International Water Resources Association
Water International, Volume 30, Number 3, Pages 270–282, September 2005

Influence of Hydroclimatology and Socioeconomic Conditions on Water-Related International Relations

Kerstin Stahl, *University of British Colombia, Canada*

***Abstract**: Climatic and environmental changes and a rising water demand have increased the competition over water resources and have made cooperation between countries that share a transboundary river an important issue in water resources management and hydropolitics. Yet in river basins around the world, international conflict and cooperation are influenced by different factors, and general conclusions about forces driving conflict and cooperation have been difficult to draw. Using global data, this study investigates how combinations of hydroclimatic, socioeconomic and political conditions influence the pattern of water-related international relations (WIR). The data complexity required several steps of statistical analysis. First, patterns of WIR were defined by subjecting the basins' aggregate history of political event data to cluster analysis. Using the classified WIR-patterns as a dependent variable, the study then explores the influence of combinations of explanatory variables on international relations by the means of a multivariate classification tree model. The obtained model, which determines a basin-country-share's most likely WIR-class by its hydroclimatic, socioeconomic, and political characteristics, correctly classifies two-thirds of the basin shares. The tree rules suggest that hydroclimatic variability and population density are most influential in arid to sub-humid basins, while socioeconomic and political factors seem to be more important in determining WIR in humid basins. The application of the model to all international basins worldwide illustrates its potential for the assessment of the global water-related international relations processes.*

***Keywords:** transboundary rivers, conflict, cooperation, international relations, hydroclimatic variability, cluster analysis, classification tree*

Introduction

Rivers are an essential natural resource closely linked to a country's well being and economic success. But rivers ignore political boundaries, and competition over the water resources has lead to political tension between countries with transboundary rivers. Integrating international cooperation and conflict resolution into the water management of transboundary rivers has therefore become an important issue in water resources management and hydropolitics. The problem requires a good understanding of the history and patterns of conflict and cooperation among nations sharing international basins worldwide and of the different factors that have influenced their international relations. However, most studies have dealt only with specific rivers. Only in recent years have researchers attempted to generalize and extrapolate by collecting data and analyzing the issue on regional and global scales (Wolf et al., 2003a; Dinar, 2004).

Transboundary rivers have been studied in several disciplines. The environmental security literature has identified transboundary rivers and freshwater scarcity as being capable of invoking enough social tension to cause international conflicts (Homer-Dixon, 1994). Studying a large number of shared rivers, Toset et al. (2000) finds that a joint river increases the probability of militarized dispute over and above the tension inherent in the mere contiguity of two countries. Rather than dealing with shared rivers' contribution to interstate conflict in general, the water resources literature is more concerned with conflict and cooperation over issues that directly relate to transboundary rivers and freshwater as a resource. Gleick (1993) elaborates possible causes of conflict and proposed four quantitative indices of water resources vulnerability that may be used in general to identify "regions at risk." The indices are the ratio of water demand to supply, per capita water availability, the fraction of water supply originating outside of the country, and dependence on hydropower. Ashton (2002) discusses how water scarcity and the unequal geographic distribution of freshwater contribute to conflict potential in Africa. While most analytical studies focus on conflict, literature on cooperation is mainly concerned with management frameworks (e.g. Savenije and van der Zaag, 2000; Kliot et al., 2001) and practical issues of water allocation (e.g. Fisher et al., 2002; van der Zaag and Vaz, 2003).

Fewer studies have analyzed the whole spectrum of international relations from conflict to cooperation at a comparative or even global level. There are 263 international river basins worldwide that cover about 45 percent of the land surface and are home to about 40 percent of the world's population (Wolf et al., 1999). The geographic dataset of the world's international river basins (Wolf et al., 1999) and the dataset of water-related political events of conflict and cooperation (Yoffe et al., 2004) of the Transboundary Freshwater Dispute Database (TFDD) at Oregon State University provide a useful basis for truly global studies of water-related conflict and cooperation. They allowed the first quantitative global-scale exploration of the history and geography of international relations over transboundary rivers and revealed an overwhelming amount of cooperation (Wolf et al., 2003a). With the Basins-at-Risk study, Yoffe et al. (2003) empirically tested hypotheses of prevailing wisdom about the influence of a number of individual indicator variables on water-related conflict and cooperation. Though a few indicators were found to be significant (e.g. high population density, low per capita GDP, and generally hostile relations indicate increased conflict), no single indicator explained a large percentage of the variability in the data. Some results were contradicting. Testing climate as an indicator, for example, showed that the average historic level of conflict was lower in humid meso-thermal regions than in other climate zones; annual precipitation anomalies for a subset of eleven basins, however, showed no consistent relation to the year's conflict level (Yoffe et al., 2003). Using another statistical approach and an enhanced global hydroclimatologic dataset showed that water-related international relations in regions with a higher hydrological variability and more extreme hydroclimatic conditions (aridity) have also been more extreme (Stahl and Wolf, 2003). These findings demonstrate that the relationship between hydroclimatic factors and events along the conflict-cooperation scale of international relations is not linear and strongly depends on multiple conditions.

With the working hypothesis that hydroclimatic, socioeconomic, and political conditions influence the patterns of water-related international relations over transboundary rivers, this study expands on the previous work of Yoffe, Wolf, and their collaborators. While Yoffe et al. (2003) investigated the influence of individual indicators on a temporally averaged conflict-cooperation measure, we now move on to study how combinations of indicators have influenced the basins' historic variability of conflict and cooperation. The objectives of this study are: (1) to incorporate the variability into the parameterization of the international basins' water-related international relations history; (2) to shed light on the conjunct factors influencing conflict and cooperation worldwide, (3) to assess the potential of classification tree modeling for analysis of and application to this global multi-disciplinary dataset.

A brief introduction is provided on the available global databases: the political events, the geography, socioeconomic and political data, and the hydroclimatology. Here, we review the hypotheses for the individual indicators' (explanatory variables) influence on water-related international relations. The hydroclimatic variables are discussed in more detail, as the dataset is new and has not previously been published. The next section deals with the identification of patterns of international relations (Objective 1). We propose a new approach to classify global water-related international relations (WIR) from the history of political event data using cluster analysis. The following section then explores the combinations of explanatory variables that influence WIR (Objective 2) using a Classification Tree Model. Finally, we apply the model globally and discuss the geographical distribution of the WIR history and modeling. This application allows the discussion assessment of the potential of the modeling approach for applications of decision and policy making (Objective 3).

Database of the World's International River Basins

Events of Water-Related Conflict and Cooperation

The Transboundary Freshwater Dispute Database (TFDD) contains a political event dataset, which was created from multiple existing political and news databases, covers the time period 1948 to 1999, and concerns all international river basins worldwide. Events are defined as instances of conflict or cooperation that occurred within an international river basin, that involve the nations riparian to the river, and that concern freshwater as a scarce or consumable resource (e.g. water quality, water quantity) or as a quantity to be managed (e.g., flooding or flood control, managing water levels for navigational purposes) (Yoffe et al., 2004). Following the Conflict and Peace Databank (COPDAB) International Cooperation and Conflict Scale, the events were categorized allowing to group them by their nature and intensity of conflict or cooperation and to deal with them as a class (Azar, 1980). For the Basins-at-Risk (BAR) study, the categories were adapted to water-related events and the intensity scale was recentered to range from a level of -7 to +7 to create the so-called BAR-scale (for details of the procedure, please refer to Yoffe & Larson [2002]). The most conflictive and most cooperative events that are recorded in the existing database have a BAR-scale of –6 and +6, respectively. For the present study the 15 BAR-scale categories were aggregated to five conflict-cooperation levels (CCL) ranging from "most conflictive" over "conflictive" and "neutral" to "cooperative" and "most cooperative" events (Table 1). The aggregated classes of each three BAR scales distinguish violent conflict, political conflict, neutral verbal interactions, moderate cooperation, and active cooperation or treaty signature.

