LEARNING WHAT TO TRANSFER? SELF-SUPERVISED DISENTANGLEMENT OF TRANSFERABLE SEMANTICS FOR SOURCE-FREE MULTIVARIATE TIME SERIES DOMAIN ADAPTATION

Anonymous authors
Paper under double-blind review

ABSTRACT

We consider source-free domain adaptation (SFDA), which transfers a source-trained model to an unlabeled target domain without retaining source samples. For multivariate time series (MTS), target self-supervision can reveal temporal, spectral, and channel-relation structure, but consistency within the target domain does not necessarily imply relevance to the source task. Without source data for direct verification, indiscriminately transferring such structure may retain target-specific variation at the expense of task-relevant semantics. We propose **self-supervised transferable semantic disentanglement (STSD)** for source-free MTS domain adaptation. During source training, we decompose temporal, spectral, and channel-relation views into a shared semantic representation and view-specific components, and calibrate normalized classifier directions as persistent class anchors. We then use the frozen anchors to form a source semantic scatter operator and disagreement among unlabeled target views to form a target self-supervised variation operator. STSD solves a generalized spectral problem whose leading eigenspace preserves source class separation while suppressing target cross-view disagreement. The resulting eigenspace induces complementary transferable and target-local projectors, so semantic disentanglement follows from the geometry of the construction rather than from an auxiliary decoupling loss. Target prototypes are matched with the frozen source anchors only in the transferable subspace through entropic optimal transport. The target encoder is finally updated with a single semantic objective, while the spectral subspace and transport plan are refreshed from the current target representations. Extensive experiments on three standard MTS benchmarks demonstrate the effectiveness and superiority of STSD. The code is publicly available at `https://github.com/Sky-byte-box/STSD_Code`.

1 INTRODUCTION

Multivariate time series (MTS) collected from multiple sensors or variables often suffer from distribution shifts across domains, caused by variations in temporal patterns, channel relationships, acquisition conditions, and subject characteristics. Multivariate time series domain adaptation aims to transfer task-relevant knowledge from labeled source domains to unlabeled target domains while preserving intrinsic temporal and inter-channel structures. Conventional unsupervised domain adaptation methods jointly access source and target data during adaptation (Ganin et al., 2016; Long et al., 2018; He et al., 2023), which limits their applicability when source data are unavailable. This motivates source-free domain adaptation (SFDA), where adaptation relies only on target data and a source-trained model (Liang et al., 2020; 2022; Li et al., 2024).

Existing multivariate time series domain adaptation methods mainly reduce domain shift by aligning representations while preserving sequence structure. AdaTime (Ragab et al., 2023a) provides unified evaluation and demonstrates the competitiveness of tuned general-purpose baselines. Convolutional deep domain adaptation for time series (CoDATS) (Wilson et al., 2020) introduces adversarial adaptation. Self-supervised autoregressive domain adaptation (SLARDA) (Ragab et al., 2024) adds autoregressive learning, while contrastive domain adaptation for time series via temporal

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

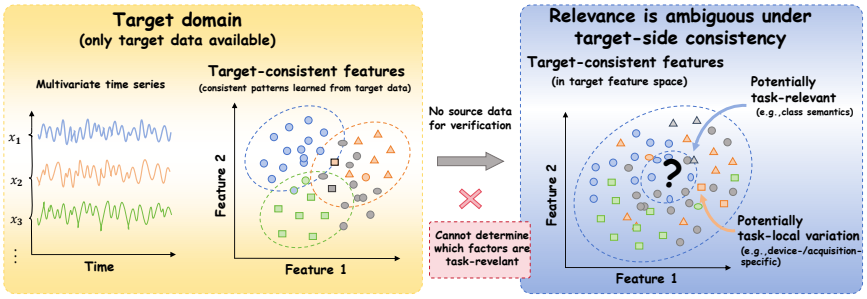

Figure 1: Target-side consistency does not necessarily imply relevance to the source task, while source data are unavailable for direct verification.

mixup (COTMix) (Eldele et al., 2023) combines temporal mixup with contrastive learning. Domain adaptation for time series under feature and label shifts (RAINCOAT) (He et al., 2023) integrates temporal and spectral information, while sensor alignment (SEA) (Wang et al., 2023) and SEA++ (Wang et al., 2024b) model sensor distributions and cross-channel interactions. Multi-source methods, including contrastive adversarial learning for multi-source time series domain adaptation (CALDA) (Wilson et al., 2023) and prompt-based domain discrimination (POND) (Wang et al., 2024a), exploit shared information across labeled sources. These approaches extend global alignment toward temporal, spectral, and channel-aware transfer but still require source observations during adaptation.

Source-free adaptation removes the need for source observations by reusing knowledge retained in a trained model. Source hypothesis transfer (SHOT) (Liang et al., 2020) and neighborhood reciprocity clustering (NRC) (Yang et al., 2021) adapt target representations using a frozen source hypothesis or target neighborhood structure without revisiting source samples. For time series, mask and impute (MAPU) (Ragab et al., 2023b) transfers temporal imputation knowledge to guide target features, while temporal restoration and spatial rewiring (TERSE) (Gong et al., 2025b) combines temporal restoration with inter-channel structural modeling. Augmented contrastive clustering with uncertainty-aware prototyping (ACCUP) (Gong et al., 2025a) uses uncertainty-aware prototypes and contrastive clustering for online target adaptation. These methods show how source-model knowledge and target-side structure can jointly support source-free MTS adaptation.

It should be noted that existing source-free MTS adaptation methods mainly rely on target-side self-supervision or structural consistency to guide adaptation after source samples become unavailable. However, although a temporal or cross-channel pattern can be highly consistent within the target domain, such consistency does not necessarily indicate that the pattern is relevant to the source task. In other words, target-side regularity alone cannot distinguish *transferable class semantics* from *target-local variation*, such as device- or acquisition-specific characteristics. Directly transferring these factors may reinforce domain-specific structure and gradually move target representations away from the source decision geometry. As illustrated in Fig. 1, the difficulty is therefore not merely how to align target features with a source model, but how to determine which semantic factors should actually be transferred in the absence of source data.

