

基于 SCE-UA 算法的多模式降雨集成预报技术研究

钟 伟¹, 朱从飞¹, 张卫国¹, 黄 峰²

(1.宁波市水利水电规划设计研究院, 浙江 宁波 315192; 2.河海大学 水文水资源学院, 江苏 南京 210098)

摘要: 高精度的降雨预报对于防汛决策具有重要指导意义. 以中国气象局 GRAPES_MESO 模式、中国气象局 T639 模式、美国国家环境预报中心 GFS 模式和上海区域中尺度 SMS-WARMS 模式为基础, 采用 SCE-UA 算法估计各预报模式权重系数, 通过加权平均建立多模式降雨集成预报, 并以宁波市 2016 年 8 月 1 日至 2016 年 10 月 31 日和 2017 年 9 月 1 日至 2017 年 11 月 30 日的降雨过程为例, 从均方根误差、相关系数、TS 评分等多个方面分析集成预报的预报效果. 结果表明: 基于 SCE-UA 算法的多模式降雨集成预报的预报效果不仅优于各单一预报模式, 而且优于遗传算法和 BP 神经网络集成预报, 表现出较好的适用性.

关键词: SCE-UA 算法; 多模式集成预报; 降雨预报

中图分类号: TP391

文献标志码: A

文章编号: 1001-5132 (2018) 06-0027-05

在防汛防台过程中, 降雨预报扮演着重要的角色. 高精度的降雨预报可有效提升防汛决策的科学性、准确性, 促进经济社会的可持续发展. 目前, 防汛部门可直接获取的预报数据精度和空间分辨率普遍较低, 难以满足实际防汛的需求. 近年来, 气象数值预报技术发展迅速, 涌现出一大批有效的数值预报模式, 提高了降雨预报的精度^[1-3]. 然而, 不同的预报模式在空间分辨率、预报时效、预报精度等诸多方面均存在差异, 如何选择最能反映当地降雨实际的预报结果是防汛决策面临的一大难题. 集成预报技术是通过一定的理论方法, 对不同数值预报模式结果进行有机整合的预报技术^[4]. 该技术充分考虑了各单一数值预报模式的预报特点, 预报效果往往优于参与集成的各单一数值预报. 李倩等^[5]采用 BP 人工神经网络法构建了 850 hPa 高度区域温度的集成预报模型, 该模型在冷空气过程的个例中取得了较好的预报效果, 且在平均绝对误差和相关系数等方面均优于其 3 个子模式. 熊聪聪等^[6]根据遗传算法和粒子群算法的特点, 采用基于二者的混合算法构建集成预报模型, 实例研究结果表明, 该集成预报模型不仅优于

单一预报模式, 而且比单一遗传算法和单一粒子群算法集成效果更好.

单纯多边形进化算法(简称 SCE-UA 算法)是一种综合了确定性搜索、随机搜索和生物竞争进化等众多方法优点的优化算法, 能快速、准确地搜寻到全局最优解, 已在水文预报领域得到广泛应用^[7-10]. 目前, 该方法在集成天气预报中鲜有涉及. 因此, 本文采用 SCE-UA 算法估计各单一预报模式的权重系数, 构建多模式降雨集成预报模型, 并以宁波市历史降雨为例, 分析集成预报模型的预报效果.

1 降雨集成预报模型构建

由于预报方法、参数设置、初始条件等方面的差异, 不同预报模式的预报精度通常存在一定的差异, 但均可为未来降雨预报提供有效信息. 为充分利用不同预报模式的优势, 多模式集成预报模型通过一定的方法, 将各机构的预报结果加以集成, 以获取更加符合实际的预报结果. 目前, 常用的集成预报方法包括权重分布法、多元回归集成法和神经网络法等, 其中权重分布法根据不同预报模式的历史预报效果, 赋予不同权重系数, 通过加

权平均得到最终的集成预报^[11], 其流程如图 1 所示.

