

PROJECTED AVERAGE SUMMER AIR TEMPERATURE INCREASES AND THE IMPLICATIONS  
FOR PHILADELPHIA'S SURFACE DRINKING WATER SUPPLY

by

Julia Rockwell  
Dr. Martin Doyle, Advisor  
December 2014

Masters project submitted in partial fulfillment of the  
requirements for the Master of Environmental Management degree in  
the Nicholas School of the Environment of  
Duke University

2014

## ABSTRACT

Water managers are faced with numerous uncertainties that need to be addressed in the development of long-term planning initiatives and large-scale investment decisions. One of the primary and perhaps most far-reaching of these uncertainties is climate change. The objective of this project is to utilize one aspect of projected climate change impacts, increasing average summer air temperature, to understand potential impacts to surface drinking water supply temperatures in the Schuylkill River at Philadelphia, PA. The project consists of three major components. As an initial step, climate model output from the Coupled Model Intercomparison Project Phase 5 (CMIP5) was evaluated for the Northeast US and Philadelphia by mapping and analyzing Network Common Data Form (NetCDF) files for near-surface air temperatures in Matlab. The evaluation of climate model output included model validation for six selected CMIP5 Global Climate Models (GCMs), as well as future projections using the Representative Concentration Pathway 8.5 (RCP8.5) climate scenario. Secondly, this project aimed to develop a statistical relationship between air and surface water temperatures in Philadelphia using publicly available data from the United States Geological Survey (USGS) and Daymet. Following the aforementioned empirical analyses, research was performed to provide insight regarding the impact of increased surface water supply temperatures on the formation of disinfection byproducts (DBPs) during drinking water treatment. The three most accurate GCMs for the Northeast US and Philadelphia indicate that the average air temperature over June, July, August and September (JJAS) will increase approximately 2°C by mid-century. Assuming the RCP8.5 climate scenario prevails beyond mid-century, the results indicate that the average JJAS near-surface air temperature may increase by as much as 5.7°C in the Northeast US and 5.3°C in Philadelphia by 2100. For the Schuylkill River at Philadelphia, statistical analyses reveal that air temperature explains the majority of variation in water temperature over the time period of analysis, from 1999-2001 and 2011-2013 for the months of JJAS. Projected increases in average JJAS air temperature are expected to increase average JJAS surface water temperature in the Schuylkill River at Philadelphia by approximately 1.69°C (3.05°F) and 4.23°C (7.62°F) by 2050 and 2100, respectively. Current climate science needs to be directly related to actionable adaptation initiatives. The outcome of this study directly links one aspect of climate change to a potential drinking water impact, with the goal of providing actionable information to inform future operational and supply management strategies.

## Table of Contents

<b>I. Introduction .....</b>	<b>5</b>
<b>II. Evaluation of Coupled Modeled Intercomparison Project 5 (CMIP5) Climate Model Output .....</b>	<b>6</b>
<b>Background.....</b>	<b>6</b>
<b>Data and Methods .....</b>	<b>8</b>
<b>Analysis &amp; Results .....</b>	<b>9</b>
Model Validation for the Six GCMs.....	9
Projected Increases in Air Temperature .....	14
<b>III. Development of a Statistical Relationship for Air and Water Temperatures in Philadelphia .....</b>	<b>16</b>
<b>Background.....</b>	<b>16</b>
Linear Regression Models .....	17
Monthly Average Nonlinear Regression Model.....	19
Weekly Average Nonlinear Regression Model.....	20
<b>Data.....</b>	<b>23</b>
Water Temperature Data.....	23
Air Temperature Data .....	25
<b>Methods.....</b>	<b>25</b>
<b>Analysis &amp; Results .....</b>	<b>26</b>
Monthly Average Data.....	26
Weekly Average Data .....	31
<b>IV. Summer Average Surface Water Temperature Projections .....</b>	<b>34</b>
<b>V. Increasing Surface Water Temperatures and Disinfection Byproduct (DBP) Formation .....</b>	<b>36</b>
<b>Background on Disinfection Byproducts (DBPs) .....</b>	<b>36</b>
<b>DBP Formation Potential .....</b>	<b>37</b>
<b>VI. Discussion and Conclusions .....</b>	<b>39</b>
<b>VII. Future Research Opportunities.....</b>	<b>41</b>
<b>Acknowledgements .....</b>	<b>42</b>
<b>References .....</b>	<b>43</b>
<b>Appendices .....</b>	<b>46</b>

## List of Tables

<b>Table 1.</b> Dataset information for the climate model evaluation process .....	9
<b>Table 2.</b> Percent errors (based on Kelvin) for output from six CMIP5 GCMs .....	14
<b>Table 3.</b> Near-surface air temperatures for the average of the three best-performing GCMs .....	15
<b>Table 4.</b> Near-surface air temperature changes relative to the historical dataset .....	15
<b>Table 5.</b> Properties of the linear regression models for weekly average and monthly average air and water temperatures in Philadelphia, PA. June, July, August, September, 1999-2001 and 2011-2013 .....	33
<b>Table 6.</b> Projected increases in average JJAS air and water temperatures in Philadelphia .....	34

## List of Figures

<b>Figure 1.</b> Radiative forcings associated with RCP climate scenarios .....	7
<b>Figure 2.</b> Estimated historical CO <sub>2</sub> emissions compared to the IS92, SRES and the RCP climate scenarios .....	8
<b>Figure 3.</b> NOAA mean monthly observed near-surface air temperatures averaged over January 1985-December 2004. ....	11
<b>Figure 4.</b> MPI-ESM-LR output for mean monthly near-surface air temperatures averaged over January 1985-December 2004.....	11
<b>Figure 5.</b> Trends in the Northeast climatological mean for six GCMs from CMIP5.....	13
<b>Figure 6.</b> Graphical representation of the parameters used in the weekly average air/water temperature nonlinear regression model .....	21
<b>Figure 7.</b> Location of the USGS Fairmount gauge relative to PWD’s drinking water treatment plants....	24
<b>Figure 8.</b> Range of observed monthly average air and water temperatures in Philadelphia, PA relative to the global nonlinear regression model developed by Punzet et al. (2012).....	26
<b>Figure 9.</b> Monthly average air/water temperature relationship, Philadelphia, PA. June, July, August and September, 1999-2001 and 2011-2013 .....	27
<b>Figure 10.</b> Air and water temperature time series data and predicted values for June, July, August and September, Philadelphia, PA. January 1999-November 2001 .....	29
<b>Figure 11.</b> Air and water temperature time series data and predicted values for June, July, August and September, Philadelphia, PA. June 2011-December 2013 .....	29
<b>Figure 12.</b> Standard deviations associated with the absolute error (predicted-observed) of available average predicted water temperatures by month, Philadelphia, PA. June, July, August, September, 1999-2001 and 2011-2013 .....	30
<b>Figure 13.</b> Weekly average air/water temperature relationship, Philadelphia, PA. June, July August, September, 1999-2001 and 2011-2013 .....	32
<b>Figure 14.</b> Observed and projected increases in average JJAS air and water temperatures in the Schuylkill River at Philadelphia.....	35

## I. Introduction

Climate change is a phenomenon that will impact populations, ecosystems, and natural resources on a global scale. The implications of climate change, however, will significantly vary on regional and local scales. In order for water utilities to ensure that current levels of service can be maintained well into the future, substantial uncertainties with respect to climate change must be addressed in some capacity. The primary objective of this Master's Project (MP) is to utilize one aspect of projected climate change, increasing air temperature, to understand the potential impacts to surface drinking water supplies in the City of Philadelphia. Specifically, this project consists of the following three major components, referred to as Parts 1-3 throughout this report:

- Part 1: Evaluate historical and Representative Concentration Pathway 8.5 (RCP8.5) climate model output from the Coupled Model Intercomparison Project Phase 5 (CMIP5) for Philadelphia, PA and the Northeast region of the United States by mapping and analyzing Network Common Data Form (NetCDF) files for near-surface air temperature in Matlab.
- Part 2: Develop a relationship between average June, July, August and September (JJAS) air and surface water temperatures at USGS gauge 01474500 using publicly available data and statistical analyses. USGS gauge 01474500 is located near the Philadelphia Water Department's (PWD's) two drinking water intakes on the Schuylkill River. Based on the relationship between air and water temperatures, provide projections for average JJAS surface water temperature increases in Philadelphia by mid and end-of-century assuming the RCP8.5 climate scenario.
- Part 3: Provide an overview of one potential implication of increasing surface water temperatures: the formation of disinfection byproducts (DBPs) during the drinking water treatment process.

Impacts from climate change are now unavoidable, necessitating effective adaptation strategies for water utilities. The objective of this study, as outlined above, is to directly link current climate change projections with potential impacts to drinking water supplies, with the goal of providing actionable information to inform future operational and supply management strategies. Each component of the project is presented separately, followed by an overarching discussion and conclusions section at the end of this report.

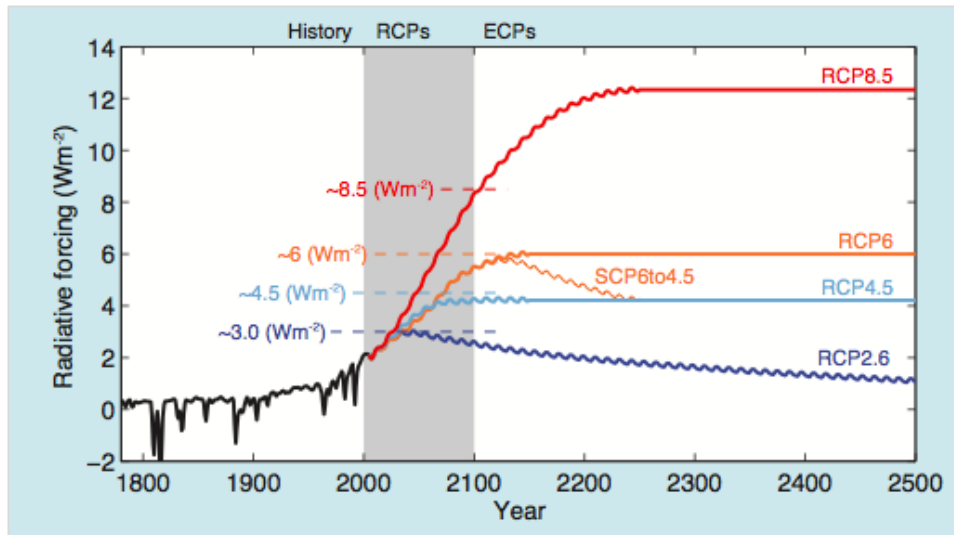
## II. Evaluation of Coupled Modeled Intercomparison Project 5 (CMIP5) Climate Model Output

### Background

The Coupled Modeled Intercomparison Project Phase 5 (CMIP5) is the fifth phase of a series of coordinated climate model assessments that were used to inform the Intergovernmental Panel on Climate Change's (IPCC's) fifth Assessment Report (AR5) (PCMDI, 2014). A total of six different global climate models (GCMs) were randomly selected for analysis in Part 1. The GCMs, which are listed below, are included in CMIP5.

- CCSM4, National Center for Atmospheric Research (NCAR)
- GFDL-CM3, National Oceanic and Atmospheric Administration (NOAA) Geophysical Fluid Dynamics Laboratory
- IPSL-CM5A-LR, Institut Pierre-Simon Laplace
- MIROC5, Atmosphere and Ocean Research Institute, University of Tokyo
- MPI-ESM-LR, Max Planck Institute for Meteorology
- MRI-CGCM3, Meteorological Research Institute (ESGF, 2014)

Part 1 of this project consists of two primary areas of analysis. The first objective of Part 1 is to validate the performance of each GCM by comparing historical modeled output to observed near-surface air temperatures. The second objective of Part 1 is to project near-surface air temperatures using the RCP8.5 output from each GCM. The RCP climate scenarios represent a new set of inputs for the modeling approaches referenced in the Intergovernmental Panel on Climate Change (IPCC) AR5 Report. There are four RCP scenarios identified as 2.6, 4.5, 6.0 and 8.5, each representing a different greenhouse gas and aerosol emissions scenario (IPCC, 2014). Unlike earlier climate sceneries, each RCP includes mitigation efforts that align with long-term policy objectives (Peters et al., 2013). The RCP scenarios are all associated with a stabilized or peak radiative forcing (RF) value for the 21<sup>st</sup> century. RCP8.5 represents the highest emissions scenario that results in a RF of approximately 8.5 W/m<sup>2</sup> by the end of the century (IPCC, 2014). Figure 1 below is included in the IPCC AR5 report and illustrates the trends in total (natural and anthropogenic) RF for the four RCP scenarios. The trends that are exhibited following year 2100 are termed extended concentration pathways (ECPs).



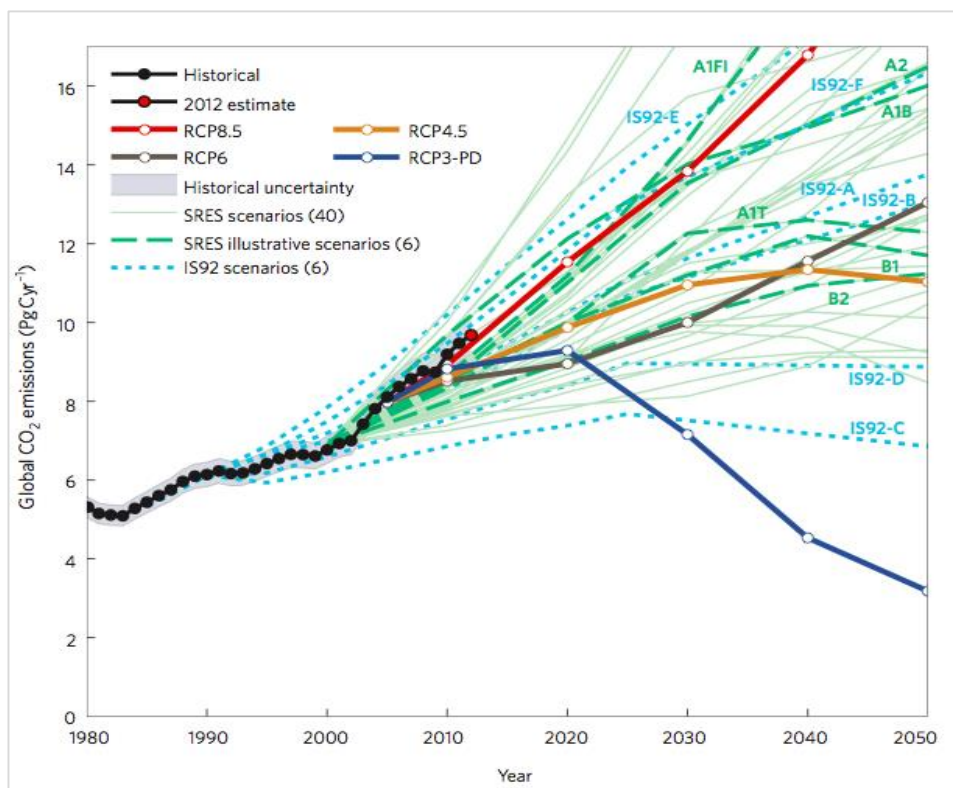
**Figure 1.** Radiative forcings associated with RCP climate scenarios  
*Image Source: Figure 4, Meinshausen et al., 2011*

The model validation process and near-surface air temperature projections were performed for four areas of analysis, consisting of different spatial and temporal domains. The four areas of analysis consist of the following: 1) climatological mean for the Northeast, 2) June, July, August and September (JJAS) mean for the Northeast, 2) climatological mean for Philadelphia, and 4) JJAS mean for Philadelphia.

Previous reports cite the implications of climate change for the Northeast US, including a 2006 report published by the Union of Concerned Scientists (NECIA, 2006). This report indicates that since 1970, temperatures in the Northeast US have been increasing at a rate of approximately 0.5°F per decade (NECIA, 2006). The warming trend is accompanied by climatic changes including a higher frequency of hotter days. From 1961-1999, Philadelphia experienced an average of 18 days per year with temperatures higher than 90.2°F (ICF, 2014). Downscaled projections for the RCP8.5 climate scenario in Philadelphia indicate that the city could see 98 days per year, on average, above 90.2°F by the end of the century (ICF, 2014).

Given the potential for notable warming in the Northeast as well as an increased frequency of hotter days, Part 1 of this project assesses both the near-surface air temperature climatological mean as well as the mean for the months of JJAS. The climatological mean is defined as the average of temperature values across all months within the time period of analysis and over the entire region or area under consideration. The high temperature and drought-prone months of JJAS were considered important to include as a separate temporal domain in the analysis.

For the four areas of analysis discussed above, near-surface air temperature projections were evaluated using the RCP8.5 scenario, the highest RCP emissions scenario. The RCP8.5 scenario was selected given recently documented trends in carbon emissions. Figure 2 below compares CO<sub>2</sub> emissions from the burning of fossil fuels, gas flaring, and cement manufacturing with the four generations of emissions scenarios that were produced as a part of the IPCC process: the Scientific Assessment 1990 (SA90), IPCC Scenarios 1992 (IS92), Special Report on Emissions Scenarios (SRES), and the RCPs (Peters et al., 2013). It is apparent from Figure 2 that observed CO<sub>2</sub> emissions track the A11 and A2 SRES scenarios, as well as the RCP8.5 climate scenario, all of which lead to the greatest projected temperature increases (4.2-5.0°C in 2100) compared to all other climate scenarios. In fact, recently observed emissions are tracking at slightly higher levels than the RCP8.5 scenario (Peters et al., 2013).



