

基于特征向量法的地基红外高光谱反演大气 温湿廓线附加影响因子分析

李宛桐¹ 黄威² 姜明¹ 史静¹

(1 天津市气象探测中心, 天津 300061; 2 中国洛阳电子装备试验中心, 济源 459000)

摘要 在地基高光谱遥感中,特征向量法获取的温湿廓线以初值的方式对物理反演进行约束,其反演精度对物理反演结果有着重要的影响。利用 AERI 的观测辐射资料 and 同站点的探空数据,基于特征向量法分析了温度廓线与湿度廓线反演的异同点;研究了主成分个数的选择问题,综合考虑反演精度和特征向量中包含的信息将反演温度廓线和湿度廓线的最优主成分个数定为 7。为提高反演精度,引入地面温度、湿度、气压作为影响因子,试验结果表明,考虑反演精度和稳定性,地面气压的引入相比于其他 2 种单一气象要素以及 3 种气象要素组成的因子集表现更好,尤其是对边界层中下部的温湿廓线有着明显的提升,并随着高度的降低提升作用更明显,温度廓线 RMSE 降低最高达到 1.5 K,湿度廓线 RMSE 降低最高达到 0.42 g/kg。同时,分析了对数反演形式对湿度廓线的影响,结果表明,以水汽混合比的形式反演时取自然对数对反演精度的影响较小;将反演得到的水汽混合比转化为相对湿度后,取自然对数对反演精度有 12% 以上的提升。

关键词 特征向量法;附加影响因子;红外高光谱;主成分分析;遥感

中图分类号: P407.6 **DOI:** 10.19517/j.1671-6345.20190060 **文献标识码:** A

引言

大气温湿廓线在辐射传输过程、层结稳定度、降水过程以及云的形成和演变过程中具有重要的作用^[1],大气底层的边界层作为地球大气动量、热量和各种物质上下输送的重要通道,受地球下垫面的影响较大,其对表面强迫力的响应时间约为 1 h 或者更短^[2],对人们的生产生活有着直接的影响。因此,更高时间分辨率的温湿廓线观测资料对于研究边界层大气中的中小尺度运动以及提高数值预报的时效性有着重要的意义。

遥感探测作为一种新兴的探测方式,具有时间分辨率高、资料一致性好、探测成本低等诸多优势,可以很好地弥补探空仪探测方式的不足。其中,地基红外高光谱辐射计(Atmospheric Emitted Radiance Interferometer,简称 AERI)^[3],能够以被动的方式获取精确的边界层温湿廓线,由于观测的是下行红外辐射,受地面的影响较小,可以很好地弥补天基遥感在近地面层精度较差的缺陷^[4],在边界层的探测精度更高;此外,AERI 观测的辐射数据分辨率

可以达到 1 cm^{-1} ,包含有丰富的吸收谱线,相比于微波遥感理论上能够获得更加精确的温湿廓线信息。基于上述优点,地基高光谱遥感获得了越来越多的关注。

AERI 目前自带的反演系统 AERIprof 采用了基于“剥洋葱”算法,能够提供高时间分辨率的大气温湿廓线资料。但是受限于反演问题的病态特性,“剥洋葱”算法非常依赖于初值问题的准确性,尤其是最底层大气初值的准确程度^[5]。统计反演法由于反演速度快、稳定性好的特点,常常扮演提供初值的角色。因此,无论是从为物理反演提供初值的角度,还是直接用于高光谱遥感的输出结果,统计反演方法精度的提升对地基高光谱遥感均有着至关重要的影响。

特征向量法作为一种常用的统计反演方法,已经在 AERI 和天基高光谱遥感中得到了应用。但是特征向量法本质是多元线性回归关系,对辐射传输方程的非线性特性考虑不足,导致其反演精度不高。为提高统计反演法的反演精度,Shinhoo Kang 等人将水汽混合比廓线取对数并与大气辐射建立回归关

<http://www.qxkj.net.cn> 气象科技

作者简介:李宛桐,女,1991年生,硕士,主要从事大气物理与大气探测方面的研究工作,Email:leona0804@live.com

收稿日期:2019年2月1日;定稿日期:2019年8月14日

系,相比于 AERI 中直接对相对湿度反演,其均方根误差提高了大约 0.2 g/kg ^[6]; Feltz 等将地面湿度信息加入统计反演中,指出引入的湿度信息对物理反演法在 1500 m 至 3000 m 处有提升,但并未分析地面湿度对统计反演的影响^[7]。我国早在 20 世纪 70 年代就开始了利用下行红外辐射数据反演大气温湿廓线的研究。曾庆存院士给出了红外遥感反演大气温度廓线的原理和方法,在理论上阐述了地基遥感与天基遥感的巨大区别:不同于天基遥感,地基红外遥感大气温度时权重函数不会在大气的某个高度达到极大值,此时无法通过权重函数对大气定位求取温度廓线^[8]。在特征向量反演法方面,我国学者的研究主要集中在天基领域,其中官莉分析了 AIRS 卫星亮度温度分类和扫描角分类对算法精度的影响^[9];李俊等将地面湿度和地面温度引入到牛顿非线性迭代法中,以期提升星载遥感在近地面的反演精度^[10];蒋德明等研究了微波探测通道、地面信息等附加影响因子对特征向量反演法的改进作用^[11];黄静等^[12-13]发展了一种物理统计反演算法,提升了算法的稳定性。

