

基于多源信息融合的能源需求预测模型研究综述

张金锁^{1, 2}, 冯雪^{2, 3}, 邹绍辉^{1, 2}

(1.西安科技大学管理学院, 陕西西安 710054; 2.西安科技大学能源经济与管理研究中心, 陕西西安 710054; 3.西安科技大学能源学院 陕西西安 710054)

摘要: 针对已有能源需求预测存在预测误差越来越突出的问题, 从信息源的角度出发, 在梳理国内外相关理论研究现状的基础上, 分别从单一模型、多源模型组合和智能信息融合三方面进行评述并找出问题, 最后给出该领域进一步的展望。

关键词: 多源信息融合; 单一模型; 多源模型组合; 智能信息融合; 能源需求预测

A review of energy demand forecasting model

based on multi-source information fusion

Jinsuo Zhang^{1, 2}, Xue Feng^{2, 3} and Shaohui Zou^{1, 2}

(1.School of Management, Xi'an University of Science and Technology, Xi'an 710054, China; 2.Research Center for Energy Economy and Management, Xi'an University of Science and Technology, Xi'an 710054, China;
3. College of energy, Xi'an University of Science and Technology, Xi'an 710054, China)

Abstract: Based on the improvement of the prediction accuracy, this paper reviewed the domestic and international relevant theory and characteristics of model of energy demand forecasting from the perspective of information source. Energy demand forecasting model can be divided into 3 kinds: single model, multi-model combination and intelligent information fusion model. The study reviewed the model's characteristics and existing problems, and presented suggestions on the future fields.

Keywords: Multi-source information fusion; Single model; Multi-model combination; Intelligent information fusion; Energy demand forecasting

1 引言

能源是各国经济发展和社会进步的基本物资保障, 能源可持续发展对一个国家的农业现代化、工业化、城市化以及居民生活水平的提高都起着决定性的作用。能源需求预测是能源供给的重要先决条件, 是维系能源供需平衡的重要预期因素, 已逐渐成为世界各国普遍关注的重要战略问题。准确地预测能源需求趋势, 有利于规避能源供给风险, 降低能源供需缺口, 减缓经济周期波动, 进而促进经济社会可持续发展。我国是一个能源消费大国, 能源需求变化对全球能源格局的影响日益增强, 科学的能源需求预测是制定相关能源发展战略的重要依据。近十

基金项目: 国家自然科学基金项目 (71273206, 71273207); 陕西省软科学计划重点项目 (2012KR2-01); 陕西省教育厅科研计划项目 (2010JK185)

通讯作者: 张金锁, 男, 1962 年 8 月生, 博士, 教授, 博士生导师, 研究方向为资源环境政策和管理系统工程。E-mail : mark56zhang@163.com

年来,我国对能源需求预测工作不断得到重视,相关文献也越来越多,然而已有能源需求预测方法普遍存在影响因素筛选简单、模型处理的信息源不完备的问题,在此基础上较大的预测误差问题难以避免。鉴于此,本文对能源需求预测领域的文献进行系统分析与述评,对预测模型存在的问题进行剖析,在此基础上提出构建包含多源信息的能源需求预测模型分析框架,以期改善和提高现有能源需求预测模型的预测精度,提高该类模型的应用价值。

2 基于信息源的我国能源需求预测模型研究现状

2.1 基于信息单一化的能源需求预测模型

信息单一化是指仅采用单一模型预测能源需求。单一模型是一种纯数学模型的拟合估计方法。根据能源需求分析方法的基本假设不同,大致可将其分为两类:第一类是单一自相关时间序列模型,是一种基于能源系统自身对能源需求进行预测的研究方式,基本假设是影响需求变化的一切因素都已反应在需求变量本身中,这些影响因素不会发生突变,因此以历史需求可以推测未来需求。以趋势外推模型、ARMA 模型、灰色模型为主要代表。第二类是多因素相关关系预测模型,是基于整个社会经济系统对能源需求进行预测,基本假设是预测值不局限于自身各滞后期的影响,还包括其它相关因素的作用甚至是相互作用。该类模型主要包括多元回归模型、协整与误差修正模型、向量自回归模型、系统动力学模型、人工神经网络模型和情景分析法^[1]。

2.2 基于信息多源化的能源需求预测模型

基于信息单一化的预测模型在能源需求预测领域得到了充分的运用,但是,当样本数据的发展趋势受到随机因素的影响时会导致较大的随机波动,单一模型的适用性及预测精度就很难提高。这主要是由于单一预测模型的信息源有限,无法考虑更多的波动因素。随后预测模型的构建者都是希望可以得到系统中尽量多的特性(或者说信息),越多的信息被纳入预测系统,预测效果就越好。但是系统的复杂特性很难将全部的信息都考虑到。通过文献梳理,以下两类基于信息多源化的能源需求预测模型得到重视。

2.2.1 多源模型组合

Reid (1968) 以及 Bates 和 Granger (1969) ^[2] 分别提出多源模型的组合预测理论。该理论认为,有不同可供预测的单一模型,同时也有不同的组合方法,预测精度往往各不相同,没有一种方法可以达到 100% 的预测精度,但是组合模型有效地集结了各种数据样本的信息以及不同模型的有用信息,比单一预测模型考虑了更为全面的信息,因而可以有效减少或者抵消单个模型中的一些随机因素的影响^[3],提高了预测模型的预测精度和可信度。按组合预测方法与各单一预测方法的函数关系,多源模型组合可分为线性组合和非线性组合。

