

مروری بر طرح شبکه بهینه فضایی

زهره صمدی^۱، هوشنگ طالبی^۱، فیروزه ریواز^۲

^۱گروه آمار، دانشکده علوم ریاضی، دانشگاه اصفهان

^۲گروه آمار، دانشکده علوم ریاضی، دانشگاه شهید بهشتی

چکیده: پایش فرایندهای محیطی از موضوعات مورد توجه در بسیاری از علوم نظیر سلامت عمومی، محیط زیست و تغییرات آب و هوایی است. یکی از مسائل مهم در این زمینه، تعیین موقعیت‌های سنجش است. از آنجا که چگونگی انتخاب موقعیت‌های نمونه‌گیری تاثیر به‌سزایی بر نتایج، تحلیل‌ها و اعتبار آن‌ها دارد؛ لازم است این موقعیت‌ها به‌گونه‌ای انتخاب شوند که بر اساس مشاهدات به‌دست آمده در آن‌ها، بیشترین میزان اطلاعات به‌دست‌آید تا بتوان تحلیل‌های بهینه‌ای ارائه کرد. این مساله که به طرح شبکه فضایی بهینه معروف است از موضوعات اساسی و پرکاربرد در آمار فضایی است. با توجه به ضرورت انجام یک پژوهش با حداکثر کارایی، باید شیوه‌ای علمی در طراحی شبکه به‌کار گرفته شود. برای این منظور با توجه به اهداف مطالعه تابع مناسبی به‌عنوان معیار بهینگی انتخاب شده و طرح نهایی بر اساس آن تعیین می‌شود. در مطالعه حاضر رویکردهای مختلف طراحی شبکه فضایی برای اهداف متفاوت به اختصار بررسی می‌شود.

واژه‌های کلیدی: داده‌های فضایی، طراحی شبکه، معیار بهینگی، پیشگویی فضایی، همبستگی فضایی.

کد موضوع‌بندی ریاضی (۲۰۱۰): 62H11, 62M30

۱ مقدمه

شیوه جمع‌آوری داده‌ها در همه مطالعات از اهمیت ویژه‌ای برخوردار است. یکی از موضوعات اساسی در این زمینه، تعیین طرح نمونه‌گیری بهینه است. در سال‌های اخیر و با افزایش روزافزون مطالعات محیطی نظیر کشاورزی، هواشناسی، اقلیم‌شناسی و به‌طور کلی مطالعاتی که در آن‌ها با داده‌های فضایی مواجهیم، پژوهشگران شیوه‌های جدیدی برای جمع‌آوری این داده‌ها ارائه کردند. این روش‌ها را می‌توان به سه رهیافت هندسه‌پایه، احتمال‌پایه و مدل‌پایه تقسیم کرد (گلفند و همکاران، ۲۰۱۰). در رهیافت هندسه‌پایه، موقعیت‌های نمونه‌ای به گونه‌ای انتخاب می‌شوند که کل فضای مورد مطالعه را پوشش دهند

^۱ نام و آدرس پست الکترونیکی ارائه‌دهنده مقاله: زهره صمدی Za.samadi@gmail.com

^۲Space filling design

(دالینوس و همکاران ، ۱۹۶۰). طرح‌های مشبکه‌ای منظم مانند شبکه‌های مثلثی و مربعی و طرح پوشاننده فضایی ^۲ (نیچکا و سالتزمن ، ۱۹۹۸؛ رویل و نیچکا ، ۱۹۹۸) از این دسته‌اند. کاکس و همکاران (۱۹۹۷) عنوان کردند که طرح‌های هندسه‌مبنا برای بسیاری از اهداف مطالعاتی کاربرد دارند. مولر (۲۰۰۵) این طرح‌ها را رویکردی مناسب برای کشف ساختار فرایند تحت مطالعه معرفی کرد.

