arXiv:2603.03247v1 [stat.ME] 3 Mar 2026

融合稀疏观测与密集模拟进行空间极值分析：在美国的应用

沿海海平面

Brian N. White[1], *, Brian Blanton[2], Rick Luettich[2,3], Richard L. Smith[1]1统计与运筹学系，北卡罗来纳大学教堂山分校2文艺复兴计算研究所（RENCI），北卡罗来纳大学教堂山分校3海洋科学研究所，北卡罗来纳大学教堂山分校*通信作者：bnw@unc.edu

摘要

从稀疏观测网络估计空间极值，会产生不确定的重现水平图；而基于物理的模拟模型通常能提供密集输出，作为补充数据源。我们提出一种两阶段频率学派框架，用于融合观测数据与模拟数据。阶段 1 中，在每个站点分别拟合广义极值（GEV）分布，并在适当情况下采用非平稳位置参数，以考虑观测到的趋势。第二阶段中，将所有数据源的参数估计值作为高维空间过程，通过共区域化线性模型（LMC）进行联合建模。对于没有共址站点、但网络在空间上交错分布的情况，估计得到的跨来源相关性可实现信息传递；为所得似然计算解析梯度，则能使估计在计算上切实可行。我们将该框架应用于美国沿海

1979–2021 年的海平面数据，融合了 29 个美国国家海洋和大气管理局（NOAA）验潮站记录与 100 个 ADCIRC 水动力模拟站点的数据。留一交叉验证显示，与仅测站模型相比，百年重现水平均方根误差降低了 35%。地理分块交叉验证证实，即使在空间外推情况下，融合效益仍然存在。该方法通过 R 软件包 evfuse 实现。

关键词：数据融合、共区域化、重现水平、风暴潮、地质统计学、交叉验证

1 引言

极值分布会随空间变化，但观测网络通常不够密集，难以精确估计这种变化。基于物理的模拟模型能够以远高于观测网络的空间密度生成关注量；将其输出与稀疏观测融合，可以提高重现水平估计的覆盖范围和精度。重现水平是指平均每隔 T 年预计会超过一次的 T 年阈值。我们提出了一个通用的融合框架，并通过结合潮汐测潮仪观测与水动力模拟输出，绘制美国东海岸与墨西哥湾沿岸极端沿海海平面的分布图，以此展示该框架的应用。

这两种数据源体现了环境监测中的一种常见结构。美国国家海洋和大气管理局的潮汐水位计网络提供可靠的长期海平面记录，但测潮仪分布稀疏，且相较于所关注的重现期，其记录时间较短，因此在单个站点的尾部参数估计中会产生较大的不确定性。基于物理的水动力模型（如 ADCIRC（Luettich 等人，1992））能够模拟数百个沿海位置的海平面，弥补测潮仪在空间密度上的不足，但其输出反映的是模型物理过程，而非直接观测。

空间极值的统计方法包括最大值稳定过程，以及参数随空间变化的分层广义极值模型（Davison et al., 2012; Cooley et al., 2007;

Sang & Gelfand, 2010）。最大稳定模型在理论上颇具吸引力，并能刻画极端事件的空间范围，但其似然即使对于中等数量的站点也难以处理。当主要关注的是边际重现水平而非联合尾部相依性时，通常更倾向于潜在过程模型：将每个地点的广义极值分布参数视为某一空间过程的实现，推断既可通过贝叶斯层次建模进行（Cooley et al., 2007; Sang & Gelfand, 2010; Gelfand et al., 2004; Finley et al., 2008），也可通过将逐点最大似然估计在空间上加以平滑的两阶段频率学派方法进行（Russell et al., 2020）。将原位观测与网格化模型输出或遥感相结合的数据融合，在空间统计学中已有丰富历史（Cressie & Wikle, 2011; Nguyen et al., 2012; Berliner, 2003）。尽管如此，融合方法很少被应用于极值参数；其目标是整合异质数据源以改进尾部推断，而非对均值场进行插值。

我们采用两阶段频率学派方法，以兼顾计算可行性与可解释性；该方法建立在Russell等人（2020）的工作之上，他们将协区域化线性模型（LMC）与针对空间变化的广义极值分布参数的两阶段频率学派推断相结合，并应用于单一的墨西哥湾沿岸降水站点网络。我们的贡献从两个方向扩展了这一框架。

首先，我们将该模型推广为用单一高维空间过程来表示来自多个数据源的广义极值分布参数。协区域化线性模型在异位采样下是可识别的（Wackernagel, 2003），因此无需共址站点，即可从空间交错网络估计跨来源相关性。选择矩阵从完整的潜在协方差中提取观测分量，从而得到标准高斯似然。由LMC 结构（S1节）导出的所得33参数对数似然的解析梯度，将台式工作站上的优化时间从约90分钟缩短至3分钟以内，从而使多起点优化和自助法推断切实可行。

其次，我们提供了交叉验证证据，量化融合带来的提升。在能够适应观测趋势的非平稳第一阶段模型下，联合模型将

百年重现水平的留一交叉验证均方根误差降低了 35%，相较于仅 NOAA 基线。地理分块交叉验证表明，尽管效果有所减弱，这一优势在空间外推情况下仍然存在。

2 数据

空间范围从得克萨斯州南部（26.1°N，97.2° W）延伸至缅因州北部（44.9° N，67.0°W）。我们分析了来自两个数据源的1979–2021年（$T_0 = 43$ 年）年最大海平面。

2.1 NOAA潮位站观测

我们于2022年5月从NOAA潮汐与洋流门户获取了逐时海平面测量数据。所有测量值均以米为单位，相对于平均海平面（MSL）潮汐基准面（1983–2001年时段）表示。

我们采用了两步质量筛选：（i）剔除预期逐时观测数据不足90%的年份；（ii）剔除在步骤（i）后年度最大值少于20个的站点。筛选后共得到 $L_N = 29$ 个站点，每个站点至少有34个年度最大值（29个站点中有24个超过40个）。90%的阈值可降低错过真正年最大值的风险。相对于关注的重现期（100年）而言，这些记录较短：根据34–43个年度最大值估计控制尾部行为的形状参数 ξ，会使单个站点的不确定性显著增加，因此有必要进行空间强度借用。

如果验潮站在极端事件期间发生故障，最大的观测值因此被删失，就会使 ξ 向下偏倚。这进一步说明了融合的必要性：模拟结果不受仪器故障影响，而第二阶段的跨来源形状相关性可以减轻这种偏倚。

2.2 ADCIRC模拟数据

先进环流（ADCIRC）模型是一种有限元水动力模型，可在非结构化网格上求解浅水方程（Luettich 等人，1992；Dietrich 等人，2011）。文艺复兴计算研究所（RENCI）在北卡罗来纳大学教堂山分校的支持下，生成了 $L_A = 100$ 沿海站点在1979–2021年间逐小时的模拟海平面值，模拟由CFSR再分析大气强迫驱动。这些历史重建仅基于模型物理过程生成。虽然ADCIRC在包含数千个节点的非结构化网格上运行，但我们使用了100个具有代表性的沿海站点；对于沿6,000公里海岸线所关注的大尺度空间格局而言，这一密度已足以进行解析。若要研究更精细尺度的问题，则需要从ADCIRC网格中抽取更密集的站点，并相应扩展空间模型（第5.4节）。由于记录在设计上是完整的，每个站点均有全部43个年度最大值。

2.3 研究设计

站点总数为 $L = L_N + L_A = 129$，没有任何站点同时拥有美国国家海洋和大气管理局和ADCIRC的观测数据。ADCIRC站点填补了NOAA站点网络中的空间空白，尤其是沿墨西哥湾沿岸和大西洋东南沿岸（图1）。NOAA站点与ADCIRC站点之间的最近邻距离中位数为21.7公里（范围：1.9–119.8公里），使两个网络得以密集交错。

2.4 趋势诊断

在拟合空间模型之前，我们使用Mann-Kendall非参数趋势检验（Mann，1945；Kendall，1975）并结合Sen斜率估计量（Sen，1968），评估各站点年度最大值的时间平稳性。在NOAA站点中，29个里有19个（66%）呈现显著正趋势（$p < 0.05$），Sen斜率中位数为4.6毫米/年。相比之下，100个ADCIRC站点中仅有1个呈现显著趋势（正如在 $\alpha = 0.05$处所预期），而ADCIRC中位数

