Adaptive Incremental Prognostics for Remaining Useful Life Under Changing Operating Conditions

Aryan Morteza, *Member, IEEE*, Mohammad Hossein Anisi, *Senior Member, IEEE*, Morteza Varasteh, Faiyaz Doctor, Mark Hadaway, and Andy Dowell

***Abstract*—Remaining Useful Life (RUL) prediction is central to reliability assurance and predictive maintenance. Deep learning models capture nonlinear degradation behaviour but often require substantial computation and retraining, while existing Broad Learning System (BLS) approaches have limited adaptability to evolving degradation patterns and operating conditions. This paper proposes the Adaptive Transfer Broad Learning System (ATBLS), a lightweight framework for incremental RUL prediction under changing operating conditions. ATBLS combines structured enhancement fusion, sequential projection, PCA-guided contribution weighting, and analytical ridge regression within a single adaptation pathway. Enhancement fusion reorganizes nonlinear responses before fixed sequential projections generate complementary degradation representations for prediction and cross-dataset comparison. Projection aggregates form normalized-life descriptors in a jointly recomputed shared PCA space to estimate participant relevance and source–incremental domain contributions. These contributions weight the participating transition matrices before one closed-form ridge solution is obtained, while the feature, enhancement, fusion, and projection mappings remain fixed. Thus, newly available labeled data are incorporated without end-to-end retraining. ATBLS is evaluated on three run-to-failure benchmarks covering rolling bearings and turbofan engines, including within-domain incremental learning and explicit cross-domain operating-condition transitions. Results demonstrate accurate RUL prediction, low analytical adaptation cost, and substantial improvements in several cross-domain transitions, although the benefit of incremental target-domain information remains condition dependent.**

***Index Terms*—Remaining Useful Life (RUL), Incremental Learning, Broad Learning System (BLS), Projection Layer, PCA-guided Contribution Weighting, Reliability Prognostics**

I. Introduction

ACCURATE prediction of Remaining Useful Life (RUL) is central to Prognostics and Health Management (PHM), supporting reliability assurance, condition-based maintenance, and operational planning [1]–[3]. Data-driven prognostic methods use industrial sensing and operational data to estimate degradation trends without requiring complete physical degradation models [2], [3]. Reliable RUL estimation can improve equipment availability, reduce downtime, and support maintenance decisions [1], [4].

Classical data-driven approaches, including Support Vector Machines (SVM), Relevance Vector Machines (RVM), autoencoders, and Restricted Boltzmann Machines (RBM), have been widely applied to degradation modelling and RUL estimation [5]–[8]. Their effectiveness can depend on feature representation, training-data coverage, and operating conditions. Changes in load, speed, sensing characteristics, or degradation trajectories may alter feature distributions and weaken generalization across operating conditions.

Deep learning architectures, including Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM) networks, attention-based models, and Transformer variants, can capture nonlinear degradation behaviour and temporal dependencies [9]–[13]. Recent research has further investigated domain adaptation, transfer learning, self-supervised time-series representation, and robustness to changing sensing conditions [12]–[15]. Despite their predictive capabilities, these approaches can involve iterative optimization, substantial computational requirements, and parameter tuning. Incorporating degradation data from different operating conditions may also require additional optimization or fine-tuning.

Broad Learning Systems (BLS) offer an alternative through flat feature and enhancement mappings with analytically estimated output weights [16]–[18]. Their architecture supports efficient model construction and incremental learning without optimizing a deep network. Existing BLS-based prognostic approaches have explored frequency-domain representations, recurrent components, temporal-cascade structures, and multiscale feature learning to improve RUL prediction and incorporate additional degradation data [4], [9], [19], [20]. Related flat-learning approaches further demonstrate the potential of computationally efficient non-deep structures [21].

However, incremental RUL learning involves more than incorporating additional training samples. Newly available degradation trajectories may originate from different loads, speeds, or fault conditions, producing representations that differ from the original source data. Consequently, participating datasets may not contribute equally to an updated predictor. This distinction motivates evaluating dataset relevance and representation differences rather than relying solely on the availability of additional training samples. Although existing BLS-based RUL methods support efficient prediction and incremental learning, the integration of complementary

Aryan Morteza (Corresponding Author) is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: aryan.morteza@essex.ac.uk).

Mohammad Hossein Anisi is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: m.anisi@essex.ac.uk).

Morteza Varasteh is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: m.varasteh@essex.ac.uk).

Faiyaz Doctor is with the School of Computer Science and Electronic Engineering, University of Essex, Colchester CO4 3SQ, U.K. (e-mail: fdocto@essex.ac.uk).

Mark Hadaway and Andy Dowell are with Lubron Water Technologies Ltd., Lubron House, 8 Challenge Way, Hythe Hill, Colchester, Essex CO1 2LY, U.K. (e-mail: mark.hadaway@lubron.co.uk; andy.dowell@lubron.co.uk).

degradation representations, dataset relevance assessment, and representation-shift-dependent contribution weighting within a single analytical adaptation pathway remains less explored. This motivates a framework that adjusts source–incremental contributions while retaining computationally efficient predictor updating.

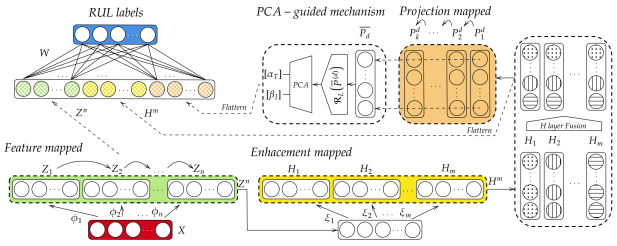


Fig. 1. Architecture of the Adaptive Transfer Broad Learning System (ATBLS).

To address these limitations, this paper proposes the Adaptive Transfer Broad Learning System (ATBLS), illustrated in Fig. 1. ATBLS extends the flat analytical structure of BLS by combining structured enhancement fusion, lightweight sequential projection, PCA-guided contribution weighting, and weighted closed-form ridge regression within a single adaptation pathway.

The feature, enhancement, and projection mappings are initialized once and remain fixed, while enhancement fusion is deterministic and parameter-free. The fused enhancement responses are processed through sequential scalar projection mappings, whose aggregated outputs provide complementary degradation representations for RUL prediction. The native projection aggregates are retained in the predictive transition matrices, while separate fixed-row copies support cross-dataset comparison through shared PCA analysis. Applying common fixed mappings across participating datasets enables consistent representation construction without re-estimating the network.

When labeled incremental data become available, source and incremental projection descriptors are jointly centered to recompute a shared PCA basis. The resulting representation estimates participant relevance and source–incremental representation shift, which determine dataset and domain contribution weights. These coefficients weight the participating transition matrices before one updated closed-form ridge solution is obtained. PCA therefore guides contribution estimation without replacing or reducing the predictive representations. ATBLS consequently adjusts dataset influence under observed representation shifts without relying on a permanently fixed source-only PCA basis or requiring end-to-end network retraining.

The framework is evaluated across three run-to-failure benchmark case studies involving rolling bearings and turbofan engines. The experiments include within-domain incremental learning and matched cross-domain transitions, where source and target-test partitions remain fixed while newly available labeled target-domain data are incorporated. These experiments evaluate predictive accuracy, computational cost, and the condition-dependent effects of incremental adaptation under changing operating conditions.

The main contributions of this work are summarized as follows:

- A structured enhancement-fusion and sequential-projection mechanism constructs complementary degradation representations within the flat BLS architecture without deep or recurrent layers.
- A representation-shift-aware PCA-guided contribution mechanism recomputes a shared PCA basis to estimate participant relevance and adaptively balance dataset and domain contributions under changing operating conditions.
- A weighted closed-form ridge formulation incorporates newly available labeled data while retaining fixed feature, enhancement, fusion, and projection mappings, avoiding end-to-end retraining.

The remainder of this paper is organized as follows: Section II reviews related BLS-based and adaptive prognostic methods; Section III presents the ATBLS architecture and adaptation formulation; Section IV evaluates the proposed framework across three benchmark case studies; and Section V summarizes the main findings and future directions.

II. Related Works

This section reviews RUL prediction methods relevant to the proposed ATBLS framework, with emphasis on Broad Learning System (BLS)-based approaches. As summarized in Table I, existing methods are considered from three perspectives: degradation representation, adaptation to newly available data or changing operating conditions, and the computational mechanism used for model updating. Deep and transfer-oriented prognostic approaches are also briefly discussed to position ATBLS relative to broader developments in data-driven PHM.

Deep architectures, including CNN–LSTM hybrids, multiscale CNNs, recurrent models, attention mechanisms, and Transformer-based predictors, have been widely investigated for RUL estimation [22]–[25]. Their hierarchical structures can capture nonlinear degradation patterns, temporal dependencies, and cross-channel relationships, while recent domain-adaptation and transfer-learning methods further aim to improve generalization across operating conditions. These capabilities are attractive for complex prognostic problems, but typically rely on iterative gradient-based optimization and may require additional training or fine-tuning when the operating domain changes. This contrasts with the flat analytical learning structure considered in the present work.

Recent PHM research has also explored cross-domain representation learning and large pre-trained models. Graph-based domain adaptation has been investigated for cross-condition RUL prediction [26], while source-free and domain-generalization approaches seek to reduce dependence on source-domain data during adaptation [27]. Industrial foundation models and large-model-based prognostic frameworks have likewise been studied for cross-task and cross-domain representation learning [28]–[30]. These approaches demonstrate the growing importance of transferable degradation

TABLE I
COMPARISON OF RECENT BLS-BASED METHODS FOR RUL PREDICTION.

Method	Temporal Modeling	Case Studies	Manual Tuning	Network Expansion	Feature Representation	Incremental Learning
Fourier-BLS [4]	No	1	Yes	No	Fourier-based	No
LSTM–BLS [9]	Partial	1	Yes	No	LSTM-based	No
TCBLS [19]	Partial	4	Yes	Yes	Linear and nonlinear	Yes
Multiscale BLS [20]	Partial	2	Yes	No	Multiresolution	Yes
Proposed ATBLS	**Structured**	**3**	**Limited**	**No**	**Fusion + projection**	**Yes**

representations, although their training and deployment mechanisms differ substantially from the lightweight closed-form adaptation targeted by ATBLS.

BLS provides an alternative flat-learning paradigm derived from random-vector functional-link principles [18], using feature and enhancement nodes together with analytically estimated output weights [16], [17]. Its shallow structure enables rapid model construction and makes BLS attractive for prognostic applications with limited computational resources. Nevertheless, conventional BLS does not explicitly model degradation evolution or operating-condition changes, motivating several RUL-oriented extensions.

Frequency-domain BLS approaches incorporate spectral degradation information into the flat network structure [4], while LSTM–BLS architectures introduce recurrent representations to improve temporal sensitivity [9]. Temporal Cascade BLS (TCBLS) further extends BLS toward incremental prognostics by incorporating newly available degradation information [19]. Multiscale BLS (MSBLS) constructs multiscale representations to capture degradation behaviour at different resolutions [20]. These developments demonstrate that BLS can be adapted for RUL prediction while retaining much of its lightweight structure.

However, these extensions mainly address individual aspects of the prognostic problem. Frequency-oriented approaches emphasize input representation, recurrent hybrids improve temporal modelling at the cost of introducing an additional trainable structure, and incremental BLS variants focus primarily on efficient model updating. A unified BLS-based mechanism that simultaneously enriches the degradation representation, estimates the relative relevance of source and newly available datasets under operating-condition change, and performs adaptation through a single analytical learning stage remains less explored.

