

## Об аппроксимации влажной тропосферной задержки в зените случайным блужданием

© А. А. Куделькин

ИПА РАН, г. Санкт-Петербург, Россия

### Реферат

Одна из важнейших проблем обработки данных космической геодезии — учет некоторых быстро меняющихся параметров, в том числе флуктуаций влажной части тропосферной задержки в зените (ZWD, Zenith Wet Delay). Учет данного параметра особенно важен при анализе данных РСДБ и ГНСС. Стандартным методом обработки данных параметров является аппроксимация некоторым стохастическим процессом, чаще всего случайным блужданием.

Цель данной работы — вычисление параметра случайного блуждания для моделей с трендом и без тренда по результатам наблюдений радиометра водяного пара (РВП) за 2019 г. для станций «Светлое», «Зеленчукская» и «Бадарь», а также анализ качества модели. Для оценки параметра случайного блуждания применялись методы Калмана и коллокации — для моделей без тренда и с трендом соответственно. Для ускорения вычисления оценок в случае модели с трендом был предложен алгоритм быстрого решения линейной системы с матрицей специального вида, представляющий собой модификацию известного метода прогонки. Наконец, для проверки качества полученной оценки использовалось сравнение спрогнозированного значения ZWD со значением, полученным РВП.

В статье вычислена оценка параметра нормального случайного блуждания стохастической части ZWD для моделей случайного блуждания с трендом и без тренда для станций «Светлое», «Зеленчукская» и «Бадарь» по результатам РВП-наблюдений. Методом Шапиро – Уилка, а также визуально проверено соответствие данных и модели. Показано, что нормальное случайное блуждание является достаточно грубым приближением и достоверно описывает данные лишь в тех случаях, когда значения ZWD малы. Кроме того, установлено, что модель блуждания с линейным трендом на коротких интервалах дает более точное приближение, чем модель без тренда.

**Ключевые слова:** влажная тропосферная задержка в зените, радиометр водяного пара, стохастическая аппроксимация, радиоинтерферометрия со сверхдлинной базой.

*Контакты для связи:* Куделькин Алексей Александрович (aa.kudelkin@iaaaras.ru).

*Статья поступила в редакцию 02.09.2020, принята к публикации 09.10.2020, опубликована 30.10.2020.*

**Для цитирования:** Куделькин А. А. Об аппроксимации влажной тропосферной задержки в зените случайным блужданием // Труды ИПА РАН. 2020. Вып. 54. С. 44–51.

<https://doi.org/10.32876/AplAstron.54.44-51>

## Zenith Wet Delay Approximation by Random Walk

A. A. Kudelkin

Institute of Applied Astronomy of the Russian Academy of Sciences, Saint Petersburg, Russia

### Abstract

One of the important problem of data processing of space geodesy is the accounting of some rapidly meant parameters, including fluctuations of zenith wet delay (ZWD). Accounting for the given parameter is especially important in analyzing VLBI and GNSS data. The standard method of processing these parameters is the approximation by some stochastic process, most often by a random walk.

The purpose of this work is to calculate the random walk parameter for models with a trend and without a trend based on the results of water vapor radiometer (WVR) observations in 2019 for the stations “Svetloe”, “Zelenchukskaya” and “Badary”, as well as an analysis of the model quality. To estimate the random walk parameter, the methods of Kalman and collocation were used for the models without a trend and with a trend, respectively. To speed up the calculation of estimates in the case of the model with a trend, an algorithm for the rapid solution of a linear system with matrix of a special appearance. This algorithm is a modification of the well-known tridiagonal matrix algorithm. Finally, a comparison of the predicted ZWD value with the value obtained by the WVR was used to check the quality of the assessment.

The paper presents the calculation of the maximum likelihood estimate for the parameter of the normal random walk of the ZWD stochastic part for random walk models with and without a trend. The obtained algorithms were applied to WVR observations at the “Svetloe”, “Zelenchukskaya”, and “Badary” stations. The compliance of the data and the model was verified using Shapiro–Wilk test, and visually. It is shown that the normal random walk is a rather rough approximation and describes adequate the data only in cases where the ZWD value is small. In addition, it was established that random walk with a linear trend in short intervals gives a more accurate approximation than the model without a trend.

**Keywords:** tropospheric zenith wet delay, water vapor radiometer, very long baseline interferometry, stochastic approximation.

*Contacts:* Aleksey A. Kudelkin (aa.kudelkin@iaaras.ru).

*Received 2 September, 2020, accepted 9 October, 2020, published 30 October, 2020.*

For citation: Kudelkin A. A. Zenith wet delay approximation by random walk // Transactions of IAA RAS. 2020. Iss. 54. P. 44–51.

<https://doi.org/10.32876/AplAstron.54.44-51>

## Введение

При решении задач обработки данных космической геодезии методами сглаживания и фильтрации Калмана, а также методом коллокации возникает необходимость рассматривать некоторые быстро меняющиеся параметры как стохастические процессы с известными ковариационными функциями  $k(t, s)$ , где  $t, s$  — временные переменные. Особенную важность представляет учет стохастических параметров атмосферы, поскольку точность вычисления тропосферной задержки сигнала оказывает существенное влияние на качество определения ПВЗ.

