Report on Activities of the Integrated National Adaptation Pilot Project (INAP)

Executive Summary

Health and Climate Workshop in Bogota Colombia, September 2008.

First Report

Stephen J. Connor

May 28th, 2010

The International Research Institute for Climate and Society

Throughout more than four years The International Research Institute for Climate and Society (IRI) of Columbia University has advised the National Institute of Health (INS) and the Institute of Hydrology, Meteorology and Environmental Studies (IDEAM) in the design and development of an Early Warning System (EWS) that allows, through the development of predictive models using climate variables and other variables of interest, detection and timely intervention for Malaria.

Activity 1: Capacity Building on the Impact of Climate Change in Public Health

Capacity building activities of Colombian personnel involved in the project was achieved through interaction with IRI staff during short visits of Colombian scientists to IRI's headquarters, workshops organized in Colombia with the participation of IRI staff, and through periodic communication and interaction between Colombian and IRI personnel. Below is a summary of the specific activities carried out during the visits of Colombian staff to IRI headquarters and during IRI staff visits to Colombia.

- *Visit of Franklyn Ruiz, meteorologist at IDEAM, to IRI, September 30th to October 30th, 2007.*

The IRI hosted IDEAM's Franklyn Ruiz during the month of October, 2007. Activities conducted during his stay included: a) work with Tony Barnston, Lead Forecaster, in the Climate Predictability Tool (CPT) using the National Center for Atmospheric Research's (NCAR) data, characterization of climate variability, and learning to use IRI's Data Library to obtain data from General Circulation Models (GCM) for use as inputs for CPT; b) work with Michael Bell, Senior Staff Associate for Climate Monitoring and Dissemination, to be trained on options to produce precipitation forecasts for the next three or six months based on models assembled in the Data Library; c) work with Jian-Hua Joshua Qiah and Liquiang Sun, Research Scientists, to learn about dynamical downscaling using the regional model RegCM3 as well as using the Regional Spectral Model (RSM); d) work with Andrew Robertson, Research Scientist for Predictability and Downscaling, to learn about the Stochastic Daily Rainfall Sequences Toolkit and the Hidden Markov Model Tool for daily precipitation and temperature data analysis. The tools were used to analyze the spatial coherence between global models such as ECHAM4P5 with rainfall data recorded by weather stations, as well as to link characteristics of the rainy season and of dry periods with the behavior of meteorological variables provided by GCMs; and e) Work with Tufa Dinku, Associate Research Scientist for Climate and Environmental Monitoring, to be trained in the use of Remote Sensing for precipitation forecasts, in particular the importance and challenges of using infrared images to get better image-based results.

- *Visit of Viviana Ceron, Consultant at INS for the INAP project, to IRI's "Summer Institute", June 2nd to 14th, 2008.*

The IRI hosted INS' Viviana Ceron for two weeks in June, 2008 as a participant of IRI's Summer Institute. The main objectives of the Summer Institute (SI) were to (i) understand the relationship between climate and human health and learn about new tools developed for studies and analysis of this relationship, taking into account the limitations of different data sources including those from remote sensing, and (ii) engage professionals who play a key role in the operational decision-making for climate-sensitive diseases in identifying and evaluating appropriate use of climate information. In this regard, the SI was especially beneficial for our Colombian partners since part "D" of the INAP project relates to the impact of climate in health.

The course curriculum was designed to help participants to enhance their knowledge of climate-sensitive diseases, such as malaria, influenza and meningococcal meningitis, and foster the use of climate information in the management of climate-sensitive disease programs. Over the course of two weeks, participants were exposed to methodologies and tools developed by IRI and partner institutions, such as the IRI Data Library.

The SI included practical exercises and work on the Climate and Health interface which allowed Viviana Ceron to use the case study of the INAP project. In her participation in the SI, Ms Ceron also interacted with several staff members of the IRI involved in the INAP project.

