УДК 519.6

АППРОКСИМАЦИЯ ЭКСПЕРИМЕНТАЛЬНЫХ ТОЧЕК ПОЛИНОМИАЛЬНОЙ ФУНКЦИЕЙ МЕТОДОМ НАИМЕНЬШИХ КВАДРАТОВ

Швалёва Анна Викторовна

канд. пед. наук
Национальный исследовательский технологический университет
«МИСиС» (филиал), Новотроицк

author@apriori-journal.ru

Аннотация. В статье приведены основные расчетные формулы, изложен алгоритм исследования зависимости механических свойств стали от химического состава с помощью программ **Mathcad**, **Excel**.

Ключевые слова: аппроксимация; метод наименьших квадратов; полином; линейная; квадратичная; кубическая парные регрессии; множественная регрессия.

APPROXIMATION OF EXPERIMENTAL POINTS BY POLYNOMIAL FUNCTION BY A METHOD OF THE SMALLEST SQUARES

Shvaleva Anna Viktorovna

candidate of pedagogic sciences
National University of Science and Technology «MISIS» (branch), Novotroitsk

Abstract. The basic settlement formulas are given in article, the algorithm of research of dependence of mechanical properties of steel from a chemical composition by means of the programs Mathcad, Excel is stated.

Key words: approximation; method of the smallest squares; polynom; linear; square; cubic pair regressions; multiple regression.

В последние годы важным аспектом математического образования становится владение компьютерными технологиями. Влияние компьютера приводит к необходимости частичного пересмотра структуры и содержания курса математики, иной расстановке акцентов в задачах, требующих объемных вычислений. Особенно актуальным является использование компьютерных технологий при изучении «Математической статистики», для обработки полученных экспериментальных данных. Рассмотрим возможности компьютерных программ в описании (аппроксимации) зависимости, представленной в виде набора чисел аналитической функцией. В качестве примера рассмотрена зависимость временного сопротивления от химического состава и размера листа стали марки С345.

Аппроксимация зависимости, представленной в виде набора чисел аналитической функцией, решается в разделе математической статистики «Регрессионный анализ». Наиболее часто используется аппроксимация с помощью полинома степени *n*:

$$y = f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + ... + a_n x^n ,$$

причем, коэффициенты полинома a_0, a_1, a_2, $\dots a_n$ подбираются таким образом, чтобы кривая, описываемая полиномом, проходила максимально близко ко всем точкам, аппроксимируемым этим полиномом. Эта близость достигается с использованием метода наименьших квадратов (сумма квадратов отклонений фактических значений результативного признака (y) от значений полинома ($\overline{y}$) должна быть наименьшей (1)):

$$S = \sum_{i=1}^{n} (y_i - \overline{y})^2 \to \min \qquad (1)$$

Для аппроксимации с помощью полинома, нужно чтобы выполнялось условие: $n<m-1$ (m – количество аппроксимируемых точек). Так, через три точки $m=3$ можно провести только прямую – полином первой степени, через 4 точки – полином первой или второй степени и т.д.

Для отыскания коэффициентов полинома, с учетом (1) условия необходимо все частные производные от *S* по коэффициентам a_0, a_1, a_2, $\ldots a_n$ приравнять к нулю. Решая систему линейных алгебраических уравнений порядка *(n+1)* найдем (*n+1*) коэффициентов полинома:

$$\begin{cases} \dfrac{\partial S}{\partial a_0} = 0 \\ \dfrac{\partial S}{\partial a_1} = 0 \\ \ldots \\ \dfrac{\partial S}{\partial a_n} = 0 \end{cases}$$

В данной работе была рассмотрена аппроксимация только полиномами первой, второй и третьей степенями.

Различают регрессию двух видов: парную и множественную. Разница между ними в количестве независимых переменных. В природе имеет место исключительно множественная регрессия, так как нельзя ограничить внешнее влияние на какое-то явление строго одним фактором.

Рассмотрим построение парных и множественных регрессионных моделей зависимости функции временного сопротивления от химического состава и размера листа.

