



# ПРОГНОЗИРОВАНИЕ ФИНАНСОВОГО СОСТОЯНИЯ КРЕДИТНЫХ ОРГАНИЗАЦИЙ НА ОСНОВЕ СИНГУЛЯРНО-СПЕКТРАЛЬНОГО АНАЛИЗА

Докладчик: Приказчикова Анастасия

Научный руководитель: Иванов Виктор Владимирович

# Цель и задачи исследования

- **Цель научной работы** – разработка методики оценки и прогнозирования финансового состояния кредитных организаций с точки зрения ПОД/ФТ
- **Задача** – провести анализ показателей деятельности кредитных организаций с использованием метода «Гусеница»-SSA

# Обзор предметной области. Кредитные организации

## ЭТАПЫ ЛЕГАЛИЗАЦИИ ПРЕСТУПНЫХ ДОХОДОВ



# Обзор предметной области. Кредитные организации



ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ФИНАНСОВОМУ  
МОНИТОРИНГУ



# Этапы работы с данными





ВЫБОРКА БАНКОВ

# Ранжирование банков

< Изменить

Банков

345

345

Сохранить настройки

В список

Отчет

№	Наименование	Регистрационный номер	Руководство	Банковский идентификационный номер	Регион регистрации	2022 - 1мес., Чистые активы, RUB ↑	2022 - 1мес., Счета в Ба
---	--------------	-----------------------	-------------	------------------------------------	--------------------	------------------------------------	--------------------------

<input type="checkbox"/>	1	<a href="#">СБЕРБАНК, ПАО</a>	1481	Греф Герман Оскарович	044525225	Москва	40 419 030 693 000	1
<input type="checkbox"/>	2	<a href="#">БАНК ВТБ (ПАО)</a>	1000	Костин Андрей Леонидович	044525187	Санкт-Петербург	20 546 524 318 000	8
<input type="checkbox"/>	3	<a href="#">БАНК ГПБ (АО)</a>	354	Акимов Андрей Игоревич	044525823	Москва	8 983 078 098 000	4
<input type="checkbox"/>	4	<a href="#">АЛЬФА-БАНК, АО</a>	1326	Соколов Андрей Борисович	044525593	Москва	6 269 550 359 000	4
<input type="checkbox"/>	5	<a href="#">НКЦ (АО), НКО</a>	3466	Марич Игорь Леонидович	044525646	Москва	6 222 149 048 000	
<input type="checkbox"/>	6	<a href="#">РОССЕЛЬХОЗБАНК, АО</a>	3349	Листов Борис Павлович	044525111	Москва	4 420 953 502 000	:
<input type="checkbox"/>	7	<a href="#">МОСКОВСКИЙ</a>	1978		044525659	Москва	3 707 421 107 000	:



списки



выборки



сервисы



события



Справка

Главная Вставка Разметка страницы Формулы Данные Рецензирование Вид Разработчик

Вставить Буфер обмена Шрифт Выравнивание Число

Общий \$ % 000

Условное форматирование Форматировать как таблицу Стиль ячеек

Вставить Удалить Формат

Сортировка и фильтр Редактирование

H88 fx 24867244000

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	№	bank_name	register_num	Отзыв	Причина	region	niy_cap	net_as	account	accou	le secu	dits am	ompani	ndividua	banks
2	151	БАЛТИЙСКИЙ БАНК, ПАО	128	1	ЛИКВ.	Санкт-Петербург	10000001	7,84E+10	1E+08	5,53E+08	0	5,53E+10	9,39E+09	27527000	3,99E+10
3	1557	СОЦИНВЕСТБАНК, АО	1132	1	ЛИКВ.	Башкортостан (Респ	10000001	5,89E+10	2,5E+08	48083000	8,42E+09	3,53E+10	2,71E+10	2,17E+09	5,52E+09
4	986	МПСБ, ПАО КБ	752	1	ЛИКВ.	Мордовия (Республ	4,85E+08	3,85E+09	4,84E+08	19334000	16023000	2,28E+09	2,15E+09	1,24E+08	6257000
5	239	БАНКХАУС ЭРБЕ (АО)	1717	1	ЛИКВ.	Москва	4,23E+08	3,23E+09	1,18E+08	1,72E+09	6,72E+08	4,78E+08	1,46E+08	0	3,32E+08
6	1370	РОСКОМСНАББАНК (ПАО)	1398	1	Эконом.	Башкортостан (Респ	9,21E+08	2,96E+10	9,2E+08	28192000	6003000	2,01E+10	1,92E+10	7,82E+08	8000000
7	695	КЕМСОЦИНБАНК, АО	96	1	Эконом.	Кемеровская област	90000000	2,63E+09	4,54E+08	480000	0	1,03E+09	8,79E+08	1,39E+08	1,7E+08
8	531	ЖИЛСТРОЙБАНК, АО КБ	2769	1	ЛИКВ.	Москва	1,1E+08	3,7E+08	2,33E+08	475000	0	92271000	35754000	56517000	
9	327	ВЗАИМОДЕЙСТВИЕ, ООО КБ	1704	1	Эконом.	Новосибирская обл	2,35E+08	6,45E+08	65622000	50201000	0	1,63E+08	91137000	66763000	2739000
10	564	ИВАНОВО, АО КБ	1763	1	Эконом.	Ивановская область	1,43E+08	3,38E+09	1,49E+08	37484000	0	2,49E+09	1,4E+09	1,1E+09	1348000
11	158	БАНК АВБ, АО	23	1	ЛИКВ.	Самарская область	3,5E+08	3,68E+11	3001000	6331000	7,67E+09	3,47E+11	2,29E+11	9,13E+09	3,53E+10
12	1229	ПРАЙМ ФИНАНС (АО), БАНК	2758	1	Эконом.	Санкт-Петербург	58196000	1,61E+09	12231000	5900000	0	1,05E+09	6,44E+08	3,1E+08	6411000
13	1796	ХОЛДИНВЕСТБАНК, ООО КБ	2837	1	Эконом.	Московская область	2,28E+08	8,49E+08	20140000	5270000	0	8,05E+08	7,05E+08	369000	1E+08
14	884	МЕЖДУНАРОДНЫЙ РАСЧЕТНЫ	3028	1	ОД	Москва	2,4E+08	1,24E+09	3,17E+08	26147000	1,77E+08	5,15E+08	2,67E+08	47702000	2E+08
15	680	КБ ДЕЛЬТАКРЕДИТ, АО	3338	1	ликв.	Москва	3,24E+09	2,01E+11	2,22E+09	2,05E+08	0	1,04E+11	1,15E+11	1,55E+11	1,28E+11
16	110	АРБ-ИНКАСС, НКО ООО	3353	1	Эконом.	Москва	18260000								
17	1396	РТС-БАНК, АО	3401	1	Эконом.	Самарская область	1,75E+08								
18	128	АСПЕКТ (АО), АБ	608	1	ОД	Москва	45300000								
19	1685	ТРОЙКА-Д БАНК, АО	3431	1	Эконом.	Москва	1,17E+09								
20	1458	СБЕРБАНК, ПАО	1481	0	0	Москва	6,78E+10								
21	1189	ПЕРВЫЙ ИНВЕСТИЦИОННЫЙ Б	604	0	0	Москва	1,98E+08								
22	273	БМ-БАНК, АО	2748	0	0	Москва	1,62E+10								
23	91	АЛЬФА-БАНК, АО	1326	0	0	Москва	5,96E+10								
24	1647	ТИМЕР БАНК, АО	1581	0	0	Москва	10000001								
25	975	МОСККОММЕРЦБАНК (АО), КБ	3365	0	0	Москва	3,57E+09								

№	Показатель	№	Показатель
1	Население региона регистрации банка	13	Кредиты физическим лицам
2	Уставный капитал	14	Кредиты другим банкам (МБК)
3	Чистые активы	15	Основные средства
4	Счета в Банке России	16	Учтённые векселя
5	Коррсчета (НОСТРО)	17	Прибыль (убыток) до налогообложения
6	Ценные бумаги	18	Коррсчета (ЛОРО)
7	Кредиты (общий показатель)	19	Кредиты других банков
8	Кредиты организациям	20	Средства клиентов (физ.лица)
9	Средства организаций на расчетных счетах	21	Депозиты юридических лиц
10	Вклады физических лиц	22	Облигации
11	Векселя	23	Резервы на возможные потери
12	Капитал		

# Прошлые исследования – база настоящего исследования

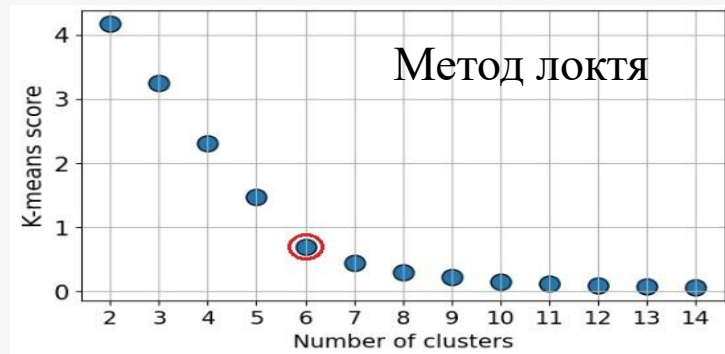
1. Многомерный анализ данных
2. Деревья и леса решений



# Формирование зоны риска для кредитных организаций

*Кластерный анализ, Метод Варда, Метод силуэтов, Метод локтя, Метод Дэвиса-Болдина, Метод главных компонент*

№ кластера	Количество благонадежных	Количество неблагонадежных	Общее количество
1	62	42	104
2	0	3	3
3	14	4	18
4	26	6	32
5	4	2	6
6	7	7	14
7	13	4	17



$$d(x_i, x_j) = \left( \sum_{p=1}^P (x_{ip} - x_{jp})^2 \right)^{1/2}$$

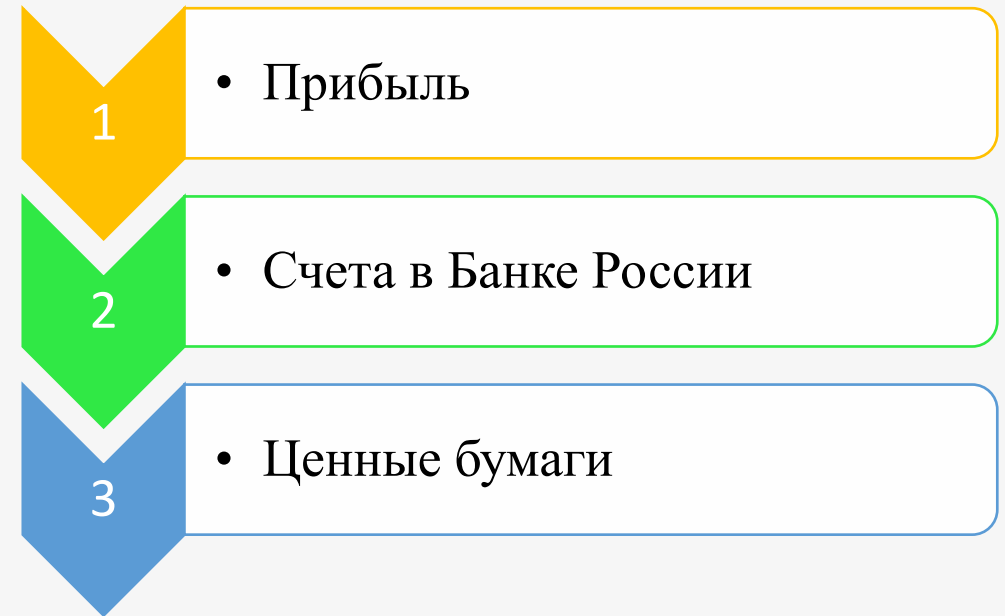
# Деревья и леса решений в задаче идентификации наиболее информативных показателей

## Дерево *CART* леса *Random Forest* *Adaboost* *Xgboost*

Метрика	CART	Random Forest	Adaboost	Xgboost
Precision	79%	93%	92%	92%
Recall	86%	64%	55%	56%
F1-score	83%	76%	69%	69%
Accuracy	85%	83%	80%	80%

Модель	Время поиска оптимальных гиперпараметров		Отношение	Кол-во обработанных объектов
	Сервер Google ( $t_1$ )	Платформа HybriLIT ( $t_2$ )		
CART	28 мин. 49 сек.	8 мин. 47 сек.	3,3	576480
Random Forest	58 мин. 6 сек.	1 мин. 14 сек.	47	3252
Adaboost	2 часа 54 мин. 39 сек.	4 мин. 5 сек.	43	2196
Xgboost	65 мин. 23 сек.	1 мин. 31 сек.	43	5388

### Наиболее информативные показатели:

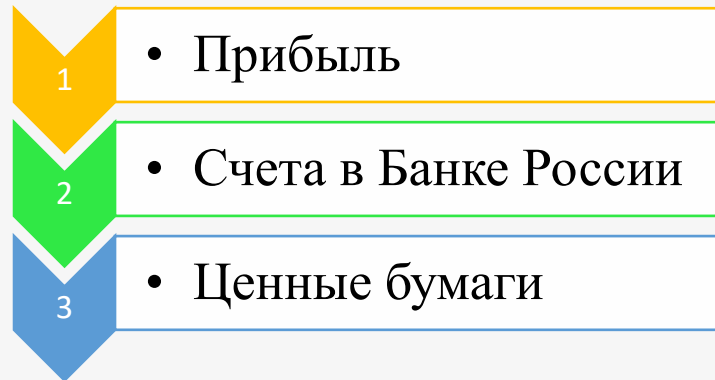


# Предварительная информация

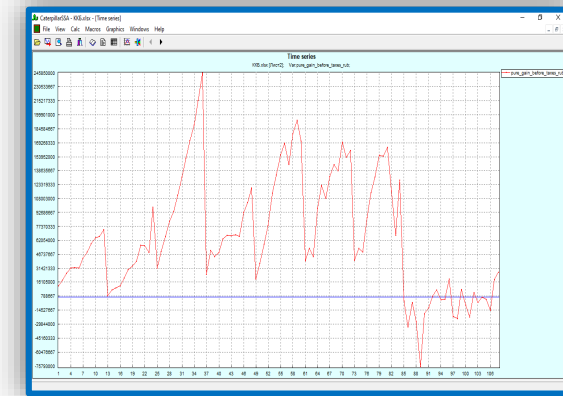
Зона риска



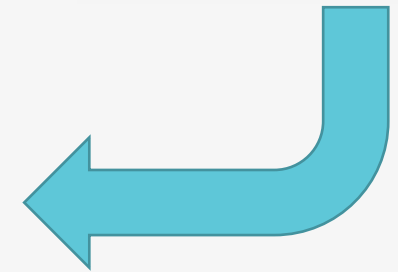
Информативные показатели



Анализ рядов



Метод «Гусеница»-SSA



# Базовый алгоритм метода «Гусеница»-SSA

- Вложение
- Сингулярное разложение

Восстановление

Разложение

- Группировка
- Диагональное усреднение



Метод «Гусеница»-SSA заключается в преобразовании одномерного ряда в многомерный при помощи сингулярного разложения траекторной матрицы, исследовании полученной многомерной траектории методом главных компонент и отборе тех компонентов, которые представляют интерес. Далее – реконструкция одномерного временного ряда на основе отобранных компонентов.

