

# 基于时空图卷积神经网络的管网产量模拟计算<sup>①</sup>



张志远

(中国石油大学(华东)青岛软件学院、计算机科学与技术学院, 青岛 266580)

通信作者: 张志远, E-mail: [z21070215@s.upc.edu.cn](mailto:z21070215@s.upc.edu.cn)

**摘要:** 针对原油集输管网的流量计测量数值偏差大, 模拟仿真软件人工校正繁琐、自适应差的问题, 提出一种自适应时空图卷积神经网络的产量计算方法, 实现原油集输管网产量的模拟计算. 以潜油电泵井为节点, 输油管道为边构建管网拓扑图. 使用图卷积神经网络提取井分布空间信息, 时间卷积神经网络获取产量数据的时间序列特征, 计算得到准确的产量模拟计算结果. 在某油田原油集输管网系统上进行了实验验证, 结果表明本文方法能够准确对管网系统内各电泵井的产量进行计算, 与其他基准网络模型相比, 各项误差指标均有下降, 平均绝对误差降至 0.87, 平均绝对百分比误差降至 4.45%, 均方误差降至 0.84, 证明了提出方法的有效性和准确性.

**关键词:** 油气集输; 产量计算; 图卷积神经网络; 时间卷积神经网络; 时空相关性

引用格式: 张志远. 基于时空图卷积神经网络的管网产量模拟计算. 计算机系统应用, 2024, 33(6): 169-176. <http://www.c-s-a.org.cn/1003-3254/9524.html>

## Simulation Calculation of Pipeline Network Yield Based on Spatio-temporal Graph Convolutional Neural Network

ZHANG Zhi-Yuan

(Qingdao Institute of Software & College of Computer Science and Technology, China University of Petroleum, Qingdao 266580, China)

**Abstract:** Flowmeter measurement values have a large deviation in crude oil gathering and transmission pipeline network, and the manual correction of simulation software is cumbersome with poor adaptive. To solve these problems, this study proposes an adaptive spatio-temporal graphic convolutional neural network production calculation method to realize the simulation calculation of crude oil gathering and transmission pipeline network production. The topology of the pipeline network is constructed with the submerged oil electric pump wells as nodes and the oil pipelines as edges. The study utilizes the graph convolutional neural network to extract the spatial information of well distribution and the temporal convolutional neural network to obtain the time series characteristics of the production data, so as to calculate the accurate production simulation results. The experimental validation is carried out on the crude oil gathering and transmission pipeline network system of an oil field. The results show that the proposed method can accurately calculate the production of each electric pump well in the pipeline network system. Compared with other baseline network models, the error indexes are reduced: the average absolute error is reduced to 0.87; the average absolute percentage error is reduced to 4.45%; the mean square error is reduced to 0.84, which proves the validity and accuracy of the proposed method.

**Key words:** oil and gas gathering and transportation; production calculation; graph convolutional neural network; temporal convolutional neural network; spatio-temporal correlation

① 基金项目: 青岛市自然科学基金 (23-2-1-162-zyyd-jch)

收稿时间: 2023-12-14; 修改时间: 2024-01-17; 采用时间: 2024-01-29; csa 在线出版时间: 2024-05-07

CNKI 网络首发时间: 2024-05-10

石油集输是指把分布在油田的各潜油电泵井所开采的石油及其混合物集中起来,经过初步分离加工,通过长距离输油管线输送到陆地原油库的全部过程.管网产量即各电泵井汇总后输出总产量,油田对管网产量进行计算主要分为两种:正算法和倒算法<sup>[1]</sup>.正算法是从生产的角度出发,计算管网中各潜油电泵井的产量后再计算石油产量;倒算法是从销售与使用的角度,根据周期销售量、使用量等,经过盘库后进行推算.正算法中,对于集输管网中各油井汇入管道的产量计算主要通过机理模型或模拟仿真软件实现.机理模型指结合油井液体密度流速公式、井口压力流速公式、管道压降等物理知识形成的计算方法.油田的滚动开发过程中,随着潜油电泵井和输油管道投入运营的数量不断增加,集输管网结构复杂度急剧升高,产量模拟计算需要考虑更多的约束条件.受电泵井启停操作、井口非线性回压、管道温降、可变摩阻等因素影响,基于传统机理模型的正向产量计算方法与实际数据偏差大,倒算法人工成本高、统计周期长、盘库存在计量与统计误差,亟需一种准确、快速的原油集输管网产量模拟计算方法.

