

Introducing geoSIMEX: A Generalized Approach To Modeling Spatial Imprecision

Daniel Runfola¹, Rob Marty¹, Seth Goodman¹, Michael LeFew¹ and
Ariel BenYishay²

¹Institute for the Theory and Practice of International Relations,
AidData, William and Mary

²Department of Economics, William and Mary

October 18, 2016

Corresponding Author:

Dr. Daniel Runfola

Institute for the Theory and Practice of International Relations, *AidData*
427 Scotland Street, Williamsburg, VA 23185

Email: dsmillerrunfol@wm.edu Telephone: 508.316.9109 Fax: 757.221.4650

1 Abstract

Spatial imprecision exists in nearly all spatial information - from census data to satellite information, the exact geographic location to which a measurement can be attributed is rarely known. Following this, there is a large and growing set of literature examining how different classes of models can integrate information on spatial imprecision in order to more accurately reflect available data. Here, we present a flexible approach - geoSIMEX - which can provide parameter and error estimates for both linear and non-linear models while adjusting for spatial imprecision. Further, this approach can provide an estimate of the potential value of additional spatial information by distinguishing between traditional model error and additional errors introduced by imprecision. We illustrate this approach through a case study leveraging a novel, publically available dataset recording the location of Chinese aid in Africa at varying levels of precision. Using a difference-in-difference modelling approach, we integrate Chinese aid information with satellite derived data on vegetation (NDVI) and a number of other spatially explicit covariates to examine if there is evidence Chinese aid has caused an increase or decrease in vegetation. Following commonly employed procedures which do not incorporate spatial imprecision, we found that Chinese Aid had a negative impact on vegetation; once spatial imprecision was incorporated into our estimates through the geoSIMEX procedure no evidence of impact was found. We use these findings to argue for the importance of incorporating spatial imprecision into analyses, and introduce new software in the R programming language to enable similar analyses.

Keywords: SIMEX, Spatial Uncertainty, Spatial Data Integration, Simulation, Ecological Fallacy

2 Software

Accompanying this analysis, we provide a new software package for researchers interested in adopting geoSimex in the R environment. This package is designed explicitly for a large body of researchers examining the impacts of international aid, but the main functions are generally adaptable to other use cases. All code is provided and made available under an open source license at <https://github.com/itpir/geoSIMEX>. Further, extensive documentation on the use of this software, as well as the ability to download datasets customized for use with this software, can be found at <http://geo.aiddata.org>.

3 Introduction

The lack of exact geographic information on where measurements are obtained presents a barrier to research. This has become increasingly evident as more scholars integrate geographic data from multiple sources - for example, census, satellite, and GPS sources - to try and establish causal or predictive relationships (c.f., Bare, Kauffman, and Miller 2015; Buntaine, Hamilton, and Millones 2015; Gallo and Goodchild 2012; Andam et al. 2008). Many scholars have engaged in research seeking to integrate information on the precision of geographic data to improve the accuracy of modeling efforts, variably referred to as spatial (im)precision and spatial uncertainty “Imprecision in Finite Resolution Spatial Data”; Aerts, Goodchild, and Heuvelink 2003; Ogryczak and Sliwinski 2009. This paper presents a generalizeable approach to integrating information on the precision of geographic data into both linear and non-linear models - geoSIMEX. We illustrate the capability of geoSIMEX to provide more accurate parameter estimates than traditional approaches through a simulation framework. Using a novel dataset, we then apply both traditional models and a geoSIMEX model to examine the causal impact of Chinese Aid on vegetation in Africa. We use this case study to illustrate the importance of including information on spatial imprecision into analyses. Finally, we provide all data and an accompanying R package for users seeking to incorporate information on spatial imprecision into analyses.

3.1 Literature

Past literature has shown that uncertainty in the locations of where measurements are taken can produce biased estimates in empirical analyses (Perez-Heydrich et al. 2013; Rettie and McLoughlin 1999). For example, Perez-Heydrich et al. show that regression coefficients will be biased when using raster data in conjunction with point data, where the true locations of the point data are only known to exist within some 5-10km radius of the measured location (Perez-Heydrich et al. 2013). One frequently cited "best practice" to overcome this challenge is to take average raster values within a buffer encompassing where the point could have fallen, instead of the single raster value associated with the point (Perez-Heydrich et al. 2013; Rettie and McLoughlin 1999). Another practice to address spatial uncertainty is to aggregate to some higher spatial scale where there is no spatial uncertainty (Perez-Heydrich et al. 2013). Yet another is to only use information for which exact (or, otherwise very precise) geographic information is known (Runfola and Napier 2016; Dreher et al. 2015).

There are many limitations, assumptions, and biases that these approaches incur. Most predominantly are the challenges highlighted by the large and well-established body of literature illustrating that analyzing data at different levels of aggregation can produce different regression coefficients and correlation coefficients (Clark and Avery 1976; SELVIN 1958; Gotway and Young 2002; Gehlke and Biehl 1934; Cramer 1964). As a simple example, using data from the 1930 US census, Robinson 2009 found a negative correlation between the proportion of immigrants in a state and average

literacy levels, but a positive correlation between being an immigrant and literacy level. A number of techniques have been proposed to address such biases, but most rely on additional assumptions (or, covariates) to aid in an effective dis-aggregation of data to finer scales (i.e., Gotway and Young 2002; Zhu et al. 2004; O’Loughlin 2000; Wong 2004).

The simulation and extrapolation method (SIMEX) provides one solution to address measurement error in covariates (Wang et al. 1998; KÃijchenhoff, Mwalili, and Lesaffre 2006; Li and Lin 2003), but has not previously been applied explicitly to spatial imprecision. In it’s traditional application, SIMEX leverages the relationship between increasing measurement error and bias following a two step process. First, SIMEX simulates additional measurement error to establish a relation between measurement error and covariate bias:

$$X(\lambda) = X + \sqrt{\lambda}U \quad (1)$$

where X is the measured covariate, U is the variance of the measurement error and λ is a parameter that simulates additional measurement error. $\lambda = 0$ corresponds to the original amount of measurement error in the measured covariate, X . SIMEX uses increasing values of λ (e.g., 0.5, 1, 1.5, and 2) to estimate models with simulated amounts of additional measurement error. Second, SIMEX estimates a trend between λ and the coefficient on X , and uses the trend to extrapolate back to $\lambda = -1$ (point of no measurement error).

4 Methods and Data

4.1 Methods

Two different approaches are followed. First, we use a monte carlo simulation procedure to examine the relative accuracy of geoSIMEX as contrasted to other procedures. Second, we apply geoSIMEX to a case study of the impact of Chinese Aid on environmental outcomes in Africa to provide an illustrative example of when the geoSIMEX approach might lead a researcher to a different conclusion than traditional approaches.

4.1.1 geoSIMEX

In this section, we detail the geoSIMEX approach using an illustrative example, in which we solve the following simplified equation:

$$NDVI = \theta * \text{Chinese aid} + \epsilon \quad (2)$$

in which Chinese aid is measured with spatial imprecision (due to, for example, limited documentation). Through using geoSIMEX we account for this spatial imprecision to accurately estimate the model coefficient θ .

