INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600



Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/595

UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

THE RELATIONSHIP BETWEEN WEATHER VARIABLES AND ELECTRICITY DEMAND TO IMPROVE SHORT-TERM LOAD FORECASTING

A Dissertation

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

degree of

Doctor of Philosophy

By

Ahsha N. Tribble
Norman, Oklahoma
2003

UMI Number: 3082952



UMI Microform 3082952

Copyright 2003 by ProQuest Information and Learning Company.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

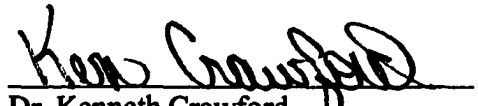
ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© Copyright by AHSHA N. TRIBBLE 2003
All Rights Reserved.

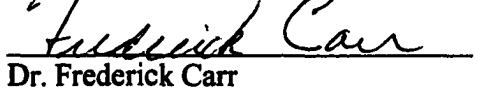
THE RELATIONSHIP BETWEEN WEATHER VARIABLES AND ELECTRICITY DEMAND TO IMPROVE SHORT-TERM LOAD FORECASTING

A Dissertation APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY



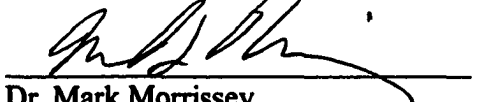
Dr. Kenneth Crawford



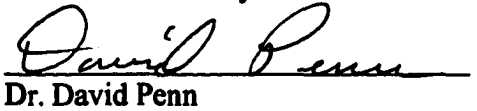
Dr. Frederick Carr



Dr. Kevin Kloesel



Dr. Mark Morrissey



Dr. David Penn

Acknowledgements

I must first thank God for blessing me with so many opportunities, for surrounding me with wonderful people, and for giving me the strength and courage to overcome the many obstacles placed in my path.

It is with the greatest respect and admiration that I thank Dr. Jeff Kimpel. My ride at OU has been a rollercoaster, but Dr. Kimpel has been my constant. No matter how busy he was, he always made time to advise, mentor, understand, and listen. He never let me give up, even when my bags were packed. He is a great leader, and truly cares about the well-being (both professionally and personally) about the people around him. Without the support of Dr. Kimpel, I do not know where I would be. Thank you for the opportunity at NOAA and for going the extra mile to help me succeed.

I know that Dr. Ken Crawford did not know what challenge was before him when he agreed to be my advisor. I cannot thank you enough for your continued support and encouragement throughout this process. I was extremely fortunate to be mentored by you, and I sincerely appreciate everything that you have taught me. I have grown a lot under your wing, and one day soon I will make you proud to have had me as a student.

I followed Dr. Kevin Kloesel to OU – enough said. Dr. Kloesel taught me my first meteorology class and served on my Masters at Florida State University. He, too, has been very supportive and very much an instrumental part of my successes. Thanks for being there for me and agreeing to serve on one last committee of mine. Dr. Mark Morrissey, Dr. Fred Carr, and Dr. David Penn, thank you for your time and effort as members of my committee. To Dr. Pete Lamb, thank you for your valuable comments and continued support.

Many others have helped me along the way. Renee, you have been awesome. Thank you for your continuous support and encouragement. Derek, I cannot thank you enough for your arduous effort in getting those Eta forecasts up and running. I must acknowledge Cerry, Amy, Kelly, Mark, Jared, Stdrovia, Billy, and Janet for their help and support as well. Celia, you are awesome; thanks for being there when I needed you. Kelly (at NSSL), thank you for the laughs and for keeping me in line at every turn.

Thanks to Western Farmers Electric Cooperative for their cooperation. Dr. Liao, Darryl, Mandi, Janelle, Bert, Mike and Josh were instrumental in supplying the electric load data and helping me understand the concerns of the electricity industry.

I must also thank all of my family and friends who have been by my side for so many years. Trevia - my best friend and lifeline for 15 years – thanks for hanging in there with me. Lisa, thanks for all of the laughs and cries and letting me vent! Thanks to my little sister, Aiyisha, for keeping me on my toes. I must also show gratitude to Dr. John Cortinas for his support and friendship, which has been so valuable throughout my stay in Oklahoma. To my husband, Tony, thank you for supporting me and putting up with me, especially near the end of this process. Hopefully, we can spend some time together now!

And last but certainly not least, words cannot express the love, respect, and admiration I have for my parents. Mommy, thanks for your unconditional love and support and being the best mother in the world. I hope we can still talk 4 times a day when I have a job! Daddy, you are a true pillar of strength and a bundle of inspiration. This one is for you, dad! The three of us have worked hard to get here, and I hope I have made you both proud. I love you both, and thank you so much for a lifetime of support.

Table of Contents

Acknowledgements vi

List of Tables ix

List of Figures xi

Abstract xvii

Chapter 1: Introduction 1

Chapter 2: Literature Survey 7
- 2.1 Importance of Weather 7
- 2.2 The Meteorologist View 7
  - 2.2.1 History of Numerical Weather Prediction (NWP) 9
  - 2.2.2 Types of NWP Models 14
    - 2.2.2.1 Nested Grid Model (NGM) and Model Output Statistics (MOS) 16
    - 2.2.2.2 Eta Model 18
    - 2.2.2.3 Rapid Update Cycle (RUC) 20
    - 2.2.2.4 Multi-Scale Models 21
- 2.3 Introduction to Electrical Load Forecasting 22
- 2.4 Short-term Load Forecasting Methods 24
- 2.5 Standard Load Curve Models 25
  - 2.5.1 Time-of-Day Models 25
  - 2.5.2 Dynamic Models 27
    - 2.5.2.1 Autoregressive Moving Averages (ARMAs) 27
    - 2.5.2.2 State Space Models and Kalman Filter 29
  - 2.5.3 Knowledge-Based Expert System (KBES) Approach 30
  - 2.5.4 Artificial Neural Networks (NNs) 30
    - 2.5.4.1 Neural Network Architecture 33
    - 2.5.4.2 Neural Networks and Meteorological Forecasting 37
  - 2.5.5 STLFs Using Weather-Load Models 38
- 2.6 Merging Load Forecasting and NWP 39

Chapter 3: Data and Sites 41
- 3.1 Weather Variables 41
- 3.2 Electric Load Data 43
- 3.3 Geography of Oklahoma 45
- 3.4 Description of Data and Sites 50
  - 3.4.1 Norman Site 52
  - 3.4.2 Woodward Site 53
  - 3.4.3 Altus/Altus AFB Site 53
  - 3.4.4 Broken Bow/Dominance Site 54

Chapter 4: Analysis of Data 56
4.1 Climatology of Oklahoma 56
4.2 Oklahoma Weather During The Study Period 61
4.2.1 Oklahoma Weather – 1998 62
4.2.2 Oklahoma Weather – 1999 63
4.2.3 Oklahoma Weather – 2000 64
4.3 Electric Load Demand 65
4.3.1 Average Load Profiles – Norman 67
4.3.2 Weekdays, Weekends, and Holidays – Norman 72
4.3.3 Average Load Profiles – Woodward 75
4.3.4 Weekdays, Weekends, and Holidays – Woodward 80
4.3.5 Average Load Profiles – Altus AFB 83
4.3.6 Weekdays, Weekends, and Holidays – Altus AFB 88
4.3.7 Average Load Profiles – Dominance 91
4.3.8 Weekdays, Weekends, and Holidays – Dominance 96
4.4 Quality Assurance of the Data 98
4.4.1 Quality Assurance of Mesonet Data 99
4.4.2 Quality Assurance of WFEC Electric Load Data 103

Chapter 5: Interrelationship Betweens Weather Variables and Electric Load Data 106
5.1 Significant Weather Variables 106
5.2 Norman Site 108
5.2.1 Temperature-Load Relationship (Norman) 108
5.2.1.1 Scatterplots 108
5.2.1.2 Line Graphs 114
5.2.1.3 Apparent Temperature-Load Relationship 122
5.2.1.4 Temporal Correlations 122
5.2.2 Relative Humidity-Load Relationship (Norman) 124
5.2.3 Solar Radiation-Load Relationship (Norman) 127
5.2.4 Wind Speed-Load Relationship (Norman) 130
5.3 Woodward Site 131
5.3.1 Temperature-Load Relationship (Woodward) 131
5.3.1.1 Scatterplots 131
5.3.1.2 Line Graphs 136
5.3.1.3 Apparent Temperature-Load Relationship 143
5.3.1.4 Temporal Correlations 143
5.3.2 Relative Humidity-Load Relationship (Woodward) 144
5.3.3 Solar Radiation-Load Relationship (Woodward) 146
5.4 Altus AFB Site 150
5.4.1 Temperature-Load Relationship (Altus AFB) 150
5.4.1.1 Scatterplots 150
5.4.1.2 Line Graphs 155
5.4.1.3 Apparent Temperature-Load Relationship 163
5.4.1.4 Temporal Correlations 163
5.4.2 Relative Humidity-Load Relationship (Altus AFB) 164
5.4.3 Solar Radiation-Load Relationship (Altus AFB) 165

5.5 Dominance/Broken Bow Site 168
5.5.1 Temperature-Load Relationship (Dominance) 168
5.5.1.1 Scatterplots 168
5.5.1.2 Line Graphs 172
5.5.2 Temporal Correlations (Dominance) 172
5.6 Summary 174

Chapter 6: Applications of Electric Load Modeling 175
6.1 MetrixND – A Load Forecasting Tool 176
6.1.1 The Neural Network Model 176
6.1.2 The Multiple Regression Model 177
6.1.3 Modeling Specifications 178
6.2 Modeling Study I 179
6.3 Modeling Study II 187
6.4 Economic Benefit of Improved Load Forecasts 190

Chapter 7: Summary and Concluding Remarks 193

References 200

Appendix A 207

Appendix B 211

Appendix C 212

Appendix D 214

Appendix E 216

Appendix F 220

List of Tables

Table 2.1 Significant milestones in the development of NWP. 9

Table 3.1 The four Mesonet sites and substations used in this study, and their separation distances. 50

Table 3.2 The site name, type of customer, average annual load, and the peak annual load for each substation (based on 2000 load data). 51

Table 3.3 The number of customers served each year (for the 3 year study period) and the miles of distribution line needed to reach these customers. 52

Table 4.1 Climatological normal temperatures (F) and annually-averaged temperatures observed at four sites during the study period. 62

Table 4.2 The number and percentage of observations interpolated, substituted or missing from the Mesonet data sets for all four study sites. (Each site has 105,216 observations, and the total data set has 420,864 observation. 100

Table 4.3 The percentage of valid observations that were within ±5 and ± 10 (°C, %, and ms^{-1}) of each other when comparing measurements from a pair of neighboring Mesonet sites. 102

Table 5.1 Hour-to-hour correlation coefficients averaged seasonally over the study period between load and the following weather variables: temperature, apparent temperature, relative humidity, and solar radiation. When a significantly higher correlation resulted from a time lag (i.e., the number of hours the load lagged the weather variable) versus a zero lag coefficient, the highest correlation coefficient was shown. (The number of hours that the load lagged the weather variable is in parenthesis.) 123

Table 6.1 The user-defined, exogenous (non-meteorological) variables used in the hourly neural network and regression models. 180

Table 6.2 The specification list for the (a) neural network and (b) regression model in Modeling Study I, including weather variables. 180

Table 6.3 A comparison of the mean absolute percentage errors (MAPEs) of the load forecast using the NN and using the regression model using temperature (L_T) versus using temperature, relative humidity, and solar radiation ($L_{T,RH,Srad}$) at the four study sites. 182

Table E.1 "Gaps" in the Norman Mesonet data set filled by temporal interpolation from its own site or by substituting data from the Washington Mesonet site, whenever possible; otherwise, observations were labeled as missing. 215

Table E.2 As in Table E.1, except for the Woodward Mesonet site with substitutions from the Seiling Mesonet site. 216

Table E.3 As in Table E.1, except for the Altus Mesonet site with substitutions from the Tipton Mesonet site. 217

Table E.4 As in Table E.1, except for the Broken Bow Mesonet site with substitutions from the Idabel Mesonet site. 218

Table E.5 Electric load data missing/corrupt from the data set for the West Norman substation. 219

Table E.6 As in Table E.5, except for the Woodward substation. 219

Table E.7 As in Table E.5, except for the Altus AFB substation. 219

Table E.8 As in Table E.5, except for the Broken Bow substation. 219

List of Figures

Figure 1.1 An illustration of the three components involved to provide electricity to customers (illustration provided by WFEC). 2

Figure 2.1 Historic evolution of the operational forecast skill of the NCEP models over North America (500 hPa). The S1 score measures the relative error in the horizontal pressure gradient, averaged over the region of interest. The S1 = 70% and S1 = 20% were empirically determined to correspond respectively to a "useless" and a "perfect" forecast when the score was designed. Note that the 72-hr forecasts are currently as skillful as the 36-hr were 10-20 years ago (data courtesy of C. Vlcek, NCEP). It shows the results from global (AVN) and regional (LFM, NGM, and Eta) forecasts. The LFM model development was "frozen" in 1986 and the NGM was frozen 1991. (Figure and caption for Kalnay 2003) 13

Figure 2.2 An artificial neuron where the $\mathbf{p}_i$s are the four inputs, b is the bias term, and the $\mathbf{w}_i$s are the weights (weight matrix), all of which are combined and fed into a function f, to produce an output , $a = F(\mathbf{w}_i\mathbf{p}_i + b)$. 32

Figure 2.3 A 3-layer network where each layer has its own weight matrix **w**, bias vector **b**, net input vector **n**, and output vector **a**. The first and second layers are the hidden layers, while the third layer is the output layer. (Figure from Hagan et al. 1996) 34

Figure 3.1 The Oklahoma Mesonet sites. 42

Figure 3.2 The physiographic regions of Oklahoma as described by Wikle (1991). 45

Figure 3.3 As in Fig. 3.2 except the agricultural regions are shown. 47

Figure 3.4 The ratio of farm income to total income by county in Oklahoma as described by Späth (1998). 47

Figure 3.5 The population per county in Oklahoma as described by Späth (1998). 48

Figure 3.6 The percentage of each county in Oklahoma that is urban (versus rural) as described by Späth (1998). 48

Figure 3.7 The various grasses and trees found across Oklahoma are displayed as described by Späth (1998). 49

Figure 3.8 The forested areas found across Oklahoma are displayed as described by Wikle (1991). 49

Figure 3.9 Four data sites used in this study: 1 - Norman Mesonet site / West Norman substation; 2 - Woodward Mesonet site / Woodward substation; 3 - Altus Mesonet site / Altus AFB substation; 4 - Broken Bow Mesonet site / Dominance substation. 51

Figure 4.1 The nine climate divisions of Oklahoma. 58

Figure 4.2 The average temperature for February based upon data from 1971-2000. 59

Figure 4.3 As in Fig. 4.2, except for May. 59

Figure 4.4 As in Fig. 4.2, except for August. 60

Figure 4.5 As in Fig 4.2, except for November. 61

Figure 4.6 Typical load profiles for the customers of rural electric cooperatives. (Figure from Rastogi and Roulet 1994) 66

Figure 4.7 Annually-averaged load profiles from Norman during the study period. 68

Figure 4.8 Average hourly electric load profiles from Norman during the winters of the study period. 69

Figure 4.9 As in Fig. 4.8, except for spring. 70

Figure 4.10 As in Fig. 4.8, except for summer. 71

Figure 4.11 As in Fig. 4.8, except for fall. 72

Figure 4.12 Annually-averaged load profiles from Norman in 1998 that illustrate the differences in the weekday versus weekend load. 74

Figure 4.13 Daily load profiles for the four Mondays in February 1998 from Norman. The holiday Monday (President's Day) produced an electrical load with a different pattern than was observed on non-holiday Mondays. 75

Figure 4.14 Annually-averaged load profiles from Woodward during the study period. 76

Figure 4.16 As in Fig. 4.15, except for spring. 78

Figure 4.17 As in Fig. 4.15, except for summer. 79

Figure 4.18 As in Fig. 4.15, except for fall. 80

Figure 4.19 Annually-averaged load profiles from Woodward in 1999 that illustrate the differences in the weekday versus weekend load. 82

Figure 4.20 Daily load profiles for the four Mondays in September 1999 from Woodward. The holiday Monday (Labor Day) produced an electrical load with a different pattern than was observed on non-holiday Mondays. 82

Figure 4.21 Annually-averaged load profiles from Altus AFB during the study period. 84

Figure 4.22 Average hourly electric load profiles from Altus AFB during the winters of the study period. 85

Figure 4.23 As in Fig. 4.22, except for spring. 86

Figure 4.24 As in Fig. 4.22, except for summer. 87

Figure 4.25 As in Fig. 4.22, except for fall. 88

Figure 4.26 Annually-averaged load profiles from Altus AFB in 1999 that illustrate the differences in the weekday versus weekend load. 89

Figure 4.27 Daily load profiles for the four Mondays in February 1999 from Altus AFB. The holiday Monday (President's Day) produced an electrical load with a different pattern than was observed on non-holiday Mondays. 90

Figure 4.28 Annually-averaged load profiles from the Dominance substation during the study period. 92

Figure 4.29 Average hourly electric load profiles from the Dominance substation during the winters of the study period. 94

Figure 4.30 As in Fig. 4.29, except for spring. 94

Figure 4.31 As in Fig. 4.29, except for summer. 95

Figure 4.32 As in Fig. 4.29, except for fall. 95

Figure 4.33 Annually-averaged load profile from the Dominance substation in 1999 that illustrates load differences in the workday versus non-workdays (e.g., when the plant was closed, had scheduled downtime, or observed holidays). 97

Figure 4.34 Daily load profiles for the four Thursdays in November 1999 from the Dominance substation. The holiday Thursday (Thanksgiving Day) produced an electrical load with a much different pattern than observed on non-holiday Thursdays. 98

Figure 5.1 Scatterplot of hourly temperature from the Norman Mesonet site versus hourly electric load from the West Norman substation for 1999. 109

Figure 5.2 As in Fig.5.1 except for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 111

Figure 5.3 Diurnal plot of hourly temperature and hourly electric load from Norman averaged during weekdays, weekends/holidays, and a month for (a) August 1998 and (b) March 1998. 116

Figure 5.4 Hourly temperature and hourly electric load from Norman during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 121

Figure 5.5 Scatterplot of hourly relative humidity from the Norman Mesonet site versus hourly electric load from the West Norman substation for 1999. 126

Figure 5.6 Scatterplot of hourly relative humidity versus hourly temperature from the Norman Mesonet site for 1999. 126

Figure 5.7 Seasonally-averaged hourly electric load and hourly solar radiation from the West Norman substation and the Norman Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 129

Figure 5.8 Scatterplot of hourly temperature from the Woodward Mesonet site versus hourly electric load from the Woodward substation for 1999. 132

Figure 5.9 As in Fig. 5.8 except for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 135

Figure 5.10 Diurnal plot of hourly temperature and hourly electric load from Woodward averaged during weekdays, weekends/holidays, and a month for (a) August 2000, (b) July 1999, and (c) March 2000. 138

Figure 5.11 Hourly temperature and hourly electric load from Woodward during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 142

Figure 5.12 Scatterplot of hourly relative humidity from the Woodward Mesonet site versus hourly electric load from the Woodward substation for 1999. 145

Figure 5.13 Scatterplot of hourly relative humidity versus hourly temperature from the Woodward Mesonet site for 1999. 145

Figure 5.14 Seasonally-averaged hourly electric load and hourly solar radiation from the Woodward substation and the Woodward Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 149

Figure 5.15 Scatterplot of hourly temperature from the Altus Mesonet site versus hourly electric load from the Altus AFB substation for 1999. 151

Figure 5.16 As in Fig. 5.15 except for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 152

Figure 5.17 Diurnal plot of hourly temperature and hourly electric load averaged from Altus AFB during weekdays, weekends/holidays, and a month for (a) August 2000 and (b) March 2000. 157

Figure 5.18 Hourly temperature and hourly electric load from Altus AFB during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 162

Figure 5.19 Scatterplot of hourly relative humidity from the Altus Mesonet site versus hourly electric load from the Altus AFB substation for 1999. 165

Figure 5.20 Seasonally-averaged hourly electric load and hourly solar radiation from the Altus AFB substation and the Altus Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 167

Figure 5.21 Scatterplot of hourly temperature from the Broken Bow Mesonet site versus hourly electric load from the Dominance substation for 1999. 169

Figure 5.22 As in Fig. 5.8 except for (a) winter, (b) spring, (c) summer, and (d) fall of 1999. 171

Figure 5.23 Diurnal plot of hourly temperature and hourly electric load averaged over a month during a period of normal from Dominance for (a) February 1999 and (b) August 1999. 173

Figure 6.1 Observed air temperature from the Norman WFO, forecast temperature from the 20-km Eta Model, and forecast temperature from NGM MOS at Norman on 4 April 2003. 189

Figure A.1 A diagram of the system that produces electricity (illustration provided by WFEC). 207

Figure B.1 A map of WFEC member system's service area (illustration provided by WFEC). 210

Abstract

The power utility industry has become highly volatile with a deregulated market on the horizon and with enormous profit and loss swings in the energy trading market. Electricity, in particular, has become a commodity that is bought and sold at market prices, where load forecasting plays a crucial role in the composition of those prices. Public and private utilities must contend with the fact that a small error in an electric load forecast can create a large financial loss for the company. For example, underpredicting electric loads frequently results in an expensive purchase from the spot market that can eventually lead to a utility's demise. Overpredicting their system loads causes an unnecessary use of generation units, increased operating costs, and the loss of an opportunity to sell bulk power profitably. On the other hand, a small reduction of error in a load forecast can prove to be highly profitable for the utility. Hence, improving the accuracy of electricity load forecasts has become necessary for the long-term viability of all power utilities.

Weather has a significant impact on load demand and load forecasting. However, the weather-load relationship is unknown at the substation-level – mostly because substation-level load data have rarely been available to those outside the corporate infrastructure. Equally as important, most utilities have made inconsistent and antiquated use of weather data.

This study used electric load data from four substations in Oklahoma and concurrent weather observations from co-located Oklahoma Mesonet sites to: (1) determine the interrelationships between weather variables and electric load demand; (2) determine the impact of weather on the consumption of electricity by different customer

classes (e.g., residential, commercial, industrial); (3) establish thresholds of temperature associated with changes in the patterns of the use of electricity; and (4) produce load model simulations to quantify the improvements in the accuracy of a load forecast. This study also links a much improved, high-resolution numerical weather prediction model to a neural network load model to quantify the economic value of improved accuracy in load forecasts. In the end, this dissertation determined that a comprehensive understanding of the relationship between weather variables and electricity demand will improve the accuracy of load forecasting. The results of this study can save a small utility in excess of $0.5 million annually. If the results are applied to the larger power companies around the United States, a decrease in operating costs could exceed millions of dollars.

Chapter 1: Introduction

Electricity is a natural force that has always existed but has not always been able to be used. To use electricity we need to make it. In 1831, Englishman Michael Faraday became the first to generate electricity. He placed a magnet inside of a coil of wire and slowly rotated the magnet and immediately detected a current. This method, on a much larger scale, is still used to generate electricity today. As a result of this profound breakthrough and because of the technological advances since that time, we expect that, upon entering our driveway after a long day at work, motion lights in front of our homes will automatically turn on and garage doors will open with a click of a button. Once inside our homes, the flip of a switch or two illuminates the house. Another push of a button and the Nightly News appears on the television. We open our refrigerators to find cold drinks and leftover food that can be heated in just minutes using electrical power and a microwave oven. We have come to expect that electricity is available for our use whenever desired. We even tend to get irritated when our electrical service is interrupted, even if only for a few seconds. Western society has become fully dependent on electricity and perhaps spoiled by the continuous service which, if interrupted, severely limits our ability to function.

To achieve reliable electric utility service at all locations across the United States, research, planning, operating proficiency and risk on the part of the provider have been required. The continuous supply of electricity to our homes or businesses involves a three-step process – generation, transmission and distribution (Fig. 1.1; Appendix A). Larger utilities normally house each component, while smaller companies usually focus

on generation and transmission (G&T) or solely on distribution. Because consumers (e.g., residents, businesses, schools, hospitals, and factories) rely on continuous, affordable electric service for day-to-day functions, utility companies have had to develop a power system that operates at maximum efficiency. Failure to do so has a negative economic impact for the utility, the consumer, and ultimately the nation at large.

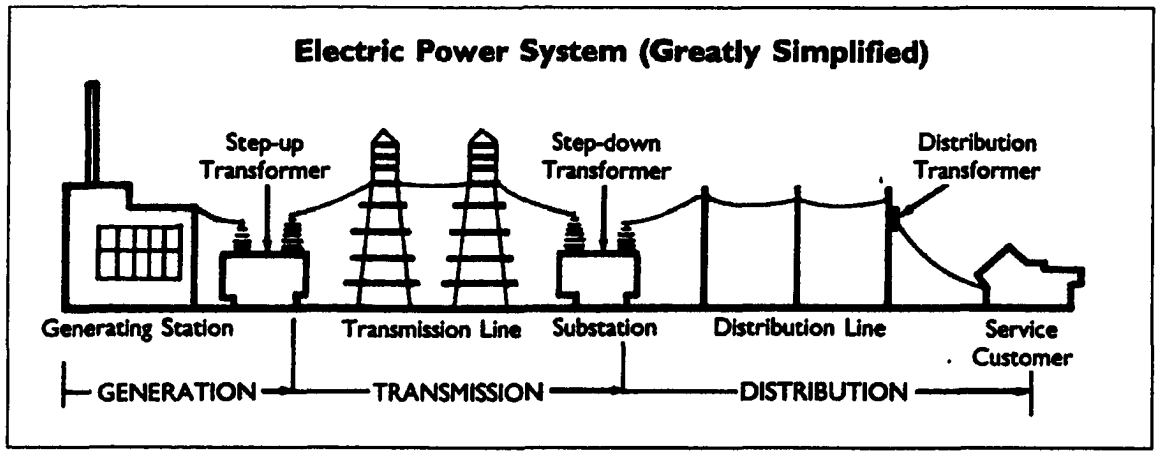


FIG. 1.1 An illustration of the three components involved to provide electricity to customers (illustration provided by WFEC).

Lack of efficiency in any component of the three-step process (to provide electricity) is costly to both the supplier and the consumer. As a consequence, heavy economic penalties result for the supplier due to increased operating costs. This economic impact is passed to the customer in the form of rate hikes or fuel adjustment charges. In addition, nationwide deregulation of the generation of electricity may be on the horizon in the not-so-distant future. By 1 January 2004, Oklahoma will allow the electrical customer to choose their generation supplier[1]. In other words, choosing our

[1] O.A.E.C (Oklahoma Association of Electric Cooperatives), 2001: O.A.E.C 2001-2002 Directory of Member Systems. 95 pp. [Available from O.A.E.C., P.O. Box 54309, Oklahoma City, OK 73111]

electricity provider will be similar to choosing a long-distance telephone provider[2]. Hence, to survive in a deregulated market, utility companies must offer continuous electric service at competitive rates. To do so, they must keep overhead costs to a minimum (i.e., they must operate efficiently). Finally, environmental issues such as air and water quality – out of necessity – must be incorporated into research and development for all utilities.

In the United States, electricity is generated primarily from the combustion of fossil fuels, namely, coal, natural gas, and oil. However, fossil fuels represent nonrenewable energy resources that eventually will be depleted. Rather than being forced into a reactive position caused by the depletion of fossil fuels, utilities must become proactive and prepare to incorporate renewable resources in their electricity generation (without sacrificing efficiency). The Renewable Energy Policy Project (REPP) has indicated that "...a 10% renewable base energy supply would lower the nation's energy bill by $15 billion per year by 2020 compared to a heavily fossil based supply mix"[3]. Wind energy, hydroelectric power, and nuclear power are among alternative generation methods being considered.

While the demand for electricity originates at the consumer level, meeting that energy demand is the sole responsibility of the utility companies. To meet these demands efficiently, utilities must predict accurately the electric load demand by their customers. In particular, short-term load forecasts, on the order of one hour to one day, have the most widespread implications for utilities and hence command much of the research

---

[2] Good Energy, 2002: http://www.goodenergy.com/electricity_deregulation/deregulation.asp

[3] Sterzinger, George, *REPP Testimony to Subcommittee on Energy Policy, June 17, 2002*

compared to all other time scales (Choueiki et al. 1997; Charytoniuk and Olinda 1998). The timeliness and accuracy of these forecasts have significant economic impacts to the operation and production costs at any utility. Thus, short-term load forecasts must provide critical information concerning the commitment of generation units, fuel allocation, real-time prices, economic allocation between plants, maintenance scheduling, load interruption and available transmission capability. For example, system dispatchers must anticipate their daily load patterns to generate sufficient electricity to meet the expected demand. Simultaneously, due to uncertainty inherent in the electrical forecast, an adequate amount of spinning reserve power must be available. The cost of accessing reserve power on demand is high because the electrical units in reserve are not necessarily operating at maximum efficiency at the time of an unexpected demand (Gross and Galiana 1987). Thus, a more accurate peak load forecast will reduce the need to access reserve power. Moreover, any error in a load forecast can have significant economic implications on operating costs. Overprediction of the load demand results in the unnecessary start-up of generation units and an increase in unused reserve power, which, in turn, increases operating costs. Furthermore, underprediction of the load demand often results in the purchase of power from the spot market which has unusually high costs. Hobbs et al. (1998) stated that, "A conservative estimate is that a 1% reduction in forecasting error for a 10,000 MW utility can save up to $1.6 million annually."

An accurate short-term load forecast (STLF) requires an understanding of the factors that cause fluctuations in the consumption of electricity. Typical load curves are well documented for the residential, commercial and industrial customer (Rastogi and

Roulet 1994). However, a major factor that creates deviations from these patterns, and hence presents difficulty in achieving accurate load predictions, is the day-to-day weather. It is well-known that weather is responsible for significant errors in load forecasts. The peak loads of most utilities have a large weather-sensitive component due to, for example, air conditioning units, space heating, agricultural irrigation and industrial demands. While forecasting the weather is a daunting task, using a weather forecast and its uncertainty to predict human behavior (electricity consumption) in response to that forecast is an enormous and complex challenge. Douglas et al. (1998) concluded that errors in a temperature forecast have a huge impact on the accuracy of a final load forecast. Therefore, accurate meteorological tools available to load forecasters should achieve a more efficient operation.

This study presents a unique research opportunity because it integrates hourly, proprietary load data at the substation level from a rural electric cooperative with nearby hourly weather data from a state-of-the-art meteorological observation network. The weather data available includes *all* of the parameters that may be important to load forecasting (e.g., temperature, humidity, wind speed, and solar radiation). Yet, many of these parameters have not been evaluated extensively in load models. Bunn and Farmer (1985) compared several models for electrical forecasting and recommended that:

> **"It would probably be useful to study the use patterns of residential, commercial, and industrial users separately...To implement this, however, requires that actual hourly load data for each user class be available. Is it? Could it?"**

Charytoniuk et al. (2000) proposed the same approach. This study will address their recommendation because substation-level data are available in three consumer categories. While the base load for electricity has been determined for each consumer class, the

effect of weather, which significantly alters the base load, is still unknown to the scientific literature.

This dissertation will evaluate whether a comprehensive understanding of the relationship between weather variables and electricity demand will improve the accuracy of load forecasting. The evaluation of this hypothesis will require an extensive analysis of weather variables using data from the Oklahoma Mesonet (Brock et al. 1995) and electric load data from Western Farmers Electric Cooperative (WFEC) for the period of 1998, 1999, and 2000. Subsequently, the economic value of these results will be quantified using a load forecast model based upon a neural network architecture.

Chapter 2 will overview the history of numerical weather prediction and the history of load forecasting and its relationship to weather variables. The data used in the study will be described in Chapter 3. Chapter 4 will document the (a) variations in the climate of Oklahoma during 1998, 1999, and 2000 and (b) the diurnal/monthly/seasonal variations of electric load demand for the same period. The interrelationship between weather variables and load data will be discussed in Chapter 5. Chapter 6 will present modeling studies, including a neural network, that incorporate results from Chapter 5 to quantify any improvement over the traditional load-modeling techniques. Any economic value from these potential improvements will be discussed as well. A summary of the results are presented in Chapter 7.

Chapter 2: Literature Survey

2.1 The Importance of Weather

Weather is an essential part of short-term load forecasting, no matter which approach to load forecasting is used. Several weather variables have been tested to determine which have the greatest impact on load demand. Variables such as temperature, humidity, wind speed, cloud cover, precipitation, visibility, and heating/cooling degree days have been used in various techniques. The essence of the work documented in more than five decades of research is that temperature has the greatest influence on load demand. Other variables associated with the variation in electric load demand are humidity in the summer and wind speed in the winter (Charytoniuk et al. 1998). However, with the arrival of competitive, possibly deregulated energy markets, load forecasters should begin to incorporate as much useful weather data as possible without losing parsimony in the load models.

2.2 The Meteorologist View

Meteorologists have a forecasting task, just as complex as that of load forecasters. Weather forecasters are responsible for predicting a multitude of weather variables (e.g., temperature, wind, relative humidity, precipitation) and storms across several time scales (e.g., one hour, one day, one week, one decade). Meteorological forecasts (and warnings) are disseminated to the public primarily for the purpose of protecting lives and property. The forecasts also provide a service to the public for the planning of work, travel, recreation and commerce.

Weather forecasts result from solving nonlinear, partial-differential equations that define the dynamics, thermodynamics, mass continuity, and moisture conservation of the atmosphere; these equations are known as the primitive equations of motion (Dutton 1986). Thus far, an analytical solution to the full governing equations has not been found. Two alternatives to this problem do exist. One option is to find an exact solution to a simplified form of the equations of motion. This task involves making assumptions and/or parameterizations for some of the unknown processes. This other method is to determine an approximate numerical solution to the full governing equations. This approach entails use of a computer to give an approximate answer to a complicated, nonlinear problem.

The origins of weather forecasting extend more than 100 years (Table 2.1). The earliest forecast methodologies included mapping weather variables and storms onto sequential charts to track the movement and acceleration/deceleration of various phenomena. The initial weather predictions were based mostly on persistence and trends, similar to the early methodologies used in load forecasting. Not surprisingly, the skill scores of weather forecasts which used persistence and trends proved to be nearly worthless. Subsequently, weather forecasting advanced during the 20th century with the introduction of numerical weather prediction (NWP) and many technological innovations.

Table 2.1 Significant milestones in the development of NWP.

DATE	IMPORTANT MILESTONES IN THE HISTORY OF NWP
1888	H. von Helmholtz formulated the primitive equations for fluid mechanics
1904	Vilhem Bjerknes first recorded the idea of NWP by discussing the application of physical laws (and primitive equations from Helmholtz) to the problem of predicting the atmosphere
1922	Lewis Fry Richardson published the first numerical weather forecast
1950	Charney produce a successful 24-hour forecast using a single-level barotropic model run on the new ENIAC
1955	Charney's Princeton three-level model was programmed for the IBM 701 and run on an operational schedule – ultimately provided little useful information
1958	Operational Barotropic model at NMC using objective analysis initial conditions and improved numerics
1962	First baroclinic model (3-level quasi-geostrophic model) became operational
1966	6-layer primitive equation (PE) model became operational due to major computer advancements; NWP was regarded as a useful forecasting tool
1971	LFM model (first regional model) implemented at NMC, remained in use for 20 years and became the basis for MOS
1980	Global spectral model (GSM) become operational
1985	Implementation of first comprehensive package of physical parameterization on GSM from Geophysical Fluid Dynamics Laboratory (GFDL)
1991	First operational 3D-Var (improvements in data assimilation methods)
1992	Ensemble forecasts became operational; improvements in 1994 and 2000
1993	First operational implementation of the Eta Model at NMC for North America at 80-km and 38-layer resolution twice daily
1994	RUC implemented for continental United States, with 3-hourly OI updates at 60-km resolution and 25 hybrid vertical levels
1996	Meso Eta model introduced with 29-km and 50-layers and improved model physics
1998	RUC upgraded to 40-km and 40 levels with extensive physics upgrades

2.2.1 History of Numerical Weather Prediction (NWP)

Through use of computing power, the goal of NWP is to provide weather forecasts beyond a few hours. NWP is an initial value problem (IVP) such that given an estimate of the present state of the atmosphere, the computer model simulates its

evolution. In order to make a skillful forecast for an IVP, it is required that the computer model be a realistic representation of the atmosphere and the initial conditions be known accurately (Kalnay 2003). The following NWP discussion is adapted from Shuman (1978), Shuman (1989), Kalnay et al. (1998), and Kalnay (2003).

In 1888, H. von Helmoltz of Germany formulated the primitive equations for fluid mechanics. Sixteen years later, Vilhelm Bjerknes of Norway suggested that those equations could be applied to the atmosphere. During World War I, Lewis Fry Richardson in England produced the first numerical weather forecast by performing a comprehensive numerical integration of the full primitive equations using hand calculators. Unfortunately, his forecast of surface pressure was in error by an order of magnitude mostly because the initial conditions were not balanced (i.e., they included fast-moving gravity waves with the slower weather-related oscillations). To no one's surprise, further (immediate) attempts were discouraged.

In 1945, electronic computers were invented, and one year later, John von Neumann organized the Electronic Computer Project at the Institute for Advanced Study (IAS) at Princeton University. Its purpose was to design and build the most powerful electronic computer to date. In 1948, Jule Charney created the Meteorology Group within that project – which included John von Neumann and renowned theoretical meteorologists Carl Rossby, Arnt Eliassen, and George Platzman – whose goal was to apply dynamic laws to the problem of weather forecasting. Within a year, Charney and Eliassen attempted to solve the problems Richardson encountered in his first prediction by deriving *filtered* equations of motion based on pressure fields alone. These equations

were based on quasi-geostrophic balance which "filtered" out the faster gravity and sound waves.

In 1949, the electronic numerical integrator and computer (ENIAC) was developed at the University of Pennsylvania and recognized as the first general-purpose electronic computer. Charney developed a single-level barotropic model (Charney et al. 1950) using the ENIAC and produced a "historic first" one-day weather forecast. This simple model was unable to convert potential energy to kinetic energy and explicitly predict storm development. Even so, the results of the first forecasts were encouraging as a pattern correlation existed between the forecasted and observed 24-hour pressure field changes, unlike Richardson's forecast attempt. It was after this successful use of a filtered model that Charney realized, to Richardson's credit, that significantly more progress in NWP would come from the use of the full primitive equations of motion. Charney also recognized the need for objective analysis of meteorological data to estimate initial conditions, replacing the labor- and time-intensive task of manual interpolation of available observations to grid points. In current models, this process has been improved through use of data assimilation techniques which uses short-term forecasts and observations to compute a sequence of initial conditions.

In 1953, IBM announced specifications for a new commercial computer, the IBM 702. The U. S. Weather Bureau (now known as the National Weather Service; NWS), the Air Weather Service of the U. S. Air Force, and the Naval Weather Service acquired an IBM 702 to launch a numerical weather prediction service. On 1 July 1954, the three agencies formed the Joint Numerical Weather Prediction Unit (JNWPU), under the authority of the Joint Meteorological Committee (JMC), to implement their strategy. The

JNWPU decided to use Charney's Princeton three-level model, a model that permitted the conversion of potential to kinetic energy and predicted the development of the Thanksgiving Day storm in 1950 over the northeastern United States (though, not in real time). Unfortunately, the numerical predictions could not compete with forecasts produced manually. Though the disappointment was great, this effort catapulted JMC to carefully attack the problems encountered and to focus on the accuracy and timeliness of forecasts for the operational community.

By 1958, research at the JNWPU and JMC lead to an automatic analysis system and automatic data handling. As a result, skillful, timely numerical forecasts were delivered to centrally-located forecasters. A single-level barotropic model was used to produce 500-mb forecasts across a limited domain (roughly North America and adjacent waters). These forecasts were more skillful (Fig. 2.1) than past modeling attempts (i.e., the mean-squared error of the forecasts had been reduced). Soon thereafter, the JNWPU was split into three organizations: the National Meteorological Center (NMC), the Global Weather Central (U. S. Air Force), and the Fleet Numerical Oceanography Center (U.S. Navy).

Ninety-five percent of the products produced by NMC were automated by 1960. In addition, the quality of products from NWP began to supercede those produced by manual methods. It was realized, though, that the model only had a single layer and was not able to predict the development of mid-latitude cyclones. This substantial problem was addressed using the three-level filtered-equation model which Cressman (1963) made operational at NMC. This baroclinic (multi-layered) model incorporated (Shuman 1989): (1) Charney's three-level model, (2) an additional term to account for the

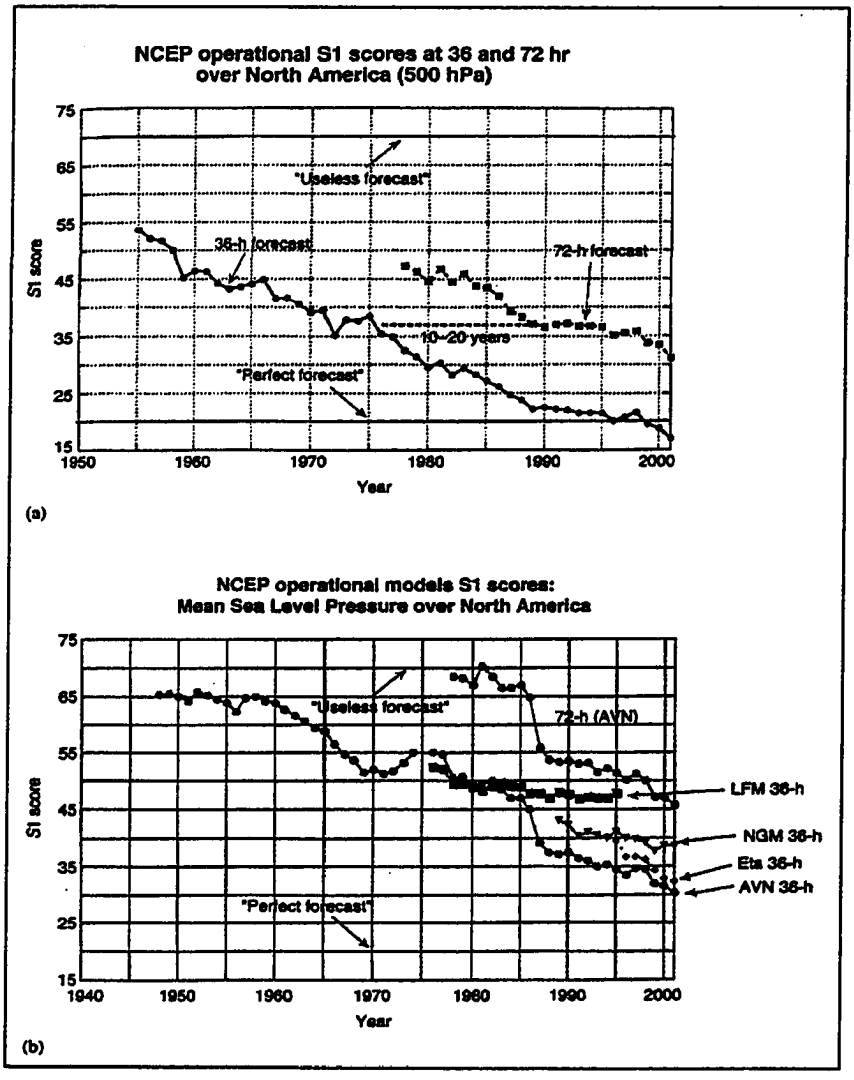


FIG. 2.1 (a) Historic evolution of the operation forecast skill of the NCEP models over North America (500 hPa). The S1 score measures the relative error in the horizontal pressure gradient, averaged over the region of interest. The values S1 = 70% and S1 = 20% were empirically determined to correspond respectively to a "useless" and a "perfect" forecast when the score was designed. Note that the 72-hr forecasts are currently as skillful as the 36-h were 10-20 years ago (data courtesy of C. Vlcek, NCEP). (b) Same as (a) but showing S1 scores for sea level pressure forecasts over North America (data courtesy C. Vlcek, NCEP). It shows results from global (AVN) and regional (LFM, NGM and Eta) forecasts. The LFM model development was "frozen" in 1986 and the NGM was frozen in 1991. (Figure and caption from Kalnay 2003)

advection of vorticity by the divergent component of the wind, (3) use of the balance equation, and (4) attention to detail in the formation of finite difference schemes and other numerical procedures to prevent a systematic accumulation of truncation error. The first baroclinic model became operational in 1962. The skill and usefulness of the forecasts at 500-mb continued to increase (Fig. 2.1). Thus, the model domain was expanded to include the entire Northern Hemisphere.

By 1963, the advancements in computer power and the discovery of relatively stable finite difference forms of the primitive equations allowed for the development of the six-layer primitive equation (PE) model. On 6 June 1966, the six-layer PE model at NMC became operational. As a result, the magnitude of the forecast errors clearly began to decrease (Fig. 2.1). Not long thereafter, the six-layer PE model became the first NWP model to produce a useful prediction at sea level. These forecasts were more skillful than those produced by manual methods. Finally, NWP had begun to be regarded as a valuable forecasting tool.

2.2.2 Types of NWP Models

Two types of models are used for NWP – global models and regional models (Kalnay 2003). Global models are typically used for guidance in medium-range forecasts (e.g., out to 16 days) as well as climate simulations. The horizontal domain of a global model is the entire earth, so these models usually cannot be run at high resolution. If a more detailed forecast is needed, it is necessary to increase the resolution and thus reduce the domain size. In that, regional models can be used. Regional models are run with a

resolution of 2 or more times higher than that of the global model and are used for shorter-range forecasts (e.g., 1-3 days). Regional models have an advantage over global models of higher accuracy along with the ability to resolve smaller-scale phenomena such as fronts, squall lines and orographic forcing. The disadvantage to regional models is that they require lateral boundary conditions at the borders of their horizontal domain, a problem that the global models do not encounter. To avoid this problem, regional models are often "nested" within more coarse models, whose forecasts can provide the necessary boundary conditions.

A further partitioning of NWP is hydrostatic versus nonhydrostatic models. A hydrostatic approximation regards the vertical accelerations in the vertical equation of motion negligible relative to the gravitational acceleration. This is an accurate approximation as long as the horizontal scales of motion (e.g., ~100 km) are much larger than the vertical scales of motion. The primary advantage to the hydrostatic assumption is that it filters sound waves (except those propagating horizontally) which allows the use of larger time steps and subsequent more timely completion of the forecast. On the other hand, the nonhydrostatic equations allow representation of smaller-scale phenomena such as convective clouds whose vertical accelerations cannot be neglected relative to buoyancy forces. Nonhydrostatic models with an efficient treatment of sound waves, however, are computationally competitive with hydrostatic models.

Beginning in 1963, Lorenz realized that NWP needed to account for the stochastic nature of the evolution of the atmosphere. Because the atmosphere has inherent instabilities, the growth of error due to the instabilities implies that a minute imperfection in a forecast model or the slightest error in initial conditions will inevitably lead to

depletion of skill in weather forecasts over a finite forecast period – Lorenz called this the "limit of weather predictability". To address the uncertainty of atmospheric model predictions, the concept of ensemble forecasting was implemented operationally in 1992. The ensemble forecast is one in which several model forecasts are performed by introducing small perturbations in the initial conditions or in the models themselves. The main achievements of ensembles were: (1) to provide an ensemble average forecast that, beyond the first few days, is more accurate than individual forecasts that occurred because the components of the forecast that were most uncertain tended to be averaged out; (2) to provide forecasters with a indication of the *reliability* of the forecast which changes from day to day and from region to region; and (3) to provide a quantitative basis for probabilistic forecasting. Several methods of ensemble forecasting have been developed since its inception.

From the 1970s to the present, several new numerical models were developed. These models included the Limited-area Fine-mesh Model (LFM), Nested Grid Model (NGM), Global Spectral Model (GSM), Aviation Model (AVN), Medium-Range Forecast Model (MRF), Rapid Update Cycle (RUC), and Eta Model. All contributed to the gradual reduction of forecast errors and provided more reliable details in the forecasts they produced.

2.2.2.1 Nested Grid Model (NGM) and Model Output Statistics (MOS)

Growth in computer power and gradual improvements in forecast quality have always gone hand-in-hand. As technology advanced, the models became more complex through the addition of more vertical layers, a finer mesh of grid points in the horizontal,

a fully global domain, detailed topography, and landscape characteristics. Parameterizations of physical processes such as radiation, clouds, precipitation and turbulence also improved (Stull 1995). In 1971, limited-area fine-mesh model (LFM) was created from the six-layer PE model at NMC. The Nested-Grid model (NGM) was created in 1979. Refinements were made through the 1980s until the model development was frozen. The NGM was developed as a grid-point model with 16 vertical levels and a horizontal resolution of 80 km for its inner grid and 160 km for its outer grid. The NGM also was a hydrostatic model that used a sigma coordinate system (a terrain-following coordinate system in which the ground is always at the lower coordinate surface).

Klein et al. (1959) introduced a statistical method of forecast refinement called the "prefect prog" method (PPM). The PPM used a best-fit multiple regression technique to develop a concurrent statistical relationship between the input fields (predictors) and the output fields (predictands). *Observations* were used for *both* the predictors and the predictands. As a result, an advantage of the PPM is that it does not depend on any particular forecast model and can be used immediately upon changing forecast models. On the other hand, the PPM produces the optimal predictand only in the rare event that the model produces perfect predictors (Stull 1995).

The PPM paved the way for its replacement, known as model output statistics (MOS; Glahn and Lowry 1972). MOS consisted of determining a statistical relationship between *model forecast fields* (predictors) and local weather observations (predictands), primarily through a multiple or stepwise regression technique (Jacks et al. 1989). The LFM was first used to develop MOS equations. An advantage to the MOS approach is that systematic model errors can be offset by the statistical regression. A disadvantage to

MOS is that a multi-year set of model output must be archived and statistically fit before the ensuing regression can be used for future forecasts (Stull 1995).

Consequently, an NGM-based MOS could not be developed until at least two years of model output data from a version of the NGM, which was not in a state of constant evolution, were available. In 1989, an NGM MOS package was implemented on a twice daily basis. This package included equations for maximum and minimum temperature, probability of precipitation (PoP), cloud amount and surface wind (e.g., parameters not directly predicted by the models). The myriad of predictors included NGM forecasts, surface observations, and climatological data.

The accuracy of the MOS guidance is strongly dependent upon the accuracy and consistent performance of the numerical model. Though the MOS technique is able to discern systematic errors in a model, it cannot resolve features that are below the resolution of the model. However, routine biases are accounted for via the MOS technique. Though statistical techniques can predict extreme events, MOS guidance tends to produce conservative forecasts and has often failed to forecast extreme events. As the forecast lead time increases, MOS produces guidance that trend toward the climatological value of a predictand. This tendency reflects the decreasing forecast accuracy that all numerical weather models suffer from as lead time increases.

2.2.2.2 Eta Model

The late 1980s and early 1990s brought an effort to increase the resolution (i.e., finer horizontal grid spacing and more vertical levels) of numerical models. The downscaling of these models provided forecasters with details on mesoscale weather

features which could not be resolved by synoptic-scale numerical models like the NGM or the AVN. In 1993, an early edition of the regional Eta model was implemented by NMC which replaced the LFM. The Eta model is a hydrostatic, grid-point mesoscale model that provides forecast guidance over North America (Black 1994). At that time, the horizontal and vertical resolution of the Eta was 80 km with 38 vertical levels. A new Eta coordinate system was defined to remove or reduce errors that occurred when computing the pressure gradient force, advection, or horizontal diffusion over steeply-sloped terrain. Each Eta surface is quasi-horizontal everywhere because it is normalized by a constant value of sea level pressure (instead of station pressure which varies greatly across mountainous terrain). Thus, the Eta coordinate system permits a more accurate calculation of the horizontal pressure gradient because errors caused by elevation changes between adjacent grid points were eliminated. Models that used sigma coordinates often produced unrealistic pressure gradients near sloped terrain because temperature changes on sigma surfaces were, in part, the result of an elevation change rather than an actual horizontal temperature change. Because temperature changes are much larger in the vertical compared to those in the horizontal, the vertical temperature gradient dominated pressure gradient calculations. Thus, the Eta model was developed to improve forecast guidance on the mesoscale, especially in areas with widely varying topography. The Eta guidance began with twice daily runs out to 36 hours.

During the late 1990s, the horizontal grid spacing of the hydrostatic Eta model was reduced significantly. In 1996, the National Centers for Environmental Prediction (NCEP; formerly NMC) introduced the Meso Eta model. This version had a horizontal resolution of 29 km with 50 vertical levels. NCEP continues to refine its Eta model – the

current operational version has a 12 km horizontal resolution with 60 vertical levels, issued every 6 hours.

2.2.2.3 Rapid Update Cycle (RUC)

Due to numerous observations obtained from the automated reporting by commercial aircraft and a demonstration network of wind profilers, the Forecast Systems Laboratory[1] (FSL) developed an analysis and data assimilation system known as the Mesoscale Analysis and Prediction System (MAPS); its forecasts were produced at 3-hour intervals. In 1994, NMC implemented an operational version of MAPS known as the Rapid Update Cycle (RUC). The first version of the RUC (RUC-1) had a horizontal grid spacing of 60 km and 25 vertical levels. The RUC produced three-dimensional (3-D) analyses and short-range forecasts out to 12 hours, every three hours. The analyses were based on a combination of observations (at asynoptic times) and a background field usually from the previous 3-hour RUC forecast.

The RUC uses a hybrid isentropic-sigma coordinate system. The isentropic coordinates are used for most of the atmosphere except near the ground where the terrain-following sigma coordinates are used. An isentropic coordinate system uses potential temperature (theta) as its vertical coordinate. When the atmospheric processes are adiabatic, air flows along isentropic surfaces. The isentropic coordinates also add resolution near frontal zones and the tropopause where theta surfaces are tightly packed. However, in regions where non-adiabatic processes occur (mixed layer; boundary layer), the same value of potential temperature may occur at more than one level above ground.

[1] FSL is a division of the National Oceanic and Atmospheric Administration/Office of Oceanic and Atmospheric Research (NOAA/OAR)

Under these circumstances, the sigma coordinate system provides the best solution for an NWP model.

On 6 April 1998, NCEP replaced RUC-1 with RUC-2, and improved the resolution to 40 km horizontally and 40 levels vertically. The RUC-2 produced 12-hour forecasts every 3 hours, and 3-hour forecasts every hour (i.e., the model runs become an hourly assimilator of observations). The RUC-2 expanded its domain east and west to cover 50% more area than did the RUC-1. The goal was to improve forecasts near coastal areas. Other improvements to RUC-2 eliminated known weaknesses in RUC-1 (Benjamin et al. 1998). On 17 April 2002, a major revision of the RUC was implemented, and the RUC20 (Benjamin et al. 2002) replaced the RUC-2. The RUC20 has a 20 km horizontal resolution and 50 vertical levels. Improvements included physics to better process moisture and better use of observations in the analysis. As with the Eta, improvements are continual.

2.2.2.4 Multi-Scale Models

The Penn State University (PSU) / National Center for Atmospheric Research (NCAR) mesoscale model (MM5) is a limited-area, compressible, nonhydrostatic, sigma-coordinate model developed to predict mesoscale and regional-scale atmospheric circulations. The MM5 is capable of real-data simulations on any scale, and the model is limited only by data quality, data resolution, and computer resources (Dudhia 1993). The MM5 is widely used for numerical weather prediction, air quality studies, and hydrological studies (Chen and Dudhia 2001) as well as for studies involving mesoscale

convective systems, fronts, land-sea breezes, mountain-valley circulations, and urban heat islands[2].

The Advanced Regional Prediction System (ARPS) is another multi-scale, atmospheric prediction system developed by the Center for Analysis and Prediction of Storms (CAPS) at the University of Oklahoma. It was primarily developed to serve as a prototype system for stormscale numerical weather prediction and has since extended its application to a multitude of idealized studies, operational analyses, and real-time forecasting[3].

2.3 Introduction to Electrical Load Forecasting

Electrical load forecasting is of paramount importance to the cost-effective operation and the day-by-day planning in the complex realm of managing electric power systems. The lead times of a load forecast can range from a few minutes to several decades (Hippert et al. 2001; Gross and Galiana 1987). This expansive time period is broadly divided into three main categories for power systems (Choueiki et al. 1997), and weather plays an important role in all three. One is long-term (econometric) forecasting with a lead time of 5 to 40 years. This forecast category is used for planning new systems, purchasing generation units, and building power plants or transmission lines. The second is the short-term load forecast with predictions that range from one day to one week. *In this study, the short-term forecast will mainly refer to electrical load forecasts on the order of one day.* The applications for short-term forecasts include fuel allocation, scheduling for generator and line maintenance, optimal commitment of

[2] www.mmm.ucar.edu/mm5/mm5-home.html
[3] http://www.caps.ou.edu/ARPS/index_flash.html

generator units, and buying/selling power. Third, very short-term load forecasts are on the order of minutes to an hour. Economic (load) dispatch is the sequence in which available generating units are called upon to serve the minute-by-minute fluctuations in load such that the cost of an operation is minimized. In addition, the very short-term forecast is used to monitor the security of the total power system, which includes proper channeling of electricity through transmission circuits and pinpointing potential overloads of the system during peak use periods.

Though the three categories of load forecasting are essential components in managing an electric power system, this study will focus on short-term load forecasting (STLF) and the meteorology that is involved. The STLF is critical in the day-to-day decisions required to operate an electric utility and the economic consequences of these decisions to consumers. Thus, STLFs have been the center of attention in the electrical engineering and power utility communities for several decades in a massive effort to improve the accuracy of operational decisions (IEEE Committee 1980). The short-term electric load is primarily dependent on nonlinear combinations of variables that have been classified by their dependence on weather, social and seasonal factors. The most important weather variable is temperature – past, current and future (Gross and Galiana 1987). Social impacts involve home, work, school, industry, and special events. Seasonal variations result from load growth, population growth and seasonal weather changes. Because of these factors and an apparent randomness to the load demand, short-term load forecasting remains an elusive challenge.

2.4 Short-term Load Forecasting Methods

A multitude of techniques have been used and are still being developed to solve the STLF problem. The early load forecast models were generally regression studies of load and weather data (Davies 1958; Heinemann et al. 1966; Matthewman and Nicholson 1968). Advanced time series techniques, such as autoregressive moving averages and the Box-Jenkins method, were soon incorporated into STLF methodologies (Gupta 1971; Galiana et al. 1974; Poysti 1984; Hagan and Behr 1987). Other methods surfaced in the 1980s that are still in use today. They include knowledge-based expert systems (Rahman and Bhatnagar 1988; Ho et al. 1990; Rahman and Hazim 1993), the state space method and the use of Kalman filters (Abu-El-Magd and Sinha 1981; Irisarri et al. 1982), and neural networks (Lee et al. 1992; Lu et al. 1993; Lamedica et al. 1996; AlFuhaid et al. 1997; Darbellay and Slama 2000).

Because no two power companies have identical load demands and load capacities, the methods used are often tailored to the needs and the computing power of the company. However, in developing a loose taxonomy for the different forecasting methods, three characteristics were identified (Bunn and Farmer 1985). Apart from other specific factors, applied short-term forecasting of electric load differs according to whether use is made of a standard load curve, how weather variables are used, and whether or not the load data are minute-by-minute spot data or data integrated over a period of time. It also is important to note that systems for forecasting electric load are divided into those which use on-line predictors versus those which use off-line predictors. This discrimination often dictates the type of method used. Systems that use on-line procedures adjust to the minute-by-minute evolution of the most current telemetered

demands and typically use time-series methods to forecast the load minutes in advance. On the other hand, systems that use off-line techniques are often applied to scheduling plant functions 2 – 7 days in advance. The off-line component typically employs regression models that use exogenous variables, such as meteorological factors.

2.5 Standard Load Curve Models

One popular forecasting tool is the standard load curve, often referred to as a 'nominal', 'base', or 'reference' load curve (Gupta and Yamada 1972; Papalexopoulos and Hesterberg 1990; Barakat and Al-Qasem 1998). The standard load curve is produced once per day and is based on a long history of seasonal and daily patterns. A basic equation to estimate the load using this method is:

$$L(t,d) = S(t,d) + R(t,d), \tag{1}$$

where L(t,d) is the actual load for the time t of the day d, S(t,d) is the standard load, and R(t,d) is the residual. The standard load curve needs re-scaling across specified time intervals; hence, the residual term is important. The standard load approach explicitly seeks to provide a daily series of residuals, which are adaptively tracked and forecasted. Thus, the conceptual idea is to update the standard load curve each day and update the residual every minute or so. Two types of load-shape models exist: (1) time-of-day models and (2) dynamic models.

2.5.1 Time-of-Day Models

The time-of-day models describe the load, z(t), at each discrete sampling time, t, over the duration of the forecast period, T, by a time series:

$$\{z(t), t = 1, 2, \ldots, T\}.$$

In its simplest form, this model stores T load values based on observed load behavior, where T is the duration of the forecast period. The past load behavior might be from the previous week or from a set of load curves for typical weeks of the year or typical weather patterns. The most common time-of-the day model takes the form of:

$$z(t) = \sum_{i=1}^{N} \alpha_i f_i(t) + \varepsilon(t), \quad t \in \tau \tag{2}$$

where z(t) is the load at time t and is considered to be the sum of a finite number of explicit time functions, f(t), which are often sinusoids with periods of 24 to 168 hours depending on the forecast lead time (Gross and Galiana 1987). The coefficient, α, is a slow-varying time constant while $\varepsilon(t)$ represents the modeling error, or white noise. The model is assumed to be valid over the time interval, τ, which must represent the recent past, the present, and a future time period otherwise known as the lead time. When explicit time functions, such as sinusoids are predetermined, the coefficients (α_i) are estimated through simple linear regression or exponential smoothing applied to a set of past load observations (Christiaanse 1971; Lijesen and Rosing 1971; Brubacher and Tunnicliffe Wilson 1976). Thus, these time series models essentially make use of historical load data for extrapolation to obtain the future load demands.

