

На правах рукописи



**Ван Ефэн**

**МЕТОДОЛОГИЯ СОВЕРШЕНСТВОВАНИЯ СИСТЕМ КРАТКОСРОЧНОГО  
ПРОГНОЗИРОВАНИЯ ЭЛЕКТРОПОТРЕБЛЕНИЯ**

Специальность: 05.13.06 – Автоматизация и управление технологическими процессами и производствами (промышленность)

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата технических наук

Санкт - Петербург – 2013

Работа выполнена в Санкт-Петербургском государственном электротехническом университете «ЛЭТИ» им. В.И.Ульянова (Ленина), на кафедре систем автоматического управления

Научный руководитель:

Приходько Ирина Аркадьевна  
кандидат технических наук, доцент  
Санкт-Петербургский государственный  
электротехнический университет «ЛЭТИ» им.  
В.И. Ульянова (Ленина) (СПбГЭТУ), доцент  
кафедры систем автоматического управления

Официальные оппоненты:

Ефимов Владимир Васильевич  
доктор технических наук, профессор,  
заместитель директора по научной работе  
ОАО "Авангард", Санкт-Петербург

Южаков Александр Анатольевич  
доктор технических наук, профессор  
Пермский национальный исследовательский  
политехнический университет, заведующий  
кафедрой автоматизации и телемеханики

Ведущая организация:

Федеральное государственное бюджетное  
образовательное учреждение высшего  
профессионального образования «Санкт-  
Петербургский национальный  
исследовательский университет  
информационных технологий, механики и  
оптики»

---

Защита состоится «24» июня 2013 г. в 14 часов на заседании диссертационного совета Д 212.238.07 при Санкт-Петербургском государственном электротехническом университете «ЛЭТИ» им. В.И.Ульянова (Ленина) по адресу: 197376, Санкт-Петербург, ул. Проф. Попова, д. 5.

С диссертацией можно ознакомиться в библиотеке университета

Автореферат разослан «  » \_\_\_\_\_ 2013 года.

Ученый секретарь  
диссертационного совета



В.В. Цехановский

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность работы.** Электроэнергетика является ключевой отраслью российской экономики и многих стран мира. К 2020 г. потребления электроэнергии в России должно увеличиться на 25%. В условиях роста электропотребления необходимо повысить показатель надежности и коэффициент полезного действия электростанций; снизить долю износа основного оборудования и потери в электросетях.

Основную информацию для планирования развития электроэнергетических систем составляют прогнозные оценки электропотребления. Нагрузка электроэнергетической системы формируется комплексом потребителей электроэнергии, состоящих из промышленных предприятий, сельского и жилищно-коммунального хозяйств. В зависимости от времени упреждения прогнозы делятся на прогнозирование долгосрочное, среднесрочное, краткосрочное и оперативное. При оперативном управлении режимами используется оперативный прогноз графика потребления на следующие час и оставшиеся до конца суток часы.

Прогнозирование потребления электроэнергии является на сегодняшний день одной из важных научных и практических задач в электроэнергетике. Повышение точности прогнозирования обеспечивает экономию энергетических ресурсов, определяет эффективность управления электроснабжением и соответствующее увеличение прибыли энергетических предприятий.

Прогнозы электропотребления закладываются в инвестпрограммы энергокомпаний. По оценкам российских специалистов, каждая оплошность в ежегодном прогнозе электропотребления на 1% – это 4 млрд. долларов дополнительных инвестиций на возведение генерирующих мощностей<sup>1</sup>. Согласно оценкам экономической эффективности, приведенным в зарубежных публикациях, для типичной региональной энергокомпании Северной Америки улучшение качества прогноза на 1% приводит к снижению издержек на десятки миллионов долларов в год.

Развитие рыночных отношений в управлении электроэнергетикой привело к появлению противоречия между рыночными конкурентными механизмами функционирования отдельных подсистем электроэнергетических систем (ЭЭС) и требованиями обеспечения надежности и качества работы ЭЭС, как единого комплекса.

Электрическая энергия в себестоимости продукции крупных предприятий, например в металлургическом производстве, составляет 11% – 30%. Одним из основных путей снижения затрат на электрическую энергию является выход предприятий в качестве участников на оптовый рынок электрической энергии.

Таким образом, в условиях современного энергорынка для крупных предприятий и добывающих производств важное значение имеет создание системы почасового потребления электрической энергии, позволяющей минимизировать отклонения потребляемой мощности от заявленной.

Решению различных аспектов задачи прогнозирования электропотребления посвящены работы Бердина А.С., Воропая Н.И., Гамма А.З., Глазуновой А.М., Готман Н.Э., Колосок И.Н., Макоклюева Б.И., Манова Н.А., Манусова В.З., Старцевой Т.Б., Чукреева Ю.Я., Шумиловой Г.П. и др.

---

<sup>1</sup> Рудаков Е., Саакян Ю., Нигматулин Б., Прохорова Н. Цена расточительности // «Эксперт» №24 / июня 2008

### **Цель и задачи исследований**

Целью диссертационной работы является увеличение точности предсказания предполагаемого потребления электроэнергии с целью оптимального использования энергоресурсов, минимизации стоимости и повышения надежности электроснабжения.

Для достижения поставленной цели в диссертационной работе решены следующие задачи:

1. Разработка нейросетевой модели прогнозирования электропотребления, включающей нечеткую обработку данных обучающей выборки, индикацию времени суток, оптимизацию числа нейронов в слоях и переобучение сети на основе текущих фактических значений электропотребления.
2. Разработка системы прогнозирования электропотребления на основе метода опорных векторов.
3. Определение оптимальных параметров ядерной функции метода опорных векторов с использованием генетического алгоритма и алгоритма роя частиц.
4. Сравнительное исследование эффективности интеллектуальных моделей прогнозирования, а также моделей на основе метода опорных векторов.
5. Разработка программной реализации предложенных систем прогнозирования, построение графических интерфейсов пользователя.

### **Методы исследования**

Для решения поставленных задач были использованы: методы математического моделирования, обработки данных и прогнозирования временных рядов; статистический и регрессионный анализ; теория нечетких множеств; теория искусственных нейронных сетей; теория гибридных сетей; метод опорных векторов; метод роя частиц; математические пакеты MatLab (*Neural Networks Toolbox*, *Statistics Toolbox*, *Fuzzy Logic Toolbox*), Orange.

### **Научные результаты, выносимые на защиту:**

1. Способ формирования обучающей выборки с предварительной обработкой исходных данных.
2. Нейросетевая модель прогнозирования электрической нагрузки с оптимизацией числа нейронов в слоях и переобучением сети на основе текущих фактических значений электропотребления.
3. Модель прогнозирования электрической нагрузки на основе метода опорных векторов с оптимизацией параметров ядерной функции.
4. Программные модули прогнозирования и два пользовательских интерфейса в пакете MatLab.

### **Новизна научных результатов**

1. Способ формирования обучающей выборки, отличающийся тем, что для улучшения точности прогноза введены известные значения нагрузки в прошлые дни *в прогнозируемый час*, индикация времени суток, а также выполнена предварительная обработка данных временного ряда методом главных компонент с добавлением шума (для нейронной сети).