针对 AERI 中特征向量反演法存在的精度有待提升的问题,借鉴卫星遥感中采用的相关改进方法,通过增加地面气象要素等非红外观测影响因子,进一步改善反演结果;此外,为减少水汽的时空变化对反演结果产生的影响,本文分析了湿度廓线表达形式对特征向量反演法的影响。

1 基本原理

假定大气参数与辐射值之间满足线性关系

$$\mathbf{X} = \mathbf{C}\mathbf{Y}^T \quad (1)$$

其中, \mathbf{X} 是 $N_p \times N_t$ 维, N_p 表示大气分层数, N_t 是训练样本个数; \mathbf{C} 是回归系数矩阵,它表征了大气参数与辐射值之间的线性关系; \mathbf{Y} 是仪器的观测值; \mathbf{T} 表示对矩阵的转置。

特征向量法反演大气廓线是基于最小二乘法,求解最合适的矩阵 \mathbf{C} 使得目标函数 $\sum (\mathbf{X} - \mathbf{C}\mathbf{Y}^T)^2$ 取得极小值。为此,将目标函数对回归系数 \mathbf{C} 求偏导,使其一阶偏导数等于零,得到系数矩阵^[14]

$$\mathbf{C}(i, n) = \mathbf{X}^T \mathbf{Y} [\mathbf{Y}^T \mathbf{Y}]^{-1} \quad (2)$$

之后利用式(2)对实时的地基红外高光谱辐射观测数据进行反演,得到大气的温湿廓线。

由于高光谱仪器通道数量较多,各通道之间的

观测信息有很大的相关性,造成数据不仅包含有丰富的的大气温湿廓线信息,同样包含有较多的冗余信息及相关噪声,直接利用式(2)反演会使反演结果产生很大的震荡。目前,常用的方法是对高光谱辐射数据做主成分分析,提取出其中占主要部分的特征向量(又称为主成分),将观测辐射投影至提取的主成分上,从而计算出投影系数

$$\mathbf{A}_{tr} = \mathbf{Y}_{tr} \mathbf{U} \quad (3)$$

其中, \mathbf{Y}_{tr} 表示训练集中的观测辐射, \mathbf{U} 表示提取的主成分。

之后将投影系数代替式(2)中的观测辐射,得到新的回归系数

$$\mathbf{C} = \mathbf{X}_{tr} \mathbf{A}_{tr} (\mathbf{A}_{tr}^T \mathbf{A}_{tr})^{-1} \quad (4)$$

可以看到,该方法不仅能够降低数据的维度从而减少计算量,更重要的是能够去除数据中大量的冗余信息及部分非相关噪声^[15],可以明显地提升算法的精度及稳定性。

2 试验数据及试验分析

2.1 试验数据

本文使用的数据是部署在南部大平原(South Great Plain,简称 SGP)站点晴空条件下的 AERI 观测辐射数据和探空数据。其中 AERI 是在 HIS 基础上研发的地基红外傅里叶变换光谱仪,该仪器具有方便运输、鲁棒性强的优点,已经在 ARM 计划中的固定站点和分布在德国、葡萄牙、中国、印度、尼日尔、加利福尼亚等移动站点进行了部署^[16]。AERI 以被动的方式自动地获取 $520 \sim 3020 \text{ cm}^{-1}$ 波段的大气下行红外辐射,其中 $520 \sim 1800 \text{ cm}^{-1}$ 波段的红外辐射由碲镉汞(MCT)探测器获得, $1800 \sim 3020 \text{ cm}^{-1}$ 波段的下行红外辐射由碲化铟(InSb)探测器获得,光谱分辨率小于 0.5 cm^{-1} 。

湿度廓线的反演一般采用水汽敏感通道,温度廓线的反演一般采用含量稳定的气体成分(如 CO_2) 的敏感波段。在反演过程中使用的通道与 AERI 中保持一致,具体波段如表 1 所示^[17];其中反演温度

表 1 AERI 反演温湿廓线使用的波段 cm^{-1}

温度	水汽
612~618	538~588
624~660	1250~1350
674~713	
2223~2260	

所使用的波段包含 244 个通道,水汽波段包含 311 个通道。D. D. Turner 等指出,如果训练数据样本个数小于使用的通道数目,那么观测数据中的噪声以及冗余信息将会影响到特征向量的取向,训练样本个数大于 2 倍通道数量最为稳妥^[18]。为获得充足的训练样本,从 2006~2012 共 7 年的探测数据中选取晴空样本数据。对于晴空的判别参考 Shinho Kang 等所使用的方法,即根据同站点释放的探空仪获取的湿度廓线,在整层大气湿度均没有超过 80% 时,将其判定为晴空^[6]。探空数据使用的是 SGP 站点 17:30 UTC 的温湿廓线,根据上述晴空判定方法找出了 933 个晴空样本,与同时刻的辐射数据组成训练样本对。待反演的数据选取的是 2013 年 SGP 站点同时刻的辐射数据和探空数据,按照上述方法寻找出 70 组晴空样本数据。

2.2 反演试验及分析

由于 AERI 观测辐射中关于大气温湿廓线的信息 90% 集中在 3 km 以下,因此其反演高度一般不超过 3 km^[19]。本文在反演时,将地面至 3 km 高度的大气分为 19 层,且分辨率成逐层递减的形式。此外,要加入地面信息等附加影响因子,数值大小与红外辐射存在明显差异。为减小奇异值带来的影响,需要对影响因子做标准化处理。对原始数据的标准化处理会改变变量的方差,使得基于原始数据和基于标准化数据所得到的主成分产生巨大变化^[20],因此不同的数据标准化方法会得到不同的主成分。经试验研究发现,标准化得到的反演精度最高。

$$y = (y_{\max} - y_{\min})(x - x_{\min}) / (x_{\max} - x_{\min}) + y_{\min} \quad (5)$$