In figure ??, we present a hypothetical country with sixteen districts for which we seek to solve equation 2. Four of these sixteen units of analysis, districts 5,6,7, and 8 are distinguished on the map. Within this study area, a hypothetical data set contains three geocoded Chinese aid project locations of various levels of spatial precision, projects A, B, and C. Project A is assigned a coordinate pair in District 5 and had strong documentation,

resulting in precise geographic information (i.e. an exact latitude and longitude). Due to weaker project documentation, location B has a precision level indicating that it was allocated anywhere in the region that includes districts 5,6,7, and 8. Project C has very uncertain spatial information, such that it may be anywhere in the country. The area (in square kilometers) of the unit of analysis in which each project location may have been allocated gives us the size of the project location’s area of coverage, which is summarized in table 2. Using the spatial overlap between each region of interest (the sixteen districts) and the area an aid project might exist in, we calculate a probability that each district contains a given project¹:

$$V_t = \sum_S^{s=1} U_s \left(\frac{a_{st}}{\sum_{t=1}^T a_{st}} \right) \quad (3)$$

These probabilities are used in the below steps of the geoSIMEX procedure to estimate model parameters while accounting for spatial imprecision.

The first step of geoSIMEX is to calculate the initial level of spatial imprecision in a given dataset, defined by λ . To reflect imprecision across a given set of aid projects and a set of units of analysis (i.e., districts), we calculate (λ) following:

$$\lambda = \frac{\sum_i^P \text{Area of Coverage}_i}{\sum_i^P \text{Total Possible Area of Coverage}_i} \quad (4)$$

¹Here, probabilities are only based on geographic overlap, as opposed to integrating other factors which might mediate where aid is allocated. This equation can be modified to incorporate more information on spatial location, thus allowing a researcher to trade off additional assumptions or information about factors that mediate spatial allocation in exchange for higher degrees of spatial precision (i.e., through dasymetric mapping approaches).

where i is an individual project out of P total Chinese Aid projects. *Area of Coverage _{i}* is project i 's known area of coverage defined by the available documentation - i.e., the geographic area across which a project could be located. *Total Possible Area of Coverage _{i}* is the area of coverage of project i under complete spatial imprecision - e.g., the geographic area of the study area.

If the latitude and longitude of every aid project was known, λ would resolve to 0—indicating zero spatial imprecision. If spatial data was only available for the entire study area (e.g., aid provided for general budget support without indication of where the project was allocated), λ would resolve to 1—indicating 100% spatial uncertainty. In practice, combinations of different levels of precision in the documentation of individual aid projects result in *lambda* values between these two extremes, resulting in a measurement of spatial imprecision between 0 and 1.

The second step involves estimating a naive model (which can be of variable functional forms; for illustration we use ordinary least squares regression) using the source data. For each unit of observation (i.e., district), the total dollars of aid is estimated by calculating the spatial overlap between the possible region aid could be in and the unit of observation (i.e., see table 2). In figure 2, we provide an example of the geoSIMEX procedure is applied to a dataset with an initial λ value of 0.4. In this figure, the x-axis represents the λ value (spatial imprecision) for a dataset, with higher values indicating more imprecision. The y-axis represents the estimated value of θ in equation ?? . In 2a, the orange line represents the 95% confidence interval of the coefficient on aid in the naive model fit in step 2. The horizontal black line represents the true model coefficient generated for this example

($\theta = 1$), which the naive model fails to capture.

In **the third step**, additional imprecision is simulated by randomly decreasing the precision of observations (in this case, Chinese Aid projects), and re-calculating the total dollars of aid for each unit of observation according to the updated areas-of-overlap between each unit and aid project (see table 2). For example, a project that has a measurement with an exact latitude and longitude will randomly be assigned a lower level of precision - i.e., a county, state, or even the entire country. Using these new, reduced levels of precision a model is fit in an identical fashion to step 1, and the estimated θ parameter, standard errors of the model, and λ value for a given permutation are saved. In figure 2b, the black points represent individual iterations, with the saved model coefficients (y axis) and their associated λ (x axis) values.

The fourth step subdivides this set of iterations into three equally-sized bins based on the level of spatial uncertainty (λ) of the aid variable (e.g., if λ values range from 0.4 to 1, coefficients are separated into bins of 0.4-0.6, 0.6-0.8, and 0.8-1). Average coefficient and λ values are calculated within each bin, represented as red dots in figure 2c. A quadratic trend is fit on the resulting average coefficient and lambda values. The trend is then extrapolated back to $\lambda = 0$, thus providing an estimate of θ with perfect spatial precision. In figure 2d, the red line represents the extrapolated trend, and the blue dot represents the extrapolated estimate of the coefficient on aid.

In the fifth step, the variance and standard errors of these estimates are calculated. We employ a bootstrapping method to calculate the component

of the standard error resulting from spatial imprecision. Here, a point from each bin is sampled, a quadratic trend is fit on the resulting values, and the trend is extrapolated back to $\lambda = 0$. This process is repeated iteratively to capture as many permutations as is computationally feasible (with the number of iterations defined as R). In figure 2e, each blue line represents one extrapolated trend and $\lambda = 0$ estimate. Following the approach outlined in Burnham and Anderson 2002, we explicitly quantify both the original standard errors and the additional error from spatial imprecision:

$$var(\hat{\theta}) = \sum_i^R \frac{1}{R} \{var(\hat{\theta}_i) + (\hat{\theta}_i - \hat{\hat{\theta}}_i)^2\} \quad (5)$$

where R is the number of extrapolated coefficients (in this example, 1000). $var(\hat{\theta}_i)$ is the standard error of each extrapolated coefficient, calculated by fitting a quadratic trend on the standard error estimates from each bin, extrapolating back to $\lambda = 0$, and collecting the resulting standard error value. $(\hat{\theta}_i - \hat{\hat{\theta}}_i)^2$ captures the remaining variance attributable to spatial uncertainty, based on the range of possible θ outcomes found in the extrapolation procedure (step 4).

4.1.2 Simulations

We employ a monte carlo simulation procedure to examine the accuracy of geoSIMEX compared to other approaches to handle spatial imprecision. Each simulation follows 7 steps:

1. One of three hypothetical countries with different administrative heirarchies is generated: (1) a country with 60 subcounties, 30 counties, 10

districts, and 2 regions, (2) a country with 120 subcounties, 40 counties, 20 districts, and 5 regions, or (3) a country with 120 subcounties, 60 counties, 30 districts, and 10 regions. Each subcounty is randomly assigned a spatial area and a probability of receiving aid.

