

文章编号: 2095-2163(2023)11-0185-07

中图分类号: TP311.13

文献标志码: A

基于 LSTM 和 BP 神经网络模型的净初级生产力预测

于明洋, 魏霖静

(甘肃农业大学 理学院, 兰州 730070)

摘要: 为研究在气候变化与人类活动影响下植物的净初级生产率(Net Primary Productivity, NPP)并加以预测,以利于研究气候变化和人类活动影响下植被净初级生产力。为此,首先收集 2001~2015 年的温度、降水、NDVI、NPP 数据,并对数据集进行预处理;进而构建起了基于时间顺序的 BP 神经网络预测模式和 LSTM 神经网络预测模式,并利用 2 种模式对 NPP 及其影响因素进行了分类与预测;最后,再将 2 种预测模式的实测结果加以比较。结果显示,BP 神经网络模型在估计 NPP 方面的偏差,较 LSTM 神经网络模型更少。

关键词: 温度;降水;NDVI;净初级生产力;BP 神经网络;LSTM 神经网络

Net Primary Productivity prediction based on LSTM and BP neural network models

YU Mingyang, WEI Linjing

(College of Science, Gansu Agricultural University, Lanzhou 730070, China)

Abstract: To study the Net Primary Productivity (NPP) of vegetation under the influence of climate change and human activities and to predict it, in order to facilitate the study of NPP of vegetation under the influence of climate change and human activities, the paper first collects the temperature, precipitation, NDVI, and NPP data from 2001 to 2015 and preprocesses the data set; then the paper constructs the BP neural network prediction model based on temporal order and the LSTM neural network prediction model, and uses the two models to classify and predict NPP and its influencing factors; finally, the paper compares the measured results of the two prediction models. The results show that the BP neural network model has less bias in estimating the NPP than the LSTM neural network model.

Key words: temperature; precipitation; NDVI; Net Primary Productivity; BP neural network; LSTM neural network

0 引言

由于工业技术的不断进步以及人类城市建设步伐的加快,人类的活动对于全球生态系统所造成的影响正日益增强,也随即导致了全球气候变暖、极端气象灾害频繁发生和海平面有所升高等一系列问题。作为人们生存和发展所依赖的系统,对气候变化这一大背景下的陆地生态系统变化进行研究已逐渐成为关注热点。植被是陆地生态系统的主要组成部分,在全球气候变化中起着非常重要作用。植被作为气候变化和人类活动最重要的影响要素之一,是人类生存离不开的物质和环境主体之一,也是链接土壤和承载水系的重要节点,因此关于植被如何

响应全球气候变化也是需要研究的核心问题。

植被生产力对人类自身的日常生活和经济发展产生着重要影响,同时也是各种人类活动中至关重要的基础资源。净初级总产量(Net Primary Productivity, NPP)是指从总产销量减去植物呼吸的需求量后剩下来的总量;也指特定时间内,经植株的组织形式或贮存产物的形态所表达而蓄积出来的有机物的总量。NPP 是指示植被环境运动的重要因素,体现了生态稳定 CO₂ 的功能。NPP 与异养循环速度的均衡(即净生态系统生产,NEP)则确定了是否有微生物环境中对过量大气 CO₂ 的积累,是大气环境中 CO₂ 含量季节变动的主要因素。通过正确地计量 NPP,可以认识全球碳循环的基本过程。另

基金项目: 国家自然科学基金(62103418);中国农业科学院农业信息研究所基本科研业务费(JBYW-AII-2022-08, JBYW-AII-2022-38)。

作者简介: 于明洋(1998-),女,硕士研究生,主要研究方向:智慧农业。

通讯作者: 魏霖静(1977-),女,博士,副教授,主要研究方向:农业信息技术。Email:1046082231@qq.com

收稿日期: 2022-11-17

外, *NPP* 也是对陆地生态系统中营养物质和能源平衡运转研究的重要基础。除供应了水生植物本身之外, 还为整个有机体生活提供了能源和物质基础, 因此对陆地 *NPP* 的研究也为人们合理开采、使用资源提供了依据。

