Chilean Journal of Statistics
Volume 13, Number 2 (December 2022), Pages 133–153
DOI: 10.32372/chjs.13-02-01

Spatial Statistics
Research Paper

Spatial autocorrelation and unorthodox random variables: the uniform distribution

Daniel A. Griffith[*]

School of Economic, Political, and Policy Sciences, University of Texas, Dallas, United States

(Received: 04 February 2022 · Accepted in final form: 21 April 2022)

Abstract

By definition, original Besag-conceived spatial auto-random variables incorporate an autoregressive spatial lag term (i.e., the sum/average of nearby attribute values) to characterize geospatial data. Very common variates include the auto-normal, auto-logistic, and auto-binomial; less common ones include the auto-beta and auto-multinomial. Some of these specifications can capture the full range of spatial autocorrelation, and others cannot. These latter variates are unorthodox in their nonconformist restrictions to either only positive or only negative spatial autocorrelation domains. The literature already offers successful modifications of the auto-Poisson and auto-negative binomial, two popular random variables for describing counts, but neither of which can encapsulate positive spatial autocorrelation. The literature dismissively mentions the auto-exponential variate, which cannot accommodate negative spatial autocorrelation situations. Meanwhile, the literature lacks any discussion about auto-uniform random variables, with implications especially from point pattern analysis publications that they solely refer to complete spatial randomness. The purpose of this paper is to postulate a productive and viable spatialized continuous uniform distribution specification that easily extends to its corresponding discrete version. A standard benchmark location-allocation simulation experiment for a simple $p = 1$ median problem, a spatial optimization circumstance that illuminates bivariate spatial median properties, illustrates its practical applicability.

Keywords: Auto-model · Moran eigenvector spatial filtering · random effects · spatial autocorrelation · uniform distribution

Mathematics Subject Classification: 62H11.

1. Introduction

Explicit spatial autocorrelation (SA) concept formation originated in the early part of the twentieth century (Griffith, 2012) with Cliff and Ord (1973) motivating its popularity within the context of spatial statistics and quantitative geography in the late 1960s, followed by Paelinck and Klaassen (1979) initiating a similar promulgation in the context of spatial econometrics a decade later, with Anselin (1988) subsequently motivating its popularity within that specialty. These two developments operated with the inverse spatial covariance

[*] Corresponding author. Email: dagriffith@utdallas.edu.

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

matrix, which is some function of a spatial weights matrix.[1] In parallel, geostatistics emerged as an advancement in the SA realm (Ecker, 2003) directly operating with the spatial covariance matrix itself. This earlier literature is replete with auto-normal spatial counterparts to the normal probability model, primarily because, commencing with de Moivre in 1733 (Steigler, 1986), the statistics discipline has a prominent normal curve (i.e., Gaussian distribution) theory based analysis history prior to Nelder and Wedderburn's formalizing and implementation of generalized linear model (GLM; McCullagh and Nelder (1989)) theory in the early 1970s (Hilbe, 2014). The relative simplicity of univariate and multivariate normal curve theory (e.g., Lohnes and Cooley, (1968)) mathematical statistics helped preserve its sustained prominence, with the advent of its affiliated normal approximation power transformation technique (Box and Cox, 1964) extending its suitability to many of the hundreds of other univariate random variable (RV) distributions that exist (e.g., Johnson et al. (1994/95)). These numerous distributions have a complex interrelationship structure (Leemis and McQueston, 2008). However, normal approximation analyses suffer from some degree of misspecification error (Griffith, 2013), wheter trivial, moderate, or severe, which helped provoke the eventual discovery and usage of GLM theory, which this paper exploits.

The geospatial sciences mostly deal with just the following few of the hundreds of possible univariate RVs: Bernoulli, beta, binomial, multinomial, negative binomial (NB), normal, lognormal, Poisson and uniform. Normal curve theory treats continuous interval/ratio measurement scale georeferenced RVs over a $(-\infty, \infty)$ support domain, with Box-Cox power transformations and other normal approximations artificially expanding its practical applicability to limited domains such as the truncated support $[0, \infty)$—the lognormal RV signifies a special case of this truncation linkage. Pollution contamination levels are a geospatial attribute exemplifying it. NB and Poisson GLM theory treat non-negative integer counts over a $[0, \infty)$ support domain; they have a natural lower bound restriction. Areal unit (e.g., government agency polygons such as census tracts) post-stratified demographic data are a geospatial attribute exemplifying them. Bernoulli (aka logistic) RVs are dichotomous, almost always measured by the parsimonious binary set {0,1} in non-physics situations, and frequently denote presence/absence. The appearance of a disease incidence in a location is a geospatial example of it. Binomal RVs effectively are percentages (i.e., 100 times categorized counts divided by their category totals) that span the interval [0,100], are an aggregation—and hence one type of generalization—of Bernoulli RVs, and have an obvious limit at both ends of their natural numbers measurement interval, respectively 0 and n_T (i.e., the maximum number of objects for each category)— in other words, their support domain is the set $\{0, 1, \ldots, n_T\}$, which converts to percentages by dividing its entries by n_T and then multiplying by 100; their prefix bi- refers to their underlying dichotomy. A population cohort percentage residing in different locations is a geospatial example of it. Further generalizing Bernoulli RVs produces a multinomial RV, which is polychotomous (re its multi- prefix), and whose individual groupings (i.e., in a specific group versus not in that specific group) relate to binomial RVs. A remotely sensed satellite image pixel land use classification is a geospatial example of it. Unlike a binomial RV, which involves discrete combinations of counts, a beta RV is continuous over a [0,1] support domain. The proportion of a given heavy metal concentration in a GPS-tagged soil sample is a geospatial example of it. Finally, a uniform RV enters geospatial work in various ways, including equally likely geographic probability sampling, and controlling the geographic distribution of some phenomenon in a simulation experiment often for benchmark reasons. Under the auspices of a spatial stationarity as-

[1] A spatial weight matrix, say $\boldsymbol{C}$, is a n-by-n matrix whose row and column labels virtually always are the same sequence of areal unit names, solely for the convenience and mostly by convention. In its simplest manifestation, namely a binary non-negative matrix, its entries are 0 and 1, where 0 denotes that the row and column areal units are not, whereas 1 denotes that they are, neighbors/adjacent/nearby. In other words, it is a collection of n n-by-1 indicator variable vectors.

sumption, if a RV has both a mean and a finite variance, then the locational average of simulation replications across a geographic landscape yields a uniform distribution of its mean value—the Cauchy RV is perhaps the most famous exception, because its mean and variance are undefined (Bagui et al., 2013). This simulation context for spatial statistical numerical experiments constitutes a primary application theme of this paper.

Besag (1974) conceived of and coined the term auto-model to label RV equation descriptions that have the dependent variable Y on both sides of their equal sign (=), with the right-hand side term usually being the spatial lag matrix expression $\boldsymbol{WY}$—where matrix $\boldsymbol{W}$ is the row-standardized (i.e., each row sums to one) version of a spatial weights matrix $\boldsymbol{C}$—calculating average neighboring attribute values and emphasizing the previously mentioned auto-normal specification. These formulations actually correspond to joint statistical distributions of n conditional marginal univariate RVs (Kaiser and Cressie, 2000). Besag (1974) devoted a profuse amount of his text to the auto-normal model, while not only commenting on especially the auto-logistic specification—which has been the subject of acute scrutiny more recently (Bardos et al., 1974)—but also the auto-binomial, auto-Poisson, and auto-exponential specifications. Cressie (1991), among others, furnishes considerably more detail about the auto-logistic and auto-binomial models, likewise discussing the auto-beta (p. 440), gamma (p. 439), NB (p. 432) and Poisson (pp. 427-429) RVs. Besag (1974) dismissed the auto-exponential and gamma specifications because he views their conditional probability structures as lacking intuitively appealing mathematical expressions. Furthermore, these RVs seem to be of little practical interest to most geospatial researchers, and hence are not subjects of this paper. Meanwhile, Besag (1974) highlights that an auto-Poisson RV is unorthodox because it is incapable of describing positive SA situations. This feature transfers to its extra-Poisson variation auto-NB relative by the very nature of its equivalency to the auto-Poisson with a gamma-distributed mean parametric mixture (e.g., a probabilistic combination of multiple probability distributions such that all but one of them describe the distribution of a different parameter in the designated other) RV. This precise construct suggests that the auto-gamma may deserve more attention from spatial scientists than it currently receives. The auto-multinomial (Kavousi et al., 2011) and auto-beta (Zikariene and Ducinskas, 2021) RVs have received more consideration in recent years (e.g., Hu et al. (2020)). As one of various explorations into more realistic real world auto-model applications (e.g., Bolin et al. (2013)), this paper contributes an additional examination to this literature of a special auto-beta specification case in order to explore the pair of unorthodox uniform RVs with regard to their relationship to SA.

2. Positive SA and the unorthodox auto-Poisson and auto-NB RVs

Spatial scientists often contend that positive SA is by far the most common spatial dependence materializing in real world phenomena. Accordingly, the auto-Poisson and auto-NB RVs are unorthodox as well as wanting because they are unable to depict this situation, even though most geographic distributions of counts data defy this limitation and display moderate-to-strong positive SA. In addition, some suitable binomial distribution (i.e., n_T is sizeable and the probability of an event occurring is exceptionally small) provides a good approximation to a particular Poisson distribution; this contention rests on the property of a Poisson RV being the limiting case of a binomial RV (Le cam, 1960). Similarly, some fitting normal distribution also furnishes a good approximation to a particular Poisson distribution (Cheng, 1948). These two relationships imply that a proper spatial Poisson probability mass function accounting for positive SA should be feasible. The literature reports three successful respecifications that remedy this auto-Poisson shortcoming.