其中, x 是原始数据, y 是标准化后得到的数据, y_{\min} 和 y_{\max} 是设定的参数。

2.2.1 主成分个数选取

在用主成分分析法对观测数据降维时,主成分个数 p 的选取对反演结果有着至关重要的影响: p 取值过大,会有较多的冗余信息保留下来,冗余信息越多,反演的温湿廓线震荡程度就越大; p 取值过小,则无法将足够的有效信息用于温湿廓线的反演,导致反演精度的下降。通常使用累计贡献率准则选取主成分个数,当累计贡献率达到 80% 或 85% 以上时,所对应的主成分个数 p 即为所选取的主成分个数^[21]。但在实际的反演试验中发现,经主成分分解

后得到首个特征值的贡献率已经在 80% 以上,前两个特征值的累计贡献率已经在 94% 以上,但是取前两个主成分进行反演的精度远远达不到要求,此时利用累计贡献率的阈值进行判别不再适用,本文通过分析反演廓线的精度、震荡特性以及特征向量的特点来综合判断最优的主成分个数。

研究发现,在震荡特性方面温度廓线与湿度廓线的反演呈现出很大的差异性:反演的温度廓线在主成分个数超过 6 时开始出现震荡,在取 7 时呈现出了一定的震荡特性,在取 8 时震荡加剧无法使用,因此温度廓线的反演呈现出明显的震荡特性且由稳定到震荡的过渡区间特别狭窄;而水汽廓线的反演呈现出相反的特点,随着主成分个数的增加并没有出现震荡显著加剧的情况,即水汽的反演由稳定到震荡的区间特别宽泛。

为刻画反演精度随主成分个数的变化,将整层大气的均方根误差(Root Mean Square Error,简称 RMSE)取平均值,得到了平均 RMSE 随主成分个数的变化图像。如图 1 所示,随着主成分个数的增加,温度廓线的平均 RMSE 值逐渐减小,主成分个数取 8 时达到最优,但是实际反演过程中取 8 时反演得到的 RMSE 廓线图像震荡较为强烈,不适合反演,因此取 7 较为合适。对于湿度廓线的反演,如图 2 所示,在主成分个数取 6 和 14 时平均 RMSE 图像出现了两个极小值:取 6 时为 1.25 K,取 14 时为 1.24 K。由于二者精度上差别很小,出于降维的考虑,取前 6 个主成分进行湿度廓线的反演更为合理。

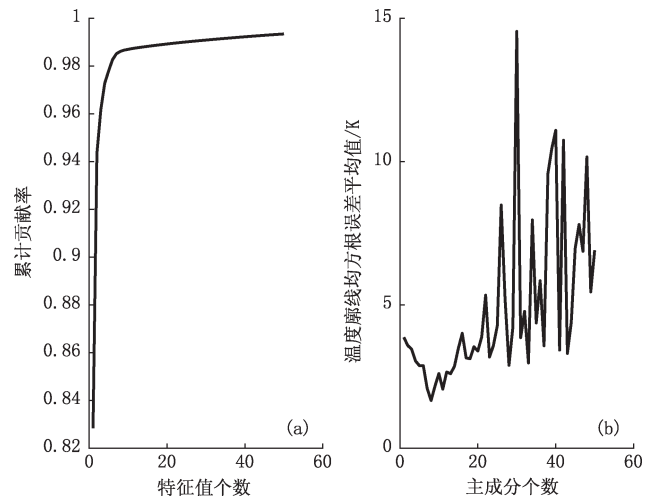


图 1 反演温度廓线时累计贡献率随特征值个数变化 (a) 及平均均方根误差随主成分个数变化 (b)

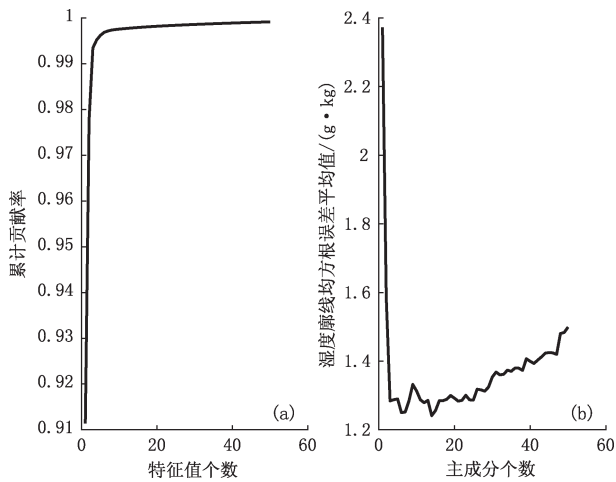


图 2 反演湿度廓线时累计贡献率随特征值个数变化 (a) 及平均均方根误差随主成分个数变化 (b)

