

# Prediction of Short-Term Heavy Fog Based on Neural Network Method

Weiyu Shi, Xia Yu, Yong Duan, Jianwei Huang

School of Information Science and Engineering, Shenyang University of Technology, Shenyang Liaoning  
Email: 1937546455@qq.com, 1643952@qq.com, duanyong0607@163.com, 845498873@qq.com

Received: May 6<sup>th</sup>, 2020; accepted: May 20<sup>th</sup>, 2020; published: May 28<sup>th</sup>, 2020

---

## Abstract

Machine learning has significantly made progress in the researches of strong convection weather, but there is little research on the prediction of short-term fog. For this problem, we adapt the model of neural network to solve the problem of short-term fog prediction. Firstly, dataset of different short-term time is established based on the meteorological definition of fog and non-fog weather. Then these data are inputted into neural network to train and develop the corresponding short-term predictive model, respectively. Meanwhile, we use some test data to test the predictive performance of the model. Eventually, we use the meteorological approach—TS Score to validate the predictive capability on neural network.

## Keywords

Short-Term Fog Prediction, Machine Learning, Neural Network

---

# 基于神经网络方法的短邻大雾预测

时玮域, 于霞, 段勇, 黄建伟

沈阳工业大学信息学院, 辽宁 沈阳  
Email: 1937546455@qq.com, 1643952@qq.com, duanyong0607@163.com, 845498873@qq.com

收稿日期: 2020年5月6日; 录用日期: 2020年5月20日; 发布日期: 2020年5月28日

---

## 摘要

机器学习已经在气象领域强对流天气的研究中取得了进步, 但是在雾短临预测的研究中较少。针对该问题, 本文利用神经网络方法建立模型实现对未来时刻大雾的预测。首先根据气象学对有雾, 无雾的定义建立不同短邻时刻样本集。然后, 利用神经网络方法对样本集进行训练建立不同短邻时刻雾

预测模型并通过测试数据集检验神经网络的预测能力。最后使用实时天气预报系统中常用的气象学方法 TS 评分来评估预测模型的性能。

## 关键词

短临预测, 机器学习, 神经网络

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

大雾是人类面临的严重的气象灾害之一, 通常伴随着低能见度情况, 严重影响了人们的出行以及国民经济。另外, 气象要素本身具有随机性以及不同地区具有不同的气象要素的特点, 这也导致了不同地区气象要素之间的差异性, 进而不同地区的出雾规律也大不相同。这影响了未来时刻是否出雾的判断。

现阶段对大雾预测的国内外研究多采用两种方式, 第一种采用数值法, 蒋大凯[1]等人利用数值法中的 MM5 模拟衍生算法(PP 算法)分析了近 10 年辽宁省区域性大雾天气的形成过程。NCAR-ANC (Auto Nowcaster)系统[2]也采用数值法对大雾进行分析, 但是由于大雾本身发生因素错综复杂, 并且产生较大的计算开销, 因此无法很好地统计出大雾的规律。第二种为利用卫星监测的方式, 包括 Hunt [3]、Ben-dix [4]、Eyre [5]、EllORD [6]等, 这类方法对卫星传递的数据进行波段分析, 实现对雾的检测和研究。但是由于受到夜间波段噪声干扰较大, 无法在夜间精准地对是否有雾进行判断。

目前, 一些学者将机器学习方法应用于气象领域, 如孙丽华[7]、修媛媛[8]、刘杰[9]分别利用机器学习方法对短时雷电预报、强对流、PM 2.5 浓度等进行预测和分析, 皆取得了较好的效果。但鲜有学者利用机器学习对大雾进行短临预测的研究。

基于此, 本文采用了机器学习中的神经网络方法对每一短临时刻样本进行训练, 得到不同短时邻近时刻雾预测训练模型, 实现对不同短时邻近时刻大雾的预测。在得到平均准确率及召回率的基础上, 利用气象领域预报标准 TS 评分验证神经网络模型的预测性能。

## 2. 大雾短临预测问题及神经网络方法

### 2.1. 大雾短临预测描述

短临预测即根据气象业务标准, 对未来 0.5 小时, 1 小时, 2 小时以及 3 小时的是否有雾情况进行预测。利用神经网络方法建立未来 0.5 小时, 1 小时, 2 小时, 3 小时大雾预测模型, 使其能够对未来时刻进行预测。

