

天气预报 技术论文集 (1998)

中国气象局业务发展与天气司



气象出版社

天气情报 技术论文集

(1988)

中国气象局编 气象出版社



天气预报技术论文集 (1998)

中国气象局业务发展与天气司

气象出版社

内容提要

本书收入了1998年在湖南长沙召开的“1997年重大灾害性天气过程总结和预报技术经验交流会”上交流的论文47篇,内容涉及到:暴雨、冰雹、台风等灾害性天气过程的个例分析,一些新理论在预报中的应用,数值预报产品的解释应用,新的业务流程的设计,评分管理系统的建立及预报业务管理的经验总结等。

本书可供气象、水利、航空等部门从事天气预报的科技人员和高等院校相关专业师生参考。

图书在版编目(CIP)数据

天气预报技术论文集:1998/中国气象局业务发展与天气司编.-北京:气象出版社,1999.5

ISBN 7-5029-2710-7

I.天… II.中… III.天气预报-文集 IV.P45-53

中国版本图书馆 CIP 数据核字(1999)第16314号

天气预报技术论文集(1998)

中国气象局业务发展与天气司

责任编辑:庞金波 俞卫平 终审:周诗健

责任校对:杨波平 责任技编:吴庭芳

气象出版社出版

(北京西郊白石桥路46号 邮编:100081)

北京宏远兴旺印刷厂印刷

新华书店总店北京发行所发行 全国各地新华书店经销

开本:787×1092 1/16 印张:14 字数:360千字

1999年5月第一版 1999年5月第一次印刷

印数:1~800

ISBN 7-5029-2710-7/P·0952

定价22.80元

编者的话

中国气象局业务发展与天气司于1998年4月7~9日在湖南省长沙市召开了“1997年全国重大灾害性天气过程总结和预报技术经验交流会”。会议共收到论文和总结240篇,经过专家初审和评选,本文集录用的共47篇;获1997年全国重大灾害性天气过程总结和预报技术经验交流会优秀奖的论文和总结共10篇;综合预报方法的共7篇;数值分析预报产品解释应用的共14篇;热带气旋预报的共6篇;气候预测的共6篇;业务系统建设和管理的共4篇。由于时间仓促,水平有限,不妥和错误之处在所难免,望广大读者批评指正。

编者

1998年11月

目 录

编者的话

第一部分 1997年全国重大灾害性天气过程总结和预报技术经验交流会优秀论文和总结

1997年中期尺度集合预报的试验应用	李小泉等(1)
不确定性理论——集对分析在 MOS 概率天气预报中的应用	王国强(6)
GMS-5分裂窗通道在降水估计方面的应用前景展望	卢乃锰 郑新江(13)
建立在 SGI 工作站 MICAPS 工作平台上的宁夏中、短期天气预报制作流程	吴岩峻等(20)
云南省自动化预报评分管理系统建立及其业务化	周云 刘劲松(25)
天气图 R. T. E. 客观预报技术及业务系统	古德 刘勇(27)
T106中期数值预报产品的性能检验和应用	万玉明 尹新怀 肖青(31)
HLAFS 资料对山东暴雨的诊断分析	闫丽凤 于仁成 迟淑芹(35)
TOVS 信息处理及其在西太平洋副热带高压分析预报中的应用	沈树勤等(40)
林火多因子综合分析及预报方法	施望芝(46)

第二部分 综合预报方法

中低层环流季节变化与江淮梅雨	陈永秀(51)
GMS-5水汽图像特点及其在梅汛期暴雨预报中的应用	朱凌云 钱鹰 岳擎宇(55)
贵州1997年7月初大暴雨过程的 α 中尺度 Q 矢量诊断分析	武文辉 田淑珍(60)
溃变理论在冬季一场异常冰雹过程中的应用	黄元森 黄光明(66)
江西省春末夏初强对流回波与地形的关系	杨梅(70)
淮河防洪减灾业务系统的建设与管理	徐慧(75)
应用区域气候参数定量预报农业年景温度	冯学洪(78)

第三部分 数值分析预报产品解释应用

夏季 T106与 EC 中期预报能力的对比分析	晁淑懿 鲍媛媛(83)
概率变换法在 MOS 预报中的应用	周信禹(91)
T106数值预报产品在赣南温度预报中的释用	钟名森 谢勇 杜世逊(94)
应用中国数值预报产品做基本要素预报的业务系统简介	胡欣(98)
螺旋度构造量在1997年沿江区域暴雨预报中的应用	王新生 张苏 李厚楹(103)
数值预报产品省级降水动力相似统计释用系统	杨小萍等(107)
暴雨及其落区预报的动力、热力参数统计释用研究	贺哲等(111)
数值预报在1997年10月23日区域寒潮预报中的运用分析	王中(116)
用卡尔曼滤波法做中期时效的逐日极端气温预报	孙欣 张凯(119)
濮阳市汛期大-暴雨短期分县指导预报业务系统	杜瑞莉等(123)

