

北方季风边缘区洞穴石笋 $\delta^{18}\text{O}$ 序列预测初步研究

杜雅娟^{①*}, 冯添^②, 康志海^①^① 中国科学院地球环境研究所, 黄土与第四纪地质国家重点实验室, 西安 710075;^② 西安交通大学人居环境与建筑工程学院, 西安 710061

* 联系人, E-mail: duyj@ieecas.cn

2012-05-07 收稿, 2012-08-15 接受

国家重点基础研究发展计划(2010CB833405)和国家自然科学基金(41023006)资助

摘要 基于时域组合模型建立了中国北方季风边缘区万象洞及黄谷洞石笋 $\delta^{18}\text{O}$ 时间序列的动态模型, 并对未来 20 年的降水变化趋势进行了预测. 研究表明, 时域组合模型较好地提取了时间序列的周期信息, 用其拟合实测数据精度较高; 用其预测时间序列的变化情况, 和现代气象观测记录相吻合, 具有一定的可信度. 同时通过中国北方季风边缘区树轮的预测结果对比显示, 该地区在 2012~2013 年前后降水最少, 随后 20 年里将呈现先升后降的变化趋势.

关键词

季风边缘区
时域组合模型
石笋 $\delta^{18}\text{O}$ 时间
序列预测

中国季风边缘区对亚洲季风的强弱变化非常敏感, 降水的年至年代际变率比东亚季风区更大^[1]. 随着人类活动的不断加剧, 季风边缘区域水资源日益匮乏, 因此利用高分辨率气候代用指标探索该区域降水的特征和规律, 研究过去预测未来尤为重要^[1]. 毫无疑问, 树轮、石笋等高分辨率的古气候替代性指标为我们研究长期气候变化提供了宝贵的数据资料. 运用这些替代性指标, 从分析时间序列的历史变化规律角度来预测未来, 是一项更为有效的探索性方法. 古气候替代性指标时间序列既储存了长期变化的信息, 同时由于受各种自然因素及采样过程等人为因素的影响又具有随机性变化. 基于序列的这种特征, 时域组合模型将其拆分为反映长期变化的周期项以及反映随机变化的随机项, 建立具有明确表达式并和时间相关的动态模型^[2], 进而达到预测的目的. 这种基于序列特征和数学理论基础的预测具有一定的可靠性和实际意义.

本文选用位于北方季风边缘区的甘肃武都万象洞的高分辨率石笋 $\delta^{18}\text{O}$ 序列作为主要研究对象. 石笋 WX42B 的 $\delta^{18}\text{O}$ 与当地器测降水资料具有良好的相关性, 其变化主要受西南(印度)季风带来的降水量效应所控制, 受温度的影响比较弱, 因此石笋的 $\delta^{18}\text{O}$

的轻重变化主要反映了季风降水 $\delta^{18}\text{O}$ 的信息, 其气候意义为数值越小表示降水越多^[3,4]. 我们将其作为该地区古降水的替代性指标曲线, 借助时域组合模型和奇异谱分析的方法, 并结合同样位于北方季风边缘区的黄谷洞石笋、内蒙古贺兰山及白音敖包树轮的预测趋势对比, 探讨了该地区未来的降水趋势.

1 甘肃万象洞石笋 $\delta^{18}\text{O}$ 时间序列的时域组合模型

时域组合模型的基本思路是将非平稳时间序列看成周期项和随机项的组合, 已有研究者应用该模型基于地质记录或现代环境开展了气候变化或环境监测的预测试验^[2,5-7]. 其中周期项为正余弦波的累加组合, 反映了序列的长期变化(包括趋势和周期性); 随机项为原始信号减去周期信号的残差, 通常为平稳时间序列, 反映了序列中由随机因素引起的变化, 数学中常用 ARMA 模型来进行模拟. 建模的思路为先将整个序列的变化分解为周期项和随机项的具体数学表达式, 预测时周期项的部分直接外延, 随机项则根据 ARMA 模型的相应参数来进行计算, 最后将这两部分进行叠加组合, 具体计算过程详见文

引用格式: 杜雅娟, 冯添, 康志海. 北方季风边缘区洞穴石笋 $\delta^{18}\text{O}$ 序列预测初步研究. 科学通报, 2013, 58: 254-259

Du Y J, Feng T, Kang Z H. Preliminary study on the stalagmite $\delta^{18}\text{O}$ record prediction of the monsoon boundary zone in the northern China (in Chinese). Chin Sci Bull (Chin Ver), 2013, 58: 254-259, doi: 10.1360/972012-437

献[2,5~8]. 因此, 本文首先应用时域组合模型将石笋氧同位素的时间序列分解为周期项和随机项.

