

O wyrażaniu i publikowaniu danych pomiarów wieloparametrowych – stan aktualny a potrzeby

Zygmunt Lech Warszawa¹, Vladimir V. Ezhela²

¹Przemysłowy Instytut Automatyki i Pomiarów PIAP

²Instytut Fizyki Wielkich Energii (IHEP), RosAtom, Protvino MR, Rosja

Streszczenie: W artykule omówiono problemy związane z wyznaczaniem, prezentacją i publikacją danych pomiarów wieloparametrowych w nawiązaniu do publikacji szczegółowych z PAR nr 2 i 6 z 2011 r. Przybliżono zasady opisu wyników pomiarów wielkości powiązanych statystycznie za pomocą algebry wektorów losowych oraz problemy występujące przy ich zaokrągleniu. Omówiono specyfikę propagacji niepewności w pomiarach jedno- i wieloparametrowych o funkcjach nieliniowych oraz wyznaczono liczby członów rozwinięcia w szereg wymagane przy dokładnym wyznaczaniu niepewności pomiarów pośrednich. Wykazano, że parametry czterech podstawowych stałych fizycznych podane w ostatnich korektach CODATA z 2010 r. nie są w pełni poprawne, gdyż ich macierze korelacji mają ujemne wartości własne. Zalecenia dla wyrażania wyników pomiarów wieloparametrowych mając przygotowywany Suplement 2 do przewodnika GUM-2008. Proponuje się, by jego niezbędnym uzupełnieniem były zasady standaryzacji publikowania danych pomiarowych w dualnej postaci: na papierze i towarzyszącej jej e-publikacji zawierającej wyniki oryginalne. Omówiono sposób gromadzenia i rozpowszechniania danych stosowany przez wiodące centra światowe.

Słowa kluczowe: wieloparametrowe dane pomiarowe, niepewność multimezurandu, stałe podstawowe, e-publikacje

1. Wprowadzenie

Wyniki jednoczesnych pomiarów wieloparametrowych, czyli dane multimezurandu, zbiera się, przetwarza i wykorzystuje w wielu dyscyplinach naukowych opartych na eksperymencie oraz w technice [19] i innych dziedzinach gospodarczych. Komputerową obróbkę danych pomiarowych rozpoczęto ok. pięćdziesiąt lat temu – mimo to wciąż brak ustalonych wspólnych metod i zasad jak wyrażać i przekazywać estymaty parametrów wielu jednocześnie mierzonych wielkości wraz z ich niepewnościami i współczynnikami korelacji. Mimo że przetwarzanie, przechowywanie i rozpowszechnianie takich pomiarów stało się powszechne, to specyfika opisu wyników pomiarów za pomocą algebry wektorów losowych nie jest szerzej znana, a numeryczne wyrażanie wielowymiarowych danych pomiarowych nie jest ujednolicone. Opracowane dotychczas przepisy metrologiczne, np. przewodnik GUM [1, 4] i poradnik NASA [14], nie nadążają za szybkim rozwojem nauki, technologii i eksploatacji. Liczni eksperymetatorzy, a nawet i metrologzy ignorują specyfikę pomiarów wieloparametrowych i w opisie wyników badań stosują ogólne

zalecenia GUM [4] opracowane dla pomiarów pojedynczych wielkości. Potrzebny od dawna Suplement 2 do GUM o wyrażaniu niepewności pomiarów wieloparametrowych nie ukazał się do dziś, chociaż wersja robocza (*draft*) jest już dostępna w Internecie (nie obejmuje całości zagadnień związanych z tymi pomiarami). Należy zapewnić wysoką jakość i pełną wiarygodność metrologiczną wyrażania i przekazywania parametrów statystycznych wyników powiązanych ze sobą, równocześnie mierzonych wielkości. Tymczasem w publikacjach i dostępnych w sieci zasobach podawanie danych pomiarowych i danych wielkości wyznaczanych pośrednio na podstawie wyników tych pomiarów bywa dalekie od tzw. dobrej praktyki. Wiele negatywnych przykładów, zaczerpniętych z publikacji w czołowych recenzowanych z wysokim *impact factor* czasopismach naukowych, opatrzone komentarzami [9]. Analiza ta pokazuje, że dotychczasowa praktyka tworzenia, przekazywania i przechowywania informacji nie gwarantuje odpowiedniej naukowej i metrologicznej jakości danych eksperymentalnych. Stosowane metody selekcji i ekstrakcji danych pomiarowych z publikacji do zastosowań w zadaniach naukowych i technicznych oraz ogólniej – gospodarczych, nie zapobiegają rozprzestrzenianiu niepełnych lub nieprawidłowych danych, nawet gdy są przekazywane drogą elektroniczną. Międzynarodowa organizacja CODATA (*Committee on Data for Science and Technology*) w danych podstawowych stałych fizycznych FPC (*Fundamental Physical Constants*) opublikowanych w korektach z lat 1998–2010 [5, 6, 9, 13] i traktowanych powszechnie jako podstawowe dane odniesienia nie uniknęła błędów. Z doświadczeń wynika, że konieczne jest sformalizowanie procedur oraz metod przesiewu i akceptacji danych, zwłaszcza do zastosowań o wysokiej precyzji, odpowiedzialności i dużym ryzyku.

Zadania wymagające realizacji to:

- popularyzacja podstaw teoretycznych oraz specyfiki przetwarzania danych pomiarów wieloparametrowych i wyrażania ich wyników w kręgach nie tylko metrologicznych, ale i wśród użytkowników takich pomiarów, co zagwarantuje stosowanie zaleceń Suplementu 2 do GUM po jego ukazaniu się; Suplement ma ujednolicić sposób opisu danych z pomiarów i wielkości pochodnych wyznaczanych pośrednio na ich podstawie;
- opracowanie zaleceń uzupełniających Suplement 2 oraz inne normy i przewodniki, niezbędnych do przetwarzania, zaokrąglenia, raportowania, publikowania dobrej jakości danych multimezurandu;
- analiza zasadności i wdrożenie do opisu pomiarów multimezurandu najnowszych narzędzi matematycznych;
- zmiana praktyki publikowania danych pomiarowych, z tradycyjnej formy prezentacji danych na papierze, na

pełne wykorzystanie możliwości e-publikacji [9] wraz z dostępnością numerycznej weryfikacji wyników przez użytkowników oraz opracowanie odpowiednich przepisów międzynarodowych.

Tematykę dwu pierwszych zadań omówiono w dwuczęściowej pracy (PAR nr 2 i 6 2011 r.). W części 1 [11] podano podstawy teoretyczne opisu przetwarzania i wyrażania wyników pomiaru multimezurandu z zastosowaniem wielowymiarowego rozkładu normalnego i algebry wektorów losowych oraz typowe przykłady uszkalowania danych przy ich przetwarzaniu przez pomijanie wpływu korelacji i nadmierne zaokrąglanie wyników. W części 2 [12] zanalizowano nieliczne fragmenty ostatniej wersji przewodnika GUM [4], w tym przykład H.2, które wiążą się z pomiarami wieloparametrowymi, a także zaproponowano kilka korekt. Podano też zależności dla progów zaokrąglania (minimalne liczby cyfr) wyników takich pomiarów, tj. ich wartości, niepewności i współczynników korelacji przy numerycznym przetwarzaniu wektora wielkości losowych o parametrach statystycznych traktowanych jako znane absolutnie dokładnie. Przedstawiono dwie propozycje zaokrąglania poniżej tych progów dla danych uzyskanych z próbek obciążonych niepewnościami statystycznymi wskutek małej liczby elementów.

