統計数理（2025）
第 73 巻 第 1 号 3–17

特集「空間統計モデリング：理論と応用」
［原著論文］

空間統計学の研究動向と今後の展望

村上 大輔†

（受付 2024 年 6 月 30 日；改訂 9 月 8 日；採択 9 月 9 日）

要　旨

（時）空間データを対象とした統計解析手法の研究が空間統計学で進められてきた．とりわけ空間データの大規模化や多様化の進む近年では，計算効率と柔軟性を両立するような解析手法が発展を遂げてきている．本研究の目的は，以上のような空間統計学の近年の研究動向を整理することである．そのために，まずは基礎的な空間統計モデルとその課題を紹介する．次に，同モデルを大規模データに応用するための近似手法を低ランク近似，共分散行列の近似，精度行列の近似に分けて，それぞれについて既往研究を整理する．その後，空間統計モデルの柔軟性を高めるための拡張手法を，共分散カーネルに基づく方法と，そうでない方法に分け，ニューラルネットワークの応用などの直近の動向も踏まえて整理する．その後，以上の手法の時空間データへの応用研究について紹介する．最後に，当該分野の手法を実装するためのソフトウェアについて解説したうえで，今後の空間統計学の課題について議論する．

キーワード：空間統計学，時空間モデリング，ガウス過程，空間相関，ニューラルネットワーク．

1. はじめに

地上・衛星からの観測技術や位置測位技術の発展に伴い，位置・時間情報を伴う時空間データが幅広く収集されている．時空間データの収集や公開は国などの公的機関だけでなく，研究グループなどの有志によっても活発に進められるようになってきている．オープンデータ化の進む今日，収集された空間データは無償公開されることも多く，例を挙げるならば全球の道路網や建物データを公開する OpenStreetMap（Vargas-Muñoz et al., 2020）や Microsoft Building Footprints（Huang et al., 2021），空間詳細な推計人口統計を公開する WorldPop（Stevens et al., 2015），幅広い衛星画像を直ちに分析可能な形で公開する GoogleEarthEngine（Gorelick et al., 2017）などがある．今後も利用可能な時空間データは増えていき，それらの統計解析はより一層重要となっていくと見込まれる．

時空間データを統計的にモデリングすることで予測，要因分析，不確実性評価などの様々な分析を行うことができる．特に，時空間データの得られる地点，集計単位，時点などは一般に限られているため，空間統計学（狭義には地球統計学）では時空間予測／補間に着目した研究が盛んに進められてきた（Cressie, 2015; Cressie and Wikle, 2015）．その中で，空間データの一般的性質とされる空間相関（近所と強く相関）や時系列相関（直前／直後と強く相関）を，その不確実性とともにモデル化する手法が確立されてきた．また近年では，時空間データの大規模化や多様化に対応するための手法の幅広い拡張が試みられてきた．しかしながら，空間統計学に関

† 統計数理研究所：〒 190–8562 東京都立川市緑町 10–3

するレビュー論文は，特に和文では，筆者の知る限り堤・瀬谷 (2012) のみであり，近年の研究動向について把握することが困難となっている．

そこで本研究では空間統計学の近年の研究動向を整理する．以降の章立ては次のとおりである．2 章では基礎的な空間統計モデルとその課題を紹介する．3，4 章では，同モデルの肝である空間過程のモデル化に関する研究動向を，計算効率化と柔軟化の 2 軸に分けてそれぞれレビューする．5 章では時空間過程に対するモデル化について整理する．6 章では空間統計モデルを様々なデータに応用するための手法についてレビューし，7 章では以上の手法を実装するためのソフトウェアについて整理する．8 章では残された課題を紹介する．

2.　空間統計モデル

2.1　基礎

地点・時点毎の観測データのベクトル $\boldsymbol{y} = [y(s_1, t_1), \ldots, y(s_N, t_N)]'$ を考える．$s_1, \ldots, s_N$ は対象地域 $D \subset \mathbb{R}^d$ (典型的には $d = 2$) 上の観測地点，$t_1, \ldots, t_N$ は時点を表す．具体例としては，観測所毎の気温や調査地点毎の地価などが挙げられる．空間統計学では時点数が 1 つ ($t_1 = t_2 = \cdots = t_N$) と仮定される場合も多く，本研究でも 2–4 章では時点は 1 つと仮定して明示的には扱わない．N はサンプルサイズであり 2–4 章では観測地点数に一致する (地点の重複がない場合)．複数時点を考慮する 5 章では，時点 t のサンプルサイズは N_t と，全時点についてのサンプルサイズ N は区別することとする．

基礎的な空間統計モデルは観測データが下式に従うと仮定する：

$$\boldsymbol{y} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I}), \tag{2.1}$$

$\boldsymbol{X}$ は説明変数行列，$\boldsymbol{\beta}$ は回帰係数ベクトル，$\boldsymbol{I}$ は単位行列である．空間統計学では行列 $\boldsymbol{K}_r$ の第 (i, j) 要素を直線距離 $d_{i,j}$ のカーネル $k_r(d_{i,j})$ で与えることで空間相関をモデル化する．例えば指数カーネル：$\exp(-\frac{d_{i,j}}{r})$，ガウスカーネル：$\exp(-(\frac{d_{i,j}}{r})^2)$，それらを包含する Matérn カーネル：$\frac{2^{1-\nu}}{\Gamma(\nu)}(\sqrt{2\nu}\frac{d_{i,j}}{r})K_\nu(\sqrt{2\nu}\frac{d_{i,j}}{r})$ ($\Gamma(\cdot)$ はガンマ関数，$K_\nu(\cdot)$ は次数 ν の第 2 種の修正ベッセル関数) は標準的に用いられている．r は空間相関の距離減衰の速さを決めるパラメータであり，その値が大きいことは大域的な空間相関パターンが誤差項にみられることを意味する．ν は空間過程のなめらかさを決めるパラメータであるが，r との識別の問題から固定値で与えられる場合も多く，本研究でも固定値とする．τ^2 と σ^2 は誤差項に対する分散パラメータであり，τ^2 が大きいことは空間相関パターンが支配的であることを，σ^2 が大きいことは空間相関では説明されないノイズパターンが支配的であることを，それぞれ意味する．なお，(2.1) 式は期待値と共分散の関数が移動不変という 2 次定常性の仮定に基づいている点に注意されたい．

多変量ガウス分布の性質から (2.1) 式の対数尤度は以下となる：

$$\log L = -\frac{1}{2}[\log(|\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I}|) - (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) - N\log(2\pi)], \tag{2.2}$$

M は説明変数の数である．パラメータ $\{\boldsymbol{\beta}, \tau^2, \sigma^2, r\}$ は尤度最大化により容易に推定できる．回帰係数の最尤推定量は $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'(\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'(\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}\boldsymbol{y}$ となり，$\{\tau^2, \sigma^2, r\}$ は $\boldsymbol{\beta}$ を $\hat{\boldsymbol{\beta}}$ に置き換えた (2.2) 式を最大化することで推定できる．

