Northeast Journal of Complex Systems (NEJCS)

Volume 6
Number 1 *Special Issue: NERCCS 2024 Papers*

Article 3

June 2024

Multivariate Analysis of the Temporal and Spatial Correlations of the Global Human Rights Dataset

Amanda Goodrick
Binghamton University, SUNY, agoodri2@binghamton.edu

Skip Mark
University of Rhode Island, skipmark@uri.edu

Mikhail G. Filippov
Binghamton University--SUNY, filippov@binghamton.edu

David Cingranelli
Binghamton University, SUNY, davidc@binghamton.edu

Hiroki Sayama
Binghamton University, SUNY, sayama@binghamton.edu

Recommended Citation

Goodrick, Amanda; Mark, Skip; Filippov, Mikhail G.; Cingranelli, David; and Sayama, Hiroki (2024) "Multivariate Analysis of the Temporal and Spatial Correlations of the Global Human Rights Dataset," *Northeast Journal of Complex Systems (NEJCS)*: Vol. 6 : No. 1 , Article 3.
DOI: 10.22191/nejcs/vol6/iss1/3
Available at: https://orb.binghamton.edu/nejcs/vol6/iss1/3

Multivariate Analysis of the Temporal and Spatial Correlations of the Global Human Rights Dataset

Amanda Goodrick[1,2*], Skip Mark [3], Mikhail Filippov [1],
David Cingranelli [1], and Hiroki Sayama[1,4]

[1] State University of New York at Binghamton, Binghamton, USA

[2] Slippery Rock University, Slippery Rock, USA

[3] University of Rhode Island, Kingston, USA

[4] Waseda Innovation Lab, Waseda University, Tokyo, Japan

[*] agoodri2@binghamton.edu

Abstract

We propose an extension of a previously proposed method for lead-lag analysis of multivariate time series to include the analysis of spatial correlations. We applied the extended spatial and temporal method to CIRIGHTS, a large global human rights dataset, in order to determine the most influential and most influenced indicators of human rights, freedoms, and atrocities over time. We consider four target countries, each from a different continent. The previously proposed method used a weighted directed network with several lags of each variable as nodes and with edges weighted by transfer entropy. In this study, that method is extended to the spatial multivariate time series by investigating geographical correlations between each target country and its surrounding region to determine which variables best represent activity between the target country and the surrounding region. The nodes in this undirected network are variables in the target country and variables in the surrounding region, with edges weighted by mutual information. Eigenvector centrality is used to determine which variables best represent activity between the target country and the surrounding region. Key findings indicate that occupational and worker rights are the most influential and most influenced in the target countries over time and by region, and laws tend to influence future activity.

1 Introduction

Human rights are essential to maintaining a reasonable quality of life. The minimal desired level of rights and freedoms allows for basic needs to be met. Lower levels of these rights and freedoms easily lead to disruption, corruption, and inability to meet those basic needs. Because we are all human, this topic impacts everyone. With these freedoms, we are then responsible to make wise choices to enhance the life experience for ourselves and others.

Human rights are complex and fairly subjective. The data can be messy and incomplete because different regions have different governments. Consequently, there are many technicalities that precede any attempt to analyze the data [1, 2]. Recent past work has greatly focused on collecting and maintaining an accurate data set [1, 3, 4]. Additional complexity is due to standards of accountability for human rights and how to best measure and code these variables, without loss of information, in order to use them for analysis [5, 6]. Human rights, in general, may be changing over time, which has been a topic of interest, including debate about whether or not human rights have improved over time [5, 6, 7, 8]. This debate occurred after an earlier study that used an additive physical rights index, based on extrajudicial killings, disappearance, torture, and political imprisonment, to investigate the likelihood of human rights improvement by weighting variables based on a proportion of spatial distances between state borders less than 50 km. [1, 3, 9]. One study found that while atrocities may be decreasing, brutality-based atrocities have been increasing [10]. Change in human rights over time is nearly inevitable, though that change complicates the data collection even more.

There has been some work comparing machine learning methods to human coding, both done by reading textual reports, though a recent study discovered divergent results of accuracy, likely due to the change in human rights over time [11]. An additional effort was made to investigate best measurement practices and extract indirect classifications of human rights variables using natural language processing in text from the US State Department, Amnesty International, and Human Rights Watch over time [12].

Obtaining reasonably accurate data and minimal information pertaining to the current status of human rights is only one element of complexity. Most statistical methods work best on quantitative data with only a few variables. Categorical or ordinal data can be more difficult to analyze adequately. Most studies have been able to compare only a few variables to reach meaningful conclusions. For instance, one study investigated the likelihood of violence, terrorism, and/or civil war as it relates to brutality-based atrocities and women's rights [13]. Another study considered the implementation of laws and promise-keeping using 3 variables: right to a fair trial, children's rights, and the right of workers to form unions [14]. Another study examined the effects of region on levels of protection of physical integrity rights by developing several ordinal logistic models to investigate the level of respect for human rights, using the previously mentioned additive physical rights index and four individual physical integrity rights as dependent variables for each model [15].

