

Анализ и прогнозирование гидрологических данных

Павлов Александр Сергеевич

Научный руководитель: Цеховая Татьяна Вячеславовна

Рецензент: Рафеенко Екатерина Дмитриевна

БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

КАФЕДРА ТЕОРИИ ВЕРОЯТНОСТЕЙ И МАТЕМАТИЧЕСКОЙ СТАТИСТИКИ

Минск

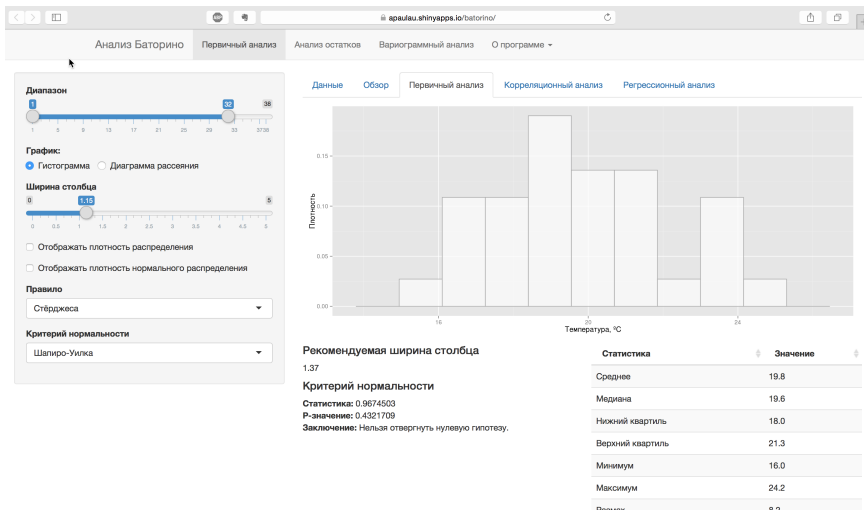
16/06/2015

Постановка задачи

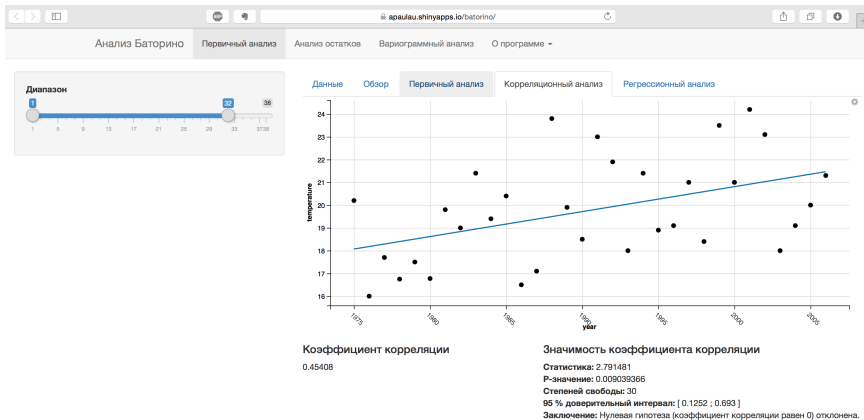
- ▶ Предварительный статистический анализ данных
- ▶ Исследование статистических свойств оценки вариограммы
- ▶ Вариограммный анализ
- ▶ Прогнозирование значений временного ряда методом Кригинг



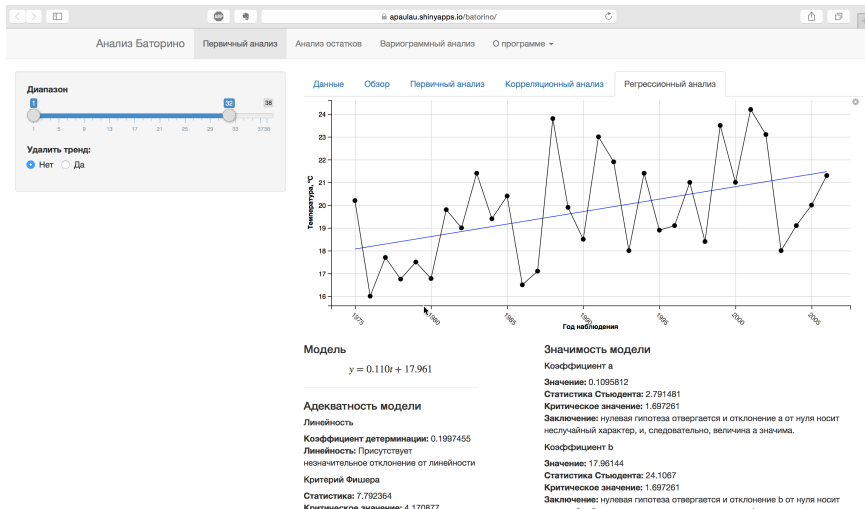
Первичный анализ и описательные статистики



