

# 引入 Himawari-8 卫星数据协变量的 能见度样条插值方法

赵春雷<sup>1,2</sup> 杨鹏<sup>3</sup> 张杏敏<sup>4</sup> 赵增保<sup>4</sup> 冯一淳<sup>5</sup>

(1 河北省气象科学研究所, 石家庄 050021; 2 河北省气象与生态环境重点实验室, 石家庄 050021;  
3 石家庄市气象局, 石家庄 050081; 4 河北省气象服务中心, 石家庄 050021; 5 河北省承德市气象局, 承德 067000)

**摘要** 在 ANUSPLIN 薄盘光滑样条插值中, 高相关协变量的选取决定了插值结果的精确性。本文选取 2017—2019 年大雾和霾能见度较差的天气过程, 利用 183 个能见度观测站点对能见度进行插值, 引入 Himawari-8 卫星的通道数据和 DEM 数据作为协变量对能见度的插值结果进行改进, 并对能见度插值结果进行对比分析。研究表明, 引入 Himawari-8 数据和 DEM 数据作为协变量的能见度插值结果在精度上有显著提高, 尤其对雾区和霾区的边界范围和纹理的反演更为准确, 基于 Himawari-8 卫星数据和气象监测站点的观测数据, 使用协变量的方法进行能见度插值可以做能见度监测网格化的一种有效途径。

**关键词** 样条插值; ANUSPLIN; Himawari-8; 协变量; 能见度

**中图分类号**: P413 **DOI**: 10.19517/j.1671-6345.20190066 **文献标识码**: A

## 引言

能见度是人眼所能将目标物从背景中分辨出来的最远水平距离<sup>[1]</sup>。由于大气中的气体和气溶胶对光的散射和吸收作用, 自然状况下的视程范围都会受到一定的限制。雾、霾、雪、沙尘暴及降水等天气现象都能导致大气低能见度事件的出现。随着工业技术的飞速发展和城市化进程的加快, 化石燃料和生物质的燃烧过程、污染物的传输以及其他燃烧源等排放出的大量气溶胶颗粒物, 使中国的很多地区都遭受到了严重的气溶胶污染。京津冀区域由于其特殊的地理位置, 经常出现大范围污染天气, 导致能见度下降, 严重影响交通安全及危害百姓生活<sup>[2-4]</sup>。

准确的能见度空间分布数据, 理论上可由高密度站网采集, 气象数据的观测主要依赖于气象台站, 但目前气象台站空间分布不均, 密度最少几十公里。因此, 站点外区域的气象数据通常只能由邻近站点的观测值进行估算, 即气象信息空间插值。近年来, 已有较多关于气象要素插值的专门探讨和研究<sup>[5]</sup>。空间数据内插就是根据一组已知的离散数据或分区数据, 按照某种数学关系推求出其他未知点或未知

区域数据的数学过程。目前常用的方法有最近邻近法(泰森多边形法)、移动平均法、样条函数法(spline 法)、空间自协方差最佳插值法(Kriging 法)等。针对气候数据曲面拟合的专用软件 ANUSPLIN 是基于样条插值理论, 同时其允许引进多元协变量线性子模型, 且模型系数可根据数据自动确定, 因此 ANUSPLIN 可以平稳地处理二维以上的样条, 这就为引入多个影响因子作为协变量进行气象要素空间插值提供了可能<sup>[6-8]</sup>。

Himawari-8(以下简称 H8)是日本气象厅的第二代静止气象卫星, 2014 年 10 月发射成功, 2015 年 7 月投入正式业务运行。H8 的常规观测频次为 10 min, 其数据的空间分辨率从 0.5~2 km, 应用此数据开展地面能见度反演对公共交通、天气过程监测和预警具有重要的现实意义。由于 H8 卫星具有监测频次高、空间分辨率高等特点, 它所提供的连续遥感数据以及地面自动监测站提供的能见度数据使得建立能见度的遥感模型成为可能<sup>[9-10]</sup>。

