

Все что вы хотели знать

о хемометрике, но стеснялись спросить

Алексей Померанцев



Научная конференция

01.04.07



Российское
хемометрическое
общество



Институт
химической
физики РАН

Лекция в ОмГУ

1

Содержание

Введение

1. Калибровка
2. Анализ изображений
3. Экология
4. Биология
5. MSPC & PAT
6. Четвертая парадигма

Хеометрика: два определения

Дедуктивное

Хеометрика - это научная дисциплина, находящаяся на стыке химии и математики, предметом которой являются математические методы исследования химических данных

сайт Российского хеометрического общества

Индуктивное

Хеометрика – это то, что делают хеометрики.

сайт Международного хеометрического общества

Хеометрики – это такие люди, которые все время пьют пиво и воруют идеи у математиков

Svante Wold

Что делают хемометрики?

- контролируют производство полупроводников, аспирина и пива;
- исследуют причины разрушения старинных документов;
- проводят допинг контроль спортсменов;
- расшифровывают состав косметики Древнего Египта;
- локализуют месторождение золота в Швеции;
- идентифицируют подозреваемого в убийстве Анны Линд;
- проводят диагностику артрита и рака на ранних стадиях;
- исследуют органические субстанции в кометном веществе;
- исследуют кормовой рацион на свинооткормочных комплексах;
- проверяют, как рацион питания влияет на умственные способности;
- определяют фальшивые лекарства;
- и еще многое, многое другое

А если серьезно

- Хеометрика имеет дело с **данными** (зачас-тую с очень большими), поэтому хеометрика - это подраздел информатики (Data mining)
- Данные, которые исследует хеометрика по большей части происходят из **химии**, поэтому хеометрика - это подраздел химии (Analytical chemistry)
- Методы, которые использует хеометрика ориентированы на **формальное** моделирование (Soft modeling)

Два «не» и три «да»

1. Хемометрика \neq химическая метрология
2. Хемометрика \neq статистика в химии

Хемометрика решает следующие задачи :

- (1) **как получить информацию**
из химических данных
- (2) **как представить** эту информацию
- (3) **как получить данные**, содержащие такую информацию.

Когда появилась?



«Каждая попытка применить математические методы для исследования химических проблем должна рассматриваться как абсолютно абсурдная и противоречащая самому духу химии.

Если математический анализ, когда - либо займет сколько-нибудь значительное место в химии – извращение, которое по счастью почти невероятно – это повлечет за собой повсеместно быстрое вырождение этой науки».

Огюст Конт, 1825

США, Сиэтл, 1974 год

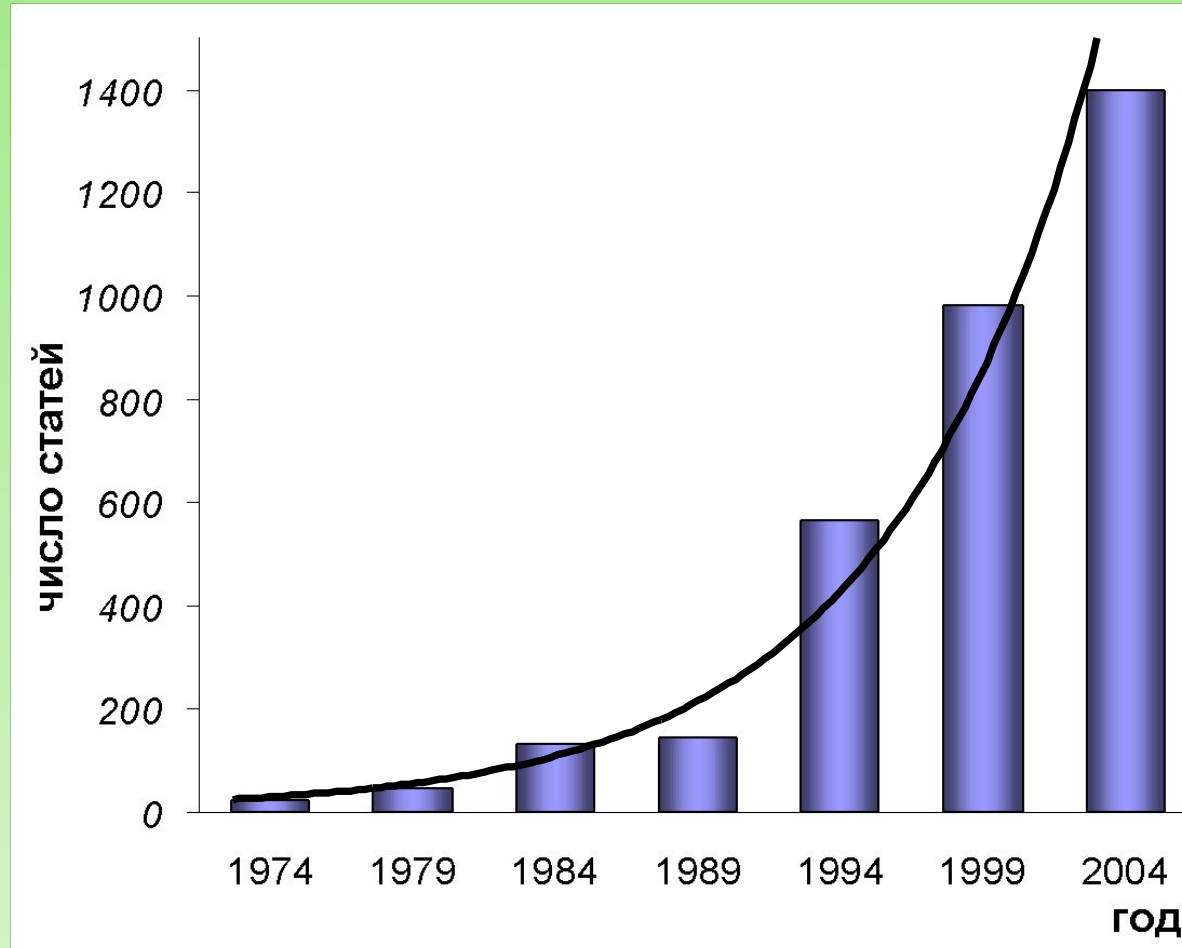


**Брюс Ковальски
(B. Kowalski)**



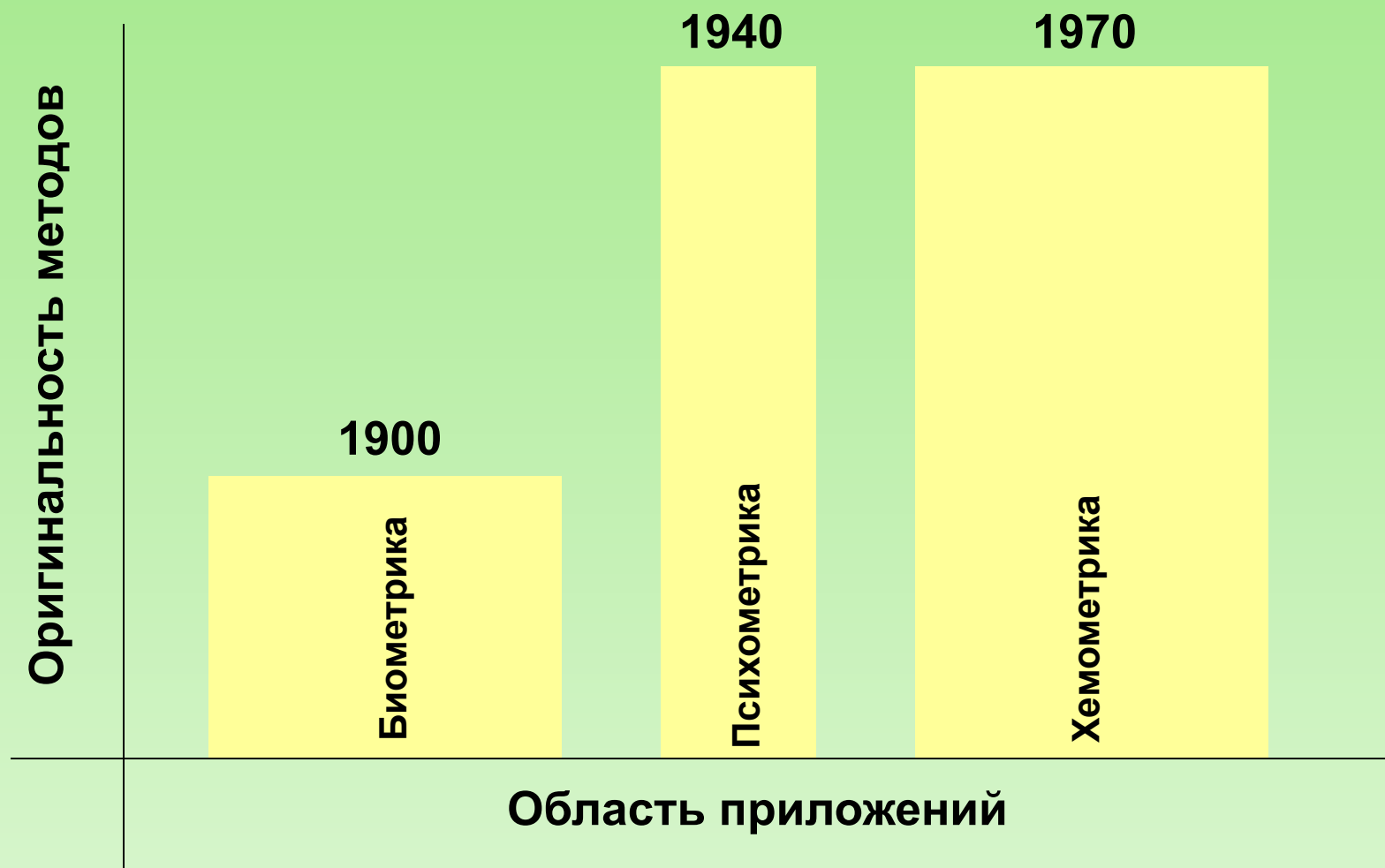
**Сванте Волд
(Svante Wold)**

Прогресс хемометрики



Статьи, опубликованные в журналах издательства Elsevier

Хеометрика среди своих сестер



Основные принципы хемометрики

1. Использование многомерного подхода при конструировании экспериментов и анализе их результатов.

2. Что считать шумом, а что – информацией, всегда решается с учетом поставленных целей и методов, используемых для ее достижения.

3. Понятие эффективного (химического) ранга и скрытых, латентных переменных, число которых равно этому рангу.

Основная задача хемометрики

Заменить **прямые** измерения, которые либо –

- невозможны
- дороги
- длительны

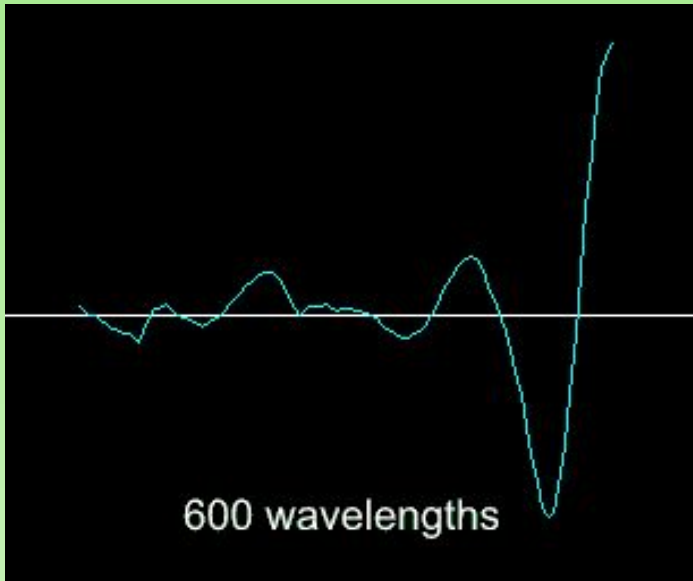
на **косвенные** измерения, которые –

- доступны
- дешевы
- быстры

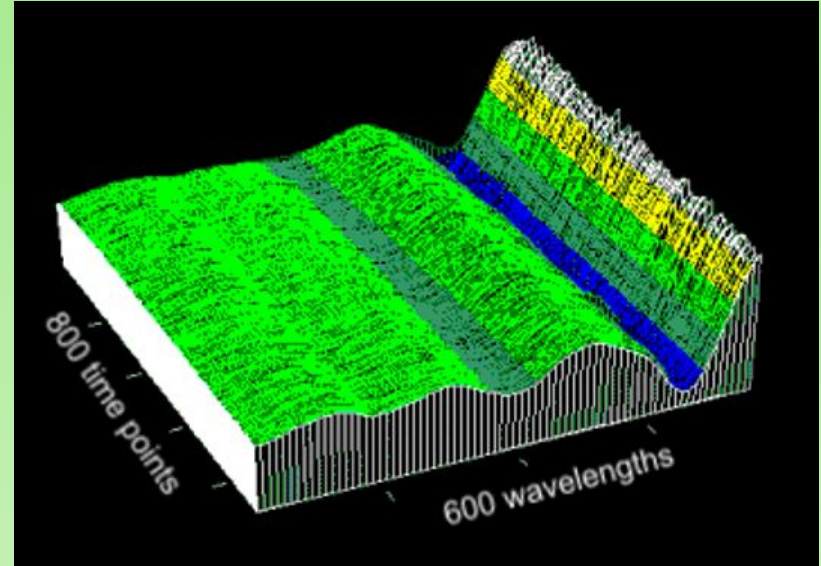
с последующей их **обработкой** (калибровкой).

Много переменных и много измерений

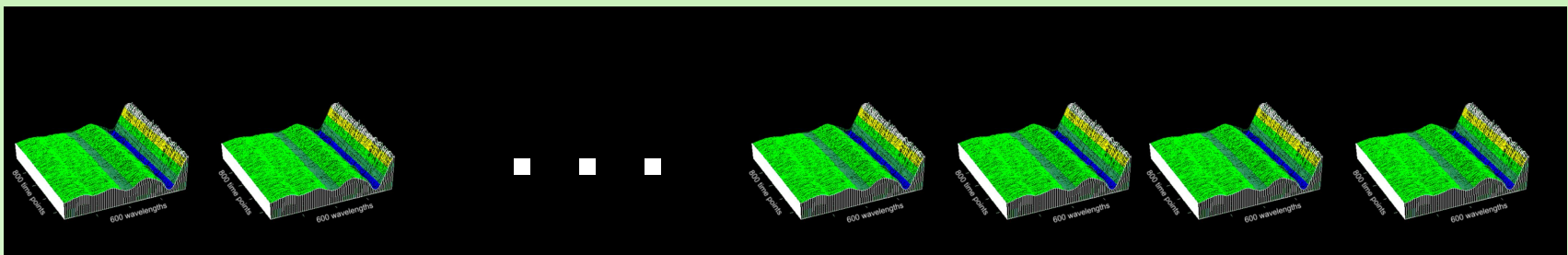
Одно измерение – спектр (600 точек)



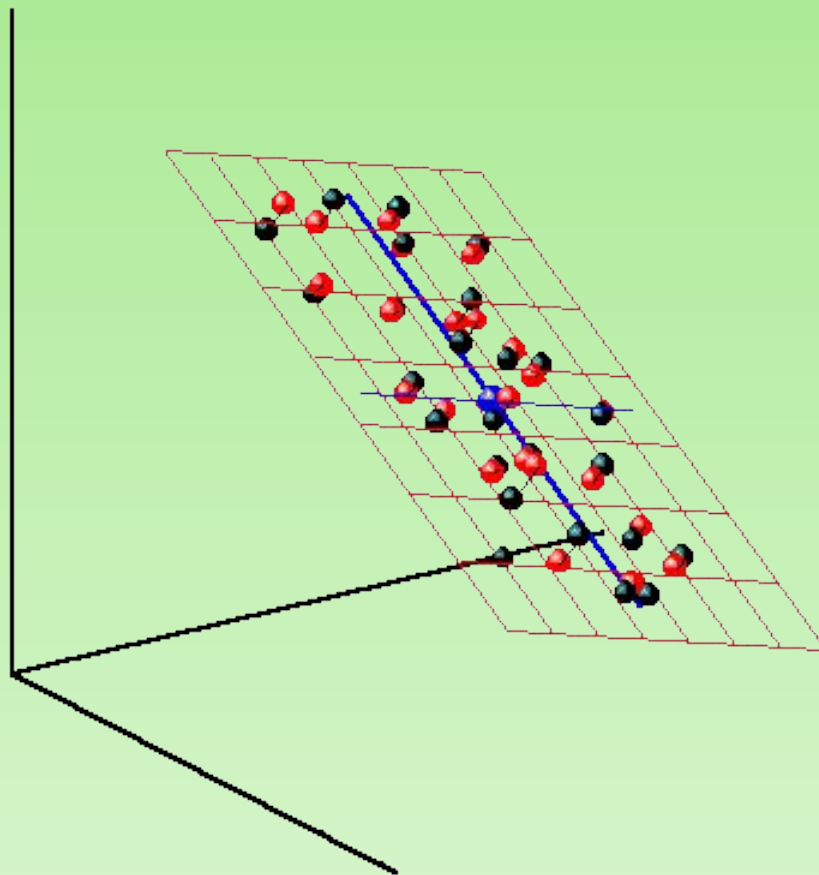
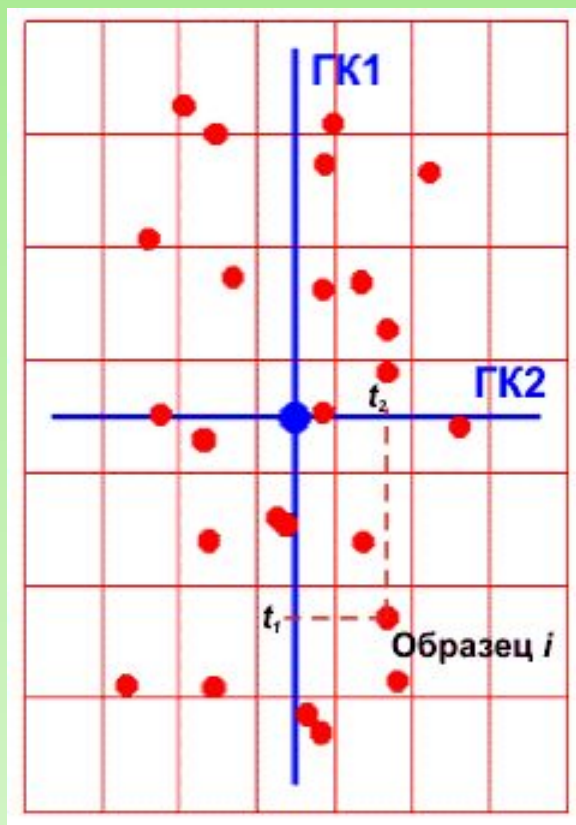
Один цикл – 800 спектров (времен)



Один массив данных – 200 образцов (циклов)



Представление данных в подпространстве



Формальные и содержательные модели

	Содержательные “Hard” models	Формальные “Soft” models
Откуда	Физика, химия,	Из данных
Формула	$y=f(\mathbf{x},\mathbf{a})+\varepsilon$	$y=\mathbf{X}\mathbf{a}+\varepsilon$
Параметры	Имеют физ. смысл	Физически бессмысленны
Проблемы	Построить модель	Обработать данные
Назначение	Экстраполяция	Интерполяция
01.04.07	Лекция в ОмГУ	15

