

## 风能和太阳能电厂的电量预测与调度策略

### Summary

随着全球对绿色能源需求的增长，风能和太阳能发电因其环保和成本效益而迅速发展。然而，这些能源的间歇性和不稳定性对电网稳定性构成挑战。通过问题的综合分析和预测模型，尝试提高风能和太阳能发电的可预测性和调度效率。

本文首先对风电场的风速和太阳能电厂的太阳辐射进行了详细分析，以预测电量的显著波动。通过描述性统计量、季节性分解和自相关函数图，尝试揭示了电力波动的模式，并运用变点检测技术预测显著波动。通过此方法不仅增强了对风能和太阳能发电波动性的理解，也为后续的预测模型提供了关键输入。

在预测模型方面，本文构建了季节性模型和长短期记忆网络（LSTM）模型，以实现短时间间隔内发电量变化的预测。季节性模型通过捕捉数据中的季节性模式，提供了一种基于历史数据的预测方法。尽管其预测结果较为集中，但它为理解数据的季节性变化提供了有价值的视角。LSTM模型则因其在处理长期依赖关系方面的优势而被采用，尽管由于数据量庞大，其结果尚未得出，但其在时间序列预测中的潜力不容忽视。

在备用电机调度策略方面，我们通过调整备用发电机的启动和关闭，以保持发电量的稳定，并使波动强度低于预设阈值的概率最大化。该策略不仅考虑了发电量显著波动的特征，还优化了发电机的运行数量，以适应发电量的变化。

本文的模型分析显示，LSTM模型在处理长期依赖问题方面具有显著优势，而季节性预测模型则能有效识别数据中的季节性波动。滑动窗口方法有助于捕捉时间序列的局部特征，但这些方法也存在计算复杂度高、对数据质量依赖性强以及窗口大小选择主观性等局限性。

针对风能和太阳能电厂的电量预测与调度策略提出了一套创新解决方案。综上所述，我们尝试提高了对风能和太阳能发电波动性的理解，而且通过预测模型和调度策略的创新，为提高绿色能源的利用效率和稳定性提供一些可行的解决方案。

**Key words:** 新能源，发电量预测，调度策略，波动模式，LSTM，季节性分解与预测，acf图

# Content

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Problem Background . . . . .	1
1.2	Restatement of the Problem . . . . .	1
1.3	Problem Analysis . . . . .	1
<b>2</b>	<b>Notations and Reasonable Assumptions</b>	<b>1</b>
2.1	Notations . . . . .	1
2.2	Reasonable Assumptions . . . . .	1
<b>3</b>	<b>Problem 1:预测显著波动的发生时间与阈值</b>	<b>2</b>
3.1	太阳能电厂电力波动模式分析 . . . . .	2
3.1.1	描述性统计量 . . . . .	2
3.1.2	季节性分解 . . . . .	2
3.1.3	季节性评判acf图 . . . . .	3
3.2	风力发电波动模式分析 . . . . .	3
3.2.1	描述性统计量 . . . . .	3
3.2.2	季节性分解 . . . . .	3
3.2.3	季节性评判acf图 . . . . .	4
3.2.4	风速和发电量的关系回归 . . . . .	4
3.3	变点检测 . . . . .	4
<b>4</b>	<b>Problem 2:发电机平均发电预测模型</b>	<b>5</b>
4.1	季节性模型 . . . . .	5
4.2	LSTM模型 . . . . .	5
<b>5</b>	<b>Problem 3:备用电机调度策略</b>	<b>5</b>
<b>6</b>	<b>Model Analysis</b>	<b>6</b>
6.1	Strengths and Weaknesses . . . . .	6
	<b>References</b>	<b>6</b>

# 1 Introduction

## 1.1 Problem Background

随着经济的发展，绿色和环保逐渐进入大众的视野，发电方式也随之带来改变。风电和太阳能发电逐渐开始迅速发展，这些能源不仅绿色环保，而且成本较低。然而，由于天气状况的不稳定，它们的电力输出非常不稳定，难以控制。为了促进风力发电的不稳定性带来的问题，需要通过预测和调节的方法来帮助公用事业和研究人员预测风速和太阳能辐射强度在短期和长期时间尺度上的变化。

