INT. J. REMOTE SENSING, 1998, VOL. 19, NO. 10, 1843–1871

On the sources of vegetation activity variation, and their relation with water balance in Mexico

F. MORA

Center for Advanced Land Management Information Technologies, University of Nebraska-Lincoln, 113 Nebraska Hall, Lincoln, Nebraska 68588-0157, USA; email: mora@tan.unl.edu

L. R. IVERSON

USDA Forest Service, 359 Main Road, Delaware, Ohio 43015, USA

(*Received 13 August 1996; in final form 18 July 1997*)

Abstract. Natural landscape surface processes are largely controlled by the relationship between climate and vegetation. Water balance integrates the effects of climate on patterns of vegetation distribution and productivity, and for that reason, functional relationships can be established using water balance variables as predictors of vegetation response. In this study, we evaluate, at the country and ecoregion level of analysis, the relationships between indicators of vegetation productivity and seasonality with several water balance variables. Vegetation indicators were derived from multitemporal analysis of satellite images, and water balance variables were obtained from ground meteorological station data. Spatial and temporal variation of climate and vegetation were evaluated with remote sensing and GIS technology, and empirical relationships were evaluated statistically via regression models. Significant non-linear relationships were established for vegetation productivity, precipitation, and actual evapotranspiration at the country level in Mexico, where the landscape is represented by a wide diversity of ecosystems. Variation of vegetation patterns of productivity and seasonality is explained less at the ecoregion scale relative to the country level, but water balance variables still account for ~50% of variation in vegetation.

1. Introduction

Land surface processes, such as primary productivity, energy balances (e.g., evapotranspiration processes), and biogeochemical cycles, are largely controlled at landscape scales by the interaction of climate with terrestrial vegetation. For that reason, vegetation disturbances can greatly modify landscape ecological processes. Presently, there is great concern that high rates of land surface modification, and therefore modification of ecological processes, are occurring due to anthropogenic causes such as deforestation and other land-use changes.

Modifications of landscape processes at the regional level are particularly important in developing countries where high rates of deforestation are occurring. Ten major fronts of active deforestation of tropical vegetation have been identified for the globe, five of which are located in Latin American countries, with Mexico ranking near the top (Myers 1993). According to Mexican officials, land-use modification in Mexico is occurring at more than 1% per year considering all vegetation types, increasing in recent years. During the last 30 years, more than 25% of the forested cover has been lost (Inventario Nacional Forestal [INF] 1985, 1991).

Long- and short-term evaluation methods are needed to monitor landscape processes and modifications under such situations. Some national and international efforts are beginning to better monitor conditions over landscapes susceptible to high rates of change, but a large scale effort is still needed to gather observations that can be used to evaluate functional relationships. In general, the current availability of data in developing countries (as in Mexico) is poor, which in turn, largely restricts the functional analysis of land-surface processes. Due to these limitations, remotely sensed data currently provides the most appropriate tool for the evaluation of landscape processes at regional scales in these countries.

The general objective in this study is to evaluate the value of remotely sensed data linked to statistical modelling to provide a tool for the analysis of ecological processes at landscape scale. The evaluation is based on the analyses of empirical relationships between integrated and seasonal measures of remotely sensed vegetation indexes with annual water balance variables that can be estimated from ground observations. First, different sources of variation in vegetation activity are analysed as a response of different scales of observation and explanatory variables. Later, a series of empirical models are fitted to explain such variation. The relationships between seasonal variations of vegetation and water balance can help elucidate those mechanisms that regulate vegetation–climate surface processes in the landscape of Mexico.

1.1. *Background*

Evidence continues to mount as to the value of remotely sensed imagery for the assessment of landscape processes. For example, land-surface processes at continental and regional scales have been related to the normalized difference vegetation index (NDVI), derived from satellite imagery. Net primary productivity, potential and actual evapotranspiration and atmospheric CO_2 dynamics have been correlated with NDVI at several scales and in different parts of the globe (Box *et al.* 1989, Chong *et al.* 1993, Choudhury 1987, Fung *et al.* 1987, Goward *et al.* 1985, 1987, Maisongrande *et al.* 1995, Running 1986, Running and Nemani 1988, Running *et al.* 1989, Tucker and Sellers 1986, Tucker *et al.* 1986). In addition, seasonal patterns of vegetation indexes can also be used to estimate climatic variability (Gallo 1989, 1990).

Direct estimation of variables associated with regional water balance is potentially a major constraint to functionally linking land surface processes for regions where data are scarce. Actual evapotranspiration and soil moisture are extremely difficult variables to estimate without making several, and sometimes very general, assumptions. Although, estimates of water balance variables can be obtained using several methods that include the use of satellite imagery along with simulation modelling, they are yet to be adequately calibrated with ground observations (Pinker 1990).

At present, water balance variables estimated via empirical methods give regional values based on a few climatic variables that are more readily available from meteorological stations, i.e., temperature, precipitation, direction and speed of wind, and relative humidity (Thornthwaite and Mather 1957, Eagleman 1980). Measurements of air temperature and precipitation are the only meteorological variables used in water balance calculations currently being gathered at most Mexican weather stations. For that reason, the Thornthwaite and Mather (1957) approach for water balance was used because other accurate formulae require data such as wind speed which is not available. Therefore, among all possible methods to use, the Thornthwaite and Mather method was the most suitable procedure for estimating water balance in this study.

Empirical functional relationships between remotely sensed vegetation parameters and water balance variables have previously been established using integrated NDVI (iNDVI) or a similar measure from the Advanced Very High Resolution Radiometer (AVHRR) sensor that can biologically interpreted. NDVI measures have been highly correlated with water balance variables, specifically with actual evapotranspiration, soil moisture, and precipitation (Davenport and Nicholson 1993, Farrar *et al.* 1994, Kustas *et al.* 1994, Di *et al.* 1994, Malo and Nicholson 1990, Nicholson and Farrar 1994, Sandholt and Andersen 1993, Seevers and Ottmann 1994).

Integrated annual and seasonal NDVI measures (which in turn, are related to vegetation activity) can be obtained by applying principal component analysis to monthly derived NDVI indexes from AVHRR data. Principal component analysis resulted in several NDVI measures that capture the seasonality in the Mexican landscape (Mora and Iverson 1997). Such components can be used as response variables in water balance processes. Integrated annual measures of NDVI can be correlated to annual variations of potential and actual evapotranspiration, soil moisture, precipitation, and the surplus or deficit of water. It is more difficult, however, to correlate seasonal NDVI measures with seasonal variations of water balance, because they vary according to specific vegetation types. Empirical relationships of seasonal patterns are therefore harder to evaluate (Chong *et al.* 1993). In such cases, the scale at which seasonal water balance controls the vegetation distribution appears to be different from the scale at which the annual water balance variation produces its effects.

When an empirical relationship between vegetation activity and water balance is found, spatial autocorrelation effects among such processes should also be considered. This analysis is particularly important because such landscape processes can be significantly correlated if they share a common spatial structure. It is therefore necessary to identify the sources of environmental variation by 'partialling out' the spatial component in correlation analysis, especially if linear regression models are used.

There are four major components of landscape variation when analysing spatially referenced data (Legendre 1993, Bocard *et al.* 1992): (1) non-spatial environmental variation; (2) spatially structured environmental variation; (3) spatial variation of the process under consideration; and (4) the unexplained, non-spatial variation. These components of variation can be identified using partial regression analysis, after empirical relationships between water balance variables and vegetation activity are established. Partial regression analysis involves the use of multiple regression models that include 'space' as an explanatory variable along with the environmental variables.

2. Methodology

The overall procedure used to evaluate the relationships between measures of NDVI and water balance in Mexico considers two scales of analysis, the country scale and the ecoregion scale. At the country scale, variability among ecosystems permits the evaluation of NDVI variations over a complete set of different ecological situations, e.g., deserts, semideserts, conifers, dry deciduous selvas (we use the term *selva* here, which can loosely be translated as 'tropical forest'), savannas, and perennial selvas. At the ecoregion scale, the variability in NDVI and water balance is largely attributed to more subtle differences among ecosystems within ecoregions.

The possible relationships at the two scales of analysis were explored through correlation, multiple regression, and non-linear regression analysis. Initially, the correlation between NDVI measures and water balance variables served to identify the variables to use in model fitting. Afterwards, non-linear regression analysis, using a previous model structure (Box *et al.* 1989) was used to fit the relationship between the most significant variables that explained vegetation indicators without considering spatial effects in water balance. Finally, partial correlation analysis (Legendre 1993) is used to explore the effects of spatial autocorrelation in the models and to identify the spatial structure of both dependent and independent variables. The models used to explore the combined effects of water balance variables over vegetation activity indicators were evaluated via multiple regression analysis.

2.1. *Integrated and seasonal NDVI measures*

In an earlier study, annual integrated and seasonal NDVI measures were obtained from principal component analysis (PCA) of the Global Vegetation Index (GVI) data produced by the National Oceanic and Atmospheric Administration (NOAA) AVHRR (Mora and Iverson 1997). The first five principal components obtained, captured more than 95% of the monthly GVI variation in Mexican data. Integrated annual vegetation activity was highly related to the first principal component which can therefore be interpreted as another measure of annual integrated NDVI (iNDVI). Seasonal variations in natural vegetation, which key on the temporal variability of chlorophyll (e.g. the July–August NDVI monthly values normally mark the peak of greenness), were mostly captured by the second principal component (sNDVI). Thus, sNDVI was a measure of natural vegetation that followed a strong seasonal pattern in Mexico. The other three components were associated with irrigated agricultural vegetation, and were not considered further in this study.

2.2. *Water balance data*

Direct observations of water balance variables in Mexico are not currently available. Climatic data is gathered in a national network of weather stations where observations of precipitation and air temperature are recorded (INEGI 1980). Water balance was estimated from these records. Potential relationships between water balance and vegetation activity were established using data for 2214 weather stations, which provided long-term monthly means (~25 years) of temperature data and precipitation. The long-term temporal variation was therefore captured. Since additional parameters such as the direction of wind and relative humidity were not available for all stations, estimates of several water balance variables (potential and actual evapotranspiration, soil moisture; water deficit and surplus) were empirically obtained according to Thornthwaite and Mather methodology (Thornthwaite and Mather 1955).

2.2.1. *Water balance estimation according to the Thornthwaite and Mather approach*

Thornthwaite and Mather's approach for water balance modelling has been implemented and used for more than 40 years. Even though it has been criticized due to its empirical approach, this method represents about the only way to estimate the water balance for places where only records of air temperature and precipitation exist. Modelling algorithms which use their equations (e.g., WATBUG, from Willmot

1977) are available to give water balance estimates from inputs of mean monthly air temperature, mean monthly precipitation, and some indication of water holding capacity for the location. Formulas described in WATBUG were used here to implement a cartographic GIS model of water balance. Five water balance (*WB*) variables were calculated from records on precipitation (*P*), air temperature (*T*), latitude, and duration of daylight. These include soil moisture (*SM*), potential evapotranspiration (*PE*), actual evapotranspiration (*AE*), water deficit (*WD*), and water surplus (*WS*). Since the water balance approach used here integrates the effects of several factors, estimates of water balance variables, except for *P* and *PE*, contain some dependency on vegetation types through the soil moisture storage (figure 1).

Soil moisture (*SM*) is the amount of water that is stored in the soil, and is available for plant growth. By definition, soil moisture is stored only when precipitation exceeds the potential evaporative demand (*PE*) of the atmosphere, otherwise the precipitation is evaporated. The maximum amount of soil moisture that can be available for plant growth on a specific site is a direct function of the water-holding capacity on that site (primarily soil structure but also rooting depth of the vegetation layer), as modified over time by the existing vegetation and *PE*. Soil maps with sufficient detail on water-holding capacities, and sufficient extent to cover all Mexico were not available. As such, inaccuracies of *SM* will occur at the fine scale. However, when calculating soil moisture here, we can assume that large-scale and long-term effects of precipitation and potential evapotranspiration will generally overwhelm the effect of variations in water-holding capacity among soil types.

