

# 基于软注意力GRU模型的堆芯瞬态热工水力参数预测方法研究

淳思琦<sup>1</sup> 冯欢<sup>3</sup> 张安妮<sup>4</sup> 赵鹏程<sup>1,2</sup>

1(南华大学核科学技术学院 衡阳 421001)

2(南华大学核燃料循环技术与装备湖南省协同创新中心 衡阳 421001)

3(南华大学资源环境与安全工程学院 衡阳 421001)

4(南华大学计算机/软件学院 衡阳 421001)

**摘要** 反应堆在各种工况下堆芯瞬态热工水力参数预测的准确性,直接影响到反应堆的安全性。质量流量和温度作为堆芯热工水力的重要参数,二者常被建模为时间序列预测问题。研究旨在解决瞬时条件下堆芯热工水力参数连续预测的精度问题,检验基于注意力机制的门控循环单元在核心参数预测中的可行性。本文采用1/2中国实验快堆(China Experimental Fast Reactor, CEFR)为研究对象,使用快堆子通道程序SUBCHANFLOW生成瞬态堆芯热工水力参数的时间序列,采用基于软注意力的门控循环单元(Gated Recurrent Unit, GRU)模型预测堆芯的质量流量和温度时间序列。结果表明:相较于自适应径向基(Radial Basis Function, RBF)神经网络,本文使用的软注意力的GRU网络模型预测结果更好,温度在步长为3的情况下平均相对误差不超过0.5%,在15 s内预测效果较好;质量流量在步长为10的情况下平均相对误差不超过5%,且在后续12 s内预测效果较好。本文构建的模型不仅在连续预测过程中表现出更高的预测精度,且能捕捉到动态时间序列中的趋势特点,这对维护反应堆安全,有效防止核电厂事故有极大的用处。基于软注意力的GRU模型能在瞬态反应堆工况下提供一段时间的连续预测,在工程应用中和提高反应堆安全性上具有一定的参考价值。

**关键词** 门控循环单元, 软注意力, 快堆, 瞬态热工水力, 参数预测

中图分类号 TL433

DOI: 10.11889/j.0253-3219.2024.hjs.47.010603

## Method of predicting transient thermal hydraulic parameters of the core based on the gated recurrent unit model of soft attention

CHUN Siqi<sup>1</sup> FENG Huan<sup>3</sup> ZHANG Anni<sup>4</sup> ZHAO Pengcheng<sup>1,2</sup>

1(School of Nuclear Science and Technology, University of South China, Hengyang 421001, China)

2(Cooperative Innovation Center for Nuclear Fuel Cycle Technology and Equipment, University of South China, Hengyang 421001, China)

3(School of Resources Environment and Safety Engineering, University of South China, Hengyang 421001, China)

4(School of Computer Science, University of South China, Hengyang 421001, China)

**Abstract** [Background] The accuracy of transient thermal hydraulic parameter prediction of reactor cores under various working conditions directly affects reactor safety. Mass flow rate and temperature are important parameters of

装备预研教育部联合基金(No.8091B032243)资助

第一作者: 淳思琦, 女, 2003年出生, 现南华大学本科在读, 研究领域为先进核能系统热工水力设计与安全分析

通信作者: 赵鹏程, E-mail: pengcheng.zhao@usc.edu.cn

收稿日期: 2023-05-29, 修回日期: 2023-09-01

Supported by Joint Fund of Ministry of Education for Equipment Pre-research (No.8091B032243)

First author: CHUN Siqi, female, born in 2003, undergraduate student, focusing on thermal-hydraulic design and safety analysis of advanced nuclear system

Corresponding author: ZHAO Pengcheng, E-mail: pengcheng.zhao@usc.edu.cn

Received date: 2023-05-29, revised date: 2023-09-01

core thermal hydraulics, which are often modeled as time-series prediction problems. **[Purpose]** This study aims to solve the accuracy problem of continuous prediction of core thermal hydraulic parameters under instantaneous conditions and to test the feasibility of a gated cycle unit based on the attention mechanism in core parameter prediction. **[Methods]** The 1/2 full core model of China Experimental Fast Reactor (CEFR) core was taken as the research object. The subchannel SUBCHANFLOW program was employed to generate the time series of transient core thermal hydraulic parameters. The gated recurrent unit (GRU) model based on soft attention was used to predict the mass flow and temperature time series of the core. **[Results]** The results show that, compared with the adaptive radial basis function (RBF) neural network, the GRU network model with soft attention offers better prediction results. The average relative error of temperature is  $<0.5\%$  when the step size is 3, and the prediction effect is quite good within 15 s. The average relative error of mass flow rate is  $<5\%$  when the step size is 10, and fairly good prediction effect is achieved in the subsequent 12 s. **[Conclusions]** The model constructed in this study not only exhibits higher prediction accuracy in the continuous prediction process but also captures the trend characteristics in the dynamic time series, which is of considerable value for maintaining reactor safety and effectively preventing nuclear power plant accidents. The GRU model based on soft attention can provide continuous prediction for a period of time under transient reactor conditions, providing a reference value in engineering applications and improving reactor safety.

