УДК 004.22:620.22

В. В. НАПРАСНИКОВ [1], Ю. В. НАПРАСНИКОВА [1], Ю. В. ПОЛОЗКОВ [2], А. В. БОРОДУЛЯ [1], Д. П. КУНКЕВИЧ [1]

ОПРЕДЕЛЕНИЕ И УСТРАНЕНИЕ ГЕТЕРОСКЕДАСТИЧНОСТИ МНОГОФАКТОРНОЙ РЕГРЕССИОННОЙ МОДЕЛИ ПРОГНОЗИРОВАНИЯ ПАРАМЕТРОВ ЯЧЕИСТЫХ СТРУКТУР СРЕДСТВАМИ MATHCAD

[1] Белорусский национальный технический университет
г. Минск, Республика Беларусь
[2] Витебский государственный технологический университет
г. Витебск, Республика Беларусь

Аннотация. *Описывается возможный подход к решению задачи получения прогноза выходных параметров ячеистых структур при тех сочетаниях значений параметров объекта, которые находятся в заданных ранее диапазонах по уже проведенным исследованиям, на основе использования полученных ранее результатов вариантных вычислений. Приводятся фрагменты набора программ на основе использования MATHCAD для реализации такого подхода. Выполнены сравнения способов устранения гетероскедастичности при различном выборе весовых коэффициентов на основе использования взвешенного метода наименьших квадратов.*

Ключевые слова: *ячеистые структуры, взвешенный метод наименьших квадратов, гетероскедастичность*

Введение

В настоящее время при проектировании всевозможных изделий все чаще применяются ячеистые конструкции. Это объясняется тем, что на передний план выдвигаются требования снижения материалоемкости изделий при сохранении жесткостных, прочностных и других характеристик этих изделий в допускаемых диапазонах. Примерами таких объектов являются подошвы облегченного типа спортивной обуви, ячеистые прокладки для летательных аппаратов и другие машиностроительные конструкции [1].

Вид подобных изделий представлен на рисунке 1.

Рисунок 1. Примеры изделий ячеистой структуры

Поскольку снижение материалоемкости изделия приемлемо только при сохранении требований прочности и жесткости, а внешние нагрузки могут изменяться как в пространстве, так и во времени, однократное выполнение анализа занимает много времени. Если же при этом предполагается и проведение оптимизационных расчетов, то решение подобных задач становится проблематичным.

В том случае, когда такие массированные вычисления по вариантному моделированию уже проведены, имеет смысл использовать полученные результаты для получения прогноза выходных параметров при тех сочетаниях значений параметров объекта, которые находятся в заданных ранее диапазонах по уже проведенным исследованиям.

Одним из возможных способов является применение модели множественной регрессии.

Особенности построения линейной модели множественной регрессии

Линейную модель множественной регрессии запишем в виде (1):

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon, \qquad (1)$$

где $\beta_0, \beta_1, \dots, \beta_k$ – коэффициенты регрессионной модели; ε – случайное слагаемое, называемое ошибкой модели или остатками.

Обозначим i-тое наблюдение зависимой переменной как y_i, а независимые переменные в этом наблюдении – $x_{i,1}$, $x_{i,2}$, …, $x_{i,k}$, в обозначении $x_{i,j}$ первый индекс i определяет номер эксперимента, а второй j – номер переменной. Тогда имеет место следующая модель наблюдений (2):

$$y_i = \beta_0 + \beta_1 x_{i,1} + \beta_2 x_{i,2} + \dots + \beta_k x_{i,k} + \varepsilon_i, \quad i = 1, 2, \dots, n. \qquad (2)$$

В матричном виде модель наблюдений примет вид (3):

$$y = X\beta + \varepsilon. \qquad (3)$$

Оценки неизвестных коэффициентов выполняются с использованием метода наименьших квадратов (МНК), на основе которого минимизируется сумма квадратов остатков:

$$\sum_{n=1}^{n} \varepsilon_i^2 \rightarrow min_\beta. \tag{4}$$

Предпосылки, которые должны выполняться при использовании линейной модели множественной регрессии (часть из них существенны именно для многомерной модели):

- объясняющие переменные являются детерминированными и линейно независимыми;
- математическое ожидание случайных ошибок равно нулю;
- дисперсия случайной ошибки одинакова для всех наблюдений;
- случайные ошибки, относящиеся к разным наблюдениям, взаимно независимы;
- случайные ошибки имеют нормальное распределение.