2. Нейросетевая модель прогнозирования электрической нагрузки отличается *способом оптимизации* (число нейронов скрытых слоев определяется на основе

генетических алгоритмов) и наличием *процедуры переобучения* в случае превышения ошибки прогноза.

3. Модель прогнозирования электрической нагрузки на основе метода опорных векторов, отличающаяся тем, что для повышения эффективности использования метода опорных векторов параметры регрессионной модели определяются на основе решения задачи квадратичного программирования с оптимизацией параметров ядерной функции на основе *генетического алгоритма и алгоритма роя частиц*.

4. Программные модули и два пользовательских интерфейса в пакете Matlab, включающие программы формирования обучающей выборки и подбора оптимальных параметров ядерных функций на основе генетических алгоритмов и метода роя частиц, обеспечивающие уменьшение ошибки прогноза, а также позволяющие использовать систему прогнозирования в качестве системы поддержки принятия решений.

### **Достоверность научных положений, результатов и выводов**

Достоверность научных положений и результатов, сформулированных в диссертации, подтверждается корректным использованием методов математического моделирования и прогнозирования временных рядов, статистического и регрессионного анализа, теории нечетких множеств, теории искусственных нейронных сетей, методов опорных векторов и главных компонент, а также эволюционных алгоритмов оптимизации.

### **Практическая ценность работы**

1. Нейросетевая модель прогнозирования позволяет учитывать дополнительные факторы и обеспечивает уменьшение ошибки прогноза на 2,7 %. Точность краткосрочного прогноза электропотребления можно улучшить: при учете дополнительной информации о времени суток на 1%; при выполнении дополнительной нечеткой обработке информации на 0,43% и при нормировании обучающей выборки на 0,3%.

2. Прогнозирующие нейросетевые модели и модели на основе опорных векторов наиболее эффективны в случаях сильно колебательной нагрузки. Точность прогнозирования при использовании нейросетевой модели по сравнению с регрессионной можно увеличить до 2 раз, при использовании модели на основе опорных векторов – до 2,5 раза. Лучшие результаты получены при использовании гауссовой радиальной базисной ядерной функции.

3. На основе полученных моделей разработаны системы прогнозирования электропотребления в виде пользовательских интерфейсов для поддержки принятия решений.

Практическая значимость подтверждается актом использования в учебном процессе СПбГЭТУ «ЛЭТИ».

### **Апробация результатов работы**

Основные теоретические и прикладные результаты диссертационной работы докладывались и получили одобрение на международном молодежном форуме «Энергоэффективные электротехнологии» (сентябрь 2011, г. Санкт-Петербург); II Международной заочной научно-практической конф. «Актуальные проблемы науки» (УСОМ–2011), (27 сентября 2011 г., Тамбов. 2011); XIII международной конференции по

мягким вычислениям и измерениям (июнь 2010, г. Санкт–Петербург); «Управление в технических, эргатических, организационных и сетевых системах» (УТЭОСС–2012, 9 –11 октября 2012 г., Санкт–Петербург), на 63 – 66 научно-технических конференциях профессорско-преподавательского состава ЭТУ (СПбГЭТУ «ЛЭТИ», февраль 2010 – 2013 гг.).

**Публикации.** Основные теоретические и практические результаты диссертации опубликованы в 7 научных работах, в том числе 3 статьи в рецензируемых и входящих в перечень ВАК и 4 публикации в материалах международных и всероссийских научно-технических конференций.

**Структура и объем работы.** Диссертационная работа состоит из введения, четырех глав с выводами, заключения, списка литературы. Основная часть работы изложена на 135 страницах машинописного текста. Работа содержит 35 рисунков и 12 таблиц, список литературы содержит 135 наименований.

### **КРАТКОЕ СОДЕРЖАНИЕ РАБОТЫ**

**Во введении** обоснована актуальность темы диссертационной работы, сформулированы цели и задачи исследований, перечислены основные научные результаты диссертации. Кратко изложено содержание диссертации по главам.

**В первой главе** представлен обзор и краткий критический анализ существующих методов прогнозирования. Рассмотрено современное состояние проблемы прогнозирования электропотребления на основе основных российских и зарубежных публикаций. В результате можно сделать обобщение: современный этап развития энергосистем характеризуется повышением требований к качеству прогнозирования потребления электроэнергии в условиях существенного ограничения на выделяемые энергоресурсы.

Поэтому актуальной становится проблема разработки и внедрения методов и технологий прогнозирования, нечувствительных к недостатку априорной информации. Наряду с методами статистического анализа и сезонных кривых многие разработчики (зарубежных – больше) используют технологии нейронных сетей.

Нейросетевые системы искусственного интеллекта, которые будучи универсальными аппроксиматорами, потенциально гарантируют возможность успешного прогнозирования. Применение нейронных сетей не накладывает ограничений на характер входной информации и позволяет успешно решать задачи, в которых затруднено или невозможно нахождение аналитических зависимостей между входными и выходными данными. Основные преимущества интеллектуальных систем прогнозирования: способность получения правильного решения задачи при наличии неполных и искаженных данных после настройки параметров (обучения) сети; возможность учета большого количества дополнительных факторов, влияющих на качество прогнозирования. Недостаток интеллектуальных систем – необходимость обучения, сложность поиска неисправности.

В последние годы при использовании метода опорных векторов новое развитие получил способ прогнозирования на основе регрессионных моделей, альтернативный интеллектуальному подходу. Характерным для метода опорных векторов является то, что параметры регрессионной модели определяются на основе решения задачи квадратичного программирования, имеющей единственное решение.

В результате утверждается, что совершенствование прогнозирующих систем электропотребления можно обеспечить путем оптимизации параметров алгоритмов.

**Во второй главе** изложено применение нейросетевого метода для построения модели прогнозирования краткосрочного прогноза электрической нагрузки.

Задача прогнозирования формулируется следующим образом: зная предыдущие  $q$  значений временного ряда, предсказать последующее значение  $x(t+d)$  :

$$\text{дано } x(t-q+1), \dots, x(t-1), x(t), \eta_1, \eta_2, \dots, \eta_g$$

$$\text{найти } x(t+d)$$

где  $d$  – шаг прогнозирования,  $g$  – количество независимых переменных  $\eta$ .

Для формирования обучающей последовательности временной ряд разбивается на окна шириной  $q$

$$\left\{ \begin{array}{l} x(t-q+1), \dots, x(t-1), x(t) \\ x(t-q+2), \dots, x(t), x(t+1) \\ \dots \\ x(t-q+1+j), \dots, x(t-1+j), x(t+j) \end{array} \right\},$$

Обучающая последовательность при этом выглядит следующим образом:

$$x(t-q+1+j), \dots, x(t-1+j), x(t+j) \rightarrow x(t+d+j).$$

Обозначим  $x = [x(t-q+1+j), \dots, x(t-1+j), x(t+j)]^T$  – вектор входа;  $y^{(k)} = x(t+j+d)$  – выход сети (выход  $k$ -го слоя),  $k$  – номер слоя. Значение прогноза  $x(t+j+d)$  получается на выходе построенной нейросетевой модели при подаче на вход вектора фактических данных  $[x(t-q+1+j), \dots, x(t-1+j), x(t+j)]^T$ .

