

DOI :10.3876/j.issn.1000-1980.2009.03.011

基于 MODIS 遥感和气象数据的 API 回归分析

黄凤辰^{1 2} 蔡志凌¹

(1.河海大学计算机及信息工程学院,江苏 南京 210098 ;2.河海大学通信与信息系统工程研究所,江苏 南京 210098)

摘要 :以江西省景德镇地区为研究区域,通过 MODIS 近红外大气反演大气水汽算法,提出了 MODIS 反演大气水汽含量联合气象数据对 API 进行回归分析的方法.仿真结果表明,该回归模型能够在一定程度上反映出 API 的变化,并可为 API 的预测提供数据参考.
关键词 :MODIS ;大气水汽 ;API ;气象数据 ;回归分析 ;景德镇地区
中图分类号 :X16 **文献标识码 :**A **文章编号 :**1000-1980(2009)03-0294-05

大气中水汽含量是一种重要的气象参数^[1].大气中水汽的分布、传输及其变化(云、雨、雪)与各种天气系统的更替变化息息相关.通过对水汽的监测可以更加准确地预报降水和灾害性天气,能够深入地了解气候特征和气候变化规律.

现有的大气水汽遥感反演方法可分为热红外、微波和近红外法^[2-3].

MODIS 传感器共包含 36 个波段,其中有 5 个近红外波段可用来反演大气水汽含量,分别是大气窗口通道 2 5 和水汽吸收通道 17、18 及 19.经严格测试,美国国家航空航天局(NASA)的 MODIS 专家 Gao 等^[4]给出了 MODIS 近红外反演大气水汽算法.孟宪红等^[5]利用 MODIS 近红外反演算法,证明了 MODIS 反演大气水汽含量和气象探空水汽含量相当接近.

空气污染指数(air pollution index ,API)是将常规监测的几种空气污染物浓度简化成为单一的概念性指数值形式,并分级表征空气污染程度和空气质量状况的参数,用于表示城市的短期空气质量状况和变化趋势.研究表明,城市空气质量好坏与季节及气象条件的关系十分密切^[6].相对于每周或每天而言,当污染源排放量没有大的变化情况下,风、雨、温度等气象条件会直接影响空气质量的好坏,使 API 产生很大的变化.例如降水稀少、无风天气多时,很容易形成雾天,而浓雾会吸附空气中漂浮的微小颗粒,形成细微的小水滴,再加上空气流动缓慢,不利于污染物的扩散,造成 API 增大.

本文以江西省景德镇地区作为研究区域,提出了 MODIS 反演大气水汽含量联合气象数据对 API 进行回归分析的方法,通过建立的回归模型来反映出 API 的变化.

1 MODIS 反演大气水汽含量算法

MODIS 反演大气水汽含量的算法有两通道比值法和三通道比值法 2 种,是根据辐射传输方程推导而来的.利用反射的太阳辐射监测水汽对辐射的吸收大小,总水汽量可用水汽的吸收通道和非吸收通道间反射的太阳辐射值获得.下面先给出近红外辐射传输方程:

$$L_S = L_{\text{gnd}}\rho T + L_{\text{atm}} \tag{1}$$

式中: L_S ——传感器接收到的辐射率; L_{gnd} ——地面直接反射和散射的辐射率; L_{atm} ——大气辐射率; ρ ——地面反射率; T ——大气透过率.其中大气透过率和大气辐射率与大气中的水汽含量有关,可以看成是水汽含量和气溶胶等大气参数的函数.将等式两边同时除以 $L_{\text{gnd}}\rho$ 得

$$T = \frac{L_S - L_{\text{atm}}}{L_{\text{gnd}}\rho} \tag{2}$$

在天气晴朗的情况下,式(2)中的 $\frac{L_{atm}}{L_{gnd}\rho}$ 非常小,因此式(2)又可以简化为

$$T \approx \frac{L_S}{L_{gnd}\rho}$$

(3)

从而大气水汽的含量主要是大气水汽透过率的函数.式(3)中的 ρ 对于不同波长是不一样的.但是 Gao 等^[7]的大量研究表明,在波长 0.85~1.25 μm 之间,各种地物反射率基本上满足线性关系,从而可以用大气窗口波段来模拟 $L_{gnd}\rho$.由此可以得出两通道比值法和三通道比值法.

