

1-4-2019

Application of Multiple-Stepwise and Kalman Filtering in Haze Forecast

Yunhao Xian

1.School of Computer and Software, Nanjing University of Information science and Technology, Nanjing 210044, China; ;

Hengde Zhang

2.Nation Meteorological Center of CMA, Beijing 100081, China; ;

Yonghua Xie

1.School of Computer and Software, Nanjing University of Information science and Technology, Nanjing 210044, China; ;3.Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information and Technology, Nanjing 210044, China;

Le Yang

1.School of Computer and Software, Nanjing University of Information science and Technology, Nanjing 210044, China; ;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research](#), [Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Application of Multiple-Stepwise and Kalman Filtering in Haze Forecast

Abstract

Abstract: Considering the importance of objective haze forecast and the low accuracy rate of haze forecast, a new haze objective forecast correction method based on the multiple stepwise regression algorithms and the Kalman filtering algorithm is proposed. The multiple stepwise regression method is used to control the physical factor of the dependent variable, and the visibility forecast equation is established. The Kalman filtering method is adopted to correct the regression coefficient in multivariate stepwise regression algorithm according to the actual data, and the haze objective forecast correction model is established. The experiments are carried out in Beijing, Guangzhou, Nanjing and Hangzhou. The experimental results show that comparing with the operational running fog - haze numerical prediction system (CUACE), the prediction accuracy of the multiple stepwise regression and Kalman filtering method is improved.

Keywords

multiple-stepwise regression, Kalman filtering, forecasting model, visibility, haze forecast

Recommended Citation

Xian Yunhao, Zhang Hengde, Xie Yonghua, Yang Le. Application of Multiple-Stepwise and Kalman Filtering in Haze Forecast[J]. Journal of System Simulation, 2018, 30(4): 1482-1489.

多元逐步回归与卡尔曼滤波法在霾预报中应用

咸云浩¹, 张恒德², 谢永华^{1,3}, 杨乐¹

(1.南京信息工程大学 计算机与软件学院, 南京 210044; 2.中国气象局 国家气象中心, 北京 100081;

3.南京信息工程大学 江苏省网络监控中心, 南京 210044)

摘要: 针对目前霾预报的重要性的客观预报准确率较低, 提出基于多元逐步回归算法和卡尔曼滤波算法的霾客观预报订正技术。利用多元逐步回归法控制因变量的物理因子, 建立能见度预报方程, 利用卡尔曼滤波法根据实况资料对多元逐步回归算法中回归系数进行订正, 建立霾客观预报订正模型。以北京站、广州站、南京站、杭州站四个站为例, 对站点进行预报实验和检验。实验结果表明, 与业务上运行的雾-霾数值预报系统(CUACE)进行对比, 提出的多元逐步回归与卡尔曼滤波法的预报准确率有所提高。

关键词: 多元逐步回归; 卡尔曼滤波; 预报模型; 能见度; 霾预报

中图分类号: TP391.9

文献标识码: A

文章编号: 1004-731X(2018)04-1482-08

DOI: 10.16182/j.issn1004731x.joss.201804033

Application of Multiple-Stepwise and Kalman Filtering in Haze Forecast

Xian Yunhao¹, Zhang Hengde², Xie Yonghua^{1,3}, Yang Le¹

(1.School of Computer and Software, Nanjing University of Information science and Technology, Nanjing 210044, China;

2.Nation Meteorological Center of CMA, Beijing 100081, China;

3.Jiangsu Engineering Center of Network Monitoring, Nanjing University of Information and Technology, Nanjing 210044, China)

Abstract: Considering the importance of objective haze forecast and the low accuracy rate of haze forecast, a new haze objective forecast correction method based on the multiple stepwise regression algorithms and the Kalman filtering algorithm is proposed. The multiple stepwise regression method is used to control the physical factor of the dependent variable, and the visibility forecast equation is established. The Kalman filtering method is adopted to correct the regression coefficient in multivariate stepwise regression algorithm according to the actual data, and the haze objective forecast correction model is established. The experiments are carried out in Beijing, Guangzhou, Nanjing and Hangzhou. The experimental results show that comparing with the operational running fog - haze numerical prediction system (CUACE), the prediction accuracy of the multiple stepwise regression and Kalman filtering method is improved.

