

Scale effects in species distribution models: implications for conservation planning under climate change

Changwan Seo^{1,*}, James H. Thorne²,
Lee Hannah³ and Wilfried Thuiller⁴

¹University of Seoul, 90 Jeonmong-dong, Dongdaemum-gu, Seoul 130-743, Korea

²University of California, Davis, CA 95616, USA

³Conservation International, Bren School, University of California, Santa Barbara, CA 93106, USA

⁴Laboratoire d'Ecologie Alpine, UMR CNRS 5553, Université J. Fourier BP 53, 38041 Grenoble, France

*Author for correspondence (dharmascw@uos.ac.kr).

Predictions of future species' ranges under climate change are needed for conservation planning, for which species distribution models (SDMs) are widely used. However, global climate model-based (GCM) output grids can bias the area identified as suitable when these are used as SDM predictor variables, because GCM outputs, typically at least 50 × 50 km, are biologically coarse. We tested the assumption that species ranges can be equally well portrayed in SDMs operating on base data of different grid sizes by comparing SDM performance statistics and area selected by four SDMs run at seven grid sizes, for nine species of contrasting range size. Area selected was disproportionately larger for SDMs run on larger grid sizes, indicating a cut-off point above which model results were less reliable. Up to 2.89 times more species range area was selected by SDMs operating on grids above 50 × 50 km, compared to SDMs operating at 1 km². Spatial congruence between areas selected as range also diverged as grid size increased, particularly for species with ranges between 20 000 and 90 000 km². These results indicate the need for caution when using such data to plan future protected areas, because an overly large predicted range could lead to inappropriate reserve location selection.

Keywords: global climate models; grid size sensitivity analysis; species range

1. INTRODUCTION

Species distribution models (SDMs) are used to estimate species ranges in conservation planning efforts (Rodríguez *et al.* 2007) and in forecasts of potential range shifts under climate change (Schwartz *et al.* 2006). Uses of SDM outputs include the selection of protected areas and identification of future reserves that mediate extinction risk from climate change. The practical importance of these applications

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rsbl.2008.0476> or via <http://journals.royalsociety.org>.

has inspired comparative reviews of SDM modelling techniques (e.g. Elith *et al.* 2006), because incorrect species' range predictions may promote spatially flawed conservation efforts.

Understanding effects related to spatial grid size is fundamental to evaluating SDM results, because SDMs are often employed in studies using large grid sizes. Particularly, future climate scenarios are often several orders of magnitude more generalized than the scale at which species experience the environment (Guisan & Thuiller 2005). Varying the operational scale (grid size) of SDMs was found to have negligible effects on SDM performance as measured by the area under the receiver operating characteristic curve (AUC) values (Guisan *et al.* 2007). Evaluation of grid size effects on SDM estimation of range size has not been examined.

This study addresses the effects of grid size on the size and location of predicted species distributions for SDMs run at seven grid sizes. We quantified the degree of spatial bias introduced by using coarser spatial grids, and identified grid size thresholds above or below which SDMs disproportionately predict more or less species range area. We tested four SDMs using BIOMOD (Thuiller 2003): generalized linear model (GLM); generalized additive model (GAM); classification tree analysis (CTA); and artificial neural networks (ANN). We assessed whether these effects are different for species of contrasting range sizes by modelling nine California endemic or near endemic tree species, which we classed into narrow (less than 20 000 km²), intermediate (20 000–90 000 km²) or broad range size (more than 90 000 km²) within California's 410 000 km² (Viers *et al.* 2006).

2. MATERIAL AND METHODS

Seven model operational grid sizes were selected; 1 × 1, 2 × 2, 4 × 4, 8 × 8, 16 × 16, 32 × 32 and 64 × 64 km. Fine resolution predictor variables (1 km²) and target species presence or absence (P/A) were resampled into these grids.

State-wide species P/A data were obtained from 32 553 geo-referenced (accurate within 50 m) vegetation plots and herbarium records and were sampled for the nine species (authority Hickman 1993). Species not listed in a plot were labelled as absent at that plot's location. Recorded presences ranged from 93 to 2285, recorded absences from 30 456 to 32 448 (see the electronic supplementary material). Species' P/A plot records were resampled into each grid, and presence per cell was assigned if one or more presence records were found, otherwise it was labelled absent. Absence cells were randomly selected from all absence cells; the absence to presence ratio was 2 : 1 (Kvamme 1985), except when not enough absence cells were available within the extent of California.

