

基于TSPU-Net的航天器测试数据分析方法及应用研究

邵寒琛¹, 尹溶森², 孙 波²

¹西北工业大学计算机学院, 陕西 西安

²北京空间飞行器总体设计部, 北京

收稿日期: 2025年7月5日; 录用日期: 2025年8月7日; 发布日期: 2025年8月13日

摘 要

针对航天器在轨温度预测中传统热网格法对瞬态温度预测效果不佳的问题, 本文提出了一种基于历史遥测数据驱动的多元时序数据预测模型——TSPU-Net。该模型通过设计时序处理单元(TSPU), 有效融合了一维卷积(Conv1D)、多尺度归一化、Dropout、GELU激活函数以及门控循环单元(GRU)等多种深度学习技术, 旨在全面提高航天器遥测数据的特征提取和表示能力, 并捕捉多尺度时间依赖关系。数值实验结果表明, TSPU-Net模型在参数量较少的情况下, 对航天器温度遥测数据具有较高的预测精度和普适性。与TCN和2D-CNN等主流模型相比, TSPU-Net在均方误差(MSE)、平均绝对误差(MAE)等评价指标上均表现出显著优势, 尤其在多步预测任务中展现出更强的鲁棒性。本研究为航天器故障预警、健康管理以及大规模星座数字化模型构建提供了新的解决方案, 具有重要的工程应用价值。

关键词

航天器, 遥测数据, 温度预测, 深度学习, TSPU-Net, 时间序列

Research on the Analysis Method and Application of Spacecraft Test Data Based on TSPU-Net

Hanchen Shao¹, Rongsen Yin², Bo Sun²

¹School of Computer Science, Northwestern Polytechnical University, Xi'an Shaanxi

²Beijing Space Vehicle Master Design Department, Beijing

Received: Jul. 5th, 2025; accepted: Aug. 7th, 2025; published: Aug. 13th, 2025

Abstract

Addressing the limitations of traditional thermal grid methods in predicting transient temperatures

文章引用: 邵寒琛, 尹溶森, 孙波. 基于 TSPU-Net 的航天器测试数据分析方法及应用研究[J]. 计算机科学与应用, 2025, 15(8): 58-72. DOI: 10.12677/csa.2025.158198

for on-orbit spacecraft, this paper proposes a multivariate time series prediction model driven by historical telemetry data, named TSPU-Net. By designing a Time Series Processing Unit (TSPU), this model effectively integrates various deep learning techniques, including one-dimensional convolutional neural networks (Conv1D), multi-scale normalization, Dropout, GELU activation function, and Gated Recurrent Unit (GRU), aiming to comprehensively enhance the feature extraction and representation capabilities of spacecraft telemetry data and capture multi-scale temporal dependencies. Numerical experimental results demonstrate that the TSPU-Net model achieves high prediction accuracy and universality for spacecraft temperature telemetry data with fewer parameters. Compared to mainstream models such as TCN and 2D-CNN, TSPU-Net exhibits significant advantages in evaluation metrics such as Mean Squared Error (MSE) and Mean Absolute Error (MAE), especially showing stronger robustness in multi-step prediction tasks. This research provides a novel solution for spacecraft fault early warning, health management, and the construction of large-scale constellation digital models, possessing significant engineering application value.

Keywords

Spacecraft, Telemetry Data, Temperature Prediction, Deep Learning, TSPU-Net, Time Series

Copyright © 2025 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 引言

航天器作为复杂的高科技系统,其在轨运行的安全性与可靠性至关重要。随着航天技术的飞速发展,航天器在轨时间不断延长,功能日益复杂,对航天器健康状态的实时监测与精准预测提出了更高要求。航天器长期安全稳定运行高度依赖对温度的精准监测与预测,因温度波动直接影响电子设备性能、材料寿命及任务成功率;传统物理模型(如热网格法)虽可预测稳态温度边界,但面对复杂空间环境(如太阳辐射、地球反照等)及航天器频繁姿态调整、载荷切换导致的瞬态非线性温度变化时,存在计算量大、建模精度不足、适应性差等局限,尤其在规模化星座任务中,单星精细化建模与实时预测的计算成本指数级增长,难以满足批量快速预测需求。

近年来,随着大数据和人工智能技术的快速发展,数据驱动的方法为解决航天器在轨温度预测问题提供了新的思路。特别是深度学习技术,凭借其强大的非线性拟合能力和从海量数据中自动提取复杂特征的优势,在时间序列预测领域展现出巨大的潜力。在航天器健康管理和故障诊断方面,深度学习模型已被广泛应用于遥测数据异常检测、寿命预测以及性能评估等任务[1][2]。

针对航天器温度预测,研究人员开始探索将深度学习模型应用于历史遥测数据,以期实现更精准、更快速的预测。例如,一些研究尝试使用循环神经网络(RNN)及其变体,如长短时记忆网络(LSTM)和门控循环单元(GRU),来捕捉航天器温度时间序列中的长期依赖关系[3][4]。卷积神经网络(CNN)也被引入,用于提取遥测数据在不同时间尺度上的局部特征[5]。然而,单一的深度学习模型往往难以全面捕捉航天器遥测数据所固有的高维度、强耦合、高动态、异质性、非平稳性等多元特性。例如,温度不仅受自身历史数据影响,还与航天器姿态、太阳帆板角度、设备功耗等非温度遥测数据存在复杂关联。如何有效融合这些多源异构数据,并同时捕捉其不同时间尺度上的依赖关系,是当前航天器温度预测领域面临的关键挑战。

为了克服上述挑战,研究者们开始尝试结合不同深度学习模型的优势,构建混合模型。例如,将 CNN

与 RNN 结合, 利用 CNN 提取空间特征, 再由 RNN 处理时间序列, 以期提升预测性能[6]。此外, 多尺度分析的思想也被引入, 通过在不同时间粒度上提取特征, 以更好地捕捉数据的复杂动态[7]。

尽管现有研究取得了一定进展, 但仍存在以下不足:

1) 多源异构数据融合的有效性: 如何更有效地融合非温度遥测数据与温度遥测数据, 并揭示它们之间复杂的非线性映射关系, 仍需深入研究。

2) 多尺度时间依赖关系的捕捉: 航天器温度变化可能受到短期瞬态事件和长期环境变化的影响, 现有模型在同时捕捉这些多尺度依赖关系方面仍有提升空间。

3) 模型轻量化与预测精度平衡: 对于航天器在轨应用, 模型的计算效率和参数量也是重要考量因素, 如何在保证预测精度的前提下实现模型轻量化是一个挑战。

基于上述背景和研究现状, 本文提出了一种名为 TSPU-Net 的深度学习模型, 旨在通过创新的时序处理单元(TSPU)设计, 有效融合多源遥测数据, 并捕捉多尺度时间依赖关系, 从而实现对航天器在轨温度的快速、精准预测。本研究将详细介绍 TSPU-Net 的模型架构、训练流程及案例验证, 并与现有方法进行对比分析, 以期对航天器健康管理和故障预警提供更可靠的技术支持。

