

文章编号:1000-694X(2006)02-0268-05

不确定性理论集对分析在沙尘暴预报中的应用研究

王繁强¹, 郭大梅²

(1. 陕西省气象局 气象科学研究所, 陕西 西安 710015; 2. 陕西省气象台, 陕西 西安 710015)

摘要: 应用集对分析(SPA)这一关于不确定性问题的数学理论和方法, 基于概率统计分析, 从 SPA 的同、异、反分析出发, 用联系度公式导出解决预报中合理地使用预报因子的方法, 实现预报模型因子结构的动态优化, 增强模型预报机制的合理性, 达到提高模型预报能力的目的。在以往对沙尘暴研究成果的基础上, 以强风、热力和沙源三大影响因子为着眼点, 结合 2001—2003 年的沙尘暴天气个例, 对沙尘暴天气进行了分类, 分别选取预报因子, 建立基于 SPA 的沙尘暴预报模型。根据数值预报产品输出结果, 于 2004 年春季进行了短期(24 h)预报试用, 结果表明, 这一方法具有较好的效果。

关键词: 集对分析; 不确定性; 联系度; 沙尘暴预报

中图分类号: P457.5

文献标识码: A

沙尘暴是一种重要的环境问题, 是沙漠化的主要过程之一, 也是在沙漠及其边缘等特殊下垫面条件下产生的一种灾害性天气。中国的沙尘暴主要发生在西北干旱、半干旱地区, 已严重干扰和影响了北方地区人们的正常活动, 对社会经济和人们的生活环境均造成了一定程度的危害, 引起了政府和科学工作者们的高度重视, 有力地推动了沙尘暴研究的广泛开展, 并取得了可喜的研究成果^[1~8]。

为了做好对沙尘暴天气的预报服务, 各地气象部门在对沙尘暴天气分析研究的基础上, 依据天气学、统计学方法建立了多元线性和非线性预报模型, 这种预报模型的优点是实用、易于操作; 缺点是在预报模型中因子的重要性是固定不变的, 是由历史样本过程决定的, 是一个统计意义上的平均概念。事实上, 对于样本中的某一个例或者应用中的某一次预报过程则并非完全如此, 有时会因因子集合中的一个或某几个“最佳”因子的表现不佳而导致该次预报的失败, 多元线性模型和非线性模型难以对这种现象进行合理的判断和动态处理。而集对分析^[9,10](Set Pair Analysis, 简记为 SPA)这一关于不确定性问题的数学理论和方法, 为研究天气预报的不确定性提供了新的思路和途径。因此, 近几年来, 在天气预报中得到了较为广泛的应用^[11,12], 笔者在应用研究的基础上, 把传统的统计、天气学方法同不确定性理论集对分析技术有机地结合起来, 探索建立一种新的沙尘暴预报方法。

1 集对分析简述

集对分析理论自 20 世纪 80 年代问世以来, 已广泛地应用于现代科学管理、人工智能、信息工程和气象等^[11~13]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:

$$u(W) = s/N + (f/N)I + (P/N)J \quad (1)$$

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。

式(1)也可记为:

$$\mu(W) = a + bI + cJ \quad (2)$$

式中: $a = s/N$, $b = f/N$, $c = P/N$ 。

公式(1)或(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。

联系度 $\mu(W)$ 的公式(1)或(2)从同、异、反 3 个方面比较两个集合的特性, 以描述事物的确定性和不确定性以及它们之间的转换关系。从表达式的形式看, 它与复数的表达式相似, 也有四则运算等运算法则。式中 I 和 J 既是差异度和对立度的标记, 同时又可赋值, 以对联系度进行计算。计算时取 $J = -1$, 而 I 的取值则根据命题的专业特点而定, 其变化区间为 $[-1, 1]$ 。