В обучающую выборку **вводятся** дополнительные значения нагрузки в прошлые дни  $D$ , в прогнозируемый час  $t+d$

$$\left\{ \begin{array}{l} x(D, t-q+1), \dots, x(D, t), x(D-1, t+d), \dots, x(D-v, t+d) \\ x(D, t-q+2), \dots, x(D, t-1), x(D-1, t+d+1), \dots, x(D-v, t+d+1) \\ \dots \\ x(D, t-q+1+j), \dots, x(D, t-1+j), x(D-1, t+d+j), \dots, x(D-v, t+d+j) \end{array} \right\},$$

*Введение дополнительных признаков.* Из рассмотрения процесса прогнозирования с позиций теории распознавания образов (метод обобщения по преобладанию), можно утверждать, что введение дополнительных признаков способствует повышению точности прогноза.

*Эффективность нечеткой обработки информации.* Предлагается использовать в нейросетевой модели на этапе подготовки обучающей выборки элементы нечеткого подхода, учитывая дополнительную информацию о принадлежности исходных данных к нечетким множествам.

Оценивается эффективность двух способов задания значений функции принадлежности. В первом случае, каждому четкому значению входной переменной  $u_0$  ставится в соответствие значение функции принадлежности. Введение дополнительной информации увеличивает число входов в два раза.

Во втором случае, более эффективном, формируется один дополнительный входной вектор, содержащий значения функции принадлежности только для данных

текущего часа. Лучший результат получен при использовании функции принадлежности  $s$ -типа

$$f(x; a, b) = \begin{cases} 0, & x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2, & a \leq x \leq \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2, & \frac{a+b}{2} \leq x \leq b \\ 1 & x \geq b \end{cases}.$$

*Влияние шума (искажений данных) в обучающей и тестовой выборках на точность прогноза.* Подбор среднеквадратичного отклонения шума, при котором действительно можно повысить качество обобщения, представляет собой самостоятельную задачу. Ее теоретическое решение весьма сложно, однако можно получить экспериментальную оценку. По результатам многочисленных тестов получено, что среднеквадратическое отклонение шума должно составлять значение порядка 0,05%. (Используется в нейросетевой модели).

*Влияние температуры на электропотребление.* Известно, что электропотребление растет в холодные и жаркие дни. Однако при краткосрочном прогнозировании (сутки вперед) учет температуры не приводит к улучшению точности прогноза.

*Нормирование обучающей выборки* выполняет предварительную обработку данных для обучения сети посредством нормирования входов и выходов так, чтобы они находились в установленном диапазоне  $[0, 1]$ .

$$\tilde{x}_i = \frac{x_i - x_{i,\min}}{x_{i,\max} - x_{i,\min}}.$$

Структура нейронной сети прогнозирования электропотребления приведена на рис.1.

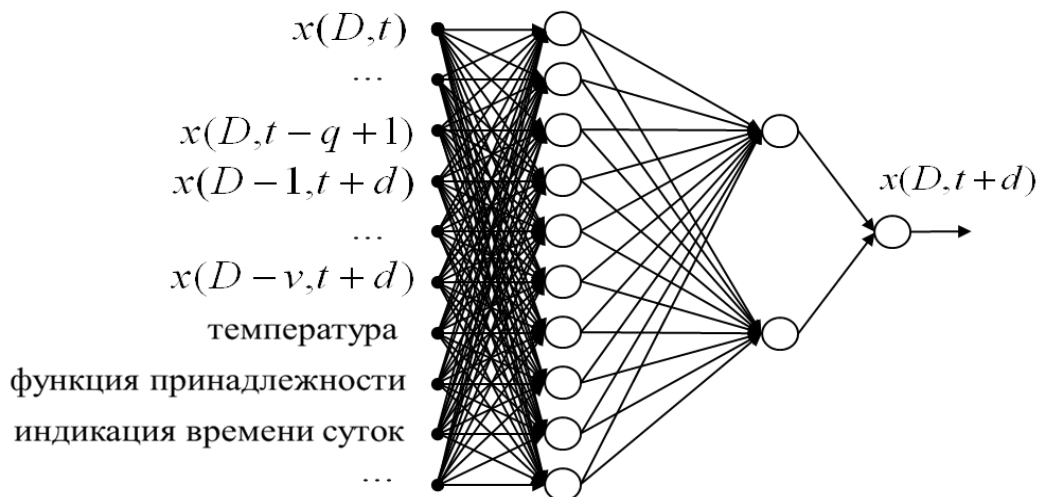


Рис.1

Для улучшения точности прогнозирования электропотребления дополнительно использован метод главных компонент. Преобразование по методу главных



компонент позволяет заменить большое количество информации, основанной на взаимно коррелирующих входных данных, множеством статистически независимых компонентов с учетом их важности.

Для ковариационной матрицы входных данных  $\mathbf{x}$ ,  $\lambda_1 > \lambda_2, \dots, > \lambda_q \geq 0$  – собственные числа,  $u_1, u_2, \dots, u_q$  – соответствующие им собственные векторы,  $q$  – размерность входного вектора,  $p$  – число образов. Метод главных компонент состоит в нахождении линейных комбинаций исходных переменных:

$$y_i = u_{i1}x_1 + u_{i2}x_2 + \dots + u_{iq}x_q, \quad i = \overline{1, s}, \quad s < q,$$

Переменные  $y_i$  упорядочены по возрастанию дисперсии, большую часть общей дисперсии характеризует подмножество  $s$  первых переменных. Относительные доли дисперсии исходных данных, определяются по формуле:  $\lambda_i \cdot 100\% / \sum_{i=1}^q \lambda_i$ .

Для выбора числа главных компонент использован график зависимости собственных значений  $\lambda_i$  от числа главных компонент (рис. 2), принято  $s = 6$ .

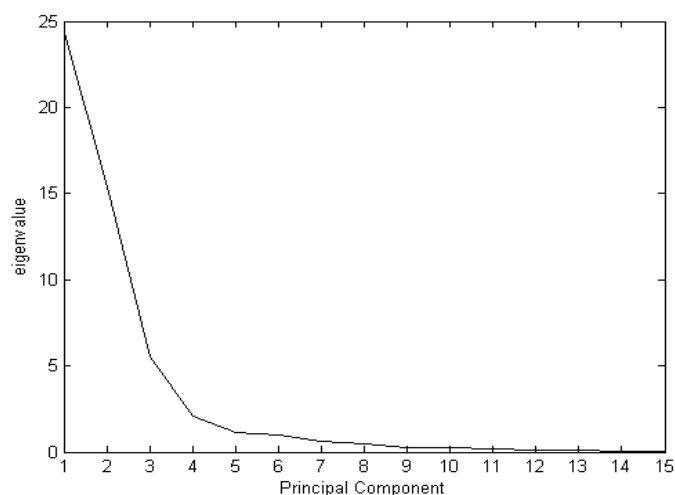


Рис.2

Для улучшения эффективности прогнозирования рассмотрена обучающая выборка с учетом предварительной обработки данных и нормирования.