اگر هدف مطالعه، برآورد پارامترهای جامعه هدف مانند میانگین، مجموع کل یا نسبت باشد، می‌توان رهیافت احتمال‌پایه را به‌کار گرفت و برآورد پارامترهای جامعه را با دقتی معلوم از یک نمونه احتمالی به‌دست آورد. در نمونه‌گیری احتمالی، واحدهای نمونه‌گیری به تصادف انتخاب می‌شوند و اگر به درستی انجام شود، احتمال انتخاب واحدها معلوم و پایه‌ای برای استنباط آماری خواهد بود. از انواع روش‌های نمونه‌گیری با این دیدگاه می‌توان به نمونه‌گیری تصادفی، نمونه‌گیری با طبقه‌بندی و نمونه‌گیری با احتمال متغیر اشاره کرد. برای مطالعه بیشتر **دوبی و همکاران (۲۰۰۷)** را ببینید.

در بسیاری از پژوهش‌ها، تحلیل‌های فضایی بخش مهمی از اهداف مطالعه هستند. به عنوان مثال در بررسی قیمت مسکن علاوه بر برآورد میانگین قیمت مسکن، شناسایی الگوی فضایی تغییرات قیمت مسکن نیز مد نظر پژوهشگر است. در چنین مواردی رویکرد مدل‌پایه به‌کار گرفته می‌شود. در این رویکرد، استنباط بر اساس مدلی تصادفی از تغییر پذیری کمیت مورد مطالعه صورت می‌پذیرد و روش انتخاب واحدها، نمونه‌گیری غیر احتمالی است. یکی از رایج‌ترین نمونه‌گیری‌های غیر احتمالی، نمونه‌گیری قصدی است. در این نمونه‌گیری سعی می‌شود واحدهای نمونه‌گیری چنان انتخاب شوند که هدفی معین را به بهترین نحو برآورده سازند.

از عمده کاربردهای طرح‌های مدل‌پایه، طراحی شبکه‌های پایش از جمله شبکه‌های سنجش فرایندهای محیطی است. این طرح‌ها ممکن است آینده‌نگر یا گذشته‌نگر باشند. در طرح‌های آینده‌نگر، موقعیت‌های نمونه‌ای از قبل وجود ندارند و باید موقعیت‌های جدید تعیین شوند. اما در طرح‌های گذشته‌نگر از قبل موقعیت‌هایی برای نمونه‌گیری وجود دارند و اطلاعاتی از آن‌ها در دسترس است. در این حالت هدف مطالعه افزایش یا کاهش موقعیت‌های نمونه‌گیری است. در چنین مواردی بر اساس رویکرد مدل‌مبنا از یک مدل آماری برای توصیف فرایندهای محیطی استفاده می‌شود و هدف مطالعه، انتخاب طرح بهینه برای پیشگویی در موقعیت‌های نامعلوم یا برآورد پارامترها است.

این مقاله به مرور رهیافت‌های مختلف در طراحی شبکه بهینه فضایی می‌پردازد. در بخش دوم، مطالعات انجام شده با هدف برآورد پارامترهای میدان تصادفی ارائه می‌شود و در بخش سوم شبکه‌های بهینه برای پیشگویی کمیت مورد مطالعه در مکان‌های فاقد مشاهده بررسی می‌شوند.

۲ طرح شبکه بهینه برای برآورد پارامترها

در آمار فضایی از یک میدان تصادفی $Z(\cdot) = \{Z(s); s \in D\}$ که در آن $D \in \mathbb{R}^d$ و $d \geq 1$ است، برای مدل‌بندی داده‌ها استفاده می‌شود. توابع میانگین و کوواریانس میدان تصادفی به ترتیب به صورت زیر تعریف می‌شوند.

$$\mu(s) = E(Z(s)); \quad s \in D$$

$$C(s, s') = Cov(Z(s), Z(s')); \quad s, s' \in D$$

اغلب فرض می‌شود $\mu(s) = f'(s)\beta$ که در آن $f(s) = (f_1(s), \dots, f_p(s))'$ یک بردار $1 \times p$ از مولفه‌های معلوم و $\beta = (\beta_1, \dots, \beta_p) \in \mathbb{R}^p$ بردار پارامترهای رگرسیونی است. هم‌چنین فرض می‌شود $C(s, s') = \sigma^2 \rho(s, s'; \theta)$ که در آن σ^2 واریانس میدان تصادفی و $\rho(\cdot, \cdot, \theta)$ تابع همبستگی میدان با پارامتر θ است. طبق این فرض، فرم تابع کوواریانس معلوم و به برداری از پارامترها مانند θ بستگی دارد. دو سوال اساسی در ارتباط با میدان تصادفی مطرح است:

۱. طرح بهینه برای برآورد تابع میانگین یا روند چگونه تعیین می‌شود.
۲. برای برآورد پارامترهای تابع کوواریانس، چه طرحی بهینه است.

