A New Open-Access HUC-8 Based Downscaled CMIP-5 Climate Model Forecast Dataset for the Conterminous United States

Dustin Woodbury

A project submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Master of Science

Daniel P. Ames
E. James Nelson
Brett Borup

Department of Civil and Environmental Engineering
Brigham Young University
March 2016

Copyright © Dustin Hanson Woodbury
All Rights Reserved
ABSTRACT

A New Open-Access HUC-8 Based Downscaled CMIP-5 Climate Model Forecast Dataset for the Conterminous United States

Dustin Woodbury
Department of Civil and Environmental Engineering, BYU
Master of Science

Watershed-scale hydrologic simulation models generally require climate data inputs including precipitation and temperature. These climate inputs can be supplied by downscaled global climate simulations which have the potential to forecast runoff at the level of local watersheds. While a simulation designed to drive a watershed model would ideally be constructed at a relatively small scale, global climate simulations are, by definition, large-scale and coarsely gridded in arbitrarily determined rectangular spatial grid patterns. This paper addresses the technical challenge of making climate simulation model results readily available in the form of downscaled data sets that can be used for watershed scale models. Specifically, we present the development and deployment of a new Coupled Model Intercomparison Project phase 5 (CMIP5) based database which has been prepared through a scaling and weighted averaging process for use at the level of USGS HUC-8 watersheds (roughly 1,800 square-km). The resulting dataset includes 2,106 virtual observation “sites” (watershed centroids) each with 698 associated time series data sets representing average monthly temperature and precipitation between 1950 and 2099 based on 234 unique climate model simulations. The new dataset is deployed on a HydroServer and distributed using WaterOneFlow web services in the WaterML format. Two example use cases for the data set are also presented.

Keywords: Climate model, CMIP5, GCM, Downscaling, HUC-8 Watersheds, database, HydroServer
ACKNOWLEDGEMENTS

This project was primarily funded by the United States Bureau of Reclamation under Cooperative Agreement, R13AC10030. I acknowledge Jiri Kadlec and the BYU Hydroinformatics laboratory students for their insights in developing scripts for this project. We also acknowledge the contributions of David Blodgett in using and understanding the Geo Data Portal.

We acknowledge the World Climate Research Programme’s Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups for producing and making available their model output. For CMIP the U.S. Department of Energy’s Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.
# TABLE OF CONTENTS

LIST OF TABLES .................................................................................................................. v
LIST OF FIGURES ............................................................................................................... vi

1 Introduction ..................................................................................................................... 1

2 Methods ......................................................................................................................... 5
   2.1 Defining Practical Watersheds .................................................................................... 5
   2.1 Rescaling Using the USGS Geo Data Portal ............................................................... 6
   2.1 Processing ................................................................................................................ 9
   2.1 Database Implementation ......................................................................................... 12

3 Results ............................................................................................................................. 19
   3.1 Discussion ................................................................................................................. 19
   3.2 Use Case 1: Examining Climate Models ................................................................. 19
   3.3 Use Case 2: Comparing Watersheds ....................................................................... 22
   3.4 Planned Implementation ......................................................................................... 24
   3.5 Future Development ............................................................................................... 24

4 Conclusion ...................................................................................................................... 27

References ......................................................................................................................... 29
LIST OF TABLES

Table 1: Inputs of submitFeatureWeightedGridStatistics Method ................................................. 9
Table 2: Climate Services Feature Table ..................................................................................... 15
Table 3: Common Climate Change Scenarios ............................................................................ 22
LIST OF FIGURES

Figure 1: USGS HUC-8 Watersheds Representative of Spatial Regions Used in Data Processing...........................................................................................................................................6

Figure 2: Example of Weighted Grid Averaging Used in Scaling of Climate Data to HUC-8 Polygons .................................................................................................................................................................................7

Figure 3 Natural and Simplified Watershed Boundary of a HUC-8 Polygon .................................................................................................................................................................................................8

Figure 4 HydroServer Google Map User Interface .................................................................................................................................................................................................16

Figure 5 Site Details User Interface on HydroServer Lite .................................................................................................................................................................................................17