(1) 线性组合

线性组合也叫并联组合,是对源自不同模型预测结果间的组合,做法是把各个竞争模型得到的预测结果赋予不同的权重来组合为一个预测结果。若用 k 个单一模型 g_1, g_2, \dots, g_k 预测某一时间序列,则对 k 个预测结果通过线性组合模型表示为:

$$\hat{y}_t = \sum_{i=1}^k \omega_i g_i$$

其中, ω_i 是预测模型 g_i ($i=1, 2, \dots, k$) 的权重值。从公式可以看出,线性

组合方法的研究重心是如何确定组合预测模型的权重，常见的方法有等权重法、最小方差法、优势矩阵法和线性回归法。在能源需求预测领域中运用的线性组合权重的求解方式通常采用非最优组合法和最优组合法。

1) 非最优组合模型

非最优组合法确定组合预测加权值的基本思想是，若某个单一预测模型预测误差序列的变异系数越大，则其在组合预测中对应的权值就越小。在统计学中，一般采用方差或标准差作为反映预测误差序列变异程度的指标。在能源需求预测领域，这类权值的确定方法对应的就有标准差法、误差平方和倒数法等。

①标准差法

标准差法是根据各个单项预测模型预测误差的变异程度来确定组合权重。

$$\omega_i = \frac{\sigma - \sigma_i}{\sigma} \cdot \frac{1}{m-1}$$

式中， σ_i 是第*i*个单项模型预测误差的标准差， $\sigma = \sqrt{\sum_{i=1}^m \sigma_i^2}$

卢奇，顾培亮^[4]分别采用灰色预测、神经网络及多元回归方法预测能源消费，并采用标准差法进行组合预测。王会强，胡丹^[5]建立了我国能源消费系统的ARMA模型和灰色预测模型，通过标准差法进行组合模型预测。高军，宋书贵等^[6]建立灰色—增长曲线组合预测模型，同样采用标准差法确定权值，对四川“十二五”期间的天然气消费量进行了预测，为提高精度还用马尔科夫链对预测结果进行了修正。

②预测误差平方和倒数法

预测误差平方和倒数法，也叫方差倒数法。预测误差平方和是反映预测精度的一个指标，值越大，表明该项预测模型的预测精度就越低，那么它在组合预测中的重要性就越低，它在组合预测中的加权系数就越小。相反，对预测误差平方和较小的单项预测模型在组合预测中就要赋予较大的权值。令：

$$l_i = \frac{E_{ii}^{-1}}{\sum_{i=1}^m E_{ii}^{-1}}, \quad i = 1, 2, \dots, m$$

显然 $\sum_{i=1}^m l_i = 1, \quad l_i \geq 0, \quad i = 1, 2, \dots, m$ ，其中 E_{ii} 为第*i*种单一预测模型的预测误差平方和

差平方和

$$E_{ii} = \sum_{t=1}^N e_{it}^2 = \sum_{t=1}^N (x_t - x_{it})^2$$

x_{it} 为第*i*种单一预测方法在第*t*时刻的预测值， x_t 为同一预测对象的某个指标序列 $\{x_t, t=1, 2, \dots, N\}$ 第*t*时刻的观测值。*N*表示时间长度， $e_{it} = (x_t - x_{it})$ 为第*i*种单一预测方法在第*t*时刻的预测误差。

张玉春，郭宁等^[7]在应用ARMA模型和GM(1, 1)模型对甘肃省的能源需求进行预测时，用方差倒数法将两个模型构成组合模型进行预测，结果发现组合模型预测精度明显高于单一模型预测精度。

③熵值法

信息论中熵值是对不确定性的一种度量,可用来判断事件的随机性及无序程度,也可用来判断某个指标的离散程度。在预测模型中,通过定义单项预测模型预测误差序列的变异系数,由此作为反映其变异程度的一个指标并确定权系数,即为熵值法的具体应用。熵值法确定组合预测加权系数的步骤如下:

第一步:把各单项预测相对误差序列归一化,也就是计算第*i*种单项预测方法第*t*时刻的预测相对误差的比重:

$$p_{it} = \frac{e_{it}}{\sum_{t=1}^N e_{it}}, \quad t = 1, 2, \dots, N$$

$$\text{此时 } \sum_{t=1}^N p_{it} = 1, \quad i = 1, 2, \dots, m$$

第二步:确定第*i*中单项预测相对误差熵值。

$$h_i = -k \sum_{t=1}^N p_{it} \ln p_{it}, \quad i = 1, 2, \dots, m$$

$k>0$ 为常数, $h_i \geq 0$, $i=1, 2, \dots, m$, 需要说明的是,对第*i*种单项预测方法而言,如果 p_{it} 全部相等,即 $p_{it}=1/N$, $t=1, 2, \dots, N$,那么 h_i 取极大值,即 $h_i=k \ln N$,这里取 $k=1/\ln N$,则 $0 \leq h_i \leq 1$ 。

第三步:确定第*i*种单项预测方法预测相对误差序列的变异系数。因为 $0 \leq h_i \leq 1$,而系统某项指标的熵值的大小与其变异系数相反,所以定义第*i*种单项预测方法的预测相对误差序列的变异系数 d_i 为

$$d_i = 1 - h_i, \quad i = 1, 2, \dots, m$$

第四步:计算各预测方法的加权系数

$$l_i = \frac{1}{m-1} \left(1 - \frac{d_i}{\sum_{i=1}^m d_i} \right), \quad i = 1, 2, \dots, m$$

由上式可以看出一个原则,那就是某个单项预测方法的预测相对误差序列的变异系数越大,则其在组合预测中对应的权值系数就越小。当然权系数要满足

$$\sum_{i=1}^m l_i = 1。$$

第五步:计算组合预测值

$$\hat{x}_t = \sum_{i=1}^m l_i x_{it}, \quad t = 1, 2, \dots, N$$

根据以上原理,秦晋栋^[8]结合灰色预测模型最新理论成果,提出了基于新陈代谢模型,GM(1,1)模型以及离散DGM模型的组合预测模型,组合预测模型中各单项方法的权重采用熵值法来确定,以此对湖北省2001-2010年能源需求总量进行实证研究。陈玉金,刘建永等^[9]在对回归分析预测模型和BP神经网络预测模型的组合预测模型研究时的权系数也是基于熵值法确定,并结合某省能源消耗总量数据进行仿真,发现结果与实际数据误差较小。

