

УДК 681.513

ПОВЫШЕНИЕ ТОЧНОСТИ ПРОГНОЗА ЗА СЧЕТ ВЫДЕЛЕНИЯ РЕГУЛЯРНОГО ТРЕНДА И СЛУЧАЙНОГО ОСТАТКА

Е.А. Савченко¹, А.А. Прачев²

¹Международный научно-учебный центр информационных технологий и систем НАН и МОН Украины,

²НТУ Украины «КПИ»
savchenko@irtc.org.ua

Запропонований метод засновано на виділенні регулярного тренда і випадкового залишку. Його можна застосовувати для підвищення точності прогнозу регулярних процесів. Цей метод полягає в розрахунку часового тренда за вибіркою експериментальних даних і прогнозі відхилення від тренду за комбінаторним алгоритмом методу групового урахування аргументів. Показано збільшення точності прогнозу на прикладі прогнозу рівня глюкози в крові хворого на діабет.

Ключеві слова: комбінаторний алгоритм МГУА, тренд, регулярний процес, прогноз, рівень глюкози в крові

The proposed method is based on calculation of a regular trend deviation. It can be applied to the forecast model accuracy enhancement in regular processes. The method consists in calculation of regular trend after real data sample and forecasting of the trend deviation by combinatorial Group Method of Data Handling algorithm. An example of blood glucose level forecasting in the task of diabetic treatment is shown.

Keywords: Combinatorial GMDH algorithm, trend, regular process, forecasting, blood glucose level.

Предложенный метод основан на выделении регулярного тренда и случайного остатка. Он может быть применен для повышения точности прогноза регулярных процессов. Этот метод состоит в расчете временного тренда по выборке экспериментальных данных и прогнозе отклонения от тренда по комбинаторному алгоритму методу группового учета аргументов. Показано увеличение точности прогноза на примере прогноза уровня глюкозы в крови больного диабетом.

Ключевые слова: комбинаторный алгоритм МГУА, тренд, регулярный процесс, прогноз, уровень глюкозы в крови.

Введение

Увеличение точности прогноза показано на примере поиска оптимальной модели для прогноза уровня глюкозы в крови (УГК) больного диабетом по реальным клиническим данным с помощью комбинаторного алгоритма МГУА [1].

В работе сравниваются три способа получения оптимальной модели для прогноза УГК [2]:

- 1) с использованием комбинаторного алгоритма МГУА;
- 2) по уравнению временного тренда;
- 3) при помощи прогнозирования отклонения от тренда по комбинаторному алгоритму МГУА.

Данные мониторинга лечения диабета представлены учеными из венгерской лаборатории логики [3]. УГК больного измеряется четыре раза в день: в 6, 12, 17 и 22 часа. Кроме того, измеряется доза инсулина, которая вводится больному в то же самое время. Требуется спрогнозировать значение УГК на три дня вперед, чтобы врач видел процесс развития болезни и мог корректировать лечение.

1. Построение модели по комбинаторному алгоритму МГУА (без выделения тренда)

Сначала построим прогнозирующую модель по комбинаторному алгоритму МГУА. Такие модели могут быть оптимальными для стационарных временных рядов. Исходные данные мониторинга представлены в таблице 1.