Figure 6 Work Flow Diagram of Translation Process .................................................................................................................................................................................................19

Figure 7 Scatter Plot of All 234 CMIP5 Models for One Watershed Produced in HydroDesktop’s Climate Analysis Toolkit .................................................................................................................................................................................................20

Figure 8 Temperature Change of HUC-8 Watershed Regions as Described by a Model From Database .................................................................................................................................................................................................23
1 INTRODUCTION

Downscaled climate change models have the potential to significantly inform regional-scale watershed runoff simulations of future scenarios for the purposes of improved short and long-term water resources management (Fonseca et al. 2014; LaFontaine et al. 2015; Pradhanang et al. 2013; Stöckle et al. 2014). A number of General Circulation Models (GCMs) have been produced and proven useful in predicting trends in future climate conditions and are of increasing interest to hydrologists and water resource managers for planning purposes (Flato 2013; Knutti and Sedláček 2013; Taylor et al. 2012). Downscaling the spatial resolution of GCMs allows hydrologists to visualize climate behavior at a more relevant spatial resolution leading to better predictions of streamflow and water resource availability (Blodgett et al. 2013; Hwang and Graham 2014; Wooten et al. 2014).

Given the growing interest in climate change research, the development and testing of GCMs continues to be a major focus of the climate modeling community (Ford et al. 2011). GCMs are capable of providing projections of ground level air temperatures and precipitation values at various locations given specified atmospheric conditions. Using a sophisticated application of the Navier-Stokes equation, climate modelers are able to predict vector paths of hot air masses around the globe (Zhang and Moore 2014). By estimating atmospheric compositions, and the corresponding thermal irradiance captured by the atmosphere, GCMs predict air temperature and precipitation conditions for long-term forecasting over the next century. GCMs vary based on the preferred modeling approach of each parent research institution and how atmospheric
composition is estimated. These variations in the GCMs themselves result in a wide variety of predicted future scenarios for any specific region of interest.

In 1995 the Coupled Model Intercomparison Project (CMIP) was established under the World Climate Research Programme (WRCP) as a standard experimental protocol for atmosphere-ocean general circulation models (Covey et al. 1996). This collection includes atmosphere-ocean general circulation models from most major research institutions in the international climate modeling community. CMIP has collected model outputs from the pre-industrial climate simulations ("control runs") and 1% per year increasing-CO2 simulations (Andrews et al. 2012). The latest collection is phase 5 (CMIP5) and includes 234 different projections for historical and future climate data at a spatial resolution of between 1/2° and 4° grid cells across the globe. Daily average air temperature and precipitation values for each month are available for these cells from 1950 through 2099 (Taylor et al. 2012).

To increase the spatial resolution of climate projections and provide more localized data, a collaborative initiative led by the Lawrence Livermore National Laboratory (LLNL) was completed in 2013 that resulted in the downscaling of general circulation models using a method known as bias-corrected spatial disaggregation (BCSD). BCSD is a process for statistically downscaling spatial data that relies on mapping GCM values by quantile onto historically observed data wherein each quantile value has received its own adjustment (Girvetz, 2013). This method, as well as the downscaling work by LLNL, has been published and shown to be consistent with historical observed data (Cavazos and Arriaga-Ramírez 2012; Hayhoe 2006; Werner 2011). This previously performed work by LLNL resulted in a collection of climate projection data at a spatial resolution of 1/8° (~12 kilometers on a side) grid cells over the contiguous United States and parts of Mexico and Canada (Wood et al., 2002 and 2004;
Reclamation, 2013). This downscaling increased the resolution from approximately 190 grid cells in the original CMIP5 data set to approximately 53,220 grid cells. This downscaled collection offers all 234 GCM projections over virtually any surface of the contiguous United States at a much finer resolution than the original CMIP5 collection and brings climate data one step closer to practical application for hydrologists. A table listing the CMIP5 models used in the LLNL downscaling work as well as their parent institutions IDs can be found at http://cmip-pcmdi.llnl.gov/cmip5/docs/CMIP5_modeling_groups.pdf.

