

INBO-opdracht

Statistisch Analyse Plan

Jorre Vannieuwenhuyze

22 februari 2026

Case 1:

Om de kwaliteit van waterplassen te evalueren worden verschillende waterkwaliteitsparameters opgevolgd (Tabel 1). Gedurende één jaar wordt maandelijks een mengstaal van de waterplas genomen. In het laboratorium wordt elke parameter vervolgens in duplo geanalyseerd. Omdat het onderzoeksteam jaarlijks de waterkwaliteit van meerdere plassen moet beoordelen, wordt aan het team BMK gevraagd om het onderzoeksdesign voor de volgende studie (gericht op één plas) te optimaliseren.

Vorig jaar werden reeds verschillende plassen onderzocht. De resultaten daarvan zijn beschikbaar in het meegeleverde CSV-bestand case1.csv. Voor elke waterplas wordt het maandelijkse gemiddelde van elke parameter vergeleken met de wettelijke streef- en grenswaarden.

Het doel van deze studie is het optimaliseren van het onderzoeksdesign voor de jaarlijkse evaluatie van de waterkwaliteit van één waterplas. De evaluatie gebeurt op basis van vergelijking van maandelijkse metingen van verschillende waterkwaliteitsparameters met wettelijke streef- en grenswaarden.

De optimalisatie van het onderzoeksdesign zoekt naar een design waarin op voldoende betrouwbare manier kan worden vastgesteld of een plas de streef- of grenswaarde overschrijdt, met zo weinig mogelijk staalnames en labokosten. Hoe meer staalnames worden getrokken en hoe meer deze geanalyseerd worden, hoe betrouwbaarder de resultaten maar ook hoe hoger de kosten.

Dit analyseplan wordt opgebroken in vijf opbouwende stappen:

- Eerst worden verschillende variantiecomponenten gekwantificeerd (maand-tot-maand variatie versus laboratoriumvariatie) waardoor we kunnen bepalen of we voor bepaalde parameters meer metingen moeten uitvoeren en voor andere minder metingen;
- Ten tweede worden de data geanalyseerd als tijdsreeksen om tijdscomponenten in kaart te brengen, inclusief autocorrelatie en seizoenseffecten;
- Ten derde kijken we naar de samenhang tussen parameters. Wanneer bepaalde parameters sterk samenhangen kan de meting van bepaalde parameters verminderd worden omdat andere parameters al informatie bevatten;

- Ten vierde worden nieuwe metingen modelgebaseerd geëvalueerd tegen grenswaarden;
- Ten vijfde wordt adaptieve monitoring toegepast om het meetplan dynamisch bij te sturen gedurende het jaar.

In de onderstaande uitwerking gaan we ervan uit dat de microdata over elke individuele staalname en parametertest toegankelijk is. Het bijgeleverde bestand case1.csv bevat enkel informatie op geaggregeerd niveau, waardoor sommige beschreven analyses minder accuraat of zelfs niet uitvoerbaar zijn.

Stap 1: Analyse van testvariatie

Elke maand wordt een staal genomen uit de waterplas. Dat staal wordt in het labo twee keer geanalyseerd (duplo-analyse). Wanneer we verschillen zien tussen metingen, kunnen die uit verschillende bronnen komen:

- Echte verschillen tussen maanden, bijvoorbeeld door weersomstandigheden of biologische processen die zorgen voor een verbeterende of verslechterende waterkwaliteit;
- Kleine verschillen die ontstaan bij het nemen van het staal (bemonsteringsvariatie);
- Meetnauwkeurigheid in het labo (laboratoriumvariatie).

Met andere woorden, niet elk verschil in meetresultaat betekent dat de waterkwaliteit echt veranderd is. Een deel van de variatie is simpelweg toe te schrijven aan toeval bij staalname of analyse.