Таким образом, последовательность этапов при проектировании таких изделий может быть следующей:

- создание параметрической геометрической модели объекта с использованием современной CAD-системы;
- формирование граничных условий, изменяющихся как в пространстве, так и во времени;
- выполнение вариантных расчетов и отладка конечно-элементной модели с использованием современной CAE-системы;
- выполнение прогноза с целью получения выходных параметров проекта с использованием модели множественной регрессии.

Рассмотрим некоторые из них для представленного ниже примера.

Проверка гомоскедастичности модели

Как отмечено выше, одним из предположений, обеспечивающим адекватность модели, является условие о том, что дисперсия случайной ошибки одинакова для всех наблюдений. В этом случае модель обладает свойством гомоскедастичности, в противном случае модель является гетероскедастичной.

Для проверки наличия гетероскедастичности используют разные подходы. Рассмотрим пример модели при наличии 150 экспериментов (наблюдений), из которых предварительно исключены 7 промахов и 3 наблюдения зарезервированы для дальнейшей валидации модели. Таким образом n=140.

В качестве инструментария выберем среду MATHCAD. Это связано с тем, что при таком подходе документ, реализующий вычисления, записывается в форме привычной инженеру, а это существенно влияет на удобство работы исследователя.

Используем визуальную проверку гипотезы о наличии гетероскедастичности. На рисунке 2 показаны значения остатков. Вид графической зависимости позволяет предположить наличие гетероскедастичности.

Используем аналитические способы проверки гипотезы о наличии гетероскедастичности.

Выполним проверку по критерию Breushc-Pagan. Фрагмент документа представлен на рисунке 3 слева.

Для проверки с использованием LM-статистики вычисляем предварительно объясненную сумму квадратов ESS по вспомогательной регрессии:

$$\frac{\varepsilon_i^2}{\hat{\sigma}^2} = \gamma_0 + x_i^T \gamma + \nu_i, \tag{5}$$

затем значение LM-статистики как половину от ESS.

Фрагмент документа, реализующего проверку, представлен на рисунке 3 справа.