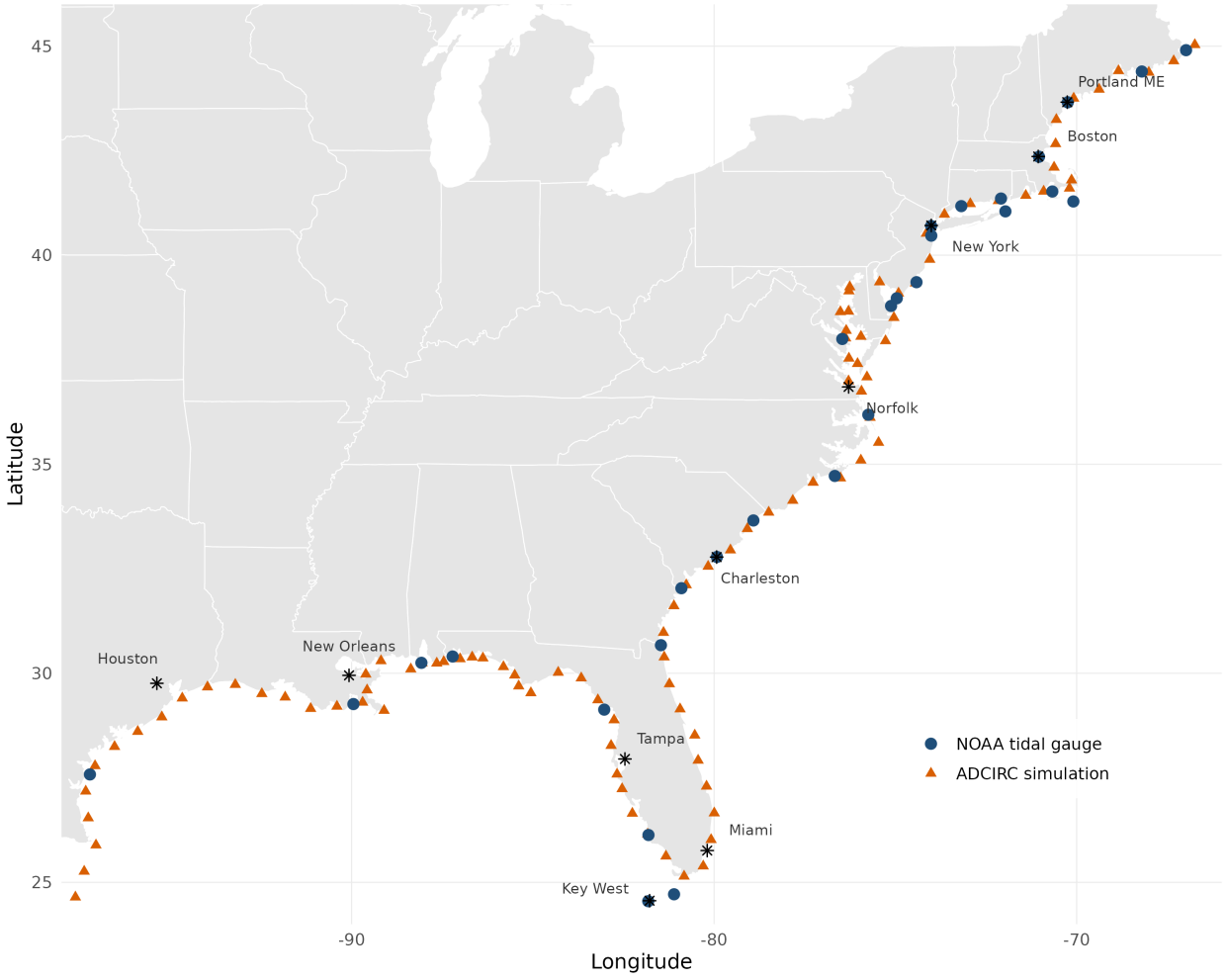

图1：研究区域，展示了美国东海岸与墨西哥湾沿岸29个美国国家海洋和大气管理局潮位计站（圆圈）和100个ADCIRC模拟站点（三角形）的位置。

斜率（0.4毫米/年）小一个数量级（表1；图S2）。ADCIRC近乎为零的趋势符合预期：模拟由CFSR再分析强迫驱动，其中未纳入长期海平面上升，因此任何趋势仅反映大气强迫的年际变率。

相对于重现水平，这些趋势较为温和：4.6毫米/年的中位数斜率在43年记录期内产生约0.2米的偏移，而100年重现水平为1–4米。尽管如此，呈趋势的美国国家海洋和大气管理局与平稳的ADCIRC之间的这种不对称性会影响第二阶段估计的跨来源相关性，因为美国国家海洋和大气管理局的年度最大值包含趋势驱动的方差，而ADCIRC的模拟则缺乏这种方差。

3 统计框架

3.1 阶段1：边缘广义极值分布拟合

已有若干成熟方法可用于对环境时间序列中的极值进行建模（Coles，2001），包括通过广义极值（GEV）分布处理年度区块最大值，以及通过广义帕累托分布处理超阈值峰值。我们在此采用区块最大值。在每个站点 s_l（$l = 1, \ldots, L$）处，我们用最大似然法将广义极值分布拟合到年度最大值（见Coles，2001，第4章），从而避免在129个异质站点上进行阈值选择，并确保与双源设计兼容。广义极值分布函数为

$$F(y \mid \mu, \sigma, \xi) = \exp\left\{-\left[1 + \xi\left(\frac{y-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}, \tag{1}$$

定义于 $\{y:\ 1 + \xi(y-\mu)/\sigma > 0\}$，其中位置参数为 $\mu \in \mathbb{R}$、尺度参数为 $\sigma > 0$、形状参数为 $\xi \in \mathbb{R}$(Coles, 2001)。T年重现水平为

$$r_T = \mu + \frac{\sigma}{\xi}\left[(-\log(1 - 1/T))^{-\xi} - 1\right]. \tag{2}$$

由于 66% 的 NOAA站点呈现显著正趋势（第2节），我们对 NOAA站点拟合非平稳广义极值分布，其中 $\mu(t) = \mu_0 + \mu_1(t - t_{\mathrm{ref}})$、 $t_{\mathrm{ref}} = 2000$ 和 σ， ξ 为平稳。以 $t_{\mathrm{ref}} = 2000$（接近记录中点）为中心，使 $\hat{}\mu_0$ 和 $\hat{}\mu_1$ 近似正交，从而将截距的标准误降至最低。在此中心化处理下， μ_0 是 2000 年的位置参数，重现水平具有明确的时间含义： T年重现水平表示在 2000 年条件下，每年以 $1/T$ 的概率被超过的水位。没有显著趋势的 ADCIRC站点仍采用平稳拟合。

在每个站点，输出为一个三元组 $\hat{}\boldsymbol{\theta}(s_l) \in \mathbb{R}^3$： NOAA站点处为 $(\hat{}\mu_0, \log \hat{}\sigma, \hat{}\xi)^\top$（年份为 2000 的截距），ADCIRC站点处为 $(\hat{}\mu, \log \hat{}\sigma, \hat{}\xi)^\top$（平稳最大似然估计）。趋势斜率 μ_1 未纳入空间模型，因为 ADCIRC 站点没有可共同建模的趋势。将μ_1 作为第七个维度纳入模型，需要将其跨源相关性固定为零，或仅根据 NOAA站点进行估计；这两种做法都无法提供 μ_0已有信息之外的空间信息。将 μ_1 纳入七维扩展模型，是气候预测应用的一个自然发展方向。对数尺度参数化确保了第二阶段中的无约束优化；在其他应用中，可以对其他参数类型采用不同的链接函数，但统计框架保持不变。

3.2 第二阶段：多元空间模型

3.2.1 潜在过程与共区域化

协区域化线性模型假设，在任意位置处，六个广义极值分布参数都是六种相互独立的空间模式的线性组合，每种模式都有各自的相关范围，因此不同的参数组合可以呈现不同尺度的空间变异。在每个位置$s \in D$，定义六维潜在向量

$$\boldsymbol{\theta}(s) = \big(\underbrace{\mu_N(s),\ \log\sigma_N(s),\ \xi_N(s)}_{\text{NOAA}},\ \underbrace{\mu_A(s),\ \log\sigma_A(s),\ \xi_A(s)}_{\text{ADCIRC}}\big)^\top. \tag{3}$$

在非平稳第一阶段模型下，$\mu_N(s)$表示美国国家海洋和大气管理局站点处的2000年截距 $\mu_0(s)$，而 $\mu_A(s)$表示ADCIRC站点处的平稳位置参数。由于ADCIRC未显示趋势，且NOAA参考年份（$t_{\text{ref}} = 2000$）接近记录中点，因此这两个量作为共同时间点的位置参数大致具有可比性。尽管在任一站点上每次只观测到一个来源的数据，但每个位置都存在NOAA和ADCIRC的广义极值分布参数作为潜在量。我们通过共区域化线性模型（LMC；Journel 和 Huijbregts 1978；Wackernagel 2003；另见 Finley 等，2008）对 $\boldsymbol{\theta}(s)$建模：

$$\boldsymbol{\theta}(s) = \boldsymbol{\beta} + \mathbf{A}\,\boldsymbol{\delta}(s), \tag{4}$$

其中，$\boldsymbol{\beta} \in \mathbb{R}^6$ 是空间均值，$\mathbf{A}$ 是一个 6×6 下三角矩阵，而 $\boldsymbol{\delta}(s) = (\delta_1(s), \ldots, \delta_6(s))^\top$ 由六个相互独立、均值为零的高斯过程组成，其协方差为指数协方差 $\mathrm{Cov}(\delta_i(s), \delta_i(s')) = \exp(-\|s - s'\|/\rho_i)$，$\rho_i > 0$。指数协方差对应光滑度为 $\nu = 1/2$ 的 Matérn 模型（见 Stein，1999）；考虑到 129 个站点对区分光滑度的能力有限，因此选择该模型以保持简约性。全文中，$\|s - s'\|$ 表示以千米为单位的大圆（haversine）距离。

任意位置参数处的边际协方差为$\mathbf{A}\mathbf{A}^\top$，其非对角块表示跨源相关性，而这些相关性正是实现信息传递的关键量。两个位置 s和 s' 之间的交叉协方差为

$$\mathrm{Cov}(\boldsymbol{\theta}(s),\, \boldsymbol{\theta}(s')) = \mathbf{A}\,\mathrm{diag}\big(\exp(-\|s - s'\|/\rho_1),\, \ldots,\, \exp(-\|s - s'\|/\rho_6)\big)\,\mathbf{A}^\top. \tag{5}$$

由于每个潜在高斯过程 δ_i 都有各自的范围参数 ρ_i，六个广义极值分布参数的不同线性组合可以表现出不同的空间相关长度。例如，若第一个高斯过程具有较大的范围 ρ_1，并主要通过$A_{1,1}$ 和$A_{4,1}$加载到 μ_N 和 μ_A 上，则位置参数在空间上较为平滑，而其他参数组合在较短距离上变化。具有单一共享范围参数的可分离模型无法刻画这种结构。

