LONG-RANGE FORECAST POSSIBILITIES FOR X-BAND RADAR CONSTRUCTION ON SHEMYA

THESIS

Brian K. Schroeder, Captain, USAF

AFIT/GM/ENP/02M-12

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
24 May 02	Final	Jun 01 - Jun 02

Title and Subtitle Long Range Forecast Possibilities for X-Band Radar Construction on Shemya	**Contract Number**
	Grant Number
	Program Element Number
Author(s) Captain Brian K. Schroeder, USAF	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Air Force Institute of Technology Graduate School of Engineering and Management (AFIT/EN) 2950 P Street WPAFB OH 45433-7765	**Performing Organization Report Number** AFIT/GM/ENP/02M-12
Sponsoring/Monitoring Agency Name(s) and Address(es) Capt Michael W. Holmes 11 Operational Weather Squadron Chief, Weather Sciences Elmendorf AFB, AK 99506	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes The original document contains color images.	

Abstract
The Missile Defense Agency (MDA) plans to construct, during a two-week period, an X-Band Radar (XBR) on Shemya, AK. Wind speeds must not, at any time during the construction, exceed a 25 knot limit set by the MDA for lifting the massive dome panels into place. The goal of this research was to explore the possibilities of long-range forecasts to determine the feasibility of predicting any upcoming two-week windows of opportunity, well in advance, that will ensure the successful completion of constructing the XBR on Shemya. In order to reach this goal, the following objectives were achieved; 1) a climatological wind study for Shemya to assess the optimal climatological window to build the XBR, 2) a detailed synoptic study over the North Pacific, to gain an understanding of how synoptic weather systems develop, move, and vary on an annual basis, 3) a traditional statistical analysis of the data followed by a Classification and Regression Tree (CART) analysis for pattern recognition of global teleconnection indices, and 4) develop forecasting decision trees to assist the 11th Operational Weather Squadron (OWS) Alaskan forecast hub in this daunting task. The Aleutian Island chain is plagued by persistent strong winds, since the Aleutian Low and expanding polar vortex affect the region in the winter, as do tropical storms and frontal passages in the summer. This, combined with Shemya being located near the exit region of the climatological storm track off the East Asian continent, makes the island one of the most challenging forecast locations in the Northern Hemisphere. This study compares surface winds and teleconnection indices, as computed by the Climate Prediction Center (CPC), to statistically analyze the data. The data were analyzed using standard statistical regression techniques, including linear and multiple linear regression methods, and then CART analysis was used for large-scale pattern recognition. The approach of the CART analysis theory used in this study was to determine which large-scale pressure patterns in the Northern Hemisphere are conducive to consistently low winds over Shemya. CART was discovered to be the best method of analysis and forecast decision trees for the 11th OWS Alaskan forecast hub were then developed to assist forecasters in providing long-range wind forecasts for the MDA, as well as the risks involved of being wrong. Since the results of the study cant offer conclusive go or no-go forecasts, an alternative wind-break proposal is included in the recommendations.

Subject Terms

Report Classification	**Classification of this page**
unclassified	unclassified
Classification of Abstract	**Limitation of Abstract**
unclassified	UU

Number of Pages
139

AFIT/GM/ENP/02M-12

LONG-RANGE FORECAST POSSIBILITIES FOR X-BAND RADAR CONSTRUCTION ON SHEMYA

THESIS

Brian K. Schroeder

Captain, USAF

AFIT/GM/ENP/02M-12

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the United States Government.

AFIT/GM/ENP/02M-12

LONG RANGE FORECAST POSSIBILITIES FOR X-BAND RADAR CONSTRUCTION ON SHEMYA

THESIS

Presented to the Faculty
Department of Engineering Physics
School of Engineering and Management
Air Force Institute of Technology
Air University
Air Education and Training Command
In Partial Fulfillment for the Requirements for the
Degree of Master of Science

Brian K. Schroeder, B.S.
Captain, USAF

May 2002

AFIT/GM/ENP/02M-12

LONG-RANGE FORECAST POSSIBILITIES FOR X-BAND RADAR CONSTRUCTION ON SHEMYA

Brian K. Schroeder, B.S.
Captain, USAF

Approved:

Ronald P. Lowther — 17 Jan 02

Lt Col Ronald P. Lowther
Advisory Committee Chairman — Date

David R. Reynolds — 17 Jan 02

Professor Daniel E. Reynolds
Advisory Committee Member — Date

Devin J. Della-Rose — 17 Jan 02

Major Devin J. Della-Rose
Advisory Committee Member — Date

Acknowledgments

I greatly appreciate the guidance and advice my faculty advisor Lt Col Ronald Lowther granted me throughout my endeavors. Professor Dan Reynolds' statistical wisdom was greatly appreciated. The assistance of Major Devin Della-Rose also deserves great praise. Thanks to Mr. Jeff Sitler for his assistance in the Meteorology lab. Thank you to all of the faculty, staff and students, in particular Capt Todd McNamara, Capt Lee Nelson, Capt Robb Randall, and 1Lt Hugh Freestrom, at the Institute for helping me achieve success. I am also grateful for the assistance provided to me by the Air Force Combat Climatology Center, specifically Mr. Dan Kerupetski and Ms. Alicia Hughes and their entire staff, and Capt Thomas Renwick and MSgt John Johnson and their entire staff. The help I received from Capt Michael Holmes and the 11th OWS should also be noted.

This work would not have been possible without the support of my entire loving family and many grateful friends. Thank you everyone for your patience and support.

Brian K. Schroeder

Table of Contents

Page

List of Figures

List of Tables

Table Page

AFIT/GM/ENP/02M-12

Abstract

The Missile Defense Agency (MDA) plans to construct, during a two-week period, an X-Band Radar (XBR) on Shemya, AK. Wind speeds must not, at any time during the construction, exceed a 25 knot limit set by the MDA for lifting the massive dome panels into place. The goal of this research was to explore the possibilities of long-range forecasts to determine the feasibility of predicting any upcoming two-week windows of opportunity, well in advance, that will ensure the successful completion of constructing the XBR on Shemya. In order to reach this goal, the following objectives were achieved; 1) a climatological wind study for Shemya to assess the optimal "climatological window" to build the XBR, 2) a detailed synoptic study over the North Pacific, to gain an understanding of how synoptic weather systems develop, move, and vary on an annual basis, 3) a traditional statistical analysis of the data followed by a Classification and Regression Tree (CART) analysis for pattern recognition of global teleconnection indices, and 4) develop forecasting decision trees to assist the 11th Operational Weather Squadron (OWS) Alaskan forecast hub in this daunting task.

The Aleutian Island chain is plagued by persistent strong winds, since the Aleutian Low and expanding polar vortex affect the region in the winter, as do tropical storms and frontal passages in the summer. This, combined with Shemya being located near the exit region of the climatological storm track off the East Asian continent, makes the island one of the most challenging forecast locations in the Northern Hemisphere.

This study compares surface winds and teleconnection indices, as computed by the Climate Prediction Center (CPC), to statistically analyze the data.

The data were analyzed using standard statistical regression techniques, including linear and multiple linear regression methods, and then CART analysis was used for large-scale pattern recognition. The approach of the CART analysis theory used in this study was to determine which large-scale pressure patterns in the Northern Hemisphere are conducive to consistently "low winds" over Shemya. CART was discovered to be the best method of analysis and forecast decision trees for the 11th OWS Alaskan forecast hub were then developed to assist forecasters in providing long-range wind forecasts for the MDA, as well as the risks involved of being wrong. Since the results of the study can't offer conclusive go or no-go forecasts, an alternative "wind-break" proposal is included in the recommendations.

LONG-RANGE FORECAST POSSIBILITIES FOR X-BAND RADAR CONSTRUCTION ON SHEMYA

I. Introduction

1.1 Background

The Missile Defense Agency (MDA) manages, directs, and executes all ballistic missile defense programs for the Department of Defense (DoD). A National Missile Defense (NMD) system has been proposed to protect the United States from the emerging threat of long-range ballistic missiles. The key NMD components include ground-based interceptors (GBI), X-band radars (XBR), upgraded early warning radars, and various other systems. The XBR, a ground-based phased array radar, will provide cued search, detection, tracking, discrimination, and kill assessment of threatening missiles and complexes. The proposed construction of XBR sites includes the mounting of the radar on its pedestal, a control and maintenance facility, power generation facility, and perimeter area controls.

In support of the MDA plan to develop a missile defense system, a system of XBR sites will be deployed. Shemya, Alaska is the most sought after location for placing the first XBR system, due to its proximity to Russia and China, as well as the overflight path trajectories of possible inbound ballistic missiles from these regions traversing the polar region. Shemya is the key location for the radar to cover all 50 U.S. states. Unfortunately, Shemya is known to have one of the harshest climates on the face of the planet, which poses a huge challenge to the construction of the XBR and aircraft operations on the island.

1.2 Problem Statement

One of the greatest challenges facing the construction of the XBR is inclement weather conditions, in particular strong persistent surface winds. Although strong winds plague the Aleutian Island chain year round, during June and July, the winds usually decrease to their lowest levels. However, tropical storms and other disturbances can drastically affect the island from June through November. In the case of the XBR construction, a wind speed limitation has been set by the MDA of less than 25 knots (kts), due to the vast size of the dome panels, which must be lifted into place (Figure 1).

Figure 1. A MDA rendering of the X-band radar, which shows the massive dome panels segmented around the entire dome structure (MDA, 2001).

Shemya's strong winds are tied mostly to strong pressure gradients surrounding the Aleutian Low. The Aleutian Low is a semi-permanent large low-pressure center, which forms off the coast of the Kamchatka Peninsula and extends eastward into the

Bering Sea and over the Aleutian Islands. It is strongest during the winter months, and moves northward and weakens during the summer. The MDA believes construction of the large XBR dome will take a period of at least two weeks to complete. During this construction time, the winds must not exceed the 25 kt limit at any time due to the possibility of damaging the dome panels as they are hoisted into place during periods of high winds. Therefore, the challenge of this research is to determine the possibility of forecasting an upcoming two-week window of opportunity within which wind speeds may not exceed this 25 kt limit threshold.

1.3 Scope of Research

In order to assist the MDA in selecting the most appropriate time to build the XBR, this research is structured to first understand the complex nature of Shemya's weather. Knowledge of the overall weather patterns and the annual movement of weather systems through the Aleutian Island chain is vital to understanding winds affecting the island. In addition, a comprehensive statistical examination of the winds is conducted in order to determine any predictability of forecasting them. After all possible methods of standard statistical analysis are exhausted, data mining techniques using classification and regression tree (CART) analyses are used to further determine if any predictability exists with current knowledge.

1.4 Research Objectives

The overall goal of this research is to determine the possibility of predicting an upcoming two-week window of opportunity that is best suited for the needs of constructing the XBR on Shemya. The model focuses on conditions that may be used to forecast the two-week window, well in advance.

To develop a forecasting technique for predicting a two-week window where wind speeds are within the MDA set limitations, a multitude of items must be examined. A flow chart is provided to help illustrate the research process (Figure 2). The following specific objectives that are necessary to achieve the overall goal of this research are to:

1. Perform an in-depth climatological study of synoptic conditions over the North Pacific. This includes an understanding of how the Aleutian Low develops, moves, and varies from year-to-year and month-to-month.

2. Examine climatological wind data for Shemya provided by the Air Force Combat Climatology Center (AFCCC) to assess the best time of year to build the XBR. Format the data in a way to show how many successive days in which wind speeds of less than 25 kts were observed, inclusive for all years of recording observations on the island (1943-1997).

3. Perform a thorough statistical examination of atmospheric teleconnection indices for that specific time of year. Some of the indices examined include the North Pacific Oscillation (NP), the Pacific-North American Oscillation (PNA), and the North Atlantic Oscillation (NAO). All teleconnection indices are obtained from the United States Climate Prediction Center (U.S. CPC), which computes them on a monthly basis and makes them available for users worldwide.

4. Perform a standard statistical study to determine any relationships between wind speeds and teleconnection indices, or a combination of teleconnection indices.

5. Use data mining CART analysis on the climatological data and teleconnection indices by data mining both sets for predictive relationships. Develop forecast decision trees to assist in choosing particular teleconnection indices that are suitable predictors (in other words, the indices with the best forecast relationship to the observed wind speeds over the island).

6. Develop a forecast decision tree to choose a two-week window of opportunity that best assists the MDA in building the XBR and the 11th OWS who will eventually produce needed forecasts. A forecast decision tree matrix that will lead forecasters from point A (location and strength of synoptic systems, sea surface temperatures, and global teleconnection indices) to point B (forecasted wind speed windows of opportunity for the island) is desired.

7. Present the results of this study as the deliverable, to the Air Force Weather Alaskan forecast hub to assist in their weather forecasting for the construction of the XBR dome, future maintenance needs, and the prospective building of other new facilities on the island, as well as overall aircraft operations over the area.

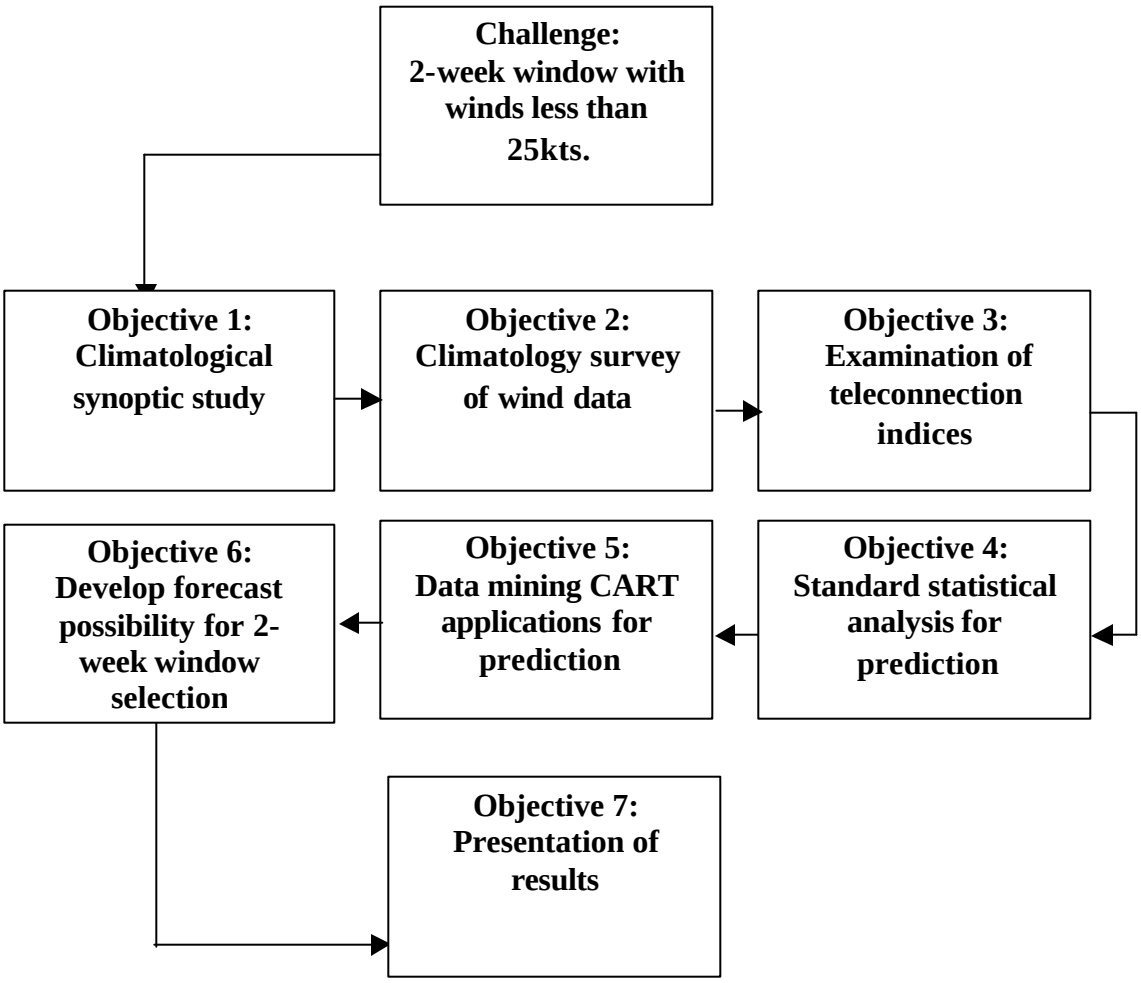


Figure 2. A flow-chart diagram of the research process.

II. Background and Literature Review

2.1 Global Patterns and Synoptic Weather Review

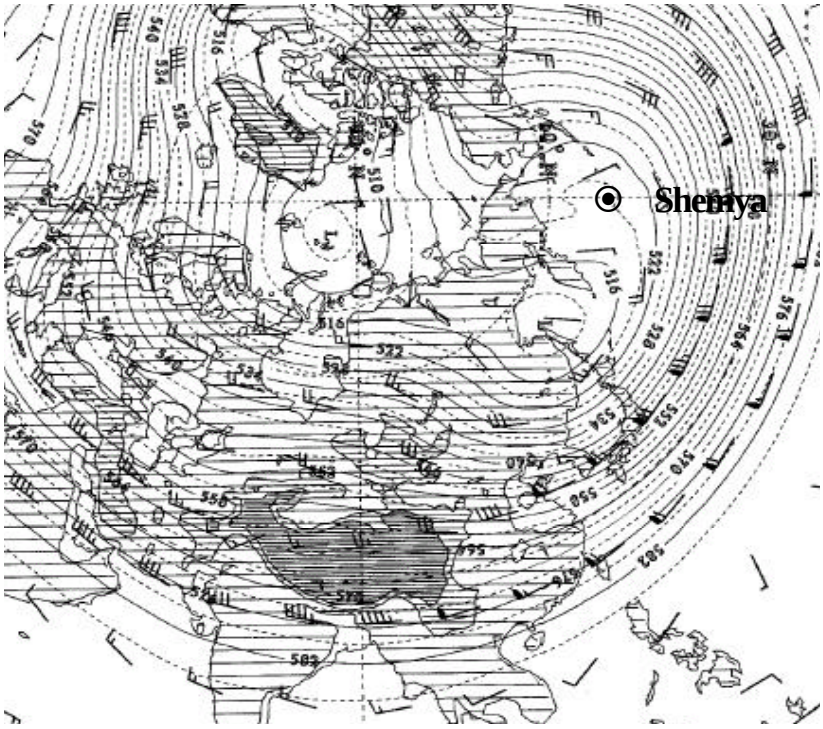
Knowledge of global weather patterns and the synoptic situation affecting the Aleutian Islands must be gained in order to understand why Shemya, Alaska is plagued year-round by strong winds. An understanding of global atmospheric circulation features and sea surface temperatures, synoptic weather conditions over the North Pacific, and variations of the Aleutian Low pressure system, are key to gaining an understanding of these phenomena.

Theorized general circulation models using the three-cell theory, along with zonal variations like the Walker Circulation give us an indication of how the global weather system functions. In the mid-latitudes, prevailing westerlies are prominent and take on a wave-like motion known as Rossby or long waves. The long-wave pattern typically seen in the Northern Hemisphere consists of high-pressure regions over land and low-pressure regions over water in the winter months and, during the summer, the opposite is apparent. The change of pressure between the regions results in pressure gradient forces.

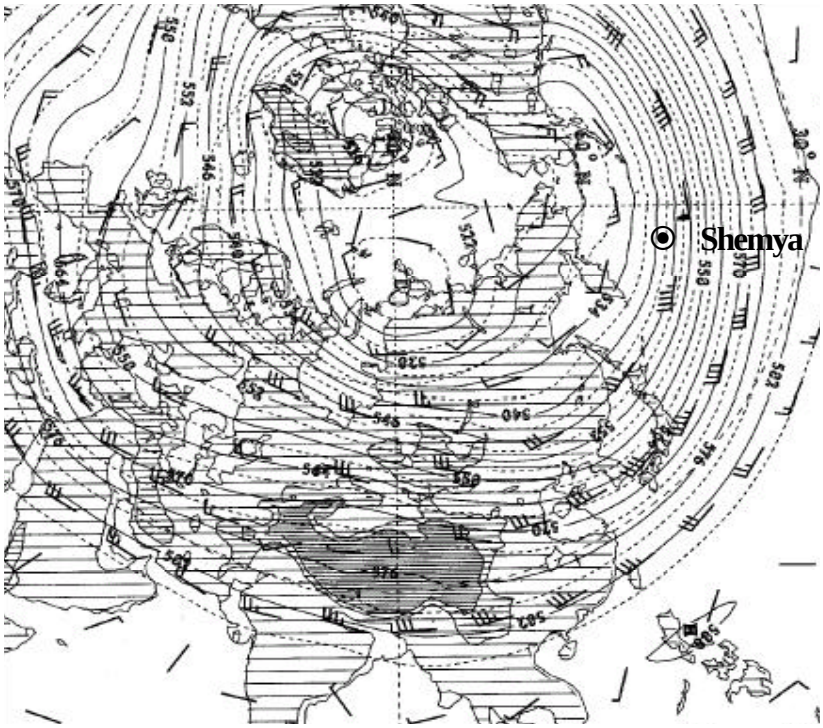
2.1.1 North Pacific Synopsis. In the North Pacific, extremely low pressure centers form on the east coast of the Asian continent near Kamchatka. They move within the prevailing westerlies following the Rossby regime flow patterns. A strong pressure gradient is created in the region between these low pressure regions and surrounding high pressure regions, which results in strong geostrophic winds between the systems. Winds from the west or southwest blow into these regions of low pressure from the equatorward

side (Blair and Fite, 1965). The mean 500 mb long-wave pattern centers a large low pressure area just north of Shemya, with a trough in Eastern Siberia and the Gulf of Alaska and a ridge predominant in the South Bering Sea, which creates a very strong pressure gradient over the Aleutian Islands. Figure 3 shows the seasonal mean 500mb heights and wind patterns, where solid and dashed lines show the heights, and winds are plotted in the traditional stem and leaf fashion. In January, the polar vortex is prominent with mainly zonal flow (Figure 3a). The polar vortex is a global scale cyclonic circulation center with two main centers of action in the Northern Hemisphere near Baffin Island and northeast of Siberia and is predominant during the winter when the north-south temperature gradient is the strongest (Glossary of Meteorology, 2000). The zonal flow leads to quickly moving weather systems passing over the island. In April, the polar vortex breaks down and a weaker temperature gradient starts to form (Figure 3b).

As summer approaches the flow starts to take on more of a meridional structure, and the pattern shifts poleward giving some relief to the strong winds experienced by the island (Figure 3c). The 500 mb isobars become less packed than during the winter months, leading to lighter winds. Slower moving storms of significant intensity tend to move across the island region during the summer, some frontal development is evident, and tropical systems (extra-tropical in nature at Shemya's latitude) dramatically affect the region on occasion. Then, as winter approaches this annual cycle of long-wave propagation repeats itself (Figure 3d).

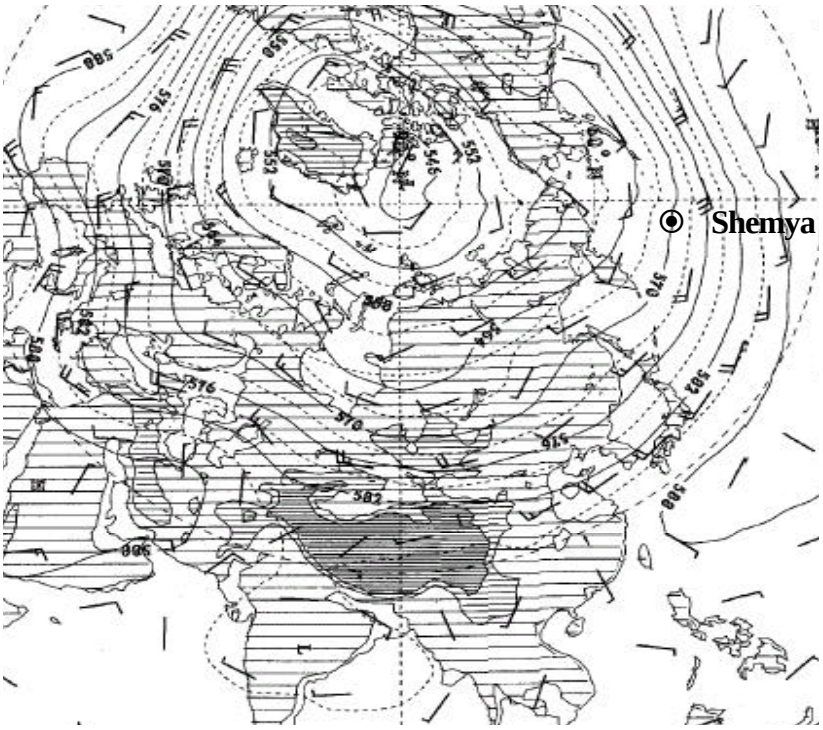


a. January

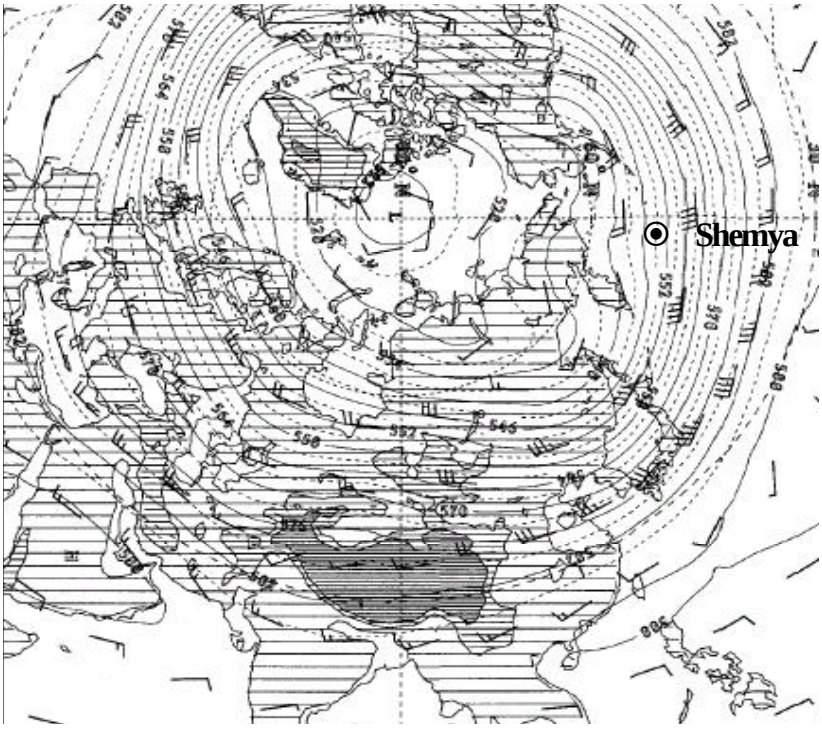


b. April

Figure 3. 500 mb mean heights (decameters) and winds (knots) for 120-year POR (modified from AFCCC/ TN-00 / 001, 2000).



c. July



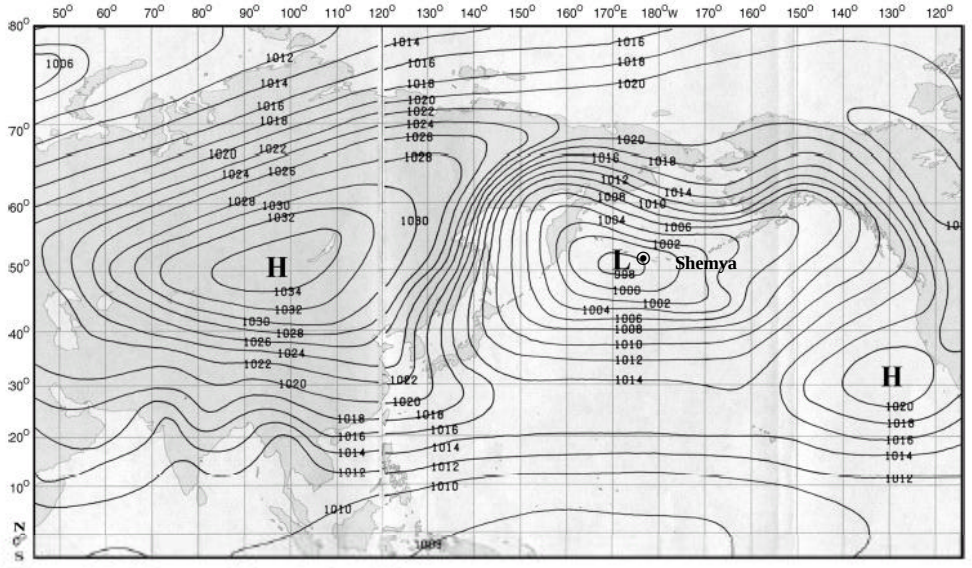
d. October

Figure 3. Continued.

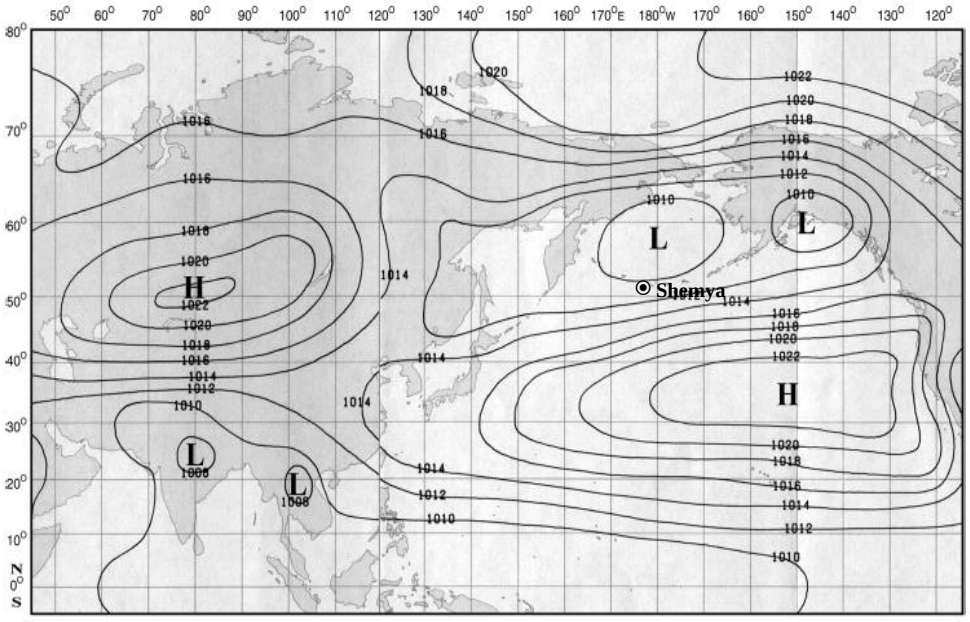
Figure 4 shows the change of mean sea-level pressure patterns over the region throughout the change of seasons. A strong northerly to westerly wind flow pattern is evident in January (Figure 4a), and the Aleutian Low and Asiatic High are discernible over the area. The winter wind regime weakens in April (Figure 4b) and disappears by May. A more southerly wind flow appears in summer and moist tropical air starts to have more of an effect on the island (Figure 4c). In July, a low pressure system known as the Asiatic Low is evident over southwest Asia and a high pressure system known as the North Pacific High is located just south of the Gulf of Alaska. In autumn, the summer pattern breaks down and a return to a more winter like pattern develops (Figure 4d).

During the summer months, frontal system passage and tropical storm activity frequent the area on many occasions. Shemya's location does not exclude it from tropical cyclones. During the summer and early fall period (June through October) the island experiences southwesterly flow (Figure 4c) that dramatically reduces ceilings and visibility. Also, tropical cyclones form in the East Pacific and track westward into the West Pacific during these months. The storms reach their most westward limit near Japan and then curve northward and eventually northeastward as they become extra-tropical. The strongest storms typically occur during the latter portion of the period from August through October. Early in the tropical season (June through August), storms are less prevalent in the North Pacific, but the average location of the tracks move further north. Tropical storm force winds (30-64 kts) are often experienced on Shemya as these systems move eastward through the area. Later in the season (September through October) storms become more numerous and are stronger in intensity. During these months, colder sea surface temperatures off the waters of the Aleutian chain weaken the

storms as they move northward, at which point they become extra-tropical and move eastward affecting the Aleutian island chain. The climatological paths of known tropical cyclones are represented in Figure 5a-f.