2) 最优组合模型

基于预测误差指标的最优组合预测模型在能源需求预测的应用中,最为广泛的是以预测误差平方和达到最小的线性组合模型组合。

①线性不变权组合模型

假定同一预测对象的某个指标序列为 $\{y_t, t=1, 2, \dots, N\}$, 有 n 种单项预测方法对其进行预测, 若第 i 种单项预测方法在第 t 时刻的预测值为 f_{it} , $i=1, 2, \dots, n, t=1, 2, \dots, N$, 其中 $e_{it}=(y_t-f_{it})$ 为第 i 种单项预测方法在第 t 时刻的预测误差, k_i 为第 i 种方法的组合加权系数。为确保组合预测无偏性, 加权系数应满足:

$$\sum_{i=1}^n k_i = 1, \quad i=1, 2, \dots, n$$

假定 $f_t = \sum_{i=1}^n k_i f_{it}$ 为 y_t 的组合预测值, 设 e_t 为组合预测在第 t 时刻的预测误差,

$$\text{则有: } e_t = y_t - f_t = \sum_{i=1}^n k_i e_{it}$$

假设 J_1 是组合预测误差平方和, 则有:

$$J_1 = \sum_{t=1}^N e_t^2 = \sum_{t=1}^N \sum_{i=1}^n \sum_{j=1}^n k_i e_{it} k_j e_{jt}$$

以此, 以预测误差平方和最小为准则的线性组合预测模型实际上是下列最优化问题:

$$\begin{cases} \min J_1 = \sum_{t=1}^N \sum_{i=1}^n \sum_{j=1}^n k_i e_{it} k_j e_{jt} \\ \sum_{i=1}^n k_i = 1 \end{cases} \quad (10)$$

负加权系数出现是没有经济意义的, 因此要确保是非负权系数组合预测模型, 这就要在模型(10)中增加一个非负约束条件, 变换为:

$$\begin{cases} \min J_1 = \sum_{t=1}^N \sum_{i=1}^n \sum_{j=1}^n k_i e_{it} k_j e_{jt} \\ \sum_{i=1}^n k_i = 1 \\ k_i \geq 0, \quad i=1, 2, \dots, n \end{cases} \quad (11)$$

模型(11)可以转化为矩阵形式, 实际上是非线性二次凸规划问题, 具有唯一最优解。

根据以上原理, 侯运炳, 张文等^[10]建立了计量技术与部门分析灰色系统结合的优化组合模型, 以河北省煤炭需求为研究对象回归分析, 并利用组合模型预测了河北省2007-2012年的煤炭需求, 取得了比较理想的效果。宋雅晴, 杨桂元^[11]根据1990-2008年我国能源消费的数据, 建立我国能源消费的组合预测模型, 发现组合预测模型精度高于各单项预测模型。曹飞^[12]分别采用指数回归模型、能

源需求弹性回归模型及灰色模型建立了陕西省能源消费系统的单项预测模型，并以预测误差平方和最小为准则解决了非负权重分配问题，由此建立了陕西省能源消费组合预测模型。

②可变加权系数组合预测模型

线性组合模型中还有一种权系数确定方法可变的，即最优非负可变加权系数的组合预测模型。该模型对各个单项预测模型在各时点赋予适当权重进行组合，因此权重是随预测时间变化的函数。

加权系数变化的原因主要基于两方面：一是不同的预测方法特点不同，“时好时坏性”反映在权重上就是“时大时小”；二是不同的预测者对事物的客观规律的认识有优劣，某种预测方法可能随时间的推移越来越优于其它的单项预测方法，客观上需要组合预测加权系数变化^[13]。

假定同一预测对象的某个指标序列为 $\{y_t, t=1, 2, \dots, N\}$ ，有 n 种单项预测方法对其进行预测，设第 i 种单项预测方法在第 t 时刻的预测值为 f_{it} ， $i=1, 2, \dots, n, t=1, 2, \dots, N$ ，令：

$$f_t = \sum_{i=1}^n k_{it} f_{it}$$

其中， f_t 是 y_t 的组合预测值， k_{it} 是第 i 种预测方法在第 t 时刻的组合加权系数，且 k_{it} 满足：

$$\begin{cases} \sum_{i=1}^n k_{it} = 1, t = 1, 2, \dots, N \\ k_{it} \geq 0, i = 1, 2, \dots, n, t = 1, 2, \dots, N \end{cases}$$

假设 e_t 为组合预测在第 t 时刻的预测误差，那么：

$$e_t = y_t - f_t = \sum_{i=1}^n k_{it} e_{it}$$

在这里， $e_{it} = (y_t - f_{it})$ 为第 i 种预测方法在第 t 时刻的预测误差。

假设 J_1 表示组合预测误差平方和，则有：

$$J_1 = \sum_{t=1}^N e_t^2 = \sum_{t=1}^N \sum_{i=1}^n \sum_{j=1}^n k_{it} e_{it} k_{jt} e_{jt}$$

因此，以预测误差平方和为准则的非负可变加权系数组合预测模型转化为下列最优化问题：

$$\begin{cases} \min J_1 = \sum_{t=1}^N \sum_{i=1}^n \sum_{j=1}^n k_{it} e_{it} k_{jt} e_{jt} \\ \sum_{i=1}^n k_{it} = 1, t = 1, 2, \dots, N \\ k_{it} \geq 0, i = 1, 2, \dots, n, t = 1, 2, \dots, N \end{cases}$$

周剑^[14]利用变权重组合预测模型对辽宁省中长期能源供给与需求进行综合预测。赵国浩，郭淑芬等^[15]分别建立灰色系统、多元回归两个单项预测模型，构建了中国1981-2008年的煤炭需求变权组合预测模型，并对中国未来12年煤炭