Structurally, these types of models are straightforward such that their parameters and the forecast can be updated easily through recursive algorithms as new load data are measured. On the other hand, these models do not represent the random nature of the load process or its relationship to weather variables. Thus, when an abrupt change in the weather occurs, the coefficients for the longer lead times create accuracy problems in the load forecast. Further, these models assume the trend is stationary and regard abnormal

data points as bad data. Abnormalities include labor strikes, major television events, or a temporary plant closure, all of which impact the load demand and should not be ignored.

Spectral decomposition is the basis for another time-of-day model. Here, the time functions, f(t), represent the eigenfunctions corresponding to the autocorrelation function of the load time series, once the regular patterns have been extracted (Farmer and Potton 1968; Pickles 1975). Residual values represent perturbations about the seasonal average due to changes in the local conditions, which in turn, impact the load demand. This method is more theoretically sound than the previous method due to its optimal choice of the time function. However, the spectral method is computationally expensive because it involves solving an eigenvalue problem (not necessarily an efficient operation when used in a real-time recursive algorithm). Finally, this method is susceptible to errors when weather conditions change rapidly because these influences are not explicitly modeled.

2.5.2 Dynamic Models

Dynamic load models recognize that, not only is the load dependent upon the time of day, it also depends on the most recent behavior of the load, on weather variables and on random inputs. Thus, dynamic models have, in essence, replaced the pure time-of-day models. Two basic types of dynamic models are pertinent to load forecasting: autoregressive moving average (ARMA) and state space models.

2.5.2.1 Autoregressive Moving Averages (ARMAs)

A general form of the ARMA(p,q) model for a stationary process y(t) with zero mean (Box et al. 1994) can be expressed as:

$$\varphi(B)y(t) = \theta(B)a(t) \quad (3)$$

where φ and θ are the parameters of the model for the autoregressive term of order p and the moving average term of order q, respectively, B is the backward shift operator, and a(t) is the white noise with zero mean and an unknown variance at equally spaced times, t. Many models are simply modifications of the ARMA (Moghram and Rahman 1989). If the process y(t) is not stationary, it can be transformed into a stationary process using a differencing transformation. The differenced, and now stationary, time series can be modeled as an autoregressive integrated moving average (ARIMA). The ARIMA(p,q,d) has an additional order term, d, denoting the number of times the series needs to be differenced to become stationary. Furthermore, if the process y(t) exhibits periodic behavior, which is common to power system processes, it can be removed using a seasonal difference operator. Thus far, the ARMA expresses y(t) in terms of its history and a white noise. However, if other factors, such as weather, impact the value of y(t), then a transfer function can be used to account for these variables in the model (Abu-El-Magd and Sinha 1982).

A number of methods used in STLFs identify the autoregressive and moving average parameters in Eq. (3). Though these methods are computationally expensive relative to the time-of-day models, they are more robust models which incorporate dynamic, weather and random processes. Ultimately, the ARIMA approach requires less parameter tuning which leads to more accurate load forecasts. One method of parameter identification is the use of the Yule-Walker equations in a recursive scheme, applied to the ARMA (Vemuri et al. 1973; Keyhani and El-Abiad 1975; Keyhani et al. 1975). This method involves obtaining a set of linear equations for the autoregressive parameters in

terms of the autocorrelations (of the stationary autoregressive process). The maximum-likelihood method (Gertler and Banyasz 1974; Hagan and Klein 1978), basically a nonlinear regression algorithm, is used to estimate the parameters. Once the parameters are determined, they must be updated. For ARMA models, daily updating is sufficient. Once the load data from the previous 24-hour period is free of anomalous behavior, it is added to the data set and data from the oldest 24-hour period are removed.

2.5.2.2 State Space Models and Kalman Filter

Because state space models can be converted into an ARMA model and vice versa, fundamental differences between the two do not exist (Ljung and Soderstrom 1983). Hence, only the basis of this model will be discussed. The state space model introduces the periodic component of the load as a random process. The load is modeled as a state variable using the state space representation which is comprised of two equations: the state (transition) equation and the observation (measurement) equation (Box et al. 1994). In its most general form, the state equation contains an unobservable state vector that summarizes the state of the dynamic system through time, and the measurement equation which indicates that observations consist of linear combinations of the state variables corrupted by white noise. The Kalman filter (Park et al. 1991; Infield and Hill 1998) is a popular and convenient method to estimate the state vector. The filtering procedure has a recursive "prediction-correction" or "updating" form and is attractive for on-line use. The difficulty in this process is that the noise covariance matrices are not easily obtained.

2.5.3 *Knowledge-Based Expert System (KBES) Approach*

Statistical techniques discussed thus far have shown that the methods require updating with changing conditions, such as weather and load demand. In some cases, new models must be developed when the dynamic of load management changes sufficiently. The knowledge-based expert system (KBES) approach (Rahman and Bhatnagar 1988; Ho et al. 1990; Rahman and Hazim 1993) represents a load forecast model built upon knowledge about the load forecast domain using human experts in the field. Once relevant knowledge about the forecast domain is extracted and built into the model, the facts are presented in an IF-THEN format as a "set of rules". This method uses a pairwise comparison technique (Saaty 1980) to prioritize and categorize the variables. For example, an algorithm has been developed that was based on the logical and syntactical relationship between the weather and the daily load shapes which have been converted to a set of rules. A rule-base was established to develop relationships between changes in the system load and changes in natural and forced condition factors, which in turn, create variation in electricity consumption (Moghram and Rahman 1989). Expert systems are responsive to changing conditions in its knowledge base which are easily expanded as new data becomes available. Fortunately, updates are fairly straightforward.

2.5.4 *Artificial Neural Networks (NNs)*

Artificial neural networks (referred to as neural networks; NNs) have undoubtedly received the most attention of all STLF techniques. This attention is mainly because NNs can learn complex and nonlinear relationships that often are difficult to model with other

statistical techniques. Using historical load and weather data, NNs are able to model correlations between factors that include weather conditions, day of the week, time of the day, and past usage patterns. Though several have enjoyed success with the use of NNs (AlFuhaid et al. 1997; Khotanzad et al. 1998; Drezga and Rahman 1998), skeptics believe that they do not outperform traditional load forecasting methods (Adya and Collopy 1998; Darbelly and Slama 2000).

Neural networks refer to a computing system which can mimic a biological neural network (Mehrotra et al. 1997). Its basic unit is the neuron, which receives information through a number of input nodes, processes it internally, and produces a response (Fig. 2.2). Within the NN, it should be noted that several stochastic techniques previously mentioned are used. The off-line component of the neural network involves a training (process of parameter estimation) algorithm based upon historical load and weather data. The on-line component uses the multilayer feed-forward technique for the forecasting application.

Zhang et al. (1998) formally demonstrated that NNs were able to approximate any continuous function to an acceptable level of accuracy. Thus, NNs model complex nonlinear relationships better than conventional nonlinear (or linear) models. In addition, researchers do not have to postulate their models and estimate parameters, for NNs are data-driven and are able to automatically map a relationship between a given sample of inputs and outputs (Hippert et al. 2001). Thus, because of the architecture, NNs are well suited to extract patterns from past events and to extrapolate into the future, a necessity in load forecasting.

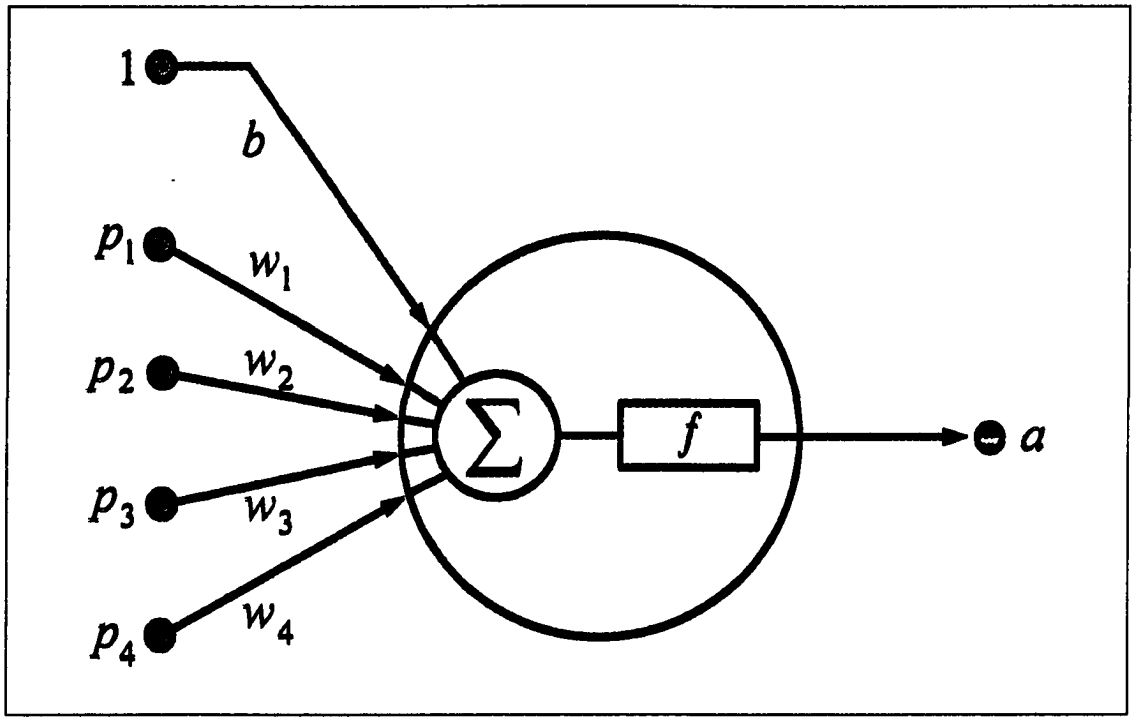


FIG. 2.2 An artificial neuron where the $\mathbf{p}_i$s are the four inputs, b is the bias term, and the $\mathbf{w}_i$s are the weights (weight matrix), all of which are combined and fed into a function, f, to produce an output, $a = f(\mathbf{w}_i\mathbf{p}_i + b)$.

It is noteworthy to mention two points of contention among skeptics. Because hourly data are necessary in short-term load forecasting, in several cases the NN architecture appeared to be too large for the data samples they intended to model. Consequently, instead of learning general relationships within the training data set, the NN tends to memorize the training data. The result is an overfitted data set that, in principle, would not produce accurate forecasts when independent data were used. Previous NN tests in electric load forecasting did not compare results with standard benchmarks of accuracy and did not make use of available graphical and statistics tools. While possible overfitting of the training set and lack of standard comparisons did not confirm that forecasts using NNs were less accurate than traditional methods, skeptical

reviewers still question the myriad of papers (Chen et al. 1992; AlFuhaid et al. 1997; Choueiki et al. 1997) which reported increased accuracy in load forecasts using NNs.

2.5.4.1 Neural Network Architecture

Neural networks, in a forecasting capacity, are extremely flexible models that can adequately model complex behavior and nonlinear relationships exhibited in the STLF problem (Choueiki et al. 1997). Because of its flexibility, the design and construction of NNs are far from trivial. Four basic tasks are required to design a neural network: (1) data preprocessing; (2) designing the NN; (3) training the NN; and (4) validating the NN. Each task, however, has components that may overlap with another.

The purpose of data preprocessing is to get a handle on the forecasting problem. It may be needed to reduce the dimension of the input vectors (i.e., avoid the "curse of dimensionality"), to clean/filter the data, and/or to classify the input data to keep the model as simple as possible. Preprocessing also is used to determine the shape of the load profile. Load profiles, typically a series of 24 hourly loads, basically are partitioned into weekday and weekend/holiday profiles. Poor results occur when the distinction is not made between these two categories (Lu et al. 1993). Other classes are sometimes made for days of the week, seasons and regularly occurring events (e.g., Super Bowl Sunday). The next most important factor that impacts the load profile is weather. Thus, days may be classified according to weather conditions by using a statistical relationship.

The completion of the preprocessing provides useful information when considering the design of the neural network. Selecting the appropriate "workhorse" for the NN architecture is the first step in the design process. The most popular workhorse

for forecasting problems is the fully-connected, feed-forward, multilayer perceptron (Fig. 2.3). Nonfully connected networks (Chen et al. 1992) and recurrent, or feedback networks (Vermaak and Botha 1998) also have been used, but those methods are not as popular and will not be discussed. Fully-connected means that every node is connected to every node in the network. A feed-forward network implies that there is no feedback between the layers. The multilayer perceptron (MLP) leaves the designer to determine the number of output nodes, the number of input nodes, the number of neurons in the hidden layer, and the number of hidden layers. Selecting the number of output nodes requires a decision of what is to be forecasted. The first option is the one-output NN, which produces one-step ahead forecasting for the next day's peak load, the next day's

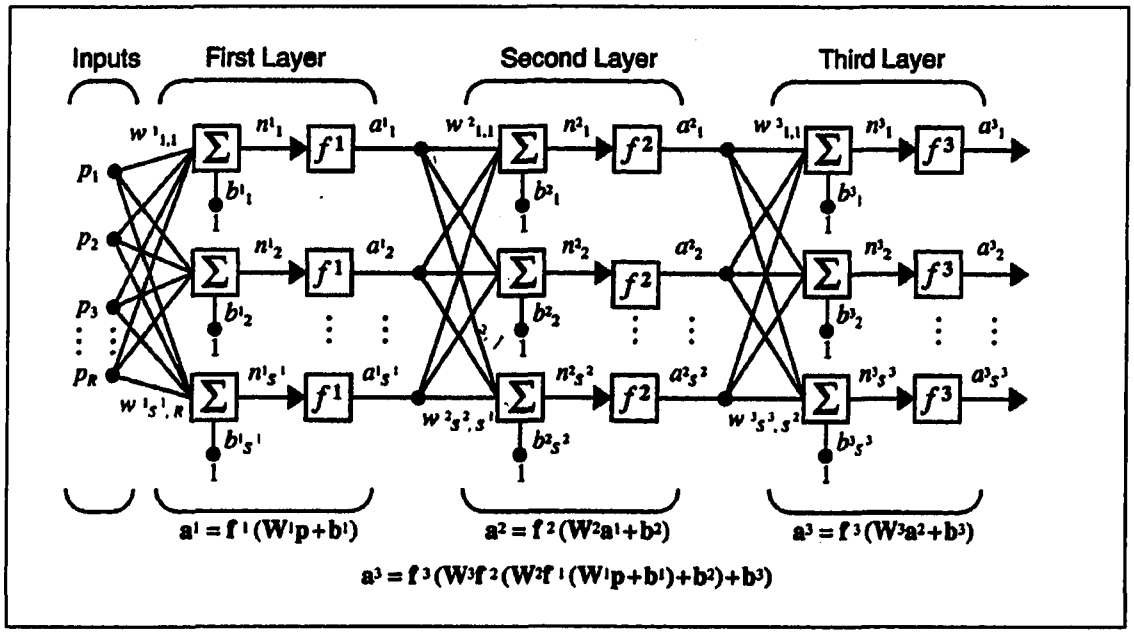


FIG. 2.3 A 3-layer network where each layer has its own weight matrix **w**, bias vector **b**, net input vector **n**, and output vector **a**. The first and second layers are the hidden layers, while the third layer is the output layer. (Figure from Hagan et al. 1996)

total load, or the next hour's load (i.e., given the load series up to hour h, forecast the load at hour $h+1$). The second option is the multi-output NN, which produces a load profile (e.g., 24 outputs – one for each hour of the day). Selecting the number of input nodes is the next step for which there is very little theoretical assistance. The designer must have some *a priori* knowledge about the behavior of the system and the factors that impact the system being studied. The most popular inputs are the load itself, the load from the past day or the past two days, and temperature data (or some statistical relationship between temperature and load). The last and perhaps most difficult step in the design process is choosing the number of hidden neurons. Again, the theoretical basis for this choice is very limited. Thus, trial and error often determines the number of hidden neurons. Note, though, that too few neurons will not be able to accurately model the load, and too many neurons will cause the network to memorize the training data. Thus, selection of the number of neurons is an important issue in optimal network design. Simulations are done and the number of hidden layers is chosen from the run with the best fitting performance, keeping in mind that too many hidden layers affect the training time.

Once the MLP is designed, the NN must be trained using a training algorithm. The most common choice of training algorithms is the backpropagation (BP) learning rule. The BP algorithm is a generalization of the least mean-squared algorithm that adjusts the weights to minimize the mean squared error between the desired and actual outputs of the network (Mehrotra et al. 1997). The BP method uses supervised learning where the network is trained using data for which the desired input and outputs are known. In other words, the output produced by the NN in response to the inputs is

repeatedly compared with the known (correct) answer. After each iteration, the weights are adjusted towards the correct answer by backpropagating the error at the output layer through the NN according to the steepest descent method. (The steepest descent has the advantage of being simple and only requires the evaluation of the gradient of the function.) Because the training step is an iterative process, however, the criterion that stops the iteration must be defined. The criterion can be defined by selecting a fixed number of iterations or choosing a specified tolerance level for the errors. However, this step of the procedure introduces overfitting which implies one of two things: (1) a model has fit the data so well that it ends by including random error which creates poor forecasts for independent data sets; or (2) the model was overparameterized or excessively complex. The first problem can be overcome by cross-validation (Khotanzad et al. 1998) or by regularization techniques (Hippert et al. 2001). The second is solved by reducing the number of neurons or eliminating some of the connections to form a less complicated model.

Finally, the NN must be validated by comparing error statistics with those from standard forecasting techniques. Often though, when comparing the NN to a regression model, as much effort is required to fit the regression model as is required to fit the MLP. Nonetheless, to determine whether or not the NN should be considered as a "well-accepted" method, it must be validated by statistical measures such as a cumulative distribution of errors, mean-squared errors, percentiles of errors, mean absolute percent errors, or serial correlations of errors through graphical means (e.g., scatterplots).

The NN architecture, like any other model, requires a complete understanding of the problem before the model should be developed. Because NNs are so flexible, several

factors should be considered in the design. The NN technique is attractive in the load forecasting process because it adequately handles nonlinear relationships, such as the temperature-load relationship, and allows the designer to choose the nature of the inputs and outputs without completely creating a new network.

2.5.4.2 Neural Networks and Meteorological Forecasting

NWP models provide guidance that help predict various weather parameters. When the forecasts contain systematic errors or biases, postprocessing of the output data can improve the raw output. Many statistical methods have been used for postprocessing meteorological data. For example, the "prefect prog" technique (Klein et al. 1959) was among the first; it was later replaced by MOS guidance (Glahn and Lowry 1972). These multiple regression methods used model output and converted it into sensible weather forecasts. Hall et al. (1999) applied the neural network scheme to develop a precipitation forecasting tool. Though NNs are relatively new to weather forecasting, the NN produced an acceptable forecast for the PoP and the quantitative precipitation forecast (QPF). Kuligowski and Barrros (1998) had similar results for a QPF that used an NN.

Thus, it is possible that NNs could become a useful tool for application to various meteorological forecast problems. While NNs are not in widespread use among the meteorological community compared to its use in the electrical engineering community, the early results are encouraging. For example, the improved forecast of thunderstorms (McCann 1992), tornadoes (Marzban and Stumpf 1996), and snowfall (Roebber et al. 2002) are early success stories.

2.5.5 STLFs Using Weather-Load Models

The fact that volatility in electric load demand is, in part, attributable to weather has been acknowledged for more than half of a century (Davies 1958; Bunn 2000). However, the use of weather information and the recognition of its importance in short-term load forecasting has varied during the past several decades. Yet, most STLF techniques include weather data in some form – historical, explicitly or via a weather-load model.

Many ARMA models include weather as an input variable (Van Meeteren and Van Son 1979; Hagan and Behr 1987). Those that do not include weather usually update various parameters automatically to take into account the impact of meteorological variations on the load. However, this inept approach creates an unsatisfactory load forecast under rapidly changing weather conditions (even as the load process is assumed to be stationary). Those techniques that do account for weather either use it as an explicit input variable (Keyhani and Miri 1983; Poysti 1984) or they rely on a heuristic approach whereby the load process is corrected for the influence of weather prior to applying the ARMA model (Ernoult and Mattatia 1984). Yet, knowledge-based expert systems are as heavily dependent on weather data - both historical and forecasted - as are the systems built around neural networks.

Some power system forecasters use an explicit weather-load or a weather sensitive model as part of their STLF regime. Gupta and Yamada (1972) and Van Meeteren and Van Son (1979) began with a stochastic model to relate future and past loads. The decomposition of the hourly load, Z(i, j), at hour (j) and day (i) had three components which were initialized off-line:

$$Z(i, j) = T(i, j) + WC(i, j) + X(i, j) \quad (4)$$

where T(i, j) is the basic component of the load at hour (j) and day (i) which more or less is considered to be constant everyday; WC(i, j) is the weekly cycle (or the day-of-the-week effect) component of the load at hour (j) and day (i) which is slowly changed to represent the weekly pattern of hourly loads; and X(i, j) is the residual component which contains the effect of weather variations. Accordingly, this third component represents a rapidly changing component and reflects hour-to-hour variations in the load due to random factors. This weather-sensitive component is typically modeled using an autoregressive technique. Heinemann et al. (1966), Stanton and Gupta (1970), and Gupta and Yamada (1972) went a step further and incorporated a linear weather-load model to forecast the peak load of the day. This approach to modeling has a basic load component (of the peak load), a weekly load component, a weather-sensitive component and a random component. (A nonlinear transformation of the temperature variable was needed to formulate the linear load model.)

2.6 Merging Load Forecasting and NWP

Meteorologists and load forecasters have been aware for several decades that variations in temperature and humidity impact both the peak load and the total load for particular power systems (McQuigg et al. 1972). Though traditional methods of load forecasting are still in use, the neural network seems capable of integrating statistical methods from NWP models of the 21st century, nonlinear relationships, and exogenous variables. The offline component requires historical hourly load and actual weather data to train the model, which are available from the National Climatic Data Center.

However, to develop the best statistical relationships between weather and electric load data, the time and spatial scales must be consistent between the two data sets. In addition, the time scale should not have a resolution less than one hour to avoid abrupt changes which can occur in the load and the weather conditions. Otherwise, errors could accumulate in training the NN and ultimately affect the online forecasting module.

The load forecast from an online component of the NN could be improved if the careful attention was given to the choice of the weather forecasting tool. While NWP models have rapidly improved in recent years, several electric companies and many load modelers have failed to keep pace with the rapidly evolving NWP models capable of producing storm-scale forecasts. Utility companies typically receive their weather forecasts from private firms, use older forms of MOS guidance, or generate weather forecasts internally. Furthermore, only temperature data have been routinely applied because little else existed (Hippert et al. 2001). The standard operating procedures remained tied to traditional synoptic-scale models and dated forms of MOS guidance to generate load forecasts. Yet, Khotanzad et al. (1997) documented that the weather forecast introduces approximately 1% additional error to the load forecast out to 1 - 2 days ahead. The load forecast errors are even greater when they are tied to long-lead weather forecasts. Because a small error in temperature forecasts can waste thousands of dollars per day, this dissertation will provide evidence that modern day NWP models can be coupled into the short-term load forecasting process and improve the economic value of the load forecast.

Chapter 3: Data and Sites

Two types of historical data were used in this study: hourly observations of weather variables and electric load data. Hourly electric load data from the three-year period of 1998, 1999, and 2000 were obtained from four substations owned by Western Farmers Electric Cooperative. Hourly meteorological data for same three-year period were obtained from four Oklahoma Mesonet sites in close proximity to the four substations. This chapter provides relevant details about the two data sets, the geography of Oklahoma, and an overview of each of the four testbed sites.

3.1 Weather Variables

Through a collaborative effort by many scientists in Oklahoma and a partnership between the University of Oklahoma and Oklahoma State University, the Oklahoma Mesonet (Brock et al. 1995) was developed. The Oklahoma Mesonet is an automated meteorological network of 115 evenly spaced stations – with at least one site in each of the 77 counties of Oklahoma (Fig. 3.1). This densely-spaced network was designed to provide research-quality data on a time and space scale appropriate to detect mesoscale[1] weather phenomena. Each station measures air temperature and relative humidity both at 1.5 m and 9 m, wind speed and direction both at 2 m and 10 m, barometric pressure, solar radiation, rainfall, and soil temperature at several depths. These measurements are acquired in real time in five-minute intervals. The Mesonet began recording nearly 1 million observations per day on 1 January 1994. The Oklahoma Climatological Survey

[1] Mesoscale refers to weather events that range from a few kilometers to a few hundred kilometers in space and several minutes to several hours in time.

(OCS) has the responsibility to collect the observations, provide data quality assurance, and share the data with the public and private sectors.

The meteorological data used in this study were obtained from four Mesonet sites: Norman, Woodward, Altus, and Broken Bow. Based upon the scientific literature, the predominant weather parameters deemed important to predict electric load demand are: temperature, relative humidity and wind speed. Thus, hourly values of air temperature (°F, °C) and relative humidity (%) at 1.5 m, wind speed (mph, ms^{-1}) at 2 m, and solar radiation (Wm^{-2}) were acquired for each day during the three-year study period (1998, 1999, and 2000).

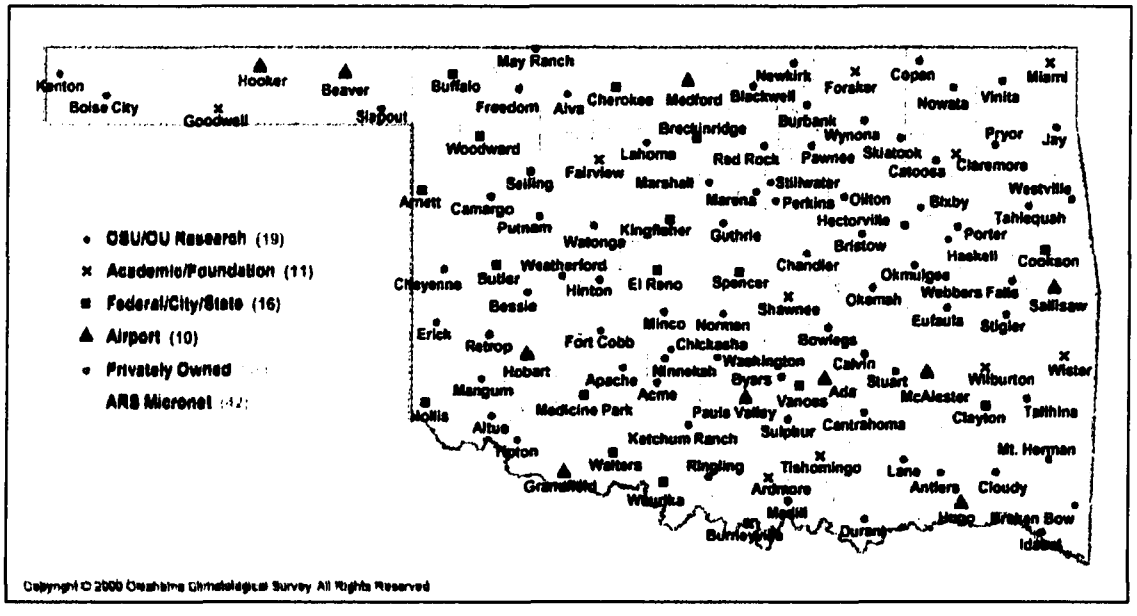


FIG 3.1. The Oklahoma Mesonet sites.

3.2 Electric Load Data

Western Farmers Electric Cooperative (WFEC) is a generation and transmission (G&T) cooperative (co-op) headquartered in Anadarko, Oklahoma. WFEC is owned by its 19 Oklahoma member distribution co-ops, all of which are members of the Touchstone Energy Alliance of electric cooperatives. WFEC also is responsible for the electricity supply to Altus Air Force Base. WFEC owns three generation plants in Oklahoma. The Hugo plant, located near Fort Towson, operates a 400 megawatt (MW) coal-fired generation unit. The Anadarko plant operates three Combined Cycle units for a total of 300 MW of generating capacity. The Combined Cycle units add efficiency to the generation process by combining natural gas and power from a steam turbine to turn its generator. The Anadarko Plant also operates three natural gas units that often are on standby. Finally, the Mooreland Plant operates three natural-gas-fired units that have a combined output of 304 MW. In total, WFEC has a power supply capacity (including generating units and hydro power allocation) of 1,300 MW. The transmission facilities of WFEC include over 3,400 miles (5,472 km) of transmission lines and more than 225 substations.

WFEC is responsible to its customers for generating or purchasing wholesale power, whichever is cheaper, and for transmitting that power to their various substations for distribution to homes and businesses. WFEC sells this power to its member distribution cooperatives, which have exclusive responsibility for the power once it leaves a substation. Hence, high voltage power sent to each substation uses *transmission lines* owned by WFEC. At the substation, the power flows into a transformer box, is powered down, and sent as low voltage power through the *distribution lines* owned by

the distribution co-ops. The distribution co-ops sell/distribute electricity in the WFEC system to its individual customers.

The second category of data in this study is the hourly electric load from each substation within the distribution cooperatives. In other words, the electrical data are the number of kilowatt hours (kWh) that WFEC sold the distribution co-ops for distribution to homes and businesses. A kWh is a unit of electrical energy which is equivalent to 1000 watts of power used for one hour. For example, an average household will consume between 800-1300 kWh per month[2] (Appendix C). It is important to note that the number of kilowatt hours sold by WFEC to the distribution co-ops normally exceeds the sum of the kilowatt hours read from individual meters of consumers. This inequity is created by line loss (i.e., energy loss and capacity loss from moving power through conductors or related equipment). Line loss results primarily from core loss within the transformers (heat loss from coils), trees near power lines that bleed power from the system, or dishonest meter readers. The distribution co-ops attempt to minimize this loss. Even so, the co-ops recoup this loss by adjusting rates. Thus, WFEC considers these small but important issues when forecasting the electrical load for its system.

The available combination of hourly, substation-level electrical load data and hourly weather parameters from co-located weather towers is unparalleled in the scientific literature. Because of the complicated issue of deregulation and increased competition between utilities and energy traders, electric load data remains proprietary. However, in the interest of learning more about the relationship between weather and electric load demand, WFEC agreed to share their load information for research purposes only. Furthermore, because stations in the Oklahoma Mesonet are spaced at a resolution

[2] WFEC Glossary of Terms: www.wfec.com/glossary

of ~12 miles (19 km), quality-assured meteorological data were within ~4.3 miles (6.9 km) of each substation. Most customers are within ~15 miles (24 km) of each substation. Hence, the combined data set represents a unique opportunity to better understand the habits of localized customers and determine how they respond to changing weather conditions.

3.3 Geography of Oklahoma

The characteristics of each Mesonet site and substation are dependent upon their location in the state. The diverse but natural environment of Oklahoma has been divided into ten regions based on the physical characteristics (Fig. 3.2) within each region (Wikle 1991). While Figures 3.3-3.8 provide supplemental information about these features, attention is focused on those regions that include the four sites used in this study.

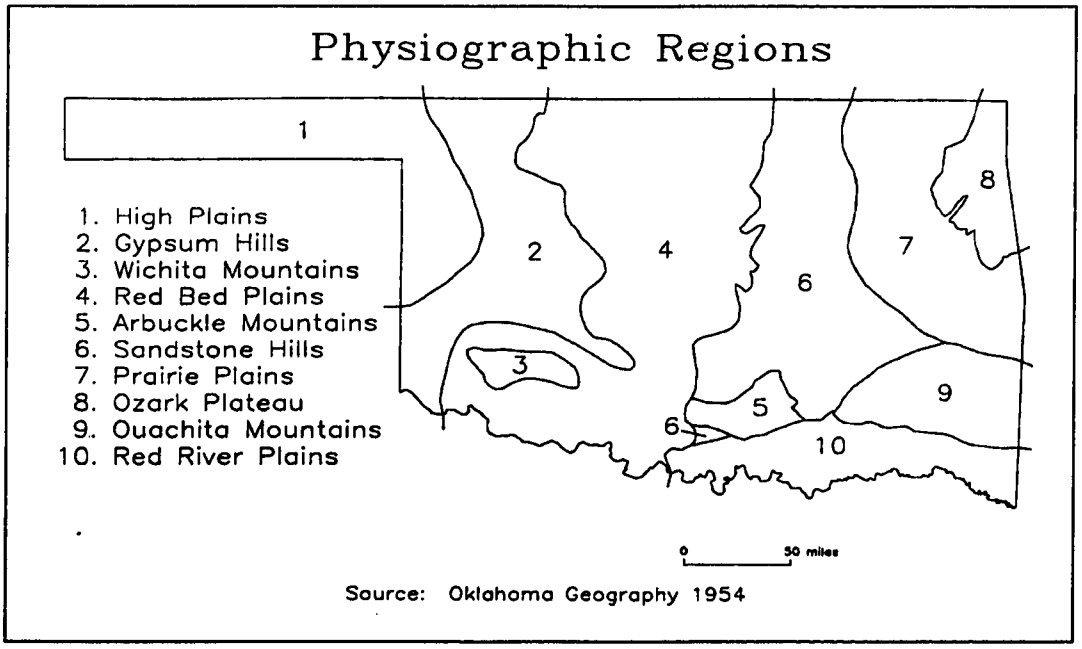


FIG 3.2. The physiographic regions of Oklahoma as described by Wikle (1991).

The plains of Oklahoma – in particular, the High Plains (Great Plains), the Red Bed Plains (Osage Plains), the Prairie Plains (Eastern Lowlands), and the Red River Plains (Coastal Plains) – are located in Regions 1, 4, 7, and 10, respectively. Though the Great Plains and Coastal Plains are characterized by a relatively flat, monotonous landscape, they are quite different. Region 1 has an average elevation of nearly 2000 feet above sea level, whereas Region 10 contains the lowest terrain in Oklahoma at 287 feet above sea level (Morris 1977). The minimum rainfall occurs across the Great Plains while the maximum rainfall and the most subtropical region in Oklahoma are found in the Coastal Plains. When water is available, the Great Plains have productive soils; thus, many Oklahoma crops are produced in northwest Oklahoma (Fig. 3.3). The ratio of farm income to total income in Figure 3.4 also confirms this characteristic of the Great Plains. However, more people live in southeast Oklahoma than live in the Oklahoma Panhandle (Fig. 3.5).

The Red Bed Plains of Region 4 is home to four major Oklahoma urban centers in the WFEC distribution area (Oklahoma City, Enid, Lawton and Altus); these centers are the most densely populated regions of Oklahoma (Figs. 3.5 and 3.6). Winter wheat is the major cash crop in this region. Soils are relatively productive compared to other sections. As a result, extensive farmlands and relatively high rural income are common.

Grasslands dominate (Fig. 3.7) Regions 2 and 6 (known as the Gypsum Hills and Sandstone Hill, respectively). The region becomes more arid and the entire landscape opens up due to low humidity (and less cloud cover) creating "the big skies of the West" (Späth et al. 1998). A portion of the winter wheat belt lies in Regions 2 and 6. The hills of Region 6 are forested areas relative to those in Region 2 (Fig. 3.8).

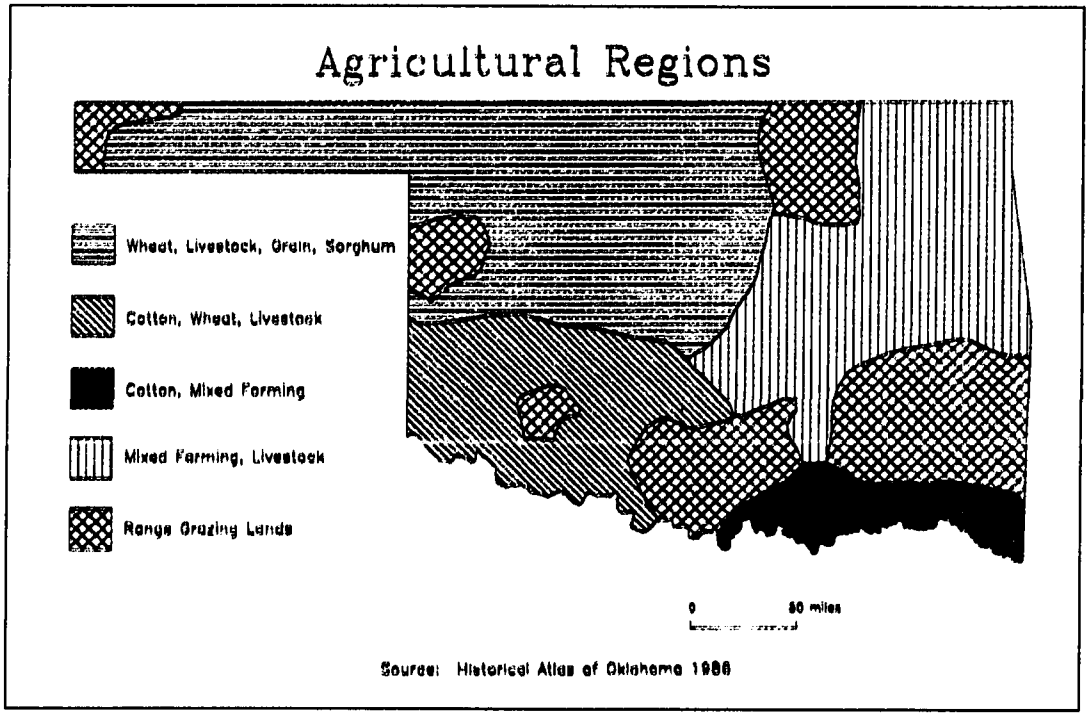


FIG 3.3. As in Fig. 3.2 except the agricultural regions are shown.

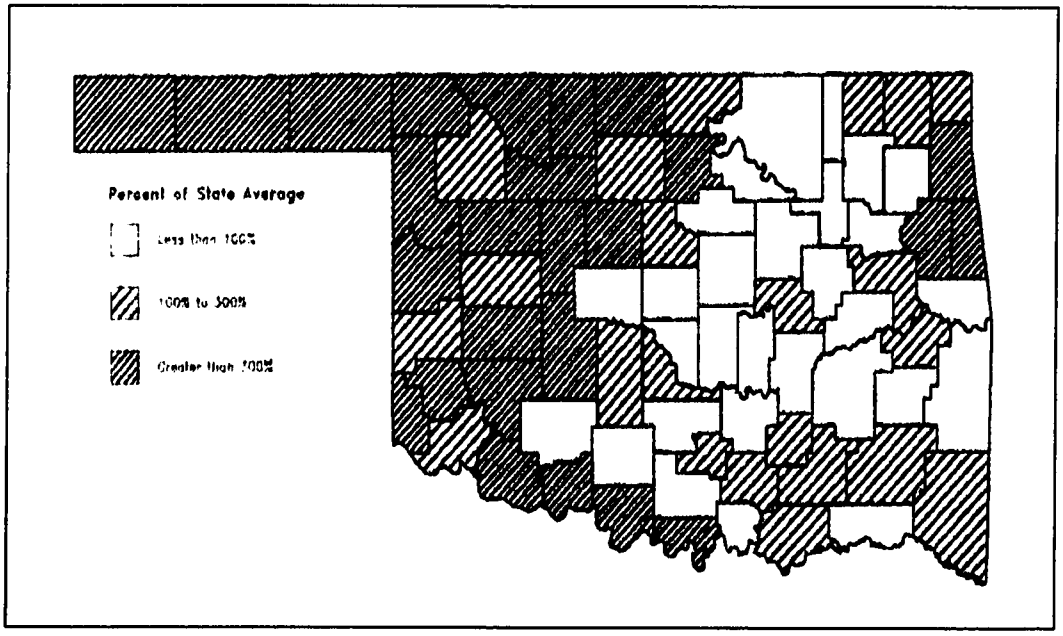


FIG 3.4. The ratio of farm income to total income by county in Oklahoma as described by Späth (1998).

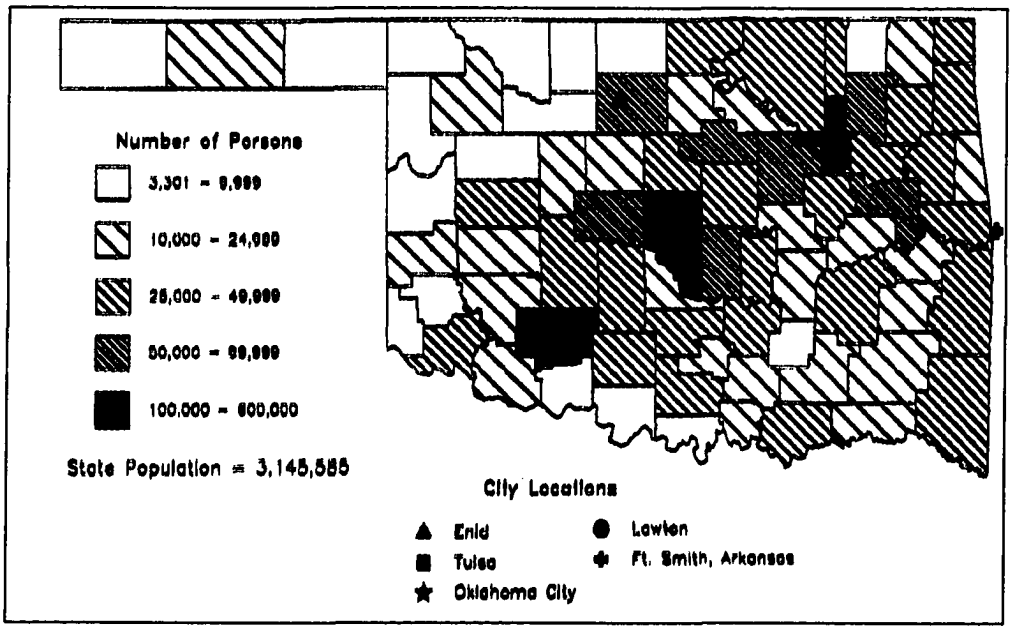


FIG 3.5. The population per county in Oklahoma as described by Späth (1998).

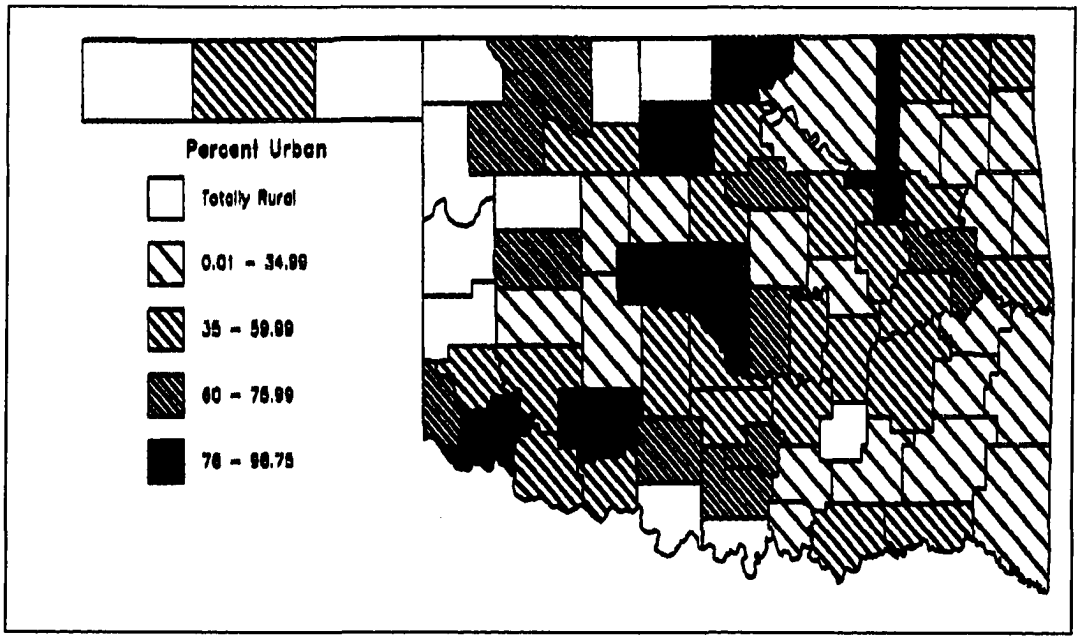


FIG 3.6. The percentage of each county in Oklahoma that is urban (versus rural) as described by Späth (1998).

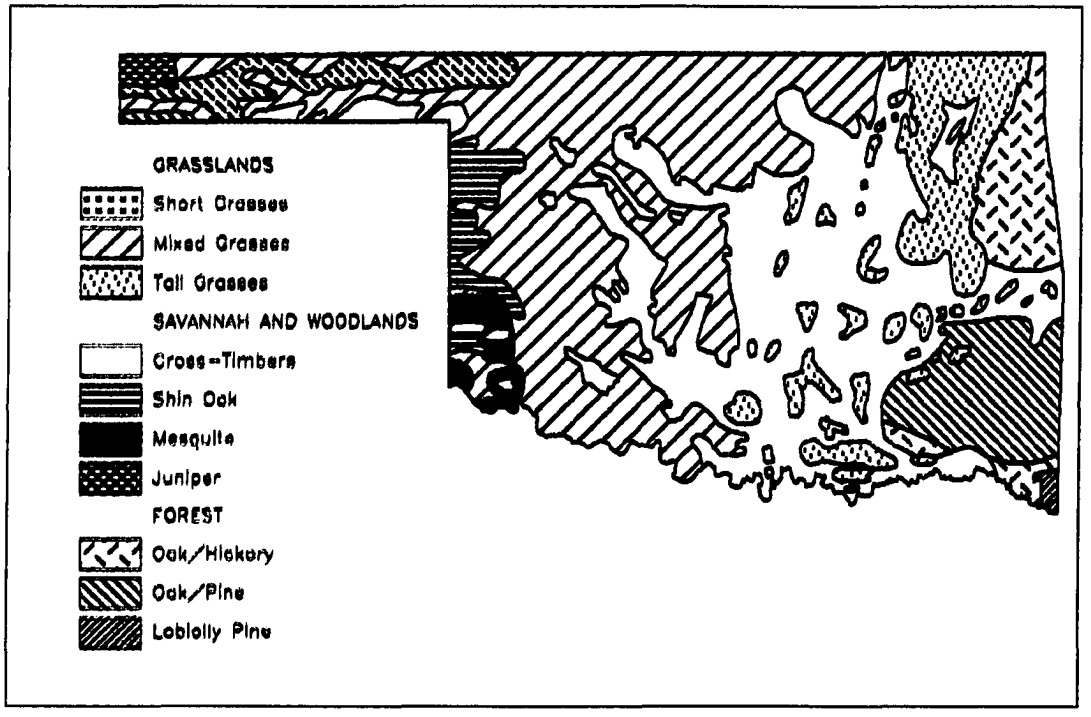


FIG 3.7. The various grasses and trees found across Oklahoma are displayed as described by Späth (1998).

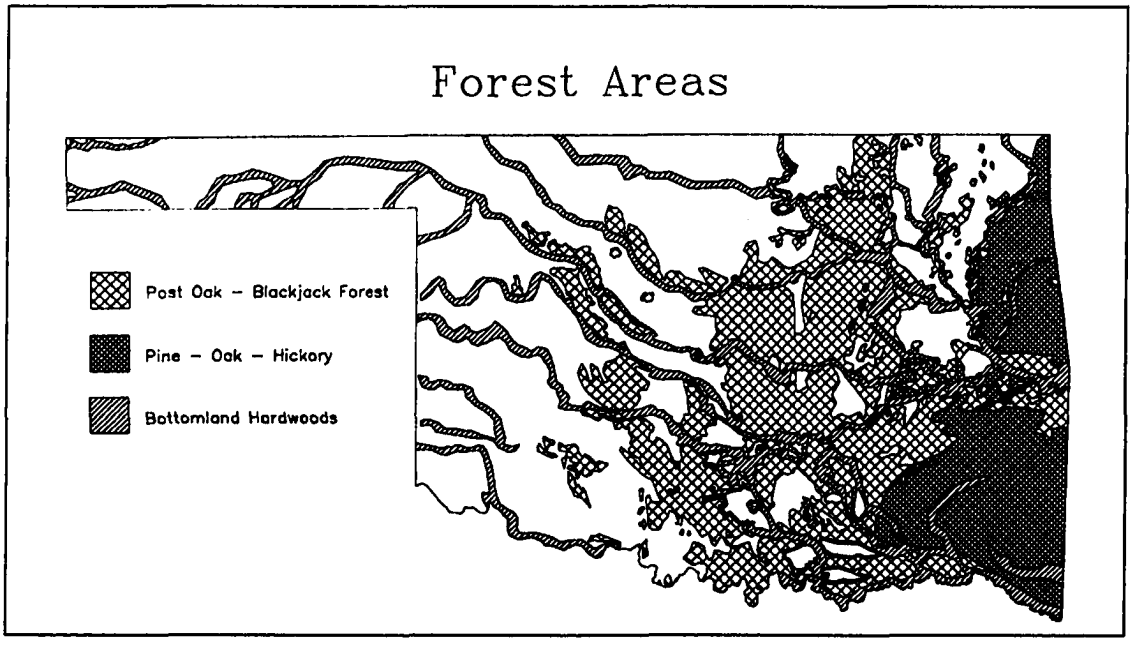


FIG 3.8. The forested areas found across Oklahoma are displayed as described by Wikle (1991).

Because Oklahoma has a diverse landscape, climate, and population, the diversity between the regions makes Oklahoma a unique study area to investigate the impact of climate/weather patterns on the demand for electricity by different groups of people. The four WFEC substations and co-located Mesonet sites chosen for this study are located in Regions 2, 4 and 10 of Figure 3.2.

3.4 Description of Data Sites

Four Mesonet sites and the nearest four WFEC substations were chosen to provide data for this study (Table 3.1; Fig. 3.9). The West Norman substation and the Norman Mesonet site are referred to as "Norman" in this study. Likewise, the Woodward substation and the Woodward Mesonet site serve Woodward. Altus AFB is served by the Altus Mesonet site and the Altus AFB substation. Finally, Dominance is an industrial location near the Broken Bow Mesonet site and the Dominance substation. These sites were chosen based on the type of electrical customer, geographical location, load demand (Table 3.2), the regional climatology, and the availability of data.

Table 3.1 The four Mesonet sites and substations used in this study, and their separation distances.

Mesonet Site	*Substation*	*Separation Distance*
Norman	West Norman	1.94 mi (3.12 km)
Woodward	Woodward	1.64 mi (2.64 km)
Altus	Altus AFB	5.23 mi (8.42 km)
Broken Bow	Dominance	8.52 mi (13.71 km)

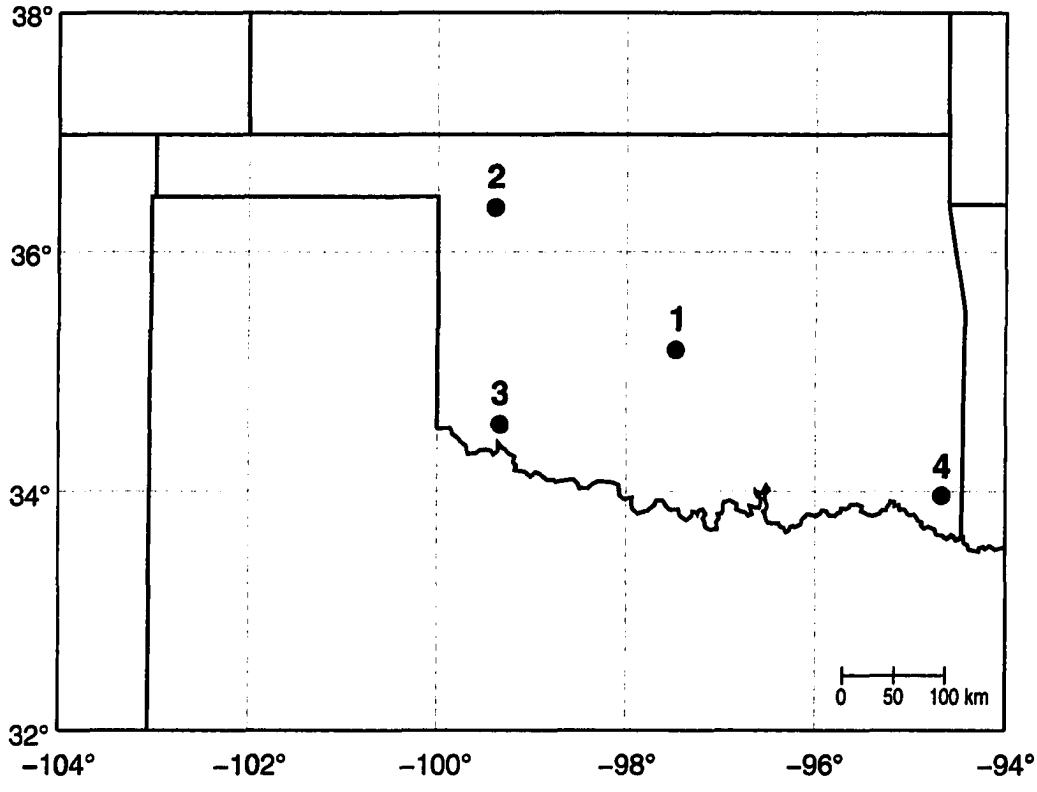


FIG. 3.9 Four data sites used in this study: 1 - Norman Mesonet site / West Norman substation; 2 - Woodward Mesonet site / Woodward substation; 3 - Altus Mesonet site / Altus AFB substation; 4 - Broken Bow Mesonet site / Dominance substation.

Table 3.2. The site name, type of customer, average annual load, and the annually peak load value for each substation (based on 2000 load data).

Substation	*Primary Customer Served*	*Average Annual Load (kWh)*	*Annual Peak Load Value (kWh)*
Norman	Urban/Residential	7392	18229
Woodward	Rural/Residential	5298	10392
Altus AFB	Air Force Base	8488	15044
Dominance	Industrial	3866	6482

3.4.1 Norman Site

Norman is located in Region 4 and is home to the University of Oklahoma. It is largely an urban community (Fig. 3.6) with thousands of single family dwellings, a large student population (28%), and several dozen apartment complexes. A sizeable group of Norman residents commute to work in Oklahoma City during the week.

Ninety-two percent of customers of the Oklahoma Electric Cooperative (OEC), which provides electricity to Norman and the surrounding communities, are residential while remaining customers are commercial. OEC has 90 miles (145 km) of distribution lines to service over 1800 customers (Table 3.3). Of the four study sites, Norman is the most densely populated per square mile. Distribution lines from the Norman substation extend approximately 1 mile (1.6 km) towards the west, 3 miles (4.8 km) towards the north, and 7 miles (11.3 km) toward the southeast.

Norman is located in central Oklahoma where the average annual temperature is ~60.1°F (15.6°C), and the mean annual precipitation amount of 37.6".

Table 3.3. The number of customers served each year (for the 3 year study period) and the miles of distribution line needed to reach these customers.

Substation	*Number of Customers in 1998*	*Number of Customers in 1999*	*Number of Customers in 2000*	*Miles of Distribution Line*	*Customers per mile of line*
Norman	1809	1874	1855	90	20.6
Woodward	1897	1929	1965	300	6.43
Altus AFB	1	1	1	-	-
Dominance	1	1	1	0.057	-

3.4.2 Woodward Site

Woodward is located east and slightly south of the Oklahoma Panhandle (Region 2). Woodward is primarily a residential community like Norman, but smaller in population (Fig. 3.6), larger in land area, and more rural in nature. This region is characterized by rural residential areas, farmland, and oil field businesses. Woodward is a major center for oil- and gas-field services, equipment repair, and well-drilling employment.

The Northwestern Electric Cooperative (NEC) services Woodward and the surrounding communities. The electrical load from the Woodward substation is consumed by customers that are 75% residential and 25% commercial. The distribution lines from this substation extend 30 miles (48.3 km) to the south and east. More than 300 miles (482.8 km) of overhead and underground distribution lines are required to serve the 1937 customers of this NEC substation, ~ 3 times more power lines than are required in Norman to serve only ~100 more customers.

The average annual temperature is ~60°F (15.6°C), the average rainfall is 23", and the average snowfall is 17". Relative to the remainder of Oklahoma, the climate in Woodward is cooler and drier.

3.4.3 Altus/Altus AFB Site

Altus is located in Region 4 of Oklahoma. It is home to Altus Air Force Base (AFB), which employs about 4800 people. Altus AFB is an Air Education and Training Command (AETC) base and is the only strategic Airlift and Air Refueling Training Center in the U.S. Air Force. During the day, Altus AFB consumes electricity similar to

that of a typical commercial/business entity (i.e., primarily between 7 AM – 4 PM). However, during the evening hours (i.e., 6 PM – 10 PM) the load profile for the AFB resembles that of a residential load profile reflecting its military base housing (i.e., 8000 family units and 400 dormitories). If necessary, this base can be converted into an operational facility for combat in a matter of days.

WFEC treats Altus AFB like one of its 19 distribution cooperatives (e.g., OEC and NEC). However, Altus AFB is different in that it has one substation which supplies power to only Altus AFB, whereas the other distribution co-ops distribute electricity from several substations to a multitude of customers in their respective service areas. Thus, Altus AFB is a single-customer substation of WFEC.

Altus has an average temperature of ~63.2°F (17.3°C) – the warmest average temperature of all four study sites. It receives 24" of rain and 8" of snow annually.

3.4.4 Broken Bow Site/Dominance

Broken Bow, which is home to the Dominance substation and the Broken Bow Mesonet site, is a rural community located in southeast Oklahoma (Region 10). Tyson Foods, Inc., the Weyerhaeuser Company, and Pan Pacific Products employ over 3000 people in Broken Bow. Farmers in this region combine cattle, poultry, tree farming and field crops into their operations.

Choctaw Electric Cooperative (CEC) distributes electricity in the Broken Bow area, some of which is delivered by the Dominance substation. The Dominance substation is provides electricity to only one (industrial) customer – Pan Pacific Products. This company manufactures wood products such as particleboard and components for the

door and moulding industry. Their electrical demand is different from most residential customers or commercial entities.

The average temperature is ~62°F (16.7°C), slightly warmer than the Norman and Woodward sites. Because the average annual precipitation is 55.1", Broken Bow is much more vegetated and humid than the other three sites.

Chapter 4: Analysis of Data

This scientific literature has established the fact that weather and climate are important in load forecasting. Hence, it is appropriate to understand the diverse weather patterns and climate across Oklahoma. Section 4.1 highlights the climatology of Oklahoma based on the last three decades of meteorological observations. In addition, the weather patterns that occurred during the three data-years used in this study (1998, 1999, and 2000) are reviewed. The focus is on the four regions of Oklahoma from which the Mesonet and load data were acquired. Significant differences in meteorological features between the 30-year climatology and those in the three-year data set are noted. Finally, the electric load data are presented for the same three-year period.

4.1 Climatology of Oklahoma

Climate is defined as a statistical accumulation of daily and/or seasonal weather events over a long period of time. Climatology is a study of the climate of a particular location, which results in "normals" of various weather parameters. A "normal" daily temperature, for example, is computed by averaging the mean daily temperature observed on each day of the year during a 30-year period[1]. Because the averaging technique tends to smooth extremes and other smaller-scale fluctuations, climatological temperature patterns applied to a demand for electricity, have a minimal impact. Yet, climatological *trends* offer load forecasters a reasonable starting point from which they can assess the load demand for each service area during monthly, seasonal, and annual periods.

[1] Oklahoma Climatological Survey (OCS): http://k12.ocs.ou.edu/teachers/lessons/accuracyofclimate.html

However, for STLFs, details that were smoothed during the climatological averaging process represent important pieces of information that better define a peak load. Thus, climatological data are often used for long-term load forecasting and as an assessment tool for short-term forecasts.

Oklahoma's climate is said to be "fundamentally transitional" (Späth et al. 1998). This terminology means weather in Oklahoma is a transition zone between the arid western United States and the humid eastern United States (in terms of precipitation). The Oklahoma climate also is transitional in terms of temperature between the cold of the north during winter and the heat of the south in the summer.

To illustrate diversity of the Oklahoma climate, Figure 4.1 presents the nine climate divisions established by the predecessor to the Climate Prediction Center during the mid-1950s. The nine divisions are: 1 – Panhandle, 2 – North Central, 3 – Northeast, 4 – West Central, 5 – Central, 6 – East Central, 7 – Southwest, 8 – South Central, and 9 – Southeast. The boundaries of these climate divisions were defined using a balance of drainage basins, crop diversity, temperature and precipitation averages, heating and cooling degree days, and drought indexes (Guttman and Quayle 1996). Although the divisions do not always enclose areas of climatological homogeneity, the size of each division also was based upon an array of criteria including crop-growing belts, electric power grids, water resources and numerical grids. From a larger perspective, states like New Mexico, Arizona, and California have only seven climate divisions to represent complex interactions between climate and the native vegetation.

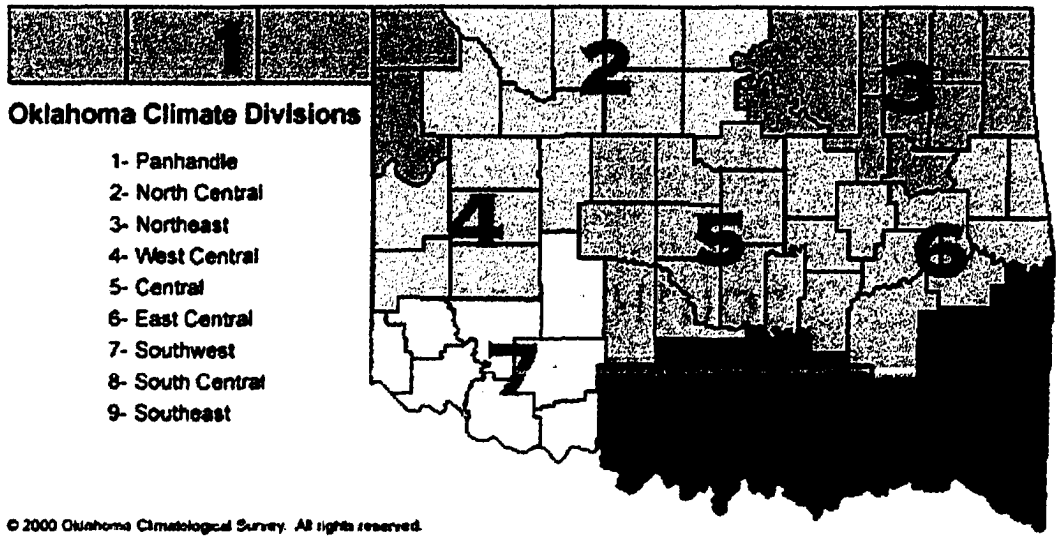


FIG. 4.1 The nine climate divisions of Oklahoma.