Таблица 1

Исходные данные мониторинга лечения диабета

День	x_6	x_{12}	x_{17}	x_{22}	z_6	z_{12}	z_{17}	z_{22}
16.08.1993	15.4000	12.4000	9.7000	7.9000	20.0	10.0	14.0	12.0
17.08.1993	10.5000	10.6000	5.5000	7.0000	20.0	12.0	14.0	12.0
18.08.1993	13.5000	24.6000	4.4000	13.6000	20.0	12.0	14.0	12.0
19.08.1993	14.9000	16.4000	5.8000	13.6500	20.0	10.0	14.0	12.0
20.08.1993	22.5000	11.4000	6.1000	13.7000	20.0	12.0	14.0	12.0
21.08.1993	10.0000	18.5000	18.1000	19.1000	20.0	12.0	12.0	12.0
22.08.1993	17.2000	14.0000	5.8000	6.4000	19.0	12.0	12.0	12.0
23.08.1993	16.3000	8.5000	7.9000	11.3000	19.0	11.0	14.0	12.0
24.08.1993	15.2000	11.5000	8.0000	6.8000	19.0	9.0	12.0	12.0
25.08.1993	14.0000	21.7000	24.4000	9.8000	19.0	11.0	14.0	12.0
26.08.1993	21.1000	10.4000	11.0000	15.3000	17.0	11.0	14.0	12.0
27.08.1993	14.2000	24.7000	20.7000	10.2000	19.0	11.0	14.0	12.0
28.08.1993	14.3000	19.6000	16.0000	9.6000	18.0	11.0	14.0	12.0
29.08.1993	26.9000	28.1000	20.9000	16.7000	18.0	11.0	13.0	12.0
30.08.1993	26.9000	24.5000	16.5000	7.2000	18.0	11.0	13.0	12.0
31.08.1993	16.5000	16.8000	19.6000	10.9000	18.0	10.0	13.0	12.0
01.09.1993	16.3000	21.4000	16.7000	13.0000	18.0	10.0	13.0	12.0
02.09.1993	15.4000	20.5000	20.4000	20.6000	17.0	10.0	13.0	12.0
03.09.1993	13.7000	15.8000	17.4000	14.4000	17.0	10.0	13.0	12.0
04.09.1993	17.0000	14.6000	17.7000	20.0000	15.0	10.0	13.0	12.0
05.09.1993	9.6000	12.1000	16.4000	12.1000	17.0	10.0	13.0	12.0
06.09.1993	8.3000	13.6000	18.4000	14.9000	18.0	10.0	13.0	12.0

В таблице обозначено:
 $x_6, x_{12}, x_{17}, x_{22}$ – содержание глюкозы, измеренное в 6,12,17 и 22 часа соответственно, ммоль/л;
 $z_6, z_{12}, z_{17}, z_{22}$ - дозы инъекций инсулина в тоже время.

Кроме заданных переменных введем две новые переменные (w_1, w_2):

$$w_1 = x_i \cdot z_i; w_2 = \frac{x_i}{z_i}. \quad (1)$$

Требуется найти следующую зависимость:

$$x_{ik+1} = f(x_{ik}, z_{ik}, w_1, w_2), i = 6, 12, 17, 22. \quad (2)$$

В таблице 2 представлен фрагмент выборки данных для прогноза по комбинаторному алгоритму МГУА с 17 по 21 августа 1993 года.

Таблица 2

Исходные данные для моделирования по МГУА

День болезни	Время измер.	$x_{ik}, \text{ ммоль/л}$	$i, \text{ МО}$	w_1	w_2	$x_{ik+1}, \text{ ммоль/л}$
17.08	6.00	10.500	20.000	210.000	0.525	13.500
	12.00	10.600	12.000	127.200	0.883	24.600
	17.00	5.500	14.000	77.000	0.393	4.400
	22.00	7.000	12.000	84.000	0.583	13.600
18.08	6.00	13.500	20.000	270.000	0.675	14.900
	12.00	24.600	12.000	295.200	2.050	16.400
	17.00	4.400	14.000	61.600	0.314	5.800
	22.00	13.600	12.000	163.200	1.133	13.650
19.08	6.00	14.900	20.000	298.000	0.745	22.500
	12.00	16.400	10.000	164.000	1.640	11.400
	17.00	5.800	14.000	81.200	0.414	6.100
	22.00	13.650	12.000	163.800	1.138	13.700
20.08	6.00	22.500	20.000	450.000	1.125	10.000
	12.00	11.400	12.000	136.800	0.950	18.500
	17.00	6.100	14.000	85.400	0.436	18.100
	22.00	13.700	12.000	164.400	1.142	19.100
21.08	6.00	10.000	20.000	200.000	0.500	17.200
	12.00	18.500	12.000	222.000	1.542	14.000
	17.00	18.100	12.000	217.200	1.508	5.800
	22.00	19.100	12.000	229.200	1.592	6.400

По комбинаторному алгоритму МГУА получена следующая модель:

$$\hat{x}_{ik+1} = 13,021 + 2,99w_2. \quad (3)$$

На рис. 1 приведены действительные значения и прогноз УГК по комбинаторному алгоритму МГУА с 19 по 21 сентября.