2 适用于天气预报问题的联系度表达式

2.1 基于预报问题的联系度表达式

在预报时希望那些重要性已下降的因子在这次预报中不要干扰模型做出正确预报结论,而让那些预报能力较强的因子发挥主要作用,也就是说当因子预报能力较强时要发挥它的作用,当因子预报能力较弱时则要抑制它的作用。概括为这样一个问题:假设某一气象因子集合由 M 个因子组成,在某次预报中有 P 个因子表现不佳,那么能否设法将这 P 个因子在这次预报中的作用降低,而使其余($m - P$)个表现优良的因子起主要作用呢?在下次预报中如果有另外 q 个因子表现不佳,同样能否将 q 个因子的作用降低,而由其余($m - q$)个表现优良的因子发挥主要作用呢?可见这一思路就是要求合理地使用预报因子,实现预报模型中因子结构的动态变化,以提高预报模型的整体预报能力。多元线性模型和非线性模型通常难以对元素在各次预测中重要性程度的变化进行合理判断和动态处理。本节给出从 SPA 的同、异、反分析出发,结合概率统计分析,用联系度公式(1)或(2)导出解决这一问题的方法。

假设对天气事件沙尘暴 W 作预报,沙尘暴 W 的出现概率为 P ,预报因子为 x ,当自变量为一维时,分析自变量与因变量的相关概率 $P = f(x)$,可知,当 $P = 50\%$ 时,则表示用 x 预测 y 的效果最差,自变量 x 关于因变量 y 的不确定性最大, P 越大或越小,表示用自变量 x 预测 y 出现或不出现的效果越好,自变量 x 关于因变量 y 的不确定性越小,也就是自变量 x 的预报能力越强。不管曲线 $P = f(x)$ 的几何图形如何,一般总可找到某一区间上的 P_{50} 、 P_{\max} 和 P_{\min} 以及对应的 x_{50} 、 x_{\max} 和 x_{\min} 。把第 i 个 X_i 与 x_{50} 、 x_{\max} 和 x_{\min} 进行对比分析可知, X_i 越接近 x_{\max} 和 x_{\min} 时,表示此时因子的预报能力越强,它对于预报量的不确定性较小,此时它对于预报量的确定性占主导地位,对预报量的不确定性占从属地位。 X_i 越接近 x_{50} 时,表示此时因子的预报能力越弱,自变量 X_i 对因变量 Y_i 的预测“态度”处于含糊状态,它对于预报量的不确定性较大,此时它对于预报量的不确定性占主导地位,对预报量的确定性占从属地位,显然, $X_i = x_{50}$ 时,它对于预报量的不确定性最大。从 SPA 的同、异、反分析出发,用联系度公式(1)或(2)导出解决这一问题的方法因此有:

$$\begin{cases} a_i = \frac{s}{N} = \frac{x_i - x_{50}}{x_{\max} - x_{50}} \\ x_i > x_{50} \text{ 时: } c_i = \frac{P}{N} = 0, u_i = a_i \\ b_i = \frac{f}{N} = 0 \end{cases} \quad (3)$$

$$\begin{cases} a_i = \frac{s}{N} = 0 \\ x_i < x_{50} \text{ 时: } c_i = \frac{P}{N} = \frac{x_i - x_{50}}{x_{50} - x_{\min}}, u_i = c_i J \\ b_i = \frac{f}{N} = 0 \end{cases} \quad (4)$$

$$\begin{cases} a_i = \frac{s}{N} = \frac{0}{x_{50}} = 0 \\ x_i = x_{50} \text{ 时: } c_i = \frac{P}{N} = \frac{0}{x_{50}} = 0, u_i = b_i I \\ b_i = \frac{f}{N} = \frac{x_{50}}{x_{50}} = 1 \end{cases} \quad (5)$$

显然,若取 $I = 0$,联系度 u 的值就表示了因子的预报能力。

式(2)中,同一度 a 和对立度 c 是对系统确定性部分的量度,差异度 b 是对系统不确定性的量度。

2.2 基于联系度表达式的预报模型

对于有 M 个因子的因子集合来说,对应有 M 个联系度组成的集合,定义: $u = \sum_{i=1}^M u_i$,为总体联系度。

如用 $u = \sum_{i=1}^M u_i$ 做数学模型(预报模型) F ,对命题 W 进行预测,显然,在每次预报中表现优良的因子将起主要作用。即在每次预报中,运用公式(3)、(4)、(5)通过 SPA 的同、异、反分析对各预报因子的预报能力进行分析和判断,对一些预报能力不强,甚至可能干扰预报模型做出正确预报结论的因子,进行抑制、甚至消除其影响,而让其他预报能力较强的因子来决定预报结论。也就是说,随着环境的变化,会有不同因子的作用被抑制或取消,从而实现预报模型的因子结构动态优化,增强模型预报机制的合理性,达到提高模型预报能力的目的。