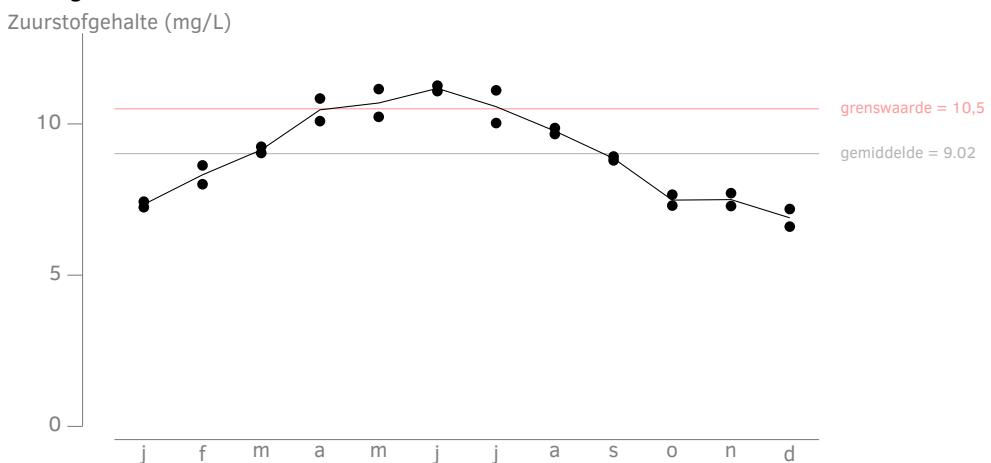
In deze stap onderzoeken we hoeveel van de totale variatie wordt veroorzaakt door de laboratoriumvariatie. Daarvoor berekenen we de zogenaamde Intraclass Correlation Coefficient (ICC). Dit geeft aan welk deel van de totale variatie toe te schrijven is aan echte verschillen tussen de maanden (inclusief natuurlijke variatie en staalnamevariatie) en hoeveel aan verschillen tussen de verschillende testen. Omdat voor elke plas slechts één staal per maand wordt genomen, kunnen natuurlijke maandvariatie en staalnamevariatie niet afzonderlijk worden onderscheiden zonder bijkomende aannames. Beide worden daarom in deze stap samen geïnterpreteerd als variatie tussen meetmomenten. De analyses in de volgende stappen bieden een antwoord op dit probleem.

De ICC kan al volgt worden geïnterpreteerd:

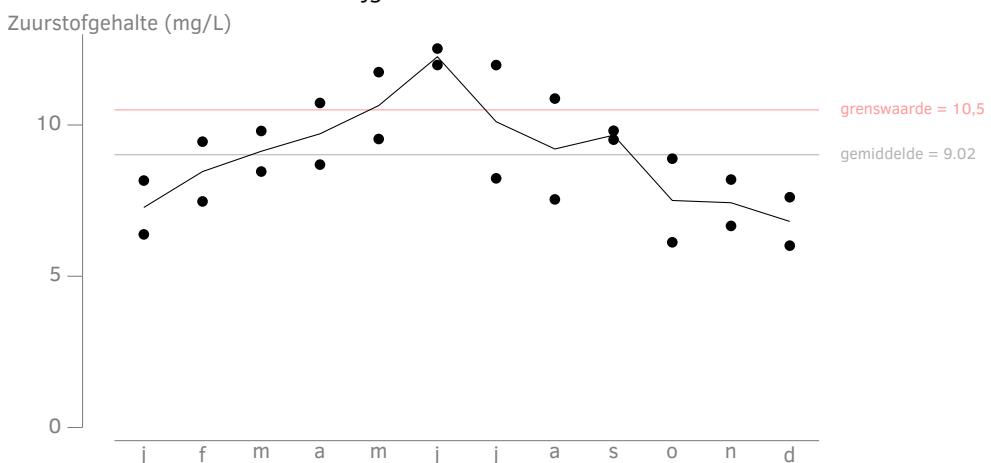
- Als de ICC hoog is (bijvoorbeeld hoger dan 0.80), betekent dit dat bijna alle variatie echte verschillen weerspiegelt tussen de meetstalen (zie Figuur 1). De bijdrage van meetfouten door de testen is dan beperkt, en duplo-analyses voegen weinig extra informatie toe.
- Als de ICC laag is (bijvoorbeeld lager dan 0.50), betekent dit dat een groot deel van de variatie afkomstig is van labo-onzekerheid (zie Figuur 2). In dat geval zijn duplo-analyses belangrijk om betrouwbare resultaten te verkrijgen. Het kan zelfs aangeraden zijn om het aantal analyses op te voeren per meetstaal.