Temperature is the weather parameter that has the most significant impact on electric load demand. During any given month across Oklahoma, a wide range of temperatures can be observed; these features are even evident in a 30-year (i.e., 1971 – 2000) climatology. During the winter months, Oklahoma is warmest in its southeast quadrant and gradually becomes colder toward the northwest and into the panhandle (Fig. 4.2). From the standpoint of electricity demand, a load forecaster can deduce from this climatology that the heating demand may be large in northwest Oklahoma.

During the spring transitional months, not only do temperatures warm, but the highest observed temperatures shift into southern Oklahoma (Fig. 4.3). The frequent intrusion of cooler air is most evident in the panhandle section of the state. Springtime temperatures are comfortable compared to other seasons. Thus, from an energy demand

point-of-view, the climatological patterns suggest that electricity consumption is minimal during this transitional season.

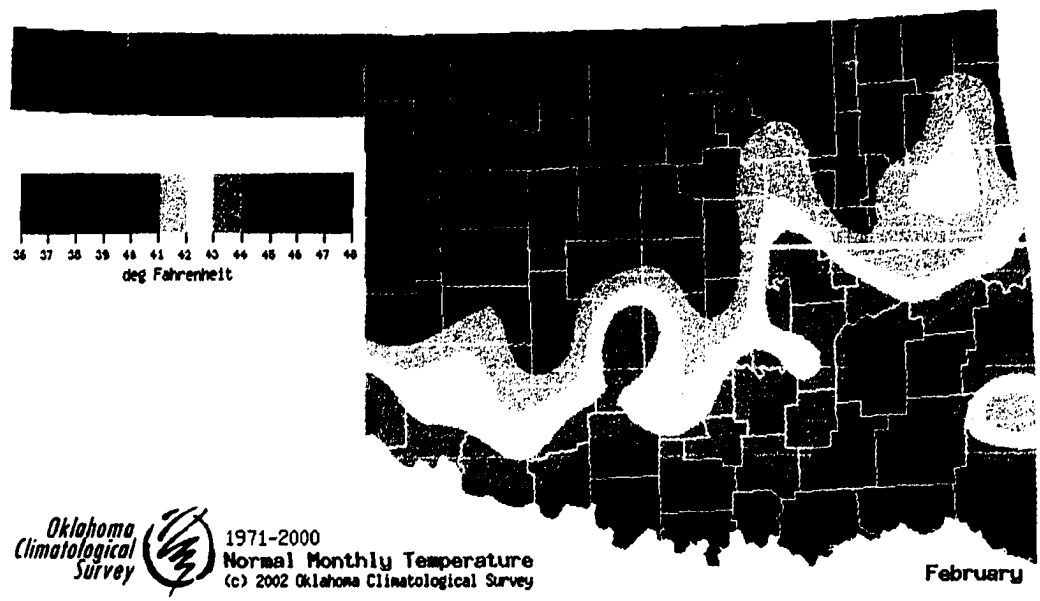


FIG. 4.2 The average temperature for February based upon data from 1971-2000.

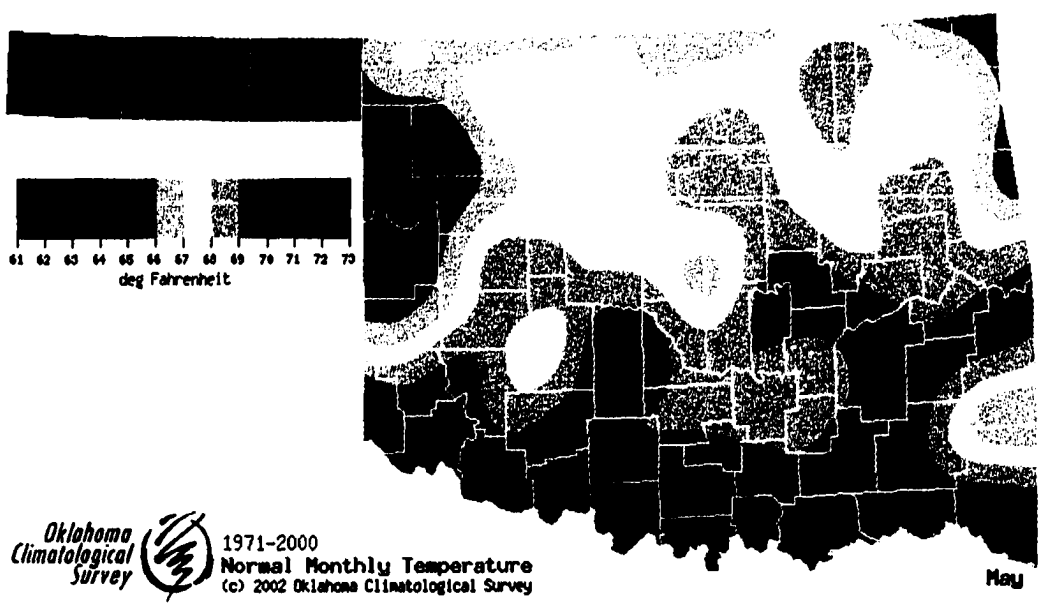


FIG. 4.3 As in Fig. 4.2, except for May.

Warm temperatures dominate a majority of the state during the summer months (Fig. 4.4). The large gradient of temperature observed during the winter is not present during summer. This climatological pattern represents "big picture" guidance to load forecasters, especially to those who deal with generation decisions.

As autumn approaches, the temperature gradient across Oklahoma begins to resemble the winter pattern; warmer temperatures are confined to the southeast and cooler ones to the northwest (Fig. 4.5). The beginning of fall, much like spring, brings comfortable temperatures to most Oklahomans. Further into the fall season, climatological temperatures decrease quickly (~0.5°F) with each passing day. Thus, heating becomes a necessity for comfort. Load forecasters must be aware that abrupt changes in temperature during the transitional seasons create load demands that are more difficult to forecast.

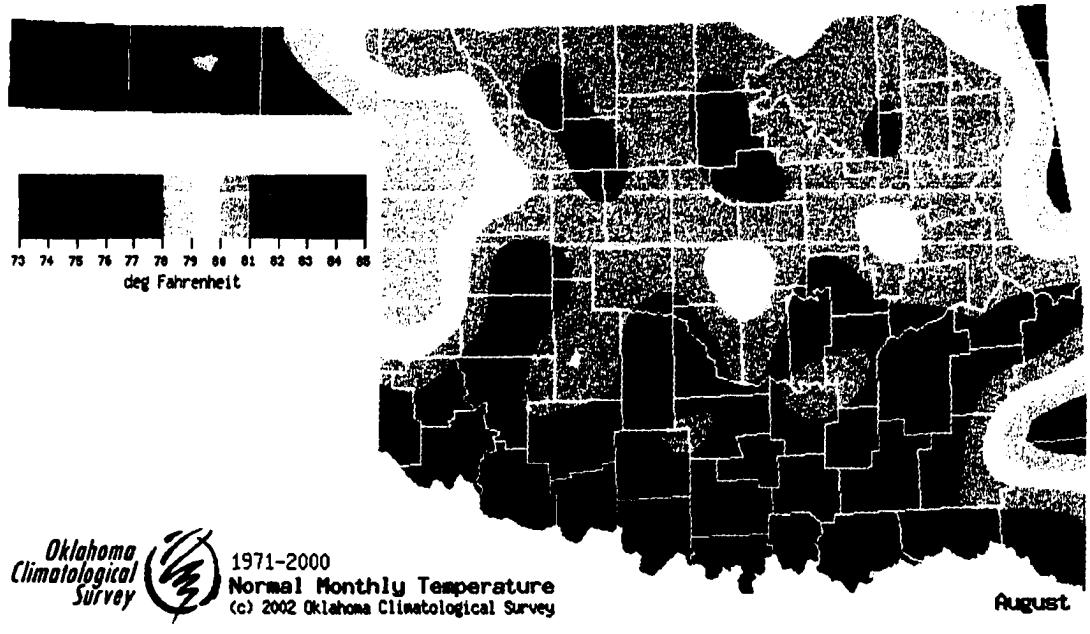


FIG. 4.4 As in Fig. 4.2, except for August.

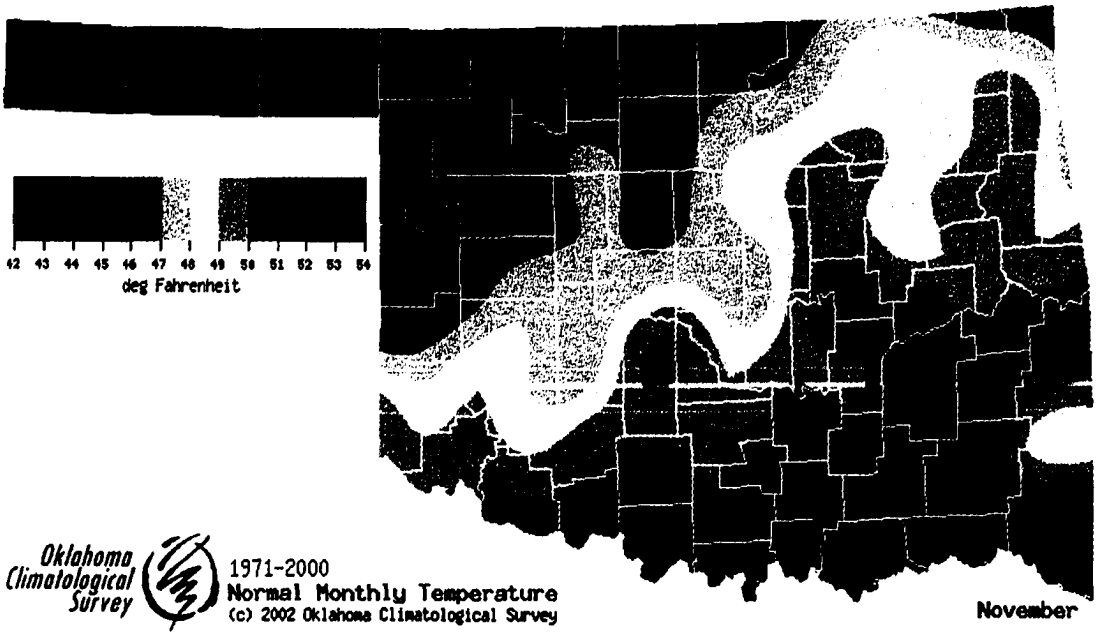


FIG. 4.5 As in Fig 4.2, except for November.

4.2 Oklahoma Weather During the Study Period

Because the actual load data and concurrent, co-located meteorological data represent pillars that underpin this work, it seemed appropriate to document the weather patterns that occurred during the study period[2]. Overall, the weather patterns were similar from year to year. Important details from each of the four study regions described in Chapter 3 are documented below. Table 4.1 displays the climatic and the observed, annually-averaged temperatures. Significant events that occurred during the study period (e.g., heat waves, cold spells, and major precipitation events) will be addressed.

[2] The Oklahoma Annual Climate Summary, produced by the Oklahoma Climatological Survey, provided a concise summary of the weather across Oklahoma during 1998, 1999 and 2000.

Table 4.1 Climatological normal temperatures (F) and annually-averaged temperatures observed at four sites during the study period.

Location	*Climate Division*	*Normal Temperature*	*Average 1998 Temperature*	*Average 1999 Temperature*	*Average 2000 Temperature*
Norman	5 (Central)	60.1	62.96	61.81	60.64
Woodward	2 (N. Central)	56.5	59.86	59.75	58.77
Altus	7 (Southwest)	61.8	63.75	62.78	61.87
Broken Bow	9 (Southeast)	60.7	63.39	61.85	61.26

4.2.1 Oklahoma Weather – 1998

The statewide-averaged temperature for 1998 was 62.3°F (16.8°C), 2.0°F (1.1°C) warmer than normal. During the year, the 4th warmest autumn and the 8th warmest summer occurred based upon temperature records that date to 1892. Each month except March and April recorded warmer than normal temperatures. At Altus Air Force Base, a daily maximum of 113°F (45°C) occurred on 1 June. By mid-June, most of western and southern Oklahoma labored under oppressively hot and dry conditions. These conditions spread statewide by month's end. From July through September, the sweltering temperatures caused 24 heat-related deaths in Oklahoma.

Record-breaking droughts, concentrated in the west central, southwest and south central climate divisions, plagued Oklahoma from April through September. Yet, 1998 ended with statewide-averaged precipitation that was above normal due to large amounts of precipitation during other seasons in other parts of the state. The state was wet and warm during January. Annual precipitation in the climate divisions which included

Norman, Broken Bow and Woodward were above normal for the year while Altus reported below normal precipitation.

4.2.2 Oklahoma Weather - 1999

Mild weather during the cold months and a wet spring created a warmer and wetter year than normal. The above normal temperatures at the end of 1998 persisted through the winter of 1999 and created an anomaly that was 3.8°F above normal for the winter. Yet, springtime temperatures registered 1.3°F below normal. In addition, the most significant snow storm of the year occurred in March. Much like the features which occurred during 1998, the summer was hot and dry across the entire state. A heat wave in late July caused 8 deaths across Oklahoma. Each day in August, the temperatures reached triple digits somewhere in the state. Some relief came in September when average temperatures for the month were lower than normal across Oklahoma. Yet, temperatures for the fall season ended warmer than normal by 1.3°F. December followed the same warm trend. The four study sites ended the year 1°F above normal.

Greater than normal precipitation occurred in January while February was drier than normal. The precipitation in April was much greater than normal across the state except for a small area in extreme south central and southeast Oklahoma. In fact, January through June recorded the 7th greatest 6-month precipitation in the history of Oklahoma. The infamous 3 May 1999 tornado also occurred, which was the most expensive tornado (i.e., an estimated $1 billion in damage) recorded to date in our nation's history[3]. While summer precipitation was near normal, dry conditions continued into September for western and southeastern Oklahoma. Locally heavy rainfall occurred in other areas to

[3] http://www.srh.noaa.gov/oun/storms/19990503/may3faqs.html

create near-normal rainfall for the month. As much as 10" of snow occurred near Woodward in early December. Near mid-December, snow and sleet occurred in many areas followed by heavy rains and relatively warm temperatures. The weather of 1999 provided an interesting contrast. Norman, Woodward, and Altus were a few inches above normal for their annual precipitation while Broken Bow was 10" below normal.

4.2.3 Oklahoma Weather - 2000

On average, the year 2000 was a warmer and wetter year than normal. However, the year also was filled with erratic temperature swings, droughts, flash floods, and winter storms. Several climatological records were set as well. The year began with temperatures that were 4.5°F (2.5°C) above normal (continuing the persistent warm pattern from December). In fact, the winter of 1999-2000 was the 9th warmest on record. The average springtime temperature was only 0.7°F (0.4°C) above normal, even with a record-setting 110°F (43.3°C) in Altus. Daytime highs exceeded 100°F (37.8°C) during May for one-third of the month. The summer season recorded slightly above normal temperatures that resulted in two heat-related deaths. The fall season displayed capricious temperature swings and more record-setting conditions. During the first week in October, triple digit temperatures were recorded about the state, with a 106°F (41.1°C) record set in southeast Oklahoma. By late-week in October, the first widespread freeze occurred across most of the state. In fact, high temperatures plummeted more than 50°F in just three days. This event was followed by the coldest November-December period on record (i.e., 7.6°F [4.2°C] below normal). During December alone, three destructive winter storms occurred. The 25-27 December ice storm destroyed 38,030 distribution

poles, 1086 transmission structures, 109 miles of transmission line, and 2113 miles of distribution lines to businesses and homes. Additionally, twenty-two deaths occurred between Christmas Day and New Year's Day as a result of the horrid weather conditions.

Though a severe drought occurred during 2000, the statewide-averaged annual precipitation was 2.55" above normal. The beginning of the year was rather dry. March became the first month to record above normal precipitation. After two more months of below average precipitation, the skies opened in June and produced the 7th wettest June on record (i.e., 3.24" greater than normal). August began a recording-setting warm season drought. These dry conditions persisted through September and proved costly and damaging (e.g., poor crop yields and destructive wildfires) for Oklahoma. The remaining three months of the year brought above normal precipitation (i.e., 3.41", 1.16", and 0.32" above normal, respectively). In addition, three of the four study sites reported above normal annual precipitation. Altus was the only site that recorded precipitation below the climatological normal.

4.3 Electric Load Demand

An electric utility serves many customers, all of whom have different electric load demands which peak at different times during the day. For example, rural utilities serve residential, commercial, industrial (or large commercial), and street lighting customers (Rastogi and Roulet 1994). Each customer class has a unique base load profile (Fig. 4.6). Residential customers consume electricity using household appliances, lights, television, and heat/cooling units. The peak load for residential customers occurs during the evening hours with a secondary peak during the early morning hours. Commercial

customers include schools, hospitals, restaurants, hotels, shopping malls, and office buildings. The commercial load profile typically peaks late in the morning and again during the evening. On the other hand, the industrial load demand parallels the work shifts of a company with the load profile remaining constant during each shift. The street lighting load, or municipal load, is constant from dusk to dawn and zero at other times.

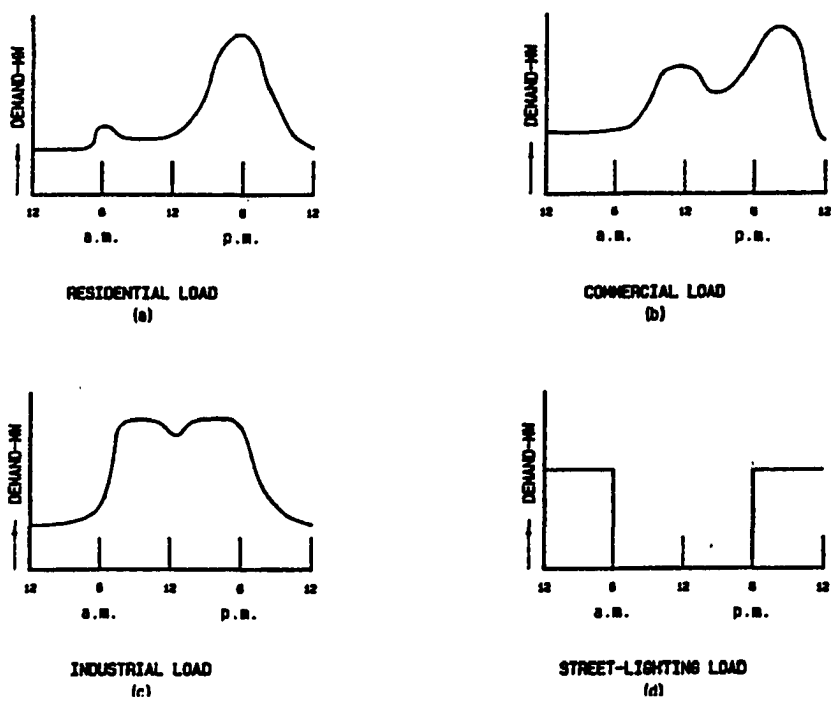


FIG. 4.6 Typical load profiles for the customers of rural electric cooperatives. (Figure from Rastogi and Roulet 1994)

Typical load profiles are commonly known as the "base demand" or "base load" for each customer class. The base demand reflects a long-term average behavior of that customer and is considered to be constant with time or change very slowly.

Superimposed on this base load is a variable component of the total load demand which is highly attributable to fluctuations in daily weather conditions. Changes in weather – seasonally, daily or hourly – cause changes in the use of electricity, thus altering the shape and magnitude of the base load profile. Other changes in load are caused by a change in human habits from weekdays to weekends and vice versa, holidays, and special events.

The four study sites were chosen to reflect a specific customer base. Norman and Woodward represent the residential urban and residential rural customers, respectively. Altus AFB characterizes the commercial customer while Dominance represents the pure industrial customer. Because changing weather impacts each customer class in a different fashion, the four unique customer classes will open the door to investigations that have never been attempted.

4.3.1 Average Load Profiles – Norman

Norman is primarily a residential community. The typical load profile illustrated in Figure 4.6 for residential customers reaches its peak around 7 PM, with a secondary peak occurring at 8 AM as residential customers prepare for their day. The annually-averaged hourly loads for Norman during the study period (Fig. 4.7)[4] resembled the profile of a residential hourly load (Fig. 4.6). However, when the load profile is analyzed on smaller time scales (e.g., seasonally, monthly, or daily), peaks in the load profile become better defined. Additionally, when the averaged load profiles combine weekdays and weekends/holidays into one graph, an otherwise significant morning peak observed

[4] Only 1998 and 1999 are shown in Figure 4.7 because the load consumption doubled at the Norman substation in 2000. This increased load consumption occurred because the capacity of the transformer was increased. The 2000 load profile had the same shape but its magnitude was twice that of 1998 and 1999.

during the winter, fall and spring seasons is masked (discussed with Fig. 4.12). Finally, and most importantly, changes in the weather can alter the magnitude of the peak values in the load profile and the shape of the profile. These characteristics were investigated.

Seasonally-averaged load profiles for Norman during the study period are shown in Figures 4.8 - 4.11. (The scale on the y-axis of the plots in the following sections was varied to retain the maximum detail possible in each load profile.) The shape of the winter load profile (Fig. 4.8) is similar to that of base load profile for a residential community. The time of the winter peak load during all three years occurred at ~8 PM, while the peak load *values* during the winter differed from year to year. During the study period, the peak load ranged from 3200 kWh – 3500 kWh. While weekend and

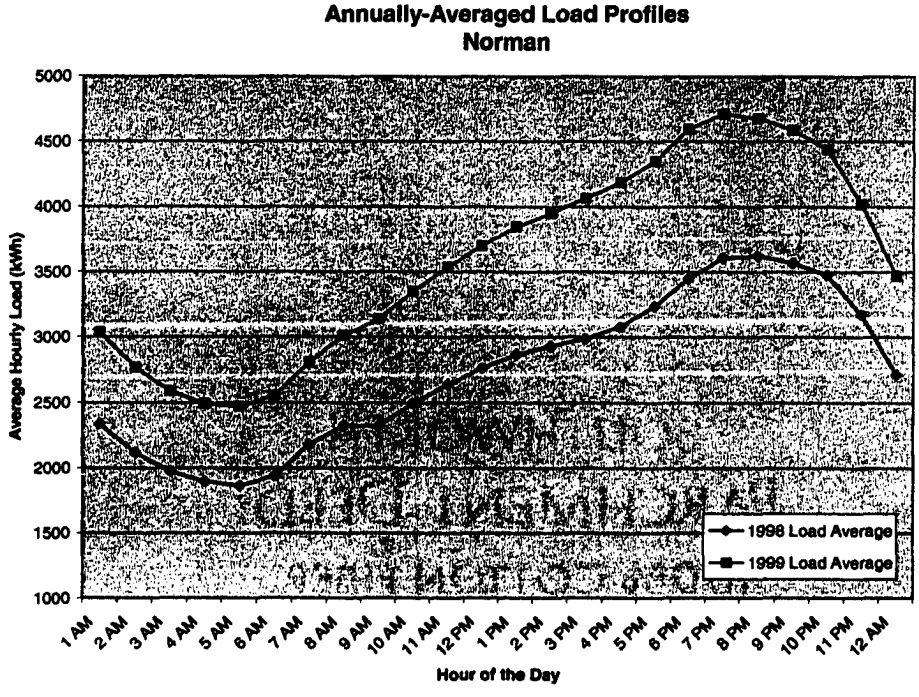


FIG. 4.7 Annually-averaged load profiles from Norman during the study period.

weekday loads were combined in all profiles shown thus far, the secondary load peak at ~8 AM was better defined in the winter load profile than in the annually-averaged load profiles. Cold mornings associated with the winter season often caused residents to increase their use of electricity. The average winter load was in excess of 2500 kWh for 18 hours in 1998 and 10 hours in 1999. (This difference is attributable to a warmer than normal winter in 1999.)

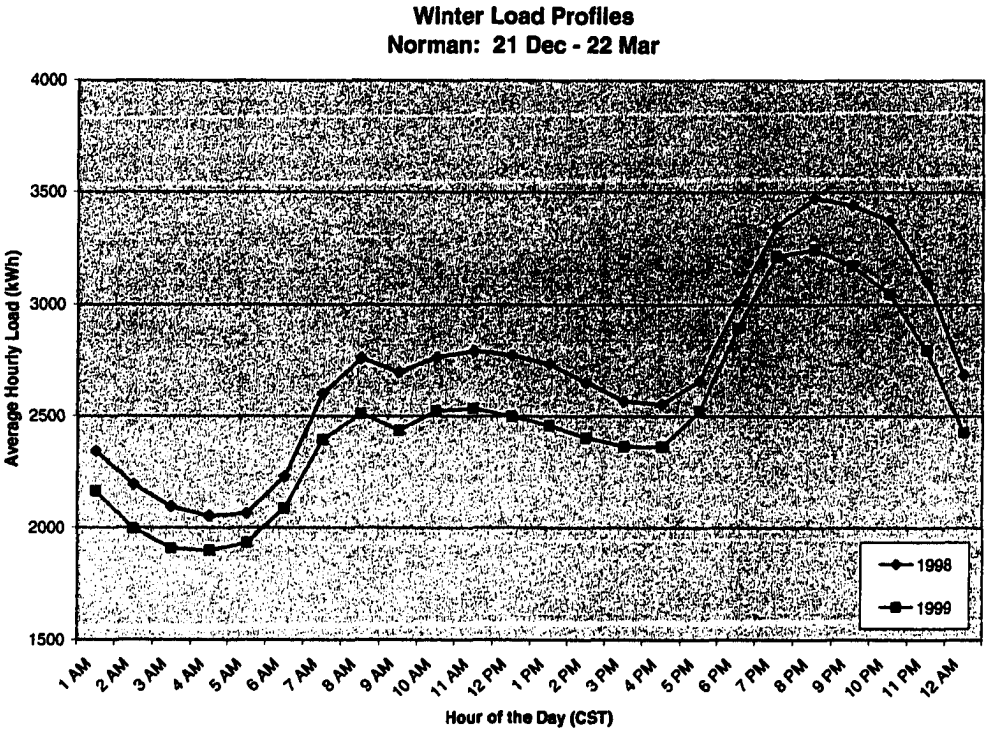


FIG. 4.8 Average hourly electric load profiles from Norman during the winters of the study period.

The springtime load profiles (Fig. 4.9) were different from the annual profile in that they had a broad evening peak that spanned ~3 hours (6 PM – 8 PM). During the study period, the springtime peak values ranged from 3000 to 3400 kWh. The annually-

averaged and spring load profiles were similar in that the secondary peak near 8 AM was ill-defined compared to that which was readily apparent in the winter profiles. Because the morning temperatures during transition seasons are more comfortable for humans (than, for example, during the winter season) less electricity is required. The average load consumption during the spring exceeded 2500 kWh for 11-12 hours throughout the afternoon and evening.

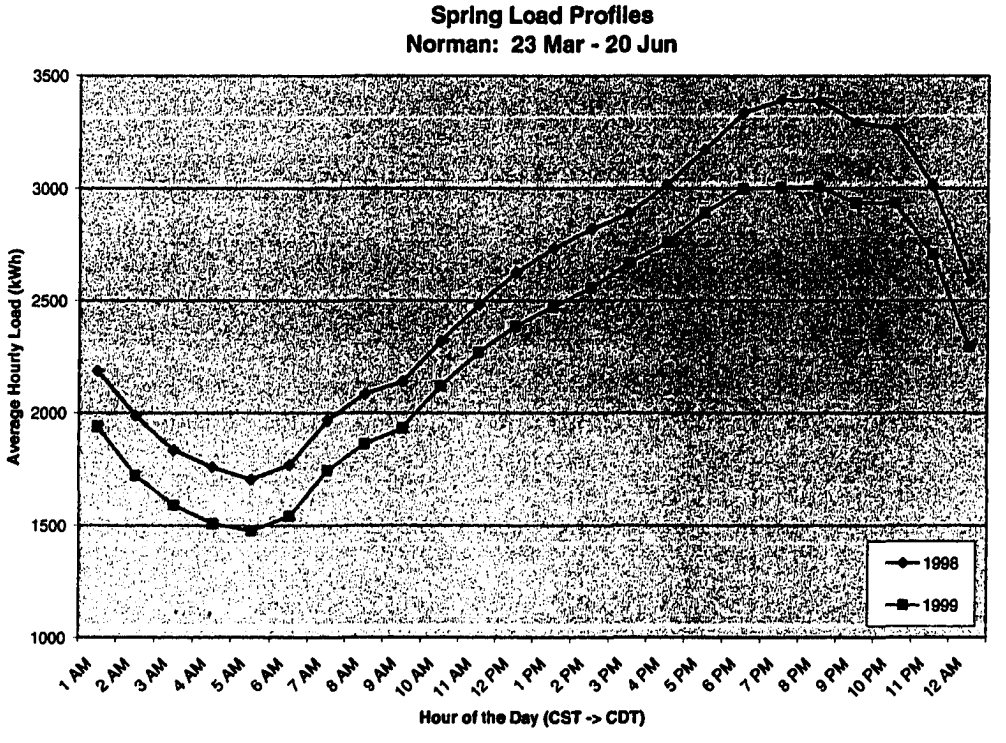


FIG. 4.9 As in Fig. 4.8, except for spring.

The summer profile (Fig. 4.10) rose steadily between 6 AM and 6 PM. Summertime peak values range from 4800 kWh to 6100 kWh, and only during the morning hours in 1998 did average load dip below 2500 kWh. The secondary peak

identified in the annually-averaged winter, was damped in the spring load profiles and was absent from the summer profile Furthermore,. The reason for the absence of this feature is explored in Chapter 5.

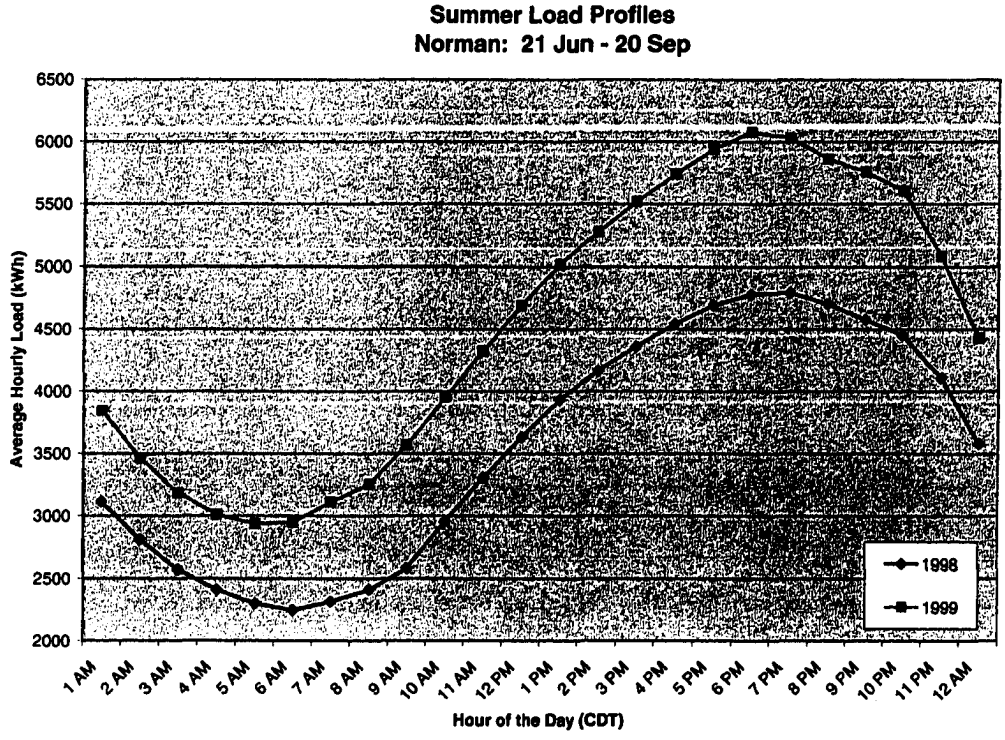


FIG. 4.10 As in Fig. 4.8, except for summer.

The load profiles for the fall season (Fig. 4.11) regained part of the shape detected in the winter and annual profiles from the study period. Peak load values ranged from 3400 to 5000 kWh, slightly higher than the peak values observed during the spring and winter season. The average load in 1998 exceeded the 2500 kWh threshold for 11 hours while the 1999 load was above 2500 kWh all day on an average day. Thus, the shape of the Norman load profiles for the winter, spring, and fall seasons mirrored the idealized

base load curve, in spite of the fact that the curves shifted (up or down) due to varying weather conditions. Only during the summer seasons did the load curves have a different morning trend and shift up or down due to changes in the weather.

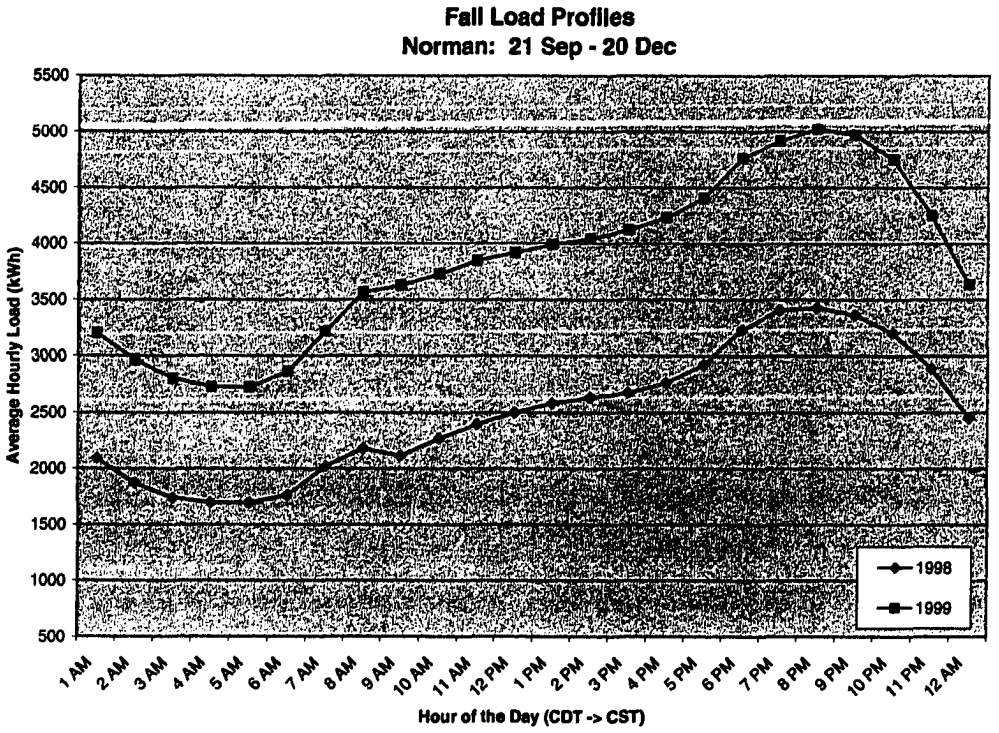


FIG. 4.11 As in Fig. 4.8, except for fall.

4.3.2 Weekdays, Weekends, and Holidays - Norman

The shape of the load profile varied dramatically when weekdays and weekends are analyzed separately (Fig. 4.12) because customers consume electricity differently on weekends than they do on weekdays. During the week, the load curve sharply increases between 5 AM and 8 AM, which suggested customers were preparing for work or school and using more electricity in the process. After 8 AM, the use of electricity decreased

slightly before increasing to the evening peak load that occurred between 7-8 PM. On weekends, nearly the same amount of electricity was consumed as on weekdays. However, the morning load curve on an average weekend reflected a reduced consumption of 15% between 5-8 AM and a 5-9% increase in consumption between 10 AM and 5 PM. Consumers appeared to remain at home on the weekends, which increased their electricity consumption during the midday hours. By evening (7-8 PM), the peak loads for weekday and weekend use of electricity were approximately the same. As midnight approached, the load demand on weekdays and weekends decreased dramatically.

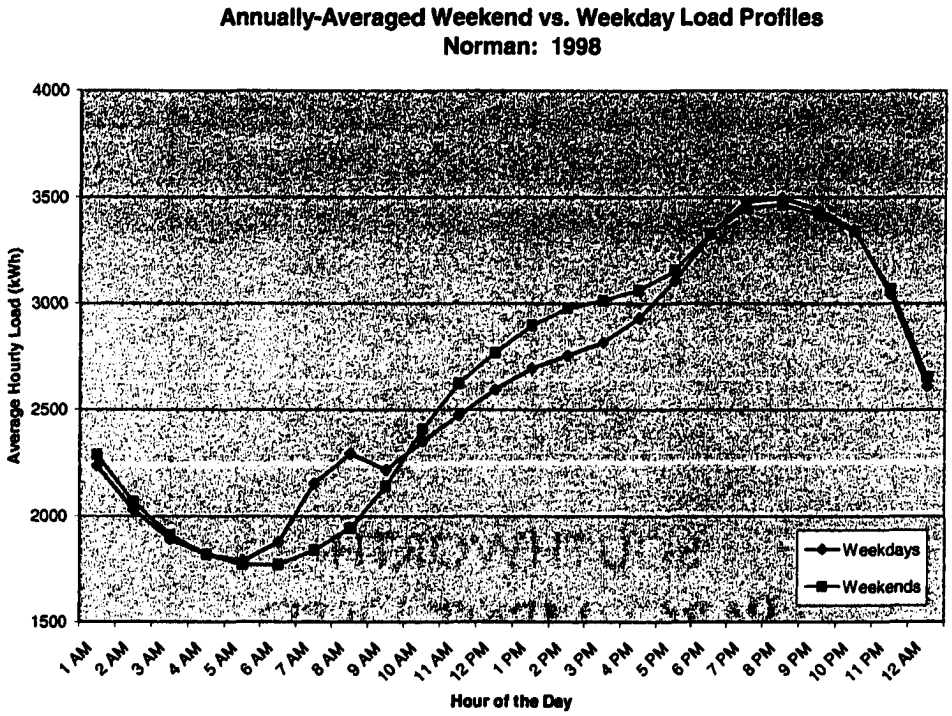


FIG. 4.12 Annually-averaged load profiles from Norman in 1998 that illustrate the differences in the weekday versus weekend load.

Holidays, especially those that occur on a weekday, significantly impacted the idealized load. Load profiles for the four Mondays in February 1998 (16 February 1998 was President's Day, a federally observed holiday) are illustrated in Figure 4.13. While three Mondays exhibited similar patterns and peak times, the holiday Monday resembled a weekend profile more than it did a weekday profile. The difference resulted from the fact that residential customers seemed not to follow their routine when Monday was a holiday. Some customers observed the holiday, while others possibly did not (e.g., The University of Oklahoma). The weekday before a holiday period begins also may experience similar changes to the load profile when individuals choose to extend their holiday weekend.

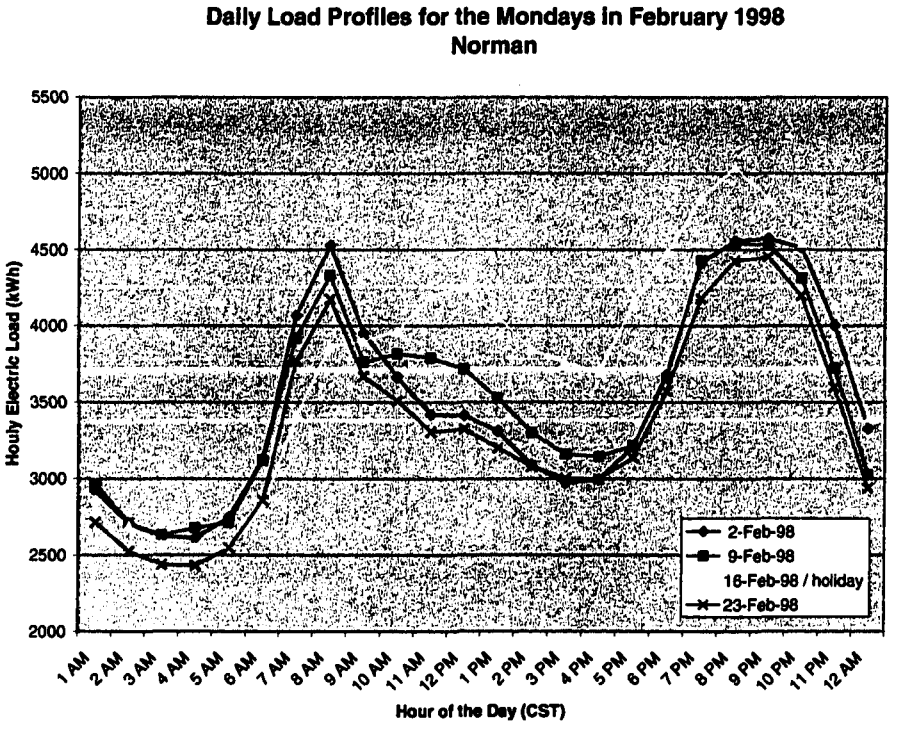


FIG. 4.13 Daily load profiles for the four Mondays in February 1998 for Norman. The holiday Monday (President's Day) produced an electrical load with a different pattern than was observed on non-holiday Mondays.

The daily load profiles (Fig. 4.13) from the winter of 1998 illustrate the different holiday pattern and exemplify the drastic morning peak (on non-holidays) which, clearly, was masked in the annually- and seasonally-averaged load profiles. Furthermore, the differences between the three non-holiday Mondays illustrate the impact of weather. For example, the average wind chill temperatures (Appendix D) during the 7 AM – 9 AM peak, were 34°F (1.1°C), 37.2°F (2.9°C), and 38.1°F (3.4°C) on 2 February, 9 February, and 23 February, respectively. The coldest morning temperature (2 February) also had paralleled the greatest morning peak load. As the temperature increased during the day, the peak load decreased. It is noteworthy that the average wind chill temperature between 7 AM – 9 AM for 16 February, the holiday Monday, was 37.3°F (2.9°C) – the same average temperature that occurred on 9 February. However, the load during the holiday period was significantly different from loads on the non-holiday Mondays due to different societal habits on holidays.

4.3.3 Average Load Profiles – Woodward

Woodward is a residential community, similar to Norman, but more rural in nature. Based upon the base load profile for a residential community (Fig. 4.6) and those created for the Norman site (Fig. 4.7), classical features included a rapid increase in electrical use towards an inflection around 8 AM followed by a steady increase toward a large evening peak. The annually-averaged load profiles for Woodward (Fig. 4.14)

during the study period[5] were not much different. However, analyses on smaller time scales revealed critical details in the profiles.

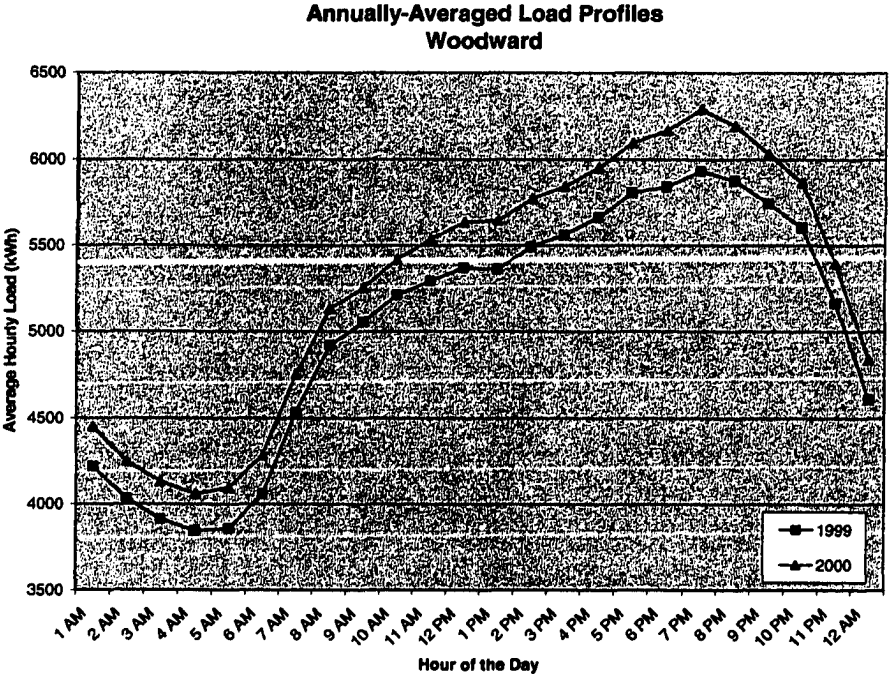


FIG. 4.14 Annually-averaged load profiles from Woodward during the study period.

Seasonally-averaged profiles for the Woodward site are shown in Figures 4.15 – 4.18. The shape of the winter profiles (Fig. 4.15) clearly define a very significant 8 AM peak as most residents prepared to leave home for the day, and an 8 PM peak which indicated that most residents had returned home and settled into their evening routine. These features were not clearly defined in the annually-averaged plots. The post-sunrise section of the load curve increased by 1300 kWh in just 3 hours to reach the 8 AM peak of ~5600 kWh. During the day (i.e., 8 AM – 4 PM), the electric load steadily decreased until 4 PM.

[5] The study period for Woodward is May 1998 – December 1998, 1999, and 2000. Because electric load data were corrupt for the first four months of 1998, neither the annual load profile nor the winter and spring load profiles will be shown.

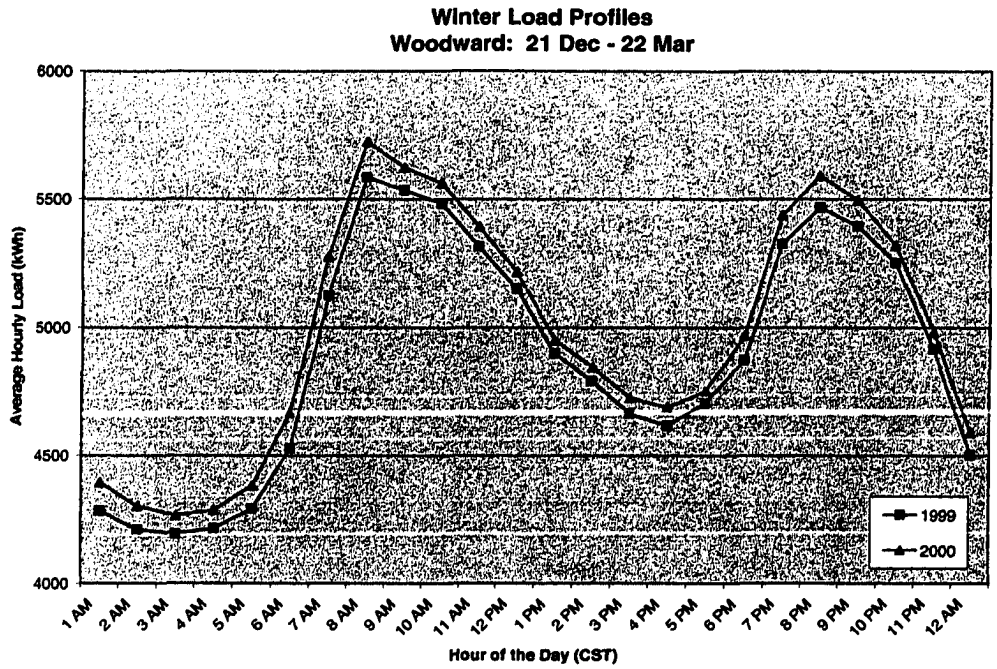


FIG. 4.15 Average hourly electric load profiles from Woodward during the winters of the study period.

Another sharp increase in electrical usage occurred between 4 PM and 8 PM. Even so, the 8 PM peak fell shy of the morning peak by ~100 kWh. Yet, considerably more electricity was needed to meet the morning power demand relative to the evening peak (e.g., ~6 hours above 5000 kWh prior to midday versus 4+ hours above 5000 kWh during the evening). The *shape* of the winter load profile to seemed result from predictable routines by customers rather than from human responses to extreme winter weather conditions. Peak values of the profile, however, were very sensitive to weather conditions.

The spring load profiles (Fig. 4.16) increased steadily from 5 AM to a broad peak at 5 PM. The load consumption increased rapidly during the period from 5 AM to 8 AM, but the increased use of electricity occurred at a slower rate after mid-morning. The

average early-evening peak was ~5300 kWh, approximately 300 kWh lower than the winter peaks. Yet, usage exceeded 5000 kWh for ~7 hours. This decrease in peak load demand from winter to spring can be attributed to warmer, more comfortable temperatures that occurred during the springtime months relative to the winter months.

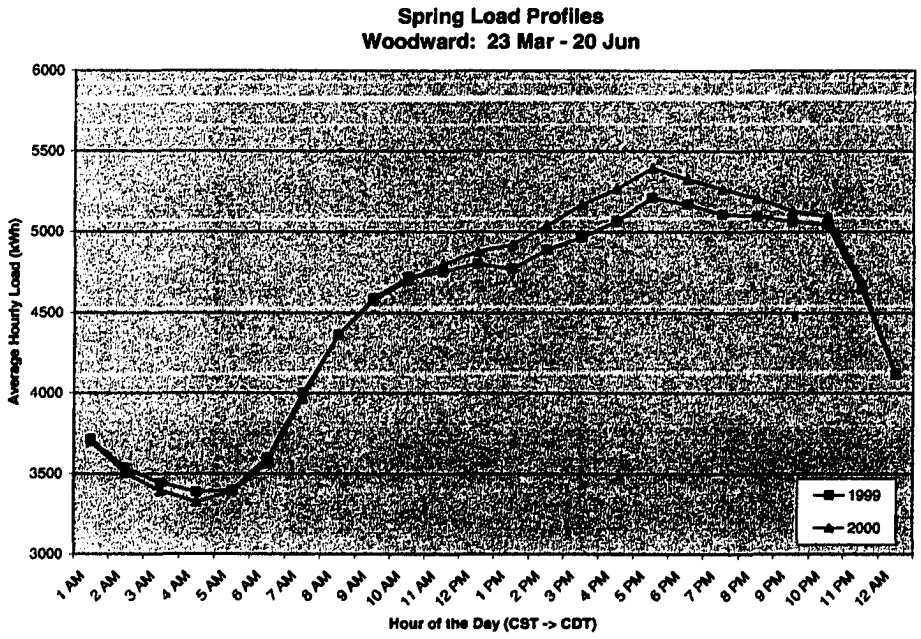


FIG. 4.16 As in Fig. 4.15, except for spring.

The summer load profiles (Fig. 4.17) climbed steadily from 5 AM (e.g., the profile minimum) to the 5 PM peak. All but 6 hours of the 24-hour day had an average load demand that exceeded 5000 kWh. The large amplitude of this summer pattern was masked in the annually-averaged load profile as was the extended period of high demand (i.e., 18 summer hours when the load exceeded 5000 kWh versus only 9 hours when the entire year was considered). The load profile during the summer months is especially important to the generation of electricity because this season is when utilities companies often operate at maximum capacity and sometimes must obtain electricity through other

sources to meet the customer demand. The 5 PM peak averaged 8500 kWh, almost 3000 kWh higher than the springtime peak. This greater demand was a direct result of the higher, often unpleasant temperatures experienced during the summer months.

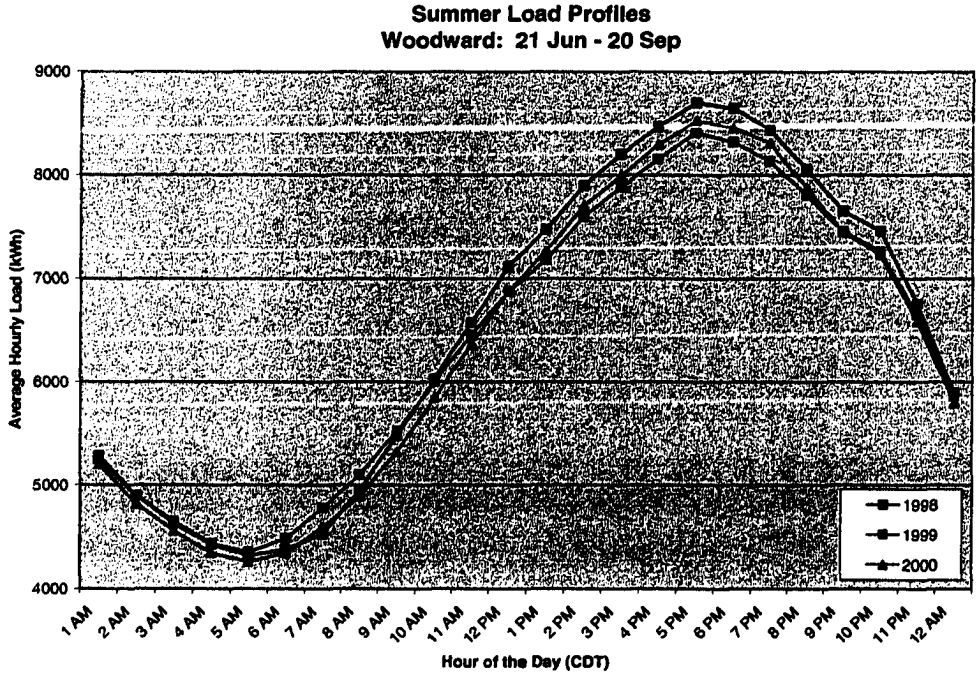


FIG. 4.17 As in Fig. 4.15, except for summer.

The bimodal structure (morning and afternoon peak) of the typical residential load demand (Fig. 4.6) returned in the fall load profiles (Fig. 4.18) as temperatures returned to more comfortable levels. A sharp incline in the load profile occurred between 5 AM and 8 AM. The use of electricity after 8 AM held steady until mid-afternoon, after which a gentle increase resumed to the 7 PM peak of ~5400 kWh. The average evening peak was about 500 kWh more than the average morning peak. Average loads were in excess of 5000 kWh for ~4-6 hours during the evening hours of 1998-1999, but exceeded this threshold for 16 hours during the dry fall of 2000.

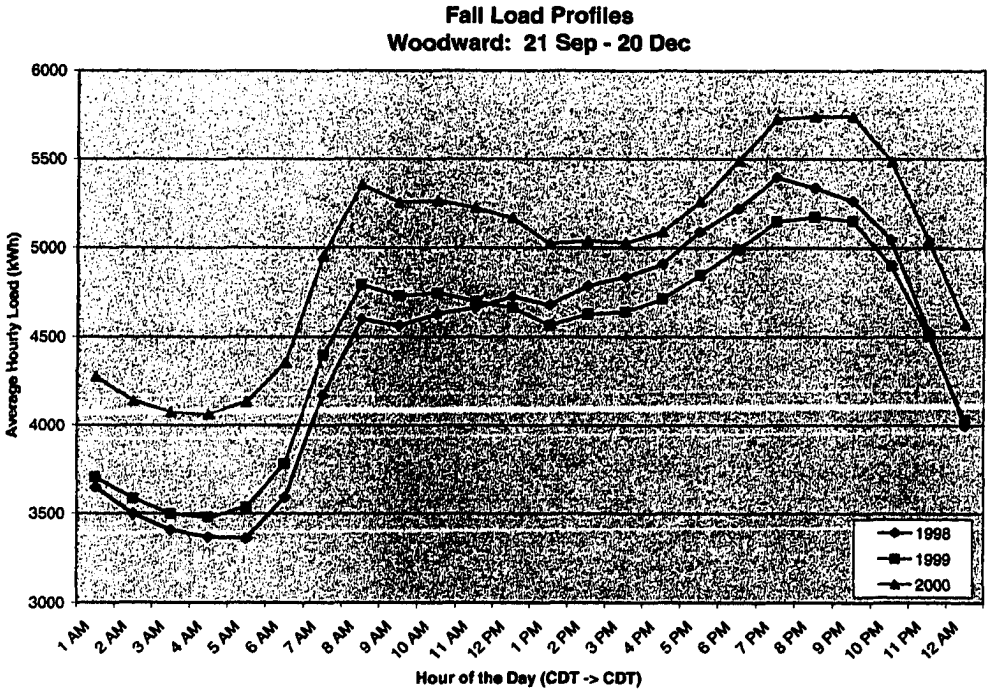


FIG. 4.18 As in Fig. 4.15, except for fall.

4.3.4 Weekdays, Weekends, and Holidays - Woodward

As illustrated in the load profiles from the Norman site (Fig. 4.12), a significant difference existed between load profiles from weekdays versus those from weekends. The rural community served by the Woodward substation clearly used electricity differently on weekdays than on weekends, especially during the morning hours (Fig. 4.19). As much as a 15% difference in electrical consumption existed between 5 AM and 6 PM. However, in contrast to electrical consumption in the suburban Norman community (Fig. 4.12), data from the Woodward substation revealed electricity consumption to be higher on the weekdays than on the weekends (between the hours of 5 AM – 6 PM). Apparently the residents of this rural community are involved in more activities outside the home on weekends than are those in the suburban community of

Norman. One consistency between these two residential sites was the weekday energy surge that occurred between 6 AM – 8 AM. Yet, as observed in Norman, the morning section of the weekend load profile from Woodward had a more gradual slope. The average load in Woodward exceeded 5000 kWh for 14-16 hours and dropped below 4000 kWh for only 2-3 hours.

Holidays, when observed, also led to a change in the shape of the load profile and the magnitude of the peak. From a residential perspective, an observed holiday implies that (employed) residents have a day off from work. While one cannot assume that the general population will stay home the entire (holiday) day, it is likely that they would be at home during more hours on a holiday than on a workday. To illustrate this point, the electrical load was analyzed for the Mondays in September of 1999 (of which 6 September 1999 was Labor Day; Fig. 4.20). More electricity was consumed on the holiday Monday than on any other Monday during the month. Not only was the magnitude of the peak 50% greater on the holiday Monday relative to the other Mondays during September, but the shape of the holiday profile was different as well. The community served by the Woodward substation began their active day later in the morning on the holiday Monday than they did on workdays on the other Mondays.

On these Mondays, the maximum afternoon temperatures were: 92°F (33.3°C) on 6 September, 75°F (23.9°C) on 13 September, 56°F (13.3°C) on 20 September, and 53°F (11.7°C) on 27 September. Thus, a 39°F (21.7°C) difference existed among the daily maximum temperatures which, no doubt, drastically altered the use of electricity. In a four week span, the temperature went from the uncomfortable 90s one week, to the

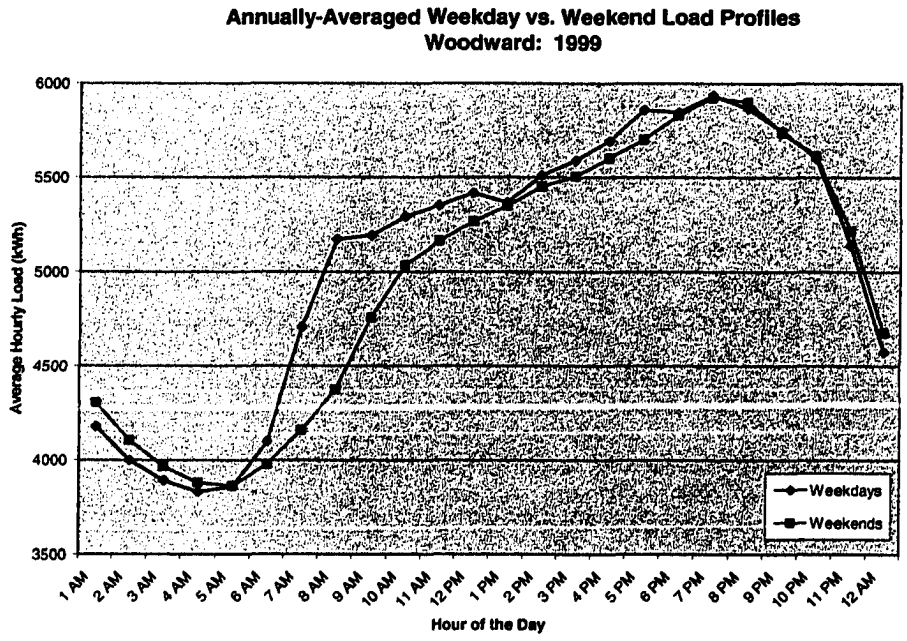


FIG. 4.19 Annually-averaged load profiles from Woodward in 1999 that illustrate the differences in the weekday versus weekend load.

