Spatial Log-Gaussian Cox process models and sampling paths: towards optimal design

Kenneth Flagg[a,*], John Borkowski[a], Andrew Hoegh[a]

[a]*Department of Mathematical Sciences, Montana State University, Bozeman, MT 59717*

Abstract

Goal of this paper (placeholder abstract—add some results when available). Evaluate a wide variety of path designs in terms design-based heuristics and model-based criteria for spatial prediction using Bayesian LGCP models. Identify promising path designs. Illuminate any relationships among design characteristics and predictive criteria that will be helpful for constrained optimization.

Keywords: log-Gaussian Cox process, optimal sampling, model-based design, spatial sampling design

1. Introduction

Spatial point process models have generally been infeasible because of their computational demands, but recent advances in Bayesian computing have made the Log-Gaussian Cox process (LGCP) an attainable model in practice (Rue et al., 2009; Lindgren et al., 2011; Illian et al., 2012; Simpson et al., 2016). These advances make it possible to fit LGCP models easily, without time-consuming Monte Carlo methods. In some applications, the entire point pattern is not fully observed due to variable sampling effort. This is referred to as a degraded point pattern (Chakraborty et al., 2011) and it is relatively simple to accommodate variable sampling effort in these models using modern Bayesian computing

*Corresponding author
Email address: kenneth.flagg@montana.edu (Kenneth Flagg)

tools (Yuan et al., 2017). However, the literature on optimal sampling for spatial point process models is in its infancy (Liu and Vanhatalo, 2020).

Point pattern data are routinely collected in species distribution studies and ordnance response projects. The data consist of the locations of events in some spatial region. These applications may use quadrat sampling or line-transect sampling, with transect sampling being more common. When the objective is to map where events occur in space, various spatial mapping procedures have been used. Traditionally these have involved aggregating the data to grid cell counts or computing moving averages. Aggregation has the downside of introducing arbitrary structure into the data by the choice of grid scheme or averaging window, and requires unnecessary computation effort (Simpson et al., 2016). Software is now available to fit spatial point process models to data acquired via distance sampling and simultaneously estimate the detection function (Johnson et al., 2014; R Core Team, 2019).

In ecological settings, sampling plans are often designed around the goal of estimating total abundance. Ordnance response surveys are typically designed to provide enough data to detect (but not necessarily map) intensity hotspots (USACE, 2015; Flagg et al., 2020). However, to our knowledge, there has been very little work done in deciding *where* to collect data when the goal is to map the intensity using a spatial point process model. While some ideas about the characteristics of a good point design apply to paths, creating an optimal path design is not as simple as connecting the points of a point design with line segments. There are many ways to connect points into a path, so optimal design criteria must apply to the whole path and not only to the waypoints. In this paper, we present a variety of sampling path designs and assess their optimality for mapping intensity using LGCP models.

1.1. Log-Gaussian Cox process

The log-Gaussian Cox process is an inhomogeneous Poisson process where the logarithm of the intensity function is a Gaussian process (Møller et al., 1998). The LGCP provides a flexible model for mapping event intensity over space using few parameters. Efficient Bayesian computation tools are available using INLA to approximate the posterior marginal distributions Rue et al. (2009), a finite element approach to represent the Gaussian process Lindgren et al. (2011), and pseudodata to approximate the point process likelihood Simpson et al. (2016).

1.2. Spatial design

Most classical sampling and design work has been done for points, or for small quadrats approximated as points, rather than for paths. In two-dimensional (geostatistical) model-based design, regularity is optimal for spatial prediction but randomness and a variety of interpoint distances are best for parameter estimation (Diggle and Lophaven, 2006). Inhibitory plus close pairs designs are a good compromise (Chipeta et al., 2017). Design-based approaches exist to spread points through high-dimensional design spaces (Borkowski and Piepel, 2009), and Latin hypercube sampling has space-filling characteristics (McKay et al., 1979; Husslage et al., 2011).

1.3. Space-filling curves

Another relevant area of research is in deterministic space-filling curves. These have used in the design of dense or stretchable circuits (Ogorzałek, 2009; Ma and Zhang, 2016) and high-dimensional data visualization in bioinformatics (Anders, 2009). The Hilbert curve is simple to construct and the Peano curve is very flexible for filling irregular shapes (Fan et al., 2014). Space-filling curves are one-dimensional paths constructed iteratively; as the number of iterations goes to infinity, the limiting path has nonzero area and actually fills the space (Sagan, 1994). For applications we stop after a finite number of iterations.

1.4. Paths as sampling designs

The small body of literature on spatial sampling design for point pattern data has focused on line transects. Pollard et al. (2002) began with line transects and adaptively added zigzags in a species abundance survey.

The Visual Sample Plan software includes features to create systematic transect plans and augment plans with additional transects in regions lacking spatial coverage (Matzke et al., 2014). It helps the user choose the transect spacing to maximize the probability of detecting the presence of a hotspot of specified size and intensity. However, it does not employ criteria to optimize spatial prediction.

Liu and Vanhatalo (2020) provided one of the first explicit discussions of design in the context of spatial LGCP models. They used narrow quadrats (swaths along line-transects) as their sampling units. The transects were short relative to the size of the study region and not connected into a path.

2. Materials and methods

With an eye toward practical considerations of data collection, we present criteria to compare sampling strategies that impact LGCP estimates. We compare plans with (approximately) fixed path lengths, most of which avoid sharp turns. Data collection equipment (e.g. metal detectors) may have limited mobility, requiring minimizing the number or angle of turns. The criteria that we evaluate are average prediction variance (APV) and mean squared prediction error (MSPE) of the Gaussian process.

2.1. Sampling design schemes

In this section, we present three variations of parallel line transect designs and three schemes that produce more complex designs. To clarify terminology, a *path* or *design* is a realized set of one or more connected components that has length but not area. The paths considered in this work are constructed as sequences of line segments. A *design scheme*, or simply *scheme*, is procedure for generating designs with some shared characteristics. Figure 1 illustrates a selection of designs from these schemes.