**Key words** Gated recurrent unit, Soft attention, Fast reactor, Transient thermal hydraulics, Parameter prediction method

由于事故或者冷却剂呈两相状态等原因,堆芯瞬态工况会导致冷却剂流量等热工参数发生明显变化,从而影响其他热工参数的变化。在这些工况中,流动不稳定性是重要的安全考虑之一,其研究对于反应堆的安全运行至关重要。

现阶段,神经网络广泛应用在堆芯热工水力中<sup>[1]</sup>。例如,焦炜<sup>[2]</sup>使用遗传算法优化的反向传播(Back Propagation, BP)神经网络来预测沸腾传热系数;靳爽等<sup>[3]</sup>利用前馈神经网络(Feedforward Neural Network, FNN)方法的机器学习方法,对不同工况下粗网格下的计算流体动力学(Computational Fluid Dynamics, CFD)数值模拟结果进行了优化,结果表明,采用前馈神经网络优化粗网格得到的结果后,子通道温度的均方根误差(Root Mean Square Error, RMSE)平均值由1.16 K降低至 $9.24 \times 10^{-3}$  K,优化效果显著;李仕鲜等<sup>[4]</sup>使用反向传播BP神经网络、Elman神经网络、径向基函数(Radial Basis Function, RBF)神经网络和支持向量机对失水事故期间的6个参数进行了训练。结果表明,经过参数优化后,支持向量机对破口位置和大小诊断精度高于其他神经网络,同时表现出良好的诊断稳定性;陈涵瀛<sup>[5]</sup>基于循环神经网络(Recurrent Neural Network, RNN)和长短期记忆网络(Long Short-term Memory, STM)等深度学习算法,提出了一种预测剩余未停堆时间和重要工况参数的模型提高了预测的精度。

在本文的研究中,温度和质量流量都是时间序

列信号,RNN善于提取序列特征,LSTM<sup>[6]</sup>和门控循环单元网络(Gated Recurrent Unit, GRU)<sup>[7]</sup>是两种改进的变体,他们都使用门机制来缓解早期RNN中的梯度消失问题。LSTM单元有3个门,GRU单元有2个门。

本文选择GRU网络,因为它具有更简单的架构,更少的参数,性能与LSTM网络<sup>[8]</sup>相当。然而,目前的热工水力工况参数预测中,还停留在简单的神经网络使用和优化阶段,预测热工水力参数的变化趋势更为重要,因为它们直接关系到反应堆运行安全<sup>[9]</sup>,现阶段很少有结合其他先进神经网络提高对变化趋势的预测效果。

目前已有一些实例<sup>[10-11]</sup>证明了注意力机制可以使模型更加聚焦,提取信息更加准确,从而提高其可解释性,增强其鲁棒性,降低信息过载的影响。

本文首先采用快堆子通道程序SUBCHANFLOW计算产生瞬态的质量流量和温度随时间变化的数据,然后对得到的数据进行小波去噪处理,使用基于注意力机制的门控循环(Gated Recurrent Unit, GRU)神经网络模型进行训练,将测试集测试结果与自适应BP神经网络和自适应RBF神经网络进行对比,进而验证基于注意力机制的门控循环神经网络在对于堆芯关键参数进行连续预测优势。此外通过消融实验验证了去噪和注意力机制的有效性。

## 1 神经网络模型构建

### 1.1 自适应RBF神经网络

自适应RBF神经网络<sup>[12]</sup>是一种基于径向基函数的神经网络(图1),它具有自适应调整RBF网络参数的能力。RBF是一种常用的非线性函数,它可以将输入映射到高维空间中进行处理。

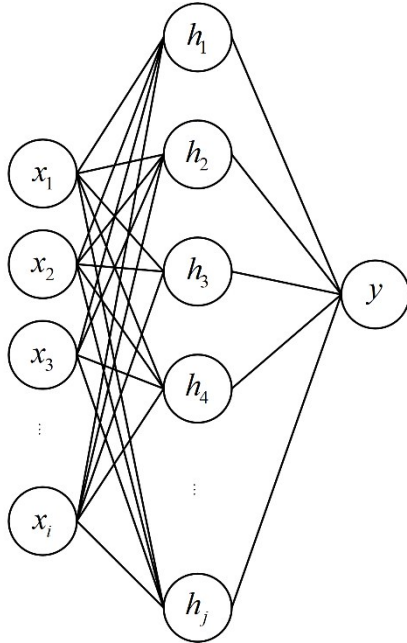


图1 自适应RBF神经网络结构  
Fig.1 Adaptive RBF neural network structure

自适应RBF神经网络首先需要确定RBF的数量和位置。通常使用聚类算法来确定RBF中心点的位置,如K-means算法<sup>[13]</sup>等。与传统的RBF神经网络相比,自适应RBF神经网络的主要优势在于其参数自适应能力,因为传统的RBF需要手动设置中心点和宽度等参数。此外,自适应RBF神经网络还具有可变的层数,因为每个样本可以添加一个新的RBF节点,从而使其更加灵活和适应性强。

设输入数据为 $x \in R$ ,隐含层由 $M$ 个径向基函数组成,输出层有 $K$ 个节点。则神经网络的输入、隐含层和输出分别可以表示为:

输入层:

$$x = [x_1, x_2, \dots, x_n] \quad (1)$$

隐含层:

$$h_j(x) = e^{-\beta \|x - c_j\|^2}, j = 1, 2, \dots, M \quad (2)$$

输出层:

$$y_k = \sum_{j=1}^M w_{kj} \phi_j(x), k = 1, 2, \dots, K \quad (3)$$

其中: $c_j$ 是第 $j$ 个径向基函数的中心; $\beta_j$ 是径向基函数宽度参数; $\phi_j$ 是径向基函数; $w_{kj}$ 是输出层中第 $k$ 个节

