THE VERIFICATION OF WEATHER FORECASTS, NUMERICAL MODELING, CLIMATOLOGY, AND VARIATIONS IN ENSO CYCLES IN FORECAST ACCURACY

A Thesis presented to the Faculty of the Graduate School
University of Missouri-Columbia

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

---

by
MICHELLE C. BOGOWITH
Dr. Anthony Lupo, Thesis Supervisor
December 2011

The undersigned, appointed by the Dean of the Graduate School,
have examined the thesis entitled

THE VERIFICATION OF WEATHER FORECASTS, NUMERICAL MODELING, CLIMATOLOGY, AND VARIATIONS IN ENSO CYCLES IN FORECAST ACCURACY

Presented by Michelle C. Bogowith

A candidate for the degree of Master of Science

And hereby certify that in their opinion it is worthy of acceptance.

Professor Dr. Anthony Lupo

Associate Professor Dr. Patrick S. Market

Assistant Professor Dr. Pat Guinan

Associate Professor Dr. Mark Palmer

ACKNOWLEDGEMENTS

At this time, I would like to say thank you to a few people who have helped me further my education in Atmospheric Science. First and foremost, Dr. Anthony Lupo has been a blessing to have not only as an advisor and mentor in my career, but also a friend. His guidance and support throughout grad school has helped to push me to better myself in this profession, and I am forever in debt to him. Without Dr. Lupo, I would not have pursued graduate school. Not to forget Dr. Patrick Market, another mentor, who has also provided encouragement and support throughout this process. Dr. Pat Guinan has been instrumental over that past few years with providing pieces of climatological data for on air use at KOMU. Dr. Mark Palmer, I appreciate the time you have taken out of your schedule to help support me in my pursuit for a graduate degree.

Aside from my thesis committee, there is one individual that has been by my side from day one. This person is not only a fellow student and co-worker of mine, but also a dear friend, Eric Aldrich. After countless hours of number crunching in Excel at the start of this project just to be told by Dr. Lupo halfway through that we used the wrong equation, you have helped to keep me sane in our pursuit for a graduate degree. The support and encouragement that we have shared through this process is the reason I made it through. It has been a pleasure working with you both at KOMU and in the classroom, and I wish you nothing but the best for the future. Can't wait for Team B.A. to walk across the stage in December!

Atmospheric Science undergraduate, Josh Kastman, played a big role in helping enter the data into the excel program. His knowledge helped provide short cuts for data analysis, that otherwise would have double or tripled the work for Eric Aldrich and myself.

The National Weather Service in St. Louis, Missouri must also be recognized for its help with the recorded data and model assistance. It's easy to maneuver website has helped with extracting data for this project.

Finally, to my family and friends, I don't know how you all put up with me. My mother and father, along with my younger brother, have been encouraging me to do my very best. Without their love and support, I would not be where I am today. To my friends, who have been patient with me and understanding that school comes first. I have given up a lot of time that could have been spent with you. I will make up for that soon. From the bottom of my heart, I love you all!

TABLE OF CONTENTS

FIGURES

TABLES

THE VERIFICATION OF WEATHER FORECASTS, NUMERICAL MODELING, CLIMATOLOGY, AND VARIATIONS IN ENSO CYCLES IN FORECAST ACCURACY

Michelle Bogowith

Dr. Anthony R. Lupo, Thesis Supervisor

ABSTRACT

Weather forecasting has been a difficult problem for meteorologist for more than one hundred years. In the early days, our understanding of atmospheric processes was limited by our incomplete understanding of the governing principles and equations. The possibility of numerical forecasting was also still many years away. The invention of the television and computers made numerical weather prediction possible by the mid-20th century. Shortly after this, television was used to broadcast the local news and weather forecasts. Understanding the atmosphere, however, was still incomplete and this made it difficult for meteorologists to make weather forecasts. In the present day, we have numerous models and advanced technologies to look at atmospheric data and analyze in order to produce the most accurate forecast possible. But how accurate are these forecasts? My research and this paper will focus on the accuracy of the human forecaster (from meteorologists at KOMU TV 8 in Columbia, Missouri) and how the forecast compares to predictions from numerical models and local climatology. Initial results suggest the human produced forecast outperformed not only climatology but also numerical models overall. However, when breaking down the data and analyzing it versus ENSO cycles, while this result was true for most seasons and ENSO phases, it was not always the case. With the analysis of the data, it is intended we will gain an understanding of systematic biases human forecasters or models possess for this region with respect to season and ENSO phase. If we understand these biases, it will improve our

forecasting skills by accounting for them and increasing accuracy. This, in turn, will boost confidence in a human generated forecast.

1 INTRODUCTION

This research focuses on the accuracy of the human forecasters at KOMU TV 8 in Columbia, Missouri. The human forecast is compared to predictions from numerical models and local climatology. In order to gain an understanding of systematic biases that models and humans possess with respect to season and ENSO phases, data from November of 2007 to 2010 was analyzed. It is expected that the human forecasters would outperform climatology, but that the model forecasts would be similar to that of the human forecasts. Also, if we can target and account for the biases, it will improve our forecasting skill by increasing accuracy. This, in turn, will boost confidence in a human generated forecast. Human forecast accuracy will outperform both models and climatology.

2 BACKGROUND

2.1 NUMERICAL WEATHER PREDICTION

The curiosity of Benjamin Franklin helped to pave the way for weather experimentation and investigation (Franklin, 1774). While it may now seem simplistic, his data extraction of the current atmospheric conditions sparked interest in forecasting similar weather for points a little farther to the east. This fascination began to grow, and questions arose of how the atmosphere worked and what were the causes of such weather phenomenon.

As we entered the 20th century, the science of meteorology became increasingly popular as more scientists began to study it. Vilhelm Bjerknes proposed that if the state of the atmosphere develops from the preceding one in accordance with physical law, then the problems of forecasting can be mended by:

1. Sufficiently accurate knowledge of the state of the atmosphere at the initial time.

2. Sufficiently accurate knowledge of the state of the laws according to which one state of the atmosphere develops from another. (Bjerknes, 1904).

The first solution called for a daily, complete diagnosis of the atmospheric conditions at a given time. The second solution called for computations of seven unknown quantities given seven equations representing our three conservation principles. The seven unknown quantities include three velocity components, density, pressure, temperature, and humidity of the air. Navier and Stokes formulation of the equation of motion for a fluid was a key component. However, by the time the computations would be completed, the information was outdated (Bjerknes, 1904).

In 1923, British meteorologist Lewis Frye Richardson proposed the idea of predicting the future state of the atmosphere by using physical equations. Regrettably, without breakthroughs in computer technologies at that time, he too couldn't pursue the idea without the proper resources. Richardson wrote about the vision he had for the future of numerical weather prediction which seemed almost impossible and unattainable:

> "Perhaps someday in the dim future it will be possible to advance the computations faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream." (Woolard, 1922).

As decades passed, what once was thought to be impossible was achieved. Breakthroughs in computer technologies allowed for the development of simple barotropic models by

University scientists and government organizations. By 1960, the first operational predictions were accessible. As scientific knowledge grew, this idea grew in decades to come (Vasquez, 2008).

The human forecaster bases a forecast on gained knowledge and experience of model and data interpretation. This data helps one to understand how the atmosphere will change over a period of time. Everyone has their own way of doing this; therefore, human forecasts will vary in comparison to the model data.

It is remarkable how far we have come with numerical weather prediction within the last few decades. Looking back on the days when it was hard to predict a few hours out let alone a few days. With the technologies and advances in computer modeling, the art of meteorology has been fine tuned to a point that forecast accuracy has increased dramatically. But in spite of this, there are still deficiencies in the models. The purpose of this research project is identify possible reasons for these deficiencies so that the human forecaster can account for them and further increase forecast accuracy. Aside from the human produced forecast, the Global Forecast System (GFS) and the North American Mesoscale (NAM) were the two models that this research focused on, in addition to the climatology data which was obtained from the National Weather Service archives (http://www.weather.gov/climate/index.php?wfo=lsx). This research will allow for us to look further into the accuracy of humans and computer models, and identify the role that ENSO cycles play in forecast accuracy.

2.2 ENSO CYCLES

As stated above, the role that El Nino Southern Oscillation, or ENSO cycles play in forecast accuracy is yet to be determined. Before we get started, let’s explain what exactly is ENSO? ENSO is the shifting of the winds and ocean currents within the Pacific Ocean. There are two phases that we will be focusing on for this project: El Nino and La Nina.

Let us first start by explaining the typical set up in the Pacific Ocean, or the neutral phase. The normal set up in the tropical Pacific, the trade winds blow westward from a high pressure in the eastern Pacific to low pressure in the western Pacific. Trade winds help produce upwelling. Upwelling brings cooler waters from the bottom of the ocean to the surface in the east. This cool water spreads out along the equator in the Pacific where it is heated by sunlight as it moves westward. Therefore, the eastern Pacific waters are cooler and the western Pacific waters are warmer. Typically, the eastern Pacific water temperatures are lower than that of the western pacific, due in part to the trade winds. This produces a thicker layer in the warmer western Pacific waters (Ahrens, 2003).

During El Nino, we see a breakdown of the pressure patterns. The pressure will begin to rise in the west and fall in the eastern Pacific. This shift in pressure causes the trade winds to weaken and reverse in direction. At this time, the thermocline, an imaginary boundary that separates the warm waters at the surface from the cold waters below, is triggered, and begins to shift aiding in the process. In doing so, the countercurrent gains strength, allowing for the warm waters to head eastward towards South America. Thus, the waters in the western Pacific begin cooling, as you can see in Figure 1.1. (Ahrens, 2003).

Shortly after an ENSO event, the trade winds will return to normal along with the thermocline. Sometimes, the trade winds are unusually strong. This causes cold water at the surface to move into the eastern Pacific. At the same time, warm waters return to the western Pacific along with stormy weather. This phenomenon is known as La Nina. (Ahrens, 2003).

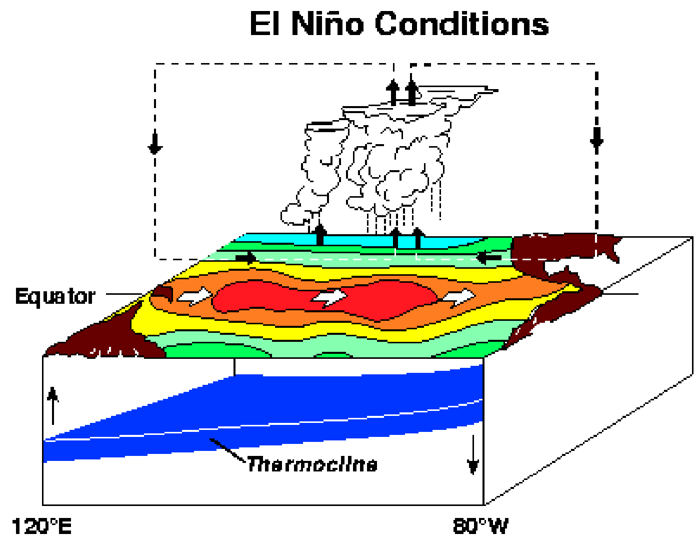

Figure 1.1 - Schematic of the tropical Pacific during El Nino
http://www.pmel.noaa.gov/tao/elnino/el-nino-story.html

During La Nina conditions, most forecasters will associate winter with mild temperatures in southern Missouri, and cooler conditions in northern Missouri. However, Birk et al. (2010) showed that generally these winters were cooler overall, and

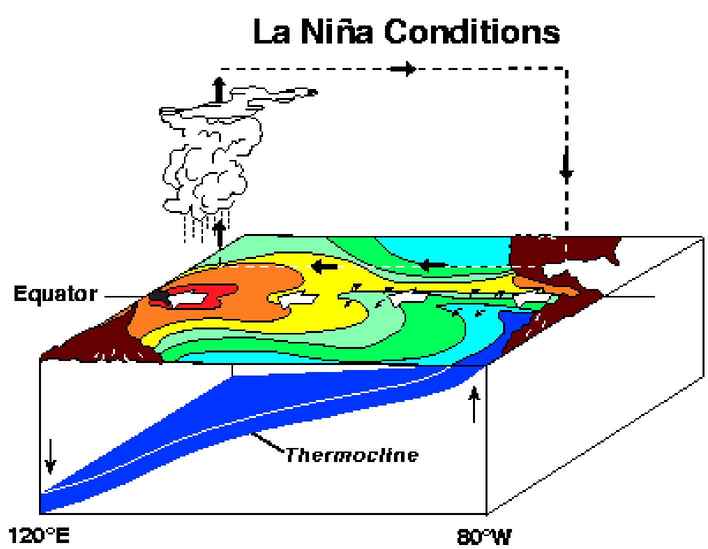

Figure 1.2 - Schematic of the tropical Pacific during La Nina.

http://www.pmel.noaa.gov/tao/proj_over/diagrams/index.html followed the pattern above during the positive phase of the Pacific Decadal Oscillation. Given what we know about La Nina, it would seem like a safe bet. There is usually a blocking High Pressure system that sits over the Gulf of Alaska. This forces the Polar jet stream to lift up over the High as it spins clockwise bringing Arctic air to the Pacific Northwest and the Northern tier of the Midwest.

