

基于粒子群算法的天气滚动预报集成技术

熊聪聪¹, 许淑祎¹, 徐 姝²

(1. 天津科技大学计算机科学与信息工程学院, 天津 300222; 2. 天津市气象局, 天津 300074)

摘 要: 针对集成预报的权重分布法, 采用具有全局寻优特点的粒子群算法优化其权重系数, 以更好地集中各种预报模式的优势. 以天津地区天气预报为实例, 建立基于粒子群算法的集成预报模型, 实现多种预报模式的集成处理, 结果表明该方法对于短时天气预报具有较好的准确度.

关键词: 集成预报; 粒子群算法; 权重分布法

中图分类号: TP301.6; P456.1

文献标志码: A

文章编号: 1672-6510(2012)01-0075-04

Integration Technology of Rolling Weather Forecast Based on PSO

XIONG Congcong¹, XU Shuyi¹, XU Shu²

(1. college of Computer Science and Information Engineering, Tianjin University of Science & Technology, Tianjin 300222, China; 2. Tianjin Meteorological Institute, Tianjin 300074, China)

Abstract: As to the weight distribution method of integrated forecast, using PSO (particle swarm optimization) which has good global optimization characteristics can optimize its weight factor and make the best of the advantages of various forecasting models. Take the weather forecast of Tianjin region as example. The integrated forecasting model based on PSO was established and the integration of several forecasting models was realized. Results show that the method has an acceptable accuracy for short-term weather forecasting.

Key words: integrated forecast; PSO; weight distribution method

集成预报是应用数学模型对同一种预报元素多种预报模式的预报结论进行综合集成, 从而得出一个统一预报结论的一种统计预报方法, 与用于预报集成的各个子预报方法相比通常能够获得更好的预报效果. 集成预报的关键在于构造具有理论依据的合理可靠的适应度函数, 并选取较好的集成预报方案, 以确保其结果在实际应用中较为客观准确^[1]. 虽然国内外的气象学者对此已有不少研究, 但是由于天气系统的复杂性, 仍没有开发出成熟而有效的集成预报技术.

目前用于集成预报的主要方法有多元回归集成法、权重分布法、神经网络集成法等^[2], 其中权重分布法建立方程简单、关系式稳定、不受样本数量的限制. 权重系数的确定是权重分布法集成预报的关键. 在实际业务中, 主要采用简单的算术平均或是根据各种方法的历史预报事先设定权重, 其主观性较强, 不能做

到客观定量, 预报效果不稳定. 遗传算法在天气集成预报中已有过研究^[3], 但是由于遗传算法需要大量计算以及早熟等缺陷, 还存在运行时间过长、计算结果不够稳定等问题.

粒子群优化算法 (PSO)^[4] 是一种基于群体智能的全局随机搜索算法, 通过个体间的协作和竞争实现复杂空间最优解的搜索, 其概念简单、实现容易、精度高、收敛快, 具有很强的发现较好解的能力, 对非线性、多峰问题均有较强的全局搜索能力, 在科学研究与工程实践中得到了广泛的关注. 目前, 粒子群算法在集成预报中的应用研究还不多. 为了提高综合预报能力, 本文提出了基于粒子群算法训练权重系数的方法, 利用粒子群算法得出各预报模式预报结果的权重系数, 进而得出某一气象元素基于权重分布法的集成预报模型, 并对集成预报结果进行分析.

收稿日期: 2011-07-06; 修回日期: 2011-11-02

基金项目: 中国气象局气象关键技术集成与应用项目 (CAMGJ2012M04)

作者简介: 熊聪聪 (1961—), 女, 四川泸州人, 教授, xiongcc@tust.edu.cn.

1 权重分布法

权重分布法的基本原理是对不同预报模式使用不同的权重,即把预报效果较好的模式赋予较高权重,相对差的权重较低,为了更好地体现各预报模式的客观预报能力,考虑到不同预报模式使用的预报因子可能不同,不同地区不同时间点的预测能力也不同,对同一种预报模式不同站点和时间点使用不同的权重,以集中各种预报模式的优势,提高综合预报水平^[5]。

权重分布法在集成预报中具体应用的方法为

$$\bar{F}_{j,t} = \sum_{i=1}^m (W_{i,j,t} \cdot F_{i,j,t}) \quad (1)$$

式中: i 为预报模式; j 为站点号; t 为一次预报时效中的预报时间点; m 为预报模式的个数; $F_{i,j,t}$ 为某一气象元素在站点 j 第 i 个模式第 t 个时间点的预报值; $W_{i,j,t}$ 为站点 j 第 i 个模式第 t 个时间点的权重系数; $\bar{F}_{j,t}$ 为该气象元素在站点 j 第 t 个时间点的集成预报值。

