Examination of the geographical variation in human West Nile virus: a spatial filtering approach

J. TEVIE[1]*, A. BOHARA[2] AND R. B. VALDEZ[3]

[1] *Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, Baltimore, MD, USA*
[2] *Department of Economics, University of New Mexico, Albuquerque, NM, USA*
[3] *Departments of Family and Community Health and Economics, University of New Mexico, Albuquerque, NM, USA*

SUMMARY

This paper examines the importance of environmental factors (mosquito pools and home foreclosures) in human West Nile virus (WNV) transmission in California and Colorado. The role of environmental factors is investigated by applying an instrumental variable technique to a spatial filtering random-effects negative binomial model to correct for both spatial autocorrelation and endogeneity. The results suggest that mosquito pools and home foreclosures are significant in explaining the prevalence of human WNV. An innovative aspect of this research is that it emphasizes the role of home foreclosures in WNV transmission and in the allocation of resources. Knowledge of the factors associated with WNV prevalence is crucial for abatement of future outbreaks. The results suggest that more resources should be allocated to areas that have a high number of home foreclosures and mosquito pools for surveillance and mitigation of the disease.

Key words: Geographical variation, spatial autocorrelation, spatial filtering, spatial modelling, West Nile virus.

INTRODUCTION

Vector-borne diseases are influenced by environmental conditions in all countries not just some. Between 1999 and 2011, more than 30 000 humans in the USA were infected with West Nile virus (WNV) and 400 of these cases resulted in death [1]. California and Colorado recorded some of the highest cases of WNV in the USA between 2003 and 2007 [1]. During this time period, these two states accounted for about 28·2% of all WNV cases and 20% of all fatalities [1]. In fact, these states have consistently ranked either first or second in terms of WNV infections during the time period under study [1]. The transmission cycle of WNV involves three organisms: a vector (mosquitoes), host (birds) and infectious agent (virus). Humans and animals are considered to be accidental or dead-end hosts and develop very low levels of viraemia which is not sufficient to infect mosquitoes [2]. The geographical distribution of the disease is determined by certain socioeconomic, climatic, anthropogenic and environmental factors [3]. It is reasonable to argue that the presence of the *Culex* genus is necessary for WNV infection, but not sufficient for an outbreak or spread of the disease.

* Author for correspondence: Dr J. Tevie, Department of Pharmaceutical Health Services Research, University of Maryland School of Pharmacy, 220 Arch Street, Baltimore, MD 21201, USA.
(Email: jtevie@rx.umaryland.edu)

The amplification of the disease is determined by the manner in which these factors impact the transmission cycle. It is intriguing why the incidence of WNV is higher in these states. These states are quite diverse in terms of geographical, climatic, demographic, socioeconomic and environmental factors so they merit a separate investigation as to the determinants of WNV incidence.

A number of WNV studies conducted in California [4, 5] focused on the role of economic conditions (*per capita* income and mortgage delinquencies) in the transmission of WNV. These studies, however, were limited to only two counties in California, i.e. Orange and Kern counties. Moreover, recent studies [6–9] which examined the role of environmental factors such as dead birds and mosquitoes in WNV transmission in Colorado and California were restricted to a few counties. A critical review of the literature reveals that these studies did not adequately address some or all of estimation issues such as spatial autocorrelation, endogeneity, and the panel nature of the data. Albeit some studies [10, 11] have used negative binomial models to address some of these problems, they failed to adequately address the panel structure of the data and endogeneity issues. The failure of prior studies to recognize the possibility that WNV counts in one county could be dependent on WNV counts in adjacent counties could be costly. Insofar as this is true, omitting such information may lead to biased estimates of the impact of county-level factors on WNV prevalence.

The spatial statistics literature has not evolved enough in terms of its application to panel-count data models. The motivation for this study is to fill such a void. Two competing methodologies available in the literature that can be employed to correct spatial autocorrelation in nonlinear models such as generalized linear models (GLM) are the spatial lag method [12] and the spatial filtering technique [13]. The spatial lag approach is a parametric method that incorporates the structure of spatial autocorrelation into the model to rectify the problem and is motivated by the spatial lag specification for linear models. It includes a spatially lagged dependent variable as an additional covariate in the nonlinear model to correct the problem. The spatial filtering approach is a non-parametric method used to correct for spatial autocorrelation. It does not assume any prior structure of the nature of spatial autocorrelation, but incorporates selected eigenvectors into the model to filter out spatial autocorrelation. It is the approach pursued in this paper.