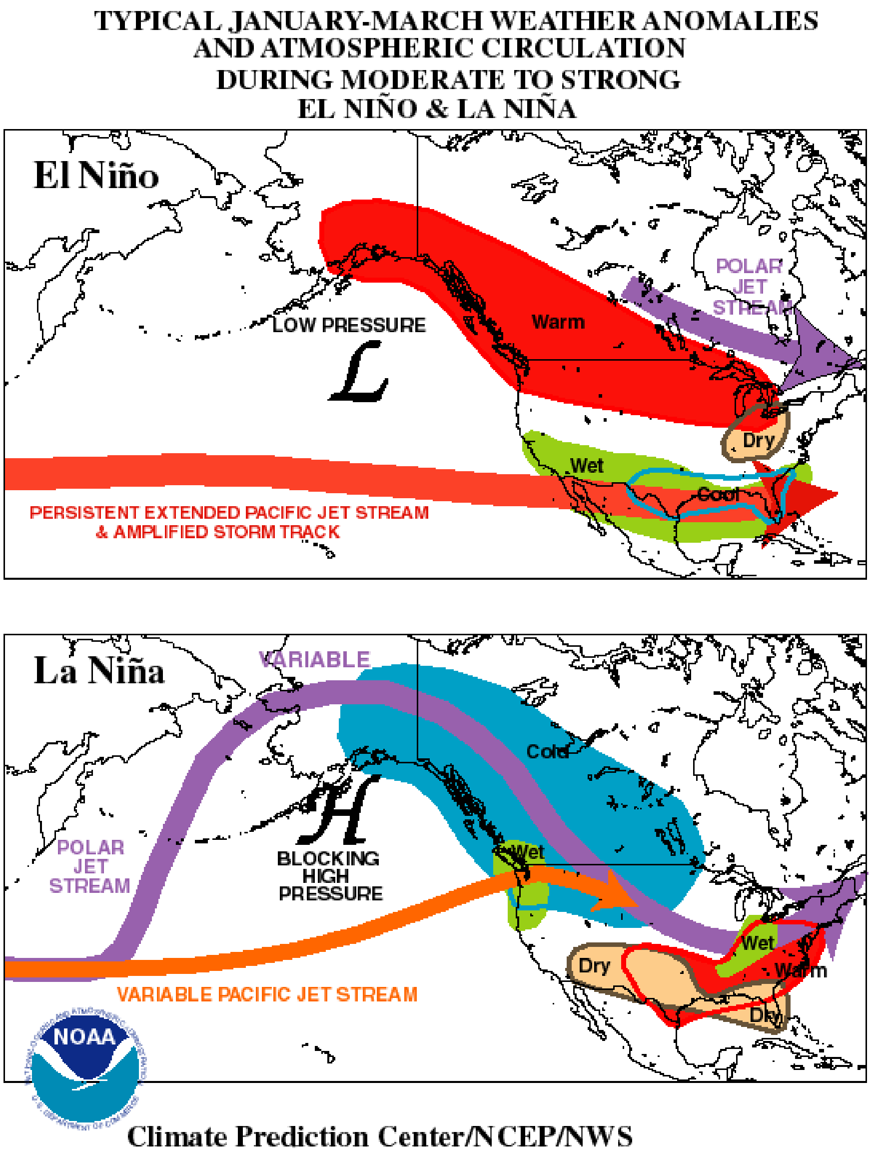

Figure 1.3 – Schematic of the upper air flow during El Nino/La Nina conditions.

http://www.srh.noaa.gov/tbw/?n=tampabayelninopage

Because of our location in the Midwest, one will generally experience mild conditions and above average precipitation during La Nina winter months. However, not all La Nina's and El Nino'a are created equal. There are differences in the magnitude and intensity of conditions that emerge in the equatorial Pacific. The publication "Interannual and interdecadal variability in the predominant Pacific region SST anomaly patterns and their impact on climate in the mid-Mississippi valley region", indicates there are different types of La Nina patterns. One brings us a more mild winter season and the other pattern does the complete opposite. It makes for a colder season than normal (Lupo et al, 2006)(Birk et al, 2010).

Alternatively, El Nino has low pressure centered in the Gulf of Alaska which forces the polar jet northward. Warm, moist air is drawn into the Pacific Northwest spreading Northeast into Canada and the upper tier of the lower forty-eight. In the southern half of the United States, temperatures are usually cooler and wet due to an amplified subtropical jetstream.

As stated above, not all El Nino's are created equal. Dependent upon the location of the warm pool and the strength of the cycle, these typical weather patterns associated with El Nino are subject to change.

These effects, while strongest in the winter season, still have impacts on our weather patterns throughout the year. The frequency, occurrence, and intensity of mid-latitude cyclones are influenced by the tropical SST patterns that can alter the prevailing wind patterns (Lupo et al, 2006). This can dramatically change the weather we see here in the Midwest, and more specifically, in Missouri. Awareness of these patterns can increase a forecasters skill. Unlike computer models, we can recognize these changes and tweak our forecasts accordingly.

3 ANALYSIS AND METHODS

3.1 OUR FORECAST

Over a period of three years, beginning in November of 2007 and ending in November of 2010, colleagues at KOMU, myself included, began a collection of model and forecast data. After careful analysis of the Model Output Statistics (MOS) and other atmospheric data by meteorologists and atmospheric science students from the University of Missouri, forecasts were put together for on air use to viewers in Mid Missouri. These forecasts were made using data from morning (12z) and evening (00z) model runs. For clarification, the morning forecasters used 00z data from the night before and the afternoon forecasters used 12z data from that morning. One must note in this research study, precipitation, sky cover, wind speed and direction data were not documented; therefore, they were not included.

When a temperature forecast was generated at KOMU, all information and data would be taken into account. The GFS and NAM models were used for guidance. From what we know as meteorologists and students of atmospheric science, we can take that knowledge and past experiences to determine if our thoughts follow the models or go in another direction. We would arrive at a specific number for the afternoon high and morning low, rather than a range of temperatures. Our viewers count on us for reliable weather coverage day in and day out. Our philosophy for forecasting is that a temperature range makes us less credible in the eyes of our viewers; it makes us look like we really are not confident what the temperature is going to be. By using a target number, it makes us appear more confident in our forecast, even if that number is not dead on all the time.

3.2 MODELS FORECAST

This research project focused on two models for data, the Global Forecast System (GFS) and North American Mesoscale (NAM) Model. The data were taken from 00z and 12z data, pulling forecasted high and low temperature for a 24 hour period. These data are most readily available to the public for viewing and can be found at the following link: http://www.nws.noaa.gov/cgi-bin/mos/getall.pl?sta=KCOU

North American Mesoscale (NAM)

The Weather Research and Forecasting (WRF) model was introduced as a prototype in 2000 after the U.S. , along with academia, wanted to imitate the European “community” numerical model. They made the model flexible, portable, and scalable in order to compete. The National Centers for Environmental Prediction (NCEP) put the model into operation in June of 2006 with the goal of replacing previous models such as the ETA and the RUC (Vasquez, 2008).

One of the largest weaknesses found with this model is the grid spacing of the NAM. It is rather small, often less than 10 km, and has tight resolutions making it difficult to predict on the large scale. In comparison to that of the GFS, the model is a short range model only extending out approximately 84 hours. Other biases found with the NAM model include aggressiveness with moisture. In order to correct the issue with aggressive moisture, the evapotranspiration parameters were adjusted to solve this problem. A more recent development in model bias is over forecasting wind speeds at the top of the troposphere (Vasquez, 2008)

Global Forecast System (GFS)

First introduced in March of 1981, The Global Spectral Model (GSM) replaced the 1966 Primitive Equations (PE) model (Vasquez, 2008). This model was the first of its kind to envision the atmosphere in mathematical waves as opposed to displaying data at grid points. To start, it had 30 waves around the hemisphere and 12 horizontal surfaces that represented the depth of the troposphere. After a little tweaking, in April of 2002, a model was designed taking into account two models: the Aviation (AVN) and the Medium Range Forecast (MRF). They consolidated these two model runs into one using all data available in order to maximize the view of global weather. This became known as the Global Forecast System, or GFS. Today, the GFS is a 382 wave spectral model with a resolution of about 35 km. It is one of the only weather models on the planetary scale made available to the public. It is run four times daily, and extends out to 384 hours. While this model is available out to 16 days, one must also take into account the accuracy deteriorates significantly beyond the 5-7 day range (Vasquez, 2008).

According to Vasquez (2008), there are biases that have emerged from this particular model. It does have issues with lee side weather systems along with cold air damming. Terrain issues are not the only problem. The GFS model also has difficulty with convective complexes without a source, also known as “precipitation bombs”. It is also found that meridional flow can be rather aggressive when it comes to amplifying the weather patterns during cold air outbreaks that move southward.

3.3 CLIMATOLOGY FORECAST

Because of the small amount of data given the grand scale of things, numbers were pulled from climatology in order to broaden the statistics. This allowed us to compare data to a baseline forecast. Comparing models and human forecasts to a 30 year daily average from 1981-2010 permitted us to do this. The National Weather Service in St. Louis, Missouri provided this information on their website: http://www.weather.gov/climate/index.php?wfo=lsx

3.4 PROCEDURES

For analysis of data, Microsoft Excel was used to enter the numbers separating the data into tabs by time frame, 00z and 12z. Within each time frame, the data were separated into columns: KOMU's forecast, GFS forecast, NAM forecast, and Climatology. All four categories were evaluated alongside the observed temperature for each given day. Once all the data were entered, analysis of the data began. As seen in Lupo and Market (2002), a skill score was generate to assess the accuracy of the forecast. This research only focused on the accuracy of temperature, unlike the 8 point skill score as seen in Lupo and Market (2002). The skill score for this research was a 2 point scale.

In order to get the skill score percentage, the following equation was used:

$$\boldsymbol{S = \frac{(F - B)}{(P - B)} \times 100}$$

This equation takes the forecasted temperature average (F) and subtracts it from the observed baseline temperature average (B) dividing by the difference between a perfect skill score (P) which is equal to 2 and the observed climatology temperature average (B).

Further breakdowns show that a 0, 1, or 2 point score was awarded. It is important to note here that if the number was negative, we made it a positive for the purpose of averaging. This is because both positive and negative numbers represent the same value on the skill score scale. Without this step, negative values will become a problem. The forecast was considered "perfect" if the score fell between the range of 2 and -2, and a 2 point score was awarded. The forecast was considered "accurate" if the score fell in the range of 4 to -4, and a 1 point score was awarded. If the skill score fell outside of this range, it was considered an "inaccurate" forecast and therefore awarded no points. Table 1.1 below is an example from the data set of the skill score computed in this research.

	KOMU Forecast		Observed Forecast		Results	
Date	High	Low	High	Low	High	Low
12/21/2007	58	42	58	41	2.00	2.00
12/24/2007	38	25	40	22	2.00	1.00
12/31/2007	25	17	26	11	2.00	0.00

Table 1.1 – Data set sample of KOMU skill scores. Adapted from Aldrich (2011)

The data from 12/24/2007 shows a forecasted high temperature of 38 and a forecasted low temperature of 25. The observed forecast high was 40 and low was 22.

