论 文

基于神经网络模型的紫外极光卵边界建模

韩冰¹, 连慧芳^{1,2}, 胡泽骏^{2*}

1. 西安电子科技大学工程学院, 西安 710071;

2. 中国极地研究中心, 极地大气和空间物理学研究室, 上海 200136

* E-mail: huzejun@pric.org.cn

收稿日期: 2018-07-04; 接受日期: 2019-01-10; 网络版发表日期: 2019-04-19

国家自然科学基金(批准号: 41874195, 41831072, 41474146, 41431072, 61572384, 61432014)、中国博士后基金(编号: 2014M560752)、陕西省博士后科学基金(编号: JBG150225)、陕西省国际合作项目(编号: 2017KW-017)、中组部“国家高层次人才特殊支持计划青年拔尖人才”和国家海洋局极地考察专项(编号: CHINARE2017-02-03, CHINARE2017-04-01)资助

摘要 极光卵的尺度大小和太阳风、磁层、电离层以及它们间的耦合过程有密切的联系, 会随着空间和地磁环境的变化而变化. 建立准确的极光卵边界预测模型对空间天气的预报以及了解日地关系具有重要意义. 本文利用误差反向传播(back propagation, BP)神经网络和广义回归神经网络(general regression neural network, GRNN)两种神经网络模型对极光卵边界进行建模. 结果显示GRNN的极光卵边界模型具有较高的准确性, 赤道向边界预测平均绝对误差在0.77~1.20磁纬度(MLAT); 极向边界预测平均绝对误差在0.83~1.39 MLAT. 基于GRNN的极光卵边界模型预测准确性分别在极向边界和赤道向边界的整个磁地方时(MLT)上比BP神经网络的极光卵边界模型平均提高了0.74和0.73 MLAT, 比多元线性回归模型平均提高了0.82和0.82 MLAT. 而在模型的外推性方面, GRNN的极光卵边界模型的外推性优于BP神经网络的极光卵边界模型, 与多元线性回归模型接近.

关键词 极光卵边界建模, 行星际环境, BP神经网络, 广义回归神经网络

1 引言

极光是由源自太阳风的能量粒子, 沿磁力线沉降进入地球南北极地区, 与极区高空大气中的中性大气粒子碰撞激发产生. 极光粒子沉降主要出现在以磁极为中心的椭圆带状区域, 称之为极光卵^[1,2]. 极光卵的极向和赤道向边界是重要的地球物理参数, 和太阳风、磁层、电离层以及它们间的耦合过程有密切的联系, 并且它会随着空间和地磁环境的变化而变化^[3], 对其研究有助于进一步理解太阳风-磁层-电离层耦合过

程. 例如, 极光卵的极向边界所封闭的区域被称为极盖区. 极盖区是开放磁力线区域, 其面积的大小变化反映出太阳风-磁层-电离层能量耦合过程中磁能在磁层中的存储、释放过程^[4,5], 而极光卵本身的尺度大小与从太阳风和磁层注入到极区电离层的能量大小有关^[6,7]. 因此建立有效的极光卵边界预测模型对电离层等空间天气的研究以及太阳风、磁层、电离层间的耦合过程的研究具有重要意义^[8-10].

在极光卵的建模研究中, 最常用的是单变量分析法对极光卵边界进行建模, 例如用单个的地磁指数

引用格式: 韩冰, 连慧芳, 胡泽骏. 基于神经网络模型的紫外极光卵边界建模. 中国科学: 技术科学, 2019, 49: 531-542

Han B, Lian H F, Hu Z J. Modeling of ultraviolet auroral oval boundaries based on neural network technology (in Chinese). Sci Sin Tech, 2019, 49: 531-542, doi: 10.1360/N092018-00227