To address the above limitation, we propose **self-supervised transferable semantic disentanglement (STSD)** for source-free MTS domain adaptation, which is visually shown in Fig. 2. More concretely, we firstly construct temporal, spectral, and channel-relation views during source training, decompose them into shared semantics and view-specific components, and calibrate normalized classifier directions as persistent class anchors. Secondly, once source samples become unavailable, we use the frozen anchors to characterize source semantic separation and the unlabeled target views to measure cross-view disagreement, and couple the two cues through a generalized spectral problem to identify the transferable semantic subspace. Thirdly, the resulting eigenspace induces complementary transferable and target-local projectors, so only the transferable component is used to form confidence-weighted target prototypes and establish source-guided semantic correspondence through entropic transport. Finally, the transported anchors guide the target encoder through a unified adaptation objective, while the transferable subspace and semantic correspondence are refreshed from the current target representations.

Our contributions are four-fold. *First*, we propose self-supervised transferable semantic disentanglement (STSD) for source-free MTS domain adaptation, which learns a shared multi-view semantic basis and calibrates classifier directions as persistent class anchors without retaining source samples.

Second, we develop a generalized spectral criterion that jointly couples source class separation with target cross-view disagreement to identify transferable semantics and their orthogonal target-local complement. *Third*, we establish confidence-weighted prototype correspondence and entropic semantic transport only in the transferable subspace, followed by a unified target objective with a margin-preservation result characterizing retained class separation. *Fourth*, extensive experiments on three standard MTS benchmarks demonstrate the effectiveness and superiority of STSD.

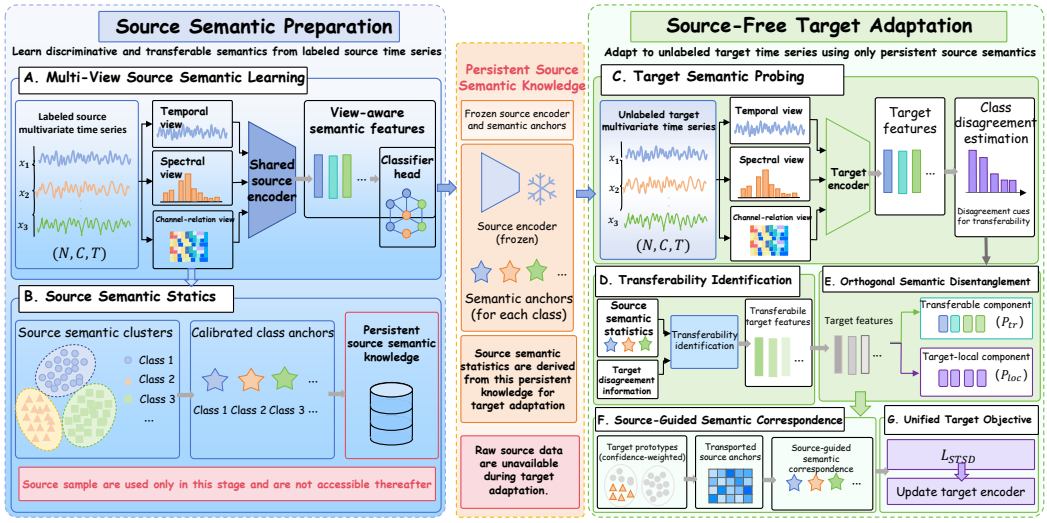

Figure 2: Overview of our proposed STSD for source-free multivariate time series domain adaptation.

2 Related Work

2.1 Self-Supervised Time-Series Representation Learning

Self-supervised learning is well suited to MTS, drawing supervision from temporal continuity, masked observations, and complementary views. TNC learns representations through local temporal neighborhoods (Tonekaboni et al., 2021), while CoST disentangles seasonal and trend components across time and frequency domains (Woo et al., 2022). SimMTM reconstructs masked observations using manifold neighbors to capture structural information (Dong et al., 2023). More recent foundation models, including including LLM-based Time-LLM (Jin et al., 2024), Timer (Liu et al., 2024), MOMENT (Goswami et al., 2024), Chronos (Ansari et al., 2024), TimeMixer++ (Wang et al., 2025), and MoE-based Time-MoE (Shi et al., 2025), demonstrate the transferability of temporal representations across tasks. These approaches exploit complementary temporal, spectral, and relational cues to enrich representations beyond a single view. However, learning such representations does not by itself establish their robustness to domain shift.

2.2 Semantic Representation Learning

Semantic representation learning organizes task-relevant features while separating shared information from view-specific factors. Prototype-based methods capture category structure without requiring instance-to-instance correspondence. PCS (Yue et al., 2021) combines in-domain prototype learning with cross-domain instance-to-prototype matching. Supervised contrastive learning (Khosla et al., 2020) forms class-consistent neighborhoods, while SwAV (Caron et al., 2020) learns prototype assignments without labels. Multi-view approaches further examine how task information is distributed across complementary observations. FactorCL (Liang et al., 2023) separates shared and view-unique task information, while multimodal domain generalization research (Fan et al., 2024) explores knowledge transfer across representation landscapes. Together, semantic organization and factorization provide complementary ways to structure class-related information. These methods primarily address how representations are organized or decomposed, leaving open how the resulting semantic structure should be reused under distribution shift.

3 METHODOLOGY

3.1 PROBLEM DEFINITION

Source-free adaptation separates source training from target adaptation. The source domain is used only to learn the initial task model and is unavailable after deployment, so adaptation must rely on the trained source hypothesis rather than direct source–target sample matching. Let the labeled source domain be $\mathcal{D}_s = \{(X_i^s, y_i^s)\}_{i=1}^{n_s}$ and the unlabeled target domain be $\mathcal{D}_t = \{X_i^t\}_{i=1}^{n_t}$, where each $X \in \mathbb{R}^{C\times L}$ contains C synchronized channels and L time steps, and $y \in \{1, \dots, K\}$. Source training learns an encoder f_{θ_s} and a cosine classifier g_W. Target adaptation then accesses only $\mathcal{D}_t$ and the exported parameters (θ_s, W), as formulated in Eq. 1 below:

$$(\theta_s, W) = \arg\min_{\theta,W} \frac{1}{n_s} \sum_i \ell_{\text{ce}}(g_W(f_\theta(X_i^s)), y_i^s), \qquad \mathcal{D}_s \notin \mathcal{A}_t, \tag{1}$$

where ℓ_{ce} denotes the cross-entropy loss, g_W is the K-class source classifier, and $\mathcal{A}_t$ denotes the information available during target adaptation, excluding $\mathcal{D}_s$. The goal is to learn θ_t without reconstructing or replaying source samples. As illustrated in Fig. 2, STSD first learns a source semantic basis, then identifies transferable semantics from target-view consistency and source-anchor geometry, and uses semantic correspondence within the resulting subspace to guide target adaptation.

