AN ABSTRACT OF THE THESIS OF

CHRISTOPHER T.W. RENTOLA for the degree of Master of Science in Electrical and Computer Engineering presented on May 21, 1984.

Title: Real-Time Data Quality Assessment Using Linear Prediction

Abstract approved: Redacted for privacy
Sigurd L. Lillevik

The U.S. Air Force models many naturally occurring phenomena. To validate such models, an automated data collection system is used. Since validation is dependent upon correct data, an inexpensive method to detect erroneous data was investigated. The method uses a Wiener filter to predict values one sample in advance. If the predicted value differs from the measured value by some constant, then erroneous data are detected. When applied to the meteorological parameter "temperature", a sixth-order filter was designed that would flag good data as correct 97.64% of the time with an error flag set at three standard deviations. This thesis outlines the software requirements, design, and testing necessary to use this method.

# Real-Time Data Quality Assessment
# Using Linear Prediction

by

Christopher T.W. Rentola

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of
Master of Science

Completed May 21, 1984
Commencement June 1985

APPROVED:

Redacted for privacy

Professor of Electrical and Computer Engineering in charge of major

Redacted for privacy

Head of department of Electrical and Computer Engineering

Redacted for privacy

Dean of Graduate School

Date Thesis is presented May 21, 1984

Typed by Modern Office Management Systems for Christopher T.W. Rentola

ACKNOWLEDGEMENT

I would like to thank my major Professor S. Lillevik for his assistance and support throughout my graduate study and the U.S. Air Force for their financial support which made this research possible. Also, I would like to thank Professor R. Rathja for the many useful insights he provided in the area of digital signal processing. Finally, many thanks to Professor J. Saugen for the many helpful suggestions he provided.

TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# REAL-TIME DATA QUALITY ASSESSMENT USING LINEAR PREDICTION

## I. INTRODUCTION

The U.S. Air Force conducts research in the modeling of many naturally occurring phenomenon to predict sensor performance under varying conditions. Validation of such a model is currently under way at the Air Force's Targeting Systems Characterization Facility (TSCF) at Wright-Patterson Air force base in Dayton, Ohio. During data collection, there exists many possibilities for collecting erroneous data. Since model verification is dependent upon correct data, the quality of data must be assured. Thus, the subject under investigation in this thesis is the detection of erroneous data.

Beneficiaries of this research will be the Air Force's Targeting Systems Characterization Facility. Others who may use the results include anyone involved in data acquisition where the cost of hardware redundancy is prohibitive.

The remainder of this thesis is organized as follows. The test facility and a review of methods to detect erroneous data are described in Chapter II. Next, chapters

III and IV provide the required background on Wiener filtering theory and power spectrum estimation for presentation of a solution to detecting erroneous data. The overall software requirements and those unique to the test facility are given in Chapter V. Lastly, Chapter VI presents results obtained by this method for the selected model parameter "temperature".

## II. THE TEST FACILITY AND DATA QUALITY ASSESSMENT

The Department of the Air Force models many physical occurrences to describe the tactical environment. For example, models of the atmosphere assist pilots in determining optimal flight plans. Other examples include Tactical Decision Aids (TDA) which provide fighters with information on when, where, and how to attack a particular target. All models approximate a complicated system and require validation to determine their accuracy and limitations. To meet this need, the Air Force Avionics Laboratory developed a Targeting Systems Characterization Facility (TSCF) composed of several minicomputers, a four-tank scene or target, and an array of sensors. Besides validating several models, the TSCF laboratory allows for extensive evaluation of sensor requirements, precisions, and ranges. Both the Air Force Armament and Geophysics Laboratories use the data collected by the TSCF and it's instruments.

### The Targeting Systems Characterization Facility

The TSCF laboratory consists of a network of loosely-coupled minicomputers configured in a star architecture. Each remote node is physically separated from the central node but linked to the central node with a serial communications line.

Building 622 houses the central node which contains HP 2113 and 2112 minicomputers. As the central node, these minicomputers coordinate all data acquisition activities and enter collected data into a data base management system. This software package provides for efficient storage, manipulation, and retrieval of data.

Located in a tower attached to Building 622, the first remote node uses an HP 6942 Multiprogrammer to monitor several meteorological sensors. In addition, a forward looking infrared (FLIR) sensor scans the target scene and an additional hot-patch. Also, this tower holds a calibrated infrared source monitored by a transmissometer at the tank site. Approximately 2.25 km separates the remote node and the target scene.

Two additional remote nodes are located at the tank site. One of them contains various meteorological sensors plus several temperature sensors mounted on a tank. These sensors and the transmissometer connect to an HP 6942 Multiprogrammer which serves as the computer for this node.

Finally, the last remote node, also located at the tank site, contains an Inframetrics 210 radiometer which monitors the tank scene. The radiometer sits on a cherry picker which swings a 180 degree half-circle to obtain different tank profiles. Under control of a DEC 11/23, this remote node determines average radiance values and sends them onto the central node for insertion into the data base.

Typical data acquisition periods last from fifty-two (52) to ninety-seven (97) hours. During this time, there exists the possibility of collecting erroneous data. Since the TSCF collects data for model verification and sensor evaluation, the quality of data must be assured.

## Detecting Erroneous Data

From the beginning of computer-aided data acquisition, designers developed methods to check and sometimes correct for erroneous data. The overall approach involves partitioning the system into several sections or components, and then looking for the presence of erroneous data within each section. Thus, designers may view a distributed data acquisition system from various levels.

To detect erroneous data, two methods are usually considered: direct, and indirect. The direct method uses hardware redundancy to compare two samples of the same parameter. If the samples do not agree, then erroneous data have been found. Conversely, the second method, indirect, relies on a mathematical treatment of the present and past data. When the data fail to exhibit the expected characteristics, then erroneous data have been found. The direct method represents a simple but expensive approach while the indirect method involves a sophisticated but inexpensive technique. As with most decisions, each alternative presents some advantages and some disadvantages, and the specific application dictates the correct choice.

Typical direct methods use Triple Modular Redundancy (TMR) to both detect and isolate the problem. For example, Evans and Price[4] reported on a reliable computer using three central processing units. To minimize costs, designers often place redundancy only in critical sections and Hertel and Clark [7] developed a technique to assist in location of the TMR. All TMR schemes require a perfect voting mechanism to decide which modules agree and to isolate the faulty module.

Indirect methods usually fall under one of two techniques: statistical inference or digital signal processing. Often, designers use estimation theory to find an approximate form for the distribution and density functions of the mean time before failure (MTBF). This becomes difficult when dealing with limited sample sizes and Dey[3] proposed a solution to this problem. Using MTBF information, equipment may be removed from service prior to their expected failure for assured data quality.

In addition to MTBF, other statistics such as means, standard deviations, and correlations may be estimated using past data. Whenever the new data exhibit statistics beyond that which is expected, then the presence of erroneous data have been detected. For some parameters, the data excursions follow a repeatable and predictable pattern. Applying regression techniques, past data may be fit to curves using, for example, the method of least squares. During data acquisition, should the data not follow their expected curves, then the presence of erroneous data have been detected. Good, et al.[5], suggested that each sensor periodically read a calibrated source. Whenever these readings deviate too far, then the presence of erroneous data have been detected. All of the above methods utilize statistical inference theory to detect erroneous data and will sometimes miss bad data and flag good data.

Another indirect method to detect erroneous data relies on digital signal processing techniques. The techniques may be thought of as a collection of computer algorithms used for analysis or modification of discrete data. Analysis may be performed in the time or frequency domain. Time-domain analysis is associated with the field of statistics while frequency-domain analysis is connected to transform theory.

Frequency-domain analysis of discrete data became more popular after the invention in 1965 of the fast Fourier transform (FFT): an algorithm to efficiently calculate the discrete Fourier transform (DFT). The economy of the FFT algorithm provides an added dimension in the analysis of time-series data.

Discrete data modification or estimation may be accomplished with infinite impulse response (IIR) or finite impulse response (FIR) digital filters. They may be used for data smoothing, prediction, or filtering. Data smoothing requires knowledge of past, present, and future data and is usually associated with offline analysis of noisy signals. Prediction incorporates knowledge of past and present data to forecast future values. This is based on the assumption that future behavior is somehow related to past behavior. Finally, the filtering problem is usually associated with data modification in some predetermined way.

Over several years, many digital filters have been developed to fill various needs. A historical perspective of linear filtering theory and least-squares estimation may be found in Kailath[8] and Sorenson[17], respectively. Perhaps the most important filters were those developed by Wiener and Kalman to estimate signals immersed in noise. Both are optimal (in the least-squares sense) linear filters that may be used for smoothing or prediction.

Although both filters provide optimal results, they use different methods. The Kalman filter requires a state-space description of the signal and noise. Because of the Kalman filter's recursive solution, it is highly suitable for implementation on digital computers. Conversely, the Wiener filter requires the signal and noise to be described by covariance data and is not recursive. The choice between the two is governed by the available data. If state-space models are available, then the Kalman filter is preferred. But in many cases only covariance data is available so the Wiener filter is selected.

Since state-space models for meteorological parameters are not readily available, the one-step ahead Wiener filter was chosen to detect the presence of erroneous data. Thus, if the measured and predicted values differ by more than some predetermined constant, erroneous data are detected.

The Wiener filter represents an inexpensive, indirect solution for detecting erroneous data. It is an optimal linear filter for processes that are normally distributed. Since many naturally occurring phenomenon exhibit distributions that are nearly normal, the Wiener filter presents a logical solution in the detection of erroneous data.

