

УДК 006.91:53.089.68

В.С. Єременко, В.М. Мокійчук, О.В. Самойліченко

Національний авіаційний університет, Київ, Україна

ДОСТОВІРНІСТЬ ОЦІНЮВАННЯ СКЛАДОВИХ НЕВИЗНАЧЕНОСТІ СТАНДАРТНИХ ЗРАЗКІВ, ОБУМОВЛЕНИХ НЕОДНОРІДНІСТЮ ТА НЕСТАБІЛЬНІСТЮ

Розглянуто метод визначення достовірності оцінювання складових невизначеності стандартних зразків, обумовлених неоднорідністю та нестабільністю досліджуваних характеристик речовини або матеріалу. Метод базується на використанні оцінок потужності статистичних критеріїв, що використовуються при оцінюванні вибіркового даних в залежності від їх статистичних параметрів та обсягу. Метод дозволяє підвищити точність оцінювання характеристик неоднорідності та нестабільності стандартних зразків та сумарної стандартної невизначеності атестованого значення складу або властивості

Ключові слова: однорідність, стабільність, статистичний критерій, статистична оцінка потужності, невизначеність.

Вступ

При атестації стандартних зразків (СЗ) встановлюють їх основні метрологічні характеристики: атестоване значення СЗ та його невизначеність. Серед складових невизначеності важливими є складові, обумовлені неоднорідністю та нестабільністю характеристик досліджуваного матеріалу (речовини). Рекомендації щодо оцінювання цих складових наведені в нормативних документах, які регламентують процедуру атестації стандартних зразків, зокрема в [1 – 3].

Основний матеріал

Відповідно до [4] під однорідністю СЗ розуміють таку його властивість, яка визначається постійністю значення характеристики зразку, що атестується, а також будь-якої його частини. Характеристика однорідності перевіряється на декількох етапах розробки СЗ. За вимогами нормативних документів необхідно проводити перевірку гомогенності маси матеріалу для виготовлення СЗ, однорідності всередині зразка, та між упакованими зразками перед атестацією партії зразків та розсилкою СЗ у лабораторії. Складова невизначеності від неоднорідності обумовлена відхиленням дійсних значень атестованої характеристики СЗ в довільній найменшій представницькій пробі. Характеристику невизначеності від неоднорідності виражають як середньоквадратичне відхилення (СКВ) з зазначенням найменшої представницької проби. Процедура оцінювання однорідності відрізняється для дисперсних та монолітних матеріалів. Характеристику однорідності СЗ складу дисперсних матеріалів оцінюють способом, заснованим на багатократних вимірювань складу компонента, що атестується в декількох пробах, відібраних випадковим чином з наступним опрацюванням отриманих даних за схемою однофакторного дисперсійного аналізу. Характеристику

однорідності СЗ складу монолітних матеріалів оцінюють методом, заснованим на багаторазових вимірюваннях складу компоненту, що атестується в декількох екземплярах СЗ, відібраних випадковим чином з наступною обробкою отриманих результатів за схемою двофакторного дисперсійного аналізу. Характеристику однорідності зазвичай оцінюють для всіх компонентів, що атестуються. Однак в певних випадках можливе оцінювання характеристики однорідності за компонентами-індикаторами. Детально процедура статистичного опрацювання отриманих результатів для отримання складової невизначеності від неоднорідності для кожного випадку описана в [2].

Висновок про однорідність (нуль-гіпотеза: досліджувані дані однорідні відносно дисперсій) СЗ приймається з використанням статистичного критерію Фішера. Статистику Фішера, розраховану як відношення факторної дисперсії до дисперсії від впливу неврахованих факторів порівнюють з табличним, взятим для відповідних степенів вільності, розрахованих за кількістю градацій фактора та обсягом даних.

Необхідною умовою застосування процедури дисперсійного аналізу є перевірка однорідності дисперсій вибірок, яка здійснюється за допомогою статистичного критерію Кохрена.

