

УДК 004.41

UDC 004.41

**РАЗРАБОТКА АЛГОРИТМА ПОИСКА
ОПТИМАЛЬНОЙ МОДЕЛИ**

**DEVELOPMENT OF OPTIMAL MODEL
SEARCH ALGORITHM**

Дьяченко Роман Александрович
к.т.н., доцент кафедры Информатики
*Кубанский Государственный Технологический
Университет, Краснодар, Россия*

Dyachenko Roman Aleksandrovich
Cand.Tech.Sci., associate professor of the Computer
Science Department
*Kuban State Technological University, Krasnodar,
Russia*

Шароватов Александр Сергеевич
инженер
ОАО Конструкторское Бюро «Селена»

Sharovатов Aleksandr Sergeevich
engineer
Open Joint Stock Company Design Office "Selena"

Лоба Инна Сергеевна
ст. преподаватель кафедры Информатики и ИТО
*Армавирская Государственная Педагогическая
Академия, Армавир, Россия*

Loba Inna Sergeevna
senior lecturer of the Computer Science and
Information Technology in Education Department
Armavir State Pedagogical Academy, Armavir, Russia

Решетняк Михаил Геннадиевич
ассистент кафедры Информатики
*Кубанский Государственный Технологический
Университет, Краснодар, Россия*

Reshetnyak Michael Gennadievich
assistant of the Computer Science Department
*Kuban State Technological University, Krasnodar,
Russia*

В статье рассмотрены вопросы прогнозирования
многопараметрических систем. Представлен
алгоритм поиска оптимальной модели

In the article, some problems of multiparameter
systems' state forecasting are reviewed. The algorithm
of optimal model searching is given

Ключевые слова: МОНИТОРИНГ,
ПРОГНОЗИРОВАНИЕ, АППРОКСИМАЦИЯ,
ИНТЕРПОЛЯЦИЯ, КРИТЕРИИ
ОПТИМАЛЬНОСТИ

Keywords: MONITORING, FORECASTING,
APPROXYMATION, INTERPOLATION,
OPTIMALITY CRITERIA

В настоящее время для решения задач мониторинга и прогнозирования состояния различных технических объектов (в том числе объектов энергетики и ЖКХ) применяют различные методики (методы линейного регрессионного анализа, нейросетевые и нечёткие методы прогнозирования и т. д.). В процессе аппроксимации создается несколько моделей прогнозирования для определенного параметра. Так как все они дают разные результаты прогноза, возникает необходимость найти оптимальную модель среди построенных.

В [1] представлены основные современные методы прогнозирования. Основным недостатком приведенных методов является использования одного критерия для поиска оптимальных параметров модели

прогнозирования. Тогда как поиск оптимальной модели – сложная и многоаспектная задача, для которой найти универсальное решение весьма затруднительно. Каждая модель имеет различные критерии, которые могут служить в качестве параметров поиска оптимальной модели. Важность каждого из критериев зависит от области, в которой решается задача прогнозирования.

Одним из вариантов решения задачи поиска оптимальной модели является выбор основного, наиболее важного критерия и нахождение модели с минимальным (либо максимальным, в зависимости от критерия) значением. Найденная модель, хотя и будет иметь наилучшее значение критерия поиска, может иметь низкое качество по другим критериям, поэтому целесообразно задать ограничения по другим критериям. В итоге поиск оптимальной модели сводится к нахождению модели с наилучшим значением основного критерия оптимальности из тех моделей, что удовлетворяют заданным ограничениям.

Стандартными аппроксимирующими зависимостями являются следующие [2]:

$$y(t) = a_0 + a_1 \cdot t \quad (1)$$

$$y(t) = a_0 + a_1 \cdot t + a_2 \cdot t^2 + \dots + a_n \cdot t^n \quad (2)$$

$$y(t) = a_0 \cdot e^{a_i \cdot t} \quad (3)$$

$$y(t) = a_0 \cdot t^{a_1} \quad (4)$$

$$y(t) = a_0 + a_1 \cdot t + a_2 \cdot \ln(t) \quad (5)$$

Каждая модель характеризуется не только коэффициентами регрессионного уравнения, а также критериями оценки их качества, то есть параметрами, которые используются для определения оптимальной модели.