2 粒子群优化算法

在粒子群算法中,每个个体被称为一个“粒子”,每个粒子代表一个潜在的解,并且都有一个被目标函数决定的适应值^[6]。设群体由 m 个粒子构成,第 i 个粒子的 D 维位置向量为 $\mathbf{z}_i = (z_{i1}, z_{i2}, \dots, z_{iD})$; 粒子 i 的飞行速度为 $\mathbf{v}_i = (v_{i1}, v_{i2}, \dots, v_{id}, \dots, v_{iD})$ 。在每次迭代中,粒子通过跟踪两个“极值”来更新自己。一个是粒子当前搜索到的最优位置 $\mathbf{P}_i = (p_{i1}, p_{i2}, \dots, p_{id}, \dots, p_{iD})$, 其相应的适应度称为个体极值 p_{best} ; 另一个是整个粒子群当前搜索到的最优位置 $\mathbf{P}_g = (p_{g1}, p_{g2}, \dots, p_{gd}, \dots, p_{gD})$, 相应的适应度叫做全局极值 g_{best} ^[7]。在搜索最优位置的过程中,粒子根据式(2)和式(3)更新自己的速度和位置:

$$v_{id}^{k+1} = \omega \cdot v_{id}^k + c_1 \cdot r_1 \cdot (p_{id} - z_{id}^k) + c_2 \cdot r_2 \cdot (p_{gd} - z_{id}^k) \quad (2)$$

$$z_{id}^{k+1} = z_{id}^k + v_{id}^{k+1} \quad (3)$$

式中: $i=1, 2, \dots, m$; $d=1, 2, \dots, D$; k 为迭代次数; ω 为惯性权重; c_1 和 c_2 为正的常数,称为学习因子或加速因子,使粒子具有自我总结和向群体中优秀个体学习的能力,通常取 $c_1 = c_2 = 2$; r_1 和 r_2 为 $[0, 1]$ 之间

的随机数。粒子群中每个粒子从初始位置和速度开始,根据式(2)和式(3)进行迭代,直到满足算法终止条件。

惯性权重^[6]起着权衡局部最优能力和全局最优能力的作用。为了使粒子群算法有更大的机会找到全局最优值,在本文中,惯性权重采用的不是固定值,而是设为一个随迭代次数线性变化的函数,具体函数式为

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{K} \times k$$

式中: $\omega_{\max}$ 为惯性权重的最大值; $\omega_{\min}$ 为惯性权重的最小值; K 为最大迭代次数; k 为当前迭代次数。

3 应用粒子群算法优化集成预报权重系数

对于天气预报或气候预测产品的集成,一般认为预报误差最小时的集成效果最好,因此可以用预报误差值来构造适应度函数^[2]。在本文中,适应度函数采用集成预报值与实况值的误差平方和,即

$$f = \sum_{t=1}^n \left(\sum_{i=1}^m (W_{i,j,t} \cdot F_{i,j,t}) - F'_{j,t} \right)^2 \quad (4)$$

式中: t 为预报时间点, n 为一次预报时效的时间点数; i 为预报模式, m 为参与集成的预报模式的个数; $F_{i,j,t}$ 为某一气象元素在站点 j 第 i 个模式第 t 个时间点的预报值; $W_{i,j,t}$ 为站点 j 第 i 个模式第 t 个时间点的权重系数; $F'_{j,t}$ 为站点 j 在第 t 个时间点的实况值。一次预报时效中某一气象元素在某站点上某种预报模式的预报值所对应的权重系数代表粒子群中一个粒子 $n \times m$ 维的位置向量。

利用粒子群算法优化集成预报权重系数的具体步骤:

- (1) 选取合适的训练样本,初始化所有的粒子,根据约束条件,随机设置粒子的初始位置和速度;
- (2) 根据式(4)计算每个粒子的适应值,取适应值最小的为最优解 z_{best} ,并且将第 i 个粒子的最优位置存为 $\mathbf{P}_i$,所有粒子经历的最优位置存为 $\mathbf{P}_g$;
- (3) 根据式(2)、式(3)对粒子的速度和位置进行更新,并重复步骤(2);
- (4) 判断是否达到最大迭代次数,如果没有达到设定的最大迭代次数,转到步骤(3)继续执行程序;否则输出最优解 z_{best} ,结束循环。

最优解 z_{best} 即为通过粒子群算法优化的权重系数,若给出几种模式下的预测值,根据其对应的权重

系数, 结合式(1)则可以求得相应的集成预报值.