Під стабільністю досліджуваних зразків розуміють таку його властивість, яка виявляється у незмінності значень атестованої характеристики протягом терміну придатності СЗ [1]. Факторами нестабільності виступають сукупність зовнішніх умов, фізичних та хімічних процесів, що протікають в матеріалі СЗ які можуть викликати зміну властивості СЗ, що атестується. Характеристика нестабільності СЗ виражається у вигляді СКВ коефіцієнту a_1 регресійної залежності $y = a_1x + a_0$, побудованої за даними перевірки стабільності СЗ за період дослідження. Дослідження стабільності проводять з метою встановлення терміну придатності СЗ, умов зберігання і застосування, за яких можливі

зміни значень атестованої характеристики СЗ знаходяться в певних межах. Розрізняють короткотривалу стабільність – стабільність властивості СЗ під час транспортування при певних умовах перевезення та довготривалу стабільність – стабільність властивості СЗ при певних умовах зберігання у постачальника або користувача СЗ. Встановлення характеристики стабільності відбувається за допомогою статистичного опрацювання результатів досліджень, порядок проведення якого наведено в [5, 6].

Висновок про короткотривалу стабільність (нуль-гіпотеза: досліджувані вибірки однорідні відносно центрів розподілу) приймається з використанням критерію Стьюдента. Додатково перевіряють статистичну однорідність досліджуваних даних відносно дисперсій з використанням статистичного критерію Фішера. Висновок про довготривалу стабільність (нуль-гіпотеза: коефіцієнт a_1 статистично значимий) приймається також з використанням критерію Стьюдента. Статистична значимість коефіцієнту буде свідчити про нестабільність зразку, тобто його показники будуть змінюватися в часі. Альтернативним методом оцінювання стабільності є застосування критерію Грабса перевірки однорідності декількох вибірок відносно математичних сподівань. Однорідність математичних сподівань буде свідчити про відсутність змін умовно дійсного значення показника в часі.

Отже, рішення про однорідність та стабільність досліджуваних зразків ґрунтується на застосуванні зазначених статистичних критеріїв, потужність яких залежить від обсягів досліджуваних даних. В роботах [7, 8] проведено дослідження потужності параметричних критеріїв в умовах, коли обсяг даних знаходиться в межах від 5 до 50 значень. У випадках, коли отримання великого обсягу досліджуваних даних можливе, статистичні критерії забезпечують достатню потужність (до 99%), тому їх використання не впливає на достовірності отримуваних даних. Однак, в тих галузях, де отримання великих обсягів вибірок неможливе, постає питання щодо потужності застосовуваних статистичних критеріїв в зазначених умовах.

Результати дослідження робастних та непараметричних критеріїв перевірки однорідності вибірок, які

виступають в якості альтернативи параметричним критеріям, приведені в [9] дозволяють зробити висновок, що робастні та непараметричні критерії мають значно складнішу процедуру обрахунку, однак, на обмежених обсягах (до 50 значень) не забезпечують кращої потужності, тому для вирішення поставлених задач їх використання не доцільне.

В табл. 1. представлено результати випробування СЗ зерна, при проведенні їх атестації, за показником «сира клейковина» отриманих відповідно методу, описаному в ДСТУ ISO 5531:2004. Результати взяті в Державному центрі сертифікації і експертизи зерна та продуктів його переробки, м. Київ. Для встановлення гомогенності матеріалу досліджувалися 10 взятих з матеріалу зразків, для кожного було отримано по 6 результатів випробувань.

Таблиця 1

Результати випробування СЗ зерна

№ зразка	Сира клейковина, %					
	1	2	3	4	5	6
1	22,5	22,7	23,2	22,6	23,0	23,3
2	22,5	23,4	23,0	22,8	23,1	23,2
3	22,8	22,8	23,1	22,6	22,6	23,0
4	22,5	22,5	22,6	22,7	22,4	22,5
5	22,6	22,8	22,8	23,0	23,0	22,9
6	23,0	22,6	23,0	22,7	22,6	23,2
7	22,4	22,7	23,1	23,1	23,0	23,0
8	22,8	23,4	22,7	22,7	22,8	22,7
9	22,5	22,7	22,9	22,6	22,6	23,0
10	22,6	23,1	22,6	23,0	23,0	22,5

Четвертий зразок має показники, що дещо відрізняються від решти, що спричиняє підозру в його неоднорідності відносно решти. Розрахований відсоток неоднорідності склав 21,3%, однорідності – 78,7%. Подальше застосування критерію Фішера показало, що матеріал гомогенний.