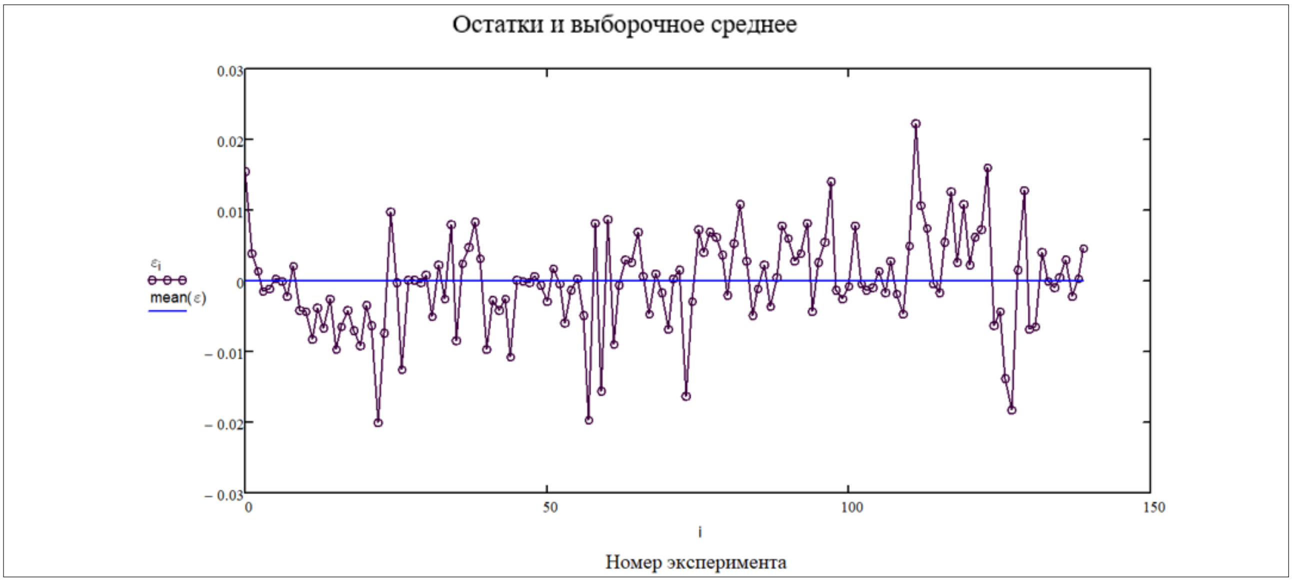


Рисунок 2. Остатки и выборочное среднее

χ_kriticeskoe := qchisq(1 − α, k) = 9.488

Наблюдаемое значение χ_nabl = 41.161

Критическое значение χ_kriticeskoe = 9.488

Если Наблюдаемое значение > критического, то присутствует ГЕТЕРОСКЕДАСТИЧНОСТЬ, иначе ГОМОСКЕДАСТИЧНОСТЬ

χ_kriticeskoe := qchisq(1 − α, k) = 9.488

Наблюдаемое значение χ_nabl = 64.607

Критическое значение χ_kriticeskoe = 9.488

Если Наблюдаемое значение > критического, то присутствует ГЕТЕРОСКЕДАСТИЧНОСТЬ, иначе ГОМОСКЕДАСТИЧНОСТЬ

Рисунок 3. Проверки по критерию Breushc-Pagan (слева) и LM-статистики (справа)

Выполним проверку на наличие гетероскедастичности для нескольких факторов по критерию Голдфелда-Квандта. Выборка при этом разбивается на три части, в которых количество компонентов вычисляется по формуле round(0.5*(n-CC)), где CC=4*n/15.

Текст основной процедуры для выполнения анализа представлен на рисунке 4.

Результаты тестирования для уровня значимости 0,05 представлены на рисунке 5.

```
Getero_GOLFILD(Tab, α) := | n ← rows(Tab)
                          | m ← cols(Tab)
                          | CC ← 4·n / 15
                          | K ← round((n − CC) / 2)
                          | y ← Tab^⟨0⟩
                          | for L ∈ 1..m − 1
                          |   | Nbeg ← 0, Nend ← K − 1
                          |   | S1_kw_L ← L_K1_K3(Tab, n, m, K, Nbeg, Nend, L)
                          |   | Nbeg ← n − K, Nend ← n − 1
                          |   | S3_kw_L ← L_K1_K3(Tab, n, m, K, Nbeg, Nend, L)
                          |   | "Строим F-статистику"
                          |   | F_L ← max(S1_kw_L, S3_kw_L) / min(S1_kw_L, S3_kw_L)
                          |   | "Критическая точка по Фишеру"
                          |   | MM ← 1, v1 ← K − MM − 1, v2 ← v1
                          |   | Fcrit ← qF(1 − α, v1, v2)
                          |   | Rez_L ← if(F_L > Fcrit, "Getero", "HOMO")
                          |   | (OTVET)_{0,0} ← "Fstatistica  "
                          |   | (OTVET)_{0,1} ← " Fcrit  "
                          |   | (OTVET)_{0,2} ← "   REZULT"
                          |   | (OTVET)_{L,0} ← F_L
                          |   | (OTVET)_{L,1} ← Fcrit
                          |   | (OTVET)_{L,2} ← Rez_L
                          | OTVET
```

