ABSTRACT

Lu, Jianbiao. Modeling Regional Evapotranspiration for Forested Watersheds across the Southern United States. (Under the direction of Dr. Ge Sun)

Evapotranspiration (ET) is the process that returns water to the atmosphere and therefore completes the hydrologic cycle. ET is a major component in the hydrological balance, and therefore is important to understanding forest water yield, sediment and nutrient movement. However, direct measurement of forest ET for a large region is not possible. The objectives of this study were to develop a model to estimate long-term annual actual evapotranspiration (AET) for forested watersheds across the southern United States (U.S.) and to compare the differences among six potential evapotranspiration (PET) methods. The developed AET model will be used to study hydrologic effects of climate and landuse changes. The six compared PET methods include three temperature-based methods (the Thornthwaite, Hamon and Hargreaves-Samani method) and three radiation-based methods (the Turc, Makkink and Priestley-Taylor method).

A GIS database including land cover, hydrology and climate was developed for thirty-nine forested watersheds across the southern U.S.. Based on these data, a long-term annual AET model was developed. The independent variables included in the model are rainfall, latitude, elevation and percentage of land cover of conifer forests and water body in the watersheds. The model has a $R^2$ of 0.85 and is sufficient to predict long-term annual AET for forested watersheds across the southern U.S.. Six PET methods were highly correlated
but significantly different from each other. Greater differences were found among the temperature-based PET methods than radiation-based PET methods. In comparisons of the six PET methods, the Priestley-Taylor, Hamon and Ture methods performed better than the Thornthwaite, Makkink and Hargreaves-Samani methods.
Modeling Regional Evapotranspiration for Forested Watersheds across the Southern United States

By

Jianbiao Lu

A thesis submitted to the Graduate Faculty of North Carolina State University in partial fulfillment of the requirements for the Degree of Master of Science

Department of Forestry

Raleigh

2002

APPROVED BY:

Dr. James D. Gregory

Dr. David A. Dickey

Dr. Steven G. McNulty

Dr. Devendra M. Amatya

Dr. Ge Sun

(Chair of Advisory Committee)
Jianbiao Lu was born on April 20, 1976 in Guangxi Province, P. R. China. He received his Bachelor degree from Beijing Forestry University, China in July 1998, and entered the Graduate School of Beijing Forestry University thereafter. In May 2000, he was admitted to North Carolina State University to continue his Master of Science studies.
ACKNOWLEDGMENTS

First and foremost, I would like to express my sincere appreciation and gratitude to my major advisor, Dr. Ge Sun, for his numerous encouragement, suggestions, guidance and assistance during all stages of this study. I would also like to thank my other committee members, Dr. Steve McNulty, Dr. Devendra Amatya, Dr. James Gregory and Dr. David Dickey for their guidance and assistance throughout the project.

I wish to express my sincere appreciation to all the people in the Southern Global Change Program (SGCP), USDA Forest Service. This project was funded by SGCP and the people in the program provided tremendous help. Their kind support and assistance made this project possible and made me have a happy experience during my study and research here. Thank you so much!

In this research project, the data sets were requested from many different sources. I would like to thank Dr. Yiguo Liang, a former graduate student at the University of Alabama, for providing the data for 17 Southeast US watersheds; Dr. Devendra Amatya and Dr. Wayne Skaggs at the Department of Biological and Agricultural Engineering, North Carolina State University, for providing data for the Carteret and Parker tract watersheds; Dr. Randy Kolka at North Central Research Station, USDA Forest Service, for providing data for the Coles Fork watershed; and Dr. Patrick Mulholland at Oak Ridge National Laboratory for providing hydrologic data for Walker Branch watershed. Also I want to thank Dr. John Parsons at the Department of Biological and Agricultural Engineering, North Carolina State University, for his help and advice.
I also feel great gratitude to all my friends for their assistance during my study in North Carolina State University.

Finally, my deep appreciation is offered to all my families in China, for their love, support, encouragement and understanding.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>1. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>2. OBJECTIVES</td>
<td>4</td>
</tr>
<tr>
<td>3. LITERATURE REVIEW–REGIONAL ET ESTIMATES</td>
<td>5</td>
</tr>
<tr>
<td>3.1. Watershed water balance method</td>
<td>5</td>
</tr>
<tr>
<td>3.2. Methods based on planetary boundary layer theory</td>
<td>6</td>
</tr>
<tr>
<td>3.3. Methods based on remote sensing data</td>
<td>7</td>
</tr>
<tr>
<td>3.4. Methods based on complementary relationship hypothesis</td>
<td>9</td>
</tr>
<tr>
<td>3.5. Hydrological model and empirical model estimates</td>
<td>10</td>
</tr>
<tr>
<td>3.6. Atmospheric moisture budget estimates</td>
<td>12</td>
</tr>
<tr>
<td>4. METHODOLOGY</td>
<td>14</td>
</tr>
<tr>
<td>4.1. Selections of six PET methods</td>
<td>14</td>
</tr>
<tr>
<td>4.1.1. Thornthwaite method</td>
<td>14</td>
</tr>
<tr>
<td>4.1.2. Hamon method</td>
<td>16</td>
</tr>
<tr>
<td>4.1.3. Turc method</td>
<td>16</td>
</tr>
<tr>
<td>4.1.4. Priestley-Taylor method</td>
<td>17</td>
</tr>
<tr>
<td>4.1.5. Makkink method</td>
<td>20</td>
</tr>
<tr>
<td>4.1.6. Hargreaves-Samani method</td>
<td>20</td>
</tr>
<tr>
<td>4.2. Database construction</td>
<td>21</td>
</tr>
<tr>
<td>4.2.1. Introduction</td>
<td>21</td>
</tr>
<tr>
<td>4.2.2. Data collection</td>
<td>22</td>
</tr>
<tr>
<td>4.2.2.1. Data of experimental forested watersheds</td>
<td>22</td>
</tr>
<tr>
<td>4.2.2.2. Data of USGS monitored watersheds</td>
<td>24</td>
</tr>
<tr>
<td>4.2.2.3. Obtain watershed elevation and land cover</td>
<td>25</td>
</tr>
<tr>
<td>4.2.2.4. Derive net radiation</td>
<td>26</td>
</tr>
<tr>
<td>4.2.2.5. Unit conversions</td>
<td>27</td>
</tr>
<tr>
<td>4.2.2.6. Data summaries</td>
<td>28</td>
</tr>
<tr>
<td>4.3. AET and PET estimates</td>
<td>32</td>
</tr>
<tr>
<td>4.3.1. AET estimates</td>
<td>32</td>
</tr>
<tr>
<td>Section</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>4.3.2. PET estimates</td>
<td>32</td>
</tr>
<tr>
<td>4.4. Data analysis</td>
<td>34</td>
</tr>
<tr>
<td>4.4.1. Model building methods</td>
<td>34</td>
</tr>
<tr>
<td>4.4.2. Regression diagnostics</td>
<td>35</td>
</tr>
<tr>
<td>4.4.3. PET comparisons</td>
<td>37</td>
</tr>
<tr>
<td>5. RESULTS</td>
<td>39</td>
</tr>
<tr>
<td>5.1. Fitting the long-term annual AET regression model</td>
<td>39</td>
</tr>
<tr>
<td>5.1.1. Model building</td>
<td>39</td>
</tr>
<tr>
<td>5.1.2. Regression diagnostics</td>
<td>43</td>
</tr>
<tr>
<td>5.1.2.1. Residual analysis</td>
<td>43</td>
</tr>
<tr>
<td>5.1.2.2. Influence statistics</td>
<td>45</td>
</tr>
<tr>
<td>5.1.2.3. Collinearity diagnostics</td>
<td>46</td>
</tr>
<tr>
<td>5.2. Six PET comparisons</td>
<td>47</td>
</tr>
<tr>
<td>5.2.1. Correlation between six PET methods</td>
<td>47</td>
</tr>
<tr>
<td>5.2.2. Correlation between six PET methods and AET estimates</td>
<td>47</td>
</tr>
<tr>
<td>5.2.3. Differences between six PET methods</td>
<td>48</td>
</tr>
<tr>
<td>5.3. Conclusions and discussion</td>
<td>57</td>
</tr>
<tr>
<td>6. REFERENCES</td>
<td>59</td>
</tr>
<tr>
<td>7. APPENDICES</td>
<td>65</td>
</tr>
<tr>
<td>APPENDIX A. Methodology of four eight-digit HUC watershed selections in South Carolina and Georgia</td>
<td>66</td>
</tr>
<tr>
<td>APPENDIX B. Procedures of delineating a watershed (site number 32) and deriving its elevation and land cover in Georgia</td>
<td>74</td>
</tr>
<tr>
<td>APPENDIX C. Summary of the gauging stations in the study watersheds</td>
<td>87</td>
</tr>
<tr>
<td>APPENDIX D. Computer program for six PET methods in Visual Basic 6.0</td>
<td>89</td>
</tr>
<tr>
<td>APPENDIX E. Formats of inputs and outputs of the PET program</td>
<td>101</td>
</tr>
<tr>
<td>APPENDIX F. SAS program for the long-term annual AET model building</td>
<td>104</td>
</tr>
<tr>
<td>APPENDIX G. SAS program for comparisons of six PET methods</td>
<td>105</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>Table 4.1. Summary of study watersheds.</td>
<td>30</td>
</tr>
<tr>
<td>Table 5.1. Variables used in the AET model building</td>
<td>39</td>
</tr>
<tr>
<td>Table 5.2. Summary of the stepwise selection method for the AET model building</td>
<td>42</td>
</tr>
<tr>
<td>Table 5.3. Analysis of variance of the AET model</td>
<td>43</td>
</tr>
<tr>
<td>Table 5.4. Parameter estimates of the AET model</td>
<td>43</td>
</tr>
<tr>
<td>Table 5.5. Tests for normality of residuals</td>
<td>44</td>
</tr>
<tr>
<td>Table 5.6. Diagnostics statistics of the AET model</td>
<td>45</td>
</tr>
<tr>
<td>Table 5.7. Collinearity diagnostics (intercept adjusted) of the AET model</td>
<td>46</td>
</tr>
<tr>
<td>Table 5.8. Pearson correlation coefficients between six PET methods</td>
<td>47</td>
</tr>
<tr>
<td>Table 5.9. Pearson correlation coefficients between six PET methods and AET estimates</td>
<td>48</td>
</tr>
<tr>
<td>Table 5.10. Simple statistics for AET and PET estimated by six methods</td>
<td>49</td>
</tr>
<tr>
<td>Table 5.11. Multivariate statistics test for PET estimated by six methods</td>
<td>49</td>
</tr>
<tr>
<td>Table 7.1. Grids used to delineate the watershed boundary</td>
<td>80</td>
</tr>
<tr>
<td>Table 7.2. Watershed land cover compositions</td>
<td>85</td>
</tr>
<tr>
<td>Table 7.3. Summary of the gauging stations in the GIS database</td>
<td>87</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 4.1. Study watersheds for analysis</td>
<td>29</td>
</tr>
<tr>
<td>Figure 4.2. Program interface for six PET calculations and data summaries</td>
<td>33</td>
</tr>
<tr>
<td>Figure 5.1. Residual plot of long-term annual AET regression analysis</td>
<td>44</td>
</tr>
<tr>
<td>Figure 5.2. Plot of residuals against site number of the watersheds</td>
<td>44</td>
</tr>
<tr>
<td>Figure 5.3. Differences between AET and six PET values</td>
<td>49</td>
</tr>
<tr>
<td>Figure 5.4. PET simulated by the three temperature-based PET methods</td>
<td>50</td>
</tr>
<tr>
<td>Figure 5.5. PET simulated by the three radiation-based PET methods</td>
<td>51</td>
</tr>
<tr>
<td>Figure 5.6. Six PET simulated at Bradford watershed, FL</td>
<td>52</td>
</tr>
<tr>
<td>Figure 5.7. Six PET simulated at Carteret watershed, NC</td>
<td>52</td>
</tr>
<tr>
<td>Figure 5.8. Six PET simulated at a watershed in Arkansas (site number 15)</td>
<td>53</td>
</tr>
<tr>
<td>Figure 5.9. Six PET simulated at a watershed in Georgia (site number 32)</td>
<td>53</td>
</tr>
<tr>
<td>Figure 5.10. AET estimated by the water balance for all the watersheds</td>
<td>55</td>
</tr>
<tr>
<td>Figure 5.11. PET estimated by the Priestley-Taylor method for all the watersheds</td>
<td>56</td>
</tr>
<tr>
<td>Figure 7.1. Available solar radiation stations in Georgia and South Carolina</td>
<td>67</td>
</tr>
<tr>
<td>Figure 7.2. Selected eight-digit HUC watersheds in Georgia and South Carolina</td>
<td>68</td>
</tr>
<tr>
<td>Figure 7.3. Available HUC watersheds and stations in GA and SC</td>
<td>71</td>
</tr>
<tr>
<td>Figure 7.4. Available watersheds in GA and SC</td>
<td>73</td>
</tr>
<tr>
<td>Figure 7.5. Neatline polygon information for grids</td>
<td>77</td>
</tr>
<tr>
<td>Figure 7.6. County information for grids</td>
<td>79</td>
</tr>
<tr>
<td>Figure 7.7. The delineated watershed</td>
<td>84</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

Evapotranspiration (ET) is the process that returns water from the earth surface to the atmosphere and therefore completes the hydrologic cycle. ET can be divided into two sub-processes: evaporation and transpiration. Evaporation essentially occurs on the surface of open water such as lakes, reservoirs, or puddles, or from vegetation and ground surfaces. Transpiration involves the removal of water from the soil by plant roots, transport of the water through the plant into the leaf, and evaporation of the water from the leaf’s interior into the atmosphere (Ward and Elliot, 1995). The key controls on ET of a watershed are rainfall interception, net radiation, advection, turbulent transport, leaf area, and plant-available water capacity (Zhang et al., 2001).

ET is a major component in the hydrological balance. On the global scale, it represents more than 60% of precipitation inputs (Vörösmarty et al., 1998), and it is more than 70% of the annual precipitation for the entire U.S. (Gay, 1993; Brooks et al., 1997). It is an especially important factor in water resources planning. However, direct measurements of ET are rarely available, and it is impossible to measure the ET values on a large scale such as regional area. In the absence of measurements, an alternative solution is to use models to predict ET.

The concept of potential evapotranspiration (PET) was developed by Thornthwaite (1948) in the mid-1940s. It was defined as “the amount of water that could evaporate and transpire from a landscape fully covered by a homogeneous stand of vegetation without any shortage of soil moisture within the rooting zone”. PET, in other words, is mainly controlled by atmosphere demands. The concept of PET has since formed a fundamental component of many water balance, drought and vegetation growth studies. In many of these studies, the
ratio of actual ET to PET plays an essential part in the calculations. Estimates of PET are also
needed as an input to hydrologic rainfall/runoff models. Terrestrial water balance, water-
management, and net primary productivity models require an accurate estimate of PET for
reliable application [Amatya et al., 1995; Vörösmarty et al., 1998].

Actual evapotranspiration (AET) depends on climatic conditions and conditions of the
plants and soil surface. AET is a measure of the amount of water actually evapotranspirated
from the plants and soil system and takes into account the conditions of plant-soil system
(Ward and Elliot, 1995). Theoretically, ET is limited by the amount of energy available to
drive the change of states and the amount of water available to be evapotranspirated.
Nevertheless, ET is often referred to the amount of water that is evapotranspirated in reality.
In this study, ET has the same meaning as AET and AET is used explicitly to emphasize its
difference from PET.

The global climate change is likely to alter the magnitude and distribution of ET, thus,
the streamflow, and the plant-available soil water. Management practices in the relatively
natural ecosystems, such as forests, may also alter ET and thus soil water and streamflow
(Federer et al., 1996). In a biology study, Currie (1991) found that, in the four vertebrate
classes studied, 80-93% of the variability in species richness could be statistically explained
by a monotonically increasing function of a single variable, PET. And in contrast, tree
richness was more closely related to AET. Both AET and PET appeared to be measures of
available environmental energies.

Many studies have shown that the Penman-Monteith equation (Monteith, 1965) is the
most reliable method to predict PET under various climatic conditions when necessary
weather and vegetation data (net radiation, wind speed, dew point temperature, air
temperature and vegetation-specific parameters) are available. But it is difficult and expensive to obtain all of the required variables in order to use this equation. Therefore, many simpler methods that are based on either temperature or radiation are often used to predict PET. Most previous comparisons among PET methods have been local or applicable only to a specified cover type (Steiner et al., 1991; Crago and Brutsaert, 1992; Amatya et al., 1995). C.A. Federer et al. (1996) compared nine methods to calculate PET on a regional and global scale. Fennessey and Vogel (1996) developed a regional multivariate regression model of average monthly reference crop evapotranspiration for the northeastern U.S.. The model was shown to be an improvement over the Linacre method and the Hargreaves-Samani method. However, there are only a limited number of studies that have tested the differences between these methods for the climatic condition in the southern U.S., especially in forested watersheds.
2. OBJECTIVES

The objectives of this study are:

1. To develop a model to predict long-term annual AET for forested watersheds across the southern U.S..

2. To compare the differences between six PET methods applied in the southern U.S..
3. LITERATURE REVIEW – REGIONAL ET ESTIMATES

As mentioned previously, it is practically impossible to measure ET on the regional scale. Most of the methods predicting ET generate only point estimates, while information for regional areas is required. This can be overcome by repetitive measurements at all the representative units in the area of interest. However, in terms of effort and money involved in this, only models can be applied in such a distributed mode (Droogers, 2000). Presently, regional ET values must be estimated by models using standard meteorological network data as input (Bussieres et al., 1997). Many methods and models have been developed to predict regional ET, and they are reviewed.

3.1. Watershed water balance method

This method is widely used to estimate long-term annual ET. It was also often used to evaluate the accuracy of the methods or models that predict ET. The watershed water balance equation can be written as

\[ P = ET + Q + G + \Delta S \]  

Where \( P \) is precipitation, \( ET \) is actual evapotranspiration, \( Q \) is surface runoff measured as streamflow, \( G \) is recharge to groundwater (loss to deep groundwater), and \( \Delta S \) is the change in soil water storage. If a watershed is watertight and the topographic divide is coincident with the groundwater divide, then over a long period of time (i.e., 5-10 years), it is reasonable to assume that recharge to groundwater and changes in soil water storage are zero (Zhang, et al., 2001; Church, et al. 1995). Thus, \( ET \) can be calculated by the simple equation \( ET = P - R \) on a long-term annual basis.

This method gives estimates only over long periods of time. Thus, it requires long-term measured data, which are difficult to obtain in reality, especially in the rural areas.
3.2. Methods based on planetary boundary layer theory

There are many methods available for calculating local ET. Regional ET has mostly been determined by water balance analyses of the drainage basins. However, there have been few studies on determining regional ET on the basis of theoretical analysis of the various transfer mechanisms. Mawdsely and Brutsaert (1973) derived a method of calculating the AET on a regional basis by considering turbulent mass transfer in the boundary layer of the atmosphere. The planetary boundary layer is the region where dynamic and thermodynamic effects resulting from the earth’s surface are directly detectable. The approach they developed made several plausible assumptions and could be performed using only standard published meteorological data. The method was tested for a number of locations in the U.S. by comparing the results with pan evaporation data. The results showed some encouraging evidence of the soundness of the model. In another paper, Brutsaert and Mawdsely (1976) further developed the approach to calculate regional ET. Several possible formulations were derived for the calculation of actual surface water vapor flux on a regional basis by making use of recent developments in planetary boundary layer theory. Then the approach was applied in Omaha, Nebraska, and the results were also compared with ET data estimated by a suitable adjustment of PET from a nearby watershed in the Treynor basins. The method was reasonably successful on a monthly basis with a correlation coefficient of the order of 0.83 but much less so on a daily basis.

Abdulmumin et al. (1987) evaluated the planetary boundary layer theory to estimate regional ET as a component of an energy balance model using only regularly recorded solar radiation and upper air rawinsonde data. They developed new procedures to estimate regional
ET, and the results indicated that the new approach performed better than a comparable procedure presented by Brutsaert and Mawdsely (1976).

3.3. Methods based on remote sensing

With the advance of remote sensing techniques, they provide a reasonably viable means of evaluating ET over large areas. Remote sensing methods are attractive to estimate ET as they cover large areas and can provide estimates at a very high resolution. Intensive field monitoring is also not required, although some ground-truth measurements can be helpful in interpreting the satellite images (Kite and Droogers, 2000). Thus, more and more researchers are taking advantage of the remote sensing data to estimate regional ET (Zhang et al., 1995; Bussieres et al., 1997; Granger, 1997). But these methods have a disadvantage that only instantaneous estimates can be obtained. And for some areas, the requirement of cloud-free days could be a limitation for methods based on remote sensing data. There are three basic methods that are based on remote sensing data to estimate regional ET (Bussieres et al., 1997).

Method one is model input parameterization, which is the most widely used remote sensing technique to estimate ET. This method adds remote sensing information into the existing hydrological (Granger, 1995), ecological (Running et al., 1989) and other engineering-type models to estimate ET. All or part of the data required to calculate ET are derived from remote sensing data. The remote sensing data in the thermal infrared spectrum can be used to determine surface temperature and this information can be incorporated in various models for estimating ET (Zhang et al., 1995). Empirical relationships are used to convert remote sensing radiances to water vapor deficits, solar radiation and leaf area indices. Zhang et al. (1995) developed an one-layer resistance model for estimating regional ET by
Method two is the iterative energy-balance model inversion. An energy balance model is used by this method (Taconet et al., 1986) to estimate the surface temperature. The iterations are made for this model until the estimated temperature corresponds, within predetermined limits, to a temperature value estimated from remote sensing data. Then, the fluxes are calculated by inverting the model. Although this method may give a good representation of reality and could be integrated into models of the atmosphere, its application to field studies has been limited due to its complexity.