- *Training workshop for INS and IDEAM's staff on the "Basic and Advanced Course in Epidemiological and Public Health Surveillance with emphasis in Climate and Health" Bogota, Colombia, September 19th to October 4th, 2008.*

Gilma Mantilla, Senior Staff Associate for Public Health, traveled to Bogota Colombia to work with some key staff members of INS and IDEAM in mid September 2008. The purpose of her visit included running the Basic and Advance Course in Epidemiological and Public Health Surveillance with emphasis in climate and health, and lecturing INS and IDEAM's staff on climate change impacts in public health and the use of climate information in public health. Gilma also followed up with activities of the INAP project, as well as participate in the First National Meeting of Climate and Health.

Gilma's visit to INS and IDEAM was a crucial activity for the INAP project in terms of training INS and IDEAM's staff, as well as developing a training module of climate and health in partnership with the IRI. Such a module will be used as training material for the Field Epidemiology Training Program in Colombia and also in other countries in South and Central America.

- *Workshop on "Health and Climate" for INS and IDEAM's staff. Bogota, Colombia, September 22nd to 24th, 2008.*

Walter Baethgen, Director of the Latin American and Caribbean Program, and Madeleine Thomson, Senior Research Scientist for Climate Information for Public Health, visited INS and IDEAM in late September to lecture in the workshop entitled "Climate and Health", sponsored by INS. Walter gave a presentation on "Climate Risk Management" in the opening session of the Workshop, and Madeleine gave a presentation on "Climate and Public Health" in a plenary session of the second day. Walter also met with Ricardo Lozano, Director of IDEAM, to discuss the need to organize and homogenize the weather databases that are available in IDEAM and the ones that the IRI has been using in the INAP project.

Walter and Madeleine met with Andres Escobar, Under Secretary General of the Department of National Planning (DNP) to discuss the development of a National Policy on Climate Change. Such a National Policy plan will establish new climate change policies by sectors, and as a result of that initial interaction, DNP and IRI are now working to establish a work agreement for exploring specific ways to incorporate climate risk management into the new policies. Meetings were also held with Giampiero Renzonni, Coordinator Environment and Sustainable Development at DNP, to coordinate his participation at the World Climate Conference-3 (WCC-3, Geneva) as one of the speakers in the health panel.

- *Visit of Luis Alfonso Lopez, Forecaster at IDEAM, to IRI, February 23rd to March 7th, 2009.*

The IRI hosted IDEAM's Luis Alfonso Lopez in late Ferbary, 2009. Luis worked with Simon Mason, Research Scientist for Forecasting and Prediction Research, Andrew Robertson, Research Scientist for Predictability and Downscaling, and Tony Barnston, Lead Forecaster, in statistical climate modeling in order to define the statistical and dynamical models for the Integrated System of Control and Surveillance for Malaria and Dengue (SIVCMD) in Colombia.

Luis worked with IRI staff to clean and organize IDEAM's climate data using the SAS statistical package in order to upload it to the Data Library. Luis also met with Michael Bell, Sr. Staff Associate for Climate Monitoring and Dissemination, to learn about the Data Library and its benefits, such as creating maproom components and developing climate forecasts. The second week was spent analyzing the data with statistical tools and programming languages, like Fortran and MatLab. Luis worked with Andy on MatLab to analyze the correlation between different climatological data and malaria data. Tony and Luis worked on running some subroutines in Fortran language in order to explore the impact of temperature and precipitation on Malaria Falciparum, as well as validate the data with statistical references.

- *Training of INS and IDEAM's staff on the use of the IRI Data Library. Bogota, Colombia, April 20th to 24th, 2009.*

Remi Cousin, Staff Associate for the Data Library, visited INS to train its staff, as well as IDEAM's staff in the use of the Data Library. Remi gave a presentation to introduce the Data Library's Maprooms and their potential uses, and to train INS staff in the use of the Colombia/ Malaria dataset of the Data Library to establish a maproom on malaria in Colombia.

The 3-day training workshop included the following activities: introduction to the IRI, its mission, its activities and the DL; learning to use the DL through hands-on exercises with epidemiological and climate data available in the IRI; hands-on DL exercises with epidemiological and climate data from Colombia, addressing specific applications such as averaging over health districts, computing lagged correlation between climate and epidemiological data, and performing sophisticated analysis such as SVD or CCA.

