

ГРАФИКИ КВАНТИЛЬ-КВАНТИЛЬ (Q – Q) ДЛЯ ГИДРОЛОГО-ГИДРОЛОГИЧЕСКИХ ДАННЫХ МОЛДОВЫ

Е.Н. МАТВЕЕВА, научный сотрудник, matveeva61@mail.ru

Л.Л. ВЭТЭМАНУ, младший научный сотрудник, liubitzaturcan@gmail.com

К.Е. МОРАРУ, старший научный сотрудник, д.г.-м.н. наб., cmoraru@yahoo.com

Институт Геологии и Сейсмологии

***Abstract.** Verification of the normal statistical law on the basis of Q-Q plot methodology for the first time has been done using a large number of hydrological and hydrogeological data (discharge and water table). This methodology includes analysis of Q-Q plots, its digitalizing, calculation of the correlation index and coefficient, and finally designing about data normality. Obtained results permits to estimate statistical reliability of the statistical mean and probability of cartographic models.*

Keywords: *Q – Q Plot, normal law, hydrological and hydrogeological data.*

ВВЕДЕНИЕ

Предположение о принадлежности случайной величины нормальному закону распределения лежит в основе многих статистических методов и критериев. В ряде случаев соблюдение данного требования является критичным для применения того или иного метода.

Статистическая проверка принадлежности гидрологических и гидрогеологических данных к нормальному закону распределения является важной задачей геостатистики. Впервые для многочисленных значений расхода рек и уровней подземных вод выполнена проверка закона распределения с использованием технологии Q – Q (квантиль-квантиль) графиков. Графики квантиль-квантиль сравнивают совокупности данных с согласованным нормальным распределением.

МАТЕРИАЛЫ И МЕТОДОЛОГИЯ ИССЛЕДОВАНИЙ

На графике квантиль-квантиль (или Q – Q график), также называемом графиком квантилей можно зрительно оценить подгонку теоретического распределения к наблюдаемым данным, т.е. это позволяет нам понять с одного взгляда, какое может быть максимальное отклонение без угрозы невозврата к историческому значению.

Q - Q график (Q - квантиль) - это график, на котором квантили из двух распределений расположены относительно друг друга. Чем ближе точки на графике к диагональной прямой, тем ближе распределение исследуемой переменной к нормальному закону. Квантиль-квантиль график, представляет собой инструмент, который помогает нам оценить правдоподобность отклонения среднего от теоретического распределения.

На квантиль-квантиль графике показана связь между наблюдаемыми значениями переменных и теоретическими квантилями. Если наблюдаемые значения попадают на прямую линию, то теоретическое распределение хорошо подходит к наблюдаемым данным.

Для построения Q - Q графика программа сначала упорядочивает по возрастанию n точек наблюдаемых данных:

$$X_1 X_2 \dots X_n$$

Эти наблюдаемые значения откладываются по одной из осей графика; по другой оси откладываются следующие значения:

$$F^{-1}((i-r_{adj})/(n+n_{adj})),$$

где i - есть ранг соответствующего наблюдения, r_{adj} и n_{adj} - корректировки, а F^{-1} есть обратный вероятностный интеграл для соответствующего стандартизованного распределения.

Получившийся график представляет собой диаграмму рассеяния наблюдаемых и ожидаемых (стандартизованных) значений при соответствующем заданном распределении.

Необходимые подготовительные расчеты и графические построения выполнялись в срезе пакета программы SPSS-22 (Statistical Package for the Social Sciences) [3], в разделе разведочная статистика. Построение и дальнейшая работа с квантиль-квантиль графиками (Q – Q plot-ами) проводилась в программе Get Data 2.26 [4]. Пример представлен на рис. 1.

Построение и работа с Q – Q графиками проводится в следующей последовательности:

1. установление системы координат;
2. выделение X_{\min} и X_{\max} ;
3. выделение Y_{\min} и Y_{\max} ;
4. проверка параметров системы координат (справа на рис. 1);
5. работа в режиме установки точек исходных данных;
6. дигитизирование (оцифровывание или преобразование в цифровую форму) данных (см. правую панель X, Y);
7. экспортирование данных в формат Excel.