Method three is the residual energy budget method. Jackson et al. (1977) first introduced this method, which is simple and may possibly be implemented in an operational weather analysis system (Gellens et al., 1994). The latent heat flux (ET) is calculated as a residual of the energy budget equation, which was proposed to use remote sensing data. Reiniger and Seguin (1986) gave an intercomparison of the iterative energy balance model inversion (TERGRA, TELL-US, TELOP and SEAL models) and the residual energy budget methods. In their study, they agreed to use the residual energy budget method as an alternative solution to the complicated modeling. Bussieres et al. (1997) used the residual energy budget method to estimate daily ET under non-cloudy conditions in an area of southern Saskatchewan, Canada. The results were compared with other computations and
indicated that this method displayed patterns with an amount of detail that can’t be duplicated from current meteorological and other ground-based observation network.

3.4. Methods based on complementary relationship hypothesis

The hypothesis of a complementary relationship between AET and PET was first introduced by Bouchet (1963) and is briefly described as follows. Bouchet (1963) considered that there exists a complementary feedback mechanism between AET and PET over a large uniform surface of a regional size. Under the hypothesis, the PET would take place if only the available energy were the limiting factor. Under conditions when AET equals PET, this rate is noted as the Wet Environment Evapotranspiration (WET). And the general complementary relationship is then expressed as

\[ AET + PET = 2WET \]  

(2)

The complementary relationship hypothesis is essentially based on empirical observations, supported by a conceptual description of the underlying interactions between evapotranspiring surfaces and the atmospheric boundary layer. Bouchet (1963) hypothesized that when, under conditions of constant energy input to a given land surface-atmosphere system, water availability becomes limited, AET falls below its potential, and a certain amount of energy becomes available. This energy excess, in the form of sensible heat and/or long wave back radiation, increases the temperature and humidity gradients of the overpassing air and leads to an increase in PET equal in magnitude to the decrease in AET. If water availability is increased, the reverse process occurs, and AET increases as PET decreases. Thus WET ceases to be an independent causal factor, or climatologically constant forcing function, and instead is predicated upon the prevailing conditions of moisture availability (Hobbins et al., 2001a; Szilagyi, 2001).
The complementary relationship areal evapotranspiration (CRAE) model (Morton, 1983) and the advection-aridity (AA) model (Brustaert and Stricker 1979) are the most widely known models that are based on the complementary relationship hypothesis. Hobbins et al. (2001a) implemented the two models to estimate monthly ET for the conterminous U.S., and the model estimations were evaluated against independent estimates of regional ET calculated from long-term, large-scale water balance (1962-1988) for 120 minimally impacted watersheds in the conterminous U.S.. The results suggested that the advective portion of the AA model must be recalibrated before it may be used successfully on a regional basis and that the CRAE model accurately predicted monthly regional ET. In order to improve the AA model (Brustaert and Stricker 1979), Hobbins et al. (2001b) compared the AA model with independent estimates of ET derived from the long-term water balance for 139 undisturbed watersheds across the conterminous U.S.. The wind function and Priestley-Taylor coefficient in the AA model were modified, which lead to significant improvement in the performance of the model.

Strong (1997) pointed out that the methods based on the complementary relationship had no theoretical basis, but yielded reasonable results for longer periods simply because of the inevitable averaging effect of the atmosphere.

3.5. Hydrological model and empirical model estimates

Researchers try to use hydrological models to simulate ET. One advantage of the hydrological models is that they can estimate actual transpiration and evaporation separately, which are the two components of ET. Thus, it can predict how much water is in the useful process (crop transpiration) and how much water is in the wasteful process (soil evaporation, considered as a non-beneficial use of water in terms of food production). Also the
hydrological models can estimate ET on quite small time-scales such as hourly or daily ET values. But they require much more data as input to run the models.

SLURP (Semi-distributed Land Use-based Runoff Processes, Kite, 1997) is a conceptual model which includes a full hydrological cycle simulation as well as inclusion of man-made factors such as reservoirs, diversions and extractions and irrigation from surface and ground-water. The model originated in Canada in 1975 for use in mid-sized (thousand of square kilometers) watersheds. Subsequently, the model has been used in many countries for watersheds ranging in size from a few hectares to a few hundred square kilometers and to macroscale waterheds with areas of millions of square kilometers (Kite, 2000).

SWAP (Soil-Water-Atmosphere-Plant, Van Dam et al., 1997) is an integrated physically based simulation model for vertical transport of water, solute and heat in the saturated-unsaturated zone in relation to crop growth. Actual crop transpiration and soil evaporation may be simulated by taking into account the crop development stage as well as limitations in soil moisture.

Kite (2000) and Droogers (2000) applied SLURP and SWAP respectively to compute daily evaporation and transpiration for a variety of crops and other land covers within the Gediz Watershed in western Turkey.

In order to simplify the process of ET, empirical models were introduced. On a watershed scale, it is assumed that ET from land surfaces is controlled by water availability and atmospheric demand. The water availability can be approximated by precipitation, and the atmospheric demand can be approximated by PET. The basic relationships between runoff, precipitation and ET are postulated. Based on these ideas, Zhang et al. [2001] developed a simple two parameter (rainfall and PET) model to calculate mean annual ET on
the global scale. The model is a practical tool that can be readily used to predict the long-term consequences of reforestation and has potential uses in watershed-scale studies of land use changes.

3.6. Atmospheric moisture budget estimates

In view of the fact that previous atmospheric moisture budget estimates of regional ET have been restricted to the atmospheric boundary layer (McNaughton and Springgs, 1986; Munley and Hipps, 1991), and the result influenced the current study in its attempt to improve radio-sonde estimates of evapotranspiration through a “total” tropospheric moisture budget. Strong (1997) proposed an atmospheric moisture budget equation that includes horizontal advection explicitly and treats vertical fluxes implicitly through a total tropospheric moisture budget. The equation is derived through the continuity equation for the conservation of mass of a fluid. By vertically integrating the equation in thin layers of say, 5 hPa or less, one can assume that each layer is well mixed so that terms involving vertical gradient of vapor mass drop out. In this way, vertical advection is treated implicitly in the method, with vapor mass changes attributed only to surface vapor flux or horizontal advective changes. There is no net effect of vertical advection since the loss from one layer equals the gain to another. Because vertical advection is not treated explicitly, the equation is strictly valid only for a “total” tropospheric budget, and is not applicable to just the atmospheric boundary layer or any other specific layers (Strong, 1997).

The equation was applied over the Canadian Prairies and compared with other ET estimations (Strong, 1997). The results suggested that ET was more variable on a daily basis than other techniques have indicated. The results also indicated that while the advection estimates were a major source of error for the “daily” estimates in this particular study
(Regional Evaporation Study), it was shown that neither advection nor moisture flux through the boundary layer could be ignored in estimating daily ET, regardless of the technique used.
4. METHODOLOGY

The main purpose of this study is to develop a model to estimate regional AET in the southern U.S. for forested watersheds. From the literature, it is already known that PET is highly correlated with AET. Thus, PET is likely to be a good variable included in the model predicting AET. Many methods are available for estimating PET, but there is no knowledge of which one is better to apply in the southern U.S.. So another purpose of this study is to compare the differences between several PET methods. If PET is included in the model, it is hoped that the best PET method would be used. Hence, a database was developed to estimate AET and PET. All the data in the database were used to derive a model to estimate AET by using the stepwise selection method. Then six PET methods were compared.

4.1. Selections of six PET methods

Many methods have been developed to predict PET. However, limited availability of the required data is the biggest problem associated with making such estimates (Samani, 1986). For this study, in terms of the availability of the data and simplicity of the equations, six PET methods are selected to be compared. There are three temperature-based methods (the Thornthwaite, Hamon and Hargreaves-Samani methods) and three radiation-based methods (the Turc, Makkink and Priestley-Taylor methods).

4.1.1. Thornthwaite method

The Thornthwaite method was developed for coastal conditions in New Jersey. This method is probably the best known and most widely used technique for estimating PET in the U.S. and the western world. PET is calculated from the mean monthly temperature (in degrees Celsius), with corrections for daylight length determined by the month of the year.
and by the latitude (Parsons, 2000). One drawback of the Thornthwaite method is that it
gives zero PET values when the mean air temperature is less than 0°C.

The equation for computing this method is (Parsons, 2000; Jensen, 1973):

$$PET = 1.6L_d \left( \frac{10T}{I} \right)^a$$  \hspace{1cm} (3)

Where,

PET = Monthly PET in cm.

$L_d$ = Daytime length, it is time from sunrise to sunset in multiples of 12 hours.

$T$ = Monthly mean air temperature in °C.

$a = 6.75 \times 10^{-7} I^3 - 7.71 \times 10^{-5} I^2 + 0.01792 I + 0.49239$  \hspace{1cm} (4)

$I$ = Annual heat index, which is computed from the monthly heat indices.

$$I = \sum_{j=1}^{12} i_j$$  \hspace{1cm} (5)

Where

$i_j$ is computed as

$$i_j = \left( \frac{T_j}{5} \right)^{1.514}$$  \hspace{1cm} (6)

$T_j$ = Mean air temperature in °C for month $j$; $j = 1, \ldots, 12$. 
4.1.2. Hamon method

Hamon (1963) developed a simple equation for which PET does not become zero when the mean air temperature is less than 0°C but provides essentially the same annual total as that of the Thornthwaite method (Federer and Lash, 1983).

The PET is estimated as (Federer and Lash, 1983):

\[
PET = 0.1651 \times L_d \times RHOSAT
\]  

(7)

Where,

\(PET\) = Daily PET in mm.

\(L_d\) = Daytime length, which is time from sunrise to sunset in multiples of 12 hours.

\(RHOSAT\) = Saturated vapor density in g/m\(^3\) at the daily mean air temperature (TEMP).

\[
RHOSAT = 216.7 \times \frac{ESAT}{(TEMP + 273.3)}
\]  

(8)

\[
ESAT = 6.108 \times \exp\left(17.26939 \times \frac{TEMP}{TEMP + 237.3}\right)
\]  

(9)

\(TEMP\) = Daily mean air temperature in °C.

\(ESAT\) = Saturated vapor pressure in mb at the given TEMP.

The ESAT equation is from Murray (1967), allowing air temperatures to fall below 0°C.

In this study, the Hamon method was calibrated with the coefficient 1.2.

4.1.3. Turc method

Turc (1961) simplified earlier versions of a PET equation for 10-day periods under general climatic conditions of Western Europe. When expressed on a daily basis in mm/day, the equations are (Jensen et al., 1990):
For RH < 50%  

\[
PET = 0.013 \left( \frac{T}{T + 15} \right) (R_s + 50) \left( 1 + \frac{50 - RH}{70} \right)
\]  \hspace{1cm} (10)

For RH > 50%  

\[
PET = 0.013 \left( \frac{T}{T + 15} \right) (R_s + 50)
\]  \hspace{1cm} (11)

Where,

\begin{align*}
PET & = \text{Daily PET in mm/day}. \\
T & = \text{Daily mean air temperature in °C}. \\
R_s & = \text{Daily solar radiation in ly/day or cal.cm}^{-2}.d^{-1}. \\
\text{cal.cm}^{-2}.d^{-1} & = (100/4.1868) \text{MJ.m}^{-2}.\text{day}^{-1}. \\
RH & = \text{Daily mean relative humidity in percentage (%).}
\end{align*}

4.1.4. Priestley-Taylor method

The Priestley-Taylor PET method (Priestley and Taylor, 1972) is a relatively simple radiation-based model that has been successfully applied in many areas. This semi-empirical model was derived from the physics-based Penman-Monteith model (Monteith, 1965). The Penman-Monteith model estimates ET as a function of available energy, vapor-pressure deficit, air temperature, pressure, aerodynamic resistance (a function of primarily wind speed, and plant-canopy height and roughness), and canopy resistance (a measure of
resistance to vapor transport from plants). In the Priestley-Taylor model, the atmosphere is assumed to be saturated, in which case the aerodynamic term is zero. In reality, atmospheric saturation generally does not occur. Therefore, an empirical multiplier is applied (the Priestley-Taylor coefficient) as an empirical correction to account for the fact that the atmosphere does not generally attain saturation. The aerodynamic component was deleted and the energy component was multiplied by a coefficient, $\alpha = 1.26$, when the general surrounding areas were wet or under humid conditions. The coefficient 1.26 was applied in this study.

The PET is estimated as (Jensen et al., 1990):

$$\lambda PET = \alpha \frac{\Delta}{\Delta + \gamma} (R_n - G)$$

(12)

Where,

$\lambda$ = Latent heat of vaporization in MJ/kg.

$$\lambda = 2.501 - 0.002361 \ T$$

(13)

$T$ = Daily mean air temperature in °C.

$\alpha$ = Calibration constant, $\alpha = 1.26$ for wet or humid conditions.

$\Delta$ = Slope of the saturation vapor pressure-temperature curve in kPa/°C.

$$\Delta = 0.200 \ (0.00738 \ T + 0.8072)^7 - 0.000116$$

(14)
\( \gamma = \text{Psychrometric constant modified by the ratio of canopy resistance to atmospheric resistance in kPa/°C.} \)

\[
\gamma = \frac{c_p P}{0.622 \lambda}
\]  
(15)

Where,

\( c_p = \text{Specific heat of moist air at constant pressure in kJ.kg}^{-1}.\,^\circ C. \)

\( c_p = 1.013 \, \text{kJ.kg}^{-1}.\,^\circ C = 0.001013 \, \text{MJ.kg}^{-1}.\,^\circ C. \)

\( p = \text{Atmospheric pressure in kPa.} \)

\( p = 101.3 - 0.01055 \, \text{EL} \)  
(16)

\( \text{EL = Elevation in m.} \)

\( R_n = \text{Net radiation in MJ.m}^{-2}.\text{day}^{-1}. \)

\( G = \text{Heat flux density to the ground in MJ.m}^{-2}.\text{day}^{-1}. \)  
It can be estimated as (Jensen et al., 1990):

\[
G = 4.2 \frac{(T_{i+1} - T_{i-1})}{\Delta t} = -4.2 \frac{(T_{i-1} - T_{i+1})}{\Delta t}
\]  
(17)

where,

\( T_i = \text{Mean air temperature in °C for the period i.} \)

\( \Delta t = \text{The difference of time in days between two periods.} \)
4.1.5. Makkink method

Makkink (1957) estimated PET in mm/day over 10-day periods for grassed lands under cool climatic conditions of the Netherlands as (Xu et al., 2000; Jensen et al., 1990):

\[
PET = 0.61 \left( \frac{\Delta}{\Delta + \gamma} \right) \frac{R_s}{58.5} - 0.12
\]  

(18)

All variables in the equation have the same meaning and units as those in the Priestley-Taylor method.

4.1.6. Hargreaves-Samani method

The method was derived from eight years of cool season Alta fescue grass lysimeter data in Davis, California. Hargreaves and Samani (1982, 1985) proposed several equations for calculating PET, and the new one is (Jensen et al., 1990):

\[
\lambda PET = 0.0023 \times R_a \times TD^{0.5} \times (T + 17.8)
\]  

(19)

Where,

PET = Daily PET in mm/day.

\(\lambda\) = Latent heat of vaporization in MJ/kg.

T = Daily mean air temperature in °C.

\(R_a\) = Extraterrestrial solar radiation in MJ.m\(^{-2}\).day\(^{-1}\).

TD = Daily difference between the maximum and minimum air temperature in °C.
4.2. Database construction

4.2.1. Introduction

The collected data are based on watershed in this study, therefore, the AET and PET values represent watershed values. Some of the data were requested from the experimental forested watersheds, and some were obtained from USGS monitored watersheds. The experimental forested watersheds are small watersheds with areas ranging from 0.1 mi² to 12 mi². The USGS monitored watersheds are generally large watersheds with drainage areas ranging from 60 mi² to 3000 mi² (Table 4.1).

To calculate the long-term annual watershed water balance and estimate PET using the six PET methods, the following database was developed from different sources in this study.

1) Watershed location (latitude, longitude) and elevation.
2) Annual precipitation, annual streamflow of the watershed.
3) Monthly mean data for the watershed: temperature (T), relative humidity (RH), maximum temperature (T$_{\text{max}}$), minimum temperature (T$_{\text{min}}$), solar radiation (R$_{s}$), extraterrestrial solar radiation (R$_{a}$) and net radiation (R$_{n}$).
4) Percentage of each land cover, including deciduous forests, conifer forests, water body, crop-grass and others in the watershed.

The time scale used in the watershed water balance calculation is annual, and the time scale in the PET calculations is monthly basis. So monthly is the minimum time scale applied in this study.

Most weather stations don’t measure solar radiation data, and freely downloadable monthly solar radiation data are available only from 1961 to 1990 (Liang, 2001). Therefore,
the time period was set to 1961-1990 in this study, except for the Carteret watershed, which has the measured solar radiation data in the watershed.

Measured net radiation ($R_n$) is seldom available for the area of interest, hence, an empirical equation was calibrated to estimate net radiation in this study. The procedures to build the database are explored in detail in the following sections.

4.2.2. Data collection

4.2.2.1. Data of experimental forested watersheds

There are six experimental forested watersheds available in this study. The data were requested from those watersheds. In the case when the required data were not available from the experimental watersheds, the data were referenced to nearby stations. The temperature data were referenced to climate stations, and relative humidity and solar radiation were referenced to solar radiation stations (Table 7.3).

1) Bradford watershed

The Bradford watershed in Florida (Sun et al., 2000a) has monthly precipitation, steamflow and mean temperature data. The monthly maximum temperature ($T_{max}$), minimum temperature ($T_{min}$), relative humidity (RH), solar radiation ($R_s$), and extraterrestrial solar radiation ($R_a$) were acquired from nearby stations.

Provided by National Oceanic and Atmospheric Administration’s (NOAA) National Climatic Data Center (NCDC), monthly maximum temperature ($T_{max}$) and minimum temperature ($T_{min}$) were downloaded from the web site (http://lwf.ncdc.noaa.gov/oa/ncdc.html).

Supported by the National Center for Photovoltaics (NCPV) and managed by the Department of Energy’s Office of Energy Efficiency and Renewable Energy, the monthly
relative humidity (RH), solar radiation (R_s) and extraterrestrial solar radiation (R_a) were obtained from nearby stations at the web site of Renewable Resource Data Center (RReDC, http://rredc.nrel.gov/solar/old_data/nsrdb/dsf/).

When particular data were not available, similar procedures were used for other watersheds to reference data to nearby stations.

2) Carteret and Parker tract watershed

The data of the Carteret and Parker tract watersheds in the coastal of North Carolina (Amatya et al., 1996; Amatya and Skaggs, 2001; Amatya et al., 2002) has monthly precipitation, streamflow and mean temperature, relative humidity, solar radiation and net radiation. The monthly maximum temperature and monthly minimum temperature, and extraterrestrial solar radiation (R_a) were acquired from nearby weather stations (Table 7.3).

3) Walker Branch watershed

The annual precipitation and streamflow data of Walker Branch watershed (Johnson and Hook, 1989) in Tennessee were requested from Oak Ridge National Laboratory (Patrick J. Mulholland, personal communication). Monthly mean temperature, monthly maximum temperature and monthly minimum temperature were obtained from the web site (http://www.esd.ornl.gov/programs/WBW/CLIMWBWA.HTM). The RH, R_s, R_a were acquired from the nearby stations (Table 7.3).

4) Coles Fork watershed

The precipitation and streamflow data of the Coles Fork watershed at the University of Kentucky’s Robinson Experimental Forest were available (Arthur et al., 1998; Randy Kolka, personal communication). The RH, R_s, R_a data were downloaded from the nearby weather stations.
5) Santee 80 watershed

The Santee 80 watershed is part of the Santee Experimental Forest watershed (Richter and Ralston, 1983; Sun et al., 2000b) located in the Francis Marion National Forest on the coastal flatwoods of South Carolina. Precipitation, streamflow and air temperature data were available from the watershed, but the unavailable data (RH, Rs, Ra) were acquired from the nearby stations (Table 7.3).

4.2.2.2. Data of USGS monitored watersheds

1) Real boundary watersheds

The twelve USGS monitored watersheds in North Carolina (Sun et al., 2000c) have real watershed boundaries. The nearby climate stations were referenced from NCDC, which have precipitation, temperature, maximum and minimum temperature data; RH, Rs, Ra data were obtained from nearby stations in RReDC.

2) Eight-digit hydrologic unit codes boundary watersheds

The seventeen USGS monitored watersheds across the southeastern U.S. used by Liang et al. (2002) do not have actual watershed boundaries. Basically, the watersheds selected satisfy three conditions: 1) Forest dominant coverage (either conifer or deciduous forest dominant). 2) Have long-term (1961-1990) streamflow gauging stations. 3) Have long-term weather stations nearby. Then solar radiation data were referenced from nearby stations. The procedures to obtain data of the seventeen watersheds were similar to those for twelve USGS monitored watersheds in North Carolina. The difference was that the real boundaries of the seventeen watersheds were unknown. Therefore, the actual land use compositions of each of the watersheds were derived from MRLC (Multi-Resolution Land Characteristics) data according to the eight-digit hydrologic unit codes (HUC) boundaries.
3) Additional four eight-digit HUC boundary watersheds in South Carolina and Georgia

Additionally, four eight-digit HUC boundary watersheds in South Carolina and Georgia were selected to give better representation of the southern U.S.. Similar to the previous seventeen USGU monitored watersheds, the real boundaries of the watersheds were unknown when the gauging stations were obtained from USGS, the eight-digit HUC boundaries were used to approximately represent the boundaries of the watersheds. First, the solar radiation stations were obtained to select the watersheds (eight-digit HUC), then the watersheds were further selected by the conditions that long-term gauging stations, long-term weather stations were available in the watersheds. Finally the watersheds were verified to be forest dominant coverage. Since all the stations were located in one eight-digit HUC, they could be used to represent the condition of that watershed. Methodology of the watershed selection is described in detail in appendix A.

4.2.2.3. Obtain watershed elevation and land cover

1) Land cover and elevation of experimental forested watersheds

The six experimental forested watersheds (site number34-39 in Table 4.1) were 100% forest covered. The land cover and elevation data were acquired from the experimental forested watersheds (Table 4.1).