需求进行了预测。索瑞霞，王福林^[16]通过离异系数法将灰色GM(1, 1)模型、BP神经网络模型和三次指数平滑模型进行优化组合，预测我国2009-2015年的能源消费。韩天锡，王彦霞^[17]建立了多变量非线性石油消费组合预测模型，并通过简单加权平均法确定组合系数，减少了不确定因素的影响。

(2) 非线性组合

非线性组合也叫串联组合，是利用非线性模型对各单一模型的预测结果构造非线性组合函数，通过非线性模型的训练学习赋权，将各预测结果加总得到最终预测结果。非线性组合预测主要包括加权调和平均组合预测、加权几何平均组合预测和神经网络组合预测。目前在能源需求领域以神经网络不变权非线性组合预测较为盛行。

Donaldson和Kamstra^[18]提出了用人工神经网络(ANN)方法来求解非线性组合预测模型。以BP(Back-Propagation)神经网络为例，先建立组合预测模型，然后将预测值预测结果 f_{it} ($i=1, 2, \dots, n, t=1, 2, \dots, N$) 作为神经网络的输入值，将能源需求序列的历史实际值 y_t 作为神经网络的输出值。各单项方法在组合中的权重通过网络的学习获得，假定 y_t 是 f_{it} 之间的最优组合，对于以 n 种预测方法的预测结果 $(X_{1t}, X_{2t}, \dots, X_{nt})$ 为输入，隐含单元为 p 个，输出为 S_t 的 $n-p-1$ 的拓扑网

络结构(如图1)存在的关系为：对任意隐含单元 j 的净输入， $u_t = \sum_{i=1}^n \omega_{ij} x_{it}$ ，隐含

单元 j 的输出 $m_t = f(u_t)$ ， $f(\cdot)$ 为传递函数，常采用对数S型 $\log \text{sig}(x) = 1/(1+e^{-x})$ ，或正切S型 $\tan \text{sig}(x) = (1+e^{-2x})/(1-e^{-2x})$ ，或纯线型函数 $\text{pure lin}(x) = x$ 。

$$S_t = \sum_{j=1}^p \omega_j m_{jt}$$

对于每个模式 (X_t, Y_t) ，其网络输出 S_t 与其实际数据 Y_t 之间的误差为

$$E_t = \frac{1}{2} (S_t - y_t)^2$$

系统误差为：

$$E_t = \frac{1}{2} \sum_{t=1}^n (S_t - y_t)^2$$

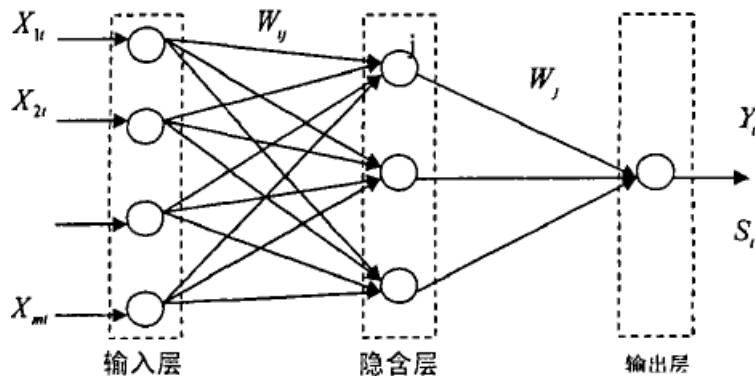


图1 最优组合预测的神经网络结构图

网络学习存在正向传播和反向传播两个过程。正向传播过程中，输入信号 X 通过隐层单元作用于输出单元，经过非线性变换后产生输出信号，一旦出现网络输出值 S 与期望输出值 Y 之间的偏差不能满足精度设定的要求，则转入反向传播，

误差信号沿原来连接通路返回,通过调整输入单元与隐层单元的联接强度取值和隐层单元与输出单元之间的联接强度以及阈值,使误差沿梯度方向下降。如此过程经过反复学习训练,确定与最小误差相对应的网络参数(权值和阈值),训练由此停止。此时经过训练的神经网络即能对类似样本的输入信息,自行处理输出误差最小的经过非线性转换的信息。

神经网络模型能够实现输入与输出之间的高度非线性映射,因此是一个强有力的学习系统。网络学习的实质是寻找一个 M ,进行 m 维输入向量空间到 n 维输出向量空间的转换,在某种意义上(通常是最小二乘意义下), M 与实际输出能够实现最佳逼近。Kolmogorov存在定理证明采用隐节点的网络的映射是存在的^[19]。Irie则证明^[20],只要网络隐单元可依要求自由设定,那么对给定任一连续函数 $F: E^m \rightarrow R^n$, $F(X)=Y$, 其中, E 为单位区间 $[0, 1]$, F 可以精确地由一个三层网络以任意精度逼近。从理论上讲,采用基于神经网络的最优组合预测方法通过精度与学习参数调节,可以获得比最优组合预测更好的效果。

国内外采用神经网络非线性组合模型进行预测的研究越来越多。Murat Kankal, Adem Akpınar^[21]等考察了土耳其4类经济指标,构建不同的组合模型,最终发现以**社会经济指标和人口指标**确定神经网络非线性组合模型预测能源需求效果最好。Pao^[22]探讨了多个单一模型和组合模型,结果表明SEGARCH-ANN和WARCH-ANN两个非线性组合模型在预测台湾的电力需求量和石油需求量时,预测精度显著提高,其中WARCH-ANN效果优于其他模型。周扬, 吴文祥^[23]将动态等维灰数递补GM(1, 1)模型、GM(1, 1)模型、无偏GM模型、基于滑动平均法改进的GM(1, 1)模型用BP神经网络非线性集成,预测江苏省未来15年煤炭和石油的需求量,取得良好效果。汪克亮, 杨力^[24]也是基于非线性组合的思想,提出了电力需求的非线性回归组合神经网络模型。

