

IOSO Аппроксимация 1.0

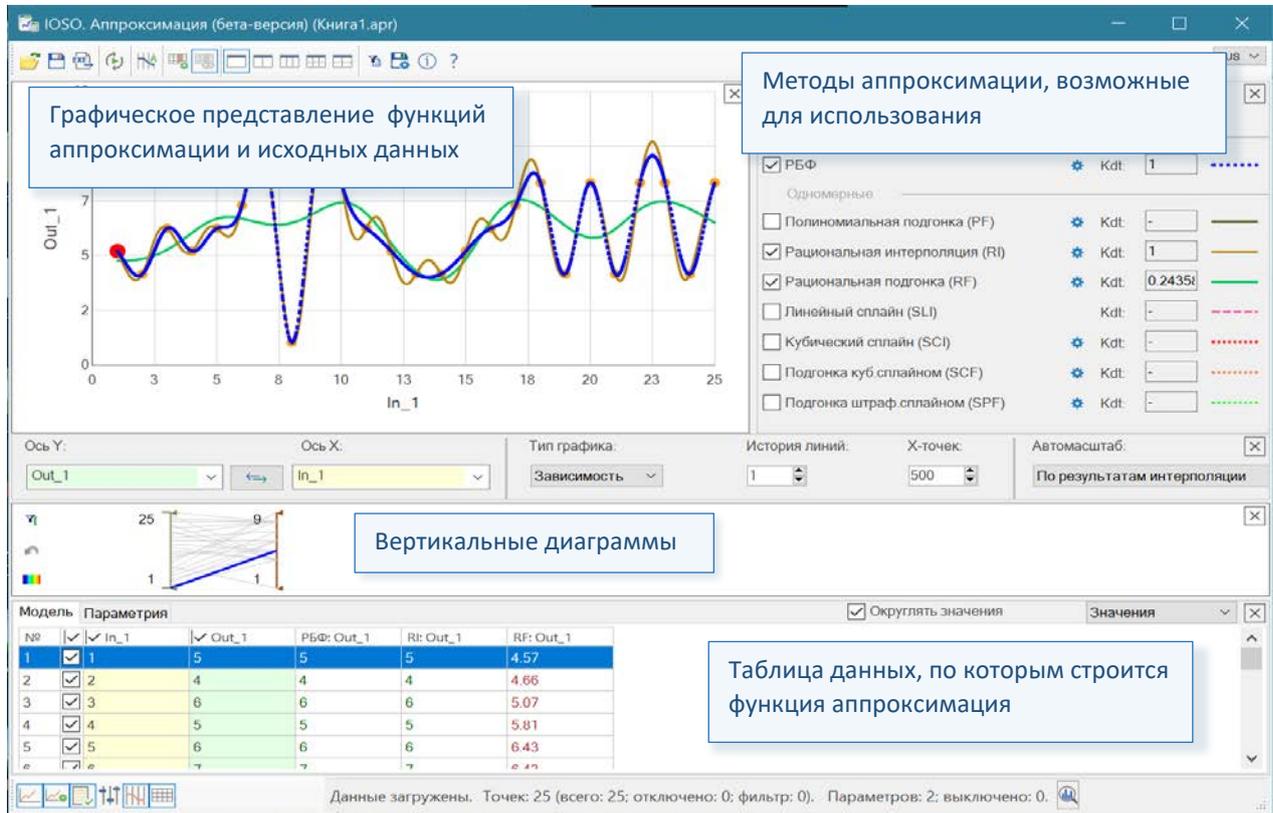
Руководство пользователя

Оглавление

Приложение “IOSO Аппроксимация” 1.0.....	3
Программные интерфейсы приложения “IOSO Аппроксимация”	5
Создание аппроксимационной модели по шагам	8
Шаг 1. Загрузка исходных данных	8
Шаг 2. Построение функции аппроксимации.....	10
Многомерные методы аппроксимации	11
Методы одномерной аппроксимации.....	15
Шаг 3 Сохранение аппроксимационной модели.....	16

Приложение “IOSO Аппроксимация” 1.0

Приложение «**IOSO Аппроксимация**» позволяет находить приближённую зависимость величин через другие, более простые величины и создать аппроксимационную модель в виде исполняемого файла (*.exe).



Приложение **IOSO Аппроксимация** позволяет пользователю создавать функции аппроксимации с использованием эффективных методов одномерной и многомерной аппроксимации.

Одномерные функции:

- Полиномиальная интерполяция (PI)
- Полиномиальная подгонка (PF)
- Рациональная интерполяция (RI)
- Рациональная подгонка (RF)
- Линейный сплайн (SLI)
- Кубический сплайн (SCI)
- Сплайн Катмулла-Рома (SCRI)
- Сплайн Акимы (SAI)
- Подгонка куб. сплайном (SCF)
- Подгонка сплайном Эрмита (SHF)
- Подгонка штраф. сплайном (SPF)

Многомерные функции:

- РБФ - радиально-базисная функция;
- IDW - метод обратных взвешенных расстояний.

Под подгонкой понимается использование семейства функций для одного массива (сочетание одного типа функций с различными параметрами на разных отрезках).

В скобках указана сокращенное название метода, используемое в таблице данных.

Данными для построения функции аппроксимации могут быть:

- протоколы проектов программы оптимизации IOSO;
- массивы данных в текстовых файлах, в том числе CSV-файлы;
- данные, вводимые в таблицу программы **IOSO Аппроксимация**.

Построенная функция аппроксимация может быть сохранена в виде исполняемой модели (exe-файле) и использована в дальнейшем как суррогатные модели в задачах моделирования систем или объектов. В программе оптимизации IOSO данные модели могут быть использованы для замены сложных ресурсоемких вычислительных моделей при решении задач оптимизации с многодисциплинарных проектах и реализована возможность автоматизированного создания, подключения и обучения аппроксимационных моделей на новых данных.