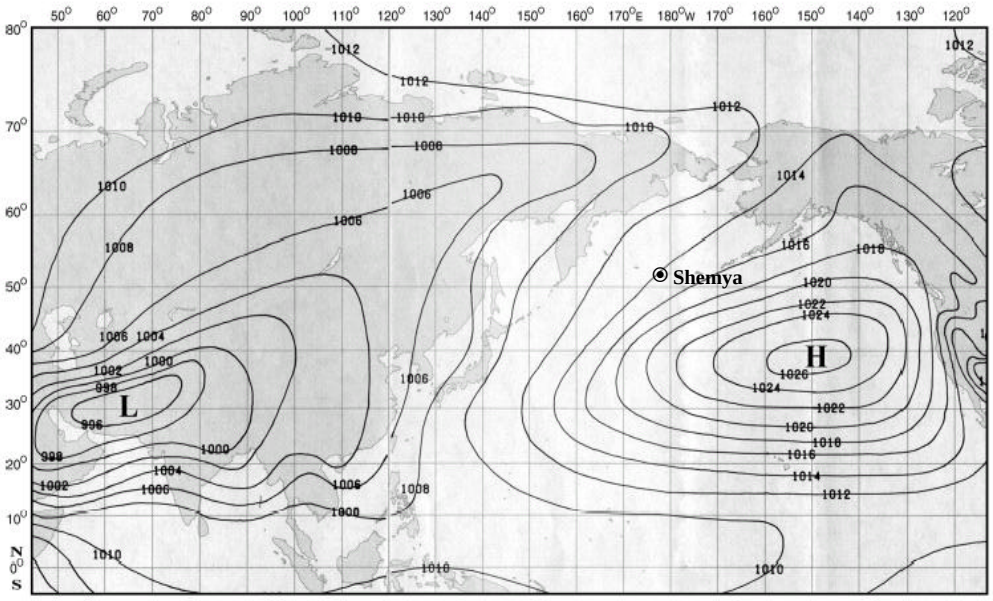


a. January

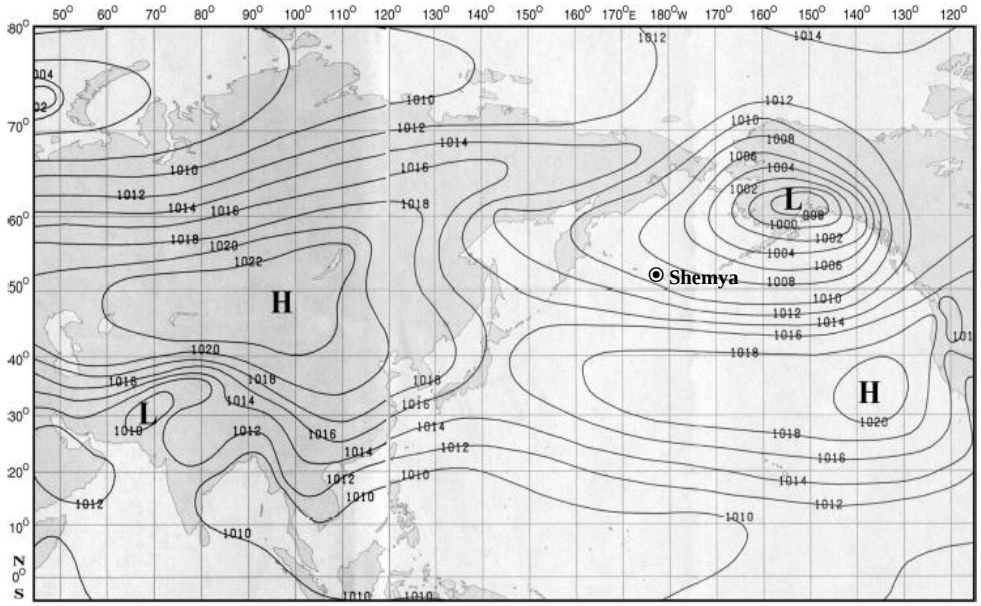


b. April

Figure 4. Asia and North Pacific mean sea-level pressure, in millibars for 120-year POR (modified from AFCCC / TN-00 / 001, 2000).

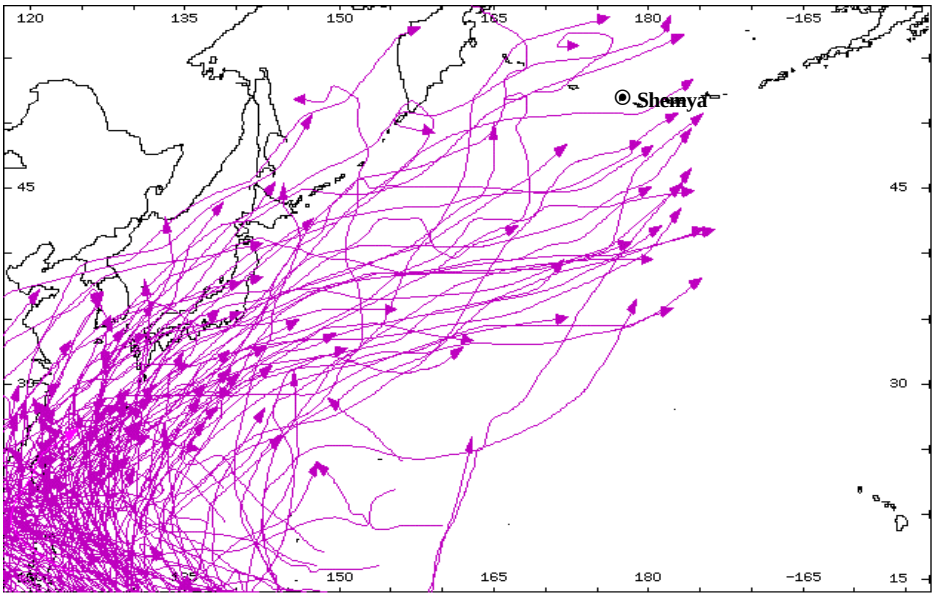


c. July

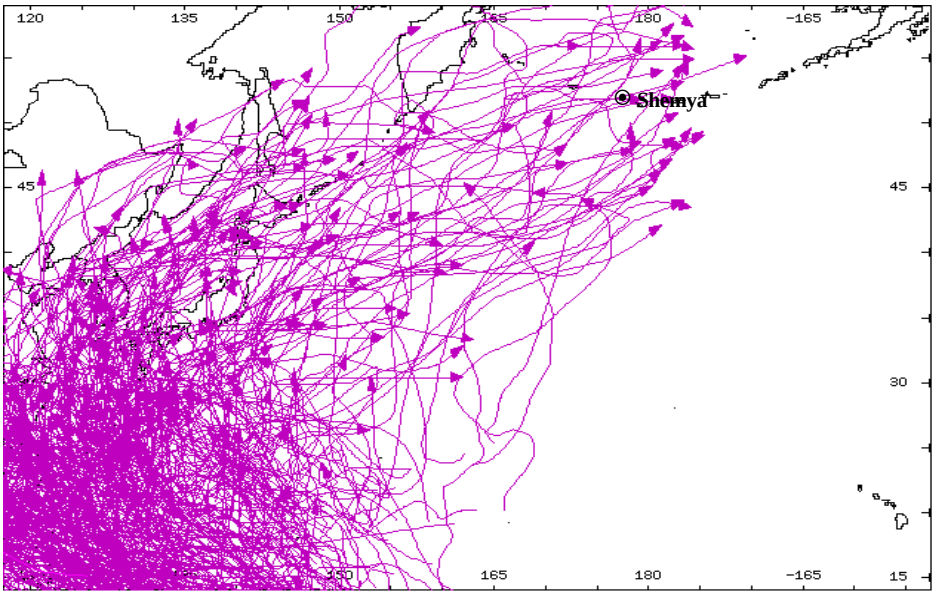


d. October

Figure 4. Continued.

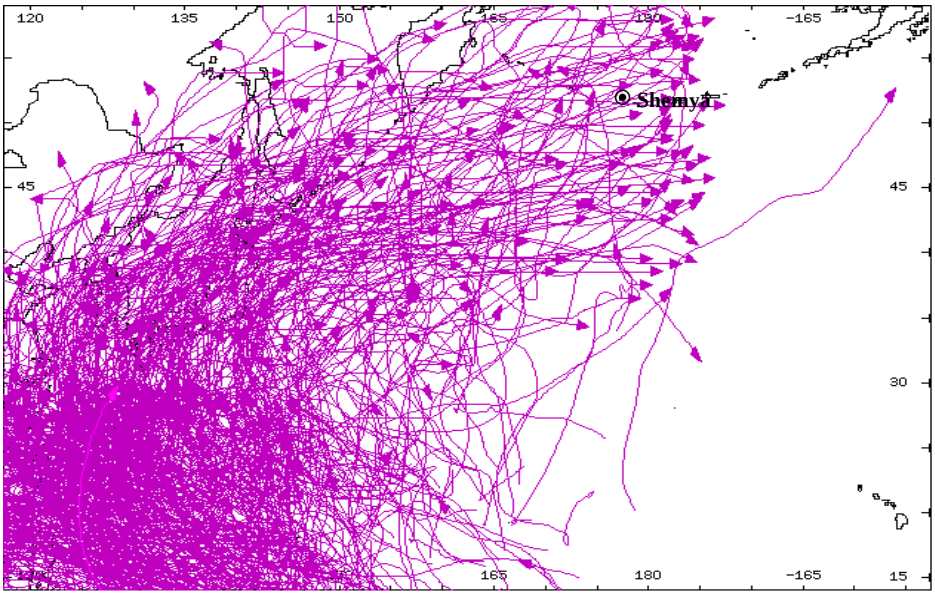


a. June

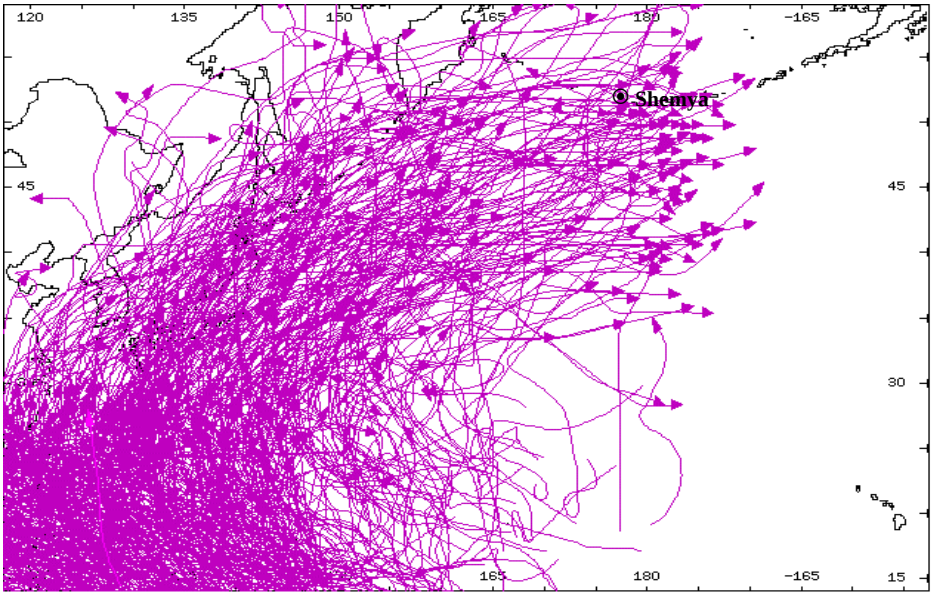


b. July

Figure 5. Path of tropical storms in the North Pacific
for the entire 160-year POR (modified from AFCCC, 1996).

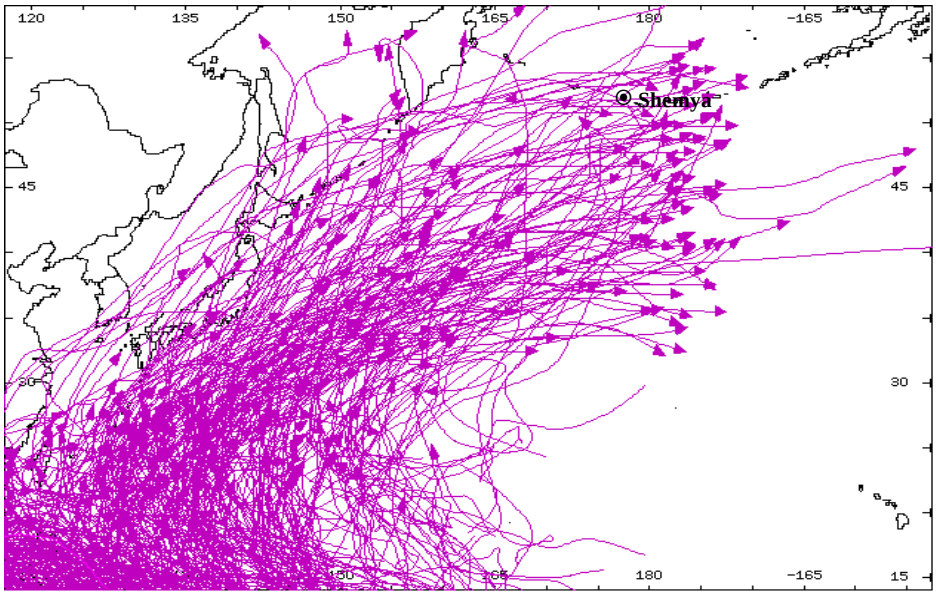


c. August

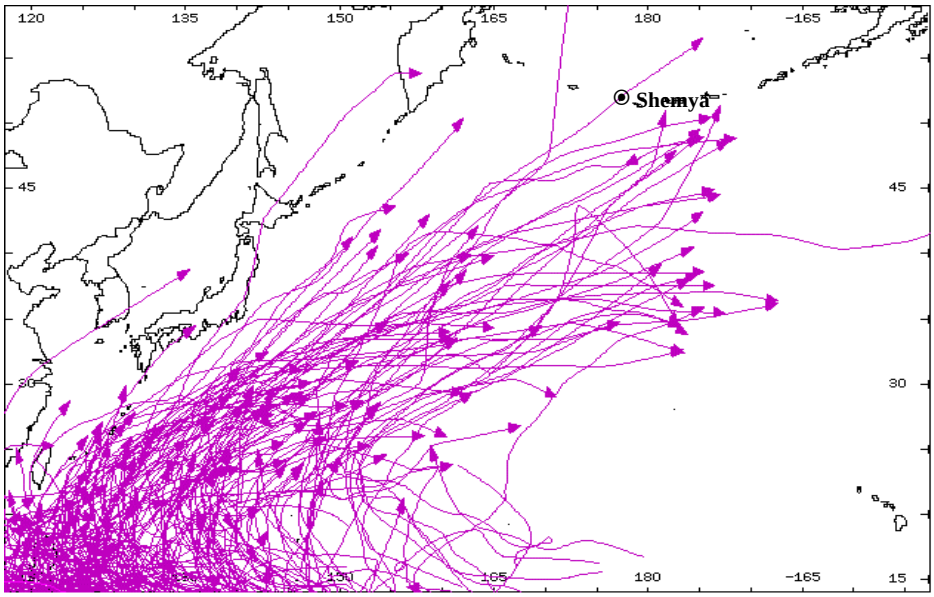


d. September

Figure 5. Continued.



e. October



f. November

Figure 5. Continued.

2.1.2 Influential North Pacific Pressure Centers. An understanding of the main influencing pressure centers in the North Pacific is beneficial for any study of the region's weather systems. The following information was obtained mainly from climatic studies performed for the region by AFCCC (AFCCC, 2000). A brief description of the Aleutian Low, Asiatic Low, North Pacific High, Asiatic High, and Mobile Polar High pressure systems follow.

Aleutian Low: The main influencing synoptic feature in the North Pacific is the Aleutian Low, which has the greatest affect on the Aleutian Islands during the winter season. The Aleutian Low is one of the main pressure centers of action in the Northern Hemisphere's atmospheric circulation. Along with the Icelandic Low, it dominates the long-wave pattern for most of the Northern Hemisphere's winter. The depth or intensity of the low-pressure area over the Bering Sea exceeds the intensity of the high-pressure areas located predominantly over central Asia . Smaller Aleutian lows form over open water in close proximity to colder snow or ice-covered land and ocean regions. The strength of these lows is dependent upon the reduction of surface friction that is evident over water, and the sensible heat sources that the open water provides. The Aleutian Low is a semi-permanent feature seen by the study of Rossby waves, as discussed earlier. The low propagates north to south seasonally and is weakest and furthest north in late spring through the early fall seasons (Hess, 1979).

Asiatic Low: Another pressure center prominent in the region is the Asiatic Low. The low joins together the eastern end of a wide climatological low-level thermal trough located through southern Pakistan and extending into northwest India. It plays a significant role in the summer mean cyclonic circulation and is prevalent in July due to

strong heating over the desert areas of Southwest Asia (Figure 4b and 4c). During the winter months these desert areas cool and high pressure dominates the region.

North Pacific High: The North Pacific High has its greatest influence in the region during the summer with its center located over the Northeast Pacific. Its circulation, when combined with the Asiatic Low, advects tropical air over Shemya bringing southwesterly flow to the island inducing low ceilings and reduced visibilities (Figure 4c and 4d).

Asiatic High: This strong yet shallow high-pressure center dominates the Asian continent from September to April (Figure 4d). It is the strongest cold type anticyclone found in the Northern Hemisphere and the central pressure is strongest in January when it is centered near Lake Baykal in Russia. It is created and supported mostly by large-scale radiational cooling over the area it encompasses.

Mobile Polar Highs: Mobile polar highs (MPHs) are transient centers of circulation that form in the Arctic and Antarctic areas. Figure 6 shows the trajectories of these MPHs. One of the main trajectories (highlighted in the figure) is directed at the Aleutian chain from the polar region. The mean number of MPHs for the entire year is one every 3.2 days with June averaging 7.8 days and 7.6 days in July (Leroux, 1998). In the wake of a MPH a void is left, or "short-lived low," that attracts accelerating masses of warm air towards the pole. As this advected air cools, low-level high pressure areas are regenerated and new MPHs are born (Leroux, 1998). These short-lived mesoscale lows are also associated with waves moving around the Aleutian and Icelandic Lows.

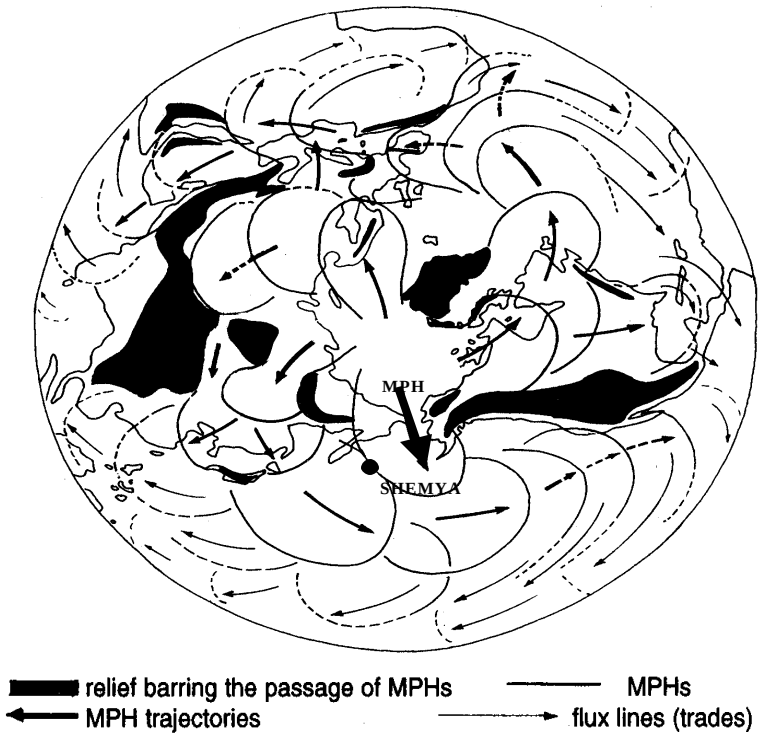


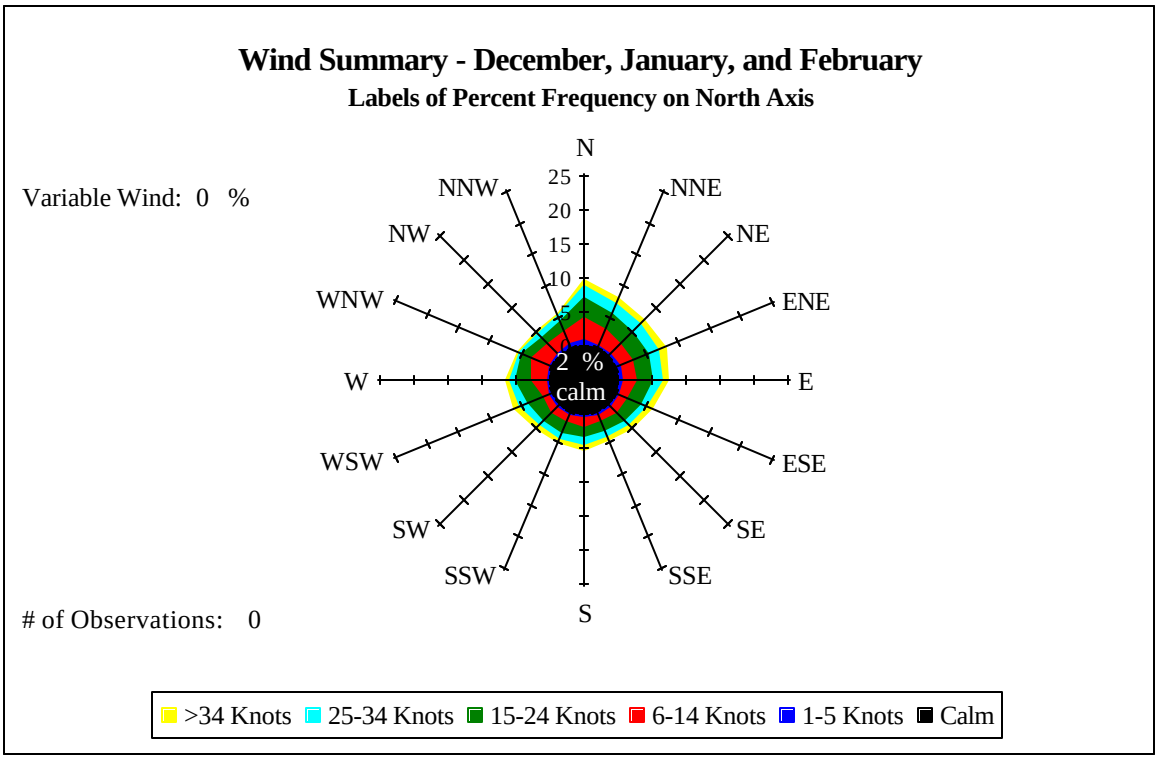
Figure 6. Trajectories of mobile polar highs in the Northern Hemisphere (modified from Leroux, 1998).

2.1.3 Shemya Winds. The island of Shemya is plagued by strong winds throughout the year. The MDA has set a minimum two-week window of low winds during which the XBR dome must be constructed. To grasp the challenges the MDA will face, it is important to consider the climatological distribution of the winds on the island. One of the best methods used to visualize winds climatologically are seasonal wind roses (Figure 7). Wind roses show the direction from which the winds blow, along with the percent frequency of occurrence. The winds on the island are directly tied to the flow associated with the synoptic weather systems and global pressure patterns previously discussed.

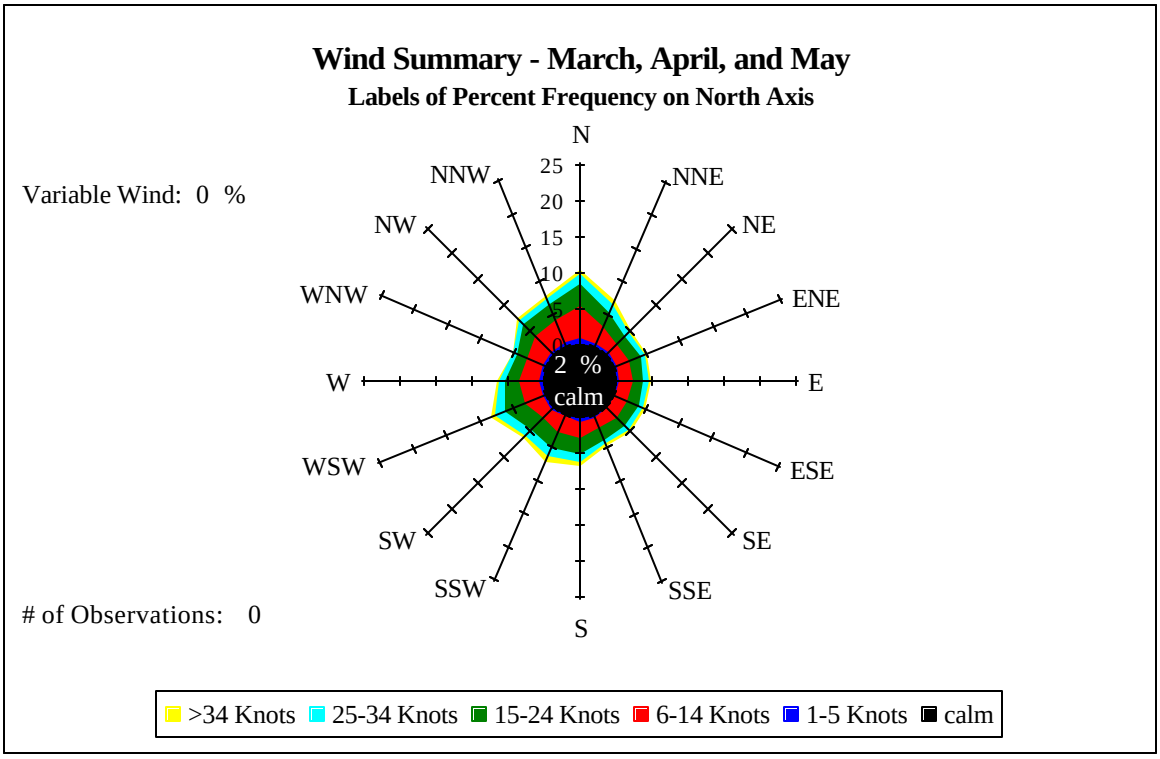
During the winter season, the placement and strength of the Aleutian Low combined with the Asiatic High, leads to very tight pressure gradient, and a flow from the north and northeast (compare Figure 4a and 7a). In springtime, the Aleutian Low moves northward and the North Pacific High expands westward across the Pacific Ocean easing these pressure gradients and producing more northerly winds (compare Figure 4b and 7b). As summer approaches, the North Pacific High expands northward and strengthens becoming the most prominent weather feature over the region and the flow shifts from a northerly to more south/southwesterly direction with significantly lighter wind speeds (compare Figure 4c and 7c). During the fall months, the North Pacific High retreats southward and the Aleutian Low and Asiatic High begin to strengthen which causes the tightening of the pressure gradient and a prevailing westerly flow over the Aleutian chain. (compare Figure 4d and 7d).

2.2 Teleconnection Patterns

Extended-range forecasting, along with macrometeorology (the study of larger-scale atmospheric circulations), has always and will continue to be a great challenge for meteorologists. Data alone, spatially and temporally, are not enough to give reliable extended-range forecasts (Namias, 1978), and current numerical weather prediction computer models are only marginally reliable for a periods extending out to 7 days (Ray, 1986). The invention of teleconnection indices (TIs) shows promising results toward our ability to produce extended-range forecasts by linking large-scale circulation pressure patterns. An explanation of how TIs are derived is included in the following chapter.

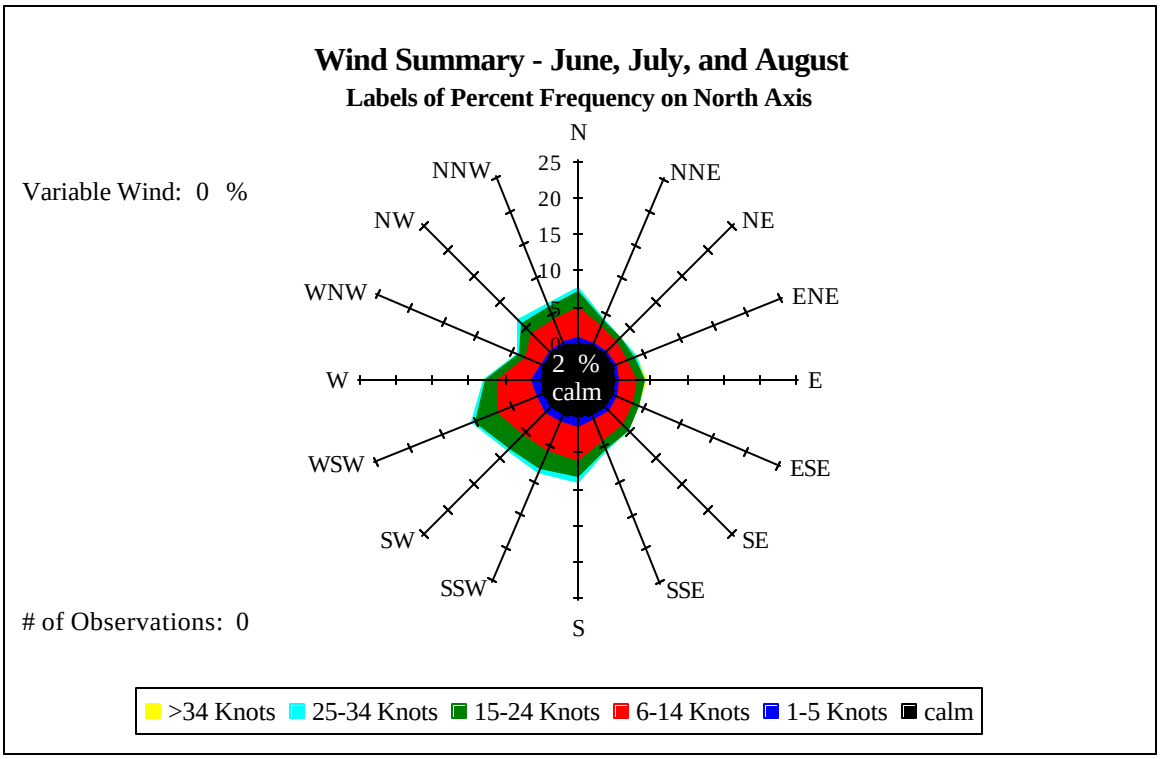


a. Winter Season

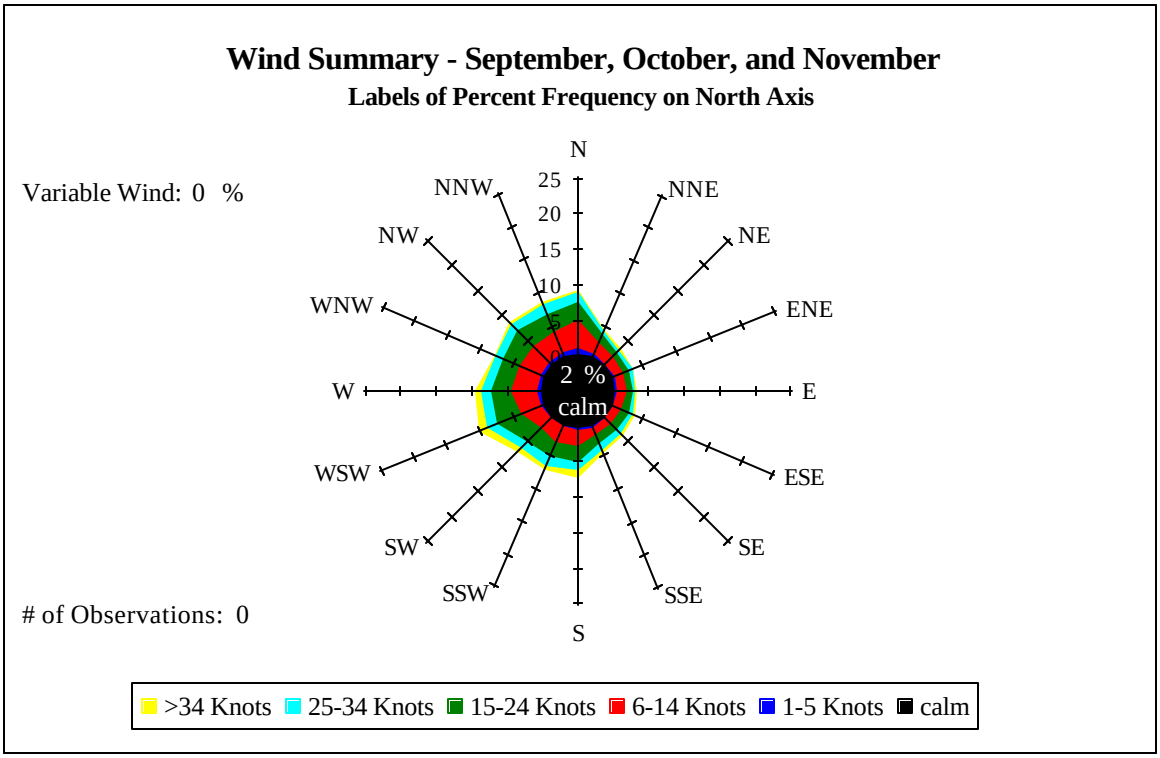


b. Spring Season

Figure 7. Wind roses for Shemya (AFCCC, 2001).



c. Summer Season



d. Fall Season

Figure 7. Continued.

III. Data Collection and Review

3.1 Data Collection

Three main sets of data were used in this study. AFCCC provided wind data, in particular maximum wind speeds, and teleconnection indices and sea surface temperature (SST) data were obtained from the U.S. CPC.

3.1.1 Wind Data. The period of record of wind data available for Shemya is from 1943 to 1997, and includes a list of year, month, day, and maximum surface winds (sustained and maximum gusts). The instruments used to record the observations were various types of simple cup anemometers. The type of transmitting and wind recording instrumentation, location of the anemometer, and its height above ground have changed significantly over the years as a result of new construction and observer requirements on the island. Appendix B contains information regarding procedures for taking wind observations. The average height of the anemometer is 27 feet above the ground; this added to the field elevation of 97 feet yields an average height of 124 feet above sea level. The location, transmitter type, type of recorder, and sensor height above ground are shown in Table 1. Due to its location in the open, and the relative sparseness of vegetation on the island, sensors are believed to give quite accurate wind readings, which is apparent in the investigation of wind speed records, which exhibited valid uniformity. Missing years of data were eliminated due to either an entire year or partial year of missing observations. Missing data can be attributed to movement of the sensor, transition from observer personnel responsibilities from the Air Force to the Navy, or changing sensor or recording devices.

Table 1. Surface wind equipment information (AFCCC, 2001).

Location Number	Date of Change	Location	Type of Transmitter	Type of Recorder	Height Above Ground
1	July 1943	Not Available	Anemometer	Ml-144B	N/A
2	June 1948	Atop Weather Stn	Same	Same	28 feet
3	March 1953	Atop Ops Building	AN/GMQ-1	ML-204	Same
4	March 1954	Same	Same	Same	25 feet
5	July 1955	Atop Weather Stn	Same	Same	25 Feet
6	September 1955	Same	Same	Same	40 Feet
7	July 1958	Atop Balloon Blg	UMQ-5C	None	37 Feet
8	September 1959	Atop Weather Stn	F420C	AW Gusts	32 Feet
9	November 1961	Same	AN/GMQ-11	RO-2	20 Feet
10	March 1968	End of Runway 28	Same	Same	20 Feet
10	Same	End of Runway 10	Same	Same	17 Feet
12	January 1988	Same	GMQ-20	Same	Same

3.1.2 Teleconnection Index Data. The U.S. CPC computes and tabulates monthly teleconnection index values, and the indices utilized in this study were computed from 1950 to 2001. Table 2 shows the calendar months when the various teleconnection indices were deemed statistically significant and computed, showing nine cold season patterns, three transition season patterns, and three warm season patterns. The NAO is included in all seasonal categories (Barnston and Livezey, 1986). This study used those

patterns (indices) that match the months when the maximum winds were observed to be below the 25 kt threshold.

