Spatial Quantification of Lifeline System Interdependencies

J. Wu & L. Dueñas-Osorio
Rice University, Texas, United States of America

M. Villagrán
Universidad Católica de la Santísima Concepción, Chile

SUMMARY:
The development of theoretical and computational models for interdependent lifeline system performance prediction under seismic hazards constitutes one of the emerging fields of research in infrastructure engineering; however, most interdependent network vulnerability models still lack the necessary validation and calibration with actual/reported data. A few models have attempted to address this gap by integrating post event data, many of which conclude that additional refinement would serve to further improve model accuracy. This paper contributes to model refinement by presenting a method using ordinary point kriging to estimate the spatial correlations or interdependence strengths of networks subjected to earthquakes from restoration data, which is then applied to the 2010 M_W 8.8 Chile Earthquake. The results are plotted on correlation maps and synthesized into correlation plots. Using kriging surfaces appears to produce both intuitive and informative results, thereby providing a feasible method of estimating spatial interdependencies for network vulnerability models.

Keywords: Spatial Correlation, Earthquake Hazard, Infrastructure Networks, Kriging Interpolation, Lifeline Interdependencies

1. INTRODUCTION

As existing infrastructure systems age and expand, network vulnerability to perturbations, such as those caused by earthquakes, becomes a growing concern. Recent incidents such as the 2010 M_W 8.8 Chile earthquake (U.S. Geological Survey 2010), the 2011 M_W 6.3 New Zealand earthquake (U.S. Geological Survey 2011a), and the 2011 M_W 9.0 Japan earthquake (U.S. Geological Survey 2011b), necessitate the understanding of network response, interdependence, and resilience to natural hazards. Thus, theoretical and computational models strive to capture realistic failure mechanisms and yield results applicable to the field. However, most lifeline network vulnerability models still lack the validation and calibration with actual/reported data which is required to assess and improve upon their predictive power and practical applications for lifeline performance assessment and decision making.

A few models have attempted to address this gap by integrating post event data and comparing the modeled predicted results with reported network response to that event. Bonneau and O'Rourke (2009) and Wang (2006) apply and calibrate the Graphical Iterative Response Analysis of Flow Following Earthquakes (GIRAFFE), a model that performs hydraulic analysis on water networks subjected to seismic damage, to the 1994 Northridge earthquake. Wu and Dueñas-Osorio (2012) calibrate the Interdependence Fragility Algorithm (IFA), a seismic damage propagation model proposed by Hernandez-Fajardo and Dueñas-Osorio (2011), to the 2010 Chile earthquake. However, the authors note that additional refinements to network component fragility estimation and interdependence strength quantification would serve to improve network vulnerability predictions.

In fact, a significant contributing factor to multi-network component fragility is the phenomenon of correlation between infrastructure system facilities or nodes. Dueñas-Osorio and Kwasinski (2012) proposed a time series analysis on the restoration of infrastructure systems to estimate the interdependent coupling strength between different networks and advocate the effect of interdependence between systems on component fragility. Lee and Kiremidjian (2007) and Rahnamay-Naeini et al. (2011) derive spatial correlations as applied to transportation networks and communication networks, respectively. This paper contributes to model refinement by presenting a method utilizing ordinary point kriging to estimate the spatial correlations of networks subjected to earthquakes from restoration data in the context of the 2010 Chile earthquake. Kriging surfaces have been implemented in structural applications as seen in Lenda and Ligas (2012); however, kriging has not yet been implemented in seismic hazard applications or interdependence quantification efforts. This paper demonstrates the potential of kriging surfaces to generate spatial correlation estimates for arbitrary points in an infrastructure system, which would contribute to more accurate network component fragility and interdependent coupling strength estimations.

The remainder of this paper is structured as follows: Section 2 discusses the correlation and kriging concepts applied in this paper for spatial correlation estimation. Section 3 discusses the application of these mathematical concepts to derive spatial correlations in the context of the 2010 Chile earthquake. Section 4 discusses the results from the methodology presented in this paper as well as new insights and limitations. Section 5 discusses the main conclusions from this study and directions for future research.

