

Małgorzata Loga

Szacowanie ufności i precyzji – podstawowych miar niepewności oceny stanu wód powierzchniowych

Przeniesienie Ramowej Dyrektywy Wodnej (RDW) [1] do polskiego prawa, najpierw w postaci rozporządzenia Ministra Środowiska z 20 sierpnia 2008 r., następnie z 9 listopada 2011 r., a obecnie z 30 października 2014 r. [2], wprowadziło sposób klasyfikacji stanu wód powierzchniowych w Polsce, obejmujący ocenę ich stanu ekologicznego i chemicznego. Procedura oceny stanu ekologicznego wód powierzchniowych opiera się na wynikach klasyfikacji elementów jakości biologicznej oraz – pełniących rolę wspierającą – elementów fizyczno-chemicznych i hydromorfologicznych. Stan chemiczny określany jest na podstawie pomiarów zawartości w wodzie substancji priorytetowych oraz innych zanieczyszczeń. Jednak w żadnym z wymienionych rozporządzeń nie wspomina się, jak szacować niepewność tych ocen, choć konieczność takiego postępowania wskazują wytyczne sformułowane w poradniku do monitoringu [3], według którego „ocenie stanu powinny towarzyszyć ufność, precyzja oraz ryzyko błędnej oceny”. Tymczasem wymagane przez RDW utrzymanie dobrego stanu wód wiąże się z ponoszeniem ryzyka środowiskowego i ryzyka ekonomicznego, których źródłem są niepewności tkwiące w monitoringowych danych pomiarowych. Wymóg, aby odpowiednie miary ufności, precyzji oraz ryzyka błędnej oceny były przypisane każdej ocenie wód jest niezaprzeczalnie potrzebny, jednak problem polega na tym, że poradnik [3] zawiera jedynie ogólne wytyczne dotyczące tych niepewności i nie określa miar niepewności, pozostawiając ich zdefiniowanie odpowiednim agendum odpowiedzialnym za monitoring wód w państwach członkowskich UE. Od momentu wejścia w życie Ramowej Dyrektywy Wodnej, w różnych ośrodkach europejskich rozpoczęły się prace nad zagadnieniem szacowania niepewności. Zagadnienie niepewności dotyczące oceny stanu wód na podstawie indeksu rybnego, opracowywanego w ramach projektu FAME, zostało przedstawione między innymi w raporcie [4]. Szczególnie ciekawe rezultaty uzyskano w przypadku wskaźnika jakości biologicznej wód, opartego na bezkręgowcach bentosowych, w wyniku wspólnych prac w projektach RIVPACS, STAR i STAR-AQEM, opisanych między innymi w publikacjach [5–7]. W przypadku makrofitów opracowano metodę szacowania niepewności oceny stanu jezior [8]. Jakkolwiek w tych projektach wykorzystano dane dotyczące polskich wód, to jednak opracowana metoda nie została wdrożona w Polsce, pozostawiając otwarte zagadnienia szacowania niepewności oceny oraz prawdopodobieństwa błędnej klasyfikacji.

Na początku wdrażania Ramowej Dyrektywy Wodnej – w Polsce oraz w tych państwach członkowskich, które stosunkowo niedawno dołączyły do wspólnoty Unii Europejskiej – zagadnienie szacowania miar niepewności oceny stanu mogło dotyczyć jedynie niewielkiej grupy jednolitych części wód, ze względu na niewielką liczbę historycznych pomiarów monitoringowych zapewniających wiarygodne oszacowanie odpowiednich parametrów statystycznych. Ustanowienie nowej sieci punktów pomiarowo-kontrolnych, zgodnej z wprowadzonym podziałem wód powierzchniowych na jednolite części wód, systematyczne wykonywanie pomiarów monitoringowych tych wód zgodnie z RDW oraz gromadzenie danych pomiarowych przez Wojewódzkie Inspektoraty Ochrony Środowiska w ciągu dziesięciu lat doprowadziły do sytuacji, w której szacowanie ufności, precyzji oraz prawdopodobieństwa błędnej oceny w przypadku poszczególnych (choć nie wszystkich) jednolitych części wód stało się możliwe.

Niniejszy artykuł zawiera definicje dwóch podstawowych miar niepewności oceny stanu wód powierzchniowych wymaganych przez RDW – ufności i precyzji. Trzecia miara tej niepewności – prawdopodobieństwo błędnej klasyfikacji – zasygnalizowana w pracy [9], jest obecnie testowana na podstawie danych Państwowego Monitoringu Środowiska. Sposób wprowadzenia miar niepewności ocen naśladuje hierarchiczną strukturę procedur klasyfikacyjnych Ramowej Dyrektywy Wodnej. Proponowane definicje miar niepewności zostały następnie zastosowane do oceny wybranych jednolitych części wód (JCW) z wykorzystaniem rzeczywistych danych pomiarowych pochodzących z polskiego systemu monitoringu wód powierzchniowych. Celem artykułu było wskazanie prostych definicji miar niepewności ocen, które mogłyby stać się elementami prawa obowiązującego w Polsce.

Hierarchiczna metoda oceny stanu jednolitych części wód (JCW) według Ramowej Dyrektywy Wodnej (RDW)