利用 T106 数值预报产品做江西暴雨动态落区预报	曹晓岗(127)
卡尔曼滤波方法应用技术	刘凤辉(131)
降水分级客观预报方法研究	李玉华等(135)
黄山旅游气象景观的概率预报	江斌(138)

第四部分 热带气旋预报

9711 号台风大暴雨诊断分析	汤桂生等(142)
9711 号台风路径与登陆后的风雨成因分析	黄玲琳(148)
9711 号台风特大风暴潮预报及致灾成因探讨	施素芬(154)
影响和登陆福建省台风路径和台风暴雨分布关系	官秀珠(158)
9711 号台风造成苏北大暴雨过程的卫星云图和雷达回波拼图合成分析	白卡娃等(163)
9711 号台风天气特征分析	邓广华 杨洪清 王统霞(168)

第五部分 气候预测

1997 年 7 月 27 日青岛历史高温成因分析	刘珍芳 韩春深(174)
湖南夏季旱涝与北半球大气环流特征	罗伯良(177)
1997 年夏季我国异常天气的 TBB 场特征	江吉喜 范梅珠 吴晓(181)
近年来黄河断流的干旱背景及其对策	江剑民 游性恬 谷湘潜(187)
影响辽宁省 30~50 天低频振荡系统特征分析	曲晓波等(192)
冷暖天气过程与热量动态相似降水预报方法	王新红 马月枝 谢晋英(197)

第六部分 天气预报业务系统建设和管理

进一步完善业务管理体系建设的探讨	高锋 董礼仁(202)
加强对下技术指导 推进预报业务改革	濮梅娟等(206)
天气预报业务的实时监控管理	郑涛(211)
广西天气预报业务和管理进展情况	陈博杰(214)

1997 年中期尺度集合预报的 试验应用

李小泉 刘金达 汪迎辉 刘国强

(国家气象中心 100081)

摘 要

本文简要介绍了 1997 年国家气象中心进行集合预报试验的情况及对预报效果的初步检验。

关键词 集合预报 蒙特卡罗方法

1. 引言

近年来,国际上天气预报领域,特别是中长期天气预报领域,在理论与实践上的一个重大发展是集合预报(Ensemble)的研究与应用^[1,6~8]。

众所周知,大气是对初始场十分敏感的系统。由于观测站点分布不均以及观测资料和客观分析必然存在着误差,使得初始场不可能完全精确获得。在大气系统内在不稳定性的作用下,预报模式积分过程中初始场的微小误差会迅速增长,使得若干时间之后,大气运动变得具有明显的随机性。这就是说,即使用完全精确的模式,从稍有差别的初始场出发,积分到若干天之后,其预报轨迹也可能发散到较大的区域,这时天气预报就不应当是寻求确定性的预报结果而成为要对它的概率分布进行适当估计的问题了。由此观点出发,中长期天气预报,便是一个集合预报的问题,就是说,要从大气初始场的不确定性和大气系统的不稳定性出发,给出大气未来发展的概率分布。但是我们既没有完全精确的大气环流模式,也无法得到初始场的全部可能样本和对所有的样本逐一进行计算。因此,集合预报通常的做法是在实际观测和分析所得的初始场基础上,叠加适当小扰动,从而形成稍有差别的多个初始场,分别进行动力延伸预报,然后进行综合,形成多种集合预报产品,以获取更多有用的信息,从而较全面地反映大气未来可能的发展状态。

早期集合预报初值的形成,大多采用蒙特卡罗方法^[3],即在客观分析的基础上叠加随机小扰动,但效果并不理想。1983 年 Hoffmann 和 Kalnay^[4]提出了时间滞后平均预报(IAF)方法,即用间隔为 6 小时或 12 小时的多个初始场分别作预报,然后对相同时刻的预报结果求平均。试验表明,IAF 方法较蒙特-卡洛方法的效果有所提高。90 年代初,美国国家气象中心和欧洲中期天气预报中心(ECMWF)分别建立了各自的集合预报业务系统,并从 1992 年 12 月起开始投入业务运行,其初始场扰动初值的形成分别采用的是增长模繁育法(BGM)和奇异向量法(SV)^[1,7]。

此外,日本、英国、法国、加拿大等国家也纷纷建立了各自的集合预报系统并投入业务应用。

近年来,由于引进了高性能、大容量的计算机,国家气象中心的计算能力有了很大的提高,特别是 SP₂ 多 CPU 并行计算机的引进和安装调试成功,为进行集合预报的试验研究和业务应用创造了有利条件。从 1996 年下半年起,按照国家气象中心领导的布署与要求,数控室与气象台合作,进行了中期时间尺度的集合预报试验。本文对 1997 年集合预报的试验应用进行了初步的总结。