Временное сопротивление *(FVS)* характеризует максимальное напряжение, предшествующее разрушению образца (предел прочности). Иными словами, временное сопротивление (предел прочности) – это максимальное механическое напряжение, выше которого происходит разрушение материала, подвергаемого деформации; предел прочности при растяжении обозначается $\sigma_в$ и измеряется в килограммах силы на квадратный сантиметр (кгс/см^2). В лаборатории были получены результаты испытаний временного сопротивления, химического состава и размера листа, объём выборки составил 366 результата замеров (рисунок 1).

Коэффициенты полинома первой степени $y = f(x) = a_0 + a_1 x$; второй $y = f(x) = a_0 + a_1 x + a_2 x^2$ и третей $y = f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$, для данного исследования можно рассчитать, например, с помощью применения функций таких программ, как **Mathcad** и **Excel**.

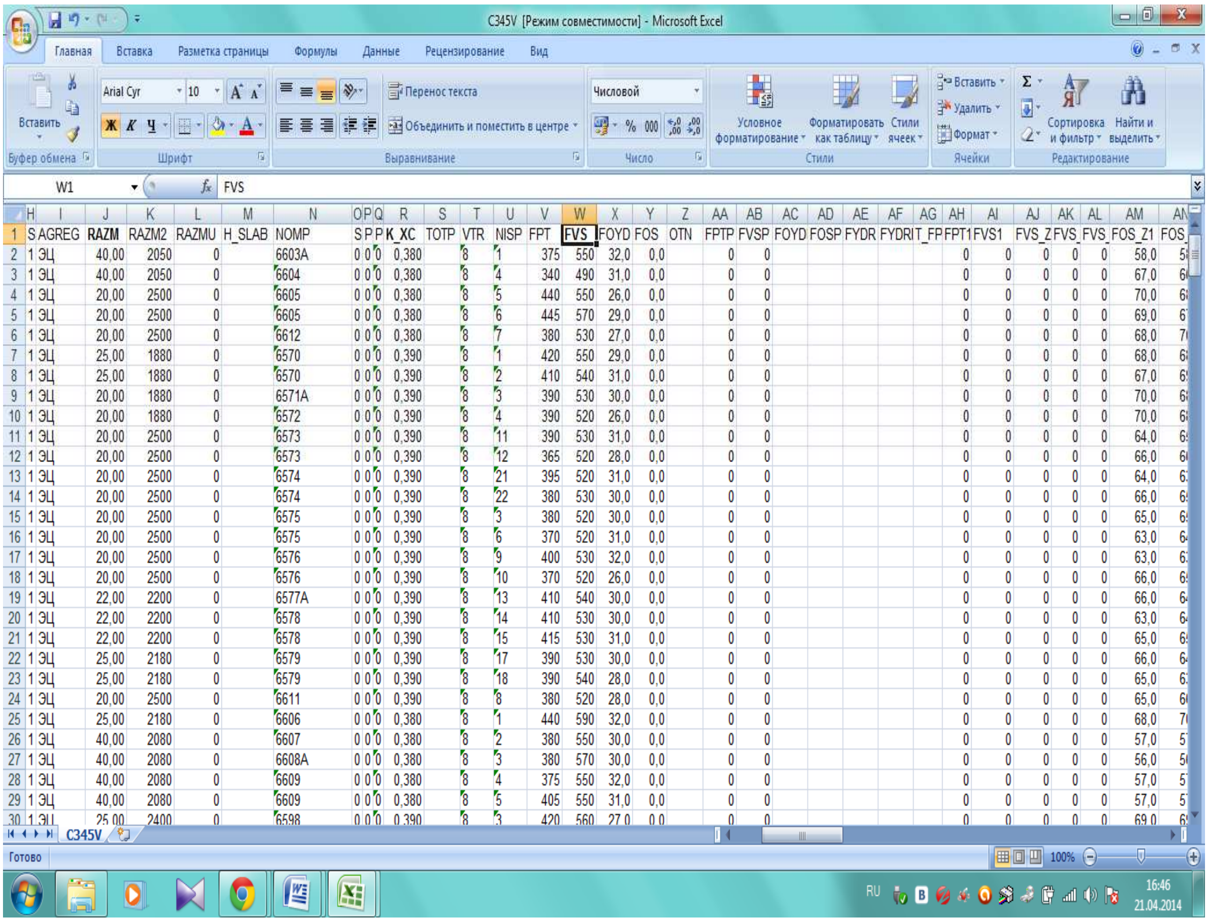

Рис. 1. Выборка

Mathcad – универсальный математический пакет, предназначенный для выполнения инженерных и научных расчетов. Основное преимущество пакета перед типичными языками программирования – естественный математический язык, на котором формулируется решаемая задача.

