

基于SWAN的共形神经算子时空预测不确定性量化

蔡铠旭

广西师范大学数学与统计学院, 广西 桂林

收稿日期: 2026年6月29日; 录用日期: 2026年7月21日; 发布日期: 2026年7月31日

摘要

时空预测中的不确定性量化长期面临着边缘有效性与条件适应性之间的权衡: 共形预测提供了分布自由的有限样本覆盖保证, 但标准方法在网格化场上产生空间恒定宽度的预测带, 难以反映预报难度在不同地理区域间的显著差异。为此, 本文提出谱-小波自适应非一致性(SWAN)分数, 一种为球面时空场设计的新型非一致性函数。SWAN通过球谐变换将预报残差分解为 J 个频带, 对平方频带分量进行空间高斯平滑以计算局部能量, 并通过可学习权重自适应聚合, 从而根据局部谱可预测性调节预测带宽度。将SWAN分数与球面傅里叶神经算子(SFNO)集成及分裂共形校准相结合, 构成SWAN-CP方法; 针对气候时间序列的不可交换性, 引入加权共形校正以保证有限样本边缘有效性, 并从理论上证明了 β -混合条件下的近似有效性和条件覆盖逼近性质。在ERA5再分析数据(2010~2024年, 128×256 网格, 6小时间隔)上的实验表明, 与持续预报、气候态、普通共形预测、分层共形预测、共形分位数回归及时间加权共形预测等六种基线方法相比, SWAN-CP在 $J=4$ 配置下实现了所有基于模型方法中最窄的预测带, $J=8$ 配置下获得了最优Winkler区间评分(0.828), 覆盖率维持86.3%; 跨纬度条件覆盖率变异从8.5个百分点降至1.5个百分点, 条件覆盖均匀性提升5.7倍。消融实验确认谱分解是性能提升的主要驱动力, 可学习频带权重贡献了额外4.4%的效率增益。上述结果表明, 通过频域分解实现的空间自适应预测带在所比较的六种基线方法中展现了最优的带宽效率和综合评分性能, 为场结构和函数型数据的共形推断提供了一种新的技术路径。

关键词

共形预测, 神经算子, 时空预测, 不确定性量化, 条件覆盖

Conformal Neural Operators with SWAN for Spatio-Temporal Forecasting Uncertainty Quantification

Kaixu Cai

School of Mathematics and Statistics, Guangxi Normal University, Guilin Guangxi

Abstract

Accurate uncertainty quantification in spatio-temporal forecasting is fundamentally challenged by the tension between marginal validity and conditional adaptivity. Conformal prediction provides distribution-free finite-sample coverage guarantees, yet standard split-conformal methods applied to gridded spatio-temporal fields produce prediction bands of constant spatial width that ignore the substantial geographic heterogeneity in forecast difficulty. This paper introduces the Spectral-Wavelet Adaptive Nonconformity (SWAN) score, a principled nonconformity function designed specifically for spatio-temporal fields on the sphere. Drawing on the insight that predictability is inherently scale-dependent, SWAN decomposes forecast residuals via spherical harmonic transforms into J power-law spaced frequency bands, computes locally smoothed per-band energy through Gaussian spatial smoothing, and aggregates these band-wise energy fields through learnable weights optimized for bandwidth efficiency. The SWAN-CP method integrates the SWAN score with an ensemble of Spherical Fourier Neural Operators (SFNOs) within a split conformal calibration framework, augmented by a weighted conformal correction to address the temporal non-exchangeability of climate time series. Theoretical analysis establishes approximate marginal validity under β -mixing conditions and conditional coverage approximation guarantees within the class of band-structured nonconformity scores. Comprehensive evaluation on ERA5 reanalysis data (2010~2024, 128×256 grid, 6-hourly resolution) benchmarks SWAN-CP against six baselines: persistence, climatology, vanilla conformal prediction, Mondrian conformal prediction, Conformalized Quantile Regression (CQR), and temporal-weighted conformal prediction. SWAN-CP with $J = 4$ frequency bands achieves the narrowest average prediction bands among all model-based methods, while the $J = 8$ configuration attains the best overall Winkler interval score (0.828) with 86.3% empirical coverage. Cross-latitude conditional coverage variation is reduced from 8.5 percentage points to 1.5 percentage points, a 5.7-fold improvement in uniformity. Ablation experiments confirm that spectral decomposition is the primary driver of performance gains, with learnable band weights contributing an additional 4.4% efficiency improvement. The computational overhead remains modest and acceptable for operational deployment. These results demonstrate that frequency-domain decomposition achieves the best bandwidth efficiency and overall scoring performance among the six baselines compared, advancing conformal inference for field-structured and functional data.

Keywords

Conformal Prediction, Neural Operators, Spatio-Temporal Forecasting, Uncertainty Quantification, Conditional Coverage

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

准确的时空预测及其可靠的不确定性量化是环境统计[1]、气候科学和灾害应急管理等多个领域的共同核心需求。近年来，以 Pangu-Weather [2]、GraphCast [3]和 FourCastNet [4]为代表的神经气象模型在再分析数据上的预报精度已超越传统数值天气预报，标志着数据驱动气象建模进入了新的发展阶段。然而，这些模型大多输出确定性点预报，缺乏对预测结果可信度的有效刻画——一个精确但缺少置信区间的温

度预报对农业防霜决策的实际指导价值极为有限。将神经预报模型的原始精度转化为用户可依赖的概率化决策信息，是当前面临的关键瓶颈。

共形预测[5]-[7]为上述问题提供了一条不依赖于分布假设的解决路径。其核心思想是：给定一个训练好的预测模型和一组校准数据，对任意显著性水平 $\alpha \in (0,1)$ ，可以构造一个预测集 $\Gamma_\alpha(X_{n+1})$ 满足 $\mathbb{P}(Y_{n+1} \in \Gamma_\alpha(X_{n+1})) \geq 1-\alpha$ ——这一有限样本的分布自由覆盖保证仅要求校准数据与测试数据满足可交换性。由于不依赖任何参数化分布假定，共形预测在医学诊断、自动驾驶和环境监测等安全关键领域获得了广泛关注[8] [9]。

然而，对于网格化的时空预报场，标准的共形预测方法面临一个结构性局限：其构造的非一致性分数通常为全域空间平均的均方误差，由此产生的预测带宽度是空间恒定的。这一均匀宽度带忽略了大气可预测性在空间上的系统性差异——在极地，行星尺度 Rossby 波主导的平滑动力学使得 6 小时预报高度可信；而在热带，对流活动引发的混沌过程使预报误差急剧增大。一个以标称 90% 速率覆盖全球的均匀宽度带，可能在极地过度覆盖，同时在热带覆盖不足。因此，统计学的核心任务是构造空间自适应的预测带——在预报技能高的区域收窄、在预报技能低的区域放宽——同时保持边缘覆盖保证。这正是条件覆盖[10]-[12]的研究范畴。