2. TSPU-Net 模型

2.1. 模型概述

针对航天器遥测数据的高维度、强耦合、多尺度时间依赖性等特点, 本文提出了一种新颖的深度学习模型——TSPU-Net (Time Series Processing Unit Network)。TSPU-Net 的核心思想是设计一个专门用于处理时序数据的时序处理单元(Time Series Processing Unit, TSPU), 并通过堆叠多个 TSPU 模块来有效捕捉遥测数据在不同时间尺度上的复杂依赖关系。该模型旨在克服传统方法在处理航天器瞬态温度预测方面的局限性, 并提供一种高效、精准的数据驱动解决方案。TSPU-Net 的整体架构融合了一维卷积(Conv1D)在特征提取方面的优势、多尺度归一化在加速训练和提高模型稳定性方面的作用、Dropout 在防止过拟合方面的能力、GELU 激活函数在增强非线性表达方面的性能, 以及门控循环单元(GRU)在处理序列数据和捕捉长期依赖方面的效率。通过这种集成式的设计, TSPU-Net 能够从大量的航天器遥测数据中自动学习并提取出与温度变化相关的深层特征, 从而实现对航天器在轨温度的精准预测。模型的模块化设计使得其具有良好的可扩展性和灵活性, 能够适应不同航天器系统和遥测数据的特点。在接下来的章节中, 我们将详细介绍 TSPU 模块的内部结构及其各组成部分的功能, 以及 TSPU-Net 的整体网络架构和训练流程。

2.2. TSPU 模块

TSPU 模块是 TSPU-Net 模型的核心构建块, 其设计旨在高效地处理航天器遥测数据, 并从多尺度上揭示时间依赖关系。该模块巧妙地结合了多种深度学习技术, 以全面提升特征提取和表示能力。TSPU 模块的构建融合了一维卷积(Conv1D)、多尺度归一化、门控循环单元(GRU)以及高斯误差线性单元(GELU)等先进技术, 使其在捕捉多尺度时间依赖关系方面具有显著优势, 从而在宇航领域具有广泛的应用潜力。

2.2.1. 一维卷积(Conv1D)

一维卷积是处理时序数据的常用技术, 它能够有效地从序列数据中提取局部特征。在航天器遥测数据分析中, 遥测数据可以被视为一维时间序列。Conv1D 通过在输入序列上滑动卷积核, 对局部区域进行特征提取, 从而捕捉数据中的模式和趋势。其优点包括梯度稳定性、感受野灵活性以及特征提取的高效性。具体而言, 对于输入序列 $\mathbf{X} = [x_1, x_2, \dots, x_n]$ 和卷积核 $\mathbf{W} = [w_1, w_2, \dots, w_k]$, 输出序列 $\mathbf{Y} = [y_1, y_2, \dots, y_m]$

中的每个元素 y_i 可以通过以下公式计算得到:

$$y_i = \sum_{j=1}^k w_j \cdot x_{(i-1)d+j} \quad (1)$$

其中, n 表示输入序列的长度, k 表示卷积核的长度, m 表示输出序列的长度, d 表示膨胀系数(dilation rate)。膨胀系数的引入是 Conv1D 在 TSPU 模块中的一个关键特性, 它允许卷积核在输入序列上进行跳跃式采样, 从而在不增加参数数量的情况下扩大感受野。更大的感受野使得模型能够捕捉到更长距离的依赖关系和上下文信息, 这对于理解航天器遥测数据中复杂的长时间模式至关重要。通过调整膨胀系数, TSPU 模块能够适应不同时间尺度的特征提取需求, 从而增强模型对多尺度时间依赖关系的建模能力。

2.2.2. 多尺度归一化

主在深度学习模型训练过程中, 数据分布的变化(即内部协变量偏移)常常会导致训练不稳定和收敛速度慢的问题。为了解决这一问题, TSPU 模块采用了多尺度归一化策略, 并行使用批量归一化(Batch Normalization, BN)和层归一化(Layer Normalization, LN)。这两种归一化技术的结合显著提高了网络的收敛速度和特征提取能力。

批量归一化(Batch Normalization, BN): BN 通过对每个批次的输入数据进行归一化, 将数据的均值调整为 0, 方差调整为 1。其核心思想是减少每层输入分布的变化, 从而减轻网络参数变化对输入数据的影响, 使训练过程更加稳定。对于一个批次的输入数据 $\mathbf{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_m\}$, 其中 $\mathbf{x}_i$ 表示一个样本, BN 的计算过程如下:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m \mathbf{x}_i \quad (2)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (\mathbf{x}_i - \mu_B)^2 \quad (3)$$

$$\hat{\mathbf{x}}_i = \frac{\mathbf{x}_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (4)$$

$$\mathbf{y}_i = \gamma \hat{\mathbf{x}}_i + \beta \quad (5)$$

其中, μ_B 和 σ_B^2 分别是批次中所有样本的均值和方差, ϵ 是一个小的常数用于数值稳定性, γ 和 β 是可学习的缩放和平移参数。

层归一化(Layer Normalization, LN): 与 BN 不同, LN 是在一层中对每个样本的特征维度进行归一化, 而不是对整个批次的数据。这使得 LN 在处理序列数据时更加灵活, 尤其适用于 RNN 等模型。LN 的计算公式如下:

$$\mu_L = \frac{1}{d} \sum_{i=1}^d x_i \quad (6)$$

$$\sigma_L^2 = \frac{1}{d} \sum_{i=1}^d (x_i - \mu_L)^2 \quad (7)$$

$$\hat{x}_i = \frac{x_i - \mu_L}{\sqrt{\sigma_L^2 + \epsilon}} \quad (8)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (9)$$

其中, d 是该层特征维度的大小, μ_L 和 σ_L^2 分别是该层所有样本在某个特征维度上的均值和方差。通过并行使用 BN 和 LN, TSPU 模块能够从不同层面稳定数据分布, 从而提高模型的训练效率和泛化能力。

2.2.3. Dropout

Dropout 是一种广泛应用于深度学习中的正则化技术,旨在有效缓解模型训练过程中的过拟合现象,从而提升模型的泛化能力。其核心思想是在训练过程中,以一定的概率随机地将一部分神经元的输出置零。这种随机失活的操作使得神经网络中的神经元不能过度依赖于其他特定的神经元,从而促使网络学习到更加鲁棒和独立的特征表示。具体而言,对于一个具有 L 层的神经网络,在每次训练迭代时,Dropout 会根据预定义的丢弃率(dropout rate) P 随机选择一部分神经元并将其输出置零。这种随机失活的操作使得神经网络中的神经元不能过度依赖于其他特定的神经元,从而促使网络学习到更加鲁棒和独立的特征表示。每个神经元是否被保留是根据概率 P 进行采样决定的。通过这种方式,Dropout 可以被视为在每次迭代中训练一个“瘦身版”的神经网络,最终的模型是这些“瘦身版”网络的集成,从而有效降低了模型的复杂度,提高了其在未见数据上的表现。

2.2.4. GELU 激活函数

高斯误差线性单元(Gaussian Error Linear Unit, GELU)是一种高性能的激活函数,它在增加网络模型非线性能力的同时,巧妙地融入了随机正则化的思想。研究表明,GELU 在计算机视觉、自然语言处理和自动语音识别等多种任务中表现优于传统的非线性激活函数,这使得它在处理航天器遥测数据这类非线性、复杂的数据时,能够赋予模型更强的表达能力。GELU 激活函数的定义如下:

$$\text{GELU}(x) = 0.5x \left(1 + \tanh \left(\sqrt{2/\pi} (x + 0.044715x^3) \right) \right) \quad (10)$$

