

钟元 著

天气与短期气候 的多元判据 综合相似预测方法

气象出版社

天气与短期气候 的多元预测 综合相似预测方法

天气与短期气候的多元判据 综合相似预测方法

钟元 著

内容简介

本书阐述了在电子计算机技术与数值天气预报技术日益发展条件下天气与短期气候的相似预测技术的新发展-多元判据综合相似预测方法的原理及其在天气和短期气候预报中的应用。

全书共分八章;第1章概述了多元判据综合相似预测方法的基本原理及其典型的预测模式;第2章为多元判据综合相似预测方法在短期天气预测中的应用;第3章为多元判据综合相似预测方法在大-暴雨天气预测中的应用;第4章为多元判据综合相似预测方法在中期天气过程预测中的应用;第5章为多元判据综合相似预测方法在热带气旋移动路径预测中的应用;第6章为多元判据综合相似预测方法在热带气旋风雨影响预测中的应用;第7章为多元判据综合相似预测方法在短期气候预测中的应用;第8章为多元判据综合相似预测方法所涉及若干问题的讨论。

本书可供气象台、站从事预报工作的气象工作者,有关科研人员和高等院校的师生参考。

图书在版编目(CIP)数据

天气与短期气候的多元判据综合相似预测方法/钟元著.

-北京:气象出版社,2003.11

ISBN 7-5029-3677-7

I. 天... II. 钟... III. ①天气预报-方法②短期

天气预报-方法 IV. P45

中国版本图书馆 CIP 数据核字(2003)第 101631 号

天气与短期气候的多元判据综合相似预测方法

钟元 著

责任编辑:俞卫平 袁信轩 终审:纪乃晋

封面设计:王伟 责任技编:吴庭芳 责任校对:杨波平

气象出版社 出版

(北京市海淀区中关村南大街 46 号 邮编:100081)

网址:<http://www.cmp.cma.gov.cn> E-mail:qxcsbs@263.net

北京市兴怀印刷厂印刷

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:17.625 字数:450 千字

2003 年 11 月第一版 2003 年 11 月第一次印刷

ISBN 7-5029-3677-7/P·1300

定价:45.00 元

前言

相似预测方法是一种历史悠久的天气预报方法。它的优点是思路明确,直观性强,可进行多种类型的天气与短期气候过程或要素的定量预报。但是,由于受到时代和资料的局限,以往的相似预测方法只能进行预测时刻之前的简易相似。而两个天气过程前期的相似并不能保证它们的后续发展也一定相似,若干有限的简易判据,不能定量地判定变化极其复杂的两个天气过程的相似程度。现在,相似预测方法能够重新得到发展是得益于大气资料的日益丰富和数值天气预报的进展。近年来,由于大气探测手段的长足进展,我们已拥有更加丰富的大气多层次多方位的观测资料。同时,电子计算机技术的巨大进步和数值天气预报的飞跃发展,大尺度天气形势数值预报的预报时效已从短期向中期拓展,而预报时效为几个月以上的动力气候预报也取得一定的进展。这些进步赋予了相似预测方法新的生命力:一方面我们现在可以进行预测时刻之后即天气过程未来发展的相似;另一方面我们能应用多元判据定量地评估两个天气过程的相似程度。这不啻相似预测方法进行一次变革,当然,随之而来的是相似预测方法的原理、内容和形式与以往相比都发生了很大的变化。作者在多年从事天气与短期气候及其预测的研究中,已将相似预测方法逐步推广应用到短期天气、中期天气(包括逐日和旬天气)、短期气候和灾害性天气(如大-暴雨,热带气旋的移动路径和风雨影响)等方面的预测^[1~7]。现以在这些工作中所发表的研究论文为素材,加以整理、充实和系统化,写成本书。希望一方面能为各级气象台站的预报人员进行天气与短期气候预测时提供一种参考,另一方面由于作者水平有限,多元判据综合相似预测方法尚有不足和纰缪之处,希望能得到同行们的批评指正。

本书一共有八章。第一章介绍多元判据综合相似预测方法的基本原理,包括相似区域与预测对象的关系及选取原则,向前与向后拓展相似时段的方法及其作用,动态相似的含义及其实施细则,相似指数的定义及其在相似预测中的作用,多元判据的构造原则及方法,综合评估的作用及方法,以及多元判据综合相似预测方法的全程操作规程。第二章至第八章为具体介绍多元判据综合相似预测方法在各种天气与短期气候预测中的应用。其中大尺度天气形势与背景的相似是各种天气与短期气候预测中共有的部分,也是最主要的相似判据。各章都选用了高、中和低的代表层次和高度(气压)、温度、湿度及风场等基本要素场,在不同的时间-空间范围构造相似判据。除了天气形势与背景的相似外,各章还针对不同的预测对象,分别构造特征判据。第二章介绍多元判据综合相似预测方法在24~48小时短期天气预测中的应用,在构造多元判据时除了大尺度天气形势判据外,还介绍了气象要素场的车贝雪夫正交多项式展开典型场在相似预测中的应用。第三章介绍大-暴雨天气的多元判据综合相似预测方法,除了大尺度天气形势相似判据外,着重介绍反映产生大-暴雨的中小尺度天气系统和局地影响的局地环境条件判据的构造。第四章为中期天气过程的多元判据综合相似预测方法,着重介绍3~10天中期天气过程的时间-空间滤波,超长波与长波天气形势背景的相似和气象要素场的纬圈谱展开波谱参数在相似预测中的应用。第五章为热带气旋移动路径的多元判据综合相似预测方法,除了大尺度天气形势相似判据外,着重介绍与热带气旋移动路径密切相关的中层引导气流及平均气候转向点判据的构造方法。第六章为热带气旋风雨影响的多元判据综合相似预测方法,除了大尺度天气形势相似判据外,着重介绍了与热带气旋风雨影响密切相关的大风权重与强降水