3.2 SOURCE SEMANTIC BASIS LEARNING

3.2.1 MULTI-VIEW CONSTRUCTION AND FACTORIZATION

Given a sequence X, we construct temporal, spectral, and relation views. The temporal view $V^{\text{tmp}}(X)$ uses patching with light masking, while $V^{\text{spc}}(X)$ encodes channel-wise Fourier magnitude and phase. The relation view $V^{\text{rel}}(X)$ is built from normalized channel affinity $G(X)_{ab} = \langle \bar{x}_a, \bar{x}_b \rangle / (\|\bar{x}_a\|_2 \, \|\bar{x}_b\|_2 + \epsilon)$ with $\bar{x}_c = x_c - L^{-1} \sum_{\ell=1}^{L} x_{c\ell}$, where a, b, c index channels, $x_c \in \mathbb{R}^L$ denotes channel c, and $\epsilon > 0$ is a numerical stabilizer. For $v \in \mathcal{V} = \{\text{tmp}, \text{spc}, \text{rel}\}$, the backbone produces $h_i^v = f_\theta(V^v(X_i))$. Shared semantics are then aggregated across the three views as:

$$s_i = \text{norm}\left(\frac{1}{|\mathcal{V}|} \sum_{v\in\mathcal{V}} P_s h_i^v \right), \tag{2}$$

where d_h is the backbone feature dimension, D is the semantic embedding dimension, $P_s, P_v \in \mathbb{R}^{D\times d_h}$ denote the shared and view-specific projectors, $s_i^v = P_s h_i^v$ is the shared semantic representation of view v, $r_i^v = P_v h_i^v$ is its view-specific residual, and $\text{norm}(z) = z/(\|z\|_2 + \epsilon)$. The source prediction is $p_i = \text{softmax}(W^\top s_i / \tau_c)$, where $W = [w_1, \dots, w_K]$ and $\tau_c > 0$ is the source temperature. Agreement is imposed only on s_i^v, leaving r_i^v to retain complementary view-specific information.

3.2.2 CLASSIFIER-ANCHOR CALIBRATION

The shared basis must preserve class information while remaining separate from view-specific content. We use $\mathcal{R}_{\text{fac}}$ to penalize the mean cross-covariance $\| \text{Cov}(S^v, R^v)\|_F^2$ and the pairwise discrepancy $\|S^v - S^u\|_F^2$ between shared views. Classifier directions may still deviate from this semantic geometry, so each normalized anchor $a_k = \text{norm}(w_k)$ is calibrated toward the source mini-batch prototype as follows: $q_k^s = \text{norm}(\sum_i \mathbb{1}[y_i^s = k] s_i / (\sum_i \mathbb{1}[y_i^s = k] + \epsilon))$. The corresponding anchor penalty $\mathcal{R}_{\text{anc}}$ combines prototype attraction $1 - a_k^\top q_k^s$ with the squared hinge $[a_k^\top a_j - m_a]_+^2$ across different classes. Source pre-training is written as:

$$\mathcal{L}_s = \frac{1}{n_s} \sum_i -\log p_{i,y_i^s} + \lambda_{\text{fac}} \mathcal{R}_{\text{fac}} + \lambda_{\text{anc}} \mathcal{R}_{\text{anc}}, \tag{3}$$

where S^v and R^v stack $\{s_i^v\}$ and $\{r_i^v\}$, $u, v \in \mathcal{V}$ index distinct views, $\text{Cov}(\cdot,\cdot)$ denotes empirical cross-covariance, $\mathbb{1}[\cdot]$ denotes the indicator function, and $[x]_+ = \max(x, 0)$. Here $\mathcal{R}_{\text{fac}}$ averages the cross-covariance and pairwise shared-view discrepancy described above, while $\mathcal{R}_{\text{anc}}$ averages prototype attraction and inter-class anchor separation over source classes. The margin m_a controls inter-anchor cosine similarity, and $\lambda_{\text{fac}}, \lambda_{\text{anc}} \geq 0$ are source-stage coefficients. After source training, the samples and mini-batch prototypes are discarded. Only $(\theta_s, W, P_s, \{P_v\})$ are retained for target adaptation.

3.3 Self-Supervised Identification of Transferable Semantics

Transferability requires a direction to preserve source class separation and remain consistent across unlabeled target views. We express both properties in the shared semantic basis. The frozen anchors define the source class scatter $S_s = K^{-1}\sum_k (a_k - \bar{a})(a_k - \bar{a})^\top$, where $\bar{a} = K^{-1}\sum_k a_k$, while target views define the disagreement operator below:

$$D_t = \frac{1}{n_t|\mathcal{V}|}\sum_{i,v}(s_i^v - \bar{s}_i)(s_i^v - \bar{s}_i)^\top + \epsilon_g I_D, \tag{4}$$

where $\bar{s}_i = |\mathcal{V}|^{-1}\sum_v s_i^v$, $I_D \in \mathbb{R}^{D\times D}$ is the identity matrix, and $\epsilon_g > 0$ is the spectral ridge. For a direction u, a large $u^\top S_s u$ indicates strong source class separation, while a small $u^\top D_t u$ indicates consistent target behavior across views. A transferable direction should satisfy both. We therefore identify the transferable subspace by contrasting these two quantities in the generalized spectral problem below:

$$U_{\mathrm{tr}}^\star = \arg\max_{\substack{U\in\mathbb{R}^{D\times r}\\ \mathrm{rank}(U)=r}} \mathrm{Tr}\big[(U^\top D_t U)^{-1}(U^\top S_s U)\big], \tag{5}$$

where $U \in \mathbb{R}^{D\times r}$ is a candidate transferable basis and the generalized eigenvalues of $S_s u = \lambda D_t u$ satisfy $\lambda_1 \geq \cdots \geq \lambda_D \geq 0$. We choose the smallest r whose cumulative generalized-eigenvalue energy reaches $\xi \in (0, 1]$. The resulting $U_{\mathrm{tr}}^\star$ spans the leading generalized eigenspace. To clarify how source discrimination and target consistency interact in this solution, consider any nonzero semantic direction u. Eq. (5) then reduces to $\mathfrak{t}(u) = u^\top S_s u/(u^\top D_t u)$, which measures source class separation relative to target disagreement. The two criteria are therefore coupled within the same spectral quantity rather than evaluated independently or combined by an additional target coefficient.