ATBLS addresses this gap through three coupled mechanisms. First, structured enhancement fusion and lightweight sequential projection construct complementary degradation representations within the flat BLS topology. Second, source and incremental projection descriptors are embedded in a shared PCA space to estimate participant relevance and source–incremental representation shift, which determine their contribution to adaptation. Third, these contribution coefficients are incorporated into one weighted closed-form ridge formulation while the feature, enhancement, fusion, and projection mappings remain fixed. This design enables incremental adaptation under changing operating conditions without requiring end-to-end network retraining.

III. METHODOLOGY

This section presents the proposed ATBLS framework for incremental RUL prediction under changing operating conditions. We first define the source, incremental, and test-data setting, then describe the ATBLS architecture, including feature mapping, enhancement-level fusion, sequential projection, PCA-guided contribution weighting, and ridge regression. Finally, the incremental adaptation and end-to-end RUL prediction procedures are summarized.

A. Problem Definition

Each sample is represented by an input-label pair $(X, Y) \in \mathbb{R}^{d_{\text{in}}} \times \mathbb{R}$, where X is the input vector and Y the corresponding RUL label. Three data roles are considered: source, incremental, and unseen test data. ATBLS maps X through feature, enhancement-fusion, and projection stages to construct the transition matrix A. When labeled incremental data become available, their contribution relative to the source data is estimated through the PCA-guided weighting mechanism, followed by one closed-form ridge solution as summarized in Algorithm 1.

B. ATBLS Framework Overview

ATBLS extends BLS through sequential feature mappings, nonlinear enhancement nodes, enhancement-level fusion, and lightweight sequential projection. The resulting representations are combined as $A^{(d)} = [Z^{n,(d)} \mid \widetilde{H}^{m,(d)} \mid \bar{P}^{(d)}]$ for each participating dataset d. The same fixed feature, enhancement, fusion, and projection mappings are used for source, incremental, and unseen test data.

When labeled incremental data become available, source and incremental projection representations are compared in a shared PCA space to determine participant relevance and relative domain contribution. PCA is used only for contribution weighting and does not replace or reduce the predictive representations Z, $\widetilde{H}$, or $\bar{P}$. The weighted participating data are then combined through one closed-form ridge solution, allowing ATBLS to adapt without end-to-end retraining.

1) Design and Structure: For a source dataset $(X^E, Y^E) \in \mathbb{R}^{N \times d_{\text{in}}} \times \mathbb{R}^{N \times q}$, where N is the number of samples, d_{in} the input dimension, and q the output dimension, ATBLS employs n sequential feature-mapping groups, m enhancement groups, and k projection mappings. The first feature representation is

$$Z_1 = \phi_1\left(X^E W_{e1} + \beta_{e1}\right), \tag{1}$$

where $W_{e1} \in \mathbb{R}^{d_{\text{in}} \times d_z}$, $\beta_{e1} \in \mathbb{R}^{1 \times d_z}$, and d_z is the width of each feature group. The remaining mappings are applied sequentially:

$$Z_k = \phi_k \left(Z_{k-1} W_{ek} + \beta_{ek} \right), \qquad k = 2, \ldots, n, \tag{2}$$

with $W_{ek} \in \mathbb{R}^{d_z \times d_z}$. Their outputs are concatenated as

$$Z^n = [Z_1, Z_2, \ldots, Z_n] \in \mathbb{R}^{N \times n d_z}. \tag{3}$$

The resulting representation is mapped to m nonlinear enhancement groups using ReLU:

$$H_j = \xi \left(Z^n W_{hj} + \beta_{hj} \right), \qquad j = 1, \ldots, m, \tag{4}$$

where $W_{hj} \in \mathbb{R}^{n d_z \times d_h}$ and $H_j \in \mathbb{R}^{N \times d_h}$. The enhancement outputs are concatenated as

$$H^m = [H_1, H_2, \ldots, H_m] \in \mathbb{R}^{N \times m d_h}. \tag{5}$$

Enhancement-level fusion then transforms H^m into $\widetilde{H}^m$ using the parameter-free mechanism described in Section III-C1. The fused representation is subsequently processed by k lightweight sequential projection mappings. For participant d,

$$\mathbf{p}_0^{(d)} = \text{vec} \left(\widetilde{H}^{m,(d)} \right), \tag{6}$$

and the projection sequence is

$$\mathbf{p}_{o+1}^{(d)} = w_{p(o+1)} \mathbf{p}_o^{(d)} + \beta_{p(o+1)} \mathbf{1}, \qquad o = 0, \ldots, k-1, \tag{7}$$

where $w_{p(o+1)}$ and $\beta_{p(o+1)}$ are scalar parameters. No nonlinear activation is introduced in this stage. Each projected vector is reshaped to

$$P_o^{(d)} \in \mathbb{R}^{N_d \times m d_h},$$

and the projection states are aggregated as

$$\bar{P}^{(d)} = \frac{1}{k} \sum_{o=1}^{k} P_o^{(d)}. \tag{8}$$

The participant transition matrix is therefore

$$A^{(d)} = \left[Z^{n,(d)} \mid \widetilde{H}^{m,(d)} \mid \bar{P}^{(d)} \right] \in \mathbb{R}^{N_d \times (n d_z + 2 m d_h)}. \tag{9}$$

A normalized-life copy of $\bar{P}^{(d)}$ is used separately for PCA-guided contribution weighting, while the native $\bar{P}^{(d)}$ remains unchanged in $A^{(d)}$. Unseen test data are excluded from PCA and contribution estimation.

All feature, enhancement, and projection parameters are initialized once from $\mathcal{N}(0, 1)$ and then kept fixed for source, incremental, and unseen test data. Adaptation therefore changes only the data-derived contribution weights and the final ridge output weights, rather than retraining the representation mappings.

2) Weight Update Formulation: For the ATBLS transition matrix $A \in \mathbb{R}^{N \times d_A}$, where $d_A = n d_z + 2 m d_h$, the output weights are obtained by ridge regression:

$$W^\star = \arg\min_W \left(\|Y - AW\|_F^2 + \lambda \|W\|_F^2 \right), \tag{10}$$

with closed-form solution

$$W = (A^\top A + \lambda I)^{-1} A^\top Y, \tag{11}$$

where λ is the fixed ridge parameter and $\widehat{Y} = AW$. During incremental adaptation, the same analytical solution is applied to the contribution-weighted design defined in Section III-C, while the representation mappings remain fixed.

C. ATBLS Incremental Adaptation

ATBLS uses a single contribution-weighted adaptation pathway. Source and available incremental datasets are passed through the same fixed feature, enhancement-fusion, and projection mappings to construct their transition matrices. Their projection representations are then compared in a shared PCA space to derive participant relevance and source–incremental domain weights. The weighted data are finally combined through one closed-form ridge solution. As new labeled incremental data arrive, only the contribution coefficients and final ridge weights are recomputed; the representation mappings remain fixed.

Algorithm 1 ATBLS contribution-weighted incremental adaptation

Input: Source datasets $\{(X_d^E, Y_d^E)\}_{d \in \mathcal{T}}$; available incremental datasets $\{(X_j^I, Y_j^I)\}_{j \in \mathcal{I}}$; unseen test data X^T.
Hyperparameters: ridge parameter λ, PCA energy threshold η, and descriptor length L.
Initialize feature, enhancement, and projection mappings once and keep them fixed.
for each participant $d \in \mathcal{T} \cup \mathcal{I}$ **do**
    Compute Z_d, H_d, and $\widetilde{H}_d$.
    Compute $P_1^{(d)}, \ldots, P_k^{(d)}$, form $\bar{P}^{(d)}$, and construct $A_d = [Z_d \mid \widetilde{H}_d \mid \bar{P}^{(d)}]$.
end for
if $\mathcal{I} = \varnothing$ **then**
    Assign equal source coefficients and set $\rho_{\text{Tr}} = 1$, $\rho_{\text{Inc}} = 0$.
else
    Form normalized-life descriptors S_d, jointly recompute the shared PCA basis, and obtain participant score clouds.
    Compute participant relevance $\boldsymbol{\alpha}^{\text{Tr}}$, $\boldsymbol{\beta}^{\text{Inc}}$, representation-shift index q, and domain weights $\rho_{\text{Tr}}, \rho_{\text{Inc}}$.
end if
Compute all participant coefficients ω_d and form one weighted stacked design $\widetilde{A}$ and targets $\widetilde{Y}$.
Solve one closed-form ridge system using $(\widetilde{A}, \widetilde{Y})$ to obtain W.
Pass unseen X^T through the same fixed mappings to form $A_{\text{test}} = [Z^T \mid \widetilde{H}^T \mid \bar{P}^T]$.
Predict $\widehat{Y}^T = A_{\text{test}} W$.
Output: $\widehat{Y}^T$

A changing operating condition is represented by the arrival of labeled incremental data from a new regime rather than by an explicit condition identifier. ATBLS therefore adapts to the resulting representation shift rather than to a predefined physical-state label. Changes in the projection distribution modify participant relevance, the shift index q, and the

source–incremental contribution weights. Accordingly, within-condition increments evaluate incremental learning, whereas cross-condition and cross-domain transitions evaluate adaptation to changing operating conditions.

1) Enhancement-Level Fusion: As illustrated in Fig. 2, enhancement fusion reorganizes nonlinear enhancement responses before sequential projection without introducing additional trainable parameters. Let $H^m \in \mathbb{R}^{N\times M}$, where $M = md_h$ is the total enhancement width. Using row-wise vectorization, denoted by $\mathrm{vec}_r(\cdot)$, ATBLS applies the deterministic one-block reordering

$$\widetilde{H}^m = \mathcal{F}(H^m) = \left[\mathrm{reshape}_{M\times N}\left(\mathrm{vec}_r(H^m)\right)\right]^\top \in \mathbb{R}^{N\times M}, \tag{12}$$

where the reshape operation fills the intermediate matrix row-wise. Let $v = \mathrm{vec}_r(H^m) \in \mathbb{R}^{NM}$. The fused element at row i and column j is therefore

$$\widetilde{H}^m_{ij} = v_{(j-1)N+i}, \qquad i = 1,\dots,N, \quad j = 1,\dots,M. \tag{13}$$

Since $(j-1)N+i$ spans every index from 1 to NM exactly once, the transformation is bijective and every enhancement activation is retained exactly once. Equivalently,

$$\mathrm{vec}_r(\widetilde{H}^m) = Q\,\mathrm{vec}_r(H^m), \qquad Q^\top Q = I, \tag{14}$$

where Q is the corresponding permutation matrix. Consequently,

$$\|\widetilde{H}^m\|_F^2 = \sum_{i=1}^{N}\sum_{j=1}^{M} v^2_{(j-1)N+i} = \sum_{k=1}^{NM} v_k^2 = \|H^m\|_F^2. \tag{15}$$

Fusion therefore preserves the complete set of scalar enhancement activations, dimensionality, and total representation energy. It does not, however, preserve their original row-wise grouping. For two rows i and ℓ, the corresponding inner products before and after fusion are

$$\begin{aligned}\langle h_i, h_\ell\rangle &= \sum_{j=1}^{M} v_{(i-1)M+j} v_{(\ell-1)M+j},\\ \langle \widetilde{h}_i, \widetilde{h}_\ell\rangle &= \sum_{j=1}^{M} v_{(j-1)N+i} v_{(j-1)N+\ell}.\end{aligned} \tag{16}$$

Because these expressions generally pair different activation coordinates, the resulting row-wise geometry is, except for special data configurations, different:

$$\widetilde{H}^m(\widetilde{H}^m)^\top \neq H^m(H^m)^\top. \tag{17}$$

Thus, fusion does not create additional activations; rather, it presents the same scalar nonlinear responses to the downstream stages through a different row-wise dependence structure. This distinction is important in ATBLS because $\widetilde{H}^m$ is retained directly in the predictive transition matrix and also supplies the sequential projection mappings. The resulting projection aggregate consequently reflects the reorganized representation and is used both in the transition matrix and, through a separate normalized-life copy, for PCA-guided contribution estimation. Hence, fusion can modify the representation seen by both prediction and contribution weighting without increasing dimensionality, introducing trainable fusion parameters, or requiring iterative optimization.