Хемометрика

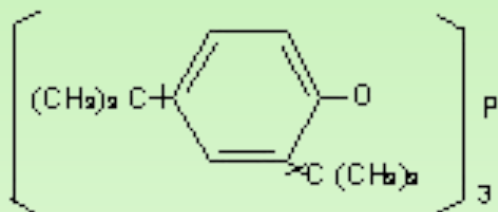
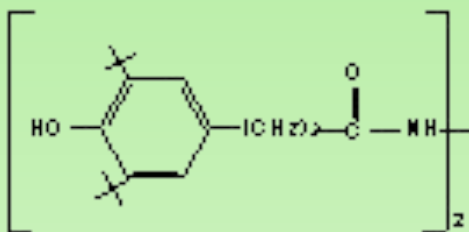
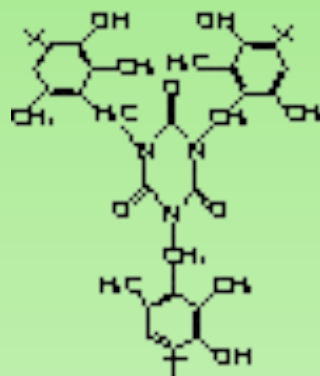
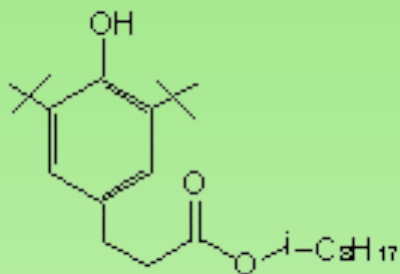


Часть 1. Калибровка



Битва снеговиков и углевиков

Антиоксиданты



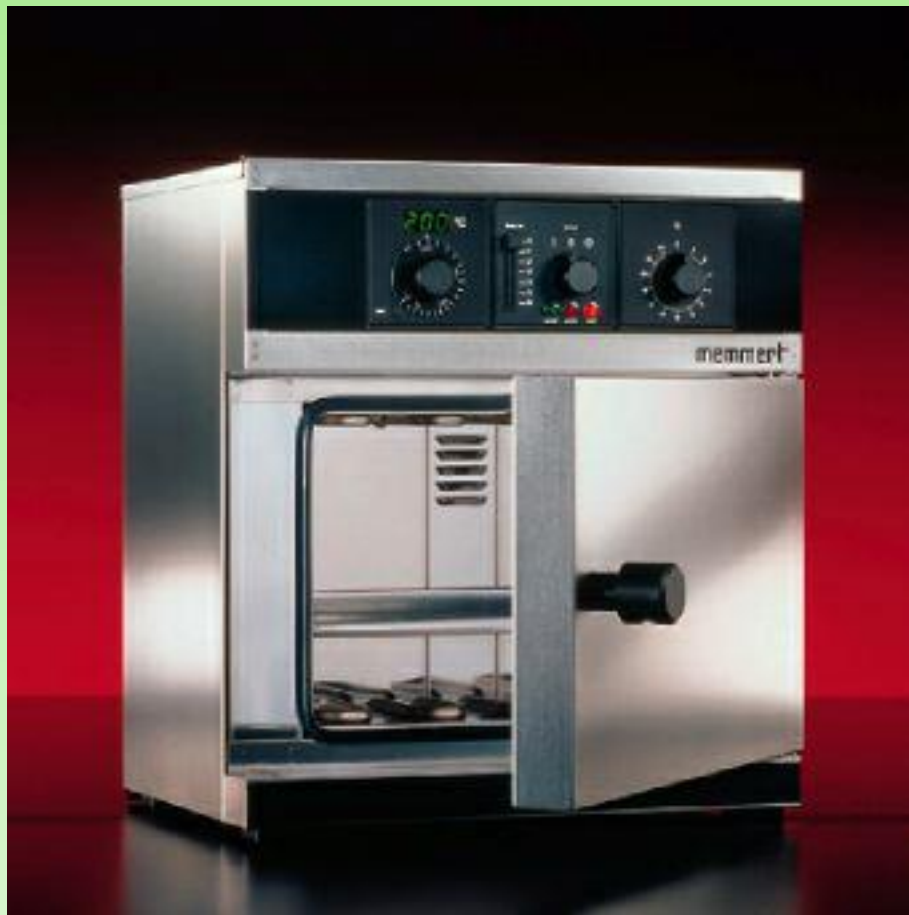
Антиоксиданты – это специальные добавки, которые замедляют термоокислительное старение полимеров.

Они защищают материал от окисления в ходе изготовления материала и при эксплуатации изделий

Задача состоит в оценке эффективности новых, перспективных химических соединений

Период индукции - OIP

Обычный метод – это длительное термическое старение

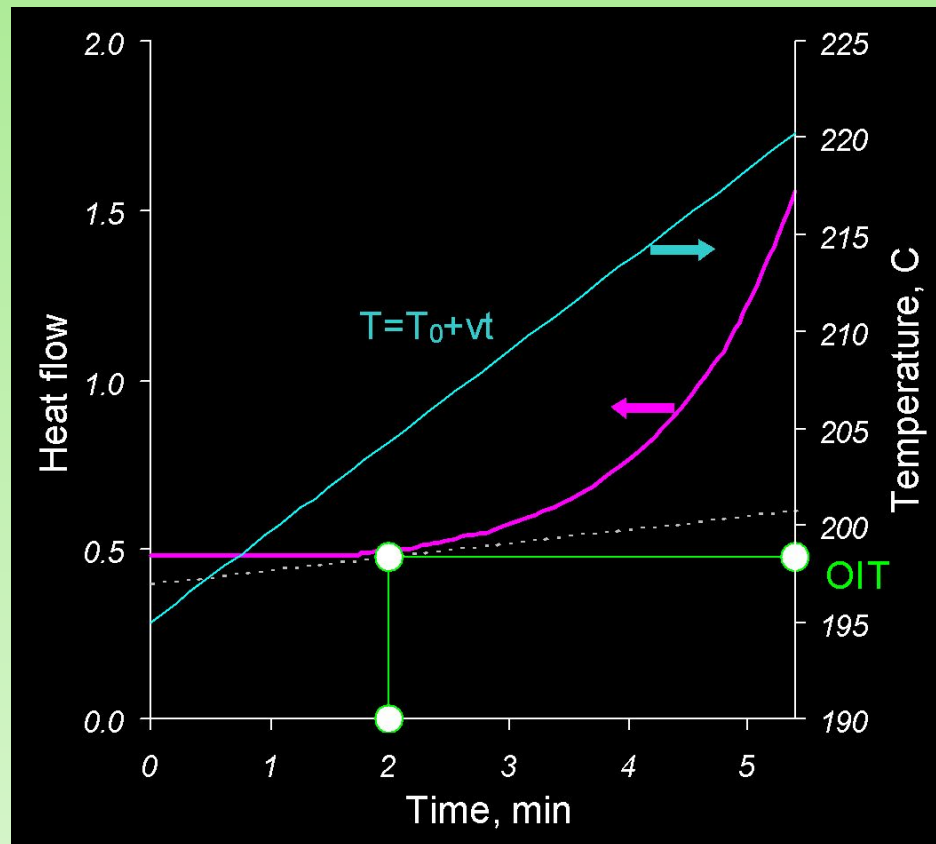


$T = 140^{\circ}\text{C}$
 $t = 1-90$ дней



Температура начала окисления - OIT

Альтернатива – это дифф. сканирующая калориметрия



Цель

Создать быстрый метод для тестирования АО

Средство

**Построить калибровку:
температура начала окисления - период индукции**

Подготовка образцов

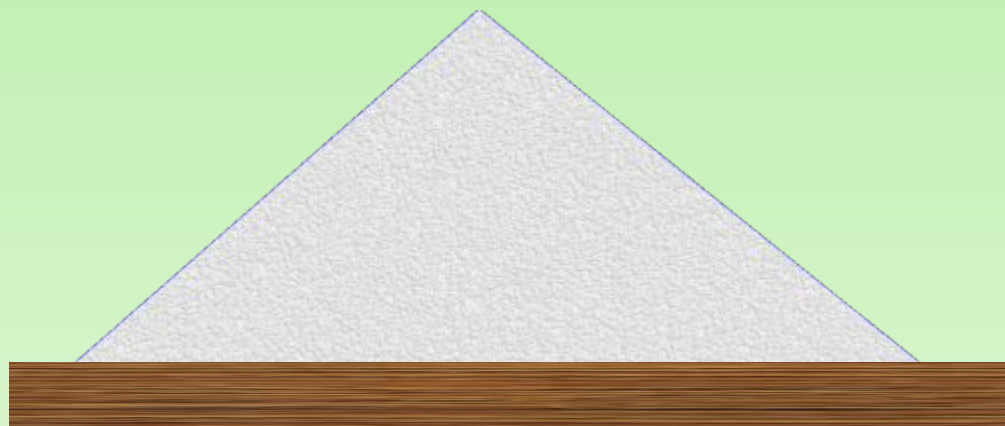
25 образцов АО



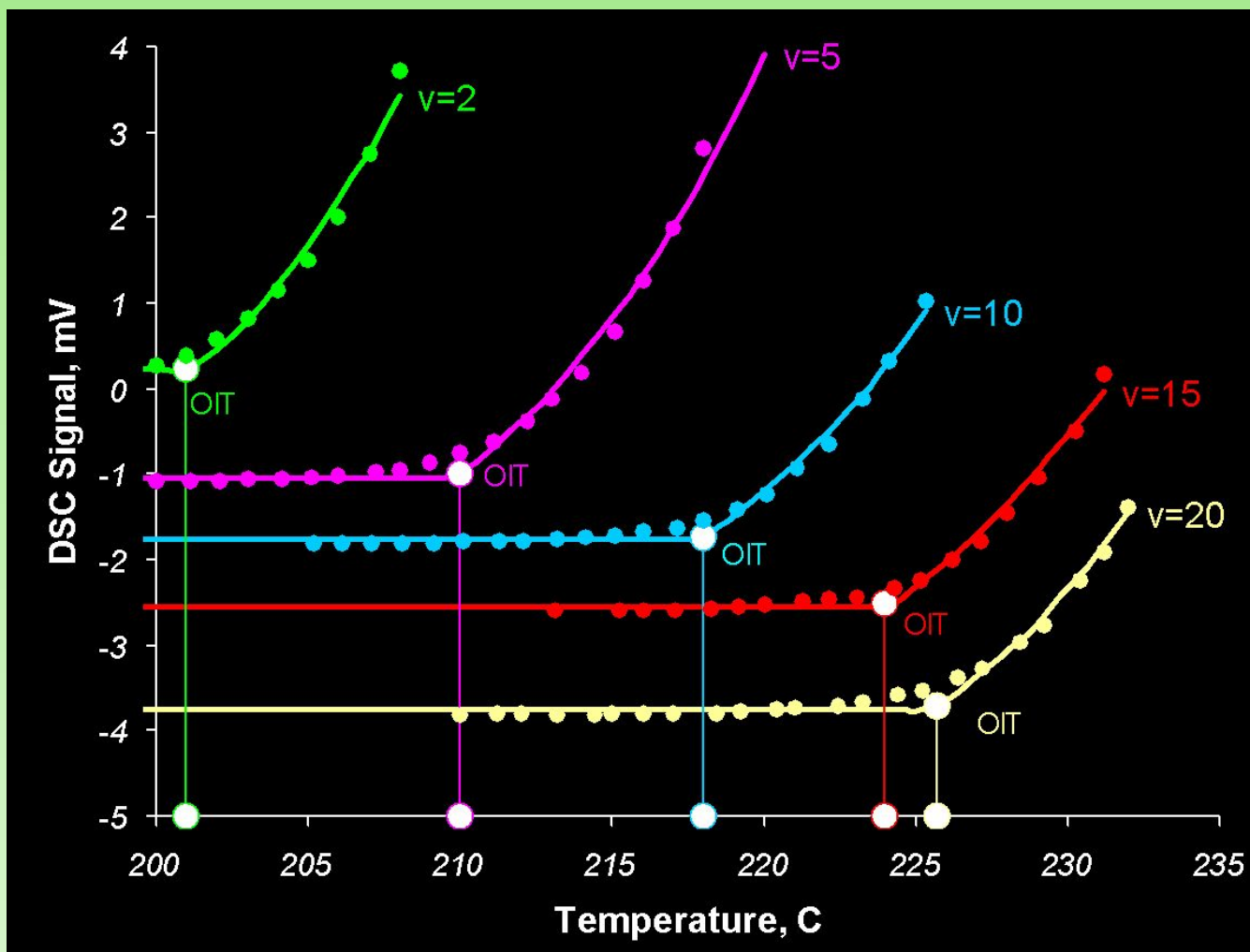
3 концентрации АО



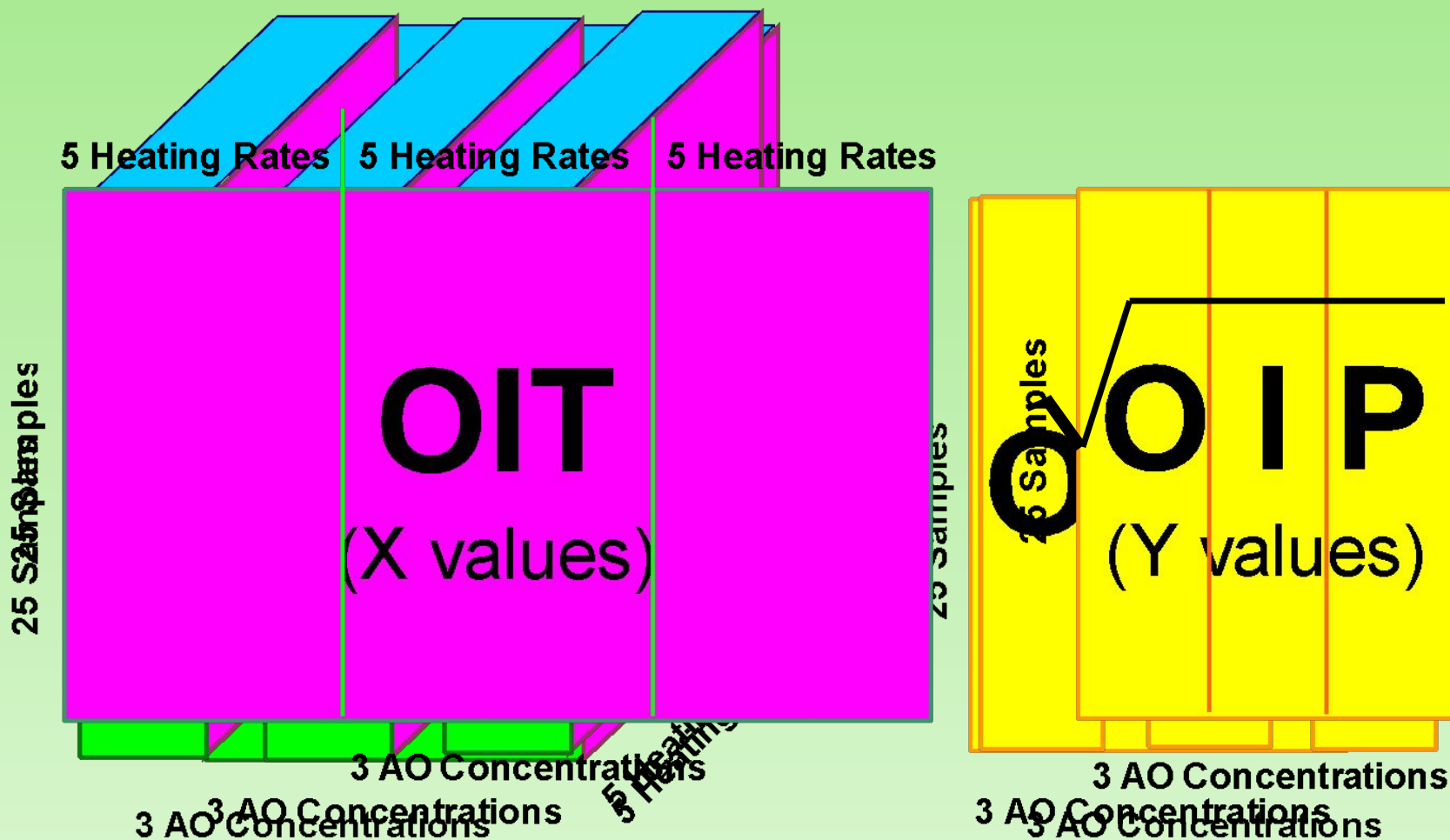
Полипропилен (PP)



DSC эксперименты. 5 скоростей нагрева



Данные и их интерпретация



PLS1 регрессия: $Xa_i = y_i$

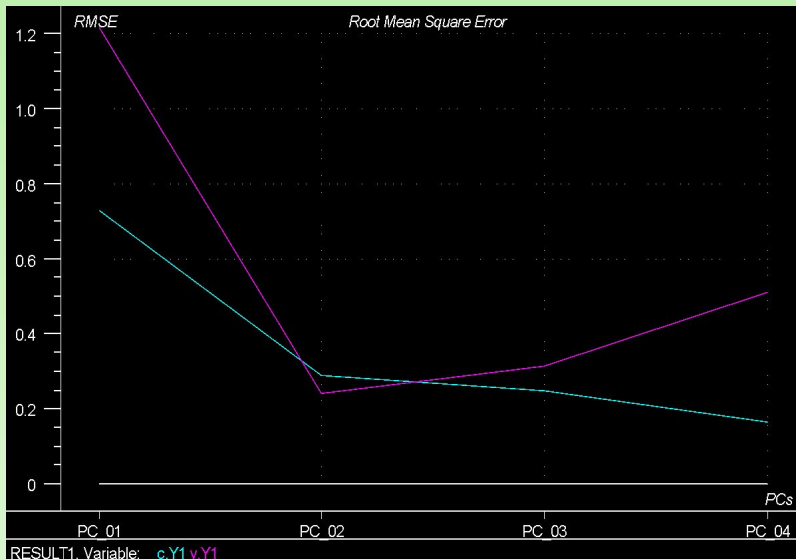
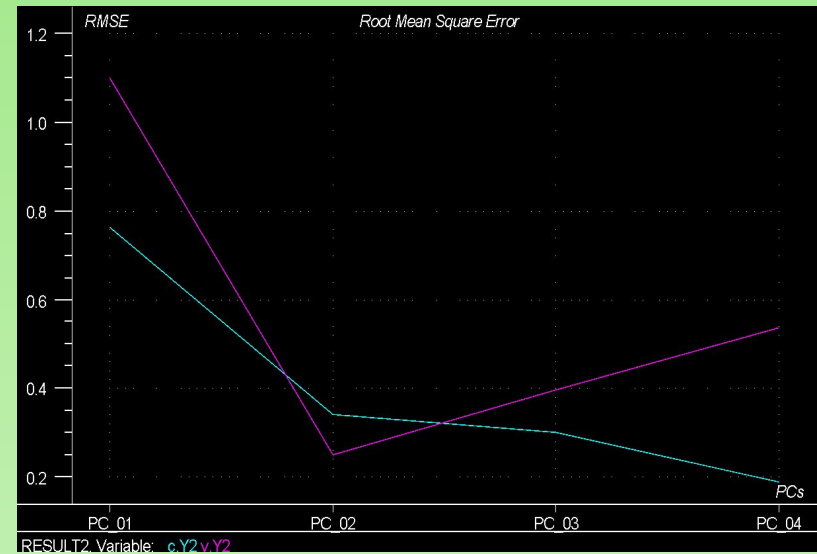
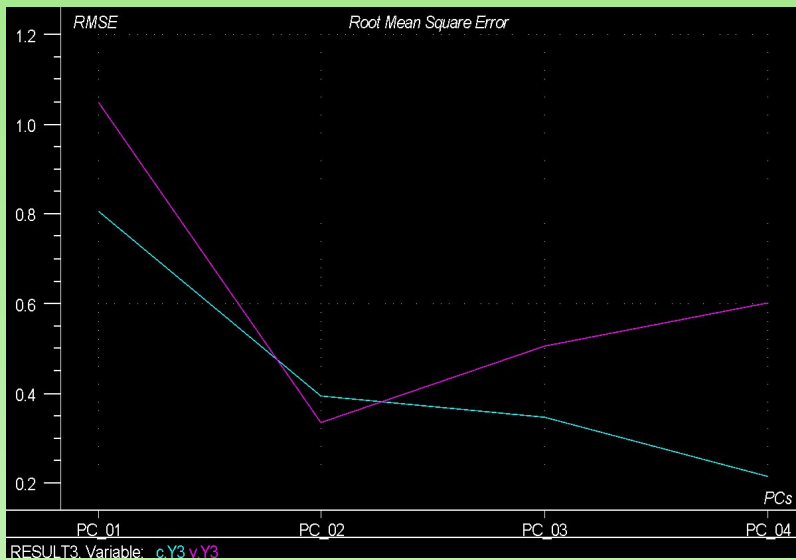
3 модели для каждой начальной концентрации АО