两通道比值法利用通道 19 和 2 上的反射率比值来计算大气透过率:

$$T_{19} = \rho_{19}/\rho_2$$

(4)

当地物反射波谱不能满足 2 个通道相等,且其反射率在近红外波段呈线性变化时,可以用一个吸收通道和 2 个窗口通道组成的三通道比值法来计算大气透过率:

$$T_{19} = \frac{\rho_{19}}{C_1\rho_2 + C_2\rho_5}$$

(5)

式中 $C_1=0.8$, $C_2=0.2$.以上 2 种方法都是利用大气水汽吸收波段和大气窗口波段的比值来计算大气水汽透过率.对于透过率与水汽含量的关系表达式,可以通过 MODTRAN 软件和 LOWTRAN 软件模拟得到.文献[7]中给出了两通道比值法与大气水汽含量的关系表达式:

$$\rho_{19}/\rho_2 = \exp(\alpha - \beta w^{\frac{1}{2}}) \quad R^2 = 0.999$$

(6)

式中: w ——总的大气水汽含量; R ——相关系数; α , β ——常数.

式(6)中的 ρ_{19}/ρ_2 可以从 MODIS 影像得出,解式(6),得到水汽表达式如下:

$$w = \left[\frac{\alpha - \ln(\rho_{19}/\rho_2)}{\beta} \right]^2$$

(7)

对于复合型地表,常数 $\alpha=0.02$,常数 $\beta=0.651$,其他情况的常数见文献[7].

2 研究区域概况

研究区域为江西省景德镇市区,位于江西省东北部,属亚热带季风气候,境内光照充足,雨量充沛,温和湿润,四季分明.市区平均海拔 320 m,年平均气温 17 $^{\circ}\text{C}$,年平均降雨量 1763.5 mm,年平均日照时数为 2009.8 h.经纬度范围为北纬 29 $^{\circ}$ 05'~29 $^{\circ}$ 26',东经 117 $^{\circ}$ 01'~117 $^{\circ}$ 22'.

本文用于研究的 API 来源于景德镇市环保局,分别是 2006 年 1 月(冬季)、4 月(春季)、7 月(夏季)、10 月(秋季)4 个月的数据.对应的气象数据、MODIS 数据分别来源于景德镇市气象局和 NASA.用于回归分析的水汽含量为景德镇地区的 MODIS 数据反演而来,考虑到数据获取受限,在仿真时增加了部分模拟值^[7-10].

3 分析及验证

通过前面的 MODIS 近红外反演大气水汽含量算法,以景德镇市区为研究区域,反演出当地上空的水汽含量,联合气象数据,利用统计分析软件 SPSS 对空气污染指数 API 进行回归分析.

3.1 一元回归分析

以 MODIS 反演大气水汽含量为自变量,API 为因变量,在 d、旬 2 种时间单位下对 API 进行回归分析.根据散点图,建立二次方曲线模型.二次回归分析统计量及回归模型见表 1.

表 1 API 二次回归分析统计量和回归模型

Table 1 Statistics and models of API quadratic regression analysis

时间单位	<i>R</i>	<i>R</i> ²	修正 <i>R</i> ²	估计标准误差	回归模型
d	0.302	0.091	0.074	8.476	$y = 0.189x - 0.266x^2 + 63.522$
旬	0.812	0.660	0.572	4.363	$y = -2.629x - 0.175x^2 + 69.585$

注: x 为 MODIS 反演大气水汽含量, y 为 API.

在 d 时间单位下,二次回归模型中的 MODIS 反演大气水汽含量与 API 的相关性都较低.当时间扩大到以旬为单位时,相关性有了显著的提高,相关系数 R 从 0.302 提高到 0.812.

3.2 多元回归分析

由 API 的性质决定,单一的 MODIS 反演大气水汽含量并不能很好地反映出 API 的变化趋势.本文联合气

象数据对 API 进行多元回归分析.