2. 50 to 250 aid projects are randomly allocated to subcounties, according to the assigned probability of a subcounty receiving aid.
3. A simulated measurement of NDVI is generated according to equation 2 (defined as the number of aid projects plus random error).
4. Each aid project is given a code indicating the spatial precision that a researcher might see. Codes range from indicating the project fell within a sub-county (no spatial imprecision), to falling somewhere within a county, district, region, or the country.
5. The expected value of aid is calculated for each subcounty. For projects with spatial imprecision, aid is disbursed to subcounties based on spatial area, with larger subcounties having a larger probability of receiving aid.²
6. Equation 2 is estimated using models that do not incorporate spatial imprecision: the first uses the expected value of aid, and the second

²For example, consider a project where \$1 million was disbursed to either subcounty A (spatial area of 400km) or subcounty B (spatial area 100km). Because subcounty A is four times as large as subcounty B, subcounty A will be assigned four times as much aid from the project as subcounty B. Specifically, subcounty A will be assigned \$800,000 and subcounty B will be assigned \$200,000. Disbursing aid using spatial area assumes that each location (e.g., each sq. kilometer) within a country has an equal chance of receiving aid; this assumption could be modified according to dasymetric mapping approaches if desired by the researcher.

only includes aid projects with complete spatial precision (dropping out other information).

7. Equation 2 is estimated using models that seek to incorporate spatial imprecision: geoSIMEX and a model averaging approach. In the model averaging approach, subcounties are assigned a probability of receiving aid according to their spatial size, and the expected value of aid is calculated iteratively according to that probability. 500 models are estimated, where the average coefficient is used with the standard error estimated using equation 5. This approach is contrasted to the geoSIMEX approach.

This process is repeated 100,000 times.

4.1.3 Chinese Aid in Africa

In an illustrative case study of geoSIMEX, we examine the causal impact that Chinese aid distributed from 2005 to 2010 had on vegetation (measured using NDVI) in southeast asia. This study leverages a novel dataset on the location of Chinese international aid in Africa available at varying levels of precision (i.e., the exact location of each aid project is not always known). It integrates this information with a variety of other ancillary datasets, including the NASA Long Term Data Record (LTDR) to examine the causal impact of Chinese aid on vegetation. Employing geoSIMEX, we use this case study to illustrate the importance of incorporating information on spatial imprecision into analyses. A difference-in-difference modeling strategy is followed, in which the average of NDVI before aid was allocated (pre-2004)

is contrasted to the average of NDVI after aid was allocated (post-2011) for each of 351 districts:

$$Y_i = \beta_0 + \theta * Aid_i + \sum_{k=1}^n (\beta_k * X_i) \quad (6)$$

Where Y_i is the difference in the average forest loss post-2011 and pre-2004 in district i ; Aid_i is the number of aid projects in each district; X_i is a vector of control variables; β_k is a vector of parameters for each beta covariate, and θ is the estimated impact of aid. We calculate the initial lambda value for our dataset following 4, and simulate additional spatial imprecision by allocating aid to increasingly coarse units of administration. We repeat this analysis using three alternative modeling approaches to contrast their outcomes to geoSIMEX: (1) an ordinary least squares model in which imprecision is ignored (representative of a traditional case of the ecological fallacy), (2) a linear model in which all coarse-resolution data is removed from the dataset, and (3) a Monte Carlo model averaging approach in which hundreds of equally-probable worlds are generated based on the initial dataset, and the average beta coefficient and standard errors are reported.³

4.2 Data

In this analysis, we examine the impact of Chinese aid in southeast asia, explicitly focusing on second-level administrative units within Cambodia, Laos, Myanmar, Thailand and Vietnam. To conduct this analysis, we leverage a

³It should be noted there are many approaches to model averaging, including AIC weighting, that are not tested in this paper. It is possible some of these alternative approaches may outperform the Monte Carlo procedure presented here.

dataset on the location of Chinese aid is derived through a methodology designed to Track Underreported Financial Flows (TUFF; Strange et al. 2015). Covariate data is collected from a variety of sources, summarized in table 1. Our outcome measure - fluctuation in NDVI - is derived from the NASA Long Term Data Record (LTDR) dataset. While relatively coarse resolution, this dataset represents the longest consistent record of NDVI available at the global scale. To facilitate our difference-in-difference modeling efforts, we further select a number of covariates we believe could also impact shifts in NDVI (other than Chinese aid). These include:

1. Long-term climate data from the University of Delaware, providing precipitation and temperature data at a monthly time-step for the full data record, which is permuted to produce yearly mean, minimum, and maximum values for each project location.
2. Population Data is retrieved from CIESIN at Columbia University, specifically leveraging the Gridded Population of the World (GPW) data record.
3. Slope and Elevation data are derived from the Shuttle Radar Topography Mission (SRTM).
4. Distance to rivers is calculated based on the USGS Hydrosheds database.
5. Distance to roads is calculated based on the Global Roads Open Access Dataset (gRoads), which represents roads circa 2010, though the actual date of datasets is highly variable by country.

6. Urban travel time, calculated by the European Commission Joint Research Centre.
7. Nighttime Lights are retrieved from the NOAA Earth Observation Group, calculated from the Department of Defense Defense Meteorological Satellite Program (DMSP). Lights values are temporally inter-calibrated following the procedure outlined in Weng 2014

Each of these datasets are processed and aggregated according to their average values within each district included in this analysis. Further, the size of districts are controlled for to mitigate the challenge of variably-sized districts across the study area. In cases where covariates were measured at a resolution coarser than the unit of observation, the relative area of overlap was used to generate a weighted mean. Further information on this decision can be found in the discussion.

5 Results

5.1 Simulations

Table 3 shows results from the simulation analysis. Results are subset according to the spatial imprecision of the simulated aid data. The naive OLS model using the expected value of aid (i.e., ignoring spatial imprecision) captures true coefficient about 50% of the time within a 95% confidence interval when there is low spatial uncertainty ($\lambda < 0.3$), and the ability of OLS to capture the true coefficient declines as spatial imprecision increases. The naive model that omits aid projects measured with spatial uncertainty captures the true coefficient 60% to 70% of the time at a 95% confidence interval, depending on the level of spatial imprecision in the data. The model averaging approach performs well under low levels of spatial imprecision; however, as spatial imprecision increases the ability for the model averaging approach to capture the true coefficient decreases. Under low levels of spatial imprecision the model averaging approach captures the true coefficient 85% of the time, but under high levels of spatial imprecision the model captures the true coefficient 46% of the time.

The geoSIMEX model outperforms the model averaging approach in its ability to capture the true coefficient at all levels of spatial imprecision. At low levels of spatial imprecision the geoSIMEX model captures the true coefficient 90% of the time. Higher spatial imprecision leads to larger standard errors in the geoSIMEX model, reflected by the percentage of time the geoSIMEX model captures the true relation as spatial imprecision grows. At high levels of spatial imprecision ($\lambda > 0.7$), the geoSIMEX model captures

the true relation 100% of the time; however, large standard errors results in it capturing the true relation and statistical significance 0% of the time.

5.2 Chinese Aid in Southeast Asia

We examine the model results from our analysis of the impact of Chinese Aid in southeast Asia to provide an illustrative example of how accounting for spatial imprecision can lead to substantively different conclusions than alternative modeling approaches. Table 4 summarizes these findings. In (1) a linear model that ignores spatial imprecision (i.e., commits an ecological fallacy), (2) a linear model that omits imprecise spatial information, and (3) a Monte Carlo model averaging approach, findings suggest that Chinese Aid has contributed to deforestation. Specifically, an additional aid project within a district is associated with a 1.1 to 1.7 percentage point increase in forest loss ($p < 0.05$). However, when the spatial imprecision of the data is explicitly modeled (4), we find that insufficient evidence exists to determine if Chinese Aid has had a positive or negative impact on vegetation ($p > 0.1$).