Нейросетевой подход прогнозирования. Процесс прогнозирования на основе нейросетевого подхода состоит из следующих этапов: выбор обучающих данных и структуры входных векторов; подбор архитектуры сети; обучение сети; тестирование сети на контрольном множестве данных и при необходимости – дообучение; использование сети для прогнозирования.

В исследовании предложено число нейронов в скрытых слоях определять на основе генетических алгоритмов оптимизации, использован Genetic Algorithm Toolbox пакета MatLab.

Точность прогнозирования оценивается как средняя абсолютная ошибка в процентах (MAPE):

$$MAPE = \frac{1}{P} \sum_{i=1}^P \frac{|y_i - \hat{y}_i|}{y_i} \times 100,$$

где  $y_i$  - фактическое значение нагрузки,  $\hat{y}_i$  - результат прогноза,  $p$  – количество данных. Из анализа публикаций, приемлемой является ошибка прогноза не более 5%.

Алгоритм переобучения параметров нейронной сети предполагает выполнение следующих шагов:

*Шаг 1.* Формирование обучающих и тестовых данных.

*Шаг 2.* Выбор параметров временного окна и структуры сети.

*Шаг 3.* Обучение нейронной сети.

*Шаг 4.* Добавление текущего значения нагрузки  $x_k$  в данные тестирования.

*Шаг 5.* Тестирование нейронных сетей и получение прогноза  $\hat{x}_{k+1}$ .

*Шаг 6.* Если качество прогноза удовлетворительное – переход к пункту 7, иначе переход к пункту 3.

*Шаг 7.* Удаление  $x_1$  из обучающей выборки и получение обучающих выборок  $x_i$  ( $i = 1, 2, \dots, n-1$ ).

*Шаг 8.*  $x_n = x_{k+1}$  и формирование новых обучающих выборок.

*Шаг 9.*  $k = k + 1$ , переход к пункту 4.

**В третьей главе** рассмотрено использование метода опорных векторов (МОВ) для прогнозирования электрической нагрузки, производится оптимизация параметров МОВ.

Метод опорных векторов – набор алгоритмов вида «обучение с учителем», использующихся для задач классификации и регрессионного анализа.

Преимуществом метода опорных векторов, является то, что параметры регрессионной модели определяются на основе решения задачи квадратичного программирования.

Задачей построения уравнения линейной регрессии является оценка неизвестной вещественной функции

$$y = f(\mathbf{x}) + e, \quad (1)$$

где  $f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + w_0$ , вектор  $\mathbf{w} = (w_1, \dots, w_q) \in R^q$  и смещение  $w_0 \in R$  – параметры уравнения,  $e$  – ошибка аппроксимации; для задачи (1)  $y = x(t+d)$ .

В методе опорных векторов задача построения нелинейной регрессии в исходном пространстве  $R^q$  рассматривается как задача построения линейной регрессии в некотором расширенном пространстве признаков  $H$ , порождаемом нелинейным отображением  $\phi: R^q \rightarrow H$

$$f(\mathbf{x}) = \langle \mathbf{w}, \phi(\mathbf{x}) \rangle + w_0.$$

Задача нахождения параметров (поиска оптимальной разделяющей поверхности) сводится к задаче квадратичной оптимизации и формулируется в виде минимизации функционала:

$$\min_{\mathbf{w}, w_0, \xi, \xi^*} \left[ \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^p (\xi_i + \xi_i^*) \right], \quad (2)$$

при ограничениях

$$y_i - \mathbf{w}^T x_i - w_0 \leq \varepsilon + \xi_i, \quad \mathbf{w}^T x_i + w_0 - y_i \leq \varepsilon + \xi_i^*, \quad \xi_i \geq 0, \quad \xi_i^* \geq 0, \quad i = \overline{1, p}$$

где параметр  $C$  – положительная константа; чем больше значение  $C$ , тем выше штраф ошибки,  $\xi_i$  – параметры для обработки неразделимых наблюдений (входов),

$\varepsilon$  – функция потерь, использована кусочно-линейной функция. Штрафное слагаемое в функционале  $\frac{1}{2}\|\mathbf{w}\|^2$  вводится согласно принципу регуляризации и означает, что среди всех векторов  $\mathbf{w}$ , минимизирующих функционал (2), наиболее предпочтительны векторы с минимальной нормой. Второе слагаемое функционала штрафует любые отклонения  $f(x)$  от  $y$  большие, чем  $\varepsilon$  для всех обучающих данных.

При использовании метода опорных векторов в работе решается двойственная к данной задача:

$$f(x) = \sum_{i=1}^p (\alpha_i - \alpha_i^*) K(\mathbf{x}', \mathbf{x}'') + w_0$$

где двойственные переменные удовлетворяют условию  $\alpha_i, \alpha_i^* \in [0, C]$ ,  $K(\mathbf{x}', \mathbf{x}'')$  – ядерная функция, если она представима в виде  $K(\mathbf{x}', \mathbf{x}'') = \langle \phi(\mathbf{x}'), \phi(\mathbf{x}'') \rangle$  при некотором отображении  $\phi: R^q \rightarrow H$ , где  $H$  – пространство со скалярным произведением.

Вместо требования максимизации величины «зазора» между распознаваемыми классами для задач распознавания в случае задач регрессии выдвигается требование минимизации вариации прогнозирующей функции на области задания переменных, которая формализуется как задача максимизации параметра  $\delta_\varepsilon = \inf_{(\mathbf{x}', \mathbf{x}'' \in X_\varepsilon)} (|\mathbf{x}' - \mathbf{x}''|)$ ,

$X_\varepsilon = \{(\mathbf{x}', \mathbf{x}'') \in X \times X \mid |f(\mathbf{x}') - f(\mathbf{x}'')| \geq 2\varepsilon\}$ , где  $\varepsilon$  – пороговый параметр,  $\mathbf{x}', \mathbf{x}''$  – значения тестовой выборки  $X_\varepsilon$ .

Для линейной регрессии  $f(\mathbf{x}) = \langle \mathbf{w}, \mathbf{x} \rangle + w_0$  можно получить  $\delta_\varepsilon = \frac{2\varepsilon}{|\mathbf{w}|}$ . Задача

максимизации  $\delta_\varepsilon = \frac{2\varepsilon}{|\mathbf{w}|}$  полностью эквивалентна задаче минимизации  $\delta_\varepsilon = \frac{1}{2} \sum_{i=1}^q w_i^2$ .

Откуда следует задача квадратичного программирования аналогичная (2).