These logistic models weighted spatial regions based on distance from the reference region, excluding regions more than 950 km from the reference region, with independent variables including population size, wealth, and other measures of turmoil and unrest [15].

Due to the complexity and dynamics of human rights over time and over space, studies have not been able to adequately capture the patterns, relations, and dependencies among the multiple types of human rights and freedoms, nor have they been able to define relevant predictors in a rigorous way. In order to advance this science and get ahead in the game, it is critical that scholars develop a "more comprehensive and theoretically sound indicator of mechanisms, in addition to indicators of causes and outcomes" [16].

In this study, we aim to determine the most influential and most influenced human rights over the time period investigated. We also aim to determine how human rights are impacted within a surrounding region. This is only the tip of the iceberg, where there is much work to be done. These results are not directly predictive, but they provide insight to allow researchers to better apply predictive and prescriptive methods that will enable more informed and responsible decisions by citizens and lawmakers.

We used the CIRIGHTS dataset to investigate relationships between human rights indicators over time and by surrounding region. This data set is the largest maintained data set of human rights in the world [1, 3, 4]. We investigated the data from the years 2001-2021, so there were 21 annual observations for each of the 31 ordinal time series variables that we included in this model analysis. The variables used have been classified into 4 categories, as indicated below, where variables notated with (L&P) indicates the inclusion of two variables, both by law and in practice [1, 3, 4].

- *Physical Integrity Rights*: Political and Other Extrajudicial Killings (later modified to Arbitrary or Unlawful Deprivation of Life), Disappearance, Torture, Political Imprisonment

- *Empowerment Rights and Freedoms*: Freedom of Speech and Press, Freedom of Religion, Freedom of Domestic Movement, Freedom of Foreign Movement and Travel, Freedom of Assembly and Association, Electoral Self-Determination, Women's Political Rights, Women's Economic Rights

- *Occupational/Workers' Rights*: Right to form Workers' Union (L&P), Right to Bargain Collectively (L&P), Right to be Free from Forced or Compulsory Labor (L&P), Children's Rights (L&P), Right to Min Wage (L&P), Right to Occupational Safety (L&P), Reasonable Limitation on Working hours (L&P), Human Trafficking (L&P)

Table 1: Countries in Surrounding Region per Target Country

Algeria	Libya, Morocco, Mauritania, Mali, Niger, Tunisia
Colombia	Brazil, Ecuador, Panama, Peru, Venezuela
Germany	Austria, Belgium, Czechia, Denmark, France, Luxembourg, Netherlands, Poland, Switzerland
India	Bhutan, Bangladesh, China, Myanmar*, Nepal, Pakistan, Sri Lanka

* Excluded due to many missing values

- *Justice Rights*: Right to a Fair Trial (L&P), Judiciary Independence

In this study, we chose Algeria, Colombia, Germany, and India, each sampled from a different continent, as four target countries of our analysis. We investigated a lead-lag network to determine influential variables in each of the four target countries. To determine if activity within the target country was best represented by the surrounding region, or vice versa, we used the same ordinal data and created networks of the target country with the surrounding region, where the surrounding region for each is given in Table 1 and can be verified by any worldwide map of countries.

In a previous study, Goodrick and Sayama proposed a robust method to investigate the lead-lag effect in multivariate time series economic indicators during the COVID-19 era, from July 2019 to December 2022, using quantitative continuous data [17]. Networks were constructed using three different relationship metrics: correlation, mutual information, and transfer entropy, and it was concluded that transfer entropy was the best metric, due to the robustness of the results from this model when determining the most influential and the most influenced economic indicators in the COVID-19 era. Transfer entropy was considered the best of the three relationship metrics because results were most insensitive to changes in a damping parameter, the variable ranking was easier to interpret because it had a larger range, and results were the least impacted by variables with greater variation [17].

Results from this method indicate the most influential and most influenced variables over a time period by investigating a lead-lag effect, or a relationship between lead variables and lagged variables, where lagged variables are shifted backward in time. A lead variable provides information about future system behavior because it acts first and has a relationship with other variables at some future time. A lagged variable provides information about past system behavior because it is a variable that has a relationship with other variables at some past time.

We refer to these variables as influential and influenced due to the relationship over time, though influence in no way implies causation. Runge lays out several density estimation methods that are useful in constructing causality networks, but our human rights variables are ordinal, so pairwise relationships are directly calculated rather than estimated [18]. We investigate pairwise relationships over time and over region, and the results do not imply causality. The previous study indicates the novelty of this proposed method, and it was the first known study to investigate a lead-lag effect with a complex network while retaining and incorporating relationship strength, time difference, and a damping parameter [17].