权重的构造方法。第七章为短期气候的多元判据综合相似预测方法,除了前期超长波天气形势背景相似判据外,着重介绍了海-气耦合的前期海水表面温度场、厄尔尼诺关键区海温和南方涛动指数,太阳黑子相对数及环流特征量,气象要素场的自然正交展开函数和球谐展开参数等相似判据的构造和应用。第八章为天气与短期气候的多元判据综合相似预测方法若干问题的探讨,主要总结了相似预测方法所涉及到的各方面的问题和解决的方法,以及尚存的不足与问题。

在本书的编著过程中,石晓伦、祝纪平和翟春华参加了资料统计工作;李鸣一、陈建敏和黄卫群参加了绘图工作;刘青、赵江山和蔡露平参加了文字输入工作。对以上参加者的热情帮助和精心工作,作者一并表示衷心的感谢。

由于作者水平有限,书中的错误和不足之处,请诸位读者不吝指教,在此深表谢意。

作者

2003年7月1日

目 录

前言

绪论	(1)
第一章 多元判据综合相似预测方法的基本原理	(3)
1.1 多元判据综合相似预测方法的基本原理	(3)
1.2 相似判据	(4)
1.3 相似指数	(8)
1.4 相似时域	(11)
1.5 预测模式	(12)
第二章 短期天气的多元判据综合相似预测模式	(14)
2.1 短期天气相似预测的样本和相似时域	(14)
2.2 短期天气相似预测的相似判据	(15)
2.3 短期天气的多元判据综合相似预测方案	(42)
2.4 短期天气的相似预测模式的预测实例	(44)
第三章 大-暴雨天气的多元判据综合相似预测模式	(46)
3.1 大-暴雨天气相似预测的原理、样本和相似时域	(46)
3.2 大-暴雨天气相似预测的相似判据	(47)
3.3 大-暴雨天气的多元判据综合相似预测方案	(76)
3.4 大-暴雨天气多判据综合相似预测模式的预测实例	(77)
第四章 中期天气的多元判据综合相似预测模式	(79)
4.1 中期天气相似预测的样本和相似时域	(79)
4.2 中期天气相似预测的相似判据	(83)
4.3 中期天气的多元判据综合相似预测方案	(126)
4.4 中期天气相似预测模式的预测实例	(128)
第五章 热带气旋路径的多元判据综合相似预测模式	(131)
5.1 热带气旋路径相似预测的原理、样本和相似时域	(131)
5.2 热带气旋路径相似预测的相似判据	(133)
5.3 热带气旋路径的多元判据综合相似预测方案	(168)
5.4 热带气旋路径相似预测模式的预测实例	(170)
第六章 热带气旋风雨影响的多元判据综合相似预测模式	(172)
6.1 热带气旋风雨影响相似预测的原理、样本和相似时域	(172)
6.2 热带气旋风雨影响相似预测的相似判据	(174)
6.3 热带气旋风雨影响预测的局地风雨权重估计	(206)
6.4 热带气旋风雨影响的多元判据综合相似预测方案	(209)
6.5 热带气旋风雨影响相似预测模式的预测实例	(210)
第七章 短期气候的多元判据综合相似预测模式	(212)
7.1 短期气候相似预测的相似区域、时域和样本	(212)
7.2 短期气候相似预测的相似判据	(214)
7.3 短期气候的多元判据综合相似预测方案	(266)
7.4 短期气候相似预测的预测实例	(267)
第八章 多元判据综合相似预测方法的若干问题	(273)
参考文献	(276)

绪 论

相似预测方法是一种历史悠久的天气预报方法。它的思路是在过去的天气纪录(历史样本)中找出尽可能与现存条件(预报样本)相近的天气过程(相似样本)。一旦这样的相似样本被找到了,今后的天气发展过程就应该与相似样本的天气发展过程相类似。当利用历史相似样本外推时,是按照大气本身规律进行运作的,它包含了大气运动所有的非线性变化^[8]。

相似预测的优点是思路明确,直观性强,可进行定量预报。由于相似预测是一种间接预报方法,预报判据不直接与预报对象建立统计关系,所以它的应用不受预报对象性质的限定,可以预报连续性的气象要素如温度、湿度等,也可以预报离散度很大的降水,还可以预报天气事件的起始与终止时间等。相似预测的预报时效较长;不仅可以作短期天气定量预报、也可以作具有持续性预报对象的中期天气过程预报,还可以作时效为几个月的短期气候预测。最后还必须指出,由于相似预测包含了大气运动所有的非线性变化,所以它还能考虑极端事件出现的可能性;如可以作灾害性、关键性天气(如暴雨、连阴雨、入梅、出梅、高温、干旱、冷害、寒潮、冻害、大雾、热带气旋的移动路径及其风雨影响等)的定量预报。而灾害性、关键性天气多数属于小概率事件,预测难度很大,其它的直接预报方法均难以奏效。

相似预测方法看起来似乎简单易行,实则不然。首先,影响天气与短期气候变化的因素太多太复杂,它是大尺度天气系统和中、小尺度天气系统相互作用及天气系统与下垫面相互作用的结果。这些条件太多,难以一一吻合。再者,即使这些条件都很相近了,却并不能保证两个天气过程的后续发展也一定相似。因为初始状态的相似并不意味着未来的发展也一定相似,仅以惯性的唯一作用从前期天气推断其后续发展不仅缺乏动力依据,且有悖事实。这是以往相似预报方法的致命点所在。由于受到时代和资料的局限,以往的相似预测方法只能进行预测时刻之前的简易相似,若干有限的简易判据,不能定量地判定变化极其复杂的两个天气过程的相似程度。因此长期以来它的预报效果不甚理想,难以提高预报准确率。

