

## АВТОМАТИЗАЦИЯ ПРОЦЕССА МОНИТОРИНГА ЖИВОТНОВОДЧЕСКОГО ПРЕДПРИЯТИЯ НА ОСНОВЕ ИССЛЕДОВАНИЯ ВРЕМЕННЫХ РЯДОВ ПАРАМЕТРОВ КРУПНОГО РОГАТОГО СКОТА

Антонов Л.В.<sup>1</sup>, Варламов А.Д.<sup>1</sup>

<sup>1</sup>Муромский институт (филиал) ФГБОУ ВПО «Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых», Россия, 602264, Владимирская область, г. Муром, ул. Орловская, д.23, is.dep.mivlgu@gmail.com

В работе представлены результаты анализа современных направлений применения информационных технологий для автоматизации животноводческих предприятий. В аналитической части рассмотрены актуальные на сегодняшний день разработки и исследования зарубежных и отечественных ученых. Особое внимание уделяется анализу работ, направленных на применение методов математической статистики при разработке алгоритмов прогнозирования непредвиденных производственных ситуаций. Анализ показал, что методы математической статистики повсеместно используются для обработки данных в автоматизированных информационных системах управления животноводческими хозяйствами и индивидуальное прогнозирование различных характеристик является важной задачей в условиях данного вида производства. Для решения задачи прогнозирования в работе выдвинута гипотеза о подчинении отклонений от среднего значения надоев животных предприятия нормальному закону распределения. Гипотеза доказана с помощью критерия согласия Пирсона при уровне значимости  $\alpha = 0.05$ . Доказательство гипотезы позволило разработать метод отслеживания всплесков внеплановых значений для системы мониторинга производственного процесса на животноводческом комплексе.

Ключевые слова: Критерий согласия Пирсона, нормальное распределение, надои, прогнозирование

## MONITORING PROCESS AUTOMATION OF DAIRY ENTERPRISE BASED ON PARAMETERS RESEARCH OF CATTLE TIME SERIES

Antonov L.V.<sup>1</sup>, Varlamov A.D.<sup>1</sup>

<sup>1</sup>Murom institute (branch) of Vladimit State University name of Aleksandr Grigoryevich and Nikolay Grigoryevich Stoletoy, Russian Federation, Valdimir region, Murom, Orlovskaya street, 23 is.dep.mivlgu@gmail.com

The results of the current trends analysis of information technology to automate the livestock enterprises are represented in the work. Current research and development of foreign and Russian scholars are considered in the analytical part. Works are directed at application of methods of mathematical statistics in the development of algorithms for predicting unforeseen production situations analyzed more thoroughly. The analysis showed that the statistical methods are commonly used for data processing in automated information systems control animal farms and individual forecasting of various characteristics is an important task for the production. To solve the problem of forecasting the been hypothesized of submission of data deviations daily from the average milk yield of animals enterprise to normal distribution. The hypothesis was proved by means of Pearson's agreement criterion with the significance level  $\alpha = 0.05$ . Proof of the conjecture allowed the development of monitoring method of values unplanned bursts for the monitoring of the production process for the livestock sector.

Keywords: Pearson's agreement criterion, normal distribution, yields, prediction

### Введение

Последние два десятилетия ситуацию в молочной промышленности существенно ухудшилась [4]. Тем не менее, в нескольких регионах России отмечается значительный прогресс и стабилизация ситуации в отрасли. Во Владимирской области наблюдается существенный рост количества сельскохозяйственных и животноводческих хозяйств. Впервые за последние годы поголовье крупного рогатого скота в хозяйствах всех категорий увеличилось на 0,8%, коров – на 1,5%. Производство мяса увеличилось к уровню 2005 года

на 14,2%, молока - на 2,9% и составило в 2010 году 63,7 тыс. тонн и 311,9 тыс. тонн соответственно [1].

### **Анализ современных направлений исследований автоматизации животноводческих предприятий**

Одной из наиболее важных задач в условиях ведения животноводства является эффективная селекция животных предприятия и сохранения качества генофонда на высшем уровне. Как показывает статистика сельскохозяйственного департамента Владимирской области, в результате неэффективной селекционной работы увеличивается падеж скота в раннем возрасте, уменьшается уровень среднего надоя за лактационный период, ухудшается иммунная защита животных, следствием чего становится высокая степень риска заболевания вирусными инфекциями, наиболее опасными из которых является лейкоз крупного рогатого скота.