Samples		OIT, C															OIP ^{1/2} , day ^{1/2}		
		Init. concentration A ₀ =0.05					Init. concentration A ₀ =0.07					Init. concentration A ₀ =0.1					0.05	0.07	0.1
		v=2	v=5	v=10	v=15	v=20	v=2	v=5	v=10	v=15	v=20	v=2	v=5	v=10	v=15	v=20			
c a l i b r a t i o n	AO-1	199.1	205.6	210.7	213.2	216.0	200.6	208.5	212.6	216.5	218.5	202.4	209.9	214.8	217.9	220.4	4.1	4.4	4.7
	AO-2	198.5	204.1	210.0	212.0	214.6	199.4	206.3	211.9	215.1	217.5	202.0	208.8	214.5	217.7	218.9	4.1	4.4	4.7
	AO-3	194.6	200.7	205.4	207.2	209.0	198.1	204.9	209.3	211.5	214.3	201.3	207.6	213.4	215.9	218.0	3.9	4.2	4.5
	AO-4	194.8	200.1	206.8	210.3	212.6	196.4	204.1	210.0	213.4	216.6	198.6	206.6	213.0	216.1	219.2	1.7	1.9	2.1
	AO-5	194.1	201.3	207.1	211.3	213.7	196.6	204.9	210.9	215.1	217.5	199.8	208.1	213.7	218.7	220.2	1.9	2.2	2.5
	AO-6	193.2	201.2	206.9	210.1	213.2	197.0	204.9	210.1	214.2	217.8	199.5	207.1	214.3	218.6	219.9	1.7	2.0	2.3
	AO-7	193.8	200.4	205.6	208.4	209.5	196.8	203.4	208.8	213.5	214.3	198.5	206.2	212.8	215.7	216.9	2.2	2.4	2.6
	AO-8	193.5	201.5	205.3	208.0	211.1	195.5	203.5	208.8	212.6	214.4	199.2	205.6	212.4	215.1	217.9	2.1	2.3	2.5
	AO-9	194.6	202.0	207.1	209.8	212.2	197.8	204.4	210.1	213.5	215.3	200.1	206.5	212.9	216.7	218.8	2.3	2.5	2.7
	AO-10	179.5	187.2	192.8	194.7	197.4	185.6	193.3	198.4	201.5	203.9	189.1	196.3	202.2	205.3	207.3	0.7	1.3	1.9
	AO-11	178.4	185.9	189.8	193.8	195.7	186.1	193.5	197.9	202.4	203.3	189.0	197.9	202.1	206.0	208.6	0.8	1.2	1.7
	AO-12	181.4	188.8	192.7	195.9	197.8	187.1	194.3	199.1	202.6	205.8	191.1	198.4	202.7	206.1	209.4	1.0	1.5	2.1
	AO-13	194.0	199.7	205.5	208.3	210.7	196.7	204.6	209.8	213.4	215.8	200.2	207.5	214.0	218.6	220.0	2.1	2.4	2.7
	AO-14	195.2	201.3	207.4	209.6	212.7	198.6	204.8	210.9	214.5	217.7	200.1	208.2	213.9	218.3	220.4	2.4	2.7	3.0
	AO-15	193.8	200.8	206.2	210.1	211.7	196.4	203.4	209.5	214.5	215.9	199.0	207.0	212.7	216.9	219.0	2.0	2.2	2.4
	AO-16	189.9	197.9	202.7	207.4	209.6	193.9	202.0	208.1	211.5	213.4	195.9	204.8	211.3	214.6	218.0	1.4	1.6	1.8
	AO-17	193.4	201.6	206.9	210.4	213.7	195.9	204.5	211.3	214.7	217.3	199.3	206.9	214.2	217.1	219.8	1.6	1.9	2.2
	AO-18	193.9	200.9	207.7	210.9	213.4	197.5	205.6	211.5	215.4	218.4	200.9	209.0	216.0	219.0	223.0	1.6	1.9	2.2
t e s t	AO-19	200.5	206.8	212.1	214.2	217.3	202.7	209.8	214.7	217.8	220.8	204.0	212.2	216.5	220.3	224.0	3.2	3.5	3.8
	AO-20	193.3	201.2	206.5	209.3	211.9	196.7	203.6	210.4	213.8	216.5	198.5	206.1	214.0	216.3	218.8	1.9	2.2	2.5
	AO-21	189.9	196.3	202.3	205.4	208.3	192.9	201.3	207.1	210.7	214.1	194.9	204.0	210.4	213.6	216.3	1.3	1.5	1.7
	AO-22	180.7	186.5	192.1	194.4	197.5	186.6	194.6	199.5	202.8	204.8	191.1	198.8	205.1	207.5	210.6	0.9	1.2	1.5
	AO-23	193.3	200.0	206.5	210.1	212.0	195.9	203.2	209.9	212.3	215.7	197.2	205.2	212.0	215.8	218.2	1.9	2.1	2.3
	AO-24	194.4	201.5	208.3	211.8	215.1	197.4	205.5	212.4	216.0	218.1	200.4	208.3	215.9	218.5	221.8	1.4	1.9	2.3
	AO-25	200.6	209.7	218.6	222.8	226.8	202.0	212.3	220.2	224.8	227.6	203.5	213.4	221.1	226.2	229.6	0.7	1.1	1.5

X

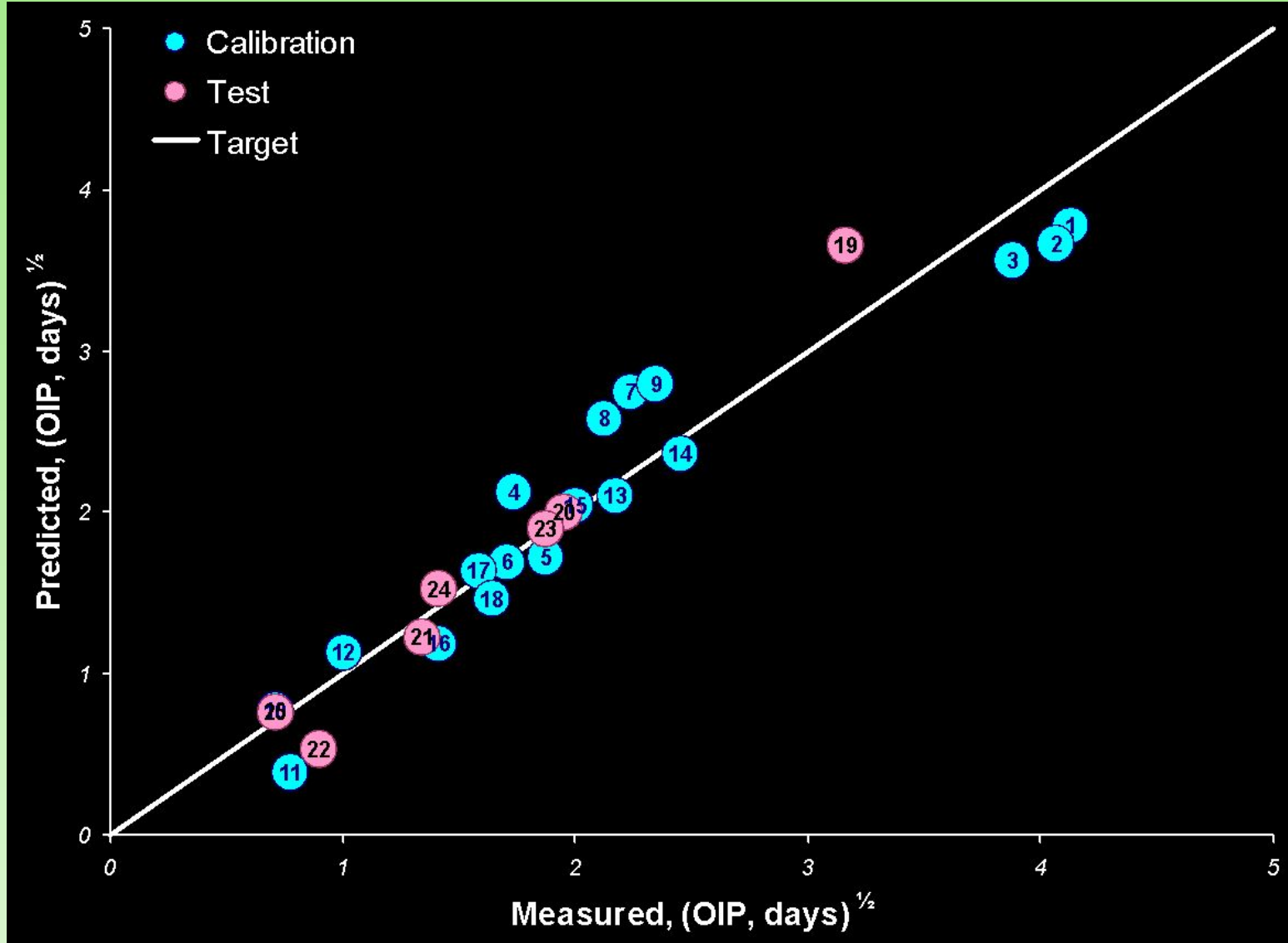
y_1, y_2, y_3

Построение модели. Три набора данных



A_0	X_{expl}	Y_{exp}	RMSEC	r^2 (cal)	RMSEP	r^2 (test)	β
0.05	99%	92%	0.29	0.96	0.24	0.99	0.8
0.07	99%	88%	0.34	0.93	0.25	0.99	1.0

Прогноз периода индукции для $A_0 = 0.05$



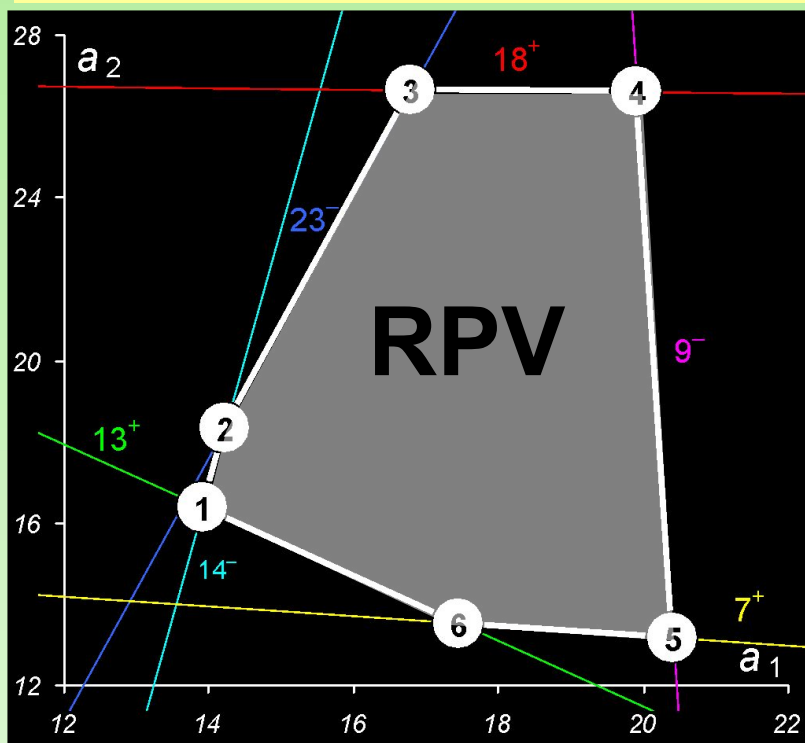
SIC Метод

Все погрешности ограничены!

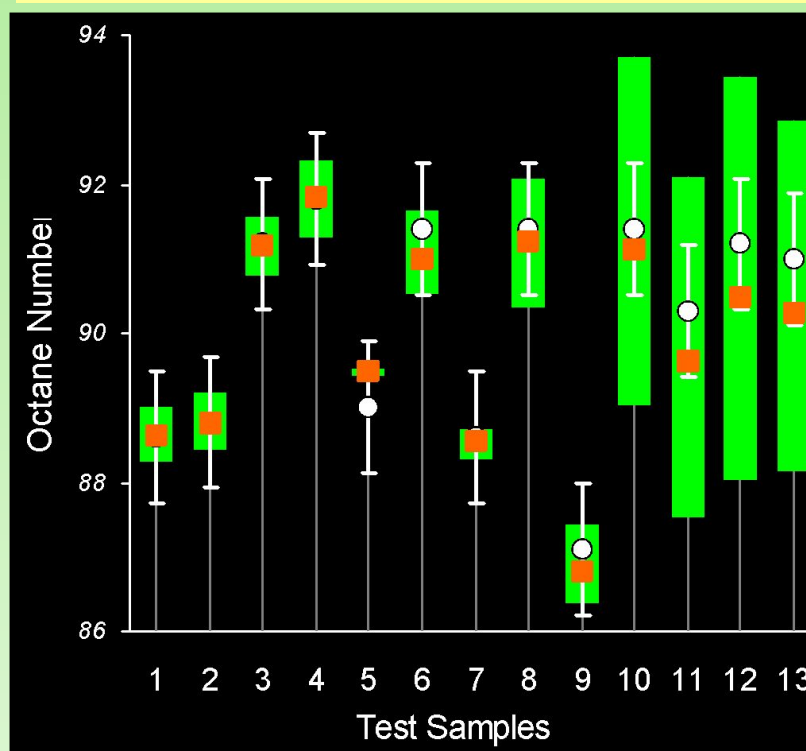
Существует максимальное отклонение, β , такое что

