A NEW APPROACH TO SPATIO-TEMPORAL KRIGING AND ITS APPLICATIONS

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Master of Science in the Graduate School of The Ohio State University

By

Abhijat Agarwal, B.Tech.

Graduate Program in Computer Science and Engineering

The Ohio State University

2011

Dissertation Committee:

Dr. Srinivasan Parthasarathy, Advisor

Dr. Gagan Agrawal
ABSTRACT

Stochastic spatio-temporal variability is often observed in naturally occurring phenomena. It had always been a challenge to predict their behavior in space and time. Statistical techniques exist that may be united to model and predict the spatio-temporal behavior of these phenomena.

In this research we present a new approach to spatio-temporal data analysis. A new Spatio-Temporal Kriging model was built to predict the spatio-temporal behavior of atmospheric temperature data, gathered from heterogeneous sensors for over 10 years at 63 locations in the US. Kriging interpolates the best linear unbiased estimate of a value at an unobserved point in space, based on the weighted linear combination of surrounding observations, minimizing the prediction error. Spatial and temporal associations in the data were initially modeled separately, using Universal Kriging (UK) and Autoregressive (AR) techniques respectively and then combined to spatio-temporally predict temperatures, \( k \) days into the future in a given spatial domain.

ARIMA (Autoregressive Integrated Moving Average) model was used to compare the performance of our Spatio-Temporal Kriging model. Our model performed twice as better with 2.47°C of average standard error (SE) in prediction estimates as compared to...
4.49°C from ARIMA. Confidence interval (95% CI) for prediction estimates from ARIMA model was ±8.80°C as compared to ±4.84°C from our Spatio-Temporal Kriging model. Uncertainty in predictions observed from both the models may be largely associated to the presence of strong temporal correlation in the observations at locations near the Great lakes, also observed from slowly decaying autocorrelation function (ACF) at these locations.

A new Space-Time linear model was also built using regression that yielded poor results, because it only captured the effect of latitude on temperature, i.e. temperature drops as we move up north.

We also introduced a novel concept of Kriging based virtual sensor (KVSense) that may be used for temporarily replacing the faulty wireless sensor(s) and also to emulate the working of a real sensor at inaccessible areas.

We concluded by discussing, possible novel energy harvesting (energy conservation and wireless sensor power rejuvenation) strategies for wireless sensor networks (WSN) configured in a spatial domain based on mined spatio-temporal knowledge on availability of ambient (sunlight, wind, etc.) energy source.
Dedicated to three extraordinary human beings who have given meaning to my life...

my Mom, my Sir and my Love.
ACKNOWLEDGMENTS

First and foremost, I would like to acknowledge and thank the invaluable support and guidance of my advisor, Dr Srinivasan Parthasarathy. From the beginning of my graduate work, he offered me the freedom to explore different research problems, and constantly motivated and challenged me to grow my perspective and deeper understanding to take this work to this level. I am deeply indebted to him for his guidance and advice. His passion, energy and work drive have been a source of inspiration. This has been a privilege to work with him and I got to learn a lot from him during this time.

I would like to acknowledge and thank my dissertation committee member Dr. Gagan Agrawal for his valued contributions and suggestions and his sincere cooperation during this work.

I would also like to acknowledge and thank Dr. Chris Holloman for his valuable support, brain storming sessions and making complex things seem simple. Also, I would like to thank Ms. Fangfang Sun and Ms. Lili Zhuang for their support and suggestions.

I would like to acknowledge and thank Raghu, Faisal, Tyler, Dave and Venu and members of Data Mining Research Lab., for their valuable suggestions and words of
encouragement that kept me motivated all this time.

I would like to acknowledge my brother Dr. Gaurav K. Agrawal for providing his unparalleled support and guidance during all these years.

I don't have enough words to express my deep gratitude towards my parents. Nothing that I will ever do can even come close to their love and support for me and the sacrifices they have made to see me here. I can't express my feelings through words for my wife to be Rachita for her unconditional love and support. I will also like to thank my relatives and friends in India, who are always there to cheer me up.

Finally, I thank all my friends (UV gang) – Banni, Danda, Dpac, J, Mutthu, Rajachan, Pattu and Tim for sharing this journey and making it a memory to be cherished forever.
VITA

April 1986 .......................................................... Born – Meerut, India

2004 ................................................................. Air Force Bal Bharati School, New Delhi

June 2007 – August 2007................................. Summer Intern, Tata Consultancy Services Ltd., New Delhi, India

2008 ................................................................. B.Tech., Amity University, U.P.

July 2009 – September 2009......................... Intern, SMC Global Securities Ltd., New Delhi, India

September 2010 – January 2011 ................. Graduate Teaching Associate, Department of Computer Science and Engineering, The Ohio State University

PUBLICATIONS


2. B. Gupta, Abhijat Agarwal, A.K. Upadhayay, Data Mining: The Complete

FIELDS OF STUDY

Major Field: Computer Science and Engineering
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>v</td>
</tr>
<tr>
<td>VITA</td>
<td>vi</td>
</tr>
<tr>
<td>PUBLICATIONS</td>
<td>vii</td>
</tr>
<tr>
<td>FIELDS OF STUDY</td>
<td>viii</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>ix</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xiii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xiv</td>
</tr>
<tr>
<td>1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1. The Intuition</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Challenges</td>
<td>2</td>
</tr>
<tr>
<td>1.3. Thesis Statement</td>
<td>4</td>
</tr>
<tr>
<td>1.4. Contributions</td>
<td>5</td>
</tr>
<tr>
<td>1.5. Organization</td>
<td>8</td>
</tr>
<tr>
<td>2. Background and Related Work</td>
<td>9</td>
</tr>
</tbody>
</table>
2.1. Geostatistics .......................................................... 9

2.2. Spatial Prediction with Kriging .................................................... 10

2.2.1. Universal Kriging (UK) ........................................ 12

2.3. Temporal Prediction with Time Series Modeling .............................. 13

2.3.1. Autoregressive (AR) Process ....................................... 14

2.3.2. Moving-Average (MA) Process ...................................... 15

2.4. Autoregressive Moving-Average (ARMA) Process ....................... 15

2.5. Auto Regressive Integrated Moving Average (ARIMA) Process ........ 16

2.6. Related Work ..................................................................... 16

3. Statistical Models .................................................................. 19

3.1. Spatio-Temporal Kriging Model ............................................. 19

3.1.1. Basic Model ............................................................ 19

3.1.2. Covariance Structure .................................................... 21

3.1.3. Mean Structure .......................................................... 30

3.1.4. Model Fitting ............................................................ 32

3.2. ARIMA Model .................................................................. 33

3.2.1. Seasonality in ARIMA models .................................... 34

3.2.2. Basic Model ............................................................. 35

3.2.3. Determining the Order of Differencing ......................... 37
3.2.4. Deciding AR and MA terms ................................................................. 38
3.2.5. Model Fitting ....................................................................................... 39

4. Results and Discussions ........................................................................ 44
  4.1. The Data .............................................................................................. 44
  4.1.1. Does Size Matter? ........................................................................... 45
  4.2. Spatio-Temporal Kriging vs. ARIMA(2, 1, 2) ..................................... 46
    4.2.1. Prediction Results ......................................................................... 48
    4.2.2. Prediction Accuracy ...................................................................... 49
      4.2.2.1. Difference in Prediction Results ............................................... 50
      4.2.2.2. Standard Errors in Prediction Estimates ................................... 52
        4.2.2.2.1. Spatio-Temporal Kriging Standard Errors ......................... 52
        4.2.2.2.2. ARIMA(2, 1, 2) Standard Errors ...................................... 54

5. Possible Novel Applications ..................................................................... 56
  5.1. KVSense: A Virtual Sensor ................................................................. 56
  5.2. Energy harvesting in wireless sensor networks .................................... 57

6. Conclusion and Future Work .................................................................. 60

Bibliography ............................................................................................... 64

Appendix A: Space-Time Linear Model ....................................................... 68
Appendix B: Glossary of Terms .................................................................... 72
Appendix C: AIC Value Table ................................................................. 74

Appendix D: Standard Errors (Spatio-Temporal Kriging Model).......................... 76

Appendix E: ACF and PACF Plots.................................................................. 84
LIST OF TABLES

Table 1: ARIMA(2,1,2) model parameters ................................................................. 42

Table 2: Summary of AIC Values .................................................................................. 75
LIST OF FIGURES

Figure 1: Kriging estimates for the year 1999 on left and Kriging SEs on right .......... 11

Figure 2: Finer spatial grid for prediction ........................................................................ 12

Figure 3: A Theoretical Variogram model ........................................................................ 23

Figure 4: Empirical Variogram ....................................................................................... 23

Figure 5: Gaussian model fitted without nugget effect ..................................................... 24

Figure 6: Gaussian model fitted semivariogram with nugget effect ................................ 24

Figure 7: ACF plots for locations 1-9 ............................................................................... 26

Figure 8: PACF plots for locations 1-9 .............................................................................. 27

Figure 9: Fitted Time Series ............................................................................................. 28

Figure 10: ACF plots for locations 1-12 .......................................................................... 41

Figure 11: Sensor grid for data ........................................................................................ 44

Figure 12: Input data (1997-99) for Spatio-Temporal Kriging model ............................... 46

Figure 13: Input Data (1990-99) for ARIMA(2, 1, 2) model ............................................ 47

Figure 14: Contour plots for Predicted vs. Observed temperature for 3 days ................. 49

Figure 15: Difference between predicted and observed values ....................................... 51

Figure 16: Spatio-Temporal Kriging SEs in prediction estimates at locations (1-12) ....... 53

Figure 17: ARIMA(2,1,2) SEs for predicted Day 1, 2 and 3 ............................................. 54

Figure 18: Prediction Results from Space-Time Linear model ....................................... 71
Chapter 1

“Everything is related to everything else, but near things are more related than distant things.”- Tobler

1. Introduction

1.1. The Intuition

The first law of Geography was given by Waldo Tobler in 1970 [5]. This is essentially synonymous to the idea of spatial and temporal dependence. This became the foundation of spatial, temporal and most importantly spatio-temporal analysis. This idea resides at the heart of our\(^1\) work. In the following sections we\(^1\) discuss the challenges we faced and overcame and also the contributions we made with our work.

