doi: 10.7690/bgzdh.2015.12.015

集成神经网络在 2.4 m 跨声速风洞马赫数预测中的应用

杜　宁，蒋婧妍，郁文山，陈　龙
（中国空气动力研究与发展中心高速所，四川 绵阳 621000）

摘要：2.4 m 跨声速风洞空气流动是复杂的三维流动，想要利用机理模型来描述马赫数的特性十分困难，所以采用数据驱动的方式建立风洞马赫数模型。提出一种基于特征子集的集成神经网络建模方法，该方法选用动态 NARMAX 模型，并采用集成神经网络的方法建立了风洞马赫数预测模型；最后，进行了单一神经网络模型与集成神经网络模型在马赫数预测上的性能对比。试验结果表明：集成神经网络模型可以在保证预测准确度和泛化性的基础上，降低模型的训练和测试时间。

关键词：2.4 m 跨声速风洞；特征子集；集成神经网络
中图分类号：TP183　　**文献标志码**：A

Application of Ensemble Neural Network in the Prediction of Maher Nnumber in 2.4 m Transonic Wind Tunnel

Du Ning, Jiang Jingyan, Yu Wenshan, Chen Long
(*High Speed Institute*, *China Aerodynamics Research & Development Center*, *Mianyang* 621000, *China*)

Abstract: Air flow is a complex three-dimensional flow in the 2.4 m transonic wind tunnel, that is very difficult to use of mechanism modeling to describe the characteristics of Mach number, so we adopt the way of data driven to set up the wind tunnel Mach number model. This paper proposes the ensemble neural networks (ENN) modeling method based on feature subset, the method chooses dynamic NARMAX model, and integrated ENN method to establish a predictive model wind tunnel Mach number. Furthermore, a comparative study among the single neural networks (NN) models and the ENN models when used to predict the Mach number is conducted. Results confirm that training time and testing time are much reduced by the ensemble neural networks models.

Keywords: 2.4 m transonic wind tunnel; feature subsets; ensemble neural networks

0　引言

2.4 m 风洞是一座暂充型引射式半回流跨声速风洞，具有非线性、强耦合、大滞后和时变等特点，给马赫数的精确预测带来很大难度。风洞试验过程中，空气流动是非常复杂的三维流动，还存在复杂的波系和非定常扰动，使得用机理模型来描述马赫数动态特性变得十分困难；所以笔者建立了基于数据驱动的风洞马赫数模型。

神经网络(artificial neural networks，ANN)广泛应用于非线性问题的求解，已经在系统辨识、预测控制和模式识别等领域得到了广泛应用[1-2]。尽管具有很好的拟合效果，但是在其设计过程中仍存在许多难点，例如如何解决过拟合、局部最小以及泛化能力差等问题。

笔者提出一种基于特征子集的集成神经网络(ensemble neural networks，ENN)模型，以提高单一神经网络的效果。ENN 中的子模型都是建立在低维的数据集上，而传统的单一模型是基于整个数据集建立的；因此，神经网络子模型的复杂度将降低。

笔者选用动态 NARMAX 模型，建立了风洞马赫数预测模型，并通过 ENN 实现。试验显示：在马赫数预测上，NARMAX-ENN 的效果要好于 NARMAX-NN，同时 NARMAX-ENN 模型的训练时间和测试时间更少。

1　风洞系统的 NARMAX 描述

1.1　2.4 m 跨声速风洞介绍

2.4 m 跨声速风洞主要承担航空航天飞行器的空气动力地面模拟试验和性能评估任务[3]。在马赫数范围为 0.3～0.9 时，该风洞主要通过栅指系统调节马赫数，以及通过主排气阀系统保持稳定段总压不变。由于受暂冲式风洞固有特性的限制(即一次试验的时间不能太长)，在试验过程中风洞系统必须满足以下条件：

1) 风洞启动充压阶段，建立稳定流场的时间要小于 7 s，同时马赫数控制精度要求在±0.002 以内；

2) 在流场调节阶段，当马赫数阶梯变化时，系统稳定时间要求在 0.8 s 以内，稳态误差小于±0.002。

根据跨声速风洞系统预测控制的严格要求，在保证精度要求的前提下希望加快马赫数预测速度。

收稿日期：2015-07-20；修回日期：2015-08-29
作者简介：杜　宁(1980—)，男，四川人，硕士，工程师，从事风洞自动控制研究。

表征空气速度的马赫数表达式是通过稳定段总压 p_{o}(kPa)和驻室静压 p_{s}(kPa)计算获得[4]：

$$Ma=\sqrt{5\left\{\left[1-\left(\frac{p_{\mathrm{o}}-p_{\mathrm{s}}}{p_{\mathrm{o}}}\right)\right]^{2/7}-1\right\}}\text{。}\tag{1}$$

