Advances in Applied Mathematics 应用数学进展, 2026, 15(11), 38-50
Published Online November 2026 in Hans. https://www.hanspub.org/journal/aam
https://doi.org/10.12677/aam.2026.1511430

Hans汉斯

基于气象场景条件风险最小化的多源迁移风速预测方法

王亚琪

辽宁师范大学数学学院，辽宁 大连

收稿日期：2026年10月4日；录用日期：2026年10月28日；发布日期：2026年11月4日

摘　要

针对新建风电场历史样本不足、跨场站分布差异及多源知识迁移易产生负迁移的问题，提出一种基于气象场景条件风险最小化的多源迁移风速预测方法。综合源域经验风险、条件分布差异、预测不确定性与梯度冲突，构建气象场景驱动的动态源选择机制；结合共享专属特征解耦、迁移门控与轻量残差适配，实现可迁移知识筛选与目标场站少样本适应。基于6个风电场开展留一场站实验，结果表明，本文方法在15 min和60 min预测任务中的MAE分别为0.448 m/s和0.972 m/s，较目标域独立训练的平均迁移增益分别达到37.34%和37.00%，负迁移率均为2.50%，验证了所提方法在少样本风速预测中的有效性。

关键词

风速预测，多源迁移学习，条件域适应，负迁移

Multi-Source Transfer Learning for Wind Speed Forecasting Based on Meteorological Scenario-Conditioned Risk Minimization

Yaqi Wang

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: October 4, 2026; accepted: October 28, 2026; published: November 4, 2026

Abstract

To address the challenges of insufficient historical samples at newly established wind farms, cross-site distribution discrepancies, and negative transfer arising from multi-source knowledge transfer, a multi-source transfer learning method for wind speed forecasting based on meteorological

文章引用: 王亚琪. 基于气象场景条件风险最小化的多源迁移风速预测方法[J]. 应用数学进展, 2026, 15(11): 38-50.
DOI: 10.12677/aam.2026.1511430

scenario-conditioned risk minimization is proposed. By jointly considering source-domain empirical risk, conditional distribution discrepancy, prediction uncertainty, and gradient conflict, a meteorological scenario-driven dynamic source selection mechanism is developed. Shared-private feature disentanglement, transfer gating, and lightweight residual adaptation are further integrated to enable transferable knowledge selection and few-sample adaptation at the target wind farm. Leave-one-site-out experiments are conducted on six wind farms. The results show that the proposed method achieves MAEs of 0.448 m/s and 0.972 m/s for the 15-min and 60-min forecasting tasks, respectively, with average transfer gains of 37.34% and 37.00% over target-domain independent training. The negative transfer rate is 2.50% for both forecasting horizons, demonstrating the effectiveness of the proposed method for few-sample wind speed forecasting.

Keywords

Wind Speed Forecasting, Multi-Source Transfer Learning, Conditional Domain Adaptation, Negative Transfer

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

风速具有显著的随机性、间歇性与非平稳性，其预测精度直接影响风电功率评估及电力系统运行调度[1]。随着深度学习在风速预测中的广泛应用，模型对历史样本规模和数据覆盖范围提出了更高要求。然而，对于投运时间较短的新建风电场，有限观测数据难以充分表征季节变化、风向转换及不同气象状态下的风速演化规律，独立训练模型容易产生过拟合和泛化能力不足等问题。

为此，迁移学习逐渐被用于风场预测任务，通过利用成熟场站的数据或模型参数提高目标场站的少样本预测能力[2]。Hu 等[2]利用数据丰富场站的隐藏层参数初始化目标模型，程凯等[3]结合特征选择与多层级深度迁移学习提高风电功率预测精度，Peng 等[4]进一步利用深度特征学习与多层级迁移机制实现跨场站知识复用。相关研究表明，源场站中蕴含的时序规律能够缓解目标场站样本不足问题。然而，不同风场受气象条件、地形及测量环境影响，其数据分布与风速演化规律存在差异，直接迁移可能导致负迁移。域适应理论通过刻画源域与目标域分布差异为跨域知识迁移提供理论基础[5]-[8]；共享－专属表征分解能够在提取域不变特征的同时保留域特有信息[9]；参数高效适应通过冻结基础模型并引入轻量模块降低目标域适配复杂度[10]。如何依据气象场景动态选择有效源知识并抑制负迁移，仍需进一步研究。

针对上述问题，本文提出一种基于气象场景条件风险最小化的多源迁移风速预测方法。综合源域经验风险、条件分布差异、预测不确定性及梯度冲突，构建气象场景驱动的动态源选择机制；通过共享专属特征解耦筛选跨场站可迁移信息，并结合迁移门控与轻量残差适配抑制有害知识传递，实现新建风电场少样本条件下的多源知识有效复用和预测模型快速适应。

2. 模型介绍

2.1. 多源迁移学习与目标风险理论

迁移学习旨在利用源任务中获得的数据结构、特征表示或模型参数，提高目标任务在样本有限条件下的泛化能力。单源迁移仅考虑一个源域，而多源迁移同时利用多个源域，其关键不在于简单扩充训练

样本，而在于判定各源域对目标任务的有效贡献。若某一源场站数据量充足但其风速生成机制、局地地形或测量条件与目标场站差异较大，该源场站仍可能增加目标预测误差。

令 K 表示候选源场站数量，f 表示预测函数，$R_T(f)$ 表示目标域风险，$R_{S_k}(f)$ 表示第 k 个源域经验风险，α_k 表示满足 $\alpha_k \geq 0$ 的源权重，$d(S_k, T)$ 表示第 k 个源域与目标域的分布差异，λ^* 表示由任务机制差异导致的联合不可约误差，$\Omega(\boldsymbol{\alpha})$ 表示权重正则项。受域适应目标风险边界思想启发，多源条件下目标风险可概括表示为

$$R_T(f) \leq \sum_{k=1}^{K} \alpha_k R_{S_k}(f) + \sum_{k=1}^{K} \alpha_k d(S_k, T) + \lambda^* + \Omega(\boldsymbol{\alpha}) \tag{1}$$