国内外针对石油产量从各方向进行研究,在基于专家系统分析法中,Almedallah 等人<sup>[2]</sup>结合使用线性逼近约束优化(COBYLA)和混合整数线性规划(MILP)来集成实际钻井路径、平台位置和井分配的优化,根据上述因素分析产量.李传亮等人<sup>[3]</sup>分析油井压裂情况,将地层中的复杂流分为径向流和线性流,根据水电相似原理,求得产量计算公式.张鑫等人<sup>[4]</sup>通过结合抽油井泵效公式、压强公式和能耗,得到了压力、耗电量与产量的关系,量化分析了集输系统中回压对产能的影响;基于统计学方法中,Airassas 等人<sup>[5]</sup>使用 Aquila 优化器改进神经模糊推理系统(ANFIS),开发 AO-ANFIS 模型,计算分析了也门两个油田生产过程中的产量变动情况.Zhao 等人<sup>[6]</sup>开发了井间数值模拟模型(INSIM)和带前沿跟踪的 INSIM(INSIM-FT),用于注水储层的历史匹配、预测、特征描述和优化,通过精确的井间连通性计算、动态流路跟踪和注水预测来分析变动后产量;基于深度学习的方法中,Fan 等人<sup>[7]</sup>融合自回归综合移动平均模型 ARIMA 和 LSTM 模型,提取时间和空间相关特征,提高了流体计算的准确性;Abdullayeva 等人<sup>[8]</sup>提出一种卷积神经网络(CNN)和长短期记忆网络(LSTM)的混合网络模型,通过卷积

操作解决时间窗口之间的连续性问题,初步实现产量计算.Chahar 等人<sup>[9]</sup>提出一种结合人工神经网络(ANN)、随机森林回归器(RF)和梯度提升回归器(GB)的方法,利用可用的生产参数来预测每日石油产量.张战敏等人<sup>[10]</sup>开发了一种由双层 GRU 网络和 3 层全连接网络组成的深度学习模型,提高了电参量油模型的准确性和泛化性,提升了产量预测的精度.Gong 等人<sup>[11]</sup>提出一种机理知识与数据双驱动建模方法,将各井口的实时生产数据输入到模型中,通过 Beggs-Bril 和 Leapienzon 方程计算管道各处的压力和温度,然后将测量值和计算值输入到储运机制约束的基于 Transformer 模型中,对方程的计算结果进行误差修正,以达到准确的目的,实现管网各处运行参数的综合预测.

但上述方法存在人工参与程度高、计算繁琐、误差较大、普适性差的问题,对管网拓扑图结构信息和时序信息关注度不足,导致计算偏差大、周期长,结果不准确.因此本文主要工作如下.

(1) 提出一种三维矩阵化电泵井数据表示学习方法,便于挖掘数据之间潜在的关系,方便模型对时空数据进行分析处理,提高模型的性能和效率.

(2) 从时空角度出发,设计一种基于时空图卷积神经网络(AST-GCN)的产量计算模型,提取电泵井生产数据中的时间和空间特征,根据管网运行状况,准确模拟计算集输管网产量.

## 1 集输管网的表示学习

考虑到集输管网内各井的生产参数会直接影响其余相邻井的产量,充分利用该井和相邻井的生产参数有利于提高产量模拟计算的准确度和可靠性,因此我们提出一种集输管网表示学习方法,为时空图卷积神经网络模型的训练及优化打下基础.

### 1.1 电泵井时空特征矩阵构建

结合电泵井空间位置、生产参数和时间序列,构造电泵井的三维时序特征矩阵.如图 1 所示,将三维矩阵图 1(a)沿时间轴切片,得到“空间-特征”二维平面数据图 1(b),其数据表示如图 1(d)所示, $WF_t$ 表示在时间  $t$  时  $W$  个电泵井的  $F$  项特征值( $c$ );将三维矩阵图 1(a)沿因素轴切片,得到“时间-特征”二维平面数据图 1(c),其数据表示如图 1(e)所示, $TF_w$ 表示电泵井  $w$  的  $F$  项特征值的时间序列.模型针对两个切面分别用图卷积神经网络和时间卷积神经网络进行特征提取和学习,提

高模拟计算的准确度和可靠性. 3D 矩阵可以用式 (1) 表示.

$$WFT = (WF, TF) = \{e(w, f, t), w, f, t \in \mathbb{N}\} \quad (1)$$

其中,  $WFT$  表示“电泵井-特征-时间”的 3D 矩阵;  $WF$

和  $TF$  分别表示“空间-特征”“时间-特征”的二维平面数据集;  $e(w, f, t)$  表示一个属性特征的浮点数值,  $w$  是电泵井编号,  $f$  是属性编号,  $t$  表示时间.