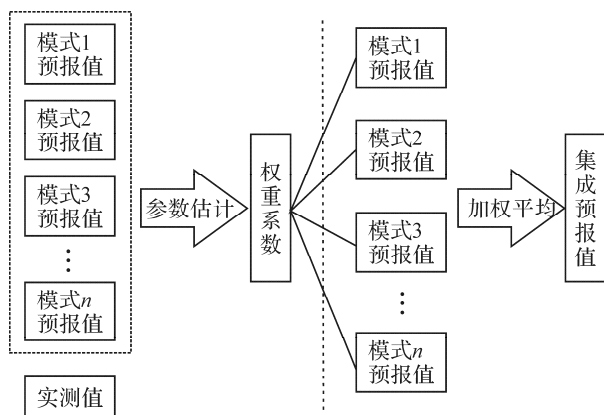


图 1 基于权重分布法的多模式降雨集成预报流程

为了真实地反映降雨的时空差异, 不同站点不同预报时间点的权重系数均有所差别. 对任一雨量站点和预报时间点, 集成公式为:

$$P' = \sum_{i=1}^m W_i P_i,$$

式中: P' 为集成预报结果; m 为预报模式个数; i 为参与集成的预报模式序号; W_i 为第 i 个预报模式的权重系数; P_i 为第 i 个预报模式的降雨预报值.

由上可知, 采用权重分布法进行多模式降雨集成的关键是计算不同模式的权重系数. 求解权重系数的方法有很多, 由此构建的集成模型也千差万别, 需采用一定的指标评价模型的优劣. 常用的评价指标有均方根误差 E_{RMS} 、相关系数 r 和 TS 评分. 均方根误差能够反映预报场和实测场的总体误差大小, 相关系数能够衡量预报场和实测场变化趋势的一致程度, 计算公式分别为^[4]:

$$E_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - P'_i)^2},$$

$$r = \frac{\sum_{i=1}^n (P_i - \bar{P})(P'_i - \bar{P}')}{\sqrt{\sum_{i=1}^n (P_i - \bar{P})^2 \sum_{i=1}^n (P'_i - \bar{P}')^2}},$$

式中: n 为样本个数; P_i 和 P'_i 分别为第 i 个样本的实测值和预报值; $\bar{P}$ 和 $\bar{P}'$ 分别是所有样本的平均实测值和平均预报值.

TS 评分可评价在一定预报时段内预报模式对预报对象超过某一临界值的预报效果, 是气象领域广泛使用的降雨预报评价方法, 其计算公式为^[6]:

$$TS_k = Na_k / (Na_k + Nb_k + Nc_k),$$

式中: k 为降雨等级; TS_k 为 k 等级降雨对应的 TS 评分; Na_k 为 k 等级降雨预报正确的次数; Nb_k 为 k 等级降雨空报(预报降雨等级大于实际降雨等级)的次数; Nc_k 为 k 等级降雨漏报(预报等级小于实际降雨等级)的次数.

2 SCE-UA 算法

Duan 等^[12]于 1965 年首先提出单纯多边形进化算法(SCE-UA), 该算法是一种基于生物竞争进化原理和基因算法等多种概念的优化算法, 可快速有效地搜索到参数的全局最优解. 该算法被认为是水文模型参数优选中最有效的方法, 在流域水文模型的参数率定中应用广泛.

SCE-UA 算法将参数寻优过程看作是一个自然生物不断竞争进化过程, 进化开始时, 算法在参数可行域内引入随机分布的点群, 并将点群划分为若干个复合形. 每个复合形根据下降单纯形算法进行进化, 定期将所有点群重新混合、分组, 产生全新的复合形. 这样不断重复进化和混合, 直到满足预先设定的收敛准则. SCE-UA 算法流程如图 2 所示.