式(1)表明，迁移性能同时取决于源模型的经验风险、源–目标分布接近程度及任务机制一致性，并受到源组合权重的影响。正迁移是指引入源知识后目标域预测误差降低，负迁移则表示误差增大。因此，多源迁移需要对源域质量进行条件化评估，并在训练过程中限制低价值或有害源域的贡献。

2.2. 条件域适应与共享–专属表示理论

域适应通过缩小源域与目标域在特征空间中的分布差异，使源域学习的表示服务于目标任务。传统方法通常对边缘分布进行全局对齐，但风速序列在不同季节、风向和风速区间具有差异化演化规律，全局对齐容易将不同气象状态的样本强制匹配。为保持气象条件一致性，在场景变量给定时约束条件分布，其中 z 表示由输入样本映射得到的隐空间特征，r 表示气象场景变量：

$$P_S(z \mid r) \approx P_T(z \mid r) \tag{2}$$

条件分布差异可采用 MMD、Wasserstein 距离或 CORAL 等度量。MMD 比较两域样本在再生核希尔伯特空间中的均值嵌入，适合小批量端到端训练；Wasserstein 距离刻画分布间最优运输代价，对支撑集不重叠情形更具解释性；CORAL 通过匹配协方差实现二阶统计对齐。实际应用中应结合模型复杂度和样本规模选择度量方式。

不同场站的隐特征并非全部可迁移。区域天气系统、共同风速变化趋势和周期性时序模式构成跨场站共享信息；地形、海拔、局地湍流、传感器偏差等反映场站专属性。本文将隐空间特征 $\boldsymbol{z}$ 分解为共享表征 $\boldsymbol{z}^{\mathrm{sh}}$ 与场站专属表征 $\boldsymbol{z}^{\mathrm{pr}}$：

$$\boldsymbol{z} = \left[\boldsymbol{z}^{\mathrm{sh}}, \boldsymbol{z}^{\mathrm{pr}}\right] \tag{3}$$

共享表征 $\boldsymbol{z}^{\mathrm{sh}}$ 用于源–目标条件对齐，专属表征 $\boldsymbol{z}^{\mathrm{pr}}$ 用于保留各场站局部信息。本文采用共享特征条件对齐与批次级正交约束，减少两类表示的冗余耦合，在提取跨域可迁移信息的同时保留目标场站局部特性。

2.3. 基础预测模型与参数适应机制

本文方法不依赖特定预测主干，可将基础预测器统一抽象为特征映射 $G_{\theta_g}(X)$ 与预测映射 $H_{\theta_h}(\cdot)$ 的组合。循环神经网络、时序卷积网络、Transformer 以及线性预测模型均可作为具体实现。为突出多源迁移策略本身的作用，本文仿真实验采用轻量预测器进行验证：

$$\hat{y} = H_{\theta_h}\left[G_{\theta_g}(X)\right] \tag{4}$$

其中，X 为长度为 L 的多变量滑动时间窗，$\hat{y}$ 为未来 τ 步风速预测值，θ_g、θ_h 分别表示特征映射与预测映射参数。该统一抽象使迁移风险估计、源权重计算及目标适配不依赖具体主干结构。

参数高效适应在固定基础预测器主体的条件下，仅利用少量目标数据更新任务相关部分，从而降低

少样本过拟合风险。式中给出的一般特征映射可与瓶颈适配器结合；本文轻量仿真实现采用标准化岭回归作为基础预测器，并以单参数残差偏置完成目标场站适配。

3. 气象场景感知的多源选择性迁移风速预测方法

针对新建风电场少样本预测问题，本文构建气象场景感知的多源选择性迁移框架。方法由三部分组成：首先，估计不同气象场景下各源场站的迁移风险并生成动态源权重；其次，将多源知识分解为共享表征和场站专属表征，仅对高价值共享表征执行双条件分布迁移；最后，从优化方向、分布距离和预测可信度三个维度抑制负迁移，并通过参数高效模块适应目标场站。三部分均以目标场站验证风险为最终评价依据。

3.1. 气象场景条件下的多源迁移风险建模与知识选择

设时刻 t 可获得的气象变量向量为 m_t，其中包含当前及历史风速、风向、温度、气压、湿度和由历史序列构造的变化特征。场景编码器 $E_\phi(\cdot)$ 用于将上述变量映射为气象场景表示 r_t：

$$r_t = E_\phi(m_t) \tag{5}$$

本文采用离散气象场景表征，根据风向扇区、风速区间和稳定度状态构建场景标签。场景编码器 $E_\phi(\cdot)$ 表示气象变量到该离散场景的映射，由 4 个风向扇区、3 个风速区间与 2 个稳定度状态组合形成 24 类气象场景。该表征便于在少样本条件下稳定估计不同天气状态对应的迁移风险。本文中的 $E_\phi(\cdot)$ 表示基于上述风向、风速及稳定度划分规则构成的场景映射，其中 ϕ 表示相应场景划分参数，并非额外训练的深度编码网络参数。