3.4 Transferable Semantic Disentanglement

The generalized eigenspace specifies which directions are transferable, while the remaining variation should not enter source-guided matching. Because coordinate-wise gating does not preserve the learned subspace geometry, we use complementary orthogonal projectors to separate the two components as following:

$$P_{\mathrm{tr}} = U(U^\top U)^{-1}U^\top, \qquad P_{\mathrm{loc}} = I_D - P_{\mathrm{tr}}, \tag{6}$$

where $U = U_{\mathrm{tr}}^\star$, and P_{tr} and P_{loc} project onto the transferable subspace and its orthogonal complement. Accordingly, $z_i^{\mathrm{tr}} = P_{\mathrm{tr}}\bar{s}_i$ and $z_i^{\mathrm{loc}} = P_{\mathrm{loc}}\bar{s}_i$ satisfy $\bar{s}_i = z_i^{\mathrm{tr}} + z_i^{\mathrm{loc}}$. Since $P_{\mathrm{tr}}P_{\mathrm{loc}} = 0$, the two components are orthogonal. The local component remains in the representation but does not participate in source-guided correspondence.

3.4.1 Transferable predictions and self-supervised confidence

After decomposition, source anchors and target features are compared only in the transferable subspace. We use $a_k^{\mathrm{tr}} = \mathrm{norm}(P_{\mathrm{tr}}a_k)$ and $z_i^{v,\mathrm{tr}} = \mathrm{norm}(P_{\mathrm{tr}}s_i^v)$, with $A^{\mathrm{tr}} = [a_1^{\mathrm{tr}}, \ldots, a_K^{\mathrm{tr}}]$. Each view predicts $p_i^v = \mathrm{softmax}((A^{\mathrm{tr}})^\top z_i^{v,\mathrm{tr}}/\tau_t)$, and $\bar{p}_i = |\mathcal{V}|^{-1}\sum_v p_i^v$ denotes their mean. We measure cross-view disagreement by $d_i = |\mathcal{V}|^{-1}\sum_v \mathrm{JS}(p_i^v \,\|\, \bar{p}_i)$ and define the confidence

$$\omega_i = \exp(-d_i/\tau_d)\left(1 - \frac{H(\bar{p}_i)}{\log K}\right), \tag{7}$$

where $\tau_t > 0$ is the target prediction temperature, JS denotes Jensen–Shannon divergence, $H(p) = -\sum_k p_k \log p_k$, and $\tau_d > 0$ is the disagreement temperature. A sample receives a larger ω_i when its views agree and the mean prediction is concentrated, so uncertain samples contribute less to subsequent prototype estimation. Semantic correspondence is also restricted to the transferable subspace because P_{tr} and P_{loc} are complementary and orthogonal. Accordingly, the full-space distance between a target representation z and a source anchor a_y decomposes as $\|z - a_y\|_2^2 = \|z^{\mathrm{tr}} - a_y^{\mathrm{tr}}\|_2^2 + \|z^{\mathrm{loc}} - a_y^{\mathrm{loc}}\|_2^2$, where $z^{\mathrm{tr}} = P_{\mathrm{tr}}z$, $z^{\mathrm{loc}} = P_{\mathrm{loc}}z$, $a_y^{\mathrm{tr}} = P_{\mathrm{tr}}a_y$, and $a_y^{\mathrm{loc}} = P_{\mathrm{loc}}a_y$. The target-local term can therefore introduce additional mismatch in full-space matching and is excluded from the subsequent correspondence.

3.5 Semantic Correspondence in the Transferable Subspace

Confidence is defined at the sample level, whereas semantic correspondence requires class-level structure. We therefore aggregate transferable target features using the cross-view confidence and obtain the class prototype below:

$$q_k^t = \mathrm{norm}\left(\frac{\sum_i \omega_i \bar{p}_{ik} \hat{z}_i^{\mathrm{tr}}}{\sum_i \omega_i \bar{p}_{ik} + \epsilon}\right), \tag{8}$$

where $\hat{z}_i^{\mathrm{tr}} = \mathrm{norm}(z_i^{\mathrm{tr}})$, $\bar{p}_{ik}$ is the kth element of $\bar{p}_i$, q_k^t is the target prototype, and $\bar{\pi}_k^t = n_t^{-1} \sum_i \bar{p}_{ik}$ denotes its estimated class mass. Here $Q^t = [q_1^t, \ldots, q_K^t]$, Q denotes the running prototype memory, and $\mu \in [0, 1)$ is the exponential-moving-average coefficient in $Q \leftarrow \mu Q + (1-\mu)Q^t$. Since target pseudo-label indices may be permuted during adaptation, direct same-index matching is unreliable. We therefore estimate their correspondence through entropic transport as follows:

$$\Pi^\star = \arg\min_{\Pi \in \mathcal{U}(\bar{\boldsymbol{\pi}}^t, \boldsymbol{\pi}^s)} \langle \Pi, C \rangle + \varepsilon \sum_{k,j} \Pi_{kj}(\log \Pi_{kj} - 1), \tag{9}$$

where $C = [C_{kj}] \in \mathbb{R}^{K \times K}$ with $C_{kj} = 1 - (q_k^t)^\top a_j^{\mathrm{tr}} + \rho |\bar{\pi}_k^t - \pi_j^s|$, $\bar{\boldsymbol{\pi}}^t = [\bar{\pi}_1^t, \ldots, \bar{\pi}_K^t]^\top$ and $\boldsymbol{\pi}^s = [\pi_1^s, \ldots, \pi_K^s]^\top$ are the target and source class-mass vectors, respectively, $\rho \geq 0$ weights their discrepancy, $\varepsilon > 0$ is the entropic regularizer, and $\langle \cdot, \cdot \rangle$ denotes the Frobenius inner product. The feasible set $\mathcal{U}(\bar{\boldsymbol{\pi}}^t, \boldsymbol{\pi}^s)$ contains $\Pi \in \mathbb{R}_+^{K \times K}$ satisfying $\Pi \mathbf{1}_K = \bar{\boldsymbol{\pi}}^t$ and $\Pi^\top \mathbf{1}_K = \boldsymbol{\pi}^s$, where $\mathbf{1}_K$ is the K-dimensional all-ones vector and $\pi_j^s = 1/K$ when no source prior is exported. The resulting plan induces the target-aligned anchor $\tilde{a}_k = \mathrm{norm}(\sum_j \Pi_{kj}^\star a_j^{\mathrm{tr}} / (\sum_j \Pi_{kj}^\star + \epsilon))$, where $\tilde{a}_k$ denotes the transported anchor associated with target prototype k. Since transport is solved after projection, both the cost matrix and transported anchors depend on the current transferable eigenspace. We next give a sufficient condition under which this correspondence preserves class separation.