2.1.1. *Parallel line transects*

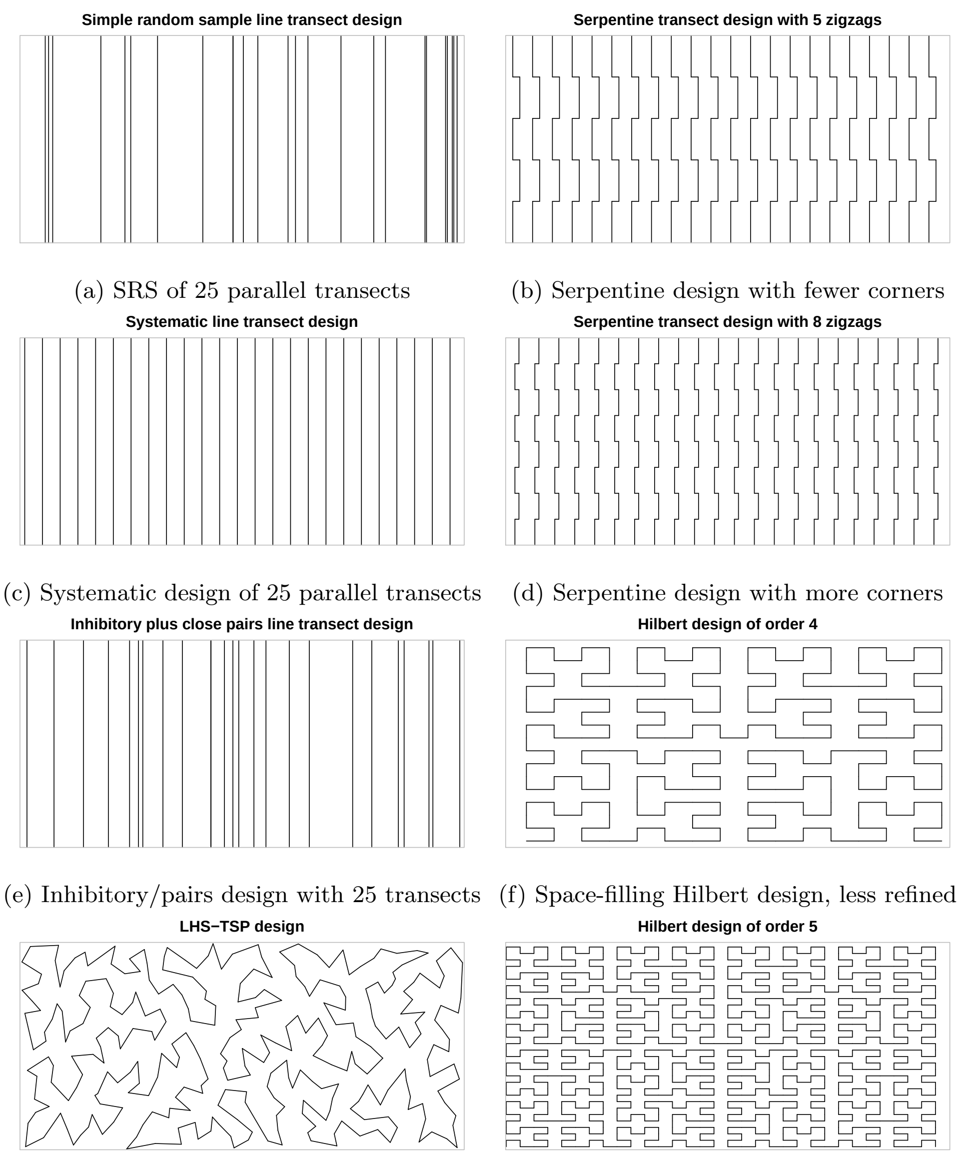
Parallel straight-line transects are common in ordnance response studies and in ecological studies using distance sampling. Systematic designs are common because they provide good spatial coverage in the sense that any point in the study region has an a priori known maximum distance from the path. For point designs, systematic designs are optimal for prediction, simple random samples are optimal for estimation, and inhibitory with close pairs designs are becoming a popular compromise. We adapt all of these to the parallel line transect setting. We use line transects running north-south, with three methods of choosing the horizontal coordinate: simple random sample (SRS), systematic with a random starting point and even spacing, and inhibitory plus close pairs. Figure 1 (left column) shows an example of each scheme with 25 transects.

2.1.2. *Parallel serpentine transects*

One simple way to observe a greater variety of locations and different directions is to add lateral zigzags to transects. We include alternate right and left turns at right angles to create serpentine transects. This could decrease prediction variance because more of the path will be close to each point in the study area than would be under a line transect design with similar total distance. They will also improve estimation of the covariance function in the presence of anisotropy. Figure 1, top right, shows two examples.

2.1.3. *Latin hypercube sampling*

Random Latin hypercube sampling (LHS) produces a design that spreads discrete points through a (potentially high-dimensional) design space, ensuring that the full range of each dimension is included while remaining balanced and keeping the number of points small (McKay et al., 1979). This is done by partitioning each dimension into a specified number k of intervals (thus stratifying a d-dimensional design space into k^d cells), selecting a Latin hypercube design to determine which k cells will contain a design point, and then drawing each design point from a uniform distribution over its cell. In two dimensions,



(a) SRS of 25 parallel transects
(b) Serpentine design with fewer corners
(c) Systematic design of 25 parallel transects
(d) Serpentine design with more corners
(e) Inhibitory/pairs design with 25 transects
(f) Space-filling Hilbert design, less refined
(g) Shortest path through a LHS design
(h) Space-filling, Hilbert design, more refined

Figure 1: Examples of plans from six design schemes. Except for the Hilbert curve of order 5, all of these plans have approximately the same total length.

this scheme produces point designs with good spatial coverage properties. We use the LHS design as waypoints for a path. Because longer distance typically brings increased costs, we treat this as a traveling salesperson problem (TSP) and use the shortest path through the waypoints as our design. This LHS-TSP scheme produces paths that have many sharp corners but leaves few large voids (example in Figure 1, bottom left). A downside of this design scheme is that the length cannot be specified directly, and only certain distances are possible depending on the number of bins used.

Waypoints are generated by the `lhs` R package Carnell (2020) and connected into a the shortest path by the `TSP` package (Hahsler and Hornik, 2020).

2.1.4. Space-filling curves

As a representative of space-filling curves, we use the Hilbert curve scaled to fit the study site. These designs have many short segments meeting at right angles. The only parameter of this design scheme is the order, or number of iterations used in refining the curve. Each iteration increases the length and complexity of the design. This is produces a deterministic design, so a random offset is added to vary which points are observed. The Hilbert curve is generated by `HilbertVis` R package (Anders, 2009).

2.2. Model fitting

We fit the spatial LGCP model using nested integrated Laplace approximations and the R-INLA package (Rue et al., 2009; Blangiardo and Cameletti, 2015). The Gaussian process is approximated using a finite element approach (Lindgren et al., 2011). The point pattern is modeled by pseudodata placed at the events and the finite element nodes (Simpson et al., 2016). This procedure allows fast and accurate approximation of the posterior distribution.

3. Simulation Study

We simulate 100 designs from each of six schemes. All events within a 2 unit radius of the path are observed. The whole experiment is repeated for 5 realizations from each of two data generating models.