---

\(^1\) Throughout this work, “we” and “our” is used instead of “I” to refer the author. It is simply a writing preference of the engineering and computer science community. Unless and otherwise noted, when using “we” the author is referring to his own work.
1.2. Challenges

Challenges we responded to during this research may be divided into challenges posed by the data and the complexities associated with building a robust spatio-temporal statistical model. First step in spatio-temporal data modeling requires identification of inherent dependencies in the data that vary spatially and cascade temporally. The spatio-temporal variability in the data must be accounted for simultaneously. We treated spatio-temporal covariance in our data to be space and time separable that, in our case may be eventually combined to get spatio-temporal predictions of daily average temperature.

First, we had to account for spatial correlations in the data. For this purpose we relied on Kriging techniques [3, 4, 6, 7]. Universal Kriging, also known as Kriging with a trend was selected to model the spatial trend in the data.

Second, we needed to model the temporal correlations in the data. While looking at temperature data over the years, a smooth cyclical trend was easily noticeable. No constant trends were visible while looking at daily temperatures from the exploratory analysis of the data. Thus, the temporal trend in the temperature was difficult to model at daily resolution. For this purpose we relied upon autoregressive techniques to account for temporal autocorrelations in the data.
Third, we needed to build a daily temperature prediction model without the knowledge of factors like snowfall, rain, storms, etc. Knowledge of what caused an event is as important as the event itself. The daily temperature data used in this work contained no information on the factors that might have accounted for the observed spatio-temporal variability. The absence of knowledge on these factors forced us to artificially induce and model the covariates into the mean structure of our Spatio-Temporal Kriging model. We modeled the historical temperature data by accounting for seasonal changes like annually and monthly repeating patterns and fluctuations occurring on daily basis. In a regression model, behavior of response variable is dependent on one or more independent variables. In our statistical model we do not assume the mean of parameters to be constant. Instead, we allow them to be linear functions of several predictor variables. These covariates include linear and higher order short term and seasonal trends. The covariate values were defined for the historical observations as well as the future observations.

Fourth, we needed to model the spatial and temporal variations and co-variations for the locations that were close to large water bodies. As discussed earlier, temperature is influenced by several factors. Great Lakes had always been a bottle-neck to build robust statistical models due to uncertainty in behavior of temperature in the surrounding regions owing to lake effects [43]. We observed a slowly decaying autocorrelation function (ACF) at the locations in the affected regions. We decided not to address these concerns at this time and investigate them further in our future endeavors.
Fifth, it was required to combine the modeled spatial and temporal covariance structures. Since it can’t be done by simply adding them together we used Kronecker product [8] to address this challenge. Kronecker product may be used to combine the separate space-time covariance structures [8].

Final challenge was to find a suitable technique that could be compared against our Spatio-Temporal Kriging model. ARIMA [12, 17, 30] was selected for this purpose. We found that the results from our Spatio-Temporal Kriging model were superior to that obtained from the ARIMA model. Results from both the models are discussed in Chapter 4. In the following sections we lay out our thesis statement and highlight our key contributions.

1.3. Thesis Statement

“We propose a new spatio-temporal model to predict the behavior of naturally occurring phenomenon such as temperature and propose novel applications for this model. Spatiotemporal covariance in the data was assumed to be space-time separable. Kriging techniques may be used to model the spatial correlation and autoregressive techniques to model the temporal dependencies in the data. Further, they can be unified to create a new spatio-temporal prediction model.
We believe that our model can be used as a virtual sensor for temporary replacement of the faulty wireless sensor(s) and also for emulating the working of a real sensor at inaccessible areas. Finally, we propose novel ways to harvest energy in wireless sensor networks by using our model as a tool to predict the spatio-temporal availability of ambient energy.”

1.4. Contributions

1. A New Spatio-Temporal Kriging Model: We begun by extracting and modeling the spatial correlations in the sensor data by using Universal Kriging. Kriging is synonymous with Best Linear Unbiased Estimator (B.L.U.E.) [3, 6]. It interpolates the best linear unbiased estimate of a value at an unobserved point in space, based on the weighted linear combination of surrounding observations, minimizing the prediction error. Universal Kriging is used in the presence of a (non-stationary) trend in the data. We studied the temporal patterns in the data and modeled the temporal correlations by building an autoregressive model. Auto Correlation Function (ACF) and Partial ACF (PACF) patterns were analyzed to study the temporal dependencies in the data. Autocorrelations for observations at locations near a water body behaved independent of rest of the observations in the spatial domain. We decided to address this behavior independently of other observations in our future efforts. We approximated the spatial and temporal
covariance together by using Kronecker product. The notion of separable approximations of space-time covariance structures has been studied earlier by Genton [8]. In Universal Kriging the mean is assumed to be unknown and non-constant. Thus, we can’t assume the parameters controlling the trend in the mean to be constant. For this reason, we allow them to be linear functions of several predictor variables. Based on exploratory analysis, we introduced seven predictor variables (covariates) to de-trend the mean and performed Kriging on the residuals. Finally, with the covariance and mean structures in place, the trend and kriged residuals were added back to our Spatio-Temporal Kriging model and the prediction estimates were obtained. We conducted extensive experimentation to test the performance of our Spatio-Temporal Kriging model by varying the size of data used to build the model. We concluded that time-series models such as ARIMA are not sufficient for spatio-temporal predictions when used independently by comparing results obtained from ARIMA and our Spatio-Temporal Kriging model. Our contribution from this research is the new approach for spatio-temporal data analysis. Our model in generic sense can be applied to spatio-temporal analysis of diverse sets of data ranging from the field of earth science to strategic marketing. Next, we discuss possible novel applications of our Spatio-Temporal Kriging model.

2. **Possible Novel Applications:** We introduce a novel application of our model as a virtual sensor (*KVSense*). With interpolation as primary objective while using
Kriging we can interpolate at the unobserved locations. The concept can be applied to observe at locations without a physical sensor. This concept can be exploited in the case of faulty sensor(s), slow response, sensing in hostile environment, etc. In the future we wish to test this concept in the real world scenarios.

We also introduce our Spatio-Temporal Kriging model as a tool to anticipate the availability of ambient energy for planned and uninterrupted functioning of autonomous wireless sensors in WSN. Predicting the availability of energy remains to be an exciting area of research. We hope to shed some light with the help of our model which can anticipate when and where the ambient energy for wireless sensors will be available based on its capability to predict in space and time. This, we believe, will assist in the efficient planning of energy harvesting, sensor rejuvenation (determining next available recharge) and energy conservation for wireless sensors in a WSN.

In this research we proposed a new spatio-temporal prediction models by combining techniques from spatial statistics and time series modeling. We showed a new approach to spatio-temporal analysis of daily temperature data with our Spatio-Temporal Kriging model. One of the goals we aimed to address was to predict the availability of ambient energy source such as sunlight to facilitate energy harvesting in WSN, based on the daily spatio-temporal prediction of temperature from our model. Focus of our research is not
weather prediction [38] but a new approach to spatio-temporal data modeling and its applications.

1.5. Organization

The rest of the thesis is organized as follows: in Chapter 2 we discuss the background of the work and techniques (spatial and temporal) related to our work. We detail our statistical models, Spatio-Temporal Kriging Model and ARIMA (2, 1, 2) formally in Chapter 3. Our dataset, experimental results and applications are presented and discussed in Chapter 4. Possible novel applications of our model are discussed in Chapter 5. Our conclusion and future courses of action are followed in Chapter 6.
Chapter 2

2. Background and Related Work

2.1. Geostatistics

Geostatistician G. Matheron [35] coined the term ‘Kriging’, inspired by previous work of Sichel[36], Krige [33], and Matern [34]. In the last 30 years the Geostatistics had been carried out and consolidated as an applied science to solve practical and concrete problems. The fundamental element is the analysis of the spatial distribution for the extractable information. The aim of Geostatistics is the characterization of a natural phenomenon. Geostatistics techniques, such as Kriging, provide estimates of the parameter, as well as a measure of their uncertainty. These techniques can help us to select the sampling points in a way that minimizes the uncertainty in estimation. Many environmental processes exhibit spatial variations that can be modeled using these
techniques. Applications of such techniques can be found in geology, mining, hydrology, meteorology, epidemiology, etc.

Geostatistics considers the variables as random functions with two components, deterministic and random, with the aim of representing the variable's spatial dependence [3]. The deterministic section represents non spatial influences while random section is interpreted as the realization of a random field. This random field (spatial stochastic process) can be characterized by statistical moments, i.e. as the covariance function or variogram. Based on the sampled values at nearby locations, Kriging techniques interpolate values of a random field at locations where samples are not available. Kriging employs the variogram model in order to generate the best linear unbiased estimate at each location. The variogram, also known as semivariogram, is used to model the spatial correlation or covariance structure in the data.