Keywords: multiple-stepwise regression; Kalman filtering; forecasting model; visibility; haze forecast

引言

霾, 又称大气棕色云, 我国香港地区也称之为



收稿日期: 2016-05-12 修回日期: 2016-07-11;
基金项目: 国家自然科学基金(61375030), 科技部大气污染专项(JFY2016ZY01002213);
作者简介: 咸云浩(1991-), 男, 江苏淮安, 硕士, 研究方向为人工智能、机器学习、污染物预报。

烟霞, 是大量极细微的干尘颗粒(直径为 μm 量级)均为悬浮在大气中, 造成光散射, 致使大气浑浊, 水平能见度不足 10 km 的一种天气现象^[1]。近些年, 随着国家经济的快速发展和日益加快的城市化进程, 人类向大气中排放的污染物逐年增加, 这样造成全国很多城市出现不同程度的极端低能见度和重度空气污染的雾霾事件^[2], 主要分布在经济发达

和工业发达地区,如京津冀、长三角和珠三角地区。霾使大气能见度降低,损害人体健康,导致农作物减产,影响交通安全和气候变化等。霾中含有上百种大气细微颗粒物,其组成非常复杂,其中主要损害人体健康的是动力学直径小于 $10\ \mu\text{m}(\text{PM}_{10})$ 和动力学直径小于 $2.5\ \mu\text{m}(\text{PM}_{2.5})$ 的大气颗粒物,它能直接进入并粘附到上下呼吸道和肺泡中,导致包括呼吸系统症状增加、肺功能降低、慢性支气管炎发病增加、心血管疾病增加等。

现有的一些研究中,有很多都是关于霾的识别、霾的成因、霾与气象要素、与污染物浓度的关系等等,这些往往都是侧重于霾的成因分析、气象机理、污染特征、典型过程研究以及霾预警的发布标准等^[3-6]。我们关注到准确地做出霾预警或预报,对于更有效地防治乃至避免霾天气的危害,具有更重要的价值。因此近年来,对霾的预报预测系统的建立与发展已经成为大气环境领域的研究热点^[7]。

目前,主要有统计学方法,数值预报方法和天气学这三种方法用于城市霾客观预报。统计预报方法主要是通过建立污染物浓度与气象要素间的定量关系,分析发展规律进行预测的一种方法,国内外都有大量相关研究应用。目前常用的污染物预报统计模型主要有回归模型、神经网络模型和分类模型等^[8]。数值预报方法以大气动力学为基础,基于对大气物理和化学过程的理解,利用数学方法建立大气污染物浓度在空气中稀释扩散的模型,通过高性能计算来预报大气污染物浓度在空气中的动态变化^[9]。天气学预报订正方法是通过霾的形成条件作天气学分析和诊断,并且需要结合预报员的经验以及实况外推法。

本文将数值方法与统计学方法相结合,运用多元逐步回归方法对数值模式产品数据中因子进行筛选,筛选出最佳预报因子,建立预报模型,在此基础上利用卡尔曼滤波,对预报方程进行客观订正,从而得到更为准确的预报模型。经过实验论证,本文选用的多元逐步回归与卡尔曼滤波方法结果与国家气象局运行的数值预报系统(CUACE)结果

相比客观预报准确率有所提高。

1 多元逐步回归与卡尔曼滤波法

1.1 多元逐步回归

多元逐步回归分析的基本思路:对全部预报因子按照其对预报项的显著程度,从大到小依次逐个地引入回归方程。与此同时,在回归方程引入一个新因子后,要对回归方程的所有已有因子逐个检查,看其是否显著。如果不显著,就将该因子剔除,直到回归方程中所有预报因子对预报项的作用都是显著的,才考虑继续引入新的预报因子。再在剩下的未选中的预报因子中选取对预报项影响最显著的预报因子,检验其显著性,将作用显著的预报因子引入方程,剔除作用不显著的预报因子。直到最后再没有显著因子可以引入,同时也没有作用不显著的预报因子需要剔除为止。

多元逐步回归是研究预报量 \hat{y} 与 p 个因子 x 之间的定量统计关系:

$$\hat{y} = b_0 + b_1x_1 + b_2x_2 + \dots + b_px_p \quad (1)$$

上述回归方程中 \hat{y} 表示对 y 的估计, b_0, b_1, \dots, b_p 线性回归系数。写成向量形式如下

$$\begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \vdots & x_{np} \end{bmatrix} \times \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix} \quad (2)$$

$$\hat{y} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}, b = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_p \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & \dots & x_{1p} \\ 1 & x_{21} & \dots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \vdots & x_{np} \end{bmatrix} \quad (3)$$