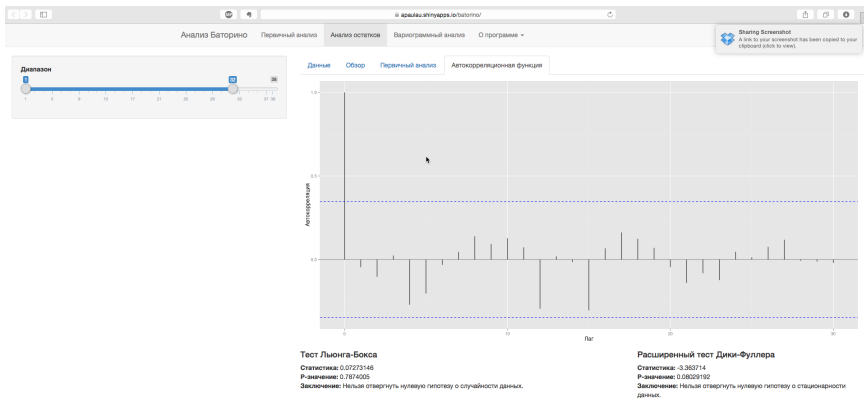
Корреляционный анализ



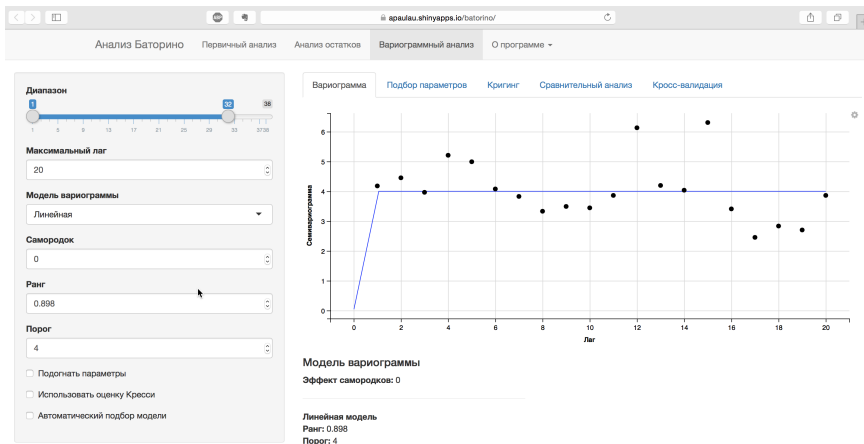
Регрессионный анализ



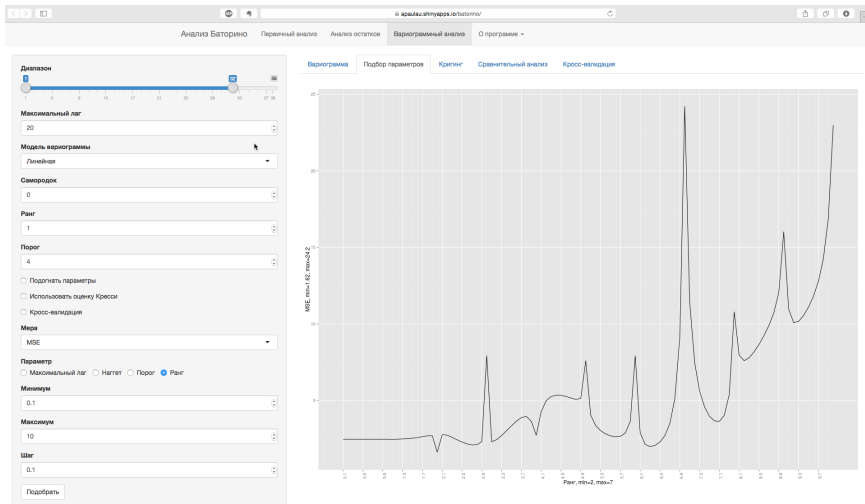
Автокорреляционная функция



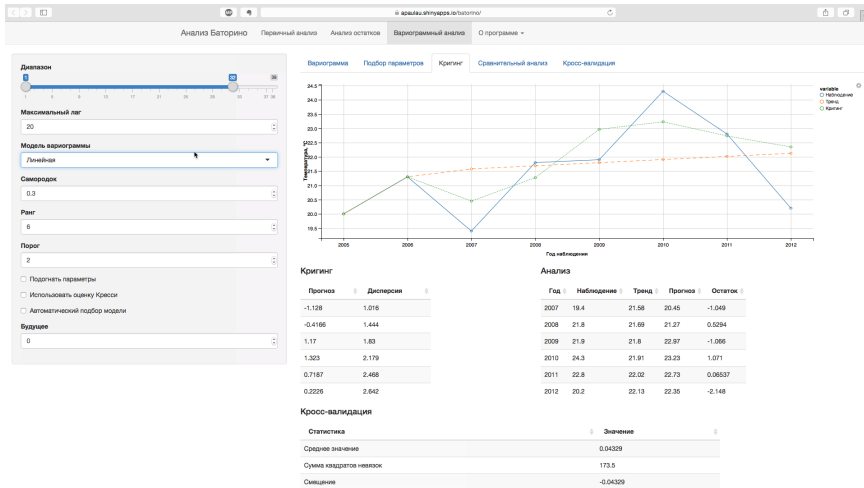
Возможности по подбору модели вариограммы



Подбор параметров модели вариограммы



Сравнение прогнозных значений



Исходные данные

Исходные данные получены от учебно-научного центра «Нарочанская биологическая станция им. Г.Г.Винберга». На рисунке представлена выборка, состоящая из наблюдений за температурой воды в июле месяце в период с 1975 по 2012 годы.



График квантилей

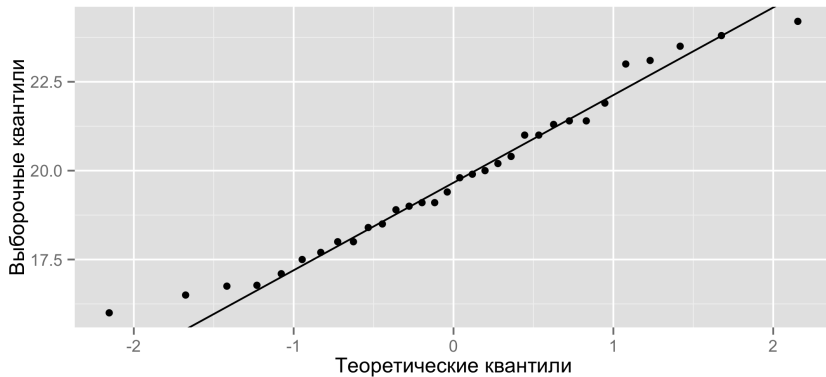
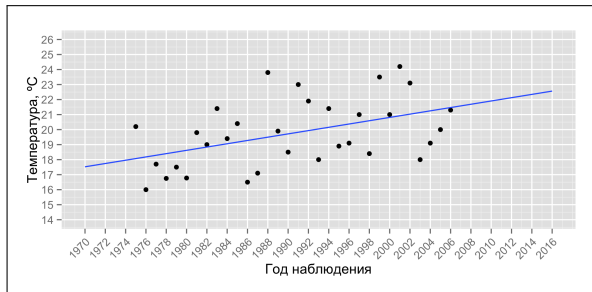


Диаграмма рассеяния

Выборочный коэффициент
корреляции: $r_{xt} = 0.454$.
При уровне значимости
