Design and development of a data warehouse for atmospheric conditions

RAÚL GUTIÉRREZ, DR. JOSÉ TORRES, DR. MICHEL ROSENGAUS, DR. RENÉ LOBATO
Computer Department
ITESM Campus Cuernavaca
Av. Paseo de la Reforma No. 182-A
Col. Lomas de Cuernavaca, CP 62589, Temixco, Morelos.
MEXICO

Abstract: - This paper considers the track error of the current tropical cyclone track forecast, and takes on an explorative approach to the problem. We took the elements that provide a hurricane with motion and direction, and built a data warehouse with those atmospheric conditions. The data were coded in a binary format called GRIB, which is a standard for meteorologists, we decode and shape them according to our needs and design a multidimensional model. This data warehouse is currently at use in the design process of a new track forecast technique taking the wind as main pattern.

Key-Words: - Data warehouse, hurricane, tropical cyclone, forecasting, track, wind, gridded data.

1 Introduction

Of all natural disasters, tropical cyclones are the most dangerous threat to life and property in the sea side areas in many parts of the world. They represent a very important matter for Mexico, each year, is affected in his coasts. Hurricanes are cyclones that develop over the warm tropical oceans, they are powerful storms that form at sea with wind speeds of 74 miles per hour or greater. Hurricanes are spotted and tracked by satellites from the moment they start to form. These severe storms cover a circular area between 200 and 480 miles in diameter. In the storm, strong winds and rain surround a central, calm eye, which is about 15 miles across. Winds in a hurricane can sometimes reach 200 mph. With hurricanes being as powerful as they are, at landfall they cause damage and destruction. However, most of the damage caused to man and nature occur as a hurricane makes landfall [1].

Atmospheric pressure and wind speed change across the diameter of a hurricane. The pressure begins to fall more rapidly while the wind speed simultaneously increases. Within the eye wall, the wind speed reaches its maximum. The surface pressure continues to drop through the eye wall and into the center of the eye, where the lowest pressure is found. Upon exiting the eye, the wind speed and pressure both increase rapidly. The wind speed again reaches a maximum in the opposite eye wall, and then quickly begins to decrease. The wind and pressure profiles inside a hurricane are roughly symmetrical, so a quick rise in winds and pressure through the eye wall followed by a slower increase in pressure and likewise decrease in wind speed would be expected [2].

The vertical wind shear in a tropical cyclone's environment is important. Wind shear is defined as the amount of change in the wind's direction or speed with increasing altitude.
When the wind shear is weak, the storms that are part of the cyclone grow vertically, and the latent heat from condensation is released into the air directly above the storm, aiding in development. When there is stronger wind shear, this means that the storms become more slanted and the latent heat release is dispersed over a much larger area [2].

The path of a hurricane mostly depends upon the wind belt in which it is located. Strong environmental steering winds throughout the entire atmosphere or the lack of any steering winds altogether, will determine the track of a tropical cyclone.

The present work on tropical cyclone track prediction includes two major subjects, addressed in more detail below: 1. the assembling of a data set with variables needed for a model; 2. the design of a track prediction model on the basis of wind circulation.

The purpose of this paper is to present the results of a research work in tropical cyclones forecast. The content of the research was to develop a data warehouse containing atmospheric conditions, relative to the guidance of tropical cyclones motion, for a track forecast model.

2 Problem Formulation

2.1 Forecast difficulty

The formation of tropical cyclones takes place mainly over oceans where there are a few meteorological stations. As a consequence, information on the storm and its environment is difficult to obtain and often track prediction models may fail even in providing early track. These models attempt to accurately predict the environmental steering and forecast movement of these severe weather events.

A number of different models of varying degrees of complexity are currently in use to predict the future path that a hurricane is likely to traverse. All regional hurricane models except for CLIPER use output from global forecast models. The global outputs are available around 4h after synoptic time, whereas the NHC (National Hurricane Center, Miami, Fl., USA) issues advisories at 3h after synoptic time. To alleviate this difficulty, some models are used with the global forecast fields from the previous run ('early'), whereas others ('late' models) wait until the recent global forecast is complete before running. Currently, the early models run 4 times per day, while the late models run twice per day [3].

Hurricane forecasters analyze large amounts of data, including conflicting computer model results, and come up with their best estimate of a three-day track and intensity forecast. At shorter forecast time periods, the forecast track error is fairly small, but when the forecast is farther in the future, the error increases significantly.

The model forecasts are evaluated by comparison with the best track positions and intensities, which are parameters based upon all available information.

Track forecast errors are determined as the great circle distance between a forecast position and a best track position for the same time. Hurricane models have increased in accuracy greatly since aircraft reconnaissance began, however 72 hour forecasts still have a forecast error of about 250 nautical miles (Table 1) [3].

