**Кригинг**

Кригинг – группа геостатистических методов для интерполяции значений случайного поля (например, высота ландшафта, как функция географических координат) в точку, в которой поле не задано, путём наблюдения величин в соседних точках, где оно известно.

Пусть имеется набор точек (векторов) , заданных в евклидовом пространстве . Так же известны значения некоторой функции  в этих точках. Требуется построить интерполяционную функцию , достаточно близкую к неизвестной . Значения неизвестной функции в данном наборе точек получаются, как правило, путем измерения. В силу этого, они могут рассматриваться как случайные величины (так как, строго говоря, разные измерения в одной и той же точке дадут различный результат), а сама функция полагается случайным процессом.

Значения интерполяционной функции в произвольной точке  будем искать как линейную комбинацию следующего вида:



где  – некоторые, пока неизвестные нам весовые функции.

Предполагаем, что интерполяционная функция  удовлетворяет следующим условиям:

1. Математическое ожидание от интерполяционной функции должно быть равно математическому ожиданию от интерполируемой функции.



1. Дисперсия разности интерполяционной и интерполируемой функций, варьируемая по весам, достигает минимального значения, т.е.:

.

Иными словами, в силу пункта 1.:



1. Математическое ожидание интерполируемой функции считается постоянным:

,

что, в силу п.1, влечет за собой следующее соотношение:

,

а значит .

Таким образом, в силу п.п.1-3, для определения величин , имеем следующую задачу на условный экстремум:

, при условии .

Введем множитель Лагранжа , тогда наша задача преобразуется к задаче на безусловный экстремум:



Определим функцию Лагранжа:

. (1)

Первое слагаемое в правой части последнего соотношения можно преобразовать к следующему виду:

 (2)

Для последнего преобразования мы воспользовались формулой:



Дифференцируя функцию Лагранжа (1) по всем переменным , учитывая при этом (2), после преобразований, получим систему линейных алгебраических уравнений:

 (3)

4. К п.п.1-3 добавим еще одно предположение:



Здесь  – некоторая, пока неизвестная нам функция, называемая вариограммой. Перепишем систему (3) с учетом введенной замены:

 (4)

Система (4) определяет неизвестные величины

,

при условии, что вариограмма  задана. Система (4) решается методом исключения Гаусса с выбором главного элемента. В качестве вариограмм предлагаются следующие функции:

* 
* 
* 
* 
* 

**Исследование метода Кригинга**

Для исследования метода Кригинга проведем ряд численных экспериментов. В качестве интерполируемой функции возьмем положительную часть сферы радиуса . Для задания точек сферы воспользуемся сферическими координатами:



где , ; при этом  и  изменяются в заданных границах с шагами  и  соответственно. Вычисляя значения ,  и  в точках  и , где , , , , мы задаем точки  и соответствующие значения  (которые следует рассматривать, как значения интерполируемой функции в указанных точках).

Заданную сферу будем интерполировать в «серединах» заданных ячеек, т.е. в точках , получаемых при следующих значениях  и :

, .

Погрешность интерполяции будем вычислять по формуле

,

где  - искомое значение в заданной точке , а  - вычисленное.

Проведем несколько численных экспериментов и попытаемся определить, какая из предложенных вариограмм показывает лучшие результаты.

В таблице 1 приведены результаты численных экспериментов с вариограммой  ( - это число точек в окрестности искомой точки). При проведении численных экспериментов с вариограммой  было замечено, что изменение параметров вариограммы практически не влияет на показываемые результаты.

Таблица 1. Вариограмма .

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | 0.814567343191323 | 0.0105384201914855 |
|  | 0.717561338332021 | 0.273434702549126 |
|  | 0.727375125629796 | 0.985873730380073 |

В таблице 1 на пересечении строк и столбцов приведены погрешности, получаемые при интерполяции сферы при указанных значениях параметров. При увеличении числа узлов, в которых задана сфера, результаты улучшаются для  и . Лучший результат достигается при  и .

В таблице 2 приведены результаты численных экспериментов для вариограммы . При проведении численных экспериментов было замечено, что на результаты, показываемые при использовании данной вариограммы, существенное влияние оказывает параметр . Причем при увеличении числа узлов в окрестности искомой точки нужно уменьшать значение параметра , чтобы погрешность уменьшалась. Оптимальным оказался следующий способ задания вариограммы .

Таблица 2. Вариограмма .

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | 0.822619936277198 | 0.01707388632299 |
|  | 0.818398814746531 | 0.820504864721681 |
|  | 1.61850745209773 | 1.61850745209773 |

Вариограмма , так же как и предыдущая, лучший результат показывает при использовании четырех узлов в окрестности искомой точки.

На результаты, показываемые при использовании вариограммы , также влияет параметр . Однако не удалось выбрать оптимальное значение параметра , так как каждый раз при изменении набора узлов, в которых задана интерполируемая функция, а также числа точек, используемых для вычисления искомой точки, приходилось подбирать параметр  так, чтобы погрешность при этом была минимальной. В таблице 3 приведены результаты численных экспериментов с вариограммой . В таблице 3 рядом с погрешностью указано соответствующее значение параметра .

Таблица 3. Вариограмма .

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | 0.81457934059104, | 0.0105384635521641, |
|  | 0.717561829340271, | 0.248114650818209, |
|  | 0.727376416887195, | 0.546670440260499, |

Из таблицы 3 видно, что для достижения лучшего результата нужно использовать четыре узла в окрестности искомой точки.

В таблице 4 приведены результаты численных экспериментов с вариограммой . При использовании данной вариограммы также необходимо варьировать значение параметра , чтобы улучшить результаты интерполяции. Остальные параметры вариограммы практически не влияют на точность. Также как и с предыдущими вариограммами лучшие результаты достигаются при использовании четырех узлов для вычисления искомого значения в требуемой точке.

Таблица 4. Вариограмма .

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | 0.814805752607466, | 0.0105392868092462, |
|  | 0.71853743408321, | 0.267760840771028, |
|  | 0.72995115535617, | 0.542606940264788, |

Наконец, проведем ряд численных экспериментов с использованием вариограммы . Результаты приведены в таблице 5.

Таблица 5. Вариограмма .

|  |  |  |
| --- | --- | --- |
|  |  |  |
|  | 0.397571488353726, | 0.0137431374647825, |
|  | 0.418754700739201, | 0.783714929183565, |
|  | 1.18603074399113, | 0.500200219125169, |

Видно, что лучшие результаты достигаются при использовании параметра . При работе с данной вариограммой также приходится варьировать значение параметра , чтобы улучшить получаемые результаты. Другие параметры вариограммы не оказывают сильного влияния на точность интерполяции.

Сравнивая результаты, показываемые при использовании каждой из вариограмм, можно заключить, что оптимальным является использование вариограммы . Во-первых, погрешности, получаемые при интерполяции, практически не отличаются от погрешностей, показываемых другими вариограммами. Во-вторых, использование вариограммы  не требует специального выбора параметров для получения хороших результатов.