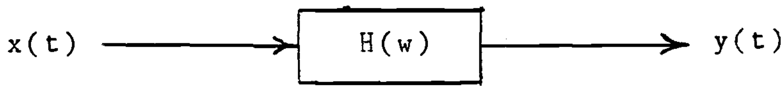
# III. THE WIENER FILTER

A fundamental filtering problem involves the estimation of a stochastic processes immersed in wide-bandwidth noise. One solution to this problem was provided by Norbert Wiener during the second World War to solve certain fire-control problems associated with anti-aircraft guns. Wiener's classical solution to the problem has come to be known as Wiener filtering theory. Fundamentally, a Wiener filter will provide the best possible linear estimate of a stochastic process immersed in noise. If all processes involved are Gaussian, then a Wiener filter provides the best estimate of all possible solutions. The following is a brief explanation of the fundamentals of Wiener filtering theory. For a more complete rigorous presentation on the continuous Wiener filter see Papoulis[13], Brown[2], or Srinath and Rajasekaran[18]. Detailed developments on the discrete Wiener filter may be found in Schwartz and Shaw[16] or Makoul[11].

In the following presentation of the continuous and discrete Wiener filter, the following assumptions will be made. First, the signal and noise are random processes with known power spectral densities or known auto and cross-correlation functions. For many practical engineering problems, these functions will not be known and must be estimated. Second, the performance criterion is the minimization of the mean-square error. This error criterion has the effect of weighting large errors heavily and small errors lightly. The effect is a spectral matching property that performs equally well over the entire frequency range. Third, the observations of the process consist of a linear combination of signal plus noise. Fourth, all random processes are assumed to be wide-sense stationary. For practical problems, the wide-sense stationarity criterion is usually only valid for short observation intervals.

## The Continuous Non-Causal Wiener Filter

The continuous non-causal Wiener filter is characterized by a knowledge of all past, present, and future observations of the process. For a sampled data system, this might be realized by having all observations stored on magnetic tape or disk. Figure 1 provides an illustration of the Wiener filter.




where:
$x(t) = f(t) + n(t)$
$y(t) \cong f(t + k)$
$f(t)$ = signal
$n(t)$ = noise
$y(t)$ = signal estimate

note:
$k > 0$ denotes the prediction problem
$k = 0$ denotes the filtering problem
$k < 0$ denotes the smoothing problem

THE CONTINUOUS NON-CAUSAL WIENER FILTER

FIGURE 1

For simplicity, only the filtering problem (k=0) will be presented for the continuous Wiener filter.

To determine the optimal transfer function H(w), an error function e(t) is defined as:

$$e(t) = f(t) - y(t) \qquad 1$$

The mean-square error is:

$$E[e^2(t)] = E[(f(t) - y(t))^2] \qquad 2$$

The optimal transfer function H(w), will be such that the mean-square error is a minimum. Expressing the output y(t) as a convolution integral yields:

$$y(t) = \int_{-\infty}^{\infty} x(t - c)h(c)dc \qquad 3$$

Where c is a dummy variable of integration.

Papoulis[13] has stated that by substituting Equation 3 into Equation 2 and extending the orthogonality principle to infinite sums, the optimal filter H(w) must be such that the observations x(t) are orthogonal to the error e(t).

$$E[[f(t) - \int_{-\infty}^{\infty} x(t - c)h(c)dc]x(t - T)] = 0 \quad \text{for all } T \qquad 4$$

This leads to:

$$Rfx(T) - \int_{-\infty}^{\infty} Rxx(T - c)h(c)dc = 0 \qquad \text{for all } T \qquad 5$$

Where Rxx(T) is the autocorrelation of x(t).

$$Rxx(T) = E[x(t)x(t + T)] \qquad 6$$

Taking the Fourier transform of Equation 5 yields:

$$Sfx(w) - Sxx(w)H(w) = 0 \qquad 7$$

Where the Fourier transform pairs are:

$$\begin{array}{lcl} Sfx(w) & \longleftrightarrow & Rfx(T) \\ Sxx(w) & \longleftrightarrow & Rxx(T) \\ H(w) & \longleftrightarrow & h(t) \end{array} \qquad 8$$

The optimum filter is found from Equation 7 to be:

$$H(w) = \frac{Sfx(w)}{Sxx(w)} \qquad 9$$

For the special case where the signal f(t) and noise n(t) are uncorrelated, that is:

$$Rfn(T) = 0 \qquad 10$$

then:

$$Sfx(w) = Sff(w) \qquad 11$$

$$Sxx(w) = Sff(w) + Snn(w)$$

and the optimal filter is:

$$H(w) = \frac{Sff(w)}{Sff(w) + Snn(w)} \qquad 12$$

The assumption of uncorrelated signal f(t) and noise n(t) is usually never strictly true, but one that must be made in many practical applications where the correlation between them is unknown.

## The Continuous Causal Wiener Filter

The causal Wiener filter is dependent upon all past and present observations. Once again, the criterion for optimality is the minimization of the mean-square error. The optimum causal Wiener filter must obey the orthogonality principle. That is, the optimal filter H(w) must provide a solution where the observations x(t) are orthogonal to the error e(t).

$$E[[f(t) - \int_{0}^{\infty} x(t - c)h(c)dc]x(t - T)] = 0 \quad \text{for } T >= 0 \qquad 13$$

Note the added constraint imposed by causality requires the solution to be valid only for T >= 0 and the limits of integration are from 0 to plus infinity. Equation 13 leads to the Wiener-Hopf equation:

$$Rfx(T) - \int_{0}^{\infty} Rxx(T - c)h(c)dc = 0 \qquad \text{for } T >=0 \qquad 14$$

The causality constraint makes the Wiener-Hopf equation considerably more difficult to solve.

One solution is based on spectral factorization. This technique, as given by Brown[2], requires that the spectral functions be divided into positive and negative time parts. The general approach proceeds by replacing the right side of Equation 14 with a function a(T) such that:

$$Rfx(T) - \int_{0}^{\infty} Rxx(T - c)h(c)dc = a(T) \qquad 15$$

where:

$a(T) = 0$ for $T >= 0$
$a(T) =$ unknown for $T < 0$
$h(t) = 0$ for $t < 0$

Taking the Laplace transform of Equation 15 yields:

$$Sfx(s) - Sxx(s)H(s) = A(s) \qquad 16$$

Using spectral factorization on Sxx(s), where the super "+" denotes positive time and the super "-" denotes negative time yields:

$$Sfx(s) - [H(s)Sxx^{+}(s)]Sxx^{-}(s) = A(s) \qquad 17$$

Rearranging Equation 17 to separate positive and negative time functions.

$$H(s)Sxx(s) = \frac{Sfx(s)}{Sxx^{-}(s)} - \frac{A(s)}{Sxx^{-}(s)} \qquad 18$$

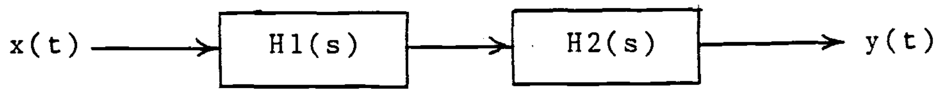
Equation 18 may be described in words as:

$$\left| \text{positive time function} \right| = \left| \text{positive and negative time functions} \right| - \left| \text{negative time function} \right|$$

The causal Wiener filter is obtained by equating the positive time functions.

$$H(s) = \frac{1}{Sxx^{+}(s)} \left| \text{positive-time part of } \frac{Sfx(s)}{Sxx^{-}(s)} \right| \qquad 19$$

Equation 19 may be viewed as two cascaded filters as shown in Figure 2.




where:

$$H1(s) = \frac{1}{Sxx^{+}(s)}$$

$$H2(s) = \text{positive-time part of } \frac{Sfx(s)}{Sxx^{-}(s)}$$

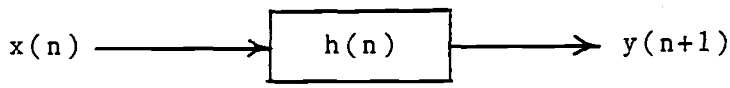
THE CONTINUOUS CAUSAL WIENER FILTER

Figure 2

H1(s) is a whittening filter that attempts to produce a flat spectrum as an input to H2(s). The second filter H2(s) is a filter whose frequency response is identical to that of the signal to be extracted from the noise. This view of Wiener filtering corresponds to the approach used in the simulation of stochastic processes driven by white noise.

## The Discrete Causal Wiener Filter

With the increasing use of digital computers, the discrete Wiener filter has become increasingly important. Again, the filter is dependent upon all past and present observations of the process. Since an infinite length filter is impossible to implement, the filter will be restricted to a length of P. The discrete, causal, one-step ahead prediction Wiener filter is shown in Figure 3.




where: $x(n) = f(n) + n(n)$
$f(n)$ = signal samples
$n(n)$ = noise samples
$y(n+1)$ = signal estimate one step in the future

THE DISCRETE WIENER FILTER

Figure 3

The signal estimate $y(n+1)$ may be written as a convolution summation.

$$y(n+1) = \sum_{k=0}^{P-1} h(k)x(n+k) \qquad 20$$

The error may be written as:

$$e(n) = f(n) - y(n) \qquad 21$$

The mean-square error is:

$$E[e^2(n)] = E[(f(n) - y(n))^2] \quad 22$$

Substituting Equation 20 into Equation 22 and changing the summation index yields:

$$E[e^2(n)] = E[(f(n) - \sum_{k=1}^{P} h(k)x(n-k))^2] \quad 23$$

The optimum filter parameters h(k), are found by taking the partial derivative of the mean-square error with respect to each of the filter parameters and setting the result equal to zero.

$$\frac{dE[e^2(n)]}{dh(k)} = 0 \quad \text{for } k = 1,2,...P \quad 24$$

This will result in a set of P normal equations in P unknowns. In matrix form these equations are:

$$\begin{vmatrix} Rxx(1,1) & Rxx(1,2) & \cdots & Rxx(1,P) \\ Rxx(2,1) & Rxx(2,2) & \cdots & Rxx(2,P) \\ & \vdots & & \\ Rxx(P,1) & Rxx(P,2) & \cdots & Rxx(P,P) \end{vmatrix} \begin{vmatrix} h(1) \\ h(2) \\ \vdots \\ h(P) \end{vmatrix} = \begin{vmatrix} g(1) \\ g(2) \\ \vdots \\ g(P) \end{vmatrix} \qquad 25$$

where: $Rxx(i,j) = E[x(i)x(j)]$
$g(i) = E[fx(i+1)]$

Since the coefficients Rxx(i,j) and g(i) are assumed to be known, the problem is reduced to solving the matrix equations for h(i). Since the Rxx(i,j) matrix is symmetric, positive definite, and Toeplitz, a recursive procedure due to Levinson[9] and written as a Fortran subroutine by Robinson[15] may be used to find the filter coefficients. This subroutine requires approximately P storage locations and will execute in about P-squared operations. Standard techniques require approximately P-squared storage locations and about P-cubed operations. An additional benefit provided by Robinson's routine is the output of a prediction error operator which may be combined with the autocorrelation coefficients to find the mean-square error.