We applied that same method to categorical data in this study, using transfer entropy to quantify relationships in multivariate ordinal time series data. This is a novel extension of that previously proposed method, as it is the first time that it is being applied to categorical or ordinal data. Influence over time with lagged variables has typically been established using Granger causality analysis, which requires stationary quantitative data, but categorical data cannot be transformed into a stationary time series. A few recent studies propose a Granger causality network for multivariate categorical data. One such method performed Granger causality analysis by applying Markov transition probabilities from both a mixture transition distribution and a multinomial logistic transition distribution [19, 20]. While this method is quite promising, it requires probability distributions and relies on many assumptions and parameter selections. Fish et al. expanded on causation entropy, a generalization of transfer entropy, to investigate causal relationships in gene expression networks that follow a Poisson distribution by efficiently approximating mutual information [21]. Another recent study investigates Bivariate Partial Information Decomposition in Poisson and multinomial systems [22].

The Goodrick and Sayama framework can handle raw data, and it uses relationship metrics of mutual information and transfer entropy, which have already been well-established as appropriate relationship metrics for categorical or ordinal data. Densities were directly calculated in the mutual information and transfer entropy calculations.

2 Methodology

For each of the four target countries, we created a lead-lag (LL) network to investigate the relationships among human rights indicators over time with the 31 variables in each target country as nodes. Relationships in this directed network were quantified by transfer entropy. We then created a target and surrounding region (TS) network for each of the four target countries to investigate which human rights indicators within the surrounding region best represented activity in the target country and which human rights indicators within the target country best represented activity in the surrounding region. This network had 62 possible nodes for each target country network, with 31 variables of the target country as nodes and 31 variables of the average of the surrounding region as nodes.

2.1 Relationship metrics

To quantify the pairwise relationships, we used transfer entropy, $T_{X\to Y}$, for the directed LL network, and mutual information $I(X;Y)$ for the undirected TS network. Mutual information is given by

$$I(X;Y) = H(X) + H(Y) - H(X,Y) = \sum\sum p(xy)\log\frac{p(xy)}{p(x)p(y)}. \tag{1}$$

In Eq. 1, $H(X)$ is defined as the entropy of a variable, $H(X,Y)$ is the joint entropy, $p(x,y)$ is the joint probability, and $p(x)$ is the marginal probability. Then $I(X;Y)$ is the amount of information obtained about a variable X by knowing about a variable Y [23]. The logarithmic function is typically base 2 and mutual information is returned in bits of information. If two variables are independent of one another, $I(X;Y) = 0$, and if two variables are dependent on one another, $I(X;Y) > 0$, so all of the edges included in the TS network had positive edge weights. Mutual information is determined by the probability that each outcome will occur, but the outcome itself is irrelevant. Mutual information can be used for categorical and ordinal relations, so it is an appropriate relationship metric.

Transfer entropy, $T_{X\to Y}$, is the amount of information gain in future values of Y by knowing past values of X in addition to past values of Y. Transfer entropy is given by

$$T_{X\to Y} = I(Y_t^{(l)} : X_{t-1}^{(l)} | Y_{t-1}^{(k)}) = H(Y_t|Y_{t-1}^{(k)}) - H(Y_t|Y_{t-1}^{(k)}X_{t-1}^{(l)}) \tag{2}$$

where l is the variable lags from 1 through fixed maximum lag l, and $H(X|Y)$ and $I(X;Y)$ are conditional entropy and mutual information, respectively [24].

Transfer entropy is not a symmetric relation, $T_{X\to Y} \neq T_{Y\to X}$, so the direction of the relation matters, and $T_{X\to Y} \geq 0$, so all edge weights in the LL network were positive, and asymmetric relations existed only for relations directed to a future time. Pairwise transfer entropy was calculated using the transfer_entropy function from the python implementation PyInform [25].

2.2 LL relationships

To create the LL network of relations over time, the first 5 lags for each variable were created, which increased the number of variables to 186, since there were 5 new variables for each of the existing 31 variables. Each lagged variable had a sample size one less than the previous lag, so the variable of the greatest lag had 16 annual values. It is relevant to note that results were robust based on the number of lags chosen. We let X_{ik} denote the lag k of variable X_i for $i = 1, 2, ..., 31$ and $k = 0, 1, ..., 5$, where $k = 0$ indicates the original variables. A directed graph was created with the 186 variables as nodes with a directed edge from node X_{ik} to node X_{jm} if $k > m$, for $i \neq j$ and $k \neq m$. Then each existing edge was directed toward the pairwise lead variable, and instantaneous relations were neglected. Edge weights were based on pairwise transfer entropy and pairwise lag difference. Pairwise lag difference is given by

$$D_{km} = |k - m|, \text{ for } k \neq m. \tag{3}$$

A parameter a was also incorporated, where $0 < a \leq 1$. This parameter represents the proportion of relationship remaining over time D_{km}. Then each edge weight is given by