植物不断进行的光合作用为生态系统的运行提供了大量的能量, 对于地球上的大多数生命来说这都是不可缺少的一部分。至今人们对于研究陆地植被净初级生产力的变化已经有很长的时间, 但是大规模的工作则于 20 世纪 60 年代开始^[1]。上世纪七十年代中, Lieth^[2] 在计算时逐步构建了首个覆盖全球范围的 *NPP* 回归模型, 并在此基础上模拟世界 *NPP* 的地理分布, 从而给出了首张全球 *NPP* 的分布图, 同时又通过大量研究证明了植被世界 *NPP* 是每年降雨量与年平均气温一致的地区函数^[3]。

植被净初级生产力是植物生态过程调控的关键指标, 在碳平衡与全球变化过程中起着关键性的影响作用^[4]。根据植被净初级生产力来确定生态系统的安全与否^[5], Rashid 等^[6] 指出, 植被 *NPP* 并不只是对地区生态演变进行评价的主要指标体系, 还在研究碳循环过程中贡献了关键的组成模块, 同时也与研究陆地生态系统固定碳能力有关的强弱有着关系。研究植被净初级生产力已经成为研究气候因素和人为因素对陆地生态系统的影响中必不可少的一项内容^[7-8]。近年来的研究表明, 中国国内的专家反演出对几十种 *NPP* 的模式。例如, 孙政国等学者^[9] 利用 BIOME-BGC 模型对中国南方不同草地类型的 *NPP* 进行了评估, 发现位于南部区域的低山丘陵草地、常见草山草坡以及高山草甸的 *NPP* 与净生态系统生产力 (NEP) 的变化有着差异。Li 等学者^[10] 研究了内蒙古草原植被生产力与 SPEI 指数之间的相关关系, 发现水分对内蒙古草原生产力的变化起主要作用。国内研究者也通过各种手段对全国及地方的净初级生产力水平展开了深入研究, 涉及 Theil-SenMedian 趋势分析、多重共线性检验、多元回归分析、残差研究和偏自相关分析等传统统计研究方法^[11]。目前研发了 CASA 模型^[12]、RBF 神经网络^[13]、Biome - BGC 模型^[14]、C - Fix 模型^[15]、MIAMI 模型^[16] 等, 这些模式都可以较为精确而详尽地说明植物发育的历程, 对 *NPP* 的评价也具有很大的可信度。然而, 这些模型无法很好地把握 *NPP* 随时间的变化特征, 缺少对于其时间维度方面的分析。因此, 需要一种捕捉其序列特征的方法来改进目前的算法, 以利于更好地对 *NPP* 进行预测。

近年来, 反传播式 (Back Propagation, BP) 神经网络算法和长短期记忆神经网络 (Long Short Memory, LSTM) 已经被应用于很多领域来解决分类和预测等问题, 也有很多研究者把 LSTM 应用于时间序列数据的预测, 并且得到相比之前方法更高的准确率^[17-19]。所以, 可以采用 LSTM 和 BP 的神经网络模式, 通过气温、雨量、归一化植被指数 (Normalized Difference Vegetation Index, NDVI) 作为影响因素对 *NPP* 进行预测。

1 模型原理和方法

1.1 BP 神经网络

BP 神经网络算法^[20] 是一个采用误差反传播方法进行的多层前馈网络系统, 是目前使用最普遍的一种神经网络模式。BP 网络能够用于了解并保存大规模的投入与输出模式的反映关联, 并且不要求学校预先公开关于这种反映关联的数学教育过程。其基本方法是选择最陡峭的下降法, 并采用反向方法来调控网络的权重值和阈值, 从而获得最低误差平方和。BP 神经网络方法已经应用在非线性建模、参数逼近、系统辨识等领域, 且有众多研究者对其进行了改进^[21-23]。

在实际问题中, BP 神经网络模型结构需由实验确定, 无规律可循。一般情况下的神经网络都先期设置好了上层的层数, 而 BP 网络系统也可能具有各种各样的隐层。理论上也已证实, 在不影响隐含节点数量的情形下, 对 2 层以上 (仅有某个隐层) 的 BP 网络, 能够进行随机本构非线性映射。在模拟样本相比较少的情形下, 较少的隐层节点也能够进行模拟样本空间的超平面分析。