2. 方案设计与业务流程

为了能够尽快获得集合预报产品,在并行计算尚未实现之前,作为第一步,先将原 Cyber 机上的 T6316 模式移植到 SP₂ 计算机上,用时间滞后平均(IAF)方法进行集合预报试验。即以连续三天 00、06、12、18Z 的四维同化资料为初始场,在不同结点上用相同的模式分别将预报时效延伸计算到 13~15 天,再把预报时期相同的 12 个预报进行集合,得到若干集合预报产品,供中央气象台中期预报科的旬天气预报会商参考使用^[8]。

具体的业务流程是:

国家气象中心数控室系统运行科在 00、06、12、18Z 四个时次的 T63 模式四维同化初始场资料形成后,启动指定作业,经过数据码转换,实现不同机型之间数据的无误传输,将初始场资料由 CRAY 机传送到 SP₂ 计算机上。

为能参加每旬 9 日下午的中期科旬预报天气会商,初始场资料取每旬 6、7、8 日,此外,也用每旬 1、2、3 日的初始场作集合预报试验,以积累资料和供必要时使用参考。以上工作均由计算机自动判断启动。

每旬 4、9 日上午 12 个集合成员的动力延伸预报完成后,人工启动集合产品计算软件,在 SP₂ 机上生成集合预报产品,然后将集合预报产品传送到会商室的工作站或微机上,利用 MIPS 或 MICAPS 的绘图功能出图并将图形打印出来。

1997 年 3 月底,SP₂ 计算机完成了磁盘扩体和系统升级,我们又进一步调整和优化了业务流程,提高系统运行的稳定性和速度,改进和逐步增加集合预报产品,基本上做到了每个月定时生成 5~6 次 10 天集合预报,并提供给中期科实时旬天气预报会商参考。1997 年 4~12 月一共进行了 46 次 10 天集合预报的计算和试验。

3. 集合预报产品

如前所述,作集合预报时,要从多个稍有差别的初始场出发,分别进行动力延伸预报计算,这样便得到了大量预报信息。怎样将众多的信息浓缩,以简捷和便于应用的形式给预报员提供反映未来大气发展主要特征的集合预报产品,是集合预报业务应用面临的一个重要问题。

目前,我们提供的集合预报产品有如下几种:

(1)将 12 个集合成员的候平均 500hPa 高度预报场作算术平均,便可得到下一旬第一候和第二候 500hPa 高度集合平均的预报图。作为对比,也给出两个候的控制预报(以最近一天 12Z 为初始场的延伸预报)500hPa 平均预报图。

(2)同样的,将下一旬逐日的 500hPa 高度预报场作算术平均,得到下一旬逐日 12 个集合成员的 500hPa 高度集合平均预报图。

(3)为反映诸集合成员之间的一致性程度,计算了下一旬两个候 12 个成员的 500hPa 高

度预报场的标准化离散度分布,即每个格点上 12 个成员预报场的均方差除以该格点的气候标准差。

(4)根据不同季节确定 500hPa 高度场的特征等高线的数值(如冬季为 544 和 580dagpm,夏季为 560 和 588dagpm),将所预报的下一句两个候 12 个成员的特征等高线重叠在一起,便可得到下一句两个候的特征等高线集合预报图(图略)。

(5)与 500hPa 类似,用求算术平均的办法,计算下一句两个候 12 个成员的 850hPa 集合平均温度及距平预报图。

(6)以 -12°C 、 0°C 、 12°C 为特征等温线的数值,给出下一句两个候 12 个成员 850hPa 特征等温线的集合预报图。

(7)将下一句逐日 12 个成员的 850hPa 温度预报场求算术平均,然后把特征等温线重叠在一起,便可得到下一句逐日的 850hPa 集合平均特征等温线预报图(见图 1)。冬半年可用以反映冷空气活动过程的集合预报意见。

4. 预报精度和误差的初步分析

集合预报的产品很多,我们不可能一一加以验证。这里,我们仅把前述集合预报产品的第一项,即下一句两个候 500hPa 高度集合平均预报图作了初步的检验和分析,按北半球和东半球分别计算距平相关系数(ACC)。

表 1 是 46 次集合预报的平均 ACC,可见:

表 1 46 次集合预报试验平均的 500hPa 高度距平相关系数

系数 旬、候	集合预报		控制预报	
	东半球	北半球	东半球	北半球
下旬第一候	0.624	0.619	0.712	0.702
下旬第二候	0.319	0.297	0.303	0.261

(1)不同个例之间预报准确率差别很大,两个候均如此。

(2)集合预报与控制预报相比较,第一候多为控制预报优于集合预报,而第二候则多为集合预报优于控制预报。从 46 次预报试验的平均结果来看,第一候北半球 500hPa 高度集合预报(相当于 3~7 天到 5~9 天预报的平均)平均的 ACC 为 0.624,低于控制预报(相当于 3~7 天预报的平均);而第二候集合预报(相当于 8~12 天到 10~14 天预报的平均)平均的 ACC 为 0.319,高于控制预报(相当于 8~12 天预报的平均,ACC=0.303),平均高 0.016。但不同个例之间的差别也很大。