**Figure 2.** Estimated historical CO<sub>2</sub> emissions compared to the IS92, SRES and the RCP climate scenarios  
*Image Source: Figure 1, Peters et al., 2013*

## Data and Methods

The data used for this project originate from two primary sources: 1) monthly average near-surface air temperature output from six GCMs for the CMIP5 (ESGF, 2014) and 2) monthly average observational near-surface air temperature data from the National Oceanic and Atmospheric Administration's (NOAA's) National Center for Environmental Prediction (NCEP)/NCAR Reanalysis Project (NOAA, 2014). Data for Part 1 were analyzed by opening and evaluating Network Common Data Form (NetCDF)

files in Matlab. The NetCDF near-surface air temperature values were evaluated over two different spatial domains. The first area of analysis encompasses the entire Northeast region, spanning 81°W to 66°W longitude and 37°N to 48°N latitude. The Philadelphia area was analyzed for each GCM and the NOAA observational data by determining the nearest grid point in the NetCDF files to a specified latitude/longitude representing the location of USGS gauge 01474500 at Fairmount Dam. The coordinates of the closest grid points to the USGS Fairmount gauge for each model and the NOAA observed dataset are included for reference in Appendix A, Table 1A. The USGS Fairmount gauge is located at approximately 75°W and 40°N, a location on the Schuylkill River near Philadelphia’s Queen Lane and Belmont drinking water intakes. It should be noted that no downscaling methods were applied to the GCM data. A summary of the datasets used in Part 1 is provided in Table 1 below.

**Table 1.** Dataset information for the climate model evaluation process

Data	Source	Variable ( <i>Label</i> )	Time Period
Historical	Six GCMs*	Near-surface air temperature ( <i>tas</i> )	Jan. 1985-Dec. 2004
Observational	NOAA NCAR/NDEP	Near-surface air temperature ( <i>air</i> )	Jan. 1985-Dec. 2004
Mid-Century Projections (RCP8.5)	Six GCMs	Near-surface air temperature ( <i>tas</i> )	Jan. 2031-Dec. 2050
End-of-Century Projections (RCP8.5)	Six GCMs	Near-surface air temperature ( <i>tas</i> )	Jan. 2081-Dec. 2100

\*Six GCMs: GFDL-CM3, CCSM4, MIROC5, MRI-CGCM3, IPSL-CM5A-LR, MPI-ESM-LR

The scope of work for Part 1 began with model validation. Each GCM was evaluated by comparing historical model output with observed NOAA data from the time period January 1985 to December 2004. The goal of the model validation process was to determine how accurately each of the six GCMs simulates near-surface air temperature conditions in the Northeast region and near Philadelphia. The second portion of Part 1 involves evaluating projected near-surface air temperature values from the six GCMs for mid-century and end-of-century data using the RCP8.5 climate scenario. Both the model validation process and projections analysis utilize the climatological mean and the mean near-surface air temperature for JJAS for analytic and comparison purposes.

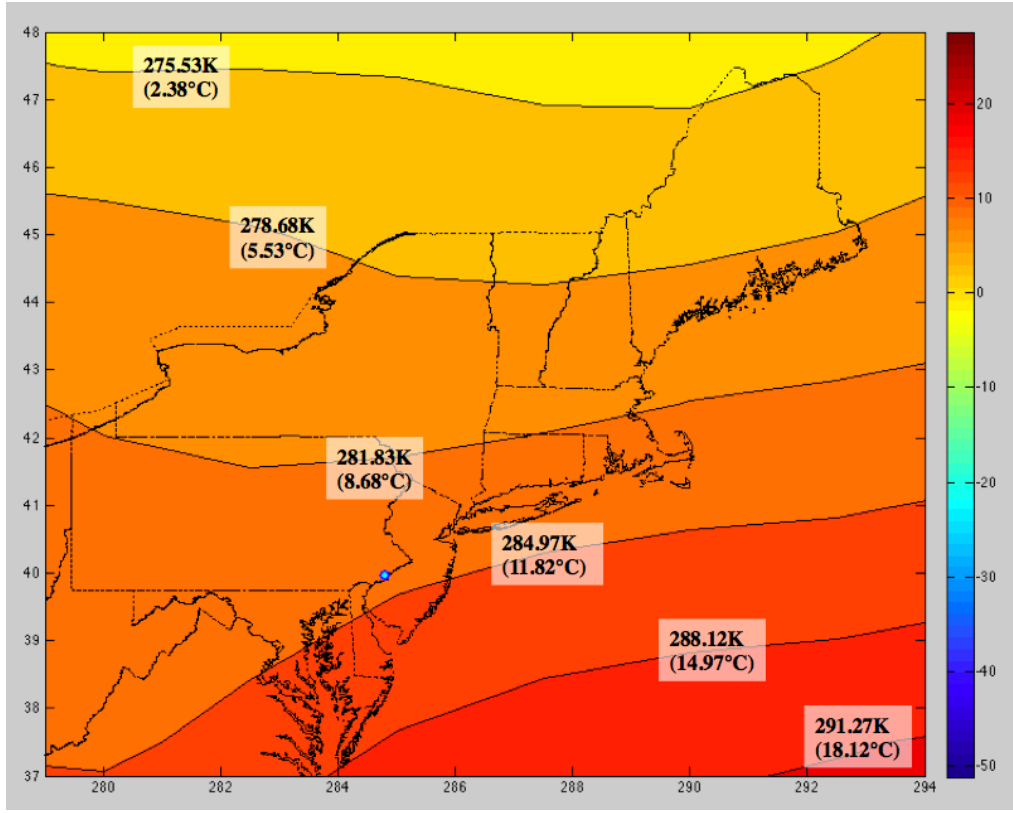
## Analysis & Results

The following section presents results for the two major components of this analysis: model validation and projected near-surface air temperatures under the RCP8.5 climate scenario.

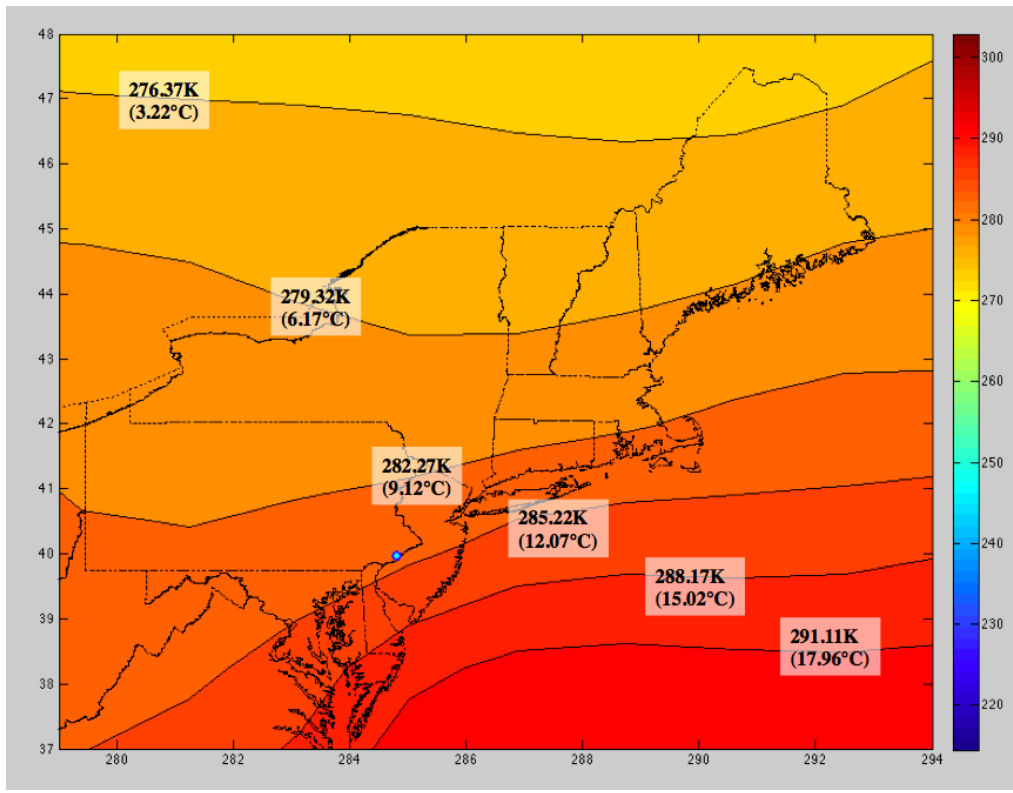
### Model Validation for the Six GCMs

Monthly mean near-surface air temperature data were analyzed for a total of six GCMs. As a result of the model validation process, the three best performing models were identified for each area of analysis. Mapping tools can be used to aid in model validation and to help visualize the spatial distribution of air temperature values. As an example of the maps that can be produced using NetCDF model output in

Matlab, two contour maps of the Northeast are presented below in Figures 3 and 4. The first map illustrates the temperature contours for the climatological mean of the NOAA monthly average observational data from January 1985 to December 2004. The second map illustrates the climatological mean obtained from averaging monthly near-surface air temperature output from the MPI-ESM-LR model over the same historical time period. The blue dot on both maps indicates the location of USGS gauge 01474500. Additional contour maps illustrating historical and projected air temperatures for the months of June, July, August, and September for both the observed data and MPI-ESM-LR model output are contained in Appendices B and C.



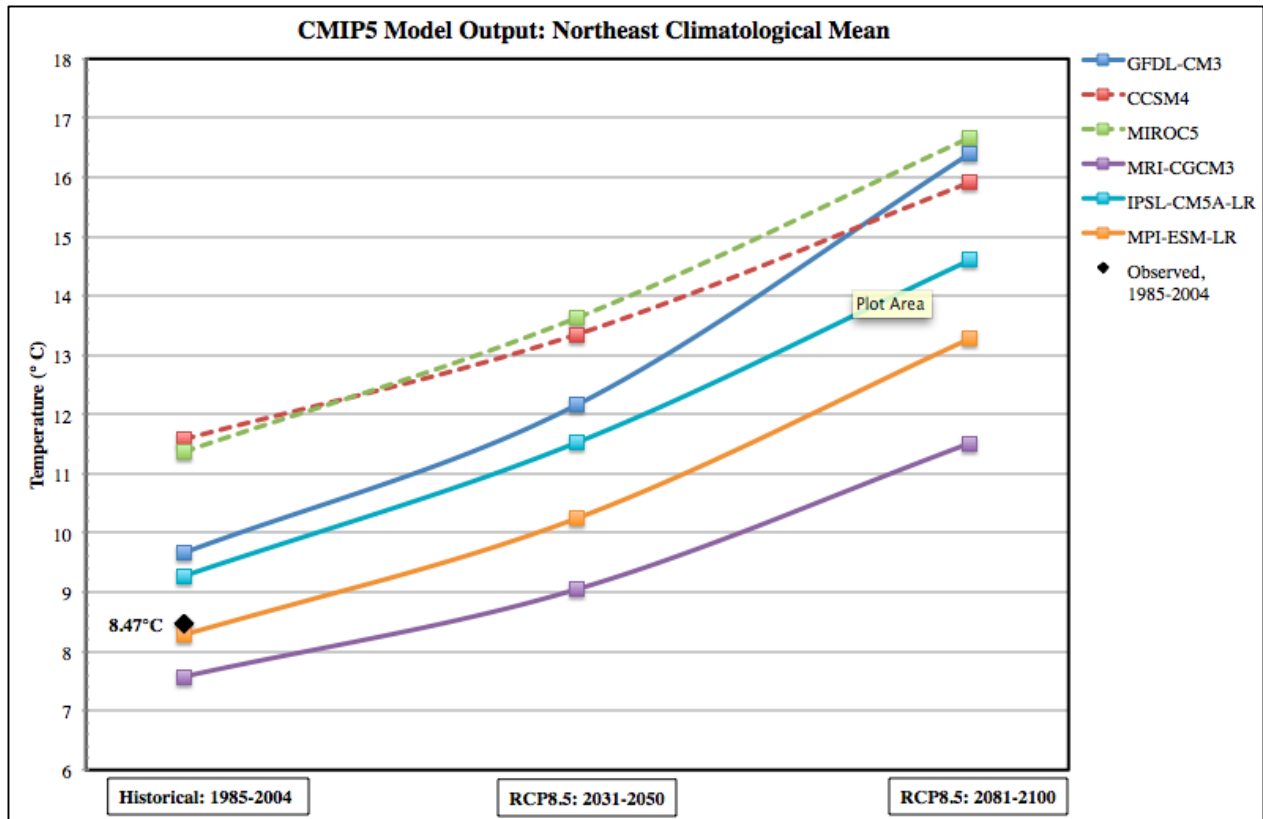
**Figure 3.** NOAA mean monthly observed near-surface air temperatures averaged over January 1985-December 2004. The blue dot indicates the location of USGS gauge 01474500.



**Figure 4.** MPI-ESM-LR output for mean monthly near-surface air temperatures averaged over January 1985-December 2004. The blue dot indicates the location of USGS gauge 01474500.

The maps above illustrate that the MPI-ESM-LR model fairly accurately represents the climatological mean of observed near-surface air temperatures for the time period from January 1985 to December 2004: the five northernmost temperature contours displayed from the MPI-ESM-LR model are only slightly higher (less than 1K) than the observed NOAA values, with the geographic distribution of contours falling in nearly the same areas throughout the region. The sixth and southernmost temperature contour is a slightly higher value in the modeled data compared to the observed data. This discrepancy may be explained by the fact that the southernmost contour on both maps is situated in slightly different geographic locations.

A complete comparison of the MPI-ESM-LR model to the five other GCMs can be viewed in Figure 5. The data plotted in Figure 5 illustrate the Northeast climatological mean for near-surface air temperatures calculated over historical, RCP8.5 mid-century and RCP8.5 end-of-century time periods using the six different GCMs. The two dashed lines indicate the poorest performing models relative to the observed value. The distribution of historical GCM output compared to the observed value of 8.47°C confirms that the MPI-ESM-LR model provides the most accurate estimate of the Northeast climatological mean from January 1985 to December 2004, while the CCSM4 and MIROC5 models clearly provide the least accurate output. Appendix A contains additional plots comparing the output from all six GCMs for all areas of analysis included in the study.



**Figure 5.** Trends in the Northeast climatological mean for six GCMs from CMIP5

Maps and plots alone, however, do not provide for consistent comparison between the observed NOAA data and each GCM. To improve the accuracy and ease of comparability among all GCMs and the observed data, the percent error relative to observed values were calculated for the near-surface air temperatures (Kelvin) for the four major areas of analysis: Northeast climatological mean, Northeast JJAS mean, Philadelphia climatological mean, and Philadelphia JJAS mean. The results from these calculations are displayed in Table 2 below.

**Table 2.** Percent errors (based on Kelvin) for output from six CMIP5 GCMs

CMIP5 Model	NE Climatological Mean	NE JJAS Mean	Phila. Climatological Mean	Phila. JJAS Mean
GFDL-CM3	0.421%	0.241%	0.243%	0.210%
CCSM4	1.101%	1.002%	0.663%	0.817%
MIROC5	1.028%	1.012%	0.824%	1.019%
MRI-CGCM3	-0.318%	-0.210%	-0.133%	0.076%
IPSL-CM5A-LR	0.285%	0.492%	-0.445%	0.032%
MPI-ESM-LR	-0.072%	-0.140%	-0.060%	0.022%

As depicted in Table 2, the CCSM4 and MIROC5 models perform least well for all four major areas of analysis. The third least accurate model differs depending on the study area under consideration. For the Northeast climatological mean, the GFDL-CM3 model results in the third least accurate simulated value. The same is true for the Philadelphia JJAS mean. For both the Northeast JJAS mean and the Philadelphia climatological mean, the IPSL-CM5A-LR model provides the third least accurate simulated set of values. As a result of the percent error calculations, the output from the following three GCMs were carried forward in the analysis:

- Northeast Climatological Mean: MRI-CGCM3, IPSL-CM5A-LR, MPI-ESM-LR
- Northeast JJAS Mean: GFDL-CM3, MRI-CGCM3, MPI-ESM-LR
- Philadelphia Climatological Mean: GFDL-CM3, MRI-CGCM3, MPI-ESM-LR
- Philadelphia JJAS Mean: MRI-CGCM3, IPSL-CM5A-LR, MPI-ESM-LR

By averaging the results from the three best performing models for each study area, the percent errors were reduced below the minimum individual model percent errors for all areas of analysis except the Philadelphia JJAS mean. The percent error for the average of the three models for JJAS in Philadelphia is 0.043%, while the lowest observed percent error when evaluating the individual models was 0.022%, from the MPI-ESM-LR model.

### Projected Increases in Air Temperature

Table 3a below summarizes the results from averaging the three most accurate models for each area of analysis. Estimated changes in near-surface air temperatures given the averaged results are provided in Table 3b.

**Table 3.** Near-surface air temperatures for the average of the three best-performing GCMs

DATA	NE Climatological Mean	NE JJAS Mean	Phila. Climatological Mean	Phila. JJAS Mean
<b>Historical</b> Jan 1985-Dec 2004	281.52K (8.37°C)	291.49K (18.34°C)	284.51K (11.36°C)	294.17K (21.02°C)
<b>RCP8.5 Mid-Century</b> Jan 2031-Dec 2050	283.42K (10.27°C)	293.66K (20.51°C)	286.43K (13.28°C)	296.29K (23.14°C)
<b>RCP8.5 End-of-Century</b> Jan 2081-Dec 2100	286.28K (13.13°C)	297.17K (24.02°C)	289.46K (16.31°C)	299.46K (26.31°C)

**Table 4.** Near-surface air temperature changes relative to the historical dataset

PROJECTION	NE Climatological Mean	NE JJAS Mean	Phila. Climatological Mean	Phila. JJAS Mean
<b>RCP8.5</b> Mid-Century (2050)	+1.90°C (+3.42°F)	+2.17°C (+3.91°F)	+1.91°C (+3.44°F)	+2.12°C (+3.82°F)
<b>RCP8.5</b> End-of-Century (2100)	+4.76°C (+8.57°F)	+5.69°C (+10.24°F)	+4.95°C (+8.91°F)	+5.30°C (+9.54°F)

The temperature changes that result from averaging the three most accurate models for the Northeast region and Philadelphia indicate that the climatological mean and summer months mean will increase by approximately 2°C by the middle of the century for all areas of analysis. Assuming the RCP8.5 climate scenario prevails beyond mid-century, the results indicate that the summer months (JJAS) mean for near-surface air temperature may increase by as much as 5.7°C (10.2°F) in the Northeast region and 5.3°C (9.5°F) in Philadelphia. All the above calculated temperature changes assume the RCP8.5 climate scenario, which, as noted earlier, closely represents the current trend in CO<sub>2</sub> emissions.