Table 1. The Conflict-Cooperation Levels for water-related political events

BAR Scale	*Description of nature of the water-related political event between riparian countries**	*Conflict-Cooperation Level (CCL)*
6	International Freshwater Treaty; Major strategic alliance	most cooperative
5	Military economic or strategic support	
4	Non-military economic, technological or industrial agreement	
3	Cultural or scientific agreement or support (non-strategic)	cooperative
2	Official verbal support of goals, values, or regime	
1	Minor official exchanges, talks or policy expressions—mild verbal support	
0	Neutral or non-significant acts for the inter-nation situation	neutral
-1	Mild verbal expressions displaying discord in interaction	
-2	Strong verbal expressions displaying hostility in interaction	
-3	Diplomatic-economic hostile actions	conflictive
-4	Political-military hostile actions	
-5	Small scale military acts	most conflictive
-6	Extensive War Acts causing deaths, dislocation or high strategic costs	

*****COPDAP scale by Azar (1980) adapted to water events by Yoffe and Larson (2002)

The event database contains approximately 1,800 events. Many basins only have one or a few events. Statistics on the events-database and the conflict-cooperation levels are summarized in Yoffe and Larson (2002) and Wolf et al. (2003a).

Geography, Socioeconomic, and Political Conditions

On the basis of the Inventory of International River Basins of the World, Fiske and Yoffe (2001) set up a Geographical Information System (GIS) of the historic and 263 current international river basins worldwide. Within the Transboundary Freshwater Dispute Database, the GIS is linked with the international freshwater treaties dataset (UNEP and OSU, 2002) and the aforementioned dataset of reported political events of water-related political conflict and cooperation. The combination of these datasets provides the most comprehensive database available in this field based on the geographic unit relevant for water resources issues: the international river basin.

The present study considers the geographic unit of the basin-country-polygon (BCP), which is defined as a country's share of an international river basin. Besides the geographical location, several biophysical, geopolitical, and socioeconomic indicators are available at the basin and country level (Yoffe et al., 2004). Table 2 summarizes the key basin characteristics and socioeconomic and political conditions used as explanatory variables in this study. Their hypothesized influences on water-related conflict and cooperation are as follows:

The percentage of a country's areal share (*Area*) of the international basin indicates the relative interests of countries in the basin. A country's relative location along the river (*Loc*, *B*) indicates the relative power and prior access to the resource. Upstream-downstream relationships and a river as a border are widely considered to increase the risk for conflict (Ashton, 2002; Toset, et al., 2000). Population Density (*Pop*), which was obtained from the LandScan2000 coverage (Dobsen, et al., 2000), is a common social indicator variable and can act as a surrogate for water demand spatial data, which is difficult to obtain. The importance of population density is underlined by scenario simulations that suggest that population growth outweighs climate change as a factor of increasing water stress (Vörösmarty et al., 2000). Water stress is assumed to increase the likelihood of conflict.

Economic and political variables are commonly used indicators for a region's economic and institutional capacity to deal with environmental stress. Such numbers are only available at the country level. The World Bank (2003) regularly publishes Per Capita Gross Domestic Product (*GDP*) per country. For this study, their annual data was averaged for the period available for each country (1960 to 2000). Cooperation is commonly hypothesized to be better among economically well-off countries that can afford technological solutions. To describe the political regime characteristics, the 1950 to 2000 democracy/autocracy indices (*DA*) of the PolityIV Project (2000) were averaged for each country. The scale ranges from –10 (most autocratic) to +10 (most democratic). Though it is often hypothesized that democratic countries cooperate better, Yoffe et al. (2003), concluded from their analysis as well as the literature that such a relation remains to be proven.

Table 2. Basin characteristics and socio-economic and political variables for basin-country-polygons

Variable	*Description*	*Data source*	*Available for number of BCPs*
P	Percentage of the basin area within the country	BAR/TFDD	654
Loc	Indicator of country being upstream, middle or downstream	BAR	656
B	River is the border for a substantial part of the country	BAR	656
Pop	Population Density	LandScan 2000	644
GDP	Per Capita Gross Domestic Product in 1995-US$ per year and person	World Bank (2003)	654
DA	Democracy/Autocracy Level	PolityIV (2000)	644

Hydroclimatology

Different variables can describe hydroclimatic conditions that may indicate resource scarcity and hence influence a region's risk for conflict or potential for cooperation over freshwater. Water availability is often combined with a social factor into supply-demand ratio (Gleick, 1993) or a water stress index (Falkenmark, 1989). The use of these indices as explanatory variables in statistical analyses implies the assumption that the variability of the individual factors in time and space is irrelevant. However, the influence of spatial and temporal variability has recently been highlighted. Ashton (2002) presents a map linking conflict to geographical regions with perennial rivers in Africa suggesting that conflicts occur where the high seasonality and inter-annual variability of water availability make an adequate preparation for dry spells difficult. In fact, many treaties between riparian countries of transboundary rivers do not include rules for extreme hydrological conditions such as floods and droughts, and although some do, there is still a risk that agreements were made during a wet climatic period. The lack of flexibility to account for changed conditions may then lead to tensions or dispute. Mexico, for example, has not been able to deliver the agreed discharge in the Rio Grande after initially taking the drought year escape clause in the agreement with the US (Kelly and Chapman, 2002). Hence, good knowledge of the hydroclimatic variability is crucial for the management and cooperation in international basins.

In the framework of this study, the existing spatial dataset was therefore complemented by detailed hydroclimatic information. Basically, two different types of data can be considered for use in global studies: original at-station data or grid-based averages often provided by climate research centers, which are produced by appropriate and well-tested interpolation strategies. For this study we decided to use the Climate Research Unit's (CRU) 0.5-degree monthly mean precipitation (New et al., 2000). For the required period, it is the global dataset with the highest spatial and temporal resolution available. From this dataset it is possible to derive time-series as well as monthly mean or annual mean precipitation for the spatial units of the basin, country or basin-country-polygon (BCP). Another global 0.5-degree grid used to calculate a time-averaged parameter in this study is the mean annual potential evapotranspiration by Ahn and Tateishi (1994), which is available to the public from UNEP GRID (2003). Figure 1 shows the world's international basins and derivation of precipitation information for the Senegal River basin illustrating the different scales and resolutions of the geographical data: the grid resolution for rainfall, the political boundaries of countries, and the units relevant for transboundary water resources management, the international basin, and the basin-country polygon (BCP). The example also illustrates the variability within the units.