Potrzeby pomiarów wieloparametrowych i stosowania ich danych można prześledzić [9, 18] na przykładzie fizyki jako najbardziej reprezentatywnej z nauk eksperymentalnych oraz przemysłu, edukacji i metrologii – rozumianej jako wspomaganą badaniami działalność służby miar i różnego typu upoważnionych laboratoriów pomiarowych. To te cztery rodzaje działalności intelektualnej SEMI (ang. *Science, Education, Metrology, Industry*) głównie tworzą dane pomiarowe i z nich korzystają w działaniach bieżących i w rozwoju wiedzy. Potrzeby pomiarowe SEMI, choć nie są identyczne, ale są ściśle ze sobą powiązane. Ich rozwój nie jest jednak synchronizowany. Jest to jedna z głównych przyczyn powodujących trudności przy ustalaniu i utrzymywaniu jednolitości pomiarów i jakości (jednorodności) danych pomiarowych. Wspólnym zadaniem nauki i metrologii jest standaryzacja pomiarów wieloparametrowych, niezbędna dla spójnego rozwoju wszystkich dziedzin SEMI.

2. Opis losowy wielkości wektorowych i wyników ich pomiarów

Dowolną wielkość skalarną X jako mezurand opisuje zmienna losowa. Minimalna struktura danych wyrażająca wynik pomiarów zawiera estymator punktu skupienia rozkładu wartości obserwacji pomiarowych i towarzyszącą mu niepewność [1]. Jest to przedział występowania wartości tego estymatora o wymaganej ufności, czyli o określonym prawdopodobieństwie i wyraża się go jako wielokrotność odchylenia standardowego σ . Oba te parametry mezurandu są jednakowo zaokrąglane. Zwykle estymatorem wartości jest średnia \bar{x} . Dla próbek zawierających obserwacje pobierane z populacji o niektórych rozkładach niegaussowskich dokładniejsze są inne estymatory, np. środek rozstępu dla rozkładu równomiernego [15] i zbliżonych doń rozkładów trapezowych, dla których jeszcze dokładniejszy jest estymator dwuelementowy [16].

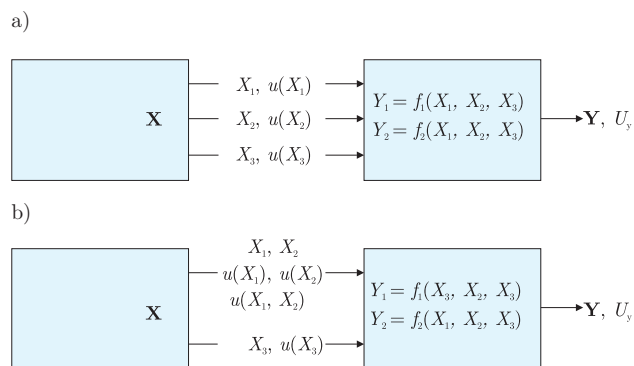
W pomiarach wieloparametrowych jako wyniki pomiarów multimezurandu należy wyznaczać nie tylko estymatory poszczególnych jego składowych wraz z niepewnościami, ale też i stopień wzajemnego ich powiązania statystycznego. Jeśli z pomiarów zbioru losowych wielkości X_i , – potraktowanych jako wejściowy wektor losowy \mathbf{X} , uzyskuje się pośrednio przez przetwarzanie danych wielkości badane (ang. *observables*) Y_j tworzące wyjściowy wektor losowy \mathbf{Y} , to taką operację wiążącą oba wektory [2, 18–20] można ogólnie zapisać jako

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}) \quad (1)$$

gdzie: \mathbf{X} – mezurand wejściowy jako wektor losowy o m elementach X_i ; \mathbf{Y} – mezurand wyjściowy jako wektor o n elementach Y_j ; $\mathbf{F}(\cdot)$ – operator, w ogólnym przypadku nieliniowy, wiążący oba te wektory.

Przy zależnościach liniowych operator \mathbf{F} staje się macierzą o wymiarach $n \times m$ i wówczas z pomiarów m wielkości wejściowych X_i otrzymać można co najwyżej $n \leq m$ wielkości wyjściowych Y_j . Przy nieliniowych zależnościach liczba wielkości wyjściowych n nie jest ograniczona przez m i jest równa liczbie niezależnych funkcji F_j wiążących składowe obu wektorów.

Na rys. 1a, b podano dwa przykłady przetwarzania wyników pomiarów pośrednich wieloparametrowych o wymiarach wektorów $m = 3$ i $n = 2$. W drugim z nich wielkości wejściowe X_1, X_2 są skorelowane.



Rys. 1. Przykłady pośredniego wyznaczania danych pomiarowych dwóch wielkości wyjściowych $\mathbf{Y} = [Y_1, Y_2]^T$ z pomiarów trzech wielkości wejściowych $\mathbf{X} = [X_1, X_2, X_3]^T$: a) nieskorelowanych, b) skorelowanych X_1, X_2

Fig. 1. Examples of indirect evaluation of measurement data of 2 jointed output variables $\mathbf{Y} = [Y_1, Y_2]^T$ from measurements of 3 input variables $\mathbf{X} = [X_1, X_2, X_3]^T$: a) no correlated, b) correlated X_1, X_2

Obserwacje w pomiarach wieloparametrowych są realizacjami wektora losowego. Gęstość prawdopodobieństwa występowania obserwacji podlega wielowymiarowemu rozkładowi w przestrzeni współrzędnych wektora. Kształt rozkładu może być różny, a jego geometria zależy od rozkładów składowych wektora i ich statystycznego powiązania. Przy przetwarzaniu danych pomiarowych wg nieliniowego operatora \mathbf{F} rozkład i jego przekroje mają inne kształty na wejściu i wyjściu [11].

Dla zadanego zakresu gęstości prawdopodobieństwa, czyli z zadanym prawdopodobieństwem rozrzutu końca wektora losowego \mathbf{X} o m elementach X_i występują w pewnym obszarze opisanym w przestrzeni m -wymiarowej. Takim obszarem jest np. hiperprostokąt o bokach równych odchyleniom stan-

dardowym składowych X_i wektora i zorientowanych równoległe do osi jego współrzędnych [11]. Hiperprostokąt otrzymuje się bezpośrednio z jednowymiarowych przedziałów rozrzutu składowych X_i traktowanych jako niezależne statystycznie, np. mierzone osobno i o równomiernych rozkładach. Koniec wektora średniego $\bar{\mathbf{X}}$ znajduje się w centrum tego hiperprostokąta.

Jeśli m wielkości X_i mierzonych jest równocześnie i rozrzut każdej z nich opisuje się rozkładem normalnym, to rozrzut końca wektora \mathbf{X} podlega m -wymiarowemu rozkładowi normalnemu. W wielu przypadkach w praktyce rozkład obserwacji z dopuszczalnym przybliżeniem modeluje się tym rozkładem. Jego opis jest bardziej złożony niż dla pojedynczej zmiennej losowej. Obok wartości współrzędnych X_i końca wektora średniego $\bar{\mathbf{X}}$ i ich odchyłeń standardowych σ_i występują też współczynniki korelacji ρ_{ij} . Skorelowanie jest nie do pominięcia, gdy wszystkie, lub kilka składowych wektora może być jeszcze potem wspólnie przetwarzanych numerycznie lub uczestniczyć wspólnie w innych eksperymentach. Jeśli niepewności składowych X_i średniego wektora $\bar{\mathbf{X}}$ różnią się liczbą cyfr po przecinku, to dąży się do ujednoczenia opisu w liczbach dziesiętnych. Obszar rozrzutu obserwacji o zadanym prawdopodobieństwie dla rozkładu m -normalnego ma kształt m -hiperelipsoidy wpisanej w m -hiperprostokąt. Bryły te są dobrze opisane analitycznie przez dodatnio określone formy kwadratowe [11, 17-19]. Dla m -wymiarowego rozkładu normalnego m -wymiarowa elipsoida opisana jest przez macierz kowariancji o wymiarach $m \times m$. Kąty nachylenia jej osi zależą od współczynników korelacji pomiędzy składowymi [11]. Jeśli obserwacje pomiarowych składowych wektora nie są skorelowane, to osie hiperelipsoidy są równoległe do osi współrzędnych wektora. W szczególnym przypadku dwuwymiarowego rozkładu normalnego ($m=2$) jego przekrojem jest elipsa, np. dla $k_{p1}=k_{p2}=1$ wpisana w prostokąt o bokach $\pm\sigma_1, \pm\sigma_2$ [11, 20].