This study, to the best of our knowledge, is one of the first empirical WNV research studies in California and Colorado to employ a spatial filtering technique.

Home foreclosures were considered a potential risk factor in the transmission of human WNV during the housing crisis that began in 2004 and culminated in the financial crisis of 2007 [4, 5]. The Federal Reserve [14] stated that the hardest hit states were California, Arizona, Nevada, Colorado, Florida and some New England states. The argument was that the economic downturn and accompanying housing market crisis adversely affected the economy of the USA in general, and particularly the states of California and Colorado. The combined effect was a growing number of neglected swimming pools, particularly across Southern California. This was attributed to the high number of home foreclosures because home owners could not afford their mortgages. Most of these neglected swimming pools on foreclosed homes collected small pockets of water and served as breeding grounds for mosquitoes [4, 5]. According to Reisen *et al.* [4], WNV cases escalated by 276% in Kern county in the summer of 2007. From the foregoing discussion, it is reasonable to hypothesize a positive relationship between home foreclosures and the transmission of WNV.

The general consensus is that the presence of the *Culex* mosquito vector is necessary for the outbreak of WNV [11]. In western USA, *C. tarsalis* and *C. quinquefasciatus* species are found in abundance [15]. *C. tarsalis* is found in both Colorado and California, but *C. quinquefasciatus* is mostly restricted to California. *C. tarsalis* is the predominant vector in rural settings [16], although its presence has been reported in urban areas of California [4]. Their habitat includes areas characterized by standing water such as irrigated fields, poor drainage, sewage treatment lagoons/pools, urban catch basins, and containers found on the compounds of many low-cost houses [17, 18]. The presence of *C. pipiens* has also been reported in Southern California [7]. They are mostly found in urban areas and lay their eggs in stagnant water [18]. Mosquitoes become infected by feeding on a bird with the virus in its bloodstream. Mosquitoes then spread the virus to new hosts by biting another bird or person. The biological evidence from the foregoing discussion suggests that the number of mosquito pools is positively related to human WNV.

The primary objective of this research is to investigate the significance of environmental (mosquito pools

Table 1. *Factors used to assess risk of human WNV in California*

Variable	Description	Mean	S.D.	Min.	Max.
HV	Human WNV case counts	8·355	29·530	0	331
INCOME	Natural log of median county income	10·750	0·246	10·258	11·341
FORCLOSE	Natural log of home foreclosures plus 1	4·004	3·668	0	10·603
POPDENSE	Natural log of population density	4·508	1·993	0·412	9·73
MOSQUITO	Number of mosquito pools or field samples that tested positive for WNV	14·679	43·901	0	408
PRECIP	Mean annual precipitation in inches	2·184	1·015	.270	4·283
TEMP	Mean annual temperature in Fahrenheit	57·912	4·029	44·733	66·4
PDSI	Palmer Drought Severity Index (range from 0 to −5)	−0·241	2·053	−5·040	−4·215

S.D., standard deviation; Min., minimum; Max., maximum.

Table 2. *Factors used to assess risk of human WNV in Colorado*

Variable	Description	Mean	S.D.	Min.	Max.
HV	Human WNV case counts	13·365	51·777	0	546
INCOME	Natural log of median county income	10·637	0·288	9·986	11·458
FORCLOSE	Natural log of home foreclosures plus 1	2·903	2·669	0	8·975
POPDENSE	Natural log of population density	2·509	1·797	−0·368	8·227
MOSQUITO	Number of mosquito pools or field samples that tested positive for WNV	5·593	21·051	0	247
PRECIP	Mean annual precipitation in inches	1·337	.185	.803	1·62
TEMP	Mean annual temperature in Fahrenheit	46·816	2·760	40·225	51·666
PDSI	Palmer Drought Severity Index (range from 0 to −5)	−0·772	2·204	−5·705	−3·005

