

РОССИЙСКАЯ АКАДЕМИЯ НАУК
ИНСТИТУТ
ПРОБЛЕМ БЕЗОПАСНОГО
РАЗВИТИЯ АТОМНОЙ ЭНЕРГЕТИКИ

RUSSIAN ACADEMY OF
SCIENCES
NUCLEAR
SAFETY INSTITUTE

Препринт № NSI-24-94

Preprint NSI-24-94

Р.В. Арутюнян, Л.А. Большов, В.В. Демьянов,
Е.В. Каневская,
М.Ф. Каневский, В.П. Киселев, И.И. Линге,
В.М. Шершаков

Методика Geo-EAS анализа
пространственно распределенных данных.
Пример исследования: Чернобыльские
выпадения

Москва
1994

Moscow
1994

Р.В. Арутюнян, Л.А. Большов, В.В. Демьянов, Е.В. Каневская, М.Ф. Каневский, В.П. Киселев, И.И. Линге, В.М. Шершаков Методика Geo-EAS анализа пространственно распределенных данных. Пример исследования: Чернобыльские выпадения. Препринт № NSI-24-94. Москва: Институт проблем безопасного развития атомной энергетики РАН, 1994. 55 с.

Аннотация

В работе представлена методика анализа пространственно распределенных данных. В работу включено подробное описание известного пакета прикладных программ Geo-EAS, созданного при участии Агентства по Охране Окружающей Среды США (EPA — Environment Protection Agency). Примеры работы с пакетом демонстрируются с использованием данных по Чернобыльским выпадениям.

©ИБРАЭ РАН, 1994

R. Arutyunyan, L. Bolshov, V. Demyanov, E. Kanevskaya, M. Kanevsky, V. Kiselyov, I. Linge, V. Shershakov. GEO-EAS SPATIAL DATA ANALYSIS METHODOLOGY. SAMPLE RESEARCH: CHERNOBYL FALLOUT. (in Russian). Preprint NSI-24-94. Moscow: Nuclear Safety Institute, August 1994. 55 p.

Abstract

There are different approaches in spatial data analysis. One of them — based on Geostatistical Environment Assessment Software (Geo-EAS) — is fully described in the work. The methodology was applied to Chernobyl fallout. The results of structural analysis and spatial interpolation clearly show abilities and restrictions of the software that was used.

©Nuclear Safety Institute, 1994

Методика Geo-EAS анализа пространственно распределенных данных. Пример исследования: Чернобыльские выпадения

*Р.В. Арутюнян, Л.А. Большов, В.В. Демьянов, Е.В. Каневская,
М.Ф. Каневский, В.П. Киселев, И.И. Линге, В.М. Шершаков*

ИНСТИТУТ ПРОБЛЕМ БЕЗОПАСНОГО РАЗВИТИЯ АТОМНОЙ ЭНЕРГЕТИКИ

113191 Москва, ул. Б.Тулская, 52

тел.: (095) 952-24-21, факс: (095) 230-20-29, эл. почта: pbl@ibrae.msk.su

Аннотация

В работе представлена методика анализа пространственно распределенных данных. В работу включено подробное описание известного пакета прикладных программ Geo-EAS, созданного при участии Агентства по Охране Окружающей Среды США (EPA — Environment Protection Agency). Примеры работы с пакетом демонстрируются с использованием данных по Чернобыльским выпадениям.

Содержание

1 Введение	3
2 Описание данных	6
3 Методика	7
4 Анализ данных с помощью пакета Geo-EAS.	8
4.1 Структура файла данных	9
4.2 Некоторые моменты для работы с пакетом	9
4.3 Работа с файлами данных (DATAPREP)	10
4.4 Модификация переменных (TRANS)	12
4.5 Карта точек измерений (POSTPLOT)	14
4.6 Статистическое описание данных (STAT1)	15
4.7 Взаимный разброс (SCATTER)	17
4.8 Подготовка файла пар (PREVAR)	18
4.9 Построение моделей вариограмм (VARIO)	19
4.10 Кросс-валидация (XVALID)	25
4.11 Кригинг (KRIGE)	28
4.12 XY График (XYGRAPH)	31
4.13 Карты изолиний (CONREC)	33
4.14 Просмотр (VIEW)	35
4.15 Плоттер (HPLOT)	35
5 Заключение	36
6 Рисунки 1-49	38

1 Введение

В работе подробно описан известный пакет прикладных программ Geo-EAS, созданный при участии Агентства по Охране Окружающей Среды США (EPA — Environment Protection Agency). Примеры работы с пакетом демонстрируются с использованием данных по Чернобыльским выпадениям.

Основные цели данной работы состояли в следующем:

- описание методики Geo-EAS анализа пространственно распределенных данных
- краткое описание основных моделей геостатистики
- анализ данных Чернобыльских выпадений с помощью пакета Geo-EAS

Ряд результатов по геостатистическому анализу Чернобыльских выпадений был получен в работах [1, 2].

При работе с пространственно распределенными данными встает проблема оценивания значений в точках, в которых измерения не проводились. Оценка значений в произвольных точках требует привлечения дополнительной информации или каких-либо предположений. Для этого нужна модель поведения данных. При наличии такой модели можно проинтерполировать данные измерений по области распределения данных.

Существующие методы интерполяции (анализ тренда поверхности, полиномы низкого порядка, сплайн, триангуляция Делоне, взвешенные движущиеся средние и более сложные) описаны в литературе [3, 4, 5, 6, 7]. Многие из них используются в ГИС картографии и анализе данных. Каждый из приведенных традиционных методов имеет свои возможности и недостатки [8]: безусловно предполагается пространственная зависимость данных; нет никакого критерия для определения сохраняется ли предположение; пространственная вариация не принимается в расчет; многие результаты грубые и склонные к флуктуациям. Еще более важно то, что ни один из методов не обеспечивает предсказание ошибок оценки.

Геостатистика является набором детерминистических и статистических инструментов, направленных на то, чтобы понять и промоделировать пространственную вариабельность. Основная идея предсказательной статистики — характеризовать любое не измеренное значение z как случайную переменную Z , плотность вероятности которой характеризует неопределенность вокруг z .

В настоящее время геостатистика находит более широкую область применения. Все виды пространственно и временно непрерывных полей переменных с пространственной и временной корреляцией данных позволяют использовать эту теорию. Подробное обсуждение геостатистики и смежных проблем приводится в работах [3, 4]. В книге Кресси (Cressie) [3] представлено современное состояние геостатистики. Ниже приведены только формулы и уравнения, использованные в исследовании. Геостатистика имеет дело со статистическим подходом. Предполагается, что данные являются реализациями случайного поля. Чтобы использовать геостатистику, нужно определить на этом поле структуру (автоковариационную функцию), используя доступные данные, и оценить данные в немеренных точках. Следует заметить, что выбор случайного поля для моделирования реализаций переменной дело аналитического удобства. Это не предполагает, что исследуемый феномен действительно имеет случайный характер.

Наиболее известной геостатистической моделью является *кригинг (kriging)*. Кригинг принадлежит к классу лучших линейных несмещенных оценивателей (BLUE — best linear unbiased estimator). Это означает, что:

1. предсказанное среднее должно равняться действительному среднему (известному при простом кригинге (simple kriging); не известному при обычном кригинге (ordinary kriging)),
2. ошибка оценки минимальна.

Кригинг — интерполяция, при которой экспериментальным данным ставится в соответствие набор весов, минимизирующих вариацию. При этом вариация программируется как функция вариограммы и взаимного местоположения

Пусть есть набор измерений $Z(x_i)$, где $i = 1, \dots, N$ — реализации случайного процесса $Z(x)$ и

$$Z(x) = m(x) + \epsilon(x) \quad (1)$$

где среднее значение $\epsilon(x) = E\{\epsilon(x)\} = 0$, x — вектор.

Среднее значение $Z(x)$ равно $m(x)$ и представляет детерминистический тренд. $m(x)$ — гладкая функция x . Все пространственные корреляции представлены $\epsilon(x)$. Обобщение этой пространственной корреляции является существенной частью геостатистики. Предполагается, что известно или можно промоделировать (используя доступные измерения) эти пространственные корреляции.

В известном пакете прикладных программ Geo-EAS используются следующие модели кригинга: простой, обычный, точечный, блочный.

Простой кригинг

Простой (simple) кригинг это оценщик с гипотезой стационарности второго порядка (см. описание программы VARIO) и известным средним значением $m(x) = \text{const}$. Оценка $Z^*(x_0)$ значения функции Z в точке x_0 проводится следующим образом.

$$Z^*(x_0) = \sum_{i=1}^N \lambda_0^i \cdot Z(x_i) \quad (2)$$

Оценщик будет оптимальным (наилучшим) при минимальной ошибке, т.е.

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = \min \quad (3)$$

$Z(x_0)$ — неизвестное действительное значение в x_0

Из этих двух уравнений можно получить N линейных уравнений для весов λ_0^i .

$$\sum_{i=1}^N \lambda_0^i \cdot C(x_i - x_j) = C(x_i - x_0) \quad (4)$$

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = \text{Var}\{Z\} - \sum_{i=1}^N \lambda_0^i \cdot C(x_i - x_0) \quad (5)$$

где $C(x_i - x_0)$ — ковариация (см. описание программы VARIO).

Обычный кригинг

Обычный (ordinary) кригинг это оценщик с внутренней (intrinsic) гипотезой (см. описание программы VARIO) и неизвестным средним значением $m(x) = \text{const}$. Оценка $Z^*(x_0)$ значения функции Z в точке x_0 проводится следующим образом.

$$Z^*(x_0) = \sum_{i=1}^N \lambda_0^i \cdot Z(x_i) \quad (6)$$

Оценщик является несмещенным если:

$$\sum_{i=1}^N \lambda_0^i = 1 \quad (7)$$

и оптимальным (лучшим) если ошибка оценки минимальна, т.е.

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = \min \quad (8)$$

или

$$\frac{1}{2} \text{Var}\{Z^*(x_0) - Z(x_0)\} - \mu \cdot \left(\sum_{i=1}^N \lambda_0^i - 1 \right) = \min \quad (9)$$

где μ — множитель Лагранжа.

Из этих уравнений $(N + 1)$ линейных уравнений для весов λ_0^i и μ

$$\sum_{i=1}^N \lambda_0^i (x_i - x_j) + \mu = \gamma(x_i - x_j) \quad (10)$$

$$\sum_{i=1}^N \lambda_0^i = 1 \quad (11)$$

где $\gamma(x_i - x_j)$ — полувариограмма (см. описание программы VARIO).

Вариация ошибки оценки

$$\text{Var}\{Z^*(x_0) - Z(x_0)\} = \sum_{i=1}^N \lambda_0^i \cdot C(x_i - x_0) + \mu \quad (12)$$

Кригинг является точным интерполятором, поскольку он дает точные значения в точках измерений, т.е. $\text{Var}\{Z^*(x_i) - Z(x_i)\} = 0$. (13)

95% доверительный интервал может быть оценен как $\pm 2\sigma$, где

$$\sigma = \sqrt{\text{Var}\{Z^*(x_0) - Z(x_0)\}} \quad (14)$$

Затем

$$Z^*(x_0) = \sum_{i=1}^N \lambda_0^i Z(x_i) \pm 2\sigma \quad (15)$$

Точечный кригинг

Точечный (point) кригинг — оценка точечного значения из набора окружающих точек. При совпадении оцениваемой точки с экспериментальной точкой значение, полученное из кригинга, совпадает с измеренным значением.

Блочный кригинг

Для более точной оценки можно усреднить несколько точечных оценок, полученных в окрестности оцениваемой точки. Так, вся рассматриваемая область разбивается на блоки, в каждом блоке производится несколько точечных оценок и усредняются. В итоге получается предсказанное локальное значение. Этот метод называется блочным (block) кригингом.

2 Описание данных

Для примера был взят файл данных по загрязнению Калужской области цезием 137 и стронцием 90 kaluga91.dat (см. рис. 1). Данные были получены при содействии НПО "Тайфун". В файле насчитывается 227 точек. Для более подробного рассмотрения выбиралась юго-восточная подобласть с наибольшей плотностью измерений. В этой подобласти находится 165 точек. Измерения попавшие в этот файл производились пространственно нерегулярно, что привело к образованию кластеров. Для декластеризации был использован метод *ячейковой декластеризации (cell declustering)*. Этот метод использует движущееся окно (moving window) для подсчета точек, попавших в ячейку. Значение в каждой точке берется с весом, пропорциональным количеству точек в данной ячейке. При подсчете кластеризованное среднее (2.7875) оказывается в два раза больше декластеризованного (1.3551). Гистограммы взвешенных и не взвешенных значений представлены на рисунке 3. При построении карты распределения весов видно, что наименьшие веса присвоены точкам сконцентрированным в области больших абсолютных значений на юге (см. рис. 2).

Следует заметить, что точечные измерения обладают собственной внутренней структурой, которая не учитывается при усреднении для получения среднего значения загрязнения

в точке. Эту структуру можно установить при анализе данных по загрязнению в отдельных населенных пунктах. Распределение данных в точке может иметь различный вид. Так, распределение может иметь нормальный (рис. 4) или логнормальный характер (рис. 5). В других случаях распределение может быть ассиметрично, причем как с положительным (рис. 6), так и с отрицательным (рис. 7) коэффициентом симметрии. Значения эксцесса так же может быть больше (рис. 8), меньше или близко к значению для нормального распределения. Были обработаны данные загрязнения в точках по Калужской и Брянской областям. По результатам анализа однопараметрической статистики были построены графики зависимости медианы и стандартного отклонения от среднего значения (рис. 9). Линейная зависимость стандартного отклонения говорит о росте вариабильности в точках с ростом среднего значения. Была получена статистика параметров распределения — эксцесса и коэффициента симметрии. Как видно из рисунка 10 эти величины имеют большой разброс вокруг значений для нормального распределения (эксцесс равен трем, а коэффициент симметрии — нулю).

3 Методика

Для обработки пространственно распределенных данных в рамках геостатистики предусмотрена определенная последовательность, которая помогает анализировать данные и получать на каждом этапе содержательную информацию.

Ниже приведена схема методики, примененной в пакете программ Geo-EAS. На каждом этапе используются соответствующие программы пакета.



4 Анализ данных с помощью пакета Geo-EAS.

Программы, представленные в данном пакете, обрабатывают экспериментальные данные в двумерной системе. Тринадцать программ этого пакета предоставляют пользователю широкие возможности при построении двумерных сеток и изолиний по результатам интерполяции (кригинга) исходя из точечных данных.

В процессе обработки данных возможно употребление следующих дополнительных процедур:

- работа с файлами данных (DATAPREP)
- модификация переменных (TRANS)
- построение карт точек измерений (POSTPLOT)
- статистическое описание данных (STAT1)
- вычисление регрессии (SCATTER)
- подготовка файла пар (.pcf) для вариографии (PREVAR)
- построение моделей вариограмм (VARJO)
- кригинг (KRIGE)
- кросс-валидация (XVALID)
- построение изолиний на картах (CONREC)
- построение XY графиков или диаграмм разброса (XYGRAPH)
- просмотр графических файлов (.met) (VIEW)
- печать графических файлов (.met) на HP плоттере (HPPLOT)

Все эти процедуры будут подробнее описаны ниже.

Для большей наглядности описание возможностей Geo-EAS сопровождается анализом данных по загрязнению Калужской области из файла kaluga91.dat.

4.1 Структура файла данных

Для работы с пакетом Geo-EAS необходим файл данных специального формата (формат Geo-EAS). В начале файла располагается область, в которой описывается содержание файла. Далее следуют сами данные записанные через пробел в виде столбцов для каждой переменной. Область описания данных имеет следующую компоновку.

Первая строка — Название

Вторая строка — Количество переменных (NVAR)

Третья — (NVAR)+2-ая строка — Название переменных (единицы измерения) и форматы переменных в Фортране.

Следующие далее данные в каждой строке должны быть отделены друг от друга по крайней мере одним пробелом. Для лучшей читабельности данные могут быть записаны в столбцах, хотя это не обязательно.

Ниже приведено начало файла выбранного для примера. Он содержит четыре переменные: две координаты в км. по осям направленным на восток и на север и данные по загрязнению Cs^{137} и Sr^{90} в Кюри на км².

KALUGA91.DAT - KALUGA REGION. CS137, SR90 (Ci/km²)

⁴	
EAST KM	G15.7
NORTH KM	G15.7
CS137	G15.7
SR90	G15.7

8.876170	-74.50600	12.81000	.2670000
3.775930	-74.84660	11.49000	.2160000
-4.771620	-76.84920	9.310000	.2200000
3.380360	-77.51830	8.270000	.2980000
18.61590	-75.03400	8.120000	.1810000
-2.916100	-77.07320	8.120000	.2330000

4.2 Некоторые моменты для работы с пакетом

Программы можно запускать как по отдельности так и из общего меню программ, которое запускается программой GEOEAS.EXE.

Практически все описанные ниже программы имеют опции **Prefix** и **Quit**, находящиеся в главном меню. В первой необходимо указать логический диск и директорию, в которой находится используемый файл. Например: *d : \geoeas\data*. Вторая является опцией выхода. Такие же опции находятся на каждой ветви вложенных меню. В меню большинства программ имеются опции указания используемого файла. Они называются либо **Data** либо **File**.

Перебор предлагаемых значений производится пробелом. Перемещения по меню программ осуществляется калашниками курсора. Вход в опции меню осуществляется нажатием клавиши <enter> или нажатием первой буквы английского названия опции. Вернуться на предыдущий экран можно с помощью клавиши <q>.

4.3 Работа с файлами данных (DATAPREP)

Программа DATAPREP предоставляет два вида операций: **DOS Utilities** и **Операции с файлами (File Operation)** а также опцию выхода (**Quit**). Все операции, входящие в меню, имеют опции **Выполнить (Execute)** и **Выход (Quit)**, они будут опускаться в дальнейшем описании. Большинство операций требует указания входного и выходного файлов (Input File, Output File). Программа работает с файлами данных формата Geo-EAS.

- В **DOS Utilities**, открывающей доступ к командам DOS, входят следующие команды:
 - Директория (Directory)
 - Печать (Print)
 - Просмотр (List)
 - Копирование (Copy)
 - Переименование (Rename)
 - Удаление (Delete)
 - Команда DOS (DOS Comand) - выход в DOS, для возврата нужно набрать *Exit* <enter>.

Все операции (кроме **Directory**) имеют трех-ветвенное меню, т.е. возможны следующие опции:

- Файл (File)
- Выполнение (Execute)
- Выход (Quit)

- В **Операции с файлами (File Operation)** входят следующие команды:

- **Присоединить (Append)** — указать три файла формата GEO-EAS; при исполнении содержимое второго файла присоединяется к содержимому первого и результат записывается в третий файл. Первый и второй файлы не могут иметь одно и тоже имя.
- **Изъять колонку (Column Extract)** — указать в опции **File** два файла формата GEO-EAS: входной и выходной; указать одну или более переменных в опции **Variable** для записи в выходной файл.

- **Изъять ряд (Row Extract)** — указать в опции **File** два файла формата GEO-EAS: входной и выходной; указать в опции **Subsetting Conditions** условия выборки: <Операнд 1 > <Соотношение> <Операнд 2 >, где <Операнд 1 > — переменная, <Операнд 2 > — переменная или постоянная, <Соотношение> может быть следующим: .LT. меньше, .LE. меньше или равно, .GT. больше, .GE. больше или равно, .EQ. равно, .NE. не равно. В выходной файл попадут записи, удовлетворяющие заданным условиям на ту или иную переменную.
- **Сжать (Compress)** — выделяются все повторяющиеся записи во входном файле, указанном в опции **File**, и оставшиеся записываются в указанный выходной файл.
- **Пронумеровать (ID Variable)** — указать в опции **File** два файла формата GEO-EAS: входной и выходной. При выполнении операции в выходной файл записывается содержание входного файла и добавляется переменная "Sequence #", значения которой являются порядковым номером строки.
- **Объединить (Merge)** — указать в опции **File** три файла формата GEO-EAS: два для объединения и третий выходной. При выполнении операции значения переменной второго файла добавляются к содержанию первого файла в виде отдельной переменной и результат записывается в выходной файл.
Одинаковые переменные, имеющиеся в обоих файлах, объединяются в одну переменную, при этом ей присваиваются значения из первого файла.
- **Доложить (Report)** — указать в опции **File** два файла формата GEO-EAS: входной и выходной. Далее при указании "On" в следующем поле в выходном файле записи (строки) будут пронумерованы. В опции **Variable** указываются переменные для занесения в выходной файл (при переборе переменных можно выбрать "yes"/ "no")
Выходной файл формируется в виде таблицы с указанием имен переменных в верхней строке. Файл делится на страницы по 50 записей на каждой. При выборе более четырех переменных только первые четыре будут напечатаны на первой странице, а следующие четыре на следующей (также имеющей номер 1).
- **Отсортировать (Sort)** — указать в опции **File** два файла формата GEO-EAS: входной и выходной. В опции **Variable** указать переменную для сортировки. При выполнении в выходном файле записи отсортированы в соответствии со значениями выбранной переменной — от меньшего к большему.