Om de ICC te schatten wordt voor elke parameter een eenvoudig lineair mixed effects model met random maandeffecten. De variantie van de maand-

Figuur 1: Als de testresultaten voor dezelfde staalname relatief dicht tegen elkaar liggen is er een lage ICC en heeft herhaald testen minder nut.



Figuur 2: Als de testresultaten voor dezelfde staalname relatief ver van elkaar liggen is er een lage ICC en veel onzekerheid over de echte waarden. Herhaald testen is dan erg nuttig om betrouwbaardere resultaten te krijgen.



effecten geven de echte variatie tussen de meetstalen weer, terwijl de variatie van de residuelen de variatie weergeeft ontstaan uit de testen. Dit model laat toe om objectief te beoordelen of dubbele labo-analyses noodzakelijk zijn of uitgebreid mogen worden. Het model zorgt er ook voor dat

- een geobserveerde stijging of daling tussen de metingen van twee opvolgende maanden met grotere betrouwbaarheid kunnen worden geschat. De standaardfout van zo'n stijging of daling kan accurater worden geschat.
- de kans op een overschrijding van drempelwaarden door louter willekeurige fluctuaties in de meetresultaten accurater berekend kan worden.

Via een poweranalyse kan bovendien het ideale aantal metingen worden vastgesteld zodat een overschrijding van de grenswaarde met een bepaalde zekerheid kan worden vastgesteld.

Stap 2: Analyse van seizoenseffecten en autocorrelatie

De geschatte maandeffecten in stap 1 kunnen seizoenpatronen vertonen. Die uiten zich als gelijkaardige waarden in dezelfde maand van opeenvolgende jaren. Indien blijkt dat sommige maanden duidelijk risicotieker zijn dan andere, kan het meetplan worden aangepast. Resultaten kunnen bijvoorbeeld tonen dat we beter vaker meten in de zomer en minder frequent in stabiele wintermaanden. Zo optimaliseren we de inspanningen waar het risico het grootst is en kunnen we vroegtijdig trends en verslechtering signaleren.

Bovendien vormen waterkwaliteitsmetingen een tijdreeks. De metingen volgen elkaar op in de tijd en zijn daardoor niet onafhankelijk. In de praktijk betekent dit dat de waarde in een bepaalde maand vaak samenhangt met de waarde in de maand ervoor. Bijvoorbeeld, als het zuurstofgehalte in juni hoog is, is de kans groot dat het in juli ook relatief hoog blijft. Dit fenomeen noemen we *autocorrelatie*, de mate waarin een meting lijkt op eerdere metingen. Wanneer autocorrelatie aanwezig is, betekent dit dat de waterkwaliteit een zekere "traagheid" vertoont. Dit is typisch voor milieudata. Seizoenspatronen kunnen zich bovendien tonen als een specifiek type autocorrelatie. Gelijkenis tussen dezelfde maanden in opeenvolgende jaren wijst op een terugkerend seizoenspatroon. Dit wordt geïllustreerd door figuren 3 en 4 waarin duidelijke seizoenseffecten waarneembaar zijn.

Autocorrelatie kunnen we analyseren op twee manieren. Enerzijds kunnen we covariantie toelaten in het model tussen de residuelen ten opzichte van het algemeen gemiddelde. Zo'n varianties kunnen heel flexibel worden gemaakteerd, bijvoorbeeld door hoge covarianties toe te laten tussen metingen van opeenvolgende maanden terwijl metingen met grotere tijdsintervallen zwak of niet gecorreleerd zijn. Anderzijds kunnen we ook gebruik maken van tijdreeksmodellen waarin het effect van een vorig meetresultaat expliciet wordt gemaakteerd. Opnieuw kan dit op een flexibele manier worden geïmplementeerd door enkel het effect van de vorige of twee vorige testresultaten toe te voegen.