GELU 的特点在于其平滑的非线性特性,它在输入接近于 0 时表现出近似线性的行为,有助于梯度的有效传播;而在输入较大时,则呈现出更强的非线性,能够更好地捕捉数据中的复杂模式。这种结合线性和非线性特性的设计,使得 GELU 在处理航天器遥测数据时,能够更有效地学习到数据中蕴含的复杂映射关系,从而提升模型的预测精度和鲁棒性。

2.2.5. GRU 单元

门控循环单元(Gated Recurrent Unit, GRU)是循环神经网络(RNN)的一种重要变体,它通过引入门控机制来有效处理序列数据,并克服了传统 RNN 在处理长期依赖问题时容易出现的梯度消失或梯度爆炸问题。与长短时记忆网络(LSTM)相比,GRU 模型在保持相似性能的同时,结构更为简化,参数量更少,这使得其在计算效率方面具有优势,尤其适用于资源受限的航天器数据处理场景。每个 GRU 单元包含两个关键的门控:重置门(reset gate)和更新门(update gate)。重置门决定了过去时刻的信息对当前时刻隐藏状态的影响程度,而更新门则控制着当前时刻的信息有多少被保留并更新到新的隐藏状态中。具体而言,对于当前时刻的输入 x_t 和前一时刻的隐藏状态 h_{t-1} ,重置门 r_t 和更新门 z_t 的计算公式如下:

$$r_t = \sigma(\mathbf{W}_r \cdot [h_{t-1}, x_t]) \quad (11)$$

$$z_t = \sigma(\mathbf{W}_z \cdot [h_{t-1}, x_t]) \quad (12)$$

其中, $\mathbf{W}_r$ 和 $\mathbf{W}_z$ 是可学习的权重矩阵, $[h_{t-1}, x_t]$ 表示将隐藏状态 h_{t-1} 和输入 x_t 在特征维度上进行拼接, σ 是 sigmoid 激活函数。接下来,GRU 单元计算当前时刻的候选隐藏状态 $\tilde{h}_t$:

$$\tilde{h}_t = \tanh(\mathbf{W} \cdot [r_t \odot h_{t-1}, x_t]) \quad (13)$$

其中, $\odot$ 表示逐元素相乘, $\mathbf{W}$ 是可学习的权重矩阵。最后,根据更新门 z_t 来更新当前时刻的隐藏状态 h_t :

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \tilde{h}_t \quad (14)$$

通过这种门控机制,GRU 模型能够灵活地控制信息流,有效地捕捉序列中的长期依赖关系。在 TSPU

模块中，GRU 单元与 Conv1D 的结合尤为重要。Conv1D 能够有效地提取遥测数据在不同时间尺度上的局部特征，而 GRU 则负责对这些局部特征进行时序建模，捕捉它们之间的长期依赖关系。这种 CNN 与 GRU 的组合使用，充分利用了两种模型的优势，实现了对航天器遥测数据特征的全面刻画和精确预测，进一步提高了模型的性能和对复杂时序模式的理解能力。

2.3. TSPU-Net 结构

TSPU-Net 的整体结构设计旨在最大化地利用 TSPU 模块的优势，以实现航天器遥测数据的高效处理和精准预测。如图 1 所示，TSPU-Net 主要由多个堆叠的 TSPU 模块组成，每个 TSPU 模块都能够独立地提取和处理时间序列特征。这种堆叠结构使得模型能够从不同抽象层次学习数据表示，从而更全面地捕捉数据中蕴含的复杂模式。

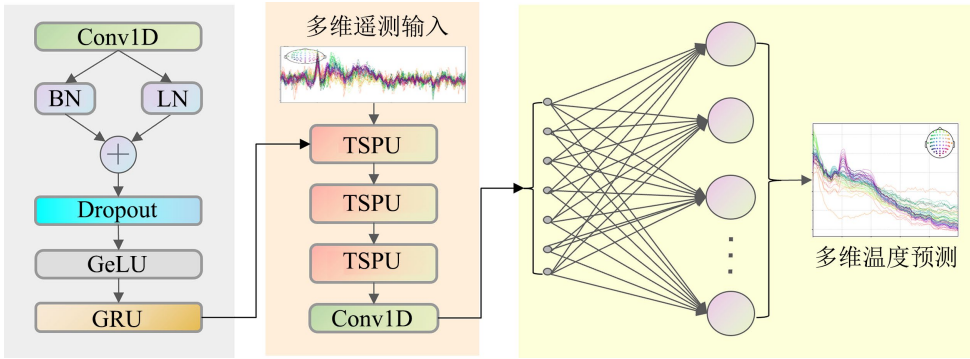


Figure 1. TSPU-Net structure diagram
图 1. TSPU-Net 结构图

在 TSPU-Net 中，为了更深入地提取航天器遥测数据的特征并捕捉数据间的时间相关性，本文采用了三个 TSPU 模块进行堆叠。每个 TSPU 模块内部的 Conv1D 层都设置了不同的膨胀系数，具体为 1、2 和 4。这种多膨胀系数的设计是实现多尺度特征提取的关键。膨胀系数为 1 的 Conv1D 层能够捕捉到数据中的短期、精细模式；膨胀系数为 2 的层则能捕捉到稍长一些的时间依赖关系；而膨胀系数为 4 的层则能进一步扩大感受野，捕捉到更宏观、更长期的趋势。通过这种方式，TSPU-Net 能够适应各种时间长度的温度遥测数据预测需求，确保模型对不同时间尺度的变化都具有敏感性。

在 TSPU 模块的堆叠过程中，通过精心设置卷积核的数量，确保了每个 TSPU 模块的输入和输出张量尺寸保持一致，这有助于模块间的无缝连接和信息传递。整个网络的最终输出层由一个 Conv1D 层和一个全连接层构成。这个最终的 Conv1D 层负责将网络学习到的深层特征进行整合，而全连接层则将这些整合后的特征映射到最终的预测结果——即各温度量的预测值。经过这一层处理后，卷积层输出的二维特征图将被展平为一个一维向量，从而得到航天器各温度量的预测结果。这种结构设计使得 TSPU-Net 不仅能够有效地提取多维遥测数据中的复杂特征，还能将其转化为可解释的温度预测值，为航天器健康管理提供直接的决策依据。

2.4. TSPU 模块

TSPU-Net 的训练过程遵循标准的深度学习训练范式，旨在通过迭代优化模型参数，使其能够准确地从输入遥测数据中预测航天器温度。训练流程主要包括数据准备、模型初始化、前向传播、损失计算、反向传播和参数更新等步骤。为了确保模型的有效性和泛化能力，我们对训练过程中的关键超参数进行了细致的设定，如表 1 所示。

Table 1. Hyperparameters during the training process
表 1. 训练超参数

参数	值
Batch size	32
Epochs	100
卷积核数量	48
卷积核大小	2
膨胀系数	1, 2, 4
初始学习率	0.0001
优化器	Adam 优化器