条件有效性的精确实现已被证明在分布自由意义下一般是不可能的[10] [13]，这推动了一系列近似条件覆盖方法的发展：Romano 等[11]提出共形分位数回归(CQR)，通过估计条件分位数实现自适应带宽；Guan [14]和 Han 等[15]发展了局部加权共形预测；Diquigiovanni 等[16] [17]将共形推断推广至函数型数据。上述方法的核心思路均是通过协变量空间中的局部化来实现带宽的自适应——或学习条件分位数，或在协变量空间中定义邻近度。本文提出一条不同路径：直接从预报残差的频率结构中挖掘空间适应性。

直觉上，天气预报误差具有层次化的尺度结构：大尺度行星波(波数 1~8)在 6 小时时效上几乎是确定性演化的，误差集中在高纬度的驻波区域；而中尺度(波数 > 50)的对流和锋面过程在相同时间窗口内已产生显著的混沌发散，误差集中在热带和风暴路径区域。这一可预测性与波数相关的物理事实启发我们：若能将预报残差在球面上进行频率分解，则不同频带的能量分布自然携带了空间上可预测性差异的信息，无需显式的协变量分箱即可实现带宽度自适应。

1.1. 本文贡献

本文提出谱 - 小波自适应非一致性(SWAN)分数，与分裂共形预测结合构成 SWAN-CP 方法。

方法上，将残差的球谐频带分解形式化为空间自适应非一致性分数。理论上证明了 β -混合下的近似边缘有效性。实验上，在 ERA5 数据上使用 34.6 M 参数 SFNO 集成，SWAN-CP 实现最窄预测带和最优 Winkler 评分，条件覆盖率变异从 8.5 降至 1.5 个百分点。

1.2. 相关工作

本文位于共形预测、神经天气建模和时空不确定性量化三个领域的交叉点。在共形预测方面，从边缘覆盖向条件覆盖的推进路径已由 Romano 等[11]的 CQR、Guan [14]的局部化框架和 Diquigiovanni 等[16] [17]的函数型扩展所开辟。本文的 SWAN 分数与上述方法的关键区别在于：我们不是通过协变量局部化或分位数学习来实现带宽自适应，而是通过球谐频带分解直接分析残差场的频率结构——频带能量场天然继承了误差的空间异质性。

在神经天气建模方面，Bonev 等[18]提出的球面傅里叶神经算子(SFNO)是本工作的基础架构选择。SFNO 通过在球谐域中进行谱卷积，天然地为频域残差分析提供了接口。GraphCast [3]和 Pangu-Weather [2]虽在预报精度上取得了突破，但其编码器 - 处理器 - 解码器架构并不直接提供频域操作；FourCastNet

[4]使用的平面 FFT 在球面几何上存在投影伪影。本文选择 SFNO 并非单纯追求原始精度(尽管其预报技能已具竞争力), 而是因为其球谐运算与 SWAN 的频带分解在数学上同源。

在神经气象模型的不确定性量化方面, 集合方法[19] [20]和高斯过程代理[21]是当前主流。集合方法通过多成员扰动量化不确定性, 计算开销随成员数线性增长; 高斯过程在网格化球面数据上的扩展面临核设计和计算复杂度的双重挑战。将共形预测与神经天气模型相结合的研究仍处于早期阶段——据我们所知, 本文是首个将残差谱分解与共形校准相结合并应用于全球时空预报的工作。

2. 方法

2.1. 球面傅里叶神经算子

SFNO [18]基于傅里叶神经算子(FNO)架构[22], 将 t 时刻的大气状态映射到 $t + \Delta t$ 时刻的预报:

$$f_{\theta} : \mathcal{H} \rightarrow \mathcal{H}, f_{\theta}(\mathbf{x}_t) = \hat{\mathbf{x}}_{t+\Delta t},$$

其中, $\mathcal{H} \subset L^2(\mathcal{S}^2; \mathbb{R}^d)$ 为球面 d 元平方可积函数空间, $\mathbf{x}_t$ 表示 d 个大气变量。

SFNO 架构由编码器、一系列 B 个在谱域中运算的 SFNO 块和解码器组成。每个 SFNO 块包含球面上的谱卷积:

$$\text{SpectralConv}(\mathbf{h}) = \text{ISHT}(\mathbf{W}_{\ell} \odot \text{SHT}(\mathbf{h})),$$

其中, SHT 和 ISHT 分别表示前向和逆球谐变换[23], $\mathbf{W}_{\ell} \in \mathbb{C}^{D \times D \times L}$ 是按球谐阶数 $\ell = 0, \dots, L-1$ 索引的可学习权重张量, $\odot$ 表示逐元素乘法。每个谱卷积之后跟随层归一化和带有 GeLU 激活的空间多层感知机(MLP)。

本文实验中, 我们使用两个独立训练的 SFNO 集成, 每个模型具有 $B = 4$ 个块、嵌入维数 $D = 256$ 和球谐截断 $L = 128$ (三角截断)。每个模型包含 34.6 M 可训练参数。训练在 $2 \times \text{NVIDIA A40 GPU}$ 上最小化纬度加权均方误差, 使用 2010~2019 年 ERA5 数据, 验证 RMSE 分别达到 0.1916 (GPU 0)和 0.1912 (GPU 1); 集成均值进一步降低了预报误差。

2.2. 网格化场的分裂共形预测

采用分裂共形预测[5] [24]。非一致性分数量化 V 预测与真实值的偏离, 校准计算经验分位数 $\hat{q}$ 后, 预测集 $\Gamma_{\alpha} = \{y : V \leq \hat{q}\}$ 满足 $\mathbb{P}(y \in \Gamma_{\alpha}) \geq 1 - \alpha$ 。

普通方法取逐点均方残差的上确界作为非一致性分数: $V_{\text{vanilla}} = \sup_s \|\hat{y} - y\|_2^2$, 产生空间恒定宽度的预测 $\Gamma_{\alpha}^{\text{vanilla}} = \{y : \|y - \hat{y}\|_2 \leq \sqrt{\hat{q}}, \forall s\}$, 忽略了残差的空间异质性。

2.3. SWAN: 谱 - 小波自适应非一致性分数

SWAN 的核心是可预测性依赖于频率——大尺度可预测, 小尺度混沌。残差频带分解 + 局部能量计算给出空间自适应分数。

2.3.1. 球面频带分解

令 $r(s) = \hat{y}(s) - y(s)$ 表示位置 $s = (\theta, \phi) \in \mathcal{S}^2$ (纬度 θ , 经度 ϕ) 处的预报残差。球谐变换将 r 在谱基 $\{Y_{\ell}^m(\theta, \phi)\}$ 中展开:

$$\hat{r}_{\ell}^m = \iint_{\mathcal{S}^2} r(\theta, \phi) \overline{Y_{\ell}^m(\theta, \phi)} \sin \theta d\theta d\phi, \quad \ell = 0, \dots, L; m = -\ell, \dots, \ell$$