除此以外,加权调和平均组合预测模型也有所应用。师亮云, 刘家军等^[25]将诱导有序加权调和平均算子和马尔科夫链相结合,提出了一种基于诱导有序加权调和平均算子和马尔科夫链的组合预测模型,对我国用电量进行了预测。

2.2.2 智能信息融合

在面对不确定性因素的影响时,能源需求预测问题总是变得扑朔迷离。人们并没有停止和满足于多源模型组合,新理论和新方法在多源模型组合中不断得到改进。随着计算机技术的日趋完善,以及人工智能模型的不断成熟,结合并利用专家系统在能源需求预测领域开拓出一条新的思路。Dennis L.Meadows(1973)等^[26]于1970年代借助计算机技术来模拟和研究能源发展问题,虽然结论并不可靠,但是第一次能源方面相关问题的系统化分析。专家系统是一个用基于知识的程序设计方法建立起来的计算机系统,它拥有某个领域内顶级专家的知识 and 经验,并存放过去几年的数据和其他各种影响因素,如图1,利用专家的知识 and 信息是对不可量化的经验进行转化的一种较好的方法,将预测模型与专家系统的融合,就是一种智能信息融合。

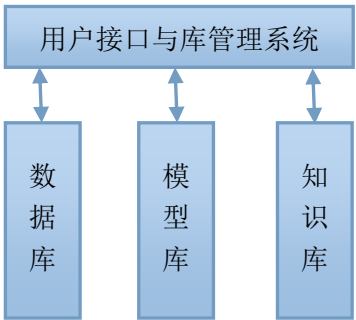


图 1 专家系统

已有智能信息融合模型在能源需求预测领域中的研究主要应用于电力负荷预测方面。中长期电力负荷预测方面，倪军，杨明志等^[27]对各个模型逐一进行评估决策，最终表明专家系统技术在城网中长期负荷预测的方法是可行的。王德金^[28]研究了国内外常用的负荷预测算法和相应的适用情况，采用熵权法解决组合预测权重，为使负荷预测更加智能化，该研究开发了一套基于 GIS 地理信息系统的负荷预测模块，设计了专家推荐模型，融合专家推荐值来不断调整不满意的预测结果。在中短期电力负荷预测方面，大多数学者^[29-32]利用神经网络的非线性逼近能力预测负荷曲线，然后通过专家系统根据天气因素或特殊事件对负荷曲线进行修正，以此提高负荷预测精度。电力负荷预测中智能信息融合的缺陷是只有神经网络模型的基础预测。

中科院余乐安博士首次将智能信息融合的理念应用于油价预测中，基本思想是：首先利用文本挖掘技术搜索影响油价波动的各种因素，然后通过粗集预处理和神经网络学习获得的相应规则形成专家系统的知识库，形成推理和解释机制，最后通过用户界面获取相应的预测信息，即提供各种数据频率（天数据、周数据、月度和年度数据）的油价预测结果。在此过程中，石油价格序列中非线性信息通过非线性模型预测，线性信息通过线性模型预测，不规则事件通过专家系统来处理，最后将不同的信息融合在一起形成最终预测结果^[33]。

两类基于信息多源化的预测模型，第一类是以模型信息为核心，通过多源模型组合的不断创新以最大限度捕捉能源需求系统的信息，另一类是以专家信息源为核心，通过专家系统，融合不规则事件的影响，突破多源模型无法捕捉的“突变点”，使预测系统智能化。

3 基于信息源的我国能源需求预测模型研究评述

3.1 单一模型

各单一模型根据各自的数学理论都有其不同的适用情况，既有不可替代的优势，也有其无法克服的缺点，不能简单、绝对地评判哪个模型的优劣，对各单一模型的适用范围及优缺点作以归纳梳理（表1）。

表 1 各单一预测模型适用性及优缺点梳理表

| 预测方法 | 时间范围 | 适用情况 | 优点 | 缺点 |
|---------|-------|--------------------------------------|------------------------|---|
| 趋势外推法 | 中期到长期 | 当被预测项目的有关变量用时间表示时，用非线性回归适用于预测周期较短的预测 | 只要历史数据，所需数据量较少 | 由于传统时间序列只能预测能源按过去的轨迹变化时的需求，无法考虑能源需求与其影响因素之间的非线性关系，当遇到历史数据起伏较大，预测精度较差。 |
| 指数平滑法 | 短期 | 具有或不具有季节变动的反复预测适用于预测量大、周期短的预测 | 简单、快速 | |
| ARMA 模型 | 短期 | 适用于平稳、变化比较均匀的短期预测 | 对非平稳序列建模时，能较好地处理随机扰动因素 | 1. 对非平稳序列进行平稳化处理即差分时，往往会丢失原序列中的重要信息，即趋势项和周期项，导致经济变量间的均衡关系研究产生谬误； 2. 没有考虑能源需求变化的因素，只致 |