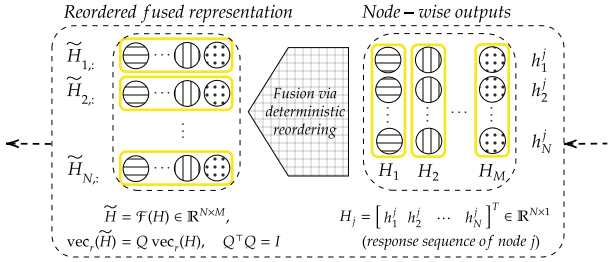


Fig. 2. Enhancement-level fusion in ATBLS. The enhancement responses in H^m are deterministically redistributed to form $\widetilde{H}^m$ without averaging, discarding, or generating activation values.

The fused representation is therefore an internal representation mechanism rather than a separate learning or updating strategy. The same deterministic operation is applied to source, incremental, and unseen test data. Its practical influence is evaluated through the controlled component ablations in Section IV-A5a rather than assumed from the transformation alone.

2) PCA-Guided Contribution Weighting: As illustrated in Fig. 3, the shared PCA mechanism estimates participant relevance and source–incremental domain contributions without reducing the predictive representation. It does not replace, reduce, or modify the predictive transition matrix $A_d = [Z^{n,(d)} \mid \widetilde{H}^{m,(d)} \mid \bar{P}^{(d)}]$. For each participant d, a normalized-life descriptor is formed from a copy of the native projection aggregate:

$$\mathcal{S}_d = \mathcal{R}_L\left(\bar{P}^{(d)}\right) \in \mathbb{R}^{L\times md_h}, \tag{18}$$

where $\mathcal{R}_L$ interpolates the row coordinate to L normalized-life positions. Each row of $\mathcal{S}_d$ is retained as one PCA observation. The native $\bar{P}^{(d)}$, RUL labels, and unseen test data are not resampled for prediction.

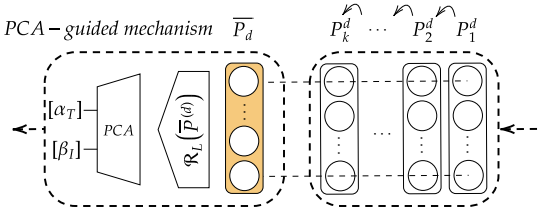


Fig. 3. PCA-guided contribution weighting. Fixed-row copies of the projection aggregates form a shared PCA space for participant relevance and source–incremental domain weighting, while the native projection representations remain in the predictive transition matrices.

All available source and incremental descriptor rows are stacked and centered:

$$S = \mathrm{vstack}_{d\in\mathcal{T}\cup\mathcal{I}}(\mathcal{S}_d), \qquad S_c = S - \mathbf{1}\boldsymbol{\mu}_s^\top. \tag{19}$$

Using $S_c = U\Sigma V^\top$, the smallest rank r satisfying the retained-energy threshold η is selected. The retained scores are then separated into participant-specific clouds $G_d \in \mathbb{R}^{L\times r}$.

For each participant,

$$\mathbf{c}_d = \frac{1}{L}\sum_{i=1}^{L}(G_d)_{i,:}, \tag{20}$$

and $\mathbf{c}_{\mathrm{Tr}}$ and $\mathbf{c}_{\mathrm{Inc}}$ denote the source- and incremental-domain score centres. Participant relevance is measured relative to the incremental-domain centre:

$$\begin{aligned} a_d^{\mathrm{Tr}} &= \|\mathbf{c}_d - \mathbf{c}_{\mathrm{Inc}}\|_2, \quad d \in \mathcal{T}, \\ a_d^{\mathrm{Inc}} &= \|\mathbf{c}_d - \mathbf{c}_{\mathrm{Inc}}\|_2, \quad d \in \mathcal{I}. \end{aligned} \tag{21}$$

The distances are converted to normalized relevance coefficients:

$$\begin{aligned} \alpha_d^{\mathrm{Tr}} &= \frac{\exp(-a_d^{\mathrm{Tr}}/\tau_{\mathrm{Tr}})}{\sum_{j\in\mathcal{T}}\exp(-a_j^{\mathrm{Tr}}/\tau_{\mathrm{Tr}})}, \\ \beta_d^{\mathrm{Inc}} &= \frac{\exp(-a_d^{\mathrm{Inc}}/\tau_{\mathrm{Inc}})}{\sum_{j\in\mathcal{I}}\exp(-a_j^{\mathrm{Inc}}/\tau_{\mathrm{Inc}})}, \end{aligned} \tag{22}$$

where each τ is the median positive distance within its domain. Thus, a larger a_d indicates lower participant relevance and produces an exponentially smaller normalized contribution coefficient. Singleton or tied participants receive equal normalized weights, avoiding the need for an additional manually selected scaling parameter.

The same shared PCA space is used to quantify the separation between the source and incremental domains:

$$\begin{aligned} D &= \|\mathbf{c}_{\mathrm{Tr}} - \mathbf{c}_{\mathrm{Inc}}\|_2, \\ S_{\mathrm{Tr}} &= \frac{1}{N_{\mathrm{Tr}}^P}\sum_{g\in G_{\mathrm{Tr}}}\|g - \mathbf{c}_{\mathrm{Tr}}\|_2, \\ S_{\mathrm{Inc}} &= \frac{1}{N_{\mathrm{Inc}}^P}\sum_{g\in G_{\mathrm{Inc}}}\|g - \mathbf{c}_{\mathrm{Inc}}\|_2. \end{aligned} \tag{23}$$

The normalized representation-shift index is

$$q = \frac{D}{D + S_{\mathrm{Tr}} + S_{\mathrm{Inc}} + \epsilon}, \qquad 0 \le q \le 1, \tag{24}$$

from which the domain contributions are obtained as

$$\rho_{\mathrm{Tr}} = \frac{1}{2+q}, \qquad \rho_{\mathrm{Inc}} = \frac{1+q}{2+q}, \qquad \rho_{\mathrm{Tr}} + \rho_{\mathrm{Inc}} = 1. \tag{25}$$

Thus, small representation shifts give approximately equal source and incremental influence, while larger shifts progressively increase the contribution of the incremental domain without discarding the source information.

The final coefficient assigned to participant d is

$$\omega_d = \begin{cases} \rho_{\mathrm{Tr}}\alpha_d^{\mathrm{Tr}}, & d \in \mathcal{T}, \\ \rho_{\mathrm{Inc}}\beta_d^{\mathrm{Inc}}, & d \in \mathcal{I} \end{cases}. \tag{26}$$

When no incremental data are available, PCA-based relevance is skipped, the source participants receive equal normalized coefficients, and $\rho_{\mathrm{Tr}} = 1$, $\rho_{\mathrm{Inc}} = 0$.

The participant transition matrices are not averaged. Instead, their relative influence is introduced through weighted stacking:

$$\widetilde{A} = \mathrm{vstack}_{d\in\mathcal{T}\cup\mathcal{I}}\left(\sqrt{\omega_d}A_d\right), \qquad \widetilde{Y} = \mathrm{vstack}_{d\in\mathcal{T}\cup\mathcal{I}}\left(\sqrt{\omega_d}Y_d\right). \tag{27}$$

The current adaptation stage then uses one closed-form ridge solution:

$$W = \left(\widetilde{A}^\top\widetilde{A} + \lambda I\right)^{-1}\widetilde{A}^\top\widetilde{Y}. \tag{28}$$

Hence, PCA changes only the relative contribution of the participating datasets; it does not filter or dimensionally reduce their predictive transition matrices.

Importantly, ATBLS does not project incremental data onto a PCA basis fitted once on the source domain. When incremental data become available, source and incremental descriptors are jointly centered and the shared PCA basis is recomputed. Consequently, a rotation of the principal subspace can alter the PCA coordinates, participant relevance, and shift index q without removing new directions from Z, $\widetilde{H}$, or $\bar{P}$.

The retained PCA rank is the smallest r satisfying

$$\frac{\sum_{i=1}^{r}\sigma_i^2}{\sum_i\sigma_i^2} \ge \eta, \tag{29}$$

where σ_i are the singular values of the jointly centered descriptor matrix. In all experiments, $\eta = 0.95$ is fixed globally. Descriptor length L, ridge parameter λ, and mapping dimensions are fixed after initial validation, while the retained PCA rank and data-derived contribution coefficients vary with the available incremental data.

PCA remains a linear, variance-based representation and therefore may not fully capture abrupt or strongly nonlinear degradation shifts. ATBLS does not assume identical degradation physics across assets, although highly irregular trajectories may weaken participant comparison after normalized-life alignment. The framework consequently adapts to observed representation changes rather than explicitly identifying their physical cause.

D. RUL Prediction Framework

Based on the proposed ATBLS, the end-to-end RUL prediction framework illustrated in Fig. 4 proceeds through the following stages:

1) *Data acquisition*: Collect run-to-failure sensor trajectories under the considered operating conditions and assign the corresponding RUL labels. Each trajectory is represented as an ordered sequence of observations; an observation may correspond to a vibration snapshot in the bearing datasets or an operating cycle in C-MAPSS.

2) *Preprocessing*: Transform each observation into the required input representation while preserving its chronological position in the degradation trajectory. For vibration data, snapshot-wise Welch analysis is used to extract frequency-domain features, whereas cycle-based datasets retain one feature row per operating cycle after the corresponding preprocessing.

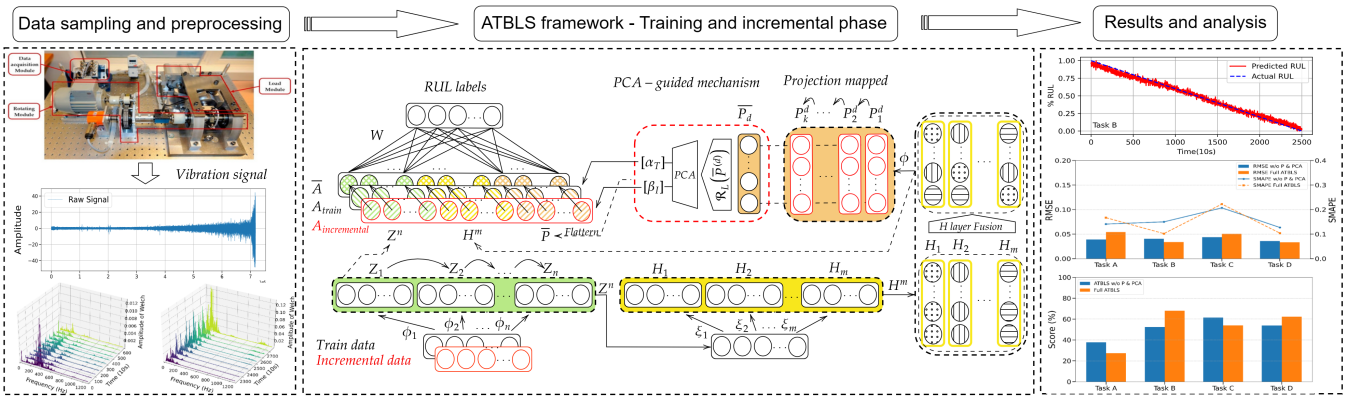


Fig. 4. Overall ATBLS architecture. Preprocessed inputs pass through fixed sequential feature mappings, enhancement fusion, and projection mappings. Projection representations provide shared PCA-based participant and domain contribution weights, followed by one analytical ridge solution for RUL prediction.