1.1 周期项

现取甘肃万象洞石笋氧同位素的变化时间序列(经 2 年等间隔插值共 905 个数据)进行时域组合模型预测. 取其前 855 个数据建立模型, 并用其预测后 50 个数据作为检验. 首先, 取 $M = (N-1)/2 = 427$ 计算各谐波分量, 做累积函数 $P(i)-i$ 图, 如图 1 所示. 根据已有研究^[2,5], 显著谐波个数应由 $P(i)-i$ 曲线上的拐点来确定, 但从直观上很难判断像图 1 这样的单调增曲线何处为拐点, 其人为主观性太大. 因此, 这里通过奇异谱(SSA)分析计算典型波形的累积方差贡献量, 以此来确定显著谐波的个数.

1.2 奇异谱分析及显著谐波个数的确定

奇异谱(SSA)方法通过分解序列中的时间主分量, 可获得不同层次上的分量序列, 所分解出的低频信号可以视为序列变化的长期趋势, 逼近程度高^[9]. 它可以简单概括为嵌套、奇异值分解以及重构 3 个步骤^[10], 具体的分析方法详见文献[11]. 原始数据选取 192~2003 年共 905 个数据标准化后进行 SSA 分析. 由于嵌套窗口长度 L 不应超过 $N/3$ ^[12], 同时根据万象石笋洞 170 a 的显著周期^[3], 选取 $L=170$. 经过计算, 序列的前 27 个典型波形向量已占总方差贡献的 85%, 它们表征了石笋时间序列 $\delta^{18}\text{O}$ 变化的主要振荡模态, 其中第一典型波形向量所占的方差贡献为 42.63%, 它表征的是序列的最主要趋势. 因此, 在选择时域组合模型的显著谐波时, 就可以根据 $P(i) \approx 85\%$

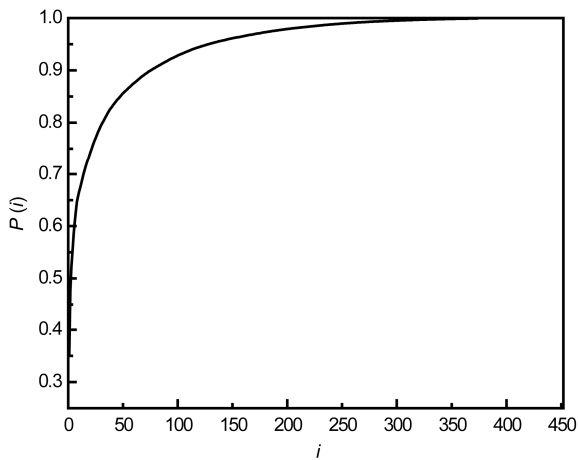


图 1 累积函数 $P(i)-i$ 图

来确定谐波个数的范围为 45~50 个.