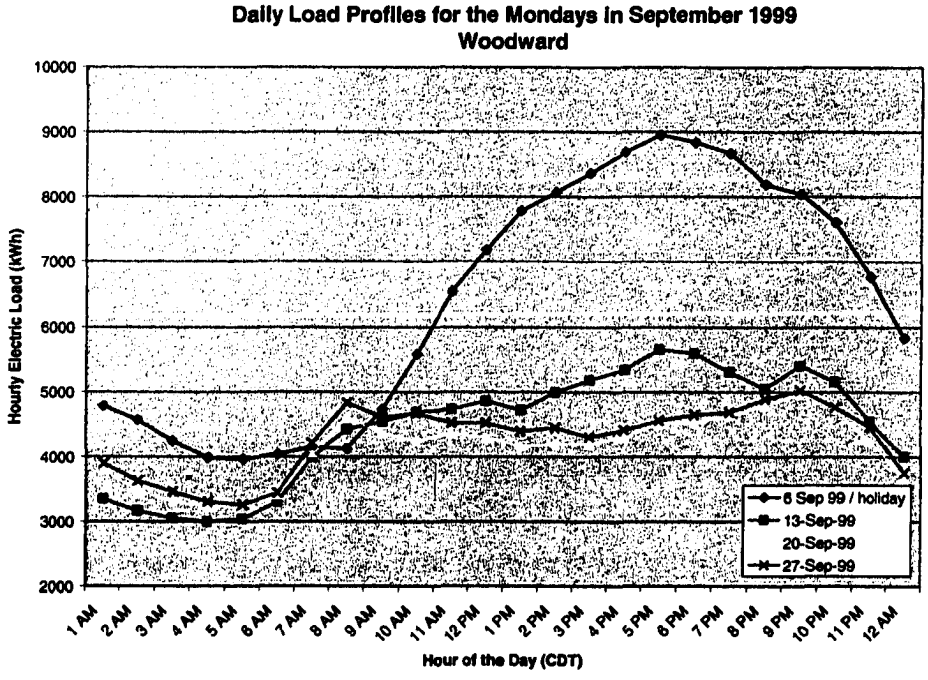


FIG. 4.20 Daily load profiles for the four Mondays in September 1999 for Woodward. The holiday Monday (Labor Day) produced an electrical load with a much different pattern than was observed on the non-holiday Mondays.

comfortable 70s the next week, to the less-comfortable 50s for the next two weeks. On the warm holiday Monday, the hourly load exceeded 5000 kWh for 15 hours of the day versus the two coolest Mondays when the load never exceeded 5000 kWh. Thus, some difference between the four Monday plots, at least in the magnitude of the peak load, can be attributed to large temperature changes between the beginning and middle of the month. It is this sort of temperature swing that compounds the difficulty of load forecasting.

4.3.5 Average Load Profiles – Altus AFB

Altus AFB is an education and training facility for the United States Air Force. The average load profiles during the study period[6] (Fig. 4.21) are similar to that of the commercial profile in Fig. 4.6. As expected, the load consumption was elevated like that during the typical 8-hour workday of a commercial entity. However, instead of having an evening primary peak (Fig. 4.6), the load profile at Altus AFB had a sharp decrease between 4 PM to 6 PM. A less rapid decline occurred between 6 PM – 10 PM, primarily because the (military) base housing mimicked the evening load peak associated with the residential load profile (Fig. 4.6). After 10 PM, the profile sharply decreased towards its overnight minimum load observed at dawn. While annually-averaged load reflects a general profile of electricity consumption, load profiles on smaller time scales reveal important details that allow the impact of weather to be analyzed. Because Altus AFB typically operates as a (Monday-Friday, 7 AM – 4 PM) commercial entity, weekdays and weekends/holidays were analyzed separately.

[6] The base had a reduction in military personnel between 1998 and 1999, which is reflected by the large reduction in the use of electricity between the 1998 and 1999 load profiles. Thus 1999 and 2000 will be considered the "study period" for Altus AFB.

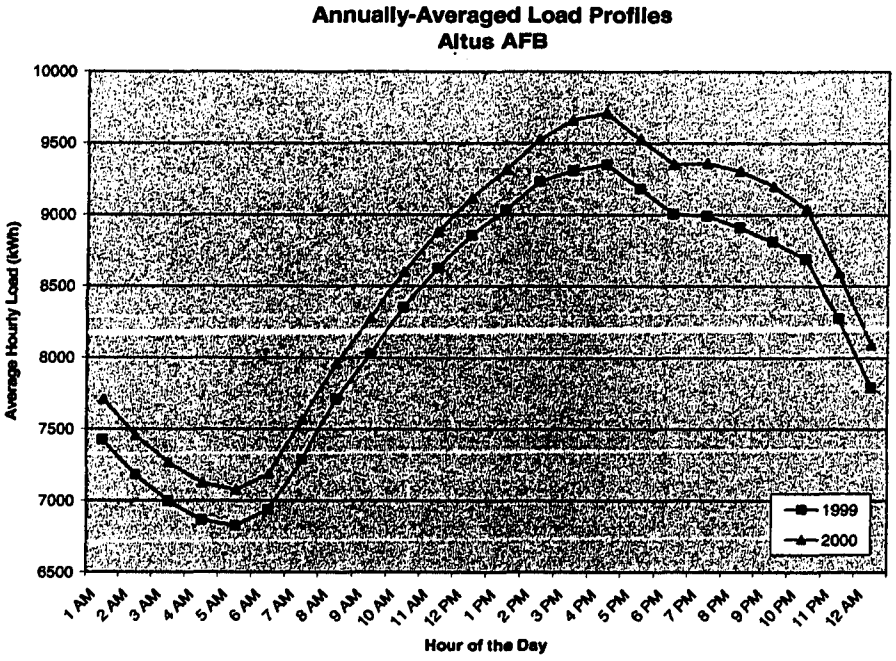


FIG. 4.21 Annually-averaged load profiles from Altus AFB during the study period.

The seasonally-averaged profiles for Altus AFB are shown in Figures 4.22 – 4.25. The winter profiles for 1999 and 2000 (Fig. 4.22) resembled the commercial load profile (Fig. 4.6), in that mid-afternoon and evening peaks existed in the use of electrical. The afternoon load curve paralleled that of a typical commercial workday, which was followed by a rapid decrease between 4 PM – 6 PM. The load profile during the workday ranged between 7000 – 7500 kWh during the study period. Outside of the 7 AM – 4 PM workday, the load profile resembled that of a winter, *residential* load profile (Figs. 4.8 and 4.15). After 6 PM, Altus AFB becomes a residential community as base operations cease for the day and base residents become the primary consumer of electricity until early the next morning. At 5 AM, a sharp increase occurred in the load; this feature was attributed to the fact that occupants of base housing began their day which included

housing military personnel arrive at work. The 8 PM winter peak coincided with the 8 PM peak load observed at substations which serve residential communities.

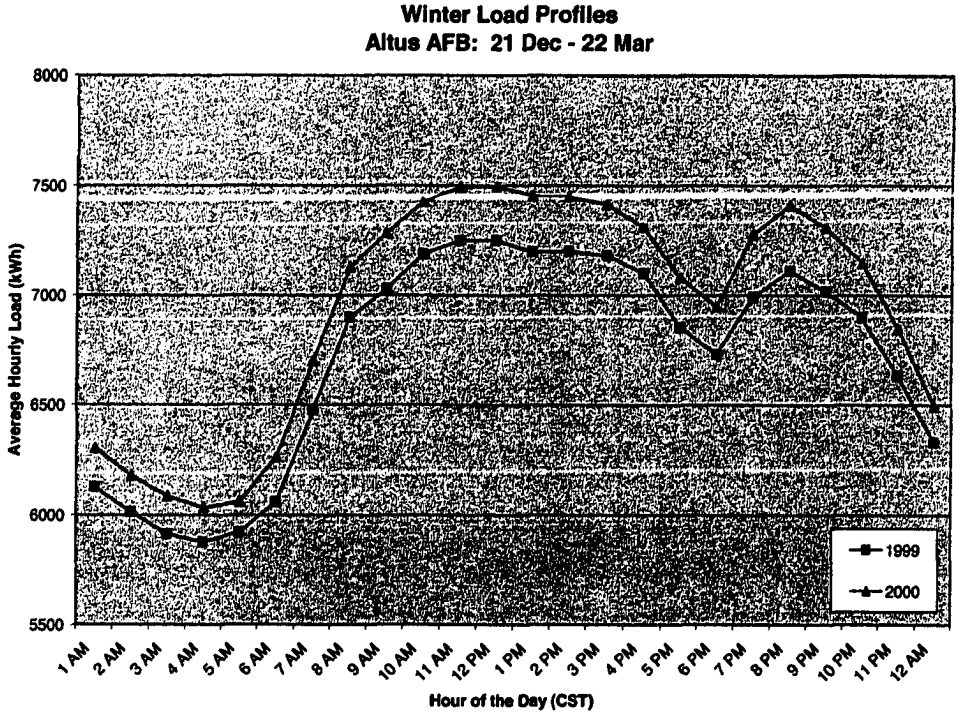


FIG. 4.22 Average hourly electric load profiles from Altus AFB during the winters of the study period.

As temperatures warmed and the season transitioned to spring, the shape of the load profile changed dramatically. Beginning at dawn, the springtime load profiles (Fig. 4.23) rose steadily until a 4 PM peak load was reached, which coincided with the end of the workday. The peak loads ranged from 9100 kWh – 9700 kWh during the study period. As occurred in the winter profiles, a late afternoon decline in electrical consumption was evident. However, between 6 PM and 10 PM a slow decline persisted in the profiles (whereas the winter profiles climbed to a secondary evening peak). After 10 PM, the use of electricity plummeted to its overnight minimum load at dawn. While

the predawn minimum was ~700 kWh greater in spring than occurred during winter, the afternoon peak was ~2000 kWh greater than its winter counterpart.

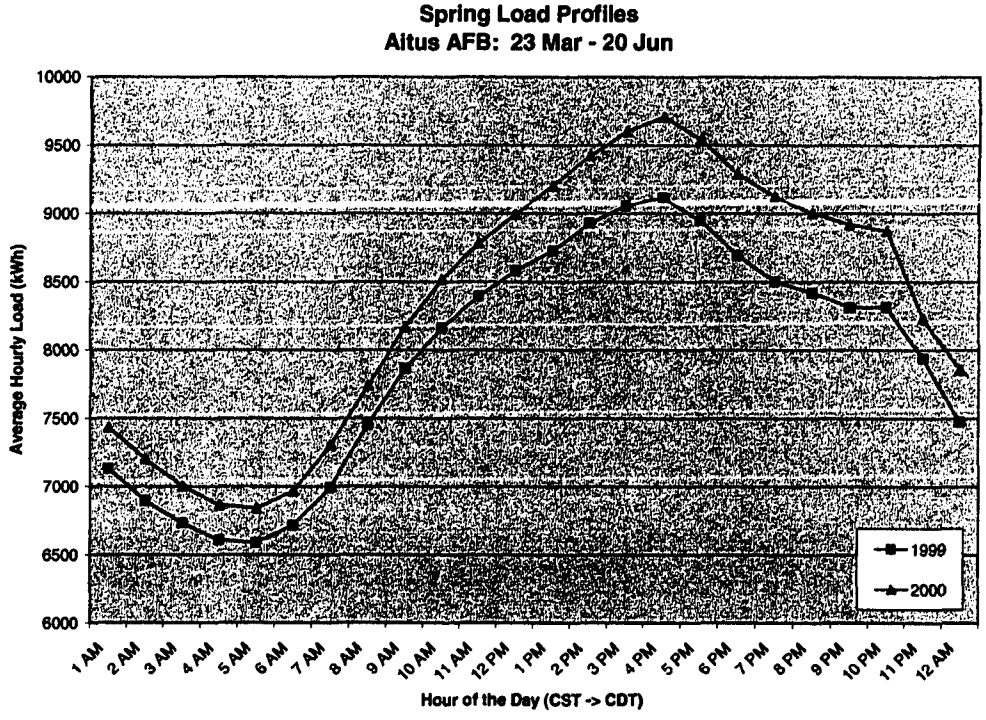


FIG. 4.23 As in Fig. 4.22, except for spring.

The *shape* of the summer load profiles (Fig. 4.24) was very similar to the spring profiles. The consumption differences between the two seasons occurred in the magnitude of the afternoon peaks and the slope of the profile between 6 AM and 4 PM. The afternoon peak load ranged between 12500 kWh and 13100 kWh, 3400 kWh greater than the springtime use of electricity. During summer months, ~20% more electricity was consumed between 6 AM and 4 PM than was used during the spring months (Fig. 4.24 versus Fig. 4.23). The predawn *minimum* of ~8500 kWh exceeded *all peak* loads that occurred during the winter.

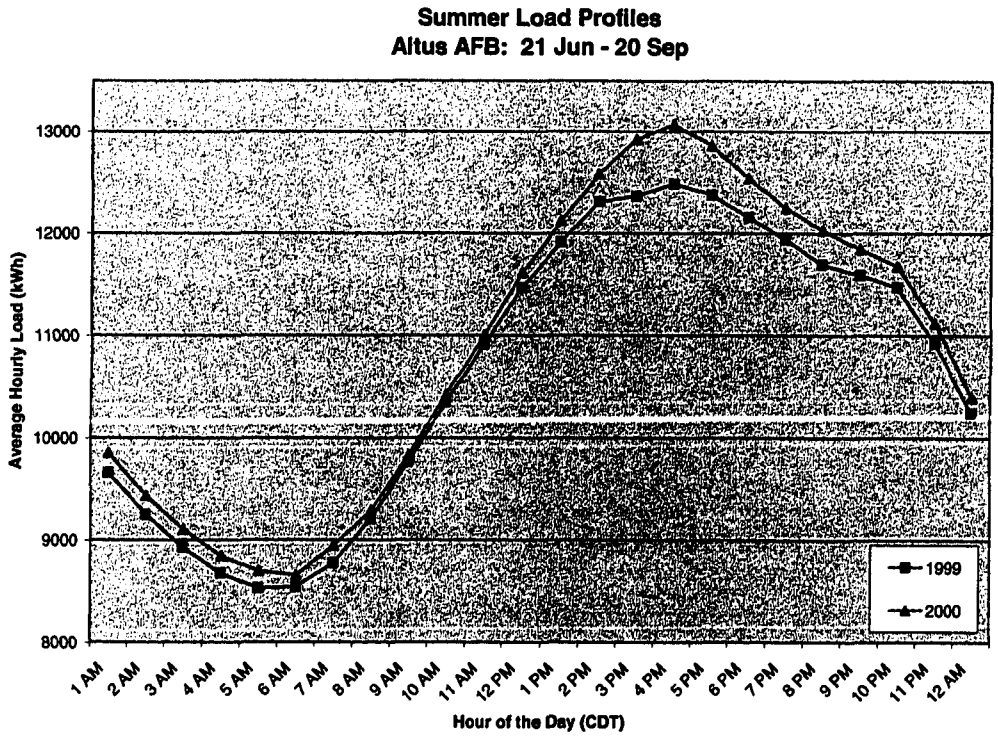


FIG. 4.24 As in Fig. 4.22, except for summer.

As the temperatures began to decrease with the arrival of a new season, the fall load for Altus AFB (Fig. 4.25) re-acquired a double peak structure similar to that from a winter profile. Even so, load consumption sharply increased from 6 AM – 8 AM, after which the profile had a gentle upward slope toward the 4 PM peak. The use of electricity decreased slightly between 4 PM – 6 PM after the typical workday had ended. An increase in load occurred from 6 PM – 9 PM, primarily due to residents on the AFB using more electricity from their homes; subsequently, a sharp decline occurred in the demand for electricity between 9 PM – 5 AM.

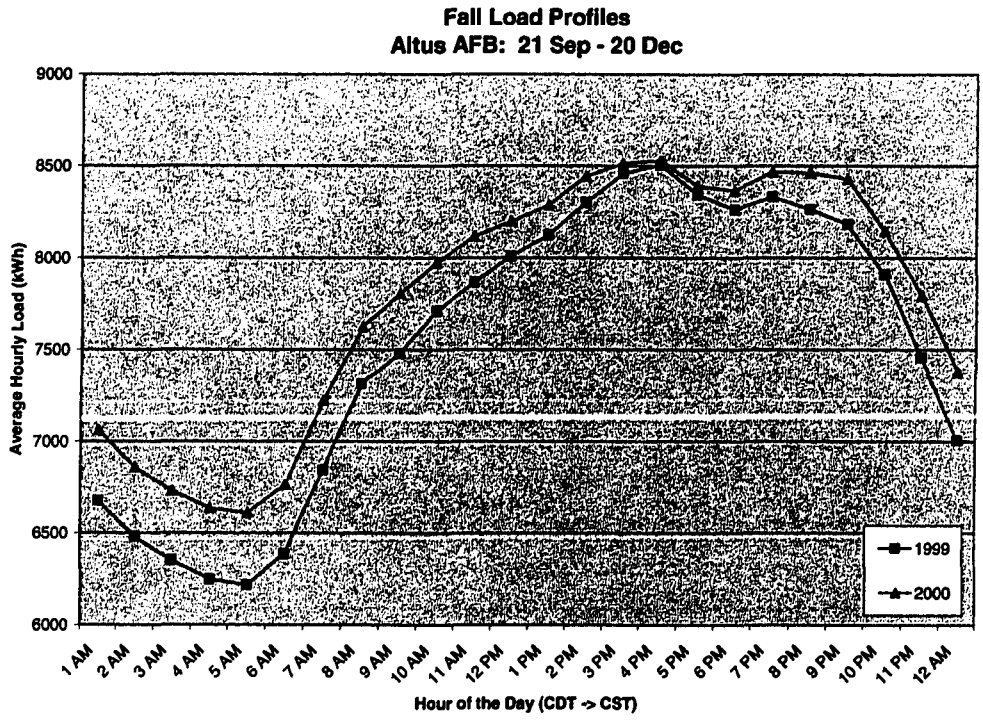


FIG. 4.25 As in Fig. 4.22, except for fall.

4.3.6 Weekdays, Weekends and Holidays – Altus AFB

For a commercial entity (e.g., operating on an 8-hour, Monday – Friday schedule), it was expected that electricity use on weekends and holidays would be lower relative to the weekdays. Thus, weekday and weekend/holiday load profiles for Altus AFB were created separately to gain a more detailed understanding of its load consumption.

Between the hours of 6 AM and 10 PM, electricity consumption was 30% higher during the week than on weekends (Fig. 4.26). It is important to remember that Altus AFB has a strong residential component (i.e., electricity consumption by the residents of base housing) which contributed to the use of electricity outside of the 8-hour workday. For example, while a typical business may cease operations at 5 PM (thereby consuming

less electricity), Altus AFB experienced an increase in the use of electricity from its residential component between 5 PM and 10 PM. A typical commercial entity would not undergo this increase in electricity consumption during the late evening hours. The weekend profile at Altus AFB closely resembled the weekend load profile from Norman (Fig. 4.12) and Woodward (Fig. 4.19), the two residential sites used in this study.

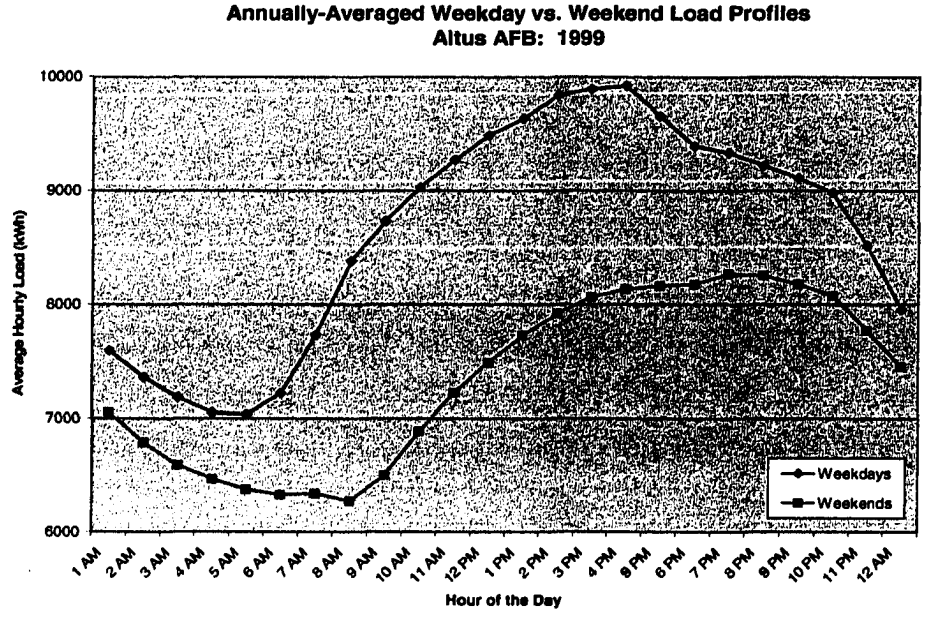


FIG. 4.26 Annually-averaged load profiles from Altus AFB in 1999 that illustrate the differences in the weekday versus weekend load.

Holidays that occurred on weekdays also impacted the load consumption at commercial entities. Daily load profiles for the four Mondays in February 1999 (15 February 1999 was President's Day, a federally observed holiday) are shown in Figure 4.27. The non-holiday Mondays mirrored the expected load pattern from a typical weekday during winter. However, during the workday hours, the use of electricity on President's Day was 30% lower relative to the other Mondays in February. The

difference was a consequence of Altus AFB closing its operation to observe a federal holiday. On a holiday, base residents might stay home, and, as a result, increase this component of electricity consumption for the AFB even though the usual weekday base activities (e.g., classes and training missions) were suspended. If base housing did not contribute to this holiday use of electricity, the load profile likely would have exhibited an even lower use of electricity. For load forecasting purposes, knowledge of how electricity is used on observed holidays must be incorporated into the load model if the accuracy of the load forecast is to be improved.

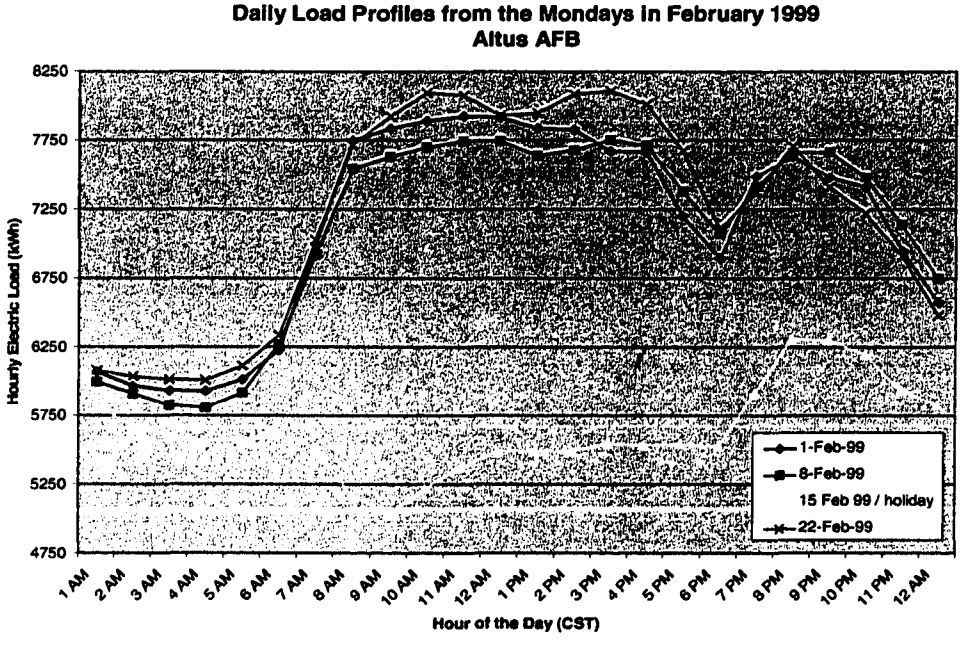


FIG. 4.27 Daily load profiles for the four Mondays in February 1999 from Altus AFB. The holiday Monday (President's Day) produced an electrical load with a different pattern than was observed on non-holiday Mondays.

Other load consumption features, evident in the daily load profiles, were masked in profiles drawn from larger time scales. For example, in Figure 4.27, the daily plot revealed a decrease in the use of electricity during the lunchtime hours. A steep slope in

the load profile between 6 AM and 8 AM increased at an average rate of 740 kWh per hour. To put that rate of increase in perspective, the average difference between the minimum and maximum (non-holiday) load on Mondays was ~2000 kWh. Thus, almost half of the growth in load (consumption) occurred within a 2-hour period (i.e., when residential use of electricity at the base was large as residents prepared for their day and as base employees arrived at work).

4.3.7 Average Load Profiles – Dominance

Dominance is the name of the substation that serves electricity to a single industrial entity, Pan Pacific Products, located in Broken Bow (southeast Oklahoma). The typical load profile for an industrial/large commercial customer (Fig. 4.6) reveals a small decrease in load consumption during the lunch hour. Otherwise electricity consumption during the period between sunrise and sunset usually is steady, especially when compared to other types of customers. Often industrial warehouses and factories work in shifts that cover the days, nights, and weekends. More specifically, Pan Pacific Products uses 8- and 12-hour shifts; they also observe 9 federally recognized holidays.[7] However, when their particular market (demand) fluctuates, the work schedule is adjusted to meet the demands of their customers. In other words, when the demand for their product increases, the company may institute mandatory overtime or may not recognize holidays. On the other hand, when the demand for their product decreases, they may institute early-release days or close the plant. This type of work schedule is a difficult, if not impossible challenge for load forecasters.

[7] Personal communication with Pan Pacific Products, January 2003

Figure 4.28 illustrates the annually-averaged load profiles for the Dominance substation[8]. Little variation is evident that relates to a pattern in a residential or commercial profile. For example, the average standard deviation in the annually-averaged load profiles at Dominance was 100 kWh, while at the residential sites of Norman and Woodward, the standard deviation was 800 kWh and 725 kWh, respectively. Additionally, the difference between the maximum and minimum values on the annually-averaged load profile was 370 kWh at Dominance, whereas both residential substations recorded a range in excess of 2000 kWh. Thus, the industrial customer at Dominance did not experience a wide range of load values as did the residential community.

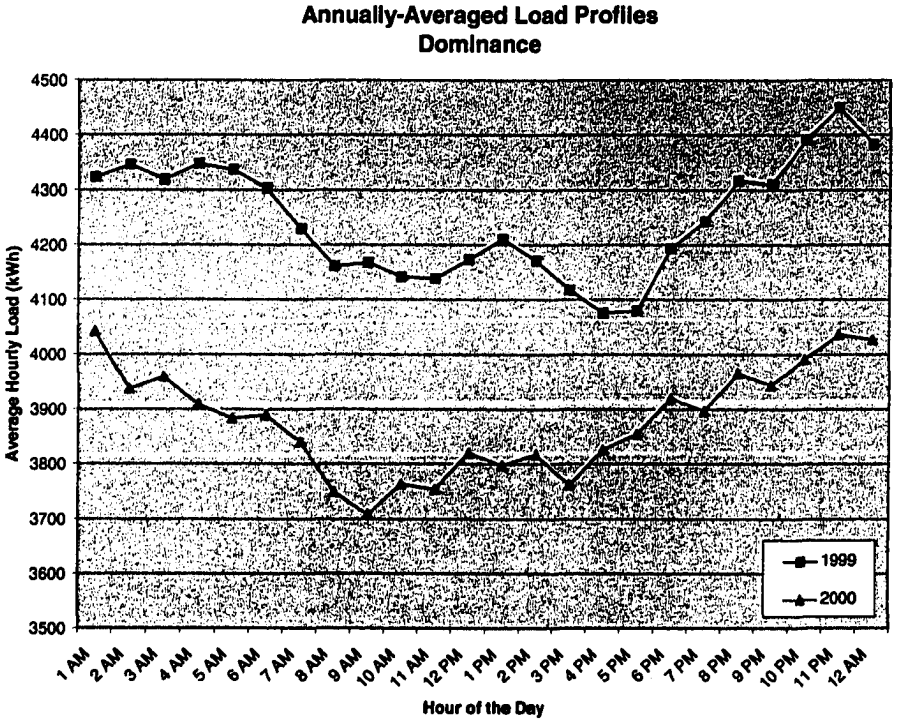


FIG. 4.28 Annually-averaged load profiles from the Dominance substation during the study period.

[8] The study period was 1999 and 2000 because a significant change in electricity consumption occurred prior to 1999.

Seasonal plots for the residential and commercial load revealed details that were masked by the process of computing annual averages. While the seasonal plots of the industrial customer (Figs. 4.29 – 4.32) were more detailed as well, the features were more difficult to explain than were details from load profiles of other customer categories. The most noticeable feature in the annually-averaged plot was that less electricity was consumed between 6 AM and 6 PM than during other times of an average day. This situation reflects the significant overnight use of electricity that was not observed in the residential or commercial profile. A reoccurring feature (e.g., consistent time of peak load or the magnitude of a peak load) was not evident in the seasonal profiles for Dominance. Thus, further analysis was necessary to characterize the profiles in hopes of improving the forecasts of electrical load.

The annually-averaged profile for Dominance (Fig. 4.28), along with the summer (Fig. 4.31) and fall (Fig. 4.32) seasonal profile indicated that the load consumption during 1999 was significantly greater than the load consumption during 2000. During the winter (Fig. 4.29) and spring (Fig. 4.30) seasons, the average hourly load was greater during 2000 than during 1999. Monthly profiles were examined to provide a more detailed explanation, but warmer than normal temperatures contributed to some of the load differences observed between the two years. The real world is more complicated.

During the first four months of 1999, Pan Pacific Products ceased plant operations for a few days up to a week at a time. In response to increased demand for products from Pan Pacific, the company began to operate continuously (e.g., 24-hours a day, 7 days a week) from May to October. Holidays were not observed (e.g., Memorial Day or Labor

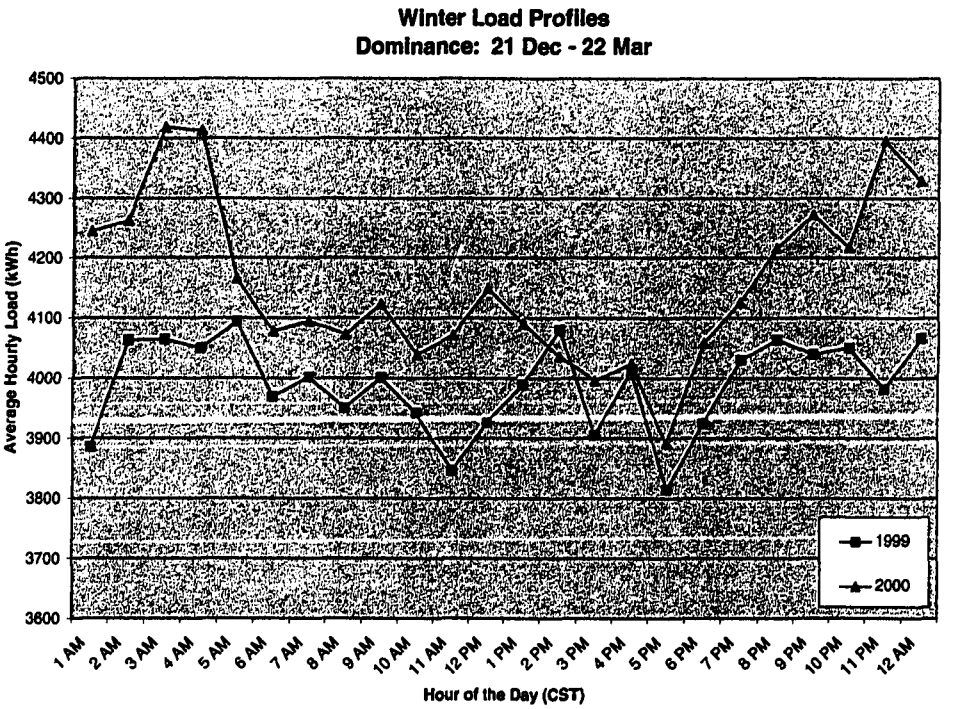


FIG. 4.29 Average hourly electric load profiles from the Dominance substation during the winters of the study period.

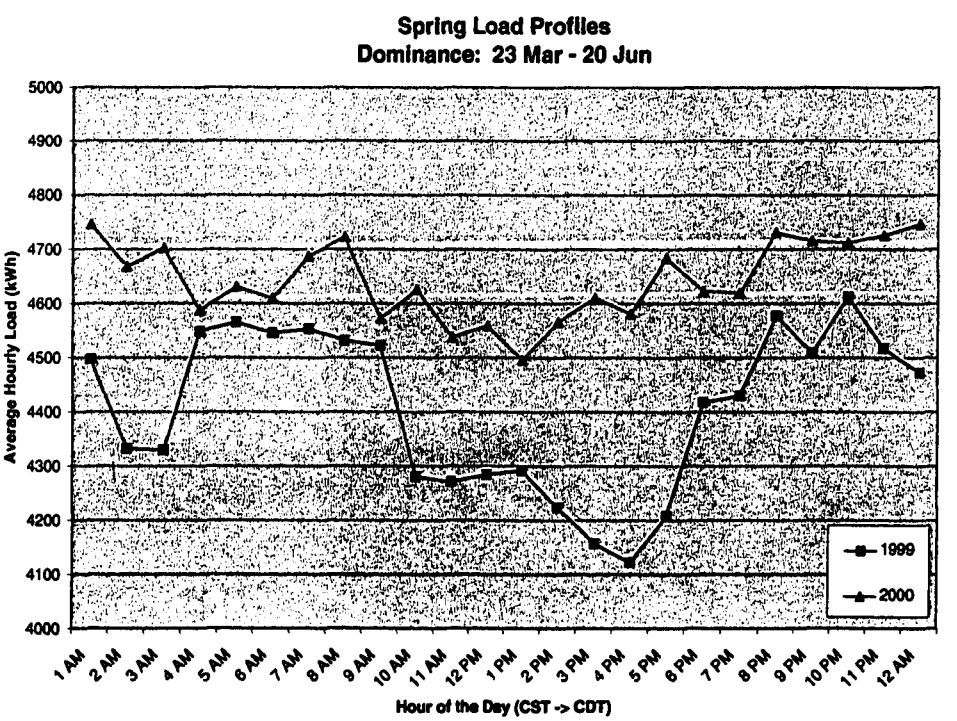


FIG. 4.30 As in Fig. 4.29, except for spring.

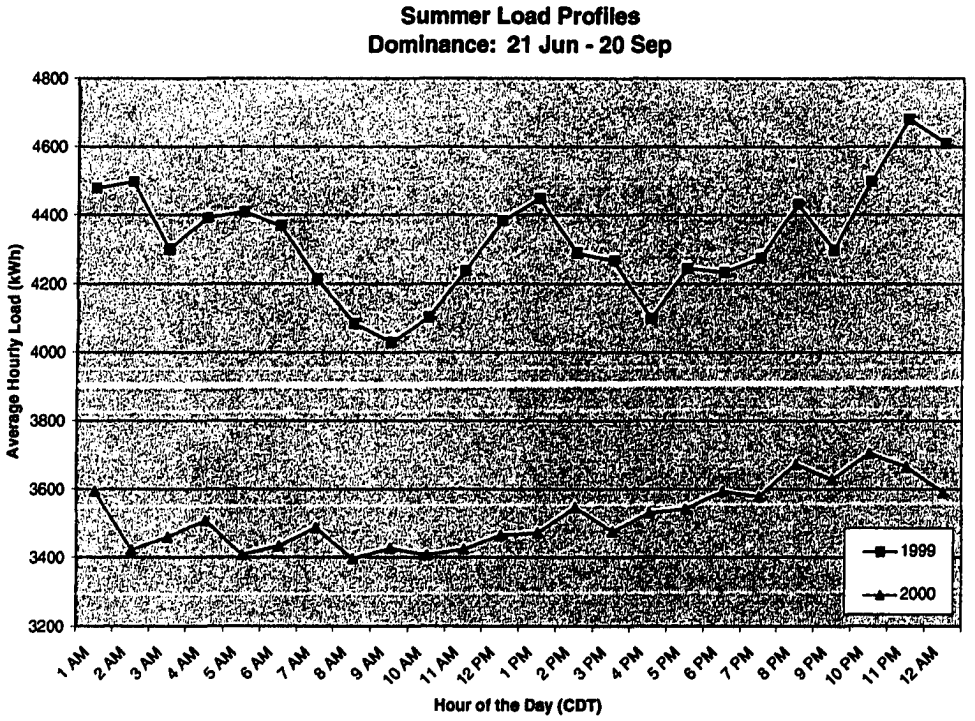


FIG. 4.31 As in Fig. 4.29, except for summer.

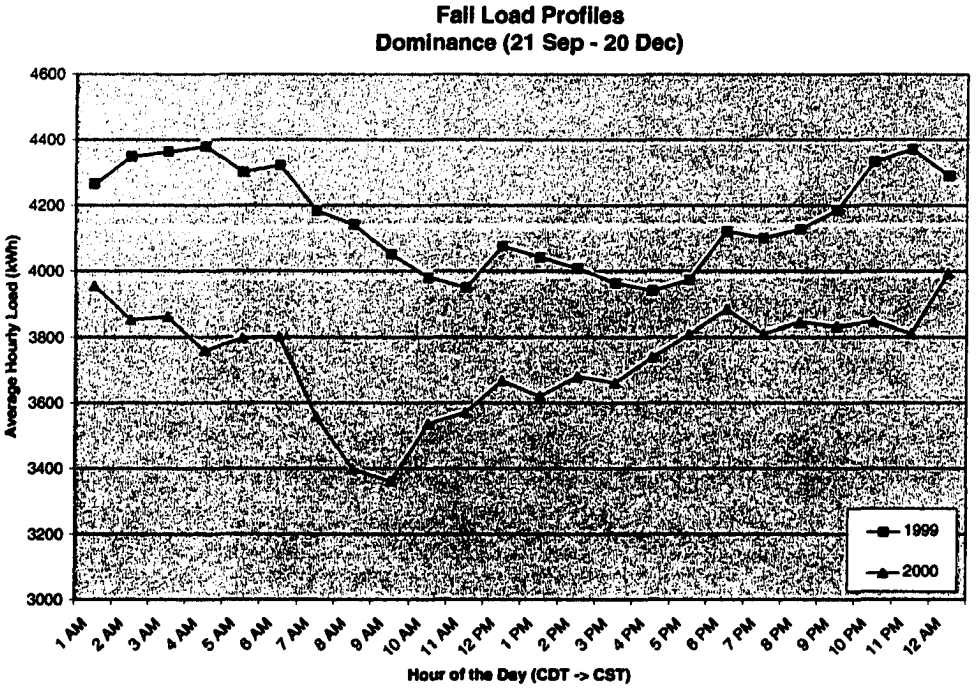


FIG. 4.32 As in Fig. 4.29, except for fall.

Day) during this period. During the last two months of 1999, the only downtime experienced by the plant was during the Thanksgiving and Christmas holidays.

This demanding work schedule persisted until July 2000 (i.e., the plant was in full operation on Easter Sunday and Memorial Day). The only break in operation prior to July was a slowdown during the last weekend in January. The opposite scenario unfolded during the final 6 months of 2000. A plant closure occurred on the 4th of July, along with intermittent downtime throughout the month. The plant also was closed for 16 days in August. From September through December, the plant closed every Sunday morning through Tuesday morning in addition to a one-week closing at Thanksgiving and at Christmas.

Unlike the electricity demand of the residential and commercial customers, electricity use at the Dominance industrial site is dependent almost entirely on the market demand for Pan Pacific's product. When the demand is high, plant operations increase, as does the electricity consumption. When the market takes a downward turn, plant operations slow or cease intermittently, and electricity use at the plant decreases. In other words, a change in seasons – weather conditions – does not alter the electricity load near as much as it did for the other three study sites.

4.3.8 Weekdays, Weekends, and Holidays - Dominance

Weekend versus weekday load profiles revealed significant differences in electricity consumption between residential and commercial customers. However, Pan Pacific Products operates just as much on weekends (unless a holiday or plant closure occurs), just as it does on weekdays. Thus, a partitioning of load profiles on weekdays

versus weekends does not add new information for a load forecaster. However, when workdays were separated from non-workdays (i.e., a stratification similar to the separation of weekdays and weekends for the residential or commercial sites), significant differences became clearly evident (Fig. 4.33). Thus, a load forecaster would greatly benefit from having knowledge of scheduled "non-workdays", especially because this one substation was dedicated to supplying electricity to Pan Pacific Products.

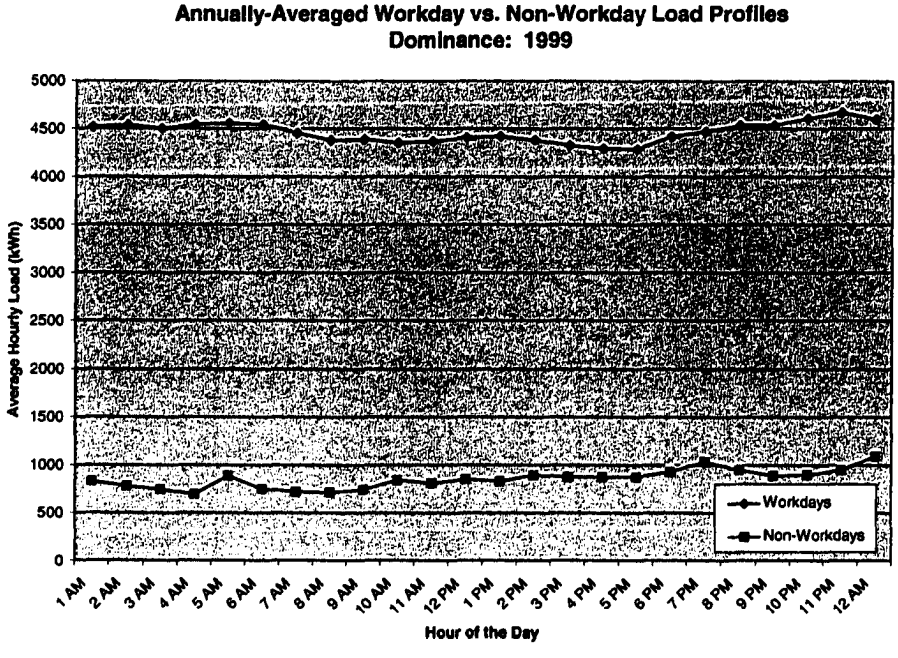


FIG. 4.33 Annually-averaged load profile from the Dominance substation in 1999 that illustrates load differences in the workday versus non-workday (e.g., when the plant was closed, had scheduled downtime, or observed holidays).

Holidays (when observed) created differences in the load profile that stand in marked contrast to the use of electricity on normal workdays. Using the load profiles from the Thursdays in November 1999 (Fig. 4.34), it was clearly evident when the plant

closed because the load profile decreased dramatically to a flat-line use of ~300 kWh by 6 AM on Thanksgiving Day. This pattern was observed on all days when the plant closed, had scheduled downtime, or observed holidays.

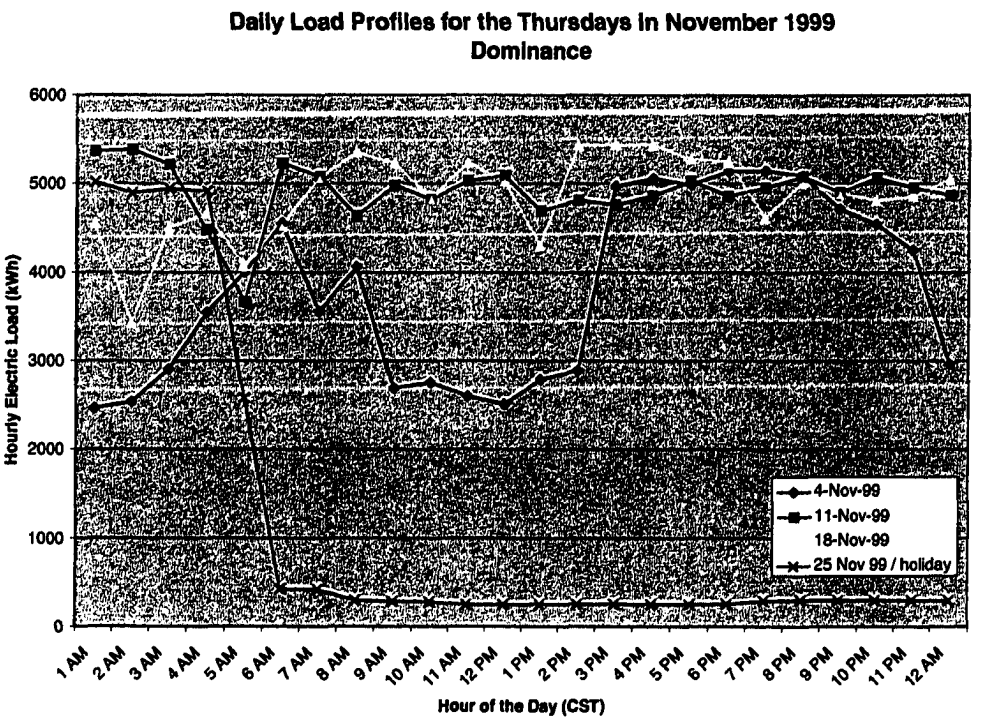


FIG. 4.34 Daily load profiles for the four Thursdays in November 1999 from the Dominance substation. The holiday Thursday (Thanksgiving Day) produced an electrical load with a much different pattern than observed on non-holiday Thursdays.

4.4 Quality Assurance of the Data

Obviously, the quality of all data will have a significant impact on research results, validation studies, assessments, and life-saving decisions. Quality can be described as a degree of excellence, while assurance is synonymous with certainty. Thus, quality assurance (QA) procedures are designed to give users a high degree of confidence

in their work because the data used were meticulously processed. QA also provides users with guidance in how best to use the data by flagging outliers or erroneous data points. Because temporary interruptions in data collection (e.g., power outages, failure of sensors, and routine maintenance) do occur in automated systems, quality assurance procedures and the processing of outliers or missing data is discussed.

4.4.1 Quality Assurance of Mesonet Data

The Oklahoma Mesonet instituted a comprehensive quality assurance system that parallels its efficient data collection and transmission process. The QA procedures include four components (Shafer et al. 2000): an instrument laboratory, field visits, automated computer routines, and manual inspection of the data by specially trained meteorologists. These procedures have been improved and updated as new insights became available about past problems in the network.

The archives of hourly data from the three-year study period were nearly complete, with only limited gaps. Because the Mesonet collects data with a five-minute time resolution, temporal interpolation was used to fill gaps in the hourly data. Spatial interpolation (substitution) of data from the next closest station also was used to eliminate gaps in this research data set. Both methods alleviated the minor problem of missing data.

The Washington Mesonet site is the nearest neighbor to the Norman Mesonet site by virtue of being ~17 miles (27 km) to the south southwest of Norman. When measurements were unavailable from the Norman site, data from the Washington site were used instead. This method was used when data gaps from the Norman site could

not be filled by temporal interpolation. Otherwise, the hourly data point was tagged as "missing". Data "gaps" (i.e., data interpolated, replaced, or missing) were less than 0.5% of the archived files during the study period for the Norman site (see Table 4.2 and Appendix E). Data from the Seiling Mesonet site were used in place of any missing data from the Woodward Mesonet site. Seiling is ~26 miles (42 km) southeast of Woodward. The "gaps" in the data archives for Woodward totaled 0.3%. When data from the Altus Mesonet site were missing, data from the Tipton Mesonet site (~15 miles or 24 km southeast of Altus) were substituted. Less than 0.3% of the archived files from Altus were incomplete during the study period. Finally, the nearest neighbor to the Broken Bow Mesonet Site is Idabel, located ~20 miles (32 km) southwest. Broken Bow had 0.5% of its data unavailable in the archived files.

Table 4.2 The number and percentage of observations interpolated, substituted or missing in the Mesonet data sets from the four study sites. Each site had 105,216 observations. The study data set had 420,864 observations.

Mesonet Site	***Interpolated Observations***		***Substituted Observations***		***Missing Observations***		***Total Observations***	
	#	%	#	%	#	%	#	%
Norman	92	0.09	271	0.26	88	0.08	451	0.43
Woodward	33	0.03	36	0.03	245	0.23	314	0.29
Altus	50	0.05	125	0.12	96	0.09	271	0.26
Broken Bow	140	0.13	262	0.25	151	0.14	553	0.52
Totals	580	0.07	694	0.16	315	0.14	1589	0.37

The scientific literature has demonstrated how load consumption is sensitive to changes in temperature. Thus, the "nearest neighbor" substitution is based on the work of

Brotzge and Richardson (2003); they determined that data from Mesonet sites up to 100 km apart were highly, spatially correlated (> 0.5) for the meteorological variables used in this study. As distance between the sites decreased, the spatial correlation increased. (The site pairs used in this study were 42 km or less apart.) Brotzge and Richardson (2003) also determined that air temperature and relative humidity were even more strongly correlated when Mesonet sites were aligned along a northwest/southeast or north/south axis (i.e., the spatial correlation was not isotropic). The Woodward-Seiling sites and the Altus-Tipton sites are favorably aligned along a northwest/southeast axis (Fig. 3.1). The Norman-Washington sites have a north/south orientation. The Broken Bow site does not have a Mesonet site to its east because of its location near the eastern border of Oklahoma. Thus, Idabel was chosen (southwest) because it is the nearest neighbor to Broken Bow.

Temperature (°C), relative humidity (%), solar radiation (Wm^{-2}), and wind speed (ms^{-1}) were collected at five-minute intervals from the four pairs of Mesonet sites during 2000. A total of 421,632[9] observations (i.e., *366* days/year x 24 hours/day x 12 observations/hour x 4 weather variables) observations were possible from each site. Table 4.3 lists the number of valid observations that occurred within ±5 or ±10 (°C, %, or ms^{-1}) of the neighboring observations (for the same time of day). For example, 99% (100%) of the valid temperature observations from the Washington Mesonet site were within ±5°C (±10°C) of the temperature observations from the Norman Mesonet site. Approximately 69% (91.8%) of the relative humidity values were within ±5% (±10%) of the neighboring observations and 99.8% (~100%) of the wind speed values were within

[9] The year 2000 was chosen to *represent* the study period. Little variation occurred from year-to-year. The year 2000 was a leap year, so there are 768 more observations in this set than non-leap years.

Table 4.3 The percentage of valid observations that were within ±5 and ±10 (°C, %, and ms^{-1}) of each other when comparing measurements from a pair of neighboring Mesonet sites.

(a) Norman Mesonet site and Washington Mesonet site

Norman & Washington Mesonet Sites	*% of Observations within ±5*	*% of Observations within ± 10*	*Total Number of Valid Observations (out of 105,408)*
Temperature (°C)	99.8	100	105,356
Relative Humidity (%)	69.0	91.8	105,356
Wind Speed (ms^{-1})	99.8	100	105,361

(b) Woodward Mesonet site and Seiling Mesonet site

Woodward & Seiling Mesonet Sites	*% of Observations within ±5*	*% of Observations within ± 10*	*Total Number of Valid Observations (out of 105,408)*
Temperature (°C)	98.7	99.9	105,207
Relative Humidity (%)	39.7	72.4	105,204
Wind Speed (ms^{-1})	98.9	100	105,207

(c) Altus Mesonet site and Tipton Mesonet site

Altus & Tipton Mesonet Sites	*% of Observations within ±5*	*% of Observations within ± 10*	*Total Number of Good Observations (out of 105,408)*
Temperature (°C)	99.4	100	104,629
Relative Humidity (%)	66.8	88.8	104,620
Wind Speed (ms^{-1})	99.6	100	104,631

(d) Broken Bow Mesonet site and Idabel Mesonet site

Broken Bow & Idabel Mesonet Sites	*% of Observations within ±5*	*% of Observations within ± 10*	*Total Number of Valid Observations (out of 105,408)*
Temperature (°C)	98.8	100	104,837
Relative Humidity (%)	54.3	81.7	104,841
Wind Speed (ms^{-1})	98.1	100	104,843

± 5 ms^{-1} (± 10 ms^{-1}) of each other. Similar results were determined using data from the other three pair of sites (Table 4.3).

Based upon Table 4.3, the percentage of relative humidity observations that were within 5 units (%) of its neighboring site was not as high as those for temperature and wind speed (°C, ms^{-1}). As shown in Chapter 5, relative humidity is most useful as a predictor of load during the summer months. Substitutions for relative humidity were used because: (1) only 0.16% of the data set (i.e., data from the four Mesonet sites) required the use of a "nearest neighbor"; and (2) a majority of the relative humidity differences between neighboring sites were less than 10%. The site pair with the largest differences in relative humidity (Woodward and Seiling) required relative humidity substitutions on only 9 occasions.

Although solar radiation was useful in this research, an analysis of spatial coherency was not performed. Even so, with the close proximity of neighboring Mesonet pairs (<42 km), Brotzge and Richardson (2003) determined that solar radiation had the least spatial variability of any Mesonet variable. They also determined that solar radiation was isotropic across the state, while temperature and relative humidity were not. Thus, a careful inspection of data from the four primary Mesonet sites and the four neighboring sites revealed that substituting neighboring data for missing or erroneous data was a sound decision.

4.4.2 Quality Assurance of WFEC Electric Load Data

WFEC has archived electric load data at 30 minute intervals for the past decade. Prior to 1994, WFEC collected and archived electric load data manually (via a meter

reader). As a result, few quality control measures were enforced. The data collection process has greatly improved at WFEC due to the implementation of new technology known as Automated Meter Reading (AMR). AMR revolutionized metering reading, and became widespread in the utility community during the 1990s. This technology uses land telephones, wireless phones or internet methods to record meter readings. Benefits of AMR include gaining access to information on the service needed by consumers, customer use patterns, and quality control techniques. More specifically, WFEC uses MV-90 software (a widely used AMR software package) which ingests meter data from virtually any solid-state[10] gas and electric metering device or data logging system. The software contains an interactive validation program that includes a routine to compare actual meter readings with the reading expected by MV-90 (i.e., load limits programmed by meter engineers at WFEC). Although AMR allowed for meter data to be transmitted electronically, the archiving of metered data remained a manual process with few quality control measures. This archiving deficiency was solved in 1999. Today, WFEC has a more efficient and effective method to receive and archive quality electric load data from meters to aid in billing, load forecasting, demand information, and research. Manual adjustments may be needed during power outages and data transmission problems.

Unlike Mesonet data, missing load data cannot be replaced using a "nearest neighbor" approach to substitute load data from a nearby substation. For example, the load demand at Substation A may not be comparable to the load demand for Substation B, even if the separation distance was only 10 miles because Substation A may service a customer base that is not comparable to those served by Substation B. On the other hand,

[10] A solid state meter is an electronic, multiple-function electric or gas meter that can measure advanced metering functions.

the weather conditions measured at Site A and Site B, separated by the same 10 miles, *typically* would represent similar atmospheric conditions. Thus, if load data were not archived, temporal interpolations or spatial substitutions were not made. Tables E.5-E.8 contain a list of electric load data missing in this data set. In summary, the percentage of missing or corrupt load data at the four substations was: 9 % for Norman, 0.1% for Woodward, ~0% (only 2 hours) for Altus AFB, and 0.15% for Dominance.

Based upon this sound foundation of quality data, Chapter 5 will document the interrelationships between weather data and concurrent load data for the study period. The purpose of this research is to use meteorological data that reflects changing weather patterns and determine the impact of those changes on load demand to improve the accuracy of electric load forecasting.

Chapter 5

Interrelationships Between Weather Variables And Electric Load Data

The scientific literature has documented how various weather phenomena have a significant impact on load demand. However, other than temperature, the literature is inconclusive as to which meteorological variables would provide value-added information to the immature science of load forecasting. Part of the problem in producing accurate load forecasts resides in the inconsistent or inadequate methods to incorporate weather data or the inability to even obtain weather data by utility companies. This chapter will document which weather variables or combinations thereof prove to be most significant in load forecasting. Hourly data from the four Mesonet study sites and the four co-located substations in the Western Farmers Electric Cooperative system are used to define the interrelationships between the variability of weather and the variability of electric load. The goal of this chapter is to quantify the weather-load relationship at the four study sites.

5.1 Significant Weather Variables

Historically, temperature has proven to be the primary weather variable that impacts electric load. It also has been the most readily available weather variable to load modelers and load forecasters. However, over time, relative humidity and wind speed have been added to the short list of weather variables used in load forecasting because the combination of temperature, relative humidity and wind speed greatly impact the daily

comfort level of humans (Stull 2000). Because the Oklahoma Mesonet has produced a rich resource in its data archives, other weather variables were considered during this study. However, to achieve parsimony in the load models, one should only use those meteorological variables which prove to have an impact on the load forecast. Yet, the use of too many exogenous variables can only increase the complexity of prediction models with little or no improvement in the resulting load forecasts. For example, too many variables given to a NN-based model only increases the size of the network, decreases the speed at which training of the model will occur, and results in insignificant changes to the load forecast (Charytoniuk and Chen 2000).

While neither humidity nor wind speed have a documented, independent relationship with electric load, they can be combined with temperature to form an apparent temperature. The apparent temperature (T_{app}) takes into account relative humidity (RH) during the warmer months (as a heat index) and takes into account wind speed (V) during the colder months (as a wind chill). The scientific formulas for heat index and wind chill are given in Appendix D. In this study, the apparent temperature is defined as:

$$T_{app} = \begin{cases} \textit{Heat Index}, & \text{if } T > 70^{\circ}\text{F and RH} > 30\% \\ \textit{Wind Chill}, & \text{if } T < 50^{\circ}\text{F and V} > 3\,\text{mph} \\ T & \text{, in all other situations} \end{cases} .$$

Other variables, such as precipitation and cloud cover, have been documented in the scientific literature to not provide significant improvements to the short-term load forecast. For example, solar radiation has not been an input variable to any load models, because the variable (indirectly) reflects cloud cover. However, solar radiation could

have psychological implications on the electric load during the fall, winter and spring months (e.g., the impact of 60°F on a cloudy day versus the impact of 60°F on a sunny day.)

Scatter plots of weather variables (i.e., temperature, apparent temperature, relative humidity, wind speed and solar radiation) versus load were created on an annual and seasonal basis for the three study years (e.g., 1998, 1999 and 2000)[1] for all four study sites. Profiles of monthly load and temperature averaged on various time scales and temporal correlation calculations also were used as a means to quantify the weather-load relationship at each study site. (*Unless otherwise stated, all three years revealed similar results. Thus, the following plots were derived from portions of the study period to represent the entire data set.*)

5.2 Norman Site

The Norman Mesonet site and the West Norman substation are located in a primarily urban, residential community. As the scientific literature has shown, the load demand is, in large part, driven by the human response to weather conditions.

5.2.1 Temperature-Load Relationship (Norman)

5.2.1.1 Scatterplots

As clearly indicated in Figure 5.1, a strong nonlinear relationship existed between temperature and load. Thus, on an annual scale, the temperature-load relationship is more effectively analyzed and modeled when nonlinear methods are used. On a seasonal scale, winter, spring and fall also exhibited a nonlinear structure (Figs. 5.2a, b, and d).

[1] All three study years were used in the analysis unless stated otherwise in Chapter 4.

During the summer season (Fig. 5.2c), the temperature-load relationship was almost linear (e.g., as the temperature increased, the load consumption increased), implying that a linear fitting technique could be used. If the load modeling results proved to be comparable regardless of which model (for the winter and summer) was used, the simpler linear model could represent the summer season while a nonlinear model or neural network could represent the more difficult-to-forecast winter and transition months (Darbelley and Slama 2000).

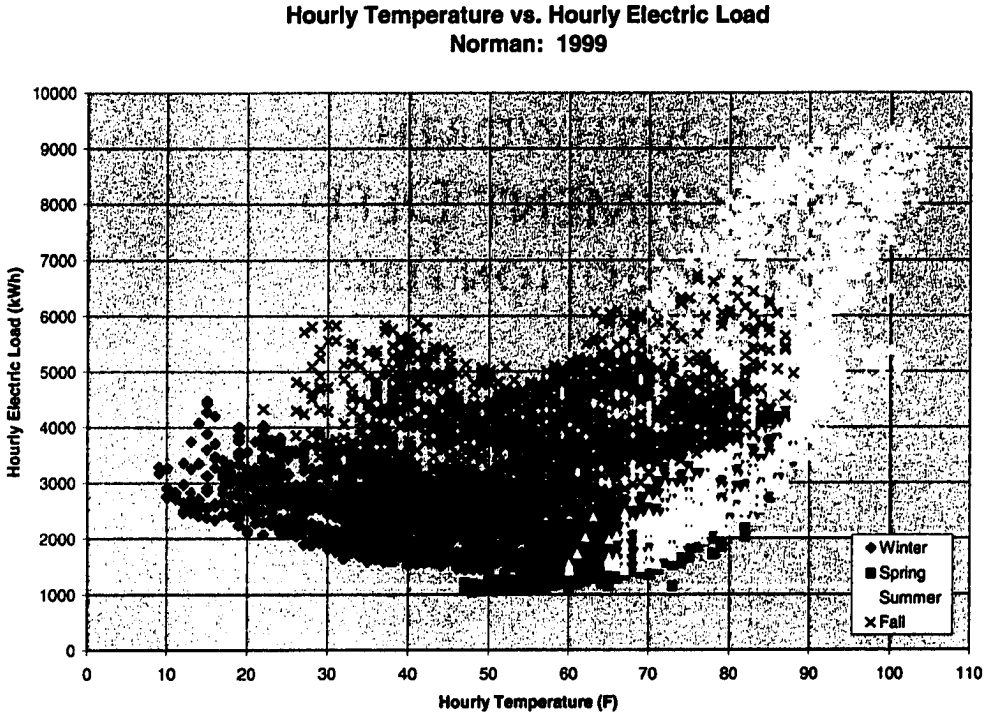
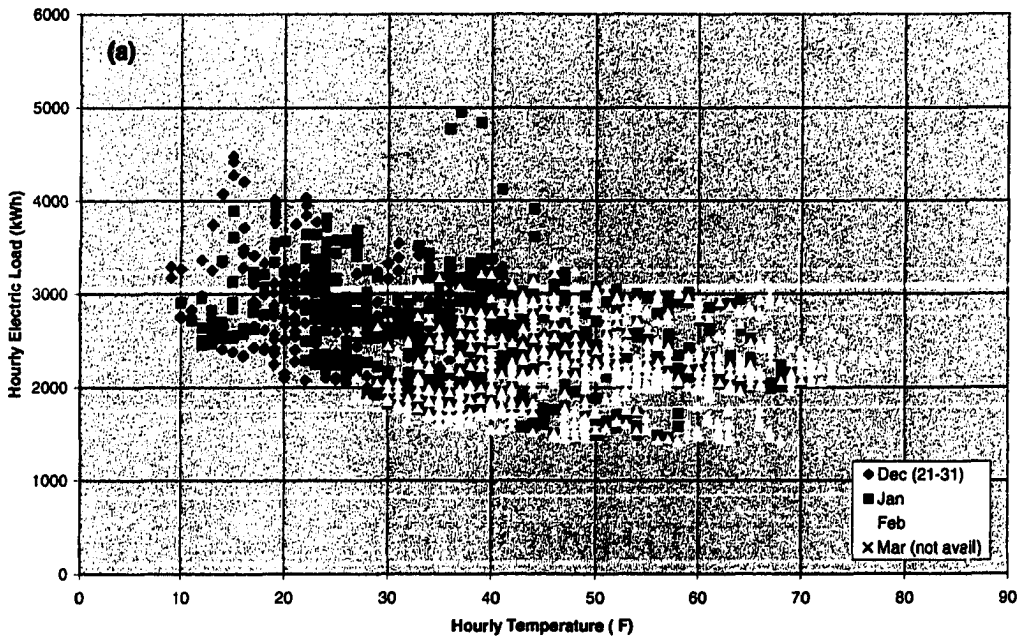
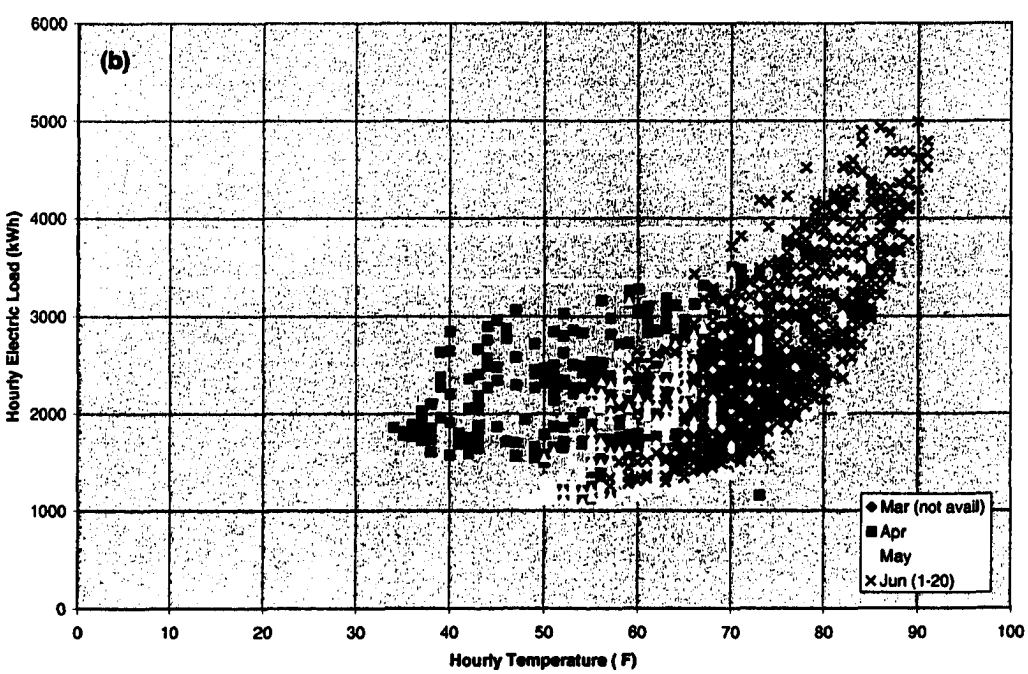


FIG. 5.1 Scatterplot of hourly temperature from the Norman Mesonet site versus hourly electric load from the West Norman substation for 1999.