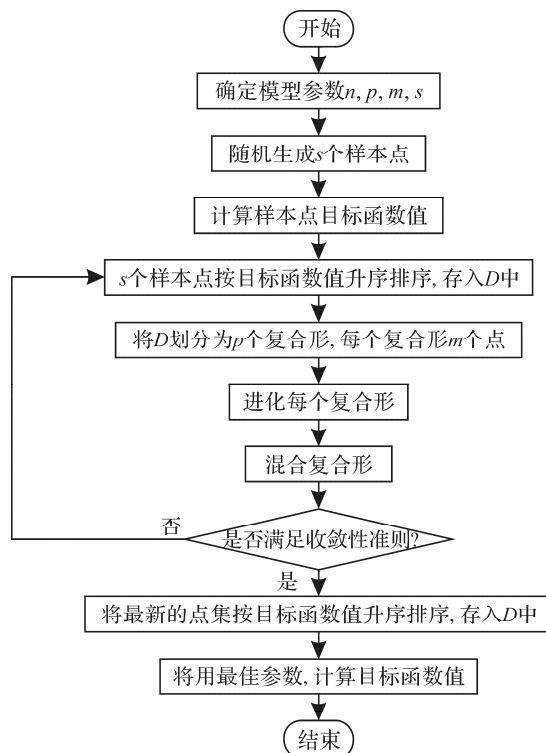


图 2 SCE-UA 算法流程

3 应用实例

实例主要考察宁波市未来 24 h 累积降雨的预报效果。根据历史预报和实测降雨,采用 SCE-UA 算法求解各单一预报模式的权重系数,构建多模式集成预报,并对预报结果进行对比分析。

3.1 资料说明与预处理

实例数据包括降雨数值预报数据和实测降雨数据。降雨数值预报数据取自中国气象局 GRAPES_MESO 模式(以下简称 GRAPES)、中国气象局 T639 模式(以下简称 T639)、美国国家环境预报中心 GFS(以下简称 GFS)和上海区域中尺度模式 SMS-WARMS 模式(以下简称 SMS)等 4 种模式,每日起报时间均取 8:00(北京时间),预报时效为 24 h。

实测降雨选择宁波市 63 个国家雨量站的逐时降雨数据,并转换为每日 8:00 至次日 8:00 累积 24 h 降雨量。若某个站点数据缺失,采用距离该站点最近的 5 个站点同时段数据的平均值代替。

2016 年 9 月的“莫兰蒂”和 2017 年 10 月的“卡努”两场台风期间,宁波市均发生较大范围的降雨,故实例以台风发生月为基准,前后各扩展 1 个月,即以 2016 年 8 月 1 日至 2016 年 10 月 31 日和 2017 年 9 月 1 日至 2017 年 11 月 30 日的降雨数据作为资料系列(简称降雨系列 1 和降雨系列 2)。考虑到各预报机构的预报能力在不同时间有所差别,较久远的历史资料对当前预报参考意义不大,故参照文献[13]中“滑动训练期”的概念,每日均取前 30 日有效数据作为训练样本,由此得到的权重系数随时间变化更符合实际。此外,由于各数值预报模式的原始预报数据均为网格数据,且分辨率不尽相同,为了便于集合预报模型的构建,需将网格预报数据通过空间插值方法转换为雨量站预报数据。

3.2 算法参数设置

根据文献[7],用经验公式确定算法参数 m 、 n 、 q 、 s 、 α 和 β ,即 $m=2n+1$, $q=n+1$, $s=pm$, $\alpha=1$, $\beta=2n+1$ 。参数 p 根据实际问题具体确定,经过多次试验,本文取 $p=2$ 。

3.3 目标函数

在天气预报中,一般认为预报值与实测值越接近预报效果越好,故选择预报降雨与实测降雨

的误差平方和作为目标函数,其公式为:

$$f = \sum_{i=1}^n (P_i - P'_i)^2,$$

式中: P_i 为实测降雨; P'_i 为预报降雨; n 为训练样本个数。

3.4 结果分析

为了检验 SCE-UA 算法在多模式集成预报中的使用效果,参考文献[5,14],分别选择遗传算法集成预报(简称 GA 集成)和基于 BP 人工神经网络的集成预报(简称 BP 集成)进行对比分析。分别从均方根误差、相关系数、TS 评分以及场次暴雨期间的应用效果等多方面进行分析,考察本文算法的适用性。