Evapotranspiration is the process of water transfer from vegetated land surfaces

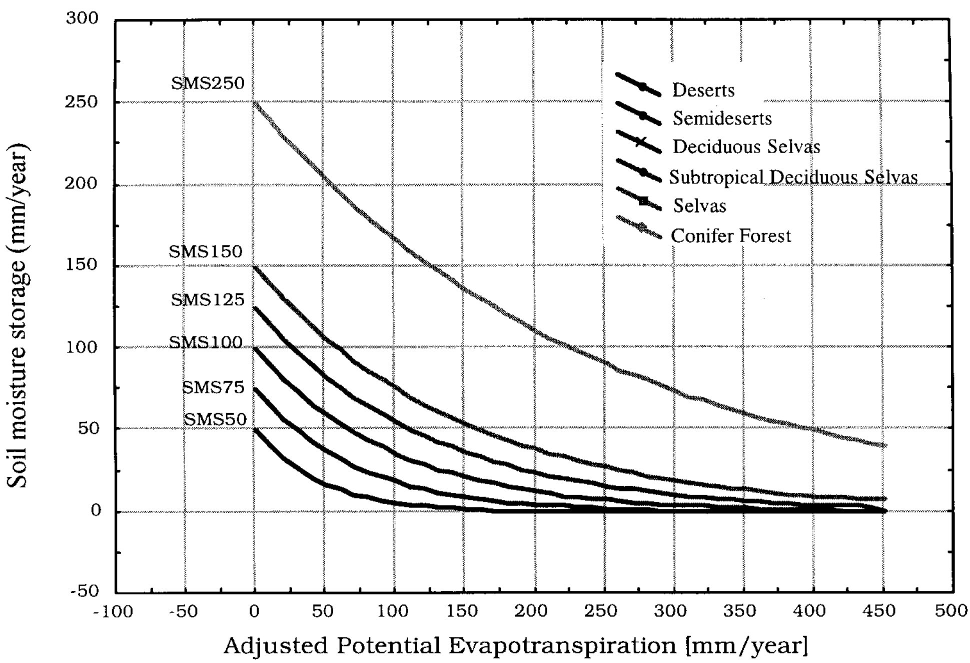


Figure 1. Exponential relationship for soil moisture retention data reported in Thornthwaite and Mather (1955). Soil moisture storage is plotted as a function of adjusted potential evapotranspiration using different vegetation types which, in turn, have various rooting zone depths.

into the atmosphere due to soil evaporation and plant transpiration (Rosenberg *et al.* 1983). In the Thornthwaite and Mather approach, evapotranspiration is mainly a direct function of air temperature, by being directly related to the amount of energy available at the soil-plant surface. As defined by Thornthwaite, potential evapotranspiration (PE) can be expressed as the evaporative water loss from a site covered by vegetation that receives unlimited amounts of water (Thornthwaite 1948). As it is directly related to both heat and radiation, PE is modified by humidity and wind speed (Stephenson 1990), but over wide geographic areas, substantially more modification results due to changes in latitude and duration of daylight (Willmott 1977). According to Thornthwaite's approach, PE is calculated here as a direct function of air temperature and heat index, and adjusted by latitude and duration of day (Willmott 1977).

Actual evapotranspiration (AE), on the other hand, is the actual water transferred from the surface to the atmosphere in accordance with present meterological, plant, and soil conditions, and depending upon available water. In an ecological context, AE can be defined as the biologically usable energy and water used by plants (Stephenson 1990). It is expressed as the amount of evaporative water loss in relation to its present availability. According to the Thornthwaite method, AE is estimated from available soil moisture (SM) and precipitation (P). When there is a water deficit in the soil (i.e., $PE > SM$) AE equals PE, otherwise AE is equal to the amount of precipitation (P) plus the moisture accumulated in the soil (SM).

Estimates of actual evapotranspiration require soil moisture (SM) estimates in advance. Unfortunately, soil water-holding capacity, a variable required for soil moisture estimation, was not available. Alternatively, an approach that used information related to the water that is retained in the soil from a series of soil moisture retention tables (Thornthwaite and Mather 1957) was used. These soil moisture retention tables were used together with land cover type information to estimate soil moisture. The soil moisture retention tables assumed that moisture accumulated in the soil is depleted exponentially as a function of PE, and varies according to the depth at which the water is held in the rooting zone. As water-holding capacity is a function of soil structure and rooting depth, there is a relationship between water-holding capacity and vegetation type (e.g., mature forests have a much deeper rooting zone, and therefore a higher water-holding capacity, than shallow-rooted crops, regardless of the soil texture). Empirical relationships like these permit the estimation of soil moisture storage, directly from soil moisture retention tables, for different dominant vegetation types within ecoregions, and for different depths of the rooting zone.

Thornthwaite and Mather (1957) previously published several soil moisture retention tables for different vegetation types and depths of the rooting system for crops and natural vegetation. From the data published in the tables, several log-transformed regression equations were fitted to describe how soil moisture is depleted by vegetation according to PE at different depths of rooting zones (see figure 1). These parameters were used to estimate the soil moisture in the different ecoregions according to dominant vegetation. Curves for water-holding capacities of 50 mm in the rooting zone were used for 'deserts', 75 mm for 'semideserts', 100 for 'deciduous selvas', 125 for 'subtropical matorrals', 150 mm for 'selvas' and 250 mm for 'conifer forests'. Obviously, if an adequate soil map was available, the use of that map in conjunction with a vegetation map could provide an improvement to the method employed here to estimate soil moisture.

A *water deficit* (*WD*) is present when water availability does not meet the vegetation evaporative demands. By definition it is the difference between monthly *PE* and *AE* (Stephenson 1990). On the other hand, a *water surplus* (*WS*) is an excess of water in the environment, and is obtained when precipitation exceeds *PE*, minus the water that is retained in the soil. *Precipitation* (*P*) is the amount of water that is received as rainfall.

2.3. *GIS implementation*

A cartographic model, which included several NDVI measures, a map of vegetation types derived from remotely sensed images (covering the whole country with a 16 km pixel size), and the water balance database for 2214 weather stations in Mexico, was implemented in Arc/Info GIS (ESRI 1995) and used to explore the climate-vegetation relationships. The implementation of a GIS cartographic model permitted the evaluation of the relationship between NDVI measures and water balance variables at two scales of analysis: (1) the country scale; and (2) the ecoregions scale.

The classification of the country into different ecoregions (figure 2) was used as a stratification criterion for the two levels of analysis. A subset of 1162 weather stations was obtained by masking their location with six 'natural' ecoregions (occupying 69% of the Mexican landscape, not including irrigation and agriculture from figure 2). A data set was thus created that included the six annual water balance variables, annual integrated and seasonal NDVI measures, ecoregions, and the geographic location of weather stations used in the analysis. These data were then used to fit linear and non-linear models.

Two special advantages are gained when seasonal water balance variables are calculated with the aid of GIS. First, potential evapotranspiration can be adjusted by latitude and duration of daylight to produce adjusted potential evapotranspiration (*ADPE*). Secondly, soil moisture can be estimated from *PE*, using the parameters of soil moisture retention tables associated with the dominant vegetation within ecoregions (figure 1). The temporal variation of water balance can thus be associated with qualitative differences in vegetation types ('natural' vs. 'non-natural'), when using a cartographic model that includes such characteristics in vegetation types.

Although the water balance-vegetation relationships were statistically assessed on a point basis for each of the weather stations, the GIS implementation also allowed the point data to be interpolated using spherical kriging models in Arc/Info (ESRI 1995). This resulted in country-level maps for each of the six variables, and allowed the exploration of their scales and forms of variation. Maps of the predicted results of the regression models permitted a visual evaluation of the forms of variation over the country.

2.4. *Statistical exploratory analysis*

Exploratory analysis was conducted on the data in order to establish potential relationships. Possible climate-vegetation relationships were established based on graphical analysis (scattergrams) and by using Pearson's product-moment correlation coefficients evaluated at $p=0{\cdot}05$.

2.5. *Mode fitting*

Non-linear relationships were tested using the model proposed by Box *et al.* (1989). Even though the model could be fitted by transformed linear regression, the

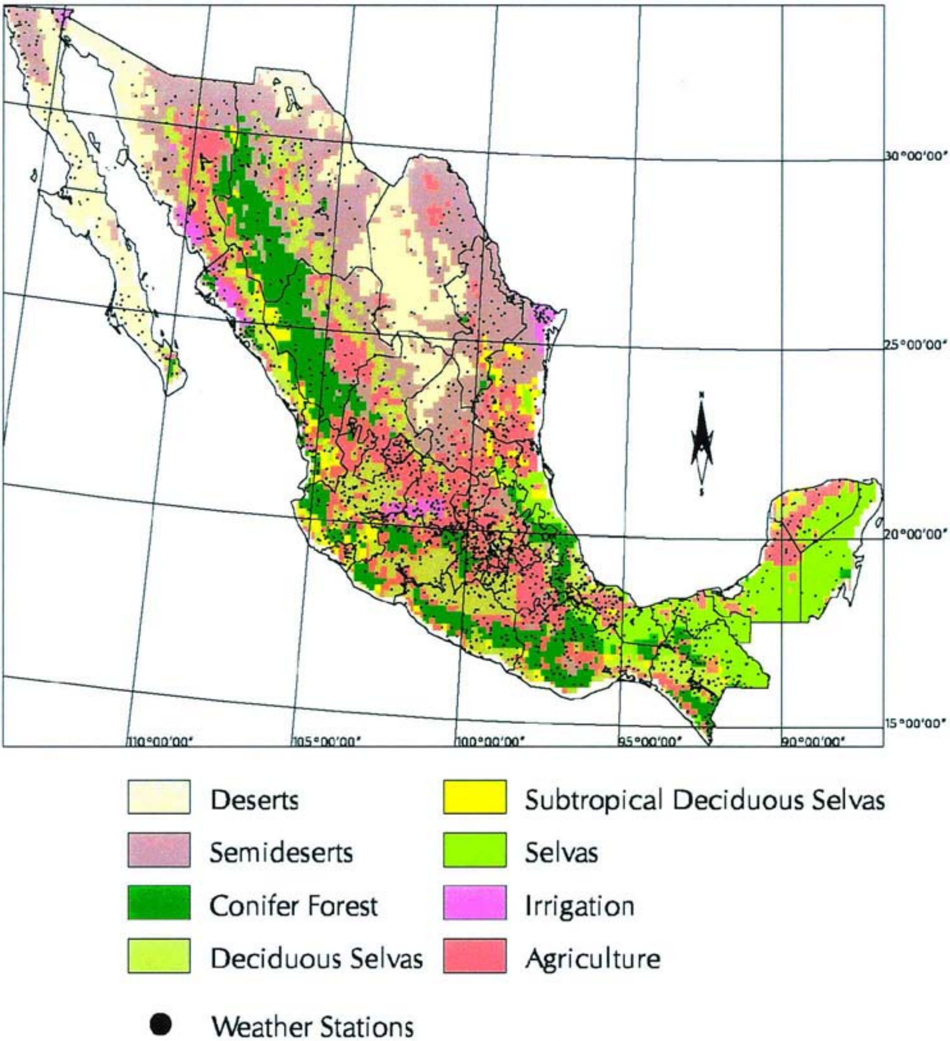


Figure 2. Distribution of weather stations within ecoregions in Mexico. The classification into ecoregions is described in Mora and Iverson (1997).

non-linear approach was used in order to prevent autocorrelation effects in the model parameters. The model has strong theoretical support, and its use permits a direct comparison between the parameters obtained here and those obtained by Box *et al.* (1989). Box's model that describes the relationship between NDVI measures and water balance variables has the following form:

$$\mathrm{iNDVI} = \alpha[1 - e^{-\beta*(WB)}] \quad (1)$$

where iNDVI=integrated annual variation of NDVI, as produced via principal component analysis for the period of 1985–1989 (the iNDVI is scaled to a 0–1 range); α=asymptote or highest iNDVI value; β=slope or iNDVI rate of change as a function of *AE* or *P* (*WB* units in mm/year^{-1}); and *WB*=water balance variable.