2. MATHEMATICAL CONCEPTS IN CORRELATION ANALYSIS

This section will discuss the theoretical and mathematic concepts involved in the estimation of the spatial correlations of the Chilean lifeline networks presented in this paper. This section begins with a description of the correlation methods performed on spatial data, followed by a discussion on the ordinary point kriging calculations to be implemented into the proposed spatial intra- and inter-dependence quantification methodology.

2.1 Correlation Methods

The correlation analysis performed in this paper uses the Pearson's product-moment coefficient as well as Kendall's tau coefficient. Pearson's coefficient ρ describes the degree of a linear correlation between two sets of data, and is a generally accepted metric for correlation in engineering (Kendall and Gibbons 1990, Ang and Tang 2007, Dueñas-Osorio and Kwasinski 2012). Let X and Y denote two datasets each with n elements denoted by x_i and y_i, respectively. The equation for the Pearson's coefficient is then described by Eqn. 2.1 (Ang and Tang 2007).

$$\rho = \frac{1}{n-1} \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{s_X s_Y} \tag{2.1}$$

where $\bar{x}$ and $\bar{y}$ are sample means of X and Y, respectively, while s_X and s_Y are sample standard deviations of X and Y.

Kendall's tau coefficient t is a rank correlation method which describes the degree of similarity between the ranks of the two sets of data. The primary advantage of using Kendall's tau coefficient is the capability to detect nonlinear correlation, a feature that Pearson's correlation lacks. Let X and Y denote two datasets each with n elements denoted by x_i and y_i, respectively. Kendall's tau coefficient is described in Eqns. 2.2 through 2.4 (Kendall and Gibbons 1990).

$$t = \frac{S}{\sqrt{\frac{1}{2}n(n-1)-U}\sqrt{\frac{1}{2}n(n-1)-V}} \tag{2.2}$$

$$U = \frac{1}{2}\sum_{j=1}^{\alpha} u_j\left(u_j - 1\right) \tag{2.3}$$

$$V = \frac{1}{2}\sum_{k=1}^{\beta} v_k(v_k - 1) \tag{2.4}$$

where u_j describes the number of elements with a tied rank (elements with equal values) for the j^{th} set of tied elements, with X having a total of α of such sets, and v_k describes the number of tied elements for the k^{th} set of tied elements, with Y having a total of β of such sets. These correlation concepts will be applied to the restoration potential, denoted ψ, of spatially distributed points to derive an estimate of spatial correlation between these points, where these points may be regarded as facilities at which lifeline system restoration may be measured. In the context of network analysis, spatial correlations of datasets from the same network describe the auto-correlation, which may be related to intra-network dependency, while spatial correlations of datasets from different networks describe the cross-correlation, which may be related to inter-network dependency. Further discussion on the significance of auto-correlation and cross-correlation values is presented later in the paper.

2.2 Ordinary Point Kriging

The critical component in the estimation of spatial correlation in this study is the formation of kriging surfaces of ψ to enable the extraction of interpolated data at particular spatial coordinates for correlation analysis. The kriging calculation first requires the formulation of the variogram for the data, which describes the spatial variability of ψ. The variogram formulation begins with the calculation of the variogram estimator derived from the raw data, expressed in Eqn. 2.5 (Trauth 2010).

$$\gamma_E(h) = \frac{1}{2*N(h)}\sum_{i=1}^{N(h)}(\psi_{zi} - \psi_{zi+h})^2 \tag{2.5}$$

where ψ_z and ψ_{z+h} denote the restoration potential evaluated at points z_i and z_i+h and $N(h)$ describes the number of pairs of points within the lag interval h, taken to be the Euclidian distances between the two points i and i+h and may be set equal to the mean minimum distance between pairs. A parametric curve is then fitted to the variogram estimator to create the variogram model suitable for the data. Once the variogram has been determined, the kriging calculations may be performed.