3) *Source learning*: Initialize the feature, enhancement, fusion, and projection mappings, construct the source transition matrices, and obtain the initial output weights through closed-form ridge regression.

4) *Incremental adaptation*: When labeled incremental trajectories become available, pass their observations through the same fixed representation mappings, update the shared PCA representation using the available source and incremental projection descriptors, and recompute the participant relevance and source–incremental domain contributions. The weighted participating data are then used to obtain one updated ridge solution without end-to-end retraining.

5) *Prediction and evaluation*: Pass the unseen test trajectory through the same fixed mappings, construct its transition matrix, and estimate RUL for each observation. Prediction performance is evaluated using RMSE, SMAPE, and Score.

IV. Experimental Results

We evaluate ATBLS on three benchmark run-to-failure datasets, reporting RMSE, SMAPE, Score, and runtime. Case Study 1 provides controlled component-ablation, sensitivity, and computational-cost analyses, while Case Studies 2 and 3 assess predictive accuracy, incremental adaptation, and cross-domain performance under different assets, operating conditions, and degradation patterns.

This experimental design separates component-level validation from broader adaptation and cross-domain assessment. For fair interpretation, comparisons are described as partition-aligned when the dataset and task partitions correspond to those of the referenced study. When such alignment cannot be established, published baselines are treated as representative literature-reported references because preprocessing, input features, tuning procedures, implementation details, and operating conditions may differ. Cross-study runtimes are reported descriptively and are not treated as directly comparable when hardware or implementation environments differ. Accordingly, the main validation evidence is provided by the within-paper component ablations, sensitivity analyses, controlled runtime measurements, within-domain incremental experiments, and explicit cross-domain adaptation studies.

A. Case Study 1: IEEE PHM Challenge 2012 Bearing Datasets

1) Dataset description and experimental arrangement: ATBLS was evaluated on the PRONOSTIA bearing dataset [31] under two operating conditions. Each vibration snapshot contains 2,560 samples acquired over 0.1 s, with successive snapshots recorded every 10 s throughout the run-to-failure trajectory. Tasks A–D use Bearing1 data acquired at 4000 N and 1800 r/min, whereas Tasks E–H use Bearing2 data acquired at 4200 N and 1650 r/min. The source, incremental, and unseen test partitions are summarized in Table II.

TABLE II
Dataset partitioning for Case Study 1.

Condition	Task	Source data	Inc. data	Test data
4000 N / 1800 r/min	A	Bearing1_1	–	Bearing1_6
	B	Bearing1_1	Bearing1_2	Bearing1_6
	C	Bearing1_1,2	Bearing1_3	Bearing1_6
	D	Bearing1_1,2,3	Bearing1_5	Bearing1_6
4200 N / 1650 r/min	E	Bearing2_1	–	Bearing2_6
	F	Bearing2_1	Bearing2_2	Bearing2_6
	G	Bearing2_1,2	Bearing2_4	Bearing2_6
	H	Bearing2_1,2,4	Bearing2_7	Bearing2_6

2) Evaluation metrics: Root-mean-square error (RMSE), symmetric mean absolute percentage error (SMAPE), and the prognostic Score are used to evaluate RUL prediction:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}, \qquad \mathrm{SMAPE} = \frac{1}{n}\sum_{i=1}^{n}\frac{|\hat{y}_i - y_i|}{\frac{1}{2}(|y_i| + |\hat{y}_i|)}, \tag{30}$$

$$\mathrm{Score} = \frac{1}{n}\sum_{i=1}^{n} A_i, \quad A_i = \begin{cases} e^{-\ln(0.5)(E_{r_i}/5)}, & E_{r_i} \le 0, \\ e^{+\ln(0.5)(E_{r_i}/20)}, & E_{r_i} > 0, \end{cases} \tag{31}$$

where $E_{r_i} = (\mathrm{RUL}_i - \widehat{\mathrm{RUL}}_i)/\mathrm{RUL}_i \times 100$. The final point, for which the actual RUL is zero, is excluded from the Score calculation. The mean Score is reported as a percentage in the experimental results. Higher Score and lower RMSE and SMAPE indicate better prediction performance.

3) Data preprocessing and label normalization: Each vibration snapshot is transformed into a compact frequency-domain representation using Welch spectral estimation while preserving its chronological position in the run-to-failure sequence. The resulting trajectory is supplied to ATBLS without using unseen test labels during model construction or contribution estimation.

Normalized RUL is assigned according to

$$\mathrm{RUL}_{\mathrm{label}} = \frac{Total_{\mathrm{life}} - Operating_{\mathrm{life}}}{Total_{\mathrm{life}}}, \tag{32}$$

where $Total_{\mathrm{life}}$ denotes the complete run-to-failure duration and $Operating_{\mathrm{life}}$ is the elapsed operating time, giving $\mathrm{RUL}_{\mathrm{label}} \in [0, 1]$.

4) RUL prediction results: Figure 5 presents the RUL predictions obtained using the complete ATBLS architecture, including enhancement-level fusion, sequential projection, and the final analytical ridge solution. For source-only Tasks A and E, PCA-based contribution estimation is inactive because no incremental data are available.

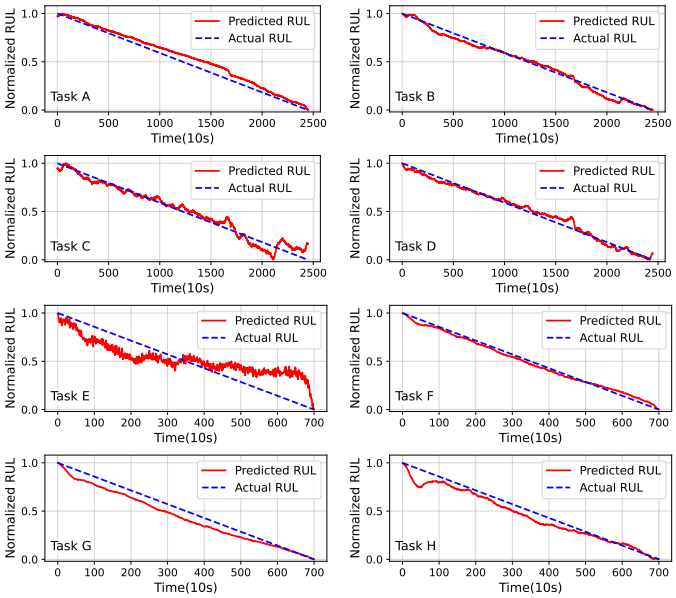


Fig. 5. RUL prediction results for Case Study 1 Tasks A–H using ATBLS.

Table III summarizes the corresponding prediction results. Tasks A and E contain source data only and therefore provide the initial reference for each operating condition. Introducing the first incremental bearing produces a substantial improvement in both protocols. From A to B, RMSE decreases by approximately 37.1% and SMAPE by 38.8%, while the Score increases from 27.31% to 67.91%. Similarly, E to F reduces RMSE by approximately 82.8% and SMAPE by 73.4%, while increasing the Score from 33.65% to 67.27%.

Later incremental stages are not strictly monotonic because different bearing trajectories are incorporated at each stage. Nevertheless, Tasks D and H retain RMSE values of 0.0334 and 0.0637, respectively, with Scores above 62%. All reported ATBLS training times remain below 2.4 s in the present implementation.

5) Ablation study:

TABLE III
ATBLS PERFORMANCE FOR CASE STUDY 1 TASKS A–H.

Task	RMSE↓	SMAPE↓	Score (%)↑	Time (s)
Protocol 1: 4000 N / 1800 r/min				
A	0.0540	0.1670	27.31	0.799
B	0.0340	**0.1021**	**67.91**	1.326
C	0.0508	0.2224	53.94	1.652
D	**0.0334**	0.1041	62.34	2.321
Protocol 2: 4200 N / 1650 r/min				
E	0.1576	0.4007	33.65	0.260
F	**0.0271**	**0.1066**	**67.27**	0.650
G	0.0689	0.1443	63.33	1.593
H	0.0637	0.1353	64.00	1.687

a) Effect of enhancement fusion and PCA-guided weighting: The roles of enhancement fusion and PCA-guided contribution weighting are examined using four controlled ATBLS configurations: PCA ON/Fusion ON (*Full ATBLS*), PCA ON/Fusion OFF, PCA OFF/Fusion ON, and PCA OFF/Fusion OFF. The feature, enhancement, projection, and ridge structures are otherwise unchanged; hence, these configurations represent component ablations of the same ATBLS framework rather than separate learning strategies.

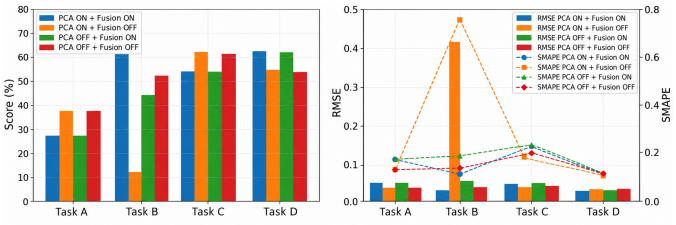


Fig. 6. Effect of enhancement fusion and PCA-guided contribution weighting on Case Study 1 Tasks A–D. Left: prognostic Score. Right: RMSE bars and SMAPE curves for the four controlled configurations.

Figure 6 shows that the two mechanisms interact rather than providing uniform independent improvements. For the source-only Task A, switching PCA on or off produces identical results for a fixed fusion setting. This is expected because no incremental dataset is available and the source–incremental PCA contribution mechanism is therefore not activated.

The strongest interaction occurs in Task B. With enhancement fusion enabled, activating PCA-guided weighting reduces RMSE from 0.0607 to 0.0340 (approximately 44.0%) and increases the Score from 44.31% to 67.91%. Conversely, PCA ON/Fusion OFF produces an RMSE of 0.4188 and a Score of only 12.41%. This indicates that the effectiveness of PCA-guided contribution weighting depends strongly on the representation from which the projection descriptors are constructed. Task D shows a similar, although smaller, benefit: Full ATBLS achieves the lowest RMSE (0.0334) and highest Score (62.34%) among the four configurations.

Task C exhibits a different response, with fusion-disabled configurations yielding lower pointwise errors and a higher Score. Thus, neither fusion nor PCA should be interpreted as providing a monotonic improvement for every degradation trajectory. Across Tasks A–D, however, Full ATBLS obtains the highest mean Score of 52.87%, compared with 41.61%, 46.80%, and 51.26% for PCA ON/Fusion OFF, PCA OFF/Fusion ON, and PCA OFF/Fusion OFF, respectively.

These results support the use of enhancement fusion and PCA-guided weighting as coupled components of the ATBLS adaptation pathway, while also showing that their influence remains task dependent.

b) Effect of ridge regularization and transition components: To further examine ATBLS, Fig. 7 reports a controlled study over the incremental Tasks B–D. The left panel evaluates the sensitivity of Full ATBLS to the ridge parameter λ using normalized RMSE, SMAPE, and Score. The right panel compares three nested configurations within the same framework: ATBLS w/o P & PCA, ATBLS w/o PCA, and Full ATBLS. These configurations represent component reductions of the same ATBLS pathway rather than separate learning strategies.