Далее приведены рисунки, демонстрирующие интерполяцию полусферы с использованием метода Кригинга.

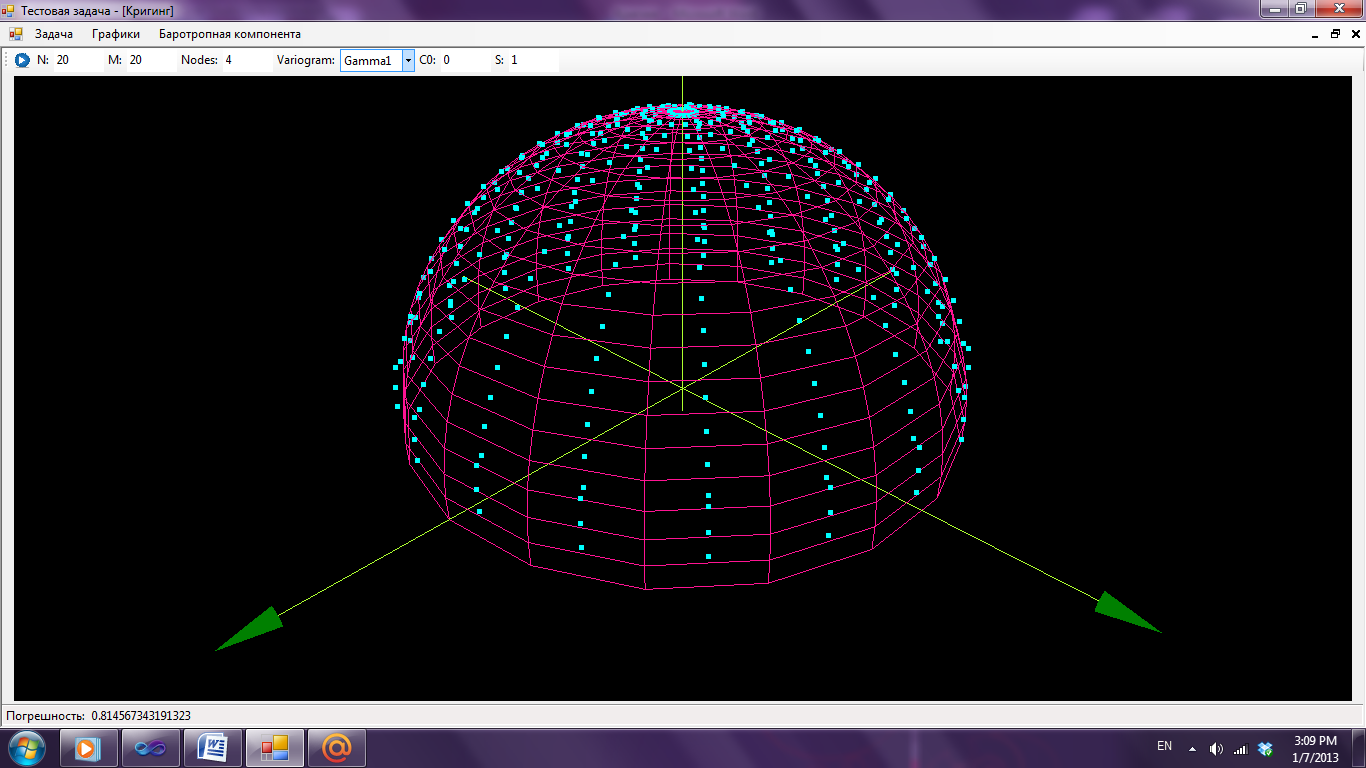


Рисунок 1. Вариограмма ,,.

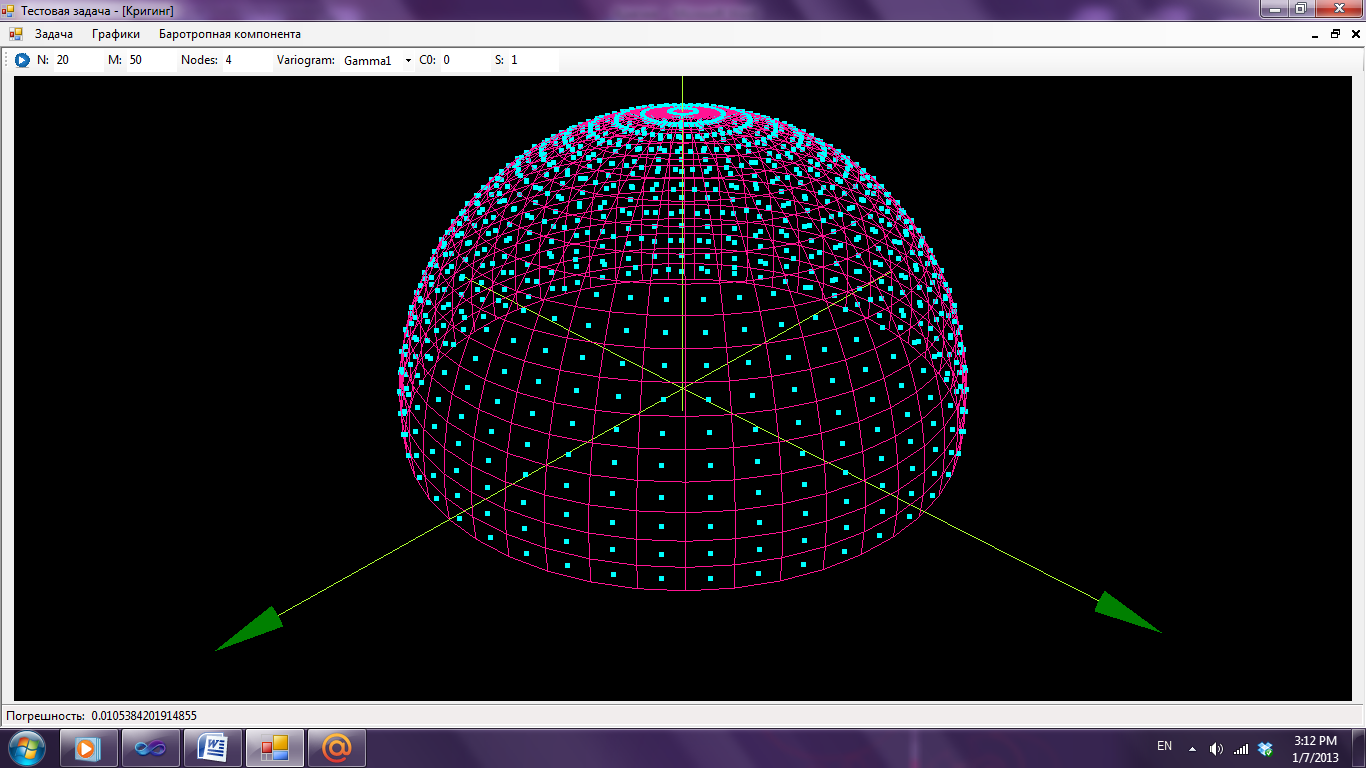


Рисунок 2. Вариограмма ,,.

