

## 二、建模与求解

### 问题三的建模与求解

#### 1. 基于气象预报数据的光伏功率预测模型构建

在问题三中，我们进一步提升光伏功率预测模型的精度与泛化能力。相比于问题二仅使用历史功率数据构建的 LSTM 模型，本问题引入了数值气象预报（Numerical Weather Prediction, NWP）信息，作为额外的输入变量，以实现对环境因子建模。NWP 数据通常由高分辨率气象模型计算获得，具有一定的时空精度，能为预测任务提供更加丰富的先验信息。

模型的输入仍然以滑动窗口的方式构造。具体地，当前时间点  $t$  的输入包含两个组成部分：一是该时间点前连续 96 个时刻的光伏功率观测值  $[P_{t-96}, P_{t-95}, \dots, P_{t-1}]$ ，用于捕捉功率时间序列的动态模式；二是当前时刻  $t$  对应的 NWP 气象要素，包括全球水平辐照度  $G_{\text{global}}$ 、直达辐照度  $G_{\text{direct}}$ 、气温  $T$ 、湿度  $H$ 、风速  $v$ 、风向  $\theta$  和大气压  $p$ 。这七个特征均来源于数据中提供的 `nwp_` 字段。为了匹配功率序列长度，这些气象变量以单点输入的形式与历史功率拼接构成最终输入向量  $x_t$ 。

模型结构方面，我们继续采用 LSTM 神经网络作为主干架构。LSTM 能够捕捉长时间序列中的依赖关系，并能有效处理光伏功率序列中的非线性变化。网络结构由一个序列输入层开始，输入维度与  $x_t$  的维度一致；接下来是一个包含 64 个单元的 LSTM 层；随后连接一个全连接层映射到输出维度；最后通过回归层计算预测输出。

训练过程采用 Adam 优化器，损失函数为均方误差  $\mathcal{L}_{\text{MSE}} = \frac{1}{N} \sum_{i=1}^N (\hat{P}_i - P_i)^2$ ，其

中  $\hat{P}_i$  表示第  $i$  个样本的模型预测值， $P_i$  表示实际功率值。最大训练轮数设置为 100 轮，并在训练过程中监控验证集误差以防止过拟合。

数据集划分方面，为方便和第二问对比，同样选择将 2、5、8、11 月的最后一周数据作为测试集，其余为训练集。在训练集中，再进一步划分出后 10% 的数据作为验证集，用于监控模型在非训练数据上的性能。所有数据均使用固定滑动窗口进行构造，确保序列特征与目标值严格对齐。

在模型评估阶段，我们采用两个标准化误差指标对预测结果进行评估，即归一化平均绝对误差  $\text{Emae}$  和归一化均方根误差  $\text{Ermse}$ 。这两个指标分别定义为：

$$\text{Emae} = \frac{1}{P_{\text{rated}}} \cdot \frac{1}{N} \sum_{i=1}^N |\hat{P}_i - P_i|$$
$$\text{Ermse} = \frac{1}{P_{\text{rated}}} \cdot \sqrt{\frac{1}{N} \sum_{i=1}^N (\hat{P}_i - P_i)^2}$$

其中  $P_{\text{rated}} = 20000kW$  表示该电站的额定装机容量。

上述两个指标均仅在白昼时段，即理论功率  $P_{\text{cs}} > 0$  的时间段中进行统计，从而剔除夜间无光照时段对评价的干扰，提升评估结果的针对性。

在模型预测完成后，为了进一步分析模型在不同天气条件下的表现，我们引入了全天空指数  $K_t$  作为气象条件分类的依据。全天空指数定义为某一时间点的实际发电功率与对应理论最大发电功率之比，即