### **III. Development of a Statistical Relationship for Air and Water Temperatures in Philadelphia**

The relationship between air temperature and water temperature has been studied extensively for a variety of reasons, most notably the impacts of temperature on a surface water system's physical, chemical and biological processes (Punzet et al., 2012). Temperature serves an important role as a water quality indicator, influencing a suite of water quality parameters, including dissolved oxygen (DO), biological oxygen demand (BOD), nitrogen concentrations and total coliform bacteria concentrations (Ducharne, 2008; Punzet et al., 2012). Surface water temperatures can also significantly impact a river's aquatic ecosystem (Punzet et al., 2012) and can even have an economic influence when temperature requirements must be met for applications in industry, agriculture and recreation (Webb et al., 2008). More recently, the air/water temperature relationship has gained newfound attention in relation to the anthropogenic impacts, including climate change, on water resources. The ability of global climate models (GCMs) to accurately predict various climate parameters under a range of carbon emissions scenarios continues to improve. Air temperature is one, if not the most, accurately predicted climate parameter (Lau et al., 1996; Punzet et al., 2012). Consequently, when assessing the impacts of climate change on water resources, specifically surface water temperature, the accuracy of air temperature projections is paramount. Furthermore, the development of an air/water temperature relationship that relies on air temperature as the only input can serve as a powerful tool to water resource managers who are trying to understand the potential impacts of climate change on their source water supplies.

The following section will explore the observed air/water temperature relationship on the Schuylkill River near two of the Philadelphia Water Department's (PWD's) drinking water intakes. The goal of this section is to first determine a predictive relationship for water temperature given air temperature using a regression model. A predictive relationship between air and water temperatures will be determined on both a monthly average and weekly average timescale for the warm and drought-prone months of June, July, August and September (JJAS). Once a relationship between air and water temperatures is determined, projected air temperature increases attributed to the Representative Concentration Pathway 8.5 (RCP8.5) carbon emissions scenario will be used to assess potential increases in average JJAS surface water supply temperatures on the Schuylkill River within this century.

#### **Background**

Previous research has tried to understand the relationship between air and water temperatures using different modeling techniques that vary considerably in complexity (Punzet et al., 2012). A complex approach to air/water temperature modeling involves the use of a deterministic, or physical, model that is founded upon thermal budget equations (Punzet et al., 2012). This method of analysis involves rigorous

data collection and extensive inputs beyond just air temperature (Punzet et al., 2012). A second approach to air/water temperature modeling involves the use of stochastic models that are based upon a statistical relationship between the response variable, water temperature, and various predictor variables that can include, but are not limited to, air temperature and stream flow (Punzet et al., 2012). A simple and straightforward application of a statistical model is a regression that requires near-surface air temperature as the only predictor variable. In this type of model, the assumption is that air temperature is the primary factor influencing surface water temperature. Regression models with air temperature as the only input variable are easily applied to climate change impact studies since GCMs accurately simulate observed near-surface air temperatures (Punzet et al., 2012).

The analysis in this section strictly focuses on the application of regression models to Philadelphia air and water temperature data for the summer and drought-prone months of JJAS. By selecting only JJAS to develop the regression models, the issue of seasonal autocorrelation in the air and water temperature data is reduced. Autocorrelation refers to the temporal patterning often observed in time series data. Specifically, future observations are correlated with past and current observations, resulting in a lack of independence between individual data points.

Based on previous research, both linear and nonlinear regression models have been developed for the relationship between monthly average and weekly average air and water temperatures. Nonlinear regression models are appropriate for air/water temperatures that range from very cold to very warm. This is due to the observation that at high and low air temperatures, the air/water temperature relationship does not remain linear (Punzet et al., 2012). At high air temperatures, generally above 25°C (77°F), water temperature tends to max out as a result of evaporative cooling. When air temperature increases it can hold a greater amount of moisture, which can further increase the rate of evaporative heat loss (Mohseni et al., 1998). At low air temperatures, the decline in water temperature typically slows down and stabilizes at a minimum temperature near 0°C (32°F) (Punzet et al., 2012). The resulting nonlinear regression models have an s-shape. Descriptions of linear and nonlinear regression models for air/water temperature data are provided below for context.

### **Linear Regression Models**

In previous research, basic linear models were used to characterize the relationship between air and water temperatures. Linear models are of the following form:

$$T_w = A + (B * T_a)$$

where

$T_w$  = simulated water temperature (°C)

$T_a$  = observed air temperature (°C)

$A$  = regression line intercept (°C)

$B$  = regression line slope

Both parameters A and B can be determined using least squares regression, a process in which the sum of squared errors between measured and predicted water temperatures is minimized. There are numerous factors that impact the values of A and B, such as regional climate, time scale and range of the data, and hydrologic influences including groundwater inflow and wastewater treatment plant (WWTP) effluent. The slope of the regression line (B) represents the significance of air temperature in predicting water temperature (Erickson and Stefan, 2000). If the value of B is large, it can be assumed that air temperature plays a significant role in predicting water temperature (Erickson and Stefan, 2000). Linear regression models are typically evaluated using the statistical parameter  $r^2$ , or the coefficient of determination.  $R^2$  indicates the proportion of the variation in the response variable (water temperature) that can be attributed to variation in the predictor variable (air temperature) (Crawley, 2011). An  $r^2$  value of 1 represents a perfect fit between the predictor and response variables (Crawley, 2011).

Two separate studies by Erickson and Stefan (1996) and Pilgrim et al. (1998) used linear regression models to characterize the relationship between air and water temperatures from 38 streams in Oklahoma and 39 streams in Minnesota, respectively. The Minnesota study included a period of record ranging from 1956-1991 for the months of April to October. For the Oklahoma study, data was analyzed for the period from March 1-November 30, with water temperature records ranging in length from the years 1969 to 1989. Both studies determined that linear regressions perform fairly well in predicting water temperatures on monthly average and weekly average timescales over the period of analysis. For monthly average temperature data, the average  $r^2$  value was 0.92 for streams in both Oklahoma and Minnesota. For weekly average temperature data, the average  $r^2$  value was 0.83 for streams in Oklahoma and 0.85 for streams in Minnesota (Erickson and Stefan, 2000).

The results from the studies in Oklahoma and Minnesota are further discussed in a paper by Erickson and Stefan (2000), particularly in regard to the influence of other factors, including timescale and time lag, on the correlation between air and water temperatures. The linear regression models for air and water temperatures in Oklahoma and Minnesota perform best when the data are averaged over monthly or weekly timescales, as opposed to daily timescales (Erickson and Stefan, 2000). The technique of averaging reduces the level of variance in the data and weakens the effects of time lag and other shorter-term, or transient, atmospheric impacts on stream temperatures (Erickson and Stefan, 2000). Other

factors that can influence the strength of the air/water temperature relationship include the presence of upstream reservoirs, groundwater inflows, wastewater treatment plant effluent, and the extent of stream shading (Erickson and Stefan, 2000). Regional climate conditions also impact the air/water temperature relationship. The application of linear regression models to air and water temperatures in Minnesota streams demonstrated that linearity was maintained throughout the range of observed air and water temperatures. In Oklahoma, however, the linearity of the relationship decreased once air temperatures reached approximately 25°C due to the effects of evaporative cooling (Erickson and Stefan, 2000). Consequently, Erickson and Stefan (2000) concluded that linear air/water temperature regression models may overestimate stream temperatures at higher air temperatures (above 25°C) and are therefore likely to be more accurate for states located in the northern portion of the country.

### Monthly Average Nonlinear Regression Model

In a study performed by Punzet et al. (2012) a standard nonlinear regression model was used to develop a global relationship between monthly average air and water temperatures. The water temperature data for this study came from a variety of sources, specifically, 935 USGS gauge stations, 570 stations from the United Nations Environmental Program (UNEP) Global Environment Monitoring Systems (GEMS), and 154 additional stream gauges in Europe (Punzet et al., 2012). Air temperature data were pulled from a gridded global meteorological dataset titled the Water and Global Change (WATCH) Forcing Data (Punzet et al., 2012). The full dataset consisted of 97,964 air/water temperature pairs from a total of 1,659 gauging stations with data spanning 36 years, from 1965-2001 (Punzet et al., 2012).

The model used to test the relationship between air and water temperatures is a logistic function that contains three coefficients,  $C_0$ ,  $C_1$  and  $C_2$ . The model's equation is provided below (Punzet et al., 2012).

$$T_w = \frac{C_0}{1 + e^{(C_1 T_a + C_2)}}$$

where

- $T_w$  = simulated monthly average water temperature (°C)
- $T_a$  = observed monthly average air temperature (°C)
- $C_0$  = upper bound water temperature (°C)
- $C_1$  = steepest slope of the function (°C<sup>-1</sup>)
- $C_2$  = measure for inflexion point of the function (°C)

The coefficient  $C_0$  represents the upper bound water temperature (°C) and was found to reach a maximum of 32°C based on prior analyses (Punzet et al., 2012). Therefore only  $C_1$  and  $C_2$  were adjusted to minimize the differences between the observed and simulated water temperature values, a technique referred to as minimizing the sum of squared errors (SSE). As part of this study, coefficients were

developed for a global regression model, in addition to separate regression models for each of the five Köppen-Geiger climate zones. The Köppen-Geiger climate zones, which are classified according to temperature, precipitation and vegetation characteristics, consist of the equatorial, arid, warm temperate, snow, and polar zones (Punzet et al., 2012). Philadelphia lies near the boundary between the warm temperate and snow climate zones. It should be noted that similar regression models account for the impact of heat storage in surface waters, referred to as hysteresis. Hysteresis was neglected in this study when finalizing the global regression model since it was found to have a negligible effect. For the purpose of simplicity, the impacts of heat storage also are not included in the regression model analyses for Philadelphia.

The study by Punzet et al. (2012) found that the global regression model performed well in many parts of the world, including the eastern United States. Through an analysis of the global dataset, it was determined that 77% of the water temperature simulations ( $n=97,964$ ) were within  $3^{\circ}\text{C}$  of the observed water temperatures (Punzet et al., 2012). In regard to the regression models for each of the Köppen-Geiger climate zones, the warm temperate, snow and arid zones, which represent a majority of the overall dataset, showed good agreement with the global model (Punzet et al., 2012).

In conclusion, this study showed it is possible to develop a global nonlinear regression model that fairly accurately (within  $3^{\circ}\text{C}$ ) simulates water temperature by using air temperature as the only predictor variable. This model can be applied to future climate scenarios to gain an understanding of potential changes in surface water temperatures.

### **Weekly Average Nonlinear Regression Model**

Multiple studies have assessed the relationship between air and water temperatures on a weekly timescale, which is the frequency period typically used to assess fish habitats (Mohseni et al., 1998). The studies referenced in this section are from Mohseni et al. (1998) and Mohseni et al. (1999). In both cases, a four-parameter, nonlinear regression model was used to estimate weekly average water temperatures given weekly average air temperatures on a multitude of streams across the contiguous United States. The logistic regression model was first developed in Mohseni et al. (1998) and then applied to a wider range of monitoring sites in Mohseni et al. (1999). This section will focus on the work in Mohseni et al. (1999), in which a total of 993 USGS stream gauging stations with data records averaging a span of 12 years were used (Mohseni et al., 1999). Air temperature data was obtained from one of 166 weather stations within the Solar and Meteorological Surface Observation Network supplied by the National Oceanic and Atmospheric Administration (NOAA) and the National Renewable Energy Laboratory (NREL) (Mohseni et al., 1999). The weather stations were selected based on proximity to the USGS gauging stations.

Surprisingly, the distance between air and water temperature stations in the Mohseni et al. (1998) study, which ranged from 1.4km to 244km, was determined to have an insignificant impact on the model's goodness of fit.

During development of the nonlinear regression model for weekly average data in Mohseni et al. (1998), several mathematical functions were tested for their efficacy in predicting water temperature from air temperature. A four-parameter logistic function, displayed below, was determined to provide the best representation of the s-shaped air/water temperature relationship.

$$T_w = \mu + \frac{\alpha - \mu}{1 + e^{\gamma(\beta - T_a)}}$$

where

$T_a$  = observed weekly average air temperature, °C

$T_w$  = simulated weekly average water temperature, °C

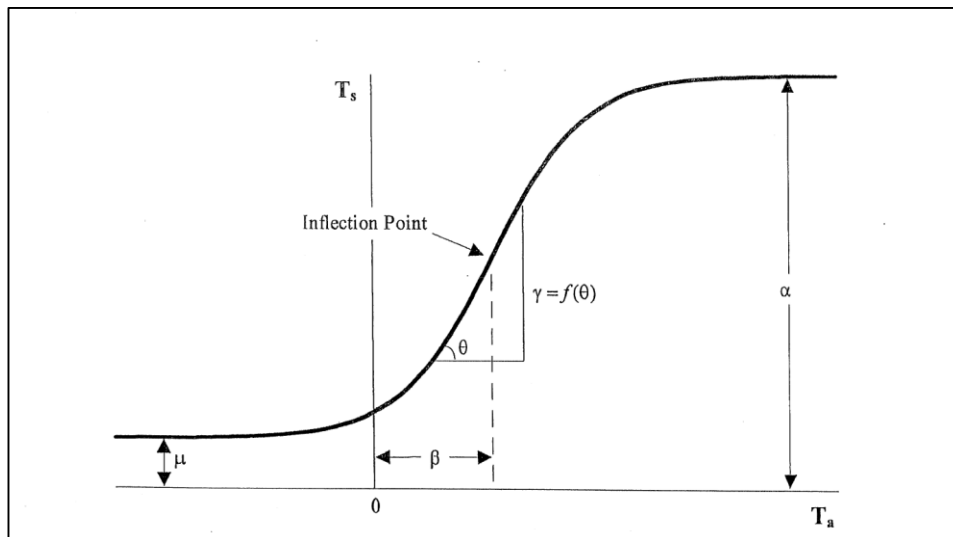
$\mu$  = estimated minimum water temperature, °C

$\alpha$  = estimated maximum water temperature, °C

$\gamma$  = measure of the steepest slope of the function

$\beta$  = air temperature at the inflection point of the function (Mohseni et al., 1998)

Figure 6 below provides a graphical representation of the four parameters used in the nonlinear regression model for weekly average air and water temperatures.



**Figure 6.** Graphical representation of the parameters used in the weekly average air/water temperature nonlinear regression model

Source: Mohseni et al., 1998

In estimating the four parameters, alpha ( $\alpha$ ), beta ( $\beta$ ), gamma ( $\gamma$ ) and mu ( $\mu$ ), the objective was to minimize the SSE between observed and simulated water temperatures (Mohseni et al., 1998). Alpha and

beta represent the maximum and minimum estimated stream temperatures, respectively, and can be approximated directly from the dataset (Mohseni et al., 1998). Beta and gamma must be approximated using a process to minimize the SSE (Mohseni et al., 1998). The Mohseni et al. (1999) study determined values for beta, gamma and mu using least squares regression analysis, while alpha, the estimated maximum stream temperature, was approximated directly from the data (Mohseni et al., 1999).

The study referenced here developed a set of regression parameters for a total of 989 streams (4 of the original streams did not produce beta and gamma values when fit to the model) (Mohseni et al., 1999). It should be noted that hysteresis, or heat storage effects, were accounted for in this study. A majority of the streams included in this analysis exhibited a good fit using the 4-parameter nonlinear regression model. The study cites that 91% of the 989 streams had predicted water temperatures that produced a Nash-Sutcliffe Coefficient (NSC) of 0.80 or greater (Mohseni et al., 1999). The mean NSC for all streams was 0.89 with a standard deviation of 0.10 (Mohseni et al., 1999). The NSC is defined as follows:

$$NSC = 1 - \frac{\sum_{t=1}^n (T_{sim_t} - T_{obs_t})^2}{\sum_{t=1}^n (\bar{T}_{obs} - T_{obs_t})^2}$$

where

- $T_{sim_t}$  = simulated water temperature at time t
- $T_{obs_t}$  = observed water temperature at time t
- $\bar{T}_{obs}$  = mean observed water temperature (Mohseni et al., 1998)

The NSC is characterized by a maximum value of 1 and no minimum value. A model with a perfect fit will have a NSC equal to 1 (Mohseni et al., 1998). Punzet et al. (2012) found that the global nonlinear regression model had an NSC greater than 0.8 in numerous parts of the world, including the eastern portion of the United States.

In addition to considering the efficacy of the 4-parameter nonlinear regression model, this study also assessed the correlation between the regression parameters at gauging stations with NSCs greater than 0.90 (n=803) and other factors, including mean annual and seasonal air temperatures. The study found only a weak relationship between the four parameters and mean annual and seasonal air temperatures, indicating that under warmer climate conditions and an unchanged physical environment, the parameter values likely will not change (Mohseni et al., 1999). Additionally, no correlation existed between the parameters and latitude or the parameters and upstream watershed drainage area (Mohseni et al., 1999). It is also interesting to note that in Mohseni et al. (1998), a set of streams with NCSs less than 0.7 and an

air/water temperature relationship that did not follow an s-shape were found to be located 1km-30km downstream of a reservoir.

