Session 2793

Object Oriented Analysis of Weather Data in Virginia

Steven Thomas, Vinod Lohani, Bevlee Watford

Senior, Computer Engineering/Assistant Professor, Engineering Fundamentals/Director, Minority Engineering Program, Virginia Tech

Abstract

An object-oriented approach is presented to analyze drought patterns in Virginia using probabilistic analysis of long - term weather data (i.e. Palmer Drought Severity Index (PDSI)). The study is a part of an undergraduate research project, funded by the office of the Minority Engineering program under a VT-GE program, at Virginia Tech. The purpose of the program is to encourage and support the efforts of under-represented engineering students to seek graduate degrees. The present paper describes research work of one of the current VT-GE scholars.

1.0 Introduction

Object oriented approach has not found many applications in analyzing hydrology and water resources problems. It may be attributed to non-exposure of professionals working in these areas to object oriented programming languages like C++. This fact is also evident from analysis of papers that were presented in recent national level conferences, organized by the ASCE, in hydrology and water resources area. The 1999 conference in Arizona had only 4 papers out of 264 presentations that used the object oriented approach. In the 2000 conference in Minnesota, this ratio was 6 out of 509 papers and in 2001 conference in Orlando, only 5 out of 443 presentations made use of this powerful programming technique[1].

The paper presents results of an ongoing undergraduate research project that has objective to analyze long- term weather records using an object oriented approach. The first author is pursuing an undergraduate degree in computer engineering at Virginia Tech. The computer engineering curriculum includes several courses that require students to learn and use object oriented programming languages like C++. The problem solving potential of the programming languages can be effectively realized by analyzing real world data, available from the Internet, for meaningful conclusions. Availability of long term weather data over the Internet definitely offers potential in this direction. Such efforts can also meet the need to have inquiry based education for undergraduates as has been proposed by a national level commission in its recent report[2].

The main objective of this study is to carry out probabilistic analysis of long - term weather data (i.e. Palmer Drought Severity Index (PDSI)) to analyze drought patterns in Virginia using an object-oriented approach. PDSI is a comprehensive numerical weather index[3] and is computed by the National Weather Service (NWS) on monthly basis for all 344 climatic divisions in the US. The state of Virginia is classified into 6 climatic divisions. PDSI is used by federal and state agencies in managing drought related management actions. Long term monthly records (~100 years) of PDSI are available from the web site of the National Climatic Data Center (NCDC), Asheville, NC. A non-homogeneous Markov chain approach[4] is used to characterize probabilistic behavior of drought using the object oriented approach. For Markov chain analysis the PDSI data is classified into various weather states based on criteria available from literature[5] (see table 1).

Table 1: PDSI Based Weather States

Weather State	Description	PDSI Range
1	Extreme Dry	(-∞ , -4.00]
2	Dry	[-4.00 , -2.99]
3	Mild Dry	[-2.99 , -1.50]
4	Normal	[-1.50 , 1.50]
5	Mild Wet	[1.50 , 2.99]
6	Wet	[2.99 , 4.00]
7	Extreme Wet	[4.00 , ∞)

2.0 C++ class for analysis

A C++ class, called Weather, is defined using object oriented technology with appropriate data members and member functions for analysis of long term PDSI data. Table 2 shows the details of the C++ class including various data members and member functions. The member functions operational at the time of this writing characterize long and short term probabilities of occurrence of drought events using non-homogeneous Markov chain approach[4]. Specifically, the member function Shortterm() is the method that calculates the probability of monthly transition from a given weather state to another in one of the climatic divisions in Virginia (see Shortterm() in Table 2). The member function Longterm() is designed to compute long term probability of occurrence of a weather state in a climatic division. In the C++ code, a user is prompted to enter a desired weather state (i.e. 1 through 7) representing different weather type (i.e. Extreme Dry, Dry, Mild Dry, Normal, Mild Wet, Wet, Extreme, Wet) and the program will return the probability of occurrence of chosen weather state in the selected climatic division (see Longterm() in Table 2). In addition, several other member functions have been defined that are used to assist in the analysis of the data (Table 2). For example, the Convert() function is used to convert the raw PDSI data, downloaded from the Internet, to weather states based on the classification given in Table 1. Also, another member function named Insert() adds the newly converted data into the storage array. To handle the file input and output, another C++ class called FileManip was created with a member function Prompt_user() that allows the program to be interactive with the user (see Table 2).