Theorem 1 (Margin preservation in the transferable subspace)**.** *Assume projected source anchors satisfy $\left\|a_y^{\mathrm{tr}} - a_j^{\mathrm{tr}}\right\|_2 \geq \gamma$ for every $j \neq y$. For a target example of class y, assume $\|z^{\mathrm{tr}} - \tilde{a}_y\|_2 \leq \delta_z$ and $\|\tilde{a}_k - a_k^{\mathrm{tr}}\|_2 \leq \delta_a$ for all k. A sufficient condition is:*

$$\gamma > 2(\delta_z + \delta_a), \tag{10}$$

where $\delta_z \geq 0$ bounds target-to-transported-anchor deviation and $\delta_a \geq 0$ bounds transport-induced anchor deviation. Under this condition, nearest-anchor prediction assigns z^{tr} to class y.

Proof. By the triangle inequality, $\left\|z^{\mathrm{tr}} - a_y^{\mathrm{tr}}\right\|_2 \leq \delta_z + \delta_a$. For any $j \neq y$, $\left\|z^{\mathrm{tr}} - a_j^{\mathrm{tr}}\right\|_2 \geq \gamma - (\delta_z + \delta_a)$. Eq. (10) makes this lower bound larger than $\delta_z + \delta_a$, so the correct anchor is strictly closer. □

Anchor calibration supports inter-class separation through γ, while spectral identification removes weak or view-dependent directions and transport controls δ_a. This establishes the quantities used by the final target objective.

3.6 Unified Source-Free Adaptation

With the transferable geometry and semantic correspondence established, target adaptation no longer needs separate objectives for each intermediate step. Eq. (5) determines P_{tr}, while Eqs. (8) and (9) determine the prototypes and their correspondence, from which the transported anchors are obtained. We update the target encoder through:

$$\mathcal{L}_{\mathrm{STSD}}(\theta_t) = \frac{1}{n_t} \sum_i \omega_i \ell_{\mathrm{sem}}\Big(\mathrm{norm}(P_{\mathrm{tr}} \bar{s}_i), \tilde{A}, \hat{y}_i\Big), \tag{11}$$

where $\tilde{A} = [\tilde{a}_1, \ldots, \tilde{a}_K]$, $\hat{y}_i = \arg\max_k \bar{p}_{ik}$, and $\ell_{\mathrm{sem}}(z, \tilde{A}, y) = -\log[\exp(\tilde{a}_y^\top z/\tau_p) / \sum_k \exp(\tilde{a}_k^\top z/\tau_p)]$ with semantic temperature $\tau_p > 0$. The quantities ω_i, P_{tr}, $\Pi^\star$, and $\tilde{A}$ are refreshed from the current target representations at each outer step. Specifically, the target features update D_t, the generalized eigenspace updates P_{tr}, and the projected prototypes update $\Pi^\star$ by Sinkhorn transport. These quantities are held fixed during the subsequent

inner encoder updates. The target stage optimizes one objective whose geometry is updated by self-supervised identification and semantic correspondence. Algorithm 1 summarizes the procedure.

Algorithm 1 Source-free adaptation with STSD

Require: Target set $\mathcal{D}_t$, source encoder θ_s, frozen anchors A, energy threshold ξ, Sinkhorn regularization ε.

$\theta_t \leftarrow \theta_s$, initialize prototype memory Q.
for each target epoch **do**
  Construct three MTS views and extract shared semantic features s_i^v.
  Form source scatter S_s from anchors and target disagreement D_t by Eq. (4).
  Solve Eq. (5) and construct $P_{\mathrm{tr}}, P_{\mathrm{loc}}$ by Eq. (6).
  Project target views and anchors, then estimate self-supervised confidence by Eq. (7).
  Estimate transferable prototypes by Eq. (8) and solve semantic correspondence by Eq. (9).
  for each mini-batch $B \subset \mathcal{D}_t$ **do**
    Minimize the objective in Eq. (11) with P_{tr} and Π fixed for the inner update.
  end for
end for
return adapted encoder θ_t and frozen classifier A.

4 Experiments

4.1 Evaluation Settings

Datasets. We adopt the three real-world benchmarks used by recent source-free MTS domain adaptation studies (Ragab et al., 2023b; Gong et al., 2025b). The University of California Irvine human activity recognition (UCIHAR) dataset contains nine inertial channels and six activities across 30 subjects (Anguita et al., 2013). The wireless sensor data mining (WISDM) dataset contains tri-axial smartphone accelerometer signals for six activities across 36 subjects (Kwapisz et al., 2011). The Sleep-EDF sleep stage classification (SSC) benchmark contains physiological channels segmented into five sleep stages and exhibits substantial subject-level shift (Goldberger et al., 2000). Each subject is treated as a domain. To keep the main comparison compact, two source-to-target subject pairs are reported for each dataset, and macro F1 (MF1) is reported as mean$\pm$standard deviation.

Baselines. We compare STSD with established UDA methods and recent source-free domain adaptation (SFDA) methods under a controlled setting using the same MTS backbone and consistent training configurations. The UDA baselines include domain-adversarial training of neural networks (DANN) (Ganin et al., 2016), conditional adversarial domain adaptation (CDAN) (Long et al., 2018), CoDATS (Wilson et al., 2020), COTMix (Eldele et al., 2023), RAINCOAT (He et al., 2023), SEA (Wang et al., 2023), and SEA++ (Wang et al., 2024b), while the SFDA baselines include SHOT (Liang et al., 2020), neighborhood reciprocity clustering (NRC) (Yang et al., 2021), attracting and dispersing (AaD) (Yang et al., 2022), source-free domain adaptation through the lens of data augmentation (SF(DA)^2) (Hwang et al., 2024), MAPU (Ragab et al., 2023b), TERSE (Gong et al., 2025b), and ACCUP (Gong et al., 2025a). We also report source-only and target-supervised performance as lower- and upper-reference bounds, respectively. Since UDA methods can access source samples during adaptation, their results are retained as an upper-information reference rather than a strictly source-free comparison.