**Интерполяция дна озера Иссык-Куль методом Кригинга**

Далее будет показано использование метода Кригинга для восстановления дна озера Иссык-Куль.

Имеются данные о рельефе дна озера Иссык-Куль, полученные в Институте водных проблем НАН Кыргызской Республики. Эти данные представляют собой массив точек , где  и  задают точку на зеркале озера Иссык-Куль, а  представляет собой глубину в этой точке. При этом крайняя левая точка имеет абсциссу , а крайняя нижняя точка имеет ординату . На рисунках 3 и 4 приведено озеро Иссык-Куль, нарисованное по указанным данным. При этом для лучшего визуального восприятия глубина озера пропорционально увеличена в 20 раз. Требуется по предоставленным данным восстановить рельеф озера Иссык-Куль (интерполировать значения глубин на некоторой достаточно плотной сетке).

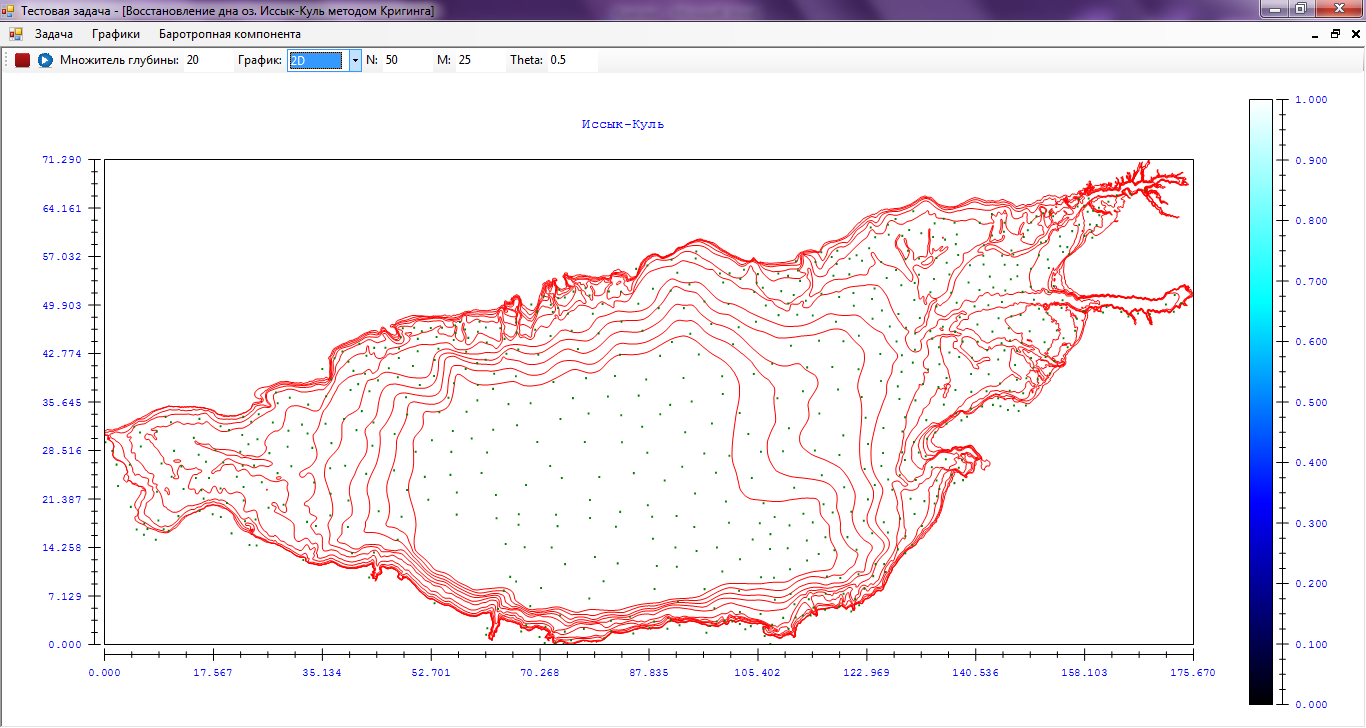


Рисунок 3. Озеро Иссык-Куль, батиметрия 2D.

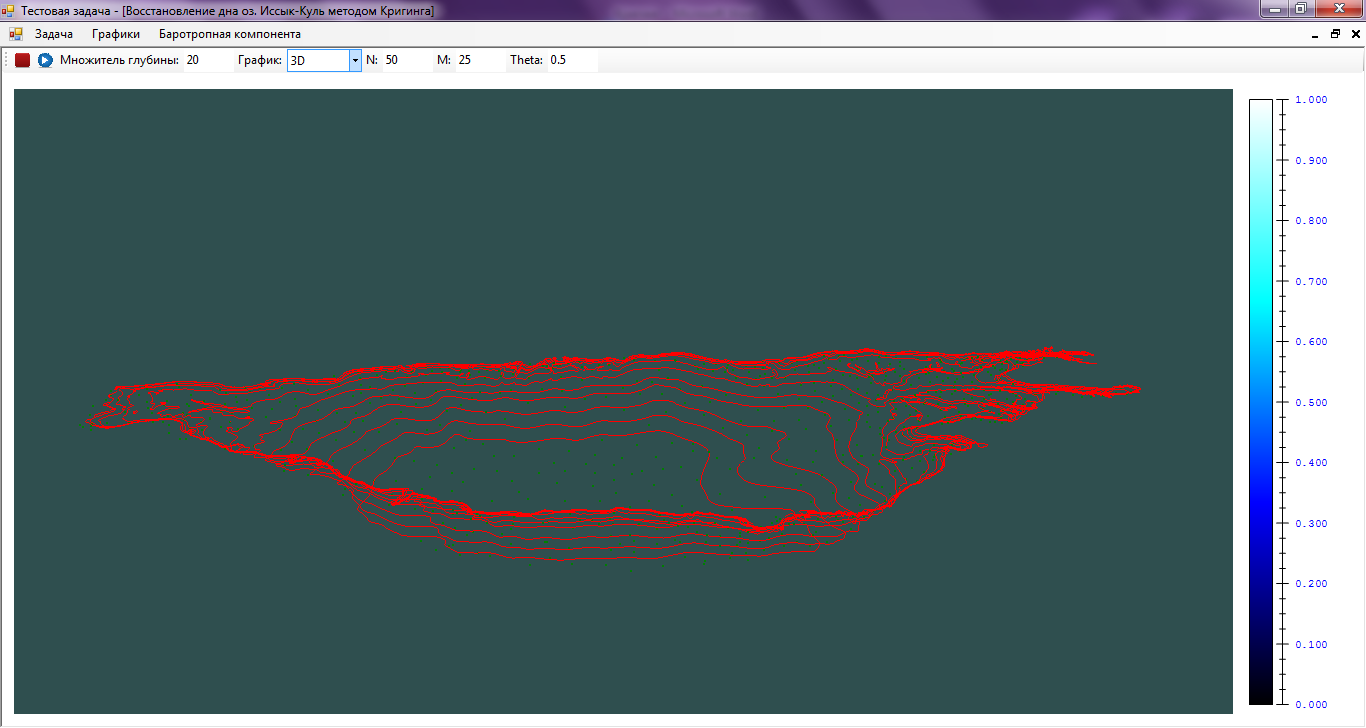


Рисунок 4. Озеро Иссык-Куль, батиметрия 3D.

