

The Integration of Geographical Information Systems and Remotely Sensed Data to Track and Predict the Migration Path of the Africanized Honey Bee

Charles Ward, Jessica Bravo, Rosalia De Luna,
Gerardo Lopez, Itza Pichardo, Danny Trejo, and Gabriel Vargas

Pan American Center for Earth and Environmental Studies (PACES)
University of Texas at El Paso
El Paso, Texas 79968

Introduction

One of the research groups at the Pan American Center for Earth and Environmental Studies (PACES) is researching the northward migration path of Africanized Honey Bees or often referred to in the popular press as killer bees. The goal of the Killer Bee Research Group (KBRG) is to set up a database in the form of a geographical information system, which will be used to track and predict the bees future migration path. Included in this paper is background information on geographical information systems, the SPANS Explorer software package which was used to implement the database, and Advanced Very High Resolution Radiometer data and how each of these is being incorporated in the research. With an accurate means of predicting future migration paths, the negative effects of the Africanized honey bees maybe reduced.

History of Africanized Honey Bees

In 1956, the African honey bee was introduced into Brazil. The hope was to create a strain of honey bee which was better suited for the tropical environment of South America and thereby improving honey production. In 1957, twenty six bee colonies escaped and began interbreeding with the native bees. These new Africanized Honey Bees (AHB) have a greater tendency to swarm and abscond than common European honey bees. Since their initial release in Brazil, the AHB have been spreading at a rate of 80 to 500 kilometers annually. In the path of their northward migration, they have reduced honey production and negatively impacted crop pollination, livestock production, tourism and public health. The migration of AHB reached the United States in October of 1990 near Hidalgo, Texas. Currently, AHB have been sighted in south central Texas, southern New Mexico and Arizona and the central valley of California. A more thorough study of AHB is reported in [4].

Geographical Information System

A geographical information system (GIS) can be defined as “a computer-based information system that enables capture, modeling, manipulation, retrieval, analysis and presentation of geographically referenced data” [7]. One of the best features of using a GIS is that it provides the user a way to effectively visualize spatial information. A GIS gives a visual reference point, which helps the user to understand.

GIS use two basic types of data : geographic and attribute. Geographic data is composed of locational or spatial information about each specific feature or entity of data. Attribute data provides non-spatial information describing the characteristics or properties of each entity [1].

There have been many similar applications of GIS. An example is the Center for Health Applications of Aerospace Related Technologies (CHAART) at NASA Ames Research Center, in which the risk of Lyme disease transmission was studied [2].

The GIS software currently being used by the PACES KBRG is SPANS Explorer, which is manufactured by TYDAC Technologies Inc. SPANS Explorer comes in several software formats. The PACES KBRG is running a Windows 95 version of the program on Pentium 100 MHz computers.

Data Layers

The foundation for the data layers is a base map. Information can then be added to the base map by overlaying data layers, which can be displayed using different graphical formats. These formats are point, line, area, raster and quadtree. Point data represents locations of geographic phenomena, such as elevation or a sampling site. Line data consists of a set of connected points. Some examples of line data include rivers, roads and non-physical boundaries, such as county and state lines. Area data consists of a region enclosed within a boundary. Our research group is storing our information in this format. The data we have collected, so far, is in the format of number of AHB sightings in a county and hence is entered as area data. Raster data is a grid data structure which represents an area of the Earth's surface as a matrix of equal sized cells. Aerial photography, satellite and radar imagery are examples of raster data. Quadtree data is a more sophisticated form of raster data. Quadtree data represents an area of the Earth's surface as a matrix of variable sized cells. Areas with more detail are composed of smaller cells to preserve the resolution. For example, a quadtree of a coastal area uses large cells to represent the ocean and smaller cells to represent the greater detail of the coastline.

Methods of Entering Data

Data can be entered into SPANS Explorer manually or through the importation of other file formats. Attribute data can be added to a data layer in a manner similar to using a common spreadsheet. This is the method that the research group has used. Geographical data, such as boundary lines, can be input with the use of a mouse. SPANS Explorer also allows for the importation of a large variety of data formats. It comes with a raster and vector translator companion programs which can translate formats such as United States Geological Survey (USGS) 7.5 Minute Digital Elevation Model.

User Interfaces

SPANS Explorer has querying and charting functions that allow for the analysis of information. The software package has two methods for performing queries of the data layers, spatially and non-spatially. A spatial query involves the analysis of data in only a user defined area of the base map. The program allows a specific area to be chosen for query, by enclosing it in a circle, rectangle or polygon. A non-spatial query involves an analysis of the data layer as a whole. A query is performed by defining an equation. An equation is made up of one or more statements, which consists of an attribute and a constant value separated by a relational operator. Several statements can be combined with the use of a logical operator, such as AND or OR. When a query is executed the entities, which fit the conditions of the defined equation, are highlighted. The query function has been used by the research group to observe which counties in Texas have had a continuous increase in the AHB population. Data contained in a data layer can also be analyzed in the form of a chart. A chart can be created by highlighting data in an open data layer. The data can be seen in a line, bar, pie or radar chart format.

Attribute Data Used in AHB Migration GIS

The present attribute data that is being used in the AHB migration GIS can be grouped into the categories of weather, land usage and sightings. The data in each of these categories are being added to the GIS in the form of data layers.

The first category, weather, consists of data on the number of frost free days, the average monthly temperature and amount of precipitation in an area(county). All three have been shown to influence the rate and possible migration paths of AHB. This information is being acquired from the National Climate Data Center (NCDC).