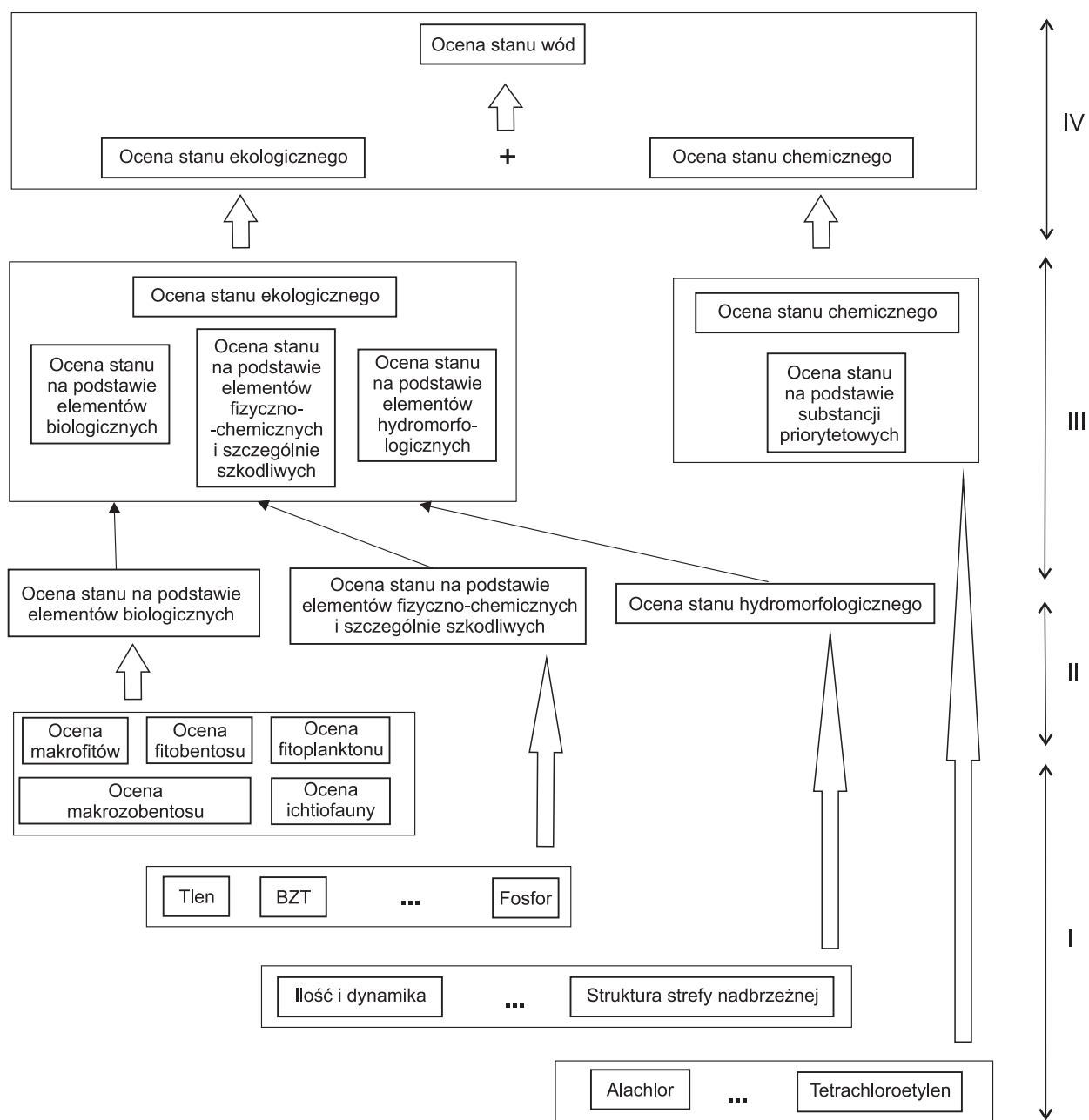
Ocenianie stanu wód wykonywane we wszystkich ich jednolitych częściach w dorzeczu na podstawie monitoringu diagnostycznego, prowadzonego co najmniej w ciągu jednego roku każdego sześcioletniego okresu obowiązywania planów gospodarowania wodami lub częstszej w oparciu o mniejszą liczbę wskaźników w przypadku części wód objętych monitoringiem operacyjnym, jest istotnym elementem prawidłowej gospodarki wodnej na terenie Unii Europejskiej. Wynik oceny stanu wód, wskazujący na nieosiągnięcie obecnie stanu dobrego lub wskazujący na pogarszanie się ich stanu albo prognoza nieosiągnięcia dobrego

stanu wód w przyszłości, powinny wywołać uruchomienie mechanizmów naprawczych, przeciwdziałających pogarszaniu się stanu danej JCW.

Procedura klasyfikacyjna, przedstawiona szczegółowo w wytycznych do prowadzenia monitoringu zgodnego z wymaganiami Ramowej Dyrektywy Wodnej [3], ma strukturę hierarchiczną składającą się z czterech poziomów. Na najniższym z poziomów, oznaczonym na rysunku 1 jako I, następuje przyporządkowanie mierzonych wskaźników jakości wody do odpowiedniej klasy. Dotyczy to wskaźników w grupach 'w' reprezentujących właściwości fizyczno-chemiczne ($w=fc$), biologiczne ($w=b$), hydromorfologiczne ($w=h$) i chemiczne ($w=ch$). Ta część klasyfikacji obejmuje – obliczenie wartości średniej i/lub wartości maksymalnej ze zbioru wartości będących wynikiem pomiarów monitoringowych w pewnym okresie czasu dla którego wykonywana jest ocena (na ogół w czasie 6 lat) oraz porównanie tych wartości średnich

lub maksymalnych, z granicami klas, które dla poszczególnych wskaźników określone są w rozporządzeniu [2]. Ten sposób klasyfikacji można określić trybem prostym. Na najwyższym poziomie, oznaczonym na rysunku 1 jako IV, dokonuje się ostateczna klasyfikacja wód, czyli wyznaczenie stanu wód w JCW w trybie, który można nazwać złożonym. Tryb złożony, występujący także na etapach II i III, polega na stosowaniu reguły „decyduje najgorszy” (one-out-all-out) oraz kilku reguł logicznych przedstawionych w wytycznych [3]. Oszacowanie miar niepewności na etapie I polega na wykonaniu dwóch pierwszych kroków algorytmu oceny stanu, w odniesieniu do wskaźników należących do wszystkich grup.

Ocena stanu JCW w odniesieniu do każdej grupy wskaźników 'w' (czyli ocena na etapie II) dokonywana jest w trybie złożonym poprzez zastosowanie zasady „decyduje najgorszy”, osobno do wskaźników każdej grupy. Podobnie przeprowadzana jest ocena na etapie III. Wówczas tę



Rys. 1. Struktura hierarchiczna klasyfikacji i oceny stanu jednolitych części wód
Fig. 1. Hierarchical structure of water body status classification and assessment

zasadę stosuje się do grup ocen zestawionych już tylko w dwie „super-grupy” – grupę ocen stanu ekologicznego oraz (jednoelementową) grupę ocen zawierającą jedynie ocenę stanu chemicznego. Na najwyższym – IV – etapie wskazanie klasy, do której należy jednolita część wód odbywa się także poprzez zastosowanie zasady „decyduje najgorszy”. Na tym etapie jednolitej części wód zostaje przypisana klasa (ocena) odpowiadająca niższej z dwóch ocen – ocenie stanu ekologicznego i ocenie stanu chemicznego. Oszacowanie miar niepewności na etapach II, III i IV polega na wykonaniu trzech ostatnich kroków algorytmu oceny stanu wód, w odniesieniu do oceny wskaźników lub grup ocen.

Źródła niepewności w ocenie stanu wód

Sposób oceny stanu wód powierzchniowych w Polsce jest procedurą dobrze zdefiniowaną, polegającą na porównaniu wartości pomierzonych wskaźników charakteryzujących jakość wody w jednolitych częściach wód lub ich prostych parametrów statystycznych z wartościami granicznymi podanymi w rozporządzeniu [2]. Jednak – mimo iż ocena i klasyfikacja stanu wód są opisane za pomocą procedur deterministycznych – losowa zmienność wskaźników wykorzystywanych do oceny stanu wód powoduje, że również wyniki poszczególnych ocen noszą cechy losowości. Czynniki przyczyniające się do losowości oceny wód można (podobnie jak samą ocenę) uporządkować hierarchicznie od czynników pierwotnych, przez czynniki wtórne [10], po czynniki wyższego rzędu. Jako założenie w dalszych rozważaniach przyjęto prostą hipotezę (model), że obserwowany losowy charakter wartości wskaźników jakości wody odzwierciedla bezpośrednio losowość procesów zachodzących w środowisku wodnym oraz losowe zaburzenia wnoszone przez (szeroko rozumiany) proces pomiarowy. W niniejszym artykule czynniki powodujące losowość oceny stanu wód dyskutowano jedynie w odniesieniu do zespołu wskaźników jakości wód określonych w rozporządzeniu [13].