Для восстановления рельефа дна озера Иссык-Куль разработан алгоритм, на вход которого подаются указанные выше данные, число сеточных узлов в горизонтальных направлениях (запад-восток (параметр *N*) и юг-север (параметр *M*)), а также параметр сглаживания . Разработанный алгоритм включает в себя следующие этапы (каждый из этапов далее будет подробно описан):

* построение прямоугольной сетки, аппроксимирующей зеркало озера Иссык-Куль;
* вычисление глубины методом Кригинга в середине каждой из ячеек, полученной на предыдущем шаге;
* применение фильтра сглаживания к полученным значениям глубин.

Было разработано программное приложение, реализующее указанный алгоритм. Приложение было разработано в среде разработки Visual Studio 2010 на языке программирования C#.

**Аппроксимация зеркала.** Имеющиеся данные по рельефу дна позволяют определить прямоугольник , в который вписано озеро Иссык-Куль (см. рисунок 3, например), где  – абсцисса крайней правой точки, а  – ордината крайней верхней точки. Данный прямоугольник разбивается на прямоугольные сеточные ячейки  (*i = 0, 1, …, N - 1*, *j = 0, 1, …, M - 1*) с размерами

, ,

и координатами левых нижних углов

, .

Далее исходный массив натурных данных распределяется по ячейкам, так что каждой ячейке ставится в соответствие набор точек из указанного массива данных, при этом если ячейка содержит хотя бы одну точку с нулевой глубиной (), то данная ячейка помечается, как ячейка берега. Для этого для каждой точки  массива данных мы определяем координаты  ячейки, в которую данная точка будет включена:

,

где  – операция округления до меньшего целого. **Замечание:** из-за операций округления (при работе с числами с плавающей точкой) при вычислении координат ячеек для правой крайней и для верхней крайней точек могут быть получены координаты  и , соответственно, которые выходят за границы возможных координат; в этом случае для этих точек в качестве координат нужно брать значения *N-1* и *M-1*.

После того, как все ячейки берега определены, определяются ячейки, не лежащие внутри области озера Иссык-Куль. Для этого используется алгоритм Заливка (<http://ru.wikipedia.org/wiki/Заливка>). Данный алгоритм используется в компьютерной графике для заполнения областей определенным цветом. В нашем случае этот алгоритм используется для определения области, которая является внешней по отношению к озеру Иссык-Куль. Для этого к каждой ячейке, лежащей на границе сеточной области, применяется алгоритм Заливки, который «закрашивает» (помечает ячейки, как внешние) ту часть ячеек, которая может быть «залита» вместе с рассматриваемой ячейкой. После того, как алгоритм Заливки будет применен к каждой ячейке границы сеточной области, все ячейки можно разбить на три категории: внешние ячейки («закрашенные» заливкой), ячейки берега и не закрашенные ячейки (внутренние ячейки, т.е. те, которые лежат внутри области, ограниченной берегом).

Ячейки берега вместе с внутренними ячейками образуют зеркало озера Иссык-Куль. На рисунке 5 приведена аппроксимация зеркала озера Иссык-Куль с использованием описанного выше алгоритма.

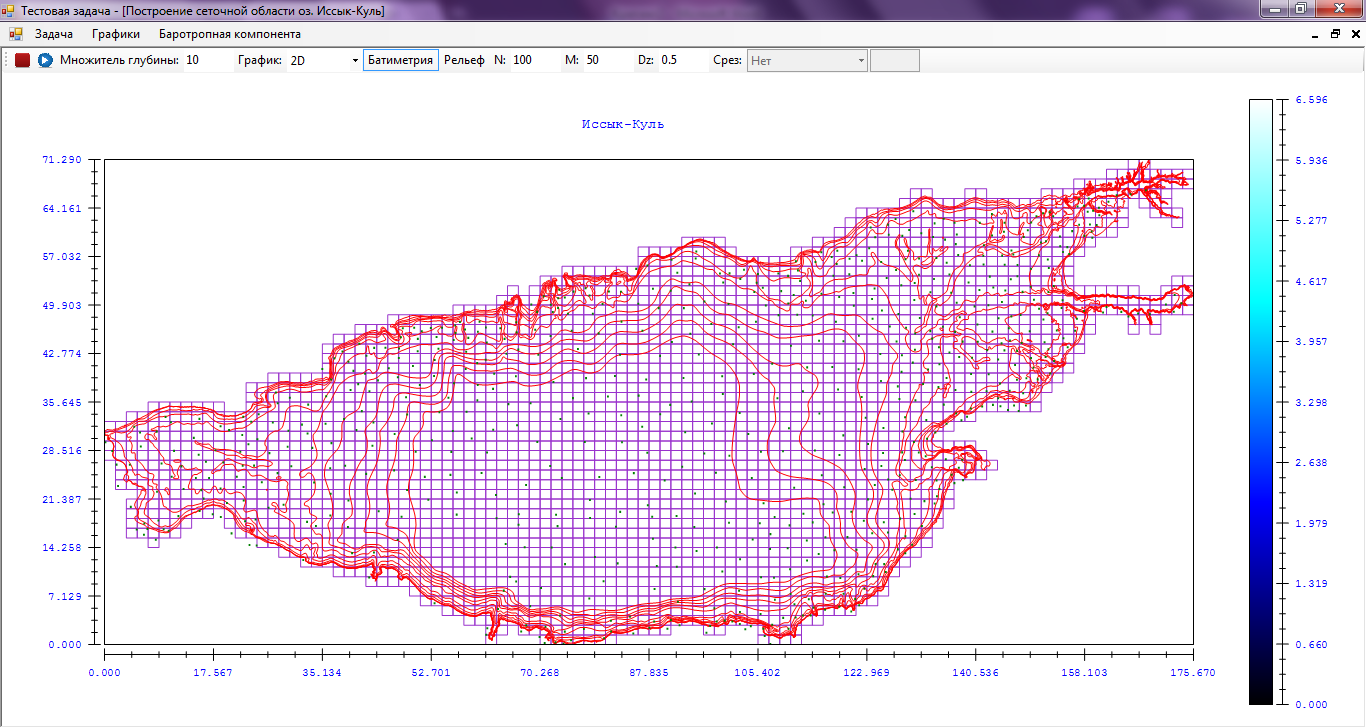


Рисунок 5. Аппроксимация зеркала, *N=100*, *M=50*.