Despite the emergence of downscaled BCSD circulation models, these data remain difficult to use at the primary spatial area of interest to hydrologists— the watershed. BCSD data are mapped to 1/8° grid cells across the United States and require processing to be re-scaled to the spatial area of a watershed. While these processing methods may be familiar to experts in the field of geoinformatics, practicing water resources engineers may be unable to invest the requisite time or resources. It is therefore relatively difficult to quickly investigate BCSD models at the spatial level of local watersheds of varying size and shape. While downscaled GCM data is available via the CMIP5 collection, the water resources engineering community has a need for a readily available database of GCM’s pre-processed to the size and shape of commonly-used watershed study areas.

This paper describes the development and deployment of a new open-access database of downscaled climate model forecasts. This database has been prepared through a downscaling and weighted averaging process for use at the level of USGS HUC-8 watersheds (between 500 and 2500 square-km). The resulting dataset includes 2,106 virtual observation sites represented by watershed mid-points that contain monthly average temperature and precipitation between 1950 and 2099 based on 234 unique climate model simulations.
2 METHODS

2.1 Defining Practical Watersheds

The USGS standard Hydrologic Unit Code (HUC) level 8 watersheds were used as a representative modeling unit for this study. Smaller watersheds (e.g. level 12 watersheds) could be processed in a similar manner as presented here. Although the size of HUC-8 watersheds are somewhat variable, HUC-8 watersheds have an average surface area of 1,800 square kilometers. The HUC system is widely used and universally recognized within the hydrologic community. Figure 1 depicts all HUC-8 watersheds within the contiguous United States and represents all of the watersheds represented in the new dataset.
2.2 Rescaling Using the USGS Geo Data Portal

To rescale the BCSD data to the spatial level of HUC-8 watersheds we employed the USGS Geo Data Portal (GDP) developed by the Center for Integrated Data Analytics (CIDA). The GDP offers a web service that is accessible via Python scripting language and is intended to “streamline the time-consuming and resource-intensive tasks associated with data access and manipulation that can inhibit the sharing and use of interdisciplinary environmental science data.” (Blodgett 2012). This process allows for the upload of HUC-8 watershed polygons which can then be matched with gridded climate data and processed to produce weighted average precipitation and temperature projections.
The GDP returns a weighted mean for the downscaled GCM across a given spatial area by using weighted grid statistical processing. In this case the given spatial areas are HUC-8 watersheds. Figure 2 demonstrates the weighted-averaging process. The dashed lines represent the downscaled GCM grid boundaries, the polygon represents a watershed, and the dots in the middle of each grid cell represent their respective centroids.

Figure 2 Example of Weighted Grid Averaging Used in Scaling of Climate Data to HUC-8 Polygons (adapted from http://cida.usgs.gov/enddat/UserGuide.jsp)

Equation 1-1 is used to calculate the weighted mean (μ). Each grid box has an area represented by $A_{grid,i}$ and a reported value (representing monthly average temperature or precipitation) represented as $P_i$. The area of the watershed polygon that lies within each grid is represented by $X_i$ and the total area of the watershed polygon is represented by $A_{shape}$. 
\[
\mu = \frac{\sum_{i=1}^{9} p_i \frac{X_i}{A_{grid,i}}}{\sum_{i=1}^{9} \frac{X_i}{A_{grid,i}}}
\] (1-1)

A polygon shapefile of the contiguous United States was used as the processing bounds for the GDP. This polygon included all HUC-8 watersheds as polygon features. The HUC-8 watershed boundaries were simplified by removing large numbers of excessive vertices that do not affect the polygon area or overall shape. Because the GDP must calculate distance between each set of vertices along a watershed boundary, the natural watershed boundary was found to require an unacceptable amount of processing time. To reduce this processing time while maintaining the spatial integrity of the watershed boundary, the number of vertices of each HUC-8 polygon was reduced in the shapefile. This process used the Ramer-Douglas-Peuker algorithm which removes vertices from a geometry. For the purpose of intersection with grid cells, the simplification of the HUC-8 polygons resulted in a negligible difference in the results. Figure 3 depicts a natural HUC-8 watershed boundary as well as the same boundary that has simplified vertices. Further information regarding the Geo Data Portal can be found at http://pubs.usgs.gov/of/2011/1157/pdf/ofr2011-1157.pdf.