Of the 19 climate variables from WorldClim at 1 km² (Hijmans *et al.* 2005), 6 were selected by running a principal components analysis over all the grids and selecting the least correlated: annual temperature range; mean temperature of the driest quarter; mean temperature of the coldest quarter; precipitation seasonality; precipitation in the wettest quarter; and precipitation of the warmest quarter. Predictor variables were resampled to each of the six other grid sizes using the mean values of the nested 1 km² predictor variables.

The SDMs were run at each grid size in BIOMOD for each species, and each probability distribution was predicted across California. A total of 252 SDMs were produced (nine species by four models by seven grid sizes). The area predicted as range for each species using a cut-off threshold determined by maximizing the percentage of presence and absence correctly predicted was recorded for each model run, and used to create binary range maps from SDM probability surfaces.

We evaluated SDM outputs in three ways: (i) area selected as habitat by each model was compared across grid sizes, (ii) a spatial congruence analysis between operational scales, and (iii) model performance at each grid size was assessed with AUC and Kappa statistics using a fivefold cross-validation (Guisan *et al.* 2006).

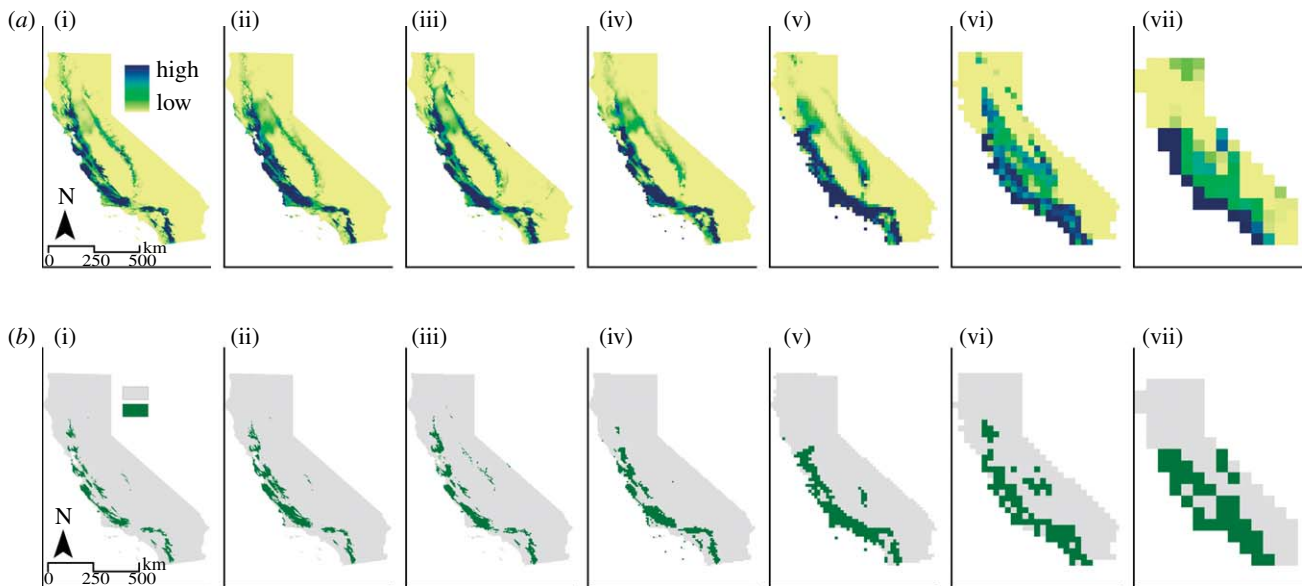


Figure 1. Range maps for *Pinus coulteri*. (a) The GAM SDM probability outputs run at grid sizes from 1×1 km to 64×64 km; (b) range selected using the AUC cut-off value for each grid size (grey, absence; green, presence). (i) 1×1 km, (ii) 2×2 km, (iii) 4×4 km, (iv) 8×8 km, (v) 16×16 km, (vi) 32×32 km, and (vii) 64×64 km.