Системные требования

Общие технические требования: ОС: Windows 7 SP 1, Windows 8.1, Windows 10.

Оперативная память 4ГБ и более.

Система доступна для 32-х и 64-разрядных операционных систем.

На компьютере должен быть установлен Microsoft .NET Framework версии 4.7 или выше.

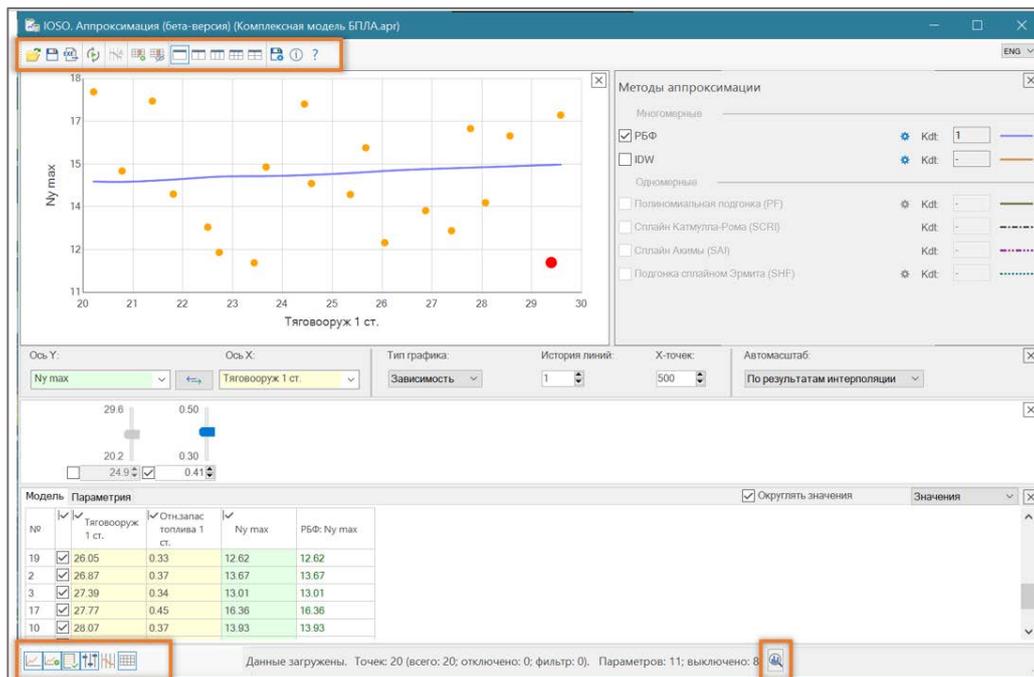


Приложение **IOSO Аппроксимация** разработано в компании СИГМА Технологии (www.iosotech.com/ru).

Предложения и замечания по программному приложению просьба направлять на company@iosotech.com с пометкой в теме письма **IOSO**.

Программные интерфейсы приложения “IOSO Аппроксимация”

Инструментальные панели “IOSO Аппроксимация”



Этот интерфейс предназначен для создания аппроксимационной модели и представления информации о ней с использованием различного таблично-графического виде.

На верхней инструментальной панели находятся команды:



- открыть файл. Для открытия доступны различные типы файлов:

- файл с данными в формате .csv или текстовом формате. Для текстового формата значения должны быть записаны в колонках с разделителями: точка с запятой или табуляция;
- ранее сохраненный файл проекта аппроксимации (*.arg);
- файл проекта программы оптимизации IOSO (*.orm, *.porm). Из данного файла могут быть взяты данные из протоколов параметрических исследований, из протокола текущей оптимизационной задачи или из общего протокола. Из этих данных могут быть выбраны данные для отображенных моделей (модели).



- сохранить проект аппроксимации (*.arg).



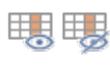
– сохранить созданную функцию аппроксимации в виде исполняемого файла (*.exe). Данный исполняемый файл может быть использован как самостоятельная расчетная модель в программе оптимизации IOSO либо для иных целей.



– ручное перестроение функции аппроксимации. Не активна при включенном режиме «Перестраивать автоматически» в разделе методов аппроксимации;

 - показывать значения функции аппроксимации в окне с вертикальными диаграммами;

 - настройка используемых параметров (столбцов) в таблице для построения функции аппроксимации ;

 - показывать или не показывать в таблице столбцы с неиспользуемыми параметрами для построения функции аппроксимации;

 - выбор количества графиков, отображаемых в окне «Графики»;

 -- настройка пользовательского интерфейса и используемых методов аппроксимации;

 - настройка фильтрации данных, используемых для построения функции аппроксимации. Настройки данного фильтра синхронизируются с фильтром на вертикальных диаграмм;

 - информация о программе IOSO Approximation;

 - справочная информация.

На нижней инструментальной панели находятся команды отображения оконных интерфейсов программы:



Данные команды используются для удобства визуализации и работы с программой с дисплеями небольшого разрешения.