Рисунок 4. Процедура проверки гетероскедастичности по критерию Голдфелда-Квандта

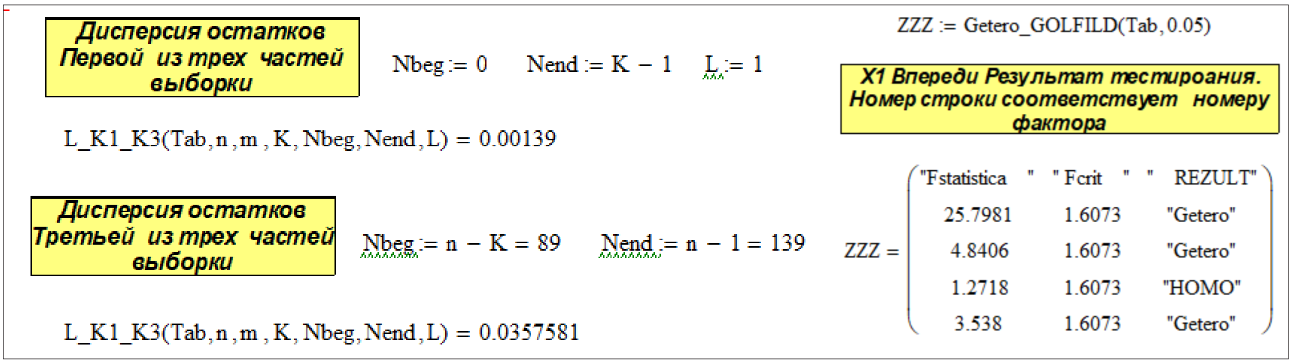


Рисунок 5. Результаты тестирования гетероскедастичности для нескольких факторов по критерию Голдфелда-Квандта для уровня значимости 0,05

Все выполненные тесты свидетельствуют о наличии гетероскедастичности.

Устранение гетероскедастичности модели

Для устранения гетероскедастичности используют взвешенный МНК таким образом, чтобы обеспечить постоянство дисперсии случайной ошибки для всех наблюдений.

Стандартный прием заключается в следующем. Сначала оценивают исходную модель при помощи обычного МНК и получают остатки регрессии. Затем оценивают эту вспомогательную модель, в которой i-тое наблюдение зависимой переменной заменяют на ε_i^2. Для нее получают свои остатки $\hat{\varepsilon}_i$. Их и используют, чтобы оценить дисперсию случайной ошибки. При этом выполняют замену переменных, используя оценку дисперсии вспомогательной модели.

Выполнив расчеты, следуя рекомендациям этого подхода, для рассматриваемого примера получены следующие остатки после возвращения к исходным переменным (рисунок 6, вверху). Как видно, размах остатков превышает 0.07, что оказалось неприемлемым для рассматриваемой задачи.

Выполним вычисления на основе взвешивания, используя оценку дисперсии непосредственно для исходной модели без привлечения вспомогательной модели. Результаты представлены на рисунке 6, внизу. В этом случае размах остатков остается в пределах 0.02, что оказалось допустимым для рассматриваемой задачи.

