

Isoliner - грид и изолинии

Содержание

1	Введение.....	1
1.1	Установка и расположение.....	2
1.2	Обновление.....	3
1.3	Вызов справки.....	3
2	Общая логика работы.....	3
3	Инструмент 2D Kriging (точки – растр).....	4
3.1	Автоматические значения.....	6
3.2	Обрезка по контуру скважин.....	6
3.3	Вариограмма и наггет.....	7
3.4	Отсев ураганных проб.....	11
3.5	Стандартная ошибка кригинга.....	13
4	Инструмент Изолинии из растра.....	14
4.1	Сглаживание изолиний.....	15
4.2	Контурные полигоны (пояса).....	15
4.3	Оформление слоёв.....	16
4.4	Типичные ситуации и решения.....	16
5	Инструмент Кросс-валидация вариограммы.....	17
6	Инструмент Создать пример скважин (демо).....	21
7	Лицензия и поддержка.....	22

1 Введение

Isoliner - провайдер инструментов Processing для интерполяции точечных данных и построения изолиний. Это порт инструмента Isoliner (ArcGIS) в QGIS. Ядро кригинга - алгоритм KB2D из GSLIB. Группа **Грид и изолинии** содержит четыре инструмента.

2D Kriging (точки – растр) - ординарный или простой кригинг по точечному слою.

Изолинии из растра - изолинии и контурные полигоны (пояса между изолиниями), границы которых совпадают с линиями.

Кросс-валидация вариограммы - скользящий контроль для проверки и подбора параметров кригинга по ошибке, а не визуально.

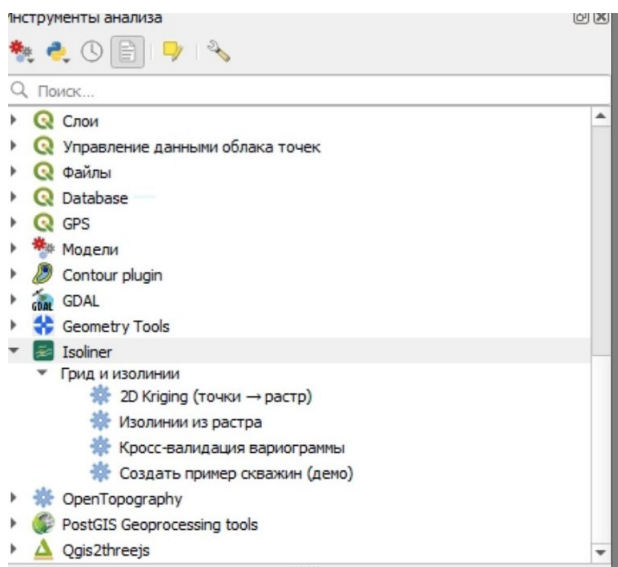
Создать пример скважин (демо) - генерация учебного точечного слоя с пространственной структурой (кровля, мощность, содержание компонента) для обучения и проверки без реальных данных.

Подходит для отметок пласта, мощностей, ФМС, химии и любых числовых атрибутов скважин.

Несколько терминов, которые встречаются далее. Вариограмма описывает, насколько сильнее различаются значения с ростом расстояния между точками. Силл - уровень, на который она выходит (близок к дисперсии данных). Наггет (от англ. nugget, «эффект самородка») - скачок вариограммы у нуля, то есть разброс на сколь угодно малых расстояниях, вызванный измерительным шумом и микроизменчивостью. Далее по тексту - просто «наггет».

1.1 Установка и расположение

Основной способ - из официального репозитория QGIS. Откройте Модули – Управление и установка модулей – вкладка **Все**, в поиске наберите **Isoliner**, выберите плагин и нажмите **Установить**. При установке из репозитория QGIS сам сообщает о выходе новых версий и обновляет плагин по кнопке.



*Инструменты модуля в панели **Обработка**: группа **Isoliner** – **Грид и изолинии** с четырьмя инструментами.*

Альтернативный способ - из ZIP-файла. Модули – Управление и установка модулей – Установить из ZIP. Это удобно для офлайн-установки и пререлизных сборок.

После установки инструменты появляются в панели **Обработка** (Processing), провайдер **Isoliner**, группа **Грид и изолинии**. Требования: QGIS 3.16+. Внешних зависимостей нет - используются NumPy, GDAL и штатные алгоритмы Processing из состава QGIS.

1.2 Обновление

При установке из репозитория QGIS показывает уведомление о доступной новой версии - значок в строке состояния и список во вкладке **Обновляемые** менеджера модулей. Обновление выполняется одной кнопкой. При установке из ZIP новая версия ставится тем же путём, поверх старой.

Плагин корректно перезагружается «на лету», перезапуск QGIS не требуется. Для быстрой перезагрузки кода при разработке удобен модуль Plugin Reloader (кнопка **Reload a plugin...**). Выберите **Isoliner** - провайдер и все инструменты перерегистрируются сразу.

1.3 Вызов справки

В диалоге каждого инструмента есть кнопка **Справка**, открывающая это руководство (PDF в комплекте плагина). Правая панель диалога дополнительно показывает краткую подсказку по инструменту.

2 Общая логика работы

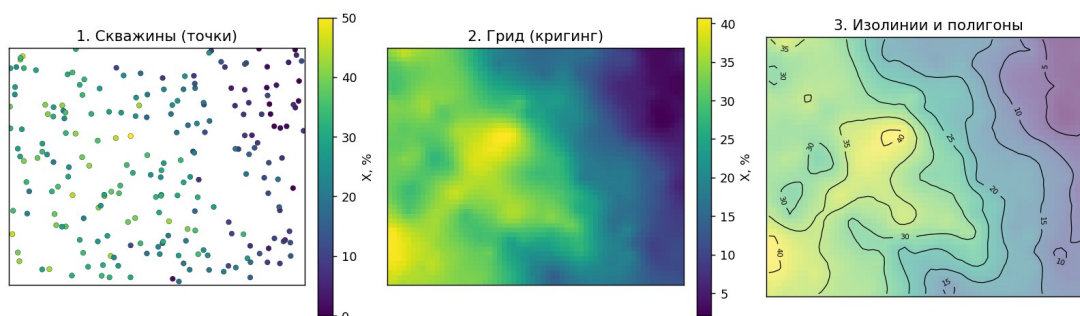
Типовой сценарий состоит из двух шагов:

2D Kriging: по точечному слою и числовому полю Z строится растр (регулярная сетка значений).

Изолинии из растра: по полученному растру строятся изолинии и, при необходимости, залитые контурные полигоны.

Шаги независимы: **Изолинии из растра** работают с любым растром, не только с результатом кригинга.

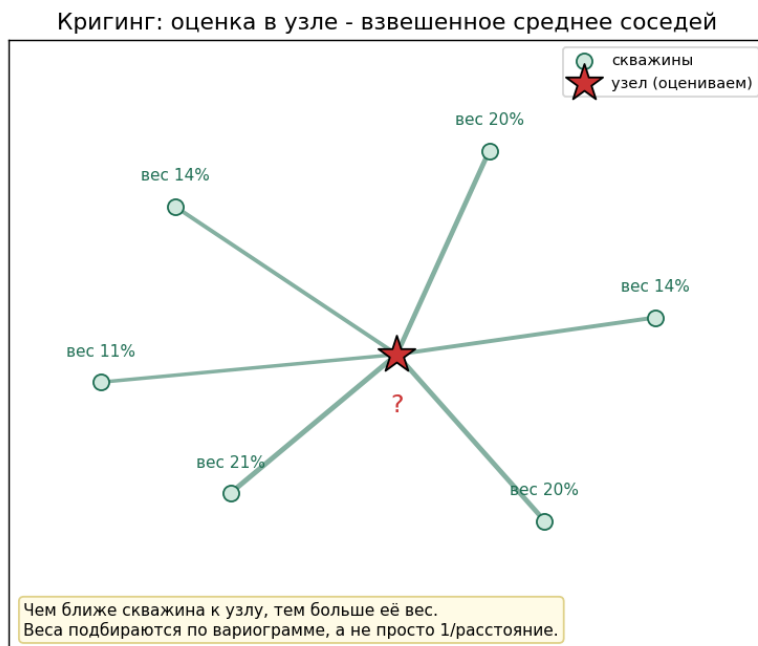
Как работает модуль: скважины → грид → изолинии



Весь процесс на сгенерированном примере: скважины с измерениями (слева) превращаются в сплошной грид кригингом (в центре), по которому строятся изолинии и контурные полигоны (справа).