对命题 W 进行有无的预测,按判别分析^[14]方法,通过二级判别分析确定预报临界值 u_c ,可对命题 W 进行有无的预测。通过多级判别分析可对命题 W 作分级预报。

3 主要环流形势

张高英等^[5]通过对 2000—2002 年春季(3~5 月)中国北方 12 次强沙尘暴的天气气候特征分析,

认为:蒙古气旋是造成我国北方强沙尘暴天气的主要影响系统,并将其归纳为西部型、中部型、东部型3种类型,气旋冷锋后的大风是强沙尘暴天气发生的主要动力因子。^[7]王式功等通过对我国西北地区“94.4”沙尘暴成因分析,认为,此次沙尘暴不是产生在冷锋过境后的西北大风区,而是产生于偏东大风区;低层中尺度辐合切变线是此次强沙尘暴的直接影响系统。

文献[1,3]对沙尘暴天气的研究分析也给出了类似的结论,在以上研究的基础上,通过对2001—2003年春季我国西北地区沙尘暴天气过程的分析,我们认为,可将我国西北地区大范围沙尘暴天气简单的分为冷锋后大风区沙尘暴和冷锋前暖区沙尘暴两类。冷锋后大风区沙尘暴的特点是:沙尘持续时间相对较短,强度强,一般是随风而至,随风而去。过程前一天500 hPa高空形势,关键区 $40^{\circ}\sim65^{\circ}\text{N}$ 、 $65^{\circ}\sim115^{\circ}\text{E}$ 内有低压槽。高空槽前24 h负变高 ΔH_{24} 大于40 gpm,槽后正变高 ΔH_{24} 超过70 gpm。低槽冷中心小于 -32°C ,最大风速大于 $16 \text{ m}\cdot\text{s}^{-1}$ 。冷锋前暖区沙尘暴的特点是:沙尘持续时间相对较长,强度相对较弱,以浮尘天气为主,一般是随东风而至,随西风而去。过程前一天500 hPa高空形势,关键区 $35^{\circ}\sim65^{\circ}\text{N}$ 、 $40^{\circ}\sim80^{\circ}\text{E}$ 内有低压槽,槽后冷空气迅速南下,使低槽不断加深,但东移缓慢。高空槽前有较强的正涡度平流和暖气流向北输送,使低层700 hPa上有低压迅速发展,导致沙源区形成偏东大风。这种天气主要出现在河西走廊,沙尘暴天气出现前期,青藏高原东部低空常出现大幅度增温,促使高原东北部低涡(如柴达木低涡)发展,若蒙古高原东南部同时也有高压发展,则可导致高原北部河西走廊地区低空东风急流的形成。

4 沙尘暴预报因子

研究表明^[1,6],沙尘暴的产生主要是受三大因子支配:强风因子、热力因子和沙源因子。因此,在对沙尘暴进行预报时,也必须综合考虑这三个因子及其相互作用问题。一般来说,沙源区前一年及冬季降水偏少,春季回暖快,且气温偏高,地表松散、干燥,抗风蚀能力差,有利于强沙尘暴的形成^[5]。强风和热力不稳定因子是影响沙尘暴的两个重要因素,与冷空气的活动有着密切的关系^[4]。

产生于冷锋过境后偏西大风区的沙尘暴,大风是沙尘暴天气发生的主要动力因子,强风因子起决定性作用。产生于冷锋前偏东大风区的沙尘暴,热力因子起决定性作用,低空常出现大幅度增温,促使