对于第 k 个源场站，在气象场景 r 下定义迁移风险 $B_k(r)$ 为

$$B_k(r) = R_k(r) + \beta_1 d_k(r) + \beta_2 U_k(r) + \beta_3 C_k(r) \tag{6}$$

考虑到经验风险、分布距离、预测不确定性和梯度冲突度具有不同量纲及取值范围，计算综合迁移风险前，分别在候选源场站范围内对各指标进行归一化处理。$R_k(r)$ 定义为第 k 个源模型在目标支持集中对应场景样本上的平均绝对预测误差，用于表征其基础适应能力；$d_k(r)$ 定义为源场站与目标场站在当前场景下标准化共享特征均值的差异，用于描述数据层面的可迁移性；$U_k(r)$ 由源模型验证残差在相应场景内的标准差表征；$C_k(r)$ 表示源任务与目标任务更新方向的不一致程度，并由第 3.3 节的梯度一致性指标确定。β_1、β_2、β_3 为非负权衡系数。

在 $\alpha_k(r) \geq 0$ 且 $\sum_k \alpha_k(r) = 1$ 的约束下，场景化源权重通过最小化目标风险上界获得：

$$\boldsymbol{\alpha}^*(r) = \arg\min_{\boldsymbol{\alpha}(r)\in\Delta_K} \left[\sum_{k=1}^{K} \alpha_k(r) B_k(r) + \lambda_1 \left\| \boldsymbol{\alpha}(r) \right\|_2^2 + \lambda_2 \Omega_s\left(\boldsymbol{\alpha}(r)\right) \right] \tag{7}$$

式(7)中的二范数正则项用于避免源权重过度集中并提高权重估计稳定性；$\Omega_s(\boldsymbol{\alpha})$ 表示稀疏约束项，用于使迁移权重集中于少量高价值源场站。实际实现中仅保留综合风险最低的两个候选源，并重新归一化其权重。式(7)给出了场景化源权重选择的风险最小化目标；为降低直接求解约束优化问题的复杂度，实际计算采用式(8)所示的温度 Softmax 形式对该目标进行连续近似。

$$\alpha_k(r) = \frac{\exp\left[-B_k(r)/T\right]}{\sum_{j=1}^{K} \exp\left[-B_j(r)/T\right]} \tag{8}$$

温度 T 控制源权重的集中程度，T 越小，权重越集中于低风险源；T 越大，权重分布越平缓。本文实验取 $T = 0.35$。当某一场景的目标支持样本少于 8 个时，相应风险分量向全局统计量收缩，以减小极少

样本条件下风险估计的方差。

3.2. 场站共享–专属表征解耦与条件分布迁移

动态源权重回答“从哪些场站迁移”，共享–专属表征解耦进一步回答“迁移哪些知识”。对第 k 个源场站输入 $X_{s,k}$ 和目标场站输入 X_t，采用共享编码器 $G_{\mathrm{sh}}(\cdot)$ 提取跨场站共性特征，并分别利用场站专属编码器 $G_{s,k}^{\mathrm{pr}}(\cdot)$ 与 $G_t^{\mathrm{pr}}(\cdot)$ 提取局地特征：

$$z_{s,k}^{\mathrm{sh}} = G_{\mathrm{sh}}\left(X_{s,k}\right), \qquad z_t^{\mathrm{sh}} = G_{\mathrm{sh}}\left(X_t\right) \tag{9}$$

$$z_{s,k}^{\mathrm{pr}} = G_{s,k}^{\mathrm{pr}}\left(X_{s,k}\right), \qquad z_t^{\mathrm{pr}} = G_t^{\mathrm{pr}}\left(X_t\right) \tag{10}$$

共享编码器重点描述区域天气系统驱动、风速变化趋势和可复用时序模式；专属编码器刻画地形、海拔、局地湍流、设备布置和测量偏差形成的场站差异。两类表示通过拼接或门控融合进入预测头，但域适应约束仅作用于共享表示，避免将局地特征误判为需要消除的域偏差。

上述共享编码器与专属编码器为统一方法框架下的特征映射表示。在本文轻量仿真实现中，共享映射与专属映射分别由预先划分的共享特征子集、场站专属特征子集及其标准化线性映射实现，不额外训练深层编码网络。

具体划分如下：共享特征由 19 维变量构成，包括 80、100 和 120 m 高度风速，100 m 风速的 1、3、7、15 和 23 步滞后量，24 步滚动均值、标准差、最小值、最大值与趋势，以及 100 m 温度、100 m 气压、2 m 相对湿度、80~120 m 风切变和 100 m 风向正余弦编码。其中，多高度与历史风速描述跨场站可复用的风速垂直结构和时间演化，温压湿、风切变及风向编码描述共同天气背景。小时与年内日的 4 维正余弦编码仅作为时间辅助变量输入预测器，不参与条件分布对齐。专属特征在轻量实现中明确为 1 维目标场站残差偏置，其数值由目标支持集预测残差的均值估计，并由目标验证集选择收缩强度；测试阶段固定该偏置，不使用测试标签。该专属量用于吸收局地地形、设备布置和测量偏差引起的系统性误差。

为保证表示分工清晰，采用正交约束降低共享特征与专属特征的线性相关性：

$$\mathcal{L}_{\mathrm{orth}} = \sum_{d\in\mathcal{D}} \left\| \left(Z_d^{\mathrm{sh}}\right)^{\mathrm{T}} Z_d^{\mathrm{pr}} \right\|_F^2 \tag{11}$$

其中，Z_d^{sh}、Z_d^{pr} 分别表示一个训练批次内第 d 个域的共享特征矩阵和专属特征矩阵，d 遍历各候选源场站及目标场站，$\|\cdot\|_F$ 表示 Frobenius 范数。该正交约束用于减少共享信息与场站专属信息之间的线性相关性。