**Восстановление дна методом Кригинга.** Следующим шагом является вычисление глубины в середине каждой сеточной ячейки зеркала озера Иссык-Куль. Для ячейки  глубина вычисляется в точке с координатами . Для вычисления глубины в указанной точке используется метод Кригинга. Выше было исследовано, что оптимальной вариограммой для метода Кригинга является вариограмма , и в окрестности искомой точки нужно использовать четыре узла. Именно такие параметры метода Кригинга и будут использоваться. В качестве указанных четырех узлов берутся ближайшие к точке  четыре точки. Для определения ближайших четырех точек используется следующий алгоритм:

* проверяем, имеются ли в текущей ячейке с координатами *(i, j)* точки; если имеются, то определяем расстояние от каждой точки до точки ; выбираем четыре ближайших, если точек четыре или более; если же точек меньше четырех, то аналогичным образом анализируем ячейки, находящиеся на один шаг левее, правее, выше и ниже рассматриваемой ячейки; продолжаем анализ окружающих ячеек до тех пор, пока не будут найдены четыре ближайших точки;
* после того, как четыре точки найдены, нужно убедиться, что данные точки являются ближайшими; для этого рассматриваются ячейки, которые окружают те ячейки, на которых был остановлен поиск на первом шаге; если в данных ячейках есть точки, которые находятся ближе к рассматриваемой точке, то они заменяют те точки из четырех найденных точек, которые находятся дальше; этот шаг повторяется до тех пор, пока не встретится уровень окружающих ячеек, в которых все точки находятся дальше от рассматриваемой точки, чем четыре найденные точки.

**Замечание:** при реализации метода Кригинга было замечено, что в некоторых случаях он выдает отрицательную глубину (мы считаем, что глубина положительна, т.е. при описании озера ось *Oz* направлена вниз). В этих случаях в искомой точке берется нулевая глубина.

Рисунки 6 и 7 демонстрируют восстановленное дно озера Иссык-Куль методом Кригинга (без сглаживания). Были выбраны параметры *N=300*, *M=300*. Глубина озера была пропорционально увеличена в 20 раз.

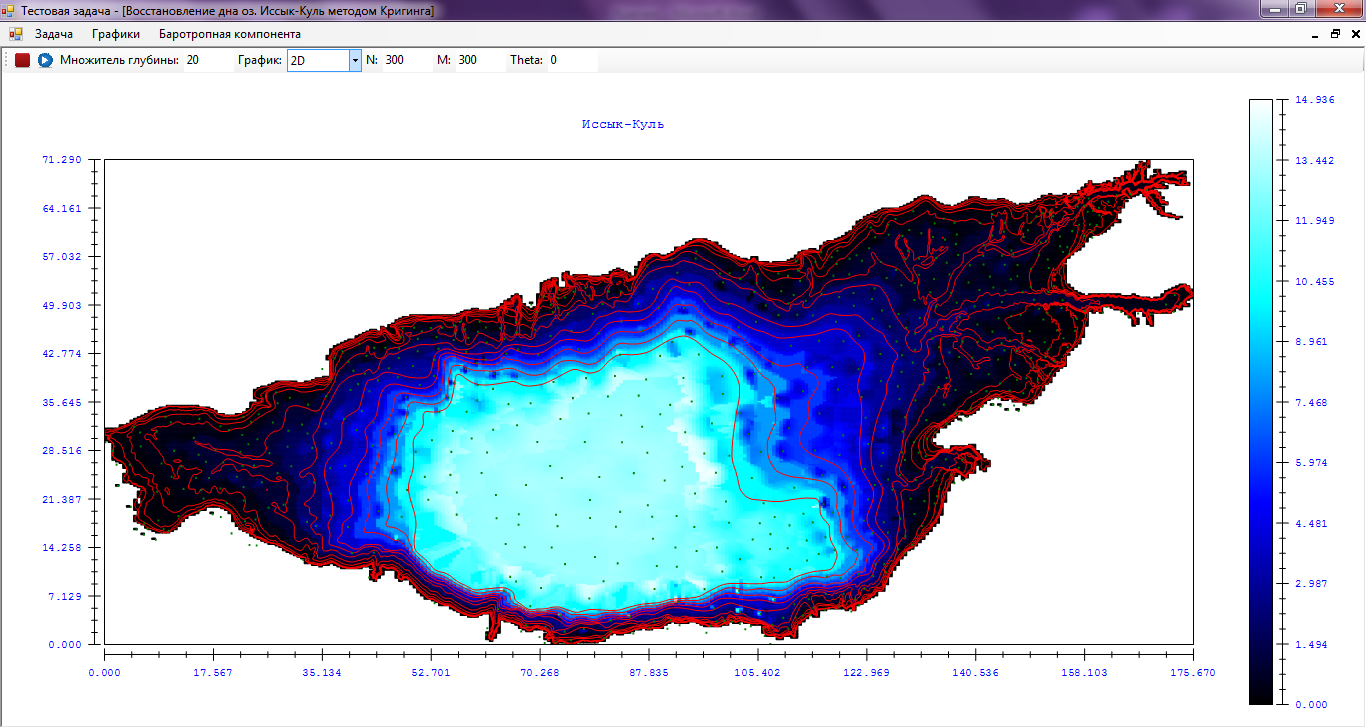


Рисунок 6. Восстановленный рельеф озера Иссык-Куль, 2D.