Ordinary point kriging interpolates the ψ value at a particular point z_P by calculating the weighted average of ψ values evaluated at an N number of neighboring points. This interpolation is expressed in Eqn. 2.6, where λ_i is the weighting coefficient associated with the point z_i that must be estimated by satisfying the constraints expressed in Eqns. 2.7 through 2.8 and minimizing the mean-squared error expressed in Eqn. 2.9.

$$\hat{\psi}_{zp} = \sum_i^N \lambda_i \psi_{zi} \tag{2.6}$$

$$\sum_i^N \lambda_i = 1 \tag{2.7}$$

$$E\left(\hat{\psi}_{zp} - \psi_{zp}\right) = 0 \tag{2.8}$$

$$E\left(\left(\hat{\psi}_{zp} - \psi_{zp}\right)^2\right) = 2\sum_{i=1}^{N}\lambda_i\gamma\left(z_i, z_p\right) - \sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j\gamma(z_i, z_j) \tag{2.9}$$

where ψ_{zp} is the true, but unknown value of ψ evaluated at a particular point z_P and $\gamma(z_i, z_j)$ is the variogram between the points z_i and z_j. More information regarding ordinary point kriging may be found in MATLAB Recipes for Earth Sciences by Trauth (2010).

3. ESTIMATING SPATIAL CORRELATIONS WITH KRIGING SURFACES

This section describes the methodology used to derive the spatial correlations between the Chilean water and power networks. This section begins with a description of the available network data from which the spatial correlations are derived, followed by a discussion of the meshing process for kriging calculation. This section concludes with correlation strategies to develop spatial intra- and inter-dependence correlation estimates.

3.1 Description of Chilean Networks

The kriging-based spatial correlation strategies pursued by this paper is applied to the electrical power and water delivery systems of the Talcahuano-Concepción area of Chile, in the context of the 27 February 2010 M_W 8.8 Offshore Maule earthquake. Available data includes the number of days of repair prior to the full restoration of lifeline system service at various locations in this area, termed "evaluation nodes." There are 94 evaluation nodes integrated into the procedure, and the length and width of the area containing the data is approximately 20 km and 10 km, respectively. Fig. 3.1 below depicts the locations of these evaluation nodes as well as the power grid and water pipelines in this area.

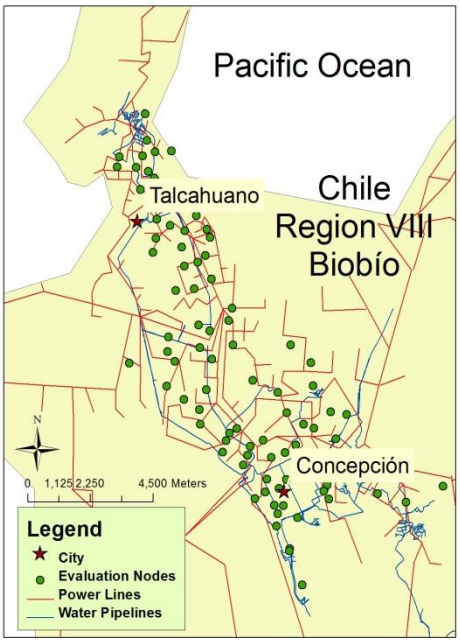


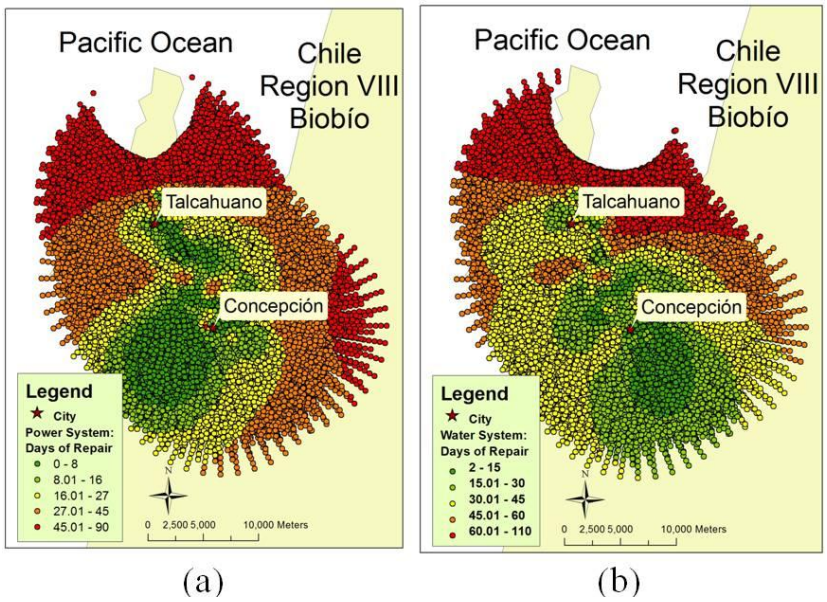
Figure 3.1. Depiction of the evaluation nodes used in this paper.