AE , AL , Kp 作为极光卵边界模型的输入参数^[11~13]。但是地磁活动和极光现象是空间环境变化引发的两种不同表现, 地磁指数只能从侧面反映极光现象, 不能从源头来解释极光现象^[14]。Shue等人^[15]研究了行星际磁场对极光亮度、极光半球能量的影响, 发现当行星际磁场三分量均为负值时, 极光半球能量最大。Hu等人^[16]和Yang等人^[17]研究得出行星际磁场三分量(B_x , B_y 和 B_z)对不同区域极光强度有不同的影响。Liou等人^[18]研究证明了不同的行星际条件会使极光卵的位置向不同的方向移动。Hu等人^[19]和Yang等人^[20]发现太阳风速度和太阳风密度增加会使日侧极光强度增加。Boudouridis等人^[21]通过事例分析说明了太阳风动压在不同的行星际磁场条件下会对极光卵的大小和亮度产生不同的影响。但是行星际磁场和太阳风参数只能反映这些参数与极光卵日侧边界部分相关的变化过程。Kauristie^[12]研究发现极光卵夜侧边界和表征亚暴磁暴活动强度的 AE 指数有密切联系。综上所述外部空间环境(如行星际磁场、太阳风参数和 AE 指数)是影响极光卵尺度大小和位置的主要因素。因此, Hu等人^[22]、杨秋菊等人^[23]直接利用行星际磁场和太阳风参数等多个参数, 采用多变量分析方法, 构造了极光卵边界的多元线性回归模型。该模型是将极光卵边界的磁纬度与所选模型的输入参数按照线性关系进行拟合。由于极光卵尺度大小与日地耦合过程中的诸多磁层动力学过程有关, 比如Siscoe和Huang^[24]研究认为极盖区的大小与磁层顶的日侧重联以及磁尾的夜侧重联有关。因此极光卵边界和空间参数间应该是复杂的非线性关系。然而目前我们还没有确切的关系式去描述这种非线性关系, 用线性拟合可以简化复杂问题, 但同时也会造成预测结果出现误差。

神经网络可以在未知数学模型内部结构的前提下, 通过随机设定初始参数并利用大量样本进行训练以更新网络内部参数, 以此获得输出对输入的非线性映射关系。由于神经网络强大的学习能力和非线性映射能力, 近年来神经网络被广泛用于预测模型的构建^[25~27], 其中应用较多的神经网络有误差反向传播(BP)神经网络^[28]、径向基函数神经网络(RBF)^[29]、广义回归神经网络(GRNN)^[30]等。

本文采用在预测模型中表现较好的BP神经网络和GRNN, 对行星际磁场三分量(IMF B_x , B_y 和 B_z)、太阳风速度(V_p)、太阳风密度(N_p)以及地磁指数(AE)为

输入参数的紫外极光卵边界进行建模, 并对神经网络构建的极光卵边界模型进行讨论总结。根据实验结果发现基于神经网络的非线性模型要比多元线性回归模型的预测准确率更高。这种通过神经网络对不确定的复杂非线性问题进行建模的方法也可以进一步应用到其他的空间物理相关研究中。

2 数据库建立及预处理

2.1 数据来源及简介

Polar卫星携带的紫外极光成像仪(Ultraviolet Imager, UVI)获得了大量的北半球的极光图像, 图像数据可以从网站Coordinated Data Analysis Web(CDAWeb, https://cdaweb.gsfc.nasa.gov/istp_public/)获取。为了避免日辉对极光图像的影响, 我们仅使用北半球冬季极夜时期(包括1996年12月, 1997年1月和12月, 1998年1月和12月以及1999年1月)获取的极光图像。本文仅使用了LBHL波段(~ 170 nm)的数据。所用的极光图像像素大小为 200×228 , 在普通观测模式下连续两帧LBHL图像之间的时间间隔在0.5~3 min之间。

本文所使用的行星际磁场、太阳风参数以及地磁指数数据来自于NASA OMNI数据库(<http://omniweb.gsfc.nasa.gov/>)。NASA OMNI数据提供了32个空间环境参数和地磁环境参数, 这里用到的是时间分辨率为1分钟的行星际磁场三分量 B_x , B_y 和 B_z , 太阳风速度 V_p , 太阳风密度 N_p , 地磁指数 AE 等6个参数。由于行星际磁场和太阳风等参数是在磁层顶前端测量得到的, 其影响到电离层的极光形态时需要一定传播时延。Weimer等人^[31]用最小方差法对行星际磁场的传播时延进行了研究, 目前传播时延一般选取10~15 min。这里将传播时延近似为经过磁鞘5 min和Alfven波传播到电离层的时间2 min之和, 即为7 min的时延。7 min的时延仅适合于日侧区域, 并不适用于其他磁地方时区域, 但目前对每个MLT区域没有合适的时延选择^[22], 并且定义不同的时延会增加模型的复杂度。因此本文近似都采用7 min的时延。 AE 指数是描述磁层亚暴强度即描述极光带电流强度的指数, 其作为模型输入参数时, 不需要延迟, 即选取和紫外极光图像同时刻的值。

2.2 极光卵边界提取及处理

利用文献^[32]对极光卵边界进行提取, 基本步骤

包括: 预处理、SFCM聚类、分割、去除小亮斑等. 图1给出极光卵边界提取各阶段的结果. 根据得到的极光卵区域可以分别提取极光卵的极向边界点和赤道向边界点.