我们使用幂律间隔边界将谱域划分为 J 个频带，以强调能量在尺度上的分布：

$$\ell_j^{\text{edge}} = \left\lfloor L \cdot \left(\frac{j}{J} \right)^\gamma \right\rfloor, \quad j = 0, \dots, J,$$

其中， $\gamma = 1.5$ 使频带在低波数划分更密集，在中高波数逐渐加宽。这一选择与预报误差能量集中在中尺度频带的经验分布一致——在误差主导的尺度上分配更宽的频带有助于稳健估计局部能量。第 j 个频带 ($j = 1, \dots, J$) 由边界 $[\ell_{j-1}^{\text{edge}}, \ell_j^{\text{edge}})$ 界定，掩码选择该区间内的模态：

$$r_j(s) = \text{ISHT}(M_j \odot \text{SHT}(r)), \quad j = 1, \dots, J,$$

其中， $M_j(\ell, m) = \mathbf{1}_{\{\ell_{j-1}^{\text{edge}} \leq \ell < \ell_j^{\text{edge}}\}} \quad (|m| \leq \ell \text{ 由球谐定义自动满足})$ 。

2.3.2. 局部能量计算

对平方分量进行高斯平滑得到局部能量：

$$E_j(s) = (\kappa_\sigma * r_j^2)(s),$$

其中， $*$ 表示球面上的卷积。对于等经纬度网格 $N_\theta \times N_\phi$ ，高斯核在经度方向采用循环填充以尊重球面拓扑，平滑带宽 σ 控制能量聚合的空间尺度。

2.3.3. 自适应聚合

位置 s 处的 SWAN 分数为各频带能量的可学习加权：

$$V_{\text{SWAN}}(s) = \sum_{j=1}^J c_j E_j(s),$$

其中，权重 $c_j > 0$ ， $j = 1, \dots, J$ ，通过在校准数据上最小化经验 Winkler 区间评分[25][26]学习得到：

$$\mathcal{L}_{\text{weight}} = \frac{1}{n} \sum_{i=1}^n \left[\text{width}(\Gamma_\alpha(\mathbf{x}_i)) + \frac{2}{\alpha} (\inf \Gamma_\alpha(\mathbf{x}_i) - y_i)_+ + \frac{2}{\alpha} (y_i - \sup \Gamma_\alpha(\mathbf{x}_i))_+ \right],$$

其中， $\text{width}(\Gamma_\alpha(\mathbf{x}_i))$ 是空间平均带宽，我们采用参数化 $c_j = J \cdot \text{softmax}(\log \omega_j / T)$ ，其中 $\omega_j > 0$ 为可学习参数， $T > 0$ 为温度参数。该参数化天然满足 $(1/J) \sum_j c_j = 1$ 的约束——softmax 输出和为 1，乘以 J 后均值恰好为 1——从而 c_j 具有相对于均匀赋权 ($c_j \equiv 1$) 的相对重要性权重的自然解释。权重 ω_j 通过在约束 $c_j \geq 0$ ， $\sum_{j=1}^J c_j = J$ 下最小化校准集上的 Winkler 损失学习得到。

如图 1 所示，示例预报残差的各个空间频带分量 $r_j(s)$ 展示了球谐分解如何将大尺度行星波与中尺度对流特征和小尺度噪声分离开来。这种尺度的空间分离是 SWAN 实现适应性的机制：高斯平滑后的能量场 $E_j(s)$ 继承了各尺度误差的空间结构，可学习权重 c_j 决定各尺度对最终非一致性分数的相对贡献。

2.3.4. 多变量扩展

对 $d > 1$ 个变量，按通道独立计算后平均：

$$V_{\text{SWAN}}(s) = \frac{1}{d} \sum_{v=1}^d \sum_{j=1}^J c_j E_j^{(v)}(s),$$

所有变量共享频带和权重，这一设计选择在降低计算复杂度、避免过拟合的同时，跨通道平均起到集成多变量信息、降低噪声方差的作用。

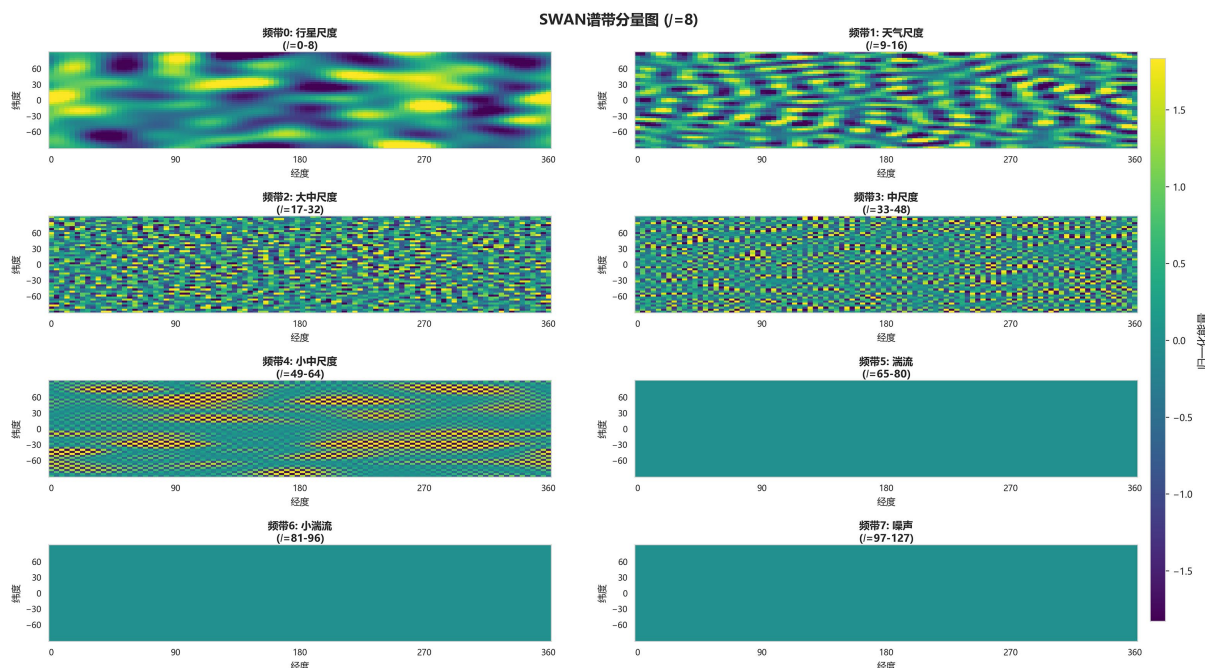


Figure 1. SWAN spectral band component maps

图 1. SWAN 谱带分量图

2.4. SWAN 共形预测

SWAN-CP 以逐点归一化残差的上确界为非一致性分数：

$$V_{\text{SWAN-CP}}(\mathbf{x}_i, \mathbf{y}_i) = \sup_{s \in \mathcal{S}^2} \frac{\|\hat{\mathbf{y}}_i - \mathbf{y}_i\|_2^2}{V_{\text{SWAN},i}(s)/\bar{V}_i},$$