主成分分析法是利用少数相互独立变量(即主成分)的线性组合来表征原始的多维数据,以此达到

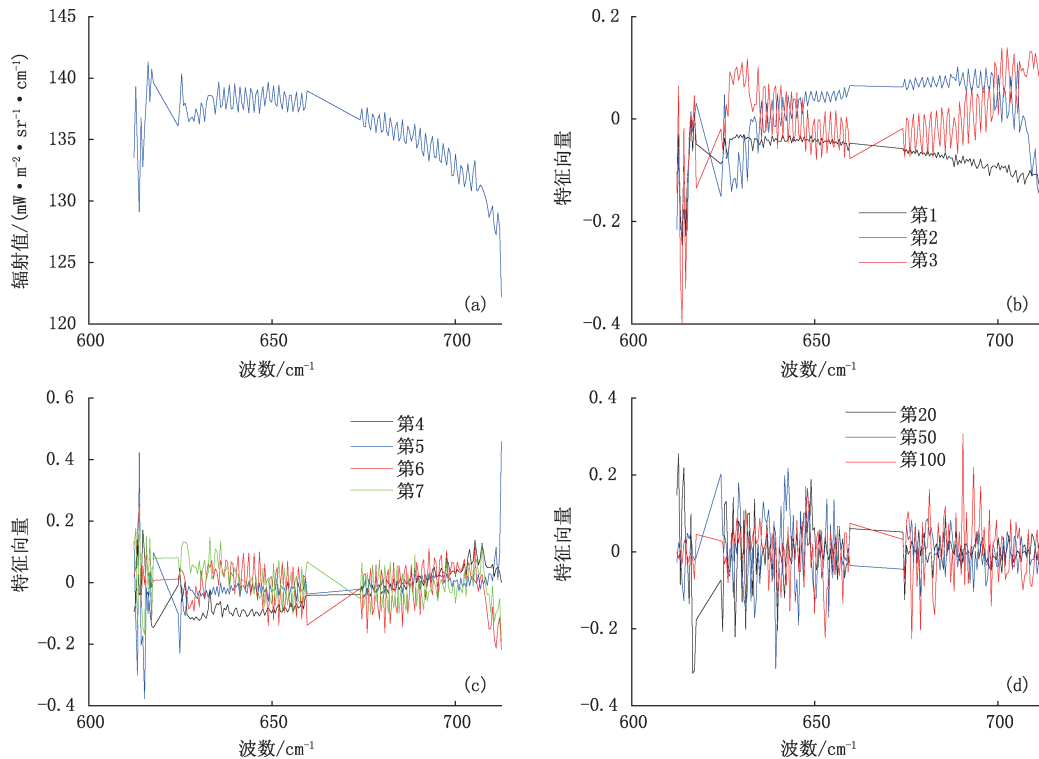


图 3 612~713 cm^{-1} 波段大气平均辐射值(a)和各特征向量(b~d)

同样地,在反演湿度时排序靠前的特征向量表征平均辐射值的信息,尤其是前 3 个特征向量,呈现出不同的大小和走势,分别表征了辐射数据的不同特点,如图 4 所示。在第 7 特征向量之后,各向量之间呈现出高度一致的特点,因此第 7 个以后的特征

降维、去除冗余信息的目的。图 3 是部分温度反演通道的观测辐射分解得到的特征向量图像,其中图 3a 表示辐射的平均值信息,图 3b、c、d 为不同序号的特征向量图像,可以看出,前 7 个特征向量的走势、均值、震荡等特性存在很大的不同,从形态上来看,前 5 个特征向量主要表征了大气辐射数据的平均值信息,尤其是第 1 个特征向量,其走势与辐射的平均值基本一致;相比于前 5 个特征向量,第 6 个特征向量的振动开始加大,表征了原始辐射中的震荡部分;第 7 个特征向量之后,随着序号的增加震荡开始加大,但是走势大致保持一致,即包含的信息含量相差无几,可以推断第 7 个之后的特征向量表征了观测数据中的冗余信息和噪声信息。因此,前 7 个特征向量中既包含有辐射数据的走势等主体信息,也包含有原始数据中的震荡信息,表明选用 7 个特征向量反演温度廓线是较为合理的。

向量主要包含的是冗余信息。因此,从特征向量的角度来考虑,取 7 个特征向量进行反演是更为合理的。从反演精度来看,取 6 个主成分反演得到的平均 RMSE 值为 1.25 K,取 7 时得到的平均 RMSE 值为 1.252 K,在精度上二者差别很小。综合考虑