利用文献[22]中边界点处理方法对得到的边界点数据进一步处理. 首先移除包含无效参数值的边界点, 然后根据磁地方时分别把得到极向边界和赤道向边界分成24组, 对每组边界点数据分别根据参数值(OMNI的6个参数及磁纬度)进行排序并去掉前5%和后5%这些统计无意义的的数据点, 最后去掉在磁子午线上5 min内磁纬度变化幅度超过5 MLAT的数据点. 经过上述预处理后共得到4849459个极向边界点, 1582569个赤道向边界点. 图2给出了边界点数在各个MLT上的分布情况. 图中菱形和矩形分别表示经过预处理后在对应的MLT上得到的极向边界点个数和赤道向边界点个数. 图中不同MLT上的边界点个数不同以及同一MLT上极向边界点个数和赤道向边界点个数不同的原因主要是由于用上述方法对提取到的边界点数据又进行了进一步处理导致的.

图1 (网络版彩图)极光卵边界提取各阶段结果

图2 (网络版彩图)边界点数在各个MLT上的分布

3 神经网络及极光卵边界建模

神经网络是对生物神经网络的近似和简化, 模仿生物神经元间的信息传递和记忆方式, 进行信息处理. 它是有由很多简单的处理单元(神经元)构成的, 这些神经元具有计算、记忆、学习等能力, 因此可以进行复杂的非线性运算, 通过对样本的学习不断调整神经网络中的连接权值和阈值等参数, 最终具备处理实际问题的能力[33]. 图3是神经元的结构图, $X_1, \dots, X_n$ 是神经元的输入, $w_1, \dots, w_n$ 是神经元输入的权重, f 是激活函数(传递函数), b 是阈值. 神经元输出和输入的关系可以表示为

$$y = f\left(\sum_{i=1}^n X_i \times w_i + b\right). \quad (1)$$

神经网络具有联想存储能力, 其中的反馈网络就是这种联想的具体体现. 它还具有快速寻找问题优化解的能力, 神经网络并行分布式系统, 可以发挥计算机高速计算的能力, 快速找到优化解. 最重要的是神经网络具有强大的自学习能力, 例如建立预测模型时, 先把训练样本输入到神经网络, 网络通过自学功能, 学习输入和输出间的对应关系, 然后根据学习到的知识近似得到网络输入对应的结果. 自学习能力对预测模型具有重要意义, 已经在经济预测、风险评估等领域有成功应用[34,35].

神经网络按照学习方式的不同可以分为有监督的神经网络和无监督的神经网络. 有监督的神经网络是利用训练样本的输出作为监督, 根据模型计算的结果和样本结果间的误差对网络的内部参数进行不断调整, 直到结果满足一定条件, 其中应用较多的有监督神经网络包括: BP神经网络和GRNN. 无监督神经网络没有训练样本的输出作为期望, 根据样本的内在规律

图3 神经元结构

调整网络的内部参数, 因此不适用于函数的拟合和逼近.

极光卵边界受太阳风-磁层-电离层耦合等一些磁层动力学过程影响, 因此它和行星际参数及地磁指数间可能是复杂的非线性关系. 而神经网络在处理输入和输出间复杂未知的非线性关系时可以表现出很好的性能. 基于此本文采用性能较好的两种神经网络(BP和GRNN)分别对极光卵进行边界建模. 本文中的两种神经网络建模方法都用6个参数(B_x , B_y , B_z , V_p , N_p , AE)作为模型的输入, 对极光卵的极向和赤道向边界进行建模, 输出是边界对应的磁纬度. 不同磁地方时内空间参数对极光卵边界的影响是不同的, 基于此分别对24个磁地方时时区上的极向和赤道向边界分别建模, 共建立48个预测模型.

3.1 误差反向传播神经网络

3.1.1 误差反向传播神经网络原理

误差反向传播(BP)神经网络是一种误差反向传输的前馈型神经网络, 它可以拟合输入和输出间的非线性关系. BP神经网络在分类模型和预测模型中都有广泛的应用. BP神经网络由输入层、隐含层(可以有一层或多层)和输出层组成. 图4是三层BP神经网络的基本结构.