2.2. Spatial Prediction with Kriging

Kriging is an interpolation method named after a South African mining engineer, D. G. Krige who developed the technique for highly accurate prediction of ore reserves. Over the past several decades, Kriging has become a fundamental tool in the field of Geostatistics. Kriging is based on the assumption that the parameter being interpolated can be treated as a regionalized variable. A regionalized variable is intermediate between
a truly random variable and a completely deterministic variable. Those are expected to vary in a continuous manner from one location to the next and therefore nearby points have a certain degree of spatial correlation, but points that are widely separated are statistically independent [37]. Kriging interpolates the best linear unbiased estimate of a value at an unobserved point in space, based on the weighted linear combination of surrounding observations, minimizing the prediction error. In (Figure 1) we can see the Kriging estimates on the left for the year 1999 and Kriging standard errors on the right. It can be clearly noticed that SEs at known locations are nearly zero and at unobserved locations are < 0.15. The Kriging surface below was produced for a finer grid (Figure 2) than the sensor grid shown in (Figure 11).

Figure 1: Kriging estimates for the year 1999 on left and Kriging SEs on right.
2.2.1. Universal Kriging (UK)

Universal Kriging (UK) is the simplest form of (Kriging with a Trend or external drift) KT systems. One of the assumptions made in Kriging is that the data being estimated are stationary. That is, as we move from one region to the next in the scatter point set, the average value of the scatter points becomes relatively constant. Whenever there is a significant spatial trend in the data values such as a sloping surface or a localized flat region, this assumption is violated. In such cases, the stationary condition can be temporarily imposed on the data by use of a drift term. The drift is a simple polynomial function that models the average value of the scatter points. The residual is the difference between the drift and the actual values of the scatter points. Because the residuals are expected to be stationary, Kriging is performed on the residuals and the interpolated residuals are added to the drift to compute the estimated values. Using a drift in this
fashion is often called “Universal Kriging”. Universal Kriging equations can be easily found in [3]. Detailed mathematical description of our Spatio-Temporal Kriging will be discussed in Chapter 4.

2.3. Temporal Prediction with Time Series Modeling

Time series modeling can be referred to as the process of modeling the temporal structures in the data. These data points may be indexed regularly or at irregular intervals. We deal with the models that assume continuously indexed data. Processes that account for temporal correlations in the data are referred to as time series models. Commonly used processes include autoregressive (AR) process, moving average (MA) process. These processes are used mutually as an autoregressive moving average (ARMA) process. Special case of ARMA is known as autoregressive integrated moving average (ARIMA) process. We discuss briefly about each of them in the following section.
2.3.1. Autoregressive (AR) Process

Observations of environmental, ecological and biological process often depend upon one or more observations that precede them immediately. A time series that models such structure is the autoregressive (AR) process [4]. A time series $X_t$ is said to be an autoregressive process of order (p), AR(p).

The notation AR(p) refers to the autoregressive model of order p. The AR(p) model is written as:

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t, \quad t = p, p + 1, ..., $$

Where $\varphi_i: i = 1, ..., p$, are fixed but unknown model parameters to be estimated, $\varepsilon_t$ is a white noise process with zero mean and variance and $c$ is a constant. The error terms that are generally assumed to be independent, identically distributed variables sampled from a normal distribution with zero mean. The constant term is usually omitted for simplicity.

The auto correlation function for an AR(p) can be written as follows:

$$\rho(\tau) = \varphi_1 \rho(\tau - 1) + \varphi_2 \rho(\tau - 2) + \varphi_3 \rho(\tau - 3) + \cdots + \varphi_p \rho(\tau - p), \quad \tau = 1, 2, ... $$
2.3.2. Moving-Average (MA) Process

The notation MA(q) refers to the moving average model of order q:

\[ X_t = \mu + \varepsilon_t + \sum_{i=1}^{q} \theta_i X_{t-i} \varepsilon_{t-i} + , \quad t = q, q + 1, ..., \]

Where the \( \theta_i : i = 1, 2, ..., q \) are fixed but unknown parameters of the model, \( \mu \) is the expected value of \( X_t \) (often assumed to equal 0), and the \( \varepsilon_t \) and \( \varepsilon_{t-1} \) are again, white noise or error terms.

2.4. Autoregressive Moving-Average (ARMA) Process

The notation ARMA(p, q) refers to the model with p autoregressive terms and q moving average terms. This model contains the AR(p) and MA(q) models and can be written as follows:

\[ X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} , \]
2.5. **Auto Regressive Integrated Moving Average (ARIMA)**

Process

An autoregressive integrated moving average (ARIMA) model is a generalization of an autoregressive moving average (ARMA) model. These models are fitted to time series data either to better understand the data or to predict future points in the series (forecasting). They are applied in some cases where data show evidence of non-stationarity (Appendix B), when an initial differencing step (corresponding to the "integrated" part of the model) can be applied to remove the non-stationarity.

The model is generally referred to as an ARIMA\((p, d, q)\) model where \(p\), \(d\), and \(q\) are non-negative integers that refer to the order of the autoregressive, integrated, and moving average parts of the model respectively. A formal description of ARIMA\((p, d, q)\) model can be found in Chapter 4.

2.6. **Related Work**

Several techniques like ‘Geographically Weighted Regression’ [39] have been developed to analyze spatially varying relationships. However these techniques do not answer the issue of temporally varying relationship. Time series modeling [12, 13, 15, 16, 19] is a popular choice for modeling relationships that cascade in time. Nonetheless, issue of daily weather prediction is not answered by any of them. Davis [37] addressed the issue
of temperature prediction but at a monthly resolution, which was accomplished by using a seasonal autoregressive integrated moving average model (SARIMA) [12, 30]. SARIMA gives good prediction result when seasonality is modeled monthly or annually. Modeling seasonality daily has not been addressed extensively in the related literature. Temperature fluctuations occur on daily basis and these fluctuations may not have strong correlations with observations that are a month or even a year apart.

Radhika and Shashi [41] used support vector machines to predict daily temperatures. The results show high inaccuracy. For predicting daily temperature Pal et al. [40] have used self organizing feature map (SOFM) and MLPs to realize a hybrid network named SOFM–MLP. Their technique only addresses the issue of temporal prediction.

Use of Kriging with an external drift for mapping temperature has been studied by Hudson and Wackernagel [42]. The authors use elevation as an external drift in the data. Presence of elevation as one of the predictor provides extra knowledge on spatial variations in temperature. Our Spatio-Temporal Kriging model was built without any information on elevation or any other naturally occurring factor. This made our work more challenging and interesting.

We could not locate any published work addressing the problem of spatio-temporal prediction of daily weather using Kriging. Also, no publications were found that used spatio-temporal Kriging for predicting the availability of ambient energy. Several solar
energy aware routing techniques [20, 26, 27, 28, 29] along with energy aware clustering of sensor nodes [25] have been extensively researched and developed. However, they fail to address the issue of predicting the availability of energy. Furthermore, no work addressed the use of Kriging as a virtual sensor.
Chapter 3

3. Statistical Models

3.1. Spatio-Temporal Kriging Model

3.1.1. Basic Model

Let a $d$ dimensional space $\mathbb{R}^d$ contain a set of observed locations $\{s_1, s_2, \ldots, s_n\}$ and unobserved locations $\{s_1, s_2, \ldots, s_u\}$. Data at $n$ locations was observed for $T_H: \{t_1, t_2, \ldots, t_T\}$ historical time points ($T_H$) and predictions were to be obtained for each of $k$ days into the future, i.e., $T+1, T+2, \ldots, T+k$ at $m$ locations. It is also assumed that the historical data is collected at the same $n$ locations during each of the $T$ historical time points.

The observed data can be represented as a vector of length $nT$ where $n$ corresponds to the number of locations and $T$ to the number of observation at each location. We denote the
temperature observations at location $s_i$ where $i = 1, 2, \ldots n$, on day $t_o$ where $o = 1, 2, \ldots T$ as $y_{s,t}$. The vector of all observations can be written as follows:

$$y_o = \begin{pmatrix} y_{1,1} \\ y_{1,2} \\ \vdots \\ y_{n,T} \end{pmatrix}.$$

During each of the $k$ future time points we wish to obtain predictions at $m$ locations, we call them $s'_1, s'_2, \ldots s'_m$. Here $\{s'_1, s'_2, \ldots s'_m\} \subseteq \{s_1, s_2, \ldots s_n\}$ or $\{s'_1, s'_2, \ldots s'_m\} \subseteq \mathbb{S}^d - \{s_1, s_2, \ldots s_n\}$. We denote the temperature observations at site $s'_p$ where $p = 1, 2, \ldots m$, on day $T + k$ as $y_{s'_p,T+k}$. Similarly, the vector of observations to be predicted can be written as follows:

$$y_p = \begin{pmatrix} y_{1,T+1} \\ y_{1,T+2} \\ \vdots \\ y_{m,T+k} \end{pmatrix}.$$

One way to conceptualize Kriging is to think of the historical observations and observations to be predicted as being jointly multivariate normally distributed with their covariance being governed by the distance (in space and time) between points. As a result, we can write the joint distribution of $y_o$ and $y_p$ as follows:
and our end goal is to derive the expected value of $y_p$ conditional on $y_o$ (Kriging estimates) i.e., $E(y_p|y_o)$ and the standard deviation of $y_p$ conditional on $y_o$ (Kriging standard variance), i.e. $V(y_p|y_o)$.