いま，未観測点 (s_0, t_0) のデータ (未知) を $y(s_0, t_0)$ とする．2 次定常性を仮定したため，(2.1) 式と同様，その期待値は $\boldsymbol{x}_0'\boldsymbol{\beta}$，各観測点との共分散は $\boldsymbol{k}_{r,0} = [k_r(d_{1,0}), \ldots, k_r(d_{N,0})]'$ で与えられることとなる．ただし $\boldsymbol{x}_0$ は同地点の説明変数ベクトル，$d_{i,0}$ は地点 i と地点 0 の間の直線距離である．以上の仮定を用いると $y(s_0, t_0)$ の最良線形不偏予測量は以下となる：

$$\hat{y}(s_0, t_0) = \boldsymbol{x}_0'\hat{\boldsymbol{\beta}} + \tau^2\boldsymbol{k}_{r,0}'(\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}). \tag{2.3}$$

空間統計学では同式を用いて(時)空間予測／補間が行わる．また，同予測量の不確実性を表す期待二乗誤差 $E[(y(s_0,t_0)-\hat{y}(s_0,t_0))^2]$ も解析的に得られる．任意地点における予測値をその不確実性とともに評価できる点は，空間統計モデルの利点の一つである．

なお，空間統計モデルは機械学習分野におけるガウス過程（Williams and Rasmussen, 2006）と同一である．従って，以降のレビューは空間統計分野におけるガウス過程の研究動向を整理したものとみなすこともできる．

2.2 課題

(2.1) - (2.3) 式からは空間統計モデルの課題がみえてくる．まず $(\tau^2\boldsymbol{K}_r+\sigma^2\boldsymbol{I})^{-1}=\boldsymbol{Q}$ の計算量は $O(N^3)$ のオーダーで指数的に増大するため大規模な空間データの解析には向かない．ビッグデータ時代である昨今においてこの課題は深刻である．そのため空間統計学では大規模データを扱うための計算コストの削減が近年のホットトピックとなっている．なお，$\boldsymbol{Q}$ は精度行列と呼ばれる．(2.2) - (2.3) 式に $\boldsymbol{Q}$ を代入すると $\boldsymbol{K}_r$ を含まない形で表せるため，その計算効率化のためには，$\boldsymbol{K}_r$ に対する近似以外に，$\boldsymbol{Q}$ を直接近似する方法も研究されている (3 章参照)．

また，ガウス分布を仮定する (2.1) 式では扱える問題が限られるという課題もある．この点に対処するために空間統計モデルは幅広く拡張されてきた．拡張されたモデルの多くは下式のように階層的に表せる（Cressie and Wikle, 2015 参照）：

$$\boldsymbol{y}|\boldsymbol{z}\sim f_{\boldsymbol{\theta}}(\boldsymbol{X},\boldsymbol{z}),\quad \boldsymbol{z}\sim g_{\boldsymbol{\theta}}(\boldsymbol{K}_r) \tag{2.4}$$

ここでパラメータは $\boldsymbol{\theta}$, $\boldsymbol{z}$ は (時) 空間過程である．$f_{\boldsymbol{\theta}}(\cdot)$ は観測データの分布，$g_{\boldsymbol{\theta}}(\cdot)$ は (時) 空間過程の分布を表す．例えば $f_{\boldsymbol{\theta}}(\boldsymbol{X},\boldsymbol{z})=N(\boldsymbol{X\beta}+\boldsymbol{z},\sigma^2\boldsymbol{I})$, $g(\boldsymbol{K})=N(0,\tau^2\boldsymbol{K}_r)$ とすると (2.1) 式となる．

一般に空間統計学では，(a) 大規模データを扱うために (時) 空間過程 $\boldsymbol{z}\sim g_{\boldsymbol{\theta}}(\boldsymbol{K}_r)$ が近似され，(b) 複雑な時空間過程を扱うために $g_{\boldsymbol{\theta}}(\boldsymbol{K}_r)$ が拡張され，(c) 幅広いデータを扱うために $f_{\boldsymbol{\theta}}(\boldsymbol{X},\boldsymbol{z})$ が拡張されてきた．以降では，(時) 空間過程に対するモデリング (a) - (b) については 3–5 章で，観測データに対するモデリング (c) について 6 章で，それぞれ研究動向を整理する．

3. 大規模データに対する空間過程の近似

空間統計モデルの計算効率を高めるための代表的な近似に，(i) 有限個の基底関数で空間過程を近似しようという低ランク近似，(ii) 密行列である $\boldsymbol{K}_r$ を疎な共分散行列に置き換える近似，(iii) $\boldsymbol{Q}$ を疎な精度行列に置き換える近似がある．本章では (i) - (iii) に関する手法を紹介する．

3.1 低ランク近似

下式のように L 個 ($L<N$) の基底関数の線形和で空間過程を近似する方法である：

$$\boldsymbol{z}=\boldsymbol{\Phi a},\quad \boldsymbol{a}\sim N(0,\boldsymbol{V_a}) \tag{3.1}$$

$\boldsymbol{\Phi}$ は L 個の基底関数を並べた行列 ($N\times L$) である．各基底関数を 2 階微分可能な位置座標の関数で与え，そのランダム係数ベクトル $\boldsymbol{a}$ を観測データから推定することで空間過程 $\boldsymbol{z}$ を近似する．基底関数には radial basis function（RBF; Cressie and Johannesson, 2008; Nychka et al., 2015），有限要素法に基づく基底（Lindgren et al., 2011），multi-resolution splines（Tzeng and Huang, 2018），random Fourier features（Miller and Reich, 2022）を含む所与の関数が用いられることが多い一方で，未知の関数とみなして期待二乗誤差の最小化に基づいて与えることもできる（例：Banerjee et al., 2008）．また気象分野では複数ケースで生成した大気拡散等のシミュ

レーションの結果を基底として用いることもある（Zammit-Mangion et al., 2022a 参照）.

低ランク近似の計算効率は複数の研究で確認されている（例：Heaton et al., 2019）．その一方で，過度な平滑化を招きやすい点や（Datta et al, 2016）ノイズ分散が小さい場合に精度が低下するなど（Stein, 2014），近似精度には課題が残されている．近似精度を高めるために，空間解像度毎の基底関数を用いて空間過程を近似する Nychka et al.（2015）や Katzfuss and Hammerling（2017）が提案されたほか，後述の疎な共分散関数と組み合わせることで表現能力を高めようという full scale approximation（Sang and Huang, 2012），multiresolution approximation（Katzfuss, 2017），modified linear projection（Hirano, 2017）なども提案されてきた．

低ランク近似は，基底を投入するだけで空間相関が考慮できるという簡便さなどの理由から，時空間モデリングや機械学習などでも広く用いられている（4–5 章参照）．

3.2 疎な共分散行列を用いる近似

$\boldsymbol{K}_r$ を疎な行列で置き換えようという近似である．$\boldsymbol{K}_r$ が疎のとき $\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I}$ も疎になるため，コレスキー分解を用いた効率的な逆行列計算が使用可能となる．