根据最小二乘法的原理^[10],计算预报量观测值与多元线性回归预报量估计值的差值平方和 Q ,选出差值平方和最小的那组回归系数作为多元回归预报方程的线性回归系数。

$$Q = (y - \hat{y})^T (y - \hat{y}) = (y - Xb)^T (y - Xb) \quad (4)$$

$$Q = y^T y - b^T X^T y - y^T Xb + b^T X^T Xb \quad (5)$$

为此我们需要将 Q 分别对 b_0, b_1, \dots, b_p 求偏导数,令偏导数为 0。经过化简线性回归系数可以得

到 b_0, b_1, \dots, b_p 必须满足下面的标准方程组:

$$\begin{cases} nb_0 + b_1 \sum_{i=1}^n x_{i1} + \dots + b_p \sum_{i=1}^n x_{ip} = \sum_{i=1}^n y_i \\ b_0 \sum_{i=1}^n x_{i1} + b_1 \sum_{i=1}^n x_{i1}^2 + \dots + b_p \sum_{i=1}^n x_{i1} x_{ip} = \sum_{i=1}^n x_{i1} y_i \\ b_0 \sum_{i=1}^n x_{i2} + b_1 \sum_{i=1}^n x_{i2} x_{i1} + \dots + b_p \sum_{i=1}^n x_{i2} x_{ip} = \sum_{i=1}^n x_{i2} y_i \\ \dots \dots \dots \\ b_0 \sum_{i=1}^n x_{ip} + b_1 \sum_{i=1}^n x_{ip} x_{i1} + \dots + b_p \sum_{i=1}^n x_{ip}^2 = \sum_{i=1}^n x_{ip} y_i \end{cases} \quad (6)$$

按照一般的线性代数求解方程组的办法就可以解出 b_0, b_1, \dots, b_p , 得到回归方程。

在此基础上对回归方程进行显著性检验, 检验预报因子与预报量是否确有线性关系。这里使用 F 检验, 来检验回归方程的显著性。 F 检验的公式如下:

$$F = \frac{S_{y_z}^2 / p}{(1 - S_{y_z}^2) / (n - p - 1)} \quad (7)$$

式中: $S_{y_z}^2$ 为标准化变量回归方差, p 为预报因子的个数, n 为样本数量。在显著性水平 $\alpha=0.05$ 下, $F > F_\alpha$ 说明该回归方程是显著的, 否则说明回归方程不显著, 需要重新选取预报因子进行计算。

1.2 卡尔曼滤波

卡尔曼滤波是数学上的一种统计方法, 该方法是由 Kalman 在 1960 年提出来的。该方法采用量测方程和状态方程组成的线性随机系统的状态空间模型来描述滤波器, 并且利用状态方程的递推性, 按线性无偏最小均方差估计准则, 采用递推算法对滤波器的状态变量做出最佳估计, 从而求得滤掉噪声的有用信号的最佳估计^[11]。卡尔曼滤波一般用来处理: 滤波、预测和平滑这 3 方面的问题。本项目研究使用卡尔曼滤波主要解决的是滤波问题, 卡尔曼滤波器通过对模型参数的估计, 实现对观测序列的滤波效果。

天气预报领域中的卡尔曼滤波的基本公式如下:

$$Y_t = X_t \beta_t + v_t \quad (8)$$

$$\beta_t = \Phi_{t-1} \beta_{t-1} + w_{t-1} \quad (9)$$

其中, (8)式称为量测方程, (9)式称为状态方程。

式中: $\beta_t = [\beta_1, \beta_2, \beta_3, \dots, \beta_m]^T$ 为卡尔曼滤波系统状态向量, 也即回归系数; Φ 为状态转移矩阵; w 为输入噪声; Y_t 为测量向量即预报量, $Y_t = [Y_1, Y_2, Y_3, \dots, Y_m]^T$; X_t 为预报因子矩阵, 即:

$$X_t = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1m} \\ x_{21} & x_{22} & \dots & x_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \dots & x_{nm} \end{bmatrix}_t \quad (10)$$

v 为状态量测噪声。

本项目中考虑到具体气象应用中季节和气候等原因^[12], 假定 Φ 为单位矩阵, 经过简化后的卡尔曼滤波的状态方程为:

$$\beta_t = \beta_{t-1} + v_{t-1} \quad (11)$$

w 和 v 分别为输入噪声和状态量测噪声, 假定它们是两个互不相关的随机向量, 方差分别为 W 、 V 且均值为零的独立白噪声, 然后应用数学上的最小二乘法, 即可得到卡尔曼滤波的递推方程组, 即:

$$\hat{Y}_t = X_t \hat{\beta}_{t-1} \quad (12)$$

$$R_t = C_{t-1} + W \quad (13)$$

$$\sigma_t = X_t R_t X_t^T + V \quad (14)$$

$$A_t = R_t X_t^T \sigma_t^{-1} \quad (15)$$

$$\hat{\beta}_t = \hat{\beta}_{t-1} + A_t (Y_t - \hat{Y}_t) \quad (16)$$

$$C_t = R_t - A_t \sigma_t A_t^T \quad (17)$$

(12)~(17)组成的递推系统中, \hat{Y}_t 为 t 时刻的预报值, X_t 为预报因子, $\hat{\beta}_{t-1}$ 为回归系数估计值, R_t 为 $\hat{\beta}_{t-1}$ 外推值得误差方差阵, C_{t-1} 为 $\hat{\beta}_{t-1}$ 滤波值得误差方差阵, σ_t 是预报误差方差阵, A_t 为增益矩阵, Y_t 为 t 时刻的观测值即能见度实况值。