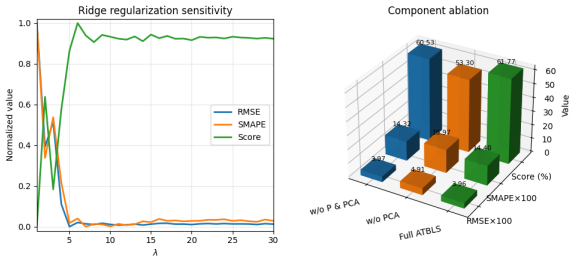


Fig. 7. Additional ablation results for Case Study 1. Left: sensitivity of Full ATBLS to the ridge parameter λ. Right: controlled component comparison among ATBLS w/o P & PCA, ATBLS w/o PCA, and Full ATBLS.

The λ-sweep shows a short sensitivity region at small regularization values, after which RMSE and SMAPE remain low and the Score stabilizes over a broad range, indicating limited sensitivity to moderate λ. The component ablation shows that introducing the projection representation without PCA-guided weighting degrades the average performance, yielding RMSE 0.0491, SMAPE 0.1697, and Score 53.30%. With PCA-guided weighting enabled, Full ATBLS reduces RMSE to 0.0396 and increases the Score to 61.77%, compared with 0.0397 and 60.53% for ATBLS w/o P & PCA, while SMAPE remains comparable at 0.1440 and 0.1432, respectively. These results indicate that the projection representation is most effective when coupled with PCA-guided contribution weighting, with the complete ATBLS providing the strongest overall balance across the evaluated metrics.

c) Impact of structural width: The sensitivity of ATBLS to model width is examined by varying the number of feature-mapping groups and enhancement nodes while retaining the remaining architecture and data partitions. The three curves correspond to nested component configurations of the same ATBLS framework: ATBLS w/o P & PCA, ATBLS w/o PCA, and Full ATBLS. These experiments therefore evaluate structural sensitivity and computational trade-offs rather than alternative learning strategies.

Figure 8 shows that increasing the number of feature-mapping groups does not produce a monotonic accuracy improvement. The best overall performance occurs in the compact region, with three feature-mapping groups giving Full ATBLS a Score above 60%, RMSE near 0.04, and the lowest SMAPE among the evaluated widths. As the number of groups increases, RMSE and SMAPE generally rise while

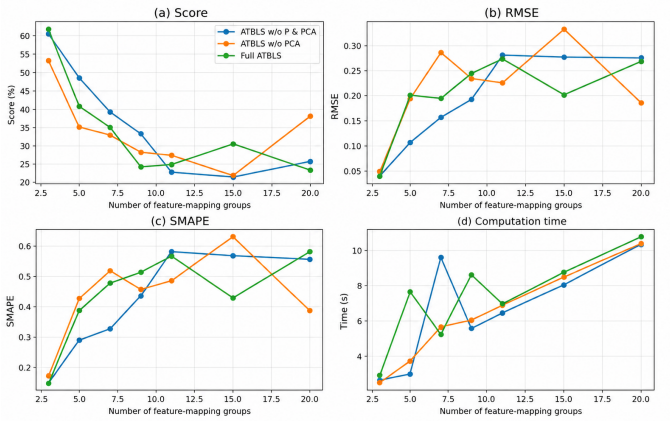


Fig. 8. Sensitivity to the number of feature-mapping groups. Panels show Score, RMSE, SMAPE, and computation time for ATBLS w/o P & PCA, ATBLS w/o PCA, and Full ATBLS.

computation time also increases. Although individual configurations recover performance at some intermediate widths, the overall trend indicates that excessive feature expansion adds computational cost without consistently improving degradation prediction.

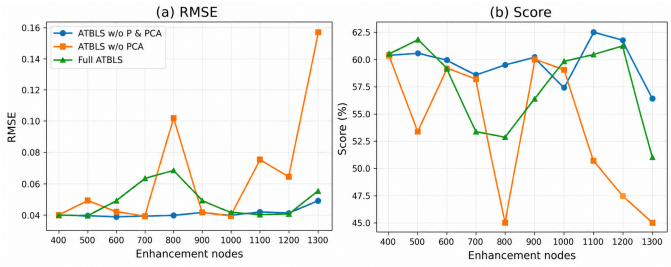


Fig. 9. Sensitivity to the number of enhancement nodes. Panels show RMSE and Score for ATBLS w/o P & PCA, ATBLS w/o PCA, and Full ATBLS.

The enhancement-node analysis in Fig. 9 similarly shows that larger enhancement layers do not guarantee better prediction. The configuration without projection remains comparatively stable in RMSE over much of the evaluated range, whereas introducing projection without PCA-guided weighting produces pronounced error increases at several widths. Full ATBLS moderates these variations and maintains comparatively stable Score values across much of the range, although the optimum remains task and width dependent. Together, the two sensitivity studies indicate that ATBLS benefits from a compact representation and that increasing structural capacity beyond an appropriate range can increase computation and variability without providing consistent accuracy gains.

d) Computational cost: At each adaptation stage, ATBLS passes the participating datasets through the fixed feature, enhancement-fusion, and projection mappings, recomputes the shared PCA representation and associated contribution coefficients, and solves one contribution-weighted ridge system. For a weighted design matrix $\widetilde{A} \in \mathbb{R}^{N_t \times d_A}$, the primal ridge formulation requires approximately $\mathcal{O}(N_t d_A^2 + d_A^3)$, while the dual formulation can be used when $d_A > N_t$, reducing the matrix inversion to the sample dimension. The shared PCA

calculation adds an SVD cost on the fixed-row projection descriptors, whereas enhancement fusion and sequential projection require no iterative optimization.

Because the feature, enhancement, fusion, and projection mappings remain fixed, adaptation requires neither backpropagation nor end-to-end network retraining. The measured ATBLS computation times in Table III remain below 2.4 s for all eight Case Study 1 tasks, including the successive incremental adaptation stages.

6) Comparison with representative reported methods: Table IV compares ATBLS Task D with representative results reported for the IEEE PHM 2012 bearing dataset. TCBLS [19] is treated as the closest partition-aligned published reference because its training and test-bearing selection is closely related to the present experiment. The remaining methods are included as contextual literature references because their preprocessing, feature construction, parameter settings, and experimental protocols differ.

TABLE IV
PARTITION-ALIGNED AND CONTEXTUAL LITERATURE-REPORTED RESULTS ON THE PHM 2012 BEARING DATASET.

Method	RMSE↓	Score (%)↑	Reported time (s)†
Closest partition-aligned published reference			
TCBLS [19]	0.1381	42.93	40.58
Contextual literature-reported references			
RVFLNN [18]	0.3912	15.08	32.15
BLS [16]	0.2485	32.62	38.10
Recurrent BLS [16]	0.2054	33.94	41.16
Vanilla LSTM [32]	0.4975	13.26	298.33
Double CNN [22]	0.3184	15.71	3825.60
Separable CNN [24]	0.2646	30.52	2724.48
Multiscale CNN [23]	0.2715	28.45	3065.05
TCN + Attention [33]	0.1142	44.70	40561.21
Sparse AE [34]	0.0294	68.50	15.20
SAL-CNN [35]	0.0278	70.10	18.45
Weibull-Loss ML [36]	0.0256	72.30	20.67
ATBLS (Task D)	**0.0334**	**62.34**	**2.32**

† Published runtimes were obtained under different hardware, software, and implementation environments and are therefore reported descriptively rather than treated as directly comparable.

Under the present ATBLS protocol, Task D achieves an RMSE of 0.0334 and a Score of 62.34%. Relative to the closest partition-aligned TCBLS result (RMSE 0.1381, Score 42.93%), ATBLS provides lower prediction error and a higher prognostic Score. It also compares favourably with several traditional and deep-learning references reported on the same benchmark. However, Sparse AE, SAL-CNN, and Weibull-informed learning report lower RMSE and higher Score under their respective experimental settings. These results are therefore presented as contextual references rather than direct controlled comparisons.

B. Case Study 2: C-MAPSS Turbofan Engine Datasets

1) Dataset description, experimental arrangement, and evaluation metrics: The C-MAPSS benchmark contains run-to-failure turbofan trajectories under different operating conditions and fault modes. Fourteen sensor variables are retained for the experiments. Each engine is represented by an ordered sequence of operating cycles, where one cycle corresponds to one observation in the degradation trajectory.

Four experimental protocols are considered. Protocols 1 and 2 evaluate within-domain incremental learning on FD001 and FD004, respectively. Protocol 3 evaluates explicit cross-domain adaptation through three matched direct-transfer/incremental pairs: FD001→FD002, FD004→FD003, and FD003→FD004. Protocol 4 is introduced separately for conventional benchmark comparison and cumulative FD002→FD004 adaptation following the corresponding incremental setting reported by MSBLS [20]. Table V summarizes Protocols 1–3, while Table VI summarizes Protocol 4.

TABLE V
CASE STUDY 2 PROTOCOLS 1–3.

Protocol	Task	Source data	Inc. data	Test data
1: FD001	A	FD001 No.1–3	–	FD001 No.10
	B	FD001 No.1–3	FD001 No.4&5	FD001 No.10
	C	FD001 No.1–5	FD001 No.6&7	FD001 No.10
	D	FD001 No.1–7	FD001 No.8&9	FD001 No.10
2: FD004	E	FD004 No.1–3	–	FD004 No.10
	F	FD004 No.1–3	FD004 No.4&5	FD004 No.10
	G	FD004 No.1–5	FD004 No.6&7	FD004 No.10
	H	FD004 No.1–7	FD004 No.8&9	FD004 No.10
3: Cross-domain	I	FD001 No.1–3	–	FD002 No.10
	J	FD001 No.1–3	FD002 No.4&5	FD002 No.10
	K	FD004 No.1–3	–	FD003 No.10
	L	FD004 No.1–3	FD003 No.4&5	FD003 No.10
	M	FD003 No.1–3	–	FD004 No.10
	N	FD003 No.1–3	FD004 No.4&5	FD004 No.10

Protocols 1–2 add data within the same C-MAPSS domain and therefore evaluate incremental learning rather than operating-condition transfer. Protocol 3 explicitly evaluates changing-condition adaptation. Tasks I/J transfer from FD001, containing one operating condition and one fault mode, to FD002, containing six operating conditions with the same fault-mode count. K/L evaluate the reverse operating-condition direction from FD004 to FD003 while retaining two fault modes, whereas M/N evaluate FD003→FD004. Within each pair, the source and unseen target test data are fixed; only the second task introduces labeled target-domain incremental data.

TABLE VI
CASE STUDY 2 PROTOCOL 4 CONVENTIONAL AND CUMULATIVE TASKS.

Task	Source data	Inc. data	Test data	Target data
P	FD002 train	–	FD002 test	–
Q	FD004 train	–	FD004 test	–
O0	FD002 train	–	FD004 test	0%
O1–O10	FD002 train	FD004 cumulative	FD004 test	10–100%

As summarized in Table VI, Tasks P and Q establish the conventional FD002 and FD004 benchmarks. O0 provides direct FD002→FD004 transfer without target-domain adaptation. Following the cumulative incremental setting reported by MSBLS [20], O1–O10 progressively introduce FD004 training data in ten cumulative stages while keeping the FD004 test set unchanged. Protocol 4 therefore separates conventional benchmark performance from cumulative cross-domain adaptation.