3.1. Study site

We consider a fictitious site $\mathcal{R}$ with the simple shape of a 1500 unit by 700 unit rectangle. In this site, we simulate two data generating models that produce random intensity functions with hotspots. First, a LGCP with latent GP mean $\mu = \log(250/|\mathcal{R}|)$ and a Matérn covariance with $\nu = 1$, $\sigma = 2$, and range $= 200$. This model produces relatively unstructured hotspots due to large variability in the GP.

Second, a two-stage cluster process and a LGCP are superposed. The cluster process (a Neyman-Scott or, more specifically, a Thomas process) is constructed as follows. The number of clusters is Poisson-distributed with mean 3. The number of events per cluster is Poisson-distributed with mean 200. The cluster centers are distributed uniformly over $\mathcal{R}$. Events come from a bivariate normal distribution with mean equal to the cluster center and variance $\mathbf{\Sigma} = \tau^2\mathbf{I}$, $\tau = 50$. The LGCP has $\mu = \log(250/|\mathcal{R}|)$ and Matérn covariance with $\nu = 1$, $\sigma = 1$, and range $= 200$. This model is based upon the typical conceptual model of a firing range, with a background process (represented by the LGCP) and a small number of higher-intensity foreground clusters containing the events of interest.

3.2. Path design schemes

The simulation uses each of the design schemes discussed in Section 2.1. The parallel transect schemes have 10, 25, 50, or 70 line transects running north-south. We expect the simple random sample scheme to produce high prediction variance and large prediction error in big gaps between transects. The systematic sample scheme uses a uniformly-distributed starting point and constant spacing between adjacent transects. We expect systematic transects to provide low bias and moderate prediction variance. However, this scheme can miss structures at certain sizes because no transects are close to each other in the east-west direction.

For the inhibitory plus close pairs line transect scheme, we vary the numbers of paired and unpaired transects. The total number of transects is 10, 25, 50, or 70, with 10% and 20% of the transects (rounded to the nearest integer) as redundant members of a pair. The remaining primary transects are placed according to a one-dimensional Strauss process (Strauss, 1975; Kelly and Ripley, 1976). The Strauss attraction parameter is set at $\gamma = 0.05$ and the radius for counting pairs is 1500 units divided by the total number of transects. Then each redundant transect is randomly paired to a primary transect, and placed within the pair radius of the primary transect according to a uniform distribution. We expect this scheme to have intermediate performance between the simple random sample and the systematic line transect schemes.

The serpentine transect scheme has 7, 22, 47, or 67 transects running north-south with constant east-west spacing and a random starting point for the first transect. The number of zigzags is 5 or 8, and the zigzag perpedicular length is set so the the total east-west distance equals the length of three north-south line transects. Thus, the serpentine designs have the same length as the line transect designs. These designs should result in smaller prediction errors and lower variance farther from path, compared to line-transect designs.

Our Latin hypercube sampling/traveling salesperson (LHS-TSP) scheme uses 50, 300, 1200, or 2400 bins to generate the waypoints. Preliminary experimentation found that these bin numbers produced total lengths similar to the line-transect schemes. The LHS-TSP scheme is expected to result in small prediction errors and low prediction variance per unit distance traveled. However, the designs will have many sharp corners and may leave some large voids.

The Hilbert curve scheme uses a random starting point and a Hilbert curve of order 3, 4, 5, or 6. The path length is a deterministic function of the order and differs greatly among curves of different orders. These orders yield lengths similar to the lenths of the transect designs. Hilbert designs should provide low prediction variance, but have lots of short segments.

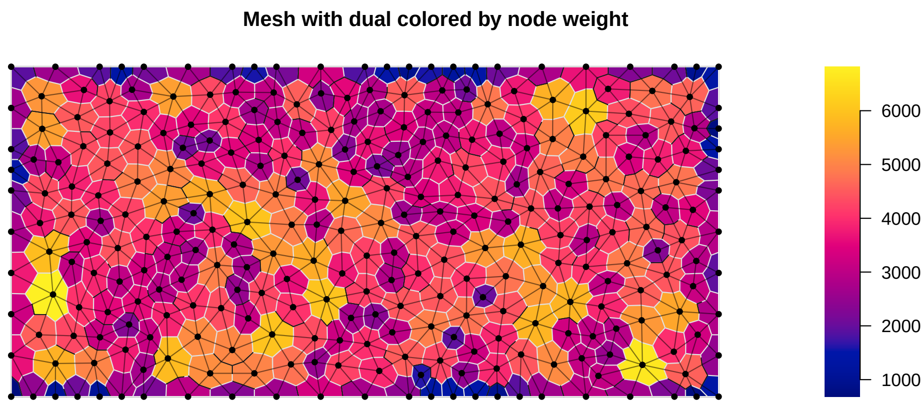


Figure 2: Illustration of the mesh and associated numerical integration weighting scheme used to approximate the latent GP.

3.3. Model specification

The same Bayesian LGCP model is fit to each observed dataset. The observed point pattern $\mathbf{x}$ is a realization of $\mathbf{X}$, a Poisson process on $\mathcal{R}$ with intensity $\lambda(u)$. The intensity is modeled as $\log[\lambda(u)] = \mu + \mathbf{e}(u)$. The spatial error term $\mathbf{e}$ is a Gaussian process with mean $\mathbf{0}$ and a Matérn covariance function with fixed $\nu = 1$.

The intercept μ has a $\text{Unif}(-\infty, \infty)$ prior. The covariance parameters σ and ρ have a PC prior with $\Pr(\sigma > 3) = 0.1$ and $\Pr(\rho < 100) = 0.1$ (Fuglstad et al., 2019; Simpson et al., 2017).

The Gaussian process prediction surface is approximated on the finite element mesh shown in Figure 2. The GP is predicted at the nodes (points) and is linearly interpolated elsewhere. The nodes are weighted according to the area of their dual cells (shading) and used for numerical integration of the likelihood (Lindgren et al., 2011).

4. Results

4.1. Initial Observations

In describing the results, we focus on one LGCP dataset and one clustered dataset (Figure 3). The results are similar for all datasets (see the online supplement.)