2) Land cover and elevation of USGS monitored watersheds

Land cover data for the thirty-three USGS watersheds were derived from the National Land Cover Dataset (NLCD) remote sensing data (http://edc.usgs.gov/glis/hyper/guide/mrlc4), which was developed by the Multi-Resolution Land Characteristics (MRLC) project and is in the form of a conterminous U.S. grid. The
U.S. land cover grid serves for 1992 and is in 30-meter resolution. The land use was assumed to be similar during 1961-1992. The USGS watershed land cover data were obtained by clipping the U.S. land cover grid with the watershed boundaries. Also, a 4-km-resolution U.S. contiguous DEM from USGS (http://edcwww.cr.usgs.gov/glis/hyper/guide/usgs_dem) was available. The DEM was clipped by the USGS watershed boundary to obtain the elevation of the USGS watersheds.

In the database, some USGS watersheds don’t have real boundaries. Appendix B lists the detailed procedures to delineate the boundary of a watershed in Georgia in order to get accurate land cover compositions in the watershed, appendix B also lists detailed procedures to derive the land cover and elevation in the watershed. Since it takes too much effort to delineate watershed boundaries, and there are more than 20 watersheds in the database that need to be delineated to meet the objective of obtaining all real watershed boundaries, an alternative solution was applied in this study. For the watersheds which real boundaries were not available, the eight-digit HUC boundaries were used to represent the boundaries of the watersheds.

After the watershed boundaries were available, they were used to clip the land cover grid and DEM to get the land cover and elevation of the watersheds (Table 4.1). The clipping procedures are similar to the step 25-36 in appendix B.

4.2.2.4. Deriving net radiation

Net radiation is a critical variable in the Priestley-Taylor method to calculate PET, but it is rarely observed. So an empirical equation (Castellvi et al, 2001) is applied in this study to calculate monthly net radiation.

\[
R_n = 0.77R_s - 2.45 \times 10^{-9} \times f \times \left(0.261 \times \exp\left(-7.7710 \times 10^{-4} T^2\right) - 0.02\right) T_{ks}^4 + T_{kn}^4
\] (20)
\[ f = \left(1.2 \times \frac{R_s}{R_a} + 0.1\right) \]  

(21)

Where,

\( R_n \) = Monthly mean net radiation in MJ.m\(^{-2}\).day\(^{-1}\).

\( R_s \) = Monthly mean solar radiation in MJ.m\(^{-2}\).day\(^{-1}\).

\( R_a \) = Monthly mean extraterrestrial solar radiation in MJ.m\(^{-2}\).day\(^{-1}\).

\( T \) = Monthly mean temperature in (K).

\( T_{km} \) = Monthly mean maximum temperature in (K).

\( T_{kn} \) = Monthly mean minimum temperature in (K).

The Carteret watershed has thirteen years of measured monthly mean net radiation data, thus, the equation was applied to estimate net radiation in Carteret watershed. The results were compared with the measured values. The average difference between measured and estimated net radiation was 0.83 MJ.m\(^{-2}\).day\(^{-1}\). The equation was modified to add the difference. Then the equation was applied to estimate monthly mean net radiation in other watersheds, assuming this correlation is valid for all other locations.

\[ R_n = 0.77R_s - 2.45 \times 10^{-9} f \times \left(0.261 \times \exp\left(-7.7710 \times 10^{-4} T^2\right) - 0.02(T_{km}^4 + T_{kn}^4)\right) + 0.83 \]  

(22)

4.2.2.5. Unit conversions

The units of data used in the PET methods are different from the units of data that are available. So unit conversions were necessary in this study.

Solar radiation and extraterrestrial solar radiation are in the units of W-hr/m\(^2\) for monthly averages of the daily values when they are downloaded on line. They were converted to MJ/m\(^2\) by the coefficient of 0.086 \(\times\) 4.1868/100. The relative humidity was converted to a percentage. The temperatures are in the unit of Celsius (C) degrees in PET
calculations, so conversion \( C = \left(\frac{5}{9}\right) F + 32 \) is necessary when the available data are in Fahrenheit (F) degrees. Also, the precipitation and streamflow are converted to mm/year when the units are not the same.

### 4.2.2.6. Data summaries

Figure 4.1 is a GIS map showing the locations of the study watersheds. The dots indicate the locations of the watersheds. More watersheds were located in North Carolina than other states, but they distribute nicely in different physiographic regions – coastal, piedmont and mountain areas. They are sufficient to represent the conditions of the southern U.S. and are included in the database for analysis. In total, there were thirty-nine watersheds obtained for this study (Table 4.1)

Appendix C is the summary of gauging stations of precipitation, streamflow and temperature, solar radiation and relative humidity of the watersheds, associated with the watershed drainage area and record period.
Figure 4.1. Study watersheds for analysis
Table 4.1. Summary of study watersheds

<table>
<thead>
<tr>
<th>Site number</th>
<th>Site name</th>
<th>Drainage area (mi²)</th>
<th>Elevation (m)</th>
<th>Mean annual rainfall (mm)</th>
<th>Mean annual runoff (mm)</th>
<th>Runoff/Rainfall</th>
<th>Mean daily Temperature (°C)</th>
<th>Land cover percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Trent R, NC</td>
<td>168</td>
<td>30</td>
<td>1320.87</td>
<td>398.06</td>
<td>0.30</td>
<td>11.73</td>
<td>34</td>
</tr>
<tr>
<td>2</td>
<td>Pocetas, NC</td>
<td>225</td>
<td>24</td>
<td>1150.76</td>
<td>349.87</td>
<td>0.30</td>
<td>14.17</td>
<td>43.2</td>
</tr>
<tr>
<td>3</td>
<td>Little, NC</td>
<td>177</td>
<td>47</td>
<td>1122.8</td>
<td>324.39</td>
<td>0.29</td>
<td>15.43</td>
<td>40.9</td>
</tr>
<tr>
<td>4</td>
<td>Eno, NC</td>
<td>141</td>
<td>193</td>
<td>1213.38</td>
<td>317.09</td>
<td>0.26</td>
<td>14.64</td>
<td>51.8</td>
</tr>
<tr>
<td>5</td>
<td>FLATR, NC</td>
<td>149</td>
<td>135</td>
<td>1122.01</td>
<td>330.64</td>
<td>0.29</td>
<td>14.27</td>
<td>54.7</td>
</tr>
<tr>
<td>6</td>
<td>Drown, NC</td>
<td>183</td>
<td>149</td>
<td>1182.68</td>
<td>479.51</td>
<td>0.41</td>
<td>12.32</td>
<td>39.2</td>
</tr>
<tr>
<td>7</td>
<td>Hunting, NC</td>
<td>155</td>
<td>322</td>
<td>1187.96</td>
<td>475.81</td>
<td>0.40</td>
<td>15.24</td>
<td>48.5</td>
</tr>
<tr>
<td>8</td>
<td>Fisher, NC</td>
<td>128</td>
<td>322</td>
<td>1158.71</td>
<td>502.17</td>
<td>0.43</td>
<td>14.56</td>
<td>51.1</td>
</tr>
<tr>
<td>9</td>
<td>New River, NC</td>
<td>205</td>
<td>955</td>
<td>1441.34</td>
<td>754.63</td>
<td>0.52</td>
<td>15.65</td>
<td>52.7</td>
</tr>
<tr>
<td>10</td>
<td>French, NC</td>
<td>296</td>
<td>999</td>
<td>1931.19</td>
<td>1239.8</td>
<td>0.64</td>
<td>12.52</td>
<td>55.2</td>
</tr>
<tr>
<td>11</td>
<td>LiTenne, NC</td>
<td>140</td>
<td>897</td>
<td>1825.07</td>
<td>970.96</td>
<td>0.53</td>
<td>10.05</td>
<td>72.5</td>
</tr>
<tr>
<td>12</td>
<td>Ocunalu, NC</td>
<td>184</td>
<td>650</td>
<td>1453.39</td>
<td>979.08</td>
<td>0.67</td>
<td>13.81</td>
<td>72.5</td>
</tr>
<tr>
<td>13</td>
<td>al03140303</td>
<td>176</td>
<td>109</td>
<td>1628.17</td>
<td>515.62</td>
<td>0.32</td>
<td>18.9</td>
<td>31.75</td>
</tr>
<tr>
<td>14</td>
<td>al03150203</td>
<td>97.5</td>
<td>70</td>
<td>1485.87</td>
<td>486.49</td>
<td>0.33</td>
<td>17.9</td>
<td>40</td>
</tr>
<tr>
<td>15</td>
<td>ar11010001</td>
<td>400</td>
<td>432</td>
<td>1123.72</td>
<td>480.41</td>
<td>0.43</td>
<td>13.97</td>
<td>53.1</td>
</tr>
<tr>
<td>16</td>
<td>fl03120003</td>
<td>102</td>
<td>41</td>
<td>1665.13</td>
<td>636.55</td>
<td>0.38</td>
<td>19.12</td>
<td>38.1</td>
</tr>
<tr>
<td>17</td>
<td>ky05070203</td>
<td>206</td>
<td>312</td>
<td>1076.09</td>
<td>412.39</td>
<td>0.38</td>
<td>11.77</td>
<td>94.4</td>
</tr>
<tr>
<td>18</td>
<td>ky05100203</td>
<td>722</td>
<td>358</td>
<td>1234.59</td>
<td>520.69</td>
<td>0.42</td>
<td>12.55</td>
<td>93.4</td>
</tr>
<tr>
<td>19</td>
<td>la08070202</td>
<td>145</td>
<td>56</td>
<td>1617.12</td>
<td>575.09</td>
<td>0.36</td>
<td>18.64</td>
<td>28.2</td>
</tr>
<tr>
<td>20</td>
<td>ms03170002</td>
<td>918</td>
<td>99</td>
<td>1426.59</td>
<td>504.67</td>
<td>0.35</td>
<td>17.37</td>
<td>36.4</td>
</tr>
<tr>
<td>21</td>
<td>ms03170009</td>
<td>96.1</td>
<td>21</td>
<td>1553.68</td>
<td>718.42</td>
<td>0.46</td>
<td>18.66</td>
<td>17.3</td>
</tr>
<tr>
<td>22</td>
<td>ms03180002</td>
<td>3171</td>
<td>110</td>
<td>1458.31</td>
<td>505.6</td>
<td>0.35</td>
<td>17.82</td>
<td>36.8</td>
</tr>
<tr>
<td>23</td>
<td>ms08060203</td>
<td>654</td>
<td>83</td>
<td>1337.63</td>
<td>478.44</td>
<td>0.36</td>
<td>18.64</td>
<td>43.9</td>
</tr>
<tr>
<td>24</td>
<td>tn06010204</td>
<td>1987</td>
<td>576</td>
<td>1516.86</td>
<td>832.43</td>
<td>0.55</td>
<td>13.55</td>
<td>50.3</td>
</tr>
<tr>
<td>25</td>
<td>tn06040004</td>
<td>447</td>
<td>238</td>
<td>1485.21</td>
<td>615.87</td>
<td>0.41</td>
<td>13.93</td>
<td>72.8</td>
</tr>
<tr>
<td>26</td>
<td>tx12030201</td>
<td>142</td>
<td>112</td>
<td>1051.4</td>
<td>200.18</td>
<td>0.19</td>
<td>18.63</td>
<td>39.1</td>
</tr>
<tr>
<td>27</td>
<td>tx12040103</td>
<td>325</td>
<td>62</td>
<td>1263</td>
<td>259.07</td>
<td>0.21</td>
<td>20.32</td>
<td>32.5</td>
</tr>
<tr>
<td>28</td>
<td>va02080201</td>
<td>329</td>
<td>633</td>
<td>1068.85</td>
<td>417.98</td>
<td>0.39</td>
<td>10.93</td>
<td>74.1</td>
</tr>
<tr>
<td>29</td>
<td>va05050002</td>
<td>223</td>
<td>760</td>
<td>1052.74</td>
<td>480.11</td>
<td>0.46</td>
<td>10.42</td>
<td>67.8</td>
</tr>
</tbody>
</table>
Table 4.1 (continued)

<table>
<thead>
<tr>
<th>Site number</th>
<th>Site name</th>
<th>Drainage area (mi²)</th>
<th>Elevation (m)</th>
<th>Mean annual rainfall (mm)</th>
<th>Mean annual runoff (mm)</th>
<th>Runoff/Rainfall</th>
<th>Mean daily Temp (°C)</th>
<th>Deciduous</th>
<th>Conifer</th>
<th>Water</th>
<th>Crop-Grass</th>
<th>Others</th>
<th>Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>ga03130005</td>
<td>272</td>
<td>213</td>
<td>1306.23</td>
<td>469.19</td>
<td>0.36</td>
<td>16.21</td>
<td>45.7</td>
<td>29.2</td>
<td>0.9</td>
<td>18.2</td>
<td>6.1</td>
<td>74.9</td>
</tr>
<tr>
<td>31</td>
<td>ga03070103</td>
<td>182</td>
<td>205</td>
<td>1133.9</td>
<td>363.9</td>
<td>0.32</td>
<td>18.12</td>
<td>42.7</td>
<td>29.8</td>
<td>1.3</td>
<td>14.9</td>
<td>11.3</td>
<td>72.5</td>
</tr>
<tr>
<td>32</td>
<td>ga03070101</td>
<td>392</td>
<td>270</td>
<td>1263.43</td>
<td>493.35</td>
<td>0.39</td>
<td>16.47</td>
<td>51.4</td>
<td>15.5</td>
<td>0.3</td>
<td>30</td>
<td>3</td>
<td>66.8</td>
</tr>
<tr>
<td>33</td>
<td>sc03050110</td>
<td>59.6</td>
<td>72</td>
<td>1196.88</td>
<td>435.39</td>
<td>0.36</td>
<td>18.39</td>
<td>34.7</td>
<td>32.3</td>
<td>1.6</td>
<td>18.5</td>
<td>13</td>
<td>66.9</td>
</tr>
<tr>
<td>34</td>
<td>Bradford, FL</td>
<td>0.54</td>
<td>44</td>
<td>1241.4</td>
<td>226.05</td>
<td>0.18</td>
<td>20.88</td>
<td>30</td>
<td>70</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>35</td>
<td>Walker Branch, TN</td>
<td>0.39</td>
<td>308</td>
<td>1331.41</td>
<td>659.77</td>
<td>0.50</td>
<td>13.88</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>36</td>
<td>Coles Fork,KY</td>
<td>6.41</td>
<td>378</td>
<td>1155.2</td>
<td>376.97</td>
<td>0.33</td>
<td>11.65</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>37</td>
<td>Carteret, NC</td>
<td>0.1</td>
<td>3</td>
<td>1538.67</td>
<td>519.68</td>
<td>0.34</td>
<td>16.29</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>38</td>
<td>Parker,NC</td>
<td>11.39</td>
<td>6</td>
<td>1248.97</td>
<td>287.98</td>
<td>0.23</td>
<td>15.03</td>
<td>0</td>
<td>100</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>39</td>
<td>Santee-80, NC</td>
<td>0.58</td>
<td>7</td>
<td>1381.56</td>
<td>246.16</td>
<td>0.18</td>
<td>18.13</td>
<td>50</td>
<td>50</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>
4.3. AET and PET estimates

4.3.1. AET estimates

Using the watershed water balance method, which was discussed in the previous section, the AET was calculated by a simple equation on a long-term annual basis:

\[ \text{AET} = \text{Rainfall} - \text{Runoff} \]  

(23)

4.3.2. PET estimates

A Visual Basic program was coded to assist PET calculations and data summaries. The source codes of the program are in appendix D. The interface of the program is showed in Figure 4.2. The program was used to calculate monthly PET by six methods and summarize data for annual and long-term annual averages for each watershed. The formats of the input and output files of the program are in appendix D and appendix E respectively.
Since many PET methods calculate monthly PET on an average daily basis and the minimum time scale of the data is monthly, the average daily PET values were multiplied by the number of days in a particular month to obtain the monthly PET values. The annual values were calculated by adding up twelve months of the year. By running the program, PET by six different methods can be calculated as monthly values and annual values.

Further, the annual data were averaged to get long-term annual values. Rainfall, runoff, AET and PET are in long-term mean annual values as mm/year; the temperature,
relative humidity and solar radiation data are averaged in long-term annual daily values. The average values are based on the available range of the data in each individual watershed and are not based on the same time period. The longest period is thirty years, and the shortest average is five years.

### 4.4. Data analysis

Statistical regression model development and diagnosis in this study were performed according to the regression analysis methods described by Rawlings et al. (1998).

#### 4.4.1. Model building methods

For modeling AET, it was assumed that a linear relationship exists between the dependent variable (AET) and the independent variables (rainfall, temperature, RH, etc.). Multiple linear regression analysis was used to fit one line to the observed data. So the goal is to develop a model to predict AET in the term of least squares.

The land cover data (Table 4.1) were joined with the long-term annual summary data and input into SAS. The SAS 8.2 was used as a tool to derive the model using the stepwise selection and R square methods.

The stepwise selection is one of the regression methods, which have been developed to identify good (although not necessarily the best) subset models, with considerably less computing than is required for all possible regressions. The stepwise selection method chooses the subset models by adding one variable at a time to the previously chosen subset. It starts by choosing one variable as the independent variable. This will be the variable having the highest simple correlation with the dependent variable. At each successive step, the variable in the subset of variables not already in the model that causes the largest decrease in the residual sum of squares is added to the subset. In the mean time, it checks the variables
that are already in the model, if the partial sums of squares for any previously included variables do not meet a minimum criterion to stay in the model, the selection procedure drops one variable at a time until all remaining variables meet the minimum criterion. Then it continues to select the next variable that meets the criterion to enter the model and checks the variables in the model to make sure they meet the minimum criterion to stay. Thus, the stepwise selection method has two criteria, one is the criterion for variables to enter the model, and the other is the criterion for variables to stay in the model. The variable selection process terminates when all variables in the model meet the criterion to stay and no variables outside the model meet the criterion to enter.

The RSQUARE option in PROC REG (SAS statement) was also used to assist the model building. This method can calculate all possible regressions, but the best m subset regressions can be showed by specifying “BEST = m” option in the procedure. In this study, m was set equal to five. For each particular number of variables in the model, the five best subset regressions are selected by $R^2$, which is defined as the ratio of the regression sum of squares to the total sum of squares and used as a standard to measure the amount of the dependent variable variation associated with the independent variables. There are thirty-nine observations (sample size = 39) available for analysis, thus, the effort was to develop a model although the model validation would be preferred if more data were available.

4.4.2. Regression diagnostics

Regression diagnostics refers to the general class of techniques for detecting problems in regression – problems with either the model or the data set. Usually, regression diagnostics is performed to check for the least squares assumptions of normality, common variance and independence of errors. Normality in residual or error distribution is a required
property for testing significance and constructing confidence interval estimates of the parameters. Common variance in least squares procedures increases the precision in the parameter estimation compared to the precision that would have been obtained with heterogeneous variances. Finally, independence of residuals or errors, like common variance, allows unbiased estimation of regression parameters (Calvo-Alvarado and Gregory, 1997). In this study, the regression diagnostics are composed of residual analysis, influence statistics and collinearity diagnostics.

Residual analysis or analysis of some transformation of the residuals is very useful for detecting inadequacies in the model or problems in the data. In this study, graphical techniques and inspection of the residuals were performed to serve the purpose of verifying the least squares assumption and detecting the presence of outliers. In this study, an outlier is defined as a data point for which the value of the dependent variable is inconsistent with the rest of the sample. Whenever an outlier is detected, it should not be dropped from the data set unless it is found to be in error and the error cannot be corrected.

The purpose of influence statistics is to detect influential points, which have negative effects on the regression results and can’t be detected by residual analysis. Observations that are farthest from the mean values have the greatest impact on regression results. And these observations are often referred to as potentially influential points, the data points that are on the fringes of the cloud of the sample points in X-space (the space of independent variables). In this study, Cook’s D was used to verify whether a potentially influential point was really an influential point. Cook’s D is designed to measure the shift in parameter values when a particular observation is omitted. It is a combined measure of the impact of that observation on all regression coefficients. If the shift is big, that observation is considered as an
influential point. Once an influential point was found in the observations, the data accuracy needed to be checked.

Collinearity diagnostics serves to detect collinearity among the independent variables in the regression model. The presence of collinearity implies that there are (near) redundancies among the independent variables, that is, the same information is being provided in more than one way. The impact of the collinearity on least squares is very serious if primary interest is in the regression coefficients per se or if the purpose is to identify “important” variables in the process, which is exactly the objective of this study. The estimates of the regression coefficients can differ greatly from the parameters they are estimating, even to the point of having incorrect signs. The collinearity will allow “important” variables to be replaced in the model with incidental variables that are involved in the near-singularity. Hence, the regression analysis provides little indication of the relative importance of the independent variables. Condition index was used to detect collinearity in this study.

4.4.3. PET comparisons

In order to compare the differences among PET values estimated by six methods, SAS multivariate statistical analysis was performed. The six PET estimates were considered as six repeated measurements on PET in the southern U.S., and the normal assumptions of repeated measurements were applied.

In order to find out which PET calculation method has better performance, two assumptions are made in this study. One is that PET is higher than AET on a long-term annual basis, which is true in most cases although the PET is referenced to grass and AET is from the forested area. Another assumption is that a linear relationship exists between PET
and AET, thus, the PET estimate that has the higher correlation with AET is considered to perform better. The second assumption is reasonable since one of the purposes of introducing PET is to estimate AET by relating it to PET.
5. RESULTS

5.1. Fitting the long-term annual AET regression model

5.1.1. Model building

In this study, the usual assumptions of ordinary least squares were made, and all of the variables were assumed to be linearly related to the dependent variable, AET. There were totally twenty-three variables (Table 5.1 and Appendix F) available for model building. The SAS program used for model building and regression diagnostics is in appendix F.