Не существует однозначно верного стохастического описания ZWD; в различных статьях и программных реализациях предлагаются разные модели. Так, Т. А. Херринг [1] предлагает использовать случайное блуждание

$$k(t, s) = \Phi \min(t - t_0, s - t_0)$$

с коэффициентом случайного блуждания  $\Phi = 5800$  мм<sup>2</sup>/день и  $t_0 = 0$  ( $t, t_0, s$  — модифицированные Юлианские даты); данная оценка используется при анализе РСДБ-наблюдений в программе «VieVS@GFZ» — локальной версии известного пакета «VieVS», применяющегося в German Research Centre for Geosciences [2]. В программе обработки РСДБ-наблюдений «ОССАМ» также используется случайное блуждание, но с коэффициентом  $\Phi = 1944$  мм<sup>2</sup>/день и с начальной ковариацией 13935 мм<sup>2</sup> (что соответствует случаю, когда случайное блуждание начинается за  $t_0 = 7.168$  сут до первого наблюдения) [3]. В. С. Губанов выводит модель ZWD, исследуя внутрисуточные флуктуации эйконала тропосферы, и определяет ZWD как стационарный случайный процесс с линейным трендом и ковариационной функцией [4, стр. 268–277]:

$$k(t, s) = \frac{\sigma^2}{\cos(\varphi)} e^{-\alpha|t-s|} \cos(\beta|t-s| + \varphi),$$

где  $\alpha, \beta, \varphi, \sigma^2$  — параметры, зависящие от станции. Данное описание используется в программном пакете «Quasar» [5].

Кроме определения ZWD методами космической геодезии, параметр влажной зенитной задержки также можно вычислить с помощью радиометра водяного пара (РВП) [6]. Данные наблюдения имеют важные особенности, существенно упрощающие их анализ. В отличие от техник космической геодезии, определяющих десятки различных параметров, наблюдения с помощью РВП

определяют только параметр зенитной задержки. Более того, ГНСС и РСДБ позволяют определить ZWD на всей небесной сфере, рассматривая данную задержку как функцию от влажной тропосферной задержки в зените и градиента влажной тропосферной задержки, что связано с распределение ГНСС-спутников или квазаров на небесной сфере, в то время как РВП определяет задержку только в направлении зенита. Эти преимущества позволяют использовать данные РВП для аналитически оптимального определения стохастических параметров атмосферы в некоторых простых случаях, в частности при случайном блуждании.

Статья построена следующим образом. В первой главе описана модель данных и математические основы вычисления параметра  $\Phi$  случайного блуждания при двух условиях: 1) случайного блуждания без тренда и 2) случайного блуждания с линейным трендом; рассмотрены модификации стандартных методов проверки соответствия модели априорным условиям для данной задачи. Во второй главе приведен анализ данных РВП-наблюдений за 2019 г. Для станций «Зеленчукская», «Светлое» и «Бадары» вычислен параметр  $\Phi$  для двух моделей. На основе прогнозов для заранее известных наблюдений проверено качество моделей. Сделан вывод о применимости описания параметра ZWD как случайного блуждания.

## Математические основы

### Уравнение оценки параметра случайного блуждания $\Phi$

Пусть  $\{y_i\}_{i=1}^N$  — наблюдения некоторого процесса в моменты времени  $\{t_i\}_{i=1}^N$ . Предположим, что наблюдаемые данные содержат: некоторый шум  $\varepsilon_i$  с известной дисперсией  $\sigma^2$ , не зависящей от времени, тренд  $e(t)$  (возможно нулевой), а также некоторую стохастическую компоненту, описываемую случайным процессом  $x(t)$  с нулевым математическим ожиданием и ковариационной функцией  $k(t, s|P)$ , где  $P$  — некоторые неизвестные параметры. В случае если  $e(t)$  непрерывна, мы можем разложить ее с нужной точностью по некоторой наперед выбранной системе функций (например, по ортогональным полиномам либо по базису тригонометрического ряда Фурье)  $\{g_i(t)\}_{i=1}^{\infty}$ :  $e(t) \approx \sum_{i=1}^n a_i g_i(t) = G_i(t)e$ . Обозначив  $x_i = x(t_i)$ ,  $G_i = G(t_i)$ , мы получим модель  $i$ -того наблюдения:

$$y_i = x_i + G_i e_i + \varepsilon_i.$$

Перейдем к векторным обозначениям, положив

$$\begin{aligned} y &= (y_1^T, \dots, y_N^T)^T, \\ x &= (x_1^T, \dots, x_N^T)^T, \\ a &= (a_1^T, \dots, a_N^T)^T, \\ \varepsilon &= (\varepsilon_1^T, \dots, \varepsilon_N^T)^T, \\ G &= (G_1^T, \dots, G_N^T)^T. \end{aligned}$$

В таком случае модель принимает следующий вид:

$$y = x + Ge + \varepsilon. \quad (1)$$

Считая шумы наблюдений независимыми, получим  $\text{cov}(\varepsilon) = \sigma^2 \text{Id}$ , где  $\text{Id}$  — единичная матрица.

