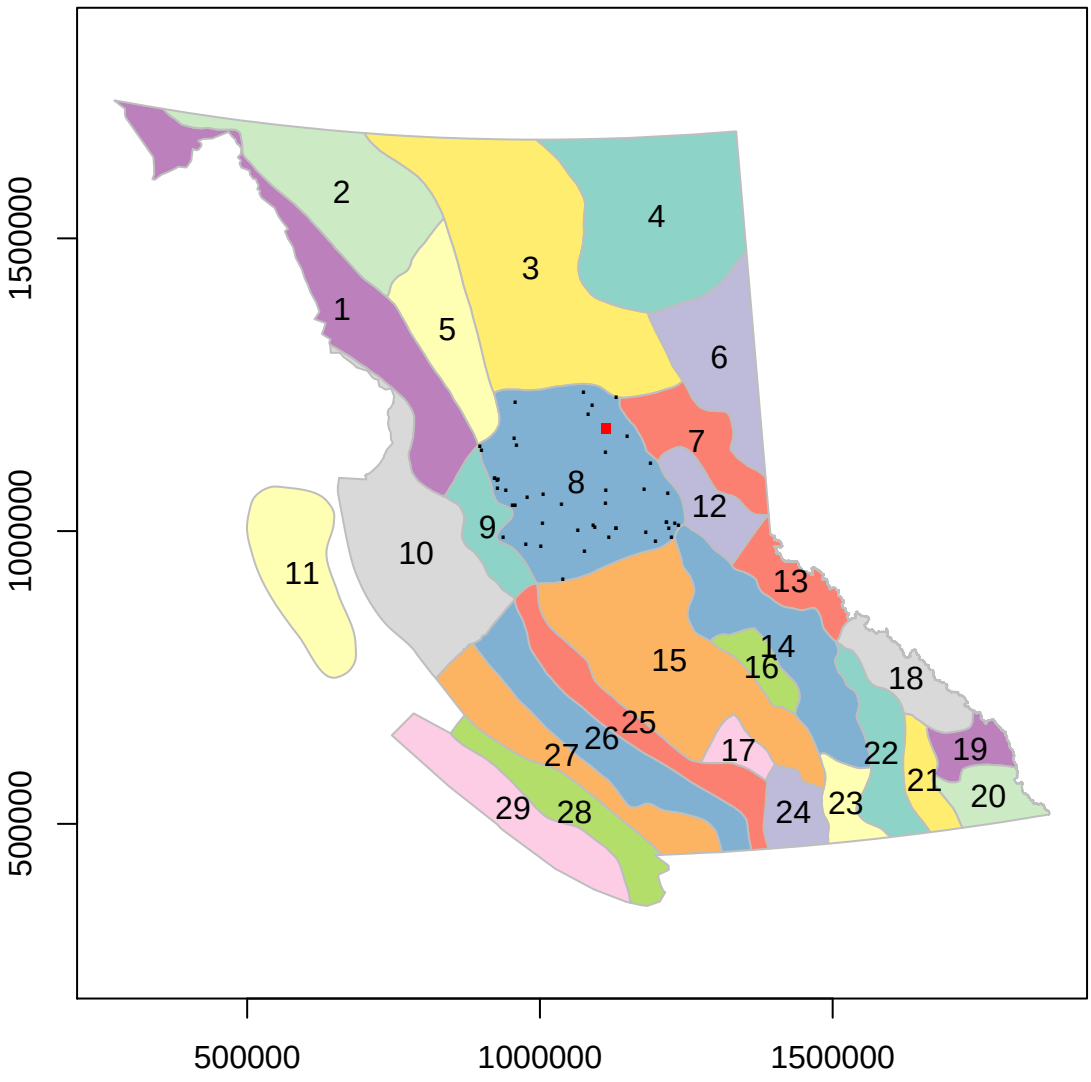
I created a short-list of suitable hydrometric stations since I'm familiar with the region around Donna Creek and Hydat database. The first step is to plot a map of hydrolocic zones for British Columbia and locate Donna Creek. The zones were (descriptively) defined by hydrologists working for the provincial government and are based on climate, physiography, and hydrometric characteristics of the landscape. Donna Creek is located in Nechako Plateau (zone 8) Hydrologic Zone and appears as a red dot in the figure below:



Next, I plotted all the hydrometric stations for BC, including both active and inactive stations (the stations appear as small black dots in the figure below):



Then, I subset all the hydrometric stations in the Nechako Plateau with at least 10 years of discharge data collected sometime between 1945 and 2012 (matching the period of the PNWNAMET data-set). These are plotted in the figure below:



The analysis identified 47 potential hydrometric stations for our analysis. For comparison, the drainage area of Donna Creek is 126 km^2 while the drainage area of subset stations are summarized below (in km^2):

```
summary(hy_zone8$DRAINAGE_AREA_GROSS)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.     NA's
##     9.08   179.00  2615.00  6313.90  6452.50 43300.00        3
```