一个通用的 BP 神经网络架构如图 1 所示。这里, 输入数据可以描述为 $\mathbf{X} = (x_1, x_2, \dots, x_n)^T$; 进入输入层后, 将其和权重向量 $\mathbf{W}_1 = (w_1, w_2, \dots, w_n)^T$ 进行位置乘积后, 再通过激活函数就可以获得下一级的输入向量; 隐藏层的层数 t 确定后, 每层的输入和输出都与输入层的操作一致, 但隐藏层的每一级神经元数量却可能和输入层有所不同; 因此, 最终将隐藏层的输出视为输出层的输入, 经过权重向量的乘积后可以得到输出数据 $\mathbf{Y} = (y_1, y_2, \dots, y_m)^T$ 。整体流程可以表示为:

(1) 输出层

$$y_j - f(\text{net}_j), j = 1, 2, \dots, m \quad (1)$$

$$\text{net}_j = \sum_{i=0}^m v_{ij}x_i, j = 1, 2, \dots, m \quad (2)$$

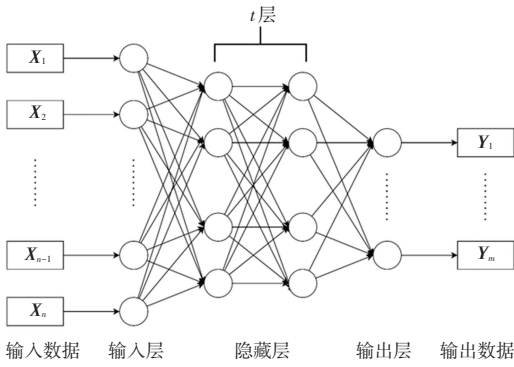


图 1 BP 神经网络结构图

Fig. 1 Neural network structure diagram

(2) 隐藏层

$$O_k = f(net_k), k = 1, 2, \dots, t \quad (3)$$

$$net_k = \sum_{j=0}^m w_{jk} Y_j, k = 1, 2, \dots, t \quad (4)$$

1.2 LSTM 神经网络

长短期记忆(LSTM)^[24]是一款已应用在人工智能和深度学习领域的人工神经网络。和标准的前馈神经网络有所不同,LSTM 有反馈接口。这样一来,循环神经网络(RNN)不但能够管理一个数据点(如图片),还能够管理一个数据序列(如语音或录像)。也因此,LSTM 已广泛应用在非分割、连接的手写辨识、语音识别、机械翻译、机器人操控、视频游戏、和医疗保健等任务中^[25-29],现已成为 20 世纪被引用最多的神经网络。

LSTM 的名字是指一个标准的 RNN 同时具有“长期记忆”和“短期记忆”的比喻。网络中的连接权重和偏差在每级训练中改变一次,类似于突触强度的生理变化如何存储长期记忆;网络中的激活模式在每个时间步中改变一次,类似于大脑中电击模式的瞬间变化如何存储短期记忆。LSTM 架构旨在为 RNN 提供一个可以持续数千时间步的短期记忆,因此是“长短期记忆”^[30]。

一个 LSTM 单元由一个单元、一个输入门、一个输出门,以及一个遗忘门所构成。单元可以在任意的时钟间隔内的数值,通过 3 个门调节一个进出单位的数量。LSTM 的基本结构如图 2 所显示,其整体流程可以表示为:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (5)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (6)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (7)$$

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (8)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (9)$$

$$h_t = o_t \odot \sigma_h(c_t) \quad (10)$$

其中, \$W_q, U_q\$ 表示输入门与递归连接的权重, \$q\$ 可以是输入门 \$i\$、输出门 \$o\$、遗忘门 \$f\$ 或记忆单元 \$c\$。