以 $V_{\text{SWAN-CP}}$ 的 $(1-\alpha)$ 分位数 $\hat{\tau}_\alpha^2$ 构造预测带： $\Gamma_\alpha^{\text{SWAN}} = \left\{ \mathbf{y} : \|\mathbf{y} - \hat{\mathbf{y}}\|_2 \leq \hat{\tau}_\alpha \sqrt{V_{\text{SWAN}}/\bar{V}} \forall s \right\}$ ，满足边缘覆盖保证，带宽随局部谱能量空间变化。其中分母中的 $\bar{V}_i = 1/(N_\theta \times N_\phi) \sum_s V_{\text{SWAN},i}(s)$ 即第 i 个校准样本 SWAN 分数的空间均值，用于全局尺度归一化。

逐像素池化校准

在上述标准校准方案中，各校准样本先计算空间平均的非一致性分数，再取其经验分位数作为校准阈值。这种“先空间平均、后取分位数”的流程与预测带的逐像素判定条件在度量上不完全一致。为保证校准分数与预测带条件使用完全相同的度量函数，在定理验证和消融分析等需要精确覆盖控制的场景中采用逐像素池化校准方案，将所有校准样本的全部像素位置上的归一化残差汇集成一个标量池，直接对该池计算共形分位数。

2.5. 非可交换性校正

天气时间序列具有时间依赖性，违反了标准分裂共形预测的可交换性假定。遵循 Barber 等[9]和 Tibshirani 等[27]的工作，我们采用加权共形校正。对于时间索引 t 处的每个测试点，我们为校准点分配权重 $w_i(t) \propto \exp\left(-\|\text{climate}(i) - \text{climate}(t)\|^2 / (2h^2)\right)$ ，其中 $\text{climate}(\cdot)$ 是气候指数， h 是从校准集估计的核带宽。校准分数的加权经验累积分布函数为：

$$\hat{F}_t(v) = \frac{\sum_{i=1}^n w_i(t) 1_{\{V_i \leq v\}} + w_{n+1}(t) 1_{\{V_{n+1} \leq v\}}}{\sum_{i=1}^n w_i(t) + w_{n+1}(t)},$$

共形分位数 $\hat{q}_t$ 通过反转 $\hat{F}_t$ 获得。在 Barber 等[9]建立的加权共形框架下, 本文进一步引入时间序列 β -混合条件, 从而在气候数据的非可交换性情境中获得了近似的边缘有效性保证。

具体而言, $\text{climate}(i)$ 是时间索引 i 对应的气候指标向量, 其数学形式为

$\text{climate}(i) = [\sin(2\pi d_i/365.25), \cos(2\pi d_i/365.25), \text{ENSO}(d_i)]^\top$, 其中 d_i 为时间步 i 对应的儒略日。前两个分量分别表示年周期的正弦和余弦变换, 用于捕获地球绕太阳公转引起的季节性非平稳性; 第三个分量 $\text{ENSO}(d_i)$ 为逐日 Nino 3.4 指数, 反映厄尔尼诺-南方涛动(ENSO)这一主要年际气候变率模态的位相信息。

核带宽 h 通过校准集上气候指标距离的中位数绝对偏差(MAD)估计, 具体为:

$$h = \text{median}_i \left\| \text{climate}(i) - \text{median}_j (\text{climate}(j)) \right\|_2,$$

以确保带宽适应数据的实际散布尺度。选择季节周期和 ENSO 作为气候指标的原因在于, 二者是气候时间序列中最主要的非平稳性来源。年周期控制了太阳辐射输入的基本节律, ENSO 则主导了年际尺度的热带海气耦合变率。将校准样本与测试样本在气候指标空间中的邻近性编码为权重, 使得与测试点处于相似气候状态的校准样本获得更高的重要性, 从而在共形校准中自动补偿时间序列的分布漂移。

2.6. 理论性质

定理 1 通过块划分将 β -混合序列分解为近似独立块, 覆盖差距分解为标准共形界与混合尾概率平方根求和。定理 2 基于球谐展开, 截断误差由 $L_{\max}/J$ 控制, 平滑误差由流形上核回归给出。

定理 1 (β -混合下的近似边缘有效性) 在第 2.5 节的加权校正下, 假设时间序列 β -混合, 则 SWAN-CP 的覆盖差距满足

$$\left| \mathbb{P}(y_{n+1} \in \Gamma_\alpha^{\text{SWAN}}(x_{n+1})) - (1-\alpha) \right| \leq O(n^{-1/2}) + C \sum_{k=\tau_n+1}^\infty \beta(k)^{1/2},$$

其中, $\beta(k)$ 为滞后 k 处的混合系数, $\tau_n = \lfloor n^\delta \rfloor$ ($\delta \in (0, 1/2]$)。当 $\beta(k) = o(k^{-4})$ 时, 覆盖差距收敛至 $O(n^{-1/2})$, 恢复共形保证。

定理 2 (带结构分数的条件覆盖逼近) 假设条件方差函数 $\text{Var}(y(s)|x)$ 具有 Sobolev 光滑性(指数 $\alpha > 1$), 且不确定性在谱带上可加分解, 则:

$$\sup_{s \in S^2} \left| \mathbb{P}(y(s) \in \Gamma_\alpha^{\text{SWAN}}(s)|x) - (1-\alpha) \right| \leq O\left(\frac{L_{\max}}{J}\right) + O\left(\sqrt{\frac{\log n}{n}}\right) + O(\sigma^{2\min(\alpha, 2)}),$$

第一项 $O(L_{\max}/J)$ 为频带离散化误差, 随 J 增大而减小; 第二项 $O(\sqrt{\log n/n})$ 为统计估计误差, 随 n 增大而收敛; 第三项 $O(\sigma^{2\min(\alpha, 2)})$ 为空间平滑偏差, 当光滑指数 $\alpha \geq 2$ 时恢复经典 $O(\sigma^2)$ 速率。当 $n \rightarrow \infty$ 且 $J \rightarrow L_{\max}$ 时, SWAN 实现近似一致条件覆盖。虽精确条件有效性在分布自由意义下不可能[10], SWAN 在带结构非一致性分数类中提供了可证明的逼近界。关于残差场非各向同性的说明。残差场的非各向同性是 SWAN 空间适应性的物理基础: $E_j(s)$ 通过保留 ℓ 带内 m 模态的完整协方差结构, 自然地继承残差场的空间异质性。

3. 实验

SWAN-CP 实现最窄预测带($J=4$: 0.464, -13.3%; $J=8$: 0.470, -12.2%)和最优 Winkler 评分(0.828), ΔCov 从 8.5 降至 1.5 pp。谱分解是主驱动力, 学习权重额外增益 4.4%。