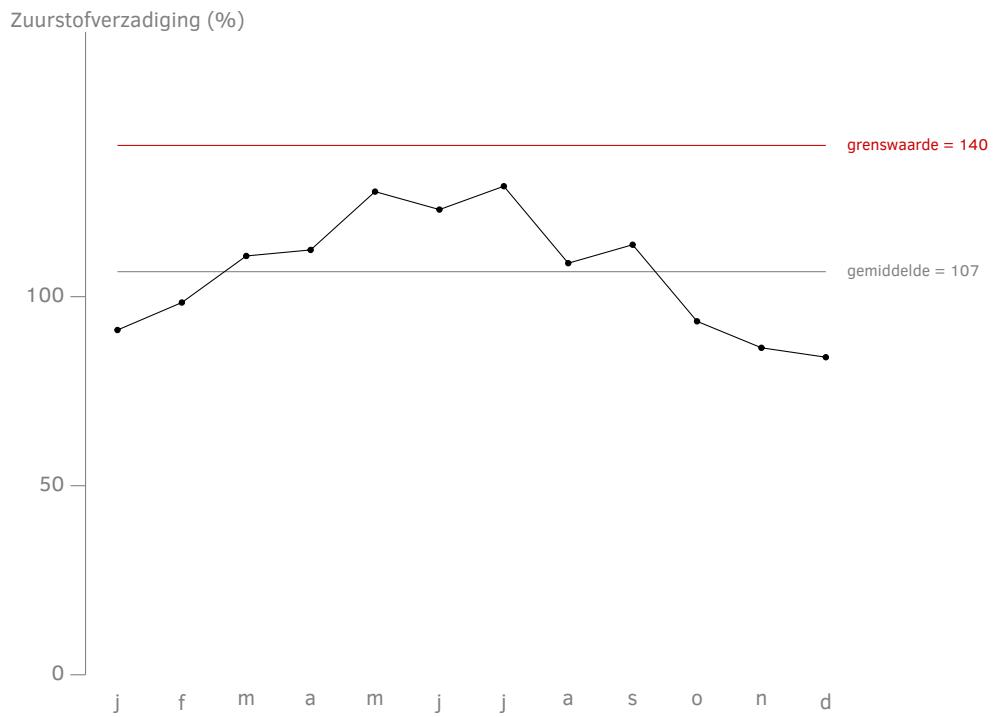
Door autocorrelatie te analyseren, kunnen we opnieuw het onderzoeksdesign optimaliseren. Als er sterke autocorrelatie aanwezig is, bevat elke meting gedeeltelijk informatie over de volgende meting, en kan daardoor een lagere meetfrequentie soms volstaan zonder veel informatieverlies. Tegelijkertijd betekent dit echter ook dat metingen minder betrouwbaar zijn omdat er minder informatie in de data aanwezig is. Als er echter weinig autocorrelatie aanwezig is fluctueert de waterkwaliteit sneller en minder voorspelbaar en is frequenter meten aangewezen.

Het modelleren van autocorrelatie leidt tot correctere onzekerheidsinschattingen en daardoor tot beter onderbouwde beslissingen over meetfrequentie. Via een poweranalyse kan opnieuw het ideale aantal metingen worden vastgesteld in verschillende maanden doorheen het jaar.