相似预报要从根本上改观,唯一的出路是要进行寻找预报时段即预报日之后未来时段的相似。因未来时段的相似才是真正意义上的相似预报。近年来,电子计算机技术的巨大进步和数值天气预报的飞跃发展,数值天气预报的预报时效已从短期向中期拓展,而预报时效为几个月乃至1年以上的动力气候预报也取得一定的进展。这些进步一方面借助于电子计算机的运算速度和容量都以惊人的速度加快和扩展,另一方面则由于数值天气预报模式本身也在以惊人的速度进步。由于中期时效内包含了若干个天气过程的交替和转换,所以用于中期预报的大气模式必须具有以逼真方式描写新扰动生成和旧扰动消失的能力,能逼真地描述大气中能量产生和耗散过程^[9]。目前数值天气预报大尺度天气形势可用预报时效上已达168小时,而短期24~72小时的预报结果已相当精确,以及近年来开展的区域气候模拟和更进一步的动力气候预报^[10,11],为相似预报方法提供了动力学基础,使得未来时段即预报时段的相似有了可能。我们可以应用数值天气预报产品和动力气候预报产品进行寻找预报时段即预报日之后的未来时段的相似,从而相似预报方法得到了根本的改观。

由于大气探测手段的长足进展,我们已拥有比以往任何时候都更加丰富的大气多层次、多方位的观测资料:这不仅仅是观测站网在空间分布和观测时间上的加密观测资料,还包括了在太空中环绕地球运转的人造卫星对大气进行的三维观测资料,以及雷达和各种自动观测站非定时观测得到再加以四维同化的资料;近年来,又有了应用新技术处理而得到历史逐日和逐月再分析资料。这些日益丰富的资料使我们能够应用多元判据较精确且定量地评估两个天气过程的相似程度。

上述这些进步已使得相似预测得到根本上改观,焕发了新的生命力,并已在天气与短期气候及多种灾害性和关键性天气预测中得到日益广泛的推广应用。

既然数值天气预报已取得很大进步,那么它是不是可以取代所有其他预报方法呢?Wicner 在他所著《控制论》中指出,气象中质点的数目浩如烟海,以至要对它们初始位置以及速度做出精确的记录是完全不可能的。他还强调,利用牛顿定律或任何其他因果律,人们能预告未来时刻的一切都是概率分布^[12]。这些精辟论述指出了天气现象和过程具有随机性。因此,虽然数值天气预报不断进步,但它并不能完全取代统计预报,至少在相当长一段时间内不可能完全取数值预报产品的统计释用预报。这一方面由于大气运动的精确规律并未完全掌握,现有的数值天气预报模式,无论是差分模式还是谱模式,它们均非大气运动和热力方程组精确的解析解而是近似解。它们在求近似解的过程中都进行了或多或少的简化,在考虑边界的卷夹作用,行星边界层的摩擦作用,气团间的质量与能量的交换过程,积云对流参数化方案等方面尚不能差强人意。而对于当前的动力气候预报而言,许多重大的关键问题仍未解决。如现有的动力气候模式还不能模拟出季风系统的低频振荡、热带风场的振荡及积云对流等观测事实,这表明动力气候模式还存在不少缺陷,有些甚至是根本性的。所有这些都使得数值天气模式与实际大气运动变化仍有较大的差距,因而系统误差是不可避免的;另一方面数值预报仅靠一个时刻的资料进行积分,即使目前发展的集合预报应用了多时刻的资料,但这些都是预报时刻后的信息,而对以往的信息数值预报则完全没有记忆。天气的变化是不能割断的,它每一个时刻的变化都基于前一时刻的状态。而对于预报员未说,以往的预报经验是至关重要的,没有这些经验往往做不好预报。然而带记忆的大气运动预报模式至今尚未出现。尽管数值天气预报模式在不断精细化,但迄今为止,它最成功的预报是大尺度的天气形势预报,而对具体的天气预报时效不超过 48 小时,且预报效果与大尺度的天气形势预报相比则不可同日而语。换句话说,在有记忆的数值预报模式开发出来之前,数值预报不能完全取代统计释用预报。而统计释用预报正好弥补了数值预报的不足,它既应用了数值预报的动力结果,又记忆了历史经验。

由于数值天气预报和动力气候预报已在一定程度上解决了未来时段相似的信息问题,我们面临的问题是如何应用这些信息去构造合理的相似判据和寻找最佳相似样本的相似预测模式。本文即根据作者多年来对相似预报方法的研究体会对这些问题展开探讨。

第一章 多元判据综合相似预测方法的基本原理

1.1 多元判据综合相似预测方法的基本原理

天气或短期气候相似预测的基本思路是:具体天气或短期气候是在一定的天气形势背景条件和一定的局地环境条件下产生的。根据这些条件确定若干判定相似与否的判据(相似判据),以当前的背景和环境条件为参照(预报样本),以相似判据为依据,从众多的历史天气与短期气候记录(历史样本)中找到与当前天气形势背景和局地环境条件的变化相似的样本(相似样本)时,当前的天气发展过程就应该与相似样本的天气发展过程大体上相同,即可以把相似样本对应的天气过程作为从当前开始的未来天气发展过程预报。

以前老的相似预报方法只寻找预报时刻前期的相似,是以过去状态的静态相似作为相似的基础。本书提出的多元判据综合相似预测方法则要寻找预报时刻前期到未来时段的演变相似,即预报时刻前后状况演变的动态相似。后者比前者更全面、更合理,也更有利找到最佳相似样本。为此,预报样本由预报时刻前后若干个时段的环境条件组成,它们由数值预报产品构造而成。历史样本同样由预报时刻前后若干个时段的环境条件组成,它们由历史的实况记录构造而成。对于不同时效的预报对象,由于天气、短期气候和环境条件的时间—空间尺度不同,前后若干个时段为适应天气和环境条件的时空尺度特征也有不同表现:如短期天气预测的预报样本和历史样本由前后两天的环境条件组成,中期天气预测的预报样本和历史样本则视不同的相似区域分别由前后两个时段(5天)或前后3个时段(3天或4天)的环境条件组成,短期气候预测的预报样本和历史样本由前期若干个时段(月)的环境条件组成。