卡尔曼滤波订正方法是通过在不断的递推过程中, 不断更新实测信息, 采用更新实测的数据对状态估计进行修正, 从而对多元回归霾预报产品进行订正, 提高霾预报准确率。

1.3 多元逐步回归与卡尔曼滤波预报订正模型建立

本方案中各个物理量资料是应用欧洲中期天

气预报中心(ECMWF)细网格数值预报产品和常规气象资料,采用双线性插值法,将 EC 格点数据转化成站点数据作为预报因子。

本方案首先使用多元逐步回归方法选取最显著的预报因子,得出最优能见度预报方程,然后结合使用卡尔曼滤波订正方法建立能见度客观预报订正方程,最后在排除降水、沙尘暴、扬沙、浮尘、吹雪、雪暴的影响视程的天气现象,结合相对湿度利用概率回归得出是否有霾。具体步骤如下:

1、首先选取 4 个具有典型的代表站(北京、广东、南京、杭州),根据对典型的霾天气的分析,利用国家气象局观测实况、EC 数值预报模式选取能够反映动力、水汽、稳定性和能量等 8 个物理量要素(T、Rh、u、v、Div、omg、Td、T-Td)作为预报因子。

2、然后利用 2012 年 10 月到 2013 年 2 月欧洲中期天气预报中心(ECMWF)的气象常规资料,基于多元回归算法,逐步选取最优化预报因子,建立能见度预报方程。首先分别计算预报因子之间及它们与预报量之间的相关系数,建立出发的增广矩阵。样本容量 131,待选因子 $p=16$ 。出发矩阵为 R ,开始逐步回归计算:①计算各因子的方差贡献,作引入的统计检验。选取方差贡献最大的因子 v_1 000 计算该因子的显著性 $F=4.852$ 2,大于在显著性水平 $\alpha=0.05$ 下的 F_α ,因此引入方程并且消去该因子在出发矩阵中所占的列。②计算余下个因子的方差贡献,选取方差贡献最大的因子 Td,同样做 F 检验, $F=9.084$ 4 $> F_\alpha$,同样引入方程。③重复第二步计算过程。④计算已引入方程因子的方差贡献,将方差贡献最小的作剔除 F 检验,经计算 $F=8.442$ 8 $> F_\alpha$,所以不作剔除操作。⑤重复上面步骤③步骤④,直到没有可引入因子和没有可剔除因子为止。建立能见度预报方程。

$$y = -1.5737 - 1.1629x_5 - 51.0931x_{10} + 49.1673x_{11} + 1.0042x_{13} + 0.7114x_{15} - 50.409x_{16} \quad (18)$$

根据 2014-10-21~2015-02-28 共 131 天

ECMWF 模式数据,利用多元回归能见度预报方程得到这 131 天的能见度预报值。

3、接着利用 2013 年 10 月到 2014 年 2 月的 ECMWF 模式的气象资料,将多元回归方程作为卡尔曼滤波的初始量测方程,然后根据 2013 年 10 月至 2014 年 2 月的 ECMWF 模式数据确定 W_0 、 V_0 、 C_0 、 β_0 。

由上文可知,本文使用多元逐步回归建立能见度最有预报方程,所以我们就将多元逐步回归方程中的回归系数作为卡尔曼滤波过程中回归系数的初始值。即:

$$\beta_0 = [-1.1629, -51.0931, 49.1673, 1.0042, 0.7114, -50.409]$$

由于 β_0 是从样本资料中通过严格的数学计算得到的,因此理论上认为它是十分精确的,即与理论值相等,所以 C_0 是 m 阶零方阵,即 $C_0 = [0]_{m \times m}$ 。

有实验证明输入噪声和状态量测噪声 W_0 、 V_0 的初值对回归系数 β 的订正并不敏感,因此将 W_0 、 V_0 都设为零方阵。

最后根据卡尔曼滤波递推公式推算出新的能见度客观预报方程。卡尔曼滤波订正能见度预报方程的过程及误差订正的订正过程设计成自动批处理程序。

4、最后根据霾天气发生时能见度、相对湿度、降水等气象要素之间的关系利用概率回归建立霾预报方程,从而得到霾的预报模型,判断是否有霾。实验方案如图 1 所示,最后再利用 2014 年 10 月至 2015 年 2 月的观测实况进行预报检验。