4 应用实例

4.1 数据预处理

根据天津市气象局提供的气象资料, 选取 2011 年 2 月 1 日—2 月 28 日(筛选后 25 d 可用)欧洲中心 EC 模式、MM5 模式、T213 模式和 Grapes 模式格点温度预报结果以及天津地区 13 个站点的地面温度实况值作为训练样本, 3 月 1 日—3 月 3 日的数据作为检测样本. 温度实况值为一日 8 次观测结果, 由于 4 种模式的预报结果存在多种预报时效和间隔, 为了与实况值一一对应, 中间时效预报通过内插获得.

各模式下的原始数据为格点温度数据, 根据集成预报的需要, 采用双线性内插法将格点预报值内插到单站^[8]. 双线性内插的原理是找到待求站点所在网格的 4 个格点所对应的气象数据, 进行二维线性插值计算.

$$\bar{A} = (1 - d_x) \times [(1 - d_y) \times A_{i,j} + d_y \times A_{i+1,j}] + d_x \times [(1 - d_y) \times A_{i,j+1} + d_y \times A_{i+1,j+1}]$$

式中: A 为各模式格点预报值; $\bar{A}$ 为插值后站点的预报值; d_x 、 d_y 为站点在网格中的相对位置. 各项的关系如图 1 所示.

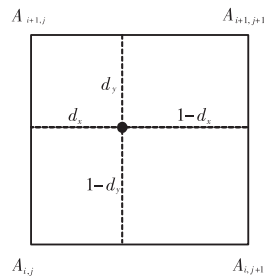


图 1 格点预报值内插到单站示意图
Fig.1 Schematic diagram of grid forecast value interpolation into a single station

4.2 集成预报结果及误差分析

结合粒子群算法的经验设置值^[9]并经过多次实验, 设置粒子群优化算法的参数为: 粒子个数 25, 迭代次数 100, 惯性权重最大值 0.9, 惯性权重最小值 0.1, 粒子初始化的下界为 -1, 粒子初始化的上界为 1, 最大速度为 0.5, 学习因子 $c_1 = c_2 = 2$.

因为遗传算法(GA)在集成预报中已经得到较好的应用^[3], 本文将基于粒子群算法的集成预报结果与其进行比较, 并分析两种算法在集成预报中的适

用性.

以站点号 54 623(塘沽)为例, 预报时效为 72 h 内 3 h 间隔, 4 种预报模式的温度 PSO 集成预报结果和 3 预报模式(去除 T213)^[3]的温度 GA 集成预报结果见图 2. 可以看出, 两种集成预报曲线与实测温度曲线基本趋于吻合; GA 集成预报需要去除预报性能较差的 T213^[3], 才能达到和 PSO 集成预报近似的效果, 可见成员个体的预报性能对 PSO 集成预报影响比 GA 集成预报小.

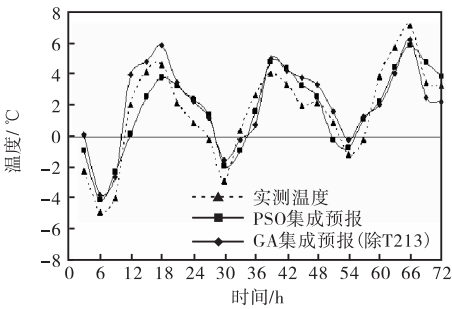


图 2 温度集成预报结果
Fig.2 Integrated temperature forecast results

4 种预报模式和 2 种集成预报模式的误差分析^[10]见表 1. 衡量预报效果指标的平均绝对误差^[11], PSO 集成预报为 1.431 7, 均低于 4 种模式的平均绝对误差, 且均方误差也优于其他模式.

表 1 集成预报与 4 种模式温度预报比较

Tab.1 Integrated forecast in temperature compared with four models

预报模式	均方差	均方根误差	平均绝对误差	平均相对误差
PSO 集成预报	1.902 5	0.309 5	1.431 7	0.112 5
GA 集成预报	1.908 9	0.447 2	1.512 0	0.191 6
欧洲中心 EC	2.339 9	0.758 9	2.231 9	0.321 7
MM5	2.367 3	1.437 4	2.494 1	0.381 2
T213	2.674 1	1.123 0	3.175 0	0.358 3
Grapes	2.586 3	1.072 9	2.167 3	0.392 6

5 结 语

本文对基于粒子群算法的集成天气预报技术进行了研究. 通过实验结果可以看出, 基于粒子群算法的集成预报结果明显优于各模式的预报结果, 也比基于遗传算法的集成预报稍好一些, 对温度预报达到了一定的精度, 实验效果较好.