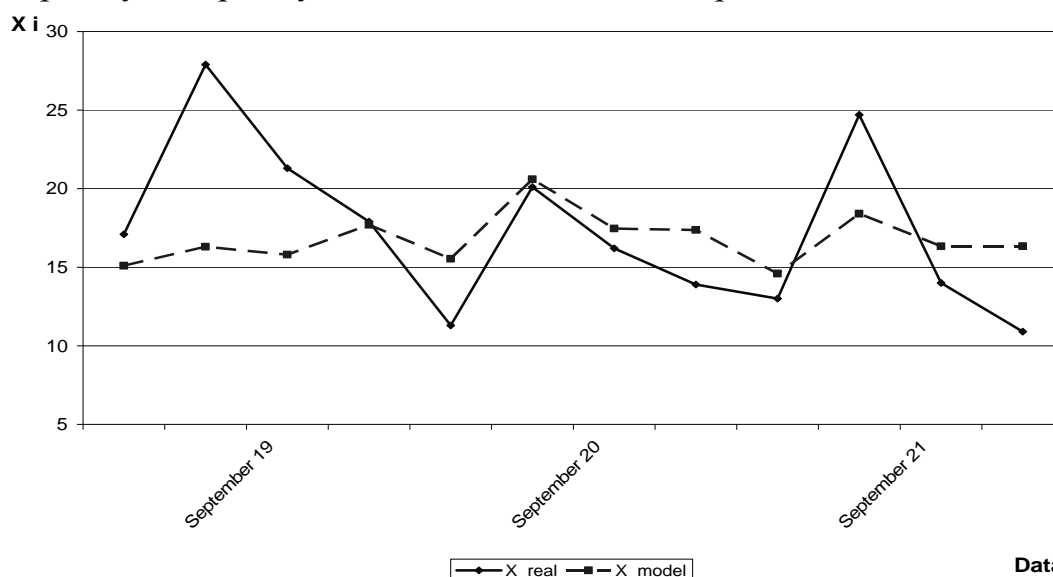


Рис.1. Действительные значения и прогноз УГК по комбинаторному алгоритму МГУА

Значения критериев: $AR = 3,61$; $BS = 1,82$, где AR - критерий регулярности, BS - критерий несмещенности ошибок [3].

Результаты моделирования представлены в таблице 3.

Таблица 3

Результаты моделирования по комбинаторному алгоритму МГУА

№	День болезни	Время измер., ч.	x_{ik} , ммоль/л	\hat{x}_{ik} , ммоль/л	Ошибка модели, %
1	19.09	6	17.1	15.10	11.7
2		12	27.9	16.30	41.6
3		17	21.3	15.80	25.8
4		22	17.9	17.71	1.1
5	20.09	6	11.3	15.54	37.5
6		12	20.1	20.60	2.5
7		17	16.2	17.46	7.8
8		22	13.9	17.37	25.0
9	21.09	6	13	14.59	12.3
10		12	24.7	18.41	25.5
11		17	14	16.33	16.7
12		22	10.9	16.34	49.9
Средняя ошибка прогнозирования (3 дня прогноза)					21.5

Можно заключить, что описанный выше алгоритм не дает достаточно малую ошибку прогноза, поэтому попробуем применить другие способы расчета прогноза. Рассчитаем прогноз по временному тренду.

2. Задание уравнения степенного тренда

В работе для расчета временного тренда использовано следующее уравнение:

$$\hat{x}_{ik+1} = a_0 + a_1T + a_2t + a_3Tt + a_4T^2 + a_5t^2, \quad (4)$$

где T - номер суток, t - время в часах.