本文尝试将 H8 卫星的通道数据和 DEM 高程数据作为协变量引入 ANUSPLIN 的插值算法进行能见度反演, 并分析不同天气条件下的反演精度。

<http://www.qxkj.net.cn> 气象科技

河北省气象局科研面上项目“基于葵花卫星的雾区和 PM<sub>2.5</sub> 浓度反演技术”(16kyd06)资助

作者简介:赵春雷,男,1967年生,正研级高级工程师,主要从事卫星遥感及生态监测技术应用,Email:308920914@qq.com

收稿日期:2019年2月15日;定稿日期:2019年7月9日

## 1 数据与方法

### 1.1 数据来源和处理

本文所使用的能见度地面监测数据来源于河北省气象信息中心,京津冀区域内能见度数据来源于各气象站点的能见度自动监测仪。182 个能见度地面监测站点的分布见图 1 所示。

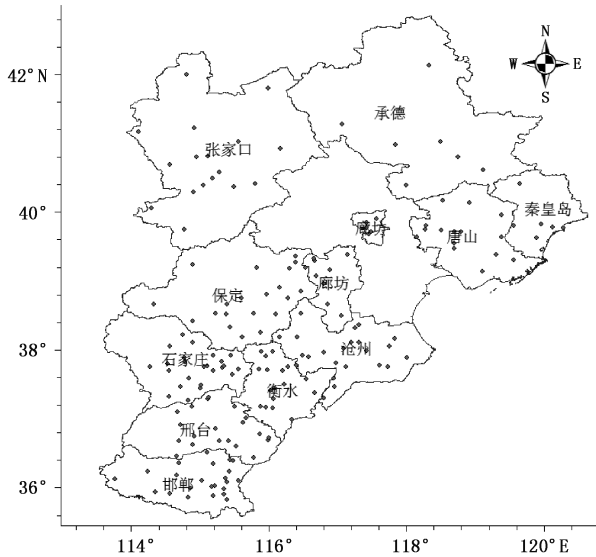


图 1 能见度地面站点分布

能见度的大小与雾、霾等天气有明显的相关变化趋势,为了更加形象、直观地看到插值结果,在 ANUSPLIN 插值制图过程中,参照了气象学上按水平能见度距离划分雾和霾的等级标准,并将插值反演得到的能见度进行分级(表 1、表 2)。

表 1 雾的等级标准

	阈值区间/km	类型
1	[0,0.05)	强浓雾
2	[0.05,0.2)	浓雾
3	[0.2,0.5)	大雾
4	[0.5,1.0)	雾
5	[1.0,2.0)	过渡期
6	[2.0,+)	非雾

表 2 霾的等级标准

	1	2	3	4
阈值区间/km	[0,2)	[2,5)	[5,10)	[10,+)
类型	重度霾	中度霾	轻度霾	非霾

H8 遥感数据来源于中国气象局,数据格式为 HSD。H8 数据有 16 个通道(表 3),包括 3 个可见光,3 个近红外、10 个红外通道,常规观测间隔为 10 min,对特殊区域或台风天气还可实施区域加密观测,数据的空间分辨率分为 0.5、1、2 km 不等。

表 3 H8 卫星(Himawari-8)AHI 成像仪参数

通道	中心波长 $\mu\text{m}$	空间分辨率 km	主要用途
可见光	1	0.47	白天陆地气溶胶,沿海水域
	2	0.51	白天水/海洋颜色,检测藻华,与蓝色和红色通道组合可以实现真彩色图像
	3	0.64	白天云,雾,风
近红外	4	0.86	白天植被,燃烧痕迹,海上气溶胶,风
	5	1.60	白天云顶相态和粒径大小,积雪
	6	2.30	白天陆地/云属性,粒径大小,植被,雪
短波红外	7	3.90	地表和云,晚上雾和低云,火,风
	8	6.20	高层大气水汽,风,降雨
中波红外	9	6.90	中层大气水汽,风,降雨
	10	7.30	低层水汽,风,二氧化硫
	11	8.60	稳定水总量,云相,粉尘,二氧化硫,降雨
长波红外	12	9.60	总柱臭氧,湍流,风
	13	10.40	地表和云
	14	11.20	影像,海面温度(SST),云,降雨
	15	12.40	总柱水蒸气,灰分,SST
	16	13.30	气温,云高和总量

ANUSPLIN 所需的 DEM 数据要求按 ASCII 形式书写,需要将栅格形式转换成 ASCII 格式。本文所用 DEM 数据来源于河北省基础地理信息中心,经重采样后形成河北省 1 km×1 km 的网格(图 2)数据,然后转换成 ANUSPLIN 要求的 ASCII 格式<sup>[11]</sup>。