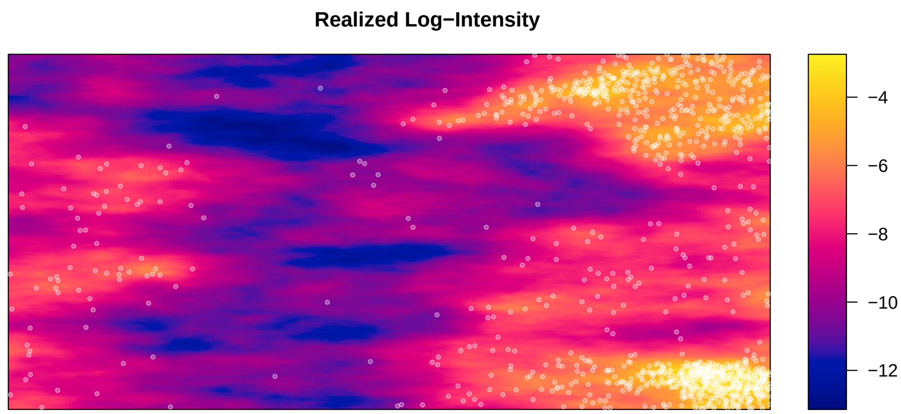
Figure 4 shows an example where the model does well at predicting the intensity of the realized LGCP from data observed along one of the SRS paths. In the figure, the path appears in white and the observed events are shown as white dots. The posterior predicted mean of the log-intensity (top panel) accurately captures the large-scale features, but smooths out much of the small-scale variation. The bottom panel shows the prediction standard deviation for each mesh node. The SD ranges from 2.0 to 3.1, and is lowest near observed events. SD increases farther from observed events, including in places where the surveyed area was observed to contain no events.

Most plans yielded similar prediction surfaces, capturing the large-scale trends, and having the least uncertainty near observed events. Results varied in accuracy at the most extreme peaks and valleys of the itensity function and in overall SD across the study region.

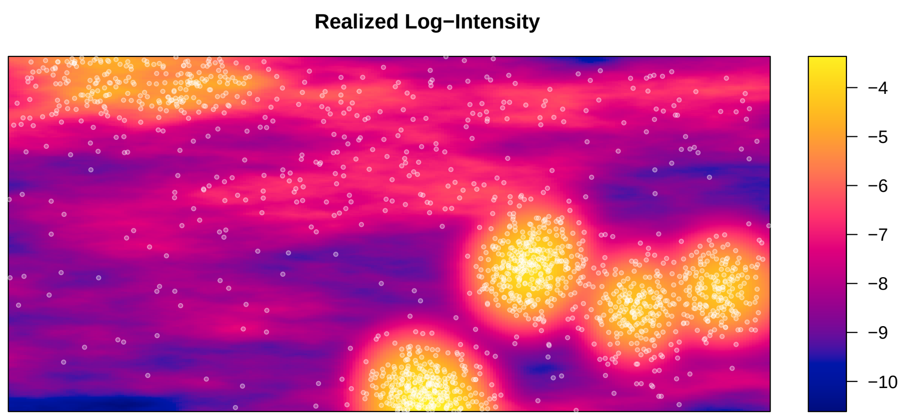
However, a small number of model fits suffered from apparent edge effects. For example, Figure 5 shows the prediction surface resulting from a serpentine transect plan. The predicted log-intensity has a hotspot of extremely large values in the southeast corner (notice the color scale). The hotspot is driven by two nodes on the boundary with very large prediction values. Another, less extreme, edge effect is present in the northeast corner.

4.2. MSPE and APV for all simulations

There are two important characteristics of a useful spatial prediction. First, deviation from the true surface should be low, and is indicated by low mean squared prediction error (MSPE). Second, the model's own assessment of prediction uncertainty will be used in practice to build trust in the inferences, so average prediction variance (APV) should be low.

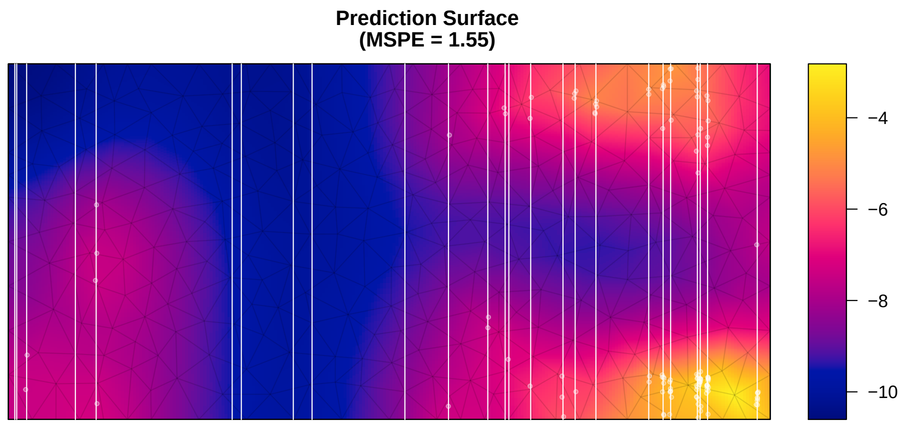


(a) Realized LGCP log-intensity

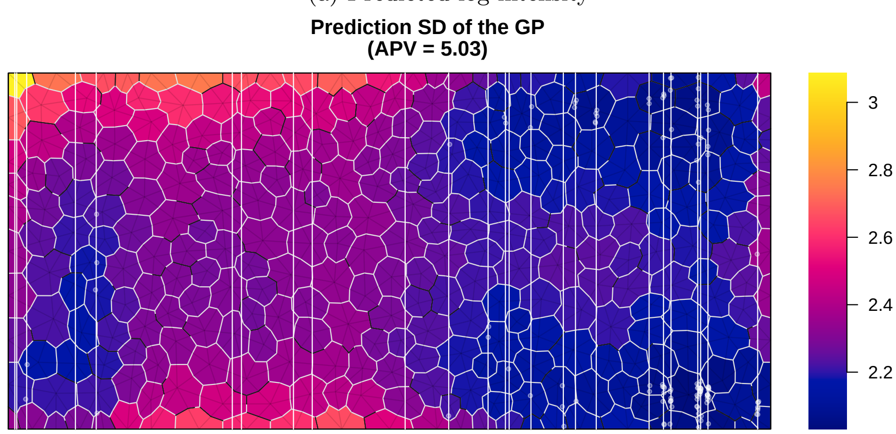


(b) Realized cluster log-intensity

Figure 3: The realized intensity function (natural log scale) and complete point pattern from a LGCP, and from a LGCP superposed with a cluster process.



(a) Predicted log-intensity



(b) Prediction SD

Figure 4: Predicted log-intensity function and prediction standard deviation using data observed via a SRS of line transects. The SD is shown for each finite element node.

Considered across all survey plans and prediction surfaces, both MSPE and APV had right-skewed distributions. Thus we use logarithmic scales for plots and summarize them using the median and interquartile range (IQR). Median MSPE decreases with increasing path distance, leveling off between 20000 and 30000 units of distance for the LGCP data but continuing to decrease through 50000 units for the clustered data (Figure 6). Variability (IQR) of MSPE also descreases as distance increases. Surfaces with edge effects form a cluster of large, outlying MSPE values. The lowest-MSPE prediction surfaces result from the longest Hilbert designs. Overall, there are no substantial differences among the different schemes with respect to median or variability in MSPE.