点和第 $j$ 个隐含单元之间的权重系数。RBF神经网络的训练过程包括确定 $c_j, \beta_j$ 和 $w_{kj}$ 等参数,最小化网络预测误差并提高泛化能力。对于分类任务,可以使用交叉熵损失函数;对于回归任务,可以使用均方误差损失函数。在确定了损失函数之后,可以使用梯度下降等优化算法进行训练。

自适应RBF是一种前馈神经网络,主要用于模式识别、数据挖掘和稳态时间序列预测等领域,由于该网络容易出现梯度消失等问题,且不能及时校正网络超参数,连续预测会造成误差堆积,从而影响后续的预测结果。

### 1.2 GRU神经网络

GRU是一种循环神经网络(Recurrent Neural Network, RNN)结构,用于学习和处理序列数据。GRU全称为门控循环单元(Gated Recurrent Unit),由Cho等<sup>[7]</sup>在2014年提出,相较于RBF神经网络,GRU在处理长期依赖问题上,能够更好地保留信息,从而更好地解决了梯度消失问题。与传统的RNN不同,GRU引入了重置门(reset gate)和更新门(update gate)机制,这两个门控制着网络中信息的流动。具体网络结构如图2所示。

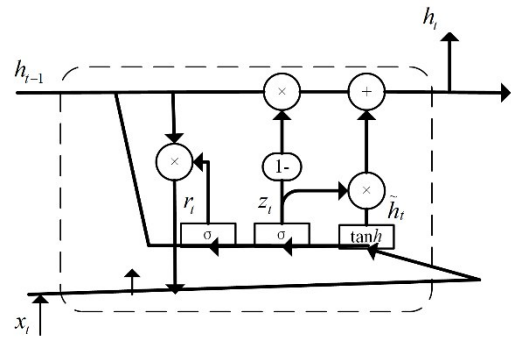


图2 GRU结构图  
Fig.2 GRU structure

具体来说,重置门控制着上一个时刻的隐藏状态 $h_{t-1}$ 如何影响当前时刻的输入 $x_t$ ,而更新门则决定了上一个时刻的隐藏状态 $h_{t-1}$ 多少被保留到当前时刻,并与当前时刻的输入一起更新隐藏状态。

$$r_1 = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (4)$$

$$r_2 = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (5)$$

$$\tilde{h}_t = \tanh(W \cdot [r_1 h_{t-1}, x_t]) \quad (6)$$

$$h = (1 - z_t) h_{t-1} + z_t \tilde{h}_t \quad (7)$$

式中: $W_r, W_z$ 和 $W$ 分别为重置门、更新门和隐藏层的权重矩阵; $\sigma$ 为激活函数,本次选择Sigmoid作为激活函数; $\tanh$ 即双曲正切激活函数;通过式(6)得到的 $\tilde{h}_t$ 为当前输入和先前隐藏层状态的总汇。为这些

门机制使得GRU能够更好地处理长期依赖关系,并且能够避免梯度消失或爆炸问题。

相比于其他RNN结构,GRU拥有更少的参数,训练速度更快,同时某些任务上也能够取得更好的性能。因此,GRU已经成为了当今深度学习领域中非常受欢迎的模型之一,广泛应用于自然语言处理<sup>[14]</sup>、语音识别<sup>[15]</sup>等任务中。不过,对于一些不涉及时间序列的问题,使用GRU可能会有些过于复杂。

### 1.3 注意力机制

实际情况下,流体在管道或容器中容易出现波动、涡旋、振荡等不稳定现象,这种现象可能会导致堆芯温度分布不均,从而导致堆芯热工水力参数的变化,例如流速、温度、压力等。在对堆芯热工水力参数的预测中,除了要求对时间序列预测误差要小,还需要重视质量流量时间序列和温度时间序列发生突变的部分<sup>[16]</sup>。借鉴人类视觉的特点,注意力机制可以选择性地将有限的信息处理资源分配给重要的信息序列段,从而实现对目标区域的专注和集中。故在网络结构中引入注意力机制,进一步提升模型的性能。目前,注意力机制广泛应用于自然语言处理、图像处理和推荐系统等各个领域。

注意力机制<sup>[17-18]</sup>是一种资源分配方案,用于帮助神经网络解决信息超载问题。通过对输入数据中的关键部分进行加权处理,注意力机制可以使模型更加精确地聚焦和提取信息,从而改善模型解释性、增强鲁棒性,并减少信息超载的影响。注意力分硬注意、软注意力和自注意力<sup>[19]</sup>,其中,硬注意力和软注意力主要的区别在于是否明确选择了关注的位置:硬注意力明确选择了一个或几个位置,而软注意力给所有位置赋予了权重。自注意力是一种特殊类型的软注意力机制,它在处理序列数据(文本)时,允许模型在每个步骤中都关注输入的所有部分。软注意力可以对输入序列中的每个位置进行加权处理,且可以计算每个位置与目标输出之间的关联度作为权重,是一种灵活,具有可解释性的注意力机制。

因此,本文引入软注意力对GRU输出的数据进行加权。本文所用软注意力如图3所示,其中, $x_1, x_2, \dots, x_t$ 为输入GRU网络的序列数据; $h_1, h_2, \dots, h_t$ 为网络对应的隐藏层状态值; $a_i$ 为对应的注意力向量; $c_i$ 为上下文向量; $\tilde{h}_t$ 为隐藏层最终状态; $y_t$ 为GRU网络的最终输出。

$$\text{score}_{ij} = h_i w_j + b_j \quad (8)$$

$$\alpha_{ij} = \frac{\exp(\text{score}_{ij})}{\sum_{i=1}^k \exp(\text{score}_{ij}) + \epsilon} \quad (9)$$