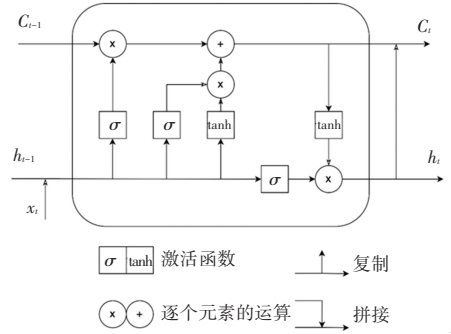


图 2 LSTM 结构图

Fig. 2 LSTM structure diagram

LSTM 系统适用于基于时间序列信息的分析、管理与估计,但在时间序列中的关键事件之间可能具有未知的滞后时间。原始的 RNN 系统在训练时,由于训练长度的增加和网络层数的增多,会产生梯度弥散爆炸或局部梯度消失的现象,使得系统无法处理更长序列信号,也因此无法得到更长距离数据的信号。针对这一状况,提出了后续的改进方案,即 LSTM 神经网络。LSTM 神经网络在深入学习后可以更好地处理序列化数据分析问题,其对于发电负载、流量等预测都有着不错的预知效果。LSTM 用于了解时间的中短期相关数据,因为神经网络中含有时钟记忆模块,所以适合于分析和检测时钟序列中的间隔和延迟事件^[31]。

1.3 激活函数

激活函数的引入主要是为了提高神经网络模型的不确定性,因为没有激活函数时每层就等于矩阵的相乘。任何一级输入/输出都是最上层的系统的线性函数,所以不论在神经网络哪一层,输入与输出都是系统的线性组合,也都是最本原的感知机。通过加入激活函数,为神经元导入非线性参数,神经网络就能够逼近一个非线性函数,而这种神经网络也能够运用在多种非线性模型上。本文采用的激活函数有 3 种: \$ReLU\$、\$Sigmoid(\sigma)\$ 以及 \$tanh\$。

1.3.1 \$ReLU\$

在人工神经网络的背景下, \$ReLU\$ (Rectified Linear Unit) 激活函数是一个定义为其参数的正部分的激活函数。\$ReLU\$ 函数的定义公式为:

$$f(x) = \max(0, x) \quad (11)$$

其中, \$x\$ 是神经元的输入。这又被称作斜率函数,类似于电气工程中的半波整流。

该激活函数从20世纪60年代末开始出现在分层神经网络的视觉特征提取中,并有专家认为该函数有着很强的生物学动机和数学理由。本文分析后发现,与在2011年以前普遍采用的激活函数比较,该激活函数可以更好地使用在更深的系统,包括 logistic sigmoid(受概率论启发,见 logistic 回归)以及要比 logistic sigmoid 更实际的对应函数方法,即双曲正切中。研究可知,该激活函数主要具有以下3个优点:

(1)稀疏的激活。在一个随机初始化的系统中,有大概50%的隐藏单元被触发(有一个非零输出)。

(2)更好的梯度传递。与双向饱和的正弦波激活函数一样,更少出现梯度消失问题^[7]。

(3)高效的计算。只有比较、加法和乘法。

1.3.2 Sigmoid

Sigmoid 函数是一个具有“S”形特征曲线的逻辑函数,其定义为:

$$S_{(x)} = \frac{1}{1 + e^{-x}} = \frac{e^x}{e^x + 1} = 1 - S(-x) \quad (12)$$

在某些领域,特别是在人工神经网络方面,“Sigmoid 函数”一词被用作 logistic 函数的别名。

Sigmoid 函数的其他情形,包括 Gompertz 曲线和 Ogee 曲线 Sigmoid 函数的值域也为实数,返回值(响应)通常是单调增长的,但也可能是下降的。Sigmoid 函数最常显示的返回值(y 轴)在0到1的范围内。另一个常用的范围是-1到1。

1.3.3 tanh

在几何上,双曲函数也是普通三角函数的近似物,但用双曲线并没有相应的基本概念。即如点 $(\cos t, \sin t)$ 构成单位半径的圆一样,点 $(\cosh t, \sinh t)$ 构成单位双曲线的右半部分。另外,与 $\sin(t)$ 和 $\cos(t)$ 的导数分别为 $\cos(t)$ 和 $-\sin(t)$ 相似, $\sinh(t)$ 和 $\cosh(t)$ 的导数分别为 $\cosh(t)$ 和 $+\sinh(t)$ 。

双曲函数出现在双曲几何的角度和距离的计算中。同时也出现在许多线性微分方程、三次方程和直角坐标的拉普拉斯方程的求解中。拉普拉斯方程在物理学的许多领域都很重要,包括电磁理论、热传递、流体动力学和狭义相对论。

tanh 是由基础的 sinh 与 cosh 计算得来,其定义式为:

$$\tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{e^{2x} - 1}{e^{2x} + 1} \quad (13)$$