Further, leveraging geoSIMEX we separately estimate the contribution of (1) spatial imprecision and (2) unexplained model variance to the standard errors estimated for θ . Here, we find that spatial imprecision contributes 98.5% of the variance around θ , while unexplained model variance contributes 1.5%.

6 Discussion

These results highlight the importance of incorporating spatial imprecision into analytic approaches which seek to establish causal relationships using spatial data. Under current "best practice" approaches, a researcher would find that Chinese aid has had a statistically significant, detrimental and causal impact on deforestation in southeast Asia, even after controlling for a wide variety of potential confounding variables. This result - which could have strong political and policy ramifications - is illustrated to have insufficient evidence to support it when the geoSIMEX procedure is followed, largely due to the relatively imprecise nature of the measurements of aid. Failing to account for spatial imprecision can thus bias results in key ways, and lead researchers to conclusions that may be inaccurate. Further, we find nearly all (98.5%) of the remaining variance can be ascribed to spatial imprecision; this offers compelling evidence that the true relationship between Chinese aid and deforestation could be better understood if more precise information was collected.

6.1 geoSimex

Broadly, this paper is related to papers that seek to overcome issues related to the Ecological Fallacy - i.e., the fallacy of measuring data at a coarse resolution and applying it to units of a finer resolution. By treating this "fallacy" as a source of explicit uncertainty in the modeling process, we argue that the use of relatively coarse resolution data not only can provide more insight into processes of interest, but that by ignoring coarse-scale in-

formation researchers may be committing a separate fallacy - that of ignoring "known unknowns". This can be particularly troublesome in analyses which seek to establish causal relationships between variables, as bias that results from spatial imprecision can lead to incorrect statements of the statistical significance of such relationships.

While geoSIMEX mitigates many issues related to spatial imprecision, it is not a silver bullet solution. First and foremost, the ability of geoSIMEX to uncover the true coefficient is heavily influenced by the initial precision of the observed data. If a researcher attempts to use data which has complete imprecision (i.e., $\lambda = 1$), the results provided by geoSIMEX will be meaningless as no information is available from which a trend can be estimated. As λ approaches 0 (perfect precision relative to the units of observation), geoSIMEX will provide increasingly accurate estimates as it has an increased number of λ units to derive data for - i.e., there are more variable simulations across observations of λ . This limitation of geoSIMEX is novel when compared to traditional SIMEX. In the traditional SIMEX process, an infinite amount of error can be simulated in any given variable by increasing the width of the error distribution. In geoSIMEX, the maximum amount of imprecision in the system is limited by the spatial configuration of the study area (i.e., the maximum imprecision is observed when all measurements are taken at the scale of the entire study area).

There are many directions for future work. First, the process described here should be generalizeable to multiple dimensions of spatial imprecision - i.e., imprecision in multiple attributes of the data. Such multidimensional approaches to SIMEX have been examined in other fields, but applying them

to the case of geographic imprecision should help to overcome many issues surrounding data integration alluded to in this piece. Second, the authors hypothesize that the SIMEX procedure could be applied not only to cases of imprecision in observed variables (the ecological fallacy), but may also provide a solution to the ever-present concern of arbitrary units of observation (the modifiable area unit problem, or MAUP). By treating the boundaries of units of observation as a unique cause of spatial imprecision (i.e., the exact - or correct - boundaries may not be known), geoSIMEX could be modified to provide estimates unbiased by covariates or boundary selection. However, both computational and methodological barriers remain to this solution. Finally, in non-hierarchical datasets the application of geoSIMEX is currently not feasible due to a lack of a relevant "weights" matrix to establish how data should be permuted at higher levels of λ . For example, in a raster dataset coarsening data can be done simply by aggregating nearby grid cells; conversely, if one sought to examine the impact of global international aid on non-contiguous protected areas, it is unclear how imprecisions should be introduced into the system.

7 Conclusion

In this piece, we introduce a flexible method to account for spatial imprecision in datasets for the purpose of fitting linear or non-linear models - geoSIMEX. We illustrated the relative accuracy of geoSIMEX as contrasted to other methods which do not account for spatial imprecision, finding that geoSIMEX outperforms all other tested approaches under cases of imprecision. Additionally, as an illustrative example we applied geoSIMEX to the case of Chinese Aid's impact on NDVI in southeast Asia. In models which did not account for spatial imprecision, it was found that Chinese Aid had a statistically significant, detrimental impact on forest cover. By accounting for spatial imprecision, we illustrated that there was insufficient evidence to draw this conclusion. We use these findings to argue for the importance of incorporating spatial imprecision into analyses, and introduce new software in the R programming language to enable similar analyses for other purposes.

8 Tables

Data Sources	
Data Name	Source
Chinese Aid Locations	AidData ⁴
Gridded Population of the World	Center for International Earth Science Information Network ⁵
Nighttime Lights	Defense Meteorological Satellite Program ⁶
Precipitation and Temperature	University of Delaware (Willmott and Matsuura 2001) ⁷
Urban Travel Time	European Commission Joint Research Centre ⁸
Distance to Rivers	World Wildlife Fund ⁹
Vegetation	NASA LTDR ¹⁰
Distance to Roads	CIESIN gRoads ¹¹

Table 1: Data sources used in this analysis.

⁴<http://china.aiddata.org>

⁵<http://sedac.ciesin.columbia.edu/data/collection/gpw-v3/sets/browse>

⁶Stable Lights retrieved from <http://ngdc.noaa.gov/eog/dmsp.html>

⁷Variables derived from these product included the average precipitation (P) and temperature (T) before a project was implemented (from 1992), the linear trend in P and T from 1992 to the project implementation, the average temperature from the date the project was implemented until the end of the temporal record(2012), and the post-project trend through 2012. Absolute measurements of each variable were also retained.

⁸<http://forobs.jrc.ec.europa.eu/products/gam/download.php>

⁹<http://hydrosheds.cr.usgs.gov/index.php>

¹⁰<http://ltdr.nascom.nasa.gov/cgi-bin/ltdr/ltdrPage.cgi>

¹¹<http://sedac.ciesin.columbia.edu/data/set/groads-global-roads-open-access-v1>

Project Location	Relative Precision	Known Aid Location (Geographic Size)
A	Very High	Populated Area in District 5 (3 km ²)
B	Moderate	Districts 5-8 (32,000 km ²)
C	Very Low	Entire Country (112,500 km ²)

Table 2: Example of the geographic area across which aid projects might be located given limited spatial information.

Table 3: Simulation Results

Spatial Uncertainty of Naive Model	Number of Simula- tions	Results Category	Naive Model	Naive Model (Sub- set)	Model Aver- age	geoSIMEX Model
$0 < \lambda < 0.3$	1262	Contains True Coef.	52.9%	62.6%	85.2%	86.9%
		Contains True Coef. & Sig.	52.9%	61.1%	84.9%	69.7%
$0.3 < \lambda < 0.7$	914	Contains True Coef.	44.4%	63.2%	75.9%	93.4%
		Contains True Coef. & Sig.	44.2%	57.5%	66.5%	25.9%
$0.7 < \lambda < 1$	678	Contains True Coef.	23.9%	72.4%	47.5%	100%
		Contains True Coef. & Sig.	23%	61.8%	17%	0%
$0 < \lambda < 1$	2854	Contains True Coef.	43.3%	65.1%	73.3%	92.1%
		Contains True Coef. & Sig.	43%	60.1%	62.9%	39.1%

“Contains True Coef” refers to the 95% confidence interval capturing the coefficient.