BP神经网络的训练由信息的正向传播和误差的反向传播组成. 信息的正向传播, 训练集中的输入数据从输入层经隐含层单元逐层处理到输出层, 每个神经元的状态只受上一层神经元状态的影响. 如果在输出层得到的输出不满足预先设定的条件, 则进行误差的反向传播, 调整各个神经元的权值和阈值, 直至输出满足预先设定的条件^[36]. 训练过程可以表示为信息的正向传播和误差的反向传播的迭代过程, 直到网络的损失函数减小到预先设定的标准. 一般地, 均方误差

图4 BP神经网络结构

作为网络的损失函数如式(2)所示:

$$M_{SE} = \frac{1}{2} \sum_{i=1}^m (t_i - y_i)^2, \quad (2)$$

其中, M_{SE} 是网络的均方误差, i 是输出节点标号, m 是输出节点个数, t_i 是样本的输出值, y_i 是网络实际输出值.

3.1.2 基于BP神经网络的极光卵边界建模

BP神经网络可以通过训练学习两个数据间的不确定关系, 特别是非线性关系. 基于此本文首先用BP神经网络对极光卵边界进行建模. 根据经验, 本文使用三层BP神经网络, 其中, 输入层神经元个数等于输入向量的维数, 即为6, 隐含层神经元个数为 m , 输出层神经元个数为输出向量维数1. 图5给出了基于BP神经网络极光卵边界建模的流程图. 首先把极光卵边界数据分为训练数据和测试数据, 然后初始化BP神经网络中的参数, 用训练数据更新BP神经网络中的参数得到训练好的网络模型, 最后把测试数据输入到训练好的网络模型中得到预测的边界位置.

其中BP神经网络模型的预测性能受BP神经网络隐含层神经元个数以及初始化参数的影响. 这里对BP神经网络隐含层神经元个数以及初始化参数寻优. 用平均绝对误差(mean absolute deviation, MAD)^[23]作为模型预测的评价标准.

一般, BP神经网络的隐含层神经元个数^[37]:

$$m = \sqrt{n+l} + \alpha, \quad (3)$$

其中, m 是隐含层神经元个数, n 是输入层神经元个数, l 是输出层神经元个数, α 是[1, 10]间的常数.

3.2 广义回归神经网络

3.2.1 广义回归神经网络原理

广义回归神经网络(GRNN)是基于数理统计的一种径向基函数网络, 它的理论基础是非线性回归分析. GRNN具有较强的非线性映射能力和学习能力, 比RBF神经网络更有优势, 预测结果更好. GRNN由输入层、模式层、求和层和输出层组成, 图6给出其网络结构图^[38].

训练样本由输入层输入到GRNN, 输入层神经元个数等于输入向量的维数, 输入层神经元是简单的分布式分布单元, 直接将输入变量传递给模式层. 模式

图5 基于BP神经网络极光卵边界建模流程图

图6 广义回归神经网络结构

层神经元个数等于训练样本的个数。求和层中有算术和与加权两种类型的神经元。输出层神经元的个数等于输出向量的维数，各神经元的输出等于求和层中对应的加权神经元输出除以算术神经元的输出。

GRNN中权值会随着训练样本值确定而确定，因此只需要调节网络中的平滑因子，减小了调参的复杂度。

3.2.2 基于GRNN的极光卵边界建模

由于BP神经网络模型训练速度慢，需要调节的参数较多，且模型性能受网络初始化影响较大。考虑到GRNN较强的非线性映射能力和学习能力且易于调

参，用GRNN建立极光卵边界预测模型。图7是基于GRNN的极光卵边界建模流程图。其建模流程和上述基于BP神经网络的极光卵边界建模流程相似，即首先将数据分为训练数据和测试数据，并用训练数据训练GRNN，更新网络中参数，最后将测试数据输入到训练好的网络中得到预测的边界值。

GRNN相对于BP神经网络也存在不足之处。比如GRNN的模式层神经元个数会随着训练样本的增加而增加，其模型的复杂度会随着增加，测试的时间也会随之增加。但是BP神经网络隐含层神经元个数预先设置，其模型的复杂度和测试时间并不会随着训练样本数目的增加而增加。但对于GRNN的这种缺陷可以通过控制训练样本的数量得到改善。

