

УДК 621.3.049.77

ПРЕДВАРИТЕЛЬНАЯ ОБРАБОТКА ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ В СТАТИСТИЧЕСКОМ ПРОЕКТИРОВАНИИ ИЗДЕЛИЙ МИКРОЭЛЕКТРОНИКИ

П.Г. ПАВЛОВ, ЧАН ТУАН ЧУНГ, А.И. КОСТРОВ, В.В. НЕЛАЕВ, В.Р. СТЕМПИЦКИЙ

*Белорусский государственный университет информатики и радиоэлектроники
П. Бровка, 6, Минск, 220013, Беларусь*

Поступила в редакцию 30 октября 2012

Представлены результаты исследований по выбору и практической реализации математических методов предварительного анализа и статистической обработки экспериментальных данных, полученных при обработке и интерпретации промышленных измерений результатов формирования элементов интегральных микросхем (ИМС). Предлагаемая методика позволяет исключать аномальные наблюдения в выборочных данных и проводить процедуру сглаживания результатов измерения или моделирования.

Ключевые слова: интегральная микросхема, статистический анализ, статистическое распределение, методы сглаживания, критерий случайности, тренд.

Введение

Достижение высокого уровня технологичности производства и максимального выхода годных изделий – основные задачи для технологов и разработчиков ИМС. В процессе опытной отработки технологических процессов изготовления ИМС, а также при проектировании прибора/схемы/системы инженеру-проектировщику важно учитывать и оптимизировать множество факторов (флуктуации параметров технологических процессов формирования ИМС, адекватность и корректная интерпретация результатов измерений тестовых структур), оказывающих влияние на выходные характеристики создаваемых изделий. Эффективным инструментом решения указанных задач является проведение статистического анализа и оптимизации на всех этапах проектирования технологии/прибора/схемы/системы [1, 2].

Результаты измерения и моделирования, используемые на различных этапах изготовления и проектирования ИМС для расчета статистических оценок, проведения статистического анализа и аппроксимации, зачастую содержат ошибки и погрешности, которые могут существенно сказываться на адекватности результатов статистических исследований. Кроме того, процедура формирования данных не всегда соответствует предъявляемым требованиям, что приводит к неоднородным и неслучайным наблюдениям. Наконец, полученные данные могут отклоняться от предполагаемой исследователем статистической модели, что исключает возможность корректного статистического анализа.

На практике для предотвращения подобных случаев необходимо проводить предварительный анализ и статистическую обработку полученных экспериментальных результатов. Предварительная обработка результатов измерений или компьютерного проектирования необходима для того, чтобы в дальнейшем с наибольшей эффективностью, а главное – корректно, использовать для построения эмпирических зависимостей статистические методы. Соответственно задача анализа сводится к выбору наилучшей процедуры обработки данных. Вопросы, решаемые при обработке результатов эксперимента, включают подбор эмпирических формул и оценку их параметров, оценку истинных значений измеряемых величин и точности измерений,

применимость методов математической статистики и теории вероятностей и, наконец, проведение статистической обработки без потери важной информации в данных.

В работе представлены результаты исследований по выбору и практической реализации математических методов предварительного анализа и статистической обработки экспериментальных данных. Основными задачами при этом являются разработка программной библиотеки статистических методов и апробация выбранных методов на данных, полученных при промышленных измерениях результатов формирования элементов ИМС.

Методика

Для пояснения роли и места основных приемов статистического моделирования и методов первичной статистической обработки исходных данных ее проведение следует условно разбить на основные этапы исследования. Предварительная статистическая обработка данных не предусматривает строго определенного алгоритма, то есть последовательность действий определяется исследователем на основании предварительного анализа экспериментальных данных, представленных графически или в виде таблицы, а также исходя из поставленных целей. Далее каждый из возможных этапов рассмотрен более подробно.

1. *Графическое представление и предварительный анализ данных.*

На первом этапе осуществляется анализ исходных данных, в частности, изучается характер распределении случайных величин, входящих в выборку. Представление данных в виде гистограмм, временных диаграмм, упорядоченных таблиц позволяет исследователю выбрать критерии оптимальные для решения конкретно поставленной задачи.