3.0 Data Source

The data for this analysis was retrieved from the following website maintained by National Climatic Data Center in Asheville, NC:

ftp://ftp.ncdc.noaa.gov/pub/data/cirs/holdfiles/

This site contains PDSI values for all the states in the US with the exception of Alaska and Hawaii. The format of available data is as follows: first two digits represent a numerical code for the states, the next two give a code for the climatic division within a state, the fifth digit is a data type number and the next four digits represent the year of record. Following this, there are 12 monthly values for that state/division/year separated by spaces. Virginia has been assigned a numerical code of 44. 100 years of monthly PDSI data, covering the period from 1900-2000, was downloaded for all six climatic divisions of Virginia, namely, Tidewater, Eastern Piedmont, Western Piedmont, Northern, Central Mountains, and Southwestern Mountains. An example of raw data, including the format, for the Tidewater division during the years 1977-1979 is as below, where the PDSI value for January 1977 is –0.03:

440151977 -0.03 -0.46 -0.65 -1.10 -0.99 -1.36 -1.96 -2.47 -2.93 0.83 1.35 2.07
440151978 3.27 2.15 3.09 3.19 3.75 3.70 3.29 2.93 1.88 1.02 1.13 1.14
440151979 2.27 2.97 2.73 2.97 3.74 3.77 3.71 3.40 5.36 4.95 5.51 -0.54

The raw data is converted to weather states (1 through 7) using the criteria given in table 1 and analysis of data is performed using the 'Weather' class. A brief description of program operation is given in the following section.

4.0 Program Operation

A stepwise procedure of running this interactive program is given in table 3. First, a user is asked to choose the climatic division in Virginia for which analysis is desired. Following this choice, the user is asked a series of questions to determine what type of analysis is desired (i.e. short term, long term). Typically (based on the type of analysis chosen) the user is asked to choose the weather state/s (1-7) for analysis. For short term analysis the choice of months is also asked. See table 3 for further details. Once all choices are made, the data file for the selected climatic division is opened and results of chosen analysis are presented in tabular form. Following section gives results of analysis of long and short term probabilities for all six climatic divisions in Virginia.

5.0 Results

Table 4 gives long term probabilities of various weather states in six climatic divisions in Virginia. It is seen that the 'normal' weather state (i.e. state 4) has the highest long term probability across all climatic divisions. This is in agreement with the weather patterns experienced in past since Virginia is traditionally not considered as one of drought prone states. Interestingly, all climatic divisions showed long term probability of about 25% for dry weather states (i.e. 1, 2, and 3) which is also in agreement with studies that have documented drought years in Virginia[6]. A comparison of long term drought probabilities for various climatic divisions indicates that the Southwest Mountains, Tidewater, and Eastern Piedmont regions are more prone to drought as compared to rest three climatic divisions. This result is in agreement with a study[6] that analyzed the frequency of drought years in Virginia over the period of 1957-87 with the exception of Southwestern Mountain division.