3 Инструмент 2D Kriging (точки → растр)

Ординарный (OK) или простой (SK) кригинг по точечному слою. Совпадающие точки (один и тот же XY) усредняются по Z. В узлах сетки значения исходных точек воспроизводятся точно (при нулевом наггете).

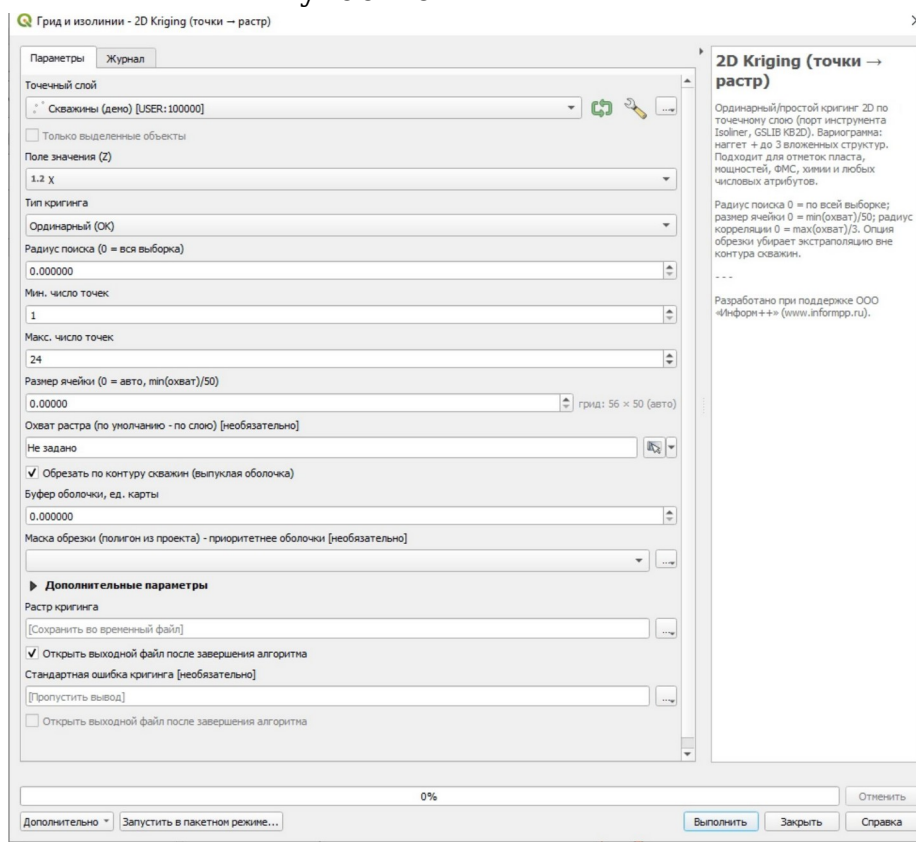


Кригинг оценивает значение в узле как взвешенное среднее ближайших скважин: чем ближе скважина, тем больше её вес. Веса подбираются по вариограмме.

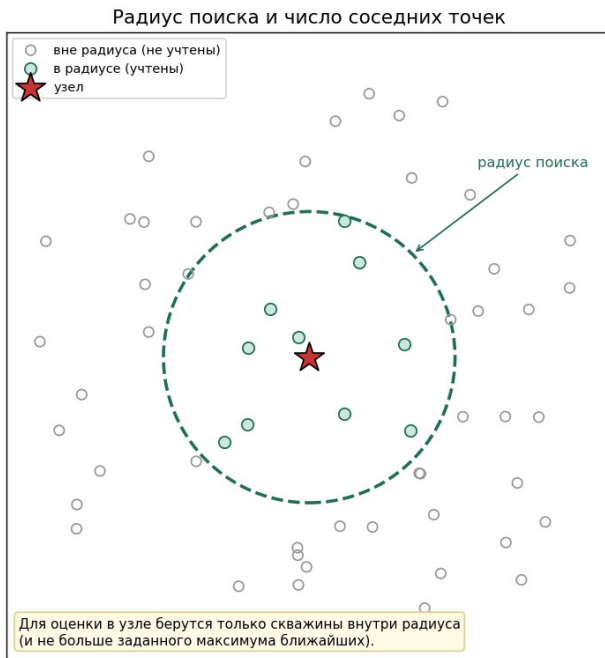
Основные параметры:

Параметр	Что задаёт	По умолчанию / совет
Точечный слой	Исходные точки (скважины) для интерполяции.	-
Только выделенные объекты	Считать лишь по выделенным точкам слоя.	выкл.
Поле значения (Z)	Числовой атрибут, который интерполируется: отметка пласта, мощность, ФМС, химия и т. п.	запоминается между запусками
Тип кригинга	Ординарный (OK) - локально оценивает среднее сам. Простой (SK) - использует заданное Среднее .	OK
Радиус поиска	Радиус окна поиска соседних точек вокруг узла. 0 = вся выборка.	0 (вся выборка)
Мин. число точек	Если в окне меньше точек - узел остаётся пустым (nodata).	1

Параметр	Что задаёт	По умолчанию / совет
Макс. число точек	Сколько ближайших точек включается в систему кригинга.	24
Размер ячейки	Шаг грида. 0 = авто = $\min(\text{охват})/50$.	мельче = плавнее, но дольше
Охват раstra	Прямоугольник расчёта. По умолчанию - по слою.	по слою
Обрезать по контуру скважин	Растр обрезается выпуклой оболочкой всех точек - убирает экстраполяцию в пустых углах.	рекомендуется вкл.
Буфер оболочки	Расширить оболочку на N единиц карты наружу.	0
Маска обрезки	Свой полигон вместо оболочки (приоритетнее) - удобно для вогнутых участков.	-



Окно инструмента **2D Kriging**: основные параметры. Справа - краткая встроенная справка.



Для каждого узла берутся только скважины внутри радиуса поиска и не более заданного числа ближайших. Точки за радиусом не участвуют.

3.1 Автоматические значения

Размер ячейки = $\min(\text{ширина, высота охвата}) / 50$.

Радиус корреляции вариограммы = $\max(\text{ширина, высота охвата}) / 3$.

Радиус поиска (при 0) = диагональ охвата, то есть берётся вся выборка.

3.2 Обрезка по контуру скважин

Кригинг считает весь прямоугольный охват, поэтому вне области данных значения являются экстраполяцией и дают артефакты (длинные «веерные» изолинии в пустых углах). Опция **Обрезать по контуру скважин** строит выпуклую оболочку всех точек (с необязательным буфером) и обрезает по ней растр. Экстраполяция исчезает. Если фактическая граница участка вогнутая, задайте свой полигон в **Маске обрезки** - он приоритетнее оболочки.