В рамках експериментального дослідження було проведено чотири додаткові дослідження (табл. 2). Розрахований відсоток неоднорідності досліджуваного матеріалу склав 30,1%, однорідності – 69,9%. Подальше застосування критерію Фішера зі збільшеним обсягом досліджуваних даних показав, що матеріал не гомогенний.

Таблиця 2

Результати чотирьох додаткових досліджень

№ зразка	Сира клейковина, %									
	1	2	3	4	5	6	7	8	9	10
1	22,5	22,7	23,2	22,6	23,0	23,3	23,2	23,6	23,0	23,1
2	22,5	23,4	23,0	22,8	23,1	23,2	23,0	22,8	23,1	23,2
3	22,8	22,8	23,1	22,6	22,6	23,0	23,1	22,6	22,6	23,0
4	22,5	22,5	22,6	22,7	22,4	22,5	22,6	22,4	22,4	22,5
5	22,6	22,8	22,8	23,0	23,0	22,9	22,8	23,0	23,0	22,9
6	23,0	22,6	23,0	22,7	22,6	23,2	23,0	22,7	22,6	23,2
7	22,4	22,7	23,1	23,1	23,0	23,0	23,1	23,1	23,0	23,0
8	22,8	23,4	22,7	22,7	22,8	22,7	22,7	22,7	22,8	22,7
9	22,5	22,7	22,9	22,6	22,6	23,0	22,9	22,6	22,6	23,0
10	22,6	23,1	22,6	23,0	23,0	22,5	22,6	23,1	23,0	23,0

Отже, матеріал для виготовлення зразків не гомогенний і потребує додаткової гомогенізації перед формуванням СЗ. Використання такого не гомогенного матеріалу може призвести до формування неоднорідних зразків. Якщо неоднорідність сформованих таким чином зразків у подальшій діяльності не буде встановлена, то це призведе до негативних результатів в раундах міжлабораторних порівнянь і можливо, до не атестації лабораторії.

Іншим прикладом є представлені в табл. 3 результати дослідження стандартних зразків зерна за показником «Масова частка білку», метод визначення білку викладений в ГОСТ 10846-91. Дослідження проводилися з метою встановлення короткотривалої стабільності в умовах транспортування.

Таблиця 3
Результати дослідження стандартних зразків зерна

Масова частка білку, %		
№	Перед транспортуванням	Після транспортуванням
1	13,01	12,98
2	12,94	13,10
3	13,04	13,07
4	13,05	13,07
5	12,98	13,05
6	13,05	12,96
7	13,02	13,11
8	13,07	13,09
9	12,98	13,00
10	12,99	13,03
Середнє арифметичне	13,013	13,046
Дисперсія	0,0016	0,0027

Згідно затверджені процедури дослідження необхідно перевірити вибірки на збіжність за допомогою F-критерію Фішера, при позитивному результаті проводиться перевірка на збіжність середніх за допомогою t-критерію Стьюдента, у протилежному випадку результати визнаються такими, що відрізняються. У випадку позитивних результатів обох перевірок, результати визнаються такими, що не відрізняються.

За результатами перевірки було встановлено, що надані до перевірки дані збіжні, розраховане значення статистики Фішера $F = 1,64$, граничне (для довірчої ймовірності 0,95) $F_{гр} = 2,98$. Результати перевірки на стабільність згідно критерію Стьюдента показали, що матеріал стабільний, розраховане значення статистики Стьюдента $t = 1,58$; граничне (для довірчої ймовірності 0,95) $t_{гр} = 2,10$.

Проведення додаткових п'яти випробувань (табл. 4) підтвердило збіжність, однак показало, що матеріал не стабільний.

Розраховане значення статистики Фішера $F = 1,81$, граничне, для довірчої ймовірності 0,95 $F_{гр} = 2,40$. Розраховане значення статистики Стьюдента $t = 2,90$; граничне (для довірчої ймовірності 0,95) $t_{гр} = 2,05$.