由式中可以看出，马赫数是通过总压和静压间接获得的。建立马赫数模型实际上就是要建立总压模型和静压模型。

根据风洞系统的结构，基于空气动力学的空气环流模型以及吹风试验特点可以获得影响总压和静压的主要因素或变量，包括主调压阀位移 S_{mc}(mm)、主排气阀位移 S_{me}(mm)、栅指位移 S_{fig}(mm)、气源压力 p_{sta}(kPa)和攻角角度 A_{att}(°)。值得注意的是，虽然总压和静压拥有相同的影响变量，但这些变量对总压和静压的影响程度不同。

1.2　NARMAX 模型

非线性自回归滑动平均模型(non-linear auto-regressive moving average model with exogenous inputs，NARMAX)是 Billings 等于 1985 年提出的[5]，并于 1985—1989 年基于 Nerode 的有限实现理论及泰勒定理对该模型描述的一般非线性系统的普适性给予了证明。NARMAX 模型因其需辨识参数少、逼近精度高、普适性强以及风洞系统自身的特点；所以笔者使用该模型来描述 2.4 m 跨声速风洞系统。

稳定段总压 NARMAX 模型形式如式 (2)，试验段静压 NARMAX 模型形式如式 (3)：

$$\hat{p}_{\mathrm{o}}(t)=f_{\mathrm{o}}\begin{pmatrix}p_{\mathrm{o}}(t-1),\cdots,p_{\mathrm{o}}(t-n_{\mathrm{o}}),S_{\mathrm{mc}}(t-1),\cdots,S_{\mathrm{mc}}(t-n_{\mathrm{mc}}),S_{\mathrm{me}}(t-1),\cdots,S_{\mathrm{me}}(t-n_{\mathrm{me}}),\\ S_{\mathrm{fig}}(t-1),\cdots,S_{\mathrm{fig}}(t-n_{\mathrm{fig}}),p_{\mathrm{sta}}(t-1),\cdots,p_{\mathrm{sta}}(t-n_{\mathrm{sta}}),A_{\mathrm{att}}(t-1),\cdots,A_{\mathrm{att}}(t-n_{\mathrm{att}}),\\ e_{\mathrm{o}}(t-1),\cdots,e_{\mathrm{o}}(t-n_{\mathrm{e}})\end{pmatrix};\tag{2}$$

$$\hat{p}_{\mathrm{s}}(t)=f_{s}\begin{pmatrix}p_{\mathrm{s}}(t-1),\cdots,p_{\mathrm{s}}(t-n_{\mathrm{s}}),S_{mc}(t-1),\cdots,S_{\mathrm{mc}}(t-n_{\mathrm{mc}}),S_{\mathrm{me}}(t-1),\cdots,S_{\mathrm{me}}(t-n_{\mathrm{me}}),\\ S_{fig}(t-1),\cdots,S_{\mathrm{fig}}(t-n_{\mathrm{fig}}),p_{\mathrm{sta}}(t-1),\cdots,p_{\mathrm{sta}}(t-n_{sta}),A_{att}(t-1),\cdots,A_{\mathrm{att}}(t-n_{\mathrm{att}}),\\ e_{\mathrm{s}}(t-1),\cdots,e_{\mathrm{s}}(t-n_{\mathrm{e}})\end{pmatrix}\text{。}\tag{3}$$

式中 n_{o}, n_{s}, n_{mc}, n_{me}, n_{fig}, n_{sta}, n_{att} 和 n_{e} 分别表示标量 p_{o}, p_{s}, S_{mc}, S_{me}, S_{fig}, p_{sta} 和 A_{att} 的阶次。

2　集成神经网络

2.1　基于特征子集的 ENN

针对回归问题，基于特征子集的 ENN 算法为：

1) 已知训练集 $\Theta_N=\{(X_k,Y_k):X_k\in R^n,Y_k\in R,k=1,\cdots,N\}$，其对应的特征空间为 $X=\{x_1,\cdots,x_n\}\in R^n$。

2) 通过到 $h(h\leqslant n)$ 划分的方式划分特征空间 $X=\{x_1,\cdots,x_n\}$，创建特征子集 $fs_i=\{X_k^{(h)},Y_k\}_{k=1}^N$，$i=1,\cdots,P$，表示第 i 个特征子集，$P=C_n^h$ 表示子模型的个数。