3.3 Вариограмма и наггет

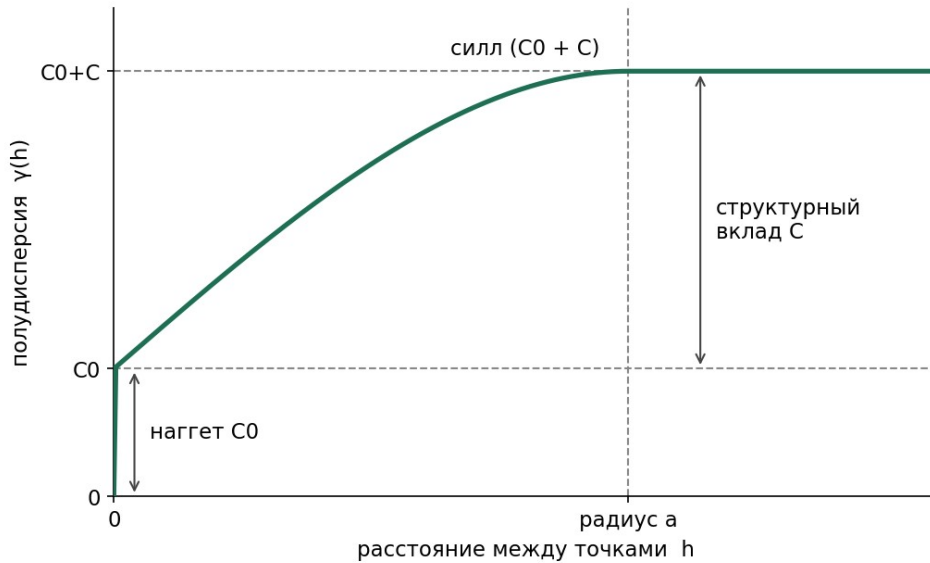
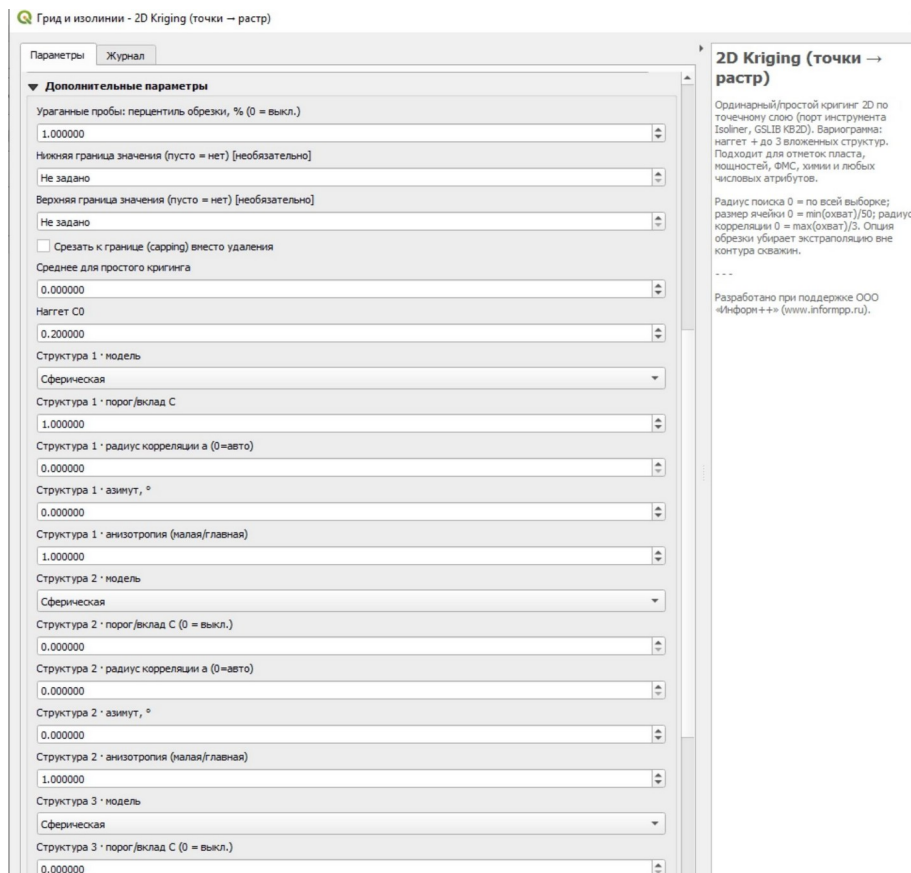


Схема вариограммы: наггет C_0 , структурный вклад C , силл $(C_0 + C)$ и радиус корреляции a .



Раздел **Дополнительные параметры** окна **2D Kriging**: отсев ураганных проб, наггет C_0 и до трёх структур вариограммы.

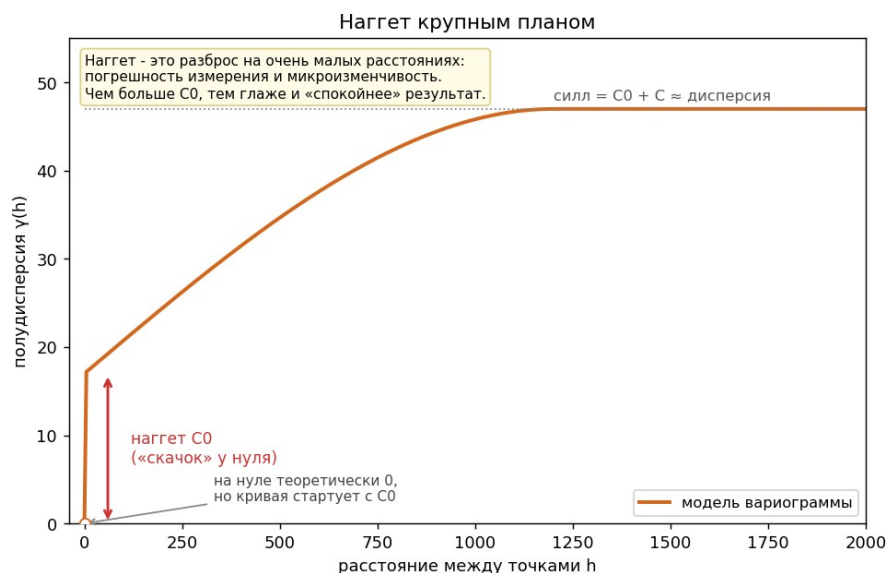
Кригинг опирается на модель вариограммы - она описывает, насколько сильно различаются значения Z в двух точках в зависимости от расстояния

между ними. По этой модели каждой соседней скважине назначается вес. Модель задаётся в разделе **Дополнительные параметры**.

Модель вариограммы: наггет C_0 , силл ($C_0 + C$) и радиус корреляции a .

Наггет C_0

Наггет - это значение, к которому стремится кривая вариограммы при расстоянии, стремящемся к нулю. Теоретически, на нулевом расстоянии расхождение должно быть нулевым (точка сравнивается сама с собой), но на практике остаётся ступенька. Она отражает то, что данные на очень малых расстояниях всё равно не совпадают: погрешность измерения и оцифровки, микроизменчивость на масштабе меньше расстояния между скважинами, расхождение дублей в одной точке.



Наггет крупным планом: модель стартует у нуля не с 0, а со «скачка» C_0 . Это разброс на сверхмалых расстояниях (погрешность измерения, микроизменчивость). Полка $C_0 + C \approx$ дисперсии данных.

Как наггет влияет на результат:

$C_0 = 0$ (по умолчанию) - кригинг является точным интерполятором: поверхность обязана пройти ровно через каждую скважину. Изолированная скважина с выбросом по Z превращается в конус («бычий глаз»).

$C_0 > 0$ - кригинг перестаёт точно восстанавливать значение в точке измерений и становится сглаживателем: вблизи скважины оценка подтягивается к локальному среднему. Чем больше доля наггета $C_0 / (C_0 + C)$, тем сильнее сглаживание.

$C_0 =$ весь силл (чистый наггет) - пространственная связь теряется, поверхность вырождается в простое среднее. Это перебор.

Важно - единицы. Наггет C_0 и силл задаются в **абсолютных единицах дисперсии данных** (квадрат единиц Z), а не в долях 0-1. Значение «1» у силла

по умолчанию - это условное значение, которое почти всегда нужно изменить: задайте суммарный силл ($C_0 + \text{вклады структур } C$) **близким к дисперсии данных**. Уровень сглаживания определяет не абсолютная величина наггета, а его **доля в силле** $C_0 / (C_0 + C)$. Практический порядок: возьмите суммарный силл \approx дисперсии, затем наггет = 0.2-0.4 от него (то есть $0.2-0.4 \times \text{дисперсии}$ - абсолютное число, а не 0.2-0.4 как таковое). Чем меньше наггет, тем больше деталей, но и больше локальных пиков. Чем больше, тем поверхность ровнее, но возможно сглаживание реальной структуры. Дисперсию данных инструмент печатает в Журнал при запуске - это и есть ваш ориентир для выбора силла.