Обозначив ковариационную матрицу  $\text{cov}(x) = Q$ , данное уравнение можно записать в терминах регрессионного уравнения с дополнительными мягкими условиями, как это делается при выводе метода коллокации [7] и сглаживания Калмана [8]. Соответствующие регрессионные уравнения в случае аппроксимации с трендом и без тренда соответственно имеют вид:

$$\begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} \text{Id} & G \\ \text{Id} & 0 \end{pmatrix} \begin{pmatrix} x \\ e \end{pmatrix} + \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix},$$

$$\begin{pmatrix} y \\ 0 \end{pmatrix} = \begin{pmatrix} \text{Id} \\ \text{Id} \end{pmatrix} x + \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix},$$

где  $\text{cov} \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix} = \begin{pmatrix} Q & 0 \\ 0 & \sigma^2 \text{Id} \end{pmatrix}$ . Нормируя шумы системы к единичной дисперсии, имеем:

$$\begin{pmatrix} \sigma^{-1}y \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma^{-1}\text{Id} & \sigma^{-1}G \\ Q^{-1/2} & 0 \end{pmatrix} \begin{pmatrix} x \\ e \end{pmatrix} + \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix}, \quad (2)$$

$$\begin{pmatrix} \sigma^{-1}y \\ 0 \end{pmatrix} = \begin{pmatrix} \sigma^{-1}\text{Id} \\ Q^{-1/2} \end{pmatrix} x + \begin{pmatrix} \varepsilon \\ \delta \end{pmatrix}. \quad (3)$$

Кроме решения данных уравнений методом наименьших квадратов, мы также можем оценить дисперсию шума  $\varepsilon$  по остаточной сумме квадратов [4, с. 46–52] — очевидно, после нормировки данная величина должна равняться единице. Таким образом, получаем следующие системы:

$$\begin{cases} (\text{Id} + Q^{-1})x + Ge = y \\ G^T x + G^T Ge = G^T y \\ \frac{1}{N - \dim(e)} (\sigma^{-2} \|y - Ge - x\|^2 + x^T Q^{-1} x) = 1, \end{cases} \quad (4)$$

$$\begin{cases} \left( \text{Id} + Q^{-1} \frac{\sigma^2}{\Phi} \right) x = y \\ \frac{1}{N} (\sigma^{-2} \|y - x\|^2 + x^T Q^{-1} x) = 1. \end{cases} \quad (5)$$

Рассмотрим теперь матрицу  $Q$  в случае, если стохастический процесс  $x$  описывается случайным блужданием. По определению  $Q_{i,j} = k(t_i, t_j)$ , где  $k(t_i, t_j) = \Phi \min(t_i, t_j)$ . Вынеся общий множитель  $\Phi$ , получим:  $Q = \Phi \hat{Q}$ ,  $\hat{Q}$  — постоянная матрица, на  $(i, j)$ -ом месте которой стоит  $\min(t_i, t_j)$ . Подставив

данное значение в формулы (4), (5), получаем итоговые уравнения для вычисления параметра  $\Phi$ :

$$\begin{cases} \left( \text{Id} + \hat{Q}^{-1} \frac{\sigma^2}{\Phi} \right) x + Ge = y \\ G^T x + G^T Ge = G^T y \\ \Phi = \frac{x^T \hat{Q}^{-1} x}{N - \dim(e) - \sigma^{-2} \|y - Ge - x\|^2} \end{cases} \quad (6)$$

$$\begin{cases} \left( \text{Id} + \hat{Q}^{-1} \frac{\sigma^2}{\Phi} \right) x = y \\ \Phi = \frac{x^T \hat{Q}^{-1} x}{N - \sigma^{-2} \|y - x\|^2} \end{cases} \quad (7)$$

в случаях модели с ненулевым трендом и без тренда соответственно. Заметим, что при решении данных уравнений мы также автоматически вычислим значения процесса  $x = (x_1, \dots, x_N)$  и тренд  $e$ .

Решение уравнений находится методом последовательных итераций. В случае уравнения 6: пусть  $(e^{(i)}, x^{(i)})$  — решение, найденное на  $i$ -ом шаге. Подставляя данные значения в третье уравнение системы (6), найдем новое значение  $\Phi^{(i)}$ . Решая линейную систему первых двух уравнений системы (6) при  $\Phi = \Phi^{(i)}$ , находим значения  $(e^{(i+1)}, x^{(i+1)})$ . Повторяем данный алгоритм до тех пор, пока значения  $\|e^{(m+1)} - e^{(m)}\|$ ,  $\|x^{(m+1)} - x^{(m)}\|$ ,  $\|(\Phi^{-1})^{(m+1)} - (\Phi^{-1})^{(m)}\|$  не превосходят некоторых наперед заданных значений — точности вычисления.

*Алгоритмы быстрого решения систем (6), (7)*

Самый медленный шаг данного алгоритма — решение линейного уравнения. При числе наблюдений  $N$  вычислительная сложность решения уравнения  $O(N^3)$ . Учитывая, что разрешение РВП-наблюдений может достигать 10 с, даже в случае часового интервала  $N = 600$ , суточного —  $N = 14400$ .

Для ускорения данного процесса при вычислении параметров модели с нулевым трендом (7) использовалось разложение Жордана матрицы  $\hat{Q}^{-1}$ :  $\hat{Q}^{-1} = J^T \Lambda J$ ,  $J^T = J^{-1}$ . Поскольку  $\hat{Q}^{-1}$  положительно определена,  $\Lambda$  — диагональная матрица со строго положительными элементами. Сделав замену координат  $x' = Jx$ ,  $y' = Jy$ , получаем следующее тривиальное уравнение с диагональной матрицей:  $(\text{Id} + \Lambda \frac{\sigma^2}{\Phi}) x' = y'$ . Таким образом каждая итерация становится линейной по  $N$ .