反演精度以及各主成分包含的信息,本文在反演湿度廓线时主成分个数选为 7。此外,湿度反演中的特征向量并没有随着序号的增加震荡逐渐加大,这

一点和温度廓线的反演有很大不同,也从侧面证明了水汽反演时过渡区域较为宽泛的特点。

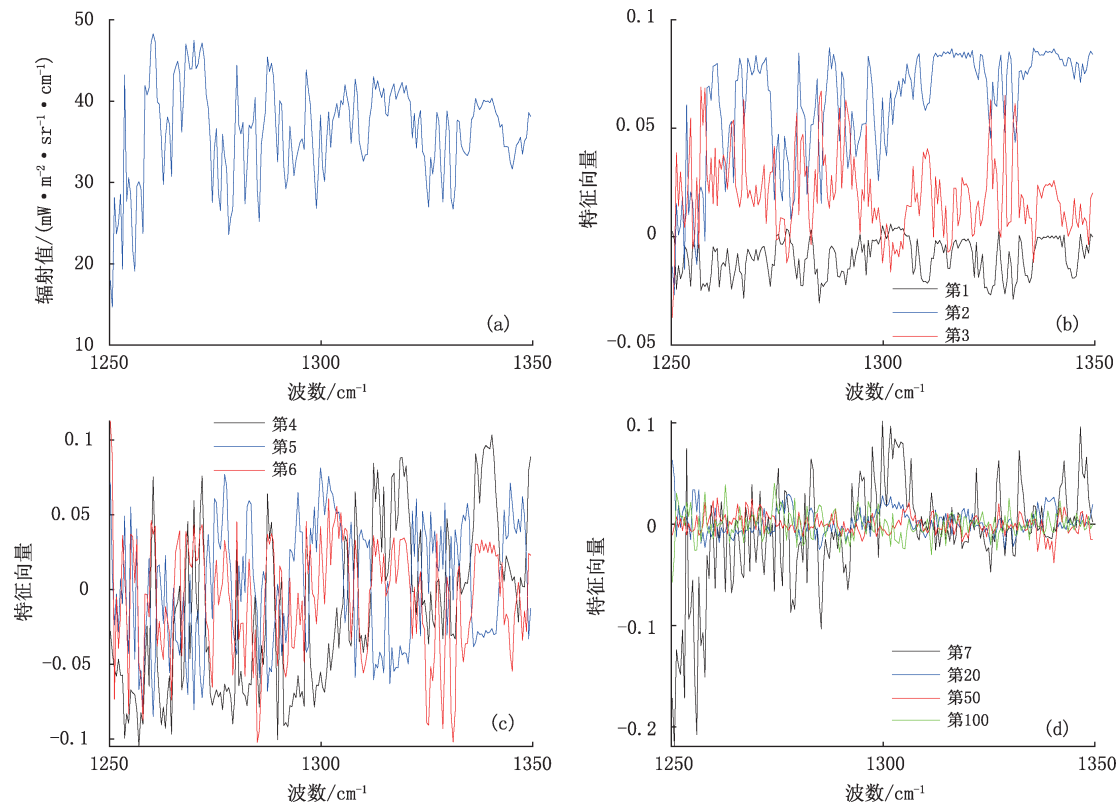


图 4 1250~1350 cm^{-1} 波段大气平均辐射值(a)和各特征向量(b~d)

2.2.2 地面气象要素的影响

在原有的红外通道基础上,增加地面气压、温度、湿度作为影响因子,经标准化处理后再进行主成分分解。为分析每个气象要素对反演产生的影响,分别将温度、气压、湿度融入到红外辐射数据中进行反演,结果如图 5 所示。对于温度廓线的反演,由于反演结果对主成分个数极为敏感,地面信息的引入使得待反演参量的数据结构发生改变,导致最优主成分个数产生差异。加入地面气压信息后,温度廓线的反演在主成分取 4 时精度达到最高,之后随着主成分个数的增加开始出现剧烈震荡,进一步压缩了由稳定到震荡的过渡区间;温度和湿度的引入对最优主成分个数影响不大,在取 6 时达到最优的反演精度。

从反演精度来看,地面温度的引入对温度廓线反演精度的提升效果最好,但是震荡过程加剧;相反,引入地面气压带来的精度提升稍逊于地面温度,但是其

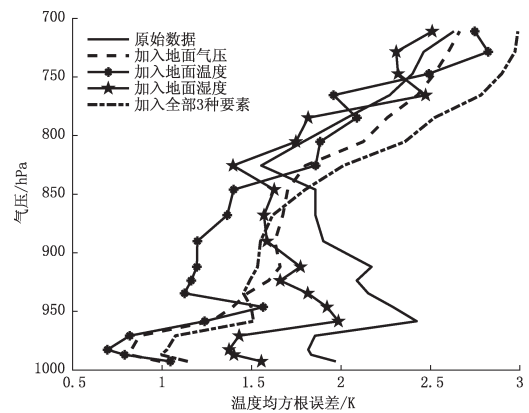


图 5 增加地面气象要素的温度廓线反演试验均方根误差图像

图像更为平滑,其提升作用主要集中在 850 hPa 以下高度,随着高度的降低提升作用越明显,最高达到了 1.5 K;引入地面水汽信息对温度反演精度有一定提升,但是提升的幅度最小;将所有地面气象要素

放入红外辐射数据中,其反演图像呈现出气压要素反演中平滑的特点,但是精度在 850 hPa 以上下降明显,由此可见,引入更多的地面气象要素并没有增加有利的信息。

基于上述分析,将地面气压加入到 SGP 站点 9 月 25 日的 178 组 AERI 辐射数据中进行反演。SGP 站点每天共有 4 组探空数据,分别是 05:30、11:30、17:30 和 23:40 UTC。在反演全天的光谱数据时,为消除不同时刻训练数据的差异对反演结果造成的影响,将 4 个不同时刻的晴空数据全部放入训练数据集中进行预测,温度廓线的反演结果如图 6 所示。其中,左侧图是对全天的光谱数据反演得到的温度廓线随时间变化图像,右侧图是同一天的探空数据以及相应时刻的反演结果,不带点缀的为探空廓线,带点缀的是反演廓线。从右侧图像来看,反演结果能够很好地反映 05:30 UTC 和 11:30 UTC 探空数据中存在的逆温信息,表明算法在反演晴空大气时具有较高的精度;从时序图上来看,在 15:00 UTC 之前存在着明显的逆温现象,15:00 UTC 之后逆温现象消失。