多元回归分析选择将 MODIS 反演大气水汽含量以及平均气压、相对湿度作为因变量.在 d、旬 2 种时间单位下,对 API 进行回归分析.

多元回归分析统计量及回归模型见表 2.分析结果可知,在 d 时间单位下,MODIS 反演大气水汽含量联合气象数据建立的多元回归模型计算结果与 API 的相关性较低.当时间扩大到以旬为单位时,相关性有显著的提高,相关系数 $R = 0.852$.

表 2 API 多元回归分析统计量及回归模型
Table 2 Statistics and models of API multiple regression analysis

时间单位	R	R^2	修正 R^2	估计标准误差	回归模型
d	0.297	0.088	0.062	8.530	$y = -1.592x_1 - 0.018x_2 + 0.955x_3 + 65.404$
旬	0.852	0.726	0.615	3.040	$y = -5.129x_1 - 0.067x_2 + 11.353x_3 + 52.429$

注: x_1 为 MODIS 反演大气水汽含量, x_2 为相对湿度, x_3 为平均风速, y 为 API. 表 3 ~ 5 同.

3.3 基于不同季节的多元回归分析

用于实验分析的 4 个月 API 分别代表了春、夏、秋、冬 4 个季节,由于 API 的变化与季节、气候变化关系密切,不同季节的 API 变化趋势有很大不同,笼统地把 4 个月的 API 放在一起分析的做法并不理想.因此,本文把 API 平均值较高、气象条件类似的 1 月、10 月数据和 API 平均值较低、气象条件类似的 4 月、7 月数据区分开来.在 d 时间单位下,用对应的 MODIS 反演大气水汽含量和气象数据进行多元回归分析.

表 3 为不同月份 API 的回归分析统计量和回归模型(时间以 d 为单位).图 1 和图 2 为 d 时间单位下不同月份 API 回归模型的累计概率.

表 3 不同月份 API 多元回归分析统计量及回归模型

Table 3 Statistics and models of API multiple regression analysis for different months

月份	R	R^2	修正 R^2	估计标准误差	回归模型
1,10	0.338	0.114	0.065	8.063	$y = -1.353x_1 + 0.136x_2 + 1.785x_3 + 51.985$
4,7	0.452	0.205	0.154	8.624	$y = -2.145x_1 - 0.364x_2 - 0.510x_3 + 95.275$

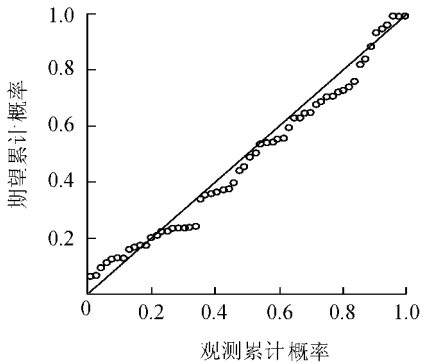


图 1 1 月、10 月 API 多元回归模型的累计概率

Fig. 1 Accumulated probabilities of multiple regression model for Jan. and Oct.

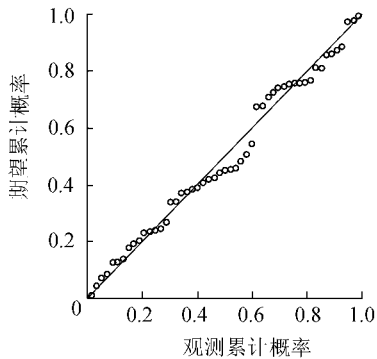


图 2 4 月、7 月 API 多元回归模型的累计概率

Fig. 2 Accumulated probabilities of multiple regression model for Apr. and July

由表 3 可知, d 时间单位下建立的 4 月、7 月回归模型的相关系数要明显高于 1 月、10 月以及整体 4 个月回归模型的相关系数.对比图 1 和图 2, 4 月、7 月的回归模型与原始数据也拟合得更好.

