Измерение риска с восстановлением пропущенных данных

Смирнов С., ГУ-ВШЭ, директор российского отделения PRMIA

Косьяненко А., лаборатория финансовой инженерии и риск менеджмента ГУ-ВШЭ

Москва 18 июня 2007 г.

Возможные применения

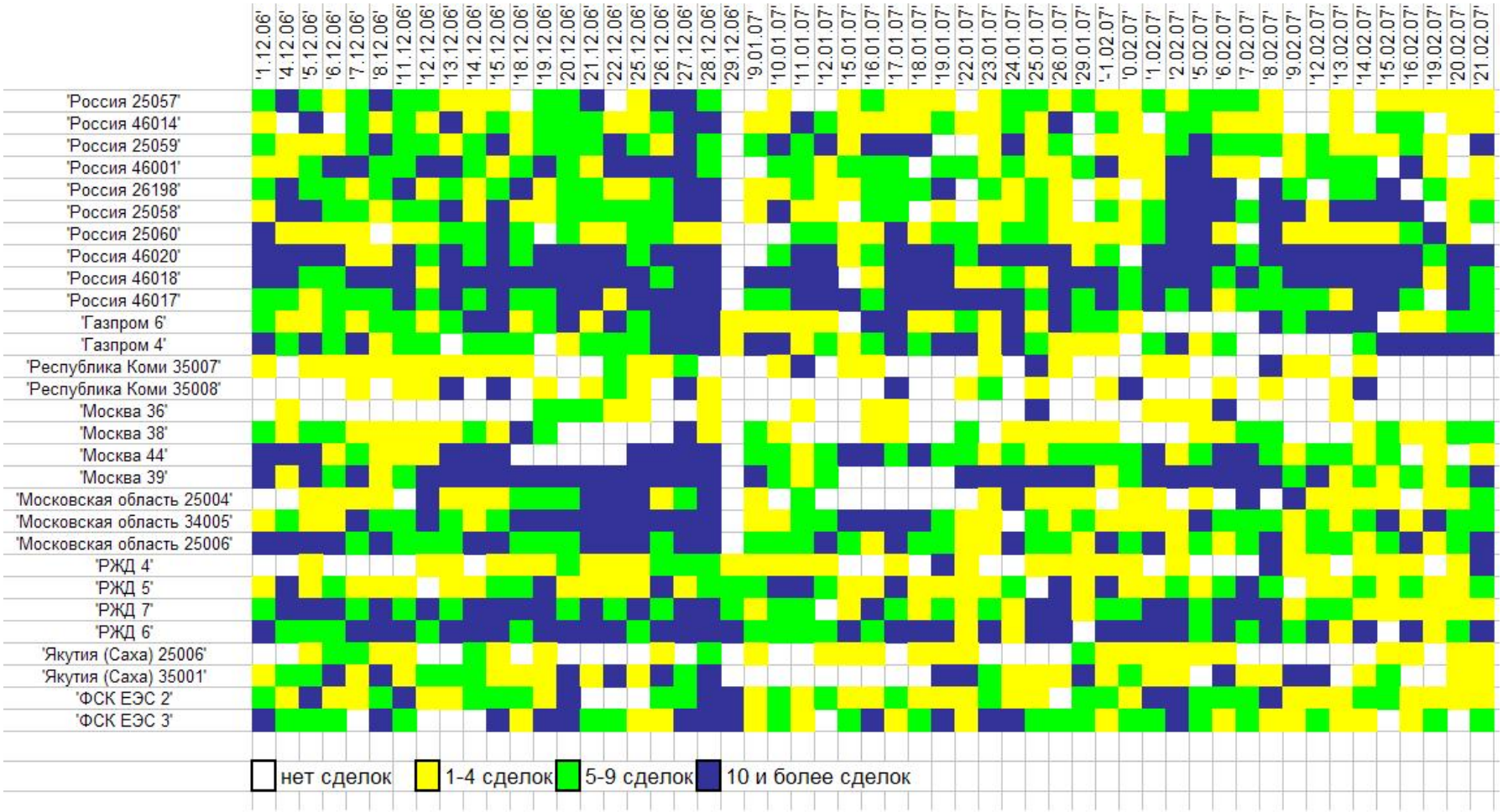
На рынке с низкой ликвидностью:

- Корректный расчет показателей риска портфеля
- Ценообразование, включая производные финансовые инструменты
- Построение бескупонных кривых доходности по группе облигаций

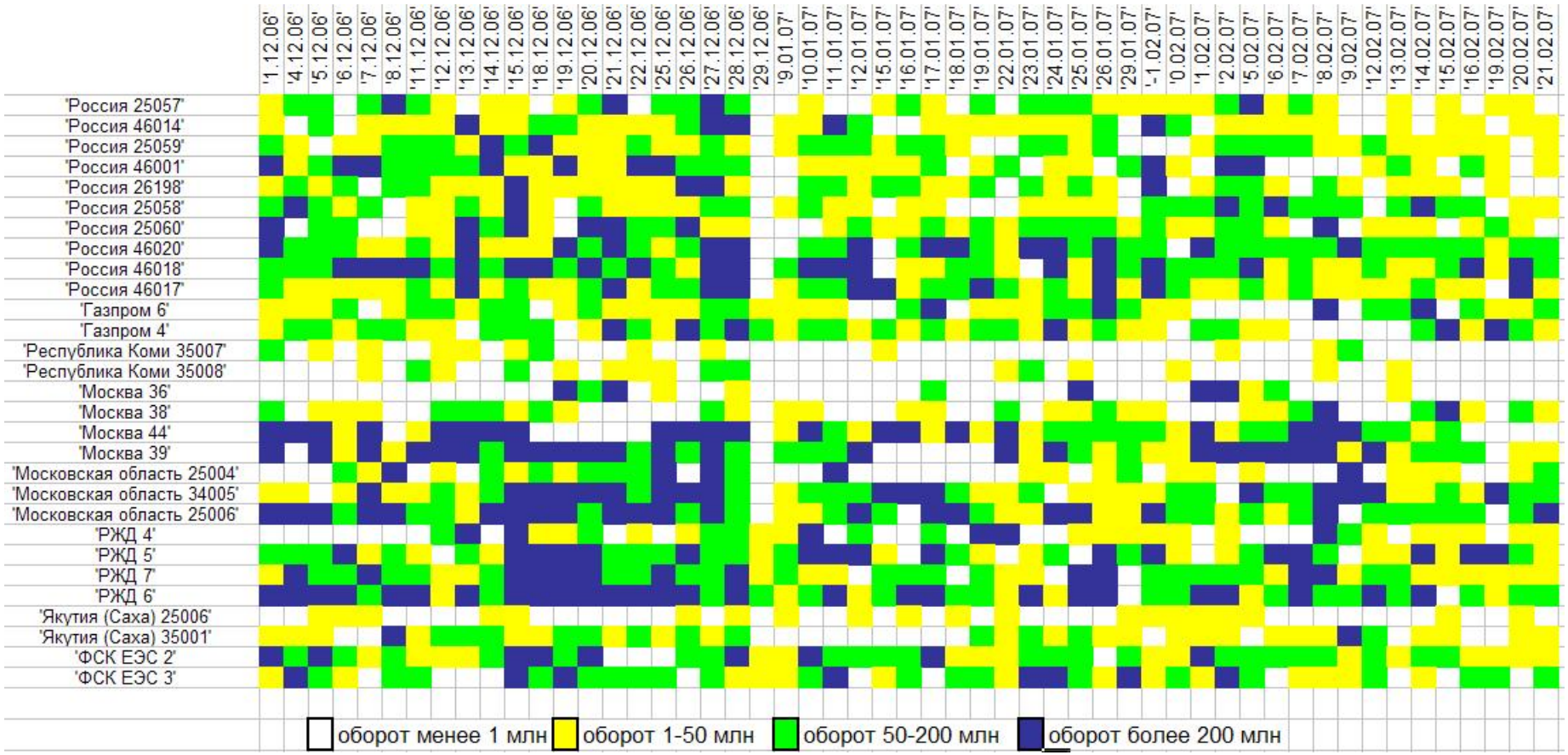
Возможные причины отсутствия рыночной информации

- Отсутствие сделок.
- Временное приостановление торгов.
- Сбои в процедуре накопления информации.
- Наличие неправдоподробных сделок и работа процедур фильтрации данных.
- Применение методики для прогнозирования будущей ситуации.

“Пропущенные” данные

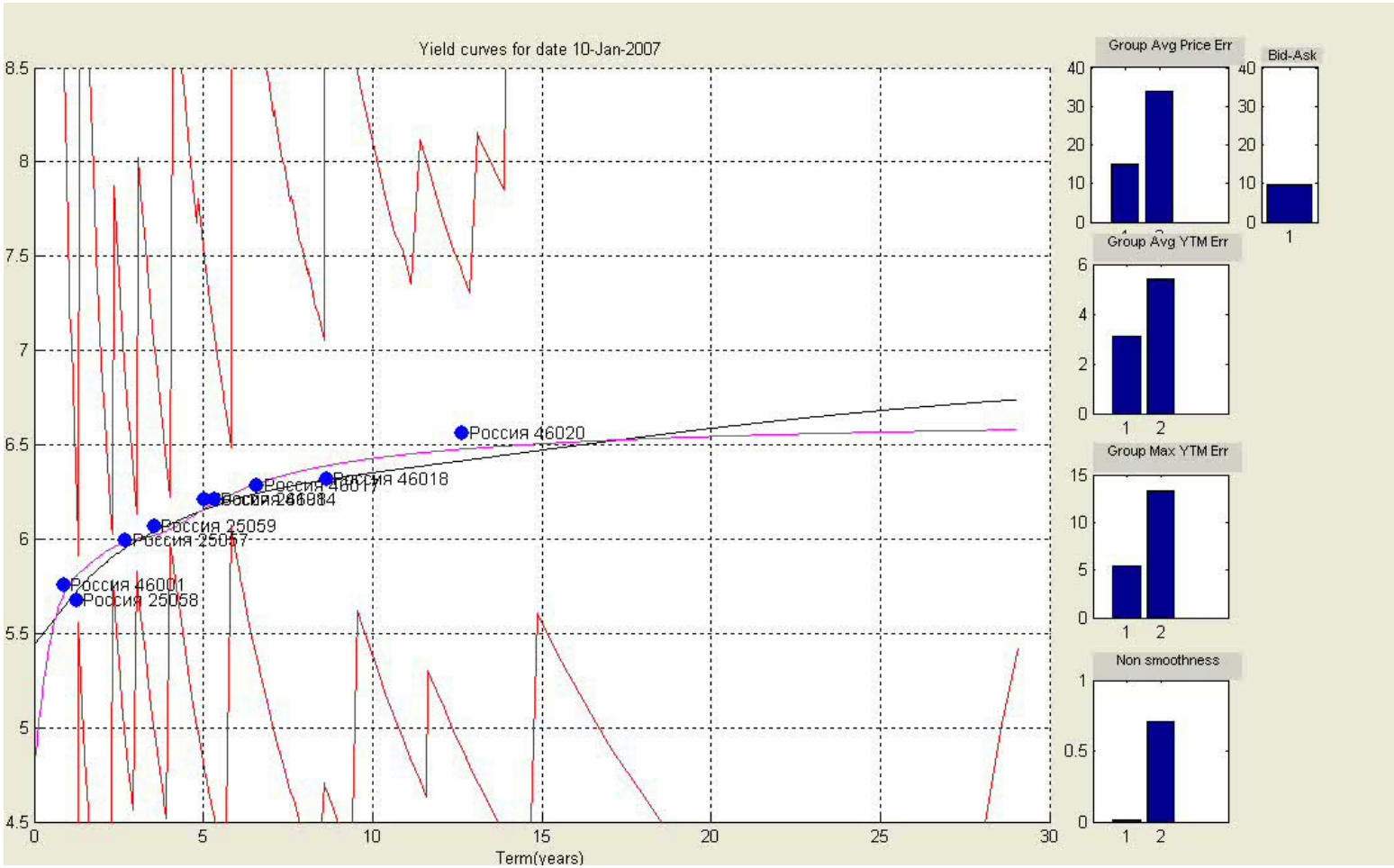


Объёмы заключенных сделок



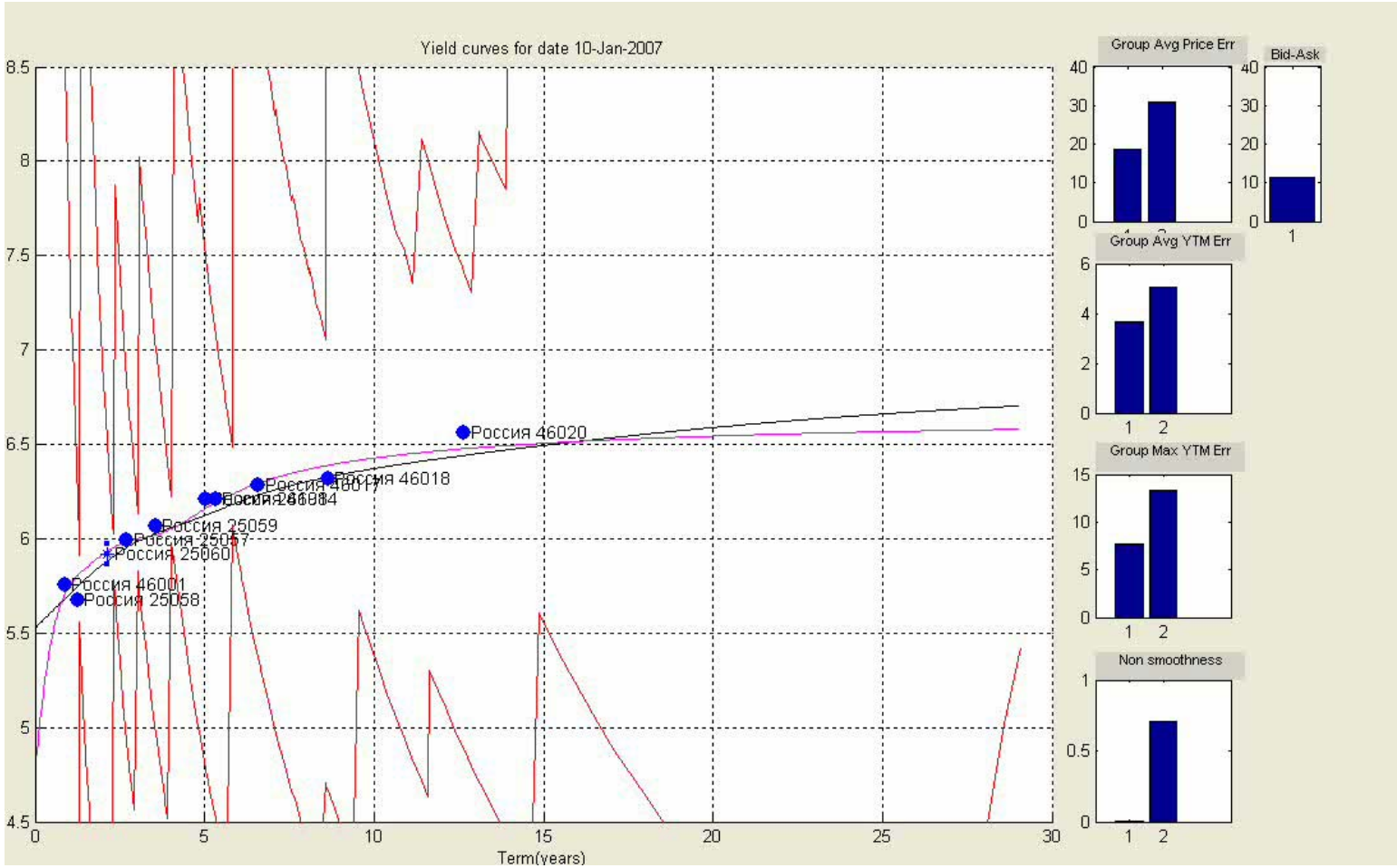
Москва 18 июня 2007 г.