可识别性。 一个自然的问题是，在没有共址观测的情况下，跨来源协方差参数是否可识别。在协区域化线性模型下，它们是可识别的：共享的范围参数 ρ_k 由源内配对识别，跨来源系数则由在跨来源距离上求值的线性系统确定（Wackernagel, 2003，第23章）。这要求范围参数彼此不同，且两个网络具有充分的空间交错；此处两者均得到满足，因为源间距离的中位数（21.7 km）远短于最短的估计范围（188 km）。

3.2.2 部分观测与选择矩阵

ˆ 在观测站点处， $\boldsymbol{\theta}(s_l) = \boldsymbol{\theta}(s_l) + \boldsymbol{\varepsilon}(s_l)$，其中 $\boldsymbol{\varepsilon}(s_l)$ 为阶段 1 估计误差。在 NOAA站点，仅观测到 $\boldsymbol{\theta}(s_l)$ 的第 1–3 个分量；在 ADCIRC站点，仅观测到第 4–6 个分量。令 $n_{\text{obs}} = 3L_N + 3L_A = 387$。每个观测值 m 均对应一个参数索引$p(m) \in \{1, \ldots, 6\}$和一个站点索引 $s(m) \in \{1, \ldots, L\}$。选择矩阵 S从完整的潜在向量 $\boldsymbol{\theta}_{\text{full}} \in \mathbb{R}^{6L}$ （按参数优先顺序存储）中提取观测条目。该矩阵的每一行恰有一个非零元素：

$$S_{m,\,(p(m)-1)L+s(m)} = 1, \qquad \text{all other entries zero.} \tag{6}$$

简单来说，S 的每一行都会从完整的潜在向量中选取一个站点上的一个观测参数。这种形式化表示适用于任意部分观测模式：包括只有一个数据源、多个数据源，或观测参数子集各不相同的站点。

3.2.3 似然

ˆ 观测参数估计值 $\boldsymbol{\Theta}_{\text{obs}}$ 是真实潜在参数的含噪声版本。因此，其协方差$\mathbf{V}$ 包含两个分量：来自共区域化模型的信号项 $\boldsymbol{\Sigma}_{\text{obs}}$ ，代表真实的空间结构；以及来自第一阶段估计误差的噪声项$\mathbf{W}$ 。矩阵$\mathbf{W}$ 仅根据自助法估计一次（如下所述），并且

在整个第 2 阶段优化过程中保持固定，从而仅通过最大似然估计共区域化参数（$\boldsymbol{\beta}, \mathbf{A}$ $\boldsymbol{\rho}$）。

在该模型下：

$$\hat{\boldsymbol{\Theta}}_{\text{obs}} \sim \mathcal{N}\big(\mathbf{S}(\boldsymbol{\beta} \otimes \mathbf{1}_L),\ \mathbf{V}\big), \qquad \mathbf{V} = \boldsymbol{\Sigma}_{\text{obs}} + \mathbf{W}, \tag{7}$$

ˆ其中， $\boldsymbol{\Theta}_{\text{obs}} \in \mathbb{R}^{n_{\text{obs}}}$ 是按与 $\boldsymbol{\theta}_{\text{full}}$相同的参数优先顺序排列的第 1 阶段最大似然估计向量：首先是全部 L位置估计值，然后是全部 L 对数尺度估计值，接着是全部 L 形状估计值；每个站点根据所观测到的来源贡献三个参数。均值向量 $\boldsymbol{\beta} =(\beta_1,\ \ldots\ldots,\ \beta_6)^\top$遵循公式 (3) 的顺序： $\beta_1 = \mathbb{E}[\mu_N]$、 $\beta_2 = \mathbb{E}[$对数 $\sigma_N]$、$\beta_3 = \mathbb{E}[\xi_N]$、 $\beta_4 = \mathbb{E}[\mu_A]$、 $\beta_5 = \mathbb{E}[$对数 $\sigma_A]$、 $\beta_6 = \mathbb{E}[\xi_A]$，而$\mathbf{A}$ 的下三角元素和范围参数 $\rho_1,\ \ldots\ldots,\ \rho_6$ 也采用相同的顺序。观测到的共区域化协方差为 $\boldsymbol{\Sigma}_{\text{obs}} = \mathbf{S}\boldsymbol{\Sigma}\mathbf{S}^\top$。对于观测值 $m,\ n$，其参数索引为 $p(m)$ $p(n)$，站点索引为 $s(m)$ $s(n)$：

$$[\boldsymbol{\Sigma}_{\text{obs}}]_{mn} = \sum_{i=1}^{6} A_{p(m),i}\, A_{p(n),i}\, \exp(-\|s_{s(m)} - s_{s(n)}\|/\rho_i), \tag{8}$$

在不构造完整的 $6L \times 6L$ 矩阵的情况下，于 $O(n_{\text{obs}}^2 \cdot 6)$中计算。

ˆ令$\mathbf{r} = \boldsymbol{\Theta}_{\text{obs}} - \mathbf{S}(\boldsymbol{\beta} \otimes \mathbf{1}_L)$，则需要在$(\boldsymbol{\beta}$,$\mathbf{A}$ $\boldsymbol{\rho})$上最小化的负对数似然为

$$\ell(\boldsymbol{\beta}, \mathbf{A}, \boldsymbol{\rho}) = \tfrac{1}{2}\big[n_{\text{obs}} \log(2\pi) + \log|\mathbf{V}| + \mathbf{r}^\top \mathbf{V}^{-1}\mathbf{r}\big], \tag{9}$$

其中，$\mathbf{V}$ 通过公式（8）和固定的$\mathbf{W}$依赖于$(\mathbf{A}$ $\boldsymbol{\rho})$。优化针对33维参数向量$(\boldsymbol{\beta}$,vech$(\mathbf{A})$, log $\boldsymbol{\rho})$ $\in \mathbb{R}^{33}$进行，其中vech$(\mathbf{A})$表示$\mathbf{A}$ 的21个下三角元素，而log-rho重新参数化确保范围参数为正。

测量误差协方差$\mathbf{W}$ 通过非参数区块自助法估计（$B = 500$ 个重复样本），在每个数据源内对所有站点的年份进行联合重抽样，以保留

站点间的风暴依赖性。由于第一阶段估计误差源自不同的数据生成机制，因此不同数据源之间的年份独立重抽样。在有限样本中，共享的时间结构可能导致残余的数据源间依赖；联合重抽样方案则可作为一种保守的替代方法。

按照 Russell 等人（2020）的做法，采用 Wendland 协方差截尾（Wendland, 1995；Furrer et al., 2006），对自助法协方差进行正则化，范围为 $\lambda = 300$km。Wend- land C_4 函数 $\psi(d/\lambda) = (1 - d/\lambda)_+^6(35(d/\lambda)^2 + 18\,d/\lambda + 3)/3$ 是一种紧支撑的正定径向基函数，当 $d \geq \lambda$时其值为零。$\mathbf{W}$ 中的每个元素都乘以 $\psi(d_{ij}/\lambda)$，以抑制有限自助法所产生的虚假远距离相关性。为选择 λ，我们按距离对站点对进行分箱，并计算每个箱内自助法相关性的绝对值均值。相关性从 0–50 km 范围内的 $\sim$0.6 降至 200–500 km 范围内的 $\sim$0.2，随后趋于平稳，达到 $\sim$0.10 的噪声底。$\lambda = 300$km 的取值既能捕捉真实信号，又能在噪声底之前将相关性衰减至零。在 $\lambda \in \{150, 300, 500\}$km 范围内，结果对这一选择几乎不敏感（表6）。

3.3 优化

负对数似然（9）使用 L-BFGS-B 进行最小化，其解析梯度由 LMC 结构导出（S1节）。从对数均匀分布的$[50, 5000]$km初始变程进行二十次随机重启，以防陷入局部最优解。

3.4 部分观测克里金法

给定已拟合的 para $\tilde{}$ $\tilde{}$米（$\boldsymbol{\beta}$，$\mathbf{A}$ $\tilde{\boldsymbol{\rho}}$），新位置 s_0 处的克里金均值和协方差为 e

$$\tilde{\boldsymbol{\theta}}(s_0) = \boldsymbol{\beta} + \mathbf{C}^\top \mathbf{V}^{-1} \mathbf{r}, \tag{10}$$

$$\mathbf{V}_{\boldsymbol{\theta}}(s_0) = \mathbf{A}\mathbf{A}^\top - \mathbf{C}^\top \mathbf{V}^{-1} \mathbf{C}, \tag{11}$$

其中 $C_{m,\ k} = \sum_{i=1}^{6} A_{p(m),\ i} A_{k,\ i} \exp(-\|s_{s(m)} - s_0\|/\rho_i)$。 $\tilde{\boldsymbol{\theta}}(s_0)$˜的前三个分量给出预测的美国国家海洋和大气管理局参数（$\tilde{\mu}_N$，log $\tilde{\sigma}_N$，ξ_N），而 ˜$\mathbf{V}_{\boldsymbol{\theta}}(s_0)$ 的左上角 3×3块给出它们的协方差。在 s_0处的 T年重现水平为 $\tilde{r}_T(s_0) = g(\tilde{\mu}_N$，log $\tilde{\sigma}_N$，$\xi_N)$，由式（2）得到，其标准误通过デルタ法计算：

$$\mathrm{SE}(\tilde{r}_T) = \sqrt{\nabla g^{\top} \left[\mathbf{V}_{\boldsymbol{\theta}}(s_0)\right]_{1:3,\,1:3} \nabla g}, \tag{12}$$

其中 $\nabla g = (\partial r_T/\partial \mu,\quad \partial r_T/\partial \log\ \sigma,\quad \partial r_T/\partial \xi)^{\top}$。令 $y_T = -\log(1 - 1/T)$，偏导数为

$$\frac{\partial r_T}{\partial \mu} = 1, \qquad \frac{\partial r_T}{\partial \log \sigma} = r_T - \mu, \qquad \frac{\partial r_T}{\partial \xi} = \frac{\sigma}{\xi^2}\left(1 - y_T^{\xi}\right) - \frac{\sigma}{\xi}\, y_T^{-\xi} \log y_T. \tag{13}$$