选用气象上衡量预报准确率的 TS 评分算法,对基于多元回归和卡尔曼滤波的霾预报结果进行评分。TS 评分的定义如下:

$$TS = \frac{t_1}{t_1 + t_2 + t_3} \times 100\% \quad (19)$$

式中: t_1 为正确预报霾的个数, t_2 为漏报霾的个数, t_3 为空报霾的个数。由定义看出, TS 越大,预报的准确率就越高,预报的效果就越好。研究步骤见图 1。

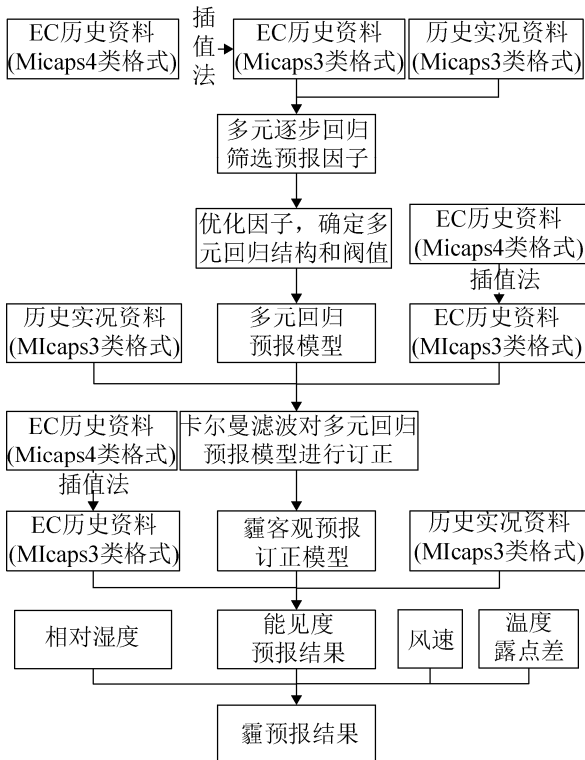


图1 实验基本步骤

Fig. 1 The basic steps of the experiment

2 实验结果

2.1 实验平台和资料

本文应用欧洲中期天气预报中心 (ECMWF) 细网格数值预报产品和常规气象资料, 具有较高的预报准确率, 已经业务化应用多年, 模式预报时效 240 h, 时间间隔为 3 h, 模式的水平分辨率为 $0.25^\circ \times 0.25^\circ$ 的经纬网格, 垂直分辨率 60 层, 模式顶 0.1 百帕。表 1 中实验数据来自雾霾数值预报业务系统 CUACE, 该系统是中国气象局自主研发的环境气象数值预报模式系统, 并且耦合了 60 余种气体成分和六类七种气溶胶组分别形成了对 PM、O₃ 和能见度的数值模拟和预报能力。

本文中应用多元回归与卡尔曼滤波算法进行预报的实验平台为 Inter(R)Core(TM) i7 -4500Cpu@1.80 GHz, 内存 32.00GB。MATLAB (R2010b) 作为编译工具, 应用 Matlab 作为编程语言编写多元回归算法和卡尔曼滤波算法, 版本为 7.11.0。

2.2 实验及结果分析

2.2.1 本文实验结果分析

在确定了多元回归能见度客观预报方程和卡尔曼滤波初始值后, 做北京、广州、南京、杭州能见度预报实验。试验日期选择 2014-10-21~2015-02-28 共计 131 天的时间, 下面对实验结果进行分析:

(1) 从上面各站的 2014-10-21~2015-02-28 共计 131 天的能见度预报值与实况的波形基本上一致, 预报结果能较好地反映能见度的变化情况。

(2) 由表 2 将基于多元逐步回归的能见度预报结果与经过卡尔曼滤波后的结果进行对比发现, 经卡尔曼滤波后的结果与能见度实况拟合情况更好。

(3) 在建立多元逐步回归和卡尔曼滤波方程时可以发现, 如果直接使用能见度来预报霾的发生情况, 则将会失去很多的因子变量对霾预报的贡献, 直接导致霾的预报效果不理想。

(4) 根据表 3 中主要预报因子与霾天气的对应统计关系, 将结合能见度, 风速, 相对湿度, 温度露点差利用概率回归算法建立霾预报模型:

$$pro = \sum_{i=1}^n \left(\frac{1}{n} * x_i * y_i \% \right) \quad (20)$$

即可得到霾发生时的概率。其中 x_i 为选取的物理因子的数据, y_i 为对应选取的物理因子在霾天气发生时, 在选定阈值中所占的比例。

(5) 检验结果: 利用 ECMWF 模式预报因子数据和霾预报模型对 2014 年 10 月 21 日至 2015 年 2 月 28 日共计 131 天进行预报试验, 来检验霾预报模型的预报准确率。利用 14 时实况进行检验, 由表 4 与图 2、3、4、5 能见度拟合情况可以看出, 北京, 南京、杭州预报值与实况拟合情况较好, 而广州拟合情况偏大。通过表 5 可以发现预报指标对于北京、南京、杭州空报次数较多, 漏报次数较少, 而广州空报次数较少, 漏报次数相对于北京、南京、杭州较多。平均准确预报率为 58%, 说明该客观预报方法具有一定的预报能力。