The U.S. CPC first determines which patterns are most prominent and then calculates the amplitudes of each pattern. RPCA is then used to identify which patterns are the most prominent. According to U.S. CPC, this method isolates the primary teleconnection patterns for all months and allows for time series of the amplitudes of the patterns to be constructed. The RPCA method is superior to the grid-point analysis, typically determined from one-point correlation maps (comparing specific locations). The Northern Hemispheric teleconnection patterns are identified based on the entire flow field, and not just from height anomalies at a few select locations (U.S. CPC, 2001). The Southern Oscillation Index (SOI), in the Southern Hemisphere, is computed seperately using the one-point correlation method by comparing central pressure readings at Tahiti and at Darwin, Australia.

Table 2. Calendar months when specific Northern Hemispheric teleconnection patterns are significant. Tabulated values indicate the mode number of the pattern for that calendar month (i.e., a 1 indicates that the pattern appears as the leading rotated mode during the month). No value is plotted when a pattern does not appear as a leading rotated mode in a given calendar month (CPC, 2001).

PATTERN	DEC	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV
NAO	2	2	3	1	1	2	3	2	2	5	1	1
EA	6	6	7	6	10	---	---	---	---	8	7	5
EA-JET	---	---	---	---	6	9	7	3	7	---	---	---
WP	4	3	4	3	4	4	6	7	8	10	4	6
EP	9	10	9	10	8	3	1	1	---	---	6	9
NP	---	---	---	2	2	1	2	6	---	---	---	---
PNA	3	1	2	5	5	10	---	---	6	6	5	2
EATL/ WRUS	7	8	10	7	9	7	---	---	---	7	3	4
SCAND	5	9	8	8	3	5	---	---	10	1	2	3
POLAR-EURASIA	1	4	1	---	---	---	---	---	---	---	---	---
TNH	8	7	---	---	---	---	---	---	---	---	---	8
PT	---	---	---	---	---	8	4	4	4	---	---	---
ASIAN SUMMER	---	---	---	---	---	---	5	5	5	---	---	---

3.2 Rotated Principle Component Analysis

An explanation of rotated principle component analysis (RPCA), also known as empirical orthogonal functions (EOF), follows and is based mainly on thoughts taken from Wilks (1995). The U.S. CPC uses RPCA methods to determine the leading nodes, actual teleconnection patterns from 700 mb height analyses, and to determine each pattern's amplitude.

3.2.1 RPCA Synopsis. Principle Component Analysis (PCA) is the most widely used multivariate statistical technique in the atmospheric sciences, and involves the process of reducing a large data set into a new data set containing fewer variables. It is conducted on centered data or anomalies, which are defined as a departure from the mean of an observation. After departures from normal (anomalies) are calculated, data sets of N observations are compared to N values of each principle component. In the case of teleconnection indices, the 700 mb gridded mean height anomalies are used as the principle components, where the mean was taken from a 30-year period of data from 1964 to 1994. Figure 8 shows the grid points as devised by Barnston and Livezey (1986) and as used by the U.S. CPC in their RPCA analysis.

Gridded data are then transformed into matrices of values of covariance. From a covariance or correlation matrix, eigenvectors are computed and then used to form a new correlation coefficient matrix. The elements of the new vector are called principle components. RPCA results in a more compact representation of the data's variations and has the potential to substantially produce better comprehension of spatial and temporal variations exhibited by the field being inspected. The data are viewed from assignment of eigenvectors to define a new coordinate system.

When physical interpretation is the goal of PCA (as is the case of this study), rotation is worthwhile in obtaining a set of new coordinate vectors. Rotating these eigenvectors creates a second set of variables called rotated principle components. There are many rotation methods, and it not clear which one the U.S. CPC uses. However, all seek to produce what is known as a simple structure, to understand the data.

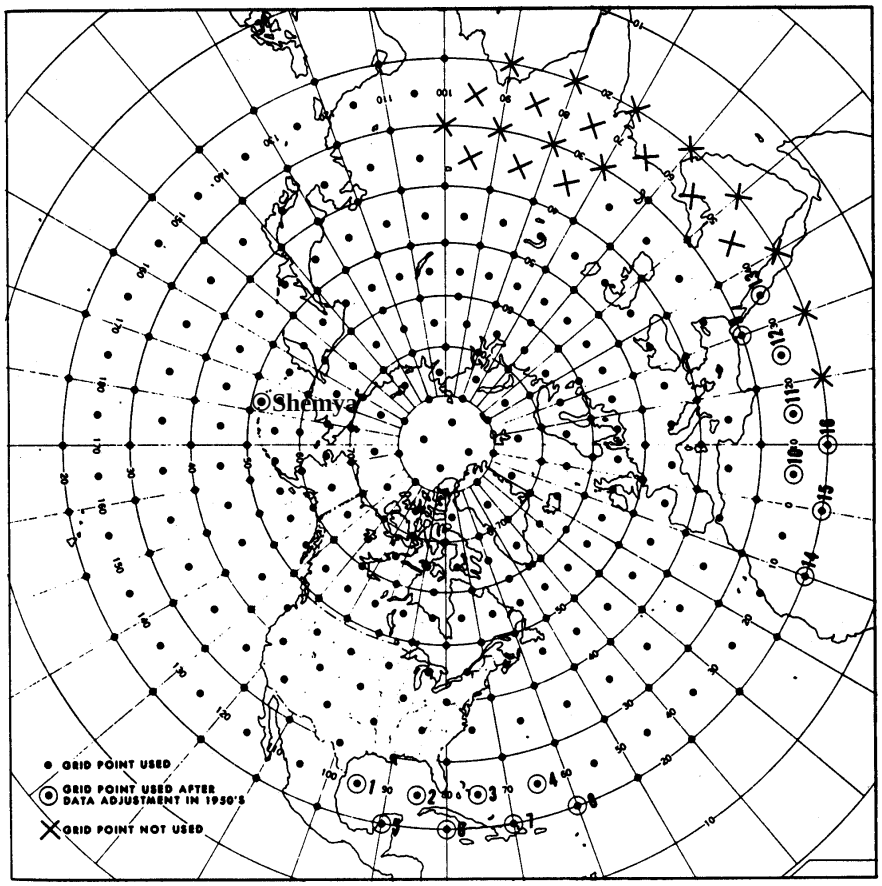


Figure 8. The 358-point Northern Hemispheric grid used to ingest time-averaged 700 mb data into rotated principle component analyses (modified from Barnston and Livezey, 1986).

3.2.2 TI Computation. Once a leading mode has been determined using the RPCA method, the individual TI values are computed. To accomplish this, CPC uses least-squares regression analysis, a type of regression analysis that fits a regression line

with the closest linear fit on the surrounding observed amplitudes. Patterns that do not appear as leading modes for a month are not computed.

3.2.3 TI Overview. There are over a dozen different teleconnection indices that are computed for the Northern Hemisphere. A brief explanation of each index is presented and is based mostly on definitions by Bell (1998), Washington, et. al. (1998) and the CPC (2001). A global depiction of the indices is provided in Figures 9-14 showing both positive and negative anomalies.

The North Atlantic Oscillation (NAO) is computed for all 12 months. The NAO consists of a north-south dipolar anomaly, with centers over Greenland/Iceland and the Azores. Its positive phase is defined as near-normal pressure near the Icelandic low and above normal pressure over the Azores. Figure 9 shows a more positive phase in January than July, which is indicated by higher values (the more positive values given in the figure's legend).

The Pacific Transition (PT) pattern is most pronounced during the Northern Hemisphere summer. The mode consists of a wave-like pattern of anomalous heights extending from the Gulf of Alaska eastward to the Labrador Sea. The main centers of action have similar signs, and are located over the inter-mountain regions of the United States and over the Labrador Sea. Weaker anomaly centers with opposing signs are located over the Gulf of Alaska and the eastern United States.

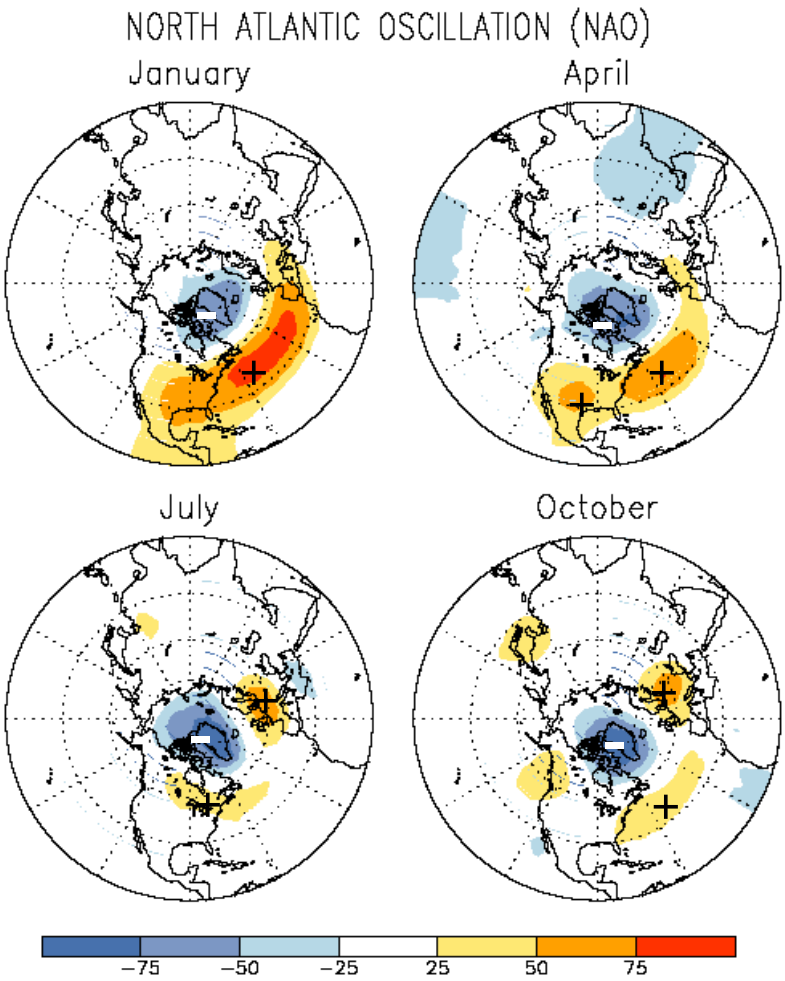


Figure 9. North Atlantic Oscillation (modified from CPC, 2001).

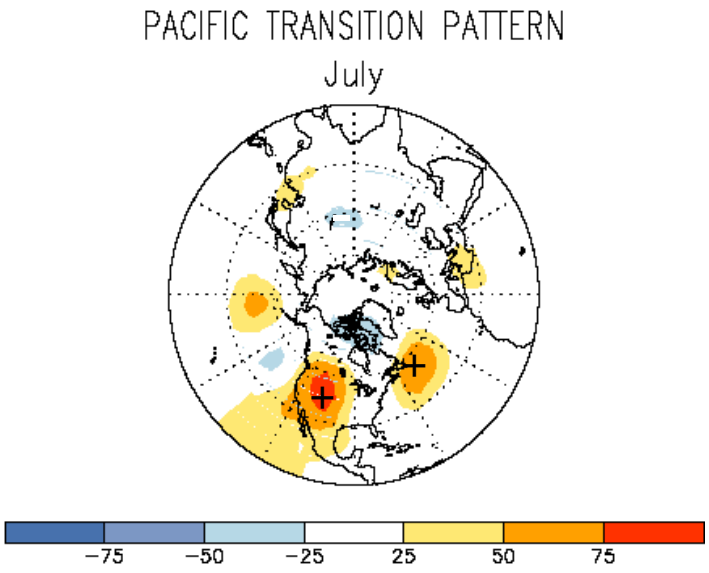


Figure 10. Pacific Transition Pattern (modified from CPC, 2001).

The West Pacific (WP) pattern is a primary mode of low-frequency variability over the North Pacific in all months (Barnston and Livezey, 1987; Wallace and Gutzler, 1981). During winter and spring, the pattern consists of a north-south dipole of anomalies, with one center located over the Kamchatka Peninsula and another broad center of opposite sign covering portions of southeastern Asia and the low latitudes of the extreme western North Pacific. Therefore, strong positive or negative phases of this pattern reflect pronounced zonal and meridional variations in the location and intensity of

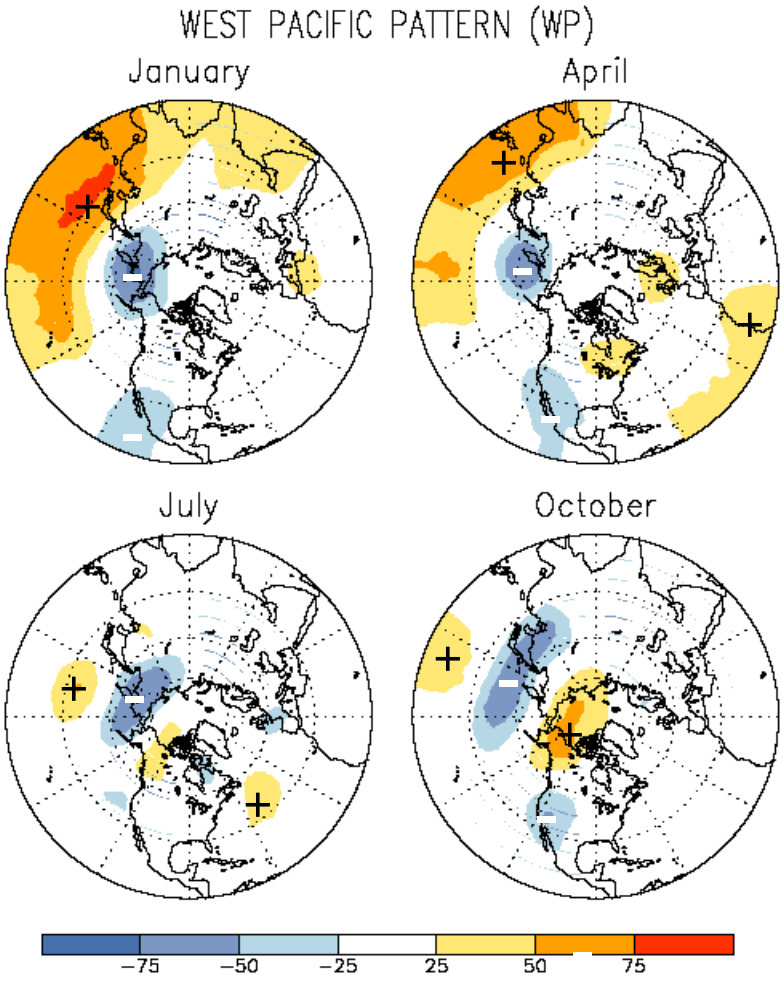


Figure 11. West Pacific Pattern (modified from CPC, 2001).

the entrance region of the Pacific (or East Asian) jet stream. During summer and autumn, a third center appears over Alaska and the Beaufort Sea, with a sign opposite to the center located over the western North Pacific.

The Asian Summer (AS) Pattern is apparent from June to August. The pattern is monopole in nature, with anomalies of the same sign observed in southern Asia and northeastern Africa. A positive phase of the pattern is indicated by above-normal heights throughout southern Asia and northeastern Africa. The AS exhibits considerable interannual and interdecadel variability, with a pattern of one sign observed for several consecutive years.

ASIA SUMMER PATTERN (AS)

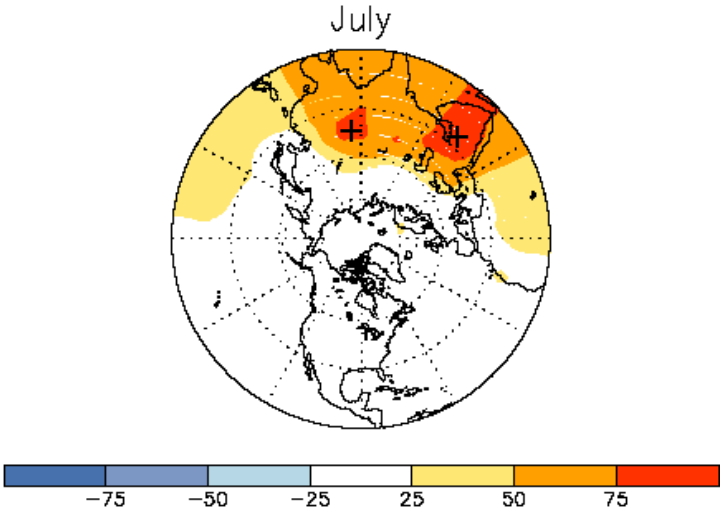


Figure 12. Asia Summer (modified from CPC, 2001).

The North Pacific (NP) pattern is most evident from March through July. This pattern consists of a primary anomaly center, which spans the central latitudes of the western and central North Pacific, and a weaker anomaly region of opposite sign that spans eastern Siberia and the intermountain region of North America. Overall,

pronounced positive phases of the NP are associated with a southward shift and intensification of the polar jet over the North Pacific from eastern Asia to the eastern North Pacific, followed downstream by an enhanced cyclonic circulation over the southeastern United States. Pronounced negative phases of the NP are associated with circulation anomalies of opposite sign in these same regions.

NORTH PACIFIC PATTERN (NP)

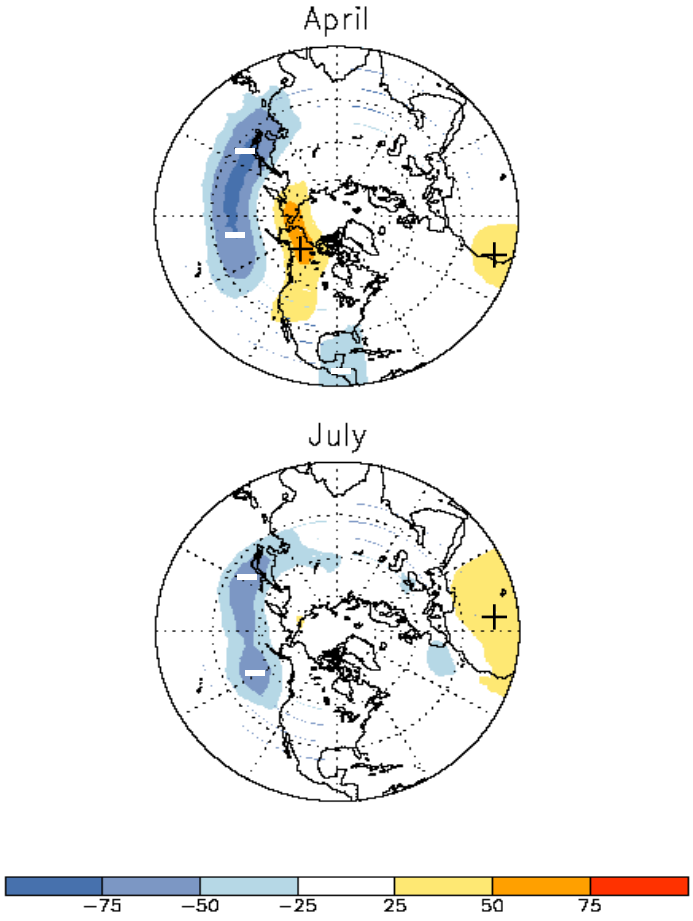


Figure 13. North Pacific Pattern (modified from CPC, 2001).

The Pacific/North American (PNA) pattern is one of the most prominent modes of low-frequency variability in the Northern Hemisphere mid-latitudes, appearing in all months except June and July. The PNA reflects a quadripole pattern of geopotential

anomalies, with anomalies of similar sign located south of the Aleutian Islands and over the southeastern United States. Anomalies with opposite sign to the Aleutian center are located near Hawaii, and over central Canada during the winter and autumn seasons (Barnston and Livezey, 1987; Bell, 1998).

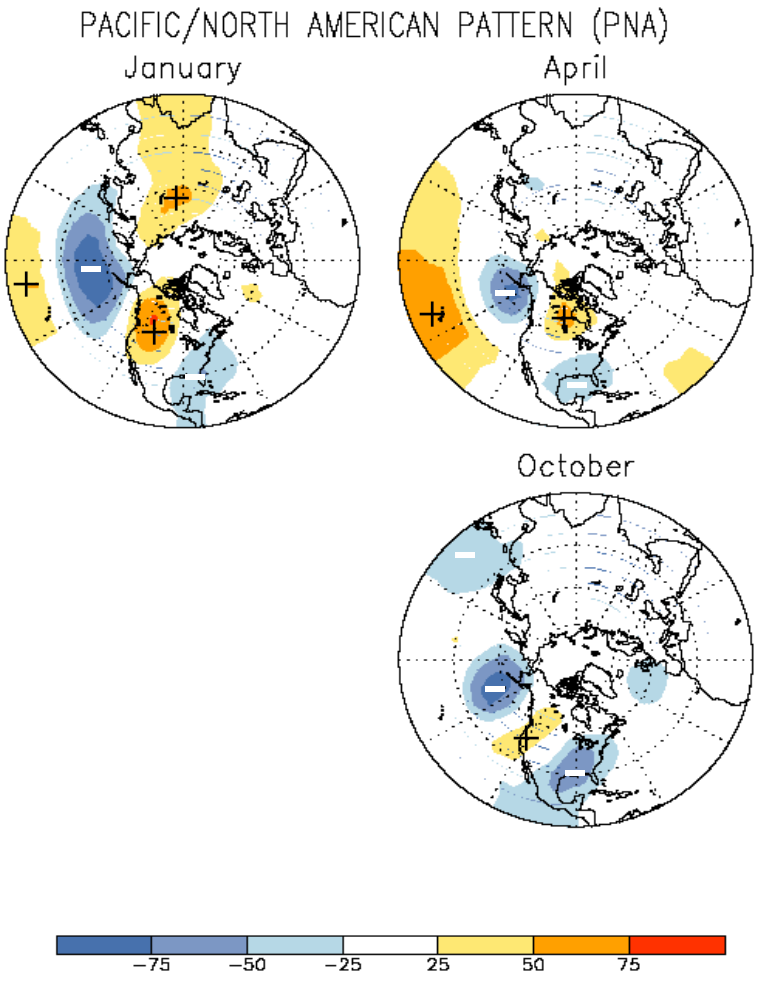


Figure 14. Pacific / North American Pattern (modified from CPC, 2001).

The PNA pattern is at its largest extent during winter, corresponding to the expanse of the Aleutian component of the mode across much of the North Pacific. The pattern undergoes major seasonal changes such that in spring, the subtropical center

covers a very large part of the Pacific but the Aleutian center becomes confined to the Gulf of Alaska. This pattern is not identifiable during June and July and is confined mainly to the landmasses of the North Pacific during the Northern Hemisphere spring. This modulation appears to link the PNA directly with Rossby waves associated with the seasonal modulation of the strength of the zonal westerlies (Barnston and Livezey, 1987; Washington et. al., 1999).

The Southern Oscillation Index (SOI), in the Southern Hemisphere, is calculated using the sea level pressures over Tahiti and Darwin, Australia. The index is modeled after the Walker Circulation.

The East Atlantic/Western Russia (EA/WR) pattern affects the Eurasian continent most of the year. It is prominent during all months, except during the June to August period. Two main centers of anomalies are located over the Caspian Sea and Western Europe during the winter. Throughout the other seasons, the anomaly centers shift to northwestern Russia and northwestern Europe, with an additional center located off the Portuguese coast.

The East Pacific (EP) pattern is evident in all months except August and September, and is based on two height anomalies over the eastern North Pacific. When the EP is in its positive phase, a deeper than normal trough is exhibited in the vicinity of the Gulf of Alaska, a more pronounced northeastward extension of the Pacific jet stream, occurs and enhanced westerlies are found over northwestern North America. Its negative phase is indicative of a shallower trough, a split flow in the jet, and a reduction of the westerly flow.

3.3 Sea Surface Temperature (SST) Data

SST data are gathered monthly by the U.S. CPC and are available for a few select regions around the globe. Two regions, called Nino 3.4 which is located from 5^{o} N to 5^{o} S and from 170^{o} to 120^{o} W and, Nino 4 which is located from 5^{o} N to 5^{o} S and from 160^{o} E to 150^{o} W, were used in this study. The SST data were compared to the start days and non-windy days in the data mining and CART analysis.

IV. Standard Statistical Methods and Results

4.1 Methods

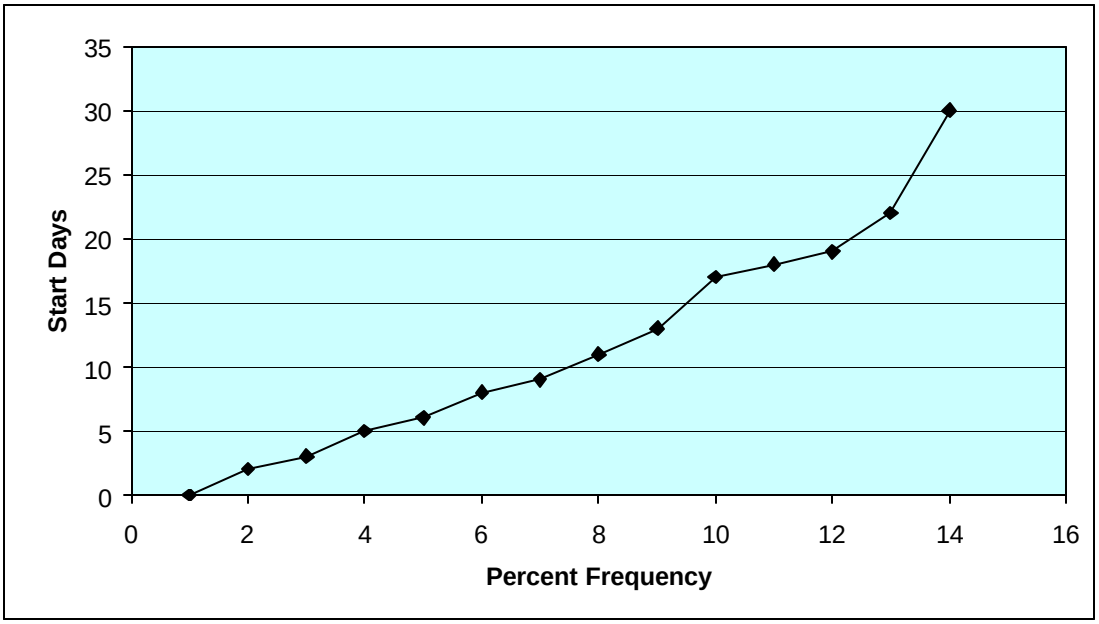
After omitting missing and non-matching years, the process of selecting days with winds less than 25 kts was performed. Two-week windows were selected and a frequency distribution of days with winds less than 25 kts was created. Months that included an excessive number of windy days (25 kts or greater) throughout the entire month were eliminated. This includes months in which all days were windy and months without two-week windows. This analysis was found to be consistent with climatology, which confirms June and July to be the least windy months on the island.

4.1.1 Selection of a Two-Week Window. In order for the MDA to complete installation of the XBR, a two-week window of winds less than 25 knots must be available. The Interactive Data Language (IDL) computing language was used to develop a program to determine these two-week windows. Appendix C contains the computer program source code and an accompanying flow chart to help explain the selection process.

The selection of two-week windows was based on determining a start day, which is defined as the first day of the construction project. In order for a day to be considered a start day it must contain 14 consecutive days with winds less than 25 kts. Once the criteria are met, the day is flagged as a start day and is written to an output file consisting of all start days included in the period of record. The total number of start days for each

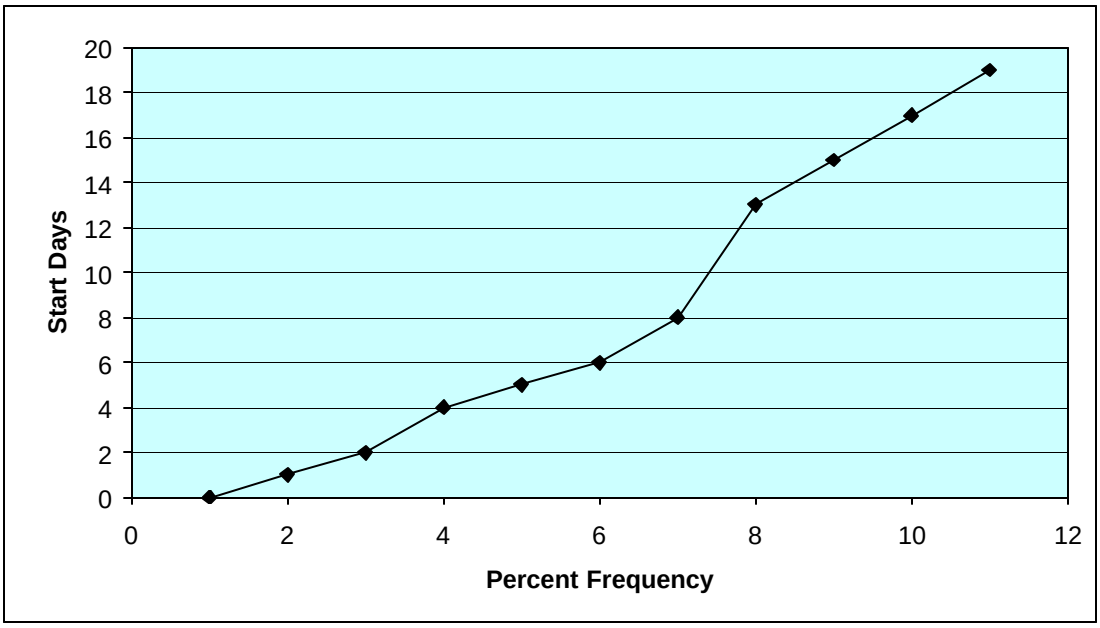
month was tabulated (Figure 15), analyzed statistically, and then matched per month to the computed teleconnection indices to develop any statistical relationship.

4.1.2 Frequency Distribution. A simple frequency distribution of non-windy days was accomplished to provide an additional data set for comparison (see Figure 16). A dichotomous analysis was accomplished by separating those days with wind 25 kts or greater (assigned a value of 1) from non-windy days (assigned a value of 0). A summation of windy days was then tabulated for each month.



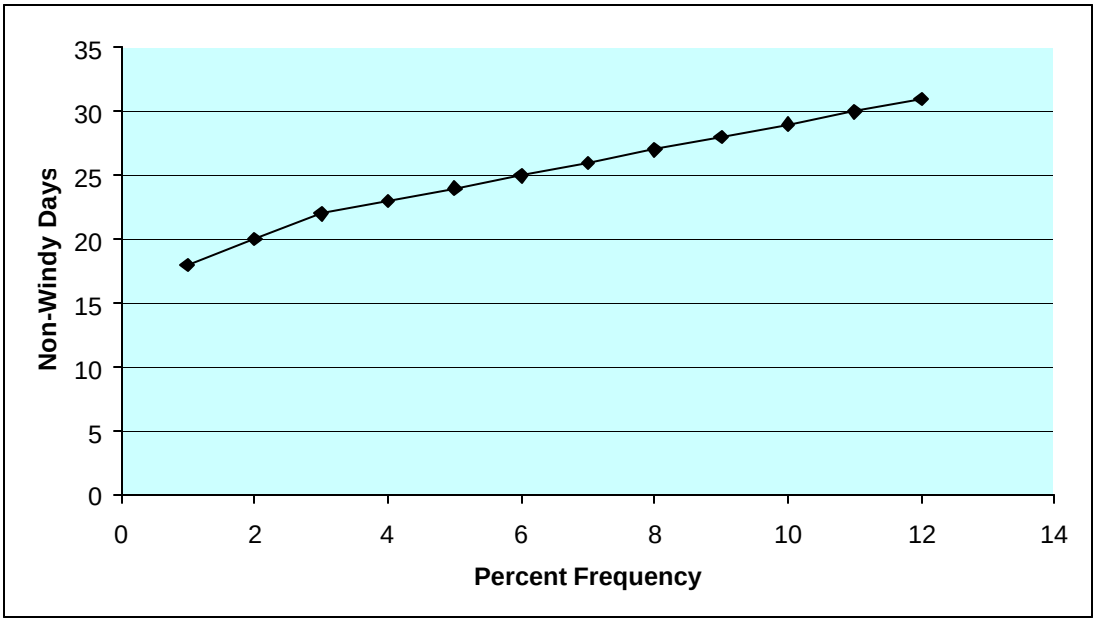
a. June

Figure 15. Frequency distribution of start days.



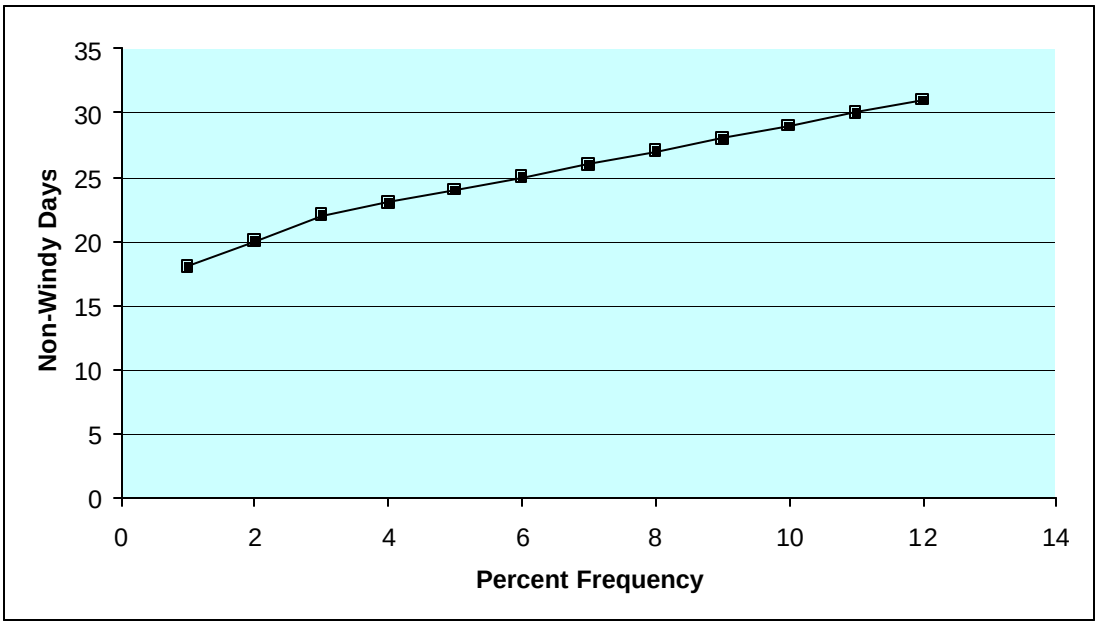
b. July

Figure 15. Continued.



a. June

Figure 16. Frequency distribution of non-windy days.



b. July

Figure 16. Continued.