Table 5 gives results of short- term analysis showing probability of monthly transition from one weather state to another for all six climatic divisions. The results are presented for monthly transition from July to August. In this table, the values in each row must add to 100 since a transition must occur from one weather state to another. Also if the desired transition did not occur, an equal probability of transition to each state is assumed. It can be seen that the weather, as characterized by PDSI, tends to stay in the same weather state over a period of month. The results for Tidewater division indicate that compared to wet states (i.e. 5, 6, and 7) the dry states (i.e. 1. 2, and 3) have higher probability of staying in current state implying that, once occurred, the dry states would last longer than wet conditions. Also, the extreme dry conditions (state 1) can be expected to continue longer than the extreme wet condition (state 7) in Tidewater region. Such results can be generated for remaining 11 pairs on months and then all 12 transition probability tables can be used to predict most likely weather states several months ahead of time. Out of all divisions it is the Tidewater region that has least probability of returning to a normal state (i.e. state 4) in August given that the weather state was normal in July.

6.0 Conclusion and Future work

An object oriented approach has been used to characterize drought probabilities in Virginia using long term PDSI data. The study is part of an ongoing undergraduate research project. The results presented agree with other drought studies in the region. Another member function for analyzing duration of drought events is under development at the time of this writing. It is also proposed to use PDSI data from drought prone states like Arizona and California for further testing the approach. Also, it is planned to repeat the analysis with weekly PDSI data, if available, to get a better characterization of drought probabilities.

Table 2: C++ Class for Analysis of PDSI Data

Class Name	Weather
Purpose	Used to perform analysis on weather data

Operations	int Convert(double);
	bool insert(double, int, int);
	double longterm(int pattern);
	double shortterm(int patternA, int patternB, int monthA, int monthB);
	double duration(int patternA);

Function Name	Convert()
Purpose	Changes the raw data (as doubles) and converts them to integers.
Data Received	Raw data from input files
Data Returned	An integer equivalent
Remarks	Conversion table: (value < -4.00) extreme dry (Class 1) (value >= -4.00 && value <= -2.99) dry (Class 2) (value >= -2.99 && value <= -1.50) mild dry (Class 3) (value >= -1.50 && value <= 1.50) normal (Class 4) (value >= 1.50 && value <= 2.99) Mild Wet (Class 5) (value >= 2.99 && value <= 4.00) Wet (Class 6) (value > 4.00) Extreme Wet (Class 7)

Function Name	Insert()
Purpose	Places converted data into a 2-D array
Data Received	An integer (1-7) which represents which weather type (Extreme Dry, Dry, Mild Dry, Normal, Mild Wet, Wet, Extreme, Wet).
Data Returned	A 2-D array of converted data.
Remarks	

Function Name	Longterm()
Purpose	Calculates the probability of occurrence of a given weather type over a period of 100 years.
Data Received	An integer (1-7) which represents the weather type (Extreme Dry, Dry, Mild Dry, Normal, Mild Wet, Wet, Extreme, Wet).
Data Returned	Percent probability of occurrence of a given weather type.
Remarks	

Function Name	Shortterm()
Purpose	Calculates the probability of monthly transition from a given weather type A to another weather type B.
Data Received	Two integers (1-7) representing the weather type (Extreme Dry, Dry, Mild Dry, Normal, Mild Wet, Wet, Extreme, Wet). Two integers representing the months involved in transition.
Data Returned	Percent probability of monthly transition.
Remarks	

Function Name	Duration()
Purpose	Calculates the expected duration of a given weather type over the entire data.
Data Received	An integer (1-7) which represents the weather type (Extreme Dry, Dry, Mild Dry, Normal, Mild Wet, Wet, Extreme, Wet).
Data Returned	An expected duration of stay for the chosen weather state.
Remarks	

Class Name	FileManip
Purpose	Used to perform I/O operations
Constructors	FileManip(); opens the output file
Operations	void prompt_user()

Function Name	Prompt_user()
Purpose	Sends prompts to screen.
Data Received	User input
Data Returned	A value based on the user's input
Remarks	A series of questions are asked of the user to determine the type of analysis desired and accordingly proper functions in the Weather class are executed.