4.4 Модификация переменных (TRANS)

Программа TRANS создает, удаляет и изменяет переменные в файле данных формата Geo-EAS. Операции, производимые над переменными, могут быть как унарные (один операнд и один оператор), так и бинарные (два операнда и один оператор). Операндом может быть переменная или постоянное значение. Оператором могут быть операции типа сложения, извлечения квадратного корня и т.д. Результаты операций могут заменять значения существующих переменных или записываться как новые переменные. Пропуск может иметь место в двух случаях: когда пропущено значение операнда или когда операция не определена (деление на ноль).

Для хранения данных используется временный файл ZZSCTCH1.FIL. При занесении модифицированных данных в новый файл данные переносятся из временного файла.

Как и другие программы TRANS работает с файлами, содержащими не более 48 переменных, и использует не более 10,000 значений из файла данных. При превышении этих параметров берутся только первые считанные значения и появляется предупреждение.

Главное меню программы имеет следующие опции:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Файл данных (Data)** — указать файл данных формата Geo-EAS
- **Заголовок (Titel)** — ввести описание выходного файла (первую строку файла — см. структуру файла данных), не более 66-ти символов
- **Создать (Create)** — переход в меню изменения данных

- Удалить (Deleat) — выбрать и удалить одну из существующих переменных из файла
- Записать (Save) — ввести имя файла, в который записываются данные
- Выход (Quit)

В опции Create можно выбрать следующее:

- Существующая переменная (Old Variable) — ввести имя существующей переменной, значения которой нужно изменить
- Новая переменная (New Variable) — завести новую переменную
- Выход (Quit)

Два первых пункта меню после нажатия `< enter >` ведут в следующее меню:

- Унарная операция (Unary Operation) — при входе в опцию Operation позволяет производить следующие операции с одним операндом:

<code>sqrt</code>	— возведение операнда в квадрат
<code>log</code>	— вычисление десятичного логарифма операнда
<code>ln</code>	— вычисление натурального логарифма операнда
<code>truncate</code>	— округление операнда
<code>exp</code>	— взятие экспоненты
<code>ranc</code>	— определение порядка значения переменной
<code>equals</code>	— приравнивание переменной к операнду
<code>abs.val.</code>	— вычисление модуля операнда
<code>sign</code>	— перемена знака операнда

Перебор всех указанных операций осуществляется клавишей пробела.

После выбора операции остается указать операнд: постоянную (Constant) или переменную (Variable); и войти в следующую опцию Execute.

- Бинарная операция (Binary Operation) — при входе в опцию Опция Operation позволяет производить следующие операции с двумя операндами:

<code>+</code>	— сложение
<code>/</code>	— деление первого операнда на второй
<code>x</code>	— умножение
<code>-</code>	— вычитание
<code>**</code>	— возведение первого операнда в степень второго
<code>MAX</code>	— определение максимального значения операндов
<code>MIN</code>	— определение минимального значения операндов
<code>Mean</code>	— определение среднего значения операндов

Перебор всех указанных операций осуществляется клавишей пробела.

После выбора операции остается указать два операнда: постоянные (Constant) или переменные (Variable); и войти в следующую опцию Execute.

- Индикаторное преобразование (Indicator Transform) — этот оператор требует двух операндов: первый операнд существующая переменная, а второй — постоянная (порог). Результат равен 1.0, если переменная больше или равна пороговому значению; и равен 0.0, если переменная меньше порогового значения.

Для выполнения операции нужно указать переменную и порог в опции Variable и войти в опцию Execute.

- Выход (Quit)

4.5 Карта точек измерений (POSTPLOT)

Программа POSTPLOT наносит на карту местоположения точечных данных, давая при этом понятие о их абсолютной величине. Это достигается разбиением всех данных на четыре равные по количеству данных части — квартили. При этом в каждую квартиль входят данные в определенных пределах. После этого точки из каждой квартили обозначаются своим условным значком и цветом на карте. Вместе с этим под графиком указываются предельные значения для каждой квартили.

При общем числе данных не кратном четырем деление на квартили происходит следующим образом: в одной или нескольких квартилях число данных на единицу больше чем в остальных. Порядок увеличения числа данных в квартилях следующий: 4-я квартиль (наибольших значений), 2-я квартиль, 3-я квартиль, 1-я квартиль (наименьших значений).

Программа POSTPLOT имеет следующее главное меню:

- Директория (Prefix) — указать диск и директорию файла данных
- Файл данных (Data) — указать файл данных формата Geo-EAS
- Переменные (Variables) — указать две переменные в качестве координат (*X и Y coordinate variable*) и переменную для нанесения на карту (*variable to plot*).
- Опции (Options) — внутри этого поля можно указать следующее:
 - *Include values* — вывести значения переменной в точках на карте (при выборе “Yes”). По умолчанию — “No”.
 - *Size* — при выборе “Yes” указать размер цифр от 1 до 10. По умолчанию для чисел — “5”.
 - *Scale factor* — указать порядок значений переменной при выведении их на карту (опция *Include values*). Возможный порядок от 10^{-4} до 10^4 . По умолчанию — 10^0 , т.е. выводятся сами значения.
 - *# Decimals* — указать количество знаков для выведения на карту (от 1 до 3) при этом значения округляются до указанного знака. По умолчанию “1”.
 - *Include symbols* — условные значки для обозначения точек. “4” — использовать четыре разных значка; “1” — использовать один значок для всех точек (желтый +); “None” — не использовать условные значки. По умолчанию “4”.
 - *Size* — при выборе “Yes” указать размер цифр от 1 до 10. По умолчанию для условных значков — “4”.

Перебор всех указанных операций осуществляется клавишей пробела.

- Графические опции (Graph options) — вход в графическое меню:
 - Параметры осей (Axis parameters) — возможен следующий выбор: “Half” — только левая и нижняя оси; “Full” — правая, левая, верхняя и нижняя оси; “Grid” — тоже что и “Full” плюс сетка натянутая на большие деления осей. По умолчанию “Full”.
 - Параметры делений осей (Tick parameters) — можно изменить параметры делений, указанные по умолчанию.
 - Пределы для осей (Graph limits) — задать минимальные и максимальные значения для осей X и Y.
 - Заголовок и названия (Titles/Labels) — ввести заголовок, два подзаголовка и название осей X и Y не более 60 символов на каждое.
 - Переустановка (Reset) — переустановить все параметры по умолчанию
 - Выход (Quit)
- Выполнение (Execute) — строится графический файл *.met (по умолчанию Meta-code.met) и выведение карты на экран.
- Выход (Quit)

На рисунке 1 приведена карта взаимного расположения точек измерения загрязнения Cs^{137} Калужской области по квартилям.

4.6 Статистическое описание данных (STAT1)

Программа производит статистическую обработку файла данных в формате Geo-EAS и рисует гистограммы и диаграммы вероятности (probability plot). Возможны вычисления в логарифмическом масштабе и учет "фактора веса" при описании распределения.

STAT1 работает с файлами, содержащими не более 48 переменных, и использует не более 10,000 значений из файла данных.

Для примера будут использоваться данные из файла kaluga91.dat и декластеризованные данные из файла kdec.dat.

Главное меню программы выглядит следующим образом:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Файл данных (Data)** — указать файл данных формата Geo-EAS
- **Переменная (Variable)** — указать переменную, для которой будет проводится статистический анализ распределения. Здесь же можно указать переменную в качестве веса (*Weight*) или выбрать логарифмический масштаб ("On" в опции *Log*). При переходе к взвешенным значениям переменная умножается на соответствующее значение веса полученное при декластеризации.

Так при работе с декластеризованными данными файла kdec.dat с помощью пробела укажем: *Variable* — Cs137 ; *Weight* — Weight.

- **Пределы (Limits)** — указать область изменения значений переменной (по умолчанию берется минимальное и максимальное значения).

Для выбранного файла: *Minimum* .019 *Maximum* 12.810

- **Итоговая статистика (Batch Statistics)** — производит запись статистических параметров для всех переменных в файл или выведение на принтер. Файл статистической информации не имеет формата Geo-EAS.
- **Выполнение (Execute)** — вход в следующее меню
- **Выход (Quit)**

Следующее меню дает возможность выбрать вид демонстрации результатов:

- **Гистограмма (Histogram)** — просмотр и вход в следующее меню;
- **Диаграмма вероятности (Probability Plot)** — просмотр нормальной бумаги;
- **Просмотр численных результатов (Examine)** — в шести колонках выведены: пронумерованные данные, их веса, накопленный (кумулятивный) процент, вероятность, номер в файле данных.

К графикам приложены диаграммы квартилей (*Box plot*), на которых отмечены минимальное и максимальное значения, среднее (mean "x") и срединное (median "|") значения, а также четыре квартили — группы с равным числом данных, расположенных в порядке возрастания. Вторая и третья квартили расположены в прямоугольнике и разделены срединным значением (медианой).

Наряду с графической информацией вычисляются следующие параметры распределения

- *Среднее значение (Mean)*
- *Вариация (Variance)*
- *Стандартное отклонение (Std.Dev.)*
- *Коэффициент вариации (% C.V.)*
- *Коэффициент симметрии (Skewness)*
- *Экцесс (Kurtosis)*

- Значения *второй и третьей квартилей* — область срединных значений (минимальное значение второй квартили — 25% и максимальное значение третьей квартили — 75%)
- *Медиана (median)* — срединное значение

Гистограмма — частота появления значений — имеет свое собственное меню:

- **Тип гистограммы (Type)** — абсолютный (Absolute) или относительный (Relative) (в процентах к общему числу данных). Устанавливается с помощью пробела, по умолчанию — абсолютный.
- **Размеры колонок (Class Limits)** — задать минимальное значение, ширину колонки и количество колонок в гистограмме.
- **Оси координат (Axes)** — задать минимальные и максимальные значения и градацию для осей X и Y.
- **Заголовок (Titels)** — ввод заголовка, подзаголовка и названий осей. Для каждого отведено до 60-ти символов. По умолчанию принимаются: "Histogram", "Weighted Histogram", "Data file:[имя входного файла]", "[имя переменной]", "Frequency".
- **Результаты (Results)** — просмотр результатов по всем колонкам: верхний предел, абсолютная и относительная частота, максимальное значение, сумма весов и среднее значение для каждой колонки (класса).
- **Просмотр гистограммы (View Graph)**
- **Выход (Quit)**

Для сравнения представлены абсолютные гистограммы кластеризованных (рис. 11) и декластеризованных данных (рис. 12) в обычных и логарифмических переменных. Видно, что декластеризация привела к уменьшению среднего значения и вариации из-за того, что в области высоких значений было больше измерений, чем в области низких значений. Декластеризованные данные менее размазаны по величине. Гистограмма логарифмов для декластеризованных данных (рис. 13) является практически симметричной — среднее значение почти совпадает с срединным и находится посередине между максимальным и минимальным значением, в отличие от гистограммы логарифмов кластеризованных данных (рис. 14).

Обратимся теперь к диаграмме вероятности (*probability plot*) (рис. 15). Это зависимость между абсолютными значениями и суммарным процентом частоты появления данных (*cumulative percent*). Диаграмма вероятности показывает насколько данное распределение близко к нормальному. При нормальном распределении график — прямая $y = x$. При работе с логарифмами прямая будет означать лог-нормальное распределение. Для лог-нормального распределения характерна высокая асимметрия, что можно заметить в распределении декластеризованных данных (рис. 17). В этом случае полезно перейти к логарифмическим переменным при построении вариограмм, что и было сделано. Видно, что данные второй и третьей квартилей в логарифмических переменных представляют собой почти прямую (рис. 16).

4.7 Взаимный разброс (SCATTER)

Программа SCATTER демонстрирует зависимость значений двух переменных из файла данных формата Geo-EAS. Возможно получить графики в логарифмическом и полу-логарифмическом масштабе. Градуирование осей и заголовки формируются автоматически.

Ограничения, налагаемые на файл данных, такие же как и для программы STAT1: не более 48 переменных, и не более 10,000 значений из файла данных.

Меню имеет следующую структуру:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Файл данных (Data)** — указать файл данных формата Geo-EAS

- **Переменные (Variables)** — указать две переменные (X,Y), для которых считается зависимость. По умолчанию это соответственно первая и вторая переменные в файле данных. Для каждой из координат можно выбрать логарифмический масштаб ("On" в опции *Log*). Все изменения производятся перебором возможных вариантов с помощью пробела.
- **Опции (Options)**
 - *Регрессия (Regression)* — вычисление регрессии: $Y = Slope * X + Intercept$ и коэффициента корреляции (*Correlation coefficient*) и построение прямой на том же графике. Возможен выбор "Да" и "Нет" (Yes/No). По умолчанию "Да".
 - *Одинаковый масштаб (Equal Scaling)* — возможность построения графика с одинаковым масштабом по обеим осям (выбор "Да"). По умолчанию — "Нет".
- **Выполнение (Execute)** — просмотр диаграммы взаимного разброса и результатов регрессии. Для выхода обратно в меню — < q >.
- **Выход (Quit)**

Рассмотрим взаимную корреляцию между данными по Cs^{137} и Sr^{90} из файла *kaluga91.dat* (рис. 19). Она достаточно велика — коэффициент корреляции равен 0.746, что объясняется спецификой источника при распространении загрязнения. Коэффициент корреляции несколько больше при обращении к логарифмическому масштабу — 0.791 (рис. 20).

4.8 Подготовка файла пар (PREVAR)

Программа создает файл пар (PCF — Pair Comparison File) для дальнейшего использования при анализе вариограмм и моделировании в программе VARIO. PCF это двоичный файл, который содержит пары данных, отсортированные по расстоянию. Он также включает в себя или копию начального файла данных или ссылку на него. Работа с файлом пар намного ускоряет вычисление вариограмм. Можно оценить количество пар в файле:

$$N_{pair} = \left(\frac{N \cdot (N - 1)}{2} \right)$$

где N — количество данных.

Программа PREVAR может создать файл, содержащий не более 16384 пары. При большом количестве данных это значение меньше числа всех возможных пар. При этом в файл пар записываются первые 16384 пары. Предусмотрена и другая процедура записи в PCF, которая регулируется опцией *Subset*. В этом случае в PCF попадают пары выбранные по случайной выборке с заданным зерном (*Seed* — первым числом при генерации "псевдослучайной" последовательности) и долей записываемых пар от общего числа (*Fraction*). При использовании этой операции каждой генерируемой паре присваивается случайное число от 0 до 1, и далее в файл записываются только пары у которых это число меньше установленного значения *fraction* (от 0 до 1).

Программа имеет следующие опции:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Файл (Files)** — указать входной файл .dat и выходной файл .pcf
- **Переменные (Variables)** — указать переменные в качестве координат X,Y пар
- **Пределы (Limits)** — указать в каких пределах выбираются пары (минимальные и максимальные координаты по осям X и Y)
- **Доустановка (Subset)** — указать *fraction* от 0 до 1 (см. выше) и зерно (*seed*) для генерации последовательности случайных чисел
- **Выполнение (Execute)**
- **Выход (Quit)**

При исполнении опции **Выполнение (Execute)** перед тем как пары записываются PCF в они сортируются по расстоянию.

При работе с файлом *kaluga91.dat* были выбраны следующие параметры для отбора пар: *fraction* = 0.6, *seed* = 515. В результате работы программы было получено 15,402 пар.

4.9 Построение моделей вариограмм (VARIO)

Данная программа дает возможность анализировать и промоделировать структуру пространственно распределенных данных в двумерии. VARIO работает с файлом пар (PCF), который создается программой PREVAR. При вариографии вычисляются значения экспериментальных вариограмм и других статистических параметров для заданного набора расстояний между парами — *лэгов (lags)*. Экспериментальные значения вариограмм могут быть представлены на графике в зависимости от расстояния. Большую помощь при вариограммном анализе может оказать просмотр результатов по отдельным лэгам, представленных в различных видах: гистограммы лэгов, диаграммы квартилей (*box plot*), карты измерений (*post plot*), диаграммы разброса значений в лэге (*lag-scatter plots*) и записаны в отдельный файл данных формата Geo-EAS. Модель вариограммы может быть подобрана с помощью от одной до четырех гнездовых (аддитивных) типов вариограмм (*nested structure*). Кроме вариограмм можно построить другие функции, описывающие пространственную непрерывность данных: мадограмма (*Madogram*), обратная ковариация (*Inverted Covariance Variogram*), относительная вариограмма (*Relative Variogram*).

Приведем краткий обзор величин (моментов первого и второго порядка), встречающихся в вариографии.

В общем нестационарном случае моменты не существуют. Но при введении некоторых предположений закон распределения случайной величины может быть описан моментами первого и второго порядка.

Момент первого порядка — математическое ожидание.

$$m(x) = E\{Z(x)\}$$

Моменты второго порядка

- **Вариация (Variance)**

$$\text{Var}\{Z(x)\} = E\{(Z(x) - m(x))^2\}$$

- **Полувариограмма (Semivariogram)** — половина вариограммы (*variogram*) — вариация разницы значений переменной в двух точках как функция от расстояния между ними и направления:

$$\gamma(x_1, x_2) = \frac{1}{2} \text{Var}\{Z(x_1) - Z(x_2)\} = E\{(Z(x_1) - Z(x_2))^2\}$$

Для *N* экспериментальных точек:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (u(x_i) - u(x_i + h))^2$$

При этом пары данных с близкими расстояниями и направлениями обычно группируются в классы (лэги).

Для существования трех выше названных величин необходимо принять *внутреннюю гипотезу (intrinsic hypothesis)*.

Внутренняя гипотеза (Intrinsic hypothesis). Случайная функция $Z(x)$ называется *внутренней (intrinsic)* если:

— математическое ожидание $m(x)$ существует и не зависит от местоположения x ;

- для любого вектора h разность $(Z(x_1) - Z(x_2))$ имеет конечную вариацию, не зависящую от x (стационарность вариограммы):

$$\frac{1}{2} \text{Var}\{Z(x+h) - Z(x)\} = E\{(Z(x+h) - Z(x))^2\} = \gamma(h) \quad \forall x$$

Для существования других моментов второго порядка необходимо дополнительно ввести стационарность второго порядка.

Стационарность второго порядка (Second order stationarity)

Случайная функция $Z(X)$ обладает стационарностью второго порядка, если:

- математическое ожидание $m(x)$ существует и не зависит от местоположения x :

$$m(x) = E\{Z(x)\} \quad \forall x$$

- для каждой пары случайных переменных $\{Z(x), Z(x+h)\}$ ковариация существует и зависит только от расстояния между ними h :

$$C(h) = E\{Z(x)Z(x+h)\} - m^2 \quad \forall x$$

- Ковариация (Covariance) — статистическая мера корреляции между двумя значениями.

$$C(h) = E\{(Z(x_1) - m(x_1))(Z(x_2) - m(x_2))\}$$

Для N экспериментальных точек:

$$C(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} (u(x_i) \cdot u(x_i+h) - m_{-h} \cdot m_{+h})$$

где:

m_{-h} среднее значение хвостовых значений — $m_{-h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} u(x_i)$

m_{+h} среднее значение головных значений — $m_{+h} = \frac{1}{N(h)} \sum_{i=1}^{N(h)} u(x_i+h)$

- Относительная вариограмма (Relative Variogram) — вариограмма, в которой значение обычной вариограммы для каждого лага делится на квадрат среднего значения величин, попавших в лаг.

$$Ral = \frac{\gamma(h)}{m_h^2}$$

Это бывает полезно, когда присутствует эффект пропорциональности (*proportional effect*), т.е. когда области с большей концентрацией чем средняя также имеют более большую ковариацию чем средняя. При использовании моделей относительных вариограмм в кригинге стандартные отклонения кригинга представляют собой десятые доли значений оценок.