Implementation. All methods use the same MTS backbone and training configuration. The temporal encoder comprises three 1-D convolutional blocks with widths 64, 128, and 128, followed by a 256-dimensional projection. The spectral branch uses an FFT front-end and shares encoder weights; the relation branch applies a one-layer graph convolution over $G(X)$. Training uses Adam with batch size 32 and learning rates 10^{-3} for source training and 3×10^{-4} for target adaptation. Source factorization and anchor coefficients are both 0.1, with anchor margin 0.2. Cosine temperatures, disagreement temperature, class-mass coefficient, spectral ridge, eigenvalue-energy threshold, Sinkhorn regularization, and EMA coefficient are respectively 0.1, 0.5, 0.1, 10^{-4}, 0.9, 0.05, and 0.9. Numerical stabilization uses 10^{-8}. Hyperparameters are selected on the source validation split and fixed across target domains without using target labels for model selection.

4.2 Approach Comparison

We evaluate two representative cross-subject transfer tasks for each benchmark under a unified setting. Table 1 summarizes the results on UCIHAR, WISDM, and SSC, where the "SF" column indicates source-free adaptation and the † symbol denotes source-only and target-supervised settings. Each result is averaged over three random seeds, with the best adaptation performance highlighted in bold. STSD achieves consistently strong performance across different cross-subject tasks and remains competitive with or surpasses existing approaches on all three benchmarks. These results demonstrate that STSD can effectively improve cross-subject adaptation without requiring access to source samples during target adaptation.

Table 1: Macro F1 (%, mean±standard deviation) on six representative source-to-target tasks from UCIHAR, WISDM, and SSC.

Method	SF	UCIHAR		WISDM		SSC	
		$12 \to 16$	$6 \to 23$	$33 \to 12$	$6 \to 19$	$16 \to 1$	$7 \to 18$
Source only	†	76.73±7.36	77.11±11.20	27.43±2.82	54.90±5.96	64.16±1.88	61.18±0.27
Target supervised	†	97.54±0.70	100.00±0.00	85.95±4.47	97.49±0.85	85.37±0.71	76.53±0.79
DANN (Ganin et al., 2016)	✗	84.86±2.18	95.31±2.33	74.67±17.72	56.29±1.64	66.55±3.84	71.46±3.10
CDAN (Long et al., 2018)	✗	88.68±0.89	93.99±0.80	64.44±7.49	33.81±1.63	69.25±2.02	72.53±1.43
CoDATS (Wilson et al., 2020)	✗	86.07±1.57	97.60±2.65	70.17±12.90	59.73±3.20	65.71±0.86	75.01±0.52
COTMix (Eldele et al., 2023)	✗	88.07±1.82	98.76±1.42	76.89±8.16	64.39±4.07	69.78±2.88	72.86±1.81
RAINCOAT (He et al., 2023)	✗	88.52±1.48	96.92±1.11	81.19±7.43	63.74±3.56	68.76±2.49	72.79±1.49
SEA (Wang et al., 2023)	✗	88.96±1.23	98.55±0.87	81.80±6.97	65.20±3.28	68.57±2.21	73.36±1.32
SEA++ (Wang et al., 2024b)	✗	89.20±1.05	98.05±0.66	82.05±6.44	63.79±3.06	70.72±2.01	72.00±1.19
SHOT (Liang et al., 2020)	✓	85.65±0.83	98.47±1.32	67.18±9.32	60.88±0.85	71.87±1.21	66.60±3.08
NRC (Yang et al., 2021)	✓	84.78±4.51	92.87±1.77	56.09±10.54	63.56±8.30	67.86±3.17	70.52±2.55
AaD (Yang et al., 2022)	✓	89.25±1.21	99.03±1.52	65.24±9.28	71.33±7.86	64.96±0.89	66.91±6.93
SF(DA)^2 (Hwang et al., 2024)	✓	88.85±0.00	97.20±0.00	60.40±4.74	69.59±4.37	66.88±1.52	67.47±1.71
MAPU (Ragab et al., 2023b)	✓	86.80±0.96	98.84±0.77	77.36±6.64	68.04±9.88	64.74±3.69	72.87±1.16
TERSE (Gong et al., 2025b)	✓	88.87±0.00	97.96±0.39	89.94±12.53	70.05±10.19	75.11±0.28	75.64±1.35
ACCUP (Gong et al., 2025a)	✓	90.35±0.96	98.43±0.44	85.28±5.61	72.56±6.12	71.61±1.38	73.61±1.01
STSD (ours)	✓	**91.01±1.09**	**99.07±0.22**	**90.16±4.20**	**74.09±3.73**	**75.28±0.71**	**75.87±0.95**

4.3 Ablation Analysis

We compare STSD with six controlled variants examining transferable-space identification, spectral evidence, semantic correspondence, objective formulation, and operation order. Table 2 reports average MF1 on UCIHAR, WISDM, and SSC, alongside the negative-transfer ratio (NTR), the proportion of domain pairs performing worse than source-only inference. Applying optimal transport before spectral identification causes the largest degradation and highest NTR, supporting the importance of identifying transferable semantics before correspondence estimation. Removing source geometry or target persistence also reduces performance, indicating that both contribute to cross-subject adaptation.

Table 2: Mechanism ablation with average MF1 (%) and negative-transfer rate (NTR).

Variant	UCIHAR	WISDM	SSC	NTR ↓
A. Full-space transfer, $P_{\mathrm{tr}} = I$	91.88	76.92	71.63	33.3%
B. Source geometry only	92.74	78.45	72.84	33.3%
C. Target persistence only	92.31	77.96	72.22	33.3%
D. Direct correspondence	93.16	79.58	73.62	16.7%
E. Naive weighted combination	91.42	76.31	71.05	50.0%
F. Optimal transport before spectral identification	90.36	74.88	69.94	66.7%
Full STSD	**95.04**	**82.13**	**75.58**	**0.0%**

4.4 Parameter Sensitivity

Fig. 3 examines the sensitivity of STSD to ξ, ϵ_g, ε, and ρ on SSC, reporting mean±standard deviation MF1 over three random seeds with other settings fixed. As ξ increases from 0.70 to 0.90, MF1 improves and remains stable between 0.90 and 0.95, indicating a broad effective range

for subspace selection. The spectral ridge ϵ_g has the smallest effect, with little variation between 10^{-5} and 10^{-3}. By contrast, ε shows greater, asymmetric sensitivity: performance declines more noticeably at 0.20, suggesting that excessive entropic smoothing weakens semantic correspondence. Introducing class-mass information through ρ improves MF1, which plateaus around 0.10–0.20 without further gains at larger values. These trends indicate that the chosen settings fall within stable operating ranges rather than isolated optima.