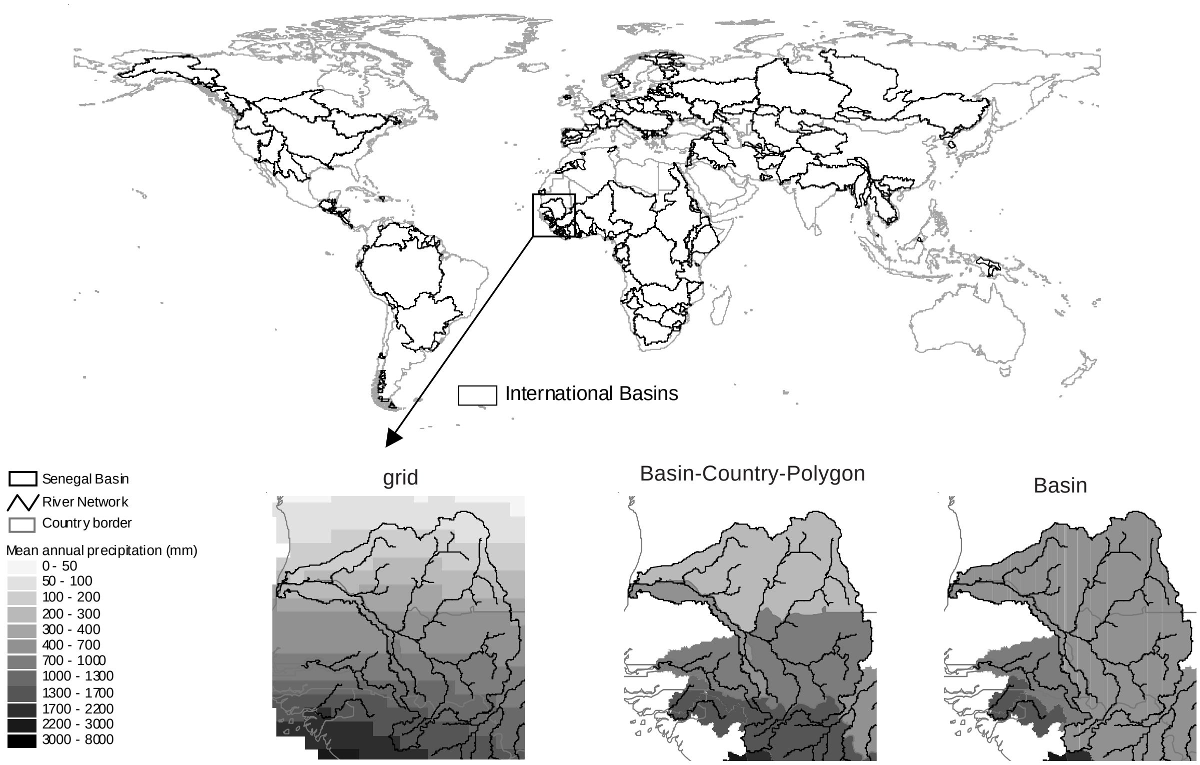

Figure 1. Data. Upper panel: GIS of international river basins (source: TFDD, 2003). Lower panel: example of the aggregation levels of precipitation in the Senegal Basin (Source: New et al., 2000)

Table 3. Hydroclimatic variables for the BCPs

Variable	*Description*	*Data Source*	*Available for # of BCPs*
A	Index of Aridity (mean annual precipitation/potential evapotranspiration)	Ahn-Tateishi (1994)	664
P	Mean annual precipitation	CRU	664
CVPs	Spatial variability: coefficient of variation of mean annual precipitation	CRU	664
CVPt	Inter-annual variability: coefficient of variation of annual precipitation 1948 to 1998	CRU	664
SP	Intra-annual variability: seasonality index of mean monthly precipitation	CRU	664

Table 3 summarizes the derived variables. The Index of Aridity (*A*) is a measure for the general climatic water balance and is often used to classify arid, semi-arid and humid regions (UNESCO, 1997). Total precipitation, but particularly its spatial variability (*CVPs*) and temporal variability (*CVPt*) indicate the general water availability and its distribution. Inter-annual variability is an important indicator of the reliability of water availability. Seasonal or intra-annual variability of precipitation is an important aspect of hydroclimatology because it determines the seasonality of other hydrologic quantities such as stream flow and groundwater recharge. It defines the possibilities of rain-fed agriculture, and if it is high, it may indicate a dependency of a country's water supply on the international river. Therefore a Seasonality Index (*SP*) was derived from circular statistics of monthly mean values according to Markham (1970). It is a measure of the concentration of the precipitation in the course of the year. An index of *SP* = 0 means that every month has the same amount of precipitation, and an index of *SP* = 1 means that all the annual precipitation is concentrated within one month of the year. Except for *CVPs*, which describe the variability of the individual grid cells that compose a BCP, BCP averages were derived from the original values of the grid cells that make up a BCP before calculating the parameters.

Water-Related International Relations

Defining Patterns of Event History

International relations scholarship is often categorized by its analytic purpose as relating to metaphor, history, theory, engineering, and pattern recognition (Chan, 2002). The event dataset of the Transboundary Freshwater Dispute Database, with its coded historic political events, provides a variety of opportunities to analyze water-related conflict-cooperation behavior globally (Yoffe et al., 2004). The sample of coded events can be used to describe a basin or country's history of conflict and cooperation over a river in quantitative terms. Used to assess the risk for conflict or potential for cooperation as shown by Wolf et al. (2003a), these may be important information for international policy decisions.

For a global study, a variable has to be found that provides an aggregate description per geographical unit. Several possibilities exist: Wolf et al. (2003a) described Basins-At-Risk by the average of the annual averages of the BAR-scale of a basin's history of events. However, as most conflicts are resolved over time (e.g. Postel and Wolf, 2001), averaging all historic events results in a reduction of variability and shift of the scale towards neutral to slightly positive average values. Another possibility is to look at events of a certain type, e.g. certain BAR-scales, only. Espey and Towfique (2004) investigated the influence of geographic, socio-cultural, and economic factors on bilateral treaty formation (equal to BAR-scale 8 events). Stahl and Wolf (2003) specifically explored the relationship between hydroclimatology and the frequency of extreme (conflictive and cooperative) political events. Both parameterizations have drawbacks. While the average of all historic events on the conflict-cooperation scale conceals a possible relation to events of specific intensity, analyzing only specific events such as the extremes ignores moderating information from the entire history. Figure 2 illustrates two different histories by showing the frequency of events at the aggregated conflict-cooperation levels (CCL) for Hungary over the Danube River and for Jordan over the Jordan River. The historic average of the original BAR-scale for both is very similar with 0.4 (133 events) and 0.2 (160 events), respectively. Although both are dominated by a maximum of neutral events, the distribution of the other events along the CCL scale is different. For instance, Hungary had no "most conflictive" (i.e. violent) events over the Danube; Jordan had a considerable number of violent conflicts over the Jordan River. With this additional knowledge, one might place the Jordan River Basin at a higher "risk" for violent conflict than might be suggested by the historic average. On the other hand, the mere look at the conflicts would underestimate the potential for cooperation.