Бескупонные кривые доходностей без фильтрации и восстановления

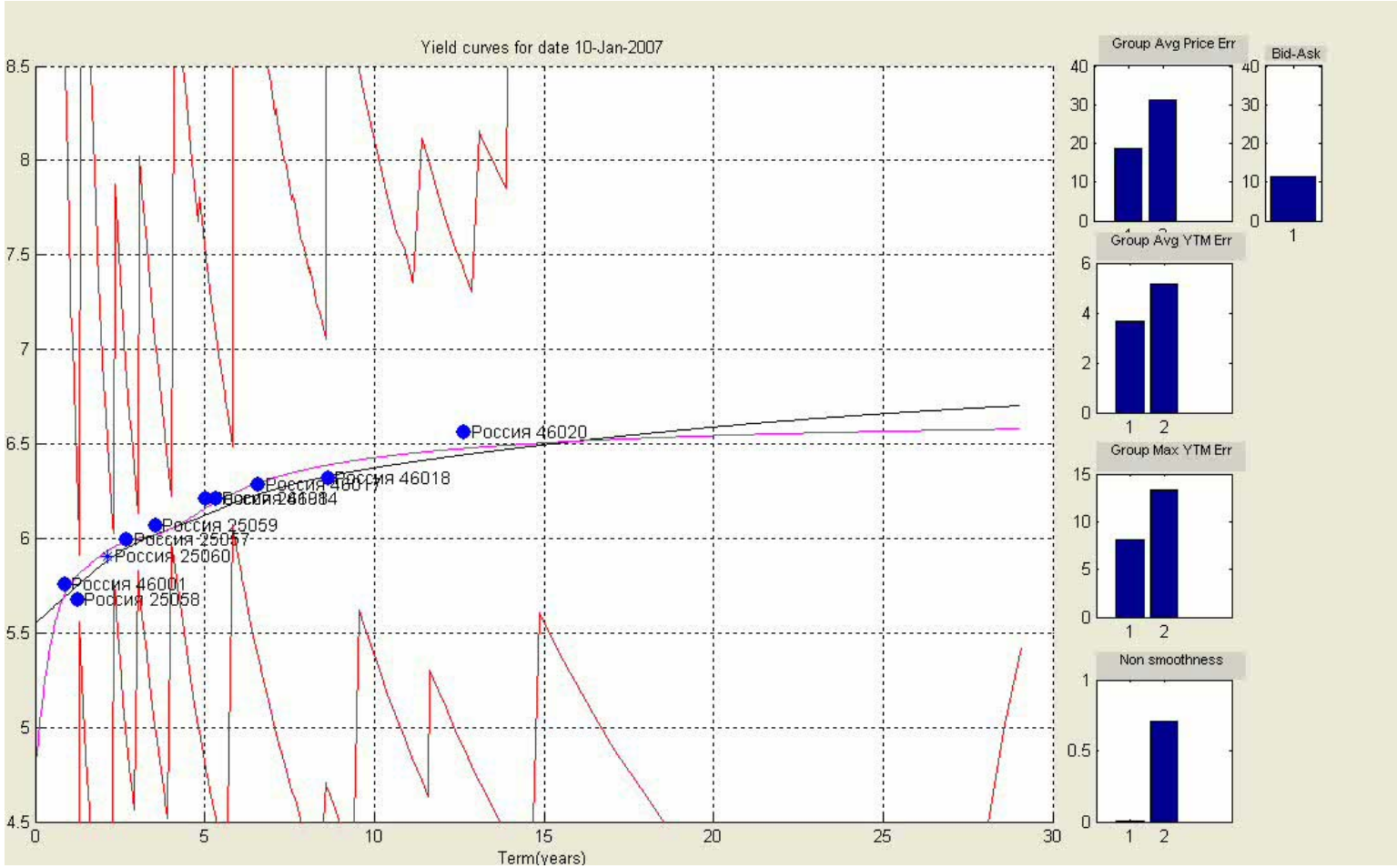


Москва 18 июня 2007 г.

Бескупонные кривые доходностей без фильтрации с восстановлением

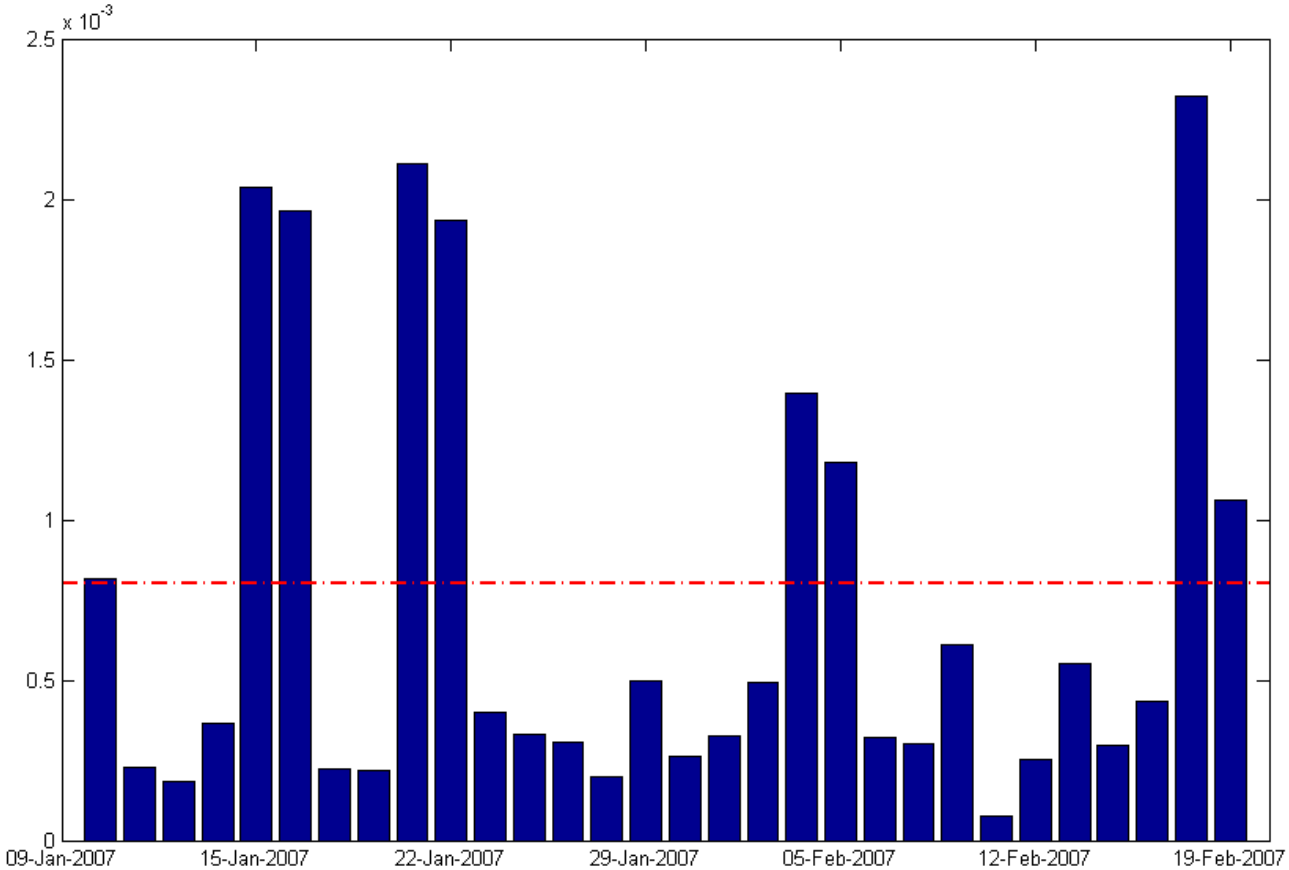


Бескупонные кривые доходностей с фильтрацией и с восстановлением

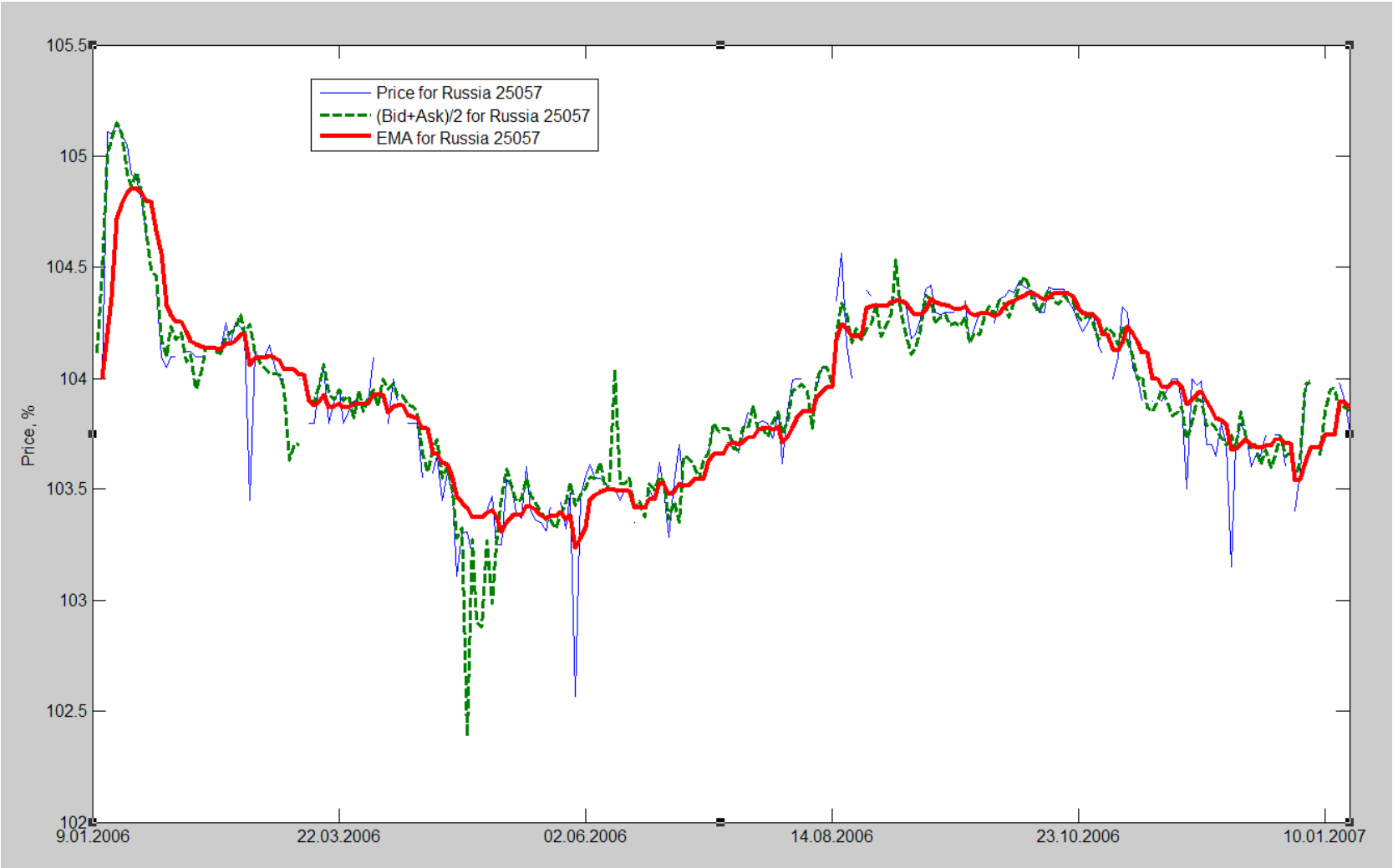


Москва 18 июня 2007 г.