1.3 残差项 ARMA(p, q)模型

原始数据序列提取周期项 P_t 后得到残差序列, $x_t = y_t - P_t$, 首先判断残差序列是否为平稳序列, 采用 Daniel 检验方法来判断平稳性^[13]: 对于显著水平 α , 由时间序列 $x_t (t = 1, 2, \dots, n)$ 计算 Spearman 秩相关系数 q_s 及相应统计量 T , 若 $|T| > t_{\alpha/2}(n-2)$, 则拒绝 H_0 , 认为序列非平稳, 且当 $q_s > 0$ 时, 认为序列有上升趋势; $q_s < 0$ 时, 认为序列有下降趋势. 仅当 $|T| \leq t_{\alpha/2}(n-2)$ 时, 接受 H_0 , 可以认为 x_t 是平稳序列. 经计算可知, $T = 0.20078$, 显著性水平 $\alpha = 0.05$ (显著性水平 95%) 的上 $\alpha/2$ 分位点的值为 1.9627, 故根据 Daniel 检验法, 该残差序列为平稳序列. 因此, 可以用 ARMA(p, q) 模型来进行拟合. 但由于 ARMA(p, q) 中的 p, q 较难通过自相关或者偏相关图来确定, 而是需要先假定一组值, 例如 (0, 1) 开始, 建立模型, 然后逐步升高 p, q 的值, 求出一系列模型, 并根据 loss function, AIC 等准则, 找出一个最优模型. 因此从 (1, 0) 开始, 采用“爬坡法”取 $p = 2, 3, \dots$, 由 AIC 准则^[14,15] 得到模型的阶数及相应的参数估计值.

1.4 时域组合模型的验证

根据前 855 个数据作为建模集对原始序列进行拟合, 并预测后 50 个数据作为校验集(检验校准期 1904~2003 年). 经过计算发现显著谐波个数在 45~50 范围内拟合效果相差很小, 从中选择拟合度最优的 $M_0 = 45$ 及 ARMA(3, 4). 统计检验结果见表 1 和图 2, 检验的主要统计量包括 r, R^2 , 标准误差(SE)和 P 值. 建模期模拟序列和原始序列相关性很好达到 0.93, 校准期模拟序列与原始序列、SSA 重构序列的相关系数分别为 0.5 和 0.71, 表明模型具有一定的可靠性. 同时, 也可以得到如下结论:

(1) 时域组合模型通过周期项和随机组合项的组合, 对于原始数据序列的拟合可以达到很好的效果, 相关性 $R = 0.93 (P = 0.0001)$.

(2) 在组合模型的建立过程中, 显著谐波的个数以及 ARMA(p, q) 模型阶数的识别, 均带有一定的人为主观性. 其中显著谐波的个数, 本文中提出通过 SSA 方差解释量来进行确定; ARMA(p, q) 模型阶数则可以依据 AIC 准则从中选择近似最优模型的方法来进行弥补^[2]. 但这种参数的识别方法仅仅是从数学

表1 时域组合模型统计检验结果

| 统计量 | 建模期(192~1902年) | | 校准期(1904~2003年) | |
|-----------|----------------|--|-----------------|----------|
| | 与原始序列 | | 与原始序列 | 与SSA重构序列 |
| r | 0.93 | | 0.5 | 0.71 |
| R^2 | 0.87 | | 0.25 | 0.51 |
| 标准误差 SE | 0.074 | | 0.13 | 0.055 |
| P | 0.0001 | | 0.0003 | <0.0001 |