2. *Проверка соответствия выбранному закону распределения.*

При практическом применении методов математической статистики знание закона распределения вероятностей для исследуемой выборки чрезвычайно важно [3, 4]. Именно поэтому любая обработка результатов наблюдений должна неизменно начинаться с определения закона распределения вероятностей исследуемого набора экспериментальных данных. Соответствие исходных данных выбранной статистической модели (распределению) является обязательным условием адекватного статистического анализа, которое дает возможность эффективно решать практические задачи, связанные с планированием производства, обеспечением качества продукции, оценкой технологичности производства, а также выбора адекватной статистической модели. Основные доверительные оценки, применяемые в математической статистике, основаны на гипотезе нормальности закона распределения случайных ошибок измерений и могут применяться лишь до тех пор, пока результаты эксперимента не противоречат этой гипотезе [5]. Наиболее эффективными и устойчивыми критериями проверки соответствия нормальному (гауссовому) закону распределения являются специальные критерии Д'Агостино и Мартинеса-Иглевича. Для получения однозначных и корректных выводов возможно использование критериев Дэвида-Хартли-Пирсона, Шапиро-Уилка, критериев симметрии (асимметрии) и эксцесса [3, 6].

3. *Проверка независимости наблюдений.*

Проверка статистической случайности исходных данных позволяет выявить ошибки, связанные с методом формирования выборки. Наличие закономерности или корреляции между наблюдениями приводит к выводу о необходимости более тщательного анализа процесса формирования данных для обработки.

4. *Исключение резко выделяющихся наблюдений (выбросов).* Экспериментальные данные необходимо проверять на наличие в выборке явных отклонений от условий, при которых предполагалось использовать процедуру статистического анализа и оптимизации с целью их правильной интерпретации [7]. Результаты измерений, содержащие грубые ошибки, обычно хорошо заметны, так как они сильно отличаются от других данных, однако выбросами могут оказаться и данные, визуально не отличимые от основной массы наблюдений. В связи с вышесказанным следует проводить дополнительный анализ резко выделяющихся наблюдений с точки зрения физической сущности наблюдаемого явления [8, 9].

Статистические процедуры выявления резко выделяющихся результатов наблюдений основаны на предположении однородности данных и необходимости исключения данных, нетипично далеко удаленных от центра распределения. При использовании статистических методов выделения выбросов следует иметь в виду, что выбросы могут оказаться наиболее суще-

ственной частью выборки, описывающей характер изменения исследуемой характеристики, и аномальные наблюдения не следует отвергать (исключать) целиком [6]. Здесь необходимо использовать специальные критерии, использующие оценки, устойчивые к отклонениям от нормальности, такие как критерии Шовене, Смирнова-Граббса, Титьена-Мура, Роснера [8]. Следует отметить, что выбор и применение конкретного критерия зависит от цели исследователя – либо это выделение самих выбросов, либо определение их количества.

5. Проверка гипотезы о наличии тренда в ряду оценок.

Важным моментом, требующим особого внимания, является учет и анализ характера зависимости между данными производимых измерений и изменение этой зависимости во времени. В этом случае эффективно использование математического аппарата, относящегося к исследованию временных рядов. Однако при изучении больших и малых последовательностей измерений как правило используются несколько различные подходы [10]. Для описания временных последовательностей применяют понятие тренда (закономерности, а не случайности появления ряда данных), который представляет собой полиномиальную зависимость исследуемой характеристики от времени [11]. Наибольшее применение критерии тренда находят при статистическом контроле и регулировании технологических процессов, позволяя заранее статистически обоснованно выявить тенденцию ухудшения качества продукции.

6. Сглаживание ряда оценок.

Многие задачи анализа вовсе не требуют единой аналитической формулы для всех данных, т.е. необходимо лишь устранить «шум» эксперимента, сохранив информацию об истинной функции путем сглаживания эмпирических данных [5]. Хотя крайние точки временного ряда обычно не представляют большого интереса, их «несглаженные» значения могут приводить к серьезным проблемам, особенно в случаях, когда для достижения необходимого сглаживания требуется большое значение экспериментальных измерений [12]. Сглаживание временного ряда означает представление тренда в данной точке посредством взвешенных средних значений, наблюдаемых в окрестности этой точки, т.е. «нерегулярный» график наблюдений заменяется, например, гладким графиком скользящего среднего, который является стандартным методом проведения процедуры сглаживания экспериментальных данных. Этот метод, несмотря на простоту реализации, имеет такой недостаток, как высокая чувствительность среднего значения к одиночным выбросам. Метод экспоненциального сглаживания позволяет решить многие проблемы, связанные с алгоритмом скользящего среднего. Наиболее оптимальным можно назвать метод скользящей медианы или медианной фильтрации. Его преимуществом является отсутствие необходимости вносить допущения относительно характера распределения шума.