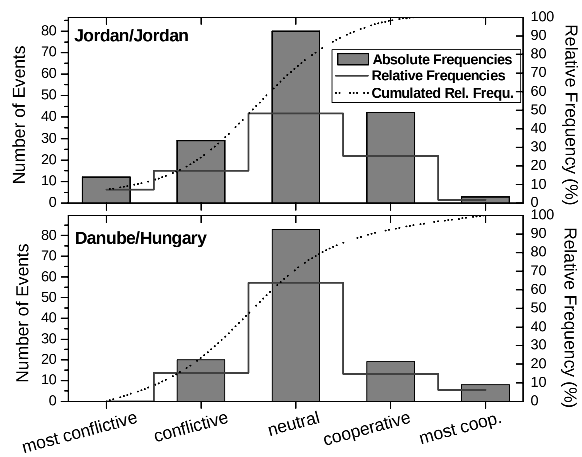

Figure 2. Distribution of water-related political events on the conflict-cooperation scale for two selected Basin-Country-Polygons (Jordan River in Jordan and Danube River in Hungary)

Classification of Event Histories by Cluster Analysis

To include as much information on conflict and cooperation as possible, the cumulative relative frequency distribution of events along the CCL scale for each BCP was used to classify different patterns of water-related international relations (WIR) history. For this purpose, the cumulative distributions of events in the world's basin-country-polygons were subjected to a cluster analysis. The purpose of a cluster analysis is to group objects (here: BCPs) in a way that a group's characteristics (here: aggregate conflict-cooperation history) are as homogeneous as possible, but the characteristics of the objects between the groups are as contrasting as possible. The common algorithms are described in current statistics textbooks such as in Hastie et al. (2001). First, a measure of similarity or dissimilarity between the characteristics of all pairs of objects is calculated, and then a clustering algorithm can be applied to search this dissimilarity matrix for homogeneous groups of objects. The most common distance measure, the Euclidean Distance, was used in the present study to define the pairwise dissimilarity of all BCPs. The dissimilarity *D* between BCP *x* and BCP *y*'s attributes *f* (the cumulative relative frequency along the five CCL levels) is calculated as:

$$D(x,y) = \sqrt{\sum_{i=1}^{5}\left(f_{x,i} - f_{y,i}\right)^2} \qquad (1)$$

The k-means algorithm classifies the observations as belonging to one of k groups. Group membership is determined iteratively by calculating the centroid for each group (the multidimensional version of the mean) and assigning each observation to the group with the closest centroid. Centroids are calculated using least-squares, and observations are assigned to the closest centroid based on least-squares.

Cluster analyses were carried out with 135 BCPs that have a record of five and more historical political events for a range of *k=3 to 12* clusters (solutions). While there is no fixed rule on the right number of clusters, some statistics can guide the decision. Figure 3 shows the total sum of squares for the 3-cluster to 11-cluster solution (left panel). The curve indicates with a kink that the overall

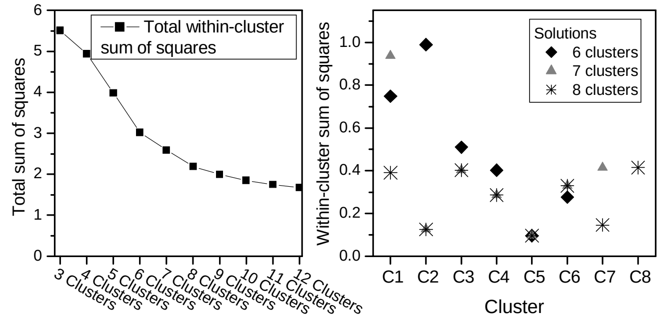

Figure 3. Sum of mean squares of several cluster solutions: total for all clusters (left panel), and individual clusters (right panel)

gain of homogeneity becomes smaller after more than six clusters and then again after eight clusters. The right panel of Figure 3, which shows within-cluster sum of squares for all individual clusters of the 6-, 7- and 8-cluster solutions, reveals very heterogeneous groups in the 6- and 7- cluster solutions. Therefore, the 8-cluster solution was chosen.

The mean sum of squares of the eight final clusters ranges from 0.1 to 0.42. Cluster C2, C5, and C7 are most homogeneous. These clusters have relatively few members. The distributions for the eight clusters can be well distinguished by the location of the maximum frequency and minimum frequency of events along the scale of conflict and cooperation. Table 4 summarizes the characteristic pattern of event history, the number of members, and the average number of events of the eight clusters. Figure 4 illustrates the eight different patterns of WIR by the mean relative frequency of events for each CCL. The maximum and minimum at each CCL adds information about the homogeneity of the cluster. Their geography will be discussed later.

Exploring Influences on Water-related International Relations

Method

After having identified them, the question is whether and how the eight classes (clusters) of WIR can be statistically explained by the combination of the hydroclimatic variables and the socioeconomic and political conditions

Table 4. Cluster characteristics

Cluster/ Class	*Description of the WIR*	*Number of BCPs*	*Average number of events per BCP*
C1	a peak of cooperative events (50%) closely followed by the percentage of neutral events (40%), some most cooperative, few conflictive events.	33	24
C2	peaks strongly with > 70% of neutral events, the rest of the events almost all conflictive	7	61
C3	an almost equal majority of most cooperative (40%) and cooperative (35%), some neutral and conflictive events	17	13
C4	flat distribution with an almost equal percentage of conflictive, neutral and cooperative events; a few events on both extremes	12	54
C5	distinct peak of most cooperative events (>70%), 15-20% of cooperative and conflictive events	5	6
C6	close to 90% of cooperative events and a few neutral ones	21	10
C7	major peak of cooperative events (> 55%); 15-20% of conflictive, neutral, and most cooperative events,	10	15
C8	peak close to 70% of neutral events, rest cooperative with a few extreme events	30	29

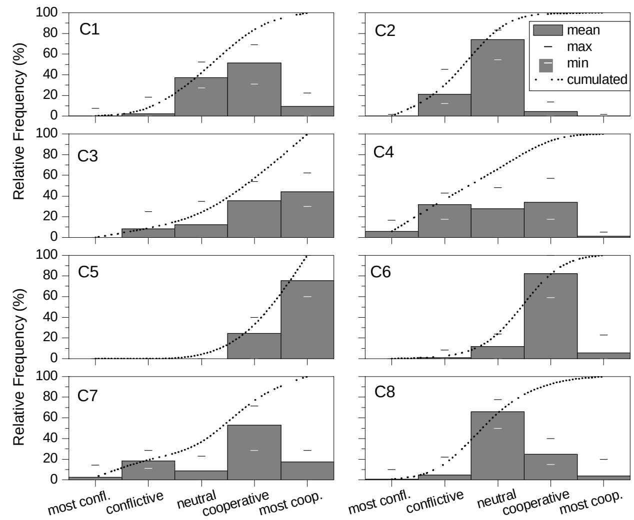

Figure 4. Mean relative frequencies and cumulative distribution of the events of conflict and cooperation for the eight clusters

within the geographic unit. We chose a classification tree model for the analysis, as it allows the inclusion of numeric as well as categorical variables that do not have to be normally distributed and can be interrelated. The method further accounts for the multiple conjunctional causality known to govern international relations; meaning that different combinations of factors can lead to the same outcome, and the same factor can have a different effect under different circumstances (Chan, 2002). A classification tree predicts a single categorical response variable (WIR Class 1 to 8) by the values of a set of predictor (explanatory) variables. Constructed by recursive partitioning of a learning sample of class values (response variable) and predictors, the tree finally represents a collection of many decision rules displayed in the form of a binary tree. Each point in the tree where a decision has to be made is called a "node": true or false leading either to the next decision rule or to a "terminal node," where a classification is assigned. For the dataset in this study the rules take the form of: "if the Index of Aridity *A* e" 1.5 and population density *Pop* e" 100 people/km^2 and the river is the border *B* = 1 and …, then the BCP is most likely in WIR Class 5."