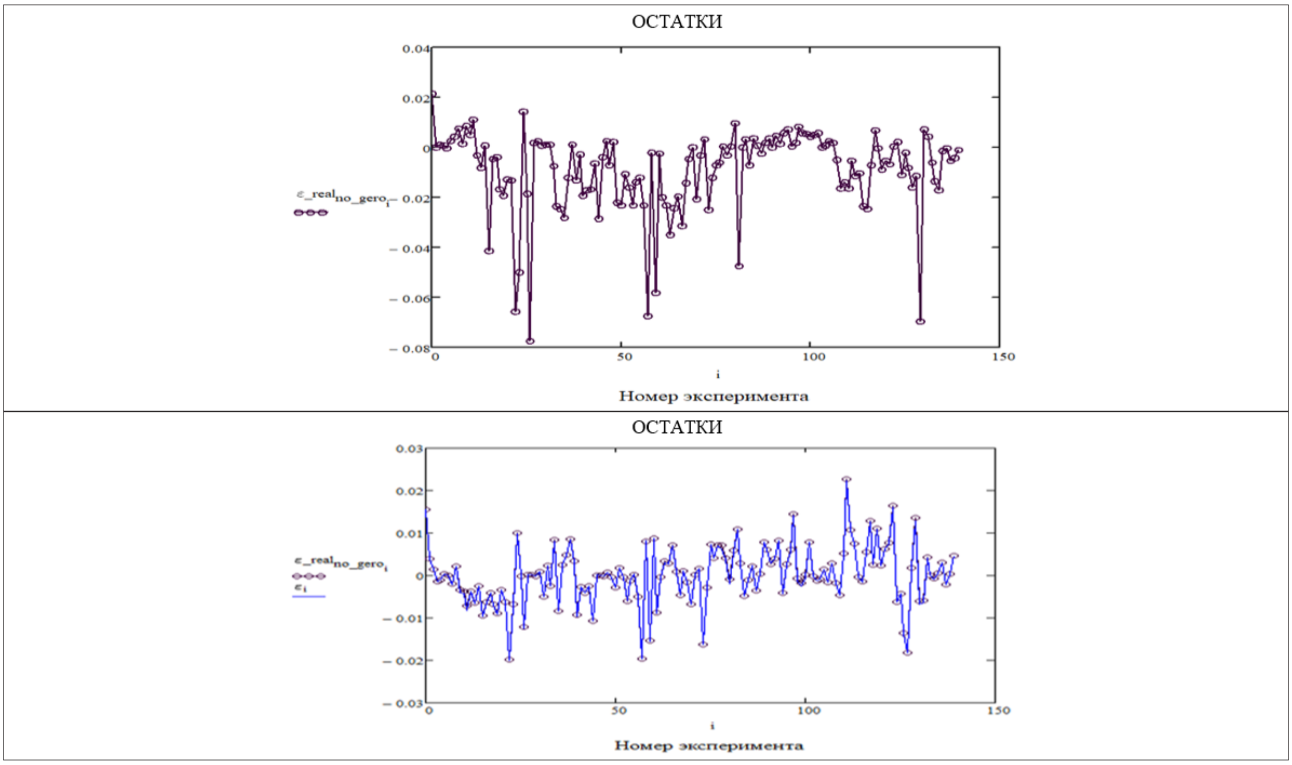


Рисунок 6. Результаты устранения гетероскедастичности модели на основе первого (вверху) и второго (внизу) выбора весов

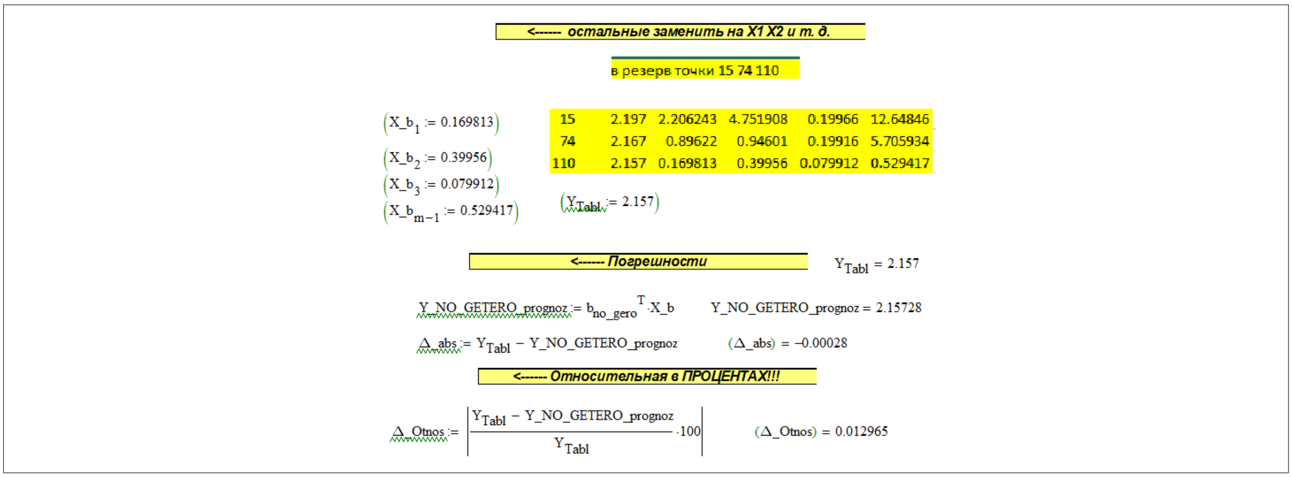


Рисунок 7. Результаты валидации модели с устраненной гетероскедастичностью

Фрагмент документа для валидации модели по трем точкам с номерами 15, 74, 110 с вычисленными погрешностями применительно к точке с номером 110 представлен на рисунке 7.