对于下文报告的结果，我们通过蒙特卡罗模拟（从克里金后验分布中抽取 10,000 次）计算标准误，这能正确传播式（2）的非线性。デルタ法提供了一种有用的闭式检验；两种方法的比较见第4.6.5节和图S6。

3.5 模型评估

我们将联合模型与仅 NOAA 和仅 ADCIRC 基线模型（均为 3 维协区域化线性模型）进行比较。联合模型在 129个站点上有 33 个自由参数（6 个均值、21 个共区域化参数、6 个范围参数）；每个单一数据源基线模型有 12 个参数。我们通过两种难度逐渐增加的交叉验证设计评估预测性能，随后进行敏感性分析。

29个NOAA站点的留一交叉验证采用闭式分块捷径，无需重新拟合模型29次，而是以全数据协方差参数估计值为条件（Rasmussen与Williams，2006）。由于每个NOAA站点贡献一个包含三项观测的区块（μ，log σ，ξ），我们采用分块推广方法（S1节）。我们报告各参数、百年重现水平、总对数预测密度（LPD）的留一均方根误差（LOO-RMSE），以及通过校准

概率积分变换（PIT）直方图和预测区间的经验覆盖率进行校准。

地理分块交叉验证通过依次留出五个连续的沿海区块（墨西哥湾沿岸， $n = 4$；佛罗里达州， $n = 5$；美国东南部， $n = 5$；美国中大西洋地区， $n = 6$；新英格兰， $n = 9$），并保留所有 ADCIRC 站点、在每一折中从头重新拟合两个模型的第二阶段，从而检验融合在空间外推下的表现。

我们评估了对三种分析者选择的稳健性：(i) 通过对子抽样 ADCIRC 站点来改变模拟网络密度；(ii) 自助法渐消范围 λ；以及 (iii) 自助法批次变化。模拟研究进一步证实，在观测到的非共址抽样设计下，可以恢复跨来源相关性（第 4.6.4 节）。

4 结果

4.1 边际广义极值分布拟合

所有 129 个广义极值分布拟合均已收敛。安德森–达令拟合优度检验仅在 129 个站点中的 1 个站点（ $\alpha = 0.05$ ）拒绝广义极值分布原假设，这与偶然情况下预期出现的 $\sim$6.5 个假阳性一致；邦费罗尼校正后没有站点达到显著水平。QQ图和自助法抽样分布证实，所有站点的拟合均足够良好（补充材料第 S2 节）。 ξ 的自助法标准误差较大（美国国家海洋和大气管理局站点的中位数为 0.21，ADCIRC 站点为 0.17），这直接源于记录时间较短，也是第二阶段采用空间平滑的主要动机。两阶段结构使第二阶段的高斯近似能够直接验证： ξ 的自助法抽样分布显示，尾部最重的墨西哥湾沿岸站点存在轻微正偏，但没有明显偏离正态性（图 S3）；概率积分变换校准（表4）则证实，这种近似不会导致预测校准不当。

估计的位置参数（美国国家海洋和大气管理局处为 $\hat{\mu}_0$ ，ADCIRC 处为 $\hat{\mu}$ ）介于 0.46 至 3.96 米之间，在大西洋沿岸呈现明显的由南向北梯度。形状参数 $\hat{\xi}$ 呈现出清晰的地理分布特征：墨西哥湾沿岸的正值（$\sim$0.2–0.5）表明

存在重尾，而哈特拉斯角以北接近零或略为负值则表明尾部较轻。美国国家海洋和大气管理局与 ADCIRC 的估计总体一致，ADCIRC站点填补了水位计网络中的空白（图 2）。

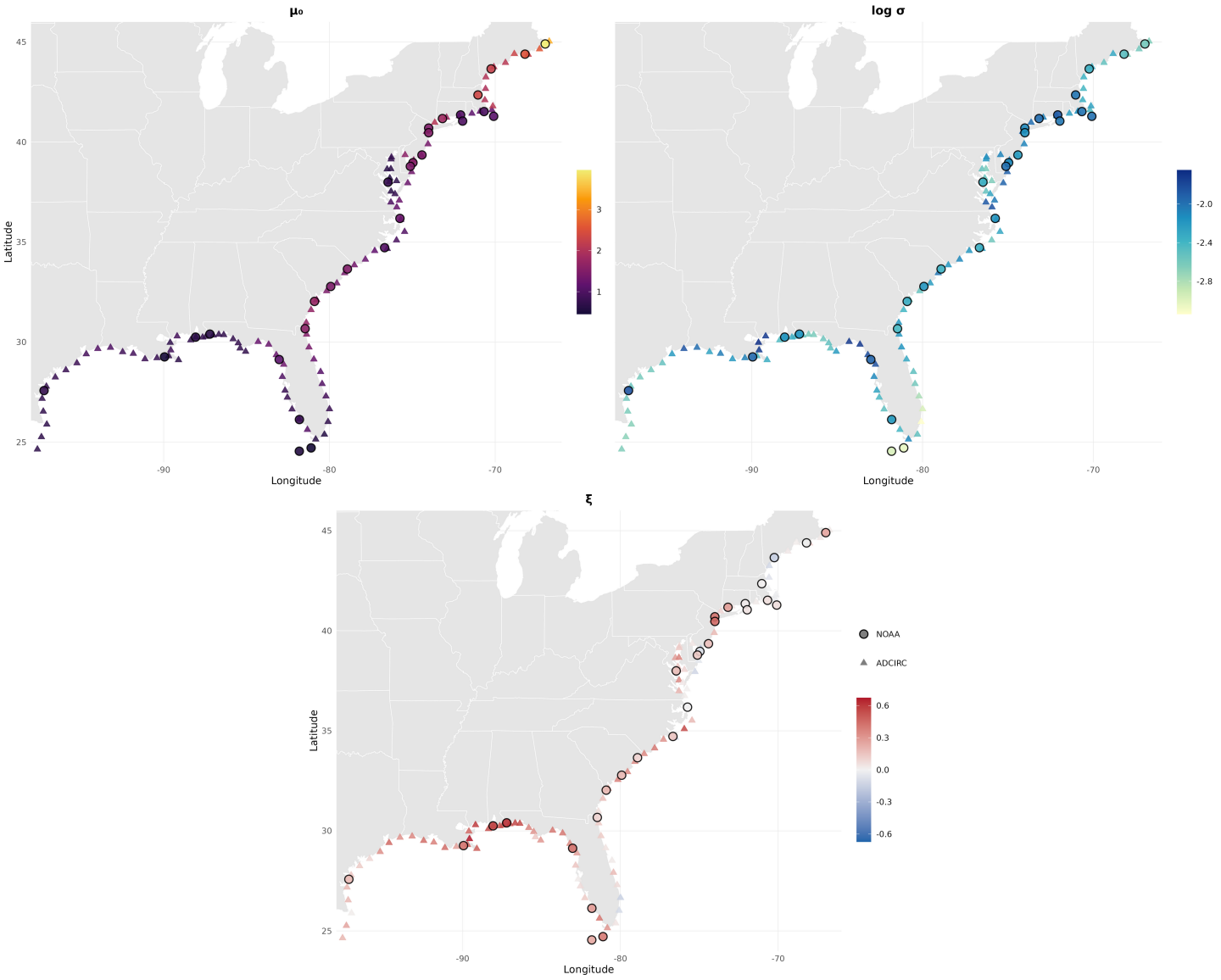

图 2：阶段 1 中 129个站点的逐点广义极值分布最大似然估计。美国国家海洋和大气管理局潮位计（带边框的圆形）和 ADCIRC 模拟站点（三角形）。左：位置参数（米）。对于 NOAA站点，这是非平稳拟合中的 $\hat{}\mu_0$，即 2000年截距；对于 ADCIRC站点，这是平稳最大似然估计 $\hat{}\mu_0$。中：$\log\sigma_0$。右：形状 ξ，采用以零为中心的发散色标。沿大西洋沿岸，位置参数从南向北递增；而墨西哥湾沿岸的 ξ 值偏高，表明具有重尾。

4.2 拟合空间模型

多起点优化（20 个起点）收敛于一致的最优解。估计的跨来源相关性如下：

$$\mathrm{Cor}(\mu_N, \mu_A) = 0.995, \qquad \mathrm{Cor}(\xi_N, \xi_A) = 0.837, \qquad \mathrm{Cor}(\log \sigma_N, \log \sigma_A) = 0.443. \tag{14}$$

近乎完美的位置相关性（$r = 0.995$）是融合的主要驱动因素：两个数据源捕捉到了相同的空间模式，因此，100 个 ADCIRC站点实际上使包含 29 个站点的水位计网络更加密集。形状相关性（$r = 0.84$）之所以重要，是因为 ξ 是估计精度最低的参数，却对重现水平具有最大的影响。中等程度的尺度相关性（$r = 0.44$）表明，与位置参数或形状参数相比，两个数据源在其尺度参数上的一致性较低。

估计的范围参数跨越一个数量级：$\hat{}\rho_1 = 2873\,\mathrm{km}$（向 μ施加主要高斯过程载荷）至 $\hat{}\rho_4 = 188\,\mathrm{km}$（细尺度变异），这支持非可分离的协区域化线性模型 LMC 结构。最长范围的高斯过程捕捉了位置参数中的大尺度纬度梯度；相比在均值函数中明确加入地理协变量，我们更倾向于采用这一做法，因为前者无需指定跨数据源共享的协变量响应。对 $\mathbf{A}$ 施加的下三角约束消除了旋转不可识别性：各个 ρ_i 值取决于 $\boldsymbol{\theta}(s)$ 中组件的排序，但所有预测和跨来源相关性均不受这一选择影响。