Структуры, радиус и анизотропия

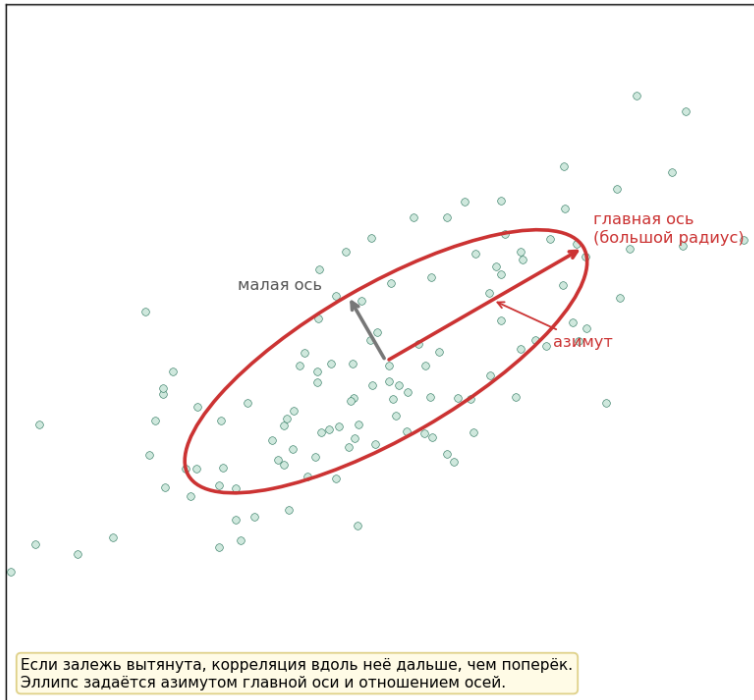
Силл (полка) это уровень, на который выходит вариограмма. Он складывается из наггета C_0 и вкладов структур C . Доступно до трёх вложенных структур (модели: сферическая, экспоненциальная, гауссова, степенная). Структуры 2 и 3 включаются ненулевым порогом C .

Силл: смысл и порядок величины. Силл - это верхний предел различий между точками: насколько в среднем отличаются далёкие друг от друга скважины. Он практически равен обычной дисперсии данных. Пример по KCl: среднее $\approx 25\%$, дисперсия $\approx 47.6 (\%^2)$, то есть $\sigma \approx 6.9\%$. Значит, суммарный силл задаём ≈ 47.6 . Если наггет $C_0 \approx 17$ (примерно 0.35 силла), то структурный вклад первой структуры $C \approx 47.6 - 17 \approx 30$. На сам грид абсолютный масштаб не влияет - для оценок важно только отношение $C_0 : C$. Но он нужен, чтобы карта стандартной ошибки и MSDR в кросс-валидации были в реальном масштабе (суммарный силл \approx дисперсии \rightarrow MSDR ≈ 1). Поэтому: не оставляйте силл = 1 по умолчанию, поднимите его до дисперсии данных.

Радиус корреляции a - расстояние, на котором вариограмма выходит на полку. Дальше этого расстояния точки практически не влияют друг на друга. При 0 берётся автоматическое значение $\max(\text{охват})/3$.

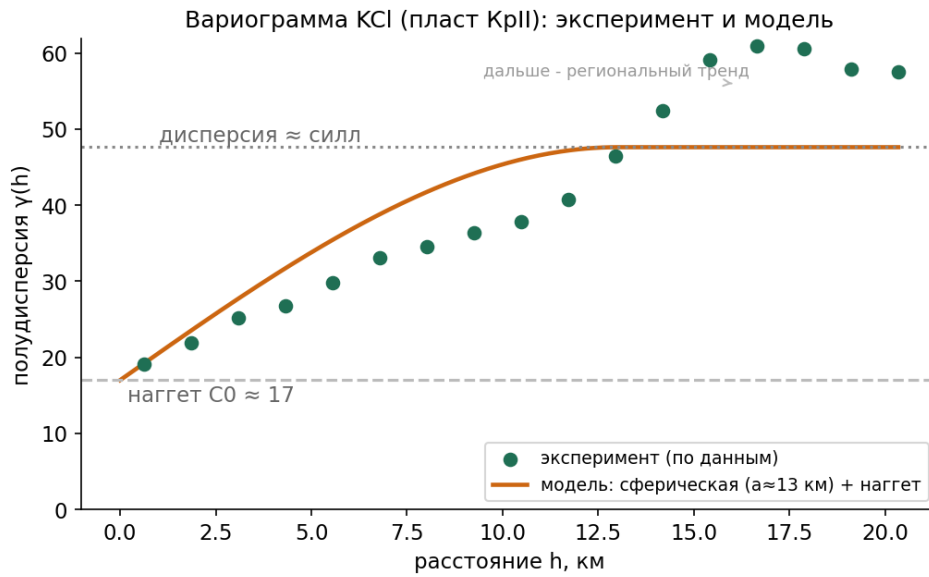
Анизотропия задаётся азимутом главной оси и отношением радиусов (малая/главная). Значение 1 - изотропно (влияние одинаково во всех направлениях). Значение меньше 1 укорачивает корреляцию поперёк главной оси - полезно для вытянутых геологических структур.

Анизотропия: вытянутая залежь и эллипс поиска



Анизотропия: если залежь вытянута, корреляция вдоль неё дальше, чем поперёк. Эллипс поиска задаётся азимутом главной оси и отношением осей (малая/главная).

Параметр	Что задаёт	По умолчанию / совет
Среднее для простого кригинга	Используется только при типе SK.	0
Наггет C0	«Шум»/скачок вариограммы у нуля. Подавляет локальные пики. В абсолютных единицах дисперсии.	0. Для сглаживания 0.2-0.4 от силла
Структура i · модель	Форма вариограммы: сферическая, экспоненциальная, гауссова, степенная.	сферическая
Структура i · порог/вклад C	Вклад структуры в силл (абс. единицы дисперсии). Сумма C0+C ≈ дисперсии данных. Для структур 2 и 3: 0 = выкл.	стр. 1 = 1 (замените на ≈ дисперсию)
Структура i · радиус корреляции a	Дистанция выхода на полку. 0 = авто = max(охват)/3.	0 (авто)
Структура i · азимут, °	Направление главной оси анизотропии.	0
Структура i · анизотропия (малая/главная)	Отношение радиусов поперёк/вдоль оси. 1 = изотропно.	1

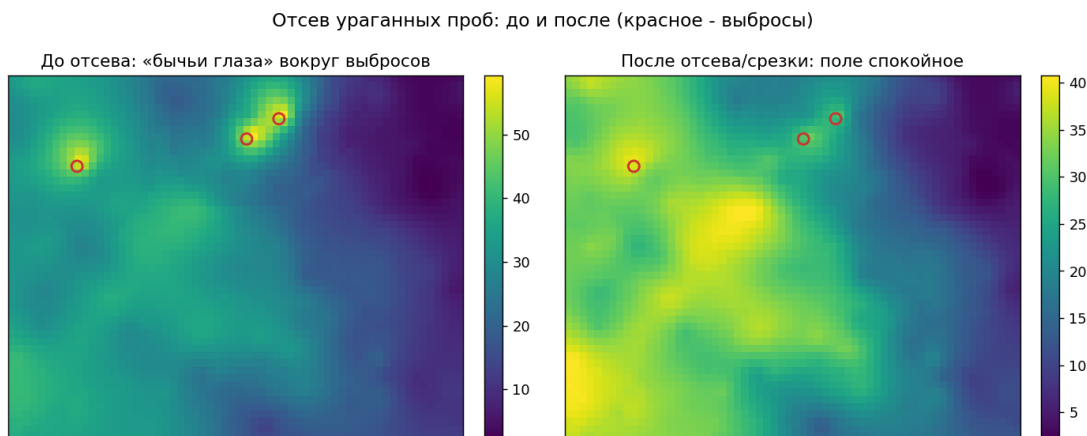


Экспериментальная вариограмма KCl по пласту KpII и подобранная модель: наггет $C_0 \approx 17$, силл совпал с дисперсией данных, радиус ≈ 13 км. Точки дальше силла - региональный тренд.

