

广义回归神经网络在日长变化 预报中的应用*

张晓红 王琪洁[†] 朱建军 张 昊

(中南大学地球科学与信息物理学院 长沙 410083)

摘要 传统的日长变化预报多是基于线性模型, 如最小二乘模型、自回归模型等, 但是日长变化包含了复杂的非线性因素, 线性模型预报的效果往往不甚理想. 所以尝试使用一种非线性神经网络——广义回归神经网络 (GRNN) 模型进行日长变化预报, 并将结果与使用 BP (Back Propagation) 神经网络模型和其它模型的预报结果进行比较. 结果表明, GRNN 用于日长变化预报是高效可行的.

关键词 天体测量, 时间, 方法: 其它诸多方面

中图分类号: P 128; **文献标识码:** A

1 引言

地球自转参数 (ERP) 包括地极移动和日长 (LOD) 变化. 天文学和测地学的许多重要任务要求地球参考系和天球参考系进行精确转换, 这就要求能够及时获得地球自转参数^[1]. 然而由于现代测地技术, 如甚长基线干涉测量技术、人工激光测距技术和全球定位技术等复杂的数据处理过程, 使得难于实时获得地球自转参数, 因此, 只能通过地球自转参数的精确预报以满足这些研究和应用部门的需要.

地球自转参数预报中, 日长变化的预报是难点之一, 尤其是在厄尔尼诺现象发生期间, 热带季风的变化, 会使日长变化信号出现大幅振荡. 高精度和实时、快速预报日长变化引起了越来越多学者的重视. 日长变化主要包括潮汐项和非潮汐项两部分, 潮汐项可以通过建模精确确定; 而非潮汐项中的亚季节性、季节性以及年际变化主要是由固体地球与全球大气和海洋之间的角动量交换引起的^[2-5].

学者们在预报 ERP 方面已经做了很多研究, 提出了各种不同的方法. 这些方法大体上分为线性模型和非线性模型两类. 线性模型如: 卡尔曼滤波^[6]、最小二乘外推、自回归模型^[7]、自协方差模型以及这些方法的不同组合等. 非线性模型包括门限自回归模型、人工神经网络^[8]、模糊推理等方法. 为了得到更精确的地球自转参数以满足实时快速预报的需求, 2005 年 Schuh 等人发起了地球自转参数预测比较计划 (EOP PCC) (起

2011-01-17 收到原稿, 2011-03-01 收到修改稿

* 国家自然科学基金委员会与中国科学院天文联合基金项目 (10878026) 资助

[†] qjwang.cn@163.com

止时间为 2005 年 10 月 1 日至 2008 年 2 月 28 日), 分别对 ERP 进行了超短期、短期和中期预报¹.

地球自转参数包含了复杂的非线性因素, 所以采用非线性预报的方法对其进行预报在理论上更为合理. Kolmogorov 定理 (1957 年) 证明: 一个 3 层前馈神经网络可以以任意精度逼近任意连续函数. 神经网络不需要事先建立数学模型, 这也就直接避免了模型误差. 对于复杂非线性系统的实现是十分可行的.

BP 神经网络是目前使用最多的一种神经网络模型, 它被广泛应用于自动控制、信号处理、时序预测及拟合等领域. 但是它存在一些缺点, 例如迭代过程容易陷入局部极小、网络拓扑结构难以确定、运算时间较长、泛化能力不高等^[9-12]. 针对这些缺点很多学者提出了改进措施, 有直接对神经网络的参数及拓扑结构进行改进的, 也有将神经网络和其他智能算法结合进行计算的, 如模糊神经网络、遗传神经网络, 但是都存在着一定的缺陷, 仍处于不断的尝试研究阶段.

GRNN 是径向基函数 (RBF) 神经网络的一种, 是在 BP 算法之后提出的一种新的函数逼近神经网络. 它的优势体现在: 具有良好的局部逼近性能, 不会陷入局部极小值; 训练过程不需要迭代, 需要调节的参数只有 1 个, 即光滑因子, 所以计算速度快; 所需的训练样本较 BP 神经网络少得多, 网络只有 2 层, 结构简单, 擅长解决逼近、预测、分类等问题^[13-14]. 越来越多的学者将其应用于实践研究并提出了改进方法, 所以理论上广义回归神经网络用于解决日长变化预报是可行的.