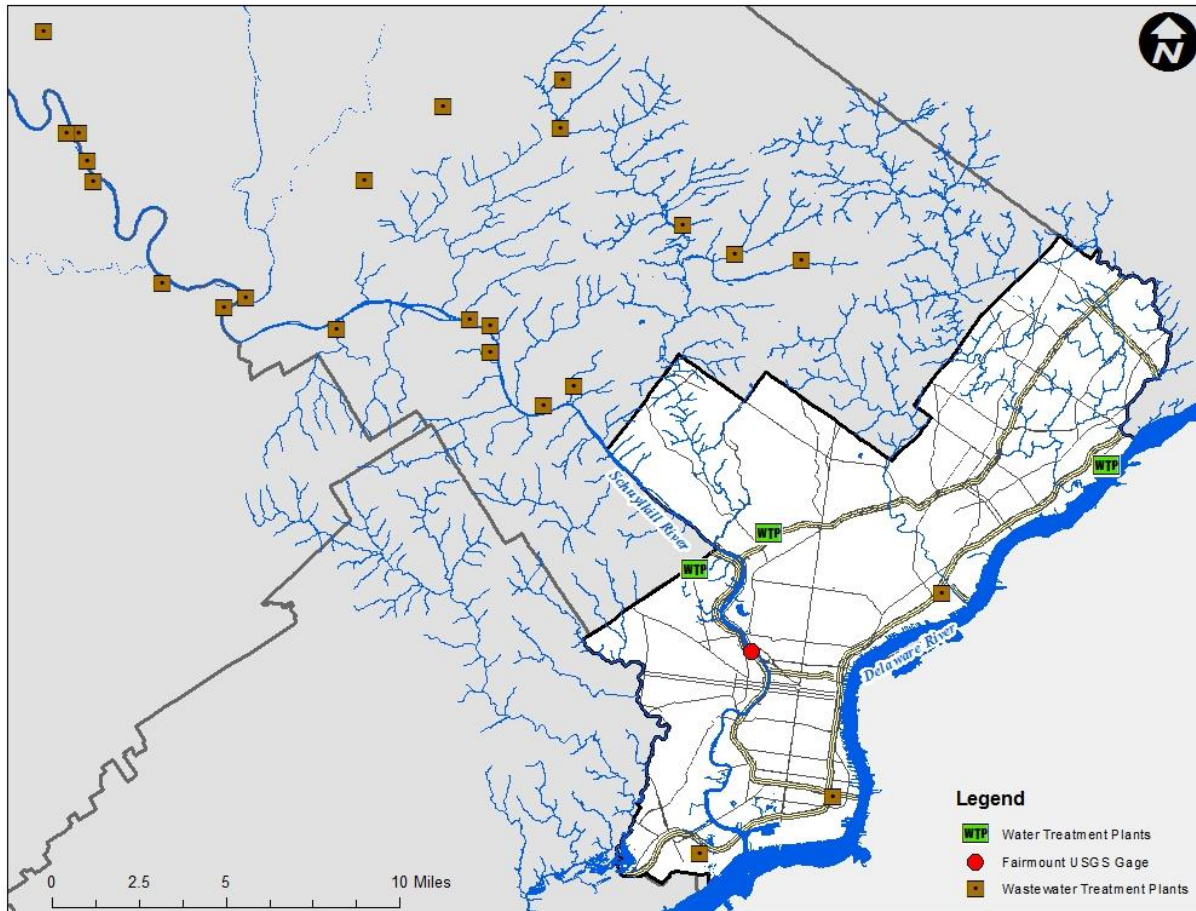
This study showed that a 4-parameter logistic regression model could accurately represent the air/water temperature relationship on a weekly timescale for the majority of streams included in the analysis. The application of monthly average and weekly average regression models to the Philadelphia data is described in the Analysis and Results section.

## **Data**

The water and air temperature data used in this analysis were from two sources, respectively: the United States Geological Survey (USGS) and Daymet, a database archived by the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (Thornton et al., 2014). Additional information regarding the water and air temperature datasets is provided below.

### **Water Temperature Data**

Water temperature data were obtained from USGS gauge 01474500 located on the Schuylkill River upstream of the Fairmount Dam (USGS, 2014). The USGS Fairmount gauge is located downstream of and in close proximity to the Philadelphia Water Department's (PWD's) Queen Lane Drinking Water Treatment Plant (DWTP) intake and Belmont DWTP intake. Both intakes and the USGS Fairmount gauge are situated at the downstream-most end of the Schuylkill River watershed, which encompasses a 1,893mi<sup>2</sup> drainage area (USGS, 2014). Since the USGS Fairmount gauge is the closest water monitoring station to PWD's Schuylkill River drinking water intakes, it provides the best representation of Philadelphia's raw water quality. The map below illustrates the location of the USGS Fairmount gauge relative to PWD's drinking water treatment plants. The map also indicates the location of the numerous wastewater treatment plants (WWTPs) upstream of the USGS Fairmount gauge in the Schuylkill River watershed.



**Figure 7.** Location of the USGS Fairmount gauge relative to PWD’s drinking water treatment plants  
*Source: Philadelphia Water Department, 2014*

The monthly and weekly average water temperature data from USGS that were used in this analysis comprise data for the months of June, July, August and September (JJAS) from 1999-2001 and 2011-2013. The two datasets were selected based on data availability and the overlap between air and water temperature observations. The results from basic linear regressions, which are described in more detail below, were used to determine whether it was necessary to treat the two inconsecutive datasets separately. A basic linear regression between observed air and water temperatures for both datasets revealed a slightly lower goodness of fit for 1999-2001 as compared to the more recent data from 2011-2013. The minimal difference in the air/water temperature relationship warranted combining the two time periods to produce a single, more robust dataset.

In order to obtain monthly and weekly average datasets, daily water temperature data were pulled from the “Daily Data” for USGS gauge 01474500 for the time period 1999-2001 (USGS, 2014). The daily averages were the used to compute monthly and weekly averages for analysis. The month of September

1999 contained no daily averages and was removed from the analysis. USGS designated all daily averages as “A”, which indicates the data processing and review are complete and the data are approved for publication.

A second dataset of JJAS water temperature values was pulled for the years 2011-2013 at the USGS Fairmount gauge. Since no daily data was available for this time period, the “Current/Historical Observations” (observed) data category was referenced (USGS, 2014). The observed dataset includes instantaneous temperature values collected every 30 seconds at the gauge. The daily means were first calculated from the 30-second observations, and then the monthly and weekly averages were calculated from the daily means. The month of September 2011 did not contain adequate data and was removed from the analysis. USGS designated all of this data as “P”, which indicates the data are provisional and subject to revision (USGS, 2014).

### **Air Temperature Data**

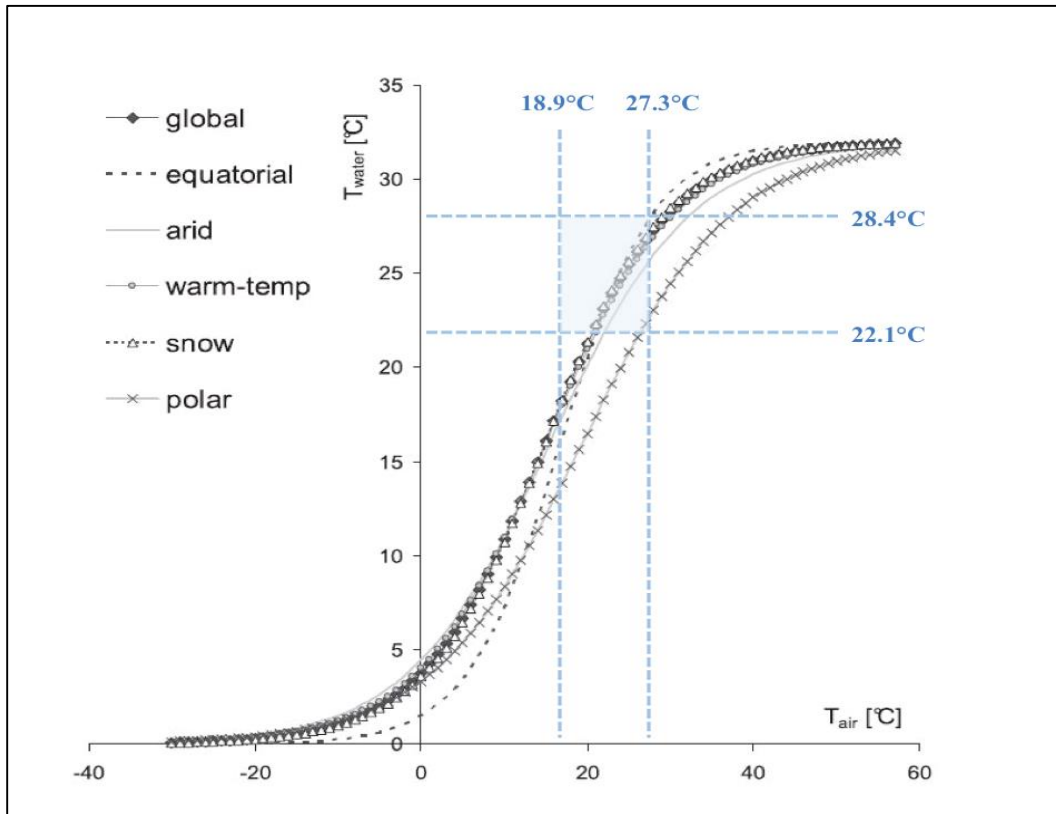
The air temperature data was extracted from the Daymet website, which contains an archived database distributed through the Oak Ridge National Laboratory Distributed Active Archive Center (ORNL DAAC) (Thornton et al., 2014). Daymet data are in continuous gridded format for North America at a resolution of 1km x 1km (Thornton et al., 2014). The gridded data include maximum and minimum near-surface (2 meter) air temperatures (Thornton et al., 2014). The Daymet air temperature data for the United States are primarily sourced from the Cooperative Summary of the Day weather station network available through the National Climate Data Center (NCDC).

Data were extracted from the database through the process of single pixel extraction, which allowed for the specification of the USGS Fairmount gauge latitude/longitude coordinates when obtaining air temperature measurements. The query returned maximum and minimum daily temperatures, which were averaged to create a dataset of daily averages for use in calculating monthly and weekly averages. Daymet data exists from January 1, 1980 through the last full calendar year, so there was no issue with data availability.

### **Methods**

After formatting the air and water temperature data to produce monthly average and weekly average datasets, two separate linear regression models were developed. Linear regression models were chosen for this analysis based on the range of observed air and water temperatures in Philadelphia during the months of JJAS. As illustrated in Figure 8, the minimum and maximum monthly average air and water

temperatures encompass a range of temperatures that are fairly linearly related given their location on the s-shaped global nonlinear regression curve.



**Figure 8.** Range of observed monthly average air and water temperatures in Philadelphia, PA relative to the global nonlinear regression model developed by Punzet et al. (2012)

The linear regression equations were developed in Excel through the least squares method. Statistical parameters including  $r^2$ , root mean square error (RMSE) and sum of squared errors (SSE) were obtained for both the monthly average and weekly average air and water temperature regression models.

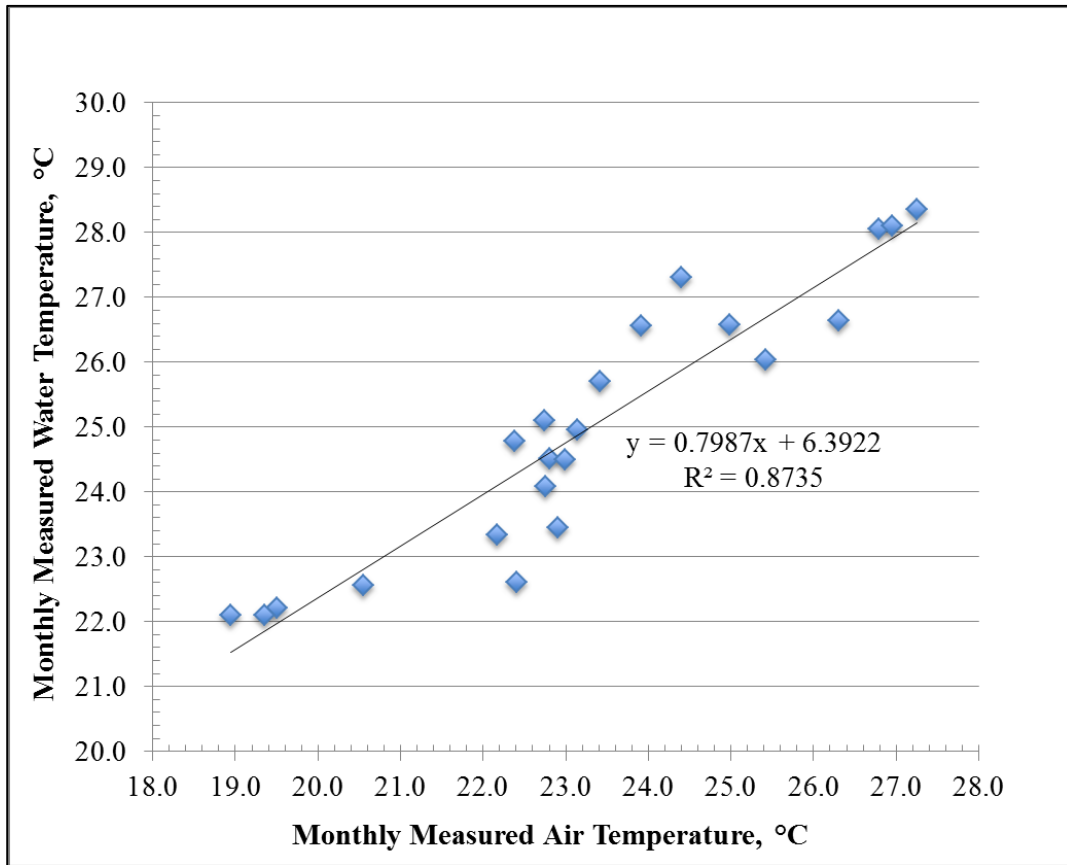
## Analysis & Results

This section outlines the results for the regression analyses from Part 2 of this project, first for the Philadelphia monthly average data followed by the weekly average data.

### Monthly Average Data

A simple linear regression equation was developed for the entire dataset, which consisted of monthly averages for June, July August and September (JJAS) from 1999-2001 and 2011-2013, with the exception of September 1999 and September 2011. In statistical modeling, regression models are used to assess the relationship between two or more continuous variables. A linear regression model was tested based on the

range of monthly average air and water temperature values that were observed during JJAS in Philadelphia, as described above. A graphical depiction of the linear relationship is provided in Figure 9.



**Figure 9.** Monthly average air/water temperature relationship, Philadelphia, PA. June, July, August and September, 1999-2001 and 2011-2013

As depicted in Figure 9, the simple linear regression model indicates that water temperature is fairly accurately estimated from air temperature using the following linear model:

$$T_w = 0.7987 * T_a + 6.3922$$

where

$T_w$  = simulated monthly average water temperature (°C)

$T_a$  = observed monthly average air temperature (°C)

The coefficient of determination,  $r^2$ , for this linear model is approximately 0.87 ( $n=22$ ). The root mean square error (RMSE), which is another indicator of goodness of fit, is 0.71°C for the linear model. The equation used to calculate RMSE in this analysis is included below.