38 (forecasted temperature) – 40 (observed temperature) = -2 A "perfect" forecast

25 (forecasted temperature) – 22 (observed temperature) = 3 An "accurate" forecast

The perfect forecast is awarded 2 points and the accurate forecast is awarded 1 point. The data from 12/31/2007 shows an example of an "inaccurate" forecast. The overnight low prediction was 17 and the observed temperature was 11. The difference was 6 degrees.

This falls out of the range therefore is an “inaccurate” forecast as is there by awarded no points. In the results column of Table 1.1, the skill scores are found. Theses scores are all averaged as seen in Table 1.2.

KOMU High	KOMU Low	Climo High	Climo Low
1.36	1.22	0.64	0.65

Table 1.2 – Averaged skill score for KOMU and Climatology.

From the data in Table 1.2, these numbers are used in the skill score equation. The average forecast skill score from KOMU is subtracted from the average skill score of climatology. Then, divide by the difference between the perfect score of 2 and the average skill score of climatology. Multiply by 100 for the skill score percentage.

$$S = \frac{(1.36-0.64)}{(2-0.64)} \times 100 = 53\%$$

This equation gives a skill score percentage. It shows the percent difference between KOMU and Climatology. As seen in Figure 1.4.

There were a few “special case” scores that were reviewed on an individual basis. At times, there were midnight high temperatures or daytime low temperatures. These special cases were either included in the data set or omitted.

Once the skill score was completed, further analyses of data were needed. The skill scores were averaged overall for both 00z and 12z, for KOMU, GFS, NAM, and Climatology. The skill scores were broken down into ENSO cycles averaging them for El Nino, La Nina, and Neutral cycles. These results can be seen in Tables 1.3 and 1.4.

	KOMU High	KOMU Low	GFS High	GFS Low	NAM High	NAM Low
La Nina 07-08'	**1.00/2.00**	**1.30/2.00**	**1.14/2.00**	**1.13/2.00**	**0.98/2.00**	**1.17/2.00**
Neutral 08-09'	**1.36/2.00**	**1.36/2.00**	**1.10/2.00**	**1.32/2.00**	**1.34/2.00**	**1.33/2.00**
El Nino 09-10'	**1.43/2.00**	**1.45/2.00**	**1.14/2.00**	**1.39/2.00**	**1.42/2.00**	**1.27/2.00**
27La Nina 10'	**1.36/2.00**	**1.25/2.00**	**1.08/2.00**	**1.04/2.00**	**1.00/2.00**	**1.05/2.00**

Table 1.3 – ENSO cycle skill score at time 12z

	KOMU High	KOMU Low	GFS High	GFS Low	NAM High	NAM Low
La Nina 07-08'	**1.55/2.00**	**1.64/2.00**	**1.09/2.00**	**1.18/2.00**	**1.27/2.00**	**1.36/2.00**
Neutral 08-09'	**1.38/2.00**	**1.20/2.00**	**1.18/2.00**	**1.15/2.00**	**1.35/2.00**	**1.16/2.00**
El Nino 09-10'	**1.31/2.00**	**1.22/2.00**	**1.12/2.00**	**1.13/2.00**	**1.25/2.00**	**1.11/2.00**
27La Nina 10'	**1.40/2.00**	**1.23/2.00**	**1.22/2.00**	**1.18/2.00**	**1.28/2.00**	**1.21/2.00**

Table 1.4 – ENSO cycle skill score at time 00z

Finally, this research project had one more step. The ENSO cycles and broke them down into seasons, averaging them to get a skill score. This allowed me to see what season of the year was the strongest/weakest forecasting period during the ENSO cycles. The results were ranked accordingly.

3.5 BREAKDOWN OF ENSO CYCLES

In order to maintain consistency, the data for sea surface temperature (SST) anomalies were pulled from the NOAA website. As previously stated, data collection began in November 2007 and runs through November 2010. Table 1.5 shows the SST anomalies, the cooling and warming of the waters in degrees Celsius.

Year	DJF	JFM	FMA	MAM	AMJ	MJJ	JJA	JAS	ASO	SON	OND	NDJ
2007	0.8	0.4	0.1	-0.1	-0.1	-0.1	-0.1	-0.4	-0.7	-1.0	-1.1	-1.3
2008	-1.4	-1.4	-1.1	-0.8	-0.6	-0.4	-0.1	0.0	0.0	0.0	-0.3	-0.6
2009	-0.8	-0.7	-0.5	-0.1	0.2	0.6	0.7	0.8	0.9	1.2	1.5	1.8
2010	1.7	1.5	1.2	0.8	0.3	-0.2	-0.6	-1.0	-1.3	-1.4	-1.4	-1.4

Table 1.5 – SST Anomalies in the tropical Eastern Pacific Ocean in Degrees Celsius

Information provided by NOAA website

http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.shtml

The warmer SST's are shown in red signifying an El Nino phase. The cooler SST's in blue represent the La Nina phase. The numbers in black are the neutral phase of the ENSO cycle. The ENSO cycle is based upon a +/- 0.5 degree Celsius threshold. When there are at least five consecutive overlapping three month periods, the anomalies for El Nino and La Nina were defined.

4 FORECAST ANALYSIS

4.1 RESULTS

The data were broken down into two time frames, 00z and 12z. As stated previously, the baseline for this research is Climatology. Climatology is an average of weather conditions over a period of time (e.g., 30 years) for a given location. The location for this experiment is Columbia, Missouri. This daily average is a number that forecasters use to provide background to the viewer on how the weather for a given location normally behaves.

However, it is known that the weather does not behave normally all the time. Several factors allow for the typical weather patterns to vary. These deviations can be caused by synoptic/meso scale events such as blocking, shifts in the jetstream, ect. This is why we will use climatology as our baseline.

The results found that the human forecaster outperforms computer models and climatology. In a handful of cases, the human forecaster lost to one of the models, and on a rare occasion, both models. More often than not, aside from the human forecaster, the model that was the strongest was the NAM. The NAM was followed by the GFS, and finally climatology. The overall results of the experiment are seen below in Tables 1.6 and 1.7.

High	Low
KOMU	**KOMU**

NAM	**GFS**
GFS	**NAM**
Climatology	**Climatology**

Table 1.6 – The overall forecast accuracy rankings for 12z

High	Low
KOMU	**KOMU**
NAM	**NAM**
GFS	**GFS**
Climatology	**Climatology**

Table 1.7 – The overall forecast accuracy rankings for 00z

The human forecaster performs best because we have the ability to synthesize information. One can use experience from past events to look beyond raw data presented by models, assess, and make our own forecasts. Models are unable to look beyond the raw data, so this allows the human forecast to improve on the model forecast. As far as climatology goes, it can't take the current conditions and use data to determine a forecast nor can it ask questions. This is why climatology performs poorly.

4.2 KOMU vs. Climatology

The overall results of this study for 00z, or the evening model run, show that KOMU's results performed better than climatology. One interesting thing to note is that there is a larger percent difference between the high temperatures than the overnight low temperatures. Climatology does a better job with predicting overnight low temperatures than afternoon high temperatures for the following day. This is a perfect example of where the human forecaster is advantageous. They are able to recognize the possibility for these changes to occur, in turn making our percentage of accuracy higher than that of climatology, as well as the models which will be discussed later. Figure 1.4 below shows the accuracy of KOMU compared to climatology. The percentages that follow are obtained by using the skill score percentage equation. For the highs, there was ~53% difference. For the lows, there was ~42% difference.

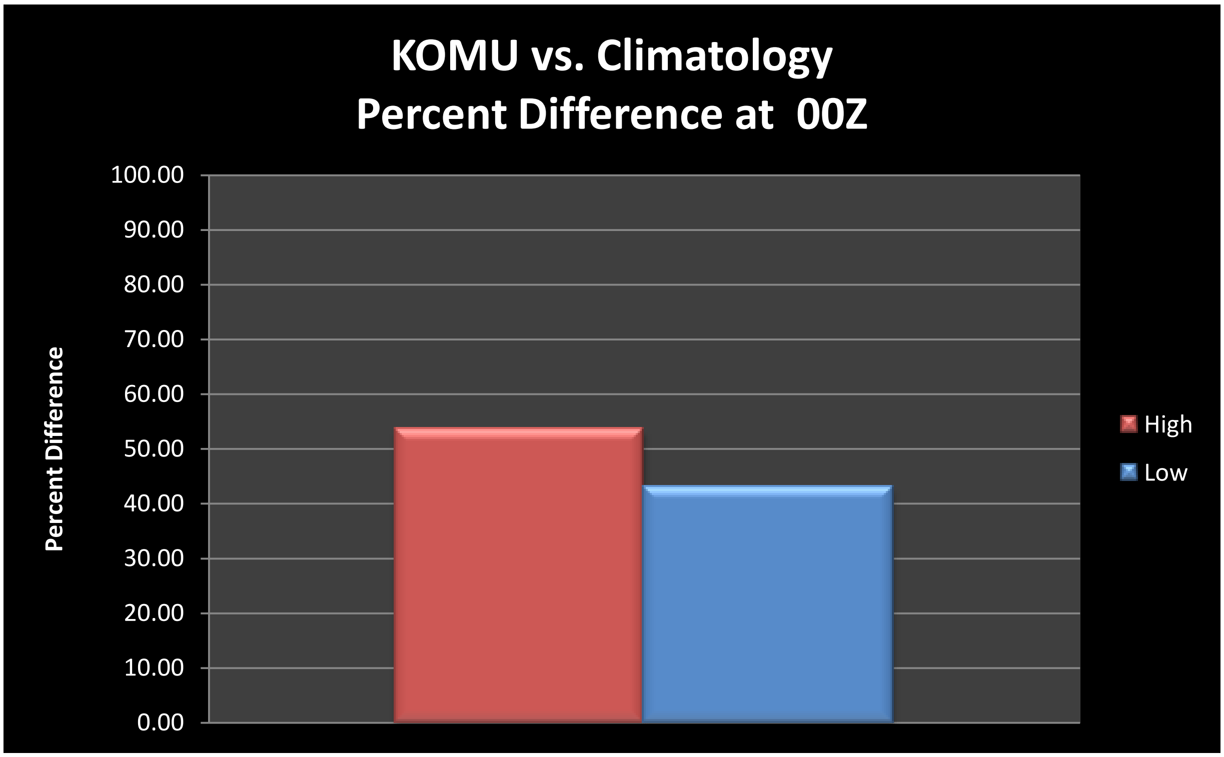

Figure 1.4 Percent difference between KOMU forecast and climatology for 00z.

As we shift to the 12z time frame, similar results are seen. KOMU exceeded the climatology forecasts by a large percentage yet again (~53% for highs and ~52% for the lows). KOMU's high temperature forecast just edged out the low temperature forecast. But the difference here is very small, only a ~1% difference between the highs and lows (Figure 1.5).