$$\text{Prob}\{|\varepsilon| > \beta\} = 0$$

Область допустимых значений



Прогнозные интервалы: SIC & PLS



Простое интервальное оценивание

Пятница 6 апреля

11:30 ауд. 208

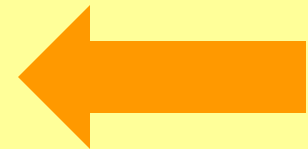
О.Е. Родионова

**Простой взгляд на простое
интервальное оценивание**

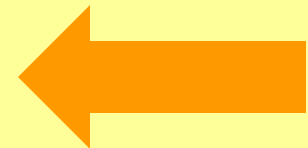
Прогнозные интервалы для $A_0=0.05$



**Проверочные
образцы
АО19-АО25**

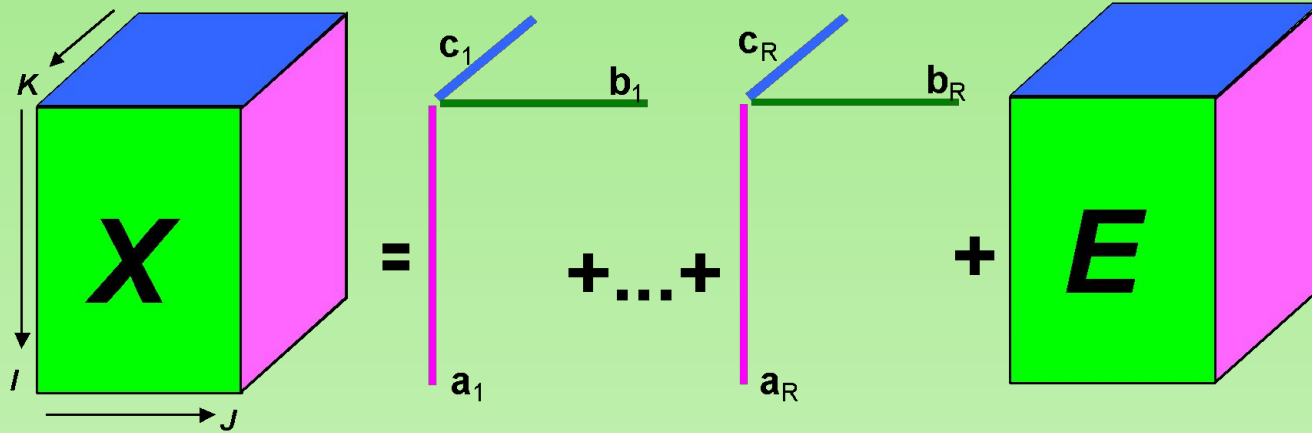


**Обучающие
образцы
АО1-АО18**

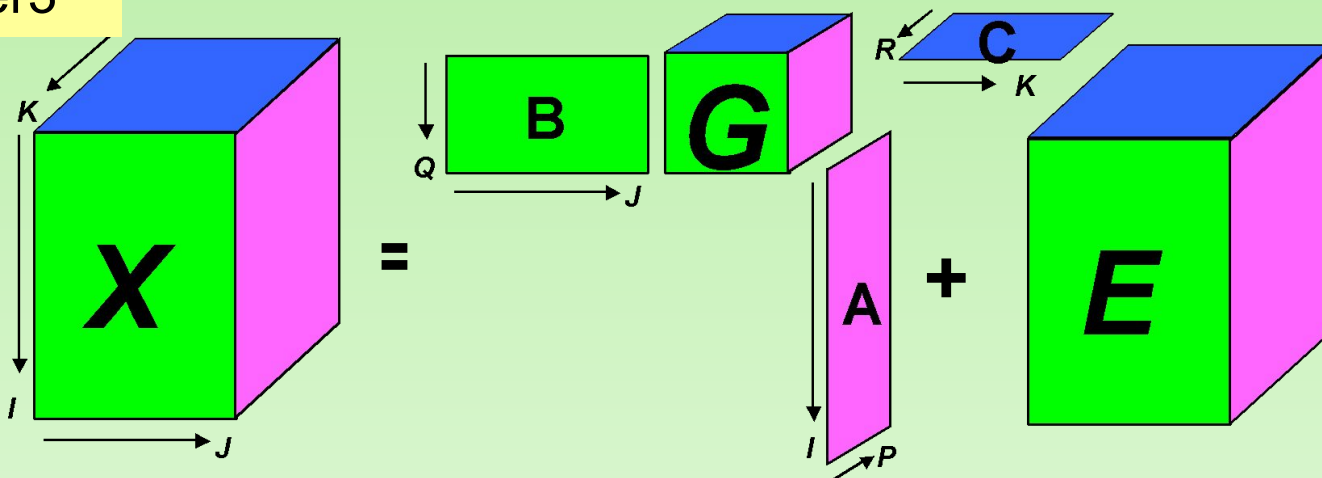


N-way методы

PARAFAC



Tucker3



Часть 2 . Анализ изображений



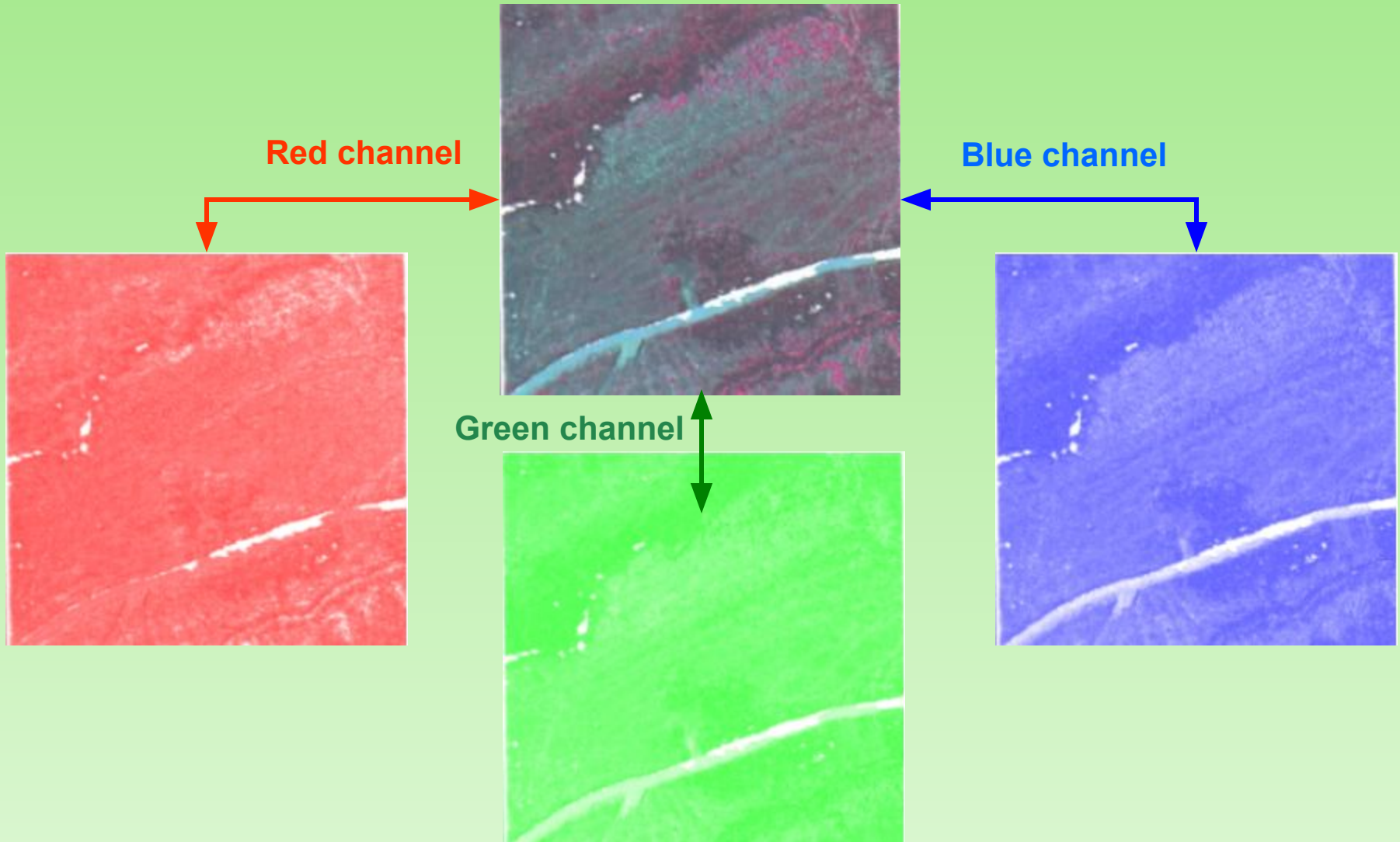
Still Life

01.04.07

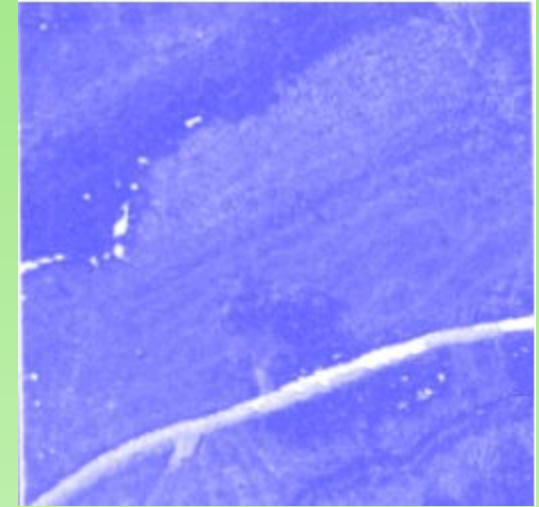
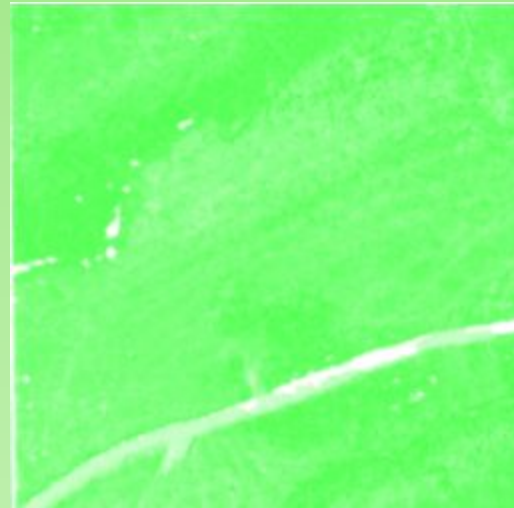
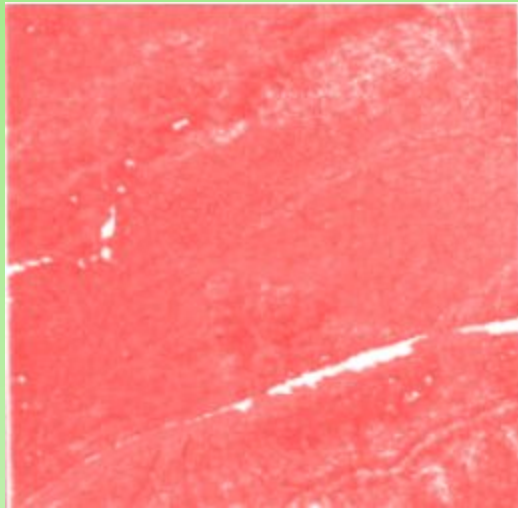
Лекция в ОмГУ

32

Разложение изображения по каналам



Оцифровка изображения



0.74	0.39	0.24	0.07	0.68	0.80	0.83	0.87	0.29	0.86	0.19	0.06
0.41	0.26	0.37	0.03	0.12	0.84	0.09	0.88	0.64	0.47	0.74	0.62
0.51	0.96	0.18	0.83	0.89	0.96	0.12	0.64	0.11	0.84	0.62	1.00
0.11	0.97	0.43	0.32	0.51	0.47	0.49	0.97	1.00	0.32	0.73	0.64
0.96	0.39	0.96	0.29	0.28	0.40	0.24	0.69	0.61	0.73	0.02	0.30
0.74	0.69	0.57	0.80	0.66	0.93	0.89	0.85	0.81	0.12	0.38	0.99
0.14	0.69	0.80	0.84	0.01	0.13	0.03	0.54	0.91	0.71	0.28	0.30
0.73	0.69	0.26	0.37	0.44	0.13	0.03	0.09	0.99	0.99	0.46	0.27
0.83	0.16	0.13	0.66	0.61	0.93	0.91	0.35	0.07	0.42	0.64	0.18
0.11	0.21	0.14	0.98	0.96	0.70	0.62	0.63	0.79	0.38	0.42	0.33
0.77	0.04	0.72	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.36
0.83	0.88	0.36	0.41	0.32	0.16	0.66	0.40	0.97	0.96	0.71	0.82
0.02	0.16	0.22	0.02	0.47	0.69	0.21	0.63	0.61	0.13	0.16	0.98
0.02	0.89	0.70	0.62	0.90	0.86	0.92	0.71	0.14	0.24	0.48	0.32
0.12	0.88	0.80	0.76	0.66	0.67	0.31	0.29	0.87	0.83	0.89	0.08
0.20	0.87	0.51	0.68	0.10	0.46	0.20	0.24	0.76	0.31	0.54	0.36
0.34	0.20	0.68	0.77	0.23	0.06	0.71	1.00	0.76	0.36	0.49	0.36
0.60	0.60	0.62	0.18	0.72	0.26	0.41	0.02	0.81	0.82	0.81	0.42
0.80	0.69	0.23	0.83	0.78	0.06	0.83	1.00	0.33	0.82	0.19	0.03
0.42	0.79	0.76	0.76	0.46	0.46	0.66	0.62	0.69	0.66	0.89	0.93
0.36	0.66	0.21	0.90	0.68	0.83	0.66	0.10	0.12	0.82	0.66	0.26
0.83	0.66	0.32	0.96	0.29	0.04	0.16	0.19	0.26	0.16	0.68	0.41
0.16	0.46	0.69	0.84	0.43	0.66	0.84	0.68	0.61	0.88	0.62	0.83
0.42	0.04	0.16	0.29	0.36	0.33	0.46	0.61	0.16	0.11	0.72	0.86
0.12	0.70	0.00	0.10	0.26	0.94	0.30	0.61	0.29	0.40	0.82	0.66
0.86	0.48	0.06	0.11	0.17	0.46	0.51	0.31	0.07	0.71	0.46	0.26
0.06	0.18	0.44	0.90	0.60	0.66	0.61	0.21	0.11	0.00	0.82	
0.19	0.46	0.06	0.76	0.93	0.68	0.28	0.69	0.82	0.42	0.63	0.66
0.79	0.18	0.47	0.26	0.77	0.82	0.07	0.10	0.44	0.09	0.57	0.74
0.28	0.17	0.61	0.33	0.73	0.76	0.62	0.83	0.68	0.07	0.33	0.76

0.74	0.39	0.24	0.07	0.68	0.80	0.83	0.87	0.29	0.86	0.19	0.06
0.41	0.26	0.37	0.03	0.12	0.84	0.09	0.88	0.64	0.47	0.74	0.62
0.51	0.96	0.18	0.83	0.89	0.96	0.12	0.64	0.11	0.84	0.62	1.00
0.11	0.97	0.43	0.32	0.51	0.47	0.49	0.97	1.00	0.32	0.73	0.64
0.96	0.39	0.96	0.29	0.28	0.40	0.24	0.69	0.61	0.73	0.02	0.30
0.74	0.69	0.57	0.80	0.66	0.93	0.89	0.85	0.81	0.12	0.38	0.99
0.14	0.69	0.80	0.84	0.01	0.13	0.03	0.54	0.91	0.71	0.28	0.30
0.73	0.69	0.26	0.37	0.44	0.13	0.03	0.09	0.99	0.99	0.46	0.27
0.83	0.16	0.13	0.66	0.61	0.93	0.91	0.35	0.07	0.42	0.64	0.18
0.11	0.21	0.14	0.98	0.96	0.70	0.62	0.63	0.79	0.38	0.42	0.33
0.77	0.04	0.72	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.36
0.83	0.88	0.36	0.41	0.32	0.16	0.66	0.40	0.97	0.96	0.71	0.82
0.02	0.16	0.22	0.02	0.47	0.69	0.21	0.63	0.61	0.13	0.16	0.98
0.02	0.89	0.70	0.62	0.90	0.86	0.92	0.71	0.14	0.24	0.48	0.32
0.12	0.88	0.80	0.76	0.66	0.67	0.31	0.29	0.87	0.83	0.89	0.08
0.20	0.87	0.51	0.68	0.10	0.46	0.20	0.24	0.76	0.31	0.54	0.36
0.34	0.20	0.68	0.77	0.23	0.06	0.71	1.00	0.76	0.36	0.49	0.36
0.60	0.60	0.62	0.18	0.72	0.26	0.41	0.02	0.81	0.82	0.81	0.42
0.80	0.69	0.23	0.83	0.78	0.06	0.83	1.00	0.33	0.82	0.19	0.03
0.42	0.79	0.76	0.76	0.46	0.46	0.66	0.62	0.69	0.66	0.89	0.93
0.36	0.66	0.21	0.90	0.68	0.83	0.66	0.10	0.12	0.82	0.66	0.26
0.83	0.66	0.32	0.96	0.29	0.04	0.16	0.19	0.26	0.16	0.68	0.41
0.16	0.46	0.69	0.84	0.43	0.66	0.84	0.68	0.61	0.88	0.62	0.83
0.42	0.04	0.16	0.29	0.36	0.33	0.46	0.61	0.16	0.11	0.72	0.86
0.12	0.70	0.00	0.10	0.26	0.94	0.30	0.61	0.29	0.40	0.82	0.66
0.86	0.48	0.06	0.11	0.17	0.46	0.51	0.31	0.07	0.71	0.46	0.26
0.06	0.18	0.44	0.90	0.60	0.66	0.61	0.21	0.11	0.00	0.82	
0.19	0.46	0.06	0.76	0.93	0.68	0.28	0.69	0.82	0.42	0.63	0.66
0.79	0.18	0.47	0.26	0.77	0.82	0.07	0.10	0.44	0.09	0.57	0.74
0.28	0.17	0.61	0.33	0.73	0.76	0.62	0.83	0.68	0.07	0.33	0.76

0.74	0.39	0.24	0.07	0.68	0.80	0.83	0.87	0.29	0.86	0.19	0.06
0.41	0.26	0.37	0.03	0.12	0.84	0.09	0.88	0.64	0.47	0.74	0.62
0.51	0.96	0.18	0.83	0.89	0.96	0.12	0.64	0.11	0.84	0.62	1.00
0.11	0.97	0.43	0.32	0.51	0.47	0.49	0.97	1.00	0.32	0.73	0.64
0.96	0.39	0.96	0.29	0.28	0.40	0.24	0.69	0.61	0.73	0.02	0.30
0.74	0.69	0.57	0.80	0.66	0.93	0.89	0.85	0.81	0.12	0.38	0.99
0.14	0.69	0.80	0.84	0.01	0.13	0.03	0.54	0.91	0.71	0.28	0.30
0.73	0.69	0.26	0.37	0.44	0.13	0.03	0.09	0.99	0.99	0.46	0.27
0.83	0.16	0.13	0.66	0.61	0.93	0.91	0.35	0.07	0.42	0.64	0.18
0.11	0.21	0.14	0.98	0.96	0.70	0.62	0.63	0.79	0.38	0.42	0.33
0.77	0.04	0.72	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.66	0.36
0.83	0.88	0.36	0.41	0.32	0.16	0.66	0.40	0.97	0.96	0.71	0.82
0.02	0.16	0.22	0.02	0.47	0.69	0.21	0.63	0.61	0.13	0.16	0.98
0.02	0.89	0.70	0.62	0.90	0.86	0.92	0.71	0.14	0.24	0.48	0.32
0.12	0.88	0.80	0.76	0.66	0.67	0.31	0.29	0.87	0.83	0.89	0.08
0.20	0.87	0.51	0.68	0.10	0.46	0.20	0.24	0.76	0.31	0.54	0.36
0.34	0.20	0.68	0.77	0.23	0.06	0.71	1.00	0.76	0.36	0.49	0.36
0.60	0.60	0.62	0.18	0.72	0.26	0.41	0.02	0.81	0.82	0.81	0.42
0.80	0.69	0.23	0.83	0.78	0.06	0.83	1.00	0.33	0.82	0.19	0.03
0.42	0.79	0.76	0.76	0.46	0.46	0.66	0.62	0.69	0.66	0.89	0.93
0.36	0.66	0.21	0.90	0.68	0.83	0.66	0.10	0.12	0.82	0.66	0.26
0.83	0.66	0.32	0.96	0.29	0.04	0.16	0.19	0.26	0.16	0.68	0.41
0.16	0.46	0.69	0.84	0.43	0.66	0.84	0.68	0.61	0.88	0.62	0.83
0.42	0.04	0.16	0.29	0.36	0.33	0.46	0.61	0.16	0.11	0.72	0.86
0.12	0.70	0.00	0.10	0.26	0.94	0.30	0.61	0.29	0.40	0.82	0.66
0.86	0.48	0.06	0.11	0.17	0.46	0.51	0.31	0.07	0.71	0.46	0.26
0.06	0.18	0.44	0.90	0.60	0.66	0.61	0.21	0.11	0.00	0.82	
0.19	0.46	0.06	0.76	0.93	0.68	0.28	0.69	0.82	0.42	0.63	0.66
0.79	0.18	0.47	0.26	0.77	0.82	0.07	0.10	0.44	0.09	0.57	0.74
0.28	0.17	0.61	0.33	0.73	0.76	0.62	0.83	0.68	0.07	0.33	0.76

Исследование состояния лесов (Канада)



Часть 3. Экология



Пухтоград

Счета и нагрузки. Свалка Безенчук

- возраст ~ 25 лет
- объем ~ 90 000 м³
- площадь ~ 60 000 м²
- состав отходов:
 1. бытовые
 2. аграрные
 3. промышленные



E.V. Mikhailov et al

Ecological assessment of landfills with multivariate analysis — a feasibility study, *Chemom. Intell. Lab. Syst.*, WSC-5 issue, In Press

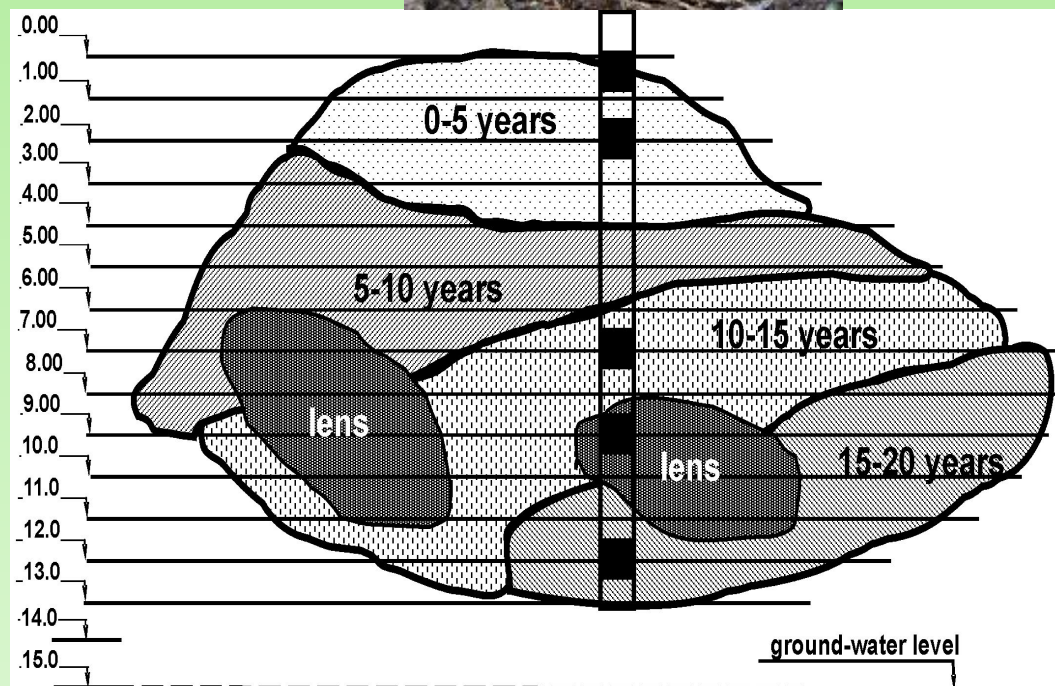
Отбор проб

Измеряемые

Глубина (Depth)
Зольность (Ash)
Влажность (Humidity)
Об. вес (Weight)
 t° летом (Summer T)
 t° зимой (Winter T)

Оцениваемые

Линза (Lens = ± 1)
Основа (Base = ± 1)