Table 5.1. Variables used in the AET model building

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rainfall</td>
<td>Long-term mean annual precipitation, mm</td>
</tr>
<tr>
<td>Temp</td>
<td>Long-term mean daily temperature, °C</td>
</tr>
<tr>
<td>( R_s )</td>
<td>Long-term mean daily solar radiation, MJ.m(^{-2}).day(^{-1})</td>
</tr>
<tr>
<td>RH</td>
<td>Long-term mean daily relative humidity, percentage</td>
</tr>
<tr>
<td>( R_n )</td>
<td>Long-term mean daily net radiation, MJ.m(^{-2}).day(^{-1})</td>
</tr>
<tr>
<td>( R_a )</td>
<td>Long-term mean daily extraterrestrial solar radiation, MJ.m(^{-2}).day(^{-1})</td>
</tr>
<tr>
<td>( T_{\text{max}} )</td>
<td>Long-term mean daily maximum temperature, °C</td>
</tr>
<tr>
<td>( T_{\text{min}} )</td>
<td>Long-term mean daily minimum temperature, °C</td>
</tr>
<tr>
<td>Thorn</td>
<td>Long-term mean annual PET estimated by the Thornthwaite method, mm</td>
</tr>
<tr>
<td>Hamon</td>
<td>Long-term mean annual PET estimated by the Hamon method, mm</td>
</tr>
<tr>
<td>Turc</td>
<td>Long-term mean annual PET estimated by the Turc method, mm</td>
</tr>
<tr>
<td>PT</td>
<td>Long-term mean annual PET estimated by the Priestley-Taylor method, mm</td>
</tr>
<tr>
<td>Makk</td>
<td>Long-term mean annual PET estimated by the Makkink method, mm</td>
</tr>
<tr>
<td>HS</td>
<td>Long-term mean annual PET estimated by the Hargreaves-Samani method, mm</td>
</tr>
<tr>
<td>Latitude</td>
<td>Watershed latitude by the stream gauging station, degree</td>
</tr>
<tr>
<td>Longitude</td>
<td>Watershed longitude by the stream gauging station, degree</td>
</tr>
<tr>
<td>Elevation</td>
<td>Mean watershed elevation, m</td>
</tr>
<tr>
<td>Deciduous</td>
<td>Long-term percentage of watershed covered by deciduous forests</td>
</tr>
<tr>
<td>Conifer</td>
<td>Long-term percentage of watershed covered by conifer forests</td>
</tr>
<tr>
<td>Water</td>
<td>Long-term percentage of watershed covered by the water body</td>
</tr>
<tr>
<td>CropGrass</td>
<td>Long-term percentage of watershed covered by the crop or the grass</td>
</tr>
<tr>
<td>Others</td>
<td>Long-term percentage of watershed covered by others</td>
</tr>
<tr>
<td>Forest</td>
<td>Long-term percentage of watershed covered by forests</td>
</tr>
</tbody>
</table>
First, all the thirty-nine observations were used to build the model and slstay (significant level to stay) in the stepwise selection was set at 0.05 as preliminary data analysis. The stepwise selection method picked up the model including Rainfall, Latitude, Elevation and Water as independent variables. After running regression diagnostics (using Rstudent and Cook’s D), one observation was found to be an outlier (site number 12) and one observation was found to be an influential point (site number 21).

The observations were checked for data accuracy (Table 4.1). The influential point (site number 21) has 30.3% land cover as water body. However, the rest of the observations have only an average of 0.74% water body as land cover and the watershed that has second largest water body as land cover is 3.2%. Thus, the watershed of site number 21 is much different from the other watersheds. It was not considered to be a forested watershed. Therefore, it was eliminated from the data set since this study focuses on forested areas. The outlier (site number 12) is located in North Carolina and very close to the Little Tennessee River watershed (site number 11), which is near the U.S. Forest Service Coweeta Hydrologic laboratory. The long-term mean rainfall in Little Tennessee River watershed is 1825mm. The outlier only has a long-term annual rainfall of 1453mm, which is too low compared with the nearby watershed. The Runoff/Rainfall ratio in the watershed is equal to 67%. It was also found out that another observation (site number 10) has extremely high long-term mean annual rainfall and streamflow values, which are 1931mm and 1240mm, respectively. Its Runoff/Rainfall ratio is equal to 64%. However, it was not considered as either an outlier or an influential point by regression diagnostics. The average Runoff/Rainfall ratio of all the thirty-nine watersheds in the database is 37%. AET is more than 70% of the annual precipitation for the entire U.S. (Gay, 1993; Brooks et al., 1997). Due to high Runoff/Rainfall
ratio compared to the rest of observations, some data problems were suspected in the two watersheds. The two watersheds (site number 10 and site number 12) are located in the mountain area with average elevations of 999m and 650m, respectively. The rainfall data were recorded outside of the watersheds as acquired from nearby stations. In the mountain areas, the nearby stations might have different rainfall values from the watersheds. Thus, these two observations were removed from the data set for the next analysis. Some watersheds have low Runoff/Rainfall ratios. For example, the Bradford watershed (site number 30) and the Santee watershed (site number 39) had ratios of 18%. But the rainfall data were recorded in the experimental watersheds and expected to be correct.

So there were thirty-six observations available for final model building and diagnostics. The same procedures of SAS were used to process the data. But the slstay (significant level to stay) in the stepwise selection was adjusted to 0.25.

The SAS program was executed again with thirty-six observations. After running the program, the variables selected by the stepwise selection method are Rainfall, Latitude, Elevation, Conifer and Water as land covers with $R^2 = 0.85$ and adjusted $R^2 = 0.82$. Checking with the RSQUARE method, it indicates that the stepwise selection method selected a model with the best five variables subset. The model also has the lowest Mallows’ $C_p$ value (Table 5.2).
Table 5.2. Summary of the stepwise selection method for the AET model building

<table>
<thead>
<tr>
<th>Step</th>
<th>Variable entered</th>
<th>Variable removed</th>
<th>Variable numbers</th>
<th>Partial $R^2$</th>
<th>Model $R^2$</th>
<th>Mallows’ $C_p$</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$R_a$</td>
<td>1</td>
<td>0.55</td>
<td>0.55</td>
<td>45.71</td>
<td>42.25</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>RH</td>
<td>2</td>
<td>0.12</td>
<td>0.67</td>
<td>26.83</td>
<td>11.54</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Elevation</td>
<td>3</td>
<td>0.04</td>
<td>0.71</td>
<td>21.80</td>
<td>4.52</td>
<td>0.041</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Rainfall</td>
<td>4</td>
<td>0.09</td>
<td>0.80</td>
<td>8.44</td>
<td>13.82</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Water</td>
<td>5</td>
<td>0.03</td>
<td>0.83</td>
<td>4.68</td>
<td>6.03</td>
<td>0.020</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>RH</td>
<td>4</td>
<td>0.005</td>
<td>0.83</td>
<td>3.56</td>
<td>0.93</td>
<td>0.344</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Conifer</td>
<td>5</td>
<td>0.01</td>
<td>0.84</td>
<td>3.80</td>
<td>1.90</td>
<td>0.178</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Latitude</td>
<td>6</td>
<td>0.01</td>
<td>0.85</td>
<td>3.85</td>
<td>2.18</td>
<td>0.150</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>$R_a$</td>
<td>5</td>
<td>0.004</td>
<td>0.85</td>
<td>2.61</td>
<td>1.09</td>
<td>0.364</td>
<td></td>
</tr>
</tbody>
</table>

Notes: All variables left in the model are significant at the 0.25 level. No other variable met the 0.25 significance level for entry into the model.

Thus, the model that estimates long-term annual forested watershed AET across the southern U.S. was obtained by the stepwise selection method. The multivariate linear regression equation takes the following form:

$$AET = 1156.90 + 0.32\text{Rainfall} - 19.90\text{Latitude} - 0.24\text{Elevation} + 0.92\text{Conifer} - 30.00\text{Water} \quad (24)$$

Where,

$AET =$ Long-term mean annual actual evapotranspiration of the watershed, mm.

Rainfall = Long-term mean annual precipitation of the watershed, mm.

Latitude = Watershed latitude by the stream gauging station, degree.

Elevation = Mean watershed elevation, m.

Conifer = Long-term percentage of watershed covered by conifer forests.

Water = Long-term percentage of watershed covered by the water body.
The model is highly significant (Table 5.3) with all the independent variables significant at $\alpha = 0.05$ level except Conifer, which is near significant at $\alpha = 0.10$ level (Table 5.4).

### Table 5.3. Analysis of variance of the AET model

<table>
<thead>
<tr>
<th>Source</th>
<th>Degree of freedom</th>
<th>Sum of squares</th>
<th>Mean square</th>
<th>F Value</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>5</td>
<td>643485</td>
<td>128697</td>
<td>32.78</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Error</td>
<td>30</td>
<td>117765</td>
<td>3925.49</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corrected Total</td>
<td>35</td>
<td>761250</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: $R^2 = 0.85$, adjusted $R^2 = 0.82$

### Table 5.4. Parameter estimates of the AET model

| Variable  | Degree of freedom | Parameter estimate | Standard error | t Value | Pr > |t| |
|-----------|-------------------|--------------------|----------------|---------|------|---|
| Intercept | 1                 | 1156.90            | 270.86         | 4.27    | 0.0002|
| Rainfall  | 1                 | 0.32               | 0.07           | 4.73    | <0.0001|
| Latitude  | 1                 | -19.90             | 6.46           | -3.08   | 0.0044|
| Elevation | 1                 | -0.24              | 0.06           | -4.17   | 0.0002|
| Conifer   | 1                 | 0.92               | 0.56           | 1.66    | 0.1081|
| Water     | 1                 | -30.00             | 12.69          | -2.36   | 0.0247|

### 5.1.2. Regression diagnostics

#### 5.1.2.1. Residual analysis

The plot (Figure 5.1) of the residuals ($e$) against predicted AET values indicates that it is a random scattering of the points above and below the line $e = 0$. The residual value for a specific watershed is also available (Figure 5.2). The normality tests of the residuals suggests that the assumptions of least squares are valid (Table 5.5).
Figure 5.1. Residual plot of long-term annual AET regression analysis

Figure 5.2. Plot of residuals against site number of the watersheds

Table 5.5. Tests for normality of residuals

<table>
<thead>
<tr>
<th>Test</th>
<th>Statistic value</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shapiro-Wilk</td>
<td>W</td>
<td>Pr &lt; W 0.95</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov</td>
<td>D</td>
<td>Pr &gt; D &gt;0.15</td>
</tr>
<tr>
<td>Cramer-von Mises</td>
<td>W-Sq</td>
<td>Pr &gt; W-Sq &gt;0.25</td>
</tr>
<tr>
<td>Anderson-Darling</td>
<td>A-Sq</td>
<td>Pr &gt; A-Sq &gt;0.25</td>
</tr>
</tbody>
</table>
The flag value of standardized residuals is 3.26, and the largest absolute value of the studentized residual (Student) and Rstudent is 2.70 (Table 5.6), which indicates there is no outlier in the observations.

### 5.1.2.2. Influence statistics

Cook’s D was used to evaluate influential points. The 25% ellipsoid flag value is 0.57; and the largest Cook’s D value in the observation is 0.12 (Table 5.6). Therefore, there is no influential point present in the data set.

**Table 5.6. Diagnostics statistics of the AET model**

<table>
<thead>
<tr>
<th>Site number</th>
<th>AET</th>
<th>Predicted AET</th>
<th>Residuals</th>
<th>Rstudent</th>
<th>Student</th>
<th>Cook’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>663.7</td>
<td>665.47</td>
<td>-1.78</td>
<td>-0.03</td>
<td>-0.03</td>
<td>0.00002</td>
</tr>
<tr>
<td>2</td>
<td>800.9</td>
<td>805.14</td>
<td>-4.24</td>
<td>-0.07</td>
<td>-0.07</td>
<td>0.00012</td>
</tr>
<tr>
<td>16</td>
<td>1028.58</td>
<td>1031.11</td>
<td>-2.53</td>
<td>-0.05</td>
<td>-0.05</td>
<td>0.00013</td>
</tr>
<tr>
<td>18</td>
<td>713.9</td>
<td>723.22</td>
<td>-9.32</td>
<td>-0.15</td>
<td>-0.16</td>
<td>0.00047</td>
</tr>
<tr>
<td>19</td>
<td>1042.02</td>
<td>1051.88</td>
<td>-9.86</td>
<td>-0.17</td>
<td>-0.17</td>
<td>0.00083</td>
</tr>
<tr>
<td>30</td>
<td>837.04</td>
<td>863.09</td>
<td>-26.05</td>
<td>-0.42</td>
<td>-0.42</td>
<td>0.00125</td>
</tr>
<tr>
<td>1</td>
<td>922.81</td>
<td>904.69</td>
<td>18.12</td>
<td>0.30</td>
<td>0.30</td>
<td>0.00145</td>
</tr>
<tr>
<td>25</td>
<td>869.34</td>
<td>857.5</td>
<td>11.84</td>
<td>0.21</td>
<td>0.21</td>
<td>0.00157</td>
</tr>
<tr>
<td>34</td>
<td>1015.35</td>
<td>1022.43</td>
<td>-7.08</td>
<td>-0.14</td>
<td>-0.14</td>
<td>0.0018</td>
</tr>
<tr>
<td>29</td>
<td>572.63</td>
<td>562.28</td>
<td>10.36</td>
<td>0.19</td>
<td>0.19</td>
<td>0.00183</td>
</tr>
<tr>
<td>3</td>
<td>798.42</td>
<td>818.05</td>
<td>-19.63</td>
<td>-0.33</td>
<td>-0.33</td>
<td>0.00219</td>
</tr>
<tr>
<td>9</td>
<td>686.71</td>
<td>675.38</td>
<td>11.33</td>
<td>0.22</td>
<td>0.22</td>
<td>0.00396</td>
</tr>
<tr>
<td>5</td>
<td>791.37</td>
<td>759.67</td>
<td>31.70</td>
<td>0.52</td>
<td>0.53</td>
<td>0.00437</td>
</tr>
<tr>
<td>14</td>
<td>999.38</td>
<td>963.68</td>
<td>35.70</td>
<td>0.59</td>
<td>0.60</td>
<td>0.0058</td>
</tr>
<tr>
<td>7</td>
<td>712.15</td>
<td>758.69</td>
<td>-46.54</td>
<td>-0.76</td>
<td>-0.77</td>
<td>0.00592</td>
</tr>
<tr>
<td>31</td>
<td>770</td>
<td>806.8</td>
<td>-36.81</td>
<td>-0.62</td>
<td>-0.63</td>
<td>0.00865</td>
</tr>
<tr>
<td>20</td>
<td>921.91</td>
<td>977.12</td>
<td>-55.21</td>
<td>-0.91</td>
<td>-0.91</td>
<td>0.00951</td>
</tr>
<tr>
<td>32</td>
<td>770.09</td>
<td>826.85</td>
<td>-56.76</td>
<td>-0.94</td>
<td>-0.94</td>
<td>0.01039</td>
</tr>
<tr>
<td>13</td>
<td>1112.55</td>
<td>1071.41</td>
<td>41.14</td>
<td>0.70</td>
<td>0.71</td>
<td>0.01446</td>
</tr>
<tr>
<td>27</td>
<td>1003.93</td>
<td>965.46</td>
<td>38.47</td>
<td>0.67</td>
<td>0.68</td>
<td>0.01608</td>
</tr>
<tr>
<td>23</td>
<td>859.19</td>
<td>929.93</td>
<td>-70.74</td>
<td>-1.18</td>
<td>-1.17</td>
<td>0.01653</td>
</tr>
<tr>
<td>11</td>
<td>854.11</td>
<td>837.77</td>
<td>16.34</td>
<td>0.35</td>
<td>0.35</td>
<td>0.01736</td>
</tr>
<tr>
<td>33</td>
<td>761.49</td>
<td>829.12</td>
<td>-67.63</td>
<td>-1.13</td>
<td>-1.12</td>
<td>0.01815</td>
</tr>
<tr>
<td>8</td>
<td>656.54</td>
<td>742.64</td>
<td>-86.11</td>
<td>-1.44</td>
<td>-1.41</td>
<td>0.02076</td>
</tr>
<tr>
<td>6</td>
<td>703.17</td>
<td>796.24</td>
<td>-93.07</td>
<td>-1.58</td>
<td>-1.54</td>
<td>0.02867</td>
</tr>
<tr>
<td>38</td>
<td>960.99</td>
<td>935.88</td>
<td>25.11</td>
<td>0.51</td>
<td>0.52</td>
<td>0.02901</td>
</tr>
</tbody>
</table>
Table 5.6 (continued)

<table>
<thead>
<tr>
<th>Site number</th>
<th>AET</th>
<th>Predicted AET</th>
<th>Residuals</th>
<th>Rstudent</th>
<th>Student</th>
<th>Cook’s D</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>851.21</td>
<td>807.84</td>
<td>43.37</td>
<td>0.78</td>
<td>0.79</td>
<td>0.02967</td>
</tr>
<tr>
<td>4</td>
<td>896.29</td>
<td>780.15</td>
<td>116.14</td>
<td>2.00</td>
<td>1.90</td>
<td>0.03693</td>
</tr>
<tr>
<td>37</td>
<td>1018.99</td>
<td>1050.15</td>
<td>-31.16</td>
<td>-0.62</td>
<td>-0.63</td>
<td>0.03917</td>
</tr>
<tr>
<td>15</td>
<td>643.3</td>
<td>606.24</td>
<td>37.06</td>
<td>0.71</td>
<td>0.72</td>
<td>0.04009</td>
</tr>
<tr>
<td>28</td>
<td>650.87</td>
<td>583.55</td>
<td>67.32</td>
<td>1.17</td>
<td>1.17</td>
<td>0.0409</td>
</tr>
<tr>
<td>36</td>
<td>778.23</td>
<td>690.36</td>
<td>87.87</td>
<td>1.53</td>
<td>1.49</td>
<td>0.05021</td>
</tr>
<tr>
<td>22</td>
<td>952.71</td>
<td>887.86</td>
<td>64.85</td>
<td>1.19</td>
<td>1.19</td>
<td>0.07294</td>
</tr>
<tr>
<td>24</td>
<td>684.43</td>
<td>741.85</td>
<td>-57.42</td>
<td>-1.10</td>
<td>-1.10</td>
<td>0.08672</td>
</tr>
<tr>
<td>39</td>
<td>1135.4</td>
<td>988.35</td>
<td>147.05</td>
<td>2.70</td>
<td>2.45</td>
<td>0.09142</td>
</tr>
<tr>
<td>35</td>
<td>671.64</td>
<td>793.49</td>
<td>-121.85</td>
<td>-2.23</td>
<td>-2.10</td>
<td>0.11827</td>
</tr>
</tbody>
</table>

5.1.2.3. Collinearity diagnostics

A condition index around 10 would indicate weak collinearity among the independent variables; a condition index of 30 to 100 would indicate moderate collinearity; and a condition index that is larger than 100 would indicate severe collinearity. The calculated condition indices of the model indicate that there is no collinearity existing among the independent variables (Table 5.7). Thus, the model developed in this study is sufficient to predict long-term annual AET for forested watersheds across the southern U.S..

Table 5.7. Collinearity diagnostics (intercept adjusted) of the AET model

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Eigenvalue</th>
<th>Condition index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.09</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1.20</td>
<td>1.32</td>
</tr>
<tr>
<td>3</td>
<td>0.96</td>
<td>1.48</td>
</tr>
<tr>
<td>4</td>
<td>0.49</td>
<td>2.06</td>
</tr>
<tr>
<td>5</td>
<td>0.26</td>
<td>2.83</td>
</tr>
</tbody>
</table>
5.2. Six PET comparisons

Since this study was focused on the regional long-term annual AET and PET, all the comparisons were made on the long-term annual basis. A SAS program was coded to assist the comparisons. The program is presented in appendix G, and parts of the program’s output are presented in this section.

5.2.1. Correlation between six PET methods

On the long-term annual basis, the six PET methods were highly correlated (Table 5.8). Especially between the Thornthwaite and Hamon methods, the correlation coefficient is 0.997, which indicates that the two methods are highly correlated since both are temperature-based methods. Compared with others, the Hargreaves-Samani PET method has the lowest correlation coefficients with all the other PET methods although the coefficients are still high.

<table>
<thead>
<tr>
<th>PET methods</th>
<th>Thornthwaite</th>
<th>Hamon</th>
<th>Turc</th>
<th>Priestley-Taylor</th>
<th>Makkink</th>
<th>Hargreaves-Samani</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thornthwaite</td>
<td>1</td>
<td>0.997</td>
<td>0.96</td>
<td>0.94</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>Hamon</td>
<td>0.997</td>
<td>1</td>
<td>0.97</td>
<td>0.94</td>
<td>0.93</td>
<td>0.89</td>
</tr>
<tr>
<td>Turc</td>
<td>0.96</td>
<td>0.97</td>
<td>1</td>
<td>0.97</td>
<td>0.98</td>
<td>0.88</td>
</tr>
<tr>
<td>Priestley-Taylor</td>
<td>0.94</td>
<td>0.94</td>
<td>0.97</td>
<td>1</td>
<td>0.95</td>
<td>0.85</td>
</tr>
<tr>
<td>Makkink</td>
<td>0.93</td>
<td>0.93</td>
<td>0.98</td>
<td>0.95</td>
<td>1</td>
<td>0.85</td>
</tr>
<tr>
<td>Hargreaves-Samani</td>
<td>0.89</td>
<td>0.89</td>
<td>0.88</td>
<td>0.85</td>
<td>0.85</td>
<td>1</td>
</tr>
</tbody>
</table>

5.2.2. Correlation between six PET methods and AET estimate

In order to evaluate the performances of six PET methods, six long-term annual PET values were correlated with the AET values (Table 5.9). All PET values are highly correlated with the AET values. However, in terms of comparisons, the Priestley-Taylor PET has the
highest correlation with the AET and the Hargreaves-Samani PET has the lowest correlation. The Makkink PET also has low correlation with AET compared with other PET methods.