# IV. POWER SPECTRUM ESTIMATION

In many practical engineering problems such as Wiener filter design, the true power spectrum is unknown. Therefore, many techniques have been developed to estimate the power spectrum. Optimum estimation techniques such as maximum likelihood estimation may be used when the general shape of the power spectrum is known. But in many cases the general form is unknown and estimation techniques must be used which have a strong empirical basis. For detailed explanations of the mathematics involved in power spectrum estimation, see Oppenheim and Schafer[12], Rabiner and Gold[14], or Schwartz and Shaw [16].

The objective of any estimation technique is to provide a consistent estimate where the bias and variance of the estimate tend to zero as the number of observations is increased. Specifically, the bias of an estimator is the difference between the true value of a parameter and the expected value of the estimate, and the variance is a measure of the width of the probability density function of the estimate. Many estimation techniques have the undesirable property that a reduction in variance is accompanied by an increase in the bias and vice versa. A good estimator may be thought of as one which produces a small mean-square error through some combination of variance and bias errors.

The Welch Method

With the invention of the fast Fourier transform (FFT), two spectral estimation techniques have come into widespread use. They are the Bartlett or correlation method and Welch's method of averaging modified periodograms. Only the Welch method will be presented due to its computational efficiency.

The periodogram is defined as:

$$In(w) = \frac{1}{N} \left| X(e^{jw}) \right|^2 \qquad 26$$

where:

$$X(e^{jw}) = \sum_{n=0}^{N-1} x(n)e^{-jwn}$$

is the discrete Fourier transform (DFT) of x(n). Oppenheim and Schafer[12] have shown that the periodogram is not a consistent estimator. As the value of N is increased, the variance of the estimate does not tend to zero. To reduce the variance, Bartlett introduced a method that reduces the variance by averaging independent periodograms.

To use the Welch method, which is a variation of Bartlett's method, the discrete data record of length N is sectioned into K segments of M samples.

$$K = \frac{N}{M} \qquad (27)$$

Welch's variation on Bartlett's method was to apply a window directly to each of the K segments to smooth out the discontinuities inherent in finite sampling. A DFT of each windowed sub-section is then calculated by use of the FFT.

$$X_k(e^{jw}) = \sum_{n=0}^{M-1} x_k(n)w(n)e^{-jwn} \qquad (28)$$

$x_k(n)$ = data subsection

$w(n)$ = data window

The periodograms of each sub-section are then averaged to obtain the estimate.

$$S_x(f_k) = \frac{1}{KU} \sum_{r=0}^{K} \left| X_r(k) \right|^2 \qquad 29$$

where:

$$U = \sum_{n=0}^{M-1} w^2(n)$$

is the energy in the window.

While appearing to be very simple, the Welch method contains many tradeoffs that must be considered in its application. For example, the selection of the sub-section length M has two conflicting requirements for a fixed record length N. The requirement for reduced bias and good spectral resolution calls for long data segments M. But increased data segment length M reduces the number of periodograms that may be averaged and increases the variance. Since variance reduction is proportional to 1/K compared to a single periodogram, a large number of segments is desired. Usually, the required spectral resolution, which is equal to fs/M, where fs is the sampling rate, is determined by prior knowledge of the signal being analyzed. If there is complete lack of knowledge about the signal being analyzed, then the sub-section length M should be varied over several orders of magnitude. This depends upon having enough data available.

Another question arises as to what window should be applied to the data. Harris[6] has provided an excellent summary of many popular windows. They may be thought of as weighting functions applied to the data which reduce the spectral leakage associated with finite-duration records. Spectral leakage, which is the output of one frequency component at a different frequency, creates a bias in the estimate of the power spectrum. This bias is most severe in the detection of small-amplitude components in the presence of large-amplitude components. One pitfall associated with bias reduction by the use of windows involves distortions introduced by the windows. For example, if a significant event occurs near the window boundaries, then it may be missed entirely due to the small values exhibited by the window near its edges. To avoid this problem, data segments are often overlapped with the added benefit of further variance reduction. The increased variance reduction occurs because overlapping increases the number of periodograms available for averaging. However, overlapping may cause the segments to become dependent which violates the premise of independent periodograms required by the Welch method.

According to Schwartz and Shaw[16], bias reduction may also be achieved through preprocessing of the signal. Preprocessing involves removing known trends or average components prior to power spectrum estimation. Removal of known trends requires that the original signal be passed through a filter whose frequency response is the inverse of the known spectral components. This process, also known as prewhittening, attempts to produce a flat spectrum. If the trends are unknown, then they may be estimated by methods previously described.

# V. ANALYSIS AND DESIGN SOFTWARE

The software package designed for Data Base Quality Assurance (DBQA) had to meet several requirements including data analysis in the time and frequency domain plus Wiener filter design and testing. Time domain analysis was necessary to find the probability distribution function and to determine if the available data was representative of the true process. While the time domain analysis was intended to answer the question of suitability, the frequency domain analysis was required to estimate the power spectrum of the process from which the Wiener filter was designed. The design of the Wiener filter involved solving a set of matrix equations to obtain the filter coefficients. These coefficients were then used to implement the filter as a difference equation from which real data was tested. Of the several meteorological parameters available, only temperature was analyzed due to time constraints. The application of the following methods may be easily extended to the other parameters. Real-time implementation and testing of the filter did not occur because of the distance between Oregon State University (OSU) and Wright-Patterson Air Force Base (WPAFB).

## Required Equipment

The implementation of the DBQA software package required certain computing equipment and capabilities. The following equipment was used:

1. Hewlett-Packard (HP) 1000 series minicomputer.
2. Image 1000 data base software.
3. Real Time Executive (RTE) 6 operating system.
4. Meteorological data from WPAFB.
5. Fortran 77 compiler.
6. Graphics terminal.
7. Graphics plotter.
8. Line printer.

Although most computers would have the capability to implement the necessary software, Hewlett-Packard equipment and software were selected because of its use at the WPAFB test facilities and its availability at OSU.

The meteorological test data, along with data base access routines, were transported from WPAFB to OSU during the Fall of 1983. One of the major problems in the adaptation of the WPAFB software was the lack of disk space on the OSU system. Of the 1000 tracks required to mount the entire data base, only 200 tracks were available. The lack of disk space considerably reduced the amount of data available for analysis. Fortunately, enough test data was placed on disk to permit the design of a Wiener filter.

A major consideration in the design of any software package centers around the operating system. OSU's HP-1000 uses the RTE-6 operating system which implements Extended Memory Access (EMA) for large data arrays required by the DBQA software package. In addition, the RTE-6 operating system provides for program segmentation or virtual memory access which allows machines with limited memory to execute large programs.

Other equipment required were a graphics terminal and plotter. Visual display of the data assists in detecting trends not easily discovered from numeric listings. A line printer was required for software debugging and listing data to obtain accuracy not obtainable with plots.

## The Program Root

The large size of the DBQA software package made segmentation necessary. At the center of the package was the program root which controls the entire package. It was responsible for passing parameters between segments as well as overlaying segments. While the number of segments required will vary at other computer installations, the functions to be performed must meet certain requirements.

Reading the Data Base

An important component of the DBQA package was the segment to read the data base and store the required parameters in data buffers for analysis by other segments. Ideally, proper analysis of the meteorological data would require many years of data periodically sampled throughout the day. The sample rate should satisfy the Nyquist criterion which requires the sample rate to be at least twice that of the highest frequency component of the signal. The analysis would then be performed on the same time slice over a period of many years. Since data of this type was unavailable, the analysis was performed on the parameter temperature over a six-week period where it was most stationary. This required the read segment to be capable of reading and sorting many consecutive days of data. Since the data were often sampled at different rates, proper placement in the data buffers was accomplished by reading the sample times at which the data were taken. The incompleteness of the data required a linear interpolation routine to estimate small amounts of missing data. Large blocks of missing data were flagged so the entire day could not be used.

## Time Domain Analysis Requirements

The next segment implemented was used to analyze the time-domain statistics. Two goals of this segment were to establish a probability distribution function and to determine if the available data were representative of the true process. If the distribution function had already been established, then this segment would not be required to verify the suitability of a Wiener filter. Specifically, this segment estimated the parameter's mean, variance, skewness, kurtosis, standard errors, and formed a histogram. Algorithms for calculating these parameters from weather data may be found in Brooks[1].

Estimation of the distribution function may be accomplished in several ways. The chi-square method requires the data to be divided into classes and the number of occurrences counted to establish a histogram. Then, the observed data are compared to the expected values for each class of the hypothesized distribution.

The two results are then compared with a chi-square test to establish confidence intervals. Since only seventeen days of reasonably complete data were available, it was felt that a simpler technique should be used. According to Brooks[1], the skewness of a random sample from a normal population will not exceed in magnitude 1.96 times its standard error with a confidence level of 95 percent. The hypothesis of normality was reinforced by the histogram.

To determine if the available data represented the true process, the United States Weather Bureau statistics were consulted. Typically, the published figures show highs, lows, and daily average temperatures that were compared to the observed values with a t-test. Although these parameters were not sufficient to establish whether the data were representative of the true process, they contributed to the analysis. Representative data would permit the designed filter to be used for the period in general.