# Разложение временного ряда

## □ Вложение

Рассмотрим временной ряд  $F = (f_1, f_2, \dots, f_N)$ , где  $N$  – длина ряда. Пусть  $L$  – длина «гусеницы»,  $1 < L < N$ .

$X = [X^1 X^2 \dots X^k] = \begin{pmatrix} f_1 & f_2 & \dots & f_k \\ f_2 & f_3 & \dots & f_{k+1} \\ \dots & \dots & \dots & \dots \\ f_L & f_{L+1} & \dots & f_N \end{pmatrix}$  – траекторная матрица. На стадии сингулярного разложения траекторной матрицы она раскладывается в сумму элементарных матриц.

## □ Разложение

$$\mathbf{X} = \mathbf{X}_1 + \dots + \mathbf{X}_d \quad \mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^T \quad S = \mathbf{X} \mathbf{X}^T \quad d = \max\{i, \text{таких, что } \lambda_i > 0\} \quad V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i}$$

$\lambda_i (i = 1, \dots, L)$  – собственные числа матрицы  $S$

$\{U_1, \dots, U_d\}$  – соответствующие системы ортонормированных собственных векторов матрицы  $S$

# Восстановление временного ряда

## □ Группировка

На основе разложение процедура группировки делит все множество индексов  $\{1, \dots, d\}$

на  $m$  непересекающихся подмножеств  $I_1, \dots, I_m$   $I = \{i_1, \dots, i_p\}$   $\mathbf{X}_I = \mathbf{X}_{i_1} + \dots + \mathbf{X}_{i_p}$

Группы: тренд, периодическая и шумовые составляющие

$$\mathbf{X} = \mathbf{X}_{I_1} + \dots + \mathbf{X}_{I_m}$$

## □ Диагональное усреднение

Пусть  $\mathbf{Y}$  – матрица  $L \times K$  с элементами  $y_{ij}$   $1 \leq i \leq L$   $1 \leq j \leq K$   $N = L + K - 1$   $L^* = \min(L, K)$

$$K^* = \max(L, K)$$

$$g_k = \begin{cases} \frac{1}{k+1} \sum_{m=1}^{k+1} y_{m, k-m+2}^* \text{ для } 0 \leq k \leq L^* - 1, \\ \frac{1}{L^*} \sum_{m=1}^{L^*} y_{m, k-m+2}^* \text{ для } L^* - 1 \leq k \leq K^*, \\ \frac{1}{N-k} \sum_{m=k-K^*+2}^{N-K^*+2} y_{m, k-m+2}^* \text{ для } K^* - 1 \leq k \leq N. \end{cases}$$

– диагональное усреднение  
переводит матрицу  $\mathbf{Y}$  в ряд

$$g_0, \dots, g_{N-1}$$

Пусть  $y_{ij}^* = y_{ij}$ , если  $L < K$   
 $y_{ij}^* = y_{ji}$ , иначе

# Прогнозирование на точку вперед

$$\tilde{x}_{N+p} = \sum_{k=1}^{L-1} a_k x_{N+1-k}$$

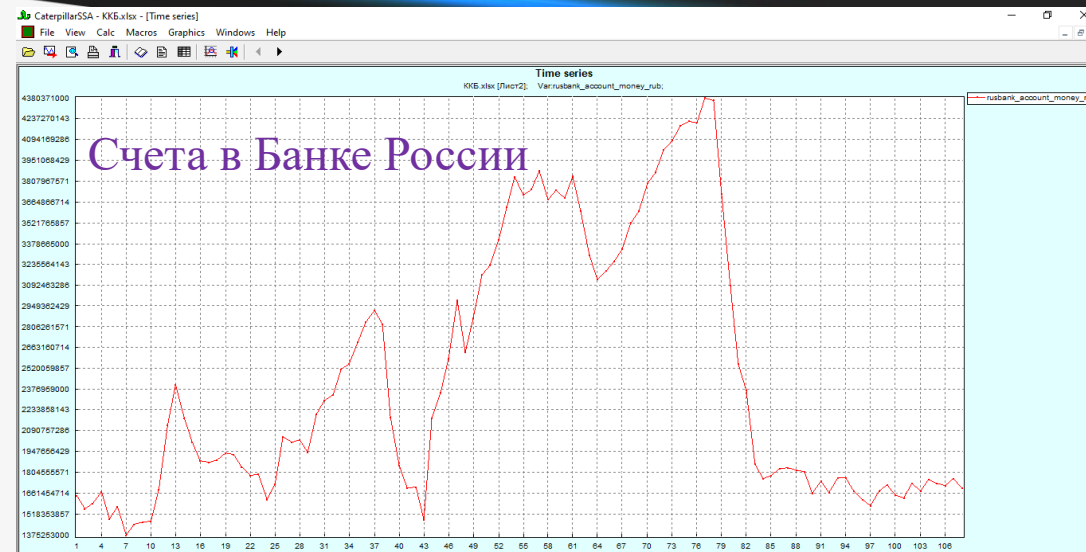
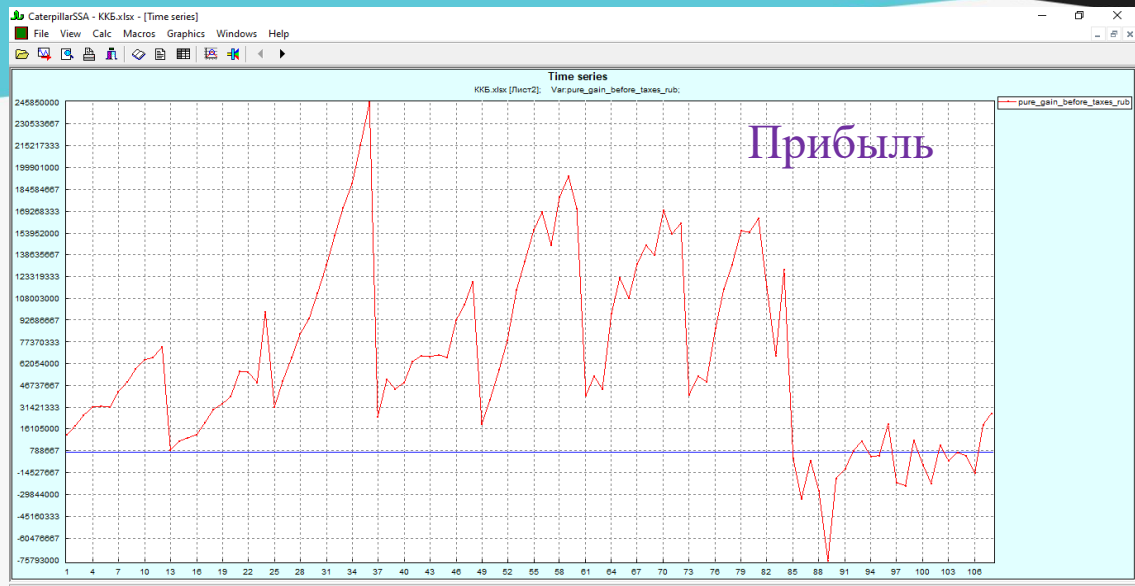
$$U = \begin{bmatrix} u_1^{i_1} & u_1^{i_2} & \dots & u_1^{i_r} \\ u_2^{i_1} & u_2^{i_2} & \dots & u_2^{i_r} \\ \vdots & \vdots & \dots & \vdots \\ u_{L-1}^{i_1} & u_{L-1}^{i_2} & \dots & u_{L-1}^{i_r} \\ u_L^{i_1} & u_L^{i_2} & \dots & u_L^{i_r} \end{bmatrix} = \begin{bmatrix} U_{L-1}^r \\ u_1^r \end{bmatrix}$$

$$(a_{L-1}, \dots, a_1)^T = \frac{1}{1-\vartheta^2} [U_{L-1}^r \cdot (u_1^r)^T]$$