相似判据是相似预测成功与否的关键。以往的相似预报方法一般只应用一个或很少几个相似判据作相似预报。须知最简单的平面三角形相似的判定尚须三个条件,而实际大气变化是异常复杂的,少数相似判据难以反映不同层次、不同时空尺度天气系统的相互作用,从而难以找到较佳的相似样本,预测的成功率较低。应用不同层次、不同时空尺度的多元判据可以较全面地描述当前的背景和环境条件,特别要根据预报对象特点构造相关的判据,如中期预报要突出行星尺度与大尺度的天气系统的演变,就要选择3~5天的平均场来构造相似判据,除了天气形势场相似外,还增加了时空尺度与中期天气一致的波谱因子作为相似判据;暴雨具有中、小尺度特征,除了天气形势场相似外,要大量增加局地环境条件的相似判据等等,这样才有利于找到较佳的相似样本,提高预测的成功率。

以往的相似预报方法对相似程度判定只是一种定性的判定,无法评估众多样本相似程度在量上的差异,这就很难在众多样本中找到最佳相似样本。本文在这一方面进行了较深入的探讨,构造了定量判别的相似指数。当我们应用多种因子构造多元相似判据时,由于诸多相似判据的值域、量纲和量级都不一致,这就使得对它们进行综合评估产生困难。此外,在进行综合评

估时由于历史样本众多(一般有 200~400 个),又如何在众多的历史样本中突出最相似的样本。这些问题不解决,相似预测就会发生困难。为此本文定义了一个非线性分布的相似指数,它为各种差异悬殊的相似判据确定了一个判定相似程度的统一标准,并把众多的历史样本在相似程度上加以分离,以突出最相似的样本。

当所有的相似判据都变换成相似指数后,我们就可以对多元判据进行综合评估,然后根据评估结论进行相似预测。由于天气与短期气候的多样性,预报对象有短期天气预测、中期天气预测、短期气候预测和灾害性天气(大-暴雨、热带气旋等)等等,它们发生的过程与原因均不相同,因而所采用的相应相似判据和预测模式也不尽相同,这在以后的章节中将逐一加以描述。

本文提出的多元判据综合相似预测方法是一种与旧的相似预报方法在思路、原理和形式上都完全不同的新方法,以下简称相似预测模式。

1.2 相似判据

相似预测以相似判据为基础,所以相似判据是寻找相似天气过程的关键,因而它们必须能反映天气过程背景的大气环流特征。同时,由于所预报的各种天气过程差别很大,具有不同的时间-空间尺度特征,为了正确预报这些天气过程,构造的相似判据一方面必须与所预报的各种天气过程的时间-空间尺度相匹配,另一方面还必须能反映影响这些天气过程的各种尺度局地天气系统和局地环境条件。这样才能针对某种特定天气与短期气候过程进行寻找相似及进一步做出预测。

1.2.1 相似判据的构成

相似判据由预报样本和历史样本两方面组成,预报样本是相似的参照样本,因此它只有一个,一般由所预测时段或再加上前后时段的环境条件构成。本书中除第七章外,预报样本均由数值天气预报产品或动力气候预报产品构造而成。历史样本是过去的天气与短期气候过程的记录,因而它是大量的。相似样本是以预报样本为参照对历史样本进行考察与评估后寻找出的与预报样本发展过程较相似的历史样本。

在构造相似判据时,根据所预测的天气与短期气候过程的时间-空间尺度特征和判据因子的时间-空间尺度特征从大体上选择判据因子样本的时间-空间(相似区域-相似时域)范围:

预测短期天气过程时,判据因子样本为某个时刻的要素场或某个要素值,场的范围,一般中层(500 hPa)以上取地球表面的 $1/4 \sim 1/2$ 半球(如 $0 \sim 90^\circ\text{N}$, $90^\circ\text{E} \sim 180^\circ$ 或 $0 \sim 90^\circ\text{N}$, $0 \sim 180^\circ$),中层以下,取地球表面的 $1/8 \sim 1/16$ 半球或更小范围(如 $0 \sim 45^\circ\text{N}$, $90^\circ\text{E} \sim 180^\circ$ 或 $0 \sim 45^\circ\text{N}$, $90^\circ \sim 135^\circ\text{E}$),对于某些气象要素展开场,如车贝雪夫正交多项式展开典型场则取能反映展开场特征的更小的区域范围。

预测中期天气过程时,判据因子样本为某个时段(3~5 天)的要素平均场或某个要素平均值,场的范围,一般中层(500 hPa)以上取地球表面的 $1/2 \sim$ 整个半球(如 $0 \sim 90^\circ\text{N}$, $0 \sim 180^\circ\text{E}$ 或 $0 \sim 90^\circ\text{N}$, $0^\circ\text{E} \sim 180^\circ \sim 0^\circ\text{W}$),中层以下,取地球表面的 $1/4 \sim 1/2$ 半球(如 $0 \sim 45^\circ\text{N}$, $90 \sim 180^\circ\text{E}$ 或 $0 \sim 45^\circ\text{N}$, $0 \sim 180^\circ\text{E}$),对于某些气象展开场,如纬圈谱分析时则取一个纬圈的小范围。

预测短期气候过程时,判据因子样本为某个月的要素平均场或某个关键区域要素月平均值,场的范围在所有空间层取地球表面的整个半球或全球(如 $0 \sim 90^\circ\text{N}$, $0^\circ\text{E} \sim 180^\circ \sim 0^\circ\text{W}$ 或

90°S~0°~90°N, 0°E~180°~0°W)。关键区域(如环流特征量,厄尔尼诺—南方涛动关键区)的大小则依要素和预测对象的特征而定,对于某些气象要素展开场,如自然正交展开函数则取能反映展开场特征的更小的区域范围;