$$w_{ik,jm} = T_{X_{ik} \to X_{jm}} a^{D_{km}}. \tag{4}$$

This 186-node network was simplified by aggregating nodes of each variable over all lag values. For each node i, X_{ik} was merged into a single node for all k, resulting in a 31 node graph. For each directed edge from node i to node j, edge weights were summed so that each node pair had exactly two directed edges, each with a single weight,

$$w_{i,j} = \sum_k \sum_m w_{ik,jm}. \tag{5}$$

This simplified graph had 31 nodes with weighted edges directed toward the lead variable. PageRank (PR) was used to determine the most influential node over the time period when edges were directed backwards in time. The PR of a node is an iterative process, and the PR of node i depends on the PR of all other nodes in the network. Since PR calculates a measure of importance based on incoming links, when edges were directed backward in time toward the pairwise lead variable, the node with the greatest PR was considered to be the most influential. Edge weights remained the same, and the direction of each edge was reversed and directed toward the lagged variable of the pair. PR was used to determine the most influenced node during the era when edges were directed toward the pairwise lag variable. It is important to emphasize that this terminology in no way implies causation.

2.3 TS relationships

To create the network of TS relations, nodes were based on the 31 variables in the target region and the 31 variables in the surrounding region. For each variable in the surrounding region, values were averaged for all surrounding countries per year to create a new mean regional annual time series variable for each of the 31 variables. These values were not necessarily whole numbers, but they were still discrete, so it was appropriate to quantify the relationship using mutual information between two discrete variables.

We let r denote the region, where $r = 0$ is the target country and $r = 1$ is the surrounding region. Then for each variable X_i for $i = 1, 2, ..., 31$, there was a node X_{ir} for $r = 0, 1$, so the graph had 62 nodes. Edges were weighted by pairwise mutual information, so

$$w_{i,j} = I(X_i; X_j), \tag{6}$$

and node X_{ir} was connected to node X_{js} if $r \neq s$ and $w_{i,j} > 0$. Then nodes were linked if they were between any two variables from the target region and from the surrounding region. Edges between nodes in the same region were neglected.

Eigenvector centrality (EC) was used to determine which human rights variables from the surrounding region were most representative of activity in the target country and which variables in the target country were most representative of activity in the surrounding region. EC and PR are variations of one another, with EC typically used for undirected graphs and PR typically used for directed graphs. Considering the node with the greatest connections as the most important, EC measures the importance of a node based on its connections with other important nodes, so similarly to calculating PR, EC is based on the EC measure of the other nodes in the network.

EC was calculated for each node in the TS network, so all nodes from the target region and the surrounding region were ranked based on EC. The EC values per node were extracted for the target country nodes and also for the surrounding region nodes. The nodes of the target country with the greatest EC values were considered to best represent activity in the surrounding region, since edge weights represented the relationship strength between nodes of the two geographical regions. Similarly, the nodes of the surrounding region with greatest EC values were considered as those variables that best represent activity in the target country.

3 Results

3.1 LL network

Each simplified graph of the LL networks had 31 nodes with edge weights based on transfer entropy and lag difference. PR was calculated for each of the nodes in each network to determine the most influential and most influenced variables in each target country over the time period investigated.

Once all nodes were merged and edge weights were summed, PR was calculated and plotted to investigate behavior of the human rights indicators over time. The PR values for each target country are shown in Figure 1, with the most influential nodes shown on the left, and the most influenced nodes shown on the right. Variables on the x-axis are sorted by average PR, greatest to least. Since PR values are based on incoming weighted edges, and outgoing links are not considered, any node that has a relatively high PR value in both lists is one that has a heavily weighted connection