В случае системы (6) метод замены координат не работает, поскольку появляются дополнительные члены. Однако в таком случае можно воспользоваться тем фактом, что  $(\text{Id} + \hat{Q}^{-1} \frac{\sigma^2}{\Phi})$  — трехдиагональная матрица и описать некоторое обобщение метода прогонки, позволяющее вычислять систему за линейное по  $N$  время.

Итак, мы имеем следующую линейную систему:

$$\begin{cases} a_1x_1 + b_2x_2 + c_1e & = r_1 \\ b_ix_{i-1} + a_ix_i + b_{i+1}x_{i+1} + c_ie & = r_i \\ b_Nx_{N-1} + a_Nx_N + c_Ne & = r_N \\ c_1^T x_1 + \dots + c_N^T x_N + ce & = r \end{cases} \quad i \in [2: N-1], \quad (8)$$

где  $\{a_i\}_{i=1}^N$ ,  $\{b_i\}_{i=2}^N$ ,  $\{r_i\}_{i=1}^N$  — некоторые числа;  $\{c_i\}_{i=1}^N$  — матрицы размера  $(1, n)$ ,  $c$  — матрица размера  $(n, n)$ ;  $r$  — вектор длины  $n$ .

Наша первая цель — обнулить коэффициент перед  $x_{i-1}^s$  в  $i$ -ом уравнении системы (8) и таким образом получить:

$$A_i x_i + b_{i+1} x_{i+1} + C_i e = R_i \quad (9)$$

для всех  $i \in [1: N]$ .

Вычисляем по индукции. Очевидно, первое уравнение системы (8) уже имеет форму (9), таким образом  $A_1 = a_1$ ,  $C_1 = c_1$ ,  $R_1 = r_1$ . Предположим, значения  $A_{i-1}$ ,  $C_{i-1}$ ,  $R_{i-1}$  вычислены. Умножая  $(i-1)$ -ое уравнение на  $(-b_i A_{i-1}^{-1})$  и прибавляя  $i$ -ое уравнение системы (8), получаем:

$$(a_i - A_{i-1}^{-1} b_i^2) x_i + b_{i+1} x_{i+1} + (c_i - b_i A_{i-1}^{-1} C_{i-1}) x^c = r_i - b_i A_{i-1}^{-1} R_{i-1}.$$

Таким образом, находим следующие рекуррентные формулы на коэффициенты уравнения (9):

$$\begin{aligned} A_1 &= a_1; & A_i &= a_i - b_i^2 A_{i-1}^{-1}, \\ C_1 &= c_1; & C_i &= c_i - b_i A_{i-1}^{-1} C_{i-1}, \\ R_1 &= r_1; & R_i &= r_i - b_i A_{i-1}^{-1} R_{i-1}. \end{aligned}$$

Далее сократим  $\hat{x}_i$ ,  $i \in [1: N]$  из последнего уравнения системы (8). Для этого отнимем уравнение (9), домноженное на некоторую константу, из  $(N+1)$ -ого уравнения. Заметим, что в таком случае меняются лишь коэффициенты при  $x_{i+1}$  и  $e$ , а также правая часть  $r$ . Таким образом, после  $(i-1)$  шага мы сократим члены  $x_1^s, \dots, x_{i-1}^s$ , а также изменим коэффициенты  $x_i^s$ ,  $x^c$  и правую часть  $r$ . Таким образом  $(N+1)$ -ое уравнение системы (8) станет:

$$c'_i x_i^s + c'_{i+1} x_{i+1}^s + \dots + c_N^T x_N^s + c'^{x^c} = r'.$$

Отнимая из этого уравнения уравнение (9), домноженное на  $(c'_i A_i^{-1})$ , получим:

$$(c'_{i+1} - c'_i A_i^{-1} b_{i+1}) x_{i+1} + \dots + (c' - c'_i A_i^{-1} C_i) x^c = (r' - c'_i A_i^{-1} R_i).$$

Очевидно, начальные значения  $c'_1 = c_1^T$ ,  $c' = c$ ,  $r' = r$ . Следовательно:

$$\begin{aligned} c'_1 &= c_1^T; & c'_i &= c_i^T - c_{i-1}' A_{i-1}^{-1} b_i, \\ c' &= c - \sum_{i=1}^N c'_{k-1} A_{k-1}^{-1} C_{k-1}, \\ r' &= r - \sum_{i=1}^N c'_{k-1} A_{k-1}^{-1} R_{k-1}. \end{aligned}$$

После обнуления всех  $x_i$  мы получим уравнение  $c'e = r'$ , таким образом  $e = c'^{-1} r'$ . Подставляя найденное значение в (9) для всех  $i \in [N: 1]$ , находим:

$$\begin{aligned} e &= c'^{-1} r' \\ x_N &= A_N^{-1} (R_N - C_N e) \\ x_i &= A_i^{-1} (R_i - b_{i+1} x_{i+1} - C_i e) \quad i \in [N-1: 1]. \end{aligned}$$

Легко найти, что данный алгоритм работает за линейное по  $N$  и кубическое по  $n$  время.

### Прогнозирование значения тропосферной задержки по двум наблюдениям

Одной из важных задач анализа ZWD является прогнозирование значения в тот момент времени  $t_*$ , когда наблюдение не производится.