В исследованиях моделированием были использованы следующие ядерные функции:

– полиномиальная  $K(\mathbf{x}', \mathbf{x}'') = (\langle \mathbf{x}', \mathbf{x}'' \rangle + 1)^d$ ;

– гауссова радиальная базисная функция  $K(\mathbf{x}', \mathbf{x}'') = \exp(-\frac{\|\mathbf{x}' - \mathbf{x}''\|^2}{2\sigma^2})$ ;

– экспоненциальная базисная функция  $K(\mathbf{x}', \mathbf{x}'') = \exp(-\frac{\|\mathbf{x}' - \mathbf{x}''\|}{2\sigma^2})$ ;

Ограничение использования метода опорных векторов связано с отсутствием рекомендаций по выбору параметров ядерной функции, наиболее подходящих для решения конкретной задачи.

Подбор управляющего параметра алгоритма  $C$  (отношение между эмпирическим риском и сложностью модели) может повлиять на способность к обобщению модели.

Для определения эффективных параметров алгоритма использованы генетический оптимизационный алгоритм и алгоритм роя частиц.

Алгоритм на основе метода роя частиц

Метод является стохастическим, не требующим вычисления градиента, может применяться для оптимизации параметров метода опорных векторов. В качестве оптимизируемой функции использована ошибка прогнозирования.

Алгоритм осуществляет передвижение частиц в пространстве поиска. Частицы передвигаются с определенной скоростью до тех пор, пока не будет найдена позиция глобального оптимума.

Расчет вектора скорости  $i$ -ой частицы на  $t+1$  шаге вычисляется в соответствии со следующими формулами:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (p_{bi}(t) - x_i(t)) + c_2 \cdot r_2 \cdot (p_{gi}(t) - x_i(t)),$$

$$x_i(t+1) = x_i(t) + \beta \cdot v_i(t+1),$$

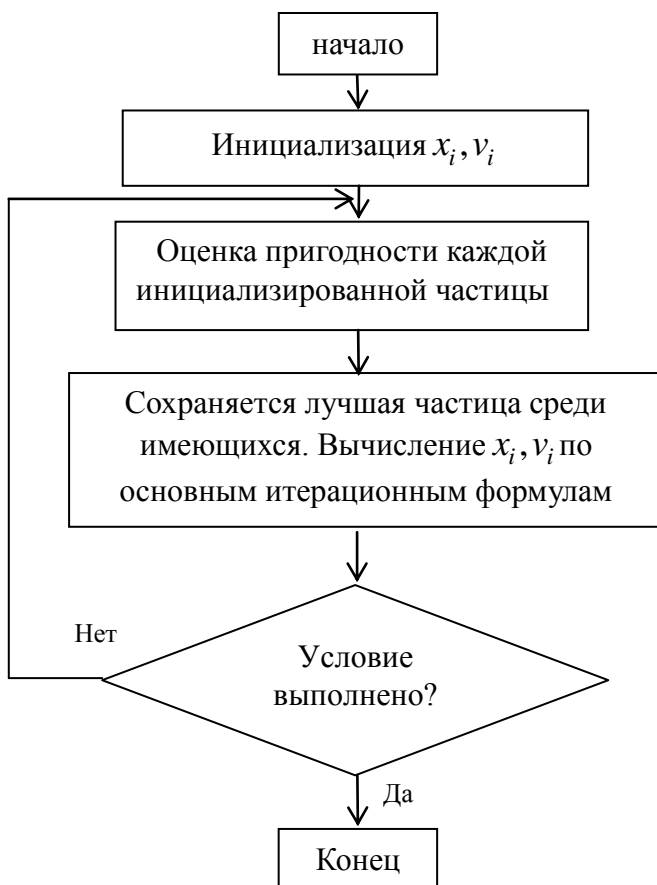


Рис.3

где  $w$  – инерционный вес и является константой,  $v_i t$  – текущая скорость  $i$ -ой частицы в момент  $t$ . Постоянный множитель  $c_1$  является константой, характеризует ускорение с положительными значениями (локальная весовая доля), задано 1,5. Множители  $r_1, r_2$  являются случайными переменными в диапазоне  $[0, 1]$ . Блок-схема алгоритма показана на рис.3. Векторная величина  $p_{bi}(t)$  – это лучшая позиция  $i$ -ой частицы, а векторная величина  $x_i t$  – текущая позиция  $i$ -ой частицы.

$$p_{bi}(t) = \begin{cases} p_{bi}(t-1) & f(x_i(t)) \geq f(p_{bi}(t-1)) \\ x_i(t) & f(x_i(t)) \leq f(p_{bi}(t-1)) \end{cases}$$

Множитель  $c_2$  является константой – глобальная весовая доля;  $p_{gi}$  – глобальная лучшая позиция всего роя частиц,  $\beta$  – фактор ограничения, используемый для контроля веса скорости.

Генетический алгоритм генерирует последовательность новых векторов, содержащих оптимизированные переменные.

Зависимости ошибки от параметра алгоритма  $C$  и параметра ядерной функции  $\sigma$  приведены на рис.4 и рис.5.

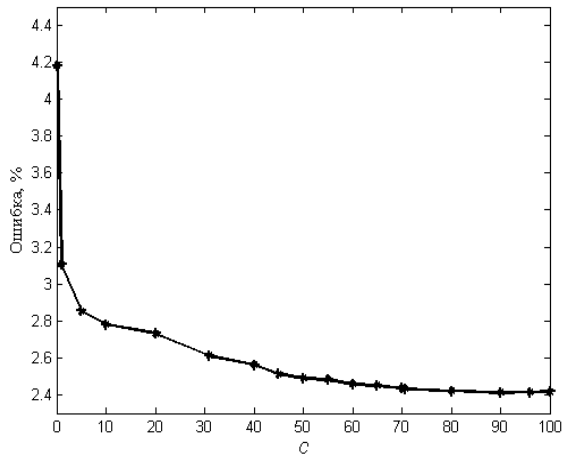


Рис. 4

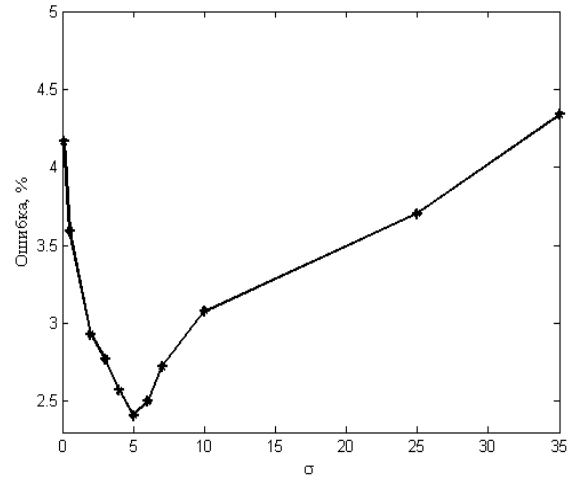


Рис. 5

В четвертой главе исследованы модели краткосрочного прогнозирования электропотребления, приведены результаты оценки точности системы, сравнительный анализ эффективности прогнозирования построенных систем, а также построены графические интерфейсы пользователя.

Обучающая выборка представляет собой почасовые данные электропотребления за все рабочие дни летнего периода 2009 года города Сюй Джоу (КНР). На рис. 6 показан график почасовых данных электропотребления.