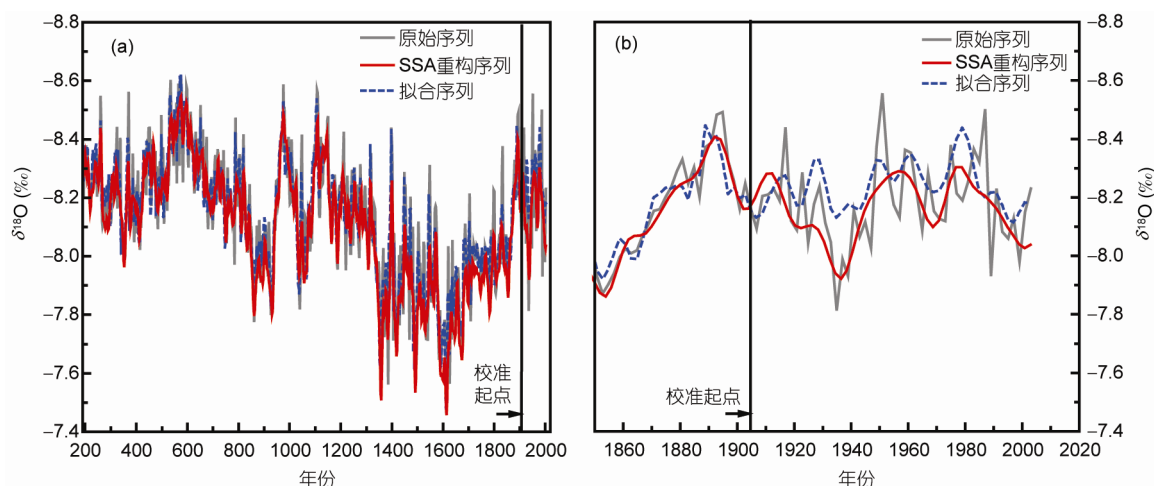


图2 万象洞石笋1800年以来原始序列与SSA重构序列、模拟序列的对比(检验校准期1904~2003年)
 (a) 1800年以来原始序列与SSA重构、时域组合模型模拟序列对比; (b) 1850~2003年局部放大图

的角度出发, 至于是否能真正符合实际情况无法确定。

(3) SSA 重构的 $\delta^{18}\text{O}$ 序列很好地逼近了原始序列的变化振荡趋势, 是研究非平稳时间序列的一个有力工具. 由原始序列构造嵌入空间并求出相关矩阵后, 可以通过特征值得到嵌入空间中数据点的主因子方向, 用这些主因子(特征矢量)作为一个单边(向前)滤波器, 能有效地将原时间序列分解成各个主成分^[16]. 本文中 SSA 为选择显著谐波个数提供了定量的依据, 降低了人为主观性, 同时其重构序列也为校验区间的验证提供了依据。

(4) 校准期内的统计检验结果显示, 较之原始序列, SSA 重构序列和模拟序列的相关性更高, 这是由于时域组合模型中周期项的解释占比很高, 校准期内的模拟序列主要反映周期项, 而 SSA 恰恰也是典型波形的叠加。

2 组合模型对未来降水的预测

由于万象洞石笋的 $\delta^{18}\text{O}$ 的轻重变化主要反映了

季风降水 $\delta^{18}\text{O}$ 的信息^[3,4], 我们可以进一步运用时域组合模型对石笋的 $\delta^{18}\text{O}$ 时间序列继续向未来预测, 并将结果和现代降水资料及同样位于北方季风边缘区的树轮预测结果进行对比. 现代降水资料选用天水站点(34°36'N, 105°41'E)的观测记录(1951~2010)(国家气候中心, <http://ncc.cma.gov.cn/cn/>), 树轮采用刘禹等人^[10]通过毛毛虫-奇异谱分析法对贺兰山北部和内蒙古白音敖包地区的降水预测, 比较结果如图3所示. 从图中可以看出, 在1994~2010期间树轮和石笋的预测均表现出先升后降的趋势, 降水在2003年附近达到峰值, 随后呈现出10年降水减少期, 并在2012~2013年前后降水最少, 较好地吻合了现代观测记录; 之后, 树轮和石笋的预测趋势均表明降水开始增加. 在2023年附近达到峰值后, 石笋指示降水又开始下降. 而树轮未来的预测趋势如何变化, 还有待于该指标的进一步研究. 总之, 未来20年的降水变化总体呈现先升后降的趋势。