Skomplikowany charakter procesów zachodzących w środowisku wodnym, wynikający z nakładania się na siebie znacznej liczby czynników powoduje, że wartości wskaźników jakości wody w rzekach i jeziorach mają charakter losowy. Czynniki zewnętrzne pośrednio wpływające na zmienność czasową i przestrzenną wskaźników jakości wód powierzchniowych mogą mieć charakter naturalny (np. czynniki meteorologiczne, szata roślinna pokrywająca zlewnię) lub wynikać z działalności człowieka (np. odprowadzanie ścieków, spływ biogenów z obszarów użytkowanych rolniczo itp.). Jakkolwiek rozdzielanie czynników antropogenicznych wpływających na stan wód powierzchniowych od czynników naturalnych jest jednym z głównych zadań, jakie Ramowa Dyrektywa Wodna stawia systemom monitoringu wód, problem ten wciąż daleki jest od ostatecznego rozwiązania. Podchodząc pragmatycznie do zagadnienia źródeł niepewności ocen w niniejszym artykule przyjęto, że w kontekście celu tych rozważań skomplikowane procesy hydrodynamiki oraz transportu i wymiany masy w systemach wodnych można traktować jako (nieznane) pola losowe bez odnoszenia się do czynników wywołujących tę losowość (tj. bez odnoszenia się do procesów w skali zlewni). Te złożone i sprzężone pola losowe są traktowane w dalszej dyskusji jako praprzyczyna losowości występującej w danych pomiarowych monitoringu wód powierzchniowych. Zaobserwowano bowiem,

że wyniki punktowych pomiarów wartości wskaźników jakości wody wykonywanych wielokrotnie w jednolitych częściach wód charakteryzują się fluktuacjami, które dają się opisać jako zmienne losowe. W dalszej części artykułu opisano czynniki, które powodują tę losowość i nazwano pierwotnymi źródłami niepewności ocen. Wyniki pomiarów punktowych w jednolitych częściach wód są zazwyczaj agregowane, a wynik tej operacji przypisywany jest danej JCW w postaci jednej (zagregowanej) wartości wskaźnika. W tego typu pomiarach określa się wtórne źródło niepewności ocen – jest nim sama metoda agregacji.

Także – niemożliwe do uniknięcia – losowe zaburzenia wielkości mierzonej (wskaźnika) powstające w wyniku czynności manualnych lub automatycznych towarzyszących poborowi próbek wody do badań laboratoryjnych, ich transportowi, magazynowaniu i przechowywaniu próbek wody, należy uznać za pierwotne źródła niepewności. Podobnie przygotowanie próbek wody do badań i wyznaczenie wartości wskaźników jakości wody w warunkach laboratoryjnych wprowadza do procesu pomiarowego kolejne zaburzenia wynikające z przyjętych metod analitycznych, czynności manualnych i instrumentalnych/aparaturowych oraz metod interpretacji laboratoryjnych wyników pomiarowych. Wszelkie czynniki związane z laboratoryjną fazą wyznaczania wartości wskaźników jakości w danej JCW są pierwotnymi źródłami losowości wartości pomiarowych wskaźników. Podsumowując, losowe fluktuacje wartości pomiarowych każdego ze wskaźników jakości wody wykonywanych wielokrotnie w danej JCW [11] są wynikiem działania wszystkich czynników jednocześnie. Złożoność tego współdziałania powoduje, że niepewność sumaryczna i -tego wskaźnika należącego do grupy elementów w , z konieczności szacowana jest za pomocą jednej liczby – empirycznego odchylenia standardowego $\hat{\sigma}_i^w$. Na przykład empiryczną wartość odchylenia standardowego zawartości chlorków w wodzie, należących do grupy elementów jakości fizyczno-chemicznej (służąca do oceny stanu ekologicznego) można zapisać w postaci $\hat{\sigma}_{Cl}^{fc}$.

Należy zauważyć, że wobec braku obowiązującej w polskim monitoringu metody oceny hydromorfologicznej, obecnie trudno wyobrazić sobie wyliczanie odchylenia standardowego jakiegokolwiek wskaźnika wchodzącego w skład tej grupy. Jednak biorąc pod uwagę np. brytyjską metodę RHS (river habitat survey), zaadaptowaną do polskich warunków przez K. Szoszkiewiczą [12], można wyznaczyć wartość średnią czy wariancję wskaźników naturalności siedliska (HQA) oraz przekształcenia siedliska (HMS). Metoda szacowania niepewności oceny stanu wód opisana w niniejszym artykule uwzględnia domyślnie przyczyny losowości wskaźników oraz ocen na wszystkich poziomach hierarchicznej struktury klasyfikacji (rys. 1), poprzez kwantyfikację niepewności związanych z oceną stanu chemicznego i wszystkich elementów stanu ekologicznego, łącznie z hydromorfologią.

Specyficznym źródłem niepewności jest sama procedura klasyfikacji, zawierająca w sobie ocenę ekspercką granic klas. Niepewność wprowadzają eksperci, którzy na podstawie swego dotychczasowego doświadczenia określają wartości granic poszczególnych klas wskaźników, za pomocą których dokonywana jest (ma być) ocena stanu JCW. Coraz większa liczba danych pomiarowych elementów biologicznych pozwala mieć do tych decyzji eksperckich coraz większe zaufanie. Granice klas, zwłaszcza najistotniejsze z punktu widzenia konsekwencji w odniesieniu do decyzji wodno-gospodarczych, np. granica pomiędzy stanem

dobrym i umiarkowanym, są w kolejnych przybliżeniach coraz lepiej udokumentowane wiedzą ekspertów i obserwacjami systemów wodnych. Z tego względu niepewność wprowadzana przez ekspertów nie była rozważana. Naturalnie decyzje eksperckie, czy granice klas odnosić się będą do 90. czy 70. percentyla z wartości pomierzonych czy też do innej statystyki, pozostają nadal decyzją arbitralną. Na wyższych poziomach hierarchicznej struktury procedur klasyfikujących stan JCW pojawiają się kolejne przyczyny, w przypadku których ocena stanu JCW charakteryzuje się losowością – jej źródłem jest stosowanie zasady „decyduje najgorszy”, która przypisuje ocenie wyższego rzędu miarę niepewności tej zmiennej (lub oceny) niższego rzędu, która decyduje o przynależności klasowej oceny wyższego rzędu.