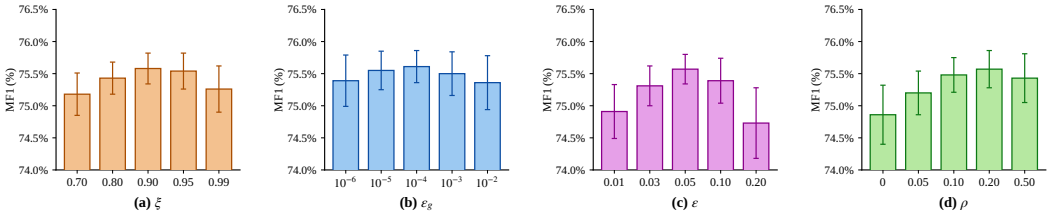

Figure 3: Parameter sensitivity of STSD on SSC.

4.5 Feature Visualization and Transport Analysis

Fig. 4 presents the visualization results on the SSC 7→18 task. As shown in Fig. 4(a), source-only inference produces highly overlapped target features across several sleep stages, indicating that source semantics cannot be directly transferred under cross-subject distribution shifts. Fig. 4(b) reduces this overlap but still shows unclear boundaries among intermediate stages, especially N1, N2, and N3. In contrast, Fig. 4(c) generates more compact class distributions with clearer boundaries while preserving moderate overlap between adjacent stages, demonstrating the effectiveness of transferable-space identification. Fig. 4(d) shows that the learned transport matrix mainly follows diagonal correspondence, with off-diagonal responses concentrated between neighboring stages. These results verify that semantic transport captures meaningful class relationships and improves cross-subject sleep staging.

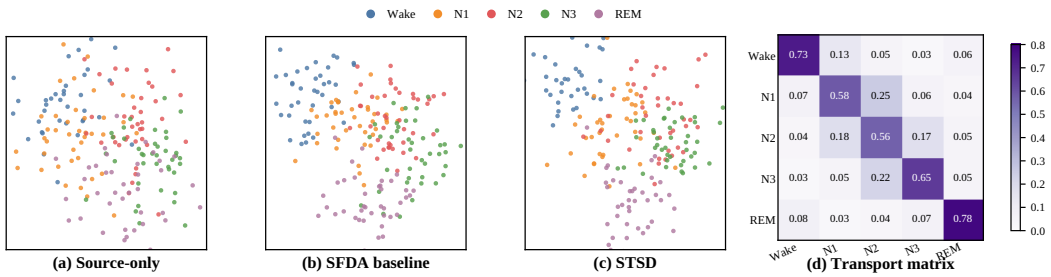

Figure 4: Representative visualization on the SSC 7→18 task. Panels (a)–(c) show t-SNE embeddings of target features obtained by source-only inference, a representative SFDA baseline (Gong et al., 2025b), and STSD, respectively. Panel (d) shows the learned semantic transport matrix across the five sleep stages in the transferable subspace.

5 Conclusion

In this study, we propose STSD for source-free multivariate time series domain adaptation. More specifically, multi-view source training first constructs a shared semantic basis and calibrates classifier directions as persistent class anchors. Second, the source anchor geometry and cross-view disagreement from unlabeled target data are jointly incorporated into a generalized spectral criterion to identify transferable semantics and separate them from target-local variation. Third, confidence-weighted target prototypes are established in the transferable subspace, and entropic semantic transport is employed to resolve class-level correspondence before target adaptation. Based on this, a unified target objective updates the target encoder, while the margin-preservation analysis characterizes the retained class separation after semantic transport. Finally, we evaluate STSD on the UCIHAR, WISDM, and SSC datasets, with the experimental results validating its effectiveness and superiority for source-free cross-subject adaptation.

REFERENCES

Davide Anguita, Alessandro Ghio, Luca Oneto, Xavier Parra, and Jorge L. Reyes-Ortiz. A public domain dataset for human activity recognition using smartphones. In *Proceedings of the European Symposium on Artificial Neural Networks*, 2013.

Abdul Fatir Ansari, Lorenzo Stella, Ali Caner Turkmen, Xiyuan Zhang, Pedro Mercado, Huibin Shen, Oleksandr Shchur, Syama Sundar Rangapuram, Sebastian Pineda Arango, Shubham Kapoor, Jasper Zschiegner, Danielle C. Maddix, Hao Wang, Michael W. Mahoney, Kari Torkkola, Andrew Gordon Wilson, Michael Bohlke-Schneider, and Bernie Wang. Chronos: Learning the language of time series. *Transactions on Machine Learning Research*, 2024.

Mathilde Caron, Ishan Misra, Julien Mairal, Priya Goyal, Piotr Bojanowski, and Armand Joulin. Unsupervised learning of visual features by contrasting cluster assignments. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 33, pp. 9912–9924, 2020.

Jiaxiang Dong, Haixu Wu, Haoran Zhang, Li Zhang, Jianmin Wang, and Mingsheng Long. Simmtm: A simple pre-training framework for masked time-series modeling. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 36, 2023.

Emadeldeen Eldele, Mohamed Ragab, Zhenghua Chen, Min Wu, Chee-Keong Kwoh, and Xiaoli Li. Contrastive domain adaptation for time-series via temporal mixup. *IEEE Transactions on Artificial Intelligence*, 2023.

Yunfeng Fan, Wenchao Xu, Haohao Wang, and Song Guo. Cross-modal representation flattening for multi-modal domain generalization. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 37, 2024.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016.

Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23):e215–e220, 2000.

Peiliang Gong, Mohamed Ragab, Min Wu, Zhenghua Chen, Yongyi Su, Xiaoli Li, and Daoqiang Zhang. Augmented contrastive clustering with uncertainty-aware prototyping for time series test time adaptation. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 390–401, 2025a.

Peiliang Gong, Yucheng Wang, Min Wu, Zhenghua Chen, Xiaoli Li, and Daoqiang Zhang. Temporal restoration and spatial rewiring for source-free multivariate time series domain adaptation. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 685–696, 2025b.

Mononito Goswami, Konrad Szafer, Arjun Choudhry, Yifu Cai, Shuo Li, and Artur Dubrawski. Moment: A family of open time-series foundation models. In *Proceedings of the International Conference on Machine Learning*, pp. 16115–16152, 2024.