Besag et al. (1991) formulated the first spatial Poisson model that efficaciously accounts for positive SA in georeferenced counts data by invoking statistical random effects mixed models theory (Pan et al., 2020) in which they introduce a Gaussian random effects term comprising a spatially structured (SSRE)—proficient in accounting for positive SA—and unstructured (SURE) component free of SA. The mixed models maneuver stratagem is an efficient way to properly handle correlated data. This formulation essentially converts a constant conditional mean response to a varying intercept term. To wit, the baseline for each areal unit may differ, with the average of the n individual intercept terms theoretically equal to the global constant conditional mean. Therefore, some individual intercepts are greater, and others are less, than the single global value. These fluctuations reflect heterogeneity, correlatedness, omitted variables, and other data and/or mathematical expression corruptions and noise. In Besag et al. (1991), the SSRE component is an auto-normal (specifically a conditional autoregression) term, and their estimation method is Bayesian implemented with Markov chain Monte Carlo (MCMC) techniques (Gilks et al. (1996); Casella and George (2008)) requiring numerically intensive calculations. Their shift from a frequentist standpoint directly extends to the auto-NB specification.

Alternatively, Kaiser and Cressie (1997) proposed truncating an auto-Poisson probability model support domain through data Winsorizing (i.e., a statistical transformation that constrains count data in order to eliminate the possibility of extreme observed values that qualify as potentially spurious outliers). Doing so causes the sum of all possible probabilities to be less tan one (their trick is to decrease this sum by a trivial amount), violating a fundamental axiom of probability theory. The trade-off is that this model specification allows positive SA to materialize in joint multivariate count data by utilizing modified conditional univariate distributions that are Winsorized Poisson probability mass functions. Again, estimation proceeds with MCMC techniques, retaining the numerical intensity associated with the original auto-Poisson specification. One advantageous consequences is that, unlike Besag's original auto-Poisson form, their model is capable of accounting for either positive or negative SA among georeferenced counts while retaining Besag's original auto- formulation. A rather cumbersome trait it display is that the interval length of the positive SA feasible parameter space often tends to be quite small. This reconstruction directly extends to the auto-NB specification, too.

Griffith (2002) devised a third way to account for positive SA in these two GLM versions by exploiting Moran eigenvector spatial filtering (MESF) theory (Griffith, 2003). This approach first extracts n synthetic variates called eigenvectors (see Abdi (2007) for a reader-friendly overview of these mathematical entities), each of size n-by-1, from the modified spatial weights matrix $(\boldsymbol{I} - \mathbf{1}\mathbf{1}^\top/n)\boldsymbol{C}(\boldsymbol{I} - \mathbf{1}\mathbf{1}^\top/n)$ appearing in the numerator of the Moran coefficient (MC) index of SA, where $\boldsymbol{I}$ is the n-by-n identity matrix, $\mathbf{1}$ is the n-by-1 vector of ones, and superscript T denotes the matrix transpose operator. Each eigenvector represents a distinct SA map pattern for which each of its n elements links to its assigned row label areal unit in its parent spatial weights matrix. Next, a selected subset of these eigenvectors best describing SA in a given response variable Y becomes a surrogate for it; screening by their nature and minimal degree—which their accompanying eigenvalues index—coupled whit stepwise regression routines accomplishes this selection task. Moreover, eigenvectors in the judiciously chosen subset become additional covariates in a GLM regression, with their resulting linear combination being a constructed eigenvector spatial filter (ESF). This implementation is analogous to incorporating just the SSRE component of a random effects term (e.g., Besag et al. (1991)) into the mean response equation. Estimation commonly is with maximum likelihood techniques, often implemented using iteratively reweighted least squares; in other words, standard GLM estimation. The greatest advantage of MESF is that it dramatically reduces the numerical intensity of accounting for SA by switching from MCMC to conventional estimation procedures while accommodating not only positive and negative SA, but also mixtures of these two. One of its weaknesses is that the number of candidate synthetic variable eigenvectors for constructing an ESF increases with n (but at a decreasing rate).

All three of these methodologies rewrite the RV mean (e.g., see Ferrari and Cribari-Neto (2004)) such that it can account for positive SA. In addition, this is the general viewpoint conveyed in Kaiser and Cressie (2000), Hardouin and Yao (2008) and Hu et al. (2020), for the auto-beta kind of conceptualization. This facet is noteworthy when considering the final unorthodox RV listed in this paper's introduction, namely the uniform distribution, the topic of the next section.

3. Positive SA and auto-uniform RVS

The only conspicuous parameterization a uniform RV may have concerns its extremes (say Ψ and/or θ), such that its continuous case probability density function, for example, is written as

$$f(y) = \frac{1}{(\theta - \Psi)}, \quad \Psi \leq y \leq \theta$$

when both are unknown. A diverse assortment of estimators exists for these two parameters (e.g., Jabeen and Zaka, (2020)). Their maximum likelihood estimators (MLEs) are simply x_1 and x_n, the extreme values of a sample. Their individual uniformly minimum variance unbiased estimators (MVUEs) are, conditional on the other being known, $[(n{+}1)x_1{-}\theta]/n$ for Ψ and $[(n{+}1)x_n{-}\Psi]/n$ for θ. This collective outcome implies the following iterative/recursive calculations, alternately assuming one parameter is known to estimate the other, employing the following two estimators that are functions stricly of x_1, x_n and n:

$$\begin{aligned}
\text{initial estimates :} \quad & \Psi_0 = x_1, \quad \text{and} \quad \theta_0 = x_n \\
\text{iteration} \quad & \tau(= 1, 2, \ldots,) \\
\text{estimates :} \quad & \Psi_{\tau+1} = \theta_\tau + (x_1 - \theta_\tau)(n+1)/n \\
& \theta_{\tau+1} = \Psi_\tau + (x_n - \Psi_\tau)(n+1)/n
\end{aligned}$$

In addition, the method of moments estimators (MMEs) for these two parameters are $\bar{y} {\pm} s\sqrt{3}$ ($\Rightarrow \hat{\theta} = 2\bar{y}$ when Ψ is known), where s denotes the sample standard deviation. Implications include $\hat{\mu} = (x_1 + x_n)/2$ (MLEs), $\hat{\mu} = [(x_1 + x_n)(n+1)/n - (\Psi_\infty + \theta_\infty)/n]/2$ (MVUEs), and, statistically speaking, that the MMEs are inadmissible[1], mean response parameter estimator expressions lacking any apparent valid operational opportunity to introduce a bone fide spatial lag term; for the first two estimators, only two of n terms can be cast as functions of weighted averages of their neighboring values in an auto- specification. However, the beta conditional conception (James, 1975), which dovetails with the preceding mixed models scheme and promoted by Kaiser and Cressie (2000) and by Hardouin and Yao (2008), offers useful insights.

The usual simulation experiment situation assuming a uniform RV sets $\Psi = 0$ and $\theta = 1$, a standard uniform RV whose extremes easily rescale and translate to any other pair of real numbers. Likewise, the following beta RV may characterize this unit interval support

[1] As a counterexample to this estimator's admissibility, nearly 40% of the output from a simple 10,000 replications simulation experiment with samples of size $n = 100$ drawn from a continuous uniform distribution over the interval [5,21] contains MMEs that do not exceed one or the other of their sample extremes, with roughly 10% of these samples failing to exceed both of their sample extremes.

domain stated as

$$f(y) = \frac{\Gamma[2y]}{(\Gamma[y])^2} y^{\gamma-1}(1-y)^{\gamma-1}, \quad 0 \le y \le 1 \quad \text{and} \quad \gamma > 0 \quad \text{is a shape parameter,}$$

where this function has only a single parameter (i.e., its two shape parameters are identical and equal to γ) to ensure the uniform distribution symmetry property, and $\gamma = 1$, whose substitution into the probability density function renders the expression given by

$$f(y) = \frac{\Gamma[2]}{(\Gamma[1])^2} y^0 (1-y)^0 = 1, \quad 0 \le y \le 1, \tag{3.1}$$

which is the unit interval form of a continuous uniform RV. Its mean and variance are the respective constants 1/2 and 1/12, quantities failing to provide an entrance into this probability density function for SA like larger values of γ supply [e.g., $\mu = \gamma/2$ and $\sigma^2 = 1/4(2\gamma+1)$]. Its theoretical frequency distribution is flat, preventing the standard SA variance inflation from materializing. However, one of its conventional probability density function revisions may be written as[1]

$$f(z) = \frac{1}{1^2}\left(\frac{1}{1+\exp(-z)}\right)^{1-1}\left(1 - \frac{1}{1+\exp(-z)}\right)^{1-1} = 1, \quad -\infty \le z \le \infty \tag{3.2}$$

where $z = LN[Y/(1-Y)]$—with Y being the RV in Equation (3.1)—the standard logit transformation historically employed in logit-linear regression assuming normally distributed errors. This is the same compound probability distribution variable argument form as for the auto-Bernoulli (i.e., logistic), binomial, and multinomial RVs. One assumption governing variate Z, in keeping with the beta conditional tradition (James, 1975), is that, besides being a logarithmic quantity, its underlying distribution is beta(γ,γ), with $\gamma \to \infty$ resulting in this beta distribution increasingly mimicking a normal distribution (Peizer and Pratt (1968),Pratt (1968))—e.g., at least the first four moments of both distributions match, satisfying the convergence in the r^{th} moment principle (e.g., Hoeffding (1952))—creating a parametric mixture situation resembling that promoted by Besag et al. (1991), and exploited by Griffith (2002). Furthermore, as the auto-logistic specification demonstrates, this mathematical alteration can contain a spatial lag term.