From the field-collected data, the number of days required to fully restore the service of lifeline networks at the evaluation node is used as a proxy for the restoration potential at each evaluation node, ψ. The maximum and minimum number of days for repair are 61 and 1, respectively, for the power system and 75 and 3, respectively, for the water system. Factors that may contribute to ψ diversity include the prioritization of the restoration of service to a given area, as well as the degree of damage inflicted on the networks that serve that area. The authors also assume that network components (e.g. water pumping stations or electrical substations) that service a particular evaluation node are sufficiently physically close

to the node such that the services experienced by the node are similar to the services experienced by those network components so that commentary on network interdependence is feasible. For example, if some point *i* experiences a level of service α from the local electrical substation, the water pumping station servicing point *i* would also experience a level of service α from the same electrical substation.

Formulation of Spatial Meshes for Kriging Calculation

Spatial correlation analyses, as described later in this paper require that the kriging surfaces must include ψ values at additional coordinates (X',Y'), describing various R Euclidian distances from each evaluation node at coordinates $(\hat{X}, \hat{Y})$ with varying θ angles from the horizontal. Thus, a mesh of (X',Y') coordinates is created using distance increments ΔR of 500 meters up to a maximum distance R_{max} of 10 km and angle increments $\Delta\theta$ of $\pi/20$ from the horizontal. Kriging surfaces of ψ values are derived within this mesh for each network. The resulting kriging surfaces are presented as a cloud of points describing the value of ψ at those points, which are depicted in Fig. 3.2.



(a) (b)

Figure 3.2. Kriging surfaces of interpolated ψ values as a cloud of spatial points for the (a) power and (b) water networks.

Spatial Correlation Analysis

Using the kriging surfaces described above, ψ values may be extracted from particular points (X'_P,Y'_P) resulting from translating the nodes a particular distance R_P at an angle θ_P from the initial evaluation nodes at $(\hat{X}, \hat{Y})$. This operation is diagramed in Fig. 3.3 below.

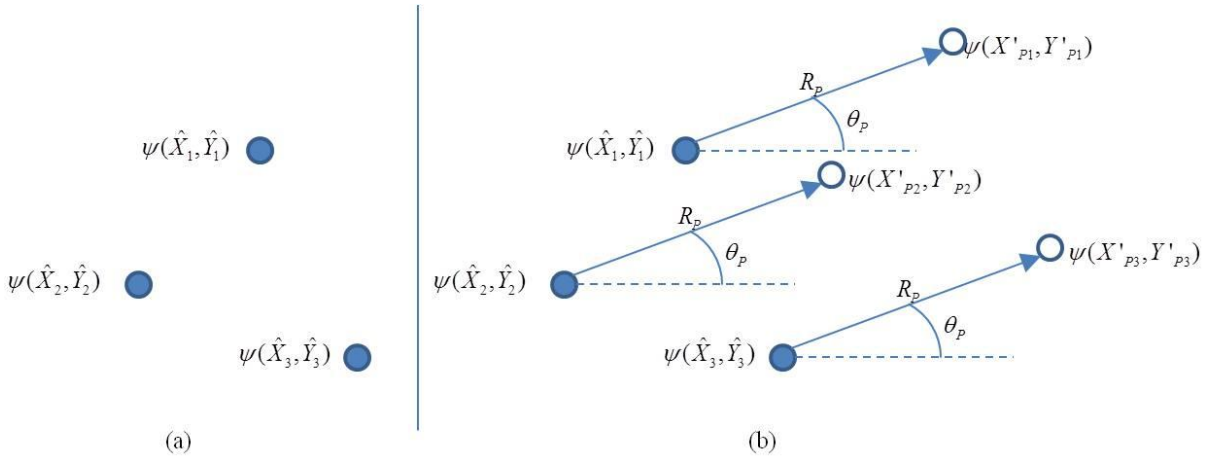


Figure 3.3. Diagram depicting an example of (a) evaluation nodes and (b) the operation performed to find ψ values at particular points using R_P and θ_P.