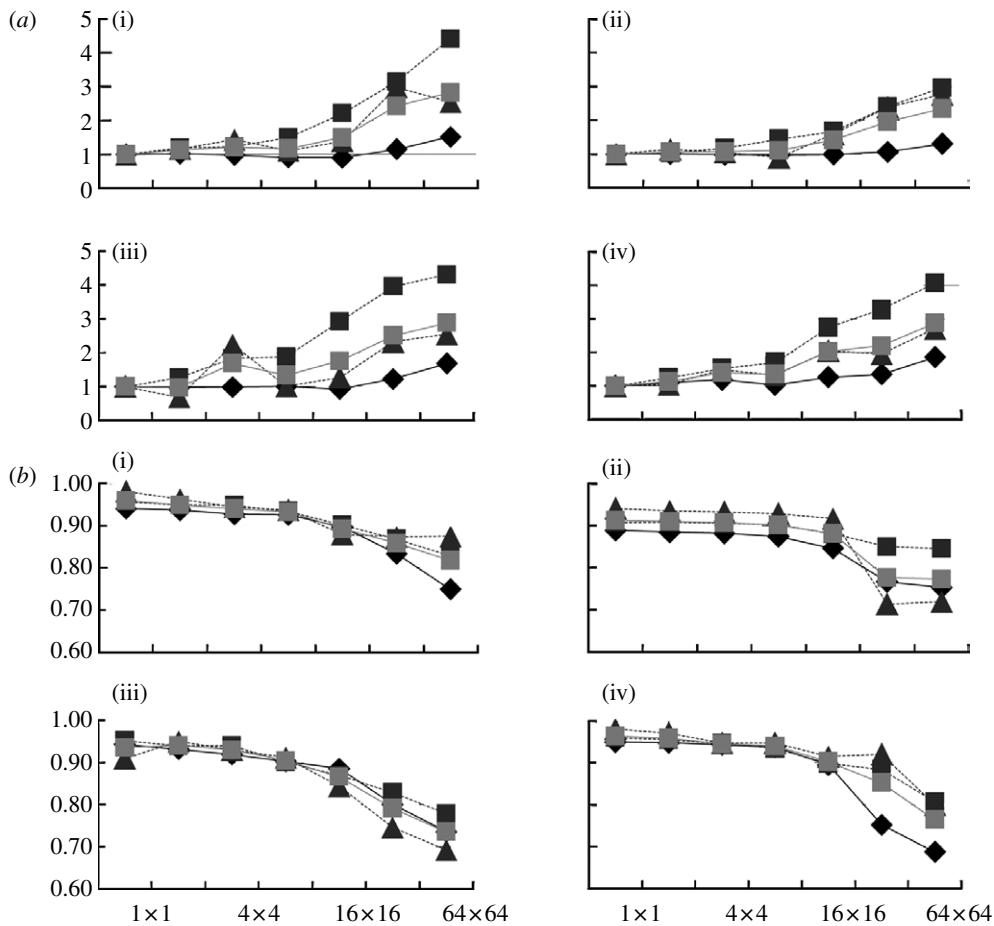


Figure 2. (a) Ratios of area selected as range between the 1×1 km grid size SDMs and each other grid size. Intermediate range size species show the greatest divergence in predicted range area as grid size increases. B, broad range size (diamonds); I, intermediate range size (black squares); N, narrow range size (triangles); A, all range sizes combined (grey squares). (b) The AUC values of SDM probability maps from each grid size, calculated using the 1×1 km presence/absence data. The decline in AUC values represents spatial bias. The 8×8 km operational scale is an optimal threshold that permits coarser grid modelling, while maintaining most of the accuracy of the finer operational scales. (i) GAM, (ii) GLM, (iii) CTA, and (iv) ANN.

Ratios of species' binary range extents across grid sizes were calculated using the 1×1 km extent as denominator. This permitted viewing of normalized differences of area selected by SDMs operating on different grid sizes. We assessed the results for species

with different range size classes to see whether the range size itself might have an impact on optimal SDM operational scale.