Исследования в данной области начали проводиться с середины 50-ых годов и продолжаются до настоящего времени. Отдельно стоит отметить современную работу [3] американских ученых из университета штата Висконсин. Используя математическую статистику для анализа данных о животных, искусственно разбитых на 28 групп по среднему ежегодному показателю надоя, авторы разрабатывают новую методику отбора животных, основанную на процентном соотношении наследуемых признаков и их генетической корреляции. В работе [9] авторы пользуются средствами аппроксимации данных о молочных надоях для определения концентрации протеина и белка для каждого животного. Используя средства математической статистики, корреляционного анализа, математического моделирования и методы аппроксимации в статье приводятся математические выкладки и графические материалы на их основе, позволяющие оценить уровень надоев, содержание протеина и белка в молоке в зависимости от времени года. В работе [10] авторы используют Байесовский анализ для оценки наследуемости и генетической корреляции между молочными надоями, интервалами между отелами и длиной лактации. Математический анализ и средства математической статистики по отношению к различным параметрам жизнедеятельности животных молочного предприятия применяются также в работах [5-8]. Стоит отметить также работы отечественных авторов. В работе [2] авторы проводят экономический анализ состояния финансового стимулирования для увеличения молочной продуктивности одного животного и приходят к выводу, что при определённых условиях всевозрастающие затраты на молокоотдачу становятся больше, чем сам эффект от продажи продукта. В работе авторы путем статистического анализа находят оптимальное значение надоя для одного животного, дальнейшее увеличение которого не будет приносить предприятию дополнительного дохода.

## Исследование закона распределения производственных данных

При разработке интеллектуального продукта автоматизации животноводческого производства требуется решать проблемы индивидуального прогнозирования состояния и качества продукции для каждой конкретной особи на предприятии. Необходимо отслеживать отклонения и резкие колебания в величине показателей надоев молока, ежедневной активности, изменения температуры, химического состава молока и др. показателей. С этой целью необходимо провести анализ реальных данных, состоящий из следующих этапов:

1. Визуальный анализ данных. Общая характеристика временного ряда исследуемых показателей.
2. Разделение временного ряда показателя на периоды, в которых присутствуют и отсутствуют отклонения значений, требующих внимание к особи.
3. Определение законов, которым подчинены данные этих периодов.
4. Построение математической модели и разработка информационно-аналитической системы на ее основе.

На рис. 1 представлен временной ряд удоя животного за один лактационный период. В рамках решения задачи индивидуального прогнозирования надоев животных на подобных диаграммах выделяются стабильные участки (A-B, C-D). На графике видны резкие нехарактерные перепады значений (B-C, D-E), что говорит о внеплановом снижении объема производства молока и проблемах в состоянии животного. Подобные ситуации необходимо отслеживать на ранних стадиях их возникновения.

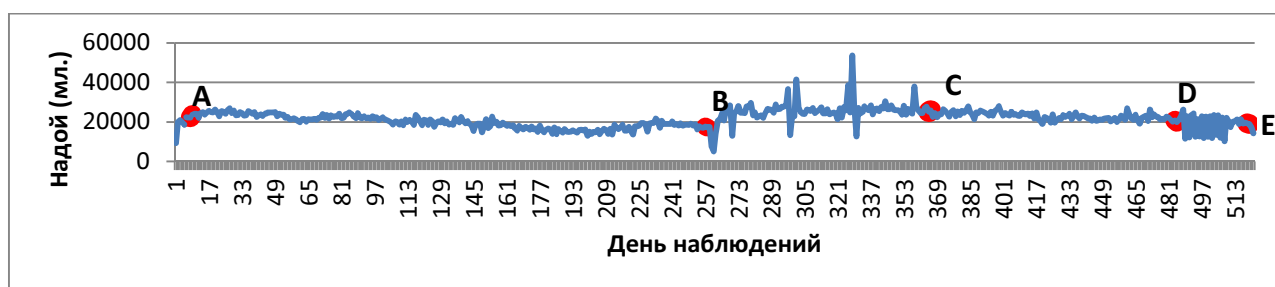


Рис. 1. – график изменения значения ежедневного надоя

Поскольку численность поголовья крупного рогатого скота (КРС) на крупных предприятиях исчисляется тысячами голов отслеживать подобные изменения для каждого животного человеком весьма проблематично. В этом случае возможен человеческий фактор возникновения ошибок, что, в конечном итоге, приведет либо к гибели животных, либо к ухудшению качества выпускаемой продукции. В связи с этим возникает актуальная задача автоматизации выявления нестабильностей удоев в период лактации.