Another type of data that influences AHB migration is land resource usage. This indicates if land is a wilderness area or if the land is currently being used for agriculture. A source of this data is still being searched for at the present. An alternative approach to filling this layer is to use the results of a classification model driven by remotely sensed data.

The final information that is being used in the set of data layers is the sightings of AHB swarms by year. AHB sightings give the important information of where the next migration move will start from, The best source of this information, to this point, has been "Detection Records of Africanized Honey Bees in Texas During 1990, 1991 and 1992" [3]. This paper gives the estimated number of trapped and free swarms of AHB in Texas, over those three years.

Remotely Sensed Data

Another form of data that the research group will be integrating into the AHB migration GIS is remotely sensed data from Advanced Very High Resolution Radiometers (AVHRR). AVHRR is a broad-band, four or five channel (depending on the model) scanner that senses in the visible, near infrared, and thermal infrared portions of the electromagnetic spectrum. The spectral ranges are as follows :

Wavelength	Channel #	in micrometers
	1	0.58-0.68
	2	0.725-1.10
	3	3.55-3.93
	4	10.3 -11.3
	5	11.5 -12.5

The AVHRR scanner is carried on the National Oceanic and Atmospheric Administration (NOAA) Polar Orbiting Environmental Satellites (POES). The POES orbit the Earth 14 times each day at an altitude of 833km(517mi). The data has a ground resolution of approximately 1.1 km. Four out of every five samples along the scan line are used to compute one average value and the data from only every third scan line are processed. This yields a 1.1km by 4-km resolution at nadir. One problem experienced by the AVHRR scanner is clouds. Several AVHRR overpasses are required to ensure cloud-free images. The duration of consecutive daily observations is called the compositing period. An image which provides a clear observation of a large ground surface area at reasonable nadir viewing angles is included in the composite.

AVHRR data is used to calculate the Normalized Difference Vegetation Index (NDVI) which gives the greenness or biomass of an area. The NDVI is calculated by taking the difference between the near-infrared (AVHRR Channel 2) and visible (AVHRR channel 1) reflectance values and dividing by the total reflectance. This calculation results in a value between -1 and 1, where a negative value tends to represent clouds, snow, water and other non-vegetated surfaces; positive values represent vegetated surfaces. The value is then scaled to fall between 0 and 200, and stored as an 8 bit number. When compositing the image, the highest value for a pixel during that period is used.

NDVI data has a problem distinguishing between clouds and water because the value over water is much lower than that of cloud. As a result of this cloudy images are taken over bodies of water. To solve this problem values less than 100 are flipped so clear observations will be chosen. For example a value of 99 becomes a value of 1.

NDVI data has a problem distinguishing between clouds and water because the value over water is much lower than that of cloud. As a result of this cloudy images are taken over bodies of water. To solve this problem values less than 100 are flipped so clear observations will be chosen. For example a value of 99 becomes a value of 1.

AVHRR data has been collected is performed by the EROS Data Center by the USGS, since May 1987. The KBRG has acquired the AVHRR data for the Conterminous United States from the USGS for the year 1989 to 1995.

From various statistical analyses, it is believed that a correlation maybe found between the NDVI data and the AHB migration paths. Preliminary research has found that there tends to be a difference between the standard deviation of the NDVI values for a given county that has Africanizedbees and those that do not. The NDVI data set contains statistical information, including the mean, median, maximum, minimum and standard deviation for every county in the conterminous United States. A classification scheme is being developed to determine whether a county is likely to be infested with bees. The correlation coefficient is being calculated between a group of test counties. The county is then classified based on whether it tends to be more closely related to counties which are known to have AHB or those that do not. If ibis can be done, a profile based on NDVI data can be created to predict future AHB migration. Results of the work to date are provided in [5].

Summary and Future Work

The KBRG has use a GIS to help track, analyze and predict the migration path of the AHB. The group continues to collect information to improve the quality of the predictions. Remotely sensed data, specifically AVHRR data, is being integrated into the GIS and analyzed by other programs to determine any correlation can be found with the AHB migration path. Correlation studies based on other measures associated with NDVI will be investigated to expand the predictive models.

Acknowledgments

We wish to thank NASA for its support of this project through the grant NCCW-0089. Also, we wish to extend gratitude to Byron Wood and Louisa Beck of the Center for Health Applications of Aerospace and Related Technologies at NASA Ames Research Center for their assistance in the formulation of this project.

References

[1] Burke, Diane, SPANS Explorer: The Future of Geographic Analysis Defined, TYDAC Research Inc. ,1996.

[2] Dister, S. W., L.R. Beck, B. L.Wood, R. Falco, and D. Fish, "The use of GIS and remote sensing technologies in a landscape approach to the study of Lyme disease transmission risk," Proceeding. GIS '93: Geographic Information Systems in Forestry. Environmental and Natural Resource Management, Vancouver, B. C., Canada, pp. 15-18 February 1993.

[4] Navarro, Hector, "An Investigation of the Migration of Africanized Honey Bees into the Southern United States", Proc. of the 1997 NASA URC-TC '97 National Technical Conference on Education, Aeronautics, Space, Autonomy, Earth, and Environment, Albuquerque, NM, Feb. 1997.

[5] Ward, Charles, "A System To Characterize Landscape Phenology Using AVHRR Data", MSEE Thesis, The University of Texas at El Paso, December 1996.

[6] Wood, Byron, Louisa Beck, Sheri Dister, Brad Lobitz, and Vince Ambrosia, Readings in Remote Sensing, Geographic Information Systems, and Related Technologies, Center for Health Applications of Aerospace Related Technologies, NASA Ames Research Center, 1995.

[7] Worboys, Michael F., GIS: A Computing Perspective, New York : Taylor & Francis Inc., 1995.