3.4 基于不同季节的分组多元回归分析

由于 API 用于表示城市的短期空气质量状况,因此扩大时间单位对其进行回归分析的实际意义不大.考虑到 API 是分级表征空气质量状况,将每天的 API 以从小到大、且按照一定的 API 单位间隔分为不同组,对组内的 API 取平均值.对应的 MODIS 反演大气水汽含量与气象数据也按照同样的方法处理.然后再联合各组 MODIS 反演大气水汽含量与气象数据对 API 进行多元回归分析.

表 4 为不同月份分组 API 回归分析统计量和回归模型, API 单位间隔为 3.图 3 和图 4 为 d 时间单位下不同月份分组 API 回归模型的累计概率.

表 4 不同月份分组 API 多元回归分析统计量及回归模型

月份	<i>R</i>	<i>R</i> ²	修正 <i>R</i> ²	估计标准误差	回归模型
1、10	0.856	0.733	0.645	7.115	$y = 1.474x_1 + 1.725x_2 + 28.760x_3 - 106.207$
4、7	0.873	0.762	0.693	8.602	$y = -6.369x_1 - 2.335x_2 - 4.430x_3 + 257.745$

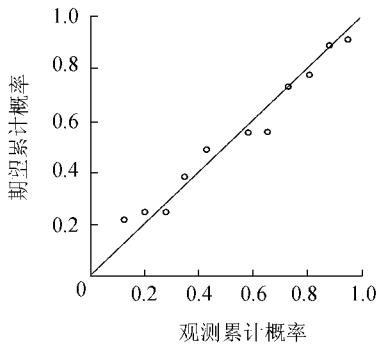


图 3 1 月、10 月分组 API 多元回归模型的累计概率

Fig. 3 Accumulated probabilities of multiple regression model for Jan. and Oct. group

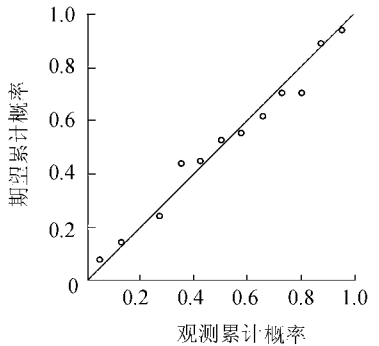


图 4 4 月、7 月分组 API 多元回归模型的累计概率

Fig. 4 Accumulated probabilities of multiple regression model for Apr. and July group

分析结果可知,将 API 间隔分组后,不同月份回归模型的相关系数均有显著提升.1 月、10 月 API 回归模型的相关系数由 0.338 提高到 0.856,4 月、7 月 API 回归模型的相关系数由 0.452 提高到 0.873.对比图 3 和图 4,4 月、7 月 API 回归模型的拟合度依然优于 1 月、10 月 API 回归模型.

3.5 加权最小二乘法优化回归分析

这里使用加权最小二乘法来优化多元回归模型.依照前面的做法,把 API 按照月份分为 2 部分,然后分别以单位间隔 3 进行分组,最后用对应的 MODIS 反演大气水汽含量和气象数据进行多元回归分析,并选择 MODIS 反演大气水汽含量作为加权变量.

表 5 为不同月份分组 API 加权优化后的回归分析统计量和回归模型.对比表 4、表 5 可知,加权回归分析后,模型的相关系数得到进一步提高.1 月、10 月 API 回归模型的相关系数由 0.856 提高到 0.941,4 月、7 月 API 回归模型的相关系数由 0.873 提高到 0.939,这说明加权回归后的优化效果还是比较明显的.

表 5 不同月份分组 API 多元加权回归分析统计量及回归模型

月份	<i>R</i>	<i>R</i> ²	修正 <i>R</i> ²	估计标准误差	回归模型
1、10	0.941	0.885	0.846	4.810	$y = 1.057x_1 + 1.663x_2 + 26.017x_3 - 97.107$
4、7	0.939	0.882	0.842	26.951	$y = -5.954x_1 - 2.835x_2 - 6.822x_3 + 296.568$

4 结 语

采用 MODIS 反演大气水汽含量建立的一元回归模型以及联合相对湿度、平均风速建立的多元回归模型,能在一定程度上反映出 API 的变化.由于 API 用于表示城市短期几天内的空气质量状况,旬时间单位 API 的回归模型可以作为短期空气质量状况的一种参考.