(3)东半球与北半球的情况基本相同,第二候东半球 500hPa 46 次集合预报的平均 ACC 要较同期控制预报的平均 ACC 高 0.036。

为考察集合平均预报图是否存在系统性的偏差,我们计算了这 46 次集合预报第一、二候的平均误差、均方根误差和正偏差(预报值大于实况值)的频率。由计算结果可见(图略):

(1)第一候和第二候的三种图形在分布上基本相同,反映出这 46 次集合平均预报不仅存在一定的系统性偏差,而且前后两个候的预报误差似有一定的稳定性。这启示我们,消除系统误差将会有助于提高集合预报的准确率。

(2)平均来说,中纬度从太平洋至美洲大陆到大西洋以正偏差为主,正偏差的概率达到 70%~90%,而欧亚大陆和鄂海附近地区为负偏差,负偏差的概率为 50%~60%。总的来看,正偏差的概率明显高于负偏差,即预报值常大于实际值,在洋面上尤为明显。

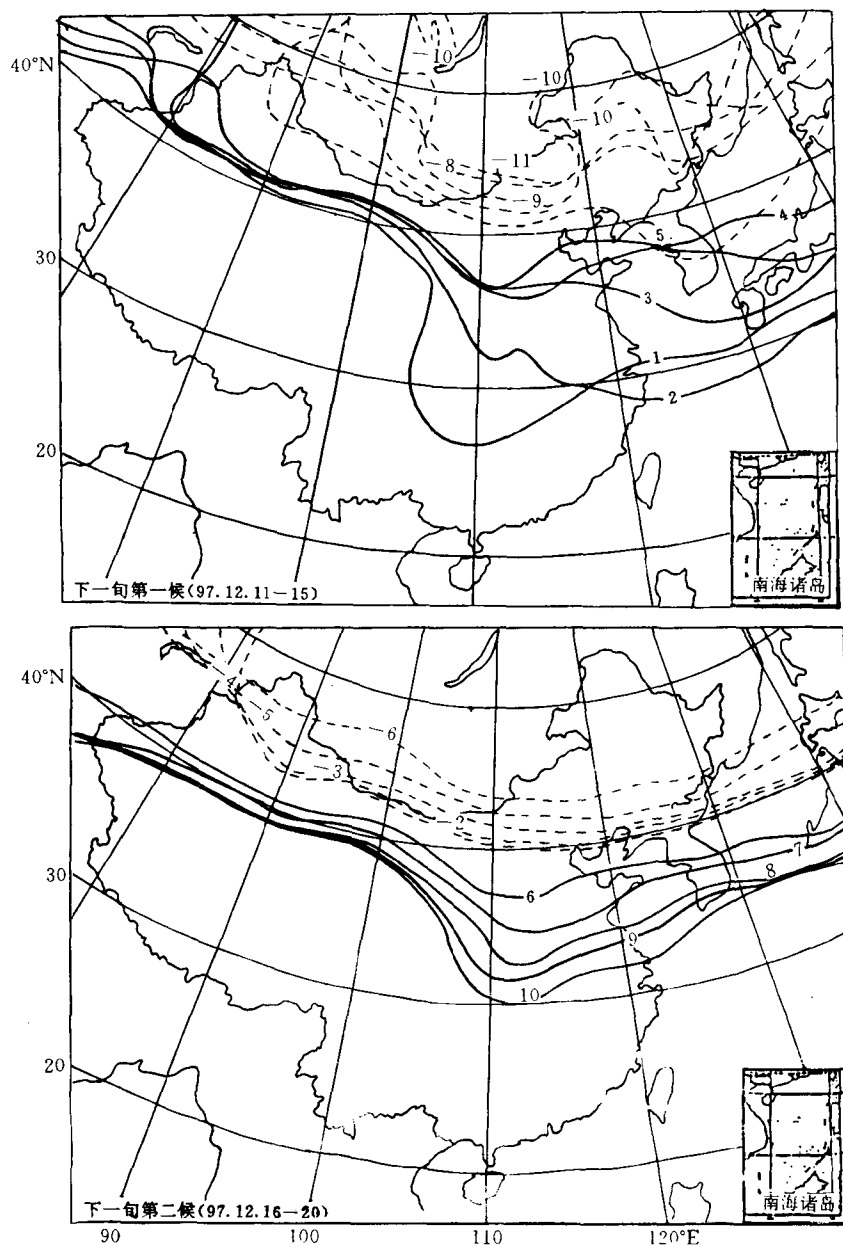


图1 逐日 850hPa 特征等温线(0°C和-12°C)集合
预报图(1997年12月9日)

此外,集合成员之间的离散度第二候明显大于第一候。不同成员之间槽脊位置和强度的差别造成集合平均图往往较实况平滑得多,第二候的预报图尤为明显。

5. 几点体会与看法

(1)从近一年的集合预报试验结果来看,对下一旬第一候来说集合预报没有多大意义,多数情况下其准确率远不如控制预报;而对第二候而言,虽然平均来说准确率不算高,但多数情况集合预报准确率要超过控制预报,对实际业务有一定参考价值。许多研究结果均表明,数值