S.D., Standard deviation; Min., minimum; Max., maximum

and home foreclosures) factors in human WNV transmission in California and Colorado. This study contributes to the literature in a number of ways. First, it employs a spatial filtering random-effects negative binomial model to study the importance of home foreclosures in the transmission of WNV; many studies have focused on the significance of climatic, geographical factors and other environmental factors. Second, it addresses the issue of spatial autocorrelation in the dependent variable within a panel-count data model context. The presence of spatial autocorrelation in the dependent variable will cause estimates of the variance to be biased if not corrected. Third, it applies an instrumental variable technique to the spatial filtering random-effects negative binomial models to correct for endogeneity in income and home foreclosures.

MATERIALS AND METHODS

Data sources

Data used in this study were collected from several sources. Data on temperature (TEMP), precipitation (PRECIP) and the drought index (PDSI) were collated from NOAA [19]. County-level climatic data was not readily available so information from weather stations in each county was used to calculate annual climate data. An arithmetic average was used to calculate the climatic variables in counties with several weather stations. Information on human WNV (HV) and the number of mosquito pools (MOSQUITO) were acquired from the CDC [1]. Data on income (INCOME) and population density (POPDENSE) were taken from the US Census Bureau [20]. Data on home foreclosures (FORCLOSE) were acquired from Data Quick News [21] and the Colorado Department of Local Affairs [22]. A detailed description of the variables used this study is provided in Tables 1 and 2.

Empirical model specification

The following random effects negative binomial model (RENB) is specified to investigate the determinants of human WNV:

$$HV_{it} = \exp(\beta_0 + \beta_1 FORCLOSE_{it} + \beta_2 MOSQUITO_{it} + Z'_{it}\theta + u_i), \qquad (1)$$

where i (1, ..., N) indexes counties, t (1, ..., T) indexes time, and u_i is the random-effects term. The choice of controls is driven by a mix of theory and empirical findings. The vector Z' comprises control variables such as annual precipitation (PRECIP), annual temperature (TEMP), annual drought (PDSI), income (INCOME), population density (POPDENSE) and time dummies for 2004–2007 (D2004, D2005, D2006 and D2007).

In order to further verify and clarify the role of home foreclosures on WNV transmission, the variable FORCLOSE is included in the model. The variable MOSQUITO is included to control for vector abundance. It should be mentioned that the environmental factors identified in this research are partly caused by economic conditions. In that regard, economic conditions could exacerbate the transmission of WNV.

Endogeneity

The model specified in equation (1) is appropriate only if all the explanatory variables are exogenous. This is potentially a problem due to the possibility of a simultaneity bias between human WNV on the one hand and income and home foreclosures on the other. In other words, human WNV, income and home foreclosures may be jointly determined. To correct for this endogeneity problem, an instrumental variable method is applied to the RENB. To implement this method the contaminated variables (income and home foreclosures) are regressed on an array of instruments and the predicted values INCHAT and FORHAT are used to replace income and home foreclosures in equations (1) and (2).

Spatial filtering technique

This technique is based on the spatial eigenvector mapping method. The rationale behind this method is that the configuration of spatial data points on a map, are reflected in covariates that capture spatial effects at different spatial scales. This is a nonparametric method for correcting spatial autocorrelation [13]. Spatial filtering techniques have been developed and implemented by several authors [23–26]. The spatial filtering technique employed in this paper is according to Griffith [27]. It is based on the eigenvector decomposition of a modified spatial weights matrix. The eigenvectors of this matrix are judiciously selected based on a stepwise selection method to filter out spatial autocorrelation from the model residuals. The selected eigenvectors are then