## 1.2 Restatement of the Problem

问题一，需要分析风电场的风速和太阳能电厂的太阳辐射，然后，对于风速和太阳辐射强度进行预测，并且根据波动公式，求解合适的显著性阈值。

问题二，对于问题二，假设数据记录频率为1Hz，需要对于1-120秒内的发电进行单独的间隔预测。

问题三，发电站有一定比例的备用发电机，发电量少时启动，发电量增加时关闭，需要设计合理比例的备用机调度策略，能够在阈值较低的情况下，使波动强度低于阈值的概率尽量高，尽量保证发电量的稳定。

## 1.3 Problem Analysis

问题一，第一步，需要分别研究风电场和太阳能发电厂的电力波动模式。我们将波动模式分为四部分，分别从描述性统计特征，趋势，周期性，季节性进行分析，得出时间序列的特征；第二步，在总发电量显著下降和增加前进行预测，我们首先筛选出显著的变点，然后，提取出所有变点的前5分钟，然后寻找这些时间序列的共同特征，之后就可以通过这些特征来进行预测；第三步，需要计算波动幅度，当波动幅度超过阈值时，定义为显著，需求解尽量小的显著性阈值。

问题二的求解，利用时间序列预测的模型，导入历史数据，预测数据为发电量，进行预测即可。

问题三，需设计一个调度方案，通过启动和关闭备用发电机，尽量使波动强度低于制定的阈值 $t$ 的概率为 $r$ 。发电厂每天需要的总发电量一定，需要尽可能保持每天的发电量不变，在问题二中，假设发电设备个数一定，所以预测的平均每台发电量乘发电机个数即为总个数，此文中，应用第二问模型，预测每台发电机发电量之后，用总发电量除以每台发电量，即可得出当时需要的台数，再加之第一问的显著降低和升高时的特征，即可求解。

# 2 Notations and Reasonable Assumptions

## 2.1 Notations

## 2.2 Reasonable Assumptions

1. 对于 $k$ ，必须在当前时刻前30分钟都存在太阳辐射，此时 $k$ 才可以作为有效的波动幅度
2. 通过观察数据，每台发电机发电量基本相同，因而假设之后每台发电机发电量也相同

$p_{total}$	总发电量
$\bar{p}$	所有发电机平均发电量
$\hat{p}_i$	各发电机预测发电量
$q$	前30分钟平均发电量
$k$	波动幅度
$t$	估计阈值
$m$	正在工作的发电机总数
$\delta m$	需要调整的发电机数量
$M$	发电厂发电机总数

Table 1: Key parameters for Problem 1

### 3 Problem 1:预测显著波动的发生时间与阈值

#### 3.1 太阳能电厂电力波动模式分析

需要研究太阳能发电厂的电力波动模式。我们将波动模式分为四部分，分别从描述性统计特征，趋势，周期性，季节性进行分析，得出时间序列的特征。

##### 3.1.1 描述性统计量

第一步，观察数据，发现时间只有一半的描述，由于光照只有白天才有，夜晚不会出现，因而我们对数据进行了补全，把夜晚的太阳辐射变为0。

第二步，绘制了太阳能电厂的发电时间变化曲线：

对于曲线的初步观察，并未发现很明显的特征，整体处于振荡状态，以昼夜交替为规律。

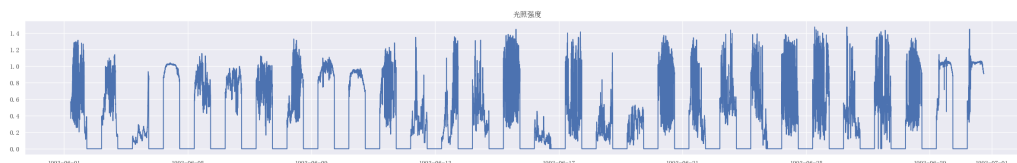


Figure 1: 太阳辐射强度变化图

##### 3.1.2 季节性分解

第三步，由于通过初步观察无法得出基本结果，尝试通过季节性分解进行分析。

季节性分解又称Holt-Winters模型，它将一个时间序列分解为趋势，周期性，残差三部分。当季节性变化在整个序列中大致恒定时，使用加法方法，季节性分量在观测序列的尺度上以绝对值表示，在水平方程中，通过减去季节性分量对序列进行季节性调整。在每年中，季节性成分的总和将大约为零。使用加法模型时，残差（模型未能解释的序列部分）往往更接近于白噪声，这表明模型已经很好地捕捉到了数据的季节性特征。绘制季节性分解图片如下：