۱.۲ برآورد پارامترهای تابع میانگین: رهیافت مدل رگرسیون

ایده اصلی این رهیافت، تقریب میدان تصادفی همسانگرد با تابع کوواریانس معلوم توسط یک مدل رگرسیون خطی با ضرایب تصادفی است. این تقریب امکان استفاده از نظریه بهینه‌سازی محدب در طراحی شبکه فضایی را فراهم می‌سازد. **فدرو و مولر (۱۹۸۹)** با استفاده از بسط کارهونن-لو تقریبی از میدان تصادفی ارائه کردند. اما استفاده از این تقریب در عمل پیچیده است و در حالت کلی نمی‌توان الگوریتم‌های عددی قابل اعتمادی برای حل این مسائل یافت. **اسپوک و پیلز (۲۰۰۸)** تقریب دیگری بر اساس نمایش طیفی قطبی میدان تصادفی پیشنهاد کردند. این روش نیز در عمل با چالش مواجه است زیرا بیشتر فرایندهای محیطی ویژگی همسانگردی را ندارند و ساختار همبستگی نیز عموماً نامعلوم است. بنابراین اگرچه رهیافت مدل رگرسیونی در طراحی شبکه، دسترسی به نظریه طرح آزمایش‌ها را فراهم می‌کند اما در عمل با مشکلات متعددی روبه‌رو است. برای مطالعه بیشتر **مولر (۲۰۰۷)** و **اسپوک (۲۰۱۱)** را ببینید.

۲.۲ برآورد پارامترهای تابع کوواریانس

در آمار فضایی برای بیان میزان وابستگی فضایی داده‌ها از تغییرنگار استفاده می‌شود. واریانس اختلاف بین مقادیر میدان تصادفی در دو موقعیت s و $s+h$ تغییرنگار نامیده شده و به صورت زیر تعریف می‌شود.

$$\gamma(h) = \text{Var}(Z(s+h) - Z(s))$$

طراحی بسیاری از شبکه‌های پایش از جمله مطالعات **روسو (۱۹۸۴)**، **واریک و مایرز (۱۹۸۷)** و **موریس (۱۹۹۱)** مبتنی بر تغییرنگار است. **مولر و زیمرمن (۱۹۹۹)** درمیان ماتریس کوواریانس برآوردگرهای کمترین توان‌های دوم پارامترهای تغییرنگار را به عنوان معیار بهینگی در نظر گرفته و برای حل مشکل وابستگی این معیار به پارامترهای نامعلوم، طرح‌های دنباله‌ای را پیشنهاد کردند. آن‌ها نشان دادند شبکه حاصل، نسبت به طرح‌های شبکه‌ای منظم و پوشاننده فضایی کارایی بهتری دارد. با این وجود **استاین (۱۹۹۹)** نشان داد این روش با محدودیت جدی مواجه است.

ژو و استاین (۲۰۰۵) میدان تصادفی را گاوسی در نظر گرفته و نشان دادند شرایط نظم برقرار است. آن‌ها لگاریتم درمیان ماتریس اطلاع فشر را به عنوان معیار بهینگی معرفی کردند و از آن‌جا که این معیار نیز به پارامترهای نامعلوم وابسته است، برای حل این مشکل سه رهیافت طرح‌های بهینه موضعی، مینی‌ماکس و بی‌زی را پیشنهاد دادند. **زیمرمن (۲۰۰۶)** با استفاده از معیار بهینگی **ژو و استاین (۲۰۰۵)** مطالعات شبیه‌سازی انجام داد و نتیجه گرفت برای دستیابی به برآوردهای دقیق از پارامترهای تابع کوواریانس، تعداد موقعیت‌های نمونه‌گیری نزدیک به هم در شبکه باید بسیار بیشتر از طرح‌های منظم یا تصادفی باشد. هم‌چنین شبکه باید شامل تعداد مناسبی از موقعیت‌های دور از هم نیز باشد. بنابراین شبکه بهینه باید به صورت خوشه‌هایی با فواصل منظم باشد.