| | | | | | |
|----------------|--------|-------------|--|---|--|
| | | | | | 力于数据的拟合，对规律性的处理不足 |
| 灰色模型 | | 中、长期 | 适用于时序较为平稳，且发展呈指数型趋势，对于其它变化趋势，则有时拟合灰度较大，导致精度难以提高。 | 1. 不需要影响预测对象的因素数据，预测结果不受因素数据的影响，能解决历史数据少、序列的完整性及可靠性低的问题，使复杂问题大大简化。不考虑分布规律，不考虑变化趋势，趋势项模拟能力较强，运算方便，短期预测精度高，易于检验 2. 在建模时都不需要计算统计特征 | 1.数据离散程度越大，即数据灰度越大，预测精度越差；2.不太适合非指数型趋势的序列，针对波动性较大的序列，需通过灰色马尔科夫模型来提高精度 |
| 回归模型 | 一元线性回归 | 短、中期适用于中期负荷 | 自变量与因变量两个变量之间存在着线性关系 | 能够研究预测对象与有关因素的关系，考虑了相关性；模型参数估计技术比较成熟，预测过程简单，操作简便，预测结果也比较科学可信 | 要求数据满足先决条件，回归变量不好考量，这样模型不好建立恰当，认为数据对预测的影响一样，不符合实际，这些都会影响到预测的精度；线性回归分析模型预测精度较低；非线性回归预测计算开销大，预测过程复杂；由于我国现在处于经济调整阶段，不具有一般规律性。因此，该方法在我国的实用性有待提高。 |
| | 多元线性回归 | | 因变量与两个或两个以上自变量之间存在着线性关系 | | |
| | 非线性回归 | | 因变量与一个或多个自变量之间存在着非线性关系 | | |
| 协整与误差修正模型 | | 长期 | 适合于非平稳时间序列 | 协整理论可以克服随机性时间序列的缺陷，提供了分析非平稳数据均衡关系的方法 | 已有研究均是将能源需求及其影响因素的变化关系假定为恒定不变，是一种静态变化关系，这种研究框架无法反映改革开放以来制度变迁对我国能源消费的内在结构特点以及动态变化的相互影响关系，进而也无法为我国以后进行的重大经济体制改革如何影响未来的能源需求提供理性的认识和科学的预测依据。 |
| 向量自回归模型和脉冲响应函数 | | | 适合于分析各影响因素的冲击对能源需求的传递效应以及各自对能源消费的贡献程度 | 向量自回归模型、系统动力学模型可以克服以上模型线性、静态的局限性，能反映多个因素对能源需求的影响，建立它们之间非线性、动态的关系。向量自回归模型还避免了依据经济理论或对经济行为的认识来建模的主观性和局限性，并且有助于更准确地把握能源系统的动态变化。 | VAR 模型的局限性在于其使用的经济理论并不要求完全规范，而系统动力学模型的运行成本很高 |
| 系统动力学模型 | | 中、长期 | 对复杂系统隐藏规律的研究发掘尤为适用 | | |
| 人工神经网络 | | 中、短期 | 适合于大样本非线性问题 | 1.可以模拟人脑的智能化处理；2.对大量非结构性、非精确性规律具有自适应功能；3 具有信息记忆、自主学习、知识推理和优化计算的特点 4 | 1.常规的神经网络要求样本数量较多，并存在局部最小、过学习以及隐层网络节点数选取缺乏理论指导等缺陷，限制了它的推广（预测）能力；2.神经网络的学习过程比较慢，对突发事 |

| | | | | |
|-------|-----|---------|---|--|
| | | | 在建模时都不需要计算统计特征 | 件的适应性差 |
| 专家系统法 | 中长期 | 适合于复杂系统 | 1.能汇集多个专家的知识 and 经验,最大限度地利用专家的能力; 2.占有的资料、信息多,考虑的因素也比较全面,有利于得出较为正确的结论 | 1.不具备自学习能力,受数据库里存放的知识总量的限制; 2.对突发性事件和不断变化的条件适应性差 |

3. 2 多源模型组合

3.2.1 按线性最优组合模型和线性非最优组合模型来看

通过文献梳理发现，线性非最优组合模型的预测精度不如最优组合预测模型。非最优正权组合预测方法是各个单一预测模型预测误差的方差和其权系数成反比的特征，给出组合预测的权系数的计算公式。因此，非最优正权组合预测方法目标函数值一般要劣于最优正权组合预测方法目标函数值。

3.2.2按线性组合模型和非线性组合模型来看

鉴于能源系统的特征，大量文献证明，线性组合预测法解决非线性问题的局限性较大、效果较差。它们预测效果产生差异的原因，主要有以下几点。

（1）从组合加权方式来看，采用线性加权的最优线性组合预测方法，对组合预测结果的影响既依赖于其单一模型预测值，又依赖于其相应的权重值，对于它们的权重值而言，取决于各种单一预测方法拟合误差的“贡献度”。也就是说，如果该方法在整个组合预测中的均方拟合误差较小，则说明该方法性质优良，因此赋予其较大的权重，反之，则赋予其较小的权重；采用非线性的加权方法（以ANN法为例），权重是按单个样本误差和总体误差满足给定精度要求，并由网络在动态学习中给出，与线性加权不同的是，某一种预测方法对组合预测结果的影响虽然与自身的预测结果和权重有关，但特殊之处是，其权重对组合预测结果的影响是非线性的，它既包括输入与隐单元之间的连接权，还包括输出与隐单元之间的连接权，在遇到某些难以处理的情况，比如指标间相关度较高、数据呈非线性变化、数据缺漏不全时仍可得到较满意的结果。因而更能适应能源预测系统的非线性特点。

（2）从建模方式和模型适应性看，线性组合更多地注重权重分配问题，很难找出对权重系数的最佳确定方式。线性最优组合预测方法采用数学规划问题模型，其求解比较繁琐，并且根据不同假设条件采用不同建模机理，这对预测模型提出了更苛刻的要求，进一步限制了组合预测的应用范围；以ANN为代表的非线性加权方法通过神经网络的训练过程直接从数据中发现规则和特征，实现任意复杂的函数映射，这种学习能力类似于“黑箱”，无需建立模型结构和设定参数估计过程，从而改善了传统组合模型中权重系数不易确定的状况。更为灵活的是，在没有输入模式先验信息的情况下，通过数据驱动就可以取得优良的结果，在整体网络结构不变的情况下，通过调整参数权值就可完成任意关系的学习，还可以通过补充训练来跟踪和适应外界环境的变化，因此对复杂的能源预测系统适应性较强。