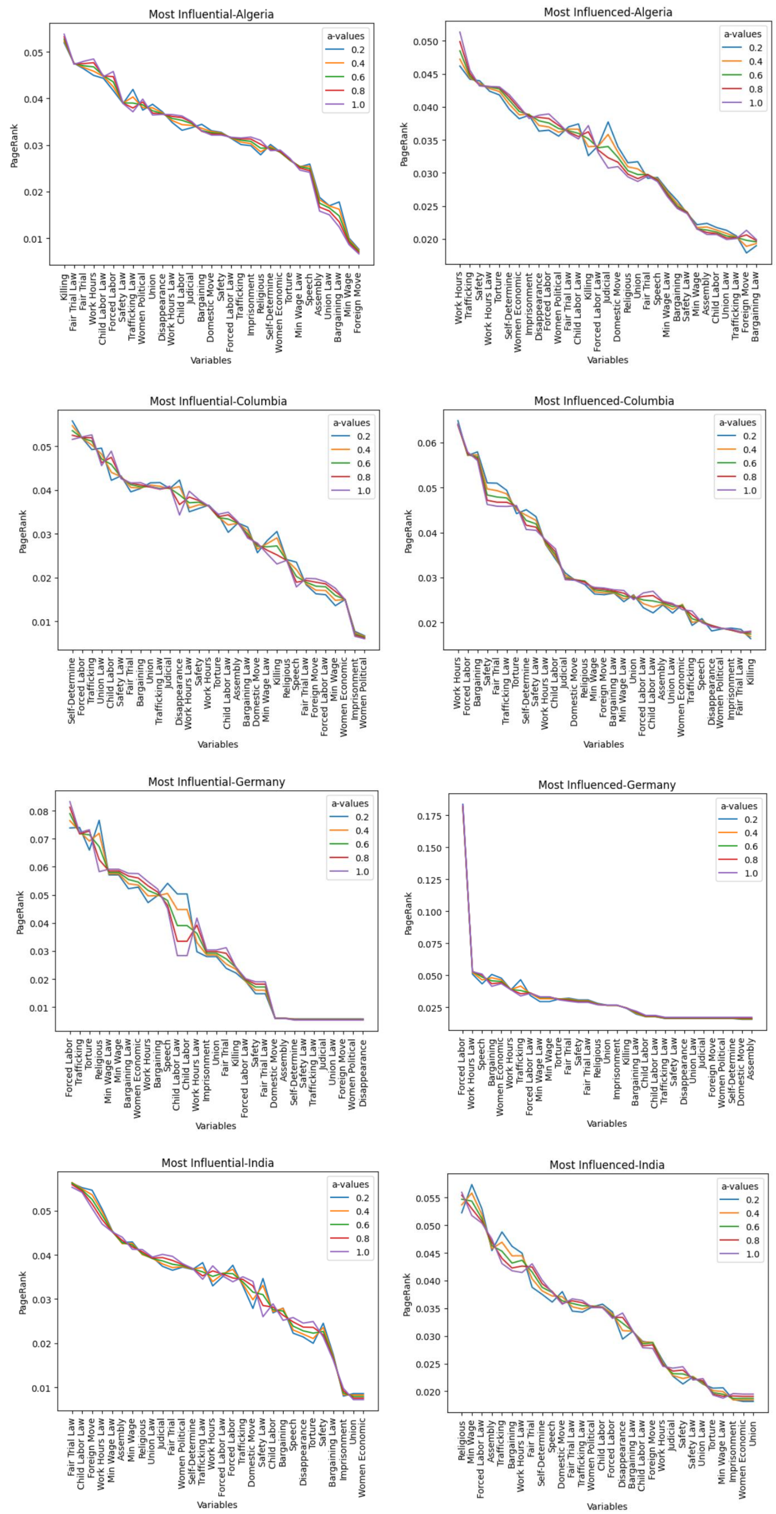

Figure 1: PR values (y-axis) per node (x-axis) for networks with edge weights based on transfer entropy and determined by select values of a in the legend. The nodes on the x-axis are sorted by average PR. The most influential node is shown on the left and has greatest PR with edges directed toward the lead variable. The most influenced node is shown on the right and has greatest PR with edges directed toward the lagged variable.

with other variables at some past time and with other variables at some future time, and influence does not mean causality. For instance, work hours in Algeria and forced labor in both Colombia and Germany are variables that are heavily weighted regardless of whether the edges were directed toward lead variables or lagged ones. As expected, results were insensitive to changes in a, which was the proportion of the relationship remaining over time, indicating the robustness of this method.

3.2 TS network

The TS networks had 62 possible nodes, where pairwise edges existed between a node of the target country and a node of the surrounding region if mutual information was positive. All resulting networks had fewer than 62 nodes, indicating that none of the networks had all variables sharing information with all other variables. In particular, Algeria is missing foreign movement, Colombia is missing women's political rights, Germany is missing foreign movement, judicial independence, safety law, electoral self-determination, trafficking law, union law, and women's political rights, while disappearance is missing from both Germany and its surrounding region. India is missing killing, union, and women's economic rights. Missing nodes in the target region indicate that the missing variables are independent of all variables in the surrounding region, and missing nodes in the surrounding region indicate that the missing variables are independent of all variables in the target country.

The EC results for the TS networks are shown in Figure 2. EC values were based on the entire network and sorted greatest to least. Then nodes of the target country and nodes of the surrounding region were separated to determine which nodes of the surrounding region most represented activity in the target country and which nodes in the target country most represented activity in the surrounding region.

3.3 Top 5 human rights indicators over time and within region per country

The five most influential and most influenced human rights indicators over the time period are provided in the second and third columns, respectively, in Table 2. The fourth and fifth columns in Table 2, respectively, are the top five indicators within the surrounding region that best represent human rights activity in the target country and the top five indicators within the target country that best represent activity in the surrounding region. Variables in each cell are listed according to PR or EC from greatest to least, with Law in bold when applicable. Each variable is notated according to the categories established within the data documentation [1]. These symbolic notations are included at the bottom of the table.