Расстояние между соседними кривыми



Типичное изменение цены облигации

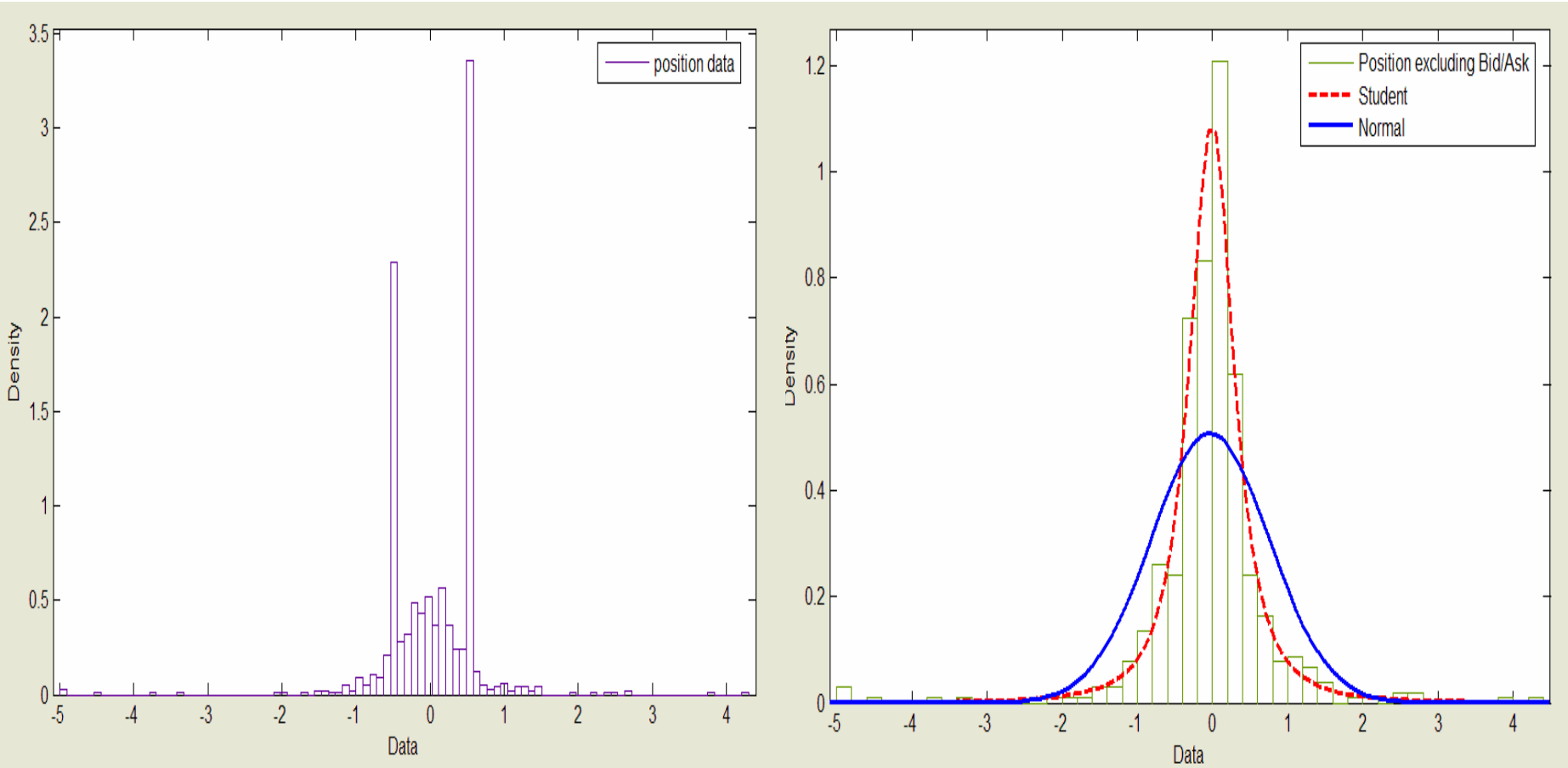


$T = 1 - \frac{\varepsilon^2}{\varepsilon_0^2}$

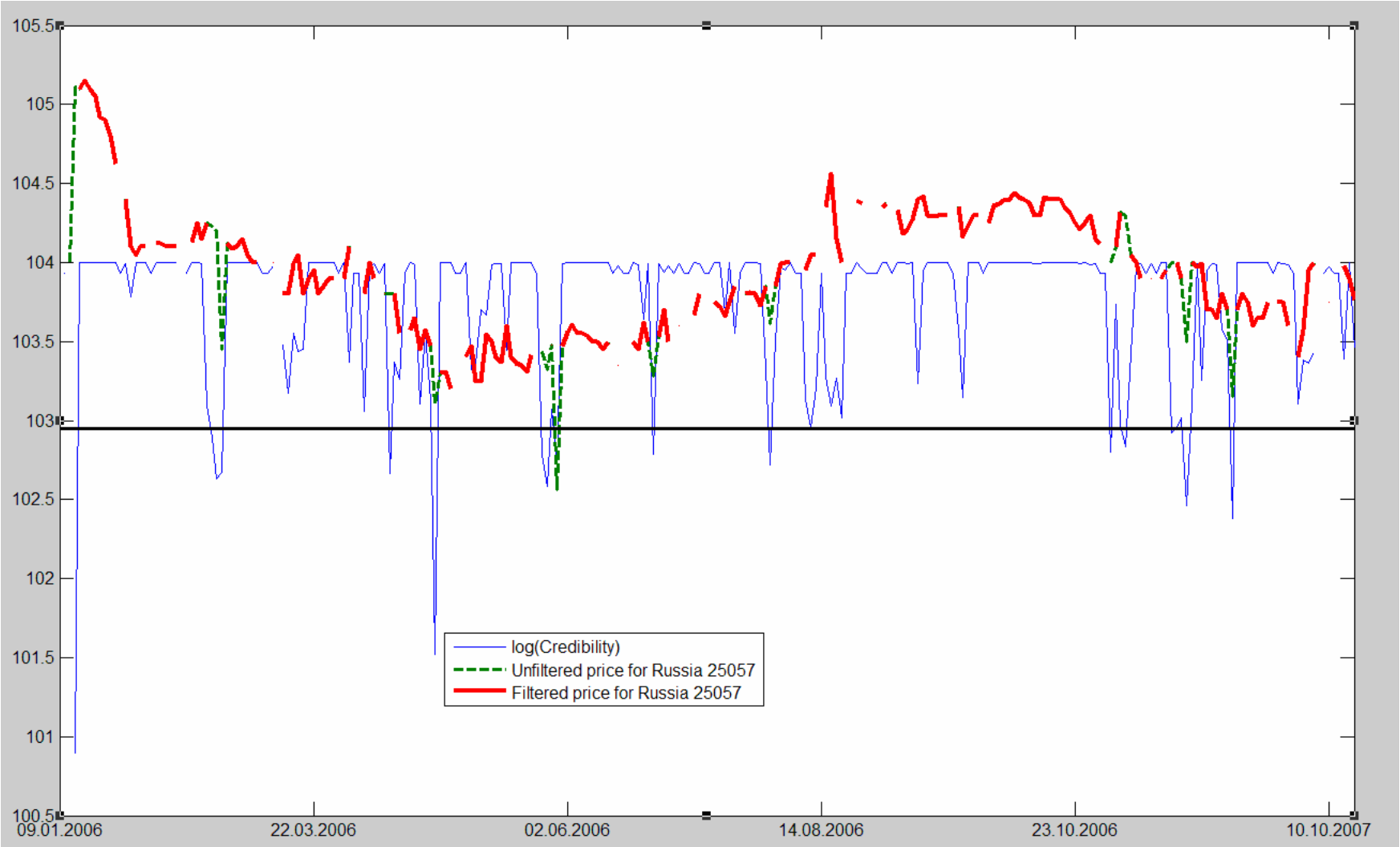
Виды использованных фильтров

- Фильтры уровня.
- Фильтры изменения значения.
- Фильтр относительного положения цены относительно котировок.
- Фильтры объёма (количество сделок, оборот).

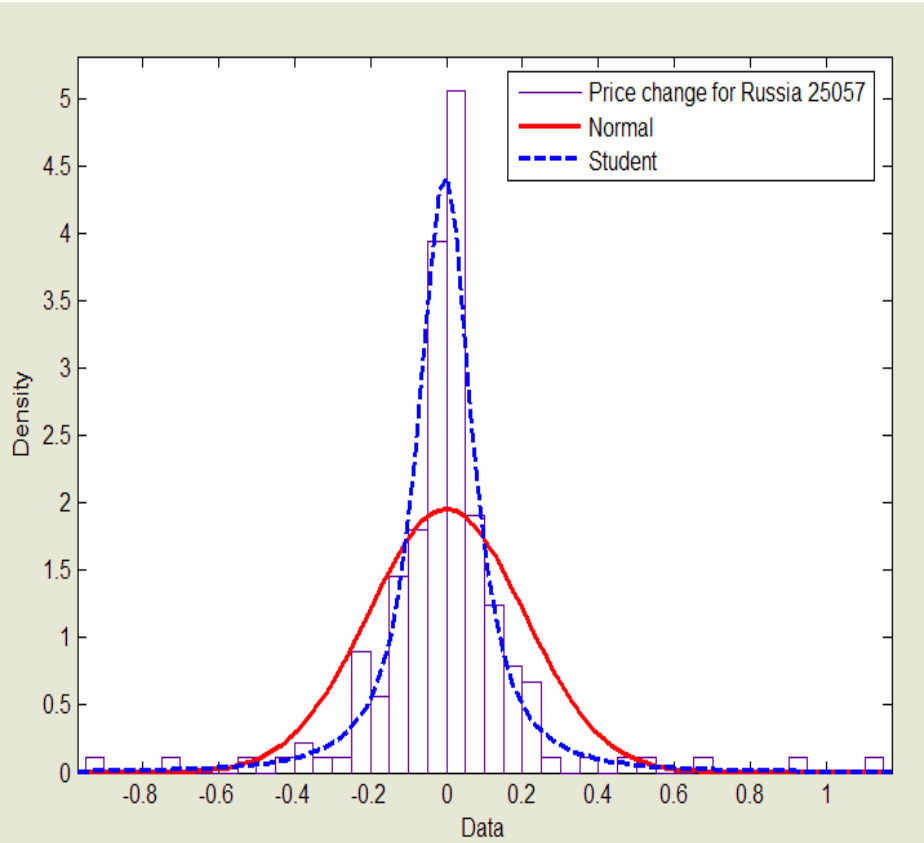
Положение цен закрытия относительно котировок