A well-known and previously noted auto-RV property induced by SA is variance inflation (also see Griffith (2011), Chun and Griffith (2018), and Hu et al. (2020)). The only way that Equation (3.2) can integrate SA and experience variance inflation is through the logistic function $1/(1+\exp(-z))$, although theoretically its zero exponent neutralizes this impact. Because the final variance is a constant (e.g., 1/12 for the standard uniform RV), the hyperprior type of RV needs to have less variance, enabling this smaller variance to inflate to the final constant variance through the presence of SA. This is a rather rare case in which the resulting posterior distribution has more variance than its input prior distribution. An approximately normal beta RV is suitable for this purpose: deflating its center while simultaneously inflating both of its tails in a complete transition from bell-shaped to perfectly flat symbolizes this variance inflation. Additionally, because variance inflation is a function of the latent positive SA's magnitude, if its parameter is $\rho = 0$, then the outcome should be a uniform distribution, whereas as $|\rho| \to 1$, the solitary beta shape parameter γ

[1] Replacing the exponent of zero by a logarithmic function in accordance with the indefinite integral equivalency $\int dy/y = LN|y|$ (Dwight, 1961) because y^0 and $(1-y)^0$ in Equation (3.1) do not produce a sensible solution, yielding $LN(y) \times LN(1-y)$—a pooled $LN[y(1-y)]$ is infeasible because it gives negative quantities—where LN denotes the natural logarithm, fails to render a uniform distribution.

should increase. Appendix A expounds upon the theoretical framework underpinning this conceptualization in some detail.

4. A numerical example: a spatially autocorrelated continuous uniform RV

The objetive pursued here is to generate a set of pseudorandom numbers spatially distributed across a geographic landscape that globally closely conform to a uniform distribution, and that contain visibly apparent SA in their map pattern. The simulation experiment executed for this section utilizes either a 40-by-40 regular square tessellation of pixels (a la the High Peak empirical example in Bailey and Gatrell (1995)) or the 2010 Dallas-Ft. Worth (DFW) Metropolitan Statistical Area (MSA) census tracts irregular lattice, a rook's chess move type of adjacent neighbors definition, the row-standardized spatial weights matrix version of this binary 0-1 matrix, an initiating random geographic distribution of sampled beta RV values (a randomly permuted representative systematic sample from across the uniform distribution support based upon Blom (1958) quantile points $(r - 3/8)/(n + 1/4), r = 1, 2, \ldots, n$), and 10 (i.e., essentially the number recommended by practical guidelines) frequency distribution bins.

Although these two specimen geographic landscapes are modest in their areal unit sizes, the methodology and conceptualizations inaugurated in this paper extend to virtually any size regular square tessellation (see Griffith and Chun (2019)), as well as to irregular surface partitonings comprising many thousands of polygons (e.g., Griffith (2018)).

4.1 Preliminary simulation experiment inputs

Figure 1 portrays the relationship between the SA parameter[1] ρ, through its much simpler and more readily available MC computation[2], and beta hyperparameter γ for this experimental design. This association summarizes data from an idealized auto-beta type of reconnaissance (Figure 1a); the second reconnaissance exploited an empirical geographic landscape (i.e., 2010 DFW MSA census tracts), furnishing confirmatory data (Figure 1b). Meanwhile, Figure 2 reveals that for the most commonly encountered georeferenced socio-economic/demographic attribute variable degree of SA (i.e., MC = 0.5, indicating a moderate level), the mixed model SSRE type of random effects term does not closely conform to a bell-shaped curve. In contrast, high SA levels like those found in remotely sensed imagery (i.e., MC = 0.9) closely mimic, and essentially are indistinguishable from, a bell-shaped curve. As an aside, the critical feature of a random effects RV seems to be symmetry, with the magnitude of kurtosis (i.e., peakedness) primarily managing the prerequisite realized degree of SA. Because both the input to and output from this mixture is a beta RV, this situation is reminiscent of that for traditional conjugate priors.

Figure 1 suggests that this γ-MC relationship is reasonably robust against varying surface partitionings. Future research needs to address this hypothesis. In addition, future research

[1] The spatial linear operator employed in this analysis is from the spatial simultaneous autoregressive (SAR), also known in the spatial econometrics literature as the spatial error, specification. Because the regression equation contains no covariates, in this simulation experimental context, this specification is equivalent to the spatial autoregressive response (AR), also known in the spatial econometrics literature as the spatial lag, specification.

[2] Luo et al. (2018) discuss the relationship between ρ and the MC. The two illustrative datasets used in this paper deliver the following approximations:

$$\text{40-by-40 regular square tessellation:} \hat{\rho} \approx -1.76317 + 2.76317/(1 - MC^2), R^2 \approx 1$$

$$\text{2010 DFW MSA:} \hat{\rho} \approx -0.75807 + 1.75807/(1 - MC^2), R^2 \approx 1$$

needs to establish whether or not any higher SA degree deviations from the theoretical trendline change with n; the difference here of 284 areal units may be sufficient to hypothesize that it does not, and that, rather, it may be a function of irregular surface partitioning effects. A similar hypothesis concerns the maximum MCs, which here are 1.018 and 1.175, respectively, for the 40-by-40 regular square tessellation and the 2010 DFW MSA geographic landscapes.

4.2 Preliminary simulation experiment outputs

The goal of the knowledge formation simulation experiments outlined in this paper is to generate globally uniform RV geographic distributions that display approximately targeted levels of positive SA. One benefit of this capability is the ability to engage in exploratory spatial analysis simulation studies that assume uniformly distributed georeferenced attribute values; SA is a requisite property of virtually all geospatial attribute data. The previous section documents that introducing SA into Equation (3.2) is no easy feat, appearing almost impossible at first glance, branding it an unorthodox spatial RV. The preceding subsection documents that incorporating SA into a beta RV through a SSRE term progressively transforms it from a uniform toward a bell-shaped frequency distribution as the degree of positive SA increases. Figure 3 visualizes specimen frequency distributions output from performing its itemized workflow.

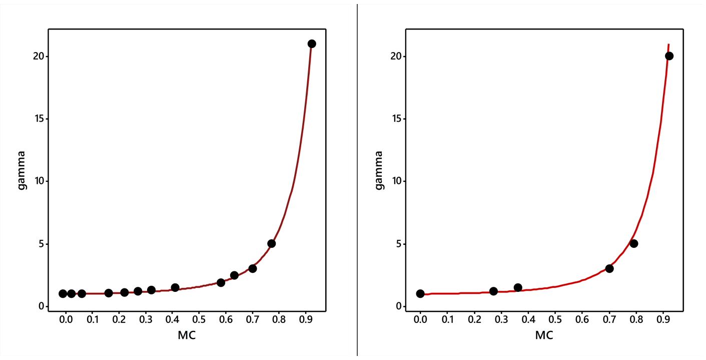

Figure 1. Relationships between the standard equal shape parameters, γ, beta RV and the induced SA level ($0 \leq \rho \leq 0.99$) MC, rook's adjacency; the gray curved lines denote the theoretical trendline given by Equation (6.6) appearing in the appendix, calibrated with data generated by Equation (6.7), and filled black circles denote observed values from numerical analyses—for both cases, the accompanying nonlinear regression pseudo-R^2 is approximately 0.99. Left (a): a regular square tessellation ($n = 40 \times 40 = 1,600$). Right (b): the 2010 DFW MSA census tracts ($n = 1,314$).

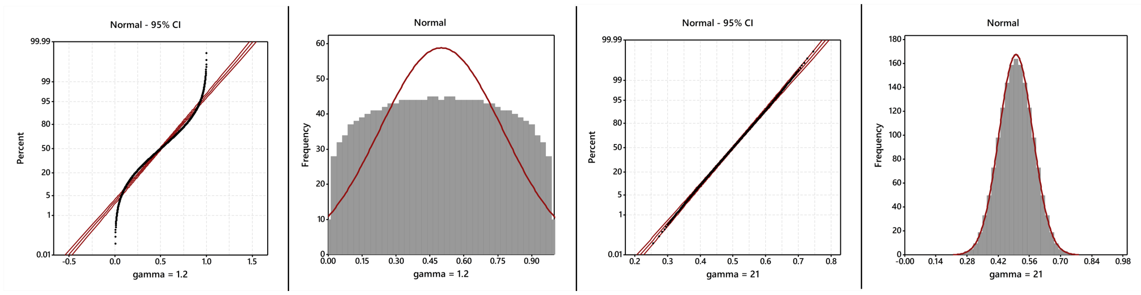

Figure 2. Comparisons of selected independent and identically distributed beta RVs with a normal RV. Left (a): normal quantile plot for $\gamma = 1.2$ [MC $\approx$ 0.5, Probability (Shapiro-Wilk) < 0.001]. Left middle (b): $\gamma = 1.2$ beta distribution histogram with superimposed normal curve. Right middle (c): normal quantile plot for $\gamma = 21$ [MC $\approx$ 0.9, Probability (Shapiro-Wilk) $\approx$ 0.997]. Right (d): $\gamma = 21$ beta distribution histogram with superimposed normal curve.

Figures 3a and 3b portray the representative systematic sample with zero SA; the uniform distribution is unchanged. These frequency distributions contain negligible variation across their 10 bins because the sample is not, although its allocation to the geographic landscape pixels is, random. This is the SA context characterizing all geospatial simulation experiments to date that involve a bivariate uniform distribution (e.g., demand across a regional economic landscape). Figures 3b and 3e reflect the modest variation across bins introduced by weak SA. This fluctuation is consistent with that attributable to simple random sampling, and, in part, reflects the small variance inflation weak SA induces. Figures 3c and 3f signal two sources of variation, the first once more being SA, which the hyperprior beta distribution successfully handles, and the second being influences of an asymmetric distribution of spatial weights matrix eigenfunctions (see Section 2), which include both more negative than positive SA ones, and positive extreme SA that is considerably greater in extent than its negative extreme SA counterpart (Griffith, 2017). A redeeming feature of this latter situation is that few georeferenced socioeconomic/demographic, although most remotely sensed image, attributes have a MC value noticeably in excess of the reported magnitude here. Nevertheless, for this particular realization, $\hat{\chi}^2 \approx 11.9 < \chi^2_{9,.09} \approx 14.7$ (i.e., fail to reject the null hypothesis that the global frequency distribution of the 1,314 values does not conform to a uniform distribution in the population). In addition, all of the beta estimated shape parameter pairs $\hat{\gamma}_1$ and $\hat{\gamma}_2$, are nearly equal, corroborating the assumption of a single γ value necessary to produce both a bell-shaped and a uniform distribution. As an aside, being able to estimate these parameters from data confirms the identifiability of the new spatialized uniform RV specification. Figure 4 visualizes the six respective maps for the set of frequency distributions displayed in Figure 3 for which the cornerstone interval [0,1] encompasses all attributes values. Its perusal from left to right discloses the increasing geographic clustering of similar values with increasing positive SA. Currently geospatial simulation experiments can use just maps like those in Figures 4a and 4d. A principal aim of this paper is to expand that possibility to a continuum of more realistic maps epitomizing various degrees of positive SA (e.g., Figures 4b 4c, 4e, and 4f) through embedding rather than constrained permutations (see Figure 5).