 $\alpha = 0.05$ коэффициент является
значимым



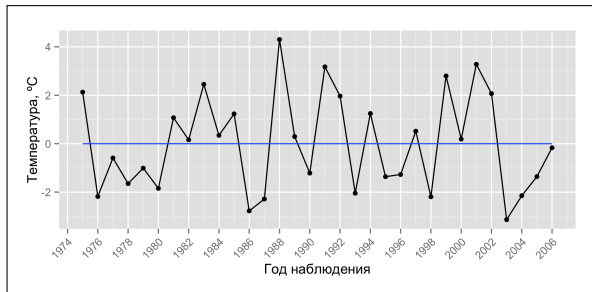
Временной ряд

Вид исходного временного ряда:

$$x(t) = y(t) + \varepsilon(t)$$

Уравнение тренда:

$$y(t) = at + b = 0.1014t + 18.0521$$



Оценка модели

- ▶ С помощью критерия Стьюдента доказана значимость коэффициентов регрессионной модели
- ▶ F-критерий Фишера при уровне значимости $\alpha = 0.05$ показал адекватность модели
- ▶ Точность модели невысока, поскольку коэффициент детерминации $\eta_{x(t)}^2 = 0.275$



Оценка вариограммы

Определение 1

Вариограммой случайного процесса $X(t)$, $t \in \mathbb{Z}$, называется функция вида

$$2\gamma(h) = V\{X(t+h) - X(t)\}, \quad t, h \in \mathbb{Z}. \quad (1)$$

При этом функция $\gamma(h)$, $h \in \mathbb{Z}$, называется семивариограммой.