判定某个因子两个样本的相似程度随因子的形式不同有不同的表达方式。对于两个要素空间平面场的相似,一般采用相似系数判定;对于两个要素空间场距离的相似,一般采用欧氏距离判定;而对空间某个固定点要素值的相似,一般采用绝对距离进行判定。

1.2.2 空间平面场的相似

当相似区域的网格中纬圈总数为 K , 经圈总数为 L 的空间平面场因子的历史样本为 $X_{j,k,l}$, 预报样本为 $X_{0,k,l}$ 时,两者之间的相似程度由两个样本的 $M(M=K \times L)$ 个变量所确定,由在 M 维空间中两个向量之间的余弦来表示,即相似系数 Sx_j 为:

$$Sx_j = \cos \theta_j = \frac{\sum_{l=1}^L \sum_{k=1}^K X_{j,k,l} \cdot X_{0,k,l}}{\sqrt{\sum_{l=1}^L \sum_{k=1}^K X_{j,k,l}^2 \sum_{l=1}^L \sum_{k=1}^K X_{0,k,l}^2}}, \quad \begin{cases} j = 1, 2, \dots, J \\ k = 1, 2, \dots, K \\ l = 1, 2, \dots, L \end{cases} \quad (1.1)$$

式中 j 为历史样本序, J 为样本总数; k 为纬圈序, K 为纬圈总数; l 为经圈序, L 为经圈总数。相似系数 Sx_j 的值越大,两样本越相似。

如果 $X_{j,k,l}, X_{0,k,l}$ 都是经过标准化变换后的变量,则标准化的 Sx_j 与相关系数 Rx_j 相同:

$$Rx_j = \frac{\sum_{l=1}^L \sum_{k=1}^K (X_{j,k,l} - \bar{X}_j)(X_{0,k,l} - \bar{X}_0)}{\sqrt{\sum_{l=1}^L \sum_{k=1}^K (X_{j,k,l} - \bar{X}_j)^2 \sum_{l=1}^L \sum_{k=1}^K (X_{0,k,l} - \bar{X}_0)^2}}, \quad \begin{cases} j = 1, 2, \dots, J \\ k = 1, 2, \dots, K \\ l = 1, 2, \dots, L \end{cases} \quad (1.2)$$

式中 $\bar{X}_j$ 和 $\bar{X}_0$ 分别是因子的历史样本 $X_{j,k,l}$ 和预报样本 $X_{0,k,l}$ 的场平均值:

$$\bar{X}_j = \frac{1}{L \cdot K} \sum_{k=1}^K \sum_{l=1}^L X_{j,k,l}$$

$$\bar{X}_0 = \frac{1}{L \cdot K} \sum_{k=1}^K \sum_{l=1}^L X_{0,k,l}$$

用 1.1 式或 1.2 式判定相似程度略嫌粗糙,因在整个场中计算相似会造成过多的平滑,不易突显特征。为此分别求每个纬圈方向标准化的相似系数 $Sz_{j,k}$, 和每个经圈方向标准化的相似系数 $Sm_{j,l}$:

$$Sz_{j,k} = \frac{\sum_{l=1}^L (X_{j,k,l} - \bar{X}_{j,k})(X_{0,k,l} - \bar{X}_{0,k})}{\sqrt{\sum_{l=1}^L (X_{j,k,l} - \bar{X}_{j,k})^2 \sum_{l=1}^L (X_{0,k,l} - \bar{X}_{0,k})^2}}, \quad \begin{cases} j = 1, 2, \dots, J \\ k = 1, 2, \dots, K \\ l = 1, 2, \dots, L \end{cases} \quad (1.3)$$

式中 $\bar{X}_{j,k}$ 和 $\bar{X}_{0,k}$ 分别是因子的历史样本 $X_{j,k,l}$ 和预报样本 $X_{0,k,l}$ 的每个纬圈方向平均值:

$$\bar{X}_{j,k} = \frac{1}{L} \sum_{l=1}^L X_{j,k,l}$$

$$\bar{X}_{0,k} = \frac{1}{L} \sum_{l=1}^L X_{0,k,l}$$

$$Sm_{j,l} = \frac{\sum_{k=1}^K (X_{j,k,l} - \bar{X}_{j,l})(X_{0,k,l} - \bar{X}_{0,l})}{\sqrt{\sum_{k=1}^K (X_{j,k,l} - \bar{X}_{j,l})^2 \sum_{k=1}^K (X_{0,k,l} - \bar{X}_{0,l})^2}} \quad (1.4)$$

式中 $\bar{X}_{j,l}$ 和 $\bar{X}_{0,l}$ 分别是因子的历史样本 $X_{j,k,l}$ 和预报样本 $X_{0,k,l}$ 的每个经圈方向平均值:

$$\bar{X}_{j,l} = \frac{1}{K} \sum_{k=1}^K X_{j,k,l}$$

$$\bar{X}_{0,l} = \frac{1}{K} \sum_{k=1}^K X_{0,k,l}$$

对所有的纬圈的 $Sz_{j,k}$ 按纬圈数平均得到场的纬向相似系数 $\bar{Sz}_j$:

$$\bar{Sz}_j = \frac{1}{K} \sum_{k=1}^K Sz_{j,k} \quad (1.5)$$

对所有的经圈的 $Sm_{j,k}$ 按经圈数平均得到场的经向相似系数 $\bar{Sm}_j$:

$$\bar{Sm}_j = \frac{1}{L} \sum_{l=1}^L Sm_{j,l} \quad (1.6)$$

因子 X 的空间平面场的相似系数 Sx_j 为:

$$Sx_j = \bar{Sz}_j \cdot \bar{Sm}_j \quad (1.7)$$

1.2.3 空间平面场距离的相似

空间平面场的历史样本 $X_{j,k,l}$ 与预报样本 $X_{0,k,l}$ 之间的距离表征了它们在性质上的差异, 该差异越小, 两样本越相似; 反之, 差异越大则越不相似。本文用欧氏距离 Dx_j 评估空间平面场距离的相似:

$$Dx_j = \sqrt{\sum_{k=1}^K \sum_{l=1}^L (X_{j,k,l} - X_{0,k,l})^2}, \quad \begin{cases} j = 1, 2, \dots, J \\ k = 1, 2, \dots, K \\ l = 1, 2, \dots, L \end{cases} \quad (1.8)$$