具体的な方法は，共分散行列の半正定値性を満たす形で提案されてきた．対象地域をボロノイ領域やグリッド毎に分割する方法はその一つである．分割により得られた空間領域間に独立性を仮定する近似（Stein et al., 2004; Hazra et al., 2024）や，隣接以外の領域との条件付き独立性を仮定する近似（Eidsvik et al., 2014）がある．いずれも共分散行列 $\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I}$ はブロック対角行列またはそれに類する疎行列となるため，その逆行列は効率よく評価できる．また，空間領域毎にモデルを推定し，それらの予測分布を合成することで全域の予測分布を得ようという nested kriging（Rullière et al., 2018）や，各領域から得られた予測分布の加重和を最適化しようという cluster kriging（Van Stein et al., 2020）など，関連したアンサンブル学習手法もある．

別の方法に，共分散カーネルに taper 関数と呼ばれる距離減衰の速い関数をかけることで共分散を疎にする covariance tapering がある（Furrer et al., 2006）．同手法には共分散カーネルにのみ taper 関数をかける one-taper 法と，データの共分散行列にも taper 関数を書ける two-taper 法があり，前者は，特に taper 関数の減衰が早すぎる場合にパラメータにバイアスがかかることが知られている（Furrer et al., 2016）．一方で，後者は不偏となるものの計算コストが増大するため一長一短である．関連手法に，ある地点の空間構造をモデル化するために，近隣の観測地点を動的に選択していくことで疎な共分散構造を実現しようという local approximate Gaussian process（laGP; Gramacy, 2016）がある．同手法は全域に対する半正定値な共分散構造を与えるものではないという批判はあるが，地点毎にパラメータが推定でき，非定常性を捉える実用手法としては便利である．

3.3 疎な精度行列を用いる近似

精度行列 $\boldsymbol{Q} = (\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}$ を直接疎にする方法である．（i）Stochastic partial differential equation（SPDE）を用いる方法と，（ii）Vecchia 近似を用いる方法が注目を浴びている．

（i）に関し，Lindgren et al.（2011）は，SPDE がグラフ上の確率場である Gauss-Markov random fields（GMRF）で近似できること，ならびに特定の SPDE の解がガウス過程（Matern カーネル）に一致することを利用して，GMRF とガウス過程が一対一対応することを示した．SPDE 法（Krainski et al., 2018 参照）は，この対応関係を利用して空間グラフ上の GMRF の線形和で空間過程を近似する方法であり，ガウス過程の代わりに GMRF を推定すればよい．GMRF の精度行列 $\boldsymbol{Q}_{GMRF}$ は空間グラフの隣接構造から自動的に決まるため，明示的な逆行列計算は不要である．結果として，GMRF で近似されるガウス過程の精度行列もまた自動的に得られ，計算コストは大幅に削減される．

この方法は単に計算効率が良くなるだけでなく，SPDE の特性を活かすことで様々な非定常・非ガウス空間過程への拡張が可能である（Bakka et al., 2018; Lindgren et al., 2022; 本論文 4 章参照）．

（ii）では観測点 $\boldsymbol{s}_{1:N} = [s_1, \ldots, s_N]$ を順位づけたうえで空間過程 $\boldsymbol{z} \sim N(0, \boldsymbol{K}_r)$ の分布が $p(\boldsymbol{z}) = p(z(s_1)) \prod_{i=2}^{N} p(z(s_i)|\boldsymbol{s}_{1:(i-1)})$ のように表せることを利用する．$p(z(s_i)|\boldsymbol{s}_{1:(i-1)})$ は地点群 $\boldsymbol{s}_{1:(i-1)} = [s_1, \ldots, s_{i-1}]$ における空間過程 $z(s_1), \ldots, z(s_{i-1})$ が与えられた下での空間過程 $z(s_i)$ の条件付き分布を表す．General Vecchia approximation（Katzfuss and Guinness, 2021）では，上位の地点群 $\boldsymbol{s}_{1:(i-1)}$ をより少数のサブサンプル $\boldsymbol{s}_{g(i)}$ で置き換えた下式で空間過程を近似する：

$$p(\tilde{\boldsymbol{z}}) = p(z(s_1)) \prod_{i=2}^{N} p(z(s_i)|\boldsymbol{s}_{g(i)}) \tag{3.2}$$

$\tilde{\boldsymbol{z}} \sim N(0, \tilde{\boldsymbol{K}}_r)$ の精度行列をコレスキー分解すると $\tilde{\boldsymbol{K}}_r^{-1} = \boldsymbol{U}\boldsymbol{U}'$ となり，$\boldsymbol{U}$ は地点 s_i と地点 $s_j \in \boldsymbol{s}_{g(i)}$ のペアについての要素のみが非ゼロの疎な上三角ブロック行列となる．この特性を活かすと，尤度と予測量が線形時間で評価できる．

既存の高速近似手法には（3.2）式の応用と見做せるものも多い（Katzfuss and Guinness, 2021）．例えば nearest neighbor Gaussian process（Datta et al., 2016）は $\boldsymbol{s}_{g(i)}$ を近隣地点のみで与え，modified predictive process（Finley et al., 2009）は $\boldsymbol{s}_{g(i)}$ を対象地域に均等に配置する地点群で与えるほか，3.1 節で紹介した full scale approximation（Sang and Huang, 2012）や multiresolution approximation（Katzfuss, 2017）も同近似の一種とみなせる．（3.2）式を用いた近似の精度は観測点の順位とサブサンプルの選び方 $\boldsymbol{s}_{g(i)}$ に依存するが，それらの決定方法については Guinness（2018; 2021）を参照されたい．Katzfuss and Guinness（2021）は精度行列のスパース性を保持しながら Kullback-Leibler ダイバージェンスの意味で精度を改良させた Sparse General Vecchia（SGV）を提案した．その精度の良さは比較研究でも確認している（Hazra et al., 2024）．

4. 複雑な空間過程を表現するための拡張

複雑な空間過程を表現するために共分散カーネルの拡張が行われてきた．それに加え，近年では共分散カーネルの代わりに neural network（NN）を用いて柔軟性を高めようという手法の研究も盛んである．そこで本章ではそれぞれについて簡単にレビューする．

4.1 共分散に基づく方法

基礎的な空間統計モデルの共分散カーネルは，どの方向も空間相関の及ぶ範囲は同じという等方性や，どの場所の空間相関も同じカーネルで記述されるという定常性の仮定に基づいているが，現実にはそれらが満たされない場合もある．例えば CO2 濃度の場合，地形や風向きの影響で空間相関の及ぶ範囲が方向毎に異なるかもしれないし，陸上と海上では空間相関パターンが異なるかもしれない．以上のような等方性や定常性を満たさない空間過程を扱うために共分散カーネルが拡張されてきた．

例えば，方向毎に空間相関パターンが異なるという異方性（anisotropy）が考慮されてきた．例えば Tsutsumi and Seya（2009）はつくばエクスプレス開業に伴う沿線の地価上昇に，同路線に沿った異方性がみられることを示した．さらに，Nychka et al.（2018）と Wiens et al.（2020）は局所尤度を用いて場所毎の局所異方性を計算効率良く推論する手法を提案した．同手法を用いることで，海岸線や山脈などによる空間相関パターンの歪みが表現できる．彼らは同手法を気候モデルのエミュレーションに応用した．