В качестве оценки вариограммы рассматривается статистика, предложенная Матероном:

$$2\tilde{\gamma}(h) = \frac{1}{n-h} \sum_{t=1}^{n-h} (X(t+h) - X(t))^2, \quad h = \overline{0, n-1}, \quad (2)$$



Первые два момента оценки вариограммы

Теорема 1

Для оценки $2\tilde{\gamma}(h)$ имеют место следующие соотношения:

$$E\{2\tilde{\gamma}(h)\} = 2\gamma(h),$$

$$\text{cov}(2\tilde{\gamma}(h_1), 2\tilde{\gamma}(h_2)) =$$

$$= \frac{2}{(n-h_1)(n-h_2)} \sum_{t=1}^{n-h_1} \sum_{s=1}^{n-h_2} (\gamma(t-h_2-s) + \gamma(t+h_1-s) - \gamma(t-s) - \gamma(t+h_1-s-h_2))^2,$$

$$V\{2\tilde{\gamma}(h)\} = \frac{2}{(n-h)^2} \sum_{t,s=1}^{n-h} (\gamma(t-h-s) + \gamma(t+h-s) - 2\gamma(t-s))^2,$$

где $\gamma(h)$, $h \in \mathbb{Z}$, — семивариограмма процесса $X(t)$, $t \in \mathbb{Z}$, $h, h_1, h_2 = \overline{0, n-1}$.



Асимптотическое поведение оценки вариограммы

Теорема 2

Если имеет место соотношение

$$\sum_{h=-\infty}^{+\infty} |\gamma(h)| < +\infty, \text{ то}$$

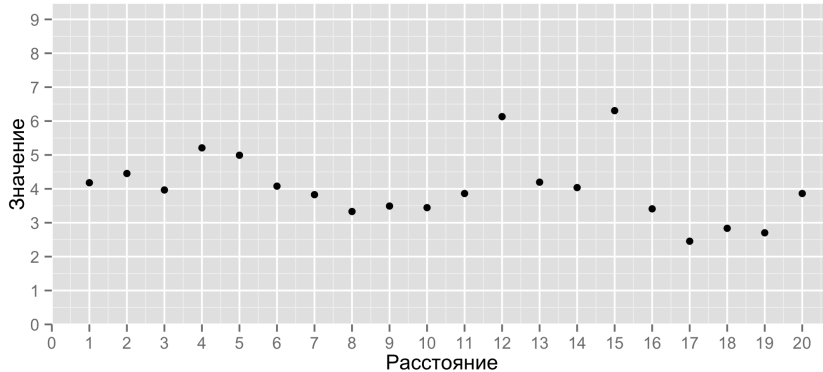
$$\lim_{n \rightarrow \infty} (n - \min\{h_1, h_2\}) \text{cov}\{2\tilde{\gamma}(h_1), 2\tilde{\gamma}(h_2)\} = 2 \sum_{m=-\infty}^{+\infty} \gamma(m - h_2) + \gamma(m + h_1) - \gamma(m) - \gamma(m + h_1 - h_2))^2,$$

$$\lim_{n \rightarrow \infty} (n - h) V\{2\tilde{\gamma}(h)\} = 2 \sum_{m=-\infty}^{+\infty} \gamma(m - h) + \gamma(m + h) - 2\gamma(m))^2.$$

где $\gamma(h)$, $h \in \mathbb{Z}$, — семивариограмма процесса $X(t)$, $t \in \mathbb{Z}$, $h, h_1, h_2 = \overline{0, n-1}$.