Таблиця 4
Результати додаткових п'яти випробувань

Масова частка білку, %		
№	Перед транспортуванням	Після транспортуванням
1	13,01	12,98
2	12,94	13,10
3	13,04	13,07
4	13,05	13,07
5	12,98	13,05
6	13,05	12,96
7	13,02	13,11
8	13,07	13,09
9	12,98	13,00
10	12,99	13,03
11	13,01	13,07
12	13,02	13,11
13	13,03	13,09
14	12,98	13,05
15	12,99	13,03
Середнє арифметичне	13,011	13,054
Дисперсія	0,0018	0,0022

Отже, в зв'язку з недостатньою потужністю критеріїв на вибірках малого обсягу, що використовувалися при встановленні однорідності та стабільності виникає ймовірність зробити помилку другого роду (прийняти нуль-гіпотезу, коли вона не вірна). При цьому буде помилково використана відповідна метрологічна характеристика, що відобразиться на точності оцінювання розширеної невизначеності. Так, СКВ, розраховане за неоднорідним або нестабільними матеріалом завжди буде завищеним. Використання такого матеріалу в подальшому може призвести до негативних наслідків в раундах міжлабораторних порівнянь, зменшить шанси лабораторії успішно пройти атестацію. Для підвищення точності оцінювання складових невизначеності, обумовлених неоднорідністю та нестабільністю запропоновано використовувати метод з використанням статистичної оцінки потужності (СОП) критеріїв [10]. СОП представлена як функція $f = f(P |_{\alpha, N, a_0 \dots a_{n-1}})$, яка описує співвідношення потужності критерію, обсягу вибірки N і значення статистичних параметрів вибірки $a_0 \dots a_{n-1}$ при заданому рівні значимості α . На основі СОП для критеріїв Фішера, Стьюдента, Кохрена та Грабса було знайдено відповідне апроксимуюче рівняння, оцінено його коефіцієнти.

При оцінюванні складових невизначеності від неоднорідності та нестабільності слід розрахувати потужність статистики критерію, що використовується в умовах отримання відповідної характеристики. Якщо потужність критерію виявиться нижчою за рівень довіри, з яким представлятимуть результат вимірювання та його невизначеність, то використання таких метрологіч-

них характеристик, як СКВ від неоднорідності та нестабільності є недоцільним. Відповідно до СОП можна розрахувати мінімально необхідний обсяг досліджуваних даних для забезпечення заданої достовірності. У випадку, коли немає можливості провести додаткові випробування, цю характеристику можливо використовувати лише для випадку, якщо поставити у відповідність довірчу ймовірність і відповідне значення потужності критерію.

Висновки

Оцінювання метрологічних характеристик однорідності та стабільності є невід'ємною частиною процесу атестації СЗ. Окрім свого внеску як складові розширеної невизначеності вони дозволяють перевірити необхідні умови подальшого використання СЗ. Оцінювання відповідних характеристик базується на статистичній обробці даних, і, за обмежених обсягів, завжди виникає ймовірність зробити помилку другого роду при застосуванні відповідного статистичного критерію. При цьому знижується точність оцінювання розширеної невизначеності, погіршується єдність результатів, отримуваних з використанням неоднорідного чи нестабільного матеріалу.

Запропоновано використовувати метод, який дозволяє встановлювати достовірність отримуваних результатів оцінювання розглядуваних метрологічних характеристик.