В точках с номерами 15 и 74 относительные погрешности в процентах составили соответственно 0,911977 и 0,090345, что приемлемо для рассматриваемой задачи.

Заключение

При реализации данного проекта получены следующие результаты:

- решение задач при моделировании сложных ячеистых конструкций становится проблематичным из-за продолжительности однократного вариантного расчета;
- предложен подход к решению этой проблемы на основе использования полученных ранее результатов вариантных вычислений для получения прогноза выходных параметров при тех сочетаниях значений параметров объекта, которые находятся в заданных ранее диапазонах по уже проведенным исследованиям;
- создан набор программ на основе использования средства MATHCAD для реализации такого подхода;
- с использованием созданного программного обеспечения построены соответствующие модели. В частности, выполнены тесты на гетероскедастичность и реализованы способы ее устранения;
- выполнены сравнения способов устранения гетероскедастичности при различном выборе весовых коэффициентов на основе использования взвешенного метода наименьших квадратов;
- представлены результаты валидации модели с устраненной гетероскедастичностью.

ЛИТЕРАТУРА

1. Построение геометрической части конечно-элементной модели одного вида пористых структур / В. В. Напрасников, Ю. В. Полозков, А. В. Бородуля, Д. П. Кункевич // Системный анализ и прикладная информатика. 2019. № 4. С. 55–61. DOI: 10.21122/2309-4923-2019-4-55-61.
2. Полозков, Ю. В. Проблемы проектирования и формообразования легковесных деталей в аддитивном производстве / Ю. В. Полозков // Математические методы в технике и технологиях: сб. тр. междунар. науч. конф. : в 12 т. / под общ. ред. А. А. Большакова. СПб., 2017. Т. 10. С. 61–65.
3. APDL-моделирование ячеистых конструктивных элементов деталей для аддитивного формообразования / А. В. Бородуля, Д. П. Кункевич, В. В. Напрасников, Ю. В. Полозков // Аддитивные технологии, материалы и конструкции : материалы научно-технической конференции, Гродно, 5-6 октября 2016 г. / редкол.: А. И. Свириденок (гл. ред.) [и др.]. Гродно, 2016. С. 146–152.

REFERENCES

1. Naprasnikov V.V., Polozkov Y.V., Borodulya A.V., Kunkevich D.P. Building a geometric part finite element model of one kind of porous structures. System analysis and applied information science. 2019;(4):55–61 (in Russian). https://doi.org/10.21122/2309-4923-2019-4-55-61.
2. Polozkov Y.V. Problems of design and forming of lightweight details in additive manufacturing. Matematicheskie metody v tehnike i tehnologijah: Sbornik trudov mezhdunarodnoj nauchnoj konferencii. T. 10 [Mathematical methods in engineering and technology: Proceedings of the international scientific conference. Vol. 10]. St. Petersburg;

2017. pp. 61–65 (in Russian).

3. Borodulya A.V., Kunkevich D.P., Naprasnikov V.V., Polozkov Y.V. APDL-modeling of the most constructive elements of the details for additive form. Additivnye tehnologii, materialy i konstrukcii: materialy nauchno-tehnicheskoj konferencii, Grodno, 5-6 oktyabrya 2016 g. [Additive technology, materials and design: Proceedings of the scientific and technical conference, Grodno, October 5-6, 2016]. Grodno; 2016. pp. 146–152 (in Russian).