- Обратная ковариация (Inverted Covariance Variogram) — вариограмма, получаемая при вычитании ковариаций лагов из экспериментальных вариаций. Такой подход учитывает случаи, когда среднее головных точек отличается от среднего хвостовых точек.

$$S.I.C.(h) = \frac{\sigma^2 - C(h)}{\sigma^2}$$

где вариация данных:

$$\sigma^2 = \text{Var}\{Z(x)\} = E\{(Z(x) - m(x))^2\} = E\{Z(x)^2\} - m(x)^2$$

Обратная ковариация может использоваться в кригинге тем же образом, что и обычная полувариограмма.

- **Мадограмма (Madogram)** — среднее значение разницы между измерениями в парах как функция от расстояния между ними и направления:

$$M(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} |u(x_i) - u(x_i + h)|$$

Мадограмма не является действительной вариограммой и обычно не должна быть использована в кригинге.

Рассмотрим возможности которые предоставляет программа VARIO. Главное меню программы выглядит следующим образом:

- **Директория (Prefix)** — указать диск и директорию файла пар
- **Файл данных (Data)** — указать файл пар (.pcf)
- **Переменная (Variable)** — указать переменную для построения вариограммы (Log Option позволяет строить вариограмму в логарифмическом масштабе, что иногда дает дополнительную информацию).
- **Пределы (Limits)** — указать пределы изменения значения выбранной переменной (по умолчанию принимаются минимальное и максимальное значения)
- **Опции/Выполнение (Options/Execute)** — вход в последующие меню
- **Выход (Quit)** — выход из программы

Выбрав файл kaluga91.pcf, проведем вариографию по переменной Cs^{137} , пределы изменения оставив по умолчанию — 0.019–12.810.

Перейдем в следующее меню через опцию **Options/Execute**. Здесь можно задать параметры для построения вариограммы и посмотреть расположение точек на карте типа post plot:

- **Направление (Direction)** — задать направление ориентации пар (от 0 до 180°), допустимый раcтвор угла (*Tolerance*) (от 0 до 180°) и ширину полосы. *Ширина полосы (Max Bandwidth)* — это максимальная длина перпендикуляра, опущенного из второй точки пары на прямую направления ориентации пар.
- **Новые лаги (New Lags)** — задать минимальное и максимальное значения расстояния между парами и длину лага (*Increment*) — разницу между минимальным и максимальным расстояниями до вторых точек пар в лаге.
- **Изменение лагов (Change Lags)** — редактирование уже введенных лагов
- **Диаграмма опытных точек (Post Plot)** — просмотр расположения точек на карте без информации о значении переменной в этих точках
Замечено, что диаграмма содержит все точки файла данных *.dat не зависимо от того какие из них были включены в файл пар *.pcf.
- **Выполнение (Execute)** — вход в меню просмотра результатов вариографии
- **Выход (Quit)** — в предыдущее меню

Вариограмма по всем направлениям — $Tolerance=90$ — имеет псевдопериодический вид (рис. 24). Это означает наличие двух или более популяций данных с различной пространственной структурой или наличие эффекта дырки (*hole effect*). Построив вариограмму в направлении $Direction=90$, можно увидеть, что в этом направлении пространственная структуризация данных наибольшая (рис. 25). Обратившись к диаграмме опытных точек (**Post Plot**)(рис. 1), увидим две большие группы точек, расположенных с запада на восток. С помощью программы DATAPREP можно выделить их в отдельные файлы данных путем разбиения файла kaluga91.dat и рассматривать по отдельности.

Подготовим файл пар из данных самой крупной юго-восточной группы данных протяженностью на восток EAST -50 – 85 км. и на север NORTH -105 – -20 км. Построим вариограммы эти данных по следующим направлениям: 0,45,90 и 135 градусов с раствором угла $Tolerance=30$ и шириной полосы $Max\ Bandwidth=30$ (км). Процесс моделирование будет описан ниже.

Результаты вариографии можно просмотреть и промоделировать, войдя в меню опции (Execute);

- Тип (Type) — указать тип вариограммы для анализа. Возможны следующие типы: вариограмма (Variogram), мадограмма (Madogram), относительная вариограмма (Relative) и обратная ковариация (Inv.Cov.). По умолчанию — Variogram.
- График (Plot) — просмотр графика вариограммы (в зависимости от расстояния).
- Диаграмма квартилей (Box Plot) — просмотр диаграмм квартилей для всех лагов (рис. 21).
- Результаты лагов (Lag Results) — вход в меню просмотра результатов лагов.
- Модель (Model) — вход в меню моделирования.
- Выход (Quit) — в предыдущее меню.

При необходимости можно проанализировать структуру данных, попавших в каждый отдельный лаг, и получить статистическую информацию о них, более полную чем в опции Box Plot. Это достигается при переходе в опцию Lag Results, имеющую свое собственное меню. При этом интересующий лаг выбирается с помощью курсора.

- Гистограмма (Histogram) — просмотр гистограммы лага (рис. 22). Над гистограммой помещается диаграмма квартилей (Box Plot) данного лага.
- Диаграмма взаимного разброса (Scatter Plot) — просмотр зависимости значений первых точек пар от значений вторых точек пар (рис. 23).

Замечено, что при работе с сортированными по возрастанию данными, все точки диаграммы расположены выше прямой $x=y$ т.е. первое значение в паре всегда меньше второго.

- Просмотр (Examine) — просмотр численных значений результатов вариографии для данного лага: номера точек в парах, значения переменной в парах, расстояние между точками в парах, направления пар, квадрат разницы между значениями переменной в парах. Результаты представлены в виде таблицы.
- Запись (Write) — запись результатов вариографии для данного лага в текстовый файл формата Geo-EAS для дальнейшего более детального анализа.
- Выход (Quit)

С правой стороны графика выводится информация о лаге: номер (от нуля), количество пар в лаге, значение вариограммы и параметры вариограммы (направление, расстояние и т.д.) заданные в меню опции Options/Execute.

Перейдем теперь к моделированию. Для чего войдем в меню опции Model.

- Модель (Model) — указать параметры модели: наггет (Nugget), тип модели, сил (Sill) и радиус корреляции (Range).

Модели бывают следующих типов: Самородок (Nugget)

$$\gamma(h) = c$$

Сферическая (Spherical)

$$\gamma(h) = \begin{cases} c \cdot \left(1.5 \frac{h}{a} - 0.5 \left(\frac{h}{a}\right)^3\right) & \text{если } h \leq a \\ c & \text{если } h > a \end{cases}$$

где a — действительный радиус корреляции (*range*).

Гауссова (Gaussian)

$$\gamma(h) = c \cdot \left(1 - \exp\left(-\frac{3h^2}{a^2}\right) \right)$$

где a — эффективный радиус корреляции (*range*).

Экспоненциальная (Exponential)

$$\gamma(h) = c \cdot \left(1 - \exp\left(-\frac{3h}{a}\right) \right)$$

где a — эффективный радиус корреляции (*range*).

Линейная (Linear)

$$\gamma(h) = c \cdot h$$

Во всех моделях c — *slope* — может быть только положительным.

Наилучшая искомая модель вариограммы может быть найдена путем комбинации нескольких моделей различных типов с одним и тем же наггетом. При этом для каждой модели надо задать сил (*Sill*) (разницу между действительным силов и наггетом) и радиус корреляции (*range*). В этом случае окончательная модель будет равна сумме моделей с различными радиусами корреляции (*range*), что свидетельствует о наличии гнездовой структуры (*nested structure*) данных.

• График (Plot) — просмотр графиков экспериментальной и модельной вариограмм.

• Опции (Options) — вход в меню изменения параметров графика:

- Заголовок (Titels) — ввод заголовка, подзаголовка и названий осей X и Y (по 66 символов каждое). По умолчанию вводятся тип вариограммы и “Distance”.
- Расстояние между делениями (Tic Spaces) — указать расстояние между делениями по осям X и Y. Можно оставить по умолчанию.
- Пределы (Limits) — указать минимальные и максимальные значения по осям X и Y. По умолчанию берутся по оси X — минимальное и максимальное значения вариограммы; по оси Y — 0 и максимальное расстояние между парами.
- Выход (Quit) — в предыдущее меню.

• Выход (Quit) — в предыдущее меню.

При рассмотрении всей области оказалось довольно сложно построить модели вариограмм в виду того, что в области находятся несколько популяций данных. Эти популяции вытянуты с востока на запад и расположены друг на другом в направлении на север. Это отражается на вариограммах (рис. 24, 25). Рассмотрим моделирование вариограмм для Cs^{137} по выделенной юго-восточной области. Вариограмму по всем направлениям (рис. 26) промоделировать не удается, поэтому были выбраны модели вариограмм по четырем направлениям (рис. 27 – 31):

Nugget (Nugget) 0.0

Направление	Тип модели	Sill	Range
0°	Spheric.	7.8	52
45°	Linear	13	26
90°	Linear	13	25
135°	Spheric.	7	35

Для направления 45° также хорошо подходит модель:

Gaussian 12 25

Можно заметить, что при рассмотрении вариограммы в направлении 90° эффект дырки (“hole effect”) является наиболее четко (рис. 25). На расстоянии 60–90 км. наблюдается появление другой структуры. Это и есть появление другой популяции — группы точек севернее рассматриваемой подобласти.

- *R Minor* — длина малой полуоси эллипса поиска экспериментальных точек.
 - *Angle* — угол ориентации эллипса (от 0 до 180).
 - *Min. Dist.* — указать наименьшее расстояние от оцениваемой точки до соседнего значения в окрестности. По умолчанию принимается ноль, при этом все опытные точки в окрестности будут использоваться при оценивании.
 - *Distance Type* — тип измерения расстояния: евклидово (Euclidian по умолчанию), значение вариограммы (Variogram).
 - *Num.Sectors* — число секторов разбиения эллипса: "1" (по умолчанию), "4", "8".
 - *Max.Pts/Sector* — максимальное число точек в секторе (24 для одного сектора и 64 в случае четырех или восьми секторов).
 - *Min Pts. to Use* — минимальное число точек в секторе (по умолчанию "1")
 - *Empty Sectors* — максимальное возможное число секторов, не содержащих точек (определяется числом секторов *Num.Sectors*).
- **Модель (Model)** — указать параметры модели вариограмм: тип модели наггет (nugget), сил (sill), радиус корреляции (range — большой (Major) и малый (Minor), близкие по смыслу к *R Major* и *R Minor*) и угол (см. описание программы VARIO). При простом (Simple) кригинге необходимо указать среднее значение (Global Mean) (см. описание программы KRIGE). При обычном (Ordinary) кригинге оно не нужно и опускается (по умолчанию). Возможно одновременное использование до четырех различных моделей, как при моделировании в программе VARIO.
 - **Выполнение (Execute)** — выполнение кросс-валидации и вход в меню просмотра результатов.
 - **Отладчик (Debug)** — вход в эту опцию не влечет за собой каких-либо действий, а только выдает подсказку о клавишах отладчика. При выполнении программы можно посмотреть на карте выбранные окрестности поиска <caps lock> (рис. 35); таблицу весов кригинга <num lock>; значения ковариаций между оцениваемыми значениями и соседними точками <scroll lock>. При работе с отладчиком для перехода на следующую точку расчета нужно нажимать <enter>. При работе без отладчика это происходит автоматически. Для выхода из отладчика во время счета следует деактивировать соответствующие клавиши и нажать <q>.
 - **Выход (Quit)** — выход в главное меню
- После выполнения кросс-валидации пользователь автоматически попадает в меню просмотра результатов:
- **Карта ошибок (Error Map)** — просмотр разницы между оценками и наблюдаемыми значениями в точках на карте (рис. 37). Если точка отмечена символом красный +, значит оценка превосходит наблюдаемое значение в этой точке; если точка отмечена символом синий x, значит оценка меньше наблюдаемой величины. Величина разницы пропорциональна размерам символа в обоих случаях. Справа от карты выводятся статистические параметры распределения: количество данных, среднее значение, стандартное отклонение, минимальное и максимальное значения, медиана и границы второй и третьей квартилей.
 - **Диаграмма разброса (Scatter Plot)** — построение диаграмм зависимости оценок кригинга от наблюдаемых значений ("Observed vs. Estimate" или ошибки в зависимости от оценки ("Estimate vs. Error"). Выбор осуществляется пробелом. Используются те же символы, что и на карте ошибок. Сверху и справа от диаграммы помещаются диаграммы квартилей для каждой из координат.
 - **Гистограмма (Histogram)** — построение гистограммы ошибки оценки (разница оценки и наблюдаемых значений "Frequency vs. $(Z^* - Z)$ ") и гистограммы отношения ошибки к стандартному отклонению кригинга "Frequency vs. $(\frac{Z^* - Z}{Z_{\text{dev.}}})$ ". Выбор осуществляется

нажатием клавиши пробела: Сверху гистограммы помещается диаграмма квартилей. Справа от карты выводятся статистические параметры распределения: количество данных, среднее значение, стандартное отклонение, минимальное и максимальное значения, медиана и границы второй и третьей квартилей.

- **Запись (Write)** — запись результатов в файл данных формата Geo-EAS. При этом создается файл с семью переменными: две координаты, переменная, выбранная для кригинга (Z), оценка кригинга (Z^*), ошибка оценки (разница между оценкой и наблюдением — $Z^* - Z$), стандартное отклонение кригинга (Z_{sdev}) и отношение ошибки оценки к стандартному отклонению кригинга ($Z_{score} = \frac{Z^* - Z}{Z_{sdev}}$). Этот файл может быть использован в других программах пакета Geo-EAS.
- **Просмотр (Examine)** — просмотр результатов кросс-валидации в виде таблицы. В таблицу входят следующие семь величин: измеренное значение переменной, выбранная для кригинга (Z), оценка кригинга (Z^*), ошибка оценки (разница между оценкой и наблюдением — $Z^* - Z$), стандартное отклонение кригинга (Z_{sdev}) и отношение ошибки оценки к стандартному отклонению кригинга ($Z_{score} = \frac{Z^* - Z}{Z_{sdev}}$). Перемещение по таблице осуществляется клавишами курсора.
- **Выход (Quit)** — выход в меню исполнения.

Проведем кросс-валидацию моделей, построенных в программе VARIO (см. описание программы VARIO). Это модели для Cs^{137} в обычном и логарифмическом масштабах. Сравним две модели для $\ln(Cs^{137})$ — оминаправленную (одна вариограмма для всех направлений) и модель по четырем направлениям. Из карт ошибок, диаграмм разброса (рис. 40, 41, 43) и гистограмм (рис. 39, 42) видно, что среднее значение ошибки меньше во второй модели. При работе в обычном масштабе среднее значение ошибки близко к медиане распределения (рис. 39). То же можно сказать и для распределения оцененных значений (рис. 40, 41).

4.11 Кригинг (KRIGE)

С помощью программы KRIGE осуществляется двумерный кригинг.

Кригинг (kriging) — интерполяция, при которой экспериментальным данным ставится в соответствие набор весов, минимизирующих вариацию. При этом вариация программируется как функция вариограммы и взаимного местоположения.

Можно выделить обычный (*ordinary*) и простой (*simple*) кригинг, точечный (*point*) и блочный (*block*) кригинг.

Простой (simple) кригинг — оценивание в рамках гипотезы стационарности второго порядка (*second-order stationarity*) (см. описание программы VARIO). При этом среднее значение постоянно, известно и используется в каждой локальной оценке вместе с точками из указанной окрестности. Этот метод не всегда самый подходящий.

Обычный (ordinary) кригинг — оценивание в рамках внутренней (*intrinsic*) гипотезы (см. описание программы VARIO). При этом среднее значение постоянно но неизвестно. Этот метод — наиболее часто используемый.

Точечный (point) кригинг — оценка точечного значения из набора окружающих точек. При совпадении оцениваемой точки с экспериментальной точкой значение, полученное из кригинга, совпадает с измеренным значением.

Блочный (block) кригинг — оценка значения по блоку из набора окружающих точек. В программе предусмотрен кригинг по блокам 2×2 , 3×3 , 4×4 точек.

Входными данными для кригинга являются файл данных формата Geo-EAS и модели вариограмм, которые можно получить из программы VARIO. Программа KRIGE создает файл данных формата Geo-EAS, в который входят оценки значений, полученные в результате кригинга, с координатами на прямоугольной сетке и стандартные отклонения кригинга (KSD — *kriging standard deviation*). Просмотр нанесенных на карту изолиний осуществляется с помощью программы CORNEC.

Для каждой переменной в кригинге можно указать тип кригинга, окрестность для поиска экспериментальных точек, модель вариограммы и размер сетки. Кригинг может проводиться

для десяти переменных одновременно. Параметры программы можно сохранить в отдельном файле параметров (*.kpf). Во время запуска программы можно использовать отладчик (Debug) для просмотра поступающих результатов кригинга.

Входной файл данных может содержать не менее трех и не более 48 переменных. Две из этих переменных должны быть координатами данных. Количество данных экспериментальных данных ограничивается 1000.

Главное меню программы выглядит следующим образом:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Прочитать параметры (Read Parameters)** — ввести имя файла *.kpf (kriging parameter file) из которого зачитать параметры для программы KRIGE
- **Выполнение (Options/Execute)** — вход в меню выполнения
- **Сохранить параметры (Save Parameters)** — ввести имя файла *.kpf (kriging parameter file) для записи параметров программы KRIGE
- **Выход (Quit)**

Меню выполнения имеет следующие опции:

- **Данные (Data)** — указать имя входного файла данных формата Geo-EAS и имя выходного файла (по умолчанию [имя файла данных].grd). Оба имени могут иметь до 14 символов
- **Полигон (Polygon)** — при необходимости указать файл границ полигона. По умолчанию "NO" — все точки из файла данных используются в кригинге.
- **Тип кригинга (Type)** — указать тип кригинга: обычный (*Ordinary*) или простой (*Simple*) и точечный (*Point*) или блочный (*Block*). Возможны блоки следующих размеров: 2 × 2, 3 × 3, 4 × 4 точек.
- **Сетка (Grid)** — указать для осей X и Y следующие параметры:
 - переменная (*Variable*)
 - координаты левого нижнего угла сетки (*Origin*)
 - размер деления (*Spacing*)
 - количество делений, не более 100 (*Number*)
- **Поиск (Search)** — указать следующие параметры окрестности, используемой в кригинге (области, из которой выбираются точки для оценивания):
 - *R Major* — длина большой полуоси эллипса поиска экспериментальных точек.
 - *R Minor* — длина малой полуоси эллипса поиска экспериментальных точек.
 - *Angle* — угол ориентации эллипса (от 0 до 180)
 - *Distance Type* — тип измерения расстояния: евклидово (*Euclidian* по умолчанию), значение вариограммы (*Variogram*)
 - *Num.Sectors* — число секторов разбиения эллипса: "1" (по умолчанию), "4", "8"
 - *Max.Pts/Sector* — максимальное число точек в секторе (24 для одного сектора и 64 в случае четырех или восьми секторов)
 - *Min Pts. to Use* — минимальное число точек в секторе (по умолчанию "1")
 - *Empty Sectors* — максимальное возможное число секторов, не содержащих точек (определяется числом секторов *Num.Sectors*)
- **Переменные/Модели (Variables/Models)** — вход в меню описания переменных и моделей вариограмм для кригинга

- **Заголовок (Title)** — указать заголовок (до 60 символов)
- **Выполнение (Execute)** — процесс счета с визуализацией в виде мозаики по квартилям. На мозаичной карте также нанесены опытные точки с разбиением на кварталы. При выполнении можно посмотреть на карте выбранные окрестности поиска <caps lock> (рис. 44); таблицу весов кригинга <num lock>; значения ковариаций между оцениваемыми значениями и соседними точками <scroll lock>. При работе с отладчиком для перехода на следующую точку расчета нужно нажимать <enter>. При работе без отладчика это происходит автоматически. Для выхода из отладчика во время счета следует деактивировать соответствующие клавиши и нажать <q>.

При необходимости можно прервать процесс счета одновременным нажатием одной из следующих комбинаций: <ctrl> и <alt>, <ctrl> и <shift>, <left shift> и <right shift>; при условии, что не нажаты клавиши отладчика. Это бывает нужно для внесения изменений в параметры окрестности или модели по результатам работы отладчика.