Hourly Temperature vs. Hourly Electric Load
Norman: Winter 1999 (21 Dec 98 - 22 Mar 99)



Hourly Temperature vs. Hourly Electric Load
Norman: Spring 1999 (23 Mar - 20 Jun)



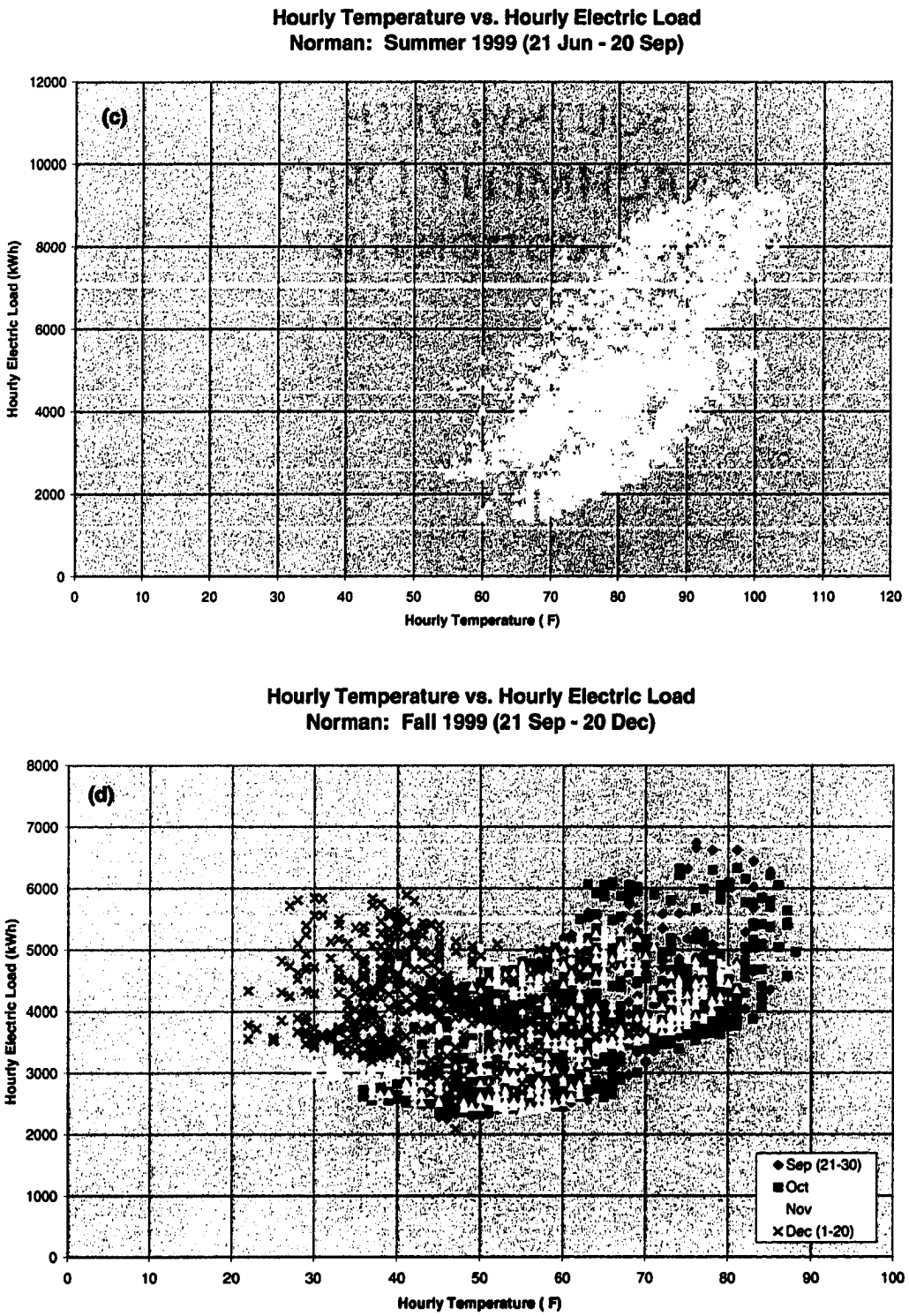


FIG. 5.2 As in Fig. 5.1 except for (a) winter, (b) spring, (c) summer, and fall of 1999.

Because the winter of 1999 was relatively mild, the winter load consumption was low (Figs. 5.1 and 5.2a). Intuitively, cold winter temperatures should create elevated levels of electricity consumption that might be comparable to the summertime peaks. However, much of the Norman community uses natural gas as their primary source of heat[2]. While central heating units[3], space heaters, electric furnaces and heat pumps require electricity to operate, more electricity is required to run an air conditioning unit, which in turn, creates a higher demand for electricity during the summer. Thus, communities which encourage the use of natural gas as a primary source of heat may not experience high electrical loads during winter (relative to summer), no matter how frigid the temperatures. Overall, the Norman load in winter appeared to have an inverse relationship with temperature that weakened as the temperatures increased (Fig. 5.2a), especially beyond 45°F (7.2°C).

An obvious division between early spring (e.g., late March[4] and April) and late spring (e.g., May and early June) was reveal in Figure 5.2b. The early spring values in the scatterplot mimicked those late in the winter season (Fig. 5.2a), when load seemed almost independent of temperature. The warmer spring temperatures and their corresponding loads of late spring were similar to those of the summer season (Fig. 5.2c). Thus, the scatterplot (Fig. 5.2b) clearly illustrated a transition in load use as customers responded to a transition in temperature during the spring season.

As an aside, a mean daily temperature value of 65°F (18.3°C) is used as the base temperature to calculate heating and cooling degree days (i.e., a quantitative index derived to reflect the demand for energy to heat or cool a residence or business;

[2] Personal communication, Oklahoma Natural Gas, 2002.
[3] Electricity is required to operate the fan inside central heating units.
[4] Load data at Norman during March 1999 was not available (Table E.5).

Appendix C). In comparison, ~62°F (16.7°C) seemed to be the benchmark temperature when load consumption transitioned into a different pattern for the Norman site during 1999 (Fig. 5.1). Thus, 62°F (16.7°C) could be considered as the base value from which site-specific heating or cooling degree days should be calculated. This threshold temperature could also signal a significant change in load demand and the need for load forecasters to shift to different forecasting models.

The temperature-load relationship during the summer season (Figs. 5.1 and 5.2c) was linear (i.e., as the temperature increased, the load consumption increased). This feature resulted from an increase in the use of air conditioners as temperatures warmed to uncomfortable levels. The summer season also had the largest variation in load relative to other seasons, spanning ~8500 kWh.

The load consumption during the fall of 1999 (Figs. 5.1 and 5.2d) appeared to be a combination of features from the other three seasons. However, the colder side of the plot was peculiar in that for the same (cold) temperature at often the same time of day, the load consumption was noticeably higher during the fall than during the winter (Fig. 5.1). The psychological impact of changing weather could have been a factor. For example, the fall season follows the sweltering "dog days" of summer. Further, Oklahoma is known for its abrupt temperature swings, especially during the fall season (e.g., a temperature drop of 40°F / 22.2°C within a few weeks during the fall of 1999; Section 4.3.4). Thus, a short period often exists during which the community acclimates to cooler fall temperatures. On the other hand, the winter season follows the "transitional" fall season, in which consumers likely experience a series of cold snaps prior to onset of winter. Another factor inherent to Norman could have been the fact that

the University of Oklahoma closed for a few weeks during winter break. As the students departed from Norman, the winter load decreased regardless of the weather. At warmer temperatures on the scatterplot (Fig. 5.1), early fall values were intertwined with summer values, which in turn, created high nonlinearity in the fall load.

5.2.1.2 Line Graphs

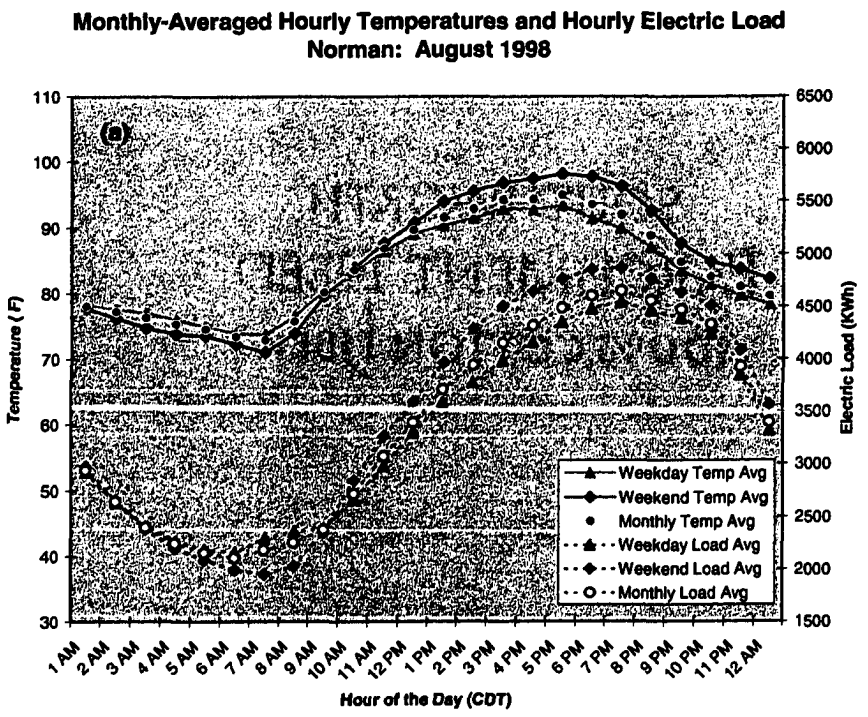
Diurnal plots of monthly-averaged temperature and electric load for the Norman site were created (Figs. 5.3a-b) to analyze the temperature-load relationship on a smaller time scale. Weekdays and weekends/holidays were plotted separately (Figs. 5.3a-b) because the weekday load demand is much different than the load demand on weekdays/holidays (Christiaanse 1971; AlFuhaid et al. 1997; Khotanzad et al. 1998; Charytoniuk and Chen 2000). Weekly data from representative weeks during each of the four seasons (Figs. 5.4a-d) also were plotted. The temperature profiles illustrated the expected diurnal cycle (Figs. 5.3a-b; Figs. 5.4a-d) and revealed how temperature varied during any particular month across the study period.

The monthly load profiles (Figs. 5.3a-b) revealed details about the residential load profile that were smoothed out of the residential *base* load profile (Fig. 4.6). Consequently, the summer months were easy to detect because they exhibited a much different load profile than those observed during other seasons (Figs. 5.3a-b). Customers used more electricity for longer periods throughout the day during the summer relative to other seasons (Fig. 5.3a). A well-defined load minimum occurred during the predawn hours before most residents awoke for the day. During the next few hours, the sun rose, air temperature increased, and residents began their day (e.g., a combination of activities

that typically increase the load demand). Air temperature increased ~2 hours before the steady rise[5] in the use of electricity. Thus, hourly load trailed the hourly temperature profile by ~2 hours. Further evidence of the 2-hour lag is discussed in subsequent sections. After a steady increase of the daily load during the summer months, peak use in the August profile occurred at ~7 PM. Thereafter, the load decreased towards its early morning minimum such that the load profile and temperature profile had similar characteristics during the overnight hours.

During the summer, a difference was observed between the weekday and weekend[6] load (Fig. 5.3a). Between 6 AM and 9 AM, the weekday load was ~15% greater than the weekend load even though the average weekday and weekend *temperature* were nearly equivalent (Fig. 5.3a). The morning portion of the two load profiles reflected a primary difference in the morning routines of customers on weekdays versus on weekends (i.e., people consume more electricity during the morning hours as they prepare for work or school, while they begin a weekend morning on a more leisurely pace). After 9 AM, the weekend load consumption exceeded the weekday load consumption. If temperature is not considered, it appears that Norman residents consumed ~12% more electricity during the afternoon and evenings on weekends than they did on weekdays. It is also possible that Norman residents were home more on the weekends than they were on the weekdays.

[5] "Steady rise" refers to the time period after 9 AM – that which is less affected by the morning spurt of electricity use as residents prepared for their day.
[6] Weekends also include federally observed holidays (e.g., New Year's Day, Martin Luther King, Jr. Day, President's Day, Memorial Day, Independence Day, Labor Day, Columbus Day, Veteran's Day, Thanksgiving Day, and Christmas Day).



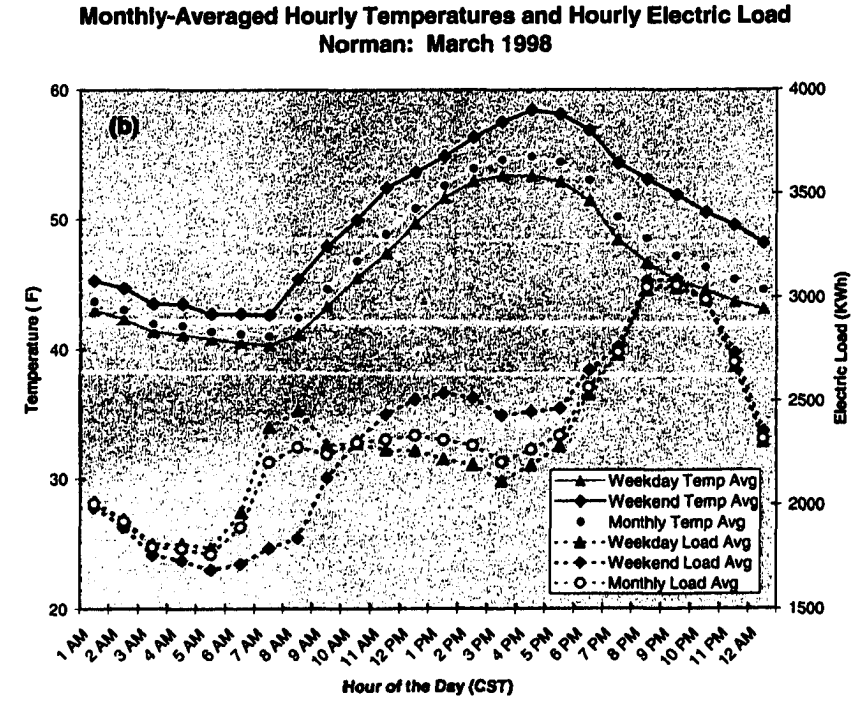


FIG. 5.3 Diurnal plot of hourly temperature and hourly load from Norman averaged during weekdays, weekends/holidays, and a month for (a) August 1998 and (b) March 1998.

However, when the effect of temperature (Fig. 5.3a) is considered, weekend temperatures during August 1998 were greater than the weekday temperatures after 11 AM. Thus, the weekend load *and* weekend temperatures were greater than those which occurred on weekdays. Was the load difference a result of hourly temperature differences, or, was the load difference a result of the fact that people were at home more during the day on weekends than they were on weekday? The weekend load exceeded the weekday load after 9 AM, prior to any significant hourly temperature differences on the weekdays versus weekends. Between 9 AM and 11 AM (e.g., when average weekend and weekday temperatures were equivalent), many residents were home on weekends, thus consuming more electricity than would the case during this same 2-hour period on weekdays. However, the greater use of electricity on the weekend during the *afternoon* likely was a combined affect of higher weekend temperatures and the fact that Norman residents likely remained at home more on weekends than they did on weekdays. The daytime temperatures peaked at 5 PM while the load peaked at 7 PM, more evidence of a ~2 hour lag in the load response. Approximately 3 hours after their respective peaks, the weekend temperatures and the weekend loads declined at a faster rate than those which occurred during the weekdays. While Norman residents may or may not have been home more throughout the day on weekends versus on weekdays, temperature differences between weekdays and weekends contributed to load differences. In fact, the temperature differences after 11 AM may have created much of the load difference that occurred on weekdays versus that which occurred on weekends.

During the spring, a bimodal distribution (Fig. 5.3b) existed in the electrical load. (A similar load profile was observed during the winter and fall seasons – not shown.)

However, the amplitude of the morning peak was not as pronounced during June (e.g., the month that included spring and summer) and September (e.g., the month that included summer and fall) because these months (not shown) were influenced by temperature and load trends associated with the summer season (Fig. 5.3a). This damping effect also is evident during May when the onset of summer temperatures occurs or during October when the summer temperatures linger into the fall season.

A major evening peak occurred at ~8 PM with the secondary peak at ~8 AM (Fig. 5.3b). A well-defined secondary morning peak was most evident on the weekday and monthly-averaged load profiles, reflecting the morning routine associated with a residential community. (The morning spurt of electricity use also occurred during the summer months, but the load profile did not level off or decrease after 8 AM as it often does during the other three seasons.)

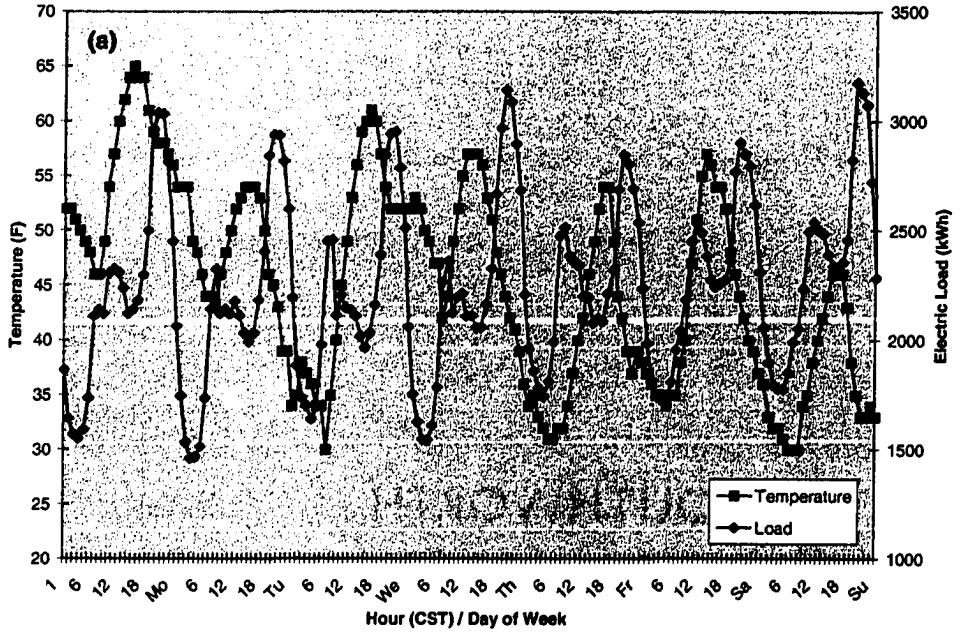
During March 1998, the average weekend temperature was *higher* than that of the weekday for the entire day (Fig. 5.3b). The morning section of the weekday load was expected to exceed that of the weekend load regardless of the temperature difference. However, after 10 AM, the same scenario unfolded during March as it did during August when weekend loads surpassed the weekday load. It was difficult to discern how much of the load difference was due to higher weekend temperatures or a larger number of residents being home on weekend afternoons. After 6 PM, the weekday load sharply increased (i.e., a byproduct of customer habit as residents returned home from work). By evening, no difference existed in the electrical loads.

The temperature and electric load profiles plotted on a weekly time scale further illustrate their interrelationships. A one-week period from each season during 1999 is

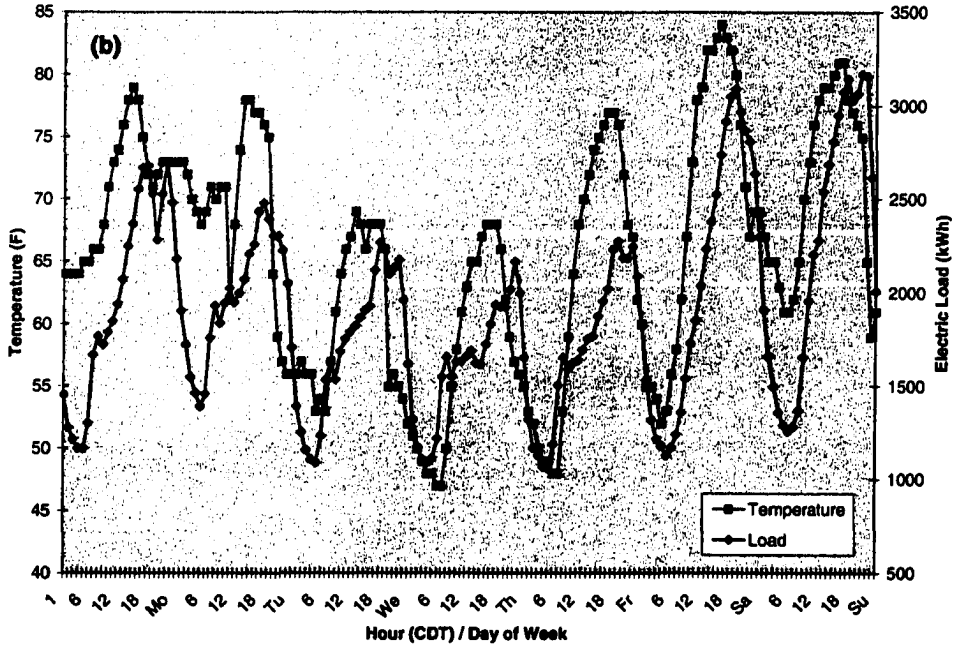
shown (Figs. 5.4a-d). During the winter, an inverse relationship between the temperature and load occurred in the one-week data set (Fig. 5.4a). This inverse relationship was most noticeable during the late afternoon hours. As the temperature increased towards its maximum daily value, load values decreased to a secondary (afternoon) minimum; the primary minimum load value occurred at ~3 AM, an hour during which almost all residential customers likely were asleep.

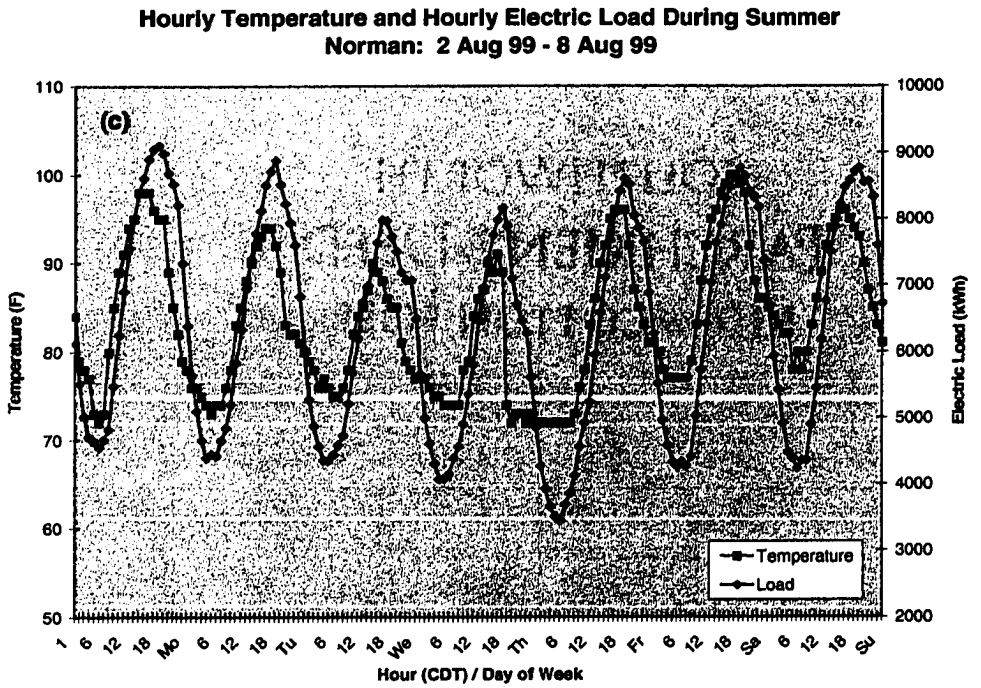
A one-week data set during the spring (Fig. 5.4b), summer (Fig. 5.4c), and fall (Fig. 5.4d) seasons revealed a direct relationship between temperature and load (i.e., as the average temperature increased between sunrise and sunset, the average load demand increased). However, the data revealed the load response to temperature was on average 2 hours. Additionally, the fall season often exhibited the most dramatic temperature swings among the four seasons. For example, a maximum temperature of 53°F (11.7°C) on 2 November was followed by 77°F (25°C) on 5 November (e.g., a 24°F difference in the daily maximum temperature separated by only *3 days*). These rapid temperature changes associated with fall caused fluctuations in the load demand as humans adjusted to these abrupt temperature changes. Thus, the fall season presents great difficulty to load forecasters.

Hourly Temperature and Hourly Electric Load During Winter
Norman: 15 Feb 99- 21 Feb 99



Hourly Temperature and Hourly Electric Load During Spring
Norman: 3 May 99 - 9 May 99





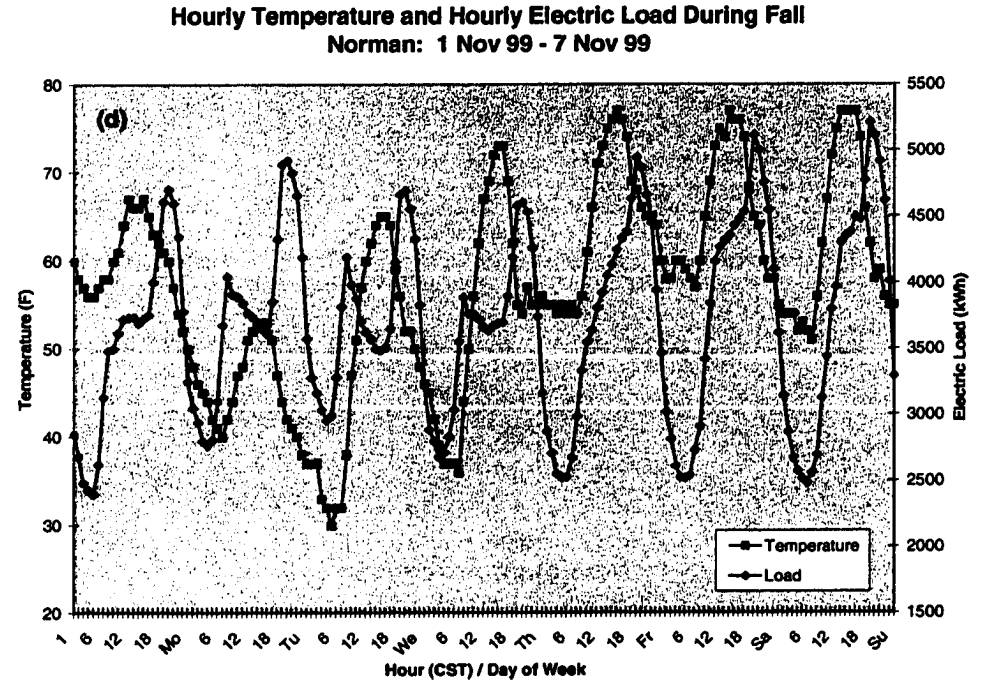


FIG. 5.4 Hourly temperature and hourly electric load from Norman during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

5.2.1.3 Apparent Temperature-Load Relationship

An analysis of the impact of apparent temperature (T_{app}) on load consumption did not reveal a significant difference when the scatterplots and profiles of temperature-load and T_{app}-load were compared (not shown). A scatterplot for T_{app} and temperature detected only minimal divergence in the scatter and then only at the extreme ends (e.g., at the hottest and coldest temperatures) of the plot.

5.2.1.4 Temporal Correlations

To quantify the temperature-load relationships, correlation coefficients (ρ) were calculated for seasonally-averaged, hourly temperature versus seasonally-averaged, hourly electric load. Hour-to-hour correlation coefficients between the two data streams (Table 5.1a) revealed a strong, positive relationship during the spring (ρ = 0.7) and summer months (ρ = 0.87). The positive relationship between the two variables was even stronger during all seasons (except winter) when the hourly electrical load lagged the hourly air temperature by 2 hours. The stronger correlation (with the 2-hour lag) reflected two processes: (1) modern homes in Norman are better insulated compared to older homes in more rural areas, and (2) the air temperature must rise (fall) above (below) a certain threshold before the human discomfort level becomes widespread among consumers who receive electricity via the West Norman substation.

The relationship weakened during the fall season (ρ = 0.29), disappeared and became moderately negative (ρ = -0.32) during winter (i.e., the colder the temperature, the greater the demand for electricity). Accordingly, the temperature-load relationship during the fall season remained difficult to describe (relative to the other three seasons).

For completeness, correlation coefficients between the T_{app} and the electrical load were included in Table 5.1a. The results revealed a relationship that was no stronger than the temperature-load relationship, and thus T_{app} offered no new information.

Table 5.1 Hour-to-hour correlation coefficients averaged seasonally over the study period between load and the following weather variables: temperature, apparent temperature, relative humidity, and solar radiation. When a higher correlation resulted from a timed lag (i.e., the number of hours the load lagged the weather variable) versus a zero lag coefficient, the highest correlation coefficient was shown. (The number of hours that the load lagged the weather variable is in parentheses.)

(a)	*Correlation Coefficients (X, Load) – Norman Site*			
Variable Correlated with Load (X)	**Temperature**	**Apparent Temperature**	**Relative Humidity**	**Solar Radiation**
Winter	-0.321	-0.338	-0.006	0.397 (7 hrs)
Spring	0.703 (2 hrs)	0.696 (2 hrs)	-0.392 (1 hr)	0.641 (5 hrs)
Summer	0.867 (2 hrs)	0.850 (2 hrs)	-0.741 (1 hr)	0.741 (4 hrs)
Fall	0.294 (3 hrs)	0.258 (2 hrs)	-0.277 (3 hrs)	0.562 (6 hrs)

(b)	*Correlation Coefficients (X, Load) –Woodward Site*			
Variable Correlated with Load (X)	**Temperature**	**Apparent Temperature**	**Relative Humidity**	**Solar Radiation**
Winter	-0.604 (1-2 hrs)	-0.613 (2 hrs)	0.290 (2 hrs)	-0.181 (3 hrs)
Spring	0.649 (0-1 hrs)	0.609	-0.307	0.534 (4 hrs)
Summer	0.921 (1-2 hrs)	0.907 (1 hr)	-0.694	0.791 (3 hrs)
Fall	-0.250 (2 hrs)	-0.255 (1-hr lag)	-0.145	0.300 (6 hrs)

(c)	*Correlation Coefficients (X, Load) –Altus AFB Site*			
Variable Correlated with Load (X)	**Temperature**	**Apparent Temperature**	**Relative Humidity**	**Solar Radiation**
Winter	0.233	0.217	-0.116	0.229
Spring	0.788	0.775	-0.421	0.496 (3 hrs)
Summer	0.857	0.839	-0.664	0.708 (3 hrs)
Fall	0.551	0.521	-0.424	0.454 (3 hrs)

(d)	*Correlation Coefficients (X, Load) –Dominance*			
Variable Correlated with Load (X)	**Temperature**	**Apparent Temperature**	**Relative Humidity**	**Solar Radiation**
Winter	0.072	0.071	0.019	-0.016
Spring	-0.112	-0.124	0.020	-0.068
Summer	-0.067	-0.054	0.030	-0.035
Fall	0.051	0.047	-0.026	-0.023

5.2.2 Relative Humidity-Load Relationship (Norman)

Other weather variables and their relationship to load were explored. Relative humidity is a ratio of the actual amount of water vapor in the air relative to the maximum amount that could be held in the air at that temperature. During the warmer months, higher relative humidity makes it feel hotter than the actual temperature. Figure 5.5 represents a scatterplot of the relative humidity-load relationship for 1999. (Scatterplots for other years are not shown, but similar results were obtained). Though the diagram seemed to show a seasonal stratification, the following inferences could be drawn.

During the winter season, the load was not responsive to changes in relative humidity (minor changes in load but major variations in relative humidity from one extreme to the other). The transitional seasons produced similar results. During the summer, as relative humidity increased, the load decreased. Because of an inverse relationship between *temperature* and relative humidity during the summers (Fig. 5.6), as temperature increased, relative humidity decreased, and the load consumption increased.

The correlation coefficients of relative humidity and load data confirmed this inverse relationship (Table 5.1a). Temporal correlations were high (ρ = -0.74) when the load lagged the relative humidity by 1 hour, but only during the summer season. During other seasons, the highest correlations resulted when the load lagged relative humidity by 1-4 hours. The smearing effect created by relatively high correlation values spread over several hours helped explain why the T_{app}-load relationship did not yield higher correlations than the temperature-load relationship. Perhaps if the strongest correlation coefficients for the temperature-load relationship and the relative humidity-load relationship occurred at the same lag time, the T_{app}-load relationship might have been stronger than temperature-load relationship during the summer. Thus, the relative humidity offered little independent information for the Norman site.

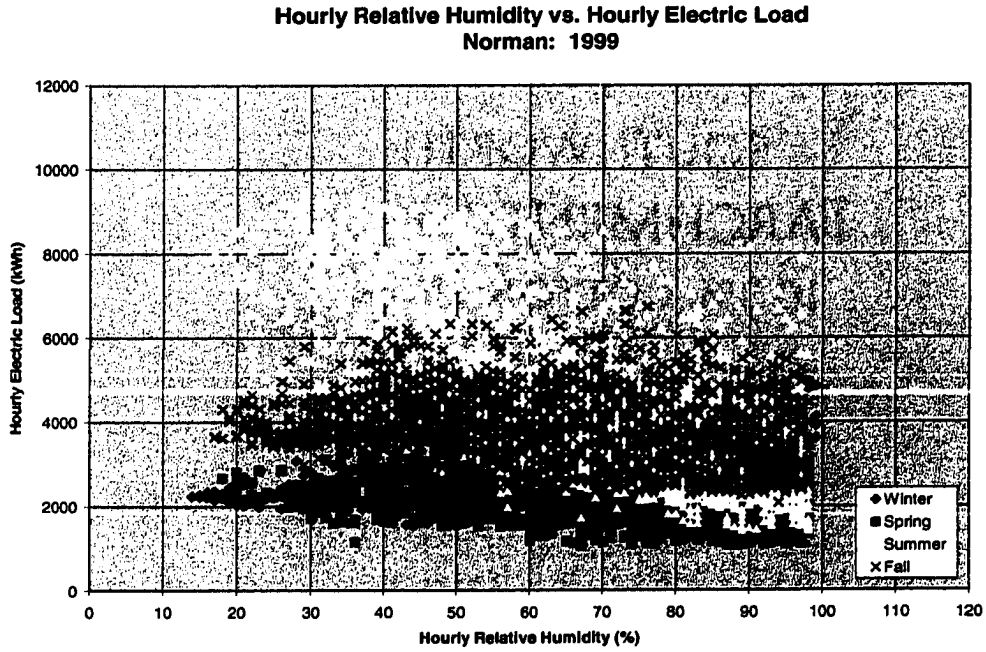


FIG. 5.5 Scatterplot of hourly relative humidity from the Norman Mesonet site versus hourly electric load from the West Norman substation for 1999.

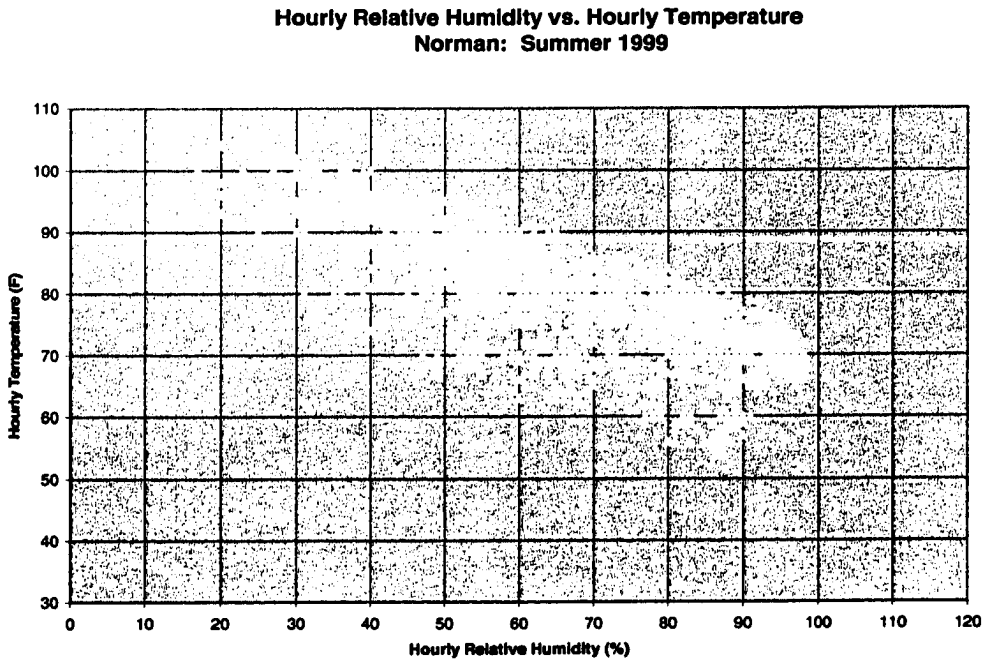


FIG. 5.6 Scatterplot of hourly relative humidity versus hourly temperature from the Norman Mesonet site for 1999.

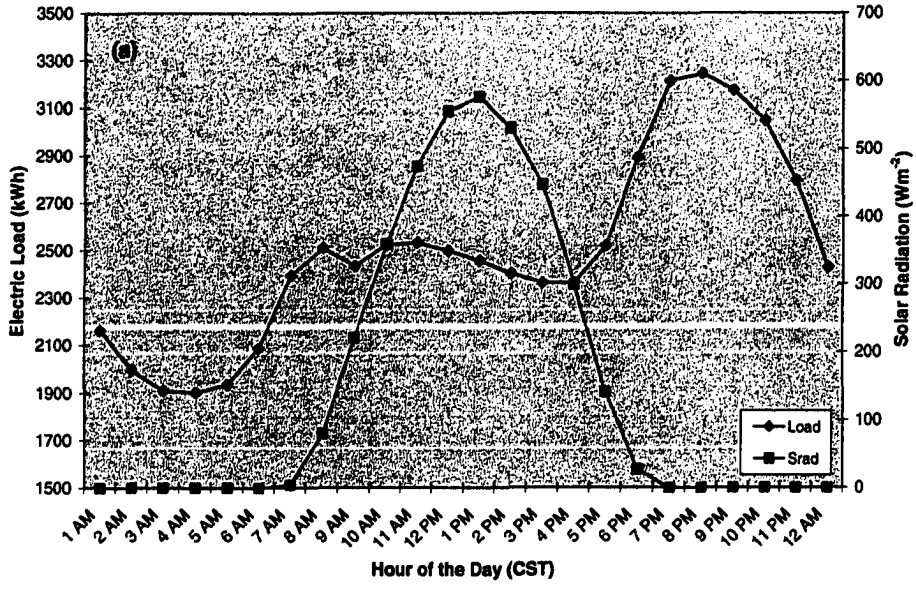
5.2.3 Solar Radiation-Load Relationship (Norman)

Solar radiation is electromagnetic radiation emitted by the sun. Mesonet observations of this parameter provide information about how much of the sun's energy strikes a particular location on earth at a particular time. Solar radiation reached it maximum value at solar noon. Yet, the peaks in solar radiation occurred hours prior to the daily maximum temperatures and the peak loads. Thus, higher correlation coefficients occurred when the load lagged solar radiation by several hours. A scatterplot of the solar radiation-load relationship (not shown) did not reveal any new insights.

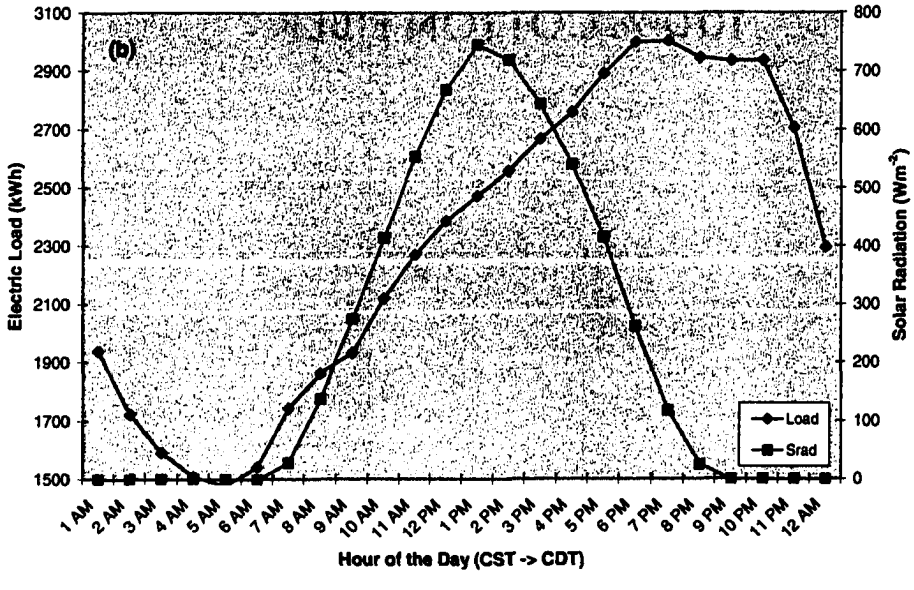
However, the seasonal temporal correlations (Table 5.1a) and seasonal line graphs (Figs. 5.7a-d) gave rise to the appropriate lags needed to obtain the strongest relationship between solar radiation and load. During the winter, solar radiation peaked at 1 PM and the load peaked at 8 PM (Fig. 5.7a); thus, they were best correlated when the load lagged solar radiation by 7 hours. The springtime correlations were highest when the load lagged solar radiation by 5 hours, as the load peaked at ~6 PM and the solar radiation peaked at 1 PM (Fig. 5.7b). The summertime correlations between load and solar radiation were the strongest of all seasons. Figure 5.7c revealed a peak load that occurred at ~6 PM and a solar radiation peak at 2 PM[7] when the entire 1999 season was considered. Thus, the lag between load and solar radiation was about ~4 hours.

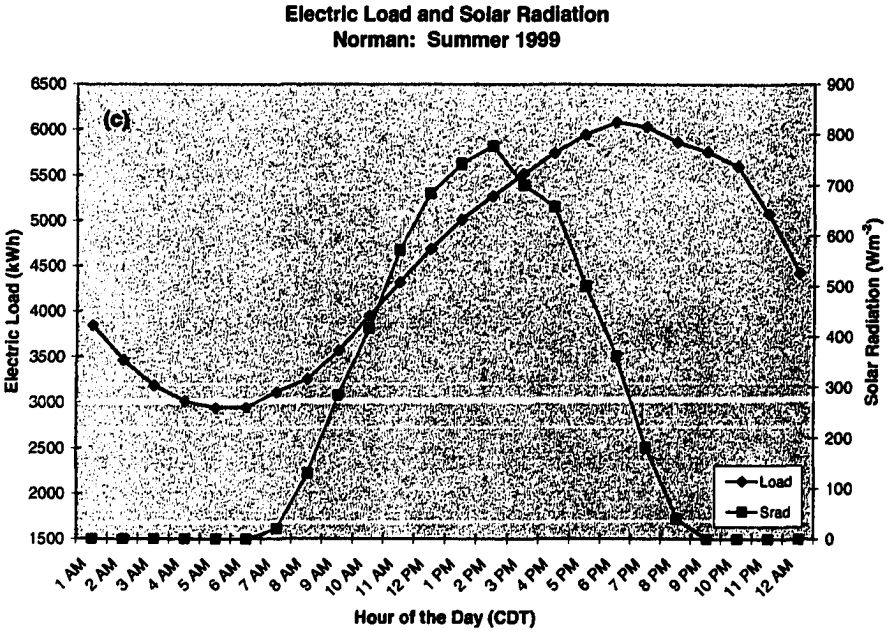
[7] Because hourly solar radiation data were used (as opposed to data on a smaller time scale), solar noon appeared to occur an hour later during the summer months.

Electric Load and Solar Radiation
Norman: Winter 1999



Electric Load and Solar Radiation
Norman: Spring 1999





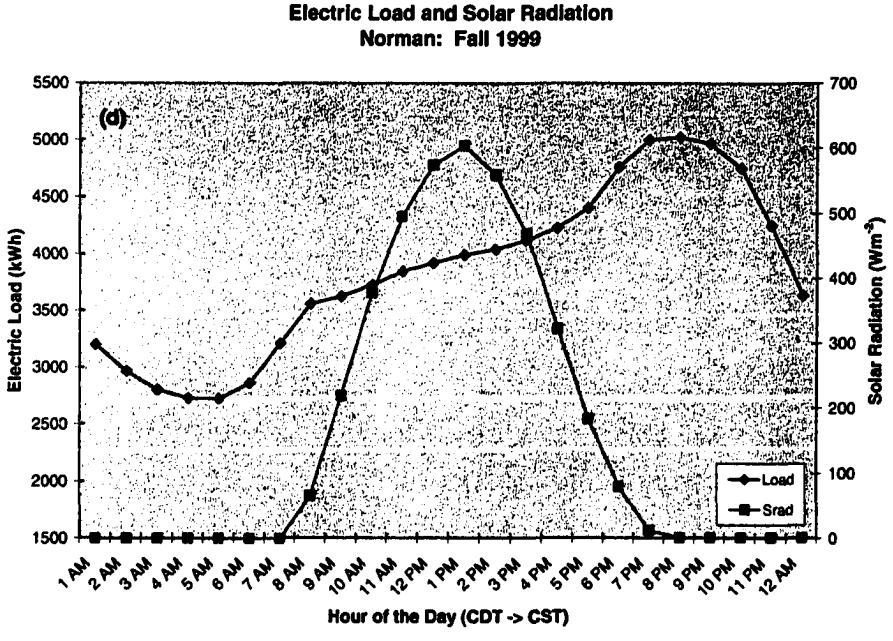


FIG. 5.7 Seasonally-averaged hourly electric load and hourly solar radiation from the West Norman substation and the Norman Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

Surprisingly, the relationship between solar radiation and load was the *strongest* relationship uncovered during the fall season ($\rho = 0.56$) among all weather variables correlated with load (Table 5.1a). Evidently, the diurnal cycle of solar radiation is more strongly correlated with load than with temperature during the fall season at Norman. The load peaked between 7 PM and 8 PM, while the solar noon occurred at 1 PM (Fig. 5.7d) during the fall season. The correlation coefficients were highest when the load lagged solar radiation by ~6 hours.

The relationship between solar radiation and electric load is not trivial, thus the lag factor (by season) must be used with care for the data to be useful to load forecasters. Overall, temperature has proven to be the most valuable weather variable for predicting load. However, the fall season has proven to be the most difficult season for load forecasting because the correlation between load and weather is weak. Thus, the use of solar radiation data could provide valuable information to load forecasters, especially during fall season.

5.2.4 Wind Speed-Load Relationship (Norman)

The analysis of the wind speed-load relationship did not produce strong relationships during any season. Accordingly, the correlation (not shown) between variables was inconsistent and was not apparent. While the wind chill value (e.g., a temperature measure calculated using air temperature and wind speed) seemed significant in defining the apparent temperature-load relationship during the winter seasons, the independent use of wind speed appeared to be of no value to load forecasters. All sites

revealed the same relationship between wind speed and load, and thus wind speed will no longer be considered in this investigation.

5.3 Woodward Site

Like Norman, Woodward is primarily a residential community. Unlike Norman, however, Woodward is much more rural in nature. While the daily activities for most residential communities are similar, a few differences between Norman and Woodward include: (1) differences in the geographic location of Norman and Woodward; (2) differences in how each community heats its homes; and (3) differences in insulation between older and more modern homes. Selections from the data set were used for illustration purposes to represent the entire data set.

5.3.1 Temperature-Load Relationship (Woodward)

5.3.1.1 Scatterplots

On an annual scale, the relationship between temperature and load at Woodward is very nonlinear (Fig. 5.8), exhibiting a 'U-shaped' appearance. Seasonal data (Figs. 5.9a-d) revealed a few characteristics that distinguished the rural residential from the urban residential sites.

Although the winter of 1999 was mild, the winter load consumption for Woodward (Figs. 5.8 and 5.9a) averaged ~50% more per hour (e.g., 2400 kWh) than occurred at the Norman site – even though the Woodward substation served only 20% more customers (Table 3.3). One reason that more electricity was used in Woodward resulted from its location in northwest Oklahoma (Fig. 3.1) where the climate is cooler

than near Norman (Figs. 4.2-4.5). Another reason Woodward used more electricity for heating resulted from the fact that natural gas is not readily available throughout Woodward and some customers appear to favor electricity over natural gas when a choice was available. A third reason for greater use of electricity is that homes in rural Woodward may not be as well-insulated as are the more modern homes that dominate Norman[8] (the third fastest growing city in Oklahoma). Any combination of these factors contributed to a greater load demand in Woodward during the winter. Like Norman, the overall temperature-load relationship for Woodward during the winter months was nonlinear and inversely related.

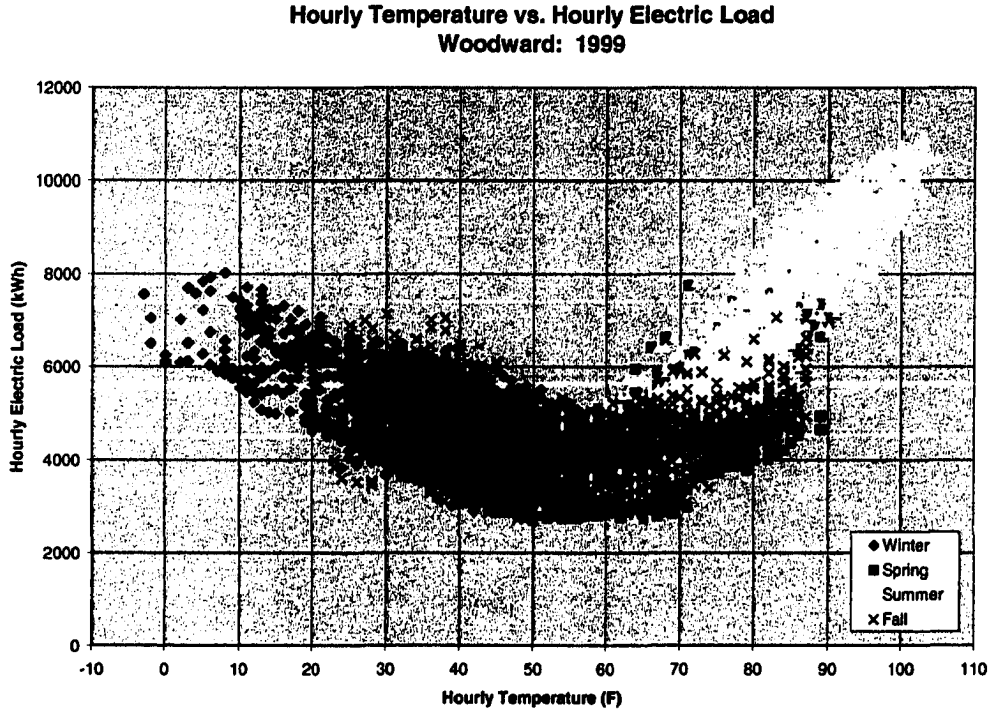


FIG. 5.8 Scatterplot of hourly temperature from the Woodward Mesonet site versus hourly electric load from the Woodward substation for 1999.

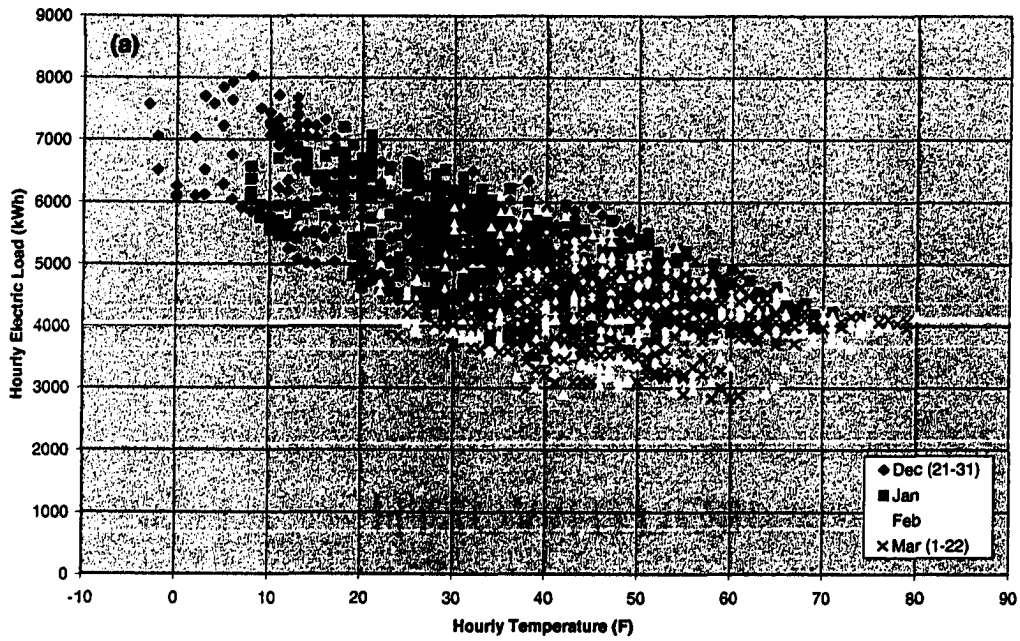
[8] Email communication, Oklahoma Natural Gas, 2002

The spring season in Woodward (Figs. 5.8 and 5.9b) was transitional in the sense that the temperature-load relationship reflected a winter pattern when (colder) early spring temperatures prevailed and a summer pattern when (warmer) late spring temperatures were dominant. This blend created a nonlinear relationship between temperature and load (as with Norman) during the spring.

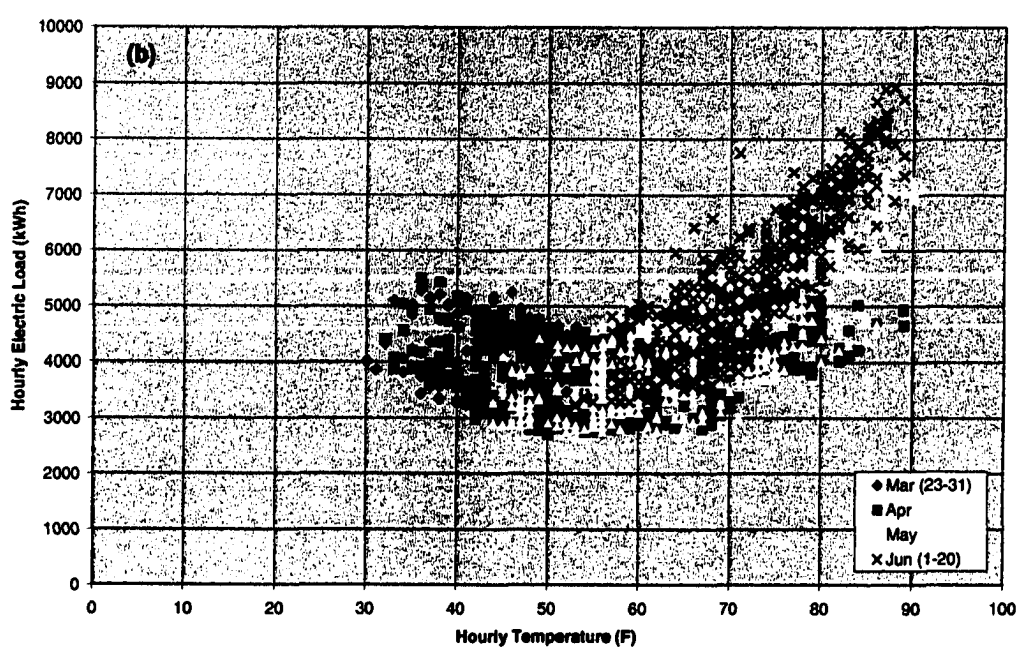
The tightly packed cluster of summertime load values for the Woodward site (Figs. 5.8 and 5.9c) illustrated a strong, positive and linear relationship between temperature and load. This relationship was stronger for the Woodward site than that discovered for Norman. Load values during the summer months in Woodward were 30% greater than those observed in Norman, even though Woodward was cooler than Norman by an average of only 1.2°F (0.67°C) cooler than Norman. A reason for this disparity likely resulted from the fact that houses in Norman are better-insulated on average than those in Woodward.

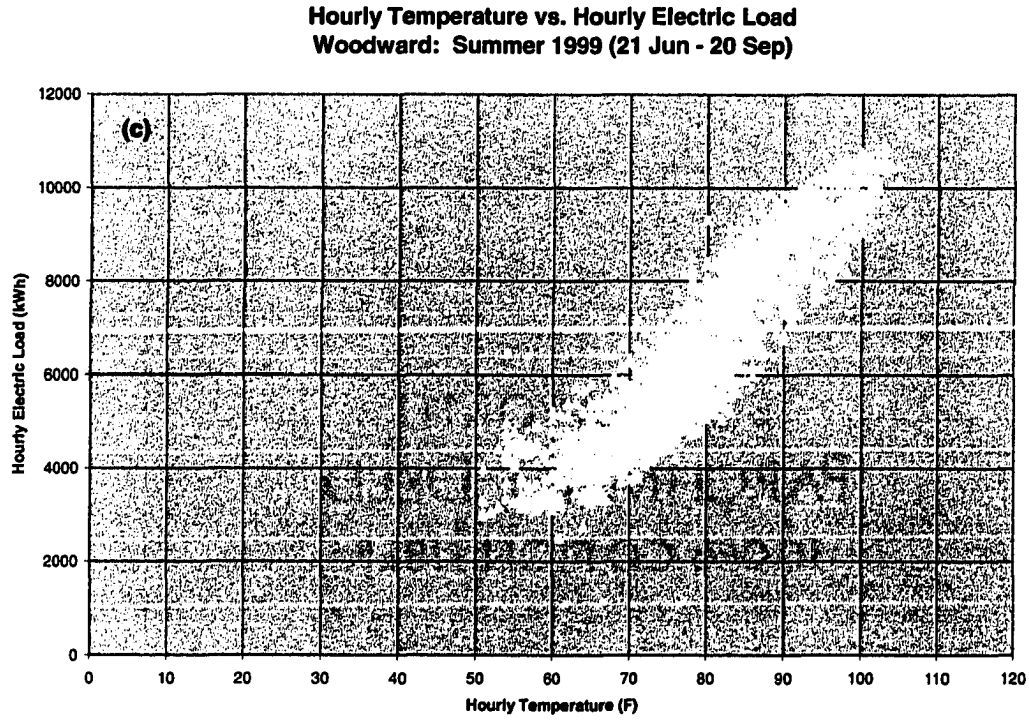
The 'U-shaped' appearance of the scatterplot during the fall season for Woodward (Figs. 5.8 and 5.9d) was very similar to that from Norman. Because the fall is a transitional season, a highly nonlinear relationship between temperature and load was created by warm temperatures (hence the need for cooling) during the early fall and cold temperatures (hence the need for heating) during the late fall. However, colder temperatures during the fall season in Woodward created essentially the same load demand that occurred during winter, in stark contrast to the load profile from Norman. Norman's fall season created a higher load demand than occurred during winter, even when the same temperatures occurred.