### 3.1.2. Covariance Structure

To simplify our statistical modeling, we assume that the covariance for our observations $\Sigma_o$ is space-time separable [4, 8, 9]. Covariance matrix of observations $\Sigma_o$ can be written as a kronecker [8, 9] product of a spatial covariance matrix $\Sigma_s$ and a temporal covariance matrix $\Sigma_{\tau_h}$.

$$\Sigma_o = \Sigma_s \otimes \Sigma_{\tau_h}.$$  

$\Sigma_s$ is an $n \times n$ matrix giving the covariance between observations at a pair of spatial locations at the same time point. First, the empirical variogram was calculated from the data and *sill* ($\sigma^2$), *range* of influence ($\phi$) and nugget effect ($\tau^2$) parameters were obtained. Nugget effect is the variance at zero lag observed from the empirical variogram and represents the micro-scale variations or measurement errors in the data. Based on
exploratory analysis a Gaussian variogram was fitted to the empirical variogram. Initially, empirical variogram was fitted without the nugget effect (Figure 5) but we found that model fits better by including the nugget effect. Also, Gaussian model should not be used without the nugget effect [3, 8] since it reflects a very smoothly varying spatial process.

Thus, Gaussian covariogram with nugget (τ²) effect (Figure 6) was chosen to describe the correlation. Σ_s is given as follows:

$$\Sigma_s[i,j] = \begin{cases} 
\tau^2 + \sigma^2 & \text{if } i = j \\
\sigma^2 \exp\left\{-\left(\frac{\text{dist}(s_i,s_j)}{\phi}\right)^2\right\} & \text{if } i \neq j
\end{cases}$$

Partial sill (σ²), range of influence (ϕ) and nugget effect (τ²) shown in (Figure 3) are the parameters estimated from the empirical semivariogram (Figure 4).
Figure 3: A Theoretical Variogram model

Figure 4: Empirical Variogram
Figure 5: Gaussian model fitted without nugget effect

Figure 6: Gaussian model fitted semivariogram with nugget effect
\( \Sigma_{T_h} \) is a \( T \times T \) matrix giving the covariance between observations at the same location over two time points. Based on exploratory analysis, an autoregressive model was chosen to describe the correlation. ACF (Figure 7) and PACF (Figure 8) pattern was observed to determine the order of the autoregressive model. For ACF and PACF observed at each location in the data set please see (Appendix E). The model is given as follows:

\[
\Sigma_{T_h}[l,j] = \rho^{\lvert i-j \rvert}.
\]
Figure 7: ACF plots for locations 1-9
Figure 8: PACF plots for locations 1-9
Figure 9: Fitted Time Series

\(\Sigma_p\), the covariance matrix for predicted observations, can be decomposed using a similar method to the one used for \(\Sigma_o\). However, the spatial covariance and temporal covariance matrices may be different since the predictions can be over a different spatial domain and a different set of time points:

\[
\Sigma_p = \Sigma_q \otimes \Sigma_{T_p}.
\]

\(\Sigma_q\) is a \(m \times m\) matrix that gives the covariance between future observations at a pair of spatial locations at the same time point.
\[ \Sigma_q[i,j] = \begin{cases} 
\tau^2 + \sigma^2 & \text{if } i = j \\
\sigma^2 \exp\left\{-\frac{(\text{dist}(q_i,q_j))^2}{\phi}\right\} & \text{if } i \neq j 
\end{cases} \]

\( \Sigma_T \) is a \( k \times k \) matrix that represents the covariance between observations at the same location over two future time points.

\[ \Sigma_T[i,j] = \rho^{|i-j|}. \]

The off-diagonal elements of the covariance matrix \( \Sigma_p \) and its transpose are slightly more complicated. These sub-matrices give the correlation between past data and future predictions. Nonetheless, they can still be expressed as a kronecker product.

\[ \Sigma_p = \Sigma_s \otimes \Sigma_T. \]

In this case, \( \Sigma_s \) is a \( m \times n \) matrix giving the covariance between observations at a pair of spatial locations using the Gaussian covariogram and \( \Sigma_T \) is a \( k \times T \) matrix giving the covariance between observations at two time points using the autoregressive model.
3.1.3. Mean Structure

In our statistical model, we do not assume that the mean parameters are constant. Instead, we allow them to be linear functions of several predictor variables. When means are allowed to be a linear function of predictor variables, the methodology being employed is typically called universal Kriging [3].

First, we define a matrix of covariate values for the historical observations as well as the future observations. For our model, this matrix of covariates includes following terms:

1. Intercept
2. Linear trend over time
3. Quadratic trend over time
4. Cubic trend over time
5. Linear latitudinal effect
6. Sine wave with an annual cycle
7. Cosine wave with an annual cycle

As a result, the covariate matrix of historical observations is a $nT \times 7$ matrix, shown as following:
And the covariate matrix of future observations is a $mk \times 7$ matrix, as shown below:

$$X_p = \begin{bmatrix}
1 & \frac{T+1}{365.25} & \left(\frac{T+1}{365.25}\right)^2 & \left(\frac{T+1}{365.25}\right)^3 & \text{lat}(s_1) & \sin\left(\frac{2\pi(T+1)}{365.25}\right) & \cos\left(\frac{2\pi(T+1)}{365.25}\right) \\
1 & \frac{T+2}{365.25} & \left(\frac{T+2}{365.25}\right)^2 & \left(\frac{T+2}{365.25}\right)^3 & \text{lat}(s_1) & \sin\left(\frac{2\pi(T+2)}{365.25}\right) & \cos\left(\frac{2\pi(T+2)}{365.25}\right) \\
1 & \frac{T+3}{365.25} & \left(\frac{T+3}{365.25}\right)^2 & \left(\frac{T+3}{365.25}\right)^3 & \text{lat}(s_1) & \sin\left(\frac{2\pi(T+3)}{365.25}\right) & \cos\left(\frac{2\pi(T+3)}{365.25}\right) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \frac{T+k}{365.25} & \left(\frac{T+k}{365.25}\right)^2 & \left(\frac{T+k}{365.25}\right)^3 & \text{lat}(s_m) & \sin\left(\frac{2\pi(T+k)}{365.25}\right) & \cos\left(\frac{2\pi(T+k)}{365.25}\right)
\end{bmatrix}.$$
3.1.4. Model Fitting

Now that we have defined the covariance structure and mean structure of the model, it is easy to write the estimators for the unknown parameters and to obtain the Kriging estimates and standard errors.

First, we calculate the best linear unbiased estimators for $\beta$:

$$\hat{\beta} = (X_o^T\Sigma_o^{-1}X_o)^{-1}X_o^T\Sigma_o^{-1}y_o.$$  

Calculating this quantity requires inversion of a very large matrix $\Sigma_0$. To perform the calculations efficiently, we can take advantage of the fact that $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$. As a result, the estimates are actually calculated by performing the following calculation:

$$\hat{\beta} = (X_o^T(\Sigma_s^{-1} \otimes \Sigma_{th}^{-1})X_o)^{-1}X_o^T(\Sigma_s^{-1} \otimes \Sigma_{th}^{-1})y_o.$$  

Next, we calculate the Kriging estimates using the following formula:

$$E(y_p | y_o) = X_p \hat{\beta} + \Sigma_{po} \Sigma_o^{-1}(y_o - X_o \hat{\beta}).$$  

The covariance matrix of the Kriging estimates can be obtained using the following formula:
\[
V(y_p|y_o) = \Sigma_p - \Sigma_{po}\Sigma^{-1}_{o}\Sigma^T_{po}.
\]

The diagonal elements of above Spatio-Temporal Kriging variance matrix \(V\) are the Spatio-Temporal Kriging variance. The square roots of the diagonal elements of the above variance matrix are the Kriging standard errors.

\[
\tilde{\sigma}_R = \sqrt{V_{diag}(y_p|y_o)},
\]

where \(V_{diag}(y_p|y_o) = (V_1, V_2, \ldots, V_k)\) are the \(k\) elements of the vector obtained by taking the diagonal elements of Spatio-Temporal Kriging variance matrix. Finally, we can calculate the 95% confidence interval for our prediction estimates at each location by:

\[
[E(y_p|y_o) \pm 1.96\tilde{\sigma}_R]
\]

The uncertainty in prediction results can also be discussed by just using \(\pm 1.96\tilde{\sigma}_R\).

### 3.2. ARIMA Model

ARIMA models have a general form of ARIMA\((p,d,q)\) where \(p\) is the order of the standard autoregressive (AR) component, \(q\) is the order of moving average (MA) term
and $d$ is the order of differencing. AR term describes the dependence of a variable $X_t$ such as daily average temperature in our case, on the previous values $X_{t-1}$ given by $AR(1)$, $X_{t-2} AR(2)$, and so on. Dependence of $X_t$ on the weighted moving average of historical values $X_{t-1}$ to $X_{t-n}$ is described by MA term in ARIMA model. For example, forecasting $k$ day(s) ahead with $AR(p)$, all weight is assigned to previous $p$ day(s) temperature, where $p = 1, 2\ldots, n$. On the contrary, forecasting $k$ day(s) ahead with $MA(q)$, partial weight is given to temperature observed on $(q - 1)$ day and lesser weight is assigned to observation on $(q - 2)$ day and so on. The value of assigned weight decays exponentially.

### 3.2.1. Seasonality in ARIMA models

Presence of strong seasonality in temperature time series often calls for fitting a seasonal component to ARIMA resulting in SARIMA or seasonal ARIMA model. In this case seasonal relationship exists between $X_t$ and $X_{t-s}$ where $s$ is the seasonal cycle. Seasonal cycle could be adjusted to capture weekly, monthly, yearly or any other seasonal relationships in the time series. This multiplicative seasonal modeling approach is given the general form of $ARIMA(p, d, q) \times S(p, D, Q)$.

However, our goal is to predict daily average temperatures and not temperatures averaged over months or years. Thus, we discard the use of seasonal component into our ARIMA
model and use non-seasonal ARIMA with higher order AR and MA terms, ARIMA (2, 1, 2). Details on model selection are discussed later with model fitting. Following sections lay out the correct basis for fitting an ARIMA model. Insights related to incorrect and over-fitting are also discussed along with subtle suggestions to avoid them.