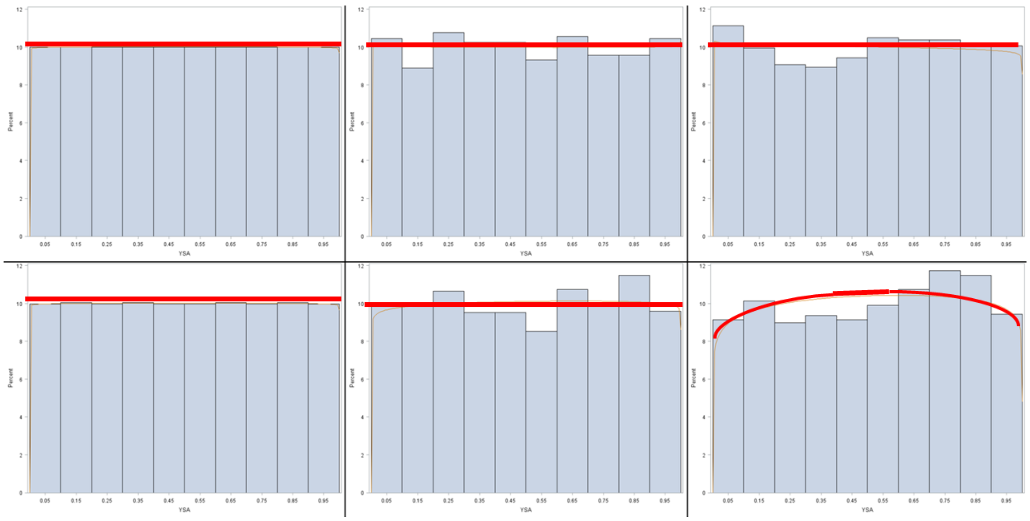

Figure 3. Simulated compound conjugate posterior uniform frequency distributions containing SA with superimposed beta distribution curves (red lines); the top row displays results for a 40-by-40 regular square tessellation, and the bottom row displays results for the 2010 DFW MSA census tracts. Top left (a): $\rho = 0.00$; MC = -0.01, $\hat{\gamma}_1 = \hat{\gamma}_2 = 1.002$. Top middle (b): $\rho = 0.50$; MC = 0.30, $\hat{\gamma}_1 = 0.997$,$\hat{\gamma}_2 = 0.998$. Top right (c): $\rho = 0.95$; MC = 0.78, $\hat{\gamma}_1 = 0.995$,$\hat{\gamma}_2 = 1.010$. Bottom left (d): $\rho = 0.00$; MC = 0.01, $\hat{\gamma}_1 = \hat{\gamma}_2 = 1.003$. Bottom middle (e): $\rho = 0.50$; MC = 0.28, $\hat{\gamma}_1 = 1.023$,$\hat{\gamma}_2 = 1.011$. Bottom right (f): $\rho = 0.95$; MC = 0.75, $\hat{\gamma}_1 = 1.081$,$\hat{\gamma}_2 = 1.054$.

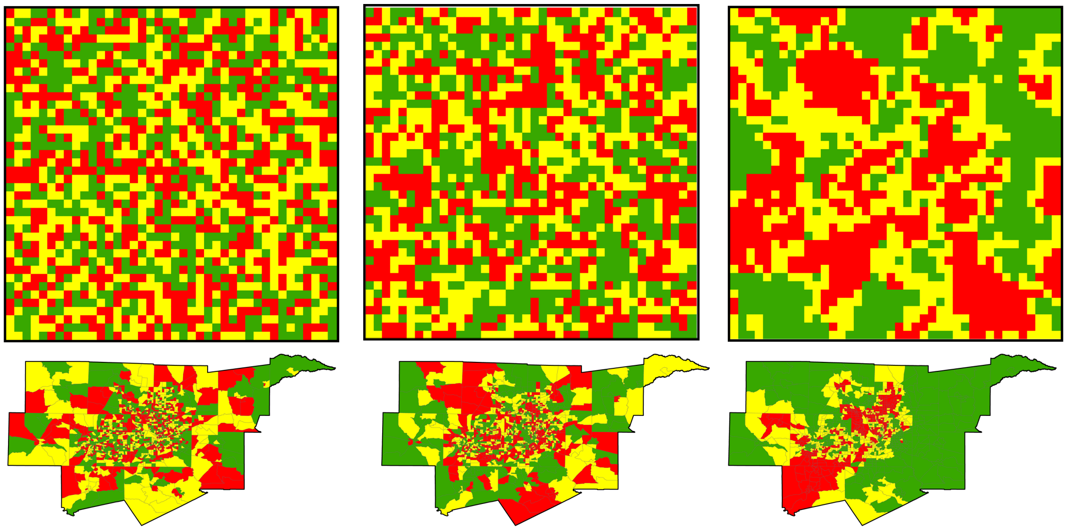

Figure 4. Specimen uniform RV geographic distributions, with green, yellow, and red respectively denoting values between 0 and 1/3 (relatively low), 1/3 and 2/3 (moderate), and 2/3 and 1 (relatively high); the top row displays results for a 40-by-40 regular square tessellation, and the bottom row displays results for the 2010 DFW MSA census tracts. Top left (a): MC = -0.01 and Geary Ratio (GR) = 1.02. Top middle (b): MC = 0.30 and GR = 0.70. Top right (c): MC = 0.78, and GR = 0.22. Bottom left (d): MC = 0.01 and GR = 0.98. Bottom middle (e): MC = 0.28 and GR = 0.73. Bottom right (f): MC = 0.75 and GR = 0.24.

4.3 A testbed demonstration

The interface between SA and spatial optimization (Griffith et al. (2022); Griffith et al. (2023)) furnishes a spatial statistics relevant as well as a sound laboratory setting for illustrating rigorous, transparent, and replicable utilization testing of the new spatially autocorrelated uniform RV specification. The specific spatial analysis problem here is to calculate a solitary spatial median (Small (1990); Ninimaa (1995); Eftelioglu (2017); i.e., the simplest Weberian location problem of economic geography and location-allocation fame, and a basic descriptive centrographic spatial statistic) given a uniform distribution of demand. The spatial mean and the standard distance centrographic measures quantifying variations in the computation of this location across a set of simulated geographic landscapes illuminates that acknowledging the prevailing degree of positive SA matters.

The simulation experiments here begin with unrestricted random sampling from a uniform distribution by location, differing from the preceding experiments because those draw a single representative systematic sample from across the full range of a uniform distribution, and then randomly allocate the carefully selected sample values to locations by randomly permuting them. Given that the intention of this exercise is to show that non-zero SA in a uniform RV has a pronounced impact on spatial analysis outcomes, these experiments executed only 100 replications, enough to expose whether or not a discrepancy exists, but not truly enough for suitable precision to accurately quantify the size of that discrepancy. In other words, they are more exploratory than explanatory. This is a proof of concept demonstration project, not an attempt to establish exact magnitudes of differences attributable to SA. Table 1 tabulates selected substantiating results from these experiments. One treatment factor manipulated across experiments is the degree of SA, which the MC and the GR index; a second is the prevailing surface partitioning. Control factors include the beta RV, its versions with equal shape parameters, and the mean and variance parameters of a standard uniform RV (i.e., 1/2 and 1/12). The chance aspect is unrestricted random sampling from a beta distribution with an assigned value to a designated single shape parameter, γ. Corresponding findings appearing in Figure 4 and Table 1 are coherent, as are the mean and variance of the SA injected posterior uniform distributions vis-à-vis their theoretical

Table 1. Simulation experiment summary statistics for the $p = 1$ (bivariate) spatial median location-allocation problem solution for SA embedded with an autoregressive spatial linear operator; 100 replications.

statistic	40-by-40 regular square tessellation			2010 DFW MSA census tracts		
	$\rho = 0.00$	$\rho = 0.50$	$\rho = 0.95$	$\rho = 0.00$	$\rho = 0.50$	$\rho = 0.95$
$\bar{y}(\mu = 0.5)$	0.499	0.503	0.501	0.500	0.499	0.508
$s_y(\sigma = 0.289)$	0.289	0.289	0.302	0.288	0.286	0.290
$\overline{\mathrm{MC}}$	−0.001	0.281	0.803	−0.000	0.242	0.779
s_{MC}	0.018	0.020	0.025	0.017	0.024	0.031
$\overline{\mathrm{GR}}$	1.000	0.718	0.193	0.998	0.755	0.213
s_{GR}	0.018	0.020	0.025	0.018	0.024	0.029
$\overline{U}$	20.518	20.507	20.397	-96.911	-96.912	-96.911
s_{U}	0.222	0.410	1.628	0.007	0.012	0.036
$\overline{V}$	20.486	20.447	20.629	32.857	32.856	32.857
s_{V}	0.197	0.419	1.574	0.003	0.006	0.025

counterparts. The average bivariate spatial medians are very close to their respective theoretically expected two-dimensional points: $(20.5, 20.5) = ((40 + 1)/2, (40 + 1)/2)$ for the regular square tessellation, and (longitude = -96.943, latitude = 32.851) for the DFW MSA; SA tends not to impact first moment calculations. However, their variability (used to compute centrographic standard distance statistics) strikingly increases with increasing SA; SA tends to spawn variance inflation.