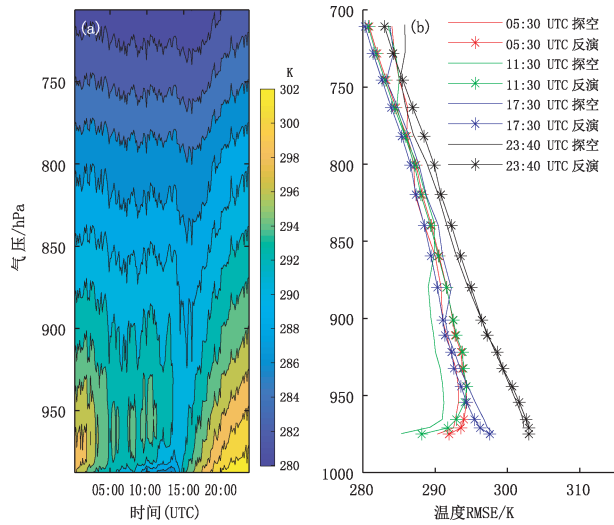


图 6 2013 年 9 月 25 日反演的温度廓线随时间变化(a)及误差(b)

对于湿度廓线的反演,地面信息的引入对主成分个数的选择影响较小,仍然在取 7 时达到最优,且过渡区间同样较为宽泛(图 7)。与温度廓线的反演图像对比来看,地面湿度引入的提升效果反而不如温压,即同类别地面气象要素的提升效果不明显,这

一点与温度廓线反演有很大区别。此外,地面温度和压强在湿度廓线的反演中呈现出高度的相似性,引入地面气压和温度后的湿度反演廓线几乎重合,反演精度的提升主要在 800 hPa 以下高度且随着高度的降低改进程度越明显,最高达到了 0.42 g/kg。同样地,引入全部地面气象要素相比于单一气象要素并没有提供额外的有用信息,其精度相比于地面温压反而有所下降。图 8 是 SGP 站点 9 月 25 日 AERI 辐射数据中引入地面气压后对大气湿度的反演结果,可见在 15:00 之后大气的水汽混合比逐渐增大,这一点和温度廓线的变化是吻合的。

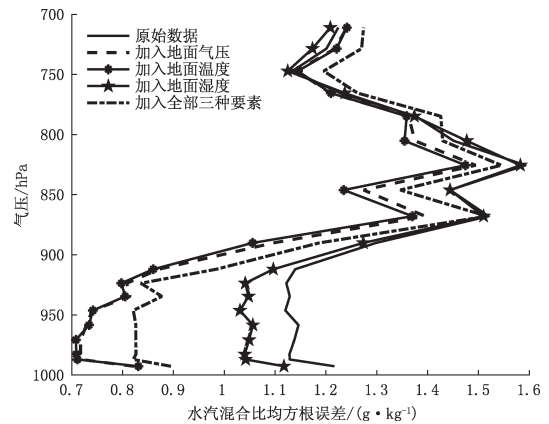


图 7 增加地面气象要素的湿度廓线反演试验均方根误差

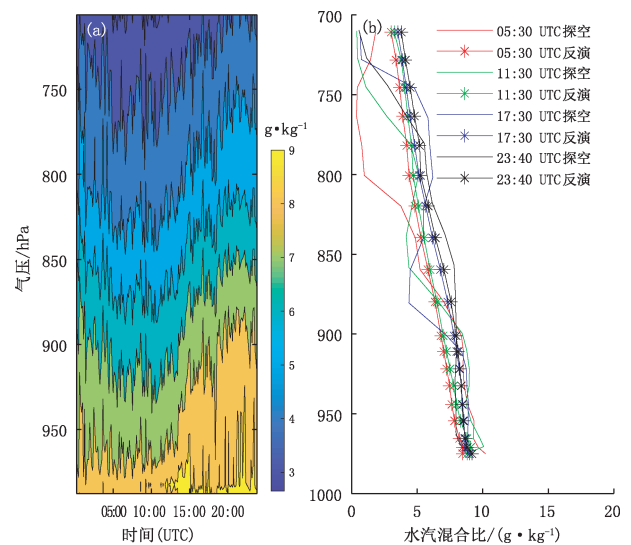


图 8 2013 年 9 月 25 日反演的湿度廓线随时间变化(a)及误差(b)

2.2.3 对数反演形式的影响

在 Shinhoo Kang 的工作中,仅分析了水汽混合

比的对数反演形式与直接对相对湿度反演的差异,目的是为了规避非物理因素导致的相对湿度大于 100% 或者小于 0 等非合理值,但是对水汽混合比取对数的益处不只如此。由于大气中的水汽分布随时空变化较大,特征向量法本质上是一种多元线性回归方法,无法较好地反映水汽含量的非线性变化;对水汽混合比取对数可以显著减小水汽含量随高度的变化量级,更加符合特征向量反演法的特点。

为分析取对数对反演结果产生的影响,分别进行了原始辐射数据和取自然对数后的辐射数据的反演试验;此外,由于目前业务运行中大多数探测设备的产品结果为相对湿度,又将反演得到的水汽混合比转化为相对湿度进行对比,结果如图 9 所示。其中图 9a 是水汽混合比的对比图像,图 9b 是转化为相对湿度后的均方根误差。可以看出,以水汽混合比形式反演时在 900 hPa 以上高度取自然对数后的反演精度有着一定程度的改进,尤其是对反演高度的顶层有着显著的改进;在底部区域反演精度有一定的下降,但整体的波动性有所减小。转化为相对湿度后,两种均方根误差图像在形态上高度吻合,但是在反演精度上,将水汽混合比转化为相对湿度后相比于原始数据有 12% 以上的提升。通过试验还发现,取自然对数和常用对数得到的均方根误差图像几乎完全重合,表明取对数的方式对反演结果没有影响。