Table 3. *Global spatial autocorrelation test on human WNV counts*

Moran's I statistic	Z value	P value
Moran's I test for California		
0·350	8·552	0·000
Moran's I test for Colorado		
0·537	14·848	0·000

used as components of the spatial filter and included as covariates in the GLM regression. This transformed spatial weights matrix reflects the latent spatial autocorrelation inherent in the dependent variable. The spatial filtering RENB model is specified as follows:

$$HV_{it} = \exp(\beta_0 + \beta_1 FORHAT_{it} + \beta_2 MOSQUITO_{it}$$
$$+ SFILTER_{i\gamma} + Z'_{it}\theta + u_i) \qquad (2)$$

where $SFILTER_i$ is an array of selected eigenvectors (spatial filter components), denoted by VEC, of the transformed spatial weights matrix. The spatial filter can be perceived as a proxy for omitted or missing variables from the regression [13]. Getis & Griffith [25] contend that all the n eigenvectors extracted represent all the possible orthogonal map patterns. In other words, they represent a kaleidoscope of all possible map patterns. Specifically the first two principal eigenvectors extracted are often associated with North–South and East–West patterns, respectively. Eigenvectors with intermediate values of Moran's I typically exhibit regional patterns, while eigenvectors with extremely low values of Moran's I are associated with local map patterns.

ESTIMATION AND RESULTS

Statement of hypotheses

The following hypotheses regarding the environmental variables are tested.

Hypothesis 1: $\beta_{FORCLOSE} > 0$

This hypothesis states that home foreclosures will have a positive effect on the prevalence of WNV. Under this hypothesis the regression coefficient on FORCLOSE is expected to be positive.

Hypothesis 2: $\beta_{MOSQUITO} > 0$

This hypothesis states that the number of mosquito pools will have a positive effect on the prevalence

Table 4. *Instrumental variable spatial filtering random-effects negative binomial regression for California*

	Model 1	Model 2	Model 3	Model 4
FORHAT[a]	0·406 (2·16)**	0·382 (2·39)**	0·413 (2·33)**	0·387 (2·52)**
INCHAT[b]		−1·206 (1·69)*	−1·234 (1·70)*	−1·012 (0·95)
MOSQUITO	0·011 (2·94)***	0·005 (2·28)**	0·005 (1·81)*	0·005 (2·36)**
POPDENSE	0·008 (0·05)	0·022 (0·11)	0·005 (0·03)	−0·001 (0·01)
D2004	3·288 (1·09)	2·870 (0·75)	2·861 (0·63)	2·743 (0·84)
D2005	4·294 (1·41)	3·652 (0·96)	3·665 (0·80)	3·968 (1·24)
D2006	4·114 (1·33)	3·413 (0·89)	3·405 (0·74)	3·454 (1·06)
D2007	4·157 (1·34)	3·335 (0·88)	3·290 (0·72)	2·899 (0·88)
PRECIP		0·037 (0·33)		
TEMP			−0·023 (0·63)	
PDSI				−0·111 (1·39)
VEC4	−3·263 (4·08)***			
VEC6	1·638 (1·20)			
VEC15	−1·630 (1·76)*			
VEC16	−2·201 (1·74)*	−1·728 (1·64)	−1·725 (1·53)	−1·696 (1·67)*
VEC46	2·402 (2·22)**			
VEC47	1·319 (1·21)			
VEC41				−0·013 (0·01)
Constant	48·989 (4·12)***	20·222 (1·65)*	21·861 (1·79)*	17·605 (1·44)
Observations	290	290	290	290
Number of locations	58	58	58	58
AIC	1044·021	1055·944	1055·721	1055·966
r (S.E.)	2·111 (0·714)	1·668 (0·492)	1·656 (0·422)	1·787 (0·958)
s (S.E.)	7·474 (4·172)	4·689 (2·413)	4·598 (4·910)	5·414 (5·846)
Moran's I on residuals (P value)	0·196	0·200	0·200	0·195

S.E., Standard error; AIC, Akaike's Information Criterion.

Z statistics in parentheses.

*Significant at 10%, ** significant at 5%, *** significant at 1%.

[a], [b] Represent instrumentalized versions of FORCLOSE and INCOME, respectively. Instruments are bird, mosquito, net migration, precip, temp, PDSI poverty, unemployment rate, education, airport, equine, elevation, urbanization, popdense, roads, log of area and D2004–D2007.

of WNV. Under this hypothesis the regression coefficient on MOSQUITO is expected to be positive.