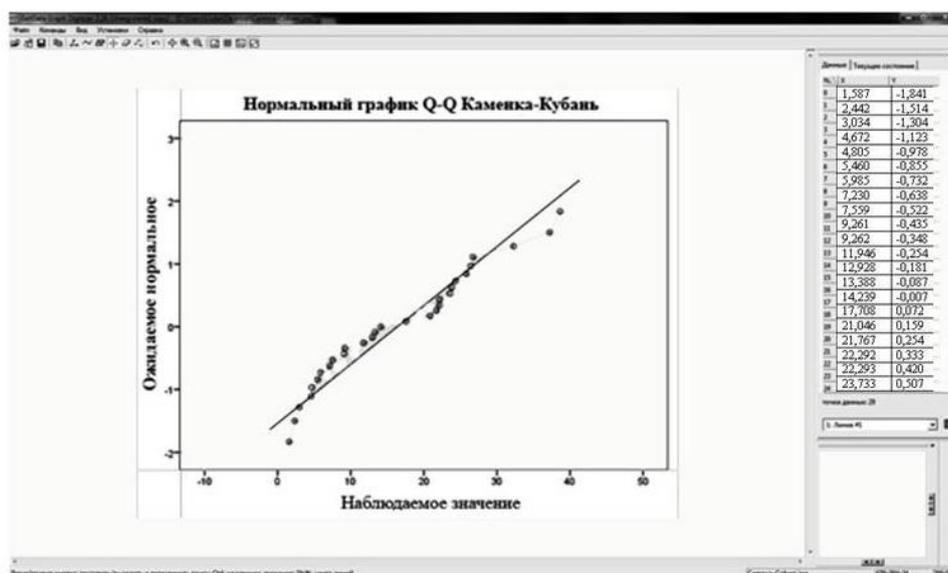


Рис. 1. Построение Q – Q графика в программе Get Data 2.26

Исходными материалами для проделанной работы послужили многолетние гидрологические и гидрогеологические данные [1, 2] и архивные материалы лаборатории Гидрогеологии Института Геологии и сейсмологии.

Предложенная методология опробована авторами для следующих статистических методов, используемых для подсчетов ресурсов подземных вод для территории Республики Молдова: Воскресенского; Баланса; Потери стока; Межени; Куделина; PART и BFI.

РЕЗУЛЬТАТЫ И ОБСУЖДЕНИЕ

Авторами предложена методология количественного сравнения Q – Q графиков, которая включает дигитизирование графиков, расчет индекса и коэффициента корреляции и определение интервала нормальности.

Программе Get Data 2.26 позволяет работать как с нормальными Q – Q графиками (см. рис. 2), так и с нормальными Q – Q графиками с удаленным трендом (см. рис. 3).

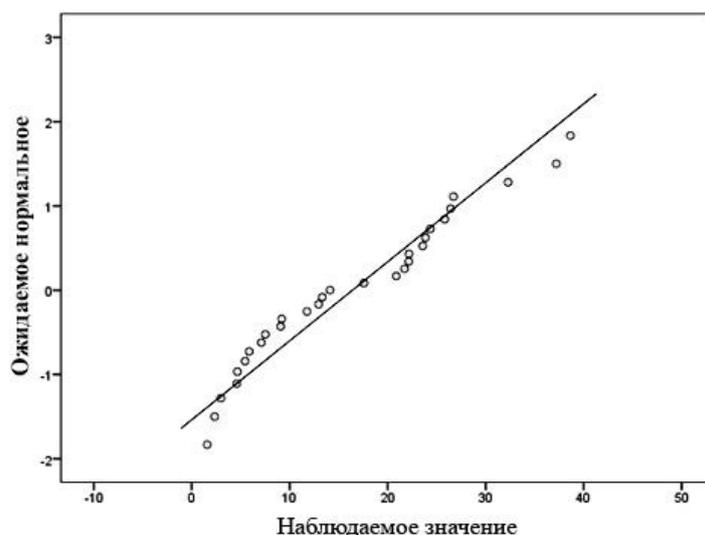


Рис. 2. Нормальный график Q – Q (Каменка - Кубань)