Hourly Temperature vs. Hourly Electric Load
Woodward: Winter 1999 (21 Dec 98 - 22 Mar 99)



Hourly Temperature vs. Hourly Electric Load
Woodward: Spring 1999 (23 Mar - 20 Jun)





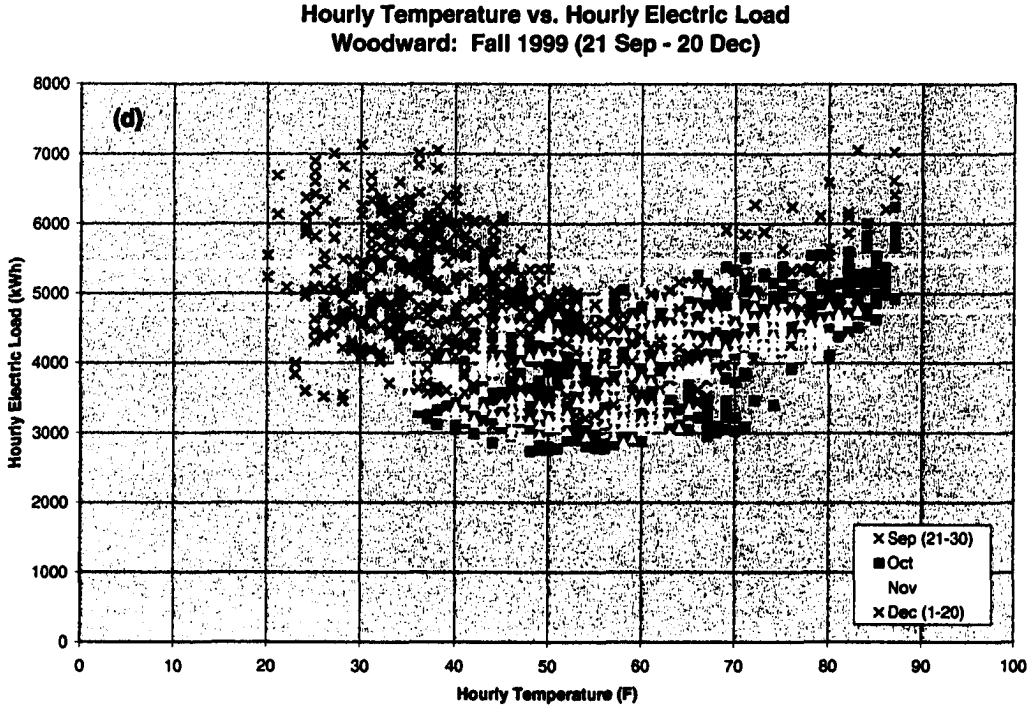


FIG. 5.9 As in Fig. 5.8 except for (a) winter, (b) spring, (c) summer, and fall of 1999.

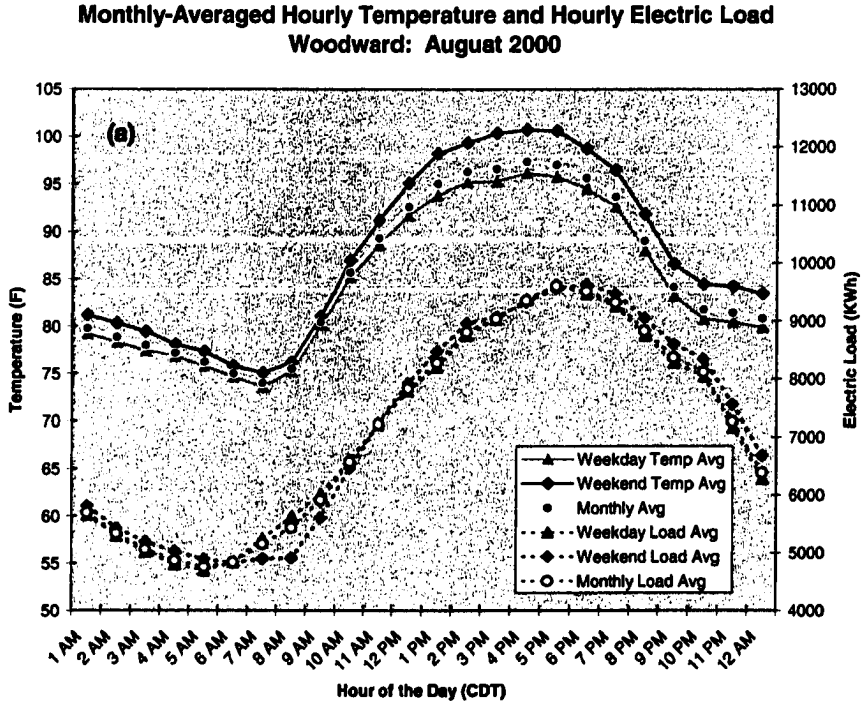
It appeared that ~57°F (13.9°C) was a threshold temperature when the load demand changed for Woodward, a cooler threshold than determined for the Norman site (i.e., the slope of clusters on the scatterplot changed at the threshold temperature).

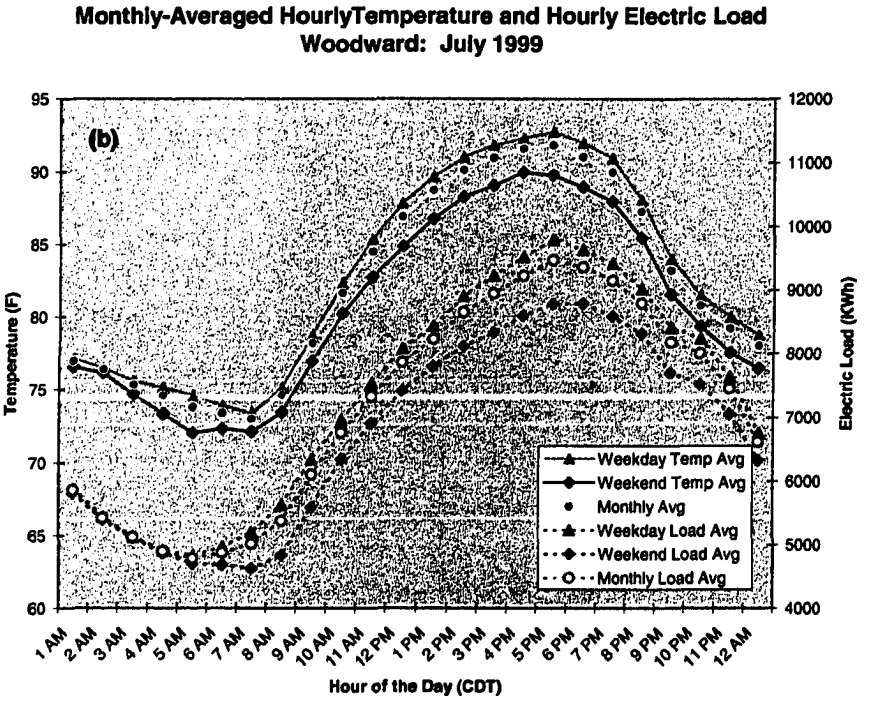
5.3.1.2 Line Graphs

Data were analyzed on a monthly scale to gain more insight into the temperature-load relationship for a rural, residential community like Woodward. The structure of the line graphs for the summer in Woodward (Fig. 5.10a) mimicked those from Norman, in that a single evening peak was observed. However, the evening peak for Woodward occurred at 5 PM, while the Norman peak occurred at 4 PM. The temperature led the load by ~1 hour at Woodward during the summer months (versus ~2 hours in Norman). The shorter lag time required for temperatures to impact load consumption can be attributed to (but not limited to) the lack of well-insulated homes in Woodward relative to Norman.

The weekend load was as much as ~13% *less* than the weekday load during the morning hours for Woodward in August 2000 (Fig. 5.10a). From noon to 3 PM and from 6 PM to midnight, the weekend load was slightly *greater* than (e.g., by < 3% at any given hour) the weekday load. The weekend *temperatures* were greater than those that occurred on an average weekday in Figure 5.10a. (This difference occurred during all August months of the study period). The greater weekend temperatures increased the afternoon load consumption as use of air conditioning increased. Thus, if the weekend and weekday temperatures had been equivalent, the weekend load may have been less, implying that Woodward residents consume less electricity on average during the

weekends than they do on weekdays. Unfortunately, identical weekend and weekday temperatures did not occur during a summer month of the study period. However, the weekend temperatures were *lower* than the weekday temperatures during July 1999 (Fig. 5.10b). During this July, the weekend load was significantly lower than the weekday load. If temperature is not considered, one might infer that Woodward residents were home less on the weekends than on weekdays. When temperatures were considered, and given that the citizens of Woodward were home just as much or even more on weekends versus on weekdays, the decreased use of electricity on the weekends was a response to lower weekend temperatures. Thus, on a monthly scale, a difference in temperatures did not change the shape of the load profile but did influence the amount of load consumed (i.e., shifted the load profile up or down on a graph).





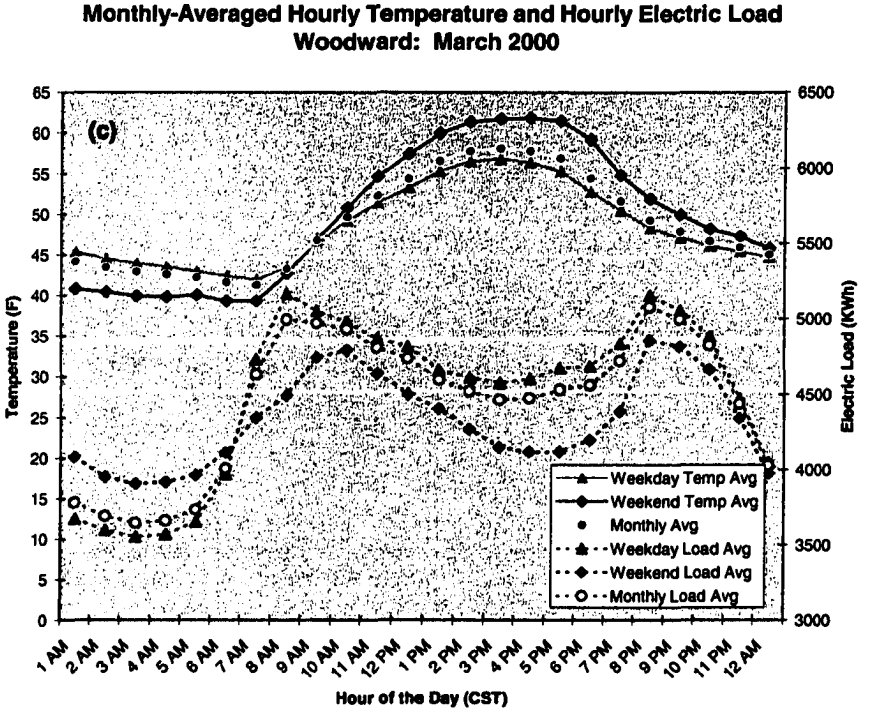


FIG. 5.10 Diurnal plot of hourly temperature and hourly electric load from Woodward averaged during weekdays, weekends/holidays, and a month for (a) August 2000, (b) July 1999, and (c) March 2000.

The winter, spring, and fall seasons exhibited a bimodal distribution in their load profiles at Woodward just as they did in Norman. Figure 5.10c illustrates this load pattern, which was most prominent during the weekdays. However, the early morning peak load in Woodward was equal to or greater than the evening peak (mostly during the winter months). On the other hand, in Norman, the morning load was always less than the evening load. The seasonally-averaged minimum temperature (just before sunrise) in Woodward was always colder than it was in Norman (both observed during the study period and climatologically). Thus, in response to colder temperatures and a greater electrical heating demand, a greater morning peak occurred in Woodward than occurred in Norman.

Weeklong profiles were created to represent each season (Figs. 5.11a-d) to study the temperature-load relationship on an even smaller time scale. During the winter season, an inverse relationship between temperature and load was observed (Fig. 5.11a). The secondary load minimum, observed at ~4 PM, occurred within 1-2 hours of the afternoon high temperature, but only 1 hour prior to most residents returning home from daily activities. The primary minimum in load occurred between 2 AM and 3 AM when most residents were asleep and 1-2 hours prior to the normal morning minimum temperature. This lull in load consumption was more a result of customer habits than the impact of air temperature. This weekly plot from winter revealed a bimodal structure observed in the seasonally-averaged winter load profile (Fig. 4.15).

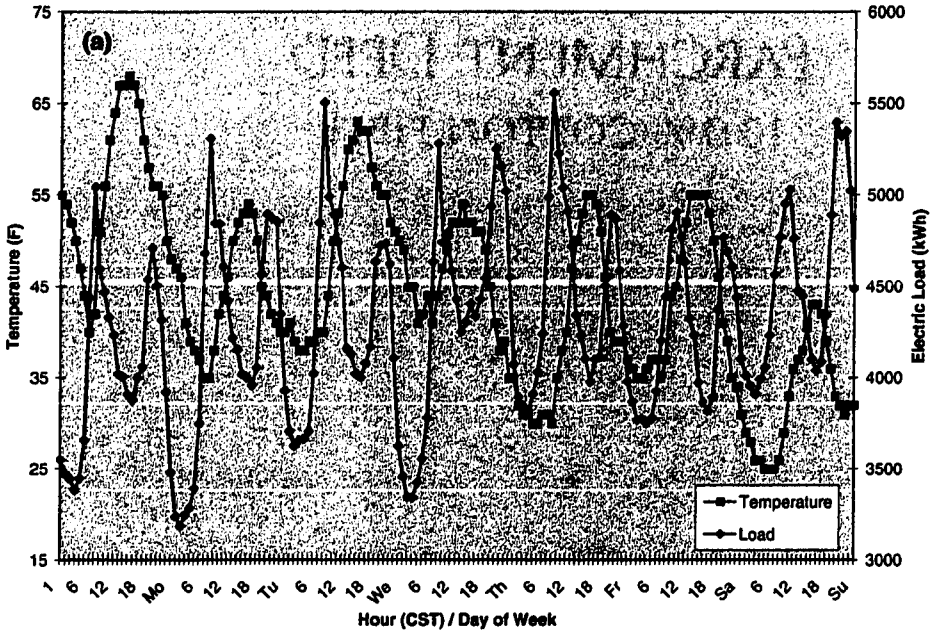
A direct relationship between temperature and load in the weekly plots is clearly evident during the spring season (Fig. 5.11b). The maximum temperature and peak load occurred at the same time each day or within 1 hour of each other (i.e., between 4 PM

and 5 PM) during the spring months. However, when summer-like temperatures occurred during spring season, the early morning (weekday) mode was suppressed and the load profiles assumed a single-mode (summertime) appearance. During the spring week in May 1999 (Fig. 5.11b), a weak bimodal distribution was observed on Monday-Thursday and was mostly damped by Friday. The load profiles on Saturday and Sunday did not contain a morning peak because the morning spurt of electricity does not occur on weekends.

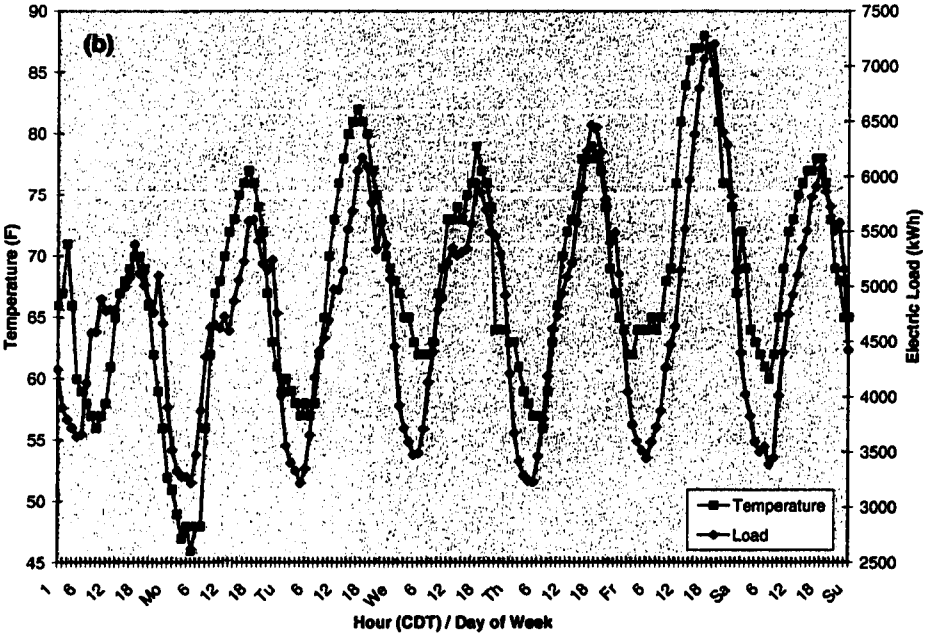
The summer season exhibited a direct temperature-load relationship as did most of the spring season. The one-week plot from the summer of 1999 (Fig. 5.11c) revealed a single-peak structure in the load that occurred at ~5 PM, one hour after the daily maximum temperature. Thus, at the Woodward site during the summer, the load peak lagged the temperature peak by 1 hour, a result consistent with features in the monthly summer load profiles (Figs. 5.10a-b).

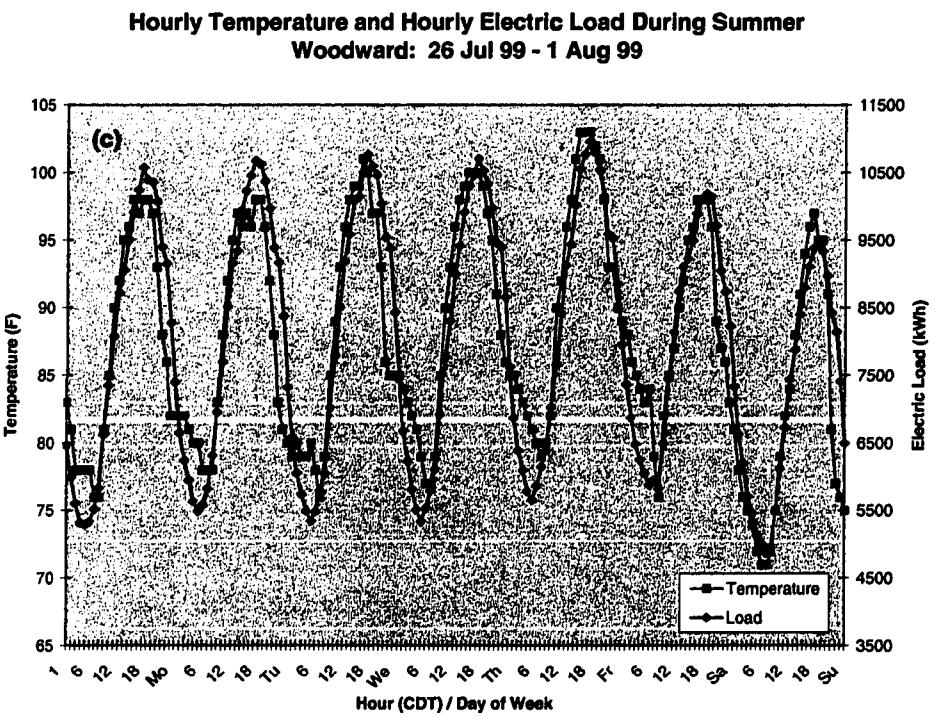
With the arrival of fall, an inverse relationship developed between temperature and load (Fig. 5.11d). Temperatures and load were somewhat chaotic during the fall season, as observed at the Norman site. For example, the maximum daily temperature decreased from 80°F (26.7°C) to 62°F (16.7°C) in one day. As a result, the afternoon (secondary) minimum load occurred about 1 hour prior to the daily maximum temperature, a feature which nicely illustrates the inverse relationship.

Hourly Temperature and Hourly Electric Load During Winter
Woodward: 15 Feb 99 - 21 Feb 99



Hourly Temperature and Hourly Electric Load During Spring
Woodward: 17 May 99 - 23 May 99





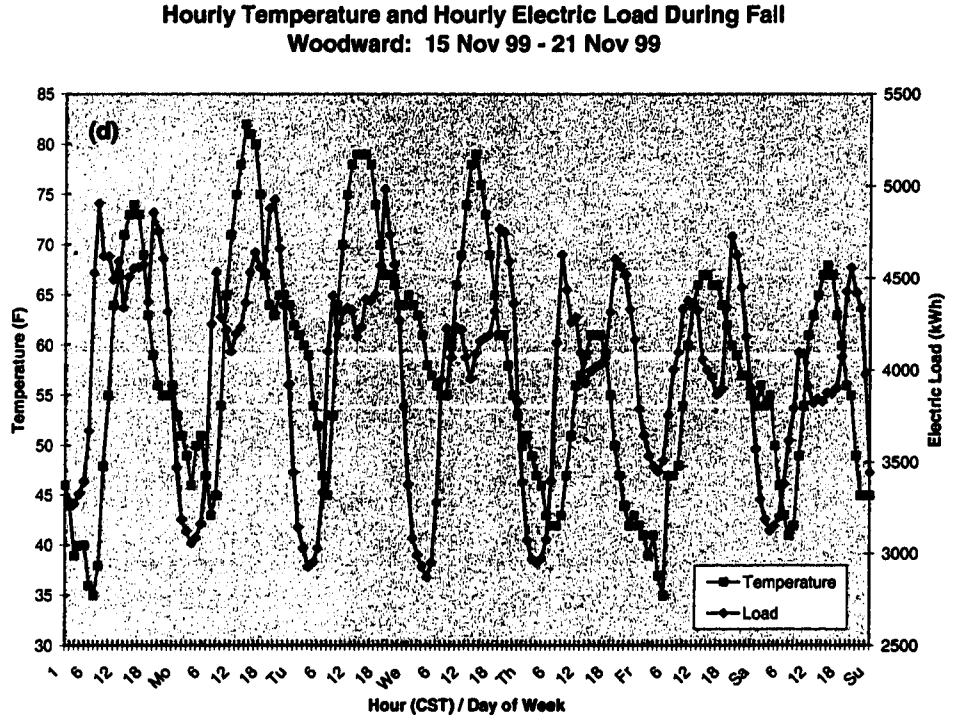


FIG. 5.11 Hourly temperature and hourly electric load from Woodward during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

5.3.1.3 Apparent Temperature-Load Relationship

The relationship between T_{app} and electric load was not significantly different from the temperature-load relationship. Differences were not observed in any scatter plot or line graph on any time scale. While apparent temperature could be used in lieu temperature in the estimating stage of load modeling, the evidence does not suggest that apparent temperature would produce a more accurate load forecast than temperature.

5.3.1.4 Temporal Correlations

Correlation coefficients were calculated (using 0-12 hour lags) to quantify the temperature-load relationship and the apparent-temperature load relationship for the Woodward site (Table 5.1b). The strongest correlations during the study period at Woodward occurred during the summer ($\rho = 0.91$), when the temperature led the load by 1 hour. The Norman site experienced the strongest correlation with a 2-hour lag during the summer ($\rho = 0.87$). A strong, positive relationship indicated that an increase in temperature implied an increase in load consumption. The T_{app}-load relationship also was strongest during the summer at a lag time of 1 hour ($\rho = 0.91$), but the correlation values were weaker than those from the temperature-load relationship.

The next highest correlation occurred during the spring months ($\rho = 0.65$), except that the demand was immediately responsive to changes in temperature (i.e. lag time of zero). The temperature-load relationship was stronger than the T_{app}-load relationship. In addition, zero lags and 1-hour lags produced nearly the same correlations (Table 5.1b). Both the winter and fall seasons experienced a negative (positive) correlation between load and temperature which meant that the maximum temperature was correlated with the

afternoon minimum (maximum) load. The T_{app}-load relationship mirrored the temperature-load relationship during the two seasons. The winter and fall relationship were strongest when ~2 hours was allowed for the load to respond to temperature. However, the fall correlation coefficients were the weakest among of all the seasons (ρ = -0.25). Thus, the graphical results and correlation results support the following relationships in rural Woodward: (1) use of an apparent temperature did not add new information to the temperature-load relationship; (2) the temperature-load relationship was strongest during the summer and weakest during the fall; (3) temperature led the load by ~2 hours during the winter and the fall seasons; (4) the response time of the load to temperature during the summer was reduced to 1 hour; and (5) an immediate load response to temperature was observed during the spring season.

5.3.2 Relative Humidity-Load Relationship (Woodward)

Figure 5.12 illustrates the relationship between relative humidity and load during 1999; it is representative of the entire study period for Woodward. During the winter season, a slight, positive relationship existed between relative humidity and load (i.e., as the relative humidity increased, so did the load). Perhaps cold, damp conditions during winter increased the heating demand. During the transitional seasons (e.g., spring and fall), the scatter was so widespread that the load seemed almost independent of relative humidity. During the summer months, the load had an inverse relationship with relative humidity. Although Norman was slightly more humid than Woodward, the same (inverse) *temperature*-relative humidity pattern occurred (Fig. 5.13), resulting in an increase in load consumption.

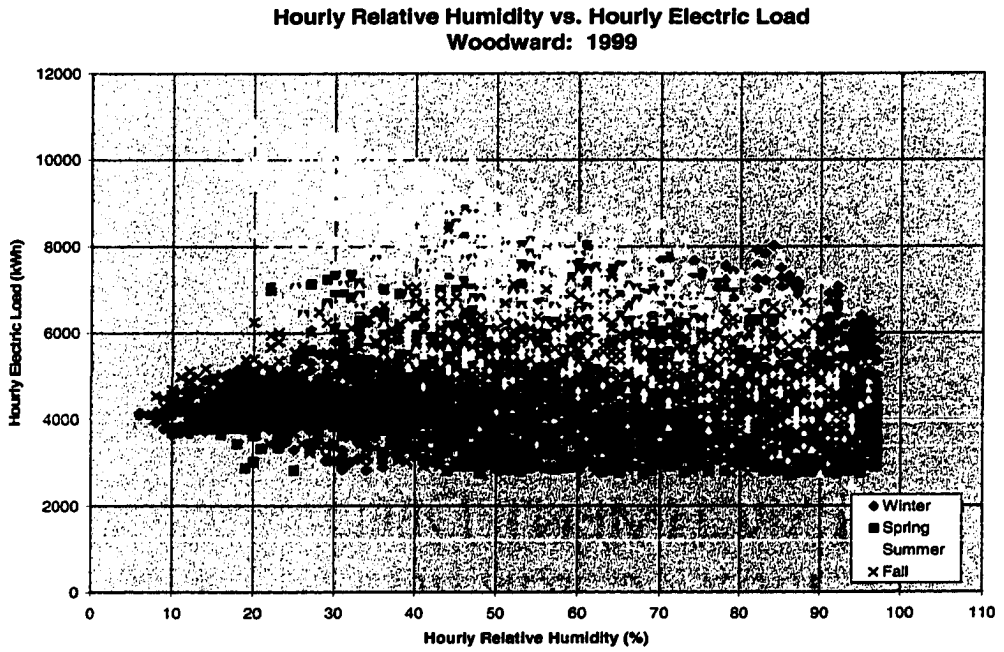


FIG. 5.12 Scatterplot of hourly relative humidity from the Woodward Mesonet site versus hourly electric load from the Woodward substation for 1999.

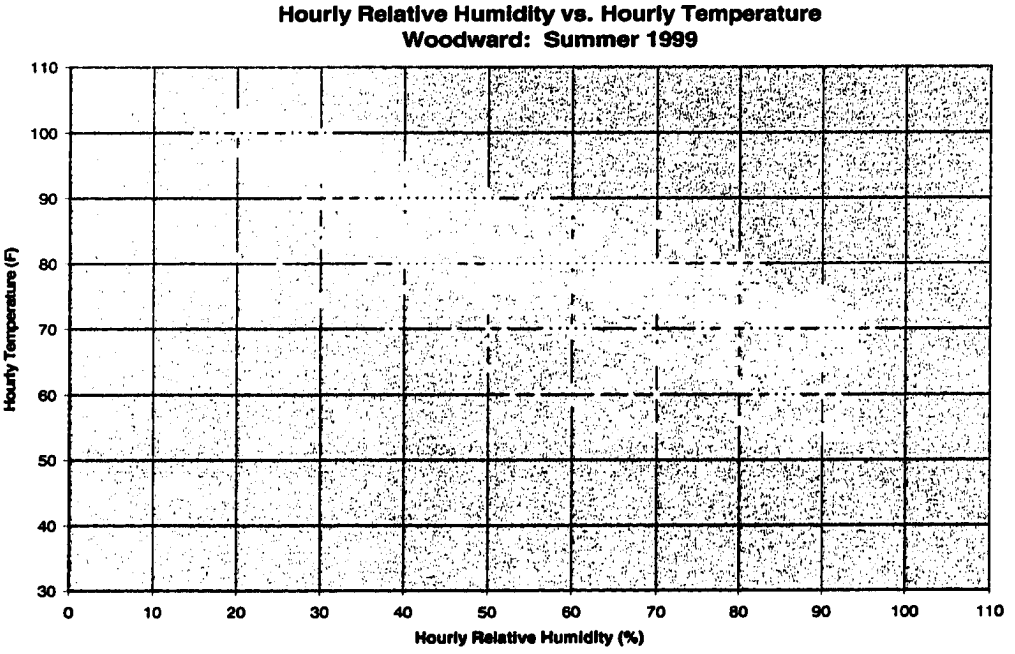


FIG. 5.13 Scatterplot of hourly relative humidity versus hourly temperature from the Woodward Mesonet site for 1999.

The temporal correlations in Table 5.1b quantify the results above. During the winter season, the relationship between relative humidity and load was positive; it was stronger when the response time of the load to relative humidity was 2 hours. However, the relative humidity-load relationship was weak ($\rho = 0.29$). During the spring, summer, and fall seasons, the correlations were negative. The strongest relationship ($\rho = -0.69$) occurred during the summer with an immediate load response to changes in relative humidity (i.e., lag time of zero) versus a 1-hour lag detected in Norman. The weakest relationship ($\rho = -0.15$) occurred during the fall. Thus, use of relative humidity as a predictor for electric load modeling may be most valuable during the summer months.

5.3.3 Solar Radiation-Load Relationship (Woodward)

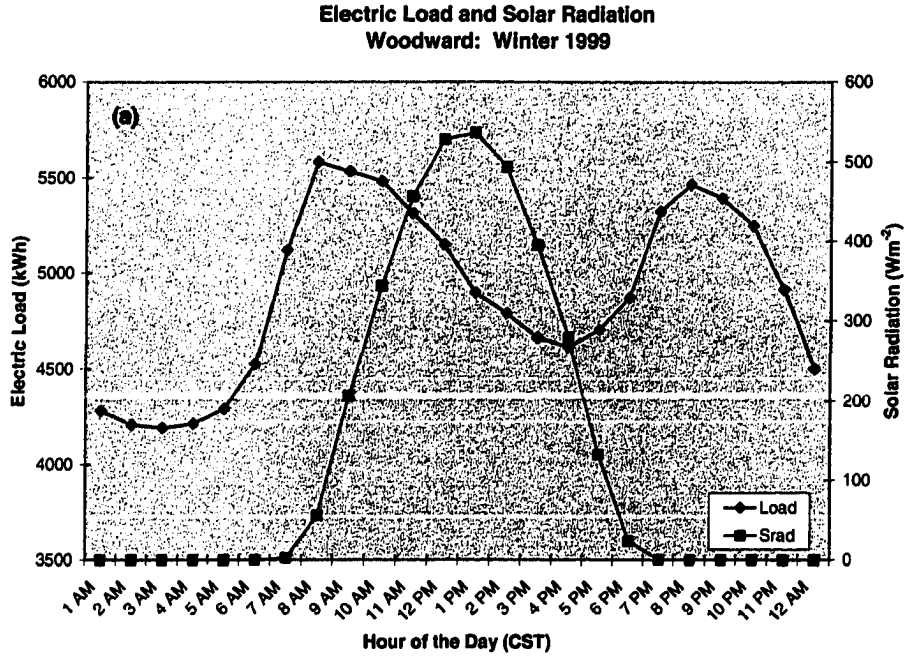
The relationship between solar radiation and load for Woodward was slightly stronger ($\rho = 0.79$) than observed in Norman, but only during the summer season. Correlations were positive for all seasons except winter. Line graphs (Figs. 5.14a-d) created to assess the lag between the two data streams and the correlation coefficients (Table 5.1b) quantified the strength of this relationship.

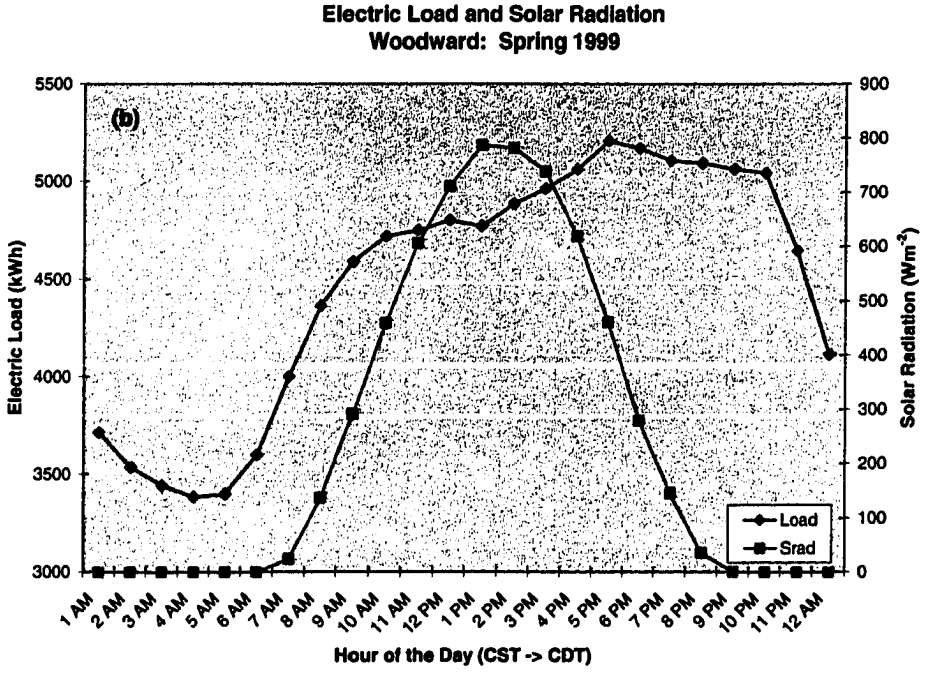
As observed in Norman, solar noon occurred at 1 PM during the winter, spring, and fall seasons and at 2 PM during the summer season. The solar radiation maxima occurred between the primary and secondary peak loads during the winter, spring, and fall and occurred hours prior to the single peak load during the summer (Fig. 5.14a-d). During the winter season, the load profile peaked twice per day at nearly the same magnitude: once at 8 AM and once at 8 PM. As a result, a secondary load minimum occurred during the afternoon (e.g., at ~4 PM). The strongest correlation between solar

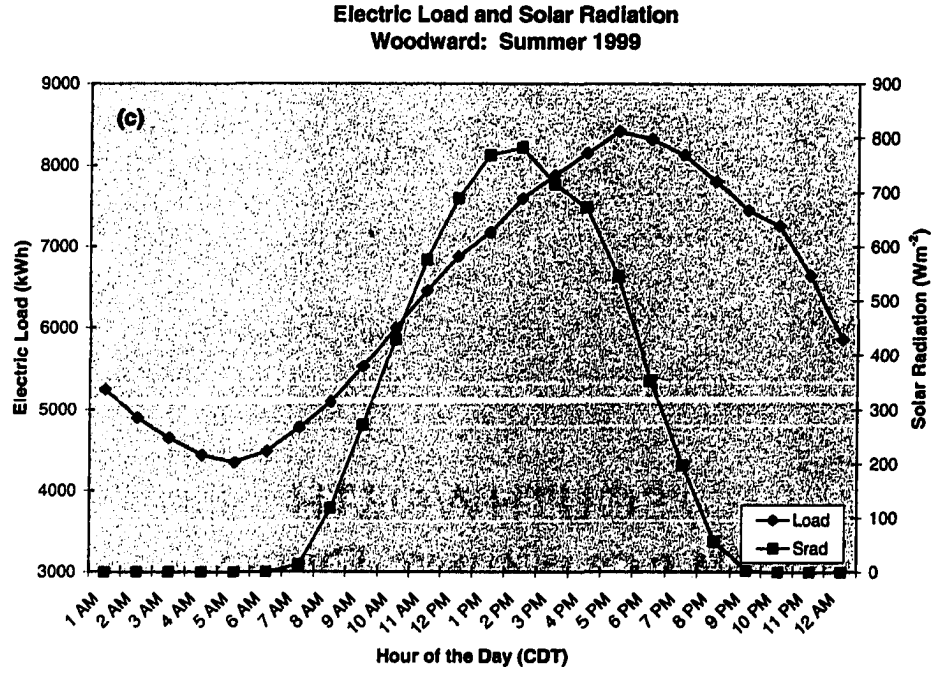
radiation and load was negative because a maximum of solar radiation preceded a minimum of electrical use by 4 hours (Fig. 5.14a and Table 5.1b). Although the correlation was weak, the relationship became weaker when trying to correlate the peak load and the solar radiation maximum with the appropriate time lag.

The interrelationship during the spring season was strongest when the maximum of solar radiation preceded the maximum load by 4 hours ($\rho = 0.53$), the number of hours between the load peak at 4 PM and solar noon (Fig. 5.13b, Table 5.1b). Similar results were uncovered from the summer season (Fig. 5.13c). The strongest summer relationship ($\rho = 0.79$) was detected when solar noon led the peak load by 3 hours, the number of hours between the load peak (e.g., 5 PM) and solar noon. In general terms for all seasons, the load always responded to the weather variable 1 hour sooner in Woodward than detected in Norman.

During the fall season for Woodward, the physical relationship for most variables were weaker than those detected in Norman ($\rho = 0.29$ versus $\rho = 0.56$). On the other hand, the magnitude of the solar radiation-load correlation for Woodward was stronger ($\rho = 0.3$) than the temperature-load correlation ($\rho = -0.25$), which indicated that solar radiation contains more predictive power than the widely used and accepted temperature data during the fall season.


Electric Load and Solar Radiation
Woodward: Winter 1999
(a)
Electric Load (kWh)
Solar Radiation (Wm-2)
6000
5500
5000
4500
4000
3500
600
500
400
300
200
100
0
Load
Srad
1 AM
2 AM
3 AM
4 AM
5 AM
6 AM
7 AM
8 AM
9 AM
10 AM
11 AM
12 PM
1 PM
2 PM
3 PM
4 PM
5 PM
6 PM
7 PM
8 PM
9 PM
10 PM
11 PM
12 AM
Hour of the Day (CST)


Electric Load and Solar Radiation
Woodward: Spring 1999
(b)
Electric Load (kWh)
Solar Radiation (Wm-2)
5500
5000
4500
4000
3500
3000
900
800
700
600
500
400
300
200
100
0
Load
Srad
1 AM
2 AM
3 AM
4 AM
5 AM
6 AM
7 AM
8 AM
9 AM
10 AM
11 AM
12 PM
1 PM
2 PM
3 PM
4 PM
5 PM
6 PM
7 PM
8 PM
9 PM
10 PM
11 PM
12 AM
Hour of the Day (CST -> CDT)



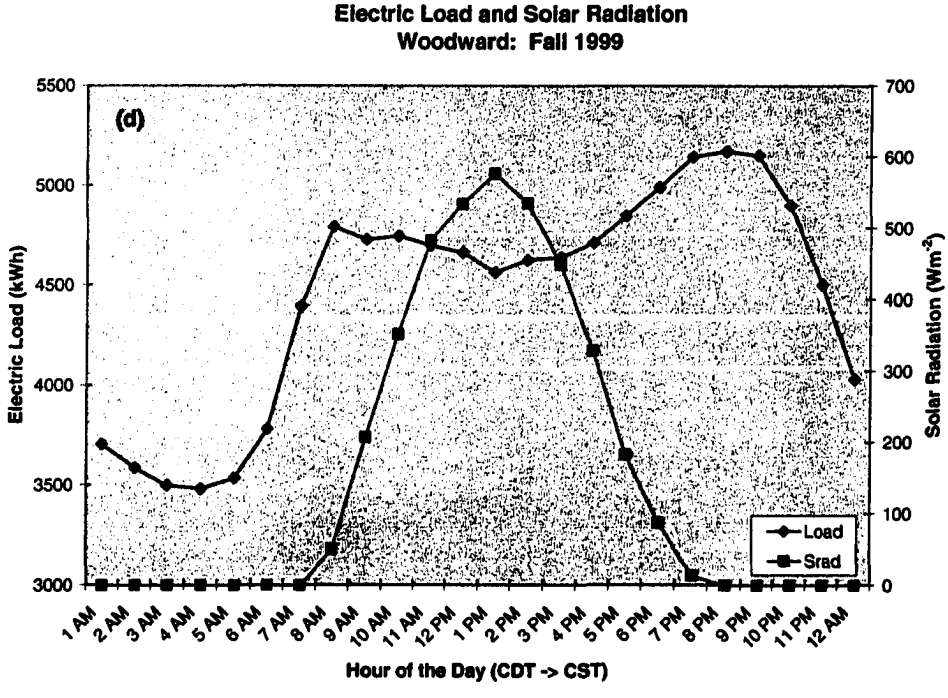


FIG. 5.14 Seasonally-averaged hourly electric load and hourly solar radiation from the Woodward substation and the Woodward Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

5.4 Altus AFB Site

The Altus AFB site is a commercial entity. This training facility operates on a Monday-Friday schedule made up of 8-hour workdays. However, some military personnel live on the base and it is that residential use of electricity which sometimes skews the use of electricity outside the workday. These relationships between weather variables and electric load are examined below.

5.4.1 Temperature-Load Relationship (Altus AFB)

5.4.1.1 Scatterplots

On an annual scale, a nonlinear relationship existed between temperature and load at Altus AFB (Fig. 5.15). The trends in the plot were such that, on either side of ~61°F (16.1°C), the temperature-load relationship had different characteristics. On the cold side of the scatterplot, the relationship was linear. More importantly, the load did not appear to have a dependence on temperature (i.e., a horizontal band of load values spanned temperatures between 13°F/-10.6°C and 50°F/12.8°C). On the warm side of the plot, the relationship appeared linear with an orientation that defined a positive relationship (i.e., as the temperature increased, the load increased).

Seasonal scatterplots (Figs. 5.16a-d) revealed more details about the piecemeal linear relationship between temperature and load. Temperatures during the winter of 1999 ranged from 13°F (-10.6°C) to 80°F (26.7°C) while the load varied between 5000 kWh and 9000 kWh (Fig 5.16a). Because the load appeared independent of temperature, other factors must have influenced the small band of electrical load to vary across the wide range of temperatures. One factor is the fact that Altus AFB uses natural gas

(versus electricity) as its primary source of heat. Thus, a decrease in winter temperatures does not necessarily increase the consumption of electricity. Because Altus AFB operates as a business from Monday through Friday, the weekday loads were separated from the weekend and holiday loads (e.g., New Years Day, Martin Luther King, Jr. Day, and President's Day). The weekdays were further stratified by day (8 AM – 8 PM) and by night (9 PM – 7 AM). Although the typical workday at Altus AFB is 8 AM – 4 PM, the 'days' were extended to 8 PM because of the early evening influence on

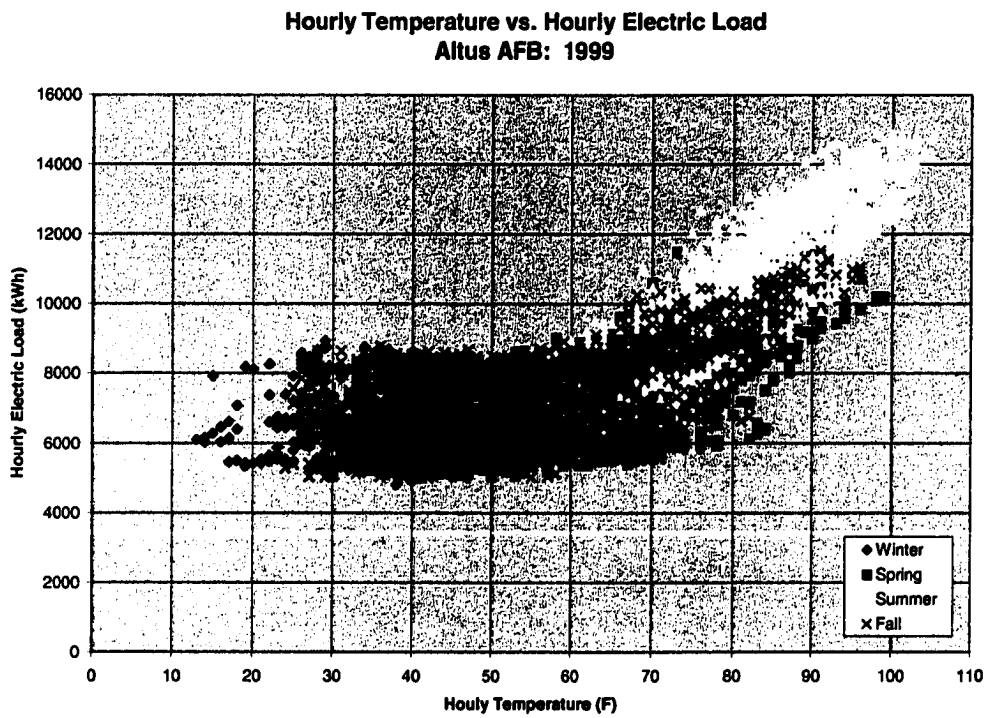
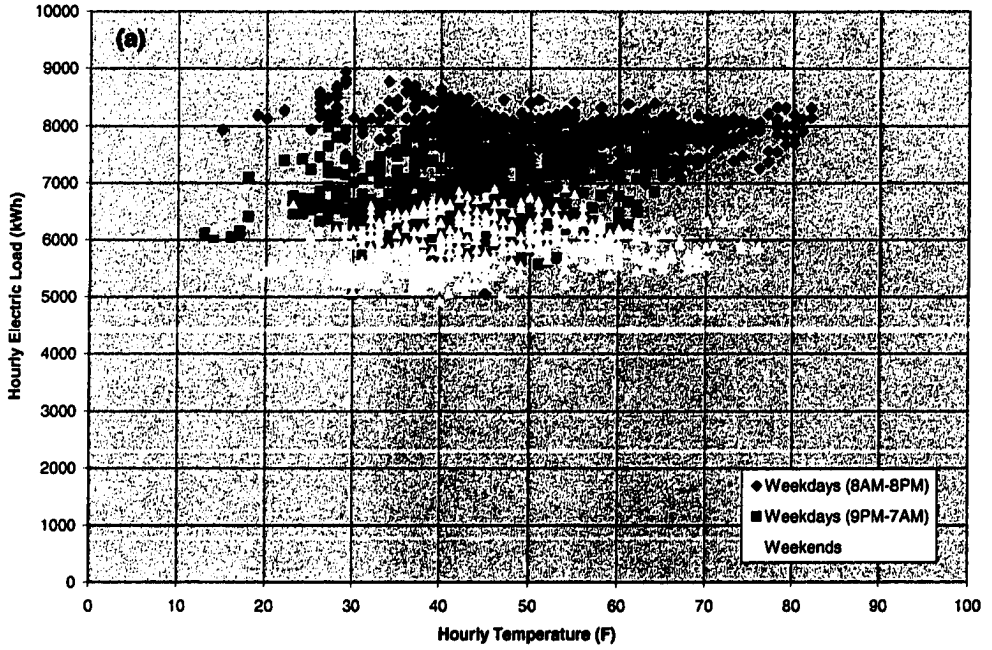
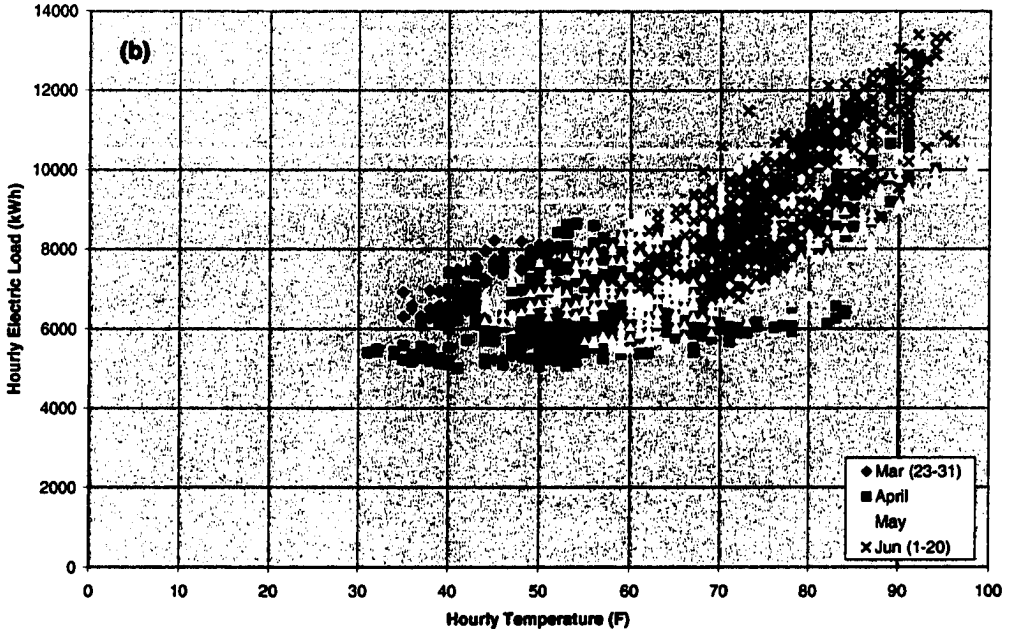


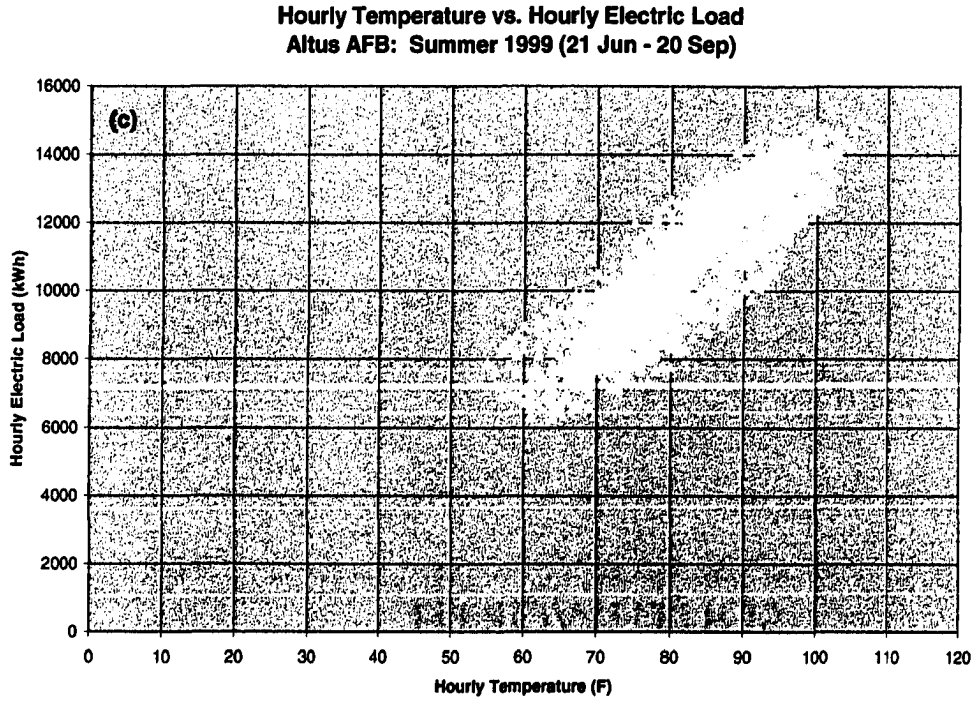
FIG. 5.15 Scatterplot of hourly temperature from the Altus Mesonet site versus hourly electric load from the Altus AFB substation for 1999.

Hourly Temperature vs. Hourly Electric Load
Altus AFB: Winter 1999 (1 Jan 99 - 22 Mar 99)



Hourly Temperature vs. Hourly Electric Load
Altus AFB: Spring 1999 (23 Mar - 20 Jun)





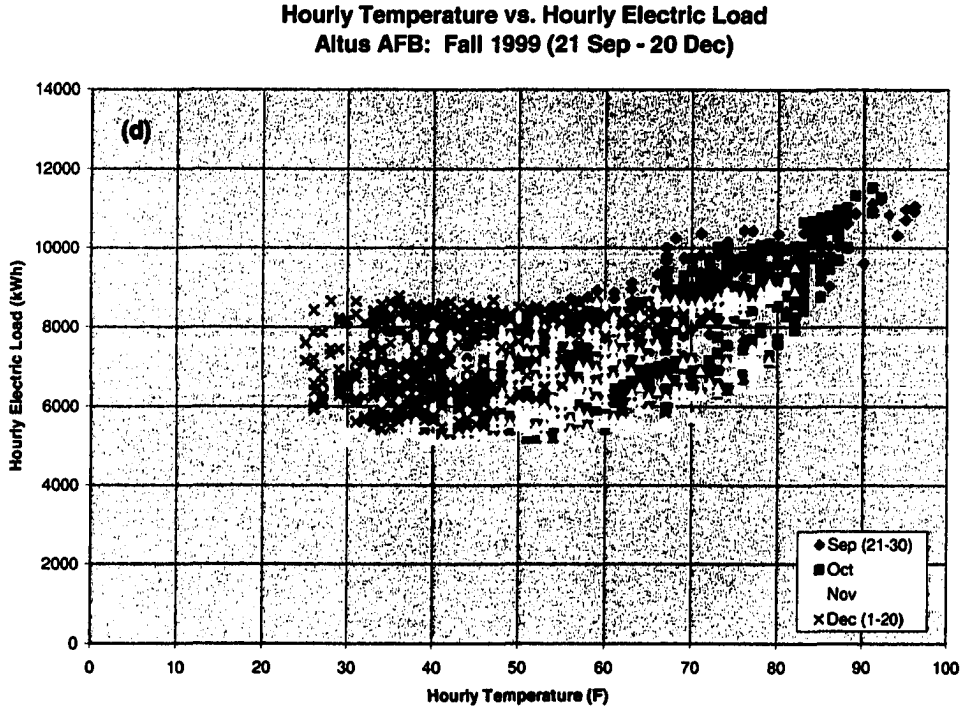


FIG. 5.16 As in Fig. 5.15 except for (a) winter, (b) spring, (c) summer, (d) and fall of 1999.

electric load from base housing. In other words, the use of electricity started to decline after 4 PM, but its use increased sharply between 6 PM and 8 PM as base residents settled in for the evening. After 8 PM, the load decreased. After stratifying weekday-days, weekday-nights, and weekends/holidays, a clear pattern developed in the scatterplot. During the winter, higher loads during the week coincided with the 8-hour workday and evening use of electricity. In contrast, smaller load values corresponded to weekends or to the overnight period. Clearly, the use of electricity to operate the base during the week dwarfed the weekend load consumption, which was created by the housing element.

In a response that was similar to the use of electricity at residential sites, the temperature and load values during the transitional spring season became a combination of winter-like and summer-like responses. An inspection of Figure 5.16b revealed that the March scatter was confined to the cool side of the plot and resembled the horizontal structure observed during the winter (Fig. 5.16a). By June, the relationship became linear and positive on the warm side of the plot as the data acquired a summertime appearance (Figs. 5.15 and 5.16c). The overlap of values was minimal from March to June. A break in the linear/nonlinear relationship appeared to occur at ~61°F (16.1°C). The temperature and load data from April and May were intermingled throughout the seasonal plot, as the spring oscillated back and forth between the cold of early spring and the warmth of late spring. With the arrival of summer, the relationship between temperature and load was strong, positive, linear, and well defined. Clearly, as temperature increased, the consumption of electricity increased.

Because the fall season is one of transition, the temperature-load relationship also represented a combination of the relationships from the summer and winter seasons.

Figure 5.16d exhibited the same nonlinear relationship between temperature and load that occurred during the fall months at the residential sites. As with the annual and spring plots, the 61°F (16.1°C) isotherm appeared to be the breakpoint temperature whereby the temperature-load relationship changed its pattern. The cooler side resembled winter scatter while the warmer side resembled the summer scatter. Based upon a monthly partitioning of the fall data (Fig.5.16d), a pattern emerged that resembled the spring pattern (Fig. 5.16b). The most significant difference between the two plots is that the fall plot contained much colder temperatures than did the spring plot which is a consequence of the fact that the fall season experienced a wide range of temperatures (e.g., a difference of 58°F [32.2°C] between extremes) each month. The large variances of temperature during the fall made it the most difficult season to produce electric loads accurately.

The largest contrast between the temperature-load relationships from a commercial site and from a residential site occurred during the winter season. The images (Fig. 5.16a) clearly revealed that the load consumption at Altus AFB during the winter was driven by the day of the week and by the time of day instead of begin modulated by air temperature. Norman was more responsive than Altus AFB to winter temperatures in that the load increased with colder temperatures (Fig. 5.2a). Woodward, the rural residential site, was the most responsive to winter temperatures, as the load increased significantly with colder temperatures (Fig. 5.9a).

5.4.1.2 Line Graphs

To analyze the impact of temperature on load consumption for a commercial customer, the temperature-load relationship was plotted on a monthly time scale (Figs.

5.17a-b). During the day while residential customers were at work (e.g., not consuming high levels of electricity at their homes), the commercial customer used elevated levels of electricity. But, after the workday ended, the commercial customer experienced a dramatic decrease in the use of electricity even as the residential customer (on the base) increased the use of electricity at home. Thus, the Altus load demand shifted from home to the place of business and back to home during a 24-hour day. Because of base housing, the load consumption between 8 PM and 10 PM had characteristic profiles that resembled residential sites. Furthermore, because residential and commercial customers were relatively inactive during the overnight hours (i.e., between 12 AM – 6 AM), both customers experienced a load minima.

The weekdays and weekends were plotted separately in the line graphs. Because Altus AFB is a weekday operation, the weekday and weekend load profiles were divergent in their shape *and* their magnitude (Figs. 5.17a-b). For Altus AFB, regardless of the temperature, the load decreased after the close of business during the week. Each late afternoon decline in the consumption of electricity was not evident on weekends. For a purely residential customer, the *shape* of the weekend load profile mirrored that of the weekday profile; the difference between the weekday and the weekend load was the magnitude of the load, a feature created by temperature differences between the weekday and weekends.

Load profiles during summer (Fig. 5.17a) distinguished themselves from those during all other seasons at Altus AFB (also observed at residential sites). While in most cases the load trends were different between the commercial and residential customers, the single-peak appearance in the load profile during summer occurred at Altus AFB,

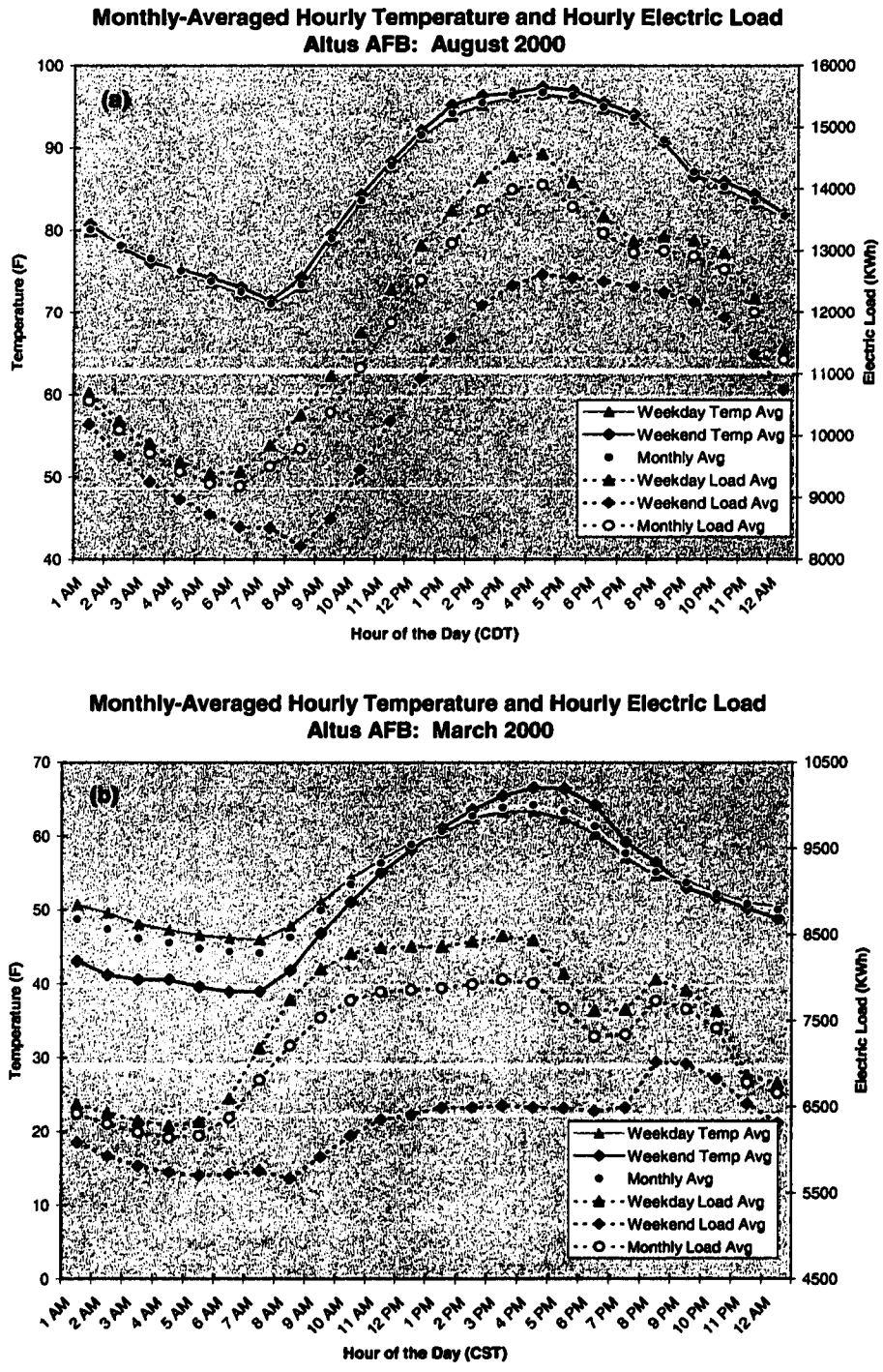


FIG. 5.17 Diurnal plot of hourly temperature and hourly electric load from Altus AFB averaged during weekdays, weekends/holidays, and a month for (a) August 2000 and (b) March 2000.