同时为避免“一孔之见”, 对该地区的甘肃黄谷洞(33°35'N, 105°07'E, 1650 m a.s.l.)石笋的氧同

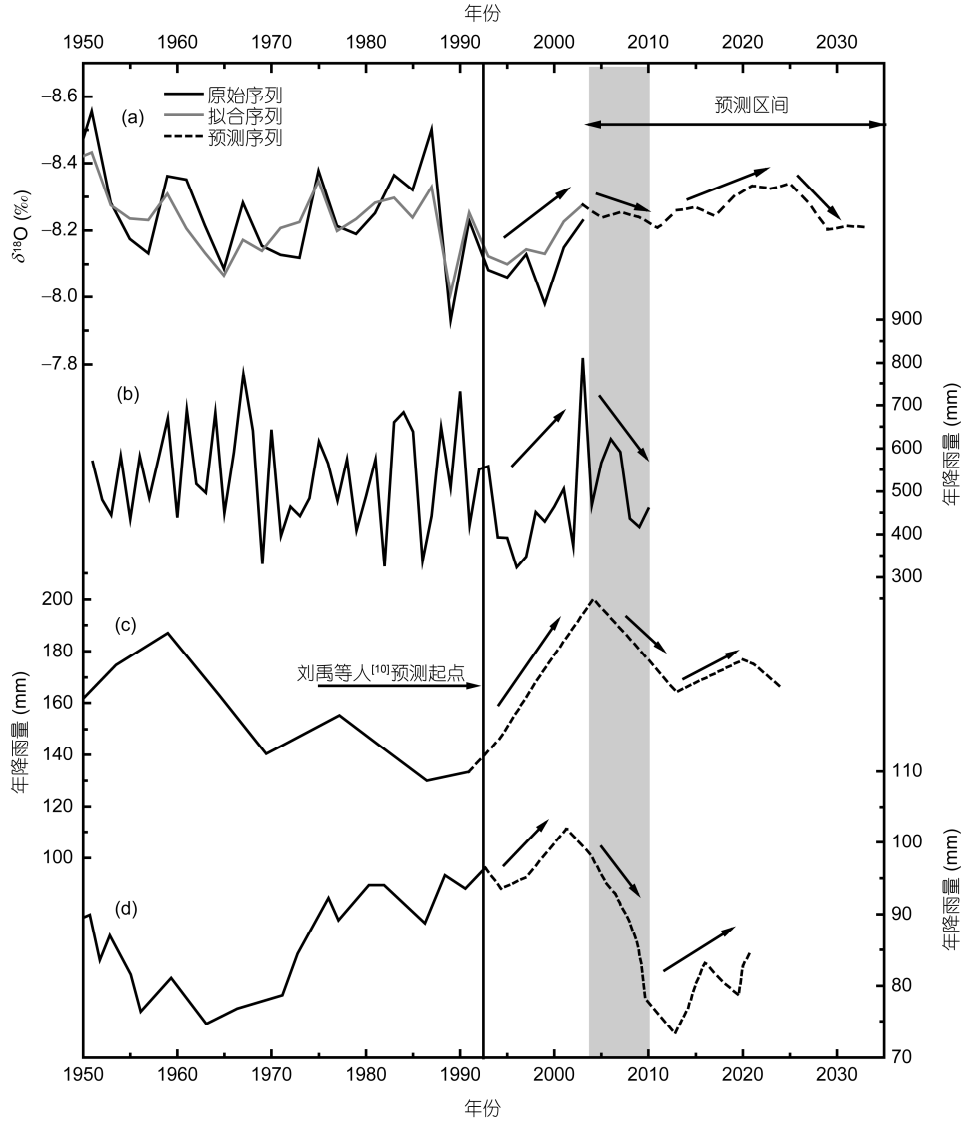


图3 万象洞石笋 $\delta^{18}\text{O}$ 预测趋势和现代降水记录、树轮预测结果对比

(a) 万象洞石笋原始序列与时域组合模型模拟、预测序列(1950~2032年); (b) 甘肃天水 1951~2010年年降水记录; (c) 白音敖包 4~7月上旬降水量(1950~2024年)^[10]; (d) 贺兰山 5~7月的降水量(1950~2022年)^[10]

位素序列也进行了预测分析. 这两个石笋洞在地理位置上非常接近, 并且黄谷洞的 $\delta^{18}\text{O}$ 也主要反映了季风降水的变化^[17], 因此对于该区域的降水趋势预测比对很有意义. 周期项和残差项 ARMA(p, q)模型的选取方法和万象洞的类似. 根据计算分析, 选取 50 个显著谐波(解释量为 85.3%), 残差项选用 ARMA(2,3)模型, 拟合和预测结果如图 4 所示, 其中模拟序列和原始序列的相关性 $r = 0.94$ ($R^2 = 0.877$, $SE = 0.08$, $P < 0.0001$). 黄谷洞的未来降水量在 2020 年附近达到一个峰值然后开始下降至 2034 年左右,

未来 20 年也呈现先升后降的总体趋势, 但并未显示出和万象洞类似的以 2003 年为起点的 10 年降水减少期, 这可能是由于黄谷洞的分辨率较低不能显示出某些细节变化. 但石笋预测(2011~2032 年)和已有的树轮预测(1994~2024 年)在 2011~2024 年期间呈现出的大趋势一致性, 表明这种趋势有可能代表该地区未来的气候变化.