	Forecast Interval (hr)				
Model	12	24	36	48	72
CLIPER	51	103	161	220	351
NHC90	46	85	129	180	285
BAMS	61	114	168	222	336
BAMM	49	91	133	177	268
BAMD	47	88	132	183	293
LBAR	41	75	111	159	284
GFDI	42	69	98	128	200
No. Cases	346	310	279	255	207

Table 1. Average Errors (nm) of the Early Track Models for Atlantic Tropical Cyclones

In the last three decades, NOAA (National Oceanic & Atmospheric Administration) has increased the accuracy of the forecasted track of a tropical storm three days in advance of possible impact from 450 nautical miles to 225 nautical miles. With the latest advancements to the AVN tracking model, that margin of error was further reduced to 120 nautical miles with Hurricane Michelle [4].

Given the complex nature of producing forecasts of hurricane motion, it has been essential to develop models to provide track forecasts. It is important that public confidence in these forecasts is high, and the reliability of the model forecasts is of course critical in a situation of potential landfall, when decisions on preparation for damage limitation and evacuation must be made.

The NHC track forecasts based on the numerical models are improving each year, however, is still limited. With the advancement in computer technology in recent years, forecasting for these storm systems has become much more effective and accurate. However, there is still much room for improvement. Complex numerical models can be represented in computer programs and provide forecasts that draw upon many different aspects of the atmosphere. Continued modifications to these models are taking place as meteorologists learn more about these complex and systems.

Significant improvement is required in the assimilation of data and the model initialization. Traditionally, hurricane models are run from the global analyses, which are interpolated onto the regional model domain. A small number of parameters that specify the size, strength and shape of the hurricane are presently used; however, increasingly abundant data from aircraft and satellites still needs to be implemented. Other areas for improvement are the detection of new sources for forecasting models develop.

2.2 Data resources

Unfortunately, strong atmospheric data has never been easy to acquire. Data are available from a few organizations and services, and knowing where to look in order to retrieve a particular set of records is, to put it nicely, complicated.

A few years ago the National Centers for Environmental Prediction and the National Center for Atmospheric Research of USA, begin a project called "Reanalysis", that is an effort to reanalyze historical data using state-of-the-art models. In other words, this project focuses on reanalyzing all the available atmospheric data, to produce accurately data. The first datasets from the Reanalysis Project are now available.

This datasets are in GRIB format. GRidded In Binary is a WMO (World Meteorological Organization) format for gridded data. GRIB is used by the operational meteorological centers for storage and the exchange of gridded fields. GRIB's major advantages are files are typically 1/2 to 1/3 of the size of normal binary files (floats), the fields are self describing, and GRIB is an open, international standard [5].

These files contain variables relative to every atmospheric condition. The data that is now available is more accurate than other years. The existing dataset is from 1960 to 2000. The reason is that until 1960 the use of weather satellites starts.

The data in the Reanalysis contains measures collected all over the world. It is produced four times a day, at 0, 6, 12, 18 hours Greenwich Meridian Time, every day of the year, from 1960. The grain in the data is about every degree in the latitude and longitude, and at different levels of the atmosphere. The amount of data is huge, that's why it is coded in a binary format (GRIB).

With the existing data sources and archives it is particularly difficult to retrieve data. This is because, most of them are in gridded formats, and it takes special tool for deploying, this kind of tools are very familiar for the meteorologists, but not for the academic community, where most of the current research work is done. For development purposes, we realize that every algorithm needs a data set in an easy and understandable format. With this data set, an application will be easy to build.

3 Problem Solution

Every algorithm needs a data set in an easy and understandable format. This leads us to develop a data set, in a common format about atmospheric variables, for the further design of a track forecast model. The data set must be a multidimensional model for an easily data access, so a data warehouse is what we needed.

This work provides a data warehouse with atmospheric conditions. The process of building the data warehouse, took us first to contacting data sources, retrieving hurricane motion relative parameters, accessing data and information pertaining to strong data reliability and after words, building up the dimensional model of a data warehouse. Also, given the amount of data to store and the time dimension matter, a data warehouse was the solution for the storage

The Reanalysis project provides data, but as previously addressed this data was in GRIB format and contains every climatology variable, and so it isn't just about tropical cyclones. At the beginning we learn about the atmospheric conditions that empower a hurricane movement, to determine which variables to obtain from the GRIB files. After that, the GRIB format of the files had to be understood, and using a decoder, we extract the data, removing the unwanted data and leave the relevant for the purpose. Once the data was in an understandable format, we define the dimension model of the data warehouse, and proceed to feed it.

At the end, there is a powerful reason why building a data warehouse makes sense: when it comes time to build a new application, with a data warehouse in place, the application can be built quickly; with no data warehouse, the infrastructure has to be built again.