3) 在每一个特征子集上建立一个神经网络子模型。

4) 计算权值向量 $\boldsymbol{\beta}=[\beta_1,\cdots,\beta_P]^{\mathrm{T}}$，其中 β_i 表示第 i 个子模型 f_i 的权值。

5) 加权平均 P 个子模型得特征子集模型 $f_E(X)$。

2.2　子模型的融合

集成模型由 P 个子模型加权平均获得，形式为

$$f_E(X)=\sum_{i=1}^{P}\beta_i f_i(X)=g(X)+\sum_{i=1}^{P}\beta_i\varepsilon_i(X)\text{。}\tag{4}$$

其中：β_i 表示第 i 个子模型的权值，而且 $\sum_{i=1}^{P}\beta_i=1$；$g(X)$ 表示期望的非线性函数；$\varepsilon_i(X)=f_i(X)-g(X)$ 表示第 i 个子模型的误差；$f_i(X)$ 是建立在第 i 个特征子集上的第 i 个子模型。考虑相关矩阵 $\boldsymbol{\varGamma}_{P\times P}$，其元素如下：

$$\varGamma_{ij}=E[\varepsilon_i(X)\varepsilon_j(X)]\text{。}\tag{5}$$

实际上是由有限个样本近似估计获得：

$$\varGamma_{ij}=\frac{1}{N}\sum_{k=1}^{N}[f_i(X_k)-Y_k][f_j(X_k)-Y_k]\text{。}\tag{6}$$

其中 N 表示原始训练集样本个数。

集成模型的平方误差可表示为

$$J_E=E[\{f_E(X)-g(X)\}^2]=E[(\sum_{i=1}^{P}\beta_i\varepsilon_i)(\sum_{j=1}^{P}\beta_j\varepsilon_j)]\approx\sum_{i=1}^{P}\sum_{j=1}^{P}\beta_i\beta_j\varGamma_{ij}=\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\varGamma}\boldsymbol{\beta}\text{。}\tag{7}$$

即，可由下式获得 $\boldsymbol{\beta}$ 的最优值：

$$\min_{\beta}\frac{1}{2}\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\varGamma}\boldsymbol{\beta}\ \text{ such that }\ \sum_{i=1}^{P}\beta_i=1\text{。}\tag{8}$$

根据 Lagrangian 变化，引入 Lagrange 乘子 λ：

$$\ell(\beta,\lambda)=\frac{1}{2}\boldsymbol{\beta}^{\mathrm{T}}\boldsymbol{\varGamma}\boldsymbol{\beta}-\lambda(\sum_{i=1}^{P}\beta_i-1)\text{。}\tag{9}$$

根据最优条件：

$$\left.\begin{aligned}\frac{\partial \ell}{\partial \beta}=\Gamma\boldsymbol{\beta}-\lambda 1_v=0\\ \frac{\partial \ell}{\partial \lambda}=1_v^{\mathrm{T}}\boldsymbol{\beta}-1\end{aligned}\right\}。\tag{10}$$

得最优解

$$\boldsymbol{\beta}=\frac{\Gamma^{-1}1_v}{1_v^{\mathrm{T}}\Gamma^{-1}1_v}。\tag{11}$$

其中 $1_v=[1;\cdots;1]^{\mathrm{T}}$。如果相关矩阵为病态矩阵(ill-conditioned)可以使用规范化调整；或者为了避免过大的正、负权值，可附加约束条件。

3　仿真试验

从风洞吹风试验中获得了 8 组数据集。表 1 给出了 8 组数据集的样本个数，其中前 3 组数据作为训练集，其他数据集作为测试集分别进行测试。

表 1　8 组数据集的样本个数

工况	Ma=0.6 p_o=110 kPa	Ma=0.54 p_o=110 kPa	Ma=0.85 p_o=130 kPa	合计
1	10 440	4 091	11 070	25 601
2	11 220	3 836	11 004	26 060
3	10 986	4 504	11 100	26 590
4	10 800	4 546	11 010	26 356
5	9 655	4 300	9 930	23 885
6	10 641	4 183	10 026	24 850
7	9 684	4 176	9 780	23 640
8	10 032	4 187	8 851	24 170

所有数据都归一化到[−1,1]之间，均方根误差(root-mean-square error，RMSE)作为模型预测效果的评价标准。基于其串行运行的特点以及 ENN 的并行结构，笔者测量了模型的总训练时间和每组测试集的总测试时间，然后将它们分别除以子模型个数，所得到的平均值作为集成神经网络模型的训练时间和测试时间，并与单一神经网络模型进行比较。