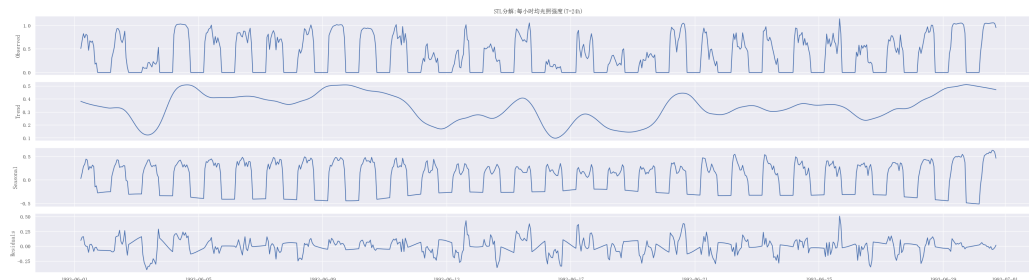


Figure 2: 太阳能发电季节性分解图

图片中，第一幅为太阳能原数据，第二幅为变化趋势，第三幅为周期性，第四幅为残差。可以观察到，太阳能在此月整体趋势大致保持平稳，但是存在明显的升降，变化较为频繁；周期性较为明显，以一天为周期，太阳在一天中东升西落，此结论也符合常理；通过观察残差，发现基本无明显规律，也基本符合正态分布，因而季节性分解效果较好。