By performing this operation to each evaluation node, a set of extracted ψ values associated with the parameters R_P and θ_P may be compared to the ψ values at the evaluation nodes using correlation analysis. If the correlation values are averaged over all θ_P values for each R_P, a global (average) spatial correlation value may be derived as a function of displacement from an evaluation node.

Auto-correlation values are obtained from comparing the ψ values for a network at the evaluation nodes with the ψ values for the same network at some displacement, R, from each evaluation node. A positive value of this metric describes the similarity between an arbitrary point and surrounding points, and may be interpreted as an assessment of the similarities of the inflicted stresses and resulting damage due to proximity, prioritization of repair, or a measure of intranetwork dependency, termed the likelihood of a "shared fate." As the auto-correlation approaches 0, the nodes are considered to be more independent (unique) with regards to inflicted damage or repair.

Cross-correlation values are obtained from comparing ψ values for a network (the reference network) at the evaluation nodes with the ψ values for a different network (the adjunct network) at some R displacement from the evaluation nodes. A positive value of this metric describes the similarities between a network at an arbitrary point and a different network at nearby points, which may be interpreted as a measure of the coupling strength between interdependent networks. As the cross-correlation approaches 0, the coupling strength approaches 0 and the networks are considered relatively independent.

4. RESULTS AND DISCUSSION

After appropriate meshing, the formulation of kriging surfaces, and the application of the aforementioned auto- and cross-correlation methods, the estimations of global (average) spatial correlations are derived and presented in correlation maps as seen in Fig. 4.1 as a function of distance and angle from the evaluation nodes. Fig. 4.2 synthesizes these results by averaging the correlations over all angles, resulting in the average spatial correlation as a function of Euclidian distance from the evaluation nodes. The error bars depict 1 standard deviation from the mean.