一方，SPDE 法(3 章参照)は非定常な空間過程のモデリングに広く用いられている（Krainski et al., 2018; Lindgren et al., 2022)．同手法は空間過程と SPDE を関連付けるものであり，SPDE のパラメータを場所毎に変えたり（Lindgren et al., 2011; Fuglstad et al., 2015)説明変数の関数で与えたりすることで（Ingebrigtsen et al., 2015)，非定常な空間過程が表現できる．また，SPDE を多様体上で定義することで，例えば球面上の空間過程や，川や国境といった障害物を考慮した空間過程などが表現できるため（Barrier model; Bakka et al., 2019)，全球を対象とした気象の分析や入り江などを対象とした生態の分析などで用いられてきた（例：Lezama-Ochoa et al., 2020; Fioravanti et al., 2023)．また，SPDE 法で扱うモデルには空間的な滑らかさを表すパラメータ ν があり，その値は Matern カーネルに基づく空間過程を近似する整数値で与えられることが多い．一方，ν に整数値以外の実数値も許容することで，より柔軟に空間過程をモデル化しようという fractional SPDE 法も提案されている（Bolin et al., 2024)．

4.2 共分散に基づかない方法

前節では共分散カーネルに基づく非定常モデルを紹介したが，現実には共分散では捉えきれない複雑なパターンがみられる場合もある．そのような場合に対処するために，近年では NN の応用が空間統計分野で活発化してきている．特に deep NN(DNN)は N の増加に対して高いスケーラビリティを有することや，空間過程が非連続となる特異性を有する場合に他手法に優ることが知られており（Imaizumi and Fukumizu, 2022)，大規模で複雑な現象の解析に役立つ．

位置情報に基づく説明変数(特徴量)の投入は，NN で空間パターンを学習するための代表的な方法の一つである．これまでに用いられてきた説明変数には位置座標（Cracknell and Reading, 2014)，経験バリオグラム（Gerber and Nychka, 2021)，kriging 予測量（Wang et al., 2019)，空間基底（Yoshida et al., 2022; Zammit-Mangion et al., 2022a; Chen et al., 2024)などがある．一方，この方法では空間パターンの非線形が間接的にしか学習されず，その柔軟性には課題もある（Zhan and Datta, 2024 参照)．

より柔軟性を高めるために，NN 内部の変換関数への空間情報の導入も試みられてきた．例えば，Zammit-Mangion et al.（2022b; 2024a)は変換関数のパラメータを空間可変とすることで，非定常性や異方性を有する複雑なパターンを捉える spatial Bayesian NN を提案した．また，Katzfuss and Schäfer（2023)は，空間相関情報を織り込んだ変換関数を用いたベイズ最適輸送により，複雑な空間パターンを計算効率よく学習する生成モデルを提案した．いずれも急激な変化を伴うような複雑なパターンが表現できる．

グラフを考慮する NN である graph NN（Wu et al., 2020 参照)の応用も盛んであり，隣接グラフで繋がれたゾーン/地点間の空間相関パターンが学習されてきた（例：Tonks et al., 2024; Cisneros et al., 2024)．また，Zhan and Datta（2024)は，疎な精度行列を用いる空間過程の近似手法である nearest neighbor Gaussian process と graph NN の理論的関係を明らかとしたうえで，両者を統合した空間過程モデル NN-GLS を提案した．

モデルの柔軟性を高める以外の NN の利点に，明記的に尤度を評価することなく尤度が近似できる点がある(likelihood-free)．この性質により，NN は，既存手法では計算コスト等の観点で解析困難であった，複雑な分布の解析を容易にした．例えば，max stable process は極端現象(例：異常気象)を扱う上で用いられるが，N の増加に伴う計算量の増大が速く大規模データには応用できなかった．そこで Sainsbury-Dale et al.（2024)は同確率過程を NN で近似することでこの課題に対処した．

以上に加え，時空間データに対する NN の応用も盛んである．この点については次章で紹介する．

5. 時空間過程に対するモデルの発展

時空間過程のモデリング方法は，(i)時空間上のカーネルを用いて全時点を同時にモデリングしようという静的な方法と，(ii)過去の状態が与えられた下で現在の状態を推論しようという動的な方法がある．

5.1 静的な方法

空間統計モデル（(2.1)式）の $\boldsymbol{K}_r$ を，時差と距離に依存して減衰する時空間カーネルに基づいて与えることで時空間過程がモデリングする方法である（Cressie and Wikle, 2015）．Porcu et al. (2021)でレビューされたように，これまでに数多くの時空間カーネルが，半正定値性を満たすという条件の下で提案されてきた．それらは時間方向のカーネルと空間方向のカーネルに線形分離可能な separable カーネルとそうでない non-separable カーネルに分類できる．separable カーネルは時間・空間カーネルの和または積をとることで得られる．それらを用いることで計算コストは抑えられるが（Porcu et al., 2021），時間と空間の相互作用は無視されるため柔軟性には難がある．そのため，product-sum カーネル（De Cesare et al., 2001），Gneiting カーネル（Gneiting, 2002），scale mixture に基づくカーネル（Porcu and Zastavnyi, 2011)を含む，non-separable カーネルが数多く提案されてきた．

静的な方法は，全時点についての大きなカーネル行列 $\boldsymbol{K}_r$ を扱うため，逆行列計算 $(\tau^2\boldsymbol{K}_r + \sigma^2\boldsymbol{I})^{-1}$ の計算コストが増大する．そのため separable カーネルを用いるか，3 章で説明した近似手法を応用することで，計算コストが軽減されてきた（例えば Wood et al., 2017; Wikle et al., 2019）．静的な方法について，詳しくは Porcu et al. (2021)を参照されたい．

5.2 動的な方法

過去からの動的な変化を明示的に記述する方法である．例えば，離散時間 $t \in \{1, \ldots, T\}$ 毎に同一の複数地点でデータが観測されている場合，時点 t のデータベクトルは状態空間モデルを用いて次のように表現できる：$\boldsymbol{y}_t = \boldsymbol{z}_t + \boldsymbol{e}_t, \boldsymbol{e}_t \sim N(0, \sigma^2\boldsymbol{I}), \boldsymbol{z}_t = \rho\boldsymbol{W}_t\boldsymbol{z}_{t-1} + \boldsymbol{u}_t, \boldsymbol{u}_t \sim N(0, \tau^2\boldsymbol{K}_t)$. $\boldsymbol{z}_t$ は動的に変化する時空間過程を表し，パラメータ ρ が 0 の場合，$\boldsymbol{z}_t$ は時点毎に独立な空間過程となり，ρ が大きくなるにつれて $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_T$ の時系列相関が強まる．また $\boldsymbol{W}_t$ は遷移行列であり，同行列を通して時空間過程の移流や拡散などの流れを柔軟に表現することができる（Integro-difference equation (IDE)モデル; Wikle et al., 2019 参照）．一般に，動的なモデルでは時点毎の共分散行列 $\boldsymbol{K}_t$ が逐次的に処理されるため，全時点についてのカーネルを用いる静的な方法に比べて計算効率が良い．また，現象の動きが明示的に表現できる点や，急激な変化をとらえやすい点，将来予測に応用しやすい点なども動的な方法の利点である（Cressie and Wikle, 2015 参照）．