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (T_{sim_t} - T_{obs_t})^2}{n}}$$

where

$T_{sim_t}$  = simulated water temperature at time t

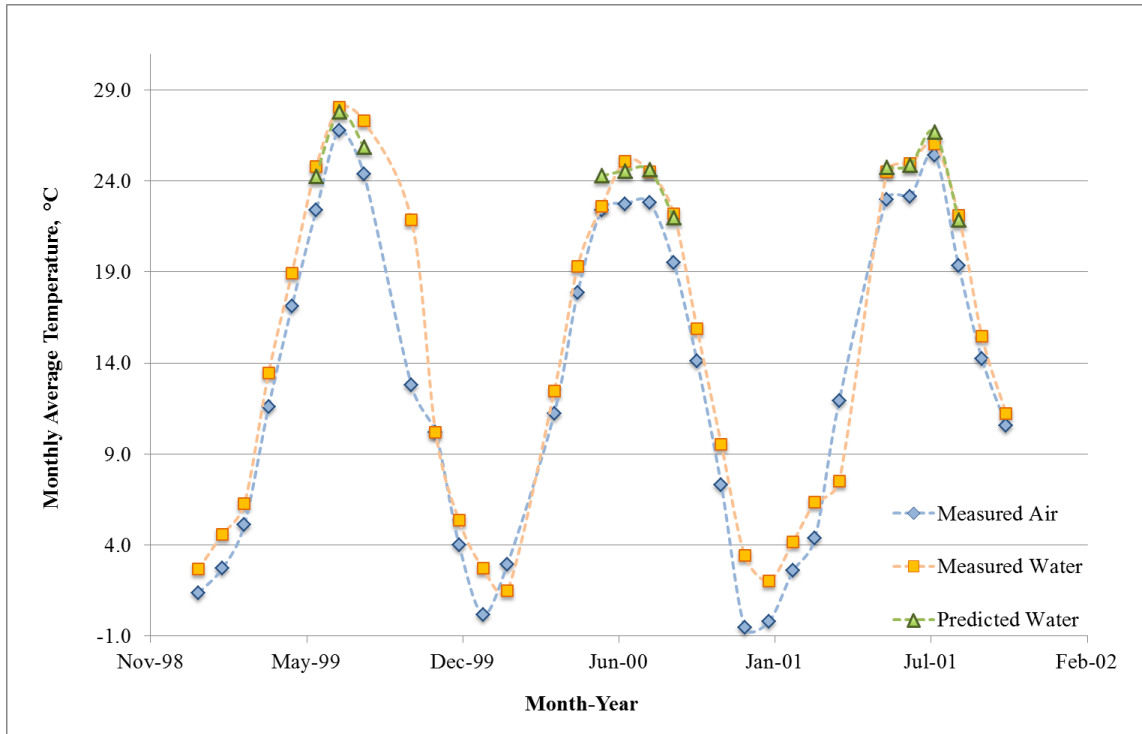
$T_{obs_t}$  = observed water temperature at time t

$n$  = number of water temperature observations

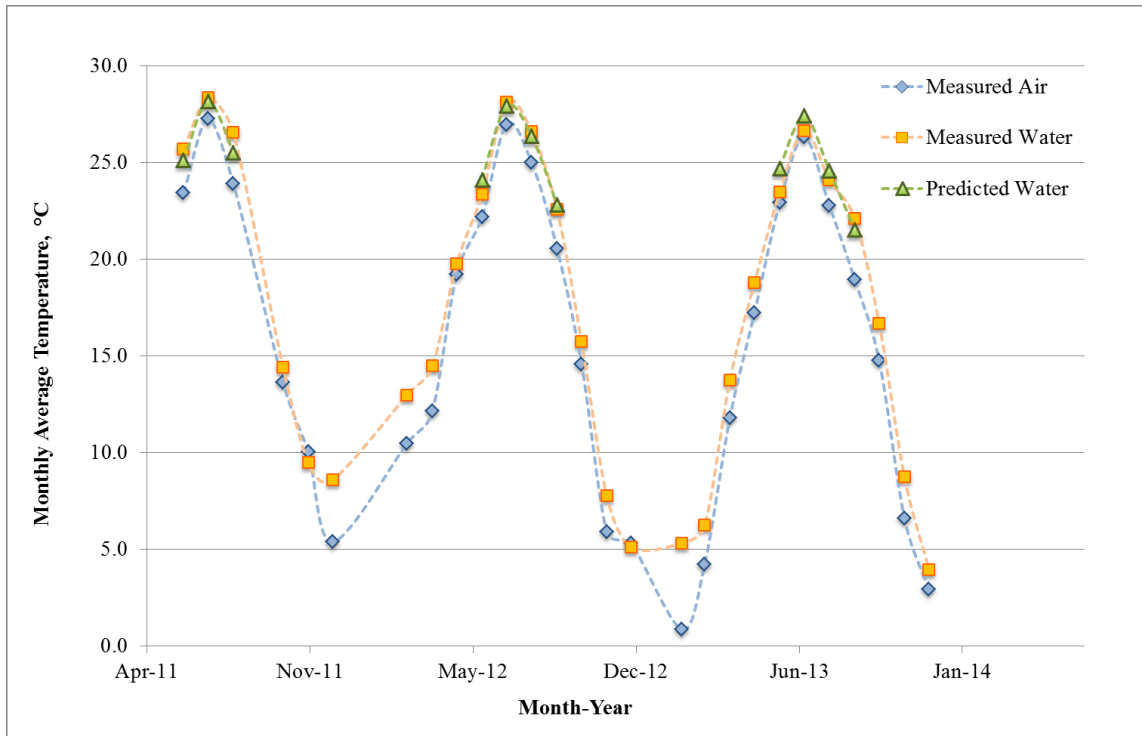
The Nash-Sutcliffe coefficient (NSC) and RMSE are commonly used to determine a model's quality of fit, particularly in studies involving water temperature (Punzet et al., 2012). For reference, the NSC for the monthly average linear regression model was equivalent to the  $r^2$  value of 0.87. Using RMSE as one indicator of model fit implies that the best model will produce the smallest difference between simulated and observed water temperatures (Punzet et al., 2012). Additionally, the sum of squared errors (SSE), representing the sum of squared differences between the measured and simulated water temperatures, was found to be approximately 11.06 for this model.

A more extensive analysis would consider potential contributing factors that impact the  $r^2$  value, including low flow conditions, wastewater treatment plant (WWTP) effluent, other point source discharges and reservoir releases. One or a combination of these factors could contribute to a deterioration of the air/water temperature relationship in the Schuylkill River near Philadelphia. Additionally, a more comprehensive dataset of monthly average air/water temperatures in Philadelphia would aid in understanding the relationship between the two variables.

To further illustrate the relationship between observed air and water temperatures on the Schuylkill River at Philadelphia, and the performance of the linear regression model in predicating JJAS Schuylkill River water temperatures, two times series graphs displayed in Figures 10 and 11 below were produced for each 3-year dataset.



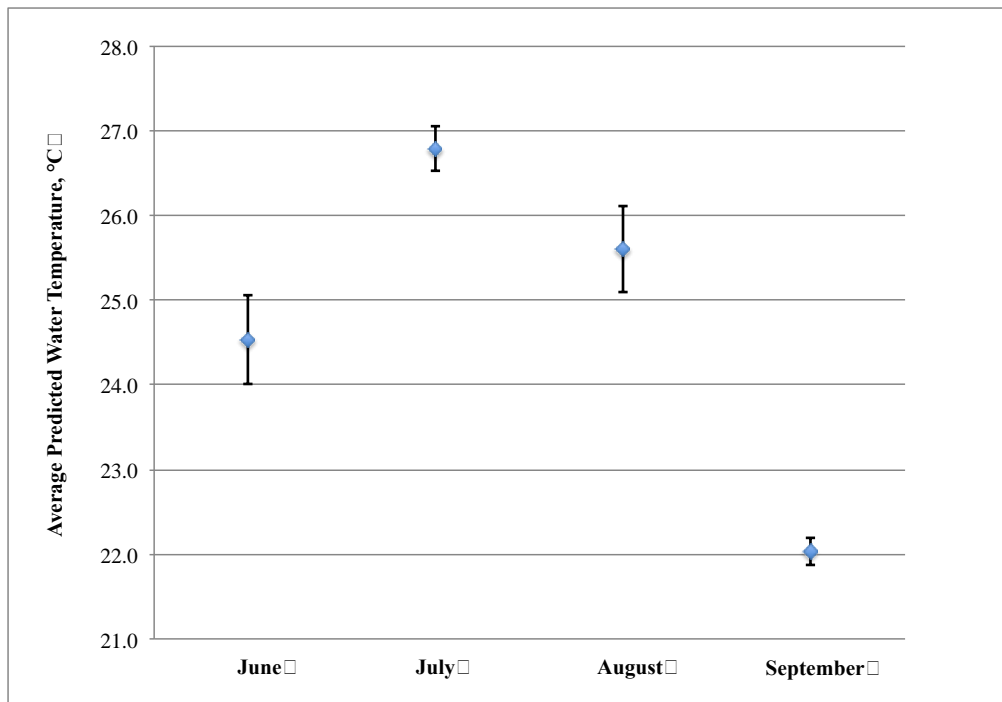
**Figure 10.** Air and water temperature time series data and predicted values for June, July, August and September, Philadelphia, PA. January 1999-November 2001



**Figure 11.** Air and water temperature time series data and predicted values for June, July, August and September, Philadelphia, PA. June 2011-December 2013

Both Figures 10 and 11 illustrate a sinusoidal shape created by seasonal patterns in the observed air and water temperature time series data. The plotted predicted values indicate that the linear regression model performs fairly well in simulating water temperature over the months of June, July, August and September. It should be noted that during the winter months, the measured air and water temperatures appear to be less well correlated than during the warmer months. Future research could focus on trying to understand the deteriorating air/water temperature relationship during colder months on the Schuylkill River at Philadelphia.

To further assess the performance of the linear regression model on a month-by-month basis, the standard deviation associated with the absolute error for each pair of measured and predicted water temperatures was plotted in Figure 12 below. The absolute error is defined in this instance as the absolute value of the difference between each observed water temperature value and the corresponding predicted value from the linear regression model.



**Figure 12.** Standard deviations associated with the absolute error (predicted-observed) of available average predicted water temperatures by month, Philadelphia, PA. June, July, August, September, 1999-2001 and 2011-2013

The standard deviations of the absolute errors for each month indicate that the linear regression model performs fairly well during the months of JJAS in Philadelphia over the period of analysis. For the dataset

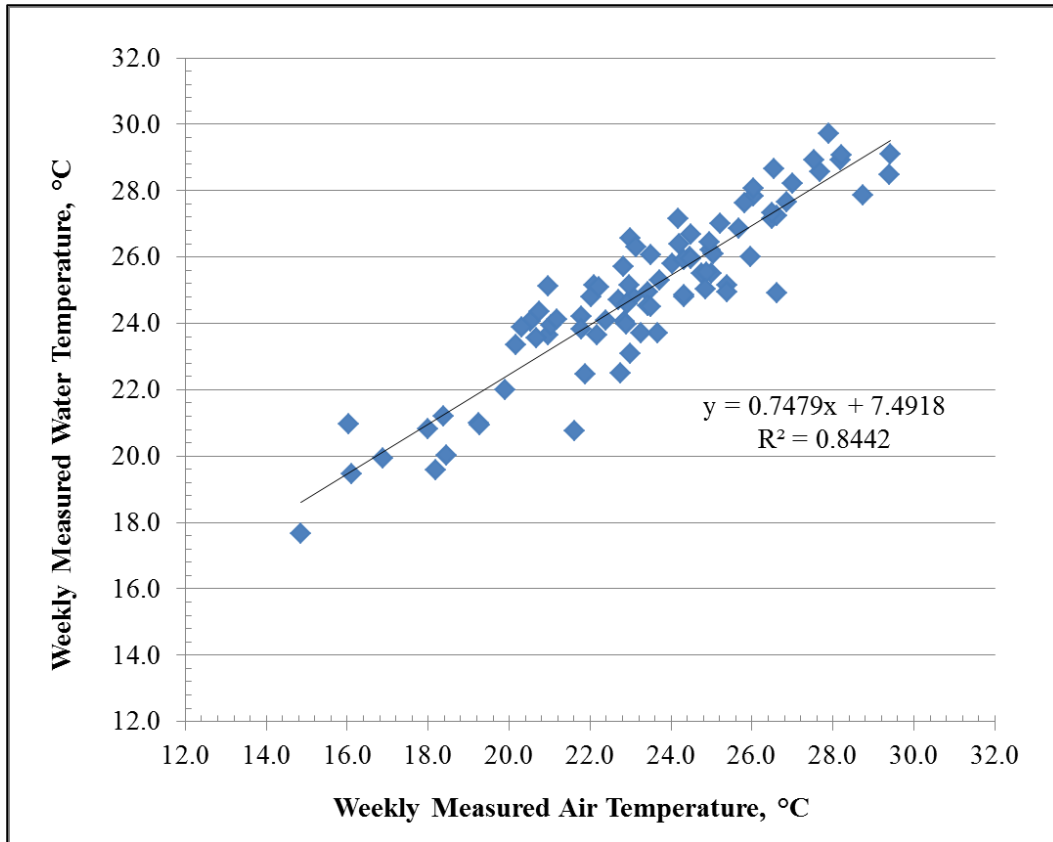
used, the months of June and August have slightly higher standard deviations associated with the absolute errors as compared to the months of July and September.

When developing a linear model, there are several assumptions that must be checked in order to assess model validity. The primary assumptions associated with a linear regression model relate to the distribution of the errors and their variances. Ideally, the errors should be normally distributed and the variance terms should be constant along the regression line, a property also referred to as homoscedasticity (Crawley, 2011). The statistical software R was used to develop a series of four diagnostic plots to test these linear model assumptions. The diagnostic plots and an analysis of the results are included in Appendix D, Figure 1D. In general, the model was found to satisfy the assumptions of a linear regression.

Based on the analysis outlined above, it can generally be assumed that the linear regression model performs fairly well for predicting monthly average water temperatures during JJAS. Specifically, approximately 87% of the variation in monthly average water temperature can be explained by monthly average air temperature during JJAS on the Schuylkill River at Philadelphia.

### **Weekly Average Data**

The same method as outlined above was used to analyze the weekly average data over the months of June, July, August and September for the years 1999-2001 and 2011-2013. A graphical depiction of the linear relationship between the weekly average temperature data is provided in Figure 13.



**Figure 13.** Weekly average air/water temperature relationship, Philadelphia, PA. June, July August, September, 1999-2001 and 2011-2013

The simple linear regression model indicates that water temperature is fairly accurately estimated from air temperature using the following equation:

$$T_w = 0.7479 * T_a + 7.4918$$

where

$T_w$  = simulated weekly average water temperature (°C)

$T_a$  = observed weekly average air temperature (°C)

The  $r^2$  value for the linear model using weekly averages is approximately 0.84. The root mean square error (RMSE) is approximately 0.48°C, and the SSE was found to equal 83.05 (n=86). The NSC for the weekly average linear model is equivalent to the  $r^2$  value of 0.84.

As with the regression model used for monthly average temperatures, the linear model using the weekly average data makes several assumptions that must be checked. The statistical software R was used to develop a series of four diagnostic plots that are included in Appendix D, Figure 2D. The regression model for weekly average temperature data generally meets the assumptions for a linear model.

Research suggests that the relationship between air and surface water temperatures deteriorates as the averaging period decreases. It is then implied that a linear regression model developed for weekly average data will perform less well than a monthly average regression model. In this case, the models performed very similarly, with the regression model for monthly average data having a slightly higher  $r^2$ . The linear regression model for weekly average temperatures indicates that approximately 84% of the variation in JJAS stream temperatures is due to air temperature, as compared with approximately 87% for the monthly average regression model.

### ***Summary of Linear Regression Models***

A summary of the statistical parameters associated with each of the air/water temperature linear models is included in Table 5 below.

**Table 5.** Properties of the linear regression models for weekly average and monthly average air and water temperatures in Philadelphia, PA. June, July, August, September, 1999-2001 and 2011-2013

<b>Timescale</b>	<b>Intercept (°C)</b>	<b>Slope</b>	<b>SSE</b>	<b>RMSE (°C)</b>	<b>R<sup>2</sup></b>	<b>n</b>
Weekly Average	7.49	0.75	83.05	0.48	0.84	86
Monthly Average	6.39	0.80	11.06	0.71	0.87	22

Based on the results from the analysis using both monthly average and weekly average temperatures, it can be concluded that a basic linear regression model fairly accurately predicts water temperatures on the Schuylkill River near Philadelphia over the time period of analysis. The results for Philadelphia are comparable to those found in literature. As noted previously, streams in Oklahoma and Minnesota exhibited an average  $r^2$  of 0.83 and 0.85 for weekly average temperatures, respectively. Using a monthly average timescale resulted in an  $r^2$  of 0.92 for both sets of streams in Oklahoma and Minnesota. The  $r^2$  value for monthly average temperatures in the Schuylkill River at Philadelphia is not as high as those observed in the Oklahoma and Minnesota streams. Extensive research would be needed to evaluate other factors that are influencing the Schuylkill River's air/water temperature relationship. Although literature notes that the relationship between air and water temperature usually becomes nonlinear at temperatures above 25°C, the Schuylkill River at Philadelphia did not exhibit nonlinearity within the range of observed JJAS air and water temperatures.

## IV. Summer Average Surface Water Temperature Projections

The ability to predict stream temperature using air temperature is critical when considering the potential impacts of climate change on surface water supplies. The implications for increasing surface water temperatures are wide-ranging, and include impacts to drinking source water quality as well as the drinking water treatment process.

The goal of the projections part of this analysis was to gain a general understanding of the potential for surface water temperature increases in Philadelphia given a higher average air temperature for the months of June, July, August and September (JJAS). The linear regression model for monthly average data was applied to projected increases in air temperature for JJAS over the two available time periods: 1999-2001 and 2011-2013. The monthly average temperatures were averaged over JJAS to maintain consistency with the calculations performed in Matlab when evaluating the climate model output.

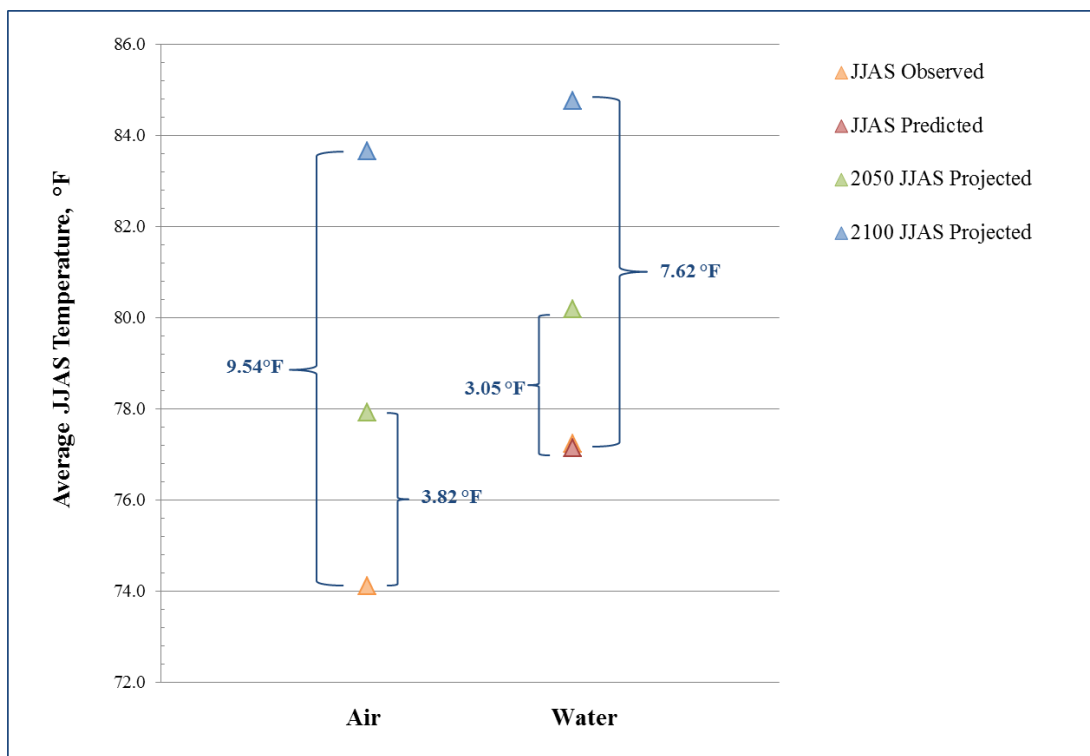
Part 1 of this project utilized the RCP8.5 climate scenario to approximate air temperature increases in the Northeast region and Philadelphia by mid-century (2050) and the end of the century (2100). By using climate model output from three different GCMs, it was determined that the Philadelphia area can expect 2.12°C (3.82°F) of warming for JJAS average air temperatures by 2050, and approximately 5.30°C (9.54°F) of warming for JJAS averages by 2100. To calculate projected changes in water temperature, the observed air temperatures for the months of JJAS were averaged together for each year over the period of analysis. The projected changes in air temperature were then added to the JJAS average over all years, which were then plugged into the linear regression model equation in order to determine projected changes in JJAS average water temperatures. The results are summarized in Table 6 below.