W niniejszym artykule założono, że niepewności wynikające z przyczyn pierwotnych są znacznie mniejsze niż te, które są wynikiem niepewności wtórnej, tj. agregacji danych. Założenie to wynika z obserwacji, że sieć monitoringu wód płynących, w porównaniu z całą siecią rzeczną, nie charakteryzuje się dużą gęstością przestrzenną, a częstość pomiarów rzadko przekracza kilkanaście w skali roku, co przemawia za niewielką reprezentatywnością zmienności przestrzennej oraz czasowej zachodzących procesów kształtujących stan wód. Niepewność pierwotna, mająca źródło np. w losowym charakterze procesów, może zostać wyznaczona w postaci odchylenia standardowego wartości poszczególnych wskaźników i uwzględniona w oszacowaniach niepewności wtórnej na poziomie I metody hierarchicznej. Ta z kolei stanowi punkt wyjścia do oszacowania niepewności na wyższych poziomach, aż do oszacowania niepewności oceny stanu JCW na poziomie IV.

Ocena stanu jednolitych części wód (JCW) na podstawie danych monitoringowych

Przedstawiony poniżej algorytm oceny stanu JCW zakłada, w okresie, w którym JCW podlega ocenie, wykonano wszystkie pomiary wskaźników wynikające z rozporządzenia monitoringowego [11]. Przyjmuje się zatem, że dostępne są wartości pomiarowe wszystkich wskaźników we wszystkich grupach elementów jakości wody, tj. wskaźników reprezentujących elementy jakości biologicznej, fizyczno-chemicznej, hydromorfologicznej wchodzące w skład oceny stanu ekologicznego oraz wszystkie wskaźniki wchodzące w skład oceny stanu chemicznego.

Niech w przypadku każdego wskaźnika należącego do grupy wskaźników elementu 'w' dostępną będzie seria pomiarowa w postaci:

$$x_{i1}^w, x_{i2}^w, x_{i3}^w, \dots, x_{in_{iw}}^w \quad (1)$$

w której:

i – i-ty wskaźnik w danej grupie elementów 'w' (i=1, ..., N_w)

N_w – liczba wskaźników należących do grupy 'w'

n_{iw} – liczba pomiarów i-tego wskaźnika należącego do grupy 'w' w czasie oceny (w dalszej części pracy przyjęto uproszczone oznaczenie: n_{iw}=n).

Na podstawie takiej serii pomiarów można wyznaczyć estymator wartości średniej danego wskaźnika 'i', a następnie przypisać go do JCW, w której zostały wykonane pomiary. Porównanie wartości średniej danego wskaźnika z wartościami granic klas [2] prowadzi do określenia przynależności tego wskaźnika do pewnej klasy c_i^w. Postępując tak w przypadku wszystkich grup wskaźników 'w' oraz naśladując hierarchiczną strukturę oceny stanu JCW można sformułować algorytm oceny stanu JCW.

Algorytm oceny stanu JCW

– Obliczenie estymatorów wartości średnich wszystkich wskaźników wchodzących w skład grupy 'w' na podstawie pomiarów monitoringowych:

$$\bar{x}_i^w = \frac{1}{n} \sum_{j=1}^n x_{ij}^w \quad (i=1, \dots, N_w; j=1, \dots, n) \quad (2)$$

– Wyznaczenie przynależności do klasy i-tego wskaźnika na podstawie estymatora wartości średniej wskaźnika obliczonego z zależności (2). Granice klas zdefiniowane są w rozporządzeniu klasyfikacyjnym [13]. Tak więc, np. w przypadku pięciu wskaźników jakości biologicznej wyznaczonych jest pięć klas, wskazujących na przynależność każdego ze wskaźników do przedziału definiującego klasę odpowiadającą obliczonej wartości estymatora jego wartości średniej. W odróżnieniu od innych grup elementów jakości, wskaźniki biologiczne przekształcane są do postaci bezwymiarowego wskaźnika jakości ekologicznej (EQR – ecological quality ratio). Klasy zdefiniowane są w rozporządzeniu [13], poprzez podanie ich granicy górnej U_i^w(k) oraz dolnej L_i^w(k); (k=1, ..., 5). W przypadku klas skrajnych, przedziały definiujące klasy są zwykle otwarte, tzn. klasa najwyższa nie ma zdefiniowanej granicy górnej, natomiast klasa najniższa (wskazująca na najgorszy stan) nie ma zdefiniowanej granicy dolnej. Numer klasy c_i^w, do której należy JCW ze względu na i-ty wskaźnik typu 'w', wskazywany jest na drodze przeszukiwania klas. Ostatecznie przyjmowany jest ten numer klasy, w przypadku której spełniona jest nierówność:

$$L_i^w(c_i^w) \leq \bar{x}_i^w < U_k^w(c_i^w); \quad (i=1, \dots, N_w) \quad (3)$$

Ten krok algorytmu odpowiada zakończeniu I etapu hierarchicznej struktury oceny JCW (rys. 1).

– Wskazanie oceny JCW (e^w) w odniesieniu do grupy wskaźników 'w' elementów jakości na podstawie znajomości przynależności do danej klasy ze wskaźników grupy 'w' odbywa się poprzez zastosowanie zasady „decyduje najgorszy”:

$$e^w = \min \{c_i^w\}; \quad (i=1, \dots, N_w) \quad (4)$$

Ten krok algorytmu odpowiada zakończeniu II etapu hierarchicznej struktury oceny JCW (rys. 1).

– Zastosowanie zasady „decyduje najgorszy” do grup ocen zestawionych w dwie „super-grupy” – grupę ocen do oceny stanu ekologicznego {e^b, e^{fc}, e^h}:

$$e^e = \min \{e^b, e^{fc}, e^h\} \quad (5)$$

oraz jednoelementową grupę ocen zawierającą jedynie ocenę stanu chemicznego {e^{ch}}.

Ten krok algorytmu odpowiada zakończeniu III etapu hierarchicznej struktury oceny JCW (rys. 1).

Kolejno, zależnie od wyniku pierwszego etapu, sprawdza się elementy fizyczno-chemiczne. W przypadku, gdy chociaż jeden z danej grupy wskaźników (elementy biologiczne, elementy fizyczno-chemiczne) wskazuje na gorszą klasę stanu ekologicznego niż pozostałe elementy, ustala się określoną przez ten wskaźnik klasę wód.

– Specjalny krok algorytmu przewidywany przez RDW uwzględnia także wyniki klasyfikacji substancji szczególnie szkodliwych dla środowiska wodnego. W przypadku, gdy wskaźniki biologiczne i/lub fizyczno-chemiczne wskazują na I klasę (najlepszą), ale co najmniej jeden wskaźnik substancji szczególnie szkodliwych dla środowiska

wodnego wskazuje klasę niższą niż II, klasa stanu ekologicznego w przypadku danej jednolitej części wód ustalana jest jako III.

– Wyznaczenie oceny stanu JCW przez zastosowanie zasady „decyduje najgorszy” do dwóch ocen (e^e oraz e^{ch}):

$$e^{JCW} = \min \{e^e, e^{ch}\} \quad (6)$$

Ten krok algorytmu kończy IV etap i cały proces oceny jednolitych części wód powierzchniowych.