Пусть  $\hat{x}$  — оценка параметров. Ее ковариация в случае модели с трендом определяется по формуле

$$\left( \begin{array}{cc} (\sigma^{-2} \text{Id} + Q^{-1}) & \sigma^{-2} G \\ \sigma^{-2} G^T & \sigma^{-2} G^T G e \end{array} \right)^{-1}$$

или по формуле

$$(\sigma^{-2} \text{Id} + Q^{-1})^{-1}$$

в случае модели без тренда.

Поскольку процесс марковский, его распределение в момент времени  $t_*$  однозначно определяется по двум ближайшим значениям  $(x_i, x_{i+1})$ ,  $t_i < t_* < t_{i+1}$ . Воспользовавшись формулой условного распределения для многомерного нормального распределения, получим:

$$\begin{aligned} p(e(t_*) + x(t_*) | e, x_i, x_{i+1}) &= \\ &= \int_{\mathbb{R}^3} \left( \frac{t_{i+1} - t_*}{t_{i+1} - t_i} x_{i+1} + \frac{t_* - t_i}{t_{i+1} - t_i} x_i + G(t_*) e \right) \cdot \\ &\quad \cdot p'(x_i, x_{i+1}, e) dx_i dx_{i+1} de, \end{aligned}$$

где  $p'(x_i, x_{i+1}, e)$  — плотность распределения полученной оценки параметров. Аналогично, для случая модели без тренда:

$$p(x(t_*) | x_i, x_{i+1}) = \int_{\mathbb{R}^2} \left( \frac{t_{i+1} - t_*}{t_{i+1} - t_i} x_{i+1} + \frac{t_* - t_i}{t_{i+1} - t_i} x_i \right) p'(x_i, x_{i+1}) dx_i dx_{i+1}.$$

Данный интеграл можно трактовать как математическое ожидание линейной функции от многомерного нормального распределение, так что он легко вычисляется методом Монте-Карло.

### Обработка данных РВП-наблюдений

Для вычисления параметра случайного блуждания были взяты ряды наблюдений РВП на станциях «Светлое», «Зеленчукская» и «Бадарь» за 2019 г. с временным разрешением 1 мин. Основные цели обработки следующие:

1) определение параметра случайного блуждания  $\Phi$  для моделей с трендом и без тренда на суточном интервале;

2) проверка удовлетворения данных априорным условиям модели.

На практике поиск параметров происходил следующим образом. На первом шаге производилась предварительная фильтрация общего массива данных. При этом удалялись данные в моменты

времени, когда значение напряжения датчика осадков РВП превышало некоторое пороговое значение. Это необходимо, поскольку показания РВП во время осадков некорректны [9]. Также удалялись наблюдения, отличающиеся от ближайших соседних наблюдений более чем на 10 мм.