**Table 6.** Projected increases in average JJAS air and water temperatures in Philadelphia

AIR		WATER	
2050	2100	2050	2100
+2.12°C (+3.82°F)	+5.30°C (+9.54°F)	+1.69°C (+3.05°F)	4.23°C (+7.62°F)

The projected increase in JJAS surface water temperatures by mid-century and end-of-century represent approximately 80% of the projected increase in JJAS air temperature. These results are higher than those found in literature. Previous research has determined that increases in surface water temperature are typically in the range of 50-70% of total air temperature increases (WHO, 2008). The difference in the results presented here could have to do with the use of a linear model in predicting water temperature from air temperature for future timeframes. The projected water temperature increases are potentially

overestimated since it is unknown if the linear or nonlinear approach is better at catching the impacts of evaporative cooling and other influencing factors at future temperature ranges and this location. As air temperature increases, the water temperature in the Schuylkill River may begin to reach a critical threshold where the temperature begins to level off, slowing its rate of increase relative to air temperature increases, as presented in the nonlinear regression functions from Punzet et al. (2012) and Mohseni et al. (1999). From a physical standpoint, the difference in the rate of air and water temperature increases can be partially attributed to the concept of heat capacity, in which it takes more heat energy to change the temperature of water as opposed to air (Blauch, 2014). Figure 14 illustrates the projected increases in air and water temperatures on a degrees Fahrenheit scale.



**Figure 14.** Observed and projected increases in average JJAS air and water temperatures in the Schuylkill River at Philadelphia

The implications of an increase in surface water temperature in the range of 1.69-4.23°C (3.05°F-7.62°F) should be examined from the perspectives of raw water supply quality, treatment and distribution. While the temperature increases will occur over long time scales (50-100 years), it is useful for PWD to be aware of potential temperature changes in their water supply when assessing long-term planning strategies. The following section will provide a brief overview of one of the potential impacts of surface water temperature increases for drinking water suppliers.

## **V. Increasing Surface Water Temperatures and Disinfection Byproduct (DBP) Formation**

Water utilities extensively monitor water quality at the source, throughout the treatment process, and during distribution before the water finally reaches a customer's tap. The implications of increasing source water temperatures will directly impact water quality at DWTP intakes, otherwise known as raw water quality. For a drinking water utility, the impact of climate change on raw water quality needs to be considered in the context of the treatment and distribution systems, creating complex operational situations that likely vary on a utility-by-utility basis. The formation of disinfection byproducts (DBPs) is one example of a direct link between raw water quality characteristics and resulting challenges to drinking water treatment. The following section describing Part 3 of this project provides an overview of the connection between water temperature and DBP formation and the subsequent operating and management challenges faced by drinking water utilities.

### **Background on Disinfection Byproducts (DBPs)**

The water industry is currently aware of several types of disinfection byproducts (DBPs), some of which are already regulated and others of which are considered emerging contaminants (McGuire et al., 2014). The first group of DBPs to be regulated was trihalomethanes (THMs). THMs were discovered in the 1970s and regulated through promulgation of the Total Trihalomethane (TTHM) Rule in 1979, which applied to drinking water providers with more than 10,000 customers (McGuire et al., 2014). Depending on water quality characteristics and the disinfection mechanism(s) that are applied during the treatment process, different types and amounts of DBPs have the potential to form. In very general terms, DBPs are created when a disinfecting agent, such as chlorine, reacts with the natural organic matter (NOM) present in water (Rodriguez & Sérodes, 2001). From a public health perspective, DBPs have been shown in laboratory studies to be carcinogenic, with the potential to cause adverse reproductive or developmental effects (US EPA, 2011). Since DBPs are formed during disinfection, near the end of the overall treatment process, traditional treatment methods provide little, if any, additional barriers before water is distributed to customers. The health implications for humans exposed to low DBP concentrations over longer time periods warrant further study, but nonetheless, DBPs are an area of concern for drinking water suppliers.

Since the 1970s, national regulations surrounding DBPs have gone through multiple iterations and revisions and are now included in a set of regulations known as the Stage 2 Disinfectants and DBP (D/DBPR) and Long-Term 2 Enhanced Surface Water Treatment Rule (LT2 ESWTR) (McGuire et al., 2014). This set of regulations strives to balance the opposing goals of limiting DBP formation during the disinfection process with that of providing adequate disinfection to deactivate microbial contaminants (McGuire et al., 2014). Compliance with these regulations has proven challenging and expensive for

water suppliers (McGuire et al., 2014). As raw water quality conditions change, the challenges associated with balancing the opposing risks of microbial contamination and DBP formation may increase.

### **DBP Formation Potential**

Several water quality characteristics are known to impact the formation of DBPs, including dissolved organic carbon (DOC), pH and bromide (Teksoy, Alkan, & Başkaya, 2008). DBP formation is also dependent on operational considerations, particularly the type of disinfection process utilized and the related doses and contact times (McGuire et al., 2014). Chlorine still presents a comparatively inexpensive and effective means of controlling microbial contamination during treatment and also within the distribution system, but other disinfection products are also used in the drinking water sector, including chloramines, ozone, and coagulation (McGuire et al., 2014). While switching from chlorine to a new type of disinfection method may prevent the formation of chlorinated DBPs, there is the potential for other DBPs to form given the complex interactions between water quality characteristics and disinfection methods. Consequently, an alternative approach to limiting DBP formation is to reduce known precursors, such as organic matter, prior to the disinfection process (Teksoy, Alkan, & Başkaya, 2008).

Water temperature, which can vary over spatial and temporal scales within the source, treatment and distribution systems, has been shown to have a significant impact on DBP formation. In a study by Rodriguez and Sérodes (2001), water temperature was found to be the dominant parameter affecting THM levels in the distribution systems of three utilities in Quebec, Canada (Rodriguez & Sérodes, 2001). Additionally, the study found that water temperatures above 15°C increase the spatial variation in THM concentrations in finished water (Rodriguez & Sérodes, 2001). As noted in Rodriguez & Sérodes (2001), the water temperature-DBP formation relationship is further complicated when high water temperatures speed up the process of chlorine decay, prompting DWTP operators to elevate chlorine doses, in turn elevating the potential for DBP formation (Rodriguez & Sérodes, 2001).

Water resource managers are now faced with regulatory implications if they are unable to balance the risks associated with microbial contamination and DBP formation. Certain management strategies are available from an operational perspective, including strategies to reduce DBP formation precursors in water prior to disinfection or changing the disinfection method based on water quality characteristics (McGuire et al., 2014; Teksoy, Alkan, & Başkaya, 2008). Even with DBP management strategies in place, it is evident that the interaction between water temperature and DBP formation creates an additional layer of complexity that can vary on spatial and temporal scales. It is important for water

resource managers to be aware of the potential for warmer water temperature to increase DBP formation rates when developing monitoring and operations strategies.

## VI. Discussion and Conclusions

In the field of climate modeling, there is consensus that air temperature is one of the most accurately modeled climate variables (Mohseni & Stefan, 1999). The results from Part 1 of this project provide evidence that GCMs accurately simulate average near-surface air temperatures in the Northeast region and Philadelphia. Assuming that current CO<sub>2</sub> emissions continue on their current trend, the RCP8.5 climate scenario provides estimates for near-surface air temperatures that can be expected by the middle and end of the 21<sup>st</sup> century. Assuming greenhouse gas emissions maintain a track in line with the RCP8.5 scenario, the GCM output used in this analysis suggests an increase in Philadelphia's June, July, August, and September (JJAS) average temperature of approximately 2.12°C by 2050 and approximately 5.30°C by 2100.

Previous research already demonstrates that air temperature increases can be strongly correlated with stream temperature increases. Factors that can deteriorate the relationship between air and water temperature in a surface water body include WWTP effluent, reservoir releases and groundwater inflow (Morrill et al., 2005; Webb et al., 2008). For the Schuylkill River at Philadelphia, it was found that air temperature explains the majority of variation in water temperature when using a linear regression model for the months of June, July, August and September from 1999-2001 and 2011-2013. The linear regression model for weekly average temperatures indicates that approximately 84% of the variation in JJAS stream temperatures is due to air temperature, as compared with approximately 87% for the monthly average regression model.

The linear regression model for monthly average temperatures in Philadelphia was used to produce stream water temperature projections based on the aforementioned projected increases in air temperature by 2050 and 2100. The results from this analysis indicate that the Schuylkill River water temperature at Philadelphia may increase by approximately 1.69°C (3.05°F) by 2050 and 4.23°C (7.62°F) by 2100. It is important to note that the projections in this analysis assume temporal stability of the air/water temperature relationship through the end of the century.

One potential implication of warming surface water temperatures could be increased formation of DBPs during the treatment process, a scenario that could have regulatory implications for utilities. The complexity of DBP formation makes it challenging to attribute increased formation rates to a single water quality parameter, but there is significant evidence that warmer water temperature is associated with higher formation rates. Aside from DBP formation potential, there are numerous other impacts to surface water quality that drinking water utilities will need to consider as temperatures increase. These impacts

include changes to surface water chemistry, including decreased dissolved oxygen (DO) concentrations, which could impact the ability of a water body to meet pre-designated water quality standards, and increasing algal blooms which can impact the taste and odor of treated drinking water. The impacts of increasing surface water temperatures could also be compounded by additional climate change-related and anthropogenic impacts, including decreased flows during drought periods and watershed land use changes. As these impacts affect aquatic ecosystems and other natural resources, drinking water suppliers may be competing with multiple interests when it comes to surface water quality upstream of their intakes. Research involving the impacts of climate change on water supply has to date largely focused on water quantity, or flow, impacts (Langan et al., 2001). More targeted research is needed to understand the impacts of climate change on surface water quality, and what future changes will present the greatest challenges to the drinking water treatment process.

The thermal regimes of rivers are determined by complex interactions between climatological, hydrological, and anthropogenic influences (Caissie, 2006). The ability to explain and predict surface water temperature from air temperature is therefore a powerful tool in accounting for temperature impacts in a utility's planning and operation strategies. As stated in Webb et al. (2008), "...streams differ in their sensitivity to human modifications and their assimilative capacity for heat." Consequently, efforts to understand air/water temperature relationships need to be examined on a stream-by-stream basis. The validity of the air/water temperature relationship on the Schuylkill River at Philadelphia could be further validated through additional data collection and analyses over longer time periods. Monitoring the air/water temperature relationship over time could also provide insight into how anthropogenic or natural changes are impacting the Schuylkill River's thermal regime. Furthermore, the climate change projections associated with the Philadelphia area should continue to be updated as the accuracy of climate models improves.

There is uncertainty associated with future climate conditions, with how these conditions will impact surface water supplies, and the level of adaptation that will be necessary on the part of drinking water providers. With the need to adapt and remain flexible to changing climatic conditions, utilities are best served by using a methodical approach based on the latest science to understand and transform potential impacts into actionable planning, operations, and management decisions.

## VII. Future Research Opportunities

There are numerous opportunities to expand upon the research included in this report. Three areas of potential future research that pertain to the air/water temperature relationship and its implications are provided below.

- 1) **Assess other impacts of the air/water temperature relationship.** This MP provides brief insight into one potential impact of increasing surface water temperature on the drinking water treatment process: the formation of disinfection byproducts. A future study could reference the air/water temperature relationship defined in this analysis, and explore other implications in greater detail, including changes in surface water chemistry, primary productivity and/or ecological impacts. A similar analysis could also be performed for maximum and/or minimum stream temperatures in the Schuylkill River and the extent of impacts at these more extreme temperatures.
- 2) **Explore different air/water temperature relationships on the Schuylkill River at Philadelphia or in another type of water body in the PA region.** If one were to continue analysis of the Schuylkill River at Philadelphia, a regression model for the air/water temperature relationship in another season besides summer could be developed and evaluated. Flow could be added as another predictor variable in developing a regression model for a different season. Additionally, other types of water bodies could be explored, such as lakes or reservoirs, to determine if there are notable differences in the air/water temperature relationship.
- 3) **Characterize and compare the water quality profiles of a river in a southern state with a river, such as the Schuylkill River, in Pennsylvania.** Previous climate change research provides analogies of Pennsylvania's future climate to present-day climates in southern U.S. states. Specifically, the Pennsylvania Climate Impact Assessment (2009), states that "Moderate climate change on the order of 1-3°C increase in average annual temperature in southeastern PA, where most of PA's agriculture is located, would make it similar to present-day Maryland. Significantly greater climate change on the order of a 5-6°C increase in average annual temperature would make it close to that of present-day northern Georgia" (Shortle et al., 2009). If one were to compare, for example, the water quality profile of a river in northern Georgia with the water quality profile of the Schuylkill River, it may be possible to identify future water quality impacts that PWD and other water users, including industry, will need to adapt to. It would be important to consider rivers that are comparable in drainage area size, watershed land use type(s) and flow.

## Acknowledgements

I would like to thank the following people for their guidance and support throughout the preparation of my Master's Project:

- Dr. Martin Doyle (Advisor), Professor, Nicholas School of the Environment, Duke University
- Dr. Lauren Patterson, Water Policy Program, Nicholas Institute, Duke University
- Dr. Laifang Li, Nicholas School doctoral graduate, Woods Hole Oceanographic Institution
- Dr. Wenhong Li, Assistant Professor of Climate, Earth and Ocean Sciences, Duke University
- Dr. Molly Hesson, Source Water Protection Program, Philadelphia Water Department
- Kelly Anderson, Manager, Source Water Protection Program, Philadelphia Water Department
- Dr. Betsy Albright, Assistant Professor of the Practice, Environmental Science and Policy Methods, Nicholas School of the Environment