Table 5.9. Pearson correlation coefficients between six PET methods and AET estimate

<table>
<thead>
<tr>
<th>PET methods</th>
<th>R</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thornthwaite</td>
<td>0.63</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hamon</td>
<td>0.63</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Turc</td>
<td>0.64</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Priestley-Taylor</td>
<td>0.65</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Makkink</td>
<td>0.60</td>
<td>0.0001</td>
</tr>
<tr>
<td>Hargreaves-Samani</td>
<td>0.57</td>
<td>0.0003</td>
</tr>
</tbody>
</table>

Notes: N = 36, Prob > |r| under H₀: R (h₀) = 0 (No correlation)

5.2.3. Differences between six PET methods

Averaged on all sites, the Thornthwaite method predicts the lowest long-term annual PET value while the Hargreaves-Samani method predicts the highest value (Figure 5.3). The PET estimated by Thornthwaite method is even slightly lower than long-term annual AET value (Table 5.10). The long-term mean annual PET predicted by six PET methods are highly correlated. However, in multivariate statistics test, they are all significantly different from each other at α = 0.05 level (Table 5.11).
Figure 5.3. Comparison between PET by six PET methods

Table 5.10. Simple statistics for AET and PET estimated by six methods (n = 36)

<table>
<thead>
<tr>
<th>Source</th>
<th>N</th>
<th>Mean (mm)</th>
<th>Std Dev (mm)</th>
<th>Sum (mm)</th>
<th>Minimum (mm)</th>
<th>Maximum (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AET</td>
<td>36</td>
<td>836.43</td>
<td>147.48</td>
<td>30111</td>
<td>572.63</td>
<td>1135</td>
</tr>
<tr>
<td>Thornthwaite</td>
<td>36</td>
<td>833.78</td>
<td>124.91</td>
<td>30016</td>
<td>619.73</td>
<td>1088</td>
</tr>
<tr>
<td>Hamon</td>
<td>36</td>
<td>1108</td>
<td>170.23</td>
<td>39878</td>
<td>789.93</td>
<td>1414</td>
</tr>
<tr>
<td>Turc</td>
<td>36</td>
<td>1031</td>
<td>134.89</td>
<td>37132</td>
<td>795.23</td>
<td>1266</td>
</tr>
<tr>
<td>Priestley-Taylor</td>
<td>36</td>
<td>1071</td>
<td>115.46</td>
<td>38547</td>
<td>867.83</td>
<td>1279</td>
</tr>
<tr>
<td>Makkink</td>
<td>36</td>
<td>943.60</td>
<td>81.49</td>
<td>33970</td>
<td>796.50</td>
<td>1100</td>
</tr>
<tr>
<td>Hargreaves-Samani</td>
<td>36</td>
<td>1336</td>
<td>130.94</td>
<td>48111</td>
<td>1020</td>
<td>1577</td>
</tr>
</tbody>
</table>

Notes: Std Dev = Standard deviation.

Table 5.11. Multivariate statistics test of PET estimated by six methods (n = 36)

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Value</th>
<th>F Value</th>
<th>Numerator DF</th>
<th>Denominator DF</th>
<th>Pr &gt; F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilks' Lambda</td>
<td>0.001</td>
<td>4621.51</td>
<td>5</td>
<td>31</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Pillai's Trace</td>
<td>0.999</td>
<td>4621.51</td>
<td>5</td>
<td>31</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Hotelling-Lawley Trace</td>
<td>745.405</td>
<td>4621.51</td>
<td>5</td>
<td>31</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Roy's Greatest Root</td>
<td>745.405</td>
<td>4621.51</td>
<td>5</td>
<td>31</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>

Notes: Hypothesis $H_0$: no PET Effect, DF = Degree of Freedom
When the long-term mean annual PET values were investigated in thirty-six sites, greater differences were found among the temperature-based PET methods than radiation-based PET methods (Figure 5.4 and Figure 5.5). The PET values predicted by the three radiation-based methods were found to be very close, especially for the Priestley-Taylor and Turc methods, which have a correlation coefficient of 0.97 (Table 5.8). The Makkink method predicts the lowest PET value among the radiation-based methods. However, the PET values predicted by all six methods are significantly different from each other ($\alpha = 0.05$). Although the Hamon and Thornthwaite methods have the highest correlation coefficient (0.997, Table 5.8), the PET values predicted by the Hamon method are much higher than those predicted by the Thornthwaite method (Figure 5.4).

![Temperature-base PET](image)

**Figure 5.4.** PET simulated by the three temperature-based PET methods.
Figure 5.5. PET simulated by the three radiation-based PET methods

Four watersheds were selected to explore the annual pattern of the PET in these particular sites (Figure 5.6, Figure 5.7, Figure 5.8 and Figure 5.9). There are two from forested experimental watersheds (Figure 5.6 and Figure 5.7) and two from USGS monitored watersheds (Figure 5.8 and Figure 5.9). The results indicate that the spatial patterns (Figure 5.4 and Figure 5.5) and temporal patterns (Figure 5.6, Figure 5.7, Figure 5.8 and Figure 5.9) of the PET estimated by 6 methods are similar. The temperature-based PET methods have bigger differences than the radiation-based methods both spatially and temporally. The argreaves-Samani PET is larger than Hamon PET, which is larger than the Thornthwaite PET. For radiation based methods, the Makkink PET is the lowest and Priestely-Taylor and Turc are very close, they are even close to the mean PET calculated by the six methods.
Six PET at Bradford watershed (site number 34), FL (13 years)

Figure 5.6. Six PET simulated at Bradford watershed (site number 34), FL

Six PET at Carteret watershed (site number 37), coastal NC (13 years)

Figure 5.7. Six PET simulated at Carteret watershed (site number 37), NC
Figure 5.8. Six PET simulated at a watershed in Arkansas (site number 15, 30 years)

Figure 5.9. Six PET simulated at a watershed in Georgia (site number 32, 30 years)
Based on the two assumptions that were made in this study, the Priestley-Taylor, Turc and Hamon methods performed better than the Thornthwaite, Makkink and Hargreaves-Samani methods. And the radiation-based PET methods performed better than temperature-based PET methods.

Spatially, AET and PET have similar distribution patterns (Figure 5.10 and Figure 5.11), their values are in a gradient from the south to the north and from the coastal zone to the mountain areas. These observations are also actually reflected in the AET model (equation 24), which includes latitude and elevation as two independent variables.
AET estimates across the Southern US

AET (mm/year)
- 572.63 - 656.54
- 656.54 - 713.9
- 713.9 - 800.9
- 800.9 - 896.29
- 896.29 - 1003.93
- 1003.93 - 1135.4

Figure 5.10. AET estimated by the water balance for all the watersheds
Figure 5.11. PET estimated by the Priestley-Taylor method for all the watersheds
5.3. Conclusions and discussion

From this study, a long-term annual AET model was developed and can be used in the southern forested watersheds to evaluate hydrologic effects of climate and land use changes. The five independent variables included in the model are rainfall, latitude, elevation, and the percentage of the land cover of conifer forests and the water body. These variables are readily available in a regional database. The PET is not included in the model since other variables perform better than the PET to explain the variation of the AET. Validation is needed in order to apply the model. The regression equation should be applied to other forested watersheds in the southern U.S.. The quality of land cover data could be improved if the real boundaries of the watersheds were available.

PET estimates from the six PET methods are highly correlated with each other and the AET estimates. However, the estimates are significantly different from each other ($\alpha = 0.05$ level). The Thornthwaite method yielded the lowest PET value while the Hargreaves-Samani method the highest. There are greater differences among the three temperature-based PET methods than among the three radiation-based PET methods. In the three radiation-based methods, the Makkink method predicted the lowest PET values, and the Priestley-Taylor and Turc methods predicted values that were very close to each other.

In order to evaluate the performances of six PET methods, two assumptions were made in this study. One was that PET should be higher than AET on a long-term annual basis. Another assumption was that a linear relationship exists between PET and AET, therefore, the higher correlation coefficient of PET with AET, the better the PET estimates. The PET estimated by the Thornthwaite method is lower than AET. The Priestley-Taylor has the highest correlation coefficient with the AET. In general, the Priestley-Taylor, Turc and
Hamon methods performed better than the Thornthwaite, Makkink and Hargreaves-Samani methods.

In this study, hydro-meteorologic data collected in one station were used to represent the condition of the entire watershed. However, the watersheds have different sizes. Some watersheds are small in area, but some others are very large. For large watersheds, using data from only one station may lead to larger errors because of the spatial variability, and the spatial relationships of data were not explored in this study.
REFERENCES


Jensen, M.E. (Editor), 1973. Consumptive use of water and irrigation water requirements. ASCE, New York, NY, USA.


APPENDICES
Appendix A. Methodology of four eight-digit HUC watershed selections in South Carolina and Georgia

State boundaries and eight-digit HUC watersheds

The U.S. state boundary coverage in shapefile format included with ArcView GIS 3.2 (ArcView) software was used. The eight-digit HUC watersheds were downloaded from the USGS web site (http://water.usgs.gov/lookup/getspatial?huc250k). The data were added to ArcView as two themes.

Radiation data of Georgia and South Carolina.

Radiation data (in units of W-hr/m²) were available from the web site of the Renewable Resource Data Center (RReDC, http://rredc.nrel.gov/solar/old_data/nsrdb/dsf/). The data used in this study are the monthly averages of the daily data. The solar radiation, extraterrestrial solar radiation, and relative humidity were obtained from this dataset for all watersheds. There were six stations in Georgia and three stations in South Carolina. The locations of the stations were input into ArcView as a table. Then from “View” menu, click “Add Event Theme…” to get a new theme that displays the stations as point data in the view window (Figure 7.1).
Figure 7.1. Available solar radiation stations in Georgia and South Carolina

Selections of eight-digit HUC watersheds

In ArcView, from “Theme” menu, click “Theme/Select by theme…” to select the HUC watersheds that completely contain at least one solar radiation station. Thus, there were
six HUC watersheds selected in Georgia and three HUC watersheds selected in South Carolina. After nine HUC watersheds were selected, they were converted to a new theme in the view window (Figure 7.2).

![Selected 8-digit HUC in GA and SC](image)

Figure 7.2. Selected eight-digit HUC watersheds in GA and SC

**Streamflow Data for Georgia and South Carolina**

Both the active and discontinued gauging stations in Georgia were available on line at: [http://ga.water.usgs.gov/publications/wdr99-1/allstreamflow.html](http://ga.water.usgs.gov/publications/wdr99-1/allstreamflow.html), which has information
about station number, station name, latitude, longitude, drainage area and periods of record of all the stations in the state of Georgia.

The data were downloaded and the gauging stations were picked up manually by setting the periods of record to be 1961 – 1990. Only those stations that satisfy the periods of record were selected for further analysis. There were fifty-four gauging stations in total available for this study in Georgia.

The available streamflow data could be obtained on line at USGS web site at http://water.usgs.gov/nwis/annual. South Carolina as a state criterion was used to search the database, there were one hundred and nine gauging stations available in South Carolina.

The gauging station information such as site number, site name, longitude, latitude, drainage area, etc. were downloaded. The file containing fifty-four gauging stations in Georgia and one hundred and nine gauging stations in South Carolina was transferred into Tab delimited text file. ArcView imported the text file as a table and converted it to a theme via the “View/Add Event Theme…” menu function. Therefore, a new streamflow station point theme showing the spatial distribution of streamflow gauging stations in South Carolina and Georgia was created and displayed in the view window.

In ArcView, from “Theme” menu, click “Theme/Select by theme…”. From the nine selected HUC theme, select the HUC watershed that completely contains at least one streamflow gauging station in the streamflow gauging station point theme. There were five HUC watersheds selected in Georgia and two HUC watersheds selected in South Carolina. The seven selected HUC watersheds were converted to a new theme.

Again, the “Theme/Select by theme…” function was used to select the stream gauging stations that were located in the seven selected HUC watersheds, there were totally
forty-two streamflow gauging stations selected. The forty-two selected gauging stations were converted to a new theme.

**Climate Data for Georgia and South Carolina.**

Web site [http://cdo.ncdc.noaa.gov/cdo/info.html](http://cdo.ncdc.noaa.gov/cdo/info.html) provides information about the database that contains the summary of the monthly type data for U.S. stations along with selected non-U.S. stations in U.S. territories. There are over 19,000 stations in the historical database. Monthly total precipitation, monthly mean temperature, and the monthly average maximum and minimum temperature are commonly available. The dataset also has the station list that contains information about all the stations such as station name, country, state, latitude, longitude, elevation, etc.

The station list was downloaded and the file was transferred into a Tab delimited text file using Microsoft Excel (Version 2000). ArcView imported the text file as a table and converted it to a theme via the “View/Add Event Theme…” menu function. Therefore, a new climate station point theme showing the spatial distribution of climate gauging stations was created and displayed in the view window.

In ArcView, from “Theme” menu, click “Theme/Select by theme…” to select the HUC watersheds in the seven selected HUC theme that completely contain at least one climate station in the climate station point theme. The result showed that all the seven selected HUC theme contain climate stations.

Again, the function was used to select the climate stations that are completely within the seven selected HUC watersheds. There are fifty-eight climate stations located in the seven selected HUC watersheds. The fifty-eight selected climate stations were converted to a new point theme (Figure 7.3).
Figure 7.3. Available HUC watersheds and stations in GA and SC
So far, there were seven potential HUC watersheds available. If a HUC has more than one streamflow gauging station or climate station, the one that was the closest to the solar radiation station was used to reference data.

The seven selected HUC watersheds were further investigated to ensure that all the required data were available. A HUC watershed that is located in Georgia has only one gauging station and its drainage area is 9850 mi². The watershed drainage area is too large and doesn’t compare well with other watersheds that are already in the database, most of which have drainage areas of less than 1000 mi². Thus, it was eliminated from consideration. Another HUC watershed in Georgia either doesn’t have the solar radiation station, streamflow gauging station and climate station that have long-term data or the stations are not close enough to represent the watershed. It was also eliminated from further consideration. Finally, there is a HUC watershed in South Carolina that has five stream gauging stations, but none of them has long-term data for the period of 1961-1990. The watershed was also eliminated from consideration. Thus, there were only four HUC watersheds that were available for the states of Georgia and South Carolina (Figure 7.4).
Figure 7.4. Available watersheds in GA and SC
There were only one solar radiation station, one gauging station and one climate station selected to acquire data to represent a HUC watershed. If a HUC watershed had more than one stream gauging station or climate station available, the one that was the closest to the solar radiation station was used in the watershed.

Appendix B. Procedures of delineating a watershed (site number 32) and deriving its elevation and land cover in Georgia

Obtain information about the stream gauging station in the watershed

1. The information about the gauging station is obtained by searching site number 02217500 at USGS web site (http://water.usgs.gov/nwis/annual). The searching result indicates that the stream gauging station is located in Clarke County, Georgia, near Athens. Hydrologic Unit Code 03070101, Latitude 33°56'48", Longitude  83°25'22", NAD27, Drainage area 392.00 square miles, Gage datum 555.66 feet (169.37 meter) above sea level NGVD29.

Obtain necessary DEM by looking up USGS Quad Neatlines

2. Download USGS Quad Neatlines coverage from NCSU library: go to web page http://www.lib.ncsu.edu/stacks/gis/themes/term0280.html. Click "ArcUSA 1:2M: USGS 1:24,000 Topographic Quadrangle Series Index" to download the neatline coverage Q_24K. The coverage is in the projection Albers, map units: meters.

3. Clip Georgia (GA) state neatline to reduce study area: Launch ArcView GIS 3.2 (ArcView), and add the neatline coverage Q_24K into the view. Query the neatline Q_24K by the expression (St_name1 = “Georgia”), convert the selected polygon to ganeatline.shp.
There are totally 948 polygons located in GA. Don’t add the new theme into the view and delete Q_24K from the view.

4. Download SDTS format DEM: Go to Geocommunity web site (http://www.gisdatadepot.com/catalog/US/61089/sublist.html) to download DEM data in GA state by counties. Since it is known that the gauging station is located in Clarke County and nears Athens, click on “Clarke”, then click “Digital Elevation Models (DEM) - 24K”, download Athens East and Athens West DEM zipped files.

5. Download a DEM translator: Go to web page http://www.cs.arizona.edu/topovista/sdts2dem.html to download the program sdts2dem.exe. Read information about how to run this program and follow the direction.

6. Unzip downloaded DEM files and convert them to USGS DEM format in separate directory: Unzip two downloaded DEM files separately and one in each directory (in separate directory). Copy file sdts2dem.exe to each directory and run the program to convert the DEM from STDS format to USGS format. Name the output eathens.dem and wathens.dem respectively.

7. Convert DEM to Grid: Launch ArcToolBox, then double click “Conversion Tools / Import to Raster / DEM to Grid” to convert DEM to Grid. Convert eathens.dem and wathens.dem to grids and name them eathens and wathens respectively. Put the output grids (eathens and wathens) in the same directory, and the following workspace is default to this directory.

8. Convert shapefile to coverage: In ArcToolBox, double click “Conversion Tools / Import to Coverage / Shapefile to Coverage”, convert the shapefile ganeatline.shp to ganeatlinecov coverage.
9. Define projection for the ganeatlinecov coverage: In ArcToolBox, double click “Data Management Tools / Projections / Define Projection Wizard (coverages, grids, TINs)” to define projection for ganeatlinecov coverage. Use “Define the coordinate system interactively…” Define it to “Projections: Albers Equal-Area; Units: meters; 1st standard parallel: 29 30 0; 2nd standard parallel: 45 30 0; Longitude of center of projection: -96 0 0; Latitude of origin of projection: 23 0 0; False easting: 0; False northing: 0; X shift: 0, Y shift: 0; Spheroid: Clarke 1866; Semimajor axis: 6378206.4; Semiminor axis: 6356583.8.”

10. Reproject ganeatlinecov to the same as Grid data: launch ArcInfo workstation. After set the workspace to eathens and wathens grid data directory (this project's default workspace), in Arc command line, type “Arc: describe eathens” to get the projection information about the grid data, which will be used to reproject ganeatlinecov coverage. In ArcToolBox, double click “Data Management Tools / Projections / Project Wizard (coverages, grids)” to reproject the coverage to the same as the grids. Use “Project my data to a specified coordinate system…”, then select “Projections: UTM; Units: meters; Zone: 17; X shift: 0, Y Shift”. Name the output coverage prjganeatline.

11. Build the polygons: In Arc command line, type “Arc: build prjganeatline poly” to build polygons for prjganeatline coverage.

12. Download Longitude/Latitude Conversion Utility: go to web page http://gis.esri.com/arcscripts/details.cfm?CFGRIDKEY=4F05C7CF-A385-11D4-943300508B0CB419 to download the Longitude/Latitude Conversion Utility extension and save it in the ArcView extension directory (Usually is C:/ESRI/AV_GIS30/ARCVIEW/EXT32).
13. Add extension and data to ArcView: Back to ArcView, add Longitude/Latitude Conversion Utility and Spatial Analysis extensions, then add data prjganeatline, eathens and wathens to the view. Set map units and distance units to meters. Auto-label theme prjganeatline using field Quad_name, the view window clearly displays where is the neatline polygon which a grid is located in (Figure 7.5).

Figure 7.5. Neatline polygon information for grids

14. Add gauging station point: In ArcView, from the View menu, click “View/New Theme…” to create a new point theme named gauging.shp and add it to the view. Then click on Longitude/Latitude Conversion Utility icon to input the gauging station location (Longitude -83°25'22", Latitude 33°56'48") to the view. In the Projection properties dialog window, select “Standard option; Category: UTM – 1927; Type Zone 17”, which is the same
15. Identify the mouth (pour point) of the watershed: The location of the gauging station is the mouth of the watershed. Zoom in the view until it clearly displays the cell of the grid that contains the point in gauging.shp. The cell right next to the cell that contains the point is the mouth of the watershed. Because its elevation is 171m, which is very close to 169.37m and lower than the cell that contains the point, which is 174m. Move the mouse to the cell that is the mouth of the watershed and get its approximate coordinate: (276105.85, 3758692.60), which will be used to delineate the watershed.

16. Getting GA state county boundary coverage: Open another new ArcView project and add ESRI USA county theme data into the view. The data is usually located in the directory “C:\ESRI\ESRIDATA\USA\counties”, which is included in ArcView software. Query the theme by the expression ( [State_name] = "Georgia" ), and convert the selected polygons to shapefile gacounties.shp. Don’t add the new theme to the view. Close the project and don’t save anything. In ArcToolBox, double click “Conversion Tools / Import to Coverage / Shapefile to Coverage” to convert the shapefile gacounties.shp to gacountiescov as a coverage. Then double click “Data Management Tools / Projections / Define Projection Wizard (coverages, grids, TINs)” to define the projection for the coverage. Use “Define the coordinate system interactively…” to define it as “Projections: Geographic; Units:DD; Spheroid: Clarke 1866; Semimajor axis: 6378206.4; Semiminor axis: 6356583.8”. Then use ArcToolBox to reproject gacountiescov to the same as prjganeatline coverage and name the new coverage prjgacounties. In Arc command line, type “Arc: build prjgacounties” to build the polygons of the coverage. Add prjgacounties to the ArcView project, remove labels of
theme prjganeatline and auto-label theme prjgacounties by selecting “Name” as the label field. Zoom in the view until it clearly shows where is the county which a grid or a neatline polygon is located in (Figure 7.6).

