

11. Podstawowe założenia systemu kontroli i weryfikacji danych pomiarowych i analitycznych w Zintegrowanym Monitoringu Środowiska Przyrodniczego

Schemat systemu gromadzenia, przepływu i kontroli danych ZMŚP przedstawiony jest na rycinie 11.1. Podstawowym założeniem systemu jest, że gromadzenie, kontrola i weryfikacja podstawowych (tzw. "surowych") danych pomiarowych i analitycznych ZMŚP dokonuje się na Stacjach Bazowych, w Lokalnych Bazach Danych. Do Centrum Informatycznego ZMŚP trafiają dane przetworzone (średnie, sumy, mody, ekstrema itp. dla stanowisk i okresów pomiarowych), które niezbędne są do analizy zmienności sezonowej i wieloletniej procesów, obliczania rocznych bilansów materii oraz jako wartości wejściowe i weryfikacyjne modeli. W Lokalnej Bazie Danych gromadzone są również archiwalne dane pomiarowe zawierające wyniki badań sprzed momentu powstania Stacji Bazowej a dotyczące zlewni ZMŚP, jej otoczenia i regionu, jak również materiały kartograficzne i zdjęcia lotnicze. Przy takim systemie organizacji gromadzenia i przepływu danych podstawą ich wiarygodności jest:

- przygotowanie prostych standardowych procedur kontrolno-weryfikacyjnych dla personelu Stacji Bazowych, a także,
 - ☐ okresowe inspekcje instalacji pomiarowych i laboratoriów oraz,
 - ✂ wewnątrzlaboratoryjna i międzylaboratoryjna kontrola jakości analiz,
 - ✂ zaawansowana weryfikacja zarówno danych podstawowych, jak i przetworzonych.

Realizacja pierwszego punktu opiera się na przygotowaniu przez specjalistów ZMŚP odpowiednich instrukcji dla personelu Stacji Bazowych i włączaniu zagadnień kontroli i weryfikacji danych do tematyki okresowych szkoleń (tzw. "Szkoły ZMŚP"). Instrukcje i szkolenia obejmują w tym przypadku głównie problem zapewnienia jakości danych na etapie obserwacji, pomiarów i analiz. Kontrola i weryfikacja na etapie gromadzenia i przetwarzania danych jest istotnym elementem corocznych szkoleń z zakresu rozwoju systemu informatycznego ZMŚP.

Drugi z wymienionych zakresów to głównie obszar działania zespołu specjalistów ZMŚP, lecz także zapraszanych okazjonalnie ekspertów. Program na najbliższe 3 lata zakłada coroczne inspekcje instalacji pomiarowych i laboratoriów.

Podstawowym celem kontroli wewnątrzlaboratoryjnej jest utrzymanie jakości rutynowych operacji. Umożliwia to obniżenie wielkości błędów systematycznych, dominujących przy analizie danych wynikających z porównań międzylaboratoryjnych (ryc. 11.2). Każde laboratorium wykonujące analizy próbek pozyskanych w programie ZMŚP jest zobowiązane do realizacji minimalnego programu kontroli wewnątrzlaboratoryjnej obejmującego między innymi (Witczak, Adamczak 1994-1995):

- Po kalibracji, a przed rozpoczęciem analiz powtórzyć oznaczenie wzorca o najwyższym stężeniu traktując go jak normalną próbkę. Należy upewnić się czy stężenie zmierzone nie odbiega od rzeczywistego więcej niż +/- 5% (lub mniej zależnie od procedury przyjętej w laboratorium). Jeżeli występuje większa różnica należy usunąć przyczynę błędów i przeprowadzić ponowną kalibrację przyrządu.
- W trakcie serii analiz przepłukiwać system pomiarowy (analityczny) wzorcem zerowym (czysta próbka - nie zawierająca analizowanej substancji) przed każdą próbką.
- Co każde 10 próbek analizować wzorzec kontrolny przyrządu. Wzorzec zawierać musi pożądaną - ze względu na skład analizowanego materiału terenowego i typ stosowanego przyrządu - zestaw pierwiastków. Wzorzec kontrolny umożliwia wykrycie dryftu przyrządu. Jeśli pomiar wzorca wykazuje różnicę większą od +/- 5% od spodziewanego (lub mniej zależnie od procedury przyjętej w laboratorium), należy przerwać analizę, wyjaśnić przyczynę błędów i przeprowadzić ponowną kalibrację przyrządu.
- Co każde 10 próbek analizować wzorzec zerowy. Wynik powinien się mieścić w przedziale 2 odchyłeń standardowych od wartości średniej. Jeśli różnica jest większa należy powtórzyć analizę wzorca zerowego jeszcze dwukrotnie i uśrednić wszystkie rezultaty. W przypadku, gdy uzyskana średnia nie mieści się w ustalonym przedziale kontrolnym, należy przerwać analizę, wyjaśnić przyczynę błędów i ponownie wykalibrować instrument.
- Co każde 20 rutynowych próbek dokonywać powtórzeń (duplikacji) analiz dla kontroli precyzji ich wykonywania.
- Na początku i na końcu serii analiz oraz przy dłuższych seriach periodycznie co pewną ilość próbek poddać analizie wzorzec kontrolny interferencji, zawierający znane stężenie oznaczanej substancji i domieszki pierwiastków wpływających znacząco na wynik analizy. Rezultaty powinny się mieścić w ustalonym przedziale kontrolnym dwa