Przypisanie stanowi JCW przynależności do określonej klasy (tj. wyznaczenie oceny JCW) jest zatem dobrze zdefiniowanym procesem obliczeniowym i technicznie nie stwarza trudności. Ponieważ jednak nie przybliżyła do rozwiązania zagadnienia kwantyfikacji niepewności obliczonej oceny JCW, dlatego w dalszej części artykułu omówiono dwie miary niepewności ocen pojedynczych wskaźników jakości wód – ufność i precyzja. Choć są one powszechnie stosowane w Europie, to jednak ich odniesienie do końcowej oceny JCW zgodnie z wymaganiami RDW nie jest już tak oczywiste. Dlatego przedstawiono propozycję, w jaki sposób można propagować niepewności ocen odniesione do pojedynczych wskaźników w górę hierarchicznej struktury ocen (rys. 1), aby ostatecznie podać ilościowe oszacowanie ufności i precyzji końcowej oceny JCW.

Miary niepewności oceny stanu wód

Oszacowanie miar niepewności oceny, tj. ufności i precyzji na najniższym poziomie procesu oceny, tj. na poziomie wyznaczania klas w przypadku pojedynczych wskaźników w każdej z grup ‘w’ elementów jakości możliwe jest wyłącznie w sytuacji, gdy statystyki pomiarów są wystarczająco liczne. Przedstawiony algorytm wyznaczania dwóch miar niepewności (ufności i precyzji) ocen wskaźników zakłada, że w ocenianym okresie wykonano w JCW dostatecznie wiele pomiarów wartości każdego ze wskaźników w każdej grupie ‘w’ elementów jakości wód. Algorytm wyznaczania miar niepewności ocen można przedstawić w trzech krokach:

♦ Pierwszą miarą niepewności – ufność – określa się przez podanie przedziału, w którym z założonym prawdopodobieństwem znajduje się prawdziwa wartości średnia i-tego wskaźnika. Miarę niepewności oceny ($\delta \bar{x}_i^w$) i-tego wskaźnika typu ‘w’ ($i=1, \dots, N_w$) oblicza się następująco:

– przyjmuje się założenie, że wartość danego wskaźnika może być opisana jako zmienna losowa o rozkładzie normalnym (jeśli wskaźnik nie jest zmienną losową o rozkładzie normalnym, należy go uprzednio znormalizować),

– zakłada się, że obliczony na podstawie pomiarów wartości danego wskaźnika estymator odchylenia standardowego $\hat{\sigma}_i^w$ jest bliski odchyleniu standardowemu σ_i^w w populacji ogólnej tego wskaźnika,

– szacuje się, że przy współczynniku ufności równym 90% prawdziwa wartość średnia wskaźnika znajduje się w przedziale:

$$\left(\bar{x}_i^w - 1,65 \frac{\sigma_i^w}{\sqrt{n_i}}, \bar{x}_i^w + 1,65 \frac{\sigma_i^w}{\sqrt{n_i}} \right) \quad (7)$$

Jeśli założenie o (przybliżonej) równości estymatora $\hat{\sigma}_i^w$ i odchylenia standardowego σ_i^w w populacji generalnej wartości danego wskaźnika nie jest spełnione (najczęściej dlatego, że wartość σ_i^w nie jest znana), wówczas zamiast wzoru (7) należy zastosować przedział ufności:

$$\left(\bar{x}_i^w - t_\alpha \frac{\hat{\sigma}_i^w}{\sqrt{n_i-1}}, \bar{x}_i^w + t_\alpha \frac{\hat{\sigma}_i^w}{\sqrt{n_i-1}} \right) \quad (8)$$

w którym t_α jest wartością kwantyla α rozkładu Studenta o n_i-1 stopniach swobody. Na przykład, w przypadku współczynnika ufności równego 0,90, wartość kwantyla przy $n_i=12$ wyniesie $t_\alpha=1,79$.

♦ Drugą miarą niepewności wartości średniej i-tego wskaźnika typu ‘w’ ($i=1, \dots, N_w$) jest precyzja, zdefiniowana jako połowa szerokości przedziału ufności:

$$\delta \bar{x}_i^w = 1,65 \frac{\sigma_i^w}{\sqrt{n_i}}, \quad (i=1, \dots, N_w) \quad (9)$$

lub w przypadku, gdy zachodzi przypadek opisany równaniem (8):

$$\delta \bar{x}_i^w = t_\alpha \frac{\hat{\sigma}_i^w}{\sqrt{n_i-1}}, \quad (i=1, \dots, N_w) \quad (10)$$

Aby uniezależnić wielkość precyzji od jednostek, w których mierzone są wskaźniki, proponuje się dokonanie standaryzacji poprzez podzielenie precyzji przez wartość średnią uzyskaną z pomiarów danego wskaźnika, czyli definiując precyzję standaryzowaną (wskaźnik zmienności) następująco:

$$P_i^w = \frac{\delta \bar{x}_i^w}{\bar{x}_i^w} \quad (11)$$

♦ Z kolei miarą precyzji klasyfikacji δe^w jednolitej części wód, w odniesieniu do grupy wskaźników ‘w’ (lub grupy ocen), jest precyzja standaryzowana wskaźnika (grupy ocen), który w danej grupie ‘w’ został sklasyfikowany najniższej:

$$P_{i_0}^w = \frac{\delta \bar{x}_{i_0}^w}{\bar{x}_{i_0}^w} \quad (12)$$

w której i_0 jest numerem wskaźnika (grupy ocen) należącego do najniższej klasy w grupie ‘w’ i który na mocy zasady „decyduje najgorszy” decyduje o tym, że grupie/super-grupie ‘w’ przypisano klasę, do której należy wskaźnik (grupa ocena) i_0 (czyli $e^w = cl_{i_0}^w$).

Ostatecznie, za niepewność klasyfikacji jednolitych części wód powierzchniowych przyjmuje się precyzję standaryzowaną obliczoną w wyniku realizacji tych trzech kroków. Klasyfikacja JCW, uzyskana na podstawie pomierzonych wartości wskaźników danego elementu jakości, jest tym lepsza (bliższa klasyfikacji prawdziwej) im niepewność klasyfikacji, a więc szerokość przedziału ufności, jest mniejsza. Duża wartość precyzji standaryzowanej wskazuje na dużą niepewność klasyfikacji, a w konsekwencji – dużą niepewność oceny stanu jakości wód.

Przykład wyznaczania ufności i precyzji oceny stanu JCW

Zaproponowaną hierarchiczną metodę szacowania ufności i precyzji przetestowano na danych monitoringowych jakości wód powierzchniowych dwóch województw – dolnośląskiego i pomorskiego [14]. Wyboru tych odległych od siebie regionów wodnych dokonano w celu porównania stanu wód należących do zlewni o zróżnicowanych charakterystykach hydromorfologicznych. Wykorzystano dane pomiarowe z lat 2006–2012, pochodzące z Państwowego Monitoringu Środowiska uzyskane z WIOŚ we Wrocławiu i Gdańsku. Obliczenia wykonano za pomocą pakietu

R (www.gnu.org/licenses/gpl-3.0.txt). Dane pomiarowe pochodziły zarówno z punktów monitoringu diagnostycznego, jak również operacyjnego. W tabeli 1 zebrano syntetyczne informacje o zbiorze danych, które wykorzystano w obliczeniach.