- *Presentation at the "XIV Conference of The Colombian Parasitology and Tropical Medicine". Medellín, Colombia, October 8th to 11th, 2009.*

Tufa Dinku, Associate Research Scientist for Climate and Environmental Monitoring, was invited to give a talk at the XIV Conference of The Colombian Parasitology and Tropical Medicine. Tufa gave a presentation on "Integration of Remotely Sensed and Conventional Climatic data in IRI's Malaria Early Warning System." This presentation described the IRI Data Library, the Health Map Room, the Victoria Capacity Mapping, and the value of remotely sensed climate information. With regards to remote sensing, Tufa presented the validation of different satellite rainfall estimates over different part parts of Colombia (a scientific paper was produced, and it is attached to this report).

Activity 2: Data Quality Control and Development of Data Sets of Malaria Incidence

A major goal of the INAP project is to characterize links between climate and malaria incidence in Colombia, toward developing an early warning system for this disease. Toward this goal, data sets of malaria incidence and two climate variables (temperature and rainfall) have been developed for four Colombia municipalities that possess adequate data for the 7-year study period (2001-2007): Puerto Libertador, Montelibano, San Jose Del Guaviare, Guapi, and Buenaventura. Guapi was omitted from the analyses because of inadequate malaria data. Additionally, Puerto Libertador lacks temperature data, so it was only used for rainfall-malaria analyses.

Data set development was challenging due to periods of missing or incorrect data. Consequently, in order to facilitate further analyses, missing entries in the daily datasets of precipitation and temperature (daily minimum, mean and maximum) were filled using a stochastic daily weather generator. Regarding the two types of malaria, it was decided to use Falciparum rather than

Plasmodium Vivax for the analyses because it is possible that Vivax cases might actually include some Falciparum cases too, but the reverse is not true.

The initial steps included defining the seasonal cycle of both rainfall and Falciparum at each of the municipalities. Results show a broad, long-lasting rainfall seasonality with relatively heavier rainfall between May and September at most of the municipalities, and lower amounts from October to March. (Buenaventura is an exception, with a short period of heavy rainfall in April and May, but heaviest rainfall from August to November.) The seasonality of mean temperature is less well-defined for most municipalities. However, in one or two municipalities with a particularly well-defined dry season at the beginning of the calendar year, there is a maximum in temperature at that same time of year (e.g. San Jose Del Guaviare). The seasonality of Falciparum roughly parallels that of rainfall, with relatively highest incidence centered in the middle of the calendar year and lower incidence near the beginning and the end of the year for several municipalities. However, the correspondence is quite rough, due to the broadness of the climate seasonality, and it is difficult to note a lag time between the peak rainfall and the peak Falciparum. In some cases temperature may play some additional role.

Secondly, we studied the association between rainfall and temperature with Falciparum on a year-to-year basis, for a given season. In other words, we explored if high-rainfall years showed, for a given time of year (e.g., June), more Falciparum, and whether the relationship is maximized using a lag time of one or two months (i.e., strongest link between rainfall and Falciparum values 1 or 2 months later). Results indicated fairly strong associations in Montelibano between temperature and Falciparum in April, both simultaneously (zero lag) and with Falciparum lagging temperature by up to 3 months. Possible associations between rainfall and Falciparum are also suggested for certain municipalities at certain times of year. A major problem with this analysis is that due to having only 7 years, there is a large sampling error problem, and correlations need to be close to 0.8 in order to be considered statistically significant. Overall, the results suggest that associations between temperature (and/or rainfall) and Falciparum occur with a shorter lag time than found in other regions, and thus may appear during the same month as the climate anomaly and persist for 1-3 months following that initial month. However, the 7-year analysis period is too short for robust conclusions to be made and further assessments are recommended.