3.2.2. Basic Model

Given a time series of data $X_t$ where $t$ is an integer index and the $X_t$ are real numbers, an ARMA($p$, $q$) model is given by:

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

Above equation can be rearranged and rewritten as follows:

$$\left(1 - \sum_{i=1}^{p} \phi_i L^i \right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i L^i \right) \varepsilon_t,$$
where $L$ is the lag operator, the $\varphi_i$ are the parameters of the autoregressive part of the model, the $\theta_i$ are the parameters of the moving average part and the $\epsilon_t$ are error terms.

Assuming that the polynomial $\left(1 - \sum_{i=1}^{p} \varphi_i L_i \right)$ has a unitary root of multiplicity $d$, it can be rewritten as:

$$\left(1 - \sum_{i=1}^{p} \varphi_i L_i \right) = \left(1 + \sum_{i=1}^{p-d} \varphi_i L_i \right) (1 - L)^d.$$

An ARIMA($p, d, q$) process is obtained by integrating an ARMA($p, q$) process. That is,

$$\left(1 - \sum_{i=1}^{p} \varphi_i L_i \right)(1 - L)^d X_t = \left(1 + \sum_{i=1}^{q} \theta_i L_i \right) \epsilon_t,$$

where $d$ is a positive integer that controls the level of differencing (or, if $d = 0$, this model is equivalent to an ARMA model) [13].
3.2.3. Determining the Order of Differencing

Determining the order of differencing is the first and most important step needed to stationarize the series and remove the gross features of seasonality before fitting an ARIMA($p, d, q$) model [12, 14, 15, 30]. Usually, the right order of differencing is the lowest order of differencing that yields a time series which fluctuates around a well-defined mean value with its autocorrelation function (ACF) plot decaying fairly rapidly to zero, either from above (positive autocorrelation) or below (negative autocorrelation). It is often possible to decide from ACF plots whether higher order of differencing is necessary. Higher order of differencing is evident if a series shows a long-term trend, or lacks a tendency to return to its mean value, or its autocorrelations are positive out to a high number of lags (e.g., 10 or more) [14].

Stationarity in the original series is assumed for a model with no ($d = 0$) order of differencing. Similarly, model with one ($d = 1$) order of differencing assumes that the original series has a constant average trend (e.g. a random walk or Simple Exponential smoothing (SES) type model, with or without growth). A model with two ($d = 2$) orders of total differencing assumes that the original series has a time-varying trend (e.g., a random trend or Linear Exponential Smoothing (LES) type model) [14].
3.2.4. Deciding AR and MA terms

After a time series has been stationarized by differencing, the next step in fitting an ARIMA model is to determine whether AR or MA terms are needed to correct any autocorrelation that remains in the differenced series. It is always easy to do this with software suites like R, where some different combinations of terms could be tested and see what works best. But by looking at the autocorrelation function (ACF) and partial autocorrelation (PACF) plots of the differenced series, the numbers of AR and/or MA terms can be tentatively identified in a more systematic way. The ACF plot is a bar chart of the coefficients of correlation between a time series and lags of itself. The PACF plots the partial correlation coefficients between the series and lags of itself [14, 15]. If the PACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is positive then adding an AR term to the model should be considered.

The lag at which the PACF cuts off is the indicated number of AR terms. In principle, any autocorrelation pattern can be removed from a stationarized series by adding enough autoregressive terms (lags of the stationarized series) to the forecasting equation, and the PACF tells you how many such terms are likely to be needed. However, this is not always the simplest way to explain a given pattern of autocorrelation. Sometimes it is more efficient to add MA terms (lags of the forecast errors) instead. The autocorrelation function (ACF) plays the same role for MA terms that the PACF plays for AR terms—that is, the ACF tells us how many MA terms are likely to be needed to remove the remaining autocorrelation from the differenced series. If the autocorrelation is significant at lag $k$
but not at any higher lags, this indicates that exactly $k$ MA terms should be used in the forecasting equation. In the latter case, we say that the stationarized series displays an “MA signature,” meaning that the autocorrelation pattern can be explained more easily by adding MA terms than by adding AR terms. An MA signature is commonly associated with negative autocorrelation at lag 1 i.e., it tends to arise in series which are slightly over-differenced. The reason for this is that an MA term can “partially cancel” an order of differencing in the forecasting equation [14, 30].

If the ACF of the differenced series displays a sharp cutoff and/or the lag-1 autocorrelation is negative then adding an MA term to the model should be considered. The lag at which the ACF cuts off is the indicated number of MA terms.

Mixed models: In most cases, the best model turns out to be a model that uses either only AR terms or only MA terms; although in some cases a “mixed” model with both AR and MA terms may provide the best fit to the data [30].

3.2.5. Model Fitting

For our data set, we fitted a time series model for each location along 1, 2, 3, 5 and 10-year period on a daily basis. Results (in Chapter 5) are presented for only model fitted to 3 years daily time-series of average daily temperature for 63 locations. To simplify the computations and due to the similarity between points of observations, each location is
fitted by an ARIMA model with the same set of parameters \((p, d, q)\). However, since coefficients \(\varphi\) and \(\theta\) may vary from one location to another, we allow some uniformity as well as some individuality in our model. Strong possibility of variation in the coefficients may be anticipated at locations over and near a large water body (Great Lakes in our case). A slowly decaying ACF (Figure 10) was observed for these locations.
Figure 10: ACF plots for locations 1-12
For a simple model, one may decide the \((p, d, q)\) value by looking at the ACF and PACF plot. But for this case, we plotted the ACF plot and found it hard to tell by just looking at these plots. Hence, we followed a stepwise incremental approach, adding terms of one kind or the other as indicated by the appearance of the ACF and PACF plots. We tried several different models on the data set and use AIC (Akaike Information Criterion) [31] value to help us to make a decision. H. Akaike in 1974 introduced AIC as a measure to solve the problem of statistical model selection. AIC criterion prefers a lower value. Summary of the AIC values for time series fitted at each location for different ARIMA models can be found in Table 2 (Appendix C).

<table>
<thead>
<tr>
<th>Model Parameters</th>
<th>(\varphi_1)</th>
<th>(\varphi_2)</th>
<th>(\theta_1)</th>
<th>(\theta_2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimates</td>
<td>0.3157</td>
<td>0.0039</td>
<td>0.5702</td>
<td>0.4137</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.0544</td>
<td>0.0388</td>
<td>0.0518</td>
<td>0.0521</td>
</tr>
<tr>
<td>(\delta^2 = 4.756)</td>
<td>log likelihood = -8022.08</td>
<td>AIC = 16058.17</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: ARIMA(2,1,2) model parameters

It is evident by looking at the AIC table that in terms of model selection based on AIC score, both ARIMA(2, 1, 2) and ARIMA(1, 1, 2) defeated ARIMA(2, 0, 2) and ARIMA(2, 1, 1) models. A closer look at the results reveal that, for some locations, performance of ARIMA(1, 1, 2) was better than ARIMA(2, 1, 2) with a fairly small
margin. For other locations, ARIMA(2, 1, 2) performs better with a relatively larger margin. This can be clearly observed for locations 50-63. Hence, we decided to experiment with ARIMA (2, 1, 2) model. Estimated parameters of model are shown in Table 1.
Chapter 4

4. Results and Discussions

4.1. The Data

Figure 11: Sensor grid for data
The point referenced or geospatial data that we have analyzed for spatio-temporal temperature forecasting consists of global daily average air temperature, amassed from heterogeneous sensors (Figure 11) for fifty years (1950-1999). Each sensor corresponds to a physical location (latitude, longitude) on the Earth's surface. The data is available in .csv format from the Oak Ridge National Laboratory website [1]. Temperature records are in Degrees Celsius (°C). The raw data has 18048 observations available for 18262 days. The 18048 observations are locations including 94 latitudes ranging from 90° N (88.5420°N) to -90°N (-88.5420°N) and 192 longitudes per latitude between 0°E and 360°E (358.1250°E); the days included are from 1950 to 1999.

In this work, we extracted daily temperature data of 10 consecutive years, 1990-1999 for 63 locations in USA as shown in Figure 11, spanning between latitudes 35.2° N and 46.7° N, and longitudes 75° W and 90° W.

4.1.1. Does Size Matter?

Sets of 1, 2, 3 and 5 years of data were used for our experiment and analysis. We selected 3 years of data for building Space-Time-Kriging model. Selecting large amount of data did not significantly improve the prediction results. Spatio-Temporal Kriging model built using 10 years of data took over an hour to produce results on a stand-alone PC/Mac with ~2/4GB of RAM and 32/64bit ~2.5 GHz. Intel Pentium Core 2 Duo. On the contrary, results were obtained in minutes (~ 4-5 minutes) for ARIMA(2,1,2) build using 10 years of data. We used software suite R to build the Spatio-Temporal Kriging and ARIMA
models. R is a free application available both for Mac and PC users for statistical computing and graphics [2].

4.2. Spatio-Temporal Kriging vs. ARIMA(2, 1, 2)

Prediction models were built using 1, 2, 3, 5 and 10 years of data. Finally, Spatio-Temporal Kriging model was built using 3 years (1997-1999) of data (Figure 12) and ARIMA(2, 1, 2) with 10 years of data (Figure 13). Daily average air temperature was predicted for last 3 days (December 29th, 30th, and 31st) of the year 1999. These years were chosen without any preference simply for illustrating the results in the document.