Tabela 1. Dane monitoringu, na podstawie których przeprowadzono ocenę stanu jednolitych części wód w województwach pomorskim i dolnośląskim

Table 1. Monitoring data serving as basis for uniform water bodies status assessment in Pomorskie and Lower Silesian voivodships

Dane monitoringu	Województwo	
	dolnośląskie	pomorskie
Liczba ppk*, w których wykonano pomiary monitoringowe	814	390
Liczba ppk*, w których wykonano pomiary elementów jakości biologicznej	265	166
Liczba ppk*, w których wykonano pomiary elementów jakości fizyczno-chemicznej	796	371
Liczba ppk*, w których wykonano pomiary elementów specyficznych	337	296
Liczba ppk*, w których wykonano pomiary elementów chemicznych (substancji priorytetowych)	219	229

*ppk – punkt pomiarowo-kontrolny

Ze względu na to, że pełne wdrożenie elementów jakości biologicznej do systemu monitoringu wód nastąpiło dopiero w latach 2007–2009, w analizowanym zbiorze danych wartości wskaźników jakości biologicznej wielu jednolitych części wód nie były dostępne. Tylko w nielicznych JCW wykonane zostały pomiary więcej niż jednego elementu jakości biologicznej, lecz także w następnych latach liczba pomiarów wskaźników elementów biologicznych nie przyrastała znacząco, gdyż przy planowaniu pomiarów wskaźników biologicznych, a następnie przy ocenie JCW na podstawie elementów jakości biologicznych, zakłada się małą zmienność struktury ekosystemów rzecznych w czasie, co jest uzasadnione długością cyklu życia organizmów. Szczególnie w przypadku makrofitów i makrobezkręgowców nie należy oczekiwać zmian wielkości populacji w czasie krótszym niż rok. Organizmy żywe charakteryzują się na ogół tolerancją na krótkotrwałe zmiany warunków środowiskowych, dlatego zmiany składu gatunkowego i liczebności populacji organizmów wodnych mogą nastąpić dopiero jako reakcja na długotrwałe zmiany warunków środowiskowych. Zgodnie z rozporządzeniem monitoringowym [11] badania wszystkich elementów jakości biologicznej wód wykonywane są tylko w ciągu jednego rocznego cyklu pomiarów w 6-letnim planie gospodarowania wodami na obszarze dorzecza. Jedynie w przypadku fitoplanktonu na roczny cykl badań składa się 6-krotny pomiar i na tej podstawie obliczana jest wartość wskaźnika fitoplanktonowego (IFPL) [15]. Istnieje więc niebezpieczeństwo, że ocena stanu ekologicznego, która dokonywana jest przede wszystkim na podstawie oceny biologicznej, może być obciążona głównie niepewnością właśnie tej oceny. Jednokrotnie wykonany pomiar w ciągu cyklu monitoringowego oznacza brak możliwości obliczenia wartości precyzji i niepewności na podstawie wskaźników innych niż fitoplankton. W latach 2006–2012 fitoplankton w analizowanych pomiarach monitoringowych

był klasyfikowany na podstawie ilości chlorofilu *a* (a nie jak obecnie – na podstawie wskaźnika fitoplanktonowego), przy czym jego zawartości (wartości średnie) były interpretowane na podstawie rozporządzenia Ministra Środowiska z 20 sierpnia 2008 r. (staus: uchylony). W przypadku wskaźników odnoszących się do innych elementów jakości biologicznej, a więc makrofitowego wskaźnika rzeczno (MIR) oraz wskaźnika okrzemkowego (IO), gdy dostępny był tylko jeden czy dwa pomiary, nie mogły być podstawą do wyznaczenia niepewności oceny tych wskaźników.

Przykładowe wyniki klasyfikacji i oceny stanu ekologicznego wód wraz z oszacowaną ufnością tej oceny w przypadku kilku rzek województwa pomorskiego zestawiono w tabeli 2. Z przeprowadzonych obliczeń wynika, że miara niepewności wskaźników, a tym samym niepewność klasyfikacji, przyjmowała wartości w bardzo szerokim zakresie od 0 do 6,19 (niektóre wartości były szczególnie wysokie). W celu łatwiejszej analizy uzyskanych wyników zastosowano oznaczenia przynależności do jednego z czterech zakresów niepewności ocen:

- 1: $0,0 \div 0,3$,
- 2: $0,3 \div 0,5$,
- 3: $0,5 \div 1,0$,
- 4: $>1,0$.

Wartości niepewności wskaźników wchodzących do oceny stanu ekologicznego, przytoczone przykładowo w tabeli 3 wskazują, że większość wskaźników fizyczno-chemicznych charakteryzowała się niepewnością poniżej 0,5. Niepewność oceny stanu ekologicznego wskazuje, że w województwie dolnośląskim ponad 56%, a pomorskim ponad 66% (w obu przypadkach większość ocen) była mniejsza niż 0,5, co można uznać jako wartość akceptowalną. W przypadku oceny stanu chemicznego wód, w żadnym punkcie pomiarowym nie były dostępne pomiary wartości wszystkich wskaźników, koniecznych do dokonania oceny, a tym samym do oszacowania niepewności oceny. Dostępne dane wskazywały, że pomiary wielu wskaźników wykonane były z bardzo małą częstością i niejednokrotnie brak było pomiaru jakiegokolwiek wskaźnika stanu chemicznego. W wielu przypadkach wskaźniki stanu chemicznego mierzone były tylko jedno- albo dwukrotnie, uniemożliwiając obliczenie przedziału ufności w przypadku danego wskaźnika, a w konsekwencji uniemożliwiając oszacowanie niepewności oceny.

Oceny stanu chemicznego wód dokonano przyjmując założenie, że wskaźniki, które nie były mierzone reprezentowały stan dobry. Taki sposób postępowania prowadził do uzyskania oceny stanu nie gorszej, niż gdyby dysponowano pomiarami większej liczby wskaźników. W przypadku, gdy wynik oceny dokonanej w oparciu o mniejszy niż wymagany zestaw substancji priorytetowych wskazywał na zły stan wód, rzeczywisty stan, a więc taki, który zostałby stwierdzony na podstawie pełnego zestawu wskaźników, byłby również zły. Natomiast w przypadku, gdy wynik oceny stanu jakości wód, dokonanej na podstawie mniejszego niż obowiązujący zestaw wskaźników, wskazywał na ich dobry stan, wówczas nie można wykluczyć, że uwzględnienie innych (niepomierzonych) wskaźników może prowadzić do oceny niższej. Naturalnie ocena stanu chemicznego wód uzyskana na podstawie mniejszego zestawu wskaźników niż wymagany w rozporządzeniu [13], powinna być postrzegana jako niepewna, bez względu na wariancję uzyskaną z pomiarów dostępnych wskaźników. W przypadku tak rozumianej niepewności trudno było jednak określić miarę ilościową.