2 GRNN 的基本原理及其用于日长变化预报的方法

2.1 GRNN 的基本原理

GRNN 神经网络是基于普通 RBF 神经网络的一种改进的 RBF 神经网络模型, 由 2 层组成, 即隐含层和输出层. 它与普通 RBF 神经网络的区别在于输出层的数据处理方式不同.

网络的第 1 层为径向基隐含层, 神经元个数等于训练样本数, 该层的权值函数为欧氏距离函数 (用 $\|distant\|$ 表示), 其作用为计算网络输入与第 1 层的权值 $LW_{1,1}$ 之间的距离. 隐含层的激励函数为:

$$R_i(x) = e^{-\frac{\|x-c\|^2}{2\sigma_i^2}}, \quad (1)$$

其中 $R_i(x)$ 表示隐含层第 i 个节点的激励函数, c 表示激励函数中心, σ_i 为光滑因子, σ_i 越大则激励函数越平滑.

网络的第 2 层为线性输出层, 其权函数为规范化点积函数. 计算网络的向量 n^2 , 它的每个元素就是向量 a^1 和权值矩阵 $LW_{1,1}$ 每行元素的点积再除以向量 a^1 的各元素之和得到的, 并将结果 n^2 提供给线性传递函数 $a^2 = \text{purelin}(n^2)$ 计算网络输出^[15-17].

¹http://www.cbk.waw.pl/EOP_PCC

2.2 GRNN 用于日长变化预报

2.2.1 数据预处理

本文所用的日长变化资料取自国际地球自转和参考系服务的 EOP C04 序列, 时间跨度为 1980~2010 年, 每日一个值. 日长变化序列中的潮汐项, 包括固体地球潮汐项和海洋潮汐项, 其中前者为 5 d~18.6 yr 周期的固体地球带谐潮项, 已经有国际公认的经验公式和数据可以进行修正^[18-20], 后者为海洋近周日、半日潮项^[21-22], 对于本文的研究可以忽略; 日长季节性变化的周年、半年项可以通过最小二乘方法精确确定, 剩下的部分为带有非线性成份的残差序列. 图 1 从上到下分别绘出了 1980~2010 年期间的日长变化的原始序列、带谐潮序列、线性模型拟合序列和残差序列.

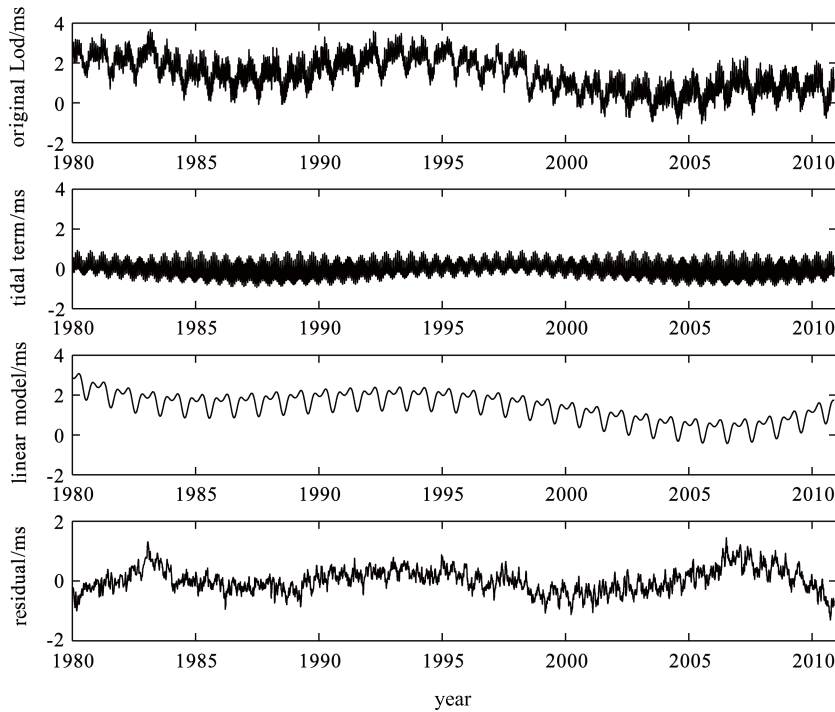


图 1 由上到下依次为: 日长变化原始数据、带谐潮项、线性拟合项、残差序列

Fig. 1 From top to bottom: Series of the original data of LOD (length of day) change, tidal term, fitted term of linear model and residual

本文采用上述残差序列作非线性神经网络的建模和预报, 将线性模型的预报结果和非线性模型预报结果相加即为最终的日长变化预报值.