In the following models $WB = AE$ when actual evapotranspiration (mm/year^{-1}) is used; and $WB = P$ when precipitation (mm/year^{-1}) is used.

The model was fitted using a loss non-linear function, which minimizes the residual variance (sum of squared deviations) around the regression line, using a quasi-Newton minimization method (first order and second order derivative of the function).

Linear regression analysis was performed using the least squares method. In both non-linear and linear regression methods, the plot of observed vs. predicted values, normal and half-normal probability plots of residuals, and the proportion of variance explained were used to evaluate the model fit.

2.6. *Multivariate regression analysis*

Multiple regression analysis was used to test the interaction between water balance variables and spatial autocorrelation, in explaining the variation of both iNDVI and the sNDVI at the country scale. Subsequently, the combined effect of water balance variables over vegetation indicators included the water balance spatial structures. At the ecoregion scale, only environmental variation (without partialling out the spatial component) was tested in the regression models. The Durbin–Watson test, histograms, normal probability, and standard residuals vs. predicted plots were used in residual analysis. Tolerance and variance inflation factor (VIF) values were used in multicollinearity diagnostics for all regression models.

2.6.1. *Partial correlation analysis*

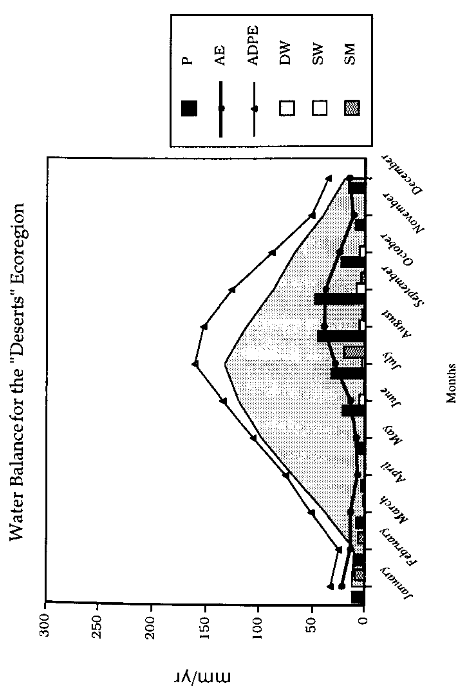
At the country scale, stepwise multiple regression analyses was used to identify four sources of iNDVI and sNDVI variation (*a*, *b*, *c*, *d*; see table 1). First, nine spatial variables that define the spatial component in the analysis were regressed on each water balance variable to determine their correspondent spatial structure. The nine spatial variables were obtained when a matrix of two-dimensional geographical co-ordinates (x=longitude and y=latitude) was completed by adding all terms for a cubic trend regression surface of the form:

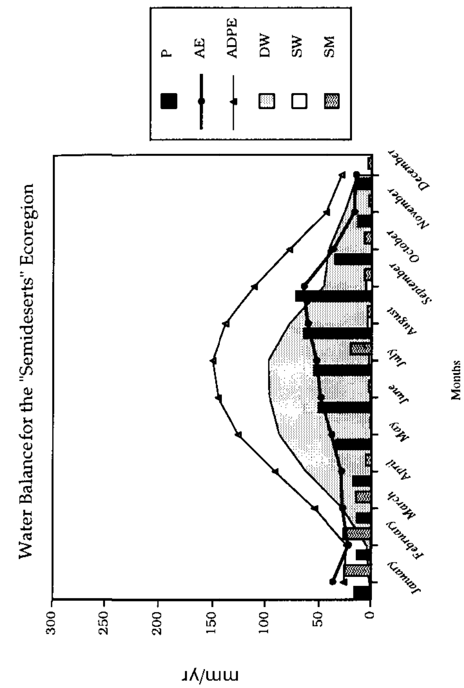
$$z = b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3 \quad (2)$$

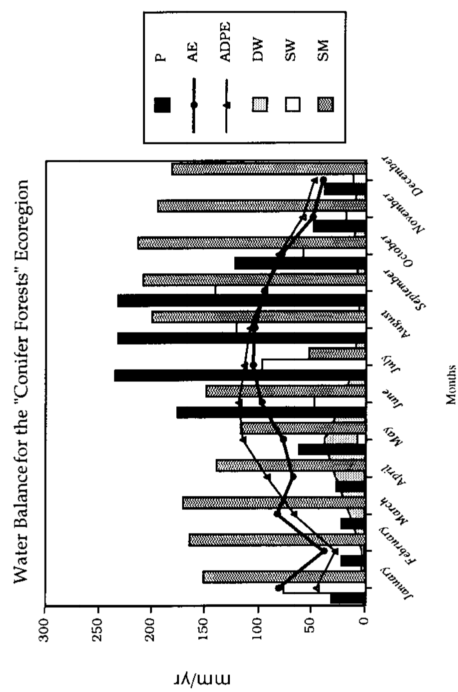
This cubic form of geographic co-ordinates accounts not only for linear gradient patterns, but also complex features such as patches or gaps (Bocard *et al.* 1992).

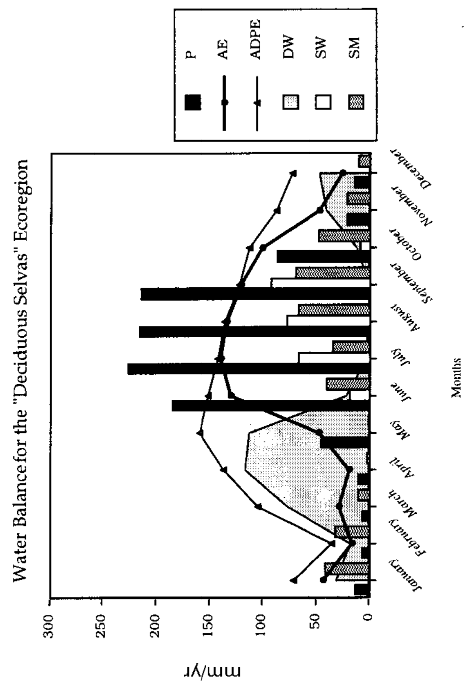
Table 1. Sources of iNDVI and sNDVI variation.

Components of variation in NDVI–space	Equations and regression models
Non-spatial *WB* variation [*a*]	iNDVI = [*AE*res, *PE*res, *P*res, *SM*res, *WD*res, *WS*res]
Spatially structured *WB* variation [*b*]	iNDVI = [*AE*, *PE*, *P*, *SM*, *WD*, *WS*]–[*a*]
iNDVI spatial variation independent of *WB* [*c*]	iNDVI = [*f*(LONG, LAT))]–[*b*]
WB (environmental) variance [*a*+*b*]	iNDVI = [*AE*, *PE*, *P*, *SM*, *WD*, *WS*]
Spatial structure [*b*+*c*]	iNDVI = [*f*(LONG, LAT)]
Environmental–spatial variation combined [*a*+*b*+*c*]	iNDVI = [*AE*, *PE*, *P*, *SM*, *WD*, *WS*, *LONG*, *LAT*]
Unexplained variance [*d*]	[*d*] = 1–[*a*]–[*b*]–[*c*]


Water Balance for the "Deserts" Ecoregion
mm/yr
300
250
200
150
100
50
0
January
February
March
April
May
June
July
August
September
October
November
December
Months
P
AE
ADPE
DW
SW
SM


Water Balance for the "Semideserts" Ecoregion
mm/yr
300
250
200
150
100
50
0
January
February
March
April
May
June
July
August
September
October
November
December
Months
P
AE
ADPE
DW
SW
SM


Water Balance for the "Conifer Forests" Ecoregion
mm/yr
300
250
200
150
100
50
0
January
February
March
April
May
June
July
August
September
October
November
December
Months
P
AE
ADPE
DW
SW
SM


Water Balance for the "Deciduous Selvas" Ecoregion
mm/yr
300
250
200
150
100
50
0
January
February
March
April
May
June
July
August
September
October
November
December
Months
P
AE
ADPE
DW
SW
SM

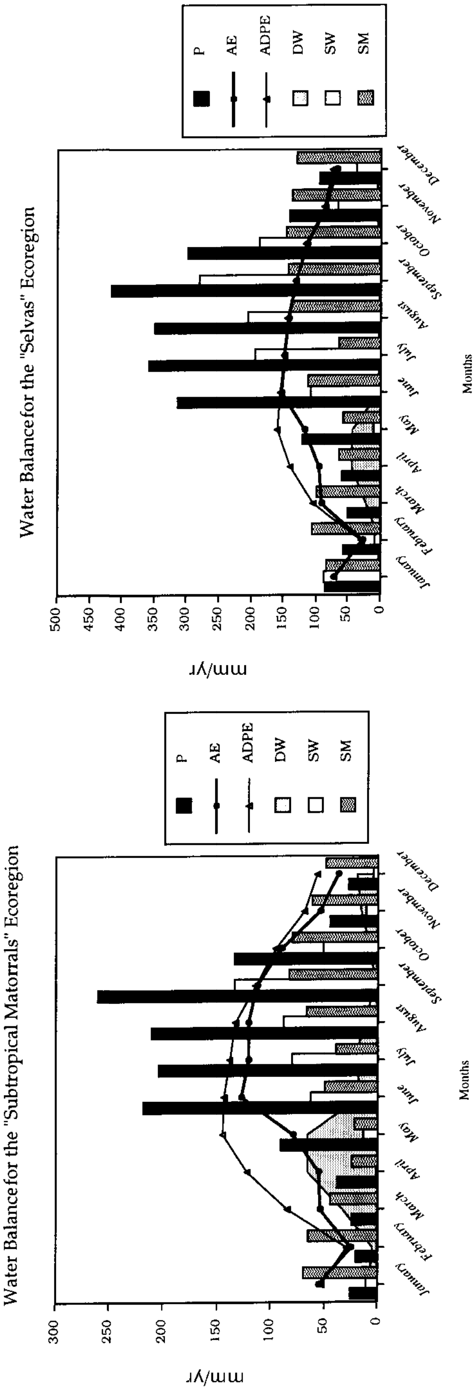


Figure 3. Climographs for the water balance in different ecoregions in Mexico. Water balance units are in mm/year^{-1}.

After the individual linear regression models for the spatial variation of each water balance were fitted, the residuals were retained and regressed on both iNDVI and sNDVI. This defined the non-spatial environmental vegetation variation (*fraction a*). A stepwise method was used also to determine the set of water balance variables that explain the most non-spatial variation in both iNDVI and sNDVI. The fraction that combines the non-spatial with the spatial water balance variation (*fraction a+b*) of the vegetation indicators was modelled by regressing simultaneously all water balance variable against iNDVI and sNDVI. The fraction is defined as $a+b$ because spatial autocorrelation has not been removed in this model. The *fraction b* (spatially structured water balance variation) is obtained by subtracting the *fraction a* from $a+b$. The spatial fraction of vegetation variation (independent from water balance or *fraction c*) is obtained when the set of nine geographical variables (derived from latitude and longitude) are regressed on iNDVI and sNDVI to obtain *fraction b+c*, and then subtracting the *fraction b* from *fraction b+c*. The remaining variances unexplained by either spatial or water balances variables is the *fraction d*. Each fraction was modelled spatially using kriging (ARC/INFO, ESRI 1995) and the resultant surfaces were implemented in a GIS cartographic model.

3. Results

3.1. *Water balance variables*

Patterns in water balance for the different ecoregions in Mexico contrast highly among the different ecoregions (figure 3). These differences lead to the hypothesis that vegetation measures can be associated with the patterns observed in water balance.