2 实证分析

2.1 数据集来源与处理

文中选取的4个数据,即3个相关因子:气温、降水、NDVI,以及所要预报的NPP值。气温和降水数据均采用了兰州市气象局资料,NDVI采用的是MOD13C2,而NPP则采用的是北纬18°以北的土地生态系统,逐年净初级生产力1km的栅格数据集(1985-2015)^[32]。选取的时间为2001年1月1日到2015年12月1日。

在开展试验前,必须对样本进行预处理。此次试验,把数据集中按照8:2的标准分类成了训练集和测试集。将数据集中输入后,先对数值进行归一化处理,并测试是否出现了异常数值和缺失数据,上述步骤完成后才能进入下一步。本研究的样本集经检测均未产生异常值和缺失值,故不要求对其做幅值归一性处理。信息的归一化能够使各种信息的计量单位一致,而且能够增加模型的准确性。进行归一化分析的方法见式(5):

$$x = \frac{x_i - \bar{x}}{\max(x) - \min(x)} \quad (14)$$

其中, x_i 为第*i*个变量; \bar{x} 为 x_i 的均值; $\max(x)$, $\min(x)$ 分别表示 x_i 的最大值和最小值。

本文选取SSE、MSE、RMSE、MAE、MAPE、SMAPE作为2种神经网络预测模型性能的评价指标。其中,SSE、MSE、RMSE、MAE、MAPE、SMAPE的数学定义及表述详见如下。

(1)和方差SSE(the Sum of Squares Due to Error)。是观测值(Observed Values)与预测值(Predicted Values)的误差的平方和,定义公式为:

$$SSE(y, \hat{y}) = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (15)$$

(2)均方误差MSE(Mean Squared Error)。是观测值(Observed Values)与预测值(Predicted Values)的差值的平方和的平均数,即 SSE/n 。MSE是误差的二次矩,是估计量的方差(Variance)及其偏差(Bias),是衡量估计量质量的指标,定义公式为:

$$MSE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (16)$$

(3)均方根误差RMSE(Root Mean Squared Error),也称作RMSD(Root Mean Square Deviation),是MSE的算术平方根。由于每个误差(Each Error)对RMSD的影响与误差的平方(Squared Error)成正比,因此较大的误差会对RMSE影响过大,RMSE对异

常值很敏感。其公式为:

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (17)$$

(4) 平均绝对值误差 *MAE* (Mean Absolute Error)。是时间序列分析中预测误差常用的指标, 由于 *MAE* 使用的是与被测数据相同的尺度 (Scale), 因此不能用于比较 2 个不同尺度的序列。*MAE* 又被称为 *L1* 范数损失函数 (就是可以作为损失函数), 是真实数据与预测数据之差的绝对值的均值。定义公式为:

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (18)$$

(5) 平均绝对值百分比误差 *MAPE* (Mean Absolute Percentage Error), 也被称为 *MAPD* (Mean Absolute Percentage Deviation)。是一种衡量预测方法的预测准确性的指标。*MAPE* 在解释相对误差 (Relative Error) 方面非常直观, 在评价模型时 *MAPE* 通常用作回归 (Regression) 问题的损失函数 (Loss Function)。从定义中可以看出, 在计算 *MAPE* 时如果出现一系列特别小的分母, 可能会出现一些问题。比如分母为 0 的奇异点、较小的误差引起结果发生非常大的变化等。解决这个问题的替代方案是, 可以将公式中的实际值, 替换为该序列的所有实际值的平均值。这种方案等效于求绝对差的总和除以实际值的总和, 也被称为加权绝对百分比误差 (*WMAPE*), 或者 *wMAPE* (Weighted Mean Absolute Percentage Error)。定义公式为:

$$MAPE(y, \hat{y}) = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (19)$$

(6) 对称平均绝对百分比误差 *SMAPE* (Symmetric Mean Absolute Percentage Error)。实际值与预测值差值的绝对值除以实际值与预测值绝对值之和的一半。定义公式为:

$$SMAPE(y, \hat{y}) = \frac{100\%}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{\frac{|y_i| + |\hat{y}_i|}{2}} \quad (20)$$