К таким критериям относятся [2]:

$$R = \sqrt{1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2}} \quad (6)$$

$$R2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{\sum (y - \hat{y})^2}{N}} \quad (8)$$

$$MAE = \sum |y - \hat{y}| \quad (9)$$

$$MAPE = \sum \left| \frac{y - \hat{y}}{y} \right| \cdot \frac{1}{N} \quad (10)$$

$$F = \frac{R2}{1 - R2} \cdot \frac{n - m - 1}{m} \quad (11)$$

$$d = \frac{\sum_{i=2}^N (e_i - e_{i-1})^2}{\sum_{i=1}^N e_i^2} \quad (12)$$

где \hat{y} - значение уравнения регрессии;

\bar{y} - среднее значение экспериментальных значений y ;

N, n – количество экспериментов;

m – количество факторов модели;

$e_i = y_i - \hat{y}_i$.

Алгоритм поиска оптимальной модели по заданному критерию оптимальности и ограничениям по другим критериям состоит из следующих шагов:

1. Получение наборов данных для аппроксимации каждого необходимого параметра (например, с датчиков технических устройств).

2. Построение необходимых моделей.
3. Вычисление необходимых критериев оптимальности для каждой модели.
4. Определение моделей, удовлетворяющих заданным ограничениям по критериям.
5. Определение модели, обладающей наилучшим значением заданного критерия оптимальности.

Далее в тексте будем использовать обозначения и названия типов данных для языка программирования С# версии 4.0 [3].

Поиск оптимальной модели осуществляется перебором всех значений критериев оптимальности выбранного типа и выбором модели, имеющей оптимальное значение выбранного критерия. При наложении ограничений по другим критериям следует сначала отсеять модели, не удовлетворяющие заданным ограничениям.

Поиск оптимального значения среди критериев, а так же проверка значения критерия на соответствие ограничению осуществляются путём сравнения двух критериев одного типа и выбора критерия с оптимальным значением, причём для одних критериев оптимальным значением считается большее, для других – меньшее. Для сравнения значений критериев целесообразно перегрузить операторы сравнения “<” и “>”. Блок-схема оператора “<” представлена на рисунке 1.

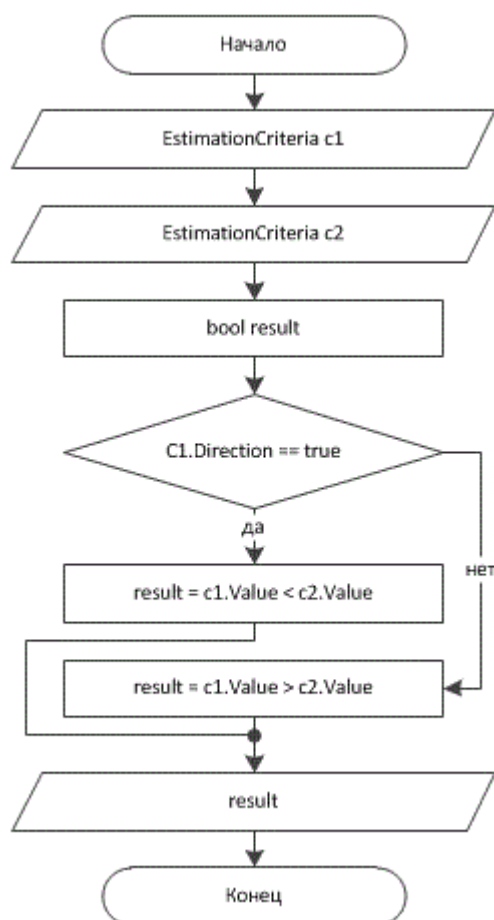


Рис. 1. Блок-схема оператора “<” сравнения значений двух критериев **c1**, **c2** – соответственно первый и второй сравниваемые критерии оптимальности (значения типа EstimationCriteria).

result – результат сравнения (переменная типа bool).

c1.Direction – “направление” сравнения, определяется типом критерия (переменная типа bool). При **Direction = true** оптимальным является критерий с наибольшим значением, при **false** - наименьшим.

c1.Value, **c2.Value** – значения критериев оптимальности (переменные типа double).

Оператор “>” отличается использованием противоположных операторов сравнения – “<” меняется на “>” и наоборот.

Блок-схема алгоритма поиска оптимальной модели представлена на рисунке 2.