$$\vartheta^2 = (u_1^{i_1})^2 + \dots + (u_1^{i_r})^2$$

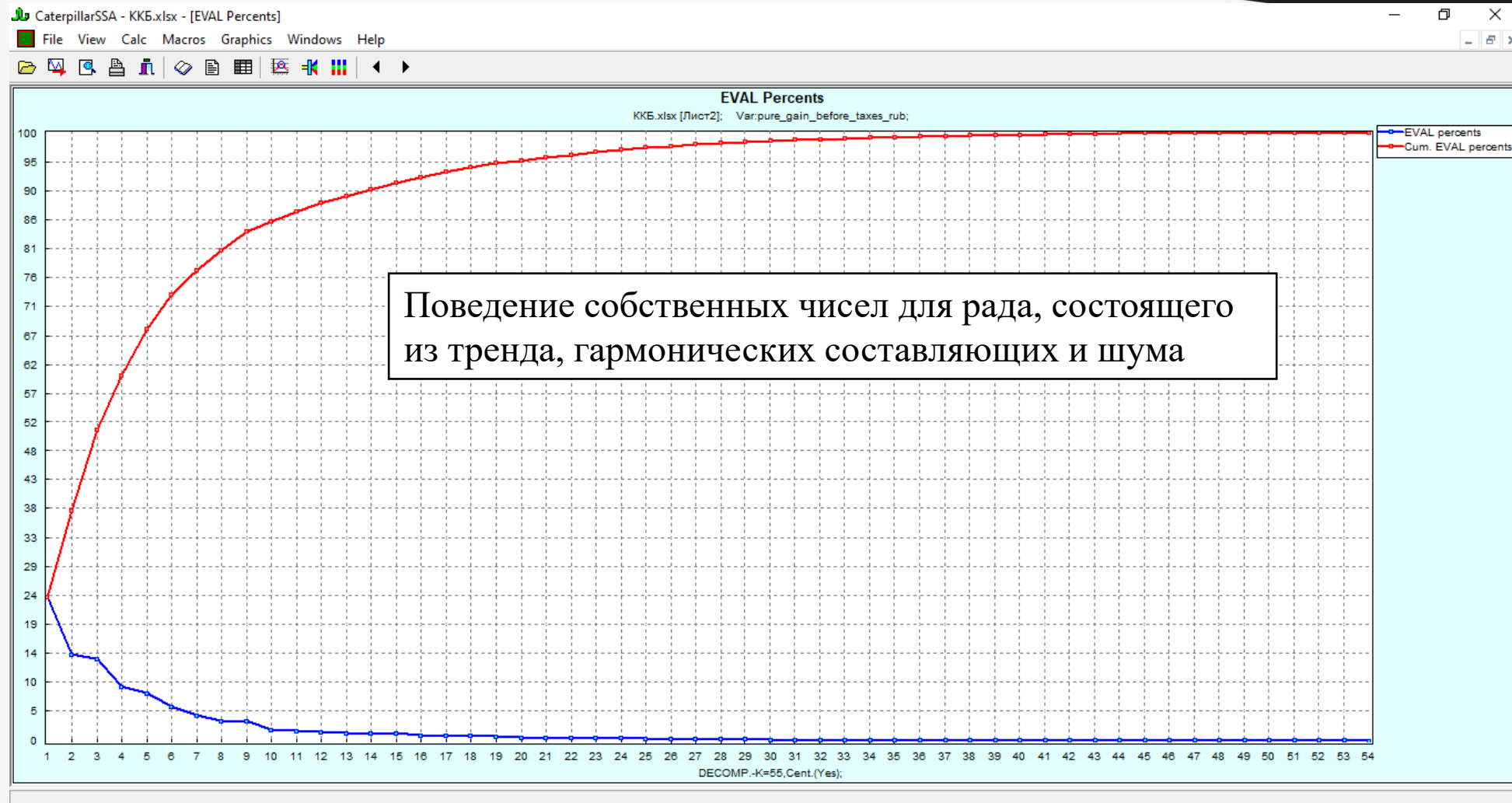
$$S = \mathbf{X}\mathbf{X}^T$$

# Анализ временных рядов ключевых показателей АО «ККБ» в программе CaterpillarSSA

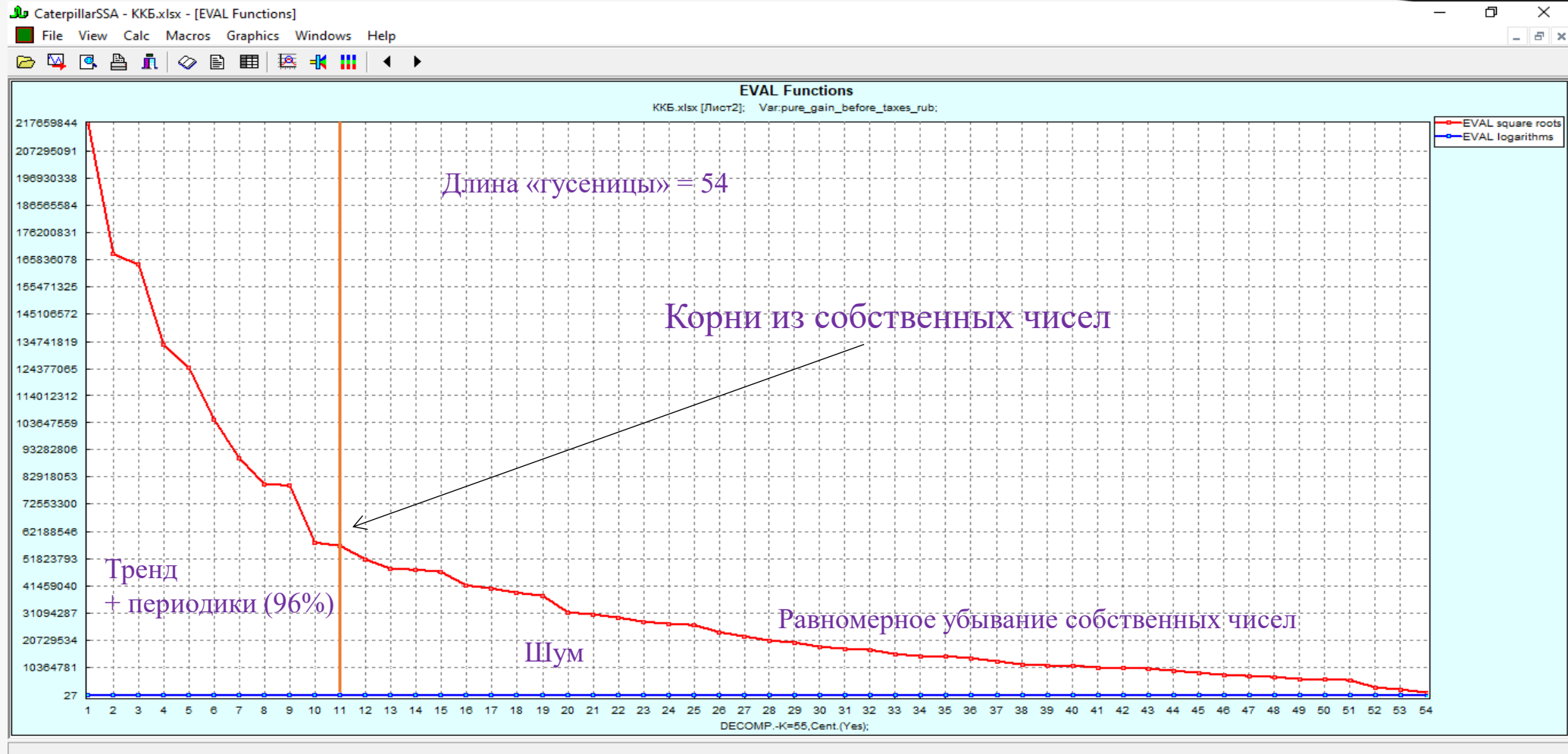




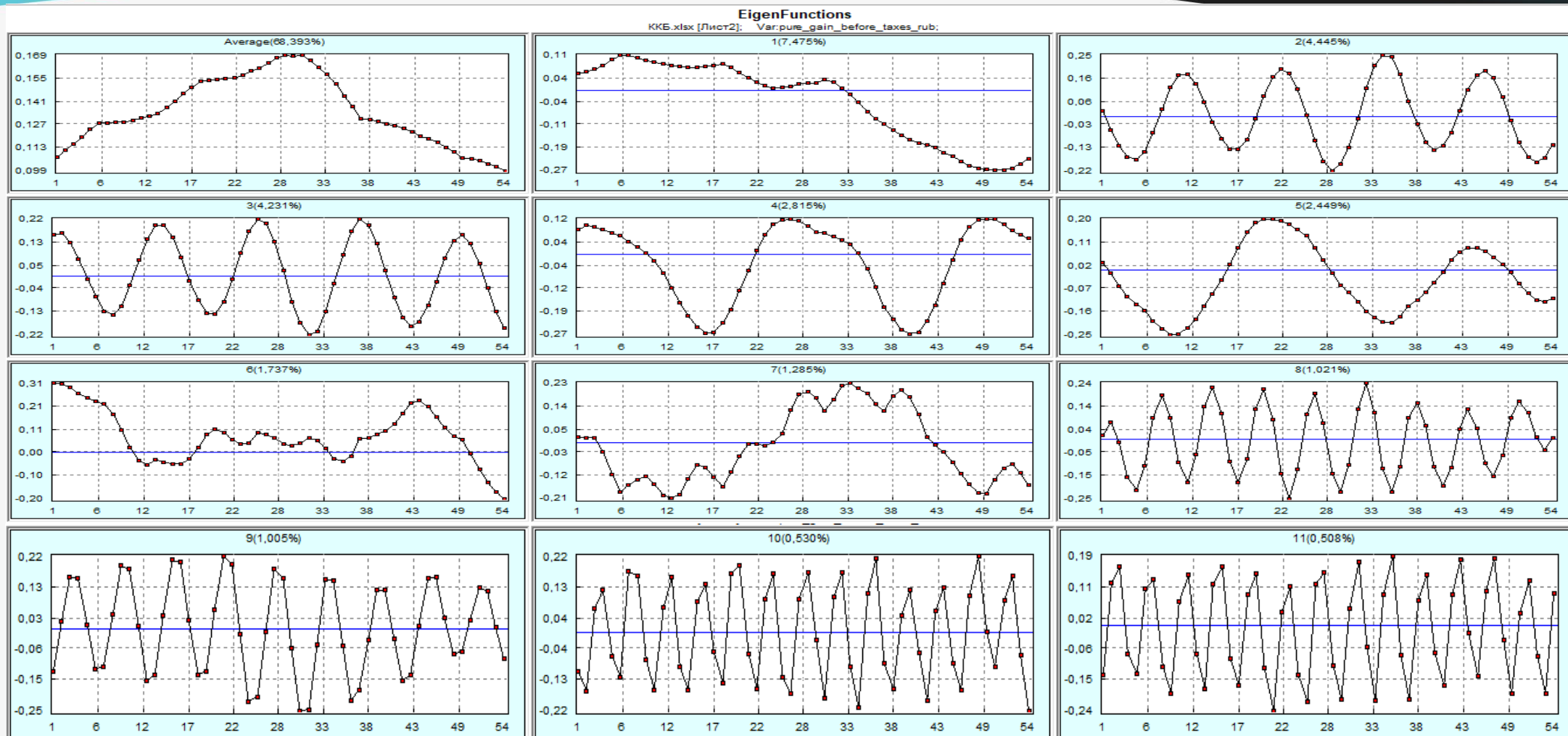
# Собственные числа в процентах



# Функции от собственных чисел ковариационной матрицы показателя «Прибыль» для банка «ККБ»



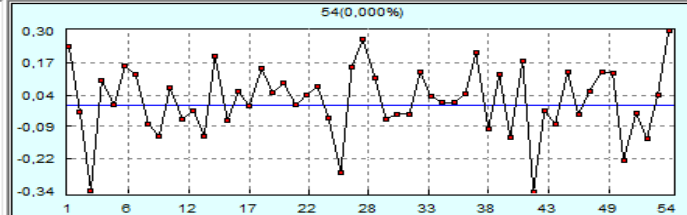
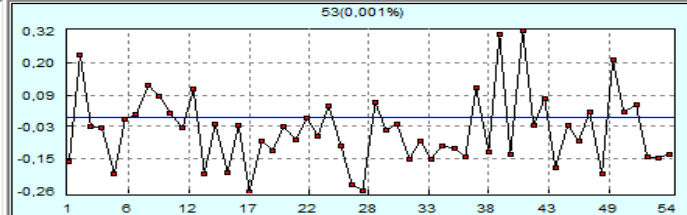
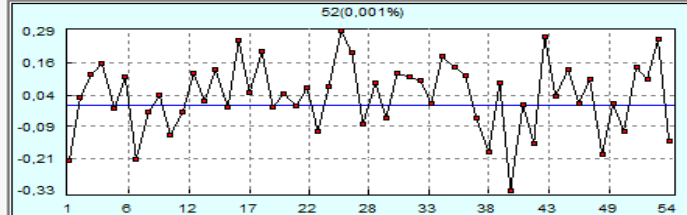
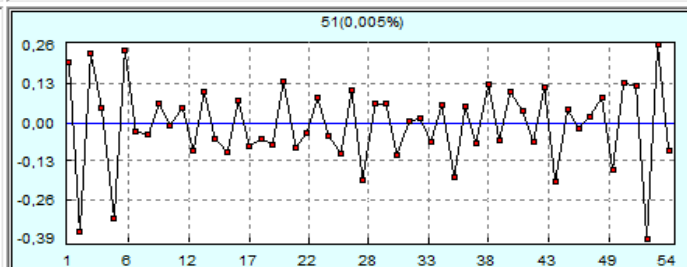
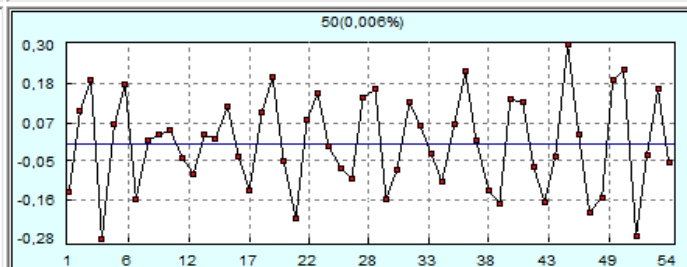
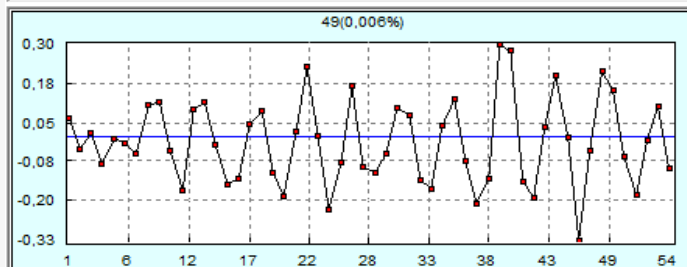
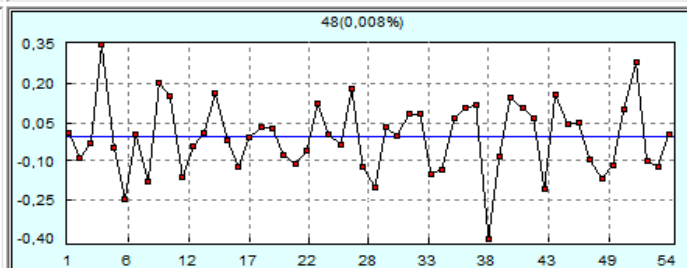
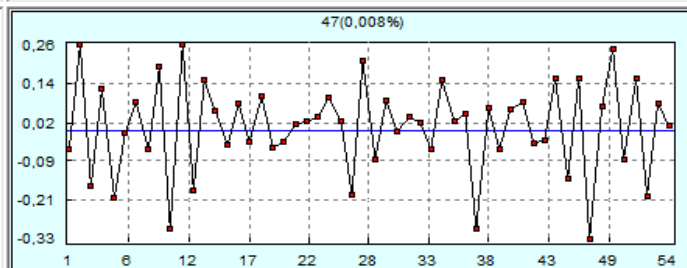
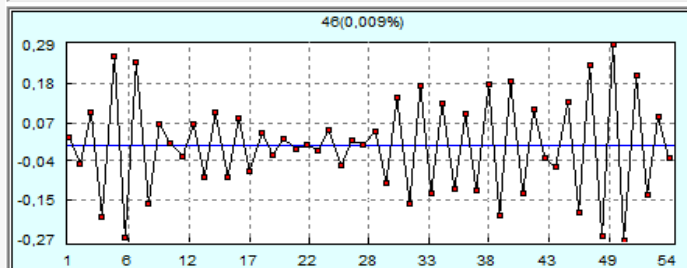
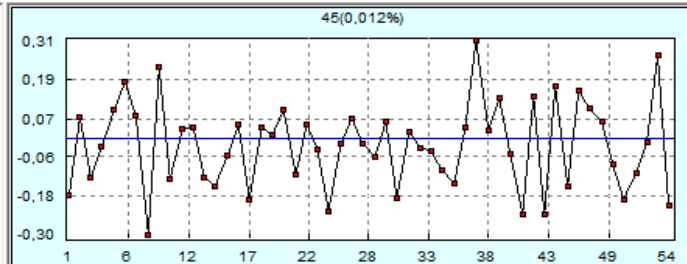
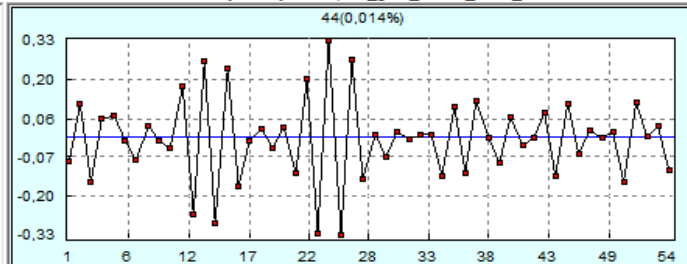
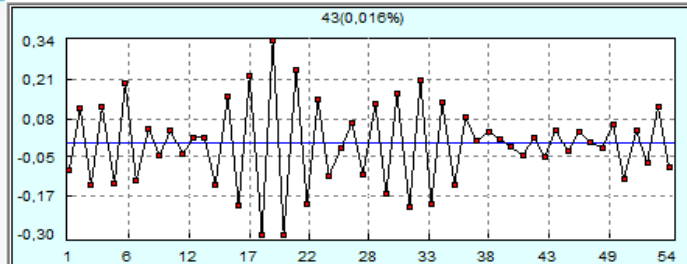
# Одномерные графики собственных векторов



# Одномерные графики собственных векторов

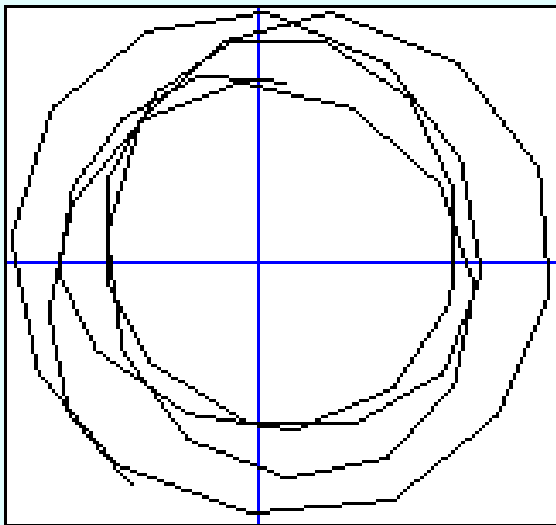
## EigenFunctions

ККБ.xlsx [Лист2]: Var: pure\_gain\_before\_taxes\_rub;