一方で，時点毎の観測データが多い場合は動的な方法でも計算量が増大するため，データベクトルを $\boldsymbol{y}_t = \boldsymbol{\Phi}_t\boldsymbol{z}_t + \boldsymbol{e}_t$ のように与えて，低ランク近似した時空間過程 $\boldsymbol{\Phi}_t\boldsymbol{z}_t$ を用いる場合も多い（$\boldsymbol{z}_t(L \times 1)$：潜在変数ベクトル，$\boldsymbol{\Phi}_t(N_t \times L)$：基底関数行列．ただし N_t は時点 t のサンプルサイズ）．fixed rank filter（Cressie et al., 2010)や multiresolution filter（Jurek and Katzfuss, 2021)はその一例である．低ランク近似は自然に時空間モデルに組み込むことができ，計算効率が良く拡張もしやすいため，動的な時空間モデリングでも幅広く用いられている（Cressie et al., 2022 参照）．

状態空間モデルは，観測データ $\boldsymbol{y}_t$ と潜在変数 $\boldsymbol{z}_t$ に対する 2 層のモデルだが，これを多層化する試みもある．例えば Katzfuss et al. (2019)は変換，観測，時間発展，パラメータを表す 4 層の階層モデルを提案した．Wikle (2019)も関連する階層モデルを提案して DNN 等との

比較を行っている．より近年は，4.2 章でも述べた通り DNN などの機械学習手法の応用が盛んとなっている．特に時空間データに関しては，複雑な時間変化が学習可能な Recurrent NN（Yu et al., 2019 参照）や echo state network（Jaeger, 2001）などが，非定常・非ガウスの時空間過程をモデリングするために応用されてきた（Nag et al., 2023; McDermott and Wikle, 2017; 2019）．（D）NN は，一度学習すれば，以降は新規データを学習済みモデルに入力するだけで，極めて小さな計算コストで解析結果が得られるという計算上の利点がある（Amortized inference; Zammit-Mangion et al., 2024b）．高頻度に観測される時空間データが分野を問わず増えてきていることを踏まえると，（D）NN は時空間データに対してもより一層重要となると考えられる．空間統計分野における（D）NN の応用・拡張に関してより詳しくは Wikle and Zammit-Mangion (2023) を参照されたい．

6. 観測データに対するモデルの発展

5 章までは空間・時空間過程をモデル化する方法についてレビューした．本章では，観測データを柔軟にモデル化するための代表的な方法である generalized linear model（GLM）を応用する方法と，変換を行う方法について紹介する．

GLM とは指数分布族であるポアソン分布や二項分布などに従うデータを扱う回帰である．GLM は空間統計分野で古くから応用されており（Gotway and Stroup, 1997），今日では GLM の潜在変数として（時）空間過程を導入した階層ベイズモデルが Stan や INLA といった R パッケージに実装されている．例えば INLA パッケージでは 2024 年 6 月現在で約 100 もの確率分布がデータに仮定できるなど（Bakka et al., 2018），幅広いデータの空間統計モデリングが誰でも実装できるようになっている．以上のように GLM に基づく空間統計モデルはすでに確立されている一方で，計算効率を改善するための推定法は今なお改良されており，例えば上述の INLA パッケージで用いられている近似ベイズ推論法 integrated nested Laplace approximation（INLA 法）に関しては，効果識別のために導入していたノイズ項を除去したより精緻で計算効率の良い定式化（Van Niekerk et al., 2023），ラプラス近似部分を変分ベイズ推論で置き換えることによる近似精度改善（Van Niekerk and Rue, 2024）などが試みられてきた．INLA 法以外では，自動微分（automatic differentiation）を応用したモデル推定の高速化も試みられており，（時）空間モデリングに対する精度と計算効率が INLA 法を含む既存手法を上回ることなどを確認している（Anderson et al., 2022）．なお，以上の研究の多くは，空間相関だけでなく時系列相関，非線形効果，グループ別の効果なども扱うことのできるセミパラメトリックモデルの枠組みを扱っており，単に精度や計算効率が良いだけでなく，汎用性も高い統計手法が確立されてきている．

一方，変換を行う方法とは，観測データを対数変換や Box-Cox 変換などで変換することで非ガウスデータを扱おうという方法である．近年は NN による変換が盛んに試みられている．5 章でも述べた通り NN は likelihood-free であり，様々な分布に従うデータが近似できるため，これまで扱うことが困難だったような複雑な分布に従うデータも分析可能である．この特性を活かしたベイズ推論である neural Bayes 推定（Sainsbury-Dale et al., 2024）は今後幅広く利用される可能性がある．また，NN をはじめとした機械学習手法を用いれば複数ドメイン（例：地域）のデータが同時に学習でき，例えば転移学習やメタ学習といったこれまで空間統計分野であまり扱われてこなかったような問題設定にも応用できる可能性がある．

7. ソフトウェア

最後に以上で紹介した手法を実装するためのソフトウェア・パッケージを R を中心に紹介す

る．まず，3 章で紹介した大規模データモデリング手法に関しては，Wikle et al.（2019）でも一部解説されている通り，低ランク近似法は FRK，LatticeKrig，mgcv パッケージなどで，疎な共分散を用いる方法は spam，laGP パッケージなどで実装できる．また，SPDE を用いて精度行列を疎にする方法は INLA パッケージで，Vecchia 近似を用いる方法は GpGp や GPvecchia で実装できるほか，spNNGP，spBayes などでも本文中で紹介した関連手法が実装できる．各空間モデルの精度や計算時間の比較については Heaton et al.（2019）や Hazra et al.（2024）を参照されたい．

4 章で紹介した，共分散カーネルに基づいて柔軟な時空間過程を扱う方法に関しては，大域的な異方性は gstat で，局所異方性は LatticeKrig でそれぞれ実装できる．また，SPDE を用いて非定常パターンを捉える方法に関しては INLA パッケージで実装できる（Krainski et al., 2018 参照）．NN を用いる方法に関しては deepspat（https://github.com/andrewzm/deepspat）パッケージで実装できるほか，Python であれば Tensorflow，PyTorch，Pyro ライブラリなどでも幅広い NN の実装が可能である．また，GPyTorch では，PyTorch の自動微分や GPU（Graphics Processing Unit）による並列計算を活かすことで，柔軟・高速なガウス過程の実装を実現している．

5 章で紹介した時空間モデリングに関しては，静的な方法は gstat，FRK，mgcv パッケージなどで実装でき，動的な方法は INLA，KFAS などで実装できるもののパッケージがやや少ない印象である．

6 章で紹介した GLM に基づく方法に関しては INLA，sdmTMB，mgcv パッケージなどで実装できる．それらの比較については Anderson et al.（2022）を参照されたい．変換／ NN に基づく方法に関してはすでに述べたとおりである