Deserts and semideserts are characterized by high annual accumulated *ADPE* and *WD* (figure 3). In these ecoregions, a deficit of water can occur for more than 10 months of the year. In addition, *AE* follows a similar seasonal pattern as precipitation in deserts. In contrast, the conifer forest ecoregion showed a high accumulation of soil moisture throughout the year, and seasonal precipitation during the summer (figure 3). Deciduous selvas and subtropical deciduous selvas showed a distinctively seasonal accumulation of *WD* during the spring, followed by a characteristic four-month rainy season in the summer, when *ADPE* approximates *AE*. Seasonal precipitation is also present in selvas, but with very little accumulation of *WD* and substantially more water surplus than conifers. *AE* roughly equals *ADPE* throughout the year in tropical rainforests (selvas).

Computation of annual water balance variables, and their interpolation across the entire country, provided a realistic picture of water balance in Mexico (figure 4). Actual evapotranspiration was highest along the coasts and in the Yucatan peninsula where precipitation was also greatest. The deserts of northern Mexico showed large annual water deficits, while only a small region of southern Mexico showed a significant water surplus. Potential evapotranspiration rates were highest along the coasts and other low elevation zones, and decreased in the higher elevations. Soil moisture was most related with water surplus and precipitation, as expected. Semivariogram parameters indicated that water balance variables are highly autocorrelated. This indicated that spatial autocorrelation should be included in the analysis when exploring relationships between NDVI indicators and water balance.

3.2. *Relating NDVI to water balance variation at country scale*

Scattergrams of annual water balance variables and NDVI measures as determined from PCA suggest that the relationship between iNDVI and *AE* or *P* is non-

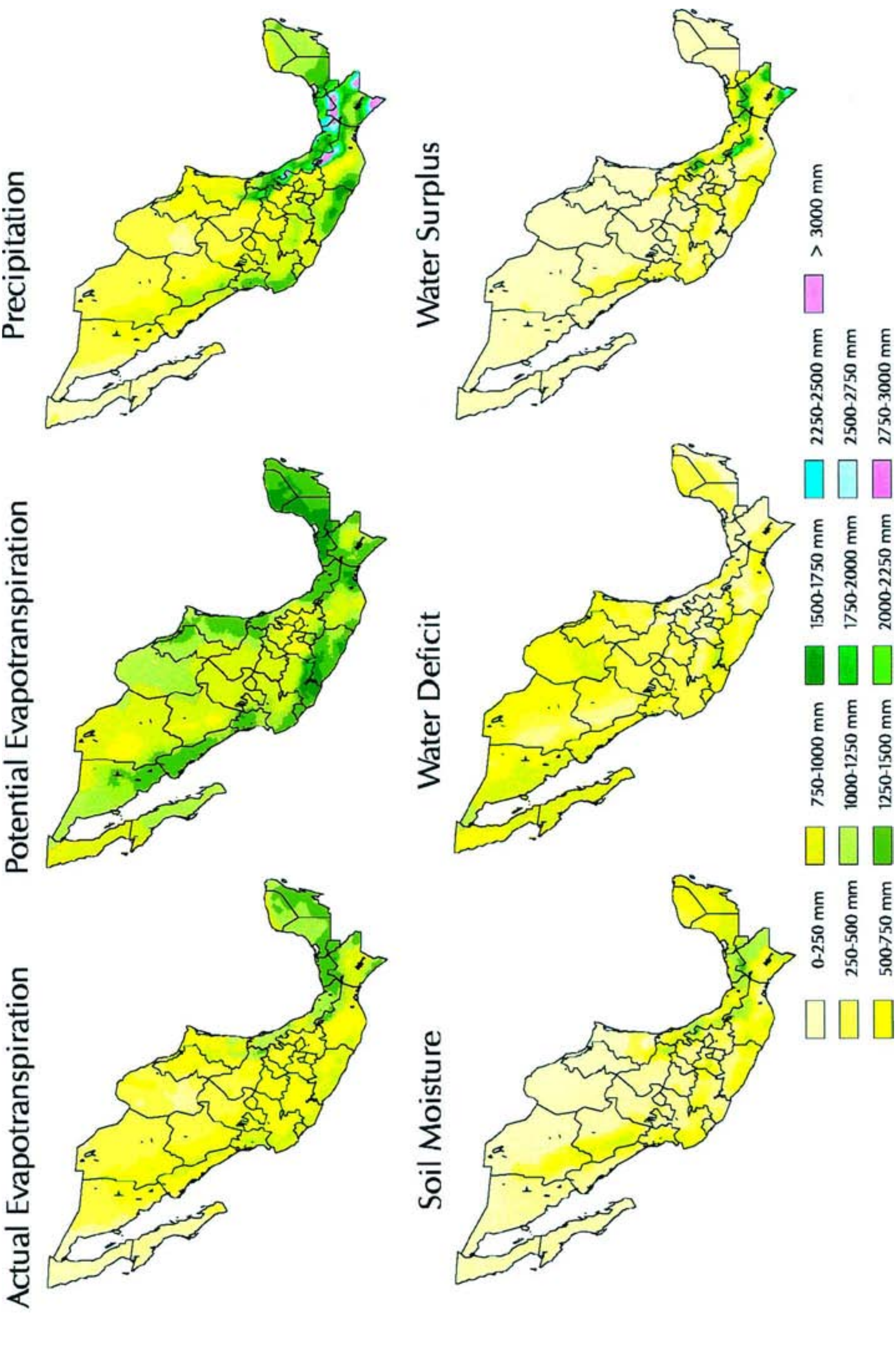


Figure 4. Spatial variation of the annual water balance variables.

linear, while the relationship between iNDVI and *WD* seems to be negative linear with an R^2 of 0·33 ($p<0·001$) (figure 5). No correlation seems to exist for *WS*, *ADPE*, and *SM*. When scattergrams of the seasonal NDVI (sNDVI) were plotted against water balance variables, no apparent correlations were present (figure 5).

A non-linear regression analysis was used to test the relationship between iNDVI and *P* or *AE* (figures 6 and 7). To test the goodness of fit, a pseudo-coefficient of determination (R^2) was obtained from the ratio between the non-linear regression sum of squares to the total sum of squares (SSR/SST), which explained the proportion of variance accounted by the dependent variable (iNDVI) in each model. This measure is helpful to evaluate the fit even if iNDVI is not normally distributed across cases; the R^2 obtained in this way was 0·77 for *AE* (figure 6) and 0·72 for *P* (figure 7).

3.3. *Relating NDVI measures and annual water balance variables at the ecoregion scale*

Correlations between NDVI measures and annual water balance variables at the ecoregions scale are shown in figures 8 to 13. Many Pearson's correlation coefficients are significant (indicated by *, $p<0·05$), where either or both iNDVI and sNDVI variability can be explained by annual water balance variables.

The highest correlation in the desert ecoregion was for the iNDVI with annual soil moisture ($r=0·33$, figure 8). The highest correlation in the semidesert ecoregion was with actual evapotranspiration for both iNDVI ($r=0·48$) and sNDVI ($r=0·59$, figure 9). In the conifer forest ecoregion, the highest correlation was between the sNDVI variation and annual soil moisture ($r=0·39$, figure 10). The deciduous selvas ecoregion had the highest correlations with water balance variables, with both iNDVI and sNDVI being highly correlated with soil moisture and precipitation ($r=0·64$ and 0·52, respectively, figure 11). Subtropical deciduous selvas showed a high negative correlation between soil moisture and the sNDVI (figure 12). Finally, the correlated measures for the selvas ecoregion were *ADPE* and *WS* ($r=0·36$ and $r=-0·35$, respectively, figure 13).

3.4. *Partial regression analysis*

The general results obtained with partial regression analysis are shown in table 2. The total iNDVI variance explained by water balance variability is ∼70% (*fraction a+b*). Of this, only 28% is related to the non-spatial water balance variation (*fraction a*), while 43% of iNDVI variation is associated with the spatial structure (*fraction b*). The spatial iNDVI variation independent of water balance (*fraction c*) is ∼13% and the unexplained variation is ∼25% (figure 14). The mapped results for iNDVI sources of variation are presented in figure 15. Since iNDVI is strongly associated with water balance variables considered here, they will be considered further in the following sections.

In contrast, only ∼49% of the seasonal NDVI variation (sNDVI) can be explained by annual water balance, from which 11% is explained by the total variation (*fraction a+b*) and ∼6% is explained by water balance variability alone (*fraction a*). The variation of sNDVI susceptible to other factors not considered is ∼38% (*fraction c*), while the unexplained variance is ∼51%. Clearly sNDVI is not predicted nearly as well by the annual water balance variables explored in this study.

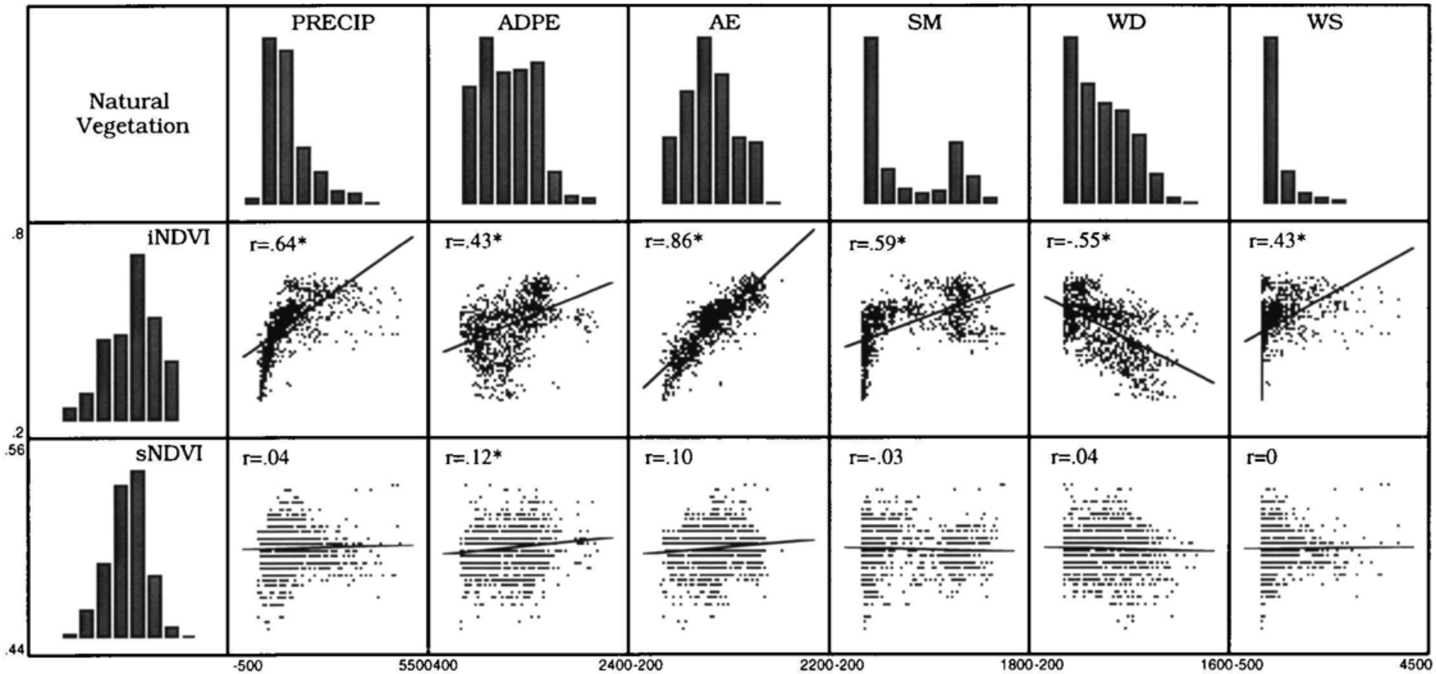


Figure 5. Scatterplot matrix for iNDVI and sNDVI with annual water balance variables for 1112 climatic stations that occur in natural vegetation areas.