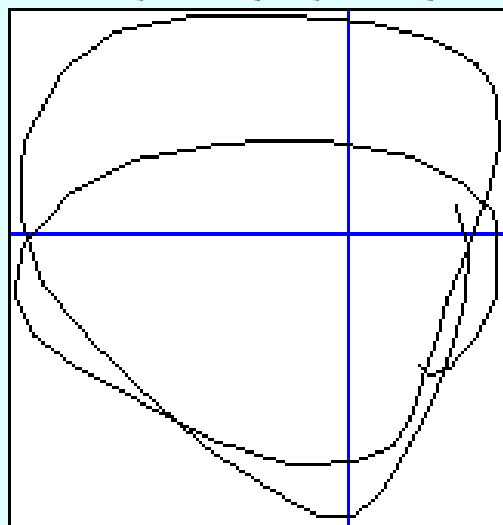
# Двумерные скаттерграммы собственных векторов

2(4,445%) - 3(4,231%)



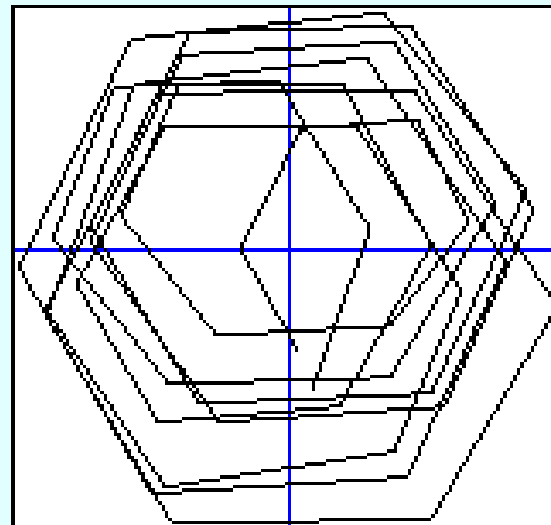
12

4(2,815%) - 5(2,449%)



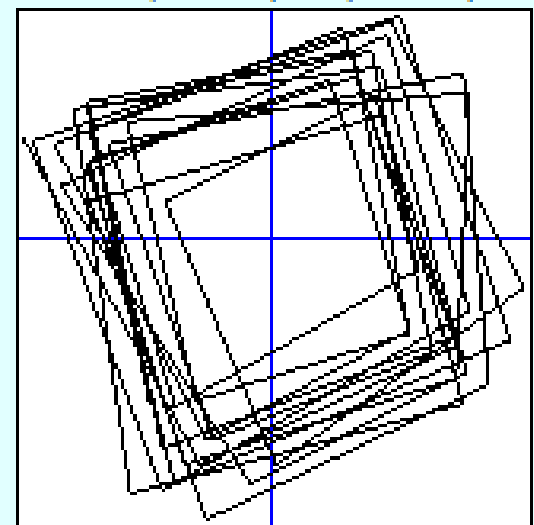
24

8(1,021%) - 9(1,005%)



6

10(0,530%) - 11(0,508%)



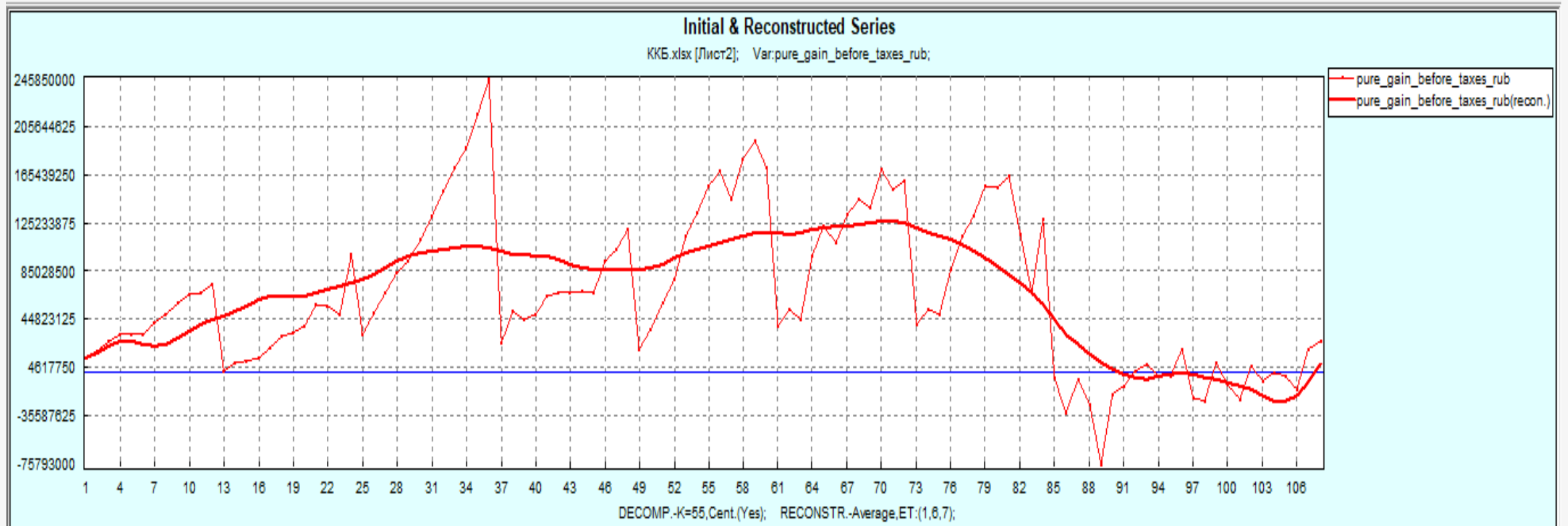
4

Период

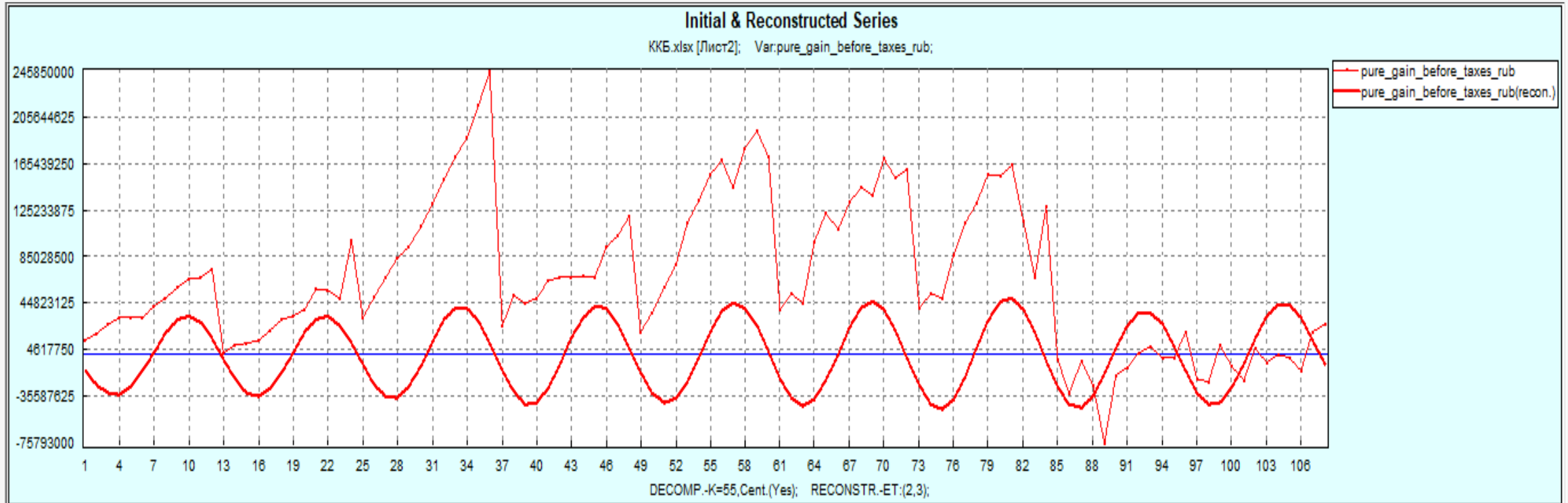
# Интерпретация главных компонент показателя «Прибыль» для банка «ККБ»

ГК, №	Фактическое содержание ГК
1, 6, 7	Нелинейный тренд
2-3	Годовая периодика (сезонная компонента)
4-5	Двухгодовая периодика
8-9	Полугодовая периодика (гармоника сезонной компоненты)
10-11	4 мес. колебание (гармоника сезонной компоненты)

# Реконструкция ряда по тренду (гусеничное среднее и 1,6,7 ГК)

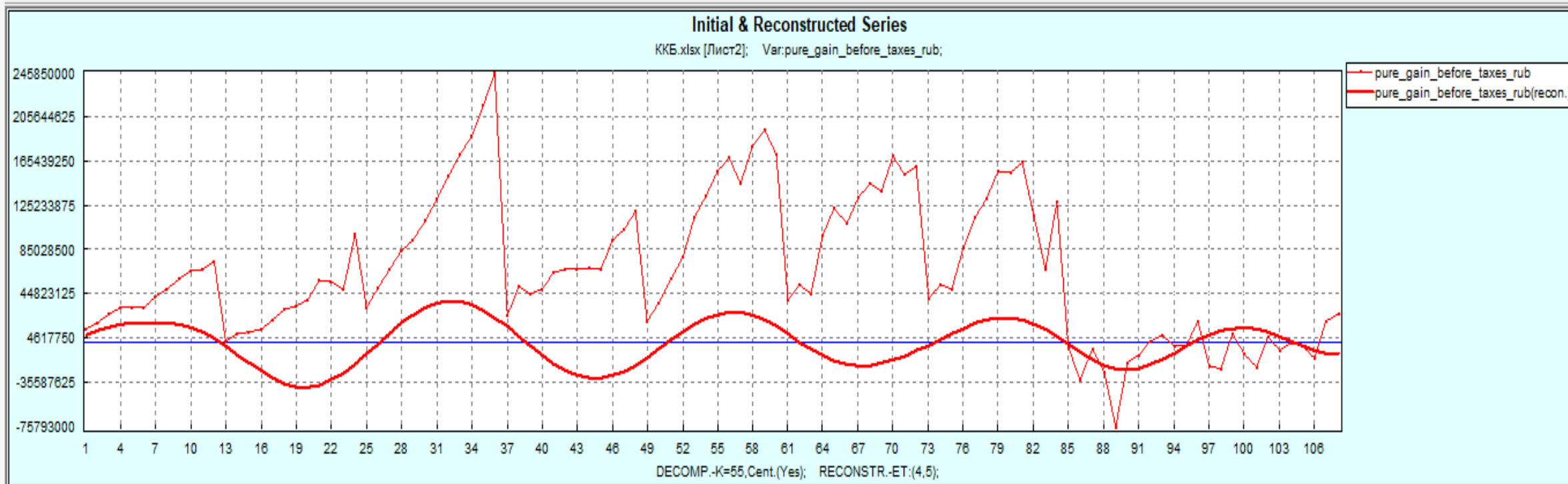


# Реконструкция ряда по 2-3 паре ГК (годовая цикличность без учета тренда)

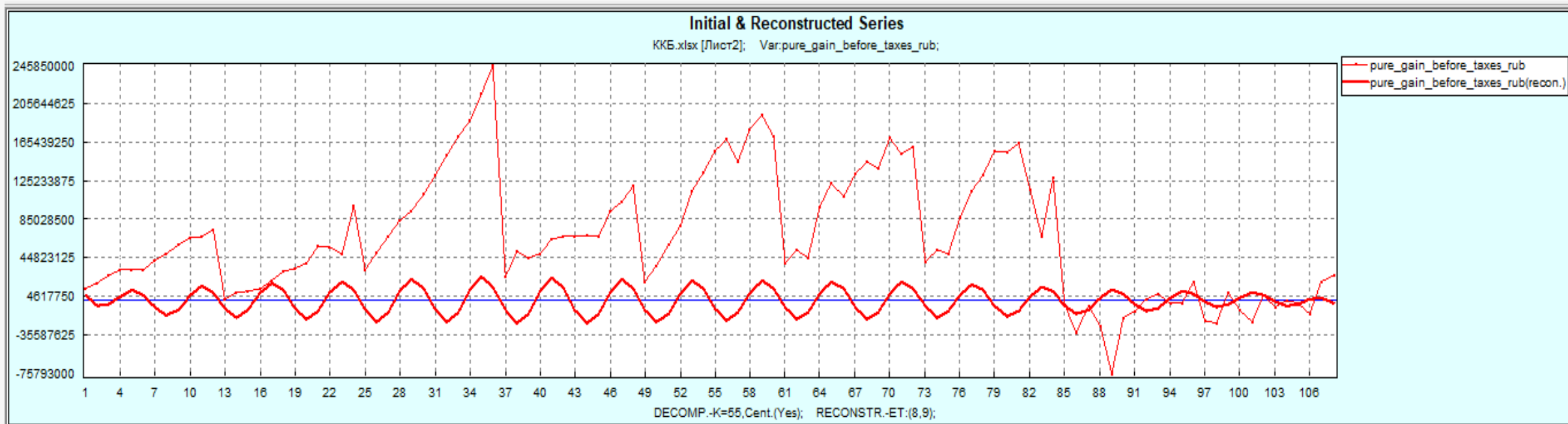




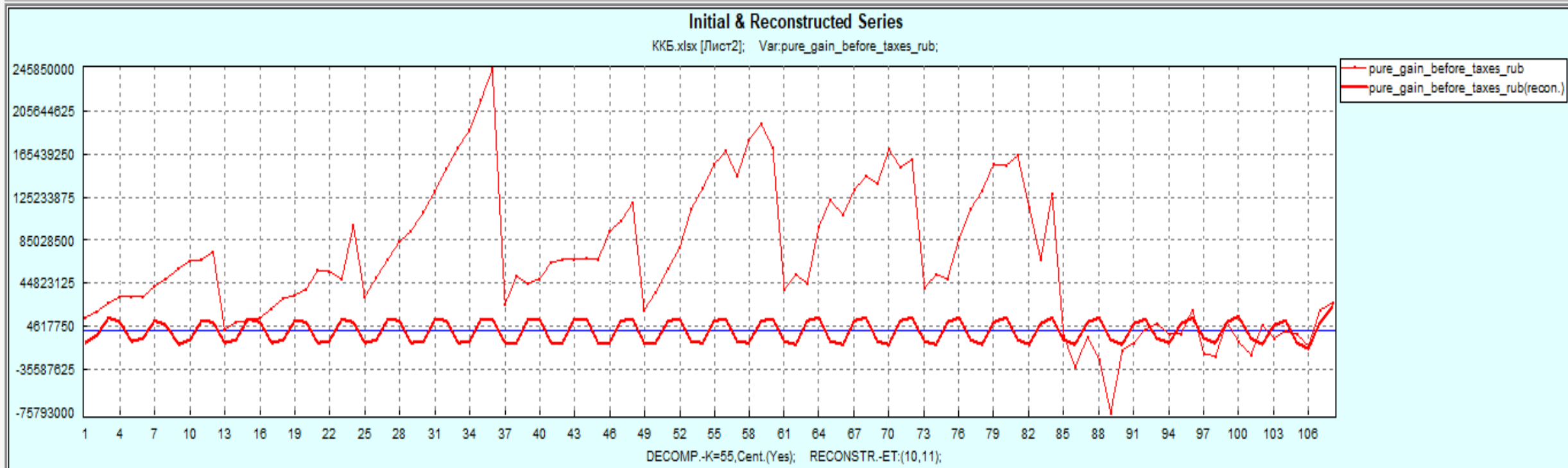
# Реконструкция ряда по 4-5 паре ГК (двухгодичная цикличность без учета тренда)