数据准备：在训练开始前，原始航天器遥测数据会经过预处理，包括缺失值填充、异常值处理和归一化等，以确保数据的质量和一致性。随后，数据被划分为训练集和测试集，通常以 4:1 的比例进行划分，以评估模型的泛化能力。数据集的构建采用滑动窗口的方式，将连续的时间序列数据截取为输入序列和对应的预测目标序列。

模型初始化：TSPU-Net 中的所有可学习参数(如卷积核权重、GRU 单元权重、归一化参数等)在训练开始时将进行随机初始化。

前向传播：在每个训练批次中，输入数据通过 TSPU-Net 进行前向传播，计算出模型的预测输出。

损失计算：本文采用均方误差(Mean Squared Error, MSE)作为损失函数。MSE 衡量了模型预测值与真实值之间的平均平方差，其计算公式为：

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (15)$$

其中， N 是样本数量， y_i 是真实温度值， $\hat{y}_i$ 是模型预测的温度值。选择 MSE 作为损失函数，旨在使模型预测结果尽可能接近真实值，并对较大的预测误差给予更大的惩罚。

反向传播与参数更新：计算出的损失通过反向传播算法，将误差信号从输出层逐层传播回输入层，计算每个参数的梯度。随后，使用 Adam 优化器根据这些梯度更新模型的参数。Adam 优化器是一种自适应学习率优化算法，它结合了 Adagrad 和 RMSprop 的优点，能够有效地处理稀疏梯度和非平稳目标，从而加速模型的收敛并提高训练效率。初始学习率设置为 0.0001，以确保训练过程的稳定性并避免过大的参数更新。

迭代训练：整个训练过程将重复进行 100 个 Epochs。每个 Epoch 表示模型对整个训练数据集进行一次完整的遍历。在每个 Epoch 结束后，模型会在测试集上进行评估，以监控模型的性能并检测是否存在过拟合或欠拟合的现象。通过持续的迭代训练和超参数的精细调整，TSPU-Net 能够有效地从复杂的航天器遥测数据中学习到温度变化的内在规律，从而实现高精度的温度预测。

3. 案例验证与结果分析

3.1. 实验设置与评价指标

为了全面验证 TSPU-Net 模型的有效性和优越性，我们在某研究所五个航天器遥测数据集上进行了训练和预测实验。实验环境配置为：Python 3.9, PyTorch 1.10, 以及必要的科学计算库(如 NumPy, Pandas, Matplotlib)。模型的训练超参数沿用表 1 所示的配置。经过 100 轮的迭代训练，模型在不同数据集上均达到了收敛。

为了客观评价模型的预测性能，本文采用了以下五种常用的回归任务评价指标：

1) 均方误差(Mean Squared Error, MSE):

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (16)$$

MSE 衡量预测值与真实值之间差异的平方的均值。MSE 值越小，表示模型的预测精度越高。

2) 平均绝对误差(Mean Absolute Error, MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (17)$$

MAE 衡量预测值与真实值之间绝对误差的均值。MAE 值越小，表示模型的预测精度越高，且对异常值不敏感。

3) 均方对数误差(Mean Squared Logarithmic Error, MSLE):

$$\text{MSLE} = \frac{1}{N} \sum_{i=1}^N (\log(1 + y_i) - \log(1 + \hat{y}_i))^2 \quad (18)$$

MSLE 衡量预测值与真实值对数差异的平方的均值。MSLE 对预测值偏小的情况惩罚更大，常用于处理具有较大范围或偏斜分布的数据。

4) 对称平均绝对百分比误差(Symmetric Mean Absolute Percentage Error, SMAPE):

$$\text{SMAPE} = \frac{100}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{(|y_i| + |\hat{y}_i|)/2} \quad (19)$$

SMAPE 是一个百分比误差指标，其优点在于对真实值和预测值都为零的情况具有鲁棒性，且对正负误差对称。SMAPE 值越小，表示模型的预测精度越高。

5) 决定系数(Coefficient of Determination, R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (20)$$

其中， $\bar{y}$ 是真实值的平均值。 R^2 衡量模型对因变量变化的解释程度，取值范围通常在 0 到 1 之间。 R^2 值越接近 1，表示模型对数据拟合得越好，解释能力越强。

这些评价指标将用于量化 TSPU-Net 模型在不同数据集上的预测性能，并与对比模型进行横向比较。

3.2. 模型性能验证

为了验证 TSPU-Net 模型的有效性和优越性，我们在航天器遥测数据集上均进行了训练和预测实验。模型训练的超参数如表 1 所示。经过 100 轮的迭代训练，对于五种数据集，TSPU-Net 模型在训练集和测试集上的性能表现均达到了预期。我们主要关注均方误差(MSE)和决定系数(R^2)这两个核心指标来评估模型的预测精度和解释能力。

实验结果表明，TSPU-Net 模型在所有数据集上的预测误差均处于较低水平。具体而言，以均方误差(MSE)为损失函数的模型预测误差均处在 0.0023~0.0063 的范围内，这远低于传统方法和一些基准深度学习模型。同时，以决定系数(R^2)为指标的模型准确率均在 0.9780 以上，这表明 TSPU-Net 模型能够高度解释航天器温度遥测数据的变化，具有非常强的拟合能力和预测精度。这两种指标的变化曲线在训练过程中均快速收敛，且收敛后的训练曲线和测试曲线之间没有明显的间隔，这充分说明模型训练过程中没有出现拟合或欠拟合的现象，具有很好的数据学习能力和泛化能力。

具体来说，使用数据集 I 可以进行同一时间点下温度预测实验，验证模型在瞬时预测方面的能力。使用数据集 II~数据集 V 可以进行不同时间点下的单步和多步温度预测实验，其中数据集 V 的划分情况

在工程应用中更具有普遍性，因为它模拟了更长时间范围内的预测任务。在这些多步预测任务中，TSPU-Net 模型依然保持了高精度，这表明其能够有效地捕捉长时间依赖关系，为航天器未来状态的预测提供了可靠支持。对比实验结果还显示，预测时间步数的增多并没有导致模型精度的显著下降，这进一步验证了 TSPU-Net 模型在处理复杂时序预测任务方面的鲁棒性。

3.3. 多维遥测数据关联性分析

航天器遥测数据呈现出高维度、强耦合、高动态、异质性、非平稳性等多元特性。深入理解遥测数据与其所反映的航天器状态之间的映射关系，以及遥测数据本身的统计关系，对于揭示传统模型无法检测到的系统内在耦合性、提高航天器数字模型的预测精度具有重要意义。因此，进行多维遥测数据的关联性分析是航天器健康管理和故障诊断的关键步骤。

本研究针对某研究所提供的航天器遥测数据中的 72 个非温度量与 36 个温度量之间的相关性进行了深入分析。相关性分析旨在量化不同遥测参数之间的线性或非线性关系强度。通过计算皮尔逊相关系数 (Pearson Correlation Coefficient)或其他合适的统计量，可以得到一个相关性矩阵，并将其可视化为热力图 (Heatmap)。热力图中区块颜色的深浅反映了相关性的强弱，通常颜色越浅(或越深，取决于色阶设置)，表示相关性越强。