- **Выход (Quit)**

Далее опишем меню моделирования:

- **Новая переменная (New Variable)** — заведение новой переменной для кригинга
- **Редактирование (Edit)** — редактирование параметров модели. Указать наггет (*nugget*), тип модели, сил (*sill*), радиус корреляции (*range* — большой (*Major*) и малый (*Minor*), близкие по смыслу к *R Major* и *R Minor*) и угол (см. описание программы VARIO). При простом (*Simple*) кригинге необходимо указать среднее значение (*Global Mean*). При обычном (*Ordinary*) кригинге оно не нужно и пропускается (по умолчанию).
- **Удаление (Delete)** — удаление переменной из списка для кригинга
- **Выход (Quit)**

Для примера был проведен кригинг по юго-западной части Калужской области. Переменная для кригинга — Cs^{137} . Были использованы соответствующие модели вариограмм, полученные из программы VARIO. Параметры окрестности поиска были выбраны так, чтобы наилучшим образом учесть различную пространственную структуру данных в направлениях 0° , 135° и в направлениях 45° , 90° (рис. 44).

Был проведен кригинг на сетке с ячейкой 1×1 км. (рис. 45) и на сетке с ячейкой 5×5 км. (рис. 46, 47). Из рисунков видно, что даже небольшое изменение в модели вариограммы (в направлении 90° использована в одном гауссова, а в другом случае линейная модель) влечет к изменениям в оценке, заметным даже на мозаичной карте кварталов. Первый расчет потребовал намного больше времени, чем два последующих.

4.12 XY График (XYGRAPH)

Программа XYGRAPH может рисовать графики от одной до шести переменных на координатной плоскости X,Y на основе данных из файлов формата Geo-EAS. При этом по оси Y откладываются от одной до шести зависимых переменных, а по оси X одна независимая. Наряду с точечными диаграммами можно построить регрессионные прямые. Диаграммы для каждой переменной имеют свой цвет и символ. График сохраняется в графическом файле (*.met). Параметры программы XYGRAPH можно сохранить в файле и использовать при повторных обращениях к программе.

Входной файл данных должен иметь не более 48 переменных и только первые 500 значений каждой переменной будут использованы.

Главное меню программы имеет следующие опции:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Прочитать параметры (Read Parameters)** — ввести имя файла из которого зачитать параметры для программы XYGRAPH

- **Выполнение (Options/Execute)** — вход в меню выполнения
- **Сохранить параметры (Save Parameters)** — ввести имя файла для записи параметров программы XYGRAPH
- **Выход (Quit)** — выход из программы

Меню выполнения (Options/Execute) выглядит следующим образом:

- **Данные (Data)** — указать имя входного файла данных
- **Переменные (Variables)** — указать одну переменную по оси X и от одной до шести переменных по оси Y. По умолчанию берутся первая и вторая переменные из файла данных.
- **Символ/Линия (Symbol/Line)** — указать для каждой из зависимых переменных следующие параметры:
 - *Symbol Type* — номер символа от “0” до “6” для рисования точек (см. таблицу символов и типов линий). По умолчанию символ имеет порядковый номер переменной (от “1” до “6”), “0” означает, что символы не рисуются.
 - *Line Type* — номер типа линии от “0” до “6” для рисования линий (см. таблицу символов и типов линий). По умолчанию принимается “0” — линии не рисуются.
 - *Color* — выбрать цвет диаграммы. Возможны следующие цвета: “Black” (черный), “Blue” (синий), “Red” (красный), “Yellow” (желтый), “Green” (зеленый), “Brown” коричневым. По умолчанию все диаграммы имеют разные цвета.

Таблица символов и типов линий.

Номер	Тип символа	Тип линии
0	нет символов	нет линий
1	+	— — — —
2	o	— — — —
3	<	— — — —
4	x	— — — —
5	□	— — — —
6	x/x	— — — —

- **Регрессия (Regression)** — вычисление регрессии: $Y = Slope * X + Intercept$ и коэффициента корреляции (*Corelation coefficient*) и построение прямой на том же графике. Возможен выбор “Да” и “Нет” (Yes/No). По умолчанию “Да”.
- **Условные обозначения (Legend)** — выбрать местоположение таблицы условных обозначений: “Bottom” внизу, “Right” в верхнем правом углу и “None” без таблицы (по умолчанию).
- **Графические опции (Graph Options)** — вход в меню графика.
- **Выполнение (Execute)** — построение графического файла (по умолчанию Metacode.met) и вывод графика на экран.
- **Выход (Quit)** — выход в главное меню.

В меню опции **Graph Options** можно указать следующие параметры графика:

- **Параметры осей (Axes Parameters)** — выбрать форматы отдельно для каждой из осей X и Y: “Half” — только внизу и слева, “Full” — внизу, вверху, слева и справа, “Grid” — тоже, что “Full” и плюс сетка, натянутая на большие деления. По умолчанию принимается “Full”.

- **Параметры делений (Tick Parameters)** — выбрать формат делений на каждой из осей X и Y. Предусмотрены следующие опции:
 - *# Major Tickmark Division* — указать количество больших делений на осях. Вычисляется по умолчанию.
 - *# Minor Division/Major* — указать количество малых делений на осях. Вычисляется по умолчанию.
 - *Tickmark Label Format* — формат Фортрана, описывающий как значение должно быть напечатано. По умолчанию принимается формат подходящий в большинстве случаев.
 - *Tickmark Label Size* — указать размер шрифтов для подписи делений (от “1” до “10”). По умолчанию принимается “3”.
- **Заголовки/Названия (Titels/Labels)** — выбрать положение заголовка (“Left” слева, “Center” в центре по умолчанию, “Right” справа), ввести заголовок, подзаголовки (до трех строк), названия осей X и Y, размер используемых шрифтов (от “0” до “10”, по умолчанию принимается “6” для заголовка, “5” и “4” для подзаголовков, и “5” для названий осей). Каждая строка может иметь до 60 символов.
- **Аннотация (Annotation)** — выбрать выводить ли названия осей и заголовки на график “Plot axes and titels” (принимается по умолчанию) или не выводить их на график “No annotation”.
- **Выход (Quit)** — выход в меню выполнения

4.13 Карты изолиний (CONREC)

Программа CONREC наносит на карту изолинии значений, полученных в результате кригинга. Входным файлом программы является файл результатов кригинга (*.grd). Программа создает графический файл (*.met) и выводит его на экран. При работе с программой можно получить немедленные результаты, оставив параметры по умолчанию. Опции позволяют изменять параметры осей, изоуровни, заголовки и названия. Параметры программы можно сохранить в файле *.cpf (conrec parameter file) и использовать при повторных обращениях к программе.

Входной файл не должен содержать более 48 переменных и 10000 значений каждой из них. Минимальное количество переменных — 3, координаты X, Y и переменная для рисования изолиний. Размер сетки не должен превосходить 100 ячеек по каждому из направлений.

Главное меню программы имеет следующий вид:

- **Директория (Prefix)** — указать диск и директорию файла данных
- **Прочитать параметры (Read Parameters)** — ввести имя файла *.cpf, из которого зачитать параметры для программы CONREC.
- **Выполнение (Options/Execute)** — вход в меню выполнения
- **Сохранить параметры (Save Parameters)** — ввести имя файла *.cpf для записи параметров программы CONREC.
- **Выход (Quit)** — выход из программы

В меню выполнения указываются входные параметры программы.

- **Данные (Data)** — указать имя входного файла данных формата Geo-EAS, полученного в результате кригинга (*.grd).
- **Переменные (Variables)** — указать две переменные в качестве координат X и Y, и переменную для рисования изолиний: оценка кригинга или стандартное отклонение кригинга (KSD — kriginng standart deviation).

- **Графические опции (Graph Options)** — вход в меню графика.
- **Опции изолиний (Contour Options)** — вход в меню изолиний.
- **Выполнение (Execute)** — построение графического файла (по умолчанию Metacode.met) и вывод карты изолиний на экран.
- **Выход (Quit)** — выход в главное меню.

В меню опции **Graph Options** можно указать следующие параметры графика:

- **Параметры осей (Axes Parameters)** — выбрать форматы отдельно для каждой из осей X и Y: "Half" — только внизу и слева, "Full" — внизу, вверху, слева и справа, "Grid" — тоже, что "Full" и плюс сетка натянутая на большие деления. По умолчанию принимается "Full".
- **Параметры делений (Tick Parameters)** — выбрать формат делений на каждой из осей X и Y. Предусмотрены следующие опции:
 - # *Major Tickmark Division* — указать количество больших делений на осях. Вычисляется по умолчанию.
 - # *Minor Division/Major* — указать количество малых делений на осях. Вычисляется по умолчанию.
 - *Tickmark Label Format* — формат Фортрана, описывающий как значение должно быть напечатано. По умолчанию принимается формат подходящий в большинстве случаев.
 - *Tickmark Label Size* — указать размер шрифтов для подписи делений (от "1" до "10"). По умолчанию принимается "3".
- **Заголовки/Названия (Titels/Labels)** — выбрать положение заголовка ("Left" слева, "Center" в центре по умолчанию, "Right" справа), ввести заголовок, подзаголовки (до трех строк), названия осей X и Y, размер используемых шрифтов (от "0" до "10", по умолчанию принимается "6" для заголовка, "5" и "4" для подзаголовков, и "5" для названий осей). Каждая строка может иметь до 60 символов.
- **Аннотация (Annotation)** — выбрать выводить ли оси и заголовки на график "Plot axes and titels" (принимается по умолчанию) или не выводить их на график "No annotation".
- **Выход (Quit)** — выход в меню выполнения

В меню опции **Contour Options** можно задать параметры изолиний:

- **Новые изоуровни (New Levels)** — ввести значение первой изолинии (*Starting Contour Value*), максимальное значение изолинии (*Maximum Contour Value*) и величину разницы между соседними изолиниями (*Contour Level Increment*). При исполнении этой опции абсолютные значения уровней и значения подписей изолиний печатаются в две колонки в отдельном поле экрана, где их можно редактировать по отдельности, войдя в следующую опцию (**Edit Levels**).
- **Редактирование изоуровней (Edit Levels)** — редактирование абсолютных значений уровней и значений подписей изолиний.
- **Подписи (Labeling)** — указать параметры подписей изолиний:
 - *Labeling Option* — выбрать выводить подписи изолиний на карту — "On" (принимается по умолчанию) или нет — "Off".
 - *Label Skip Factor* — указать периодичность появления подписей на изолиниях (от "1" до "5"). По умолчанию принимается "2", что означает подпись каждой второй изолинии.

- *First Level to Label* — указать номер первой изолинии с подписью (от “1” до “10”). По умолчанию принимается “1”.
- *Label Size* — указать размер символов подписи (от “1” до “10”). По умолчанию принимается “6”.

- **Образец черточек (Dash Pattern)** — указать параметры рисования изоуровней со значением меньшим значения *Dash Cutoff*. Отрезок линии как комбинация символов и пробелов (символ означает непрерывную линию; по умолчанию принимается линия “— —”);

Dash Cutoff — значение изоуровня, при котором изоуровни с меньшими значениями состоят из отрезков указанных в *Dash Pattern*.

Rep/Label — указать количество отрезков линии (*Dash Pattern*) между подписями на изолинии от “2” до “10”. По умолчанию “5”.

- **Аннотация (Annotation)** — выбрать отмечать ли на карте области высоких и низких значений буквами “H” и “L” (по умолчанию “Plot Min/Max”); или нет (“No Annotation”).
- **Сплайн (Spline)** — указать величину сглаженности линий (от 2.5 до 30). Чем больше значение тем менее гладки изолинии. По умолчанию принимается 15.
- **Выход (Quit)** — выход в меню выполнения

На рисунках построены изолинии C_s^{137} после кригинга на сетке с ячейкой 1×1 км. (рис. 48) и на сетке с ячейкой 5×5 км. (рис. 49). Первая карта учитывает более мелкие особенности распределения данных.

4.14 Просмотр (VIEW)

Программа VIEW дает возможность выводить на экран графические файлы (*.met), которые могут быть созданы в процессе выполнения программ CORNEC, POSTPLOT и XYGRAPH. Графики могут демонстрироваться на мониторах SVGA, VGA, EGA в цвете и на мониторах CGA и Hercules в черно-белом варианте.

Главное меню программы имеет следующие опции:

- **Директория (Prefix)** — указать диск и директорию графического файла (*.met)
- **Файл (File)** — указать имя графического файла (*.met)
- **Масштаб (Scale)** — выбрать тип рисования графика: в исходном масштабе “Square” или так, чтобы график занял весь экран “Rectangular”. “Square” принимается по умолчанию.
- **Выполнение (Execute)** — просмотр графика на экране
- **Выход (Quit)** — выход из программы

4.15 Плоттер (HPPLOT)

Программа HPPLOT позволяет транслировать графические файлы (*.met) в файл команд HPGL плоттера. График выводится на плоттер командой DOS PRINT. Плоттер присоединяется через порт (COM1) и посредством запуска файла HPSETUP.BAT. Поддерживаются следующие плоттеры: HP 7470 (двух стержневой), HP 7475 (шести стержневой), HP 7450 (восемью стержневой).

Графические файлы создаются в процессе выполнения программ CORNEC, POSTPLOT и XYGRAPH.

Меню программы имеет следующие опции:

- **Директория (Prefix)** — указать диск и директорию графического файла (*.met)
- **Файл (File)** — указать имя графического файла (*.met)

- **Масштаб (Scale)** — выбрать тип рисования графика: в исходном масштабе "Square" или так, чтобы график занял весь экран "Rectangular". "Square" принимается по умолчанию.
- **Выполнение (Execute)** — указать имя файла команд плоттера
- **Выход (Quit)** — выход из программы

Для вывода подготовленного файла на плоттер сначала нужно запустить файл HPSETUP.BAT и набрать команду PRINT.

5 Заключение

В работе описан один из подходов к анализу пространственно распределенных данных с использованием известного пакета программ Geo-EAS. В рамках этой методики реализованы основные аспекты геостатистики: структурный анализ, кросс-валидация, кригинг. В рамках исследования была ограничена возможность промоделировать анизотропную структуру данных из-за особенностей программ пакета.

Литература

- [1] M. Kanevsky, R. Arutyunyan, L. Bolshov, I. Linge, E. Savel'eva, T. Haas. Spatial Data Analysis of Chernobyl Fallout. Preliminary Results. Preprint NSI-23-93, Moscow 1993.
- [2] M. Kanevsky, R. Arutyunyan, L. Bolshov, I. Linge, E. Savel'eva, T. Haas. Spatial Data Analysis of Chernobyl Fallout. Monitoring network description. Univariate statistics and declustering. Cross-validation. Preprint NSI-1-94, Moscow 1994.
- [3] N. Cressie. Statistics for Spatial Data. John Wiley & Sons, N.Y. 1991.
- [4] C. Deutsch, A.G. Journel. GSLIB, Geostatistical Software Library and User's Guide. Oxford University Press. 1992.
- [5] B.D. Ripley. Spatial Statistics. John Wiley & Sons, N.Y. 1985.
- [6] M. David. Handbook of Applied Advanced Geostatistical Ore Reserve Estimation. Elsevier Applied Science Publishers, Amsterdam B.V., 1988.
- [7] I. Clark. Practical Geostatistics. Elsevier Applied Science Publishers, London and New York, repr. 1984.
- [8] M.A. Oliver, R. Webster. Kriging: a method of interpolation for geographical information systems. Int. J. Geographical Information Systems, 1990, v.4, No.3, 313-332.
- [9] Geo-EAS 1.2.1 — Geostatistical Environmental Assessment Software, User's Guide, E. Englund A. Sparks U.S. EPA, Las Vegas 1991.
- [10] Ed.H. Isaaks, R.M. Srivastava, An Introduction to Applied Geostatistics, Oxford, Oxford University Press, 1989.
- [11] A.G. Journel, C.J. Huijbregts, Mining Geostatistics, London, Academic Press, 1978.

6 Рисунки 1-49

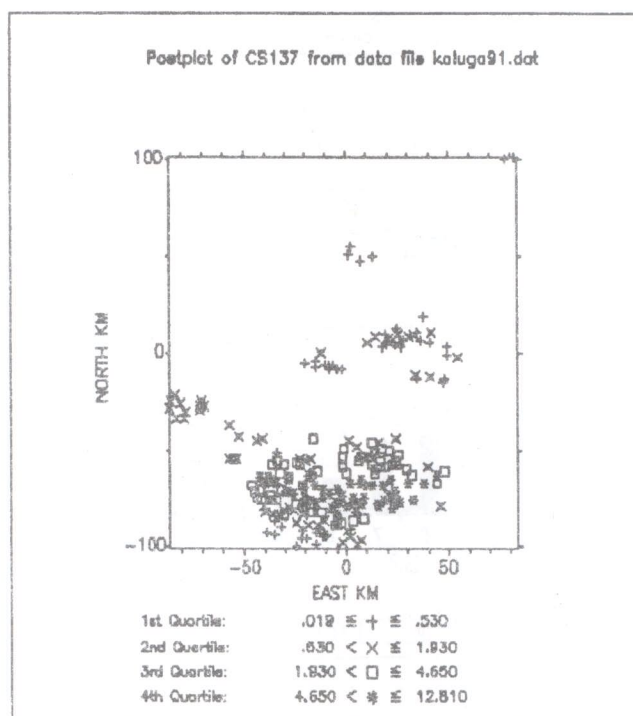


Рис. 1: Карта расположения точек измерений.

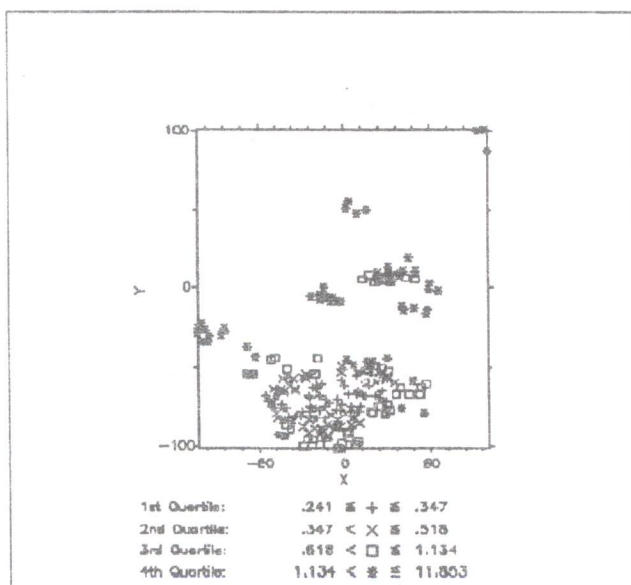


Рис. 2: Карта весов декластеризации по квантилям.

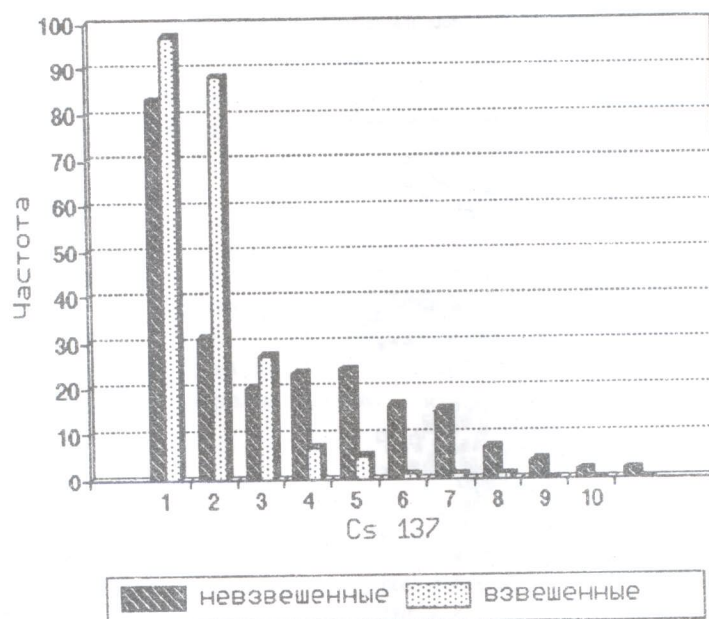


Рис. 3: Гистограмма распределения опытных точек и взвешенных значений.