۳ طراحی شبکه بهینه با هدف پیشگویی

یکی از اهداف اصلی تحلیل‌های فضایی، پیشگویی مقدار نامعلوم میدان تصادفی در موقعیت s بر اساس بردار مشاهدات است. به دلیل اهمیت پیشگویی، مطالعات زیادی برای طراحی شبکه با این هدف صورت گرفته است. ایده اصلی بخشی از این مطالعات روش کریگیدن است که پیشگوی ناریب خطی با واریانس کمینه را برای پیشگویی $Z(s_0)$ بر اساس

$Z = (Z(s_1), \dots, Z(s_n))$ ارائه می‌دهد. پیشگوی کریگی عام و واریانس آن به ترتیب به صورت زیر است.

$$\hat{Z}(s_0) = \lambda'Z$$

$$\sigma^2(s_0) = \sigma^2 - \lambda'c + x'm$$

که در آن $X, m = (X'\Sigma^{-1}X)^{-1}(x' - X'\Sigma^{-1}c)$, $\lambda = (c + X(X'\Sigma^{-1}X)^{-1}(x' - X'\Sigma^{-1}c))\Sigma^{-1}$ ماتریس متغیرهای کمکی در موقعیت‌های s_1 و s_2 و \dots و بردار متغیرهای کمکی در موقعیت s_0 است (کرسی، ۱۹۹۳). با توجه به وابستگی پیشگویی فضایی به پارامترهای تابع کوواریانس، رده‌ای از مطالعات شبکه‌های بهینه با فرض معلوم بودن تابع کوواریانس طراحی شده‌اند. پژوهش‌های مک‌براتی و همکاران (۱۹۸۱)، سو و کابانیز (۱۹۹۳) و ریتز (۱۹۹۶) از این دست است. دو معیار بهینگی پرکاربرد در این مطالعات، متوسط واریانس کریگیدن^۳ و بیشینه واریانس کریگیدن^۴ است:

$$AKV = \int_D \sigma^2(s_0) w(s_0) ds_0$$

$$MKV = \sup_D \{ \sigma^2(s_0) w(s_0), s_0 \in D \}$$

که در آن $w(s_0)$ وزن تعیین شده برای موقعیت s_0 است. بروس و هیولینک (۲۰۰۷) کمینه متوسط واریانس کریگیدن را به عنوان معیار بهینگی معرفی کرده و شبکه‌ای برای پایش آب‌های زیرزمینی طراحی کردند. این معیارها به پارامترهای ساختار همبستگی وابسته‌اند و در مسائلی می‌توان آن‌ها را به‌کار برد که این ساختار معلوم باشد، اما در بیشتر کاربردهای عملی این فرض برقرار نیست. در چنین مواردی برای لحاظ کردن تاثیر عدم قطعیت پیشگویی، معیارها باید به نحوی تعدیل شوند. ژو و استاین (۲۰۰۶) طراحی شبکه را در حالتی بررسی کردند که میدان تصادفی گاوسی و ساختار همبستگی نامعلوم است. آن‌ها از روش جایگذاری^۵ در برآورد پارامترهای مدل استفاده کردند. این شیوه یک رهیافت کلاسیک برای پیشگویی فضایی است که توسط محققان زیادی پیشنهاد شده است. در ادامه برای لحاظ کردن خطای برآورد پارامترها، معیار AKV را تعمیم داده و معیار جدیدی به نام EA^۶ معرفی کردند. همچنین با استفاده از این معیار بهینگی شبکه پایش آلاینده SO_۲ را بازطراحی نموده و تعداد ایستگاه‌های سنجش را از ۱۰۱ به ۵۰ ایستگاه کاهش دادند. زیمرمن (۲۰۰۶) با استفاده از تقریب هارویل و جسک (۱۹۹۲) و زیمرمن و کرسی (۱۹۹۲) برای واریانس پیشگویی کریگیدن، یک معیار بهینگی برای پیشگویی میدان تصادفی با ساختار همبستگی نامعلوم به نام کریگیدن تجربی^۷ ارائه و عملکرد این معیار را در افزایش ایستگاه‌های شبکه پایش باران‌های اسیدی بررسی کرد.