图3 展示了 29 个测站处经克里金插值得到的 NOAA 尺度广义极值分布参数。与第 1 阶段最大似然估计（图 2）相比，克里金插值场在空间上更加平滑，反映了联合空间模型所带来的正则化。位置参数和形状参数与最大似然估计模式密切吻合；对数尺度参数与阶段 1 数值的偏差最小，这与中等程度的跨源相关性一致。

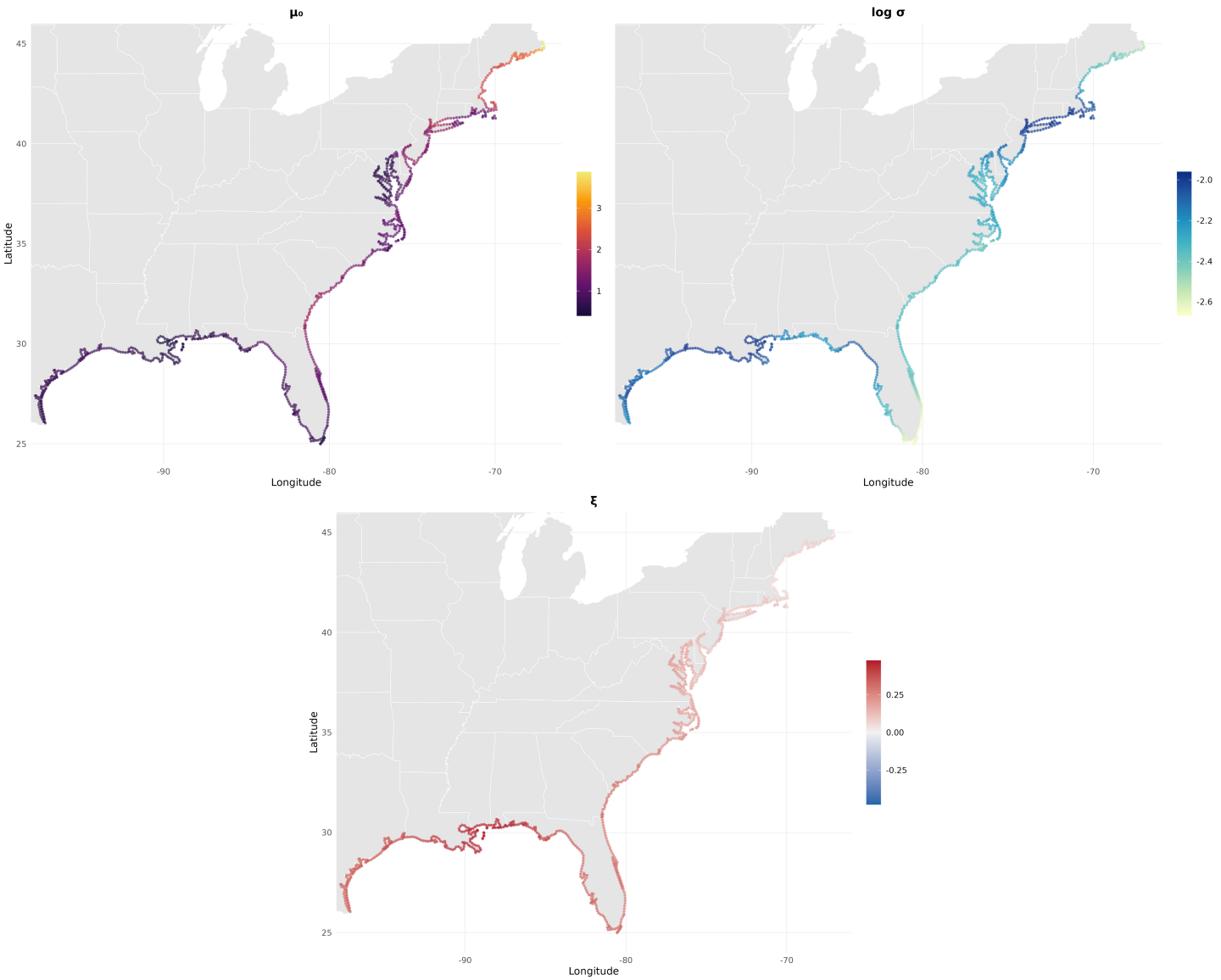

图3：来自联合模型、位于29个NOAA站点的克里金插值NOAA尺度广义极值分布参数。左：位置参数 μ_0 （2000年截距）。中：$\log \sigma$。右：形状参数 ξ，发散色标以零为中心。请与图 2中的第 1 阶段最大似然估计进行比较；由于借鉴了ADCIRC，克里金插值场在空间上更为平滑。

4.3 重现水平图与融合效益

沿15千米沿海网格经克里金插值得到的百年重现水平，从墨西哥湾的1.28米到新英格兰北部的4.29米不等（图4左），代表2000年参考条件。大西洋沿岸的标准误最低（0.10–0.19米），此处观测最为密集；墨西哥湾沿岸的标准误最高（0.16–0.54米），测潮仪较为稀疏且尾部更厚，二者共同加剧了不确定性（图4右）。

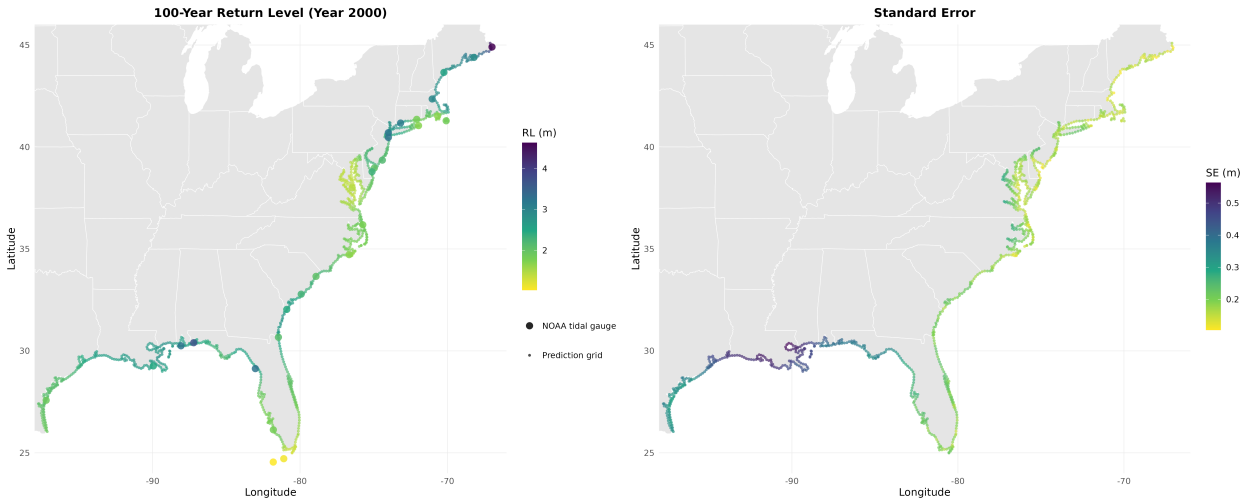

图4：左：美国东海岸与墨西哥湾沿岸在2000年参考条件下，经克里金插值得到的百年重现水平（高于平均海平面的米数）。大圆表示美国国家海洋和大气管理局站点级最大似然估计；小点表示15千米沿海网格上的联合模型预测。右：通过对克里金后验分布进行蒙特卡罗模拟得到的百年重现水平标准误（米）。

比较沿海地区的联合模型与仅NOAA模型（图5）可见，融合带来两项不同的优势，其相对重要性因地理位置而异。在大西洋沿岸，ADCIRC覆盖最密集，置信带显著变窄（标准误比中位数为1.34；图6左图）。在墨西哥湾沿岸，标准误比接近1（中位数为0.92），融合的作用则是使点估计上移，因为在测潮站覆盖范围稀疏的地区，ADCIRC重尾形状参数估计会将 ξ 拉高（图6右图）。百年重现水平标准误的比值（仅NOAA模型与联合模型之比）中位数为1.17，以下称为标准误比；这意味着在典型位置，联合模型的重现水平不确定性约为仅NOAA模型的85%。标准误比大于1表示融合降低了不确定性；约24%

海岸线上的标准误比超过 1.5。仅使用 ADCIRC 的模型估计值系统性地低于基于潮位站的估计。

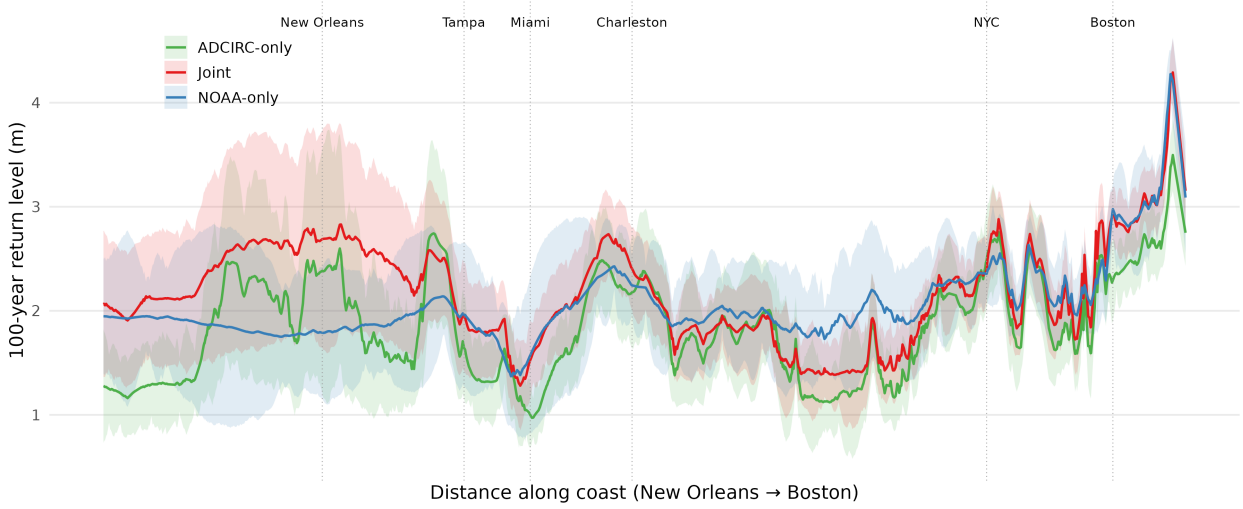

图5：新奥尔良至波士顿沿岸在2000年参考条件下的百年重现水平估计及其95%置信区间。沿大西洋沿岸，联合模型（红色）的不确定性小于仅 NOAA 模型（蓝色）；墨西哥湾沿岸的置信区间则相近。仅使用 ADCIRC 的模型（绿色）估计值系统性地偏低。