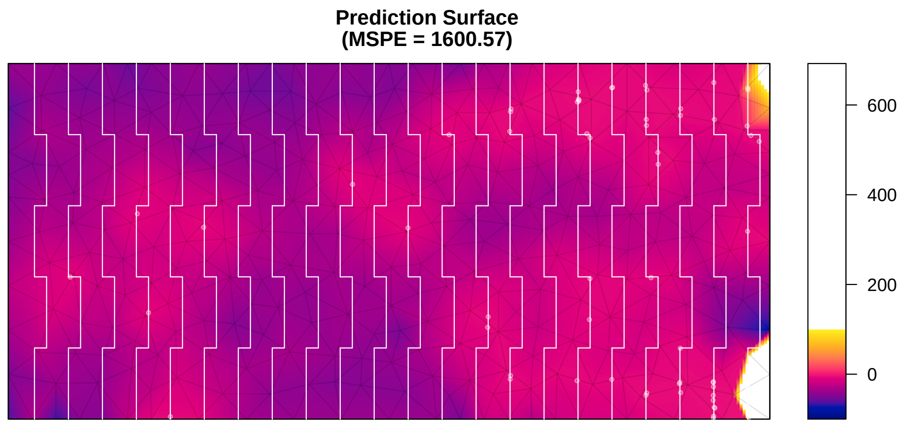
The relationship between APV and MSPE differs from realization to realization (Figure 7). Surfaces in the high-MSPE cluster have high APV, but otherwise there is little association between APV and MSPE.

4.3. Augmenting a poor-performing design

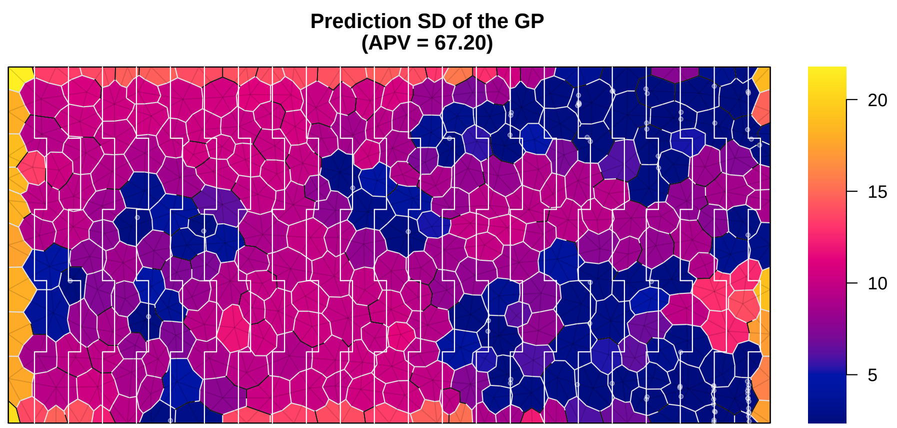
Even a poor-performing design could be used as a starting point for sequential design. As a simple illustration, we augment the design from Figure 5 with some additional sampling effort in the eastern part of the site, where the edge effects were seen in the prediction surface. The total distance surveyed increases from 17500 to 20180 units, while the predicted log-intensity surface is much more accurate (Figure 8). MSPE decreases from 1600.57 to 2.48 and APV improves from 67.20 to 11.02. Across the site, the prediction standard deviation is lower than before, and is now highest around the edge of the western half. If we were to continue adding segments to the path, giving some attention to the western portion of the site could further improve the prediction.

4.4. Spatial coverage

While the above results suggest the choice of design is relatively unimportant and the distance traveled is the main driver of the quality of the spatial predictions, it is important to consider that all of the design schemes ensure

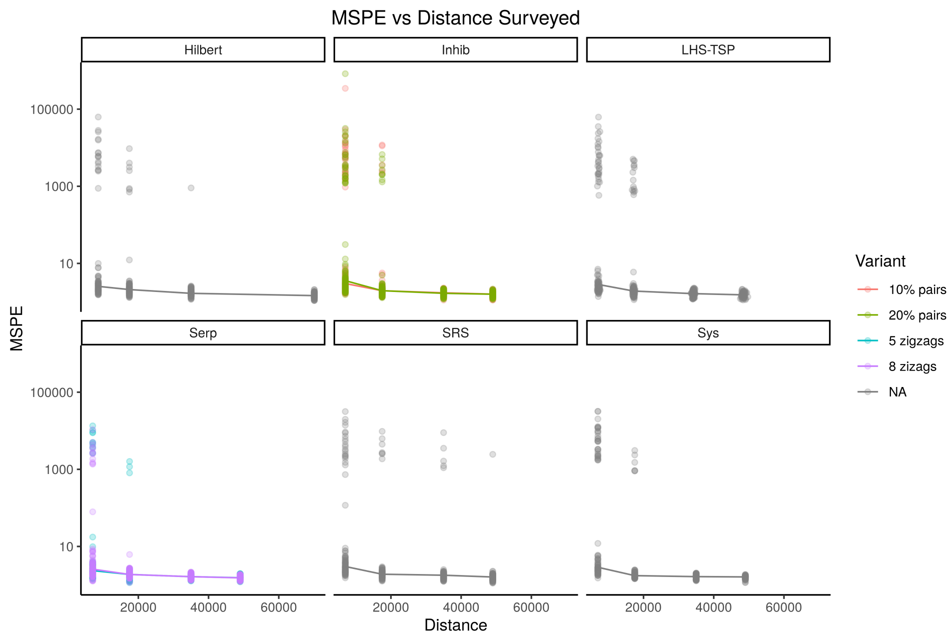


(a) Predicted log-intensity. The color scale is truncated at 100, but reaches a maximum of 692.