9 Figures

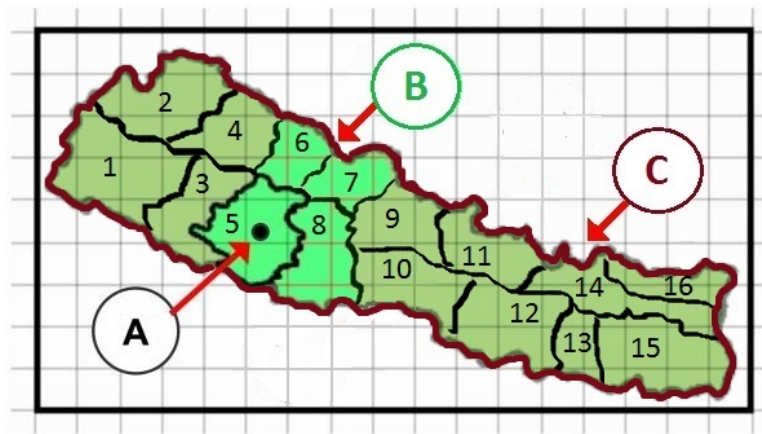


Figure 1: Hypothetical example of spatial imprecision in aid allocation.

Table 4: Impact of Infrastructure Aid on Forest Loss

	Forest Loss			
	Naive (1)	Naive (Low PC) (2)	Model Avg (3)	geoSIMEX (4)
Aid	1.745*** (0.523)	1.126** (0.553)	1.694*** (0.535)	2.218 (4.595)
Air Temp (Max)	0.107 (0.248)	0.116 (0.248)	0.107 (0.248)	0.099 (0.264)
Air Temp (Min)	1.037** (0.521)	0.997* (0.522)	1.037** (0.522)	1.014 (0.648)
Air Temp (Avg)	3.487*** (0.925)	3.572*** (0.926)	3.500*** (0.925)	3.498*** (1.247)
Precip (Max)	0.286*** (0.029)	0.287*** (0.029)	0.286*** (0.029)	0.287*** (0.030)
Precip (Min)	0.007** (0.003)	0.008*** (0.003)	0.008** (0.003)	0.007 (0.005)
Precip (Avg)	−0.032** (0.015)	−0.035** (0.015)	−0.033** (0.015)	−0.032 (0.020)
lnyx	0.0001 (0.0002)	0.0001 (0.0002)	0.0001 (0.0002)	0.0001 (0.0002)
Nighttime Lights	−0.021 (0.033)	−0.021 (0.033)	−0.021 (0.033)	−0.018 (0.034)
Constant	1.999*** (0.335)	2.035*** (0.335)	2.004*** (0.335)	1.989*** (0.389)
Observations	1,869	1,869	1,869	1,869

*Note:**p<0.1; **p<0.05; ***p<0.01
Standard errors in parentheses

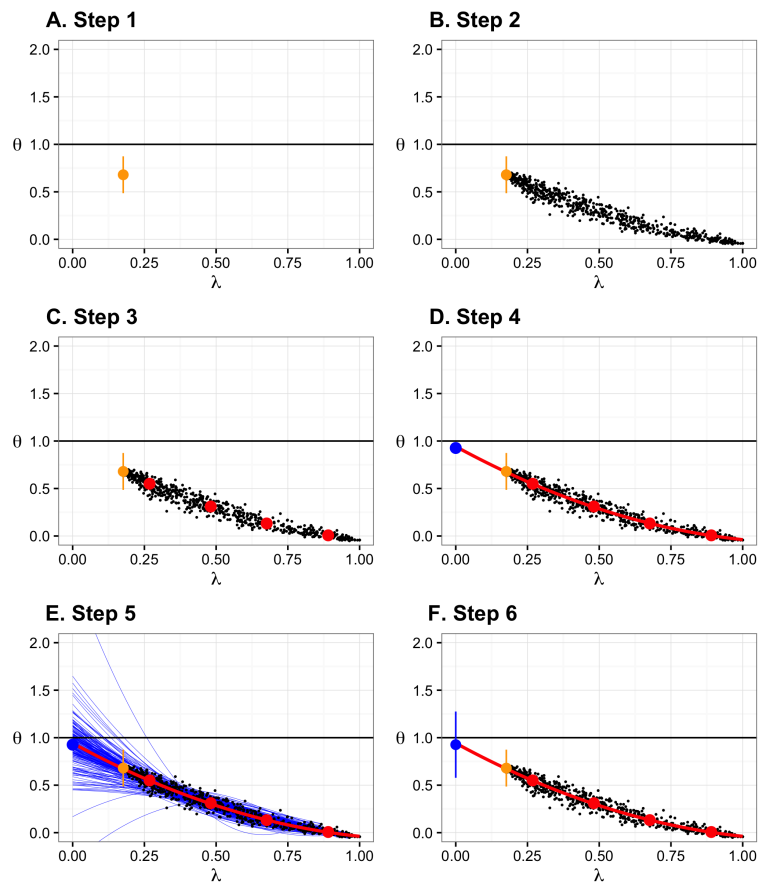


Figure 2: Steps of the geoSIMEX procedure.

Figure 3: Study area and units of observation.

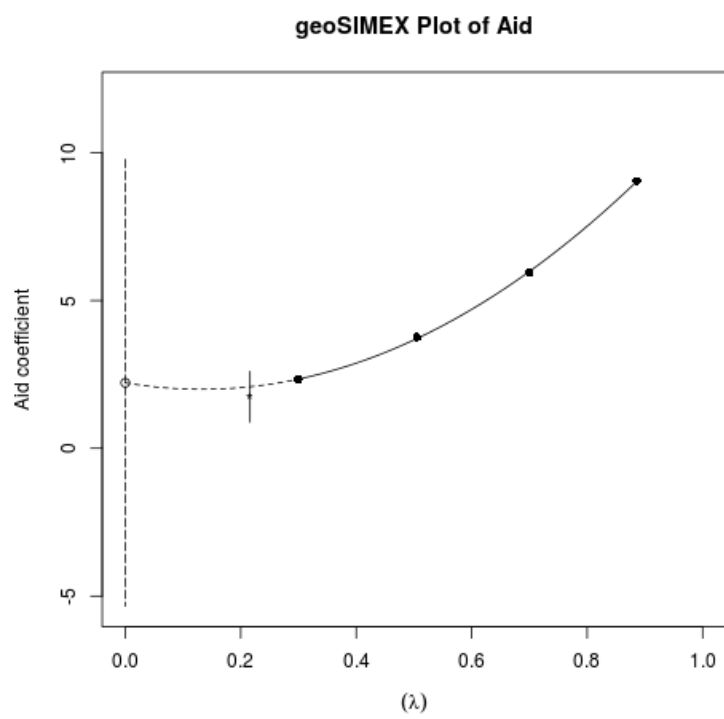


Figure 4: geoSIMEX Plot of Aid.