Struktura opisująca dane multimezurandu wejściowego \mathbf{X} o rozkładzie normalnym zawiera wektor wartości średnich $\bar{\mathbf{X}}$ i macierz kowariancji \mathbf{c} . Macierz ta jest powiązana z macierzą korelacji \mathbf{r} (korelatorem) i z macierzą diagonalną odchyłeń standardowych $\boldsymbol{\sigma}$ [9, 11, 18–20] następującą zależnością

$$\mathbf{c} = \boldsymbol{\sigma} \cdot \mathbf{r} \cdot \boldsymbol{\sigma}^T \quad (1)$$

Elipsoidalny opis obszaru rozrzutu danych pomiarowych występuje dla dodatnio określonej macierzy kowariancji \mathbf{c} (i korelatora \mathbf{r}), tj. dla dodatnich wartości własnych λ_p tej macierzy, czyli jednokrotnych pierwiastków równania charakterystycznego

$$\det [\mathbf{c} - \lambda \mathbf{1}] = 0.$$

Wynik pomiaru losowego wektora wejściowego \mathbf{X} można wyznaczać w postaci dwuelementowej – jako wektor wartości średnich i macierz kowariancji, lub w postaci trójelementowej – jako wektor wartości średnich, wektor odchyłeń standardowych i macierz korelacji (korelator). Dla wektora \mathbf{X} o liczbie elementów $m=2$ opis będzie następujący:

$$\left(\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & \sigma_1 \sigma_2 \cdot r_{12} \\ \sigma_1 \sigma_2 \cdot r_{12} & \sigma_2^2 \end{bmatrix} \right) \Rightarrow \left(\begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \end{bmatrix}, \begin{bmatrix} \sigma_1 \\ \sigma_2 \end{bmatrix}, \begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix} \right) \quad (2)$$

Forma przedstawiania wyników wektora wyjściowego \mathbf{Y} jest identyczna jak wektora \mathbf{X} . Przy założeniu liniowej lub linearyzowanej propagacji wariancji macierze kowariancji \mathbf{c}_Y i \mathbf{c}_X wektorów wejściowego i wyjściowego są powiązane następującym równaniem liniowym:

$$\mathbf{c}_Y = \mathbf{S} \cdot \mathbf{c}_X \cdot \mathbf{S}^T \quad (3)$$

gdzie: macierz wrażliwości

$$\mathbf{S} \equiv \frac{\partial(\mathbf{Y})}{\partial(\mathbf{X})} \equiv \begin{bmatrix} \frac{\partial y_1}{\partial x_1}, & \dots & \frac{\partial y_1}{\partial x_m} \\ \dots & \dots & \dots \\ \frac{\partial y_m}{\partial x_1}, & \dots & \frac{\partial y_m}{\partial x_m} \end{bmatrix}$$

Przy przetwarzaniu wektora losowego \mathbf{X} wg nieliniowego operatora F geometria obszaru rozrzutu końca wektora \mathbf{Y} zmienia się zależnie od F . Dlatego należy śledzić zmiany granic tego obszaru i położenie wierzchołka wektora średniego $\bar{\mathbf{Y}}$. W przypadku małych nieliniowości lub małych wartości niepewności składowych wektora \mathbf{X} można przy przetwarzaniu, podobnie jak dla pojedynczej wielkości mierzonej, stosować w przybliżeniu liniową propagację wariancji wg (3) [18]. Przykłady obliczeń niepewności dla liniowego i nieliniowego przetwarzania danych multimezurandu podano w [11, 12].

Zestaw danych, który umożliwi poprawne oszacowanie parametrów wyniku pomiaru wyjściowej wielkości wektorowej i ich zaokrąglenie po przetwarzaniu, powinien być rozszerzony w stosunku do (2) i obejmować:

- wektor średni – wartości średnie jego składowych,
- wektor odchyłeń standardowych tych wartości i ich dokładność wyznaczaną statystycznie, która zależy od liczby pomiarów w próbkach mierzonych wielkości,
- dodatnio określoną macierz kowariancji lub korelacji oraz ponadto:
- obliczoną minimalną wartość własną macierzy korelacji (korelatora),
- informację o stosowanej precyzji obliczeń.

Przy takiej strukturze danych uzyskuje się pełną informację o wartościach, dokładności i skorelowaniu danych, niezbędną do planowania badań i do kontroli możliwości bezpiecznego korzystania z tych danych w innych pomiarach i w obliczeniach przy ich przetwarzaniu.

3. Zaokrąglenie wyników pomiarów wieloparametrowych

Najczęściej stosowanym nieliniowym przetwarzaniem danych pomiarowych jest zaokrąglenie wyrażeń liczbowych opisujących surowe wyniki pomiarów. Sposób zaokrąglenia danych multimezurandu, tj. wartości średnich jego składowych, ich niepewności i kowariancji (lub korelacji), nie jest dotąd znormalizowany. Operację tę realizuje się z dużą dowolnością. Sposób zaokrąglenia zależy od rodzaju zadania pomiarowego i często od tego, jak zamierza się ostatecznie przedstawić wyniki pomiarów do ich interpretacji i stosowania. Żaden z dotychczasowych podręczników i przewodników nie omawia

wystarczająco wnikliwie, jak poprawnie przetwarzać i okrąglać dane wektorowe.

Przyczyny losowych rozrzutów wartości wieloparametrowych obserwacji pomiarowych są dwójakiego rodzaju, tj.:

- same wielkości mierzone w obiekcie badanym mają charakter losowy, lub
- wartości tych wielkości są stałe, a wskutek zakłóceń w obiekcie badanym i torach pomiarowych pojawiają się rozrzuty.

Mogą też występować różne kombinacje obu tych źródeł losowości. Przy dominacji pierwszego z tych przypadków zadanie pomiarowe może polegać na dokładnym zbadaniu nie tylko samych estymat wartości średnich składowych mezurandu \mathbf{X} , ale i ich parametrów statystycznych, np. modelu rozkładu, standardowych odchyłeń i współczynników korelacji oraz ich niepewności. Należy wtedy pobierać próbki o bardzo dużej liczbie obserwacji, by ich parametry statystyczne były bliskie parametrom populacji wektora \mathbf{X} . Otrzymane w ten sposób dane wejściowe można traktować w przybliżeniu jako absolutnie dokładne statystycznie. Przy pośrednim wyznaczaniu z nich danych wektora \mathbf{Y} , zminimalizuje się utratę informacji, gdy koniec tego wektora po cyfrowym przetwarzaniu i zaokrągleniu pozostanie wewnątrz obszaru opisującego rozrzut przetworzonych, ale niezaokrąglonych danych wejściowych. Powinno się więc śledzić zmiany granic obszaru rozrzutu oraz położenie wierzchołka wektora średniego względem takiego obszaru rozproszenia, np. wykorzystując miarę odległości Mahalanobisa [11]. Dla spełnienia powyższego warunku i wymagania, by zachował się elipsoidalny obszar rozrzutu, V. Ezhela wyznaczył wzory dla progów zaokrąglania wartości, niepewności i współczynników korelacji mezurandu wyjściowego \mathbf{Y} , tj. minimalne liczb ich cyfr [9, 12]. Przy obu powyższych założeniach konieczne staje się wyznaczanie dużej liczby znaków dla parametrów wektora.