razy większym od odchylenia standardowego od wartości średniej. W przypadku, gdy uzyskany wynik nie mieści się w ustalonym przedziale kontrolnym, należy przerwać analizę, wyjaśnić przyczynę błędów i ponownie wykalibrować instrument.

- Co pewien ustalony okres (raz na kilka dni, raz na tydzień lub rzadziej) zależny od ilości wykonywanych serii analiz (i częstości kalibracji przyrządu) wykonać analizę próbki wzorcowej, uzyskanej ze źródła zewnętrznego, a zawierającej znane stężenia składników. Umożliwia to weryfikację roztworów wzorcowych. Jeśli wyniki analizy próbki wzorcowej różnią się więcej niż $\pm 5\%$ od spodziewanego, należy przygotować nowe roztwory wzorcowe i ponownie wykalibrować przyrząd. Jeśli i to nie poprawi dokładności należy przygotować lub zakupić nowe wzorce podstawowe, z nich sporządzić nowe roztwory wzorcowe i powtórzyć kalibrację przyrządu.

Stacje Bazowe włączone zostaną w najbliższym czasie, na równorzędnych zasadach, do programu międzylaboratoryjnej kontroli analiz prowadzonej przez PIOŚ.

Zakres czwarty obejmuje dwa poziomy: Lokalnej Bazy Danych w Stacji Bazowej i Centralnej Bazy Danych w Centrum Informatycznym ZMŚP, różnią się one jednakże stopniem szczegółowości wynikającym z operowania na danych "surowych" w przypadku pierwszym, a danych "przetworzonych" w przypadku drugim. Podstawowym warunkiem sprawnej realizacji systemu zaawansowanej kontroli danych na Stacjach Bazowych jest instalacja jednolitego, standardowego oprogramowania bazy danych, co warunkowane jest możliwościami finansowymi. Olbrzymie zróżnicowanie parametrów mierzonych w ZMŚP powoduje, że jedynie automatyzacja procesu weryfikacji danych zapewnia szybką i obiektywną identyfikację wartości błędnych, lub wątpliwych. System zaawansowanej kontroli danych na poziomie Centrum Informatycznego ZMŚP umożliwi identyfikację wartości błędnych jedynie wtedy, gdy analizowane będą wyniki poszczególnych pomiarów (podawane w niektórych programach pomiarowych). W przypadkach danych przetworzonych możliwa będzie jedynie selekcja tych wartości wątpliwych, gdzie grube błędy (o różnej genezie) wpływają w istotny sposób na odchylenie wartości średnich i sum okresowych lub przestrzennych. Odpowiednie procedury kontrolne są aktualnie opracowywane, a ich algorytmy będą wbudowane w bazę danych. Zidentyfikowane błędne i wątpliwe wyniki przekazywane będą z Centrum do Stacji Bazowej celem sprawdzenia i wyjaśnienia. W przypadku, kiedy błąd powstał na etapie gromadzenia i przetwarzania materiałów pomiarowych, skorygowane dane przesyłane są ponownie do CI ZMŚP. Jeśli błąd był efektem analityki laboratoryjnej, a materiał jej poddany jest przechowywany na Stacji Bazowej w "archiwum próbek środowiskowych" możliwe jest powtórzenie analizy w tym samym lub w w razie potrzeby innym laboratorium. Powtórzenie błędnego pomiaru w terenie nawet w przypadku, gdy jego częstotliwość wykonywania jest wieloletnia najczęściej nie jest możliwe. Sprawdzenie takich danych w Lokalnej Bazie umożliwi jednakże w większości przypadków jednoznaczną interpretację wartości wątpliwych.