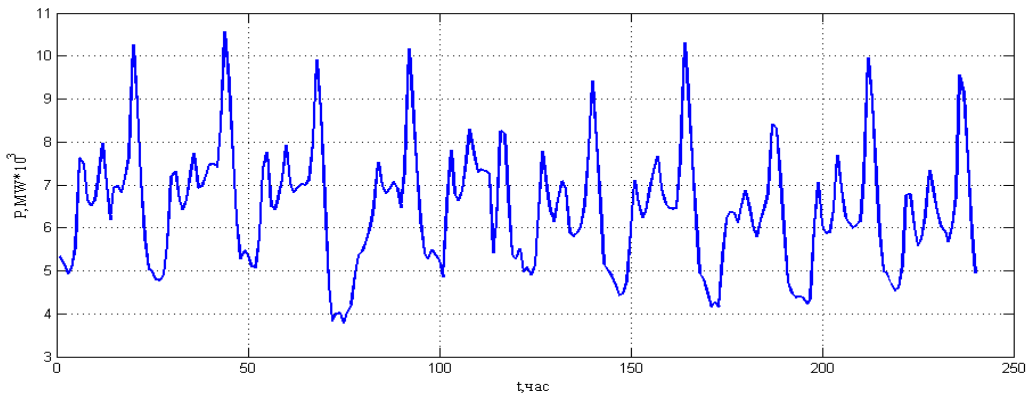


Рис. 6

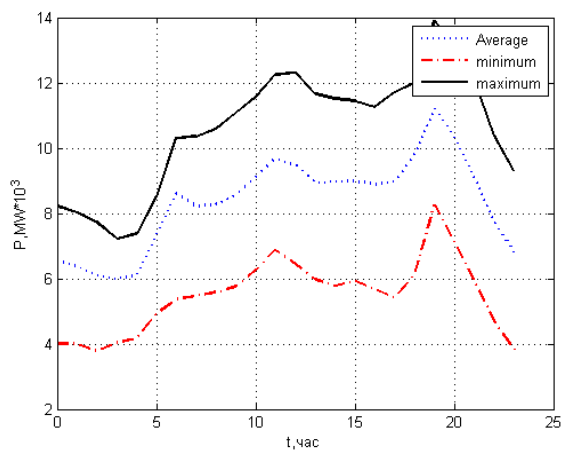


Рис. 7. Максимальные, среднечасовые и минимальные значения нагрузки

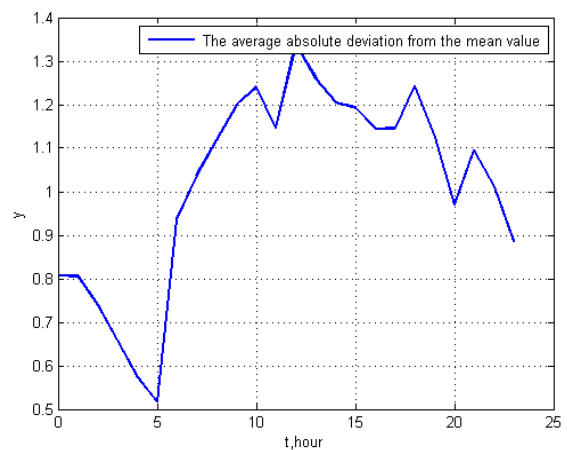


Рис. 8. Среднее отклонение от среднего значения нагрузки

Из рис. 7, 8 следует, что максимальное отклонение от среднего значения нагрузки составляет более 130%.

Прогнозирование проведено с помощью математического пакета MatLab Toolbox Statistics, который поддерживает 4 вида регрессионных модели. Ошибки по прогнозу для построенных моделей представлены в табл. 1.

Таблица 1

Прогноз	Модель прогнозирования (MatLab)			
	linear	interaction	purequadratic	quadratic
Ошибка, %	7,1067	6,6094	6,7973	6,0180

Минимальная ошибка из них более 5%, поэтому полученные регрессионные модели не могут найти свое применение.

Нейронечеткая модель прогнозирования построена с помощью пакета Fuzzy Logic Toolbox системы MatLab. Каждому входу системы соответствуют по две функции принадлежности. При использовании нейронечеткой модели получены значения ошибок прогноза – 4,91%. По сравнению с регрессионными моделями точность увеличилась на 1,1%.

При использовании нейросетевой модели получены следующие результаты:

- оптимальной по способности к обобщению является сеть с тремя слоями (2 скрытых слоя);
- точность выше, если учитываются три предыдущих значения данных электропотребления;
- введение в обучающую выборку дополнительных значений нагрузки в прошлые дни в прогнозируемый час повышает точность на 0.52%;
- введение дополнительных признаков (разделение суток на 24 часа), повышает точность прогнозирования на 1% (рис. 9, 10);
- для усиления способности к обобщению необходимо оперировать достаточно большим объемом обучающих данных (более 40 дней);
- предварительная обработка данных обучающей выборки улучшает точность прогноза;
- добавление шума в обучающую выборку позволяет повысить точность прогноза 0.48%.

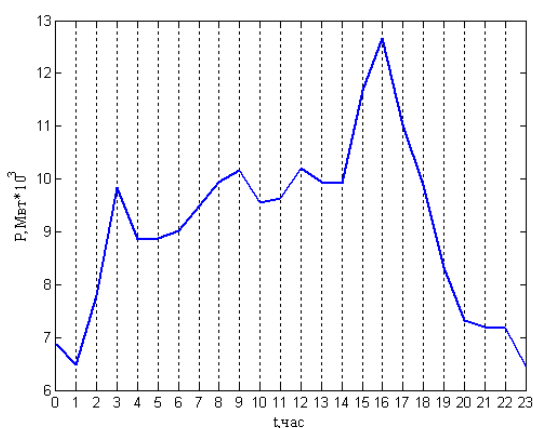


Рис. 9. Разделение суток на 24 часа

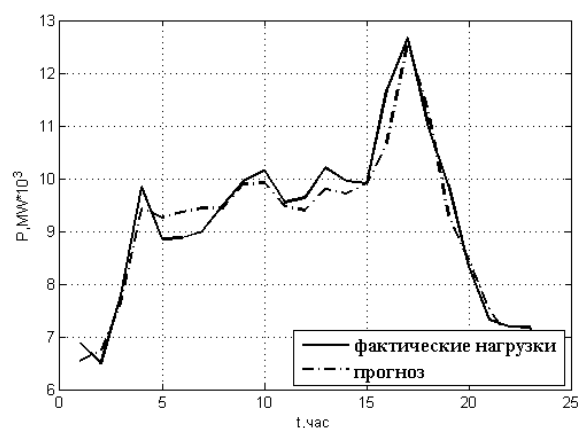


Рис. 10. Фактическое и прогнозируемое электропотребление

Сравнительный анализ построенных моделей прогнозирования электропотребления представлен на рис. 11. Ошибка прогноза с помощью нейронной сети – 3,31%, точность повысилась на 2,7%.