W praktyce pomiarowej, nawet gdy pomijalne są niepewności typu B składowych wektora, związane z niemożliwością wyeliminowania wpływu wszystkich systematycznych błędów instrumentalnych i błędów metody pomiaru, to wymagana dokładność cyfrowego przetwarzania danych wektora wejściowego \mathbf{X} będzie zależeć od liczności jego próbek pomiarowych. Liczność ta determinuje wielkość niepewności typu A (wyznaczanych statystycznie) nie tylko dla wartości składowych, ale i dla wszystkich parametrów statystycznych rozkładu każdego z wektorów \mathbf{X} i \mathbf{Y} . W tabeli E.1 przewodnika GUM [4] podano, że np. przy pięciu tylko pomiarach, jak w rozpatrywanym w nim przykładzie H.2, niepewność wyznaczonej z tych pomiarów niepewności standardowej nie przekracza 36 %, a niepewność współczynników korelacji jest zależna od ich wartości i będzie jeszcze większa. Precyzja przetwarzania danych nie musi być w takim przypadku nadmierna i przy małej liczności powinna wystarczyć by była o rząd większa niż najmniejsza cyfra niepewności. Progów zaokrąglania podanych w [12] dla przetwarzania nie należy stosować dla próbek niezbyt licznych lub gdy dokładność odchyłeń standardowych (i wyznaczanych z nich niepewności) i współczynników korelacji wektora wejściowego \mathbf{X} jest mała ze względu na nieusunięte nieznanne błędy instrumentalne (niepomijalna niepewność B).

Dla takich próbek zadanie pomiarowe polegać może jedynie na wyznaczeniu wartości średnich składowych wektora \mathbf{Y}

i to z takimi niepewnościami, jakie wynikają z niepewności danych wejściowych. Wyniki końcowe powinno się zaokrąglać tak, by nie wprowadzić zbyt dużych dodatkowych błędów. Niezależne zaokrąglenie składowych wektora wg zasad GUM dla skalarnych wielkości, może spowodować pojawienie się wyników niespełniających tych wymagań. Wyjaśni to przykład 1.

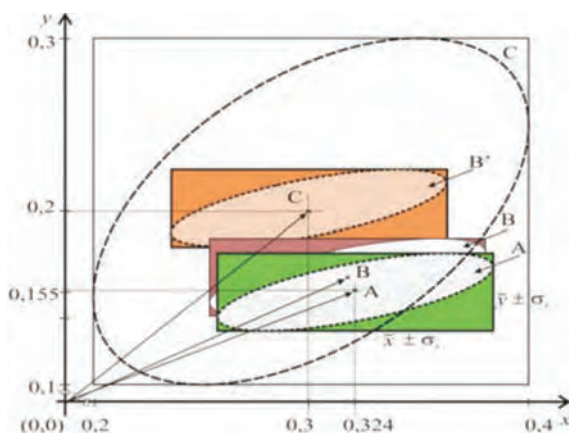
Przykład 1.

Rys. 2 ilustruje wyniki zaokrąglania danych surowych (tab. 1) opisywanych dwuwymiarowym rozkładem normalnym. Po jednolitym zaokrągleniu do trzech, dwóch i jednej cyfry wg zasad GUM jak dla skalara, tj. wartości średnich \bar{x} , \bar{y} do liczby najbliższej, otrzymano punkty położenia A, B i C końców wektora średniego. Natomiast elipsy A, B i C o środkach w tych końcach odpowiadają takim samym zaokrągleniom w górę odchyłeń standardowych $\sigma(\bar{x})$, $\sigma(\bar{y})$ i niezmiennej wartości współczynnika korelacji r_{xy} . Elipsa A pokrywa się praktycznie z elipsą dla surowych tj. niezaokrąglonych danych. Punkty B i C są przesunięte względem środka elipsy A. Elipsa B' jest przeniesioną do punktu C elipsą B. Wszystkie elipsy znajdują się wewnątrz prostokąta C. Elipsy B i C są większe od elipsy danych niezaokrąglonych A i powinny ją obejmować, ale ją przecinają, gdyż ich środki są przesunięte. Zmiana nachylenia osi elips B i C dokonana poprzez zmianę współczynnika korelacji r_{xy} tego nie rozwiąże. Opis elipsoidalnego obszaru rozrzutu umieszczonego niesymetrycznie względem końca zaokrąglonego wektora nie wydaje się być wygodny do stosowania w praktyce. Szczęśliwie sytuacja poprawia się jeśli rozpatrzeć elipsy (nie podane na rys. 2) odpowiadające niepewnościom rozszerzonym, np. o dwukrotne powiększonych wymiarach, tj. o współczynniku rozszerzenia $k_p = 2$ dla obu składowych. Z danych tab. 1 wynika, że obie takie elipsy obejmą już podwójnie powiększoną elipsę A, np. $2\sigma(\bar{x}_B) - (\bar{x} - \bar{x}_B) > 2\sigma(\bar{x}_B)$ Zaokrąglenie współczynników korelacji o wartościach bliskich ± 1 wymaga osobnego potraktowania. Wynika stąd, że należy opracować i przyjąć odpowiednie umowne ustalenia o jednolitym zaokrągleniu wartości, niepewności i współczynników korelacji mierzonych wielkości wektorowych i o precyzji wymaganej przy ich przetwarzaniu. Powinny się uwzględnić niepewności tych parametrów wynikające z ograniczonej liczby pomiarów w próbce i niepewności typu B od nieusuwalnych błędów instrumentalnych. Dwie takie propozycje dla danych przykładu H.2 z GUM przedstawiono w [12].

Tab. 1. Zaokrąglenia składowych i odchyłeń standardowych składowych wektora losowego [x, y]

Tab. 1. Rounding of components and standard deviations of the random vector [x, y]

Zaokrąglenia	\bar{x}_i	$\sigma(\bar{x}_i)$	\bar{y}_i	$\sigma(\bar{y}_i)$	$2\sigma(\bar{x}_i)$	$2\sigma(\bar{y}_i)$
Dane surowe	0,3242	0,0664	0,1555	0,0256	0,1328	0,0512
A. do 3 cyfr	0,324	0,067	0,156	0,026	0,133	0,051
B. do 2 cyfr	0,32	0,07	0,16	0,03	0,14	0,05
C. do 1 cyfry	0,3	< 0,1	< 0,2	< 0,1	< 0,2	< 0,1



Rys. 2. Elipsy A, B, C rozrzutu wektora $[x, y]^T$ o dwuwymiarowym rozkładzie normalnym uzyskane przy jednolitym (jak wg GUM dla skalarów) zaokrągleniu składowych i ich niepewności odpowiednio do 3, 2 i 1 cyfry oraz stałym współczynnikiem korelacji $r_{xy} = \text{const}$.

Fig. 2. Scattering ellipses A, B, C of vector $[x, y]^T$ of two variable Normal distribution obtained after recommended by GUM for scalars unified rounding of its components and uncertainty to 3, 2 and 1 digit and of constant correlation coefficient $r_{xy} = \text{const}$.