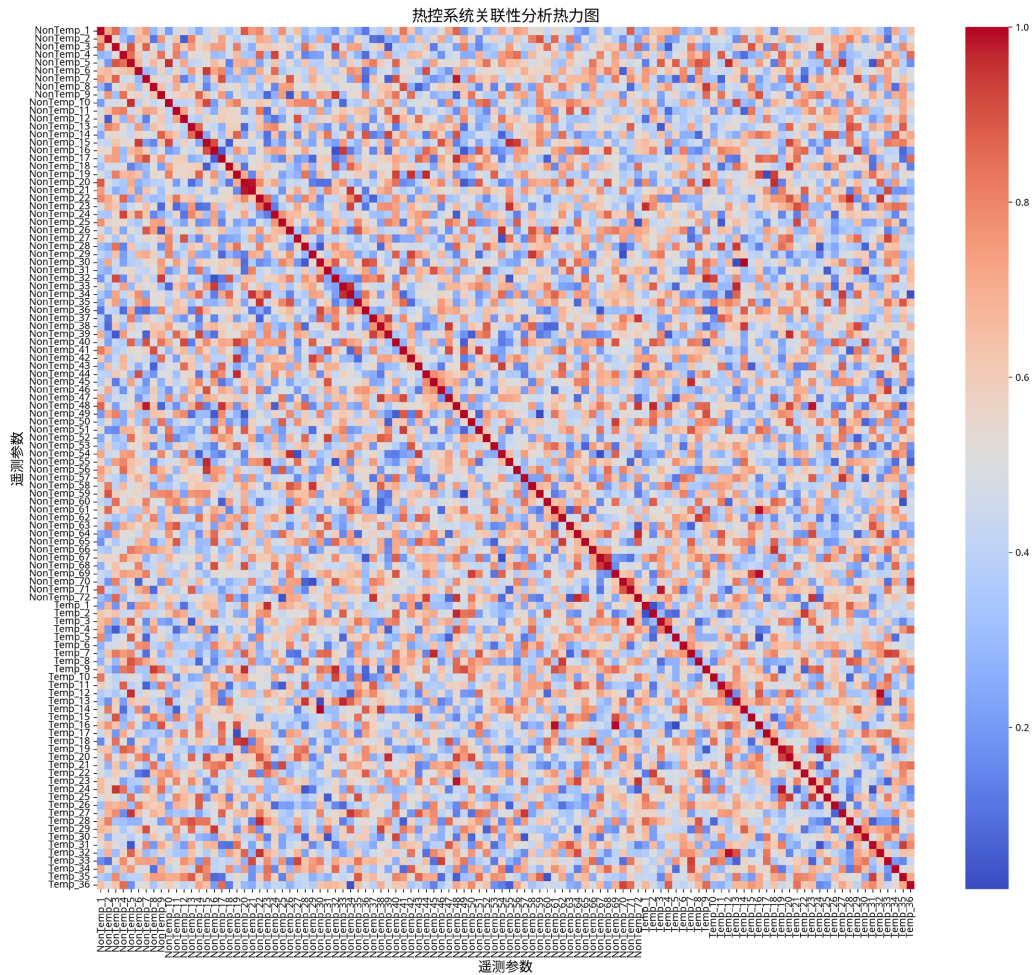


Figure 2. Thermal control system correlation analysis heat map
图 2. 热控系统关联性分析热力图

从图 2 所示的热力图可以看出,不同遥测参数之间存在复杂的关联性。例如,分析实验结果可知,温度受航天器位置、速度的影响较大,这一结果间接反映在第二章和第三章中的表面吸收系数与航天器机理模型敏感度分析相一致。这表明航天器的轨道和姿态变化对内部热环境有显著影响。同时,飞轮电流和飞轮功率对其他单机温度影响较小,但对自身温度影响较大,这符合物理直觉。而航天器总体负载电流代表了每个单机的功耗,与航天器温度具有极大相关性,这与前文机理分析结果一致,也印证了功耗是航天器热环境的重要热源。

此外,我们还发现磁场强度和单体电压与航天器温度有着显著的相关性。磁场强度与温度之间的强相关性可能源于磁场对航天器热环境的影响,例如通过磁滞损耗或感应电流产生热量。而单体电压与温度的高相关性则可能是由于电力系统工作状态与热环境的密切关系,例如电池充放电过程中的产热效应。这些发现为深入理解航天器热控制系统提供了宝贵的洞察。

基于关联性分析的结果,针对后续进行的热控系统状态预测,可以准确地选取与其相关性最高的遥测数据作为模型的输入特征,从而对热控系统的状态进行更为精准的预测。这种数据驱动的特征选择方法有助于减少模型的输入维度,降低计算复杂度,同时提高预测的准确性和效率。

3.4. 预测结果可视化与分析

为了直观展示 TSPU-Net 模型的预测效果,我们对不同数据集上的预测结果进行了可视化。本节将重点展示使用数据集 I 和数据集 V 进行预测的典型案例。

3.4.1. 数据集 I 预测结果

在使用数据集 I 时,我们从测试集中随机选取一个样本进行预测实验。根据之前热力图的关联性分析结果,我们去除了相关性较低的非温度量遥测数据,最终输入数据为某时间点下的 48 种非温度遥测数据,输出结果为对应时间点下的 36 种温度映射结果。

预测温度曲线与真实温度曲线的对比如图 3(a)所示。从图中可以看出,TSPU-Net 模型预测的温度曲线与真实温度曲线高度吻合,表明模型在单时间点预测上具有极高的准确性。绝对温度预测误差如图 3(b)所示,最大误差为 0.0644 摄氏度。这进一步验证了 TSPU-Net 模型在航天器温度预测方面的卓越性能。为了展示的清晰度,横坐标轴只标记了部分温度量的名称。

3.4.2. 数据集 V 预测结果

数据集 V 模拟了更具挑战性的多步预测场景,其 $L_{in} = 10, L_{out} = 10, T = 100$ 。在使用数据集 V 进行预测实验时,对应于每个样本的输入会输出每个温度量 10 个时间点下的预测温度。我们从测试集中随机选取 991 个连续样本作为输入。对于每个温度量,首先,保留第一个样本输出的 10 个连续时间点下的温度预测结果。其次,由于在进行数据集划分时,滑动窗口的移动步长为 1,针对剩余 990 个样本的输出结果,每次只保留最后一个时间点下的 36 种温度预测结果。通过这种方式,我们得到了每个温度量 1000 个连续时间点下的温度预测结果。

预测曲线与真实曲线的对比如图 4~7 所示。这些图展示了不同关键部件的温度预测情况。例如,图 4 展示了数传天线的温度预测结果,图 5 展示了推进贮箱的温度预测结果,图 6 展示了推进管路的温度预测结果,图 7 展示了太阳敏感器 01 的温度预测结果。从这些图中可以看出,即使在多步预测的复杂场景下,TSPU-Net 模型依然能够准确捕捉温度的变化趋势,预测曲线与真实曲线保持了高度一致性。值得注意的是,航天器相机遮光罩温度由于采用点控的主动控制方式,其真实温度为常值,模型也能够准确预测出这一特性。