For Protocols 1–3, RMSE and SMAPE follow Eq. (30), and the normalized prognostic Score (%) follows Eq. (31). Protocol 4 additionally reports the conventional PHM08 Score.

For prediction error $e_i = \widehat{y}_i - y_i$, the PHM08 Score is defined in Eq. (33) as

$$S_{\mathrm{PHM08}} = \sum_{i=1}^{N} \begin{cases} \exp(-e_i/13) - 1, & e_i < 0, \\ \exp(e_i/10) - 1, & e_i \geq 0, \end{cases} \tag{33}$$

where lower values indicate better performance and late RUL predictions receive a larger penalty.

2) Data preprocessing and label normalization: For Protocols 1–3, each engine is represented as an ordered trajectory of the 14 retained sensor variables. Complete trajectories are aligned to a common ordered RUL grid, and causal Welch-based features are extracted while preserving the chronological degradation structure. The normalized RUL trajectory remains at unity for the initial 135 aligned positions and then decreases linearly to zero.

Protocol 4 follows a separate conventional C-MAPSS setting to support comparison with published methods, particularly MSBLS [20]. Consistent with the MSBLS experimental setting, a 30-cycle temporal window is used to construct the ATBLS input representation. In the present Protocol 4 implementation, a maximum RUL of 125 cycles is adopted, and predictions are returned to the cycle scale before RMSE and PHM08 evaluation.

3) RUL prediction results: Table VII reports the ATBLS performance for Protocols 1–3.

TABLE VII
ATBLS PERFORMANCE FOR CASE STUDY 2 PROTOCOLS 1–3.

Task	RMSE↓	SMAPE↓	Score (%)↑	Time (s)
Protocol 1: FD001 within-domain				
A	0.0754	0.1691	70.20	0.154
B	0.0569	0.0611	**81.15**	1.601
C	0.0406	0.0590	70.69	0.861
D	**0.0341**	**0.0541**	73.54	0.926
Protocol 2: FD004 within-domain				
E	0.0468	0.0768	79.75	0.155
F	**0.0218**	**0.0391**	**86.07**	0.752
G	0.0262	0.0516	82.64	0.875
H	0.0448	0.0851	75.72	1.682
Protocol 3: Cross-domain adaptation				
I	0.3840	0.6227	32.55	0.150
J	**0.1404**	**0.1464**	**63.60**	0.341
K	0.4338	0.6870	29.93	0.143
L	**0.0829**	**0.2315**	**59.73**	1.650
M	0.2732	0.5946	30.90	0.376
N	**0.0583**	**0.0900**	**74.63**	0.410

Within-domain Tasks A–H evaluate incremental updating under an unchanged dataset domain. As reported in Table VII, Protocol 1 reduces RMSE from 0.0754 in Task A to 0.0341 in Task D, while Task B obtains the highest Protocol 1 Score of 81.15%. In Protocol 2, Task F obtains the lowest RMSE of 0.0218 and the highest Score of 86.07%. The results are not required to vary monotonically with each increment because each stage introduces different engine trajectories.

4) Comparison with the partition-aligned TCBLS baseline: Protocols 1–2 permit a partition-aligned comparison with the published TCBLS experiments [19], since the FD001 and FD004 source, incremental, and test-unit progressions correspond to the TCBLS task definitions. Table VIII compares the current ATBLS results with the published TCBLS values. This is not a controlled reimplementation because the two approaches use different preprocessing, feature representations, and learning architectures.

TABLE VIII
PARTITION-ALIGNED COMPARISON OF ATBLS WITH PUBLISHED TCBLS RESULTS ON C-MAPSS PROTOCOLS 1–2.

Task / Method	RMSE↓	SMAPE↓	Score (%)↑
Protocol 1: FD001			
A (ATBLS)	**0.0754**	**0.1691**	**70.20**
A (TCBLS [19])	0.2834	0.3706	34.88
B (ATBLS)	**0.0569**	**0.0611**	**81.15**
B (TCBLS [19])	0.2313	0.3176	36.17
C (ATBLS)	**0.0406**	**0.0590**	**70.69**
C (TCBLS [19])	0.1878	0.2176	58.89
D (ATBLS)	**0.0341**	**0.0541**	**73.54**
D (TCBLS [19])	0.1039	0.1766	60.08
Protocol 2: FD004			
E (ATBLS)	**0.0468**	**0.0768**	**79.75**
E (TCBLS [19])	0.2422	0.4176	30.90
F (ATBLS)	**0.0218**	**0.0391**	**86.07**
F (TCBLS [19])	0.1903	0.2553	40.23
G (ATBLS)	**0.0262**	**0.0516**	**82.64**
G (TCBLS [19])	0.1287	0.2341	40.62
H (ATBLS)	**0.0448**	**0.0851**	**75.72**
H (TCBLS [19])	0.1184	0.2270	41.87

As shown in Table VIII, ATBLS obtains lower RMSE and SMAPE and a higher prognostic Score than the published TCBLS results across all Tasks A–H. These results provide a partition-aligned comparison with a closely related incremental BLS method while avoiding claims of a fully controlled architectural comparison.

The corresponding within-domain ATBLS prediction trajectories are shown in Fig. 10.

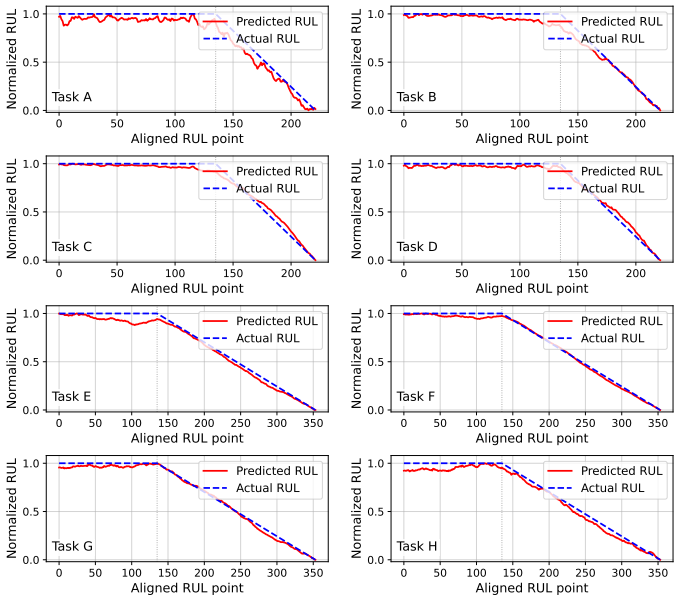


Fig. 10. RUL prediction results for Case Study 2 within-domain Tasks A–H. Tasks A–D correspond to Protocol 1 and Tasks E–H to Protocol 2.

Protocol 3 provides the explicit changing-condition evidence. As reported in Table VII, I/J use the same FD001 source and FD002 target test trajectory, but J additionally uses labeled FD002 incremental data. This reduces RMSE from 0.3840 to 0.1404, a 63.45% reduction, while Score increases from 32.55% to 63.60%. Similarly, K→L reduces RMSE by 80.88% when FD003 incremental information is introduced, and M→N reduces RMSE by 78.65%. The corresponding Score increases are 29.80 and 43.73 percentage points, respectively. Because the source and target-test partitions are

fixed within each pair, these changes isolate the effect of the available target-domain incremental data.

The corresponding cross-domain prediction trajectories are shown in Fig. 11.

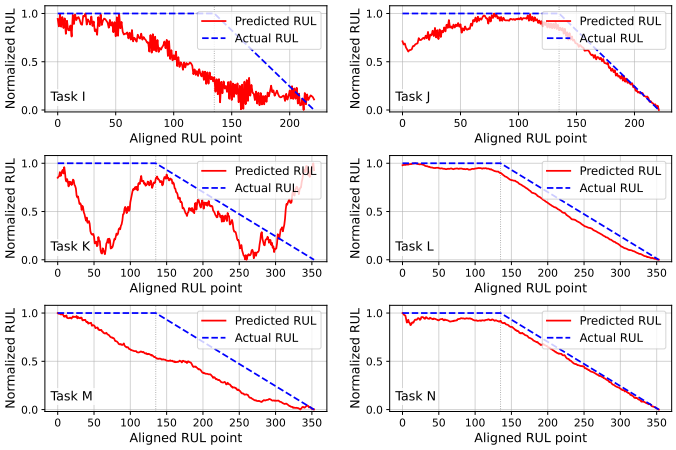


Fig. 11. RUL prediction results for Case Study 2 cross-domain Tasks I–N. I/K/M are direct-transfer references, while J/L/N include labeled target-domain incremental data.

Across Protocols 1–3, the measured ATBLS training time remains below 1.7 s for all reported tasks in the present implementation. These timings are implementation-specific and are reported only to characterize the computational cost of the analytical adaptation procedure.

5) Protocol 4 benchmark and cumulative adaptation: Protocol 4 evaluates the separate conventional benchmark and cumulative adaptation setting defined in Table VI. Tasks P and Q establish conventional FD002 and FD004 predictive performance, O0 provides the direct FD002→FD004 transfer baseline, and O1–O10 progressively introduce target-domain FD004 training data.

The RMSE distributions over ten runs are shown in Fig. 12.

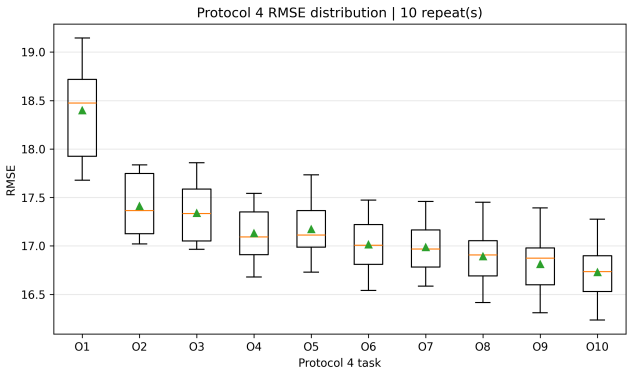


Fig. 12. RMSE distributions over ten runs for Case Study 2 Protocol 4 cumulative adaptation Tasks O1–O10. The tasks progressively incorporate 10%–100% of the available FD004 target-domain training data while the FD004 test set remains fixed. Lower values are better.

6) Comparison with representative recent baselines: Table IX compares the conventional P/Q experiments with MSBLS [20], HDA-GNN [37], and LVDACNN [38]. Because these methods use different model structures and implementations, the comparison is interpreted as benchmark-level evidence rather than a controlled architectural or runtime comparison.

TABLE IX
CONVENTIONAL C-MAPSS COMPARISON. LOWER VALUES ARE BETTER.

Method	FD002		FD004	
	RMSE↓	PHM08↓	RMSE↓	PHM08↓
MSBLS [20]	16.39	1434.27	17.75	1785.33
HDA-GNN [37]	19.06	1343.24	18.27	1214.81
LVDACNN [38]	12.19	643.27	14.46	979.92
ATBLS (P/Q)	15.41 ± 0.32	941.38 ± 68.05	16.74 ± 0.38	1536.60 ± 219.97

As shown in Table IX, ATBLS obtains lower RMSE than MSBLS and HDA-GNN on both FD002 and FD004. Its PHM08 Score is also lower than MSBLS on both datasets and lower than HDA-GNN on FD002, while HDA-GNN and LVDACNN obtain lower FD004 PHM08 Scores. LVDACNN gives the strongest conventional benchmark results among the compared methods. These comparisons concern conventional prediction and are distinct from the cumulative adaptation experiment.