Inne przykłady ilustrujące możliwe zniekształcenia oszacowań wyników pomiarów wieloparametrowych przy stosowaniu procedury zalecanej dla przypadku skalarnego, omówiono w [11, 12]. W literaturze naukowej i technicznej występują wiele przykładów wyników pomiarów, które zawierają niepełne dane, nieprawidłowe ich oszacowania i niewłaściwe procedury obliczeniowe [9–12]. Oto główne przyczyny:

1. podaje się tylko średnie wartości składowych wektora i ich odchylenia standardowe, a pomija macierz korelacji;
2. dane elementów korelatora są „za bardzo zaokrąglone”, tj. wyznacznik macierzy i najmniejsza jej wartość własna przestają być dodatnie;
3. końcowy, zbytnio zaokrąglony wektor średni znacznie wychodzi poza granice obszaru rozrzutu surowych wyników obserwacji pomiarowych na wiele odchyleń standardowych. Jest to szkodliwa ukryta dezinformacja, gdyż to ten obszar określa niepewność oszacowania składowych wektora;
4. elipsa otrzymana dla rozszerzonej niepewności po zaokrągleniach nie obejmuje elipsy danych niezaokrąglonych.

4. Statystyczna analiza danych a prawa fizyczne

Głównym celem badań w fizyce jest ograniczenie liczby wielkości podstawowych i tworzenie matematycznych modeli służących wyjaśnianiu lub przewidywaniu wyników pomiarów jak największej liczby wielkości w oparciu o dane kilku wielkości podstawowych i prawa fizyki łączące je z wieloma różnymi wielkościami obserwowanymi (obserwabliami). Specyfikę pośrednich pomiarów wieloparametrowych uwidacznia przykład wyznaczania wartości n pochodnych stałych

fizycznych (lub wielkości powiązanych z nimi analitycznie) z danych m podstawowych stałych fizycznych FPC (*Fundamental Physical Constans*). Multimezurand wyjściowy jest zbiorem stałych pochodnych jako *obserwabli*. Większość tych stałych ma wymiary w jednostkach SI i jest zwykle nieliniowo zależna od dwóch lub więcej stałych FPC o różnych wymiarach. Odzworowanie obszaru rozproszenia wyników obserwacji w przestrzeni FPC w dziedzinę *obserwabli* jest więc operacją nieliniową. Wówczas powierzchnie bryły ograniczającej obszar rozrzutu wyników wielkości wyjściowych o określonym prawdopodobieństwie w wielowymiarowej euklidesowej przestrzeni multimezurandu wyjściowego podlegają zakrzywieniu. Stosując liniową propagację wariancji dla estymat rozrzutu danych, ignoruje się tę zależność i narusza prawa fizyki.

Wniosek: liniowa procedura propagacji jest w pełni prawidłowa tylko dla liniowych zależności. Wówczas liczba n obserwabli wyjściowych i mierzonych bezpośrednio m wielkości wejściowych spełnia zależność $n \leq m$, a odpowiadająca im macierz wrażliwości [9–12] nie jest zdegenerowana. Przy wystarczającej precyzji oszacowania nie wystąpią problemy z zapewnieniem dodatniej określoności macierzy i z oceną ufności obszaru rozproszenia w przestrzeni obserwabli. We wszystkich nieliniowych przypadkach w pełni poprawną jest tylko nieliniowa procedura propagacji niepewności, a liniowa jest dopuszczalna jako przybliżenie tylko dla niektórych z funkcji nieliniowych i małych wartości niepewności. W badaniach naukowych i metrologii należałoby obecnie w wielu opisach pomiarów przejść do używania znacznie już rozwiniętych teoretycznie niegaussowskich statystyk wielowymiarowych i teorii kopuł [2, 20], a dla zależności nieliniowych – do statystyk na bryle obrotowej [7].

5. Propagacja niepewności w pomiarach pośrednich o funkcjach nieliniowych

W Przewodniku GUM [4] podaje się zalecenie dotyczące propagacji niepewności u_c dla pojedynczej wielkości mierzonoj o nieliniowej funkcji przetwarzania $f(\cdot)$. W Uwadze do punktu 5.1.2 zaleca się, aby do wyrażenia

$$u_c^2(y) = \sum_{i=1}^N \left(\frac{\partial f}{\partial x_i} \right)^2 u_c^2(x_i) \tag{4}$$

wynikającego z liniowej propagacji wariancji stosowanej też powszechnie przy zależnościach nieliniowych o małych niepewnościach, dodawać dodatkowy człon o postaci

$$\sum_{i=1}^N \sum_{j=1}^N \left[\frac{1}{2} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \cdot \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_j} \right] u^2(x_i) u^2(x_j) \tag{5}$$

W wyrażeniu (5) w nawiasie występuje składnik z trzecią pochodną. Powoduje on, że dla wariancji $u^2(F)$ obliczonej wg tak poprawionego wzoru (4) można otrzymać wartość ujemną (przykład 2 [10]), niewystępującą w rzeczywistości.

Przykład 2

Należy wyznaczyć niepewność standardową wielomianu $F(x) = 1 - x + 2x^2 + 3x^3 + 4x^4$ dla zmiennej losowej x o rozkładzie normalnym z wariancją σ^2 wokół $x=0$.

Dla liniowej propagacji z (4) otrzymuje się: $u'(F) = \sigma^2$. Wg rozszerzonego o (5) wzoru (4) wariancja $u^2(F)$ wynosi:

$$u^2(F) = \sigma^2\{F'(0)^2 + \sigma^2[(1/2) F''(0)^2 + F'(0) F'''(0)]\} = \sigma^2(1 - 10\sigma^2).$$

Drugi składnik w nawiasie powoduje, że przy $\sigma^2 > 0,1$ otrzymamy ujemną wartość wariancji $u^2(F)$.

Dodatnią wariancję $u^2(F) > 0$ otrzyma się z zależności (4) uzupełnionej o (5), ale z pominięciem składnika z $F'''(0)$, tj.

$$u^2(F) = \sigma^2\{F'(0)^2 + \sigma^2(1/2) F''(0)^2\} = \sigma^2(1 + 4,5 \sigma^2)$$

Tego faktu, że zalecany w Uwadze do punktu 5.1.2 GUM dodatkowy składnik (5) do wzoru (4) jest nieprawidłowy, nie zauważyło ani środowisko naukowe, ani metrologiczne. Wzór ten wielokrotnie powtórzono w wielu dokumentach metrologicznych i podręcznikach oraz w publikacjach w czołowych światowych czasopismach naukowych i metrologicznych. O problemach wynikłych ze stosowania liniowego prawa propagacji niepewności nie wspomina się też w podręcznikach i poradnikach z fizyki i metrologii, lub nie analizuje się jego istoty [9].

W pomiarach wieloparametrowych o nieliniowych funkcjach przetwarzania wymiary wektorów wejściowego i wyjściowego oraz rząd T wyrazów szeregu Taylora są ze sobą następująco powiązane:

$$n \leq n_{th} = \frac{(m + T)!}{m! \times T!} - 1 \tag{6}$$

gdzie: n i m – wymiary wektorów mierzandu wyjściowego i wejściowego, T – rząd szeregu Taylora wielkości wyjściowej.