معیارهای بهینگی معرفی شده، موجب کاهش عدم حتمیت برآورد پارامترهای مدل یا پیشگویی در مکان‌های فاقد مشاهده می‌شود. حال این سوال مطرح است آیا معیاری وجود دارد که علاوه بر این‌ها، عدم حتمیت مرتبط با اندازه‌گیری کمیت مورد مطالعه را نیز کنترل کند؟

۱.۳ معیار بهینگی آنتروپی

عدم حتمیت مرتبط با اندازه‌گیری متغیر مورد بررسی را با استفاده از آنتروپی می‌توان مدل‌بندی نمود. آنتروپی، دو هدف دیگر یعنی کنترل عدم حتمیت مرتبط با برآورد پارامترهای تابع کوواریانس و پیشگویی را نیز تامین می‌کند. آنتروپی بردار تصادفی $Z = (Z(s_1), \dots, Z(s_n))$ که دارای توزیع f است به صورت $H(Z) = E(-\log f(z))$ تعریف می‌شود.

³ Average Kriging Variance

⁴ Maximum Kriging Variance

⁵ Plug-in

⁶ Estimation Adjusted

⁷ Empirical Kriging

ایده استفاده از آنتروپی در طرح آزمایش‌ها به لیندلی (۱۹۵۶) برمی‌گردد. پژوهش‌های کسلتون و زیدک (۱۹۸۴)، گاترپ و همکاران (۱۹۹۳) و زیدک و همکاران (۲۰۰۰) دیدگاه طرح ماکسیمم آنتروپی را توسعه داد که منجر به بیشترین اطلاع از ایستگاه‌های بالقوه فاقد مشاهده شد. فوئنتس و همکاران (۲۰۰۷) با استفاده از نظریه آنتروپی، طرح بهینه را با هدف کاهش ایستگاه‌های اندازه‌گیری آلاینده‌های هوا به نصف ایستگاه‌های موجود ارائه دادند. در مطالعات کاهش شبکه می‌توان بردار متغیر مورد مطالعه Z را به صورت $Z = (Z_1, Z_2)$ افزایش کرد که Z_2 بردار مشاهدات در موقعیت‌هایی است که قرار است حذف شوند. اگر آنتروپی کل Z با $H(Z_1, Z_2)$ نمایش داده شود، با توجه به ویژگی‌های آنتروپی خواهیم داشت:

$$H(Z_1, Z_2) = H(Z_1 | Z_2) + H(Z_2).$$

از آنجا که $H(Z_1, Z_2)$ ثابت است، افزایش $H(Z_2)$ معادل با کاهش $H(Z_1 | Z_2)$ است. بنابراین طرحی بهینه است که در آن آنتروپی توزیع $Z_1 | Z_2$ مینیمم یا به‌طور معادل آنتروپی توزیع Z_2 ماکسیمم شود. اگر میدان تصادفی گاوسی باشد، ماکسیمم کردن آنتروپی Z_2 با ماکسیمم کردن لگاریتم دترمینان ماتریس کوواریانس آن معادل است.

اگر هدف مطالعه افزایش تعداد موقعیت‌ها باشد، آنتروپی کل Z را می‌توان به صورت

$$H(Z) = H(Z_1, Z_2) = H(Z_2 | Z_1) + H(Z_1)$$

نشان داد که Z_2 بردار مشاهدات در موقعیت‌های جدید است. با توجه به ثابت بودن موقعیت‌های کنونی، $H(Z_1)$ ثابت است. بنابراین باید آنتروپی توزیع $Z_2 | Z_1$ ماکسیمم شود که در میدان تصادفی گاوسی معادل با ماکسیمم $\log |\Sigma_{Z_2|Z_1}|$ است. لازم به ذکر است که این معیارها به پارامترهای نامعلوم وابسته‌اند و در مواجهه با این مساله رهیافت جایگذاری یا بیزی به کار می‌رود.