Figure 3 Natural (Left) and Simplified (Right) Watershed Boundary of a HUC-8 Polygon
2.3 Processing

The GDP web service is accessible via a Python package published by the USGS called pyGDP. The package is openly available at the following internet address:

https://github.com/USGS-CIDA/pyGDP.

To use the pyGDP package, a simple Python script was constructed which imports the package, declares needed variables, and calls the “submitFeatureWeightedGridStatistics” method. This function call is central to the operation of the GDP and is provided here for reference purposes:

```python
def submitFeatureWeightedGridStatistics(self, geoType, dataSetURI, varID, startTime, endTime, attribute='the_geom', value=None, gmlIDs=None, verbose=None, coverage=True, delim='COMMA', stat='MEAN', grpby='STATISTIC', timeStep=False, summAttr=False, weighted=True)
```

The arguments which the method accepts are defined in Table 1. The method returns a delimited text file that contains the requested processed data.

<table>
<thead>
<tr>
<th>Argument</th>
<th>Title</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>geoType</td>
<td>WFS Feature Collection</td>
<td>A feature collection encoded as a WFS request or one of the supported GML profiles. This would be the simplified HUC-8 shapefile containing all watersheds as feature polygons.</td>
</tr>
<tr>
<td>dataSetURI</td>
<td>Dataset URI</td>
<td>The base data web service URI for the dataset of interest.</td>
</tr>
<tr>
<td>varID</td>
<td>Dataset Identifier</td>
<td>The unique identifier for the data type or variable of interest. This would be the downscaled model time-series of interest.</td>
</tr>
<tr>
<td>---------</td>
<td>--------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>startTime</td>
<td>Time Start</td>
<td>The date to begin analysis. All available dates were collected.</td>
</tr>
<tr>
<td>endTime</td>
<td>Time End</td>
<td>The date to end analysis. All available dates were collected.</td>
</tr>
<tr>
<td>Attribute</td>
<td>Feature Attribute Name</td>
<td>The attribute that will be used to label column headers in processing output.</td>
</tr>
<tr>
<td>Delim</td>
<td>Delimiter</td>
<td>The delimiter that will be used to separate columns in the processing output (i.e. a comma).</td>
</tr>
<tr>
<td>Stat</td>
<td>Statistics</td>
<td>Statistics that will be returned for each feature in the processing output, such as “MEAN” for the average value of the weighted grid statistic calculation.</td>
</tr>
<tr>
<td>Grpby</td>
<td>Group By</td>
<td>If multiple features and statistics are selected, this will change whether the processing output columns are sorted according to statistics or feature attributes.</td>
</tr>
<tr>
<td>timeStep</td>
<td>Summarize Time Step</td>
<td>If selected, processing output will include columns with summarized statistics for all feature attributes for each time step.</td>
</tr>
<tr>
<td>summAttr</td>
<td>Summarize Feature Attribute</td>
<td>If selected, processing output will include a final row of statistics summarizing all time steps for each feature attribute value.</td>
</tr>
<tr>
<td>weighted</td>
<td>Weighting</td>
<td>If selected, processing output will include weighting of grids who partially intersect with watershed polygon. This is desired in processing downscaled models.</td>
</tr>
</tbody>
</table>
The GDP divides the 234 BCSD models into different time-series for time period and variables. The historical time-series spans the years 1950 through 2006 while the future time-series spans the years 2006 through 2099. Each BCSD model contains a time-series for historical precipitation, historical temperature, future temperature, and future precipitation. A historical time-series is considered a run for a particular model while a future time-series for the same model may contain multiple runs. As a result, many models contain multiple future time-series and only one historical time-series. This results in 698 time-series needed to collect a complete database for all CMIP5 models. The 698 time series are represented as follows:

- **Historical Time Series (1950-2006)**
  - Precipitation: **115**
  - Temperature: **115**

- **Future Time Series (2006-2099)**
  - Precipitation: **234**
  - Temperature: **234**

- **Total Time Series: **698**

While a batch of polygons (HUC-8 watersheds) may be processed within a single request to the GDP without additional processing time, each time-series must be processed individually. Requesting more than one time series per request is beyond the limits of the web service. Therefore, each request would eventually process all 2,106 HUC-8 watersheds for a single time series.