（3）从个体样本失误对组合预测整体效果的影响来看，采用线性组合加权的最优线性组合预测方法，个别样本的较大误差会对整个组合预测产生较大的影响；以ANN为代表的非线性加权方法因具有无后效性的特点，即每一层神经元的状态只会影响下一层神经元的状态，不会产生累积不良影响，加上神经网络的神经元众多，且由多层组成，即使个别单元产生失误，其对整体的影响也能起到

缓冲作用。这种对小失误的宽容性，有效减少了组合预测结果对真实值的偏离。

(4) 从目标函数来看，最优组合预测方法追求的是在约束条件下总体误差平方和最小，目标最优解唯一，这种要求刚性很大；以ANN为代表的非线性加权方法不仅要求总体误差平方和最小，还对每个学习单元的拟合误差也有要求，因而能够有效保证总体误差平方和相对最小目标的实现，其目标最优解不唯一恰好成为一种优良特征，可视拟合精度的要求灵活调节，从而能更好地满足预测要求。

总之，从理论基础和模型构造来看，非线性组合预测方法总体预测效果优于线性组合预测方法，以神经网络模型为代表的非线性组合预测方法具有更高的预测精度和实用性、简便性，是一种解决非线性组合问题的有效工具。

3.2.3 按不变权组合预测方法和可变权组合预测方法来看

不变权组合预测方法是通过最优化规划模型或其他方法计算出各个单项预测方法在组合预测中的权系数。权系数只与预测方法有关，而与时间无关。也就是说同一种单项预测方法在各个时点的权系数是一样的。需要强调的是，各种预测方法对于不同的预测时间段表现出不同的预测能力，比如有的方法对瞬态变化敏感，那么适用于中短期预测；有的方法善于考察长期趋势，那就表现出优越的中长期预测能力。这时，如果将不同时间组合的权系数设定为常值，就无法各取所长获得最佳预报结果。在能源需求预测领域，除了非负可变加权系数的组合预测模型之外，所有的线性组合模型和神经网络非线性组合模型均属于不变权组合预测。

然而在预测实践中，不同的预测方法在不同时刻有不同的预测精度，也就是说同一种单项预测方法在不同的时刻表现并不是完全相同，有可能在某一点预测精度较高，但是另外一点预测精度则较低^[34-37]。因此，变权重组合预测更符合实际，具有较高的预测精度和预测稳定性，能比较合理地描述系统的客观现实。因此，可变权组合预测方法显然比不变权组合预测方法更科学。

此外，多源模型组合的预测方式面临一个共同的问题，均假设未来的能源消费变化趋势仍延续着历史的规律，并从历史消费数据中寻找能源需求的演变规律。但是，突发事件作为能源需求情景因素之一，势必引起能源消费的突变，从而出现“异常点”或转折性变化，各类组合模型对“异常点”或转折性变化的拟合，则显得无能为力。

3.3 智能信息融合

智能信息融合是预测模型中的一个新突破，并逐渐开始应用于能源预测相关的领域。其优势在于考虑到系统的复杂性和动态性，融合了专家系统，注入了专家系统的知识、信息，同时考虑了不规则事件的影响度，实现了人与预测技术的融合，不仅能克服模型方法的片面性，还能使得预测结果更准确，具有快速决断的优点。智能信息融合的客观性和科学性已得到广泛的认可。

但是，以往智能信息融合模型文献中，以电力负荷、石油价格为对象进行分散研究的居多，能源需求预测领域研究的较少，更缺乏将煤炭、石油、天然气三个重要能源资源集成为一个整体构建智能信息融合模型。此外，在电力负荷预测中，大多以神经网络单一模型作为预测技术，而在石油价格、外汇汇率等其他预测中，组合预测技术也只停留在不变权线性和不变权非线性技术层面。

4 进一步展望

从文献归纳和梳理中可以看出，能源需求预测方法逐渐向规范性、可行性及

智能化转变, 尽管国内外学者对能源需求预测问题进行了有益的探索和研究, 但在能源需求预测模型中可从以下两方面改进:

(1) 构建基于非线性变权重的能源需求组合预测模型

目前在能源需求预测领域, 已有线性、非线性组合预测模型中, 大多采用定权系数加以组合, 非负可变加权系数的非线性可变权组合预测模型较少。非线性组合模型中, 神经网络模型既有效地利用了神经网络的非线性映射能力, 又克服了通常对不同预测方法适用性强调不足的缺陷。因此, 关于非线性不变权的神经网络组合模型得到广泛的应用, 但是神经网络非线性可变权组合预测仍是空白。可变权组合预测方法尤其是非线性可变权方法比较复杂, 因此研究成果并不多见, 未来有进一步研究空间, 这也是本文能源需求预测考虑的重要研究方法。

(2) 构建基于多源信息融合的能源需求预测模型

在分析智能信息融合的基础上, 进一步拓展信息源, 构建基于多源信息融合的能源需求预测模型 (MIF 组合预测模型)。设想的思路是: 首先将非线性变权重的组合模型预测结果视为第一个信息源, 然后通过专家经验判断的方式, 量化各种突发事件或政策调整的影响, 诊断异常点对预测结果的影响, 将影响结果视为第二个信息源, 最后将多个预测机构的预测数据作为第三个信息源。在此多源信息基础上通过变权非线性融合技术赋予这三个信息源不同的权重, 从而得到最终的预测结果。该思路实质上也是一种智能信息融合系统, 当然, 模型的实用性和可行性既要考虑到研究对象的数据特征, 也要考虑到模型求解的具体技术方法。

