Duncan Golicher, John; Ramirez-Marcial, Neptali; Levy Tacher, Samuel Israel
Correlations between precipitation patterns in southern Mexico and the el niño sea surface
temperature index
Interciencia, vol. 31, núm. 2, febrero, 2006, pp. 80-86
Asociación Interciencia
Caracas, Venezuela

Available in: http://www.redalyc.org/articulo.oa?id=33911302
CORRELATIONS BETWEEN PRECIPITATION PATTERNS IN SOUTHERN MEXICO AND THE EL NIÑO SEA SURFACE TEMPERATURE INDEX

JOHN DUNCAN GOLICHER, NEPTALÍ RAMÍREZ-MARCIAL and SAMUEL ISRAEL LEVY TACHER

SUMMARY

Historical data for regional climate research in the tropics is typified by extreme unevenness in both temporal and spatial coverage. Spatial interpolation techniques were applied in order to fill the gaps in time series of rainfall records for the state of Chiapas. The method involved iterated universal Kriging that combines a spatial covariance function with a polynomial trend surface. Automated outlier removal was used to prevent spurious records distorting the results. The procedure was applied 612 times in order to produce complete monthly time series from 1950 to 2000. In order to trace temporal trends the time series were decomposed into seasonal, trend and irregular components and analyzed using loess smoothing (STL). The seasonal values were removed, and the remainder smoothed to find the trend. An identical procedure was applied to the El Niño3.4 index. The trend component of each data set was analyzed for autocorrelation and cross correlation. The autocorrelation function for the standardized number of days with rainfall shows significant positive correlations between data points around three to four months apart. There is significant negative cross correlation between the standardized El Niño sea surface temperature index and rainfall. The technique thus led to a clear description of a pattern that might be used in order to partly predict precipitation driven events such as floods and wildfire.

KEYWORDS / Chiapas / Correlation / Interpolation / El Niño / Rainfall / Time Series /

can be cleaned by selective usage of only the most complete records. Ad-hoc interpolation or corrections can be made. However, there are a variety of additional problems that they do not confront. Missing or erratic data is not simply imprecise and noisy. Biased and misleading signals can be introduced due to instrument or recorder lack of homogeneity combined with changes in regional coverage (Hulme and New, 1997). It is not always clear if the quantity and completeness of a climate station’s records are positively correlated with their accuracy. Operational details that would help clarify sources of systematic error can be impossible or costly to obtain.

To solve some of these problems methods have been devised that permit individual data points to either "borrow strength" or become censored by the rest of the available information. Iterated kriging has been suggested by Sherwood (2001). The approach is elegant but computational intensive. A related but simpler method was used herein, based on repeated universal block kriging. The method was loosely based on that used by Kittel et al. (1997). It was chosen in order to fulfil three key objectives. Firstly, to assess the extent to which rainfall patterns in the Southern Mexican state of Chiapas are linked to the El Niño Southern Oscillation, using as much of the available data as possible without prejudging its quality. Secondly, to produce detailed, yet easily communicated breakdown of the available information on temporal and spatial rainfall trends in order that it may be better incorporated into both regional decision making and agricultural and ecological research. Thirdly, to produce robust time series of monthly mean precipitations for use as input for stochastic simulation models that can be incorporated into agricultural and ecological models being developed for the region.

**Methods**

The input for the statistical model was compiled by the Instituto Mexicano de Tecnologia del Agua (IMTA) from data provided by the Comisión Nacional del Agua (CNA) through the network of climatological stations CLICOM and the Comisión Federal de Electricidad (CFE). This recompiled data was also checked against and augmented by more recent data obtained directly from the CNA. Contemporary developments are leading to improvements in the quality of meteorological data for the state. However the description of the methods by which most of the data were collected and compiled reveals many potential causes of loss and errors. Values were taken daily at 08:00am. Minimum and maximum temperatures and rainfall refer to the previous day. The data were handwritten in notebooks which were collected at sporadic intervals of weeks or months and either sent by mail or collected by authorized personnel. The notebook data was then copied into sheets with the format 212-50 in the case of the CNA, the format CFF-VII-176 or some other unspecified format in the case of stations associated with irrigation districts or agricultural associations. Four standard measurements (maximum and minimum temperature, precipitation, and pan evaporation) were available on a daily basis since 1950 for 15 to 212 climate stations in the state. Temperature and evaporation records appeared to be the most unreliable and this present study thus concentrated on rainfall records alone. Records were available at a daily time scale for 221 stations distributed across the state. However, most of the time series are erratic and incomplete, meaning that for any single day records are available for between 31 and 188 stations.