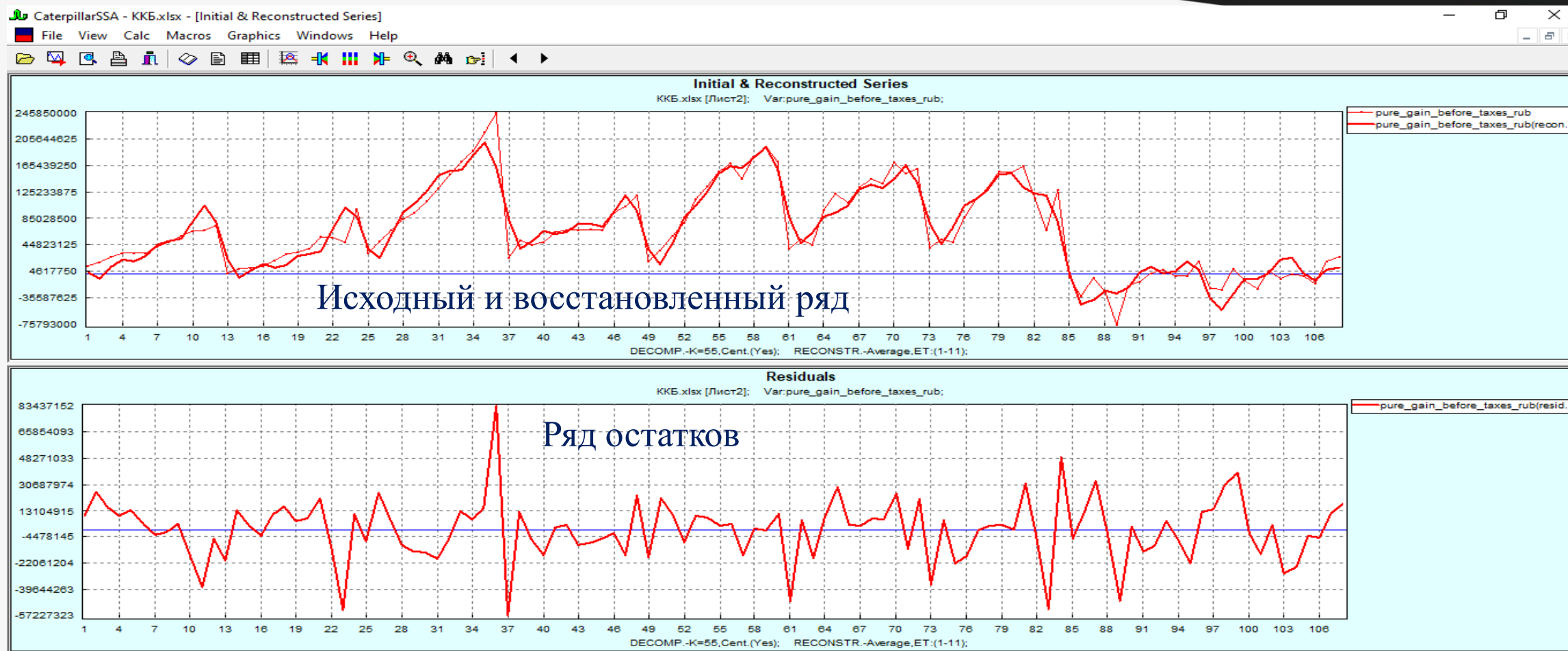
# Реконструкция ряда по 8-9 паре ГК (полугодовая цикличность без учета тренда)



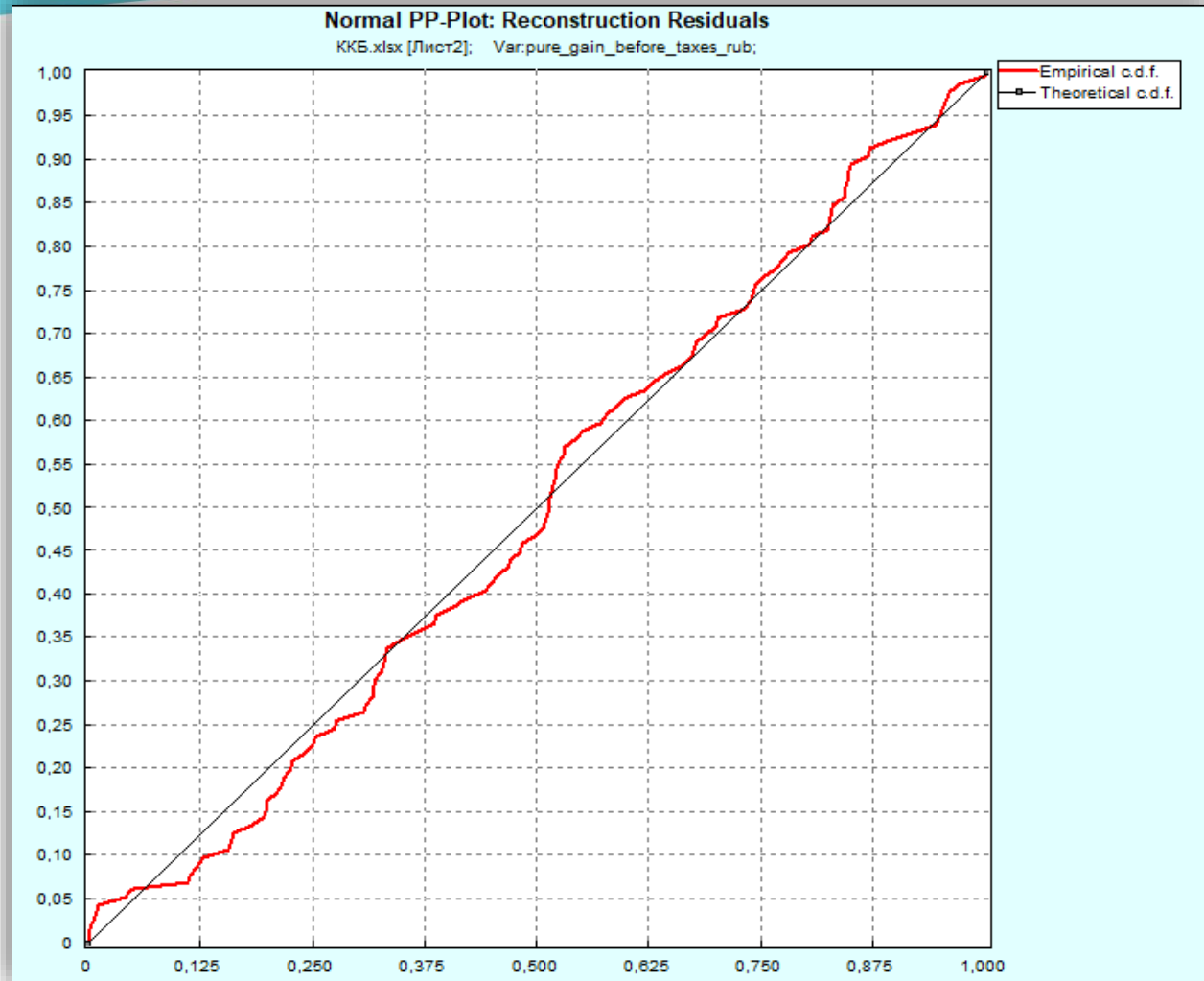
# Реконструкция ряда по 10-11 паре ГК (4 мес. колебание без учета тренда)



# Исходный ряд «Прибыль» и восстановленный ряд по первым 11-ти компонентам (вклад 96%)

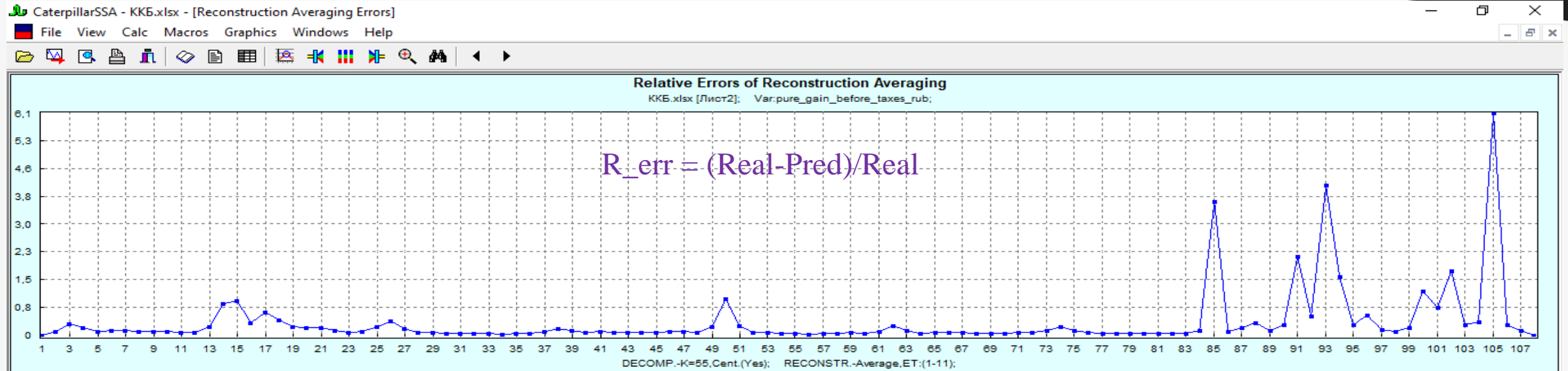


# Normal-pp-plot распределение, отвечающее отброшенным компонентам

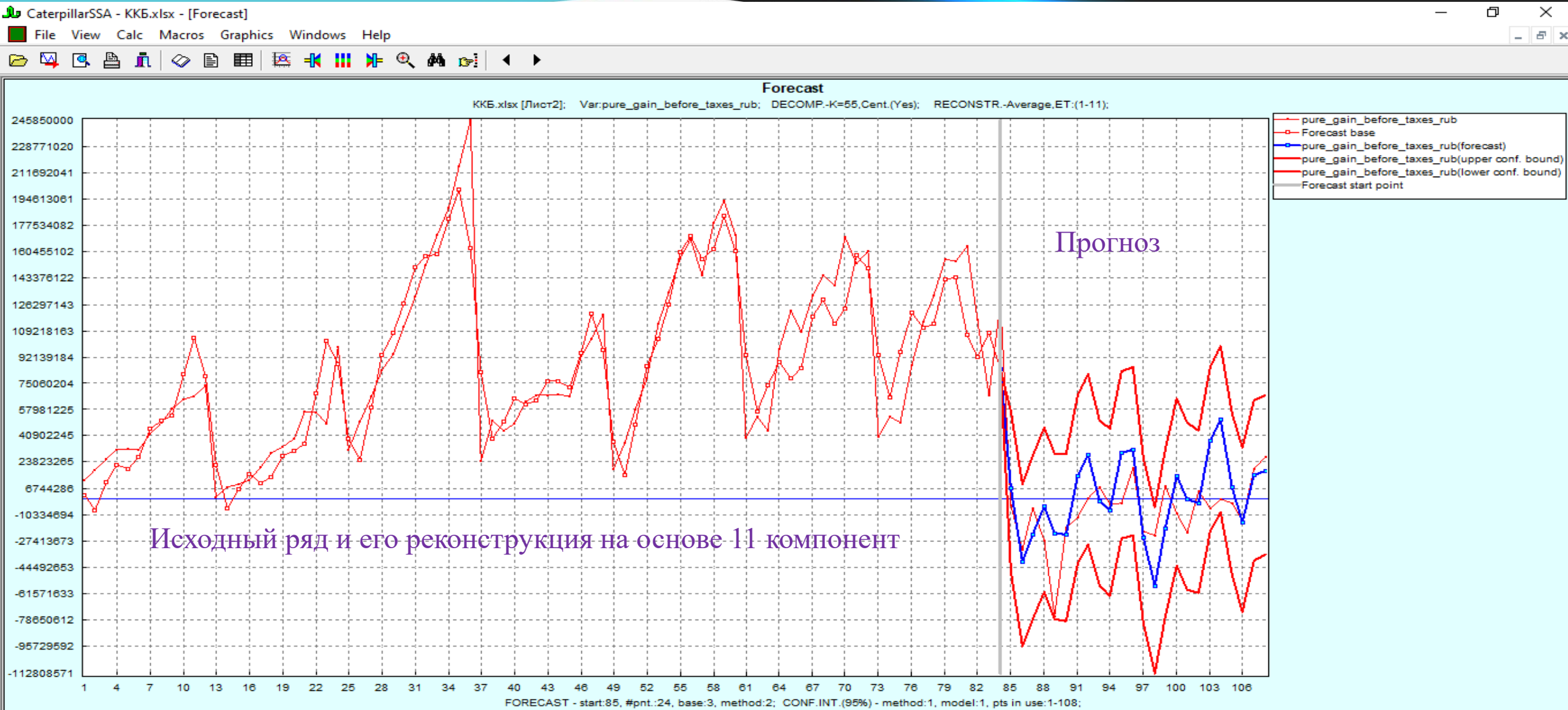


На рисунке представлен PP-plot функций накопленной вероятности для нормального распределения (изображенного отрезком прямой) и для данных, отвечающих отброшенным компонентам. Видно, что отбрасываемые данные согласуются с нормальным распределением, из чего можно предположить, что отвечающий им процесс по своему поведению близок к гауссовскому шуму.