$$K_t = \frac{P_t}{P_{\text{cs},t}}$$

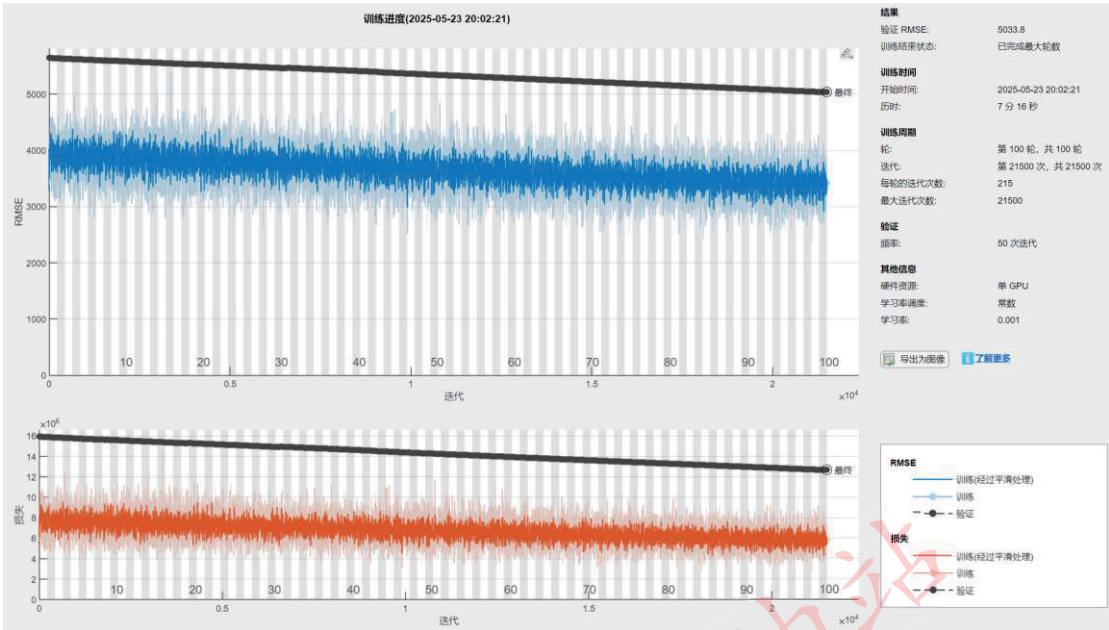
其中  $P_t$  表示该时刻的实际发电功率， $P_{\text{cs},t}$  为基于太阳高度角估算得到的理论最大功率。 $K_t$  是衡量光照条件是否理想的关键指标。理想情况下，当太阳高度较高、无云遮挡时， $K_t$  接近 1；而在多云或阴天等弱辐照条件下，该指数会显著下降。

为了系统性评估模型的泛化能力，我们将测试集样本按其全天空指数进行分类。参考文献与实际经验，我们采用如下分类标准：当  $K_t \geq 0.75$  时，视为晴天；当  $0.35 \leq K_t < 0.75$  时，视为多云；而当  $K_t < 0.35$  时，视为阴天。

对于每一类天气条件，我们分别提取对应的测试样本，并重新计算归一化平均绝对误差（Emae）与归一化均方根误差（Ermse）。

## 2. 模型求解与分析

- (1) 与单纯时间序列对比
- 将模型放入 matlab 中求解，求解图示：



从训练过程的表现来看，第三问中的 LSTM 模型表现出良好的收敛趋势。训练误差和验证误差曲线均保持稳定下降，尤其是验证集的 RMSE 从初始的约 5500 下降至最终的 5033.8，说明模型具备较好的泛化能力。整个训练过程中并未出现验证误差突然上升的现象，说明模型没有明显的过拟合问题，能够较好地适应新的天气样本。与第二问相比，这种稳步优化的趋势表明，NWP 信息确实为模型提供了有效的额外信息维度，有助于学习更复杂的非线性规律。

在模型预测误差方面，我们采用了归一化的平均绝对误差（EAMSE）与归一化的均方根误差（ERMSE）作为衡量指标。第二问模型的 EAMSE 为 0.2204，ERMSE 为 0.2745；第三问模型的 EAMSE 为 0.2155，ERMSE 为 0.2720。这些结果清晰地表明，引入 NWP 数据可以有效提高模型在测试集上的预测精度，尤其是在光照波动剧烈或气象条件多变的场景下，模型借助外部气象预报因子更好地拟合了输出目标。这种误差的细微改善是实质性的，反映出模型能力在更复杂输入信息下的增强。

### （2）不同天气场景划分结果

首先，在晴天条件下（ $K_t \geq 0.7$ ），RMSE 达到 6467.88kW，是三种天气中误差最大的情况。这是一个表面上看起来反直觉的结果，因为晴天天气稳定、辐照强、光伏出力模式也相对规律，理论上模型应该更容易学习其规律。然而，正是由于晴天时光伏出力峰值高、振幅大，因此一旦模型预测出现偏差，其数值误差会被放大。例如，当真实功率为 12000kW，而模型预测为 11000 kW，仅 8.3% 的相对误差就会导致 1000kW 的 RMSE 贡献。因此，晴天的高 RMSE 并不代表模型表现差，而是由于功率基值高导致的“放大效应”。