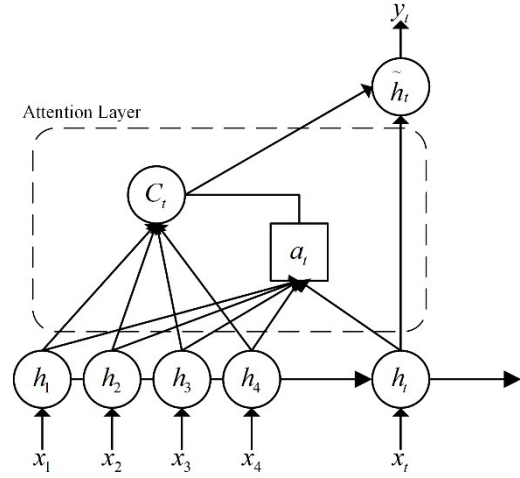


图3 软注意力示意图  
Fig.3 Soft attention diagram

$$c_i = \sum_j \alpha_{ij} \cdot h_j \quad (10)$$

注意力机制的计算如下:通过计算输入张量与学习到的权重矩阵的点积,并添加偏置项,然后通过非线性变换( $\tanh$ 函数)将结果映射到对应的得分上。接着,它通过指数函数将得分进行正规化,最后将得到的注意力权重与输入张量相乘,最终输出加权平均值。

其中,分母对所有得分进行求和并加上一个小的 $\epsilon$ 值,以避免除以零错误。这个过程可以看作是一个softmax运算的变体,将得分转换为概率分布。式中: $h_i (i \in [1, t])$ 为第*i*时刻的隐藏层变量; $w_j$ 为在注意力机制中第*j*个位置的权重向量; $b_j$ 是偏置向量; $\text{score}_{ij}$ 为注意力得分; $\alpha_{ij}$ 为注意力权重。

### 1.4 基于注意力机制的GRU预测模型

本次选用Attention+GRU进行时间序列预测,GRU神经网络是一种常用于时间序列预测的循环神经网络模型,它可以通过对时序数据进行建模来捕捉时间序列中的长期依赖关系。然而,在实际应用中,不同时间步的输入数据往往存在着不同的重要性和影响程度,因此需要在GRU模型中引入注意力机制来对每个时间步的输入进行加权,以更好地捕捉时序数据特征。

首先将选定步长的数据组输入一个节点数为64的GRU网络中,本网络使用了更新门和重置门来控制信息的流动,能够有效地解决传统RNN在处理长序列时会遇到的梯度消失或梯度爆炸的问题;然后将结果进行一次注意力操作,模型计算得出一个权重分布,这个权重分布决定了模型对于每个时间步的输出应该给予多大的关注;之后通过Flatten层将序列张量按照顺序展平成一个一维向量,最后通

过一个全连接层,计算得到最终的输出结果。网络层如图4所示。

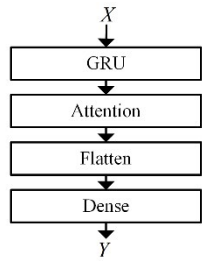


图4 基于注意力机制的GRU预测拓扑结构  
Fig.4 GRU prediction topological structure based on the attention mechanism

## 2 仿真实验

### 2.1 数据预处理

#### 2.1.1 数据来源

为对比自适应RBF神经网络在堆芯热工水力参数预测的效果,采用与冀南等在2022年发表文章<sup>[9]</sup>中同样的数据样本获取方法,即使用子通道程序SUBCHANFLOW构建中国实验快堆(China

Experimental Fast Reactor, CEFR)堆芯的1/2全堆芯模型,用同样的方式对堆芯子通道进行简化,得到质量流量时间序列和温度时间序列。

#### 2.1.2 DB8小波去噪处理

考虑到实际工作中由于各种不稳定性因素的影响,热工水力参数常常会受到噪声的干扰,主要噪声包括高斯白噪声等。因此,需要采取相应的措施来降低噪声对预测结果的影响。

本次选用DB8小波对时间序列进行去噪处理。Daubechies 8小波<sup>[20]</sup>,也称为DB8小波,是一种基于小波分析的信号处理技术中常用的小波函数,在进行信号降噪时,DB8小波能够更好地捕捉到它们的高频部分,同时保留低频部分的重要信息,通常适用于具有渐变或者突变特性的信号。同时,DB8小波也提供了较高的信号压缩效率,可以在保证去噪效果的同时减少数据量。此外,DB8小波可以有效去除高斯白噪声信号。因为它可以提供相对尖锐的细节系数,这有助于更好地捕捉信号中的细节和噪声。

质量流量时间序列与温度时间序列去噪前后对比如图5所示。

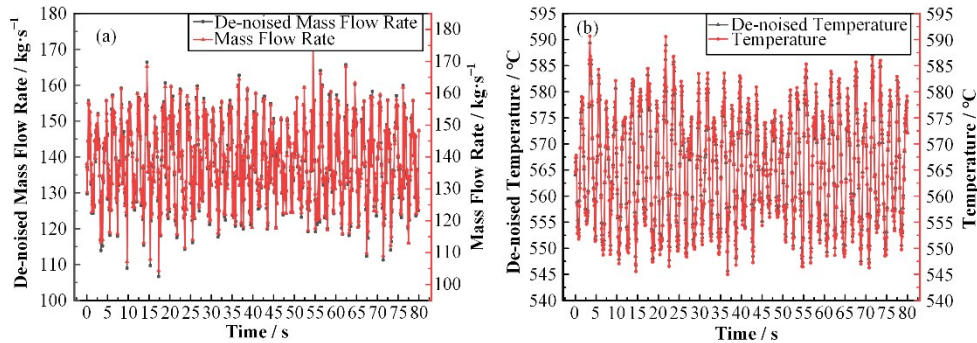


图5 时间序列的去噪前后对比 (a) 未去噪和去噪后的质量流量时间序列,(b) 未去噪和去噪后的温度时间序列  
Fig.5 Comparison of time series before and after denoising (a) Un-denoised and denoised mass flow time series, (b) Un-denoised and denoised temperature time series