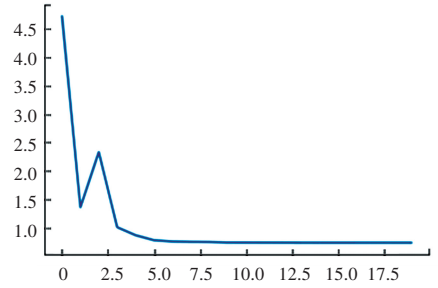
2.2 实验模型介绍

本文采用 BP 神经网络与 LSTM 对于 *NPP* 数据进行预测。模型统一输入 48 天的数据作为历史信息, 预测未来 1 天的 *NPP* 值。对于 BP 神经网络, 模型设计为输入层为 48 个神经元; 隐藏层为 1 层, 共计 4 个神经元; 输出层为 1 个神经元。BP 神经网络采用 *ReLU* 激活函数。对于 LSTM 网络, 模型设计为

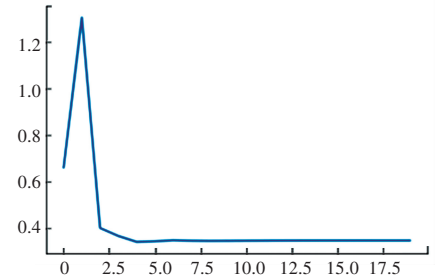
48 个神经元; LSTM 层选择 1 层, 输出层采用 48 到 1 的线性映射, 并采用 *sigmoid* 激活函数。LSTM 层中采用 *ReLU* 与 *tanh* 激活函数。

2.3 实验结果及分析

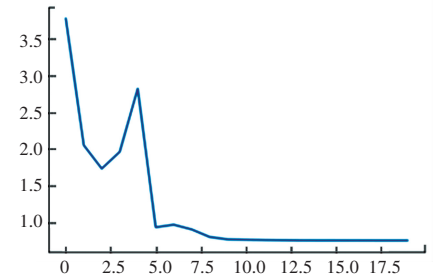
考虑到深度学习模型可以无限制地迭代, 就会希望在即将过拟合时, 或训练效果微乎其微时停止训练, 因此研究采用早停 (Early Stop) 策略防止模型过拟合。LSTM 和 BP 神经网络损失函数如图 3 所示。通过图 3 可以看出, LSTM 神经网络模型和 BP 神经网络模型的损失函数在迭代 20 次后达到最小值。



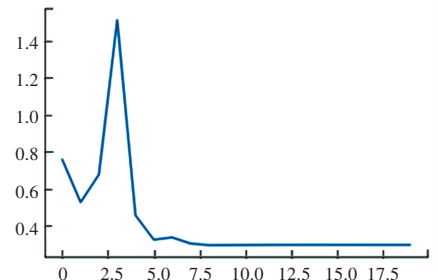
(a) LSTM 训练集损失函数



(b) LSTM 测试集损失函数



(c) BP 神经网络训练集损失函数

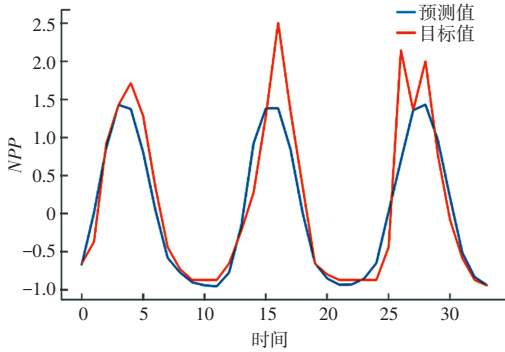


(d) BP 神经网络测试集损失函数

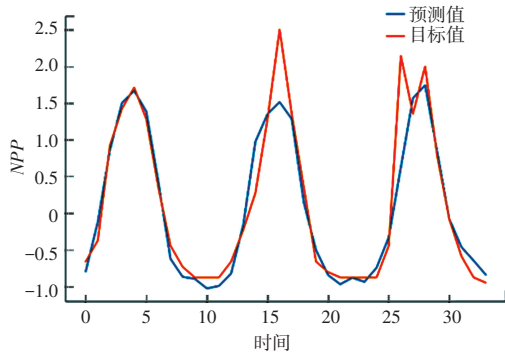
图 3 LSTM 和 BP 神经网络损失函数

Fig. 3 Loss functions of LSTM and BP neural network

LSTM 和 BP 神经网络在测试集上的预测值如图 4 所示。由图 4 可以发现, BP、LSTM 等神经网络预测模型的预测结果与 *NPP* 的实际走向情况大体相符, 误差也较小; 而且很明显地, BP 神经网络的拟合能力更佳。由图 4 还会发现 BP 神经网络预测模型的预测值与 *NPP* 实际数值最为符合的, 而且预测值与 *NPP* 实际数值之间的偏差也很小。