10 Acknowledgements

This work was performed in part using computational facilities at the College of William and Mary which were provided with the assistance of the National Science Foundation, Virginia Port Authority, Virginia's Commonwealth Technology Research Fund, and the Office of Naval Research. The authors would also like to thank Ben Dykstra and Miranda Lv for their valuable insights.

References

- Aerts, Jeroen C. J. H., Michael F. Goodchild, and Gerard B. M. Heuvelink (2003). “Accounting for Spatial Uncertainty in Optimization with Spatial Decision Support Systems”. In: *Transactions in GIS* 7.2, pp. 211–230. ISSN: 1467-9671. DOI: 10.1111/1467-9671.00141. URL: <http://onlinelibrary.wiley.com/doi/10.1111/1467-9671.00141/abstract> (visited on 07/24/2014).
- Andam, Kwaw S. et al. (2008). “Measuring the effectiveness of protected area networks in reducing deforestation”. In: *Proceedings of the National Academy of Sciences of the United States of America* 105.42, pp. 16089–16094. ISSN: 1091-6490. DOI: 10.1073/pnas.0800437105.
- Bare, Matthew, Craig Kauffman, and Daniel C. Miller (2015). “Assessing the impact of international conservation aid on deforestation in sub-Saharan Africa”. In: *Environmental Research Letters* 10.12, p. 125010. ISSN: 1748-9326. DOI: 10.1088/1748-9326/10/12/125010. URL: <http://stacks.iop.org/1748-9326/10/i=12/a=125010> (visited on 05/23/2016).
- Buntaine, Mark T., Stuart E. Hamilton, and Marco Millones (2015). “Titling community land to prevent deforestation: An evaluation of a best-case program in Morona-Santiago, Ecuador”. In: *Global Environmental Change* 33, pp. 32–43. ISSN: 0959-3780. DOI: 10.1016/j.gloenvcha.2015.04.001. URL: <http://www.sciencedirect.com/science/article/pii/S0959378015000503> (visited on 03/16/2016).
- Burnham, KP and DR Anderson (2002). “Information and likelihood theory: a basis for model selection and inference”. In: *Model selection and*

multimodel inference: a practical information-theoretic approach, pp. 49–97.

Clark, W. A. V. and Karen L. Avery (1976). “The Effects of Data Aggregation in Statistical Analysis”. In: *Geographical Analysis* 8.4, pp. 428–438. ISSN: 0016-7363. DOI: 10.1111/j.1538-4632.1976.tb00549.x. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMw3V3PS8MwFA7Ti4KIP3E6IScvo6Vp0qQ9eChT8ebBeS5Jk8wdNkUnuv_epE3abjIEj15LeKR5X9-P8r0vAOA4jIK1mEAjnfKEC0wSmzC1UohpqoRGMZe0lGujVv7m0vbZf3H8bcvSueNUSbpVqrsZ_fnnAs_oqah54TAPW5Jsbs5m2cyXuWkGWQ_YMdw05oYiWjAsAs7qg2LhOKVuJl24EE6MZCkc1yTyHsyuPvZa2GjJhp40xtgprq7C2Cmer-LqyMuozOS0X12oePD1umR49tZ06ehg1aTtm9V3a_vWdQq2jfw3Y5Wo1s5rMq0Q8PgD7rrWaeQ2JQ9BT8y0w4325PAbMQAM6aMAXDS00YAcacDqHHWhA7_kTML67HY_uA3dxRlCyLAoUKTMmqKJcaWS1Xq36ayYTK9yEkIgEiwnTGcZjwjXSVRGkqZYJBjzUy8ifAr2qOUx5BqPCWHDEsFCmAicCm-VUilgzKVTky6wPsD-I4rXWSSl-d0wfDPyZFe4Tei-sAJRZalqJ8z8ZvQC7La4HYHvx9qEuQe9z9g00s2ax.

Cramer, J. S. (1964). “Efficient Grouping, Regression and Correlation in Engel Curve Analysis”. In: *Journal of the American Statistical Association* 59.305, pp. 233–250. ISSN: 0162-1459. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMw1V3LTsMwEFxVnCoh3ohHK_kDCMQb03ZPCFWt0KNyjhW_EIhGpbT09_EjgQo0iA9IFNnJ7Gw80wNQ4HWe_cAELYwZaW4cjhgV2pZaa4dUGaasEVH8vjVqBV2oYxBZRpVgPNP3dKl-tTcYugSBt4u3LIRHhUPWNknDIzH1Ie_kLfijDTfXjsBnOSkmQz3xJXb7k0_vPhDmCv5ZjkLr0Uh9CzzRH0A61MrszHMJ1E5whfcEj8-eTr1xV5sE9JFNsQ1RgyDrkdSSlHnhsyCepXM14vPyzpnExOYDadzMb3WZuokL1ISrMRK30nkSnp1iI3BjFe9Ch5AMpjz4cUfl19CP1iKJVXYAHZWy7UdQm8z_wSas74h.

Dreher, Axel et al. (2015). *Aid on Demand: African Leaders and the Geography of China’s Foreign Assistance*. SSRN Scholarly Paper ID 2630152.

Rochester, NY: Social Science Research Network. URL: <https://papers.ssrn.com/abstract=2630152> (visited on 10/14/2016).

Gallo, John and Michael Goodchild (2012). "Mapping Uncertainty in Conservation Assessment as a Means Toward Improved Conservation Planning and Implementation". In: *Society & Natural Resources* 25.1, pp. 22–36. ISSN: 0894-1920. DOI: 10.1080/08941920.2011.578119. URL: <http://dx.doi.org/10.1080/08941920.2011.578119> (visited on 07/24/2014).

Gehlke, C. E. and Katherine Biehl (1934). "Certain Effects of Grouping Upon the Size of the Correlation Coefficient in Census Tract Material". In: *Journal of the American Statistical Association* 29.185, pp. 169–170. ISSN: 0162-1459. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMw1V3NS8MwFH8MTwPxW_wavIPXatNkTXeSURxePLmdS2IzEFxXXWXgX7_3kg63HRRvgUBIacj7te_3ASCTuzjauRMIVWeDMnudaqWcFKyHlKxMsCYu-8K6bakVrEMdmWtpWYK-p09wyb67-yTRVNb0Q_0RcXgUN1nbJA26iQUVuGCT-3Mfa58-Sdgm1YTq7zAPfR0ZHcJkrcYJ9JH1bINRvWM8dHcFBCyxxGE7CMXRcdQJdxpLBivkUTB7a_xgcixc4n6L_9UTVCyf1vEICg_jy9u14hsc5R3cEshyNnXeboCKFtEb0ERkLHLPIcP9N4w_yGdy0Hsf5U-T3XtTByaKgLwAmghftA8hz2DdMXfPqj_lr6BKGVIHWdQN7zeeX60Fn0VsB3VWtVQ.