Additional analyses that could shed more light on climate-Falciparum relationships, and that could help lead to a Falciparum early warning system, might be:

- Multiple regression of Falciparum, using both rainfall and temperature as predictors (this would require even longer time series)
- Broaden the study to include relationships between ENSO and the climate variables (rainfall and temperature) at each municipality, based on a much longer time period than 7 years. If one assumes the typical relationships between climate and Falciparum (greater

temperature increases it, and greater precipitation also increases it), then climate forecasts could be made on the basis of the expected ENSO state for one season ahead, making possible a preliminary version of an early warning system for Falciparum at municipalities where the disease is affected by the climate.

- Gather more socioeconomic data so that an early warning system could be better suited to the on-the-ground vulnerability. (So far we lack such data.)

For more information on these datasets please visit the following link:
http://wiki.iri.columbia.edu/index.php?n=HealthAndClimate.Colombia.

Activity 3: Analysis for a Statistical Model

A document that discusses a procedure to build time series statistical models for malaria outcomes is included in Annex 1 of this report.

Activity 4: Analysis for a Dynamical Model

Activities conducted over the period January 2010-April 2010 included the study of two additional process-based models, the simulation of their baseline scenarios, the design of an online tool of the mathematical model proposed by Worrall, Connor, and Thomson (2007), and the documentation of up-to-date activities. A brief description of these advances is presented below.

Modeling: Two additional process-based models, namely the Chiyaka, Garira and Dube endemic human malaria transmission model (2007) and the Gomero's malaria-sickle-cell model (2008), were considered to analyze malaria transmission dynamics in the Colombian pilot sites. Their stock-flow models (Forrester models) were created on the computer software Powersim Constructor, Version 2.51, as depicted in Figure 1. A detailed description of these mathematical tools is presented in Ruiz et al. (2010).

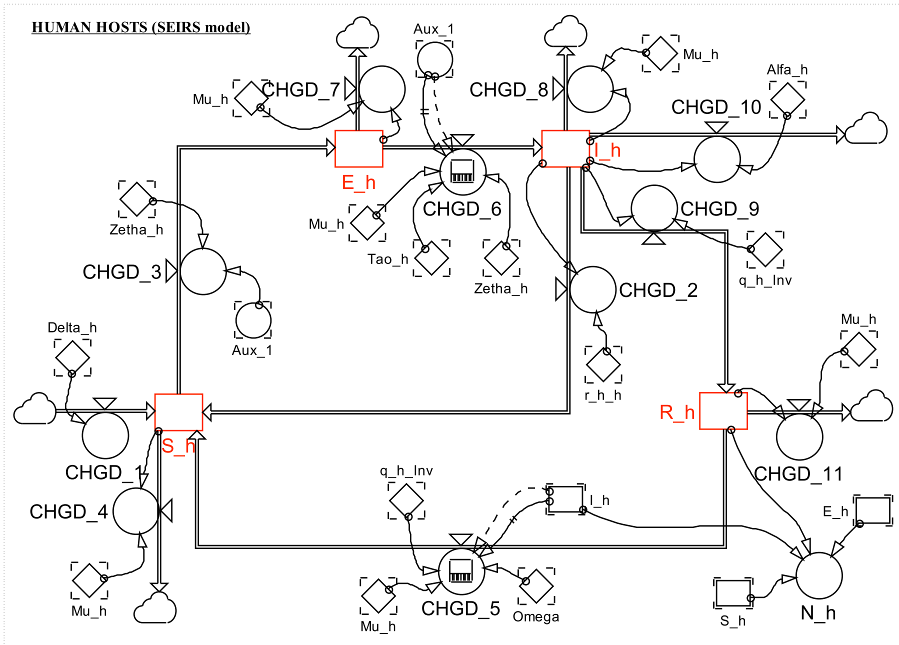

Figure 1. Stock-flow model of the human host component of the malaria transmission model proposed by Chiyaka, Garira and Dube (2007)