(a) LSTM 神经网络



(b) BP 神经网络

图 4 LSTM 和 BP 神经网络在测试集上的预测值

Fig. 4 Predicted values of LSTM and BP neural networks on the test set

BP 神经网络和 LSTM 神经网络评价指标结果见表 1。由表 1 可以看出, BP 神经网络模型所预测的 *SSE*、*MSE*、*RMSE*、*MAE*、*MAPE*、 *SMAPE* 等的数值, 显然都要比 LSTM 的神经网络模型的小, 大小依次 836.058、24.59、4.958 83、2.755 98、0.218 6、0.152 19, 说明了 BP 神经网络的预测结果比较精确, 并由此再次表明了 BP 神经网络对该数据的检测结果, 相比于 LSTM 神经网络要更佳。

表 1 BP 神经网络和 LSTM 神经网络评价指标

Tab. 1 Evaluation metrics of BP neural network and LSTM neural network

	<i>SSE</i>	<i>MSE</i>	<i>RMSE</i>	<i>MAE</i>	<i>MAPE</i>	<i>SMAPE</i>
LSTM	1 070.770	31.493 3	5.611 88	3.418 99	0.201 89	0.133 84
BP	836.058	24.590 0	4.958 83	2.755 98	0.218 60	0.152 19

3 结束语

全球气候变化逐渐影响到生态系统的平和稳定, 这一现象已日渐成为各学科进行学术研究的重要背景。通过研究温度、降水、*NDVI* 来预测兰州市 *NPP*, 能为全球气候变化及环境保护提供重要依据。本文采用 BP 算法和 LSTM 算法, 对 *NPP* 进行预测。通过对历史数据进行归一化处理, 再比较 5 种数据的预期结果, 从预测结果图的前 5 种预测指标来看, BP 神经网络的预期结果相比于 LSTM 神经网络更为出色, 其 *SSE* 能够达到 836 左右。未来相关研究可以考虑将基于统计学的时间序列预测方法与深度学习方法相结合, 从而更好地提取 *NPP* 序列数据中的相关特征, 并对其进行预测。

参考文献

- [1] 俞静芳. 基于 CASA 模型的浙江省植被净初级生产力估算[D]. 杭州: 浙江农林大学, 2012.
- [2] LIETH H. Primary production; Terrestrial ecosystems[J]. 1973, 1(4): 303-332.
- [3] LIETH H. Modeling the primary productivity of the world[J]. Primary Productivity of the Biosphere, Human Ecology, 1975, 13: 125-130.
- [4] 宗玮, 林文鹏, 周云轩, 等. 基于遥感的上海崇明东滩湿地典型植被净初级生产力估算[J]. 长江流域资源与环境, 2011, 20(11): 1355-1360.
- [5] 朱文泉, 陈云浩, 徐丹, 等. 陆地植被净初级生产力计算模型研究进展[J]. 生态学杂志, 2005, 24(3): 296-300.
- [6] RASHID R, ZHAO F, de JONG R, et al. Global and regional variability and change in terrestrial ecosystems net primary production and *NDVI*: A model-data comparison[J]. Remote Sensing, 2016, 8(3): 177.
- [7] 张凤英, 张增信, 田佳西, 等. 长江流域森林 *NPP* 模拟及其对气候变化的响应[J]. 南京林业大学学报(自然科学版), 2021, 45(1): 175-181.
- [8] ZHANG F, ZHANG Z, KONG R, et al. Changes in forest net primary productivity in the Yangtze River basin and its relationship with climate change and human activities[J]. Remote Sensing, 2019, 11(12): 1451.
- [9] 孙政国, 陈奕兆, 居为民, 等. 我国南方不同类型草地生产力及对气候变化的响应[J]. 长江流域资源与环境, 2015, 24(4): 609-616.
- [10] LI Chunlan, WALTER L F, YIN Jie, et al. Assessing vegetation response to multi-time-scale drought across inner mongolia plateau[J]. Journal of Cleaner Production, 2018, 179: 210-216.
- [11] 徐勇, 郑志威, 戴强玉, 等. 顾及时滞效应的西南地区植被 *NPP* 变化归因分析[J]. 农业工程学报, 2022, 38(9): 297-305.
- [12] 杨丹, 王晓峰. 黄土高原气候和人类活动对植被 *NPP* 变化的影响[J]. 干旱区研究, 2022, 39(2): 584-593.
- [13] 张黎明, 蔡琦, 宋梅村. 基于 RBF 神经网络的 *NPP* 运行状态趋势预测[J]. 原子能科学技术, 2013(11): 2103-2107.
- [14] 韩红珠. 黄土高原植被物候和净初级生产力(*NPP*)的关系及其对气候变化的响应[D]. 西安: 陕西师范大学, 2020.