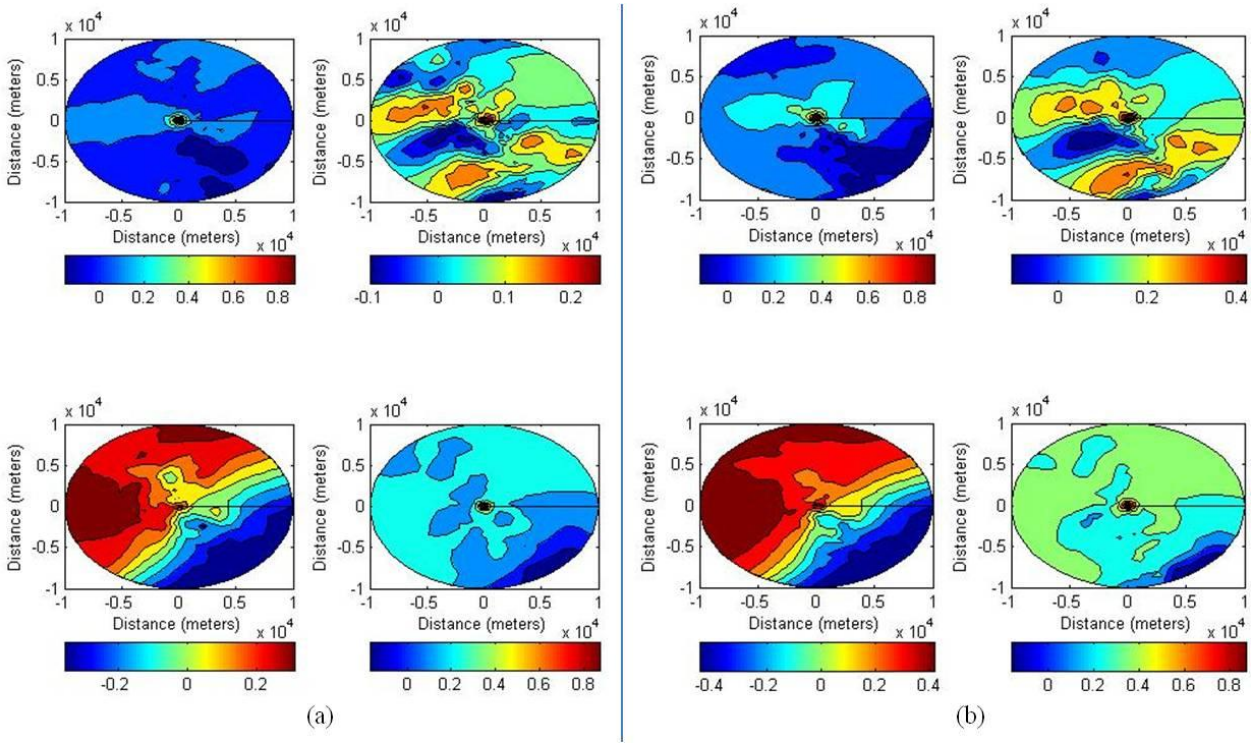


Figure 4.1. Spatial correlation maps between an arbitrary point and another point as a function of distance from the arbitrary point at the center using (a) Pearson's coefficient (b) Kendall's tau coefficient. For both sets of maps: top left, auto-correlation for power system; bottom right, auto-correlation for water system; top right, cross-correlation with power as reference system and water as adjunct system; bottom left, cross-correlation with water as reference system and power as adjunct system.

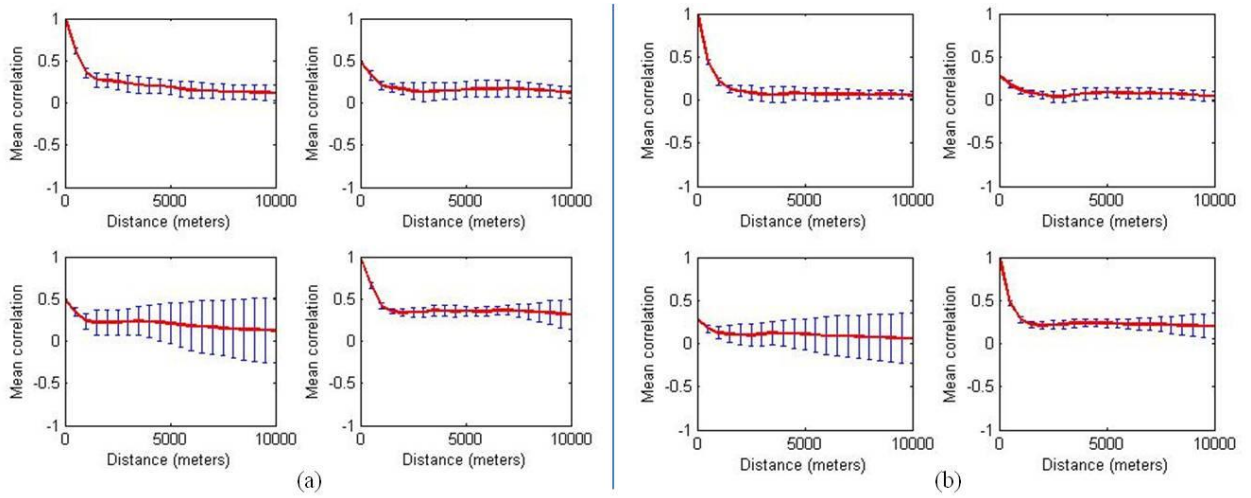


Figure 4.2. Spatial correlation plots between an arbitrary point and another point as a function of distance from the arbitrary point, averaged over all directions, using (a) Pearson's coefficient and (b) Kendall's tau coefficient. For both sets of plots: top left, auto-correlation for power system; bottom right, auto-correlation for water system; top right, cross-correlation with power as reference system and water as adjunct system; bottom left, cross-correlation with water as reference system and power as adjunct system.