Moran's I [28] was employed to verify the presence of spatial autocorrelation in HV in both states. It is a coefficient to measure the strength of spatial autocorrelation in regional data and the results are given in Table 3. The calculation of Moran's I is based on Pearson's correlation coefficient formula and was derived using neighbouring values of HV. The Z value is 8·552 for California, while that for Colorado is 14·848. Both are positive and significant indicating the presence of positive spatial autocorrelation.

Estimation of the parameters of the RENB model was undertaken by the technique of maximum likelihood and estimations were performed in Stata v. 9.2 (StataCorp, USA). Pearson's correlation was employed to ascertain the degree of multi-collinearity among the potential risk factors identified in this study. In conformity with the literature, only those

risk factors whose correlations were not in excess of 0·8 were included [11]. The econometric analysis estimates four different regressions models using HV as the dependent variable. The baseline model (model 1) excludes climatic and economic variables. In model 2, PRECIP and INCOME are included. In model 3, only the climatic variable TEMP is included. Finally in model 4, only the PDSI climatic variable is included. The results of models 2–4 are used as a robustness check of the stated hypotheses.

The results of the instrumental variable spatial filtering RENB for California are presented in Table 4. The estimated coefficients on FORHAT are positive and statistically significant at the 5% level in all models, with values ranging from 0·382 to 0·413. The expected positive relationship suggests that WNV prevalence is higher in counties with a higher number of foreclosed homes. The coefficients on MOSQUITO are highly significant at either the 1%, 5%

Table 5. *Instrumental variable spatial filtering random-effects negative binomial regression for Colorado*

	Model 1	Model 2	Model 3	Model 4
FORHAT[a]	2·631 (9·80)***	2·854 (6·36)***	2·926 (6·89)***	2·961 (7·95)***
INCHAT[b]		−4·573 (3·83)***	−3·706 (2·46)**	−4·071 (3·91)***
MOSQUITO	0·009 (2·36)**	0·006 (1·07)	0·007 (1·85)*	0·007 (1·96)*
POPDENSE	0·242 (1·90)*	0·135 (0·60)	0·205 (1·15)	0·055 (0·36)
D2004	−2·692 (10·96)***	−2·453 (7·01)***	−2·643 (9·73)***	−2·733 (13·88)***
D2005	−2·968 (14·49)***	−2·854 (11·53)***	−2·953 (13·43)***	−3·066 (11·75)***
D2006	−2·923 (9·46)***	−2·768 (10·30)***	−2·934 (11·82)***	−2·985 (11·80)***
D2007	−2·689 (12·21)***	−2·736 (9·18)***	−3·023 (9·89)***	−3·066 (10·87)***
PRECIP		−0·637 (0·68)		
TEMP			−0·030 (0·31)	
PDSI				0·024 (0·30)
VEC2	4·392 (3·56)***		4·707 (3·83)***	4·005 (2·94)***
VEC4	3·269 (2·98)***	2·554 (1·68)*	3·320 (2·41)**	
VEC6	3·319 (2·86)***	1·937 (2·02)**	3·175 (3·06)***	2·753 (2·48)**
VEC22	1·866 (1·82)*	1·597 (1·76)*	1·607 (1·80)*	1·643 (1·68)*
Constant	35·232 (3·02)***	45·016 (3·74)***	36·123 (2·02)**	39·195 (3·65)***
Observations	315	315	315	315
Number of locations	63	63	63	63
AIC	1221·647	1221·554	1205·827	1213·728
r (S.E.)	2·438 (1·351)	2·552 (4·443)	2·856 (7·898)	2·464 (0·905)
s (S.E.)	6·376 (3·206)	6·075 (10·964)	7·300 (21·131)	5·042 (1·873)
Moran's I on residuals (P value)	0·125	0·130	0·125	0·123

S.E., Standard error; AIC, Akaike's Information Criterion.

Z statistics in parentheses.

*Significant at 10%, ** significant at 5% *** significant at 1%.