表 1 北京站部分实验数据
Tab. 1 Experimental Data of Beijing Station

	1	2	3	4	5	6	7	8	9	10	11	12	13
Div	1.0	-3.1	-12.2	4.0	10.8	-1.7	0.2	-4.8	0.4	-9.9	-6.0	-1.7	-5.9
Omg	-1.5	-0.1	-0.5	0.9	1.4	1.1	-0.2	-1.4	0.9	-0.8	-0.2	0.7	0.5
Rh	19.2	40.2	32.5	59.9	12.5	23.3	26.1	21.1	35.2	29.0	66.7	48.2	64.8
Td	-22.0	-14.7	-12.7	-6.4	-24.2	-18.4	-12.6	-17.9	-10.8	-15.0	-8.0	-8.0	-8.2
T	-2.1	-4.0	1.7	0.5	-1.7	-1.6	-1.7	-0.2	0.3	-1.9	-3.7	-2.2	-0.9
U	4.4	5.4	-2.3	2.8	6.2	-0.1	1.9	8.2	3.4	-5.7	2.7	6.0	2.0
V	-9.3	-0.1	-1.4	-2.6	-7.6	-4.9	0.9	-11.6	0.2	0.5	3.7	-0.1	-1.6
T-Td	19.9	10.7	14.3	6.9	22.5	16.7	10.8	17.7	11.1	13.1	4.3	5.8	7.3
Visib	10	5	2	2	2	23	5	2	14	0.9	14	4	1

表 2 部分实验结果
Tab.2 The partial experiment results

	多元回归	卡尔曼滤波	实况
1	5.3	17	19
2	5.2	9	6
3	5.5	5	3
4	8.2	3	1
5	19.3	2	0.9
6	26.0	28	33
7	5.9	25	19
8	5.6	8	11
9	5.6	5	4
10	5.5	3	2
11	2.9	1	2
12	29.3	20	16
13	26.6	24	28
14	11.6	26	35

表 3 主要预报因子与霾天气的对应统计关系
Tab. 3 The statistics of correspondence between the main predictors and the haze weather

	阈值(km)	比例(%)	平均
能见度	5~8	87.5	6.17
风速	1~3	91	2.11
相对湿度	35~55	74.2	51
温度露点差	9~15	78.8	11.9

表 4 能见度拟合结果与实况偏差统计
Tab. 4 The statistics of visibility fitting results and live deviation