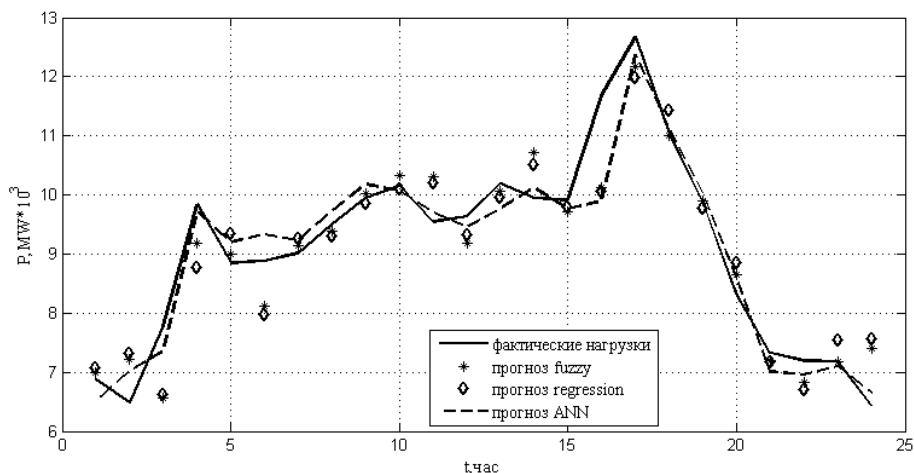


Рис. 11. Фактическое и прогнозируемое электропотребление различными методами  
Иллюстрацией результатов прогнозирования полученных с использованием метода опорных векторов, где в качестве ядрой выбрана гауссова радиальная базисная функция, служит рис. 12. Результаты исследований приведены в табл. 2.

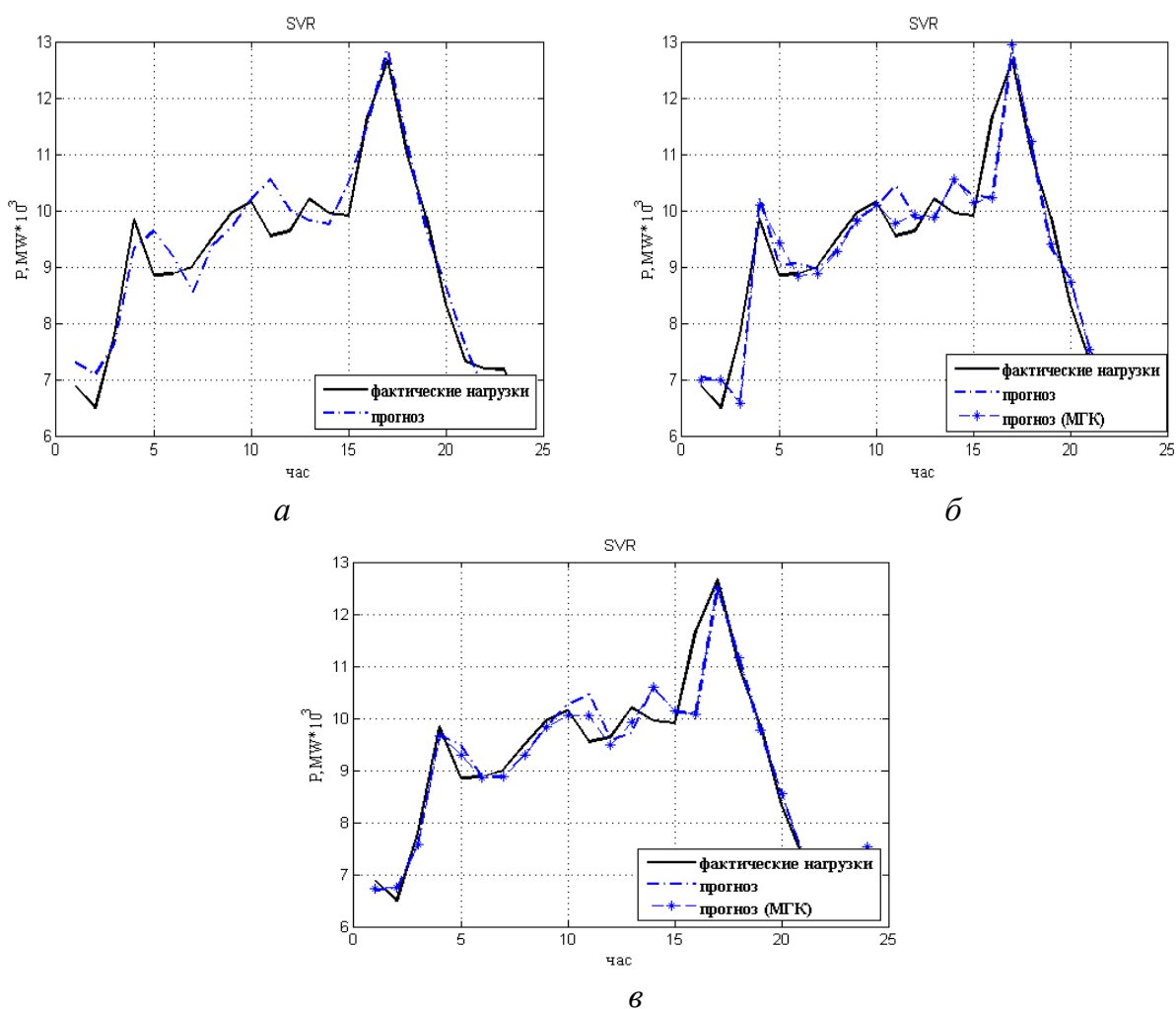


Рис. 12. Прогнозирование электропотребления на основе метода опорных векторов

Таблица 2

Метод	Параметры		Ошибка прогноза %
	$\sigma$	$C$	
Метод опорных векторов, (рис. 12 а)	13	65	3,3451
Метод опорных векторов + генетический алгоритм (+ метод главных компонент, МГК), (рис. 12 б)	6	55	2,99 (2,77)
Метод опорных векторов + метод роя частиц (+ метод главных компонент), (рис. 12 в)	5	90	2,9 (2,398)

Наилучший результат получен для модели оптимизации с помощью метода роя частиц при использовании метода главных компонент. Точность прогноза улучшилась на 3,6%.

Для построения графического интерфейса использован Toolbox GUI MatLab (рис. 13).