 - включить / отключить окно «Графики»;

 - включить / отключить панель «Настройка графиков»;

 - включить / отключить окно «Методы построения функции аппроксимации»;

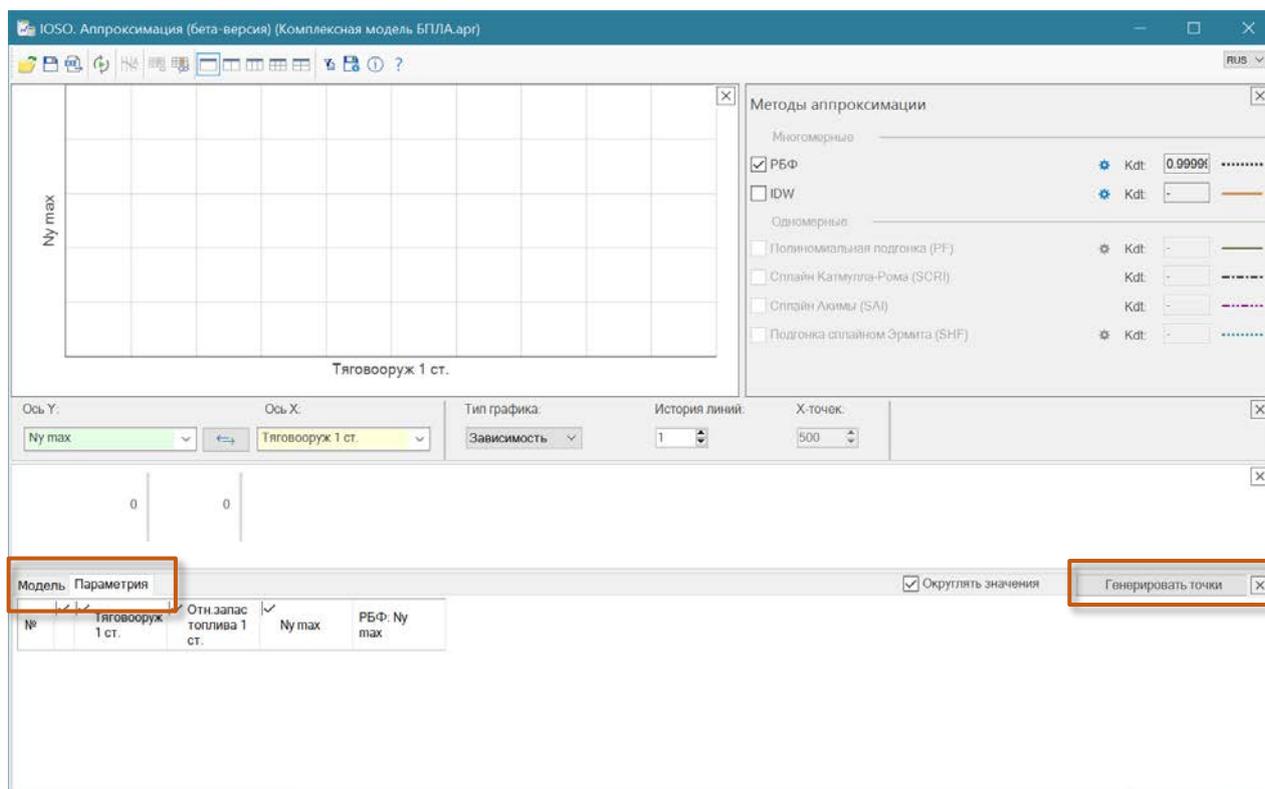
 включить / отключить окно «Задание константных величин». Задание константных величин используется для оценки поведения многомерной функции аппроксимации для изменении выбранных параметров (входных значений);;

 включить / отключить окно «Вертикальные диаграммы»;

 включить / отключить окно «Таблица»;

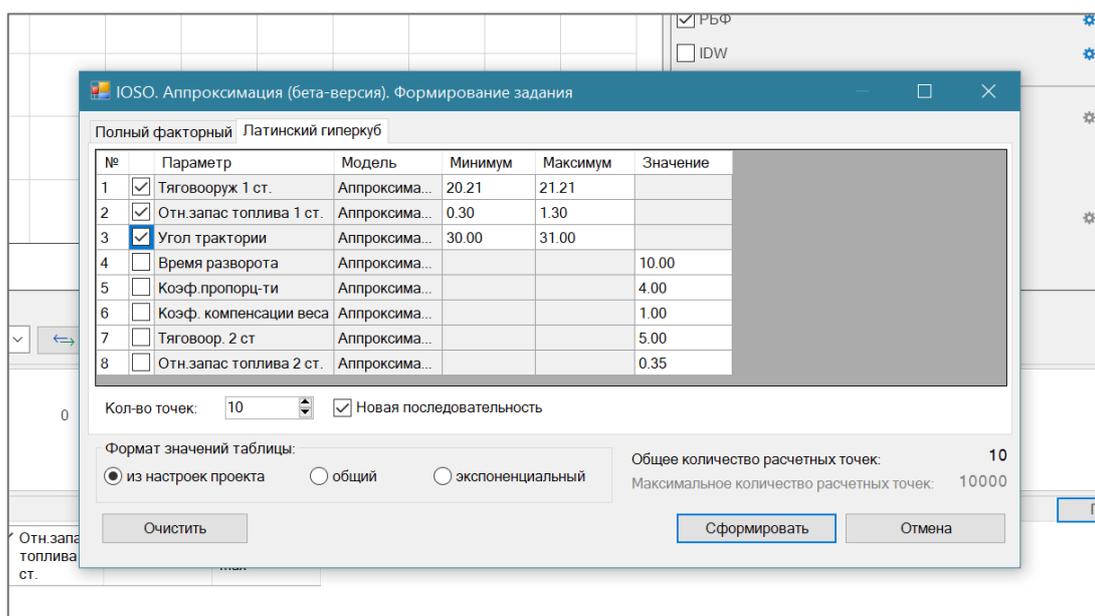
 - статистическая информация о массиве исходных данных.

Имеется возможность проведения параметрических исследований созданной аппроксимационной модели с использованием интерфейса «Параметрия». Переключение на этот режим осуществляется на закладке (смотри рис. Ниже)



Формирования массива расчетных точек осуществляется с использованием формы создания плана расчетов. Доступны режимы создания массива данных:

- полнофакторный эксперимента;
- латинский гиперкуб.