条件分布迁移仅对相同或相近气象场景下的共享表征实施对齐。预测状态 v 由预测时刻能够获得的历史风速状态、历史变化趋势或基础预测器当前状态构成，不使用未来时刻真实风速标签。设第 k 个源场站样本 i 与目标场站样本 j 的场景表示分别为 $r_{s,i}^{(k)}$、$r_{t,j}$，预测状态表示分别为 $v_{s,i}^{(k)}$、$v_{t,j}$，则软匹配权重定义为：

$$w_{ij}^{(k)} = \exp\left[-\frac{\left\| r_{s,i}^{(k)} - r_{t,j} \right\|_2^2}{\sigma_r^2} - \frac{\left\| v_{s,i}^{(k)} - v_{t,j} \right\|_2^2}{\sigma_v^2} \right] \tag{12}$$

$$\tilde{w}_{ij}^{(k)} = \frac{w_{ij}^{(k)}}{\sum_{p,q} w_{pq}^{(k)}} \tag{13}$$

其中，σ_r、σ_v 分别为气象场景距离与预测状态距离的正带宽参数。按所有源–目标样本对进行归一化后

得到式(13)中的 $w_{ij}^{(k)}$，并据此定义第 k 个源场站与目标场站共享表征的局部条件分布距离。

$$D_k^{\mathrm{cond}} = \left\| \sum_{i,j} w_{ij}^{(k)} \left[\phi\left(z_{s,i}^{(k),\mathrm{sh}}\right) - \phi\left(z_{t,j}^{\mathrm{sh}}\right) \right] \right\|_{\mathcal{H}}^2 \tag{14}$$

式中，$\Phi(\cdot)$ 表示将共享特征映射至再生核希尔伯特空间的特征映射。$d_k(r)$ 用于源场站选择阶段，描述气象特征层面的源–目标差异；D_k^{cond} 用于特征迁移阶段，衡量共享隐空间中的局部条件分布差异，二者分别服务于源选择和共享特征对齐。

$$\mathcal{L}_{\mathrm{cda}} = \sum_{k=1}^{K} \alpha_k(r) D_k^{\mathrm{cond}}(r) \tag{15}$$

最终目标表征由加权共享知识与目标专属知识融合得到：

$$z_t = F_g\left(\left[\sum_{k=1}^{K} \alpha_k(r) A_k\left(z_{s,k}^{\mathrm{sh}}\right), z_t^{\mathrm{sh}}, z_t^{\mathrm{pr}} \right] \right) \tag{16}$$

其中，$A_k(\cdot)$ 表示将第 k 个源场站共享表征映射至统一空间的映射，$F_g(\cdot)$ 为门控融合函数。推理阶段不需要重新访问完整源场站训练数据，仅调用训练阶段已经学习并固化的源域共享知识、模型参数及目标场站适配模块；源共享表示以场景统计量和模型参数的形式固化。

3.3. 梯度风险控制与目标适应

分布接近并不必然意味着优化方向一致。令 $g_{s,k}(r) = \nabla_\theta \mathcal{L}_{s,k}(r)$ 表示第 k 个源场站在气象场景 r 下的任务梯度，$g_t(r) = \nabla_\theta \mathcal{L}_t(r)$ 表示目标场站支持集梯度，采用余弦相似度定义梯度一致性 $C_k(r)$：

$$c_k = \frac{\boldsymbol{g}_{s,k}^{\mathrm{T}} \boldsymbol{g}_t}{\|\boldsymbol{g}_{s,k}\|_2 \|\boldsymbol{g}_t\|_2 + \varepsilon} \tag{17}$$

$$C_k(r) = \max\{0, -c_k(r)\} \tag{18}$$

其中，ε 为防止梯度范数为零而设置的正小量。$C_k(r) = \max\{0, -c_k(r)\}$，用于式(6)中的综合迁移风险。当 $C_k(r) < 0$ 时，源梯度与目标梯度发生方向冲突。参考冲突梯度投影思想[11]，将源梯度在目标梯度方向上的冲突分量投影消除：

$$\boldsymbol{g}_{s,k} = \boldsymbol{g}_{s,k} - \frac{\min\left(0, \boldsymbol{g}_{s,k}^{\mathrm{T}} \boldsymbol{g}_t\right)}{\|\boldsymbol{g}_t\|_2^2 + \varepsilon} \boldsymbol{g}_t \tag{19}$$

投影后的源更新不再包含与目标梯度直接对立的分量。本文轻量线性预测器采用源预测修正量与目标支持集残差方向的余弦一致性作为梯度一致性的可计算近似，并将负一致部分计入 $C_k(r)$，从而与一般梯度投影形式保持一致。

在本文轻量实现中，梯度冲突度主要用于综合迁移风险与门控权重计算；当检测到负一致性时，通过降低相应源场站的有效迁移权重实现冲突抑制，式中的梯度投影形式用于说明一般模型下的冲突消除机制。

轻量实现的梯度冲突近似过程如下。输入为目标适应池真实风速、持续性预测和各源模型预测，输出为场景冲突度及有效源权重。步骤 1：在不含测试样本的目标适应池内，按气象场景划分样本；步骤 2：以真实风速相对持续性预测的残差作为目标更新方向，以第 k 个源模型预测相对持续性预测的修正量作为源更新方向；步骤 3：按式(17)计算二者的余弦一致性，并按式(18)取负一致部分得到冲突度；步骤 4：将冲突度计入式(6)的综合迁移风险，并由式(20)计算门控系数；步骤 5：按式(21)重新归一化得到有效源

权重。全过程仅使用适应阶段可用标签，不访问测试标签。

进一步联合梯度一致性、条件分布差异与预测不确定性构造迁移门控：

$$q_k(r)=\sigma\left[\gamma_1 c_k(r)-\gamma_2 d_k(r)-\gamma_3 U_k(r)+b\right] \tag{20}$$