Выполним построение полинома первой степени (парной регрессионной модели) $y = f(x) = a_0 + a_1 x$ (в качестве зависимой переменной *у* – временное сопротивление, независимой переменной *х* – коэффициент химического состава). Для начала в отдельной папке создаем два текстовых документа. Например, в блокнотах записываем массивы *X* и *У*. В файле Mathcad с помощью команды READPRN выводим массивы *X* и *У*. Далее используем команду SLOPE, которая возвращает коэффициент при *х* в линейной регрессионной модели, а так же INTERCEPT, которая возвращает свободный член в линейной регрессионной модели. Получаем полином первой степени (парную линейную регрессионную модель) (рисунок 2).

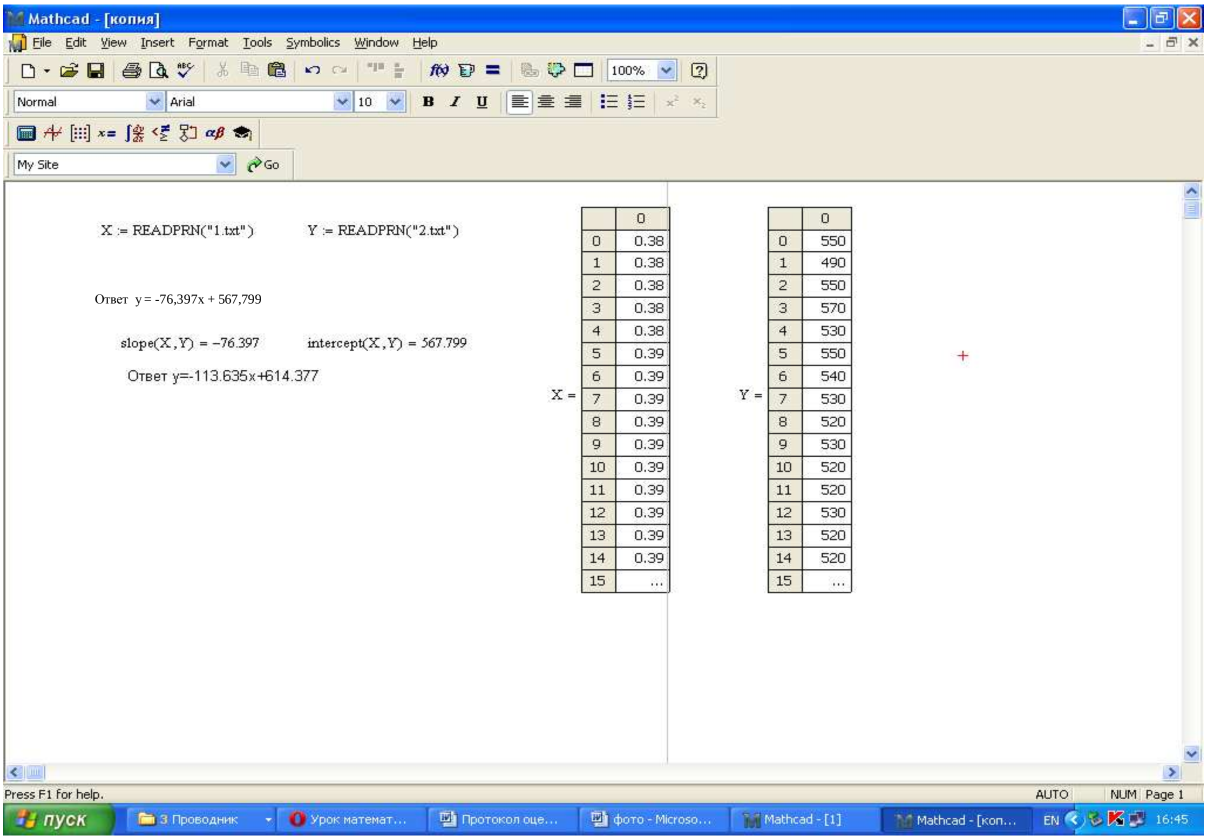

Рис. 2. Аппроксимация зависимости временного сопротивления от коэффициента химического состава (полином первой степени)

