**开题报告资料**

1. 选题背景和意义：（在论文里面摘抄摘抄）

太阳能发电的必要性：太阳能在全球范围内的开发不足，虽然最近几年势头很猛，而太阳可以被看作是一个巨大的天然核聚变反应堆，太阳能资源本身清洁环保，活动产生的废物不需要处理，不会造成任何环境威胁，基本不需要任何维护和燃料供应。另外，太阳将继续向地球表面发送稳定数量的太阳能。同时，由于经济和环境原因，迫切需要人类尽快发展可持续的无气体排放动力资源。事实上，化石燃料正在变得越来越稀少和昂贵。

南极昆仑站无人电厂的特殊性：南极昆仑站无人电厂要求配备400KW装机容量的太阳能发电系统，并在极昼晴空时实际输出达到40KW。考虑到昆仑站特殊的地理位置和环境条件，太阳能发电系统现有的基于经验的初步设计参数需进行优化设计。

预测的必要性：光伏电站发电功率受诸多天气因素的影响，环境因素影响而呈现出间歇性、随机性、波动性。大规模光伏发电介入电网会给系统的功率平衡和电网安全运行带来挑战。同时在南极存在极昼极夜的极端天气，对电网稳定性提出了更高的要求，只有能源可靠稳定才能保证南极科考站内的各项工作能够顺利展开。因此，对太阳辐照度的精准预测有利于电网调度决策，从而保障了科考站的安全和高效运行。

2、课题关键问题及难点：

（1）功率预测的关键就是对地表光伏的预测。（GHI DNI 等介绍）

（2）地表辐照度预测的关键在于不确定因素的耦合。外界辐照度可以根据太阳天平角等数据直接计算出来，传过大气层垂直照射到地表，在晴空的情况下，可以计算出DNI，一种类似抛物线的包络线。但由于天气因素的影响，例如云、雾霾等状况使得DNI曲线发生剧烈突变。如何在预测过程中，加入不确定因素以及不确定因素之间的耦合从而得到较为贴近实际的结果。

（3）不确定的环境因素使得地表辐照度呈现随机性间歇性，短期数据平稳性较差。

（4）过拟合，往往模型设计对数据集要求过高，导致在某些数据集上的表现很好，单着对预测并不是非常有利。

3、方案（设计方案、或研究方案、研制方案）论证：（填充自己的方案）

首先对昆仑站现有太阳能进行评估，更具体的是对GHI的评估。由于南极气候的特殊情况，所以排除极夜天气。接着对每天、每月、每季度、每年和近22年年平均分别进行分析，并且根据对太阳能资源评估的原则以及昆仑无人电厂对能源的需求整体分析，优化设计。

接着是对辐照度的预测，辐照度和温度直接影响到光伏电站的功率输出，所以更好的预测有助于保证电网工作的稳定平稳运行。辐照度预测一般分为四步走：（1）首先是对预测期的定义，因为预测期的不同决定着采用如何的时空策略。现有的预测一般以短期以及超短期为主，这次的目标是给出步长为5分钟的预测，更短更及时的预测有效的预防了潜在波动带来的负面影响。（2）第二步，充足的数据源，数据源的充足例如周年的温度、风速，GHI，DHI，TSI图，卫星云图等等，有效维度越多构造的模型越立体，越能抑制模型在单个参数方向的累计和突变，当然这要建立在模型合理的基础上。（3）第三步，数据预处理。包括数据的清洗，插值，相关性分析甚至降维等。关键参数的选择对于预测至关重要，冗余的参数不仅增加了计算的负担，同时将模型预测引导向了低效的甚至错误的方向上，这是不必要的。同时不仅仅有一维数据的预处理，图像数据的取噪，插值也是必不可少的，有时候可能成为关键指标上不去的原因。（4）将处理好的数据输入到模型当中去，并最终转化功率模型。模型的选择有好坏之分，模型的结构也对预测的结果至关重要，其中的关键在于模型与数据集特征的匹配，以及与数据集维度的契合。

4、现有方案的缺陷：

（1）大部分方案的缺陷关注点在于数据特征的缺失以及模型的特性，但具体都表现为预测的准确度不高或者是存在过拟合问题。

（2）对于数据特征的缺失，一般是经典概率模型、物理模型对云图信息问题考虑不足从而导致对于波动性预测精度不高的问题。当然即使引入了云图信息，也并没有完全解决预测精度不高的问题（虽然在精度上面有着永无止境的指标），原因又可能是对气溶胶等大气因素考量的缺失。

（3）对于模型特性，可以理解为模型的先验与数据集特征的不匹配。单一神经网络模型在非线性逼近上有着较好的表现，但是通常存在着过拟合的现象。经典时间序列预测模型，或者数据挖掘方法，又存在着模型简单，精度不高的问题。现在比较常用的就是对有不同特征的数据集采用不同的模型，并最终集成起来。如何选用不同模型，如何耦合不同模型又成为了新的难题。