Norman, and Woodward (Figs. 5.17a, 5.3a, and 5.10a). The differences in the summer profiles between the residential and commercial customers included the time of the peak load and the profile structure several hours after the peak. During the summer, the load and the temperature peaked at ~4 PM. Thus, the load peaked at the end of the workday rather than being modulated totally by the maximum afternoon temperature. However, the magnitude of the peak load *was* influenced by temperature. The summertime peak loads at Altus AFB averaged 12800 kWh, which was ~5400 kWh greater than peak loads during other seasons (Fig 5.17a). After 4 PM, the load curve for Altus AFB decreased quickly until 7 PM. If base housing were not a factor, it is probable that the decrease in load demand would have continued at a constant rate towards the overnight minimum load. However, because of base housing, load consumption increased to a secondary peak (e.g., the same hour of peak loads for residential customers at 8 PM). Between 8 PM and 10 PM, the load consumption decreased gradually. After 10 PM, electricity use at Altus AFB decreased rapidly to its overnight minimum – much like what occurred at a substation serving residential users. Thus, base housing added an evening spurt of electricity consumption to the profile for Altus AFB, which typical commercial entities would not experience.

The weekends during the summer months at Altus AFB produced a slightly different load profile, both in magnitude and shape, compared to the weekday load profile (Fig. 5.17a). The overnight load minimum occurred at 8 AM on the weekends and at 5 AM during the week, whereas the low *temperature* occurred at 7 AM. Because the base operations were closed on weekends, the pattern for the load profile was influenced primarily by base housing. Thus, the load on weekends primarily responded to changes

in temperature, much like residential users, as military residents began their day. The weekend profile peaked at 4 PM (at the same hour as the weekday profile) when the weekend temperature reached its maxima. However, as temperatures decreased, the weekend load decreased gradually during the next several hours, and, thereby, exhibited a very different behavior than observed in the weekday load profile. Thus, during summer, the shape of the weekday load profile was influenced by the workday schedule while the shape of the weekend load profile was influenced by the response of military residents to changes in temperature. The magnitude of both profiles was also influenced by the summertime temperatures.

All seasons (other than summer) exhibited monthly-averaged load profiles like those shown in Figure 5.17b. Between 5 AM and 8 AM, a steep increase in the electric load occurred during the week. This feature reflected how base residents began a morning routine during the weekdays that was like those in residential communities. In addition, military personnel may have trickled into work and increased the use of electricity. After 8 AM, the load profile continued to increase for the commercial entity, but it load decreased in the residential community (Figs. 5.3b and 5.10c). The greatest use of electricity occurred during workday hours (i.e., 9 AM – 4 PM), interrupted by a small decrease in load during the lunch hour (Fig. 5.17b). The peak load occurred at 3 PM during March 2000, ~1 hour *before* the maximum daily temperature. This unexpected scenario implied that temperature did not drive the *timing* of the peak load as it did in residential communities. However, the temperature did impact the *magnitude* of the peak load which averaged 7400 kWh during winter, 9400 kWh during spring, 12800 kWh during summer, and 8500 kWh during fall (Figs. 4.22-4.25). A greater load

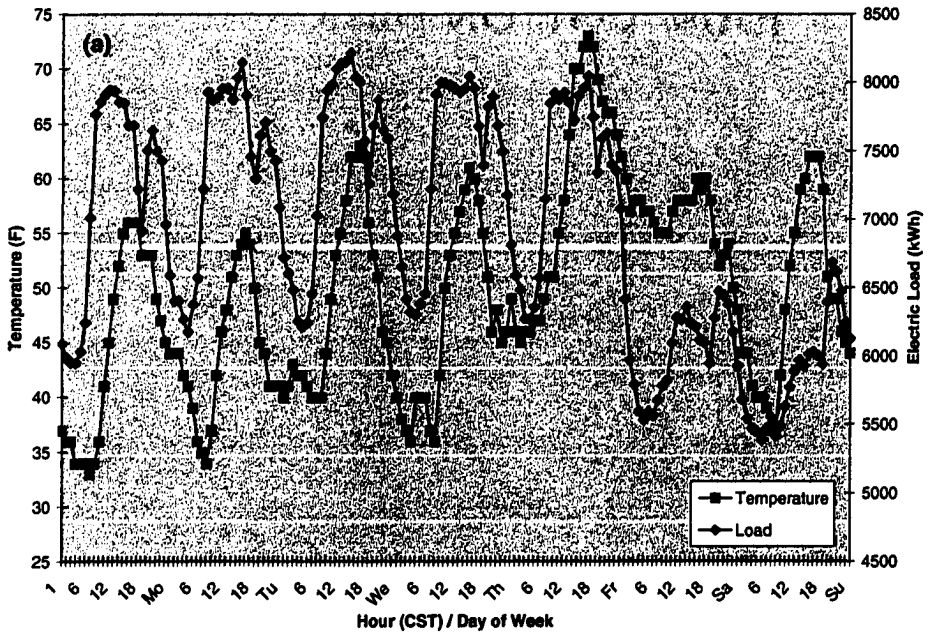
occurred during the warm season while a lesser load occurred during the cold season, as the primary source of heat at Altus AFB is natural gas, not electricity.

Weekend load profiles at Altus AFB were less variable than those which occurred during the week during all but the summer season (e.g., Fig. 5.17). The only similarity between the two load profiles was an 8 PM spike in the use of electricity. Obviously, base house residents were at home ~ 8 PM regardless of the day of the week. Otherwise, the weekend and weekday load profile diverged during the day. This difference reflected changes in the use of electricity due to daily activities of a commercial entity like Altus AFB versus the residential community at the base. Thus, for the commercial entity, the monthly-averaged load consumption always was significantly greater on weekdays than on weekends.

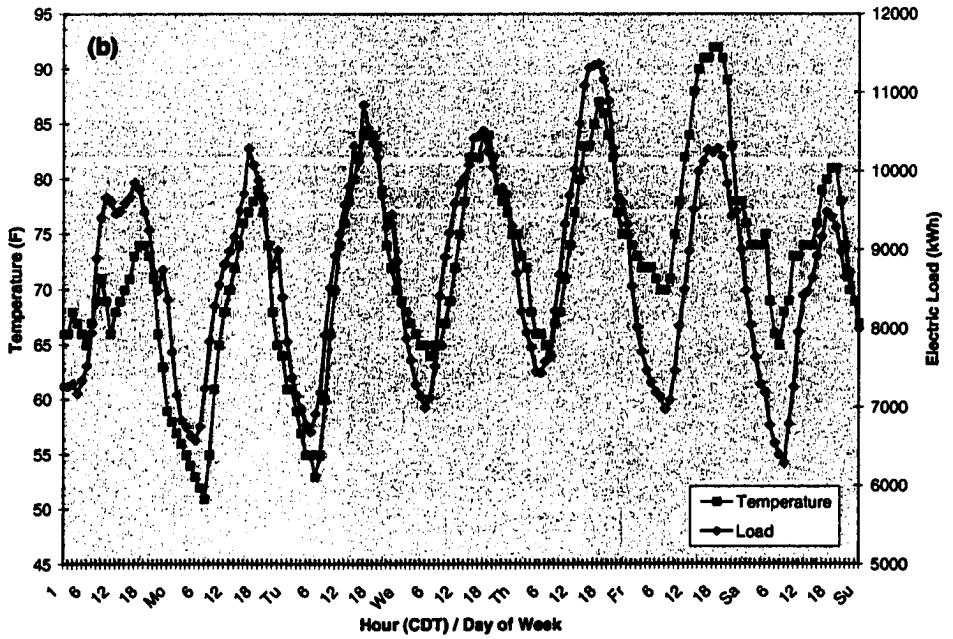
Graphs of weekly temperature and load data for each season (Figs. 5.18a-d) were analyzed to undercover detailed information about the use of electricity at Altus AFB which was smoothed by the monthly analysis. During the winter week, the load and temperature peaked within 1-2 hours of each other except on the weekend when an afternoon lull in load consumption coincided with the maximum temperature (Fig. 5.18a). The weekend trend reversed the relationship between temperature and load that had been observed during the winter. Electricity use on weekdays far exceeded that on the weekends.

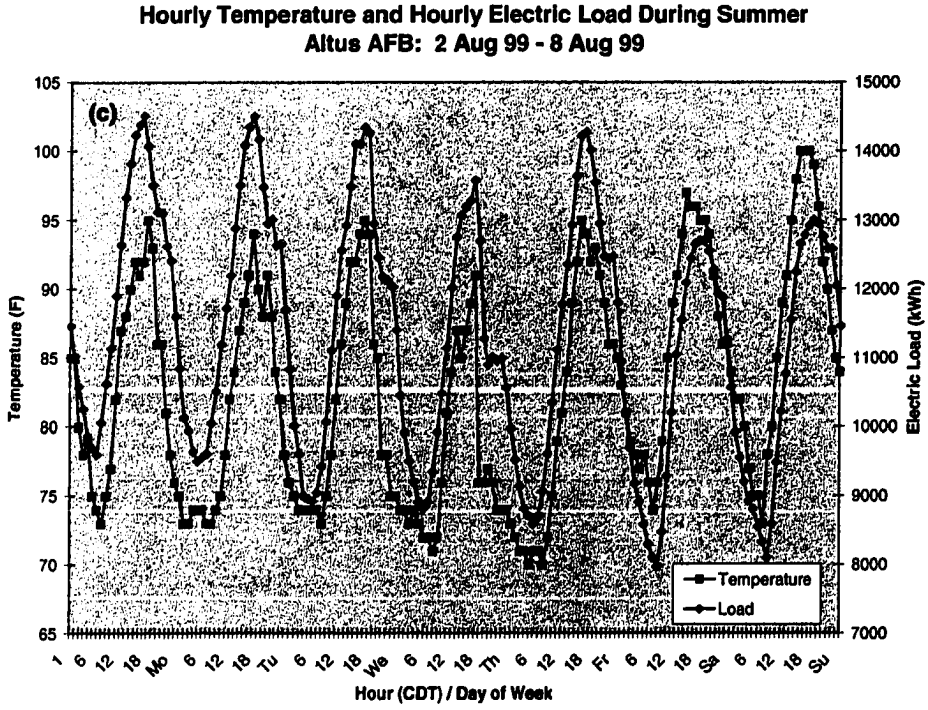
The weekly plot from the spring (Fig. 5.18b) and summer (Fig. 5.18c) seasons maintained a direct relationship between temperature and load. The peak load and temperature occurred simultaneously during both seasons. The weekend load demand was less than the weekday load demand, even when weekend temperatures exceeded the

Hourly Temperature and Hourly Electric Load During Winter
Altus AFB: 1 Feb 99 - 7 Feb 99



Hourly Temperature and Hourly Electric Load During Spring
Altus AFB: 17 May 99 - 23 May 99





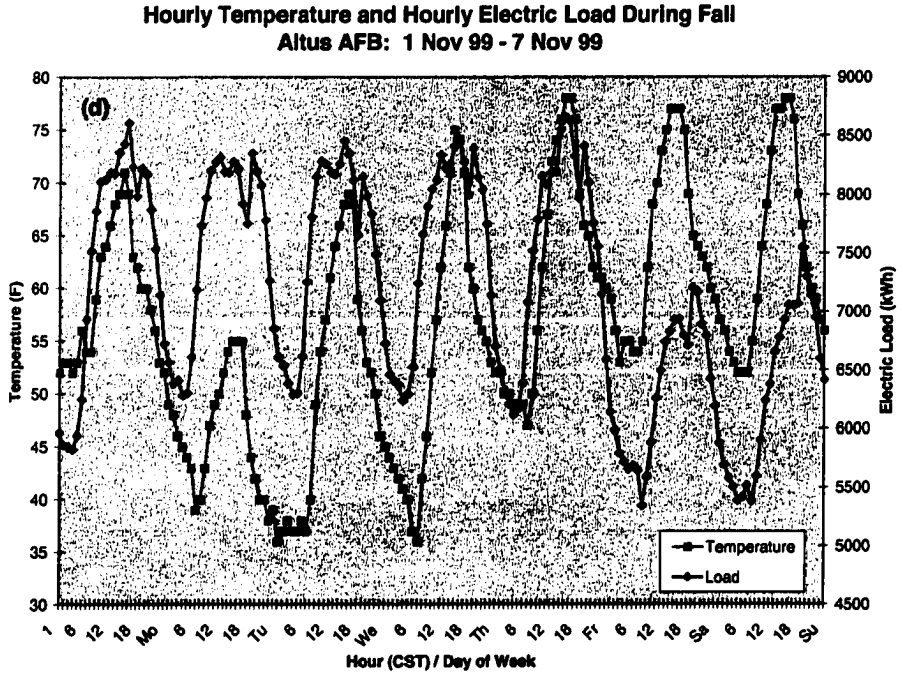


FIG. 5.18 Hourly temperature and hourly electric load from Altus AFB during a single week from the (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

weekday temperatures (e.g., Saturday and Sunday in August [Fig. 5.18a] and Saturday in May [Fig. 5.18b]). However, the use of electricity on a weekend was elevated when warmer temperatures occurred. In other words, the load demand on weekends would have been less when comparable temperatures occurred during the week.

During the fall season (Fig. 5.18d), a structure similar to that of a winter season occurred. A typical fall temperature swing occurred such that the difference between the maximum temperature on Monday and that on Tuesday was 16°F (8.9°C). However, because the load was unaffected by this temperature swing, the weekday load seems much more dependent on the work schedule rather than on air temperature at this commercial site.

5.4.1.3 Apparent Temperature-Load Relationship

The use of T_{app} did not provide any new information for use in better understanding the variation of electrical load to Altus AFB. Because of its strong dependence on the schedule for a given work week, the heat index and wind chill temperatures did not significantly impact the load any more than did the air temperature.

5.4.1.4 Temporal Correlations

Although the shape of the load profile seemed to depend on the work schedule at the AFB instead of temperature, the temperature-load relationship was still strong (ρ = 0.6; Table 5.1c). Furthermore, because the peaks load and maximum temperatures occurred nearly simultaneously throughout most a given year, a temporal lag between the load and temperature did not increase the correlation maxima. While part of the load at

Altus AFB came from residential users, that fraction of the load did not influence daily trends dominated by the weekday work schedule.

The temporal correlation during winter was positive, yet weak ($\rho = 0.23$). Both residential sites had stronger winter relationships between temperature and load ($\rho_{Norm} = -0.32$ and $\rho_{Wood} = -0.6$). During winter, the weekend profiles exhibited a strong inverse relationship between temperature and load (Fig. 5.18a). When data containing that inverse relationship was merged with the winter data that had a direct relationship from the weekday period, the resulting correlation coefficient was weak. The temporal correlation during spring and summer were high ($\rho = 0.79$ and $\rho = 0.86$, respectively) as the maximum load and temperature occurred at the same time. The fall season produced higher temporal correlations at Altus AFB than occurred at either residential site. The load and temperature reached a simultaneous maximum that, in turn, created a strong correlation ($\rho = 0.55$).

5.4.2 Relative Humidity-Load Relationship (Altus AFB)

A seasonal distribution of relative humidity and load for Altus AFB is shown in Figure 5.19. The scatter in this plot resembled that for the residential sites. A negative relationship existed between relative humidity and load, especially during the summer months. As relative humidity decreased, temperature increased followed by an increase in load during the summer months. The correlation was consistent with this inverse relationship. In fact, the correlations were strongest during the summer ($\rho = -0.66$) and weakest during the winter ($\rho = -0.12$). The spring and fall temporal correlations ranked between those from the summer and winter, as did the seasonal temperature.

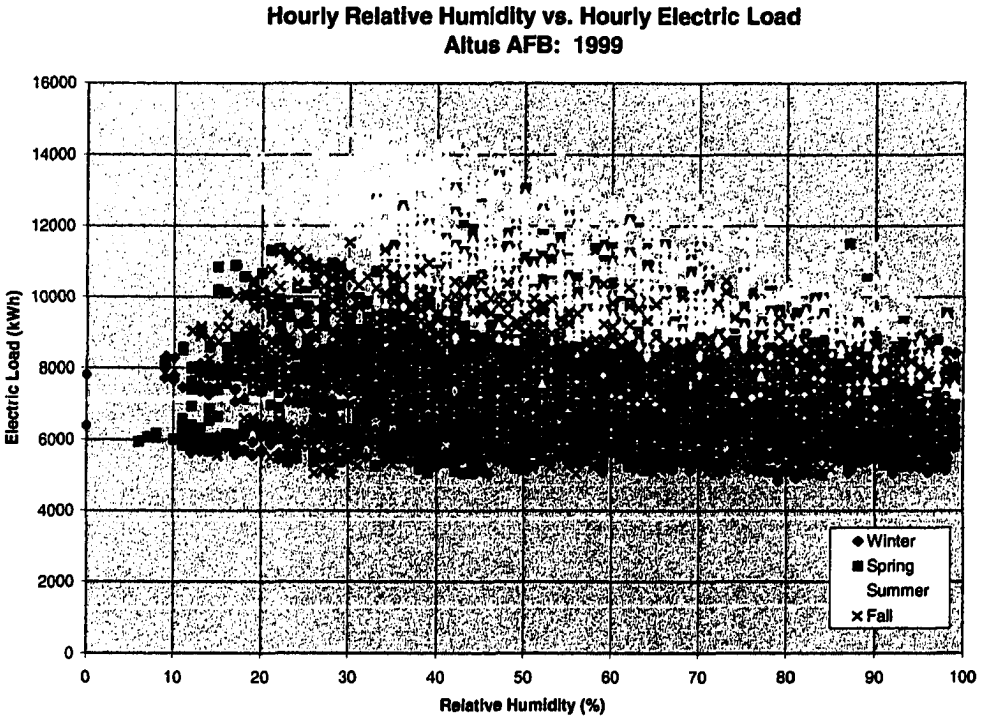
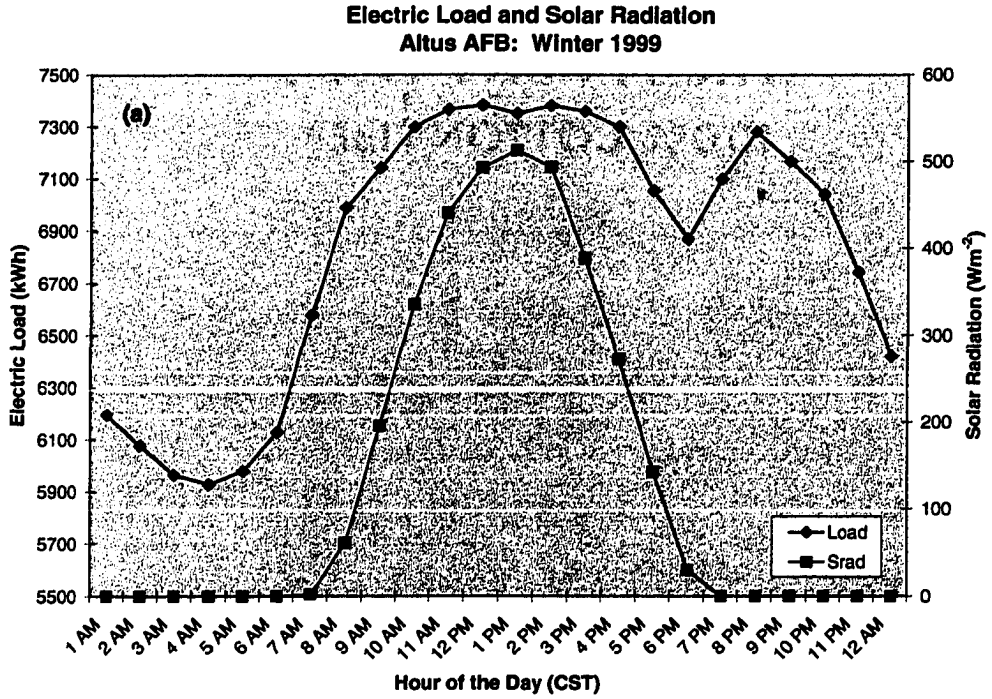
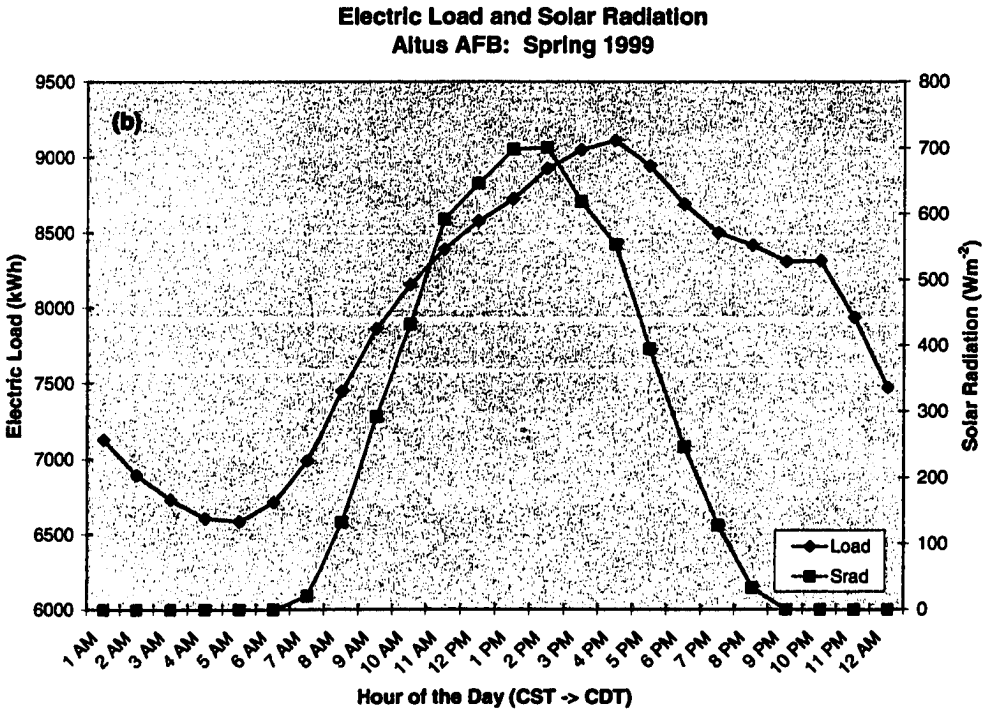


FIG. 5.19 Scatterplot of hourly relative humidity from the Altus Mesonet site versus hourly electric load from the Altus AFB substation for 1999.

5.4.3 Solar Radiation-Load Relationship (Altus AFB)

The solar radiation-load relationship at Altus AFB also was investigated. A scatterplot of solar radiation and load (not shown) provided inconclusive evidence. Yet, the temporal correlation (Table 5.1c) and graphs of seasonally-averaged solar radiation and load (Figs. 5.20a-d) reveal a strong relationship. The winter season produced the weakest correlation between solar radiation and load ($\rho = 0.23$). The multi-scale features of the load profile were difficult to reconcile with the well-defined plot of solar radiation. Thus, the strongest winter correlation, though weak, occurred at a temporal lag of zero.


Electric Load and Solar Radiation
Altus AFB: Winter 1999
(a)
Electric Load (kWh)
Solar Radiation (Wm-2)
7500
7300
7100
6900
6700
6500
6300
6100
5900
5700
5500
600
500
400
300
200
100
0
Load
Srad
1 AM
2 AM
3 AM
4 AM
5 AM
6 AM
7 AM
8 AM
9 AM
10 AM
11 AM
12 PM
1 PM
2 PM
3 PM
4 PM
5 PM
6 PM
7 PM
8 PM
9 PM
10 PM
11 PM
12 AM
Hour of the Day (CST)


Electric Load and Solar Radiation
Altus AFB: Spring 1999
(b)
Electric Load (kWh)
Solar Radiation (Wm-2)
9500
9000
8500
8000
7500
7000
6500
6000
800
700
600
500
400
300
200
100
0
Load
Srad
1 AM
2 AM
3 AM
4 AM
5 AM
6 AM
7 AM
8 AM
9 AM
10 AM
11 AM
12 PM
1 PM
2 PM
3 PM
4 PM
5 PM
6 PM
7 PM
8 PM
9 PM
10 PM
11 PM
12 AM
Hour of the Day (CST -> CDT)

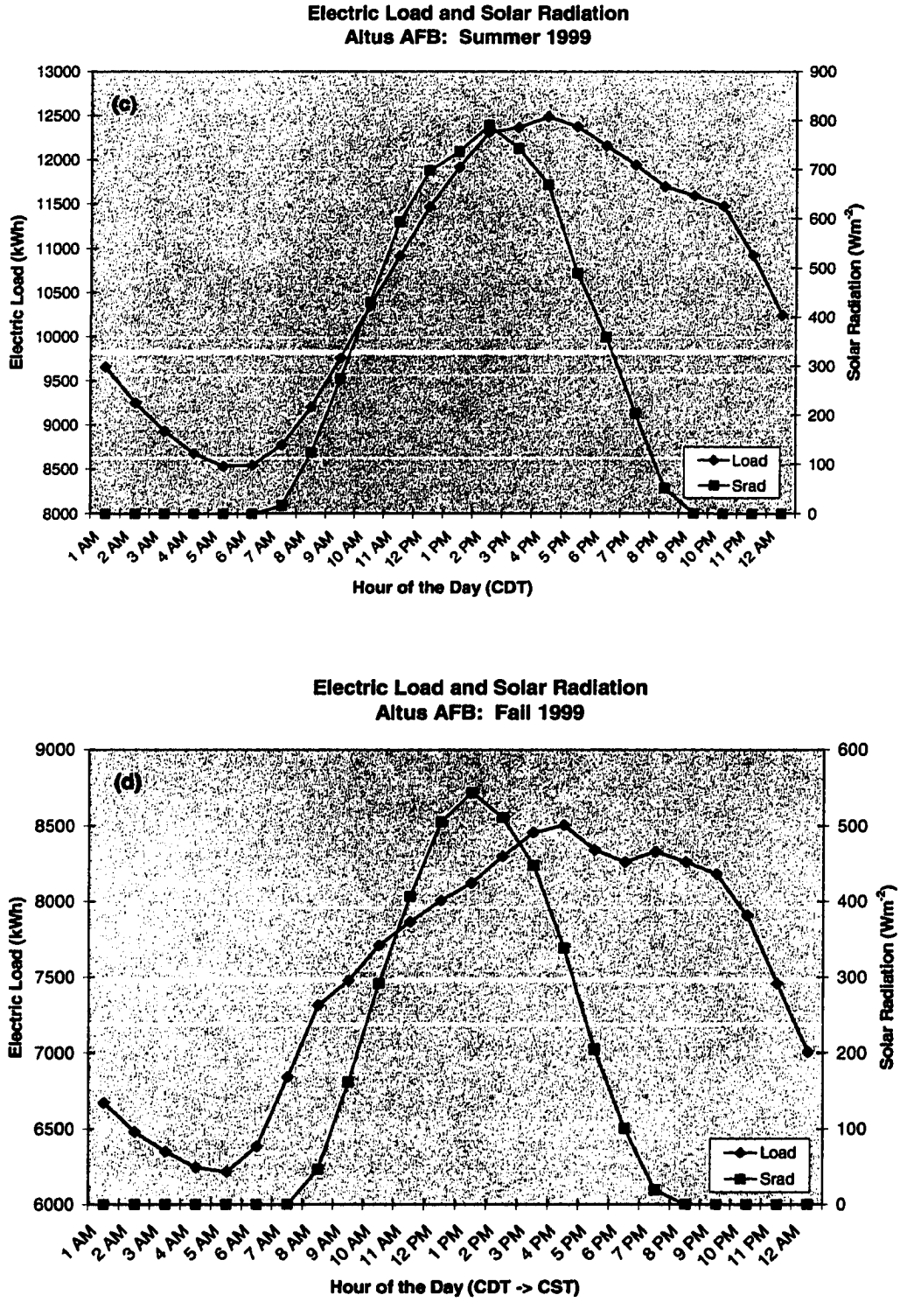


FIG. 5.20 Seasonally-averaged hourly electric load and hourly solar radiation from the Altus AFB substation and the Altus Mesonet site for (a) winter, (b) spring, (c) summer, and (d) fall of 1999.

The strongest correlation between solar radiation and load occurred during the spring when the sunshine preceded the electrical load using a lag of 3 hours ($\rho = 0.5$; Fig. 5.20b). The peak load at 4 PM trailed solar noon by 2-3 hours but was consistent with the time of maximum temperature. During the summer, solar noon at 2 PM preceded the peak load at 4 PM (Fig. 5.2c). However, the temporal correlation was strongest when the load trailed solar radiation by ~3 hours ($\rho = 0.71$). The load declined less rapidly than did the solar radiation at the end of each day. The fall season (Fig. 5.20d) produced results similar to that of the spring season. Thus, solar radiation may be a valuable weather variable used in predicting electrical load at the commercial site and at the residential sites.

5.5 Dominance/Broken Bow Site

The Dominance substation supplies electricity to an industrial warehouse that can operate 24-hours a day, 7 days a week, or it can shut down for weeks based on market demand. Furthermore, employees work either 8- or 12-hour shifts during the day or overnight hours. Thus, the load demand at Dominance appeared as random numbers over time.

5.5.1 Temperature-Load Relationship (Dominance)

5.5.1.1 Scatterplots

The annual scatterplot for Dominance (Fig. 5.21) had little structure when compared with data from the residential and commercial sites. Thus, load was

independent of temperature at this industrial site, consistent with the seasonal plots (Figs. 5.22a-d). Additionally, these plots revealed that a majority of plant closures (when the load was < 1000 kWh) occurred during the spring and fall seasons of 1999. During 2000 (not shown), most plant closures occurred during the summer and winter seasons, primarily as a result of market demand.

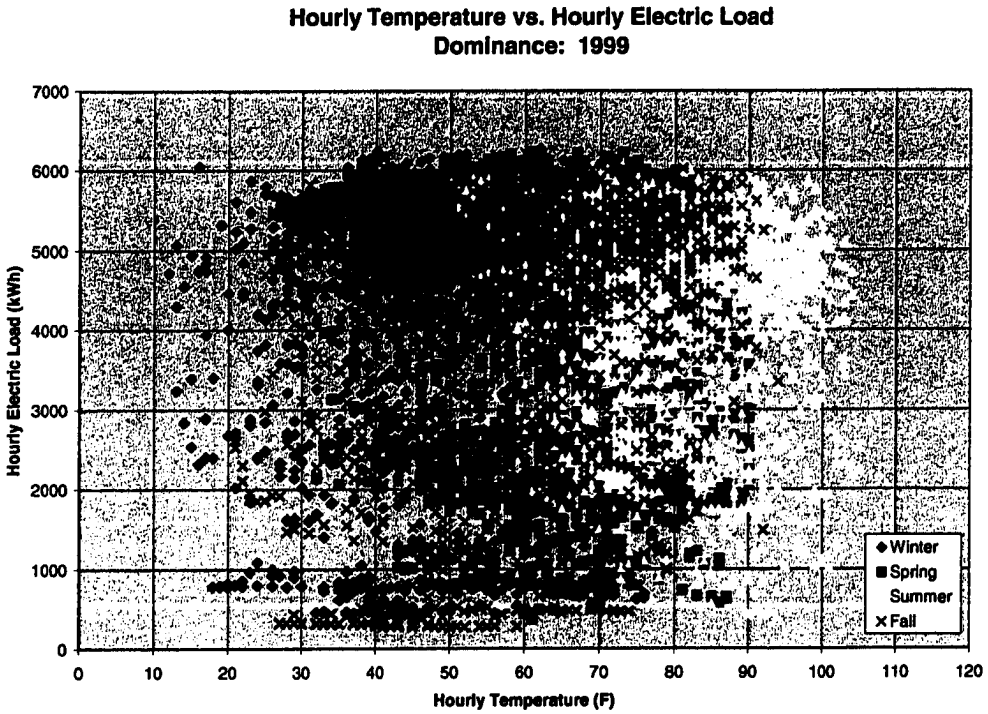
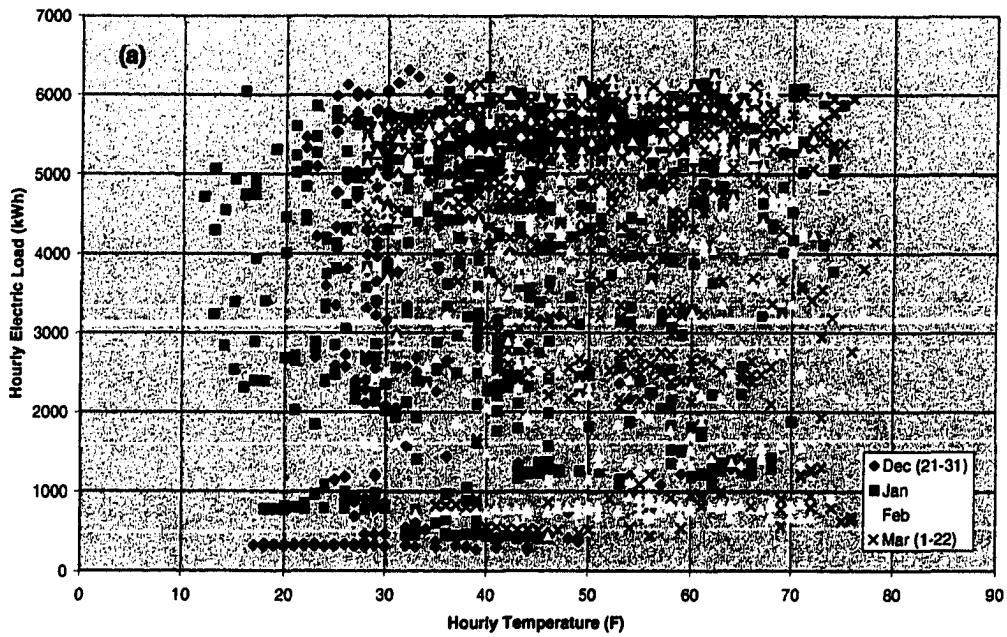
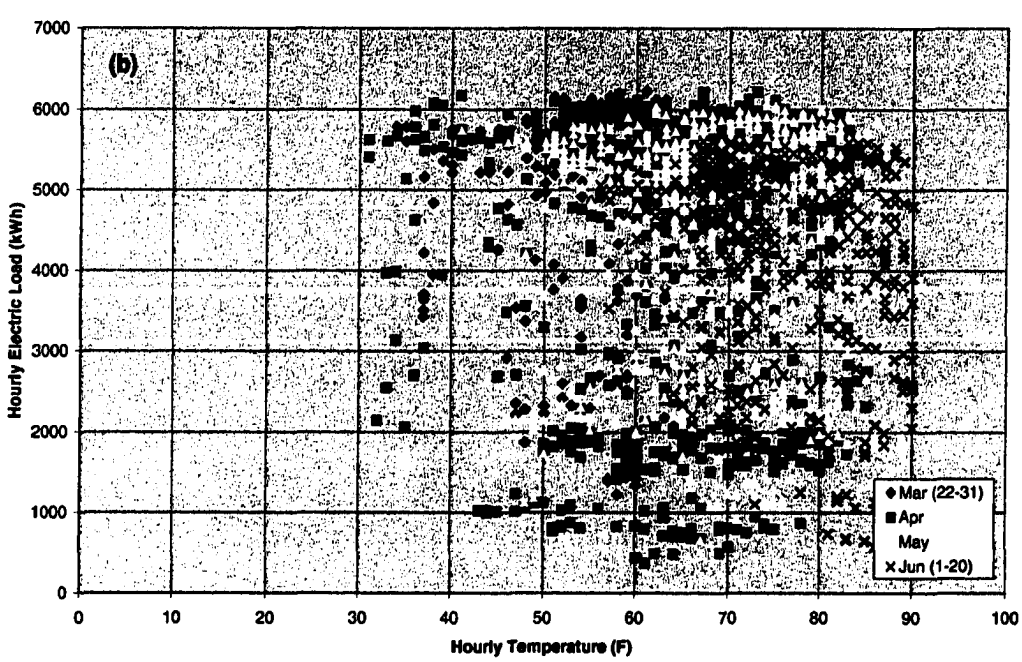


FIG. 5.21 Scatterplot of hourly temperature from the Broken Bow Mesonet site versus hourly electric load from the Dominance substation for year 1999.

Hourly Temperature vs. Hourly Electric Load
Dominance: Winter 1999 (21 Dec 98 - 22 Mar 99)



Hourly Temperature vs. Hourly Electric Load
Dominance: Spring 1999 (23 Mar - 20 Jun)



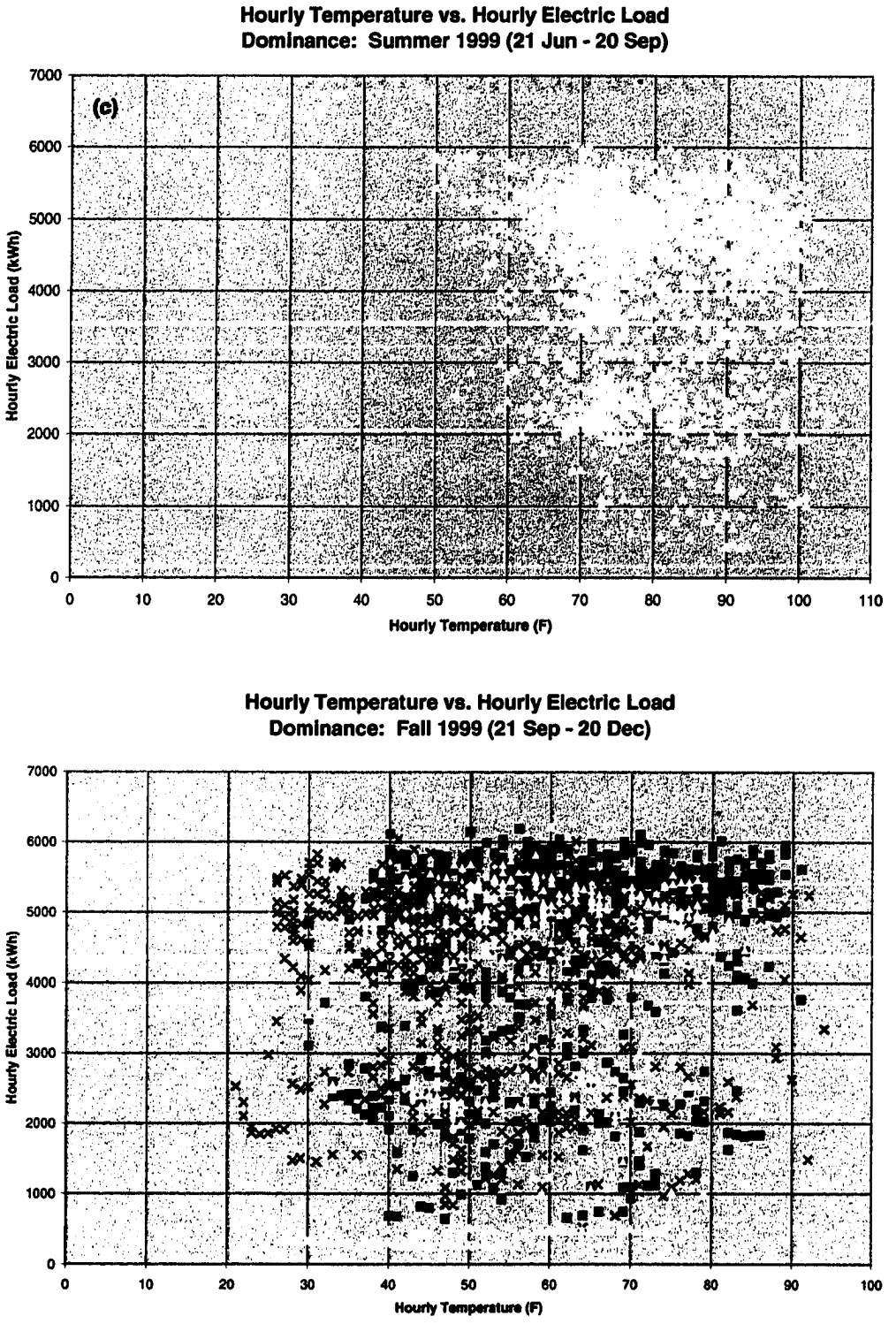


FIG. 5.22 As in Fig. 5.15 except for (a) winter, (b) spring, (c) summer, and fall of 1999.

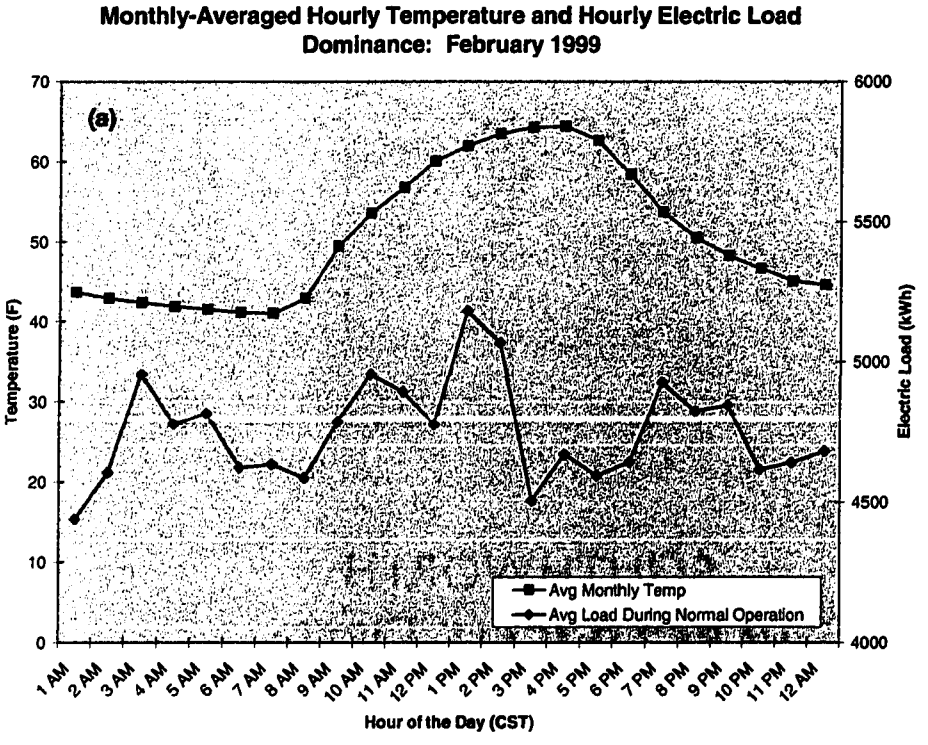
5.5.1.2 Line Graphs

Figures 5.23a-b displays hourly temperature and hourly load averaged over a month during a period of normal operation at Dominance (i.e., when the plant was not closed). When the warehouse was closed, its load was constant at ~1000 kWh each hour. The load was erratic during February and April and completely independent of temperature. Throughout the study period, no consistent load pattern was observed, even when the data were analyzed on a weekly time scale.

5.5.2 Temporal Correlations (Dominance)

Clearly, the load consumption at the industrial site was not influenced by temperature. The correlation between load and temperature was near zero (Table 5.1d). Other weather variables (e.g., apparent temperature, relative humidity, and solar radiation) that influenced the load for residential and commercial customers also had no bearing on the load at the industrial site. As a result, temporal lags were irrelevant.

It appears that load forecasting for Dominance would be totally influenced by advance notice that a plant closure was about to occur and minimal loads would result. Otherwise, in full operation, the variation in load was minimal. Throughout each year in the study period, more than 90% of the hourly load values were between 2000 kWh – 6500 kWh. These load values had no apparent relationship to any weather variable.



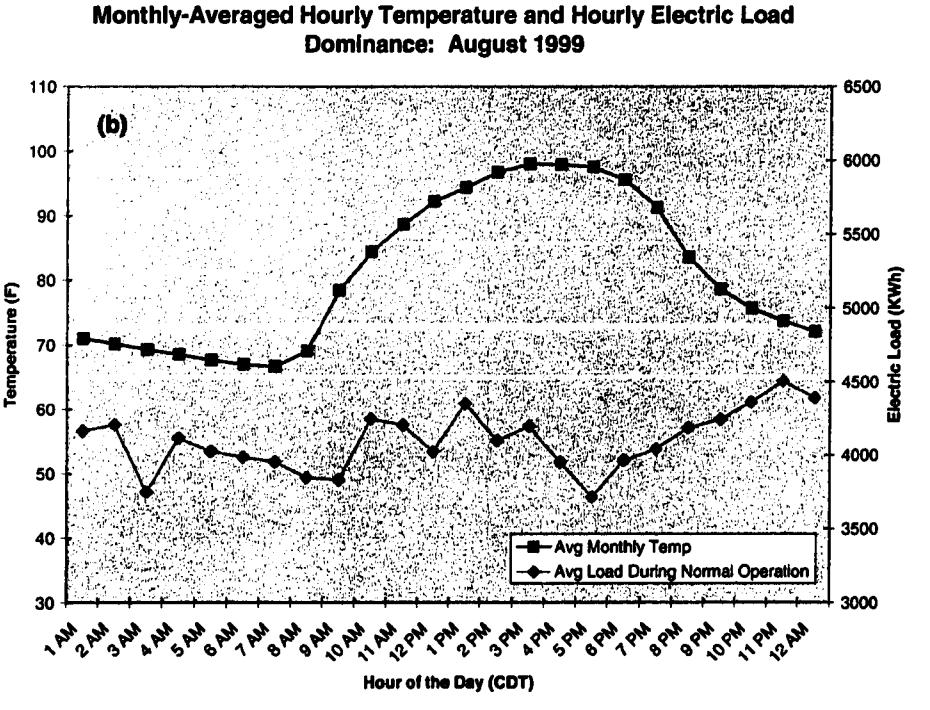


FIG. 5.23 Diurnal plot of hourly temperature and hourly electric load averaged over a month during a period of normal operation from Dominance for (a) February 1999 and (b) August 1999.

5.6 Summary

Evidence in this chapter was consistent with prior results documented in the scientific literature which clearly reveals that temperature is the best predictor of load consumption for all types of customers, except industrial. However, on a seasonal time scale, other variables proved to be valuable because they had stronger relationships than did temperature and load. Perhaps other weather variables, when available, can reduce load forecasting errors, and save money for the utility company and its customers. Modeling studies in Chapter 6 will test this concept.

Chapter 6: Applications of Electric Load Modeling

Two load modeling simulations were performed as another means to evaluate the null hypothesis (i.e., to determine if a comprehensive understanding of weather-load relationships will improve the accuracy of a load forecast). In other words, the interrelationships between weather and load developed in Chapter 5 were incorporated into a load model to reduce load forecasting errors. Modeling Study I produced a set of day-ahead load predictions for each substation using a neural network (NN) load model and a regression-based load model. The goal was to determine which combination of weather variables resulted in a more accurate load forecast using two load model strategies. The purpose of Modeling Study II was to determine the economic value of using a 21-century, high-resolution weather forecast versus an antiquated approach known as NGM MOS guidance (the technique used at WFEC).

Accuracy in this study was measured by a reduction in the forecast error which, in turn, became economically significant to a small utility like WFEC. The mean absolute percentage error (MAPE) was used as a performance index in both modeling studies; it is the average of the absolute value of the percentage residuals and is calculated as:

$$\text{MAPE} = \frac{\sum_{t=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t} \right|}{N} \times 100$$

where y_t is the observed load, $\hat{y}_t$ is the forecasted load, and N is the number of observations in the estimation period, at hour t. MAPE is a common statistical measure

used by electrical load modelers (Barakat and Al-Qasem 1998; Hobbs et al. 1999; and Alves da Silva 2000), in part because the error can be readily converted into an economic benefit (e.g., a dollar amount). For example, Hobbs et al. (1999) determined that for a 5000 MW peak generating system with a $20 mean variable generation cost and a 5% annual MAPE, "improved accuracy is worth about $0.6M to 1.6M annually per 1% improvement [in MAPE]." Other performance measures (e.g., root mean squared error, mean absolute error, and the sum of squared errors) were computed during this study, but those results are not discussed.

6.1 MetrixND – A Load Forecasting Tool

The MetrixND model, chosen for a comparative load forecasting analysis, was developed by the Regional Economic Research (recently acquired by Itron, Incorporated). MetrixND allows the load forecaster to choose from among the following modeling techniques: a neural network, multiple regression, exponential smoothing, and ARIMA. This study focused on load forecasts produced by the neural network and regression-based models. MetrixND also computes univariate statistics (e.g., means, standard deviations, and correlation coefficients for each predictor variable with the dependent variable) and a correlation matrix for all variables used in the model. Sample-period statistics, forecast statistics and error statistics are provided by MetrixND as well.

6.1.1 The Neural Network Model

The neural network model within MetrixND is an interactive model that provides the load forecaster with advanced specification options. The load forecaster also can

choose to eliminate bad observations, to eliminate observations to calculate forecast test statistics, and to place limits on the training period and forecast period. Finally, the NN model allows the load forecaster to enter the parameters that control the NN model's training and learning processes. (Specifications of the neural network are shown in Appendix F.)

The purpose of the training algorithm is to estimate the unknown parameters (i.e., to learn relationships among the variables). Once the variables are defined by the load forecaster, an initial set of coefficients is selected by using a random number generator. The estimation algorithm automatically adjusts the parameters to reduce the estimation (training) error until the process ultimately settles into one of many local optima (because it is virtually impossible to find the true global optimum). However, the estimation process is not hindered because a majority of the local optima have been shown to work well in both testing and forecasting modes. Once the parameters have been estimated by the NN model, the time period for a load forecast can be altered as new data become available, at which time the model parameters are updated.

In the end, the neural network model produces estimates of its coefficients, the predicted and residual values of electric load from the sample period, the predicted load values for the forecast period, and residual values of electrical load for the forecast period when actual load values become available.

6.1.2 The Multiple Regression Model

The linear (least-squares) regression technique has been used for decades in load forecasting. Although more advanced techniques have since been developed, the least-

squares method still finds its way into load models of today. The linear regression model, however, required transformations of weather variables to account for the nonlinearity between weather and load data. For example, transformations of temperature (e.g., temperature^2 and temperature^3) were used as predictor variables in this study. Hence, a clear understanding of the weather-load relationship is necessary prior to use of the regression model. In contrast, the neural network had the ability to learn complex relationships between variables, and thus, the pre-processing stage was not as laborious as occurred with the regression model. However, once the transformations were developed, the estimation period during which the regression coefficients were calculated was 3 times faster (using the regression model) than was the training period using the NN.

Once the variables in the regression model are specified, the coefficients are estimated by minimizing the sum of squares of the residuals (e.g., the deviations of the observed response from the fitted response) between the load forecasts and the observations. Finally, a check is performed for linearly dependent regressors; if identified, the parameters and associated statistics are set to zero. Once the estimation is complete, the execution of the forecast requires the same amount of time as does the NN.

6.1.3 Modeling Specifications

Twenty-four load models (i.e., one for each hour of the day) were built to produce day-ahead forecasts for both the NN and regression model. The output from the 24 models was used either as hourly forecasts during a specified period or they were combined to produce a daily load forecast (useful for generation and transmission

companies like WFEC). This hourly model approach allows the load modeler to choose a combination of variables that are inherent to a specific hour of the day. For example, solar radiation values are zero during the night. Thus, solar radiation may be eliminated as a variable during overnight hours.

Previous scientific works determined that several exogenous factors (e.g., the day of the week and holidays) were fundamental to load forecasting. Consequently, these variables (Table 6.1) were used in the hourly NNs during the 'training' stage to develop their coefficients and in the hourly regression models to develop regression coefficients. Each hourly model (e.g., the 1 AM regression model and the 1 AM NN model) used the same (non-weather related) exogenous variables to isolate the impact of weather. However, the variables (non-weather related) used in the 1 AM models were sometimes different than the variables used in 3 PM models. Other user-defined model specifications for the NN and regression model are listed in Tables 6.2a-b.

6.2 Modeling Study I

A neural network-based model (e.g., a popular yet controversial modeling technique) and a multiple regression-based model (e.g., a traditional load modeling technique) were chosen to produce load forecasts at each study site using (1) temperature (LF_T), and (2) temperature, relative humidity and solar radiation ($LF_{T,RH,Srad}$). MAPEs were developed over a specified forecast period, and the results from each model were compared. Finally, the errors were quantified into their economic impacts.

Table 6.1 The user-defined, exogenous (non-meteorological) variables used in the hourly neural network and regression models.

Exogenous Variables for the NN and the Regression Model
❑ Years (1998, 1999, and 2000)
❑ Seasons (winter, spring, summer, and fall)
❑ Months (January, February, ..., December)
❑ Days of the week (Monday, Tuesday, ..., Sunday)
❑ Holidays (all federally recognized holidays)
❑ Heating Degree Days (calculated using threshold temperatures developed in this study)
❑ Cooling Degree Days (as calculated by the NWS)

Table 6.2 The specification list for the (a) neural network and (b) regression model in Modeling Study I, including weather variables.

(a) ***Specification List for the Neural Network***
❑ Single output feedforward neural net
❑ Linear activation function at the output layer
❑ Sigmoidal activation function at the hidden layer
❑ 3 nodes in the hidden layer
❑ 3 trials in the training module
❑ 100 iterations for each trial (training and learning)
❑ Convergence criterion for trial is 0.0001
❑ Bad spots
❑ Test periods
❑ Estimation period
❑ Forecast period
❑ Hourly load
❑ Hourly temperature (T)
❑ Hourly relative humidity (RH)
❑ Hourly solar radiation (Srad)

(b) Specification List for the Regression Model

- Bad spots
- Test periods
- Estimation period
- Forecast period
- Hourly load
- Hourly temperature transformations
- Hourly relative humidity transformations
- Hourly solar radiation transformations

Results of Modeling Study I

Monthly-averaged MAPEs were computed from hourly load forecasts using the NN and regression models. A "good" load forecast by (electric utility) industry standards is when the MAPE is less than 5% (Khotanzad et al. 1997; Hobbs et al. 1999). Accordingly, the results in this section are "good" by industry standard.

For illustration purposes, one month was chosen to represent each season at each of the four study sites (Table 6.3). At the residential and commercial sites, the model errors from both simulations were nearly equal or reduced when temperature, relative humidity, and solar radiation (compared to using temperature alone) were used in the NN and regression load models. The electrical load at the industrial site did not in any way resemble a load profile from a residential or commercial site (e.g., the load was not dependent on the exogenous variables used in the models); at the Dominance substation, the errors were large and inconsistent. The differences in the MAPEs for LF_T and $LF_{T,RH,Srad}$ at all four sites ranged between ~0% - 0.55% for the NN model and between ~0% - 0.86% for the regression model. In these simulations, 0% implied that the forecast

Table 6.3 A comparison of the mean absolute percentage errors (MAPEs) of the load forecast at the four study sites from the NN and from the regression model using temperature (LF_T) versus using temperature, relative humidity, and solar radiation ($LF_{T,RH,Srad}$).

NORMAN	*NN*	*Regression*
Feb-99		
Temp / Temp, RH, Srad	3.66% / 3.40%	3.46% / 3.22%
Difference	0.26%	0.24%
May-99		
Temp / Temp, RH, Srad	5.41% / 4.86%	5.63% / 4.77%
Difference	0.55%	0.86%
Jul-99		
Temp / Temp, RH, Srad	3.49% / 3.25%	3.45% / 3.02%
Difference	0.24%	0.43%
Oct-99		
Temp / Temp, RH, Srad	3.68% / 3.32%	3.49% / 3.22%
Difference	0.25%	0.27%

WOODWARD	*NN*	*Regression*
Feb-99		
Temp / Temp, RH, Srad	3.73% / 3.68%	3.83% / 2.96%
Difference	0.05%	0.87%
May-99		
Temp / Temp, RH, Srad	4.34% / 4.29%	4.26% / 4.24%
Difference	0.05%	0.02%
Jul-99		
Temp / Temp, RH, Srad	3.46% / 3.45%	3.43% / 3.43%
Difference	0.01%	0.00%
Oct-99		
Temp / Temp, RH, Srad	3.89% / 3.70%	3.81% / 3.77%
Difference	0.19%	0.04%

ALTUS AFB	*NN*	*Regression*
Feb-99		
Temp / Temp, RH, Srad	1.80% / 1.66%	1.65% / 1.44%
Difference	0.14%	0.21%
May-99		
Temp / Temp, RH, Srad	3.18% / 2.88%	3.19% / 2.88%
Difference	0.30%	0.31%
Jul-99		
Temp / Temp, RH, Srad	2.04% / 1.91%	2.24% / 2.06%
Difference	0.13%	0.43%
Oct-99		
Temp / Temp, RH, Srad	2.44% / 2.43%	2.72% / 2.37%
Difference	0.01%	0.35%

DOMINANCE	*NN*	*Regression*
Feb-99		
Temp / Temp, RH, Srad	25.95% / 25.91%	25.84% / 26.13%
Difference	0.04%	-0.29%
May-99		
Temp / Temp, RH, Srad	26.34% / 26.41%	26.41% / 26.42%
Difference	-0.07%	-0.01%
Jul-99		
Temp / Temp, RH, Srad	23.54% / 23.63%	23.80% / 23.08%
Difference	-0.09%	0.72%
Oct-99		
Temp / Temp, RH, Srad	21.74% / 21.47%	21.99% / 21.73%
Difference	0.27%	0.26%

error was the same regardless of whether using temperature, relative humidity and solar radiation served as predictors versus temperature alone.

Statistical significance of the differences between the MAPEs from the two forecasts (LF_T and $LF_{T,RH,Srad}$) were analyzed using a Student t-test. Each season was tested separately (such that the degrees of freedom = the number of days in the season minus 2). At Norman, the differences between the means of the MAPEs for LF_T and $LF_{T,RH,Srad}$ were statistically significant 90% of the time. This result occurred for both models during all seasons. At Woodward, the only significant differences (at a 95% confidence interval) in MAPEs between the two load forecasts (LF_T versus $LF_{T,RH,Srad}$) occurred during the winter season using the regression model. The regression model also produced significant differences at Altus AFB with 90% confidence for all seasons except winter. Otherwise, the differences between the MAPEs for the two sets of load forecasts were insignificant at Woodward and Altus AFB. Statistical significance was not relevant at Dominance because load forecasts were not related to meteorological phenomena. Additionally, high values of R^2 (i.e., greater than 0.85) were produced by the load model simulations all sites except for Dominance[1]. The bottom line is that a small improvement in a load forecast can be *economically* significant to a utility company (Hobbs et al. 1999), which means a decrease in operating costs or an increase in operating revenues.

The modeling results revealed that the use of electricity at Norman responded to more than one meteorological variable. For example, when the temperature is 60°F

[1] R^2 is an accepted statistical tool to measure 'goodness of fit', where values close to 1.0 (0.0) indicate that the model explains the majority (very little) of the variation of the dependent variable. Using temperature, relative humidity and solar radiation versus using temperature alone was judged to produce an insignificant improvement in the R^2 value.

(15.6°C) and the sky is sunny versus the conditions of 60°F (15.6°C) and cloudy skies, temperature *and* solar radiation would model this scenario better than using temperature alone. Throughout the 24-hour forecast period, the regression model uncovered a greater benefit for using three weather variables than did the NN; in addition, the regression model always produced more accurate load forecasts. Seasonally, in both models, the smallest errors occurred during the summer month when the weather-load correlations were the strongest ($\rho = 0.87$). However, the least accurate forecasts occurred during the spring month, even though the weather-load relationship was strong ($\rho = \sim 0.7$). This surprising result was derived from the fact that load data from Norman were missing for the months of April 1998, March 1999, and April 1999; hence, the estimation of the model coefficients may have suffered, especially during the spring months. Thus, at Norman, the load was most effectively modeled using the regression model with temperature, relative humidity and solar radiation.

The benefit of including relative humidity and solar radiation data in the load models was minimal at Woodward, except when the regression model was used during the winter season (Table 6.3). The most accurate load forecast at Woodward occurred during the summer in both models when the strongest weather-load relationship occurred. However, like Norman, the weakest weather-load correlations occurred during the fall season while the least accurate forecast occurred during the spring season. Surprisingly, both models learned the weather-load relationship of the fall season more effectively than might be implied through the strength of the temporal correlations. From a modeling standpoint, the load at Woodward was most accurately predicted by both models using temperature as the sole predictor.

Relative to the other study sites, both models at Altus AFB produced the smallest MAPEs. The pattern of electricity use for this commercial entity was effectively modeled in both simulations, even though a portion of the Altus load had a residential component. The benefit of using additional weather variables occurred throughout the forecast period when the regression model was used, but they were insignificant when the NN was used. However, the accuracy of load forecasts from both models was comparable for each month. Seasonally, the load prediction at Altus AFB was most *accurate* during the winter month, when the *weakest* weather-load relationships occurred (ρ = ~0.23). This surprising result occurred because the heating demand is met by natural gas instead of electricity (i.e., the response to colder temperatures was not observed in the load demand).