Создание аппроксимационной модели по шагам

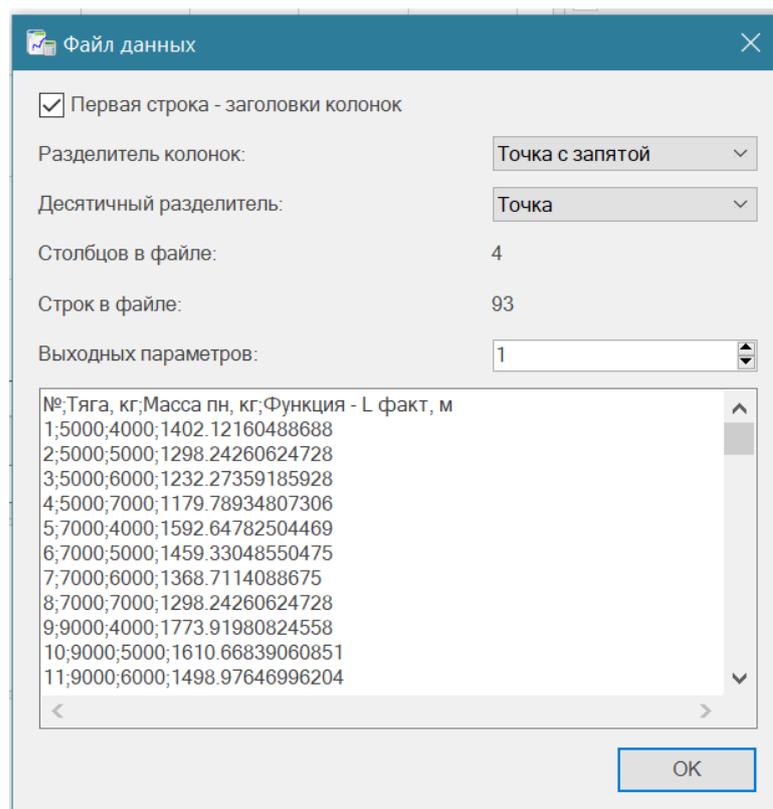
Шаг 1. Загрузка исходных данных

Для загрузки массива данных, для которого необходимо построить функцию аппроксимации использовать команду «Открыть файл», иконка  на инструментальной панели.

Доступны различные форматы файлов с массивами данных:

- файл с данными в формате .csv или текстовом формате. Для текстового формата значения должны быть записаны в колонках с разделителями: точка с запятой или табуляция;
- ранее сохраненный файл проекта аппроксимации (*.arg);
- файл проекта программы оптимизации IOSO (*.opt, *.port). Из данного файла могут быть взяты данные из протоколов параметрических исследований, из протокола текущей оптимизационной задачи или из общего протокола. Из этих данных могут быть выбраны данные для отобранных моделей (модели).

При загрузке массива данных в текстовом формате в окне «Выходных параметров» необходимо указать количество колонок со значениями функции. Предполагается, что в первых колонках записаны аргументы функции, в последних - значения функции. Количество значений функций может быть 1 и более.



Данные загружаются в таблицу и отображаются на графиках. На желтом фоне в таблице – аргументы функции (на данном рисунке их 2), на зеленом – значения функции (на рисунке их 4).

Предусмотрена возможность использования *Check box* для выбора данных (аргументов функции и функций) для которых необходимо построить функции аппроксимации. На рисунке показано, что для функции «Y2_Belegundu» функция аппроксимации не строится и значения в строках 5,7 не будут использованы для построения функций аппроксимации.

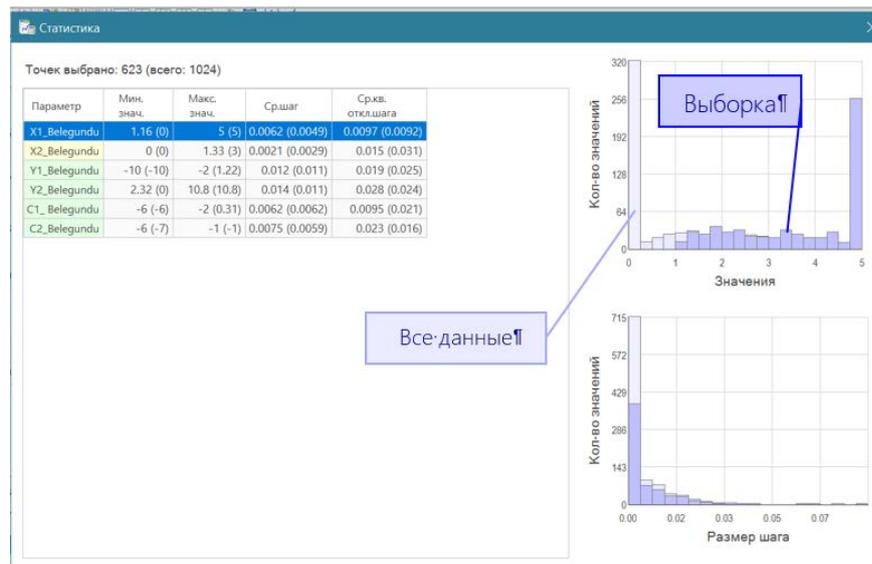
Модель		Параметрия					
№	<input type="checkbox"/>	<input checked="" type="checkbox"/> X1_Belegundu	<input checked="" type="checkbox"/> X2_Belegundu	<input checked="" type="checkbox"/> Y1_Belegundu	<input checked="" type="checkbox"/> Y2_Belegundu	<input checked="" type="checkbox"/> C1_Belegundu	<input checked="" type="checkbox"/> C2_Belegundu
1	<input checked="" type="checkbox"/>	2.71	3	-2.4	8.4	-0.71	-1.29
2	<input checked="" type="checkbox"/>	2.72	1.69	-3.7	7.1	-2.03	-2.59
4	<input checked="" type="checkbox"/>	3.08	2.94	-3.2	9.1	-1.14	-0.98
5	<input type="checkbox"/>	3.45	0.97	-5.9	7.9	-3.48	-2.58
6	<input checked="" type="checkbox"/>	2.8	2.28	-3.3	7.9	-1.52	-1.92
7	<input type="checkbox"/>	1.54	1.21	-1.9	4.3	-1.33	-4.25
8	<input checked="" type="checkbox"/>	3.76	1.16	-6.4	8.7	-3.61	-2.08