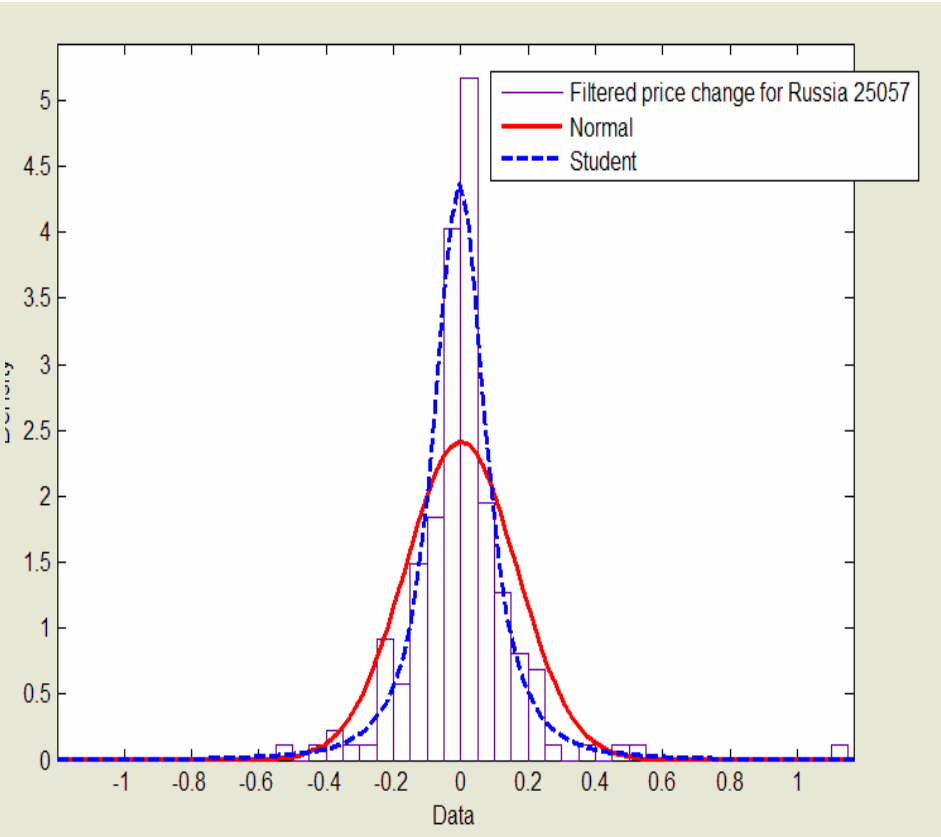


Результаты фильтрации

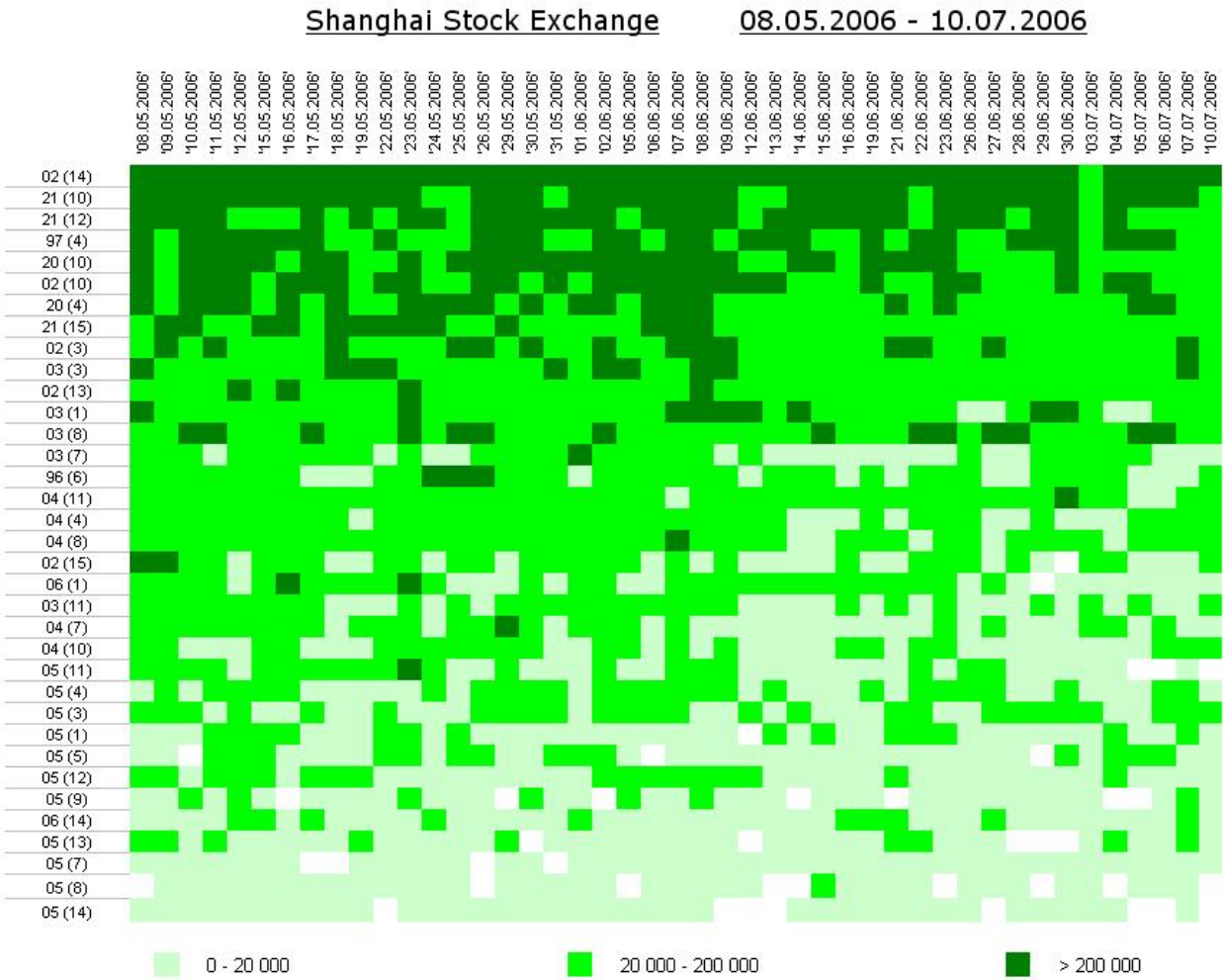


Результаты фильтрации



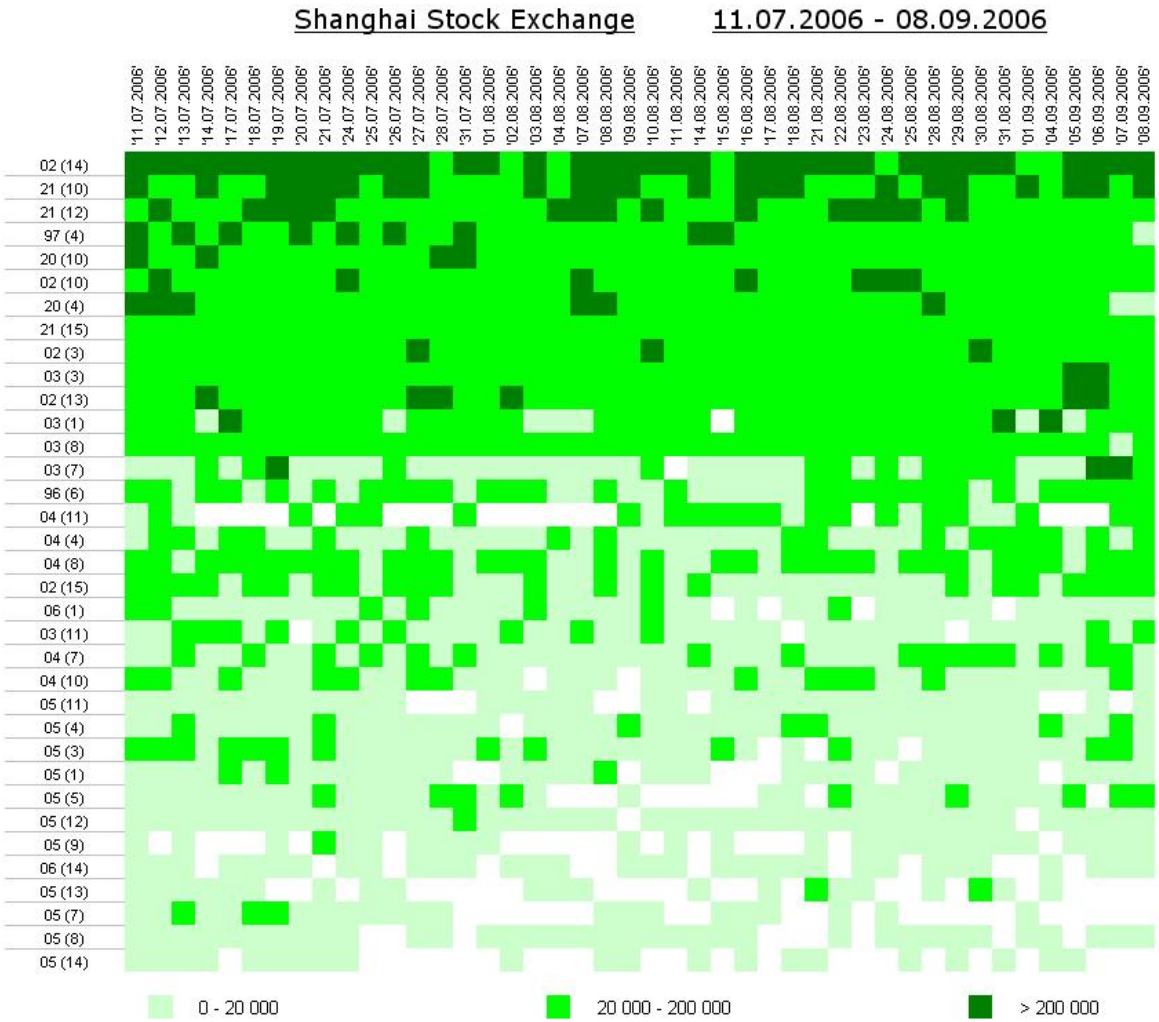


Ликвидность рынка Shanghai SE



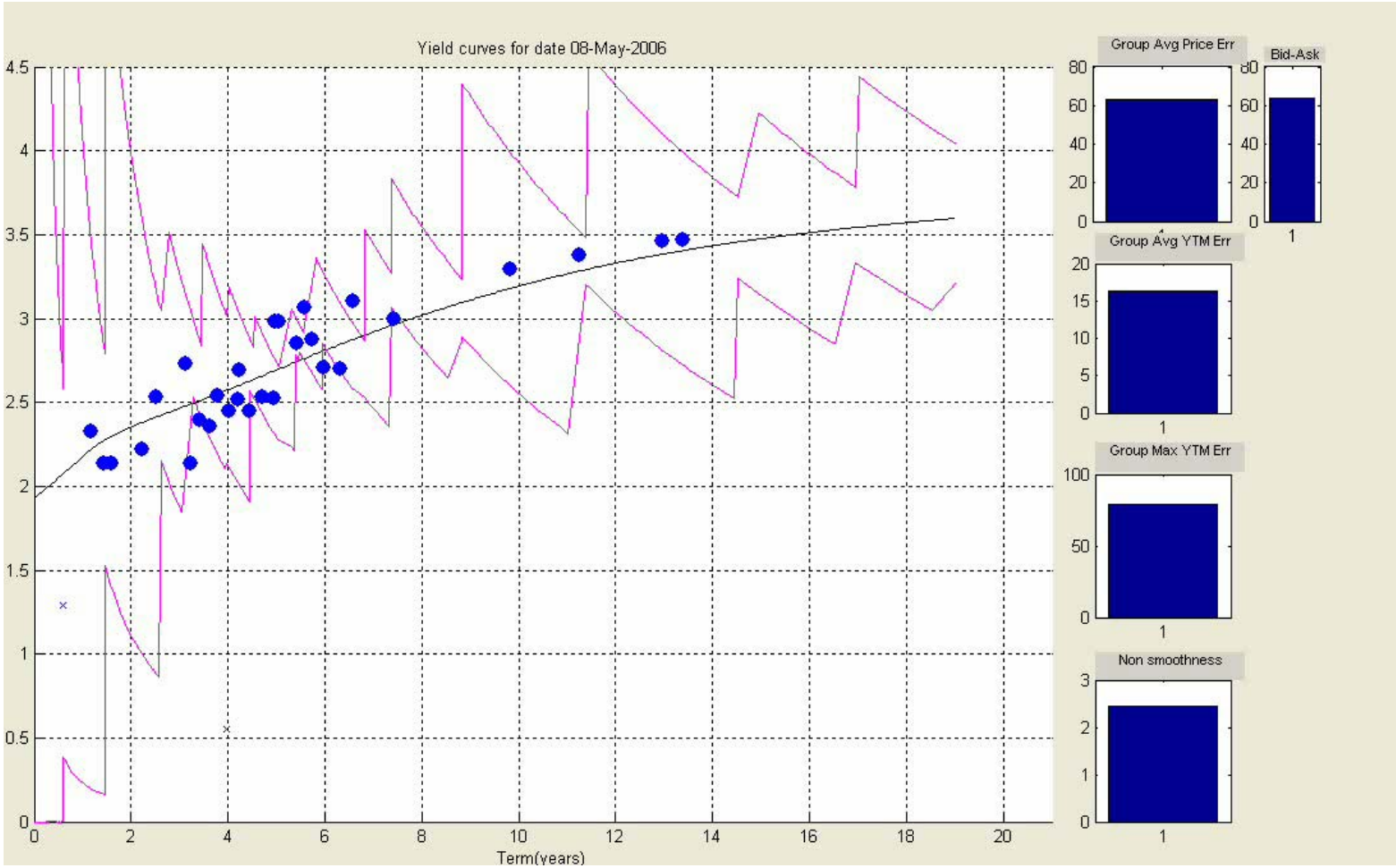
Москва 18 июня 2007 г.

Ликвидность рынка Shanghai SE



Москва 18 июня 2007 г.

Китайский рынок облигаций



Москва 18 июня 2007 г.

Модель динамики цен

Относительные (логарифмические) цены облигаций удовлетворяют стохастическому дифференциальному уравнению:

$$dS_t = \mu dt + \gamma dW_t,$$

$$W_t = (W_t^1, W_t^2, \ldots, W_t^k)^T$$

$W_t^1, W_t^2, \ldots, W_t^k$ - независимые стандартные винеровские процессы

$\mu \in R^k, \sigma \in R^{k \times k}$ - параметры процесса

Проблема: как оценивать параметры при наличии пропущенных данных?

Байесовский подход

Апостериорная плотность

Априорная плотность

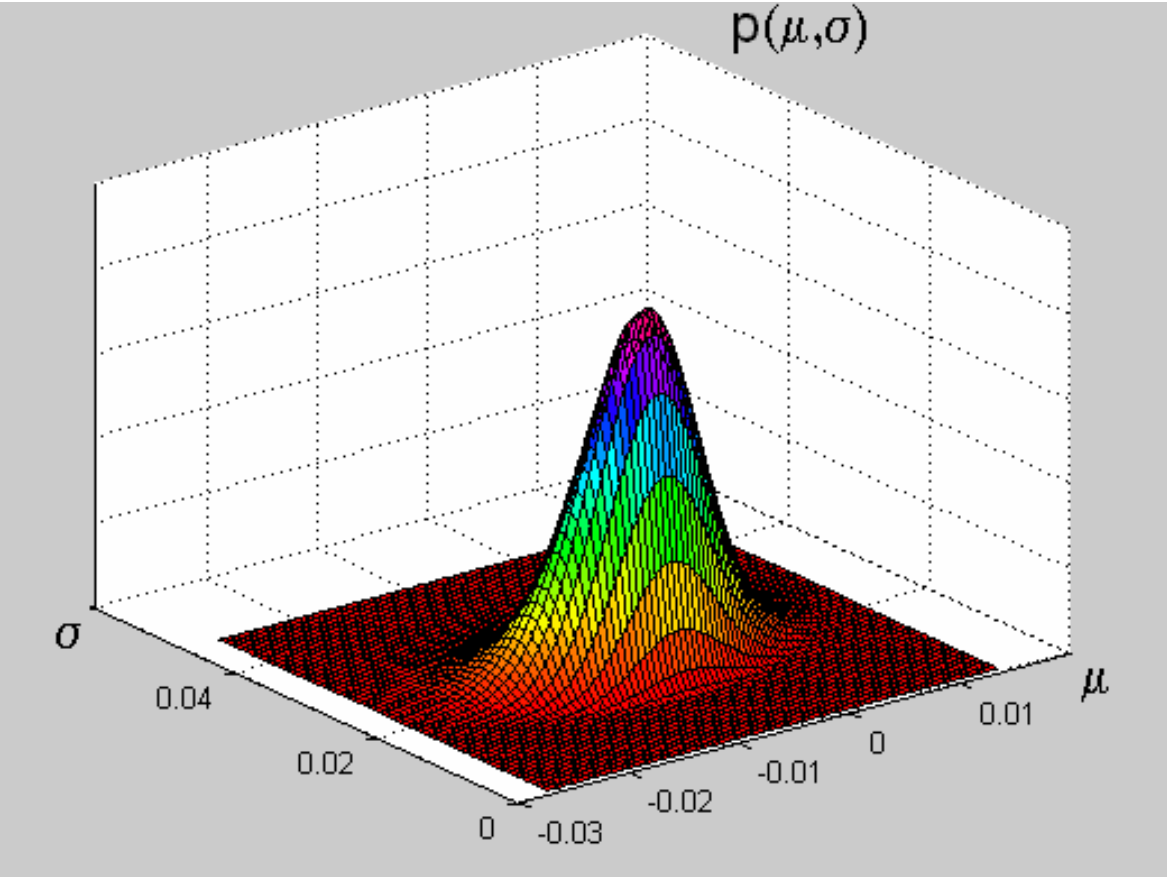
$$p(\theta \mid X) \propto p(\theta)\, p(X \mid \theta)$$

Функция правдоподобия

θ - параметры (случайные величины)

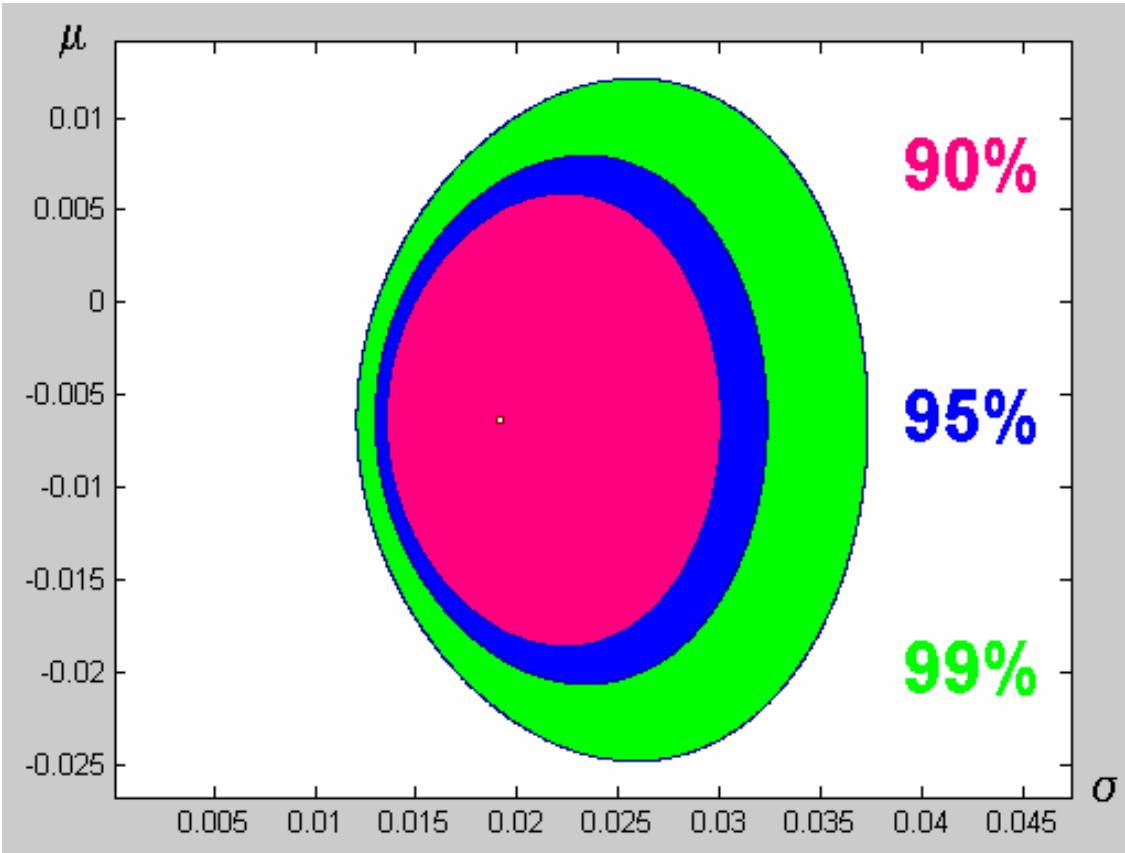
X - наблюдаемые данные

Совместное апостериорное распределение тренда и волатильности



Москва 18 июня 2007 г.