The load forecasting errors at the industrial site were a magnitude greater than those at the residential and commercial sites. The poor forecast of load at the Dominance substation could not be modeled effectively by use of predictor variables normally used for the other customer categories. A purely industrial load like occurred at Dominance must be estimated after consideration of the market demand for the product and any scheduled closures and downtime at the warehouse. Thus, comparing the differences between the MAPEs for LF_T and $LF_{T,RH,Srad}$ were not relevant for Dominance.

The use of temperature, relative humidity and temperature improved the accuracy of the load forecast at Norman no matter which modeling technique was chosen. The accuracy of the load forecasts at Altus AFB was improved (using three weather variables) during all seasons using the regression model. On the other hand, temperature can be used as the sole predictor at Altus AFB when using the NN. Except for Norman, the NN

apparently "learned" the intricacies of the relationship between temperature and load such that other weather variables were not useful to reduce errors in the load forecast. However, the accuracy of the load forecasts from the regression model (in most cases) benefited from the use of relative humidity and solar radiation. Either model is capable of producing an acceptable load forecast. However, the careful choice of weather variables in a model can improve the accuracy of a load forecast.

The load forecast at the urban, residential site was most impacted by the use of additional weather variables. *More importantly, the methods used to improve the load forecast for an urban, residential community in Oklahoma are applicable across the United States and perhaps worldwide.*

6.3 Modeling Study II

Because temperature is essential to an accurate load forecast, the most accurate temperature data are required. Modeling Study I used historical weather data, such that error in the weather forecasts were not a factor in the load model forecast. However, Khotanzad et al. (1997) documented that 1-2 day weather forecasts were responsible for ~1% additional (MAPE) error in load forecasts. Thus, the purpose of Modeling Study II was to determine the value of improved temperature forecasts to a load forecast. Load forecasts were compared based upon a high-resolution temperature forecast (20-km Eta Model) versus the temperature forecast from NGM MOS.

Spatial interpolation of the Eta Model output was used to produce a temperature forecast at Norman (35.24N, 97.47W), whereas the nearest available NGM MOS guidance was the Oklahoma City airport site (35.23N, 97.36W). In addition, forecasts

from the Eta Model are updated every 6 hours while the NGM MOS is issued twice daily. Hence, a higher-resolution model with a greater temporal frequency should produce a more accurate weather forecast that, in turn, should reduce the MAPE in a load forecast.

The NN model from MetrixND was used to produce hourly forecasts out to 24 hours at Norman on 4 April 2003 (a typical spring day). Three data sets were used: observed, hourly temperature from the NWS Weather Forecast Office in Norman; temperature forecasts from the 20-km Eta Model, and temperature forecasts from NGM MOS. Because solar radiation forecasts are not available from weather forecast models, this study demonstrated the impact of improved temperature forecasts on the accuracy of load forecast.

Results of Modeling Study II

Twenty-four hourly NNs were trained with observed hourly, electric load data from the West Norman substation and hourly temperature data from the Norman Mesonet site from 2001 and 2002. A load forecast was produced at Norman on 4 April 2003 using the NWS observed air temperature. This load forecast was considered a "perfect" load forecast in that error resulting from an imperfect weather forecast was eliminated from the load forecast. Two additional load forecasts were produced; one used temperature forecasts from the Eta model and the other used forecast temperatures from NGM MOS. The three sets of temperature data are displayed in Figure 6.1.

Temperature forecasts inherently introduce error into the load forecast model. On 4 April 2003, Eta Model forecasts had a mean temperature error of 3.4°F (1.9°C) versus a 6°F (3.3°C) error in temperature forecasts from NGM MOS, when they were compared to

the observed air temperature. However, a 2.6°F (1.4°C) improvement in the temperature forecast (Eta versus NGM MOS) reduced the MAPE of the load forecast by 0.55% at Norman. This improvement to the load forecast was clearly dependent on the accuracy the temperature forecast for the online component of the load model.

Load forecasts are imperfect even if a perfect weather forecast was available, in part due to the stochastic component of the load. In an effort to minimize load errors in the face of deregulation or to minimize operating expenses for any company, accurate weather forecasts are essential. While high-resolution weather forecasts are not perfect, forecasts produced by the Eta and RUC Models, for example, have demonstrated significant improvements in skill over forecasts in the NGM MOS guidance (Kalnay 2003).

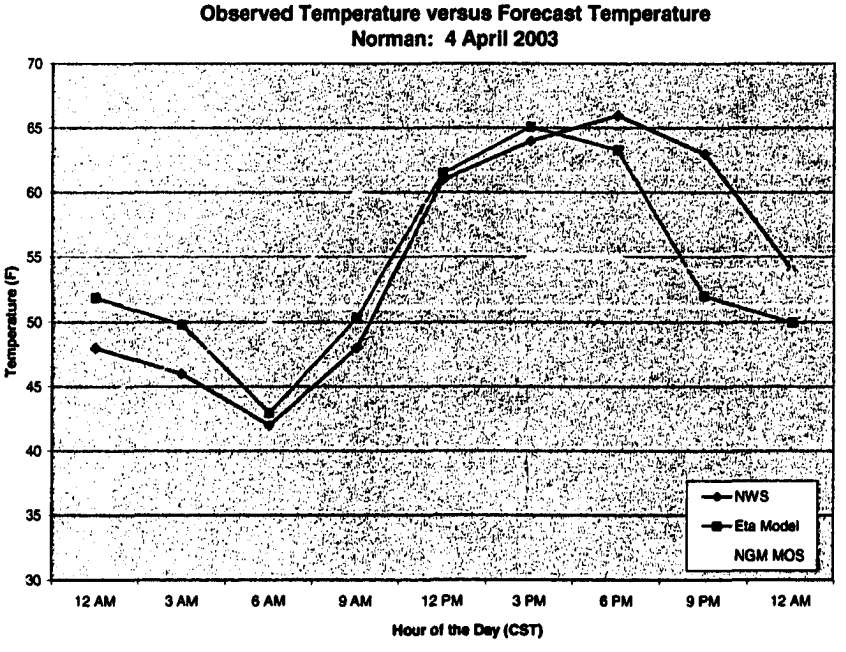


FIG. 6.1 Observed air temperature from the Norman WFO, forecast temperature from the 20-km Eta Model, and forecast temperature from NGM MOS at Norman on 4 April 2003.

6.4 Economic Benefit of Improved Load Forecasts

The impact of the two modeling studies was quantified by calculating a dollar amount saved by WFEC if they could reduce the MAPE in their load forecasts. The economic benefits to WFEC are unique because of: the size and generation capacity of WFEC; the agreements WFEC may have with other electricity generators; and the market prices at the time of the study. Thus, based on this study, the impact of weather variables on electric load should be transportable across the United States.

WFEC has a generation capacity of 1300 MW. Hugo, the coal-fired plant owned and operated by WFEC, can generate an additional ~400 MW. Because coal is the least expensive generation method available at WFEC, WFEC daily operates the Hugo plant to its full capacity. Each day, WFEC must determine the most cost effective method to fulfill the remainder of the electricity demand above the generation capacity of Hugo. Their choices include whether to generate power using natural gas, to purchase power at market prices, or to interchange power with another utility (e.g., Oklahoma Gas and Electric).

During 2002, the system load averaged 800 MW (MWh/hour) at WFEC. During April of 2003, the cost to generate electricity using coal averaged $13/MW and averaged $50/MW if natural gas were used[2]. In addition, the market prices for electricity ranged from $12.68 during off-peak hours to $74.60 during peak hours[3]. In Modeling Study I, an average 0.3% reduction of the MAPE was detected at Norman when temperature, relative humidity, and solar radiation were used in the load models.

[2] Prices supplied by WFEC on April 4, 2003 (personal communication).
[3] www.NrgStream.com, an online source for current energy market information

Suppose this average reduction in MAPE was realizable across the entire WFEC system, and that the load system analyst decided to purchase electricity at a market price of $45/MW because it was cheaper than generating the power. The use of three weather variables in a new-generation load forecast could have saved WFEC nearly $0.5M annually (800 MW – 400 MW = 400 MWh/hour x 24 hours/day x 365 days/year = 3,504,000 MW/year x 0.3% = 10512 MW/year x $45/MW).

The results in Modeling Study II also were quantified. On 4 April, 2003, the reduction in MAPE was 0.55% at Norman when an improved temperature forecast was used in the load model. The Norman substation distributed 133 MW for the day, which would translate into a savings of $16.50/day (provided that half the load was generated using coal and the other half was purchased at a market price of $45/MW). However, Norman is only one of 225 substations owned by WFEC, and is one of the larger substations in the WFEC system. The reduction in error in the load forecast will vary throughout the year and vary at each substation across the state as a result of using improved the temperature forecasts. Thus, suppose the savings averaged only $6.50/day at each substation; this could result in an annual savings of over $500,000 for the WFEC system.

Therefore, the value of a comprehensive understanding of the interrelationships between weather and load data combined with improved temperature forecasts can save a small utility at least $0.5M annually. Larger power companies like Virginia Dominion Power and Florida Power and Light generate more than 10 times the amount of electricity produced by WFEC and have operating revenues that are more than 4 times that of WFEC. If the results are multiplicative and if 100 large electric utilities applied these

results to their operational decisions, the reduced operating expenses could approach $0.25 billion per year. Regardless, the potential improvements in the load forecasts at WFEC, if projected to the larger power companies, would create almost astronomical reductions in annual operating expenses!

Chapter 7: Summary and Concluding Remarks

The hypothesis of this dissertation was that *a comprehensive understanding of the relationship between weather variables and electricity demand will improve the accuracy of a load forecast.* It was evaluated using three years of hourly weather variables from the Oklahoma Mesonet and hourly electrical load from co-located WFEC substations. The two data sets permitted a unique examination of the weather-load relationship. Load modeling simulations incorporated the interrelationships established in this study to determine the validity of this hypothesis.

This research is unparalleled in that the spatially-dense network offered by the Oklahoma Mesonet obtained weather observations which were within ~4 miles of each substation. In addition, each substation represented four customer classes: urban residential – Norman; rural residential – Woodward; commercial – Altus AFB; and industrial – Dominance. Previous scientific work focused on the system load of an entire utility company (which combined all customer categories and substations) and the relationship of those loads with temperature observed at an NWS airport site. Unfortunately, the observations used in previous work also were from unmaintained and spatially inadequate networks while concurrent forecasts were from antiquated weather models. While a few studies were conducted using other weather variables, those studies produced inconclusive results. Even though this study focused on the load consumption and weather patterns in Oklahoma, the methodology and general results of this analysis are transportable to other states.

An extensive analysis of the relationship between weather variables and substation-level load revealed that weather clearly influenced the use of electricity, though the impacts varied based on factors such as time of year, time of day, and customer class. Principal results about the interrelationships between weather and load (excluding the industrial customer) are:

1. Temperature is the most important weather variable to load forecasting. Temporal correlations were strongest between temperature and load during summer (ρ = ~0.9) for residential and commercial customers. The weakest relationship between temperature and load occurred during the fall (ρ = ~0.3) for the residential user and during winter (ρ = 0.2) for the commercial user. Because of the important role of temperature as a predictor of load, careful attention must be given to the sources and quality of the historical and forecasted temperature data.
2. Solar radiation revealed predictive power for load in the residential community, especially during the fall season when the temporal correlation between solar radiation and load (ρ = ~0.4) was *stronger* than that of temperature and load. However, the solar radiation-load relationship was weakest during the winter. Solar radiation also had a strong relationship with load at the commercial site, but the temporal correlations never exceeded those between temperature and load. Because the solar radiation-load relationship is not trivial, a comprehensive understanding of this relationship is necessary to maximize its potential as a predictor of load. Even so, solar radiation is not readily available from weather forecast models; hence its ability to improve a load forecast in real-time is unknown.

3. Relative humidity and load were inversely related (except at Woodward during the winter), and their relationship was strongest during summer ($\rho = \sim 0.7$). However, the temporal correlations between relative humidity and load were not as strong as those using temperature or solar radiation. Thus, relative humidity did not provide new independent information for load modeling.
4. While the apparent temperature and load relationship was similar to that for temperature and load, the evidence did not suggest that use of apparent temperature would produce a more accurate load forecast than one which used only temperature. In addition, the use of wind speed appeared to have no value to load modeling.
5. On an annual basis, the weather-load relationship was highly nonlinear. However, a seasonal investigation revealed near-linear relationships during winter and summer while the relationships during the transitional seasons remained nonlinear. A seasonal analysis also revealed that variations in weather impacted the shape *and* magnitude of the load profiles. Within a particular season, fluctuations in weather altered the magnitude of load profiles while the shape for that season was maintained.
6. Variations in the magnitude and shape of the load profile were observed when weekday and weekend loads were compared. For example, a sharp increase in the use of electricity occurred during the morning hours of a week, while the increase was gradual on weekend mornings. Temperatures influenced the magnitude of the morning use of electricity, while the shape was determined by the day of week.
7. The age, building materials, and insulation quality of a home appeared to impact load consumption and the response time of electricity use to fluctuations in temperature.

In particular, older homes responded quickly to temperature changes while new construction (with better-quality insulation) delayed the load response to ~2 hours following temperature changes.

8. Threshold temperatures were discovered that revealed how the characteristics of the load demand changed from the cold side to the warm side of the threshold. On the warm (cold) side, the temperature-load relationship mimicked a summer (winter) scatterplot. The threshold temperatures were: 62°F (16.7°C) at Norman, 57°F (13.9°C) at Woodward, and 61°F (16.1°C) at Altus AFB. These thresholds should prove to be a more accurate representation of the true heating and cooling degree days instead of the current NWS standard which is based on a daily mean temperature of 65°F (18.3°C).
9. The load profile for an industrial user was completely different from residential and commercial users and completely different from profiles documented in the scientific literature. At Dominance, the load was unaffected by weather or other factors deemed appropriate for load forecasting. Instead, the market demand for the industrial product drives the plant schedule, which, in turn, dictates the demand for electricity. Accordingly, an industrial user like Dominance should be modeled separately from the residential and commercial users (which were influenced by weather).

Two modeling studies were conducted to determine the economic value of using an improved understanding of interrelationships between weather and load to a load forecast. A neural network (NN) and a regression-based load model were used to

produce the load forecasts. The observed weather variables were obtained from the Oklahoma Mesonet, while temperature forecasts were acquired from the Eta Model and NGM MOS guidance. The key results include:

1. The use of temperature, relative humidity and solar radiation (instead of temperature alone) in the two load models improved forecast accuracy at Norman during each season. Otherwise, the improvements were statistically insignificant. It was not surprising that the commercial and industrial site did not benefit from the additional weather variables. However, it was interesting to determine that the load forecasts at Woodward and Norman required different predictors for the most accurate load forecast.
2. The accuracy of the load forecasts was comparable using the NN and regression model. However, the differences between the load modeling techniques were in the pre-processing (i.e, obtaining the optimal combination of predictors and transformations to achieve the most accurate load forecast) and the required computing power. The NN required less pre-processing and more computing power. On the other hand, the regression model required less computing power and more pre-processing. Thus, the load modeler must determine how much time is available for pre-processing and the availability of adequate computing resources.
3. Improved temperature forecasts were tested in a load forecast simulation for the Norman substation. The result was an improved load forecast that translated into a $16.50/day savings at the Norman substation. However, Norman is only one of 225 substations in the WFEC system. Because it was the largest substation (2000)[1], the

[1] The Norman substation recorded a yearly peak of nearly 20000 kWh while the yearly peak of the smallest substations in the WFEC system averaged ~1000 kWh in 2000.

savings at smaller substations are likely to be less than $16.50/day. Even so, on a system-wide basis, improved temperature forecasts could save WFEC in excess of $500,000 annually, or a 10-15% improvement to the bottom line of an operating for-profit-company. As a result of this study, WFEC is empowered to make smarter generation decisions that decrease operating costs, allow profitable sales of bulk power, and reduce the number of purchases of power from the spot market.

4. *The value of understanding the interrelationships between weather and load and the value of improved temperature forecasts as explored in this study can save a small utility at least $0.5M annually.* However, the generation capacity and net margin of typical electric cooperatives like WFEC is small (~1000 MW and $3.4M, respectively) compared to those of investor-owned powerhouses like Florida Power and Light (~19,000 MW and $473M)[2]. Khotanzad et al. (1998) found that "weather forecasts introduce 1% of additional error in load forecasts", and Hobbs et al. (1999) followed with "a conservative estimate is that a 1% reduction in forecasting error can save [one utility] up to $1.6 million annually." *Therefore, if these reductions in operating expenses are replicated nationwide, it is easy to see how improved weather forecasts on the short term can translate into millions of dollars in annual savings for utility companies in the United States.*

This dissertation has identified several relationships between weather and load that had not been studied due to the fact that substation-level load data is proprietary and

[2] The generation capacity and net margin for WFEC and Florida Power and Light were obtained from their 2002 Annual Reports.

had not been readily available, the weather resources used in practice had been limited, and the cooperation between utility corporations and the (scientific) meteorological community had occurred on a hit or miss basis. Clearly, an improvement to our understanding of the weather-load relationship will help determine parameters that should be used to develop load models that produce more accurate forecasts. The methods used in this study to improve load forecasts at WFEC are applicable across the United States and perhaps worldwide.

A more accurate load forecast allows the generation utilities to produce the optimal amount of electricity to meet the load demands of it customers. These generation decisions by experienced load forecasters minimize the need to buy electricity on the spot market at inflated prices (Hunn 2000) due to an underproduction or to be forced to sell electricity at bargain prices due to an overproduction. However, a more accurate load forecast requires increased knowledge of the factors which impact the load demand. Aside from the typical base load and weekday versus weekend differences, the fluctuations in the actual load are primarily caused by rapidly changing the weather – especially during transition seasons. Because a load forecast model is only as good as its developmental data sets, an improved understanding of the relationship between weather and electric load demand are crucial to load forecasting, in that the best combination of historical and forecasted information must be used. The hypothesis of this dissertation stated that *a comprehensive understanding of the relationship between weather variables and electricity demand will improve the accuracy of load forecasting and have a positive economic benefit.* Based upon the results described herein, this hypothesis is accepted.

References

Abu-El-Magd, M. A., and N. Sinha, 1981: Two new algorithms for on-line modeling and forecasting of the load demand of a multinode power system. *IEEE Trans. Power Apparatus and Systems,* **PAS-100,** 3246-3253.

Abu-El-Magd, M. A., and N. Sinha, 1982: Short-term load demand modeling and forecasting: A review. *IEEE Trans. Syst. Man Cybern.,* **SMC-12,** 370-382.

Adya, M., and F. Collopy, 1998: How effective are neural networks at forecasting and prediction? A review and evaluation. *J. Forecast.,* **17,** 481-495.

AlFuhaid, A. S., M. A. El-Sayed, and M. S. Mahmoud, 1997: Cascaded artificial neural networks for short-term load forecasting. *IEEE Trans. Power Systems,* **12,** 1524-1529.

Alves da Silva, A. P., 2000: Confidence intervals for neural network based short-term load forecasting. *IEEE Trans. Power Systems,* **15,** 1191-1196.

Barakat, E. H., and J. M. Al-Qasem, 1998: Methodology for weekly load forecasting. *IEEE Trans. Power Systems,* **13,** 1548-1555.

Benjamin, S. G., J. M. Brown, K. J. Brundage, B. E. Schwartz, T. B. Smirnova, T. L. Smith, and L. L. Morone, 1988: RUC-2 – The Rapid Update Cycle Version 2. *NWS Technical Procedure Bulletin No. 448.* NOAA/NWS, 18 pp.

Benjamin, S. G., J. M. Brown, K. J. Brundage, D. Dévényi, G. A. Grell, D. Kim, B. E. Schwartz, T. B. Smirnova, T. L. Smith, and S. S. Weygandt, 2002: RUC20 – The 20-km version of the Rapid Update Cycle. *NWS Technical Procedure Bulletin No. 490.* NOAA/NWS, 30 pp.

Black, T. L., 1994: The new NMC mesoscale Eta model: description and forecast examples. *Wea. Forecasting,* **9,** 265-278.

Box, G. E. P., J. M. Jenkins, and G. C. Reinsel, 1994: *Time Series Analysis: Forecasting and Control.* Prentice-Hall, Inc., Upper Saddle River, NJ, 598 pp.

Brock, F. V., K. C. Crawford, R. L. Elliott, G. W. Cuperus, S. J. Stadler, H. L. Johnson, and M. D. Eilts, 1995: The Oklahoma Mesonet: a technical overview. *J. Atmos. Oceanic Technol.,* **12,** 5-19.

Brotzge, J. A., and S. J. Richardson, 2003: Spatial and temporal correlation among Oklahoma Mesonet and OASIS surface-layer measurements. *J. Appl. Meteor.,* **42,** 5-19.

Brubacher, S. R., and G. T. Wilson, 1976: Interpolating time series with applications to the estimation of the holiday effects on electricity demand. *Appl. Stat.*, **25,** 107-116.

Bunn, D. W., and E. D. Farmer (eds.), 1985: *Comparative Models for Electric Load Forecasting*. John Wiley & Sons Ltd., New York, NY, 232 pp.

Bunn, D. W., 2000: Forecasting loads and prices in competitive power markets. *Proc. IEEE,* **88,** 163-169.

Charney, J. G., R. Fjørtoft, and J. von Neumann, 1950: Numerical integration of the barotropic vorticity equation. *Tellus*, **2**, 237-254.

Charney, J. G., 1954: Numerical prediction of cyclogenesis. *Proc. Nat. Acad. Sci. U. S.,* **40,** 99-110.

Charytoniuk, W., M. S. Chen, and P. V. Olinda, 1998: Nonparametric regression based short-term load forecasting. *IEEE Trans. Power Systems,* **13,** 725-730.

Charytoniuk, W., E. D. Box, W. Lee, M. Chen, P. Kotas, and P. V. Olinda, 2000: Neural-network-based demand forecasting in a deregulated environment. *IEEE Trans. Industry Applications,* **36,** 893-898.

Chen, F., and Dudhia, J., 2001: Coupling an advanced land-surface hydrology model with the Penn State-NCAR MM5 modeling system. Part II: Preliminary model validation. *Mon. Wea. Rev.*, **129,** 587-604.

Chen, S., D. C. Yu, and A. R. Moghaddamio, 1992: Weather sensitive short-term forecasting using nonfully connected artificial neural network. *IEEE Trans. Power Systems,* **7,** 1098-1105.

Choueiki, M. H., C. A. Mount-Campbell, and S. C. Ahalt, 1997: Building a 'quasi optimal' neural network to solve the short-term load forecasting problem. *IEEE Trans. Power Systems,* **12,** 1432-1439.

Cressman, G. P., 1963: A three-level model suitable for daily numerical forecasting. NMC Tech. Memo. No. 22, Weather Bureau, ESSA, U.S. Department of Commerce, 22 pp. [NOAA Central Library, 6009 Executive Boulevard, Rockville, MD 20862.]

Cristiaanse, W. R., 1971: Short-term load forecasting using general exponential smoothing. *IEEE Trans. Power Apparatus and Systems,* **PAS-90,** 900-911.

Darbellay, G. A., and M. Slama, 2000: Forecasting the short-term demand for electricity - Do neural networks stand a better chance? *Int. J. Forecast.*, **16,** 71-83.

Davies, M., 1958: The relationship between weather and electricity demand. *Proc. IEEE,* **106C,** 27-37.

Douglas, A. P., A. M. Breipohl, F. N. Lee, and R. Adapa, 1998: Impacts of temperature forecast uncertainty on Baysian Load Forecasting. *IEEE Trans. Power Systems,* **13,** 1507-1513.

Drezga, I., and S. Rahman, 1998: Input variable selection for ANN-based short-term load forecasting. *IEEE Trans. Power Systems,* **13,** 1238-1244.

Dudhia, J., 1993: A nonhydrostatic version of the Penn State-NCAR mesoscale model: Validation tests and simulation of an Atlantic cyclone and cold front. *Mon. Wea. Rev.*, **121,** 1493-1513.

Dutton, J. A., 1986: *Dynamics of Atmospheric Motion.* Dover Publications, Inc., New York, NY, 617 pp.

Ernoult, M., and R. Mattatia, 1984: Short term load forecasting: new developments at E.D.F. *Proc. 8th Power Systems Computation Conf.*, PSCC'84, 369-375.

Farmer, E. D., and M. J. Potton, 1968: Development of on-line load prediction techniques with results from trials in the south-west region of the CEGB. *Proc. Inst. Elec. Eng.*, **115,** 1549-1558.

Gertler, J., and C. S. Banyasz, 1974: A recursive (on-line) maximum likelihood identification method. *IEEE Trans. Automat. Contr.*, **AC-19,** 816-820.

Galiana, F. D., E. Handschin, and A. R. Fiechter, 1974: Identification of stochastic load models from physical data. *IEEE Trans. Automat. Contr.*, **AC-19,** 887-893.

Glahn, H. R., and D. A. Lowry, 1972: The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteor.*, **11,** 1203-1211.

Gross, G., and F. D. Galiana, 1987: Short-term load forecasting. *Proc. IEEE,* **75,** 1558-1573.

Gupta, P. C., 1971: A stochastic approach to peak power demand forecasting in electric utility systems. *IEEE Trans. Power Apparatus and Systems,* **PAS-90,** 825-832.

Gupta, P. C., and K. Yamada, 1972: Adaptive short-term forecasting of hourly loads using weather information. *IEEE Trans. Power Apparatus and Systems,* **PAS-91,** 2085-2094.

Guttman, N. B., and R. G. Quayle, 1996: A historical perspective of U.S. climate divisions. *Bull. Amer. Meteor. Soc.*, **77,** 293-303.

Hagan, M. T., and R. Klein, 1978: On line maximum likelihood estimation for load forecasting. *IEEE Trans. Syst. Man Cybern.*, **SMC-8,** 711-715.

Hagan, M. T., and S. M. Behr, 1987: The time series approach to short term load forecasting. *IEEE PES Winter Meeting*, New Orleans, LA, Paper 87 WM 044-1.

Hagan, M. T., H. B. Demuth, and M. Beale, 1996: *Neural Network Design.* PWS Publishing Company, Boston, MA, 680 pp.

Hall, T., H. E. Brooks, and C. A. Doswell III, 1999: Precipitation forecasting using a neural network. *Wea. Forecasting,* **14,** 338-345.

Heinemann, G. T., D. A. Nordman, and E. C. Plant, 1966: The relationship between summer weather and summer loads – a regression analysis. *IEEE Trans. Power Apparatus and Systems,* **PAS-85,** 1144-1154.

Hippert, H. S., C. E. Pedreira, and R. C. Souza, 2001: Neural networks for short-term load forecasting: a review and evaluation. *IEEE Trans. Power Systems,* **16,** 44-55.

Ho, K., Y. Hsu, C. Chen, T. Lee, T. Lai, and K. Chen, 1990: Short-term load forecasting of Taiwan power system using a knowledge-based expert system. *IEEE Trans. Power Systems,* **5,** 1214-1221.

Hobbs, B. F., S. Jitprapaikulsarn, S. Konda, V. Chankong, K. A. Loparo, and D. J. Maratukulam, 1998: Analysis of the value for unit commitment of improved load forecasts. *IEEE Trans. Power Systems,* **14,** 1342-1348.

IEEE Committee, 1980: Load forecast bibliography. *IEEE Trans. Power Apparatus and Systems,* **PAS-99,** 53-58.

Infield, D. G., and D. C. Hill, 1998: Optimal smoothing for trend removal in short term electricity demand forecasting. *IEEE Trans. Power Systems,* **13,** 1115-1120.

Irisarri, G. B., S. E. Widergren, and P. D. Yehsakul, 1982: On-line load forecasting for energy control center application. *IEEE Trans. Power Apparatus and Systems,* **PAS-101,** 71-78.

Jacks, E., J. B. Bower, V. J. Dagostaro, J. P. Dallavalle, M. C. Erickson, and J. C. Su, 1989: New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind. *Wea. Forecasting,* **5,** 128-138.

Kalnay, E., S. J. Lord, and R. D. McPherson, 1998: Maturity of operational numerical weather prediction: medium range. *Bull. Amer. Meteort. Soc.*, **79,** 2753-2769.

Kalnay, E., 2003: *Atmospheric Modeling, Data Assimilation and Predictability.* Cambridge University Press, New York, NY, 341pp.

Keyhani, A., and S. M. Miri, 1983: On-line weather sensitive and industrial group bus load forecasting for microprocessor based applications. *IEEE Trans. Power Apparatus and Systems,* **PAS-102,** 3868-3876.

Keyhani, A., and A. El-Abiad, 1975: One-step ahead load forecasting for on-line applications. *IEEE Winter Power Meeting,* IEEE, 1092.

Keyhani, A., A. El-Abiad, and E. G. Lansing, 1975: Dynamic system load generation control using variable pressure steam generators. *IEEE Summer Power Meeting,* IEEE, 1924-1925.

Khotanzad, A., R. Afkhami-Rohani, and D. Maratukulam, 1998: ANNSTLF – Artificial neural network short-term load forecaster – generation three. *IEEE Trans. Power Systems,* **13,** 1413-1422.

Klein, W. H., B. M. Lewis, and I. Enger, 1959: Objective prediction of five-day mean temperature during winter. *J. Meteor.,* **16,** 672-682.

Kuligowkski, R. J., and A. P. Barros, 1998: Localized precipitation forecasts from a numerical weather prediction model using artificial neural networks. *Wea. Forecasting,* **13,** 1194-1204.

Lamedica, R., A. Prudenzi, M. Sforna, M. Caciotta, and V. O. Cencelli, 1996: A neural network based technique for short-term forecasting of anomalous load periods. *IEEE Trans. Power Systems,* **11,** 1749-1756.

Lee, K. Y., Y. T. Cha, and J. H. Park, 1992: Short-term load forecasting using an artificial neural network. *IEEE Trans. Power Systems,* **7,** 124-132.

Lijesen, D. P., and J. Rosing, 1971: Adaptive forecasting of hourly loads based on load measurement and weather information. *IEEE Trans. Power Apparatus and Systems,* **PAS-90,** 1757-1767.

Ljung, L., and T. Soderstrom, 1983: *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 529 pp.

Lu, C. N., H. T. Wu, and S. Vemuri, 1993: Neural network based short term load forecasting. *IEEE Trans. Power Systems,* **8,** 336-342.

Marzban, C., and G. J. Stumpf, 1996: A neural network for tornado prediction based on Doppler radar – Derived attributes. *J. Appl. Meteor.,* **35,** 617-626.

Matthewman, P. D., and H. Nicholson, 1968: Techniques for load prediction in the electricity-supply. *Proc. Inst. Elec. Eng.*, **115,** 1451-1457.

McCann, D. W., 1992: A neural network short-term forecast of significant thunderstorms. *Wea. Forecasting,* **7,** 525-534.

McQuigg, J. D., S. R. Johnson, and J. R. Tudor, 1972: Meteorological diversity – Load diversity, a fresh look at an old problem. *J. Appl. Meteor.,* **11,** 561-566.

Mehrotra, K., C. K. Mohan, and S. Ranka, 1997: *Elements of Artificial Neural Networks.* MIT Press, Cambridge, MA, 344 pp.

Moghram, I., and S. Rahman, 1989: Analysis and evaluation of five short-term load forecasting techniques. *IEEE Trans. Power Systems,* **4,** 1484-1491.

Morris, J. W., 1977: *Geography of Oklahoma.* Oklahoma Historical Society, Oklahoma City, OK, 182 pp.

Papalexopoulos, A. D., and T. C. Hesterberg, 1990: A regression-based approach to short- term system load forecasting. *IEEE Trans. Power Systems,* **5,** 1535-1547.

Park, J. H., Y. M. Park, and K. Y. Lee, 1991: Composite modeling for adaptive short-term load forecasting. *IEEE Trans. Power Systems,* **6,** 450-457.

Pickles, J. H., 1975: Automatic load prediction by the spectral analysis method. *Proc. 5th Power Systems Computations Conf.,* PSCC'75, 2.4/1.

Poysti, K., 1984: Box-Jenkins method in short-term forecasting of grid load in Finland. *Proc. 8th Power Systems Computations Conf,* PSCC'84, 357-368.

Rahman S., and O. Hazim, 1993: A generalized knowledge-based short-term load-forecasting technique. *IEEE Trans. Power Systems,* **8,** 508-514.

Rahman, S., and R. Bhatnagar, 1988: An expert system based algorithm for short term load forecast. *IEEE Trans. Power Systems,* **3,** 392-399.

Rastogi, S. and G. Roulet, 1994: Effect of weather on rural electric cooperative demand. *Rural Electric Power Conference, 38th Annual Conference,* IEEE, A3/1-A3/6.

Roebber, P. J., S. L. Bruening, D. M. Schultz, and J. V. Cortinas Jr., 2002: Improving snowfall forecasting by diagnosing snow density. *Wea. Forecasting,* In review.

Saaty, T. L., 1980: *The Analytic Hierarchy Process.* McGraw-Hill, New York, NY, 296 pp.

Shafer, M. A., C. A. Fiebrich, D. S. Arndt, S. E. Fredickson, and T. W. Hughes, 2000: Quality assurance procedures in the Oklahoma Mesonetwork. *J. Atmos. Oceanic Technol.*, **17,** 474-494.

Shuman, F. G., 1978: Numerical weather prediction. *Bull. Amer. Meteor. Soc.*, **59,** 5-17.

Shuman, F. G., 1989: History of numerical weather prediction at the National Meteorological Center. *Wea. Forecasting,* **4,** 286-296.

Späth, H., G. L. Thompson, and H. Eisenhart, 1998: *Oklahoma Resources for Economic Development.* Oklahoma Geological Survey, The University of Oklahoma, Norman, OK, 266 pp.

Stanton, K. N., and P. C. Gupta, 1970: Forecasting annual or seasonal peak demand in electric utility systems. *IEEE Trans. Power Apparatus and Systems,* **PAS-89,** 951-959.

Stull, R. B., 1995: *Meteorology Today for Scientists and Engineers.* West Publishing Company, St. Paul, MN, 385 pp.

Thompson, P. D., and W. L. Gates, 1956: A test of numerical prediction methods based on the barotropic and two-parameter baroclinic models. *J. Meteor.*, **13,** 127-141.

Vermuri, S., D. F. Hill, and R. Balasubramanian, 1973: Load forecasting using stochastic models. *IEEE PICA Conf.*, IEEE, 31-37.

Van Meeteren, H. P. and P. J. M. Van Son, 1979: Short-term load prediction with a combination of different models. *IEEE PICA Conf.*, IEEE, 192-197.

Vermaak, J., and E. C. Botha, 1998: Recurrent neural networks for short-term load forecasting. *IEEE Trans. Power Systems,* **13,** 126-132.

Wikle, T., 1991: *Atlas of Oklahoma* (Classroom Edition). Computer Cartography Laboratory, Dept. of Geography, Oklahoma State University, Stillwater, OK, 180 pp.

Zhang, G., B. E. Patuwo, and M. Y. Hu, 1998: Forecasting with artificial neural networks: the state of the art. *Int. J. Forecast.*, **14,** 35-62.

Appendix A: How Electricity Is Produced

The electric power system is composed of three stages before electricity reaches the customer: (1) generation, (2), transmission, and (3) distribution. First, the electricity must be produced (Fig. A.1). This is called "generation". At any power plant, the "generator" is the devise that actually makes the electricity. Generators have two main parts, a rotor and a stator. The rotor is a long, heavy cylinder-shaped magnet that spins inside a housing of coiled copper wire, known as the stator. As the rotor spins, electric current is created in each coil. Electricity flows from the coils by other wires and begins its trek along various transmission lines.

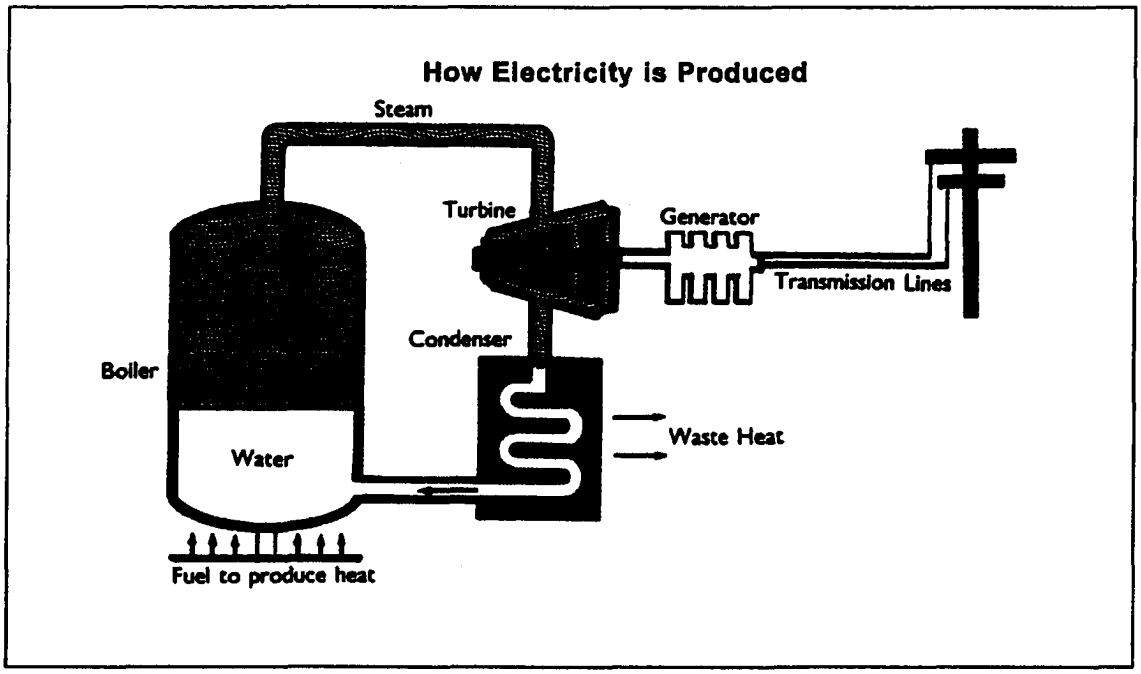


FIG. A.1 A diagram of the system that produces electricity (illustration provided by WFEC).

Getting the rotor to spin requires a lot of power. Most plants generate this power using high-pressure steam. Power plants often use coal, natural gas and/or oil to heat water until it boils and produces steam. The fire from the burning fuel is contained inside a boiler, whose walls are made of hundreds of hollow pipes welded together side-by-side and connected on the ends. Water is dispersed inside the boiler walls and the burning gas/coal/oil heats the water until it becomes steam.

The steam is piped into the turbine, a windmill or fan-like structure composed of many blades tough enough to withstand the hot, high-pressure steam that turns it. The steam turbine is covered in a thick steel casing. The steam forces the blades to move very rapidly (i.e., the tips of the blades move faster than the speed of sound). As the turbine spins, so does the rotor because they are connected by a thick shaft.

Once the turbine is rotating, the steam is piped to a condenser where it is cooled and undergoes a phase change back to liquid water. The condenser removes heat from the steam by discharging that heat into the air through cooling towers. As the heat rises through the cooling tower, large white clouds form above power plants. Meanwhile, the resultant liquid water is piped back into the boiler and the process is repeated.

Electricity has been generated and is moving through the transmission lines. This step is called transmission (i.e., the transfer of electricity or natural gas from a generation plant or pipeline to another facility). The electricity moves through a transformer to increase the voltage for a more economical trip to a substation. Switches in a substation direct the flow of electricity into the distribution system of member co-ops. Technicians and computers determine if the correct amount of electricity is delivered to each substation by constantly monitoring the transmission system. When a storm damages a

transmission line, WFEC technicians use remote-controlled switches to remove the damaged portion of the network from the rest of the system and re-route the electricity to the appropriate substation. Each substation is composed of transformers that reduce the voltage for the distribution cooperative.

Once the electricity passes into the substation, the distribution process begins. Distribution is defined as the delivery of electricity to an end-user through low-voltage lines or natural gas through pipeline systems. The voltage transformers on poles near businesses reduce the voltage to either 240 or 120 volts so it can be used in machinery or appliances. The electricity passes through a user's meter which records the amount of electricity used, then into the member's own distribution panel and home electrical system. At this point, the power is divided into several circuits that serve different locations in a facility. For example, when a light switch is turned on, the entire system is affected. A little more current flows through the user's meter from the distribution transformer, which was fed by the substation and transmission line from the generator.

To maintain a constant output of current (rate at which electricity flows), the rotor must turn at a constant speed. The demand for current makes the rotor slow. To maintain its speed more steam must be supplied to the turbine. Computers sense the demand for more electricity and open valves that supply fuel to a boiler to maintain the speed of a rotor.

Thus, hundreds of people work around the clock to supply power to our homes or businesses. Linemen who climb the poles, and substation technicians who work with huge transformers, during transmission, risk their lives to repair components that stopped flow of electricity.

Appendix B: Oklahoma Rural Electric Cooperatives

A co-op is a nonprofit utility owned by its members and generally serve rural communities. In Oklahoma, co-ops serve 75% of the territorial domain (Fig. B.1). The remaining 25% of Oklahoma is served either by Oklahoma Gas and Electric (OG&E) or the Public Service Company of Oklahoma (PSO), which are the investor-owned utilities. The Oklahoma cooperatives have 106,000 miles of distribution lines and serve about 500,000 customers. On the other hand, OG&E and PSO combined have 46,000 miles of distribution lines and serve nearly 1.3 million customers. Hence, the electrical rates are slightly higher in a co-op versus an investor-owned utility. However, because the members/customers of the co-op have ownership in the co-op, any "profit" that is made by the company at the end of the year is returned to its members.

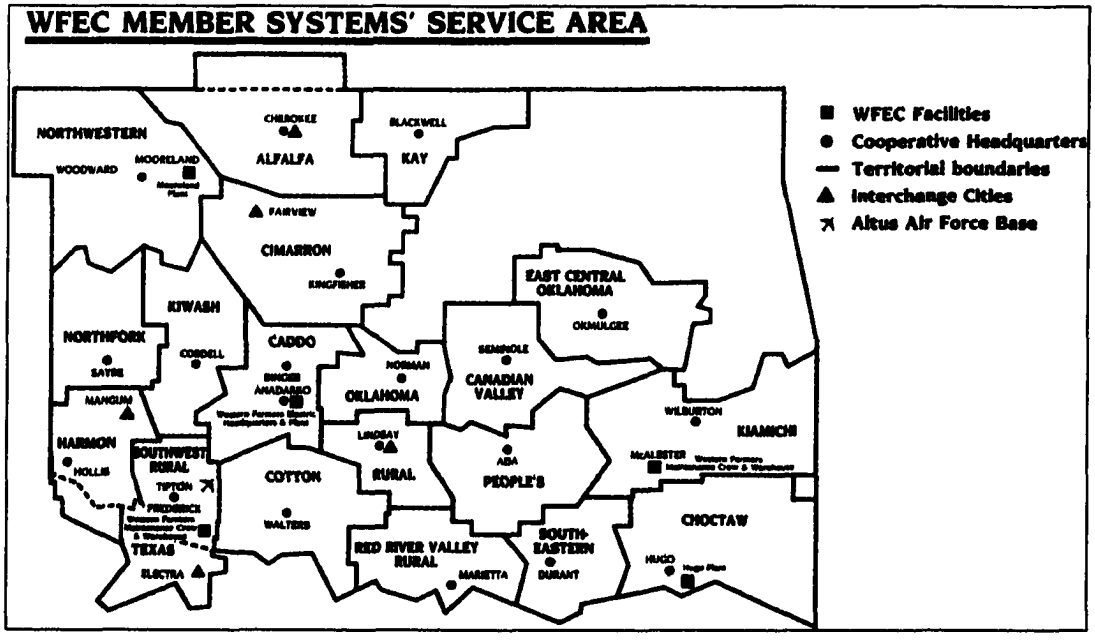


FIG B.1 A map of WFEC member system's service area in Oklahoma (illustration provided by WFEC).

Appendix C: Glossary of Electric Utility Terms

Capacity – The real power output rating of a generator or system, typically in megawatts, measured on an instantaneous basis. The amount of electric power delivered or required for which a generator, transformer, transmission circuit, station or system is rated by the manufacturer. The maximum power that can be produced by a generating resource at specified times under specified conditions. Capacity is also used synonymously with capability.

Degree day - a quantitative index demonstrated to reflect demand for energy to heat or cool houses and businesses. This index is derived from daily temperature observations at nearly 200 major weather stations in the contiguous United States. The "heating year" during which heating degree days are accumulated extends from July 1st to June 30th and the "cooling year" during which cooling degree data are accumulated extends from January 1st to December 31st. A mean daily temperature (average of the daily maximum and minimum temperatures) of 65°F is the base for both heating and cooling degree day computations. Heating degree days are summations of negative differences between the mean daily temperature and the 65°F base; cooling degree days are summations of positive differences from the same base. (www.cpc.ncep.noaa.gov)

Generating Unit – any combination of physically connected generator(s), reactor(s), boiler(s), combustion turbine(s), or other prime mover(s) operated together to produce electric power.

Interruptible Load (interruptible energy) – Energy flow that can be reduced or completely stopped with little or no notice. Interruptible energy is a type of electrical service sold to firms that might only be able to operate profitably when energy prices remain below a certain level. When interruptible energy is purchased, the purchaser voluntarily assumes the risk of loss of access to that energy, which usually occurs only during peak demand periods or during periods when market prices rise above the agreed-upon rate. Commercial, industrial and agricultural customers tend to be the first affected by this type of interruption.

Load – The amount of electric power delivered or required at any specific point or points on a system. The requirement originates at the energy-consuming equipment of the consumers. The load of an electric utility system is affected by many factors and changes on a daily, seasonal and annual basis, typically following a pattern. System load is usually measured in megawatts (MW).

Load Curve – A curve of power versus time showing the level of a load for each time period covered. The horizontal axis is time and the vertical axis is load (kW).

Load Factor – The ratio of average load to peak load during a specified period of time, expressed as a percent. The load factor indicates to what degree energy has been consumed compared to maximum demand or the utilization of units relative to total system capability. An electric system's load factor shows the variability in all customer's demands.

Load Shape – The variation in the magnitude of the power load over a daily, weekly or annual period.

Load Shedding – blocking of customer access to energy, usually due to temporary shortage of supply. Load shedding is rare, and is most commonly applied during times of emergency or severe shortage. In most cases, the first loads a utility will shed in these conditions are loads required by industrial and commercial customers. Institutional loads are typically the last to be shed since public institutions (hospitals, schools, municipal lighting authorities, etc.) are considered to be a utility's most essential customers.

Kilowatt (kW) – A unit of electrical power equal to one thousand watts.

Kilowatt-hour (kWh) – A unit of electrical energy which is equivalent to one kilowatt of power used for one hour. One kilowatt-hour is equal to 1,000 watt-hours. An average household will use between 800 - 1,300 kWh per month depending upon geographical area.

Peak Load – Denotes the maximum power requirement of a system at a given time, or the amount of power required to supply customers at times when need is greatest. It can refer either to the load at a given moment (e.g. a specific time of day) or to averaged load over a given period of time (e.g. a specific day or hour of the day).

Real-time Pricing – electricity pricing based on the actual, fluctuating price of electricity at time throughout the day, impacted by demand and weather

Short-term load forecast (STLF) – Predictions of electric load on the order of one day to one week. *For this research, STLFs will refer to forecasts on the order of one day.*

Spinning Reserve – Unused capacity available from units connected to and synchronized with the grid to serve additional demand. The spinning reserve must be under automatic governor control to quickly (within minutes) respond to system requirements. Spinning is derived from hydroelectric and combustion turbine terminology. Reserve generator turbines can literally be kept spinning without producing any energy as a way to reduce the length of time required to bring them online when needed.

Spot Market – A market where (electrical) goods are traded for immediate delivery.

Substation – facility equipment that switches, changes or regulated voltage

Volt – The unit of measurement of electromotive force. It is equivalent to the force required to produce a current of one ampere through a resistance of one ohm. The unit of measure for electrical potential. Generally measured in kilovolts or kV. Typical transmission level voltages are 69kV or 138kV.

Watt – A measurement of real power production or usage equal to one joule per second. The rate of energy transfer equivalent to one (1) ampere flowing under a pressure of one (1) volt at unity power factor. An electric unit of power or a rate of doing work.

Watt-hour (Wh) – An electrical energy unit of measure equal to one (1) watt of power supplied to, or taken from, an electric circuit steadily for one hour.

Appendix D: Heat Index and Wind Chill

Heat Index

The heat index is the temperature a body feels when air temperature and relative humidity are combined. Our human bodies dissipate heat through a loss of water from the skin and sweat glands because evaporation is a cooling process. During the warmer months, high humidity makes the air feel hotter by reducing the evaporation/cooling process. The heat index formula is a multiple linear regression equation which combines relative humidity and temperature to ascertain a human-perceived temperature:

$$\begin{aligned}\text{Heat Index } (^\circ\text{F}) = & -42.379 + 2.04901523 * T + 10.14333127 * RH \\ & - 0.22475541 * T * RH - 6.83783 * 10^{-3} * T^2 \\ & - 5.481717 * 10^{-2} * RH^2 + 1.22874 * 10^{-3} * T^2 * RH \\ & + 8.5282 * 10^{-4} * T * RH^2 - 1.99 * 10^{-6} * T^2 * RH^2 ,\end{aligned}$$

where T is temperature (°F) and RH is relative humidity (%)[1].

Wind Chill

Wind chill is a hypothetical air temperature that measures of how cold people and animals feel. During the winter, faster winds make the air feel colder because it removes heat from our bodies faster than would occur if the winds were calm. This rapid loss of heat from the body decreases the skin temperature and ultimately the body's internal temperature. The wind chill formula was derived by modifying the heat transfer equation

[1] http://www.usatoday.com/weather/whumcalc.htm

for a flux across a surface. The National Weather Service defines the wind chill formula as:

$$\text{Wind Chill}(^\circ\text{F}) = 35.74 + 0.6215 * T - 35.75 * V^{0.16} + 0.4275 * T * V^{0.16},$$

where T is temperature (°F) and V is the wind speed (mph)[2].

[2] http://www.crh.noaa.gov/dtx/New_Wind_Chill.htm

Appendix E: Missing Mesonet and Electric Load Data

Table E.1 Gaps in the Mesonet data set from Norman filled by temporal interpolation or by substituting data from the Washington Mesonet site. Otherwise, observations were labeled as missing.

Date	*Time (24-hour clock)*	*Interpolated (I) Substituted (S) Missing (M)*	*Weather Variables*
15 Jun 1998	0800	I / S	T, Wspd, Srad / RH
17 Jun 1998	1800	I	T, RH, Wspd, Srad
18 Jun 1998	1200-1500	S	T, RH
	1600-2400	S	RH
19 Jun 1998	0100-1100	S	RH
	1200	I	RH
29 Jul 1998	0900,1000	S	T, RH, Wspd, Srad
10 Sep 1998	0800-1000	M	T, RH, Wspd, Srad
18 Sep 1998	2000	I	T, RH, Wspd, Srad
5 Oct 1998	0800-1300	S	T, RH, Wspd, Srad
4 Nov 1998	1600	I	T, RH, Wspd, Srad
	1700-2400	S	T, RH, Wspd, Srad
5 Nov 1998	0100-0400	S	T, RH, Wspd, Srad
	0500-1000	M	T, RH, Wspd, Srad
11 Nov 1998	1400, 1500	S	T, RH, Wspd, Srad
17 Nov 1998	0500	I	T, RH, Wspd, Srad
	0600, 0700	S	T, RH, Wspd, Srad
30 Mar 1999	1200	I	T, RH, Wspd, Srad
31 Mar 1999	1200	M	RH
	1300	S	T, RH, Wspd, Srad
1 Apr 1999	0500-1000	I	T, RH, Wspd, Srad
3 May 1999	1000, 1100	S	T, RH, Wspd, Srad
9 May 1999	1900-2400	S	Wspd
10 May 1999	0100-1500	S	Wspd
16 May 1999	0800	I	T, RH, Wspd, Srad
11 Jun 1999	0900-2400	S	Wspd
12 Jun 1999	0100-0500	S	Wspd
16-17 Jun 1999	0100-2400	S	Wspd
18 Jun 1999	0100-0600	S	Wspd
	0700-2000	S	T, RH, Wspd, Srad
23 Jun 1999	1200-1500	I	T, RH, Wspd, Srad
25 Jun 1999	1700	I	T, RH, Wspd, Srad
26 Jun 1999	2300, 2400	S	T, RH, Wspd, Srad
27 Jun 1999	0100-1100	S	T, RH, Wspd, Srad
27 Jun 1999	1400, 1700	I	T, RH, Wspd, Srad
1 Jul 1999	1400	I	T, RH, Wspd, Srad
19 Jul 1999	0600, 0700	S	T, RH, Wspd, Srad
	1000, 1200	I	T, RH, Wspd, Srad
25 Dec 2000	1000-2400	M	Wspd
26 – 27 Dec 2000	0100-2400	M	Wspd
28 Dec 2000	0100-1300	M	Wspd

Table E.2 As in Table E.1, except for the Woodward Mesonet Site with substitutions of data from the Seiling Mesonet Site.

Date	*Time (24-hour clock)*	*Interpolated (I) Substituted (S) Missing (M)*	*Weather Variables*
8 Jul 1998	0300-1900	M	T, RH, Wspd, Srad
	2000	S	T, RH, Wspd, Srad
23 Jul 1998	1000	I	Srad
10 Sep 1998	0900-1100	M	T, RH, Wspd, Srad
17 Oct 1998	0900-1200	M	T, RH, Wspd, Srad
22 Oct 1998	1100	I	T, RH, Wspd, Srad
18 Dec 1998	2200	I	T, RH, Wspd, Srad
24 Aug 1999	0900-2400	M	T, RH, Wspd, Srad
25 Aug 1999	0100	M	T, RH, Wspd, Srad
28 Aug 1999	1200	I	T, RH, Wspd, Srad
29 Aug 1999	0700	I	T, RH, Wspd, Srad
17 Sep 1999	1000, 1100	M	T, RH, Wspd, Srad
18 Sep 1999	2100	S	T, RH, Wspd, Srad
28 Sep 1999	0500	I	T, RH, Wspd, Srad
29 Sep 1999	0800, 0900	M	T, RH, Wspd, Srad
	1000	S	T, RH, Wspd, Srad
13 Oct 1999	1600, 1700	S	T, RH, Wspd, Srad
19 Oct 1999	0800-1800	M	Srad
4 Nov 1999	1100	I	T, RH, Wspd, Srad
26 Dec 1999	0800-1700	M	Srad
27 Dec 1999	0800-1700	M	Srad
28 Dec 1999	0800-1700	M	Srad
13 Mar 2000	1300	I	T, RH, Wspd, Srad
13 Apr 2000	1600	I	T, RH, Wspd, Srad
16 Mar 2000	1200-1400	S	T, RH, Wspd, Srad
22 Aug 2000	1600	S	T, RH, Wspd, Srad

Table E.3 As in Table E.1, except for the Altus Mesonet Site with substitutions of data from the Tipton Mesonet Site.

Date	*Time (24-hour clock)*	*Interpolated (I) Substituted (S) Missing (M)*	*Weather Variables*
24 Aug 1999	1400	I	T, RH, Wspd, Srad
28 Aug 1999	1200	I	T, RH, Wspd
18 Oct 1999	1900	M	Srad
19 Oct 1999	0800-1800	M	Srad
10 Nov 1999	1000, 1100, 1700	S	T, RH, Wspd, Srad
26 Dec 1999	0800-1700	M	Srad
27 Dec 1999	0800-1700	M	Srad
28 Dec 1999	0800, 1700	M	Srad
2 Feb 2000	1200-1600	S	T, RH, Wspd, Srad
3 Mar 2000	1200	I	T, RH, Wspd, Srad
24 Mar 2000	1200-1500	S	T, RH, Wspd, Srad
26 May 2000	1400-1600	S	T, RH, Wspd, Srad
30 Aug 2000	1600, 1700	S	T, RH, Wspd, Srad
26 Oct 2000	0600-0800	S	RH
20 Nov 2000	0700	M	T
		S	RH, Wspd, Srad
25 Dec 2000	1400-1600	S	Wspd
	1700-2400	M	Wspd
26 Dec 2000	0100-2400	M	Wspd
27 Dec 2000	0100-1300	M	Wspd

Table E.4 As in Table E.1, except for the Broken Bow Mesonet Site with substitutions of data from the Idabel Mesonet Site.

Date	*Time (24-hour clock)*	*Interpolated (I) Substituted (S) Missing (M)*	*Weather Variables*
25 Apr 1999	2000	I	T, RH, Wspd, Srad
24 Aug 1999	0900-1300, 2000	M / I	T, RH, Wspd, Srad
28 Aug 1999	1200	I	T, RH, Wspd, Srad
19 Oct 1999	0800-1800	M	Srad
2 Nov 1999	1500	I	T, RH, Wspd, Srad
7 Nov 1999	1500	I	T, RH, Wspd, Srad
26 Dec 1999	0800-1700	M	Srad
27 Dec 1999	0800-1700	M	Srad
28 Dec 1999	0800-1700	M	Srad
10 Feb 2000	1100	I	T, RH, Wspd, Srad
21 Jun 2000	1100	I	T, RH, Wspd, Srad
23 Jun 2000	1900-2400	S	RH
24-27 Jun 2000	0100-2400	S	RH
28 Jun 2000	0100-0800, 1300,		
	1700- 2400	S	RH
	0900-1100	M	RH
	1400-1600	S	T, RH, Wspd, Srad
29 Jun – 25 Jul 2000	0100-2400	S	RH
26 July 2000	0100-1600	S	RH
	1700	I / S	T, Wspd, Srad / RH
29 Aug 2000	1600-1800	S	T, RH, Wspd, Srad
31 Oct 2000	1800-1900	S	T, RH, Wspd, Srad
1 Nov 2000	2100 / 2300, 2400	I / S	T, RH, Wspd, Srad
9 Nov 2000	1600	I	T, RH, Wspd, Srad
24 Nov 2000	2300, 2400	S	T, RH, Wspd, Srad
25 Nov 2000	0400, 1800, 2000, 2300	I	T, RH, Wspd, Srad
26 Nov 2000	0100, 0200, 0700, 0900,		
	1200, 2200, 2300	I	T, RH, Wspd, Srad
	0300-0500, 2000	S	T, RH, Wspd, Srad
27 Nov 2000	0600, 1000 / 2300, 2400	I / S	T, RH, Wspd, Srad
28 Nov 2000	1300	I	T, RH, Wspd, Srad
29 Nov 2000	0600, 0900, 1100, 1200,		
	2000-2200 / 1400-1900	I / S	T, RH, Wspd, Srad
30 Nov 2000	0100, 0500	I	T, RH, Wspd, Srad
11 Dec 2000	2100-2400	S	Wspd
12 Dec 2000	0100-2400	S	Wspd
13 Dec 2000	0100-0600 / 0700-2400	S / M	Wspd
14 Dec 2000	0100-1000	M	Wspd
25 Dec 2000	1100-2400	M	Wspd
26 Dec 2000	0100-2400	M	Wspd
27 Dec 2000	0100-0900 / 1000-2400	M / S	Wspd
28 Dec 2000	0100-1300	S	Wspd

Table E.5 Missing or corrupt electric load data from the data set for the West Norman substation.

Date(s)	*Hours (24-Hour Clock)*	*Number of Missing Hourly Obs*
1-30 Apr 1998	0100-2400	720
1 Mar – 26 Apr 1999	0100-2400	1464
27 Apr 1999	0100-1500	15
1 Apr 2000	0600	1
22 May 2000	1100-2400	11
23-24 May 2000	0100-2400	48
25 May 2000	0100-1500	15

Table E.6 As in Table E.5, except for the Woodward substation.

Date(s)	*Hours (24-Hour Clock)*	*Number of Missing Hourly Obs*
9 May 2000	0100-2400	24

Table E.7 As in Table E.5, except for the Altus AFB substation.

Date(s)	*Hours (24-Hour Clock)*	*Number of Missing Hourly Obs*
29 Mar 2000	2300	1
1 May 2000	2400	1

Table E.8 As in Table E.5, except for the Dominance substation.

Date(s)	*Hours (24-Hour Clock)*	*Number of Missing Hourly Obs*
12 Feb 1999	1100, 1200	2
28 Mar 2000	0500-2400	20
27 May 2000	2000-2300	4

Appendix F: Neural Network Specifications of MetrixND

Artificial neural networks are flexible nonlinear models which make them attractive in forecasting. MetrixND expands upon a general form of a neural network for a single-variable forecasting problem function:

$$y_t = F(X_t, \beta) + u_t$$

$$= B_0 + \sum_{h=1}^{H} B_h \text{ x } H(X_t, \alpha) + u_t$$

$$= B_0 + \sum_{h=1}^{H} B_h \text{ x } \frac{1}{1 + \exp\left(-\left(a_{h,0} + \sum_{k=1}^{K} a_{h,k} X_{k,t}\right)\right)} + u_t$$

where y is the dependent variable, X's are the predictors of the dependent variable, B's are the parameters in the output layer, a's are the parameters in the hidden layer activation functions.

This neural network equation has a few specific properties. First, it is a single-output feedforward neural network. A feedfoward system implies an absence of feedbacks between layers and an absence of node-level interactions. Next, this NN equation has a single hidden layer with N nodes. Although the hidden layer (i.e., where specific algebraic transformations occur) is not explicitly hidden, it is a level of computation that has little meaning to the forecaster. Last, logistic (sigmoid) activation functions are used in the hidden layer and a linear activation function in the output layer.

The process of parameter estimation is called training in NN literature. The goal of this process is to determine network parameters that result in small model errors. The estimation process is more complicated than for a regression model because the model is nonlinear and because the objection function is relatively complicated. A conventional nonlinear least squares algorithm (Levenberg-Marquat algorithm/IMSL library) is used to find optimal parameters.