Dla wartości m i T o początkowych liczbach naturalnych z (6) wynikają następujące progowe wymiary wektora wyjściowego – liczba składowych n_{th} :

m	1	1	1	2	2	2	3	3	3	4	4	4	5	5	5
T	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
n_{th}	1	2	3	2	5	6	3	9	19	4	14	4	5	20	55

Do obliczenia niepewności wg nieliniowej propagacji dla n wielkości wyjściowych z danych m wielkości wejściowych, np.:

dla $m=2 \rightarrow n=4$ należy przyjąć $T=2$,

a dla $m=2 \rightarrow n=6$: $T=3$.

Podane tu wartości T liczby wyrazów rozwinięcia w szereg Taylora warto też (podobnie jak w przykładzie 2) zweryfikować, analizując stosowane w praktyce wzory i porównując otrzymywane niepewności dla przyjętej nieliniowej i liniowej propagacji.

Macierze wrażliwości wykorzystywane w opisie pomiarów wieloparametrowych podaje się zwykle bez sprawdzenia stosowalności aproksymacji liniowej. Jest to dodatkowe źródło

niepoprawnych danych w literaturze naukowej i technicznej, świadczące o braku koordynacji między metrologią i nauką.

W dokumentach metrologicznych nie są też jeszcze rozpatrywane następujące problemy:

- jak szacować estymaty składowych średniego wektora zbioru wektorów losowych o końcu przemieszczającym się po zakrzywionej powierzchni?
- jak wówczas scharakteryzować obszar rozpraszania i jego zasięg o zadanym prawdopodobieństwie?

Stosowanie nieliniowej propagacji niepewności jest istotne np. przy wyznaczaniu dokładności pomiarów wielkości fizycznych powiązanych nieliniowymi, nawet prostymi, np. iloczynowymi zależnościami algebraicznymi. Ujednolicone zalecenia, wraz z przykładami, powinna zawierać poprawiona wersja draftu Suplementu 2 do GUM o wyrażaniu wyników pomiarów wielowymiarowych, będącego obecnie w końcowym stadium opracowywania, lub należy opracować kolejny Suplement poświęcony tylko niepewności w pomiarach pośrednich o funkcjach nieliniowych.

7. Podstawowe dane odniesienia i ich związki

Centrum Danych o Stałych Podstawowych (*Fundamental Constants Data Center, FCDC*) w National Institute of Science & Technology (NIST, USA) wraz z postępowaniem regularnie koryguje wartości, niepewności i współczynniki korelacji podstawowych stałych fizycznych FPC i publikuje je w sprawozdaniach i czasopiśmie międzynarodowej organizacji CODATA. Jest to powszechnie uznane źródło o bieżących parametrach FPC zalecanych do stosowania w nauce, metrologii i innych dziedzinach. Korekcja dotyczy wartości liczbowych aż 325 wielkości, z których 79 to niezależne algebraicznie stałe fizyczne, czyli tzw. stałe podstawowe C_a^B . Do ich oszacowania, według równań wiążących je z wielkościami mierzonymi, stosuje się metodę najmniejszych kwadratów (LSA). Pozostałe stałe C_i^D , nazywane pochodnymi, szacuje się według zależności funkcyjnych wiążących je ze stałymi podstawowymi:

$$C_i^D = \Phi_i(C_a^B) \tag{7}$$

Zwykle zależności Φ_i są prostymi funkcjami algebraicznymi, ale w większości przypadków są one nieliniowe i mogą występować współzależności między stałymi pochodnymi. Przy dużych dokładnościach niewłaściwe jest stosowanie liniowej propagacji wariancji dla niepewności i nadmierne zaokrąglanie. W prezentacji wielkości C_i^D występuje globalna macierz korelacji o wymiarach 325×325 . Obliczanie elementów tej macierzy według liniowej propagacji niepewności powoduje znaczną jej degenerację i otrzymuje się około $325 - 79 = 246$ równych zeru wartości własnych.

Jako przykład w [9, 17] podano numeryczną postać wektora czterech często stosowanych stałych fizycznych, tj. ładunku elementarnego e , stałej Plancka h , masy elektronu m_e i stałej struktury subtelnej $\alpha(0)^{-1}$ w jednostkach SI, dla korekt wartości tych stałych publikowanych przez międzynarodową organizację CODATA w okresie 1986–2006. W ich

Tab. 2. Wartości i niepewności wybranych stałych fizycznych [e , h , m_e , $1/\alpha(0)$] według CODATA 2010**Tab. 2.** Values and their uncertainty of chosen physical constants [e , h , m_e , $1/\alpha(0)$] according to CODATA 2010

Rok korekty	Stałe fizyczne	Symbol	Miara	Wartość (niepewność) · mnożnik	Współczynniki korelacji
CODATA: 2010 (WWW) [13]	ładunek elementarny	e	C	$1,602\ 176\ 565(35) \cdot 10^{-19}$	
	stała Plancka	h	Js	$6,626\ 069\ 57(29) \cdot 10^{-34}$	1,0000
	masa elektronu	m_e	kg	$9,109\ 382\ 91(40) \cdot 10^{-31}$	0,9998 0,9999
	1/stała struktury subtelnej	$1/\alpha(0)$		137,035 999 074(44)	-0,0145 -0,0072 0.0075
	Wartości własne macierzy korelacji: [2,99942 1,00006 -0,000451912 -5,92939 · 10 ⁻⁶] (ujemne pogrubiono)				

danych wykryto kilka nieprawidłowości, w tym we wszystkich tych korektach ujemne wartości własne macierzy korelacji [9] (tab. 3). Wartości w jednostkach SI tych stałych wg danych ostatniej korekty CODATA 2010 z poprawkami z 2011 r. [13] zestawiono w tab. 2.

Podkreślone stałe fizyczne h i $\alpha(0)$ są wyznaczone z pomiarów metodą najmniejszych kwadratów, natomiast masę elektronu m_e oraz jego ładunek e obliczono wg metody CODATA [6] z zależności iloczynowych typu (7), tj.

$$m_e = \frac{2R_\infty h}{c_0 \alpha^2}, \quad e^2 = \frac{2h\alpha}{\mu_0 c_0}$$

gdzie: $R_\infty = 10,973\ 731\ 568\ 539(55) \cdot 10^6\ \text{m}^{-1}$ – stała Rydberga, $\mu_0 = 4\pi \times 10^{-7}\ \text{NA}^{-2}$ – przenikalność magnetyczna i $c_0 = 299\ 792\ 458\ \text{m/s}$ – prędkość światła w próżni.

Macierz korelacji (tab. 2) ma ujemne wartości własne. Tak więc dane FPC-2010 nie mogą być bez zastrzeżeń używane do szacowania wartości wielkości wyrażanych przez więcej niż dwie z tych stałych fizycznych aż do czasu, gdy stanowiące ich podstawę dane numeryczne stałych fizycznych zostaną ujawnione i oszacowane metodą najmniejszych kwadratów o wystarczającej liczbie cyfr znaczących. Jest to jeszcze jeden dowód, że potrzebne są normy dla numerycznego wyrażania i przekazywania skorelowanych danych oraz ich recenzowania. Brak takich norm doprowadził do pojawienia się wewnątrz społeczności naukowej i technicznej ukrytego „binarnego podziału” na twórców, których aktywność zawodowa służy pozyskiwaniu jak najlepszej jakości danych pomiarowych i na nie w pełni usatysfakcjonowanych odbiorców tej wiedzy. Dotychczasowa forma publikacji stwarza bowiem odbiorcom swoistą zaporę (*firewall phantom*), gdyż nie mogą oni sami ocenić prawidłowości udostępnianych im danych.