4 实验结果与分析

为了验证神经网络模型预测极光卵边界的有效性，我们将上述两种神经网络模型与现有预测效果最好的多元线性回归模型^[22]进行对比实验。

4.1 网络参数选取

在BP神经网络的极光卵边界建模中，为了选取合适的隐含层神经元个数，本文进行了隐含层神经元个数与模型训练的时间复杂度对比实验以及隐含层神经

图7 基于GRNN极光卵边界建模流程图

图8 (网络版彩图)不同隐含层神经元数目与时间复杂度对比

元个数与预测模型准确率的对比实验。由于模型训练的时间复杂度以及模型预测的准确率受模型初始化参数影响较大,这里取3次实验的平均值。从图8中可以看出,总体上时间复杂度随着隐含层神经元个数增加而增加,但在隐含层神经元个数为4和8时时间复杂度会略微下降,造成这种略微震荡的原因主要是模型初始化参数随机性造成的。图9给出不同隐含层神经元个数对应模型在整个MLT范围内赤道向边界预测准确率。图中12条线分别表示隐含层神经元个数为1~12时模型在各个MLT上赤道向边界预测结果的平均绝对误差。从图中可以看出,模型预测的平均绝对误差随着隐

含层神经元个数增加而减小,但减少的幅度也会随着隐含层神经元个数增加而变小。由于增加隐含层神经元个数可以增加模型的中间信息,从而提高模型预测结果的准确率,但是在隐含层神经元个数大于一定值后,中间信息会出现冗余从而使模型结果改善并不明显。同时隐含层神经元个数的增加会导致模型时间复杂度的增加。因此,综合考虑模型准确率和模型时间复杂度两个方面,实验中设定隐含层神经元个数为4。

BP神经网络的初始化参数会影响预测结果的精确度,这里对每个MLT初始化了10个BP神经网络,根据预测结果的准确率选取了最优的初始化参数。

对于GRNN,当平滑因子太大时,测试样本的估计值会被近似为所有训练样本的平均值,出现欠拟合的情况,而当平滑因子趋于0时,测试样本会被近似为训练样本的值,从而出现过拟合的情况,所以选择合适的平滑因子才能训练出较为准确的预测模型^[39]。经过实验模型预测的平均绝对误差随着平滑因子的减小而减小,但是为了避免出现过拟合设定网络平滑因子为0.1。

4.2 主观评价结果

图10给出极坐标下3种模型的预测边界结果。图中红点表示由文献[32]中的SFCM聚类算法分割得到的边界点,黄色圆圈表示各个模型预测得到的边界点。第一列是基于多元线性回归模型的预测结果,第二列是

图 9 神经网络隐含层神经元数对准确率的影响

基于BP神经网络模型的预测结果, 第三列的是基于GRNN模型的预测结果. 三行分别表示不同时刻的测试样本. 多元线性回归模型对第二和第三个样本的预

测结果和真实结果比较接近, 但对第一个样本的预测结果和真实结果又有很明显的误差. BP神经网络模型对第一个样本的预测结果较为准确, 但在第二个样本的21 MLT位置以及第三个样本的21 MLT和6 MLT位置误差较大. GRNN模型对3个样本极光卵边界的预测都和真实边界较为接近, 特别是对于第一个样本21 MLT位置的特殊结构边界位置的预测要优于其他两个模型. 从主观评价可以看出本文中的GRNN模型预测得到的边界更接近极光卵的真实边界.

4.3 客观评价结果

本文采用五重交叉验证法和平均绝对误差^[22]对模型进行客观评价. 图11给出3种模型预测边界的平均绝对误差. 图中红虚线实心点和蓝虚线实心点分别表示多元线性回归模型对极向边界和赤道向边界预测的平均绝对误差, 红实线星号点和蓝实线星号点分别表

图 10 三种模型预测结果和SFCM分割结果对比

图 11 (网络版彩图)三种模型预测误差对比

示BP神经网络模型对极向边界和赤道向边界预测的平均绝对误差。红实线空心点和蓝实线空心点分别表示GRNN模型对极向边界和赤道向边界预测的平均绝对误差, 结果显示GRNN的极光卵边界模型性能最好, 其在赤道向边界预测的平均绝对误差为0.77~1.20 MLAT, 在极向边界预测的平均绝对误差为0.84~1.39 MLAT。在极向边界和赤道向边界的整个磁地方时上, GRNN极光卵边界模型预测的平均准确率, 比BP神经网络模型分别提高了0.74和0.73 MLAT, 比多元线性回归模型分别提高了0.82和0.82 MLAT。说明GRNN在极光卵边界建模有较好的性能。