其中，$\sigma(\cdot)$为 Sigmoid 函数，γ_1、γ_2、γ_3及偏置b为由目标验证集确定的门控系数。$q_k(r)\in(0,1)$表示第k个源场站的迁移门控系数；结合原始源权重后，经门控重新归一化得到有效源权重$\tilde{\alpha}_k(r)$：

$$\tilde{\alpha}_k(r)=\frac{\alpha_k(r)q_k(r)}{\sum_j \alpha_j(r)q_j(r)} \tag{21}$$

目标场站适配采用瓶颈残差结构的一般形式。对基础模型中间状态h，$W_{\text{down}}h$、W_{up}分别表示降维与升维映射，适配后的状态为

$$h'=h+W_{\text{up}}\sigma(W_{\text{down}}h),\qquad \dim(W_{\text{down}}h)\ll\dim(h) \tag{22}$$

式(22)给出了参数高效适配器的一般残差形式。考虑本文仿真实验采用轻量预测模型，为避免少样本条件下引入过多自由参数，实验中进一步采用单参数残差偏置作为简化实现，即仅学习目标场站预测输出的残差校正量δ_t：

$$\hat{y}_{\text{ad}}=\hat{y}+\delta_t \tag{23}$$

由预测误差、条件分布迁移和共享－专属解耦构成的统一目标函数为

$$\mathcal{L}=\mathcal{L}_{\text{pred}}^{T}+\lambda_s\sum_{k=1}^{K}\tilde{\alpha}_k(r)\mathcal{L}_{\text{pred}}^{S_k}+\lambda_c\mathcal{L}_{\text{cda}}+\lambda_o\mathcal{L}_{\text{orth}} \tag{24}$$

式(24)中，$\mathcal{L}_{\text{pred}}^{T}$为目标支持集预测误差损失，$\mathcal{L}_{\text{pred}}^{S_k}$为第$k$个源场站预测误差损失，$\mathcal{L}_{\text{cda}}$为由条件分布距离$D_k^{\text{cond}}(r)$构成的对齐损失，$\mathcal{L}_{\text{orth}}$为式(11)定义的正交约束，$\lambda_s$、$\lambda_c$、$\lambda_o$为相应非负权衡系数，$\tilde{\alpha}_k(r)$为式(21)定义的有效源权重。本文轻量仿真实现以固定特征划分和显式统计映射实现条件对齐与表征约束，联合目标仅包含式(24)所列各项。需要说明的是，式中联合目标用于统一描述预测误差与各迁移约束之间的关系；在本文轻量仿真实现中，条件对齐与共享－专属约束主要通过预先定义的特征划分和统计映射显式实现，并不与岭回归系数进行端到端联合梯度优化。

$$\theta_t^*(\psi)=\arg\min_{\theta_t}\mathcal{L}_{\text{sup}}(\theta_0,\theta_t;\psi),\qquad \psi^*=\arg\min_{\psi}\mathcal{L}_{\text{val}}\left(\theta_0,\theta_t^*(\psi);\psi\right) \tag{25}$$

本文采用支持集与验证集分层的参数选择机制。θ_0表示由源场站数据获得的基础参数，θ_t表示利用目标支持集更新的目标适配参数，ψ表示源权重、条件对齐和门控等迁移策略参数。利用目标支持集更新少量目标适应参数，并利用目标验证集评价迁移策略参数，以避免迁移策略仅适配训练样本。推理阶段固定基础预测器与迁移策略，仅调用固化的源域模型参数、场景统计量和目标场站残差适配参数。

式(25)的轻量实现流程分为四步：① 利用源域训练集和源域验证集确定基础预测参数θ_0，进入目标场站适应后保持固定；② 利用目标支持集估计单参数残差偏置δ_t，即目标适配参数θ_t；③ 利用目标验证集评价候选迁移策略，确定由场景源权重、门控统计、偏置收缩强度和动态迁移－等权锚点融合系数组成的策略参数ψ；④ 参数确定后，利用完整目标适应池重新估计δ_t，并固定θ_0、θ_t和ψ完成测试预测。该过程采用分层参数选择，不进行高阶可微超梯度求解。

4. 仿真实验与结果分析

4.1. 实验设置

实验数据由美国风电机组数据库(United States Wind Turbine Database, USWTDB)和 WIND Toolkit

Data V2 整理获得。USWTDB 用于筛选美国得克萨斯州西部六个风电场及其地理中心，WIND Toolkit Data V2 提供相邻 18 个网格点的 15 min 多高度气象序列；每个风电场取距离中心最近的 3 个唯一网格点，对标量变量求算术平均，对风向采用单位向量平均。数据时间范围为 2011~2013 年。输入窗口包含过去 24 个 15 min 时刻，即 6 h；15 min 预测为一步超前，60 min 预测为四步超前。2011 年至 2012 年 10 月数据用于源模型训练，2012 年 11~12 月用于源模型验证，2013 年数据按时间顺序构造目标支持集、验证集和后续测试集。数据处理不进行统计插补，特征构造后剔除包含非有限值的窗口；标准化参数仅由相应训练数据估计，以避免测试信息泄漏。