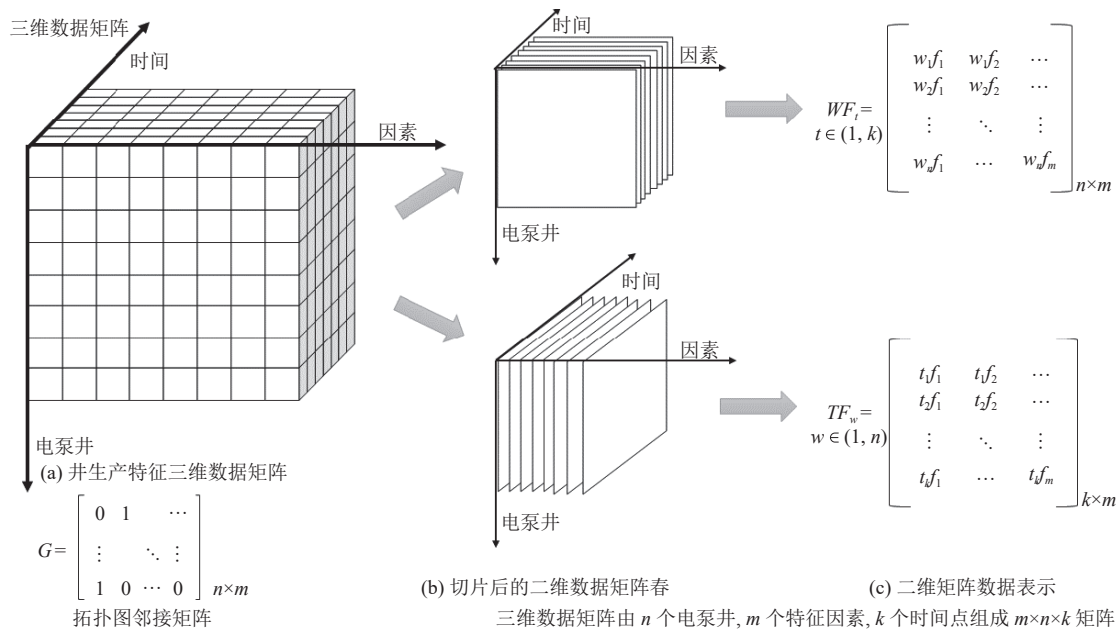


图 1 集输管网学习表示与切片操作流程

### 1.2 管网时空图构建

集输管网由电泵井和输油管道组成, 其拓扑结构错综复杂, 可以抽象定义为非欧图数据. 我们定义一个连续时刻序列, 其中在  $t$  时刻及其相邻时刻中连续观察到多个时刻的图结构数据没有发生任何变化, 将此时间序列下的集输管网进行图论化, 和特征矩阵一起合并为一个管网数据快照 (data snapshots), 用式 (2) 表示:

$$G_t = \{V_t, E_t, WFT\} \quad (2)$$

其中,  $G_t$  为电泵井节点数据、管网拓扑图数据和特征矩阵数据的集合,  $V_t = \{v_t^1, v_t^2, \dots, v_t^n\}$  表示在  $t$  时刻  $n$  个电泵井列表;  $E_t = (e_{ij})_{n \times n}$  表示  $n$  个电泵井通过管道连通情况的二维矩阵, 若两井通过管道连通, 则  $e_{ij} = 1$ , 否则  $e_{ij} = 0$ ;  $WFT$  表示在  $t$  时刻后的连续时刻组成的电泵井时空特征矩阵.

### 1.3 问题定义

在电泵井实际生产过程中, 时常根据生产计划调整油井生产时间, 调整某口井会对整体管网系统产生一定影响, 时空图卷积神经网络模型可对停开井前的集输管网时间快照进行分析, 预测产量数据集  $Y =$

$\{y_0, y_1, \dots, y_m\}$  其中管网总产量为  $y_0$ , 其余为各电泵井或井组平台的产量数据. 使用管网数据快照  $G = (V, E, WFT)$  表示管网的时空信息,  $V$  表示图的点信息集合,  $E$  表示边信息集合. 综上, 任务描述为: 基于若干管网数据快照, 训练得到一个映射  $f: \mathbb{R}^{N \times F \times T} \rightarrow \mathbb{R}^{N \times 1}$ , 去对产量进行预测, 即  $\{V, E, WFT\} \xrightarrow{f} \{y_0, y_1, \dots, y_m\}$ .

## 2 时空图卷积神经网络

### 2.1 网络架构

各潜油电泵井和输油管道组成的管网图数据为非欧几里德结构数据, 建立在欧氏数据上的卷积神经网络 (CNN) 并不能发挥有效作用, 提取空间结构信息. 为有效利用管网数据的空间结构信息和时间序列信息, 本文基于图卷积神经网络 (GCN)<sup>[12]</sup>、时间卷积神经网络 (TCN)<sup>[13]</sup> 和长短期记忆网络 (LSTM)<sup>[14]</sup> 搭建了一种 3 层时空交替的网络模型 AST-GCN 如图 2 所示.

图 2 中自下而上, 时空特征矩阵和邻接矩阵作为输入; 第 1 层为空间特征提取层, 由 adaptive module 组成, 内部使用 LSTM-GCN 提取管网拓扑图的空间信息, GCN 使各节点聚合一阶邻域的信息, 同时 LSTM

对 GCN 处理拓扑不同的图带来的权重变化进行长短期记忆并更新, 最后形成各节点带有空间信息的高质量嵌入向量输出到下一层; 第 2 层为时间特征提取层, 使用因果扩张卷积的 TCN 对时间序列数据进行扩张因果卷积, 使每个节点获取自身历史信息; 第 3 层为时空信息融合层, 此时传入的节点向量, 分别包含当前时刻自身和一阶邻域节点信息, 历史时刻自身信息, 本模

块使用 GCN 进行一次一阶邻域图卷积, 使每个节点包含当前时刻自身、相邻节点的信息, 历史时刻自身、相邻节点的信息, 形成蕴含丰富时空信息的向量, 最后通过全连接输出为管网产量数据. 模型通过 3 个模块的时空交替学习, 能够深度利用时间和空间信息, 进行准确的管网产量模拟计算, 得到管网系统中各电泵井和总输出产量情况.