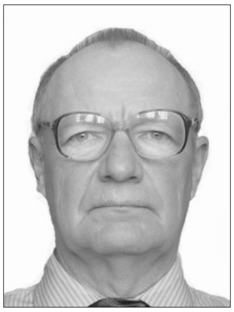
V. V. NAPRASNIKOV [1], Y. V. NAPRASNIKOVA [1], Y. V. POLOZKOV [2], A. V. BORODULYA [1], D. P. KUNKEVICH [1]

DETERMINATION AND ELIMINATION OF HETEROSCEDASTICITY OF A MULTIVARIATE REGRESSION MODEL FOR PREDICTING THE PARAMETERS OF HONEYCOMB STRUCTURES USING MATHCAD

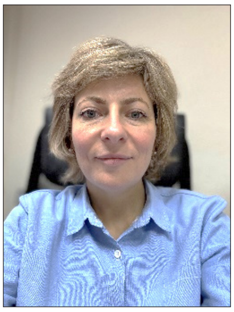
[1] Belarusian National Technical University
Minsk, Republic of Belarus
[2] Vitebsk State Technological University
Vitebsk, Republic of Belarus

Abstract. *Described is a possible approach to solving the problem of obtaining a forecast of output parameters of cellular structures with those combinations of values of parameters of the object which are in previously given ranges based on studies already carried out, based on use of previously obtained results of variant calculations. Provides portions of a suite of programs that use MATHCAD to implement this approach. Comparisons of methods for eliminating heteroskedasticity with different selection of weighting factors based on the use of the weighted least squares method were made.*

Keywords: *cellular structures, weighted least squares, heteroskedasticity*



Напрасников Владимир Владимирович
Белорусский национальный технический университет (БНТУ), г. Минск, Республика Беларусь.
Кандидат технических наук, доцент кафедры «Программное обеспечения информационных систем и технологий» БНТУ. Научные интересы – конечно-элементное моделирование, компьютерные средства инженерного анализа технических систем.
Vladimir V. Naprasnikov
Belarusian National Technical University, Minsk, Republic of Belarus.
PhD in Engineering, Associate Professor of the Department of Software of Information Systems and Technologies. His research interest focus on finit-element computer aided engineering.
E-mail: vnaprasnikov@bntu.by



Напрасникова Юлиана Владимировна
Магистр технических наук, ведущий инженер УП «Калийпроект». Научные интересы – конечно-элементное моделирование, компьютерные средства инженерного анализа технических систем.
Juliana V. Naprasnikova
Master of Engineering Sciences, Leading Engineer at unitary enterprise «Kaliyproekt». Research interest focus on finit-element computer aided engineering.
E-mail: vnaprasnikov@bntu.by

Полозков Юрий Владимирович
Витебский государственный технологический университет (ВГТУ), г. Витебск, Республика Беларусь.
Доцент, кандидат технических наук, проректор по научной работе ВГТУ. Научные интересы – методы и алгоритмы автоматизации проектирования технических объектов и производственных процессов; информационные, обучающие, тестирующие методы и программные средства в образовании.
Yury V. Polozkov
Vitebsk State Technological University (VSTU), Vitebsk, Republic of Belarus.
Associate Professor, PhD in Engineering, Vice-Rector for Science at VSTU. Research interests focus on methods and algorithms for automating the design of technical objects and production processes; information, training, testing methods, and software in education.
E-mail: vstu@vstu.by

Кункевич Дмитрий Петрович
Белорусский национальный технический университет (БНТУ), г. Минск, Республика Беларусь.
Кандидат технических наук, доцент кафедры «Программное обеспечения информационных систем и технологий» БНТУ. Научные интересы – автоматизация конструкторско-технологического проектирования и инженерного анализа механических систем.
Dmitry P. Kunkevich
Belarusian National Technical University, Minsk, Republic of Belarus.
PhD in Engineering, Associate Professor of the Department of Software of Information Systems and Technologies. His research interest focus on computer aided design and engineering of mechanical systems.
E-mail: povt@bntu.by

Бородуля Алексей Валентинович
Белорусский национальный технический университет (БНТУ), г. Минск, Республика Беларусь.
Кандидат технических наук, доцент кафедры «Программное обеспечения информационных систем и технологий» БНТУ. Научные интересы – CALS-технологии.
Aleksei V. Borodulya
Belarusian National Technical University, Minsk, Republic of Belarus.
PhD in Engineering, Associate Professor of the Department of Software of Information Systems and Technologies. His research interest focus on CALS-technologies**.**
E-mail: povt@bntu.by