Gotway, Carol A. and Linda J. Young (2002). "Combining Incompatible Spatial Data". In: *Journal of the American Statistical Association* 97.458, pp. 632–648. ISSN: 0162-1459. DOI: 10.1198/016214502760047140. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwnV1bS8MwFD5MfRmI92HVwUDIYjDH6CITYFpEl_c1LVjf9-cNNvqBRFFSiG9pD3pyXd0vn4HgGQXSfzFJ9DEcjmUiurDnDCtMSmz1CibZ11QoWqhDB2gsn6T23fi0xaX6J4RSjLMmv3t5jrCKFq62hpMYabGRZQnCU8enlWdmvg6lQz1Zn0YoT9qai4DmDz0Sd6qj7abEauVFC5F0Mh-3eNgtQcf_Psc0bAVc0rhuBtIudMxkD7oIRRs15304da5D-

XISA1ROGCMVW72YQYU771Tkmh7Aw-j2_uYuDiUW4jIrUh4TriwdGmIKmnFpqNWUUeMgAEegwOvN0dG4gKI
 VMF7y0Cd0yyBPSg02JVPxJ7X_Z04cwMIQomTKijANruXIhsKJWZZZpZbgswjOF29dvDWSGsKHIgUXaBpn
 h66vV-CzRCazX05npQ2c-_gCDnhRs.

KÃijchenhoff, Helmut, Samuel M. Mwalili, and Emmanuel Lesaffre (2006).

“A General Method for Dealing with Misclassification in Regression: The Misclassification SIMEX”. In: *Biometrics* 62.1, pp. 85–96. ISSN: 1541-0420. DOI: 10.1111/j.1541-0420.2005.00396.x. URL: <http://onlinelibrary.wiley.com/doi/10.1111/j.1541-0420.2005.00396.x/abstract> (visited on 10/14/2016).

Li, Yi and Xihong Lin (2003). “Functional Inference in Frailty Measurement Error Models for Clustered Survival Data Using the SIMEX Approach”. In: *Journal of the American Statistical Association* 98.461, pp. 191–203. ISSN: 0162-1459. DOI: 10.1198/016214503388619210. URL: <http://dx.doi.org/10.1198/016214503388619210> (visited on 10/14/2016).

Ogryczak, W. and T. Sliwinski (2009). “On efficient WOVA optimization for decision support under risk”. In: *INTERNATIONAL JOURNAL OF APPROXIMATE REASONING* 50.6, pp. 915–928. ISSN: 0888-613X. DOI: 10.1016/j.ijar.2009.02.010. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMw3V1Na9swGBbdTo0x7w-vHegwdgkujiTb8qGFtLT0lkNScjSyJA-HxS20S7f8-umVZMdJ0th51yQmivTkOf08vB8IUXIahXucYESpEfqaxEpqwmihMiqymAjNmblxbPemYalVlOPhpDUkalS10HC2mi8noztDCytdb2rRC5QfrjNYP96C_7TDcxqaZD8XqYbRwK12bu1-VEbsw9UGsbUi3D9f-aH7LjbAsu5hM1cZ8t242PbJ-Vo_VxE_LnkN20mYn8pBtM6Rc00ygJMazFjd-1A75NReMY1We0pClnu497bp-sx5eQw7NXH3nAbe7MMPytFqKxjcaNUjzSbE7jbQvZjfeK0-20gZgE0Mo4-_QV32lKtme6Tq8nTOzpp2DdY9m1_09DmNYnAdxP8iXXLnsWP1F7Mia_cv9SVFjBcz8DXrljw9PHGLEoiNdV00vu6ke2JP8e3Q-rXEPIAwAwkMAYQMg3AEIewBhCyAMAPqAbq-

v5pc3oZ-yEcoxj3go44IVojQ8zKgQiaKca1ISJsapoCWHfoRa6VhEGSkTmcSpjFQsaKwFmGdZOI_
opYBqjLq1VZvqM8JUqDSK1JIyEcbfRkVmbA0jkuiCFCVjARp1u5Xfu64qeZduuMxhb2E8apZHJDd7G6BPs
DU-kXAiIhSY3TINZQBWj8Lx-79C31oZVE--WvKz5GL7Z_pRP0vG0e9Fd09Lj6A1KeovA.

O’Loughlin, John (2000). “Can King’s Ecological Inference Method Answer
a Social Scientific Puzzle: Who Voted for the Nazi Party in Weimar Ger-
many?” In: *Annals of the Association of American Geographers* 90.3,
pp. 592–601. ISSN: 0004-5608. DOI: 10.1111/0004-5608.00213. URL:
<http://dx.doi.org/10.1111/0004-5608.00213> (visited on 10/14/2016).

Perez-Heydrich, Carolina et al. (2013). *Guidelines on the use of DHS GPS
Data: DHS Spatial Analysis Reports 8*. URL: [https://dhsprogram.com/
pubs/pdf/SAR8/SAR8.pdf](https://dhsprogram.com/pubs/pdf/SAR8/SAR8.pdf).

Rettie, WJ and PD McLoughlin (1999). “Overcoming radiotelemetry bias in
habitat-selection studies”. In: *CANADIAN JOURNAL OF ZOOLOGY-
REVUE CANADIENNE DE ZOOLOGIE* 77.8, pp. 1175–1184. ISSN:
0008-4301. DOI: 10.1139/cjz-77-8-1175. URL: [http://wm.summon.
serialssolutions.com/2.0.0/link/0/eLvHCXMwnV1NS8QwEB1WRRDEb7GuQg_
irdo2bdLiQWRx8eZFL15KmQSwwn64rSz6651J27WKiHhum5RMMvOSvHkDwMIL3_
vmE6LAL5goYtKnCiT3i1Cj7Q1nMi1kUZivqVbQqkIO1m6dpPXceqro0PySMkQRDghxPXvxqIoU3bY2JTVW_
TMXAR1SoqfeBFR6sAsY9EnwPwhBNlwM9yua6qWVqWQWCaLcYd43VFw_0-P78BWA0Tdm3rm7ELPTPZg_
Wlqj9n34eoe5zi2h5HNNUs9mlaGi0bV_M3NR7J0Rx0XJL4Rqnq1LaWD9nXLmpR4AI_
D24fBndcUWvAUBvzASwi0SMH91JFqMdOI4bhQEvdmeZTmvEh1qBKjQh3qVMfGRBJxWxrWGphZME0YVMSI.
VKITjhWopEYlw2Dpy3Y5_NamGNzG5IWJqp5_dMiCwhJfLYgX47jJnM6fxHVWVGG6kYfUv3abP6ymw5wg6c.
mp0obcYfwC2kvED](http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwnV1NS8QwEB1WRRDEb7GuQg_irdo2bdLiQWRx8eZFL15KmQSwwn64rSz6651J27WKiHhum5RMMvOSvHkDwMIL3_vmE6LAL5goYtKnCiT3i1Cj7Q1nMi1kUZivqVbQqkIO1m6dpPXceqro0PySMkQRDghxPXvxqIoU3bY2JTVW_TMXAR1SoqfeBFR6sAsY9EnwPwhBNlwM9yua6qWVqWQWCaLcYd43VFw_0-P78BWA0Tdm3rm7ELPTPZg_Wlqj9n34eoe5zi2h5HNNUs9mlaGi0bV_M3NR7J0Rx0XJL4Rqnq1LaWD9nXLmpR4AI_D24fBndcUWvAUBvzASwi0SMH91JFqMdOI4bhQEvdmeZTmvEh1qBKjQh3qVMfGRBJxWxrWGphZME0YVMSI.VKITjhWopEYlw2Dpy3Y5_NamGNzG5IWJqp5_dMiCwhJfLYgX47jJnM6fxHVWVGG6kYfUv3abP6ymw5wg6c.mp0obcYfwC2kvED).