## Frequency Domain Analysis Requirements

Perhaps the most important program component was the segment to estimate the power spectrum. It made extensive use of the Signal Processing Library routines on the

HP-1000 to calculate FFT's and inverse FFT's. These routines are available in Rabiner and Gold[14] and Robinson[15]. Since the power spectrum of temperature was unknown, this segment was designed to use different length periodograms. The incompleteness of the data limited their length to 144 sample points obtained from 10-minute samples over a 24-hour period. Also, the effects of averaging different numbers of periodograms was studied. The presence of correlated noise will deteriorate the estimate if too many periodograms are averaged.

Another consideration for estimation of unknown power spectrums involves window selection. Different windows permit various features of the power spectrum to be exhibited. Although there are many possible windows, those selected were the rectangle, triangle, and four-term Blackman-Harris. Selection of the four-term Blackman-Harris was based on a study by Harris[6] to resolve two nearby peaks. Of the windows evaluated, the performance of the four-term Blackman-Harris was found to be superior in the resolution of two nearby peaks. The rectangular window was selected because it does not introduce distortions in the time-series data. This was important because overlapped periodograms were not used.

Finally, the triangular window was selected because of the simplicity in generating its coefficients.

Another essential feature of this program segment was the capability to estimate the autocorrelation function. The autocorrelation estimate was obtained by taking the inverse DFT of the averaged spectral estimate.

Estimation of an unknown power spectrum is an art-form and proper use of this segment is difficult. Generally, different length periodograms are calculated and the results compared until a general shape emerges. To smooth the spectrum, different windows should be used and the results compared. Also, the effect of averaging different numbers of periodograms should be studied. Once the estimate of the power spectrum is obtained, the autocorrelation coefficients can be calculated from which the Wiener filter is designed.

## Wiener Filter Design Requirements

Once an estimate of the autocorrelation function was obtained, the process of designing the Wiener filter was straight forward. Since the signal f(t) and noise n(t) were assumed to be uncorrelated, the autocorrelation coefficients obtained for temperature represented g(i) in Equation 25. The one-step ahead prediction filter required that:

$$g(i) = Rff(i+1) \qquad 26$$

The left side matrix coefficients were obtained by an inverse DFT of the power spectrum of temperature plus noise. That is:

$$Rxx(i,j) = IDFT[\ Sff(w) + Snn(w)\ ] \qquad 27$$

Since the noise was assumed to be of unit intensity and white, the value one was added to each spectral sample of Sff(w). Therefore, the problem was reduced to solving the matrix equations for the unknown filter coefficients by use of Robinson's[15] subroutine EUREKA.

An added benefit of EUREKA is the output of error coefficients which were used to obtain the optimal filter length. The optimal length was obtained from a plot of error versus the filter length. Typically, the curve exhibits a large decrease in error for the first several coefficients and flattens out as the number of coefficients increases. The optimal filter length is then obtained from the knee of the curve.

Wiener Filter Testing Requirements

Another essential component of the DBQA software package involves testing the filter on real data. This segment's purpose was to determine performance levels of the filter and establish levels to flag erroneous data. To accomplish this task, the filter was implemented as a difference equation using the previously obtained filter coefficients. Before testing could proceed, it was necessary to remove the mean value from the data to reduce the bias of the prediction error. Since it was desired to simulate the performance in a real-time environment, the mean of the current test point was estimated by a moving average whose length was equal to the optimal filter length. The filter was then tested on this data while the errors were stored in a file for analysis. Restoration of the mean values would be required to obtain the actual predicted values. Statistics such as the mean, variance, skewness, and kurtosis, were compiled to determine the levels to flag erroneous data.

# VI. A WIENER FILTER FOR TEMPERATURE PREDICTION

While the previous chapter outlined software requirements to implement a Wiener filter for detection of erroneous data, this chapter presents results obtained from the analysis of temperature.

## Time Domain Results

Time-domain analysis of temperature was performed with two goals in mind. The first was to find the probability distribution function to determine the appropriateness of a Wiener filter as a solution to the prediction problem. The second was to determine if the available data were representative of the true process.

According to Brooks[1], air temperature is best described by a normal distribution function. Therefore, a simple test was used to validate this hypothesis. Using a 95% confidence level, the magnitude of the data skewness will not exceed 1.96 times its standard error for a normal distribution. For the seventeen days analyzed in Dayton, Ohio from December 18, 1982 to February 1, 1983, the air temperature could be considered normal from 0000 Greenwich

Mean Time (GMT) to 1250 GMT and not normal from 1300 GMT to 2350 GMT. Therefore, the Wiener filter offered a slightly sub-optimal solution.

Another important consideration was to determine if the available data were representative of the true process. Representative data would permit a filter to be designed that would be valid every year for the period being analyzed. Intuitively, one would not expect temperature to be stationary over a period of several months. Therefore, a period near the winter solstice was selected were the temperature was most stationary. Table 1 shows the days analyzed along with normal temperatures for Dayton, Ohio compiled by the United States Weather Bureau for the period 1921 to 1950.

TABLE 1

Average Daily Temperatures in Degrees Centigrade for Dayton, Ohio for the Years 1921-1950

| DATE | MAXIMUM | MINIMUM | AVERAGE |
|---|---|---|---|
| DEC 18 | 3.3 | -5.0 | -0.6 |
| DEC 25 | 3.3 | -5.0 | -0.6 |
| DEC 27 | 3.3 | -5.0 | -0.6 |
| DEC 29 | 3.3 | -5.0 | -0.6 |
| DEC 31 | 3.3 | -5.0 | -0.6 |
| JAN 1 | 3.3 | -5.0 | -0.6 |
| JAN 2 | 3.3 | -5.0 | -0.6 |
| JAN 4 | 3.3 | -5.6 | -1.1 |
| JAN 6 | 3.3 | -5.6 | -1.1 |
| JAN 20 | 2.8 | -5.6 | -1.1 |
| JAN 22 | 2.8 | -5.6 | -1.1 |
| JAN 23 | 2.8 | -5.6 | -1.1 |
| JAN 24 | 2.8 | -5.6 | -1.1 |
| JAN 25 | 2.8 | -5.6 | -1.1 |
| JAN 29 | 2.8 | -5.6 | -1.1 |
| JAN 30 | 2.8 | -5.6 | -1.1 |
| FEB 1 | 2.8 | -5.6 | -1.1 |

During the 1982-1983 period, these days were found to have an average temperature of 1.46°C, a maximum of 4.45°C, and a minimum of -0.45°C. To compare these values for those from the U.S. Weather Bureau, a t-test was used.

The t-test is used to accept or reject the hypothesis that two normally distributed populations have equal means. To use this test the quantity t is found as:

$$t = \frac{\text{deviate}}{\text{standard error}} = \frac{M - A}{S/\sqrt{n}} \qquad 30$$

where:

M = postulated mean
A = sample mean
S = sample standard deviation
n = number of samples

Using the data from Table 1 and the calculated averages, the following values of t were found:

Average High: t = 1.06
Average Low: t = 3.14 31
Daily Average: t = 1.70

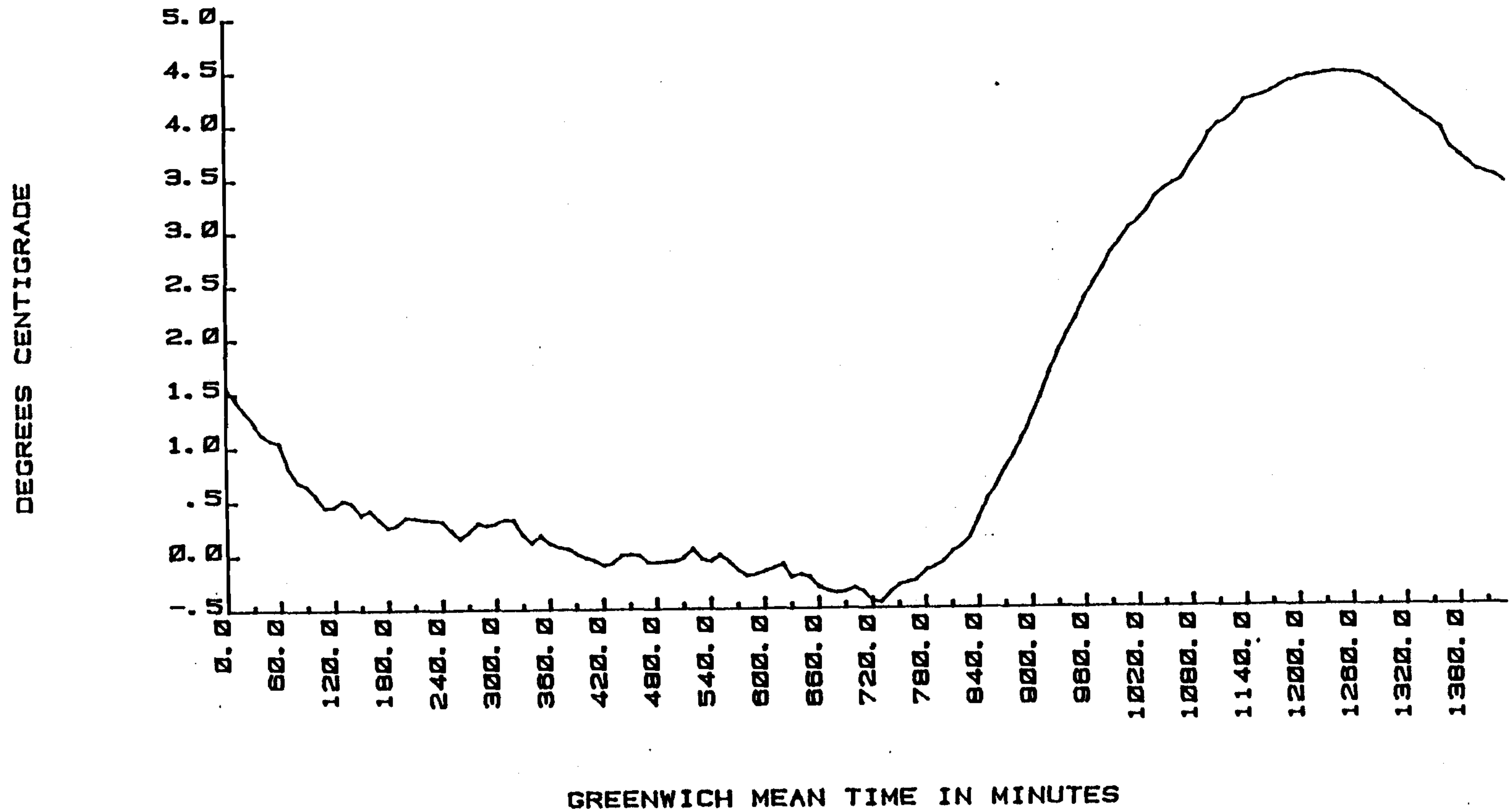
With a significance level of 95% and 16 degrees of freedom (obtained from 17 data points minus 1), the critical value of t for a two sided test was found in Maisel and Gnugnoli[10] to be 2.120. If the magnitude of t is smaller than 2.120 the hypothesis was accepted while a larger magnitude would indicate rejection. The values of t for the daily average and average high are in agreement with the hypothesis while the average low temperature is not.