# Зависимость относительной ошибки от порядкового номера месяца при восстановлении исходного ряда «Прибыль» по 11 ГК



# Фактические и прогнозируемые значения показателя «Прибыль» для АО «ККБ» на период в 24 месяца (значение вероятности, соответствующей доверительному интервалу = 0,95)

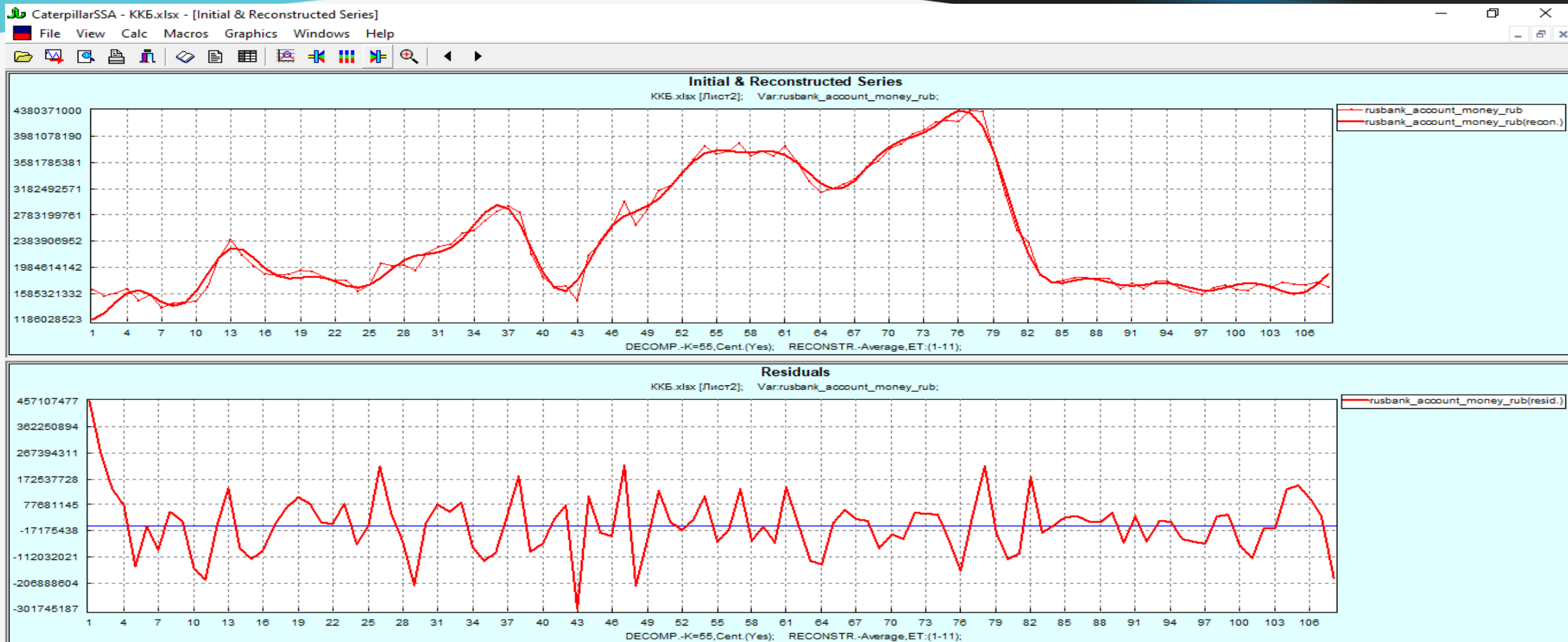


# Интерпретация главных компонент показателя «Счета в Банке России» для банка «ККБ»

ГК, №	Фактическое содержание ГК
1, 2, 5	Нелинейный тренд
3-4	Годовая периодика (сезонная компонента)
6-7	9 мес. колебание (гармоника сезонной компоненты)
7-8	9 мес. колебание (гармоника сезонной компоненты)
8-9	9 мес. колебание (гармоника сезонной компоненты)
10-11	10 мес. колебание (гармоника сезонной компоненты)



# Исходный и восстановленный ряд показателя «Счета в Банке России»



# Фактические и прогнозируемые значения показателя «Счета в Банке России» для банка «ККБ» на период в 24 месяца

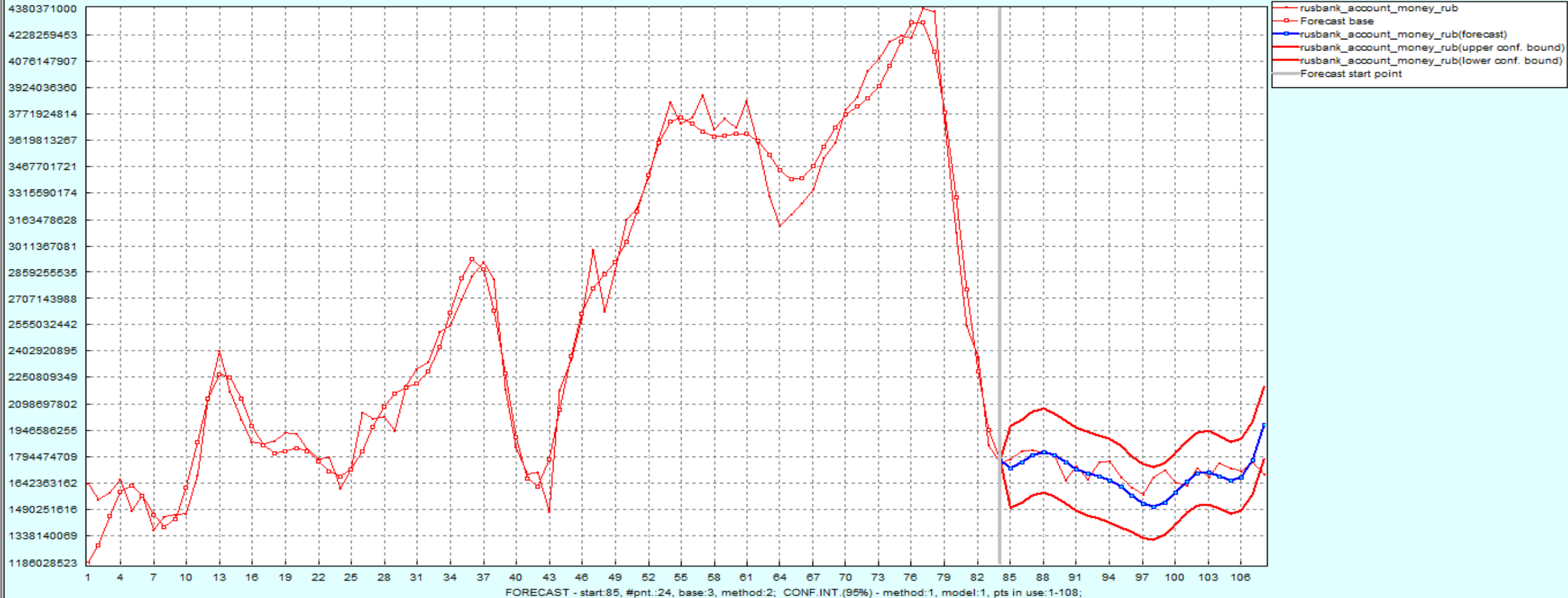
CaterpillarSSA - ККБ.xlsx - [Forecast]

File View Calc Macros Graphics Windows Help



Forecast

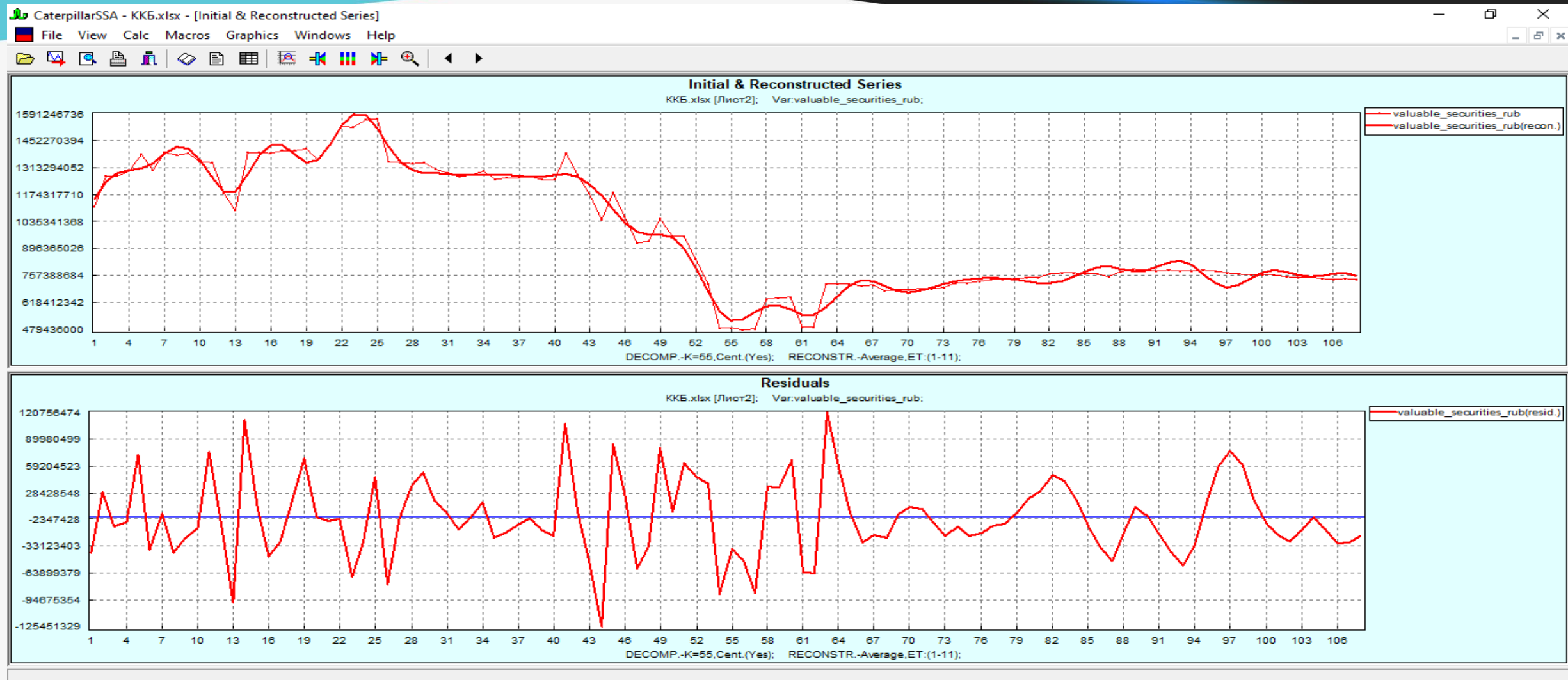
ККБ.xlsx [Лист2]; Var:rusbank\_account\_money\_rub; DECOMP-K=55,Cent.(Yes); RECONSTR.-Average,ET:(1-11);



# Интерпретация главных компонент показателя «Ценные бумаги» для АО «ККБ»

ГК, №	Фактическое содержание ГК
1, 2, 3	Нелинейный тренд
4-5	Двухгодичная периодика
6-7	8 мес. колебание (гармоника сезонной компоненты)
7-8	9 мес. колебание (гармоника сезонной компоненты)
9-10	Полугодовая периодика (гармоника сезонной компоненты)
10-11	Полугодовая периодика (гармоника сезонной компоненты)

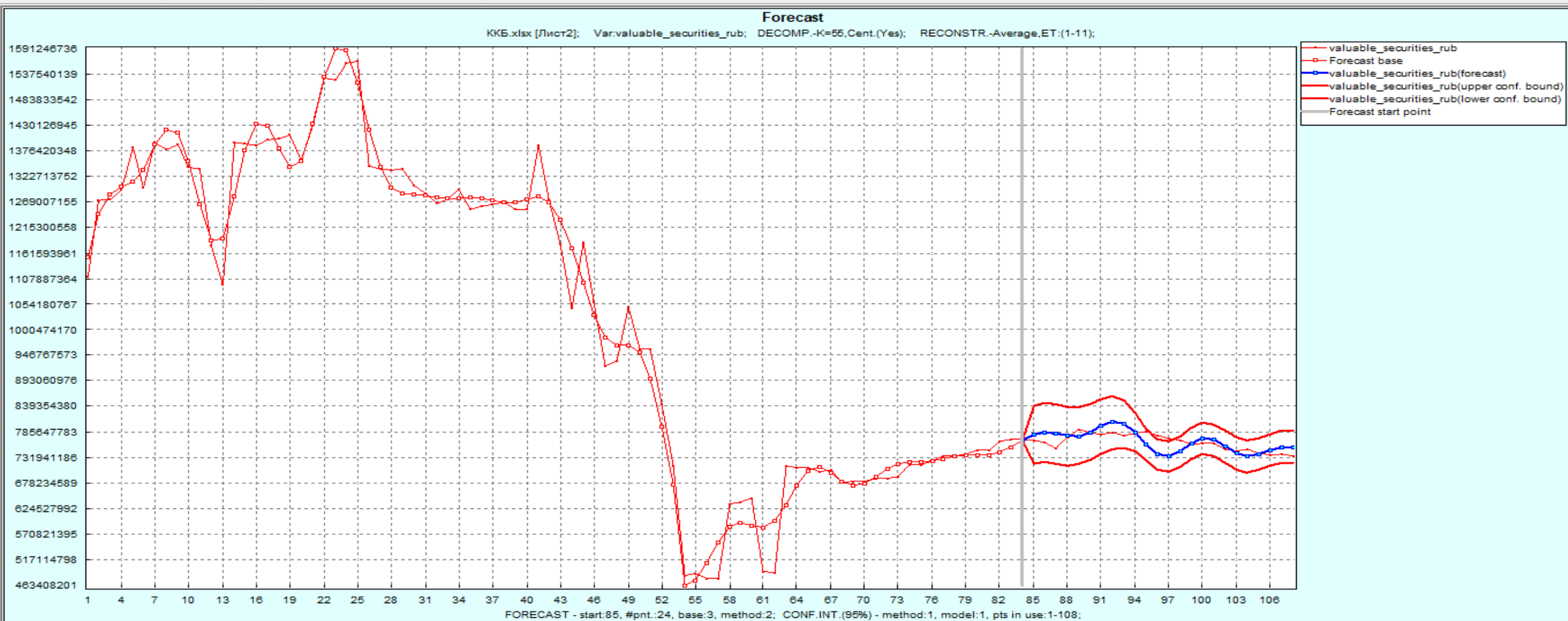
# Исходный и восстановленный ряд значений показателя «Ценные бумаги»