In both Fig. 4.1 and Fig. 4.2, the left set of plots (a) depict the correlation results using Pearson's coefficient, and the right set of plots (b) depict the correlation results using Kendall's tau coefficient. In

each set of plots, the main diagonal describes the auto-correlation in each network; the top left depicts this value for the power network, and the bottom right depicts this value for the water network. The remaining plots represent cross-correlation between different networks; the top right depicts this value for the power network as the reference network with the water network as the adjunct network, and the bottom left depicts this value for water as the reference network and power as the adjunct network.

Regarding Fig. 4.1, all plots have a very small region in the center that depicts very high correlation, while correlation generally decreases as the distance from the center increases. However, it can be observed that the correlations are not radially symmetric (anisotropic); points equidistant from the center may vary greatly. This variability is reflected in the error bars of Fig. 4.2. Additionally, it appears that correlation generally tends to decrease towards the south-southeast and increases north-northwest. This trend may be explained through observing the kriging surfaces in Fig. 3.2 from Section 3, which depicts ψ generally increasing going northward—service in the southern portion of the area tends to be restored before the northern portion. This implies that there may be a south-to-north repair scheme, or that seismic damage is more extensive in the north relative to the south.

Regarding Fig. 4.2, all plots have positive initial correlation values that decay towards 0 until a distance of approximately 1 000 – 1 500 meters, after which the correlation values are relatively stagnant and independent of distance. This implies that spatial correlation only plays a significant impact between points within 1.5 kilometers of each other, which is intuitive as spatial correlation should weaken as distance between points increases. The auto-correlation plots have initial values of 1.0 correlation, as this expresses the correlation of identical datasets (same network, same values of ψ evaulated at the same points). The cross-correlation plots have initial values of 0.483 and 0.284 for Pearson's coefficient and Kendall's tau coefficient, respectively; the ψ values of the power and water networks have a discernible linear correlation and slightly weak rank correlation. With the distance from the evaluation node equal to 0, this value may be interpreted a measure of the coupling strength between the local power and water network components. The authors compare these values with the cross correlation values of 0.35, 0.53, 0.50, and 0.75 calculated from a time series analysis applied in Dueñas-Osorio and Kwasinski (2012). The values calculated in this paper are reasonable in comparison with the results from the time series method; thus, the methods utilized in this work appear to be viable.

Current limitations to the application of ordinary point kriging to spatial correlation estimation include the issue regarding the kriging surface evaluated outside the cloud of evaluation nodes. As can be seen in the kriging surfaces in Fig. 3.2, the values outside the cloud are fairly constant, being simply extrapolations from the evaluation nodes at the edges of the cloud, which may not reflect the true spatial correlation found at those points, especially upon observing the intricacy of the central part of the surface where evaluation nodes exist. In correlation analysis, any displacement R_P results in the inclusion of these extrapolated values outside the cloud and introduces a subset of elements that will be falsely correlated (due to the consistency of the extrapolations) that is integrated into the global (average) correlation associated with R_P. At sufficiently large R_P, the impact of the extrapolated values will become significant, thus implying that the spatial correlation estimation becomes less accurate for larger distances. On the other hand, the spatial correlation at large distances is expected to be small and possibly disregarded, thereby minimizing this issue.

Another issue regarding ordinary point kriging is the potential of creating too fine of a mesh for kriging surface calculation given the granularity of given data (evaluation nodes). The interpolated ψ values very close to an evaluation node (and very far from neighboring nodes) will be heavily weighted towards the ψ values at the evaluation node, thereby creating nearly identical datasets and potentially introducing false correlations at sufficiently small R_P. While spatial correlation at small distances is expected to be high, correlation estimates would only be significant if there is sufficient contribution from neighboring nodes. One option is to consider the mean minimum distance between pairs of evaluation nodes (as seen in the

variogram calculation) when determining a proper ΔR for mesh creation. The mean minimum distance between pairs of evaluation nodes in this paper is approximately 1 200 meters.