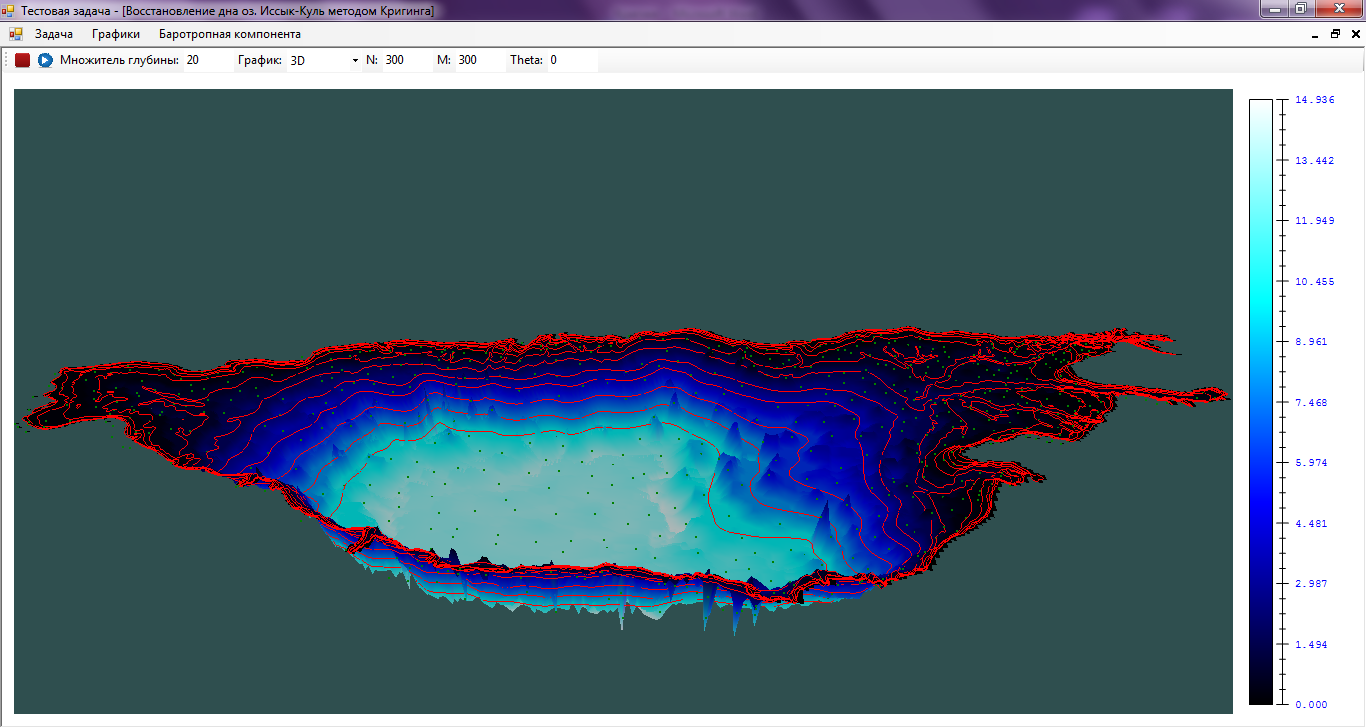


Рисунок 7. Восстановленный рельеф озера Иссык-Куль, 3D.

**Сглаживание.** После того как в каждой ячейке рассчитана глубина, к полученным данным применяется сглаживание. Следующий фильтр сглаживания



применяется три раза (сначала к полученным данным, а затем к результатам применения данного фильтра), где  – значения сглаживаемой функции,  – параметр сглаживания, . Сглаживание к ячейке с координатами *(i, j)* применяется только в том случае, если все четыре соседние ячейки (сверху, снизу, слева и справа) принадлежат к зеркалу озера.

Рисунки 8 и 9 демонстрируют применение сглаживания к результату восстановления дна озера Иссык-Куль, приведенному на рисунках 6 и 7.

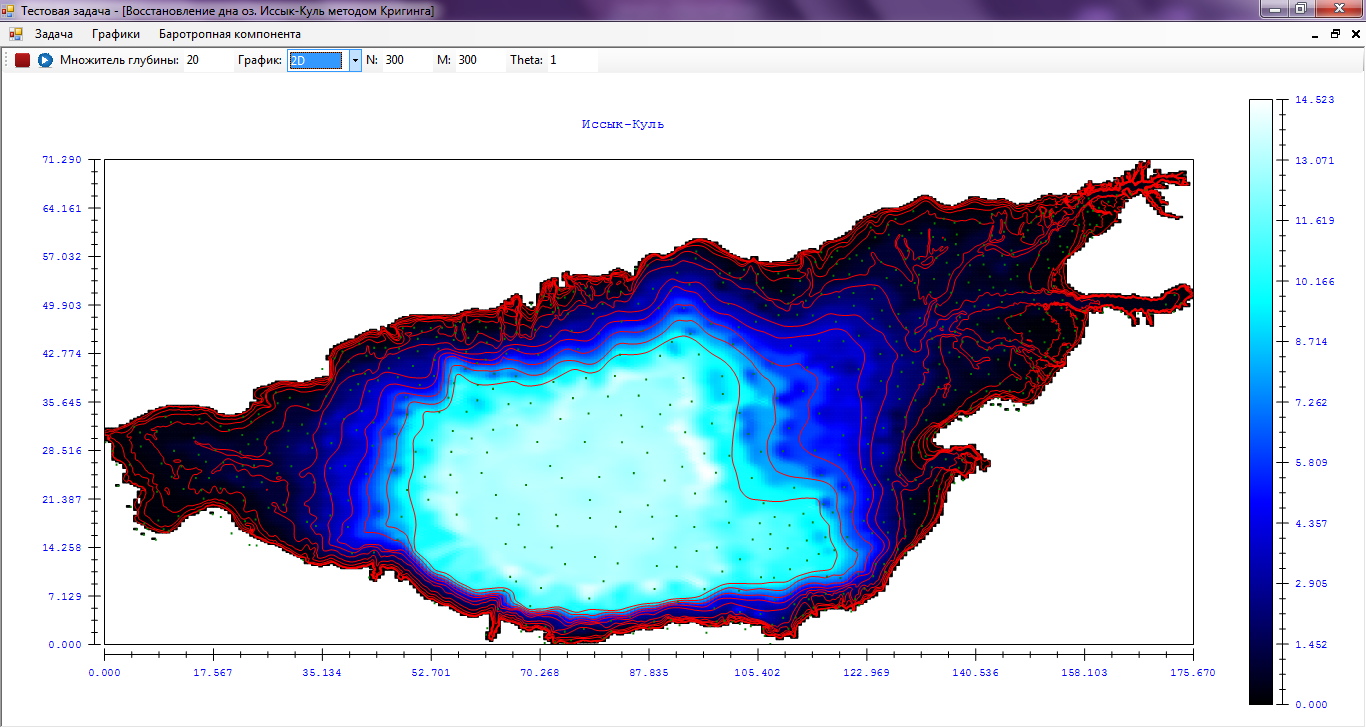


Рисунок 8. Сглаженный рельеф озера Иссык-Куль, 2D.