文中具体地选择了 BP 神经网络。NARMAX 总压模型和 NARMAX 静压模型都分别使用单一神经网络和集成神经网络实现，并作出比较。文献[1]中已经证明：在 BP 神经网络中，如果只有一层 sigmoid 或线性隐层，而且这一层隐层中拥有足够多的神经元，那么这个单层的 BP 神经网络足以描述任意的输入输出变量之间的非线性函数关系；因此，笔者只考虑拥有一层隐层的 BP 神经网络，以 sigmoid 激活函数作为神经元。另外，文献[1]给出了当 BP 神经网络只拥有一层隐层时，隐层中神经元个数的确定方法，如下式所示：

$$NN=\log_2(n)。\tag{12}$$

其中 n 表示 BP 网络的输入特征(变量)维数。

根据式 (12)，如果 BP 子模型只有 2 个输入特征，那么神经元的个数 NN 等于 1。无论是单一 BP 模型还是 FSE-BPs 模型，训练 BP 神经网络的过程中都使用均方误差作为网络性能的评价准则。

对于只有一层隐层的 BP 神经网络，NARMAX-NN 模型中 BP 子模型所需要的神经元个数明显要比单一 BP 模型少得多，从而大大降低模型的复杂度。这意味着 FSE-BPs 模型的训练时间不仅要比单一 BP 模型短，实际使用预测速度也要比单一 BP 模型快。表 2 和表 3 中的试验结果也证实了该结论。

表 2　NARMAX-NN 模型结果

NN 模型结果	#数据集	总压(p_o/kPa) RMSE	总压 时间/s	静压(p_s/kPa) RMSE	静压 时间/s	马赫数 RMSE
训练	1～3	0.021 5	2 032.7	0.018 6	2 031.5	0.001 0
测试	4	0.024 8	0.188	0.018 4	0.188	0.001 1
	5	0.020 9	0.172	0.018 3	0.188	0.001 1
	6	0.021 4	0.172	0.018 4	0.172	0.001 1
	7	0.022 5	0.172	0.019 1	0.187	0.001 0
	8	0.024 4	0.187	0.020 0	0.172	0.000 9

表 3　NARMAX-ENN 模型结果

ENN 模型结果	#数据集	总压(p_o/kPa) RMSE	总压 时间/s	静压(p_s/kPa) RMSE	静压 时间/s	马赫数 RMSE
训练	1～3	0.021 4	320.19	0.019 3	349.52	0.000 9
测试	4	0.041 0	0.037 6	0.019 1	0.040 5	0.001 1
	5	0.021 2	0.039 1	0.018 8	0.042 6	0.001 1
	6	0.045 3	0.038 4	0.018 8	0.037 7	0.001 1
	7	0.021 4	0.038 4	0.019 5	0.037 6	0.001 0
	8	0.045 7	0.039 0	0.020 5	0.037 6	0.001 1

文中只关注马赫数的预测性能。表 2 和表 3 的试验结果显示：与 NARMAX-NN 模型相比，NARMAX-ENN 模型在马赫数预测准确度和泛化能力的基础上，大大减少了模型的训练时间和测试时间，同时也提高了马赫数预测的稳定性。

4　结束语

笔者建立了基于特征子集的集成神经网络建模方法，用于风洞马赫数的预测。神经网络子模型建立在低维数的特征子集上，与建立在整个数据集上的单一神经网络相比，具有更低的复杂度，从而大大提高了模型的测试速度。试验结果证明：使用 NARMAX-ENN 获得的总静压模型的训练时间和测试时间明显缩短，马赫数预测模型性能也显著提高。

参考文献：

[1] Chai M L, Song S, Li N N. A review of some main improved models for neural network forecasting in time series[J]. Proc. IEEE Intel. Veh. Symp, 2005, 68: 866–868.

[2] 海金. 神经网络原理[M]. 北京: 机械工业出版社, 2004: 15–18.

[3] 战培国, 杨炯. 国外风洞试验的新机制、新概念、新技术[J]. 流体力学实验与测量, 2004, 18(4): 1–4.

[4] 恽起麟. 实验空气动力学[M]. 北京: 国防工业出版社, 2000: 10–20.

[5] 唐亮, 徐晓鸣. 一种基于前馈神经网络的 NARMAX 模型辨识新方法[J]. 电机与控制学报, 1998, 2(3): 141–144.