8. 今後の展望

以上で空間統計分野の最近の動向について整理した．無論，本研究で紹介できたのは当該分野の一部のトピックのみであり，例えば空間領域別の離散データのモデリング（例：Lee, 2013），多変量モデリング（例：Zhang et al., 2021），空間交絡（例：Hughes and Haran, 2013）など，活発に議論されているトピックは他にも複数存在する点には注意されたい．

空間統計学はまだまだ課題が残されている．例えば NN の応用は活発化してきているが，NN と空間統計モデルの融合モデルの統計的性質はほぼ明らかとされておらず（Zhan and Datta, 2024），今後，両分野をまたぐ統計理論の確立が必要である．

また，地球物理学を中心に発展を遂げてきたデータ同化手法の空間統計学での応用が，基礎的なカルマンフィルタ以外は進んでいないという課題もある．実際，代表的なデータ同化手法の一つである ensemble Kalman filter（Evensen, 2003）は複雑な時空間パターンを捉えるうえで有用であるものの Katzfuss et al.（2016）によれば統計コミュニティでは largely unknown な手法であり適用例が少なかった．数多くのデータ同化手法が提案されてきているとともに（Carrassi et al., 2018 参照），近年は S4（Structured State Spaces for Sequence Modeling; Gu et al., 2022）や Mamba（Gu and Dao, 2023）といったより柔軟な時空間モデルが機械学習分野でも登場しており，空間統計モデルとの融合の余地がある．

また，大量で多様なデータで学習され，幅広いタスク（例えば気温の将来予測，風速の空間内挿，土地被覆の分類）に適応可能な大規模モデルである基盤モデル（Bommasani et al., 2021）が，空間統計分野の主要応用分野である気象・環境分野や都市分野で活用され始めている（Nguyen et al., 2023; Zhang et al., 2024 参照）．以上のような機械学習分野の研究動向も踏まえながら，より大規模で複雑なモデリングも扱えるように，空間統計学の手法の高度化や実用化を進めて

いくことが今後期待される．

参 考 文 献

Anderson, S. C., Ward, E. J., English, P. A. and Barnett, L. A. (2022). sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields, *BioRxiv*, https://doi.org/10.1101/2022.03.24.485545.

Bakka, H., Rue, H., Fuglstad, G. A., Riebler, A., Bolin, D., Illian, J., Krainski, E., Simpson, D. and Lindgren, F. (2018). Spatial modeling with R-INLA: A review, *Wiley Interdisciplinary Reviews*: *Computational Statistics*, **10**(6), e1443.

Bakka, H., Vanhatalo, J., Illian, J. B., Simpson, D. and Rue, H. (2019). Non-stationary Gaussian models with physical barriers, *Spatial Statistics*, **29**, 268–288.

Banerjee, S., Gelfand, A. E., Finley, A. O. and Sang, H. (2008). Gaussian predictive process models for large spatial data sets, *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **70**(4), 825–848.

Bolin, D., Simas, A. B. and Xiong, Z. (2024). Covariance–based rational approximations of fractional SPDEs for computationally efficient Bayesian inference, *Journal of Computational and Graphical Statistics*, **33**(1), 64–74.

Bommasani, R., Hudson, D. A., Adeli, E., Altman, R., Arora, S., von Arx, S., Bernstein, M. S., Bohg, J., Bosselut, A., Brunskill, E., et al. (2021). On the opportunities and risks of foundation models, ArXiv, https://doi.org/10.48550/arXiv.2108.07258.

Carrassi, A., Bocquet, M., Bertino, L. and Evensen, G. (2018). Data assimilation in the geosciences: An overview of methods, issues, and perspectives, *Wiley Interdisciplinary Reviews*: *Climate Change*, **9**(5), e535.

Chen, W., Li, Y., Reich, B. J. and Sun, Y. (2024). Deepkriging: Spatially dependent deep neural networks for spatial prediction, *Statistica Sinica*, **34**, 291–311.

Cisneros, D., Richards, J., Dahal, A., Lombardo, L. and Huser, R. (2024). Deep graphical regression for jointly moderate and extreme Australian wildfires, *Spatial Statistics*, https://doi.org/10.1016/j.spasta.2024.100811.

Cracknell, M. J. and Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information, *Computers and Geosciences*, **63**, 22–33.

Cressie, N. (2015), *Statistics for Spatial Data*, John Wiley, New York.

Cressie, N. and Johannesson, G. (2008). Fixed rank kriging for very large spatial data sets, *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **70**(1), 209–226.

Cressie, N. and Wikle, C. K. (2015), *Statistics for Spatio-Temporal Data*, Wiley, New Jersey.

Cressie, N., Shi, T. and Kang, E. L. (2010). Fixed rank filtering for spatio-temporal data, *Journal of Computational and Graphical Statistics*, **19**(3), 724–745.

Cressie, N., Sainsbury-Dale, M. and Zammit-Mangion, A. (2022). Basis-function models in spatial statistics, *Annual Review of Statistics and Its Application*, **9**(1), 373–400.

Datta, A., Banerjee, S., Finley, A. O. and Gelfand, A. E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets, *Journal of the American Statistical Association*, **111**(514), 800–812.

De Cesare, L., Myers, D. E. and Posa, D. (2001). Product-sum covariance for space-time modeling: An environmental application, *Environmetrics*: *The Official Journal of the International En-*

vironmetrics Society, **12**(1), 11–23.

Eidsvik, J., Shaby, B. A., Reich, B. J., Wheeler, M. and Niemi, J. (2014). Estimation and prediction in spatial models with block composite likelihoods, *Journal of Computational and Graphical Statistics*, **23**(2), 295–315.

Evensen, G. (2003). The ensemble Kalman filter: Theoretical formulation and practical implementation, *Ocean Dynamics*, **53**, 343–367.

Finley, A. O., Sang, H., Banerjee, S. and Gelfand, A. E. (2009). Improving the performance of predictive process modeling for large datasets, *Computational Statistics and Data Analysis*, **53**(8), 2873–2884.

Fioravanti, G., Martino, S., Cameletti, M. and Toreti, A. (2023). Interpolating climate variables by using INLA and the SPDE approach, *International Journal of Climatology*, **43**(14), 6866–6886.

Fuglstad, G. A., Simpson, D., Lindgren, F. and Rue, H. (2015). Does non-stationary spatial data always require non-stationary random fields?, *Spatial Statistics*, **14**, 505–531.

Furrer, R., Genton, M. G. and Nychka, D. (2006). Covariance tapering for interpolation of large spatial datasets, *Journal of Computational and Graphical Statistics*, **15**(3), 502–523.

Furrer, R., Bachoc, F. and Du, J. (2016). Asymptotic properties of multivariate tapering for estimation and prediction, *Journal of Multivariate Analysis*, **149**, 177–191.

Gerber, F. and Nychka, D. (2021). Fast covariance parameter estimation of spatial Gaussian process models using neural networks, *Stat*, **10**(1), e382.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data, *Journal of the American Statistical Association*, **97**(458), 590–600.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone, *Remote Sensing of Environment*, **202**, 18–27.