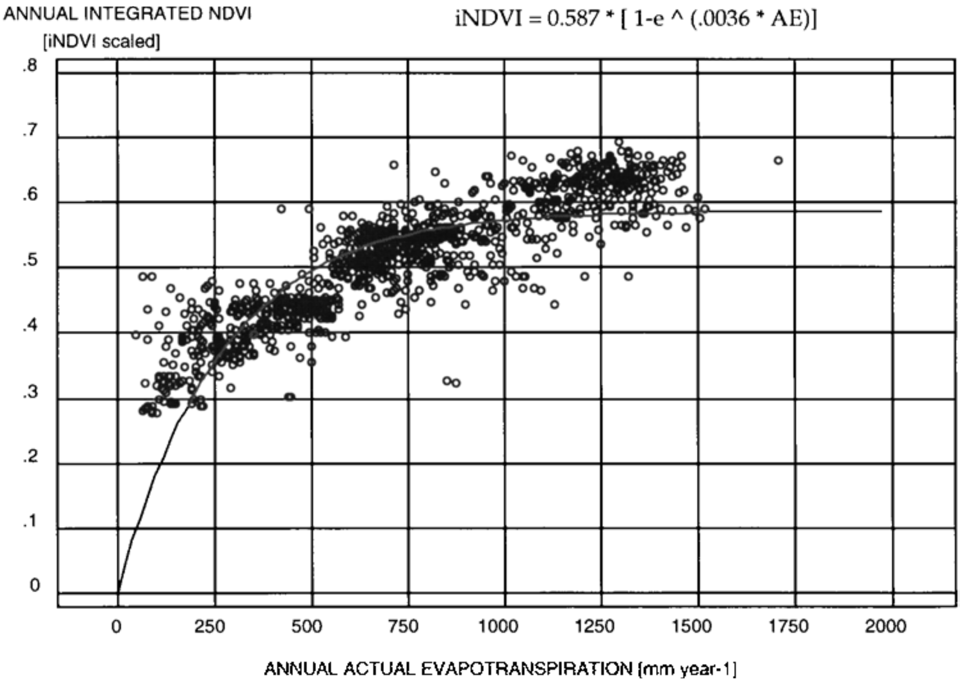


Figure 6. Non-linear regression model [1] for iNDVI as a function of annual actual evapotranspiration. The NDVI measure is scaled to 0–1 units and AE is in mm/year^{-1}.

3.4.1. *Non-spatial (fraction a) and spatial (fraction b) iNDVI variation controlled by water balance*

An analysis of residuals in linear models showed that only actual evapotranspiration (after the spatial effect is partialled out) was significant to explain the iNDVI variation ($p<0{\cdot}01$). As previously observed, *AE* variability has an increasing effect on iNDVI (when space is partialled out), which strongly reinforces the idea of a functional relationship between them.

The best combination of water balance variables in explaining iNDVI accounts for 75% ($R^2=0{\cdot}755$, $p<0{\cdot}05$) of its respective variance (table 2). The stepwise partial regression method identified *AE*, *WD*, *WS*, *P*, and *ADPE* (according to the order in which they were entered in the model) as the most important variables in explaining the iNDVI variation. *AE* alone explained 74% of the total variation, but *WD*, *SW*,

Table 2. Partial multiple regression analysis results (R^2) for NDVI measures as a function of water balance. Percentages of variance explained.

Components of variation in iNDVI–space	iNDVI	sNDVI
Non-spatial *WB* variation [*a*]	0·276	0·060
Spatially structured *WB* variation [*b*]	0·429	0·048
iNDVI spatial variation independent of *WB* [c]	0·135	0·378
WB (environmental) variance [*a*+*b*]	0·705	0·108
Spatial structure [*b*+*c*]	0·461	0·426
Environmental–spatial variation combined [*a*+*b*+*c*]	0·746	0·486
Unexplained variance [*d*]	0·254	0·514

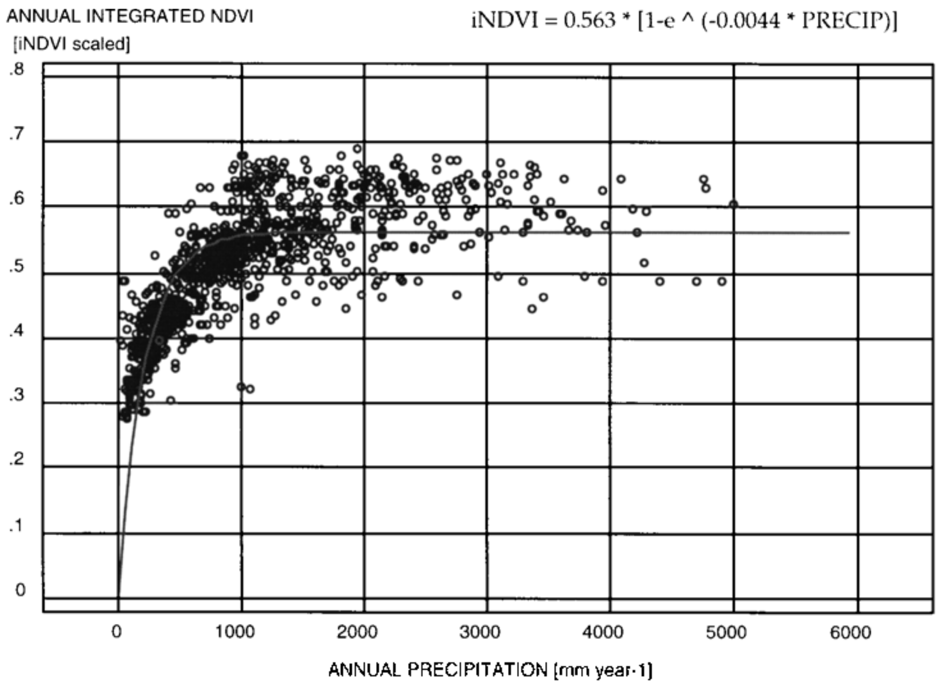


Figure 7. Non-linear regression model [2] for iNDVI as a function of annual precipitation. The NDVI measure is scaled to 0–1 units and P is in mm/year^{-1}.

P, and *ADPE* still explained significant iNDVI variance ($p<0{\cdot}001$). *AE*, *P*, and *ADPE* have an increasing effect on iNDVI, with *AE* and *P* having the relatively most important effects (regression coefficients βs$=0{\cdot}86$ and $0{\cdot}57$, respectively). *WD* and *WS* have a decreasing effect on iNDVI, but interestingly, *WS* has a greater decreasing effect than *WD*.

According to results obtained by modelling the sources of iNDVI variation, the best predictions for the *fraction a* were obtained for intermediate iNDVI values, when compared with the remotely sensed iNDVI (figure 15 (a)). These iNDVI values correspond to conifer and deciduous vegetation, especially along the Pacific coast and in portions of the Gulf of Mexico coast (figure 15 (c)). Predictions at extreme iNDVI values (associated with deserts and rainforests) were not as good. Predictions of iNDVI variation in the Mexican landscape were greatly enhanced when water balance spatial structure was considered in the model (figure 15 (d)). Better predictions were obtained for deserts and selvas with the models based on both *fractions* $a+b$. However, vegetation types in the Mexican Plateau (mostly semideserts) were still poorly represented by this model. This result indicates that there are some important factors (other than water balance) contributing to the distribution of semideserts in Mexico.

Particular attention is given to the *fraction c* that represents the spatial variation of iNDVI (independent from water balance) which is more likely to be explained by factors not considered (Legendre 1993). In the iNDVI case, the variance that can be potentially explained is ~13% (table 2). In contrast, in the sNDVI case, the unexplained spatial variation is ~38% that could potentially be explained by other

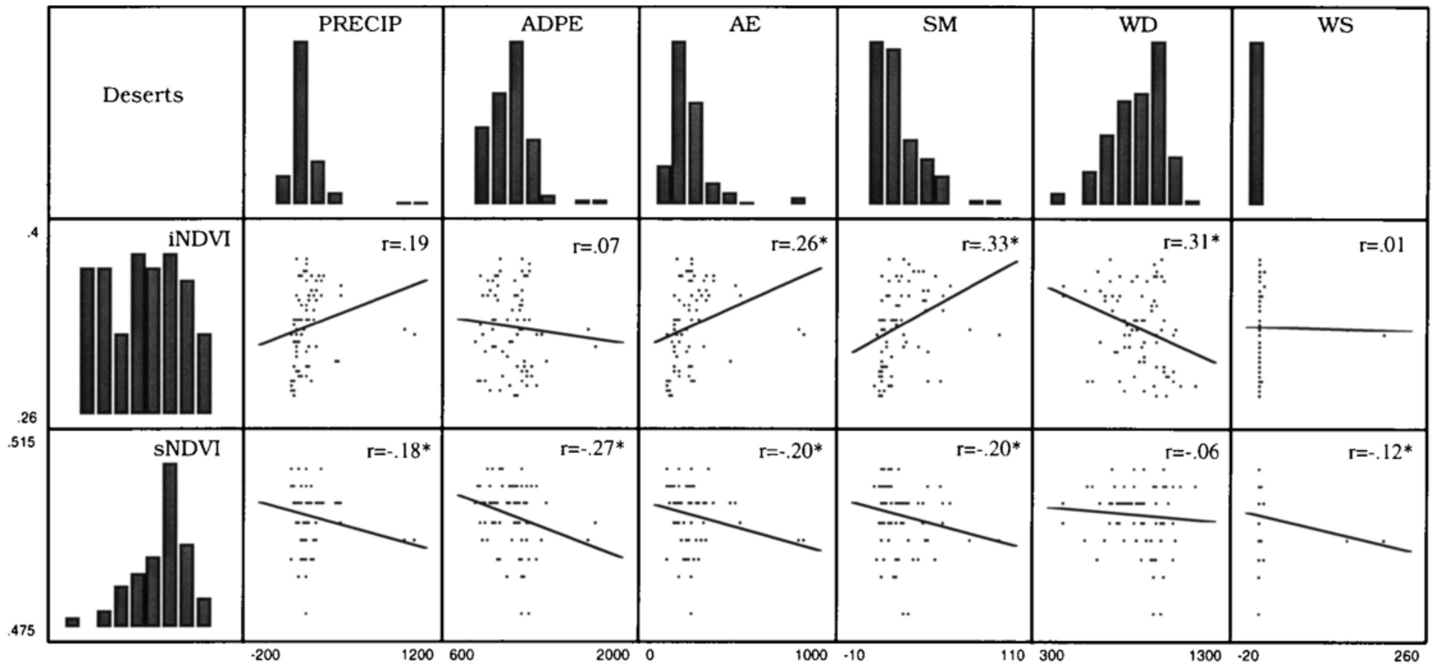


Figure 8. The relationship of iNDVI and sNDVI to annual water balance variables for the 'desert ecoregion' in Mexico. Correlation coefficients (r^*) are also given indicating significance at the 0·05 level.