其次，在多云条件下（ $0.4 \leq K_t < 0.7$ ），RMSE 显著降低至 2526.91kW。这表明模型对多云天气具有较好的适应能力。在多云天气下，虽然短时辐照变化具有一定不确定性，但平均功率水平中等，模型预测的绝对值相对较低，从而减小了 RMSE 的整体量级。此外，多云天气下的功率曲线往往较为平滑，可能被模型较好地捕捉。

最后，在阴天条件下（ $K_t < 0.4$ ），RMSE 最低，仅为 471.82kW。这种结果主要源于两个方面：一是阴天光伏出力极低，大部分时间处于接近于零的功率状态；

二是低光照条件下即使预测存在偏差，其绝对误差也非常有限。例如，从真实值 200kW 误差到 300kW 的预测也仅产生 100kW 的 RMSE 贡献。因此，即便模型预测误差率可能并不小，但在功率绝对值很低的情况下，RMSE 很难被拉高。

具体代码如下：

```
clc; clear;

%% 1. 站点参数
eta_STC = 270 / 1.6635e3; % 组件效率 ≈0.1623
A_total = 74000;          % 组件总面积 m²
G_STC   = 1000;           % STC 辐照 W/m²
P_rated = 20000;          % 装机容量 kW
lat      = 38.18306;

%% 2. 读取数据
T = readtable('station01.csv','TextType','string');
rawStr = string(T.date_time);
rawStr = replace(rawStr, ' ', ' ');
parts  = split(rawStr, ' ');
dateP  = split(parts(:,1), '-'); timeP = split(parts(:,2), ':');
T.date_time = datetime(str2double(dateP(:,1)),str2double(dateP(:,2)), ...
    str2double(dateP(:,3)),str2double(timeP(:,1)), ...
    str2double(timeP(:,2)),str2double(timeP(:,3))) + hours(8);
T.power = T.power * 1000;

%% 3. 理论功率 P_cs
my_doy = day(T.date_time,'dayofyear');
my_hourf = hour(T.date_time) + minute(T.date_time)/60;
delta = deg2rad(23.45) .* sin(deg2rad(360 * (284 + my_doy) ./ 365));
H      = deg2rad(15 .* (my_hourf - 12));
phi     = deg2rad(lat);
sin_h = sin(phi).*sin(delta) + cos(phi).*cos(delta).*cos(H);
sin_h(sin_h < 0) = 0;
G_cs = G_STC .* sin_h;
T.P_cs = eta_STC .* G_cs .* A_total ./ G_STC;
T.power_clipped = min(T.power, T.P_cs);

%% 4. 构建序列输入 (含 NWP 特征)
win = 96;
featureNames =
{'nwp_globalirrad','nwp_directirrad','nwp_temperature','nwp_humidity', ...
    'nwp_windspeed','nwp_winddirection','nwp_pressure'};
F = T{:,featureNames}; % NWP 特征
```

```
P = T.power_clipped;
t = T.date_time;

X = {}; Y = [];
for i = win+1:length(P)
    if all(~isnan(P(i-win:i))) && all(all(~isnan(F(i-win:i,:))))
        seq = [P(i-win:i-1), F(i-win+1:i,:)]'; % 功率 + NWP 滑窗
        X{end+1} = seq'; % [feature x timestep]
        Y(end+1,1) = P(i);
    end
end

%% 5. 训练集/验证集/测试集划分
ts = t(win+1:end);
mon = month(ts); dayN = day(ts);
is_test = false(size(ts));
for m = [2 5 8 11]
    idx = find(mon == m);
    last7 = idx(end-7*96+1:end);
    is_test(last7) = true;
end

X = X(~isnan(Y)); Y = Y(~isnan(Y)); % 清理 NaN

XTest = X(is_test); YTest = Y(is_test);
XTrain = X(~is_test); YTrain = Y(~is_test);

% 取训练集后 10%作为验证集
nTrain = numel(XTrain);
nVal = round(0.1 * nTrain);
XVal = XTrain(end-nVal+1:end); YVal = YTrain(end-nVal+1:end);
XTrain = XTrain(1:end-nVal); YTrain = YTrain(1:end-nVal);