## References

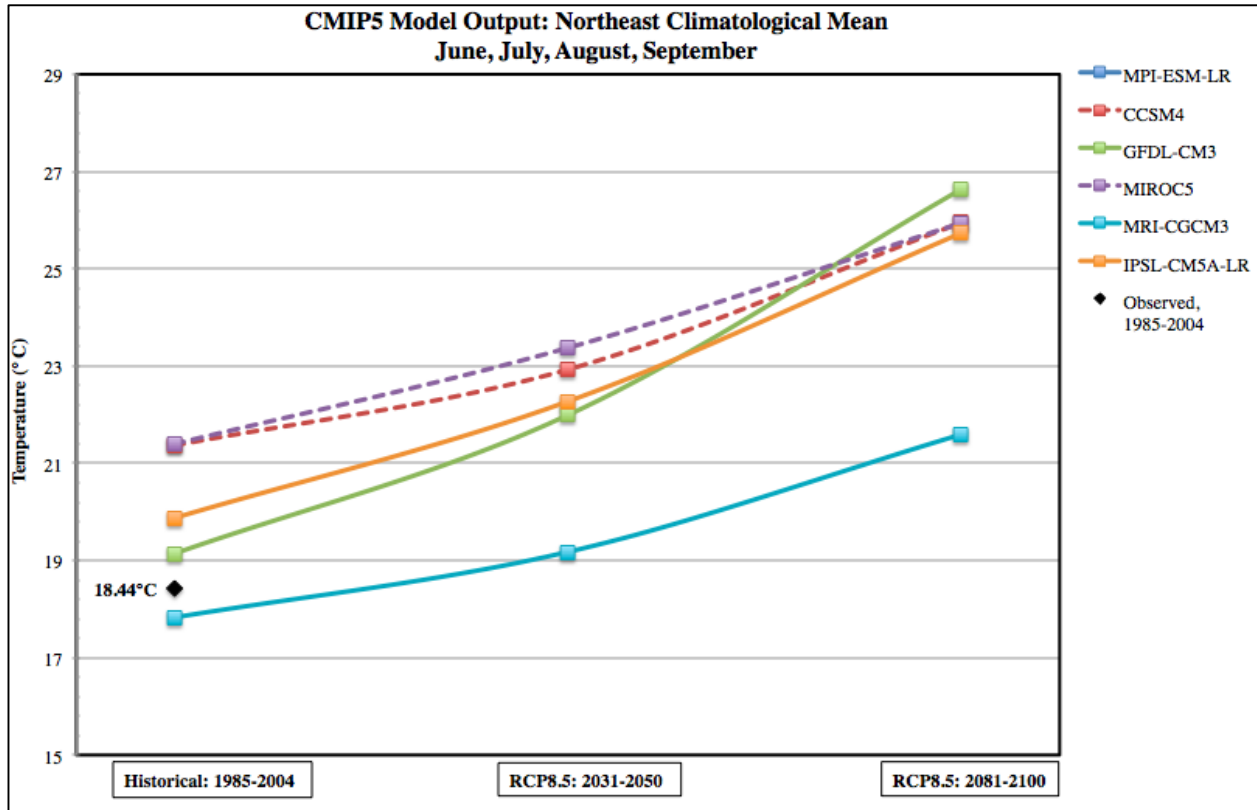
- Blauch, David N. (2014). Calorimetry; Heat Capacity Concepts. Retrieved October 6, 2014, from <http://www.chm.davidson.edu/vce/calorimetry/heatcapacity.html>
- Caissie, Daniel. (2006). The thermal regime of rivers: a review. *Freshwater Biology*, 51(8), 1389-1406.
- Crawley, M.J. (2011). *Statistics: An Introduction using R*: Wiley.
- Ducharne, Agnès. (2008). Importance of stream temperature to climate change impact on water quality. *Hydrology and Earth System Sciences Discussions*, 12(3), 797-810.
- Earth System Grid Federation (ESGF). Retrieved June 3, 2014. [pcmdi9.llnl.gov/esfg-web-fe/live](http://pcmdi9.llnl.gov/esfg-web-fe/live).
- Erickson, Troy R, & Stefan, Heinz G. (1996). Correlations of Oklahoma stream temperatures with air temperatures.
- Erickson, Troy R, Mohseni, Omid, & Stefan, Heinz G. (1998). Estimation of an upper bound for weekly stream temperatures.
- Erickson, Troy R, & Stefan, Heinz G. (2000). Linear air/water temperature correlations for streams during open water periods. *Journal of Hydrologic Engineering*, 5(3), 317-321.
- ICF Incorporated, LLC. (2014). Useful Climate Information for Philadelphia: Past and Future. Philadelphia, PA.
- Intergovernmental Panel on Climate Change (IPCC). (2014). Fifth Assessment Report (AR5): Climate Change 2014. New York, NY.
- Langan, S. J., Johnston, L., Donaghy, M. J., Youngson, A.F., Hay, D.W., & Soulsby, C. (2001). Variation in river water temperatures in an upland stream over a 30-year period. *Science of the Total Environment*, 265(1), 195-207.
- Lau, KM, Sud, Y, & Kim, JH. (1996). Intercomparison of hydrologic processes in AMIP GCMs. *Bulletin of the American Meteorological Society*, 77(10), 2209-2227.
- McGuire, Michael J, Karanfil, Tanju, Krasner, Stuart W, Reckhow, David A, Roberson, J Alan, Summers, R Scott, . . . Xie, Yuefeng. (2014). Not your granddad's disinfection by-product problems and solutions (PDF). *Journal-American Water Works Association*, 106(8), 54-73.
- Meinshausen, Malte et al, (2011). The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Climatic Change*, 109(1-2), 213-241.
- Mohseni, Omid, Stefan, Heinz G, & Erickson, Troy R. (1998). A nonlinear regression model for weekly stream temperatures. *Water Resources Research*, 34(10), 2685-2692.
- Mohseni, Omid, Erickson, Troy R, & Stefan, Heinz G. (1999). Sensitivity of stream temperatures in the United States to air temperatures projected under a global warming scenario. *Water Resources Research*, 35(12), 3723-3733.

- Mohseni, O, & Stefan, HG. (1999). Stream temperature/air temperature relationship: a physical interpretation. *Journal of Hydrology*, 218(3), 128-141.
- Morrill, Jean C, Bales, Roger C, & Conklin, Martha H. (2005). Estimating stream temperature from air temperature: implications for future water quality. *Journal of Environmental Engineering*, 131(1), 139-146.
- National Oceanic and Atmospheric Administration (NOAA), Earth Systems Research Laboratory, Physical Sciences Division. CDC Derived NCEP Reanalysis Products Surface Level. Retrieved April 25, 2014 [http://www.esrl.noaa.gov/psd/cgi-bin/db\\_search/DBSearch.pl?Dataset=CDC+Derived+NCEP+Reanalysis+Products+Surface+Level&Variable=Air+Temperature&Statistic=Mean&group=0&submit=Search](http://www.esrl.noaa.gov/psd/cgi-bin/db_search/DBSearch.pl?Dataset=CDC+Derived+NCEP+Reanalysis+Products+Surface+Level&Variable=Air+Temperature&Statistic=Mean&group=0&submit=Search)
- Northeast Climate Impacts Assessment (NECIA). (2006). Climate Change in the U.S. Northeast. Cambridge, Massachusetts: Union of Concerned Scientists.
- Peters, Glen P. et al, (2013). The challenge to keep global warming below 2°C. *Nature Climate Change*, 3(1), 4-6.
- Pilgrim, John M, Fang, Xing, & Stefan, Heinz G. (1998). Stream Temperature Correlations with Air Temperature in Minnesota: Implications for Climate Warming: Wiley Online Library.
- Program for Climate Model Diagnosis and Intercomparison (PCMDI). CMIP5 - Coupled Model Intercomparison Project Phase 5 - Overview. *CMIP5 Coupled Model Intercomparison Project*. Retrieved October 6, 2014, from [cmip-pcmdi.llnl.gov/cmip5/](http://cmip-pcmdi.llnl.gov/cmip5/)
- Punzet, Manuel, Voß, Frank, Voß, Anja, Kynast, Ellen, & Bärlund, Ilona. (2012). A Global Approach to Assess the Potential Impact of Climate Change on Stream Water Temperatures and Related In-Stream First-Order Decay Rates. *Journal of Hydrometeorology*, 13(3), 1052-1065. doi: 10.1175/JHM-D-11-0138.1
- Rodriguez, Manuel J, & Sérodes, Jean-B. (2001). Spatial and temporal evolution of trihalomethanes in three water distribution systems. *Water Research*, 35(6), 1572-1586.
- Shortle, James et al. (2009). Pennsylvania Climate Impact Assessment Report to the Department of Environmental Protection: Environment and Natural Resources Institute, Pennsylvania State University. Retrieved November 4, 2013 from <http://www.elibrary.dep.state.pa.us/dsweb/Get/Document-75375/7000-BK-DEP4252.pdf>
- Teksoy, Arzu, Alkan, Ufuk, & Başkaya, Hüseyin Savaş. (2008). Influence of the treatment process combinations on the formation of THM species in water. *Separation and Purification Technology*, 61(3), 447-454.
- Thornton, P.E., M.M. Thornton, B.W. Mayer, N. Wilhelmi, Y. Wei, R. Devarakonda, and R.B. Cook. 2014. Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 2. Data set. Available on-line [<http://daac.ornl.gov>] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA. Date accessed: YYYY/MM/DD. Temporal range: YYYY/MM/DD-YYYY/MM/DD. Spatial range: N=DD.DD, S=DD.DD, E=DDD.DD, W=DDD.DD. <http://dx.doi.org/10.3334/ORNLDAAAC/1219>

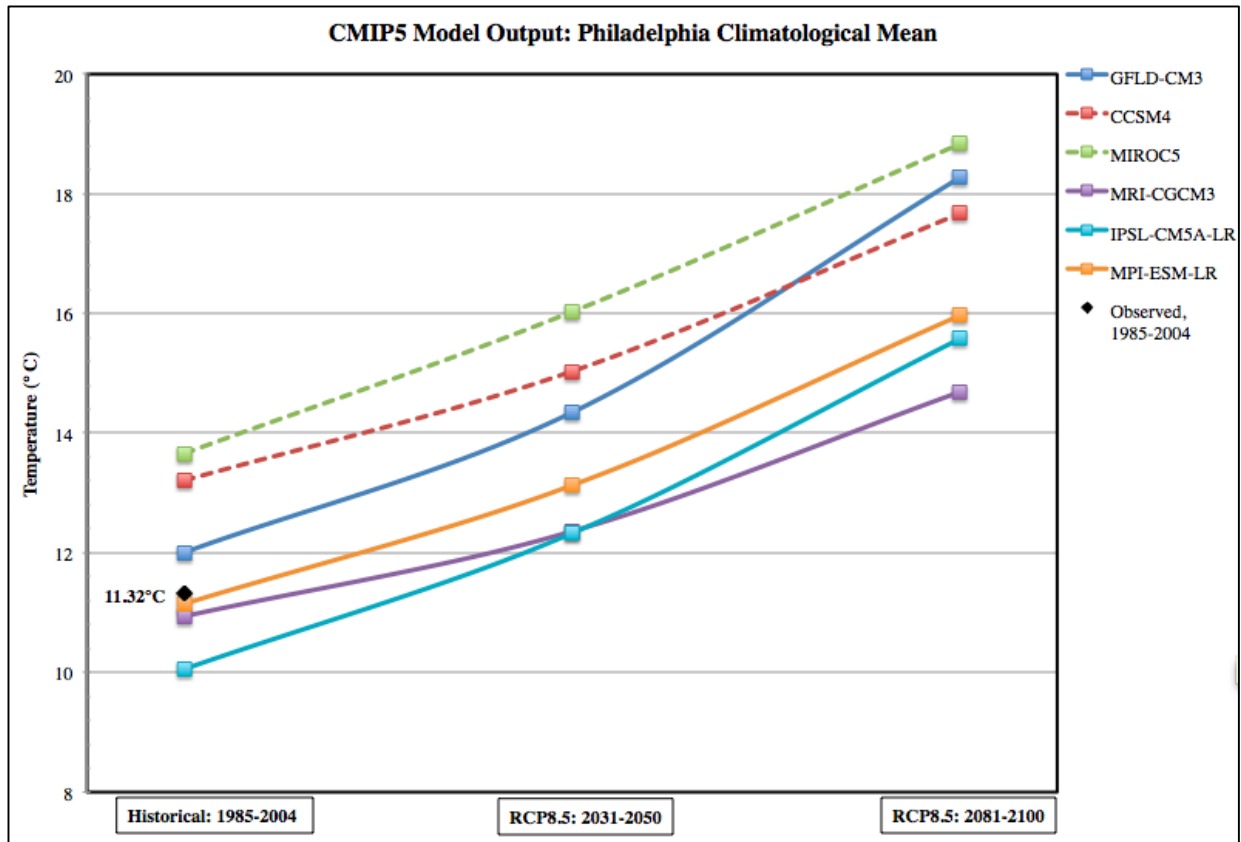
- U.S. Environmental Protection Agency. (2011). Disinfection Byproduct Health Effects. Retrieved April 27, 2014, from [http://www.epa.gov/enviro/html/icr/dbp\\_health.html](http://www.epa.gov/enviro/html/icr/dbp_health.html)
- United States Geological Survey (USGS). (2014). USGS 01474500 Schuylkill River at Philadelphia, PA. *National Water Information System: Web Interface*. Retrieved September 3, 2014, from [http://waterdata.usgs.gov/pa/nwis/uv/?site\\_no=01474500&PARAMeter\\_cd=00065,00060,00010](http://waterdata.usgs.gov/pa/nwis/uv/?site_no=01474500&PARAMeter_cd=00065,00060,00010)
- Webb, Bruce W, Hannah, David M, Moore, R Dan, Brown, Lee E, & Nobilis, Franz. (2008). Recent advances in stream and river temperature research. *Hydrological Processes*, 22(7), 902-918.
- World Health Organization (WHO). (2008). Impacts of Europe's changing climate: 2008 indicator-based assessment.

# Appendices

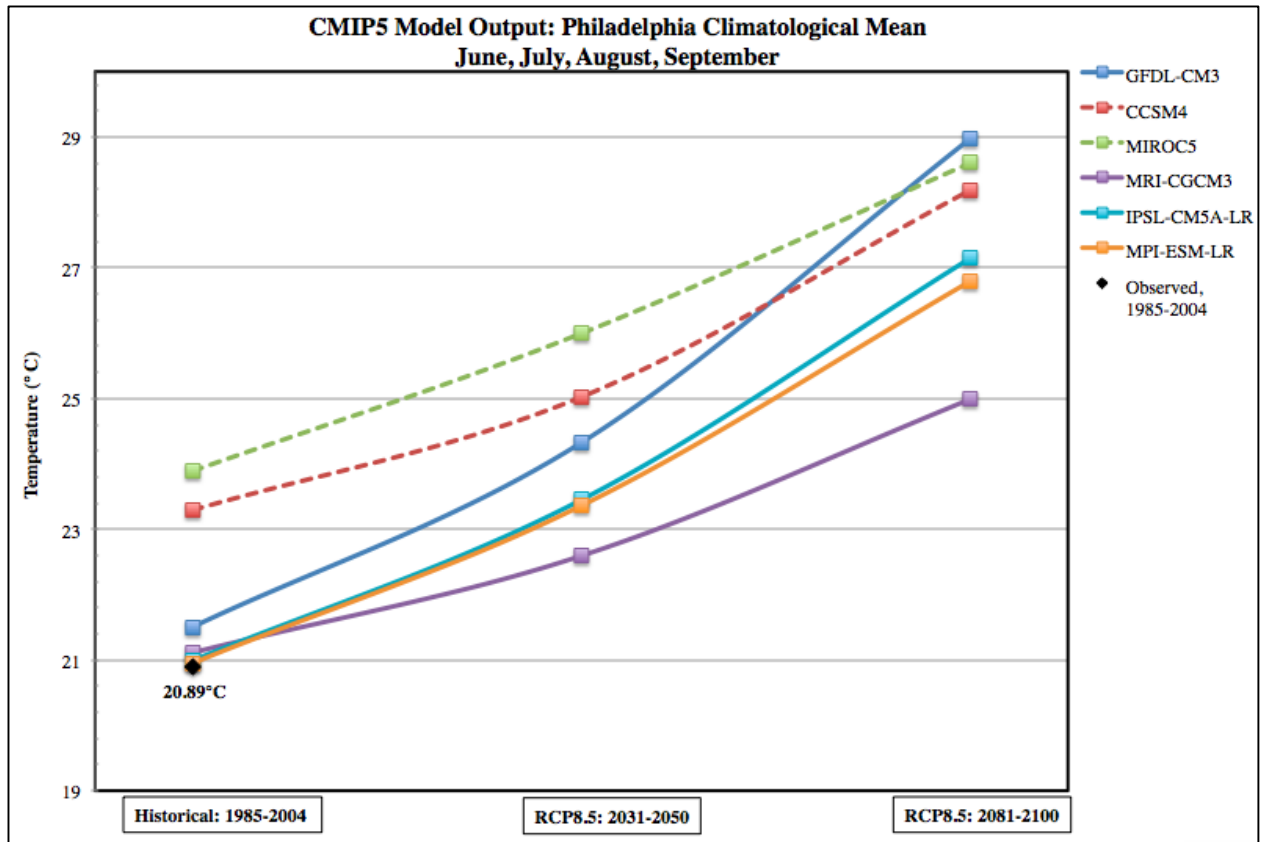
## APPENDIX A: Climate Model Evaluation



**Figure 1A.** Northeast climatological mean from June-September for near-surface air temperature calculated over historical, RCP8.5 mid-century and RCP8.5 end-of-century time periods using six different global climate models (GCMs). The observed June-September mean for near-surface air temperature from 1985-2004 is also included to illustrate performance of the models. The two dashed lines indicate the poorest performing models relative to the observed value.



**Figure 2A.** Philadelphia climatological mean for near-surface air temperature calculated over historical, RCP8.5 mid-century and RCP8.5 end-of-century time periods using six different global climate models (GCMs). The observed climatological mean for near-surface air temperature near Philadelphia from 1985-2004 is also included to illustrate performance of the models. The two dashed lines indicate the poorest performing models relative to the observed value.

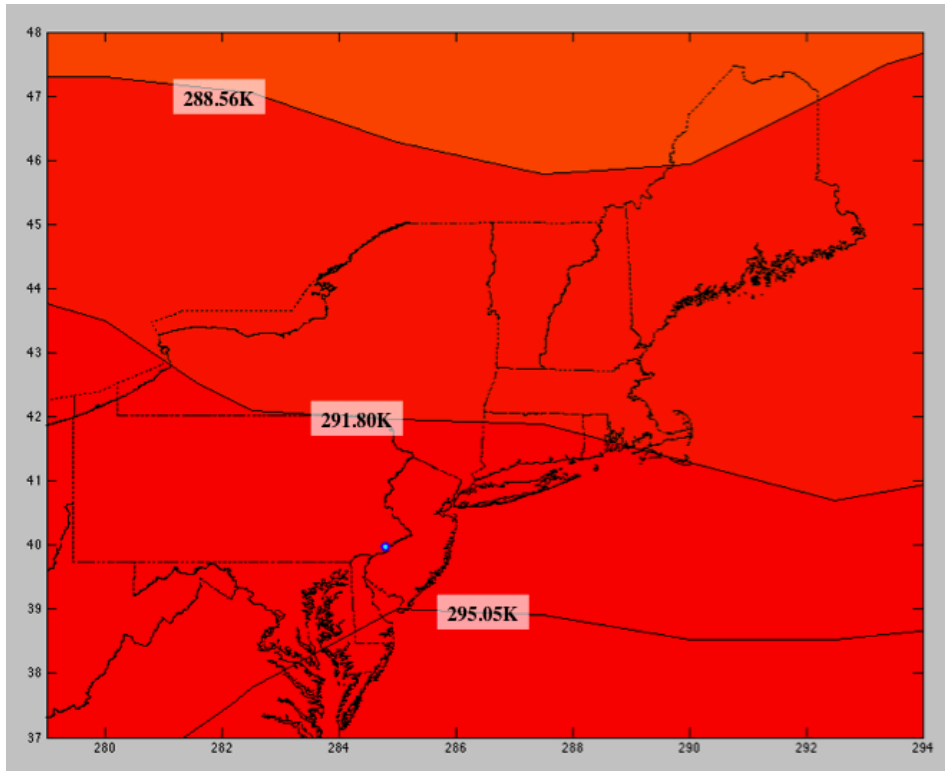


**Figure 3A.** Philadelphia climatological mean from June-September for near-surface air temperature calculated over historical, RCP8.5 mid-century and RCP8.5 end-of-century time periods using six different global climate models (GCMs). The observed June-September mean for near-surface air temperature near Philadelphia from 1985-2004 is also included to illustrate performance of the models. The two dashed lines indicate the poorest performing models relative to the observed value.