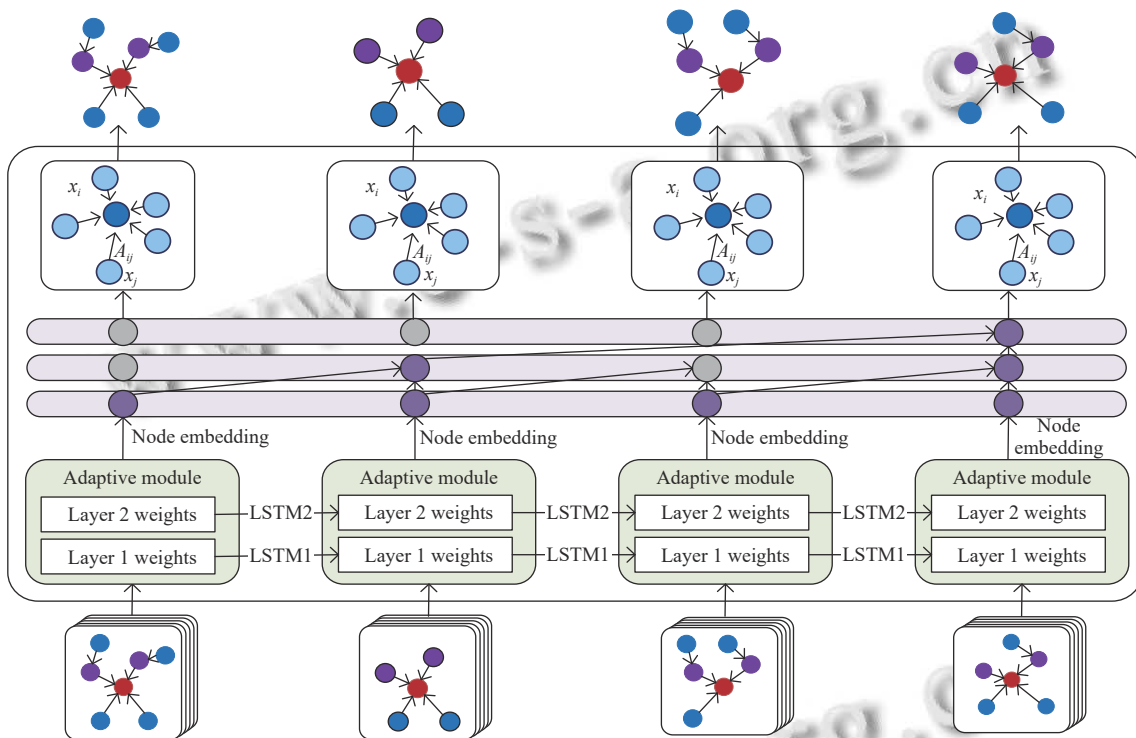


图 2 时空图卷积神经网络模型架构图

### 2.2 自适应图卷积空间特征提取

图论化的集输管网拓扑图具有非欧氏的空间结构, 节点之间存在复杂的空间依赖关系, 图中的每个结点均受到相邻节点影响, 直到最终平衡. 同时, 受生产计划影响, 管网拓扑图会随时间产生变化, 节点和边会被添加或删除.

本文采用 GCN 提取图中的节点特征, 将特征表示为向量; LSTM 学习演化 GCN 中的参数, 加强 GCN 对不同图的适应性. 两者结合完成自适应空间特征提取任务, 过程如图 3(a) 所示. 图 3(a) 为 GCN-LSTM 更新流程图, 其中  $WF_t$  为输入的特征矩阵,  $WF_{t+1}$  为特征提取后的矩阵,  $W_t^l$  为 GCN 的第  $l$  层权重参数. 图 3(a) 中, 首先 LSTM 对图卷积神经网络的权重参数  $W_t^l$  进行更新, 并且记忆当前状态,  $w_t^l = LSTM(w_{t-1}^l)$ ; 然后 GCN 使

用更新后的权重矩阵对各节点的嵌入向量进行图卷积操作, 即  $H^{l+1} = \sigma(AH^l W_t^l)$ . 图 3(b) 为图卷积原理, GCN 对每个节点的一阶邻域进行卷积计算, 节点嵌入的更新公式如式 (3) 所示:

$$H_i^{l+1} = \sigma \left( \frac{1}{c_i} \sum_{j \in N_i} H_j^l W_t^l \right) \quad (3)$$

其中,  $N_i$  表示节点  $i$  的所有邻居节点,  $c_i$  是归因子, 每个节点  $i$  结合自身以及所有邻居节点进行计算更新. 图 3(c) 为 LSTM 原理图, LSTM 通过遗忘门、输入门和输出门这 3 个门控制信息流动或遗忘之前的信息, 使其可以处理长短序列信息. 各门控单元的计算方式如式 (4)~式 (7) 所示. 其中,  $x_t$  是来自 GCN 的权重, 其余为 LSTM 内部权重等参数.