Tabela 2. Klasyfikacja i ocena stanu ekologicznego wybranych rzek województwa pomorskiego oraz oszacowana niepewność tej oceny
 Table 2. Ecological status classification and assessment of selected rivers in Pomorskie voivodship and estimated uncertainty of that assessment

Numer JCW rzeka	Rok	Elementy biologiczne			Elementy fizyczno-chemiczne			Elementy specyficzne			Stan ekologiczny		Niepewność oceny wg wzoru (12)	Zakres niepew- ności
		klasa	wskaźnik	niepewność	klasa	wskaźnik	niepewność	klasa	wskaźnik	niepewność	klasa	wskaźnik		
PLRW20002247699 Łeba	2006	I	fitoplankton	0,0314		utleniałość	0,1108	glin	1,2202	utleniałość		0,111	0,0÷0,3	
	2007			0,0811		azot ogólny*	0,2740	fluorki	0,4584	azot ogólny*		0,274		
	2010	-	-	-	II	fosforany	0,1398	fenole	0,7206	fosforany	II	0,140		
	2011	I	fitoplankton	0,3349		utleniałość	0,1860	cynek	0,2912	utleniałość		0,186		
	2012	II	kilka wskaźników	-		OWO	0,3781	glin	0,4539	OWO		0,376		
PLRW6000174624 Pokrzywna	2006	I	fitoplankton	0,0811	I	azotany	0,3517	selen	0,9146	selen	I	0,915	0,5÷1,0	
	2007	-	-	-		zasadowość	2,1639		0,0574	zasadowość		2,164		
PLRW6000194639 Wieprza	2006	I	fitoplankton	0,0314	II	azot ogólny*	0,2048	glin	0,9051	azot ogólny*	II	0,205	0,0÷0,3	
	2007			0,0821		utleniałość	0,2704	fluorki	0,5312	utleniałość		0,270		
	2010	-	-	-	I	azot amonowy	0,4270	miedź	0,4605	miedź	I	0,460		
PLRW6000174619 Wieprza	2012	II	kilka wskaźników	-		utleniałość	0,5098	glin	0,9387	utleniałość		0,510	0,5÷1,0	
	2006	I	fitoplankton	0,0314	II	azot ogólny*	0,2048		0,9051	azot ogólny*	II	0,205		
	2007			0,0821		utleniałość	0,2704	fluorki	0,5312	utleniałość		0,270		
	2010	-	-	-	I	azot amonowy	0,4270	miedź	0,4605	miedź	I	0,460		
	2012	II	kilka wskaźników	-		utleniałość	0,5098	glin	0,9387	utleniałość	II	0,510		
PLRW20002129999 Wisła	2006				III	ChZT	0,1178	selen	3,9461	ChZT	III	0,118	0,0÷0,3	
	2007					fosforany	0,2385	kilka wskaźników		fosforany		0,239		
	2008	-	-	-		chlorki	0,9332	-	-	chlorki		0,933		
	2009				II	azot ogólny*	0,1106	kilka wskaźników		azot ogólny*	II	0,111		
	2010					fosforany	0,2074	cynek	0,4819	fosforany		0,207		
2011	II	MIR			fosforany	0,3075	fenole	1,1962	fosforany		0,307	0,3÷0,5		
2012	-	-	-			0,2497	miedź	0,4104			0,250			

* metoda Kjeldahla

MIR – makrofitowy wskaźnik rzeczny

OWO – ogólny węgiel organiczny

ChZT – chemiczne zapotrzebowanie na tlen

Tabela 3. Niepewność oceny stanu wód powierzchniowych w województwach pomorskim i dolnośląskim na podstawie danych z lat 2006–2012

Table 3. Uncertainty of surface water status assessment in Pomorskie and Lower Silesian voivodships based on data from the period 2006–2012

Zakres niepewności	Oceny elementów biologicznych		Oceny elementów fizyczno-chemicznych		Oceny elementów specyficznych		Oceny stanu ekologicznego	
	liczba	%	liczba	%	liczba	%	liczba	%
województwo pomorskie								
0,0÷0,3	43	93,5	179	48,4	40	15,0	166	46,4
0,3÷0,5	2	4,3	92	24,9	57	21,4	73	20,4
0,5÷1,0	1	2,2	76	20,5	73	27,3	86	24,0
>1	0	0,0	23	6,2	97	36,3	33	9,2
Suma	46	100	370	100	267	100	358	100
województwo dolnośląskie								
0,0÷0,3	64	97,0	258	32,4	46	12,2	212	30,2
0,3÷0,5	1	1,5	212	26,6	85	22,6	185	26,4
0,5÷1,0	0	0,0	244	30,7	138	36,6	203	28,9
>1	1	1,5	82	10,3	108	28,6	102	14,5
Suma	66	100	796	100	377	100	702	100

W pracy [14] przyjęto założenie, że wskaźniki, które nie były mierzone reprezentowały stan dobry. Zastosowanie tak liberalnego sposobu oceny doprowadziło do przypisania JCW w województwie pomorskim dobrego stanu 30% wód, a w województwie dolnośląskim 49% wód. Należy podkreślić, że ze względu na brak informacji o substancjach priorytetowych i innych zanieczyszczeniach z załącznika 9 [13], 41% JCW w województwie pomorskim i 62% JCW w województwie dolnośląskim nie mogło być w ogóle ocenione pod względem stanu chemicznego.

Analizując dane dotyczące wskaźników chemicznych, które były mierzone w danym punkcie pomiarowo-kontrolnym przynajmniej trzykrotnie w ciągu roku, obliczono niepewność oceny stanu chemicznego wód (tab. 4). W tym przypadku niepewność mniejszą niż 0,5 reprezentowała zdecydowana mniejszość ocen, niecałe 28% i 22%, odpowiednio w przypadku województw pomorskiego i dolnośląskiego.

Tabela 4. Niepewność oceny stanu chemicznego wód w województwach pomorskim i dolnośląskim
Table 4. Uncertainty of water chemical status assessment in Pomorskie and Lower Silesian voivodships

Zakres niepewności	Województwo			
	pomorskie		dolnośląskie	
	liczba	%	liczba	%
0,0÷0,3	30	17,1	27	12,3
0,3÷0,5	19	10,8	20	9,1
0,5÷1,0	49	27,8	78	35,6
>1	78	44,3	94	43,0
Suma	176	100	219	100

Na podstawie uzyskanych wyników można przyjąć, że jeśli klasyfikacja została dokonana za pomocą wskaźników o poziomie niepewności poniżej 0,5, to stan JCW został oceniony wiarygodnie i odzwierciedla stan rzeczywisty. Przytoczona wartość jest postulowaną propozycją wartości granicznej niepewności ocen jakości JCW.