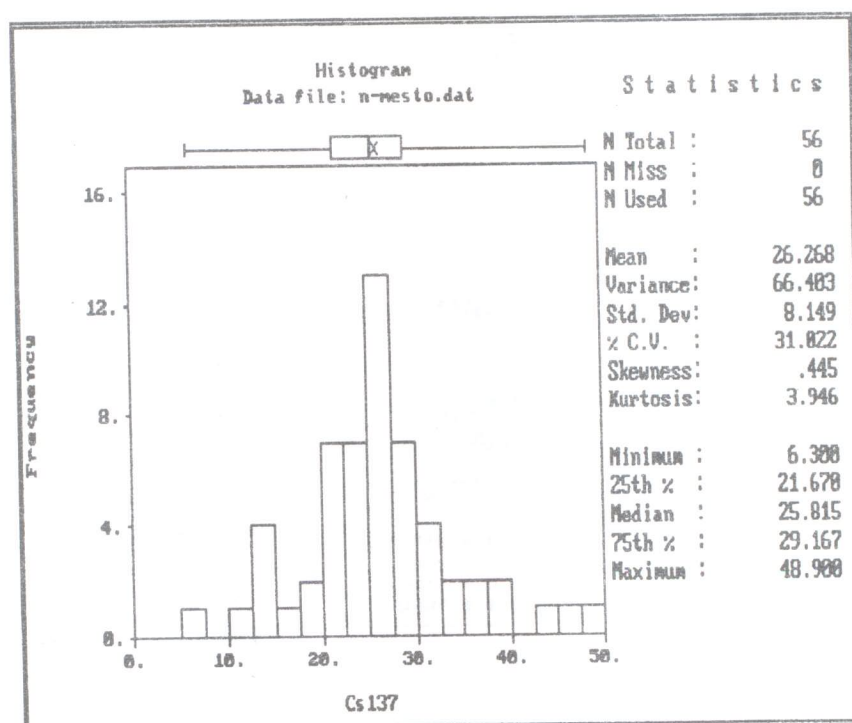


Рис. 4: Распределение данных близкое к нормальному.

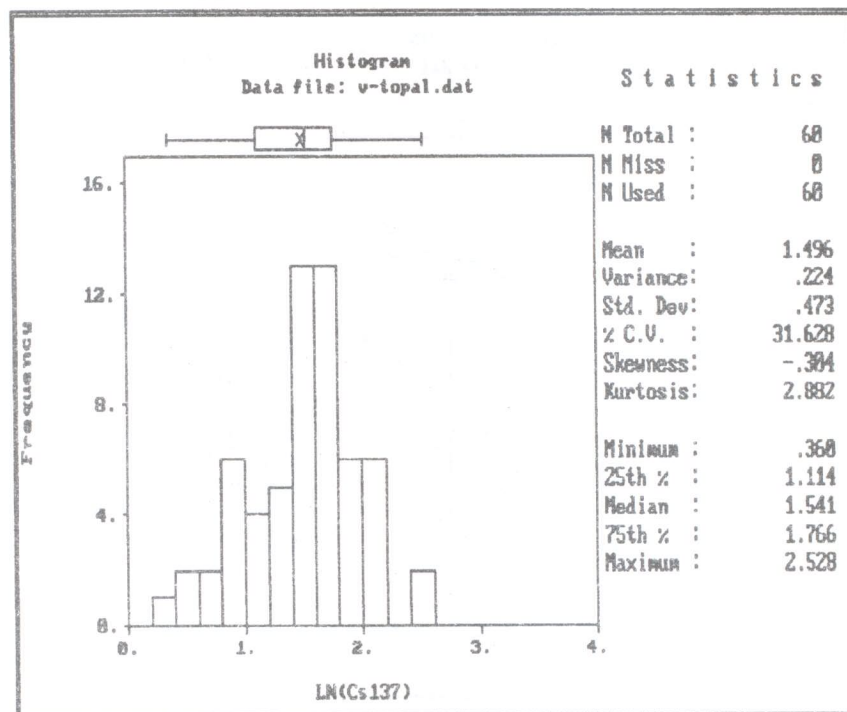


Рис. 5: Распределение данных близкое к логнормальному.

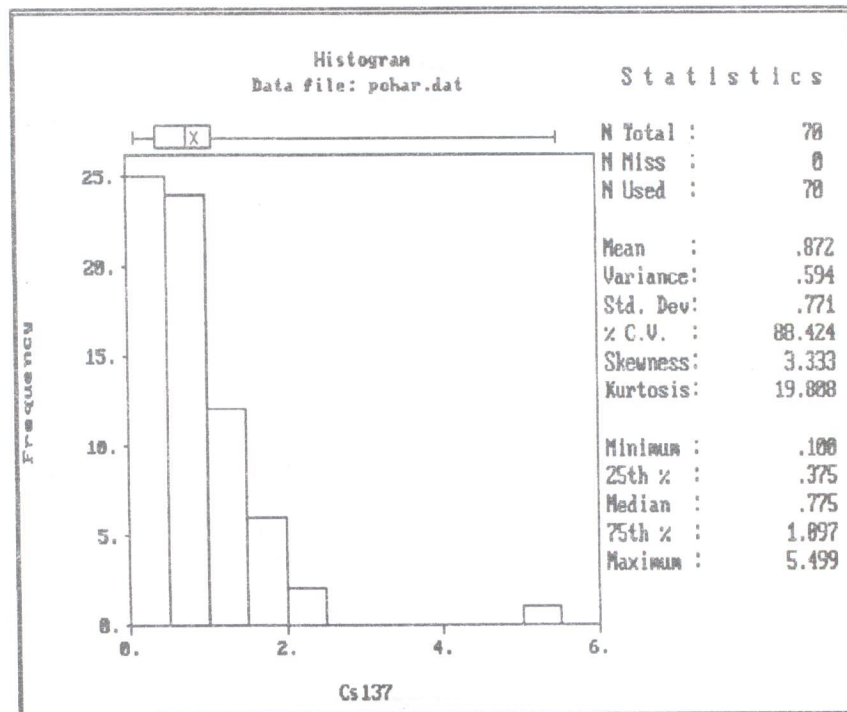


Рис. 6: Распределение данных с положительным коэффициентом симметрии.

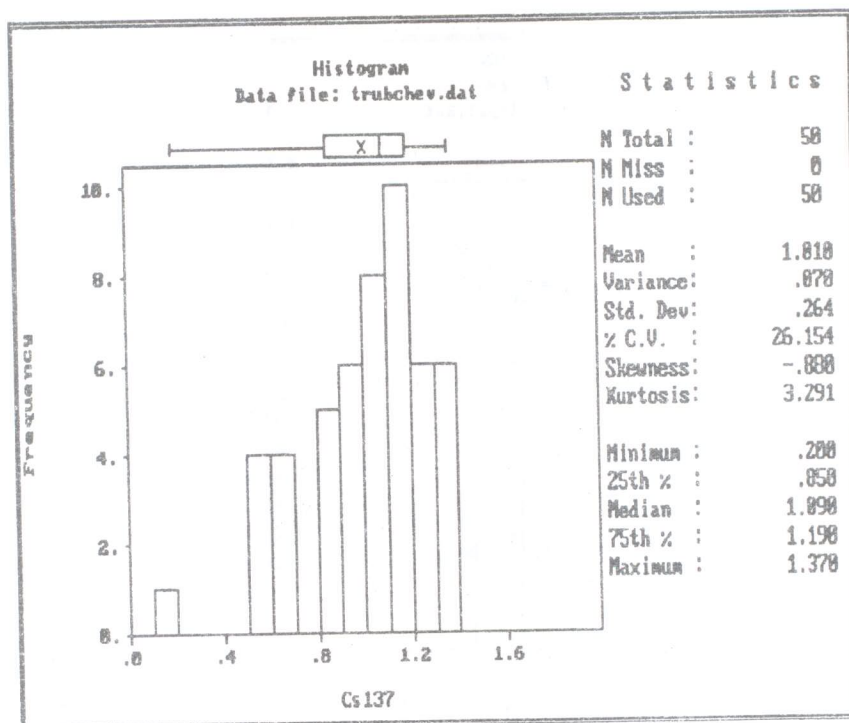


Рис. 7: Распределение данных с отрицательным коэффициентом симметрии.

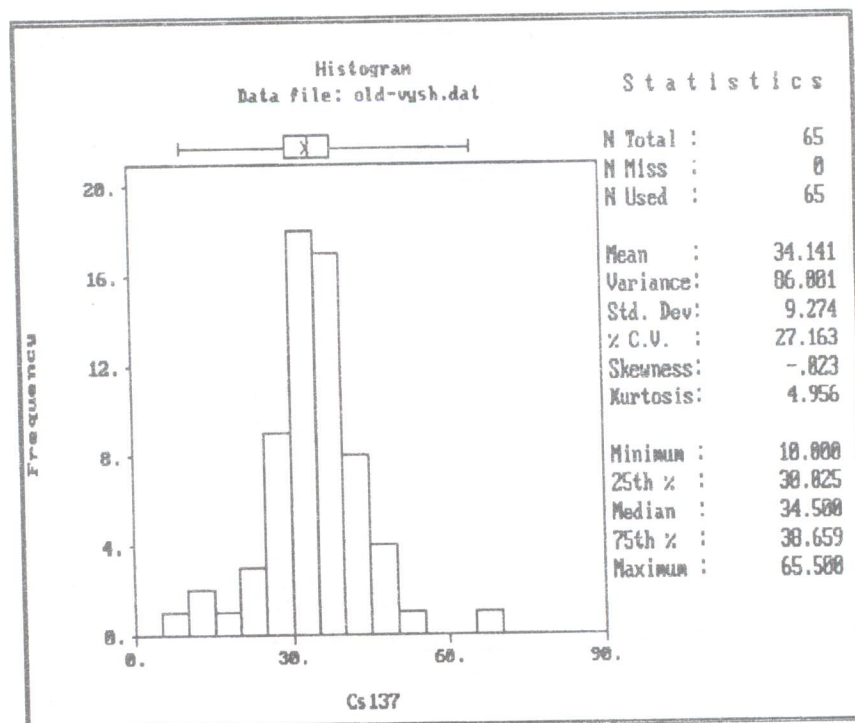


Рис. 8: Распределение данных с большим эксцессом.

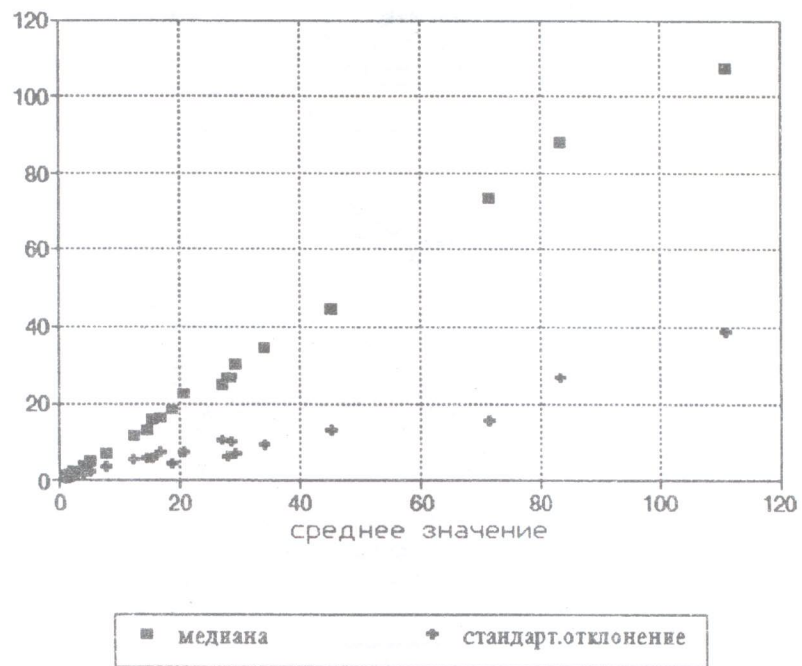


Рис. 9: Графики зависимости медианы и стандартного отклонения от среднего значения распределений в точках измерений.

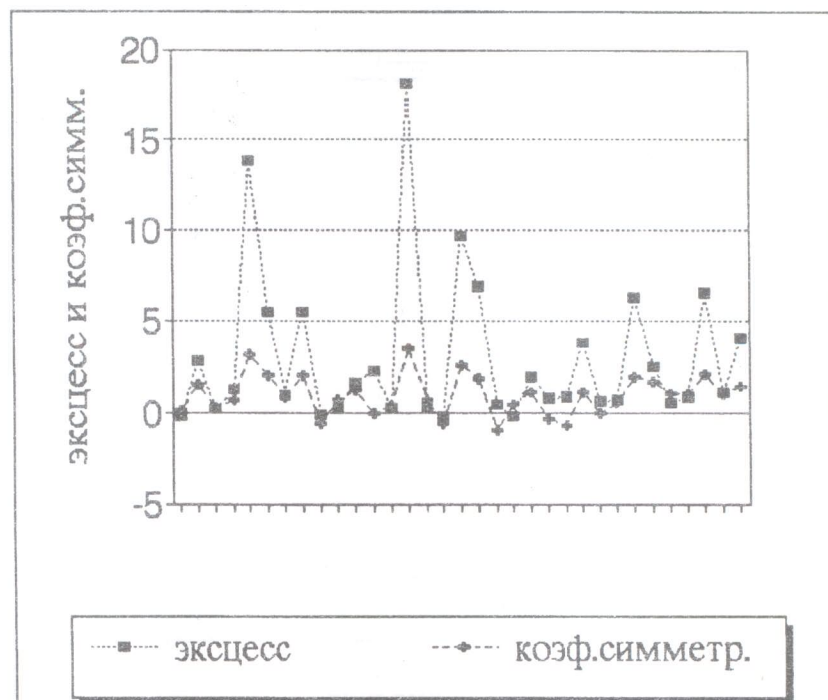


Рис. 10: Диаграмма эксцесса и коэффициента симметрии распределений в точках измерений.

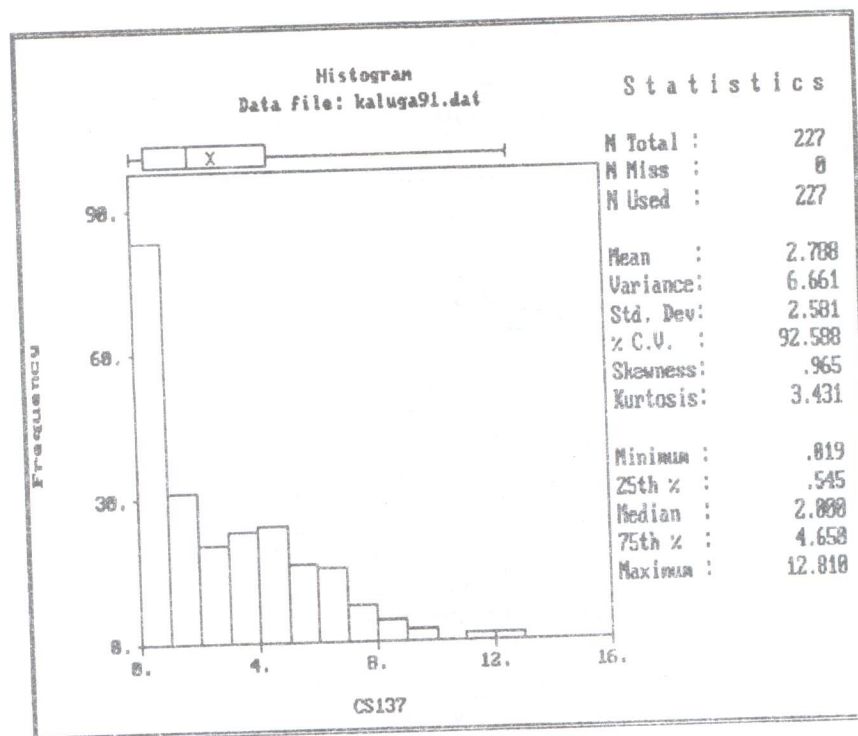


Рис. 11: Гистограмма распределения измеренных данных.

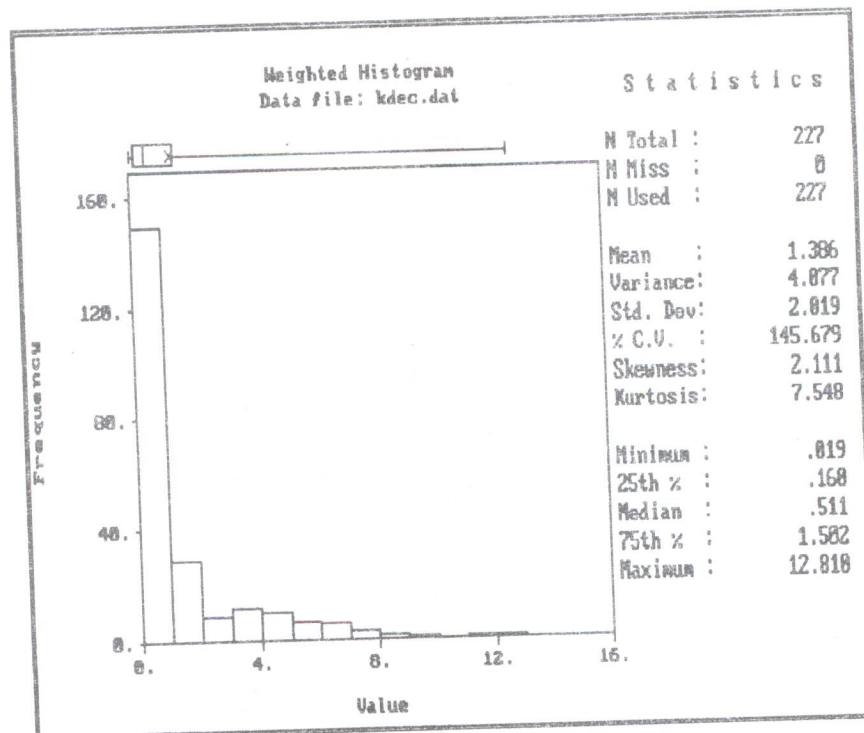


Рис. 12: Гистограмма распределения декластеризованных данных.

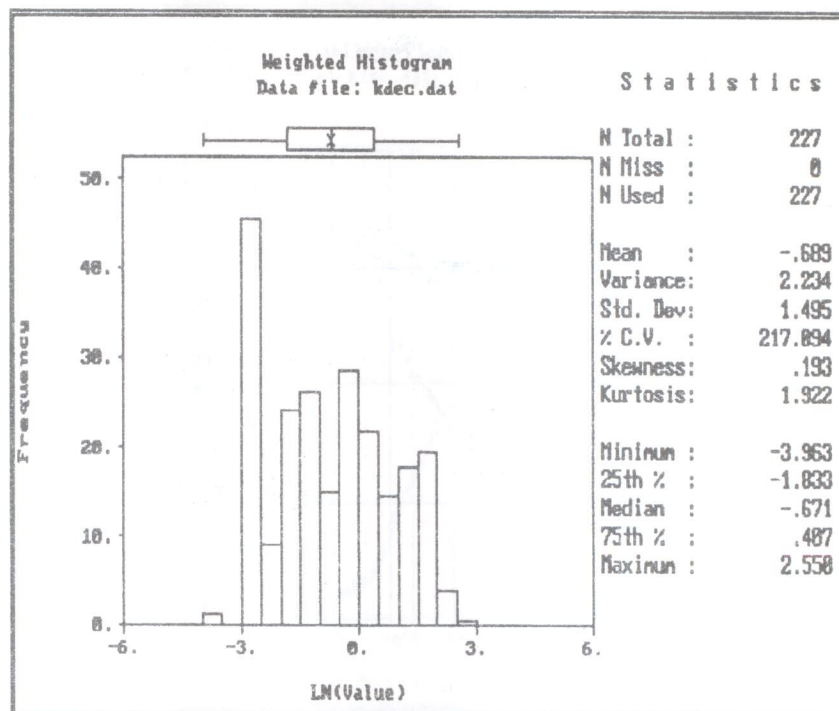


Рис. 13: Гистограмма распределения логарифмов декластеризованных данных.

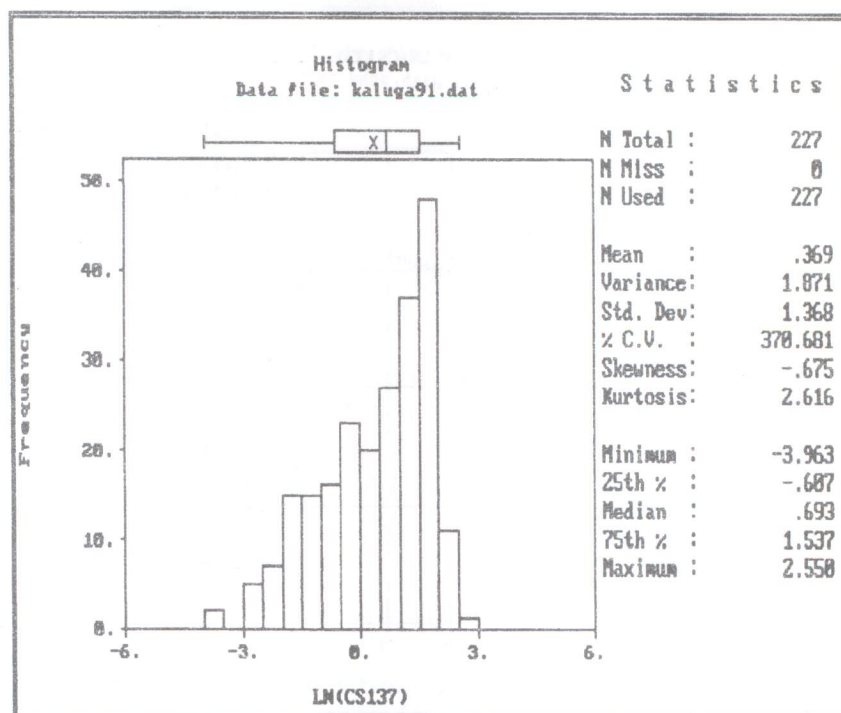


Рис. 14: Гистограмма распределения логарифмов измеренных данных.