В таблице 4 представлен фрагмент выборки данных для поиска значения тренда. Сутки (T) имеют номера от «-34» до «-1» для рабочей выборки и от «1» до «3» экзаменационной выборки, которая использована для проверки точности прогноза.

Таблица 4

Фрагмент выборки данных для прогнозирования по тренду

№ п/п	Дата	T	t	$T \cdot t$	T^2	t^2
1	17.08	-34.000	6.000	-204.000	1156.000	36.000
2		-34.000	12.000	-408.000	1156.000	144.000
3		-34.000	17.000	-578.000	1156.000	289.000
4		-34.000	22.000	-748.000	1156.000	484.000
5	18.08	-33.000	6.000	-198.000	1089.000	36.000
6		-33.000	12.000	-396.000	1089.000	144.000
7		-33.000	17.000	-561.000	1089.000	289.000
8		-33.000	22.000	-726.000	1089.000	484.000
9	19.08	-32.000	6.000	-192.000	1024.000	36.000
10		-32.000	12.000	-384.000	1024.000	144.000
11		-32.000	17.000	-544.000	1024.000	289.000
12		-32.000	22.000	-704.000	1024.000	484.000

133	19.09	-2.000	6.000	-12.000	4.000	36.000
134		-2.000	12.000	-24.000	4.000	144.000
135		-2.000	17.000	-34.000	4.000	289.000
136		-2.000	22.000	-44.000	4.000	484.000
137	20.09	-1.000	6.000	-6.000	1.000	36.000
138		-1.000	12.000	-12.000	1.000	144.000
139		-1.000	17.000	-17.000	1.000	289.000
140		-1.000	22.000	-22.000	1.000	484.000

Уравнение тренда описывает регулярную часть процесса лечения диабета. Оценки коэффициентов находим с использованием метода наименьших квадратов по данным таблицы 4.

Получим следующее уравнение тренда:

$$\hat{x}_{ik+1} = 8,69 - 0,23T + 1,45t + 0,017Tt - 0,0006T^2 - 0,0047t^2. \quad (5)$$

Рассчитаем прогноз УГК по полученному уравнению для трех дней прогноза с 19 по 21 сентября (таблица 5).

Таблица 5

Результаты моделирования по тренду

№ п/п	День	T	t	x_{ik} , ммоль/л	\hat{x}_{ik} , ммоль/л	Ошибка модели, %
1	19.09	1.0	6	17.1	15.568	9.0
2		1.0	12	27.9	19.240	31.0
3		1.0	17	21.3	19.679	7.6
4		1.0	22	17.9	17.736	0.9
5	20.09	2.0	6	11.3	15.434	36.6
6		2.0	12	20.1	19.208	4.4
7		2.0	17	16.2	19.732	21.8
8		2.0	22	13.9	17.874	28.6
9	21.09	3.0	6	13	15.299	17.7
10		3.0	12	24.7	19.175	22.4
11		3.0	17	14	19.784	41.3
12		3.0	22	10.9	18.010	65.2
Средняя ошибка прогнозирования (3 дня прогноза)						23.9

Действительные значения и прогноз УГК по тренду, приведенные в табл.5, показаны на рис.2.