Gotway, C. A. and Stroup, W. W. (1997). A generalized linear model approach to spatial data analysis and prediction, *Journal of Agricultural, Biological, and Environmental Statistics*, **2**(2), 157–178.

Gramacy, R. B. (2016). laGP: large-scale spatial modeling via local approximate Gaussian processes in R, *Journal of Statistical Software*, **72**, https:/doi.org/10.18637/jss.v072.i01.

Gu, A. and Dao, T. (2023). Mamba: Linear-time sequence modeling with selective state spaces, ArXiv, https://doi.org/10.48550/arXiv.2312.00752.

Gu, A., Goel, K., Gupta, A. and Ré, C. (2022). On the parameterization and initialization of diagonal state space models, *Advances in Neural Information Processing Systems*, **35**, 35971–35983.

Guinness, J. (2018). Permutation and grouping methods for sharpening Gaussian process approximations, *Technometrics*, **60**(4), 415–429.

Guinness, J. (2021). Gaussian process learning via Fisher scoring of Vecchia's approximation, *Statistics and Computing*, **31**(3), 25.

Hazra, A., Nag, P., Yadav, R. and Sun, Y. (2024). Exploring the efficacy of statistical and deep learning methods for large spatial datasets: A case study, *Journal of Agricultural, Biological and Environmental Statistics*, https://doi.org/10.1007/s13253-024-00602-4.

Heaton, M. J., Datta, A., Finley, A. O., Furrer, R., Guinness, J., Guhaniyogi, R., Gerber, F., Gramacy, R. B., Hammerling, D., Katzfuss, M., Lindgren, F., Nychka, D. W., Sun, F. and Zammit-Mangion, A. (2019). A case study competition among methods for analyzing large spatial data, *Journal of Agricultural, Biological and Environmental Statistics*, **24**, 398–425.

Hirano, T. (2017). Modified linear projection for large spatial datasets, *Communications in Statistics-Simulation and Computation*, **46**(2), 870–889.

Huang, X., Wang, C., Li, Z. and Ning, H. (2021). A 100 m population grid in the CONUS by disaggregating census data with open-source Microsoft building footprints, *Big Earth Data*, **5**(1), 112–133.

Hughes, J. and Haran, M. (2013). Dimension reduction and alleviation of confounding for spatial generalized linear mixed models, *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **75**(1), 139–159.

Imaizumi, M. and Fukumizu, K. (2022). Advantage of deep neural networks for estimating functions with singularity on hypersurfaces, *Journal of Machine Learning Research*, **23**(111), 1–54.

Ingebrigtsen, R., Lindgren, F., Steinsland, I. and Martino, S. (2015). Estimation of a non-stationary model for annual precipitation in southern Norway using replicates of the spatial field, *Spatial Statistics*, **14**, 338–364.

Jaeger, H. (2001). The "echo state" approach to analysing and training recurrent neural networks—with an erratum note, GMD Technical Report, 148, German National Research Center for Information Technology, Bonn, Germany.

Jurek, M. and Katzfuss, M. (2021). Multi-resolution filters for massive spatio-temporal data, *Journal of Computational and Graphical Statistics*, **30**(4), 1095–1110.

Katzfuss, M. (2017). A multi-resolution approximation for massive spatial datasets, *Journal of the American Statistical Association*, **112**(517), 201–214.

Katzfuss, M. and Guinness, J. (2021). A general framework for Vecchia approximations of Gaussian processes, *Statistical Science*, **36**(1), 124–141.

Katzfuss, M. and Hammerling, D. (2017). Parallel inference for massive distributed spatial data using low-rank models, *Statistics and Computing*, **27**, 363–375.

Katzfuss, M. and Schäfer, F. (2023). Scalable Bayesian transport maps for high-dimensional non-Gaussian spatial fields, *Journal of the American Statistical Association*, **119**, 1409–1423, https://doi.org/10.1080/01621459.2023.2197158.

Katzfuss, M., Stroud, J. R. and Wikle, C. K. (2016). Understanding the ensemble Kalman filter, *The American Statistician*, **70**(4), 350–357.

Katzfuss, M., Stroud, J. R. and Wikle, C. K. (2019). Ensemble Kalman methods for high-dimensional hierarchical dynamic space-time models, *Journal of the American Statistical Association*, **115**(530), 866–885, https://doi.org/10.1080/01621459.2019.1592753.

Krainski, E., Gómez-Rubio, V., Bakka, H., Lenzi, A., Castro-Camilo, D., Simpson, D., Lindgren, F. and Rue, H. (2018), *Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*, Chapman and Hall/CRC, Boca Raton.

Lee, D. (2013). CARBayes: an R package for Bayesian spatial modeling with conditional autoregressive priors, *Journal of Statistical Software*, **55**(13), https://doi.org/10.18637/jss.v055.i13.

Lezama-Ochoa, N., Pennino, M. G., Hall, M. A., Lopez, J. and Murua, H. (2020). Using a Bayesian modelling approach (INLA-SPDE) to predict the occurrence of the Spinetail Devil Ray (Mobular mobular), *Scientific Reports*, **10**(1), 18822.

Lindgren, F., Rue, H. and 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**(4), 423–498.

Lindgren, F., Bolin, D. and Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running, *Spatial Statistics*, **50**, https://doi.org/10.1016/j.spasta.2022.100599.

McDermott, P. L. and Wikle, C. K. (2017). An ensemble quadratic echo state network for non-linear spatio-temporal forecasting, *Stat*, **6**(1), 315–330.

McDermott, P. L. and Wikle, C. K. (2019). Deep echo state networks with uncertainty quantification for spatio-temporal forecasting, *Environmetrics*, **30**(3), e2553.

Miller, M. J. and Reich, B. J. (2022). Bayesian spatial modeling using random Fourier frequencies, *Spatial Statistics*, **48**, https://doi.org/10.1016/j.spasta.2022.100598.

Nag, P., Sun, Y. and Reich, B. J. (2023). Spatio-temporal DeepKriging for interpolation and probabilistic forecasting, *Spatial Statistics*, **57**, https://doi.org/10.1016/j.spasta.2023.100773.

Nguyen, T., Brandstetter, J., Kapoor, A., Gupta, J. K. and Grover, A. (2023). ClimaX: A foundation model for weather and climate, ArXiv, https://doi.org/10.48550/arXiv.2301.10343.

Nychka, D., Bandyopadhyay, S., Hammerling, D., Lindgren, F. and Sain, S. (2015). A multiresolution Gaussian process model for the analysis of large spatial datasets, *Journal of Computational and Graphical Statistics*, **24**(2), 579–599.

Nychka, D., Hammerling, D., Krock, M. and Wiens, A. (2018). Modeling and emulation of nonstationary Gaussian fields, *Spatial Statistics*, **28**, 21–38.

Porcu, E. and Zastavnyi, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields, *Journal of Multivariate Analysis*, **102**(9), 1293–1301.

Porcu, E., Furrer, R. and Nychka, D. (2021). 30 Years of space–time covariance functions, *Wiley Interdisciplinary Reviews*: *Computational Statistics*, **13**(2), e1512.