低涡、切变发展。低层低涡、辐合切变是强沙尘暴的直接影响系统。但对流层低层强烈的对流不稳定是两类沙尘暴天气“起沙成暴”的共同点,对沙尘暴的预报非常重要。因此,依据上述分析,通过对2001—2003年春季我国西北地区沙尘暴天气个例与对应的当日20时(北京时,下同)micaps系统分析处理的实况格点资料的相关统计分析,针对两类沙尘暴天气分别选取20个相关较好预报因子,如:锋前暖区沙尘暴预报因子,700 hPa 24 h 变温 ΔT_{24} 、地面24 h 变压 ΔP_{24} 、地面24 h 变温 ΔT_{24} 、850~700 hPa 的温度差 ΔT_{87} 等;锋后大风区沙尘暴预报因子,700 hPa 24 h u 分量风速差 ΔU_{24} 、850 hPa 24 h u 分量风速差 ΔU_{24} 、500 hPa 24 h 变高 ΔH_{24} 、地面24 h 变压 ΔP_{24} 等。

5 基于SPA的预报模型的应用分析

预报采用PP法,以当日12时为起报点预报未来24 h西北地区出现沙尘暴的区域,预报因子从国家气象台下发的T213数值模式输出的24 h预报资料中获取。

5.1 x_{50} 、 x_{\max} 和 x_{\min} 及临界值 u_c 的确定

利用2001—2003年春季3~5月逐日20时实况分析资料,首先,分析计算两类沙尘暴天气下各预报因子的 x_{50} 、 x_{\max} 和 x_{\min} 值,具体计算可依据沙尘暴个例求得。如关于500 hPa ΔH_{24} 变压因子的 x_{50} 、 x_{\max} 和 x_{\min} 值的计算,分析历史资料共有锋后大风区沙尘暴天气23次,其中,1次 $\Delta H_{24}=20 \text{ gpm}$,12次 $\Delta H_{24}\leqslant46 \text{ gpm}$,23次 $\Delta H_{24}\leqslant75 \text{ gpm}$ 。因此,沙尘暴天气在 $\Delta H_{24}=20 \text{ gpm}$ 条件下出现的概率 $P(\Delta H_{24}=20 \text{ gpm})=\frac{1}{23}=4\%$;沙尘暴天气在 $\Delta H_{24}=46 \text{ gpm}$ 条件下出现的概率 $P(\Delta H_{24}=46 \text{ gpm})=\frac{12}{23}=52\%$;沙尘暴天气在 $\Delta H_{24}=75 \text{ gpm}$ 条件下出现的概率 $P(\Delta H_{24}=75 \text{ gpm})=\frac{23}{23}=100\%$,得:预报因子

ΔH_{24} 的 $x_{50}=46 \text{ gpm}$ 、 $x_{\max}=75 \text{ gpm}$ 、 $x_{\min}=20 \text{ gpm}$ 。其次,按二级判别分析,依据历史资料分别确定对两类沙尘暴天气进行预报的临界值 u_c 。锋前暖区沙尘暴预报临界值 $u_{c1}=7\sqrt{15}$,对样本进行回代,准确率=72%,空报率=17%,漏报率=11%;锋后大风区沙尘暴预报临界值 $u_{c2}=6.83$,回代,准确率=79%,空报率=13%,漏报率=8%。

5.2 沙尘暴预报

预报时,首先,根据环流形势分析判断可能发生

沙尘暴天气的类型,从数值预报产品资料获取预报区内各格点(站点)的预报因子。计算其联系度 u_i ,对各预报因子 x_i 的预报能力进行分析和判断,取前 10 个具有较好判断能力的因子的联系度,计算总体联系度 u 。确定 $u > u_c$ 的区域即为沙尘暴预报区

域。预报流程如图 1 所示。

图 2 中 a、b 分别为 2004 年 3 月 8 日、4 月 29 日沙尘暴天气的预报,图 3 中 a、b 分别为 2004 年 3 月 8 日、4 月 29 日沙尘暴天气的卫星监测结果。可以看出,沙尘暴的预报区域与实况基本一致,效果是满意的。