Использование методов сглаживания эмпирических данных эффективно при подготовке полученных статистических оценок для дальнейшего регрессионного анализа и аппроксимации. Так сглаживание выходных данных в методе поверхности отклика позволяет получать аппроксимирующий полином с меньшими относительными и абсолютными погрешностями. Конкретный метод сглаживания или фильтрации выбирается исходя из характера экспериментальных данных и требуемой степени обработки данных [11–14].

Результаты и их обсуждение

Эффективность предложенной методики проверена на примере предварительной статистической обработки результатов измерения параметров изделий микроэлектроники.

В производственных условиях филиала «Завод полупроводниковых приборов» ОАО «Интеграл» разрабатывалась и исследовалась стойкая к воздействию субмикронная технология и элементная база для изготовления КМОП ИМС с проектными нормами 0,35 мкм, предназначенными для использования в устройствах специального применения.

На одном из этапов отработки технологии был предусмотрен эксперимент по реализации диффузионных резисторов в различных слоях (n^+ , p^+ , нелегированный поликристаллический кремний (ПКК), легированный ПКК) с формированием силицида (полицида) титана для снижения контактного сопротивления. Анализ начинался с контроля внешнего вида изготовленных пластин на наличие явных дефектов, на следующем этапе проводился замер вольт-амперных характеристик (ВАХ). После ручной проверки правильности замеров проводилась

обработка результатов измерений и выбор образцов для более глубокого анализа ВАХ.

Следует отметить, что для измерений цепочек контактов и сопротивлений диффузионных резисторов типовая методика измерений ВАХ с отбраковкой по нормам, соответствующим технической документации, оказалась непригодной. Вместо предполагаемого измерения путем задания тока и снятия напряжения было введено измерение на напряжении 1,5 В и 3 В с дальнейшим сравнением полученных результатов между собой (считая сопротивление линейным при разнице менее 10 %). Вследствие этого потребовался углубленный анализ и статистическая обработка результатов измерений с целью выделения области воспроизводимости характеристик приборов.

Из всей совокупности экспериментальных данных было отобрано 6 выборок из 50 наблюдений и 6 выборок из 10 наблюдений, при этом выбирались измерения с резко различающимися свойствами.

А. Статистическая обработка выборок большого объема.

На первом этапе проверки предложенной методики предварительной обработки данных использовались выборки относительно большого объема. В качестве исходных данных использовались результаты измерения поверхностного сопротивления резисторов в слое «База», Ом (выборки X1, X4, X6); поверхностного сопротивления резисторов в слое « n^+ стоки», Ом (выборки X2, X5); поверхностного сопротивления резисторов в слое «Нелегированный ПКК», кОм (выборка X3).

Для оценки закона распределения данных выборок строятся гистограммы распределения. Однако без предварительной обработки результатов измерений данные могут содержать выбросы, что существенно искажает их графическое представление (рис. 1).

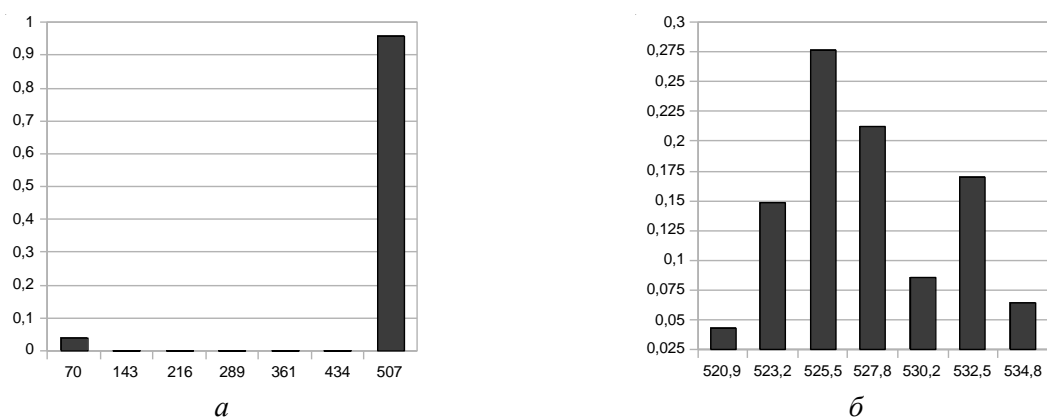


Рис. 1. Гистограммы распределения выборки X4
а – до исключения выбросов; б – после исключения выбросов