![Figure 7.6. County information for grids](image)

17. Search and download DEM data: After arrange the themes in the view in an appropriate order (Figure 7.6), the view clearly displays the counties where the DEM need to be downloaded. Make theme prjganeatline active and select the Identify tool, click on the neatline polygons that near the mouth of the watershed. In the Identify Results dialog, find out the corresponding USGS quadratic DEM by field name “Usgs_qd_id” or “Quad_name”, then go to Geocommunity web site (http://www.gisdatadepot.com/catalog/US/61089/sublist.html) to download the corresponding DEM. All needed DEM data are in the following table (Table 7.1).
Table 7.1. Grids used to delineate the watershed boundary

<table>
<thead>
<tr>
<th>NO.</th>
<th>County Name</th>
<th>Quad_name</th>
<th>Usgs_qd_id</th>
<th>Grid Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Clarke</td>
<td>Athens East</td>
<td>33083h3</td>
<td>Eathens</td>
</tr>
<tr>
<td>2</td>
<td>Clarke</td>
<td>Athens West</td>
<td>33083h4</td>
<td>wathens</td>
</tr>
<tr>
<td>3</td>
<td>Clarke</td>
<td>Hull</td>
<td>34083a3</td>
<td>Hull</td>
</tr>
<tr>
<td>4</td>
<td>Clarke</td>
<td>Barnett Shoals</td>
<td>33083g3</td>
<td>Barnett</td>
</tr>
<tr>
<td>5</td>
<td>Clarke</td>
<td>Jefferson</td>
<td>34083a5</td>
<td>Jefferson</td>
</tr>
<tr>
<td>6</td>
<td>Clarke</td>
<td>Nicholson</td>
<td>34083a4</td>
<td>Nicholson</td>
</tr>
<tr>
<td>7</td>
<td>Clarke</td>
<td>Statham</td>
<td>33083h5</td>
<td>Statham</td>
</tr>
<tr>
<td>8</td>
<td>Oconee</td>
<td>High Shoals</td>
<td>33083g5</td>
<td>High</td>
</tr>
<tr>
<td>9</td>
<td>Oconee</td>
<td>Watkinsville</td>
<td>33083g4</td>
<td>watkins</td>
</tr>
<tr>
<td>10</td>
<td>Oconee</td>
<td>Winder South</td>
<td>33083h6</td>
<td>swinder</td>
</tr>
<tr>
<td>11</td>
<td>Barrow</td>
<td>Auburn</td>
<td>34083a7</td>
<td>Auburn</td>
</tr>
<tr>
<td>12</td>
<td>Barrow</td>
<td>Bold Springs</td>
<td>33083h7</td>
<td>Bold</td>
</tr>
<tr>
<td>13</td>
<td>Barrow</td>
<td>Chestnut Mountain</td>
<td>34083b7</td>
<td>Chestnut</td>
</tr>
<tr>
<td>14</td>
<td>Barrow</td>
<td>Winder North</td>
<td>34083a6</td>
<td>Nwinder</td>
</tr>
<tr>
<td>15</td>
<td>Walton</td>
<td>Between</td>
<td>33083g7</td>
<td>Between</td>
</tr>
<tr>
<td>16</td>
<td>Walton</td>
<td>Loganville</td>
<td>33083g8</td>
<td>Logan</td>
</tr>
<tr>
<td>17</td>
<td>Walton</td>
<td>Monroe</td>
<td>33083g6</td>
<td>monroe</td>
</tr>
<tr>
<td>18</td>
<td>Gwinnett</td>
<td>Lawrenceville</td>
<td>33083h8</td>
<td>Lawren</td>
</tr>
<tr>
<td>19</td>
<td>Jackson</td>
<td>Apple Valley</td>
<td>34083b5</td>
<td>apple</td>
</tr>
<tr>
<td>20</td>
<td>Jackson</td>
<td>Commerce</td>
<td>34083b4</td>
<td>Commerce</td>
</tr>
<tr>
<td>21</td>
<td>Jackson</td>
<td>Gillsville</td>
<td>34083c6</td>
<td>gillsville</td>
</tr>
<tr>
<td>22</td>
<td>Jackson</td>
<td>Homer</td>
<td>34083c4</td>
<td>homer</td>
</tr>
<tr>
<td>23</td>
<td>Jackson</td>
<td>Ila</td>
<td>34093b3</td>
<td>Ila</td>
</tr>
<tr>
<td>24</td>
<td>Jackson</td>
<td>Maysville</td>
<td>34083c5</td>
<td>May</td>
</tr>
<tr>
<td>25</td>
<td>Jackson</td>
<td>Pendergrass</td>
<td>34083b6</td>
<td>Pender</td>
</tr>
<tr>
<td>26</td>
<td>Hall</td>
<td>Chestatee</td>
<td>34083c8</td>
<td>Chesta</td>
</tr>
<tr>
<td>27</td>
<td>Hall</td>
<td>Gainesville</td>
<td>34083c7</td>
<td>Gain</td>
</tr>
<tr>
<td>28</td>
<td>Hall</td>
<td>Flowery Branch</td>
<td>34083b8</td>
<td>Flower</td>
</tr>
<tr>
<td>29</td>
<td>Hall</td>
<td>Hog Mountain</td>
<td>34083a8</td>
<td>Hog</td>
</tr>
<tr>
<td>30</td>
<td>Madison</td>
<td>Ashland</td>
<td>34083c3</td>
<td>Ash</td>
</tr>
</tbody>
</table>

Totally, there are 30 DEM data downloaded from the Geocommunity web site. After download all the DEM data, follow the same procedures as step 6 and step 7 to obtain the corresponding grid, which are: unzip the file, convert SDTS to DEM, then convert DEM to Grid. Save all the output grids in one directory.
Watershed and stream network delineation using ArcInfo 8.1

18. Merge 30 grids to one grid: In Arc command line, type “Arc: Grid” to launch the Grid model. In Grid command line, type the following command (30 grids inside the ( ) ) to Merge 30 grids to one grid.

Grid: pregagrid = mosaic ( eathens, wathens, hull…, hog, Ash )

19. Fill the gap with neighboring cells for the output grid

Grid: gagrid = con ( isnull (pregagrid), focalmean (pregagrid, rectangle, 3, 3), pregagrid )

20. Set the window of the analysis environment

Grid: setwindow gagrid

21. Fill the sinks in the grid and name output filledgagrid

Grid: fill gagrid filledgagrid sink

22. Getting the flow direction grid

Grid: gaflow = flowdirection ( filledgagrid )

23. Delineate the watershed by specifying the mouth (pour point) of the watershed

Grid: gawshed = watershed ( gaflow, selectpoint (gaflow, 276105.85, 3758692.60) )

24. Make a polygon coverage of the watershed boundary

Grid: gawshedcov = gridpoly ( gawshed )

Acquire the watershed elevation

25. Clip the watershed DEM

Grid: gridclip gagrid wshedgrid cover gawshedcov

26. Acquire watershed average elevation

Grid: describe wshedgrid
The output indicates that the average elevation of the watershed is 270 m.

**Derive the watershed land cover**


   The land cover grid is in a projection different from the gawshedcov. Therefore, reprojection is needed for gawshedcov.

   **Getting the projection of landcover grid**

   Grid: describe landcover

   Use ArcToolBox to reproject gawshedcov to the same projection as landcover grid (select option: Project my data to match existing data - landcover). Name the output coverage prjgawshedcov.

   Then quit the Grid model and back to Arc model in ArcInfo workstation. Build the coverage and clip out the land cover grid of the watershed.

   Grid: q

   Arc: build prjgawshedcov

   Arc: grid

   Grid: gridclip landcover.type wshedlandcov cover prjgawshedcov

28. Quit ArcInfo workstation
Grid: q

Arc: q

29. In ArcToolBox, reproject grid wshedlandcov to the same projection as gagrid. Name the output grid prjwshedlandcov.

30. Add watershed data to ArcView: Back to ArcView, add data gawshedcov, wshedgrid, wshedstrorder, wshedstrcov, prjwshedlandcov to the view. Add landcover.txt as table to the project. Rearrange the theme and the legend in an appropriate manner (Figure 7.7) to clearly display the themes in the view window.
31. Open theme prjlandcover attribute table and landcover.txt table. Select “Value” as the common field in both tables. Then join tables to prjlandcover attribute table.

32. Make sure prjlandcover attribute table is active. Select column “Count” in the table. Then from the menu, click “Field/Statistics…” In the Statistics for Count field dialog window, the Sum is 1126850. Click OK button to close the dialog window.

Figure 7.7. The delineated watershed
33. From menu, click “Table/Start Editing”, then click menu “Edit/Add Field…” in the “Field Definition” dialog window, set Field Name: LandcoverRatio; Type: Number; Width: 6; Decimal Places: 4. Then click OK.

34. Click menu “Field/Calculate…”, in the Field Calculator dialog window, input the following expression:

\[
\text{[Count]} / 1126850
\]

35. Click OK, which calculate the ratio of each landcove type in the watershed.

36. From menu, click “Table/Stop Editing” and click “Yes” to save the edits. In the attribute table, sort the table by column LandcoverRatio as descendent. The percentage of each land cover type is sorted by descendent in the table (Table 72). Close the attribute table, and back to the view.

Table 7.2. Watershed land cover compositions

<table>
<thead>
<tr>
<th>Value</th>
<th>Count</th>
<th>Landcover</th>
<th>LandcoverRatio</th>
</tr>
</thead>
<tbody>
<tr>
<td>41</td>
<td>493401</td>
<td>Decidous Forest</td>
<td>0.4379</td>
</tr>
<tr>
<td>81</td>
<td>237142</td>
<td>Pasture/Hay</td>
<td>0.2104</td>
</tr>
<tr>
<td>43</td>
<td>147621</td>
<td>Mixed Forest</td>
<td>0.1310</td>
</tr>
<tr>
<td>42</td>
<td>99880</td>
<td>Evergreen Forest</td>
<td>0.0886</td>
</tr>
<tr>
<td>82</td>
<td>90975</td>
<td>Row Crops</td>
<td>0.0807</td>
</tr>
<tr>
<td>21</td>
<td>16280</td>
<td>Low Intensity Residential</td>
<td>0.0144</td>
</tr>
<tr>
<td>91</td>
<td>11320</td>
<td>Woody Wetlands</td>
<td>0.0100</td>
</tr>
<tr>
<td>85</td>
<td>9703</td>
<td>Urban/Recreational Grasses</td>
<td>0.0086</td>
</tr>
<tr>
<td>23</td>
<td>7865</td>
<td>Commercial/Industrial/Transpotation</td>
<td>0.0070</td>
</tr>
<tr>
<td>22</td>
<td>3293</td>
<td>High Intensity Residential</td>
<td>0.0029</td>
</tr>
<tr>
<td>11</td>
<td>3205</td>
<td>Open water</td>
<td>0.0028</td>
</tr>
<tr>
<td>33</td>
<td>3055</td>
<td>Transitional</td>
<td>0.0027</td>
</tr>
<tr>
<td>32</td>
<td>2356</td>
<td>Quarries/Strip Mines/Gravel Pits</td>
<td>0.0021</td>
</tr>
<tr>
<td>92</td>
<td>509</td>
<td>Emergent Herbaceous Wetlands</td>
<td>0.0005</td>
</tr>
<tr>
<td>31</td>
<td>245</td>
<td>Bare Rock/Sand/clay</td>
<td>0.0002</td>
</tr>
</tbody>
</table>
Thus, the percentage of forest coverage (Decidous Forest, Mixed Forest, Evergreen Forest, Woody Wetlands) of the watershed = $0.4379 + 0.1310 + 0.0886 + 0.0100 = 0.6675 = 66.75\%$. The watershed is a forest dominant coverage watershed.

Land cover is classified as deciduous, conifer, mixed, water, crop-grass and others in the database. Woody wetland was classified as deciduous.

37. Open the attribute table of theme Gawshedcov and edit the table. Create a new field named wshedArea and keep others as default. Click menu “Field/Calculate…”, in the Field Calculator dialog window, type in the following expression to obtain the area of the watershed:

\[ \text{[Shape].returnarea} \]

Then stop editing and save edits and close the attribute table. The area of the watershed = $1014146100 \text{ m}^2 = 1014.14 \text{ km}^2 = 391.57 \text{ mi}^2$. It is approximately 392 $\text{mi}^2$, which is the same as drainage area in the USGS gauging station dataset.
Appendix C. Summary of the gauging stations in the study watersheds

Table 7.3. Summary of the gauging stations in the GIS database

<table>
<thead>
<tr>
<th>Site number</th>
<th>Site name</th>
<th>Drainage area (mi²)</th>
<th>Elevation (m)</th>
<th>Year</th>
<th>Record period</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Streamflow Precip., Tmax, Tmin</th>
<th>Solar radiation, RH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TrentR, NC</td>
<td>168</td>
<td>30</td>
<td>29</td>
<td>miss 90</td>
<td>35.07</td>
<td>77.46</td>
<td>35.07/35.22</td>
<td>77.78/77.9</td>
</tr>
<tr>
<td>2</td>
<td>Pocetas, NC</td>
<td>225</td>
<td>24</td>
<td>30</td>
<td>61-90</td>
<td>36.37</td>
<td>77.03</td>
<td>36.4</td>
<td>77.42</td>
</tr>
<tr>
<td>3</td>
<td>Littlef, NC</td>
<td>177</td>
<td>47</td>
<td>30</td>
<td>61-90</td>
<td>36.19</td>
<td>77.88</td>
<td>36.28</td>
<td>77.98</td>
</tr>
<tr>
<td>4</td>
<td>Eno, NC</td>
<td>141</td>
<td>193</td>
<td>27</td>
<td>64-90</td>
<td>36.07</td>
<td>78.91</td>
<td>36.03</td>
<td>78.97</td>
</tr>
<tr>
<td>5</td>
<td>FLATR, NC</td>
<td>149</td>
<td>135</td>
<td>30</td>
<td>61-90</td>
<td>36.18</td>
<td>78.88</td>
<td>36.32</td>
<td>78.9</td>
</tr>
<tr>
<td>6</td>
<td>Drown, NC</td>
<td>183</td>
<td>149</td>
<td>30</td>
<td>61-90</td>
<td>35.06</td>
<td>79.49</td>
<td>35.22</td>
<td>79.73</td>
</tr>
<tr>
<td>7</td>
<td>Hunting, NC</td>
<td>155</td>
<td>322</td>
<td>30</td>
<td>61-90</td>
<td>36.00</td>
<td>80.75</td>
<td>35.81/36.82</td>
<td>80.88/81.19</td>
</tr>
<tr>
<td>8</td>
<td>Fisher, NC</td>
<td>128</td>
<td>322</td>
<td>30</td>
<td>61-90</td>
<td>36.36</td>
<td>80.69</td>
<td>36.48</td>
<td>80.67</td>
</tr>
<tr>
<td>9</td>
<td>New River, NC</td>
<td>205</td>
<td>955</td>
<td>29</td>
<td>Miss 87</td>
<td>36.39</td>
<td>81.41</td>
<td>36.25/36.15</td>
<td>81.26/81.7</td>
</tr>
<tr>
<td>10</td>
<td>French, NC</td>
<td>296</td>
<td>999</td>
<td>30</td>
<td>61-90</td>
<td>35.30</td>
<td>82.62</td>
<td>35.11/35.27</td>
<td>82.82/82.7</td>
</tr>
<tr>
<td>11</td>
<td>LiTenne, NC</td>
<td>140</td>
<td>897</td>
<td>30</td>
<td>61-90</td>
<td>35.15</td>
<td>83.38</td>
<td>35.07</td>
<td>83.43</td>
</tr>
<tr>
<td>12</td>
<td>Ocunalu, NC</td>
<td>184</td>
<td>650</td>
<td>30</td>
<td>61-90</td>
<td>35.46</td>
<td>83.35</td>
<td>35.52</td>
<td>83.31</td>
</tr>
<tr>
<td>13</td>
<td>al03140303</td>
<td>176</td>
<td>109</td>
<td>16</td>
<td>75-90</td>
<td>31.40</td>
<td>87.00</td>
<td>31.1</td>
<td>87.1</td>
</tr>
<tr>
<td>14</td>
<td>al03150203</td>
<td>97.5</td>
<td>70</td>
<td>30</td>
<td>61-90</td>
<td>32.00</td>
<td>87.60</td>
<td>31.9</td>
<td>97.7</td>
</tr>
<tr>
<td>15</td>
<td>ar11010001</td>
<td>400</td>
<td>432</td>
<td>27</td>
<td>64-90</td>
<td>36.10</td>
<td>94.20</td>
<td>36.1</td>
<td>94.2</td>
</tr>
<tr>
<td>16</td>
<td>fi03120003</td>
<td>102</td>
<td>41</td>
<td>26</td>
<td>65-90</td>
<td>30.20</td>
<td>84.60</td>
<td>30.4</td>
<td>84.4</td>
</tr>
<tr>
<td>17</td>
<td>ky05070203</td>
<td>206</td>
<td>312</td>
<td>30</td>
<td>61-90</td>
<td>37.60</td>
<td>82.50</td>
<td>38.5</td>
<td>82.6</td>
</tr>
<tr>
<td>18</td>
<td>ky05100203</td>
<td>722</td>
<td>358</td>
<td>30</td>
<td>61-90</td>
<td>37.20</td>
<td>83.70</td>
<td>37.6/36.6</td>
<td>84.3/83.7</td>
</tr>
<tr>
<td>19</td>
<td>la08070202</td>
<td>145</td>
<td>56</td>
<td>30</td>
<td>61-90</td>
<td>30.80</td>
<td>91.10</td>
<td>30.5/31.1</td>
<td>91.1/91.2</td>
</tr>
<tr>
<td>20</td>
<td>ms03170002</td>
<td>918</td>
<td>99</td>
<td>30</td>
<td>61-90</td>
<td>32.20</td>
<td>88.80</td>
<td>32.8/32.3</td>
<td>88.1/89.5</td>
</tr>
<tr>
<td>21</td>
<td>ms03170009</td>
<td>96.1</td>
<td>21</td>
<td>30</td>
<td>61-90</td>
<td>30.50</td>
<td>89.30</td>
<td>30.3</td>
<td>89.38</td>
</tr>
<tr>
<td>22</td>
<td>ms03180002</td>
<td>3171</td>
<td>110</td>
<td>30</td>
<td>61-90</td>
<td>32.40</td>
<td>90.20</td>
<td>32.6/32</td>
<td>90.0/90.4</td>
</tr>
<tr>
<td>23</td>
<td>ms08060203</td>
<td>654</td>
<td>83</td>
<td>29</td>
<td>62-90</td>
<td>32.10</td>
<td>90.90</td>
<td>31.95/32</td>
<td>91.2/91.0</td>
</tr>
<tr>
<td>24</td>
<td>tn06010204</td>
<td>1987</td>
<td>576</td>
<td>30</td>
<td>61-90</td>
<td>35.60</td>
<td>84.10</td>
<td>35</td>
<td>84.4</td>
</tr>
<tr>
<td>25</td>
<td>tn06040004</td>
<td>447</td>
<td>238</td>
<td>30</td>
<td>61-90</td>
<td>35.60</td>
<td>87.90</td>
<td>35.3</td>
<td>87.8</td>
</tr>
</tbody>
</table>
Table 7.3 (continued)

<table>
<thead>
<tr>
<th>Site number</th>
<th>Site name</th>
<th>Drainage area (mi²)</th>
<th>Elevation (m)</th>
<th>Year Record period</th>
<th>Streamflow</th>
<th>Precip., Tmax, Tmin</th>
<th>Solar radiation, RH</th>
</tr>
</thead>
<tbody>
<tr>
<td>26</td>
<td>tx12030201</td>
<td>142</td>
<td>112</td>
<td>22 69-90</td>
<td>31.90</td>
<td>96.40</td>
<td>32.1/31.7</td>
</tr>
<tr>
<td>27</td>
<td>tx12040103</td>
<td>325</td>
<td>62</td>
<td>30 61-90</td>
<td>30.40</td>
<td>95.10</td>
<td>30.2/30.1</td>
</tr>
<tr>
<td>28</td>
<td>va02080201</td>
<td>329</td>
<td>633</td>
<td>30 61-90</td>
<td>38.00</td>
<td>80.00</td>
<td>38</td>
</tr>
<tr>
<td>29</td>
<td>va05050002</td>
<td>223</td>
<td>760</td>
<td>30 61-90</td>
<td>37.30</td>
<td>80.60</td>
<td>37.1</td>
</tr>
<tr>
<td>30</td>
<td>ga03130005</td>
<td>272</td>
<td>213</td>
<td>30 61-90</td>
<td>33.23</td>
<td>84.42</td>
<td>33.27</td>
</tr>
<tr>
<td>31</td>
<td>ga03070103</td>
<td>182</td>
<td>205</td>
<td>30 61-90</td>
<td>32.80</td>
<td>83.75</td>
<td>32.68</td>
</tr>
<tr>
<td>32</td>
<td>ga03070101</td>
<td>392</td>
<td>270</td>
<td>30 61-90</td>
<td>33.93</td>
<td>83.42</td>
<td>33.95</td>
</tr>
<tr>
<td>33</td>
<td>sc03050110</td>
<td>59.6</td>
<td>72</td>
<td>24 67-90</td>
<td>33.99</td>
<td>80.97</td>
<td>34</td>
</tr>
<tr>
<td>34</td>
<td>Bradford, FL</td>
<td>0.54</td>
<td>44</td>
<td>13 78-90</td>
<td>29.50</td>
<td>82.11</td>
<td>29.75</td>
</tr>
<tr>
<td>35</td>
<td>Walker Branch, TN</td>
<td>0.39</td>
<td>308</td>
<td>22 69-90</td>
<td>35.97</td>
<td>84.28</td>
<td>35.97</td>
</tr>
<tr>
<td>36</td>
<td>Coles Fork, KY</td>
<td>6.41</td>
<td>378</td>
<td>18 73-90</td>
<td>37.45</td>
<td>83.13</td>
<td>37.45</td>
</tr>
<tr>
<td>37</td>
<td>Carteret, NC</td>
<td>0.1</td>
<td>3</td>
<td>13 88-00</td>
<td>34.80</td>
<td>76.70</td>
<td>34.80/34.73</td>
</tr>
<tr>
<td>38</td>
<td>Parker, NC</td>
<td>11.39</td>
<td>6</td>
<td>5 96-00</td>
<td>35.83</td>
<td>76.75</td>
<td>35.83/35.87</td>
</tr>
<tr>
<td>39</td>
<td>Santee-80, SC</td>
<td>0.58</td>
<td>7</td>
<td>5 76-80</td>
<td>33.00</td>
<td>80.00</td>
<td>33.00</td>
</tr>
</tbody>
</table>
Appendix D. Computer program for six PET methods in Visual Basic 6.0

'Function calculate monthly Thornthwaite PET

Private Function monThornPET(wsLat As Double, Ivalue As Double, year As Integer, month As Integer, temp As Double) As Double
    Dim Ld As Double, Jday As Integer, aValue As Double
    Jday = JdayInMonth(year, month)
    Ld = daylength(Jday, wsLat)
    aValue = 6.75 * (10 ^ (-7)) * (Ivalue ^ 3) - 7.71 * (10 ^ (-5)) * (Ivalue ^ 2) + 0.01792 * Ivalue + 0.49239
    When temperature is below 0, PET = 0
    If temp > 0 Then
        monThornPET = 10 * 1.6 * Ld * ((10 * temp / Ivalue) ^ aValue)
    Else
        monThornPET = 0
    End If
End Function