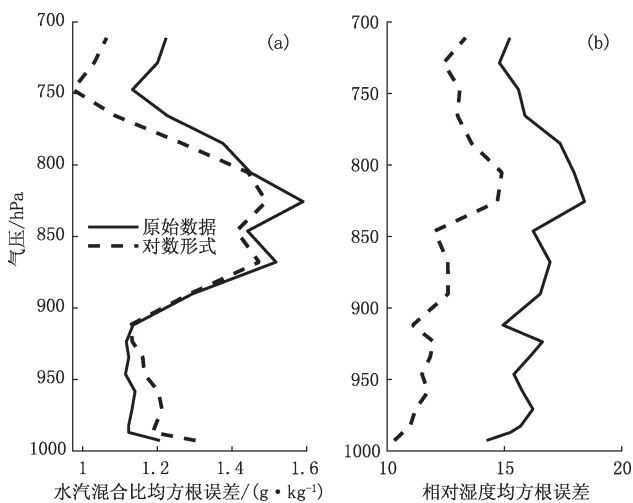


图 9 取自然对数后的湿度廓线反演试验均方根误差:

(a)水汽混合比,(b)相对湿度

3 结论与讨论

本文针对特征向量反演法中存在的问题,提出利用地面气象要素作为附加影响因子引入至辐射数据中,为地基高光谱反演提供了更多有效的信息;此外,对大气水汽混合比廓线取对数进行反演,抑制了湿度廓线在大气中的非线性变化。为验证上述措施的改进效果,根据美国 SGP 站点 2006—2012 年间晴空条件下的 AERI 观测辐射数据和探空数据建立模型,利用 2013 年同站点的 70 组晴空数据进行实验验证。

(1) 试验结果显示,温度廓线与湿度廓线的反演呈现出不同的特点:随着主成分个数的增加,温度廓线的反演呈现出明显的震荡特性且由稳定到震荡的过渡区间特别狭窄;湿度廓线的反演则较为稳定,没有出现剧烈震荡的情况。在反演精度方面,随着主成分个数的增加,二者的均方根误差均呈现出先减小后增大的趋势,即存在最优的主成分个数使得反演结果达到最优。

(2) 地面气象要素的引入对反演精度和稳定性有着明显的改善作用,这种改善作用主要表现在边界层中下部,且越靠近地面改进作用越明显。对于温度廓线的反演,地面温度的引入对反演精度的提升最大,但是反演廓线的震荡有所增加,引入地面气压则能很好地兼顾反演精度与稳定性;对于湿度廓线的反演,地面温度和压强的引入对反演精度和稳定性的提升最强,且二者在湿度廓线反演中的作用基本一致;无论是温度廓线的反演还是湿度廓线的反演,引入全部的 3 种地面气象要素均不能提供更多的有效信息。因此在地基红外高光谱反演中,增加地面气压作为影响因子不失为一种改善反演精度和稳定性的有效方法。

(3) 对水汽混合比取自然对数反演能够抑制水汽含量的剧烈变化对反演结果产生的影响,通过试验发现,以水汽混合比的形式进行反演时,在 900 hPa 以上高度的反演精度有着一定程度的改进,整体改进幅度较小;将反演得到的水汽混合比转化为相对湿度后,反演精度有了大幅提升。由于在业务运行时获取的大都为相对湿度,因此取对数反演仍然具有很重要的意义。

虽然上述方法对反演结果的精度以及稳定性有一定程度改进,但仍然存在以下不足。首先,地基红外遥感的权重函数分布较低,大都在 1 km 以下,无

法为边界层上层提供足够的有效信息,导致 3 km 附近大气廓线的精度无法获得有效的提升,下一步拟通过对大气廓线做经验正交展开来减少待反演的参数,从而提高反演的精度和稳定性。其次,由于缺少全天空测云仪等判别晴空的辅助设备,在挑选晴空大气廓线时,将整层相对湿度均小于 80% 的判别为晴天,这种方法虽然能保证获得的廓线均为晴空数据,但是条件过于苛刻,使得一部分晴空数据被剔除掉,不利于业务运行。为研究一种更加准确的无需额外辅助设备的晴空判别方法,拟利用 AERI 观测的红外高光谱数据判定晴空,选取湿度敏感通道和地面气象要素信息作为综合指标,通过同站点的云图图像将云量信息量化,利用神经网络模型建立综合指标与云量的相关关系。最后,红外波段受云的影响较大,这就造成了云下温湿廓线反演精度的下降以及无法获取云上的温湿廓线信息,而微波波段由于波长较长,受云的发射以及散射的影响大为减少,因此引入微波通道理论而言可大为改善有云时温湿廓线的反演精度,这也是下一步的工作重点。