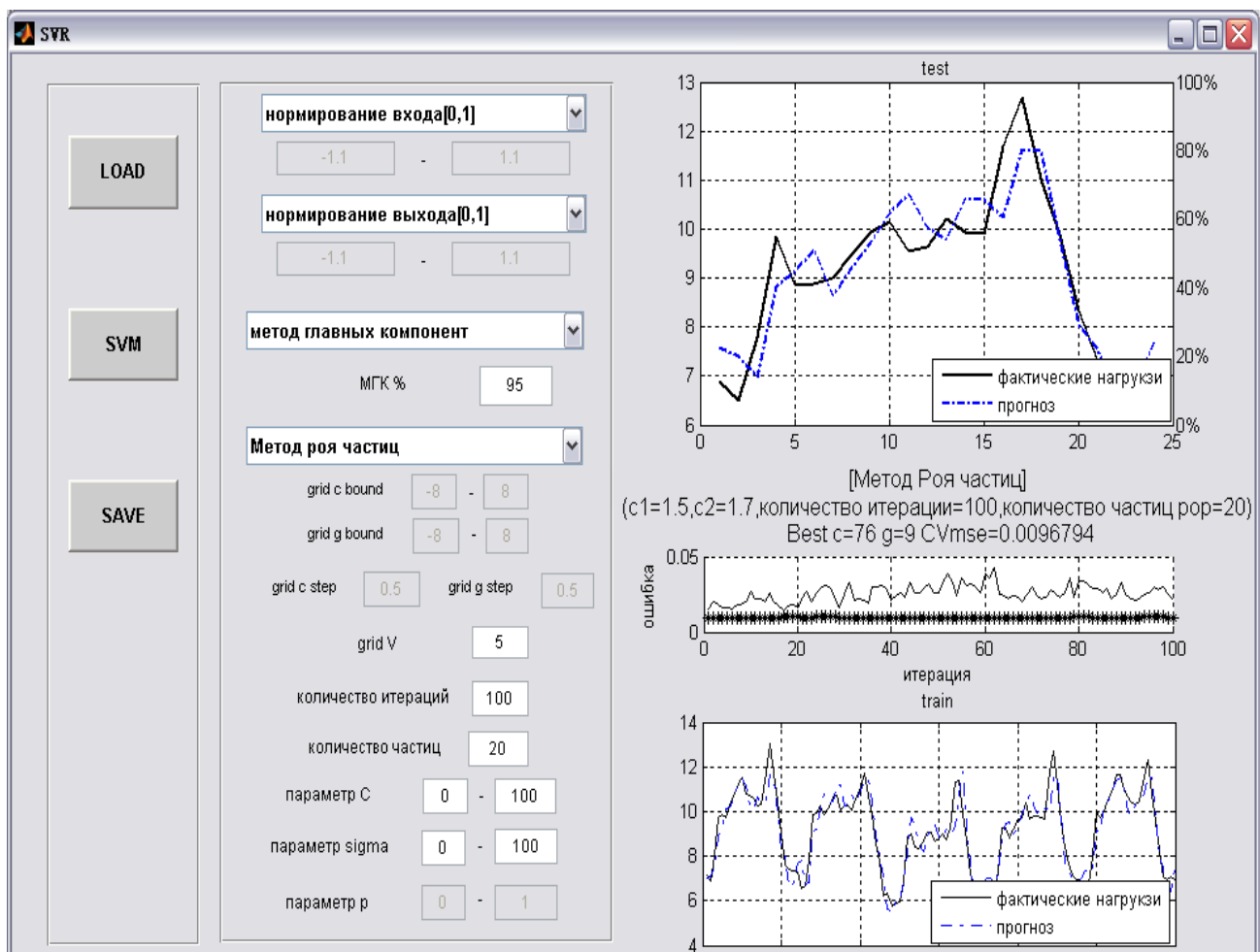


Рис. 13



## Заключение

1. В способе формирования обучающей выборки наиболее эффективным для повышения точности прогноза является уменьшение размерности исходных данных с использованием метода главных компонент, индикация времени суток, дополнительное введение значений нагрузки в прошлые дни в прогнозируемый час. Каждое из этих мероприятий обеспечивает уменьшение ошибки прогноза на 0,5 – 1%. Применение нормирования и добавление шума в обучающую выборку позволяют снизить ошибку на 0,3–0,5%.

2. В модели прогнозирования электрической нагрузки на основе нейросетевого подхода число нейронов скрытого слоя определяется на основе генетических алгоритмов, а при превышении допустимой погрешности предусмотрена возможность переобучения (улучшение точности в 0,5%).

3. Модель прогнозирования электрической нагрузки на основе метода опорных векторов, включающая предварительную обработку исходных данных, программу оптимизации выбора параметров ядерной функции на основе генетического алгоритма и алгоритма роя, позволяет улучшить прогноз на 1%.

Точность прогнозирования зависит от типа электрической нагрузки. При разбросе данных обучающей выборки 130% точность прогноза составила 2,4%.

4. Программные модули систем прогнозирования и пользовательские интерфейсы в пакете MatLab ориентированы на использовании диспетчерами энергообъединений в качестве систем поддержки принятия решений.

### Публикации по теме диссертации

Статьи, опубликованные в изданиях, включенных в перечень ВАК:

1. Поляхов, Н.Д., Приходько, И.А., Анушина, Е.С., Ван Ефэн. Оценка эффективности интеллектуальных и классических моделей краткосрочного прогнозирования электропотребления/ Н.Д. Поляхов, И.А. Приходько, Е.С. Анушина, Ван Ефэн //Естественные и технические науки. – 2011. – №3. – С. 304–309.

2. Поляхов, Н.Д., Приходько, И.А., Ван Ефэн, Д.В. Ли. Повышение точности прогнозирования на основе метода опорных векторов/ Н.Д. Поляхов, И.А. Приходько, Ван Ефэн, Д.В. Ли.// Известия СПбГЭТУ «ЛЭТИ», Сер. «Автоматизация и управление». – 2012. – № 7. – С. 64-68.

3. Поляхов, Н.Д., Приходько, И.А., Ван Ефэн. Прогнозирование электропотребления на основе метода опорных векторов с использованием эволюционных алгоритмов оптимизации/ Н.Д. Поляхов, И.А.Приходько, Ван Ефэн // Современные проблемы науки и образования. – 2013. – № 2. – URL: <http://www.science-education.ru/108-8962>

Другие работы:

4. Приходько, И.А. Ван Ефэн, Панюкова, Д.В. Решение задачи прогнозирования на основе нейросетевой модели с нечеткой обработкой информации/ И.А. Приходько, Ван Ефэн,

Д.В. Панюкова // Докл. междунар. конф. по мягким вычислениям и измерениям SCM, 2010. – Saint-Petersburg. – 23–25 June. – 2010. Т.1. – С.306–310.

5. Приходько, И.А. Ван Ефэн. Прогнозирование электропотребления на основе метода опорных векторов/ И.А. Приходько, Ван Ефэн // Материалы II Международной заочной научно-практической конф. «Актуальные проблемы науки» (UCOM–2011), 27 сентября, 2011 г. – Тамбов. – 2011. –Т.4. –С. 122–124.

6. Ван Ефэн, Ли, Д.В. Метод опорных векторов в задачах прогнозирования/ Ван Ефэн, Д.В. Ли// Докл. междунар. молодеж. форума «Энергоэффективные электротехнологии» Saint-Petersburg, 19-23 Sep. –2011. – С. 62.

7. Ван Ефэн, Ли, Д.В. Улучшение точности прогнозирования временных рядов на основе метода опорных векторов/ Ван Ефэн, Д.В. Ли// Докл. междунар. конф. «Управление в технических, эргатических, организационных и сетевых системах» (УТЭОСС–2012), (9 –11 октября 2012 г. Санкт–Петербург). – С. 545–548.