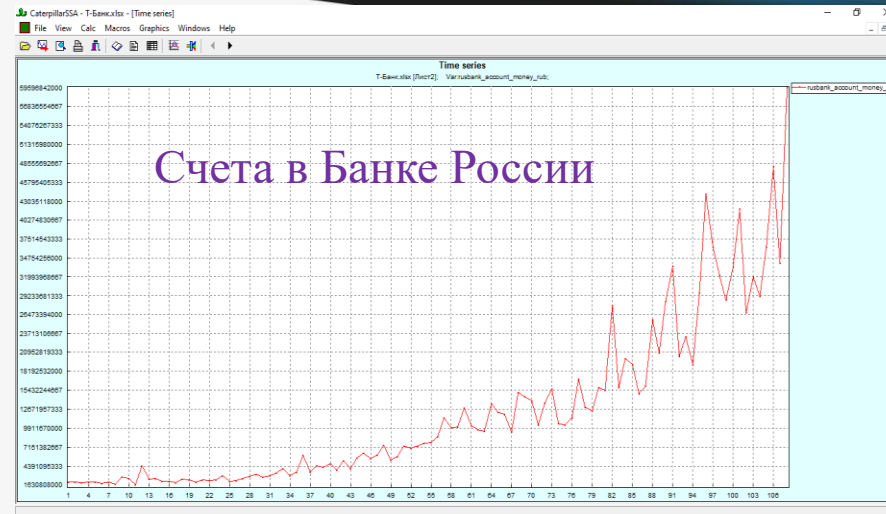
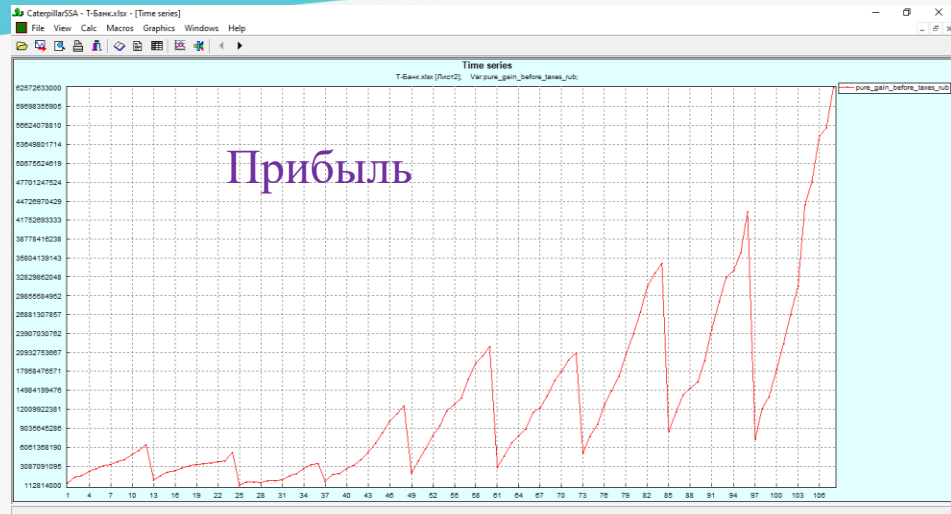
# Фактические и прогнозируемые значения показателя «Ценные бумаги» для банка «ККБ» на период в 24 месяца

CaterpillarSSA - ККБ.xlsx - [Forecast]

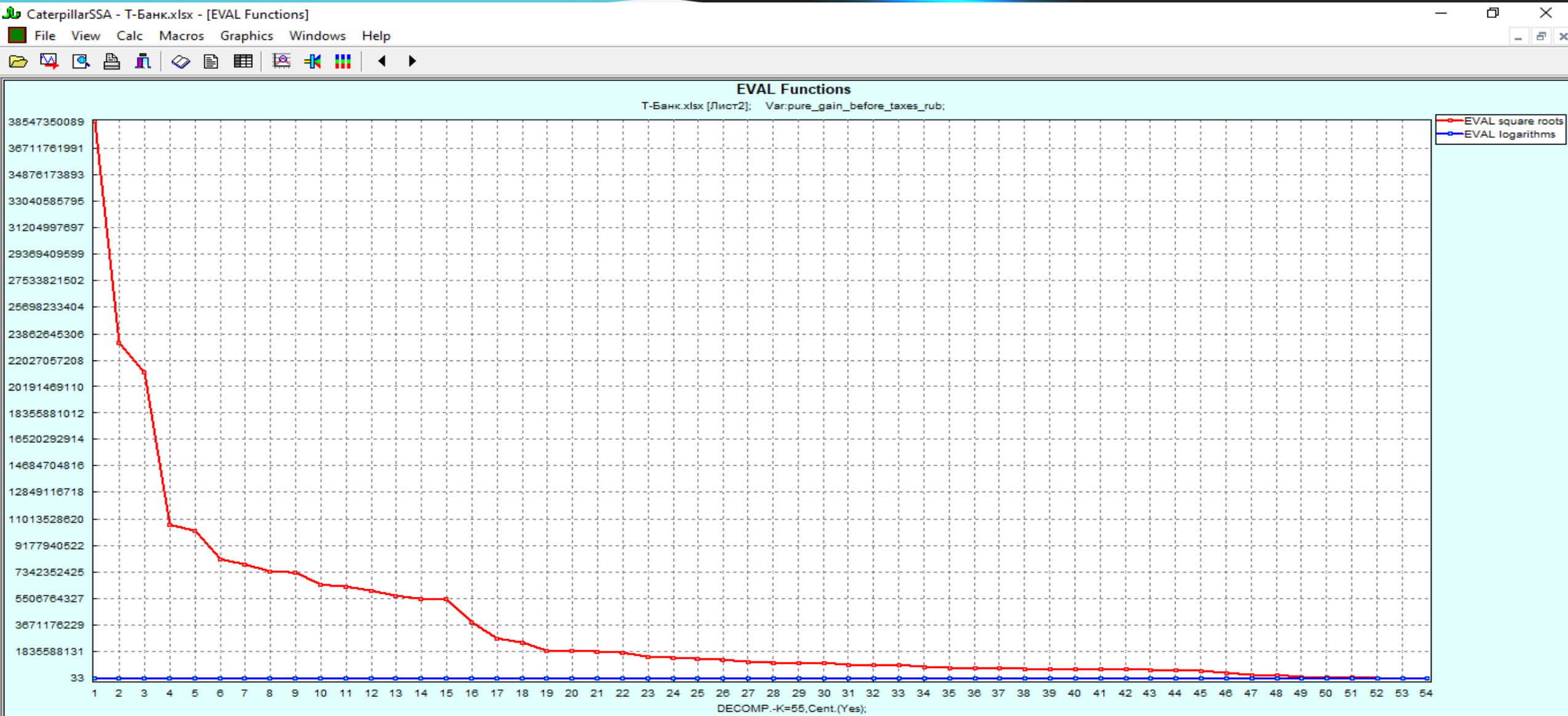
File View Calc Macros Graphics Windows Help



# Анализ временных рядов ключевых показателей АО «ТБанк»



# Функции от собственных чисел ковариационной матрицы показателя «Прибыль» для АО «ТБанк»



# Интерпретация главных компонент показателя «Прибыль» для АО «ТБанк»

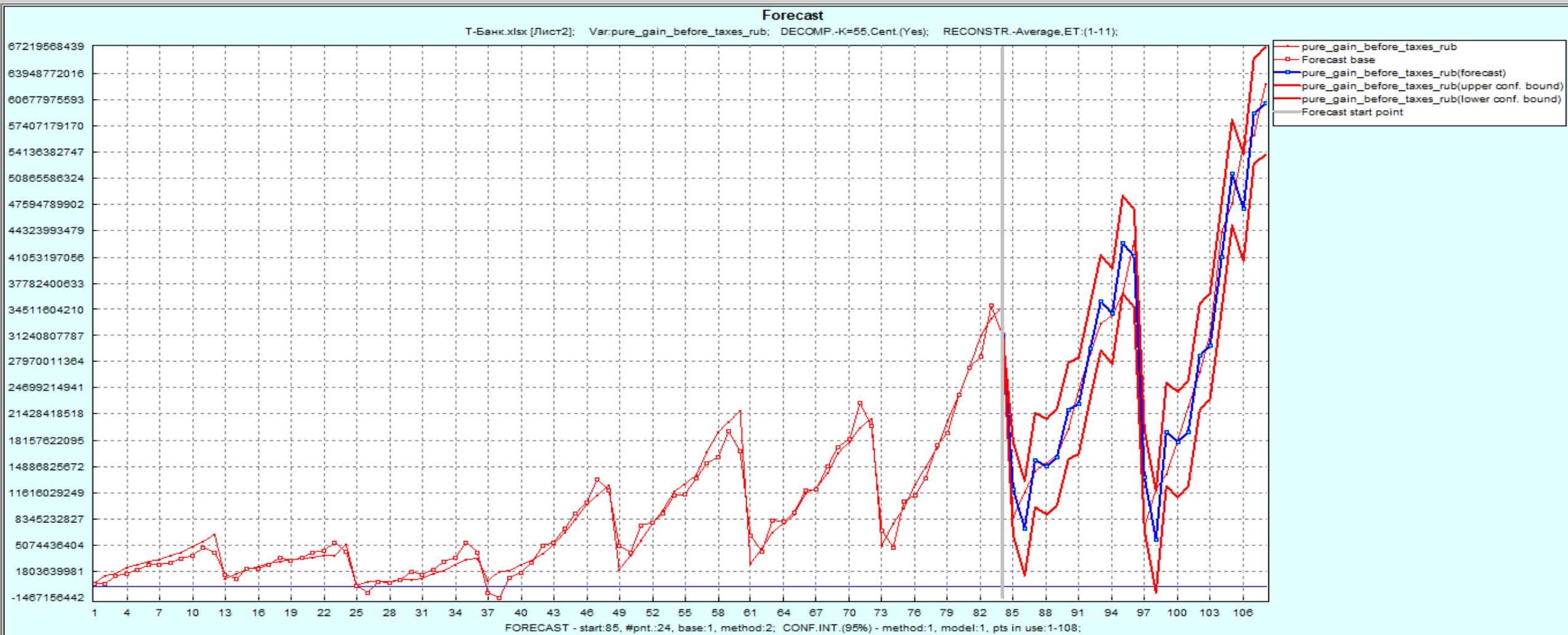
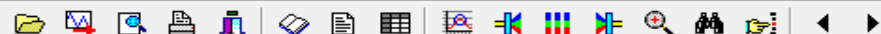
ГК, №	Фактическое содержание ГК
1, 6, 9	Нелинейный тренд
2-3	Годовая периодика (сезонная компонента)
4-5	Полугодовая периодика (гармоника сезонной компоненты)
7-8	4 мес. колебание (гармоника сезонной компоненты)
10-11	3 мес. колебание (гармоника сезонной компоненты)



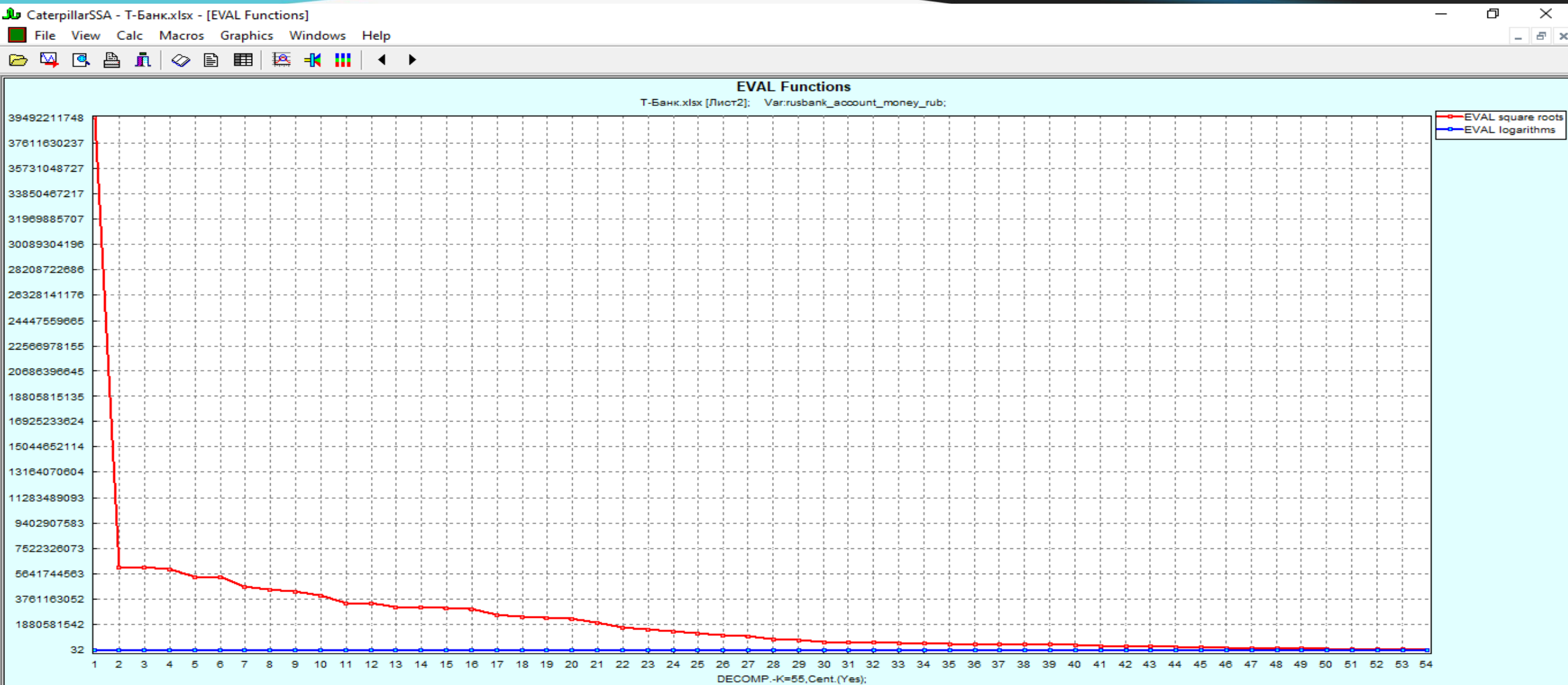
# Фактические и прогнозируемые значения показателя «Прибыль» для АО «ТБанк» на период в 24 месяца