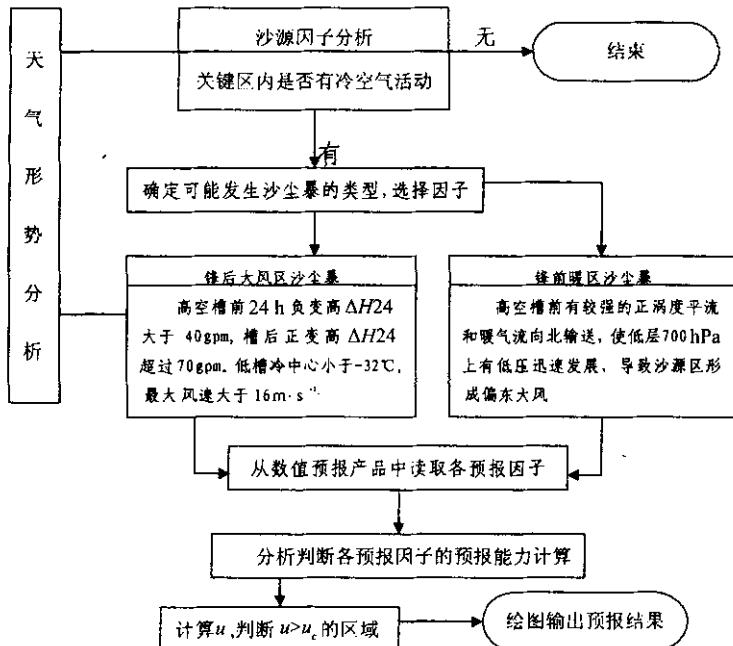


图 1 沙尘暴预报流程

Fig. 1 The flow chart of sandstorm forecasting

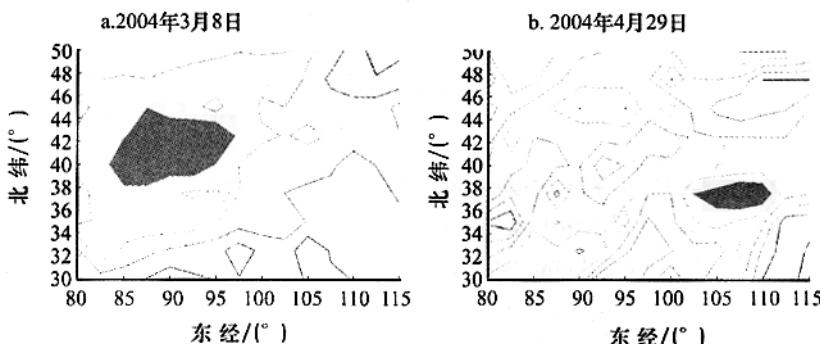
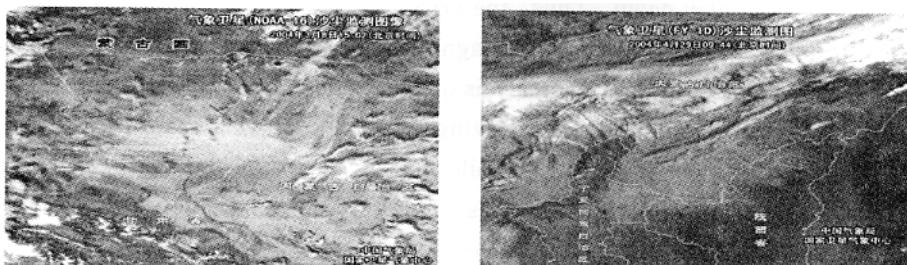


图 2 沙尘暴预报

Fig. 2 Sand storm forecasting map



a.2004年3月8日

b.2004年4月29日

图 3 卫星监测图像

Fig. 3 Satellite monitoring figures for sandstorm

6 小结

应用集对分析原理, 基于概率统计分析, 从 SPA 的同、异、反分析出发, 用联系度公式导出解决预报中合理地使用预报因子的一种新的预报方法。预报时对所建立的预报模型中的预报因子进行 SPA 分析, 即在因子参与作出预报结论之前先对各因子的预报能力做同、异、反分析, 根据判断结果, 对一些预报能力不强, 甚至可能干扰预报模型作出正确预报结论的因子, 进行抑制、甚至消除其影响, 让表现优良的因子在预报中发挥主要作用, 因而实现预报模型因子结构的动态优化, 增强了模型预报机制的合理性。

通过对造成西北地区沙尘暴天气的分析, 从强风、热力和沙源三大影响因子着眼, 分别分析总结了影响两类沙尘暴天气的“最佳”预报因子, 建立了基于 SPA 的预报模型。试报结果看, 具有较好的效果。

参考文献(References):

- [1] 钱正安, 宋敏红, 李万元. 近 50 a 来中国北方沙尘暴的分布及变化趋势分析[J]. 中国沙漠, 2002, 22(2): 107—111.
- [2] 王式功, 董光荣, 陈惠忠, 等. 沙尘暴研究的进展[J]. 中国沙漠, 2000, 20(4): 349—356.