在神经网络建模和预报之前需将残差序列分为训练段、测试段和验证段, 每个数据段都要根据神经网络建模和训练函数的要求构建输入矩阵和相应的目标矩阵 (即期望的输出矩阵), 最后为了减小数据处理误差, 还要对这些输入矩阵和目标矩阵进行归一化处理. 在建模和预报之后, 对仿真和预报结果必须作相应的逆归一化处理, 才能将这些数据用于模型性能评估、预报与观测的比较和精度评定.

2.2.2 建模和预报

以跨度为 1 d 的日长变化预报为例: 输入样本矢量的个数由 2 取到 10, 隐含层节点个数和训练样本个数相等, 输出层节点个数根据需要设定, 它等于目标矩阵中输出矢量的个数. 每次计算需要调节的是光滑因子, GRNN 通常用均方根误差 (RMSE) 作为衡量网络性能好坏的指标. 使用时, 光滑因子值需要不断尝试, 以实验经验来看, 预报跨度较小、输入节点个数较少时, 光滑因子取较小值可使 RMSE 最小. 设网络第 i 个节点实际输出为 t_i , 目标输出为 e_i , 则均方根误差的数学表达式为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (t_i - e_i)^2}, \quad (2)$$

N 表示输出节点的个数.

选取光滑因子时可以先在较大的范围内选择, 如先在 (0.05, 1.0) 的范围内选取, 步长取 0.05, 这样需要计算 20 次; 当计算结果表明光滑因子在一个较小区域, 例如在 (0.20, 0.35) 的范围内 RMSE 较小时, 下一次计算就可将光滑因子的取值范围缩小到 (0.20, 0.35), 步长也可适当缩减为 0.01, 那么再计算 15 次就可以找到适合的光滑因子值. 实际应用中, 对于特定的数据, 光滑因子值的调节过程有一定的规律性. 对于日长变化预报而言, 在计算完跨度为 1 d、2 d 的预报后, 后面的预报在不同输入节点的情况下, 可根据前面相应输入节点条件下确定的光滑因子值, 基本确定一个小的范围, 很快求出使 RMSE 最小时的光滑因子值.

为了更好地衡量预报结果的精度, 并同 EOP PCC 预报的结果进行比较, 我们也计算了平均绝对误差 (MAE) 值. 其数学计算公式如下:

$$MAE = \frac{1}{N} \sum_{i=1}^N (t_i - e_i), \quad (3)$$

式中 t_i 、 e_i 及 N 的意义同 (2) 式.

3 基于 GRNN 的日长变化预报

本文首先将 1980 年 1 月 1 日至 1999 年 12 月 31 日的日长变化残差序列用于广义回归神经网络的训练. 训练过程中, 训练段、验证段和测试段分别按 80%、10%、10% 的比例进行分配. 然后用训练好的网络对 2000 ~ 2001 年的日长变化残差序列进行 1 ~ 10 d、15 d、20 d、25 d、30 d、60 d、90 d、...、360 d 的预报 (与 Schuh 等^[2]的时间段相同), 并将预报结果与 Schuh 等^[2]预报结果进行比较, 见表 1. 由于 Schuh 等^[2]的预报是 10 yr 以前的预报, 这期间 BP 神经网络也得以不断发展进步, 所以在使用 GRNN 预报前我们也尝试使用改进的 BP 神经网络方法对日长变化进行预报, 并将此方法预报所得均方根误差值一并列于表 1.