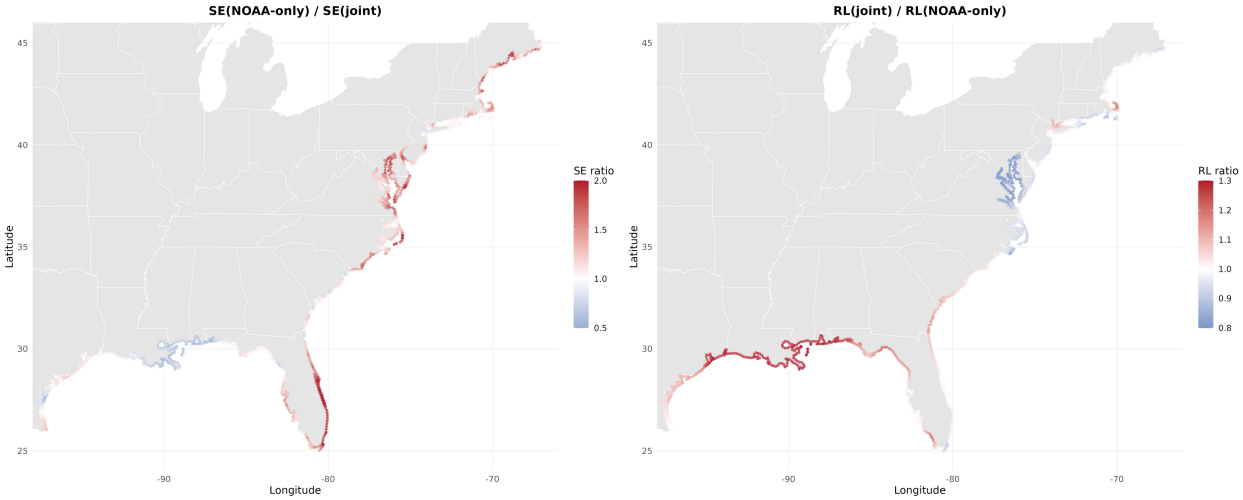

图6：左：百年重现水平标准误之比（2000年参考条件下；仅 NOAA 模型与联合模型之比）。数值大于1表示融合降低了不确定性。右：百年重现水平之比，联合模型与仅 NOAA 模型之比。不确定性降幅最大的是大西洋东南沿岸；点估计偏移最大的是墨西哥湾沿岸。

4.4 留一交叉验证

在 29 个 NOAA 站点进行留一交叉验证时，联合模型将百年重现水平的均方根误差降低了 35%（0.467 m 对比 0.724 m；表 2），评估基准为每个留出站点的第一阶段极大似然估计。改进主要体现在位置参数上，这与几乎完美的跨源相关性一致（$r = 0.995$）。形状参数的均方根误差小幅降低了 9%；对数尺度参数则没有变化（0.205 对比 0.205），这符合中等程度的尺度相关性。29 个站点的总对数预测密度增益为 $+32.4$；联合模型在 29 个站点中的 28 个上表现更优。

参数层面的散点图（图 S5）证实，与 1:1 线偏差最大的情况出现在 ξ，此处 ADCIRC 会使墨西哥湾沿岸的估计值偏高；另一个出现在 $\log\sigma$，这与中等程度的跨源相关性一致（$r = 0.44$）。

为单独评估各参数对重现水平改善的贡献，我们每次将仅 NOAA 的留一预测中的一个参数替换为联合模型的留一值，并重新计算重现水平（表3）。位置参数约占均方根误差降低幅度的一半（18.1%），形状参数进一步贡献了 10.5%，而对数尺度参数则没有贡献（-1.0%）。综合效应（35.5%）超过各项单独贡献之和，这是因为这些参数在重现水平公式（2）中存在非线性交互作用。

概率积分变换（PIT）值用于评估 LOO 预测分布的校准情况（表4；PIT 直方图见图 S4）。经验覆盖率达到或高于名义水平：预测区间比必要范围更宽，这符合预期，因为将 $\mathbf{W}$ 视为固定值会略微高估预测方差。μ 和重现水平的覆盖率最接近名义水平，而 ξ 和 $\log\sigma$ 即使在 50% 水平下也呈现过度覆盖（24/29 和 23/29，而预期值为 14.5）。柯尔莫哥洛夫-斯米尔诺夫检验均未拒绝原假设（$p = 0.96$、0.15、0.22、0.21，分别对应 μ、$\log\sigma$ ξ和重现水平），不过由于 $n = 29$，检验效能有限。

4.5 地理分块交叉验证

表5总结了地理分块交叉验证，其中依次留出五个相连的沿海区块，并重新拟合第二阶段模型。在全部29个站点中，联合模型将百年重现水平均方根误差降低了18%（0.665米，对比0.810米）。信息量最大的折来自墨西哥湾沿岸：移除全部四个测潮仪后，最近的保留美国国家海洋和大气管理局站点也在500多公里之外；此时，仅使用美国国家海洋和大气管理局数据的模型遇到近奇异的克里金协方差，预测能力下降，而联合模型借助 ADCIRC 覆盖范围维持了合理的预测。排除墨西哥湾沿岸的四个站点后，可单独考察其余25个站点的效果；联合模型将重现水平均方根误差降低了36%（0.496米，对比0.776米），证实即使在空间外推情况下，融合仍然有益。

对数尺度参数是唯一一个融合会增加误差的情况：仅 NOAA 的对数尺度均方根误差为 0.298，而联合模型为 0.406。鉴于跨源尺度相关性中等（$r = 0.44$），这一结果在预料之中：在空间外推时，借用与 NOAA 数值差异显著的 ADCIRC 尺度参数会引入噪声，而非提供信息。这一参数特定结果表明，实践者在假设融合能带来一致收益之前，应先检查分参数相关性。

4.6 敏感性分析

4.6.1 模拟网络密度

纳入约30个ADCIRC站点后，均方根误差降低幅度趋于饱和，接近35%；这大致相当于每个最小相关范围配置一个站点（在 $\sim$6000 公里海岸线上，$\hat{}\rho_{\min} = 188$ 公里；图7）。仅需15个站点，均方根误差便可降低20%。这一阈值反映了当前分析的空间分辨率：在6,000公里海岸线上，$\sim$30 个站点已足以完整刻画大尺度海岸线格局。若要分析更精细尺度的结构，则需要从ADCIRC网格中更密集地抽取站点，也可能会提高饱和阈值。这项饱和分析进一步支持了本研究采用的100站点ADCIRC网络：

在达到100个站点之前，融合的全部效益便已基本实现，这表明当前网络对于该项分析已绰绰有余。

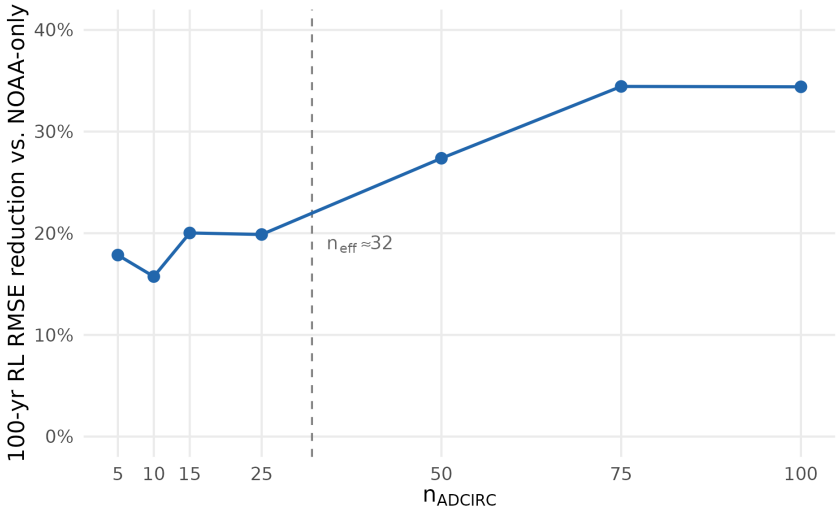

图7：纳入的 ADCIRC 站点数量与百年重现水平留一交叉验证均方根误差的降幅（相对于仅 NOAA）之间的关系。虚线标示 $n_{\text{eff}} \approx 32$，即每个最小相关范围（海岸线上 $\hat{\rho}_{\min} = 188\,\text{km}$/ $\sim$6000 km 约有一个站点），超过此范围后，在当前空间分辨率下，边际增益会逐渐减小。仅需 15 个站点，误差便可降低 20%。

4.6.2 渐消范围与分析者选择

在 $\lambda \in \{150, 300, 500\}$km 范围内，结果对渐消范围的变化不敏感（表6）：位置相关性稳定（$r \approx 0.995$），重现水平均方根误差几乎相同（0.463–0.467 m），对数预测密度增益的变化约为 10%。形状相关性略有变化（$r = 0.81$–0.90）。选择指数协方差是出于简约性考虑；129 个站点不足以有效区分光滑度。其他选择（重采样单位、多起点次数、网格分辨率）的影响也微乎其微。