预报的准确率随着预报时效的延长而降低,这种现象在一周以内表现得最明显,一周到两周则差别不大。因此,用 IAF 方法所作的集合预报,在第一候由于将较短预报时效的预报场与较长时效的预报场相平均,使得集合预报的平均 ACC 低于控制预报是很自然的。对于第二候而言,由于不同时效的准确率差别不大,集合平均可以在一定程度上平滑掉随机扰动,从而使得集合预报的 ACC 高于控制预报。

(2)集合预报场往往较实况平滑得多,这是由于集合成员之间明显差异所形成的。这就使得目前实际应用时集合预报往往只能作趋势性预报参考,集合成员之间的离散程度与预报可信度之间的关系还需要进一步分析、检验。怎样才能用好集合预报也还需要不断总结经验。

(3)无论是集合预报还是控制预报,其准确率对不同个例很不相同,集合预报与控制预报准确率之差也是如此,它们取决于何种因素尚不清楚。如能事前对其作出估计和判断,对于用好集合预报是很有好处的。

(4)目前集合产品的生成完全是借用 MIPS 和 MICAPS 的绘图功能,出图速度和能力均不能令人满意,有些种类的产品(如邮票图、概率烟羽图等)则无法制作。如不进行图形开发,要增加新品种的集合产品是很困难的。

(5)目前还只限于对形势预报场进行集合处理,如何由形势集合预报转化为要素(降水、温度)的集合预报仍是待解决的问题。

(6)用 IAF 方法进行集合,相当于要丧失一定的预报时效,因此并非集合预报的理想方法。在控制预报基础上叠加合理的小扰动,用并行计算机方法或在多个结点上同时运算以产生多个预报场,再采用适当方法进行集合,给出集合预报产品,看来可能是今后改进集合预报的方向。

参考文献

- [1] Tracton, M. S. and E. Kalnay, Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Weather and Forecasting*. 1993, 8(3): 378—398
- [2] Bromkovic, C. et al., Extended-range predictions with ECMWF models: Time-lagged ensemble forecasting. *Q. J. R. Meteor. Soc.* 1990, 116: 867—912
- [3] Leifh, C. E., Theoretical skill of Monte Carlo forecasts. *Mon. Wea. Rev.*, 1974, 102: 409—118
- [4] Hoffmann, R. N. and E. Kalnay, Lagged average forecasting, an alternative to Monte Carlo forecasting, *Tellus*, 1983, 35A: 100—118
- [5] Toth, E. and E. Kalnay, Ensemble forecasting at NMC: The generation of Perturbations. *Bull. Amer. Meteor.*, 1993, 74(12): 2317—2330
- [6] 李小泉, 美国国家气象中心中期预报时段内的集合预报. *气象科技* 1994, 2: 7~11
- [7] 张小礼, 集合预报简介, *气象科技* 1996, 2: 9~14
- [8] 李小泉, 刘金达, 张迎辉: 集合预报及其在预报中期天气预报中的应用. *气象*, 1997, 23(8): 3~9

不确定性理论——集对分析在 MOS 概率天气预报中的应用

王国强

(浙江省绍兴市气象台 312000)

摘 要

概率天气预报的关键问题之一是天气预报的不确定性,而集对分析(SPA)为研究天气预报的不确定性问题提供了新的思路。本文针对概率天气预报的特点,应用 SPA 中联系度的概念,对数值预报产品的预报因子设立中相关概率区间,进行同异反分析,建立了基于 SPA 的 MOS 概率天气预报模型。试验表明,联系度概念的引进增强了概率天气预报模型结构的合理性,提高了用数值预报产品制作概率天气预报的准确率。

关键词 概率天气预报 不确定性 集对分析 联系度

1. 引言

概率天气预报在国外已普遍投入业务应用,近年在国内的一些业务应用也收到良好的效果。这些预报大多采用数值预报产品的释用方法制作^[3]。开展概率天气预报的基本依据有两条:其一,天气预报本身存在着不确定性;其二,概率天气预报能进而为用户作出最佳气象-经济决策,以取得更好的经济效益。大气运动同时具有确定性和随机性的双重特性,是人们在天气预报的长期实践中得到的辨证认识,概率天气预报是这种认识深化的必然结果^[2]。显然,大气运动的随机性和天气预报的不确定性是研究概率天气预报的关键问题之一,但是我们对此还缺乏行之有效的分析处理方法。而集对分析(SPA)正是关于不确定性问题的数学理论和方法,它为研究天气预报的不确定性提供了新的思路和途径。