同时我们给出了基于 GRNN 方法的跨度为 1 d 的日长变化残差序列预报拟合曲线图及预报残差图, 见图 2. 图 2 中横坐标代表预报结果所对应的日期, 将 1 yr 分为 365 d,

则每一天对应的日期可以写成年号加上 $1/365$. 由于横坐标的范围有限, 这里只保留到了小数点后一位. 图 2(a) 中实线代表日长变化序列原残差值, 图中用 lodr 表示, 虚线为预报残差值, 图中用 prediction 表示. 图 2(b) 中给出了预报的误差.

表 1 GRNN 预报结果与 BP 神经网络预报结果 (Schuh 等^[2] 的预报结果) 及改进的 BP 神经网络预报结果的比较 (单位: ms)

Table 1 Comparison of the prediction results of GRNN, BP (Schuh et al.^[2]) and modified BP neural networks (unit: ms)

Days ahead	GRNN	Schuh et al. ^[2]	Modified BP
1	0.037	0.019	0.027
2	0.074	0.049	0.073
3	0.097	0.074	0.093
4	0.117	0.097	0.110
5	0.134	0.121	0.131
6	0.151	0.142	0.148
7	0.164	0.159	0.162
8	0.174	0.174	0.170
9	0.179	0.184	0.176
10	0.187	0.193	0.185
15	0.204	0.246	0.211
20	0.210	0.251	0.217
25	0.211	0.249	0.215
30	0.217	0.245	0.219
60	0.222	0.292	0.219
90	0.226	0.306	0.231
120	0.226	0.314	0.229
150	0.233	0.330	0.237
180	0.234	0.361	0.234
210	0.236	0.397	0.241
240	0.236	0.377	0.236
270	0.240	0.386	0.231
300	0.247	0.402	0.249
330	0.254	0.372	0.262
360	0.250	0.347	0.245

从表 1 和图 2 拟合误差图 (a) 及 (b) 中可以看出, GRNN 用于日长变化预报是比较成功的, 随着跨度的增大, 预报精度有所降低.

为了使结果更加直观, 图 3 绘出了不同跨度的预报精度, 十字符号为 GRNN 预报结果, 圆圈为 Schuh 等^[2] 的预报结果, 五角星代表改进的 BP 神经网络预报结果.

从图 3 可以看出, 对于 $1 \sim 8$ d 的预报, GRNN 的预报精度低于 Schuh 等^[2] 的预报精度, 从第 9 d 开始, GRNN 的预报精度优于 Schuh 等^[2] 的预报精度并逐渐拉开距离. 在超短期 $1 \sim 10$ d 的预报中, GRNN 预报精度低于改进的 BP 神经网络的预报精度, 在短期和中期的预报中除了 60 d、270 d 及 360 d 的预报精度稍弱于改进的 BP 神经网络方法, 其余跨度的预报精度均优于 BP 神经网络的预报精度或与之相当. 但是在预报效率上, GRNN 所用时间仅为 BP 神经网络预报所用时间的十几分之一. 在图中也

