Getting Data:

To download daily data from OCCP, navigate to the their data store which is different than the main website (<https://cmip5.lamps.yorku.ca:5001/sharing/H8pWZgk18>), navigate to LAMPS\_10KM\\Grids, then click on the appropriate variable and emissions scenario, then click “Download Folder” in the top-left and proceed. Unzip the files into a folder of your choice. Python code is then used to parse the data in that folder from a .mat file into a .npy file. This format takes ~3x as much memory on disk, although it allows the use of memory maps to reduce the load on RAM during processing. Once the code is complete, all .mat files will be deleted. This code is provided in the appendix and requires the *scipy* and *numpy* modules included with ArcGIS.

The process for downloading daily data from CCDP is using HTML requests to query data based on the parameters. Python code to download this data will be provided, and requires Python 2.7.14, in addition to the *arcpy* and *requests* modules to run (both included with current installations of ArcGIS). Future implementations may make those requirements obsolete although it is not a priority.

Processing Data:

* Ensemble mean for best estimate, range of values would preferably be given by 10th/90th percentiles (York study used those ranges) – in interests of computation time the ranges will not be calculated
* Reference period from 1986 to 2005, projections from 2020 to 2099

OCCP:

* Download .mat files for each climate model at RCP8.5
* Convert into a .npy data structure
* Use a memory-map to load every structure into Python
* Take the mean across each model for each element in array
* Unravel 3D array into 2D
* Load into Pandas as a dataframe

CCDP:

* Use GET requests to download text-files for each grid square
* Parse text files into numerical values
* Align data across consecutive time periods
* Take the mean of those values across different models
* Load into Pandas as a dataframe

Combining Data:

* ArcGIS was used to get the intersection of the different grids to create a “master grid.” Each grid was constructed by analyzing the data files behind the interactive maps present on each data portal’s website. The attributes of this grid give us a handy index to link data, as well as get the area of each grid cell, using distance measurements based on the [] projection. [choose projection for accuracy]
* Join OCCP and CCDP data to large master dataframe
* Use equations to get humidex values
* Do math with weighted average to return data to OCCP format