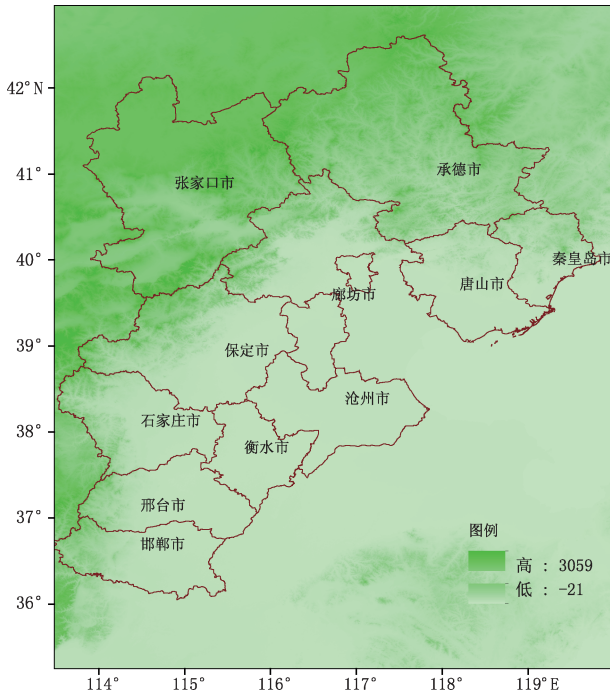


图 2 河北省高程图(DEM)

## 1.2 ANUSPLIN 插值原理与方法

ANUSPLIN 软件基于普通薄盘和局部薄盘样条函数插值理论<sup>[12-13]</sup>,除了可以引入自变量外,还允许引入协变量。

局部薄盘光滑样条的理论统计模型公式:

$$z_i = f(\mathbf{x}_i) + \mathbf{b}^T \mathbf{y}_i + e_i \quad (i = 1, \dots, N) \quad (1)$$

式中, $z_i$  是位于空间  $i$  点的因变量; $\mathbf{x}_i$  为  $p$  维样条独立变量矢量; $f$  为要估算的关于  $\mathbf{x}_i$  未知光滑函数; $\mathbf{y}_i$  为  $p$  维独立协变量矢量; $\mathbf{b}$  为  $\mathbf{y}_i$  的  $p$  维系数; $e_i$  为具有期望值为 0 和方差为  $\omega_i \sigma^2$  的自变量随机误差; $\omega_i$  为作为权重的已知相对误差方差; $\sigma^2$  为所有数据点上一常数据误差方差,但通常未知; $T$  为转置符号。

由式(1)可见,当式中缺少第 2 项,即协变量( $p=0$ )时,模型简化为普通薄盘光滑样条;当缺少第 1 项独立自变量时,模型变为多元线性回归(ANUSPLIN 中不允许这种情况出现),函数  $f$  和系数  $b$  通过最小二乘估计来确定:

$$\sum_{i=1}^N \left[ \frac{z_i - f(\mathbf{x}_i) - \mathbf{b}^T \mathbf{y}_i}{\omega_i} \right]^2 + \rho J_m(f) \quad (2)$$

其中, $J_m(f)$  为函数  $f(\mathbf{x}_i)$  的粗糙度测度函数, $m$  在 ANUSPLIN 中称为样条次数,也叫糙度次数; $\rho$  为正的光滑参数,在数据保真度与曲面的粗糙度之间起平衡作用,在 ANUSPLIN 中通常用广义交叉验证 GCV 的最小化以及最大似然法 GML 的最小化来确定。

对于时间序列气象要素的空间插值,既要保证每个表面的插值精度,又要保证插值模型的相对稳定性,使其在时间连续性上具有可比性。在对研究区不同时间节点的数据进行曲面空间插值过程中,分别选取不同的 DEM 数据与 H8 数据的波段进行组合,参照 ANUSPLIN 的原理及 H8 数据通道特征,经多次试验及验证最终确定以经纬度作为自变量,白天选用 DEM(高程)、CH3、CH5、CH7、CH14 作为协变量(夜间无 CH3、CH5)作为最优方案。