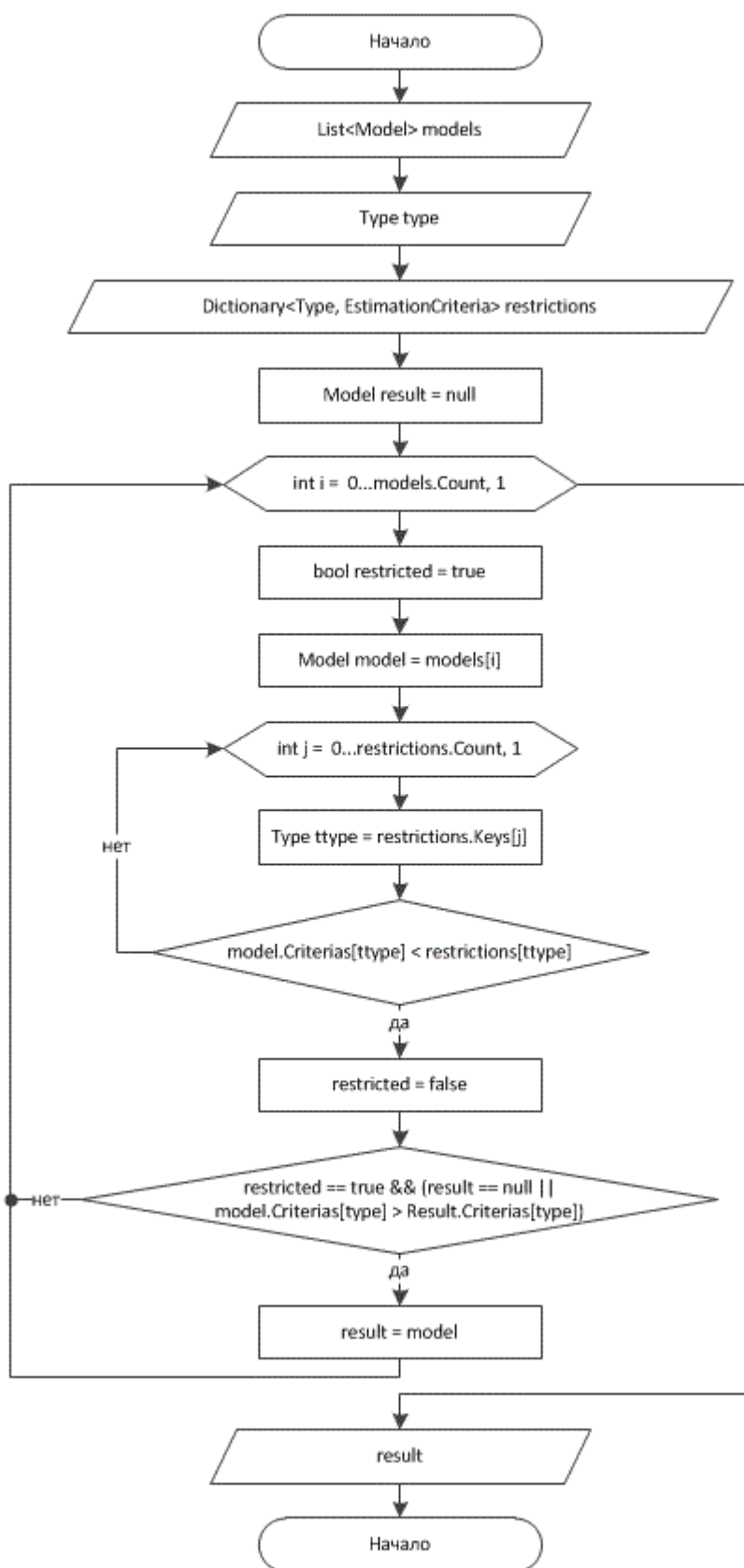


Рис. 2. Блок-схема алгоритма поиска оптимальной модели

models – массив моделей, среди которых необходимо найти оптимальную (переменная типа Model).

models.Count – количество моделей в списке (переменная типа int).

type – тип критерия, по которому следует выбирать оптимальную модель (переменная типа Type).

restrictions – словарь ограничений, где ключом каждого значения является его тип (ассоциативный массив пар типа Type, EstimationCriteria).

restrictions.Count – количество заданных ограничений (переменная типа int).

restrictions.Keys – типы заданных ограничений (массив элементов типа Type).

model – рассматриваемая модель (переменная типа Model).

restricted – переменная, отражающая соответствие рассматриваемой модели заданным ограничениям (переменная типа bool).

ttype – тип рассматриваемого ограничения (переменная типа Type).

model.Criteria – список критериев модели (массив элементов типа EstimationCriteria).