The results of the lead-lag effect analysis with PR indicate that occupational and

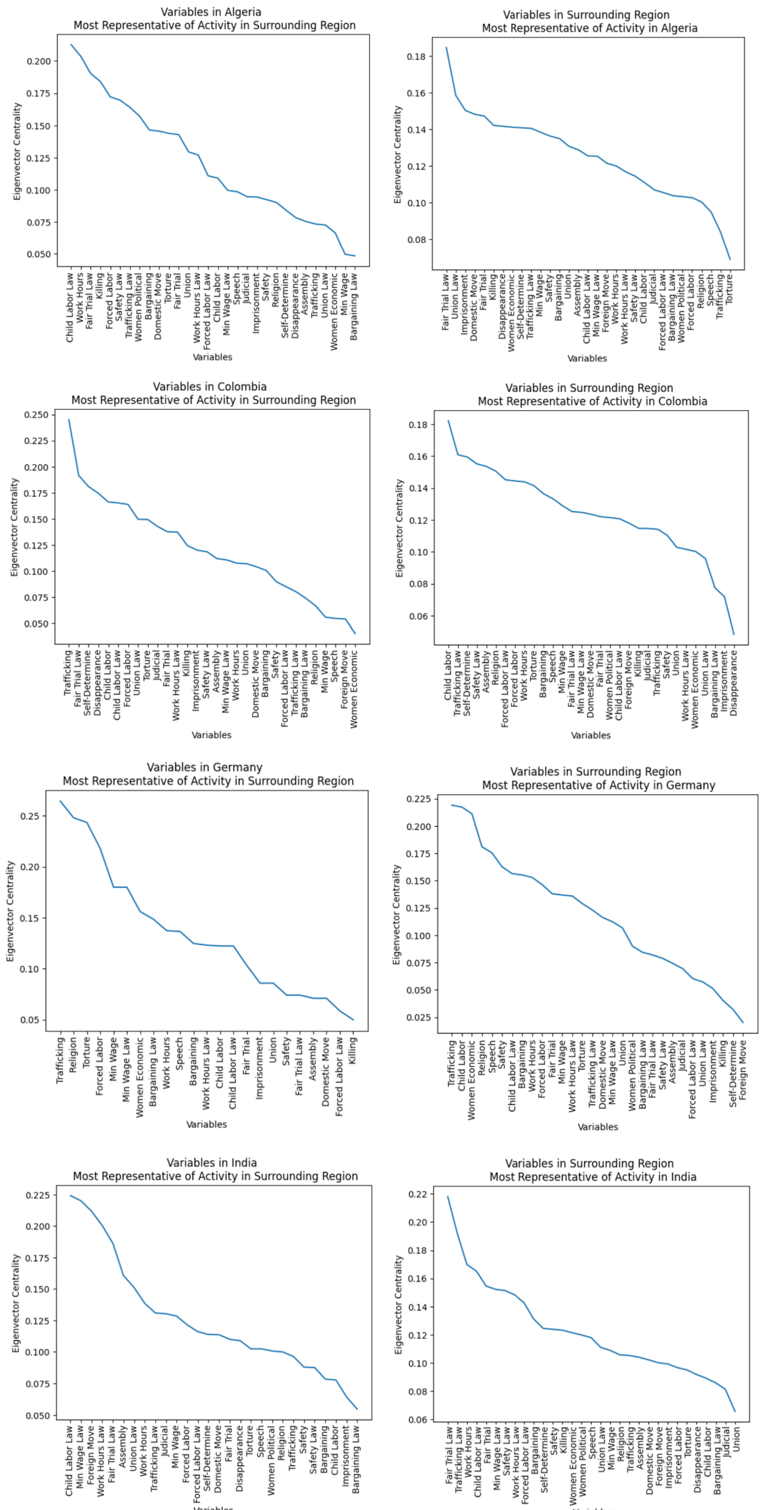

Figure 2: EC values (y-axis) per node (x-axis) for TS networks with edges based on mutual information. Nodes on the x-axis are sorted greatest to least. Nodes representing activity in the surrounding region are shown on the left and are nodes within the target country. Nodes representing activity in the target country are shown on the right and are nodes within the surrounding region.