利用粒子群优化算法训练权重系数过程中, 其参数的设置直接影响优化结果, 并且需要大量数据和反复进行实验, 由于天津市气象台提供的原始有效数据

有限,所以集成结果还有待进一步优化.虽然 PSO 集成预报的结果对成员模式的预报性能要求比 GA 集成预报低,但是各模式的预报性能对集成结果也有一定影响,因此可以从此方面做出改进,选取预报效果更好的预报模式.由于粒子群优化算法本身的一些缺陷,常常会受 p_{best} 和 g_{best} 的影响而陷入非最优区域,算法容易发生早熟收敛现象,因此还可以考虑基于遗传思想改进粒子群算法.

参考文献:

- [1] 吉延艳,熊方,黄继用.一种短期气候预测集成技术的应用研究[J].高原气象,2002,21(5):522-525.
- [2] 杞明辉,许美玲,程建刚,等.天气预报集成技术和方法应用研究[M].北京:气象出版社,2006:77-87.
- [3] 熊聪聪,王静,宋鹏,等.遗传算法在多模式集成天气预报中的应用[J].天津科技大学学报,2008,23(4):80-84.
- [4] 纪震,廖惠连,吴青华.粒子群算法及应用[M].北

京:科学出版社,2009:14-21.

- [5] 陈桂英.权重分布法集成预报试验[J].应用气象学报,2000,11(增刊):51-57.
- [6] 杨志鹏,朱丽莉,袁华.粒子群优化算法研究与发展[J].计算机工程与科学,2007,29(6):61-64.
- [7] Zhang Li-ping, Yu Huan-jun, Hu Shang-xu. Optimal choice of parameters for particle swarm optimization[J]. Journal of Zhejiang University Science, 2005(6A):528-534.
- [8] 李军龙,张剑,张丛,等.气象要素空间插值方法的比较分析[J].草业科学,2006,23(8):6-11.
- [9] 张雯雯,王刚,朱朝晖,等.粒子群优化算法种群规模的选择[J].计算机系统应用,2010,19(5):125-128.
- [10] 吴建生,刘丽萍,金龙.粒子群-神经网络集成学习算法气象预报建模研究[J].热带气象学报,2008,24(6):679-686.
- [11] 赵声蓉.多模式温度集成预报[J].应用气象学报,2006,17(1):52-58.

责任编辑:常涛

(上接第65页)

升温阶段也可以保持温度按照人工设定的升温速率进行变化.在满足了加速量热仪对温度控制的要求的同时,也为其他场合的电加热温度控制提供了一种可行的方法.

参考文献:

- [1] 崔席勇,花福安,李建平,等.连续退火模拟实验机的模糊专家温度控制系统[J].东北大学学报:自然科学版,2010,31(7):957-961.
- [2] 刘桂英,杭进.动态矩阵控制在电阻炉温度控制系统中的应用[J].北华大学学报:自然科学版,2009,10(3):285-288.
- [3] 陆平,顾子明,郭汉清.新型预测函数控制及其在电加热炉中的应用[J].信息技术,2009,33(8):10-13.
- [4] 帅建忠.PLC在热弯炉加热器功率调节控制中的应用

[J].可编程控制器与工厂自动化,2009(3):56-57.

- [5] 蒋小平,田元,刘建,等.基于非线性模型的电加热温度控制器的设计[J].自动化与仪表,2010,25(3):29-32.
- [6] 侯建德,傅智敏,黄金印.物质热稳定性评估方法研究[J].消防科学与技术,2002,21(2):21-24.
- [7] Lee Chuen Chien. Fuzzy logic in control systems: Fuzzy logic controller-part I[J]. IEEE Transactions on Systems, Man and Cybernetics, 1990, 20(2):404-418.
- [8] Abonyi J, Babouska R, Szeifert F. Fuzzy expert system for supervision in adaptive control[C]//Artificial Intelligence in Real-Time Control (AIRC-2000) Proceedings Volume from 9th IFAC Symposium. Kidlington, UK: Elsevier Science, 2001:241-246.

责任编辑:常涛