### 2.2 步长超参数确定

在使用GRU模型时,步长的选择对结果影响是非常重要的。步长(stride)指的是卷积核或池化窗口在每一次滑动时所跨越的像素数目。如果步长太小,模型会计算更多的输出,并且需要更长的时间来训练和处理数据;如果步长太大,则可能会导致信息损失和降采样,从而影响模型的准确性<sup>[21]</sup>。因此我们通过调整网络步长来找到最佳步长,并据此提高预测准确度。具体操作如下:首先,从1开始逐步增大输入步长,在每个步长下进行50次实验,计算测试集的平均相对误差。然后,绘制不同步长与平均相对误差的折线图,误差越小说明此步长越合理,也就是网络预测效果越佳,如图6所示。

从调试结果可知,质量流量时间序列在步长为10的情况下,在本网络中的效果最佳;温度时间序列在步长为3的情况下,在本网络中的效果最佳。

### 2.3 模型验证

本次实验的流量和温度数据均为800条,直接按照80%、20%的比例划分训练集和测试集。根据§2.2中选定的步长进行相关实验,质量流量时间序列在步长为10的情况下,即每十个连续值预测当前时间的质量流量,总体特征从800变成790×10,故训练集为630,测试集为150。同理,温度时间序列预测网络步长设置为3,总体特征从800变成797×3,故训练集为637,测试集为157。

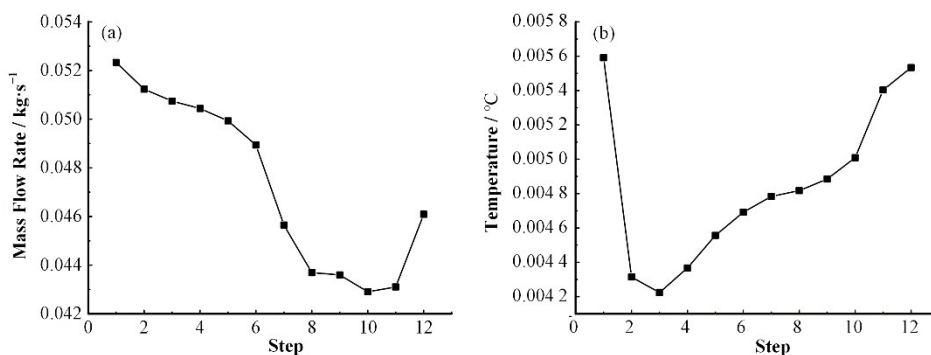


图6 步长超参数与平均相对误差关系 (a) 质量流量时间序列, (b) 温度时间序列  
Fig.6 Relationship between the super parameter of step length and average relative error  
(a) Mass flow time series, (b) Temperature time series

### 2.3.1 预测结果分析

根据上面的步长选择,使用注意力+GRU网络进行预测,预测结果如图7所示。由图7可知,对于堆芯入口流量的预测,由于其本身存在较多白噪声,因此预测效果不如燃料包壳最高温度预测的效果好。在本次实验中,选择了步长为10进行预测。然而,步长过大会使模型对历史数据的抽象程度过高,从而受到未来数据的影响,导致预测中期可能存在

轻微延时的问题。虽然本次实验成功地学习到了数据的变化趋势,并能够很好地捕捉到数据的变化趋势,但是在准确获得当前数据的情况方面仍有欠缺。

对于燃料包壳最高温度而言,本次选择的步长是3,步长选择适中,预测数据很好地学习了当前时刻以及未来的数据变化情况,以及未来的变化情况,从图8可以看出,预测效果与实际值基本上完美拟合,预测效果极佳。

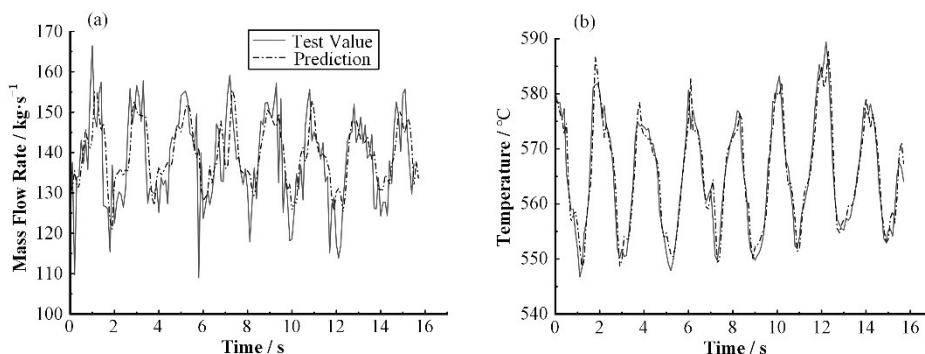


图7 时间序列预测结果 (a) 质量流量时间序列, (b) 温度时间序列  
Fig.7 Prediction results (a) Mass flow time series, (b) Temperature time series