Table 3: Program Execution - Stepwise Procedure

FileManip Class Objects
Weather Class Objects

Prompt 1	"Please select the region (1-6) within the state of Virginia:" Region_name[] : 1 Tidewater 2 Eastern Piedmont 3 Western Piedmont 4 Northern 5 Central Mtns 6 Southwestern Mtns
Input 1	*User input = Choice*
Operation	Open data file and insert() is called
Prompt 2	"What type of analysis would you like to do?" 1 Long term 2 Short term 3 Duration *(not operational yet)*
Input 2	*User input = analysis*
Input 2, Case1	**Long term Analysis**
Prompt 2.1	"Would you like to determine the Long term probability for: " 1 weather pattern 2 All the weather patterns
Input 2.1	*User input = times*
Input 2.1, Case 1	**Long term Analysis for 1 pattern** "What will that weather pattern be?" Weather_pattern[]: 1 Extreme Dry 2 Dry 3 Mild Dry 4 Normal 5 Mild Wet 6 Wet 7 Extreme Wet
Input 2.11	*User input = pattern*
Operation	call to Longterm() and output displayed to a file
Input 2.1, Case 2	**Long term Analysis for all weather patterns**
Operation	call to Longterm() and output displayed to a file
Input 2, Case 2	**Short term Analysis**
Prompt 2.2	"Would you like to determine the Short term probability for: " 1 One weather pattern to another One weather pattern 2 One weather pattern to All the other weather patterns 3 All weather patterns to All the other weather patterns 4 All weather patterns to All the other weather patterns for all months
Input 2.2	*User input = times*
Input 2.2 ,Case 1	**Short term Analysis for One weather pattern to another weather pattern** "Choose first weather pattern"

Input 2.21	*User input = patternA*
Prompt 2.22	"Choose second weather pattern"
Input 2.22	*User input = patternB*
Prompt 2.23	"Choose initial month"
Input 2.23	*User input = monthA*
Operation	call to Shortterm() and output displayed to a file
Input 2.2 ,Case 2	**Short term Analysis for One weather pattern to All the other weather patterns** "Choose first weather pattern"
Input 2.221	*User input = patternA*
Prompt 2.222	"Choose initial month"
Input 2.222	*User input = monthA*
Operation	call to Shortterm() and output displayed to a file
Input 2.2 ,Case 3	**Short term Analysis for All weather patterns to All the other weather patterns** "Choose first weather pattern"
Prompt 2.2222	"Choose initial month"
Input 2.2222	*User input = monthA*
Operation	call to Shortterm() and output displayed to file
Input 2.2 ,Case 4	**Short term Analysis for All weather patterns to All the other weather patterns for all months** "Choose first weather pattern"
Operation	call to Shortterm() and output displayed to file
Input 2, Case 3	**Duration** *(work in progress)*
Prompt 2.3	"You have chosen to determine the expected Duration of a weather state " "Choose weather pattern"
Input 2.3	*User input = patternA*
Operation	call to Duration() and output displayed to a file

Table 4: Long Term Probability (%) of Various Weather States in Virginia

Weather State	Tidewater	Eastern Piedmont	Western Piedmont	Northern	Central Mountains	Southwestern Mountains
1	1.78	1.07	0.911	1.94	1.41	0.990
2	3.69	5.85	4.47	3.40	4.75	4.65
3	19.2	18.2	17.6	16.4	17.5	19.5
4	49.1	53.8	56.6	58.0	55.6	53.5
5	19.5	15.9	15.5	14.6	16.7	17.4
6	5.82	3.83	4.09	3.68	2.64	3.40
7	0.910	1.35	0.829	1.98	1.40	0.560

Table 5: Short Term Probability of Monthly Transition: *July Þ August*

1) Climatic Division: Tidewater

Weather States	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	0	60	20	20	0	0	0
3	0	4.30	74.0	21.7	0	0	0
4	0	0	21.4	61.9	16. 7	0	0
5	0	0	0	26.3	57.9	15.8	0
6	0	0	0	0	44.4	33.4	22.2
7	0	0	0	0	50	0	50