Проверка отклонения от нормального закона распределения исследуемых выборок осуществлялась с помощью приведенных выше критериев с различным уровнем доверительной вероятности. В целом можно отметить отклонение от нормальности закона распределения выборок X3, X4, X5, X6, однако при этом критерии Дэвида-Хартли-Пирсона и Шапиро-Франция иногда показывали результаты отличные от общей массы оценок.

Представление о характере наблюдений также можно получить, представив их в виде временного ряда (рис. 2, а). Вместе с временным рядом горизонтальной линией отмечено среднее арифметическое значение всей выборки. При таком представлении данных иногда возникают проблемы, связанные с наличием резко выделяющихся результатов наблюдений (рис. 2, б).

Тем не менее, диаграммы позволяют судить о некоторых свойствах опытных данных. Так для выборок X2 и X3 заметно наличие характерного тренда, а результаты измерений в выборке X6 явно неоднородны.

Эффективным средством для проверки гипотезы о наличии тренда являются критерии случайности. Применение критериев Аббе-Линника, Кокса-Стюарта, автокорреляции и критериев «восходящих» и «нисходящих» серий для различных уровней значимости позволяет сделать вывод о наличии тренда в выборках X2 и X3. Кроме того для низкого уровня значимости не подтвердилась гипотеза о случайном характере выборки X5.

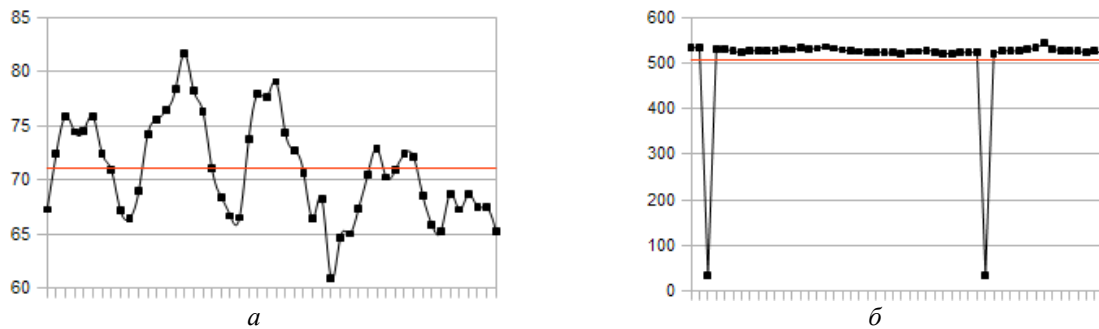


Рис. 2. Временной ряд исследуемых выборок
 а – выборка X2; б – выборка X4

Следует отметить, что использование критериев случайности не всегда позволяет получить справедливые оценки. Это обусловлено тем, что характеристики полупроводниковых приборов не распределены на пластине случайным образом, а имеют четкую пространственную зависимость, связанную с технологией изготовления приборов, а также со свойствами самой пластины. Часто применяемый метод последовательного измерения случайных точек может приводить к получению циклического характера таких характеристик. Таким образом, критерии тренда и случайности скорее позволяют исследовать свойства метода формирования выборки, нежели свойства самих наблюдений.

Даже при визуальной оценке временных рядов становится очевидным наличие выбросов в выборках X4 и X5. Для получения объективной информации о наличии выбросов в выборках следует воспользоваться критериями Шовене, Смирнова-Грabbса, Титьена-Мура, и Роснера. Для методов исключения одного резко выделяющегося результата наблюдения был применен подход, при котором сначала отсеивалось одно anomальное значение, после чего процедура проверки повторялась. При реализации метода Титьена-Мура количество подозрительных наблюдений выбиралось исходя из оценок других критериев.