针对堆芯入口流量预测轻微延时情况,我们在实验过程中绘制了步长较小情况下网络的预测图,如图8所示,发现步长选择为3,预测趋势虽然延迟不明显,但结果存在较大偏差,综合考虑预测趋势和预测结果偏差两个因素,我们最终以偏差为导向,选择步长10输入网络进行训练。

将最终预测的结果与冀南等使用BP神经网络和自适应RBF神经网络的预测结果进行对比,考虑到瞬时工况中,进行点对点单步预测容易受到噪声和异常值的影响,导致预测精度降低。此外,单步预测通常只考虑当前时刻的数据,无法充分利用历史数据的信息,学习未来趋势。本次重点与连续预测的结果进行对比,如表1所示。

平均相对误差(Mean Relative Error, MRE)是一

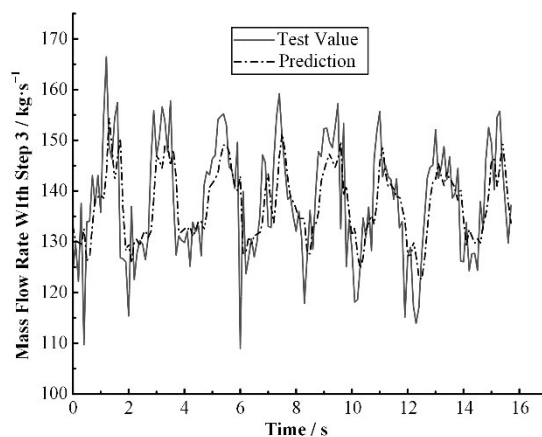


图8 步长为3的质量流量时间预测结果  
Fig.8 Prediction results of mass flow with a step of 3

表1 神经网络预测结果对比  
Table 1 Comparison of neural network prediction results

预测变量 Predictive variables	神经网络算法 Neural network algorithm	平均相对误差 Mean relative error
质量流量 Mass flow	自适应BP神经网络 Adaptive BP neural network	0.062 9(连续预测 Continuous prediction)
	自适应RBF神经网络 Adaptive RBF neural network	0.059 2(连续预测 Continuous prediction)
温度 Temperature	注意力+GRU Attention+GRU	0.042 9
	自适应BP神经网络 Adaptive BP neural network	0.011 3(连续预测 Continuous prediction)
	自适应RBF神经网络 Adaptive RBF neural network	0.008 8(连续预测 Continuous prediction)
	注意力+GRU Attention+GRU	0.004 2

种常用的衡量预测模型精度的指标,通常用于衡量预测值与实际值之间的差异程度。MRE的值越小,说明预测模型的精度越高。该指标可以用于评价各种预测模型的优劣,并帮助选择最佳的预测模型。

$$MRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_{\text{pred},i} - y_{\text{true},i}}{y_{\text{true},i}} \right| \quad (11)$$

从表1看出,本文使用的 Attention+GRU 算法在预测堆芯入口流量的平均相对误差为 0.042 9,比自适应 BP 神经网络和自适应 RBF 神经网络分别低了约 32.1% 和 27.6%,由图 7 可以看出,质量流量时间序列在后续 12 s 内预测效果较好,12 s 后出现明显滞后偏差。而在预测燃料包壳表面最高温度方面,本文使用的 Attention+GRU 算法的平均相对误差为 0.004 2,比自适应 BP 神经网络和自适应 RBF 神经网络分别低了约 62.8% 和 52.3%,由图 7 可以看出,温度时间序列在后续 15 s 内预测效果较好,15 s 后出现较为明显的结果偏差。因此,可以认为本文使用的 Attention+GRU 算法在这些预测变量上的预测

效果都有较大的提升。

与此同时,我们可以观察到,质量流量时间序列没有温度时间序列的预测效果好,是因为相较于温度而言,质量流量的数据会受到多种复杂因素的影响,比如流体的流动不稳定性,环境因素(温度、压力、湿度)等,这使得质量流量的数据可能更复杂,有更多的噪声,使得预测更困难,所以精度稍低。

### 2.3.2 消融实验

在深度学习领域的论文中,消融对比实验<sup>[22]</sup>是指通过对神经网络结构中的某一部分进行删减或禁用来研究该部分对于网络性能的贡献。消融对比实验是一种重要的评估技术,可以帮助我们了解网络结构的组成和作用,并且可以帮助我们更好地设计和优化神经网络。

为了更好地验证去噪结果的有效性,本文通过比较有无进行去噪处理的数据集对模型的影响来分析去噪操作的重要性和效果。定量实验结果如表 2 所示。

表2 小波去噪的消融实验对比  
Table 2 Comparison of DB8 wavelet denoising ablation experiments

预测变量 Predictive variables	质量流量 Mass flow	温度 Temperature
原始数据 Raw data	0.051 6	0.004 7
去噪数据 Denoising data	0.042 9	0.004 2

从表2看出,对于堆芯入口流量,原始数据与去噪数据的相对误差为 16.8%,说明去噪操作对该预测变量有比较显著的影响。对于燃料包壳最高温度,原始数据与去噪数据的相对误差为 10.4%,说明去噪操作对该预测变量的影响比质量流量小一些。由于数据本身温度噪声就比流量噪声小,所以去噪前后变化小属于正常情况。综上所述,进行去噪操