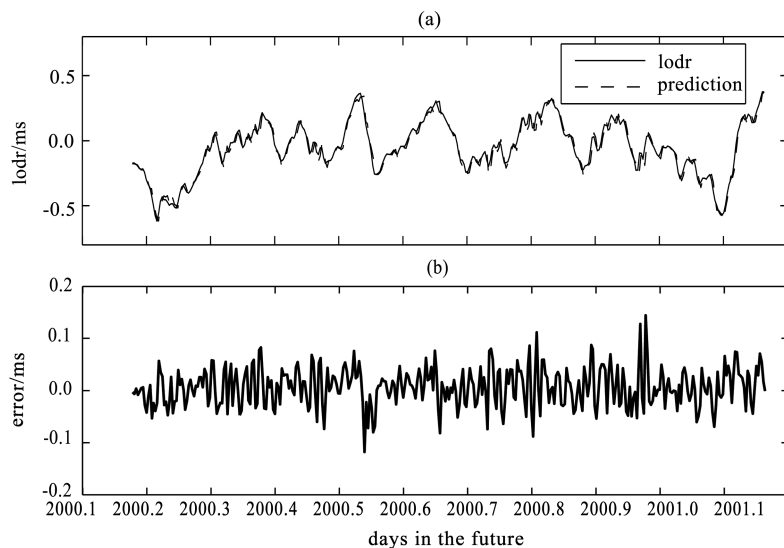
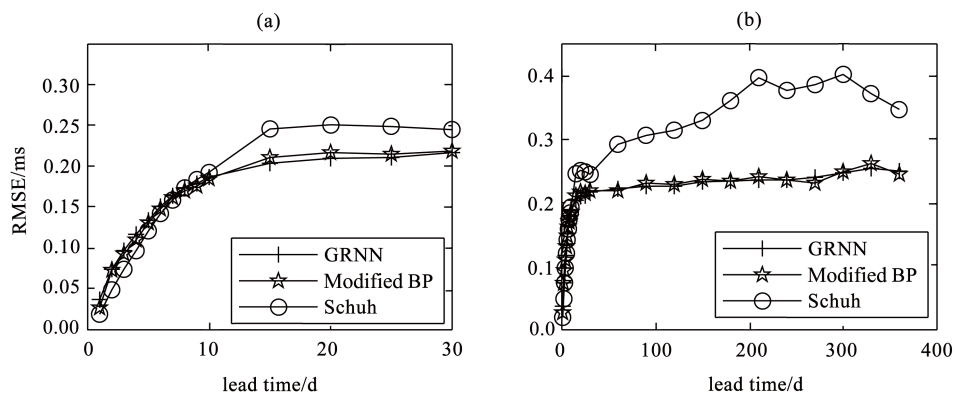


图 2 跨度为 1 d 的预报拟合结果及拟合误差图

Fig. 2 Fitted result and error with the prediction time span of 1 d

图 3 GRNN 预报精度与 Schuh 等^[2]及改进的 BP 神经网络预报精度比较. (a) 短期 (1 ~ 30 d) 预报, (b) 中期 (1 ~ 365 d) 预报Fig. 3 Comparison of the prediction accuracies of GRNN, Schuh et al.^[2] and modified BP neural networks. (a) Short-term (1 ~ 30 d) prediction, (b) medium-term (1 ~ 365 d) prediction

可以发现,随着时间跨度的增大,GRNN 预报误差曲线较光滑,预报误差呈逐渐增大趋势,而 BP 神经网络预报误差不稳定,跳跃较明显,这也是导致跨度为 60 d、270 d 及 360 d 的预报中,GRNN 方法略弱于改进的 BP 神经网络方法的直接原因之一。

为了进一步探讨 GRNN 用于日长变化预报的性能,我们也选取了 1980 年 1 月 1 日至 2005 年 9 月 30 日的日长变化残差序列作为训练样本,预报 2005 年 10 月 1 日至 2008 年 2 月 28 日(与 EOP PCC 时间段相同)不同跨度的日长变化残差值,计算了预报结果的平均绝对误差,并与 EOP PCC 的结果进行了比较^[23],比较结果见图 4~6。

图 4(a)、图 5(a)及图 6(a)表示使用 GRNN 进行日长变化预报所得 MAE 值,图 4(b)、图 5(b)及图 6(b)中不同颜色线条分别为参与 EOP PCC 不同小组所得的预报成果.黑色圆圈连线表示的是联合几种方法的预报结果,其中精度较好的几个研究小组是:黑色实线的 Gross R. 的预报结果,橘黄色实线的 Kalarus M. 的预报结果,粉红色实线的 EOP 产品中心的预报结果.预报结果的评价指标为 MAE。

从图 4~6 的对比中可以看出,对于超短期 1~10 d 的预报,GRNN 方法预报的 MAE 精度仅次于排在第 1 位(Gross R.)和第 2 位(联合预报)的预报精度;对于短期 1~30 d 的预报,GRNN 的预报精度和排在第 1(Gross R.)和第 2(联合预报)的预报精度大致相当;对于中期的预报,GRNN 的预报精度则优于 EOP PCC 预报精度。