3 结论

由于石笋 $\delta^{18}\text{O}$ 时间序列的内在机理影响因素多,

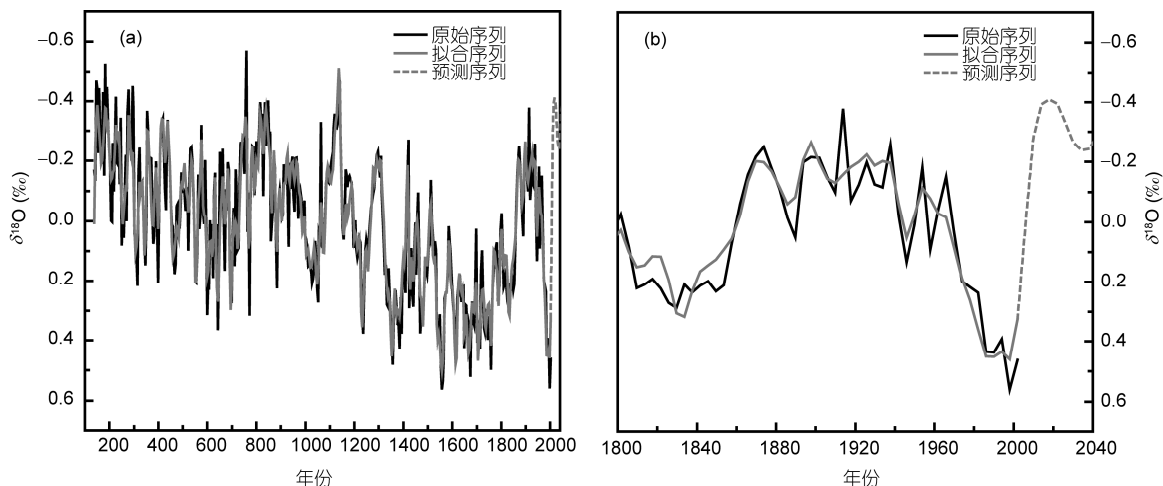


图4 黄爷洞石笋 $\delta^{18}\text{O}$ 原始序列、拟合序列及预测序列

(a) 1800年以来原始序列与时域组合模型模拟序列、预测序列; (b) 1800~2040年局部放大图

因此数据分析目前尚有一定的困难. 本文将万象石笋 WX42B 作为季风降水的替代指标, 通过建立时域组合模型在一定程度上给未来气候变化趋势的预测提供了依据. 通过建模和预测的整个过程, 可以得到以下认识:

(1) 时域组合模型在建立的过程中, 其模型参数的辨识仍带有一定的主观性. 文中通过对校准期内预测结果和实际序列的统计检验, 在一定程度上保证了模型的有效性.

(2) 运用时域组合模型外推得到的预测序列反

映出时间序列变化的趋势, 目前还无法从理论或动力因素及机理上对预测模型的结果给予分析和解释. 但是从万象洞和黄爷洞在未来 20 年所表现出的先升后降的趋势一致性, 以及和现代降水资料、树轮预测结果的对比表明, 该结果具有一定的可信度.

(3) 然而, 由于数学模型本身的限制, 也可能会导致预测的适用性和效果存在差异, 本文也仅仅是一个初步的尝试. 因此古气候替代性指标的预测应集多种方法之长, 尝试构建不同序列的预测方法体系, 并在此基础上进行综合对比, 实现进一步验证.

致谢 兰州大学张平中教授对本文提出了中肯的建议; 万象洞、树轮及黄爷洞数据分别引自张平中教授、刘禹教授和谭亮成博士已发表论文, 在此一并深表感谢.