作确实可以在一定程度上提高预测结果的准确性。

为了验证 Attention 机制的有效性,本次在 GRU 的基准网络上,基于去噪后的数据集,采用与上述实验相同的超参数设置,通过比较基准模型和加入 Attention 机制的模型的表现,观察模型的性能变化。定量实验结果如表 3 所示。

表3 注意力机制的消融实验对比  
Table 3 Comparison of attention mechanism ablation experiments

预测变量 Predictive variables	质量流量 Mass flow	温度 Temperature
GRU 基准网络 GRU benchmark network	0.049 0	0.004 3
GRU 基准网络+Attention GRU benchmark network + Attention	0.042 9	0.004 2

从表3看出,对于堆芯入口流量的预测,相较于加了注意力的网络,GRU基准网络的预测效果略差。在对于燃料包壳最高温度预测时,两者相差不大。但是均优于连续预测的自适应BP神经网络和自适应RBF神经网络,说明GRU和注意力机制在一定程度上均能提高模型的性能。

### 3 结语

本文针对瞬态工况下CEFR燃料包壳最高温度以及质量流量预测问题,建立基于软注意力的GRU模型的反应堆瞬态热工参数预测方法,与目前常用的两种自适应神经网络算法的分析对比,得出如下结论:

1)选用1/2的CEFR堆芯为研究对象,开展反应堆瞬态热工工况下燃料包壳表面的最高温度和质量流量预测,单步预测在实际工程应用中没有较大意义。在对堆芯温度和质量流量的连续预测中,相较于自适应RBF神经网络和自适应BP神经网络,本文使用的基于软注意力的GRU网络模型连续预测结果更好,温度的平均相对误差不超过0.5%,在15 s内预测效果较好;质量流量的平均相对误差不超过5%,在12 s内预测效果较好。

2)基于注意力机制的神经网络能够有效捕捉到动态时序的趋势特征,帮助预测反应堆热工水力参数的不稳定趋势,对于提高反应堆安全性有重要价值。

3)本文提出的模型,仅针对1/2中国实验快堆进行质量流量预测和温度预测,在选定合适的步长后具有很好的瞬时预测效果。对于步长超参数确定、预测效率以及后续提高模型的泛化能力,还需要做进一步的实验。

**作者贡献声明** 淳思琦负责论文整体设计, SUBCHANFLOW程序设计,起草文章;冯欢负责获取CEFR堆芯的有效数据,数据分析;张安妮负责神经网络算法设计,实验实现;赵鹏程负责对文章作批评性审阅,研究经费支持,思想政治指导。

### 参考文献

- 1 Lu Q, Yuan Y, Li F C, *et al.* Prediction method for thermal-hydraulic parameters of nuclear reactor system based on deep learning algorithm[J]. Applied Thermal Engineering, 2021, **196**: 117272. DOI: [10.1016/j.applthermaleng.2021.117272](https://doi.org/10.1016/j.applthermaleng.2021.117272).
- 2 焦炜. 机器学习算法在多相热工水力参数快速预测的应用研究[D]. 北京: 华北电力大学, 2022.  
JIAO Wei. Application of machine learning algorithm in

- rapid prediction of multiphase thermal-hydraulic parameters[D]. Beijing: North China Electric Power University, 2022.
- 3 靳爽, 刘晓晶, 程旭. 基于神经网络的CFD粗网格模拟优化方法研究[J]. 核技术, 2021, **44**(6): 060601. DOI: [10.11889/j.0253-3219.2021.hjs.44.060601](https://doi.org/10.11889/j.0253-3219.2021.hjs.44.060601).  
JIN Shuang, LIU Xiaojing, CHENG Xu. Optimization method of CFD coarse grid numerical simulation based on neural network[J]. Nuclear Techniques, 2021, **44**(6): 060601. DOI: [10.11889/j.0253-3219.2021.hjs.44.060601](https://doi.org/10.11889/j.0253-3219.2021.hjs.44.060601).
- 4 李仕鲜, 刘井泉, 沈永刚. 基于神经网络方法的LOCA事故诊断[J]. 核技术, 2017, **40**(8): 080604. DOI: [10.11889/j.0253-3219.2017.hjs.40.080604](https://doi.org/10.11889/j.0253-3219.2017.hjs.40.080604).  
LI Shixian, LIU Jingquan, SHEN Yonggang. Fault diagnosis of LOCA based on ANN methods[J]. Nuclear Techniques, 2017, **40**(8): 080604. DOI: [10.11889/j.0253-3219.2017.hjs.40.080604](https://doi.org/10.11889/j.0253-3219.2017.hjs.40.080604).
- 5 陈涵瀛. 核电站热工水力系统工况预测与诊断方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2018.  
CHEN Hanying. Study on working condition prediction and diagnosis method of thermal hydraulic system in nuclear power plant[D]. Harbin: Harbin Engineering University, 2018.
- 6 Hochreiter S, Schmidhuber J. Long short-term memory [J]. Neural Computation, 1997, **9**(8): 1735 - 1780. DOI: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735).
- 7 Cho K, Van Merriënboer B, Bahdanau D, *et al.* On the properties of neural machine translation: Encoder-decoder approaches[EB/OL]. arXiv preprint arXiv: 1409.1259, 2014.
- 8 Chung J, Gulcehre C, Cho K, *et al.* Empirical evaluation of gated recurrent neural networks on sequence modeling [EB/OL]. 2014: arXiv: 1412.3555. <https://arxiv.org/abs/1412.3555>.
- 9 靳爽, 刘晓晶, 程旭. 基于神经网络的CFD粗网格模拟优化方法研究[J]. 核技术, 2021, **44**(6): 060601. DOI: [10.11889/j.0253-3219.2021.hjs.44.060601](https://doi.org/10.11889/j.0253-3219.2021.hjs.44.060601).  
JIN Shuang, LIU Xiaojing, CHENG Xu. Optimization method of CFD coarse grid numerical simulation based on neural network[J]. Nuclear Techniques, 2021, **44**(6): 060601. DOI: [10.11889/j.0253-3219.2021.hjs.44.060601](https://doi.org/10.11889/j.0253-3219.2021.hjs.44.060601).
- 10 Chen S, Ge L. Exploring the attention mechanism in LSTM-based Hong Kong stock price movement prediction[J]. Quantitative Finance, 2019, **19**(9): 1507-1515.
- 11 Qin Y, Song D, Chen H, *et al.* A dual-stage attention-