4.2 Standard Statistical Analysis

Multiple types of regression analyses, investigating the relationships between two or more variables in a non-deterministic fashion, were performed to show any relationships between two-week windows (start days) and teleconnection indices. They also show any relationships between frequency of windy days and the teleconnection indices.

4.2.1 Simple Linear Regression. The simplest way to compare an X value to a Y value is to use simple linear regression. In this study, a relationship comparison of two-week windows (start days) and frequency distributions of non-windy days to each separate teleconnection index was performed. To determine if a significant statistical relationship exists, examination of correlations and p-values were executed, whereas the

correlation explains the proportion of X values that accurately describe the Y values and, a p-value defines how much confidence there is in the experiment (in this case maximum wind vs teleconnection index). In order for a relationship to be determined as a viable solution to the problem, the p-value must be below a value of 0.05 (95% confidence interval) and the correlation should be high enough to show a strong linear relationship. In order to establish the degree of any relationship, the value of the coefficient of determination is investigated, which is defined as the proportion of observed variation in the Y values that can be attributed to an approximate linear relationship between the values of X and Y.

4.2.2 Multiple Linear Regression. A more complex method to analyze data statistically is to perform multiple linear regression, which involves comparing one X value to many, or all Y values, or comparing Y values with themselves. Comparisons of relationships between teleconnection indices show the feasibility of using all of the indices in the study. The wind data were then compared with all of the teleconnection indices and the same goal for both probability and correlation was established.

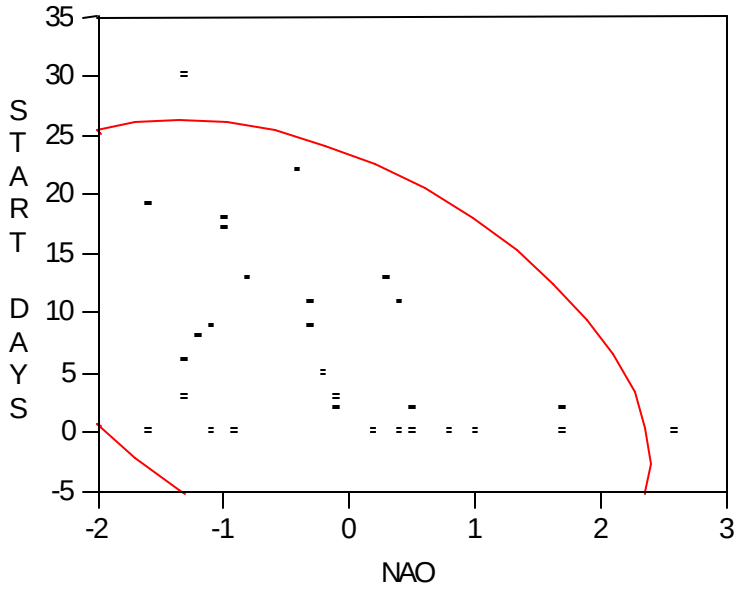
4.3 Standard Statistical Results

The data were analyzed from a predictive and observational standpoint. Table 3 shows the comparisons that were made for standard statistical and CART analysis, discussed in Chapter 5. June start days and non-windy days were compared to the May TIs (data set 1) and June TIs (data set 2). July start days and non-windy days were compared to June TIs (data set 3) and July TIs (data set 4).

Table 3. Statistical comparisons.

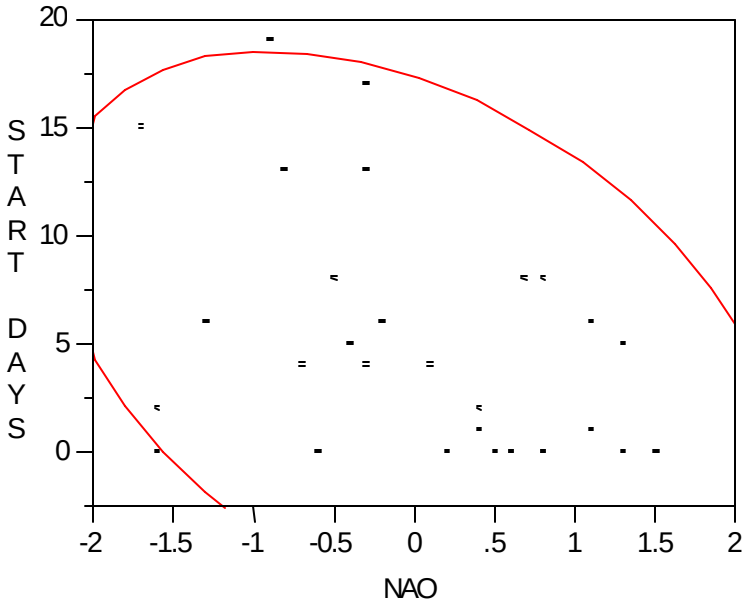
Data Type	TI Month	Wind Data Month
Start Day and Non-Windy Day Data Set 1	MAY	JUNE
Start Day and Non-Windy Day Data Set 2	JUNE	JUNE
Start Day and Non-Windy Day Data Set 3	JUNE	JULY
Start Day and Non-Windy Day Data Set 4	JULY	JULY

4.3.1 Simple Linear Regression. Using Simple linear regression, the only TI with promising results was the NAO. A graphical representation of the relationship of these data sets is displayed in Figure 17a-f. The correlation coefficient, r, explains how strong a relationship exists between the wind data and TIs. The coefficient of determination, R-squared, is the proportion of observed variation in one set of variables that can be explained by the simple linear regression model (relationship between x and y).

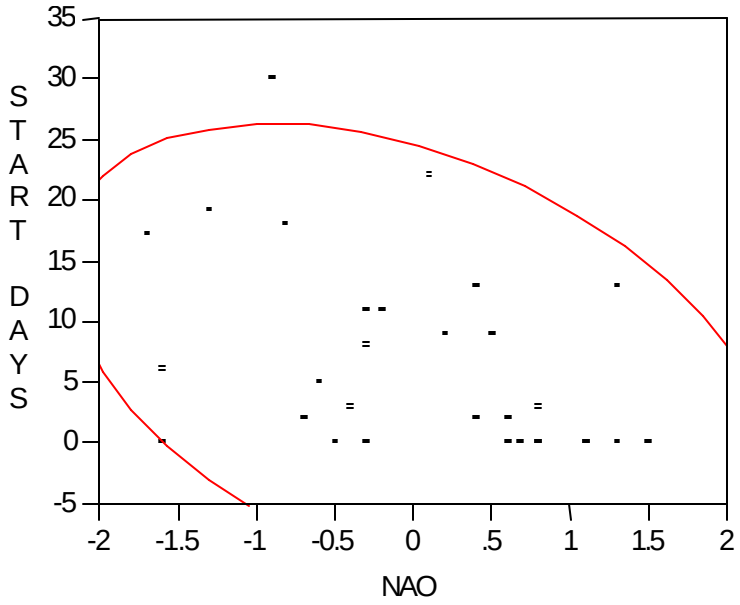


a. June Start Days vs. May NAO TIs.

Figure 17. Bivariate fit of wind data to TIs. The 95% confidence interval ellipse is drawn.

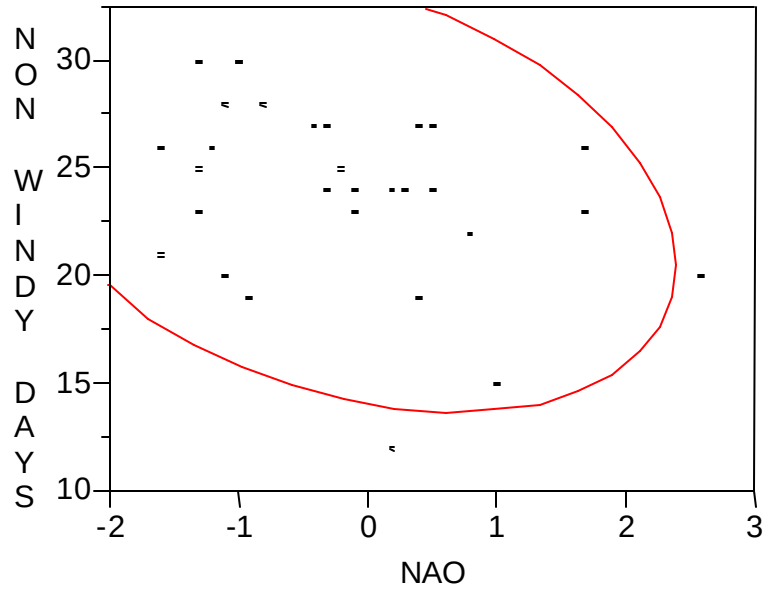


b. June Start Days vs. June NAO TIs.

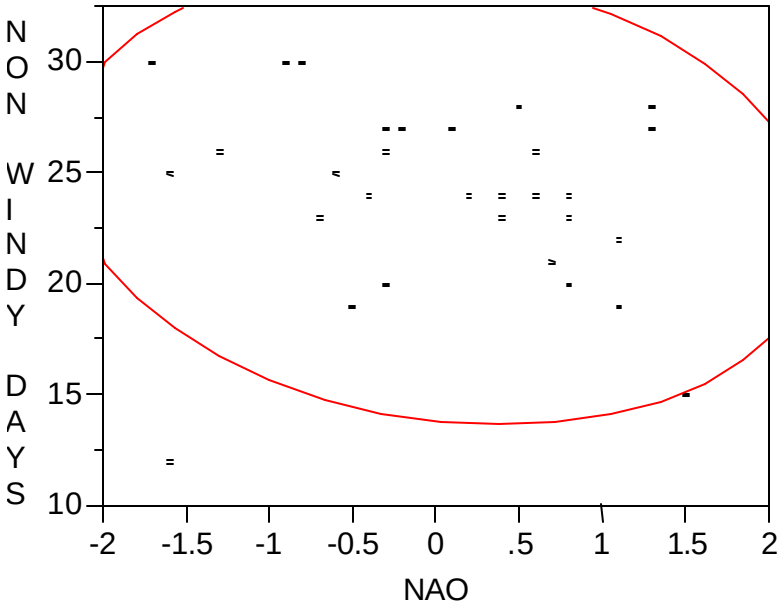


c. July Start Days vs. June NAO TIs.

Figure 17. Continued.

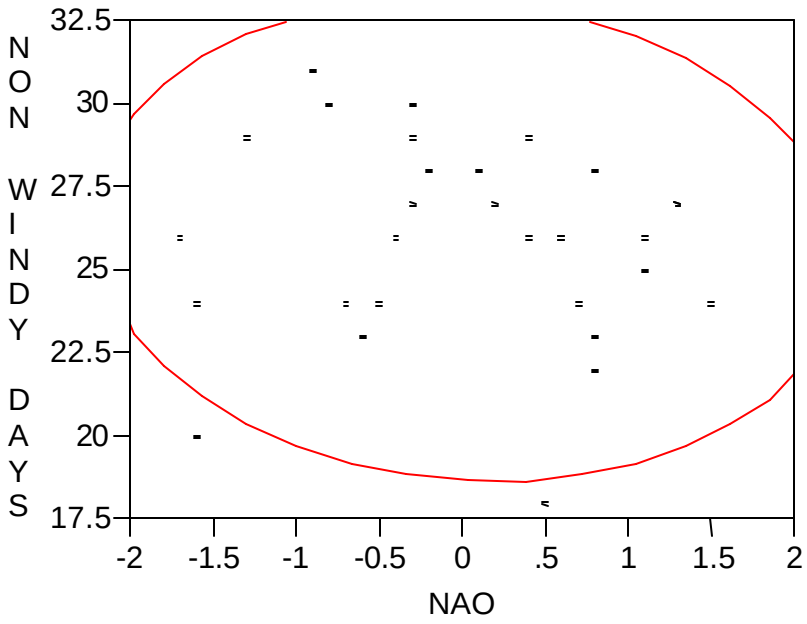


d. June Non-Windy Days vs. May NAO TIs.



e. June Non-Windy Days vs. June NAO TIs.

Figure 17. Continued.



f. July Non-Windy Days vs. June NAO TIs.

Figure 17. Continued.

A discussion of the findings from each of the previous figures follows, and the relationship strength and probabilities (P-values) for each are discussed and included in Table 4. Figure 17a has an r-value of –0.4625, R-squared of 0.2139, and a high predictability (P-value of 0.0101). For probability to be high, the P-value should be below 0.05, at the 95% confidence interval. These numbers suggest somewhat of a relationship between the start days and the NAO TI. A higher number of start days leads to a lower NAO TI number. In Figure 17b, an r-value of –0.4187, R-squared of 0.1753, and a P-value of 0.0213 was exhibited; the same relationship as discussed for Figure 17a was concluded. Figure 17c shows the same relationship as the previous two, but the relationship is not as significant. An r-value of –0.4123, R-squared of 0.1620, and high probability are the results. Of Figures 17d-f, d is the only one that showed any

relationship. An r-value of 0.3324, R-squared of 0.1105, and moderate probability were the results. Figures 17e-f have high P-values and low R-squared values, and should not be used to show a relationship between start days/non-windy days and NAO TIs.

The relationship between start days and the NAO can be explained by the strong connection it has to the rest of the Northern Hemispheric long wave pattern. When the NAO is in a more positive phase, higher pressure can be found over the Azores. It is common for the high pressure centers in this region to be strong enough to "block" usually zonal flow. This leads to a more meridional flow pattern over the globe and the slowing and strengthening of weather systems in their normal west to east migration. This in-turn could be responsible for stronger winds in and near associated frontal boundaries set-up due to more meridional flow, which could help explain why Shemya experiences fewer start days (i.e. more windy conditions) during years when the NAO is in a positive phase.

Table 4. Computations from Figure 17.

Figure Number	Correlation (r)	R-squared	P-value
15a.	-0.4626	0.2139	0.0101
15b.	-0.4187	0.1753	0.0213
15c.	-0.4128	0.1620	0.0234
15d.	-0.3324	0.1105	0.0727
15e.	-0.1675	0.0281	0.3765
15f.	-0.0836	0.0070	0.6606

4.3.2 Multiple Linear Regression. A multiple linear regression model was fit to the data, and involves comparing a single predictand (start days and non-windy days) to multiple predictors (TIs). Multiple regression models were created in hopes of showing better results than single regression. Model runs similar to the previous single regression models were performed. Poor relationships, with low correlations and P-values, were

discovered from all model runs. The most promising results were from the June wind data vs. May TI model runs. Tables 5 and 6 show the coefficient of determination (R-squared), predictability (P-values), and root mean square errors (RMSE) for each model run. RMSE describes how accurately a model can be used to forecast results. The closer the RMSE is to zero, the more precise a forecast is assumed to be. The best results were obtained from the comparison of the July wind data to the July TIs, with the lowest RMSE. However, this run did not have the highest R-squared or P-values.

Table 5. Multiple regression model run results, start days vs. monthly TIs.

Model Run	R-squared	P-value	RMSE
June vs. May	0.3296	0.2384	7.5592
June vs. June	0.3635	0.1390	7.3777
July vs. June	0.2596	0.2881	5.3703
July vs. July	0.2666	0.3738	5.4798

Table 6. Multiple regression model run results, non-windy days vs. monthly TIs.

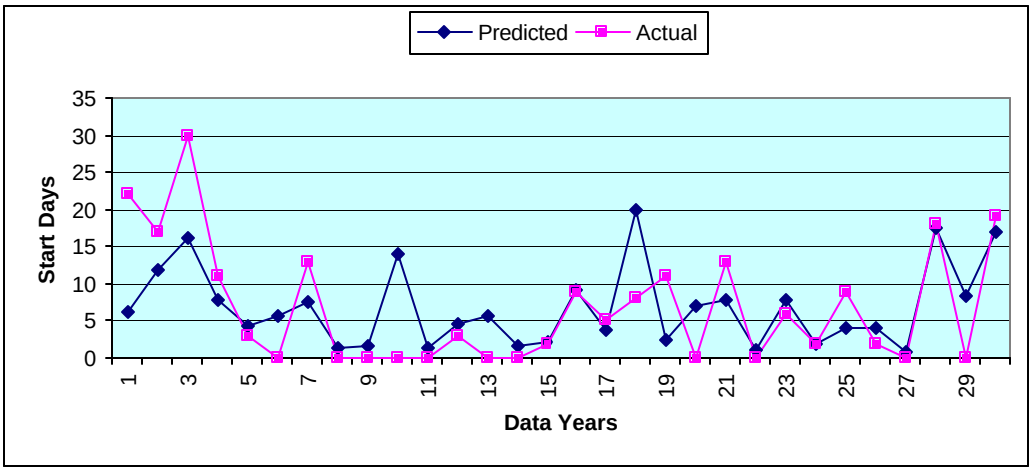
Model Run	R-squared	P-value	RMSE
June vs. May	0.2498	0.6693	4.3865
June vs. June	0.1730	0.7049	4.3913
July vs. June	0.1192	0.8759	3.1933
July vs. July	0.2932	0.2946	2.8605

4.3.3. Poisson Regression. An investigation of the start day models revealed that Poisson regression might show a more significant relationship between start days and TIs. Poisson regression assumes that data follow a Poisson distribution, which is evident here when counting a number of events (start days in this case). The Poisson distribution must be non-negative, skewed, and the variance of the distribution must increase along with the mean (McCullagh and Nelder, 1983). Poisson regression was used to model the number or rate of occurrences of start days as a function of the TIs and, a Poisson

regression model was run for each of the previous cases. The model is a log-linear based model where expected values $E[y_i|X_i]$, are given by the equation:

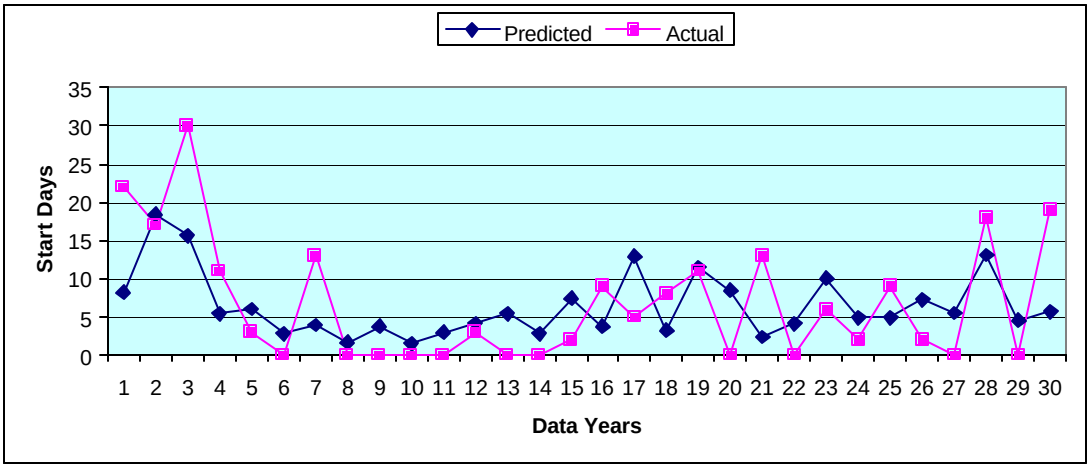
$$E[y_i|X_i] = \exp(\beta_1+\beta_2 x_{2i}+\cdots\beta_k x_{ki}) \qquad (1)$$

In this research, the y values (expected values) are the two-week windows that are forecast by the model. The estimated values given by the model run are $\beta_1, \beta_2, \beta_k$ (where k is the total number of TIs used in the model and i is the number of outcomes), and the known predictors (TIs) are the x values. The result of each model run is presented in Figure 18, which shows the predicted values, computed using the equation above and plotted along with the actual values. A summary interpreting these results using count data analysis is listed in Tables 7-10. The tables show the parameter estimates that were used in each model run, along with P-values and Pseudo R-squared values. Pseudo R-squared values coincide with the interpretation of R-squared values in the linear model, however there is no universal definition for them in non-linear models (Cameron and Trivedi, 1998).

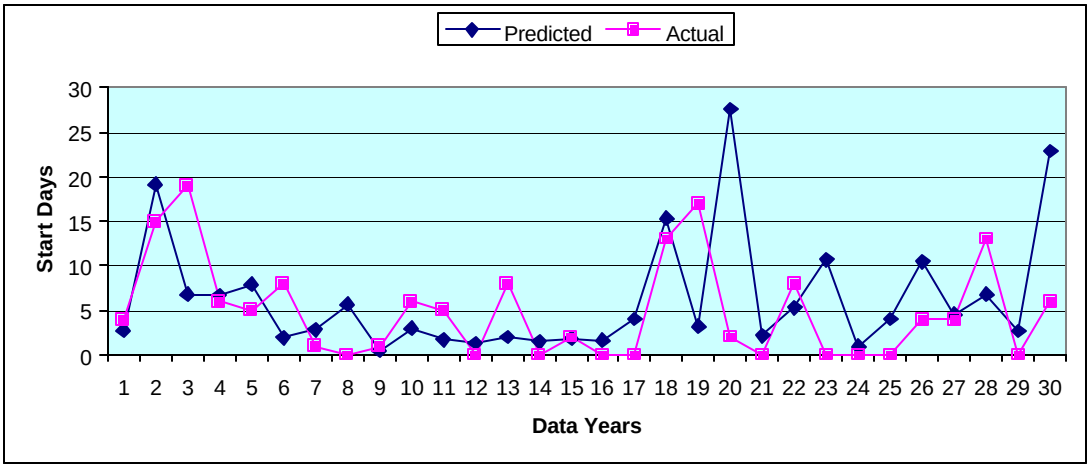


a. June start days vs. May TIs.

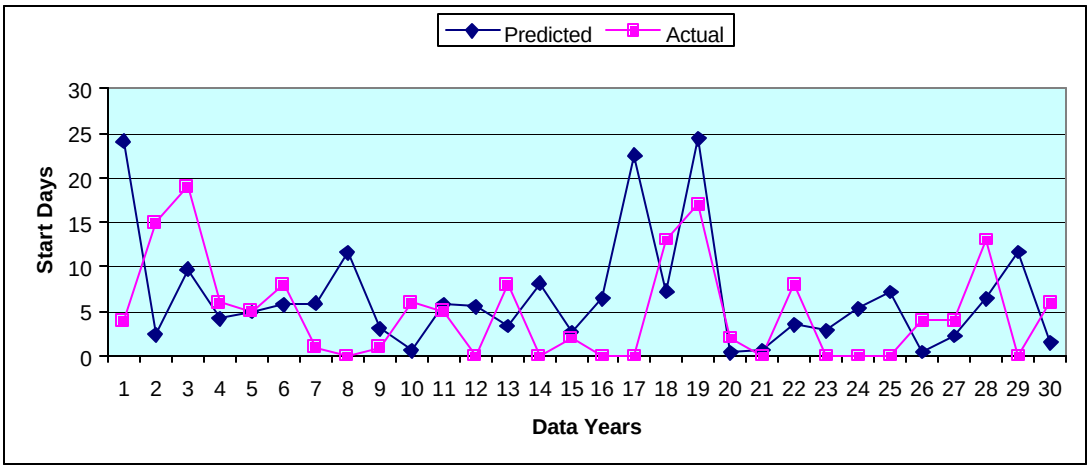
Figure 18. Predicted and actual number of start days using Poisson regression.



b. June start days vs. June TIs.



c. July start days vs. June TIs.



d. July start days vs. July TIs.

Figure 18. Continued.

Table 7. Results from analysis of Poisson regression model,
June start days vs. May TIs.

Parameter	Estimate	Test (P-Value)	Pseudo R^2
Intercept	1.3824	< 0.0001	
NAO	-0.6289	< 0.0001	
WP	0.3531	< 0.0001	
PNA	-0.2153	< 0.0174	
PT	-0.2259	< 0.0049	0.5592

Table 8. Results from analysis of Poisson regression model,
June start days vs. June TIs.

Parameter	Estimate	Test (P-Value)	Pseudo R^2
Intercept	1.7309	< 0.0001	
NAO	-0.8194	< 0.0001	
NP	0.2544	0.0140	
SOI	0.3332	< 0.0001	0.5071

Table 9. Results from analysis of Poisson regression model,
July start days vs. June TIs.

Parameter	Estimate	Test (P-Value)	Pseudo R^2
Intercept	1.3767	< 0.0001	
NAO	-0.6975	< 0.0001	
NP	0.1490	0.1923	
ASU	0.4704	0.0005	
SOI	0.2441	0.0009	0.4474

Table 10. Results from analysis of Poisson regression model,
July start days vs. July TIs.

Parameter	Estimate	Test (P-Value)	Pseudo R^2
Intercept	1.3245	< 0.0001	
NAO	-0.2751	0.0088	
WP	0.2468	0.0080	
NP	-0.6855	< 0.0001	
SOI	1.0000	< 0.0001	0.4646

The results of Poisson regression show slight improvement over multiple regression results. Lower P-values and higher R-squared values are evident by examination of Tables 7-10. As in the previous cases (single and multiple regression analyses), using the May TIs to predict the June start days shows the best results. This model run shows the highest number of predicted values to observed values; however, a high number of days that the model either over-forecasted or under-forecasted for that period still exists.

The challenge of using TIs to forecast much smaller-scale weather parameters (Shemya winds) is apparent in the results obtained using standard statistics. Therefore, data mining through CART analysis is the next logical step to try in hopes of finding any prediction.

V. CART Overview, Methods, Results and Application

5.1 CART Overview.

Classification and regression tree (CART) analysis is based on using tree diagrams to show binary decision rules (or decision points), has no parameters, and is non-linear. At each decision point in the tree (called a node), the tree splits into two separate branches (called child nodes). The direction of the branched value is based on a test against a threshold predictor value (in this study, TI and SST data), and continues until a final point is reached (Burrows, 1992). This splitting process is known as binary recursive partitioning (continually splitting the tree into two branches) and continues until the ideal outcome, or full-sized tree is reached (a tree with greatest improvement in predictive accuracy). Building a tree from start to finish involves selecting a set of data for each node of the tree, splitting each node in the tree, deciding when each tree is finished growing, validating the accuracy of the tree (cross-validation), pruning the tree, and (when using regression trees) to assign each node a predicted value for regression (mean, standard deviation, and number of observations).

5.1.1 Splitting the Tree. CART splits each node into two child nodes using either *Gini* or *Twoing* methods to reduce misclassification errors and the child nodes are assigned a yes or no answer. The tree contains a finite number of splits, which CART continues to scour and rank order. *Gini* takes the largest class in the database and isolates it from the other classes. Then, the split is created to search for the best division in the tree that produces the highest amount of purity; i.e., the most statistically significant tree (Breiman, 1984). CART produces an improvement value for each split, which is a

measure of how well the split improves the predictive performance of the tree. *Twoing* splits each class of each node into two superclasses, in order to reach the greatest decrease in impurity. *Twoing* incorporates a "strategic" splitting of each node (Breiman, 1984), in that it groups together a large number of classes at the tree top and isolates single classes at the tree base. The *Gini* method of splitting was incorporated in this research.

5.1.2 Class Assignment and Pruning. Once the best split is found for a node, CART repeats its search for each additional node that it develops. This continues until the splitting of nodes is no longer possible, CART has exhausted all statistical computations for the assigned model run, or it is determined the node has too few cases (low number of observations). After building the tree, CART examines smaller versions of the tree to see if more accurate information can be gained from these smaller versions of the larger tree. If more accurate information can be gained from the smaller version, CART may prune the tree to resemble a smaller tree. Accomplishing manual pruning on each tree gives even more accurate results, and is accomplished by setting limitations on the total number of n outcomes in each parent and child node. A student "t" test was also accomplished on each split to ensure that the tree was producing sound statistical results.

5.1.3 Cross-Validation. In this study, the data set is not large enough to allow for a test sample to yield promising results by the use of classical regression. CART implements cross-validation to grow the tree, and when data are in short supply (usually less than 50 for a sample size), grows the tree using the entire sample of data. Then, CART divides this sample into a given number of separate and equal parts known as folds, or sub-trees. In a 10-fold analysis, 10 sub-trees are grown and 1 out of 10 of the

sub-trees is left out of the computation and compared to the other sub-trees. For example, CART holds the first sub-tree and compares it to the remaining nine sub-trees, then holds the second tree comparing it to the first and all remaining sub-trees. This process is repeated until each of the folds has been tested and the error rates of all are applied to the entire tree. This process is accomplished to allow for maximum tree accuracy when CART constructs the final full-sized tree. An estimate of the risk statistics and cross-validation is computed to show the predictability level of the final tree, and continues until the highest improvement is found. Recall that improvement is computed for each node and is defined as the measure of decrease in impurity for each predictor, or more simply the measure of how well a certain split improves the overall predictive performance of the tree. The goal is to have the lowest impurity number, and thus a high improvement score.

5.2 CART Methodology and Results

CART is implemented in this study to develop a relationship between TIs and SSTs with start days and non-windy days on Shemya, as previously defined. A tree is constructed for each month using both classification and regression tree analysis. The results of the trees were then compared to a set of selected synoptic charts. In the case of this study, the nodes were split based on a TI or SST value that CART found to hold the best predictive solution for the tree. Pruning each tree was accomplished to produce a more effective forecast process by finding the optimal pruning number, which is defined as the most effective number of n observations in each parent and child node. In this study, cross-validation was incorporated using 10 folds, due to the small data set (30

years of data). CART computes a risk estimate for each cross-validated tree, whereas lower cross-validation risk estimate values indicate a stronger relationship between the input data (TIs and SSTs) and the model output (the tree itself) exists.

5.2.1 Classification Trees. In the case of classification trees, a tree was used for each month studied and each node consisted of multiple wind categories. The start days were separated into three categories; a category of two was assigned for above-normal start days, a category of one assigned for near-normal start days, and a category of zero assigned for below-normal days. The same was carried out using the data for non-windy days. After growing full-sized trees, they were individually pruned by selecting the most effective value of n (number of years) to assess which tree would yield the most significant results. In this case, a minimum value of n=4 for both the parent and child nodes was found to generate the strongest association between data and predictors (highest improvement scores).

An example of one of the classification trees used in this study is presented in Figure 19, which was grown using June start day categorical data compared to May TI and SST continuous data. To follow the tree from start to finish, one would use the split criteria for each node. In this case, starting with the PNA at a split value of 0.05, proceed to the resulting node of 1 that splits again, or 2 that ends. If it is node 1, then use the index split criteria for that node (SOI value of –1.45), and continue until an end node is reached. Cross-validation risk estimates were tabulated for each tree, and the results for all model runs are presented in Table 11.

After individually pruning the trees using the criteria previously stated, running a risk estimate synopsis for each tree, and evaluating each node's categorical data set; the

classification trees did not prove to be useful predictive methods of forecasting the number of below-normal, near-normal, or above-normal start days or non-windy days. There was no improvement seen over climatology for all four trees that were run using the following tree accuracy method (Freestrom, 2002): 1) subtract the cross-validation risk estimate from 1 to compute the accuracy of the tree, 2) compare the tree accuracy to climatology which is defined as 33% of the data, and 3) if the tree accuracy is greater than climatology (a cross-validation risk estimate less than 0.66) then the model shows more significance than climatology. The cross-validation risk estimate for all classification tree model runs were not less than 0.66, so there was less significance than climatology. Improvement scores were also extremely low, which indicates a high amount of impurity in the trees being created. The small sample size of n=30, and the relatively high number of below-normal start days (category 0) and non-windy days evident in the data set, along with relatively high cross-validation risk estimates, contribute to this dilemma. Because of this result, regression tree analysis was next performed to more effectively mine the data for predictive relationships.

5.2.2 Regression Trees. Regression trees use continuous data, and were developed for each month studied in the same manner as the classification trees. When grown, the trees created nodes, with each node containing statistical information such as the mean and standard deviation. Figures 20 and 21 represent the start day regression trees that were created for this study. Following guidance in Randall (2002), a goodness of fit test was accomplished to see how normally distributed the start days and non-windy days were, using the Shapiro-Wilks method (See Table 12). A normal distribution can be identified when the P-value is greater than 0.05 (at the 95% confidence level). According

to this method, the non-windy days were assumed to be normally distributed. The start days did not appear to be normally distributed when examining the Shapiro-Wilks results, however, by investigation, the distribution has no outliers and has the physical appearance of a normal distribution. When working with relatively small data sets, which indeed were the case in this study, the Shapiro-Wilks test is not the sole means available to examine normally distributed data, and therefore it was assumed that the start days were also normally distributed.