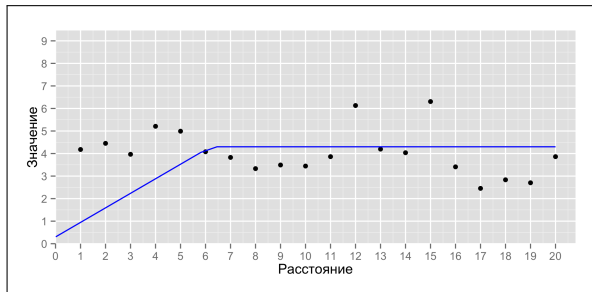


График экспериментальной вариограммы

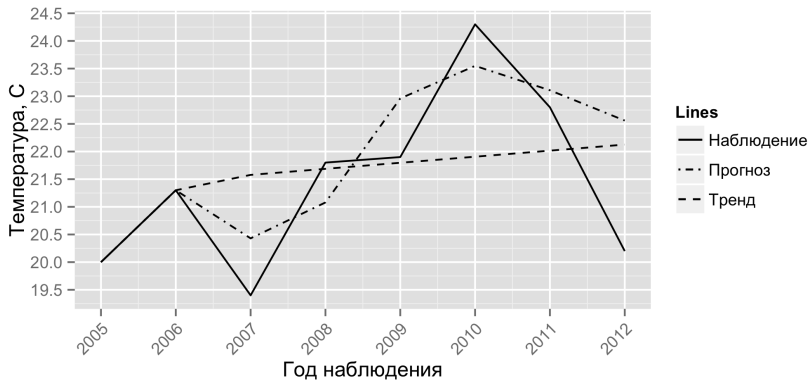


Линейная модель с порогом

Подобранная модель:
 $0.3 + 4 \cdot \text{Lin}(h, 6.2)$



Прогнозирование методом ординарного кригинга

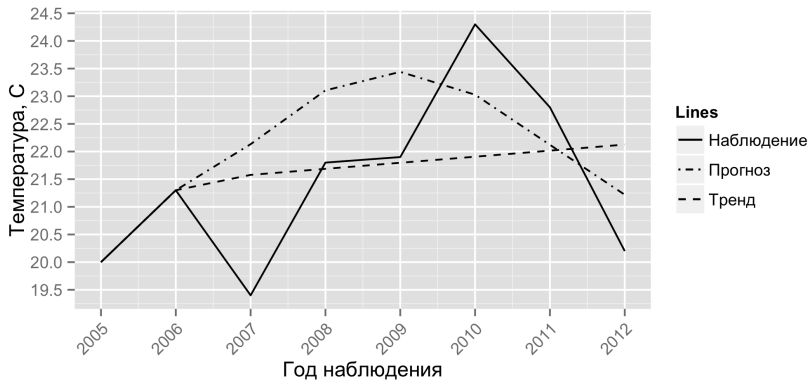


Периодическая модель

Подобранная модель:
 $4 \cdot \text{Per}(h, 0.898)$

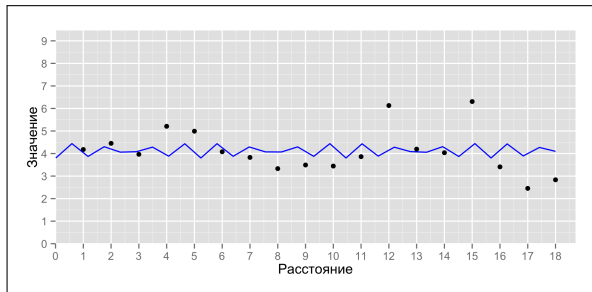


Прогнозирование методом ординарного кригинга

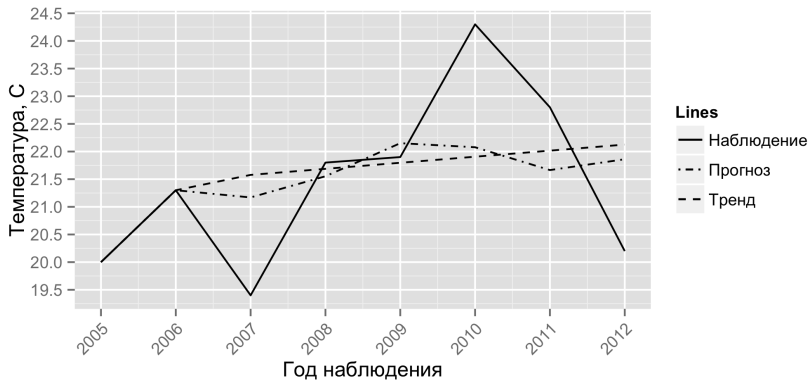


Периодическая модель

Подобранная модель:
 $3.8 + 0.32 \cdot \text{Per}(h, 1.3)$

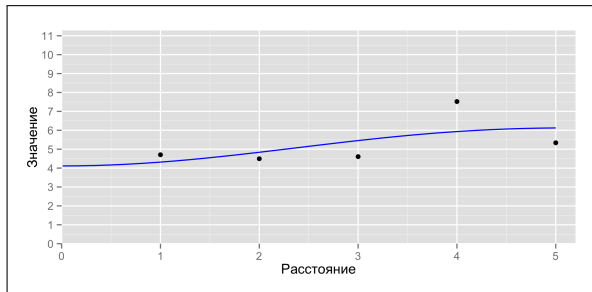


Прогнозирование методом ординарного кригинга



Волновая модель

Подобранная модель:
 $4.11 + 1.65 \cdot Wav(h, 3.59)$



Прогнозирование методом ординарного кригинга