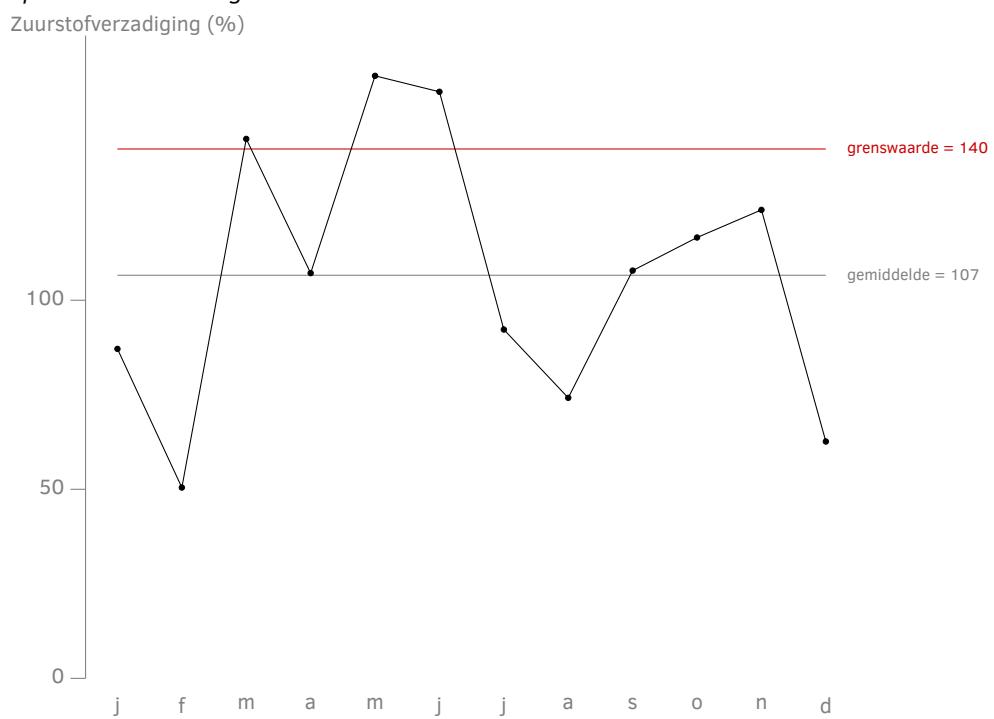
Stap 3: Samenhang tussen waterkwaliteitsparameters

Verschillende waterkwaliteitsparameters hangen vaak samen. Figuur 5 toont bijvoorbeeld een sterke relatie tussen de pH-waarde en de zuurstofverzadiging. Dit betekent dat metingen van de ene parameter informatie bevatten

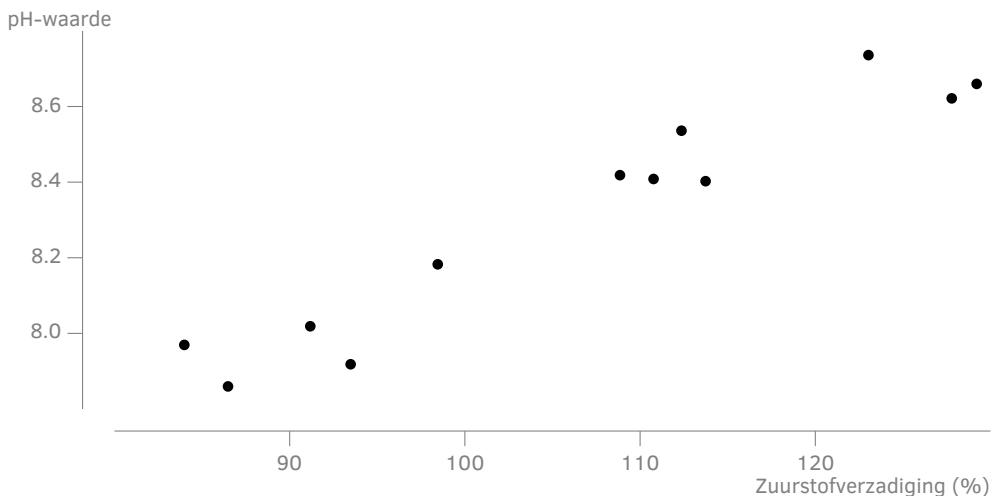
Figuur 3: Testresultaten met hoge autocorrelatie. Elk resultaat lijkt sterk op het vorig gemeten resultaat.



Figuur 4: Testresultaten met lage autocorrelatie. Een resultaat kan moeilijk voorspeld worden op basis van het vorige resultaat.