CaterpillarSSA - Т-Банк.xlsx - [Forecast]

File View Calc Macros Graphics Windows Help



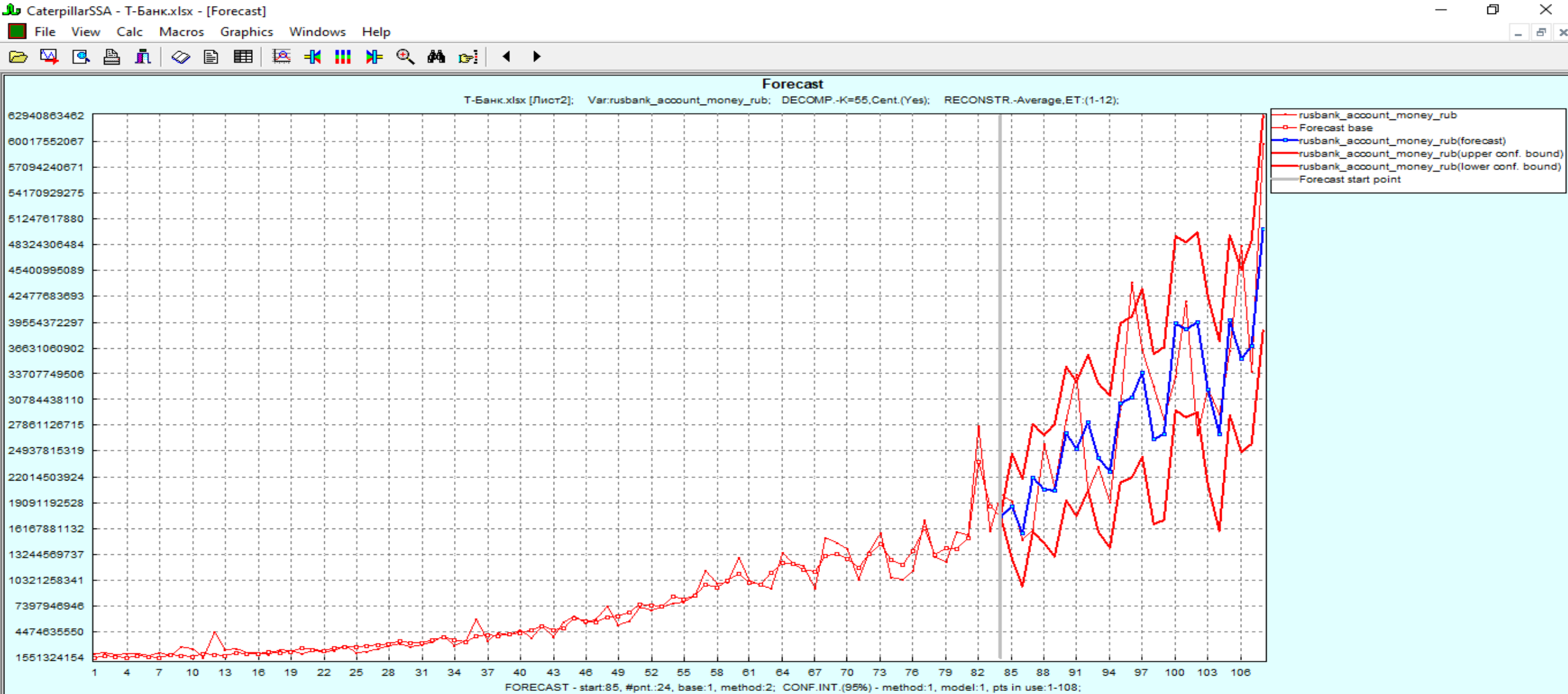
# Функции от собственных чисел ковариационной матрицы показателя «Счета в Банке России» для АО «ТБанк»



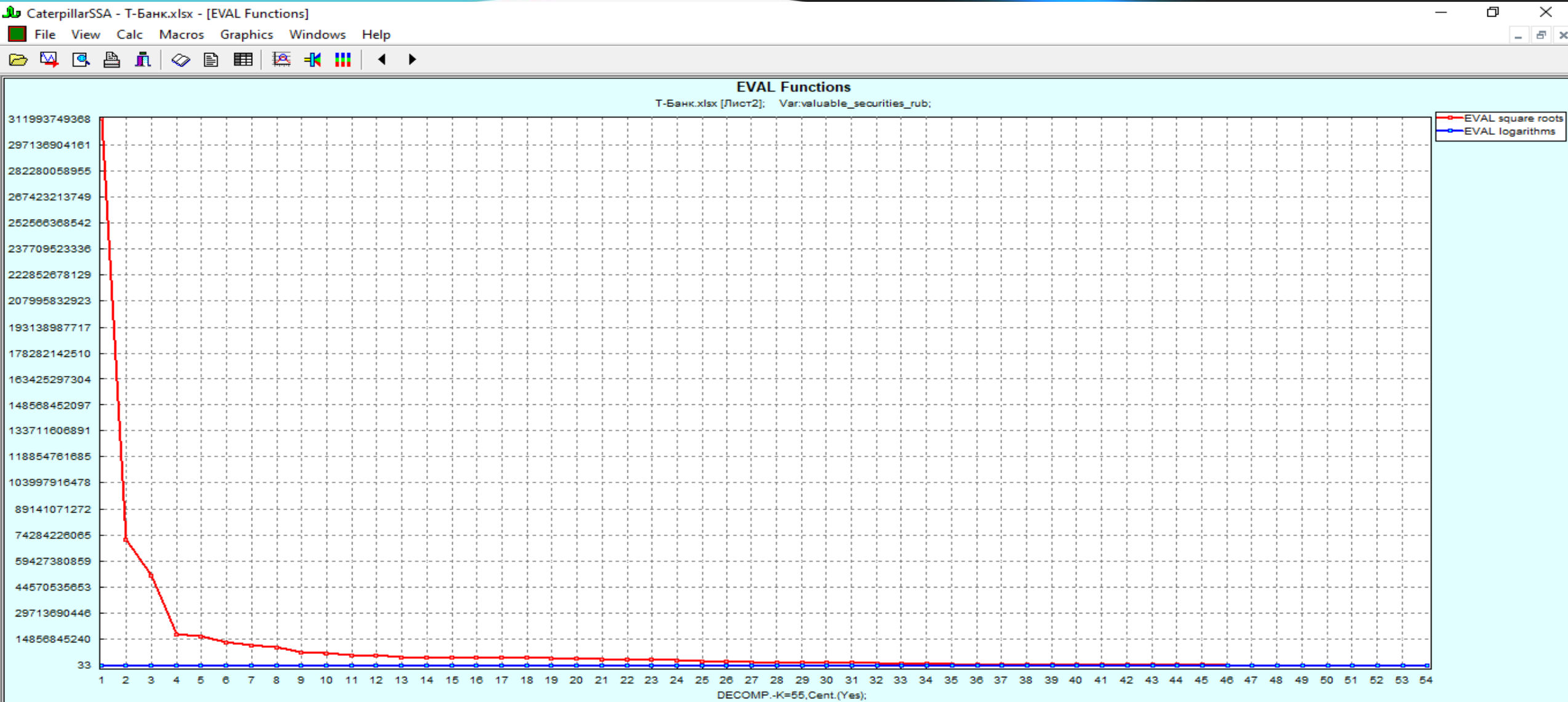
# Интерпретация главных компонент показателя «Счета в Банке России» для АО «ТБанк»

ГК, №	Фактическое содержание ГК
1,7,10	Нелинейный тренд
2-3	5 мес. колебание
3-4	5 мес. колебание
5-6	7 мес. колебание
8-9	3 мес. колебание
11-12	3 мес. колебание

# Фактические и прогнозируемые значения показателя «Счета в Банке России» для АО «ТБанк» на период в 24 месяца



# Функции от собственных чисел ковариационной матрицы показателя «Ценные бумаги» для АО «ТБанк»



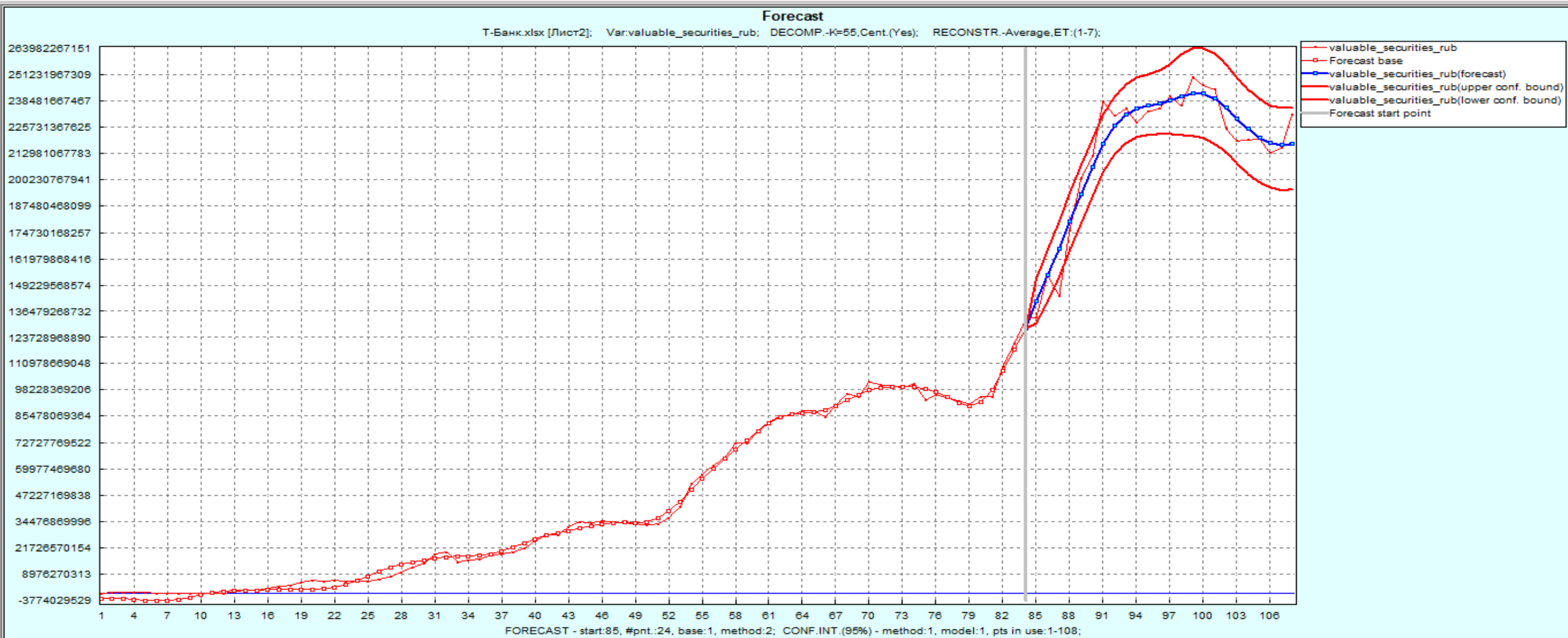
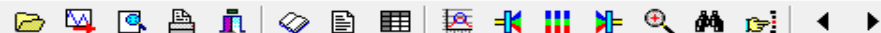
# Интерпретация главных компонент показателя «Ценные бумаги» для АО «ККБ»

ГК, №	Фактическое содержание ГК
1, 2, 3	Нелинейный тренд
4-5	Двухгодовая периодика
6-7	8 мес. колебание (гармоника сезонной компоненты)
7-8	9 мес. колебание (гармоника сезонной компоненты)
9-10	Полугодовая периодика (гармоника сезонной компоненты)
10-11	Полугодовая периодика (гармоника сезонной компоненты)

# Фактические и прогнозируемые значения показателя «Ценные бумаги» для АО «ТБанк» на период в 24 месяца

CaterpillarSSA - Т-Банк.xlsx - [Forecast]

File View Calc Macros Graphics Windows Help



# Анализ результатов и выводы

Представлена новая математическая методика анализа и прогнозирования финансового состояния кредитных организаций

В основе методики - анализ временных рядов ключевых показателей («Прибыль», «Счета в Банке России», «Ценные бумаги») на основе сингулярно-спектрального анализа в подходе «Гусеница»-SSA

Наглядно продемонстрирована возможность декомпозиции временных рядов ключевых показателей на их составляющие: трендовые, гармонические и случайный шум. На основе трендовых и гармонических компонент удалось с высокой точностью выполнить реконструкцию и прогнозирование анализируемых временных рядов

Показано, что для благонадежного банка тренд временных рядов «Прибыль», «Счета в Банке России» и «Ценные бумаги» имеет тенденцию к росту, в то время как для неблагонадежного банка – тренд «спадает» к моменту отзыва лицензии



# Благодарности

Хотела бы поблагодарить моего научного руководителя доктора физико-математических наук, профессора Иванова Виктора Владимировича за помощь и ценные указания в подготовке данной научной работы.

Также выражаю благодарность моему научному консультанту, кандидату физико-математических наук Акишиной Елене Павловне за поддержку в продвижении исследований в сфере ПОД/ФТ.



Спасибо за внимание!  
Вопросы?