Полином второй и третьей степеней определим с помощью команды REGRESS, которая возвращает вектор, последние *l* координат которого – коэффициенты степенной регрессионной модели степени *l – 1* (рисунок 3).

Microsoft Excel – является широко распространенной компьютерной программой, с помощью которой производятся расчеты, составляются таблицы и диаграммы, вычисляются простые и сложные функции.

Для того чтобы построить полином первой степени (парную линейную регрессионную модель) зависимости временного сопротивления от химического состава нужно записать два массива значений зависимой и независимой переменных.

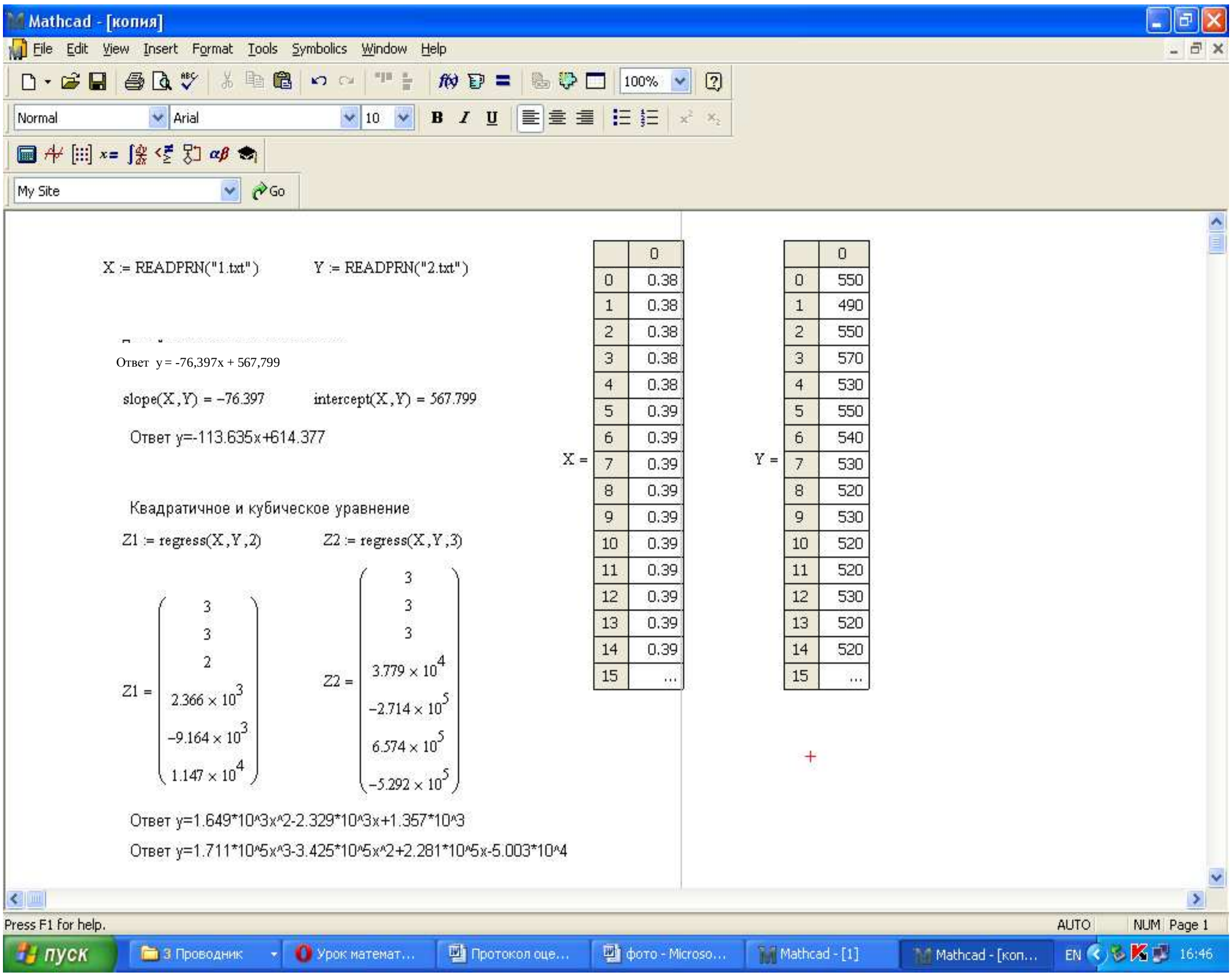

Рис. 3. Аппроксимация зависимости временного сопротивления от коэффициента химического состава (полиномы второй и третей степеней)

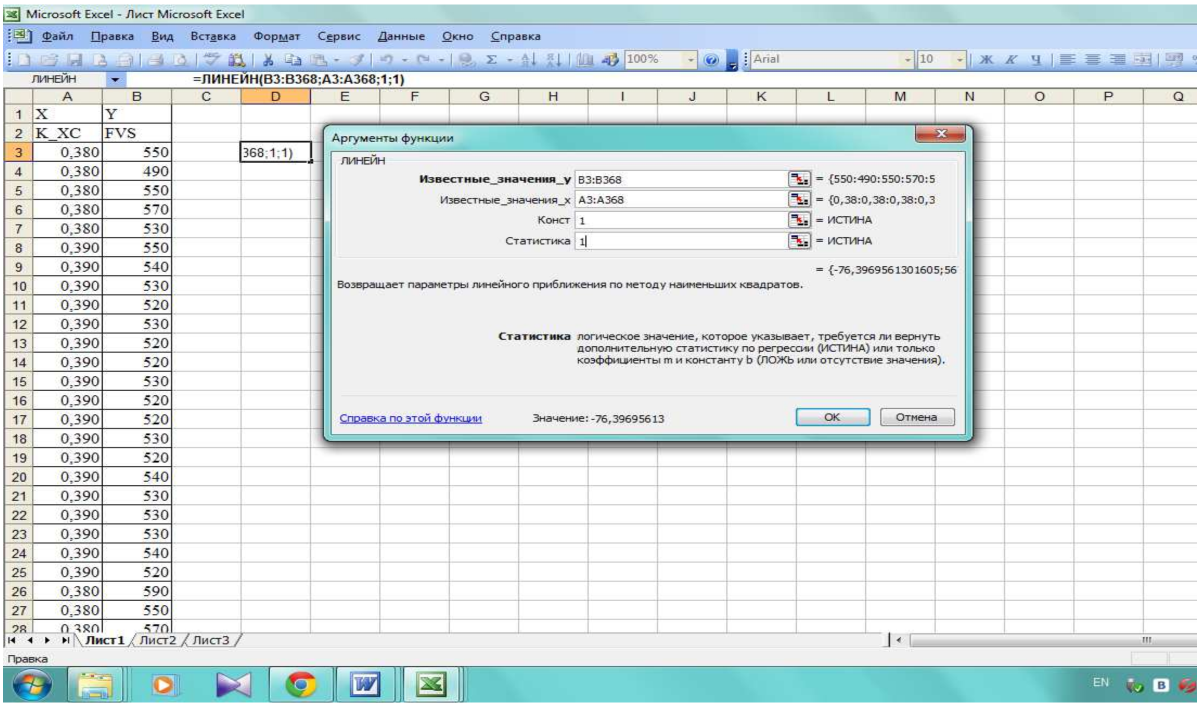

Рис. 4. Использование функции ЛИНЕЙН

Для того чтобы определиться с формой связи (линейная, квадратичная и т.д.) можно построить точечную диаграмму. Затем воспользуемся функцией ЛИНЕЙН (рисунок 4).