Table X compares the cumulative FD002→FD004 adaptation results with the corresponding incremental trend reported by MSBLS [20]. The MSBLS values are approximate readings from its published figure and are therefore interpreted at trend level rather than as results from a common implementation.

TABLE X
CUMULATIVE FD002→FD004 COMPARISON FOR CASE STUDY 2 PROTOCOL 4.

Task	Inc. (%)	ATBLS RMSE↓	MSBLS RMSE↓
O0	0	41.43 ± 0.57	–
O1	10	18.40 ± 0.52	≈ 21.2
O2	20	17.41 ± 0.34	≈ 21.4
O3	30	17.34 ± 0.31	≈ 21.3
O4	40	17.13 ± 0.29	≈ 21.0
O5	50	17.17 ± 0.30	≈ 20.4
O6	60	17.02 ± 0.29	≈ 19.5
O7	70	16.99 ± 0.27	≈ 22.0
O8	80	16.89 ± 0.31	≈ 24.4
O9	90	16.81 ± 0.31	≈ 24.3
O10	100	$\mathbf{16.73 \pm 0.30}$	≈ 24.2

As reported in Table X, direct FD002→FD004 transfer in O0 gives an RMSE of 41.43 ± 0.57. Introducing only 10% of the target-domain training data reduces RMSE to 18.40 ± 0.52, a 55.59% reduction. With all target-domain training data available, O10 reaches 16.73 ± 0.30, corresponding to a 59.62% reduction relative to O0. The progression is not strictly monotonic, but the overall trend demonstrates that target-domain information substantially reduces the direct-transfer error.

ATBLS is numerically below the approximate MSBLS RMSE trend at every reported O1–O10 stage. Because the MSBLS values are obtained from its published figure rather than from a common implementation, this comparison is interpreted at trend level.

C. Case Study 3: XJTU-SY Bearing Accelerated Degradation Dataset

1) Dataset description and experimental arrangement: The XJTU-SY run-to-failure bearing dataset was used to evaluate ATBLS under bearing-to-bearing degradation variability and explicit operating-condition changes. Two operating conditions

are considered: 2250 r/min (37.5 Hz) with an 11 kN radial load, and 2400 r/min (40 Hz) with a 10 kN radial load. Tasks A–D evaluate unseen bearings within their original operating condition, whereas Tasks E–H evaluate cross-condition transfer and adaptation between the two operating regimes, as summarized in Table XI. The XJTU-SY accelerated bearing degradation testbed is shown in Fig. 13.

TABLE XI
EXPERIMENTAL ARRANGEMENT FOR CASE STUDY 3.

Condition	Task	Source data	Inc. data	Test data
37.5 Hz / 11 kN	A	B2.1–B2.3	–	B2.4
	B	B2.1–B2.3	–	B2.5
40 Hz / 10 kN	C	B3.1–B3.3	–	B3.4
	D	B3.1–B3.3	–	B3.5
37.5 → 40 Hz	E	B2.1–B2.3	–	B3.4
	F	B2.1–B2.3	B3.1, B3.2	B3.4
40 → 37.5 Hz	G	B3.1–B3.3	–	B2.4
	H	B3.1–B3.3	B2.1, B2.2	B2.4

Each physical record contains horizontal and vertical vibration measurements sampled at 25.6 kHz. A record contains 32,768 samples, corresponding to 1.28 s of vibration data, and successive records are acquired at one-minute intervals throughout the run-to-failure experiment [39].

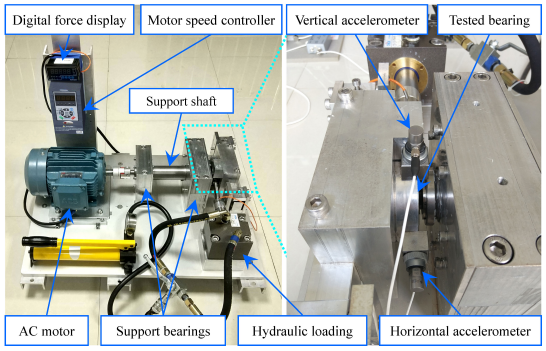


Fig. 13. XJTU-SY accelerated bearing degradation testbed.

Consequently, one CSV record is treated as one physical vibration snapshot and one RUL observation; snapshots from an entire bearing are not concatenated before feature extraction.

Tasks E/F examine transfer from the 37.5 Hz/11 kN condition to the 40 Hz/10 kN condition, whereas Tasks G/H evaluate the reverse direction. Within each matched pair, the source and unseen target test bearings are fixed; only Tasks F and H additionally introduce labeled degradation data from the target operating condition. The paired experiments therefore isolate the effect of newly available target-condition information under changes in both rotational speed and radial load. RMSE, SMAPE, and the prognostic Score are evaluated using Eqs. (30) and (31).

2) Data preprocessing and label normalization: The native chronological order and number of snapshots are preserved, and the same feature construction is applied to source, incremental, and unseen test bearings. Unseen test labels are not used during preprocessing, PCA-guided contribution estimation, or model adaptation. Normalized RUL decreases from 1 at the beginning of each processed run-to-failure trajectory to 0 at its final observation.

3) RUL prediction results: Tasks A–D evaluate bearing-to-bearing generalization without introducing incremental data. The corresponding ATBLS trajectories are shown in Fig. 14. All experiments use the full ATBLS architecture with three sequential projection mappings; for Tasks A–D, PCA-based contribution estimation is inactive because no incremental data are available.

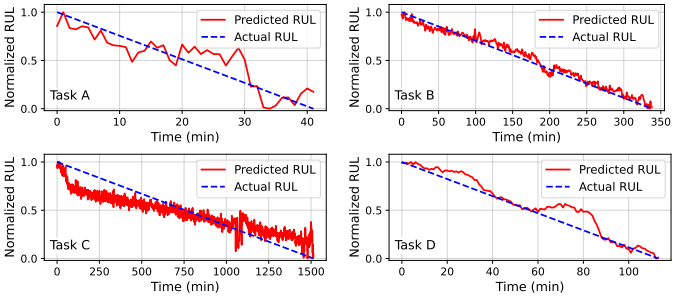


Fig. 14. ATBLS RUL prediction results for the within-condition XJTU-SY Tasks A–D.

The within-condition results in Table XII show RMSE values between 0.0472 and 0.1360 across bearings with different degradation durations. Task B provides the lowest within-condition error, with RMSE 0.0472 and SMAPE 0.1534, while obtaining a Score of 45.89%. These results demonstrate that prediction performance varies across individual bearing trajectories even when load and rotational speed are unchanged.

The cross-domain trajectories are shown in Fig. 15. Tasks E and G provide direct-transfer references before labeled target-condition bearings are incorporated, while F and H include available target-domain degradation data.

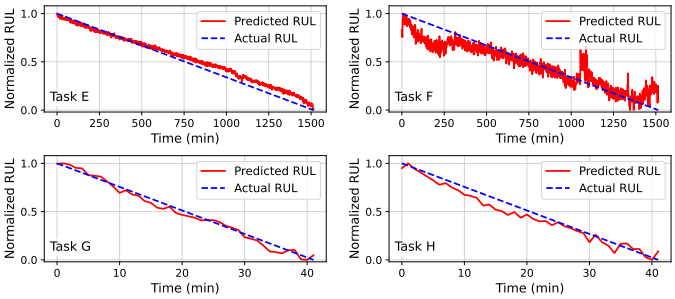


Fig. 15. ATBLS cross-domain RUL prediction results. Tasks E/F evaluate the 37.5 Hz/11 kN → 40 Hz/10 kN transition, whereas G/H evaluate the reverse transition. Tasks E and G contain no target-domain incremental data; F and H additionally incorporate target-condition bearings.

TABLE XII
ATBLS PERFORMANCE FOR CASE STUDY 3 TASKS A–H.

Task	RMSE↓	SMAPE↓	Score (%)↑	Time (s)
Within-condition evaluation				
A	0.1360	0.4161	35.21	1.706
B	**0.0472**	**0.1534**	**45.89**	1.532
C	0.1281	0.3202	41.60	4.481
D	0.0919	0.2213	36.95	2.336
Cross-domain evaluation				
E	0.0539	**0.2069**	41.03	4.096
F	0.1043	0.2384	51.26	5.103
G	**0.0370**	0.2339	**61.87**	1.812
H	0.0620	0.2428	58.38	2.202

The two cross-domain directions show different adaptation behaviour. In the forward 37.5 Hz/11 kN $\rightarrow$ 40 Hz/10 kN experiment, incorporating target-condition data increases the Score from 41.03% in Task E to 51.26% in Task F, corresponding to a 10.23 percentage-point, or 24.92% relative, increase. However, RMSE and SMAPE increase from 0.0539 and 0.2069 to 0.1043 and 0.2384, respectively. Thus, the forward transition does not provide a uniform improvement across all error measures; instead, the added target-domain information improves the prognostic Score while increasing the pointwise prediction errors.

The reverse 40 Hz/10 kN $\rightarrow$ 37.5 Hz/11 kN experiment exhibits a different response. Task G, representing direct transfer without labeled target-condition incremental data, achieves an RMSE of 0.0370, SMAPE of 0.2339, and Score of 61.87%. After incorporating target-condition incremental bearings in Task H, RMSE increases to 0.0620 and SMAPE to 0.2428, corresponding to relative increases of approximately 67.6% and 3.8%, respectively. The Score decreases to 58.38%, a reduction of 3.49 percentage points, or approximately 5.6% relative to the direct-transfer reference.

Taken together, the two transfer directions show that the effect of target-domain incremental information is condition dependent rather than uniformly beneficial. In the forward transition, the added target-domain data improve the prognostic Score but not RMSE or SMAPE, whereas in the reverse transition the direct-transfer baseline remains stronger across all three reported metrics. These results indicate that ATBLS can incorporate target-condition information under operating-condition changes, while the resulting predictive benefit depends on the representativeness and compatibility of the newly available incremental data.

4) Comparison with representative reported methods: Table XIII places the within-condition ATBLS results alongside representative XJTU-SY results reported in the literature. These values are used as contextual references because preprocessing, bearing partitions, feature construction, and evaluation protocols differ across studies; consequently, they are not treated as controlled head-to-head comparisons.

TABLE XIII
ATBLS WITHIN-CONDITION RESULTS AND REPRESENTATIVE LITERATURE-REPORTED XJTU-SY RMSE VALUES.

Task / Method	RMSE↓	SMAPE↓	Score (%)↑
ATBLS within-condition tasks			
Task A	0.1360	0.4161	35.21
Task B	**0.0472**	**0.1534**	**45.89**
Task C	0.1281	0.3202	41.60
Task D	0.0919	0.2213	36.95
Representative literature-reported results			
CNN-GRU-MHA [40]	0.0691	–	–
TCNMABG [41]	0.0822	–	–
CT-MLP + CWT [42]	≈ 0.195	–	–
Wiener + EIKF [43]	0.09–0.12	–	–

Among the within-condition ATBLS experiments, Task B attains an RMSE of 0.0472, numerically below the representative literature-reported RMSE values listed in Table XIII. The remaining ATBLS tasks illustrate the substantial bearing-to-bearing variability of XJTU-SY, with RMSE ranging from 0.0919 to 0.1360. Because the cited studies employ different experimental arrangements, these values provide only contextual evidence rather than a direct method ranking.

The cross-domain experiments serve a different purpose from these literature comparisons: they explicitly examine the response of ATBLS when the operating condition changes and labeled degradation information from the target condition becomes available.