集对分析^[3]理论自 80 年代问世以来,已广泛地应用于现代科学管理、系统控制、决策分析、人工智能、工业设计、军事指挥和社会经济学等诸多领域。所谓“集对”是指具有一定联系的两个集合所组成的对子。集对分析认为,对这样两个集合的特性可以作同异反分析,并可定量地用联系度表达式来描述:

$$\mu(W) = S/N + (F/N)i + (P/N)j \quad (1)$$

式中 S/N 为在命题 W 下的同一度, F/N 为差异度, P/N 为对立度。(1)式也可记为

$$\mu(W) = a + bi + cj \quad (2)$$

其中 $a = S/N, b = F/N, c = P/N$ 。公式(1)和(2)表示在命题 W 下对某集对作分析,共得到 N 个特性,其中有 S 个特性为两个集合所共有,有 P 个特性为两个集合相对立,有 F 个特性既不对立又不同一。从公式的形式看,它与复数表达式十分相似,它也有四则运算等运算法则。式中的 i 和 j 既是差异度和对立度的标记,同时又可给予赋值,以便计算联系度的数值。计算时取 $j = 1$,而 i 则要根据命题 W 的专业特点而定,其变化区间为 $[-1, 1]$ 。关于集对分析对不确定

性问题的详述可参阅文献[4]。

本文试图在数值预报产品释用中,引入集对分析中的联系度概念,针对概率天气预报的特点进行同异反分析,从而建立新的概率天气预报模型。

2. 集对分析在概率天气预报中应用的思路

假设我们对天气事件 W 作预报,事件 W 的出现概率为 p ,预报因子为 x 。当自变量为一维时,自变量与因变量的相关概率曲线如图 1。在图 1a 中曲线 $\hat{P} = p(x)$ 近似于线性关系,其中 a 段为低相关概率区间,指示事件 W 不出现的主要特性; c 段为高相关概率区间,指示事件 W 出现的主权特性;而 b 段则介于二者之间,为中相关概率区间,对事件 W 出现或不出现表示“不确定”(以下对三个相关概率区间分别简称为低概率区、高概率区和中概率区)。在图 1b 中 $\hat{P} = p(x)$ 为一种单峰曲线。不管曲线 $\hat{P} = p(x)$ 的几何图形如何,一般总可找到低概率区、高概率区和中概率区。图 1b 的低概率有两段—— a_1 和 a_2 ,中概率区也有两段—— b_1 和 b_2 。

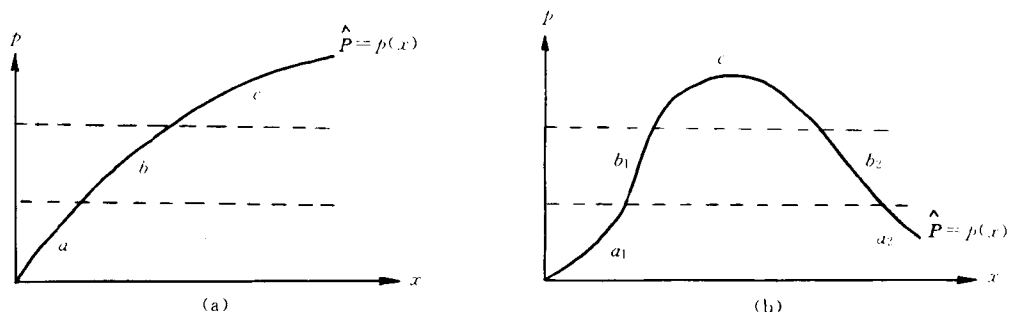


图 1 相关概率曲线的区间划分示意图

在某高概率区(或低概率区),如果共有 N 次个例,其中天气事件 W 出现次数为 S ,不出现次数为 P ,那么同一度 $a = S/N$,对立度 $c = P/N$,而差异度 $b = F/N = 0/N$ 。在某中概率区,既然自变量对事件 W 的出现或不出现表示不确定,因此不再按照在高概率区(或低概率区)的方式来表示联系度。如在中概率区共有 N 次个例,那么差异度 $b = N/N$,而同一度 $a = 0/N$ 。此时联系度表示了一种既不倾向于事件 W 出现,也不倾向于 W 不出现的“含糊”态度。这是十分符合预报实际情况的。在用数理统计方法处理这类问题时,往往不问自变量对于事件 W 出现或不出现的不确定状态,而勉强给出一个出现频率(即 SPA 中的同一度 a),这样往往可能导致预报结论的失败。这种情况在第 5 节将作一些具体解析。天气预报模型一般由多元预报因子构成,当某一因子位于中概率区,既然它处于不确定状态,那么与其勉强“表态”,还不如让其它因子去决定预报结论。这样在处理上就显得更合理些,其效果也可能更好些。

3. 基于 SPA 的概率天气预报模型的操作方法和试验结果

现以绍兴市气象台 12 月份 120 小时降水概率预报(POP)为例,来说明基于 SPA 的概率天气预报模型(以下简称 SPA 概率预报)的操作方法。样本容量为 93,预报因子为 5 元,因子均由数值预报产品的格点资料组合而成,并已作天气学方面的分析加工^[5]。预报因子与预报量的名称如下:

x_1 —— 96 小时地面气压预报场的锋区气压差;

x_2 —— 120 小时地面气压预报场的相关区平均气压;
 x_3 —— 72 小时 500hPa 高度预报场的锋区高度差;
 x_4 —— 72 小时 500hPa 高度预报场的相关区平均高度;
 x_5 —— 96 小时 500hPa 高度预报场的锋区高度差;
 P —— 96~120 小时时段的降水概率。