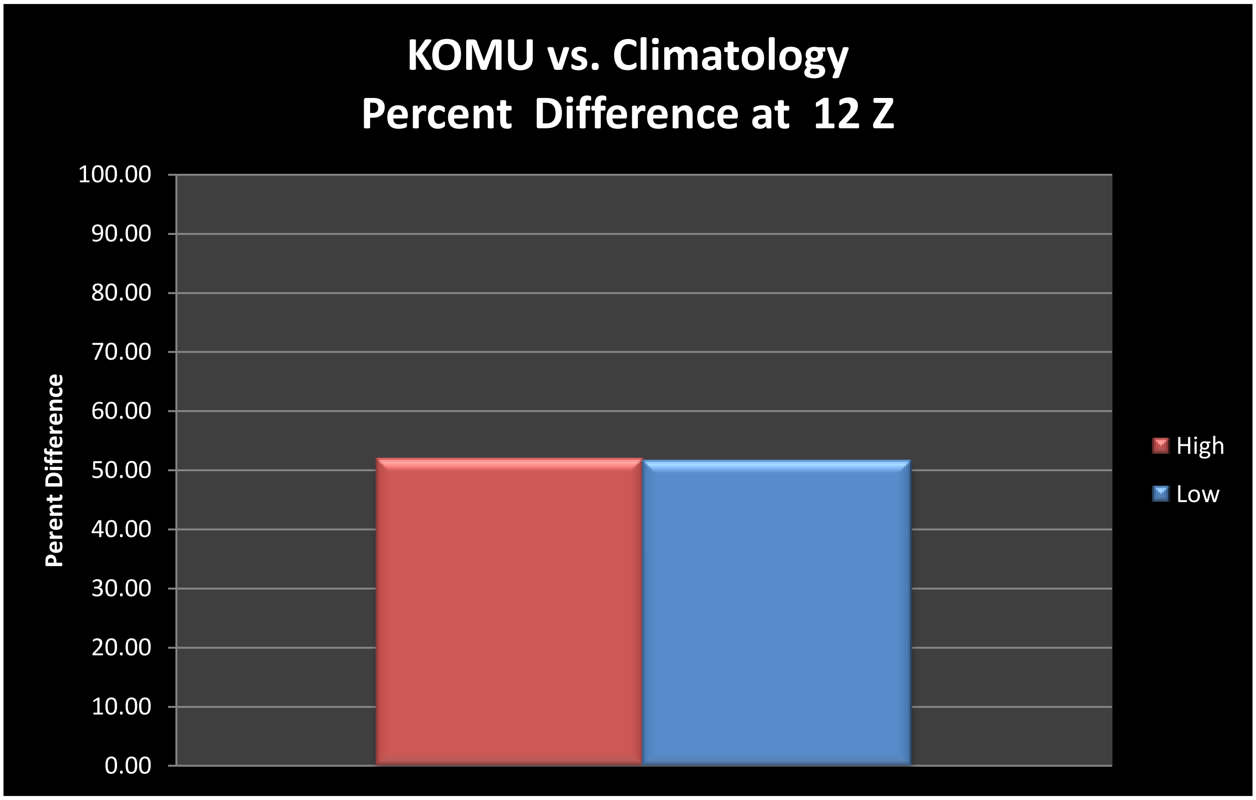

Figure 1.5 – Percent difference between KOMU forecast and climatology for 12z.

4.3 KOMU vs. NAM

As seen previously in Tables 1.6 and 1.7, when looking at the NAM, more often than not, the model performs best and is the most reliable model to forecast with. When looking at the results in Figure 1.6, notice the percent difference between the KOMU forecast and the NAM forecast has decreased significantly, in comparison to climatology.

Taking a look at the 00z data, the advantage that KOMU has over the NAM is ~10% for the high temperatures and for the low temperatures it drops to about ~8%. Even though the human forecaster still beats the model overall, the proximity of the numbers would imply that the NAM model was well engineered.

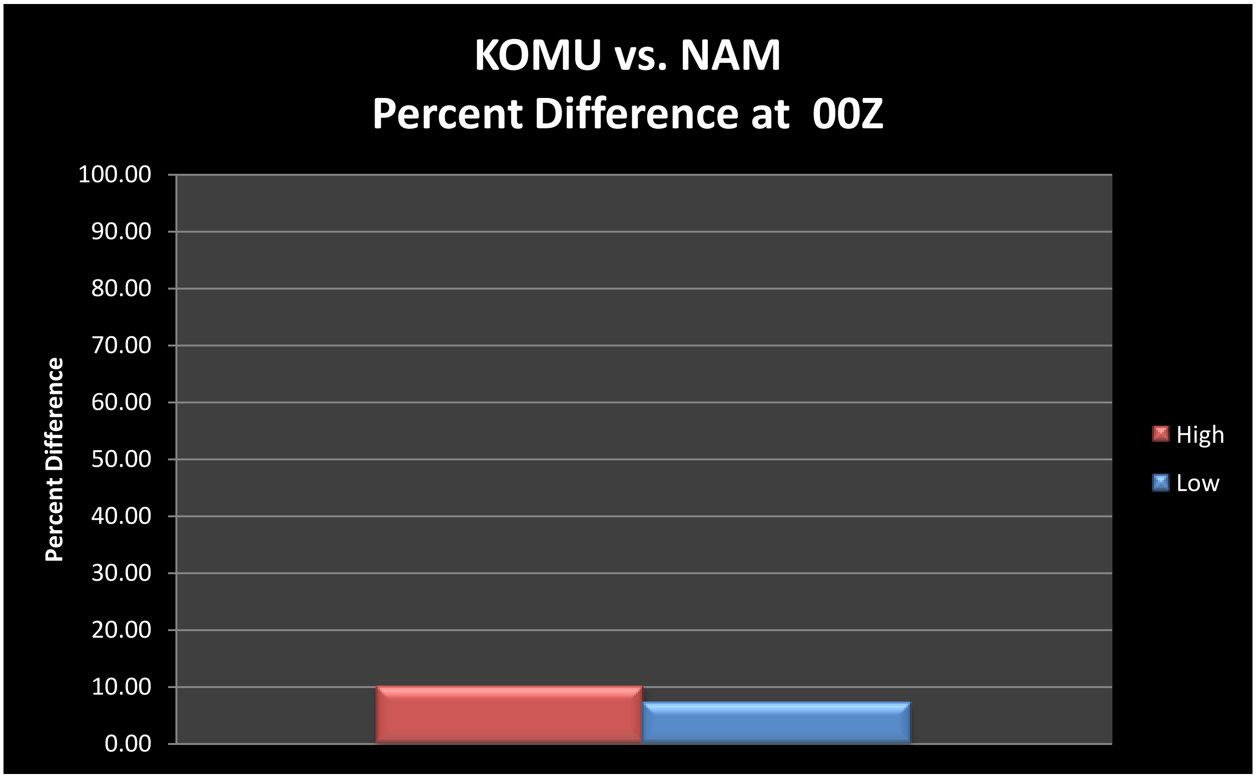

Figure 1.6 – Percent difference between KOMU forecast and the NAM at 00z.

When looking at the 12z NAM data in Figure 1.7, there was a slight increase in the percentage difference of KOMU over the model for both the high and low temperatures. The human forecaster's high temperature was ~11% more accurate than the NAM and the low temperature was ~16%. Just like climatology, the end result here is reversed. Where the high temperature percentage decreases and the low temperature increases.

The reversal of results is not surprising. To note here, the 00z model runs for the evenings are the numbers that the morning forecasters use to make their forecasts. The morning forecast is created prior to that mornings' latest model run, so it is using the data from the model run the evening before. The same goes for the evening forecaster. By the time the numbers come out in the early evening, the forecast had already been created using that mornings' model run from 12z. So, the 12z model runs that take place in the mornings are the numbers that the afternoon forecasters use when producing forecasts.

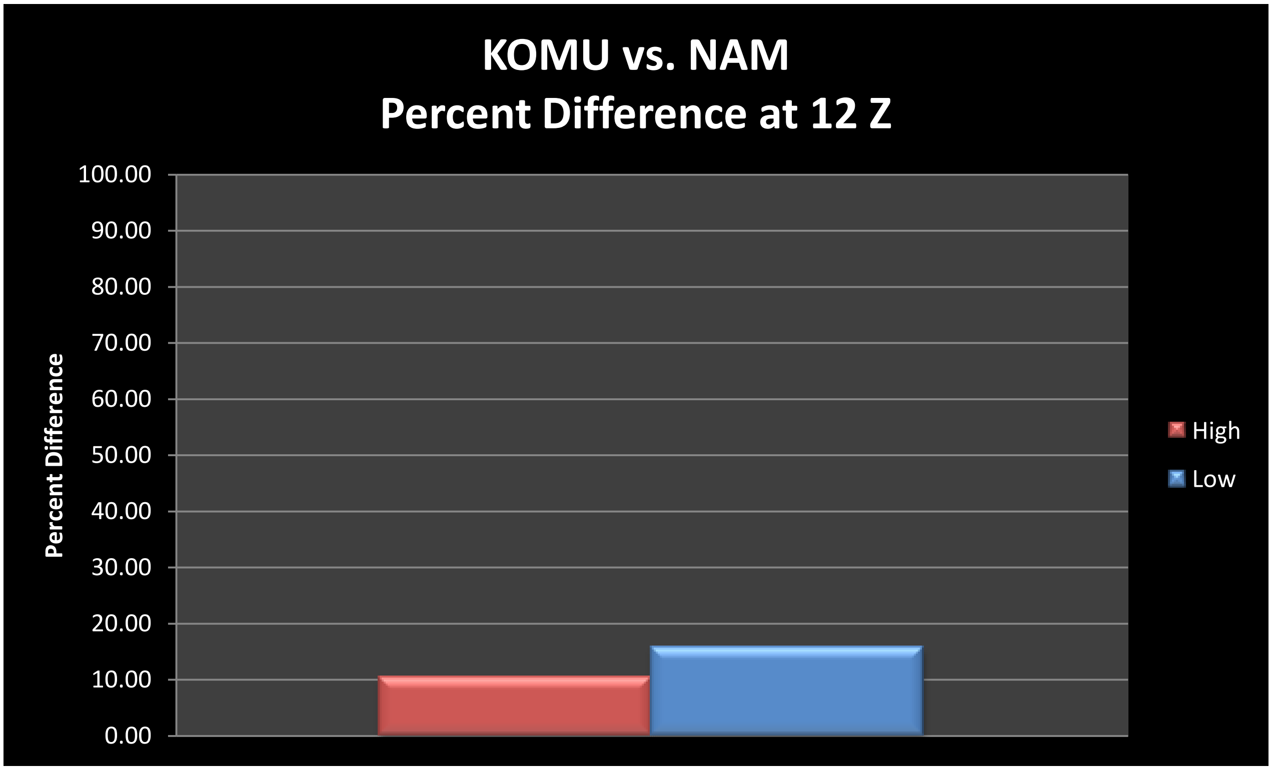

Figure 1.7 – Percent difference between KOMU forecast and the NAM at 12z.

This would explain the reversal of the results. The data from the model run (00z or 12z) that is closest to the initial forecast period is going to be the most accurate. So for morning forecasters, the day time highs are going to be more accurate than the overnight lows. For the evening forecasters, overnight lows are more accurate than the next day's high temperature. As a forecaster, we know, that the further out your forecast is, accuracy will be decreasing, even over a short period of time as in this case. For the 00z data used by the morning forecaster, that day's high temperature has a higher percentage of accuracy compared to that of the overnight low. For the 12z data used by the afternoon forecaster, the overnight low temperature will have a higher percentage of accuracy compared to that of the next days' high temperature.

4.4 KOMU vs. GFS

Overall, the GFS is the weaker model of the two used in this research project, but consistently beats climatology as seen in Tables 1.6 and 1.7. One will also take note of one instance where the GFS model outperforms the NAM in the 12z low temperatures. The results are as follows.

At 00z, KOMU outperforms the GFS model by ~23% for the high temperature but only ~8% for the low temperature as see in Figure 1.8. The high temperature percent difference is greater than the NAM model. The NAM model is the more accurate model of the two; therefore, the percent difference is much smaller.