****************************************************************************

'Function calculates monthly Hamon PET

Private Function monHamonPET(pec As Double, wsLat As Double, year As Integer, month As Integer, temp As Double) As Double
    Dim Jday As Integer, dayLen As Double
    Dim Esat As Double, RHOSAT As Double
    Jday = JdayInMonth(year, month) 'average Jday in a month, which is the middle julian day in month
dayLen = daylength(Jday, wsLat)
    Esat = 6.108 * Exp(17.26939 * temp / (temp + 237.3))
    RHOSAT = 216.7 * Esat / (temp + 273.3)
    monHamonPET = monthDay(year, month) * pec * 0.1651 * dayLen * RHOSAT 'Monthly Hamon PET Value
End Function

*******************************************************************************

'Function calculates monthly Turc PET

Private Function monTurcPET(year As Integer, month As Integer, temp As Double, Rs As Double, RH As Double) As Double
    If temp > 0 Then
        If RH < 0.5 Then 'PET depents on RH, assume each month has 30 days 'MJ/m^2.day is transfer to cal/cm^2.day: MJ*100/4.1868 = cal
            monTurcPET = monthDay(year, month) * 0.013 * (temp / (temp + 15)) * ((Rs * 100 / 4.1868) + 50) * (1 + (50 - RH) / 70)
        Else
            monTurcPET = monthDay(year, month) * 0.013 * (temp / (temp + 15)) * ((Rs * 100 / 4.1868) + 50)
        End If 'If block for RH
    Else
        monTurcPET = 0
End If 'if block for temp > 0
End Function

***************************************************************************

'Function calculates monthly Makkink PET

Private Function monMakkPET(Pvalue As Double, year As Integer, month As Integer, temp As Double, Rs As Double) As Double
Dim Lambda As Double, Delta As Double, Gamma As Double
Lambda = 2.501 - 0.002361 * temp
Delta = 0.2 * ((0.00738 * temp + 0.8072) ^ 7) - 0.000116
Gamma = 0.001013 * Pvalue / (0.622 * Lambda)
monMakkPET = monthDay(year, month) * 0.25 * (Delta / (Delta + Gamma)) * Rs - 0.12
End Function

***************************************************************************

'Function calculates monthly Hargreaves-Samani PET

Private Function monHSPET(year As Integer, month As Integer, temp As Double, Ra As Double, Tmax As Double, Tmin As Double) As Double
Dim Lambda As Double
Lambda = 2.501 - 0.002361 * temp
monHSPET = monthDay(year, month) * 0.0023 * Ra * (Abs(Tmax - Tmin) ^ 0.5) * (temp + 17.8) / Lambda
End Function

***************************************************************************

'Function calculates I (Annual heat index) for each site(watershed) in Thornthwaite method
'by getting long-term monthly temperature

Private Function Icalculation(Iparameter() As Double) As Boolean
Dim dummyRa As Double, dummyTmax As Double, dummyTmin As Double
Dim watershedNum As Integer, wsNumLoop As Integer
Dim dummyTitle As String, annHeatIndex As Double
Dim I As Integer, J As Integer, Nyears As Integer, year As Integer, month As Integer
Dim avgTemp(1 To 12) As Double, totalTemp(1 To 12) As Double
Dim dummyRs As Double, dummyRH As Double, dummyRn As Double
Dim lstart As Integer, rfinish As Integer
Dim pec As Double, siteElevation As Double, siteLatitude As Double
Open txtInput.Text For Input As #1
'get boundary of the array
lstart = LBound(Iparameter)
rfinish = UBound(Iparameter)
Input #1, watershedNum
If watershedNum > (rfinish - lstart) Then
    Icalculation = False
Else
    'site(watershed) loop
    For wsNumLoop = 1 To watershedNum
        'get Site info
        Line Input #1, dummyTitle

        Open txtSiteInfo.Text For Input As #1
        Input #1, dummyRa, dummyTmax, dummyTmin, dummyRs, dummyRH, dummyRn, pec, siteElevation, siteLatitude
        Input #1, annHeatIndex
        Close #1

        For I = 1 To 12
            Line Input #1, avgTemp(I)
            Input #1, totalTemp(I)
        Next I
        Close #1

        Input #1, Nyears
        For year = 1 To Nyears
            line Input #1, month
            For month = 1 To 12
                Input #1, avgTemp(I)
                Input #1, totalTemp(I)
            Next month
        Next year
        Close #1

        Input #1, year, month
        Open txtSediment.Text For Input As #1
        Input #1, sediment
        Close #1

        Open txtRunoff.Text For Input As #1
        Input #1, runoff
        Close #1

        Open txtErosion.Text For Input As #1
        Input #1, erosion
        Close #1

        Open txtAvailability.Text For Input As #1
        Input #1, availability
        Close #1

        Open txtPrecipitation.Text For Input As #1
        Input #1, precipitation
        Close #1

        Open txtEvaporation.Text For Input As #1
        Input #1, evaporation
        Close #1

        Open txtVegetation.Text For Input As #1
        Input #1, vegetation
        Close #1

        Open txtSoil.Text For Input As #1
        Input #1, soil
        Close #1

        Open txtTopography.Text For Input As #1
        Input #1, topography
        Close #1

        Open txtHumanActivity.Text For Input As #1
        Input #1, humanActivity
        Close #1

        Open txtInitialConditions.Text For Input As #1
        Input #1, initialConditions
        Close #1

        Open txtBoundaryConditions.Text For Input As #1
        Input #1, boundaryConditions
        Close #1

        Open txtInitialConditions.Text For Input As #1
        Input #1, initialConditions
        Close #1

        'calculate I
        annHeatIndex = (totalTemp - avgTemp) / year

        'write to file
        Open OutFile.Text For Output As #2
        Line Output #2, dummyTitle
        Line Output #2, annHeatIndex
        Close #2

        'close files
        Close #1, #2
    Next wsNumLoop
    Close #1
End If
End Function
' get site parameters
Input #1, Nyears, pec, siteElevation, siteLatitude
' get data title
Line Input #1, dummyTitle

'initialize Annual Heat Index for each site
annHeatIndex = 0

'initialize totalTemp(T) in each site
For I = 1 To 12
  totalTemp(I) = 0
Next I
For J = 1 To Nyears
  'get every year's input data
  For I = 1 To 12
    Input #1, year, month, avgTemp(I), dummyRs, dummyRH, dummyRn, dummyRa, dummyTmax, dummyTmin
    totalTemp(I) = totalTemp(I) + avgTemp(I)
  Next I
Next J

'get the average temperature for each month in N years and calculate I
For I = 1 To 12
  avgTemp(I) = totalTemp(I) / Nyears
  If avgTemp(I) > 0 Then
    annHeatIndex = annHeatIndex + ((avgTemp(I) / 5) ^ 1.514)
  End If
Next I
Iparameter(wsNumLoop) = annHeatIndex
Next wsNumLoop 'sitely loop
Close #1
Icalculation = True
End If ' If block to test site limitation
End Function

***************************************************************************
'Main function to calculate six monthly and annual PET

Private Sub PETcalculation()
  Dim myI(1 To 100) As Double
  'call Icaculation function to calculate parameter I
  'up to 100 watersheds as maximum input
  If (Icalculation(myI()) = False) Then
    MsgBox "Too many sites in the input file! Largest = 100, decrease the input and try again"
  Else
    'Variable Declaration Section
    Dim watershedNum As Integer, wsNumLoop As Integer
    Dim sitelInfo As String, titleLine As String
    Dim Nyear As Integer, pec As Double, wsElevation As Double
    Dim wsLat As Double, Ivalue As Double, Pvalue As Double
    Dim year As Integer, month As Integer, I As Integer, J As Integer
    Dim temp As Double, Rs As Double, RH As Double, Ra As Double, Tmax As Double, Tmin As Double
    'Rn As Double
    Dim ThornMonth(12) As Double, ThornAnn As Double, HamonMonth(12) As Double, HamonAnn As Double
    Dim TurcMonth(12) As Double, TurcAnn As Double, PTmonth As Double, PTAnn As Double
  End If
End Sub
Dim MakkMonth(12) As Double, MakkAnn As Double, HSMonth(12) As Double, HSAnn As Double
Dim annThorn As Double, annHamon As Double, annTurc As Double, annMakk As Double, annHS As Double

'additional parameters used in Priestley-Taylor method
Dim lastYear As Integer, myTemp(-1 To 12) As Double, Rn(12) As Double
Dim Lambda(12) As Double, Delta(12) As Double, Gamma(12) As Double, Gvalue(12) As Double

'Variable Declaration End

Open txtInput.Text For Input As #1
Open txtOutput.Text For Output As #2
Open txtAnnSum.Text For Output As #3

Input #1, watershedNum
For wsNumLoop = 1 To watershedNum
    Line Input #1, siteInfo
    Print #2, siteInfo
    Print #3, siteInfo

    Input #1, Nyear, pec, wsElevation, wsLat
    'each site has one Pvalue and Ivalue
    Pvalue = 101.3 - 0.01055 * wsElevation
    Ivalue = myI(wsNumLoop)
    Line Input #1, titleLine

    siteInfo = "Monthly PET Calculated by 6 PET methods"
    Print #2, siteInfo
    Write #2, "Year", "Month", "Thorn", "Hamon", "Turc", "PT", "Makkind", "HS"

    siteInfo = "Annaul PET Calculated by 6 PET methods"
    Print #3, siteInfo
    Write #3, "Year", "Thorn", "Hamon", "Turc", "PT", "Makkind", "HS"
    PTAnn = 0

    'Year loop
    For I = 1 To Nyear
        'Initialize annual PETs = 0
        ThornAnn = 0
        HamonAnn = 0
        TurcAnn = 0
        MakkAnn = 0
        HSAnn = 0

        'Monthly loop - get input data and calculation PET in each month for 5 PETs
        'Priestley-Taylor PET will calculate later
        For J = 1 To 12
            Input #1, year, month, myTemp(J), Rs, RH, Rn(J), Ra, Tmax, Tmin

            'calculate other 5 PETs
            temp = myTemp(J)
            ThornMonth(J) = monThornPET(wsLat, Ivalue, year, month, temp)
            HamonMonth(J) = monHamonPET(pec, wsLat, year, month, temp)
            TurcMonth(J) = monTurcPET(year, month, temp, Rs, RH)
            MakkMonth(J) = monMakkPET(Pvalue, year, month, temp, Rs)
            HSMonth(J) = monHSPET(year, month, temp, Ra, Tmax, Tmin)

            'add up annual values
            ThornAnn = ThornAnn + ThornMonth(J)
            HamonAnn = HamonAnn + HamonMonth(J)
            TurcAnn = TurcAnn + TurcMonth(J)
            MakkAnn = MakkAnn + MakkMonth(J)
HSAnn = HSAnn + HSMonth(J)
'calculate each month parameter
Lambda(J) = 2.501 - 0.002361 * temp
Delta(J) = 0.2 * ((0.00738 * temp + 0.8072)^7) - 0.000116
Gamma(J) = 0.001013 * Pvalue / (0.622 * Lambda(J))
Next J
'If the site has only one year data
If Nyear = 1 Then
  'a loop to calculate G value for each month
  'G = 4.19(Ti+1 - Ti-1)/60
  'for Feb to November, and Jan uses Feb and December uses November values
  For J = 2 To 11
    Gvalue(J) = 4.19 * (myTemp(J + 1) - myTemp(J - 1)) / 60
  Next J
  Gvalue(1) = Gvalue(2)
  Gvalue(12) = Gvalue(11)
'calculate PTmonth and print 6 PETs out
  For J = 1 To 12
    PTmonth = monthDay(year, J) * 1.26 * Delta(J) * (Rn(J) - Gvalue(J)) / (Lambda(J) * (Delta(J)
    + Gamma(J)))
  Next J
  PTAnn = PTAnn + PTmonth
  Write #2, year, J, Round(ThornMonth(J)), Round(HamonMonth(J)), Round(TurcMonth(J)),
  Round(PTmonth(J)), Round(MakkMonth(J)), Round(HSMonth(J))
'when the site have many years data
Else
  'for the first year but not the final year in the data set
  If I = 1 Then
    'calculate first 11 months Gvalue
    For J = 2 To 11
      Gvalue(J) = 4.19 * (myTemp(J + 1) - myTemp(J - 1)) / 60
    Next J
    'Jan uses Feb G value for the first year
    Gvalue(1) = Gvalue(2)
    'leave Gvalue(12) uncalculated
    'Gvalue(12) = Gvalue(11)
    For J = 1 To 11
      PTmonth = monthDay(year, J) * 1.26 * Delta(J) * (Rn(J) - Gvalue(J)) / (Lambda(J) * (Delta(J)
      + Gamma(J)))
  Next J
  PTAnn = PTAnn + PTmonth
  Write #2, year, J, Round(ThornMonth(J)), Round(HamonMonth(J)), Round(TurcMonth(J)),
  Round(PTmonth(J)), Round(MakkMonth(J)), Round(HSMonth(J))
  Next J
  'save December's data&parameters
  Else
    'begin with last year December
    '0 is the 12 for the last year, and -1 represents November
    For J = 0 To 11
      Gvalue(J) = 4.19 * (myTemp(J + 1) - myTemp(J - 1)) / 60
      PTmonth = monthDay(year, J) * 1.26 * Delta(J) * (Rn(J) - Gvalue(J)) / (Lambda(J) * (Delta(J)
      + Gamma(J)))
    Next J
    PTAnn = PTAnn + PTmonth
  The December for the last Year
  If J = 0 Then
Write #2, lastYear, 12, Round(ThornMonth(J)), Round(HamonMonth(J)), Round(TurcMonth(J)), Round(PTmonth), Round(MakkMonth(J)), Round(HSMonth(J))
Write #3, lastYear, Round(annThorn), Round(annHamon), Round(annTurc), Round(PTAnn), Round(annMakk), Round(annHS)
   PTAnn = 0 'reset Priestley-Taylor annual PET
   Else
      Write #2, year, J, Round(ThornMonth(J)), Round(HamonMonth(J)), Round(TurcMonth(J)), Round(PTmonth), Round(MakkMonth(J)), Round(HSMonth(J))
   End If 'December If
Next J
' The final year for the input data, output the December's PET and Annual PETs directly
If I = Nyear Then
   'December uses the November's value for the G value
   PTmonth = 31 * 1.26 * Delta(12) * (Rn(12) - Gvalue(11)) / (Lambda(12) * (Delta(12) + Gamma(12)))
   PTAnn = PTAnn + PTmonth
   Write #2, year, 12, Round(ThornMonth(J)), Round(HamonMonth(J)), Round(TurcMonth(J)), Round(PTmonth), Round(MakkMonth(J)), Round(HSMonth(J))
   Write #3, year, Round(ThornAnn), Round(HamonAnn), Round(TurcAnn), Round(PTAnn), Round(MakkAnn), Round(HSAnn)
End If 'If for the last year
End If 'If for only-one-year/many-year in the site
Next I 'year loop
Next wsNumLoop 'site(watershed) loop
Close #1: Close #2: Close #3
End If 'If for test >100 watersheds a time
End Sub

***************************************************************************
' Sub averages monthly input to get annual data

' Save November temperature to calculate last year December G value in next loop
myTemp(-1) = myTemp(11)
myTemp(0) = myTemp(12)
Rn(0) = Rn(12)
Lambda(0) = Lambda(12)
Delta(0) = Delta(12)
Gamma(0) = Gamma(12)
lastYear = year
' Save PET values
ThornMonth(0) = ThornMonth(12)
HamonMonth(0) = HamonMonth(12)
TurcMonth(0) = TurcMonth(12)
' PTmonth = monPTPET(Pvalue, year, month, temp, Rn)
MakkMonth(0) = MakkMonth(12)
HSMonth(0) = HSMonth(12)
' Save annual data
annThorn = ThornAnn
annHamon = HamonAnn
annTurc = TurcAnn
' PTAnn = PTAnn + monPTMonth
annMakk = MakkAnn
annHS = HSAnn
End If 'If for only-one-year/many-year in the site

Next I 'year loop
Next wsNumLoop 'site(watershed) loop
Close #1: Close #2: Close #3
End If 'If for test >100 watersheds a time
End Sub

***************************************************************************

' Sub averages monthly input to get annual data
Private Sub annualSum()
  Dim temp As Double, Rs As Double, RH As Double
  Dim Rn As Double, Ra As Double, Tmax As Double, Tmin As Double
  Dim sumTemp As Double, sumRs As Double, sumRH As Double
  Dim sumRn As Double, sumRa As Double, sumTmax As Double, sumTmin As Double
  Dim watershedNum As Integer, wsNumLoop As Integer
  Dim siteElevation As Double
  Dim Nyear As Integer
  Dim siteInfo As String, titleLine As String
  Dim dayLen As Double, pec As Double, siteLat As Double
  Dim I As Integer, J As Integer, year As Integer, month As Integer
  Open txtInput.Text For Input As #1
  Open txtAnnSum.Text For Output As #3
  Input #1, watershedNum
  For wsNumLoop = 1 To watershedNum
    Line Input #1, siteInfo$
    Print #3, siteInfo$
    Input #1, Nyear, pec, siteElevation, siteLat
    Line Input #1, siteInfo$
    Print #3, "Summarize Input data to annual monthly"
    Write #3, "Year", "Temp", "Rs", "RH", "Rn", "Ra", "Tmax", "Tmin"
    'Year loop
    For I = 1 To Nyear
      sumTemp = 0
      sumRs = 0
      sumRH = 0
      sumRn = 0
      sumRa = 0
      sumTmax = 0
      sumTmin = 0
      'Monthly loop - 12 months/year
      For J = 1 To 12
        Input #1, year, month, temp, Rs, RH, Rn, Ra, Tmax, Tmin
        sumTemp = sumTemp + temp
        sumRs = sumRs + Rs
        sumRH = sumRH + RH
        sumRn = sumRn + Rn
        sumRa = sumRa + Ra
        sumTmax = sumTmax + Tmax
        sumTmin = sumTmin + Tmin
      Next J 'monthly loop
      'average data
      sumTemp = sumTemp / 12
      sumRs = sumRs / 12
      sumRH = sumRH / 12
      sumRn = sumRn / 12
      sumRa = sumRa / 12
      sumTmax = sumTmax / 12
      sumTmin = sumTmin / 12
      Write #3, year, Round(sumTemp, 2), Round(sumRs, 2), Round(sumRH, 2), Round(sumRn, 2),
        Round(sumRa, 2), Round(sumTmax, 2), Round(sumTmin, 2)
    Next I 'yearly loop
  Next wsNumLoop 'sitely loop
Private Sub longtermAnnualSum()
    Dim I As Integer, J As Integer, year As Integer, siteID As Integer 'Month As Integer
    Dim siteNum As Integer, currID As Integer, preID As Integer, yearCount As Integer
    Dim myData(1 To 16) As Double, myDataSum(1 To 16) As Double, myDataPre(1 To 16) As Double
    Dim dataTitle As String
    siteNum = 0
    preID = 0
    currID = 0

    For I = 1 To 16
        myDataPre(I) = 0
    Next I

    Open txtInput.Text For Input As #1
    Open txtAnnSum.Text For Output As #3

    Line Input #1, dataTitle
    Write #3, dataTitle

    Do Until EOF(1)
        'initiate Summary data
        yearCount = 0
        For I = 1 To 16
            myDataSum(I) = 0
        Next I

        Do While (preID = currID And Not EOF(1))
            Input #1, currID, year, myData(1), myData(2), myData(3), myData(4), myData(5), myData(6),
            myData(7), myData(8), myData(9), myData(10), myData(11), myData(12), myData(13), myData(14),
            myData(15), myData(16)
            yearCount = yearCount + 1
        End If

        If preID = currID Then
            For I = 1 To 16
                myDataSum(I) = myDataSum(I) + myDataPre(I)
            Next I

        End If

        'set the ID and add the last data after read first record
        If yearCount = 1 Then
            preID = currID
            siteNum = 1

        End If

    If preID = currID Then

End Sub
For I = 1 To 16
    myDataSum(I) = myDataSum(I) + myData(I)
    Next I
Else
' save the read data, it will be added in the next loop
    For I = 1 To 16
        myDataPre(I) = myData(I)
    Next I
End If
Loop 'while loop

' adjust yearCount to correct number due to the loop property (the first and the last loop)
    siteNum = siteNum + 1
    If siteNum = 2 Then
        yearCount = yearCount - 1
    End If
    If EOF(1) Then
        yearCount = yearCount + 1
    End If
siteID = preID
    preID = currID
    For J = 1 To 16
        myData(J) = myDataSum(J) / yearCount
    Next J
    Write #3, siteID, yearCount, Round(myData(1), 2), Round(myData(2), 2), Round(myData(3), 2),
    Round(myData(4), 2), Round(myData(5), 2), Round(myData(6), 2), Round(myData(7), 2), Round(myData(8),
    2), Round(myData(9), 2), Round(myData(10), 2), Round(myData(11), 2), Round(myData(12), 2),
    Round(myData(13), 2), Round(myData(14), 2), Round(myData(15), 2), Round(myData(16), 2)
Loop 'until loop

dataTitle = "your have " & siteNum - 1 & " watersheds in your output data!"
MsgBox (dataTitle)
Close #1: Close #3
End Sub

***************************************************************************

' Function to test whether a year is a leap year by specifying the year
' A year is a Leap year (Feb has 29 days and a year has 366 days) if it is divisible
' by 4 but not by 100, except that years divisible by 400 are leap year

Private Function isLeapYear(year As Integer) As Boolean
    Dim fourDivident As Integer, hundredDivident As Integer, fHunDivident As Integer

    fourDivident = year Mod 4
    hundredDivident = year Mod 100
    fHunDivident = year Mod 400
    ' if the year is divisible by 400
    If fHunDivident = 0 Then
        isLeapYear = True
        ' if the year is divisible by 4 but not 100
        ElseIf fourDivident = 0 And hundredDivident <> 0 Then
            isLeapYear = True
        Else
            isLeapYear = False
        End If
End Function
End Function