2) Climatic Division: Eastern Piedmont

Weather States	1	2	3	4	5	6	7
1	14.3	14.3	14.3	14.3	14.3	14.3	14.3
2	0	66.7	33.3	0	0	0	0
3	0	21.0	57.9	21.1	0	0	0
4	0	0	18.9	67.9	9.43	3.77	0
5	0	0	0	27.8	61.1	11.1	0
6	0	0	0	50	25	25	0
7	0	0	0	0	0	100	0

3) Climatic Division: Western Piedmont

Weather States	1	2	3	4	5	6	7
1	0	100	0	0	0	0	0
2	0	80	20	0	0	0	0
3	0	21.4	57.1	21.5	0	0	0
4	0	0	12.3	70.2	17.5	0	0
5	0	0	0	42.9	38.1	14.3	4.70
6	0	0	0	0	50	50	0
7	0	0	0	0	0	100	0

4) Climatic Division: Northern

Weather States	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	25	25	50	0	0	0	0
3	0	13.3	66. 7	13. 3	0	6.70	0
4	0	0	12.1	72.4	13.8	1.70	0
5	0	0	0	21.0	63.2	15.8	0
6	0	0	0	0	33.3	66.7	0
7	14.3	14.3	14.3	14.3	14.3	14.3	14.3

5) Climatic Division: Central Mountains

Weather States	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	20	80	0	0	0	0	0
3	0	16. 7	61.1	22.2	0	0	0
4	0	0	9.10	78.2	12.7	0	0
5	0	0	0	36.84	47.4	10.5	5.30
6	0	0	0	0	100	0	0
7	0	0	0	0	0	0	100

6) Climatic Division: Southwestern Mountains

Weather States	1	2	3	4	5	6	7
1	100	0	0	0	0	0	0
2	0	66.7	0	33.3	0	0	0
3	0	9.10	68.2	22.7	0	0	0
4	0	0	7.60	77.0	15.4	0	0
5	0	0	0	33.3	52.4	14.3	0
6	0	0	0	0	0	100	0
7	14.3	14.3	14.3	14.3	14.3	14.3	14.3

Bibliography

1. ASCE Conference Proceedings, Arizona, 1999, Minnesota, 2000, and Orlando, 2001.

2. Boyer Commission Report, 1998, Reinventing Undergraduate Education: A Blueprint for America's Research Universities, Available from: http://notes.cc.sunysb.edu/Pres/boyer.nsf/.

3. Palmer, W. C., 1965, Meteorological drought, U.S. Weather Bureau, Research Paper no.45, Washington, D.C.

4. Lohani, V. K., and Loganathan, G. V., 1997, An Early Waning System for Drought Management Using the Palmer Drought Index, *Journal of the American Water Resources Association, Vol. 33, No. 6.*

5. Karl, T. R., 1986, The sensitivity of the Palmer drought index and Palmer's Z index to their calibration coefficients including potential E. T., *Journal of Climate and Applied Meteorology, Vol. 25*, pp. 77-86.

6. Drought Indicator Parameters, State Water Control Board, Information Bulletin # 582, April 1990.

Biographical Information

Steven Thomas is a senior pursuing a Bachelor's of Science in Computer Engineering. The Educational and Political Action Chairperson for the NAACP, President Virginia Tech Committee for African American Student Issues and a member of IEEE and NSBE.

Vinod Lohani is an assistant professor in the Engineering Fundamentals Department at Virginia Tech. He obtained a Ph.D. in civil engineering from Virginia Tech in 1995 and has research interests in hydrology and water resources and engineering education.

Bevlee Watford is an associate professor in the Industrial and Systems Engineering program in the College of Engineering at Virginia Tech. Also, she is the director of the Minority Engineering program and also holds the position of Associate Dean for Academic Affairs.