- based recurrent neural network for time series prediction [EB/OL]. arXiv preprint arXiv: 1704.02971, 2017.
- 12 Fei J T, Ding H F. Adaptive sliding mode control of dynamic system using RBF neural network[J]. *Nonlinear Dynamics*, 2012, **70**(2): 1563 - 1573. DOI: [10.1007/s11071-012-0556-2](https://doi.org/10.1007/s11071-012-0556-2).
  - 13 谭新. 基于RBF神经网络的谐波电流检测与抑制技术研究[D]. 吉林: 东北电力大学, 2022.  
TAN Xin. Research on harmonic current detection and suppression technology based on RBF neural network[D]. Jilin: Northeast Dianli University, 2022.
  - 14 Rana R. Gated recurrent unit (GRU) for emotion classification from noisy speech[EB/OL]. arXiv preprint arXiv:1612.07778, 2016.
  - 15 Prakash P R, Anuradha D, Iqbal J, *et al.* A novel convolutional neural network with gated recurrent unit for automated speech emotion recognition and classification [J]. *Journal of Control and Decision*, 2023, **10**(1): 54-63.
  - 16 陈先兵. 模拟核热耦合的低压自然循环流动不稳定性研究[D]. 哈尔滨: 哈尔滨工程大学, 2019.  
CHEN Xianbing. Study of simulated coupled neutronic-thermo-hydraulic low pressure natural circulation flow instability[D]. Harbin: Harbin Engineering University, 2019.
  - 17 吴晓建, 危一华, 王爱春, 等. 基于融合Dropout与注意力机制的LSTM-GRU车辆轨迹预测[J]. *湖南大学学报(自然科学版)*, 2023, **50**(4): 65 - 75. DOI: [10.16339/j.cnki.hdxzbzkb.2023155](https://doi.org/10.16339/j.cnki.hdxzbzkb.2023155).  
WU Xiaojian, WEI Yihua, WANG Aichun, *et al.* Vehicle trajectory prediction based on LSTM-GRU integrating dropout and attention mechanism[J]. *Journal of Hunan University (Natural Sciences)*, 2023, **50**(4): 65 - 75. DOI: [10.16339/j.cnki.hdxzbzkb.2023155](https://doi.org/10.16339/j.cnki.hdxzbzkb.2023155).
  - 18 Zhu X Z, Cheng D Z, Zhang Z, *et al.* An empirical study of spatial attention mechanisms in deep networks[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). October 27 - November 2, 2019, Seoul, Korea (South). IEEE, 2020: 6687 - 6696. DOI: [10.1109/ICCV.2019.00679](https://doi.org/10.1109/ICCV.2019.00679).
  - 19 皇甫晓瑛, 钱惠敏, 黄敏. 结合注意力机制的深度神经网络综述[J]. *计算机与现代化*, 2023(2): 40 - 49, 57. DOI: [10.3969/j.issn.1006-2475.2023.02.008](https://doi.org/10.3969/j.issn.1006-2475.2023.02.008).  
HUANGFU Xiaoying, QIAN Huimin, HUANG Min. A review of deep neural networks combined with attention mechanism[J]. *Computer and Modernization*, 2023(2): 40 - 49, 57. DOI: [10.3969/j.issn.1006-2475.2023.02.008](https://doi.org/10.3969/j.issn.1006-2475.2023.02.008).
  - 20 Shoaib M, Shamseldin A Y, Melville B W. Comparative study of different wavelet based neural network models for rainfall-runoff modeling[J]. *Journal of Hydrology*, 2014, **515**: 47 - 58. DOI: [10.1016/j.jhydrol.2014.04.055](https://doi.org/10.1016/j.jhydrol.2014.04.055).
  - 21 孙岩, 雷超, 董恒, 等. 历史数据下优化GRU模型的测控设备相位预测[J]. *测控技术*, 2023, **42**(1): 16 - 20. DOI: [10.19708/j.ckjs.2023.01.003](https://doi.org/10.19708/j.ckjs.2023.01.003).  
SUN Yan, LEI Chao, DONG Heng, *et al.* Phase prediction of measurement and control equipment based on optimizing GRU model under historical data[J]. *Measurement & Control Technology*, 2023, **42**(1): 16 - 20. DOI: [10.19708/j.ckjs.2023.01.003](https://doi.org/10.19708/j.ckjs.2023.01.003).
  - 22 Du L. How much deep learning does neural style transfer really need? an ablation study[C]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV). March 1-5, 2020, Snowmass, CO, USA. IEEE, 2020: 3139 - 3148. DOI: [10.1109/WACV45572.2020.9093537](https://doi.org/10.1109/WACV45572.2020.9093537).