Так схема выглядит на реальных данных. Вариограмму строят по скважинам: для пар точек считают полудисперсию и усредняют по расстояниям - получается облако (зелёные точки), под которое подбирают модель (кривая). По нему и задают параметры кригинга: высота «скачка» у нуля - наггет C_0 , полка - силл (обычно близок к дисперсии данных), расстояние выхода на полку - радиус a . Если на больших расстояниях точки уходят выше силла, как здесь, - это региональный тренд (нестационарность). Его либо учитывают отдельно, либо ограничивают радиус поиска.

3.4 Отсев ураганных проб

Ураганные пробы - аномально высокие (или ошибочные) значения, которые искажают оценку: несколько «бонанц» по содержанию могут перетянуть на себя всю карту грейда, а явные ошибки (например, отрицательная мощность) портят поверхность. Инструмент **2D Kriging** позволяет ограничить такие пробы прямо при расчёте, без правки исходных данных. Параметры - в разделе **Дополнительно**.



Отсев ураганных проб на примере: слева три выброса дают «бычьи глаза» (горячие пятна), справа после срезки к верхней границе поле спокойное.

Два режима. Удалить - пробы вне допустимого диапазона выбрасываются (для явно битых записей). **Срезать (capping)** - значения вне диапазона прижимаются к границе, а сама точка остаётся в расчёте. Срезка - классический приём для ураганных проб по содержанию: положение точки не теряется, но её влияние ограничивается. Режим переключается флажком **Срезать к границе (capping) вместо удаления**.

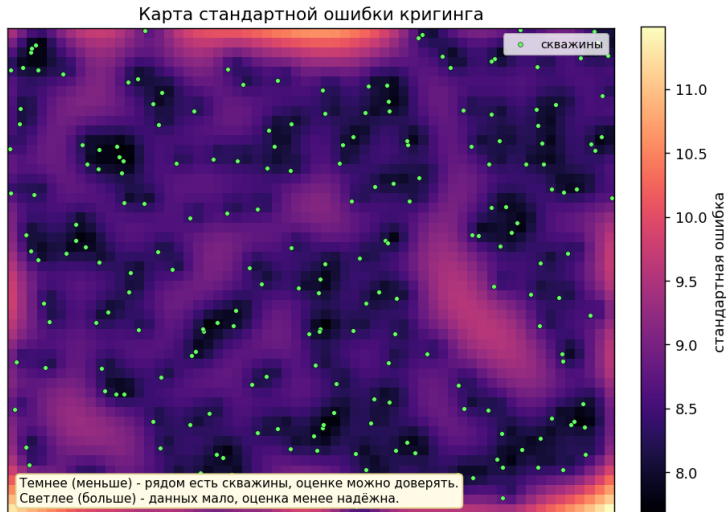
Границы по абсолюту. Нижняя граница и Верхняя граница задают пороги в единицах Z напрямую. Пустое поле - граница не задана. Они приоритетнее перцентилей. Пример: для мощности поставьте нижнюю границу 0 - уйдут отрицательные значения, верхнюю, скажем, 30 - уйдёт явный выброс в 122 м.

Границы по перцентилю. r -й перцентиль - это значение, ниже которого лежит $r\%$ всех проб. Например, 5-й перцентиль - порог, ниже которого только 5% самых малых значений. 95-й - порог, выше которого 5% самых больших. Параметр **перцентиль обрезки**, % задаёт число r , и границы берутся симметрично: от r -го до $(100-r)$ -го перцентилей. То есть $r = 2$ означает «считать ураганными 2% самых низких и 2% самых высоких проб»: всё ниже 2-го и выше 98-го перцентилей либо удаляется, либо срезается. Чем больше r , тем агрессивнее обрезка. $R = 0$ выключает перцентильный режим. Удобство в том, что абсолютные пороги знать не нужно - они вычисляются по самим данным и подходят к любому распределению и масштабу.

Двусторонность - важно для химии. Перцентильный режим режет оба хвоста - и верхний, и нижний. Для содержаний это опасно: $KCl = 0$ в зонах замещения - реальная геология, и обрезка нижнего хвоста ошибочно поднимет «пустые» участки. Поэтому для грейда отсекайте только сверху: оставьте **Нижнюю границу** пустой и задайте **Верхнюю** абсолютном (или применяйте перцентиль, понимая, что низ тоже будет затронут). Для отметок и мощностей двусторонняя обрезка обычно уместна.

Порядок и журнал. Фильтр применяется до усреднения совпадающих точек. В Журнал инструмента выводится, сколько проб удалено или срезано и в каких границах - это удобно для контроля.

3.5 Стандартная ошибка кригинга



Карта стандартной ошибки: тёмное у скважин (зелёные точки) - оценке можно доверять, светлое в пустых углах - данных мало.

Кроме самой оценки, кригинг даёт в каждом узле дисперсию ошибки - меру неопределённости. Её корень, стандартная ошибка, выводится необязательным вторым растром (параметр **Стандартная ошибка кригинга** инструмента **2D Kriging**). Единицы - те же, что у интерполируемой величины Z .

Ключевое свойство: стандартная ошибка зависит от геометрии расположения скважин и модели вариограммы, но не от самих значений Z . Поэтому это карта надёжности сети наблюдений, а не разброса данных. В точке скважины (при нулевом наггете) ошибка равна нулю - там значение известно точно. По мере удаления от скважин она растёт, а в областях без данных достигает максимума (примерно корень из силла).

Как читать. Тёмные (малые) значения - оценке можно доверять: рядом достаточно скважин. Светлые (большие) - оценка держится на далёких точках, фактически экстраполяция. Это первые кандидаты на доразведку. Сравнить удобнее относительно (где больше, где меньше), потому что абсолютная величина зависит от масштаба вариограммы (силла $S1_SILL$).

Важно. Это модельная оценка: она верна настолько, насколько верна заданная вариограмма (наггет, радиус, анизотропия). При наггете больше нуля ошибка у скважин не нулевая - наггет задаёт нижний «пол» неопределённости. Строгим доверительным интервалом стандартная ошибка не является, но как относительная карта неопределённости очень полезна.

Оформление. Задайте слою градуированную символику по значению (например, от тёмного к красному) - и сразу видно, где карта надёжна, а где нет.

4 Инструмент Изолинии из растра

Строит изолинии (линии) и, по умолчанию, контурные полигоны. Уровни задаются равномерным шагом или явным списком. Параметры:

Грид и изолинии - Изолинии из растра

Параметры Журнал

Растр

Шаг изолиний (0 = задать уровни ниже)

1.000000

Начальный уровень (offset)

0.000000

Явные уровни (через пробел) - приоритетнее шага (необязательно)

Главная изолиния каждая N-я (0 = выкл.)

5

Мин. длина линии, ед. карты (0 = без фильтра)

0.000000

☒ Сглаживать

Радиус сглаживания поля, ячеек (0 = без него)

2.000000

Скругление линий, итераций (0 = выкл.)

1

Имя поля значения

ELEV

Дополнительные параметры

Канал

1

Изолинии (линии)

[Сохранить во временный файл]

☒ Открыть выходной файл после завершения алгоритма

Контурные полигоны [необязательно]

[Сохранить во временный файл]

☒ Открыть выходной файл после завершения алгоритма

0%

Дополнительно * Запустить в пакетном режиме...

Выполнить Закрыть Справка

Изолинии из растра

Строит изолинии из растра: равномерный шаг или явные уровни (через пробел), главные (утолщенные) изолинии флагом is_index, фильтр коротких линий.