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (4)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (5)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (6)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (7)$$

其中,  $W_f, W_i, W_o, W_c$  是 LSTM 的权重参数;  $b_f, b_i, b_o, b_c$  是 LSTM 状态更新的偏置;  $f_t, i_t, o_t$  分别代表遗忘门、输入门和输出门,  $c_t$  是单元状态, 3 个门控和单元状态形成对 GCN 权重矩阵的长短期记忆。

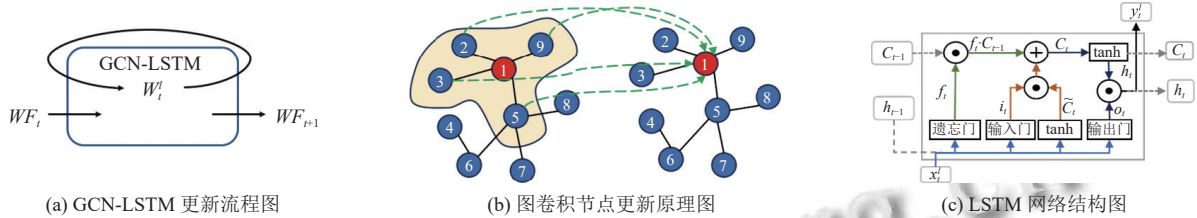


图3 GCN-LSTM 自适应更新机制流程图

对于集输管网 3D 特征矩阵  $WFT$ , 按照时间轴切片操作后, 每一个  $t$  时刻的切片  $WF_t = H_t$ , 代表一个管网拓扑图在  $t$  时间的数据快照. 对快照中的拓扑图进行图卷积, 方法如式 (3) 所示. GCN 聚合电泵井节点的 1 阶邻居节点特征以更新节点的特征表示  $H^{l+1}$ .  $D^{-1/2} \hat{A} D^{-1/2}$  为正则化拉普拉斯矩阵.  $W^l$  为第  $l$  层的可训练权重参数矩阵.

$$H^{l+1} = \sigma \left( D^{-\frac{1}{2}} \hat{A} D^{-\frac{1}{2}} H^l W^l \right) \quad (8)$$

$$h_t = \text{LSTM}(h_{t-1}, W_h, W_x) \quad (9)$$

$$h_t = \sigma \left( \sum_{u \in \mathcal{N}(v)} W_h h_u + W_x x_v \right) \quad (10)$$

当图结构发生变化时, 需要重新计算节点之间的特征关系, 调整 GCN 的权重参数, 以适应新的拓扑图结构<sup>[15]</sup>. 在表示学习中, 拓扑图的变化直接体现在不同快照之间邻接矩阵  $A$  的变化和 3D 数据矩阵  $WF_t$  的变化. 使用 LSTM 在时间维度上向后传递了两份信息: (1) 包含 GCN 权重变动长期信息的单元状态  $C_t$ ; (2) 包含序列短期信息的状态  $h_t$ . LSTM 接收自身的权重矩阵  $W_h$ 、GCN 的权重  $W_x$ , 和其先前的状态向量  $h_{t-1}$ , 计算出新的状态向量  $h_t$ , 如式 (9) 所示. 经过引入 LSTM 模型, GCN 参数传递更新机制如式 (10) 所示. 其中,  $u \in \mathcal{N}(v)$  是  $v$  的邻居节点集合,  $W_h$  和  $W_x$  分别是 LSTM 和 GCN 的权重矩阵. 在迭代过程中, 不仅包含节点邻域信息, 还受历史权重的影响.

### 2.3 井参数时间序列特征提取

原油集输管网的生产参数有时间依赖性, 因此我们使用 TCN 提取时序特征. 通过因果卷积确保对时间

序列进行建模时, 只依赖  $0 \sim t-1$  时刻的值来预测第  $t$  时刻的特征值; 采用扩张卷积以保持计算效率的同时扩大卷积核视野, 捕捉序列中中长期关系. TCN 结构的扩张因果卷积如图 4 所示.

$$F(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d-i} \quad (11)$$

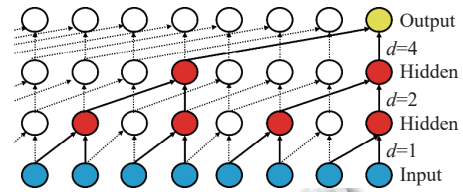


图4 TCN 扩张因果卷积

对于当前的时间步  $s$ , 在时间步  $s$  处的输出值  $F(s)$  根据输入序列  $x$  和权重系数  $f$  进行计算得到, 如式 (11) 所示. 式中  $k$  为卷积核大小,  $i$  表示权重系数的索引,  $f(i)$  代表在卷积核中的第  $i$  个位置上的权重值.  $x_{s-d-i}$  表示输入序列  $x$  中与当前时间步的偏移量  $d \cdot i$  相关的值, 这里的  $d$  是一个常数, 用于控制每个时间步之间的距离. 此方法使有效窗口的大小随着层数呈指数型增长, 网络用比较少的层获得较大的感受野. 将构造的 3D 数据矩阵沿特征轴切片后得到“时间-特征”二维平面数据  $TF_w$ , 使用 TCN 对  $TF_w$  中每个特征的时间序列  $x$  进行特征提取.