Доверительные области максимального правдоподобия



Байесовское оценивание – случай полных данных

Выборка из многомерного нормального распределения с неизвестным вектором средних и матрицей ковариаций

↓

Сопряжённые априорные распределения

↓

Аналитические решения; высокая скорость вычислений

Байесовское оценивание – случай неполных данных

Непостоянная размерность наблюдений (в зависимости от количества наблюдаемых цен)

↓

Отсутствие сопряжённых априорных семейств

↓

Численные решения; низкая скорость

Методы Markov Chain Monte-Carlo

$$X_{all} = (X_{obs}, X_{mis})$$

X_{all} - все данные (наблюдаемые+отсутствующие)

X_{obs} - наблюдаемые данные

X_{mis} - отсутствующие данные

$p(\theta \mid X_{obs})$ - сложное распределение

$p(\theta \mid X_{obs}, X_{mis})$ - простое распределение

Методы Markov Chain Monte-Carlo

Imputation Step

Генерация отсутствующих данных в наблюдениях

$$X_{mis}^{(t)} \sim p(X_{mis} \mid X_{obs}, \theta^{(t-1)})$$

Posterior Step

Генерация параметров из апостериорного распределения

$$\theta^{(t)} \sim p(\theta \mid X_{obs}, X_{mis}^{(t)})$$

Марковская цепь

$$(X_{mis}^{(1)}, \theta^{(1)}), (X_{mis}^{(2)}, \theta^{(2)}), \ldots \xrightarrow{d} p(X_{mis}, \theta \mid X_{obs})$$

EM алгоритм

Пополнение данных путём заполнения пропусков их условными математическими ожиданиями

$$y_{ij}^{old} = \begin{cases} y_{ij}, \ \ y_{ij} \in y_i^o, \\ E\left(\sum_{i=1}^{n} y_{ij} \middle| y^o, \theta^{old} \right), \ \ y_{ij} \in y_i^u; \end{cases} \quad c_{ijk}^{old} = \begin{cases} 0, \ \ y_{ij} \in y_i^o \ \ u \ \ y_{ik} \in y_k^o, \\ \operatorname{cov}\left(y_{ij}, y_{ik} \middle| y^o, \theta^{old} \right), \ \ \textit{в противном случае.} \end{cases}$$

Перерасчет мод совместного апостериорного распределения параметров

$$\mu_j^{new} = \frac{1}{n} \sum_{i=1}^{n} y_{ij}^{old}, \ \ j = 1, \ldots, d$$

$$\sigma_{jk}^{new} = \frac{1}{n} \left(\sum_{i=1}^{n} y_{ij}^{old} \cdot y_{ik}^{old} + c_{ijk}^{old} \right) - \mu_j^{new} \mu_k^{new}, \ \ j, k = 1, \ldots, d$$

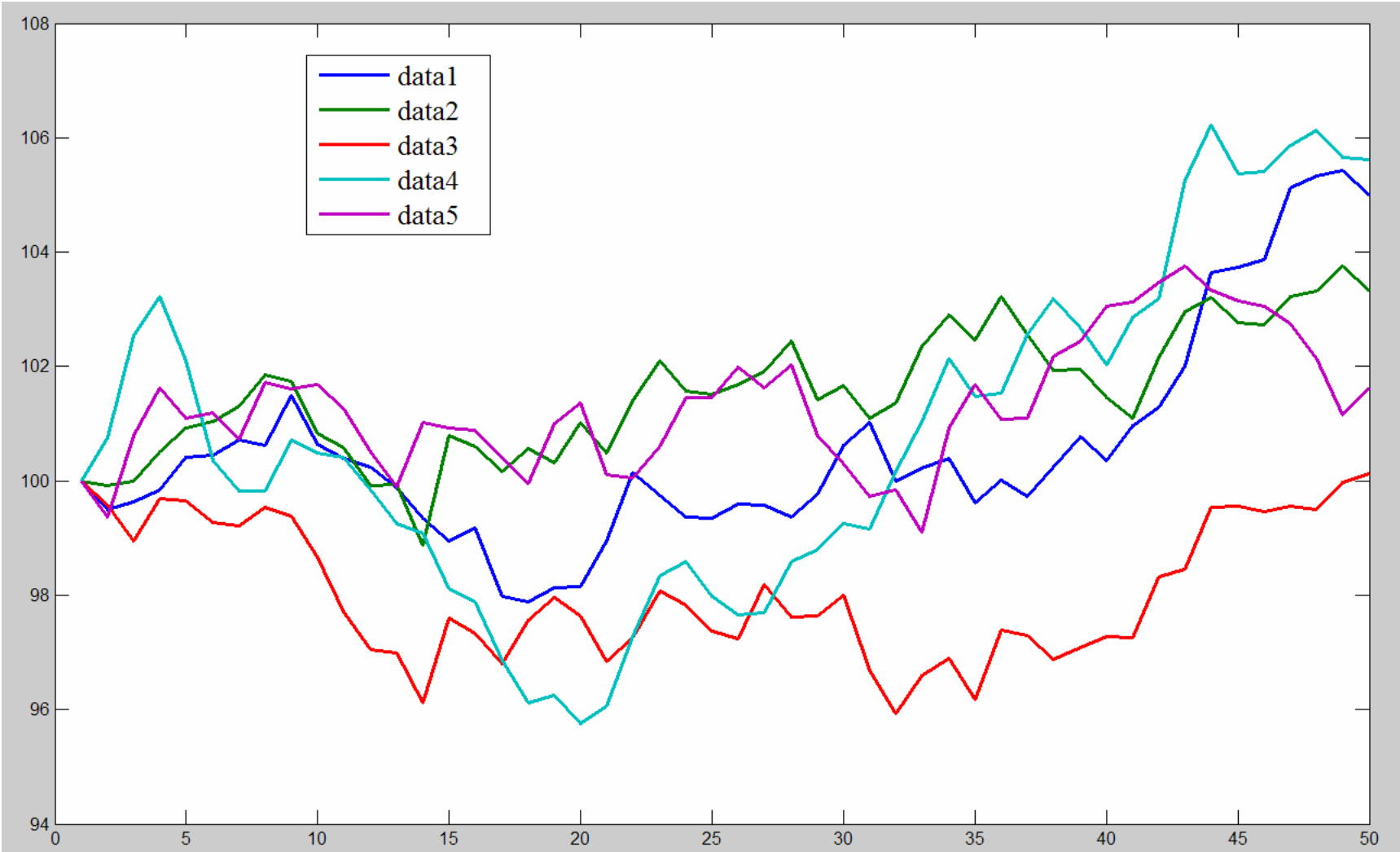
Моды совместного апостериорного распределения параметров и пропущенных данных

Тестовый пример

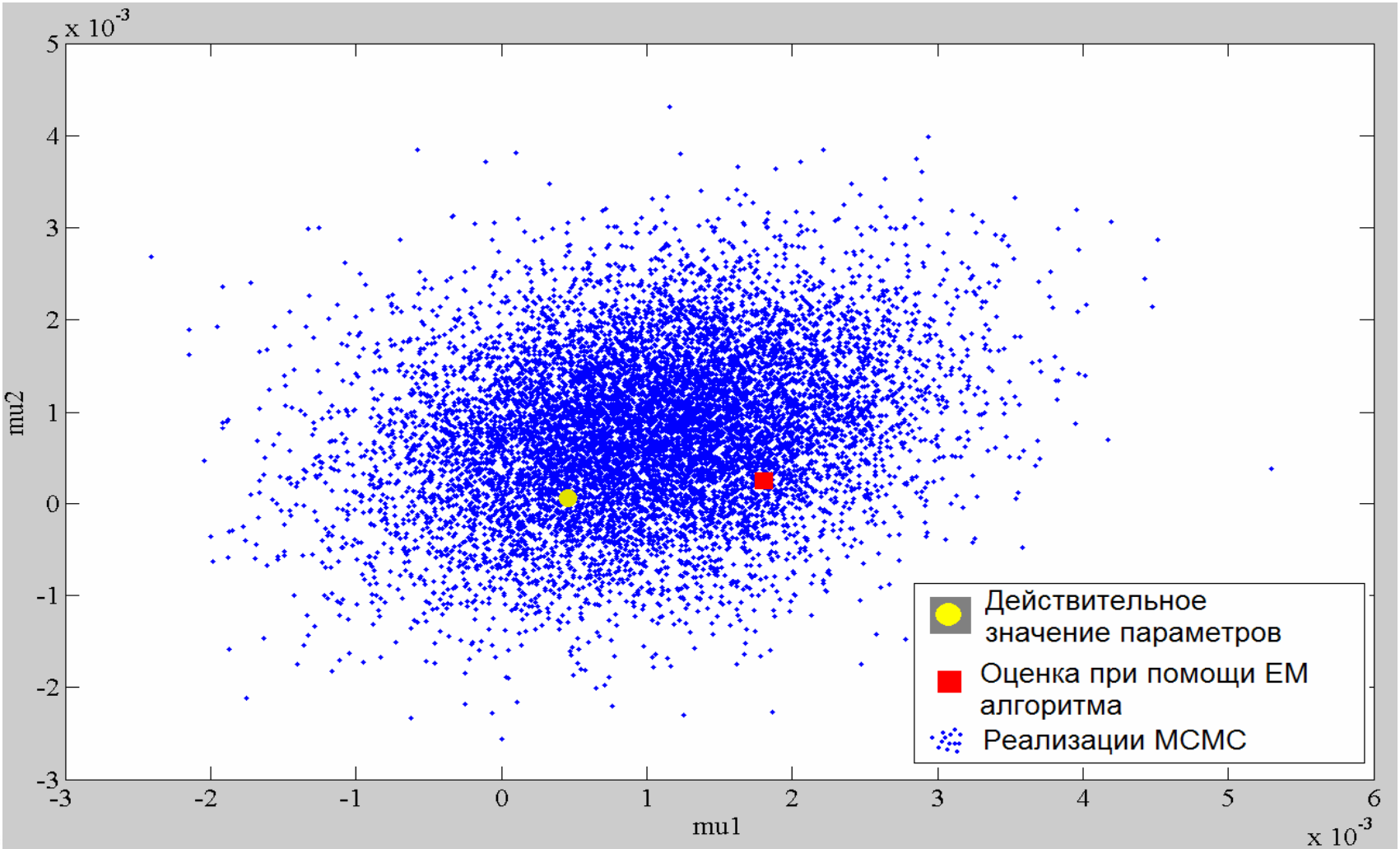
$$\mu = \begin{pmatrix} 0{,}0008 \\ 0 \\ 0{,}0004 \\ 0{,}0002 \\ 0{,}001 \end{pmatrix}, \quad \sigma^2 = \begin{pmatrix} 5 & 0 & -0{,}3 & 2 & 1 \\ 0 & 5 & 4 & 0{,}5 & 0 \\ -0{,}3 & 2 & 5 & 0 & -0{,}25 \\ 2 & 0{,}5 & 0 & 6 & 0 \\ 1 & 0 & -0{,}25 & 0 & 5 \end{pmatrix} \cdot 10^{-5}$$

- 50 случайных векторов.
- 30% случайно распределенных пропусков данных.
- Горизонт прогнозирования 10.
- 10 000 итераций Markov Chain Monte-Carlo.
- Условный портфель по 1 каждого актива.