## 2 结果分析

本文选取大雾和霾两种天气过程,对比分析两种天气过程的能见度插值效果,具体选择日期见表 4。

表 4 大雾和霾天气过程选择日期

天气过程	选择日期
大雾	2017 年 4 月 20 日、7 月 30 日、10 月 11 日
霾	2018 年 3 月 16 日至 3 月 19 日

### 2.1 大雾过程

图 3 分别为 2017 年 4 月 20 日、7 月 30 日和 10 月 11 日 3 天的实况数据、未加协变量和加入协变量 3 种情况下的分布图。图 4 给出了加入协变量和未加入协变量与实况数据的相关系数,从图 4 可以看出,2017 年 4 月 20 日、7 月 30 日和 10 月 11 日未加入协变量时的相关系数分别为 0.38、0.38 和 0.36,加入协变量后相关系数分别为 0.91、0.78 和 0.71,加入协变量后的结果与实况数据的相关系数有明显提高,可以得出,加入协变量后的插值效果优于未加协变量的插值效果。

图 5 为 2019 年 1 月 12 日 13:00 未加入协变量和加入协变量的能见度分布图,可以看出:加入协变量反演出衡水、沧州、邢台部分地区已达大雾级别,而未加入协变量反演全省只达到雾级别。通过实况数据发现,13:00 能见度最小值出现在邢台市宁晋站(站号 53796),监测值为 362 m,属大雾级别,未加入协变量的反演结果为 670 m(雾),加入协变量的反演结果为 446 m(大雾),对极值区域有更好的反映。