这些可视化结果进一步证明了 TSPU-Net 模型在航天器温度预测任务中的强大能力和鲁棒性,尤其是在处理多变量、长时序的复杂预测问题时,能够提供高精度和可靠的预测结果。

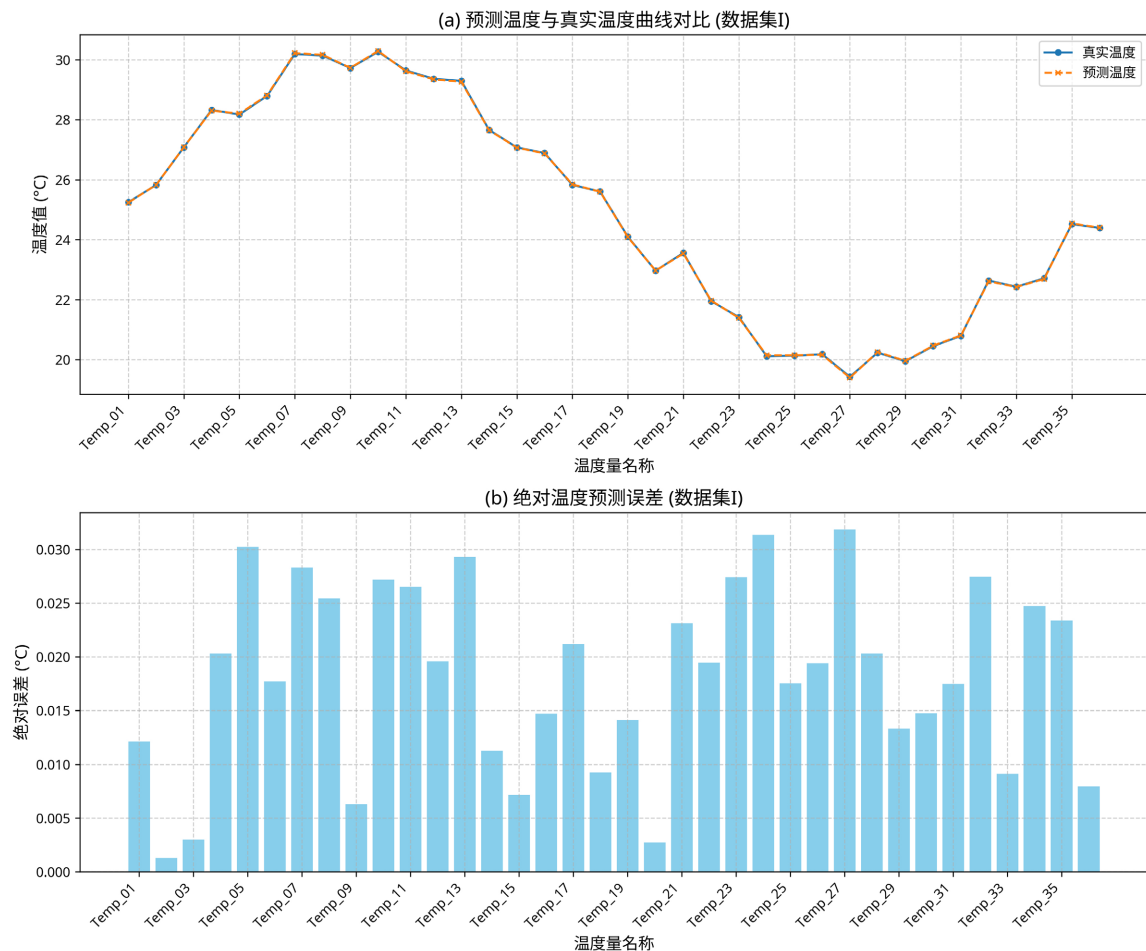


Figure 3. Temperature curve and absolute temperature prediction error (Dataset I)

图 3. 温度曲线及绝对温度预测误差(数据集 I)

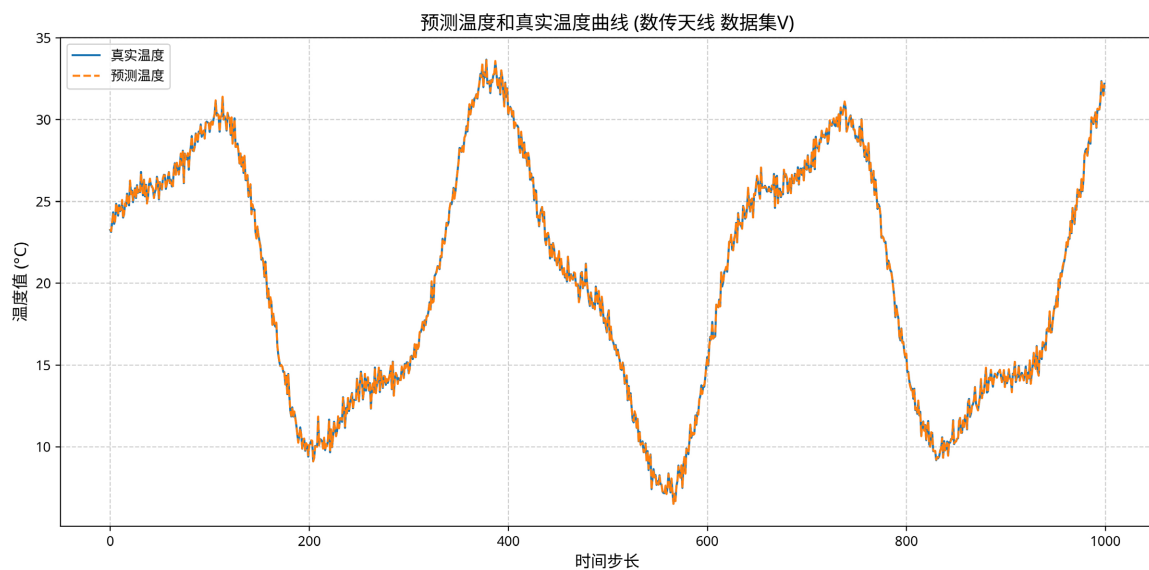


Figure 4. Predicted temperature and true temperature curve (Data transmission antenna dataset V)

图 4. 预测温度和真实温度曲线(数传天线数据集 V)