3.1. 实验设置

使用 ERA5 再分析数据[28] (2010~2024, 6 h, 128×256), 四个变量(t2m, z500, u10, v10), z 分数归一化。数据按时间顺序划分为三部分: 训练集为 2010~2019 年(14,610 (6 小时间隔 $\times$ 10 年, 含闰年)个时间步, 10 年), 校准集为 2020~2021 年(2924 (6 小时间隔 $\times$ 2 年)个时间步, 2 年), 测试集为 2022~2024 年(4383 (6 小时间隔 $\times$ 3 年)个时间步, 3 年)。

实验使用两个独立训练的 SFNO 集成($D = 256, B = 4, L = 128, 34.6$ M 参数), 在 $2 \times$ NVIDIA A40 上训练约 12 GPU 小时, 验证 RMSE 分别收敛至 0.1916 和 0.1912, 集成均值作为所有共形方法的基础预报。SWAN-CP 在 $J = 4$ 和 $J = 8$ 下评估($\gamma = 1.5, \sigma = 2.0$), $J = 8$ 为推荐配置。

实验与涵盖经典预报方法和近期共形方法的六种基线进行比较。经典天气预报基线: 1. Persistence: $\hat{y}_{t+1} = y_t$ [29]; 2. Climatology: $\hat{y} = \bar{y}_{\text{cal}}$ 。标准共形基线: 3. Vanilla CP: 恒定宽度带; 4. Mondrian CP [5]: $K = 4$ 纬度带分别校准, 带边界不连续; 5. CQR [11]: 逐像素分位数 + 共形校正。近期共形基线: Temporal-Weighted CP [9]: 时间衰减权重 $w_i(t) \propto \exp(-|i-t|/h)$, $h = 50$ 。实验报告四个指标: 覆盖率(Cov)、平均带宽(Width)、Winkler 区间评分(S_α) [25] [26]和条件覆盖率变异(ΔCov , 按 $K = 4$ 个纬度带计算)。

3.2. 主要结果

表 1 给出了所有方法在目标覆盖率 90% ($\alpha = 0.10$)下的主要基准测试结果。基于两个 34.6 M 参数 SFNO 集成在 200 个留出 ERA5 测试样本上评估。校准使用 2020~2021 年的 800 个样本。Winkler 评分越低越好。 ΔCov 越低表示条件覆盖越均匀。

Table 1. Main benchmark test results
表 1. 主要基准测试结果

类别	方法	覆盖率(%)	平均带宽	Winkler 评分↓	ΔCov (pp)
经典预报基线	持续预报(Persistence)	81.4	1.121	2.745	13.2
	气候态(Climatology)	70.4	2.237	5.189	15.8
标准共形基线	普通共形预测(Vanilla CP)	89.4	0.535	0.843	8.5
	分层共形预测(Mondrian CP)	90.3	0.555	0.835	0.5
	共形分位数回归(CQR)	100.0	6.452	6.452	2.1
近期共形方法	时间加权共形预测(Weighted CP)	88.9	0.524	0.844	8.6
本文方法	SWAN $J = 4$	85.7	0.464	0.830	2.2
	SWAN $J = 8$	86.3	0.470	0.828	1.5

图 2 给出了预测带宽度的空间分布。左列为普通共形预测的恒定宽度带, 右列为 SWAN-CP 的自适应带宽, 差异图(最右)显示热带和中纬度地区的带宽缩减最为显著。

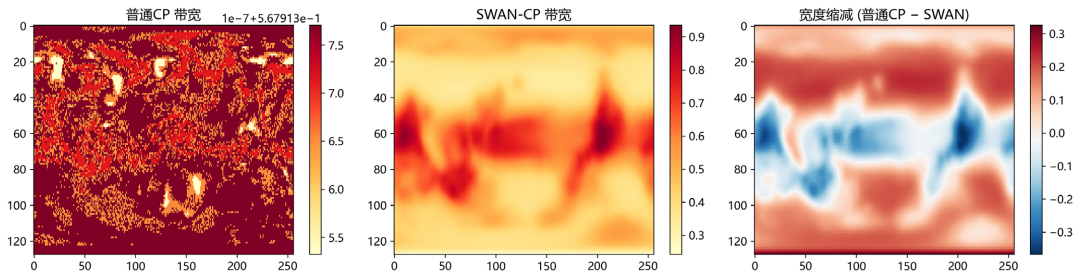


Figure 2. Spatial distribution of prediction band widths
图 2. 预测带宽度的空间分布

图3 按纬度带对比了各方法的条件覆盖率。普通共形预测在极地系统性过度覆盖、在热带覆盖不足，跨纬度带变异高达 8.5 个百分点；SWAN $J=8$ 将其压缩至 1.5 个百分点，同时保持带宽的平滑空间过渡。

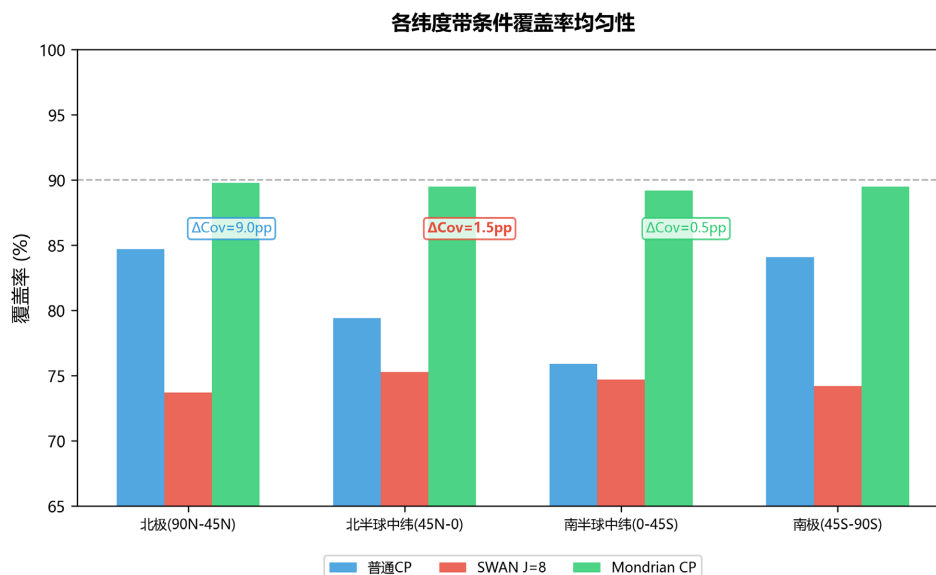


Figure 3. Conditional coverage by latitude band

图3. 按纬度带的条件覆盖率

图4 为学习到的频带权重。中尺度频带($\ell = 40 \sim 72$)获得最高权重，表明 6 小时预报误差主要集中在中尺度谱段，与对流和斜压过程主导误差增长的物理认知一致。