式中 j, k, l, J, K, L 的说明同 1.2.2 节。

1.2.4 空间点要素值的相似

对空间某个固定点要素值的历史样本 X_j 与预报样本 X_0 之间的距离表征了它们在性质上的差异, 该差异越小, 两样本越相似; 反之, 差异越大则越不相似。本文用绝对距离 Δx_j 判定空间点要素值的相似。

$$\Delta x_j = |X_j - X_0|, \quad (j = 1, 2, \dots, J) \quad (1.9)$$

式中 j, J 的说明同 1.2.2 节。

对于暴雨和台风风雨影响等灾害性天气的影响因素众多而且复杂。由于目前观测网点的大尺度布局和定时观测得到的预测信息有限, 构造多元判据从多个层面要素的整体场的相似以捕获大尺度的场信息。整体场的相似虽然体现了整体场特征, 但平滑了局地特征, 而正是这些局地特征(如某要素场的某些点的特征值与灾害性天气发生事件之间的高相关)隐含了中、小尺度天气系统的预报信息。

为了从大尺度天气形势信息中提取灾害性天气的局地要素变化的信息, 本文将采用相关与相似相结合的方法构造局地灾害性天气特征判据。

如果在预测关键区域要素 $X_{j,k,l}$ 因子与灾害性天气发生事件 Y_j 之间的相关系数:

$$R_{x_{k,l}} = \frac{\sum_{j=1}^J (X_{j,k,l} - \bar{X}_{k,l})(Y_j - \bar{Y})}{\sqrt{\sum_{j=1}^J (X_{j,k,l} - \bar{X}_{k,l})^2 \sum_{j=1}^J (Y_j - \bar{Y})^2}}, \quad \begin{cases} j = 1, 2, \dots, J \\ k = 1, 2, \dots, K_r \\ l = 1, 2, \dots, L_r \end{cases} \quad (1.10)$$

式中

$$\bar{X}_{k,j} = \frac{1}{J} \sum_{j=1}^J X_{j,k,l}$$

$$\bar{Y} = \frac{1}{J} \sum_{j=1}^J Y_j$$

K_r, L_r 分别为预测关键区的纬圈、经圈总数。

当空间坐标为 (p, q) 的点的绝对相关系数 $R_{x_{p,q}}$ 为预测关键区的最高绝对相关系数即:

$$R_{x_{p,q}} = |\max(R_{x_{k,l}})|, \quad \begin{cases} p \in [1, 2, \dots, K_r] \\ q \in [1, 2, \dots, L_r] \end{cases} \quad (1.11)$$

则用绝对距离 Δx_j 判定空间点要素值的相似:

$$\Delta x_j = |X_{j,p,q} - X_{0,p,q}| \quad (1.12)$$

式中 j, k, l, J, K, L 的说明同 1.2.2 节。

1.2.5 空间平面场的环流相似

对空间平面场相似和空间平面场距离相似进行综合可得到空间平面场的环流相似 C_j [1]:

$$C_j = \frac{1}{Dx_j} S z_j \cdot S m_j, \quad (j = 1, 2, \dots, J) \quad (1.13)$$

1.2.6 构造相似判据的若干原则

在进行相似判定时,根据不同的判据因子分别采用以上空间平面场的相似、空间平面场距离的相似、空间点要素值的相似和空间平面场的环流相似等 4 种形式的相似判据。式(1.1)~(1.12)为典型的表达式,对于不同的判据因子,即使采用相同形式的判据,表达式在细节上也有不同的形式,以下章节将逐一介绍。

构造相似判据的因子有多种,常用的有高度场、气压场、温度场、湿度场、水平风场和垂直运动等要素场,以及它们的导出因子场如变量场、散度场、涡度场、通量场、平流场,还可以应用要素场的正交展开如自然正交展开的 EOF 函数、切贝雪夫正交多项式展开的典型场、以及应用傅里叶级数进行谐谱分析展开的纬圈谱因子和球谐展开参数等等。而对于灾害性天气预测,还可以应用一些特别构造的因子,如暴雨预测时与暴雨发生高相关的空间特定点的要素及变量,台风路径预测时的引导气流、以台风为中心的移动网格因子和气候平均转向点等。

对不同的预报对象,需构造不同的相似判据。对于较简单的天气过程(如一般的短期天气过程),仅需要多元天气形势相似就能进行相似预测,判据的总数相对可以少一些,而对于较复杂的天气过程,如暴雨、中期天气过程、台风路径和风雨影响等,除了各种尺度天气系统相似外,还需要判定局地环境条件相似,才能进行相似预测,判据的总数则相对要多一些。究竟要构造多少相似判据,需视预报对象而定。

天气相似预测以数值天气预报为基础,其预报样本由数值天气预报输出产品构成。所以,

相似判据还受数值天气预报输出产品的制约,即使对于预报对象而言是至关重要的因子,只要该因子的要素无数值天气预报输出产品,就不能作为相似判据。目前,数值天气预报输出产品的品种有限,很多对于预报对象是至关重要的因子要素没有数值天气预报输出产品。由于现有的数值天气预报模式还很不完善,预报的系统误差不可避免,因此输出产品精度不是很高,这种制约显然要影响相似预测的效果。