### 2.2. 神经网络基本理论

与雾预测相关的数据本身具有随机性, 并且随区域的不同而不同, 因此很难有一个准确的方法实现短临雾预测。而神经网络方法是基于数据的方法, 通过采用误差反向传播、正则化等方法在学习到数据所反映的规律的同时又使得学习的模型具有很好的泛化性, 因此可用于对大雾天气现象预测的研究中。

在研究中选择神经网络作为预测模型学习不同短时邻近时刻的数据样本建立预测模型。神经网络对输入的数据, 经过多层神经元得到预测结果, 再根据预测结果和实际数据建立目标函数。通过目标函数

进行反向误差传播以优化目标损失，得到训练模型。其训练方法最早由 Rosenblatt, Frank [10]提出，并不断的改进与优化，例如应用于图像识别的 CNN [11]，应用于序列分析的 RNN [12]等，其网络结构和训练方式使得模型对预测数据有泛化性。

神经网络的建立模型分为 2 个阶段，包括前向传递和反向误差更新。在前向传递时，每一层的每个神经元节点值是由连接上一层全部节点的权重值与上一层节点值的期望得到。公式(1)如下所示，其反映了各神经元的强度，式中连接权重值是由反向训练得到。

$$x_{l,p} = \sum_{i=1}^n w_{pi} x_{l-1,i} \quad (1)$$

式中  $x_{l,p}$  表示第  $l$  层第  $p$  个神经元节点值。 $w_{pi}$  表示与  $x_{l,p}$  相连接的第  $l-1$  层的第  $i$  个神经元的权重值， $x_{l-1,i}$  表示与  $x_{l,p}$  相连接的神经元， $n$  表示与第  $l$  层神经元  $x_{l,p}$  相连接的数量。

通过激活函数进行前向信息传递，并利用关于预测值与实际值间的损失函数作为目标函数，通过反向求导进行残差传递实现权重系数的更新，并建立神经网络训练模型。

神经网络的学习权重表示每个神经元在训练过程中对训练样本的学习能力。激活函数用于学习数据间的非线性关系，将线性关系进行非线性变化更有利于学习数据。根据训练学习误差进行反向传播从而得到误差的最小值，并学习到各神经元的学习权重。其最小值方法通过梯度求导直至达到最优目标函数。通过多轮前向传播与反向传播建立神经网络训练模型。

本文的基于短临大雾预测的神经网络模型的输出结果采用 Softmax 方法来判断是否有雾情况，在训练中选择自适应矩估计的神经网络参数梯度更新机制。通过以上组成，实现并建立不同短时邻近时刻大雾预测模型。本文采用自适应矩估计，作为神经网络训练短临大雾预测数据的反向传播更新参数的方法以实现短时邻近时刻雾的预测模型。

由于神经网络训练时步长为超参数，人为设定不同的超参数可能会影响模型的训练效果，又由于气象数据包含大量的气象要素，因此为了避免训练过程中步长对梯度的影响。选择自适应矩估计方法作为短临雾预测的参数更新机制，这样能够减少超参数带来的困难。另外，自适应矩阵估计通过对应梯度的偏方差二阶力矩以及平均一阶力矩进行更新，即考虑了当前梯度，又考虑了上一轮训练梯度情况，一阶力矩，二阶力矩分别如式(2)和式(3)所示。

$$m_t = 0.9m_{t-1} + 0.1Grad_t \quad (2)$$

$$v_t = 0.9v_{t-1} + 0.1Grad_t^2 \quad (3)$$

其中  $m_t$ ， $v_t$  表示  $t$  时刻的一阶力矩，二阶力矩， $Grad_t$  表示  $t$  时刻神经网络训练参数梯度向量。

### 3. 基于神经网络的短临大雾预测研究

短临大雾预测方法的研究包括对原始大雾数据建立不同短临时刻大雾预测样本，并根据样本对神经网络进行训练得出预测模型，最后通过测试数据及气象业务数据检验模型效果。

#### 3.1. 短临大雾预测样本的建立

在输入神经网络进行训练之前，需要对原始数据进行处理以得到不同短时临近时刻大雾数据集。本文在中国综合信息数据共享平台(CIMISS)中获取气象属性数据并进行数据预处理。在这个过程中，需要对下载的数据中相同站点的数据进行合并，并将同一站点特定时间间隔的气象要素进行特征合并，以得到不同时刻属性要素。将这些气象要素的不同时刻值进行运算以获得每个气象要素的变化量。由此构成不同时刻的气象要素，并作为属性数据共同对未来时刻有雾及无雾情况进行预测。根据上述处理得到的