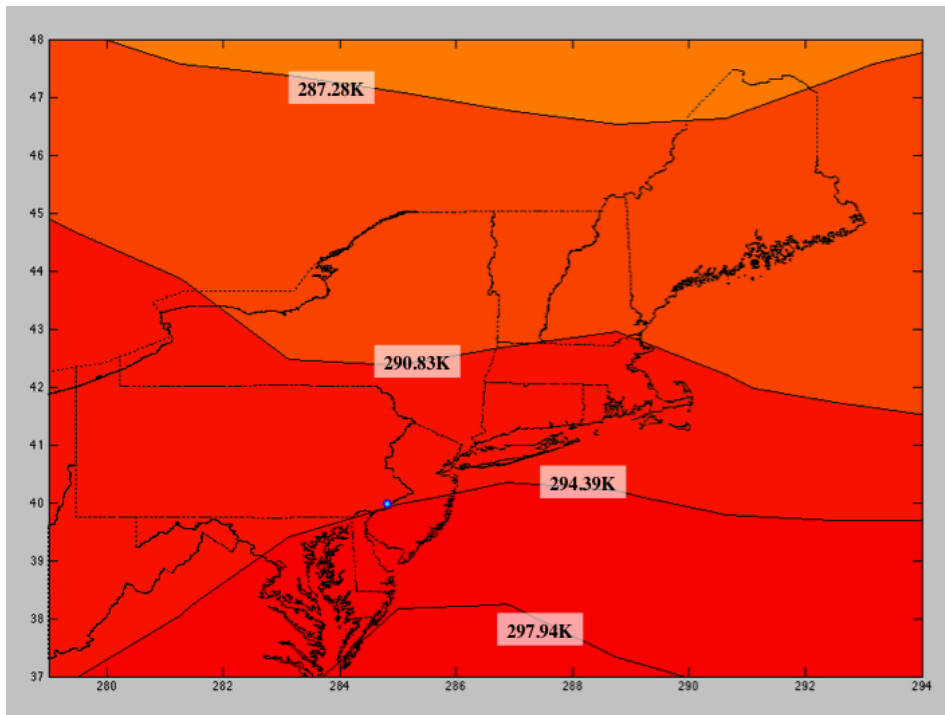
**Table 1A.** Locations of grid points closest to USGS Fairmount Gauge 01474500 for each NetCDF file used in the analysis

	Latitude	Longitude
<b>USGS Gauge</b>		
1474500	39.97	284.81
<b>Observed</b>		
NOAA	40.00	285.00
<b>Modeled</b>		
MPI-ESM-LR	40.10	285.00
MRI-CGCM3	39.81	284.63
IPSL-CM5A-LR	40.74	285.00
GFDL-CM3	39.00	283.75
CCSM4	40.05	285.00
MIROC5	39.92	285.47

**APPENDIX B: Near-Surface Air Temperature Contour Maps, Historical, MPI-ESM-LR Model**

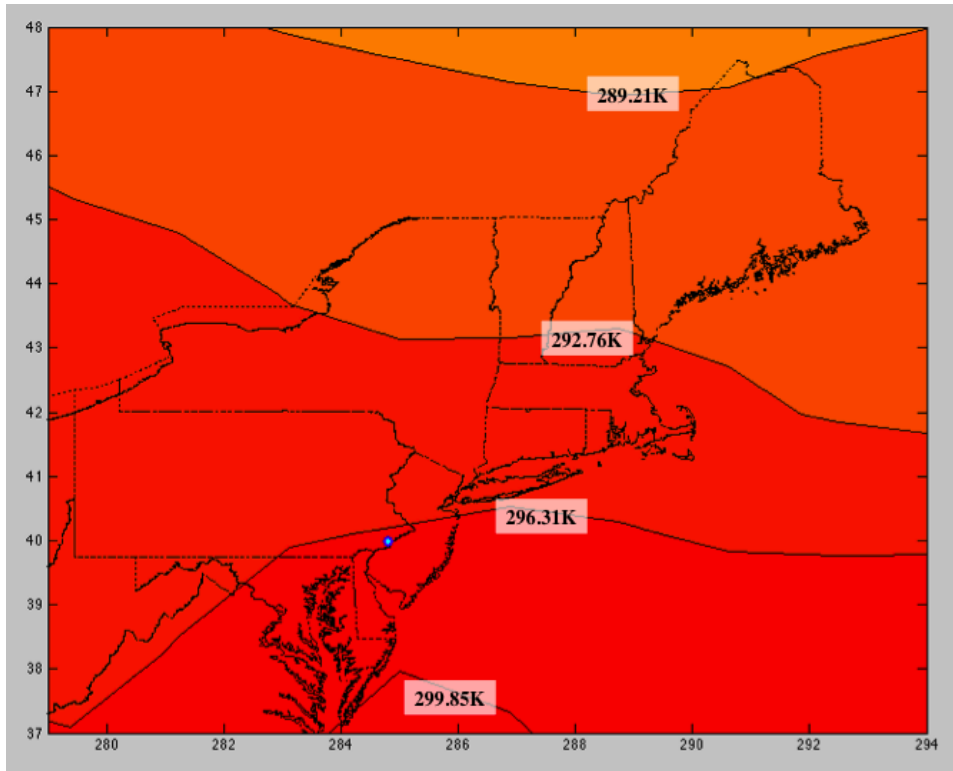


**Figure 1B .** NOAA observed near-surface air temperatures for the months of June through September, 1985-2004. The blue dot indicates the location of USGS Gauge 01474500.

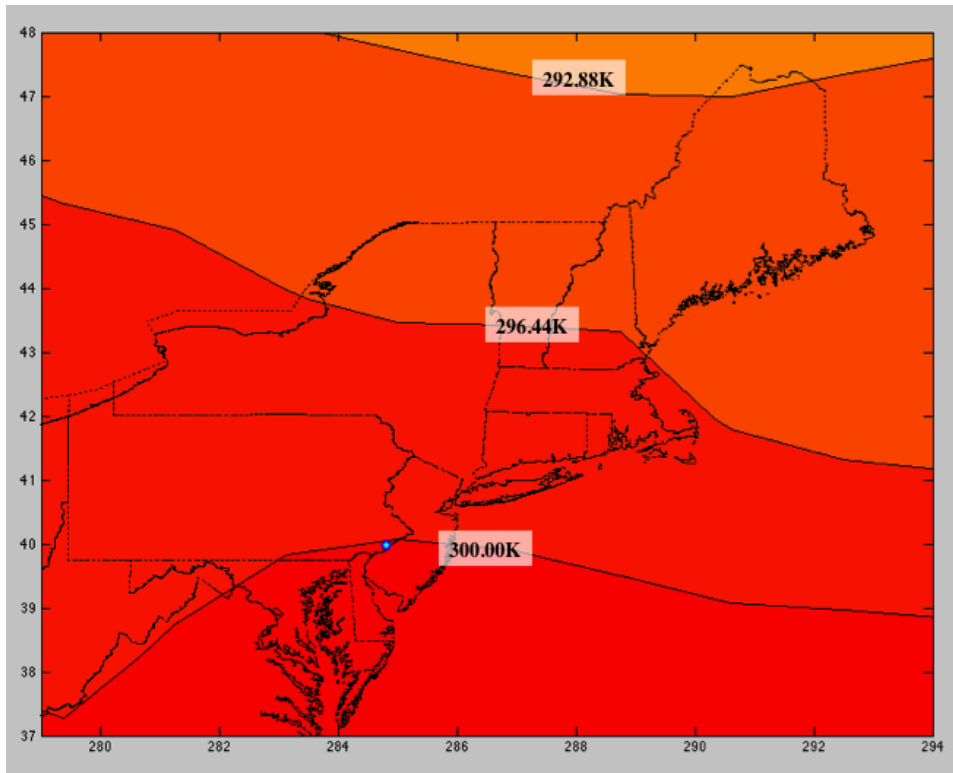


**Figure 2B.** MPI-ESM-LR near-surface air temperatures for the months of June through September, 1985-2004. The blue dot indicates the location of USGS Gauge 01474500.

**APPENDIX C: Near-Surface Air Temperature Contour Maps, RCP8.5, MPI-ESM-LR Model**

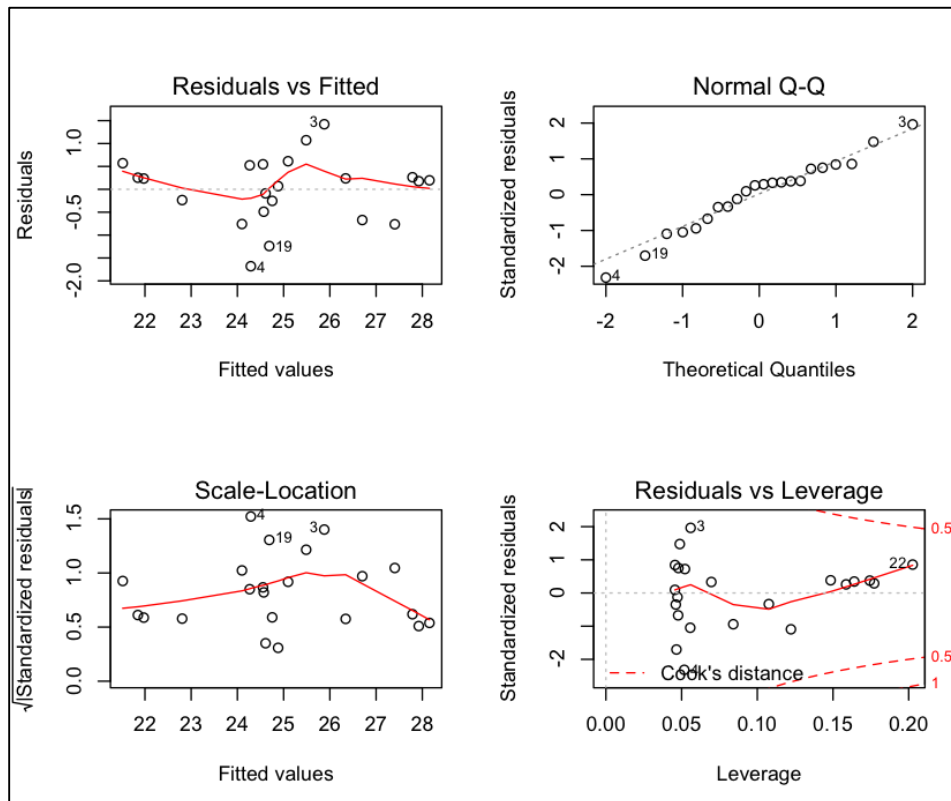


**Figure 1C.** RCP8.5 MPI-ESM-LR near-surface air temperatures for the months of June through September, 2031-2050. The blue dot indicates the location of USGS Gauge 01474500.



**Figure 2C.** RCP8.5 MPI-ESM-LR near-surface air temperatures for the months of June through September 2081-2100. The blue dot indicates the location of USGS Gauge 01474500.

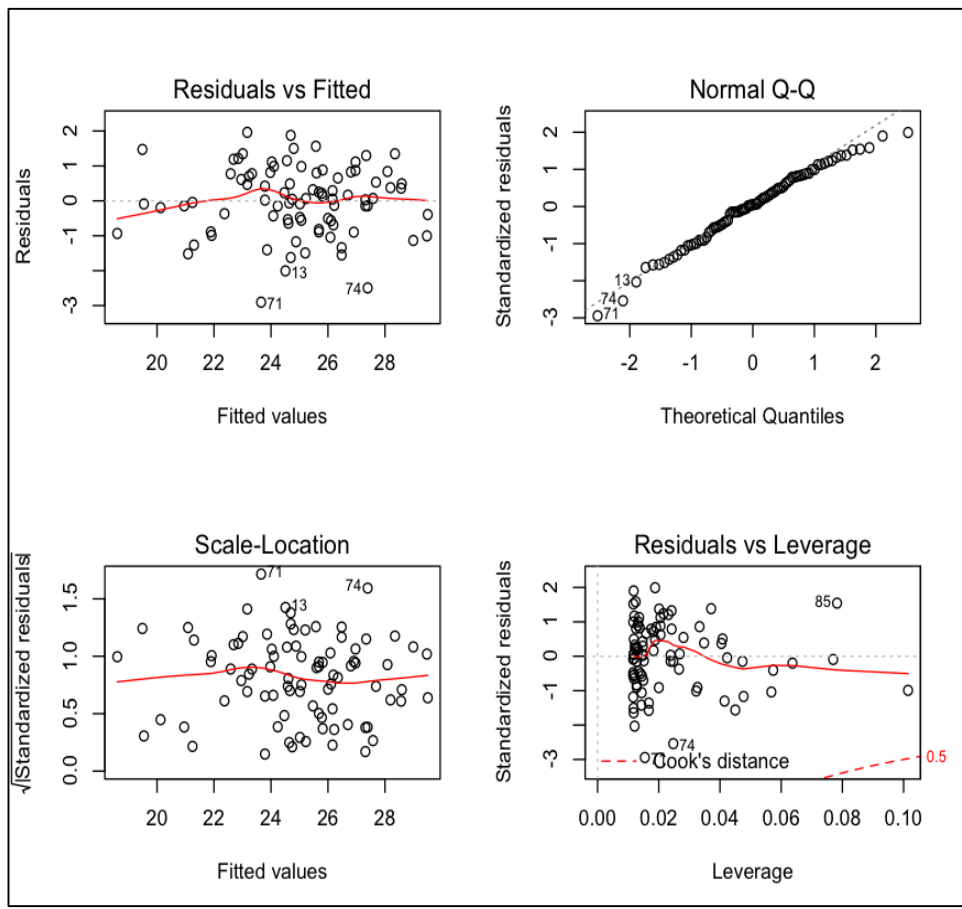
## APPENDIX D: Diagnostic Plots for Air/water Temperature Linear Regressions



**Figure 1D.** Diagnostic plots for the monthly average JJAS air/water temperature linear regression model

The purpose of the above four plots is to test the two major assumptions in a linear regression model: constant variance, or homoscedasticity, and normally distributed errors (Crawley, 2011). The first plot titled “Residuals vs. Fitted” in Figure 1D illustrates a fairly even scattering of points across all fitted values. This figure is checking for constant variance, or homoscedasticity. There is an issue with the distribution of variance if the variances increase as the fitted values increase (Crawley, 2011). Since a clustered pattern of points is not illustrated in this plot, and the scatter of residuals is fairly constant across fitted values, it can be inferred that the assumption of homoscedasticity is not violated. The second plot in the upper right-hand corner is referred to as the normal Q-Q plot. In cases where the errors follow a normal distribution, the plotted points will form a straight line (Crawley, 2011). The slight curvatures in the Q-Q plot indicate that the errors may not be perfectly normally distributed, which could have to do with the fact that the dataset for monthly average temperatures is quite limited, amplifying the impact of any outliers. The third plot in the lower left-hand corner titled “Scale-Location,” is similar in concept to the first plot in the series, but uses the square root of the standardized residuals on the y-axis (Crawley, 2011). The ideal plot in this case will also illustrate an even scattering of points. Since there is no clustering of points for this example, it can again be inferred that the assumption of homoscedasticity is

met. The final plot titled “Residuals vs. Leverage” illustrates the Cook’s distance for each value of the response variable (Crawley, 2011). The objective of this plot is to indicate any response (y) values that have a large impact on the parameter estimate for the model (Crawley, 2011). None of the plotted values on this last chart fall outside a Cook distance of 1, indicating that no single response value significantly influences the model structure. Based on the four plots contained in Figure 1D, it can be concluded that the linear model for monthly average JJAS air/water temperatures in Philadelphia performs adequately well given the limited dataset. The only potential concern has to do with the normality of the errors, which may be resolved through use of a more extensive dataset.



**Figure 2D.** Diagnostic plots for the weekly average JJAS air/water temperature linear regression model

The purpose of the above four plots is to test the two major assumptions in a linear regression model: constant variance, or homoscedasticity, and normally distributed errors (Crawley, 2011). The first plot titled “Residuals vs. Fitted” illustrates a fairly even scattering of points, making it possible to infer that the assumption of homoscedasticity, or constant variance, is not violated in the case of the weekly average air/water linear regression model. The Q-Q plot in the upper right-hand corner is testing for normality of

the errors. The slight curvature at the ends of the Q-Q plot indicate that the errors may not be perfectly normally distributed, which could have to do with the fact that the dataset for weekly average temperatures is quite limited, amplifying the impact of any outliers. The third plot in the lower left-hand corner titled "Scale-Location," also illustrates an even scattering of points, very similar to the first plot in the series. Since there is no clustering of points, it can again be inferred that the assumption of homoscedasticity is met. The final plot titled "Residuals vs. Leverage" illustrates the Cook's distance for each value of the response variable (Crawley, 2011). The objective of this plot, as stated above, is to indicate any response (y) values that have a large impact on the parameter estimate for the model (Crawley, 2011). None of the plotted values on this last chart fall outside a Cook distance of 1, indicating that no single response value significantly influences the model structure. Based on the four plots contained in Figure 2D, it can be concluded that the linear model for weekly average air/water temperature data generally meets all the required assumptions.