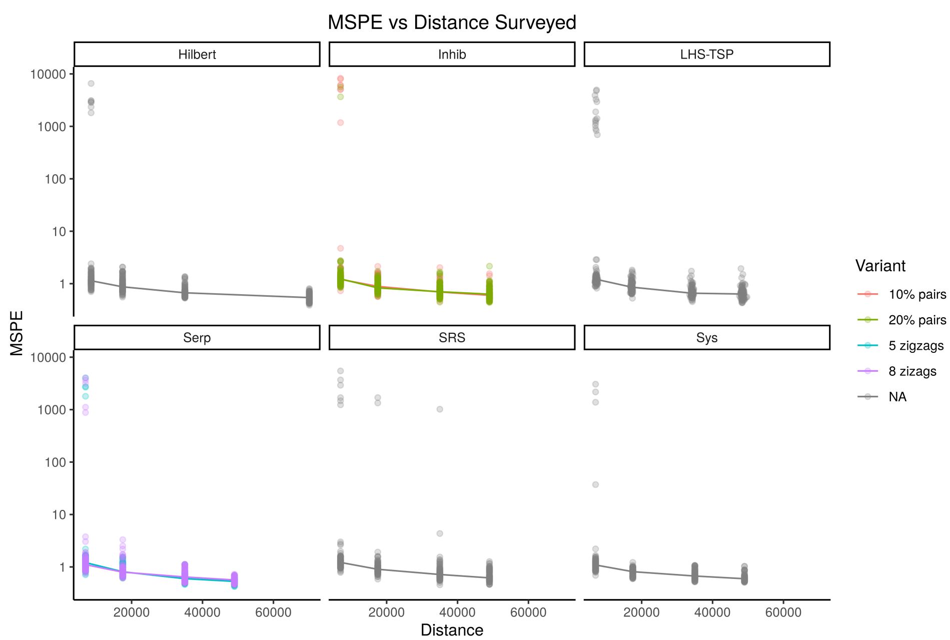


(b) Prediction SD

Figure 5: Predicted GP surface and prediction SD using data observed via a serpentine transect plan. The prediction has an apparent edge effect in the southeastern corner. The SD is high across much of the site.

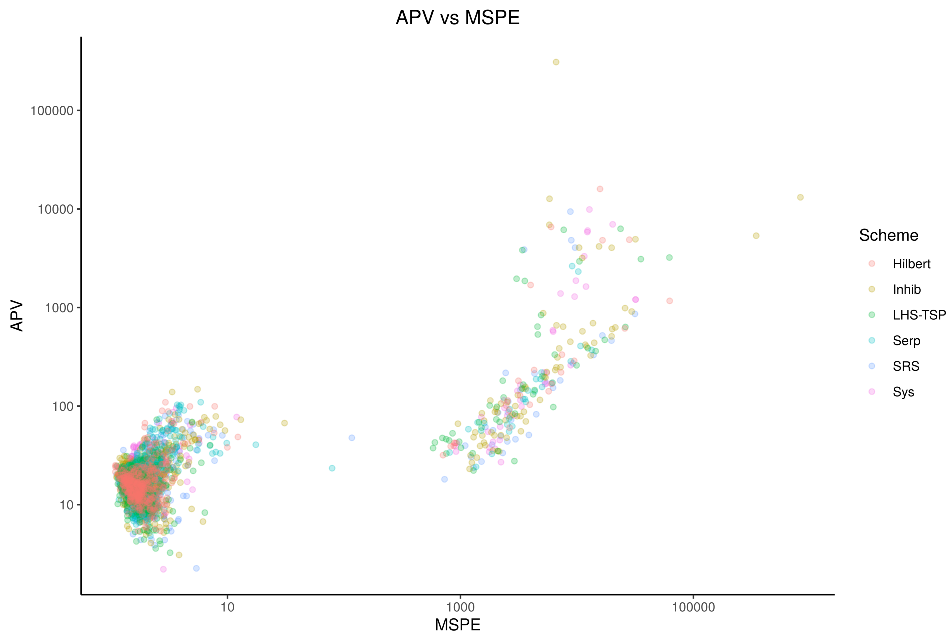


(a) MSPE vs length of the path for one LGCP dataset

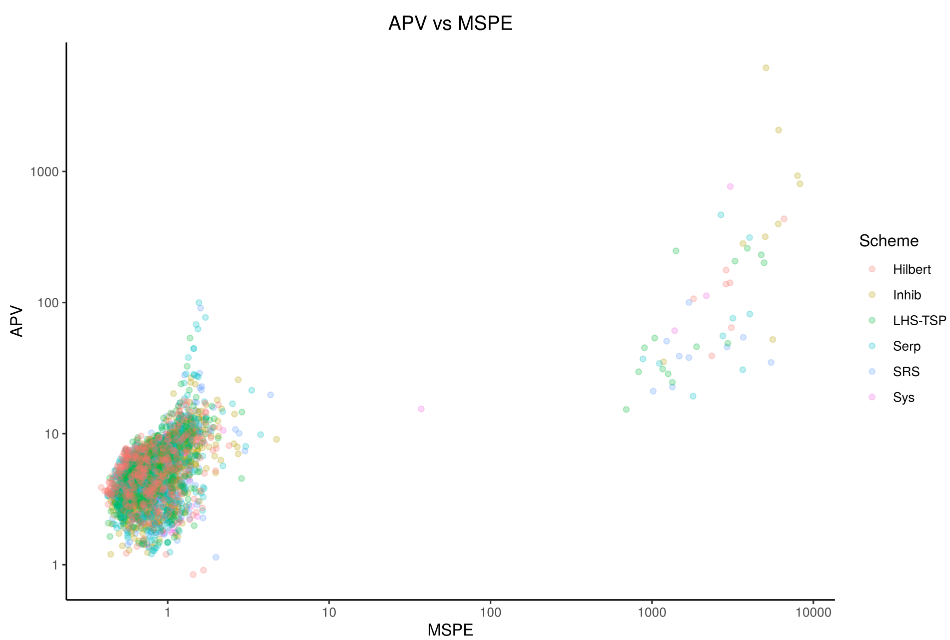


(b) MSPE vs length of the path for one clustered dataset

Figure 6: Plots of mean squared prediction error (MSPE) vs length of the path for each plan applied to one realization of a LGCP and one realization of an LGCP with a cluster process overlaid. Line segments connect the median MSPE at each group of distances. The plots are paneled by design scheme.