每个短时临近雾预测的属性包括相对湿度、露点温度等 42 个雾预测属性。

在得到属性后需要标记每条数据在对应未来时刻是否有雾情况以方便神经网络的训练。这需要进行有无雾的判别并分别与短临时刻建立关系，即将得到的有雾或无雾时刻与短临时刻相减，最终得到的样本标记表示为该数据在未来时刻是否有雾情况。

对有雾及无雾的判别是根据专家知识，即对有雾样本取特定时间范围内满足以下条件：1) 能见度小于特定阈值；2) 相对湿度大于特定阈值。当条件 1)和 2)满足并达到指定的持续时间范围的要求，则将首次出现的时刻定义为有雾时刻。对无雾样本的判别同样对能见度进行阈值判定，并达到指定的持续时间的条件，取中间时刻夜间的数据作为无雾时刻，并进行标记。以上数据构成了用于神经网络训练的短时临近能见度预测样本。

### 3.2. 基于神经网络的大雾分类模型建立

将建立的不同短临时刻大雾数据集按照一定比例划分成训练样本、测试样本。并将训练样本输入多层神经网络中进行训练，训练过程中神经网络通过自适应矩估计的方式更新参数。由于模型需要输出有雾及无雾的概率情况，因此输出层后面增加 Softmax 层输出得到预测结果。

## 4. 实验结果及分析

论文利用 CIMISS 现有资料数据对已有的辽宁省 62 个站点数据进行处理和标注得到不同短临时间大雾样本集。实验用到的训练数据以及测试数据是由上述样本集按照一定比例划分而成。实验分别按照每一短时临近时刻的数据进行训练和测试。

利用上述划分的训练数据训练神经网络。神经网络中激活函数选择 sigmoid，更新参数选择自适应矩估计方法。在得到训练模型后选取测试数据进行预测，并根据预测结果并用召回率及平均准确率对模型性能进行评估。召回率及平均准确率分别如式(4)和(5)所示。

$$\text{Recall}_{-} \text{Rate} = \frac{N_a}{N_a + N_b} \quad (4)$$

其中  $N_a$  表示在实际有雾(或无雾样本)中对应之前时刻预测正确的数量， $N_b$  表示对应预测错误的样本数量。

$$\text{precision}_{-} = \frac{N_1 + N_2}{N_1 + N_2 + N_3 + N_4} \quad (5)$$

其中  $N_1$  表示预测正确的有雾样本个数。 $N_2$  表示预测正确的无雾样本个数。 $N_3$  表示实际有雾预测无雾的样本个数。 $N_4$  表示实际无雾预测有雾的样本个数。

由测试样本的预测结果得到的各短时临近神经网络有雾召回率、无雾召回率及平均准确率图 1 所示。图 1 中横轴表示不同的短临时刻，每个短时邻近时刻的柱形分别表示有雾召回率、无雾召回率和平均准确率。这表明研究的神经网络模型能够通过大量气象要素数据训练以取得较准确的不同短临时刻是否有雾的预测结果。

对上述训练的各短时邻近雾预测模型应用于 CIMISS 实时系统中，通过气象学评分对神经网络的预测表现进行评价。这里的评分是气象学中用到的 TS 评分方法。该方法能够衡量模型方法的预报能力，值为 0 到 100，值越大表示该方法预测能力越强。如式(6)所示。

$$\text{TS}_{-} \text{Score} = \frac{NA_k}{NA_k + NB_k + NC_k} \times 100\% \quad (6)$$

式中  $NA_k$  为预报正确站的次数,  $NB_k$  为空报站的次数(预测有雾实际为无雾),  $NC_k$  为漏报次数(实际有雾预测无雾)。

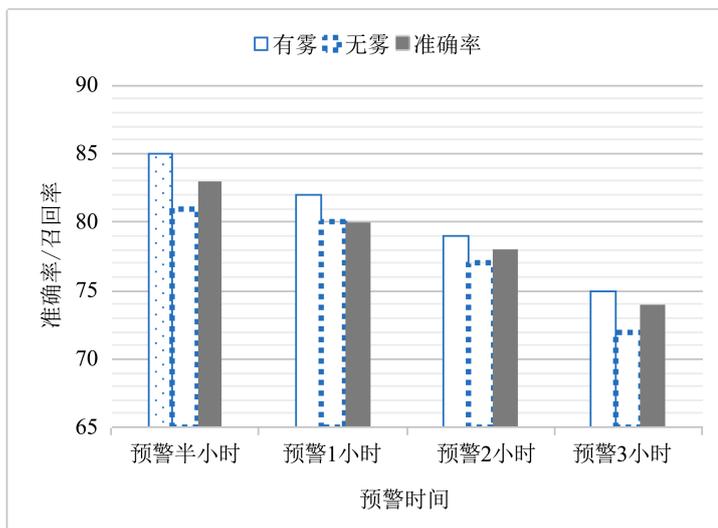


Figure 1. Comparison of various recall rates and average accuracy at different short-term and imminent times