### 3.1.3 季节性评判acf图

通过观察自相关图，可以帮助判断周期，太阳能发电绘制acf图如下：

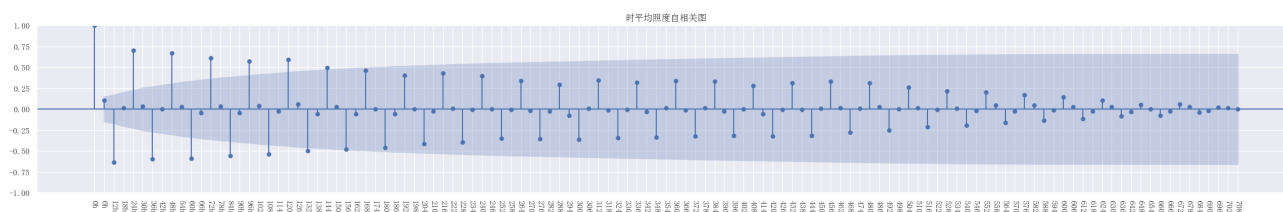


Figure 3: 太阳能发电acf图

观察acf自相关回归图，也可以明显看出，周期为24小时。

## 3.2 风力发电波动模式分析

需要研究风电场发电的电力波动模式。同样将波动模式分为四部分，分别从描述性统计特征，趋势，周期性，季节性进行分析，得出时间序列的特征。

### 3.2.1 描述性统计量

第一步，观察数据，发现时间中有不少缺失值，因为不可随意补全，所以暂且全部赋值为0。

第二步，绘制了风力发电厂的发电时间变化曲线：

对于曲线的初步观察，并未发现很明显的特征，整体大多发电量位于0.4-0.6之间，无明显规律。

### 3.2.2 季节性分解

第三步，由于通过初步观察无法得出基本结果，尝试通过季节性分解进行分析。绘制季节性分解图片如下：

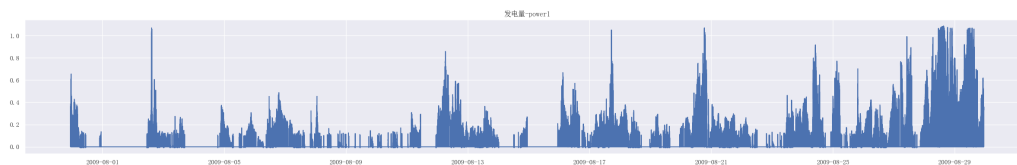


Figure 4: 风力发电电力变化图

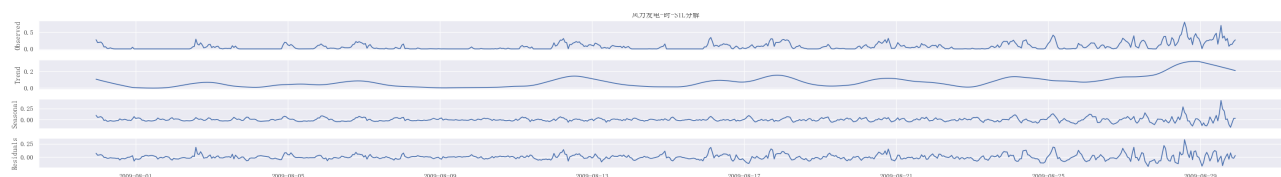


Figure 5: 风能发电季节性分解图

通过观察图片可以发现，周期分解最后才能勉强分解出周期。而且观察残差图，可以明显发现，残差在周期性固定后，存在一定的规律，所以季节性分解效果较差。

### 3.2.3 季节性评判acf图

通过观察自相关图，可以帮助判断周期，风能发电绘制acf图如下：

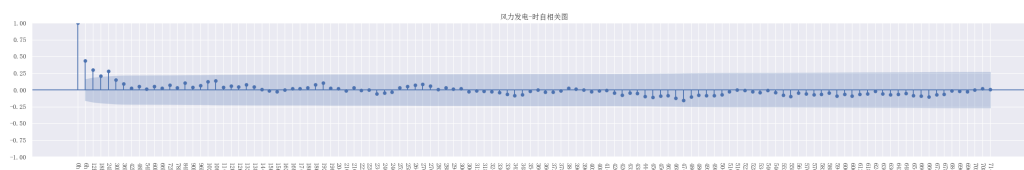


Figure 6: 风能发电acf图

观察acf自相关回归图，也可以看出，周期为24小时，其它基本为结尾，在置信度为95%的水平下，基本可以认为周期为24小时，经过分析，应该也和昼夜交替有一定关系。

### 3.2.4 风速和发电量的关系回归

由于拥有风速和电力图，我们需要对于风速和发电量进行回归，得出风速和发电量之间的关系，通过风速的变化，辅助电力发电的规律推断。

## 3.3 变点检测

在总发电量显著下降和增加前进行预测，首先使用变点检测[1]的滑动窗口方法筛选出显著的变点，然后，提取出所有变点的前5分钟，然后寻找这些时间序列的共同特征，之后就可以通过这些特征来进行预测。

尝试使用滑动窗口法进行异常点的查找，滑动窗口会基于数据前后均值，方差等统计量的变化进行评判，如果产生较大的变化，就会认为滑动窗口中出现了显著的变化点。接着，通过观察图片，可以找到一

些明显的变点。

## 4 Problem 2:发电机平均发电预测模型

### 4.1 季节性模型

季节性模型[2]是一种用于时间序列预测的统计模型，它特别适用于具有季节性模式的数据。这种模型通过结合自回归（AR）、差分（I）和移动平均（MA）的概念，以及季节性分析，来捕捉数据中周期性的季节性变化。我们通过季节性模型，预测得出未来一年的发电量数据。

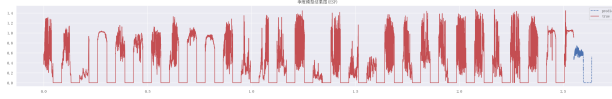


Figure 7: 太阳能发电预测图



Figure 8: 风能发电预测图

观察效果，发现结果过于集中，因为相当于取所有天的均值和每天每秒的均值之比作为季节性系数，最后乘每天均值，所以最终导致较为集中。接着，我们尝试用arima模型[3]尝试，发现效果和季节性预测类似，结果也不理想。

### 4.2 LSTM模型

我们使用LSTM[4]对于发电量进行预测。但由于数据量较大，暂时还未运算出准确结果。

## 5 Problem 3:备用电机调度策略

需设计一个调度方案，通过启动和关闭备用发电机，尽量使波动强度低于制定的阈值 $t$ 的概率为 $r$ 。发电厂每天需要的总发电量一定，需要尽可能保持每天的发电量不变，在问题二中，假设发电设备个数一定，所以预测的平均每台发电量乘发电机个数即为总个数，此文中，应用第二问模型，预测每台发电机发电量之后，用总发电量除以每台发电量，即可得出当时需要的台数，再加之第一问的显著降低和升高时的特征，即可求解。