This increased locational variance occurs because relatively large weights (i.e., the georeferenced attribute values) tend to form influential geographic clusters under the influence of positive SA, which, in turn, pull a bivariate spatial median optimal location toward them. Each Gaussian random field generates positive SA clusters that randomly disperse across a geographic landscape, such that the average of their replications is the bivariate midpoint of the given surface (a la the Law of the Large Numbers). A particular spatial linear operator, $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}$, preserves the approximate level of SA within random sampling error bounds, but not any particular map pattern realization. In contrast, simulation experiments by means of MESF maintain a specific created ESF map pattern, as well as an approximate SA level, across replications (Griffith, 2017). Therefore, as Table 2 discloses, the collective pull of more peripheral large weight geographic clusters preserved by an ESF moves the average bivariate median away from a given landscape's midpoint (i.e., the bivariate spatial median in the presence of zero SA in the weights is the center of gravity for a geographic landscape) and toward such a map pattern's prominent weight clusters, markedly reducing its standard distance through an ESF's perpetuation of the same map pattern from one replication to another. In other words, results become conditional on an observed map pattern.

Impacts of SA on their answers is an important feature for location-allocation problems, and alludes to a relationship between SA hot spots and these location-allocation solutions (Griffith (2021); Griffith et al. (2022) Griffith et al. (2023)). Consequently, as anticipated and stated previously, SA in a uniform RV matters! This state of affairs is completely ignored in the literature.

5. Estimating SA in uniform RVs with MESF methodology

MESF furnishes one way to avoid inserting spatial lag terms into spatial statistical model specifications, enabling an unrestricted utilization of GLM theory. This circumstance ex-

Table 2. Simulation experiment summary statistics for the $p = 1$ (bivariate) spatial median location-allocation problem solution with 100 replications.

statistic	40-by-40 regular square tessellation			2010 DFW MSA census tracts		
	weak SA	moderate SA	strong SA	weak SA	moderate SA	strong SA
ESF eigenvectors	1st 10	1st 10	1st 10	1st 10	1st 10	1st 10
beta γ, ESF weight	1.2, 8	2.0, 17	3.5, 21	1.2, 8	2.0, 17	3.5, 20
: $\bar{y}(\mu = 0.5)$	0.502	0.514	0.526	0.500	0.500	0.500
$s_y(\sigma = 0.289)$	0.284	0.286	0.287	0.285	0.287	0.280
$\overline{\mathrm{MC}}$	0.140	0.529	0.747	0.164	0.552	0.748
s_{MC}	0.021	0.017	0.012	0.022	0.018	0.012
$\overline{\mathrm{GR}}$	0.861	0.480	0.266	0.857	0.503	0.323
s_{GR}	0.021	0.017	0.012	0.021	0.016	0.012
$\overline{U}$	20.209	19.770	19.442	32.882	32.906	32.912
s_{U}	0.206	0.155	0.114	0.003	0.002	0.002
$\overline{V}$	19.215	18.286	18.016	-96.902	-96.891	-96.887
s_{V}	0.195	0.119	0.083	0.005	0.003	0.002

tends to the uniform distriburion via its formulation as a beta RV. In addition, MESF methodology confirms the identifiability of the spatially autocorrelated uniform RV specification: the values of its parameters (e.g., μ and σ) are ascertainable from empirical observations with MESF estimation techniques (e.g., Table 3 reports a correct estimated mean of roughly 0.500).

Equation (3.2) divulges that the spatialized beta probability model lacks an explicit conventional autoregressive spatial lag term, deviating from Besag's inaugural blueprint. Table 3 uncovers the following SA effects exhibited as changes with increasing uniform RV MC values: an increasing number of statistically significance eigenvectors selected by stepwise regression to describe latent SA; a redistribution of eigenvector frequencies from a lesser (i.e., near 10%) to a greater (i.e., < 1%) statistical significance category; an increasing pseudo-R^2 (i.e., the squared correlation between the simulated Y values and their MESF beta regression fitted values)—redundant information in nearby weights (i.e., observed attribute values) increases with increasing SA; and, an increasing beta regression scale estimate—SA induced variance inflation. Diagnostic metrics imply that the selected eigenvector percentages in the presence of zero SA (i.e., roughly 13.4% and 12.5%) are very close to their 10% expectation; the magnitudes of their differences are not substantively significant, with one barely being statistically significant (40-by-40 regular square tessellation: $z \approx 2.53, p \approx 0.006$), whereas many quantitative researchers usually would consider the other not statistically significant (DFW MSA: $z = 1.47, p = 0.071$); this stepwise selection issue is the topic of vital contemporary research (e.g., G'Sell et al. (2016)). These computations rest on a single simulation experiment replication involving a nearflawlessly representative sample; conferring more precision upon this deviation requires thousands of additional replications. The overdispersion deviance statistics essentially match their expectations (of 1), and all intercepts essentially are the same (implying correctly that $\mu = 1/2$). Overall, then, this tabulation reinforces the contention that the procedure outlined in this paper successfully embeds SA in uniform RVs, and that identifiability accompanies this proposed spatialized uniform RV specification.

Table 3. Beta regression results for intercept-only and constructed ESFs for the Figure 3 specimen geographic distributions.

statistic		simple	with covariates	simple	with covariates	simple	with covariates
			40-by-40 regular tessellation				
intercept		0.000	0.003	−0.001	−0.001	−0.015	−0.014
scale		2.004	2.398	1.994	4.193	2.005	22.811
deviance		1	1.01	1	1.02	1.03	1.01
	< 0.01	0	6	0	93	0	216
#vectors†:	0.01-0.05	0	31	0	58	0	45
	0.05-0.10	0	29	0	37	0	13
pseudo-R^2		0	0.152	0	0.494	0	0.866
a: bivariate regression intercept		***‡	0.499	***	0.500	***	0.503
b: bivariate regression slope		***	0.245	***	0.219	***	0.177
			2010 DFW MSA ($n = 1,314$)				
intercept		0.000	−0.001	0.001	0.013	0.025	0.037
scale		2.005	2.311	2.034	3.516	2.135	16.202
deviance		1.00	1.02	1.02	1.03	1.01	
	< 0.01	0	4	0	46	0	131
#vectors*:	0.01-0.05	0	23	0	41	0	33
	0.05-0.10	0	13	0	29	0	17
pseudo-R^2		0	0.123	0	0.400	0	0.788
a: bivariate regression intercept		***	0.500	***	0.501	***	0.509
b: bivariate regression slope		***	0.247	***	0.229	***	0.171

‡ denotes not applicable

† selected from 492 (*321) candidate positive SA eigenvectors ($\text{MC}_j/\text{MC}_{max} > 0.25$)

6. Conclusions and implications

Major findings conveyed in this paper are that spatial autocorrelation can materialize in a uniform random variable—and appropriate permutation of map values [e.g., (near-)perfectly uniformly distributed integers from the set {1, 2, . . . , 10} portrayed in Figure 5] shows the feasibility of this possibility, too—and its presence can make a difference in an analysis (e.g., the bivariate spatial median equivalence of the Weberian location problem). The intellectual product presented in this paper should prove useful to spatial scientists conducting simulation experiments involving a uniform random variable distributed across two dimensions, especially for the frequency distribution of attribute values; extending the continuous case treated here to its discrete case companion is straightforward (e.g., Figure 5). A more general implication from these conclusions is that a respecification should exist for any unorthodox random variable that introduces spatial autocorrelation into it. Key ingredients for such a transformation include a parametric mixture perspective, a spatially structured random effect component, and a compatible hyperprior that most likely relates to a normal random variable, the approach pioneered by Besag et al. (1991).

Besides some of the nine random variables considered in this paper, Besag (1974) similarly discusses auto-exponential and auto-gamma random variables, varieties not commonly found in geospatial science analysis descriptions, although the Poisson-gamma mixture rendering a negative binomial random variable normatively suggests otherwise for the auto-gamma specification. Nevertheless, in keeping with the theme of this paper, because the auto-exponential can describe only positive spatial autocorrelation situations, it also qualifies as an unorthodox random variable, expressly in light of arguments given in Griffith (2019) concerning negative spatial autocorrelation. These two random variables merit subsequent spatial statistical research attention along the lines of the conceptualization advanced in this article.

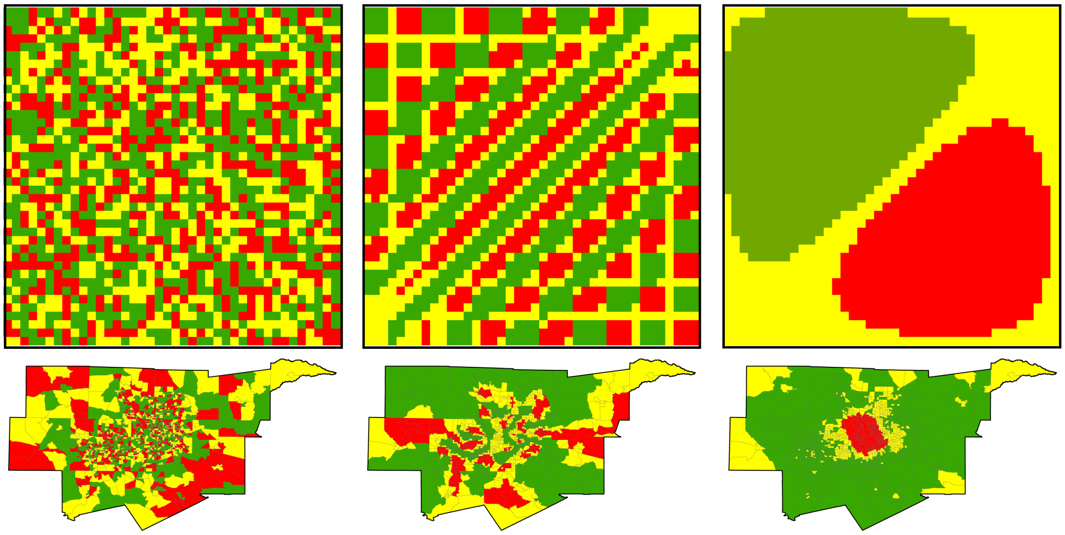

Figure 5. Judiciously permuted discrete uniform random variable geographic distributions, with green, yellow, and red respectively denoting values 1-4 (relatively low), 5-7 (moderate), and 8-10 (relatively high); the top row displays results for a 40-by-40 regular square tessellation, and the bottom row displays results for the 2010 DFW MSA census tracts. Top left (a): complete spatial randomness: MC = -0.01 and GR = 1.01. Top middle (b): $\boldsymbol{E}_{169}$ rank ordering: MC = 0.69 and GR = 0.31. Top right (c): $\boldsymbol{E}_1$ rank ordering: MC = 1.01 and GR = 0.02. Bottom left (d): complete spatial randomness: MC = -0.00 and GR = 1.01. Bottom middle (e): $\boldsymbol{E}_{75}$ rank ordering: MC = 0.79 and GR = 0.26. Bottom right (f): $\boldsymbol{E}_1$ rank ordering: MC = 0.96 and GR = 0.13.