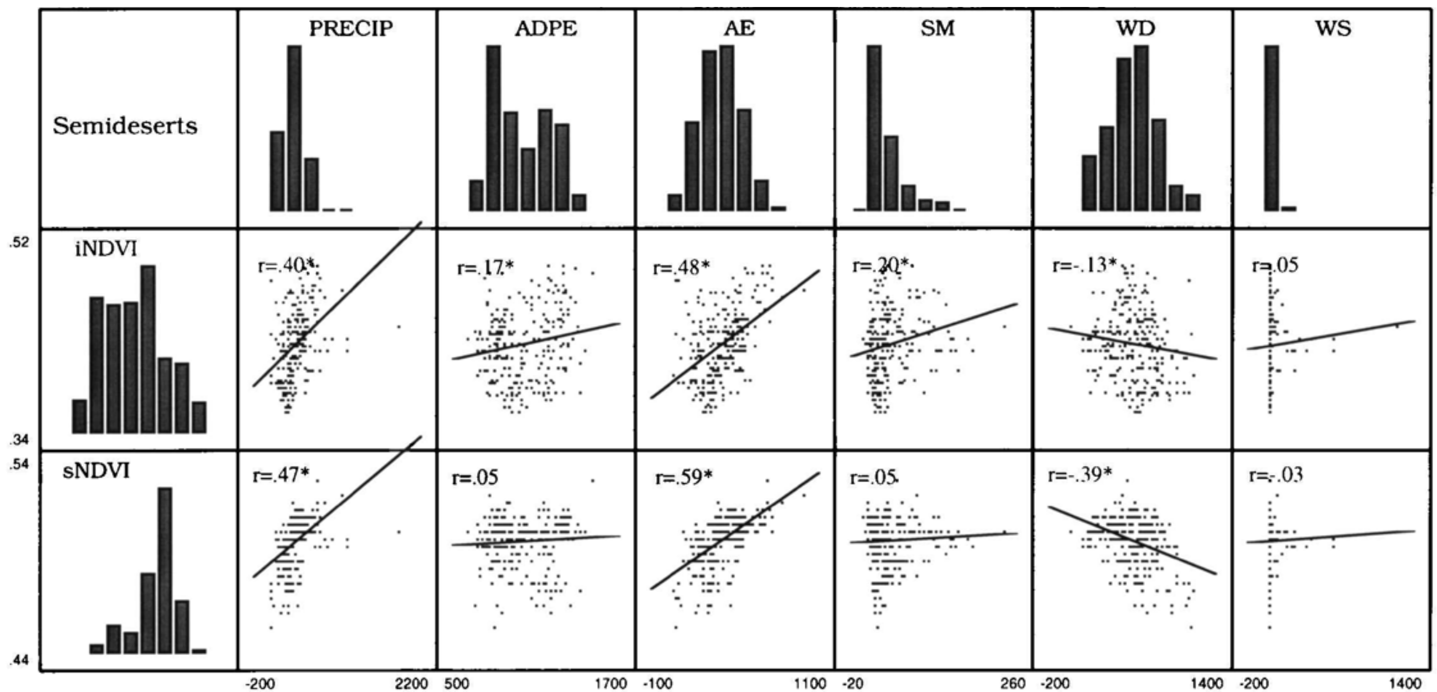


Figure 9. The relationship of iNDVI and sNDVI to annual water balance variables for the 'semidesert ecoregion' in Mexico. Correlation coefficients (r*) are also given indicating significance at the 0·05 level.

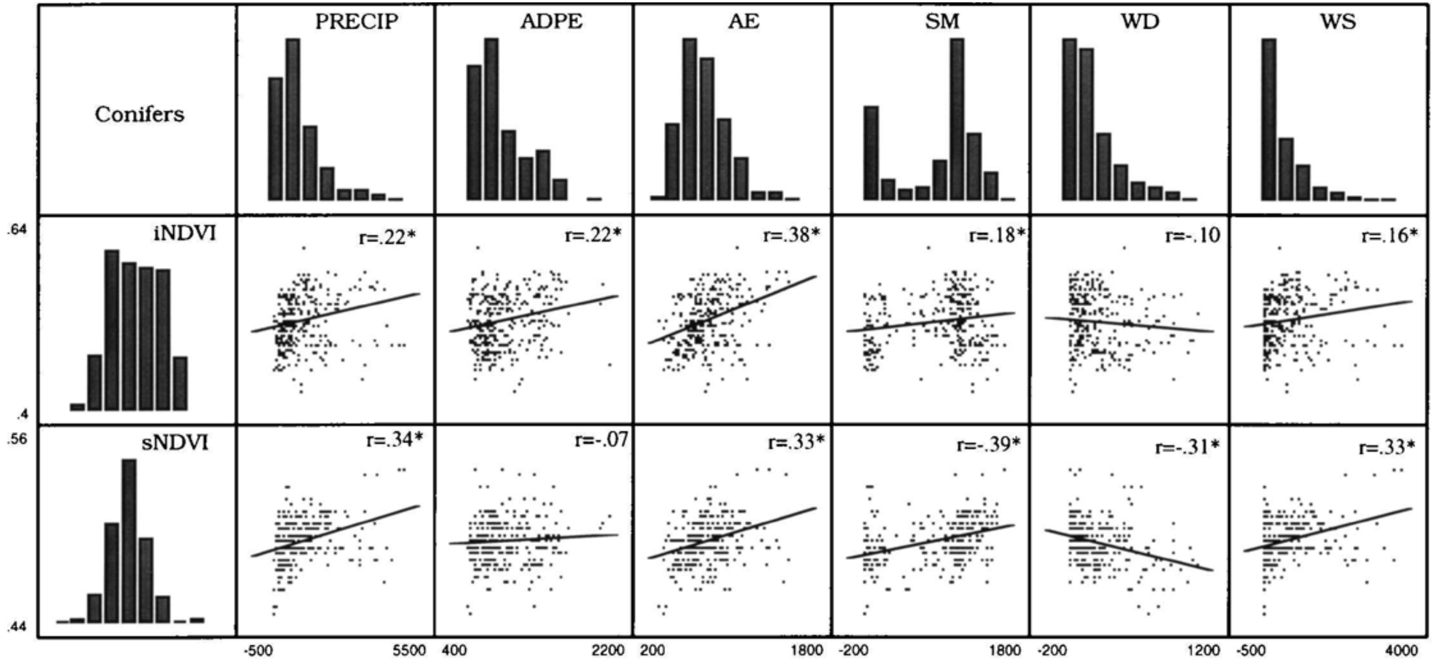


Figure 10. The relationship of iNDVI and sNDVI to annual water balance variables for 'conifer forests' ecoregion in Mexico. Correlation coefficients (r^*) are also given indicating significance at the 0·05 level.

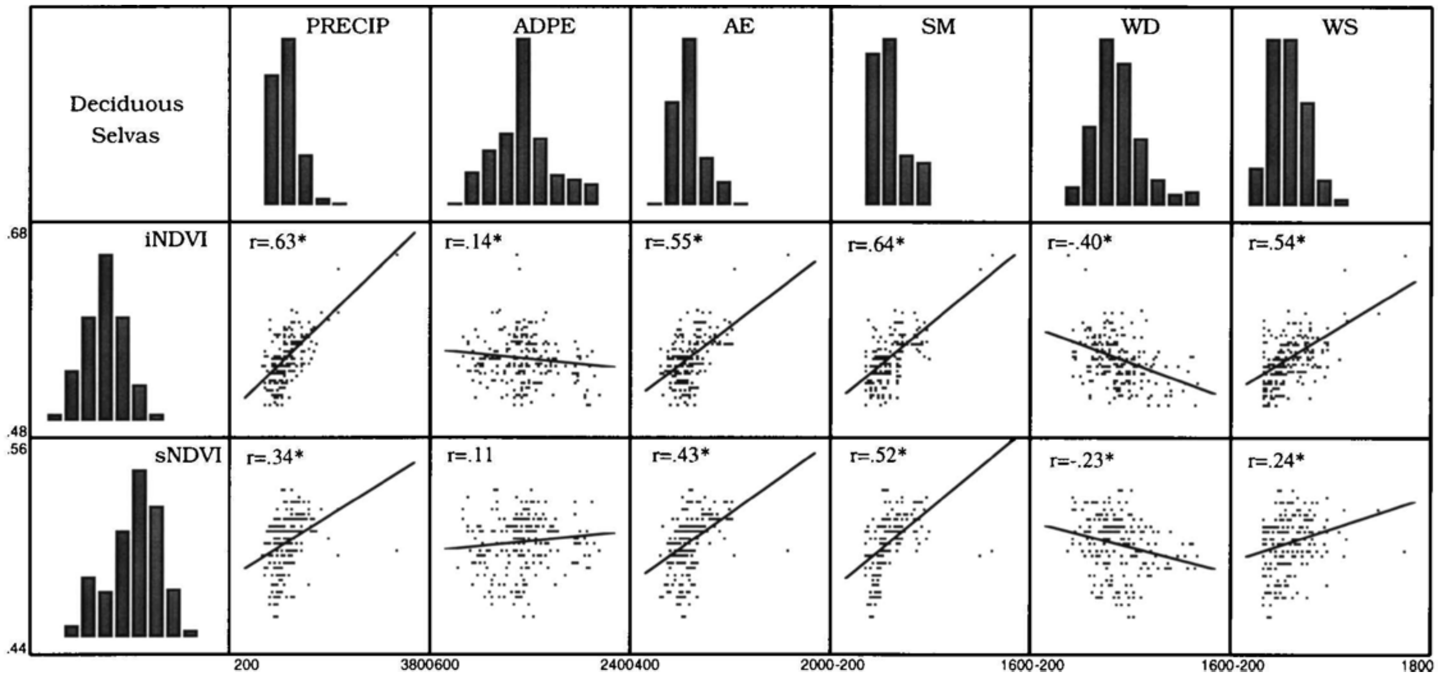


Figure 11. The relationship of iNDVI and sNDVI to annual water balance variables for 'deciduous selvas' ecoregion in Mexico. Correlation coefficients (r^*) are also given indicating significance at the 0·05 level.

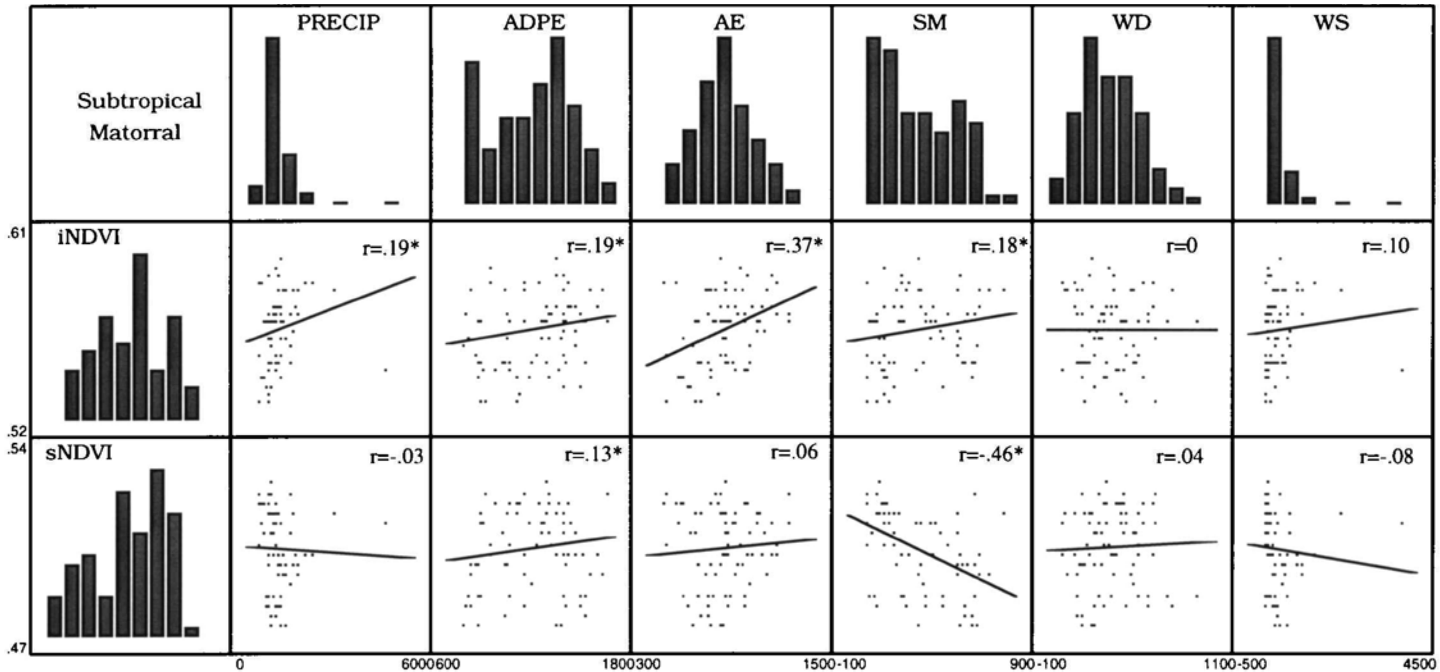


Figure 12. The relationship of iNDVI and sNDVI to annual water balance variables for 'subtropical deciduous selvas' ecoregion in Mexico. Correlation coefficients (r^*) are also given indicating significance at the 0·05 level.