表 1 预报因子 x_1 的区间划分和联系度分析

名 称	区 间	同一度 $a = S/N$	差异度 $b = F/N$	对立度 $c = P/N$
高概率区	$(13, \infty)$	9/15	0/15	6/15
中概率区	$[-1, 13]$	0/24	24/24	0/24
低概率区	$(-\infty, -1)$	6/54	0/54	48/54

表 1 为预报因子 x_1 的区间划分和联系度分析的基本情况。以样本为依据,把 x_1 的定义域划分为三个区间——高概率区、中概率区和低概率区。高概率区为 $x_1 \in (13, \infty)$,在此区间中共有 15 次个例,其中降水个例 9 次,无降水个例 6 次,则同一度 $a = 9/15$,对立度 $c = 6/15$,差异度 $b = 0/15$ 。同样可得到低概率区为 $x_1 \in (-\infty, -1)$, $a = 6/54$, $c = 48/54$, $b = 0/54$ 。中概率区为 $x_1 \in [-1, 13]$,在此区间中 x_1 对降水事件的出现与不出现没有明确的态度,这是一个不确定区域,因此有 $a = 0/24$, $c = 0/24$,而 $b = 24/24$ 。

利用表 1 的分析结果,可以逐次统计关于 x_1 的 93 次个例的同一度、差异度和对立度,结果列于表 2。对其它 4 个因子也同样作 3 类区间的划分和联系度分析,同样计算各因子的 93 次个例的同一度、差异度和对立度(表略)。然后在此基础上,再逐次计算 93 次个例的 5 个因子的同一度、差异度和对立度的平均值。平均值数据见表 3。

表 2 预报因子 x_1 的各个例的联系度统计

序号	同一度 a	差异度 b	对立度 c
1	0.600	0.000	0.400
2	0.600	0.000	0.400
3	0.600	0.000	0.400
4	0.000	1.000	0.000
5	0.111	0.000	0.889
6	0.111	0.000	0.889
7	0.111	0.000	0.889
8	0.111	0.000	0.889
9	0.111	0.000	0.889
10	0.111	0.000	0.889
⋮	⋮	⋮	⋮

表 3 5 个因子平均联系度

序号	平均同一度 $\bar{a}$	平均差异度 $\bar{b}$	平均对立度 $\bar{c}$
1	0.407	0.400	0.193
2	0.407	0.400	0.193
3	0.336	0.400	0.264
4	0.024	0.800	0.176
5	0.062	0.400	0.538
6	0.139	0.400	0.461
7	0.062	0.400	0.538
8	0.178	0.200	0.622
9	0.039	0.600	0.361
10	0.072	0.200	0.728
⋮	⋮	⋮	⋮

在联系度表达式(1)中差异度 $b = F/N$,它表示在 N 个特性中有 F 个特性表现为既不同一又不对立,即有 F 个特性对天气事件 W 表示“含糊”态度。前文已分析,此时该因子还不如把预报结论的决定权让给其它因子更为合理些。因此令

$$i = \sum S / (\sum S + \sum P) + [\sum P / (\sum S + \sum P)]j \quad (3)$$

或

$$i = \sum a / (\sum a + \sum c) + [\sum c / (\sum a + \sum c)]j \quad (4)$$

式中 $\sum$ 表示对所有因子求和,在上式中求和与求平均等价。公式(3)和(4)的含义是当因子落入中概率区时,它的差异度的 F 个特性按一定的比分配给它的同一度和对立度,这个比就是其它几个因子的平均同一度和平均对立度之比。把(3)式和(4)式分别代入(1)式和(2)式可得

$$\mu = [S/N + (F/N) \sum S / (\sum S + \sum P)] + [P/N + (F/N) \sum P / (\sum S + \sum P)]j \quad (5)$$

或

$$\mu = [\alpha + b \sum \alpha / (\sum \alpha + \sum b)] + [c + b \sum c / (\sum \alpha + \sum c)]j \quad (6)$$

按 SPA 的规则,取 $j = -1$,故有

$$\mu = (S - P)/N + F/N(\sum S - \sum P)/(\sum S + \sum P) \quad (7)$$

或

$$\mu = (\alpha - c) + b(\sum \alpha - \sum c)/(\sum \alpha + \sum c) \quad (8)$$

(7)式和(8)式的基本意义一致,差别在于平均频率的定义不同。比如当频率 $p_1 = m_1/n_1$ 和 $p_2 = m_2/n_2$, 其平均频率

$$p = (1/2)[(m_1/n_1) + (m_2/n_2)]$$

或

$$p = (m_1 + m_2)/(n_1 + n_2)$$

后者考虑了样本大小所引起的影响,一般样本越大,用频率代替概率的稳定性越好。这里不再详细讨论。用(8)式可以分别计算 5 个因子 93 次个例的联系度 μ 值,计算结果见表 4。