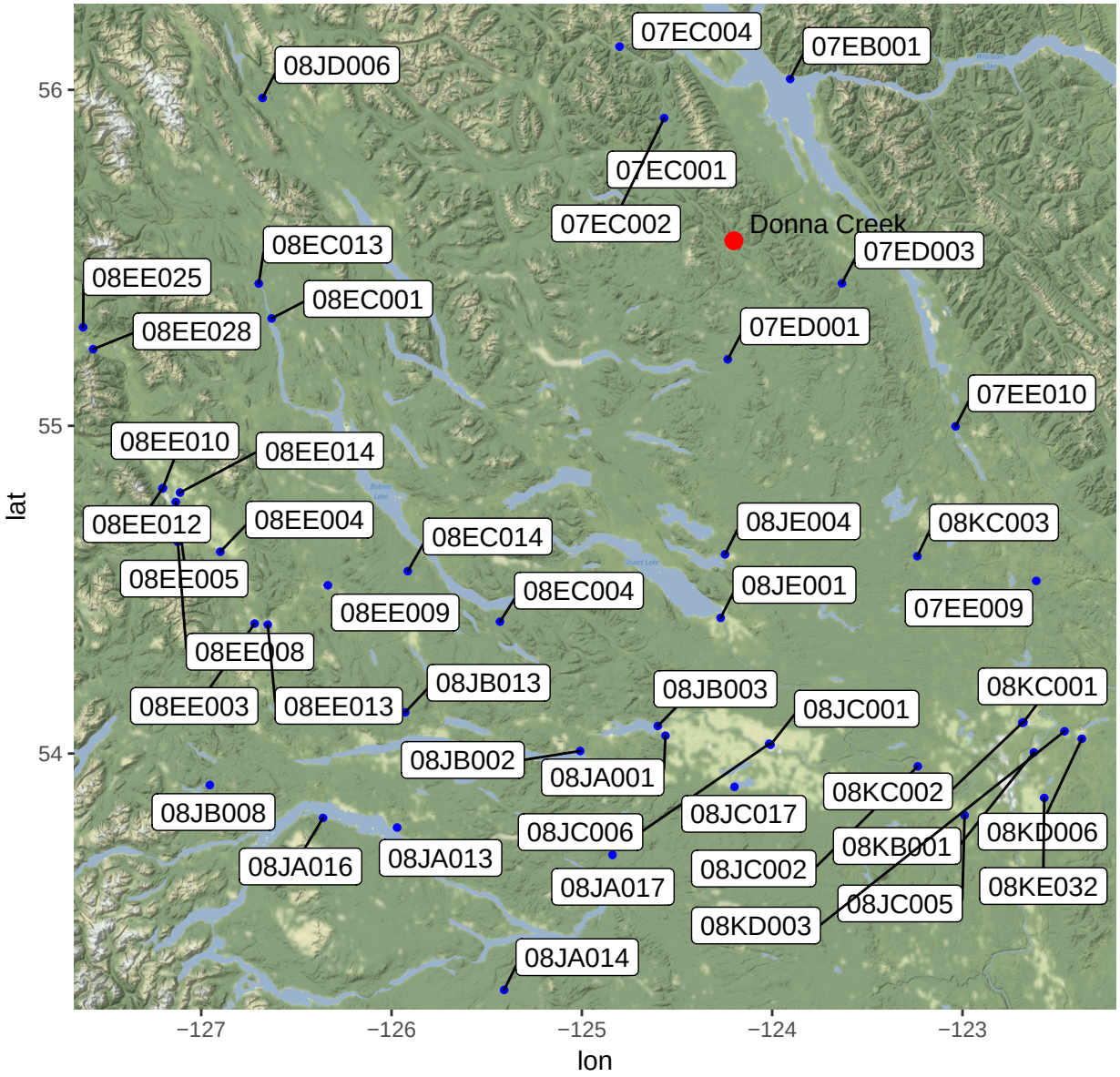
Most records started in 1965 (median Year_from) and terminated in 2015 (median Year_to) and were at least 10 years in length but most were about 35 years in length (see below):

```
summary(hy_zone8[,16:18])
```

```
##    Year_from       Year_to     RECORD_LENGTH            geometry
##  Min.   :1915   Min.   :1954   Min.   :10.00   POINT        :47
##  1st Qu.:1950   1st Qu.:1996   1st Qu.:20.50   epsg:3005    : 0
##  Median :1965   Median :2015   Median :35.00   +proj=aea ...: 0
##  Mean   :1962   Mean   :2004   Mean   :38.85
##  3rd Qu.:1976   3rd Qu.:2017   3rd Qu.:55.50
##  Max.   :1998   Max.   :2019   Max.   :89.00
```

Finally, for visual interpretation, here are the hydrometric stations plotted over a shaded terrain map with

some water bodies present. The station numbers are plotted in the white boxes (use these to cross reference with the .csv file of raw data).



I don't know how computationally intensive the analysis will be, but we can choose all 47 stations or perhaps a sub-set of these sites to give a range of basin areas, terrain types, lakes, etc (something to consider).