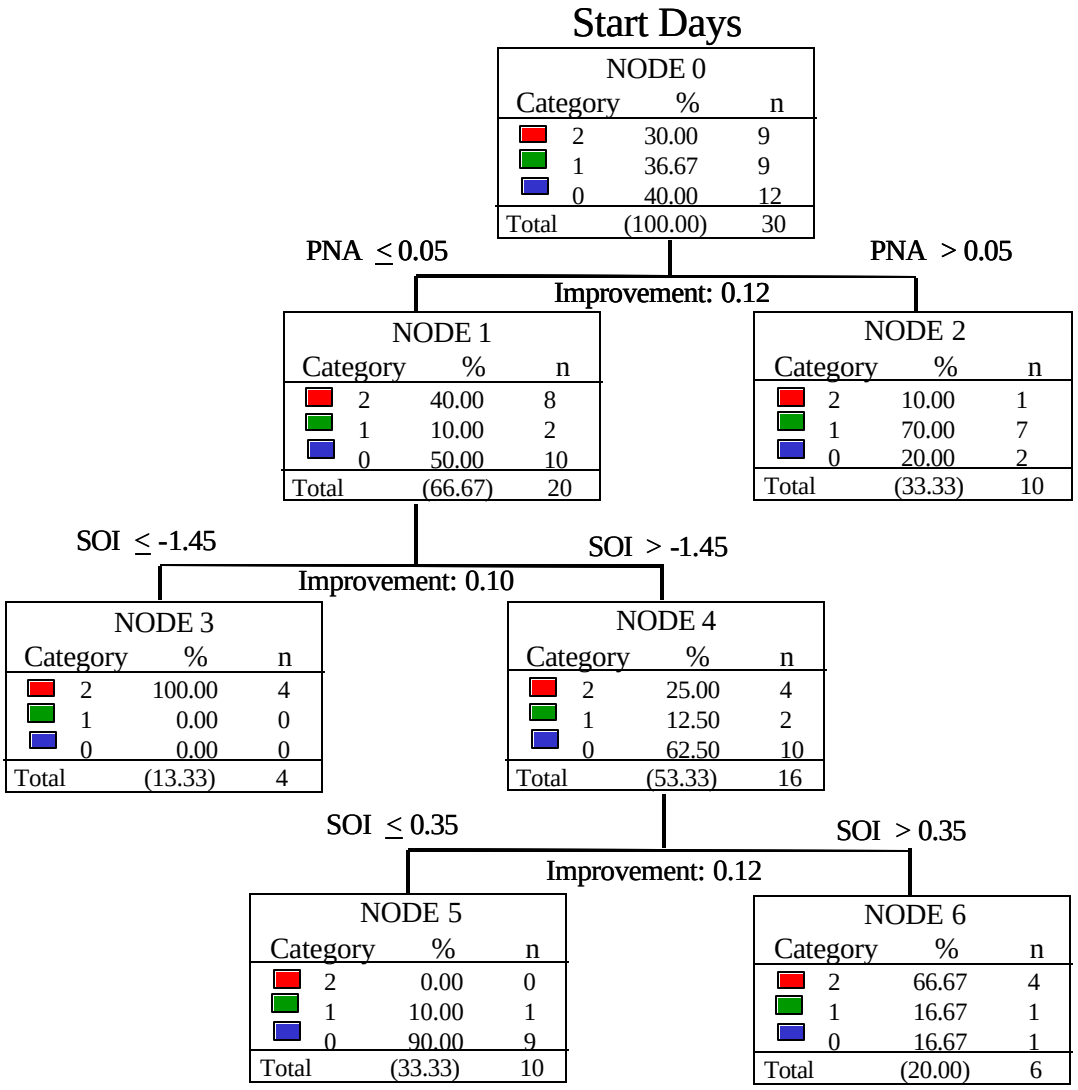


Figure 19. Example of a classification tree with start days from June compared to TIs and SSTs from May.

Table 11. Cross-validation risk estimates for each model run.

Model Run	May TIs and SSTs Vs. June Data		June TIs and SSTs Vs. July Data	
	start days	non-windy days	start days	non-windy days
Risk Estimate	0.67	0.67	0.67	0.73

Table 12. Shapiro-Wilks test for normality.

Model Run	May TIs and SSTs Vs. June Data		June TIs and SSTs Vs. July Data	
	start days	non-windy days	start days	non-windy days
P-Value	0.0001	0.0733	0.0001	0.3787

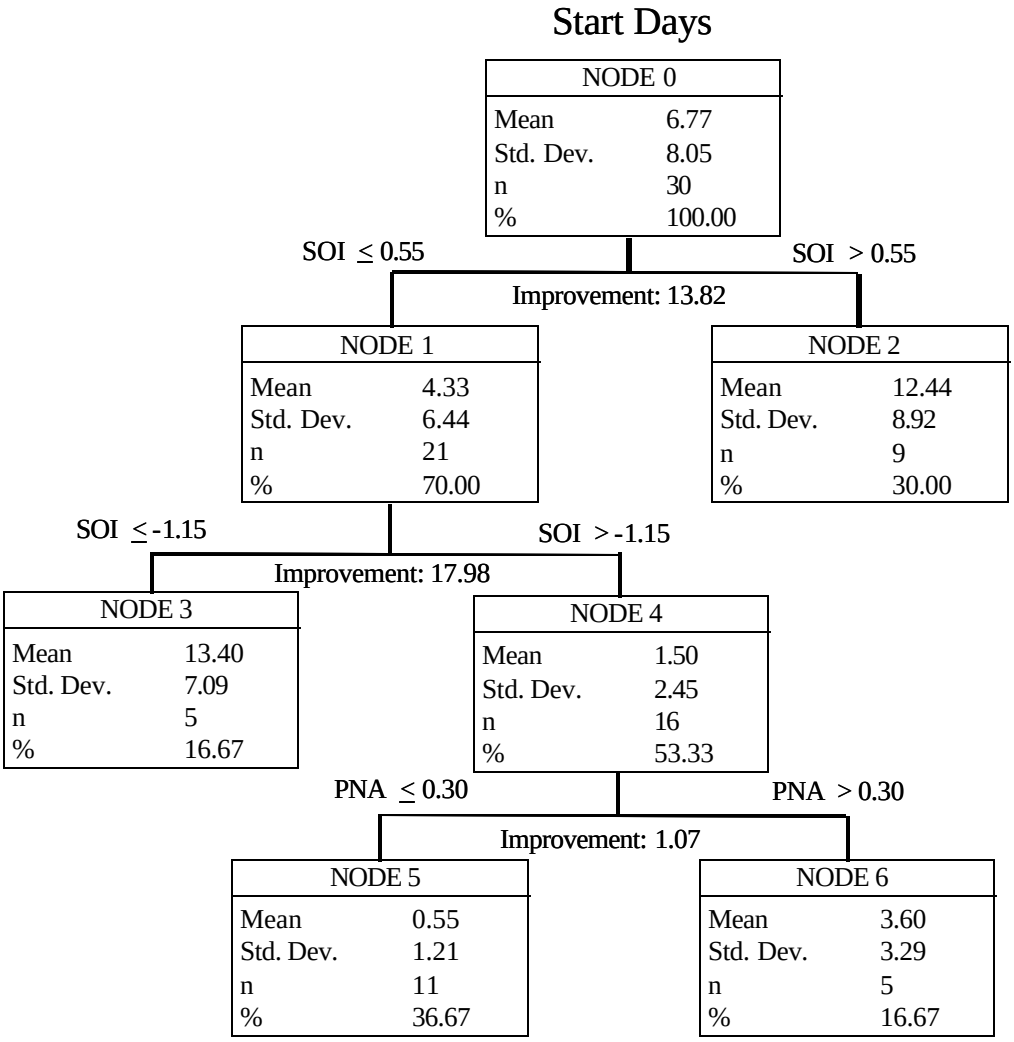


Figure 20. Example of a regression tree with start days from June compared to TIs and SSTs from May. This tree has been manually pruned.

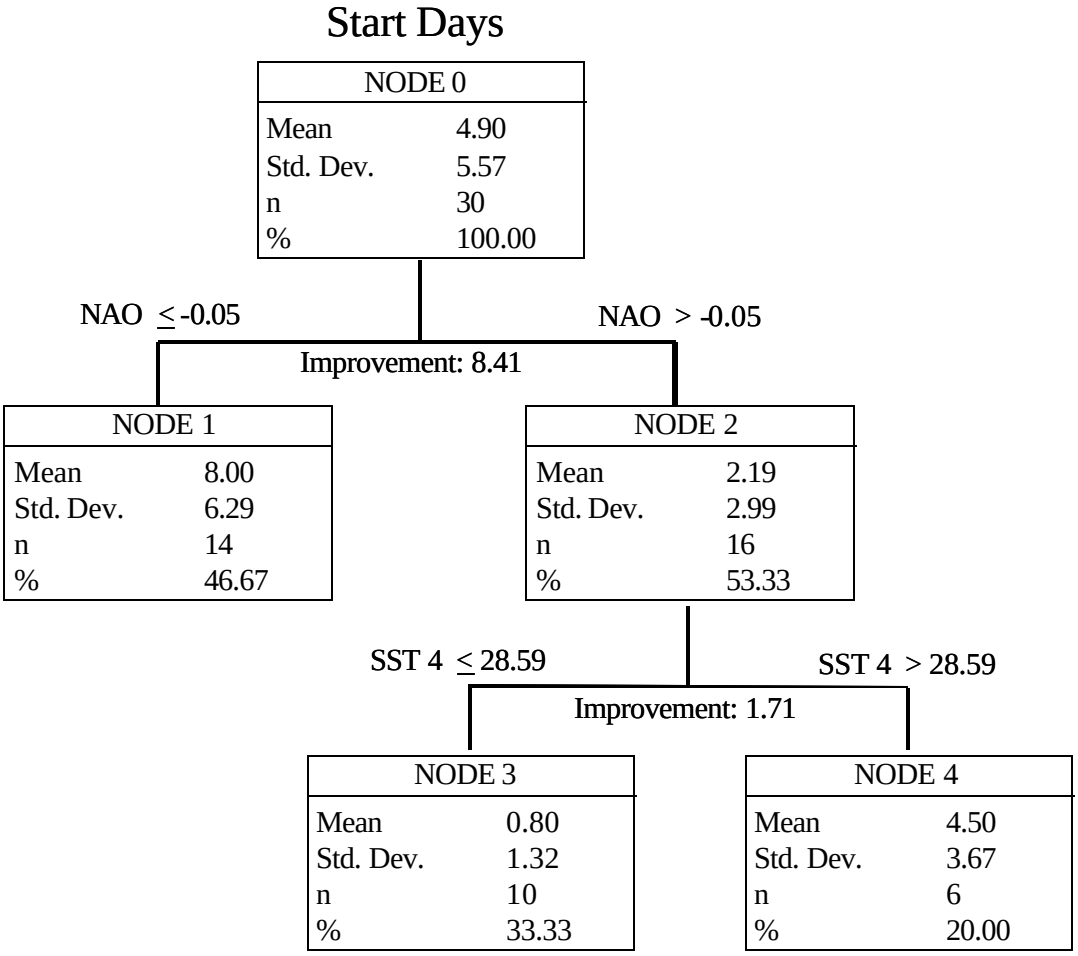


Figure 21. Example of a regression tree with start days from July compared to TIs and SSTs from June. This tree has been manually pruned.

The statistical information computed by each tree was used to calculate 95% confidence and prediction intervals for all models (see Tables 13 and 14). Each year of data was then run through each tree model to assess what the ending node would be for that year, and then the ending node was compared to the confidence and prediction interval for that node. An evaluation was then made to assess how much of the data fell within the computed prediction interval for that year, and this in turn determines the accuracy of the tree. Very promising results with high tree accuracies and significant improvement scores were discovered for all model runs (Tables 15 and 16).

Table 13. Confidence and prediction interval computations for ending nodes of start-day tree model runs.

May Vs June	95% Confidence		95% Prediction	
Node:	Lower	Upper	Lower	Upper
2	6.6	18.3	1.6	23.3
3	7.2	19.6	2.6	24.2
5	0.0	1.3	0.0	2.0
6	0.7	6.5	0.0	8.6

June Vs July	95% Confidence		95% Prediction	
Node:	Lower	Upper	Lower	Upper
1	4.7	11.3	1.0	15.0
3	0.0	1.6	0.0	2.4
4	1.6	7.4	0.0	9.6

Table 14. Confidence and prediction interval computations for ending nodes of non-windy days tree model runs.

May Vs June	95% Confidence		95% Prediction	
Node:	Lower	Upper	Lower	Upper
2	16.1	25.4	12.2	29.3
3	25.1	29.3	23.5	30.8
5	20.7	24.2	19.2	25.7
6	24.5	27.3	23.2	28.6

June Vs July	95% Confidence		95% Prediction	
Node:	Lower	Upper	Lower	Upper
1	21.6	26.1	19.8	27.9
2	25.4	27.6	23.7	29.4

Table 15. Tree accuracy at the 95% confidence (CI) and prediction (PI) intervals for start days.

	May TIs and SSTs Vs. June Data		June TIs and SSTs Vs. July Data	
	95% CI	95% PI	95% CI	95% PI
Accuracy	0.73	0.77	0.63	0.83

Table 16. Tree accuracy at the 95% confidence (CI) and prediction (PI) intervals for non-windy days.

	May TIs and SSTs Vs. June Data		June TIs and SSTs Vs. July Data	
	95% CI	95% PI	95% CI	95% PI
Accuracy	0.57	0.90	0.47	0.73

To determine the level of the tree's predictability, a test of independence of the child nodes from each parent node split was accomplished by using the two child nodes created by the split. A student "t" test was then used to examine the predictability level of each split node created by CART, where

$$t = \frac{X_1 - X_2}{\sqrt{\frac{N_1 \cdot s_1^2 + N_2 \cdot s_2^2}{N_1 + N_2 - 2} \cdot \left(\frac{1}{N_1} + \frac{1}{N_2} \right)}} \qquad (2)$$

and, the values X_1 and X_2 are the computed means for each node, N_1 and N_2 are the total number of observations in each node, and s_1 and s_2 are the standard deviations for each node. The "t" values were compared to the tree's improvement score to determine what level of improvement denoted significant tree predictability. If the majority of the model

shows both high improvement and a t-value lower than the t-table value, it is assumed to show ample predictability (See Table 17).

Table 17. Student "t" value improvement table computed for each model run. Acceptable values and model runs are bold-faced.

Model Run	days	split nodes	t-value	t-table	improvement
May Vs June	**start**	1 and 2	**-2.703**	2.048	**13.82**
		3 and 4	5.431	2.093	**17.98**
		5 and 6	**-2.529**	2.145	1.07
	non-windy	1 and 2	2.215	2.052	2.04
		3 and 4	**2.040**	2.069	1.28
		5 and 6	**-2.859**	2.110	1.89
June Vs July	**start**	1 and 2	**0.861**	2.048	**8.41**
		3 and 4	**-1.058**	2.145	1.71
	non-windy	1 and 2	**-1.058**	2.048	1.27

5.3 Forecast Guidance

The results obtained from regression tree analysis are used to produce a forecast decision aid for a two-week window of opportunity for MDA construction operations for the XBR. The forecaster should follow the regression tree using the split criteria to reach the ending node in that tree. Once the ending node is determined, the forecaster should use the table of confidence interval values (Table 13) to assess the range of start days that is expected. Only the trees with the highest predictability (start day model runs) are used as decision forecast models. Before presenting the forecast guidance, an explanation of the occurrence of TIs and SSTs in the trees and a comparison between tree results and synoptic conditions is necessary. CART analysis splits the tree based on the strongest improvement of the TI and SST data. Of the TIs and SSTs used in this study, CART chose the SOI and PNA for the May vs June model run and the NAO and SST 4 for the June vs July model run in its splitting of each node (Figures 20 and 21).

Regression tree analysis shows the SOI splitting at 0.55 and –1.15 for the May vs June model run. In one split out of the two, the resulting nodes indicated less windy conditions for SOI values above 0.55, and the other split indicated less windy conditions for SOI values greater than –1.15. The split criteria indicate more windy conditions for SOI values that approach zero. A more negative SOI represents below-normal air pressure over Tahiti and above normal air pressure over Darwin, which can be linked to warmer than normal SSTs in the tropical East Pacific. A more positive value would represent colder than normal SSTs.

During the pruning process, it was discovered that one SST value out of the 30 was negatively affecting the tree's improvement at the first split. This value occurred in 1982, which is known to be one of the strongest and most active ENSO events on record, and resulted in those child nodes exclusively related to the event being pruned out of the tree.

The PNA was split at a value of 0.30 for the May vs June start day model run, with a less windy result for values greater than 0.30. A more negative PNA is related to a stronger Aleutian Low over the North Pacific, and thus stronger winds over Shemya.

The NAO, the only TI shown to have a significant direct linear relationship with the start days, had a split at –0.05 (nearly zero). When the NAO is in a more positive phase, less start days are exhibited (more windy conditions). An explanation of why this is indeed the case is given in Chapter 4, where a strong bivariate relationship between start days and the NAO was revealed. The NAO is one of the most influential of the TIs and dramatically affects long-wave patterns in the Northern Hemisphere due to its

coverage over the polar region. The coverage of the NAO pattern over the poles might be tied to MPH activity and their frequent movement over Shemya.

SST 4 had a split at 28.59 (degrees C), which means temperatures below this value indicated more windy conditions (less start days) and occurred in the start day model for the June vs July run. It is usually assumed that higher SST values are tied to more tropical activity and higher wind speeds.

The model runs were then compared to synoptic conditions and tropical systems for years of data with no start days and again for years with a high number of start days (possible outliers). Upper-air and surface charts from June and July of 1995 and 1996 were analyzed to find any association of pressure centers, fronts, and jet locations to the start days or indices. These years were chosen because of the availability of the data, and the significant differences between the two years (1995 being relatively calm, and 1996 being windy). Tropical activity was also investigated to compare the number of storms to the TIs.

During 1995, which was relatively calm during June and July with a total of 18 start days in June and 13 in July (see Figure 22 and 23), the main storm track was just south of Shemya with fairly light winds in the vicinity of passing centers of low pressure. Very strong winds (30-35 kts) were indicated just south of the islands (approx. 100nm) in proximity to the active jet region and associated storm track. The number of tropical systems was higher in July with three, compared to only one in June. Later in July, the main storm track begins to move southward, closer to Shemya. For the May vs June model run comparison, the stronger high pressure influence over the North Pacific combined with the relatively weak low pressure center over the Gulf of Alaska, can be

connected to the PNA which had a value of –1.1. A SOI of 0.2 was registered for that time period and matches the relatively benign tropical activity. For the June vs July model run, the NAO value of –0.8 matches well with the strong low pressure that was analyzed over the north polar region.

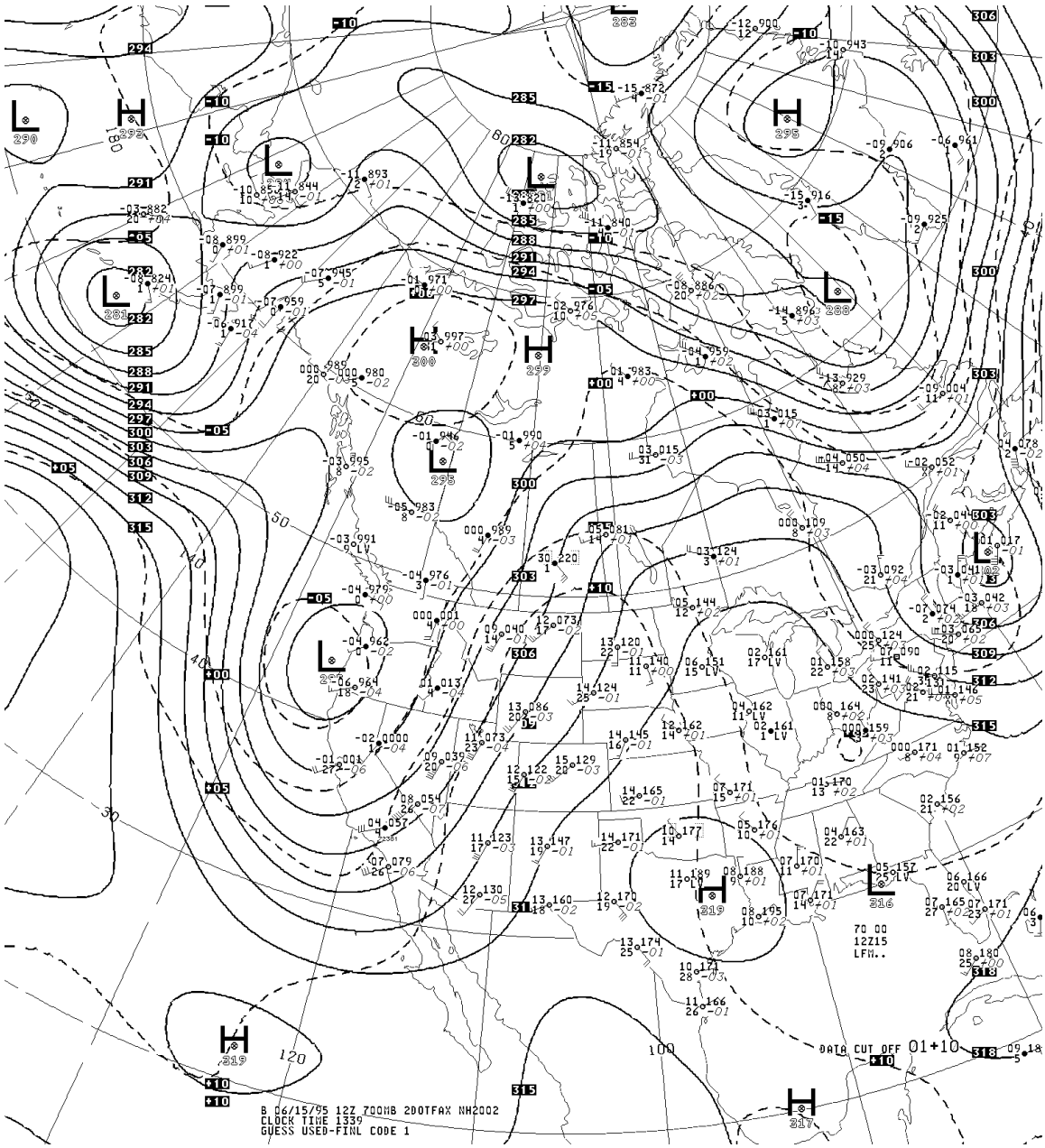


Figure 22. 700 mb analysis of heights and temperature for 15 June 1995 at 1200Z (National Climatic Data Center (NCDC, 1995).

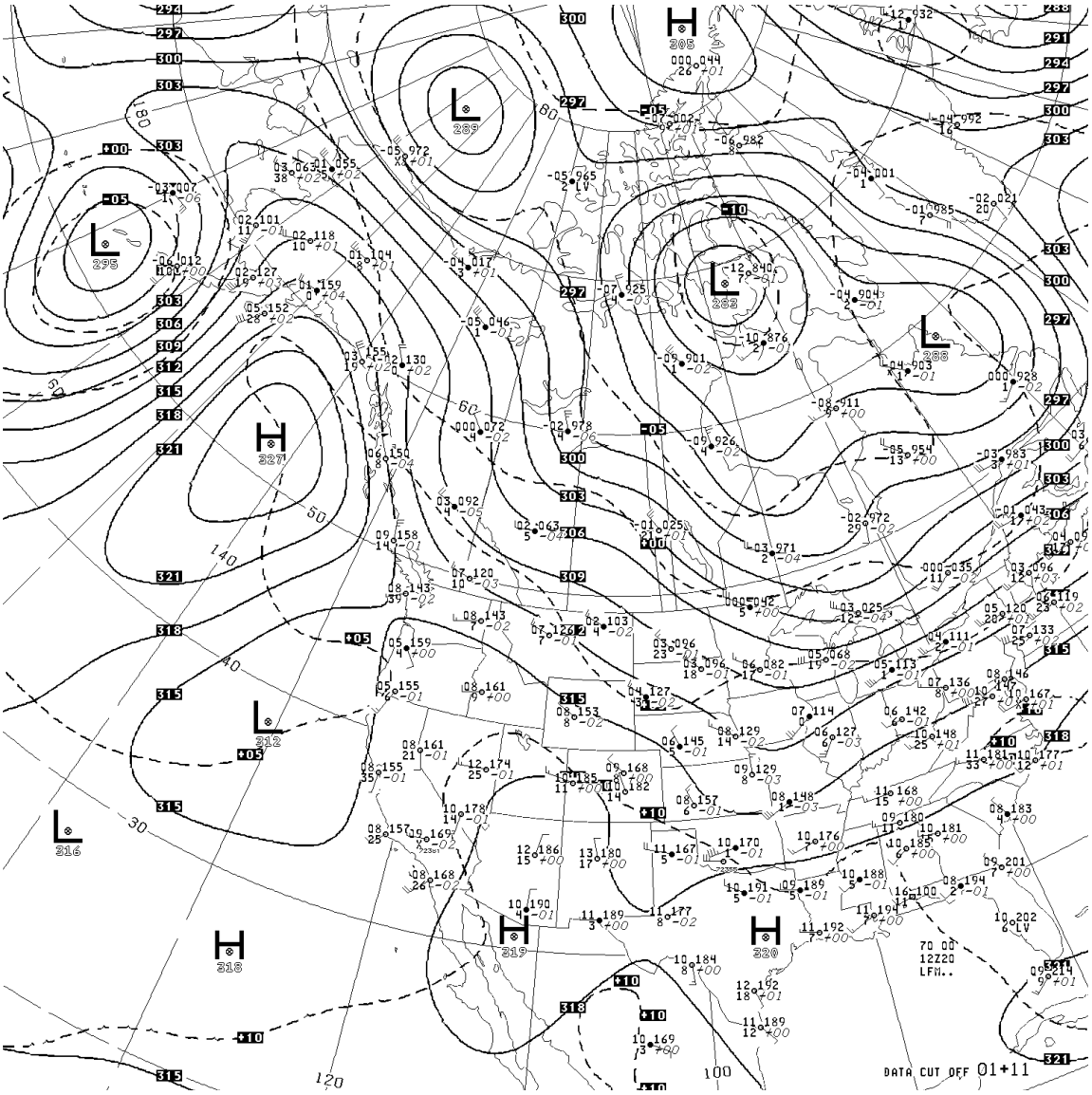


Figure 23. 700 mb analysis of heights and temperature for 20 July 1995 at 1200Z (NCDC, 1995).

The year 1996 was characterized by quite windy conditions with no start days for either June or July, a strong North Pacific High was present, and the storm track and related jet stream moved directly over the island (see Figures 24 and 25). A very deep low pressure center was evident moving from the Bering Sea into the Gulf of Alaska during the June to July period. A comparison of the June 700 mb chart (Figure 24) was

made to the May vs June model run. The SOI value was 0.2, which is indicative of windy conditions (very near zero), and the PNA value of –2.2 shows very windy conditions would exist due to the high negative value. More negative PNA values denote a stronger low pressure center over the Aleutians and therefore stronger winds. The June vs July model run was compared to the July 700 mb chart (Figure 24), and the link between synoptic features and the indices computed for that month was accomplished. It was discovered that the NAO value of 0.8, which indicates windy conditions, matches well with the synoptic features on the chart and the findings from Chapter 4. The positive NAO value can be coupled to the meridional flow and strong pressure centers exhibited on the chart. As in the prior case, tropical systems for this year were also more numerous during July.

By understanding the Northern Hemisphere long wave pattern and synoptic conditions over the region along with the results from the regression tree models, forecasters can assess the risk of windy conditions on the island. This information, along with other local forecasting tools, may be used to help planners determine the best opportunity to complete the XBR construction project.

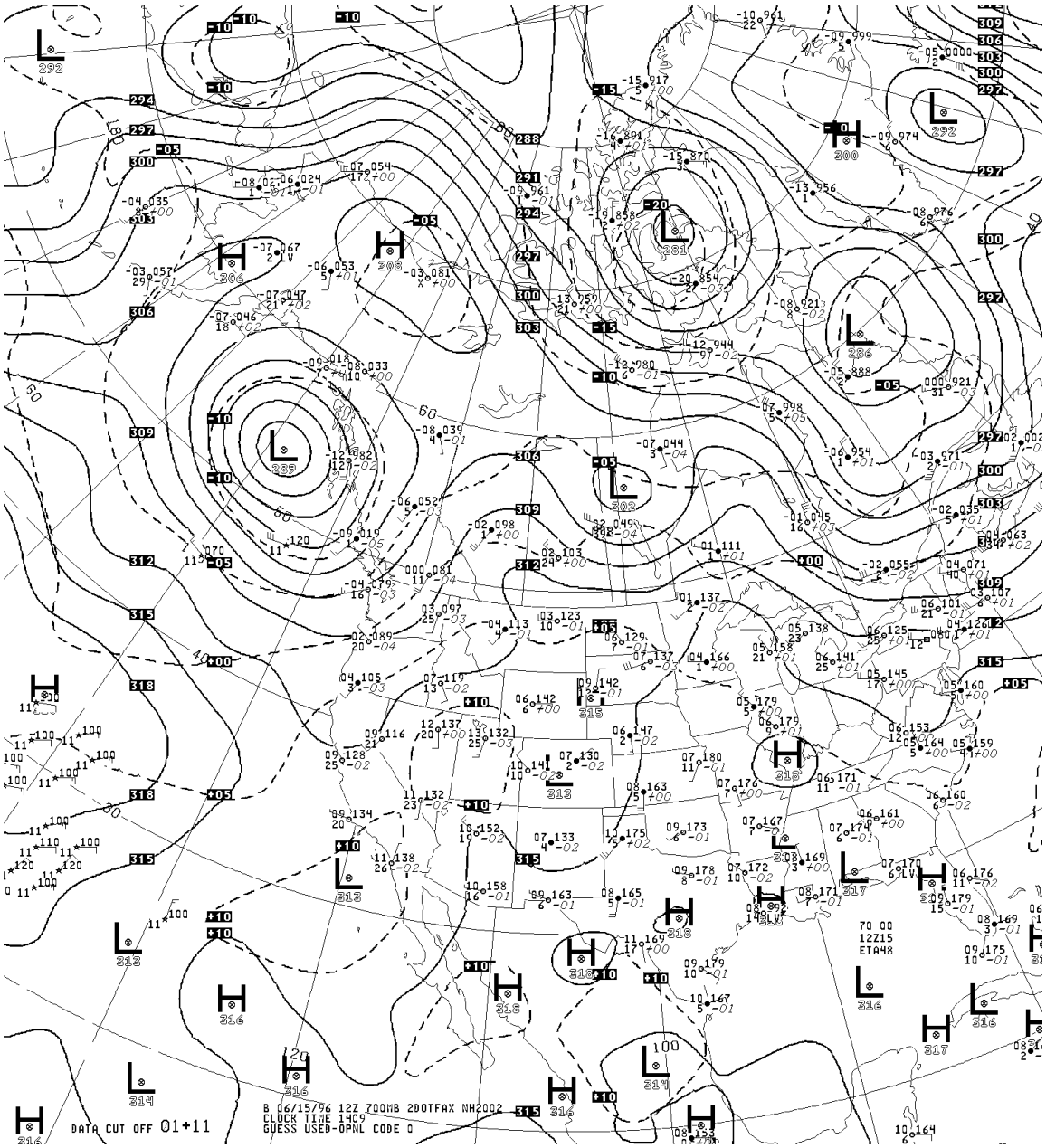


Figure 24. 700 mb analysis of heights and temperature for 15 June 1996 at 1200Z (NCDC, 1996).

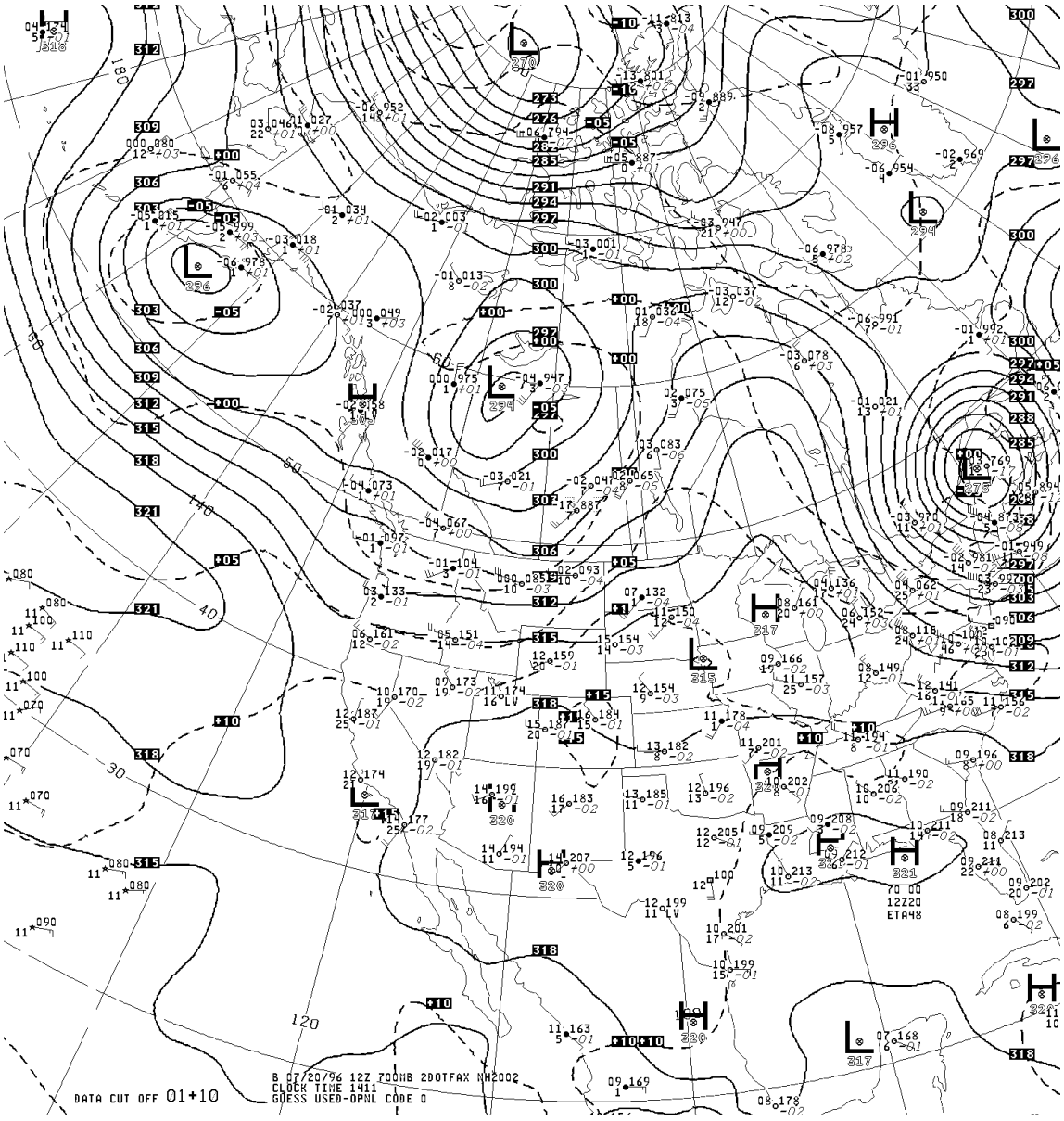


Figure 25. 700 mb analysis of heights and temperature for 20 July 1996 at 1200Z (NCDC, 1996).