Huan He, Owen Queen, Teddy Koker, Consuelo Cuevas, Theodoros Tsiligkaridis, and Marinka Zitnik. Domain adaptation for time series under feature and label shifts. In *Proceedings of the International Conference on Machine Learning*, pp. 12746–12774, 2023.

Uiwon Hwang, Jonghyun Lee, Juhyeon Shin, and Sungroh Yoon. SF(DA)^2: Source-free domain adaptation through the lens of data augmentation. In *Proceedings of the International Conference on Learning Representations*, 2024.

Ming Jin, Shiyu Wang, Lintao Ma, Zhixuan Chu, James Y. Zhang, Xiaoming Shi, Pin-Yu Chen, Yuxuan Liang, Yuan-Fang Li, Shirui Pan, and Qingsong Wen. Time-llm: Time series forecasting by reprogramming large language models. In *Proceedings of the International Conference on Learning Representations*, 2024.

Prannay Khosla, Piotr Teterwak, Chen Wang, Aaron Sarna, Yonglong Tian, Phillip Isola, Aaron Maschinot, Ce Liu, and Dilip Krishnan. Supervised contrastive learning. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 33, pp. 18661–18673, 2020.

Jennifer R. Kwapisz, Gary M. Weiss, and Samuel A. Moore. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12(2):74–82, 2011.

Jingjing Li, Zhiqi Yu, Zhekai Du, Lei Zhu, and Heng Tao Shen. A comprehensive survey on source-free domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

Jian Liang, Dapeng Hu, and Jiashi Feng. Do we really need to access the source data? source hypothesis transfer for unsupervised domain adaptation. In *Proceedings of the International Conference on Machine Learning*, pp. 6028–6039, 2020.

Jian Liang, Dapeng Hu, Yunbo Wang, Ran He, and Jiashi Feng. Source data-absent unsupervised domain adaptation through hypothesis transfer and labeling transfer. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(11):8602–8617, 2022.

Paul Pu Liang, Zihao Deng, Martin Q. Ma, James Y. Zou, Louis-Philippe Morency, and Ruslan Salakhutdinov. Factorized contrastive learning: Going beyond multi-view redundancy. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 36, 2023.

Yong Liu, Haoran Zhang, Chenyu Li, Xiangdong Huang, Jianmin Wang, and Mingsheng Long. Timer: Generative pre-trained transformers are large time series models. In *Proceedings of the International Conference on Machine Learning*, pp. 32369–32399, 2024.

Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I. Jordan. Conditional adversarial domain adaptation. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 31, 2018.

Mohamed Ragab, Emadeldeen Eldele, Wee Ling Tan, Chuan-Sheng Foo, Zhenghua Chen, Min Wu, Chee-Keong Kwoh, and Xiaoli Li. Adatime: A benchmarking suite for domain adaptation on time series data. *ACM Transactions on Knowledge Discovery from Data*, 17(8):1–18, 2023a.

Mohamed Ragab, Emadeldeen Eldele, Min Wu, Chuan-Sheng Foo, Xiaoli Li, and Zhenghua Chen. Source-free domain adaptation with temporal imputation for time series data. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1989–1998, 2023b.

Mohamed Ragab, Emadeldeen Eldele, Zhenghua Chen, Min Wu, Chee-Keong Kwoh, and Xiaoli Li. Self-supervised autoregressive domain adaptation for time series data. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):1341–1351, 2024.

Xiaoming Shi, Shiyu Wang, Yuqi Nie, Dianqi Li, Zhou Ye, Qingsong Wen, and Ming Jin. Time-moe: Billion-scale time series foundation models with mixture of experts. In *Proceedings of the International Conference on Learning Representations*, 2025.

Sana Tonekaboni, Danny Eytan, and Anna Goldenberg. Unsupervised representation learning for time series with temporal neighborhood coding. In *Proceedings of the International Conference on Learning Representations*, 2021.

Junxiang Wang, Guangji Bai, Wei Cheng, Zhengzhang Chen, Liang Zhao, and Haifeng Chen. Pond: Multi-source time series domain adaptation with information-aware prompt tuning. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3140–3151, 2024a.

Shiyu Wang, Jiawei Li, Xiaoming Shi, Zhou Ye, Baichuan Mo, Wenze Lin, Shengtong Ju, Zhixuan Chu, and Ming Jin. Timemixer++: A general time series pattern machine for universal predictive analysis. In *Proceedings of the International Conference on Learning Representations*, 2025.

Yucheng Wang, Yuecong Xu, Jianfei Yang, Zhenghua Chen, Min Wu, Xiaoli Li, and Lihua Xie. Sensor alignment for multivariate time-series unsupervised domain adaptation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pp. 10253–10261, 2023.

Yucheng Wang, Yuecong Xu, Jianfei Yang, Min Wu, Xiaoli Li, Lihua Xie, and Zhenghua Chen. Sea++: Multi-graph-based higher-order sensor alignment for multivariate time-series unsupervised domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(12): 10781–10796, 2024b.

Garrett Wilson, Janardhan Rao Doppa, and Diane J. Cook. Multi-source deep domain adaptation with weak supervision for time-series sensor data. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 1768–1778, 2020.

Garrett Wilson, Janardhan Rao Doppa, and Diane J. Cook. Calda: Improving multi-source time series domain adaptation with contrastive adversarial learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(12):14208–14221, 2023.

Gerald Woo, Chenghao Liu, Doyen Sahoo, Akshat Kumar, and Steven Hoi. Cost: Contrastive learning of disentangled seasonal-trend representations for time series forecasting. In *Proceedings of the International Conference on Learning Representations*, 2022.

Shiqi Yang, Joost van de Weijer, Luis Herranz, Shangling Jui, et al. Exploiting the intrinsic neighborhood structure for source-free domain adaptation. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 34, pp. 29393–29405, 2021.

Shiqi Yang, Yaxing Wang, Kai Wang, Shangling Jui, and Joost van de Weijer. Attracting and dispersing: A simple approach for source-free domain adaptation. In *Proceedings of the Conference on Neural Information Processing Systems*, volume 35, 2022.

Xiangyu Yue, Zangwei Zheng, Shanghang Zhang, Yang Gao, Trevor Darrell, Kurt Keutzer, and Alberto Sangiovanni-Vincentelli. Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13834–13844, 2021.