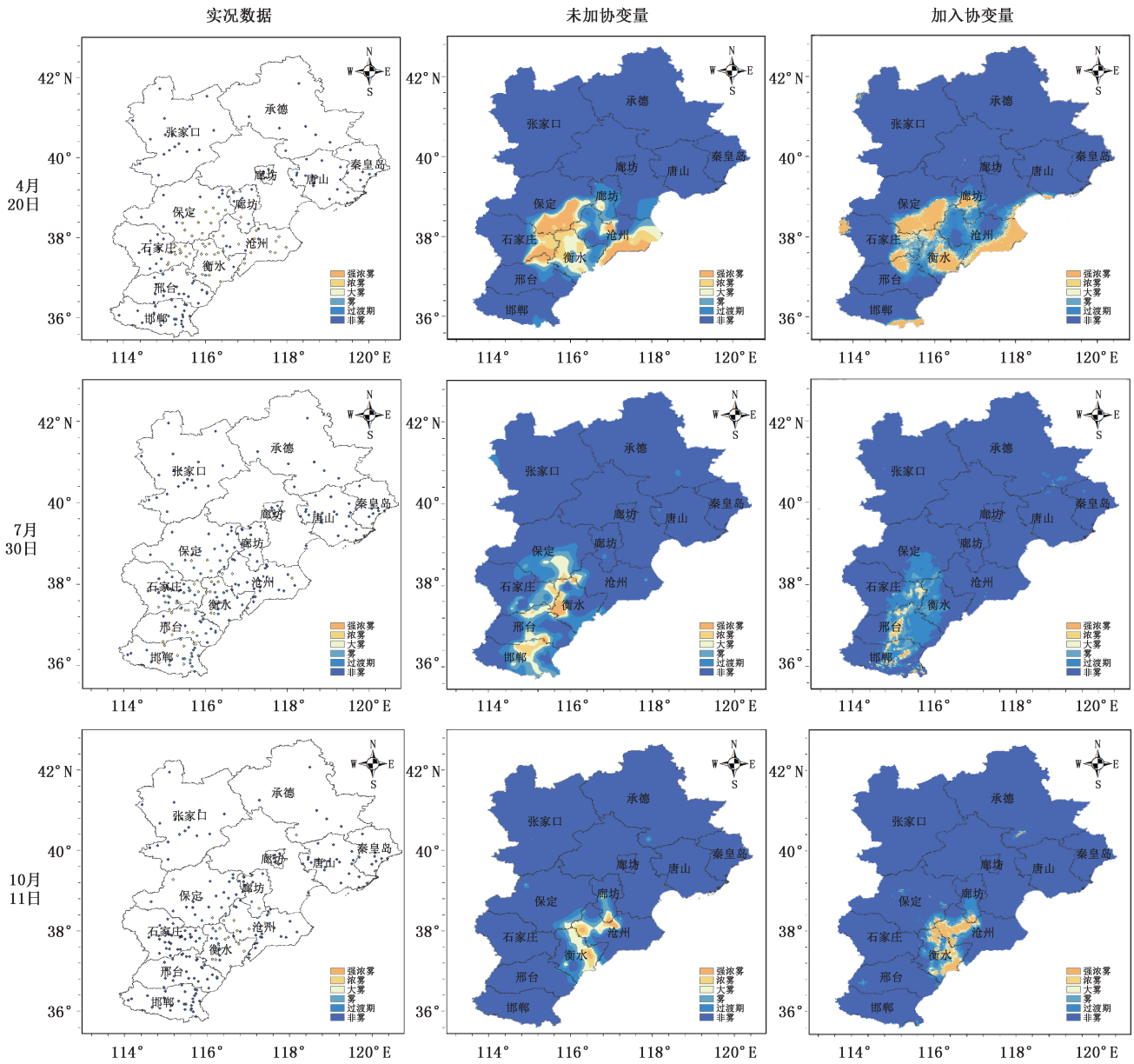


图 3 2017 年 4 月 20 日、7 月 30 日和 10 月 11 日实况数据、未加协变量和加入协变量的大雾分布

图 6 为饶阳站的能见度时间变化曲线,红色曲线代表加入协变量的反演结果,绿色曲线代表实况值。可以看出,03:30 左右饶阳站开始从轻雾转向雾发展,反演结果在时间上略有提前,对能见度低于 500 m 时,反演结果良好,误差在 100 m 以内,对于能见度极低值,反演结果较实际结果略偏低。对所有站点统计这次雾过程期间的准确率,结果见表 5,表中等级差代表反演结果等级与实况等级的差值。

可以看出,在出现强浓雾的站点和时次中(等级 1),反演结果为强浓雾的共占 40%(准确),反演结果为浓雾的占 30%(基本准确),反演结果为 1.0~2.0 km 的占到 30%(错误),准确率接近 70%;在实况为浓雾的情况,反演结果准确的占 8.5%,反演结果为大雾或强浓雾的占 40.13%,反演结果为雾的占 17.0%,准确率接近 65.6%;对过渡期及非雾情况反演准确率均较高。