Далее выделяем диапазон ячеек с двумя столбцами и пятью строками, после чего получаем следующие значения (рисунок 5).

Значения дополнительной статистики соответствуют следующим характеристикам (таблица 1).

Таблица 1

Дополнительная статистика

Коэффициент a_1 линейной модели	Коэффициент a_0 линейной модели
Стандартная ошибка коэффициента a_1	Стандартная ошибка коэффициента a_0
Коэффициент детерминации	Стандартная ошибка для *y* ($s_{a\partial}$)
F-наблюдаемое ($F_э$)	Степени свободы ($k_{a\partial} = n - l$)
Регрессионная сумма квадратов $((\Delta\overline{Y}_i)^2)$	Остаточная сумма квадратов $((\Delta Y_i)^2)$

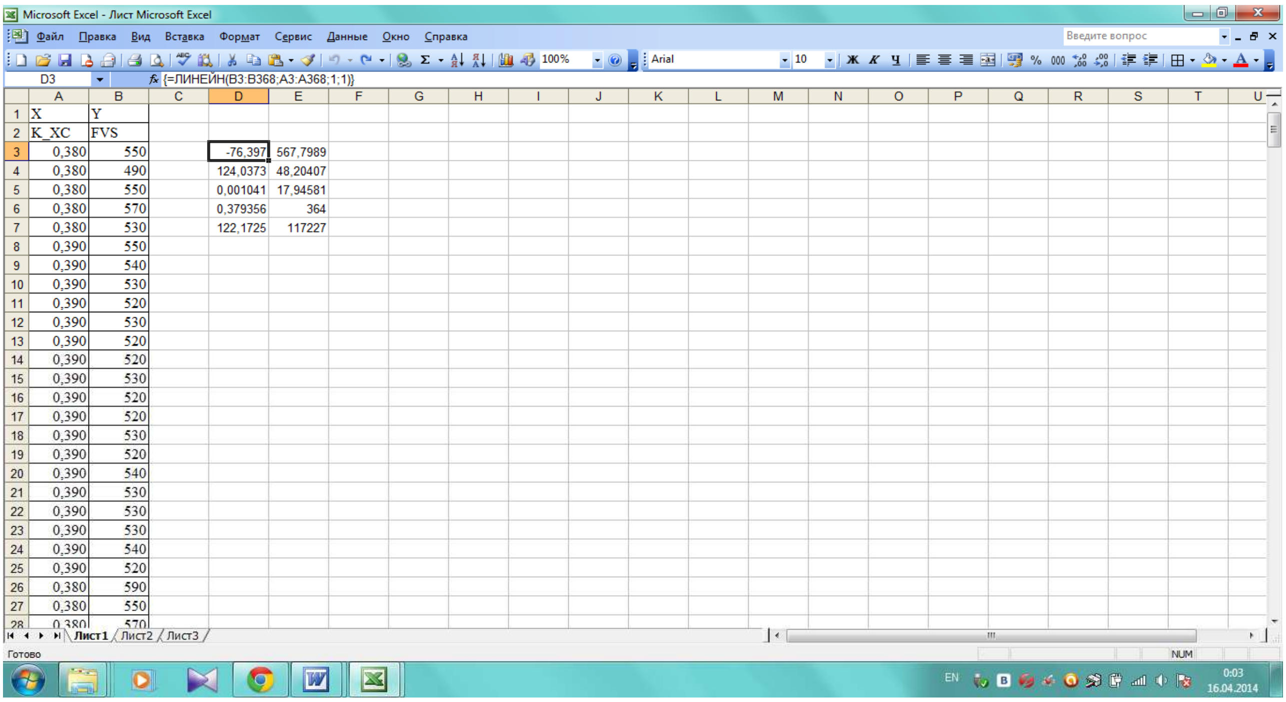

Рис. 5. Дополнительная статистика

Коэффициент детерминации R^2 равен *0,1 %,* то есть построенная при таких условиях регрессионная модель зависимости временного сопротивления от коэффициента химического состава имеет низкое практическое значение. Коэффициент a_1 линейной модели равен -76,397. Стандартная ошибка коэффициента a_1 равна 124,0373.