4.6.3 自助法批次稳定性

将 $B = 2{,}000$ 个自助法重复抽样分成四个独立批次后，位置参数相关性基本保持不变（$r = 0.994$–0.996），形状相关性也很稳定（$r = 0.81$–0.85），

而重现水平的留一均方根误差介于 0.454 至 0.491 米之间。尺度相关性最不稳定（$r = 0.44$–0.78），这反映出似然面较为平坦；由于尺度参数对重现水平变化的贡献最小，这种不稳定性对预测的影响也微乎其微。不同批次的克里金重现水平图在视觉上无法区分。

4.6.4 非共址条件下的可识别性

为了评估在本研究的抽样设计下能否恢复跨源相关性，我们从拟合模型（真实的 $\boldsymbol{\beta}$、$\mathbf{A}$、$\boldsymbol{\rho}$）中生成了 100 个合成数据集，并从固定的自助法估计 $\mathbf{W}$ 中抽取测量误差（将阶段 1 不确定性视为已知），同时采用相同的部分观测结构（29 个站点观测参数 1–3，100 个站点观测参数 4–6）。对每个数据集重新拟合第二阶段模型（表7）。

位置相关性得到了精确恢复。对数尺度相关性存在噪声，但近似无偏。形状相关性的估计不确定性较高：报告的 0.84 点估计值位于恢复值的分布范围内，但对于 ξ，在 29 个源内观测值的情况下，该量的似然面实际上是平坦的。尽管存在这种不确定性，留一交叉验证的增益仍然稳健：均方根误差分解（表3）显示，ξ 的贡献（降幅 10.5%）次于估计精确的 μ相关性（18.1%）。

4.6.5 不确定性组成

报告的标准误采用蒙特卡罗模拟计算（从克里金后验分布中抽取 10,000 次），其中未纳入$\mathbf{W}$ 和第二阶段参数（$\mathbf{A}$ $\boldsymbol{\rho}$）的不确定性。作为比较，我们还根据式（12）计算了德尔塔法标准误。在留一交叉验证中，德尔塔法会大幅低估不确定性：留出整个站点会迫使模型进行远距离外推，从而产生较宽的克里金后验分布，而非线性重现水平函数（2）在该分布上的一阶泰勒展开近似效果不佳。在预测网格上，插值使克里金后验分布保持较窄，两种方法的结果十分接近（图S6）。

表4中观察到的经验性过度覆盖——预测区间在所有名义水平上都比所需范围更宽——间接表明，忽略$\mathbf{W}$ 和第二阶段参数的不确定性，并不会导致置信区间过窄。

5 讨论

近乎完美的跨源位置相关性（$r = 0.995$）带来了融合效益，相当于利用 100 个模拟站点提高了 NOAA 29 个站点水位计网络的密度。效益的地理分布格局（墨西哥湾沿岸的点估计偏移，与美国东南部大西洋沿岸的不确定性降低）反映了网络密度与尾部行为之间的相互作用。

5.1 与现有方法的比较

本研究在建模方面的主要贡献，是将 Russell 等人（2020）的单源线性模型协同定位扩展到多源情境。跨源协方差块为信息传递提供了机制，即使没有共址数据也能实现。选择矩阵在标准线性模型协同定位似然中处理部分观测结构，无需额外的分布假设。该建模形式无需修改结构，即可纳入更多数据源或共址观测。

与较简单的替代方法（例如使用 ADCIRC 站点进行最近邻偏差校正）相比，线性模型协同定位的主要优势在于，它能联合建模全部六个参数的空间相关性，并将不确定性连贯地传递到重现水平置信区间中。临时校正可以使点估计发生偏移，但无法生成校准预测区间，而这对于基于风险的决策至关重要。

重现水平均方根误差降低幅度是相对于仅使用由 29 个站点组成的水位计网络拟合的频率学派线性模型协同定位基线来衡量的。用于空间极值的完整贝叶斯分层模型（Huser & Wadsworth, 2019; Zhang et al., 2022, 2024）是主要的方

法论替代方案，它们在观测层面建模依赖关系，而不是通过潜在高斯过程建模。这些方法计算量很大，需要针对数百个相互耦合的参数进行马尔可夫链蒙特卡洛计算，并采用非共轭的广义极值似然；此外，它们仍然是参数模型，在这一尺度下很难验证其假设。相比之下，我们的两阶段结构使阶段 1 诊断可以直接验证，无需在两个阶段之间反复迭代，并能生成可直接解释的跨来源相关性。如何将完整贝叶斯空间极值模型适配到具有部分观测的多源融合场景，仍是一项开放挑战。

其他近期针对美国无验潮站沿海站点的极端海平面估计方法，包括拟合于潮位计年最大值的最大值稳定过程（Rashid 等，2024）以及时空贝叶斯层级框架（Morim 等，2025），完全依赖观测数据。本框架的独特之处在于将观测与模拟输出相融合；相对于我们的仅潮位计基线，重现水平均方根误差改善了35%，量化了这一收益。

5.2 模拟器偏差与模型差异

协区域化线性模型假定美国国家海洋和大气管理局与 ADCIRC 的广义极值分布参数是同一空间过程的实现，二者的差异来自共享潜在高斯过程上的来源特定载荷，而非系统性的空间偏差场。仅基于 ADCIRC 的重现水平系统性地低于基于潮位计的估计（图5）。在协区域化线性模型中，这一水平差异被吸收进 **A** 中的来源特定载荷，从而在没有单独偏差项的情况下产生较低的 ADCIRC μ 估计。

关键假设是，ADCIRC 的广义极值分布参数空间模式 与现实一致，即便其水平发生了偏移。近乎完美的位置相关性（$r = 0.995$）支持这一点：两个数据源按极端海平面对各个位置的排序完全相同。留一交叉验证提供了经验检验：若 ADCIRC 引入了空间结构化偏差，联合模型就会在靠近 ADCIRC 聚类的站点上出现系统性误判，但并未出现此类模式

并不明显。形式化差异模型（例如源特定加性高斯过程）需要共址观测，才能分别识别偏差场与跨源协方差；在此类数据可得之处，这一扩展将十分直接。

5.3 实践建议

在以下情况下，该框架最有益：（i）相对于空间相关变程，真值观测网较为稀疏；（ii）模拟数据源能够重现关注量的空间模式，即使其水平存在偏差；（iii）对于最难估计的参数（通常为 ξ），跨源相关性足够高。在我们的应用中，饱和分析（第 4.6 节）提供了一条实用经验法则：每个最小相关范围（$n_{\text{eff}} \approx L_{\text{domain}}/\rho_{\min}$，其中 L_{domain} 为区域长度）约配置一个模拟站点，即足以获得全部融合效益；超过这一数量后，边际收益会递减。仅使用 15 个 ADCIRC站点，就使重现水平均方根误差降低了 20%。

两个网络的空间交错程度比各自的密度更重要：源间距离应远小于 $\rho_{\min}$，才能识别跨源协方差块。实践者应先考察分参数相关性，再判断所有分量是否都能同等受益；对于相关性中等的对数尺度参数（$r = 0.44$），在空间外推时，借用模拟数据反而增加了预测误差，而非降低误差（第 4.5 节）。

除了沿海海平面之外，只要稀疏的地面实测观测与密集的模型输出并存，同一框架便适用：例如，利用雷达合成产品或再分析数据研究极端降水，或利用卫星反演进行空气质量监测。选择矩阵形式可自然扩展至包含共址观测、额外数据源，或不同数据源间参数维数不对称的情形。

5.4 局限性与未来工作

第二阶段假设第一阶段极大似然估计服从联合高斯分布。对于 $\hat{}\mu_0$ 和 $\log \hat{}\sigma$，若有 $40+$ 个年度 $\hat{}$ 极大值，这一假设得到渐近理论的有力支持。对于 ξ，在 $\xi = 0$ 附近收敛较慢（Coles，2001），但阶段 1 诊断证实其正态性足够良好（第 4.1 节）。

模拟研究（第 4.6.4 节）表明，在当前抽样设计下，跨来源形状相关性虽为正，但其估计具有较高的不确定性。应谨慎解读报告的 $r = 0.84$ 点估计；真实值可能显著更低。这并不影响重现水平增益，因为这些增益主要由估计精确的位置相关性驱动，但这限制了我们对跨来源尾部依赖得出结论的力度。

仍有若干方向有待探索。第二阶段的空间变化趋势将使未来参考年份的重现水平预测成为可能。中等的尺度相关性（$r = 0.44$）意味着ADCIRC对$\log \sigma$ 所提供的信息少于对 μ 或 ξ的信息，这很可能反映了两个数据源在该参数上的真实差异。本分析利用6,000公里范围内的100个ADCIRC站点解析了大尺度海岸线格局。ADCIRC在分辨率高得多的非结构化网格上运行；利用这一点可开展次区域分析，但需要对由此产生的高维协方差矩阵采用稀疏似然近似，并可能需要非平稳协方差结构来刻画空间变化依赖。即便在当前分辨率下，融合也能使重现水平改善35%，表明该框架能够以计算上可行的分析有效支持大尺度风险评估。完全的贝叶斯表述将消除自助法协方差$\mathbf{W}$ 并自动传播阶段 1 不确定性，但需要对约420个耦合参数、在非共轭广义极值似然下进行马尔可夫链蒙特卡洛，所得模型也将更难解释。