با توجه به تاثیر شبکه بهینه در پژوهش‌ها، محققان روش‌های جدیدی برای گسترش این حوزه ارائه کرده‌اند. لی و زیمرمن (۲۰۱۵) طراحی شبکه برای مشاهدات فضایی چندمتغیره‌ی گاوسی را مطالعه و برای برآورد ساختار همبستگی و پیشگویی میدان تصادفی با ساختار همبستگی معلوم و نامعلوم، شبکه بهینه را طراحی کرده و استواری این طرح‌ها را نسبت به میزان همبستگی فضایی بررسی کردند. ریواز (۲۰۱۶) با استفاده از معیار بهینگی آنتروپی، طراحی شبکه بهینه برای پیشگویی داده‌های فضایی چندمتغیره ناگوسی را بررسی کرد. وادوکس و همکاران (۲۰۱۷) با تعمیم روش پیشگویی کریگیدن، یک معیار بهینگی برای طراحی شبکه فرایندهای نامانا معرفی کردند. بیات و همکاران (۲۰۱۹) برای طراحی شبکه با هدف پیشگویی، دو معیار مبتنی بر کریگیدن عام و آنتروپی را بررسی کرده و نشان دادند شبکه طراحی شده بر اساس آنتروپی کارا تر است. سنارانت و همکاران (۲۰۲۰) با استفاده از تابع زیان، معیار جدیدی برای کاهش عدم حتمیت ارائه دادند. تحقیقات آنها نشان می‌دهد که این معیار جدید در مقایسه با معیارهایی که برای برآورد پارامترها یا پیشگویی به کار می‌روند، عملکرد بسیار بهتری دارد.

بحث و نتیجه‌گیری

با توجه به وابستگی تحلیل‌های آماری به الگوی موقعیت‌های نمونه‌ای، طراحی شبکه بهینه برای پایش فرایندها اهمیت بسیاری دارد. شبکه‌های پایش با اهداف مختلفی نظیر برآورد پارامترهای تابع میانگین، کوواریانس و پیشگویی طراحی می‌شوند و انتخاب معیار بهینگی نیز متناسب با هدف مطالعه انجام می‌گیرد. در این مقاله، مطالعات انجام شده در این زمینه و معیارهای بهینگی مناسب با هر یک از اهداف مورد بررسی قرار گرفت. یک نکته مهم در مورد شبکه‌های پایش آن است که این شبکه‌ها فارغ از چگونگی طراحی باید به‌صورت دوره‌ای بازبینی شده و در صورت لزوم بازطراحی شوند. چیزی که در عمل به ندرت اتفاق می‌افتد.

مراجع

- Bayat, B., Hosseini, K., Nasser, M., and Karami, H. (2019), Challenge of Rainfall Network Design Considering Spatial Versus Spatiotemporal Variations, *Journal of Hydrology*, doi: <https://doi.org/10.1016/j.jhydrol.2019.04.091>
- Brus, D.J. and Heuvelink, G. B.M. (2007), Optimization of sample patterns for universal kriging of environmental variables, *Geoderma*, **138**, 86-95.
- Caseltan, W. F. and Zidek, J. V. (1984), Optimal design for variogram estimation, *Statistics and Probability Letters*, **2**, 223–227.
- Cox, D.D., Cox, L.H., and Ensor, K.B. (1997), Spatial sampling and the environment: some issues and directions, *Environmental and Ecological Statistics*, **4**, 219–233.
- Cressie, N. (1993), *Statistics for Spatial Data*, Wiley, New York.
- Dalenius, T., Hajek, J., and Zubrzycki, S. (1960), On plane sampling and related geometrical problems, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, Vol I, Berkeley: University of California Press, 125–150.
- Dobbie, M.J., Henderson, B.L., and Stevens Jr., D.L. (2007), Sparse sampling: spatial design for aquatic monitoring, *Statis Surv*, **2**, 113–153.
- Fedorov, V.V., and Muller, W. (1989), Comparison of two approaches in the optimal design of an observation network, *Statistics*, **3**, 339–351.
- Fuentes, M., Chaudhuri, A., and Holland, D. (2007), Bayesian entropy for spatial sampling design of environmental data, *Environmetrics Ecological statistics*, **14**, 323–340.
- Gelfand A.E., Diggle, P.J., Fuentes, M. and Guttorp P. (2010), *Handbook of Spatial Statistics*, Chapman and Hall, New York.
- Guttorp, P., Le, N., Sampson, P. D., and Zidek, J. V. (1993), Using entropy in the redesign of an environmental monitoring network, *Applied Statistics*, **49**, 63–79.
- Harville, D.A., and Jeske, D.R. (1992), Mean squared error of estimation or prediction under a general linear model, *Journal of the American Statistical Association*, **87**, 724–731.
- Li, J. and Zimmerman, D.L. (2015), Model-based sampling design for multivariate geostatistics, *Technometrics*, **57**, 75–86.
- Lindley, D. (1956), On the measure of the information provided by an experiment, *Ann. Math. Statist.*, **27**, 968–1005.