На стабильном участке график ряда значений удоев можно разложить на две составляющих: тренд и отклонение от тренда. Тренд характеризует основную тенденцию

изменения надоев во времени. Взаимодействие между собой регулярных составляющих может быть представлено как аддитивная комбинация:

$$Y(t) = T(t) + E(t), \quad (1)$$

где  $Y(t)$  – значение надоя в момент  $t$ ;

$T(t)$  – тренд изменения надоев;

$E(t)$  – отклонение от тренда.

Для разработки математической модели поиска нестабильных участков, которые могут свидетельствовать о критических изменениях здоровья животного, необходимо знать закон распределения значений  $E(t)$  на стабильных и нестабильных участках. В данной статье ставится задача определения вида этого закона.

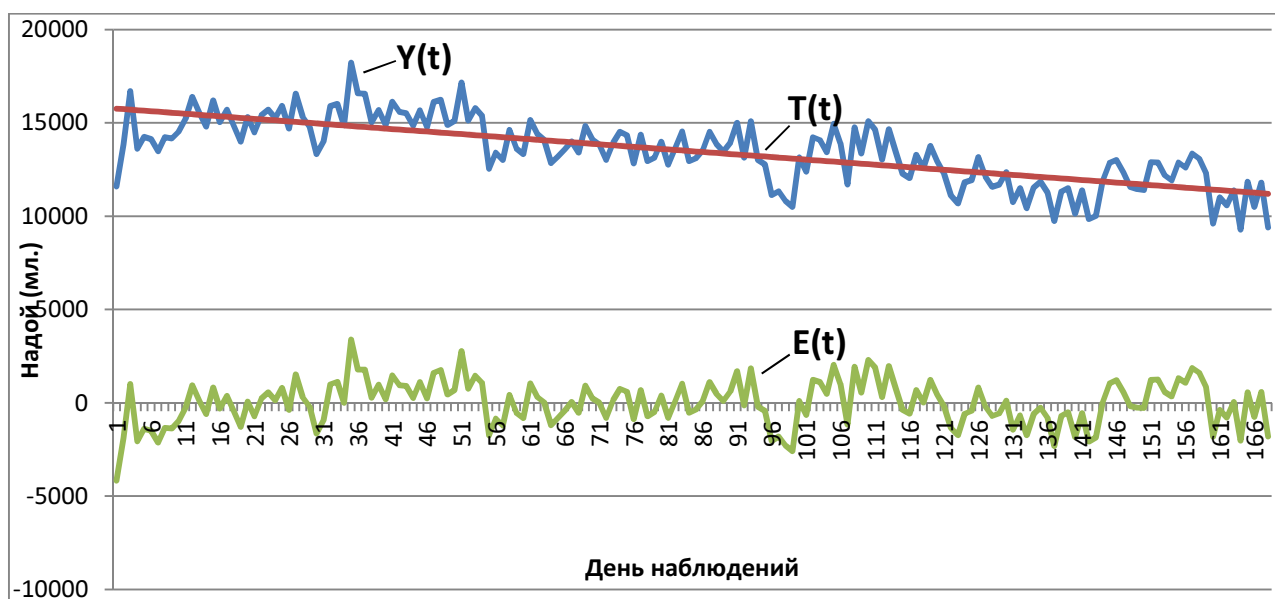


Рис. 2. –  $Y(t)$  - График изменения ежедневных надоев на стабильном участке;  $T(t)$  – трендовая линия аппроксимации  $Y(t)$ ;  $E(t)$  – График отклонения ежедневных надоев от тренда

Используя метод наименьших квадратов, построили трендовую линию изменения надоев  $T(t)$ . Функция  $E(t)$  находится из формулы (1). Построим гистограмму значений эмпирического распределения функции  $E(t)$ , разделенных на 17 периодов (рис 3).