Simulation: Simulation activities included the base scenarios discussed by Chiyaka, Garira and Dube (2007) and Gomero (2008); see Figure 2. The group also supported the analysis of the Colombian INAP pilot sites, conducted by the School of Engineering in Antioquia (Colombia), the National University of Colombia and the University CES (Colombia) for the Colombian National Institute of Health. Activities included the characterization of local eco-epidemiological settings and the simulation of local malaria transmission dynamics. The characterization included: general aspects (location, population at risk, natural resources, economic activities), climate (climatology and historical trends), entomology (primary and secondary vectors, feeding frequencies and preferences, human biting rates, breeding sites), total malaria positive cases (annual cycles of malaria incidence, stability conditions), non-climatic factors (demographic conditions, control campaigns and surveillance activities), and level of understanding. Numerical simulations included the implementation of the mathematical tools proposed by Macdonald (1957), Martens (1997), Ruiz et al. (2002, 2003, 2006), and Worrall, Connor and Thomson (2007) in the proposed 5 Colombian pilot sites. These experiments included retrospective analyses of at least 2,700-day simulation periods (base scenarios, changes in initial conditions, simulation of local climatic, entomological and socioeconomic conditions, sensitivity analyses, and uncertainties) and future scenarios (short-, medium- and long-term changing climatic and socioeconomic conditions)

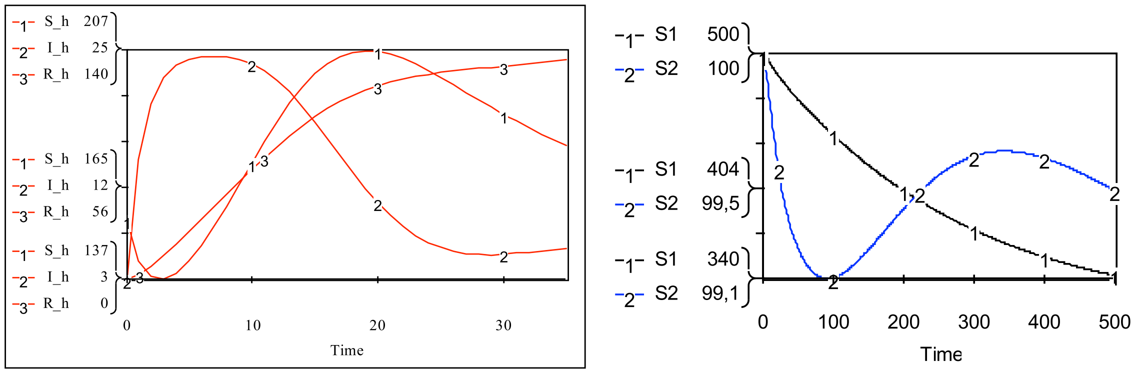

Figure 2. Time series of the Chiyaka, Garira and Dube endemic human malaria transmission model (left) and the Gomero malaria-sickle-cell model (right) for the proposed base scenarios or 'baseline parameter values'

Interactive platforms: The group has also worked on the design of an online version of the temperature- and rainfall-driven dynamic model of malaria transmission proposed by Worral, Connor and Thomson (2007). The first version of the online WCT model (see Figure 3) can be found at:
http://iridl.ldeo.columbia.edu/
home/.remic/.maproom/.Health/.Regional/.S_America/.Malaria/.WCT_Model/.

The mathematical model is now running for three INAP pilot sites (Buenaventura, Puerto Libertador and San Jose del Guaviare), using the temperature, rainfall and relative humidity records gathered at nearby weather stations. Records were provided by the Colombian Institute of Hydrology, Meteorology and Environmental Studies – IDEAM and are available at a daily time scale; see the following link:
http://iridl.ldeo.columbia.edu/home/.remic/.CIPH/.Colombia/.pilot_sites/.climate/.station/.daily/.

The group is currently working on the organization of the individual interactive platforms (focusing on the MAR, SimulMal and WCT models) and the operative multi-model ensemble platforms.