Использование кнопки «Статистика о массиве исходных данных»  позволяет получить следующую информацию о массиве исходных данных:

- минимальное и максимальное значение параметра;
- средне-арифметическое значение шага между значениями параметра («Ср. шаг»);
- средне квадратичное отклонение для значений шага параметра («Ср. кв. откл. шага»). Для построения функции аппроксимации рекомендуется использование массивов данных с более равномерным построением параметров в исследуемом диапазоне;
- графическое распределение - гистограмма количества значений в 20 равных отрезках искомого диапазона;
- гистограмма количества значений шага параметра в 20 равных отрезках имеющегося диапазона шага параметра.

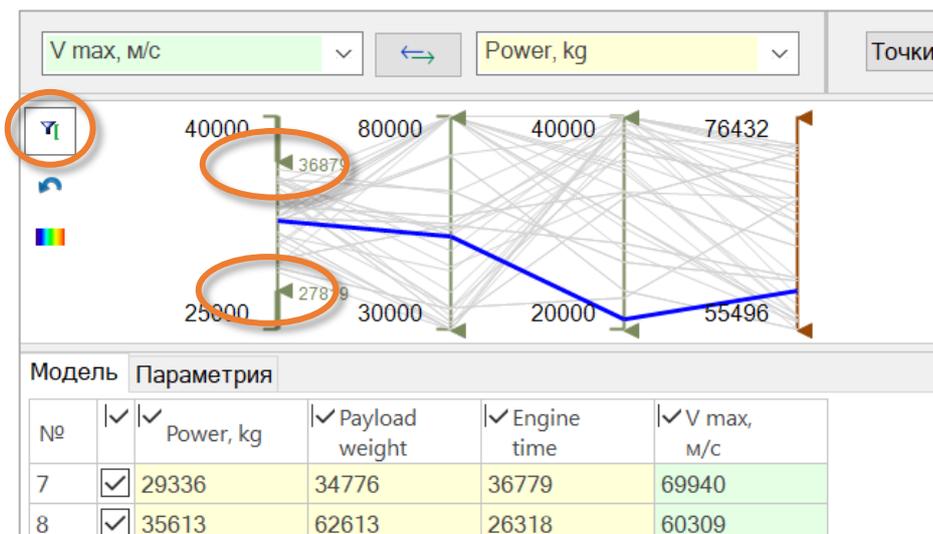
В таблице указываются значения для текущей и всей выборки. Значения для всей выборки указываются в скобках.

Светлый цвет на гистограмме соответствует значениям всего массива данных, темный цвет соответствует значениям текущей выборки (при использовании фильтрации массива данных).

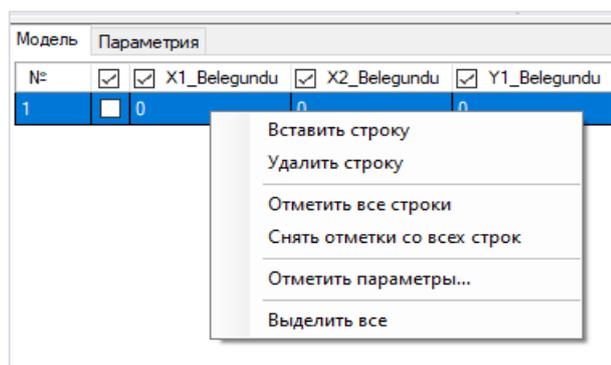


Выборку массива данных также можно фильтровать, используя фильтр в окне «вертикальные диаграммы». Включение фильтра осуществляется при нажатии значка , изменение диапазона

параметров в выборке осуществляется передвижением мышкой индексов на вертикальной шкале параметра.

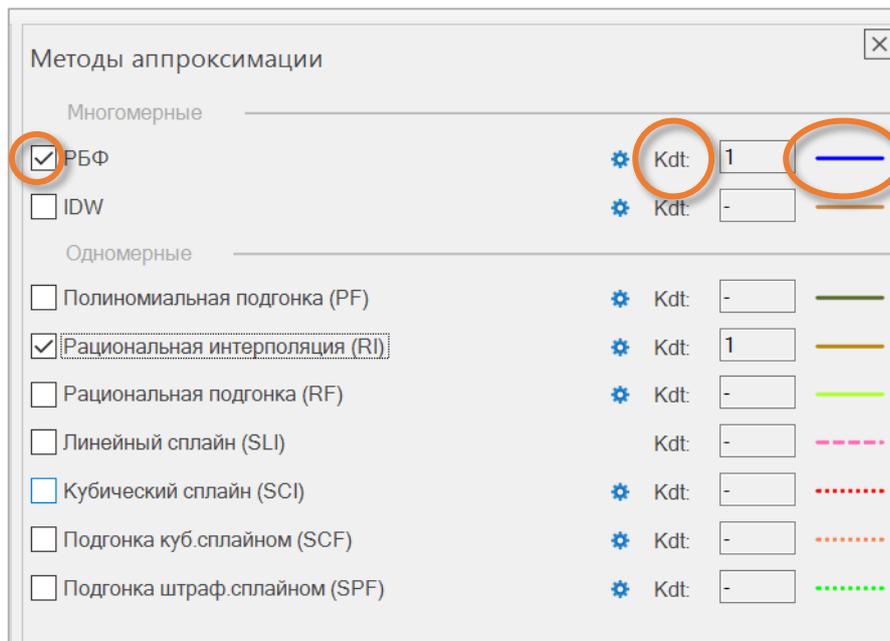


Предусмотрена возможность «вручную» сформировать массив данных в таблице. Меню редактирования таблицы вызывается правой кнопкой мыши.