Rullière, D., Durrande, N., Bachoc, F. and Chevalier, C. (2018). Nested Kriging predictions for datasets with a large number of observations, *Statistics and Computing*, **28**, 849–867.

Sainsbury-Dale, M., Zammit-Mangion, A. and Huser, R. (2024). Likelihood-free parameter estimation with neural Bayes estimators, *The American Statistician*, **78**(1), 1–14.

Sang, H. and Huang, J. Z. (2012). A full scale approximation of covariance functions for large spatial data sets, *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **74**(1), 111–132.

Stein, M. L. (2014). Limitations on low rank approximations for covariance matrices of spatial data, *Spatial Statistics*, **8**, 1–19.

Stein, M. L., Chi, Z. and Welty, L. J. (2004). Approximating likelihoods for large spatial data sets, *Journal of the Royal Statistical Society Series B*: *Statistical Methodology*, **66**(2), 275–296.

Stevens, F. R., Gaughan, A. E., Linard, C. and Tatem, A. J. (2015). Disaggregating census data for population mapping using random forests with remotely-sensed and ancillary data, *PLoS ONE*, **10**, e0107042.

Tonks, A., Harris, T., Li, B., Brown, W. and Smith, R. (2024). Forecasting West Nile virus with graph neural networks: Harnessing spatial dependence in irregularly sampled geospatial data, *GeoHealth*, **8**(7), e2023GH000784, https://doi.org/10.1029/2023GH000784.

Tsutsumi, M. and Seya, H. (2009). Hedonic approaches based on spatial econometrics and spatial statistics: Application to evaluation of project benefits, *Journal of Geographical Systems*, **11**, 357–380.

堤盛人, 瀬谷創 (2012). 応用空間統計学の二つの潮流：空間統計学と空間計量経済学, 統計数理, **60**(1), 3–25.

Tzeng, S. and Huang, H. C. (2018). Resolution adaptive fixed rank kriging, *Technometrics*, **60**(2), 198–208.

Van Niekerk, J. and Rue, H. (2024). Low-rank variational Bayes correction to the Laplace method, *Journal of Machine Learning Research*, **25**(62), 1–25.

Van Niekerk, J., Krainski, E., Rustand, D. and Rue, H. (2023). A new avenue for Bayesian inference with INLA, *Computational Statistics and Data Analysis*, **181**, https://doi.org/10.1016/j.csda.2023.107692.

Van Stein, B., Wang, H., Kowalczyk, W., Emmerich, M. and Bäck, T. (2020). Cluster-based Kriging approximation algorithms for complexity reduction, *Applied Intelligence*, **50**(3), 778–791.

Vargas-Muñoz, J. E., Srivastava, S., Tuia, D. and Falcão, A. X. (2020). OpenStreetMap: Challenges and opportunities in machine learning and remote sensing, *IEEE Geoscience and Remote Sensing*

Magazine, **9**, 184–199.

Wang, H., Guan, Y. and Reich, B. (2019). Nearest-neighbor neural networks for geostatistics, *Proceedings of International Conference on Data Mining Workshops*, 196–205, https://doi.org/10.1109/ICDMW.2019.00038.

Wiens, A., Nychka, D. and Kleiber, W. (2020). Modeling spatial data using local likelihood estimation and a Matérn to spatial autoregressive translation, *Environmetrics*, **31**(6), e2652.

Wikle, C. K. (2019). Comparison of deep neural networks and deep hierarchical models for spatio-temporal data, *Journal of Agricultural, Biological and Environmental Statistics*, **24**(2), 175–203.

Wikle, C. K. and Zammit-Mangion, A. (2023). Statistical deep learning for spatial and spatiotemporal data, *Annual Review of Statistics and Its Application*, **10**(1), 247–270.

Wikle, C. K., Zammit-Mangion, A. and Cressie, N. (2019). *Spatio-temporal Statistics with R*, Chapman and Hall/CRC, Boca Raton.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian Processes for Machine Learning*, MIT Press, Cambridge, Massachusetts.

Wood, S. N., Li, Z., Shaddick, G. and Augustin, N. H. (2017). Generalized additive models for gigadata: modeling the UK black smoke network daily data, *Journal of the American Statistical Association*, **112**(519), 1199–1210.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C. and Philip, S. Y. (2020). A comprehensive survey on graph neural networks, *IEEE Transactions on Neural Networks and Learning Systems*, **32**(1), 4–24.

Yoshida, T., Murakami, D. and Seya, H. (2022). Spatial prediction of apartment rent using regression-based and machine learning-based approaches with a large dataset, *The Journal of Real Estate Finance and Economics*, **69**, 1–28.

Yu, Y., Si, X., Hu, C. and Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures, *Neural Computation*, **31**(7), 1235–1270.

Zammit-Mangion, A., Bertolacci, M., Fisher, J., Stavert, A., Rigby, M. L., Cao, Y. and Cressie, N. (2022a). WOMBAT v1.0: A fully Bayesian global flux-inversion framework, *Geoscientific Model Development*, **15**(1), 45–73.

Zammit-Mangion, A., Ng, T. L. J., Vu, Q. and Filippone, M. (2022b). Deep compositional spatial models, *Journal of the American Statistical Association*, **117**(540), 1787–1808.

Zammit-Mangion, A., Kaminski, M. D., Tran, B. H., Filippone, M. and Cressie, N. (2024a). Spatial Bayesian neural networks, *Spatial Statistics*, https://doi.org/10.1016/j.spasta.2024.100825.

Zammit-Mangion, A., Sainsbury-Dale, M. and Huser, R. (2024b). Neural Methods for Amortised Parameter Inference, ArXiv, https://doi.org/10.48550/arXiv.2404.12484.

Zhan, W. and Datta, A. (2024). Neural networks for geospatial data, *Journal of the American Statistical Association*, https://doi.org/10.1080/01621459.2024.2356293.

Zhang, L., Banerjee, S. and Finley, A. O. (2021). High-dimensional multivariate geostatistics: A Bayesian matrix-normal approach, *Environmetrics*, **32**(4), e2675.

Zhang, W., Han, J., Xu, Z., Ni, H., Liu, H. and Xiong, H. (2024). Towards urban general intelligence: A review and outlook of urban foundation models, ArXiv, https://doi.org/10.48550/arXiv.2402.01749.

Recent Advances in Spatial Statistics

Daisuke Murakami

The Institute of Statistical Mathematics

Statistical methods for geospatial data have been studied in spatial statistics. Especially in recent years, when spatio-temporal data have become larger and more diverse, methods that are both computationally efficient and flexible have been developed rapidly. This study therefore summarizes recent advances in spatial statistics. We first introduce basic spatial statistical models and their challenges. Next, approximations for large samples are classified into low-rank approximation, covariance approximation, and approximation of the precision matrix, and studies are reviewed for each. Then, methods for flexibility modeling spatial processes and observations are reviewed respectively. Software packages for implementing spatial statistical methods are explained after that. Finally, future directions in spatial statistics are discussed.

Key words: Spatial statistics, spatiotemporal modeling, Gaussian process, spatial correlation, neural networks.