Outreach activities: A two-page description of dynamical models of malaria transmission and a fourteen-page description of the Colombian Case Study were prepared for the forthcoming IRI Malaria Manual. The Colombian case study includes, in particular, a brief description of the country profile, its climate and malaria situation, the available local evidence of malaria-climate links, the current policies, concerns and priorities, and a local study/experiment. The local study focuses on one of the Colombian INAP pilot sites, the municipality of San José del Guaviare, and includes a description of the malaria-climate situation and the available malaria, climate, entomological and socioeconomic variables and datasets, as well as the ongoing experimentation-validation-analysis process (exploratory analyses of malaria incidence and climatic drivers, and malaria modeling simulation outputs).

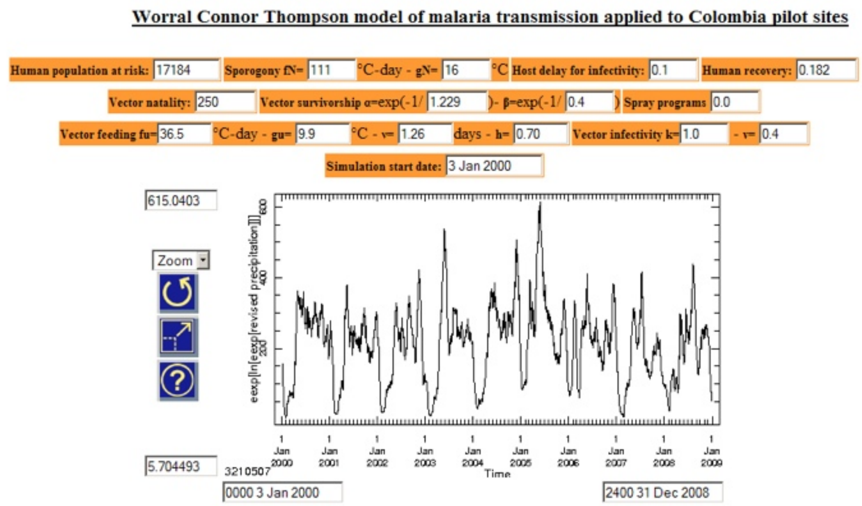

Figure 3. First version of the online tool of the mathematical model (WCT model) proposed by Worral, Connor and Thomson (2007)

For information on the full report please visit the following link:
http://iri.columbia.edu/lac/docs/simulating_malaria_transmission.pdf

Activity 5: Validation and Intercomparison of Satellite Rainfall Estimates over Colombia

Over the past decade, a number of precipitation products with high spatial and temporal resolution and near-global coverage have been developed. These products combine precipitation information from multiple sensors and multiple algorithms to produce estimates of rainfall over the globe at spatial resolutions of 0.25° latitude/longitude (or finer) and three-hour temporal resolution (or less). Because these products are constructed from satellite data, they supply crucial rainfall information over parts of the world where conventional surface-based observations of rainfall (rain gauges, radars) are very sparse or non-existent.

Evaluation of the different satellite rainfall estimates over different climatic and geographic regions is very important. This will be useful in identifying specific weaknesses and strengths of the different products under different circumstances. However, the evaluation of these products, particularly over Africa and South America, has been very limited. Yet it may be argued that these regions are where the satellite products are needed most because of the sparse station networks over most parts of the two regions.

A station network of about 600 gauges was used to evaluate seven satellite rainfall products over Colombia. The rain gauge data were obtained from IDEAM and the evaluated satellite rainfall products were Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation Analysis products, NOAA-CPC morphing technique (CMORPH) Precipitation Estimation from Remotely Sensed Information using Artificial Neural Network (PERSIANN), the Naval Research Laboratory's blended product (NRLB) as well as the Global Satellite Mapping of Precipitation (GSMaP). These products were evaluated over different parts of Colombia at daily and ten-daily time scales and spatial resolution of 0.25o latitude/longitude. This work has been accepted for publication in the Journal of Applied Meteorology and Climatology.

Several statistics were used to assess the performance of the different satellite tools: linear correlation coefficient (CC), multiplicative bias (Bias), mean absolute error (MAE), frequency bias (FBS), probability of detection (POD), false alarm ratio (FAR),) and the Heidke Skill Score (HSS) for validation at daily time scale.