Шаг 2. Построение функции аппроксимации

Для построения функции аппроксимации необходимо выбрать метод/методы аппроксимации и функция будет построена автоматически.



Цвет функции аппроксимации можно изменить - дважды нажать на образец линии функции аппроксимации.

Для методов аппроксимации имеется интерфейс редактирования настроек метода -  .

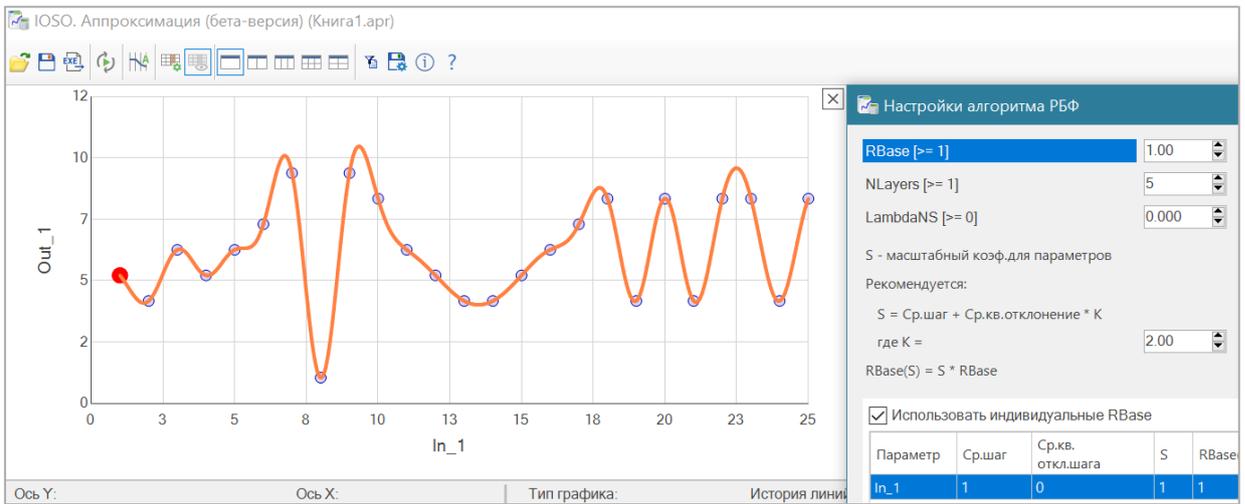
Многомерные методы аппроксимации

- РБФ - радиально-базисная функция;
- IDW - метод обратных взвешенных расстояний.

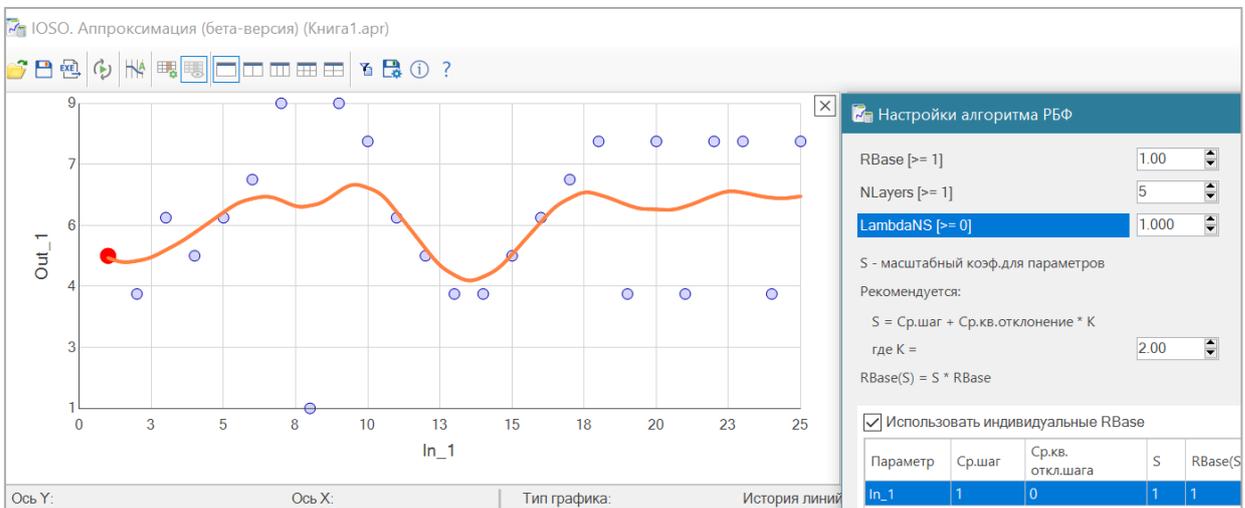
РБФ - Радиально-базисная функция

Алгоритм RBF-является значительным прорывом в интерполяции. Он может решать проблемы интерполяции и сглаживания и работать с большими массивами данных (десятки тысяч точек), с которыми другие алгоритмы не работают. Наиболее важной особенностью является его высокая производительность - время построения модели для N точек увеличивается, как $N \times \log N$. RBF может работать с произвольным распределением узлов, включая плохо разделенные или не различимые узлы. Этот алгоритм рекомендован как основной для интерполяции/аппроксимации.