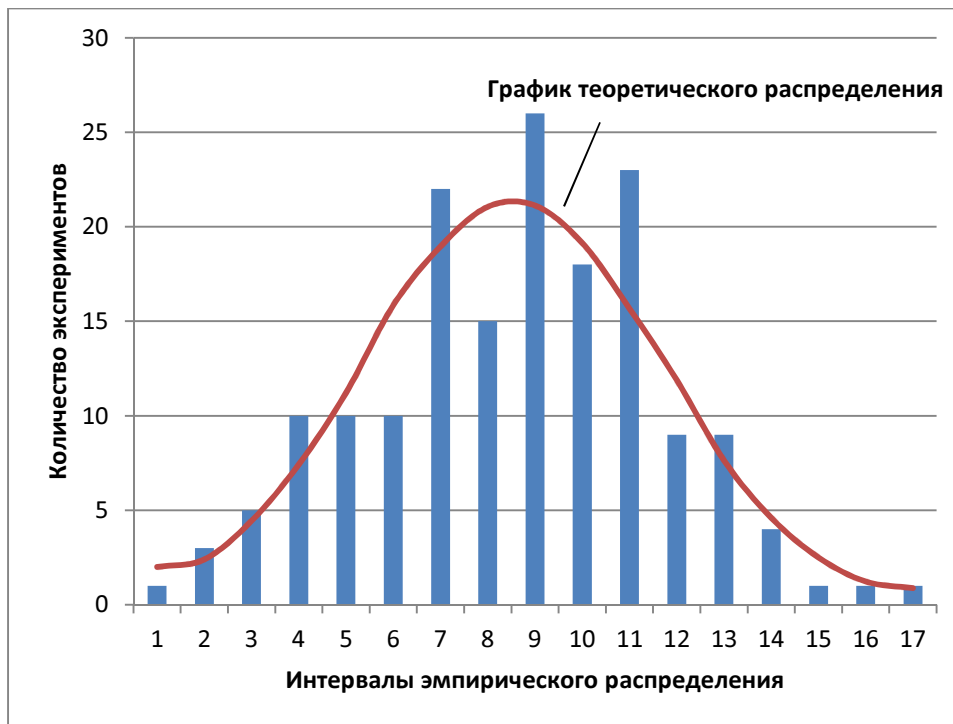


Рис. 3. – Эмпирическое распределение экспериментальных данных отклонения от тренда

На рисунке 3 визуально наблюдаем подобие эмпирического распределения нормальному закону распределения случайных величин. Это предположение не является фактом, поэтому выдвинем его в качестве гипотезы  $H_0$ .

**Доказательство гипотезы подчинения производственных данных нормальному закону распределения**

Для проверки этой гипотезы используем критерий согласия Пирсона  $\chi^2$ . Использование критерия  $\chi^2$  предусматривает разбиение размаха варьирования выборки на интервалы и определения числа наблюдений (частоты)  $n_j$  для каждого из  $e$  интервалов. Для удобства оценок параметров распределения интервалы выбирают одинаковой длины. Для экспериментального ряда, включающего от 100 до 200 наблюдений, рекомендуют выбирать количество интервалов от 15 до 20. Количество экспериментальных отчетов, в нашем случае, составляет 168, поэтому возьмём 17 интервалов. Интервалы, содержащие менее пяти наблюдений, объединяют с соседними. Однако, если число таких интервалов составляет менее 20% от их общего количества, допускаются интервалы с частотой  $n_j \geq 2$ . Оценка  $\chi^2$  вычисляется по формуле:

$$\chi^2 = \sum_{j=0}^e \frac{(n_j - np_j)^2}{np_j} \quad (2)$$

где  $p_j$  - вероятность попадания изучаемой случайной величины в  $j$ -й интервал, вычисляемая в соответствии с теоретическим законом распределения. При вычислении вероятности  $p_j$

нужно иметь в виду, что левая граница первого интервала и правая последнего должны совпадать с границами области возможных значений случайной величины. При нормальном распределении первый интервал изменяется до  $-\infty$ , а последний - до  $+\infty$ .

Нулевую гипотезу о соответствии выборочного распределения теоретическому закону проверяют путем сравнения вычисленной по формуле (2) величины с критическим значением  $\chi^2_\alpha$ , найденным по таблице критических точек распределения Пирсона для уровня значимости  $\alpha$  и числа степеней свободы  $k = e_1 - m - 1$ . Здесь  $e_1$  - число интервалов после объединения;  $m$  - число параметров, оцениваемых по рассматриваемой выборке (для нормального распределения  $m = 2$ ). Если выполняется неравенство:

$$\chi^2 \leq \chi^2_\alpha \quad (3)$$

то нулевую гипотезу не отвергают. При несоблюдении указанного неравенства принимают альтернативную гипотезу о принадлежности выборки неизвестному распределению.