Figure 4 shows the calculated temperature averages for each ten-minute sample point. Although the endpoints of the curve are separated by 1.9°C, the general shape agrees with that given by Wenstrom[19]. Since there is considerable doubt as to whether the data analyzed was representative of the December 18 to February 1 period in general, the filter designed was valid only for the days analyzed.

There are several possible explanations for these results. The most probable reason is that average daily temperatures are highly correlated. Although the period analyzed spanned six weeks, 47% of the data came from a period less than two weeks. This was unavoidable due to the inconsistencies of the available data.

## Frequency Domain Results

Analysis of temperature in the frequency domain focused on estimation of the power spectrum. Using the estimated power spectrum, the autocorrelation function was obtained from which the Wiener filter was designed.




ESTIMATED AVERAGE TEMPERATURE

FIGURE 4

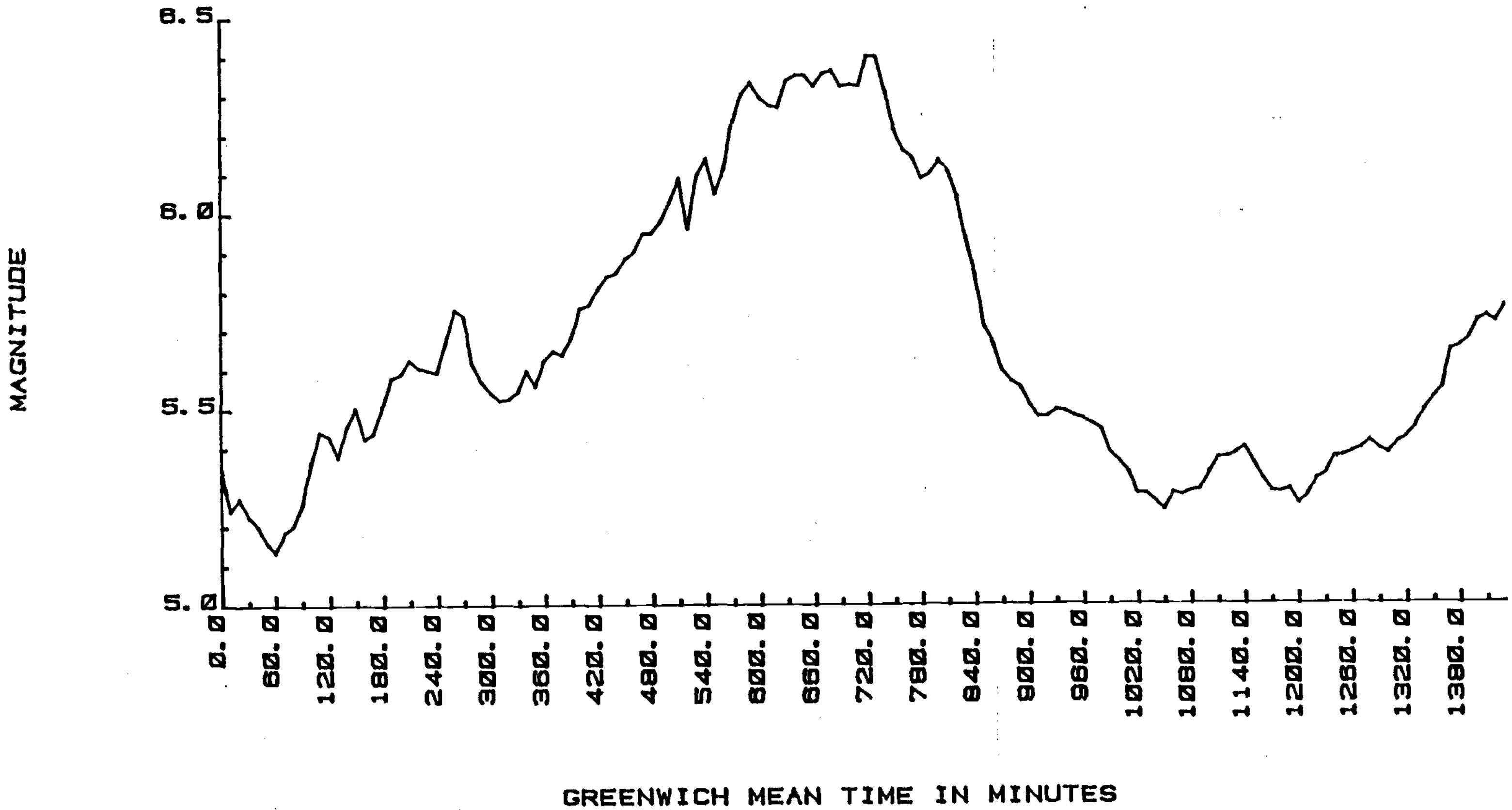
Welch's method of modified periodograms was used to estimate the power spectrum without first preprocessing the signal. Seventeen periodograms were calculated from the days shown in Table 1 for the years 1982 and 1983. Different numbers of periodograms were averaged to check for the presence of correlated noise. As more periodograms were averaged, the variance of the estimate decreased while the magnitude remained fairly constant. This helped support the assumption that the noise was uncorrelated.

For reduced spectral leakage, three windows were used to obtain three separate estimates. The Blackman-Harris window produced the smoothest estimate which showed the presence of a large peak near a large DC component.

To reduce the bias of the estimate near DC, the average daily temperature was removed from each day before the periodograms were calculated. Again, different numbers of periodograms were averaged with similar results. The three windows produced results somewhat different from those previously obtained. When seventeen periodograms were averaged, the rectangular window produced the smoothest estimate. A plausible explanation is that once the average was removed there was less of a transition between the presence and absence of data thereby decreasing the proportion of variance due to finite sampling. Since both the triangular and Blackman-Harris windows introduced distortions in the data to smooth the transition, they overemphasized the temperatures near the center of the periodogram. The standard deviation of temperature for the days analyzed is shown in Figure 5. It indicates that the temperatures near the center of the periodogram have a larger standard deviation thereby causing the increased variance in the power spectral estimate obtained from these two windows.

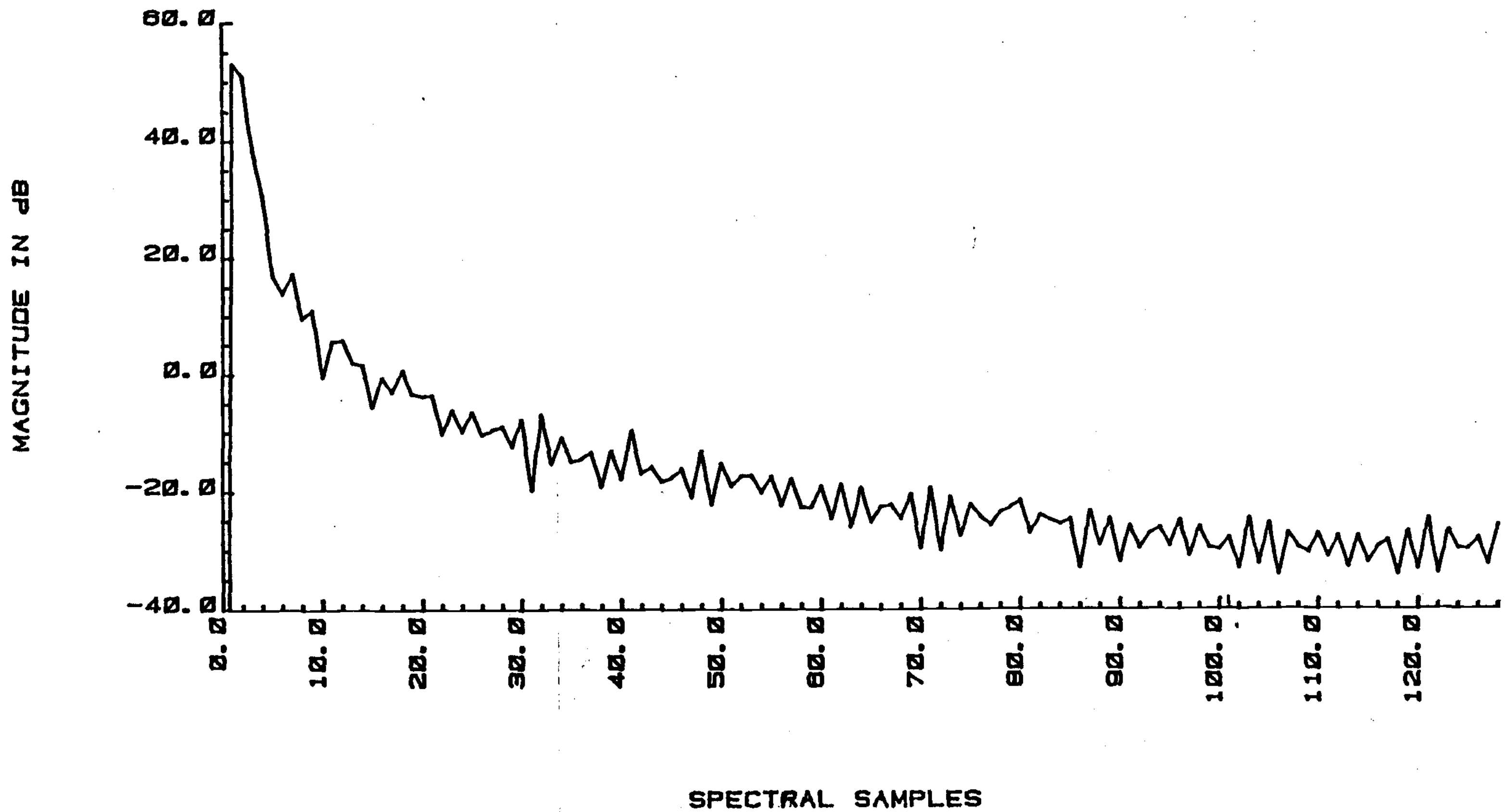
The power spectrum estimate for temperature obtained with a rectangular window is shown in Figure 6. It was obtained by averaging seventeen periodograms calculated with 256-point DFT's. Since each periodogram was based on 144 data points, it was necessary to append the data with zeros to get the next power of two required by the FFT algorithm.