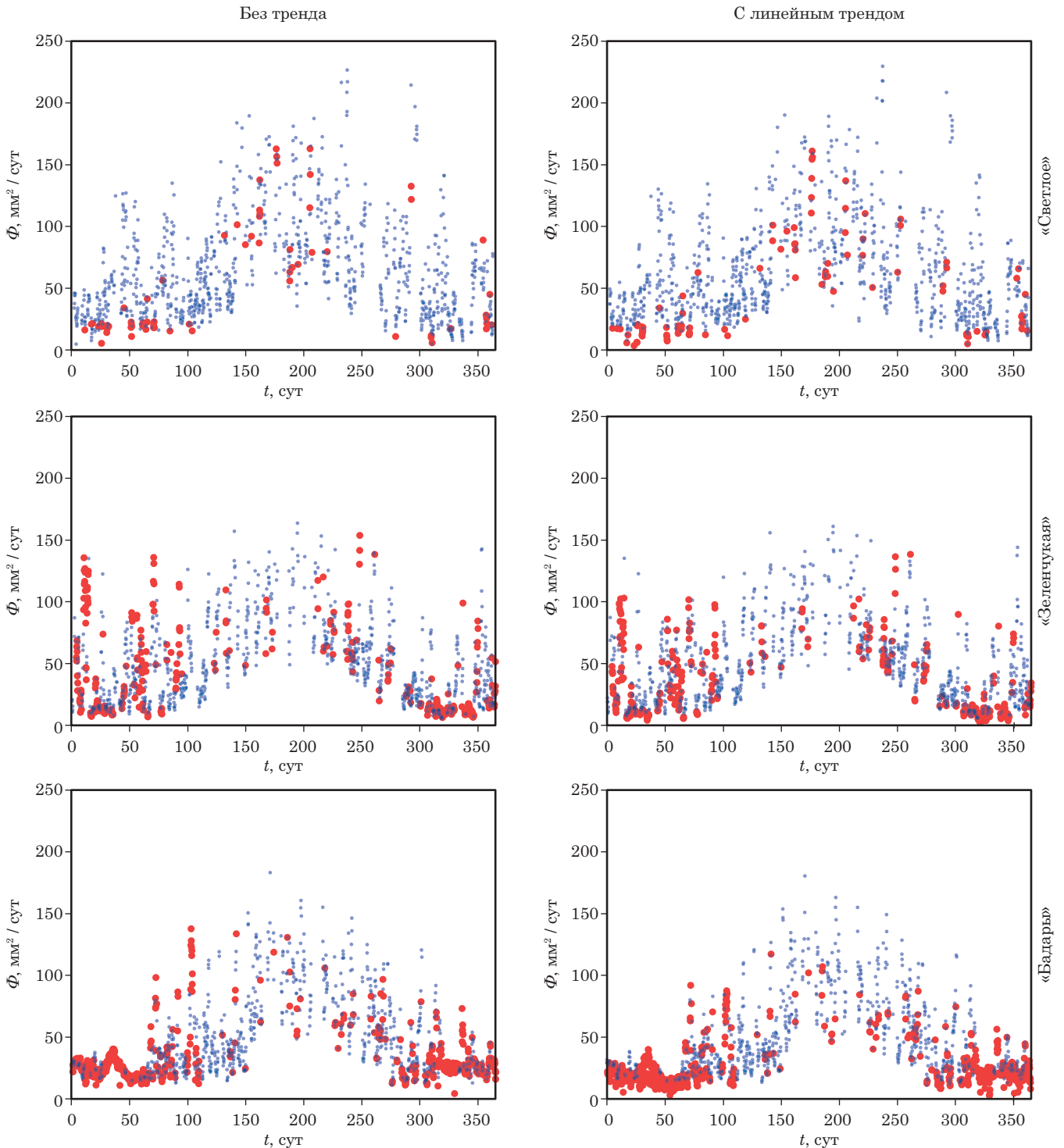


Рис. 1. Значение параметра  $\Phi$  для станций «Светлое», «Зеленчукская», «Бадарь». Красными точками отмечены значения, удовлетворяющие гипотезе о нормальном распределении, синими — не удовлетворяющие (тест Шапиро — Уилка, уровень значимости  $\alpha = 0.001$ )

После предварительной фильтрации данные разбивались на (пересекающиеся) суточные интервалы, начинающиеся в 0, 4, 8, 12, 16, 20 часов. Для каждого интервала производилась априорная оценка сверху дисперсии инструментального шума  $\sigma_{\max}^2$ . Далее, для каждого значения дисперсии

$\sigma^2$  из множества  $\{(0.01n)^2\}_{n=1}^k$  ( $k$  — максимальное целое, не превосходящее  $100\sigma_{\max}$ ) производился поиск параметра  $\Phi$  для моделей с трендом и без. После этого производился анализ остатков методом построения 90 %-ого доверительного интервала для невязок и наблюдения-выбросы убирались