3.4.1 均方根误差与相关系数

表 1 和表 2 列出了各单一预报模式、SCE-UA 集成、GA 集成和 BP 集成等 7 种预报模式对不同降雨系列预报结果的均方根误差和相关系数。在对降雨系列 1 的 4 种单一预报模式中, GFS 模式的均方根误差最小,相关系数达到 0.95,为降雨系列 1 期间的最佳预报模式。在降雨系列 2 的 4 种单一预报模式中, T639 模式的均方根误差最小,相关系数与 GFS、GRAPES 模式相同,为降雨系列 2 的最佳预报模式。可见,对于同一降雨系列,各单一预报模式的预报效果不尽相同,即使同一预报模式,对于不同降雨场景的预报效果也有一定的差异。

表 1 降雨系列 1 各预报模式的均方根误差和相关系数

评价 指标	预报模式						
	模式 1	模式 2	模式 3	模式 4	模式 5	模式 6	模式 7
E_{RMS}/mm	3.53	4.86	3.98	4.47	3.13	3.28	3.49
r	0.95	0.93	0.94	0.94	0.97	0.95	0.96

注: 模式 1 为 GFS; 模式 2 为 GRAPES; 模式 3 为 T639; 模式 4 为 SMS; 模式 5 为 SCE-UA 集成; 模式 6 为 GA 集成; 模式 7 为 BP 集成。下表同。

表 2 降雨系列 2 各预报模式的均方根误差和相关系数

评价 指标	预报模式						
	模式 1	模式 2	模式 3	模式 4	模式 5	模式 6	模式 7
E_{RMS}/mm	4.45	3.77	3.52	6.84	2.98	3.15	3.07
r	0.95	0.95	0.95	0.86	0.96	0.96	0.95

综合表 1 和表 2 可知,无论均方根误差还是相关系数,集成预报模式的预报效果均优于单一预报模式,这与目前已有相关研究结果相一致^[3-6]。此外可以看出, SCE-UA 集成预报的均方根误差略

小于 GA 集成和 BP 集成预报, 相关系数略高于 GA 集成和 BP 集成预报, 表现出较好的预报效果。

3.4.2 TS 评分

TS 评分可定量衡量各预报模式对不同降雨等级的预报效果, TS 评分越高, 表示该等级的预报效果越好, 本次降雨等级划分参考文献[6]。图 3 和图 4 为不同降雨系列期间 7 种预报模式的 TS 评分结果。由图 3 和图 4 可知, 各预报模式对小雨的预报效果最好, 暴雨的预报效果相对较差, 个别预报模式暴雨的 TS 评分甚至为 0 (如降雨系列 2 的 GRAPES 和 T639 模式)。对不同降雨等级, 各预报模式的 TS 评分虽有差异, 但 3 种集成预报的预报效果均在一定程度上优于 4 种单一预报模式, 且 SCE-UA 集成预报的预报效果略优于 GA 集成和 BP 集成预报。