System zaawansowanej kontroli danych obejmuje:

- kontrolę kodów identyfikujących program i punkt pomiarowy oraz datę pomiarów,
- kontrolę poprawności kodowania parametrów pomiarowych gleb, taksonów flory, fauny, typu i jakości danych itp.,
- kontrolę kompletności danych,
- kontrolę wartości ekstremalnych,
- kontrolę definicyjną i logiczną,
- kontrolę związków funkcyjnych,
- kontrolę przebiegów czasowych

Kontrolę kodów identyfikujących punkt pomiarowy oraz datę pomiarów przeprowadza się w stosunku do zakresu i frekwencji monitoringu przyjętych dla poszczególnych programów pomiarowych na konkretnej Stacji Bazowej. Inne kody zapisane w pliku z danymi źródłowymi porównuje się z zestawieniem wszystkich dopuszczalnych kodów dla danego zakresu (organizmy, gleby itp.). Umożliwia to głównie selekcję błędów powstałych na etapie przenoszenia danych z dokumentów źródłowych na magnetyczne lub optyczne nośniki informacji.

Kontrola wartości ekstremalnych polega na porównaniu zmierzonych wartości z maksymalnym możliwym ich zakresem dla danego stanowiska i daty pomiaru. Procedura taka umożliwia selekcję zarówno danych ewidentnie błędnych, jak i wątpliwych, ze względu na skrajnie małe prawdopodobieństwo wystąpienia zjawiska o zmierzonym natężeniu. Podstawowe znaczenie dla określenia kryteriów kontroli wartości ekstremalnych ma dostępność archiwalnych danych obserwacyjnych i pomiarowych.

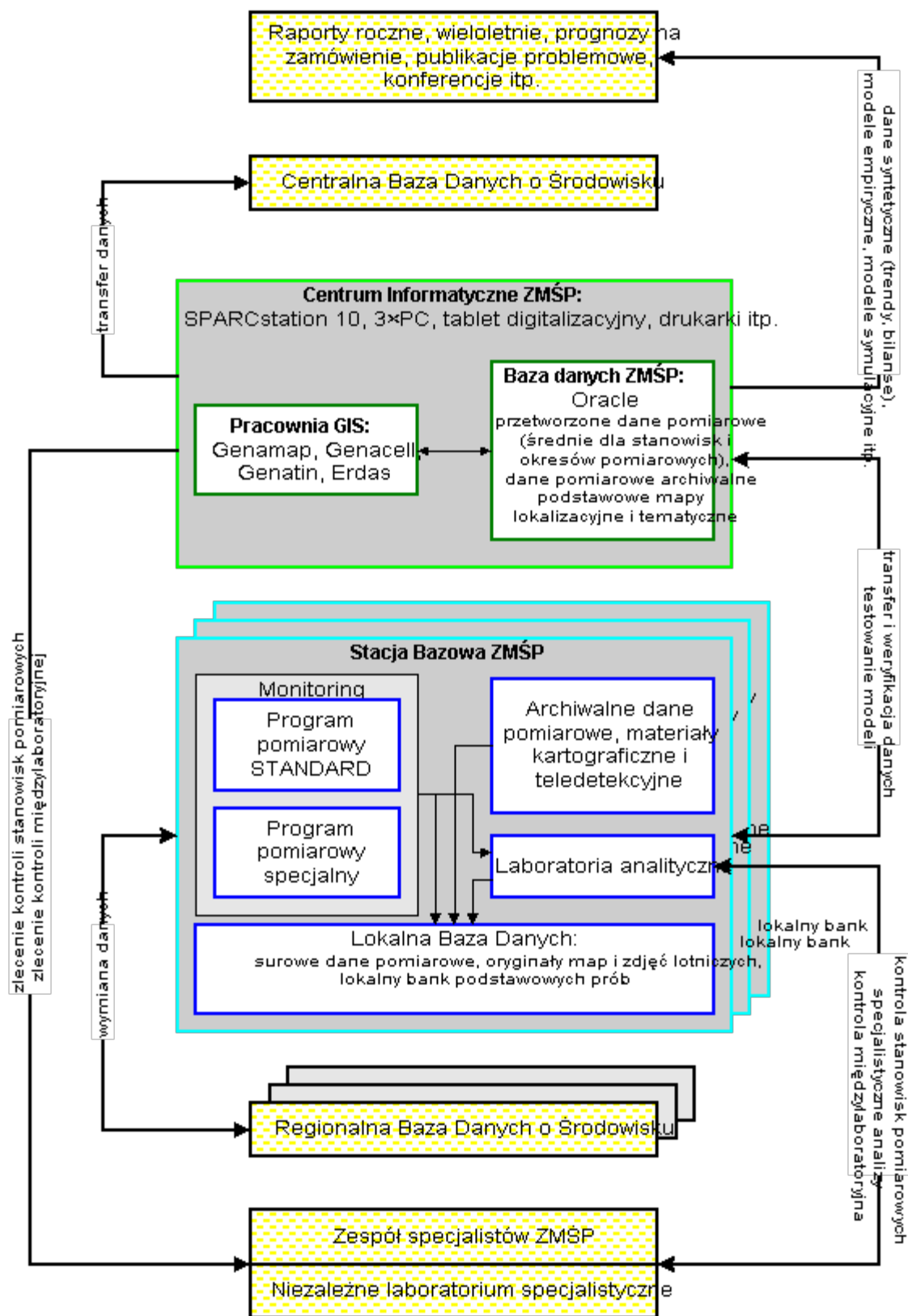
Kontrola związków definicyjnych i logicznych polega na znajomości współwystępowania, lub też wzajemnego wykluczania się niektórych zjawisk. Podawanie dla terminu pomiarowego parametrów charakteryzujących chemizm opadów, i jednocześnie wykazywanie braku opadu może być przykładem takich sprzeczności.