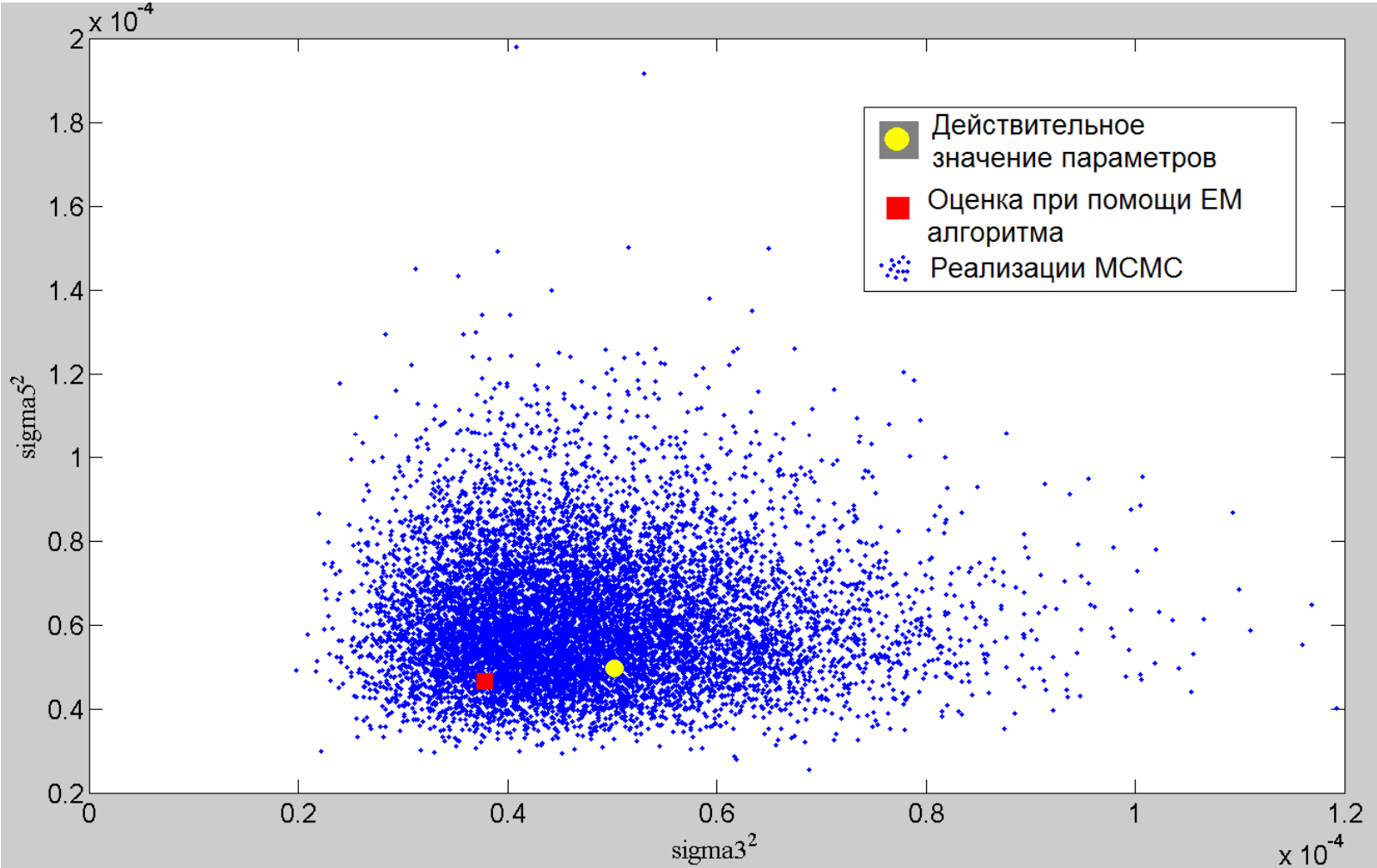
Тестовый пример



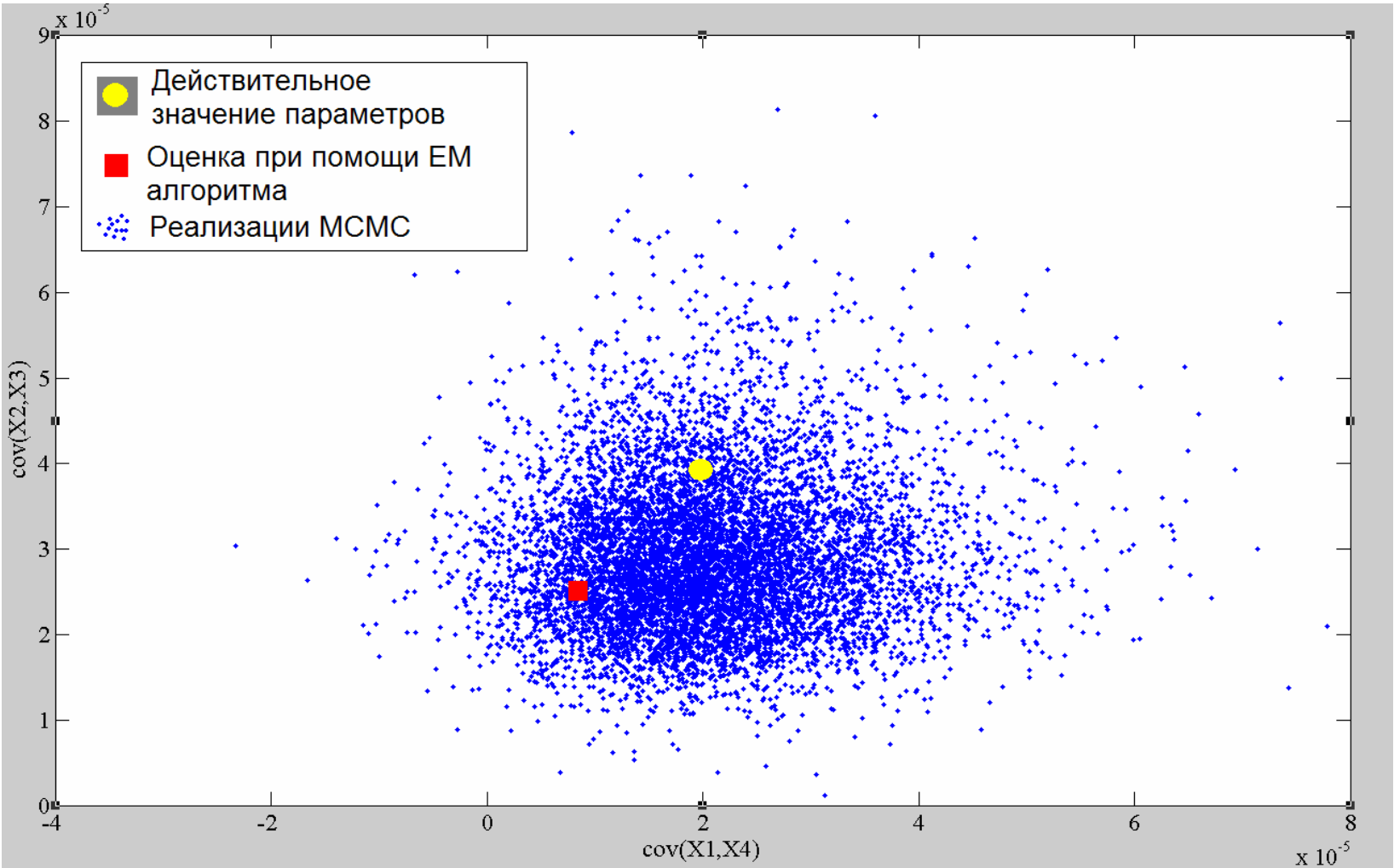
Результаты оценки параметров распределения (средние значения)



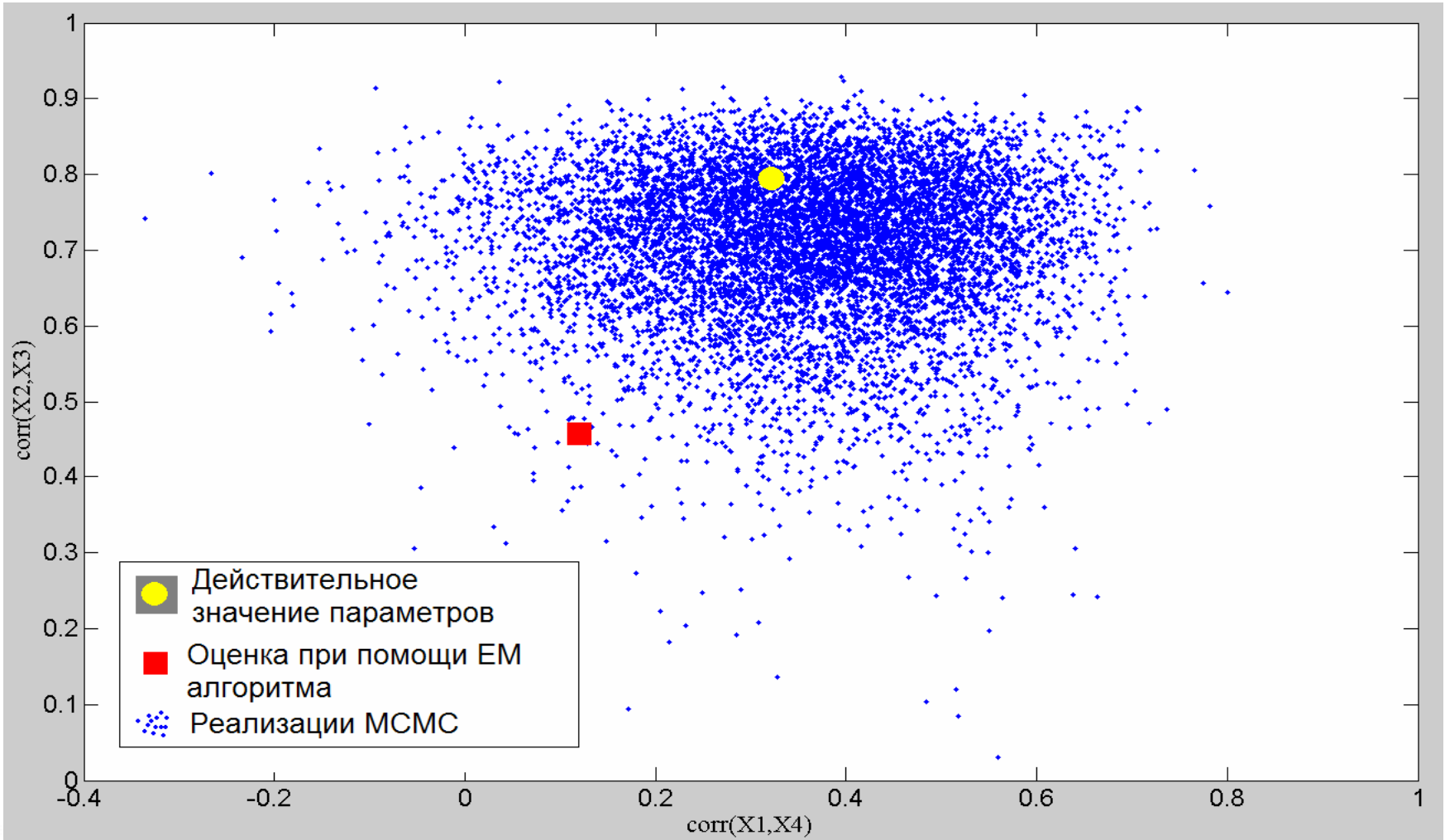
Результаты оценки параметров распределения (дисперсии)



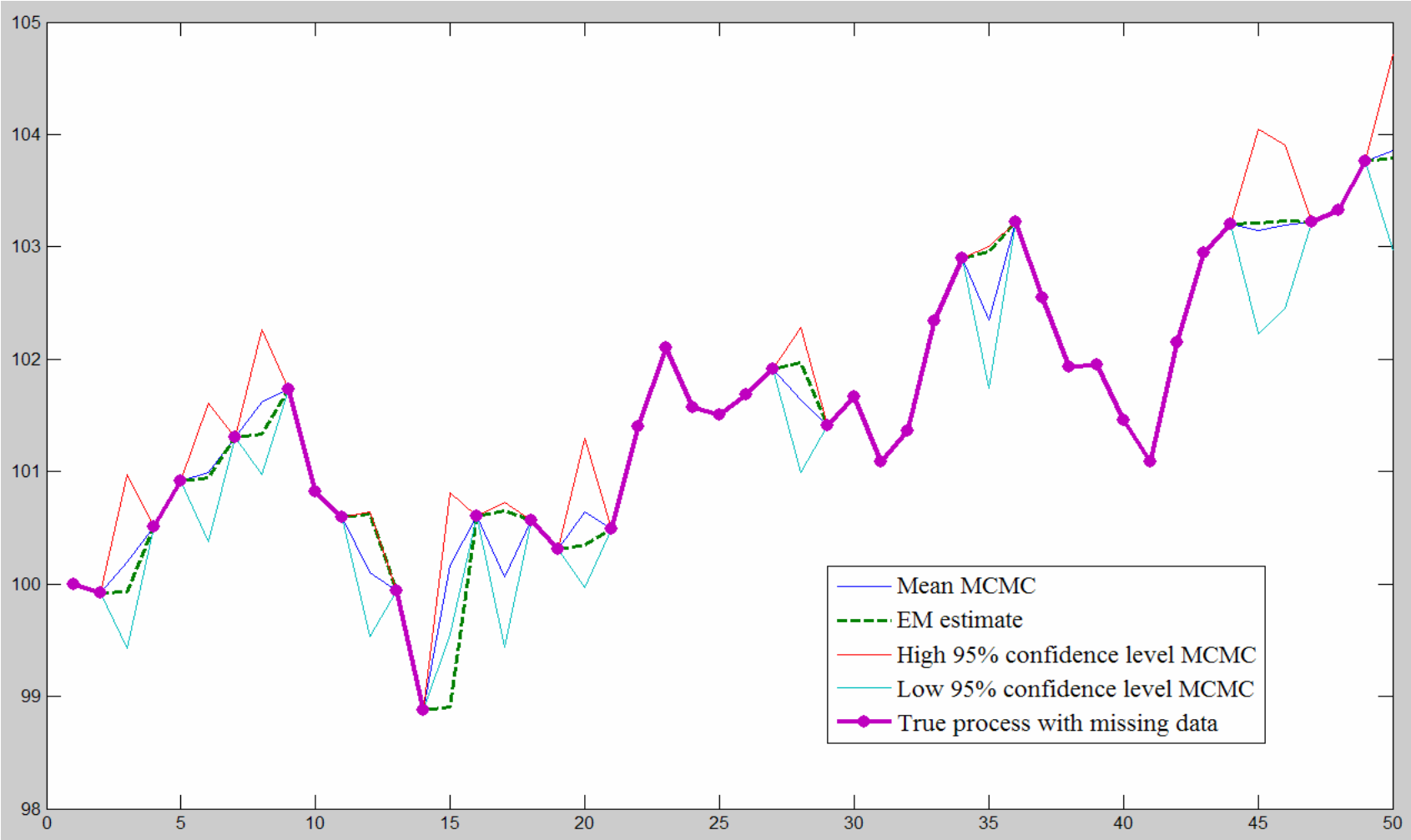
Результаты оценки параметров распределения (коэффициенты ковариации)



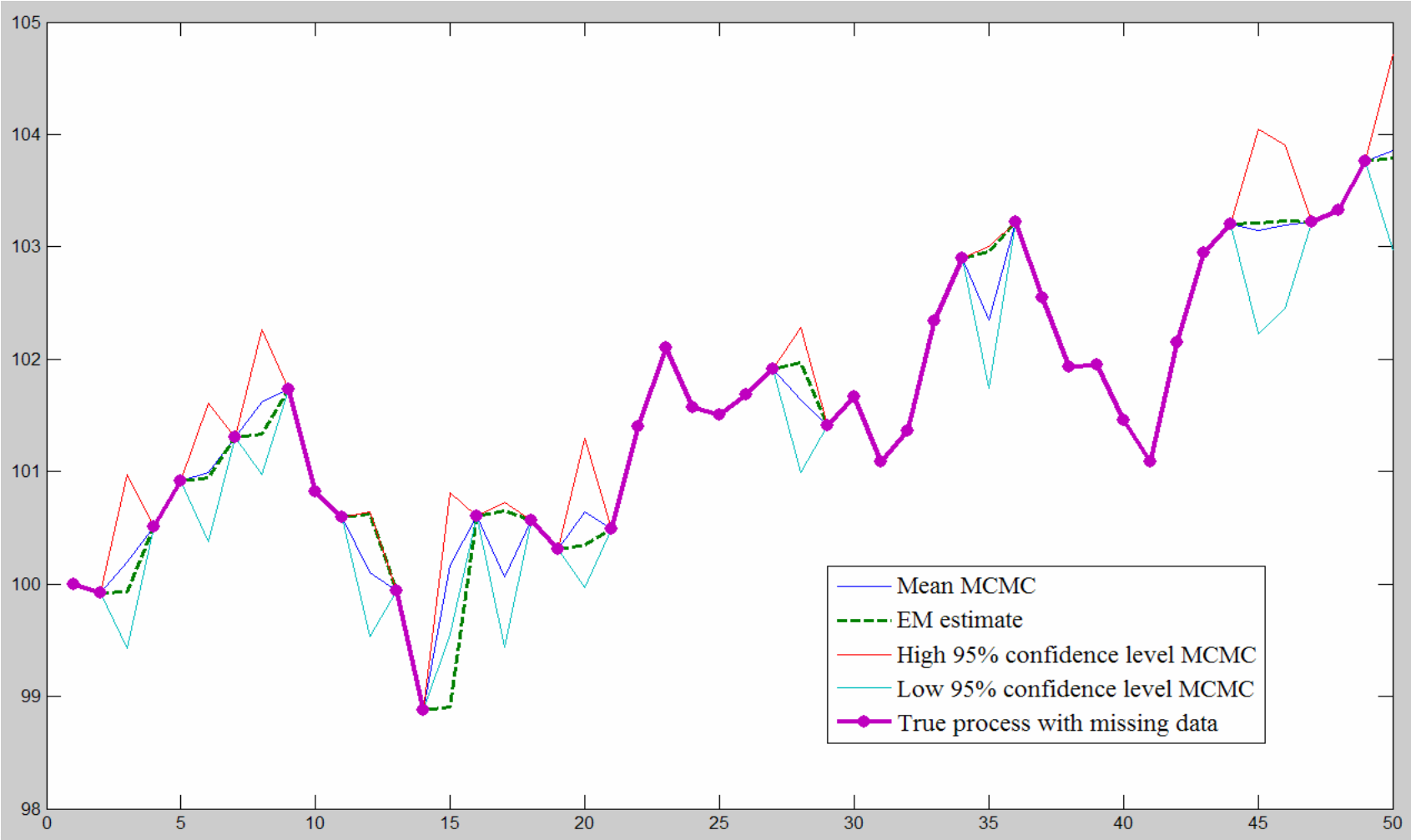
Результаты оценки параметров распределения (коэффициенты корреляции)