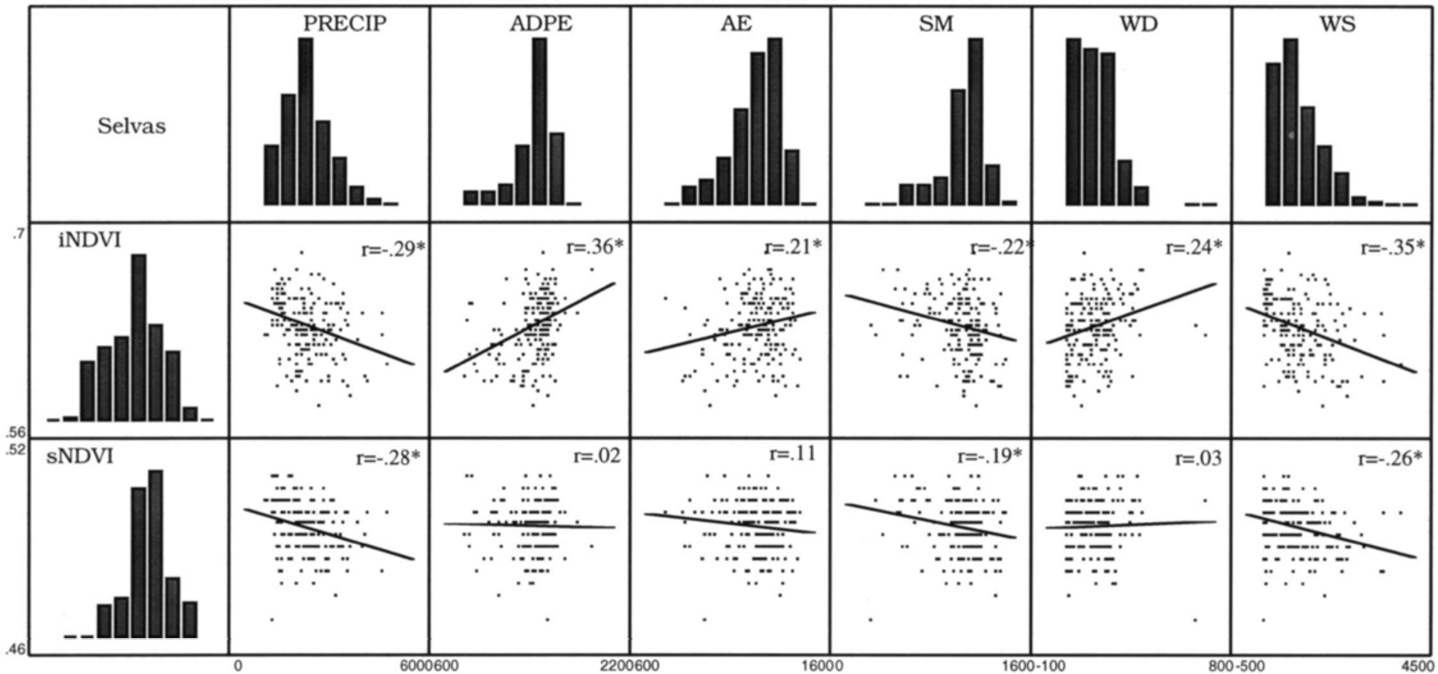


Figure 13. The relationship of iNDVI and sNDVI to annual water balance variables for ‘selvas’ ecoregion in Mexico. Correlation coefficients (r^*) are also given indicating significance at the 0·05 level.

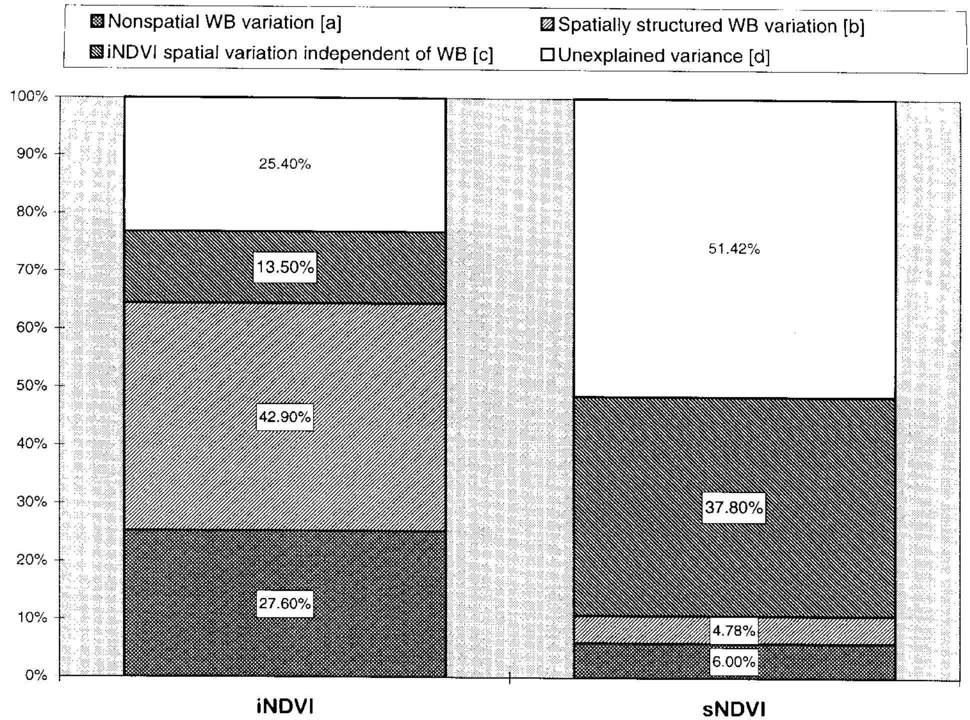


Figure 14. Proportion of variance explained by the sources of spatial variation (fractions) in modelling the iNDVI.

factors such as plant species phenology, soils, topography, and historical disturbance factors.

4. Discussion

Because of Mexico's wide diversity of ecosystems, it is a particularly suitable place to evaluate of the relationships between vegetation patterns and water balance variables. Its location in the neartic and neotropical zones promotes a great diversity of ecological conditions, ranging from deserts to tropical selvas, with a great variety of phenological life-forms and vegetation types. Georeferenced climatic databases and remotely sensed information allows a quantitative evaluation of the functional relationships between climate and vegetation.

The patterns observed by the remotely sensed images can be reasonably reproduced by spatially modelling pixel-NDVI values associated with the weather stations using spatial-autocorrelated functions such as kriging (figures 15 (a) and (b)). This leads us to believe that the (partial and non-partial) regression analysis captured the patterns and relationships among variables considered.

Annual integrated NDVI measures are significantly correlated to each of six annual water balance variables at the country scale in Mexico. Annual actual evapotranspiration and precipitation (figures 6 and 7) were the best variables which predict the variation of iNDVI ($R^2 > 0{\cdot}7$), on a nonlinear basis, when spatial patterns of water balance are not excluded from the models. Overall, actual evapotranspiration is the best predictor of the annual accumulation of 'greenness' in vegetation. This

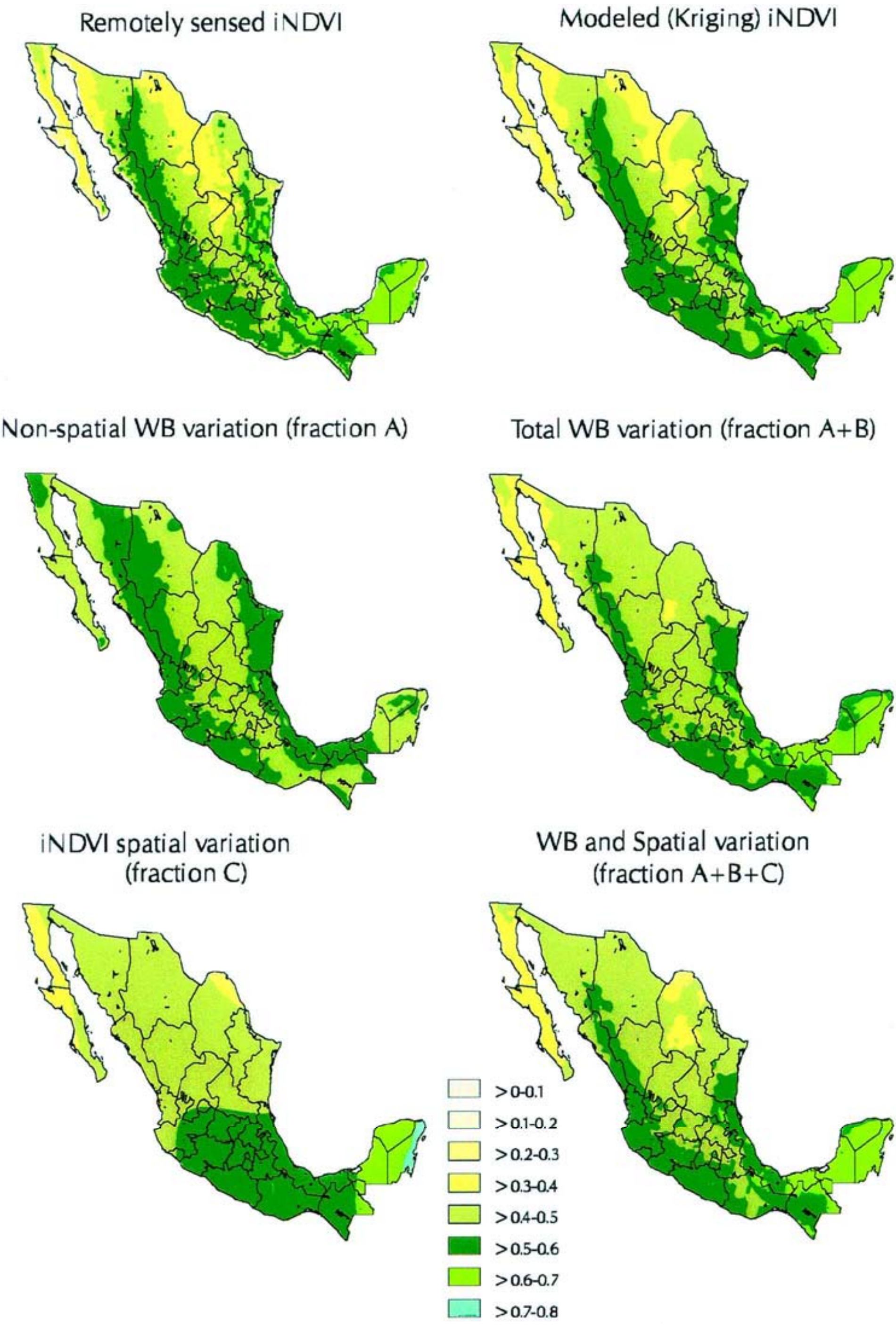


Figure 15. Predicted sources of spatial variation (fractions) in modelling the iNDVI.

pattern is consistent with several previous studies conducted elsewhere in the world (Box *et al.* 1989, Chong *et al.* 1993, Goward *et al.* 1985, 1987).

The relationship between iNDVI and *AE* or *P* can be described as non-linear when the spatial variation is not partialled out. The maximum accumulation of greenness in Mexico during the year (standardized to a 0–1 interval) was in the order of 0·57 iNDVI units using both actual evapotranspiration and precipitation in the models. Box *et al.* (1989) reported a maximum saturation of annual NDVI around 0·4 for the global trend observed in their study. However, they recognized that different values could result from different patterns of spectral response in different regions or cover types. Beta parameters in both models (*AE* and *P*) indicated a rate of greenness accumulation of 0·004 units of iNDVI for each mm of *AE* or *P* annually. These values are also different from the global trend reported for Box *et al.* (beta in Box's model=0·0012), indicating that, in Mexican landscapes, larger rates of greenness accumulation occur than the average over the globe.