参考文献

- [1] Wulfmeyer V, Hardesty R, Turner D D, et al. A Review of the Remote Sensing of Lower-Tropospheric Thermodynamic Profiles and its Indispensable Role for the Understanding and the Simulation of Water and Energy Cycles [J]. *Reviews of Geophysics*, 2015, 53(3):819-895.
- [2] 张碧辉, 刘树华, 马雁军. MYJ 和 YSU 方案对 WRF 边界层气象要素模拟的影响 [J]. *地球物理学报*, 2012(7): 2239-2248.
- [3] Knuteson R O, Revercomb H E, Best F A, et al. Atmospheric emitted radiance interferometer. Part I: Instrument design [J]. *Journal of Atmospheric and Oceanic Technology*, 2004(12): 1763-1776.
- [4] Ebell K, Orlandi E, Hünerbein, A, et al. Combining ground-based with satellite-based measurements in the atmospheric state retrieval: Assessment of the information content [J]. *Journal of Geophysical Research*, 2013, 118(13):6940-6956.
- [5] Turner D D, Löhnert U. Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI) [J]. *Journal of Applied Meteorology and Climatology*, 2013, 53(3): 752-771.
- [6] Kang S, Goo T, Ou M. Improvement of AERI T/q retrievals and their validation at Anmyeon-Do, South Korea [J]. *Journal of Atmospheric and Oceanic Technology*, 2013(7): 1433-1446.
- [7] Feltz W F, Smith W L, Knuteson R O, et al. Meteorological applications of temperature and water vapor retrievals from the Ground-Based Atmospheric Emitted Radiance Interferometer (AERI) [J]. *Journal of Applied Meteorology*, 1998(9): 857-875.
- [8] 曾庆存. 大气红外遥测原理[M]. 北京: 科学出版社, 1974.
- [9] 官莉. 利用 AIRS 卫星资料反演大气廓线 I. 特征向量统计反演法[J]. *南京气象学院学报*, 2006(06):756-761.
- [10] 李俊, 曾庆存. 有云时大气红外遥感及其反演问题 II. 反演试验研究[J]. *大气科学*, 1997(4):13-17.
- [11] 蒋德明, 董超华, 曹思沁. 附加影响因子对红外遥感资料反演大气温湿廓线的辅助作用[J]. *热带气象学报*, 2009, 25(增刊 1):79-84.
- [12] 张蕾, 邱崇践. 利用卫星红外遥感资料反演大气温湿廓线的三维物理统计方法[J]. *红外与毫米波学报*, 2008(2): 109-114.
- [13] 黄静, 邱崇践, 张艳武. 一种利用卫星红外遥感资料反演晴空大气参数的物理统计方法[J]. *红外与毫米波学报*, 2007(2): 102-106.
- [14] Smith W L, Woolf H M, Jacob W J. A regression method for obtaining real-time temperature and geopotential height profiles from satellite spectrometer measurements and its application to NIMBUS 3 "SIRS" observations [J]. *Monthly Weather Review*, 1970, 98(98): 582.
- [15] Huang H, Antonelli P. Application of principal component analysis to high-resolution infrared measurement compression and retrieval [J]. *Journal of Applied Meteorology*, 2001(3): 365-388.
- [16] Jian Y. Retrieval of temperature and water vapor from combined satellite and ground based ultra-spectral measurements [D]. Hampton University, 2013.
- [17] Smith W L, Feltz W F, Knuteson R O, et al. The retrieval of planetary boundary layer structure using ground-based infrared spectral radiance measurements [J]. *Journal of Atmospheric and Oceanic Technology*, 1999(2-3): 323-333.
- [18] Turner D D, Knuteson R O, Revercomb H E, et al. Noise reduction of Atmospheric Emitted Radiance Interferometer (AERI) observations using principal component analysis [J]. *Journal of Atmospheric and Oceanic Technology*, 2006(9): 1223-1238.
- [19] Turner D D, Löhnert U. Information content and uncertainties in thermodynamic profiles and liquid cloud properties retrieved from the ground-based Atmospheric Emitted Radiance Interferometer (AERI) [J]. *Journal of Applied Meteorology and Climatology*, 2014(3): 752-771.
- [20] 魏登云, 张文俊. 主成分分析法应用中原始数据的标准化辨析[J]. *安徽师范大学学报(自然科学版)*, 2018, 41(2): 189-194.
- [21] 李广正. 关于选取主成分个数的探讨[D]. 兰州: 兰州财经大学, 2015.

Impact of Additional Surface Meteorological Elements on Thermodynamic Profiles Retrieved by Ground-Based Hyper-Spectral Spectrum Based on Eigenvector Regression Algorithm

LI Wantong¹ HUANG Wei² JIANG Ming¹ SHI Jing¹

(1 Tianjin Meteorological Observation Center, Tianjin 300061;

2 Luoyang Electric Equipment Test Center of China, Henan, Jiyuan 459000)

Abstract: The initial profile, which is calculated by the eigenvector regression algorithm in the ground-based hyper-spectral remote sensing, has a significant impact on the accuracy of the physical retrieval. Based on the eigenvector regression algorithm, the similarities and differences between the retrievals of temperature profiles and water vapor profiles are analyzed using the radiance data observed by AERI and coincident radiosonde profiles. The optimal number of principal components is analyzed when retrieving the temperature and water vapor profiles. Considering both the accuracy and the information contained in the eigenvectors, the optimal numbers of principal components are both set to 7. In order to improve the accuracy of remote sensing, the surface temperature, humidity and pressure are introduced as the influence factors. The experiment results show that the introduction of surface pressure has a better performance than the other two single meteorological elements and the assemble of factors composed of all three meteorological elements, especially for the accuracy and stability of temperature and water vapor profiles in the middle and lower parts of the boundary layer. With the decrease of altitude, the RMSE of temperature profiles decreases to a maximum of 1.5 K, and the RMSE of water vapor profiles decreases to a maximum of 0.42 g/kg. At the same time, the impact of logarithmic water vapor mixing ratio on the retrieval of water vapor profiles is analyzed. The result shows that the introduction of the logarithmic profile has little effect on the accuracy of the retrieval. However, the accuracy of the water vapor profile gains more than 12% when converting mixing ratio to relative humidity.

Keywords: eigenvector regression algorithm; additional predictor; infrared hyper spectral; principal component analysis; remote sensing