从3种模型预测结果的平均绝对误差可以看出, 3种模型对赤道向边界预测结果的平均绝对误差都在日侧时达到最大值, 对极向边界预测结果的平均绝对误差均在夜侧时达到最大值。根据文献[19]中的分析, 日侧极光卵强度较弱并且还有日辉的影响, 这导致在用SFCM方法提取极光卵边界时出现误差进而导致日侧极光卵赤道向边界预测误差较大。由于亚暴影响在夜侧, 从而使得极光卵极向边界会有很多复杂结构, 对极光卵边界的提取和预测造成影响, 进一步使极光卵极向边界在夜侧预测误差增大。

4.4 极光卵边界模型的外推性

一般地, 外推性是指用行星际参数或地磁指数在一定范围内的训练数据训练得到的模型, 对训练数据范围外的数据预测结果的可靠性。为了验证模型的有效性和鲁棒性, 我们进一步对比了3种模型的外推性。这里用平均绝对误差作为模型外推性的度量标准。

外推性的实验设置: 分别针对 B_x , B_y , B_z , V_p , N_p , AE 这6个空间环境参数进行6组外推性实验。例如针对 B_x 参数的外推性实验设置如下: 数据集中所有 B_x 值的范围在 $-7.11 \sim 7.11$ nT区间内, 我们选择 B_x 在 $-2.78 \sim 2.78$ nT区间的对应数据作为训练集, 利用该训练集构建极光卵边界模型。利用 B_x 不在 $-2.78 \sim 2.78$ nT区间但在我们数据集中的对应数据作为测试集。图12是 B_x 参数训练集数据和测试集数据的分布情况。

图13给出了3种模型的外推性结果。其中图13(a)~(f)分别是3种模型的 B_x , B_y , B_z , V_p , N_p , AE 外推性对比。图中e-Linear和p-Linear分别表示多元线性回归模型在赤道向和极向预测边界的平均绝对误差; e-BP和p-BP分别表示BP神经网络模型在赤道向和极向预测边界的平均绝对误差; e-GRNN和p-GRNN分别表示GRNN模型在赤道向和极向预测边界的平均绝对误差。另外, 以上3个模型对行星际磁场三分量 B_x , B_y , B_z 的外推性的结果没有明显的差别。总体上, GRNN模型和非线性回归模型的外推性相近, 均优于BP神经网络模型的外推性。对于太阳风速度 V_p 、太阳风密度 N_p 、地磁指数 AE 的外推性, BP神经网络模型要比GRNN模型和非线性回归模型差很多。分析原因可能是因为 B_x , B_y , B_z 值的分布范围较小, 而 V_p , N_p , AE 值的分布范围较大, 而BP神经网络的推广能力有限, 它只能学习到训练样本范围内的特征。

5 结论

本文利用神经网络良好的预测性能, 构建了基于BP神经网络和GRNN的极光卵边界预测模型。实验证明, 两种神经网络模型的预测准确性均优于现有的多元线性回归模型的预测。主要原因是极光卵的边界位置受多种磁层动力学过程影响, 与空间参数是复杂未知的非线性关系。神经网络可以很好地学习到输入参数和输出参数间的非线性映射关系, 而多元线性回归模型将极光卵的边界位置和空间参数的关系简单化为线性映射关系, 这与实际情况不符合。

两种神经网络模型中基于GRNN的预测准确率优

图 12 B_x 外推性实验中训练集数据和测试集数据分布情况

图 13 (网络版彩图)三种模型的外推性

于基于BP神经网络的预测. 主要原因是BP神经网络内参数是随机初始化的, 并且BP神经网络在训练过程中又容易陷入局部最优, 而GRNN只需要调整网络的平滑因子, 因此更容易寻找到全局最优. 在外推性实验中, GRNN模型和多元线性回归模型的性能相似, 优于BP神经网络模型. 原因主要是BP神经网络的推广能力

有限, 它只能学习到训练样本范围内的特征.

综上所述, GRNN在极光卵边界建模中具有较好的性能, 能够较为准确地预测出极光卵边界的位置, 同时这种通过神经网络对不确定的复杂非线性问题进行建模的方法也可以进一步应用到其他的空间物理相关研究中.