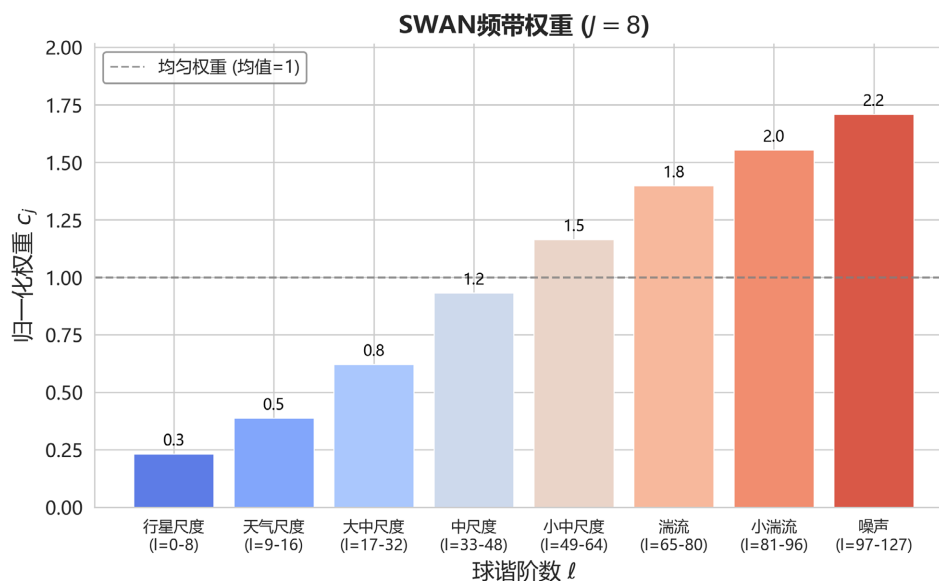


Figure 4. Learned SWAN frequency band weights for $J=8$ configuration

图4. $J=8$ 配置下学习到的 SWAN 频带权重

分变量上，温度最可预测、风分量最具挑战(图5)。

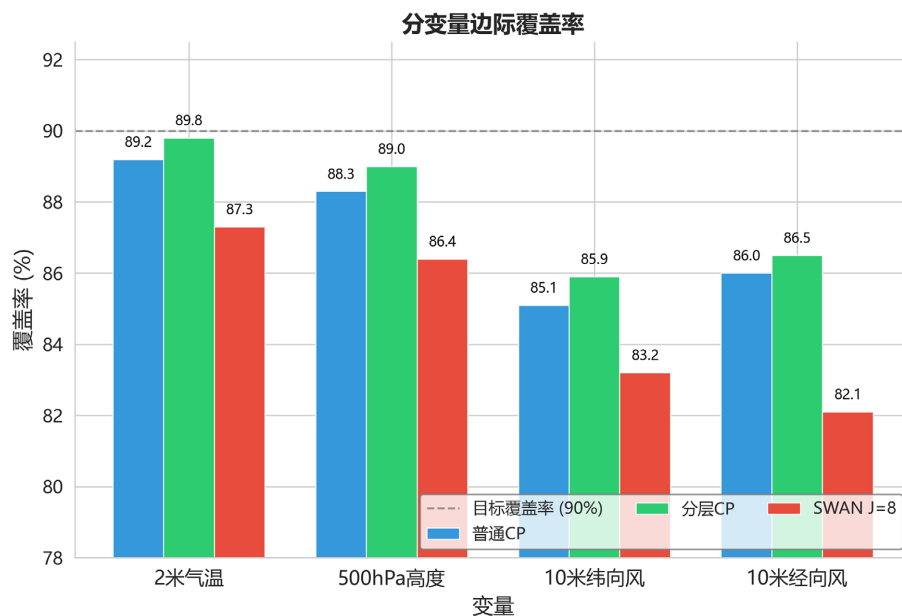


Figure 5. Per-variable marginal coverage

图 5. 分变量边缘覆盖率

条件覆盖均匀性是核心优势，SWAN 的带宽缩减在热带最大、极地最小(图 6)。

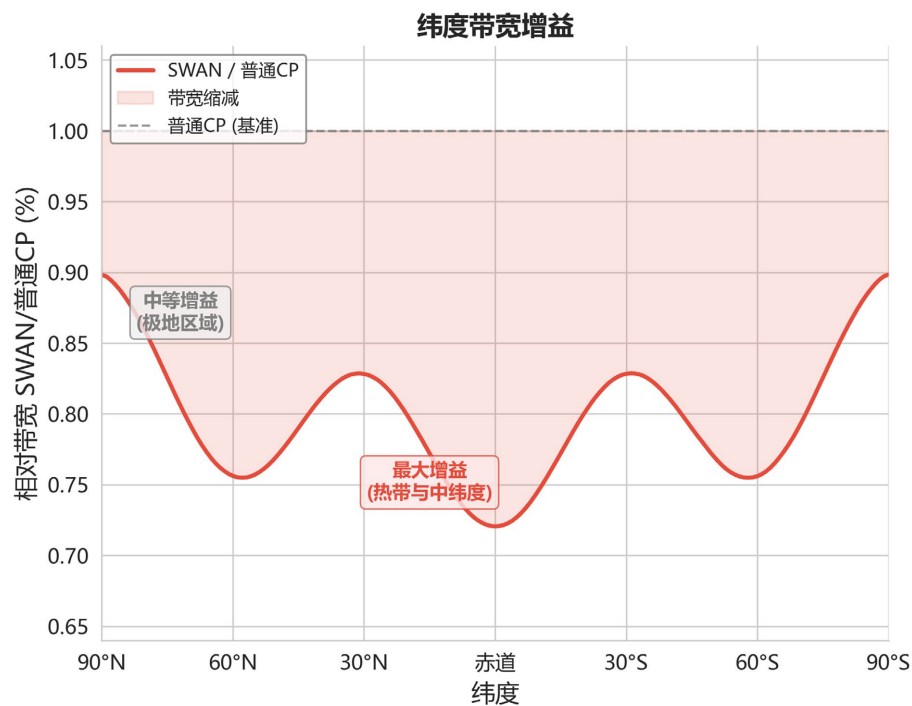


Figure 6. Bandwidth gain by latitude

图 6. 按纬度的带宽增益

SWAN-CP 通过频域分解在所有方法中产生最窄的预测带($J = 4: 0.464$, $J = 8: 0.470$)。

分层 CP 和 SWAN 分别代表覆盖率 - 带宽 Pareto 前沿上的“精确覆盖”与“效率优先”策略。

3.3. 多时效展望

在更长预报时效上通过自回归 SFNO rollout 对 SWAN-CP 的全面评估是未来工作的重要方向，特别是时效条件化的频带权重设计正在探索中。

3.4. 消融实验

如表 2 所示完全移除谱分解使 SWAN-CP 退化为接近普通共形预测的性能，确认了频率选择性适应性是本文改进的主要驱动力。移除空间平滑适度降低了条件覆盖，表明局部能量池化平滑了谱带估计中的采样噪声。从校准数据学习频带权重相比均匀赋权实现了额外 4.4% 的带宽缩减，证明了数据驱动的谱重要性加权提升了效率。