- [15] 杨晶. 大沽河自然保护区红松与落叶松 NPP 变化趋势与其树轮年表间相关研究[D]. 哈尔滨: 哈尔滨师范大学, 2019.
- [16] 刘智勇, 张鑫, 周平. 广东省未来温度、降水及陆地生态系统 NPP 预测分析[J]. 广东林业科技, 2011, 27(1): 59-65.
- [17] KAREVAN Z, JAK S. Transductive LSTM for time - series prediction; An application to weather forecasting [J]. Neural Networks, 2020, 125: 1-9.
- [18] GERS F A, ECK D, SCHMIDHUBER J. Applying LSTM to time series predictable through time - window approaches [C]// Proceedings of ICANN2001 International Conference on Artificial Neural Networks. Berlin/Heidelberg: Springer, 2001: 669-676.
- [19] YADAV A, JHA C K, SHARAN A. Optimizing LSTM for time series prediction in Indian stock market [J]. Procedia Computer Science, 2020, 167: 2091-2100.
- [20] YALN M E, TUBA A, RAMAZAN Y. Artificial neural network models [EB/OL]. [2002-11-23]. <https://doi.org/10.1046/j.1469-0705.1999.13010017.x>.
- [21] 周永进. BP 网络的改进及其应用[D]. 南京: 南京信息工程大学, 2007.
- [22] 黄庆斌. BP 算法的改进及其应用研究[D]. 成都: 西南交通大学, 2010.
- [23] 刘天舒. BP 神经网络的改进研究及应用[D]. 哈尔滨: 东北农业大学, 2011.
- [24] HOCHREITER S, SCHMIDHUBER J. Long short-term memory [J]. Neural Computation, 1997, 9(8): 1735-1780.
- [25] SCHMIDHUBER J, BLOG A. The 2010s; our decade of deep learning/outlook on the 2020s [EB/OL]. [2020]. <http://people.idsia.ch/~juergen/2010s-our-decade-of-deep-learning.html>.
- [26] GRAVES A, LIWICKI M, FERNANDEZ S, et al. A novel connectionist system for unconstrained handwriting recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 31(5): 855-868.
- [27] MAYER H, GOMEZ F, WIERSTRA D, et al. A System for Robotic Heart Surgery that Learns to Tie Knots Using Recurrent Neural Networks [J]. Advanced Robotics, 2008, 22(13-14): 1521-1537.
- [28] LI Xiangang, WU Xihong. Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition [J]. arXiv preprint arXiv:1410.4281v2, 2015.
- [29] WU Yonghui, SCHUSTER M, CHEN Zhifeng, et al. Google's neural machine translation system: Bridging the gap between human and machine translation [J]. arXiv preprint arXiv:1609.08144, 2016.
- [30] ELMAN J. Finding structure in time - cognitive science [J]. Cognitive Science, 1990, 14(2): 179-211.
- [31] 陆泽楠, 商玉林. 基于 LSTM 神经网络模型的钢铁价格预测 [J]. 科技视界, 2017(13): 116-117.
- [32] 陈鹏飞. 北纬 18° 以北中国陆地生态系统逐月净初级生产力 1 公里栅格数据集 (1985-2015) [J]. 全球变化数据学报 (中英文), 2019(1): 16.