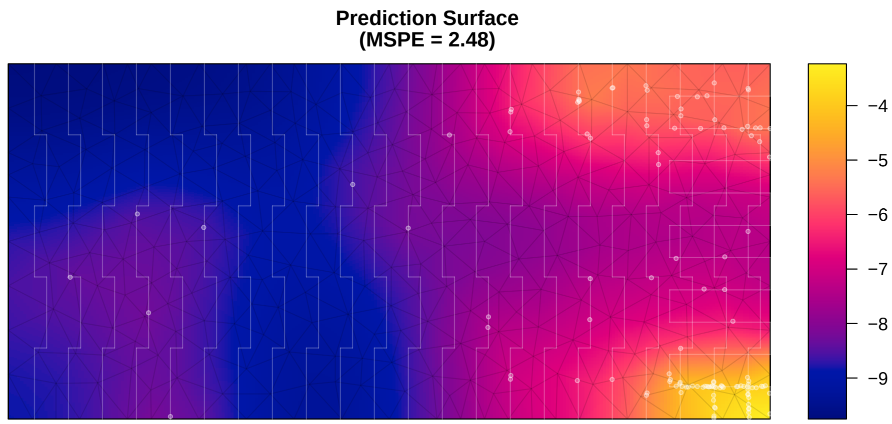


(a) APV vs MSPE for one LGCP dataset

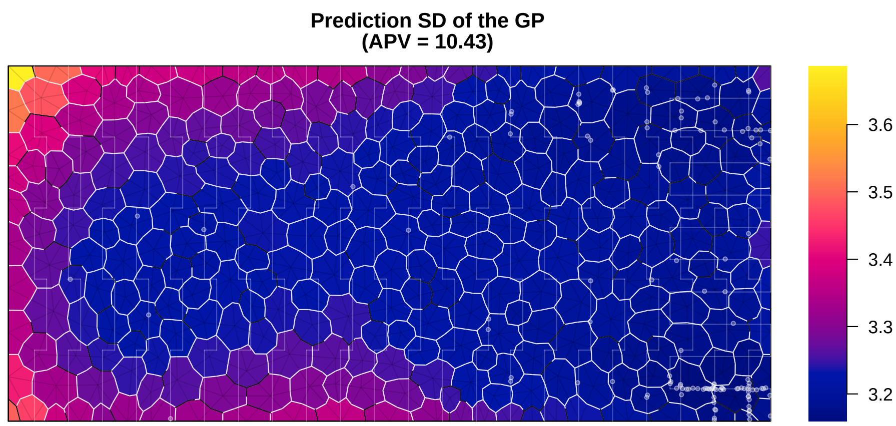


(b) APV vs MSPE for one clustered dataset

Figure 7: Plots of average prediction variance (APV) vs MSPE for each plan applied to one realization of a LGCP and one realization of an LGCP with a cluster process overlaid.



(a) Predicted log-intensity



(b) Prediction SD

Figure 8: Predicted log-intensity function and prediction standard deviation using data observed via a serpentine transect design augmented post-hoc. The SD is shown for each finite element node.

the path is distributed across the entire study region. Designs that leave large unexplored voids will not perform as well.

As an example, Figure 9 shows the results of using a systematic sample of 50 parallel transects in the western 20% of the study site. This design traverses a distance of 35000 units but leaves most of the site far from the observed path. As a result, the predicted log-intensity is flat near the GP posterior mean of -8.71 over most of the site, rendering the prediction mostly useless.

5. Discussion

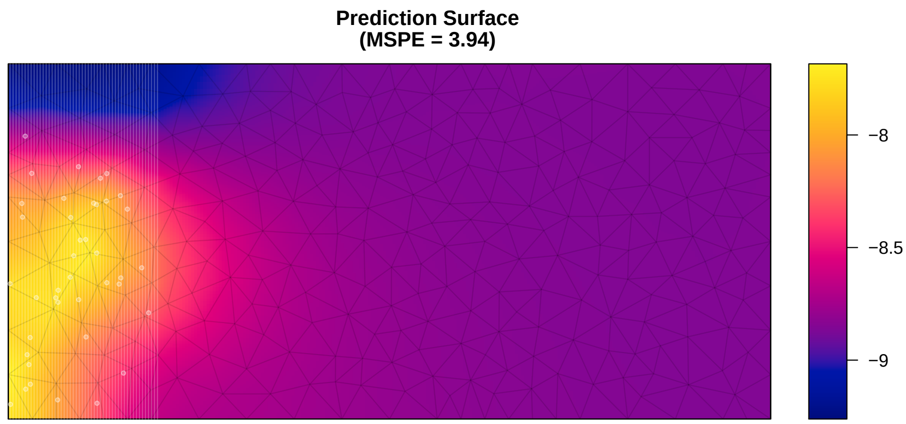
- increasing uncertainty along path — reasonable, path is narrow, could be events just out of detection range
- convergence problems/large variance solution is more data collection?
- discuss starting points for optimization and sequential design
- practical issue: path will be smoothed, no instantaneous direction changes at corners, equipment may have limitations which is why we looked at number and distribution or turn angles
- could incorporate turns into loss function or use multi-objective optimization (Lark, 2016)

6. Conclusions

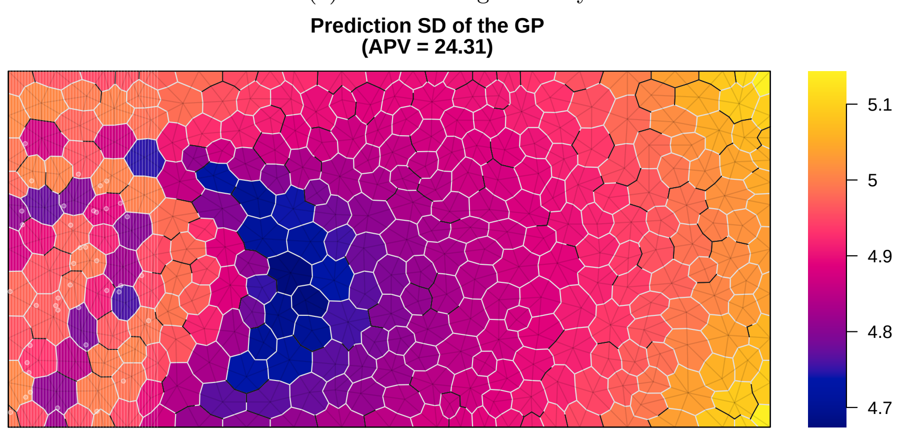
- choice of scheme does not matter much as long as it provides spatial coverage
- of the schemes considered here, only transect schemes have flexiblility in distance and/or a priori known distance

References

Anders, S., 2009. Visualisation of genomic data with the Hilbert curve. Bioinformatics doi:`10.1093/bioinformatics/btp152`.



(a) Predicted log-intensity



(b) Prediction SD

Figure 9: Predicted log-intensity function and prediction standard deviation using data observed via a systematic sample of a small section of the site. The SD is shown for each finite element node.