- McBratney, A.B., Webster, R., and Burgess, T.M. (1981), The design of optimal sampling schemes for local estimation and mapping of regionalized variables. I—Theory and method, *Comp Geosci*, **7**, 331–334.
- Morris, M. D. (1991), On counting the number of data pairs for semivariogram estimation, *Mathematical Geology*, **23**, 929-943.
- Muller, W.G. (2007), *Collecting Spatial Data: Optimum Design of Experiments for Random Fields*, 3rd ed, Heidelberg: Physica-Verlag.
- Muller, W.G. (2005), A comparison of spatial design methods for correlated observations, *Environmetrics*, **16**, 495–506.
- Muller, W.G., and Zimmerman, D.L. (1999), Optimal design for variogram estimation, *Environmetrics*, **10**, 23–37.
- Nychka, D., and Saltzman, N. (1998), Design of air-quality monitoring designs, In *Case Studies in Environmental Statistics*, Eds. D. Nychka, W. Piegorisch, and L. Cox. New York: Springer, 51–76.
- Ritter, K. (1996), Asymptotic Optimality of Regular Sequence Design, *Ann. Statist*, **24**, 2081–2096.
- Rivaz, F. (2016), Optimal network design for Bayesian spatial prediction of multivariate non-Gaussian environmental data, *Journal of Applied Statistics*, **43**, 1335-1348.
- Royle, J.A., and Nychka, D. (1998), An algorithm for the construction of spatial coverage designs with implementation in Splus, *Comp and Geosci*, **24**, 479–488.
- Russo, D. (1984), Design of an optimal sampling network for estimating the variogram, *Soil Science Society of America Journal*, **48**, 708-716.
- Senarathne S. G. J., W.G. Muller, and J.M. McGree. (2020), Bayesian design for minimising uncertainty in spatial processes, arXiv:2001.08308v2 [stat.CO].
- Spock, G. (2012), The application of Bayesian statistics and convex design methodologies to geo-statistical prediction and sampling design, PhD Thesis, Alpen-Adria University Klagenfurt.
- Spock, G., and Pilz, J. (2008), Spatial sampling design and covariance-robust minimax prediction based on convex design ideas, *Stochastic Environmental Research and Risk Assessment*, **24**, 463-482.
- Stein, M. L. (1999), *Interpolation of Spatial Data: Some Theory for Kriging*, Springer-Verlag, New York.

- Su, Y., and Cambanis, S. (1993), Sampling designs for estimation of a random process, *Stoch Proc Appl*, **46**, 47–89.
- Wadoux, A.M.J-C., Brus, D.J., Rico-Ramirez, M.A., and Heuvelink, G. B.M. (2017), Sampling design optimisation for rainfall prediction using a non-stationary geostatistical model, *Advances in Water Resources*, **107**, 126–138.
- Warrick, A. W., and Myers, D. E. (1987), Optimization of sampling locations for variogram calculations, *Water Resources Research*, **23**, 496-500.
- Zhu, Z., and Stein, M.L. (2006), Spatial sampling design for prediction with estimated parameters, *Journal of Agricultural, Biological and Environmental Statistics*, **11**, 24–44.
- Zhu, Z., and Stein, M.L. (2005), Spatial Sampling Design for Parameter Estimation of the Covariance Function, *J. Statist. Plan and Inference*, **134**, 583–603.
- Zidek, J., Sun, W., and Le, N. (2000), Designing and integrating composite networks for monitoring multivariate Gaussian pollution fields, *Applied Statistics*, **49**, 63–79.
- Zimmerman, D.L. (2006), Optimal network design for spatial prediction, covariance parameter estimation, and empirical prediction, *Environmetrics*, **17**, 635–652.
- Zimmerman, D.L., and Cressie, N. (1992), Mean squared prediction error in the spatial linear model with estimated covariance parameters, *Annals of the Institute for Statistical Mathematics*, **44**, 27–43.