Матрица данных

восемь переменных: $j=1\dots J$

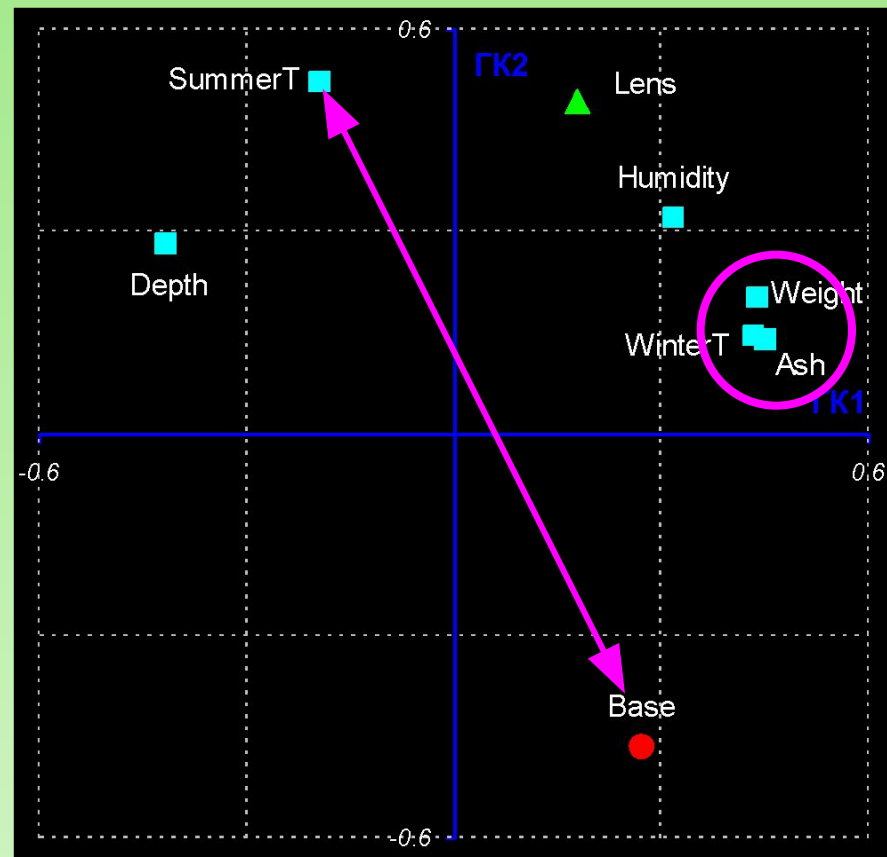
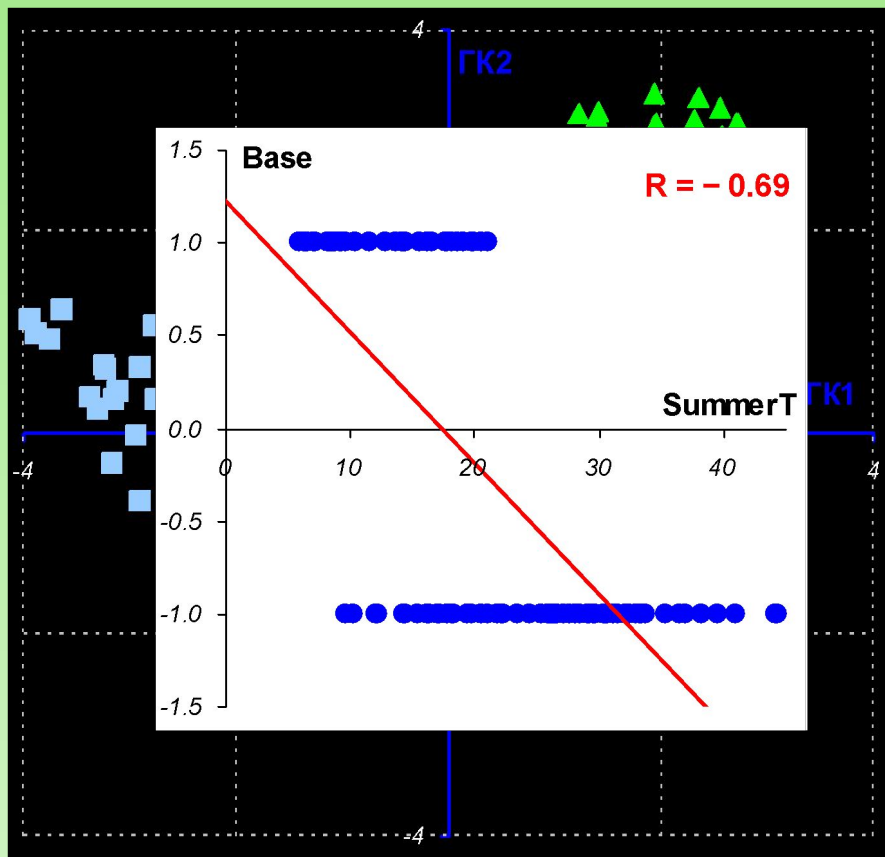
123 образца: $i=1,\dots,I$

$$\mathbf{X} = \{X_{ij}\}$$

$\bar{\mathbf{X}}$

s^2

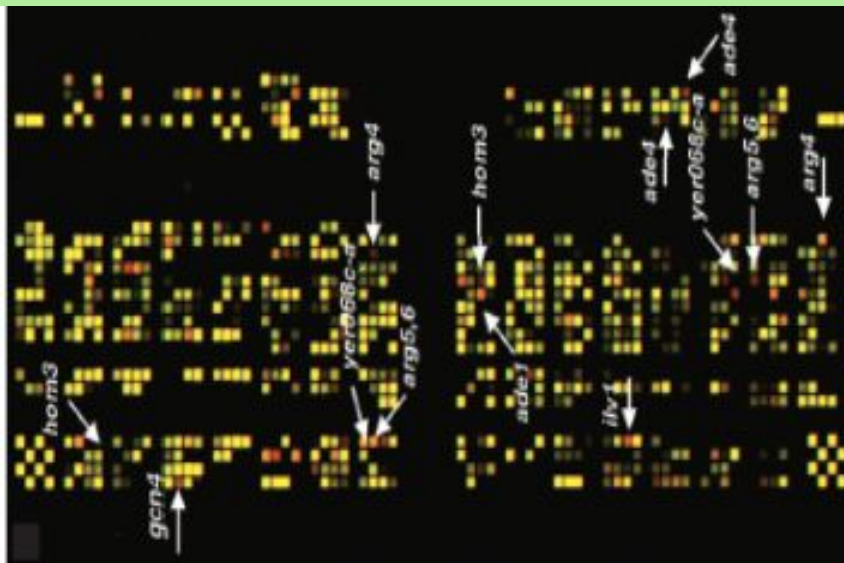
Счета и нагрузки



▲ ЛИНЗЫ

● ОСНОВА

Часть 4. Биология



Новая биология: X- omics

Традиционный подход:

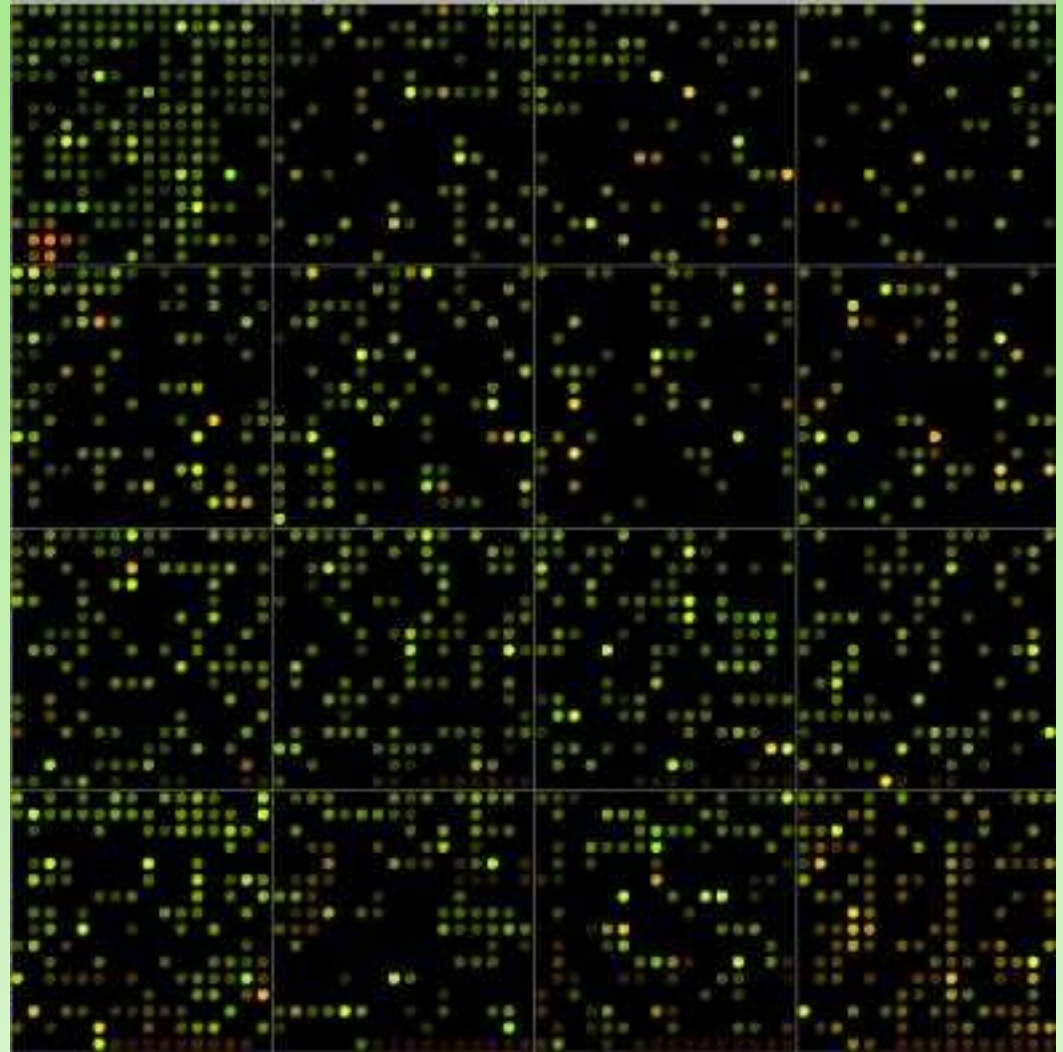
- Один ген/белок/реакция за раз
- Исследование отдельной модели

Новый «системный» подход:

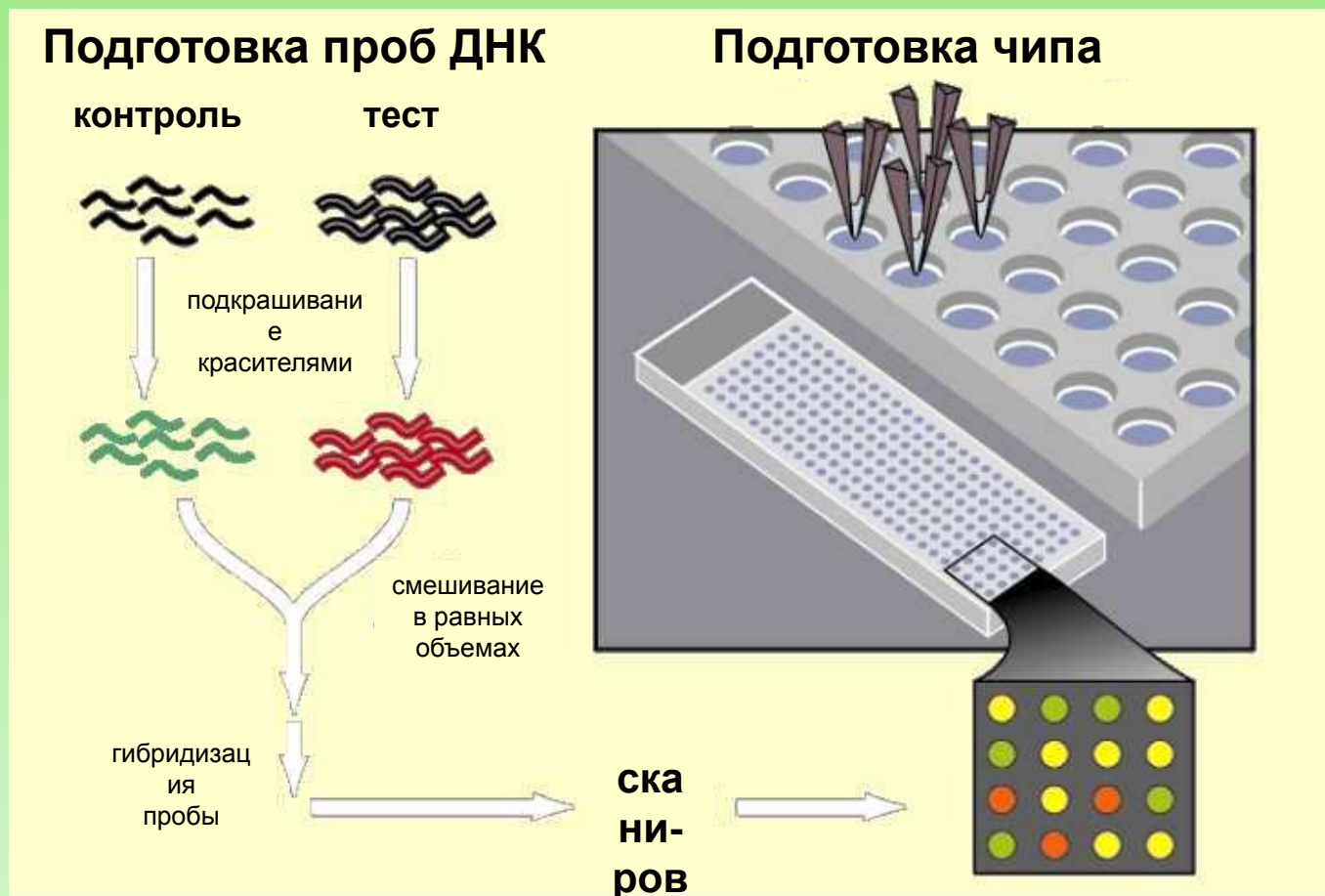
- Все ДНК/РНК/белки исследуются вместе
- Глобальный анализ больших данных

Что такое microarray - чип?

- Каждый микрореактор содержит ДНК (или белок)
- Обычно в каждом реакторе разные ДНК, с повторами для контроля точности
- За один раз несколько микрореакторов гибридизируются с экстрактами тканей
- Результат виден как образ, состоящий из множества цветных пятен

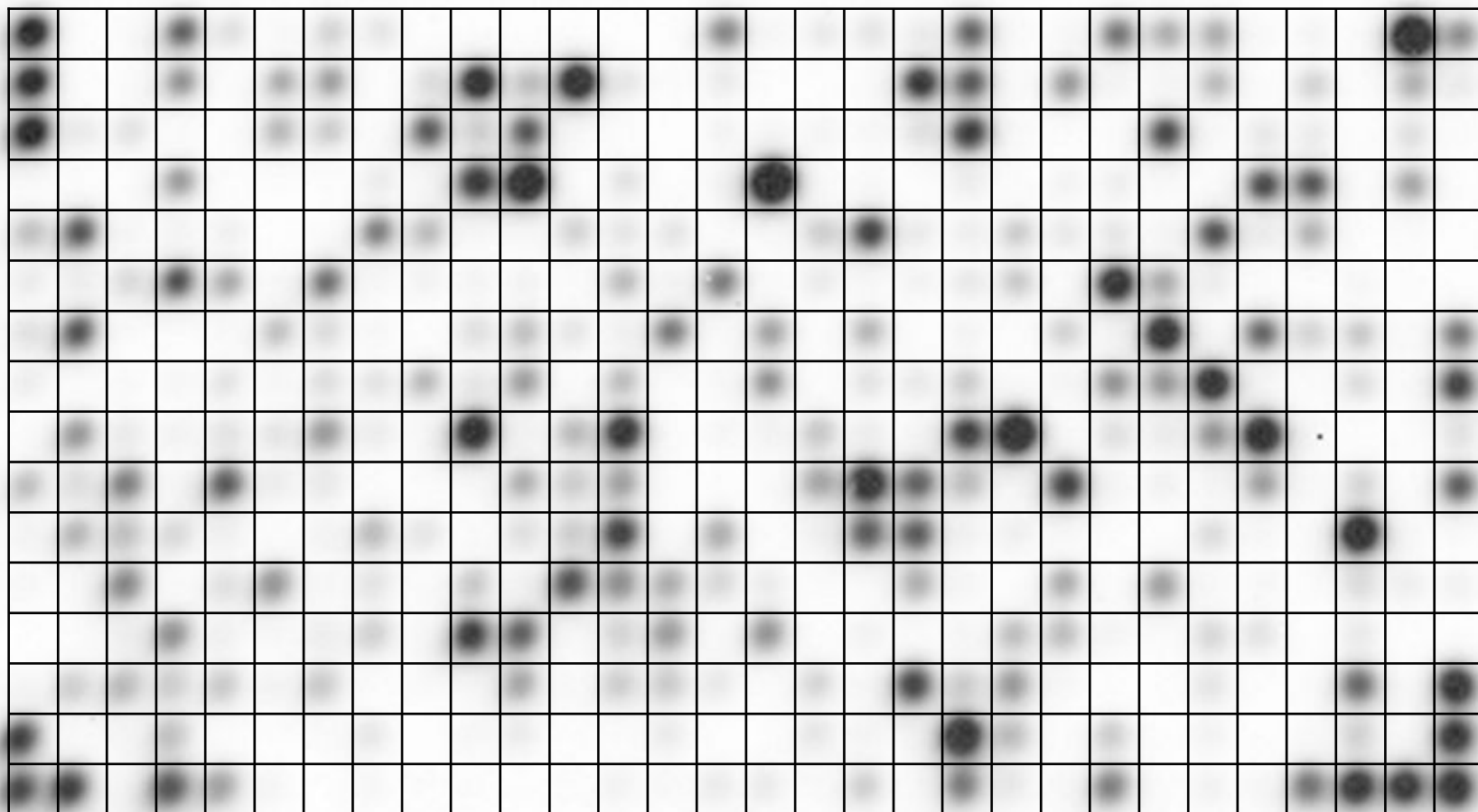


Эксперименты с ДНК

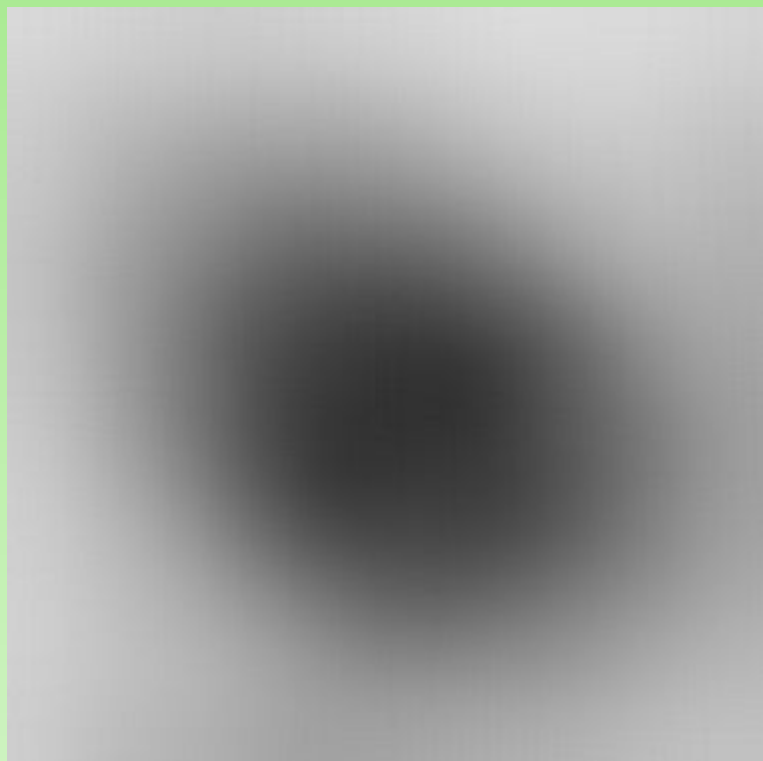


Сравниваются две пробы ДНК. Они подкрашиваются разными красителями
Результат эксперимента это интенсивности через два фильтра - красный и зеленый