表 4 各预报因子的 μ 值

序号	μ_1	μ_2	μ_3	μ_4	μ_5
1	0.200	0.158	0.357	0.714	0.357
2	0.200	0.158	0.357	0.714	0.357
3	0.200	0.158	0.119	0.119	0.000
4	-0.765	-0.765	-0.765	-0.765	-0.765
5	-0.778	-0.792	-0.833	-0.765	-0.792
6	-0.778	-0.537	-0.833	-0.537	0.000
7	-0.778	-0.792	-0.833	-0.765	-0.192
8	-0.778	0.158	-0.833	-0.765	-0.554
9	-0.778	-0.806	-0.833	-0.806	-0.806
10	-0.778	-0.820	-0.833	-0.765	-0.905
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

表 4 中 $x_1 \sim x_5$ 的 μ 值分别记为 $\mu_1 \sim \mu_5$,这是新的因子值。用 $\mu_1 \sim \mu_5$ 作因子,对降水概率 p 作 5 元回归计算,得到回归方程为

$$\begin{cases} \hat{p} = \alpha + \beta^T \mu \\ \alpha = 0.664 \\ \beta = (0.216, 0.076, 0.193, 0.313, -0.017)^T \\ \mu = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5)^T \end{cases} \quad (9)$$

其中 τ 为向量转置符号。回归方程的复相关系数为 0.680,剩余方差为 9.023。

利用方程(9)可计算 93 次个例的降水概率 $\hat{p}$,从此可归纳关于 120 小时降水的预报指标:

$$\hat{Y} = \begin{cases} 1(\text{降水出现}) & \text{当 } \hat{p} > 0.340 \\ 0(\text{降水不出现}) & \text{当 } \hat{p} \leq 0.340 \end{cases} \quad (10)$$

对样本拟合的成败频数的统计结果见表 5。降水预报的准确率与降水预报的技巧评分公式分别为：

$$TS = N_1 / (N_1 + N_2 + N_3) \tag{11}$$

$$SS = (TS - AC) / (1 - AC) \tag{12}$$

其中 N_1 为预报降水的正确次数, N_2 为空报次数, N_3 为漏报次数, AC 为降水气候概率。应用表 5 的统计数据可计算得到 $TS = 0.600, SS = 0.447$ 。

4. SPA 概率预报的效果

为了鉴定 SPA 概率预报的效果,现在用概率回归法再制作降水概率预报,所用因子不变,以供比较。

(1)在 SPA 概率预报中把因子的定义域划分成 3 类区间,现在取消不确定的区间——中概率区,而划分成两类区间——高概率区和中概率区。并且分析各区间的联系度,分别读出各因子各个例的同一度、差异度和对立度,再用公式(8)计算各因子各个例的 μ 值,最后进行多元回归计算。整个处理过程除了划分区间外,与 SPA 概率预报完全一致。得到概率回归预报方程：

$$\begin{cases} \hat{p} = a + \beta^T \mu \\ \alpha = 0.820 \\ \beta = (0.243, 0.270, 0.236, 0.302, 0.058)^T \\ \mu = (\mu_1, \mu_2, \mu_3, \mu_4, \mu_5,) \end{cases} \tag{13}$$

回归方程的复相关系数为 0.593,剩余平方和为 10.896,同样可得到关于降水的预报指标

$$\hat{Y} = \begin{cases} 1(\text{降水出现}) & \text{当 } \hat{p} > 0.250 \\ 0(\text{降水不出现}) & \text{当 } \hat{p} \leq 0.250 \end{cases} \tag{14}$$

(14)式对样本拟合的成败频数的统计结果见表 6。应用(11)和(12)式可计算得到概率回归预报的准确率和技巧评分分别为 $TS = 0.450, SS = 0.279$ 。

表 5 SPA 概率预报的质量统计列联表(单位:次)

预 报 实 况	降水	无降水	合计
降 水	18	4	22
无降水	8	63	71
合 计	26	67	93

表 6 概率回归预报的质量评定列联表(单位:次)

预 报 实 况	降水	无降水	合计
降 水	18	4	22
无降水	18	53	71
合 计	36	57	93

(2)表 7 例出了 SPA 概率预报和概率回归预报两种预报模型的质量情况。虽然两种模型所用的因子相同,但结果却不同,4 种质量指标均以 SPA 概率预报为优。其中前者的剩余平方和比后者减少 17.2%,前者的准确率比后者提高 33.3%。可见集对分析理论的应用提高了概率天气预报的质量。

表 7 两种预报模型的质量比较

	复相关系数	剩余平方和	准确率	技巧评分
SPA 概率预报	0.680	9.023	0.600	0.477
概率回归预报	0.593	10.896	0.450	0.279