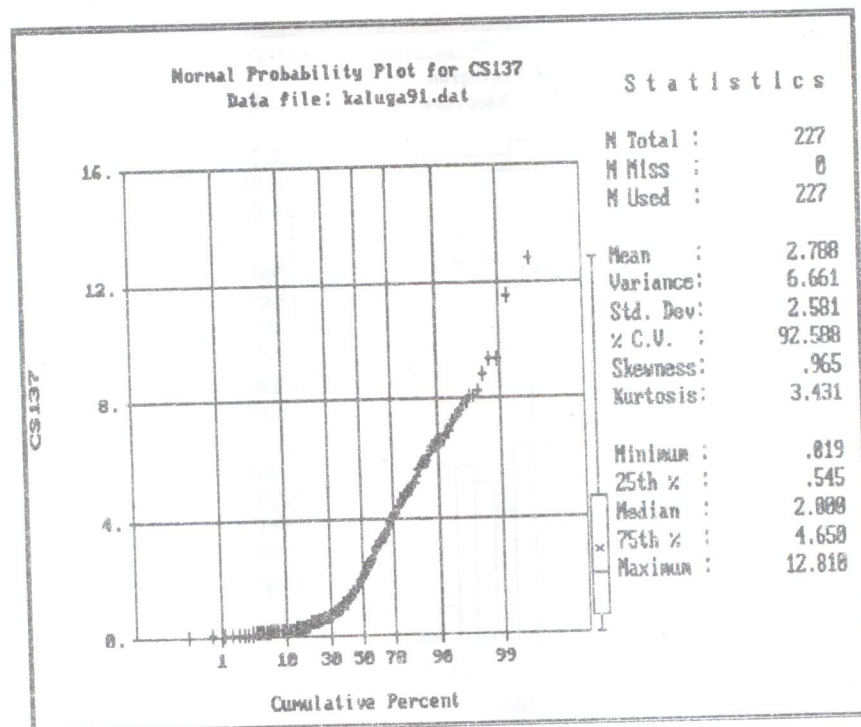


Рис. 15: Диаграмма вероятности измеренных данных.

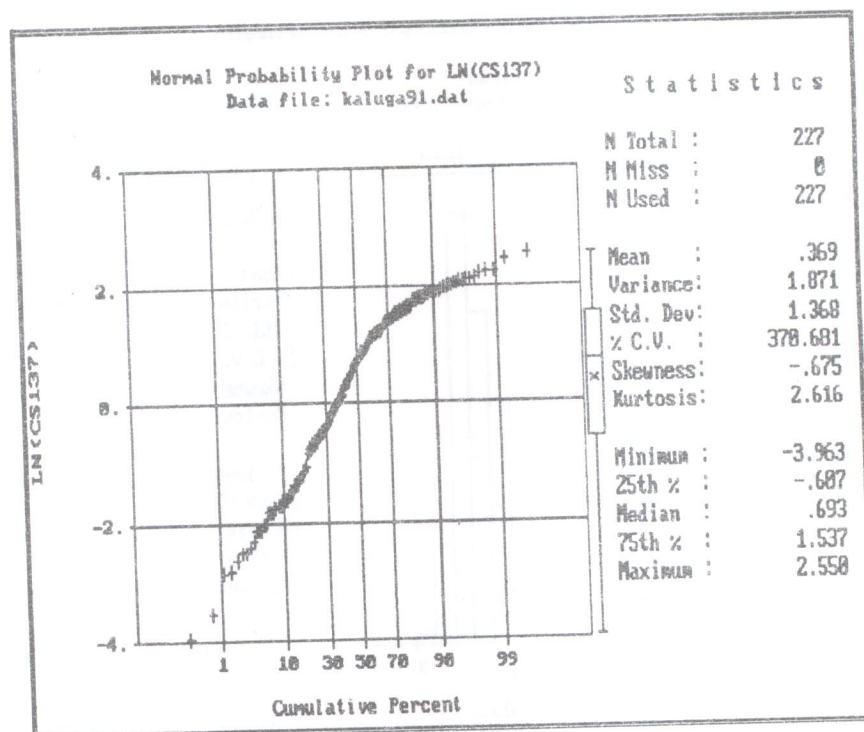


Рис. 16: Диаграмма вероятности логарифмов измеренных данных.

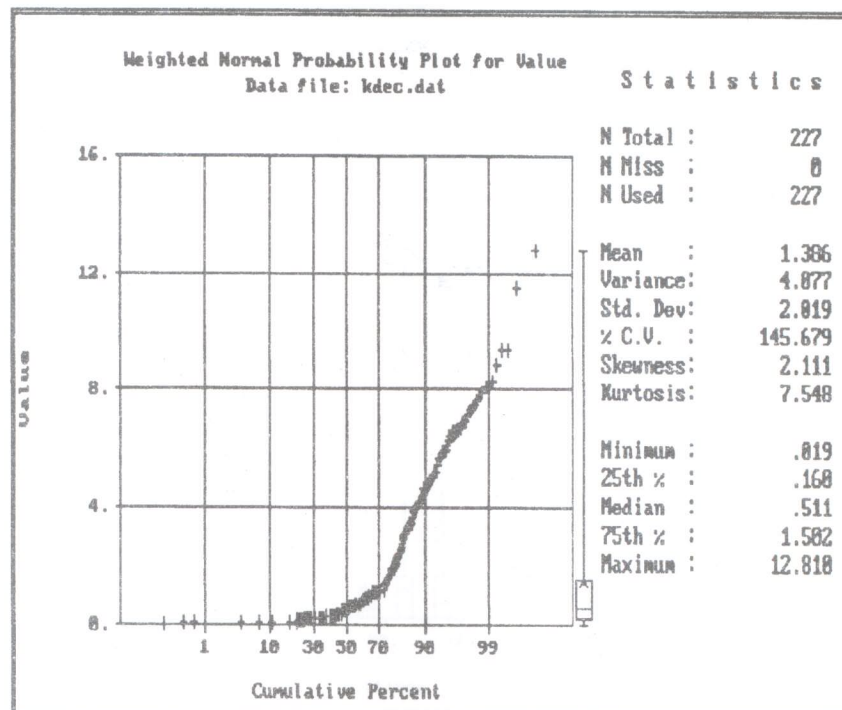


Рис. 17: Диаграмма вероятности декластеризованных данных.

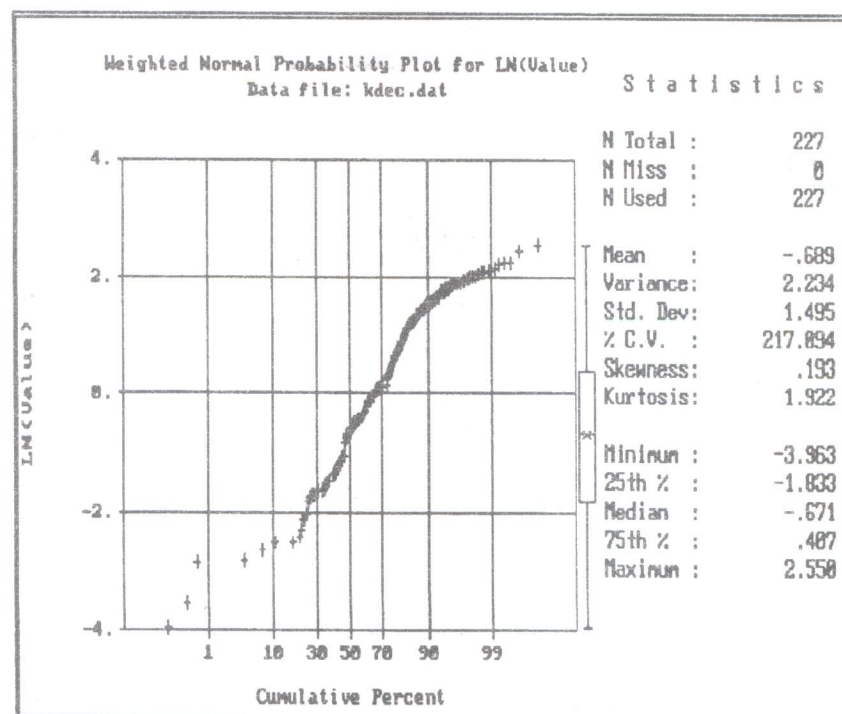


Рис. 18: Диаграмма вероятности логарифмов декластеризованных данных.

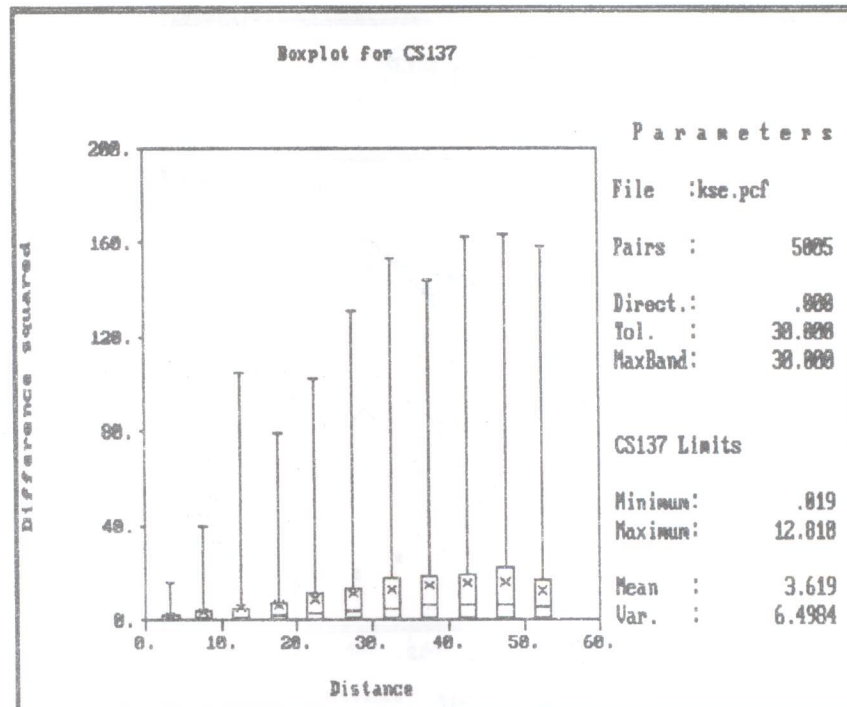


Рис. 21: Диаграмма квартилей для всех лагов.

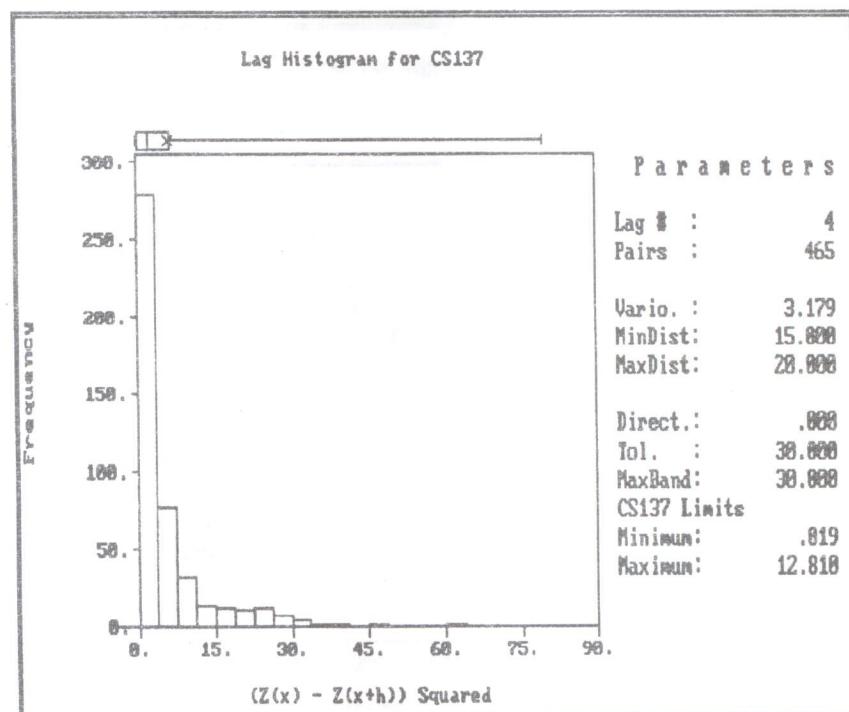


Рис. 22: Гистограмма распределения значений для одного лага.

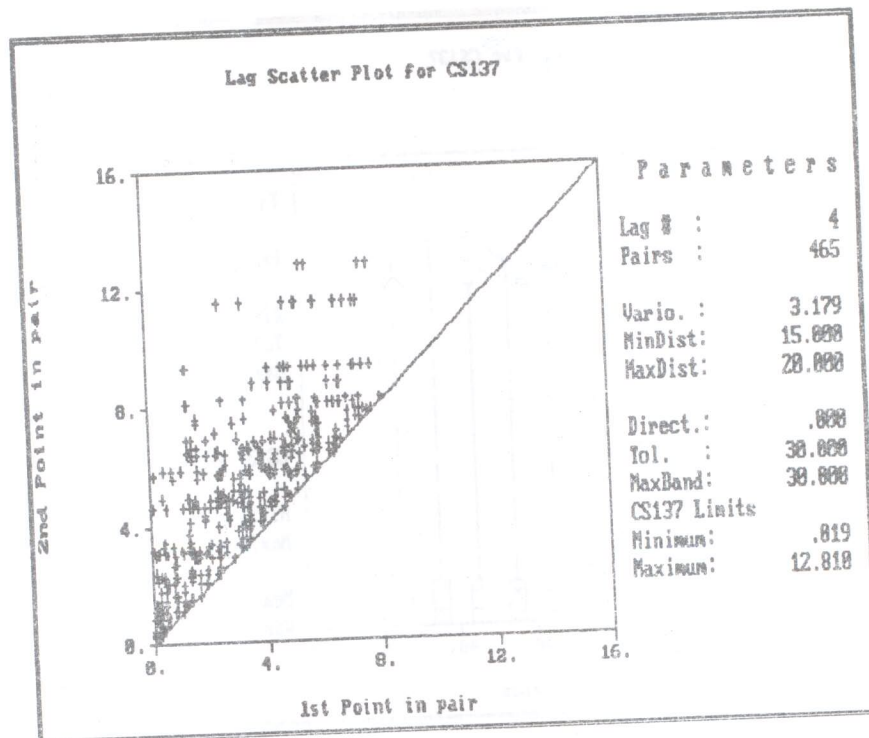


Рис. 23: Диаграмма разброса значений в парах для одного лага.

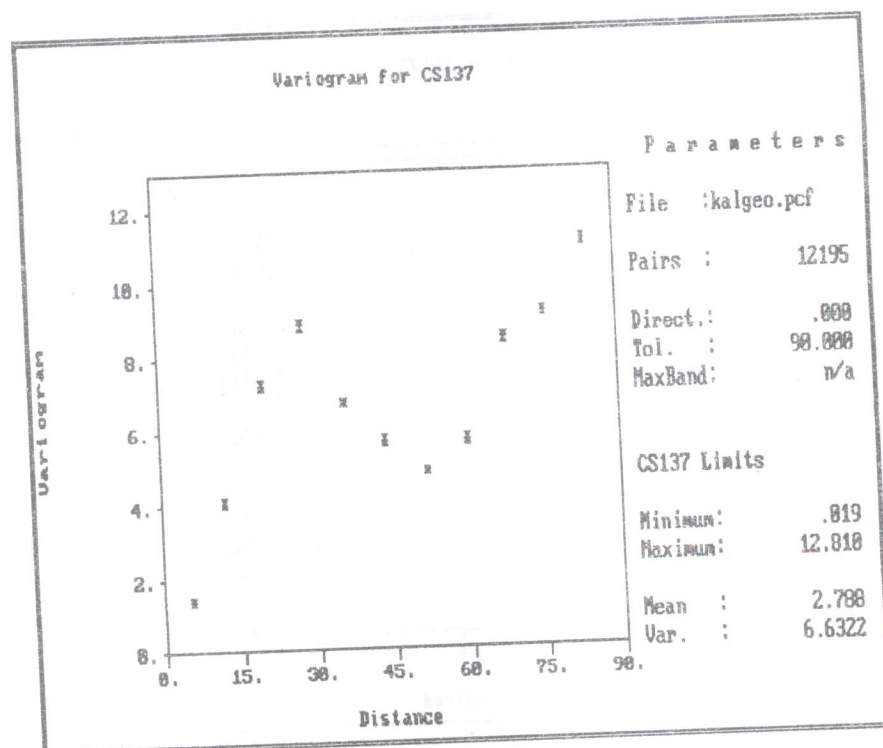


Рис. 24: Вариограмма по всем направлениям по всей области.

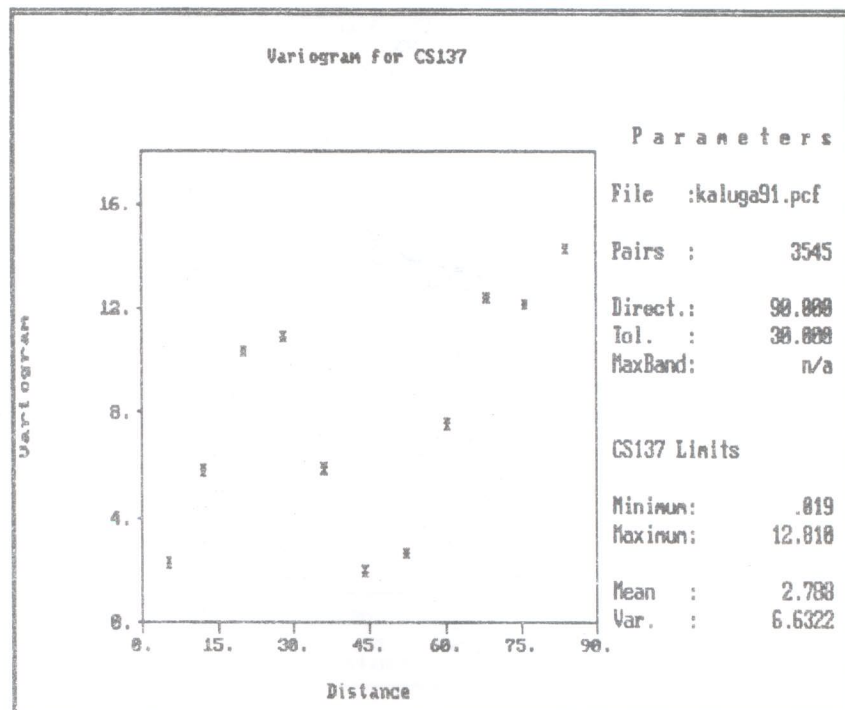


Рис. 25: Вариограмма в направлении 90° по всей области.