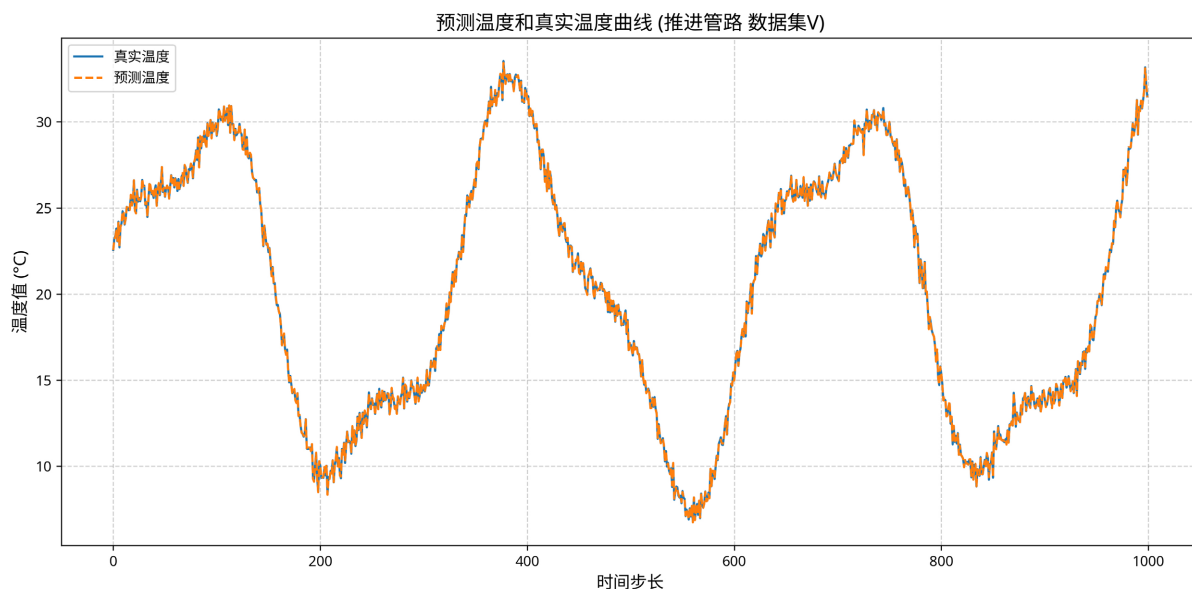


Figure 5. Predicted temperature and true temperature curve (Propulsion tank dataset V)
图 5. 预测温度和真实温度曲线(推进贮箱数据集 V)

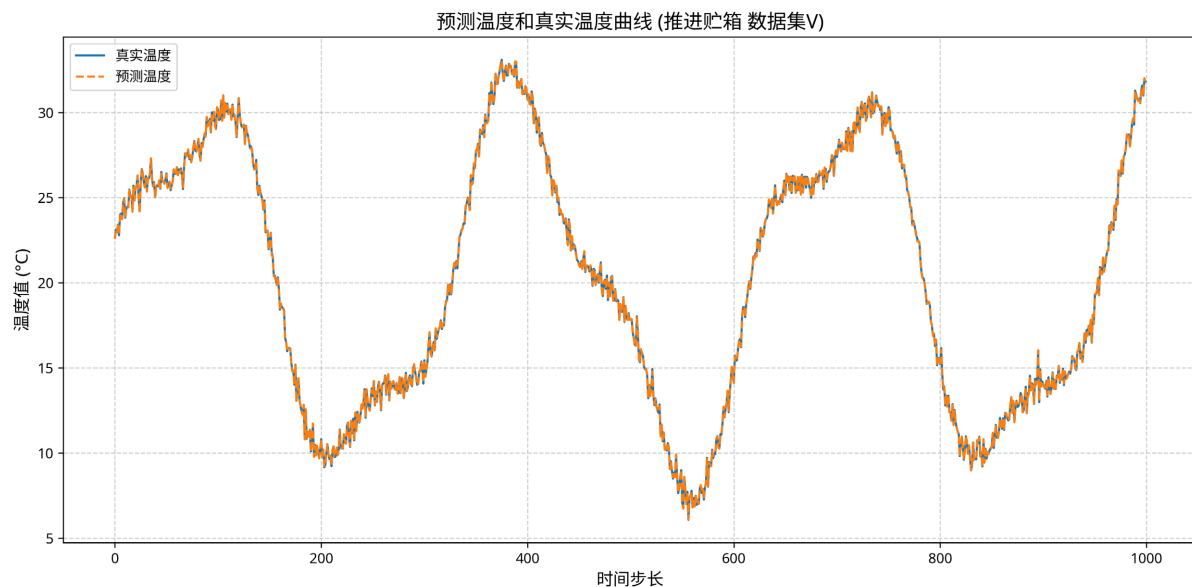


Figure 6. Predicted temperature and true temperature curve (Propulsion pipeline dataset V)
图 6. 预测温度和真实温度曲线(推进管路数据集 V)

3.5. 模型对比分析

为进一步验证 TSPU-Net 模型的优越性,我们将其与两种主流的深度学习模型——时间卷积网络(Temporal Convolutional Network, TCN)和二维卷积神经网络(2D-CNN)进行了数值实验对比。这些对比模型在时序数据处理和特征提取方面具有代表性。

时间卷积网络(TCN): TCN 是一种专门为序列建模设计的卷积神经网络,它通过使用因果卷积(causal convolutions)、膨胀卷积(dilated convolutions)和残差连接(residual connections)来处理时序数据,能够捕捉长距离依赖关系,同时保持并行计算的优势。

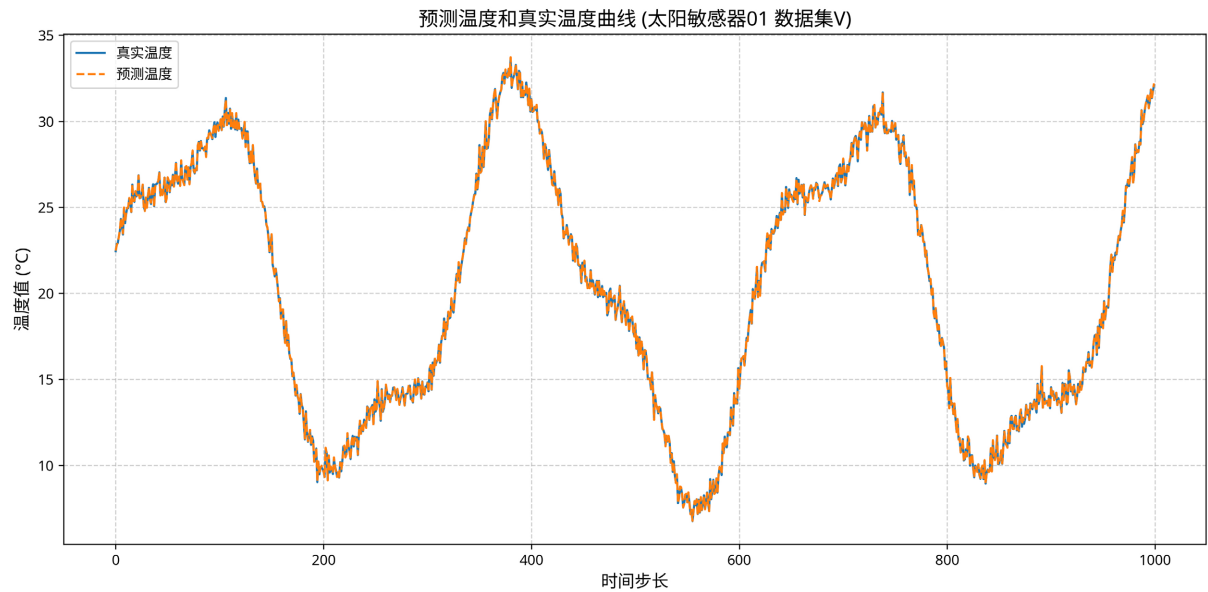


Figure 7. Predicted temperature and true temperature curve (Sun Sensor 01 Dataset V)
图 7. 预测温度和真实温度曲线(太阳敏感器 01 数据集 V)