### 2.4 全局特征融合

$t$  时刻自适应图卷积层聚合相邻点特征信息后, 使节点  $v_i^t$  拥有自身当前时刻和相邻节点  $v_{i+1}^t$  当前时刻的信息, 即  $v_i^t = v_i^t + v_{i+1}^t$ , 如图 5(a) 所示的变化过程; 时间卷积层聚合每个节点当前时刻  $v_i^t$  和历史时刻  $v_i^{t-1}$  信息,

即  $v_i^t = v_i^t + v_{i+1}^{t-1}$ , 如图 5(b) 所示. 由此,  $v_i^t = v_i^t + v_{i+1}^t + v_{i+1}^{t-1}$ , 每个节点缺少相邻节点的历史信息  $v_{i+1}^{t-1}$ .

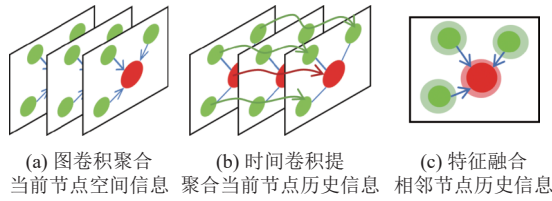


图 5 时空特征提取全流程

全局特征融合的目标是使每个节点拥有当前时间自身和邻居节点信息, 历史时间自身和邻居节点信息, 共 4 部分时空信息,  $v_i^t = v_i^t + v_{i+1}^t + v_{i+1}^{t-1} + v_{i+1}^{t-2}$ , 如图 5(c) 所示. 由于只缺少邻居节点的历史信息, 特征融合使用一阶邻域图卷积操作进行时空信息的最终融合, 其更新方式如式 (11) 所示, 但此时的  $H$  是包含历史特征的, 经此计算后的每个节点向量都包含邻居节点的历史信息, 形式化的表示为  $X_i^t = f(X_i^t, A_t, W_t^t)$ , 其中  $X_i^t$  为包含历史特征的节点向量. 最后通过全连接层, 将 8 维的向量

转换为浮点实数输出为产量数据.

### 3 实验分析

为验证提出方法在原油集输管网进行产量模拟计算的有效性和先进性, 我们在某油田区域的集输管网系统上进行了实验. 结果显示, 本文方法的产量计算获得了更准确的结果, 证明了该方法的有效性和先进性.

#### 3.1 数据集

选取某油田 2018 年 1 月 1 日–2022 年 12 月 31 日共计 5 年的生产数据, 构建数据集. 原始数据如表 1 所示. 每行数据包含井号、日期和 9 种传感器记录数据和最终人工统计的真实结果, 分别为: 油压 (YY)、回套 (HY)、套压 (TY)、井口温度 (WD)、耗电量 (HDL)、电压 (DY)、电流 (DL)、油嘴 (YZ)、日产油量 (RU), 日产油量是我们模拟计算的目标. 涉及保密数据, 表 1 中为脱密后的样例. 将其命名为 SLYTHR 数据集.

表 1 数据集样本

JH	DATE	YY	HY	TY	WD	DY	DL	YZ	HDL	RU
CA-G-1201	2018/01/01	4.0	1.5	0.8	65	1250	40	6	1236	70
CA-G-1201	2018/01/02	4.1	1.4	0.8	69	1200	45	6	980	60
CA-G-1201	2018/01/03	4.1	1.5	0.7	66	1250	50	6	1332	75
...	...	...	...	...	...	...	...	...	...	...
CA-G-2210	2022/07/20	6.5	3.5	0.8	63	1200	40	8	1023	67

将管网的每个数据快照作为一个训练数据单元, 所有训练单元组合成为实验所需数据集. 最终数据集共有 1 823 天的电泵井生产记录, 有 325 次电泵井停开造成拓扑变化, 产生 325 个时空特征 3D 数据矩阵, 每个数据矩阵中时间序列数为 1–71 不等, 将其按照 80%、20% 的比例划分训练集、测试集.

#### 3.2 对比实验设计

目前尚未有对集输管网产量计算方法, 对比方法选取现有采取不同核心技术且能够处理时空图数据的深度学习方法, 形成对比实验, 通过计算平均误差等评价指标来进行评价, 如表 2 所示.

表 2 对比实验方法

Reference	Prediction technique	Capability	
		Spatial	Temporal
Geom-GCN	GCN	√	—
SSA-TCN	TCN	—	√
TCN-GCN	GCN+TCN	√	√
PP-STNN	GCN+GRU	√	√

(1) Geom-GCN<sup>[16]</sup>: 模型通过利用连续空间来构建结构邻域, 克服传统消息传递类神经网络在处理图结构数据时的部分局限性.

(2) SSA-TCN<sup>[17]</sup>: 基于时间卷积神经网络的产量计算方法. 其使用麻雀搜索算法快速筛选出神经网络的最优超参数, 对网络进行优化.