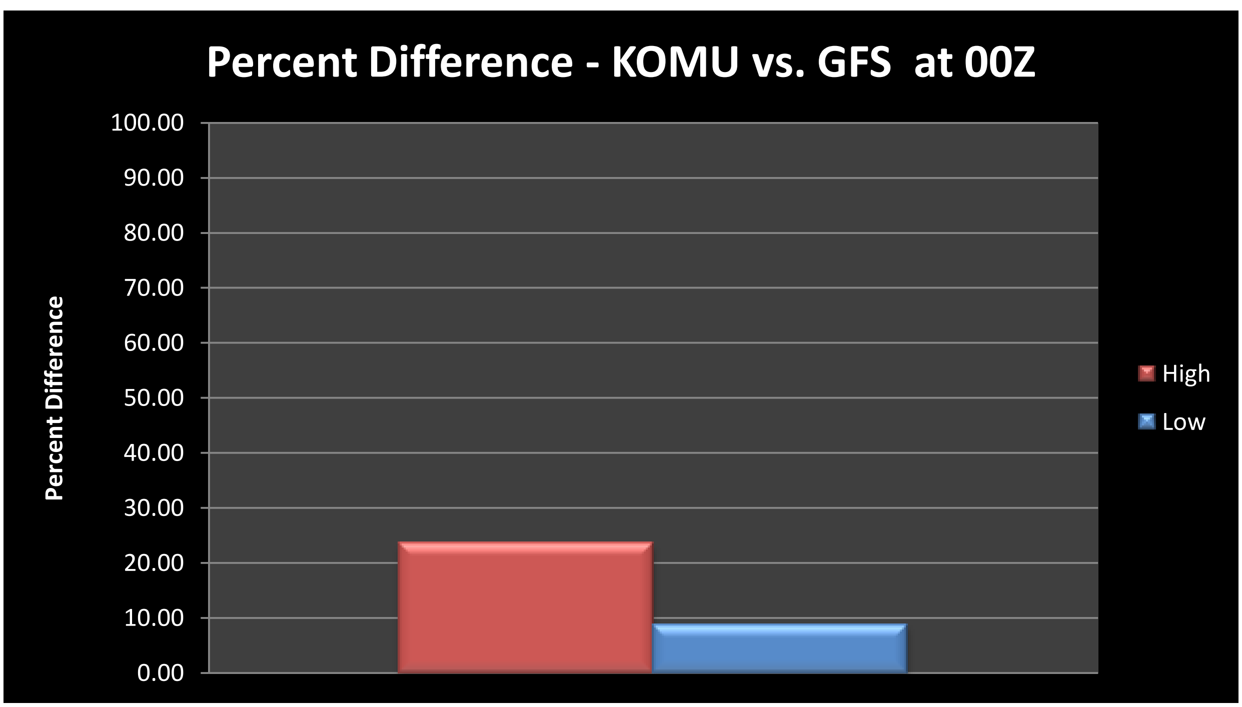

Figure 1.8 – Percent difference between KOMU forecast and the GFS at 00z.

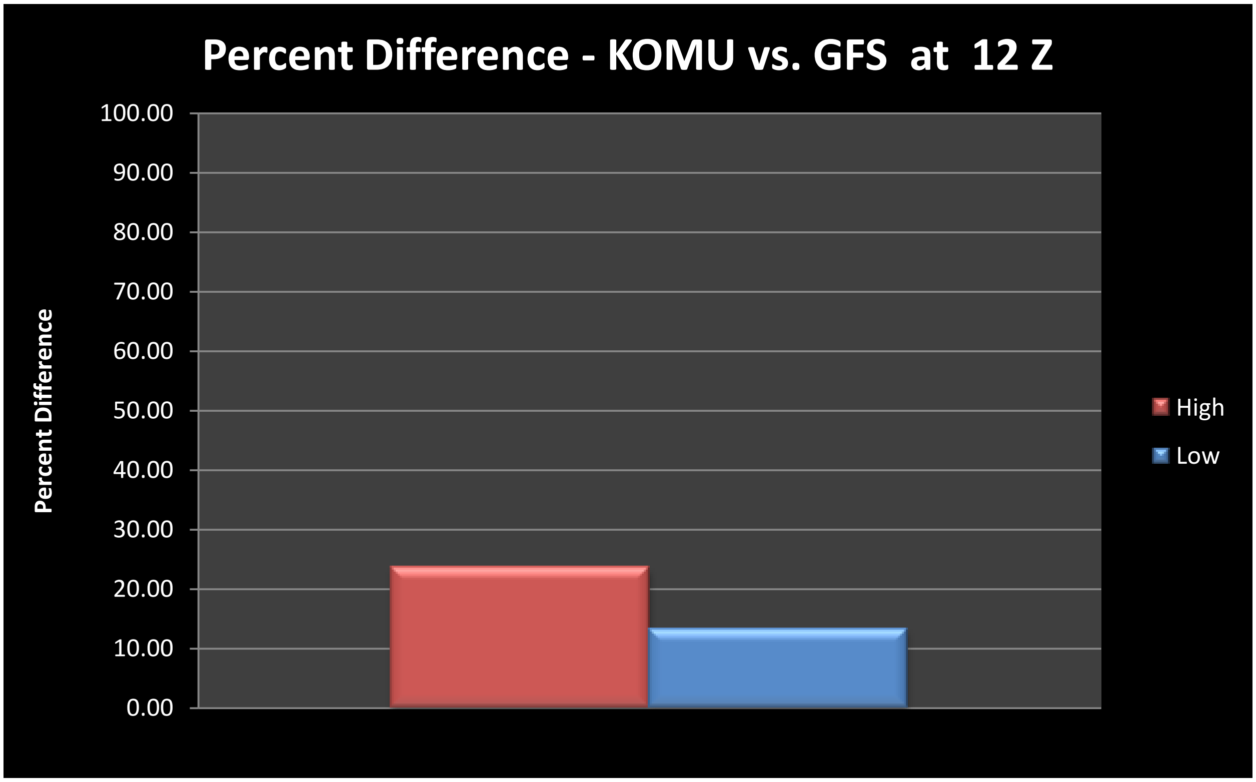

Figure 1.9 – Percent difference between KOMU forecast and the GFS at 12z.

When evaluating the 12z model run for the GFS, it was found the KOMU had a slightly larger percent difference over the model. There was not much of a change with the high temperatures projected by KOMU in comparison to the GFS. The percentage difference was still ~23%. But when looking at the low temperature, the percent difference increased with the 12z model run. It improved to a ~12% difference as seen in Figure 1.9.

After assessing the GFS model, the percent difference, between the high temperature projected by the models and what the human forecaster predicted, is more significant than that of the NAM model. However, when you compare either model in this research to climatology, the small significance that one of the models has over the other is overshadowed.

In the end, the most accurate forecast is created by the human forecasters at KOMU-TV 8 in Columbia, Missouri. As a forecaster, we are able to pick up on various changes in weather events that differ from the model data that is available, unlike the computer models and climatology. The initial conditions may vary significantly in a short time frame from what the model originally pulls in, thus throwing the data off. In this aspect, the human forecaster can analyze all characteristics of the atmosphere by using various tools in order to produce the most accurate forecast possible. However, in a handful of cases, the computer models had a better handle on a weather situation than the human forecasters did. This data will be shown in the next section. But more often than not, the human forecaster had the best results.

4.5 ENSO CYCLE RESULTS

Since the results are compared to climatology, take a look at the percent difference between KOMU and the baseline at 00z. The KOMU forecast outperformed climatology by a large margin for the varying ENSO cycles (between 40-75% differences) as seen in Figure 1.10. But will this margin be large enough to be significant? This will be discussed a little later in this research.

For the 00z timeframe, the morning forecaster, using the evening data from the night before, tends to have a better grasp on predicting the high temperature for the day in comparison to the overnight low. This is because the closer you are to the initialization of the model, the more accurate the forecast should be. (The first twelve hours should be more accurate than twenty four hours out.)

It is also important to look at the percentage differences between KOMU and the climatology. The larger the percent difference, the more variance seen from the baseline (climatology). For the La Nina period from 2007-2008, the percent difference that KOMU has over climatology is higher than the other ENSO cycle periods. This would signify that the conditions experienced during La Nina seem to deviate from the normal (climatology) due to the volatility of the La Nina phase compared to the Neutral or El Nino phases. Also, the high temperature forecasted by KOMU tends to be better than the overnight low temperature forecast. As stated previously, this is because the daytime high forecast period is closer to the model initialization than the overnight low temperature.

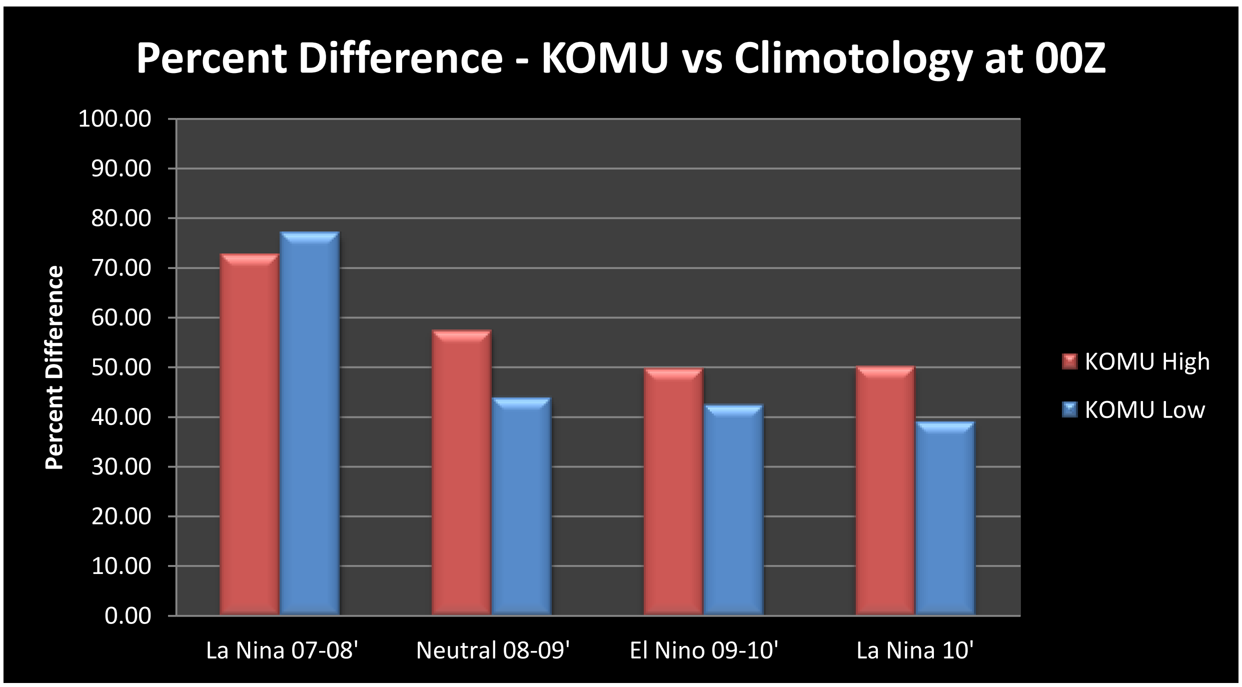

Figure 1.10 – Percent difference between KOMU forecast and climatology for ENSO cycles at 00z.

Moving on to the 12z data taken in the mornings, the data were used by forecasters making their predictions in the afternoon hours. There is an increase in the accuracy of the forecast for overnight low temperatures. The KOMU forecast out performed climatology by a large margin yet again but the percentage difference has dropped significantly (between 35-60% differences) as seen in Figure 1.11. El Nino, is the most accurate ENSO cycle with the temperatures deviating the most for our baseline, climatology.

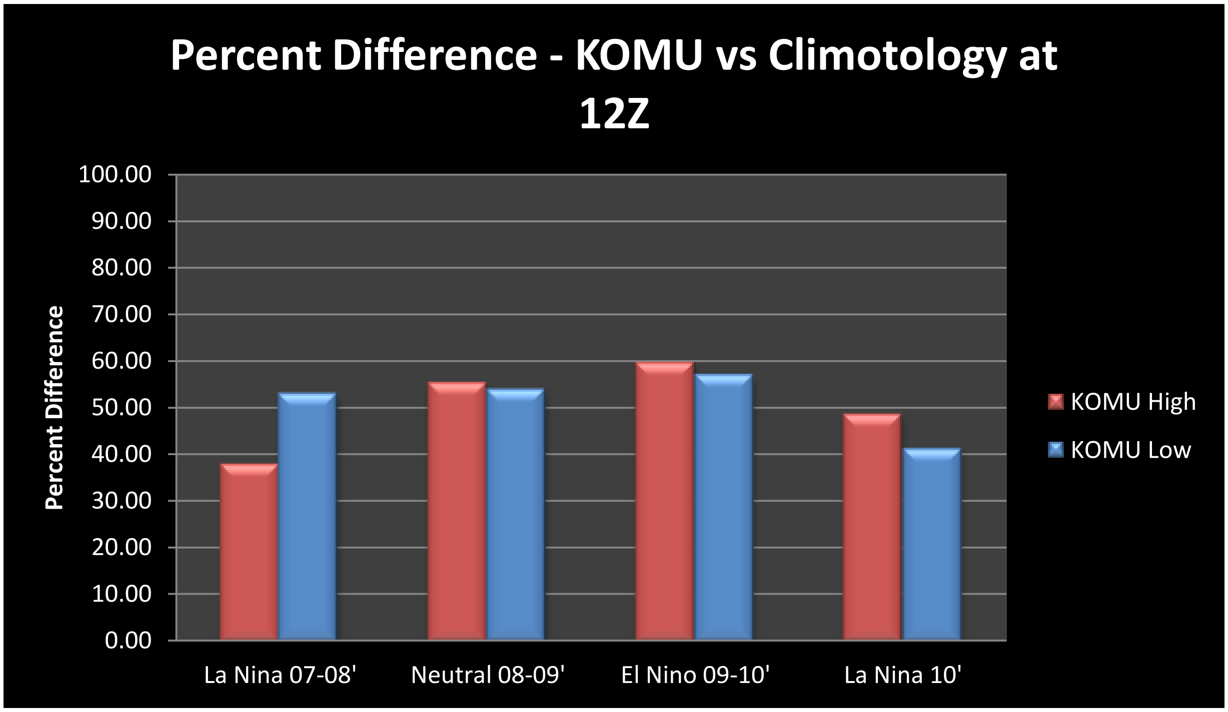

Figure 1.11 – Percent difference between KOMU forecast and climatology for ENSO cycles at 12z.