VI. Conclusions and Recommendations

6.1 Conclusions

The goal of this research was to explain the possibility of developing a forecast aid that would assist forecasters in determining the optimum time (in particular, a two-week window of winds less than 25 kts) for construction of the MDA's XBR on Shemya. Specific research objectives were carried out in order to accomplish this goal.

North Pacific synoptic conditions were studied in order to understand the various weather features that influence Shemya's winds on an annual basis. An examination of global weather patterns using seasonal 500 mb heights was accomplished. Mean sea level charts were used to illustrate the placement and propagation of various North Pacific pressure centers, to include the Aleutian Low, Asiatic Low, North Pacific High, Asiatic High, and MPH tracks. Tropical storm patterns and climatological wind roses were also studied to gain an understanding of the challenge of forecasting winds on the island.

After gaining an understanding of global and synoptic systems that influence the island's weather; a collection of wind, teleconnection, and sea surface temperature data were then assembled. The wind data were examined to create two-week windows (start days) and a frequency distribution of non-windy days. This information was implemented in the standard, multiple, and Poisson regression analyses, as well as, CART analyses.

Standard regression analysis uncovered a significant relationship between start days and the NAO, with a more positive NAO leading to more windy conditions, which

proved to be the only TI with a significant correlation to the wind data. Multiple regression analyses were also accomplished to examine relationships between all indices and wind data, however, these analyses yielded low R-squared numbers. Since the wind data (start days and non-windy days) indicated an increase in variance over time when compared to the TI data, Poisson regression analyses were performed. Poisson regression only explained about 50% of the variance of the data (pseudo R-squared values), so data mining through CART analysis was undertaken as an additional exploratory tool.

The use of data mining through CART analysis was accomplished in order to extract any useful predictive information from the data that was not uncovered through standard statistical means. Both classification and regression tree analyses were performed, whereas regression trees proved more useful a predictive method than classification trees, which had low tree accuracy. The statistical information, computed by regression tree growth, was used to compute confidence and prediction intervals for both start days and non-windy days. Conducting student "t" tests and investigating improvement scores, showed the start day models to yield the most significant results. The CART information was combined with selected 700 mb charts to assess any relationship between the two.

Shemya is known to have one of the harshest climates with some of the strongest persistent winds on the earth, which presents a great challenge to the construction of the XBR and aircraft operations on the island. The accomplishment of this research, to include an in-depth climatological study of the region, statistical and CART data mining analyses should assist forecasters in alleviating some of the concerns associated with forecasting for this location.

The main findings of the research, however, can be formulated from Figures 20 and 21. By using Figure 20 and the May SOI and PNA values from U.S. CPC, the forecasters can assess the average number of start days for June. For example, with an SOI of less than 0.55 but greater than –1.15 and a PNA value of more than 0.30, the average number of start days is 3.60 for the month of June, from ending node number 6. If the SOI were greater than 0.55, less windy conditions could be forecast with an average of 12.44 start days for the month of June, from ending node number 2 (i.e., the possibility of less windy days than average). Figure 21 is used in the same manner to forecast the number of start days for July using TI and SST data from June.

6.2 Recommendations

This research has shown that the possibility of using teleconnection indices and sea-surface temperatures to long-range forecast atmospheric winds may be possible. It has shown that CART data mining tools can be used as forecast decision aids, when traditional statistical analyses fail to produce predictive results. However, there is still a significant level of uncertainty in long-range forecasting of winds, particularly for such a challenging location as Shemya.

Further research on using CART data mining for atmospheric sciences applications should be pursued and the following are suggestions for future research endeavors:

1. Compare TI and SST data to other operational weather parameters such as precipitation, temperatures, turbulence, and icing to develop an operational long-range forecast model for Shemya.

2. Investigate any significance between the strength and location of main pressure centers and the TIs and SSTs, for example, relationships between the strength and location of the Aleutian Low given specific TI and SST information.
3. Further investigate any relationships between TI and SST data and the movement and strength of MPHs as they transit from polar regions.
4. Develop a forecast decision model that incorporates the findings of this study. An example of such a forecast decision model is presented in Appendix E. It should be noted, that this model is not intended for operational use until extensive field verification is accomplished.
5. Examination of seasonal wind roses for the area indicates a south/southwesterly persistence in wind direction during June and July. The placement of a wind protection device (wind break) to the south/southwest of the construction site could help alleviate concerns about strong winds, or reduce the chances of construction failure.

This research, along with these recommendations, will hopefully become a foothold for future CART data mining studies. By using the forecast decision trees developed in this research (Figures 20 and 21), along with other forecast tools, forecasters may be able to determine the risk of windy conditions on the island. This research will hopefully aid forecasters at the 11th OWS in forecasting windy conditions on Shemya, leading to successful completion of the XBR construction project.

APPENDIX A: ENGINEERING WEATHER DATA (EWD)

This appendix contains the set of engineering weather data compiled by and modified from the Air Force Combat Climatology Center (AFCCC). A brief description of each category of data included in the set follows along with charts created by AFCCC.

CLIMATE SUMMARY

Location Information. This section contains a summary table, which includes site name, location, elevation (above mean sea level), World Meteorological Organization (WMO) number, period of record (POR), and average (atmospheric) pressure not corrected to sea level (higher elevations result in lower pressures). The WMO number is a unique number assigned to every location in the world that takes and transmits regular weather observations. The POR is the time frame over which the data used to compute the statistics in this handbook was compiled.

Design Values. Design values are provided for dry bulb temperature, wet bulb temperature, and humidity ratio at specific percentile frequencies of occurrence.

The old EWD summer design values of 1, 2.5, and 5 percent were based on the warmest four months of the year. In the United States this was standardized as June through September. The new design values of 0.4, 1, and 2 percent are based on the entire year. The old winter design values of 99 and 97.5 percent were based on the three coldest months of the year (December through February). The new winter design values of 97.5, 99.6 and 99 percent are based on the entire year. In other words the new design values are annual values not seasonal values.

In general, for mid-latitude locations with continental climates (hot summer – cold winter), there are some *generalizations* that can be made about the differences between the old and new values. The new 0.4% annual value is comparable to the old 1% seasonal value. The new 1% annual value is usually about a degree cooler than the old 2.5% seasonal value. The new 2% annual value is similar to the old 5% seasonal value. The new 99.6% and 99% annual values are generally cooler than the old 99% and 97.5% seasonal values, however there is more variability between stations.

The new design values were instituted for several reasons. At some locations, the warmest or coldest months of the year do not fall into the months listed above. It is easier to compare locations that are in tropical or marine climates where there is less seasonal variability. It is also more straightforward to compare Southern Hemisphere locations. Finally, this is the same convention used by the American Society of Heating, Refrigeration, and Air Conditioning Engineers (ASHRAE) in their 1997 Handbook of Fundamentals.

Dry Bulb Temperature:

Median of Extreme Highs (or Lows). The dry bulb temperature extreme high (or low) is determined for each calendar year of the POR along with the coincident values for wet bulb temperature, humidity ratio, wind speed, and prevailing wind direction. Median values are determined from the distribution of extreme highs (or lows).

0.4%, 1.0%, 2.0%, 97.5%, 99.0%, and 99.6% Occurrence Design Values. Listed is the dry bulb temperature corresponding to a given annual cumulative frequency of occurrence and its respective mean coincident values for wet bulb temperature, humidity

ratio, wind speed, and prevailing wind direction. This represents the dry bulb threshold, which exceeded its respective percent of time, taking into account the entire POR. For example, the 1.0% occurrence design value temperature has been exceeded only 1% of the time during the entire POR. All the observations occurring within one degree of the design value are grouped, and the coincident mean values for wet bulb temperature, humidity ratio, and wind speed are calculated. The prevailing wind direction (the 'mode' of the wind direction distribution) is also calculated.

Mean Daily Range. The mean daily range (difference between daily maximum and daily minimum temperatures) is the average of all daily dry bulb temperature ranges for the POR.

Wet Bulb Temperature:

Median of Extreme Highs for wet bulb temperature is the highest annual extreme wet bulb temperature averaged over the POR. The corresponding mean coincident values are determined the same way as for dry bulb temperature. 0.4%, 1.0 %, 2.0% occurrence wet bulb temperature design values and the corresponding mean coincident values for dry bulb temperature are determined the same way as for dry bulb temperature.

Humidity Ratio:

Median of Extreme Highs for humidity ratio is the highest annual extreme averaged over the POR. The corresponding mean coincident values are determined the same way as for dry bulb temperature. Design values are provided for "Humidity Ratio"

at the 0.4%, 1.0%, and 2.0% occurrence and the corresponding mean coincident values for dry bulb temperature, vapor pressure, wind speed, and wind prevailing direction.

Air Conditioning/Humid Area Criteria:

These are the number of hours, on average, that dry bulb temperatures of 93 ºF (34 ºC) and 80 ºF (27 ºC) and wet bulb temperatures of 73 ºF (23 ºC) and 67 ºF (19 ºC) are equaled or exceeded during the year.

Other Site Data:

This information is provided *for general reference only, and should NOT be used as the basis for design.* There are some locations for which this data is not available. In these cases, that portion of the table will be left blank.

Weather Region. There are eleven weather regions developed by the Department of Energy. They are defined by the range of cooling-degree days and heating-degree days.

Ventilation Cooling Load Index. The VCLI is a two-part index which defines the total annual cooling load for ventilation air by calculating sensible heat load separately from the latent heat load (moisture). The results are expressed in ton-hours per cubic feet per minute per year of latent and sensible load. Values for sensible heat load are calculated by comparing the outdoor temperature to indoor conditions (75 ºF and 60% relative humidity [RH]), and calculating how much energy is required to bring the outdoor air to the indoor temperature. The latent load is calculated similarly. Separate calculations are

made for each hour of the year, and them summed to form the annual VCLI (Harriman 1997).

Average Annual Freeze-Thaw Cycles. This is simply the average number of times per year that the air temperature first drops below freezing and then rises above freezing, regardless of the duration of either the freezing or thawing. The number of cycles is summed per year, and averaged over the entire POR. Days with high temperatures or low temperatures at 32 °F (0 ºC) are not counted for a freeze-thaw cycle. A cycle is counted only when the temperature drops below freezing (31 °F [-0.5 ºC] or colder) or goes above freezing (33 °F [0.5 ºC] or warmer).

Other Values. The following are derived from sources other than the AFCCC. Engineers and architects should contact the organizations listed below for current values, including background information and complete guidelines for use of these data elements.

Groundwater:

The National Groundwater Educational Foundation

601 Dempsey Road

Westerville OH

(800) 551-7379

Note: Average groundwater temperature parallels long-term average air temperature, because soil at a depth of 50 feet (15 meters) does not undergo significant temperature change over the course of a year. Soil temperature at 50 feet stays slightly

warmer than average annual air temperature by about 2.5 degrees Fahrenheit (1.4 degrees Celsius).

Rain Rate:

International Plumbing Code

BOCA International

4051 West Flossmoor Road

Country Club Hills IL 60478

(708) 799-2300

Frost Depth, Basic Wind Speed, Ground Snow Loads:

ANSI/ASCE 7-95

American Society of Civil Engineers

1015 15th Street NW, Suite 600

Washington DC 20005

(800) 548-2723

Note: Frost depth penetration data was obtained from TI 809-01, Load Assumptions for Buildings (1986) that is published by the Army Corps of Engineers. Wind and snow load data are provided by the American Society of Civil Engineers (1995); where snow load data was not available from ASCE, TI-809-01 (1986) was used. However, since the completion of this project, a new version of TI-809-01 has also been completed. Many of the new snow loads have changed. Current values can be obtained at http://www.hnd.usace.army.mil/techinfo/ti/809-01.pdf.

Suggestions for Use. The dry bulb, wet bulb, and humidity ratio values shown are peak load conditions and are used for sizing mechanical equipment. Design guidance determines the level of occurrence applied. The 0.4% Dry Bulb Temperature value is seldom used for sizing conventional comfort control systems, but is sometimes appropriate for mission-critical systems where equipment failure due to high heat would be unacceptable. Using the 0.4% value for equipment sizing requires that the engineer consider its operation at less-than-peak design conditions. In the past, oversized cooling equipment has been incapable of modulating during the more common range of operating conditions, yielding comfort control problems. Also, over-sized equipment cycles on and off more frequently, increasing maintenance costs and failing to remove enough moisture to maintain humidity control.

Similar cautionary notes apply to the extreme low dry bulb temperature. Heating equipment designed for extreme conditions must be carefully evaluated to ensure that they will modulate properly to maintain
comfort at less extreme outdoor temperatures that occur 99.6% of the hours during the year.

The mean coincident value for humidity at the 0.4% peak dry bulb temperature is not the highest moisture value, and must not be used for design of humidity control systems. The mean coincident value is the arithmetic average of all the moisture levels which occur when the dry bulb temperature is high. However, the highest moisture values typically occur when the dry bulb temperatures are lower.

High wet bulb temperature is used for sizing cooling towers and other evaporative equipment.

Peak humidity ratio is used for sizing dehumidification systems. Peak moisture conditions usually represent a higher enthalpy (total heat) than peak dry bulb conditions. Consequently, engineers use the peak moisture condition to cross-check operation of a system which may be primarily intended to control temperature.

Coincident wind speed allows the engineer to accurately estimate latent loads due to infiltration of humid air in the summer and infiltration of dry air during the winter.

Cautionary Note: The same precautions that apply to heating and cooling equipment also apply to dehumidification and humidification systems. Oversized equipment may not control properly under typical operating conditions without special attention from the engineer.

SHEMYA/EARECKSON AFS AK

Latitude = 52.72 N — WMO No. 704140

Longitude =174.12 E — Elevation = 98 Feet

Period of Record = 1943 TO 1997 — Average Pressure = 29.57 inches Hg

		Mean Coincident (Average) Values			
	Design Value	Wet Bulb Temperature	Humidity Ratio	Wind Speed	Prevailing Direction
Dry Bulb Temperature (T)	**(^oF)**	(^oF)	(gr/lb)	(mph)	(NSEW)
Median of Extreme Highs	57	53	53	12.6	W
0.4% Occurrence	54	51	51	14.1	WSW
1.0% Occurrence	53	50	50	14.5	WSW
2.0% Occurrence	52	50	49	14.8	WSW
Mean Daily Range	5	-	-	-	-
97.5% Occurrence	26	24	14	17.1	N
99.0% Occurrence	24	22	13	16.3	N
99.6% Occurrence	22	20	12	15.2	N
Median of Extreme Lows	19	18	10	13.5	NNW
		Mean Coincident (Average) Values			
	Design Value	Dry Bulb Temperature	Humidity Ratio	Wind Speed	Prevailing Direction
Wet Bulb Temperature (T_{wb})	**(^oF)**	(^oF)	(gr/lb)	(mph)	(NSEW)
Median of Extreme Highs	54	55	57	13.6	W
0.4% Occurrence	52	53	54	14.6	WSW
1.0% Occurrence	51	52	52	14.9	WSW
2.0% Occurrence	50	51	50	15.1	WSW
		Mean Coincident (Average) Values			
	Design Value	Dry Bulb Temperature	Vapor Pressure	Wind Speed	Prevailing Direction
Humidity Ratio (HR)	**(gr/lb)**	(^oF)	(in. Hg)	(mph)	(NSEW)
Median of Extreme Highs	59	54	0.39	14.5	WSW
0.4% Occurrence	56	53	0.38	14.9	SW
1.0% Occurrence	54	52	0.36	15.2	S
2.0% Occurrence	52	51	0.35	15.5	SSW
Air Conditioning/		$T \geq 93^oF$	$T \geq 80^oF$	$T_{wb} \geq 73^oF$	$T_{wb} \geq 67^oF$
Humid Area Criteria	# of Hours	0	0	0	0

Other Site Data

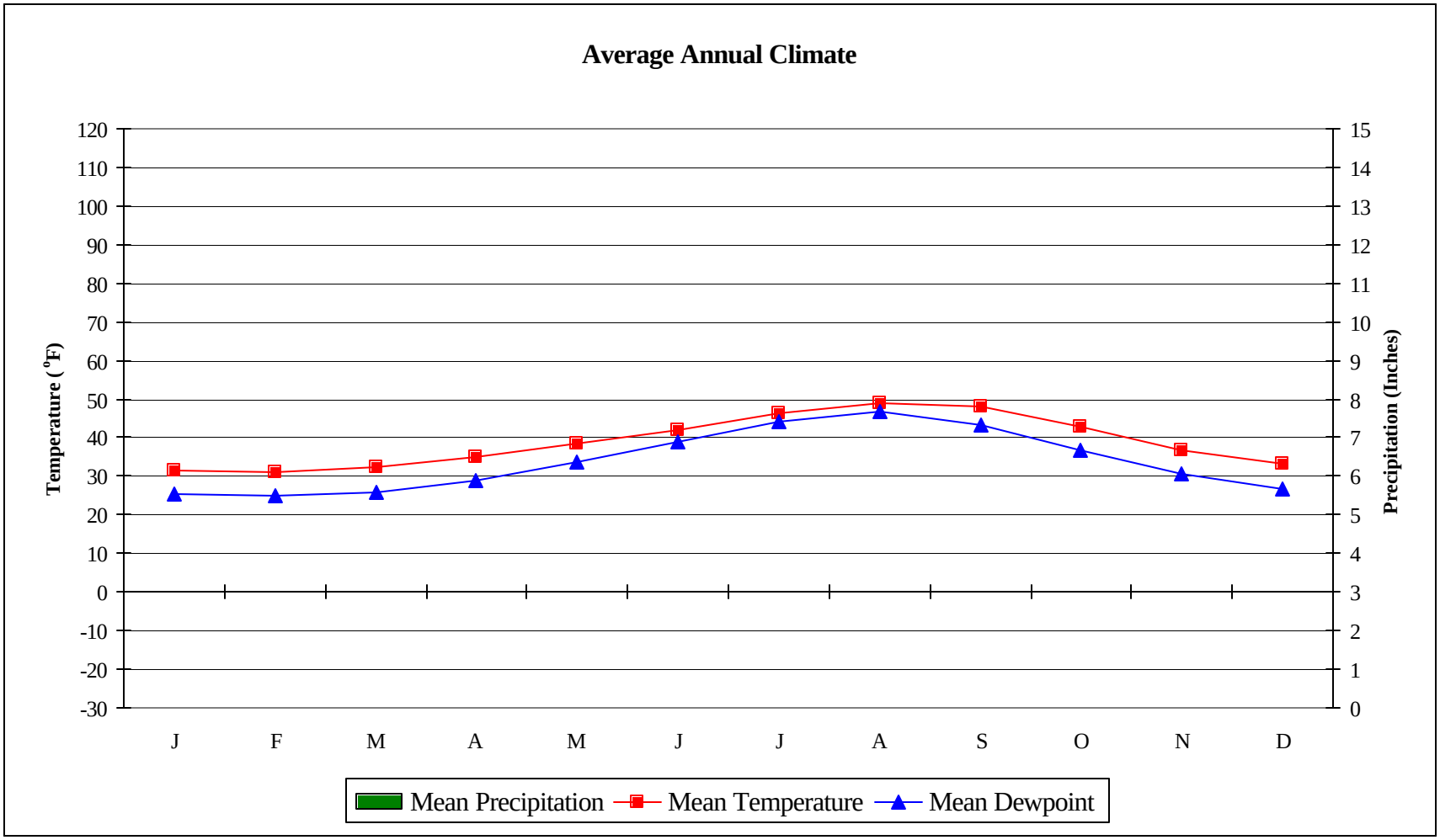
Weather Region	Rain Rate 100 Year Recurrence (in./hr)	Basic Wind Speed 3 sec gust @ 33 ft 50 Year Recurrence (mph)	Ventilation Cooling Load Index (Ton-hr/cfm/yr) Base 75°F-RH 60% Latent + Sensible
5	.	130	0.0 + 0.0
Ground Water Temperature (°F) 50 Foot Depth *	Frost Depth 50 Year Recurrence (in.)	Ground Snow Load 50 Year Recurrence (lb/ft^2)	Average Annual Freeze-Thaw Cycles (#)
41.4	52	20	58

*Note: Temperatures at greater depths can be estimated by adding 1.5°F per 100 feet additional depth.

Average Annual Climate

Explanation of Graph. The graph shows the monthly mean temperature, dewpoint, and precipitation. The bar graph representing precipitation uses the scale on the right side of the chart (inches or centimeters). Lines of temperature and dew point use the scale on the left side of the chart (degrees Fahrenheit or Celsius). These charts have fixed maximum and minimum values on their axes for easy comparison between different sites. The precipitation chart is capped at a maximum of 15 inches (45 centimeters) per month. A few sites may exceed this value; but to keep the graph readable, a fixed maximum value was used. There are a number of sites for which accurate precipitation data was not available. If this is the case, then no bars are printed on the chart.

SHEMYA/EARECKSON AFS AK WMO No. 704140



No Precipitation Data Available

Suggestions for Use:

This graph displays the average behavior of weather over a single year. An architect can compare rainfall patterns at one station with another to evaluate differences in gutter and drain sizing, and also the relative importance of water resistance for the exterior envelope. An engineer can compare the temperature and moisture patterns to understand the relative importance of sensible heat loads vice latent loads at this location.

With averages displayed by month, it is relatively easy to comprehend seasonal variations of each variable, and also understand which specific months are likely to be hot or cold; humid or dry, or have high precipitation. This can be helpful for mission planning, as well as for planning construction and building operation.

Cautionary Note: This graph displays averages, not extreme values. Data shown should not be used to size equipment or building envelopes for peak loads. Peak load data

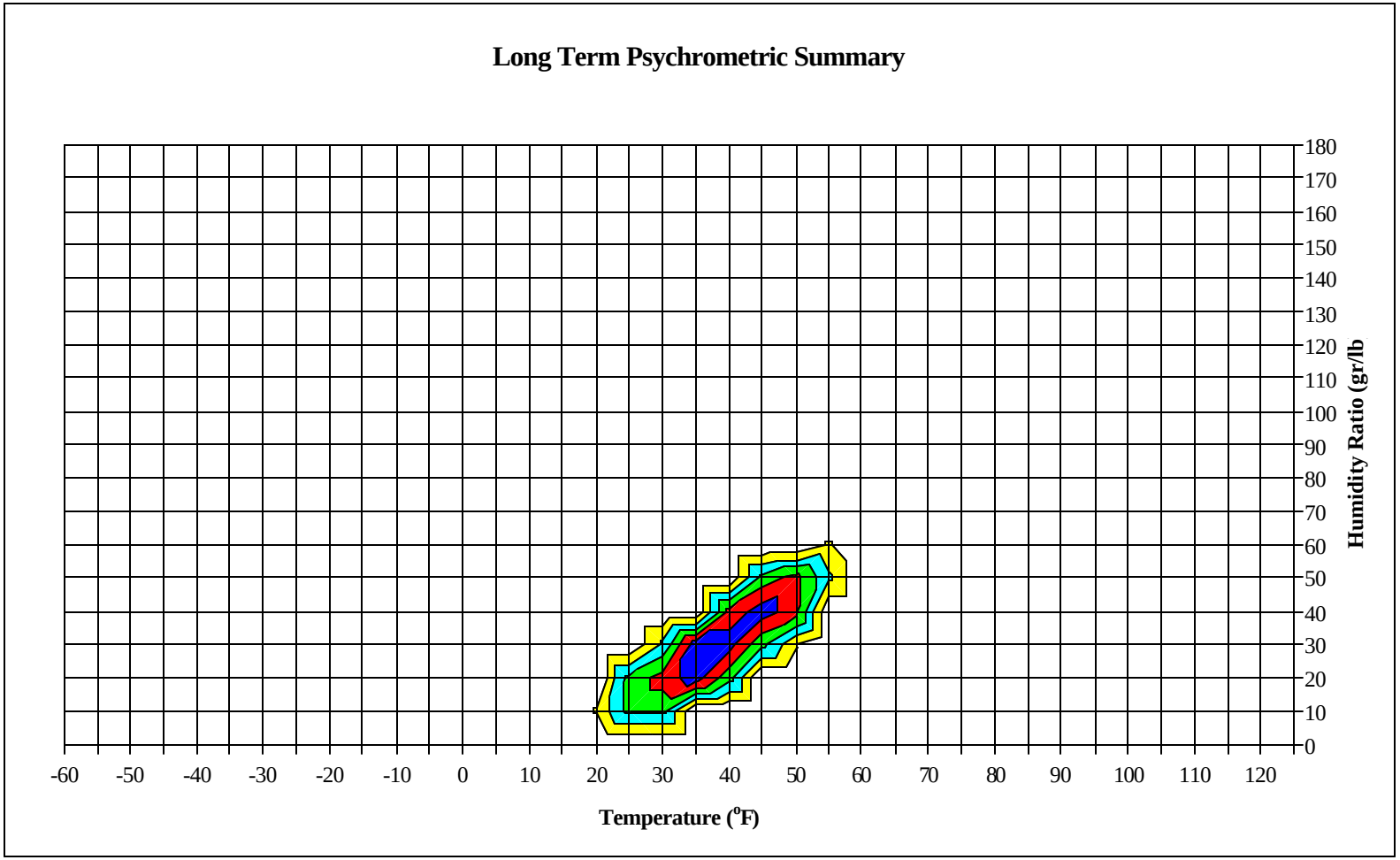
Long- term psychrometric summary

Explanation of Graph:

The graph displays the joint cumulative percent frequency of temperature and humidity ratio. Hourly observations are binned into groups of 5 ◦F and 10 grains per pound (gr/lb) (or 3 ◦C and 1.5 grams per kilogram [g/kg]), centered on each value of temperature or humidity ratio. For example, the 70 ◦F temperature bin collects all observations between 67.5 ◦F and 72.5 ◦F. The bin is depicted as a gridline on the chart; the vertical lines represent the temperature bins and the horizontal lines represent the humidity ratio bins. The intersection of temperature and humidity ratio lines represent a

further subsetting of the observations into groups meeting both temperature and humidity ratio criteria. For example, the intersection of the 70 ◦F bin line and the 40 gr/lb bin line represent the observations when temperature was between 67.5 ◦F and 72.4 ◦F and the humidity ratio was between 35 gr/lb and 44 gr/lb. Thus, a joint-frequency table is created for all temperature and humidity ratio bin combinations.

SHEMYA/EARECKSON AFS AK WMO No. 704140



of all observations
of all observations
of all observations
% of all observations
of all observations

Suggestions for Use:

This graphic displays the long-term history of temperature and moisture at each station (a total of 262,800 hourly observations if the POR is 30 years and if the data is complete over that period). The engineer can use this graph to ascertain the most common temperature and moisture conditions which will be encountered over the operating life of mechanical equipment.

It is often useful to calculate the behavior of the proposed system at "most-common" conditions, in addition to the traditional peak design calculations. This will help ensure that the selected equipment and controls are capable of modulation and control at all points of operation rather than simply at extreme conditions.

Cautionary Note: The psychrometric graph is intended as a visual tool only. Its purpose is to allow quick visual comparison between climates at different locations. Extrapolation of data directly from the graph is not advised due to the approximate plotting routine used to generate the graph from the binned data. This is evident where values of humidity appear past their saturation point. This discrepancy between the actual data and the graph is the result of the plotting routine used to generate the graph and not from errors in the original hourly data used to create the binned summary.

The contours on this chart represent the areas containing 99%, 97.5%, 95%, 80%, and 50% of all observations (cumulative percent frequency or percentiles). The contours are centered on the most frequently occurring bins (50% contour), spreading outward until almost all observations (99%) are grouped. Contours are defined by calculating a percent frequency for each bin (relative to the others), and then accumulating these

percent frequencies (from most frequent to least frequent) until the 50% value is passed, and thus the first set of bins is grouped. The accumulating continues until the 80% value is passed, and the second group of bins is grouped. This continues until the 95%, 97.5%, and 99% values are passed.

Thus, the least frequent (most extreme) bins, which when accumulated amount to less than 1% of the total observations, are outside of the 99% contour. Any bins outside the 99% contour thus have either not occurred, or have occurred so infrequently that they should not be taken into consideration for sizing equipment.

PSYCHROMETRIC DISPLAY OF DESIGN VALUES

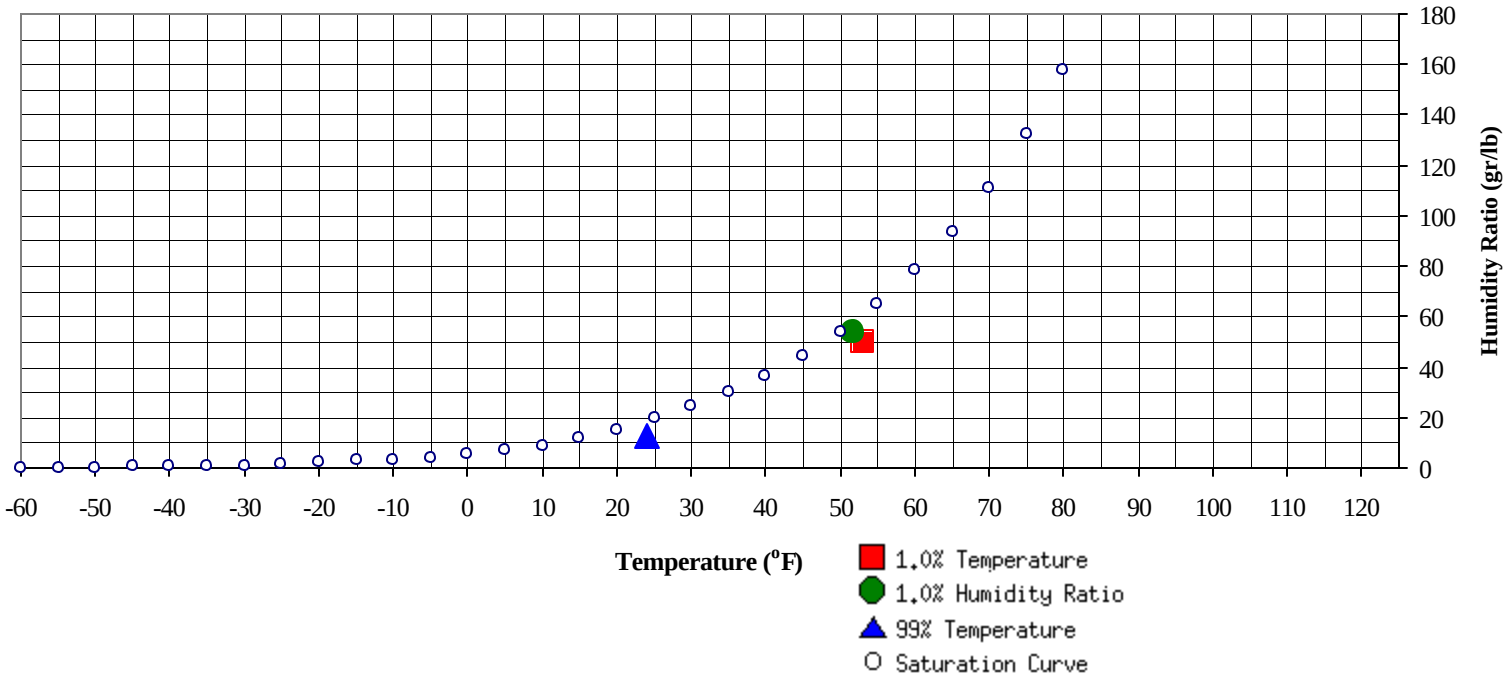
Explanation of Chart. Similar to Page 3, this chart depicts the saturation curve (when RH = 100%) along with peak design values. The design values are calculated as in the table on Page 1, but this chart shows their relationships graphically, depicting their position relative to each other and relative to the saturation curve.

Above and to the left of saturation curve, RH would be greater than 100 percent (not possible). The area below and to the right of the curve (including the points on the curve itself) represent the area where RH is less than or equal to 100 percent, and thus where all observations occur. Note that since the humidity ratio is a function of pressure, and pressure varies with elevation, different sites will have different saturation curves. The dry bulb temperature is the horizontal coordinate on this scatter plot, and the humidity ratio is the vertical coordinate. Peak design values are depicted by the red square (1.0% Dry Bulb Temperature), the green circle (1.0% Humidity Ratio), and the blue diamond (99% Dry Bulb Temperature).

The table below the chart shows the exact values of 99% dry bulb temp, 1.0% humidity ratio, and 1.0% dry bulb temperature, along with calculated values of enthalpy, mean coincident wet bulb temperature, and humidity ratio (as applicable). The value of enthalpy coincident to each temperature/humidity ratio is created using the psychometric functions provided by the Linric Company, Bedford, New Hampshire. The dry bulb temperature and humidity ratios are used to calculate enthalpy using the Linric algorithms.

SHEMYA/EARECKSON AFS AK WMO No. 704140

Psychrometric Summary of Peak Design Values



	(°F)	MCHR (gr/lb)	Enthalpy (btu/lb)		(gr/lb)	MCDB (°F)	MCWB (°F)	MC Dewpt (°F)	Enthalpy (btu/lb)
99% Dry Bulb	24	12.7	7.7	**1.0% Humidity Ratio**	53.9	51.7	50.8	50	20.8

		MCHR	MCWB	Enthalpy
	(°F)	(gr/lb)	(°F)	(btu/lb)
1.0% Dry Bulb	53	49.8	50.5	20.4

BINNED TEMPERATURE DATA

Explanation of Tables. Identical to those in AFM 88-29, these tables show the number of hours that temperatures in 5 °F (3 °C) bins occur during a given month, and during 8-hour periods during the days of that month. The 8-hour periods are based upon a 24-hour clock and displayed in Local Standard Time (LST). The total numbers of observations (hours) in each temperature bin are summed horizontally in the "Total Obs" column for the month. The mean coincident wet bulb temperature is the mean value of all those wet bulb temperatures that occur coincidentally with the dry bulb temperatures in the particular 5-degree temperature interval. At the upper or warmer end of the mean coincident wet bulb distribution, the values occasionally reverse their trend because the highest wet bulb temperatures do not necessarily occur with the highest dry bulb temperatures. There are thirteen such tables, one for each month, and one representing the overall annual summary.