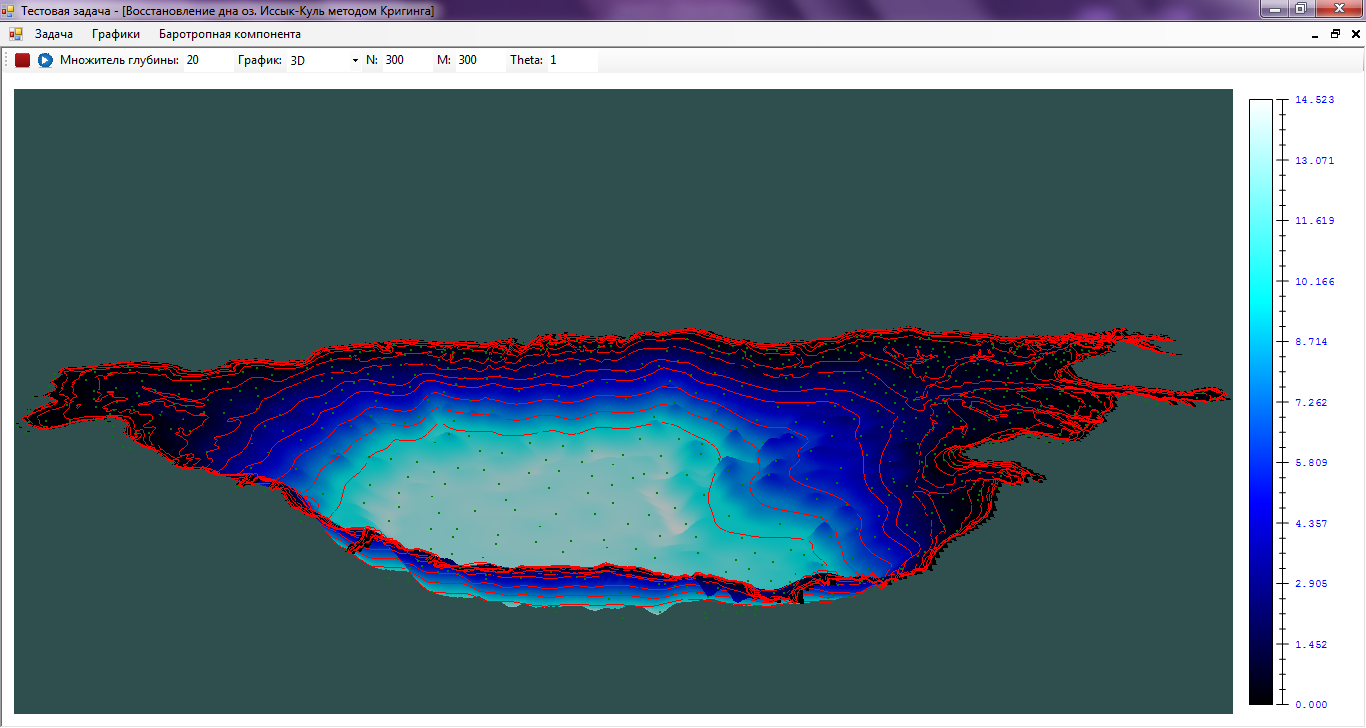


Рисунок 9. Сглаженный рельеф озера Иссык-Куль, 3D.

**Построение сеточной области озера Иссык-Куль**

На вход алгоритма, генерирующего сеточную область озера Иссык-Куль, подаются следующие данные: необходимое число сеточных узлов в горизонтальных направлениях (запад-восток (параметр *N*) и юг-север (параметр *M*)) и значения вертикальных уровней (в данный момент задается равномерная сетка по глубине с шагом *Dz*). Алгоритм построения сеточной области озера Иссык-Куль включает в себя следующие этапы:

* аппроксимация зеркала озера Иссык-Куль;
* восстановление дна озера Иссык-Куль с использованием метода Кригинга;
* сглаживание результата интерполяции дна озера Иссык-Куль;
* построение трехмерной сеточной области на основе сетки зеркала и интерполяции рельефа дна.

Первый, второй и третий этапы описаны в предыдущем параграфе. На третьем этапе происходит сглаживание результата интерполяции. При этом используется значение параметра .

На последнем этапе происходит построение трехмерной сеточной области. Для каждой ячейки зеркала водоема на втором этапе рассчитана глубина. По данной глубине вычисляется необходимое количество уровней с шагом *Dz* (так, что первый уровень представляет собой плоскость, перпендикулярную оси *Oz* при *z=0*, второй – при *z=Dz*, третий – при *z=2Dz*, и т.д.) для того, чтобы точка с рассматриваемой глубиной была заключена между двумя последними уровнями. Вычисляя количество уровней для каждой ячейки зеркала водоема, получаем трехмерную сеточную область Иссык-Куля.

По описанному выше алгоритму была разработана программа в среде разработки Visual Studio 2010 на языке программирования C#.

На рисунке 10 приведена сеточная область, построенная при . Глубина озера увеличена в 20 раз.

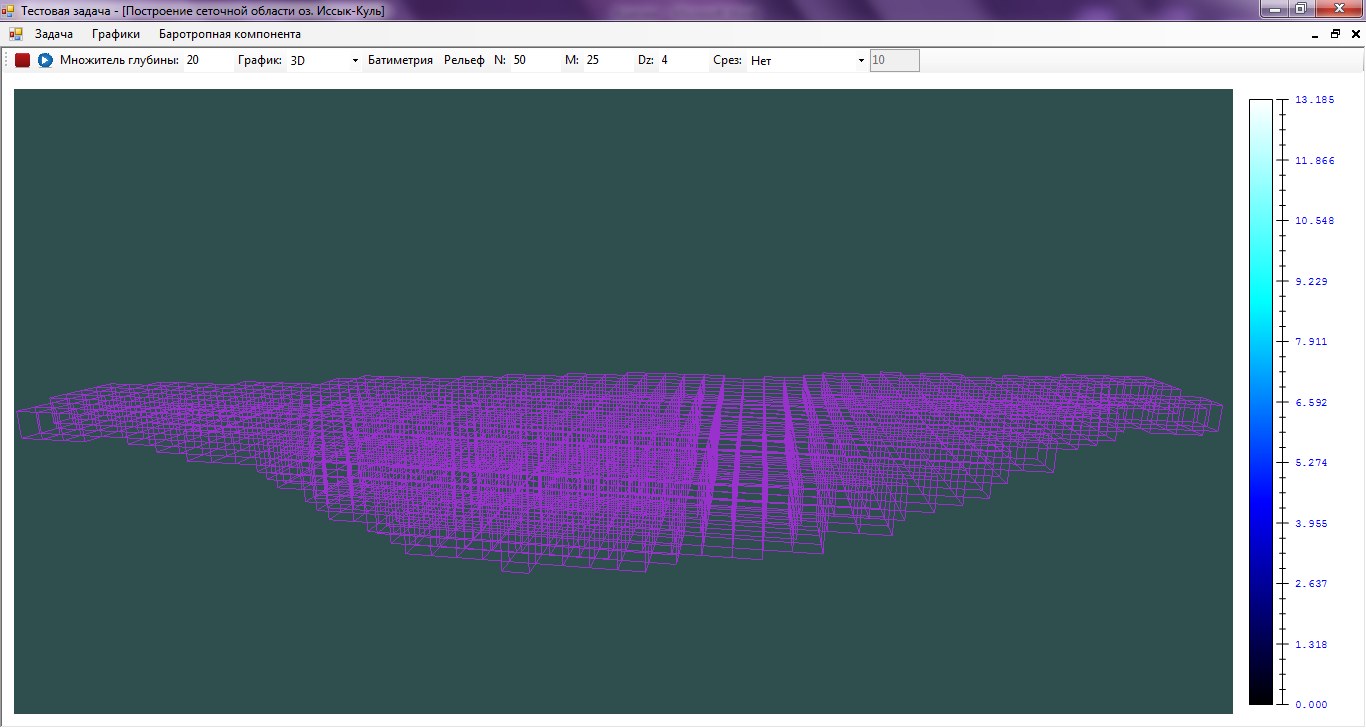


Рисунок 10. Сеточная область озера Иссык-Куль.