To assess spatial congruence (agreement on predicted range locations) between different grid size SDMs, we developed AUC

Table 1. Model performance by scale. Fivefold cross-validation AUC and Kappa statistics are shown.

	species range size	scale										grand mean	s.d.	
		1 × 1 km	2 × 2 km	4 × 4 km	8 × 8 km	16 × 16 km	32 × 32 km	64 × 64 km						
<i>AUC</i>														
GAM	broad	0.94	0.94	0.91	0.90	0.90	0.90	0.90	0.90	0.86	0.84	0.90	0.036	
	intermediate	0.95	0.94	0.95	0.95	0.94	0.94	0.94	0.94	0.94	0.91	0.94	0.013	
	narrow	0.95	0.96	0.97	0.93	0.88	0.88	0.88	0.88	0.88	0.66	0.89	0.105	
	scale mean	0.95	0.95	0.94	0.92	0.91	0.91	0.91	0.91	0.89	0.81	0.91	0.050	
GLM	broad	0.89	0.89	0.88	0.86	0.87	0.83	0.83	0.83	0.83	0.85	0.86	0.020	
	intermediate	0.91	0.91	0.92	0.90	0.93	0.93	0.93	0.93	0.93	0.93	0.92	0.014	
	narrow	0.93	0.93	0.94	0.91	0.88	0.74	0.74	0.74	0.74	0.64	0.85	0.117	
	scale mean	0.91	0.91	0.91	0.89	0.89	0.84	0.84	0.84	0.84	0.81	0.88	0.042	
CTA	broad	0.92	0.91	0.88	0.86	0.84	0.83	0.83	0.83	0.83	0.81	0.86	0.040	
	intermediate	0.93	0.91	0.91	0.89	0.89	0.86	0.86	0.86	0.86	0.80	0.89	0.045	
	narrow	0.90	0.88	0.89	0.85	0.76	0.81	0.81	0.81	0.81	0.56	0.81	0.119	
	scale mean	0.91	0.90	0.90	0.87	0.83	0.83	0.83	0.83	0.83	0.72	0.85	0.066	
ANN	broad	0.94	0.93	0.92	0.89	0.89	0.86	0.86	0.86	0.86	0.74	0.88	0.070	
	intermediate	0.95	0.94	0.94	0.92	0.93	0.93	0.93	0.93	0.93	0.89	0.93	0.018	
	narrow	0.95	0.93	0.92	0.91	0.86	0.90	0.90	0.90	0.90	0.60	0.87	0.121	
	scale mean	0.94	0.93	0.93	0.91	0.89	0.89	0.89	0.89	0.89	0.74	0.89	0.068	
<i>Kappa</i>														
GAM	broad	0.76	0.75	0.68	0.64	0.62	0.57	0.57	0.62	0.57	0.64	0.67	0.068	
	intermediate	0.77	0.74	0.75	0.76	0.79	0.79	0.79	0.79	0.79	0.72	0.76	0.027	
	narrow	0.82	0.84	0.86	0.77	0.68	0.71	0.71	0.68	0.71	0.36	0.72	0.173	
	scale mean	0.78	0.78	0.77	0.72	0.70	0.69	0.69	0.70	0.69	0.57	0.72	0.073	
GLM	broad	0.63	0.62	0.60	0.55	0.58	0.52	0.52	0.58	0.52	0.65	0.59	0.045	
	intermediate	0.68	0.69	0.71	0.66	0.76	0.75	0.75	0.76	0.75	0.76	0.72	0.041	
	narrow	0.77	0.77	0.81	0.74	0.73	0.52	0.52	0.73	0.52	0.30	0.67	0.186	
	scale mean	0.70	0.69	0.71	0.65	0.69	0.60	0.60	0.69	0.60	0.57	0.66	0.054	
CTA	broad	0.74	0.73	0.68	0.60	0.57	0.53	0.53	0.57	0.53	0.63	0.64	0.079	
	intermediate	0.77	0.72	0.73	0.69	0.73	0.67	0.67	0.73	0.67	0.61	0.70	0.051	
	narrow	0.76	0.72	0.72	0.62	0.48	0.57	0.57	0.48	0.57	0.23	0.59	0.186	
	scale mean	0.76	0.72	0.71	0.64	0.60	0.59	0.59	0.60	0.59	0.49	0.64	0.094	
ANN	broad	0.76	0.76	0.71	0.63	0.63	0.59	0.59	0.63	0.59	0.47	0.65	0.105	
	intermediate	0.77	0.75	0.75	0.73	0.75	0.78	0.78	0.75	0.78	0.64	0.74	0.045	
	narrow	0.79	0.78	0.80	0.76	0.70	0.68	0.68	0.70	0.68	0.33	0.69	0.164	
	scale mean	0.77	0.76	0.75	0.71	0.69	0.68	0.68	0.69	0.68	0.48	0.69	0.101	