1.3 相似指数

由于影响天气变化的因素众多而且复杂,单一的判据难以判断两个天气过程的相似性。只有通过综合评估多元判据的相似性,才能找到较佳的相似天气过程。当我们应用多种因子构造相似判据时,由于诸多相似判据的值域、量纲和量级都不一致,这就使得对它们进行综合评估产生困难。另外,在进行综合评估时由于历史样本众多,又如何众多的历史样本中突出最相似的样本,这也是关系到相似预测成败的关键问题。为此,本文定义一个相似指数,它为各种差异悬殊的相似判据确定了一个判定相似程度的统一标准;同时,它还把众多的历史样本在相似程度上加以分离,以突出最相似的样本。相似指数可以有两种形式:相似指数 I 和 相似指数 II 。相似指数 I 以历史样本在相似空间的点与相似函数界限值的绝对距离进行定义,相似指数 II 以历史样本依相似函数的值重新排序进行定义。两种形式的相似指数大体上相同,相似指数 II 受历史样本的总量限制,而相似指数 I 无此限制。

1.3.1 相似指数 I

对于某个作为判据的因子序列,历史样本 $x_{i,j}$ 与预报样本 $x_{i,0}$ 确定了某种相似关系的相似函数为 $f_{i,j}(x, x_0)$ 。解析式 $f_{i,j}(x, x_0)$ 可以是 1.2 节所述的空间平面场的相似系数 Sx_j 、空间平面场的欧氏距离 Dx_j 或空间点要素值的绝对距离 Δx_j 等函数关系,下标 i 为判据序,下标 j 为判据因子的样本序。

当第 i 个判据因子序列的最大值 $\max(f)_i$ 与最小值 $\min(f)_i$ 确定后,将极差十等分:

$$\Delta D_i = [\max(f)_i - \min(f)_i] / 10 \quad (1.14)$$

判据因子序列的界限值 C_i (视相似函数 $f_{i,j}(x, x_0)$ 不同可以为极大值或极小值) 确定后,以 ΔD_i 为距离单位,当 $Df_{i,j}$ 为相似函数 $f_{i,j}(x, x_0)$ 在相似空间的点 (i, j) 与界限值 C_i 的绝对距离,即

$$Df_{i,j} = |f_{i,j}(x, x_0) - C_i| \quad (1.15)$$

由此定义相似指数 $SI_{i,j}$:

$$SI_{i,j} = \begin{cases} 10, & Df_{i,j} = 0 \\ 10 - k, & (k-1) \frac{\Delta D_i}{4} < Df_{i,j} \leq k \frac{\Delta D_i}{4}, \quad k = 1, 2, 3, 4 \\ 6 - k, & (k+1) \frac{\Delta D_i}{2} < Df_{i,j} \leq (k+2) \frac{\Delta D_i}{2}, \quad k = 1, 2 \\ 4 - k, & (3k+1) \frac{\Delta D_i}{2} < Df_{i,j} \leq (3k+4) \frac{\Delta D_i}{2}, \quad k = 1, 2 \\ 1, & 5\Delta D_i < Df_{i,j} \leq 7\Delta D_i \\ 0, & 7\Delta D_i < Df_{i,j} \leq 10\Delta D_i \end{cases} \quad \begin{pmatrix} i = 1, 2, \dots, I \\ j = 1, 2, \dots, J \end{pmatrix} \quad (1.16)$$

式中 i 为判据因子序, j 为历史样本序, I 为判据总数, J 为样本总数。

图 1.1 为相似指数 $SI_{i,j}$ 随 $Df_{i,j}$ 的分布。定义表明, 相似指数 $SI_{i,j}$ 不与 ΔD_i 的均等区间成比例。其结果仅有少量 $Df_{i,j}$ 甚小即与预报样本接近的若干历史样本获得高 $SI_{i,j}$ 值, 大多数不甚相似的历史样本被赋予低 $SI_{i,j}$ 值。这使得该判据因子历史样本序列变换成非线性分布的相似指数 $SI_{i,j}$ 序列时, 样本间的差别得到扩大即样本在相似程度上被分离。

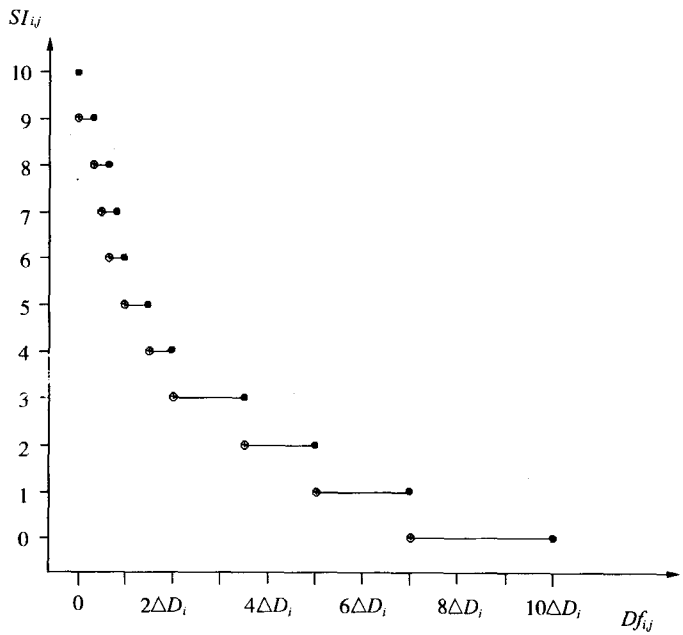


图 1.1 相似指数的 $SI_{i,j}$ — $Df_{i,j}$ 分布

作为比较, 定义以下按均等 ΔD_i 区间分布的相似指数 $SI'_{i,j}$:

$$SI'_{i,j} = \begin{cases} 10 - k, & k\Delta D_i < Df_{i,j} \leq (k + 1)\Delta D_i, \quad k = 1, 2, \dots, 9 \\ 10, & 0 \leq Df_{i,j} \leq \Delta D_i \end{cases} \quad (1.17)$$