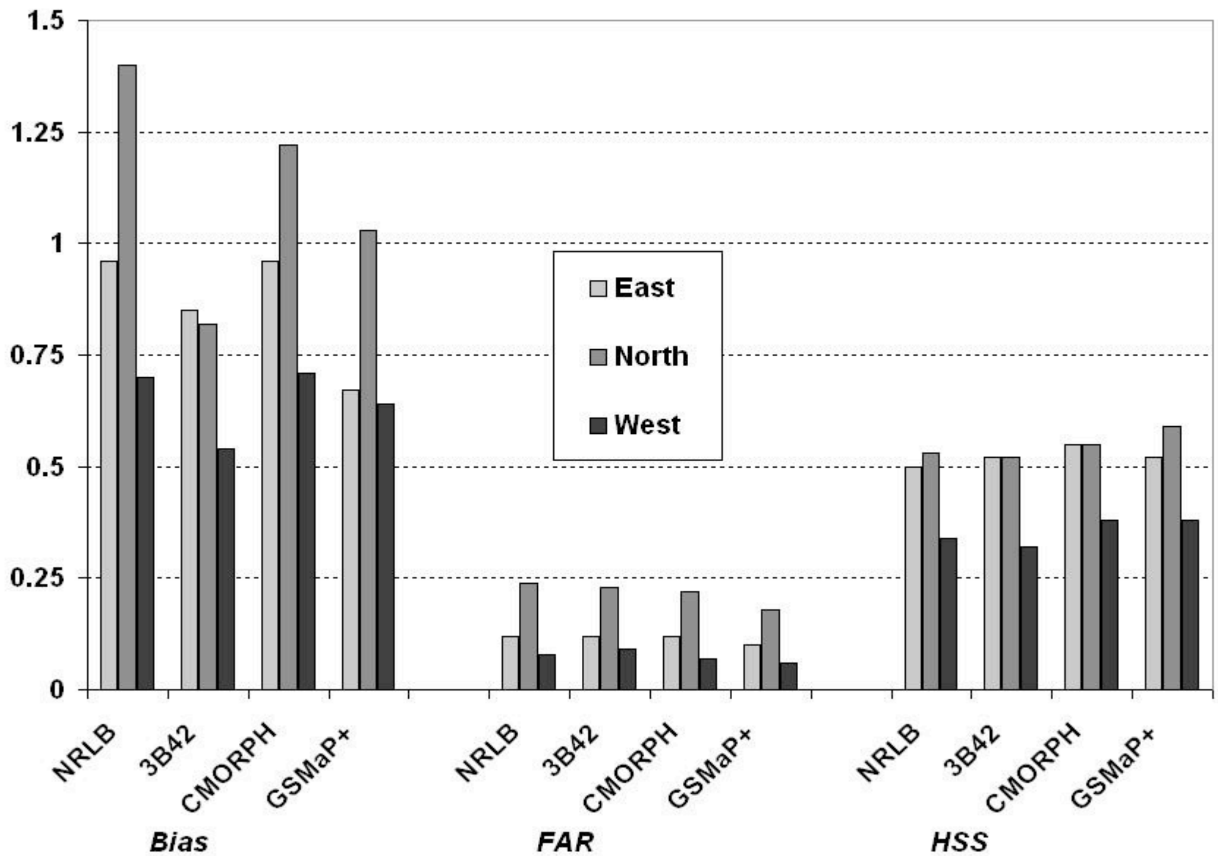

Figure 4. Comparing the performance of the different daily satellite rainfall estimates over eastern, northern, and western parts of Colombia.