Таким образом, принимаем предположение о подчинении данных о надоях (рис. 3) нормальному распределению за гипотезу  $H_0$ , требующую доказательства. Проверяя гипотезу с помощью критерия согласия  $\chi^2$  примем уровень значимости  $\alpha = 0,05$ .

Результаты расчетов приведены в таблице 1.

Таблица 1 – Расчет значения критерия согласия Пирсона  $\chi^2$

Отчеты эмпирического распределения		Пределы эмпирического распределения		Пределы теоретического распределения		Значение функции Лапласа на границах интервала $\Phi(z_j)$		Оценка вероятности попадания в интервал $p_j$	$\frac{(n_j - np_j)^2}{np_j}$
Исходное	С объединением	Левый предел	Правый предел	Левый предел	Правый предел	Левый предел	Правый предел		
1	2	3	4	5	6	7	8	9	10
1	9	-4172	-2660	$-\infty$	-2,26	0,000	0,012	0,053	0,003
3		-2660	-2282	-2,26	-1,94	0,012	0,026		
5		-2282	-1903	-1,94	-1,62	0,026	0,053		
10	10	-1903	-1525	-1,62	-1,30	0,053	0,097	0,044	0,895
10	10	-1525	-1147	-1,30	-0,98	0,097	0,164	0,067	0,131
10	10	-1147	-769	-0,98	-0,65	0,164	0,258	0,094	2,156
22	22	-769	-391	-0,65	-0,33	0,258	0,371	0,113	0,488
15	15	-391	-13	-0,33	-0,01	0,371	0,496	0,125	1,740
26	26	-13	365	-0,01	0,31	0,496	0,622	0,126	1,128
18	18	365	743	0,31	0,63	0,622	0,736	0,114	0,068
23	23	743	1121	0,63	0,95	0,736	0,829	0,093	3,426
9	9	1121	1499	0,95	1,28	0,829	0,900	0,071	0,704
9	9	1499	1877	1,28	1,60	0,900	0,945	0,045	0,242
4	7	1877	2255	1,60	1,92	0,945	0,973	0,055	0,529
1		2255	2633	1,92	2,24	0,973	0,987		
1		2633	3011	2,24	2,56	0,987	0,995		
1		3011	3389	2,56	$+\infty$	0,995	1,000		
Всего $n = 168$						$\Sigma = 1$		$\chi^2 = 11,506$	

Изначально исходные данные о надоях были предварительно обработаны: рассчитаны параметры линейной функции  $Y(t)$ . Разность данных о наблюдениях и значений функции тренда в точках наблюдения позволила получить ряд значений функции  $E(t)$ . Отсчеты этой функции с помощью операции квантования разбиты на промежутки (столбец 1, таблица 2), и получены соответствующие границы интервалов (столбец 3, 4, таблица 2) и количество вхождений наблюдений в эти интервалы (столбец 2 Исходное). В 5-ом, 6-ом столбцах показаны результаты вычисления границ интервалов, выраженные через нормированную случайную величину:

$$z_j = \sum_{j=0}^e \frac{x_j - x_{cp}}{s} \quad (4)$$

где  $x_{cp}$  и  $s$  – выборочное среднее значение и среднее квадратическое отклонение выборочных данных, соответственно. Значения этих оценок были найдены для исходных данных наблюдений заранее и представлены в таблице 2.

Таблица 2 – Основные статистические параметры, необходимые для вычисления значения критерия согласия Пирсона  $\chi^2$

Минимальное значение отклонения от тренда надоев	-4171,80729
Максимальное значение отклонения от тренда надоев	3389,225627
Среднеквадратическое отклонение линии отклонения от тренда надоев	1175,355677
Среднее значение отклонений от тренда надоев	-0,00096104

Находим значения функции Лапласа для границ интервалов и заносим их в 7-ой, 8-ой столбцы. Оценка вероятности попадания значений в интервалы (10-ый столбец) представляет собой разность значений функции Лапласа на правой и левой границе интервала. Если интервалы объединяются, вычисляют разность значений функции на границах объединенного интервала. Сумма значений  $p_j$ , в столбце 6 всегда будет равна единице. В 7-ой столбец заносят оценки математических ожиданий числа наблюдений по интервалам, которые определяем умножением оценки вероятности  $p_j$  на общее число образцов в выборке  $n = 100$ .

### Результаты

Таким образом, рассчитаем значение  $\chi^2$  по формуле, используя значения 9, 10 столбцы при уровне значимости  $\alpha = 0.05$  и количестве степеней свободы  $k = 11 - 2$ , где 2 – число параметров функции нормального распределения.