values that compared probability maps, computed at every grid size, with the baseline 1×1 km P/A records. For example, a species' 1×1 km P/A grid was overlaid with its 4×4 km probability surface, and an AUC value was computed. The difference between these AUC values from those derived from the 1×1 km SDM outputs represents the possible spatial bias introduced by using scaled-up P/A data and predictor variables. This approach tested the assumption that high SDM AUC values across grid sizes mean high spatial agreement.

Cross-validated AUC values were summarized to present overall model performance by taking the mean AUC values of all model accuracies, and of the species in each range size class. The same was done with the Kappa statistics that indicate what proportion of model performance is attributable to correct model selection.

3. RESULTS

Operational scale affected all SDM runs, causing selection of approximately three times more area when models were run on 64×64 km grids than when they were run at 1×1 km. The SDM spatial outputs of both probability (figure 1a) and binary range maps (figure 1b) differed increasingly as grids became coarser. Bias in area selected was most pronounced in species with intermediate-sized ranges, for which coarse resolution models selected four times more area. Area for narrowly distributed species also increased 2.5–3.0 times under the coarsest grids. Broadly distributed species showed the lowest change, with area selected increasing approximately 1.5 times.

The ratio analysis of range size (figure 2a) showed that the bias due to grid size occurred in all model types, particularly for intermediate range size species. From 16×16 km to 64×64 km grid sizes, all SDMs selected on average 1.42–2.89 times more area than the 1 km² model outputs. Model agreement on range size (area selected) was good for broad, but diverged for narrow and intermediate range size species as operational scale increased. At spatial grid size below 16×16 km, there was good agreement among model area estimates for species of all range sizes.

Fivefold AUC cross-validation values declined as grid size increased (table 1): combined model values were 0.92, 0.88 and 0.85 for intermediate, broad and narrow range size species, respectively. Decline between 1×1 and 64×64 km AUC values on average was 0.05, 0.08 and 0.31 for intermediate, broad and narrow range size species, respectively. Kappa values correspondingly declined by on average 0.05, 0.17 and 0.52 for broad, intermediate and narrow range size species, respectively.

The spatial congruence AUC values also declined with increasing grid size (figure 2b). For the 1×1 km P/A grid to 1×1 km probability surfaces, average AUC values were GAM 0.96, GLM 0.91, CTA 0.94 and ANN 0.96. For the 1×1 km to 64×64 km comparison, the values were GAM 0.82, GLM 0.77, CTA 0.74 and ANN 0.76. This decline indicates a divergence in area identified as range between the finer and coarser grid sizes.

4. DISCUSSION

Landscape ecologists and geographers have called for multiple tests to evaluate the impact of scale on analyses of landscape studies (Qi & Wu 1996). Because we tested the spatial outputs from SDMs, as well as measures of SDM performance, our approach differed

from other studies assessing SDM performance. Previous work has focused on identifying what model produced the highest performance statistics (Elith *et al.* 2006; Lawler *et al.* 2006), and when scale was used as a test factor the response variable was model performance rather than model spatial output. In a study of scale effects on SDMs using model performance statistics, Guisan *et al.* (2007) found that SDMs were not greatly affected by a 10-fold change of grid size. This study examined SDM dynamics across a 64-fold change in scale, and found that SDM model accuracy and spatial output agreement declined when grid size increased, with the decline generally accelerating between 8 and 16 times the initial grid size.

We also found that species' SDM-derived spatial distributions were not equivalent across grid sizes, between model types or between species with different distribution characteristics. Model divergence suggests that for best range size representation from SDMs, operational grid size should be limited, with cut-offs for narrow range size species at 4×4 km, between 4×4 and 8×8 km for intermediate range size species and between 8×8 and 16×16 km for broad range size species.