(3) TCN-GCN<sup>[18]</sup>: 模型首先使用基于扩张因果卷积算法的 TCN, 从时序数据提取特征信息; 其次再构建拓扑结构图, 使用 GCN 从拓扑结构图中获得的信息来进行预测. GCN 的邻接矩阵中的元素被 Pearson 相关系数替代, 以提高预测的准确性.

(4) PP-STNN<sup>[19]</sup>: 使用图卷积网络 (GCN) 捕获管网的时空相关性, 门控循环单元 (GRU) 用于捕获时间相关性, 结合二者计算结果.

#### 3.3 实验环境与模型设置

本文方法所用模型运行在配有 Xeon E5-2686 v4

处理器和两块 GeForce RTX 3080 Ti 的 Torch 环境, 部分采用了图学习的扩展库 PyTorch-geometric<sup>[20]</sup>. 训练中, Adam 优化器设置为 0.001, 迭代次数为 500 次. 本文所提模型参数设置为: 自适应图卷积模块所需过滤器数量  $in\_channels$  设置为 8; 时间卷积层中输入的时间序列数据的输入、输出通道数  $input\_channels$ 、 $output\_channels$  设置为 8, 每个层级的通道数合为一个数组  $num\_channels$  设置为 [32, 64, 128, 8], 卷积核大小  $kernel\_size$  设置为 2,  $dropout$  设置为 0.2; 全局特征融合图卷积层节点维度特征  $features$  设置为 8, 隐藏层数  $hidden$  设置为 16.

评估体系采用平均绝对误差 MAE、平均绝对百分比误差 MAPE、均方差 MSE、均方根 RMSE 作为评价指标.

### 3.4 实验结果与分析

表 3 显示了 AST-GCN 模型与对比模型的计算结果. 从表 3 中可以看出, 本文提出的方法模型在 SLYTHR 数据的所有仿真中都达到了最好的计算能力. 图 6 和图 7 展示了各个方法的计算结果与真实值的对比.

表 3 不同方法在 SLYTHR 数据集上的实验结果

Model	MAE	MAPE (%)	MSE	RMSE
GCN	3.13	15.31	11.08	3.33
TCN	4.27	21.61	19.21	4.38
GCN+GRU	2.46	11.92	8.84	2.97
GCN+TCN	1.42	7.20	3.29	1.82
Ours	0.87	4.45	0.84	0.92

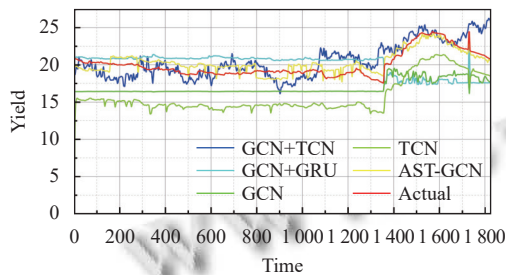


图 6 计算结果对比曲线图

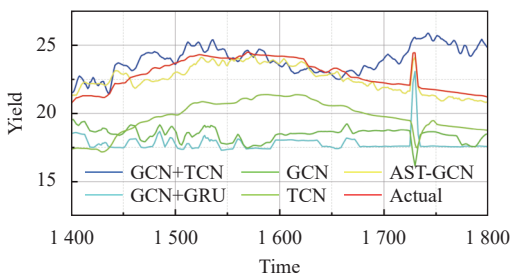


图 7 计算结果对比曲线图

GCN 的计算结果如图中深绿色曲线所示, 由于只考虑了空间结构, 因此计算结果差异较大; TCN 的计算结果如草绿色曲线所示, 它只考虑了数据之间的时间相关性, 忽略了图结构; GCN+GRU 被用作对比模型, 其计算结果如青色曲线所示, 其计算结果不如 GCN+TCN 精确; GCN+TCN 的计算结果如蓝色曲线所示, 模型考虑了时空因素, 计算结果较准确, 但在数据变动较大的时间点, 不能很好地计算结果. 本文所提出的方法结果标注为黄色曲线, 和真实值 (红色曲线) 差距最小.

此外, 通过观察图 7 中 1700–1800 时刻内的曲线可以看出, 本文提出的 AST-GCN 模型能够应对电泵井变化导致的产量数据瞬变的情况, 这主要得益于自适应图卷积模块中的 LSTM 所具备的长短期记忆能力, 而 LSTM 在 GCN 参数的演化过程中起到了重要的作用.

## 4 结论与展望

本文提出了一种通过自适应时空图卷积神经网络进行原油集输管网产量模拟计算方法. 从集输管网数据的时空特征出发, 进行特征提取和融合计算; 通过时空反复结合, 彻底捕捉管网数据的时空特征, 准确进行集输管网产量模拟计算, 能够为油田集输过程提供参考.

下一步工作应考虑如何使用更少的数据样本进行小样本训练, 保持结果准确性的同时减少资源消耗.