8. Propozycja dualnej formy publikacji danych

Analiza podstawowej literatury [1, 2, 4, 18] i wielu omówionych w [9] prac z fizyki oraz poradników metrologicznych o przetwarzaniu danych pomiarowych (np. [14]) wykazała braki i konieczność szybkiego przyjęcia nowych przepisów. W metrologii nie ma dotąd zaleceń dla procedur numerycznego wyrażania wyników i przekazywania danych multimedialnych. Suplement 2 do GUM, który jest ciągle w stadium roboczym (*draft*) i nie zaspokoi w pełni tej potrzeby. Używanie w badaniach naukowych i stosowanych niewłaści-

wych procedur powoduje, że opublikowane dane często nie są w pełni prawidłowe. Na podstawie analizy [3, 9] przykładów tzw. „złych praktyk” starano się sformułować wstępne wymagania dla poprawnej prezentacji i przekazywania danych pomiarowych.

Forma komunikacji naukowej i technicznej jest w coraz to większym stopniu oparta na e-publikacjach, ale jeszcze zorientowanych na tradycyjną formę przekazu – na papierze. Nie wystarcza to już do wymiany danych z doświadczeń wieloparametrowych oraz w metrologii o najwyższych dokładnościach. Jednakże dzięki elektronicznym multimedialnym środkom publikowania, czyli e-publikatorom, otwierają się nowe możliwości, które pozwalają uniknąć ograniczeń techniki publikowania na papierze. Nieuniknione jest przejście na dwuczściową formę publikacji [9]. Tradycyjnemu tekstowi opisowemu, dobrze już sformalizowanemu redakcyjnie, towarzyszyłyby związane z nim pliki komputerowe przedstawiające dane źródłowe o odpowiedniej rozdzielczości cyfrowej, w pełni czytelne dla komputera i przeznaczone do długotrwałego przechowywania i swobodnego dostępu. Dzięki temu umożliwi się recenzentom i użytkownikom danych sprawdzanie ich kompletności i spójności raportowanych danych, ułatwi się ich pobieranie oraz zapobiegnie degradacji przy przekształcaniu z jednej formy w drugą oraz zapewni praktycznie wieczne ich przechowywanie. To proste i logiczne postępowanie zbyt wolno przebija się do praktyki pomimo intensywnych dyskusji o jakości danych, zarządzaniu, ochronie wiedzy itp. Podobną opinię opublikowano z kręgów organizacji OECD [8] (cyt. w [9]).

Przy stosowaniu plików komputerowych pojawia się też kilka nowych problemów:

- jaka powinna być minimalna struktura danych, aby prawidłowo zapisać wyniki badań?
- jak kontrolować precyzję numeryczną danych liczbowych, aby nie utracić wyników przy transformacji ich formy?

Nauka i metrologia nie są jeszcze przygotowane do pełnego wykorzystywania ogromnych możliwości e-publikatorów w zachowywaniu poprawnych wyników pomiarów o praktycznie nieograniczonej objętości z hyperlinkami i z szybką transmisją. Trzeba sformułować odpowiednie standardy dla plików odczytywanych i zrozumiałych dla komputera, związanych z publikacjami na papierze i zawierających pełne dane pomiarowe prawidłowo przedstawione liczbowo. Metrologia i inne dziedziny działalności eksperymentalnej stoją więc przed wielkim wyzwaniem – opracowaniem zaleceń do wyrażania i przekazywania wielowymiarowych danych pomiarowych za pomocą e-publikacji i wdrożenie praktyki me-

to numerycznej weryfikacji tych danych. Istnieje już kilka przykładów standardów roboczych stosowanych w wiodących centrach danych. Mogą one stanowić podstawę do projektu takich przepisów. Przykłady można zaczerpnąć i uogólnić np. ze struktur danych w centrach NNDC (*National Nuclear Data Center*), FCDC, AME, PDG (*Particle Data Group*).

W korekcie FPC-2002 [5] dane pomiarowe przedstawione były w plikach, w sposób pełny i poprawny wraz z rezultatami oszacowania metodą najmniejszych kwadratów jednocześnie dla 61 wartości podstawowych stałych fizycznych. Pliki FPC podano wraz z publikacją [5] na internetowej stronie NIST jako LSA-2002. Jest to pierwszy przykład właściwej prezentacji wielowymiarowych danych pomiarowych. Można go potraktować jako podstawę do opracowania projektu standardu międzynarodowego. Są to pliki czytelne komputerowo w ASCII i nie wymagają ręcznej interwencji do reedycji strony w liczby. Jeśli pliki o formie LSA-2006 zostałyby powszechnie zaakceptowane, byłby to istotny krok do opracowania nowoczesnych zaleceń dla dualnej formy prezentacji danych wielowymiarowych. Baza danych FPC CODATA w centrum FCDC (NIST) oraz dostosowanie technologii tam stosowanych może stanowić poligon badawczy do opracowywania, doskonalenia i ewolucji takich zaleceń dla pomiarów wieloparametrowych przed powszechnym jego stosowaniem.

9. Podsumowanie i wnioski ogólne

Praca ta miała za zadanie przybliżyć ujęty ogólnie stan całości problemów występujących przy wyrażaniu i publikowaniu wyników pomiarów pośrednich wieloparametrowych. Scharakteryzowano krótko opis wyników tych pomiarów w ujęciu wektorowym, skorygowano omyłkę w zaleceniu 5.1.2 GUM dotyczącą wyznaczania niepewności dla funkcji nieliniowych oraz wskazano na nieścisłości w ostatniej korekcie danych czterech podstawowych stałych fizycznych opublikowanych przez NIST i CODATA w 2010 r. z uwzględnieniem ostatnich poprawek z czerwca 2011 r.

Nawiązano też do propozycji [9, 17], aby zakresem przepisów międzynarodowych objąć dualny sposób prezentacji i publikowania danych pomiarów wieloparametrowych – na papierze i w formie elektronicznej. Propozycja ta powstała jako rezultat analizy dotychczas stosowanych sposobów numerycznego wyznaczaniu i przedstawianiu skorelowanych wieloparametrowych danych pomiarowych w publikacjach i plikach komputerowych [9]. Wspólne dla całej społeczności naukowej i technicznej trudności występują wskutek braku akceptowanych powszechnie standardów numerycznego wyrażania takich danych i niemożności ich numerycznej weryfikacji przy tradycyjnej formie publikacji „na papierze” oraz wskutek w dość dowolny sposób stosowania obecnie formie elektronicznej. Tych trudności uniknie się po ujednoczeniu sposobu wyrażania danych wielowymiarowych w formie „czytelnej i zrozumiałej” dla komputera.

Przewiduje się, że dzięki temu mogłyby nastąpić istotne zmiany w sposobie posługiwania się danymi pomiarowymi. Do rozwiązania jest jeszcze szereg problemów, by pracę z danymi pomiarowymi w nauce, metrologii i technice, bankowości oraz w wielu innych dziedzinach uczynić w przyszłości bardziej komfortową i użyteczną.

Należałoby zbadać, jak stosowane w badaniach fizycznych metody i procedury precyzyjnego przetwarzania ogromnych zbiorów danych wieloparametrowych można zaadaptować do metrologii najwyższych dokładności i wieloparametrowych pomiarów użytkowych o bardzo odpowiedzialnych zastosowaniach, a po niezbędnych uproszczeniach – i do pomiarów z wieloparametrowymi próbkami o małej liczbie obserwacji. Dotyczy to w szczególności pomiarów o rozrzutach wyników opisywanych niepewnością typu A porównywalną lub większą od niepewności typu B ujmującej występowanie nieusuwalnych przez poprawki resztek błędów systematycznych, w tym instrumentalnych. Metoda ta powinna też znaleźć się w zaleceniach Suplementu 2 do GUM.