Операторы “<” и “>” – операторы сравнения критериев.

Вышеописанный алгоритм реализован в программе “PredictPro”, написанной на языке C# в среде Microsoft Visual Studio 2010 на технологии .NET Framework 4.0 [2].

Принципы работы системы состоят в следующем:

- пользователь работает с проектом, в который можно добавлять наборы данных (представляют собой параметры).
- Наборы данных аппроксимируются моделями – в набор данных можно добавить любое количество моделей следующих типов: линейная, полиномиальная, логарифмическая, степенная, экспоненциальная.
- По выбранному набору данных можно осуществить поиск оптимальной модели по одному из критериев (при этом можно задать ограничения на значения других критериев).
- Для наборов данных, либо для отдельных моделей можно создавать прогнозы.

- После задания верхней и нижней границы допустимых значений, прогноз отобразит время выхода модели за эти границы.
- Программа автоматически осуществляет поиск прогнозов с наименьшим временем выхода моделей за границы допустимых значений и помечает их красным цветом, а также выводит их наименования в строке состояния.
- Наименьшее время выхода моделей за границы допустимых значений отображается в строке состояния.
- Проект может быть сохранён на любой носитель информации в виде XML-файла.
- Наборы данных могут быть экспортированы в Excel.
- Для любой модели может быть сгенерирован отчёт в виде PDF-файла.

Интерфейс системы представлен на рисунках 3, 4, 5.