Table 2: Results: LL and TS

	LL Network				**TS Network**			
Country	Most Influential		Most Influenced		Variables in Target Most Representative of Activity in Surrounding		Variables in Surrounding Most Representative of Activity in Target	
Algeria	Killing	$\circ$	Work Hours	$\star$	Child Labor **Law**	$\star$	Fair Trial **Law**	τ
	Fair Trial **Law**	τ	Trafficking	$\star$	Work Hours	$\star$	Union **Law**	$\star$
	Fair Trial	τ	Safety	$\star$	Fair Trial **Law**	τ	Imprisonment	$\circ$
	Work Hours	$\star$	Work Hours **Law**	$\star$	Killing	$\circ$	Domestic Move	$+$
	Child Labor **Law**	$\star$	Torture	$\circ$	Forced Labor	$\star$	Fair Trial	τ
Colombia	Self-Determine	$+$	Work Hours	$\star$	Trafficking	$\star$	Child Labor	$\star$
	Forced Labor	$\star$	Forced Labor	$\star$	Fair Trial **Law**	τ	Trafficking **Law**	$\star$
	Trafficking	$\star$	Bargaining	$\star$	Self-Determine	$+$	Self-Determine	$+$
	Union **Law**	$\star$	Safety	$\star$	Disappearance	$\circ$	Safety Law	$\star$
	Child Labor	$\star$	Fair Trial	τ	Child Labor	$\star$	Assembly	$+$
Germany	Forced Labor	$\star$	Forced Labor	$\star$	Trafficking	$\star$	Trafficking	$\star$
	Trafficking	$\star$	Work Hours **Law**	$\star$	Religion	$+$	Child Labor	$\star$
	Torture	$\circ$	Speech	$+$	Torture	$\circ$	Women Economic	$+$
	Religious	$+$	Bargaining	$\star$	Forced Labor	$\star$	Religious	$+$
	Min Wage **Law**	$\star$	Women Economic	$+$	Min Wage	$\star$	Speech	$+$
India	Fair Trial **Law**	τ	Religious	$+$	Child Labor **Law**	$\star$	Fair Trial **Law**	τ
	Child Labor **Law**	$\star$	Min Wage	$\star$	Min Wage **Law**	$\star$	Trafficking **Law**	$\star$
	Foreign Move	$+$	Forced Labor **Law**	$\star$	Foreign Move	$+$	Work Hours	$\star$
	Work Hours **Law**	$\star$	Assembly	$+$	Work Hours **Law**	$\star$	Child Labor **Law**	$\star$
	Min Wage **Law**	$\star$	Trafficking	$\star$	Fair Trial **Law**	τ	Fair Trial	τ
	$+$ **Empowerment Rights and Freedoms**				$\star$ **Occupational/Workers' Rights**			
	$\circ$ **Physical Integrity Rights**				τ **Justice Rights**			

worker rights predominantly influence future activity and are, in most countries, among the most influenced by past activity. Empowerment rights and freedoms are most prevalent in Germany and India, with each having one influential and two influenced. Physical integrity rights were less prevalent, indicating that they are not among the most influential or most influenced, nor do they have a relatively great relationship between the surrounding region and the target country. Algeria had one physical integrity right as influential and one as influenced, while the only other physical integrity right was influential in Germany. Laws tend to influence future activity, particularly in India. Justice rights influence future activity, except in Colombia, where a single justice right was influenced by past activity. A few human rights indicators exhibit a strong relationship both to past and future indicators, particularly work hours in Algeria and forced labor in Colombia and Germany.

Considering the results of the TS network with EC, occupational and worker rights are also quite predominant in representing activity in both the target country and the surrounding region, while laws represent activity in both the target and the surrounding region for all target countries except Germany. Imprisonment activity in Algeria is represented by surrounding activity, and disappearance and torture activity are represented in surrounding regions in Colombia and Germany, respectively. Several regional relationships are symmetric, representing activity in both the target country and in the surrounding region, such as fair trial law in Algeria, child labor and electoral self-determination in Colombia, trafficking in Germany, and laws for fair trial and child labor in India. Overall, Algeria and India both have laws influential over time and within region, though the laws in India are more prevalent over both time and region than those in Algeria.

4 Conclusion & Limitations

In general, key findings indicate that occupational and worker rights are the most influential and most influenced in the target countries over time and within region, and laws tend to influence future activity, indicating an attempt to protect human rights. Justice rights tend to be influential and represented within region. Physical integrity rights are less prevalent in the top 5 human rights indicators rankings, indicating that while they are related to other variables, they do not have the strongest relative relationships over time or region. Of the countries investigated, Germany had the most variables that were independent of other variables. Results did vary by target country, so it would be interesting to complete this analysis with countries neighboring each target country to compare results. However, results were not sensitive to variations in the value of the input parameter a, verifying the robustness of the method. Results were less impacted by parameter variations than those in the previous study since ordinal variables did not require density estimation when calculating relationship metrics.

These results are not predictive, though they do establish relationships over time for the countries investigated, and they establish geographical relationships between target countries and the surrounding region. The results of this rigorous study can help lawmakers and citizens establish policies and make responsible decisions obtaining to human rights.

Results were limited by data availability. While the data set is quite comprehensive and large, there were limited number of categories per variable, so some information may not be apparent, though there is much work to be done in this area. There are undoubtedly confounding factors that are not considered in this study. The relationship metrics used only incorporate the strength of the relationship, but they do not provide the direction of the relationship. We only considered pairwise relationships, so a model incorporating higher order relationships may lead to more accurate results, though perhaps at the sacrifice of interpretability. Future work includes creating and interpreting a network that includes both temporal and regional relationships in the same network. It would also be interesting to investigate results of a cluster analysis and a principal components or factor analysis.