Robinson, WS (2009). “Ecological Correlations and the Behavior of Indi-
viduals”. In: *International journal of epidemiology* 38.2, pp. 337–341.

ISSN: 0300-5771. DOI: 10.1093/ije/dyn357. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwjV1LS8NAEB6KBxGKj_qKVVMw19gku9kkRwktehBBFLyFfYJCo7SV4r93t5vURLR6zmZJdjYz32Tm-xYAR5eB_80n0ECnLGyck9gGTK1UmGiquA4jJqmQbaoVDH4p6Gd4-PyihvKjxPGSQU6wTbju755W1QNCH0TFQeDHSRLDLH0PR0qu0dyCvyTuu22QxaTRgN5Qc177ALmxXqBNduW2yBx1V9mDztqqr96Dr_t4hR0rah2QkaqeIcnt-R9Uxh1gpkYGMqJJVnKJXjW5WnK7ZATyORw_5tV8dseALK9HpYy0YVSKW3MBGbfBhC6zrfjlfEnZk8eAFMachQnmyoA1wk0KzKnmkU4kVyKtMqC9eIXb05So3C1cFyYxSnc4nhwbuyydkC_N1lRfXezIrLpp_GboQdHzopfM2TWfXHqwaBp1tV1GxsoST07c4yr8yD8z7C8U1K3CgLzk7-euQ9bdTUqCE9hYz59V2fQWUw-Afbk7-M.

Runfolo, Daniel Miller and Ashley Napier (2016). “Migration, climate, and international aid: examining evidence of satellite, aid, and micro-census data”. In: *Migration and Development* 5.2, pp. 275–292. ISSN: 2163-2324. DOI: 10.1080/21632324.2015.1022969. URL: <http://dx.doi.org/10.1080/21632324.2015.1022969> (visited on 10/14/2016).

SELVIN, HANAN C. (1958). “Durkheim’s Suicide and Problems of Empirical Research”. In: *American Journal of Sociology* 63.5, p. 607. ISSN: 0002-9602. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMw3V1dS8MwFL1MnwQRP9E5ISDMp0rbrE334EOZlQkyRKevo2kSLNjN1g7x39s0TVoV-gN8baEp0XDuvck95wJg99q2fnGcb4sg9mKKR54MmIJzhwifU-G4MfMT9lNqZayfm2f_Afjbzfr9jcschySQveSFNUmY0AXJ2TNW7EWufqTIH0a137SzVX00001WjbDFnMtHD633VJDANZ-GsPm5lS1HnBU27XsuMtqsBRBNnWewo4uSaK4nlYTXtU5NpzVZp-7q6YkbfJkPpZp6xNClu-NJ6ed4qS-VAFsxPk8eW5Zft6pJFLvgnRlZxbL4Pe_UWoFDt9AH0-PIQ-mY70BApiTNSjitfR-BpCK5yVA0ASgCQBgcTBDIAIA3AMVzeRfPJ1NI_sqjhzhdlb9LC0GP4BPYjaUoYV1U4kV2CohjTGOHYMrLAb3P7BvEYB4u.

- Strange, Austin M. et al. (2015). “Tracking Underreported Financial Flows China’s Development Finance and the Aid’s Conflict Nexus Revisited”. In: *Journal of Conflict Resolution*, p. 0022002715604363. ISSN: 0022-0027, 1552-8766. DOI: 10.1177/0022002715604363. URL: <http://jcr.sagepub.com/content/early/2015/09/18/0022002715604363> (visited on 10/18/2016).
- Wang, Naisyin et al. (1998). “Bias Analysis and SIMEX Approach in Generalized Linear Mixed Measurement Error Models”. In: *Journal of the American Statistical Association* 93.441, pp. 249–261. ISSN: 0162-1459. DOI: 10.1080/01621459.1998.10474106. URL: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1998.10474106> (visited on 10/14/2016).
- Weng, Qihao (2014). *Global Urban Monitoring and Assessment through Earth Observation*. CRC Press. 412 pp. ISBN: 978-1-4665-6450-3.
- Willmott, C.J. and K Matsuura (2001). *Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950-1999)*. URL: http://climate.geog.udel.edu/~climate/html_pages/README_ghcn_ts2.html.
- Wong, David W. S. (2004). “The Modifiable Areal Unit Problem (MAUP)”. In: *WorldMinds: Geographical Perspectives on 100 Problems*. Ed. by Donald G. Janelle, Barney Warf, and Kathy Hansen. DOI: 10.1007/978-1-4020-2352-1_93. Springer Netherlands, pp. 571–575. ISBN: 978-1-4020-1613-4 978-1-4020-2352-1. URL: http://link.springer.com/chapter/10.1007/978-1-4020-2352-1_93 (visited on 10/14/2016).
- Worboys, Michael. “Imprecision in Finite Resolution Spatial Data”. In: *GeoInformatica* 2.3, pp. 257–279. ISSN: 1384-6175, 1573-7624. DOI: 10.1023/A:

1009769705164. URL: <http://link.springer.com/article/10.1023/A:1009769705164> (visited on 07/06/2016).

Zhu, Jun et al. (2004). “Combined mapping of soil properties using a multi-scale tree-structured spatial model”. In: *Geoderma* 118.3, pp. 321–334.

ISSN: 0016-7061. DOI: 10.1016/S0016-7061(03)00217-9. URL: http://wm.summon.serialssolutions.com/2.0.0/link/0/eLvHCXMwpV1LT8MwDLYG4AQb8R4SDlwgEOhjyHzvNxkAgEFx6iJqktS3bST5_Acjs8zj65BMoarqsk2iVWlu1FSplM40G0h0ZFMavGxdKrerSmACyDJGg9v1icwp0JHRcpqtlqDJ5kuG3ry8ub_tzY8W8jhwNSYy4g7vZT31IL7xNM70_DiR-i5gfRGnfEzqb4CeVfbUUJTX8QI6e4Hm8R9_twnrIWEV3fq9LWjg4zasdofPgbQDd2BAToUW2GjFWDPdw1A1PzNnq2CA_VBo4RGM0ZRMawWfiEc1gy0NZMWU4d00mb-eZxcG_au73nUUrmuIKkobswhV7BKXt6WrnHTSx0LL_-w-CM3sQGidS51tG0lTkPoq107jqkql0y04nymnnNT0HOU7bo2fLLUyzkovt1K1oJipsPwg-5JCyE9dTxZVPp-P02RapyhZ-IS6BclvXusFCnamHng5-PtXHcJKDSbiXaEjWCad4TEOXsdvQPELRw