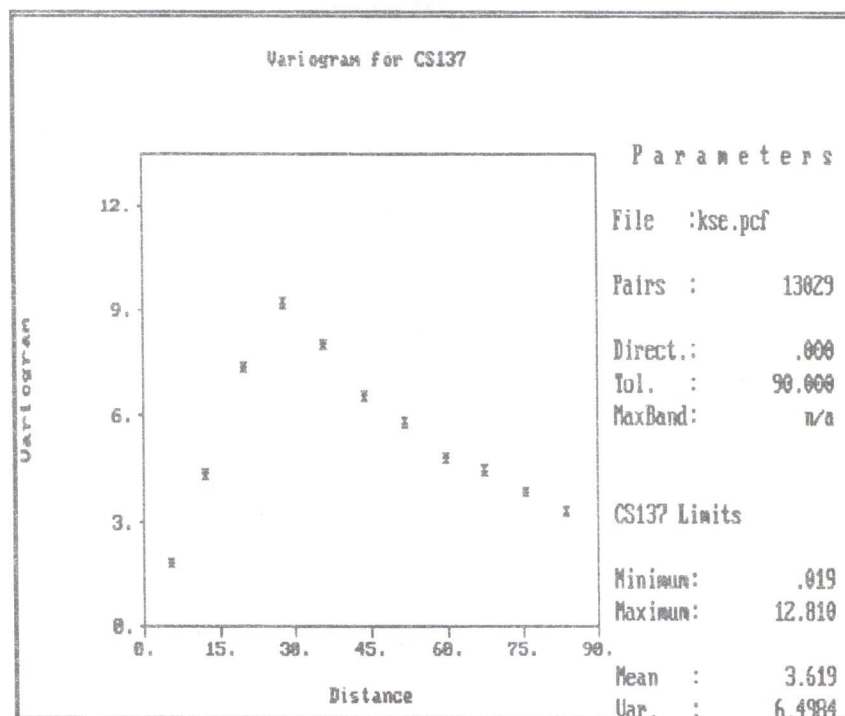


Рис. 26: Вариограмма по всем направлениям по юго-восточной области.

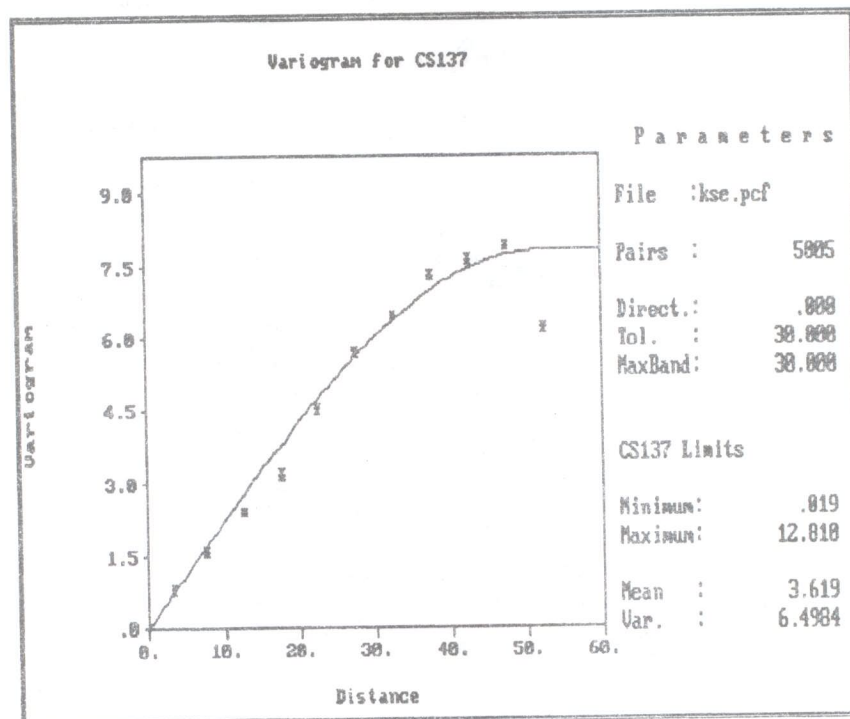


Рис. 27: Модель вариограммы в направлении 0° по юго-восточной области.

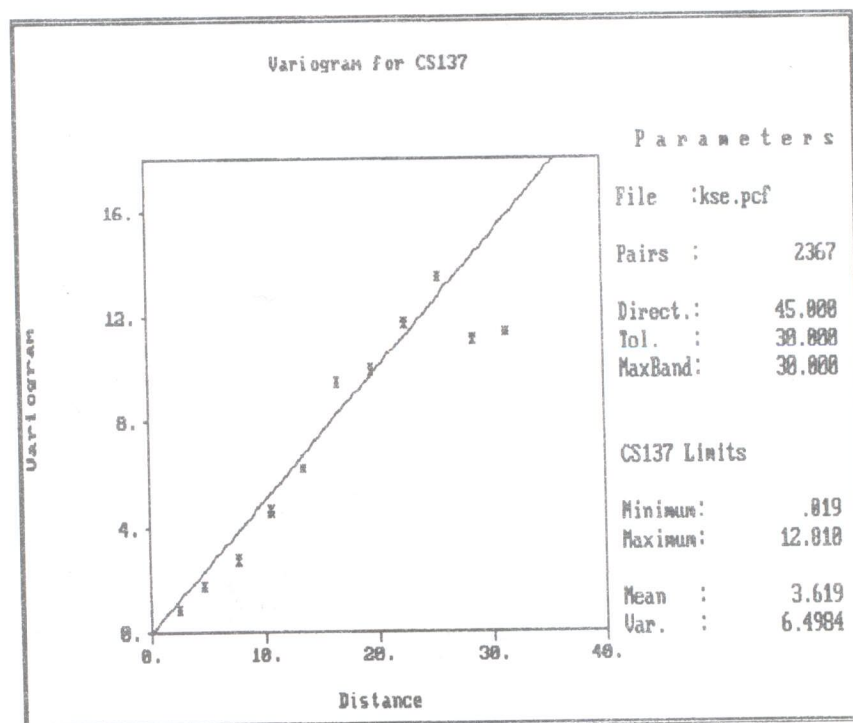


Рис. 28: Линейная модель вариограммы в направлении 45° по юго-восточной области.

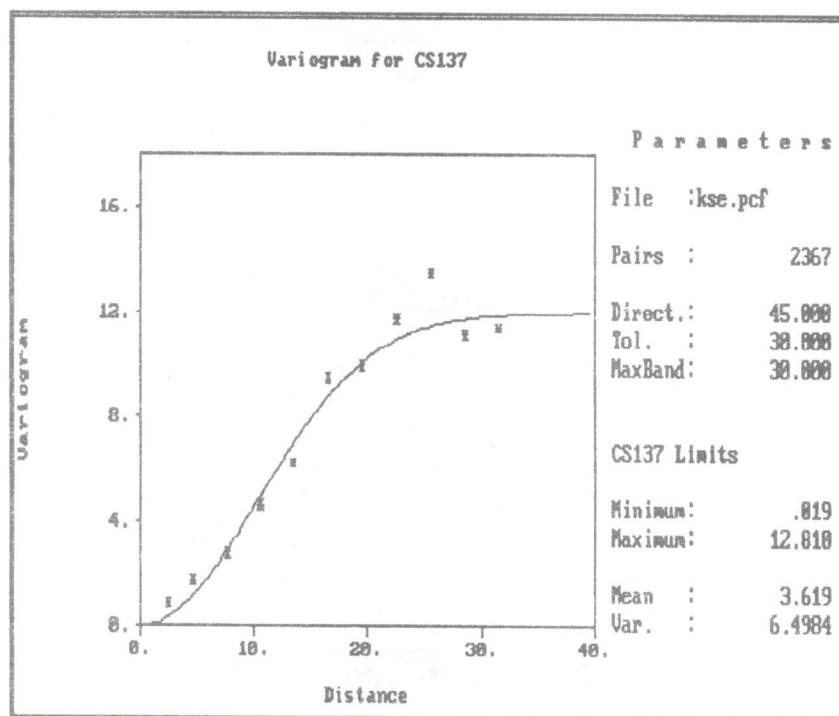


Рис. 29: Гауссова модель вариограммы в направлении 45°.

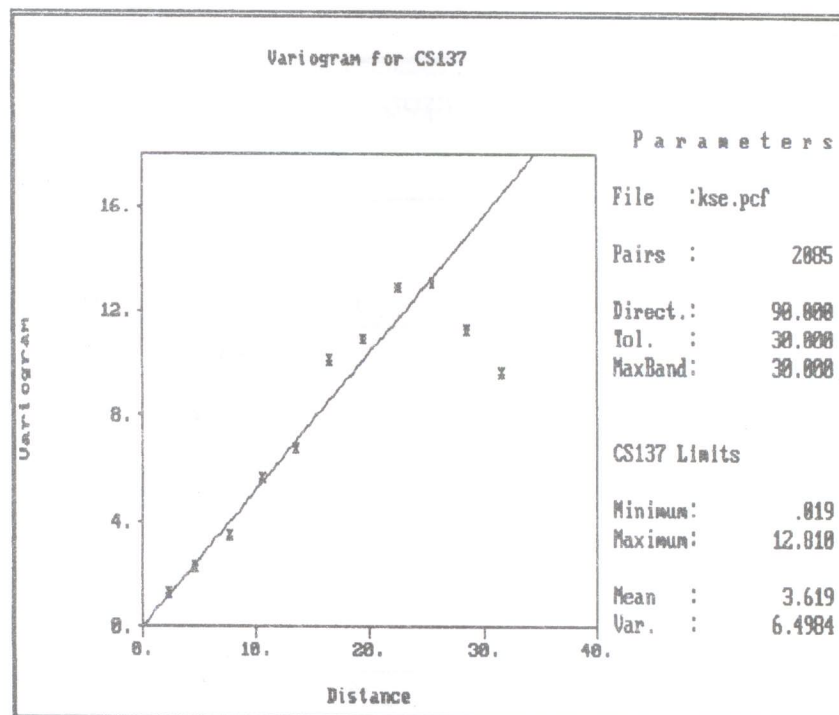


Рис. 30: Модель вариограммы в направлении 90° по юго-восточной области.

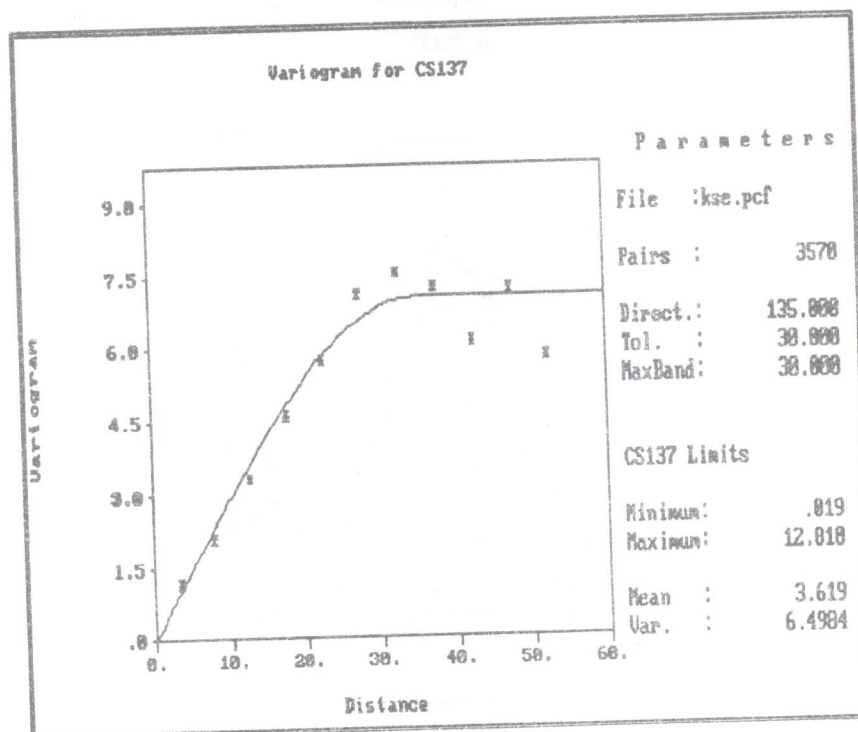


Рис. 31: Модель вариограммы в направлении 135° по юго-восточной области.

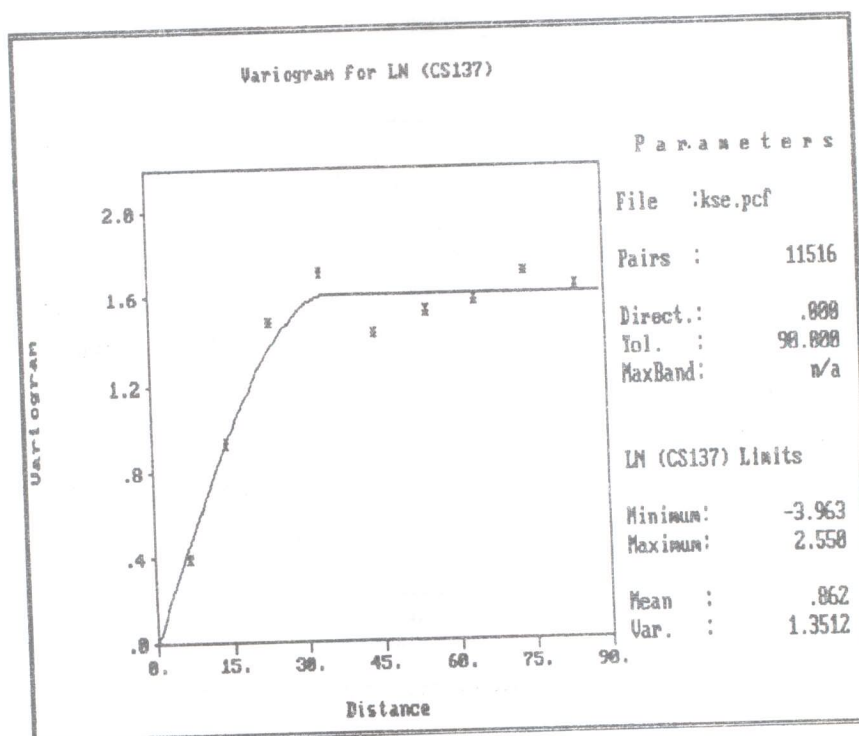


Рис. 32: Модель вариограммы логарифмов по всем направлениям по юго-восточной области.

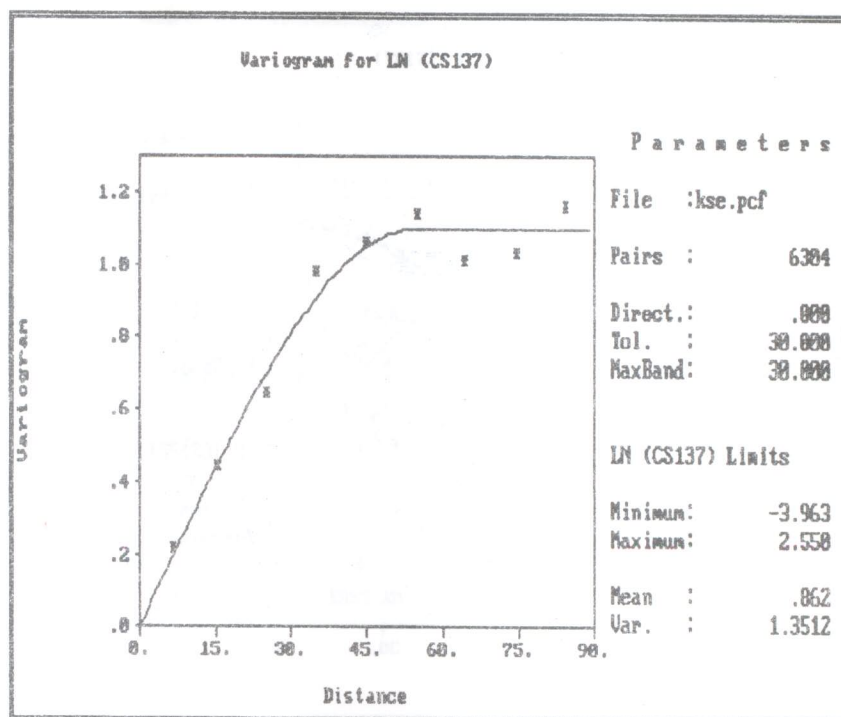


Рис. 33: Модель вариограммы логарифмов в направлении 0° по юго-восточной области.

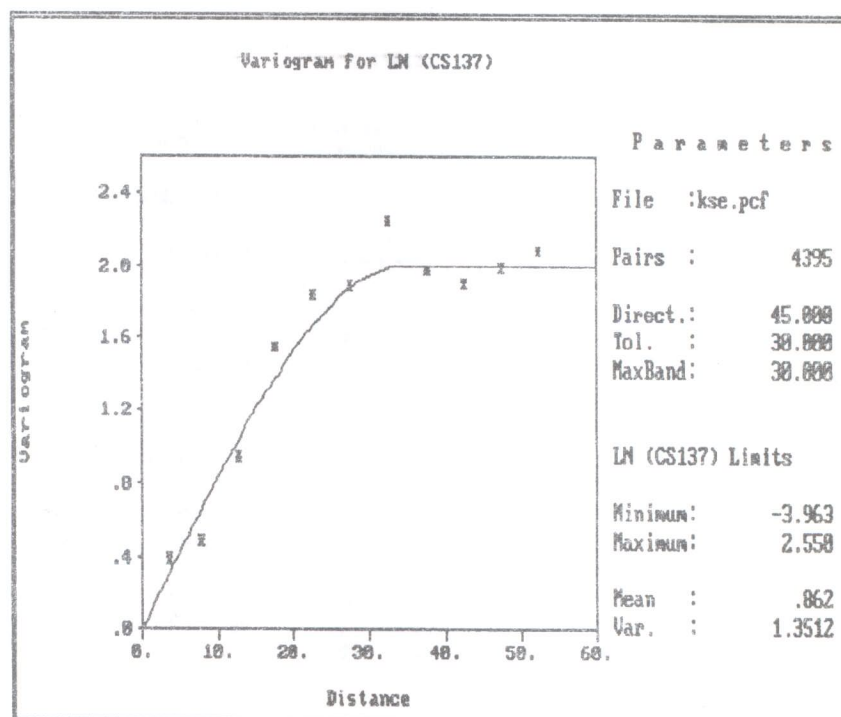


Рис. 34: Модель вариограммы логарифмов в направлении 45° по юго-восточной области.

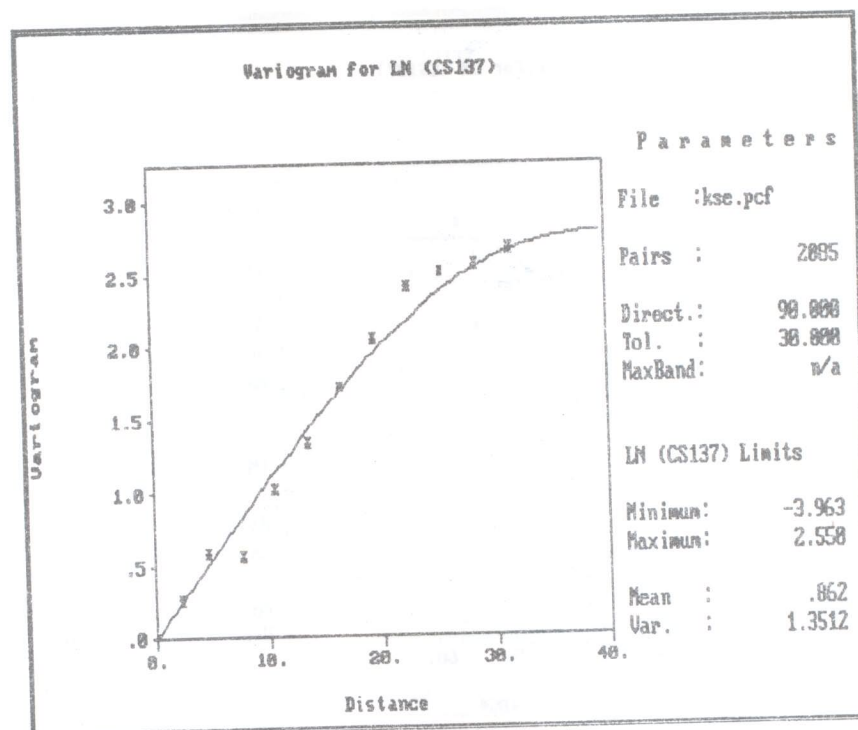


Рис. 35: Модель вариограммы логарифмов в направлении 90° по юго-восточной области.