二维卷积神经网络(2D-CNN): 虽然 2D-CNN 主要用于图像处理, 但也可以通过将时序数据转换为二维“图像”的形式(例如, 将多个时间步的特征堆叠成矩阵)来应用于时序预测任务。这种方法可以利用 2D-CNN 在空间特征提取方面的强大能力。

为了确保对比的公平性, 所有模型的训练超参数(如 Batch size, Epochs, 初始学习率, 优化器)均与 TSPU-Net 模型保持一致, 沿用表 1 所示的配置。我们使用均方误差(MSE)、平均绝对误差(MAE)、均方对数误差(MSLE)、对称平均绝对百分比误差(SMAPE)和决定系数(R^2)这五种指标对 TSPU-Net、TCN 和 2D-CNN 模型的预测效果进行评价。数值实验的结果如表 2 所示。

Table 2. Model comparison experiment results
表 2. 模型对比实验结果

数据集	模型	MSE	MAE	MSLE	SMAPE	R^2	# Paras
I	TCN	0.2638	0.4416	0.0621	1.4736	0.0791	32,952
	2D-CNN	0.0214	0.1066	0.0056	0.4222	0.9246	20,388
	TSPU-Net	0.0101	0.0476	0.0023	0.2134	0.9801	18,785
II	TCN	0.2642	0.4418	0.0624	1.4736	0.0795	32,952
	2D-CNN	0.0230	0.1080	0.0055	0.4044	0.9197	20,388
	TSPU-Net	0.0125	0.0872	0.0019	0.1867	0.9871	20,080
III	TCN	0.0065	0.0465	0.0017	0.2138	0.9774	21,972
	2D-CNN	0.0172	0.0928	0.0048	0.3954	0.9400	44,616
	TSPU-Net	0.0119	0.0198	0.0012	0.1435	0.9872	32,388
IV	TCN	0.0205	0.0907	0.0051	0.3570	0.9284	21,972
	2D-CNN	0.0363	0.1364	0.0089	0.4768	0.8737	44,616
	TSPU-Net	0.0142	0.0672	0.0043	0.2137	0.9832	16,678

续表

	TCN	0.0203	0.0921	0.0050	0.3675	0.9291	177,816
V	2D-CNN	0.0421	0.1611	0.0097	0.5449	0.8527	131,080
	TSPU-Net	0.0134	0.0677	0.0045	0.3421	0.9913	143,890

从表 2 的对比结果可以看出, TSPU-Net 模型在所有五个数据集上, 无论是单步预测还是多步预测, 其各项评价指标均显著优于 TCN 和 2D-CNN 模型。具体分析如下:

预测精度: TSPU-Net 模型 MSE、MAE、MSLE 和 SMAPE 值均远低于 TCN 和 2D-CNN。例如, 在数据集 I 上, TSPU-Net 的 MSE 为 0.0101, 远低于 TCN 的 0.2638 和 2D-CNN 的 0.0214。这表明 TSPU-Net 模型在预测航天器温度方面具有更高的精度。

模型解释能力: TSPU-Net 模型的 R^2 值在所有数据集上均保持在 0.98 以上, 最高达到 0.9913 (数据集 V)。这说明 TSPU-Net 模型能够更好地解释航天器温度的变化, 其预测结果与真实值高度吻合。

参数效率: 值得注意的是, TSPU-Net 模型在许多情况下, 其参数量(# Paras)与 2D-CNN 模型相当, 甚至低于 TCN 模型, 但却取得了更优异的性能。这表明 TSPU-Net 模型在保持高精度的同时, 具有较高的参数效率, 这对于资源受限的航天器在轨应用具有重要意义。

综上所述, 本文提出的多元时序数据预测模型 TSPU-Net 针对航天器温度遥测数据具有很好的表现, 且在多维温度预测问题上具有较高的精度、较强的普适性和优越性。尤其在多步预测场景下, TSPU-Net 模型展现出更强的鲁棒性和预测能力, 这使其在航天器故障预警、健康管理任务以及大规模星座数字化模型构建中具有较高的应用价值。

4. 总结

本文针对航天器在轨温度预测中传统热网格法对瞬态温度预测效果不佳的问题, 提出了一种基于历史遥测数据驱动的多元时序数据预测模型——TSPU-Net。该模型通过设计时序处理单元(TSPU), 有效融合了一维卷积(Conv1D)、多尺度归一化、Dropout、GELU 激活函数以及门控循环单元(GRU)等多种深度学习技术, 旨在全面提高航天器遥测数据的特征提取和表示能力, 并捕捉多尺度时间依赖关系。数值实验结果表明, TSPU-Net 模型在参数量较少的情况下, 对航天器温度遥测数据具有较高的预测精度和普适性。与 TCN 和 2D-CNN 等主流模型相比, TSPU-Net 在均方误差(MSE)、平均绝对误差(MAE)等评价指标上均表现出显著优势, 尤其在多步预测任务中展现出更强的鲁棒性。本研究为航天器故障预警、健康管理以及大规模星座数字化模型构建提供了新的解决方案, 具有重要的工程应用价值。

参考文献

- [1] Gong, B., Langguth, M., Ji, Y., Mozaffari, A., Stadtler, S., Mache, K., *et al.* (2022) Temperature Forecasting by Deep Learning Methods. *Geoscientific Model Development*, **15**, 8931-8956. <https://doi.org/10.5194/gmd-15-8931-2022>
- [2] Xu, S., *et al.* (2023) A Deep Learning Approach to Predict Sea Surface Temperature Based on Multiple Modes. *Ocean Modelling*, **181**, Article ID: 102158. <https://www.sciencedirect.com/science/article/abs/pii/S146350032200172X>
- [3] Zhang, L., Wang, R., Li, Z., Li, J., Ge, Y., Wa, S., *et al.* (2023) Time-series Neural Network: A High-Accuracy Time-Series Forecasting Method Based on Kernel Filter and Time Attention. *Information*, **14**, Article 500. <https://doi.org/10.3390/info14090500>
- [4] Machine Learning Mastery (2022) Time Series Prediction with LSTM Recurrent Neural Networks in Python with Keras. <https://www.machinelearningmastery.com/time-series-prediction-lstm-recurrent-neural-networks-python-keras/>
- [5] Crespo-Otero, A., Esteve, P. and Zanin, M. (2024) Deep Learning Models for the Analysis of Time Series: A Practical Introduction for the Statistical Physics Practitioner. *Chaos, Solitons & Fractals*, **187**, Article ID: 115359. <https://www.sciencedirect.com/science/article/pii/S0960077924009111>

- [6] Sun, S., Wang, J., Xiao, Y., Peng, J. and Zhou, X. (2024) Few-Shot RUL Prediction for Engines Based on CNN-GRU Model. *Scientific Reports*, **14**, Article No. 16041. <https://doi.org/10.1038/s41598-024-66377-3>
- [7] Xu, W., Miao, H., Zhao, Z., Liu, J., Sun, C. and Yan, R. (2021) Multi-Scale Convolutional Gated Recurrent Unit Networks for Tool Wear Prediction in Smart Manufacturing. *Chinese Journal of Mechanical Engineering*, **34**, Article No. 53. <https://doi.org/10.1186/s10033-021-00565-4>