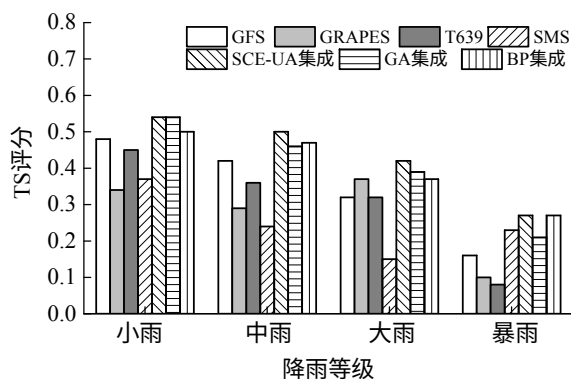


图 3 降雨系列 1 期间不同降雨等级条件下各预报模式的 TS 评分

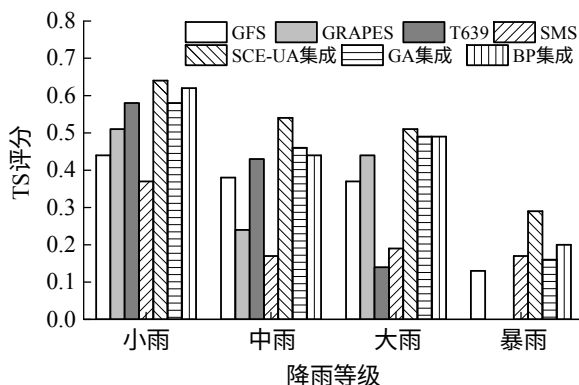


图 4 降雨系列 2 期间不同降雨等级条件下各预报模式的 TS 评分

3.4.3 典型暴雨期间预报效果

以上分析考虑了整体预报效果。现分析暴雨预报效果, 限于篇幅, 仅以 2017 年“卡努”台风期

间(取 2017 年 10 月 13 日 8 时至 2017 年 10 月 18 日 8 时)的暴雨过程为例, 考察各预报模式在典型暴雨期间的预报效果。2017 年“卡努”台风期间, 由于倒槽的影响, 宁波市发生较大范围的暴雨, 最大 3 日面雨量达 144 mm。图 5 给出了台风期间逐日实测暴雨和各预报模式预报降雨过程, 表 3 列出了各过程的误差分析结果。由图 5 可知, T639 和 3 种集成模式的预报过程与实测降雨过程整体趋势基本一致。从表 3 可知, 4 种预报模式的确定性系数均超过 0.95, 可见这 4 种预报模式均取得了较好的预报效果。进一步考察总降雨量误差、最大 1 日降雨和最大 3 日降雨相对误差, SCE-UA 集成预报的总降雨量误差和最大 3 日降雨误差均小于 T639、GA 集成和 BP 集成, 最大 1 日误差与 T639 和 GA 集成接近, 小于 BP 集成。综上可知, 对于本次暴雨过程, SCE-UA 算法集成预报模式总体预报效果最佳。

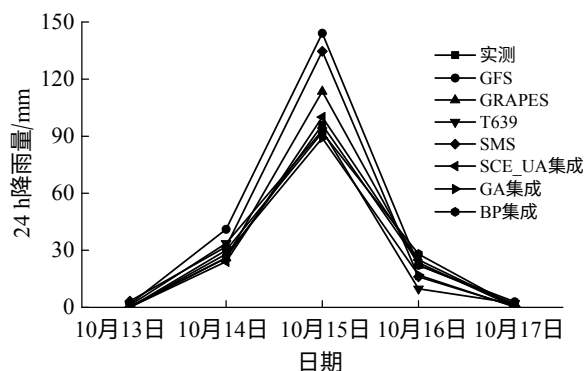


图 5 “卡努”台风期间逐日实测降雨与各预报模式逐日预报降雨过程线

表 3 “卡努”台风期间各预报模式降雨预报误差分析

评价指标	预报模式						
	模式 1	模式 2	模式 3	模式 4	模式 5	模式 6	模式 7
总降雨量	46.82	13.23	-3.30	29.11	3.26	-4.13	4.30
相对误差/%							
最大 1 日降雨	50.16	17.83	-3.02	39.73	3.28	-3.19	-4.07
相对误差/%							
最大 3 日降雨	43.75	11.81	-5.56	26.39	3.47	-3.94	4.17
相对误差/%							
确定性系数	0.59	0.95	0.96	0.74	0.98	0.98	0.96