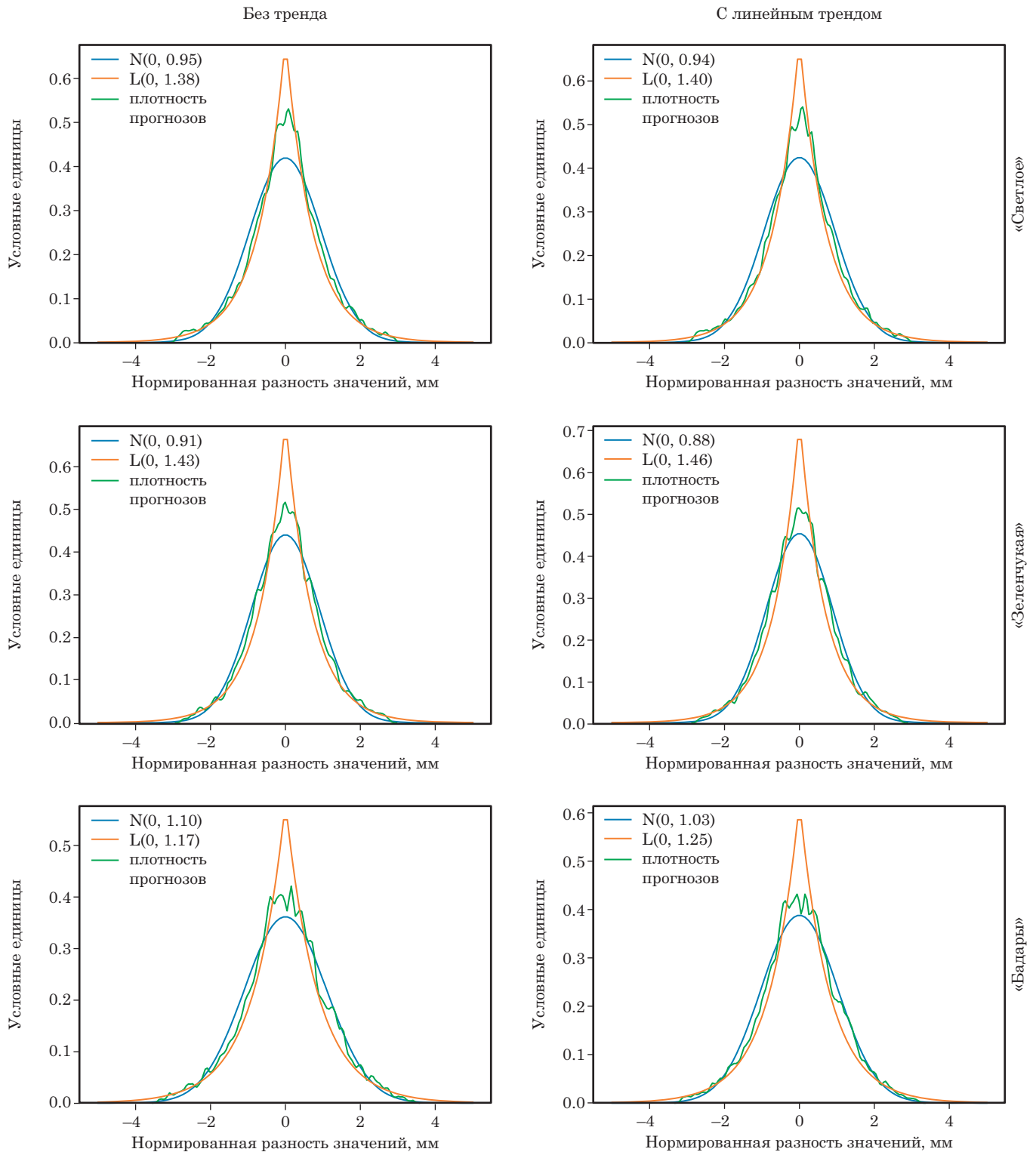


Рис. 2. Плотность разности спрогнозированного на одну минуту и реального значений влажной задержки в зенит; для сравнения также указаны нормальное распределение  $\mathcal{N}$  и распределение Лапласа  $\mathcal{L}$

из последовательности. Данные действия проводились до тех пор, пока оставались выбросы.

На последнем шаге для каждой пары  $(\sigma^2, \Phi)$  вычислялось  $p$ -значение критерия нормальности Шапиро – Уилка. Среди всех пар  $(\sigma^2, \Phi)$  выбира-

лась та пара, которая давала максимальное  $p$ -значение. В случае, если данное значение превышало  $10^{-3}$ , интервал признавался удовлетворяющим модели, в противном случае модель признавалась не соответствующей наблюдениям.

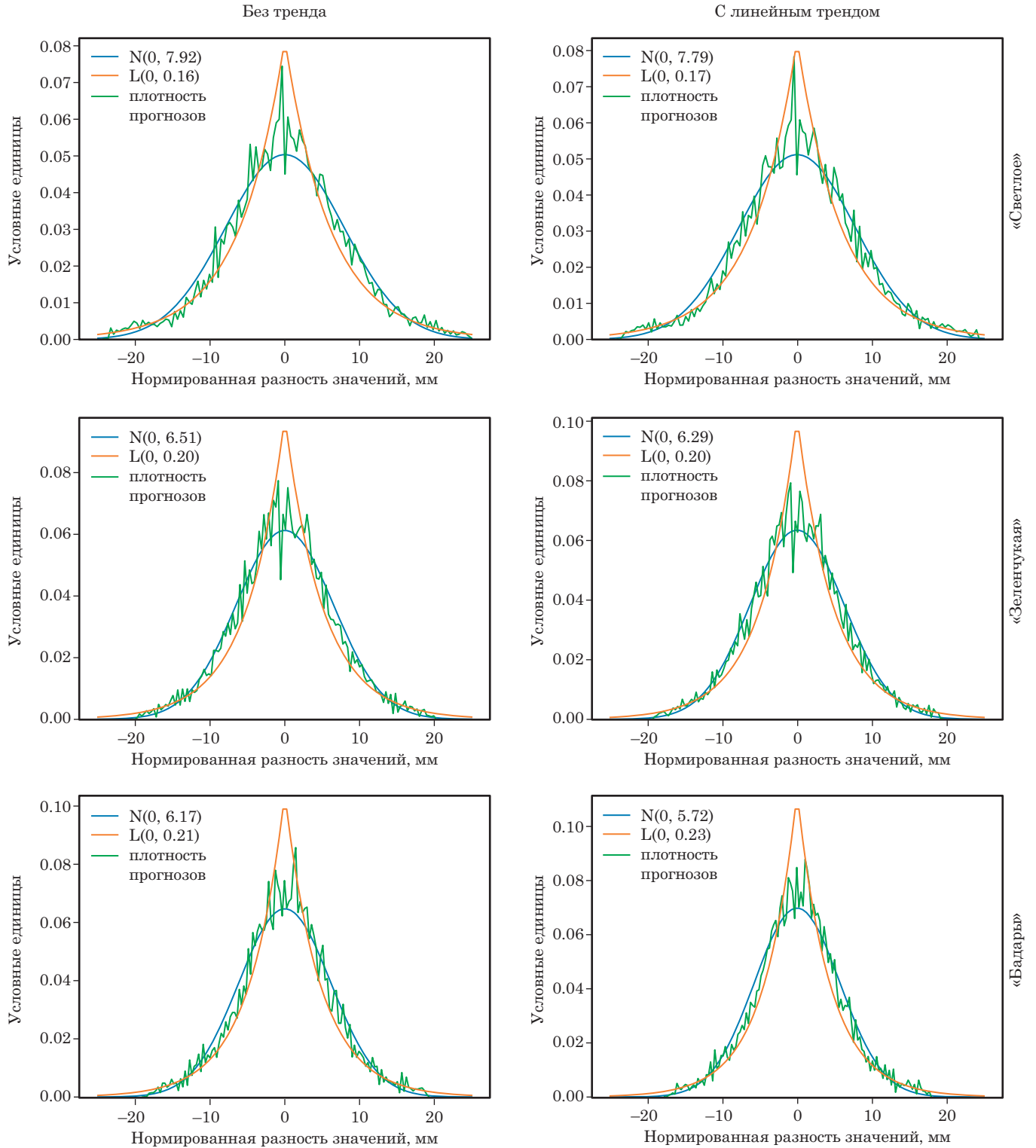


Рис. 3. Плотность разности спрогнозированного на один час и реального значений влажной задержки в зените. Для сравнения также указаны нормальное распределение  $\mathcal{N}$  и распределение Лапласа  $\mathcal{L}$

Полученные результаты изображены на рис. 1. Как видно из рисунка, нормальность распределения остатков в большинстве случаев не выполняется как для модели без тренда, так и для модели с трендом. В частности, для станции «Бадарь» данное условие не верно почти всегда в течение года; для станций «Зеленчукская» и «Светлое» условие не выполняется с середины апреля до конца октября. Чаще всего условие нормальности выполнено поздней осенью, зимой и ранней весной (с ноября по апрель) на станции «Бадарь».