**Statistical analysis**

Monthly mean rainfall was calculated together with the number of days with measurable (>1mm) precipitation for each month between Jan 1950 and Dec 2000, inclusive, for all stations with complete monthly records. The localities of the stations and completeness of the records are summarized in Figure 1. Notice the concentration of stations in the Soconusco region, close to the Guatemalan border. Most stations have coverage for less than half the period of interest. The two spatially referenced time series formed the input for our geostatistical models. In order to use spatial interpolation to produce complete time series, a statistical analysis was required that allowed repeated spatial interpolation of the rainfall over the entire study area for each month from 1950 to 2000. It was also important that the method was as robust to the effect of outliers arising from erroneous data recording as possible.

The first step of modeling thus involved fitting 612 spatial surfaces for each variable, while recording the predicted value for a region of interest for each model in order to produce a complete time series. Fitting geostatistical models is a time consuming art that is not easily turned into a fully automated repeated procedure (Cressie, 1993). No interpolation method is universally superior to any other. The quality of the interpolation depends on details of each data set rather than being an inherent property of the method chosen (Helsel, 1990). Two commonly used techniques are thin plate spline (Hutchinson, 1995) and kriging (Cressie, 1993). Thin plate spline (TPS) has two advantages over kriging: i- It uses fewer parameters. ii- Fit is evaluated through generalized crossvalidation. TPS thus appeared best suited to the task of repeated fitting of many surfaces (Hutchinson, 1993). The results of TPS were compared with those of kriging by fitting surfaces with both methods using the package “fields” (National Centre for Atmospheric Research) implemented in the statistical language and environment R (Ihaka and Gentleman, 1996). Because meteorological stations in Chiapas are rather tightly clustered into local neighbourhoods, short range correlation and inaccurately recorded station
positions were found to cause difficulties in reliably fitting surfaces using TPS. The automated crossvalidation procedure resulted in oversmoothed surfaces that did not show sufficient local pattern in precipitation. These problems can be partly resolved by the removal of data points (Hutchinson, 1998). However, because the surfaces suggested by TPS were very similar to fitting a second order polynomial trend surface, it was decided to use universal Kriging models that combine the spatial covariance function with a polynomial trend surface. These models were fitted through generalized least squares using the package “spatial” in R (Venables and Ripley, 2002).

As Katz et al. (2002) points out, precipitation data are typically right skewed and heavy tailed. The degree of skewness is somewhat reduced when monthly means are taken, nevertheless the variance remains proportional to the mean. A square root transform of the mean monthly precipitation was therefore carried out prior to kriging for both data sets. Back transformation on the predictive surfaces then led to the spatial models.

Applying repeated universal kriging is complicated by the fact that changes in empirical autocorrelation are related to changes in spatial coverage. Some abrupt changes in spatial covariance were also attributable to erroneous data. After visual analysis of the 612 variograms for each model a single covariance function was chosen over the whole time frame for each model. This introduced a potential source of error that could be eliminated by fitting separate models to each variogram. However, the drift modeled by the polynomial trend surface is fitted without reference to the variogram. The trend surface draws on all the data points and is potentially a more robust global element of a universal kriging model than the local autocorrelation component. Thus, fitting the best empirical covariance function does not necessarily lead to improvement in the model if some of the variability is noise derived from errors and omissions. Sensitivity analysis showed that the surfaces produced by universal kriging were not greatly affected when monthly means are taken, nevertheless the variance remains proportional to the mean. A square root transform of the mean monthly precipitation was therefore carried out prior to kriging for both data sets. Back transformation on the predictive surfaces then led to the spatial models.

To summarize, the steps in the analysis were

1. Calculate mean daily rainfall for each month from 1950 to 2000 for all stations with over one year’s worth of records.
2. Calculate the number of days with appreciable (>1mm) rainfall for each month for each station with a complete monthly record.
3. Fit polynomial trend surfaces to each data set for each month and remove extreme outliers.
4. Fit predictive surfaces to the data using universal kriging.
5. Calculate mean values for 10km² polygons at key positions in the state.
6. Breakdown the data into three components seasonal using STL.
7. Analyze the trend component of each data set (number of days with rain, amount of rainfall and El Niño index) looking at autocorrelation and cross correlation between them.
Results

Figure 2 illustrates the results of the kriging technique that was applied for each month. It shows the typical pattern for four months of a single year with a comparatively full set of data (110 stations with at least partial records). The typical pattern is that the dry season (Dec to Apr), is less marked in the humid North and West of the state where 2000-3500mm of rain falls annually. The seasonal rains begin earlier and are more intense on the Pacific coast as compared with central regions. The driest region is the central valley (600-1100masl) which receives around 700-1100mm per year. The highlands (1600-3000masl) have a gradient of rainfall, being drier in the South and East. The kriging surfaces allow both the spatial and temporal component of these patterns to be investigated in some detail. Here only the temporal component is evaluated.