[a], [b] Represent instrumentalized versions of FORCLOSE and INCOME, respectively. Instruments are bird, mosquito, net migration precip, temp, PDSI poverty, unemployment rate, education, airport, equine, elevation, urbanization, popdense, roads, log of area and D2004–D2007.

or 10% level in all models with values ranging from 0·005 to 0·11. This result provides evidence that WNV prevalence is higher in counties with a higher number of mosquito pools. Only the eigenvectors VEC4, VEC15 and VEC16 are significant at either the 1%, 5% or 10% level. The results of the instrumental variable spatial filtering RENB for Colorado are presented in Table 5. The coefficients on FORHAT are consistently positive and statistically significant at the 1% level, with values ranging from 2·631 to 2·926. This provides evidence that home foreclosures contributed significantly to the high prevalence of human WNV in Colorado. The coefficients on MOSQUITO are positive and statistically significant at the 5% and 10% levels with values ranging from 0·006 to 0·009 in models 1, 3 and 4. These results provide evidence that human WNV is higher in counties with a higher number of mosquito pools. All the components of SFILTER are significant at either the 1% or 10% levels. Based on Akaike's Information Criterion (AIC), models 1 and 2 are the most parsimonious for California and Colorado, respectively.

DISCUSSION AND CONCLUSIONS

This study set out to investigate the environmental determinants of human WNV in the states of California and Colorado from 2003 to 2007. The presence of spatial autocorrelation is not surprising at all because the geographical distribution of WNV is determined by climatic, environmental and anthropogenic factors [3, 29]. These are factors that do not adhere to county boundaries so events in one location are dependent on events in another location. In particular, the prevalence of WNV in one county may spill over into an adjoining county. This study consequently employs a spatial filtering random-effects negative binomial model to test a series of hypotheses relating to mosquito pools and home foreclosures. We find that a high number of home foreclosures lead to a greater number of unmaintained properties that could serve as breeding grounds for the mosquito vector thus leading to a greater number of WNV cases. This finding is consistent with previous results [4, 5] that delinquent mortgages and neglected swimming pools

contributed to the outbreak of WNV in Kern and Orange counties, respectively. This finding is quite novel because the impact of home foreclosures on the transmission of human WNV has hitherto not been explicitly investigated. Mosquitoes serve as the primary vector for WNV. A large number of mosquito pools means an increased vector population, hence a high prevalence of WNV and an increased risk of West Nile infection in humans. This is consistent with the findings of Patnaik *et al.* [6] that mosquito pools were significant in predicting WNV prevalence in Colorado. These results are robust to a variety of model specifications. In terms of policy implications, interventions such as the control of the mosquito vector and increased attention to the maintenance of foreclosed homes are implied by our findings. The results of this study also suggest that counties that exhibit characteristics such as a high number of mosquito pools and a high number of home foreclosures should be allocated more resources for WNV surveillance and abatement.

There are two possible limitations of this study. First, this study was conducted at the county-level. Using data at a finer spatial scale, such as census tract or block level, may produce different results. Second, since data on home foreclosures reveal that they peaked in 2010, it would be worthwhile extending the data to 2012 to further verify the home foreclosure hypothesis examined in this study. Despite these limitations, the findings of this study are informative and a significant contribution to current knowledge on WNV transmission.

DECLARATION OF INTEREST

None.