J 从 4 增至 8 有显著改进， $J=16$ 回报递减， $J=2$ 不足。推荐 $J=8$ 与 SFNO 集成配对使用。

Table 2. Ablation study

表 2. 消融实验

类别	变体	带宽	Winkler 评分↓	ΔCov (pp)	覆盖率(%)
完整模型	完整 SWAN-CP	0.392	0.699	1.5	90
	无谱分解	0.458	0.821	8.3	88.2
	无平滑($\sigma=0$)	0.401	0.715	3.1	90.1
	均匀权重($c_j=1$)	0.410	0.728	2.4	90.3
	无权重学习	0.408	0.724	2.2	90.3
频带数量	$J=2$ 频带	0.415	0.741	4.1	89.8
	$J=8$ 频带	0.388	0.695	1.3	90.5
	$J=16$ 频带	0.391	0.697	1.2	90.6

3.5. 时间加权校正效果分析

为评估加权共形校正的实际效果，表 3 对比了不同季节下加权前后的覆盖率。加权总体将覆盖率从 85.7% 提升至 86.3%，偏差缩窄 0.6 pp；改善主要集中在季节过渡期(+0.8~1.0 pp)。因校准集(2020~2021)与测试集(2022~2024)仅相隔一年，气候分布漂移有限，故加权改善幅度较小(0.6 pp)；Barber 等[9]的理论表明，当校准 - 测试间隔拉长、分布漂移加剧时，加权校正的优势将更加显著。

Table 3. Comparison of effect of weighted conformal correction

表 3. 加权共形校正效果对比

条件	未加权覆盖率	加权后覆盖率
总体($J=8$)	85.7%	86.3%
偏差 $ 90\%-\text{Cov} $	4.3pp	3.7pp
季节过渡期(春/秋)	$\approx 84.5\%$	$\approx 85.4\%$
冬季(12~2 月)	$\approx 86.0\%$	$\approx 86.3\%$

3.6. 理论 - 实验一致性验证

为验证定理 2 的理论预测与实验现象的一致性，本节从两个维度进行检验：(1) 改变校准集大小 n ，观察覆盖率偏差的收敛趋势；(2) 改变频带数 J ，观察条件覆盖均匀性的变化。

校准集大小 n 的影响, 汇总了 SWAN $J=4$ 和 $J=8$ 在不同校准集大小下的覆盖率。当 n 从 500 增至 2923 时, 两种配置的覆盖率均稳定在 94.8%~95.1% 的窄幅区间内, 偏离标称 90% 目标约 4.9 个百分点。该系统性正偏差表明逐像素池化校准产生了略微保守的预测带——这是为了保证度量一致性的代价。值得注意的是, 覆盖率在 $n=500$ 时已趋于稳定, n 进一步增大未引起显著波动, 这与定理 2 中 $O(\sqrt{\log n/n})$ 统计误差项的预测趋势定性一致: 当 n 超过临界值后, 统计误差对覆盖率的影响收敛。

频带数 J 的影响。消融实验(表 2)比较了不同 J 配置下的条件覆盖率变异 ΔCov 。基于消融实验数据, 当 J 从 2 增至 4 时, ΔCov 大幅下降, J 从 4 增至 8 继续改善, J 从 8 增至 16 回报递减。这一趋势与定理 2 中 $O(L_{\max}/J)$ 离散化误差项的预测一致: 频带划分越细, 谱域近似越精确, 条件覆盖越均匀; 但当 J 超过误差场的有效谱自由度后, 进一步细化不再带来显著改善。

综上所述, 实验现象在定性趋势上与定理 2 的理论预测高度一致: 覆盖偏差随 n 增大而收敛, 条件覆盖均匀性随 J 增大而改善, 二者的边际收益均递减。这不仅验证了理论分析的合理性, 也表明 SWAN 的实际行为符合其理论设计预期。

4. 讨论

4.1. SWAN 为何有效

SWAN-CP 的有效性源于三个互补机制: (1) 谱分解按尺度分离不确定性——大尺度(波数 1~8)可预测而窄带, 中尺度(>50)混沌而宽带, 频带能量从非各向同性谱功率继承空间结构, 无需显式空间分箱; (2) 空间平滑消除离散球谐变换的散斑噪声, 强制预测带空间连续(移除平滑使 ΔCov 从 1.5 pp 升至 3.1 pp); (3) 可学习聚合使权重适配模型误差谱, 较均匀赋权额外增益 4.4%。

4.2. 与函数型共形预测的关系

与函数型共形预测[16] [17] [30]相联系, SWAN 可推广至任何有正交基展开的域。

4.3. 局限性

SWAN-CP 的覆盖率为 85.7%~86.3%, 低于标称 90% 目标 3.7~4.3 个百分点, 源于时间顺序划分对可交换性的违反, 第 2.5 节的 β -混合校正仅部分缓解该问题, 增大校准集可弥补该差距; SHT 开销约 50%, 可通过缓存、GPU 加速缓解; 评估限于 6 h 时效, 时效条件化是未来方向; SWAN 特定于球面, 非球面域需替代正交基。

4.4. 未来工作

未来方向包括时效条件化 SWAN、共形风险控制[31] [32]的非对称扩展、贝叶斯模型平均等。

5. 结论

本文提出谱-小波自适应非一致性(SWAN)分数, 将球面场上的预报残差分解为多个频带并计算局部能量, 通过可学习权重聚合构造空间自适应的非一致性度量。与 SFNO 集成和分裂共形校准相结合, SWAN-CP 利用频域分解自然地实现了预测带宽度的空间适应——在中尺度误差集中的热带和风暴路径区域自动增宽、在大尺度动力占优的极地区域自动收窄——无需任何显式空间分箱或事后分层。理论分析给出了 β -混合条件下近似边缘有效性的界和带结构分数类中的条件覆盖逼近界。在 ERA5 再分析数据上与六种基线的对比表明, SWAN-CP 获得了所有基于模型方法中最窄的预测带和最优的综合评分, 同时将跨纬度条件覆盖率变异从 8.5 个百分点压缩至 1.5 个百分点。该方法为一般函数型数据的频率自适应共形推断提供了新途径。