The best supported time series were obtained from the central region of the state. Thus, an STL decomposition of the standardized mean monthly rainfall for the area around the city of San Cristóbal de Las Casas in the highlands of Chiapas (2200masl) was used in order to analyze the periodicity in the regional rainfall. The pattern is shown in Figure 3. The periodicity in the trend after seasonality has been removed is clear. There also is a slight downward trend in rainfall since the early 50s although this shows up more clearly if a larger window is used for the STL analysis. Time series produced in the same manner for other regions of Chiapas show similar overall trends, although only incomplete time series were derived for the North, Pacific coast and Lacandon regions due to lack of support for points at the margins of the region.

Figure 4 shows only the trend component of the time series for mean monthly rainfall and number of days of rain per month for San Cristóbal in more detail. For most of the series the two measures are very closely correlated, suggesting that the average size of rainfall events has remained rather constant. Thus, when higher than average rainfall occurs, it is attributable to more days with rain, rather than simply to larger individual rainfall events. It is pointed out that under this particular analytical technique large single events tend to be placed in the white noise component of the seasonal breakdown. While it would be interesting, as a separate exercise to look at extreme events alone, the argument is based on a more robust general trend, rather than single direct results of individual tropical depressions. There is, however, an apparent link-
The earliest data shows considerable discrepancies from this pattern. Data from the 50s does however need to be interpreted with a great deal of caution, as these earlier records were inevitably less reliable and very sparse. The fact that a similar pattern, although less pronounced, is present in time series derived for the other regions may hint at either a genuine anomaly in the rainfall pattern in the early 50s, a period in which the state was affected by some major hurricanes, or simply be due to changes in the way data was collected at that time.

A caveat must be mentioned regarding the analysis technique adopted. Although spatial smoothing and outlier removal was chosen in order to reduce bias due to observer error, the method may also smooth over genuine anomalies. Thus, if extreme single events such as single storms or hurricanes are of interest this technique is not optimal. However, there are alternative reliable sources for data on hurricane frequency and distribution available (Goldenberg et al., 2001).

The data source used in the present work deserves further attention. Applying rigorous non-automated quality control was too costly and time consuming for this particular study. Further detailed investigation of the quality of the data on a station by station basis should be carried out in order to extract reliable additional information from this valuable data set.

Discussion

A caveat must be mentioned regarding the analysis technique adopted. Although spatial smoothing and outlier removal was chosen in order to reduce bias due to observer error, the method may also smooth over genuine anomalies. Thus, if extreme single events such as single storms or hurricanes are of interest this technique is not optimal. However, there are alternative reliable sources for data on hurricane frequency and distribution available (Goldenberg et al., 2001).

The data analysis can be interpreted as confirming the remote effects of El Niño Southern Oscillation (ENSO) in this region of Mexico. The correlation between the El Niño 3.4 time series and the standardized rainfall is significantly negative for around 6 months on either side (Figure 7). In other words, when sea surface temperature has been higher than usual, rainfall tends to be lower than usual during that year. Significant positive correlations occur when the lag between the two time series reaches 2 years. The periodicity observed in the time series thus show a clear sign that it may be coupled with ENSO fluctuations in sea surface temperature, although interestingly, the two time series do not always appear well synchronized. The lack of synchronization is particularly noticeable in records prior to 1980.
relations shown up by the analysis match patterns previously reported for Mexico as a whole (Magaña et al., 2003). Nevertheless, because local impacts of El Niño have sometimes been found to vary quite markedly from those expected over a wider region (Trenberth and Caron, 2000) this confirmation of a general pattern for the state of Chiapas incorporating a larger number of local data points than those used in the study by Magaña et al. (2003) is particularly useful.

The remote effects of the ENSO cycle, referred to as teleconnections, are rather well understood both empirically and through modeling studies. Teleconnections are typically most pronounced during the northern hemisphere winter. During El Niño, the Aleutian low pressure centre over the North Pacific deepens, high pressure develops over western North America, and low pressure tends to prevail over the southeastern US. The pressure changes drive warmer air masses from southern latitudes into the Pacific Northwest and southern Canada, leading to mild winters in the northern part of the North American continent (Cole and Cook, 1998). Low pressure in the southeastern US brings wetter, cooler conditions to the states bordering the northern half of the Gulf of Mexico and drier conditions to the south. In the state of Chiapas the winter months are typically very dry. Given that the lower bound for monthly precipitation is zero, any slight negative change in precipitation during this period does not show up clearly in any analysis. It is the reduction in the previous year’s late season rainfall and prolongation of the dry season during El Niño years which when combined with a strong increase in precipitation during the latter part of the year associated with cooling mid Pacific sea surface temperatures that produces the very pronounced cyclical precipitation pattern.