%% 6. LSTM 模型定义
inputSize = size(XTrain{1},1);
layers = [
    sequenceInputLayer(inputSize)
    lstmLayer(64, 'OutputMode', 'last')
    fullyConnectedLayer(1)
    regressionLayer
];
opts = trainingOptions('adam', ...
```



```
'MaxEpochs', 100, ...
'MiniBatchSize', 128, ...
'Shuffle','every-epoch', ...
'Plots','training-progress', ...
'ValidationData',{XVal, YVal}, ...
'ValidationFrequency', 50, ...
'Verbose',false);

%% 7. 模型训练

net = trainNetwork(XTrain, YTrain, layers, opts);

%% 8. 测试集预测与误差
YPred = predict(net, XTest, 'MiniBatchSize', 1);
Pcs = T.P_cs(win+1:end);
Pcs = Pcs(is_test);
is_day = Pcs > 0;

nMAE = mean(abs(YPred(is_day) - YTest(is_day))) / P_rated;
nRMSE = sqrt(mean((YPred(is_day) - YTest(is_day)).^2)) / P_rated;
fprintf('nMAE = %.4f, nRMSE = %.4f\n', EMAE, ERMSE);

%% 9. 绘图
figure;
plot(YTest,'k'); hold on;
plot(YPred,'r--'); legend('实际','预测');
xlabel('样本'); ylabel('功率(kW)'); title('LSTM + NWP'); grid on;

%% === 10. 结果表格生成与保存 ===
ts_all = T.date_time(win+1:end);
ts_test = ts_all(is_test);

forecast_start_time = NaT(size(ts_test));
samples_per_week = 96 * 7;
for i = 1:samples_per_week:length(ts_test)
    i_end = min(i+samples_per_week-1, length(ts_test));
    forecast_start_time(i:i_end) = ts_test(i);
end

result_table = table;
result_table.StartTime      = forecast_start_time - days(1);
result_table.ForecastTime   = ts_test;
result_table.ActualPower_MW = YTest / 1000;
result_table.PredictedPower_MW = YPred / 1000;
```

```
result_table.StartTime = datestr(result_table.StartTime, 'yyyy/mm/dd
HH:MM:ss');
result_table.ForecastTime = datestr(result_table.ForecastTime, 'yyyy/mm/dd
HH:MM:ss');

writetable(result_table, 'lstm_prediction_result_with_nwp.csv');
fprintf('预测结果已保存至 lstm_prediction_result_with_nwp.csv\n');

%% === 11. 按全天空指数分类，计算不同天气下的 RMSE ===
% 提取对应测试集的实际功率、预测功率和理论功率
P_actual = YTest;
P_pred = YPred;
P_theory = Pcs;

% 计算全天空指数 Kt
Kt = P_actual ./ P_theory;
Kt(P_theory == 0) = NaN; % 避免除零

% 分类阈值
is_clear = Kt >= 0.7;
is_cloudy = (Kt >= 0.4) & (Kt < 0.7);
is_overcast = Kt < 0.4;

% 分别计算 RMSE（仅限于有效时间点）
rmse_clear = sqrt(mean((P_actual(is_clear) - P_pred(is_clear)).^2,
'omitnan'));
rmse_cloudy = sqrt(mean((P_actual(is_cloudy) - P_pred(is_cloudy)).^2,
'omitnan'));
rmse_overcast = sqrt(mean((P_actual(is_overcast) - P_pred(is_overcast)).^2,
'omitnan'));

% 打印结果
fprintf('\n 不同天气条件下的预测 RMSE（单位：kW）：\n');
fprintf('晴天（Kt ≥ 0.7）： %.4f kW\n', rmse_clear);
fprintf('多云（0.4 ≤ Kt < 0.7）： %.4f kW\n', rmse_cloudy);
fprintf('阴天（Kt < 0.4）： %.4f kW\n', rmse_overcast);
```