参考文献

- [1] 徐国强, 胡清友. 统计预测与决策[Z]. 上海: 上海财经大学出版社, 1998.
- [2] Bates J M, Granger C W. The combination of forecasts[J]. OR, 1969:451-468.
- [3] 汪同三, 张涛. 组合预测: 理论, 方法及应用[M]. 社会科学文献出版社, 2008:
- [4] 卢奇, 顾培亮, 邱世明. 组合预测模型在我国能源消费系统中的建构及应用 | : | 系统工程理论与实践, 2003. 3: 24-30.
- [5] 王会强, 胡丹. 能源需求组合预测模型的构建和应用研究 | : | 中国能源, 2007. 29(8): 38-40.
- [6] 高军, 宋书贵, 吴鹏等. 天然气消费量组合预测建模及应用 | : | 西南石油大学学报 (社会科学版), 2012. 14(1): 6-10.
- [7] 张玉春, 郭宁, 任剑翔. 基于组合模型的甘肃省能源需求预测研究 | : | 生产力研究, 2012(11): 31-33.
- [8] 秦晋栋. 基于熵权灰色组合预测模型的区域能源需求预测研究 | : | 价值工程, 2012. 4: 289-291.
- [9] 陈玉金, 刘建永, 李凌等. 基于神经网络回归分析组合模型的能源消耗预测研究 | : | 兵工自动化, 2008. 11: 59-60.
- [10] 侯运炳, 张文, 张浩等. 煤炭需求最优组合预测 | : | 煤炭工程, 2008(11): 75-78.
- [11] 宋雅晴, 杨桂元. 我国能源消费的组合预测 | : | 市场经济与价格, 2011(7): 34-37.
- [12] 曹飞. 陕西省能源消费的组合预测 | : | 环境科学与技术, 2013. 36(005): 147-150.
- [13] Diebold F X, Lopez J A. Forecast evaluation and combination[Z]. National Bureau of Economic Research Cambridge, Mass., USA, 1996.
- [14] 周剑. 辽宁省中长期能源供需预测与可持续发展研究[D]. 长春: 吉林大学综合信息矿产预

测研究所, 2010.

- [15] 赵国浩, 郭淑芬, 申屠菁等. 基于变权组合预测模型探讨中国煤炭需求 | : | 资源与生态学报 (英文版), 2011. 2(2).
- [16] 索瑞霞, 王福林. 组合预测模型在能源消费预测中的应用 | : | 数学的实践与认识, 2010(018): 80-85.
- [17] 韩天锡, 王彦霞. 多变量非线性组合模型在石油消费预测中的应用 | : | 天津理工大学学报, 2006. 22(4): 69-71.
- [18] Donaldson R G, Kamstra M. Forecast combining with neural networks[J]. Journal of Forecasting, 1996,15(1):49-61.
- [19] 焦李成. 神经网络系统导论[Z]. 西安: 西安电子科技大学出版社, 1992.
- [20] Irie B, Miyake S. Capabilities of three-layered perceptrons: Neural Networks, 1988., IEEE International Conference on, 1988[C]. IEEE.
- [21] Kankal M, Akpınar A, Kömürçü M İt al. Modeling and forecasting of Turkey' s energy consumption using socio-economic and demographic variables[J]. Applied Energy, 2011,88(5):1927-1939.
- [22] Pao H. Comparing linear and nonlinear forecasts for Taiwan's electricity consumption[J]. Energy, 2006,31(12):2129-2141.
- [23] 周扬, 吴文祥, 胡莹等. 基于组合模型的能源需求预测 | : | 中国人口. 资源与环境, 2010. 20(4): 63-68.
- [24] 汪克亮, 杨力. 电力需求的非线性回归组合神经网络预测研究 | : | 计算机工程与应用, 2010. 46(028): 225-227.
- [25] 师亮云, 刘家军, 王明军等. 基于诱导有序加权调和平均算子和马尔科夫链的电量组合预测模型的研究 | : | 电网与清洁能源, 2010(009): 38-43.
- [26] Meadows D L, Meadows D H, Zahn Eet al. Die Grenzen des Wachstums: Bericht des Club of Rome zur Lage der Menschheit[M]. Deutsche Verlags-Anstalt Stuttgart, 1972.
- [27] 倪军, 杨明志, 杨期余. 专家系统技术应用于城网负荷预测 | : | 供用电, 1994. 5: 7.
- [28] 王德金. 基于熵权法的电力中长期负荷组合预测 [D][D]. 华北电力大学 (北京), 2007.
- [29] 于希宁, 牛成林, 李建强. 基于决策树和专家系统的短期电力负荷预测系统 | : | 华北电力大学学报, 2005. 32(5): 57-61.
- [30] 夏道止, 王锡凡, 周琳等. 基于 RBF 神经网络和专家系统的短期负荷预测方法 | : | 西安交通大学学报, 2001. 35(4): 331-334.
- [31] 招海丹, 余得伟. 电力负荷短期预测的模糊专家系统修正方法 | : | 广东电力, 2001. 14(1): 1-4.
- [32] Kim K, Park J, Hwang Ket al. Implementation of hybrid short-term load forecasting system using artificial neural networks and fuzzy expert systems[J]. Power Systems, IEEE Transactions on, 1995,10(3):1534-1539.
- [33] 余乐安, 汪寿阳, 黎建强. 外汇汇率与国际原油价格波动预测—TEI@ I 方法论[M]. 湖南大学出版社, 2006:
- [34] 陈华友, 刘春林. 基于 IOWA 算子的组合预测方法 [J] | : | 预测, 2003. 22(6): 61-65.
- [35] 陈华友, 程蕾, 张倩. 基于相关系数的 IOWA 算子最优组合预测模型 | : | 第四届中国不确定系统年会论文集, 2006: 235-241.
- [36] 胡彦, 李秀美, 陈华友. 基于 IOWA 算子的税收组合预测模型 | : | 统计与决策, 2009. 10: 33-35.
- [37] 周礼刚, 陈华友, 丁子千等. 基于 Theil 不等系数的 IOWGA 算子组合预测模型 | : | 安

徽大学学报: 自然科学版, 2010. 1: 1-6.