4 结论

在分析多模式集成预报已有研究成果的基础上, 提出了采用 SCE-UA 算法优化权重分布法的权

重系数构建多模式降雨集成预报模型,并对宁波市2016年8月1日至2016年10月31日和2017年9月1日至2017年11月30日的降雨过程进行分析,通过均方根误差、相关系数和TS评分等多方面的指标比较,得出SCE-UA算法集成模型预报结果不仅好于各单一预报模式,同时也优于GA集成和BP集成预报模型,具有较好的预报效果。

参考文献:

- [1] 吴振玲, 潘璇, 董昊, 等. 天津市多模式气温集成预报方法[J]. 应用气象学报, 2014, 25(3):293-301.
- [2] 章祖迪, 陈淑琴, 何文岳, 等. 海上雷雨大风形成机制的分析[J]. 宁波大学学报(理工版), 2004, 17(2):134-137.
- [3] 董海萍, 张秀丽, 郭卫东. 多模式超级集合预报的应用研究[EB/OL]. [2018-01-18]. <http://xueshu.baidu.com>.
- [4] 赵成武. 短时降水的超级集合预报算法研究[D]. 长沙: 国防科学技术大学, 2014.
- [5] 李倩, 胡邦辉, 王学忠, 等. 基于BP人工神经网络的区域温度多模式集成预报试验[J]. 干旱气象, 2011, 29(2):231-235.
- [6] 熊聪聪, 耿世洁, 董昊. 基于混合算法的降水集成预报研究[J]. 计算机工程与科学, 2016, 38(10):2147-2152.
- [7] 李致家. 现代水文模拟与预报技术[M]. 南京: 河海大学出版社, 2010:258-262.
- [8] 董洁平, 李致家, 戴健男. 基于SCE-UA算法的新安江模型参数优化及应用[J]. 河海大学学报(自然科学版), 2012, 40(5):485-490.
- [9] 朱兵, 何晓燕, 王素芬, 等. 基于SCE-UA算法的新安江日模型参数优化稳定性分析[J]. 水电能源科学, 2016(11):9-12.
- [10] 邓元倩, 李致家, 刘甲奇, 等. 基于SCE-UA算法新安江模型在沔河流域的应用[J]. 水资源与水工程学报, 2017, 28(3):27-31.
- [11] 陈桂英. 权重分布法集成预报试验[J]. 应用气象学报, 2000, 11(S1):51-57.
- [12] Duan Q Y, Gupta V K, Sorooshian S. Shuffled complex evolution approach for effective and efficient global minimization[J]. Journal of Optimization Theory & Applications, 1993, 76(3):501-521.
- [13] 董海萍, 张秀丽, 郭卫东, 等. 2010年8月环流形势的多模式超级集合预报[J]. 应用气象学报, 2013, 24(5):606-616.
- [14] 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.

Multi-model integrated precipitation forecast based on SCE-UA algorithm

ZHONG Wei¹, ZHU Cong-fei¹, ZHANG Wei-guo¹, HUANG Feng²

(1. Ningbo Hydraulic Waterpower Planning Designing Research Institution, Ningbo 315192, China;

2. College of Hydrology and Water Resources, Hohai University, Nanjing 210098, China)

Abstract: Predicting with high-precision precipitation is a key for flood control decision-making. Based on the data collected from GRAPES_MESO and T639 of China Meteorological Administration, GFS of National Centers for Environmental Prediction and SMS-WARMS of Shanghai, SCE-UA algorithm is used to estimate the weighting of each forecast model. With the weighting estimated, the multi-model integrated precipitation is established. Using the data from Ningbo precipitation occurrences between August 1, 2016 and October 31, 2016, as well as from September 1, 2017 to November 30, 2017, the integrated forecast result is analyzed from a variety of indices such as root-mean-square error, correlation coefficient and TS grade. The result indicates that the integrated forecast based on SCE-UA is more accurate and applicable than not only those from individual forecast models but also those based on genetic algorithm and BP integrated forecast models.

Key words: SCE-UA algorithm; multi-model integrated forecast; precipitation forecast

(责任编辑 史小丽)