Implementation Details: All ATBLS experiments were implemented in Python using NumPy, SciPy, and scikit-learn on a consumer-grade AMD Ryzen 9 laptop with 16 GB RAM. GPU acceleration was not used. Runtimes were CPU-based averages over ten runs.

V. CONCLUSION AND FUTURE WORK

This paper proposed the Adaptive Transfer Broad Learning System (ATBLS) for incremental RUL prediction under changing operating conditions. ATBLS combines three complementary mechanisms within a single flat learning framework: structured enhancement fusion with lightweight sequential projection for degradation representation, shared PCA-based participant and domain contribution estimation, and weighted closed-form ridge adaptation with fixed representation mappings. By retaining the feature, enhancement, fusion, and projection mappings during adaptation, newly available labeled data can be incorporated without end-to-end network retraining.

ATBLS was evaluated across bearing and turbofan case studies involving within-domain incremental learning, operating-condition variation, and explicit cross-domain transitions. The results show that the framework can exploit newly available degradation data while maintaining low analytical adaptation cost. The ablation studies further demonstrate that enhancement fusion, sequential projection, and PCA-guided contribution weighting have complementary and task-dependent effects rather than acting as independent learning strategies. Several limitations remain. The current experiments primarily consider structured run-to-failure trajectories with incremental data available after operating-condition changes.

Performance under abrupt nonstationary drift, unseen fault mechanisms, severe sensor degradation, missing or asynchronous measurements, and major changes in degradation characteristics requires further investigation. Under substantial domain shift, the shared PCA representation and associated contribution coefficients may need to be recomputed as new representative data become available. Future work will therefore investigate online drift detection, uncertainty-aware prediction, robustness to missing and corrupted sensor data, multimodal degradation representations, and longer-term industrial deployment under continuously changing operating conditions.

REFERENCES

[1] Y. Lei, N. Li, L. Guo, N. Li, T. Yan, and J. Lin, "Machinery health prognostics: A systematic review from data acquisition to RUL prediction," *Mechanical Systems and Signal Processing*, vol. 104, pp. 799–834, 2018.

[2] C. Zhang, W. Zhang, H. Chen, and C. Wang, "A review on deep learning applications in prognostics and health management," *IEEE Access*, vol. 8, pp. 18 768–18 784, 2020.

[3] Y. Wang, H. Yu, and L. Chen, "A comprehensive survey on remaining useful life prediction with deep learning," *IEEE Access*, vol. 7, pp. 115 908–115 931, 2019.

[4] X. Zhang, Z. Wang, J. Ma, S. Li, and M. Suo, "State-of-health prediction for reaction wheel of on-orbit satellite based on fourier broad learning system," *IEEE Access*, vol. 9, pp. 125 691–125 705, 2021.

[5] M. Yan, X. Wang, B. Wang, M. Chang, and I. Muhammad, "Bearing remaining useful life prediction using support vector machine and hybrid degradation tracking model," *ISA Transactions*, vol. 98, pp. 471–482, 2020.

[6] Y. Song, D. Liu, Y. Hou, J. Yu, and Y. Peng, "Satellite lithium-ion battery remaining useful life estimation with an iterative updated RVM fused with the KF algorithm," *Chinese Journal of Aeronautics*, vol. 31, no. 1, pp. 31–40, 2018.

[7] X. Wang, B. Jiang, and N. Lu, "Adaptive relevant vector machine based RUL prediction under uncertain conditions," *ISA Transactions*, vol. 87, pp. 217–224, 2019.

[8] W. Yu, I. Y. Kim, and C. Mechefske, "Remaining useful life estimation using a bidirectional recurrent neural network based autoencoder scheme," *Mechanical Systems and Signal Processing*, vol. 129, pp. 764–780, 2019.

[9] X. Wang, T. Huang, K. Zhu, and X. Zhao, "LSTM-based broad learning system for remaining useful life prediction," *Mathematics*, vol. 10, no. 12, p. 2066, 2022.

[10] H. Liu, Z. Liu, W. Jia, and X. Lin, "Remaining useful life prediction using a novel feature-attention-based end-to-end approach," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 2, pp. 1197–1207, 2020.

[11] Z. Chen, M. Wu, R. Zhao, F. Guretno, R. Yan, and X. Li, "Machine remaining useful life prediction via an attention-based deep learning approach," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 3, pp. 2521–2531, 2020.

[12] X. Li, J. Li, L. Zuo, L. Zhu, and H. T. Shen, "Domain adaptive remaining useful life prediction with transformer," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–13, 2022.

[13] T. Zhang, L. Jiang, R. Huang, and X. Zhang, "A multiscale cross-channel attention network for remaining useful life prediction with variable sensors," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–10, 2025.

[14] Z. Jia, L. Ren, and Y. Tang, "A contrastive representation domain adaptation method for industrial time-series cross-domain prediction," *IEEE Transactions on Industrial Informatics*, vol. 21, no. 4, pp. 3236–3245, 2025.

[15] K. Zhang, Q. Wen, C. Zhang, R. Cai, M. Jin, Y. Liu, J. Zhang, Y. Liang, G. Pang, D. Song, and S. Pan, "Self-supervised learning for time series analysis: Taxonomy, progress, and prospects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.

[16] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 1, pp. 10–24, 2017.

[17] C. L. P. Chen, Z. Liu, and S. Feng, "Universal approximation capability of broad learning system and its structural variations," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 4, pp. 1191–1204, 2018.

[18] Y.-H. Pao and Y. Takefuji, "Learning and generalization characteristics of the random vector functional-link net," *Neurocomputing*, vol. 6, no. 2, pp. 163–180, 1994.

[19] Y. Cao, M. Jia, P. Ding, X. Zhao, and Y. Ding, "Incremental learning for remaining useful life prediction via temporal cascade broad learning system with newly acquired data," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 4, pp. 6234–6245, 2022.

[20] T. Xu, G. Han, H. Zhu, C. Lin, and J. Peng, "Multiscale BLS-based lightweight prediction model for remaining useful life of aero-engine," *IEEE Transactions on Reliability*, 2024.

[21] P. Liu, X. Wang, L. Yin, and B. Liu, "Flat random forest: A new ensemble learning method towards better training efficiency and adaptive model size to deep forest," *International Journal of Machine Learning and Cybernetics*, vol. 11, pp. 2501–2513, 2020.

[22] B. Yang, R. Liu, and E. Zio, "Remaining useful life prediction based on a double-convolutional neural network architecture," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 12, pp. 9521–9530, 2019.

[23] J. Zhu, N. Chen, and W. Peng, "Estimation of bearing remaining useful life based on multiscale convolutional neural network," *IEEE Transactions on Industrial Electronics*, vol. 66, no. 4, pp. 3208–3216, 2018.

[24] B. Wang, Y. Lei, N. Li, and T. Yan, "Deep separable convolutional network for remaining useful life prediction of machinery," *Mechanical Systems and Signal Processing*, vol. 134, p. 106330, 2019.

[25] Z. Zhang, W. Song, and Q. Li, "Dual-aspect self-attention based on transformer for remaining useful life prediction," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–10, 2022.

[26] J. Zhang, H. Luo, J. Hu, C. Cheng, S. Wu, P. Yan, and J. Tian, "A transferable topology-aware graph pooling network for remaining useful life prediction under cross-domain conditions," *IEEE Transactions on Reliability*, 2025, early access.

[27] J. Zhang, K. Chen, F. Wu, G. Qiu, Y. Cheng, and S. Yin, "Open-circuit fault diagnosis for power converters under variable operational conditions: from a source-free domain adaptation perspective," *IEEE Transactions on Industrial Electronics*, 2026, early access.

[28] L. Ren, H. Wang, J. Dong, Z. Jia, S. Li, Y. Wang, Y. Laili, D. Huang, L. Zhang, and B. Li, "Industrial foundation model," *IEEE Transactions on Cybernetics*, 2025.

[29] Y. Ma, X. Guo, J. Wang, and J. Yang, "Mutual information optimization using large language model for domain-generalized remaining useful life prediction in IIoT," *IEEE Internet of Things Journal*, 2026.

[30] J. Yao and T. Han, "Utilizing large-scale foundation models for prognostics and health management in wind turbines: Techniques, challenges, and future directions," *Renewable and Sustainable Energy Reviews*, vol. 227, p. 116527, 2026.

[31] P. Nectoux, R. Gouriveau, K. Medjaher, E. Ramasso, B. Chebel-Morello, N. Zerhouni, and C. Varnier, "PRONOSTIA: An experimental platform for bearings accelerated degradation tests," in *Proc. IEEE International Conference on Prognostics and Health Management (PHM)*, 2012, pp. 1–8.

[32] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla LSTM neural networks," *Neurocomputing*, vol. 275, pp. 167–179, 2018.

[33] Y. Cao, Y. Ding, M. Jia, and R. Tian, "A novel temporal convolutional network with residual self-attention mechanism for remaining useful life prediction of rolling bearings," *Reliability Engineering and System Safety*, vol. 215, p. 107813, 2021.

[34] D. She, M. Jia, and M. Pecht, "Sparse autoencoder with regularization method for health indicator construction and remaining useful life prediction of rolling bearing," *Measurement Science and Technology*, vol. 31, no. 9, p. 095102, 2020.

[35] B. Liu, Z. Gao, B. Lu, H. Dong, and Z. An, "Deep learning-based remaining useful life estimation of bearings with time-frequency information," *Sensors*, vol. 22, no. 19, p. 7402, 2022.

[36] T. von Hahn and C. K. Mechefske, "Knowledge-informed machine learning using a Weibull-based loss function," *Journal of Prognostics and Health Management*, vol. 2, no. 1, pp. 9–44, 2022.

[37] Y. Jin, Z. Meng, Q. Hu, and D. Chen, "Heterogeneous dynamic-aware gnn for rul prediction of aeroengine," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–14, 2025.

[38] Y. Hou, C. Lu, Q. Xia, W. Abbas, H. H. Lee, and K.-H. Loo, "Lvdacnn: a lightweight variable dependency aware cnn for remaining useful life prediction," *IEEE Transactions on Instrumentation and Measurement*, 2026.

[39] B. Wang, Y. Lei, N. Li, and N. Li, "A hybrid prognostics approach for estimating remaining useful life of rolling element bearings," *IEEE Transactions on Reliability*, vol. 69, no. 1, pp. 401–412, 2020.

[40] J. Yu, J. Shao, X. Peng, T. Liu, and Q. Yao, "Remaining useful life of the rolling bearings prediction method based on transfer learning integrated with CNN-GRU-MHA," *Applied Sciences*, vol. 14, no. 19, p. 9039, 2024.

[41] X. Cao, F. Zhang, J. Zhao, Y. Duan, and X. Guo, "Remaining useful life prediction of rolling bearing based on multi-domain mixed features and temporal convolutional networks," *Applied Sciences*, vol. 14, no. 6, p. 2354, 2024.

[42] Z. Jin, C. Chen, A. Syntetos, and Y. Liu, "Enhanced bearing RUL prediction based on dynamic temporal attention and mixed MLP," *Autonomous Intelligent Systems*, vol. 5, no. 1, p. 2, 2025.

[43] J. Li, J. Fan, Z. Wang, Z. Zhang, X. Pang, and M. Qiu, "Enhanced RUL predictions of rolling bearings using a nonlinear Wiener model with an extended incremental Kalman filter," *Measurement Science and Technology*, vol. 36, no. 1, p. 016196, 2025.