图 1. 不同短临时刻各类召回率, 平均准确率比较

具体实验步骤如下。

1) 对 2019 年 1 月 12 日, 3 月 2 日, 3 月 3 日 CIMISS 实时系统获取的数据经过处理得到与训练样本和测试样本一致的气象要素, 并分别输入至上述的预测模型中得到每条数据在不同预测时刻是否有雾情况。

2) 根据气象学 TS 评分进行评估, 得出神经网络模型在各时刻预报性能。其 TS 评分结果如图 2、图 3、图 4 所示, 分别表示 2019 年 1 月 12 日, 3 月 2 日以及 3 月 3 日各时间段 TS 评分。

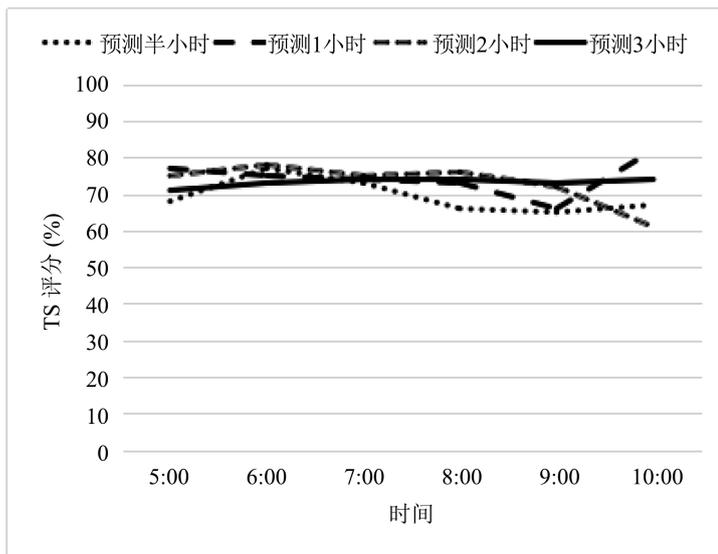


Figure 2. TS score changes at each time on January 12, 2019

图 2. 2019 年 1 月 12 日各时刻 TS 评分

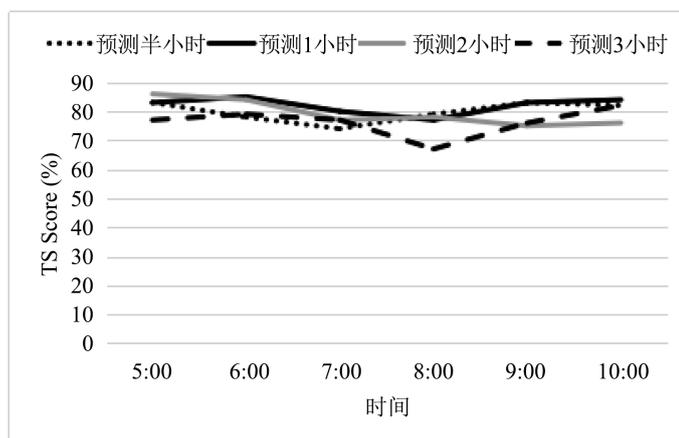


Figure 3. TS score changes at each time on March 2, 2019

图 3. 2019 年 3 月 2 日各时刻 TS 评分

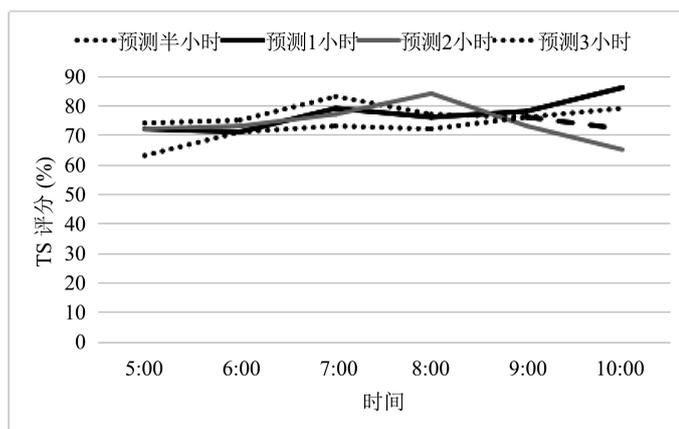


Figure 4. TS score changes at each time on March 3, 2019

图 4. 2019 年 3 月 3 日各时刻 TS 评分变化折线图

图 2、图 3、图 4 中横轴表示各日期对应的时间。不同折线分别表示在短临预测 0.5 小时，1 小时，2 小时，3 小时预测范围内对应的关于各时刻的 TS 评分变化范围。这里 TS 评分值由预测该时刻是否有雾情况和该时刻实际情况进行计算得到。

由图中折线 TS 评分值变化情况表明，随着短时邻近时间的增加，模型的 TS 评分表现在降低，体现的是越是邻近时刻预测越准确。实验结果说明本文研究的基于神经网络的短临大雾模型能够适应于实时气象数据的大雾预测。

## 5. 结论

本文采用机器学习的神经网络方法进行短临大雾预测研究，神经网络训练采用自适应矩估计的方式对不同短临时刻雾预测样本进行训练，建立神经网络模型。通过 TS 评分评价模型的性能，表明面对实时环境神经网络能够区分有雾，无雾气象要素数据。进一步表明神经网络模型训练并学习到辽宁省大雾预测的基本规律，并对不同短临时刻可以取得较好的大雾预测结果。

## 参考文献

- [1] 蒋大凯, 闵锦忠, 陈传雷, 等. 辽宁省区域性大雾预报研究[J]. 气象科学, 2007, 27(5): 578-583.

- 
- [2] Wilson, J.W., Crook, N.A., Muller, C.K., Sun, J. and Dixon, M. (1998) Nowcasting Thunderstorms: A Status Report. *Bulletin of the American Meteorological Society*, **79**, 2079-2099. [https://doi.org/10.1175/1520-0477\(1998\)079<2079:NTASR>2.0.CO;2](https://doi.org/10.1175/1520-0477(1998)079<2079:NTASR>2.0.CO;2)