6 结论

我们提出了一种两阶段频率学框架，用于在空间极值分析中融合稀疏观测与密集模拟，并通过协区域化线性模型将来自两个数据源的广义极值分布参数建模为单一的六维空间过程。该方法应用于美国沿海海平面，并采用非平稳第一阶段模型，相对于仅测站模型，在2000年参考条件下将百年重现水平的留一交叉验证均方根误差降低了35%，且该收益在空间外推下仍然存在（分块交叉验证降幅为18%）。这一增益由近乎完美的跨源位置相关性和可观的形状相关性驱动，表现为测站密度最高处置信带更窄，以及测潮仪最稀疏处点估计得到校正；不过，中等相关的尺度参数表明，融合并非对所有GEV分量都普遍有益。该框架适用于任何稀疏真值观测与密集模型输出并存的领域。

补充信息

本文的更多信息和补充材料可在该期刊的站点在线获取。

数据可获取性声明

实现所有方法的 R 包 `evfuse`可在 `https://github.com/briannathanwhite/evfuse` 获取。该软件包包含全部 129个站点（29 个 NOAA 站点、100 个 ADCIRC 站点）经过处理的年度最大值，足以重现本文中的所有结果。原始的每小时海平面观测数据可从 NOAA 潮汐与洋流门户（`https://tidesandcurrents.noaa.gov`）公开获取；原始 ADCIRC模拟输出由北卡罗来纳大学

教堂山分校的文艺复兴计算研究所（RENCI）生成，可向作者索取。软件包仓库中包含用于重现所有分析的脚本。

致谢

ADCIRC模拟数据由北卡罗来纳大学教堂山分校的文艺复兴计算研究所（RENCI）提供。本研究的初步版本曾在北卡罗来纳州教堂山举行的第14届极值分析国际会议（EVA 2025）上发表。

资助

本研究未获得外部资助。

利益冲突

作者声明不存在利益冲突。

参考文献

Berliner, L.M. (2003). 地球物理学中的物理统计建模。《地球物理研究杂志：大气》，108(D24), 8776.

Coles, S. (20 01).《极值统计建模导论》. 施普林格.

Cooley, D., Nychka, D., & Naveau, P.(2007). 贝叶斯空间建模极端降水重现水平。《美国统计协会会刊》, 102(479), 824–840.

Cressie, N. & Wikle, C.K. (2011).《时空数据统计学》. 威立。

Davison, A.C., Padoan, S.A., & Ribatet, M. (2012). 空间极值的统计建模。《统计科学》，27(2)，161–186。

Dietrich, J.C., Zijlema, M., Westerink, J.J., 等（2011）。采用整体耦合、可扩展计算对飓风波浪和风暴潮进行建模。《海岸工程》，58(1)，45–65。

Finley, A.O., Banerjee, S., & Gelfand, A.E.（2008）。空间变化协方差的贝叶斯动态建模。《农业、生物与环境统计学杂志》，13(1)，60–83。

Furrer, R., Genton, M.G., & Nychka, D.（2006）。用于大型空间数据集插值的协方差锥化。《计算与图形统计学杂志》，15(3)，502–523。

Gelfand, A.E., Schmidt, A.M., Banerjee, S., & Sirmans, C.F.（2004）。通过空间变化协同区域化对非平稳多变量过程进行建模。《Test》，13(2)，263–312。

Huser, R. & Wadsworth, J.L. (2019). 极值相依类别未知时的空间过程建模。《美国统计协会会刊》，114(525)，434–444。

Journel, A.G. & Huijbregts, C.J. (1978). 《矿业地质统计学》。学术出版社。

Kendall, M.G. (1975). 秩相关方法。Griffin，第4版。

Luettich, R.A., Westerink, J.J., & Scheffner, N.W. (1992). ADCIRC：一种先进的陆架、海岸和河口三维环流模型。报告1：ADCIRC-2DDI和ADCIRC-3DL的理论与方法。疏浚研究计划技术报告DRP-92-6，美国陆军工程兵团航道实验站。

Mann, H.B. (1945). 针对趋势的非参数检验。《计量经济学》，13(3)，245–259。

Morim, J., Rasmussen, D.J., Wahl, T., Calafat, F.M., Kopp, R.E., Oppenheimer, M., & Dangendorf, S. (2025). US-CoastEX：基于观测的美国风暴潮和海平面极值概率再分析。《科学数据》, 12, 1395.

Nguyen, H., Cressie, N., & Braverman, A. (2012). 面向遥感应用的空间统计数据融合。《美国统计协会会刊》, 107(499), 1004–1018.

Rashid, M.M., Moftakhari, H., & Moradkhani, H. (2024). 使用最大稳定过程对美国本土海岸线沿线的风暴潮极值进行随机模拟。《*Communications Earth & Environment*》, 5, 39.

Rasmussen, C.E. & Williams, C.K.I. (2006). 《机器学习中的高斯过程》。麻省理工学院出版社。

Russell, B.T., Risser, M.D., Smith, R.L., & Kunkel, K.E. (2020). 探究春末墨西哥湾海表温度与美国墨西哥湾沿岸降水极值之间的关联，重点关注哈维飓风。《*Environ metrics*》, 31(2),e2595.

Sang, H. & Gelfand, A.E. (2010). 用于空间极值的连续空间过程模型。《农业、生物与环境统计学杂志》, 15(1), 49–65.

Sen, P.K.（1968）。基于肯德尔 τ 的回归系数估计。《美国统计协会会刊》，63(324), 1379–1389。

Stein, M.L.（1999）。《空间数据插值：克里金法理论》。施普林格。

Wackernagel, H.（2003）。《多元地统计学：应用导论》。施普林格，第3版。

Wendland, H.（1995）。最小次数的分段多项式正定紧支撑径向函数。《计算数学进展》，4(1)，389–396。

Zhang, L.、Shaby, B.A.和Wadsworth, J.L.（2022）。用于在包含大量位置的数据集上灵活建模空间极值的分层变换尺度混合。《美国统计协会会刊》，117(539)，1357–1369。

Zhang, L.、Risser, M.D.、Wehner, M.F.和O’Brien, T.A.（2024）。美国本土降水数据的贝叶斯空间极值分析。《农业、生物与环境统计学杂志》。

表1：数据源之间的趋势诊断比较。

	美国国家海洋和大气管理局（观测值）	ADCIRC（模拟值）
测试的站点	29	100
显著（MK $p < 0.05$）	19（66%）	1（1%）
Sen斜率中位数（毫米/年）	4.6	0.4
斜率范围（毫米/年）	2.0 至 13.7	-2.2 至 5.3

表 2：Leave-o 在 29 个 NOAA站点进行留出交叉验证（非平稳第一阶段模型）。

指标	联合	仅 NOAA	降幅
μ_0 留一均方根误差	0.142	0.434	67.3%
$\log\sigma$ 留一均方根误差	0.205	0.205	0.0%
ξ 留一均方根误差	0.147	0.161	8.7%
100 年重现水平 留一均方根误差（m）	0.467	0.724	35.5%
总对数预测密度	$+23.84$	-8.53	
站点联合模型胜出（对数预测密度）		28 / 29（97%）	

表3：百年重现水平留一交叉验证均方根误差降低幅度的分解。每一行“混合”方法仅将指定参数替换为联合模型的留一预测，其他参数则保持为仅 NOAA 模型的取值。

方法	重现水平均方根误差（米）	降幅（%）
仅 NOAA（所有参数）	0.724	–
混合模型：μ_0 来自联合模型	0.594	18.1
混合模型：ξ 来自联合模型	0.648	10.5
混合模型：$\log\sigma$ 来自联合模型	0.732	-1.0
联合（所有参数）	0.467	35.5

表4：29 个 NOAA 站点的留一预测区间经验覆盖率（联合模型）。接近名义水平的数值表示不确定性校准良好。KS p 值用于检验概率积分变换的均匀性。

参数	50%	80%	95%	KS p
μ	17/29	23/29	28/29	0.962
日志 σ	23/29	27/29	28/29	0.145
ξ	24/29	28/29	29/29	0.220
100 年重现水平	19/29	26/29	29/29	0.212
预期	*14.5*	*23.2*	*27.6*	

表5：29个NOAA站点的地理分块交叉验证。依次留出五个连续的沿海区块，并针对每一折重新拟合第二阶段。“不含墨西哥湾”这一行排除了墨西哥湾沿岸的四个站点；在这些站点上，仅NOAA模型遇到了近奇异的克里金协方差，从而可以单独考察其对其余25个站点的影响。

指标	联合	仅 NOAA	降幅
μ_0 均方根误差	0.255	0.781	67.3%
$\log\sigma$ 均方根误差	0.406	0.298	−36.2%
ξ 均方根误差	0.194	0.213	8.9%
100 年重现水平均方根误差，全部 29（m）	0.665	0.810	17.9%
100年重现水平RMSE（不含墨西哥湾）（m）	0.496	0.776	36.1%
站点联合模型胜出（对数预测密度）		24 / 29 (83%)	

表6：关键结果对 Wendland截断函数范围 λ （km）的敏感性。

	$\lambda = 150$	$\lambda = 300$	$\lambda = 500$
Cor(μ_N， μ_A)	0.995	0.995	0.994
相关系数(log σ_N，log σ_A)	0.381	0.443	0.405
相关系数(ξ_N， ξ_A)	0.899	0.837	0.812
100 年重现水平重现水平均方根误差，联合（米）	0.463	0.467	0.467
重现水平均方根误差降低幅度（%）	36.1	35.5	35.5
对数预测密度增益（联合 −美国国家海洋和大气管理局）	30.7	32.4	33.9

表7：在观测抽样设计（无共同定位）下，从拟合模型生成的 100 个模拟数据集中恢复跨来源相关性的情况。

参数	True	中位数	SD	IQR
Cor(μ_N， μ_A)	0.995	0.989	0.014	0.014
Cor(log σ_N, log σ_A)	0.443	0.466	0.329	0.497
Cor(ξ_N， ξ_A)	0.837	0.448	0.451	0.695