SHEMYA/EARECKSON AFS AK WMO No. 704140

Dry-Bulb Temperature Hours For An Average Year (Sheet 1 of 5)

Period of Record = 1943 TO 1997

Temperature Range (°F)	January Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	February Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	March Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)
65 / 69															
60 / 64															
55 / 59															
50 / 54															
45 / 49							0		0	42.0		0	0	0	39.9
40 / 44	2	3	3	8	39.5	1	3	1	5	38.4	0	4	1	5	37.9
35 / 39	53	62	59	174	34.9	38	52	46	136	34.3	53	86	75	214	34.4
30 / 34	111	114	111	336	30.1	108	107	108	322	30.1	126	119	126	371	30.1
25 / 29	67	59	62	188	25.1	61	52	56	168	25.0	58	34	41	133	25.3
20 / 24	14	10	12	36	21.0	16	10	12	38	20.7	10	5	5	20	20.6
15 / 19	1	0	1	2	16.5	1	1	1	3	16.1	1	0	0	1	16.0
10 / 14	0	0	0	0	11.7	0		0	0	11.2	0	0		0	10.5
5 / 9															

SHEMYA/EARECKSON AFS AK WMO No. 704140

Dry-Bulb Temperature Hours For An Average Year (Sheet 2 of 5)

Period of Record = 1943 TO 1997

Temperature Range (°F)	April Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	May Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	June Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)
65 / 69															
60 / 64															
55 / 59												0	0	0	50.1
50 / 54							0		0	44.8	0	4	2	6	47.4
45 / 49		0	0	0	41.1	0	4	1	5	42.2	9	56	42	107	44.0
40 / 44	2	22	11	35	38.2	29	113	87	229	39.1	166	162	173	501	40.9
35 / 39	108	149	144	401	34.6	193	124	149	466	35.9	64	18	23	105	37.5
30 / 34	116	65	80	261	30.4	25	7	11	43	31.2	1	0	0	1	32.6
25 / 29	13	4	5	22	25.7	1	0	0	1	25.9	0			0	27.8
20 / 24	1	0	0	1	21.0	0			0	23.0					
15 / 19	0			0	16.5										
10 / 14															
5 / 9															

SHEMYA/EARECKSON AFS AK WMO No. 704140

Dry-Bulb Temperature Hours For An Average Year (Sheet 3 of 5)

Period of Record = 1943 TO 1997

Temperature Range (°F)	July Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	August Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	September Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)
65 / 69					58.0										
60 / 64		0	0	0	55.7	0	1	0	1	55.5		0		0	53.0
55 / 59		4	2	6	52.3	1	12	7	20	52.6	0	4	1	5	51.5
50 / 54	7	41	29	77	49.2	51	120	103	274	49.6	29	95	66	190	48.7
45 / 49	142	166	172	480	45.7	187	114	136	437	46.9	168	129	149	446	45.5
40 / 44	98	37	45	180	42.8	9	1	2	12	43.0	40	12	23	75	40.9
35 / 39	1	0	0	1	38.3	0			0	39.0	3	0	1	4	36.6
30 / 34											0			0	32.3
25 / 29															
20 / 24															
15 / 19															
10 / 14															
5 / 9															

SHEMYA/EARECKSON AFS AK WMO No. 704140

Dry-Bulb Temperature Hours For An Average Year (Sheet 4 of 5)

Period of Record = 1943 TO 1997

Temperature Range (°F)	October Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	November Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)	December Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)
65 / 69															
60 / 64															
55 / 59															
50 / 54	1	7	2	10	47.6		0		0	42.0					
45 / 49	60	98	75	234	43.9	3	5	3	11	43.6	0	0	0	0	43.9
40 / 44	129	111	123	363	39.7	48	64	53	165	39.5	8	10	7	25	39.4
35 / 39	52	29	42	123	35.1	111	110	112	333	34.6	79	90	84	253	34.7
30 / 34	6	3	5	14	31.4	65	53	62	180	30.2	112	110	113	335	30.0
25 / 29	0		0	0	27.3	11	7	10	28	25.7	42	33	38	113	25.4
20 / 24						1	1	0	2	21.1	6	3	5	14	20.8
15 / 19						0			0	15.9	1	1	1	3	16.4
10 / 14											0	0	0	0	10.5
5 / 9											0	0		0	6.3

SHEMYA/EARECKSON AFS AK WMO No. 704140
Dry-Bulb Temperature Hours For An Average Year (Sheet 5 of 5)
Period of Record = 1943 TO 1997

Annual Totals

Temperature Range (°F)	Hour Group (LST) 01 To 08	09 To 16	17 To 00	Total Obs	MCWB (°F)
65 / 69					
60 / 64	0	1	0	1	55.3
55 / 59	1	21	10	32	52.3
50 / 54	87	265	200	552	49.2
45 / 49	569	571	575	1715	45.6
40 / 44	533	543	529	1605	40.3
35 / 39	757	720	735	2212	35.0
30 / 34	670	579	619	1868	30.2
25 / 29	252	189	214	655	25.2
20 / 24	48	29	35	112	20.8
15 / 19	4	2	3	9	16.2
10 / 14	0	0	0	0	11.0
5 / 9	0	0		0	6.3

Caution: This summary reflects the typical distribution of temperature in a typical year. It does not reflect the typical moisture distribution. Because wet bulb temperatures are averaged, this summary understates the annual moisture load. For accurate moisture load data, see the long-term humidity summary and the ventilation and infiltration load pages in this manual.

Suggestions for Use. Binned summaries are used by many different technical disciplines for different purposes. They are useful in making informal estimates of energy consumption by cooling and heating equipment, and for gaining a general understanding of patterns of temperature and moisture at different times of the day, month, and year.

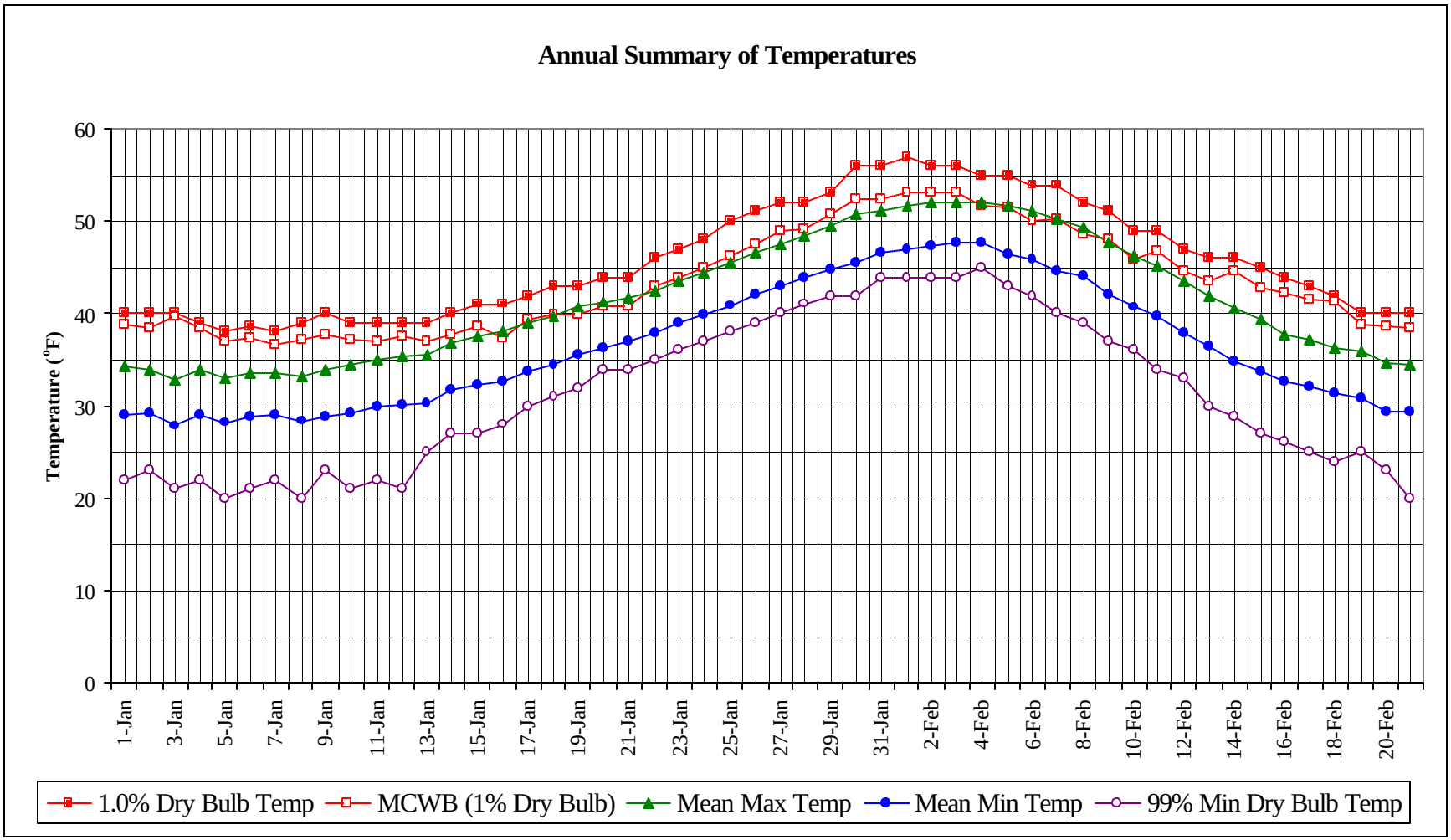
Cautionary Note: Do not use these binned summaries to calculate moisture loads. These particular summaries are based on the dry bulb temperature. After each of the one-hour observations has been placed into a dry bulb BIN, the average humidity ratio is calculated for all observations in each BIN. Consequently, dry bulb BINs underestimate

the magnitude of dehumidification and humidification loads, because the averaging calculation “flattens” the peaks and valleys of humidity ratios. The amount of the underestimation varies according to the desired humidity control level.

ANNUAL TEMPERATURE SUMMARY

Explanation of Chart. This chart shows a week-by-week summary of dry bulb temperatures for the given site. The observations are grouped into seven-day periods (approximate calendar weeks). For example, observations from 1-7 January from all years are grouped, 8-14 January are grouped, and so on, overlapping the end of one month and beginning of the next month where necessary. For each of the seven-day periods, the following statistics are shown. *1% Temperature* is the dry bulb temperature that is exceeded one percent of the time during that calendar week. *MCWB @ 1% Temp* is the mean of wet bulb temperatures coincident with 1% dry bulb temperatures during the same week *Mean Max Temp* is the daily maximum dry bulb temperature, averaged by week over the POR. *Mean Min Temp* is the daily minimum dry bulb temperature, averaged by week over the POR. *99% Temp* is the daily dry bulb temperature that is at or above this value 99 percent of the time, or below this value one percent of the time.

SHEMYA/EARECKSON AFS AK WMO No. 704140



Note: The information in this chart is calculated on a weekly basis; information on a climate summary (Data Set Page 1) is calculated on an annual basis.

Suggestions for Use. The weekly 1% and 99% temperatures are useful for understanding the probable temperature extremes that can occur during a given week of the year. The weekly dry bulb temperatures are useful for understanding the change of seasons at a given location. The display is helpful for mission planning and construction project planning.

Cautionary Notes:

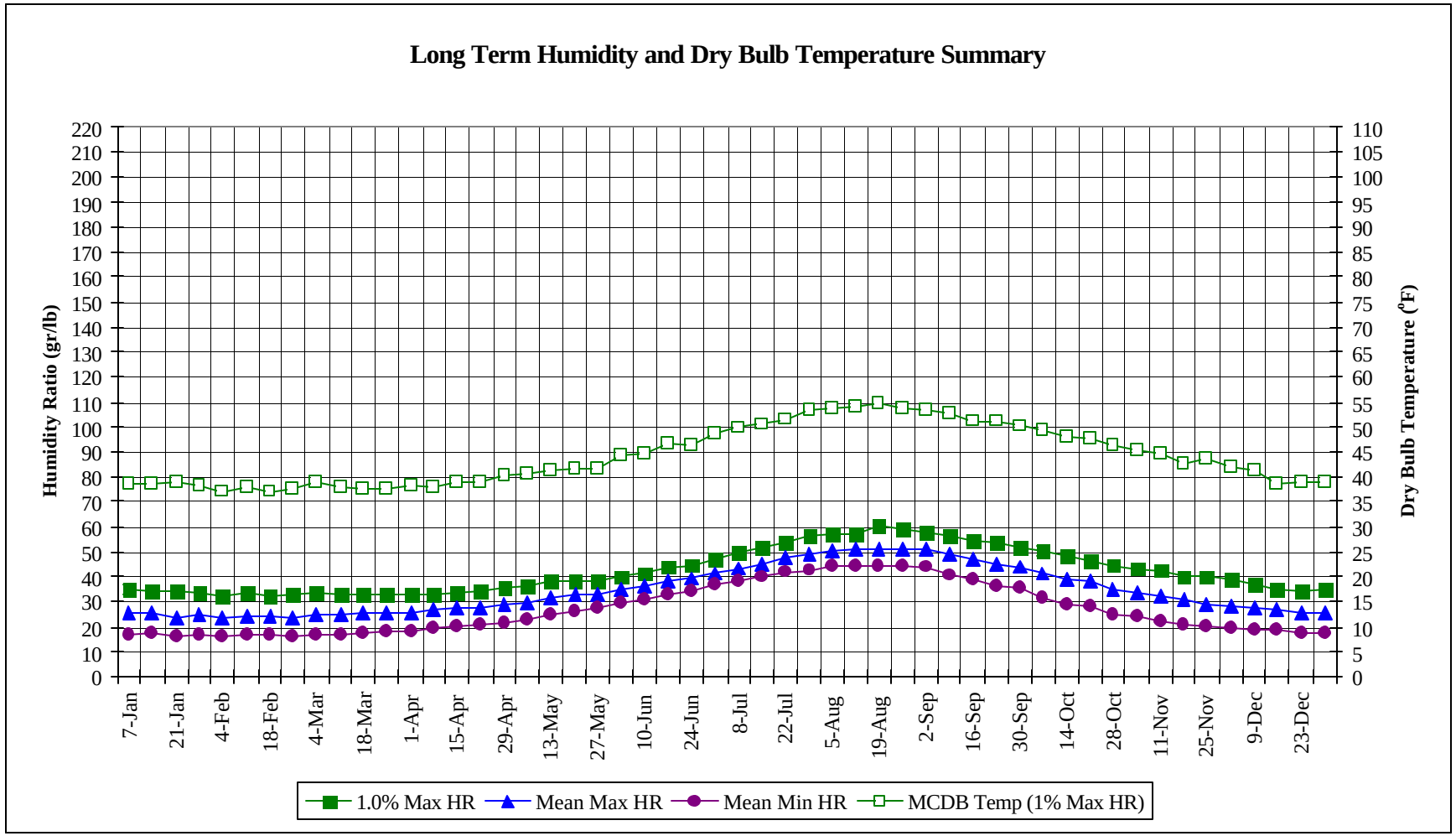
Designers. The values displayed here are based on the 30-year record. It is important that designers NOT base equipment selection on the "highest" or "lowest" recorded temperature at the station. That error would result in selecting equipment extremely costly to install, which would operate inefficiently for all but the very hottest or coldest single hour in 30 years. See the design criteria data page in this handbook for appropriate maximum and minimum temperatures for sizing equipment.

Construction and Operation Planners. The mean maximum and minimum temperatures shown for each week seldom occur in the same year. Keep in mind these are mean values useful for understanding the typical range of temperatures in a given week. The difference does NOT represent the actual day-night temperature swing in a given week.

ANNUAL HUMIDITY SUMMARY

Explanation of Chart. Similar to the annual temperature summary (Data Set Page 10), this chart depicts mean maximum and minimum values of humidity ratio, plus the 1% maximum humidity ratio, along with its mean coincident dry bulb temperature, summarized by calendar week. The chart uses two vertical axes: On the left are the humidity ratio values and on the right is a temperature scale for the mean coincident dry bulb temperature.

SHEMYA/EARECKSON AFS AK WMO No. 704140



Suggestions for Use. Weekly humidity ratios are useful for understanding the change of seasons at a given location, and the probable high and low moisture levels during a given week of the year. The display is helpful for planning humidity-controlled storage projects, and for understanding factors contributing to atmospheric corrosion. Humidity also affects the deterioration rate of building materials and weathering of military equipment and structures exposed to the elements.

Cautionary Notes:

Designers. The values displayed here are based on the 30-year record. It is important that designers NOT base equipment selection on the "highest" or "lowest" recorded humidity at the station. That error would result in selecting oversized equipment, which would increase costs and may result in control problems at other than extreme conditions. Use design values on Data Set Page 1 for equipment sizing.

Construction and Operation Planners. The high and low humidity ratios shown for each week seldom occur in the same year. Keep in mind that these are mean values that are useful for understanding the typical range of humidity ratio in a given week. The difference does NOT represent the actual day-night humidity ratio swing in a given week.

ANNUAL DRY BULB TEMPERATURE AND HUMIDITY SUMMARY

1.22. Explanation of Tables. These tables show the values used to plot the charts on Data Set Pages 10 and 11. The left half of the table uses Data Set Page 10 and the right half uses Data Set Page 11.

SHEMYA/EARECKSON AFS AK WMO No. 704140

Long Term Dry Bulb Temperature and Humidity Summary

Week Ending	1.0% Temp (°F)	MCWB @ 1% Temp (°F)	Mean Max Temp (°F)	Mean Min Temp (°F)	99% Temp (°F)	1.0% HR (gr/lb)	MCDB @ 1% HR (°F)	Mean Max HR (gr/lb)	Mean Min HR (gr/lb)
7-Jan	40.0	38.7	34.3	29.1	22.0	35.0	38.8	25.6	17.4
14-Jan	40.0	38.4	34.0	29.3	23.0	34.3	38.8	25.4	17.8
21-Jan	40.0	39.6	32.9	27.9	21.0	34.3	39.0	23.7	16.2
28-Jan	39.0	38.4	34.0	29.2	22.0	33.6	38.3	25.1	17.4
4-Feb	38.0	37.0	33.0	28.3	20.0	32.2	37.2	23.7	16.6
11-Feb	38.5	37.4	33.6	29.0	21.0	33.6	38.1	24.6	17.2
18-Feb	38.0	36.5	33.6	29.1	22.0	32.2	37.1	24.3	17.1
25-Feb	39.0	37.1	33.2	28.4	20.0	32.9	37.7	23.7	16.1
4-Mar	40.0	37.8	34.0	28.9	23.0	33.6	39.1	24.8	17.2
11-Mar	39.0	37.1	34.5	29.3	21.0	32.9	38.0	25.0	17.0
18-Mar	39.0	36.9	35.0	30.0	22.0	32.9	37.7	25.4	17.9
25-Mar	39.0	37.5	35.3	30.2	21.0	32.9	37.7	25.5	18.4
1-Apr	39.0	37.0	35.6	30.3	25.0	32.9	38.4	25.7	18.1
8-Apr	40.0	37.7	36.8	31.9	27.0	32.9	37.9	27.0	19.9
15-Apr	41.0	38.6	37.6	32.3	27.0	33.6	39.0	27.6	20.5
22-Apr	41.0	37.3	38.1	32.8	28.0	34.3	39.2	27.9	20.8
29-Apr	42.0	39.2	38.8	33.7	30.0	35.7	40.4	28.8	22.0
6-May	43.0	39.8	39.7	34.5	31.0	36.4	40.6	29.8	23.1
13-May	43.0	39.9	40.7	35.6	32.0	37.8	41.4	31.5	25.0
20-May	44.0	40.7	41.2	36.3	34.0	37.8	41.7	32.6	26.6
27-May	44.0	40.8	41.8	37.0	34.0	37.8	41.7	33.1	27.4
3-Jun	46.0	43.0	42.6	37.9	35.0	39.9	44.4	35.0	29.3
10-Jun	47.0	43.9	43.6	39.0	36.0	41.3	44.7	36.2	30.7
17-Jun	48.0	45.0	44.6	39.9	37.0	44.1	46.7	38.3	33.0
24-Jun	50.0	46.2	45.5	40.8	38.0	44.8	46.2	39.5	34.2
1-Jul	51.0	47.5	46.6	42.1	39.0	47.6	48.6	41.7	36.6
8-Jul	52.0	48.9	47.5	43.0	40.0	49.7	50.0	43.6	38.2
15-Jul	52.0	49.1	48.4	44.0	41.0	51.8	50.5	45.4	40.2
22-Jul	53.0	50.7	49.5	44.9	42.0	53.9	51.4	47.8	42.1
29-Jul	56.0	52.4	50.7	45.6	42.0	56.7	53.7	49.1	43.0
5-Aug	56.0	52.3	51.1	46.6	44.0	57.4	53.8	50.9	45.0
12-Aug	57.0	53.1	51.5	47.0	44.0	57.4	54.0	51.1	44.9
19-Aug	56.0	53.0	51.9	47.4	44.0	60.2	55.0	51.3	44.7
26-Aug	56.0	53.0	52.1	47.6	44.0	58.8	53.8	51.4	44.6
2-Sep	55.0	51.6	52.1	47.6	45.0	58.1	53.6	50.9	43.9
9-Sep	55.0	51.4	51.5	46.5	43.0	56.7	52.9	49.0	40.9
16-Sep	54.0	50.0	51.2	45.9	42.0	54.6	51.2	47.4	38.9
23-Sep	54.0	50.2	50.2	44.7	40.0	53.9	51.2	45.1	35.8
30-Sep	52.0	48.5	49.3	44.0	39.0	51.8	50.1	44.0	35.2
7-Oct	51.0	48.0	47.7	42.1	37.0	50.4	49.2	41.5	31.8
14-Oct	49.0	46.0	46.3	40.7	36.0	48.3	47.9	39.0	29.1
21-Oct	49.0	46.7	45.2	39.6	34.0	46.9	47.5	38.1	28.2
28-Oct	47.0	44.6	43.5	37.9	33.0	44.8	46.3	35.1	25.3
4-Nov	46.0	43.6	42.0	36.5	30.0	43.4	45.2	33.7	24.3
11-Nov	46.0	44.6	40.5	34.8	29.0	42.7	44.5	32.4	22.3
18-Nov	45.0	42.9	39.2	33.8	27.0	39.9	42.8	30.8	21.3
25-Nov	44.0	42.4	37.7	32.7	26.0	39.9	43.6	29.2	20.2
2-Dec	43.0	41.7	37.1	32.1	25.0	38.5	42.0	28.3	19.6
9-Dec	42.0	41.4	36.2	31.4	24.0	37.1	41.5	27.4	18.9
16-Dec	40.0	38.8	35.8	30.9	25.0	35.0	38.8	27.2	18.8
23-Dec	40.0	38.5	34.7	29.5	23.0	34.3	39.0	25.9	17.7
31-Dec	40.0	38.4	34.4	29.4	20.0	35.0	39.2	25.8	17.8

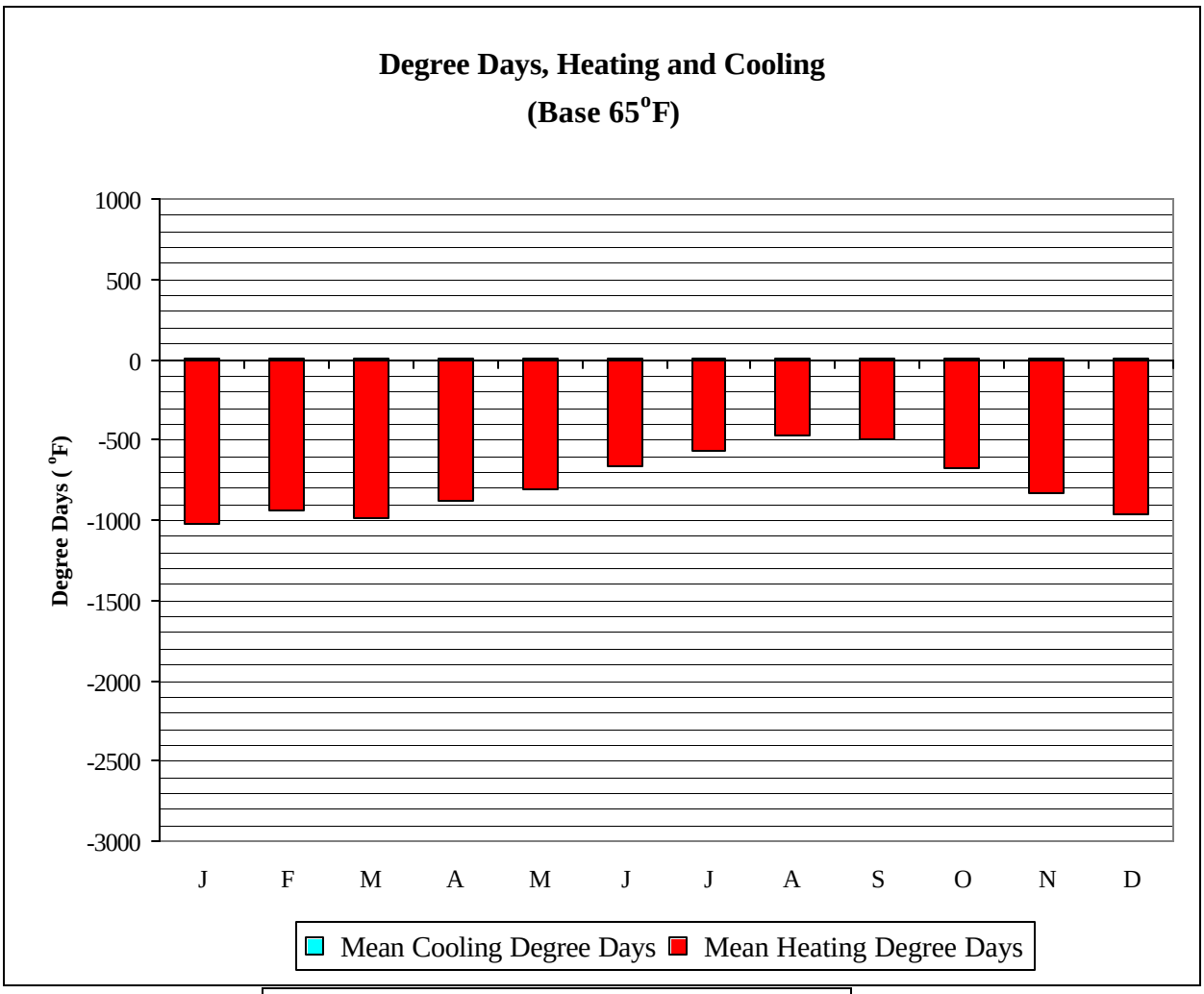
BUILDING ENVELOPE LOADS

Explanation of Charts:

Cooling degree-days are derived by multiplying the number of hours that the outdoor temperature is above 65 °F (18 ºC) times the number of degrees of that temperature difference. For example, if one hour was observed at a temperature of 78 °F, that observation adds 13 degree-hours to the annual total. The sum of the degree-hours is divided by 24 to yield degree-days.

Heating degree-days are calculated similarly, against an inside temperature of 65 °F. So a one-hour observation of 62 °F adds 3 degree-hours to the annual total. Heating degree-days are summed separately from the cooling degree-days. Hot and cold hours do not cancel each other out, as both heating and cooling conditions may occur over the course of a given day.

SHEMYA/EARECKSON AFS AK WMO No. 704140



	Mean Cooling Degree Days (°F)	Mean Heating Degree Days (°F)
JAN	0	1025
FEB	0	941
MAR	0	991
APR	0	880
MAY	0	809
JUN	0	665
JUL	0	567
AUG	0	474
SEP	0	495
OCT	0	679
NOV	0	832
DEC	0	969
ANN	0	9327

Suggestions for Use. Degree-days are used to estimate the sensible heat and sensible cooling loads on the building envelope. Degree-day loads can be used to estimate the annual energy consumption of a building, provided that the loads from ventilation and infiltration air are also considered (see next section).

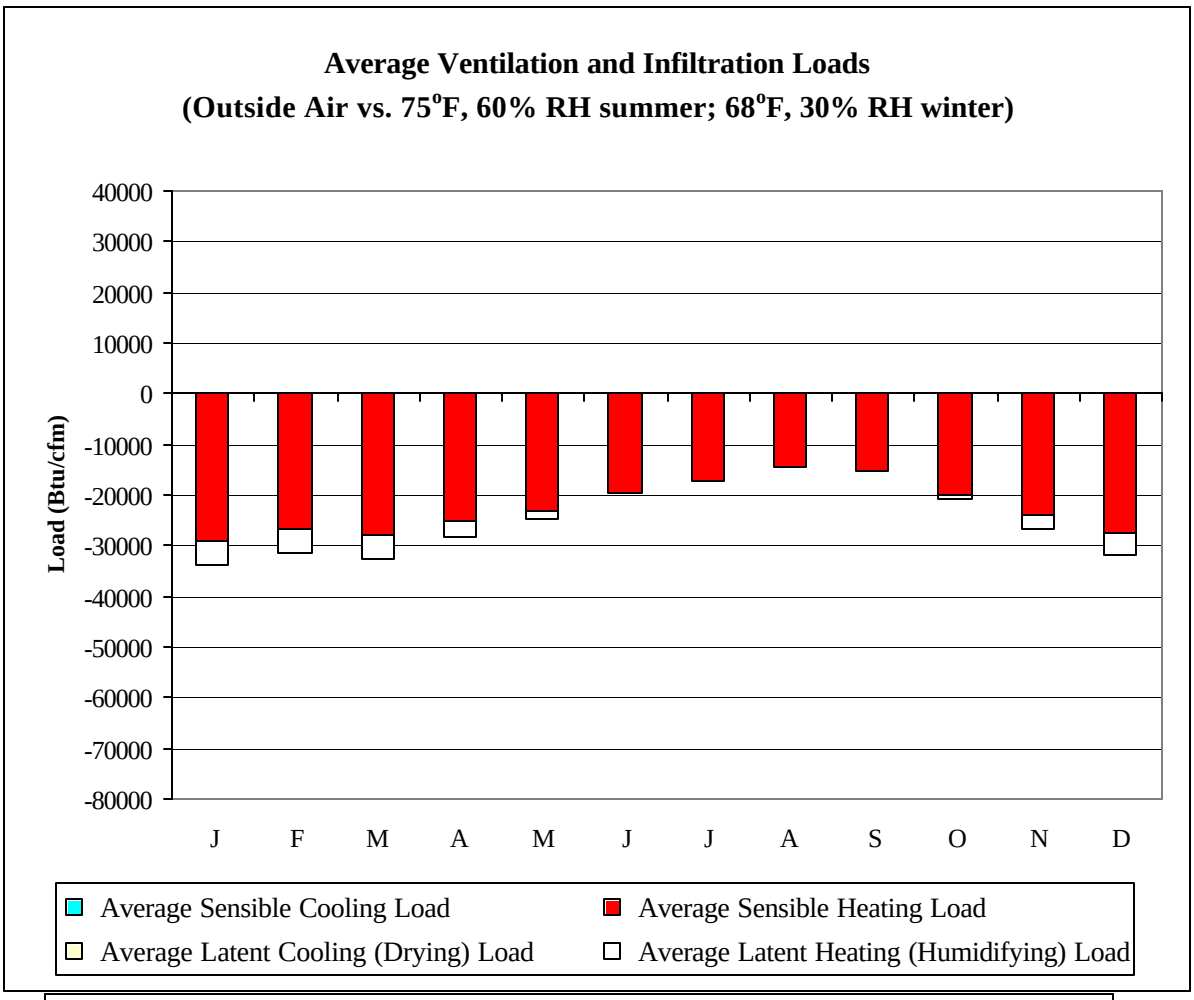
VENTILATION AND INFILTRATION LOADS

Explanation of Charts:

The graph and table display the independent loads imposed by heating, cooling, humidifying, and dehumidifying outside air as it is brought into a building. The calculation assumes that air inside the building is maintained at 68 ºF (20 ºC)/30% RH during the winter and 75 ºF (24 ºC)/60% RH during the summer. For the purposes of these calculations, when the outside air is within that range of temperature and moisture, any incoming air is assumed not to impose any load.

These values are calculated with the methodology used to calculate the annual VCLI Index on page one, except that values on this page are computed by month, and the result is displayed as British thermal units (Btu) per cubic foot per minute (cfm) rather than as ton-hours per cfm per year. The heating and humidifying loads are shown as negative values. Cooling and dehumidifying loads are displayed as positive values.

SHEMYA/EARECKSON AFS AK WMO No. 704140



	Average Sensible Cooling Load (Btu/cfm)	Average Sensible Heating Load (Btu/cfm)	Average Latent Cooling Load (Btu/cfm)	Average Latent Heating Load (Btu/cfm)
JAN	0	-28941	0	-5008
FEB	0	-26563	0	-4757
MAR	0	-28051	0	-4562
APR	0	-25104	0	-3121
MAY	0	-23325	0	-1372
JUN	0	-19493	0	-114
JUL	0	-17065	0	0
AUG	0	-14607	0	-1
SEP	0	-15075	0	-57
OCT	0	-19966	0	-883
NOV	0	-23861	0	-2793
DEC	0	-27475	0	-4294
ANN	0	-269526	0	-26962

Suggestions for Use. Bringing fresh ventilation air into a building, or allowing air to infiltrate into buildings through cracks imposes heating, cooling, dehumidification, and humidification loads on the mechanical system. This display helps the architect, engineers, and operating personnel understand the nature and magnitude of those loads on an annual basis. It also shows how the loads vary from month to month throughout the year.

Comments. These calculations are based on the load created when one cubic foot of fresh air is brought into the building each minute. The results of the calculation include the moisture load or deficit, and the sensible heat load or deficit created by that cubic foot of air during each month of the year. Note that most months have both a load and a deficit for temperature and moisture. The monthly deficit and load do not "cancel" from the perspective of the mechanical system, because temperature and moisture loads will often occur at different times of the day.