图 1.2 为中期天气相似预测中历史样本在 16 个判据下, 按均等与不均等 ΔD_i 区间划分得到相似指数 $SI'_{i,j}$ 和 $SI_{i,j}$ 出现的平均频率 f (%) 分布。在均等区间情况下, 历史样本 $SI'_{i,j}$ 的频率分布为准正态分布: $SI'_{i,j}$ 值为 5~7 区间历史样本出现频率最高, 从两边往高、低指数缓慢降低。在不均等区间情况下, 0~3 区间的低 $SI_{i,j}$ 值的历史样本出现频率最高, 而后出现频率随着 $SI_{i,j}$ 值的升高而陡降。

表 1.1 两种划分区间相似指数的历史样本频率总数分布

频率总数	低指数值(0~3)	中指数值(4~7)	高指数值(8~10)
均等区间 ($\sum SI'_{i,j}$)	14%	61%	25%
不均等区间 ($\sum SI_{i,j}$)	84%	11%	5%

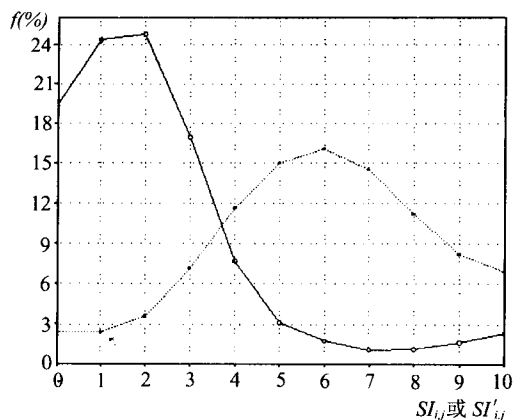


图 1.2 不同划分区间相似指数的历史样本频率分布

虚线为均等区间 $SI'_{i,j}$, 实线为不均等区间 $SI_{i,j}$

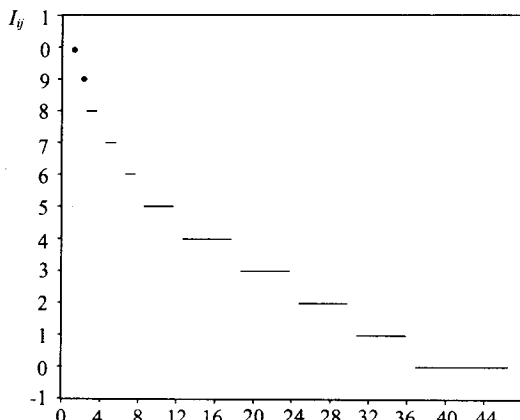


图 1.3 相似指数 $I_{i,j}$ 随 $m_{i,j}$ 排序的分布

为了比较不同划分区间的历史样本频率总数分布,这里分别统计了 $SI_{i,j}$ 与 $SI'_{i,j}$ 为低值(0~3)、中值(4~7)和高值(8~10)三个值段历史样本频率总数。统计结果如表 1.1。从表中可见, $SI'_{i,j}$ 为低值、中值和高值的历史样本频率总数分别为 14%、61% 和 25%。而在不均等区间情况下, $SI_{i,j}$ 为低值、中值和高值的历史样本频率总数分别为 84%、11% 和 5%。比较可见,判据因子经过按不均等区间分布函数非线性变换后,高值指数历史样本出现频率总数下降了 20%, 中值指数历史样本出现频率总数下降了 50%, 而低值指数频率历史样本出现频率总数上升了 70%。由此可见,判据因子经过按不均等区间分布函数的非线性变换后,仅有少量历史样本获得高 $SI_{i,j}$ 值,大多数历史样本被赋予低 $SI_{i,j}$ 值。从而相应加大了相似样本的相似权重,减小了不相似样本的相似权重,实现了历史样本在相似程度上的分离。

1.3.2 相似指数 II

对于某个作为判据的因子序列,历史样本 $x_{i,j}$ 与预报样本值 $x_{i,0}$ 确定了某种相似关系的相似函数为 $f_{i,j}(x, x_0)$ 。解析式 $f_{i,j}(x, x_0)$ 可以是 1.2 节所述的空间平面场的相似系数 Sx_j 、空间平面场的欧氏距离 Dx_j 或空间点要素值的绝对距离 dx_j 等函数关系,下标 i 为判据序,下标 j 为判据因子的样本序。

对于第 $i(i=1, 2, \dots, I)$ 个因子,其样本序为 j 的历史样本 $x_{i,j}$ 与预报样本 $x_{i,0}$,若以年份递增排序的历史年份序列 $\{N_j\}$ ($j=1, 2, \dots, J$),按照 $f_{i,j}(x_{i,j}, x_{i,0})$ 的值即 $x_{i,j}$ 与 $x_{i,0}$ 的相似程度进行重新排列,使得到新序列 $\{m_{i,j}\}$ 所对应的 $f_{i,j}(x_{i,j}, x_{i,0})$ 的值从大到小成单调递减排列,则定义相似指数 $I_{i,j}$:

$$I_{i,j} = \begin{cases} 10, & m_{i,j} = 1 \\ 10 - k, & 2^{k-1} < m_{i,j} \leq 2^k \quad (k = 1, 2) \\ 8 - k, & n_0 + 2^{k-1} < m_{i,j} \leq 4 + 2^k \quad (k = 1, 2, 3; \quad n_0 = \begin{cases} 3, k = 1 \\ 4, k \neq 1 \end{cases}) \\ 5 - k, & 6(k+1) < m_{i,j} \leq 6(k+2) \quad (k = 1, 2, 3, 4) \\ 0, & 36 < m_{i,j} \leq J \end{cases}$$

(1.18)