The relationship between precipitation and the ENSO cycle varies in strength and form throughout the North American continent (Cole and Cook, 1998). Thus, the apparent change in the synchrony of the two cycles are not unusual. The results are significant for decision makers because local impacts of El Niño have sometimes been found to vary quite markedly from those expected over a longer time frame. The results are significant for decision makers where. The results confirm and reinforce other modelling and empirical studies.

Conclusion

The method used allowed a clear cyclical pattern to be extracted from the data. This pattern is correlated with changes in mid Pacific sea surface temperatures. The extent to which these associations are causal has been investigated elsewhere. The results confirm and reinforce other modelling and empirical studies.

ACKNOWLEDGEMENTS

This work was financed by CONACYT and SEMARNAT through “Fondos Sectoriales” code N°831, as part of the project “Uso sustentable de los Recursos Naturales en la Frontera Sur de Mexico”.

REFERENCES


PATRONES DE PRECIPITACIÓN EN EL SURESTE DE MÉXICO Y SU CORRELACIÓN CON EL ÍNDICE DE TEMPERATURA SUPERFICIAL MARINA DE LA CORRIENTE EL NIÑO
John Duncan Golicher, Neptalí Ramírez-Marcial y Samuel Israel Levy Tacher

RESUMEN
Los datos históricos disponibles para la investigación de cambios climatológicos en los trópicos son típicamente escasos y mal distribuidos en espacio y tiempo. En este estudio fueron aplicados métodos de interpolación espacial para llenar la ausencia de datos de precipitación en el estado de Chiapas. Se eligió el método de Kriging universal iterado, combinando una función para la covarianza espacial con una superficie de tendencia polinomial. Para evitar que puntos extremos de calidad dudosa influyeran en el análisis fueron eliminados en forma automatizada. Se replicó el procedimiento 612 veces para producir series de tiempo completas desde 1950 hasta 2000. Para detectar tendencias temporales se descompusieron los datos en tres componentes, estacional, tendencia y ruido, usando el método de curvas suavizadas localmente ponderadas (loess smoothing, STL). El componente estacional fue removido y los restantes ajustados para encontrar una tendencia. Un procedimiento idéntico fue aplicado al índice Niño3.4 de temperatura de la superficie marina. El componente de tendencia de cada arreglo de datos fue analizado para identificar auto-correlación y correlación cruzada. La función de auto-correlación para el número de días con lluvia mostró correlación significativa entre puntos con separación entre tres y cuatro meses. Se encontró una correlación negativa significativa entre el índice de la temperatura de la superficie marina y la lluvia. La técnica condujo a una clara descripción de un patrón que podría emplearse para predecir parcialmente eventos relacionados con la precipitación como inundaciones e incendios.

CORRELACIÓN ENTRE LOS DATOS DE PRECIPITACIÓN NO SUL DE MÉXICO Y EL ÍNDICE DE TEMPERATURA NA SUPERFÍCIE MARINA DURANTE O FENÔMENO “EL NIÑO”
John Duncan Golicher, Neptalí Ramírez-Marcial y Samuel Israel Levy Tacher

RESUMO
Os dados históricos disponíveis para a investigação de mudanças climatológicas nos trópicos são tipicamente escassos e mal distribuídos no espaço e tempo. Neste estudo foram aplicados métodos de interpolação espacial para preencher a ausência de dados de precipitação no estado de Chiapas. Elegeu-se o método de Kriging universal iterado, combinando uma função para a covariança espacial com uma superfície de tendência polinomial. Para evitar que pontos extremos de qualidade duvidosa influissem na análise, foram eliminados em forma automatizada. Se repetiu o procedimento 612 vezes para produzir séries de tempo completas desde 1950 até 2000. Para detectar tendências temporais se descompuseram dos dados em três componentes, estacional, tendência e ruído, usando o método de curvas suavizadas localmente ponderadas (loess smoothing, STL). O componente estacional foi removido e os restantes ajustados para encontrar uma tendência. Um procedimento idéntico foi aplicado ao índice Niño 3.4 de temperatura da superfície marinha. O componente de tendência de cada arranjo de dados foi analisado para identificar auto-correlação e correlação cruzada. A função de auto-correlação para o número de dias com chuva mostrou correlação significativa entre pontos com separação entre três e quatro meses. Encontrou-se uma correlação negativa significativa entre o índice da temperatura da superfície marinha e a chuva. A técnica conduziu a uma clara descrição de um padrão que poderia empregar-se para prever parcialmente eventos relacionados com a precipitação como inundações e incêndios.