In order to determine accuracy during the varying ENSO cycles, the human forecaster, the NAM, the GFS, or climatology, we look to Figure 1.12. In this figure, the skill scores that were calculated for the high temperatures at 00z are shown. KOMU has the best skill score in comparison to the model data and climatology. One will notice, trailing not too far behind the KOMU (human) forecast is the NAM model, with the GFS trailing. Climatology has the lowest skill score. Overall, the scores tend to vary for the different ENSO cycles. But as we will see in the 00z low temperature results, they remain rather consistent with one another.

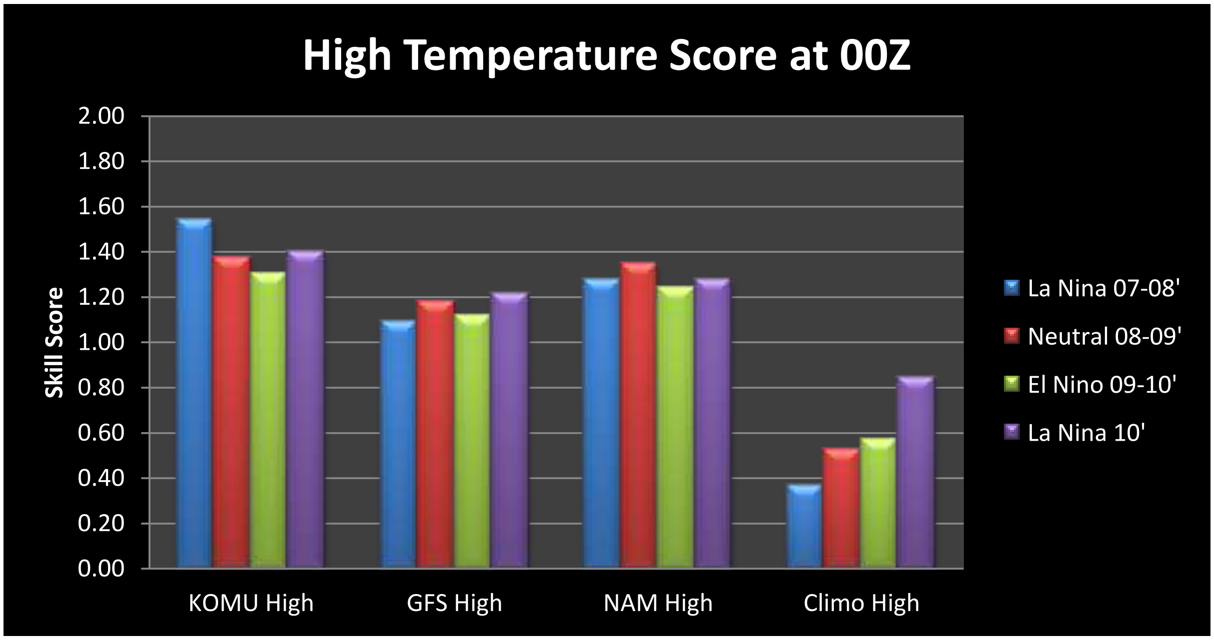

Figure 1.12 – ENSO cycle skill scores for high temperatures at 00z.

Figure 1.13 shows the low temperature skill score for the given ENSO cycles. As stated above, they do not vary nearly as much, but where they do, it's rather significant. KOMU dominates with forecast accuracy, especially with La Nina conditions from 2007-2008 boasting a 1.64 skill score. Even with the changing of the ENSO cycles, there are not many variations in the skill scores for the 00z time frame.

Now let us turn our attention to the 12z time frame high temperatures show in Figure 1.14. Even though climatology remains in last place, there are some interesting results between the human produced forecast at KOMU and the two models, NAM and GFS, in this research. The NAM has gained a competitive edge just barely falling short of the KOMU skill score in three of the four ENSO cycles listed: La Nina 07-08', Neutral 08-09', and El Nino 09-10'. The GFS model improved in accuracy outperforming both the NAM and KOMU's high temperature forecast during the La Nina ENSO cycle from 2007-2008.

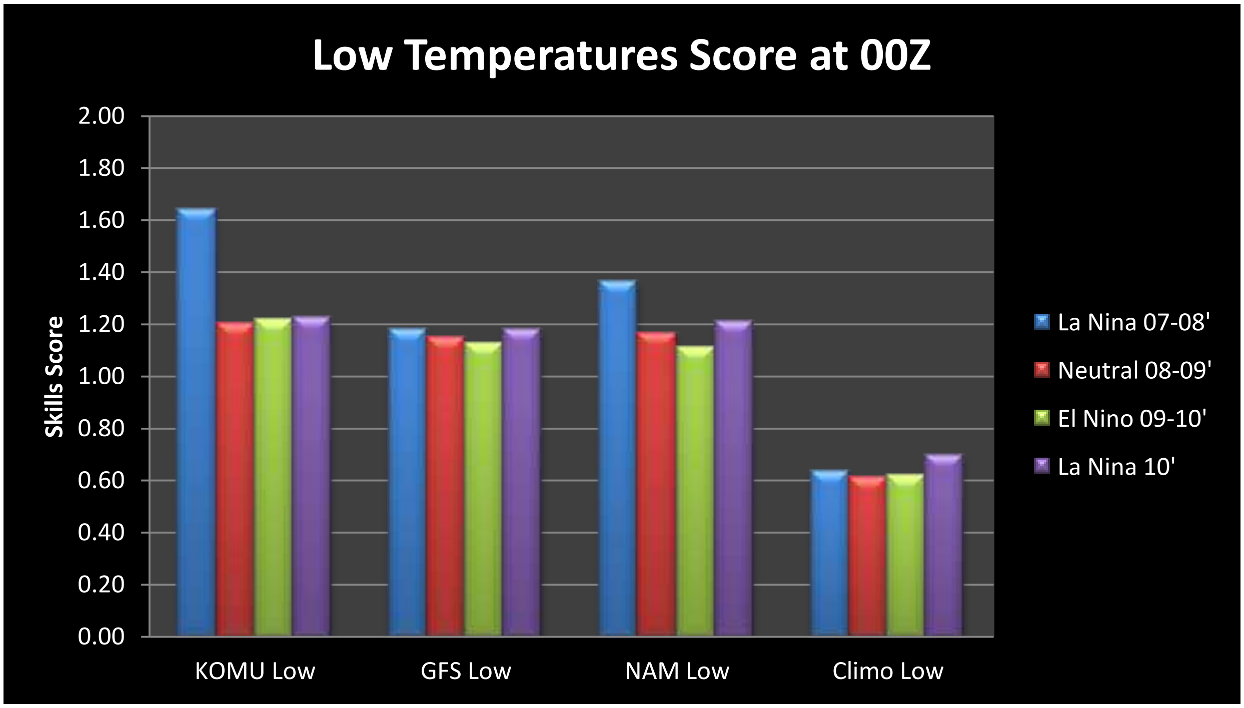

Figure 1.13 – ENSO cycle skill scores for high temperatures at 12z.

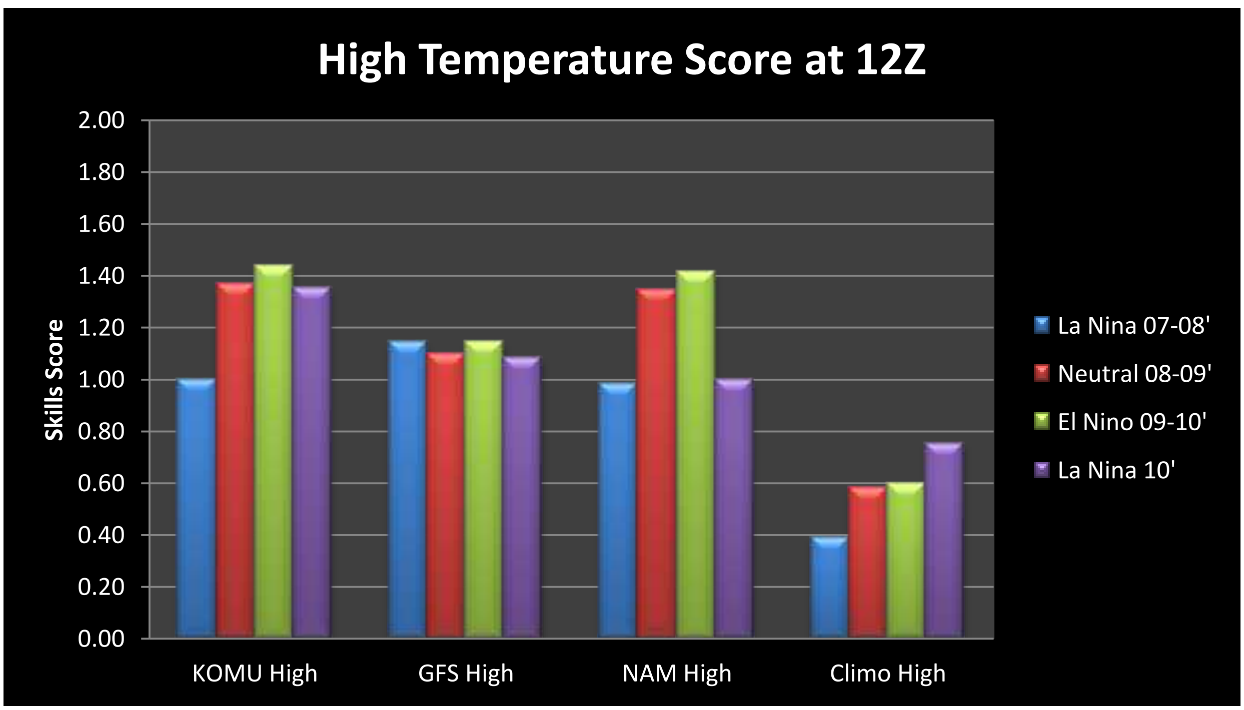

Figure 1.14 – ENSO cycle skill scores for high temperatures at 12z

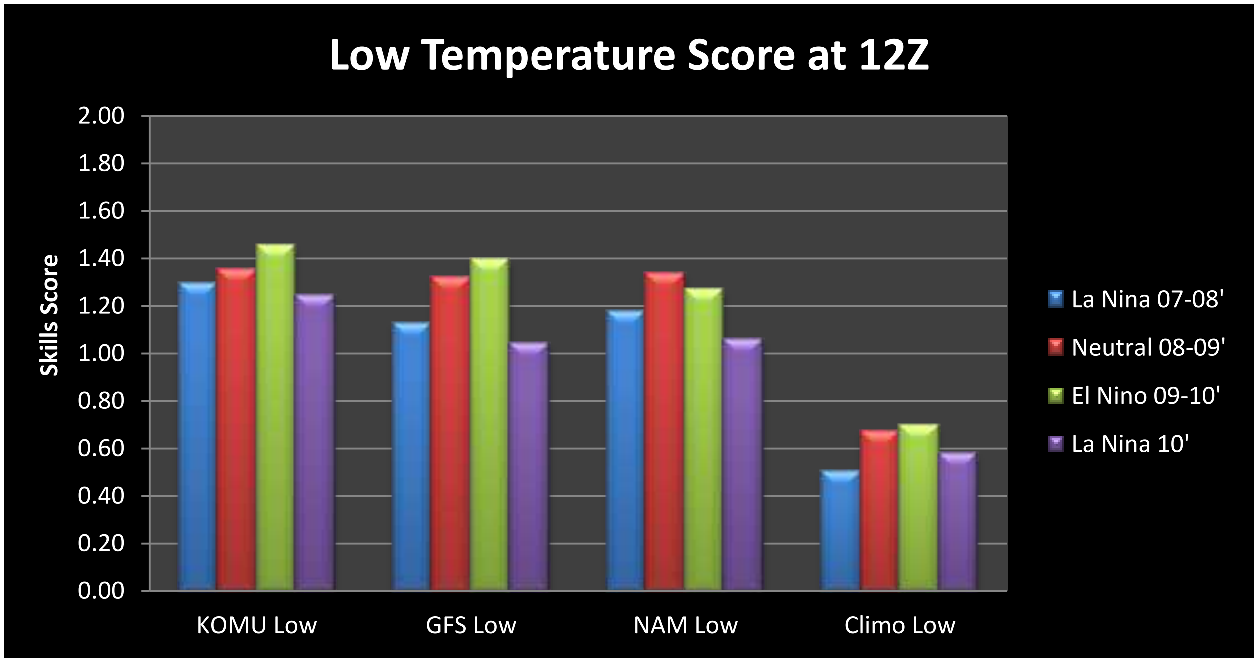

Figure 1.15 – ENSO cycle skill scores for low temperatures at 12z

Unlike in Figure 1.13, Figure 1.15 shows a little more variation with the skill scores for the different ENSO cycles. Again, the KOMU forecast just barely edges out both the NAM and GFS models for all four ENSO cycles in this research. One interesting thing to note here is that during the Neutral 08-09' phase and El Nino 09-10', GFS outperformed the NAM model which has consistently been the more accurate of the two models. Climatology falls behind KOMU and the models.