После расчетов значение величины  $\chi^2 = 11,5060317$ .

При  $e = 12$  и  $\alpha = 0,05$  для доказательства заявленной гипотезы о подчинении исходных данных нормальному закону распределения необходимо исходя из таблицы значений  $\chi^2[2]$ , чтобы полученное значение было меньше 16,9.

Таким образом, неравенство:

$$\chi^2 \leq 16,9$$

верно, а следовательно верно и утверждение:

$$\chi^2 \leq \chi^2_{\alpha}$$

Что позволяет сделать вывод о доказательстве выдвинутой в начале исследований гипотезы.

### **Заключение**

Доказательство факта подчинения числового ряда наблюдаемых значений нормальному закону распределения позволяет разработать математическую модель отслеживания случайных всплесков значений. На практике внеплановые всплески представляют собой резкие отклонения наблюдаемых значений от нормы, что в условиях реального производства может означать резкое ухудшение здоровья животных, качества выпускаемой продукции, и, как следствие этого, финансовые потери. На основе проведенного исследования и подобных исследований других важных параметров животных был разработан метод отслеживания случайных всплесков в значениях различных параметров. Данный метод лежит в основе алгоритма мониторинга отклонений значений характеристик производственного процесса и состояния животных. Разработанный алгоритм реализован в системе в рамках интеллектуального программного модуля контроля производственного процесса.

### **Список литературы**

1. Администрация владимирской области: постановление губернатора №170 [Электронный ресурс] : (О прогнозе социально-экономического развития Владимирской области на 2012 год и плановый период 2013 и 2014 годов) - 2011. - Режим доступа: [www.avo.ru/documents/docs](http://www.avo.ru/documents/docs)
2. Белялетдинова М., Светлов Н. Экономические стимулы к росту надоев // Известия ТСХА. 2010. в.6, 138-146 с.
3. Ваттио М. Принципы отбора / М. Ваттио. – Висконсин 1996. – 150 с.
4. Стряпихин А. Молочная отрасль: пути повышения рентабельности [Электронный ресурс] // DairyNews: сайт. - URL:



[http://www.dairynews.ru/news/molochnaja\\_otrasl\\_puti\\_povyshenija\\_rentabelnosti.html?PAGEN\\_2=2&PAGEN\\_3=1904](http://www.dairynews.ru/news/molochnaja_otrasl_puti_povyshenija_rentabelnosti.html?PAGEN_2=2&PAGEN_3=1904) (дата обращения: 17.09.2013).

5. Abdallah J.M., McDaniel B.T. Genetic parameters and trends of milk, fat, days open and body weight after calving in North Carolina experimental herds // *Journal of Dairy Science*. 2000. Vol. 83, P.1364-1370.
6. Cappio-Borlino A., Macciotta. N.P.P., Pulina G. The shape of Sarda ewe lactation curve analyzed with a compartmental model. // *Livest. Prod. Sci.*. 1991. Vol. 51, P. 89-96.
7. Grzesiak W., Lacroix R., Wojcik J., and Blaszczyk A. A comparison of neural network and multiple regression predictions for 305-day lactation yield using partial lactation records // *Can. J. Animal Sci.*. 2003. Vol. 83, P. 307–310.
8. Hossein-Zadeh N.G. Estimation of genetic and phenotypic relationships between age at first calving and productive performance in Iranian Holsteins // *Tropical Animal Health Production*. 2011. Vol. 43, P.967-973.
9. Khazaei J., Nikosiar M. Approximating milk yield and milk fat and protein concentration of cows through the use of mathematical and artificial neural networks models. // *World conference on agricultural information and IT, IAALD AFITA WCCA 2008 (Tokyo University of Agriculture, Tokyo, Japan, 24 - 27 August 2008)*. – Tokyo, P. 91-105.
10. Malhado C., Malhado A., Ramos A., Carneiro P., Souza J., Pala A. Genetic parameters for milk yield, lactation length and calving intervals of Murrah buffaloes from Brazil // *R. Bras. Zootec.* 2013. Vol. 42, no.8.

**Рецензенты:**

Жизняков А.Л., д.т.н., профессор, зав. кафедрой САПР МИ (ф), г. Муром.

Андрианов Д.Е., д.т.н., доцент, зав. кафедрой ИС МИ (ф), г. Муром.