Cautionary Note: The values displayed here assume that the inside air is maintained at 68 °F/30% RH during the winter and at 75 °F, 60% RH during the summer. If the inside conditions are held in a different range of temperature or moisture, the loads will be different. For example, in calculating loads for humidity-controlled, but unheated storage, the loads vary according to the change in both temperature and humidity, since the inside temperature varies, but the inside humidity is held constant. For estimating loads in that or similar applications, the engineer may obtain better results from using the average maximum weekly humidity data shown on sample pages 11 and 12.

WIND SUMMARY

Explanation of Charts:

These charts depict the frequency of different wind direction and wind speed combinations. The observations are binned into the sixteen cardinal compass directions and five speed categories (1-5knots, 6-14 knots, 15-24 knots, 25-34 knots, and greater than 34 knots). The frequency of direction and the tick marks indicating that value lie along each 'spoke' of the wind chart. The wind speed bins for each direction are color-coded by the legend at the bottom of the chart.

To determine the percent frequency of a particular wind direction, look for the tick mark bounding the outer edge of a colored (wind speed) area. In the case of the first wind speed bin (1-5 knots), the percent frequency is simply the value of the tick mark on the outer edge of the 1-5 knot region. For the higher speed bins (6-14 knots or greater), subtract the earlier spoke values from the value shown to get the frequency for the speed bin in question.

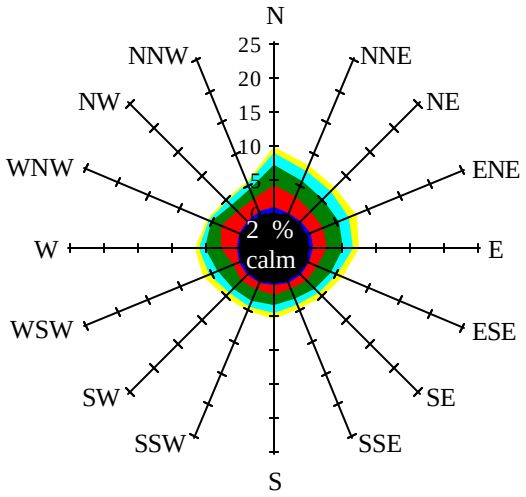
The values for percent frequency have been summed by direction, so to determine the total percent frequency for all speeds from a particular direction, look up the tick mark (or interpolated value) bounding the outermost colored area along that spoke. That tick mark represents the total percent frequency of wind from that direction.

Since the calm condition has no direction, the percent occurrence of calm conditions is displayed immediately below the chart.

Wind Summary - December, January, and February

Labels of Percent Frequency on North Axis

Variable Wind: 0 %



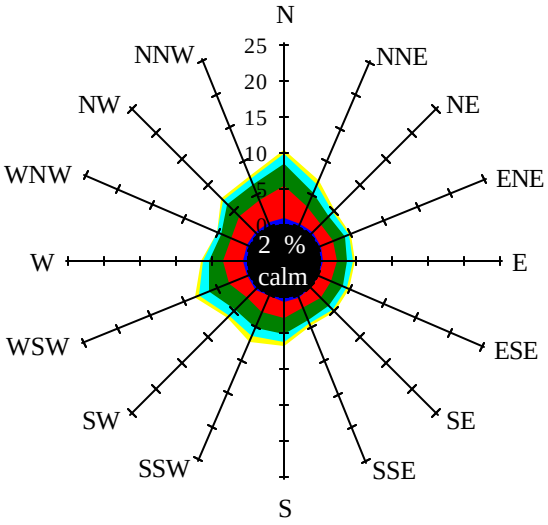
# of Observations: 0

>34 Knots | 25-34 Knots | 15-24 Knots | 6-14 Knots | 1-5 Knots | Calm

Wind Summary - March, April, and May

Labels of Percent Frequency on North Axis

Variable Wind: 0 %



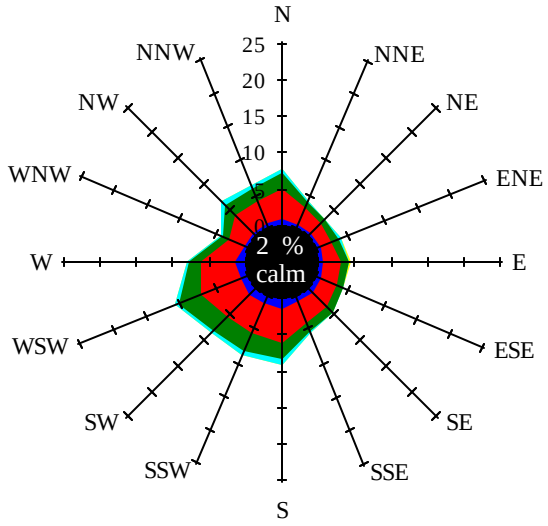
# of Observations: 0

>34 Knots | 25-34 Knots | 15-24 Knots | 6-14 Knots | 1-5 Knots | calm

Wind Summary - June, July, and August

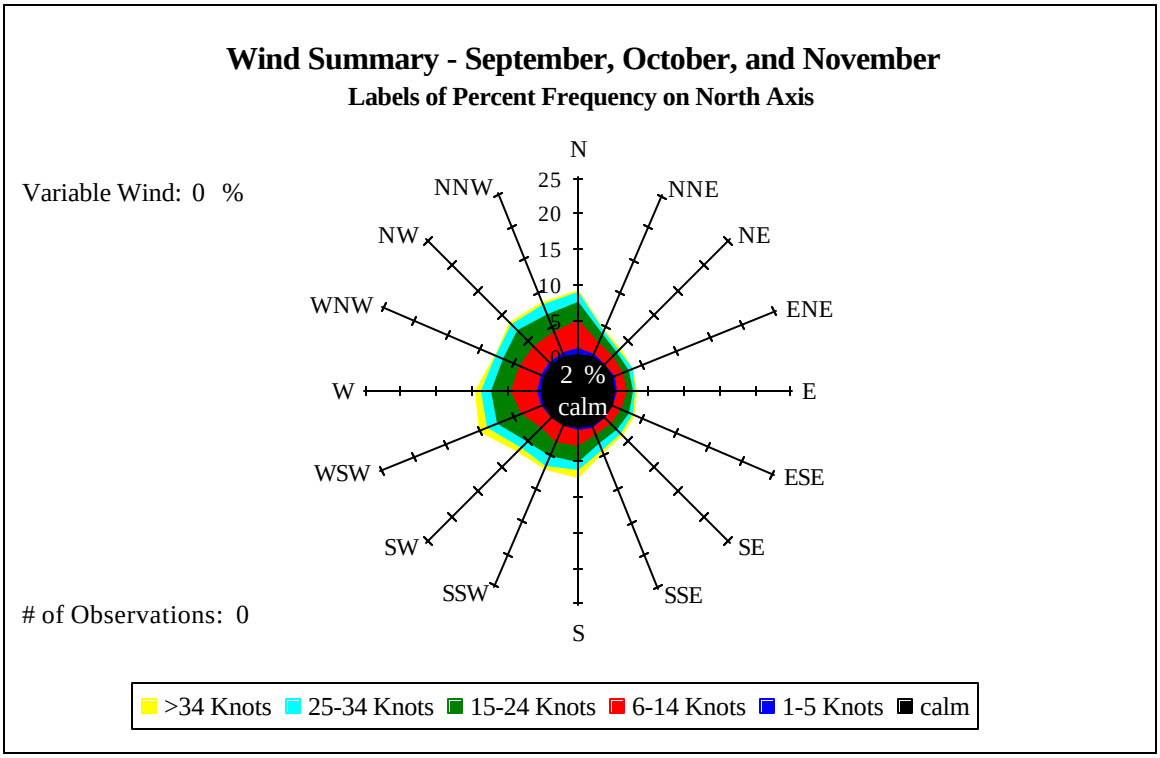
Labels of Percent Frequency on North Axis

Variable Wind: 0 %



# of Observations: 0

>34 Knots | 25-34 Knots | 15-24 Knots | 6-14 Knots | 1-5 Knots | calm



Suggestions for Use. Knowing the probable wind speed and direction in a particular season can be helpful in construction and mission planning as well as in designing structures which must face severe wind-driven rain or drifting snow. Engineers designing heating and air conditioning systems which draw fresh air from the weather, and exhaust-contaminated building air can use these data to minimize the potential for cross-contamination between supply and exhaust air streams. Also, when accumulation on roofs of drifting snow is likely, this information can be helpful for locating inlet and exhaust ducts so they are less likely to be covered by snowdrifts.

Cautionary Note: The wind currents around any building are strongly affected by the geometry of the building and the topography of the site as well as any surrounding buildings. The wind data used for these wind summaries are typical of flat and open airfields, where there are no obstructions near the observation point.

REFERENCES AND SUPPORTING INFORMATION FOR APPENDIX A

Section A—References

1997 ASHRAE Handbook: Fundamentals, American Society of Heating, Refrigerating, and Air Conditioning Engineers, Inc., Atlanta, GA

Anderson, Keith E., *Ground Water Handbook*, National Ground Water Association, Dublin, OH

Fact Book of the Society of American Military Engineers, Society of American Military Engineers, June1995 (Prince Street, Bldg 607, Alexandria, VA 22314 [Tel: 1-800-336-3097])

Harriman, Plager, and Kosar, "Dehumidification and Cooling Loads from Ventilation Air," *ASHRAE Journal*, November, 1997, pp: 37-45, American Society of Heating, Air Conditioning and Refrigerating Engineers, Atlanta, GA, 1997

International Plumbing Code, Building Officials and Code Administrators (BOCA) International, Inc., Country Club Hills, IL, 1995

Judge, James, P.E., *PSYFUNC Algorithms*, LINRIC Co., Bedford, NH, 1996

Minimum Design Loads for Buildings and Other Structures, ANSI/ASCE 7-95, American Society ofCivil Engineers, New York, NY *Solar Radiation Data Manual for Buildings*, National Renewable Energy Laboratory, Golden, CO, 1995

TI 809-01, Load Assumptions for Buildings, 1986, U.S. Army Corps of Engineers

Section B— Abbreviations and Acronyms

Btu/lb—British thermal units per pound of air (enthalpy)

Btu/sq ft/day—Btu per square foot per day (solar radiation)

cm—Centimeter (frost depth)

cm/hr—Centimeters per hour (rain rate)

gr/lb—Grains per pound (humidity ratio, grains of water vapor per pound of air)

gr/kg—Grams per kilogram (humidity ratio, grams of water vapor per kilogram of air)

in Hg—Inches of mercury (atmospheric pressure)

in—Inches (frost depth)

in/hr—Inches per hour (rain rate)

kBtu/cfm —Thousands of Btu per cubic foot per minute (sensible or latent heating or cooling loads)

klux-hr—Thousands of lux-hours (average incident illuminance)

lb/sq ft—pounds per square foot (snow load)

mb Hg—millibars of mercury (atmospheric pressure)

mph—miles per hour (wind speed)

ton-hrs/cfm/yr—ton-hours of load per cubic foot per minute per year (Btu÷12,000)

APPENDIX B: WIND OBSERVATION INFORMATION

This appendix includes information on observation procedures for the wind data used in this research.

Wind observations are taken as a representation of the touchdown area of the active runway. Wind speeds are reported in nautical miles per hour (knots), to the nearest whole knot.

Average wind speed is determined by obtaining a two-minute average (10-minutes overseas) for the period immediately preceding the time of observation.

Wind Gusts are determined by taking the maximum wind speed observed during the ten- minute period before the actual time of the observation and by using the following procedures:

1) Report the character as a wind gust when the speed observed varies during the 10-minute observational period by 10 knots or more between peaks and lulls. The value reported is the maximum wind speed. Squalls are reported as present weather.

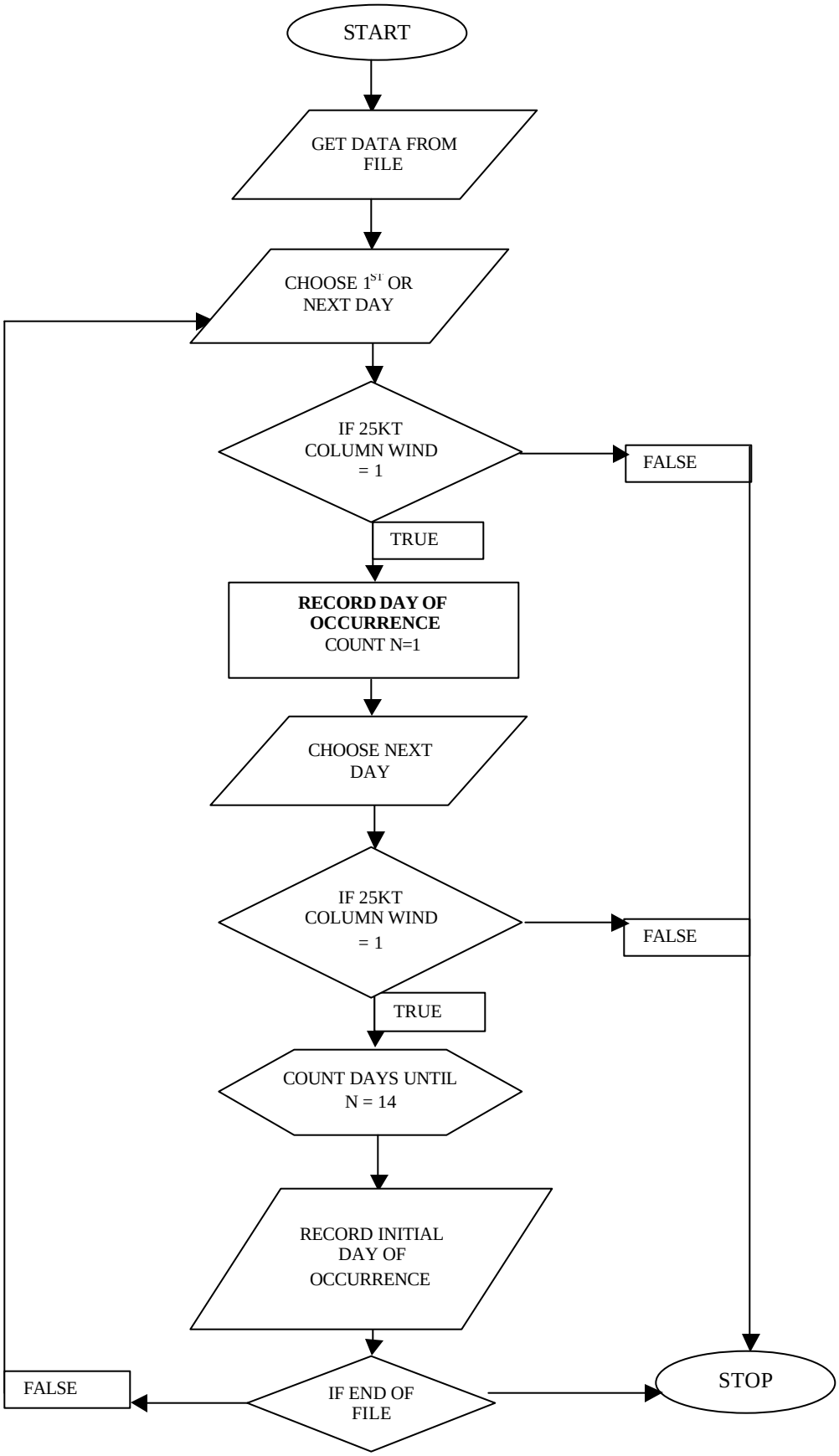
2) When available, use a recorder or digital readout to determine the occurrence and speed of gusts, squalls, and maximum wind speed. When necessary to use a direct dial indicator, monitor it as closely as practical to observe the highest attained value. Wind instruments should be located in an unsheltered area, which is most representative of the runway conditions and is least affected by local obstructions.

APPENDIX C: COMPUTER PROGRAM: CHOOSING TWO-WEEK WINDOWS

This appendix includes computer code used to choose two-week windows from the AFCCC wind observational data. A flow chart of the process involved in obtaining the two-week window of data is also provided.

```
FUNCTION dailymaxstruxr
     datline = {datline, year:0, $
                       month:0, $
                    day:0, $
                    maxwnd:0, $
                    wndind:0}
     RETURN, datline
END


PRO DATA
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;
; THESIS DATA COMPUTING NUMBER OF SUBSEQUENT DAYS OF WIND
GT 25KTS      ;
; CREATED: 18 SEP 01          VERSION: 1          UPDATED:
;
;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;;
;;;;;;;;;;;;;;

close, /all

;OPEN TEXT FILE FROM THESIS DATA


;opens a file to read from, filename:input, 'location' then
get-lun allocates the file to a unit

openr, input, '/home/kramer2/users/schroeder/thesis
data/dailymax.txt' , /get_lun


;data=dialog_pickfile
(Path="/home/kramer2/users/schroeder/thesis data" ,
Filter='*.txt' , /read)
;openr, input, filename, /get_lun
```

```
s='   '                                        ;set variable
n=0L                                       ;start counter
while not (eof(input)) do begin
     readf, input, s                              ;read
first and subsequent lines
     n=n+1L                                    ;keep on
counting till the end is reached
endwhile


;IMPORT AND SET-UP DATA TABLE

point_lun, input, 0                             ;moves pointer
to start of file/ byte 0
format1='(I4, 1x, I2, 1x, I2, 1x, I2, 1x, I1, 1x, I1, 1x,
I1, 1x, 1I)'
line = dailymaxstruxr()
wndata = replicate(line, n)
readf, input, wndata
close, input
free_lun, input


openw, output, '/home/kramer2/users/schroeder/thesis
data/wndout.txt', /get_lun


zero=intarr(14)


count=0

for i=0, n-14 do begin
     IF ARRAY_EQUAL(wndata[i:i+13].wndind, zero) THEN begin
     count=count+1
     printf, output,  wndata[i].year, wndata[i].month,
wndata[i].day
     endif
endfor

print, count
close, output
free_lun, output

end
```


START
GET DATA FROM FILE
CHOOSE 1ST OR NEXT DAY
IF 25KT COLUMN WIND = 1
FALSE
TRUE
RECORD DAY OF OCCURRENCE
COUNT N=1
CHOOSE NEXT DAY
IF 25KT COLUMN WIND = 1
FALSE
TRUE
COUNT DAYS UNTIL N = 14
RECORD INITIAL DAY OF OCCURRENCE
IF END OF FILE
FALSE
STOP

APPENDIX D: FORECAST DECISION MODEL

After analyzing the results from the model run's TI and SST splits along with the "t" test and the synoptic comparison study, development of an experimental forecast tool was accomplished. This forecast decision model is only an example, and should not be used for operational purposes before extensive field evaluation. However, it could be used to assist forecasters in developing an operational forecast tool for Shemya. The model was broken down into three sub-sections, each with a total possible score of 100.

Section I implements the results that were obtained from the May vs June or June vs July linear relationship and CART models. This section totals 100 points, with three subsections, each of which assesses scores based on TI and SST split values.

Section II examines the ending nodes of each tree. The score was assessed based on the spread of the computed confidence intervals.

Section III focuses on the locations of synoptic features, strengths of low and high pressure centers, and a comparison of pressures between the lowest and highest pressure centers over the North Pacific.

The values in the model were produced by subjectively examining the parameters for each section, and assigning the value of each score based on the amount of influence that particular parameter had on the outcome of the final score.

In section I, the NAO had the most points due to its strong linear relationship, and the other TI values were given scores based on improvement and student "t" values. Higher improvement scores and affirmative student "t" values lead to higher scores for that subsection, indicating more influence for that subsection over the outcome of the final score.

Section II scores were assessed by investigating the improvements of the splits and widths of each of the ending node's confidence intervals (CI). A high improvement and smaller CI indicates more confidence in the data, and therefore a lower score; a higher score is recorded for low improvement scores and larger CIs . For May vs June, ending nodes two and three had the largest and similar intervals, six the second largest, and five the least. The four ending node scores were assessed based on this rank and order, and the scores were compiled for a total of 100 for the entire section. June vs July was assessed in the same manner, with nodes one and four having the largest and similar intervals.

Section III scores were based solely on subjective meteorological reasoning. The forecaster can use the final score of this forecast model, along with other forecasting tools, such as medium-range model runs, to predict the best time to begin the construction project. A final score of 300 indicates a low probability of successfully completing the construction project with the possibility of very windy conditions existing on the island. Lower scores designate a lower risk of windy conditions and a higher probability of successfully completing the project.

Forecast decision model

Section 1	May Vs June		June Vs July	
	NAO	Score	NAO	Score
	Positive (or 0):	50	Positive (or 0)	50
	Negative:	10	Negative	10
	PNA	Score	NAO	Score
	≤ 0.30	25	> -0.05	25
	> 0.30	5	≤ -0.05	10
	SOI	Score	SST 4	Score
	> -1.15 & ≤ 0.55	25	≤ 28.585	25
	else	10	> 28.585	5
	Total Score (TS1)		Total Score (TS1)	

Section 2	May Vs June		June Vs July	
	End Node	score	End Node	score
	2	35	1	45
	3	35	3	10
	5	10	4	45
	6	20		
	Total Score (TS2)		Total Score (TS2)	

Section 3	Synoptic Set-up (mean position of systems)	For both models
	Storm track (movement and location)	Score
	Over Shemya	20
	North of Shemya	10
	South of Shemya	5
	Location of North Pacific High	Score
	U.S. Pacific NW	20
	Gulf of Alaska	10
	Bering Sea	5
	Location of Icelandic Low	Score
	Over Gulf of St. Lawrence	20
	Over Northen Quebec / Ontario	10
	Poleward (above 60 deg N)	5
	Long wave pattern	Score
	meridional	20
	zonal	5
	Pressure difference between Aleutian Low and North Pacific High	
	North Pacific High (NPH) - Aleutian Low (AL) = Pressure Difference (PD)	
	50 % of PD = surface wind speed (SW)	
	80 % of PD = surface gust speed (SG)	
	Criteria	Score
	IF SW and SG > 25 kts	20
	IF SW < 25 kts & SG > 25 kts	10
	IF SW and SG < 25 kts	5
	Total Score (TS3):	

Final score from forecast decision model

Final		TS1 + TS2 + TS3 = FS:	
Section			
		FS value	Risk Level
		75 or less	Low
		76-150	moderate
		151-225	high
		226-300	extreme

An example of how to use this subjective model follows by using TI and SST data from May and June 1999, to forecast windy conditions for June and July, respectively.

Sections I and II: SST and TI data for the PNA, SOI, NAO, and SST 4 are collected and put into the model for 1999.

Table: SST and TI data from U.S. CPC for 1999.

	May 1999	June 1999
PNA	-1.0	Not Calculated
SOI	0.1	-0.1
NAO	0.9	0.7
SST 4	Not Included	28.00

Table: Results from model section I.

May Vs June

NAO	Score
Positive (or 0):	**50**
Negative:	10

PNA	Score
≤ 0.30	**25**
> 0.30	5

SOI	Score
> -1.15 & < 0.55	**25**
else	10

Total Score (TS1)	**100**

June Vs July

NAO	Score
Positive (or 0)	**50**
Negative	10

NAO	Score
> -0.05	**25**
≤ -0.05	10

SST 4	Score
< 28.585	**25**
> 28.585	5

Total Score (TS1)	**100**

The bold-faced values indicate the results from the TI and SST input information

Section II uses the regression trees that were presented in Chapter 6 and Appendix D. The ending node is found and the score is obtained by using the criteria in model section II.

Table: Results from model section II.

May Vs June

End Node	score
2	35
3	35
5	**10**
6	20
Total Score (TS2)	**10**

June Vs July

End Node	score
1	45
3	**10**
4	45
Total Score (TS2)	**10**

Section III information is obtained by studying the mean synoptic patterns for the same month that the TI and SST data are collected.

Table: Section III results.

Synoptic Set-up (mean position of systems) For both models

Storm track (movement and location)			Score
Over Shemya			**20**
North of Shemya			10
South of Shemya			5

Location of North Pacific High			Score
U.S. Pacific NW			20
Gulf of Alaska			**10**
Bering Sea			5

Location of Icelandic Low	Score
Over Gulf of St. Lawrence	**20**
Over Northen Quebec / Ontario	10
Poleward (above 60 deg N)	5

Long wave pattern		Score
meridional		**20**
zonal		5

Pressure difference between Aleutian Low and North Pacific High

North Pacific High (NPH) - Aleutian Low (AL) = Pressure Difference (PD)
50 % of PD = surface wind speed (SW)
80 % of PD = surface gust speed (SG)

Criteria		Score
IF SW and SG > 25 kts		**20**
IF SW < 25 kts & SG > 25 kts		10
IF SW and SG < 25 kts		5

Total Score (TS3):	**90**

By adding the scores from all three sections, the final score is obtained which shows the risk level of windy conditions on the island. A score of 200 was calculated for both May and June, which indicates a high risk of seeing windy conditions on Shemya for June and July, respectively. The observations for these months showed seven days of winds greater than 25 kts for June, and five days for July in 1999.

Table: Final score results.

TS1 + TS2 + TS3 = FS:	**200**

FS value	Risk Level
75 or less	Low
76-150	moderate
151-225	**high**
226-300	extreme

This tool is simply an example of how forecasters may exploit the results that were uncovered by this research. It still requires verification and validation before any consideration for implementation. The tool should not be used operationally without extensive validation by experienced field meteorologists. Until then, forecast guidance offered through Figures 20 and 21 in Chapter 5 of this research should be used.

Bibliography

Air Force Combat Climatology Center (AFCCC), "Siberia: A Climatological Study, Volume I: Eastern Siberia," Technical Note TN-00/001, 2000.

Air Force Combat Climatology Center (AFCCC), Global Tropical/Extratropical Cyclone Climatic Atlas (GTECCA), CD, 1996.

Air Force Weather Agency (AFWA), "Meteorological Techniques," Technical Note TN-98/002, 1998.

Barnston, Anthony G. and Livezey, Robert E. "Classification, Seasonality and Persistence of Low-Frequency Atmospheric Circulation Patterns," American Meteorological Society: Monthly Weather Review, 1083-1126, 1987.

Baur, Franz. "Extended-Range Weather Forecasting," Beitr. Geophysics, 814-816, 1951.

Bell, Northern Hemisphere Teleconnections, 1998. http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.html

Blair, T.A. and Fite, R.C. Weather Elements, Prentice Hall, pp.364, 1965.

Breiman, Classification and Regression Trees, Wadsworth, pp. 368, 1984.

Burrows, William R. and Assel, Raymond A. "Use of CART for Diagnostic and Prediction Problems in the Atmospheric Sciences," American Meteorological Society: 12th Conference on Probability and Statistics in the Atmospheric Sciences, 161-166, 1992.

Cameron, Colin A. and Trivedi, Pravin K. Regression Analysis of Count Data, Cambridge, pp. 432, 153, 1998.

Department of the Air Force. Air Force Manual 15-111, 01 September 2000.

Devore, Jay L. Probability and Statistics: For Engineering and the Sciences, Brooks/Cole pp. 750, 2000.

Djuric, Dusan. Weather Analysis, Prentice Hall, pp. 304, 1994.

Freestrom, Hugh J. "Designing an Algorithm to Predict the Intensity of the Severe Weather Season," Master's Thesis, Air Force Institute of Technology, Department of Engineering Physics, 29-51, 2002.

Fritz, S. "The Aleutian Low in January and February- Relation to Tropical Pacific Sea Surface Temperatures," American Meteorological Society: Monthly Weather Review, 271-273, 1984.

Glossary of Meteorology, American Meteorological Society, pp. 855, 2001.

Hess, S. L. Introduction to Theoretical Meteorology, Krieger, pp. 362, 1979.

Huang, Rong-Hul. "The East Asia/Pacific Pattern Teleconnections of Summer Circulation and the Convective Activities Over the Western Pacific Warm Pool," Institute of Atmospheric Physics, Academia Sinica, Beijing, China, 121-124, 1989.

Katz, Richard W. and Brown, Barbara G. "The Problem of Multiplicity in Research on Teleconnections," International Journal of Climatology, Volume 2, 505-513, 1991.

Leroux, Marcel. Dynamic Analysis of Weather and Climate, Wiley Praxis, pp. 365, 1998.

Missile Defense Agency, Fact Sheet, 2001. http://www.acq.osd.mil/BMDO/BMDOlink/pdf/jn0019.pdf

McCullagh, P. and Nelder, J.A. Generalized Linear Models, Chapman and Hall, pp. 532, 1995.

Namias, Jerome. "Long-range Weather and Climate Predictions," National Research Council Assembly of Mathematical and Physical Sciences Geophysical Predictions, Washington D.C. Academy of Science, 103-114, 1978.

National Climatic Data Center, Surface and Upper Air Weather Charts, Series A, Volume 2, Number 6, CD, June 1995.

National Climatic Data Center, Surface and Upper Air Weather Charts, Series A, Volume 2, Number 7, CD, July 1995.

National Climatic Data Center, Surface and Upper Air Weather Charts, Series A, Volume 3, Number 6, CD, June 1996.

National Climatic Data Center, Surface and Upper Air Weather Charts, Series A, Volume 3, Number 7, CD, July 1996.

Panofsky, Hans A. and Brier, Glenn W. Some Applications of Statistics to Meteorology, The Pennsylvania State University, pp. 224, 1968.

Randall, Robb M. "Exploration of Teleconnection Indices for Long-Range Seasonal Temperature Forecasts," Master's Thesis, Air Force Institute of Technology, Department of Engineering Physics, 31-42, 2002.

Robinson, Peter J. and Henderson-Sellers, Ann. Contemporary Climatology, Longman, pp. 384 , 111, 1999.

Rogers, Jeffery C. "The North Pacific Oscillation," American Meteorological Society: Journal of Climatology, Volume 1, 39-57, 1981.

SPSS Inc. Answer Tree 3.0 User's Guide, pp. 226, 2001.

United States Climate Prediction Center, Atmospheric Teleconnections, 2001. http://www.cpc.ncep.noaa.gov/data/teledoc/telecontents.html http://www.cpc.ncep.noaa.gov/data/indices/index.html

Wallace, John M. and Gutzler, David S. "Teleconnections in the Geopotential Height Field During the Northern Hemisphere Winter," American Meteorological Society: Monthly Weather Review, Volume 109, 784-812, 1980.

Washington, Richard, et. al. "Northern Hemisphere Teleconnection Indices and the Mass Balance of Svalbard Glaciers," International Journal of Climatology, 473-487, 2000.

Wilks, Daniel S. Statistical Methods in the Atmospheric Sciences, Academic Press, pp. 467, 1995.

Vita

Brian Schroeder was born in 1971 in Lincoln, Nebraska. He graduated from Waverly High School in Waverly, Nebraska in 1989. Brian graduated from the University of Nebraska at Lincoln (UNL) with his B.S. in Meteorology / Climatology in 1994. He then attended UNL as a graduate student before receiving his commission in 1995 through ROTC. While attending school full-time, he was involved in many activities to include AFROTC, the American Meteorological Society, Theta-Chi fraternity, UNL rugby club, and the UNL marching band. He worked as an on-air meteorologist for a local television station in Lincoln before beginning his Air Force career. His Air Force career began as a Wing Weather Officer at Charleston Air Force Base in South Carolina. He was then assigned to the United States Strategic Command, at Offutt Air Force Base, Nebraska, where he flew as a battlestaff member of the "Looking Glass" strategic defense aircraft. Brian is now attending the Graduate Meteorology program, Department of Engineering Physics, Air Force Institute of Technology at Wright-Patterson Air Force Base, Ohio. Upon graduation, he will be assigned as Flight Commander to the USAFE weather hub at Sembach Air Base, Germany.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
24-05-2002	**Master's Thesis**	Jun 2001 – Jun 2002

4. TITLE AND SUBTITLE

Long Range Forecast Possibilities for X-Band Radar Construction on Shemya

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Schroeder, Brian, K, Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)

Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 P Street, Building 640
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GM/ENP/02M-12

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Michael W. Holmes, Capt, US AF
11 Operational Weather Squadron DSN: 317-552-0246
Chief, weather Sciences Comm: 907-552-0246
Elmendorf AFB, AK 99506 e-mail: michael.holmes@elmendorf.af.mil

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Missile Defense Agency (MDA) plans to construct, during a two-week period, an X-Band Radar (XBR) on Shemya, AK. Wind speeds must not, at any time during the construction, exceed a 25 knot limit set by the MDA for lifting the massive dome panels into place. The goal of this research was to explore the possibilities of long-range forecasts to determine the feasibility of predicting any upcoming two-week windows of opportunity, well in advance, that will ensure the successful completion of constructing the XBR on Shemya. In order to reach this goal, the following objectives were achieved; 1) a climatological wind study for Shemya to assess the optimal "climatological window" to build the XBR, 2) a detailed synoptic study over the North Pacific, to gain an understanding of how synoptic weather systems develop, move, and vary on an annual basis, 3) a traditional statistical analysis of the data followed by a Classification and Regression Tree (CART) analysis for pattern recognition of global teleconnection indices, and 4) develop forecasting decision trees to assist the 11th Operational Weather Squadron (OWS) Alaskan forecast hub in this daunting task.

The Aleutian Island chain is plagued by persistent strong winds, since the Aleutian Low and expanding polar vortex affect the region in the winter, as do tropical storms and frontal passages in the summer. This, combined with Shemya being located near the exit region of the climatological storm track off the East Asian continent, makes the island one of the most challenging forecast locations in the Northern Hemisphere. This study compares surface winds and teleconnection indices, as computed by the Climate Prediction Center (CPC), to statistically analyze the data.

The data were analyzed using standard statistical regression techniques, including linear and multiple linear regression methods, and then CART analysis was used for large-scale pattern recognition. The approach of the CART analysis theory used in this study was to determine which large-scale pressure patterns in the Northern Hemisphere are conducive to consistently "low winds" over Shemya. CART was discovered to be the best method of analysis and forecast decision trees for the 11th OWS Alaskan forecast hub were then developed to assist forecasters in providing long-range wind forecasts for the MDA, as well as the risks involved of being wrong. Since the results of the study can't offer conclusive go or no-go forecasts, an alternative "wind-break" proposal is included in the recommendations.

15. SUBJECT TERMS

NONE

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Brian K. Schroeder, Capt, USAF (ENP)
U	U	U		139	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-3636, ext 4626; e-mail: brian.schroeder@afit.edu

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18