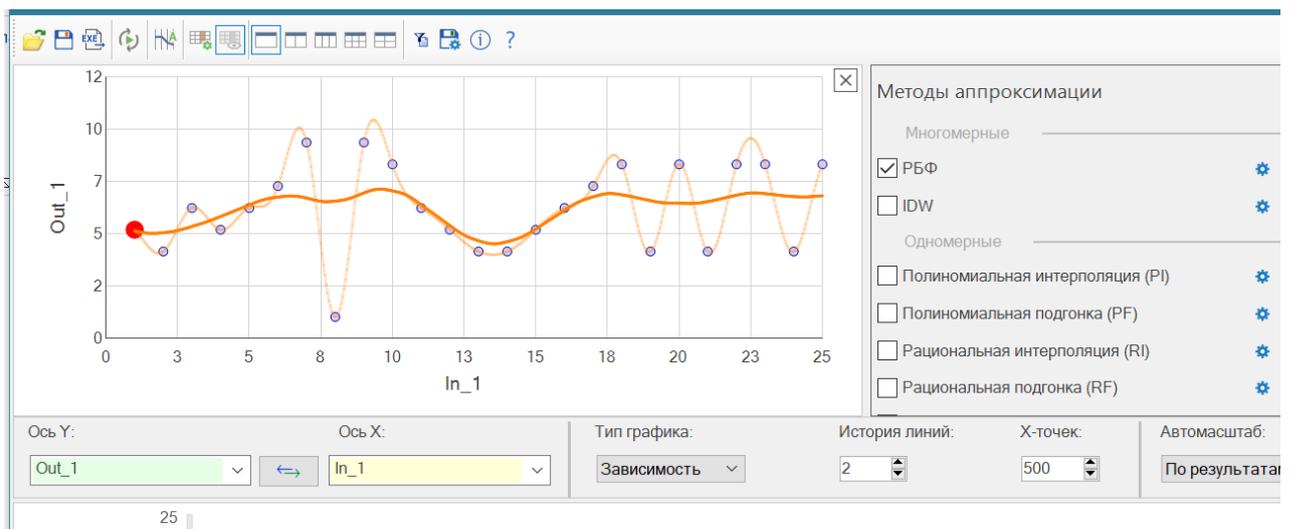
РБФ алгоритм имеет управляемое сглаживание с помощью штрафов за нелинейность.



Штраф за нелинейность = 0.



Штраф за нелинейность = 1.



Параметр «История линий» позволяет отслеживать историю изменений – показываются предыдущие графики (указанное значение соответствует количеству линий в истории построения).

Общие рекомендации настройки параметров РБФ метода.

Значение **RBase** указывает какие «дырки» необходимо заделывать (например, шаг параметра 200 и 250, если **RBase** указать 200 -> значения с шагом 250 будут усредняться).

Рекомендации для **Rbase** – в 2-3 раза выше максимального шага сетки для охвата всех областей.

О передаче мелких деталей функции свидетельствует параметра **RBase/2**(Nlayers -1)**. Сглаживаться будут значения при шаге меньше значения этого параметра.

При значении параметра **LamdaNS = 0** сглаживания осуществляются за счет управления параметрами **RBase** и **Nlayers** (рекомендуется).

Увеличение значения **LamdaNS** приводит к увеличению линейности (для некоторых функций линейность достигается уже при **LamdaNS = 1**).

Для РБФ метода реализована процедура автоматической подбора коэффициентов метода для каждого параметра массива данных. Суть адаптивной настройки заключается в определении значения **RBase** таким, что бы он охватывал практически все «дырки» массива. Значение **RBase(S)** для каждого параметра определяется умножением базового значения **RBase** (задаваемого в первой строке экранной формы) на значение **S**, равного сумме среднего значения шага параметра и среднеквадратичному отклонению шага умноженному на коэффициент **K**. Значение коэффициента **K** целесообразно задавать в диапазоне от 2 до 3.

Если значение **Rbase * S** превышает максимальное значение шага - значение **Rbase(S)** для параметра устанавливается равны максимальному значению шага.

При необходимости – значение параметра **S** можно редактировать в таблице. При этом расчет **Rbase(S)** производится на основе введенного пользователем значения.

Настройки алгоритма РБФ

RBase [>= 1] 1,00

NLayers [>= 1] 5

LambdaNS [>= 0] 0,000

S - масштабный коэф. для параметров

Рекомендуется:

$S = \text{Ср. шаг} + \text{Ср. кв. отклонение} * K$

где K = 2,60

$RBase(S) = S * RBase$

RBase [>= 1]

Начальный радиус. Должен быть ≥ 1 .
 Указывает расстояние от точки, на котором работает метод аппроксимации (какие «дырки» между точками необходимо заделывать).
 О передаче «мелких деталей» функции свидетельствует отношение $RBase/2^{(Nlayers - 1)}$.
 Сглаживаться будут значения с меньшим шагом, чем величина этого отношения.
 Рекомендация: установить значение Rbase – в 2-3 раза выше максимального расстояния между точками по аргументу.

Использовать индивидуальные RBase

Параметр	Ср. шаг	Ср. кв. откл. шага	S	RBase(S)	RBase(S) / $2^{(Nlayers)}$
Power, kg	4145	3320	12353	12353	772.1
Payload	18237	13679	50000	50000	3125
Engine time	4049	4140	14813.34	14813	925.8

Кол-во значений

Размер шага

50000

Рекомендуемые

Принять

OK

Отмена

IDW - мет од обрат ных взвешенных расст ояний.

Используется для сильно зашумленных данных или для больших, с которыми РБФ «не справляется». Используется квадратичная модель узловой функции наименьших квадратов. Подходит для достаточно большого массива зашумленных данных. **В отличие от РБФ используемый массив точек для IDW метода должен иметь параметры примерно одинакового масштаба. Поэтому массивы данных необходимо предварительно «подготовить», используя масштабные коэффициенты для параметров.**

Полученная функция чувствительна к кластеризации исходных данных и наличию выпадающих значений. Лучшие результаты получены с квадратичными моделями, линейные модели не рекомендуется использовать, если вы не уверены, что это то, что вы хотите.