SDMs run on large size grids are being used in many biological forecasts of species' response to climate change (e.g. Bakkenes *et al.* 2006). Conservation planners need to assess the scale of such SDM outputs when developing conservation landscape designs to enhance protection of future suitable habitats and dispersal corridors. Particularly for species with ranges under 90 000 km², there is a chance of identifying inappropriate regions if predictor data used are at or greater than 50×50 km (approx. 0.48° latitude). These findings are relevant for climate change studies, since global and regional climate model (RCM) outputs are generally at scales greater than 100 km and 30–70 km, respectively. Thus, while both global climate model (GCMs) and RCMs may be suitable for SDM examination of biodiversity trends at ecoregional or continental scales, the use of these data in identifying suitable reserves to shelter species under future climate change risks the misidentification of such locations. Statistical downscaling of GCMs for use in SDM conservation planning may provide a solution to this scale problem. For current SDM simulations, these results suggest the use of 1 km² grid sizes, since 1 km² climatologies are available globally.

The authors would like to thank US National Forest Service, US National Park Service, California State Department of Fish and Game, and Susan Harrison (serpentine) for the plot data. W.T. acknowledges support from IFB-ANR DIVERSITALP (ANR_07_BDIV_014-02) and by EU_FP6 ECOCHANGE (GOCE-CT-2003-506675).

- Bakkenes, M., Eickhout, B. & Alkemade, R. 2006 Impacts on different climate stabilization scenarios on plant species in Europe. *Global Environ. Change* **16**, 19–28. (doi:10.1016/j.gloenvcha.2005.11.001)
- Elith, J. *et al.* 2006 Novel methods improve prediction of species distributions from occurrence data. *Ecography* **29**, 129–151. (doi:10.1111/j.2006.0906-7590.04596.x)

- Guisan, A., Graham, C. H., Elith, J. & Huettmann, F. 2007 Sensitivity of predictive species distribution models to change in grain size. *Divers. Distributions* **13**, 332–340.
- Guisan, A., Broennimann, O., Engler, E., Vust, M., Yoccoz, N. G., Lehmann, A. & Zimmermann, N. E. 2006 Using niche-based models to improve sampling of rare species. *Conserv. Biol.* **20**, 501–511. (doi:10.1111/j.1523-1739.2006.00354.x)
- Guisan, A. & Thuiller, W. 2005 Predicting species distributions: offering more than simple habitat modeling. *Ecol. Lett.* **8**, 993–1003. (doi:10.1111/j.1461-0248.2005.00792.x)
- Hickman, J. C. E. 1993 *The Jepson manual: higher plants of California*. Berkeley, CA: University of California Press.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. & Jarvis, A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. (doi:10.1002/joc.1276)
- Kvamme, K. L. 1985 Determining empirical relationships between the natural environment and pre-historic site locations: a hunter-gatherer example. In *For concordance in archeological analysis* (ed. C. Carr), pp. 208–238. Kansas City, KS: Westport.
- Lawler, J. J., White, D., Neilson, R. P. & Blaustein, A. R. 2006 Predicting climate-induced range shifts: model differences and model reliability. *Global Change Biol.* **12**, 1568–1584. (doi:10.1111/j.1365-2486.2006.01191.x)
- Qi, Y. & Wu, J. 1996 Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelational indices. *Landscape Ecol.* **11**, 39–49. (doi:10.1007/BF02087112)
- Rodriguez, J. P., Brotons, L., Bustamante, J. & Seoane, J. 2007 The application of predictive modelling of species distribution to biodiversity conservation. *Divers. Distrib.* **13**, 243–251.
- Schwartz, M. W., Iverson, L. R., Prasad, A. M., Matthews, S. N. & O'Connor, R. J. 2006 Predicting extinctions as a result of climate change. *Ecology* **87**, 1611–1615. (doi:10.1890/0012-9658(2006)87[1611:PEAARO]2.0.CO;2)
- Thuiller, W. 2003 BIOMOD-optimizing predictions of species distributions and projecting potential future shifts under global change. *Global Change Biol.* **9**, 1353–1362. (doi:10.1046/j.1365-2486.2003.00666.x)
- Viers, J. H., Thorne, J. H., & Quinn, J. F. 2006. CalJep: a spatial distribution database of CalFlora and Jepson plant species. *San Francisco Estuary and Watershed Science* **4**, 1–18. See <http://repositories.cdlib.org/jmie/sfews/vol4/iss1/art1>.