Podsumowanie

Proponowana hierarchiczna metoda szacowania niepewności klasyfikacji stanu wód powierzchniowych, wraz z przyjętą definicją precyzji standaryzowanej, wyczerpująco wymagania Ramowej Dyrektywy Wodnej w zakresie określenia ufności i precyzji oceny jakości jednolitych części wód. Zastosowanie tej koncepcji w przypadku jednolitych części wód powierzchniowych w województwach pomorskim i dolnośląskim wskazuje, że może stać się ona przydatnym narzędziem umożliwiającym realistyczną interpretację wyników monitoringu i oceny stanu wód. Analiza występowania poszczególnych przedziałów miary niepewności w ocenie stanu ekologicznego wód wskazuje, że oceny przeważającej większości jednolitych części wód w obu województwach charakteryzowały się niewielką niepewnością w zakresie 0,0÷0,3, a ponad połowa ocen wszystkich JCW w obu województwach charakteryzowała się niepewnością poniżej 0,5. Taki wynik dobrze świadczy o jakości danych pomiarowych uzyskiwanych w polskim systemie monitoringu wód.

Miara niepewności – precyzja standaryzowana, obliczona w przypadku wskaźników i ocen stanu chemicznego wód przyjmowała duże wartości, co wynikało z uwarunkowań logistycznych i ekonomicznych systemu monitoringu wód (liczba mierzonych wskaźników stanu chemicznego była niewielka, a same pomiary drogie; ponadto część z nich reprezentowana była pojedynczymi pomiarami w odstępach jednorocznych). Wydaje się, że ze względu na brak statystycznie znaczącej liczby pomiarów wymaganych substancji priorytetowych, tę część oceny stanu jednolitych części wód należy traktować wyłącznie jako przybliżenie stanu faktycznego, a nie precyzyjną ocenę. W tym przypadku miara niepewności wyliczona na podstawie tak małej liczby danych miała znaczenie drugorzędne. Problem szacowania niepewności ocen wynikających z pomiarów elementów jakości biologicznej wód wymaga odrębnego podejścia.

LITERATURA

1. Dyrektywa 2000/60/WE Parlamentu Europejskiego i Rady z 23 października 2000 r. ustanawiająca ramy wspólnotowego działania w dziedzinie polityki wodnej (Ramowa Dyrektywa Wodna). Dz. U. UE L 327 z 22 grudnia 2000 r.
2. Rozporządzenie Ministra Środowiska z 22 października 2014 r. w sprawie sposobu klasyfikacji stanu jednolitych części wód powierzchniowych oraz środowiskowych norm jakości dla substancji priorytetowych. Dz. U. z 30 października 2014 r., poz. 1482.
3. Common Implementation Strategy for the Water Framework Directive (2000/60/EC). Guidance Document No. 13. Overall Approach to the Classification of Ecological Status and Ecological Potential. European Communities, Luxembourg 2005.
4. J. ELLIS, V. ADRAENSSENS: Uncertainty Estimation for Monitoring Results by WFD Biological Classification Tools. WFD Report. Environment Agency, Bristol 2006.
5. R. T. CLARKE, D. HERING: Errors and uncertainty in bioassessment methods-major results and conclusions from the STAR project and their application using STARBUGS. *Hydrobiologia* 2006, Vol. 566, pp. 433–439.
6. R. T. CLARKE, A. LORENZ, L. SANDIN, A. SCHMIDT-KLOIBER, J. STRACKBEIN, N. T. KNEEBONE, P. HAASE: Effects of sampling and sub-sampling variation using the STAR-AQUEM sampling protocol on the precision of macroinvertebrate metrics. *Hydrobiologia* 2006, Vol. 566, pp. 441–459.
7. R. T. CLARKE: Estimating of European WFD ecological status class and WISER Bioassessment Uncertainty Guidance Software (WISERBUGS). *Hydrobiologia* 2013, Vol. 704, pp. 39–56.
8. B. DUDLEY, M. DUNBAR, E. PENNING, A. KOLADA, S. HELLSTEN, A. OGGIONI, V. BERTRIN, F. ECKE, M. SONDERGAARD: Measurements of uncertainty in macrophyte metrics used to assess European lake water quality. *Hydrobiologia* 2013, Vol. 704, pp. 179–191.
9. M. LOGA: Hierarchical approach to water body status misclassification. In: C. A. BREBBIA [Ed.]: Water Pollution XI, WIT Press, Ashurst 2012, pp. 97–109.
10. W. S. INGERSOLL: Environmental Analytical Measurement Uncertainty Estimation. Nested Hierarchical Approach. Defense Technical Information Center, 2001.
11. Rozporządzenie Ministra Środowiska z 15 listopada 2011 r. w sprawie form i sposobu prowadzenia monitoringu jednolitych części wód powierzchniowych i podziemnych. Dz. U. nr 258, poz. 1550.
12. K. SZOSZKIEWICZ, T. ZGOŁA, S. JUSIK, B. HRYC-JUSIK, F. H. DAWSON, P. RAVEN: Hydromorfologiczna ocena wód płynących. Podręcznik do badań terenowych według metody River Habitat Survey. Bogucki Wydawnictwo Naukowe, Poznań–Warrington 2007.
13. Rozporządzenie Ministra Środowiska z 9 listopada 2011 r. w sprawie sposobu klasyfikacji stanu jednolitych części wód powierzchniowych oraz środowiskowych norm jakości dla substancji priorytetowych. Dz. U. nr 257, poz. 1545 (status: uchylony).
14. A. WIERZCHOŁOWSKA: Zastosowanie hierarchicznej metody szacowania precyzji oceny. Praca dyplomowa. Politechnika Warszawska, Wydział Inżynierii Środowiska, Warszawa 2014.
15. J. PICIŃSKA-FAŁTYNOWICZ, J. BŁACHUTA: Wytoczne metodyczne do przeprowadzania badań fitoplanktonu i oceny stanu ekologicznego rzek na jego podstawie. Główny Inspektorat Ochrony Środowiska, Warszawa 2012.

Loga, M. An Estimation of Uncertainty Measures, Confidence and Precision, of Surface Water Status Assessment. *Ochrona Środowiska* 2016, Vol. 38, No. 1, pp. 15–23.

Abstract: A proposal for estimation method of two uncertainty measures of water status assessment, confidence and precision, required by Water Framework Directive was discussed in the article. Uncertainty measures estimation process in terms of water quality indicators and groups of indicators, and then the ecological and chemical status assessment, involves a hierarchy of classification procedures. Standardized precision is a proposed uncertainty measure that enables propagation of uncertainty of the measured water quality indicator values toward the top of the hierarchical assessment structure. The proposed definitions of uncertainty measures and the way of their estimation were illustrated with an example based

on the selected water bodies employing the real measurement data from Polish surface water monitoring system in two voivodships: Pomorskie and Lower Silesian. For most of the water bodies in the two voivodships the prevalence analysis of particular uncertainty intervals in water status assessment indicated only small uncertainty (within the range of 0.0–0.3) in the ecological status, while in over half of the assessments it was below 0.5. Due to the lack of statistically significant number of priority substance measurements for the chemical status assessment that part of water bodies assessment should be considered as an approximation only and not an accurate assessment.

Keywords: Water Framework Directive, water status assessment, classification uncertainty, classification confidence and precision, hierarchical assessment structure.