Метод IDW целесообразно применять в следующих случаях:

1. Самое главное показание - когда не можете использовать РБФ из-за больших расходов памяти на решение системы линейных уравнений при построении функции аппроксимации. Как правило, это бывает, если исходный массив размером в миллионы точек сочетается со сравнительно большим RBase или исходный массив в десятки тысяч точек - и RBase, равный размеру всего облака точек.
2. Более редкий случай, когда из-за странных свойств данных (например, точки расположены вдоль траектории оптимизатора) РБФ ведет себя непредсказуемо, и надо построить более менее достоверную модель. Консервативность и "резиновость" IDW в этом случае играет в плюс, а не в минус.

Если сравнивать РБФ и IDW, то РБФ – имеет свойства «упругого металлического листа», а IDW – «неупругого резинового листа». Слабое место IDW подхода в целом – провисание функции по краям массива, если массив

имеет недостаточное количество точек, - "неупругий лист" по краям провисает, IDW методу не хватает опоры со стороны окружающих точек.

Методы одномерной аппроксимации.

- Полиномиальная интерполяция (PI)
- Полиномиальная подгонка (PF)
- Рациональная интерполяция (RI)
- Рациональная подгонка (RF)
- Линейный сплайн (SLI)
- Кубический сплайн (SCI)
- Сплайн Катмулла-Рома (SCRI)
- Сплайн Акимы (SAI)
- Подгонка куб. сплайном (SCF)
- Подгонка сплайном Эрмита (SHF)
- Подгонка штраф. сплайном (SPF)

Полиномиальная интерполяция (PI)

Интерполяция полиномом Лагранжа.

Параметры:

Полином – варианты полинома. Может принимать одно из значений:

- o Обычный - полином Лагранжа.
- o С равноотстоящими узлами - полином Лагранжа с равномерным распределением узлов интерполяции.
- o Чебышева 1 - полином Чебышева первого рода.
- o Чебышева 2 - полином Чебышева второго рода.

Полиномиальная подгонка (PF)

Подгонка полиномами.

Параметры:

- Количество базовых функций - должно быть больше или равно 1.

Рациональная интерполяция (RI)

Интерполяция рациональными функциями.

Параметры:

- D - порядок схемы интерполяции, $0 \leq D \leq N - 1$.
Где N - количество узлов, $N > 0$.
Если вы не уверены, какое значение D выбрать, используйте небольшое значение (от 3 до 5).

Рациональная подгонка (RF)

Подгонка рациональными функциями.

Параметры:

- Количество базовых функций – должно быть больше или равно 2.

Линейный сплайн (SLI)

Интерполяция линейным сплайном.

Кубический сплайн (SLC)

Интерполяция кубическим сплайном.

Параметры:

- . Монотонная кубическая интерполяция сохраняет монотонность интерполируемых данных.

Слайн Кат мулла-Рома (SCRI)

Интерполяция сплайном Катмулла-Рома.

Слайн Акимы (SAI)

Интерполяция сплайном Акимы.

Подгонка кубическим сплайном (SCF)

Подгонка методом наименьших квадратов с использованием кубических сплайнов.

Параметры:

- Количество базовых функций – должно быть больше или равно 4.

Подгонка сплайном Эрмит а (SHF)

Подгонка методом наименьших квадратов с использованием сплайна Эрмита.

Параметры:

- Количество базовых функций – должно быть больше или равно 4.

Подгонка штрафным сплайном (SPF)

Подгонка сглаживающим (штрафным) кубическим сплайном.

Аппроксимирует N разбросанных точек (некоторые из X [] могут быть равны друг другу) кубическим сплайном с M узлами на эквидистантном интервале охвата сетки [min (x, xc), max (x, xc)].

Параметры:

- Количество базовых функций – должно быть больше или равно 4.
- Константа регуляризации – должна быть больше или равна 0. Сглаживает нелинейность сплайн-регрессии. Возможные значения для начала: 0.00001, 0.1, 1

Для оценки качества построенной функции аппроксимации используется коэффициент детерминации- Kdt

При отсутствии связи функции аппроксимации с массивом данных коэффициент детерминации равен нулю, а при полной функциональной связи – единице, т.е. определяет какая доля вариации признака Y учтена в модели аппроксимации и обусловлена влиянием на него факторов.

Коэффициент детерминации (R^2) определяется следующим образом:

$$R^2 = \frac{\text{Объясненная сумма квадратов}}{\text{Общая сумма квадратов}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = 1 - \frac{\sum_{i=1}^n e_i^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Шаг 3 Сохранение аппроксимационной модели

Сохранение проекта аппроксимации осуществляется командой «Сохранить» - .

Сохранение созданной функции аппроксимации в виде исполняемого файла (*.exe) осуществляется командой «Создать аппроксимационную модель» - .

Данный исполняемый файл может быть использован как самостоятельная расчетная модель в программе оптимизации IOSO либо для иных целей.

При первом запуске Исполняемого файла (exe-- модель запускается без параметра *rw* и без входного файла данных) генерируется шаблон входного и выходного файлов данных. Название файлов формируется из названия модели добавлением *_in* и *_out* с расширением *.dat*. Эти названия должны строго соблюдаться.

При использовании в командной строке Исполняемого файла ключа «*-rw*» или «*/rw*» запуск происходит в режиме "run and wait". В этом режиме Модель запускается и висит в процессах, ожидая появления входного файла. Когда появляется новый входной файл (определяется по времени создания файла), модель рассчитывает значения функции и записывает их в новый выходной файл.

Результаты расчёта также выводятся в консоль Windows (поскольку модель висит в памяти, консоль продолжает работать).

При запуске Модели без ключа - модель будет выгружаться после завершения расчета.