The relationship between iNDVI and the water deficit was significantly linear ($p<0{\cdot}001$), with a negative slope. This trend showed that water deficits are a limiting factor in the accumulation of greenness during the year. However, the amount of iNDVI variation that is explained by water deficit is quite low (~33%). This indicates that there are additional factors limiting the accumulation of greenness (e.g., vegetation productivity) at the country scale other than annual water deficits.

The number of weather stations used in this analysis seems adequate to reproduce the patterns observed when analysing the relationship between vegetation indexes and water balance, especially when the reconstruction of the remotely sensed iNDVI was compared with the krigged surfaces that used only the values of 2214 points, representing the weather stations (figures 15 (a) and (b)). We believe that patterns of the different sources of iNDVI variation obtained with partial regression analysis (and mapped with kriging) are meaningful. Partial regression analysis demonstrated that most of the significance in the correlations obtained for iNDVI variation according to water balance is due to spatial structures in both iNDVI and predictor variables. Although the iNDVI variance explained by water balance is only ~28% (after partialling out the spatial component) it is still highly significant. Spatial structures in water balance significantly explained ~43% of the iNDVI variation at country scale, when the total water balance variation is considered. The fact that spatial structures explained more iNDVI variance than water balance variability alone, raises questions about the effects that the spatial patterns of water balance could have over vegetation activity processes. Significant changes in the spatial pattern of water balance (like those occurring as a result of deforestation) could have more dramatic effects in productivity processes than significant changes in water balance variability alone (such as those in global warming). Spatial and water balance variability explained ~70% of iNDVI variation when considered together.

At the country scale, seasonal variations of NDVI (sNDVI) are poorly correlated with annual water balance variables. Even when a seasonal analysis of water balance variables is necessary to explain this lack of correlation among seasonal vegetation indicators and water balance, it seems that the variation in NDVI is greater between ecoregions than the seasonality variation across the entire country. In fact, the variance integrated in the iNDVI measure (captured by PCA when using monthly NDVI) is five orders of magnitude greater than the NDVI seasonal measure (Mora and Iverson 1997). However, it is still necessary to explain why seasonal variations occur.

The variation captured by the seasonal NDVI component seems to be better correlated with water balance for specific vegetation types, but the overall fit is low (usually below 30% of variation accounted for by the best fit variables). At the ecoregion scale, seasonal variation in vegetation measured by NDVI is necessary to differentiate several land cover and vegetation types (i.e., conifer forests from low selvas), but annual variations in climate are not enough to explain seasonal vegetation patterns. Future work will address the effects of seasonal variations of climate on vegetation seasonality as well.

5. Conclusion

From these results, it has been shown that effects of annual water balance variables on vegetation are mainly apparent at the country scale (i.e., where the landscape is represented with a wide spectrum of vegetation types). Annual variations of actual evapotranspiration and precipitation can significantly predict annual integrated measures of NDVI, when spatial structures are considered in the models. The maximum accumulation of greenness, and their rates of accumulation in the Mexican landscape, are above the global trend shown in other studies. If functional attributes of the landscape, related to net primary productivity, are well correlated with annual NDVI measures, their values can also be reasonably well predicted by annual AE and P at this (country) scale.

Annual variations of NDVI for individual ecoregions cannot be predicted so well from water balance variables, even though they were often significantly correlated. The predictive models of measures associated with productivity (iNDVI) and seasonality (sNDVI) generally explained less than 30% of the variance, and other factors (i.e., seasonal climate, soils, disturbances, historical factors, etc.) should be considered to explain more of the residual variance not accounted for by annual climate variation.

Acknowledgment

The authors thank Dr James W. Merchant for a comprehensive revision of the manuscript. We also acknowledge Dr Sandra Brown for reviewing an earlier version. Special thanks to the Illinois Natural History Survey for providing the equipment and facilities to conduct this research. Final preparation of this manuscript was completed at the Center for Advanced Land Management Information Technologies (CALMIT), University of Nebraska, Lincoln. This research was conducted while the first author held a Fulbright-CONACYT-IIE scholarship for his MS degree. This research was partially funded by the Tinker foundation of the University of Illinois at Urbana-Champaign.

References

Bocard, D., 1992, Partialling out the spatial component of ecological variation. *Ecology*, **73,** 1045–1056.

Box, E. O., Holben, B. N., and Kalb, V., 1989, Accuracy of the AVHRR vegetation index as predictor of biomass, primary productivity and net CO_2 flux. *Vegetatio*, **80,** 71–89.

Chong, D. L., Mougin, E., and Gastellu-Etchegorry, J. P., 1993, Relating the global vegetation index to net primary productivity and actual evapotranspiration over Africa. *International Journal of Remote Sensing*, **14,** 1517–1546.

Choudhury, B. J., 1987, Relationships between vegetation indices, radiation absorption and net photosynthesis evaluated by sensitivity analysis. *Remote Sensing of Environment*, **22,** 209–234.

Davenport, M. L., and Nicholson, S. E., 1993, On the relation between rainfall and the normalized difference vegetation index for diverse vegetation types in East Africa. *International Journal of Remote Sensing*, **14,** 2369–2389.

Di, L., Rundquist, D. C., and Han, L., 1994, Modelling relationships between NDVI and precipitation during vegetative growth cycles. *International Journal of Remote Sensing*, **15,** 2121–2136.

Eagleman, J. R., 1976, *The Visualization of Climate* (Lexington, MA: Lexington Books).

ESRI, 1995, ARC/INFO User's Manual. Environmental Systems Research Institute, Inc., Redlands, California.

Farrar, T. J., Nicholson, S. E., and Lare, A. R., 1994, The influence of soil type on the relationships between NDVI, rainfall and soil moisture in semiarid Botswana. II. NDVI response to soil moisture. *Remote Sensing of Environment*, **50,** 121–133.

Fung, I. Y., Tucker, C. J., and Prentice, K. C.,, 1987, Application of advanced very high resolution radiometer vegetation index to study atmosphere-biosphere exchange of CO_2. *Journal of Geophysical Research*, **92,** 2999–3015.

Gallo, K. P., 1989, The use of satellite-derived vegetation indices as indicators of climatic variability. *Proceedings of the 6th Conference on Applied Climatology*, 7–10 March (Charleston, SC: American Meteorological Society).

Gallo, K. P., 1990, Satellite derived vegetation indices: a new climatic variable? *Proceedings of the Symposium on Global Change Systems, Special Sessions on Climate Variations and Hydrology*, 5–9 February (Anaheim: American Meteorological Society).

Goward, S. N., 1989, Satellite bioclimatology. *Journal of Climate*, **2,** 710–720.

Goward, S. N., and Dye, D. G., 1987, Evaluating North American net primary productivity with satellite observations. *Advances in Space Research*, **7,** 165–174.

Goward, S. N., Tucker, C. J., and Dye, D., 1985, North American vegetation patterns observed with NOAA-7 advanced very high resolution radiometer. *Vegetatio*, **64,** 3–14.

Goward, S. N., Kerber, A., Dye, D. G., and Kalb, V., 1987, Comparison of North American and South American biomes from AVHRR observations. *Geocarto International*, **2,** 27–40.

Inegi, 1980, Instituto Nacional de Estadistica Geografia e Informatica. Carta de climas, escala 1:1000 000. Direccion General de Geografia del territorio nacional. Secretaria de Programacion y Presupuesto, Mexico.

Inventario Nacional Forestal, 1964, *Primer Inventario Nacional Forestal*. Direccion General del INF, Subsecretaria Forestal, Secretaria de Agricultura, Mexico.

Inventario Nacional Forestal, 1991, *Inventario Nacional Forestal de Gran Vision*. Direccion General del INF, Subsecretaria Forestal, Secretaria de Agricultura, Mexico.

Kustas, W. P., Perry, E. M., and Moran, M. S., 1994, Using satellite remote sensing to extrapolate evapotranspiration estimates in time and space over a semiarid rangeland basin. *Remote Sensing of Environment*, **49,** 275–286.

Legendre, P., 1993, Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74,** 1659–1674.

Maisongrande, P., Ruimy, A., and Saugier, B., 1995, Monitoring seasonal and interannual variations of gross primary productivity, net primary productivity and net ecosystem productivity using a diagnostic model and remotely-sensed data. *Tellus*, **47B,** 178–190.

Malo, A., and Nicholson, S. E., 1990, A study of rainfall and vegetation dynamics in the African Sahel using normalized difference vegetation index. *Journal of Arid Environments*, **19,** 1–24.

Mora, F., and Iverson, L. R., 1997, Dynamic stratification of the landscape of Mexico: Analysis of vegetation patterns observed with multitemporal remotely sensed images. *Geocarto International*, **12,** 1–15.

Myers, N., 1993, Tropical forests: the main deforestation fronts. *Environmental Conservation*, **20,** 9–16.

Nicholson, S. E., and Farrar, T. J., 1994, The influence of soil type on the relationships between NDVI, rainfall, and soil moisture in semiarid Botswana. I. NDVI response to rainfall. *Remote Sensing of Environment*, **50,** 107–120.

Pinker, R. T., 1990, Satellites and our understanding of the surface energy balance. *Paleogeography, Paleoclimatology, Paleoecology*, **82,** 321–342.

Rosenberg, N. J., Blad, B. L., and Verma, S. B., 1983, *Microclimate: The Biological Environment* (New York: John Wiley and Sons).

Running, S. W., 1986, Global primary production from terrestrial vegetation: estimates integrating satellite remote sensing and computer simulation technology. *Science of the Total Environment*, **56,** 233–242.

RUNNING, S. W., and NEMANI, R. R., 1988, Relating seasonal patterns of the AVHRR vegetation index to stimulated photosynthesis and transpiration of forests in different climates. *Remote Sensing of Environment*, **24,** 347–367.

RUNNING, S. W., NEMANI, R. R., PETERSON, D. L., BAND, L. E., POTTS, D. F., PIERCE, L. L., and SPANNER, M. A., 1989, Mapping regional forests evapotranspiration and photosynthesis by coupling satellite data with ecosystem simulation. *Ecology*, **70,** 1090–1101.

SANDHOLT, I., and ANDERSEN, H. S., 1993, Derivation of actual evapotranspiration in the Senegalese Sahel, using NOAA-AVHRR data during the 1987 growing season. *Remote Sensing of Environment*, **46,** 164–172.

SEEVERS, P. M., and OTTMANN, R. W., 1994, Evapotranspiration estimation using a normalized difference vegetation index transformation of satellite data. *Hydrological Sciences*, **39,** 333–345.

STEPHENSON, N. L., 1990, Climatic control of vegetation distribution: The role of water balance. *The American Naturalist*, **135,** 649–670.

THORNTHWAITE, C. W., 1948, An approach towards a rational classification of climate. *Geographical Review*, **38,** 55–94.

THORNTHWAITE, C. W., and MATHER, J. R., 1955, The water balance. *Publications in Climatology*, **8,** 1–104.

THORNTHWAITE, C. W., and MATHER, J. R., 1957, Instructions and tables for computing potential evapotranspiration and the water balance. *Publications in Climatology*, **10,** 185–311.

TUCKER, C. J., and SELLERS, P. J., 1986, Satellite remote sensing of primary production. *International Journal of Remote Sensing*, **7,** 1395–1416.

TUCKER, C. J., FUNG, I. Y., KEELING, C. D., and GAMMON, R. H., 1986, Relationships between atmospheric CO_2 variations and a satellite-derived vegetation index. *Nature*, **319,** 195–199.

WILLMOTT, C. J., 1977, WATBUG: a Fortran IV algorithm for calculating the climatic water budget. *Publications in Climatology*, **30,** 1–52.