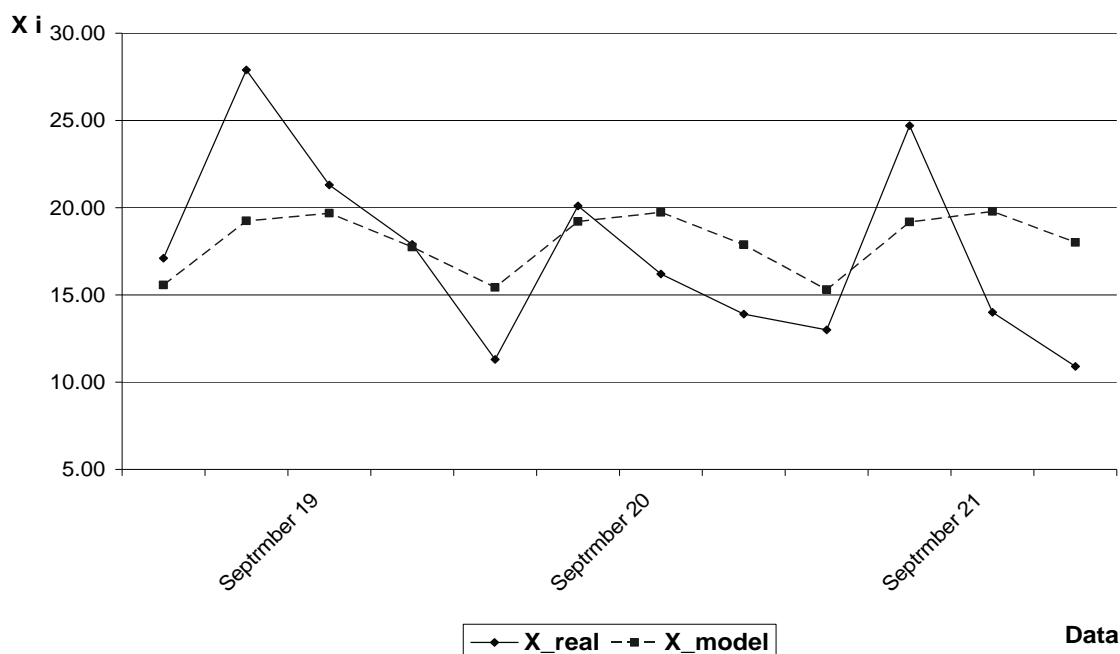


Рис.2. Действительные значения и прогноз УГК по тренду

3. Построение модели отклонения от тренда

Попробуем увеличить точность прогноза, спрогнозировав отклонение от тренда, т.е. построив модель остатка, полученный вычитанием значения тренда от реального значения УГК:

$$\Delta x_i = x_{i\text{real}} - x_{itr}, \quad (6)$$

где $x_{i\text{real}}$ - табличные данные УГК,

x_{itr} - значение УГК, полученного по уравнению тренда.

Найдя тренд по описанному выше алгоритму, определим отклонение Δx_i . По комбинаторному алгоритму МГУА рассчитаем прогноз отклонения от тренда, а затем рассчитаем прогноз УГК, просуммировав значения тренда и прогноз остатка. Результаты приведены в таблице 6.

Таблица 6

Прогнозирование по отклонению от тренда

№ п/п	День	t, ч.	x_{ik} , ммоль/л	x_{itr} , ммоль/л	Δx_i ммоль/л	$\Delta \hat{x}_i$, ммоль/л	\hat{x}_i , ммоль/л	Ошибка модели, %
1	19.09	6	17.10	15.568	-4.13	-1.08	14.513	16.1
2		12	27.90	19.240	0.89	1.31	20.578	26.5
3		17	21.30	19.679	-3.53	-0.92	18.790	11.7
4		22	17.90	17.736	-3.97	-0.42	17.346	2.5
5	20.09	6	11.30	15.434	-2.30	-1.66	13.799	20.7
6		12	20.10	19.208	5.53	-0.18	19.064	5.5
7		17	16.20	19.732	-5.78	-1.66	18.102	11.9
8		22	13.90	17.874	-7.11	-6.08	11.820	14.2
9	21.09	6	13.00	15.299	0.64	1.34	16.666	27.0
10		12	24.70	19.175	-9.64	-0.67	18.538	25.2
11		17	14.00	19.784	-4.93	-1.88	17.939	28.3
12		22	10.90	18.010	1.15	-5.19	12.852	18.9
Средняя ошибка прогнозирования (3 дня прогноза)								17,4

По комбинаторному алгоритму МГУА получена следующая модель остатка:

$$\Delta \hat{x}_i = 0,114 + 0,430w_2 - 0,452x_{itr}. \quad (7)$$

Результаты моделирования показаны на рис.3.

Как видно из табл.1,3 и 6, прогноз УГК при помощи расчета отклонения от тренда позволяет не только уменьшить среднюю ошибку прогноза с 21 до 17%. Кроме того в примере тенденция изменения УГК показана правильно. Рассчитанный таким образом прогноз, может оказать помощь врачу при принятии решения в оценке состояния больного и подбору его дальнейшего лечения.