Kontrola związków funkcyjnych polega na wykorzystaniu równań opisujących fizyczne/chemiczne (deterministyczne), lub też empiryczne zależności pomiędzy mierzonymi parametrami.

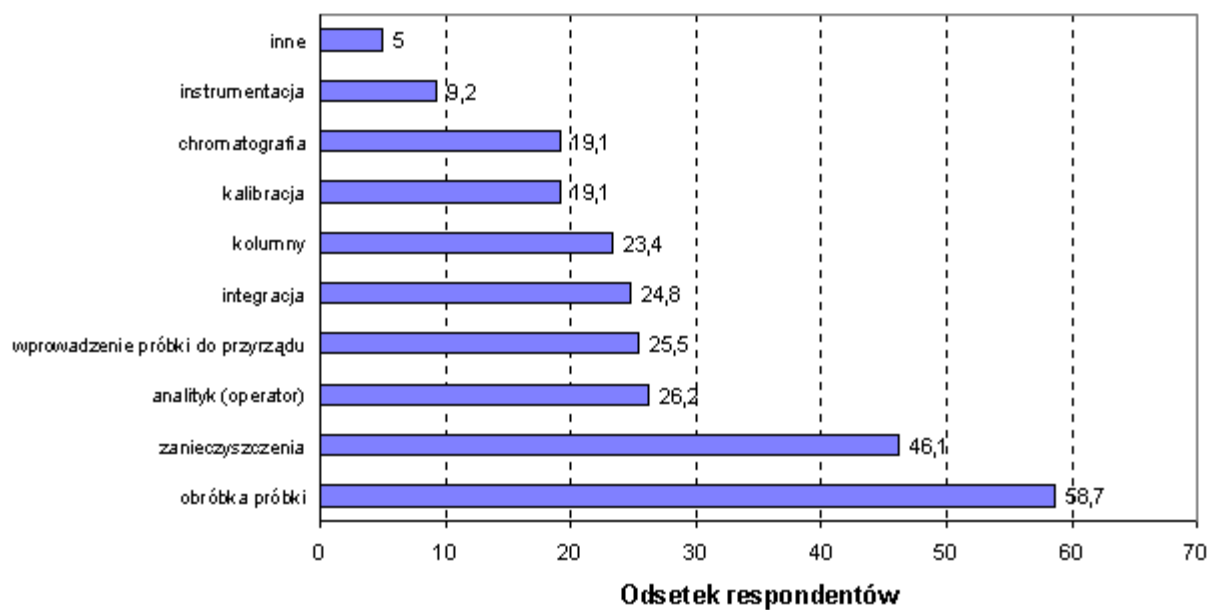
Kontrola przebiegów czasowych polega na wykorzystaniu faktu silnej autokorelacji kolejnych pomiarów niektórych parametrów. Zmiany są najczęściej jednokierunkowe, a ich gradient nie przekracza pewnych granic. Procedura opiera się najczęściej na porównaniu wartości zmierzonych z przebiegiem szacowanym w oparciu o funkcję wielomianową.

Oprócz procedur ogólnie stosowanych kontrola danych pomiarowych i analitycznych uwzględniać musi specyficzne uwarunkowania związane z lokalizacją zlewni, a także konkretnej powierzchni i stanowiska pomiarowego.

Kontrola danych na Stacjach Bazowych powinna dokonywać się w miesięcznym/rocznym odstępie. Weryfikacja w CI ZMŚP wykonywana jest w odstępie rocznym. W obu przypadkach uwzględniane muszą być całe, narastające zbiory danych. Po poprawieniu lub usunięciu wartości błędnej algorytm weryfikacji uruchamia się powtórnie, i dopiero przy braku sygnalizacji dalszych błędów (wartości wątpliwych) nowe dane mogą być dołączone do bazy.



Ryc. 11.1. Schemat systemu gromadzenia, kontroli i weryfikacji danych pomiarowych i analitycznych w ramach Zintegrowanego Monitoringu Środowiska Przyrodniczego.



Ryc. 11.2. Najczęściej występujące źródła błędów związane z etapem przygotowania próbki i analizą (Namieśnik i in. 1995).