To find the frequency of the peak in Figure 6, it was first necessary to find the maximum frequency. This was found by dividing the sampling rate of ten minutes (1.67E-3 Hz) by 2 which gives 3 cycles-per-hour or 72




ESTIMATED STANDARD DEVIATION OF THE TEMPERATURE

FIGURE 5



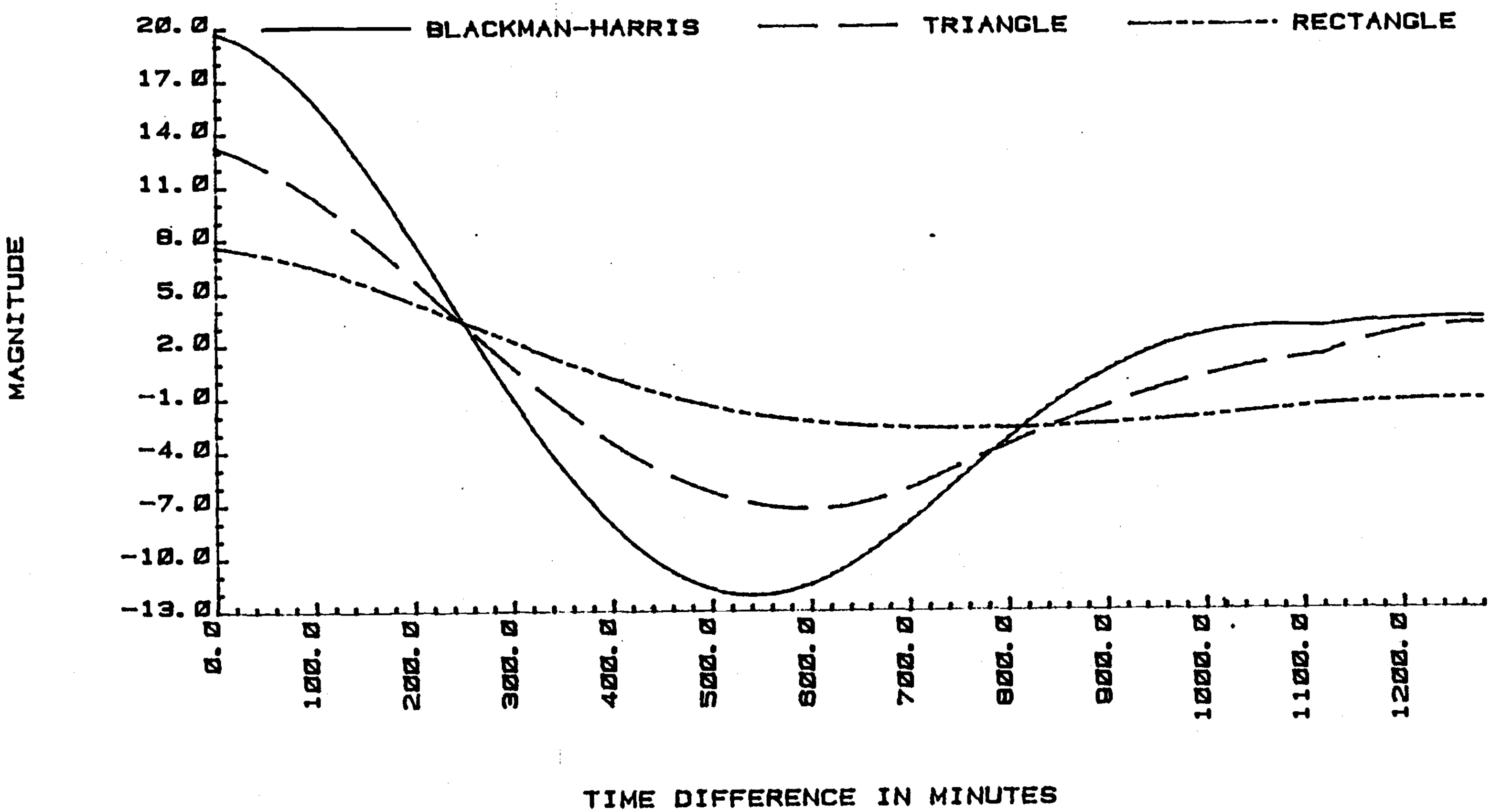

ESTIMATED POWER SPECTRUM OF THE TEMPERATURE

FIGURE 6

cycles-per-day. The frequency represented by each division was found by dividing 72 cycles-per-day by 128 for 0.56 cycles-per-day, per division. Most of the energy in the large peak of Figure 6 is between one and four samples which corresponds to the daily cycle associated with the sun. Better resolution of this peak could have been obtained with longer periodograms but would have required many days of uninterrupted data that was not available. Also, any frequency component present in the data higher than 3 cycles-per-hour would result in alaising which also distorts the estimate.

Averaging periodograms has the effect of reducing the noise in the specrtal estimate, but noise associated with temperature has a meaning much different than that found in electrical circuits. It may be thought of as anything affecting the temperature not correlated with the daily cycle. Examples might include cloud cover, rain, or wind speed.

Once an estimate of the power spectrum was found, the autocorrelation function was obtained by use of an inverse DFT. Three functions obtained from spectral estimates utilizing different windows are shown in Figure 7. As expected, the temperature was found to be highly




ESTIMATED AUTOCORRELATION OF THE TEMPERATURE

FIGURE 7

correlated for small time differences. An unexpected result was the negative correlation shown for large time differences. According to Schwartz and Shaw[16], the subtraction of the estimated mean from each data point was responsible for the negative dip. Since the autocorrelation function for temperature was not expected to have negative values, the one obtained by use of the rectangular window was selected for the design of the Winer filter.

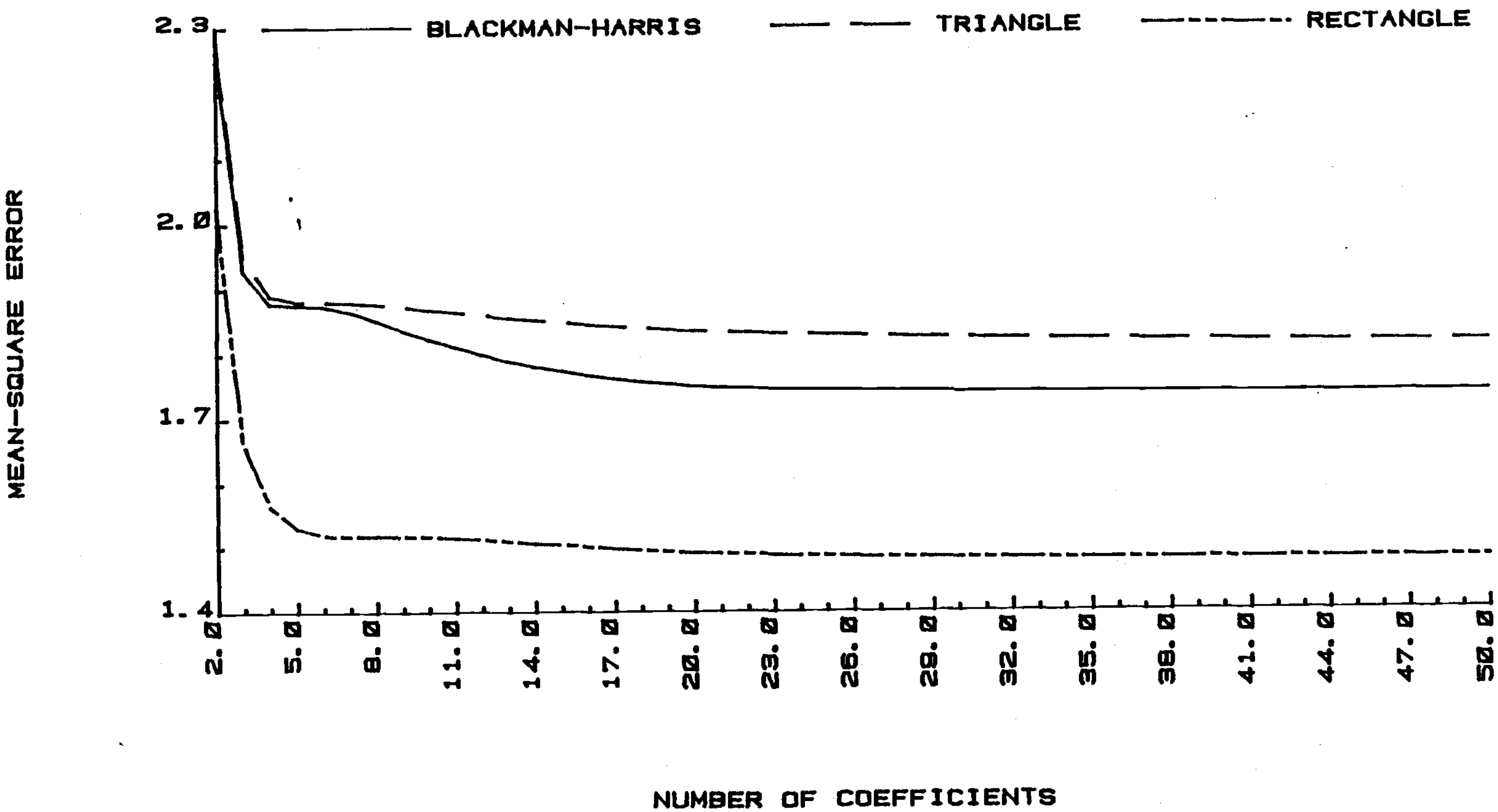
## Wiener Filter Design Results

Design of the Wiener filter required solving Equation 25 by use of Robinson's[15] subroutine EUREKA. Since the signal and noise were assumed to be uncorrelated, the right side coefficients of Equation 25 were obtained from the spectral estimate of temperature. To simulate signal plus white noise, the left side coefficients were found by an inverse DFT of the temperature spectrum with one added to each sample. The design process not only required solving the matrix equations but the determination of the optimal filter length for implementation.