Company Profile

# 睿森科研简介



## 关于我们

睿森科研 深耕论文辅导领域5年  
为广大学子提供专业化、个性化的论文咨询服务

### 坚持初心，砥砺前行

我们始终秉持“授人以鱼不如授人以渔”的初心，为广大师生提供专业化、高水平的论文教育产品以及咨询服务。自19年以来，年均辅导学员人数达数千人，并呈现迅速上升趋势。



### 国内学术能力提升领导品牌，师资雄厚

提供会议论文辅导与发表、科研论文辅导与发表、硕博核心/S刊辅导、本硕博毕业论文辅导、以及各类大学生竞赛辅导等项目。我们的师资团队由2000余位专业论文咨询师组成。其中海内外高校博士及大学教授1000多人。



## 业务内容

科研论文、本硕博毕业论文辅导  
各类大学生竞赛辅导

### 科研论文，毕业论文辅导

我们提供SCI、SSCI、CSSCI、EI 源刊、中文核心、学报等科研论文辅导；本硕博毕业论文、课题辅导。已成功助力数千名学员拿到相应辅导的录用通知，因此保研、申博成果的学员不计其数。



### 大学生竞赛辅导

各类数学建模竞赛、数学竞赛、英语竞赛、互联网+、挑战杯、力学竞赛、创青春等大学生竞赛辅导，已成功助力数百名学员荣获国奖！



rui sen ke yan & shu mo jia you zhan  
安徽省合肥市高速中央广场A座

了解更多内容，请扫码咨询科研助理





资深师资团队

丰富教学经验

助你轻松拿奖

# 大学生创新创业大赛

## 精品辅导

互联网+ | 挑战杯 | 创青春 | 三创赛等

### 我们的优势

- 强大的师资力量
- 多对一全程服务
- 辅导前试听机制
- 无限次在线答疑
- 定制化课程内容
- 学员奖学金激励

### 课程内容

#### 01 项目诊断

根据不同的项目，结合各方面背景，提供项目改进意见和项目方向规划。

#### 02 参赛规划

依据学校、专业以及项目特点，制定参赛路线。

#### 03 商业计划书修改

提供针对性的书写指导，并在完成后逐页提供修改意见。

#### 04 PPT指导与修改

提供针对性的制作指导，并在完成后提供逐页提供修改意见。

#### 05 答辩指导与训练

对答辩进行训练，并提供针对性的指导意见。

#### 06 全程无限次答疑

比赛中遇见的各个问题，在辅导期间全程免费答疑。

### 辅导成绩

互联网+省银以上10余项

创青春省二以上10余项

三创赛国奖3项

扫码右侧二维码咨询报名&gt;&gt;&gt;



数模加油站

# 新学期 科研论文新规划

试听机制 合同保障 全科覆盖 实力师资

## 雏鹰计划

- 全过程辅导（到论文定稿）：  
高质量中文/英文期刊、EI/CPCI会议
- 辅导加发表一体化（到论文发表）：  
一对一：高质量中文/英文期刊、EI/CPCI会议  
双人团（两篇文章）：EI会议
- 时间周期：定稿2-4个月，录用1个月内，见刊2-6个月，检索1-3个月

## 卓研计划

- 全过程辅导（到论文定稿）：  
SCI、EI源刊、中文核心、学报
- 辅导加发表一体化（到论文发表）：  
一对一：SCI、EI源刊  
二人小班（共同完成一篇论文）：SCI、EI源刊  
三人小班（共同完成一篇论文）：SCI、EI源刊
- 时间周期：定稿3-6个月，录用2-8个月，见刊0.5-2个月，检索0.5-2个月

详情请扫描二维码咨询学术顾问



# 大学生学科类竞赛 保奖班

数学/英语/物理等

火热招生中

### 我们的优势

- 强大的师资力量
- 多对一全程服务
- 辅导前试听机制
- 无限次在线答疑
- 定制化课程内容
- 学员奖学金激励

### 课程大纲

- 基础知识讲解培训  
依据相关竞赛大纲，逐点讲解
- 竞赛考点难点分析  
针对竞赛难点，重点突破
- 真题选讲点评  
结合历年真题，精选例题详解
- 全真模拟练习  
竞赛全真模拟，赛后详细解析

# 大学生计算机类 竞赛保奖班

ACM/蓝桥杯等

国奖导师带你冲！！



### 我们的优势

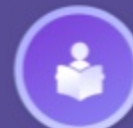
- 强大的师资力量
- 多对一全程服务
- 辅导前试听机制
- 无限次在线答疑
- 定制化课程内容
- 学员奖学金激励

### 课程设置



#### 定制学习方案

根据学员基础，定制个性化培训方案



#### 算法及编程基础培训

根据方案，开展基础培训



#### 刷题特训

导师精选题目，特训练习



#### 全真模拟练习

竞赛限时全真模拟，体验竞赛氛围

### 课程亮点



大牛授课  
干货十足



全程伴学  
无限答疑



绝密押题  
赛前助力

扫码立即报名&gt;&gt;&gt;&gt;&gt;&gt;&gt;





# 电工杯



## 转学术论文发表



### 服务内容



- 可转为EI会议/CPCI会议/高质量中英文期刊
- 免费提供论文方向评估及指导服务

### 含金量



- 发表一篇EI / CPCI会议堪比省一含金量
- 发表一篇高质量中/英文期刊堪比省奖含金量

### 我们承诺



- 收费透明，包含版面费，无任何二次收费
- 定金制，成功录用再补齐尾款，不录用全额退款

扫码添加学术顾问了解详情