Для того чтобы построить *множественную регрессионную модель* зависимости временного сопротивления от химического состава и технологических показателей нужно записать три (и более) массива зависимых и независимых переменных. При построении множественной регрессии требуется «Пакет анализа», который является довольно мощным инструментом в помощь аналитику. Этот инструментарий, помимо всего прочего, позволяет рассчитывать параметры регрессии по тому же методу наименьших квадратов, всего в несколько кликов. В активном окошке инструмента Анализа данных из списка возможностей ищем и выбираем Регрессия (здесь нужно указать интервалы исходных данных, а именно описываемого параметра (y) и влияющих на него факторов (x_i)). Данные представлены на рисунке 6.

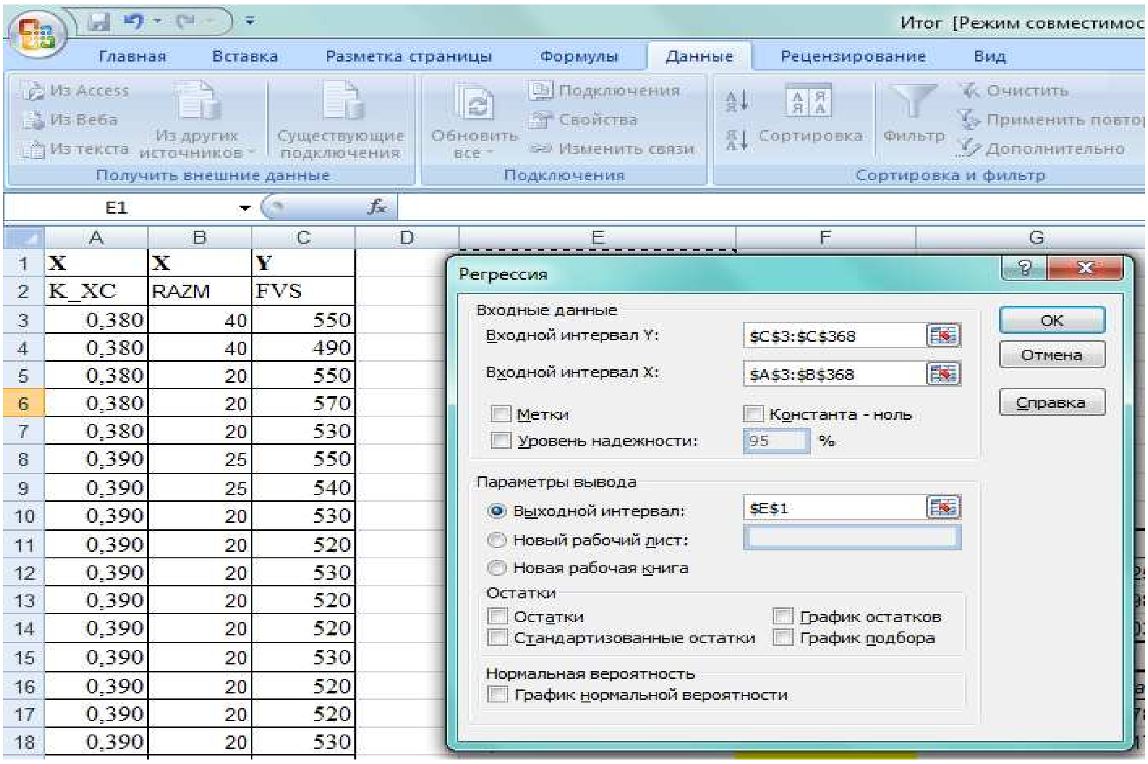

Рис. 6. Вычисление параметров регрессионной модели

Эти расчеты имеют следующий вид (рисунок 7).