- [3] 徐启运, 胡敬松. 我国西北地区沙尘暴天气时空分布特征分析[A]. 中国沙尘暴研究[C]. 北京: 气象出版社, 1997. 1—10.
- [4] 丁瑞强, 王式功, 尚可政. 近 45 a 我国沙尘暴和扬沙天气变化趋势和突变分析[J]. 中国沙漠, 2003, 23(3): 306—310.
- [5] 张高英, 赵思雄, 孙建华. 近年来强沙尘暴天气气候特征的分析研究[J]. 气候与环境研究, 2004, 9(1): 102—115.
- [6] 杨晓玲, 王式功, 李岩瑛. 河西走廊东部近 50 a 沙尘暴成因、危害及防御对策[J]. 中国沙漠, 2002, 22(3): 284—287.
- [7] 王式功, 杨德保, 周玉素, 等. 我国西北地区“94.4”沙尘暴成因探讨[J]. 中国沙漠, 2004, 15(2): 332—338.
- [8] 赵翠光, 刘还珠. 我国北方沙尘暴发生的环流形势分析[J]. 应用气象学报, 2004, 15(2): 245—250.
- [9] 赵克勤. 集对分析——一种新的不确定性理论方法与应用[J]. 系统工程, 1996, 14(1): 18—23.
- [10] 赵克勤. 集对分析对不确定性的描述和处理[J]. 信息和工程, 1995, 24(3): 165—168.
- [11] 王繁强. 集对分析在暴雨预报集成配套方法中的应用[A]. 暴雨落区预报使用方法[C]. 北京: 气象出版社, 2000. 44—47.
- [12] 薛根元, 王国强. 不确定性理论集对分析在预报模型建立中的应用研究[J]. 气象学报, 2003, (5): 592—599.
- [13] 李凡修, 陈武. 海水水质富营养化评价的集对分析方法[J]. 海洋环境科学, 2003, (2): 72—74.
- [14] 黄嘉佑. 气象统计分析与预报方法[M]. 北京: 气象出版社, 2003. 98—108.

Application of Uncertainty-Set Pair Analysis for Sandstorm Forecast in Northwest of China

WANG Fan-qiang¹, GUO Da-mei²

(1. Shaanxi Meteorological Institute, Xi'an 710015, China; 2. Shaanxi Meteorological Forecasting Center, Xi'an 710015, China)

Abstract: The paper, on the basis of formation and dynamics diagnosis, has developed a synoptic concept model about sandstorms' happening. According to this model, combined with the theory of the uncertainty-set pair analysis, the paper has designed the sandstorm short-range forecasting method.

The forecast factors in the synoptic concept model have been selected strictly by means of the characteristics of forecasted objects, the physical significance of the forecasting factors, the experiences in forecasting sandstorm weather, and some technical ways. Generally speaking, all the selected forecast factors have better forecast abilities, but the better abilities of these factors aren't always unchanged. Sometimes, the better ability of one factor may play down. Thus, the error is generated in forecast results. Set Pair Analysis (SPA) is a systemic theory and method used in diagnosing non-authenticity. Using the theory and method, this paper makes a judgment of status and analysis of same-difference-reverse about the factors that will be used to forecast the sandstorm weather. In analyzing, the method weakens effect of those factors with badness abilities in the sandstorm forecast model, while those with better abilities in the model will contribute greatly to forecast it. As a result, optimizing the structure of factors can be realized in the forecast model. The rationality of forecast mechanism in the model can be strengthened.

The applications for 24 h sandstorm forecasting in spring of 2004 interpreting and using from the numerical forecast products indicate that the method has better effect.

Key words: 集对分析 (SPA); 不确定性; 关系; 沙尘暴预报