Requests to the GDP take up to 15 minutes to process and return a time-series. Therefore,
the processing of the entire collection demands a significant amount of time. To manage this process, a program was written in the C# programming language to dispatch requests to the GDP as well as to retrieve and save the files returned. This program uses IronPython, an open-source implementation of the Python language and is integrated into the Microsoft .NET Framework (Foord, 2009). This allows for the creation of variables and file paths in C# that can be used directly in a Python script called from C#. Additionally, variable and file paths created in a Python script can be passed back to be used in the original C# script. Thus the Python scripting to communicate to the GDP can be managed within the overlying C# program. This code as well as others used in this work can be found at the following repository: https://github.com/CMIP5/HUC8Climate.

This program initially reads a comma separated value file (CSV) containing the names of all time-series which are then stored within an array. These data types will later be passed into the Python script and included in each request to the GDP. The program then calls a GDP request using the first data type in the array and waits for the GDP request to finish before continuing. Once the request has finished, an internet address where the time-series can be downloaded is returned. The program uses the address to retrieve the resulting file and stores it as a comma separated value file in the specified location. The process is then reiterated until all 698 time series are processed.

2.4 Database Implementation

To reduce the processing time required to retrieve the BCSD climate model time series for any HUC-8 watershed and any time period, we stored the processed HUC-8 level time series
data in an online database compatible with the Hydrologic Information System (HIS) of the Consortium of Universities for the Advancement of Hydrologic Science (CUAHSI). This particular database presented unique challenges in how it would be organized and hosted due to its large size. The final database contained approximately 1.5 billion data values as well as the associated metadata. A traditional CUAHSI Observations Data Model (ODM) database was originally created to host this data and required over 500 GB of disk space. This was concerning as it would not only require a significant amount of storage to host the data, but the large size of the database led to performance issues in trying to retrieve time series.

To address this issue a non-traditional file-based approach was tested and ultimately implemented successfully. This approach leverages a table-based structure to store all metadata and in so doing retains the ease of access and query that ODM offers. The difference is that the data values are not stored within a table, but are instead stored as comma separated value (CSV) files in a special folder on the same server as the metadata. The series catalog table contains links to the associated CSV file. Each CSV file contains a time series associated with one HUC-8 watershed, one variable, and one climate model run. This results in 1.5 million files. Compared with the previous 1.5 billion values that were assigned using the traditional approach, the disk storage size was reduced to 40 GB or by a factor of 10. The performance was also improved by using this new approach. The naming of the CSV files follows a predefined convention:

[HUC8 code]-[Method]_[Start year][Start month]-[End year][End month].csv

For example in the file name:

```
01010001-BCSD_0-125deg_pr_Amon_HadGEM2-CC_historical_r1i1p1_195001-200511.csv,
```
01010001 is the HUC-8 code number, 0-125deg_pr_Amon_HadGEM2-CC_historical_r1i1p1 is the method, 1950 is the start year, 01 is the start month, 2005 is the end year, and 11 is the end month. The method identifier contains the downscaling method name, variable, climate model name, and model run. The methods table in the database contains the method identifiers that are used to locate the corresponding CSV file.

The resulting database was installed on a web server and made available online through a CUAHSI HydroServer. Specifically we adapted the HydroServer Lite software using a MySQL database engine and Linux operating system. HydroServer Lite is open source software written in the PHP programming language and is available for download from http://hydroserverlite.codeplex.com (Conner et al. 2013; Kadlec and Ames 2012). The design of HydroServer Lite uses a Model View Controller (MVC) architecture pattern to separate the persistent storage (model), user interface design (view), and application logic (controller). Enabling the new CSV file data storage in HydroServer Lite required modifying the Method and DataValue classes in the model. The HydroServer with a WaterOneFlow web service (Kadlec et al. 2015; OGC 2012) and a graphical user interface was published on the server of Brigham Young University in Provo, Utah as part of the World Water Project (http://worldwater.byu.edu). The web service was also registered on the CUAHSI Water Data Center (WDC) catalog (hiscentral.cuahsi.org) which makes the database widely available through all CUAHSI products and services. The resulting database has been made available for wide distribution to hydrologists and water resources managers and educators via the WaterML format and standard CUAHSI data clients such as HydroDesktop (Ames et al. 2012) and the WaterML R package (Kadlec et al. 2015).
The address of the web services end point is:

http://worldwater.byu.edu/climate/services/ and has the following methods to query the climate data (Table 2):

**Table 2 Climate Services Feature Table**

<table>
<thead>
<tr>
<th>Method</th>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>GetSites</td>
<td>Geographic bounding rectangle</td>
<td>Show coordinates and details about all HUC-8 watersheds</td>
</tr>
<tr>
<td></td>
<td>(optional)</td>
<td></td>
</tr>
<tr>
<td>GetVariables</td>
<td>(none)</td>
<td>Show metadata about the variables (units, time units)</td>
</tr>
<tr>
<td>GetSiteInfo</td>
<td>HUC-8 code</td>
<td>Show what models and time range of data is available for a given HUC</td>
</tr>
<tr>
<td>GetValues</td>
<td>Variable code (tas or pr), HUC-8</td>
<td>Get the climate forecast values for a given HUC, variable, time range and</td>
</tr>
<tr>
<td></td>
<td>code, Start date (optional), End</td>
<td>model</td>
</tr>
<tr>
<td></td>
<td>data (optional), Model code</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(optional)</td>
<td></td>
</tr>
</tbody>
</table>

In addition to the web service interface to our new climate database, it is accessible through the CUAHSI network including the open source GIS-based software package HydroDesktop (Ames et al. 2012) which allows users to access the datasets in time series and tabular form as well as to download the datasets locally. The WaterML R package can also be used to access the data for local analysis within the R statistical software system (Kadlec et al. 2015). For quick access to climate model predictions for a specific HUC-8 watershed, the web-based HydroServer user interface on http://worldwater.byu.edu/climate/ui/ can also be used. Retrieving the data using this interface has two steps. First the user selects the HUC-8 mid-point
on a Google map (Figure 4) following which the site details page can be used to select the climate model and time range, and to download the data in CSV format (Figure 5).

Figure 4 HydroServer Google Map User Interface for selecting the HUC-8 Watershed of Interest
Figure 5 Site Details User Interface on HydroServer Lite Where Climate Data can be accessed
3 RESULTS

3.1 Discussion

This effort has resulted in CMIP5 GCM data being scaled to the level of HUC-8 watersheds and made openly available for water resource engineers and managers. Through a variety of interfaces users can access this data from any computer with a web browser. Figure 6 represents a workflow diagram showing the progression of this work from GCMs to web services.

Figure 6 Work Flow Diagram of Process Translating CMIP5 Data to Openly Available HUC-8 Scaled Data

3.2 Use Case 1: Examining Climate Models

The database has been designed for use with the Climate Analysis Toolkit (an extension to HydroDesktop) which allows for the statistical analysis of CMIP5 datasets via hybrid-delta
ensemble analysis (Hamlet et al. 2010) and for the processing to macro-scaled hydrologic models. With the addition of the HUC-8 downscaled climate model database users can use the tool to compare model predictions for any HUC-8 watershed.

Given that the CMIP-5 collection contains 234 models it is often useful to compare the anticipated changes in temperature and precipitation that each model provides. Using the Climate Analysis Toolkit we have compared all models for a particular HUC-8 watershed the data for which was downloaded from the database. Figure 7 demonstrates a scatterplot of all 234 climate models.