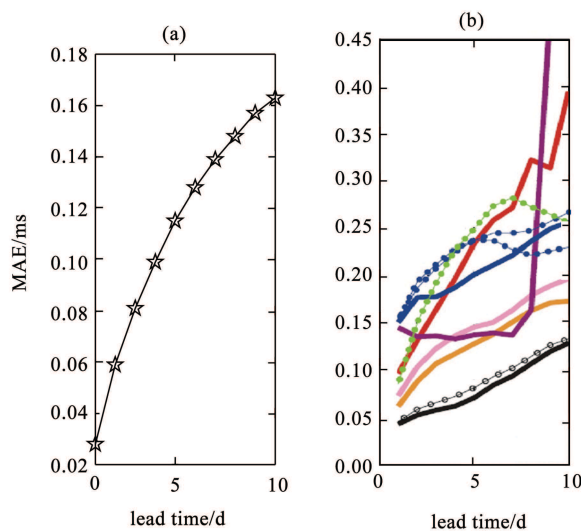


图 4 超短期(1~10 d)预报 MAE 比较. (a) GRNN 预报结果, (b) EOP PCC 预报结果^[22]

Fig. 4 Comparison of the predicted MAE for ultra short-term (1~10 d). Prediction results of GRNN (a) and EOP PCC^[22] (b)

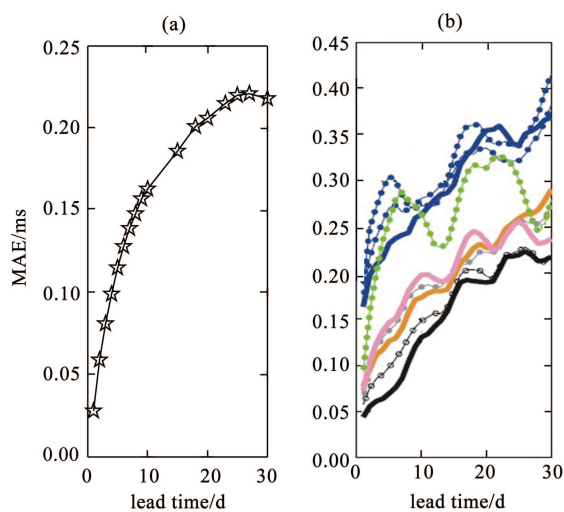


图 5 短期 (1 ~ 30 d) MAE 比较. (a) GRNN 预报结果, (b) EOP PCC 预报结果^[22]

Fig. 5 Comparison of the predicted MAE for short-term (1 ~ 30 d). Prediction results of GRNN (a) and EOP PCC^[22] (b)

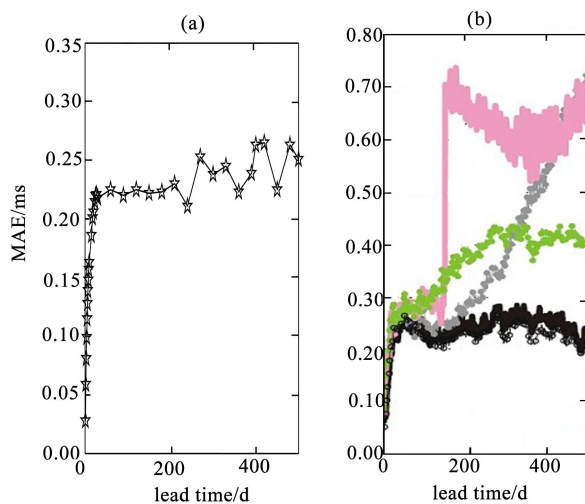


图 6 中期 (1 ~ 500 d) MAE 比较. (a) GRNN 预报结果, (b) EOP PCC 预报结果^[22]

Fig. 6 Comparison of the predicted MAE for medium-term (1 ~ 500 d). Prediction results of GRNN (a) and EOP PCC^[22] (b)

4 总结和展望

本文通过实例验证了 GRNN 用于日长变化预报的可行性. 通过与 Schuh 等^[2] 以及 EOP PCC^[23] 日长变化预报结果比较, 我们发现, 非线性的 GRNN 方法用于日长变化预报不仅能得到高精度的预报结果, 同时由于网络计算需要调节的参数只有平滑因子, 网络训练所需样本数量远远小于 BP 神经网络训练所需样本数量, 所有的计算可以通过计算机程序快速实现, 大大提高了预报效率. 以跨度为 1 d 的日长变化预报为例, 应用 BP 神经网络的方法预报日长变化需要 10 h 左右的时间, 应用 GRNN 则只要 1 h 左右, 预报效率极大提高, 这对于地球自转参数的实时快速预报有着更高的现实意义.