REFERENCES

1. **CDC.** West Nile virus – statistics, surveillance and control archive (http://www.cdc.gov/ncidod/dvbid/westnile/city_states.htm). Accessed 16 May 2011.
2. **Hayes EB, et al.** Virology, pathology, and clinical manifestations of West Nile virus disease. *Emerging and Infectious Diseases* 2005; **11**: 1174–1179.
3. **Kiszewski A, et al.** A global index representing the stability of malaria transmission. *American Journal of Tropical Medicine and Hygiene* 2004; **70**: 486–498.
4. **Reisen WK, et al.** Delinquent mortgages, neglected swimming pools, and West Nile virus, California. *Emerging Infectious Diseases* 2004; **14**: 1747–1749.
5. **Harrigan RJ, et al.** Economic conditions predict the prevalence of West Nile virus. *PLoS One* 2010; **5**: 1–8.
6. **Patnaik JL, Juliusson L, Vogt RL.** Environmental predictors of human West Nile virus infections, Colorado. *Emerging Infectious Diseases* 2007; **13**: 1788–1790.
7. **Kwan JL, et al.** West Nile virus emergence and persistence in Los Angeles, California, 2003–2008. *American Journal of Tropical Medicine and Hygiene* 2010; **83**: 400–412.
8. **Nielson CF, et al.** Risk factors associated with human infection during the 2006 West Nile virus outbreak in Davis, a residential community in northern California. *American Journal of Tropical Medicine and Hygiene* 2008; **78**: 53–62.
9. **Carney RM, et al.** Early warning system for West Nile virus risk areas, California, USA. *Emerging Infectious Diseases* 2011; **17**: 1445–1454.
10. **Messina JP, et al.** West Nile virus in the greater Chicago area: a geographic examination of human illness and risk from 2002 to 2006. *Urisa Journal* 2011; **23**: 5–18.
11. **Linard C, et al.** Determinants of the geographic distribution of Puumala virus and Lyme borreliosis infections in Belgium. *International Journal of Health Geographics* 2007; **6**: 1–14.
12. **Lambert DM, Brown JP, Florax RJGM.** A two-step estimator for a spatial lag model of counts: Theory, small sample performance and an application. *Regional Science and Urban Economics* 2010; **40**: 241–252.
13. **Patuelli R, et al.** the use of spatial filtering techniques: the spatial and space-time structure of German unemployment data – 2009. Tinbergen Institute Discussion Paper: TI 2006-049/3.
14. **Federal Reserve.** Mortgage delinquencies and foreclosures (http://ww.federalreserve.gov/newsevents/speech/bernanke2008050a.htm). Accessed 26 September 2011.
15. **DeGroote JP, et al.** Landscape, demographic, entomological, and climatic associations with human disease incidence of West Nile virus in the state of Iowa, USA. *International Journal of Health Geographics* 2008; **7**: 1–16.
16. **Epstein PR, Delfilippo C.** West Nile virus and drought. *Global Change and Human Health* 2001: **2**: 2–4.
17. **Savage H, Miller B.** House mosquitoes of the USA, *Culex pipiens* complex. *Wing Beats* 1995; **6**: 8–9.
18. **Huhn GD, et al.** The emergence of West Nile virus during a large outbreak in Illinois in 2002. *American Journal of Tropical Medicine and Hygiene* 2005; **72**: 768–776.
19. **NOAA.** Index of public data (http://www.ncdc.noaa.gov/pub/data/cirs). Accessed 10 October 2011.
20. **US Census Bureau.** People and households (http://census.gov/popest/data/historical/index.html). Accessed 2 October 2011.
21. **Data Quick News.** Real estate news and custom data (http://dqnews.com/Articles/2011/News/California/CA-Foreclosures/RRFor111018.aspx). Accessed 11 December 2011.
22. **Colorado Department of Local Affairs.** Foreclosure reports and statistics (http://www.colorado.gov/cs/Satellite?c=Page&childpagename=DOLA-Main%2FCBONLayout&cid=1251596868532&pagename=CBONWrapper). Accessed 10 December 10 2011.

23. **Griffith DA.** Towards a theory of spatial statistics: a rejoinder. *Geographical Analysis* 1981; **13**: 91–93.

24. **Haining R.** Bivariate correlation and spatial data. *Geographical Analysis* 1991; **23**: 210–227.

25. **Getis A, Griffith DA.** Comparative spatial filtering in regression analysis. *Geographical Analysis* 2002; **36**: 90–114.

26. **Tiefelsdorf M, Griffith DA.** Semi-parametric filtering of spatial autocorrelation: the eigenvector approach. *Environment and Planning A* 2007; **39**: 1193–1221.

27. **Griffith DA.** A linear regression solution to the spatial autocorrelation problem. *Journal of Geographical Systems* 2000; **2**: 141–156.

28. **Moran P.** The interpretation of statistical maps. *Journal of the Royal Statistical Society* 1948; **10**: 243–251.

29. **Sattler MA,** *et al.* Habitat characterization and spatial distribution of *Anopheles* sp. Mosquito larvae in Dar es Salaam (Tanzania) during an extended period. *Malaria Journal* 2005; **4**: 4.