	北京	广州	南京	杭州
能见度偏差 (拟合结果—实况)	1.9	2.92	1.87	2.09

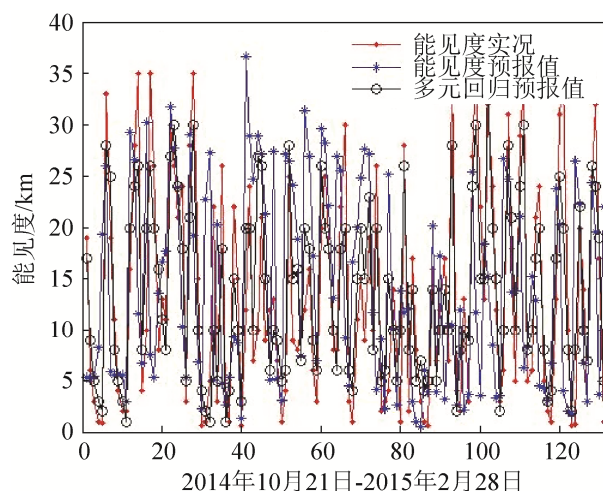


图 2 北京站实验结果对比

Fig. 2 Comparison of experimental results of Beijing Station

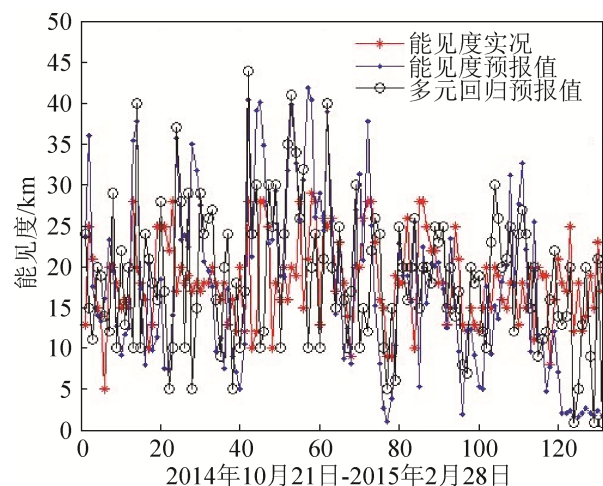


图 3 广州站实验结果对比

Fig. 3 Comparison of experimental results of Guangzhou Station

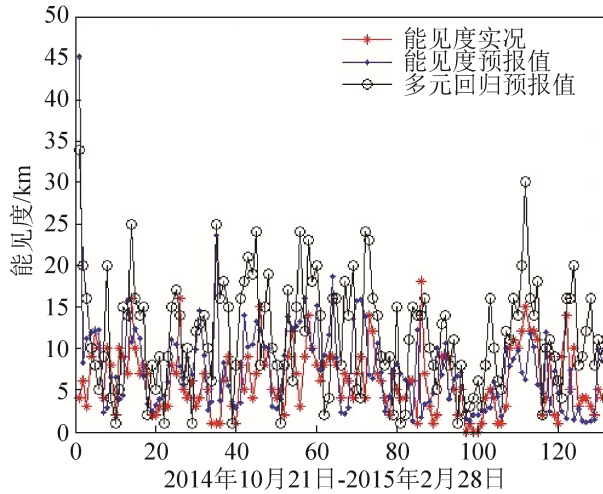


图4 杭州站实验结果对比

Fig. 4 Comparison of experimental results of Hangzhou Station

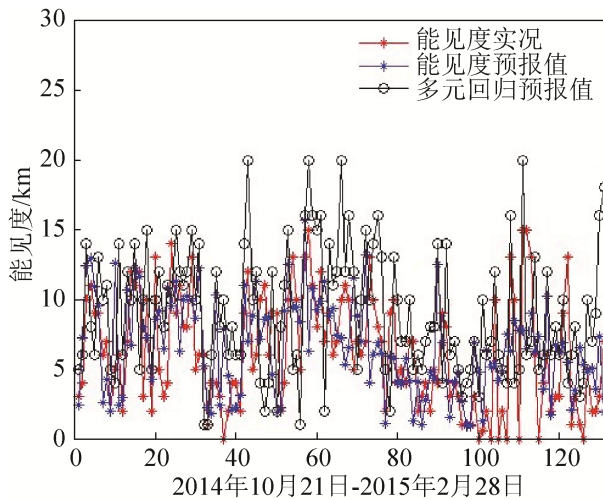


图5 南京站实验结果对比

Fig. 5 Comparison of experimental results of Nanjing Station

表5 本文方法预报检验结果

Tab. 5 The test results of the prediction method of this paper

	空报 次数	空报 率(%)	漏报 次数	漏报 率(%)	准报 次数	TS 评分
北京	61	41	3	2	86	0.57
广州	3	2	36	24	111	0.74
南京	50	34	14	9	86	0.57
杭州	74	50	8	5	68	0.45
平均	47	32	15	10	88	0.58

2.2.2 与数值预报系统结果对比分析

通过国家气象局 CUACE 模式对 2014 年 10 月 21 日至 2015 年 2 月 28 日共计 131 天时间的 4 个

代表站的霾天气进行预报实验, 利用 14 时资料进行检验。

由表 6 检验的对比结果可以看出利用多元逐步回归与卡尔曼滤波建立的能见度预报模型在排除降水、沙尘暴、扬沙、浮尘、吹雪、雪暴的影响视程的天气现象, 结合相对湿度、风速和温度露点差得出的霾预报模型进行霾预报时, 预报准确率相比于目前业务上应用的 CUACE 方法准确率提高了 15%。空报率和漏报率则分别降低了 13% 和 2%。进一步验证多元逐步回归与卡尔曼滤波方法在霾预报上有着更好的预报效果。

表6 本文方法与 CUACE 方法预报检验对比

Tab. 6 Comparison of the prediction results between the method of this paper and CUACE

	空报率(%)		漏报率(%)		TS 评分	
	CUACE	本文方法	CUACE	本文方法	CUACE	本文方法
北京	50	41	17	2	0.33	0.57
广州	10	2	20	24	0.70	0.74
南京	55	34	10	9	0.35	0.57
杭州	64	50	2	5	0.34	0.45
平均	45	32	12	10	0.43	0.58

3 结论

霾预报事关百姓生活和大气污染防治, 客观预报对现实应用有着非常重要的意义。随着数值预报技术和计算机能力的提高, 各种数值预报模式对霾的客观预报已经有了一定技术支撑, 包括欧洲中期数值预报模式(ECMWF)和中国的雾霾数值预报系统(CUACE)在内的诸多模式。本文通过结合数值预报模式和统计预报模式, 选取四个典型的代表站数据, 利用多元逐步回归与卡尔曼滤波算法进行霾预报模型的建立和实验。结果表明, 相对于目前业务上主要应用的 CUACE 模式预报分析方法, 本文所用方法能有效的提高霾客观预报的准确率, 体现了该方法的可行性和有效性。