***************************************************************************
'Function to calculate number of days in a month by
'specify the year and the month

Private Function monthDay(myYear As Integer, myMonth As Integer) As Integer
    Dim myLDay As Integer
    'if the year is a leap year, Feb will have 29 days, otherwise 28
    If isLeapYear(myYear) Then
        myLDay = 1
    Else
        myLDay = 0
    End If

    If myMonth = 2 Then
        monthDay = 28 + myLDay
    ElseIf myMonth = 4 Or myMonth = 6 Or myMonth = 9 Or myMonth = 11 Then
        monthDay = 30
    Else
        monthDay = 31
    End If
End Function

***************************************************************************
'Function calculates
'the average Julian day for a particular month in a year

Private Function JdayInMonth(myYear As Integer, month As Integer) As Integer
    Dim leapDay As Integer
    If isLeapYear(myYear) Then
        leapDay = 1
    Else
        leapDay = 0
    End If

    If month = 1 Then
        JdayInMonth = 16
    ElseIf month = 2 Then
        JdayInMonth = 45 + leapDay
    ElseIf month = 3 Then
        JdayInMonth = 75 + leapDay
    ElseIf month = 4 Then
        JdayInMonth = 106 + leapDay
    ElseIf month = 5 Then
        JdayInMonth = 136 + leapDay
    ElseIf month = 6 Then
        JdayInMonth = 167 + leapDay
    ElseIf month = 7 Then
        JdayInMonth = 197 + leapDay
    ElseIf month = 8 Then
        JdayInMonth = 228 + leapDay
    ElseIf month = 9 Then
        JdayInMonth = 258 + leapDay
    End If
End Function
ElseIf month = 10 Then  
  JdayInMonth = 289 + leapDay
ElseIf month = 11 Then  
  JdayInMonth = 320 + leapDay
Else  
  JdayInMonth = 350 + leapDay
End If  
End Function

***************************************************************************  
'This function and it associated function ZCos ( ) were programmed by  
'Dr. Ge Sun at Department of Forestry, North Carolina State University

Private Function daylength (DOY As Integer, Lat As Double) As Double
'################################################  
'   daylength and horizontal potential insolation for day  
'   and Sellers (1965) Physical Climatology  
'   DOY day of the year  
'   LAT latitude, radians  
'   I0 potential insolation (horizontal) J/m2  
'   DECL declination of the sun, radians  
'   R radius vector of the sun, dimensionless  
'   h half daylength  
'   ISC solar constant, W/m2  
'   Half daylength and declination equations from:  
'################################################  
Dim LatRad, Pi, r, Z, Decl, h, z2 As Double  
Dim hr As Double

LatRad = Lat * (2 * 3.1416) / 360
Pi = 3.14159265
r = 1! - 0.0167 * Cos(0.0172 * (DOY - 3))  
         'radius vector of the sun
Z = 0.39785 * Sin(4.868961 + 0.017203 * DOY + 0.033446 * Sin(6.224111 + 0.017202 * DOY))  
         'temporary variable
If Abs(Z) < 0.7 Then  
  Decl = Atn(Z / (Sqr(1 - Z ^ 2)))
Else  
  Decl = Pi / 2 - Atn(Sqr(1 - Z ^ 2) / Z)
End If
If Abs(LatRad) >= Pi / 2 Then LatRad = Sgn(Lat) * (Pi / 2 - 0.01)
  z2 = -Tan(Decl) * Tan(LatRad)  
         'temporary variable
If z2 >= 1 Then  
  'sun stays below horizon
  h = 0
ElseIf z2 <= -1 Then  
  'sun stays above the horizon
  h = Pi
Else  
  'call zCos() function
  h = ZCos(z2)
End If

    hr = 2 * (h * 24) / (2 * 3.1416)       ' length of day in hours
    daylength = hr / 12
End Function

***************************************************************************

'Function assists to calculate day length

Private Function ZCos (T) As Double
    Dim TA, AC
    TA = Abs(T)
    If (TA > 1!) Then
        'Print "|arg| for arccos > 1"
        Stop
    End If
    If (TA < 0.7) Then
        AC = 1.570796 - Atn(TA / Sqr(1 - TA * TA))
    Else
        AC = Atn(Sqr(1 - TA * TA) / TA)
    End If
    If T < 0 Then
        ZCos = 3.141593 - AC
    Else
        ZCos = AC
    End If
End Function

***************************************************************************
Appendix E. Formats of inputs and outputs of the PET program

1) Formats of inputs and outputs of PET calculations and annual summaries

* The formats of the input file is as following, the values are in average daily.

Number of watersheds in the file
Site Information (for watershed #1)
Number of Years, Pec, elevation, latitude
Year Month Temp Rs(MJ/m².d) RH(decimal) Rn(MJ/m².d) Ra(MJ/m².d) Tmax Tmin
...
Site Information (for watershed #2)
Number of Years, Pec, elevation, latitude
Year Month Temp Rs (MJ/m².d) RH (decimal) Rn (MJ/m².d)
...

Example,

<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Temp</th>
<th>Rs</th>
<th>RH</th>
<th>Rn</th>
<th>Ra</th>
<th>Tmax</th>
<th>Tmin</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961</td>
<td>1</td>
<td>-0.65</td>
<td>10.24</td>
<td>0.65</td>
<td>3.75</td>
<td>18.7</td>
<td>7.8</td>
<td>-9.1</td>
</tr>
<tr>
<td>1961</td>
<td>2</td>
<td>6.4</td>
<td>10.95</td>
<td>0.74</td>
<td>4.66</td>
<td>23.58</td>
<td>14.1</td>
<td>-1.3</td>
</tr>
<tr>
<td>1961</td>
<td>3</td>
<td>9.1</td>
<td>15.69</td>
<td>0.69</td>
<td>7.83</td>
<td>30.21</td>
<td>16.8</td>
<td>1.4</td>
</tr>
<tr>
<td>1961</td>
<td>4</td>
<td>9.45</td>
<td>19.81</td>
<td>0.62</td>
<td>10.78</td>
<td>36.24</td>
<td>16.5</td>
<td>2.4</td>
</tr>
<tr>
<td>1961</td>
<td>5</td>
<td>14.8</td>
<td>21.37</td>
<td>0.72</td>
<td>12.29</td>
<td>40.21</td>
<td>22</td>
<td>7.6</td>
</tr>
<tr>
<td>1961</td>
<td>6</td>
<td>18.85</td>
<td>21.52</td>
<td>0.78</td>
<td>12.83</td>
<td>41.8</td>
<td>25.6</td>
<td>12.1</td>
</tr>
<tr>
<td>1961</td>
<td>7</td>
<td>20.9</td>
<td>23.27</td>
<td>0.78</td>
<td>13.98</td>
<td>41.02</td>
<td>27.8</td>
<td>14</td>
</tr>
<tr>
<td>1961</td>
<td>8</td>
<td>21.1</td>
<td>18.85</td>
<td>0.82</td>
<td>11.1</td>
<td>37.78</td>
<td>28.1</td>
<td>14.1</td>
</tr>
<tr>
<td>1961</td>
<td>9</td>
<td>19.8</td>
<td>18.01</td>
<td>0.8</td>
<td>9.91</td>
<td>32.49</td>
<td>27.1</td>
<td>12.5</td>
</tr>
<tr>
<td>1961</td>
<td>10</td>
<td>11.3</td>
<td>15.16</td>
<td>0.75</td>
<td>6.92</td>
<td>26.08</td>
<td>21.2</td>
<td>1.4</td>
</tr>
<tr>
<td>1961</td>
<td>11</td>
<td>9.4</td>
<td>9.9</td>
<td>0.73</td>
<td>3.6</td>
<td>20.11</td>
<td>17.6</td>
<td>1.2</td>
</tr>
<tr>
<td>1961</td>
<td>12</td>
<td>3.35</td>
<td>8.08</td>
<td>0.72</td>
<td>2.48</td>
<td>17.16</td>
<td>9.8</td>
<td>-3.1</td>
</tr>
<tr>
<td>1962</td>
<td>1</td>
<td>1.95</td>
<td>8.07</td>
<td>0.77</td>
<td>2.86</td>
<td>18.7</td>
<td>9</td>
<td>-5.1</td>
</tr>
<tr>
<td>1962</td>
<td>2</td>
<td>6.2</td>
<td>11.14</td>
<td>0.72</td>
<td>4.73</td>
<td>23.58</td>
<td>14.6</td>
<td>-2.2</td>
</tr>
<tr>
<td>1962</td>
<td>3</td>
<td>6.95</td>
<td>14.97</td>
<td>0.66</td>
<td>7.49</td>
<td>30.21</td>
<td>13.6</td>
<td>0.3</td>
</tr>
<tr>
<td>1962</td>
<td>4</td>
<td>10.2</td>
<td>20.75</td>
<td>0.63</td>
<td>11.29</td>
<td>36.24</td>
<td>18</td>
<td>2.4</td>
</tr>
</tbody>
</table>

* PET Calculations monthly and annual output

* Monthly output

2092500 -1201 TrentRiver
Monthly PET Calculated by 6 PET methods
"Year", "Month", "Thorn", "Hamon", "Turc", "PT", "Makkind", "HS"
1961,1,0.23,0.17,31,38
1961,2,19,37,34,27,38,58
1961,3,33,54,65,63,66,92
1961,4,38,58,79,84,81,103
1961,5,74,90,112,112,102,144
1961,6,104,115,122,125,108,158
1961,7,117,132,142,150,125,172
<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Thorn</th>
<th>Hamon</th>
<th>Turc</th>
<th>PT</th>
<th>Makkind</th>
<th>HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1961</td>
<td>1</td>
<td>27</td>
<td>11</td>
<td>16</td>
<td>32</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td>1961</td>
<td>2</td>
<td>13</td>
<td>36</td>
<td>31</td>
<td>24</td>
<td>36</td>
<td>55</td>
</tr>
<tr>
<td>1961</td>
<td>3</td>
<td>36</td>
<td>61</td>
<td>69</td>
<td>60</td>
<td>65</td>
<td>100</td>
</tr>
<tr>
<td>1961</td>
<td>4</td>
<td>41</td>
<td>66</td>
<td>85</td>
<td>82</td>
<td>117</td>
<td></td>
</tr>
</tbody>
</table>

* Annual output

2092500 - 1201 TrentRiver
Annual PET Calculated by 6 PET methods
"Year","Thorn","Hamon","Turc","PT","Makkind","HS"
1961,674,873,911,928,880,1251
1962,699,904,902,936,870,1292
1963,634,835,844,908,855,1284
1964,632,840,829,871,829,1276
1965,636,846,863,912,861,1307

* Annual Summary output

2092500 - 1201 TrentRiver
Summarize Input data to annual monthly
"Year","Temp","Rs","RH","Rn","Ra","Tmax","Tmin"
1961,11.98,16.07,.73,8.34,30.45,19.53,4.43
1962,12.17,15.67,.73,8.22,30.45,19.78,4.56
1963,10.75,15.94,.7,8.28,30.45,19.09,2.4
1964,11.14,15.41,.73,7.98,30.43,19.28,3

…
2) The format of input and output of long-term annual summary

* Input

SiteID Year Rainfall Streamflow AET Temp Rs RH Rn Ra Tmax Tmin Thorn Hamon Turc PT Makk HS
(Actually, it can summarize any data that is in the format as: siteID, year, 16 numeric data items)

Example,

<table>
<thead>
<tr>
<th>SiteID</th>
<th>Year</th>
<th>Rainfall(mm)</th>
<th>StreamFlow(mm)</th>
<th>AET(mm)</th>
<th>Temp</th>
<th>Rs</th>
<th>RH</th>
<th>Rn</th>
<th>Ra</th>
<th>Tmax</th>
<th>Tmin</th>
<th>Thorn</th>
<th>Hamon</th>
<th>Turc</th>
<th>PT</th>
<th>Makk</th>
<th>HS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1201</td>
<td>1961</td>
<td>1450.1</td>
<td>480.2421783</td>
<td>969.8578217</td>
<td>11.98</td>
<td>16.07</td>
<td>0.73</td>
<td>8.34</td>
<td>30.45</td>
<td>19.53</td>
<td>4.43</td>
<td>473</td>
<td>872</td>
<td>911</td>
<td>953</td>
<td>880</td>
<td>1251</td>
</tr>
<tr>
<td>1201</td>
<td>1962</td>
<td>1485.5</td>
<td>634.1659534</td>
<td>851.3340466</td>
<td>12.17</td>
<td>15.67</td>
<td>0.73</td>
<td>8.22</td>
<td>30.45</td>
<td>19.78</td>
<td>4.56</td>
<td>699</td>
<td>904</td>
<td>902</td>
<td>961</td>
<td>870</td>
<td>1292</td>
</tr>
<tr>
<td>1201</td>
<td>1963</td>
<td>1244.6</td>
<td>320.1614522</td>
<td>924.4385478</td>
<td>10.75</td>
<td>15.94</td>
<td>0.7</td>
<td>8.28</td>
<td>30.45</td>
<td>19.09</td>
<td>2.4</td>
<td>634</td>
<td>834</td>
<td>844</td>
<td>933</td>
<td>855</td>
<td>1284</td>
</tr>
<tr>
<td>1201</td>
<td>1964</td>
<td>1554.9</td>
<td>582.3969599</td>
<td>972.5030401</td>
<td>11.14</td>
<td>15.41</td>
<td>0.73</td>
<td>7.98</td>
<td>30.43</td>
<td>19.28</td>
<td>3</td>
<td>631</td>
<td>840</td>
<td>829</td>
<td>896</td>
<td>829</td>
<td>1276</td>
</tr>
<tr>
<td>1201</td>
<td>1965</td>
<td>1195.9</td>
<td>453.5620573</td>
<td>742.3379427</td>
<td>11.24</td>
<td>15.93</td>
<td>0.71</td>
<td>8.25</td>
<td>30.45</td>
<td>19.72</td>
<td>2.76</td>
<td>636</td>
<td>846</td>
<td>863</td>
<td>937</td>
<td>861</td>
<td>1307</td>
</tr>
<tr>
<td>1201</td>
<td>1966</td>
<td>1289.2</td>
<td>348.8938902</td>
<td>940.3061098</td>
<td>11.17</td>
<td>16.24</td>
<td>0.69</td>
<td>8.42</td>
<td>30.45</td>
<td>18.53</td>
<td>3.81</td>
<td>634</td>
<td>844</td>
<td>877</td>
<td>951</td>
<td>876</td>
<td>1211</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1202</td>
<td>1961</td>
<td>1217.1</td>
<td>317.2061157</td>
<td>899.8938843</td>
<td>14.01</td>
<td>15.57</td>
<td>0.71</td>
<td>8.16</td>
<td>30.02</td>
<td>21.64</td>
<td>6.38</td>
<td>745</td>
<td>1004</td>
<td>969</td>
<td>983</td>
<td>895</td>
<td>1344</td>
</tr>
<tr>
<td>1202</td>
<td>1962</td>
<td>1265.1</td>
<td>361.6456198</td>
<td>903.4534802</td>
<td>14.2</td>
<td>15.25</td>
<td>0.73</td>
<td>8.08</td>
<td>30.02</td>
<td>21.31</td>
<td>7.1</td>
<td>778</td>
<td>1037</td>
<td>963</td>
<td>993</td>
<td>887</td>
<td>1329</td>
</tr>
<tr>
<td>1202</td>
<td>1963</td>
<td>1048.2</td>
<td>262.0398347</td>
<td>786.1601653</td>
<td>13.55</td>
<td>15.78</td>
<td>0.68</td>
<td>8.28</td>
<td>30.02</td>
<td>20.91</td>
<td>6.18</td>
<td>745</td>
<td>1001</td>
<td>951</td>
<td>1001</td>
<td>904</td>
<td>1322</td>
</tr>
<tr>
<td>1202</td>
<td>1964</td>
<td>1224.4</td>
<td>442.5393719</td>
<td>781.8606281</td>
<td>14</td>
<td>15.1</td>
<td>0.72</td>
<td>7.89</td>
<td>30</td>
<td>20.92</td>
<td>7.08</td>
<td>752</td>
<td>1013</td>
<td>940</td>
<td>954</td>
<td>868</td>
<td>1276</td>
</tr>
<tr>
<td>1202</td>
<td>1965</td>
<td>1060.7</td>
<td>292.6877686</td>
<td>768.0122314</td>
<td>14.25</td>
<td>15.57</td>
<td>0.7</td>
<td>8.13</td>
<td>30.02</td>
<td>21.31</td>
<td>7.18</td>
<td>765</td>
<td>1024</td>
<td>979</td>
<td>994</td>
<td>900</td>
<td>1292</td>
</tr>
<tr>
<td>1202</td>
<td>1966</td>
<td>1182.2</td>
<td>245.1834711</td>
<td>937.0165289</td>
<td>13.26</td>
<td>15.72</td>
<td>0.69</td>
<td>8.22</td>
<td>30.02</td>
<td>20.39</td>
<td>6.13</td>
<td>709</td>
<td>973</td>
<td>944</td>
<td>982</td>
<td>892</td>
<td>1296</td>
</tr>
<tr>
<td>1202</td>
<td>1967</td>
<td>1089.3</td>
<td>285.0257851</td>
<td>804.2742149</td>
<td>13.58</td>
<td>15.53</td>
<td>0.69</td>
<td>8.04</td>
<td>30.02</td>
<td>20.93</td>
<td>6.22</td>
<td>708</td>
<td>970</td>
<td>953</td>
<td>961</td>
<td>881</td>
<td>1300</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Output

SiteID Year Rainfall(mm) StreamFlow(mm) AET(mm) Temp Rs RH Rn Ra Tmax Tmin Thorn Hamon Turc PT Makk HS

1201,29,1320,87,398,86,922.81,11.73,15.99,.72,8.34,30.45,19.71,3.76,671,1,875.55,890.69,955.86,874,1296.1
1202,30,1150,76,349,87,800.9,14.17,15.44,.7,8.13,30.02,21.38,6.96,770.03,1028.47,965.33,399.0,892.47,1325.1
1203,30,1122.8,324.39,798.42,15.43,15.44,.7,8.19,30.02,21.62,9.23,821.43,1102.7,1010.17,1022.87,914.07,1276.1
1204,27,1213.38,317.09,896.29,14.64,15.74,.7,8.33,30.15,21.32,7.97,791.0592,22.997.59,1028.7,922.11,1279
1205,30,1122.01,330.64,791.37,14.27,15.81,.7,8.35,30.19,20.81,7.74,774.87,1034.87,989.87,1021.7,918.27,1277
...
Appendix F. SAS program for the long-term annual AET model building

```sas
options pagesize = 80;

data thesis;
/* read input data */
infile 'h:\AnalysisData\Data\JianbiaoLuThesisData.csv' dlm = ',' dsd missover firstobs=2;
input Name $ SiteNumber Year Rainfall Streamflow AET Temp Rs RH Rn Ra Tmax Tmin Thorn Hamon Turc PT Makk HS Latitude Longitude Elevation Deciduous Conifer Water CropGrass Others Forest;
/* check out the data with a look */
proc print data = thesis;
run;
/* look at the linear correlation of all variables with AET */
proc corr nosimple data = thesis; var AET; with rainfall -- Forest;
run;
/******* With RSQUARE *******
proc reg data = thesis;
model AET = Rainfall Temp -- Forest / selection = RSQUARE BEST = 5 CP ADJRSQ AIC SBC;
title 'AET model building with R square';
run;
/******* With Stepwise *******
proc reg data = thesis;
*model AET = Rainfall Temp -- Forest / selection = STEPWISE collinoint slentry = 0.25 slstay = 0.05;
model AET = Rainfall Temp -- Forest / selection = STEPWISE collinoint slentry = 0.25 slstay = 0.25;
title 'AET model building with stepwise';
run;
/* Regression Diagnostics */
* Get the model after run the stepwise selection RSQUARE to select model;
proc reg data = thesis; model AET = Rainfall Latitude Elevation Conifer Water / influence partial r vif collinoint;
output out = two rstudent = rstudent cookd = cookd press = press student = student predicted = p residual = r;
id SiteNumber;
run;
proc sort data = two; by cookd;
proc print data = two; var AET p r rstudent student press cookd;id SiteNumber;
run;
proc gplot data= two; plot r*p;
run;
proc univariate data = two normal plot; var r press cookd; id SiteNumber;
run;
proc iml;
/* Calculate critical values */
T = tinv(1-0.05/36, 30); /* for Rstudent and Studendize */
F = finv(0.25, 6, 30); /* for Cook's D cutoff value F(0.25, p`, n-p`)*/
print T F;
quit;
```
Appendix G. SAS program for comparisons of six PET methods

```sas
options pagesize = 80;
data thesis;
infile 'h:\AnalysisData\SAS\PETsComparison\36LongTermAnnualPETs.csv' dlm = ',' dsd missover firstobs=2;
   input name $ SiteNumber AET Thorn Hamon Turc PT Makk HS;
   a1 = Thorn - Hamon;
   a2 = Hamon - Turc;
   a3 = Turc - PT;
   a4 = PT - Makk;
   a5 = Makk - HS;
/* check out the data with a look */
proc print data = thesis;
/* calculate correlation coefficients between 6 PET */
proc corr cov data = thesis;
   var Thorn -- HS;
run;
/* calculate correlation coefficient between AET and 6 PET */
proc corr nosimple data = thesis; var AET; with Thorn -- HS;
run;
/* multivariate statistical test */
proc reg data = thesis;
   model a1 a2 a3 a4 a5 = ;
   mtest intercept /print;
   title 'Proc Reg Contrast test - T-sq on 5 contrast';
run;
proc glm data = thesis;
   model Thorn Hamon Turc PT Makk HS = /nouni;
   repeated PET contrast(1) /summary;
   repeated PET contrast(2) /summary;
   repeated PET contrast(3) /summary;
   repeated PET contrast(4) /summary;
   repeated PET contrast(5) /summary;
   repeated PET contrast(6) /summary;
   title '6 PET comparisons';
run;
```

105