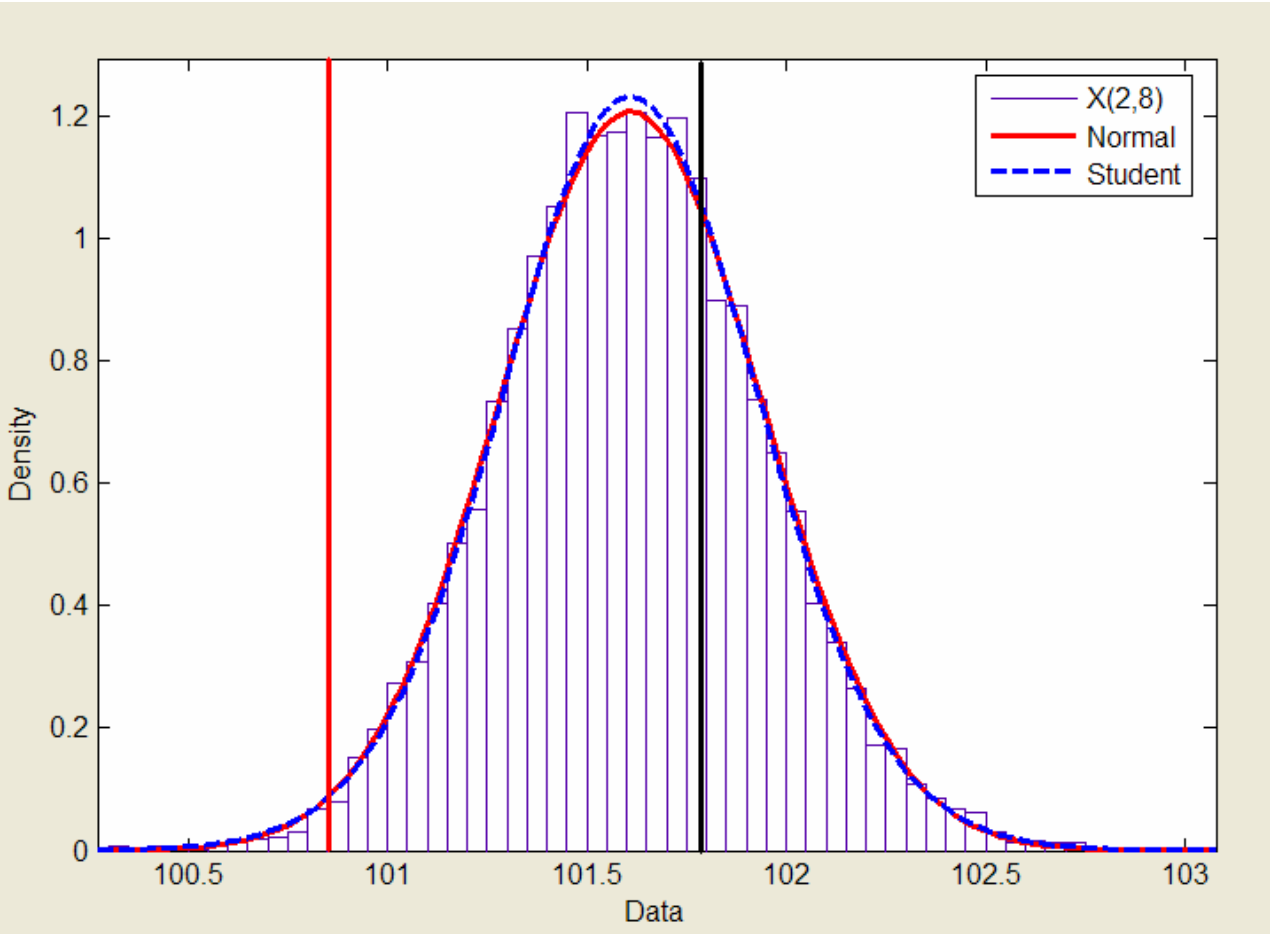
Результаты восстановления значений процесса



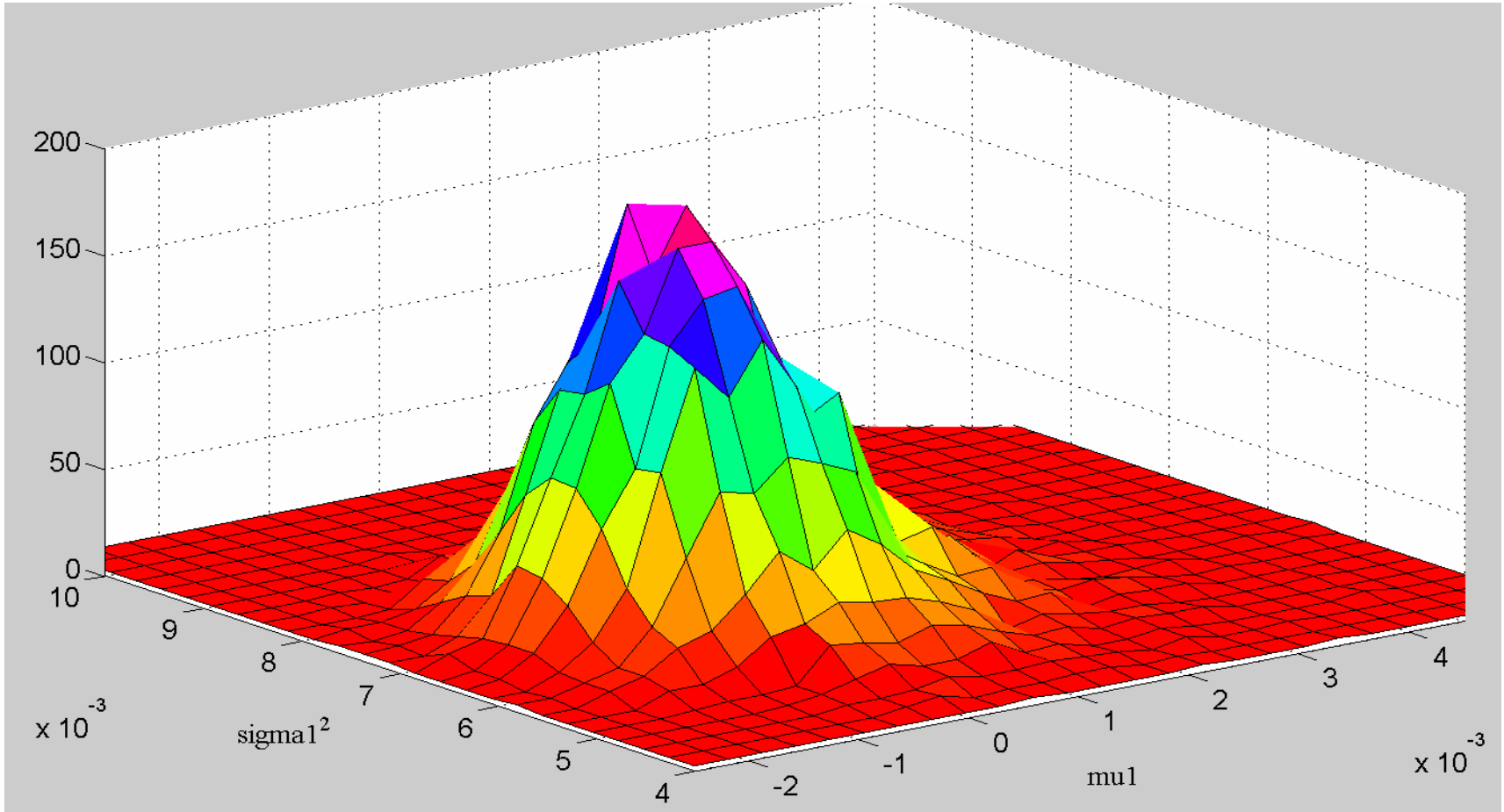
Результаты восстановления значений процесса



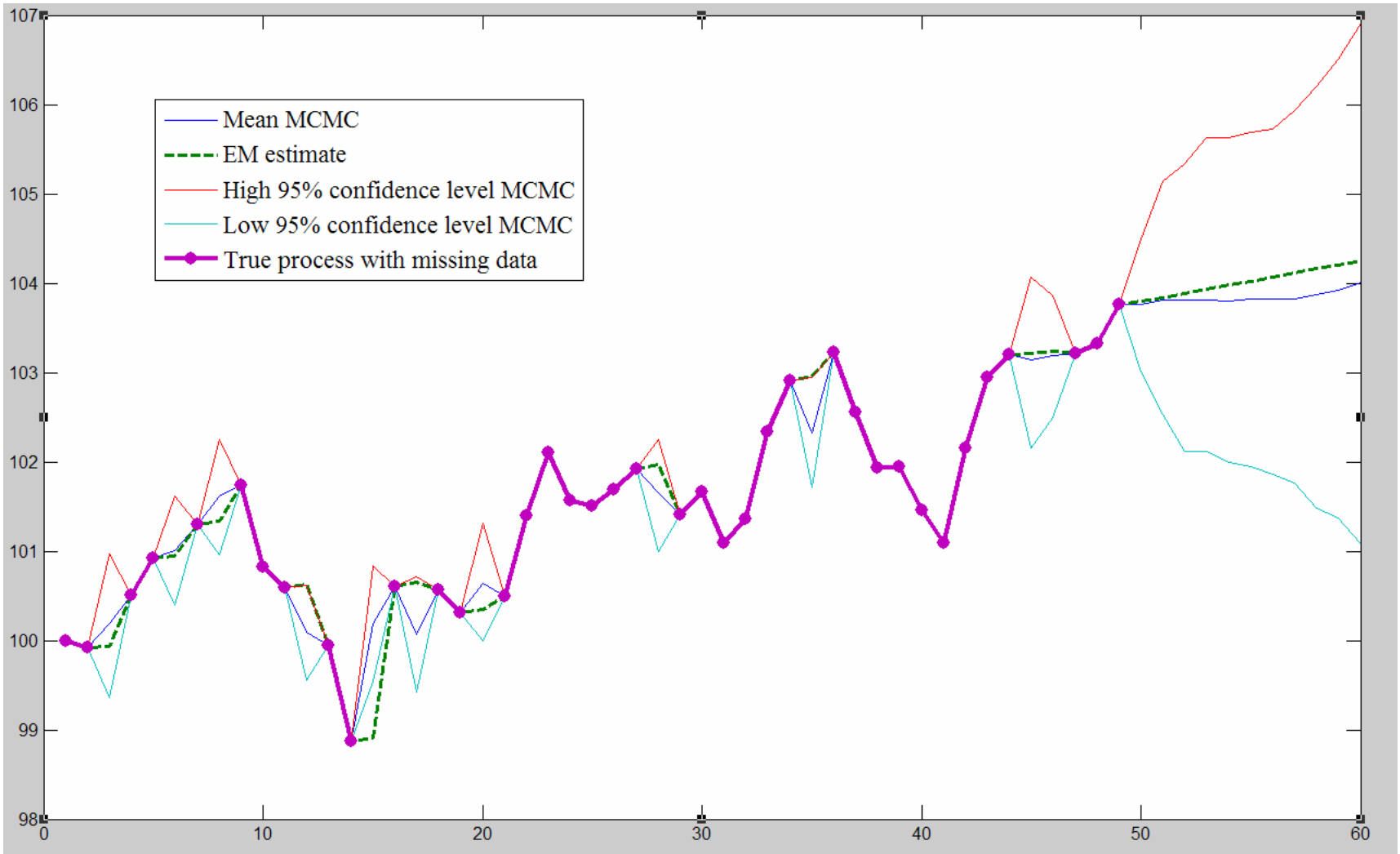
Результаты восстановления пропуска данных



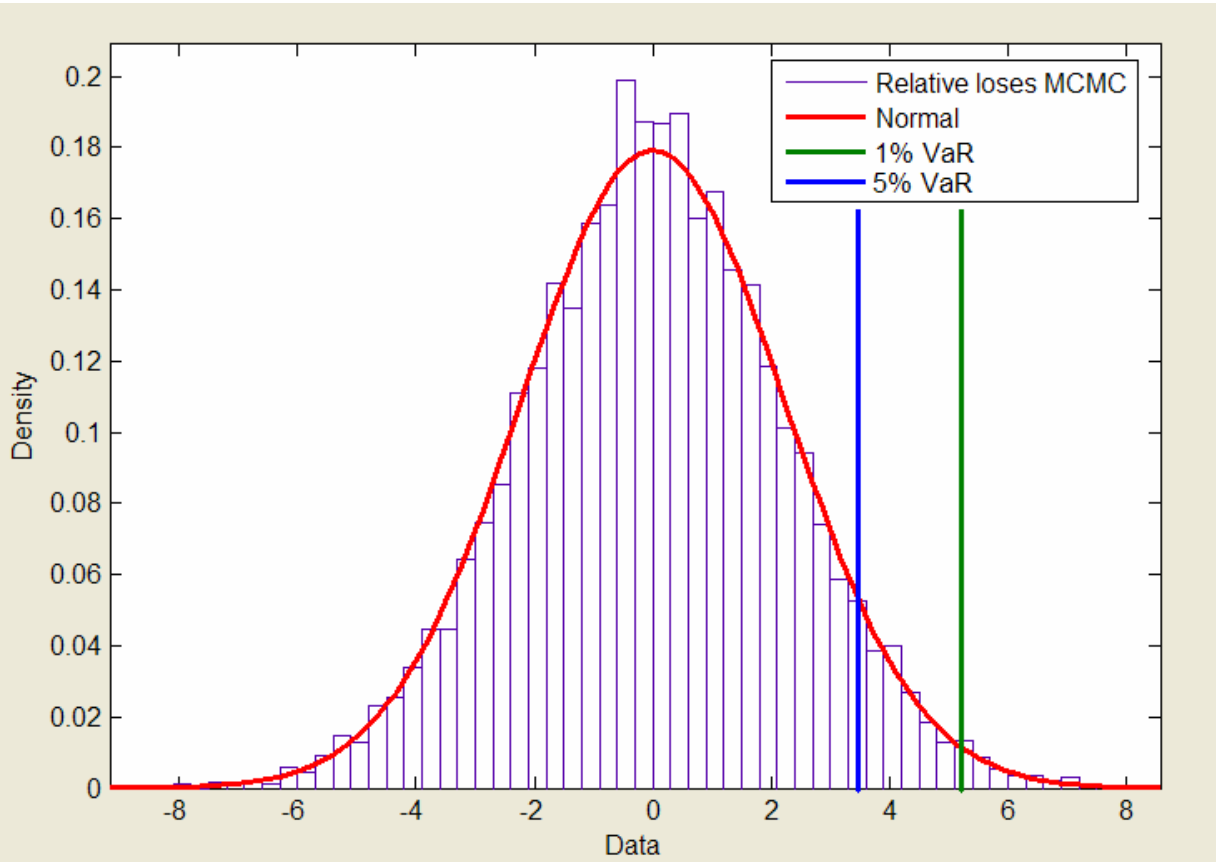
Эмпирическая плотность совместного апостериорное распределение тренда и волатильности