Expanding upon this preceding payoff notion, a major general implication stemming from the novel work summarized in this paper pertains to the probability integral transform (aka the universality of the uniform) theorem, which asserts that random variable values from any continuous distribution are transformable to a standard uniform random variable (Quesenberry, 2006). Given this perspective, the framework devised for and outlined in this paper potentially could enable a very broad treatment of spatial dependence embedding in random variables through this theorem, and hence the uniform random variable. In other words, beginning with any of the continuous random variables as a prior distribution may produce a posterior spatialized uniform distribution. This conjecture warrants careful, intensive, and thorough future consideration.

A final specific takeaway from this paper is that spatial autocorrelation may be integrated into probability density/mass functions through either their parameters—e.g., spatial autoregression, mixed models theory, and Moran eigenvector spatial filtering give preference to the mean response, whereas geostatistics gives preference to the variance-covariance structure—or strictly their support argument(s). The former almost always is the strategy taken by theorists and methodologists; convoluting a subtle version of it with the latter appears to be the only viable option for a uniform random variable. Perhaps this is the approach to use with an auto-gamma specification, in keeping with the probability integral transform revelation, and given Besag's criticism about its nonintuitive and cumbersome form when written using his traditional auto-specification. The Poisson-gamma parametric mixture model favors at least an examination of this possibility.

Author contributions Conceptualization, D.G.; methodology, D.G.; software, D.G.; validation, D.G.; formal analysis, D.G.; investigation, D.G.; data curation, D.G.; writing-original draft preparation, D.G.; writing-review and editing, D.G.; visualization, D.G.; supervision, D.G. All authors have read and agreed to the published version of the manuscript.

Acknowledgements The author thanks the editors, associate editor, and the anonymous referees as well as J. K. Ord for their valuable comments to improve this manuscript. The author is an Ashbel Smith Professor.

Funding The author received no financial support for the research, authorship, and/or publication of this manuscript.

Conflicts of interest The author declares no conflict of interest.

References

Abdi, A., 2007. The eigen-decomposition: eigenvalues and eigenvectors. Encyclopedia of Measurement and Statistics. Sage, Thousand Oaks, California, USA, pp. 305-309.

Anselin, L., 1988. Spatial Econometrics: Methods and Models. Kruger, Dordrecht, Netherlands.

Bailey, T. and Gatrell, A., 1995. Interactive Spatial Data Analysis. Longman Scientific, Harlow, UK.

Bardos, D., Guillera-Arroita, G., and Wintle, B., 2015. Valid auto-models for spatially autocorrelated occupancy and abundance data. Methods in Ecology and Evolution, 6, 1137-1149.

Besag, J., 1974. Spatial interaction and the statistical analysis of lattice systems. Journal of the Royal Statistical Society B, 36, 192-225.

Besag, J., York, J., and Mollié, A., 1991. Bayesian image restoration with two applications in spatial statistics. Annals of the Institute of Statistical Mathematics, 43, 1-59.

Bagui, S., Bhaumik, D., and Mehra, K., 2013. A few counter examples useful in teaching central limit theorems. The American Statistician, 67(1), 49-56.

Bolin, D., Wallin, J., and Lindgren, F., 2019. Latent Gaussian random field mixture models. Computational Statistics and Data Analysis, 130, 80-93.

Blom, G., 1958. Statistical Estimates and Transformed Beta-variables. Wiley, New York, USA.

Box, G. and Cox, D., 1964. An analysis of transformations. Journal of the Royal Statistical Society B, 26, 211-252.

Casella, G. and George, E., 2008. Explaining, the Gibbs sampler. The American Statistician, 46(3), 167-174.

Cheng, T-T., 1948. The normal approximation to the Poisson distribution and a proof of a conjecture Ramanujan. Bulletin of the American Mathematical Society, 55, 396-401.

Chun, Y. and Griffith, D., 2018. Impacts of negative spatial autocorrelation on frequency distributions. Chilean Journal of Statistics, 9(1), 3-17.

Cliff, A. and Ord, J., 1973. Spatial Autocorrelation. Pion, London, UK.

Cressie, N., 1991. Statistics for Spatial Data. Wiley, New York, USA.

Dwight, H., 1961. Tables of Integrals and Other Mathematical Data. Macmillan, New York, USA.

Ecker, M., 2003. Geostatistics: Past, present and future. In: Environmetrics developed under the auspices of the UNESCOS, Encyclopedia of Life Support Systems, El-Shaarawi, A., and Jureckova, J. (eds). EOLSS Publishers, EOLSS, Oxford, UK.

Eftelioglu, E., 2017. Geometric median, in Shekhar, S., Xiong, H., and Zhou, X., (eds.), Encyclopedia of GIS, 2^{nd} ed., 701-704. Springer, Cham, Switzerland.

Ferrari, S. and Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions, Journal of Applied Statistics, 31, 799-815.

Freund, J. 1992. Mathematical Statistics, 5^{th} ed. Pretience Hall, Englewood Cliffs, New Jersey, USA.

Gilks, W., Richardson, S., and Spiegelhalter, D., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

Griffith, D., 2002. A spatial filtering specification for the auto-Poisson model. Statistics and Probability Letters, 58, 245-251.

Griffith, D., 2003. ASpatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization. Springer-Verlag, Berlin, Germany.

Griffith, D., 2011. Positive spatial autocorrelation impacts on attribute variable frequency distributions. Chilean Journal of Statistics, 2(2), 3-28.

Griffith, D., 2012. Spatial statistics: A quantitative geographer's perspective. Spatial Statistics, 1, 3-15.

Griffith, D., 2013. Better articulating normal curve theory for introductory mathematical statistics students: power transformations and their back-transformations. The American Statistician, 67, 157-169.

Griffith, D., 2017. Some robustness assessments of Moran eigenvector spatial filtering. Spatial Statistics, 22, 155-179.

Griffith, D., 2018. Generating random connected planar graphs. Geoinformatica, 22, 767-782.

Griffith, D., 2019. Negative spatial autocorrelation: one of the most neglected concepts in spatial statistics. Stats, 2, 388-415.

Griffith, D., 2021. Articulating spatial statistics and spatial optimization relationships: Expanding the relevance of statistics. Stats, 4, 850-867.

Griffith, D. and Chun, Y., 2019. Implementing Moran eigenvector spatial filtering for massively large georeferenced datasets. International J. of Geographical Information Science, 33, 1703-1717.

Griffith, D., Chun, Y., and Kim, H., 2022. Spatial autocorrelation informed approaches to solving location-allocation problems, Spatial Statistics, 50, 100612.

Griffith, D., Chun, Y., and Kim, H., 2023. Spatial autocorrelation informed approaches to solving location-allocation problems, Geographical Analysis. 55, 107-124.

G'Sell, M., Wager, S., Chouldechova, A., and Tibshirani, R., 2016. Sequential selection procedures and false discovery rate control. Journal of the Royal Statistical Society, Series B, 78, 423-444.

Hardouin, C. and Yao, J-F., 2008. Multi-parameter automodels and their applications, Biometrika, 95(2), 335-349.

Hilbe, J., 2014. Generalized linear models. In: International Encycolpedia of Statistical Science, M. Lovric (ed.). Springer, New York, USA.

Hoeffding, W., 1952. The large-sample power of test based on permutations of observations. Annals of Mathematical Statistics, 23, 169-192.

Hu, L., Griffith, D., and Chun, Y., 2020. Impacts of spatial autocorrelation in georeferenced beta and multinomial random variables. Geographical Analysis, 52, 278-298.

Jabeen, R. and Zaka, A., 2020. Estimation of parameters of the continuous uniform distribution: Different classical methods. Journal of Statistics and Management Systems 23(3), 529-547.

James, I., 1975. Multivariate distributions which have beta conditional distributions. Journal of the American Statistical Association, 70, 681-684.

Johnson, N., Kotz, S., and Balakrishnan, N. 1994/95. Continuous Univariate Distributions, vols. 1 & 2, 2^{nd} ed. Wiley, New York, USA.

Johnson, N., Kotz, S., and Balakrishnan, N., 2005. Univariate Discrete Distributions, 3^{ed} ed. Wiley, New York, USA.

Kaiser, M. and Cressie, N., 1997. Modeling Poisson variables with positive spatial dependence. Statistical Papers, 35, 423-432

Kaiser, M. and Cressie, N., 2000. The construction of multivariate distributions from Markov random fields. Journal of Multivariate Analysis, 73, 199-220.