### 参考文献

- 刘丽. 浅析原油盘库统计中的误差因素. 中国石油和化工标准与质量, 2012, 32(1): 243. [doi: 10.3969/j.issn.1673-4076.2012.01.213]
- Almedallah MK, Branch G, Walsh SDC. Combined well path, submarine pipeline network, route and flow rate optimization for shallow-water offshore fields. Applied Ocean Research, 2020, 105: 102396. [doi: 10.1016/j.apor.2020.102396]
- 李传亮, 庞彦明, 周永炳, 等. 压裂直井产量公式. 新疆石油地质, 2023, 44(6): 683–689.
- 张鑫, 高星星, 王繁荣. 油井集输回压对产能影响及控制策略. 化学工程与装备, 2023(7): 80–81, 114. [doi: 10.19566/j.cnki.cn35-1285/tq.2023.07.074]
- AlRassas AM, Al-Qaness MAA, Ewees AA, et al. Optimized ANFIS model using Aquila Optimizer for oil production forecasting. Processes, 2021, 9(7): 1194. [doi: 10.3390/pr907

- 1194]
- 6 Zhao H, Xu LF, Guo ZY, *et al.* Flow-path tracking strategy in a data-driven interwell numerical simulation model for waterflooding history matching and performance prediction with infill wells. *SPE Journal*, 2020, 25(2): 1007–1025. [doi: [10.2118/199361-PA](https://doi.org/10.2118/199361-PA)]
  - 7 Fan DY, Sun H, Yao J, *et al.* Well production forecasting based on ARIMA-LSTM model considering manual operations. *Energy*, 2021, 220: 119708. [doi: [10.1016/j.energy.2020.119708](https://doi.org/10.1016/j.energy.2020.119708)]
  - 8 Abdullayeva F, Imamverdiyev Y. Development of oil production forecasting method based on deep learning. *Statistics, Optimization & Information Computing*, 2019, 7(4): 826–839. [doi: [10.19139/soic-2310-5070-651](https://doi.org/10.19139/soic-2310-5070-651)]
  - 9 Chahar J, Verma J, Vyas D, *et al.* Data-driven approach for hydrocarbon production forecasting using machine learning techniques. *Journal of Petroleum Science and Engineering*, 2022, 217: 110757. [doi: [10.1016/j.petrol.2022.110757](https://doi.org/10.1016/j.petrol.2022.110757)]
  - 10 张战敏, 朱丹丹, 刘重伯, 等. 抽油机井原油产量计量新方法. *钻采工艺*, 2023, 46(2): 71–76. [doi: [10.3969/J.ISSN.1006-768X.2023.02.12](https://doi.org/10.3969/J.ISSN.1006-768X.2023.02.12)]
  - 11 Gong FM, Du CZ, Ji XF, *et al.* Mechanistic and data-driven modelling of operational parameters prediction on oil and gas transportation pipeline network. *Proceedings of the 2023 International Conference on New Trends in Computational Intelligence (NTCI)*. Qingdao: IEEE, 2023. 329–335. [doi: [10.1109/NTCI60157.2023.10403749](https://doi.org/10.1109/NTCI60157.2023.10403749)]
  - 12 Kipf TN, Welling M. Semi-supervised classification with graph convolutional networks. *Proceedings of the 5th International Conference on Learning Representations*. Toulon: OpenReview.net, 2017.
  - 13 Bai SJ, Kolter JZ, Koltun V. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv:1803.01271*, 2018.
  - 14 Sagheer A, Kotb M. Time series forecasting of petroleum production using deep LSTM recurrent networks. *Neurocomputing*, 2019, 323: 203–213. [doi: [10.1016/j.neucom.2018.09.082](https://doi.org/10.1016/j.neucom.2018.09.082)]
  - 15 Manessi F, Rozza A, Manzo M. Dynamic graph convolutional networks. *Pattern Recognition*, 2020, 97: 107000. [doi: [10.1016/j.patcog.2019.107000](https://doi.org/10.1016/j.patcog.2019.107000)]
  - 16 Pei HB, Wei BZ, Chang KCC, *et al.* Geom-GCN: Geometric graph convolutional networks. *Proceedings of the 8th International Conference on Learning Representations*. Addis Ababa: OpenReview.net, 2020.
  - 17 Zhang L, Dou H, Wang TZ, *et al.* A production prediction method of single well in water flooding oilfield based on integrated temporal convolutional network model. *Petroleum Exploration and Development*, 2022, 49(5): 1150–1160. [doi: [10.1016/S1876-3804\(22\)60339-2](https://doi.org/10.1016/S1876-3804(22)60339-2)]
  - 18 Luo DS, Fang J, He HY, *et al.* Prediction for dissolved gas in power transformer oil based on TCN and GCN. *IEEE Transactions on Industry Applications*, 2022, 58(6): 7818–7826. [doi: [10.1109/TIA.2022.3197565](https://doi.org/10.1109/TIA.2022.3197565)]
  - 19 Liao ZY, Liu MH, Du BW, *et al.* A temporal and spatial prediction method for urban pipeline network based on deep learning. *Physica A: Statistical Mechanics and Its Applications*, 2022, 608: 128299. [doi: [10.1016/j.physa.2022.128299](https://doi.org/10.1016/j.physa.2022.128299)]
  - 20 Fey M, Lenssen JE. Fast graph representation learning with PyTorch geometric. *arXiv:1903.02428v3*, 2019.

(校对责编: 孙君艳)