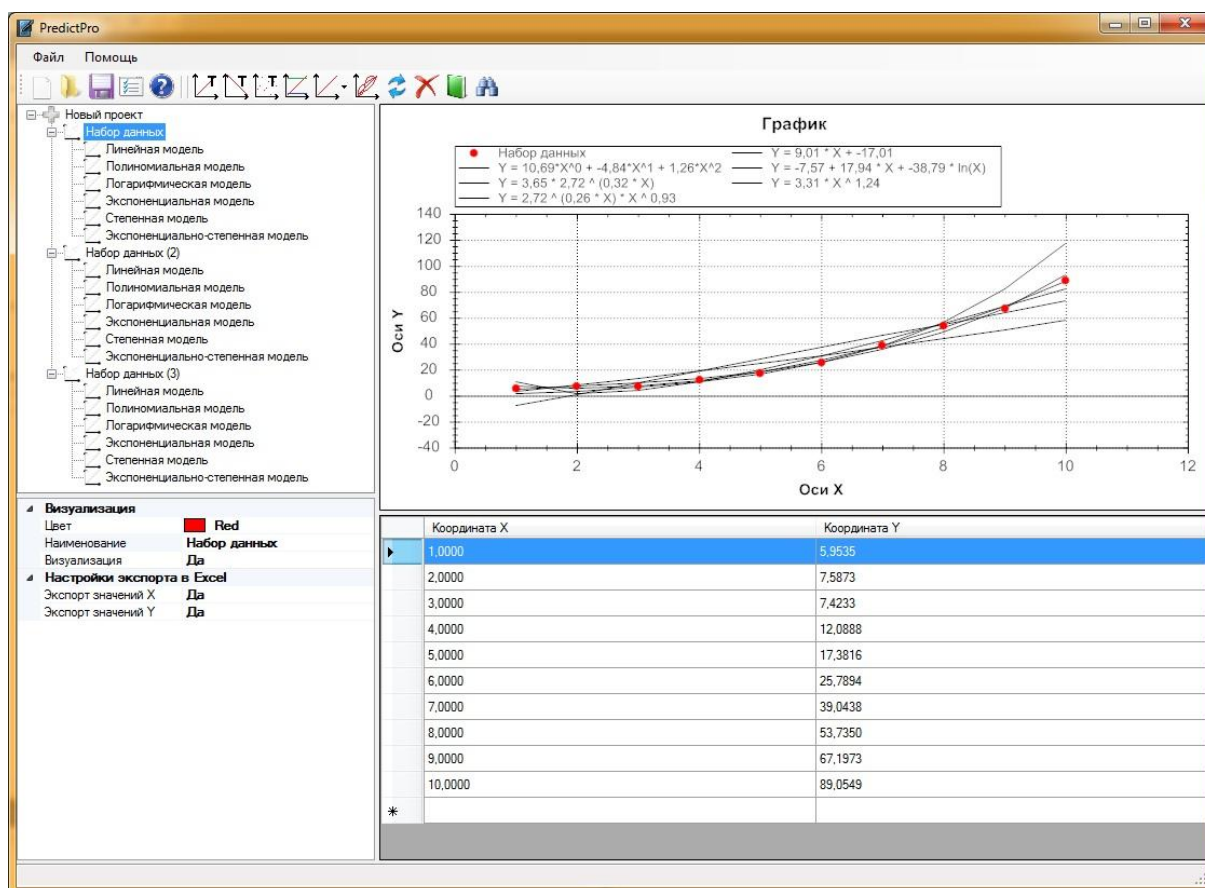


Рис. 3. – Интерфейс системы “PredictPro”