目前该研究只是基于部分典型站点, 并且对霾天气学成因分析, 只是基于气象条件的结果, 没有结合大气成分的监测分析, 无法区分出沙尘等同样

能造成能见度降低的天气事件。天气条件必须和大气成分结合才能提供准确的霾预报。后续的工作将结合大气成分监测, 同时增加预报站点, 改进预报方法, 减少空报率。

参考文献:

- [1] 俞剑蔚, 孙燕, 张备, 等. 江苏沿江一次重霾天气成因分析[J]. 气象科学, 2009, 29(5): 664-669.
Jianwei Yu, Yan Sun, Bei Zhang, et al. Analysis on the cause of a heavy haze along the Yangtze River in Jiangsu [J]. Meteorological Science, 2009, 29(5): 664-669.
- [2] 穆泉, 张世秋. 2013年1月中国大面积雾霾事件直接社会经济损失评估[J]. 中国环境科学, 2013, 33(11): 2087-2094.
Quan Mu, Shiqiu Zhang. In January of 2013, Assessment of direct social and economic losses in China's large area haze events [J]. Environmental Science in china, 2013, 33(11): 2087-2094.
- [3] 吴兑, 毕雪岩, 邓雪娇, 等. 珠江三角洲大气灰霾导致能见度下降问题研究[J]. 气象学报, 2006, 64(4): 510-516.
Dui Wu, Xueyan Bi, Xuejiao Deng, et al. Study on the visibility reduction caused by atmospheric haze over the Pearl River Delta [J]. Journal of Meteorology, 2006, 64(4): 510-516.
- [4] 张运英, 黄菲, 杜鹃, 等. 广东雾霾天气能见度时空特征分析——年际年代际变化[J]. 热带地理, 2009, 29(4): 324-327.
Yunying Zhang, Fei Huang, Juan Du, et al. Spatial and temporal distribution of visibility under Fog and Haze conditions in Guangdong Province-Annual and Inter-Annual change [J]. Tropical Geography, 2009, 29(4): 324-327.
- [5] 姚剑, 王广华, 林俊, 等. 上海市大气颗粒物与能见度的关系[J]. 气象与环境学报, 2010, 26(4): 17.
Jian Yao, Guanghua Wang, Jun Lin, et al. The relationship between atmospheric particulates and visibility in Shanghai [J]. Journal of Meteorology and Environment, 2010, 26(4): 17.
- [6] 余锡刚, 吴建, 郇颖丽, 等. 灰霾天气与大气颗粒物的相关性研究综述[J]. 环境污染与防治, 2009, 20(2): 252-256.
Xigang Yu, Jian Wu, Yingli Li, et al. Review of researches on the correlation between atmospheric and particulate matter [J]. Environmental Pollution and Prevention, 2009, 20(2): 252-256.
- [7] 刘红年, 胡荣章, 张美根. 城市灰霾数值预报模式的建立与应用[J]. 环境科学研究, 2009, 22(6): 631-636.
Hongnian Liu, Rongzhang Hu, Meigen Zhang. Development and application of Urban Haze Numerical Forecast Model [J]. Environmental Science Research, 2009, 22(6): 631-636.
- [8] 尤佳红. 上海地区霾集合预报研究[D]. 上海: 华东师范大学, 2014.
Jiahong You. Study on haze ensemble forecast in Shanghai area [J]. Shanghai: East China Normal University, 2014.
- [9] 任万辉, 苏枫枫, 赵宏德. 城市环境空气污染预报研究进展[J]. 环境保护科学, 2010, 36(3): 9-11.
Wanhui Ren, Congcong Su, Hongde Zhao. Research progress on air pollution forecast in urban environment[J]. Environmental Protection Science, 2010, 36(3): 9-11.
- [10] 任建伟, 李永峰, 韩生梅. 基于灰色线性回归模型的预测控制[J]. 系统仿真技术, 2012, 8(2): 144-148.
Jianwei Ren, Yongfeng Li, Shengmei Han. Predictive control based on grey linear regression model [J]. System Simulation Technology, 2012, 8(2): 144-148.
- [11] 黄凤新. 基于卡尔曼滤波的复杂地形WRF模式预报风速订正[D]. 南京: 南京信息工程大学, 2013.
Fengxin Huang. Prediction of wind speed correction for complex terrain WRF model based on Kalman filtering[D]. Nanjing: Nanjing University of Information Science and Technology, 2013.