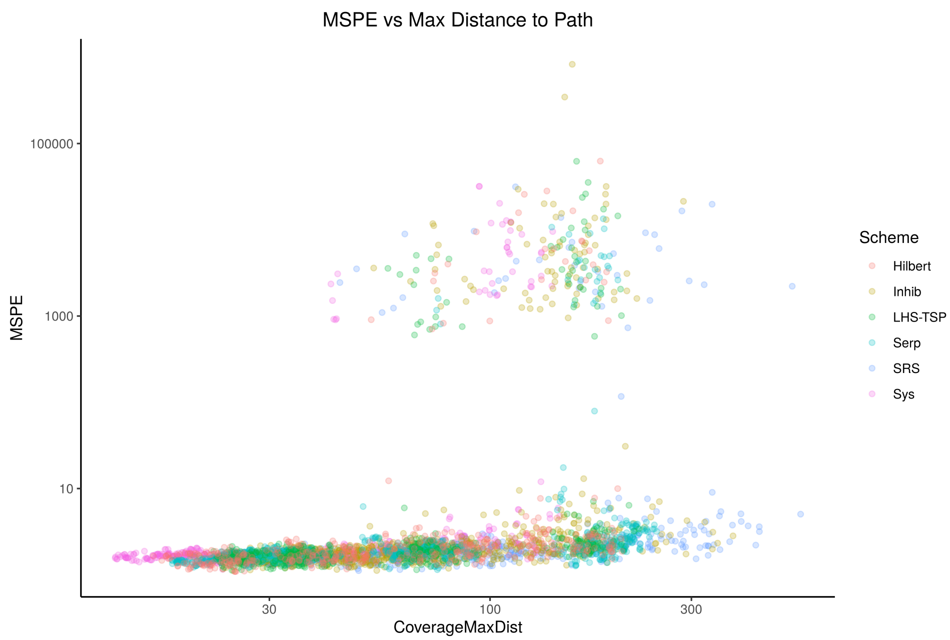


Figure 10: MSPE plotted against the maximum distance from any point in the site to the nearest point on the path.

Blangiardo, M., Cameletti, M., 2015. Spatial and Spatio-temporal Bayesian Models with R-INLA. Wiley.

Borkowski, J.J., Piepel, G.F., 2009. Uniform designs for highly constrained mixture experiments. Journal of Quality Technology 41, 35–47.

Carnell, R., 2020. lhs: Latin Hypercube Samples. URL: `https://CRAN.R-project.org/package=lhs`. R package version 1.0.2.

Chakraborty, A., Gelfand, A.E., Wilson, A.M., Latimer, A.M., Silander, J.A., 2011. Point pattern modelling for degraded presence-only data over large regions. Journal of the Royal Statistical Society: Series C (Applied Statistics) 60, 757–776.

Chipeta, M., Terlouw, D., Phiri, K., Diggle, P., 2017. Inhibitory geostatistical designs for spatial prediction taking account of uncertain covariance structure. Environmetrics 28.

Diggle, P., Lophaven, S., 2006. Bayesian geostatistical design. Scandinavian Journal of Statistics 33, 53–64.

Fan, J.A., Yeo, W.H., Su, Y., Hattori, Y., Lee, W., Jung, S.Y., Zhang, Y., Liu, Z., Cheng, H., Falgout, L., Bajema, M., Coleman, T., Gregoire, D., Larsen, R.J., Huang, Y., Rogers, J.A., 2014. Fractal design concepts for stretchable electronics. Nature communications 5, 3266.

Flagg, K.A., Hoegh, A., Borkowski, J.J., 2020. Modeling partially surveyed point process data: Inferring spatial point intensity of geomagnetic anomalies. Journal of Agricultural, Biological and Environmental Statistics 25, 186–205.

Fuglstad, G.A., Simpson, D., Lindgren, F., Rue, H., 2019. Constructing priors that penalize the complexity of Gaussian random fields. Journal of the American Statistical Association 114, 445–452.

Hahsler, M., Hornik, K., 2020. TSP: Traveling Salesperson Problem (TSP). URL: `https://CRAN.R-project.org/package=TSP`. R package version 1.1-10.

Husslage, B.G., Rennen, G., Van Dam, E.R., Den Hertog, D., 2011. Space-filling latin hypercube designs for computer experiments. Optimization and Engineering 12, 611–630.

Illian, J.B., Sørbye, S.H., Rue, H., 2012. A toolbox for fitting complex spatial point process models using integrated nested laplace approximation (inla). The Annals of Applied Statistics , 1499–1530.

Johnson, D., Laake, J., VerHoef, J., 2014. DSpat: Spatial Modelling for Distance Sampling Data. URL: https://CRAN.R-project.org/package=DSpat. R package version 0.1.6.

Kelly, F.P., Ripley, B.D., 1976. A note on Strauss's model for clustering. Biometrika 63, 357–360.

Lark, R., 2016. Multi-objective optimization of spatial sampling. Spatial Statistics 18, 412–430.

Lindgren, F., Rue, H., Lindström, J., 2011. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73, 423–498.

Liu, J., Vanhatalo, J., 2020. Bayesian model based spatiotemporal survey designs and partially observed log Gaussian Cox process. Spatial statistics 35, 100392.

Ma, Q., Zhang, Y., 2016. Mechanics of fractal-inspired horseshoe microstructures for applications in stretchable electronics. Journal of Applied Mechanics 83.

Matzke, B., Wilson, J., Newburn, L., Dowson, S., Hathaway, J., Sego, L., Bramer, L., Pulsipher, B., 2014. Visual Sample Plan Version 7.0 User's Guide. Pacific Northwest National Laboratory. Richland, Washington. URL: http://vsp.pnnl.gov/docs/PNNL-23211.pdf.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21, 239–245.

Møller, J., Syversveen, A.R., Waagepetersen, R.P., 1998. Log Gaussian Cox processes. Scandinavian journal of statistics 25, 451–482.

Ogorzałek, M.J., 2009. Fundamentals of fractal sets, space-filling curves and their applications in electronics and communications, in: Intelligent Computing Based on Chaos. Springer, pp. 53–72.

Pollard, J., Palka, D., Buckland, S., 2002. Adaptive line transect sampling. Biometrics 58, 862–870.

R Core Team, 2019. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: `https://www.R-project.org/`.

Rue, H., Martino, S., Chopin, N., 2009. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the royal statistical society: Series b (statistical methodology) 71, 319–392.

Sagan, H., 1994. Space-filling curves. Springer.

Simpson, D., Illian, J.B., Lindgren, F., Sørbye, S.H., Rue, H., 2016. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. Biometrika 103, 49–70.

Simpson, D., Rue, H., Riebler, A., Martins, T.G., Sørbye, S.H., et al., 2017. Penalising model component complexity: A principled, practical approach to constructing priors. Statistical science 32, 1–28.

Strauss, D.J., 1975. A model for clustering. Biometrika 62, 467–475.

USACE, 2015. Technical Guidance for Military Munitions Response Actions. Technical Report EM 200-1-15. United States Army Corps of Engineers. URL: http://www.publications.usace.army.mil/Portals/76/Publications/EngineerManuals/EM_200-1-15.pdf.

Yuan, Y., Bachl, F.E., Lindgren, F., Borchers, D.L., Illian, J.B., Buckland, S.T., Rue, H., Gerrodette, T., et al., 2017. Point process models for spatio-temporal distance sampling data from a large-scale survey of blue whales. The Annals of Applied Statistics 11, 2270–2297.