Следует отметить, что ни один из предложенных критериев не обеспечивает получение адекватных результатов для всех рассматриваемых выборок, хотя предпочтительнее других выглядит критерий Роснера, который кроме того удобен тем, что не требует итеративного подхода и позволяет автоматически оценивать количество выбросов в выборке. Тем не менее, при его использовании возникают трудности при анализе выборок с незначительной дисперсией основной массы (выборка X4). Наиболее строгим из предложенных критериев является критерий Смирнова-Грabbса, исключающий только существенные отклонения в ряду данных. Критерий Шовене представляет особый интерес в силу того, что для выборки X6 ему удалось выделить неоднородность в наблюдениях.

Таким образом, корректное решение об исключении результатов резко выделяющихся наблюдений можно принять только на основании анализа всех критериев. Исключив из выборки anomальные наблюдения, можно получить адекватное представление выборки X4.

Для анализа же выборки X6 необходимо, исходя из интуитивных соображений, разделить всех данных на две подвыборки без исключения каких-либо подозрительных наблюдений. Дальнейший анализ подтвердил предположение о сложном характере экспериментальных данных и наличии разрыва в ряду наблюдений.

Б. Статистическая обработка выборок малого объема.

На втором этапе проверки предложенной методики предварительной обработки данных использовались выборки малого объема. В качестве исходных данных использовались результаты измерения поверхностного сопротивления резисторов в слое «Доп. охрана», кОм (выборка X1); поверхностного сопротивления резисторов в слое «База», Ом (выборки X2); тока стока n -канального МОП-транзистора, мА (выборка X3); порогового напряжения n -канального МОП-транзистора, мВ (выборки X4, X6); порогового напряжения p -канального МОП-транзистора, мВ (выборка X5).

Следует отметить, что графическое представление результатов измерений как в виде гистограмм распределения, так и в виде временных рядов не дает представление о свойствах данных: законе распределения, случайности выборочных данных и наличии неоднородности.

Применение статистических процедур для проверки соответствия нормальному закону распределения, а также гипотезы о наличии тренда также неэффективно для выборок малого объема. Во всех случаях, независимо от структуры данных, отклонение от нормального закона распределения не было подтверждено, кроме того с высокой доверительной вероятностью принята гипотеза о случайности наблюдений для всех выборок. Исключение из набора данных с резко выделяющимися результатами наблюдений принципиально не меняет статистические оценки и выводы.

Однако процедура выделения выбросов в ряду данных имеет свою специфику для выборок малого объема. Так критерий Роснера в данном случае оказался наименее эффективным, при использовании данного критерия возникает ошибка исключения результатов наблюдений, не являющихся аномальными для рассматриваемой выборки. В свою очередь точные результаты продемонстрировали критерии Граббса и Смирнова-Граббса, которые к тому же оказались весьма чувствительными к уровню доверительной вероятности, что позволяет проводить процедуру исключения резко выделяющихся наблюдений максимально гибко.

В. Сглаживание экспериментальных данных.

На заключительном этапе был проведен сравнительный анализ различных методов сглаживания и фильтрации экспериментальных данных. Кроме того, в ходе данного анализа были адаптированы теоретические методы для обработки данных измерения параметров приборов микроэлектроники.

Базовым методом для сглаживания данных является метод скользящего среднего. Управление процедурой сглаживания для этого метода осуществлялось посредством задания размера окна сглаживания. Метод скользящего среднего оказался эффективным лишь для временных рядов, в которых отсутствует резкий тренд, при этом получается линейная функция сглаживания в значительной степени приближенная к исходной форме ряда. При использовании метода скользящего среднего невозможно эффективно устранять резкие отклонения (выбросы) во временном ряду, кроме того несглаженными остаются краевые точки в начале и конце выборки.

Назначение различных весов для наблюдений в пределах окна сглаживания позволяет в меньшей степени исказить исходный временной ряд, но не решает основные проблемы метода. В некотором смысле, экспоненциальное сглаживание является разновидностью метода скользящего среднего с экспоненциально убывающими весом. Метод априорно позволяет сглаживать краевые точки ряда, а также обладает наибольшей гибкостью в управлении процессом сглаживания. Однако данный метод также не позволяет исключать резко выделяющиеся наблюдения во временном ряду. Наиболее эффективен метод экспоненциального сглаживания для относительно плавных временных рядов.

Анализ метода скользящей медианы приводит к выводу, что обладая недостатками метода скользящего среднего он с более высокой эффективностью устраняет выбросы в наборе данных, а кроме того позволяет использовать окна сглаживания четной размерности.