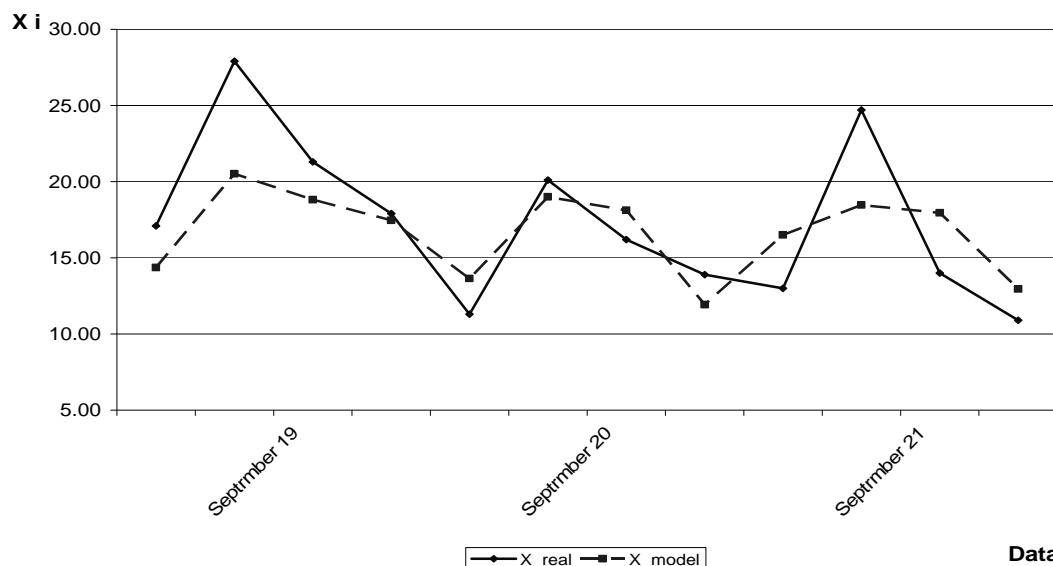


Рис.3. Действительные значения и прогноз УГК с расчетом отклонения от тренда по комбинаторному алгоритму МГУА

Заключение

Рассмотренные примеры показывают, что выделение тренда и прогноз отклонения от него приводит к уменьшению ошибки в задаче лечения больного диабетом. Для того, чтобы получить долгосрочный прогноз можно получить краткосрочный прогноз многократно.

Литература

1. Ивахненко А.Г., Савченко Е.А. Индуктивное моделирование в задаче мониторинга лечения диабета // Праці міжнародного семінару з індуктивного моделювання (МСІМ-2005). – К. : МННЦ ІТС, 2005. – С. 136–139.
2. Savchenko I., Ivakhnenko G., Somina L. Forecasting Model Accuracy Enhancement Using Trend Extraction // Proceedings of 3rd International Conference on Inductive Modelling ICIM'2010, May 16-22, 2010, Yevpatoria. – ХНТУ. – P.152-157.
3. Ивахненко А.Г., Савченко Е.А. Исследование эффективности метода доопределения выбора модели в задачах моделированием с применением МГУА // Проблемы управления и информатики, 2008, №2, С. 65 – 76.
4. Ивахненко А.Г., Степашко В.С. Помехоустойчивость моделирования. - Киев: Наук. думка. – 1985. - 215 с.
5. Ивахненко А.Г., Савченко Е.А., Ивахненко Г.А., Гергей Т. Применение алгоритмов МГУА для восстановления пропущенных данных и прогноза уровня глюкозы в крови при домашнем мониторинге диабета // Проблемы управления и информатики.– 2002. – №3. – С. 123 – 133.
6. Ивахненко А.Г., Савченко Е.А. Исследование эффективности метода доопределения выбора модели в задачах моделированием с применением МГУА // Проблемы управления и информатики. – 2008.– №2. – С. 65–76.