The above figure presents some of the validation statistics (Bias, FAR, and HSS) for three lowland regions of Colombia. The FAR shows how often the satellite products detect rainfall when rain gauge measurements are zero. HSS measures the rainfall detection accuracy of the satellite estimates relative to matches due to random chance. In terms of bias, the performance of the satellite products is relatively better over the eastern region, while there is significant overestimation over the northern region for some of the products, and serious underestimation over the western region (Pacific coast). All products underestimate rainfall amounts over the western region, while NRL and CMORPH overestimate rainfall amounts over the northern region. Rainfall amounts are also underestimated over the eastern region, but the underestimation is not as severe as that observed over the Pacific coast. The FAR are higher over the northern (and drier) region. The HSS values are significantly higher for the northern and eastern regions than for the western region. The underestimation of rainfall over the western region may be attributed to warm rainfall process associated with orographic uplift caused by onshore flow toward the Andes, while the overestimation over the northern region could be a result of sub-cloud evaporation.

In summary, the performance of the satellite rainfall products over Colombia was similar to what has been obtained for other similar regions. Overall, the satellite products are reasonably good in detecting the occurrence of rainfall, but are poor in estimating the amount of daily rainfall. The products have good skill in estimating rainfall amounts on a 10-day time scale. Results are different for the different parts of the country. There are significant overestimations of daily rainfall frequency and amount over the relatively dry northern region, while significant underestimations are observed over the mountainous regions and the Pacific coast. The overestimation over the northern region has been ascribed to possible sub-cloud evaporation, while the underestimations over the mountainous and coastal regions may be associated with warm rain process over the two regions. Considering collectively the three performance diagnostics, the best performance is observed for the eastern region. Comparisons of the different satellite products suggest that CMORPH and GSMaP-MVK+ are the best products, while PERSIANN and GSMap-MVK exhibited the poorest performance among the products evaluated here.

For information on the full report please visit the following link:
http://journals.ametsoc.org/doi/abs/10.1175/2009JAMC2260.1

Annex:

Statistical modeling of Malaria[1]

Hugo Oliveros: oliveros@iri.columbia.edu

This document discusses a procedure to build time series statistical models for malaria outcomes. The suggested approach is linked to dynamic linear models (DLM) and its generalization dynamic generalized linear models (DGLM). These models can be used to represent the evolution of malaria indicators –prevalence, incidence, number of cases – based on observed and unobserved factors under a dynamic scheme where the sampling distribution of both malaria indicators and the dynamic structure can be Gaussian or non-Gaussian. This mechanism can be used to match trends and expected impacts of explanatory factors such as weather (climate) variables supported by malaria transmission mechanisms (MTM) and dynamic models.

The document discusses the importance of using information coming from MTM and dynamic models of malaria to guide the process of assembling statistical models. Although dynamic models try to reproduce MTM conditions as well as host-mosquito interactions by means of a non-linear mesh, the use of information coming from ongoing efforts to gauge dynamic models is encouraged due to positive spillovers to statistical models.

On the other hand, several issues usually come up when building statistical models in real life. The roots of those issues are mostly associated with unexpected data set features and problems. These difficulties have an important effect in the process of building and estimating the models. Some of the usual problems have been discussed throughout the document matching methodological benefits with data drawbacks and/or unexpected features. Some of the problems considered are the following: the presence of missing values, outliers and missing variables; the need to use transitory and permanent effects to portray malaria trend and explanatory variable effects; the use of seasonality, trend and cycle components; and the need to use malaria dynamic model stylized facts and MTM to develop or test the statistical model.

The document also provides initial background elements for those interested in the estimation process. The estimation process is discussed in some detail and potential software solutions are introduced and suggested. The software solutions use Gaussian SSM properties for DLM and Gaussian approximations for OE and/or SE under non-Gaussian sampling distribution, DGLM. In both cases, Kalman filter (KF) recursions and non-linear optimization procedures are used to get the maximum likelihood estimators of the parameters of interest. KFS or IKFS processes are used to obtain estimates of the state vector depending on the sampling distribution of the dependent variable and the state equation.

[1] Full document available at: http://iri.columbia.edu/lac/docs/statistical_modeling_of_malaria.pdf

SAS, MATLAB and R options have been tested partially under simulated Gaussian shocks for the observation and the state equations of a SSM; i.e., DLM, with comparable results. DGLM software features have been partially tested for observations coming from sampling distributions that belong to the exponential family.