![Figure 7 Scatter Plot of All 234 CMIP5 Models for One Watershed Produced in HydroDesktop’s Climate Analysis Toolkit](image-url)
These figures represent the temperature change and percent precipitation change when comparing two thirty year spans across the model (1980-2010 and 2030-2060). The changes are then plotted on the above graph and grouped into ensembles that represent scenarios. The scenarios are determined by quantile and represent the scenarios shown in Table 3.
### Table 3 Common Climate Change Scenarios

<table>
<thead>
<tr>
<th>Climate Change Scenario</th>
<th>Percentile in terms of temperature change</th>
<th>Percentile in terms of precipitation change</th>
</tr>
</thead>
<tbody>
<tr>
<td>More Warming – Drier</td>
<td>80</td>
<td>20</td>
</tr>
<tr>
<td>More Warming – Wetter</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Median</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>Less Warming – Drier</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>Less Warming – Wetter</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>

These scenarios represent potential changes in climate for a particular HUC-8 watershed and can be helpful in water resource planning and study.

#### 3.3 Use Case 2: Comparing Watersheds

The database can also be used to quickly compare a single model’s forecast across watersheds in the United States. To demonstrate this potential use case, we downloaded one model for temperature projections across all HUC-8 watersheds in the United States from the newly created database. This data was downloaded in a comma separated value format which allowed us to display it in Microsoft Excel. Within Excel average temperatures for each watershed were then calculated for a 30 year period at the beginning of the time series (2010-2040) and the end of the time series (2069-2099). The change in average temperature between these two time periods was then calculated for each watershed. Within ESRI ArcMap a polygon shapefile of simplified HUC-8s was added as a layer. By accessing the attribute table we were
able to add an average temperature attribute and then add the data from Excel matching each HUC-8 code in the attribute table to its corresponding calculated change in average temperature. Finally, the polygon’s properties were adjusted to assign a color to each watershed feature based on the intensity of the average temperature change. The result of this exercise is shown in Figure 8 as a simple HUC-8 based view of temperature as a heat map. This map represents average temperature change when comparing two periods in time and graphically illustrates temperature change according to the model “BCSD_0-125deg_tas_Amon_ACCESS1-0_rcp45_r1i1p1_200601-210012”.

Figure 8 Temperature Change of HUC-8 Watershed Regions as Described by a Model From Database
This particular model predicted greater temperature change in the Upper Midwestern watersheds and less change along coastal watersheds.

3.4 Planned implementation

The Columbia River Basin Climate Impacts Assessment, which addresses Section 9503 of the SECURE Water Act (SECURE), authorizes the United States Bureau of Reclamation to assess the risks to water supplies and demands posed by climate change within major river basins in the Western United States. The study uses macroscale hydrologic models to predict future trends of water availability and climate behavior. The hydrologic models used in the study need to be adjusted by change factors generated using hybrid delta ensembles of downscaled GCMs. Using the newly created HUC-8 watershed GCM model database, Bureau of Reclamation hydrologists will be able to input downscaled GCM projection and successfully generate watershed models that are being used in this study.

3.5 Future Development

This database is limited to HUC-8 watersheds within the contiguous United States. However the methods employed can be applied to other areas (e.g. outside the conterminous United States) as rescaling services for these areas become available. Further development of downscaled GCM gridded data would make this a possibility. Also, if individuals need to develop downscaled GCM results for specific drainage areas smaller than a HUC-8, they can access the GDP directly. The methods presented in this paper can be adapted to such a need.

Other future work may include using the GDP to scale GCM projections to an even finer spatial resolution. In keeping congruent with the use of USGS HUC watersheds, a future
database could be processed through the same algorithm published herein but with HUC-12 or even HUC-16 watershed polygons. While this would employ a greater number of polygons and therefore take more processing time, it would nonetheless be feasible.
4 CONCLUSION

By pre-processing climate models through the GDP a significant amount of effort has been completed for users. Indeed, the processing time for all 234 climate models did in fact take 116 hours. By making these pre-processed model datasets available via CUAHSI HIS, hydrologists, engineers and other users can now access this information in a few seconds from any web browser, resulting in significant savings of both time and money.

The HUC-8 Based Downscaled CMIP-5 Climate Model Forecast Dataset is the first to offer downscaled global circulation models at the scale of nationally predefined watersheds. Through this database water resource managers and engineers have access to hundreds of climate projections for each HUC-8 basin through the year 2099, allowing for the study of, and preparation for, climate change.
REFERENCES