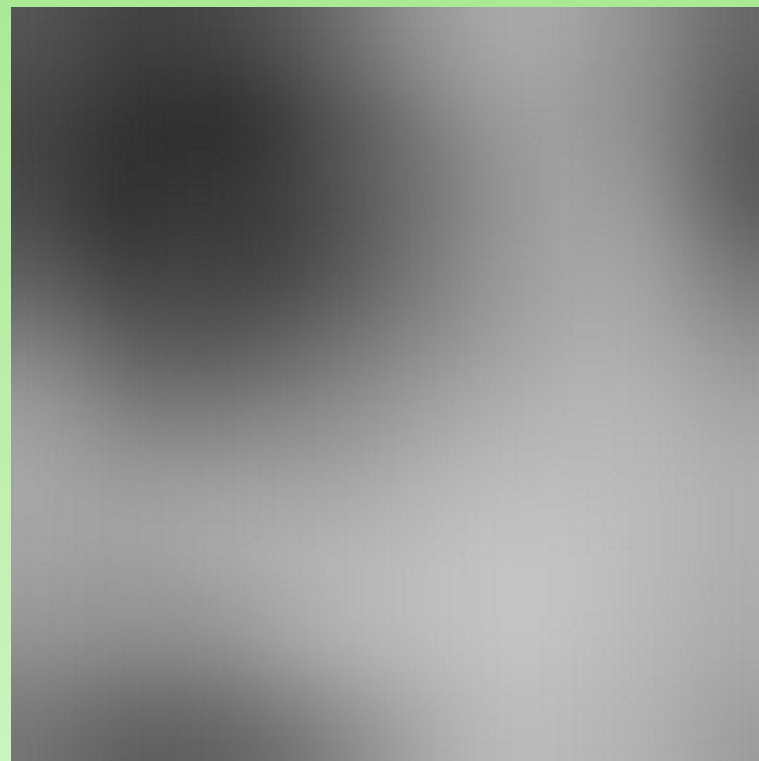
Сегментация



Результат сегментирования

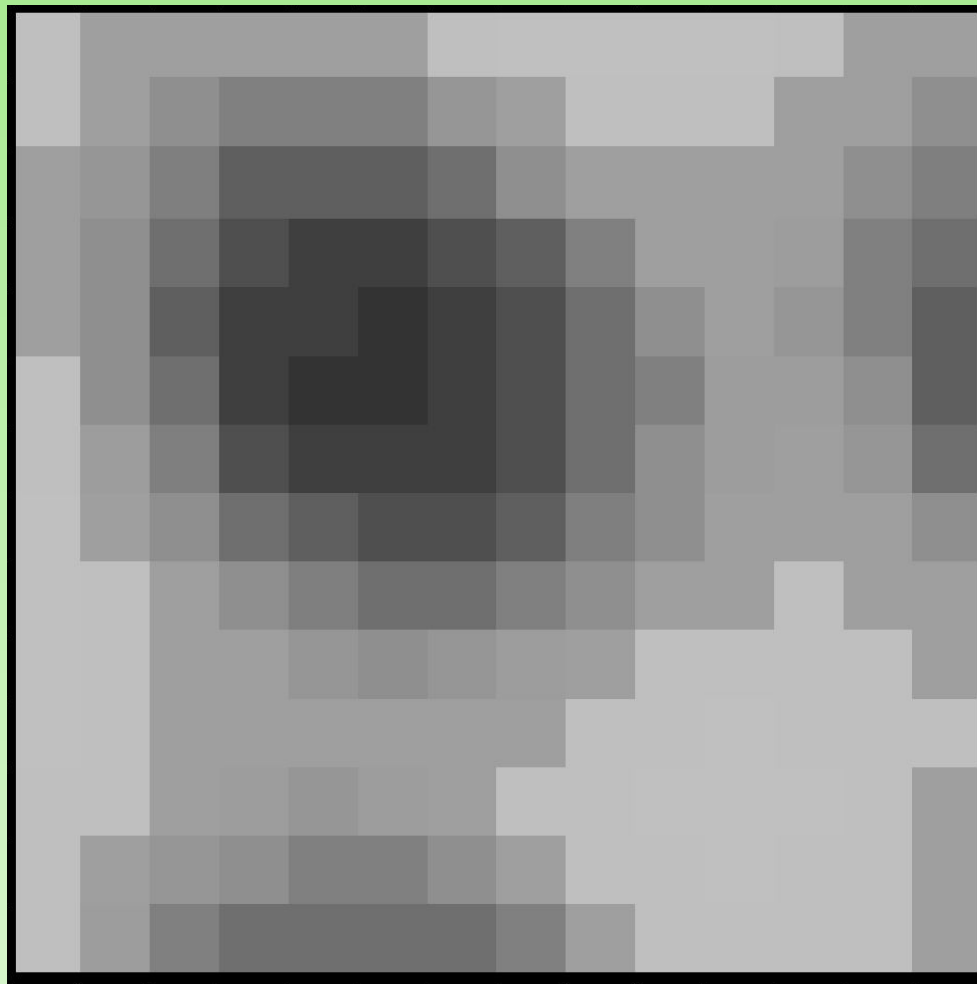


Хороший

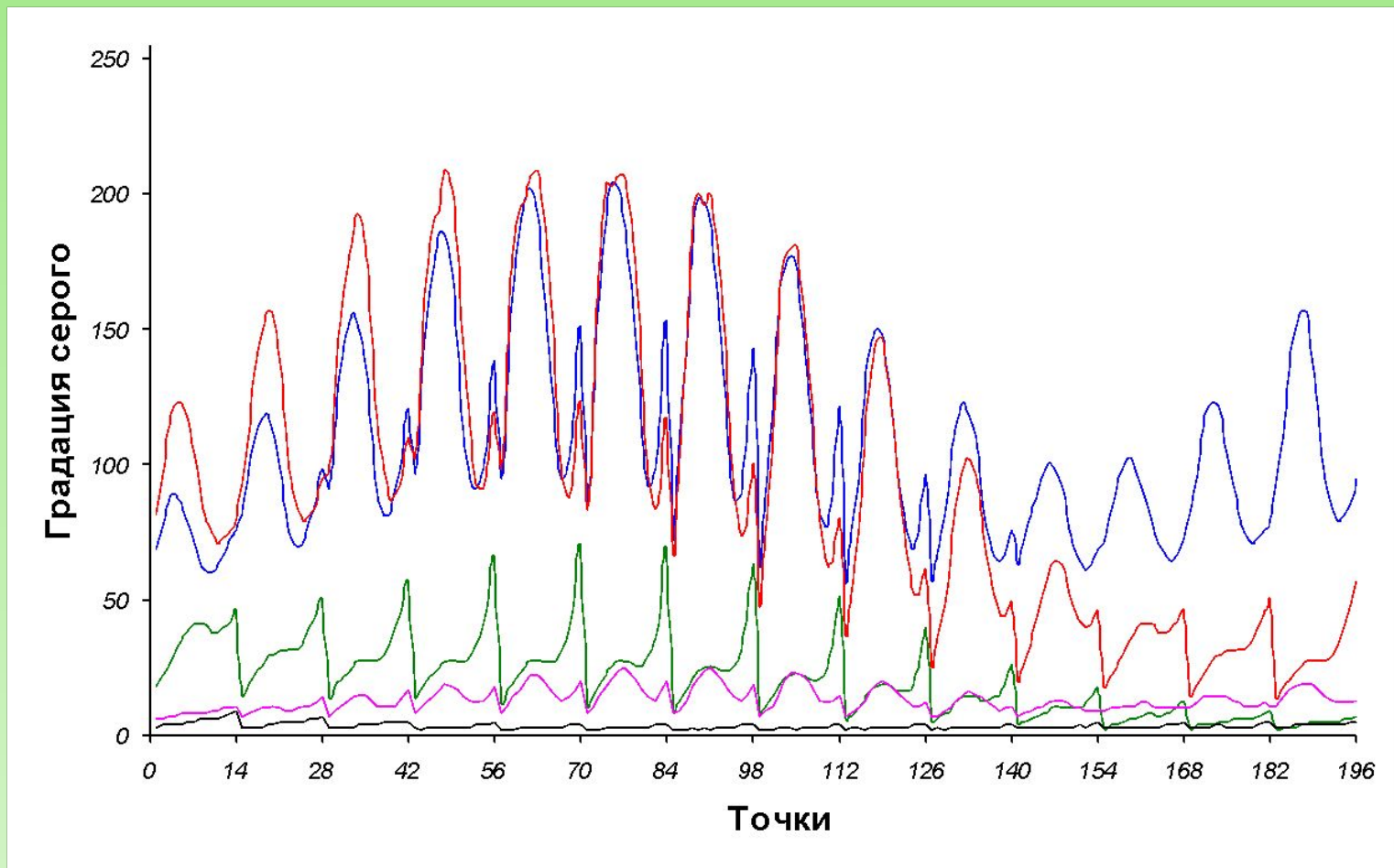


Плохой

Оцифровка изображения

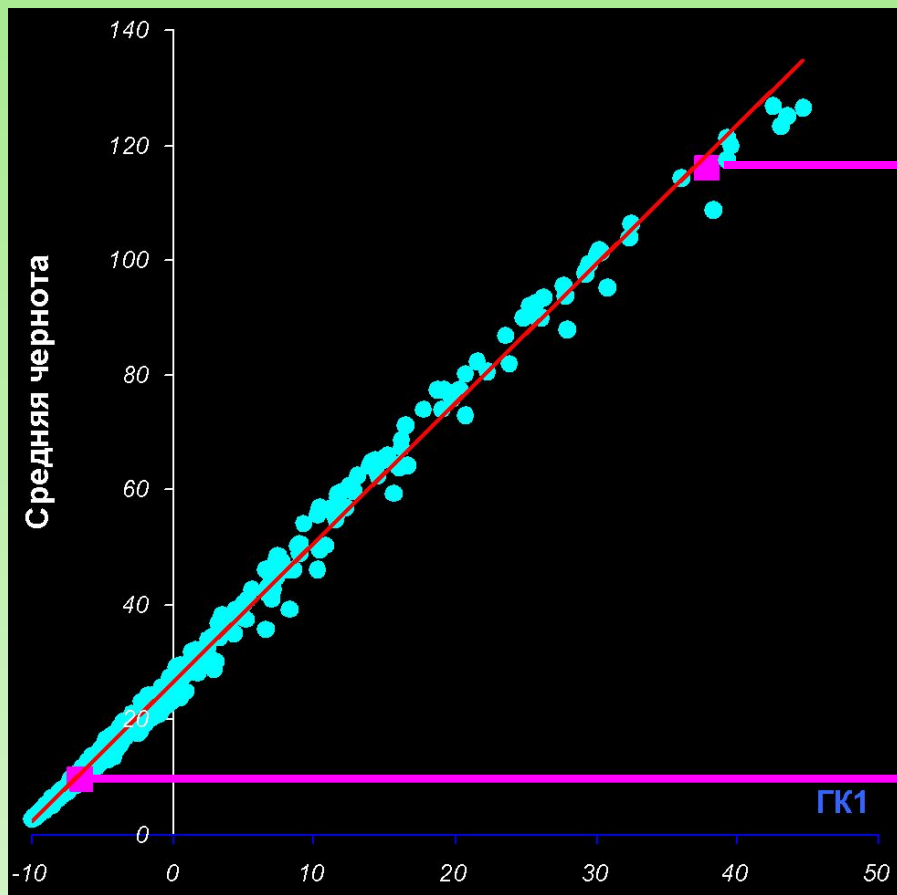


Псевдо - спектры

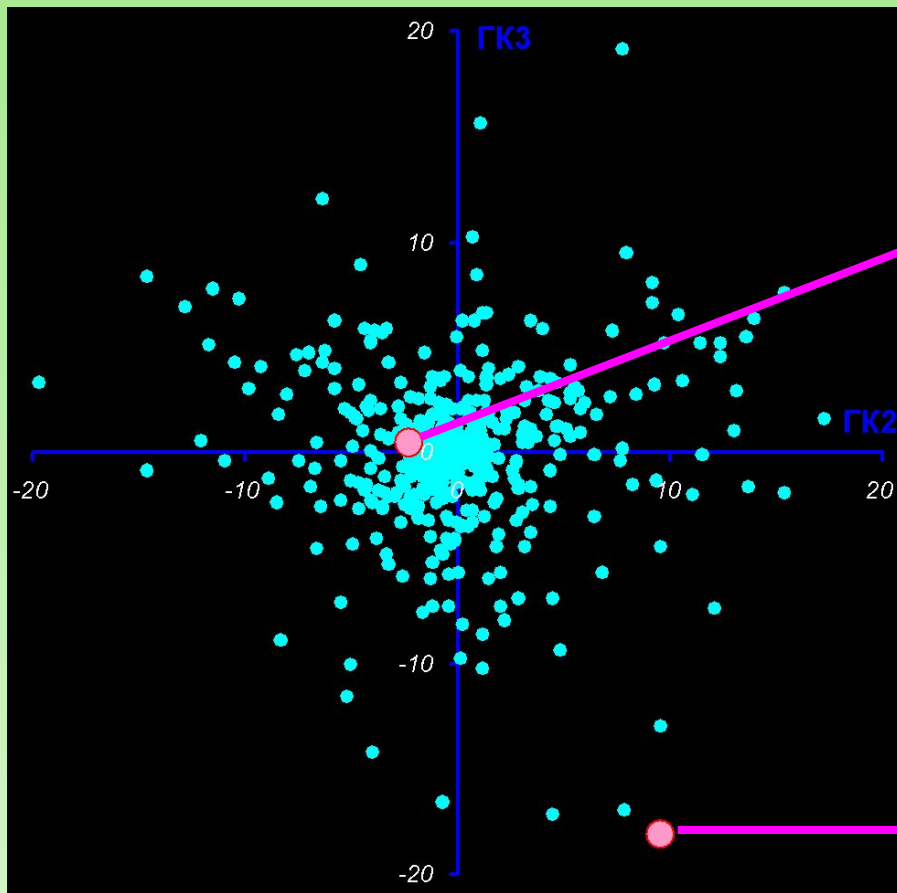


480 образцов (строк) × 196 переменных (столбцов)

Первая ГК



Вторая и третья ГК



Часть 5. Промышленность



01.04.07

Лекция в ОмГУ

52

Что такое PAT?

Process Analytical Technology (PAT) =
Технология (методы) анализа процессов

**FDA = U.S. Department of Health and Human
Services Food and Drug Administration**

Guidance for Industry PAT — A Framework for Innovative
Pharmaceutical Development, Manufacturing, and Quality Assurance
Pharmaceutical CGMPs, September 2004

PAT: Определение бюрократа

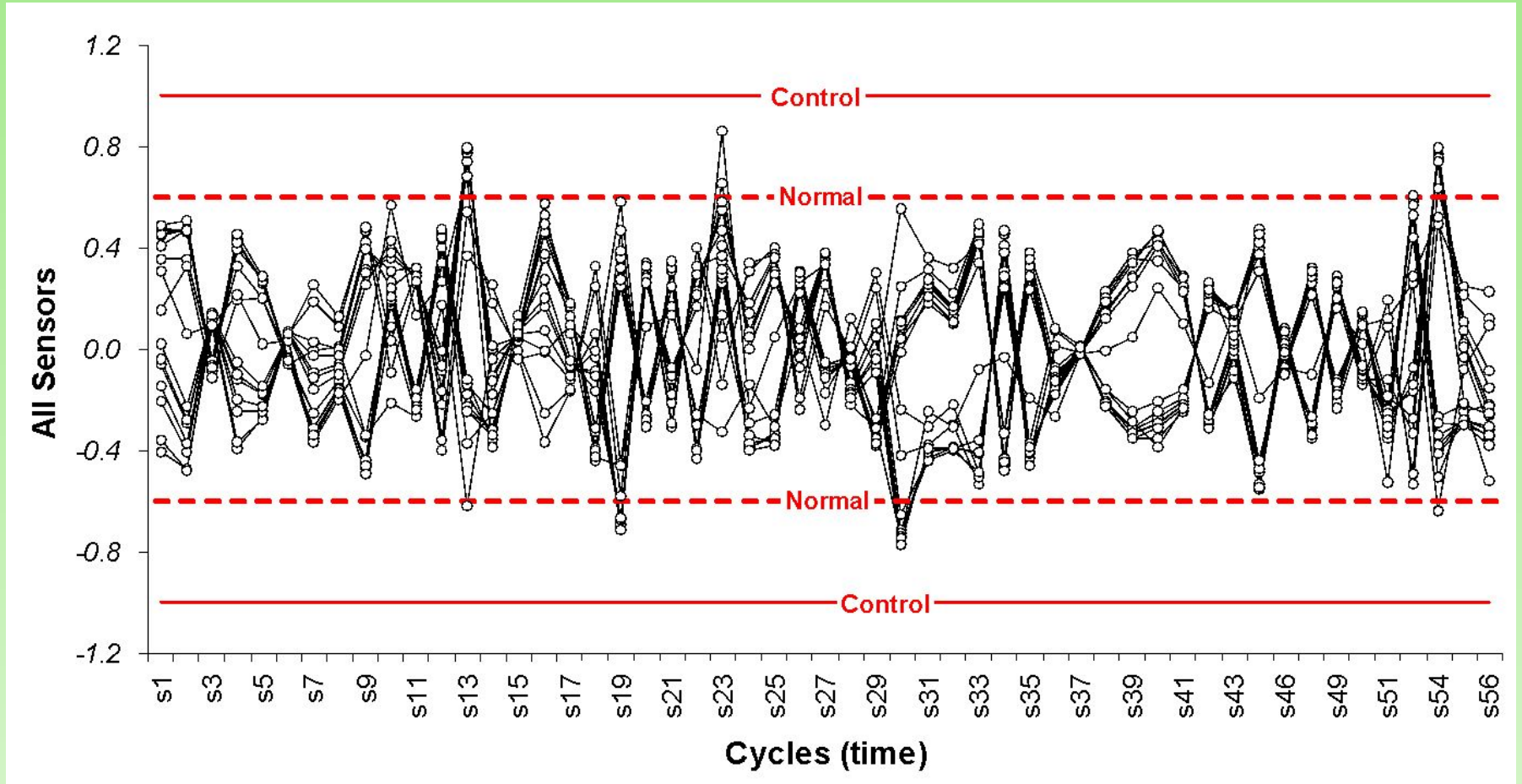
PAT – это система планирования, анализа и контроля критических переменных, характеризующих состояние производственных материалов и процессов в реальном времени (т.е. по ходу производства), с целью подтверждения качества производимого продукта.