Zracjonalizowane sposoby opisu niepewności pomiarowych danych wektorowych mogą być zastosowane w wielu dziedzinach technicznych, m.in. w identyfikacji obiektów sterowania, badaniach materiałów i powiązanych ze sobą zmian parametrów różnych obiektów pod wpływem wielu równoczesnych oddziaływań oraz w diagnostyce technicznej urządzeń i procesów, a ponadto w badaniach medycznych i monitoringu środowiska.

Bibliografia

1. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, and OIML: *Guide to the expression of uncertainty in measurement*, ISO/IEC Guide 98:1995, 2nd ed. (Wyrażanie Niepewności Pomiaru Przewodnik), Główny Urząd Miar, Warszawa, Wyd. Alfavero 2002.
2. Muciek A.: *Matematyczny model propagacji niepewności w pomiarach pośrednich*, Podstawowe Problemy Metrologii, Materiały Sympozjum PPM'03, seria: Konferencje nr 5, Oddz. PAN w Katowicach, 2003, s. 593–604.
3. Ezhela V.: *A multi-measurand ISO GUM supplement is urgent*, „Data Science Journal” 6, s. 676–789 [Errata: CODATA DSJ 7, 2007 E2–21].
4. *Evaluation of measurement data – Guide to the expression of uncertainty in measurement* (GUM 1995 with minor corrections), [www.bipm.org/utis/common/documents/jcgm/JCGM_100_2008_E.pdf].
5. Mohr P.J., Taylor B. N.: *CODATA recommended values of the fundamental physical constants* 2002, „Rev. Mod. Phys.” 77, 1/2005, 1CODATA-2002.
6. Mohr P.J., Taylor B.N., Newell D.B.: *CODATA recommended values of the fundamental physical constants: 2006*, „Reviews of Modern Physics” 80 2008, s. 633–730; CODATA-2006.
7. Pennec X.: *Intrinsic Statistics on Riemannian Manifolds: Basic Tools for Geometric Measurements*, „Journal of Mathematical Imaging and Vision” 25, 2006, s. 127–154.
8. Green T.: *We Need Publishing Standards for Data Sets and Data Tables*, OECD Publishing White Paper, OECD 2009, [http://dx.doi.org/10.1787/603233448430].
9. Ezhela V.: *Physics and Metrology*, Materiały V Kongresu Metrologii KM 2010, Politechnika Łódzka, CD.
10. Ezhela V.: *Comments on some clauses of GUM which provoking the incorrect presentation of measured data in scientific literature*. Materiały V Kongresu Metrologii KM 2010, Politechnika Łódzka, CD.

11. Warsza Z., Ezhela V.: *Wyznaczanie parametrów multi-mezurandu z pomiarów wieloparametrowych, Część 1. Podstawy teoretyczne w zarysie*. PAR 2/2011, s. 55–61.
12. Warsza Z., Ezhela V.: *Wyznaczanie parametrów multi-mezurandu z pomiarów wieloparametrowych, Część 2. Reguły zaokrąglania, nieściśłości w przewodniku GUM*, PAR 6/2011, s. 64–70.
13. Mohr P.J., Taylor B.N., Newell D.B. (2011), „The 2010 CODATA Recommended Values of the Fundamental Physical Constants”, [http://physics.nist.gov/cuu/Constants/index.html], NIST (web version 6.2).
14. *Measurement Uncertainty Analysis Principles and Methods*, NASA Measurement Quality Assurance Handbook – ANNEX 3, NASA-HDBK-8739.19-3, July 2010.
15. Dorozhovetz M., Warsza Z.L.: *Propozycje rozszerzenia metod wyznaczania niepewności wyniku pomiarów wg Przewodnika GUM (1) Uwzględnianie wpływu autokorelacji i nieadekwatności rozkładu wyników obserwacji w niepewności typu A*. PAR 1/2007 s. 16–25.
16. Galovska M., Warsza Z.L.: *Estymatory wartości mezurandu próbek danych o rozkładach niegaussowskich*, [w:] *Metrologia dziś i jutro*, Oficyna Wyd. Pol. Wrocławskiej, Wrocław 2010, s. 59–72.
17. Warsza Z., Ezhela V.: *Nieściśłości stałych podstawowych i propozycja standaryzacji dualnego sposobu publikowania wyników pomiaru multi-mezurandu*, PAK vol. 57, 5/2011 s. 486–490.

Uzupełniająca literatura polska

18. Szydłowski H. i inni: *Teoria pomiarów*, PWN Warszawa 1981 (rozdz. 11. Zastosowanie macierzy... str. 354-379)
19. Kukielka L., *Podstawy badań inżynierskich*, Wydawnictwo Naukowe PWN, Warszawa 2002.
20. Pawłowski J.: *Wprowadzenie do teorii kopuł*. Kraków, marzec 2009, Internet. ■

About evaluation and publication of multivariate measurement data – current status and needs

Abstract: Discussed is the issue of status of the determination, presentation and publication of multivariate measurement data as continuation of the publications of PAR 2 and 6 of 2011. Basic rules for the description of the statistically related measurement results by random vectors algebra are given. Problems becoming with rounding of vector measurands are discussed. Propagation of uncertainty of nonlinear functions in the measurement of single- and multi-parameter is presented. Determined is the number of members of Taylor series of nonlinear functions required for accurate determination of uncertainty. It was shown that the four fundamental physical constants published in the recent revisions

of the CODATA 2010 are still incorrectly given, as they correlation matrix has negative eigenvalues. Corrected recommendations for expressing the results of multivariate measurements should be included in Supplement 2 to guide GUM-2008, which is in preparation and be published in the near future. It is proposed that as necessary complement to it was the urgently needed standardization of the publishing measurement data in dual form: so far – on paper and on the accompanying e-publication containing the results of the original measurements. Discusses a way of collecting and disseminating data already taken by the world's leading data centers. Given are also a brief conclusions and bibliography.

Keywords: multivariate measurements, correlated data uncertainty evaluation, fundamental constants, e-publishing

doc. (em.) dr inż. Zygmunt Lech WARSZA

Ukończył Miernictwo Elektryczne na Politechnice Warszawskiej 1959, doktorat 1967, docent od 1970. Zorganizował i prowadził: Wydział Transportu Pol. Świętokrzyskiej 1970–76, Ośrodek Aparatury Pomiarowej IMGW 1978–83, Zakład Automatyki i Techniki Pomiarowej Instytutu Chemii Przemysłowej 1983–92. Był doradcą Ministra Edukacji Narodowej 1992–94. Wykładał elektrotechnikę, metrologię i diagnostykę techniczną na Pol. Radomskiej do 2002. Pracował w Instytucie Elektrotechniki i Pol. Warszawskiej, obecnie pracuje w PIAP. Autor ok. 130 publikacji, 2 monografii, kilkudziesięciu prac badawczych i konstrukcyjnych, 11 patentów oraz promotor 2 doktorów.

e-mail: zlw@op.pl

dr Vladimir V. EZHELA

W 1968 ukończył Moskiewski Instytut Fizyki i Techniki (MPTI) ze specjalnością fizyka jądrowa. Od 1968 pracuje w Instytucie Fizyki Wielkich Energii IHEP. W 1971 odbył aspiranturę w MPTI pod kierunkiem akademika A. A. Logunova, od 1971 w pionie teoretycznym IHEP, 1977–2003 starszy pracownik naukowy, kierownik Centrum Danych Fizyki Cząstek (PPDC), 2004 – główny pracownik naukowy w PPDC IHEP. Autor kilkudziesięciu publikacji naukowych, współpracuje z CODATA.

e-mail: vladimir.ezhela@ihep.ru