5. CONCLUSION

As shown by the results presented in this paper, the application of ordinary point kriging to spatial correlation estimation appears to be feasible. The correlation maps and plots demonstrate that spatial correlation is significant for points within a distance of 1.5 kilometers of each other, after which correlation stagnates and approaches 0. The correlation maps reveal a trend in restoration in the north-south directions, which may be verified from observing the kriging surfaces. This suggests a relationship between location and restoration scheme or seismic damage distribution. From observing the correlation plots, the cross-correlation between the power and water systems at the evaluation nodes yields values of 0.483 and 0.284 for Pearson's coefficient and Kendall's tau coefficient, respectively. These correlation values appear reasonable in comparison to cross-correlation values derived in Dueñas-Osorio and Kwasinski (2012).

Possible ventures for future research include the consideration of the state of restoration at varying points in time, as found in the work by Dueñas-Osorio and Kwasinski (2011). Such an additional dimension would allow for comparisons of the state of restoration in time between two points, resulting in a more accurate assessment of the correlation between those two points. This would also allow for the formulation of correlation maps particularized for an arbitrary point, thereby enhancing the granularity of correlation estimates toward the local level, as current methods only allow for correlation maps averaged over all evaluation nodes to describe estimated global spatial correlation.

The primary objective of the authors is to promote the validation and calibration of theoretical and computational models toward reality with reported data from actual seismic events. Refined and realistic models enable effective planning, design, maintenance, and retrofit of infrastructure networks; utility companies and governments may more precisely identify vulnerable components of an infrastructure network and allocate resources accordingly. The methods presented in this paper allow for the integration of reported restoration data into hazard estimations through spatial correlations, resulting in a more accurate prediction of network vulnerability by interdependence models.

ACKNOWLEDGEMENTS

The present work has been funded in part by the National Science Foundation through grants CMMI-0728040 and CMMI-0748231. This support is gratefully acknowledged. Any opinions, findings and conclusions or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Ang, A.H-S., and Tang, W.H. (2007). Probability Concepts in Engineering, John Wiley & Sons, Inc, United States.

Bonneau, A. and O'Rourke, T.D. (2009). Water Supply Performance During Earthquakes and Extreme Events, MCEER, University at Buffalo, Buffalo, New York.

Dueñas-Osorio, L. and Kwasinski, A. (2012). Quantification of Lifeline System Interdependencies after the 27 February 2010 M_W 8.8 Offshore Maule, Chile Earthquake. In press manuscript.

Hernandez-Fajardo, I., and Dueñas-Osorio, L. (2011). Sequential Propagation of Seismic Fragility Across Interdependent Lifeline Systems, *Earthquake Spectra*. **27**, 23-43.

Kendall, M. and Gibbons, J.D. (1990). Rank Correlation Methods, Oxford University Press, New York City, New York.

Lee, R. and Kiremidjian, A.S. (2007). Uncertainty and Correlation for Loss Assessment of Spatially Distributed Systems, *Earthquake Spectra*. **23:4**, 753-770.

Lenda, G. and Ligas, M. (2012). Application of Splines Supported by Kriging for Precise Shape Analysis of Incompletely Measured Structures, *Journal of Computing in Civil Engineering*. **26:2**, 214-224.

Rahnamay-Naeini, M., Pezoa, J.E., Azar, G., Ghani, N., and Hayat, M.M. (2011). Modeling Stochastic Correlated Failures and their Effects on Network Reliability, *2011 20th International Conference on Computer Communication Networks*. 1-6.

Trauth, M.H. (2010). MATLAB Recipes for Earth Sciences, Springer.

U.S. Geological Survey. (2010). Shakemap us2010tfan. http://earthquake.usgs.gov.

U.S. Geological Survey. (2011). Magnitude 6.1 – South Island of New Zealand. http://earthquake.usgs.gov.

U.S. Geological Survey. (2011). Magnitude 9.0 – Near the East Cost of Honshu, Japan. http://earthquake.usgs.gov.

Wang, Y. (2006). Seismic Performance Evaluation of Water Supply Systems, Cornell University, Dissertation in partial fulfillment of Ph.D.

Wu, J. and Dueñas-Osorio, L. (2012). Calibration and Validation of a Seismic Damage Propagation Model for Interdependent Infrastructure Systems. In press manuscript.