Kavousi, A., Meshkani, M., and Mohammadzadeh, M., 2011. Spatial analysis of auto-multivariate lattice data. Statistical Papers, 52(4), 937-952.

Le Cam, L., 1960. An approximation theorem for the Poisson binomial distribution. Pacific Journal of Mathematics, 10, 1181-1197.

Leemis, L. and McQueston, J., 2008. Univariate distribution relationships. The American Statistician, 62, 45-53.

Lindsay, B., 1995. Mixture Models: Theory, Geometry and Applications. Hayward, CA: Institute of Mathematical Statistics, NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 5.

Lohnes, P. and Cooley, W., 1968. Chapter 7: Normal Curve Theory, Introduction to Statistical Procedures: With Computer Exercises. Wiley, New York, USA, pp. 107-125.

Luo, Q., Griffith, D., and Wu, H., 2018. On the statistical distribution of the nonzero spatial autocorrelation parameter in a simultaneous autoregressive model. ISPRS International Journal of Geo-Information, 7(12), 476.

McCullagh, P. and Nelder, J., 1989. Generalized Linear Models, 2^{nd} ed. Chapman & Hall/CRC., London, UK.

McCulloch, C., 2008. Generalized, Linear and Mixed Models, 2^{nd} ed. Wiley, New York, USA.

Ninimaa, A., 1995. Bivariate generalizations of the median, in Multivariate Statistics and Matrices in Statistics: Procedings of the 5^{th} Tartu Conference, vol. 3, Tartu-Pühajärve, Estonia, 23-28 May, 1994, edited by Tiit, E., Kollo, T., and Niemi, H. Hoston: De Gruyter, pp. 163-180.

Paelinck, J. and Klaassen, L., 1979. Spatial Econometrics. Saxon House, Farnborough, UK.

Pan, L., Li, Y., He, K., Li, Y., and Li, Y., 2020. Generalized linear mixed models with Gaussian mixture random effects: Inference and application. Journal of Multivariate Analysis, 175, 104555.

Peizer, D., and Pratt, J., 1968. A normal approximation for binomial, F, Beta and other common, related tail probabilities, I. Journal of the American Statistical Association, 63, 1416-1456.

Pratt, J., 1968. A normal approximation for binomial, F, Beta and other common, related tail probabilities, II. Journal of the American Statistical Association, 63, 1457-1483.

Quesenberry, C., 2006. Probability integral transforms. Encyclopedia of Statistical Sciences, vol. 10, 2^{nd},. Kotz, S., Balakrishnan, N., Read, C., and Vidakovic, B. (eds.). Wiley, New York, USA, pp. 6476-6481.

Rahman, M., and Govindarajulu, Z., 1997. A modification of the test of Shapiro and Wilk for normality. Journal of Applied Statistics, 24(2), 219-236.

Small, C., 1990. A survey of multidimensional medians. International Statistical Review, 58(3), 263-277.

Steigler, S., 1986. The History of Statistics: The measurement of Uncertainty Before 1900. Harvard University Press, Cambridge, Massachusetts, USA.

Zikariene, E. and Ducinskas, K., 2021. Application of spatial auto-beta models in statistical classification, Lietuvos Matematikos Rinkinys. Proceedings of the Lithuanian Mathematical Society, Series A, 62, 38-42.

Appendix A
Mathematical statistical theory underlying a uniform RV containing SA

Merging statistical distributions can occur in at least three different ways: Bayesian analysis (Freund (1992), pp. 398-397); hierarchical/parametric mixture combinations (e.g., Lindsay (1995)); and, the variable transformation technique (Freund (1992), pp. 264-272; Griffith (2013)). All three have the same independence assumption, almost always involve standard RVs (e.g., Johnson et al. (1994/95),Johnson et al. (2005); Leemis and McQueston (2008)), and posit some kind of (hyper)prior distribution mechanism. The naive Bayesian approach postulates independent variables (i.e., a product of marginal univariate densities characterize its joint probability density). In contrast, a mixture model approximates a probability density with either a linear combination of component RVs accompanied by a full covariance matrix indicating pairwise independence if and only if it is diagonal, or sometimes a conjugate prior closed form expression, and other times messy open-form expressions yielding posterior distributions solely via numerical integration, when parameters become RVs. Meanwhile, a variable transformation—the most popular being the Box and Cox (1964) power kind—frequently allows a preferable output RV to approximate another undesirable input RV. These procedures furnish powerful techniques for integrating multiple data generating processes into a single RV specification. Furthermore, their (effective) posterior distributions most often experience variance reduction.

The theoretical basis for this paper explicitly builds upon the variable transformation technique, predicating a specific auto-uniform type RV upon a prior beta RV with equal shape parameters (i.e., $\gamma_1 = \gamma_2 = \gamma$), ensuring its symmetry, and hence a mean of 1/2, coupled with variance inflation attributable to SA that shrinks in the ensuing posterior distribution (i.e., this inflation is from the posterior uniform to the prior bell-shaped beta). The prior distribution is the standard beta RV probability density function

$$f(x) = \frac{\Gamma[2\gamma]}{(\Gamma[\gamma])^2} X^{\gamma-1}(1-X)^{\gamma-1}, \quad 0 \leq X \leq 1. \tag{6.3}$$

Variable X may be rewritten in its equivalent logistic function expression $1/\{1 + \exp(-LN[(1-x)/x])\}$. This is the form allowing the logit function $-LN[(1-x)/x]$ to become part of a random effects term, enabling the embedding of SA in a manner paralleling that by Besag et al. (1991) for Bayesian map analysis. The auto-normal logit-linear regression specification permits a convolution of SA and probabilities using, for example, the simultaneous autoregressive (SAR) spatial linear operator $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}$ extracted from the following description of a georeferenced sample of n proportion/probability values:

$$< LN[(1-x_i)/x_i] >= \rho\boldsymbol{W} < LN[(1-x_i)/x_i] > +\epsilon, \tag{6.4}$$

where $< \bullet >$ denotes an n-by-1 vector, $\bullet$ is a wild card character, and ϵ is an n-by-1 vector of independent and identically distributed normal errors. The matrix expression $TR[(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{I} - \rho\boldsymbol{W}^T)^{-1}]/n$ quantifies the variance inflation induced by this data generating mechanism, where TR denotes the transpose operator. Applying the trapezoidal rule of elementary integral calculus to approximate this quantity, as n increases, this variance inflation term converges on

$$\frac{1-(-1)}{2n}\{2TR[(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\boldsymbol{I} - \rho\boldsymbol{W}^T)^{-1}] - 2\} \approx \int_{-1}^{1} \frac{1}{(1-\rho\lambda)^2} d\lambda, \tag{6.5}$$

where $-1 \leq \lambda \leq 1$ denotes the variable representing the n eigenvalues of spatial weights matrix $\boldsymbol{W}$, which partition their domain[1], [-1,1], into increasingly finer intervals as n increases. The solution to Equation (6.5) yields variance inflation factor $\lceil 1/\sqrt{1-\rho^2}\rceil^2$, a limiting result that rescales[2] each initial beta RV value by $1/\sqrt{1-\rho^2}$. This is the variance inflation generated by SA for which a prior beta RV must compensate through deflation in order to have a posterior type uniform RV. Table 4 tabulates the values that a single shape parameter prior beta distribution takes on to render the posterior uniform RV. The affiliated trendline is, approximately,

$$\gamma = 0.2961 + \frac{0.7039}{1-\rho^2}, \quad R^2 \approx 1, \tag{6.6}$$

whose bivariate regression equation calculating $\hat{\gamma}$ with its predicted values has an intercept of 0.02538, indicating some possible bias in the specification of Equation (6.6), although it might be compensating for the beta shape parameter γ being strictly positive rather than non-negative, and a slope of 0.99939, which is almost identical to 1. Figure 6 portrays selected theoretical prior distributions listed in Table 4.

Table 4. Summary statistics for selected representative systematic samples drawn from uniform RVs containing SA via a beta RV specification.

ρ	Mathematica 12.3		40-by-40 square tessellation						2010 DFW MSA					
	γ	$\sigma=\sqrt{1/12}$	γ	$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	MC	GR	γ	$\hat{\sigma}$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	MC	GR
0.00	1.00	0.289	1.00	0.289	1.00	1.00	-0.02	1.02	1.00	0.289	1.00	1.00	-0.02	1.02
0.10	1.01	0.289	1.01	0.288	1.01	1.01	0.04	0.96	1.01	0.289	1.00	1.00	0.08	0.91
0.20	1.03	0.289	1.03	0.287	1.01	1.01	0.08	0.92	1.03	0.288	1.01	1.01	0.09	0.90
0.30	1.07	0.289	1.07	0.289	1.00	1.00	0.19	0.81	1.07	0.287	1.02	1.02	0.14	0.86
0.40	1.14	0.289	1.14	0.288	1.02	1.02	0.24	0.76	1.14	0.287	1.04	1.04	0.21	0.79
0.50	1.25	0.289	1.24	0.289	1.03	1.03	0.31	0.69	1.23	0.286	1.06	1.07	0.25	0.74
0.60	1.41	0.289	1.38	0.289	1.03	1.03	0.38	0.63	1.34	0.288	1.04	1.05	0.32	0.67
0.70	1.70	0.289	1.68	0.287	1.06	1.06	0.47	0.53	1.57	0.291	1.03	1.03	0.44	0.56
0.80	2.29	0.289	2.20	0.288	1.08	1.09	0.58	0.42	2.01	0.284	1.14	1.07	0.51	0.48
0.85	2.87	0.289	2.40	0.291	1.05	1.05	0.60	0.40	2.07	0.293	1.02	0.94	0.57	0.43
0.90	4.04	0.289	3.80	0.284	1.11	1.11	0.71	0.28	3.70	0.288	1.05	1.05	0.72	0.27
0.95	7.56	0.289	5.93	0.283	1.12	1.11	0.80	0.20	5.95	0.286	1.02	1.12	0.79	0.19
0.99	35.68	0.289	33.50	0.287	1.13	1.14	0.96	0.05	25.00	0.291	1.06	0.98	0.93	0.06

square tessellation: MC_{max} = 1.02, MC_{min} = -1.02; GR_{min} = 0.01, GR_{max} 2.05
DFW MSA: MC_{max} = 1.18, MC_{min} = -0.69; GR_{min} = 0.33, GR_{max} 2.53, $\lambda_n(W)$ = -0.79651