References

[1] Mark, S., Cingranelli, D., Filippov, M., Richards, D.: The CIRIGHTS Data Project Scoring Manual V2.11.06.23 SSRN:https://ssrn.com/abstract=4625036 or https://dx.doi.org/10.2139/ssrn.4625036, (2023)

[2] Cingranelli, D., Filippov, M.: Problems of Model Specification and Improper Data Extrapolation. British Journal of Political Science **48**(1), 273-274 (2018)

[3] Cingranelli, D., Richards, D., Clay, K.: The CIRI human rights dataset. (2014)

[4] Barry, C., Cingranelli, D., Clay, K.: Labor rights in comparative perspective: The WorkR dataset. International Interactions **48**(2), 327-344 (2022)

[5] Fariss, C. J.: Respect for Human Rights has Improved Over Time: Modeling the Changing Standard of Accountability. The American Political Science Review **108**(2), 297–318 (2014)

[6] Cingranelli, D., Filippov, M., Fariss, C.J.: The Concept of Physical Integrity Rights : One or Two Dimensions? (2019)

[7] Cingranelli, D., Filippov, M.: Are Human Rights Practices Improving? American Political Science Review **112**(4), 1083–1089 (2018)

[8] Fariss, C.J.: Yes, Human Rights Practices Are Improving Over Time. American Political Science Review **113**(3), 868-881 (2019)

[9] Bell, S., Clay, K., Murdie, A.: Neighborhood watch: Spatial effects of human rights INGOs.The Journal of Politics **74**(2), 354-368 (2012)

[10] Cingranelli, D., Mark, S., Garvey, J. B., Hutt, J., Lee, Y.: A Brutality-Based Approach to Identifying State-Led Atrocities. Journal of Conflict Resolution **66**(9), 1676-1702 (2022)

[11] Greene KT, Park B, Colaresi M.: Machine Learning Human Rights and Wrongs: How the Successes and Failures of Supervised Learning Algorithms Can Inform the Debate About Information Effects. Political Analysis **27**(2), 223-230 (2019)

[12] Park, B., Greene, K., Colaresi, M.: Human rights are (increasingly) plural: Learning the changing taxonomy of human rights from large-scale text reveals information effects. American Political Science Review **114**(3), 888-910 (2020)

[13] Cingranelli, D., Mark, S., Gibney, M., Haschke, P., Wood, R., Arnon, D.: Human Rights Violations and Violent Internal Conflict. Social Sciences **8**(2), 41 (2019)

[14] Cingranelli, D., Mark, S., Sadykova-DuMond, A.: Democracy, Capacity, and the Implementation of Laws Protecting Human Rights. Laws **12**(1), 6 (2023)

[15] Richards, D., Webb, A., Clay, K.: Respect for physical-integrity rights in the twenty-first century: Evaluating Poe and Tate's model 20 years later. Journal of Human Rights **14**(3), 291-311 (2015)

[16] Fariss, C.J., Dancy, G.: Measuring the Impact of Human Rights: Conceptual and Methodological Debates. Annual Review of Law and Social Science **13**(1), 273-294 (2017)

[17] Goodrick, A., Sayama, H.,: Modeling and Analysis of the Lead-Lag Network of Economic Indicators. Data Science and Security: IDSCS 2023, Lecture Notes in Networks and Systems **922**, 213-226 (2024)

[18] Runge. J.: Causal network reconstruction from time series: From theoretical assumptions to practical estimation. Chaos **28**(7), (2018)

[19] Tank, A., Xiudi, L., Fox, E., Shojaie, A.: The Convex Mixture Distribution: Granger Causality for Categorical Time Series. SIAM Journal on Mathematics of Data Science **3**(1), 83-112 (2021)

[20] Shojaie, A., Fox, E.: Granger Causality: A Review and Recent Advances. Annual Review of Statistics and Its Application **9**(1), 289-319 (2022)

[21] Fish, J., Sun, J. & Bollt, E.: Interaction networks from discrete event data by Poisson multivariate mutual information estimation and information flow with applications from gene expression data. Applied Network Science **7**(70), (2022)

[22] Goswami, C., Merkley, A., Grover, P.: Computing Unique Information for Poisson and Multinomial Systems. arXiv preprint arXiv:2305.07013v1 (2023)

[23] Train In Data, https://www.blog.trainindata.com/mutual-information-with-python/. Last accessed 15 Dec 2023

[24] Bossomaier, T., Barnett, L., Harré, M., Lizier. J.: An Introduction to Transfer Entropy. Springer International Publishing, Switzerland (2016)

[25] Time Series Measures, https://elife-asu.github.io/PyInform/timeseries.html\#module-pyinform.transferentropy. Last accessed 15 Dec 2023