考虑到季节、气候变化对 API 的影响,本文把 4 个月的 API 分为 1 月(冬季)、10 月(秋季)和 4 月(春季)、7 月(夏季)2 部分进行分析.因为降水较少,再加上城市大量使用取暖设备,污染物浓度上升,致使 1 月、10 月的 API 平均值高于 4 月、7 月.由于受到更多污染物排放的影响,气候变化对 API 的影响明显小于 4 月、7 月.因此,旬时间单位下 4 月、7 月 API 多元回归模型的相关系数要明显高于 1 月、10 月 API 多元回归模型的相关系数.

选择一定的 API 单位间隔对 API 分组后,1 月、10 月与 4 月、7 月 2 组多元回归模型的相关系数均有显著提升.使用加权回归优化后,回归模型的相关系数再次得到不同程度的提高.

总体来说,MODIS 反演大气水汽含量、相对湿度、平均风速的变化会影响到空气质量,造成 API 的上下波

动.分组后建立的4月、7月与1月、10月API多元回归模型能够准确地反映出API的变化趋势,可以为城市日常空气质量、API的预测提供数据参考.由于实验数据的限制,本文中的部分水汽含量为模拟值,但分析得到的结果仍具有适用性.在以后的研究中,需要进一步增加MODIS反演的水汽含量数据来修正建立的回归模型.

参考文献:

- [1] 盛裴轩,毛节泰,李建国,等.大气物理学[M].北京:北京大学出版社,2003:5.
- [2] 宋正方,魏合理,吴晓庆.大气水汽的红外遥感[J].环境遥感,1996,11(2):130-137.
- [3] 陈洪滨,吕达仁,魏伟,等.空基微波辐射计遥感晴天大气可降水量:不同通道组合和亮温函数形式的效果的比较分析[J].大气科学,1996,20(6):757-762.
- [4] GAO Bo-cai, KAUFMAN Y J. The MODIS near-IR water vapor algorithm, product ID: MOD05-total precipitable water[EB/OL].[2008-05-05]. http://modis-atmos.gsfc.nasa.gov/_docs/atbd_mod03.pdf.
- [5] 孟宪红,吕世华,张宇,等.基于MODIS数据的金塔绿洲上空大气水汽含量反演研究[J].水科学进展,2007,18(2):264-269.
- [6] 何莉.南宁市空气污染特征及预报模型研究[D].南宁:广西大学,2007.
- [7] GAO BO-cai, KAUFMAN Y J. Remote sensing of water vapor in the near IR from EOS/MODIS[J]. IEEE Transactions on Geoscience and Remote Sensing, 1992, 30(5):871-884.
- [8] 朱文明.基于MODIS大气水汽的时间变化及其与气象因子的相关分析[D].南京:南京师范大学,2007.
- [9] 王艳红,马荣华,邓正栋.基于分区的太湖叶绿素a遥感估测模型[J].河海大学学报:自然科学版,2007,35(1):86-91.
- [10] 李士进,艾萍.基于MODIS数据的洪涝灾情监测关键技术研究进展[J].水利水电科技进展,2007,27(增刊2):1-4.

Regression analysis of air pollution index based on MODIS remote sensing and meteorological data

HUANG Feng-chen^{1,2}, CAI Zhi-ling¹

(1. College of Computer and Information Engineering, Hohai University, Nanjing 210098, China;

2. Institute of Communications and Information System Engineering, Hohai University, Nanjing 210098, China)

Abstract: Taking Jingdezhen City of Jiangxi Province as an example, a new method for regression analysis of the air pollution index(API) was put forward using the meteorological data and the atmospheric water vapor retrieved by the MODIS near-IR water vapor algorithm. The simulated results show that the present regression model can reflect the change of API to a certain extent, and it can provide a reference for the forecast of API.

Key words: MODIS; atmospheric water vapor; air pollution index; meteorological data; regression analysis; Jingdezhen City