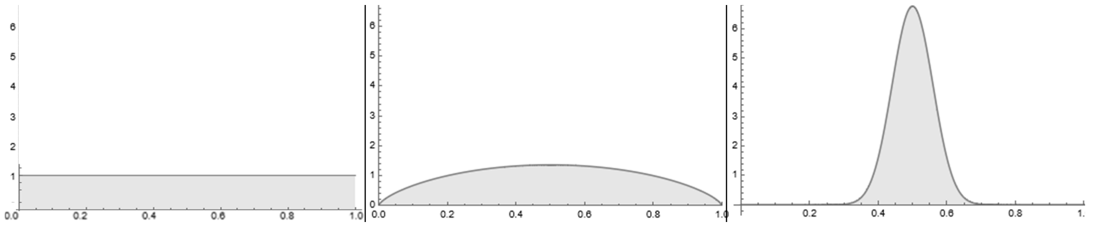

Figure 6. Shrinking variance frequency distributions of selected hyperpriors across increasing levels of positive SA. Left (a): $\rho = 0$, $\gamma = 1$. Middle (b): $\rho = 0.7$, $\gamma = 1.70$. Right (c): $\rho = 0.99$, $\gamma = 35.68$.

[1] The largest eigenvalue of matrix $\boldsymbol{W}$ always is one, by the Perron-Frobinius theorem, whereas its opposite extreme varies within the relatively wide interval [-1,-1/2], depending upon surface partitioning idiosyncrasies and nuances and the employed neighbors definition. A regular square tessellation coupled with a rook adjacency definition gives the ideal full interval [-1,1].

[2] Multiplicative rescaling proportionally decreases/increases each value, moving them forward/away from zero. Exponentiation rescaling differentially decreases/increases each value, moving them toward/away from one.

Inducing SA in the logit function via Equation (6.4), following standard spatial autoregressive SAR manipulations, produces individual observation values $-LN[(1-x)/x]/\sqrt{1-\rho^2}$, with the denominator deriving from Equation (6.4). Rewriting this term as the variable transformation

$$y = x^{\frac{1}{\sqrt{1-\rho^2}}} \Big/ \left[x^{\frac{1}{\sqrt{1-\rho^2}}} + (1-x)^{\frac{1}{\sqrt{1-\rho^2}}} \right] \Rightarrow x = y^{\sqrt{1-\rho^2}} \Big/ \left[y^{\sqrt{1-\rho^2}} + (1-y)^{\sqrt{1-\rho^2}} \right]$$

—with the left-hand equation reducing to x if $\rho = 0$—followed by making the appropriate substitutions, replaces the probability density given in Equation (6.3) with

$$g(y) = \frac{\Gamma[2\gamma]}{(\Gamma[\gamma])^2} \frac{\sqrt{1-\rho^2}[y(1-y)]^{\gamma\sqrt{1-\rho^2}}}{[y(1-y)][y^{\sqrt{1-\rho^2}} + (1+y)^{\sqrt{1-\rho^2}}]^{2\gamma}}, \quad 0 \le y \le 1, \tag{6.7}$$

which includes the Jacobian, $|dx/dy| = (\sqrt{1-\rho^2}[y(1-y)]^{\sqrt{1-\rho^2}-1})/([y^{\sqrt{1-\rho^2}} + (1-y)^{\sqrt{1-\rho^2}}]^2)$, for each individual observation, and which reduces to the functional form of Equation (6.3) if $\rho = 0$. This hypothetical asymptotic outcome furnishes guidelines for empirical cases, such as the 40-by-40 regular tessellation, and the 2010 DFW MSA census tracts. An iterative computer program script initiated with the appropriate theoretical value converges to its case specific empirical surface equivalent based upon an objective function defined as, for example, the squared difference between the empirical variance and 1/12, the theoretical variance (see Table 4 and Figure 1 in the main body of this paper); this stated procedure is a mean squared errors criterion variety of minimization. The empirical surface counterparts begin to deviate somewhat from the reported theoretical trajectory in the presence of a moderate-to-strong degree of SA (see Figure 1). Differences are attributable, in part, to the finite number as well as any skewness of the empirical eigenvalue distributions (Figure 7). All of the empirical cases reported in Table 4 conform to a beta distribution whose two shape parameters are nearly equal, and very close to one.

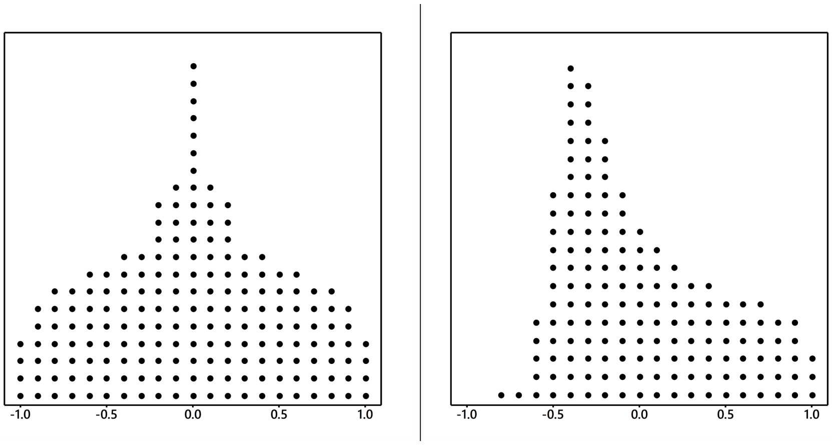

Figure 7. Dot plot (each dot represents five eigenvalues) portrayal of the matrix $\boldsymbol{W}$ eigenvalue frequency distributions. Left (a): 40-by-40 regular square tessellation; skewness = 0, kurtosis = 2.23. Right (b): 2010 DFW MSA; skewness = 0.64 (with a normal curve theory based z = 9.49), kurtosis = 2.34.

Conceptually, the implementation here is analogous to a SSRE term inserted into a beta RV. Mixed models theory (e.g., McCulloch (2008)) regularly links its random effects term to a normal distribution. Accordingly, evaluation of the entries in Table 4 reveals, based upon kurtosis—this beta measurement for equal shape parameters γ equals $4[3-6/(2\gamma+3)]$, versus 3 for a normal RV—has a test statistic of $|z| = 6/[(2\gamma+3)\sigma_K]$, where σ_K denotes kurtosis variance, which for the normal distribution is $\sigma_K^2 = 24n(n-1)^2/[(n-3)(n-2)(n+3)(n+5)]$.

The corresponding 95% confidence interval critical values attained are as follows:

$$n = 1,600 : \gamma \approx 11(|Z_{\text{kurtosis}}| = 1.96), \text{Prob(Shapiro-Wilk)} = 0.59$$
$$n = 1,314 : \gamma \approx 10(|Z_{\text{kurtosis}}| = 1.96), \text{Prob(Shapiro-Wilk)} = 0.58$$

In other words, positive SA needs to be at or beyond the level computed for remotely sensed images before the necessary prior beta RV closely mimics a bell-shaped curve. Less sensitive Shapiro-Wilk normality diagnostic statistics decrease this threshold to

$$n = 1,600 : \gamma \approx 7.61, \text{P(S-W)} = 0.10$$
$$n = 1,314 : \gamma \approx 6.91, \text{P(S-W)} = 0.10$$

which, nevertheless, corroborates the preceding finding. Figure 8 uncovers modest deviations in both tails of the transformed distributions, the principal source of any detected discrepancies between a beta and the normal distribution it attempts to mimic. Regardless, Figure 6c visually resembles a bell-shaped curve, implying that perhaps symmetry is more important than unimodal peakedness (i.e., kurtosis). This conjecture merits future investigation.

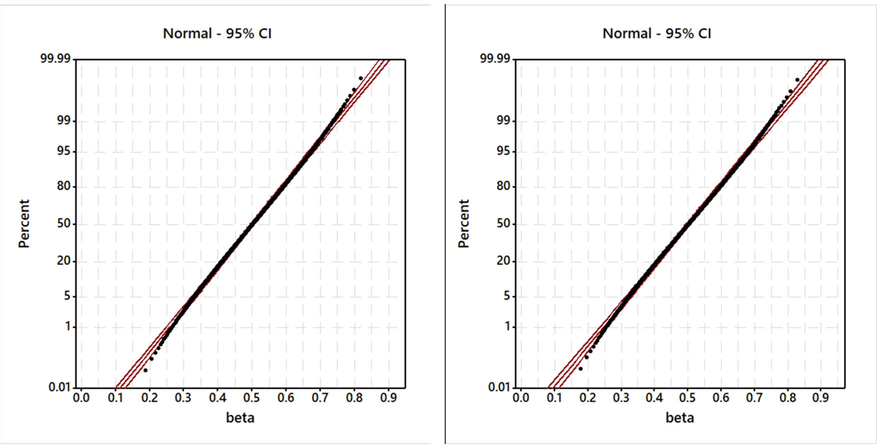

Figure 8. Beta-normal quantile plots. Left (a): square tessellation ($n = 40 \times 40 = 1,600$, $\gamma = 11$). Right (b): 2010 DFW MSA ($n = 1,314$, $\gamma = 10$).

In conclusion, the variable transformation technique offers a fruitful tool for constructing a uniform RV in which SA can materialize. Although much mixed models theory relies heavily upon the normal distribution for a random effects term, among the findings enumerated in this appendix is that symmetry alone may be the necessary quality for establishing a successful prior distribution.