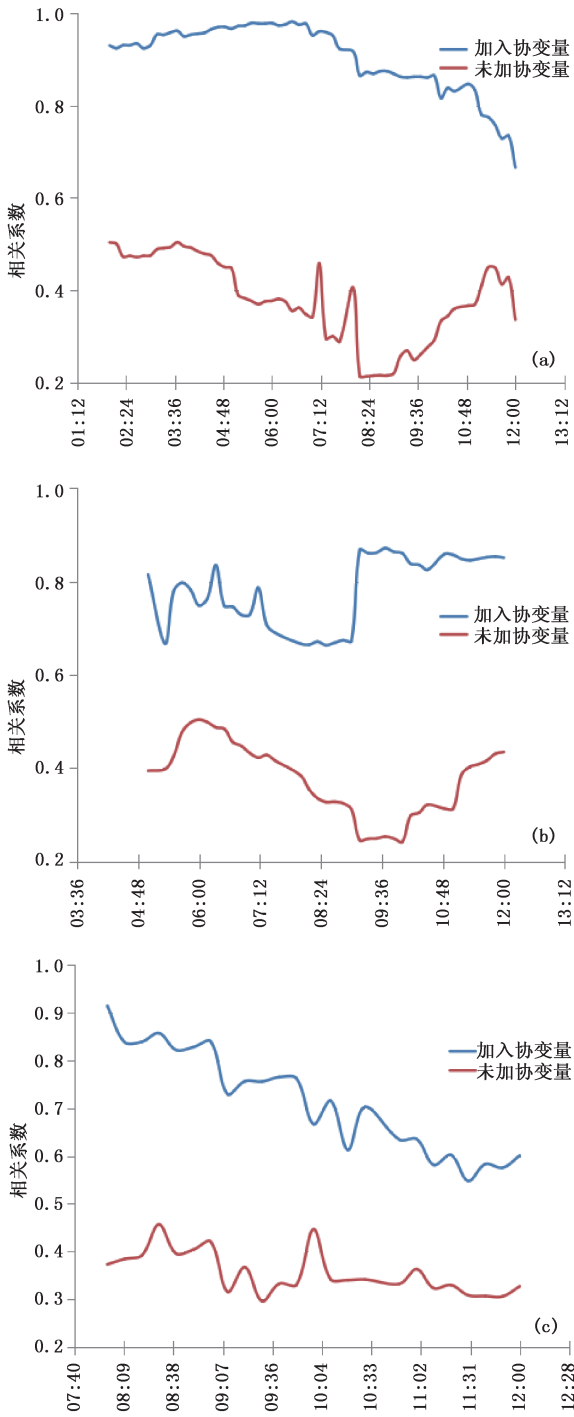


图 4 加入协变量和未加入协变量与实况的相关系数：  
(a)2017年4月20日,(b)2017年7月30日，  
(c)2017年10月11日

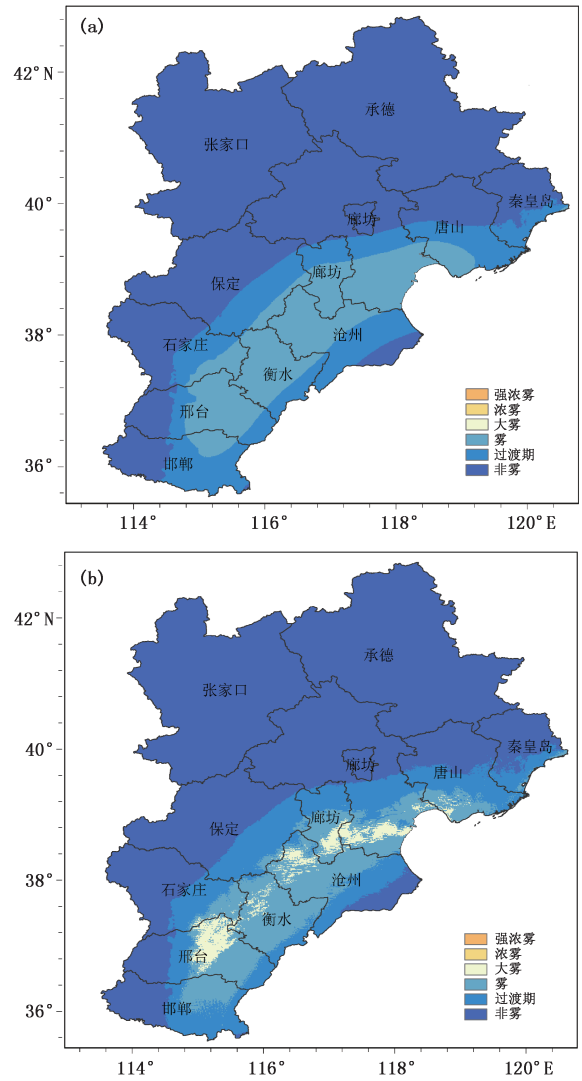


图 5 2019年1月12日未加入协变量(a)和加入协变量反演结果(b)对比

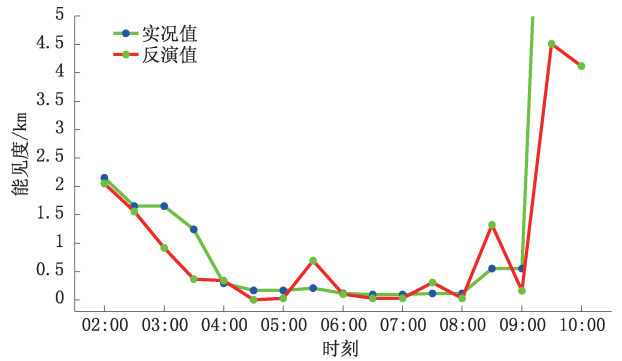


图 6 2017年4月20日饶阳站的能见度时间变化曲线

表 5 不同等级能见度下准确率 %

等 级	等级差							样本数
	0	1	2	3	4	5	6	
1	40.00	30.00	0.00	0.00	30.00	0.00	0.00	10
2	8.50	40.13	17.00	20.40	13.94	0.00	0.00	294
3	9.61	10.57	44.23	35.57	0.00	0.00	0.00	104
4	16.66	38.88	42.59	1.85	0.00	0.00	0.00	54
5	23.18	63.76	8.69	1.44	2.89	0.00	0.00	138
6	96.24	2.75	0.56	0.30	0.04	0.08	0.00	2290

2.2 霾天气过程

针对 2018 年 3 月 16—19 日霾天气过程,利用获得的 183 个站点的观测数据,其中 161 个站点的观测数据分别使用协变量方法(通道 3、5、7、14 及 DEM 作为协变量)和只使用观测数据方法进行插