Guidance for Industry PAT — A Framework for
Innovative Pharmaceutical Development,
Manufacturing, and Quality Assurance. Pharmaceutical
CGMPs, September 2004

Исторические данные

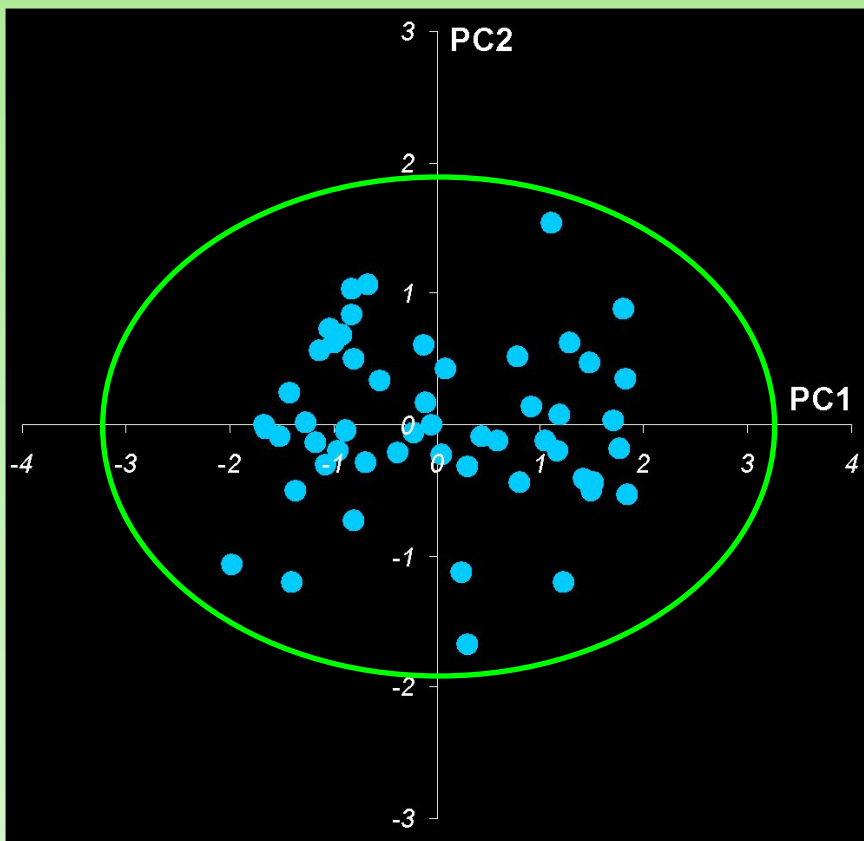
		Ключевые переменные процесса (датчики) X1, X2, ... , X17										
		X1	X2	X3	X4	X5	X6	X7	X8	X9		X17
Реализации процесса s1, s2, ... ,s54	s1	-1.19E-01	7.28E-01	-2.15E-02	5.22E-01	7.06E-04	7.32E-01	3.10E-04	-6.13E-04	-5.92E-05		9.74E-03
	s2	-1.37E-01	7.28E-01	-2.89E-02	6.08E-01	7.09E-04	7.02E-01	6.58E-04	-1.22E-03	-1.49E-04		1.01E-02
	s3	2.51E-02	-9.15E-02	6.73E-03	-1.13E-01	-9.07E-05	-7.58E-02	-2.29E-04	4.10E-04	5.65E-05		-1.43E-03
	s4	-1.14E-01	6.70E-01	-2.18E-02	5.04E-01	6.50E-04	6.65E-01	3.83E-04	-7.34E-04	-7.96E-05		9.07E-03
	s5	-7.93E-02	4.14E-01	-1.69E-02	3.51E-01	4.04E-04	3.98E-01	3.96E-04	-7.35E-04	-9.05E-05		5.78E-03
	s6	1.51E-02	-6.38E-02	3.74E-03	-6.75E-02	-6.28E-05	-5.67E-02	-1.15E-04	2.07E-04	2.78E-05		-9.49E-04
	s7	7.44E-02	-5.24E-01	1.11E-02	-3.24E-01	-5.06E-04	-5.45E-01	-1.73E-05	7.92E-05	-1.07E-05		-6.79E-03
	s8	3.65E-02	-2.66E-01	5.12E-03	-1.59E-01	-2.56E-04	-2.78E-01	1.43E-05	-3.95E-07	-1.14E-05		-3.42E-03
	s9	1.36E-01	-7.06E-01	2.89E-02	-6.01E-01	-6.88E-04	-6.77E-01	-6.83E-04	1.26E-03	1.56E-04	● ● ●	-9.86E-03
	s10	-2.74E-02	3.60E-01	1.82E-03	1.12E-01	3.42E-04	4.12E-01	-4.31E-04	7.24E-04	1.22E-04		4.18E-03
	s11	7.47E-02	-3.31E-01	1.80E-02	-3.34E-01	-3.25E-04	-2.99E-01	-5.30E-04	9.62E-04	1.28E-04		-4.84E-03
	s12	-1.17E-01	7.02E-01	-2.16E-02	5.13E-01	6.81E-04	7.03E-01	3.40E-04	-6.63E-04	-6.76E-05		9.44E-03
	s13	1.06E-01	-2.82E-01	3.23E-02	-4.82E-01	-2.85E-04	-1.87E-01	-1.25E-03	2.21E-03	3.14E-04		-4.99E-03
	s14	7.39E-02	-5.28E-01	1.07E-02	-3.21E-01	-5.09E-04	-5.50E-01	2.49E-06	4.48E-05	-1.59E-05		-6.81E-03
	s15	-9.87E-03	1.02E-01	-3.21E-04	4.17E-02	9.75E-05	1.13E-01	-8.29E-05	1.36E-04	2.44E-05		1.23E-03
	s16	-1.06E-01	7.68E-01	-1.52E-02	4.62E-01	7.41E-04	8.03E-01	-2.54E-05	-2.68E-05	2.88E-05		9.90E-03
	s17	-4.76E-02	2.66E-01	-9.52E-03	2.10E-01	2.59E-04	2.61E-01	1.92E-04	-3.61E-04	-4.19E-05		3.65E-03
● ● ●												
s54	6.61E-02	-5.40E-01	7.19E-03	-2.85E-01	-5.19E-04	-5.78E-01	1.81E-04	-2.67E-04	-6.23E-05		-6.78E-03	

Карты Шухарта SPS (1927)

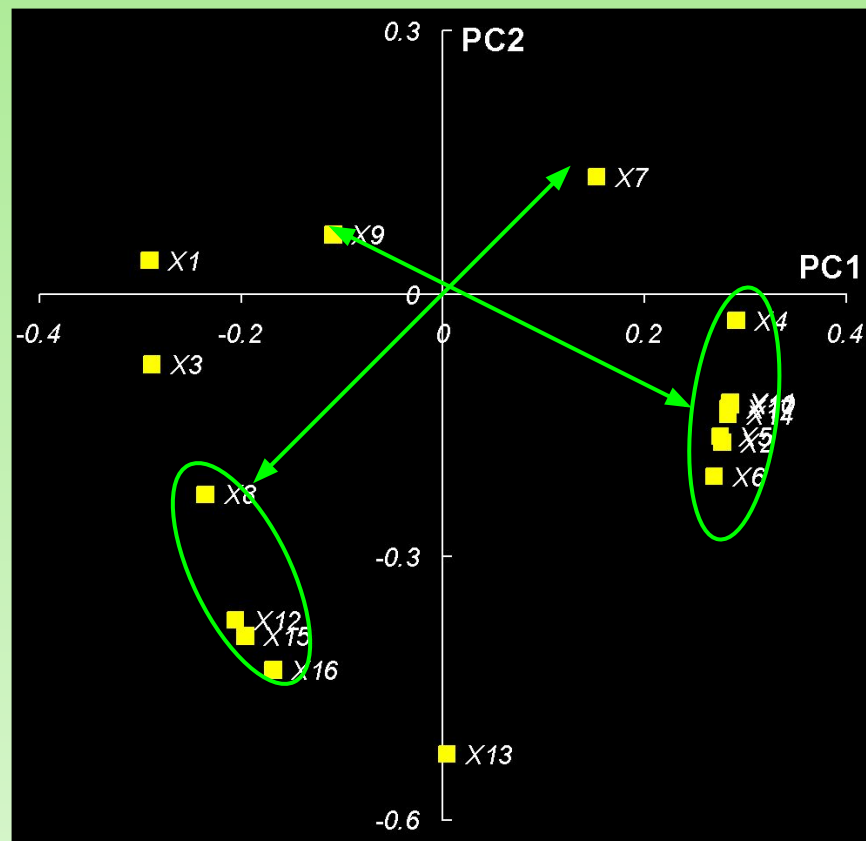


МГК на исторических данных

Образцы = Счета Т

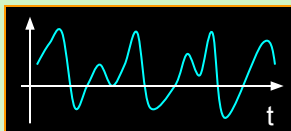
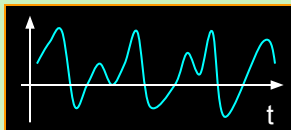
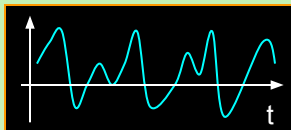
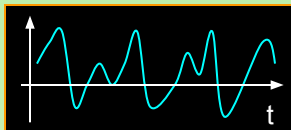
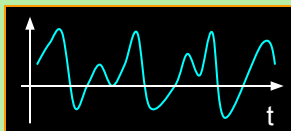
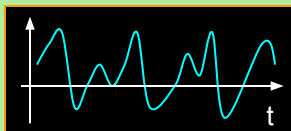
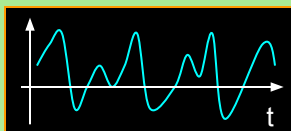


Переменные = Нагрузки Р

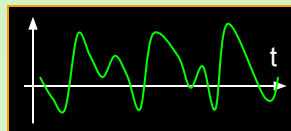
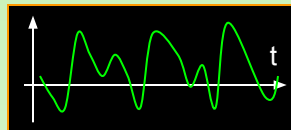
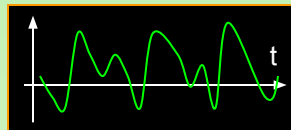
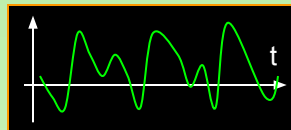
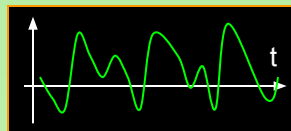
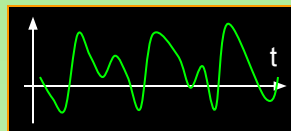
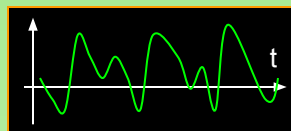


Стат. контроль процессов (SPC)

X1

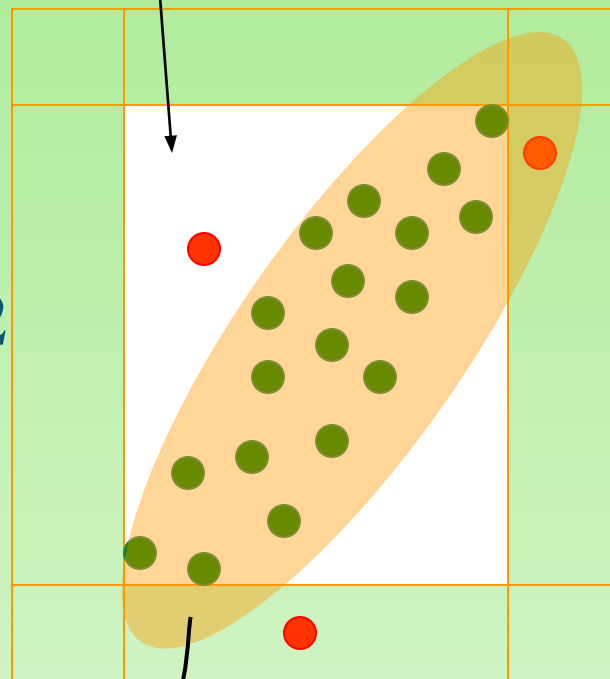


X2



Одномерная область контроля

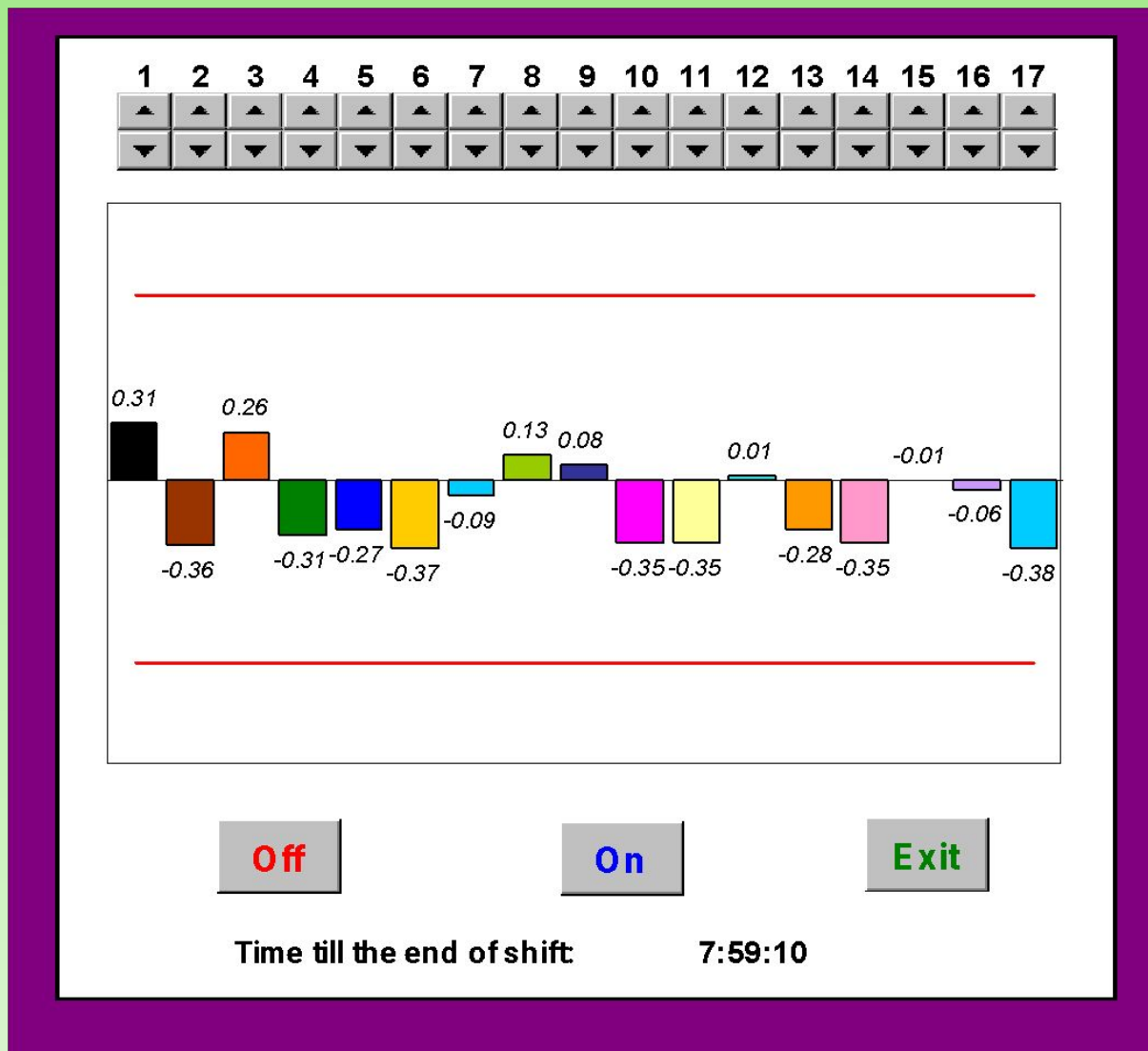
X2



X1

Многомерная область контроля

Контроль процесса (не просто игра)



MSPC & PAT

**Пятница 6 апреля
9:45 ауд. 208
А.Л. Померанцев
Аналитический контроль
процессов**

Часть 6. Четвертая парадигма



Академик Кадыров

01.04.07

Лекция в ОмГУ

Парадигма –
концептуальная схема,
используемая для
постановки проблем и их
решения, господствующая
в научном сообществе в
течение длительного
исторического периода.

Энцикл. словарь

Парадигмы науки

VII с. BC

Пифагор, Евклид

Геометрическая парадигма науки. Теорема Пифагора

XV с. AD

Лука Пачоли, папа Сильвестр II

Арифметическая парадигма науки. Таблица умножения

XVII с.
AD

Исаак Ньютон, Готфрид Лейбниц

Дифференциальная парадигма науки. Производная

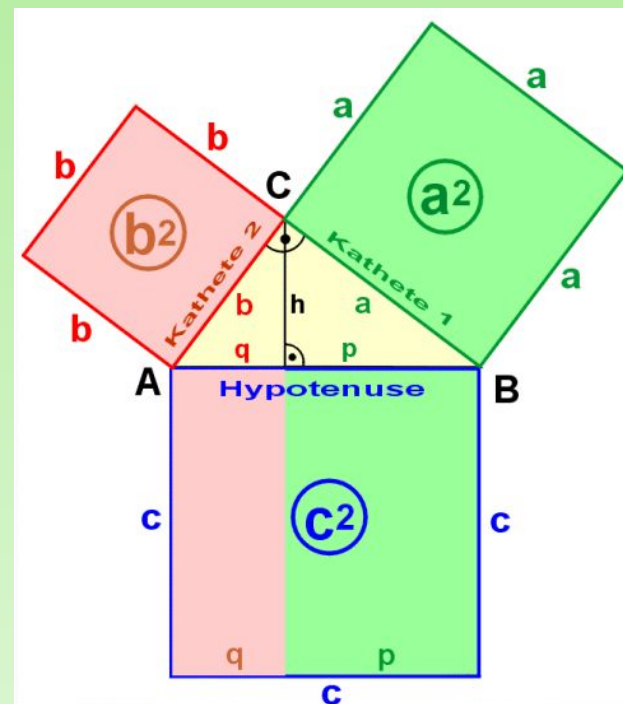
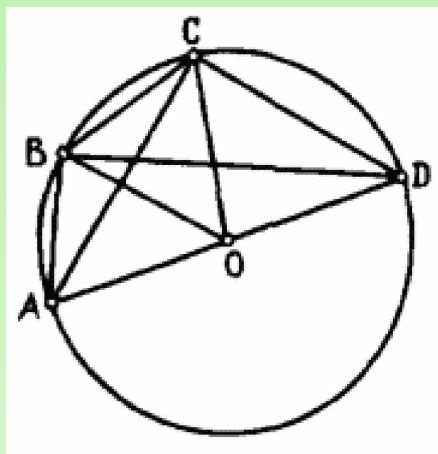
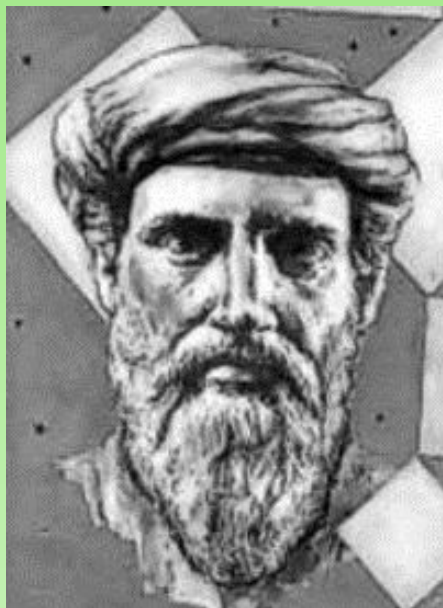
XXI с. AD

?????

Новая парадигма науки. Неформальное моделирование

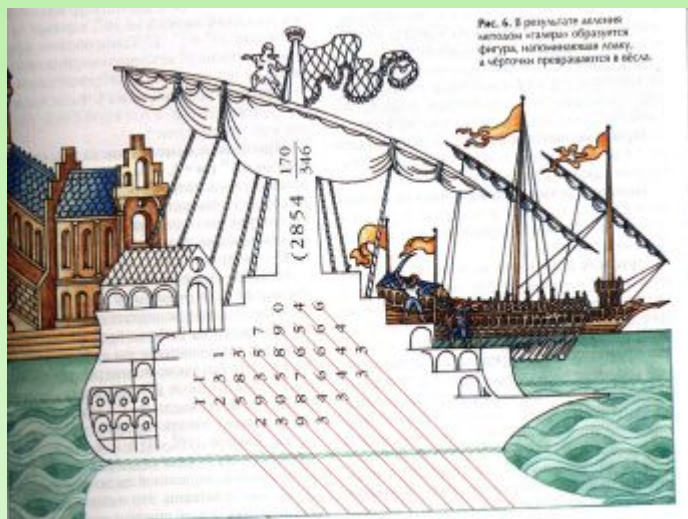
I. Геометрическая

V- III вв до н.э.



II. Алгебраическая

X-XV н.э.