Для того, чтобы устранить отсутствие сглаживания по краям исследуемого набора экспериментальных данных, был разработан метод оптимизированной медианной фильтрации, при которой набор данных рассматривается не как временной ряд, а как замкнутая последовательность данных (рис. 3). Как и в случае стандартного метода скользящей медианы, процесс сглаживания контролируется шириной окна сглаживания (параметр ω).

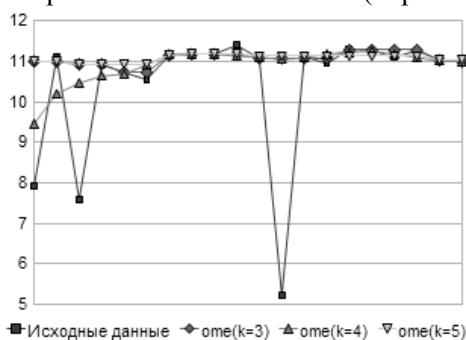


Рис. 3. Результат применения метода оптимизированной скользящей медианы

Заключение

В результате проведенных исследований проанализированы и выбраны наиболее эффективные с точки зрения алгоритмизации и применения в микроэлектронике методы предварительного анализа и статистической обработки данных, получаемых как в процессе проектирования, так и в процессе изготовления ИМС.

Используемая методика предварительного анализа и статистической обработки позволяет не только исключать аномальные результаты наблюдений в выборочных данных, но и проводить процедуру сглаживания результатов измерения или моделирования. Это дает возможность сохранить экспериментальные данные в полном объеме, что зачастую является обязательным условием проводимого исследования.

Эффективность предложенной методики продемонстрирована при обработке и интерпретации результатов измерения параметров полупроводниковых приборов. Следует отметить, что учет полученных результатов, а также статистическая обработка на всех последующих этапах проектирования и изготовления ИМС, позволяют значительно повысить эффективность и точность статистического анализа и оптимизации в сквозном цикле проектирования технологии/прибора/схемы/системы.

PRELIMINARY EXPERIMENTAL DATA PROCESSING IN STATISTICAL IC DESIGN

P.G. PAVLOV, TRAN TUAN TRUNG, A.I. KOSTROV, V.V. NELAYEV, V.R. STEMPIISKY

Abstract

The results of researches on a choice and implementation of mathematical methods of the preliminary analysis and statistical processing of experimental data are presented. Data are received at processing and interpretation of industrial measurements of results from integrated circuits production. This technique allows not only to exclude the abnormal observations in the sample data, but also to carry out the procedure of smoothing the results of measurements or simulations.

Список литературы

1. *Krasikov M., Nelayev V., Stempitsky V. et. al.* // Proc. of SPIE. 2009. Vol. 7377-40.
2. *Кулешов А.А., Нелаев В.В., Красиков М.Г. и др.* // Матер. IV-й МНТК «Проблемы проектирования и производства радиоэлектронных средств», Новополюцк, 2008. С. 96.
3. *Кобзарь А.И.* Прикладная математическая статистика. М., 2006.
4. *Гайдышев И.П.* Анализ и обработка данных: специальный справочник. СПб, 2001.
5. *Румшинский Л.З.* Математическая обработка эксперимента. М., 1971.
6. *Дэйвид Г.А.* Порядковые статистики. М., 1979.
7. *Beale E.M. L., Little R.J.A.* // Journ. Royal. Statist. Soc. 1975. Ser. B. Vol. 37. P. 129–145.
8. *Norman H.N., Bent D.H., Hull C.H.* Statistical Package for the Social Sciences (SPSS). New York, 1975.
9. *Эбенсен К.* Анализ многомерных данных. Избранные главы. Черногловка, 2005.
10. Справочник по прикладной статистике. Т.2 / Под ред. Э. Ллойда. М., 1990.
11. *Андерсон Т.* Статистический анализ временных рядов. М., 1976.
12. *Lima J. N., Casaca J.* Smoothing GNSS Time Series with Asymmetric Simple Moving Averages. Symposium on Deformation Measurement and Analysis. Lisbon, 2008. P. 1–8.
13. *Тьюки Дж.* Анализ результатов наблюдений. Разведочный анализ. М., 1981.
14. *Ruey-Chyn Tsaur* // Journal of the Chinese Institute of Industrial Engineers. 2005. Vol. 22, Is. 6. P. 521–530.