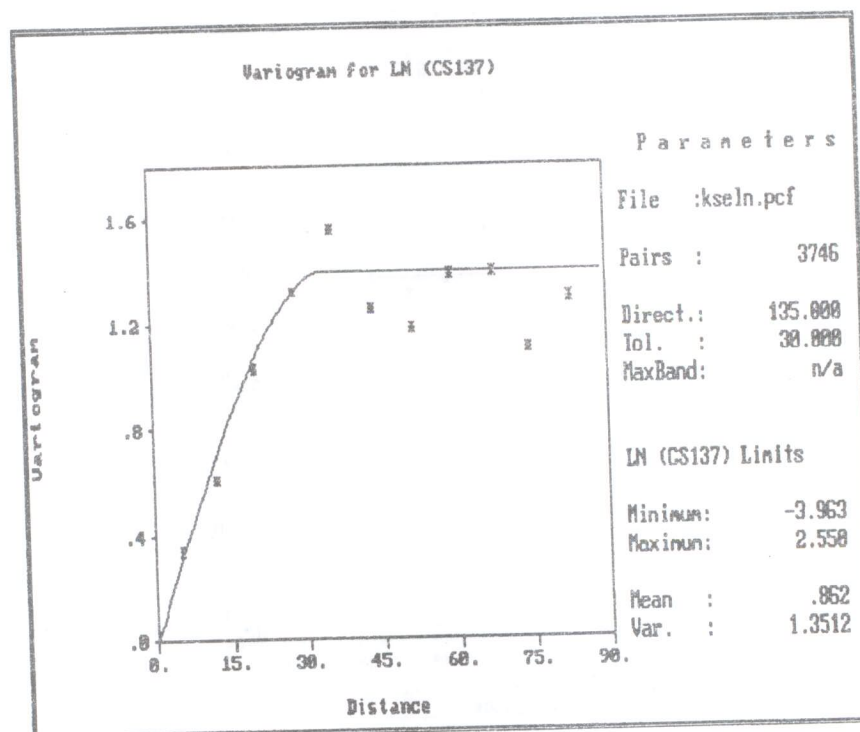


Рис. 36: Модель вариограммы логарифмов в направлении 135° по юго-восточной области.

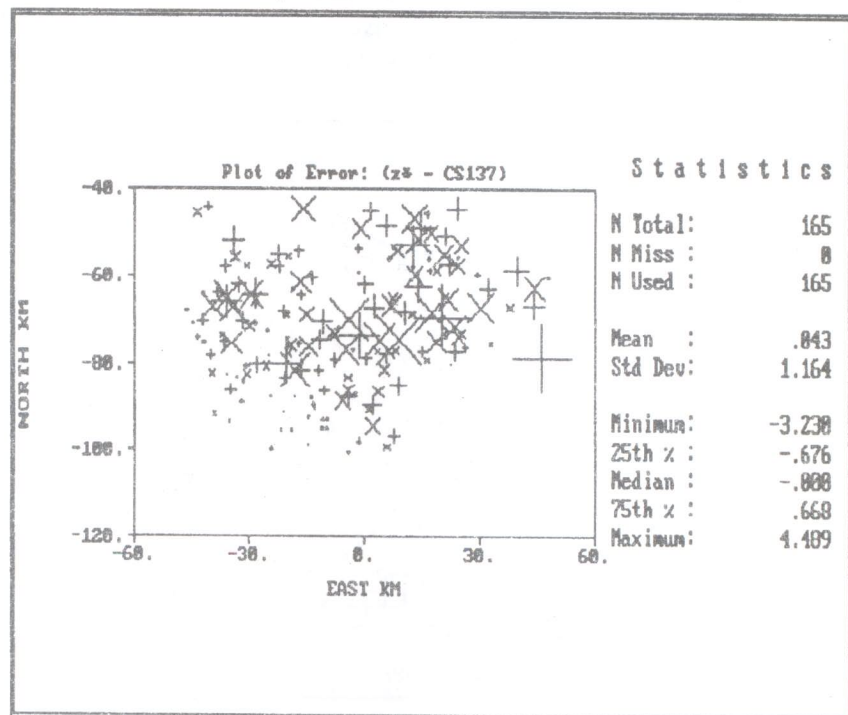


Рис. 37: Карта ошибок кросс-валидации.

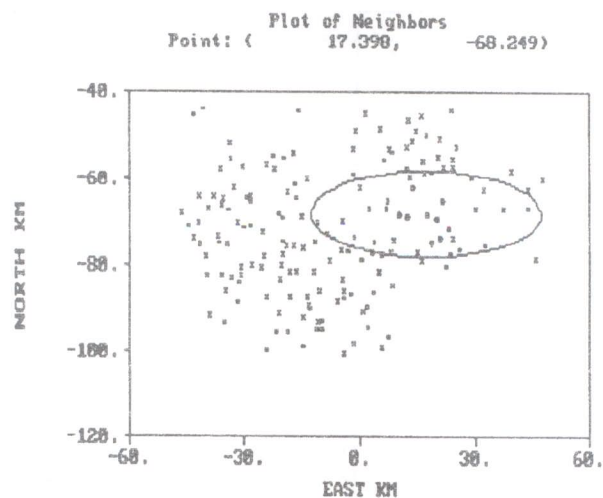


Рис. 38: Окрестность поиска для кригинга.

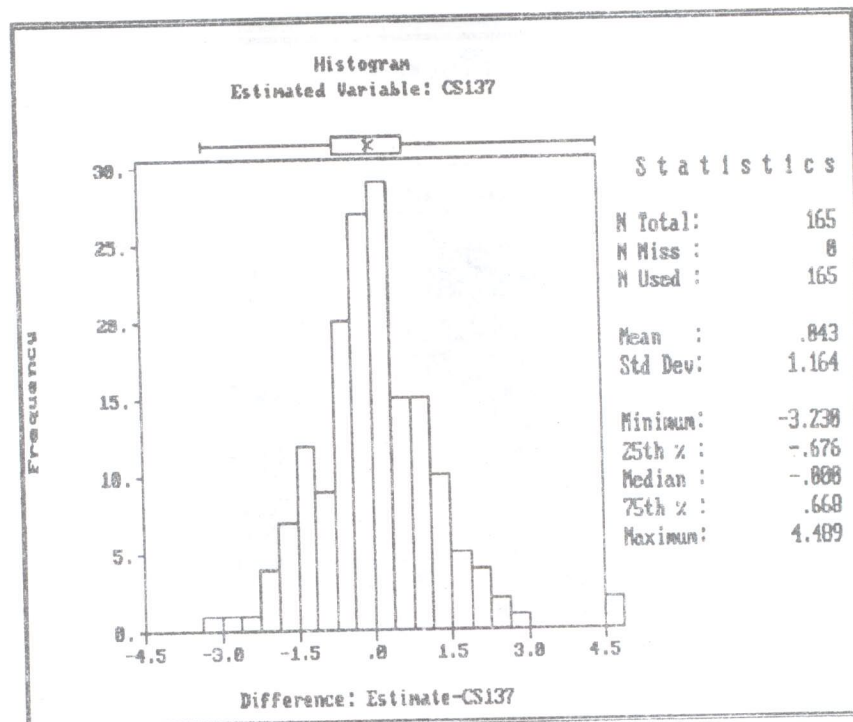


Рис. 39: Гистограмма распределения разницы между оценкой и наблюдением.

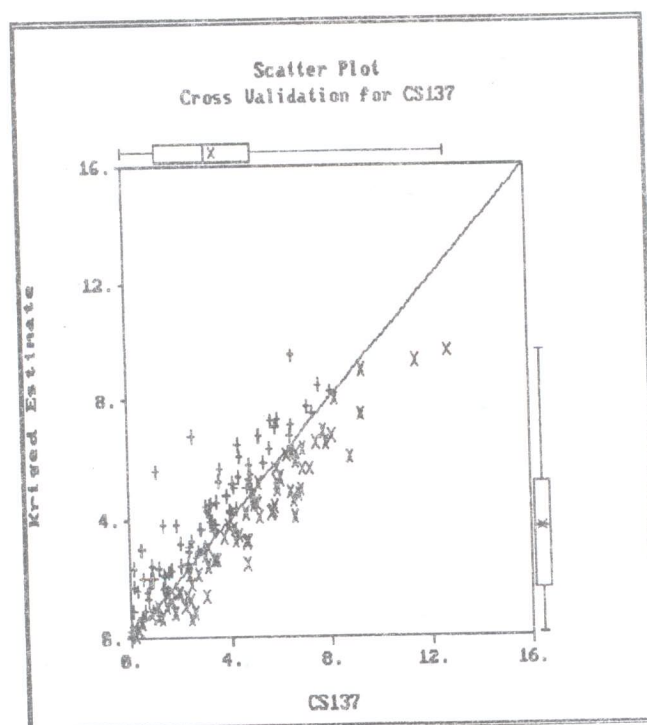


Рис. 40: Диаграмма разброса оценки в зависимости от наблюдаемых значений.

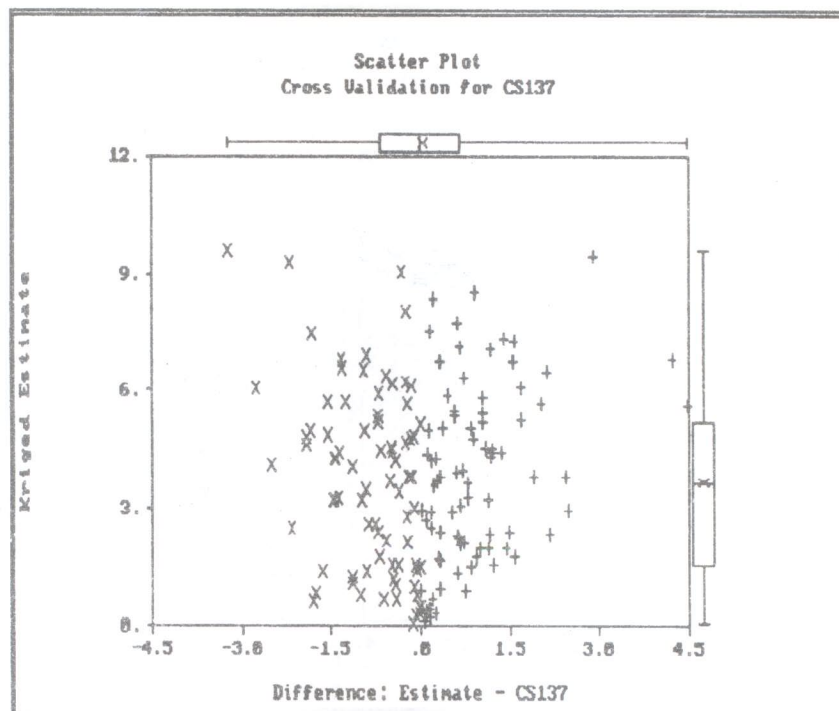


Рис. 41: Диаграмма разброса разницы между оценкой и наблюдением в зависимости от наблюдаемых значений.

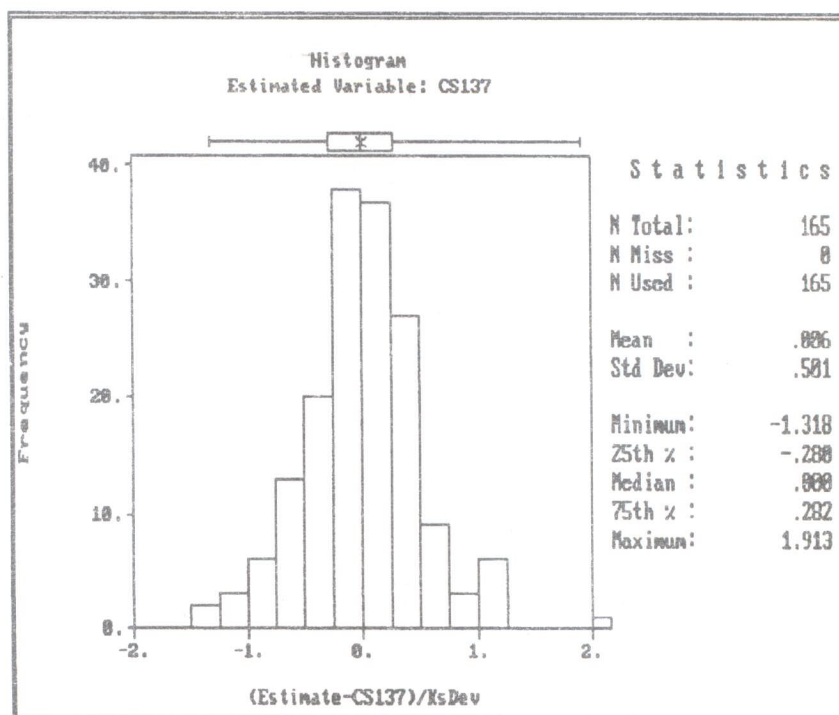


Рис. 42: Гистограмма значений Z_{score} .

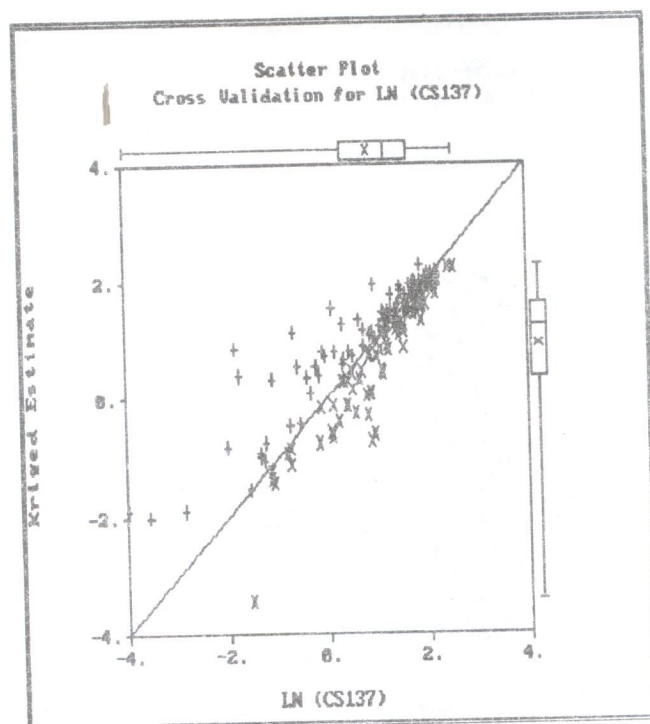


Рис. 43: Диаграмма разброса оценки кригинга зависимости от наблюдаемых значений для логарифмов.

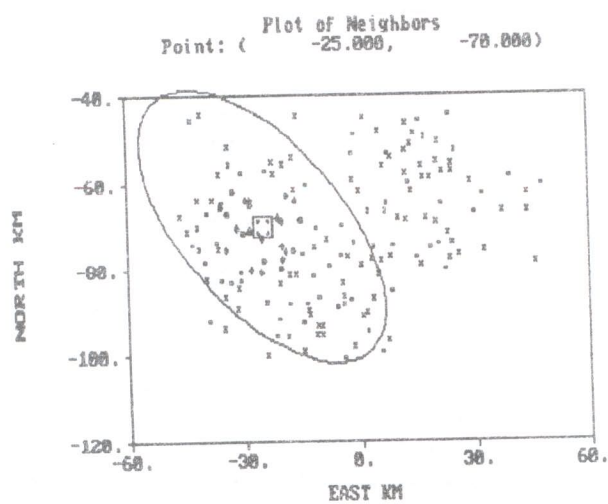
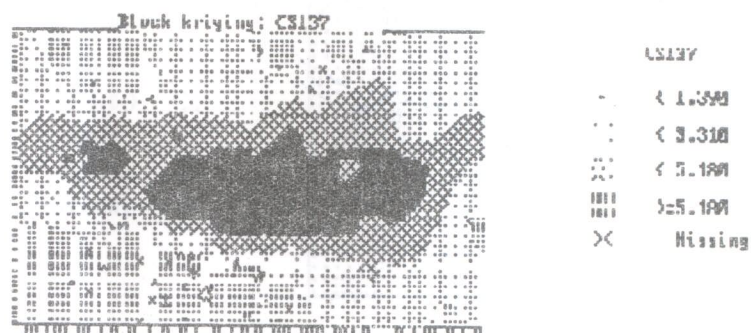


Рис. 44: Окрестность поиска при блочном 2×2 кригинге.



U	V	Y	Z*	xk	N
6888	49.8	-41.8	2.15	6.85	16

Рис. 45: Мозаичная карта кригинга на сетке 1×1 км.

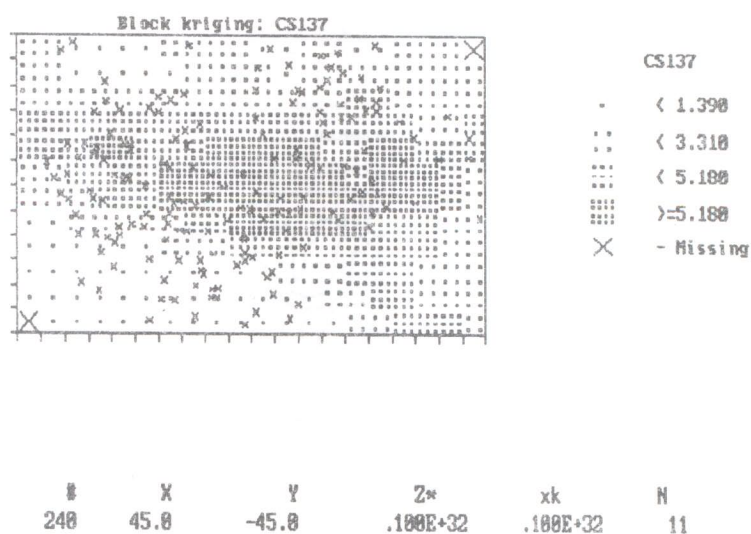


Рис. 46: Мозаичная карта кригинга с гауссовой составляющей в направлении 45° на сетке 5×5 км.

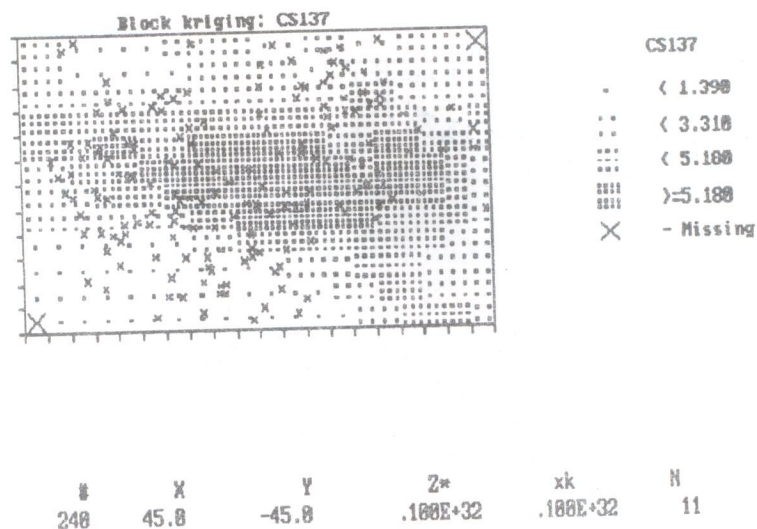


Рис. 47: Мозаичная карта кригинга по модели с линейной составляющей в направлении 45° на сетке 5×5 км.

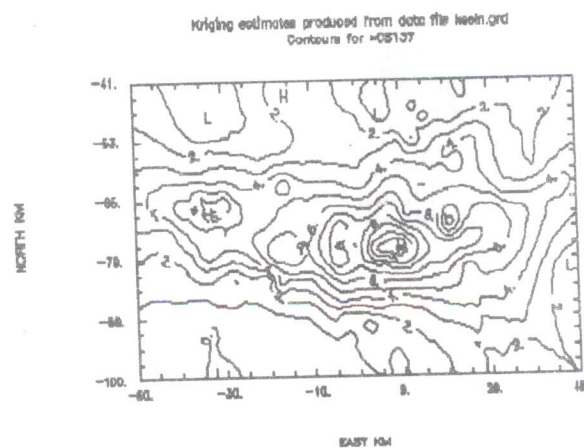


Рис. 48: Карта изолиний Cs^{137} после кригинга на сетке с ячейкой 1×1 км.

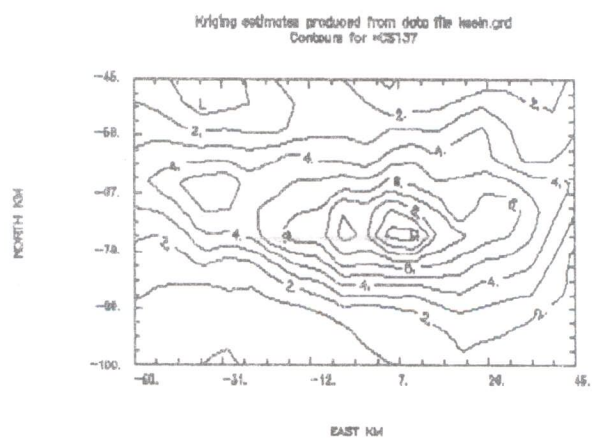


Рис. 49: Карта изолиний ^{137}Cs после кригинга на сетке с ячейкой 5×5 км.