参考文献

- 1 刘敬华, 张平中, 孟彩红, 等. 季风区边缘近 500 年的降水变化特征. 地理科学, 2011, 31: 401-407
- 2 杨永国, 余志伟, 郭正堂, 等. 基于地质记录用时域组合模型预测气候变化趋势的初步研究. 地球物理学报, 1996, 39: 38-45
- 3 Zhang P Z, Cheng H, Edwards R L, et al. A test of climate, sun, and culture relationships from an 1810-year Chinese cave record. Science, 2008, 322: 940-942
- 4 刘敬华, 张平中, 程海, 等. 黄土高原西缘在 AD 1875-2003 期间石笋氧同位素记录的季风降水变化与海气系统的联系. 科学通报, 2008, 53: 2801-2808
- 5 杨永国, 冉隆振. 时间序列叠合模型及其在矿井涌水量预测中的应用. 大学数学, 1989, 4: 18-21
- 6 钱家忠, 朱学愚. 地下水资源评价中降水量的时间序列——马尔可夫模型. 地理科学, 2001, 21: 350-353
- 7 吴晓娟, 孙根年, 冯煜. 时域组合模型与西安市城区地下水污染预测. 干旱区地理, 2007, 30: 939-945
- 8 杨位钦, 顾岚. 时间序列分析与动态数据建模. 北京: 北京理工大学出版社, 1988. 324-375
- 9 俞肇元, 袁林旺, 谢志仁, 等. 基于 SSA 和 AR 模型的海面变化预测试验. 海洋湖沼通报, 2008, 4: 14-20
- 10 刘禹, Shishov V, 史江峰, 等. 内蒙古西部贺兰山和东部白音敖包未来 20 年降水趋势预测. 科学通报, 2004, 49: 270-274

- 11 Ghil M, Allen M, Dettinger M, et al. Advanced spectral methods for climatic time series. *Rev Geophys*, 2002, 40: 1003
- 12 汪芸, 郭生练, 李响. 奇异谱分析在中长期径流预测中的应用研究. *人民长江*, 2011, 42: 4-7
- 13 王莉, 胥鹏海, 宋涛, 等. 沱河水质变化趋势及原因分析. *西北大学学报*, 2011, 41: 503-506
- 14 王丽娜, 肖冬荣. 基于 ARMA 模型的经济非平稳时间序列的预测分析. *武汉理工大学学报*, 2004, 28: 133-136
- 15 宋贵彩. 基于 ARMA 模型的结构识别. 硕士学位论文. 重庆: 重庆大学, 2005. 15-16
- 16 余志伟, 丁仲礼, 刘东生. 第四纪古气候变化非线性动力学初步研究——黄土粒度曲线的奇异谱分析. *第四纪研究*, 1993, 3: 214-230
- 17 Tan L C, Cai Y J, An S Z, et al. Centennial- to decadal-scale monsoon precipitation variability in the semi-humid region, northern China during the last 1860 years: Records from stalagmites in Huangye Cave. *Holocene*, 2011, 21: 287

Preliminary study on the stalagmite $\delta^{18}\text{O}$ record prediction of the monsoon boundary zone in the northern China

DU YaJuan¹, FENG Tian² & KANG ZhiHai¹

¹State Key Laboratory of Loess and Quaternary Geology, Institute of Earth Environment, CAS, Xi'an 710075, China;

²Department of Environmental Science and Technology, School of Human Settlements and Civil Engineering, Xi'an Jiaotong University, Xi'an 710061, China

Based on high-resolution oxygen isotope records of stalagmites collected from Wanxiang and Huangye Caves, a prediction experiment for precipitation over the next 20 years is performed, using the time domain combined model. The study indicates that this model can be used to determine the significant period of the time series and fit the original record well. The prediction accords with modern meteorological observation, and it is thus credible to a certain extent. Compared with tree-ring prediction of the monsoon boundary zone in northern China, the result indicates that precipitation may reach a minimum during 2012-2013, and first increases and later decreases during the subsequent 20 years.

monsoon boundary zone, time domain combined model, stalagmite $\delta^{18}\text{O}$, time series prediction

doi: 10.1360/972012-437