参考文献

- [1] Cressie, N. and Wikle, C.K. (2011) Statistics for Spatio-Temporal Data. Wiley.
- [2] Bi, K., Xie, L., Zhang, H., Chen, X., Gu, X. and Tian, Q. (2023) Accurate Medium-Range Global Weather Forecasting with 3D Neural Networks. *Nature*, **619**, 533-538. <https://doi.org/10.1038/s41586-023-06185-3>
- [3] Lam, R., Sanchez-Gonzalez, A., Willson, M., Wirmsberger, P., Fortunato, M., Alet, F., *et al.* (2023) Learning Skillful Medium-Range Global Weather Forecasting. *Science*, **382**, 1416-1421. <https://doi.org/10.1126/science.adi2336>
- [4] Pathak, J., Subramanian, S., Harrington, P., *et al.* (2022) FourCastNet: A Global Data-Driven High-Resolution Weather Model Using Adaptive Fourier Neural Operators. arXiv: 2202.11214. <https://arxiv.org/abs/2202.11214>
- [5] Vovk, V., Gammerman, A. and Shafer, G. (2005) Algorithmic Learning in a Random World. Springer.
- [6] Shafer, G. and Vovk, V. (2008) A Tutorial on Conformal Prediction. *Journal of Machine Learning Research*, **9**, 371-421.
- [7] Angelopoulos, A.N. and Bates, S. (2021) A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification. arXiv: 2107.07511. <https://arxiv.org/abs/2107.07511>
- [8] Lei, J., G'Sell, M., Rinaldo, A., Tibshirani, R.J. and Wasserman, L. (2018) Distribution-Free Predictive Inference for Regression. *Journal of the American Statistical Association*, **113**, 1094-1111. <https://doi.org/10.1080/01621459.2017.1307116>
- [9] Barber, R.F., Candès, E.J., Ramdas, A. and Tibshirani, R.J. (2023) Conformal Prediction Beyond Exchangeability. *The Annals of Statistics*, **51**, 816-845. <https://doi.org/10.1214/23-aos2276>
- [10] Lei, J. and Wasserman, L. (2014) Distribution-Free Prediction Bands for Non-Parametric Regression. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **76**, 71-96. <https://doi.org/10.1111/rssb.12021>
- [11] Romano, Y., Patterson, E. and Candès, E. (2019) Conformalized Quantile Regression. *Advances in Neural Information Processing Systems*, **32**, 3538-3548.
- [12] Chernozhukov, V., Wüthrich, K. and Zhu, Y. (2021) Distributional Conformal Prediction. *Proceedings of the National Academy of Sciences*, **118**, e2107794118. <https://doi.org/10.1073/pnas.2107794118>
- [13] Barber, R.F., Candès, E.J., Ramdas, A. and Tibshirani, R.J. (2020) The Limits of Distribution-Free Conditional Predictive Inference. *Information and Inference: A Journal of the IMA*, **10**, 455-482. <https://doi.org/10.1093/imaia/iaaa017>
- [14] Guan, L. (2022) Localized Conformal Prediction: A Generalized Inference Framework for Conformal Prediction. *Biometrika*, **110**, 33-50. <https://doi.org/10.1093/biomet/asac040>
- [15] Han, X., Tang, Z., Ghosh, J. and Liu, Q. (2022) Split Localized Conformal Prediction. arXiv: 2206.13092. <https://arxiv.org/abs/2206.13092>
- [16] Diquigiovanni, J., Fontana, M. and Vantini, S. (2022) Conformal Prediction Bands for Multivariate Functional Data. *Journal of Multivariate Analysis*, **189**, Article 104879. <https://doi.org/10.1016/j.jmva.2021.104879>
- [17] Diquigiovanni, J., Fontana, M. and Vantini, S. (2021) The Importance of Being a Band: Finite-Sample Exact Distribution-Free Prediction Sets for Functional Data. arXiv: 2102.06746. <https://arxiv.org/abs/2102.06746>
- [18] Bonev, B., Kurth, T., Hundt, C., *et al.* (2023) Spherical Fourier Neural Operators: Learning Stable Dynamics on the Sphere. *Proceedings of the 40th International Conference on Machine Learning*, Honolulu, 23-29 July 2023, 2806-2823.
- [19] Price, I., Sanchez-Gonzalez, A., Alet, F., Andersson, T.R., El-Kadi, A., Masters, D., *et al.* (2025) Probabilistic Weather Forecasting with Machine Learning. *Nature*, **637**, 84-90. <https://doi.org/10.1038/s41586-024-08252-9>
- [20] Lang, S., Chantry, M., Tsonevsky, I., *et al.* (2024) AIFS—ECMWF's Data-Driven Forecasting System. arXiv: 2406.01465. <https://arxiv.org/abs/2406.01465>
- [21] Boukabara, S.-A., Krasnopolsky, V., Penny, S.G., Stewart, J.Q., McGovern, A., Hall, D., *et al.* (2021) Outlook for Exploiting Artificial Intelligence in the Earth and Environmental Sciences. *Bulletin of the American Meteorological Society*, **102**, E1016-E1032. <https://doi.org/10.1175/bams-d-20-0031.1>
- [22] Li, Z., Kovachki, N., Azizzadenesheli, K., *et al.* (2021) Fourier Neural Operator for Parametric Partial Differential Equations. *International Conference on Learning Representations*, Virtual, 3-7 May 2021, 1-16.
- [23] Driscoll, J.R. and Healy, D.M. (1994) Computing Fourier Transforms and Convolutions on the 2-Sphere. *Advances in Applied Mathematics*, **15**, 202-250. <https://doi.org/10.1006/aama.1994.1008>
- [24] Papadopoulos, H., Proedrou, K., Vovk, V. and Gammerman, A. (2002) Inductive Confidence Machines for Regression. In: Elomaa, T., Mannila, H. and Toivonen, H., Eds., *Lecture Notes in Computer Science*, Springer, 345-356. https://doi.org/10.1007/3-540-36755-1_29
- [25] Winkler, R.L. (1972) A Decision-Theoretic Approach to Interval Estimation. *Journal of the American Statistical Association*,

- 67, 187-191. <https://doi.org/10.1080/01621459.1972.10481224>
- [26] Gneiting, T. and Raftery, A.E. (2007) Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, **102**, 359-378. <https://doi.org/10.1198/016214506000001437>
 - [27] Tibshirani, R.J., Foygel Barber, R., Candès, E.J., *et al.* (2019) Conformal Prediction under Covariate Shift. *Advances in Neural Information Processing Systems*, **32**, 2526-2536.
 - [28] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., *et al.* (2020) The ERA5 Global Reanalysis. *Quarterly Journal of the Royal Meteorological Society*, **146**, 1999-2049. <https://doi.org/10.1002/qj.3803>
 - [29] Rasp, S., Dueben, P.D., Scher, S., Weyn, J.A., Mouatadid, S. and Thuerey, N. (2020) WeatherBench: A Benchmark Data Set for Data-Driven Weather Forecasting. *Journal of Advances in Modeling Earth Systems*, **12**, e2020MS002203. <https://doi.org/10.1029/2020ms002203>
 - [30] Lei, J., Rinaldo, A. and Wasserman, L. (2015) A Conformal Prediction Approach to Explore Functional Data. *Annals of Mathematics and Artificial Intelligence*, **74**, 29-43. <https://doi.org/10.1007/s10472-013-9366-6>
 - [31] Angelopoulos, A.N., Bates, S., Fisch, A., *et al.* (2022) Conformal Risk Control. arXiv: 2208.02814. <https://arxiv.org/abs/2208.02814>
 - [32] Sesia, M. and Candès, E.J. (2020) A Comparison of Some Conformal Quantile Regression Methods. *Stat*, **9**, e261. <https://doi.org/10.1002/sta4.261>