实验采用留一场站法，依次将六个场站设为少样本目标域，其余场站作为候选源域。目标支持集规模设为 1、3、7、14 和 30 d，并按时间顺序以 80%和 20%划分支持集与验证集，随后一个月作为测试集。为突出多源迁移策略本身而降低复杂预测主干对结果的影响，本文采用标准化岭回归作为基础预测器，正则系数由验证集选择。输入包含多高度风速、历史滞后、24 步滚动统计、温度、气压、湿度、风切变、风向周期编码及时刻周期编码，共 23 个特征。气象场景由 4 个等宽风向扇区(分别以 0˚、90˚、180˚和 270˚为中心)、3 个风速区间和 2 个稳定度状态组合为 24 类；风速区间划分为小于 5 m/s、5~10 m/s 和不小于 10 m/s，稳定度由相邻时刻风速变化量是否达到 0.5 m/s 区分。综合风险系数依次为 0.55、0.25、0.10 和 0.10，温度 $T=0.35$，仅保留风险最低的两个候选源；门控系数 γ_1、γ_2、γ_3 及偏置 b 分别取 2.5、1.0、0.5 和 0。理论软匹配中的 σ_r、σ_v 均表示正带宽；仿真实现采用离散场景精确匹配，因而不另设带宽值。条件对齐仅作用于温度、气压、湿度、风切变及风向周期编码等共享气象特征，并采用 15%的统计映射强度；目标专属适配器仅包含一个残差偏置参数，占轻量预测器可训练参数的 4.17%。评价指标包括 MAE、RMSE、sMAPE、R^2、场景平均迁移增益和负迁移率。

温度 T 控制 Softmax 源权重集中度，综合风险权重 β 控制预测误差、分布距离、不确定性和冲突度的相对贡献。本文在原 120 个目标测试场景上设置 $T\in\{0.20,0.35,0.50,0.75\}$，并以 $\beta=(0.55,0.25,0.10,0.10)$ 为基准，对四个分量分别实施 ± 20%的单因素扰动，扰动后重新归一化；所有配置沿用相同训练–验证–测试划分，不根据测试结果重新选择参数。场站异质性依据目标场站至候选源场站的平均地理距离和共享特征分布距离的标准分数等权合成，并按三分位数划分为低、中、高三组。预测主干兼容性实验采用过去 24 个时刻的 9 维气象序列和单层紧凑 Transformer，仅选取 7 d 目标样本场景，以控制补充实验规模。

设 N 为全部目标测试场景数，场景平均迁移增益定义为

$$G_{\mathrm{tr}}=\frac{1}{N}\sum_{n=1}^{N}\left[\frac{\mathrm{MAE}_{n,\mathrm{target}}-\mathrm{MAE}_{n,\mathrm{transfer}}}{\mathrm{MAE}_{n,\mathrm{target}}}\right]\times 100\% \tag{26}$$

该指标先针对每个目标测试场景分别计算迁移增益，再对全部场景求平均，因此不等于表 1 中平均 MAE 之间直接计算得到的相对下降率。负迁移率定义为

$$R_{\mathrm{neg}}=\frac{1}{N}\sum_{n=1}^{N}\mathrm{I}\left(\mathrm{MAE}_{n,\mathrm{transfer}}>\mathrm{MAE}_{n,\mathrm{target}}\right)\times 100\% \tag{27}$$

其中，$\mathrm{I}(\cdot)$ 为示性函数。

4.2. 总体预测性能

表 1 给出了 120 个目标测试场景上的平均结果。本文方法在 15 min 和 60 min 任务上的 MAE 分别为 0.448 m/s 和 0.972 m/s，相对目标域独立训练的场景平均迁移增益分别为 37.34%和 37.00%。相较静态风险加权，MAE 分别降低 0.74%和 1.84%。多源等权迁移在 15 min 任务与本文方法基本相当，在 60 min 任

务上的平均 MAE 略低于本文方法。因此，本文方法的主要优势体现为相对静态风险加权方案的稳定改进以及对低价值源迁移影响的约束，而非在全部场景下均取得最低平均误差。

Table 1. Overall prediction results of the main method
表 1. 主要方法的总体预测结果

方法	15 min MAE	15 min RMSE	15 min R^2	60 min MAE	60 min RMSE	60 min R^2
持续性预测	0.455	0.778	0.943	1.012	1.521	0.791
目标域独立训练	0.978	1.290	0.801	2.168	2.676	−0.080
最近场站单源迁移	0.449	0.766	0.945	0.971	1.445	0.814
多源等权迁移	0.448	0.766	0.945	0.969	1.443	0.814
静态风险加权	0.452	0.769	0.944	0.990	1.464	0.810
气象场景风险加权	0.451	0.769	0.944	0.986	1.459	0.811
本文方法	0.448	0.767	0.945	0.972	1.446	0.814

多源等权迁移的 MAE 分别为 0.448 m/s 和 0.969 m/s，说明同一气候区域内源场站差异较小时，等权集成仍是较强基线。本文方法在保持总体误差接近该基线的同时，通过场景风险、门控与残差适配提供更明确的负迁移控制。

4.3. 目标样本规模与场景效应

表 2 比较了不同目标样本规模下三类代表性方法。在 1~3 d 极少样本条件下，目标独立模型误差受样本时段和气象状态影响较大，并未呈严格单调变化；两类迁移方法保持较低且稳定的 MAE。目标样本达到 7 d 以上后，目标独立模型的预测误差整体随样本规模增加而下降。

Table 2. MAE under different target sample sizes
表 2. 不同目标样本规模下的 MAE

目标样本/d	15 min MAE/(m/s)			60 min MAE/(m/s)		
	目标独立	多源等权	本文方法	目标独立	多源等权	本文方法
1	1.511	0.448	0.449	3.264	0.969	0.974
3	1.647	0.448	0.449	3.566	0.969	0.974
7	0.720	0.448	0.449	1.756	0.969	0.975
14	0.551	0.448	0.448	1.228	0.969	0.972
30	0.461	0.448	0.447	1.025	0.969	0.964

目标样本达到 30 d 时，本文方法在 15 min 和 60 min 任务上的 MAE 分别为 0.447 m/s 和 0.964 m/s，均低于多源等权迁移。该结果表明，目标验证信息增加后，场景风险估计和适配参数选择更为可靠；在极少样本条件下，验证收缩则使本文方法保持与等权迁移接近的误差水平。