假设 $p_{total}$ 数量恒定 $p_{total} = (m + \delta m)\bar{p}$ ，其中 $m$ 为正常运行发电机数量， $\delta m$ 为发电量显著增加或者娴熟减少时，减少或增加的发电机数量。 $\bar{p}$ 为各个发电机预测发电量 $\hat{p}_i$ 求和的均值。

$$\bar{p} = \frac{\sum_{i=1}^{m+\delta m} \hat{p}_i}{m + \delta m}$$

$$r = P(k|k < t) = P\left(\frac{|(m + \delta m)\bar{p} - q|}{q} < t\right)$$

$$\frac{|(m + \delta m)\bar{p} - q|}{q} \leq t \rightarrow (1 - t)\frac{q}{p} \leq m + \delta m \leq (1 + t)\frac{q}{p}$$

根据题目条件， $m + \delta m$ 的范围为 $[0, M]$ ，由此我们可将可能发生事件分为以下三种：

$$\begin{cases} \text{显著上升} & 0 < m + \delta m \leq (1+t)\frac{q}{p}, \\ \text{显著下降} & \begin{cases} \text{可调事件} & (1-t)\frac{q}{p} \leq m + \delta m \leq M, \\ \text{不可调事件} & m + \delta m < (1-t)\frac{q}{p}. \end{cases} \end{cases}$$

当发电量显著上升时对应的发电机数量为 $\lceil (1+t)\frac{q}{p} \rceil$ 。当发电量显著下降且为可调事件时，对应的发电机数量为 $\lceil (1-t)\frac{q}{p} \rceil$ ；不可调事件时发电机数量为 $M$ ，因此可调事件的比例为 $r = 1 - P(M < (1-t)\frac{q}{p})$ 。同时，我们分析了当 $[(1-t)\frac{q}{p}, (1+t)\frac{q}{p}]$ 中间不存在整数，即 $2t\frac{q}{p} < 1$ 是否存在，我们假定 $k$ 的取值范围在正常区间的条件是该时间点前30分钟都存在发电数据。在这一条件下，可求得 $K_{max} = 4.5$ 。由于 $k = \frac{|p-q|}{q}$ ， $0 < \frac{|p-q|}{q} < K_{max}$ ， $\frac{q}{p}(1 - K_{max}) < 1 < \frac{q}{p}(1 + K_{max})$ 。当 $(1 - K_{max})$ 大于0时，即 $K_{max} < 1$ 时，与计算出的条件相悖，可排除。

当 $(1 - K_{max})$ 小于0时，即 $K_{max} > 1$ 时， $\frac{q}{p} > \frac{1}{(1+K_{max})}$ ，由此可以推出 $\frac{2t}{p}(1 + K_{max}) < 2t\frac{q}{p} < 1$ ，最终可得 $t < 2.75$ 。

## 6 Model Analysis

### 6.1 Strengths and Weaknesses

#### Strengths:

1. LSTM:LSTM模型可以处理长期依赖问题，能够有效记忆信息超过长时间跨度，其门控机制使其能够根据不同的数据特点和任务需求，灵活地调整细胞状态中的信息保留与更新。
2. 季节性预测:通过引入季节性差分和季节性自回归等机制，可以识别并处理数据中的季节性波动，让预测结果更符合实际。
3. 滑动窗口:此方法可以捕捉时间序列中的局部特征，有助于更好地理解数据的短期动态计算局限于每个窗口内的数据，并大大降低了计算复杂度。能提供直观、易于理解的结果。

#### Weaknesses:

1. LSTM:由于LSTM的细胞结构和门控机制相对复杂，其计算复杂度较高。
2. 季节性预测:虽然能处理季节性数据，但如果数据不够“干净”——比如有缺失值，或者数据噪声太大，模型的效果就会大打折扣。并且，季节性预测需要额外调整季节性部分的参数，较为复杂。
3. 滑动平均窗口的大小需要人为确定，主观性较强，会导致缺少对不同大小窗口的对比分析。

## References

- [1] Samaneh Aminikhanghahi and Diane J Cook. “A survey of methods for time series change point detection”. In: *Knowledge and information systems* 51.2 (2017), pp. 339–367.



- 
- [2] Phillip G Gould et al. “Forecasting time series with multiple seasonal patterns”. In: European Journal of Operational Research 191.1 (2008), pp. 207–222.
  - [3] Siu Lau Ho and Min Xie. “The use of ARIMA models for reliability forecasting and analysis”. In: Computers & industrial engineering 35.1-2 (1998), pp. 213–216.
  - [4] Yong Yu et al. “A review of recurrent neural networks: LSTM cells and network architectures”. In: Neural computation 31.7 (2019), pp. 1235–1270.