Список літератури

1. Метрологія. Стандартні зразки складу речовин та матеріалів. Міжлабораторна метрологічна атестація. Зміст і порядок проведення робіт (ГОСТ 8.532-2002, IDT): ДСТУ ГОСТ 8.532-2003. – [Чинний від 13-05-03]. – К.: Даржстандарт, 2003. – 15 с. – (Національний стандарт України).
2. Государственная система обеспечения единства измерений. Стандартные образцы состава монолитных и дисперсных материалов. Способы оценивание однородности: ГОСТ 8.531-2002. – [Действителен от 01.07.2003]. – М.: Издательство стандартов, 2002. – 13 с. – (Межгосударственный стандарт).
3. Государственная система обеспечения единства измерений. Рекомендации по метрологии. Стандартные образцы состава и свойств веществ и материалов. Методика оценивание характеристики стабильности: Р 50.2.031-2003. – Офиц. изд. – М.: Стандартиформ, 2004. – 10 с. – (ФГУП "ВНИИМ". Рекомендации).
4. Государственная система обеспечения единства измерений. Стандартные образцы состава и свойств веществ и материалов. Основные положения: ГОСТ 8.315-97. – [Действителен от 1997-07-01]. – М.: ИПК Издательство стандартов, 2004. – 20 с. – (Межгос. стандарт).
5. Атестація стандартних зразків. Загальні статистичні принципи (ISO GUIDE 35:1989, IDT): ДСТУ-Н ISO GUIDE 35:2006. – [Чинний від 2007-10-01]. – К.: Даржстандарт, 2006. – 64 с. – (Нац. стандарт України).
6. Государственная система обеспечения единства измерений. Рекомендации по метрологии. Стандартные образцы состава и свойств веществ и материалов. Методика оценивание характеристики стабильности: Р 50.2.031-2003. – Офиц. изд. – М.: Стандартиформ, 2004. – 10 с. – (ФГУП "ВНИИМ". Рекомендации).
7. Еременко В.С. Особенности использования статистических критериев при обработке результатов аналитических измерений / В.С. Еременко, В.М. Мокійчук, О.В. Самойличенко // Качество и безопасность. Вопросы компетентности и профессионального тестирования персонала испытательной лаборатории: труды ИПК, 26-28 мая 2008 г.: тезисы докл. – Одесса: Одесский национальный университет им. Мечникова, 2008. – С. 26-31.
8. Еременко В.С. Исследование мощности критерия Кохрена при ограниченном числе наблюдений / В.С. Еременко, В.М. Мокійчук, О.В. Самойличенко // Системи обробки інформації. – Х.: ХУПС, 2007. – Вип. 6 (64). – С. 35-39.
9. Еременко В.С. Аналіз використання непараметричних критеріїв згоди при опрацюванні результатів вимірювань / В.С. Еременко, В.М. Мокійчук, О.В. Самойличенко // Метрологія та вимірювальна техніка: наук. праці конф., тези доп. – Х.: ХФ ДП «УкрНДНЦ», 2008. – Т. 1. – С. 40-44.
10. Еременко В.С. Определение статистических оценок мощности критериев однородности на выборках малого объема / В.С. Еременко, В.М. Мокійчук, О.В. Самойличенко // Системи обробки інформації. – Х.: ХУПС, 2008. – Вип. 4 (71). – С. 30-33.

Надійшла до редколегії 20.12.2010

Рецензент: д-р техн. наук, проф. Л.М. Щербак, Національний авіаційний університет, Київ.

ДОСТОВЕРНОСТЬ ОЦЕНИВАНИЯ СОСТАВЛЯЮЩИХ НЕОПРЕДЕЛЕННОСТИ СТАНДАРТНЫХ ОБРАЗЦОВ, ОБУСЛОВЛЕННЫХ НЕОДНОРОДНОСТЬЮ И НЕСТАБИЛЬНОСТЬЮ

В.С. Еременко, В.М. Мокійчук, О.В. Самойличенко

Предложен метод определения достоверности оценки составляющих неопределенности стандартных образцов, обусловленных неоднородностью и нестабильностью исследуемых характеристик веществ или материалов. Метод базируется на использовании оценок мощности статистических критериев, которые используются при оценивании выборочных данных в зависимости от их статистических параметров и объема. Метод позволяет увеличить точность оценивания характеристик неоднородности и нестабильности стандартных образцов и суммарной стандартной неопределенности аттестованного значения состава или свойства.

Ключевые слова: однородность, статистический критерий, статистическая оценка мощности, неопределенность.

RELIABILITY OF REFERENCE STANDARD UNCERTAINTY COMPONENTS ARISE FROM HETEROGENEITY AND INSTABILITY

V.S. Yeremenko, V.M. Mokyichuk, O.V. Samoylichenko

The method to determine the reliability of reference material uncertainty components estimations caused of heterogeneity and instability of substance or material tested properties has given. The method is based on using the functions of power for statistical criterion, used for sample data estimations against the statistical parameters and sample size. Method makes it possible to increase the accuracy of the evaluation of heterogeneity and instability characteristics of reference material and combined standard uncertainty of certified value of substance or material.

Keywords: homogeneity, stability, statistical criteria, function of power, uncertainty.