![Figure 12: Input data (1997-99) for Spatio-Temporal Kriging model](image-url)
For validation purpose, ground truth data from these days was removed and stored separately. Both the models were then constructed using the rest of the data. Temperature forecasts for 3 days were obtained from both Spatio-Temporal Kriging and ARIMA(2, 1, 2) models. The prediction results were then analyzed and validated against the ground truth data.
4.2.1. Prediction Results

Temperature forecasts obtained from Spatio-Temporal Kriging and ARIMA(2, 1, 2) models are shown in (Figure 14). Colored contours in the left column in of the image represent predictions from our Spatio-Temporal Kriging model. Contours for held out data are placed in between predictions from Spatio-Temporal Kriging and ARIMA(2, 1, 2) models. Even though both the models performed their best while predicting for Day1, the predictions from Spatio-Temporal Kriging model were closer to the actual observations. Further, due to modeled spatial information, our Spatio-Temporal Kriging model produced superior forecasts for Day 2 and 3.
4.2.2. Prediction Accuracy

To discuss the accuracy and reliability in prediction results from both the models we plot in (Figure 15), the difference between predicted and observed values for daily average
temperatures. For both prediction models we also compare the standard errors in prediction estimates with the help of side-by-side colored contour graphs.

4.2.2.1. **Difference in Prediction Results**

Prediction results obtained from both the models were subtracted from the observed values for the predictions. The difference was plotted with contour maps shown in (Figure 15). Differences in predictions for days 1 to 3 obtained from Spatio-Temporal Kriging model are shown on the left and from ARIMA on the right. Green area corresponds to the difference being zero or closer to zero. It may be easily observed that larger area in green is covered by Kriging plots. Larger difference in predictions can be noticed in ARIMA results. High amount of variability in predictions is observed in north and south east areas of the plots. As discussed earlier these areas correspond to the areas near or at a water body.
Figure 15: Difference between predicted and observed values.
4.2.2.2. **Standard Errors in Prediction Estimates**

Errors and uncertainty in prediction estimates are commonly reported with the help of standard error ($\bar{\sigma}_R$) for estimate and confidence intervals ($\bar{\nu} \pm \bar{\sigma}_R$). For 95% confidence interval we used $\pm 1.96\bar{\sigma}_R$. (Figure 16) displays the standard errors in prediction estimates obtained from Spatio-Temporal Kriging Model. Standard errors from ARIMA(2, 1, 2) model are shown in (Figure 17).

4.2.2.2.1. **Spatio-Temporal Kriging Standard Errors**

Plots in (Figure 16) correspond to location 1-12 (spanning between latitudes 44.8° N and 46.7° N, and longitudes 86.2°W and 90°W) respectively, with location 1 at Longitude = 86.2°W, Latitude = 44.8°N and location 12 at Longitude = 90°W, Latitude = 46.7°N. It shows 30 days of temperature values extracted from 3 years of data used for building the Space-Time Prediction model. Since we are predicting for 3 days in the future, we use 3 segments for each location. Tail of each segment indicates the standard error in prediction estimates and center represents the predicted daily average temperature. Innermost segment from the left correspond to first (day 1) predicted day and outermost segment being the last (day 3) predicted day.
Figure 16: Spatio-Temporal Kriging SEs in prediction estimates at locations (1-12).

From the plots we can easily notice the extra long-tailed segments at locations 3 and 12. Effect of lakes on temperature was observed at locations near and above water (locations 3 and 12) which concords with the observed slowly-decaying autocorrelation function.
We believe this effect to be the only known factor behind ±4.84°C of possible uncertainty in the predictions from Spatio-Temporal Kriging model. Our Spatio-Temporal Kriging model performed unexpectedly well when compared to ARIMA(2, 1, 2) model and Space-Time Linear model (please see (Appendix A) for formal description and results).

4.2.2.2.2.  **ARIMA(2, 1, 2) Standard Errors**

Standard errors for prediction estimates from ARIMA(2, 1, 2) are shown in (Figure 17). It is clearly visible from contour plots that minimum SEs are observed for day 1 prediction estimates. ARIMA(2, 1, 2) loses its accuracy in predictions as it advances into days 2 and 3. This is observed due to presence of spatial dependence in the data. As shown before, our Spatio-Temporal Kriging model is able to account for strong as well as subtle spatial covariance in the data.

![Figure 17: ARIMA(2,1,2) SEs for predicted Day 1, 2 and 3](image-url)
Average standard error (SEs) in predictions was 4.49°C for ARIMA(2, 1, 2) model. It can be concluded that the uncertainty in predictions from ARIMA(2, 1, 2) was ±8.80°C for 95% confidence interval. Best temperature forecasts were seen when predicting 1 day into the future at a time.
Chapter 5

5. Possible Novel Applications

In this section we discuss a few possible novel applications of our Spatio-Temporal Kriging model. Section 5.1 presents our model to be used as a virtual sensor. In section 5.2 we discuss the use of our model to predict the availability of natural energy such as sunlight. Harvesting of ambient energy is vital for sustaining a WSN.

5.1. KVSense: A Virtual Sensor

Our Spatio-Temporal Kriging model can be used to predict values at locations where observations don’t exist. It is an inherent property modeled into Kriging techniques that enable them to interpolate between the points of observations. There is no evidence of any published work utilizing Kriging or Spatiotemporal Kriging predictor to emulate the working of a real sensor. The spatial and temporal components of our Spatio-Temporal Kriging model can be used in conjunction or separately to replace faulty sensor(s). Due to
embedded spatial knowledge in our model, it can predict at unknown locations by deriving the value of a known random field, such as temperature. Thus we introduce a novel concept of virtual sensor or *KVSense* (Kriging based Virtual Sensor). *KVSense* can easily emulate the working of a real sensor at inaccessible areas. The idea is very motivating to extend this research and test our prediction model in the real world scenarios. We discuss this idea a bit further as one of the future research topics, in Chapter 6.

### 5.2. Energy harvesting in wireless sensor networks

A wireless sensor network (WSN) typically consists of a large number of small low-cost spatially distributed autonomous sensors to cooperatively monitor physical or environmental conditions over time without human intervention. The true autonomy of such systems depends on their reliable and unattended operability for extended times without maintenance. Thus, energy supply is a major design constraint of sensor network applications; the performance and lifetime of these networks are limited by the size, quantity and capacity of available finite-power sources.

Usual practices [23, 22] of energy harvesting in sensor networks address this issue at a systems level. Their primary focus is to exploit the energy resources efficiently rather than predicting the availability these resources. We believe that predicting the spatio-
temporal availability of ambient energy resources is vital to plan their optimum utilization, conservation and harvesting efficiently. Predicting the future availability of ambient energy in a region will also help us plan sensor rejuvenation. Sensor rejuvenation (recharging) may be understood as a process of fully utilizing the sensor’s stored energy with the assurance that new energy will be available in the future. Seasonal (SARIMA) models with monthly seasonality may be used in addition to our Spatio-Temporal Kriging model for more advanced planning of sensor rejuvenation.

Another approach to energy conservation [24] is scheduling node activity to use a subset of sensor nodes and putting the remaining nodes to sleep for reducing overall energy consumption.

Solar energy harvesting is the major source of power to the autonomous wireless sensors. Several solar energy aware routing techniques [20, 26, 27, 28, 29] along with energy aware clustering of sensor nodes [25] have been extensively researched and developed. All these techniques are well equipped to aid energy conservation. However, they fail to address the issue of predicting the spatial and temporal availability of solar energy. To address the energy challenge in WSN properly, it is necessary to unite these techniques with the prior knowledge on spatio-temporal availability of ambient energy. Our Spatio-Temporal Kriging model addresses this issue of energy prediction.
Our model can also apply if sensors are detecting sunlight intensity rather than temperature. We explain this idea by assuming that temperature is directly proportional to available ambient energy. Thus we can better predict the ambient energy for the sensors in the areas anticipated to have higher temperature. Spatial contour maps (Figure 14) from our results can be used to outline ‘good’ and ‘bad’ or ‘hot’ and ‘cold’ regions in the spatial domain constituting the sensor grid. Thus our Spatio-Temporal Kriging weather prediction model can assist a node to automatically adapt to the changing conditions and use its harvesting and scheduling abilities intelligently, thus increasing its task performance. Our model can also apply if sensors are detecting sunlight intensity rather than temperature values.

While this work was accomplished with the data associated to weather related parameters. These applications discussed above are only a glimpse of what our approach can offer. The approach and model we suggested here can find its applications in other real world scenarios, e.g., environmental data modeling, strategic and targeted marketing, agriculture, epidemiology, mining, etc.
Chapter 6

6. Conclusion and Future Work

“Neither the road ends here nor is it the destination, keep walking my friend the journey has just begun...”

In this research we modeled the spatio-temporal relationship in daily average temperature data collected over the period of 10 years at 63 locations in USA. We exploited the concept of modeling the space and time separately to fabricate Spatio-Temporal Kriging model. Spatial information was captured using Universal Kriging and temporal correlation by fitting the first order AR(1) autoregressive model. Further, to challenge the extent on reliability of forecasts from our model, we compared its results against the predictions obtained by building and fitting ARIMA model to the time series of the daily average temperature data. Relatively poor performance of commonly used ARIMA
process can be understood due to absence of seasonal component in the ARIMA model. Nature of problem addressed by this research was daily prediction of temperature rather than its seasonal behavior. Thus a non-seasonal ARIMA model was fitted and results were compared with Spatio-Temporal Kriging model.

Both the models were built using 1, 2, 3, 5 and 10 years of data. No significant gain was obtained in the accuracy of prediction results by using 10 years of data to build the model as compared to the model built with 2 or 3 years of data. Exponential increase in the time to build Spatio-Temporal Kriging model was observed as we increased the size of data used from 1 year to 3, 5 and 10 years. The experiments reveal that in terms of accuracy and reliability of temperature forecasts, Spatio-Temporal Kriging model performed twice as better than ARIMA(2, 1, 2) model. Average standard errors (SEs) in predictions for the former were 2.47°C and 4.49°C for the latter. In other words it may be stated that the uncertainty in predictions from our Spatio-Temporal Kriging model was ±4.84°C and ARIMA(2, 1, 2) was ±8.80°C for 95% confidence interval. Best temperature forecasts for both the models were observed when predicting 1 or 2 days into the future.