Сглаживание: поле (растр) сглаживается гауссовым фильтром (радиус в ячейках), затем линии слегка округляются (Smooth). Изолинии плавные и не пересекаются даже в густых местах; округление убирает «октагоны» от грубого грида.

По умолчанию строит и контурные полигоны (полоса между изолиниями) во временный слой - их границы совпадают с изолиниями, покрытие сплошное. Чтобы их не строить - очистите поле «Контурные полигоны».

Поля: линии - значение уровня (по умолчанию ELEV) и is_index (1 у главных); полигоны - ELEV_MIN/ELEV_MAX (диапазон пояса).

Разработано при поддержке ООО «Информ-++» (www.informpp.ru).

Окно инструмента **Изолинии из растра**.

Параметр	Что задаёт	По умолчанию / совет
Растр	Входной растр (например, результат кригинга).	-
Шаг изолиний	Равномерный шаг по Z. 0 = задать Явные уровни .	-
Начальный уровень (offset)	Привязка сетки уровней (уровни кратны шагу от offset).	0
Явные уровни	Список уровней через пробел. Приоритетнее шага. Десятичный разделитель - запятая или точка.	-
Главная изолиния каждая N-я	Каждая N-я линия помечается is_index = 1 (для утолщения). 0 = выкл.	5
Мин. длина линии	Отбрасывать линии короче порога (ед. карты). 0 = без фильтра.	-
Сглаживать	Главный выключатель сглаживания	вкл.

Параметр	Что задаёт	По умолчанию / совет
Радиус сглаживания поля, ячеек	(поле + скругление линий). Гауссово сглаживание раstra ДО контуринга. 0 = без него.	1-2.5
Скругление линий, итераций	Лёгкое скругление линий (Chaikin) поверх гладкого поля. Убирает «октагоны». 0 = выкл.	2 (3 на грубом гриде)
Имя поля значения	Имя атрибута уровня в выходных линиях.	ELEV
Канал (доп.)	Номер канала входного раstra.	1
Изолинии / Контурные полигоны	Выходные слои. Полигоны строятся по умолчанию во временный слой.	-

Выходные поля: у линий - значение уровня (по умолчанию ELEV) и is_index (1 у главных изолиний). У полигонов - ELEV_MIN и ELEV_MAX (диапазон пояса).

4.1 Сглаживание изолиний

Сглаживание выполняется в два приёма, и оба включаются галочкой **Сглаживать**:

Сглаживание поля. Перед построением изолиний растр сглаживается гауссовым фильтром (радиус в ячейках). Изолинии строятся уже по сглаженному полю, поэтому они плавные и принципиально не пересекаются даже в густых местах. Сам растр кригинга при этом не меняется - сглаживается лишь временная копия.

Скругление линий. Поверх этого линии слегка скругляются (алгоритм Chaikin, число итераций). На грубом гриде изолинии иначе выглядят «октагонами», так как вершины ставятся по краям ячеек. Поскольку поле уже гладкое, скругление не создаёт пересечений.

Если углы всё ещё резковаты - увеличьте число итераций скругления (до 3) или уменьшите размер ячейки в кригинге. Если линии заметно «уезжают» от исходной поверхности - уменьшите радиус сглаживания поля или снимите галочку.

4.2 Контурные полигоны (пояса)

Контурные полигоны - это залитые пояса между соседними изолиниями. Они строятся не классификацией «ступенек» раstra, а полигонизацией самих сглаженных изолиний вместе с контуром валидной области раstra: концы линий притягиваются к контуру, сеть нодируется и полигонизуется. Диапазон уровней каждого пояса определяется выборкой раstra в репрезентативной точке полигона.

Благодаря этому границы полигонов совпадают с изолиниями, а покрытие сплошное (без дыр). Полигоны несут поля ELEV_MIN и ELEV_MAX. По умолчанию они строятся во временный слой. Чтобы их не строить, очистите поле **Контурные полигоны**.

4.3 Оформление слоёв

Линии: задайте символику по правилу на основе is_index - главным изолиниям (is_index = 1) дайте бóльшую толщину. Подпись - по полю уровня (ELEV).

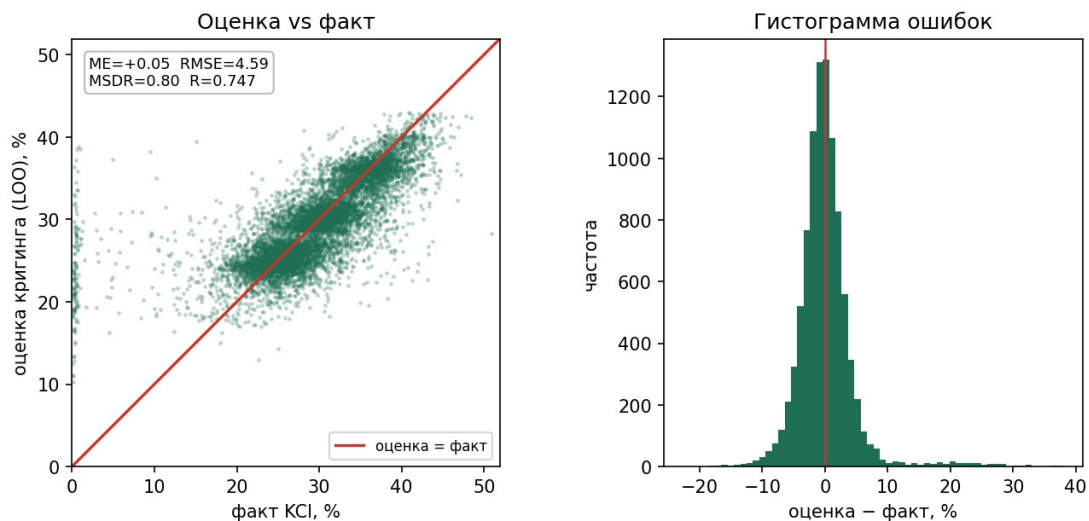
Полигоны создаются с одним символом. Для заливки по диапазонам задайте градуированную символику по ELEV_MIN (или ELEV_MAX).

Слой изолиний автоматически помещается над слоем полигонов, чтобы линии были видны поверх заливки.

4.4 Типичные ситуации и решения

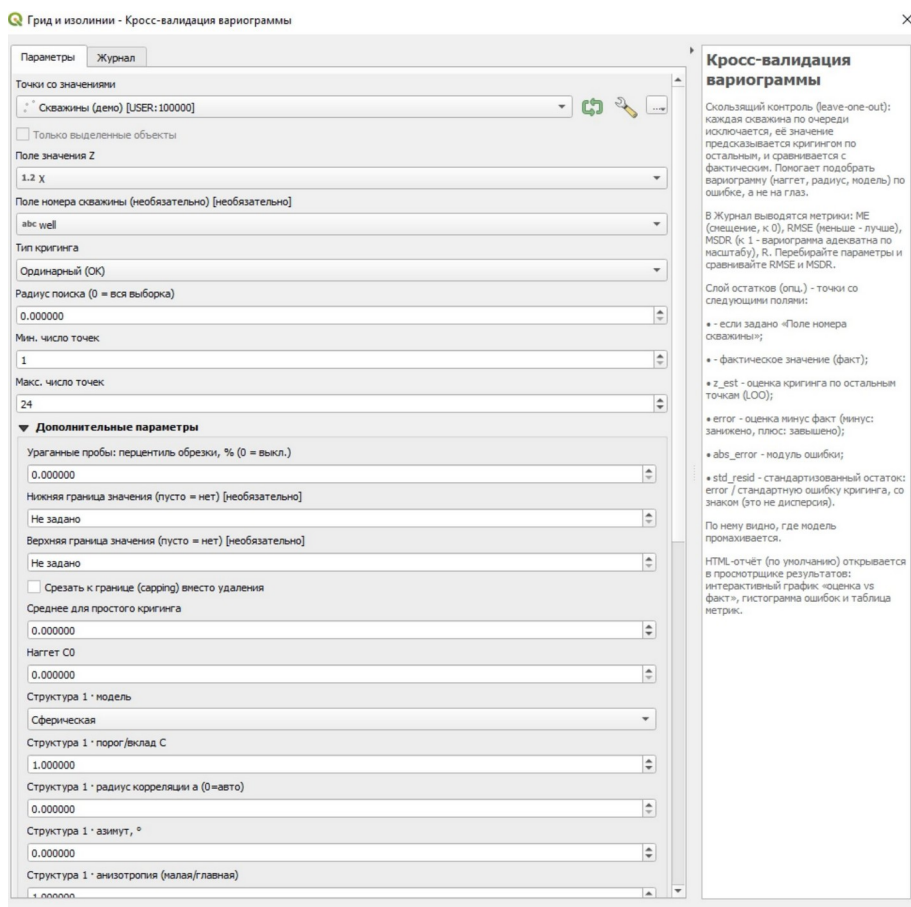
Что видно	Причина	Решение
Концентрические «бычьи глаза», конусы	Кригинг точно протягивает значение через скважины-выбросы (наггет 0).	Задать наггет C0 (0.2-0.4 от силла, в абсолютных единицах дисперсии). И/или увеличить радиус сглаживания поля.
Угловатые изолинии («октагоны»)	Грубый грид: вершины ставятся по краям ячеек.	Увеличить Скругление линий до 3 или уменьшить размер ячейки в кригинге.
Радиальные/веерные линии в пустых углах	Экстраполяция за пределами данных.	Включить Обрезать по контуру скважин или задать маску обрезки.
Изолинии пересекаются в густых местах	Раньше - следствие сглаживания каждой линии.	Сглаживание выполняется по полю. Увеличить радиус сглаживания поля.
Полигоны одного цвета	По умолчанию слой создаётся с одним символом.	Задать градуированную символику по ELEV_MIN.