- [3] Hunt, G.E. (1973) Radiative Properties of Terrestrial Clouds at Visible and Infrared Thermal Wavelengths. *Quarterly Journal of the Royal Meteorological Society*, **99**, 346-369. <https://doi.org/10.1002/qj.49709942013>
- [4] Bendix, J., Thies, B., Nauß, T. and Cermak, J. (2006) A Feasibility Study of Day-Time Fog and Low Stratus Detection with TERRA/AQUA-MODI Soverland. *Meteorological Applications*, **13**, 111-125. <https://doi.org/10.1017/S1350482706002180>
- [5] Eyre, J.R., Brownscombe, J.L. and Allam, R.J. (1984) Detection of Fog at Night Using Advanced Resolution Radiometer (AVH-RR) Imagery. *Meteorology, Magazine*, **113**, 266-271.
- [6] Ellrod, G.P. (1995) Advances in the Detection and Analysis of Fog at Night Using GOES Multispectral Infrared Imagery. *Weather and Forecasting*, **10**, 606-619. [https://doi.org/10.1175/1520-0434\(1995\)010<0606:AITDAA>2.0.CO;2](https://doi.org/10.1175/1520-0434(1995)010<0606:AITDAA>2.0.CO;2)
- [7] 孙丽华, 严军峰, 徐健锋. 基于多机器学习竞争策略的短时雷电预报[J]. 计算机应用, 2016, 36(9): 2555-2559.
- [8] 修媛媛, 韩雷, 冯海磊. 基于机器学习方法的强对流天气识别研究[J]. 电子设计工程, 2016, 24(9): 4-8.
- [9] 刘杰, 杨鹏, 吕文生, 等. 基于气象因素的 PM2.5 质量浓度预测模型[J]. 山东大学学报(工学版), 2015, 45(6): 76-83.
- [10] Rosenblatt, F. (1988) The Perception: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, **65**, 386-408. <https://doi.org/10.1037/h0042519>
- [11] LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998) Gradient-Based Learning Applied to Document Recognition. *Proceedings of the IEEE*, **86**, 2278-2324. <https://doi.org/10.1109/5.726791>
- [12] Hochreiter, S. and Schmidhuber, J. (1997) Long Short-Term Memory. *Neural Computation*, **9**, 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>