4.4. 动态源权重与负迁移分析

表 3 同时列出负迁移率和本文方法相对各比较方法的配对 MAE 优势。本文方法在两个预测步长上的负迁移率均为 2.50%，即 120 个目标测试场景中仅有少量场景出现负迁移，低于静态风险加权和场景

风险门控在 60 min 任务上的 4.17%。在当前测试场景下，配对精确符号检验进一步表明本文方法相对静态风险加权具有较稳定的配对优势，两个预测步长的 P 值均为 0.0013。

Table 3. Transfer stability and advantage of paired error
表 3. 迁移稳定性与配对误差优势

方法	负迁移率/%		本文方法配对 MAE 优势/(m/s)	
	15 min	60 min	15 min	60 min
最近场站单源迁移	2.50	3.33	0.0004	−0.0008
多源等权迁移	2.50	2.50	−0.0004	−0.0029
静态风险加权	2.50	4.17	0.0033	0.0182
气象场景风险加权	2.50	3.33	0.0029	0.0143
场景风险加权与门控	2.50	4.17	0.0029	0.0147
本文方法	2.50	2.50	—	—

本文方法相对静态风险加权的配对 MAE 优势为 0.0033 m/s 和 0.0182 m/s，相对气象场景风险加权的优势为 0.0029 m/s 和 0.0143 m/s。相对多源等权迁移的优势为负，进一步说明场站同质性限制了动态选源的总体收益；上述检验结论仅针对当前测试场景。

4.5. 消融实验与讨论

表 4 给出了完整方法与各消融方案的结果。去除气象场景、目标专属适配器以及门控与冲突抑制后，15 min 任务的 MAE 分别上升 0.53%、0.64%和 0.51%，60 min 任务分别上升 1.33%、1.51%和 1.15%，说明三类机制在当前实验中产生正向贡献，且较长预测步长下作用更明显。去除条件特征对齐后，两项任务的误差反而分别下降 0.17%和 0.45%。

Table 4. Ablation experiment results
表 4. 消融实验结果

方法	15 min		60 min	
	MAE/(m/s)	相对变化/%	MAE/(m/s)	相对变化/%
本文方法	0.448	0.00	0.972	0.00
去除气象场景条件	0.451	0.53	0.985	1.33
去除条件特征对齐	0.448	−0.17	0.968	−0.45
去除专属适配器	0.451	0.64	0.987	1.51
去除门控与冲突抑制	0.451	0.51	0.983	1.15

消融结果表明，显式条件对齐在同一区域高同质性场站上未取得净收益，说明该机制更适用于源–目标差异较明显的场景，其对齐强度应依据场站异质程度进行调整。本文实验保留该模块用于验证方法边界，并未将其解释为当前数据上的精度增益来源。

4.6. 关键超参数敏感性分析

表 5 给出了温度参数和综合风险权重的单因素敏感性结果。基准复算与原实验逐场景 MAE 的最大

差异为 4.994×10^{-9}，表明补充分析与原结果口径一致。T 在 0.20~0.75 之间变化时，15 min 和 60 min 平均 MAE 相对基准的最大绝对变化分别为 0.0115%和 0.0019%；四项风险权重分别扰动±20%后，两项任务的最大绝对变化分别为 0.0097%和 0.0059%。全部配置下负迁移率均保持 2.50%，说明在所检验范围内模型对关键超参数扰动不敏感。

Table 5. Sensitivity analysis of key hyperparameters
表 5. 关键超参数敏感性分析

调整项	设置	15 min MAE/(m/s)	15 min 相对变化/%	60 min MAE/(m/s)	60 min 相对变化/%
温度 T	0.20	0.448372	+0.0115	0.971839	−0.0019
温度 T	0.35	0.448320	+0.0000	0.971857	+0.0000
温度 T	0.50	0.448339	+0.0042	0.971864	+0.0007
温度 T	0.75	0.448340	+0.0044	0.971872	+0.0016
风险项 β_1	±20%	0.448277~0.448326	≤0.0097	0.971805~0.971907	≤0.0053
距离项 β_2	±20%	0.448313~0.448322	≤0.0016	0.971800~0.971907	≤0.0059
不确定性项 β_3	±20%	0.448310~0.448322	≤0.0024	0.971853~0.971866	≤0.0009
冲突项 β_4	±20%	0.448320~0.448321	≤0.0001	0.971838~0.971856	≤0.0019

4.7. 场站异质性与预测主干适用性

为考察场站差异对动态选源效果的影响，依据 4.1.1 节定义的综合异质性指标对 120 个目标测试场景分层。由表 6 可见，本文方法相对静态风险加权的 MAE 降幅随异质性提高而增大：低异质性组在 15 min 和 60 min 任务上的降幅分别为 0.0513%和 0.0181%，高异质性组扩大至 1.8771%和 4.7688%。该结果与场景风险控制更适合源–目标差异明显情形的设计动机一致。多源等权在各层仍保持较强竞争力，且当前分层来自同一区域数据，不能替代跨区域、多气候数据集验证。

Table 6. Results under different station heterogeneity levels
表 6. 不同场站异质性等级下的 MAE

异质性等级	预测时域	多源等权/(m/s)	静态风险加权/(m/s)	本文方法/(m/s)	相对静态降幅/%
低	15 min	0.447732	0.448279	0.448049	0.0513
低	60 min	0.973496	0.975754	0.975577	0.0181
中	15 min	0.465391	0.467141	0.465621	0.3253
中	60 min	0.976570	0.984212	0.977935	0.6378
高	15 min	0.430543	0.439542	0.431291	1.8771
高	60 min	0.956910	1.010235	0.962059	4.7688