参考文献

- 1 Akasofu S I. The latitudinal shift of the auroral belt. *J Atmos Terr Phys*, 1964, 26: 1167–1174
- 2 Feldstein Y I. Some problems concerning the morphology of auroras and magnetic disturbances at high latitudes. *Geomagn Aeron*, 1963, 3: 183–192
- 3 Germany G A, Parks G K, Ranganath H, et al. Analysis of auroral morphology: Substorm precursor and onset on January 10, 1997. *Geophys Res Lett*, 1998, 25: 3043–3046
- 4 Kamide Y, Kokubun S, Bargatze L F, et al. The size of the polar cap as an indicator of substorm energy. *Phys Chem Earth Part C-Sol Terr Planet Sci*, 1999, 24: 119–127
- 5 Milan S E, Provan G, Hubert B. Magnetic flux transport in the Dungey cycle: A survey of dayside and nightside reconnection rates. *J Geophys Res*, 2007, 112: A01209
- 6 Newell P T, Ruohoniemi J M, Meng C I. Maps of precipitation by source region, binned by IMF, with inertial convection streamlines. *J Geophys Res*, 2004, 109: A10206
- 7 Newell P T, Sotirelis T, Liou K, et al. Predictive ability of four auroral precipitation models as evaluated using Polar UVI global images. *Space Weather*, 2010, 8: S12004
- 8 胡泽骏, 杨惠根, 艾勇, 等. 日侧极光卵的可见光多波段观测特征——中国北极黄河站首次极光观测初步分析. *极地研究*, 2005, 17: 107–114
- 9 Hu Z J, Yang H, Liang J, et al. The 4-emission-core structure of dayside aurora oval observed by all-sky imager at 557.7 nm in Ny-Ålesund, Svalbard. *J Atmos Sol-Terr Phys*, 2010, 72: 638–642
- 10 Hu Z J, Yang H, Huang D, et al. Synoptic distribution of dayside aurora: Multiple-wavelength all-sky observation at Yellow River Station in Ny-Ålesund, Svalbard. *J Atmos Sol-Terr Phys*, 2009, 71: 794–804
- 11 Starkov G V. Mathematical model of the auroral boundaries. *Geomagn Aeron*, 1994, 34: 331–336
- 12 Kauristie K. Statistical fits for auroral oval boundaries during the substorm sequence. *J Geophys Res*, 1995, 100: 21885–21895
- 13 Carbary J F. A *Kp*-based model of auroral boundaries. *Space Weather*, 2005, 3: S10001
- 14 徐文耀. 太阳风-磁层-电离层耦合过程中的能量收支. *空间科学学报*, 2011, 31: 1–14
- 15 Shue J H, Newell P T, Liou K, et al. Interplanetary magnetic field B_z asymmetry effect on auroral brightness. *J Geophys Res*, 2002, 107: S1A 16-1–S1A 16-10
- 16 Hu Z J, Yang H G, Han D S, et al. Dayside auroral emissions controlled by IMF: A survey for dayside auroral excitation at 557.7 and 630.0 nm in Ny-Ålesund, Svalbard. *J Geophys Res*, 2012, 117: A02201
- 17 Yang Y F, Lu J Y, Wang J S, et al. Influence of interplanetary magnetic field and solar wind on auroral brightness in different regions. *J Geophys Res Space Phys*, 2013, 118: 209–217
- 18 Liou K, Newell P T, Sibeck D G, et al. Observation of IMF and seasonal effects in the location of auroral substorm onset. *J Geophys Res*, 2001, 106: 5799–5810
- 19 Hu Z J, Yang H G, Ebihara Y, et al. Surveys of 557.7/630.0 nm Dayside Auroral Emissions in Ny-Ålesund, Svalbard, and South Pole Station. In: Haaland S, Runov A, Forsyth C, eds. *Dawn-Dusk Asymmetries in Planetary Plasma Environments*, Geophysical Monograph 230. Geophysical Monograph Series. Hoboken: John Wiley & Sons, Inc., 2017
- 20 Yang Y F, Lu J Y, Wang J S, et al. Different response of dayside auroras to increases in solar wind dynamic pressure. *J Geophys Res*, 2011, 116: A08314
- 21 Boudouridis A, Zesta E, Lyons R, et al. Effect of solar wind pressure pulses on the size and strength of the auroral oval. *J Geophys Res*, 2003, 108: 8012
- 22 Hu Z J, Yang Q J, Liang J M, et al. Variation and modeling of ultraviolet auroral oval boundaries associated with interplanetary and geomagnetic parameters. *Space Weather*, 2017, 15: 606–622
- 23 杨秋菊, 胡泽骏, 韩德胜, 等. 基于行星际/太阳风和地磁条件的紫外极光卵边界建模和预测. *地球物理学报*, 2016, 59: 426–439
- 24 Siscoe G L, Huang T S. Polar cap inflation and deflation. *J Geophys Res*, 1985, 90: 543–547