На рисунках 11 и 12 представлены вертикальные сечения сеточной области озера Иссык-Куль. Параметры сетки: .

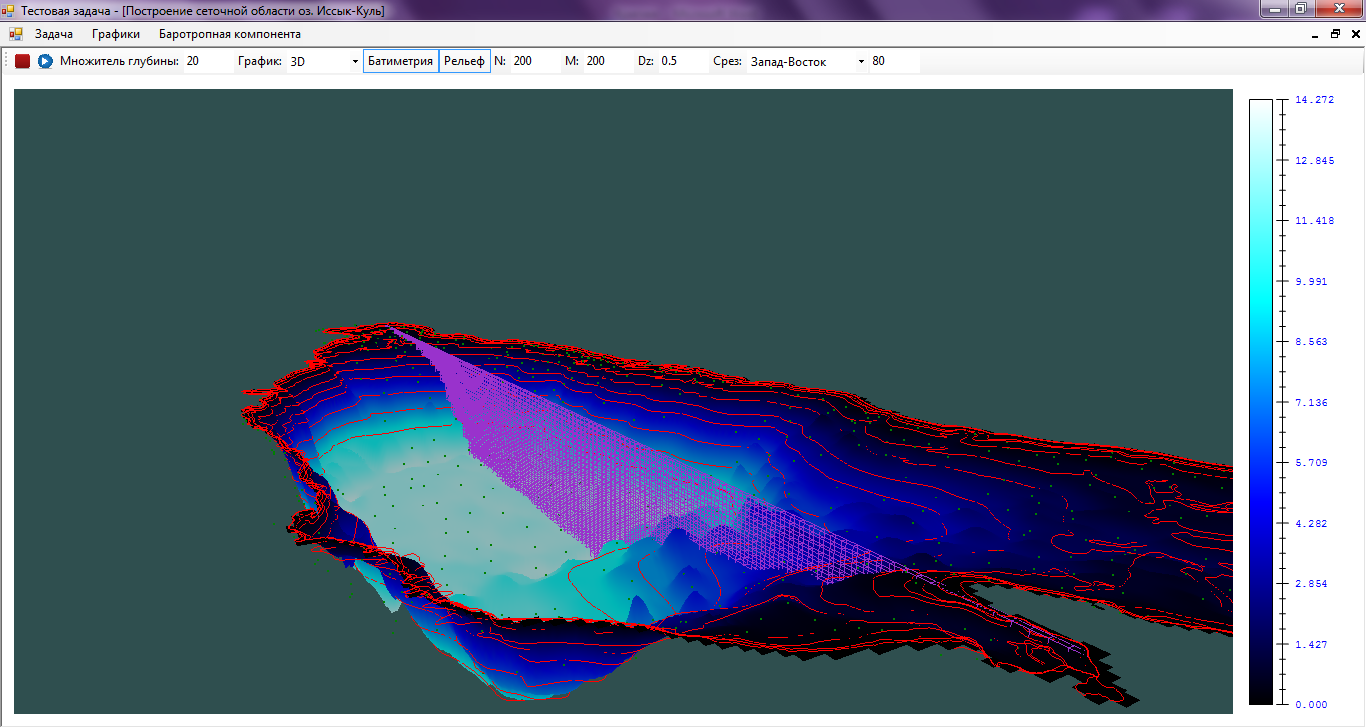


Рисунок 11. Вертикальное сечение в направлении запад-восток.

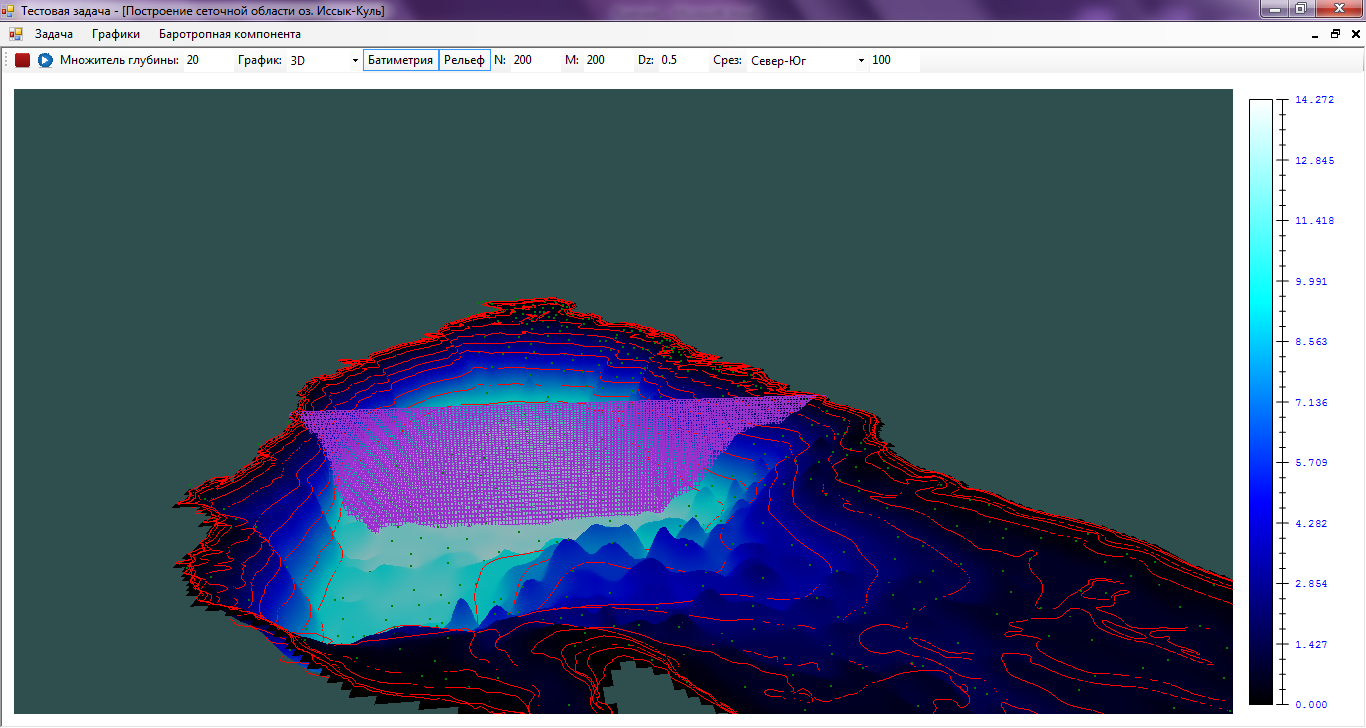


Рисунок 12. Вертикальное сечение в направлении север-юг.