From these results, conclusions can be drawn about who performs best during the different ENSO phases. During the 00z time frame, KOMU performs best during La Nina conditions, both for high and low temperatures. However, for the 12z time frame, KOMU has a better handle on the high and low temperatures during El Nino. For the 12z time frame, the NAM model does best during Neural and El Nino cycles. The GFS also performs quite well at 12z for those cycles as well; however, the results are slightly closer. Lastly, the climatology forecasts are by far the most inaccurate.

This is consistent throughout this research. These are the only significant differences that stand out from the data when looking at the skill scores.

Upon completion of this data, the real question is now whether or not the differences found in this research project are significant. Utilizing statistical methods, we were able to find the standard deviation and confidence intervals. This helped to determine if the data was significant on the 95% confidence level or the 90% confidence level. The results are as follows.

	KOMU High	KOMU Low	GFS High	GFS Low	NAM High	NAM Low
La Nina 07-08'	**1.55/0.69***	1.64/0.50	1.09/0.83	1.18/0.75	1.27/0.79	1.36/0.81
Neutral 08-09'	1.38/0.80	1.20/0.84	1.18/0.84	1.15/0.84	1.35/0.80	1.16/0.86
El Nino 09-10'	1.31/0.81	1.22/0.86	1.12/0.84	1.13/0.82	1.25/0.84	1.11/0.89
La Nina 10'	1.40/0.79	1.22/0.84	1.22/0.87	1.18/0.87	1.28/0.85	1.21/0.88

Table 1.8 –The mean forecast score and standard deviation with respect to ENSO cycles 00z data

*denotes the forecast was better than climatology at the 90% confidence level.
**denotes the forecast was better than climatology at the 95% confidence level.

Tables 1.8 and 1.9, denote the mean skill score and standard deviations found for the varying ENSO cycles associated with the KOMU forecast, the GFS and the NAM. The data were all tested for the confidence intervals of 90% and 95% to determine whether or not the results were significant. Looking at the data set, not a single piece of data tested at the 95% significance level. However, there was one piece of data that did score at or above the 90% significant level. The data were denoted with a star.

While these results may seem rather disappointing, one must refer back to Lupo and Market (2002) to find that the results are rather typical for temperature analysis. In their research, most all of the temperature data points did not reach the 90% confidence level, let alone the desired 95% confidence level. What also needs to be taken into account is the fact that in Lupo and Market (2002), there were more meteorological variables than just temperature included in their study. These variables encompassed precipitation, cloud cover, wind speed/direction, ect. This helped to increase their overall performance with forecasting.

	KOMU High	KOMU Low	GFS High	GFS Low	NAM High	NAM Low
La Nina 07-08'	1.00/0.85	1.30/0.81	1.14/0.91	1.13/0.86	0.98/0.88	1.17/0.88
Neutral 08-09'	1.36/0.80	1.36/0.81	1.10/0.85	1.32/0.83	1.34/0.79	1.33/0.79
El Nino 09-10'	1.43/0.74	1.45/0.72	1.14/0.87	1.39/0.79	1.42/0.78	1.27/0.83
La Nina 10'	1.36/0.77	1.25/0.80	1.08/0.92	1.04/0.86	1.00/0.87	1.05/0.93

Table 1.9 – The mean forecast score and standard deviation with respect to ENSO cycles 12z data

*denotes the forecast was better than climatology at the 90% confidence level.
**denotes the forecast was better than climatology at the 95% confidence level.

4.6 SEASONAL ENSO CYCLE RESULTS

Different ENSO cycles produce different conditions across the United States. Sometimes it enhances weather events, such as extreme heat in the summer, and other times it downplays them, such as mild conditions in the winter. Regardless, the varying ENSO cycles do play a role in forecast accuracy. When there are extremes, climatology, the models, and even human forecasters have difficulty with forecasting accurately.

The 00z time frame and the skill scores for the ENSO cycles is shown in Figure 1.16. You begin to notice a roller coaster like pattern with the KOMU forecast. Aldrich (2011) explains that the human forecaster tends to be a little more accurate during the transitional seasons of fall and spring. Figure 1.17 agrees with his findings as well; however, there does not seem to be any significance of one ENSO cycle over another. Another interesting fact to note here is that climatology has a similar roller coaster like pattern as well. Climatology tends to be more accurate during the summers and decreases in accuracy during winter, particularly in the La Nina cycle.

In Figure 1.17, the 00z skill scores for the seasonal ENSO cycles are shown for the low temperatures. Here, one will notice that the weakest low temperature forecast period for KOMU, the GFS, the NAM, and climatology is during the Neutral winter from 08-09'. Again, climatology appears to be weakest during the winter months and gradually increases reaching its accuracy peak during the summer months.

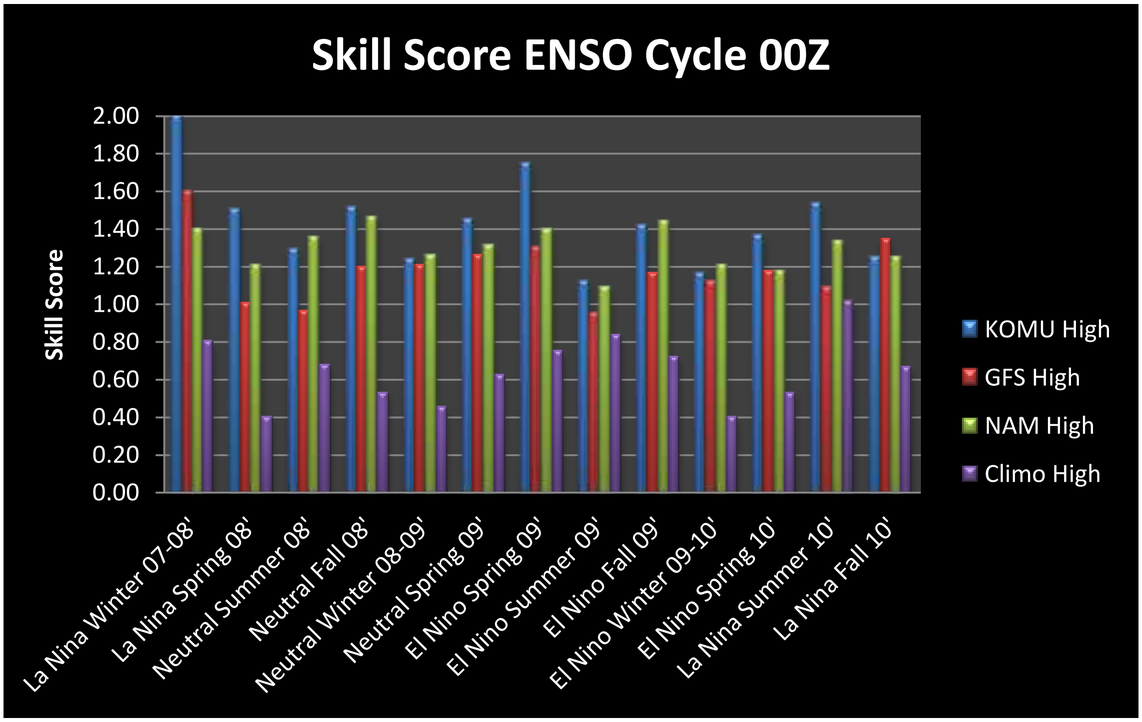

Figure 1.16 – Skill Scores for the seasonal ENSO cycles high temperatures at 00z.

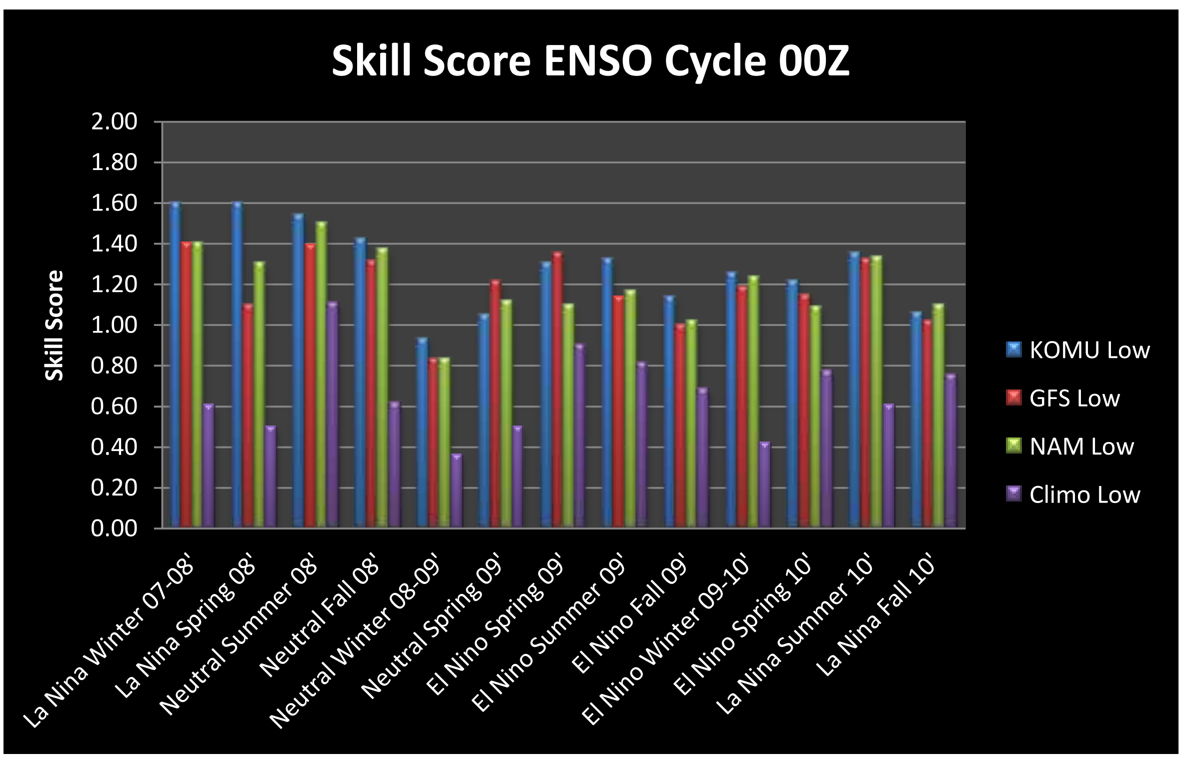

Figure 1.17 – Skill Scores for the Seasonal ENSO cycles low temperatures at 00z.