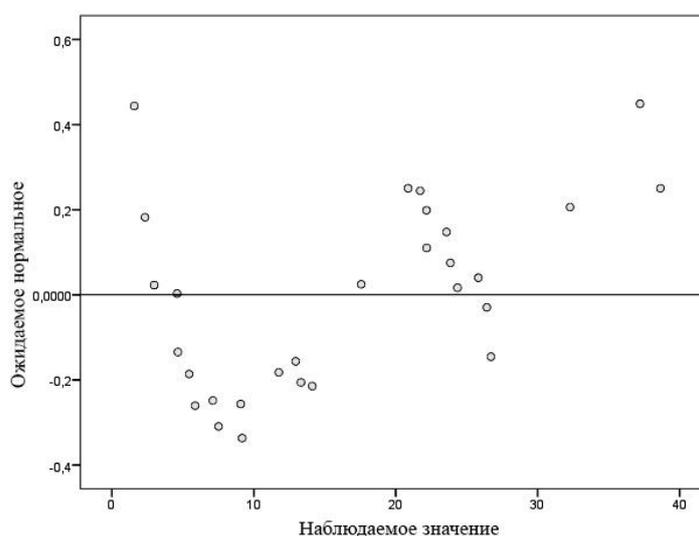


Рис. 3. Нормальный график Q – Q с удаленным трендом (Каменка - Кубань)

Метод квантиль-квантиль графиков опробован на данных по 43-м малым рекам Молдовы, а также на семи и трех гидрологических створах на р. Днестр и р. Прут соответственно [1, 2]. Для всех перечисленных пунктов метод применялся по всем семи указанным выше статистическим методам.

На рис. 4, как пример, приведен дигитизированный квантиль-квантиль график по р. Каменка (гидрологический пост Кубань).

После экспортирования дигитизированных данных Q – Q графиков в формат Excel по всем указанным методам для всех пунктов рассчитывался индекс и коэффициента корреляции среднего значения, а также определялся интервал нормальности.

Проведенная работа позволила определить статистическую надежность среднего значения исходных данных за многолетний период наблюдений.

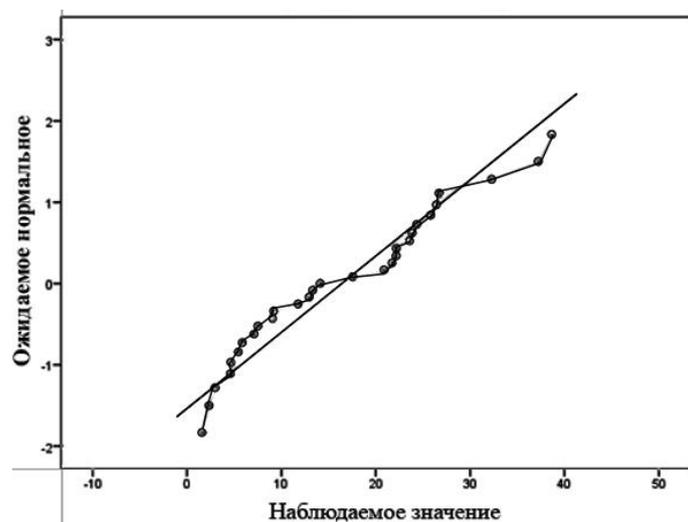


Рис. 4. Дигитизированный нормальный график Q – Q (Каменка - Кубань)

ВЫВОДЫ

Полученные результаты использования метода Q – Q графиков показали довольно высокую надежность исходного материала, что позволяет использовать его в дальнейшем для построения вероятностных картографических моделей гидрологического и гидрогеологического профиля.

СПИСОК ИСПОЛЬЗУЕМЫХ ИСТОЧНИКОВ

1. *Гидрологический Ежегодник 1958 - 1978 г. Том 2 Бассейн Черного моря (без Кавказа) выпуск 0, 1 Бассейн Черного моря без бассейнов р. Днепр и р. Дон.* - Ленинград: Гидрометеиздат, 1958 - 1978.
2. *Государственный водный кадастр: Ежегодные данные о режиме и ресурсах поверхностных вод суши 1979 – 1986 г. Части 1 и 2. Том IX.* - Кишинев, 1979 - 1990.
3. *Statistical Package for the Social Sciences (SPSS 22).* www.ibm.com/us-en/marketplace/spss-statistics (accesat 22.11.2017).
4. *Программа Get Data 2.26.* <http://www.getdata-graph-digitizer.com/download.php>