Classification and Regression Tree analysis is incorporated in most commercial statistics software packages and several shareware programs are available. We used the program QUEST (Quick Unbiased Efficient Statistical Trees) by Shih (2003). The algorithm, which is described in detail in Loh and Shih (1997) and has been used in social and natural sciences problems, performs binary splits that are unbiased towards variable selection. At each node, a set of statistics is applied. First, the analysis of variance (ANOVA) F-statistic (significance level a= 0.05) determines the variable to be used to split the node. Then, a 2-means cluster analysis carries out an initial grouping into two superclasses, which is followed by a modified form of quadratic discriminate analysis (QDA) on the two superclasses to find the split point. A tree grown to correctly predict most of the learning sample usually has an impractically large number of nodes and often performs badly on an independent test sample. To find the optimum between tree size and predictive accuracy, QUEST includes the option for v-fold cross-validation pruning after Breiman et al. (1984).

Results

The classification tree was obtained for the learning sample of 135 BCPs that have data for the variables *A*, *P*, *CVPs*, *CVPt*, *SP*, *Pop*, *GDP*, and *DA*. Since the variables *Area*, *Loc*, *and B* were found to have no significant association with the response class, but would further reduce the learning sample size, they were not used for the final tree construction. The tree for the learning sample was grown all the way to correctly classify 87 percent of the BCPs. A more practical tree, however, is the (5-fold) CV-pruned tree with the smallest cross-validation cost, which is shown in Figure 5. It has 26 terminal nodes and a total number of 51 nodes. The tree correctly classifies 67 percent of the BCPs, with much better classification rates for Classes 1, 2, 3, and 8 (greater than 75 percent correct) and lower classification rates for Classes 4, 5, and 6. The cross-validation samples on average classify 43 percent correctly.

The root node divides the BCPs by their Index of Aridity into arid to sub-humid BCPs (left hand side, *A*d"0.8) and humid BCPs (right hand side, A>0.8). The most frequent variable in both sub-trees is the per capita GDP of the countries, a variable also found significant in the test of individual indicators by Yoffe et al. (2003). However, aridity, annual precipitation, and the seasonality of precipitation again play a major role in both sub-trees. Population density and temporal precipitation variability only partition classes on the arid to sub-humid (left) side of the tree, while democracy/autocracy level and spatial rainfall variability exclusively appear on the humid (right) side of the tree. In terms of the response, all classes are present on the left hand side of the root node, while on the right hand (humid) side only Classes 1, 3, 6, and 8 are classified.

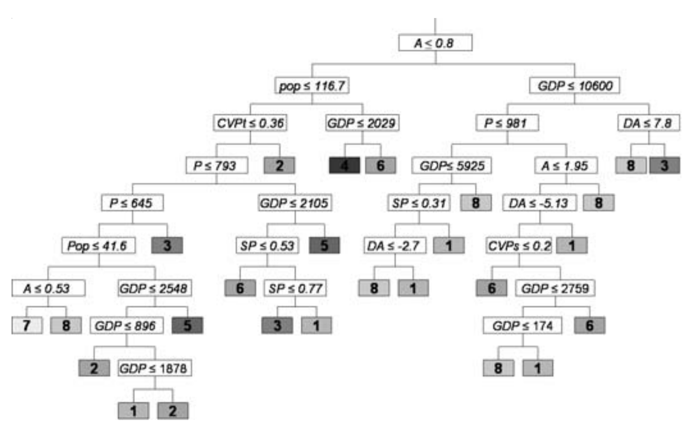

Figure 5. Classification tree with class value

Classes 4 and 7 have only one terminal node in the tree. Classes 1 and 8, which are also the largest groups, have the most terminal nodes, i.e. the most possible explanatory variable combinations.

The arid to sub-humid (left) side of the tree describes the more conflictive relations with Classes 2, 4, and 7 only occurring here. The most conflictive international relations type, Class 4, is split off by a relatively high population density and low per capita GDP, both found to be significant individual indicators of conflict by Yoffe et al. (2003). A substantial part of Class 2 is characterized by high inter-annual precipitation variability; the rest is characterized by medium population density, low annual precipitation, and a low to medium GDP. A lower population density but an even more arid climate distinguishes Class 7, which has a bimodal CCL distribution. Together, these factors contribute to social water stress, which as can be seen, may not necessarily be effective as a linear combination. Higher seasonality splits off classes with less cooperative events. The humid (right) side of the tree immediately splits off class 3 by high GDP and most democratic political conditions, and then is only concerned with the partitioning of Classes 1, 6, and 8.

A closer look at the misclassification shows that BCPs are often classified into classes with a similar CCL distribution to the real one. Most misclassified members of Class 1 and Class 4 are classified into Class 7 instead. Both combinations have similarities. Class 1 and 7 share a peak of cooperative events; Class 4 and 7 are characterized by events along the whole CCL scale. The three misclassified members of Class 5 were classified as Class 3. Both have a most cooperative maximum, and Class 5 is rather small and only contains countries riparian to three different basins.

Discussion

This study aimed to develop and test a first multivariate approach to analyze the influence of indicator variables on WIR patterns. We closely followed statistical rules to objectively determine the WIR classes and to develop the classification tree model. The model rules correctly explain two thirds of the BCPs WIR classes. Though very complex at first sight, the classification tree identified and confirmed important indicators and elucidates how they conjunct. Hydroclimatology and population density generally have a greater influence in dryer climates while economic conditions and political regime dominate the WIR classification in wetter climates. WIR with violent conflict are only found on the “dry” side of the classification tree, but other groups are found there as well and the classification depends on other combined influences. In humid regions, combinations of *GDP*, *DA* and hydroclimatic indicators determine mainly whether neutral or cooperative events dominate the international relations over water.

Some classes are difficult to distinguish and to explain. The most cooperative basins, for example, were difficult to classify, as the group is small and its characteristics are divers. Also, the large groups of WIR with mostly neutral and slightly conflictive events were difficult to be distinguished by the variables used in the model.

In a future application, a reduction of the number of WIR classes could be considered. Furthermore, additional and modified variables may improve the model. We put great emphasis on the new hydroclimatology dataset, but relied on only the most commonly accepted socioeconomic and political indicators. The hypothesized importance of climatic variability was confirmed by the model. However, the river water is the topic of the political events, and its quantity and variability might be altered by human influence or because of its exotic origin and therefore be somewhat different from the climate in the geographic unit of interest. A test with a smaller sample, for which data was available, indicated a great influence of river and discharge-based hydrological variables. However, for surface hydrology, no comparable global dataset was available. The Global Runoff Data Center (GRDC) provided data, but as the coverage would have substantially reduced the sample size (Wolf et al., 2003b), hydrologic parameters based on river discharge were not included in this study. Furthermore, qualitative issues related to the role of the river in a country might add explanatory power. An advantage of tree based models is that they can incorporate categorical variables such as the “model of institution” (Kliot et al., 2001) (if any) or the “benefit” that could be obtained through cooperation (Sadoff and Grey, 2002). Yet to date, such variables are not available globally.

The basin-country-polygon unit imposed some restrictions on variable selection. While it is an improvement towards a finer spatial scale, it only allowed the geographical position of the riparian country, and not the relations between countries to be considered. Furthermore, variables describing the overall political relations between countries (e.g. friendship-hostility index, trade relations), and differences or asymmetries in socioeconomic and political variables have shown explanatory potential at the dyad level (Yoffe et al., 2003a; Espey and Tofique, 2004; Song and Whittington, 2004).