We also noticed the effect of lakes on temperature at locations near and above water. This was backed by the slowly decaying autocorrelation function observed at these locations. We expect it to be the only possible reason behind ±4.84°C of possible uncertainty in the predictions from our Spatio-Temporal Kriging model. Predictions could have been improved further by modeling the spatial and temporal variation and co-variation at these
locations separately. However, due already complex nature of the model, we did not
model them separately and decided to resolve this issue in our future endeavors.

A Space-Time linear model was also built using linear regression. As expected,
prediction results were quite different. This model was only able to account for the effect
of latitude on temperature, i.e. higher latitudes (Northern hemisphere) correspond to
lower temperature

We also introduced a new concept of virtual sensor or *KVSense* (Kriging based Virtual
Sensor). This virtual sensor may be used for temporarily replacing the faulty wireless
sensor(s) and for emulating the working of a real sensor at inaccessible areas. We would
like to test it on a configuration of a real WSN. We recommend building a virtual sensor
test-bed that can be configured to emulate network assembly of heterogeneous sensors.
Ideas addressing the energy issues in autonomous wireless sensors were also explored in
this research. In the future, we would like to build a framework for a decision support
system for WSN. This decision support system based on spatio-temporal knowledge of
energy availability will identify candidate cluster of sensors in a WSN recommend or
deploy the correct strategies such as virtual sensor (*KVSense*), adaptive topology, etc. for
optimum energy conservation, harvesting and utilization in WSN. This will result in high
and uninterrupted availability of a WSN by minimizing data loss due to sensor failure,
faulty sensor and also due to non-availability of sensor at inaccessible area.
Recent research has extended the work in developing Hierarchical Dynamical Spatio-Temporal Models HDSTMs [4]. These models exploit the notion of dynamic spatio-temporal process, modeling and hierarchical state-space framework. We would like to test these methods to model our as well as other interesting data. In this research we used only temperature data. We would like to model various other natural phenomena for our future work.

"It is our belief that nature has many more phenomena that can be modeled in harmony to reveal several hidden and fascinating patterns."
Bibliography


Appendix A: Space-Time Linear Model

Basic Model

The data used is same for Spatio-Temporal Kriging model and Space-Time linear model. Description on how the vector of observation and prediction was constructed can be found in Chapter 3.

The vector of observed values can be written as follows:

\[
y_0 = \begin{pmatrix}
y_{11} \\
y_{12} \\
\vdots \\
y_{nT}
\end{pmatrix}.
\]

The vector of observations to be predicted can be written as follows:

\[
y_p = \begin{pmatrix}
y_{1,T+1} \\
y_{1,T+2} \\
\vdots \\
y_{m,T+k}
\end{pmatrix}.
\]
One way to conceptualize Kriging is to think of the historical observations and observations to be predicted as being jointly multivariate normally distributed independently between points and time. As a result, we can write the joint distribution of $y_0$ and $y_p$ as

$$
\begin{pmatrix}
y_0 \\
y_p
\end{pmatrix}
\sim
N
\left( 
\begin{pmatrix}
\mu_0 \\
\mu_p
\end{pmatrix},
\sigma^2 I
\right)
$$

and our ultimate goal is to derive the expected value of $y_p$ (linear estimates).

First, we define a matrix of covariate values for the historical observations and the future observations. For the current model, this matrix of covariates includes terms for the following:

1. Intercept
2. Linear trend over time
3. Quadratic trend over time
4. Cubic trend over time
5. Linear longitudinal effect
6. Sine wave with an annual cycle
7. Cosine wave with an annual cycle

As a result, the covariate matrix of historical observations is an $nT \times 7$ matrix that looks like:
And the covariate matrix of future observations is an \( mk \times 7 \) matrix that looks like this:

\[
X_p = \begin{bmatrix}
1 & \frac{T+1}{365.25} & \left( \frac{T+1}{365.25} \right)^2 & \left( \frac{T+1}{365.25} \right)^3 & \text{lat}(s_1) & \sin \left( \frac{2\pi(T+1)}{365.25} \right) & \cos \left( \frac{2\pi(T+1)}{365.25} \right) \\
1 & \frac{T+2}{365.25} & \left( \frac{T+2}{365.25} \right)^2 & \left( \frac{T+2}{365.25} \right)^3 & \text{lat}(s_1) & \sin \left( \frac{2\pi(T+2)}{365.25} \right) & \cos \left( \frac{2\pi(T+2)}{365.25} \right) \\
1 & \frac{T+3}{365.25} & \left( \frac{T+3}{365.25} \right)^2 & \left( \frac{T+3}{365.25} \right)^3 & \text{lat}(s_1) & \sin \left( \frac{2\pi(T+3)}{365.25} \right) & \cos \left( \frac{2\pi(T+3)}{365.25} \right) \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
1 & \frac{T+k}{365.25} & \left( \frac{T+k}{365.25} \right)^2 & \left( \frac{T+k}{365.25} \right)^3 & \text{lat}(s_m) & \sin \left( \frac{2\pi(T+k)}{365.25} \right) & \cos \left( \frac{2\pi(T+k)}{365.25} \right)
\end{bmatrix}
\]

With these covariate matrices defined, we can write the mean vectors as a matrix product of the covariate matrices and a vector, \( \beta \), of seven unknown parameters:

\[
\mu_0 = X_0 \beta, \text{ and } \mu_p = X_p \beta.
\]

**Model Fitting**

First, we calculate the best linear unbiased estimators for \( \beta \):

\[
\hat{\beta} = (X_0^T X_0)^{-1} X_0^T y_0.
\]
Next, we calculate the linear regression estimates using the following formula:

\[ E(y_p) = \mathbf{X}_p \mathbf{\hat{B}}. \]

**Results**

The Space-Time linear model predictions for 3 days in future are compared against the actual observed temperature for same 3 days in (Figure 18). It is evident from the contour plot of predictions that linear model takes into account only the ‘Latitude’ effect on the temperature for given the spatial domain.

Figure 18: Prediction Results from Space-Time Linear model
Appendix B: Glossary of Terms

1. **Point Reference Spatial Data**: The variable (e.g. Temperature, precipitation, etc) studied in a geo-spatial region is present at each point in that region, but is observed only on few limited locations.

2. **Geostatistics**: Class of statistical methods for the analysis of data generated from a continuously-indexed spatial process.

3. **Intrinsically Stationary Process**: A spatial process with a constant mean function that has the property that the variance of the difference of the process at any two locations is a function of the lag between the two locations.

4. **Isotropic**: A property of a covariance function that indicates that the covariance of a process at any two locations is only a function of the distance between the two locations.

5. **Non-Negative-Definite**: A necessary and sufficient condition for a real-valued function to be a valid covariance function.

6. **Nugget Effect**: A non-negative quantity consisting of the measurement-error variance and any additional micro scale variation that cannot be attributed to measurement error.

7. **Second-Order Stationary Process**: A spatial process with a constant mean function and stationary covariance function.
8. **Separability**: A property of a space-time covariance function that indicates that the covariance function can be factored into the product of a spatial covariance function and a temporal covariance function.

9. **Stationarity**: A property of a covariance function that indicates that the covariance of a process at any two locations is only a function of the displacement between the two locations.

*Note: Most of the glossary terms have been adapted from [6].*
Appendix C: AIC Value Table