如果 GRNN 中平滑因子以及输入节点的个数能够通过合适的调整方法自动找到使 RMSE 最小时所对应的值, 那么 GRNN 用于日长变化预报的效率将会有更大的提高, 同时预报也会更加精确. 这一问题的研究有望通过遗传算法的全局搜索能力得以解决. 由于日长变化和大气角动量轴向分量有着密切的关系^[24-25], 下一步的研究我们将会联合日长变化残差序列以及大气角动量轴向分量进行日长变化预报.

参 考 文 献

- [1] 叶叔华, 黄 斌. 天文地球动力学. 济南: 山东科学技术出版社, 2000: 478-479
- [2] Schuh H, Ulrich M, Egger D, et al. JGeod, 2002, 76: 247
- [3] McCarthy D D. The Earth's Rotation and Reference Frames for Geodesy and Geodynamics. Dordrecht: Kluwer Academic Publisher, 1988: 275-280
- [4] McCarthy D D, Luzum B J. JGeod, 1991, 65: 18
- [5] 郑大伟. 天文学进展, 1989, 7: 120
- [6] Morabito D D, Eubanks T M, Steppe J A. The Earth's Rotation and Reference Frames for Geodesy and Geodynamics. Dordrecht: Kluwer Academic Publisher, 1988: 257-262
- [7] Niedzielski T, Kosek W. JGeod, 2008, 82: 83
- [8] Kalarus M, Schuh H, Kosek W, et al. Geop, 2005, 7: 1753
- [9] 王琪洁. 基于神经网络技术的地球自转变化预报. 上海: 中国科学院上海天文台, 2007
- [10] 王琪洁, 廖德春, 周永宏, 等. 中国科学院上海天文台年刊, 2007, 28: 23
- [11] 廖德春, 王琪洁, 周永宏, 等. 中国科学院上海天文台年刊, 2008, 29: 19
- [12] 许雪晴, 周永宏. 飞行器测控学报, 2010, 29: 70
- [13] Haykin S. 叶世伟, 史忠植, 译. 神经网络原理. 北京: 机械工业出版社, 2004: 183-200
- [14] 邹阿金, 张雨浓. 基函数神经网络及应用. 广州: 中山大学出版社, 2009
- [15] 李杰, 王科, 王航. 交通与计算机, 2007, 25: 131
- [16] Specht D F. ITNN, 1991, 2: 568
- [17] Tomoko N. Jch, 2003, 43: 113
- [18] Yoder C F, Williams J G, Parke M E. JGR, 1981, 86: 881
- [19] McCarthy D D, Petit G. ITN, 2003, 13: 605
- [20] Defraigne P, Smits I. GeoJI, 1999, 139: 563
- [21] Ray R D, Steinberg D J, Chao B F, et al. Sci, 1994, 264: 830
- [22] Seiler U, Wünsch J. ANS, 1995, 316: 419
- [23] Kalarus M, Schuh H, Kosek W, et al. JGeod, 2010, 84: 587
- [24] 王琪洁, 廖德春, 周永宏, 等. 天文学报, 2008, 49: 93
- [25] Wang Q J, Liao D C, Zhou Y H, et al. ChA&A, 2008, 32: 342

An Application to the Prediction of LOD Change Based on General Regression Neural Network

ZHANG Xiao-hong WANG Qi-jie ZHU Jian-jun ZHANG Hao
(*School of Geosciences and Info-Physics, Central South University, Changsha 410083*)

ABSTRACT Traditional prediction of the LOD (length of day) change was based on linear models, such as the least square model and the autoregressive technique, etc. Due to the complex non-linear features of the LOD variation, the performances of the linear model predictors are not fully satisfactory. This paper applies a non-linear neural network — general regression neural network (GRNN) model to forecast the LOD change, and the results are analyzed and compared with those obtained with the back propagation neural network and other models. The comparison shows that the performance of the GRNN model in the prediction of the LOD change is efficient and feasible.

Key words astrometry, time, methods: miscellaneous