Figuur 5: Er is een correlatie tussen de pH-waarde en de zuurstofverzadiging.



over de andere parameter. Zo'n sterke en stabiele correlaties kunnen benut worden om duurdere parameters minder frequent te meten maar te schatten op basis van goedkopere parameters. Anderzijds kan correlatie tussen parameters gebruikt worden om betrouwbaardere schattingen te krijgen van waterkwaliteit omdat de samenhang zorgt voor een hogere statistische power.

Correlaties tussen parameters kunnen gemodelleerd worden met regressiemodellen voor multivariate uitkomsten. De resultaten van deze modellen kunnen opnieuw gebruikt worden om het onderzoeksdesign te optimaliseren. Bijvoorbeeld, goedkope parameters kunnen maandelijks worden geregistreerd terwijl duurdere parameters minder frequent (bv. per kwartaal) worden geregistreerd. In dit scenario moet er wel een jaarlijkse controle blijven van correlatiestabiliteit. Indien blijkt dat de parameters minder sterk samenhangen dan verwacht, kan tijdelijk opnieuw een hogere meetfrequentie worden ingevoerd.

Stap 4: Modelgebaseerde evaluatie van nieuwe data

De bovenstaande stappen worden gebruikt om aan datamodel te schatten op basis van historische data waarin rekening wordt gehouden met variabiliteit en tijdsstructuur. Dit model kan vervolgens gebruikt worden om de betrouwbaarheid van nieuwe observaties in een waterplaats te evalueren gegeven het tijdstip van die meting. Nieuwe metingen kunnen immers de grenswaarden overschrijden door toevallige schommelingen tijdens staalname en labo-analyse. Als de overschrijding niet statistisch significant is, kan niet met zekerheid een overschrijding worden geconcludeerd en is voorzichtigheid geboden.

Zo'n evaluatie gebeurt door predictie-intervallen te schatten. Zo'n predictieinterval geeft het statistisch bereik waarin je met een bepaalde zekerheid verwacht dat een toekomstige observatie zal vallen. Hierdoor kwantificeert het de onzekerheid van nieuwe metingen. Wanneer nieuwe observaties buiten de predictie-intervallen vallen is dat een signaal voor mogelijke veranderingen in waterkwaliteit. Dit geldt ook voor observaties die de grenswaarden

niet overschrijden of periodes waarin grenswaarden doorgaans niet overschreden worden.

Stap 5: Adaptieve monitoring

Het meetplan hoeft niet volledig vast te liggen. We kunnen er ook voor kiezen om het plan doorheen het jaar dynamisch aan te passen op basis van vooraf bepaalde criteria of waarschuwingssignalen. Mogelijk signalen zijn bijvoorbeeld

- een significante stijgende of dalende trend in een parameter,
- abnormale afwijkingen ten opzichte van historische patronen,
- een meting dicht bij een grenswaarde, of
- een kans op overschrijding boven een drempel (bv. $P > 0.30$).

In dit soort situaties kunnen verschillende beslissingen worden genomen zoals

- een tijdelijke verhoogde meetfrequentie;
- extra duplo-analyses indien laboratoriumvariatie hoog blijkt;
- bijkomende metingen van sterk gecorreleerde parameters;

Wanneer metingen stabiliseren, wordt het meetplan teruggebracht naar het basisniveau, volgens vooraf vastgelegde beslisregels.

Referenties

- Gilbert, R. O. (1987). *Statistical methods for environmental pollution monitoring*. John Wiley & Sons.
- Helsel, D. R., & Hirsch, R. M. (1993). *Statistical methods in water resources* (Vol. 49). Elsevier.
- Helsel, D. R. (2011). *Statistics for censored environmental data using Minitab and R* (Vol. 77). John Wiley & Sons.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. New York: Springer.
- Lindenmayer, D. B., & Likens, G. E. (2009). Adaptive monitoring: a new paradigm for long-term research and monitoring. *Trends in ecology & evolution*, 24(9), 482-486.
- Shumway, R. H., & Stoffer, D. S. (2006). *Time series analysis and its applications: with R examples*. New York, NY: Springer New York.