When assessing the results, one should keep in mind that the dependent and independent variables were derived as an aggregate description over a common historic time period. Hence, the causal role of hydroclimatic and socioeconomic variables in conflict and cooperation over water cannot be interpreted as an “event-trigger.” They only reflect the combinations of conditions underlying different patterns of historic international relations. Temporal developments of the explanatory variables (e.g. drought, political regime-change, etc.) as well as temporal developments in the dependent variables such as treaty negotiation, adaptation, and learning processes cannot be taken into account. Their influence, which has been qualitatively mentioned in Yoffe et al. (2003) and Wolf et al. (2003a) (e.g. BAR scales improved after treaty; relations were deteriorating WIR at the time of the break-up of the Soviet Union) has to be tested within another methodological framework. There will always be a limit to explanatory modeling without the consideration of temporal and spatial interactions. Considering these restrictions to the aggregate approach, two thirds of correctly classified BCPs is a satisfying result.

Global Application – Global Implications

The study focuses on global WIR patterns and their influencing variables. For an appraisal of the classification and the model, one should therefore look at the result in a global geographic context. Figure 6 shows the spatial distribution of the WIR classes at the three stages of the analysis and modeling: a) the result of the cluster analysis of the historic WIR data (135 BCPs); b) the fit of the multivariate classification tree model (135 BCPs); and c) the application of the model to all BCPs of the 263 international basins worldwide (including the ones that were not used in a) and b) because they had less than five records of political events).

The cluster analysis distinguished eight groups of BCPs that represent the different patterns of aggregate WIR histories shown in Figure 4. Figure 6a illustrates their geographic distribution:

- Cluster 4 consists of India and Pakistan's share of the Indus River, most of the countries sharing the Jordan River, Armenia and Azerbaijan's WIR over the Kura-Araks River System, Myanmar and Thailand over the Salween, and the USA and Mexico over the Yaqui River. These BCPs represent the WIR-class with the highest percentage of most conflictive and conflictive events. They also have a high number of events indicating a long history of dispute and cooperation over their international rivers.
- Cluster 2 includes Turkey, Syria, and Iraq's WIR over the Tigris and Euphrates River Basin, Chile and Bolivia's WIR over the Cancoso (Lauca), and Greece and Bulgaria' WIR over the Vardar. The class represented by these BCPs also has a substantial amount of conflictive events. The cluster is small, but its members record many events. Compared to Cluster 4, however, there was less violent conflict, but also almost no cooperative events.
- Cluster 7 includes Kyrgyzstan and Uzbekistan's Aral Sea Tributaries, the USA and Mexico's WIR over the Colorado River Basin, Spain and France's WIR over the Ebro, and South Africa and Lesotho's WIR over the Orange River. They represent a WIR pattern with considerable conflict, the frequency of which however is outweighed by the many reported efforts of cooperation.
- Cluster 3 includes most riparians to the Danube River, Canada and the USA's events over the Columbia and the St. Lawrence, South Africa, Swaziland, and Mozambique's WIR over the Incomati, and Mozambique, Zambia, and Zimbabwe's WIR over the Zambezi River. The members of Cluster 3 reported issues of dispute, but the overall relations were dominated by cooperative and most cooperative events.
- Cluster 8 includes Spain and Portugal's WIR over several rivers, India and Bangladesh's history over the Ganges, Karnaphuli, and Fenney Rivers, Lebanon and Syria's WIR over the Asi, as well as Egypt, Sudan, and Ethiopia's events over the Nile, and the events of Hungary, the Czech Republic and Slovakia concerning the Danube. The group comprising Cluster 8 is large and the WIR pattern is dominated by a majority of neutral events. The members furthermore had a few conflictive events and about 20 percent of cooperative events.
- Cluster 1 includes the Mekong riparian countries, most riparian's history of events over the Rio de La Plata, USA and Mexico's WIR over the Rio Grande, Gambia, Senegal and Guinea's events over the Gambia River, Mali's over the Senegal River and Poland's events over the Odra River. This is the largest cluster. Its members have a significant proportion of neutral verbal events, which is slightly exceeded by cooperative events. This pattern indicates a moderately positive cooperation, yet without much concrete action.
- Cluster 6 includes the Amur River, Laos and Vietnam's WIR over the Ma and Red Rivers, Mexico and the USA's relations over the Tijuana, Kazachstan and China's history over the Ili River, and Uganda and Tanzania over the Nile. It is a large cluster, but its member had a small number of events per BCP. These were mostly cooperative (over 80 percent of all events). However, only a very small percentage of events were taking action (most cooperative).
- Cluster 5 consists of Germany and France's events over the Rhine River, and surprisingly South Africa and Mozambique's history over the Limpopo, and South Africa's events over the Maputo. The cluster describes a WIR pattern of mainly most cooperative events with the rest being cooperative.

Modeling the WIR clusters with the classification tree model resulted in a few misclassifications (Figure 6b). With the exception of the Ganges-Brahmaputra, the model classification always determined a class with a higher proportion of conflictive events. This overestimation of the degree of conflict in BCPs such as Mexico's share of the Rio Grande, Egypt's and Sudan's shares of the Nile and all riparians of the Aral Sea by the model, may point to unstable WIR in these basins.

Geographically, the BCPs used in a) and b) are distributed over all continents. They represent about 20 percent of the world's international basin-country-polygons and cover about 36 percent of the whole area of international river basins. Of the remaining BCPs, 30 percent (48 percent of the area) have less than five events, 50 percent (but only 16 percent of the area) have no records of events. The model was applied to those last two groups to esti-