5 Инструмент Кросс-валидация вариограммы



Так выглядит HTML-отчёт кросс-валидации: слева график «оценка vs факт» с диагональю и метриками (пример - KCl по пласту KpII), справа гистограмма ошибок. Плотное облако вдоль диагонали - модель работает. Полоса при факте около 0 - зоны замещения. Гистограмма симметрична относительно 0 - смещения нет.

Инструмент проверяет, насколько удачно подобрана вариограмма, методом скользящего контроля (leave-one-out): каждая скважина по очереди исключается, её значение предсказывается кригингом по всем остальным, и сравнивается с фактическим. Так параметры (наггет, радиус, модель) подбираются по ошибке, а не субъективно.



Окно инструмента **Кросс-валидация вариограммы**. Поле **Поле номера скважины** включает подписи скважин в отчёте.

В Журнал выводятся метрики:

ME (среднее смещение) - систематическая ошибка. Должна быть близка к 0 (несмещённость).

MAE и RMSE - средняя и среднеквадратичная ошибка предсказания. Чем меньше, тем точнее. Но одной RMSE недостаточно: она минимальна при нулевом наггете (переобучение), хотя неопределённость при этом оценена неверно.

MSDR (стандартизованная ошибка) - средний квадрат ошибки, делённой на стандартную ошибку кригинга. Должен быть близок к 1. Если MSDR заметно больше 1 - дисперсия недооценена (наггет или силл малы). Если меньше 1 - переоценена.

R - коэффициент корреляции «оценка - факт».

На практике переберите несколько вариантов вариограммы и сравните. Хорошая модель даёт ME около 0, малую RMSE и MSDR около 1. Если RMSE тянет к нулевому наггету, а MSDR при этом огромный - это признак переобучения. Небольшой наггет калибрует неопределённость.

Опциональный слой остатков (точки с полями: факт - под именем проверяемого поля, z_est, error, abs_error и std_resid. Плюс номер скважины, если задано поле ID) показывает, где модель промахивается: крупные по модулю остатки - проблемные участки, систематические знаки остатков - локальный тренд. Слой автоматически называется по проверяемому полю и источнику, а у полей заданы псевдонимы - понятные названия (видны в таблице атрибутов и свойствах поля). std_resid - это стандартизованный остаток (оценка - факт) / стандартную ошибку кригинга, со знаком: минус - кригинг занижил, плюс - завысил (это не дисперсия, дисперсия всегда ≥ 0).

Поля слоя остатков:

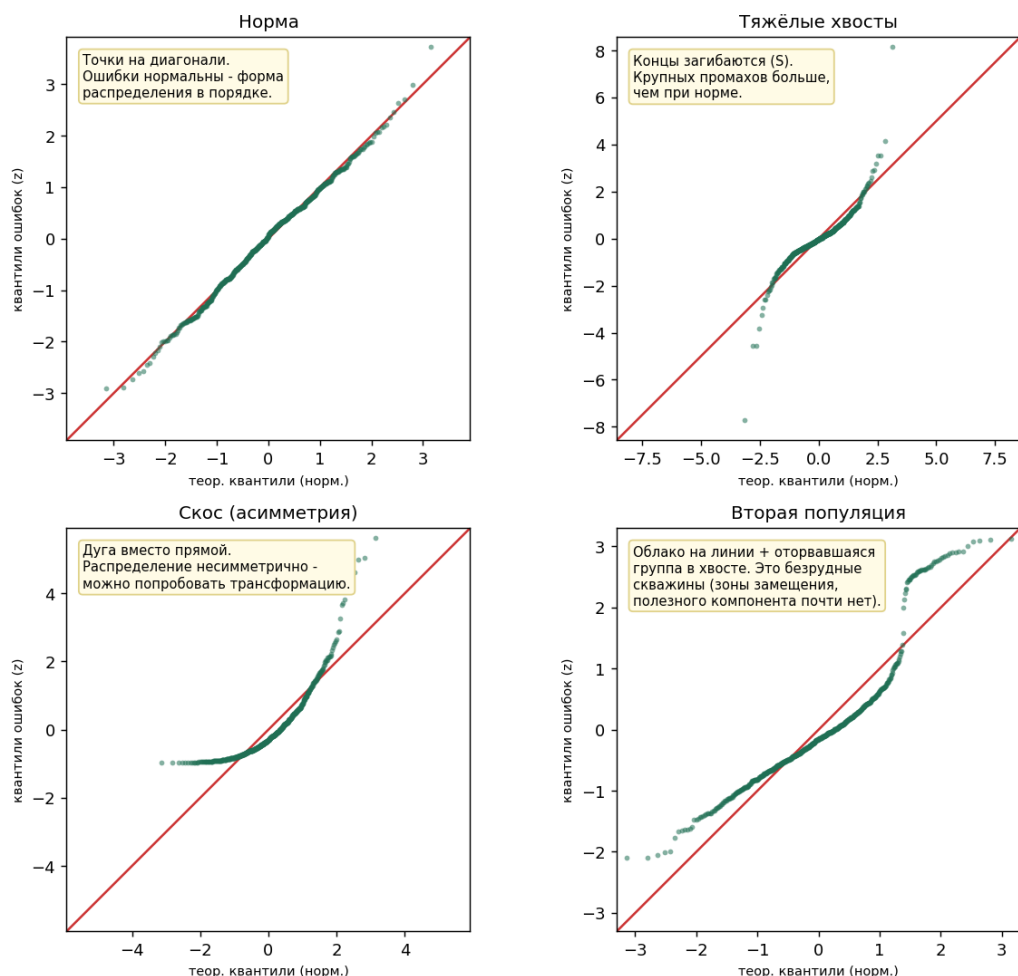
Поле	Псевдоним	Описание
<номер скважины>	Номер скважины	Значение выбранного поля ID (если задано Поле номера скважины).
<имя поля>	Факт (имя поля)	Фактическое значение проверяемого поля.
z_est	Оценка кригинга (LOO)	Оценка по остальным точкам (leave-one-out).
error	Ошибка (оценка - факт)	Оценка минус факт. Минус - занижено, плюс - завышено.
abs_error	Модуль ошибки	Абсолютная величина ошибки, error .
std_resid	Станд. остаток	(оценка - факт) / стандартную ошибку кригинга, со знаком. Не дисперсия (она ≥ 0).