值,利用 22 个观测站点对两种插值方法的结果进行检验。可以看出,使用协变量的方法进行插值与仅对观测值进行插值在精度上有显著提高(图 7)。参与检验站点对使用协变量方法和未使用协变量方法的反演结果误差统计见表 6、表 7 所示,统计误差可以看出:在反演霾级差为 0 的情况下,使用协变量的反演结果准确率在 1、2、3、4 四个霾等级分别为 27%、62%、64% 和 69%,均高于未使用协变量的 18%、45%、37% 和 49%;在反演霾级差 ≤ 1 的情况下,使用协变量的反演准确率在 1、2、3、4 四个霾等级均比未使用协变量高 19%、12%、2% 和 7%。因此,使用协变量比未使用协变量的插值结果与观测值更为接近。

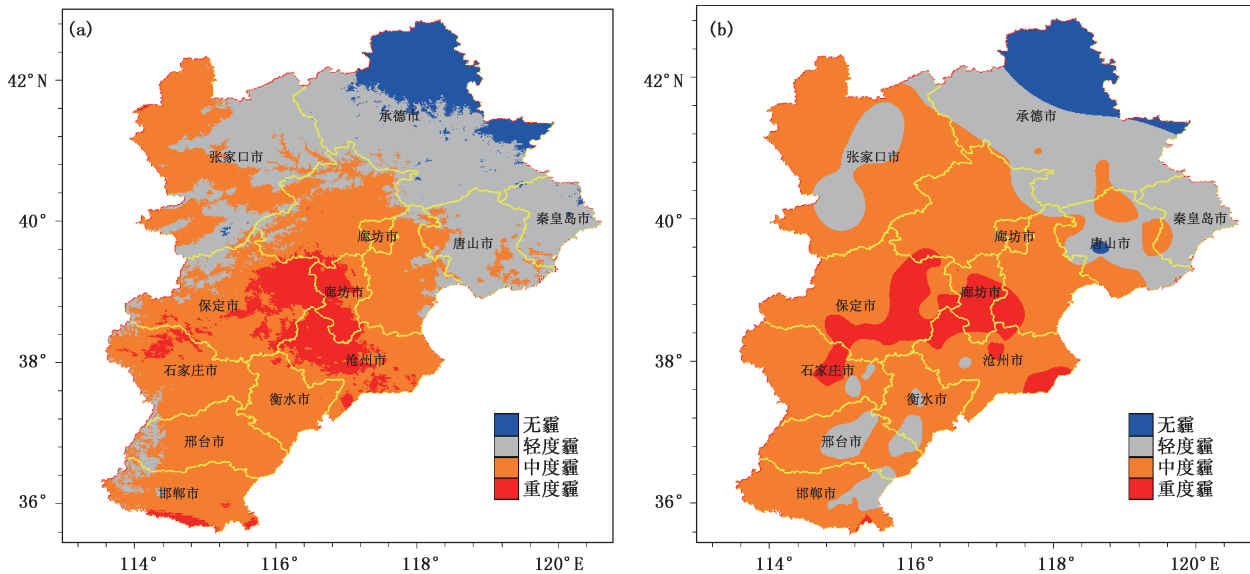


图 7 2018 年 3 月 16—19 日霾天气采用通道 3、5、7、14 及 DEM 数据插值(a)和只使用观测数据插值(b)反演结果

表 6 使用协变量方法检验结果准确率

监测等级	预测级差				样本数
	0	1	2	3	
1	27[3]	55[6]	0[0]	18[2]	11
2	62[115]	36[67]	2[3]	0[0]	185
3	64[226]	36[126]	0[1]	0[0]	353
4	69[288]	29[122]	2[9]	0[0]	419

表 7 未使用协变量方法检验结果准确率

监测等级	预测级差				样本数
	0	1	2	3	
1	18[2]	45[5]	18[2]	18[2]	11
2	45[84]	34[63]	21[38]	0[0]	185
3	37[129]	61[217]	2[7]	0[0]	353
4	49[206]	42[175]	9[36]	0[2]	419

3 结论

本文尝试将 H8 卫星的通道数据和 DEM 高程数据作为协变量引入 ANUSPLIN 的插值算法进行

能见度反演,选取 2017—2019 年大雾和霾能见度稍差的天气过程,利用气象站点观测值对能见度进行了插值反演,主要结论如下:

(1)在应用 ANUSPLIN 软件进行能见度网格插值过程中,将 H8 数据和 DEM 高程数据作为协变量引入,能有效提高插值精度。加入协变量后的插值效果优于未加协变量的插值效果,使用协变量的方法进行插值在精度上有显著提高,对雾区的极值有较好的反映。

(2)基于 H8 数据和 DEM 高程数据的能见度反演,在雾区范围的反演效果较好,对雾的起始时间往往略有提前,对结束时间把握较为准确。对霾区的插值,使用协变量的方法进行插值与仅对观测值进行插值在精度上有显著提高。

### 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.
- [2] Levy R C, Remer L A, Mattins J V, et al. Evaluation of the MODIS aerosol retrievals over ocean and land during CLAMS [J]. *J Atmos Sci*, 2005, 62(4): 974-992.
- [3] Levy R C, Mattoo S, Munchak L A, et al. The Collection 6 MODIS aerosol products over land and ocean [J]. *Atmos Chem Phys*. 2013, 6: 2989-3034.
- [4] 赵秀娟,陈长和,张武,等. 利用 MODIS 资料反演兰州地区气溶胶光学厚度 [J]. *高原气象*, 2005, 24(1): 97-103.
- [5] 林忠辉,莫兴国,李宏轩,等. 中国陆地区域气象要素的空间插值[J]. *地理学报*, 2002, 57(1): 47-56.
- [6] Peter M, Chris D. Mapping precipitation in Switzerland with ordinary and indicator Kriging [J]. *Journal of Geographic Information and Decision Analysis*, 1998, 2(2): 65-76.
- [7] Hutchinson M F, Gessler P E. Splines-more than just a smooth interpolator [J]. *Geoderma*, 1994, 62: 45-67.
- [8] 刘志红,McVicar T R, Li Lingtao. 基于 5 变量局部薄盘光滑样条函数的蒸发空间插值[J]. *中国水土保持科学*, 2006, 4(6): 23-30.
- [9] 刘超,历华,杜永明,等. Himawari 8 AHI 数据地表温度反演的实用劈窗算法[J]. *遥感学报*, 2017, 21(5): 702-714.
- [10] 王萌,郑伟,李峰. Himawari-8 气象卫星黄海浒苔动态监测 [J]. *应用气象学报*, 2017, 28(6): 714-723.
- [11] 钱永兰,吕厚荃,张艳红. 基于 ANUSPLIN 软件的逐日气象要素插值方法应用与评估[J]. *气象与环境学报*, 2010, 26(2): 7-15.
- [12] 刘志红,McVicar T R, Li Lingtao,等. 基于 ANUSPLIN 的时间序列气象要素空间插值[J]. *西北农林科技大学(自然科学版)*, 2008, 36(10): 227-234.
- [13] Bates D, Lindstrom M, Wahba G. Gcvpack-routines for generalized cross validation [J]. *Communications in Statistics B-Simulation and Computation*, 1987, 16: 263-297.

## Visibility Spline Interpolation Method for Introducing Himawari-8 Satellite Data Covariate

ZHAO Chunlei<sup>1,2</sup> YANG Peng<sup>3</sup> ZHANG Xingmin<sup>4</sup> ZHAO Zengbao<sup>4</sup> FENG Yichun<sup>5</sup>

(1 Meteorological Institute of Hebei Province, Shijiazhuang 050021; 2 Key Laboratory of Meteorology and Ecological Environment of Hebei Province, Shijiazhuang 050021; 3 Shijiazhuang Meteorological service, Shijiazhuang, 050081; 4 Meteorological Service Centre of Hebei Province, Shijiazhuang 050021; 5 Chengde Meteorological Service, Hebei, Chengde 067000)

**Abstract:** In the ANUSPLIN thin-plate smooth spline interpolation, the accuracy of interpolation results is mainly determined by choosing the independent covariates. This article selected the weather processes with poor visibility in heavy fogs and hazes from 2017 to 2019, using 183 visibility observation sites to interpolate visibility, and introduced the Himawari-8 satellite channel data and DEM data as covariables to improve the visibility interpolation results. The visibility interpolation effects are compared and analyzed. The results show that the visibility interpolation effect of Himawari-8 data and DEM data as covariables is significantly improved in accuracy, especially in the inversion of the boundary range and texture of fog and haze areas. The accuracy of interpolation using the covariate method and interpolation only using the observed values is greatly improved.

**Keywords:** spline interpolation; ANUSPLIN; Himawari-8; covariable; visibility