Actually, we are working in the designing of a forecast model relative to the wind circulation and affection on a tropical cyclone. The idea is to create a neural network, with the data warehouse, which learn from the data and establish a pattern to predict the track of a hurricane [6].

At this point, we realize that hurricane models have increased significantly in accuracy; however 72 hour forecasts still have a forecast error of about 250 nautical miles and improvement is definitely needed.

The first priority of a hurricane forecast is to predict the track to high accuracy. Once track forecasts improve, the next goal is to obtain good estimates of the intensity. Taking the hurricane track in mind, we design a new approach technique in the track forecast. The main idea is to improve the existing track error of the current models.

The model at design takes that the movement of hurricanes is steered by the global winds. If we put a feather in the air, it will move where the wind lifts it, supposing the hurricane is like a feather in the wind, we are trying to understand the winds direction and atmospheric pressure that empower them.

When a hurricane is spotted, we can get the current position of it, them we obtain all the wind and pressure relative variables in a radius of 2,500 kilometers (Fig.1A). Within this imaginary circle, surrounding low and high pressure systems, that creates the wind circulation, can be identified y take in concern those who affect more to the phenomenon (Fig.1B).

The wind circulates in opposite directions in a high and in a low pressure system (Fig.1C). This is because the wind in a high pressure system, heads to a low pressure. This movement, in addition with the Coriolis force (deflection angle produced by the rotation movement of the Earth), provides the global circulation in the atmosphere.

Once we have spotted the systems affecting our hurricane, we will try to predict the hurricane direction, by taking in concern the current direction, the distance with all the pressure systems, the intensity and directions of the winds and the degree within the hurricane direction and the position of the systems (Fig.1D). In other words, the wind is going to tell us where the hurricane is heading.

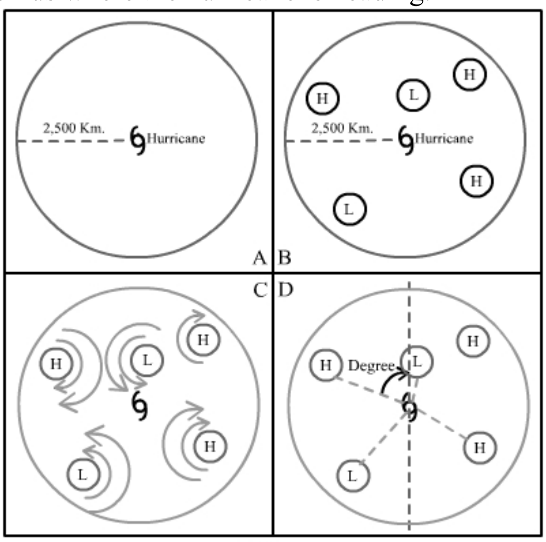

Figure 1

4 Conclusion

As previously addressed, the work on tropical cyclone track prediction included two major subjects: 1. the assembling of a data set with variables needed for the model; 2.the design of a track prediction model.

We took the first step, at this point we have a data warehouse, full with atmospheric conditions relative to a hurricane movement, the use of this data set, could give us the ability to analyze the present conditions and estimate how they will change.

The next step is to use this data and design a track prediction model, this is our current work, and we are designing the new approach and analyzing the existing tools like neural networks to apply them to the model.

After working with hurricanes and tracking models, we realize that track forecasting of tropical cyclones remains difficult at best and requires substantial amounts of information and data on the storm and the environment in which it is found. There are a lot of tools available in the information technology, relative to the analysis of data that has to be taken in concern. In this case, the combination of the meteorology and computer people has provide a fresh approach in the forecasting area. This should be taken as reference for further work.

References:

[1] Dr. Michel Rosengaus, "An analysis of the 2000 Tropical Cyclone Season for Mexico", Hydraulic Engineering in Mexico (in Spanish), vol. XVI, num. 4, pages 5-23, October-December, 2001.

[2] University of Illinois, "Hurricanes: online meteorology guide", Weather World 2010 project, 1999.

[3] Mark DeMaria, "Summary of the Tropical Prediction Center/National Hurricane Center Tropical Cyclone Track and Intensity Guidance Models", National Hurricane Center, 1997

[4] NOAA Magazine, "Three day hurricane forecast becoming more accurate with better environmental modeling", National Oceanic & Atmospheric Administration, January 14, 2002.

[5] NOAA Home page, “NCEP/NCAR CDAS, Reanalysis Project home page”, National Oceanic & Atmospheric Administration, 2000.

[6] Dr. Jose Torres Jimenez, "Short-term ozone forecasting by artificial neural networks". Advances in Engineering Software, vol. 23-3, pages 143 to 150. England, 1995.