Кроме слоя остатков инструмент по умолчанию формирует HTML-отчёт (на plotly): интерактивный график «оценка vs факт» с диагональю, гистограмма ошибок, QQ-график остатков и таблица метрик с блоком рекомендаций. В таблицу добавлена дисперсия данных - ориентир для суммарного силла C0+C. Рядом с таблицей метрик показывается блок **Параметры кригинга**: перечислены только настройки, отличающиеся от стандартных (наггет, силл, радиус, отсеив и так далее), чтобы было видно, какими параметрами получены эти метрики. На графике «оценка vs факт» при наведении на точку видны номер скважины и значения, а восемь скважин с наибольшими по модулю остатками подписаны прямо на графике - их удобно проверить в первую очередь. Отчёт открывается в просмотрщике результатов QGIS (или в браузере). Если plotly в сборке QGIS недоступен, отчёт всё равно создаётся - с таблицей метрик, но без графиков.

QQ-график остатков. Показывает форму распределения ошибок. Ошибки нормируются на их собственную дисперсию (z-оценка) и сравниваются с нормальным распределением, поэтому график читается по форме при любой калибровке. За масштаб неопределённости отвечает отдельно MSDR в таблице метрик. По горизонтали - квантили нормального распределения, по вертикали - нормированная ошибка. Если ошибки нормальны, точки ложатся на красную диагональ. Отклонения читаются сразу. Загнутые концы (S-образно) - тяжёлые хвосты, то есть крупных промахов больше, чем

при норме. Общий изгиб дугой - асимметрия, стоит подумать о трансформации значений. Отдельная группа, оторвавшаяся от линии, - чужая совокупность в данных, например безрудные пробы из зон замещения (где полезного компонента практически нет). Нормальность важна потому, что на ней основаны MSDR и карта стандартной ошибки.

QQ-график ошибок: как читать форму



Четыре типичных вида QQ-графика остатков: норма (точки на диагонали), тяжёлые хвосты, асимметрия и вторая популяция - оторвавшаяся группа в хвосте.

Главное - что делать с результатами. Смысл инструмента в том, чтобы перед построением грида утвердить или поправить весь набор параметров, который вы затем зададите в **2D Kriging**. Это и вариограмма (наггет, силл, радиус, модель, анизотропия), и настройки самого кригинга (радиус поиска, минимум/максимум точек, тип - ординарный или простой): кросс-валидация считает кригинг ровно с теми же настройками, поэтому удачный набор переносится в инструмент **2D Kriging** без изменений. Порядок решений:

- ME около 0, MSDR около 1, RMSE и R вас устраивают - набор можно утверждать: переносите эти же параметры (вариограмму и настройки поиска) в **2D Kriging** и стройте поверхность.
- MSDR заметно больше 1 - кригинг слишком «уверен в себе», карта стандартной ошибки будет занижена: увеличьте наггет C0 или силл и проверьте снова.
- MSDR меньше 1 - неопределённость завышена: уменьшите наггет или силл.
- ME заметно отличается от 0 - систематический сдвиг: проверьте данные и тип кригинга (для простого кригинга - заданное среднее).
- Большая RMSE и низкий R - модель плохо предсказывает: попробуйте другой радиус, модель или анизотропию (азимут и отношение осей). Если ничего не помогает - это предел данных: короткомасштабная изменчивость, которую сеть не ловит (например, зоны замещения по руде - на графике выше это вертикальная полоса при факте около 0).

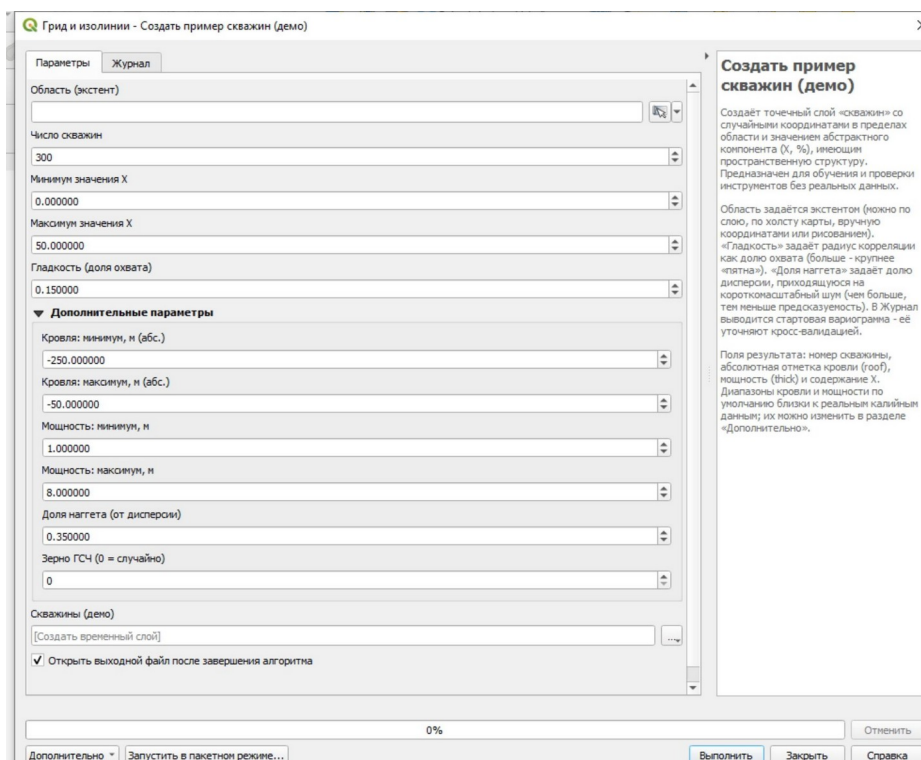
Слой остатков подсказывает точно: где остатки крупные - там стоит сгустить сеть (добавить скважины) или проверить пробы. Где остатки систематически одного знака по площади - там локальный тренд, который кригинг не учёл.

Итог: этот инструмент - последний шаг перед финальным кригингом. Сначала вы калибруете вариограмму здесь по ошибке, затем те же параметры ставите в **2D Kriging** - и поверхность вместе с картой стандартной ошибки получаются обоснованными, а не подобранными субъективно.

Замечание о скорости: контроль решает кригинг столько раз, сколько точек, поэтому на больших наборах (десятки тысяч скважин) выполняется заметно дольше. При необходимости уменьшите выборку.

6 Инструмент Создать пример скважин (демо)

Инструмент **Создать пример скважин (демо)** формирует точечный слой со случайными координатами и тремя структурированными полями: абсолютная отметка кровли пласта (roof), мощность (thick) и содержание абстрактного компонента X (%). Диапазоны кровли и мощности заданы по образцу промышленного пласта (KpII). Инструмент предназначен для обучения и проверки кригинга, изолиний и кросс-валидации без реальных данных.



Окно инструмента *Создать пример скважин (демо)*.

Параметры: область (экстент - можно задать по слою, по холсту карты, координатами или рисованием на карте). Число скважин. Минимум и максимум значения X. Диапазоны кровли и мощности (по умолчанию - как у КрII). Гладкость (доля охвата - задаёт радиус корреляции: больше значение - крупнее области однородности). Доля наггета (доля дисперсии, приходящаяся на короткомасштабный шум. Чем больше, тем ниже предсказуемость). В разделе **Дополнительно** - зерно генератора случайных чисел для воспроизводимости.

При запуске в Журнал выводится стартовая вариограмма (суммарный силл \approx дисперсии данных, наггет, радиус). Сгенерированные данные имеют восстановимую вариограмму, поэтому на них удобно освоить весь цикл: построить грид в **2D Kriging**, затем изолинии, и проверить параметры кросс-валидацией.

7 Лицензия и поддержка

Плагин распространяется под лицензией GNU GPL v2 или новее (GPL-2.0-or-later) - той же, что и сам QGIS. Полный текст в файле LICENSE в комплекте. © ООО «Информ++», www.informpp.ru.