In determining the optimal filter length, it was required to solve the matrix equations using different filter lengths while recording the mean-square error each solution produced. The graph of filter length versus mean-square error is shown in Figure 8. This graph shows a sharp decrease in mean-square error for the first few coefficients and a gradual decrease thereafter.

The optimal filter length was selected to be six which is on the lower part of the curve's knee for the rectangular window spectral estimate. Even though the graph indicates that a longer filter would have less error, this is not always true. Since the spectrum used to generate the coefficients contained noise, increasing the filter length would cause its frequency response to follow the noise. Even if the model's spectrum were completely smooth, too long a filter length would provide a good match at the sample points but a poor match between them.

The optimal filter coefficients provided by the subroutine EUREKA were:




WIENER FILTER LENGTH VERSUS MEAN-SQUARE ERROR

FIGURE 8

TABLE 2

Optimal Filter Coefficients for Temperature Prediction

$f1 = 0.3517$
$f2 = 0.2346$
$f3 = 0.1596$
$f4 = 0.1076$
$f5 = 0.0678$
$f6 = 0.0374$

Using these coefficients, the difference equation implementing the one-step-ahead prediction Wiener filter is:

$$y(n) = 0.3517x(n-1) + 0.2346x(n-2) + 0.1596x(n-3) + 0.1076x(n-4) + 0.0678x(n-5) + 0.0374x(n-6) \qquad 32$$

where: $y(n)$ = current data point being predicted
$x(n-1)$ = last observed data point

Since Equation 32 uses the last six values to predict the next value, all values of temperature past one hour were disregarded in the prediction.

To determine the transfer function of the filter, it was necessary to take the z-transform of Equation 32.

$$H(z) = \frac{f1z^5 + f2z^4 + f3z^3 + f4z^2 + f5z + f6}{z^6} \qquad 33$$

Equation 33 shows the structure of the filter is FIR, which is always stable since its poles are located inside the unit circle at zero. Then, using IBM's subroutine POLRT given by Robinson[15], the zeros of the filter were found to be:

TABLE 3

Zeros of the Optimal Filter

$$R1 = -0.6270$$

$$R2 = 0.3044 + j0.5794 = 0.6545e^{+j1.087}$$

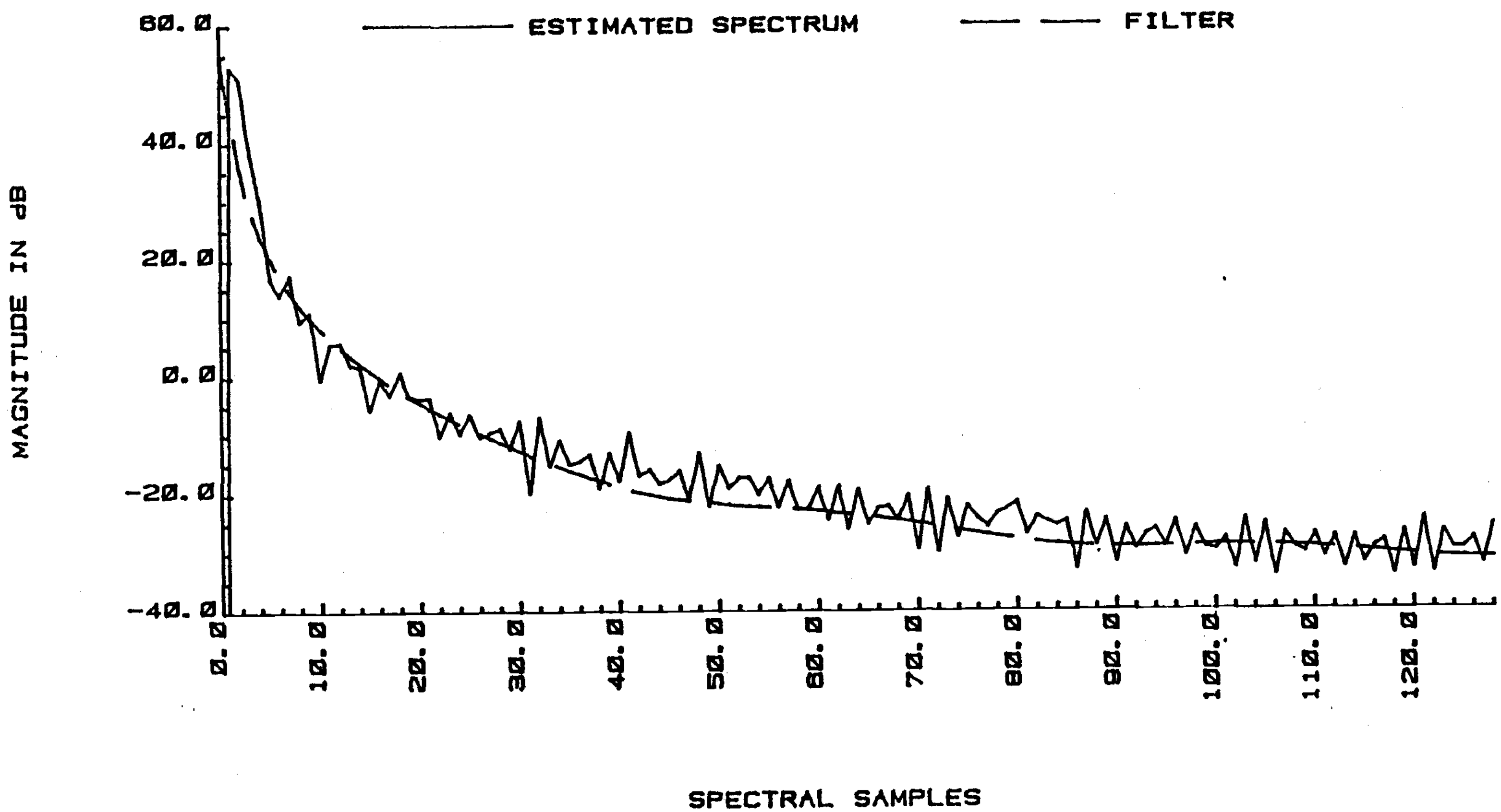
$$R3 = 0.3044 - j0.5794 = 0.6545e^{-j1.087}$$

$$R4 = -0.3244 + j0.5387 = 0.6288e^{+j1.029}$$

$$R5 = -0.3244 - j0.5387 = 0.6288e^{-j1.029}$$

which show the causal filter is minimum phase since all the zeros are inside the unit circle.

Once the filter's transfer function was found, its frequency response was calculated. This was accomplished by finding the impulse response of Equation 32, then calculating its DFT. The filter's frequency response is shown in Figure 9 with the estimate of the power spectrum of temperature.




THE WIENER FILTER'S FREQUENCY RESPONSE

FIGURE 9

Figure 9 shows that Weiner filter design corresponds to matching the frequency response of the filter to the power spectrum of the process. The closer the match, the smaller the prediction error will be. Also, Figure 9 shows the implemented filter for temperature prediction is low pass. This makes sense, since most natural phenomenon resist change. Better spectral matching around the peak could have been achieved by the use of longer periodograms to estimate the power spectrum of temperature.

## Wiener Filter Testing Results

Testing of the Wiener filter was performed to check the accuracy of the filter and to determine were error limits should be set to flag erroneous data. A major consideration in testing was to simulate the filter's performance in a real-time environment using actual test data.

To reduce the bias of the prediction error, it was necessary to remove the mean value from the data. The mean was estimated by a moving average whose length was equal to that of the filter. After the mean was removed, Equation 32 was used to predict the next value while storing the error in a file for analysis.