Отметим, что именно в эти месяцы наблюдаются малые значения ZWD, в том время как с середины весны до середины осени значения ZWD существенно больше, максимальное значение достигается в июле. Таким образом, можно предположить, что описание ZWD как нормального случайного блуждания хорошо работает только в том случае, если ее величина не превосходит некоторого значения.

Другим методом проверки соответствия данных описанным моделям было исследование соответствия предсказательной способности модели. Для этого из анализируемых данных заранее были удалены некоторые наблюдения  $x_{real}$ , значение которых требовалось восстановить по выборке. При использовании описанного способа прогнозирования были оценены значения удаленных наблюдений  $x_{est}$  и дисперсия оценки  $\sigma_{est}^2$ . В случае если модель верна, набор данных  $\frac{x_{real}-x_{est}}{\sigma_{est}}$  должен быть выборкой из несмещенного нормального распределения с единичной дисперсией.

Полученные результаты изображены на рис. 2–3 для точек, удаленных на минуту и час соответственно. Как легко заметить, полученные распределения не являются стандартным нормальным распределением ни для какой из моделей: в полученных распределениях можно заметить некоторую компоненту, похожую на распределение Лапласа с плотностью  $p_\alpha(x) = \frac{\alpha}{2} e^{-\alpha\|x\|}$ , где большая вероятность отклонений, близких к нулю, и более тяжелые хвосты. Кроме того, прогноз на час не дает единичную дисперсию, как это должно следовать в случае, если бы данные соответствовали нормальному случайному блужданию.

## Заключение

Таким образом, показано, что применяющаяся в некоторых программных пакетах модель описания влажной компоненты задержки радиосигнала в зените как случайного блуждания адекватно

описывает данные лишь в тех случаях, когда значения ZWD малы. В ином случае аппроксимация не дает достаточной точности прогнозов и, следовательно, должна с осторожностью применяться для стохастической аппроксимации данного параметра в методе фильтрации и сглаживания Калмана. Наконец, в случае если данное приближение применяется, следует отдать предпочтение модели с трендом, которая дает несколько лучшее приближение.

Также, исходя из графиков плотности прогнозов 2 и 3, можно сделать предположение, что ZWD является не нормальным случайным блужданием, а суммой нормального и лапласовского случайных блужданий — то есть процессом, чьи значения  $x_1$ ,  $x_2$  в моменты времени  $t_1$ ,  $t_2$  связаны соотношением  $x_2 - x_1 \sim \mathcal{N}(0, \Phi_n) + \mathcal{L}(0, \frac{\sqrt{2}}{\Phi_1 \sqrt{\min(t_1, t_2)}})$ . Данное предположение требует дальнейшего изучения.

## Литература

1. Herring T. A., Davis J. L., Shapiro I. I. Geodesy by radio interferometry: the application of kalman filtering to the analysis of very long baseline interferometry data // Journal of Geophysical Research: Solid Earth. 1990. Vol. 95, no. B8. P. 12561–12581.
2. Nilsson T., Soja B., Karbon M., et al. Application of kalman filtering in VLBI data analysis // Earth, Planets and Space. 2015. Vol. 67, no. 1. P. 136.
3. Titov O., Soja B., Tesmer V., Böhm J. OCCAM v.6.0 Software for VLBI Data Analysis // IVS 2004 General Meeting Proceedings. 2004. P. 267–271.
4. Губанов В. С. Обобщенный метод наименьших квадратов. Теория и применение в астрометрии. СПб.: Наука, 1997. 318 с.
5. Губанов В. С., Суркис И. Ф., Козлова И. Ф., Русин Ю. Л. Обработка РСДБ-наблюдений: программный пакет «QUASAR» // Сообщения ИПА РАН. 2002. № 141–145.
6. Ильин Г. Н., Быков В. Ю., Стэмповский В. Г., Шишкин А. М. Высокостабильный двухканальный радиометр водяного пара для измерений тропосферной задержки в реальном времени // Труды ИПА РАН. 2013. Вып. 27. С. 210–215.
7. Gubanov V. S. Parametric adjustment of astrometric data // Astronomical & Astrophysical Transactions. Taylor & Francis. 1993. Vol. 4, no. 2. P. 117–142.
8. Aravkin A., Burke J. V., Ljung L., et al. Generalized Kalman smoothing: modeling and algorithms // Automatica. 2017. Vol. 86. P. 63–86.
9. Суворкин В. В., Ильин Г. Н., Курдубов С. Л. и др. О точности определений тропосферной задержки сигнала // Труды ИПА РАН. 2018. № 46. С. 100–107.