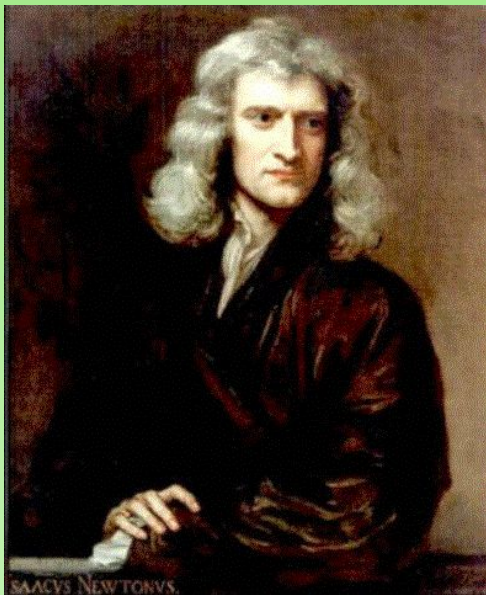


01.04.07

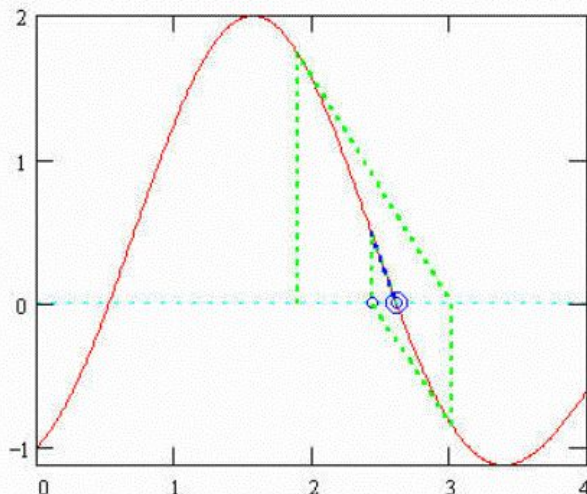
Лекция в ОмГУ

64

III. Дифференциальная XVII в. – н.в.

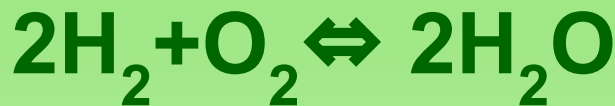


Newton's Method



r	P_r
0	1.900000000000
1	3.022987584372
2	2.439464338938
3	2.614942624257





№	Реакция	Предполагаемая ошибка, %	Диапазон температур, К	A	
				—	—
1	$\text{H}_2 + \text{O}_2 \rightleftharpoons 2\text{OH}$	250	900—2500	$(0,25-2,7) \cdot 10^{14}$	
2	$\text{OH} + \text{H}_2 \rightleftharpoons \text{H}_2\text{O} + \text{H}$	50	500—2500	$(1-3) \cdot 10^{19}$	
3	$\text{H} + \text{O}_2 \rightleftharpoons \text{OH} + \text{O}$	30—50	700—2500	$(2-3) \cdot 10^{11}$	
4	$\text{O} + \text{H}_2 \rightleftharpoons \text{OH} + \text{H}$	30—80	500—2000	$(1,8-2) \cdot 10^7$	
5	$\text{O} + \text{H}_2\text{O} \rightleftharpoons 2\text{OH}$	150	500—2500	$(6-8) \cdot 10^{14}$	
6	$2\text{H} + \text{M} \rightleftharpoons \text{H}_2 + \text{M}$	50—70	1000—2500	$(0,3-1,0) \cdot 10^8$	
7	$3\text{O} + \text{M} \rightleftharpoons \text{O}_3 + \text{M}$	70	900—2500	$(5-8) \cdot 10^8$	
8	$\text{H} + \text{OH} + \text{M} \rightleftharpoons \text{H}_2\text{O} + \text{M}$	100—120	1000—3000	$(1-2) \cdot 10^{17}$	
9	$2\text{OH} + \text{M} \rightleftharpoons \text{H}_2\text{O}_2 + \text{M}$	200—300	700—2200	$(8,4-9,1) \cdot 10^8$	
10	$\text{OH} + \text{O} + \text{M} \rightleftharpoons \text{HO}_2 + \text{M}$	400	1000—2500	$(1,0-8,5) \cdot 10^{12}$	
11	$\text{H} + \text{O}_2 + \text{M} \rightleftharpoons \text{HO}_2 + \text{M}$	30—70	500—2500	$(1,5-3,9) \cdot 10^8$	
12	$\text{HO}_2 + \text{H}_2 \rightleftharpoons \text{H}_2\text{O}_2 + \text{H}$	300—400	800—2500	$(5,0-9,5) \cdot 10^8$	
13	$\text{HO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}_2\text{O}_2 + \text{OH}$	500—800	900—2300	$(0,8-1,5) \cdot 10^8$	
14	$\text{HO}_2 + \text{H}_2\text{O} \rightleftharpoons \text{H}_2\text{O}_2 + \text{OH}$	50—70	1000—2200	$(1-4) \cdot 10^{12}$	
15	$2\text{HO}_2 \rightleftharpoons \text{H}_2\text{O}_2 + \text{O}_2$	300—500	500—2200	$(2-4) \cdot 10^8$	
16	$\text{H} + \text{HO}_2 \rightleftharpoons 2\text{OH}$	200—450	500—2000	$(1-3) \cdot 10^{11}$	
17	$\text{H} + \text{HO}_2 \rightleftharpoons \text{H}_2\text{O} + \text{O}$	800	900—2500	$(0,8-2,0) \cdot 10^{10}$	
18	$\text{H} + \text{HO}_2 \rightleftharpoons \text{H}_2 + \text{O}_2$	250—350	500—2000	$(3-5) \cdot 10^{10}$	
19	$\text{O} + \text{HO}_2 \rightleftharpoons \text{OH} + \text{O}_2$	500—800	900—2500	$(1-6) \cdot 10^{10}$	
20	$\text{H} + \text{H}_2\text{O}_2 \rightleftharpoons \text{H}_2\text{O} + \text{OH}$	300—500	500—2000	$(0,7-1,3) \cdot 10^{12}$	
21	$\text{O} + \text{H}_2\text{O}_2 \rightleftharpoons \text{OH} + \text{HO}_2$	800	800—2000	$(1-4) \cdot 10^{10}$	
22	$\text{H}_2 + \text{O}_2 \rightleftharpoons \text{H}_2\text{O} + \text{O}$	800	1000—2500	$(3-8) \cdot 10^{10}$	
23	$\text{H}_2 + \text{O}_2 + \text{M} \rightleftharpoons \text{H}_2\text{O}_2 + \text{M}$	800	1000—2500	$(1-5) \cdot 10^7$	
24	$\text{OH} + \text{M} \rightleftharpoons \text{O} + \text{H} + \text{M}$	100—150	750—2000	$(1-4) \cdot 10^{12}$	
25	$\text{HO}_2 + \text{OH} \rightleftharpoons \text{H}_2\text{O} + \text{O}_2$	800	900—2500	$(1-3) \cdot 10^{10}$	
26	$\text{H}_2 + \text{O} + \text{M} \rightleftharpoons \text{H}_2\text{O} + \text{M}$	800	900—2500	$(0,5-5,0) \cdot 10^8$	
27	$\text{O} + \text{H}_2\text{O} + \text{M} \rightleftharpoons \text{H}_2\text{O}_2 + \text{M}$	1000	1000—2000	$(6,0-9,0) \cdot 10^7$	
28	$\text{O} + \text{H}_2\text{O}_2 \rightleftharpoons \text{H}_2\text{O} + \text{O}_2$	1000	1000—2000	$(0,6-2,0) \cdot 10^8$	
29	$\text{H}_2 + \text{H}_2\text{O}_2 \rightleftharpoons 2\text{H}_2\text{O}$	1000	1000—2000	$(0,6-2,0) \cdot 10^{10}$	
30	$\text{H} + \text{HO}_2 + \text{M} \rightleftharpoons \text{H}_2\text{O}_2 + \text{M}$	1000	1000—2000	$(0,5-3,0) \cdot 10^8$	

Примечание. Коэффициент скорости реакции второго порядка по данным [26—29, 78—82]; Б — по [26—29, 78—82] с учетом [15]; В — [11]; Д — по [26—29, 78—82] с учетом [14]; Е — по [26—29, 78—82] с учетом [11]; Ж — значения k_2 , полученные решением ОИЗ.

За исключением альтернативной стадии 18, реакция 1 является одним из самых быстрых процессов зарождения, что не удивительно, так как это типичная обменная реак-

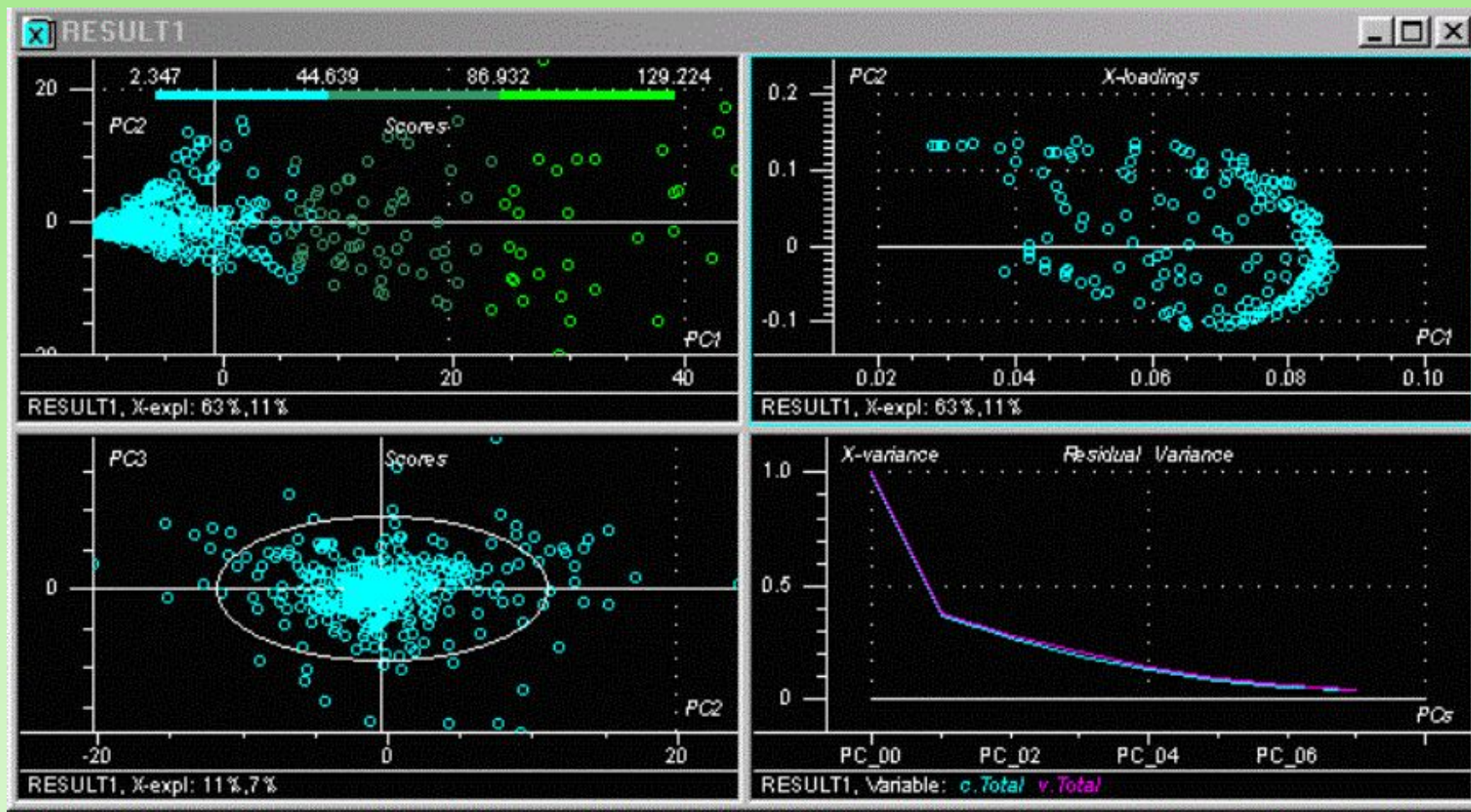
Таблица 4

n	n ⁺		n ⁻		Примечание
	Б	А	н	Е	
—	38100—38900	$(0,8-8,4) \cdot 10^8$	—	20360—20560	А
—	5000—5400	$(0,7-1,0) \cdot 10^{11}$	—	20300—21000	Г
—	16700—17600	$(1,9-3,2) \cdot 10^{10}$	—	0—750	Б, Г
—	8550—8950	$(8,0-8,5) \cdot 10^8$	+1	6800—7200	Г
—	18300—18800	$(5-8) \cdot 10^8$	—	1000—1300	А, Б, Г
—	—	$(1,8-2,2) \cdot 10^{10}$	-1	101500—103500	Б
—	—	$(5-9) \cdot 10^{10}$	-1	117200—122200	Б
—	—	$(2-3) \cdot 10^{10}$	—	103800—107800	Г
—	(4370—5960)	$(1,1-1,3) \cdot 10^{11}$	—	44700—46700	А, Г
—	—	$(0,7-1,3) \cdot 10^{12}$	—	60500—73500	В, Ж
—	(990—1890)	$(2,1-2,4) \cdot 10^{10}$	—	45700—47100	А—Е
—	17800—21800	$(1,4-1,7) \cdot 10^8$	—	3000—5000	А, Г
—	22800—24800	$(1-2,2) \cdot 10^8$	+1/2	53600—57600	А
—	31000—34000	$(3-1,3) \cdot 10^{10}$	—	1590—1950	Г
—	—	$(1-1,5) \cdot 10^8$	+1/2	39700—42700	Е, Ж
—	990—2580	$(0,8-1,5) \cdot 10^{10}$	—	39700—40900	Г
—	600—3580	$(0,1-1,0) \cdot 10^{10}$	—	57000—59000	А
—	0—1200	$(3-8) \cdot 10^{10}$	—	56000—56600	Г
—	—	$(1-6) \cdot 10^{10}$	—	53500—56500	А
—	9500—11900	$(0,8-1,5) \cdot 10^{11}$	—	77700—81700	А
—	900—1300	$(2,6-5,0) \cdot 10^{10}$	—	11000—13000	В, Е
—	52500—56500	$(3-8) \cdot 10^{10}$	—	52500—56500	В
—	17900—21900	$(0,5-5,0) \cdot 10^{10}$	—	48500—51500	В
—	101300—105300	$(0,6-4,5) \cdot 10^{10}$	—	—	А, Б
—	0—600	$(0,8-1,4) \cdot 10^{10}$	+1/2	69500—77500	Е
—	—	$(0,5-5,0) \cdot 10^{10}$	—	107000—111000	А, Ж
—	11000—13000	$(0,9-1,2) \cdot 10^{12}$	—	44600—46600	В, Ж
—	25000—29000	$(0,5-2,0) \cdot 10^8$	—	102000—111000	В
—	20000—22000	$(1-6) \cdot 10^8$	—	108000—112000	В
—	1000—1500	$(0,5-2,5) \cdot 10^{12}$	—	88000—94000	В

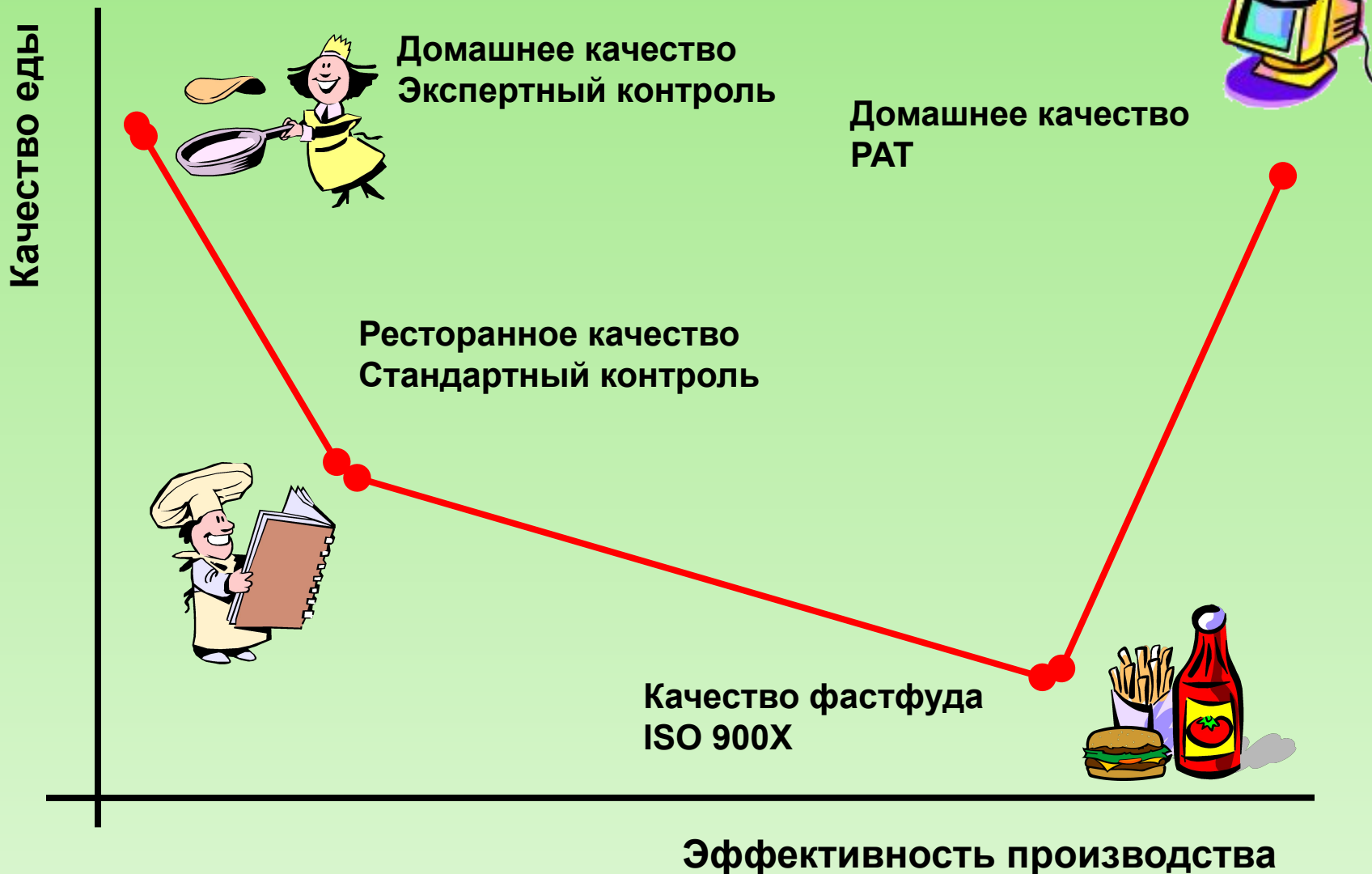
выражается в л/моль·с, третьего порядка — в л²/моль²·с. А — значения k_2 квантово-механические оценки; Г — значения k_2 по [26—29, 78—82] с учетом [11]; Ж — значения k_2 , полученные решением ОИЗ.

ция умеренной (для реакций этого класса) эндотермичности по сравнению, например, с сильно эндотермическими диссоциативными стадиями зарождения 6—8. При этом аб-

IV. Новая парадигма



Наивный пример РАТ



Sixth Winter Symposium on Chemometrics

Казань

Школа: 16-17 февраля 2008
Конференция: 18-22 февраля 2008



01.04.07

Лекция в ОмГУ

69

Участники WSC-5 в Самаре



01.04.07

Лекция в ОмГУ

70

Благодарность

*Автопортрет в
костюме Деда Мороза*



В презентации использованы
репродукции картин Питерского
художника

Николая Копейкина

www.nomzhir.spb.ru/koreykin

За это ему – большое спасибо!

Спасибо за внимание!