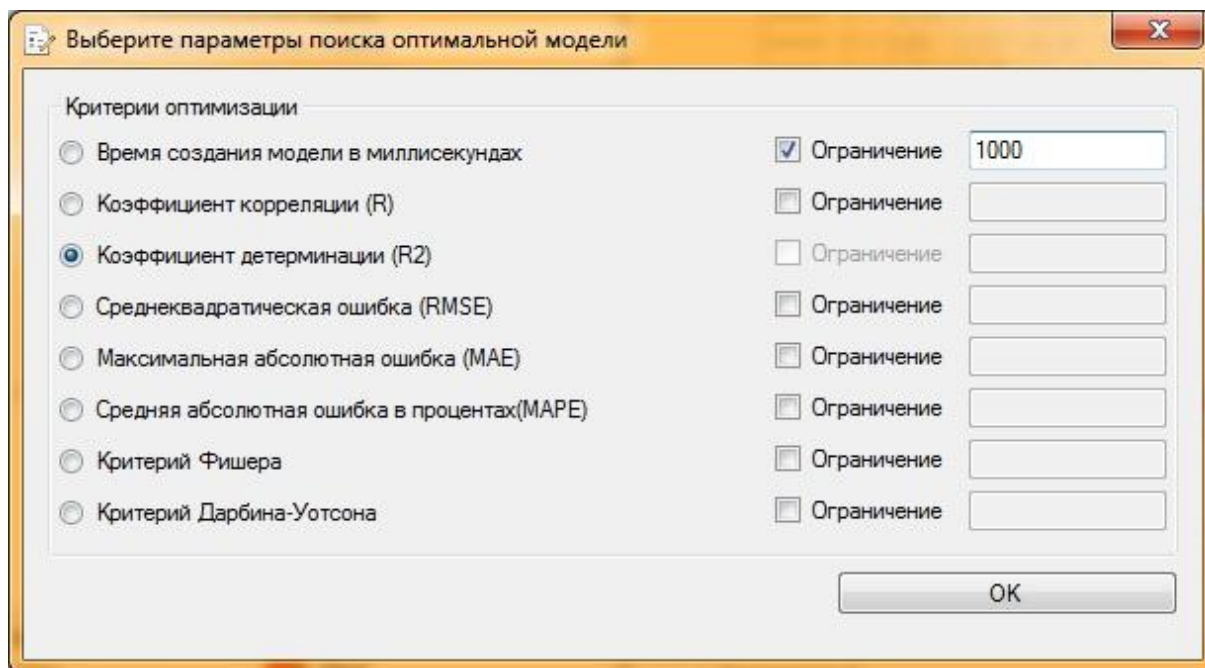


Рис. 4. Меню выбора параметров поиска оптимальной модели

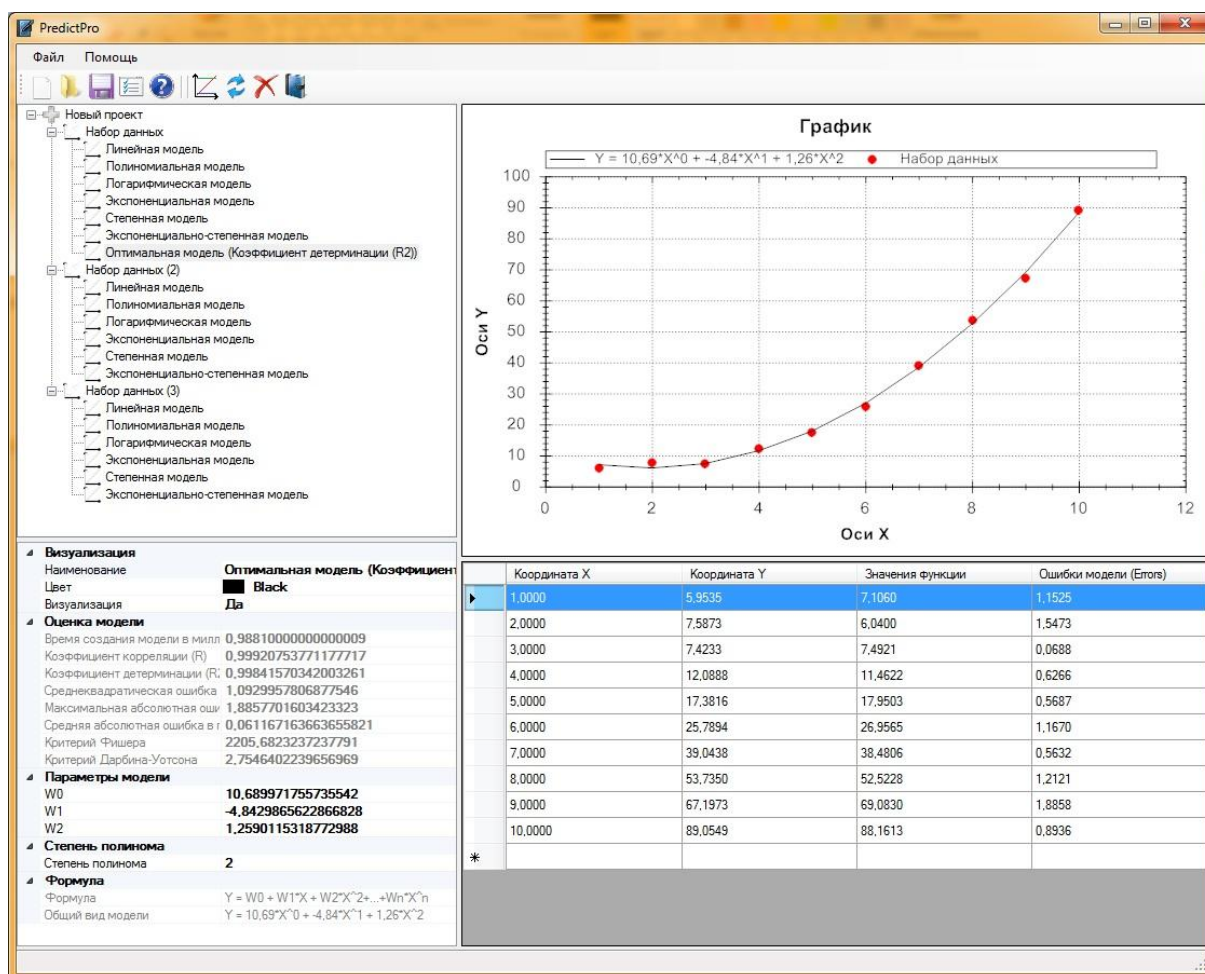


Рис. 5. Результат поиска оптимальной модели

На рисунке 5 продемонстрирован результат поиска оптимальной модели по критерию R2.

Эффективность внедрения такого программного продукта состоит в том, что стоимость разработки данного ПО значительно ниже тех затрат, нежели чем эффект (экономия) от применения прогнозирования (например прогнозирование отказа от выполнения избыточного (дорогостоящего) технического обслуживания дизельных электростанций).

В результате проведённых исследований был составлен алгоритм и разработано программное обеспечение, которое без сложных математических моделей позволяет осуществлять поиск оптимальных моделей для любых областей применения.

Литература

1. Меламед А.М. Современные методы анализа и прогнозирования режимов электропотребления в электроэнергетических системах // Итоги науки и техники. Энерг. системы и их автоматизация. 1988. Т. 4.
2. Гайдышев И.П. Анализ и обработка данных: спец. справ. — СПб.: Питер, 2001. — 752 с.
3. Атрощенко В.А., Дьяченко Р.А., Шароватов А.С., Литвинов Ю.Н. Разработка алгоритма определения критического параметра ресурса технических систем // Научно-технические ведомости СПбГПУ. 2011. № 3(126). С. 89-93.