Применение методов восстановления данных для прогнозирования цен



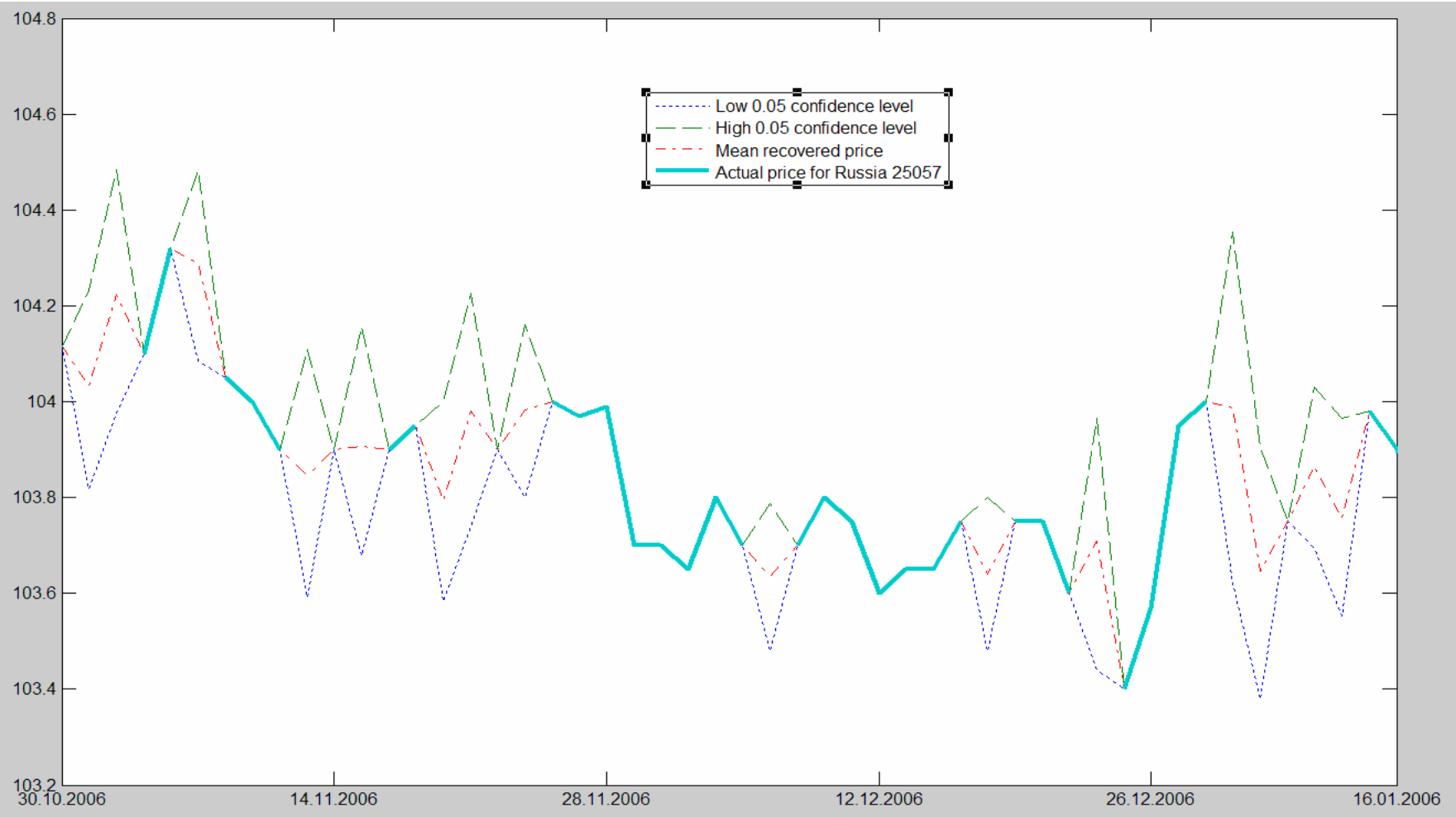
Распределение прогнозируемых относительных потерь портфеля



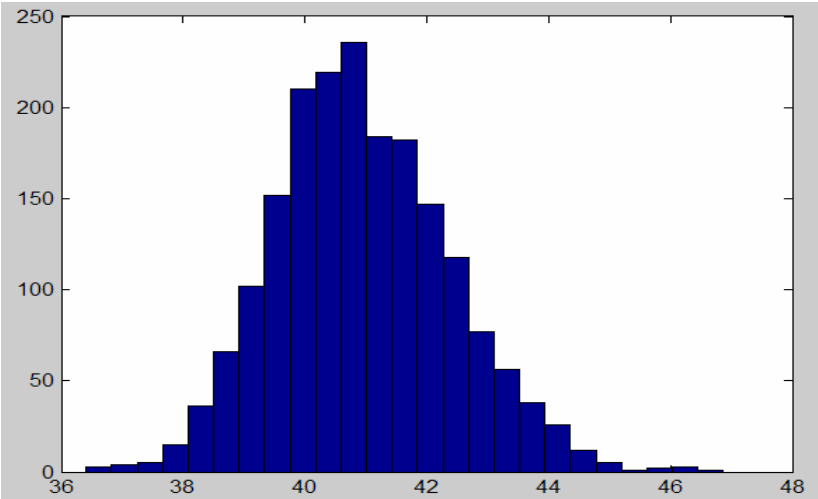
Ожидаемая стоимость портфеля и уровни соответствующие различным уровням значимости VaR

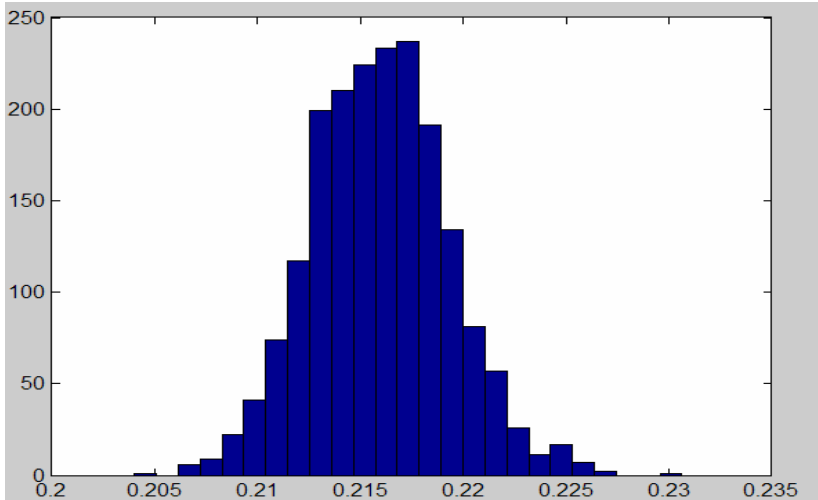
	Real process			Markov Chain Monte-Carlo				EM алгоритм			
Уровень значимости	Ожидаемая стоимость	Относительный VaR	Квантиль стоимости портфеля	Ожидаемая стоимость	Относительный VaR	Квантиль стоимости портфеля	Частота превышения потерь	Ожидаемая стоимость	Относительный VaR	Квантиль стоимости портфеля	Частота превышения потерь
5%	515,95	3,45	512,47	516,56	3,68	512,88	0,0676	516,12	2,95	513,16	0,0873
1%	515,95	4,86	511,10	516,56	5,26	511,29	0,0101	516,12	4,15	511,88	0,0253

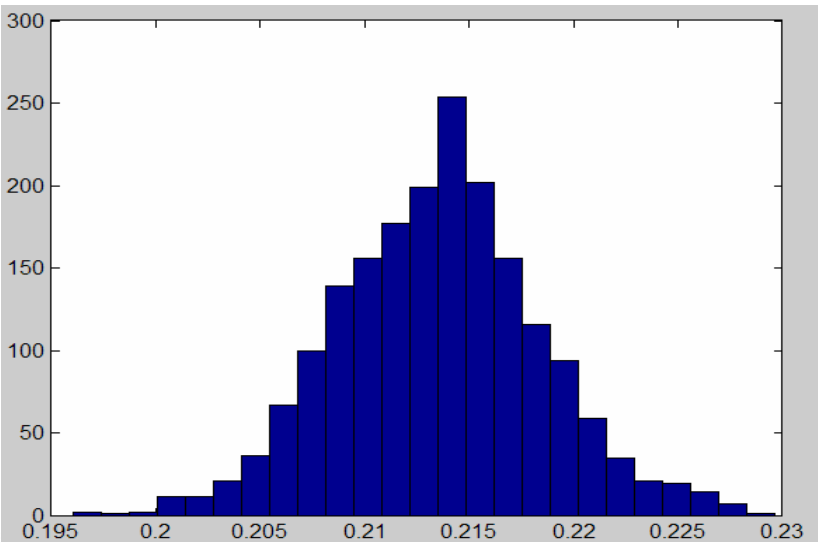
Результаты восстановления цены облигации

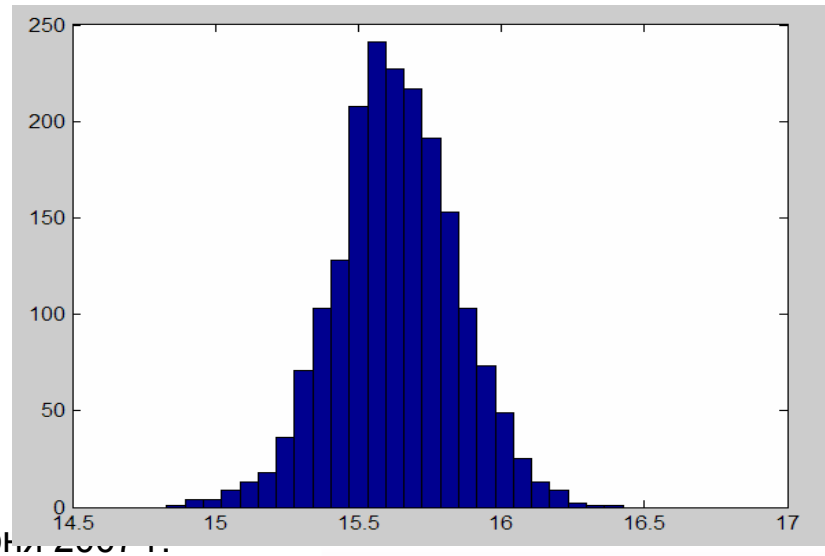
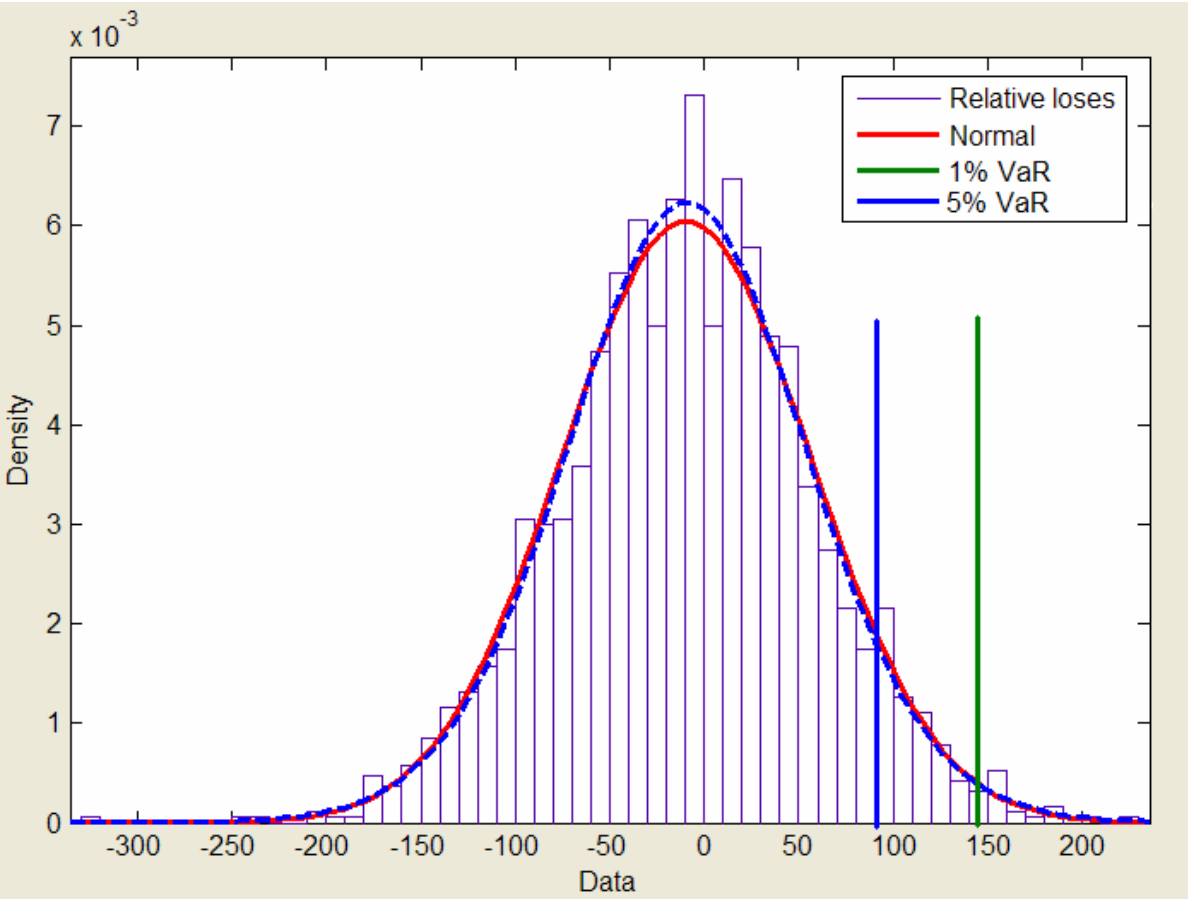


Результаты восстановления пропущенных значений цен акций









Распределение прогнозируемых относительных потерь портфеля из акций российского рынка