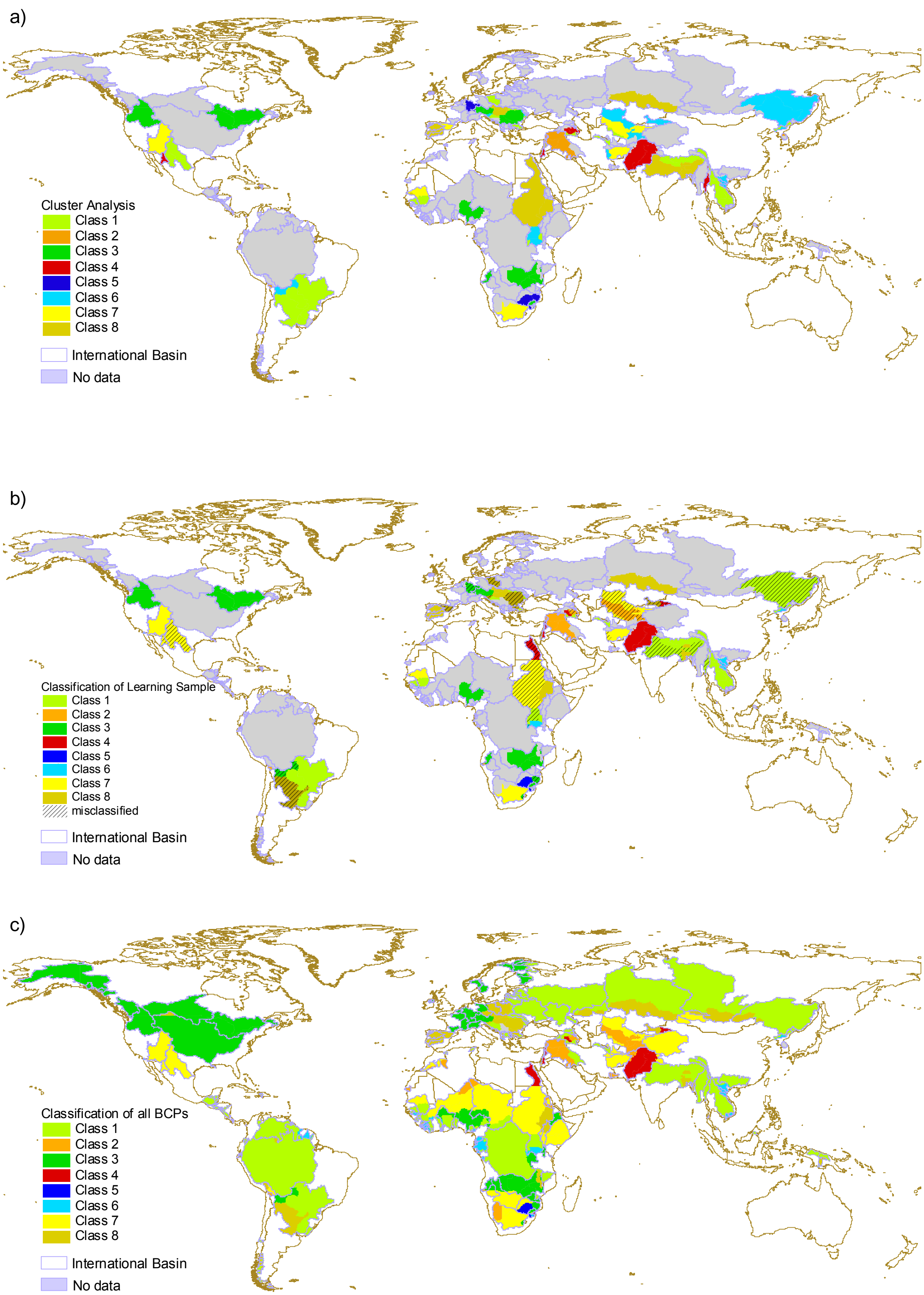

Figure 6. Spatial Distribution of the classes of WIR history; a) Result of the cluster analysis on political event data; b) Fit of classification tree; c) Application of tree rules to all BCPs

mate their most likely WIR class. The map in Figure 6c shows the geographic distribution of this model application.

Together with the Middle East, Central Asia emerges as the most conflictive region. Both are known and widely recognized as critical regions (e.g. Yoffe et al., 2003). The model also estimated considerable risk for conflict for the Umbelozi in Mozambique and the Zarumilla in Ecuador; both are small basins barely visible on the map. Cooperation is estimated to dominate in North America and Central Europe, but also in some smaller South American and African basins. The majority of BCPs are estimated to be WIR Class 1 (mainly cooperative and neutral), particularly if located in a humid climate. In fact, the global picture shows the strong influences of climate. An example is the geographically dominant pattern of Class 7 (conflicts but also considerable efforts of cooperation). The class is estimated exclusively for semi-arid regions (i.e. outer tropics in Africa, the South of North America, and continental Asia). These are regions where resource scarcity stresses but also constantly calls for a strengthening of past and recent cooperation. Another important result from the model application experiment is the classification of not openly conflictive but also not very actively cooperative basins (Class 8). In several basins in southern and eastern Europe, large basin-country shares in South America (e.g. La Plata), Central Africa (e.g. Congo), and across Asia (e.g. Ob and Amur) it will be important to move beyond verbal cooperation toward the formation of treaties, as with only weakly established cooperation in place, climatic, environmental or political changes might easily result in deteriorating international relations.

Conclusion

Water-related international relations (WIR) between countries that share the world's 263 transboundary rivers show strong patterns. We tested a new approach to objectively classify the aggregate histories of conflict and cooperation in the world's international river basins from available political event data from 1950 to 2000. The approach is contrasting to conventional approaches that classify a basin as either "conflictive" or "peaceful" or concentrate on the occurrence of conflict. The patterns determined are mostly dominated by neutral and cooperative events, but differ in the level and proportion of cooperation to conflict. Their geographic coherence varies. The classes with the highest conflictive proportion occur in the Middle East and in Asia, confirming the global patterns of Category 1 Basins-at-Risk (current conflicts) by Yoffe et al. (2003). Most other WIR patterns are distributed over several continents.

In order to statistically explore and model combinations of factors influencing these patterns and their global distribution, a classification tree approach was chosen. Using hydroclimatic, socioeconomic, and political variables to estimate the aggregate historic WIR class, the model classifies two-thirds of the BCPs correctly. The results

- confirm the main hypothesis of multiple conjunctural causality and thresholds
- specify underlying causality such as the influence of: 1) water stress due to high population density on violent conflict dominating in arid and semi-arid regions; or 2) economic and political factors on cooperation dominating in humid regions
- illustrate the importance of hydroclimatology and hydroclimatic variability (and hence resource scarcity and reliability) for both conflict and cooperation

Applying the model to all international basins worldwide provides a global geographic coverage of estimated most likely WIR classes. The result emphasizes some previously mentioned geographic patterns with the best cooperation in North America and Central Europe, less concrete cooperation in temperate humid and tropical areas of South America, Africa, and Asia and critical regions in most semi-arid climatic transition zones. Here, water scarcity and uncertainty in the availability, indicated by hydroclimatic factors, has a major influence on the WIR. If it is not addressed in cooperation efforts, these regions may be prone to join the known conflict areas in Africa, the Middle East, and Asia. A verification of the global model is difficult due to the lack of data. However, comparing the estimated class with the CCL scale of the few events that were recorded in some basins gives confidence in the classification.

The study illustrates how complex the causality of water-related international relations around the world is and the chosen modeling approach suggest directions for disparate emphasis on processes of conflict prevention and resolution. Nuances of cooperation are more difficult to distinguish and explain than the occurrence of conflict. Considering that in many international basins, the riparian countries interact and show interest in cooperation, it is important for national and international policy to address the right issues to move towards successful cooperation and agreements over international waters. The approach presented in this article was a first attempt towards a tree-based model that can assist in such decision making. Though not suited for assessments or decision making of individual basins at this stage, it provided a good framework for a global test and assessment and shows potential for further development. Future attempts should specifically address the underlying and immediate causes for successful cooperation and hence conflict prevention. The underlying causality investigated here can ultimately only explain a part of the aggregate WIR patterns. More research is needed to shed light on the many aspects of temporal causality and how the combination of underlying stressor and trigger variables can be used to estimate or predict risk and potential for conflict and cooperation.

Acknowledgements

The research was funded by a scholarship from the German Research Foundation (DFG) and carried out at the Department of Geosciences at Oregon State University. Data and support from the Department and the Transboundary Freshwater Dispute Database and Team is gratefully acknowledged. The CRU kindly provided the precipitation data. The author is particularly grateful for Shira Yoffe's introduction to the political event data and Aaron Wolf's support, advice and critical review of the manuscript.

About the Author

Kerstin Stahl is currently a postdoctoral research fellow at the Department of Geography at the University of British Columbia. Her research focuses on regional to global-scale hydroclimatic datasets and their use to study the influence of hydroclimatic variability on environment and society. Current and recent projects include the investigation of the influence of climatic variability on the Mountain Pine Beetle Epidemic in B.C., a study of the influence of hydrologic conditions and hydroclimatic variability on water-related political conflict and cooperation in global international river basins, and several aspects of hydrologic drought in Europe. She has a Ph.D. in Hydrology from the University of Freiburg in Germany.