进一步将本文的场景风险、门控、冲突抑制和单参数残差适配模块接入紧凑 Transformer 主干。该主干使用过去 24 个 15 min 时刻的多高度风速、温度、气压、湿度、风切变和风向编码，共 9 维序列输入；编码层设置 16 维隐状态、2 个注意力头和 1 层 Transformer 编码器。为控制补充实验规模，仅采用 7 d 目标样本的 24 个场景。表 7 显示，相对 Transformer 目标独立训练，结合本文迁移框架后 15 min 和 60 min MAE 分别下降 68.54%和 52.02%。与 Transformer 多源等权相比，本文框架在 15 min 任务上 MAE 略高，

在 60 min 任务上略低，表明该策略能够接入非线性时序主干并缓解少样本训练不足，但不意味着在所有预测时域均优于等权集成。

Table 7. Compatibility results with a compact Transformer backbone
表 7. 紧凑 Transformer 预测主干兼容性结果

方法	15 min MAE/(m/s)	15 min RMSE/(m/s)	60 min MAE/(m/s)	60 min RMSE/(m/s)
Transformer 目标独立	1.871	2.463	2.156	2.808
Transformer 多源等权	0.578	0.913	1.035	1.508
Transformer 结合本文迁移框架	0.589	0.928	1.034	1.507

5. 结论

针对新建风电场少样本条件下的多源知识复用与负迁移问题，本文提出一种基于气象场景条件风险最小化的多源选择性迁移风速预测方法。该方法综合源域经验风险、条件分布差异、预测不确定性和更新方向冲突生成场景化动态权重，并结合共享–专属表征、迁移门控和单参数残差适配实现目标场站快速适应。基于六个得克萨斯州西部风电场的留一场站实验表明，本文方法在 15 min 和 60 min 任务上的 MAE 分别为 0.448 m/s 和 0.972 m/s，相对目标域独立训练的场景平均迁移增益分别为 37.34%和 37.00%，负迁移率均为 2.50%；相较静态风险加权取得了稳定的小幅改善。目标专属适配器仅包含 1 个可训练参数，占轻量预测器可训练参数的 4.17%，表明该方法能够以较低的适配成本缓解目标场站极少样本问题。

在 $T = 0.20 \sim 0.75$ 以及各综合风险权重单因素 ± 20%的范围内，平均 MAE 的最大相对变化不超过 0.012%，负迁移率保持 2.50%，说明该方法对关键参数扰动具有较好的稳定性。随着场站异质性提高，本文方法相对静态风险加权的优势进一步增大；紧凑 Transformer 实验也表明，所提出的迁移策略能够与非线性时序预测主干结合。然而，在同一区域高同质性场站中，多源等权迁移仍具有较强竞争力，固定强度的条件对齐亦未获得净收益，说明动态源选择更适用于源域质量差异明显或局部气象状态变化显著的场景。后续研究将面向地理、地形和气候差异更显著的多区域风电场开展验证，并研究基于目标验证集的温度参数、综合风险权重及条件对齐强度自适应确定方法。

参考文献

[1] Foley, A.M., Leahy, P.G., Marvuglia, A. and McKeogh, E.J. (2012) Current Methods and Advances in Forecasting of Wind Power Generation. *Renewable Energy*, **37**, 1-8. https://doi.org/10.1016/j.renene.2011.05.033

[2] Hu, Q.H., Zhang, R.J. and Zhou, Y.C. (2016) Transfer Learning for Short-Term Wind Speed Prediction with Deep Neural Networks. *Renewable Energy*, **85**, 83-95. https://doi.org/10.1016/j.renene.2015.06.034

[3] 程凯, 彭小圣, 徐其友, 等. 基于特征选择与多层级深度迁移学习的风电场短期功率预测[J]. 高电压技术, 2022, 48(2): 497-503.

[4] Peng, X.S., Yang, Z.M., Li, Y., *et al*. (2023) Short-Term Wind Power Prediction Based on Stacked Denoised Auto-encoder Deep Learning and Multi-Level Transfer Learning. *Wind Energy*, **26**, 1066-1081. https://doi.org/10.1002/we.2856

[5] Ben-David, S., Blitzer, J., Crammer, K., Kulesza, A., Pereira, F. and Vaughan, J.W. (2010) A Theory of Learning from Different Domains. *Machine Learning*, **79**, 151-175. https://doi.org/10.1007/s10994-009-5152-4

[6] Gretton, A., Borgwardt, K.M., Rasch, M.J., *et al*. (2012) A Kernel Two-Sample Test. *Journal of Machine Learning Research*, **13**, 723-773.

[7] Long, M.S., Cao, Y., Wang, J.M., *et al*. (2015) Learning Transferable Features with Deep Adaptation Networks. *Proceedings of the* 32*nd International Conference on Machine Learning*, Lille, 6-11 July 2015, 97-105.

[8] Ganin, Y., Ustinova, E., Ajakan, H., *et al*. (2016) Domain-Adversarial Training of Neural Networks. *Journal of Machine*

Learning Research, **17**, 1-35.

[9] Bousmalis, K., Trigeorgis, G., Silberman, N., *et al*. (2016) Domain Separation Networks. *Advances in Neural Information Processing Systems* 29, Barcelona, 5-10 December 2016, 343-351.

[10] Houlsby, N., Giurgiu, A., Jastrzebski, S., *et al*. (2019) Parameter-Efficient Transfer Learning for NLP. 2019 *Proceedings of the* 36*th International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 2790-2799.

[11] Yu, T.H., Kumar, S., Gupta, A., *et al*. (2020) Gradient Surgery for Multi-Task Learning. 2020 *Advances in Neural Information Processing Systems* 33, Online, 6-12 December 2020, 5824-5836.