Moving on to Figure 1.18, the skill score for 12z high temperatures is shown. The most accurate forecasting takes place during El Nino conditions in the spring, summer, and fall. By winter, forecast accuracy begins to drop off. Both KOMU and the NAM's skill scores are very close during those months. Even climatology increases during this period as well. This is most likely due to the fact that while in an El Nino cycle, there is less volatility than you would see during La Nina conditions. This would allow for conditions to be a little closer to average for the daytime high temperatures.

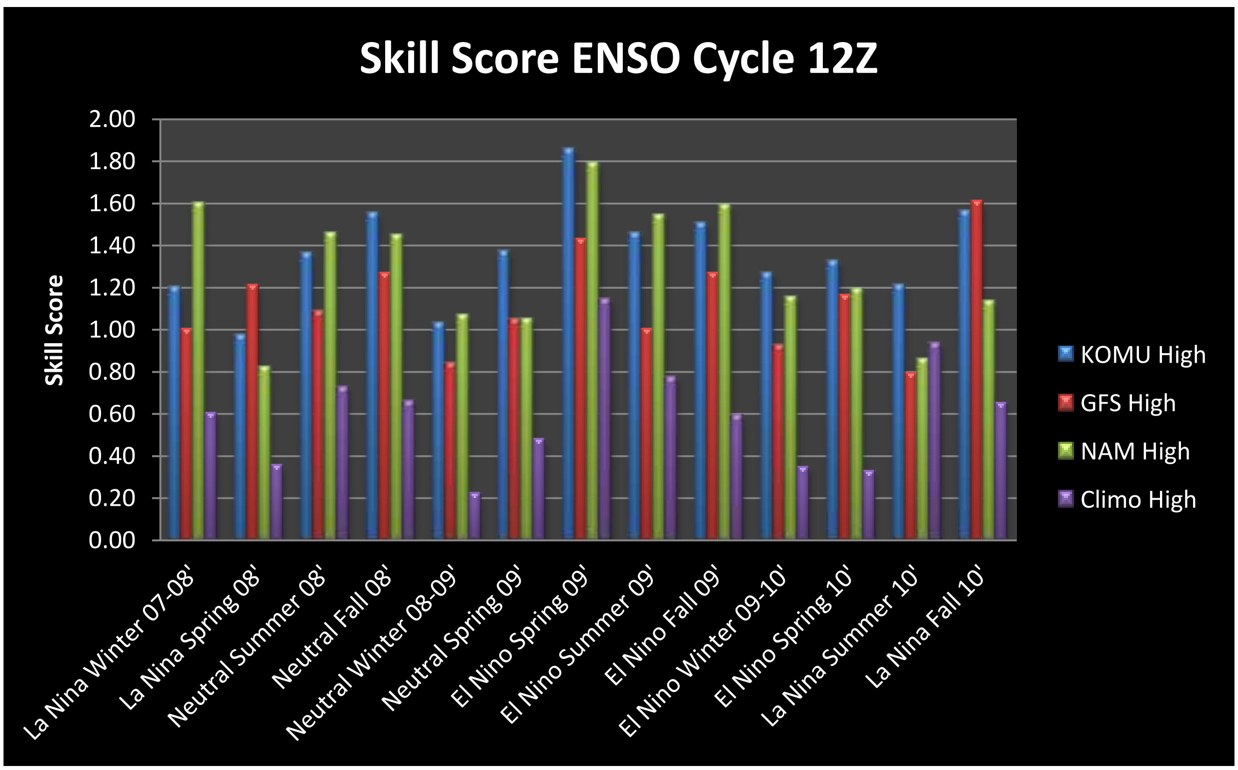

Figure 1.18 – Skill Scores for the Seasonal ENSO cycles high temperatures at 12z.

Figure 1.19, takes a look at the low temperatures predicted during the 12z time frame for the given seasonal ENSO cycles. Interestingly enough, the low temperatures during El Nino conditions for spring and summer tend to be the most accurate, with the Neutral conditions during summer coming up close behind. This is very similar to the findings seen above in Figure 1.18. However, this time instead of the NAM competing with KOMU, it is the GFS model that competes. You will notice a rather large spike in the skill score in these seasons during the El Nino/Neutral cycles. KOMU continues to outperform with the GFS following closely behind.

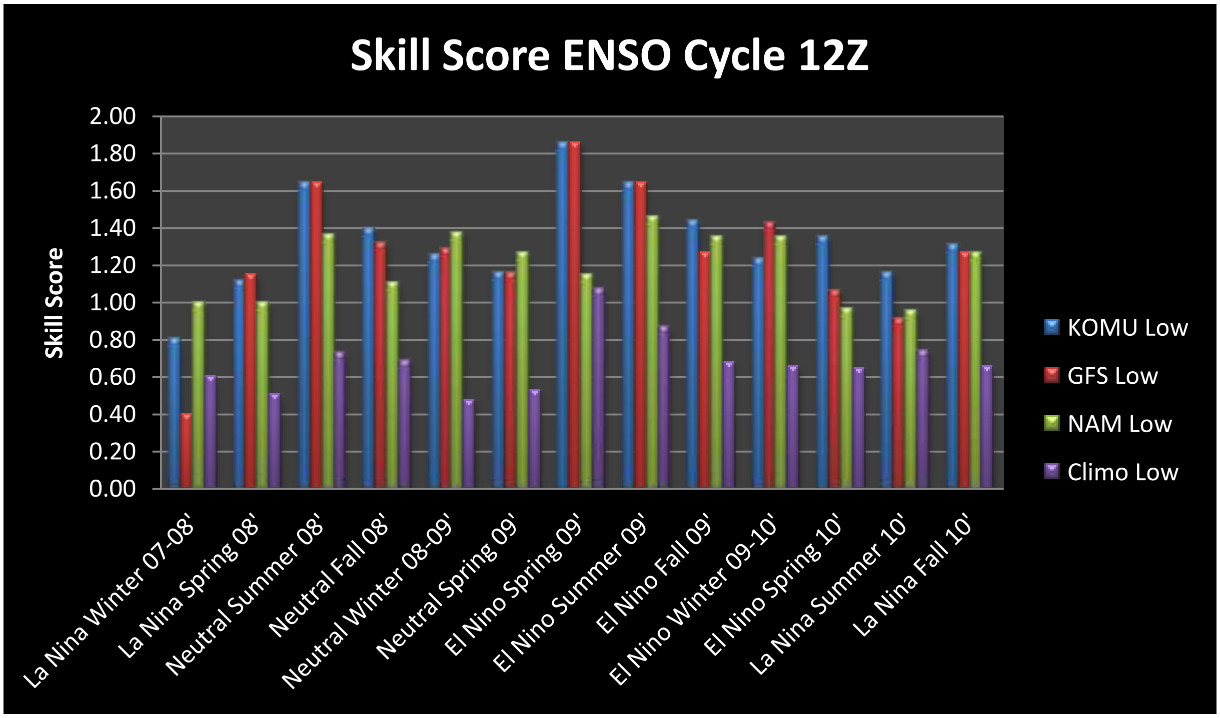

Figure 1.19 – Skill Scores for the Seasonal ENSO cycles low temperatures at 12z.

Unlike during 00z, climatology does not have a consistent pattern that it follows. But, we see a consistent spike during the spring and summer months during El Nino. As mentioned above, this is most likely due to the consistency of temperatures during this time. La Nina conditions are more volatile, and temperatures vary. Therefore, there is less consistency and the skill score for climatology suffers as a result.

5 CONCLUSIONS

5.1 SUMMARY & CONCLUSIONS

Over a three year time span, from November 2007 to November 2010, the staff at KOMU-TV 8 in Columbia, Missouri observed and collected data used in on-air forecasts. Forecast records of the projected high and low temperatures were recorded by staff meteorologists and University of Missouri atmospheric science students. Aside from the KOMU numbers, the data used were MOS data from two computer models, the NAM and GFS, along with climatological data acquired from the National Weather Service office in St. Louis, Missouri. The findings in this research project show the necessity and usefulness of the human forecaster. By far, the human produced forecast is the most accurate over the NAM, GFS, and climatology as seen in Table 1.6 and 1.7.

The NAM and GFS models along with climatology, used in this research project, are tools for the human forecaster to use as guidance. With this guidance, forecasters are able to use the knowledge they have and question this data. This allows them to come to the conclusions they need in order to determine the forecast. Understanding biases in the models increases the accuracy of the individual forecaster.

From this research data, conclusions can be drawn in order to help improve forecasting skills for the future. The first twelve hours of a forecast from a given model run is going to be the most accurate. The further out you go, the accuracy will begin to decrease rather significantly. For example, a morning forecaster is going to be more accurate predicting that day's high temperature as opposed to the overnight low temperature. The afternoon forecaster is going to be more accurate with the overnight low temperature than they will the next day's high temperature.

As far as accuracy goes, KOMU forecasters are the most accurate overall. When looking at the models, the NAM performs better than the GFS model. The least accurate is climatology. Which is the baseline used throughout this research project.

When analyzing the ENSO cycles during the 00z time frame, it was found that KOMU performs best during La Nina conditions. This includes both the high and low temperatures. But during the 12z time frame, KOMU has a better grasp on predicting high and low temperatures during El Nino. The NAM model does best during Neutral and El Nino cycles at 12z. Lastly, the GFS also performs at its best for those cycles at 12z as well; however, the results are slightly closer than that of the NAM.

Going a step further, when breaking down the ENSO cycles into seasons helped to show when exactly the models were performing at their best, and when they were at their worst. During 00z, the seasons of spring and fall tend to have the most accurate forecasts. There is no significance though of a particular ENSO cycle performing better during this time. Climatology at 00z, also yielded some interesting results. During the summer months, climatology was more accurate than any other season, particularly in the La Nina cycle.

When looking at the 12z data, a different pattern emerges. The most accurate forecasting takes place during El Nino conditions in the spring, summer, and fall.

Even climatology improves during this time period. The NAM is also found to be more competitive with KOMU during this time for high temperatures. The opposite results though for the low temperatures. The GFS is the stronger model

The findings of this research continue to enforce the idea that human forecasters are needed if you want the most accurate forecast possible. One should never put their complete trust in the models nor climatology. It is important to have human interaction with this data. Nowadays, people rely heavily on accurate forecasts to be able to plan their days. My hope is that this research helps to point us in the right direction, helping us to further our knowledge of the available resources we have. This will then help the human forecaster to continue to produce the most accurate forecast possible.

REFERENCES

Ahrens, C.D., 2003: *Meteorology Today: An introduction to weather, climate, and the environment*. Pacific Grove, CA: Brooks/Cole – Thomson Learning. 554 pp.

Aldrich, E., 2011. *The verification of weather forecasts compared to numerical model guidance, climatology, human forecasts and seasonal variations*. A thesis in the Department of Soil, Environmental, Atmospheric Science. The University of Missouri, Columbia, MO.

Birk, K.A., A.R. Lupo, P.E. Guinan, and C.E. Barbieri, 2010: The interannual variability of Midwestern temperatures and precipitation as related to the ENSO and PDO. *Atmofera,* **23***,* 95 - 128.

Bjerkness, V., 1914: *Meteorology as an Exact Science*. Mon. Wea. Rev., 43, 14.

Franklin, B., 1774: *Experiements and Observations on Electriciy*. Reprinted 1941. *Benjamin Franklin's Experiements,* I.B. Cohen, Ed., Harvard University Press, 453 pp.

Lupo, A.R., Kelsey, E.P., D.K. Weitlich, I.I. Mokhov, F.A. Akyuz, Guinan, P.E., J.E. Woolard, 2007: Interannual and interdecadal variability in the predominant Pacific Region SST anomaly patterns and their impact on a local climate. *Atmosfera,* **20,** 171-196.

Lupo, A.R., Market, P.S., 2002: *The Application of a Simple Method for the Verification of Weather Forecasts and Seasonal Variation in Forecast Accuracy*. Wea. Forecasting, 17, 891-897.

Vasquez, T. 2008. *Weather Map Handbook*. Garland,TX: Weather Graphics Publisher.

Woolard, E.W., 1922: *L.F. Richardson on Weather Prediction By Numerical Process*. Mon. Wea. Rev., February, 1922. 72.