Discussions open until February 1, 2006.

References

Ahn, C.H. and R. Tateishi. 1994. "Development of a Global 30-minute grid Potential Evapotranspiration Data Set." *Journal of the Japan Soc. Photogrammetry and Remote Sensing* 33, No. 2: 12-21.

Ashton, P.J. 2002. "Avoiding Conflicts over Africa's Water Resources." *Ambio* 31, No. 3: 236-42.

Azar, E.E. 1980. "The Conflict and Peace Data Base (COPDAB)." *Journal of Conflict Resolution* 24, No. 1: 143-52.

Breiman, L., J.H. Friedman, R.A. Ohlshen, and C.J. Stone. 1984. *Classification and Regression Tree*s. Belmont, California: Wadsworth.

Chan, S. 2002. "On different types of international relations scholarship." *Journal of Peace Research* 39, No. 6: 747-56.

Dinar, A. 2004. "Exploring Transboundary Conflict and Cooperation." *Water Resources Research* 40, W05S01, doi:10.1029/2003WR002598.

Dobson, J.E., E.A. Bright, P.R. Coleman, R.C. Durfee, and B.A. Worley. 2000. "LandScan: a global population database for estimating populations and risk." *Photogrammetric Engineering & Remote Sensing* 66, No. 7: 849–57.

Espey, M. and B. Towfique. 2004. "International bilateral water treaty formation. *Water Resources Research* 40, W05S05, doi:10.1029/2003WR002534.

Falkenmark, M. 1989. "The massive water scarcity now threatening Africa—why isn't it being addressed?" *Ambio* 18, No. 2: 112–18.

Fiske, G. and S.B. Yoffe. 2001. "Use of GIS for Analysis of Indicators of Conflict and Cooperation over International Freshwater Resources." http://transboundarywaters.orst.edu/documents/fiske-yoffe.pdf.

Fisher, F.M., S. Arlosoroff, Z. Eckstein, M. Haddadin, S.G. Hamati, A. Huber-Lee, A. Jarrar, A. Jayyousi, U. Shamir, and H. Wesseling. 2002. "Optimal water management and conflict resolution: The Middle East Water Project." *Water Resources Research* 38, No. 11: 1243.

Gleick, P. 1993. "Water and conflict: fresh water resources and international security." *International Security* 18, NO. 1: 79-112.

Hastie, T., R. Tibshirani, and J. Friedman, 2001. *The elements of statistical learning. Data mining, inference, and prediction*. New York: Springer.

Homer-Dixon, T. F. 1994. "Environmental scarcities and violent conflict: evidence from cases." *International Security* 19: 5-40.

Kelly, M. and K. Chapman. 2002. "Sharing the Waters." *Americas Program Investigative Article*. Silver City, NM: Interhemispheric Resource Center.

Kliot, N., D. Shmueli and U. Shamir. 2001. "Development of institutional frameworks for the management of transboundary water resources." *International Journal of Global Environmental Issues* 1, No. 3-4: 306-28.

Loh, W.-Y. and Y.-S. Shih. 1997. "Split selection methods for classification trees." *Statistica Sinica* 7: 815-40.

Markham, C.G. 1970. "Seasonality of precipitation in the United States." *Annals of the Association of American Geographers* 60: 593-597.

New, M.G., M. Hulme and P. D. Jones, 2000. "Representing twentieth-century space-time climate variability. Part II: Development of 1901-1996 monthly grids of terrestrial surface climate." *Journal of Climate* 13: 2217-38.

PolityIV Project. 2000. *Polity IV Dataset*. (Computer file; version p4v2000). College Park, MD: Center for International Development and Conflict Management, University of Maryland. Accessed 8 May 2003. http://www.cidcm.umd.edu/inscr/polity/

Postel, S. and A.T. Wolf. 2001. "Dehydrating Conflict." *Foreign Policy* Sep/Oct: 60-67.

Sadoff, C.W. and D. Grey. 2002. "Beyond the river: the benefits of cooperation on international rivers." *Water Policy* 2: 389-403.

Savenije, H.H.G., and P. van der Zaag. 2000. "Conceptual Framework for the management of shared river basins; with special reference to the SADC and EU." *Water Policy* 2: 9-45.

Shih, Y.-S. 2003. *QUEST User Manual*. Accessed 3 Sep. 2003. http://www.stat.wisc.edu/~loh/quest.html

Song, J. and D. Whittington. 2004. "Why have some countries on international rivers been successful negotiating treaties? A global perspective." *Water Resources Research* 40, W05S06, doi:10.1029/2003WR002536.

Stahl, K. and A.T. Wolf. 2003. "Does hydro-climatic variability influence water-related political conflict and cooperation in international river basins?" *Proceedings CD of the Int. Conference on Hydrology of the Mediterranean and Semi-Arid Regions*. April 2003. Montpellier, France.

TFDD. 2003. http://www.transboundarywaters.orst.edu. Accessed 15 June 2005.

Toset, H.P.W., N.P. Gleditsch and H. Håvard. 2000. "Shared Rivers and Interstate Conflict." *Political Geography* 19, No. 8: 971-6.

UNEP and OSU. 2002. *Atlas of International Freshwater Agreements*. Nairobi: UNEP.

UNEP GRID. 2003. http://www.grid.unep.ch/data/grid/climate.php. Accessed: 3 Sep 2003.

UNESCO. 1997. "Map of the world distribution of arid regions. Accompanied by explanatory note." *MAB Technical Notes No. 7*. Paris: UNESCO.

van der Zaag, P. and A.C. Vaz. 2003. "Sharing the Incomati waters: cooperation and competition in the balance." *Water Policy* 5, No. 4: 349-68.

Vörösmarty, C.J., P. Green, J. Salisbury, and R.B. Lammers. 2000. "Global Water Resources: Vulnerability from Global Climate Change and Population Growth." *Science* 289: 284-88.

Wolf, A.T., J. Natharius, J. Danielson, B. Ward, and J. Pender. 1999. "International River Basins of the World." *International Journal of Water Resources Development* 15, No. 4: 387-427.

Wolf, A., S.B. Yoffe, and M. Giordano. 2003a. "International Waters: Identifying Basins at Risk." *Water Policy* 5: 29-60.

Wolf, A.T., K. Stahl, and M.F. Macomber. 2003b. "Conflict and cooperation within international river basins: The importance of institutional capacity." *Water Resources Update* 125: 31-40.

World Bank. 2003. Online databases: GDP data. Accessed 3 Sep 2003. http://www.worldbank.org/data/

Yoffe, S.B. and K. Larson. 2002. "Basins at Risk: water event database methodology." In S. B. Yoffe *Basins at Risk: Conflict and Cooperation Over International Freshwater Resources*. Dissertation, Department of Geosciences. Oregon State University, Corvallis, 2002.

Yoffe, S.B., A.T. Wolf, M. Giordano. 2003. "Conflict and Cooperation over International Freshwater Resources: Indicators of Basins at Risk." *Journal of American Water Resources Association* 39, No. 5: 1109-26.

Yoffe, S., Giordano, M., Giordano M.A., Larson, K., Stahl, K., Wolf, A.T. 2004. "Geography of international water conflict and cooperation: Data sets and applications." *Water Resources Research* 40, W05S04, doi:10.1029/2003WR002530.