<table>
<thead>
<tr>
<th>Location</th>
<th>ARIMA(2,1,2)</th>
<th>ARIMA(1,1,2)</th>
<th>ARIMA(2,0,2)</th>
<th>ARIMA(2,1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>20088.27</td>
<td>20089.77</td>
<td>20091.46</td>
<td>20108.69</td>
</tr>
<tr>
<td>2.</td>
<td>19470.77</td>
<td>19470.38</td>
<td>19472.26</td>
<td>19490.5</td>
</tr>
<tr>
<td>3.</td>
<td>17461.05</td>
<td>17462.23</td>
<td>17463.09</td>
<td>17481.85</td>
</tr>
<tr>
<td>4.</td>
<td>19508.05</td>
<td>19507.63</td>
<td>19509.83</td>
<td>19535.96</td>
</tr>
<tr>
<td>5.</td>
<td>20292.82</td>
<td>20292.9</td>
<td>20295.33</td>
<td>20326.2</td>
</tr>
<tr>
<td>6.</td>
<td>21242.88</td>
<td>21245.14</td>
<td>21247.68</td>
<td>21275.78</td>
</tr>
<tr>
<td>7.</td>
<td>21404.15</td>
<td>21406.28</td>
<td>21408.81</td>
<td>21441.01</td>
</tr>
<tr>
<td>8.</td>
<td>21315.11</td>
<td>21317</td>
<td>21319.53</td>
<td>21356.15</td>
</tr>
<tr>
<td>9.</td>
<td>21325.99</td>
<td>21326.14</td>
<td>21328.68</td>
<td>21358.57</td>
</tr>
<tr>
<td>10.</td>
<td>20673.83</td>
<td>20075.24</td>
<td>20076.63</td>
<td>20099.15</td>
</tr>
<tr>
<td>11.</td>
<td>19333.61</td>
<td>19332.65</td>
<td>19334.41</td>
<td>19366.26</td>
</tr>
<tr>
<td>12.</td>
<td>17037.99</td>
<td>17037.06</td>
<td>17038.54</td>
<td>17058.98</td>
</tr>
<tr>
<td>13.</td>
<td>19201.12</td>
<td>19199.13</td>
<td>19201.22</td>
<td>19226.91</td>
</tr>
<tr>
<td>14.</td>
<td>19826.63</td>
<td>19824.83</td>
<td>19827.22</td>
<td>19853.46</td>
</tr>
<tr>
<td>15.</td>
<td>20302</td>
<td>20300.02</td>
<td>20302.46</td>
<td>20330.13</td>
</tr>
<tr>
<td>16.</td>
<td>20545.35</td>
<td>20543.91</td>
<td>20546.29</td>
<td>20577.58</td>
</tr>
<tr>
<td>17.</td>
<td>20614.96</td>
<td>20613.65</td>
<td>20615.98</td>
<td>20647.35</td>
</tr>
<tr>
<td>18.</td>
<td>21032.65</td>
<td>21030.83</td>
<td>21033.2</td>
<td>21051.01</td>
</tr>
<tr>
<td>19.</td>
<td>20149.74</td>
<td>20149</td>
<td>20150.14</td>
<td>20172.57</td>
</tr>
<tr>
<td>20.</td>
<td>19663.11</td>
<td>19661.32</td>
<td>19662.85</td>
<td>19691.94</td>
</tr>
<tr>
<td>21.</td>
<td>19568.29</td>
<td>19566.43</td>
<td>19568.05</td>
<td>19592.14</td>
</tr>
<tr>
<td>22.</td>
<td>19738.16</td>
<td>19736.17</td>
<td>19737.89</td>
<td>19763.67</td>
</tr>
<tr>
<td>23.</td>
<td>19886.89</td>
<td>19885.19</td>
<td>19887.16</td>
<td>19912.63</td>
</tr>
<tr>
<td>24.</td>
<td>20040.18</td>
<td>20038.29</td>
<td>20040.24</td>
<td>20069.95</td>
</tr>
<tr>
<td>25.</td>
<td>20224.9</td>
<td>20222.93</td>
<td>20224.44</td>
<td>20248.93</td>
</tr>
<tr>
<td>26.</td>
<td>20320.89</td>
<td>20319.22</td>
<td>20320.87</td>
<td>20343.92</td>
</tr>
<tr>
<td>27.</td>
<td>20269.43</td>
<td>20267.54</td>
<td>20269.46</td>
<td>20286.69</td>
</tr>
<tr>
<td>28.</td>
<td>20147.63</td>
<td>20145.65</td>
<td>20146.52</td>
<td>20170.09</td>
</tr>
<tr>
<td>29.</td>
<td>20180.23</td>
<td>20178.23</td>
<td>20179.34</td>
<td>20202.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>30.</td>
<td>20171.02</td>
<td>20169.13</td>
<td>20170.32</td>
<td>20198.46</td>
</tr>
<tr>
<td>31.</td>
<td>20135.42</td>
<td>20133.94</td>
<td>20135.16</td>
<td>20171.25</td>
</tr>
<tr>
<td>32.</td>
<td>20041.44</td>
<td>20040.39</td>
<td>20041.76</td>
<td>20087.5</td>
</tr>
<tr>
<td>33.</td>
<td>19896.88</td>
<td>19895.27</td>
<td>19896.48</td>
<td>19940.55</td>
</tr>
<tr>
<td>34.</td>
<td>19897.06</td>
<td>19895.06</td>
<td>19896.25</td>
<td>19923.18</td>
</tr>
<tr>
<td>35.</td>
<td>19866.75</td>
<td>19864.78</td>
<td>19866.18</td>
<td>19887.18</td>
</tr>
<tr>
<td>36.</td>
<td>19488.76</td>
<td>19486.83</td>
<td>19488.65</td>
<td>19503.18</td>
</tr>
<tr>
<td>37.</td>
<td>20050.14</td>
<td>20048.19</td>
<td>20048.65</td>
<td>20076.4</td>
</tr>
<tr>
<td>38.</td>
<td>20095.94</td>
<td>20094.17</td>
<td>20094.88</td>
<td>20120.63</td>
</tr>
<tr>
<td>39.</td>
<td>19986.99</td>
<td>19985.04</td>
<td>19985.83</td>
<td>20018.86</td>
</tr>
<tr>
<td>40.</td>
<td>19851.21</td>
<td>19849.88</td>
<td>19850.85</td>
<td>19891.17</td>
</tr>
<tr>
<td>41.</td>
<td>19714.1</td>
<td>19714</td>
<td>19714.97</td>
<td>19767.95</td>
</tr>
<tr>
<td>42.</td>
<td>19723.25</td>
<td>19721.29</td>
<td>19722.29</td>
<td>19766.22</td>
</tr>
<tr>
<td>43.</td>
<td>19946.24</td>
<td>19945.51</td>
<td>19946.51</td>
<td>19967.59</td>
</tr>
<tr>
<td>44.</td>
<td>19757.75</td>
<td>19757.43</td>
<td>19758.71</td>
<td>19774.31</td>
</tr>
<tr>
<td>45.</td>
<td>19150.93</td>
<td>19150.29</td>
<td>19151.78</td>
<td>19159.74</td>
</tr>
<tr>
<td>46.</td>
<td>19508.93</td>
<td>19509.63</td>
<td>19510.37</td>
<td>19538.11</td>
</tr>
<tr>
<td>47.</td>
<td>19441.6</td>
<td>19443.52</td>
<td>19444.46</td>
<td>19467.8</td>
</tr>
<tr>
<td>48.</td>
<td>19232.41</td>
<td>19232.03</td>
<td>19232.89</td>
<td>19263.5</td>
</tr>
<tr>
<td>49.</td>
<td>19024.4</td>
<td>19022.61</td>
<td>19023.37</td>
<td>19069.65</td>
</tr>
<tr>
<td>50.</td>
<td>19108.75</td>
<td>19106.85</td>
<td>19107.61</td>
<td>19170.25</td>
</tr>
<tr>
<td>51.</td>
<td>19521.37</td>
<td>19521.41</td>
<td>19522.15</td>
<td>19568.63</td>
</tr>
<tr>
<td>52.</td>
<td>19855.65</td>
<td>19857.89</td>
<td>19858.8</td>
<td>19890.32</td>
</tr>
<tr>
<td>53.</td>
<td>19635.15</td>
<td>19637.19</td>
<td>19638.03</td>
<td>19652.54</td>
</tr>
<tr>
<td>54.</td>
<td>18972.89</td>
<td>18976.61</td>
<td>18978.15</td>
<td>18982.99</td>
</tr>
<tr>
<td>55.</td>
<td>18899.36</td>
<td>18916.07</td>
<td>18917.15</td>
<td>18929.34</td>
</tr>
<tr>
<td>56.</td>
<td>18730.07</td>
<td>18744.98</td>
<td>18745.89</td>
<td>18762.02</td>
</tr>
<tr>
<td>57.</td>
<td>18534.73</td>
<td>18544.04</td>
<td>18544.85</td>
<td>18573.22</td>
</tr>
<tr>
<td>58.</td>
<td>18514.6</td>
<td>18518.83</td>
<td>18519.53</td>
<td>18569.62</td>
</tr>
<tr>
<td>59.</td>
<td>18731.21</td>
<td>18737.43</td>
<td>18738.05</td>
<td>18788.52</td>
</tr>
<tr>
<td>60.</td>
<td>18981.31</td>
<td>18988.11</td>
<td>18988.54</td>
<td>19028.85</td>
</tr>
<tr>
<td>61.</td>
<td>18979.89</td>
<td>18985.75</td>
<td>18986.02</td>
<td>19006.91</td>
</tr>
<tr>
<td>62.</td>
<td>18552.11</td>
<td>18556.22</td>
<td>18556.23</td>
<td>18570.54</td>
</tr>
<tr>
<td>63.</td>
<td>16254.88</td>
<td>16262.26</td>
<td>16262.33</td>
<td>16269.3</td>
</tr>
</tbody>
</table>

Table 2: Summary of AIC Values
Appendix D: Standard Errors (Spatio-Temporal Kriging Model)

The plots correspond to location 1-63 (spanning between latitudes 35.2° N and 46.7° N, and longitudes 75° W and 90° W) respectively. Each plot shows 30 days readings of temperature extracted from 3 years of data used for building the Space-Time Prediction model. Since we are predicting for 3 days in the future, 3 red segments are shown, indicating the standard error in prediction estimates at each location and for each predicted day. Innermost segment from the left correspond to first predicted day and outermost segment being the KSE for last predicted day.
Appendix E: ACF and PACF Plots

**ACF Plots:** The plots correspond to Auto Correlation Function (ACF) observed at location 1-63 (spanning between latitudes 35.2° N and 46.7° N, and longitudes 75° W and 90° W) respectively.
Long = -90, Lat = 39

Long = -88.1, Lat = 39

Long = -86.2, Lat = 39

Long = -84.4, Lat = 39

Long = -82.5, Lat = 39

Long = -80.6, Lat = 39

Long = -78.8, Lat = 39

Long = -76.9, Lat = 39

Long = -75, Lat = 39
**PACF Plots:** The plots correspond to Partial Autocorrelation Function (PACF) observed at location 1-63 (spanning between latitudes 35.2° N and 46.7° N, and longitudes 75° W and 90° W) respectively.
Long = -90, Lat = 42.9
Long = -88.1, Lat = 42.9
Long = -86.2, Lat = 42.9
Long = -84.4, Lat = 42.9
Long = -82.5, Lat = 42.9
Long = -80.6, Lat = 42.9
Long = -78.8, Lat = 42.9
Long = -76.9, Lat = 42.9
Long = -75, Lat = 42.9