- 25 LeCun Y, Bengio Y, Hinton G. Deep learning. [Nature](#), 2015, 521: 436–444
- 26 Li Q, Qin Y, Wang Z Y, et al. Prediction of urban rail transit sectional passenger flow based on elman neural network. [Appl Mech Mater](#), 2014, 505-506: 1023–1027
- 27 Wang J J, Hu S G, Zhan X T, et al. Predicting house price with a memristor-based artificial neural network. [IEEE Access](#), 2018, 6: 16523–16528
- 28 Artini M, Natoli C, Tinari N, et al. Elevated serum levels of 90K/MAC-2 BP predict unresponsiveness to α -interferon therapy in chronic HCV hepatitis patients. [J Hepatol](#), 1996, 25: 212–217
- 29 Melagraki G, Afantitis A, Sarimveis H, et al. A novel RBF neural network training methodology to predict toxicity to vibrio fischeri. [Mol Divers](#), 2006, 10: 213–221
- 30 Chen X, Wang N. Modeling a delayed coking process with GRNN and Double-Chain based DNA genetic algorithm. [Int J Chem Reactor Eng](#), 2010, 8: 47–54
- 31 Weimer D R, Ober D M, Maynard N C, et al. Predicting interplanetary magnetic field (IMF) propagation delay times using the minimum variance technique. [J Geophys Res](#), 2003, 108: 1026
- 32 王倩, 孟庆虎, 胡泽骏, 等. 紫外极光图像极光卵提取方法及其评估. [极地研究](#), 2011, 23: 168–177
- 33 Park D C, El-Sharkawi M A, Marks R J, et al. Electric load forecasting using an artificial neural network. [IEEE Trans Power Syst](#), 1991, 6: 442–449
- 34 Jang H, Lee J. An empirical study on modeling and prediction of bitcoin prices with Bayesian neural networks based on blockchain information. [IEEE Access](#), 2018, 6: 5427–5437
- 35 Ji G, Li F, Li Q, et al. A comparative study on Arrhenius-type constitutive model and artificial neural network model to predict high-temperature deformation behaviour in Aermet100 steel. [Mater Sci Eng-A](#), 2011, 528: 4774–4782
- 36 Li X Y, Cheng C T, Lin J Y. Bayesian probabilistic forecasting model based on BP ANN. [J Hydraul Eng](#), 2006, 37: 354–359
- 37 Niu D, Lv J, Liu H. Evaluating electric environmental issues using BP neural network with optimised hidden layer nodes. [IJGENVI](#), 2009, 9: 227–238
- 38 Polat O, Yildirim T. Genetic optimization of GRNN for pattern recognition without feature extraction. [Expert Syst Appl](#), 2008, 34: 2444–2448
- 39 Specht D F. A general regression neural network. [IEEE Trans Neural Netw](#), 1991, 2: 568–576

Modeling of ultraviolet auroral oval boundaries based on neural network technology

HAN Bing¹, LIAN HuiFang^{1,2} & HU ZeJun²

¹ School of Electronic Engineering, Xidian University, Xi'an 710071, China;

² Polar Atmosphere and Space Physics Laboratory, Polar Research Institute of China, Shanghai 200136, China

The size of auroral oval is closely related to the solar wind, magnetosphere, ionosphere and the coupling process between them. It is very important to establish an accurate boundary model for the space weather prediction and the understanding of the Solar-terrestrial interactions. In this paper, BP neural network and generalized regression neural network (GRNN) is used to construct auroral oval boundary model. The result shows that the auroral oval boundary predict model based on GRNN is more reliable, of which the mean absolute error in equatorial boundary is 0.77–1.20 magnetic latitude (MLAT), and the mean absolute error in poleward boundary is 0.83–1.39 MLAT. Compared with the result of the BP network model, the accuracy of the GRNN model has improved 0.74 and 0.73 MLAT in poleward boundary and equad boundary respectively. Compared with the result of the multiple linear regression model, the accuracy of the GRNN model has improved 0.82 and 0.82 MLAT in poleward boundary and equatoward boundary respectively. In terms of the extrapolation of the model, the extrapolation of GRNN model is better than that of BP model, and is close to the multiple linear regression model.

modeling of auroral oval boundary, interplanetary environment, BP neural network, generalized regression neural network

doi: [10.1360/N092018-00227](https://doi.org/10.1360/N092018-00227)