Error analysis that resulted from testing 1656 data points showed the following statistics:

TABLE 4

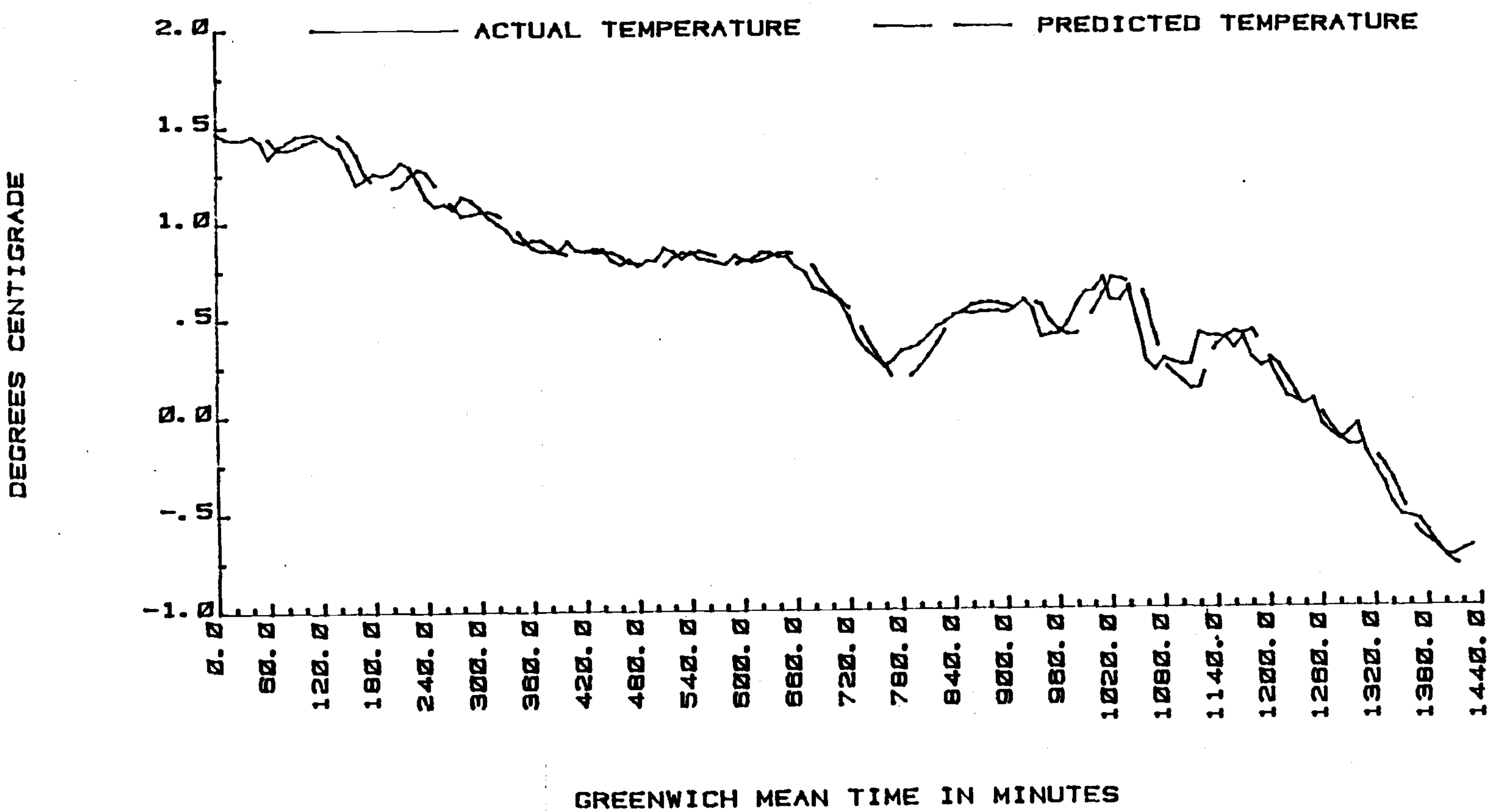
Prediction Error Statistics

| | | |
|---|---|---|
| mean | = | 0.000294672 |
| standard deviation | = | 0.26919973 |
| skewness | = | 0.069693908 |
| kurtosis | = | 10.914679 |
| percent within 1 standard deviation | = | 82.07% |
| percent within 2 standard deviations | = | 94.26% |
| percent within 3 standard deviations | = | 97.64% |

If the limits to flag erroneous data were set at three standard deviations, the filter would flag good data as correct 97.64% of the time for the days listed in Table 1 during the years 1982 and 1983.

Figure 10 shows the actual temperatures for January 25, 1983 and the predicted values. When compared to the average temperatures shown in Figure 4, the filter is seen to perform well on a day when the temperature was exhibiting abnormal behavior. As expected, the greatest errors occurred when the temperature reversed direction.

If the accuracy requirements could not be met, there are several possible explanations. First, the requirements were not realistic. An extreme example of this would be




PREDICTED TEMPERATURE VALUES FOR JAN. 25, 1983

FIGURE 10

data that was completely uncorrelated. In such cases, the best prediction would be the mean value with tolerance levels determined by the variance. Second, the estimate of the power spectrum was inadequate. Lastly, the assumptions about the noise were incorrect.

These results indicate that a Wiener filter will provide an inexpensive and efficient solution to the problem of detecting erroneous values of temperature. The solution presented could easily be adapted to a real-time environment since the computations between sample times are minimal.

# VII. CONCLUSION

The U.S. Air Force describes the tactical environment with numerous models. To access the accuracy of such models and the performance of sensors which drive the models, the Air force has developed a distributed data acquisition system which monitors environmental parameters of a test site. During data collection, there exists the possibility of collecting erroneous data. Therefore, a method to detect incoming erroneous data was found. Detection of erroneous data may be accomplished by two methods: direct and indirect. The direct method requires hardware redundancy and is simple but expensive. Indirect methods require a mathematical treatment of the data and are inexpensive and sophisticated.

An indirect method to detect erroneous values of one environmental parameter, "temperature", was investigated. This method, which uses the Wiener filter, required a software package to meet the needs of time and frequency-domain analysis, design, and testing. First, the time-domain statistics were investigated to determine the appropriateness of a Wiener filter as a solution to the problem. Next, the power spectrum of the parameter had to be estimated from which the filter was designed.

Then, the filter was designed by solving a set of equations to obtain the filter coefficients while recording the error. A plot of error versus length was used to find the optimal length. Finally, the filter had to be tested to access performance and find limits where erroneous data should be flagged.

These methods were applied to the parameter "temperature" near the winter solstice when it was most stationary. When seventeen days in Dayton, Ohio were analyzed (from December 18, 1982 to February 1, 1983), the air temperature was found to be normal (95% level of significance) from 0000 GMT to 1250 GMT and not normal (95% level of significance) from 1300 GMT to 2350 GMT. Thus, the Wiener filter provided a solution that was slightly sub-optimal. A t-test was used to show these days were not representative of the period in general, so the filter is valid only for the days analyzed.

The power spectrum was estimated by averaging seventeen periodograms that were calculated from 256-point DFT's. To reduce the variance of the estimate, three windows (the rectangle, triangle, and Blackman-Harris) were used to calculate three seperate results. When the average value was removed from the data, the smoothest estimate was produced by the rectangular window because the variance of the temperature was not constant. The power spectral estimate showed a sharp peak of 53.1dB near 1.12 cycles-per-day.

Autocorrelation coefficients, used to design the Wiener filter, were found from the power spectral estimate. An optimal filter length of six was found by solving the matrix equations for different filter lengths while recording the mean-square error. When tested, the filter would flag good data as correct 97.64% of the time with an error flag set at three standard deviations.

While this research showed that a Wiener filter will predict future values of temperature for detection of erroneous values, several topics remain to be investigated. Since the accuracy of the prediction is dependent upon the power spectrum estimate, it shoud be calculated with longer periodograms. This is dependent upon the availability of large quantities of correct data. Another question remains concerning the adequacy of the sampling rate. The analysis should be repeated with data sampled at higher rates to determine whether higher frequencies are present. Finally, other parameters should be investigated to determine if these methods may be applied to them.

# BIBLIOGRAPHY

1. C.E.P. Brooks and N. Carruthers, Handbook of Statistical Methods in Meteorology. Her Majesty's Stationary Office, London, England, 1953, pp. 13-135.

2. R.G. Brown, Introduction to Random Signal Analysis and Kalman Filtering. John Wiley and Sons, Inc., New York, N.Y., 1983, pp. 147-173.

3. K. A. Dey, "Statistical Analysis of Noisy and Incomplete Failure Data," 1982 Proc. Annual Reliab. and Maintain. Symp., Jan. 1982, pp. 246-251.

4. R.J. Evans and S.H. Price, "Fault Tolerant Microprocessor System Design," 11th Annual Intl. Symp. on Fault Tolerant Compt., June 1981, pp. 248-250.

5. R.R. Good, et al., "Automated Data Qualification," Proc. 27th Intl. Instrum. Symp., Vol. 27, April 1981, pp. 467-471.

6. F.J. Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," Proc. of the IEEE, Vol. 66, No. 1, Jan 1978, pp. 51-83.

7. J.E. Hertel and R.N. Clark, "Instrument Failure Detection in Partially Observable Systems," IEEE Trans. on Aerospace and Elect. Syst., Vol. AES-18, May 1982, pp. 310-317.

8. T. Kailath, "A View of Three Decades of Linear Filtering Theory," IEEE Trans. on Information Theory, Vol. IT-20, No. 2, March 1974, pp. 146-181.

9. N. Levinson, "The Wiener RMS (Root Mean Square) Error Criterion in Filter Design and Prediction," Jour. of Math. and Phys., Vol. 25, No. 4, Jan. 1947, pp. 261-278.

10. H. Maisel and G. Gnugnoli, Simulation of Discrete Stochastic Systems. Science Research Associates, Inc., Chicago, IL., 1972, pp. 71-75.

11. J. Makhoul, "Linear Prediction: A Tutorial Review," Proc. of the IEEE, Vol. 63, No. 4, Apr. 1975, pp. 561-580.

12. A.V. Oppenhein and R.W. Schafer, Digital Signal Processing. Prentice Hall, Inc., New York, N.Y., 1975, pp. 532-571.

13. A. Papoulis, Signal Analysis. Mc Graw-Hill, Inc., New York, N.Y., 1977, pp. 336-349.

14. L.W. Rabiner and B. Gold, Theory and Application of Digital Signal Processing. Prentice-Hall, Inc., New York, N.Y., 1975, pp. 366-419.

15. E.A. Robinson, Multichannel Time Series Analysis with Digital Computer Programs. Holden-Day, Inc., San Francisco, CA., 1967, pp. 28-104.

16. M. Schwartz and L. Shaw, Signal Processing: Discrete Spectral Analysis, Detection, and Estimation. McGraw-Hill, Inc., New York, N.Y., 1975, pp. 148-194, 274-286.

17. H.W. Sorenson, "Least-Squares Estimation: from Gauss to Kalman," IEEE Spectrum, July 1970, pp. 63-68.

18. M.D. Srinath and P.K. Rajasekaran, An Introduction to Statistical Signal Processing with Applications. John Wiley and Sons, Inc., New York, N.Y., 1979, pp. 208-224.

19. W.H. Wenstrom, Weather and the Ocean of Air. Houghton Mifflin, Co., Boston, MA., 1942, pp. 52-55.