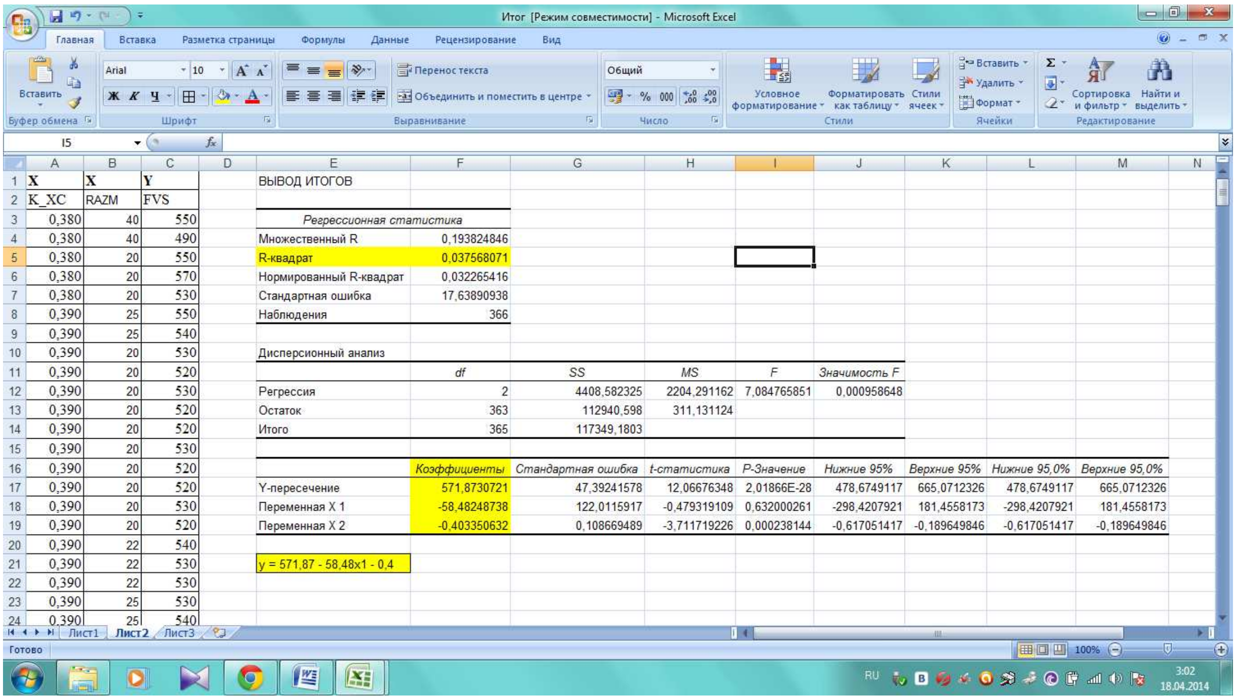

Рис. 7. Вывод значений

Ключевые ячейки залиты желтым цветом, именно на них нужно обращать внимание в первую очередь. Остановимся на них.

0,04 – это **R^2 –** коэффициент детерминации, показывающий что на 4 % расчетные параметры модели, то есть сама модель, объясняют зависимость изменения изучаемого параметра – *y* от исследуемых факторов – x_i. Чем ближе коэффициент детерминации к единице, тем ближе модель к эмпирическим наблюдениям.

Теперь перейдем к коэффициентам модели:

1) **571,87** – это коэффициент, который показывает какой будет *Y* в случае, если все используемые в модели факторы будут равны 0, подразумевается что это зависимость от других неописанных в модели факторов;

2) **-58,48** – коэффициент при x_1. Знак минус показывает, что это влияние отрицательно, то есть чем больше коэффициент химического состава, тем меньше временное сопротивление.

3) -**0,4** – коэффициент при x_2. Знак минус показывает, что это влияние отрицательно, то есть чем больше коэффициент химического состава, тем меньше временное сопротивление. Соберем рассчитанные коэффициенты в модель:

$$y=571,87-58,48x_1-0,4x_2$$

Собственно, это и есть линейная множественная регрессионная модель, которая для исходных данных, используемых в нашем исследовании, выглядит именно так.

Рассмотрение подобных задач при изучении дисциплин математического цикла, позволяет студентам демонстрировать многие важные приложения математики. Студенты начинают понимать необходимость изучения математических методов не только при изучении специальных дисциплин, в которых рассматривается решение прикладных задач, но значительно раньше.