



VIDREUDVIKLING AF MODEL FOR MILJØFARLIGE FORURENENDE STOFFER

Forbedring af statistik model og udvidelse af stofgruppe

Teknisk rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 329

2025



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

VIDREUDVIKLING AF MODEL FOR MILJØFARLIGE FORURENENDE STOFFER

Forbedring af statistik model og udvidelse af stofgruppe

Teknisk rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 329

2025

Peter B Sørensen¹
Christian F Damgaard¹
Anders Nielsen²
Hans Thodsen¹
Henrik Tornbjerg¹
Lærke Thorling³
Poul B Løgstrup⁴
Peter E Holm⁵
Dorte Rasmussen⁶

¹Aarhus Universitet, Institut for Ecoscience

²DTU AQUA, Institut for Akvatiske Ressourcer

³GEUS, Afdeling for Geokemi

⁴DTU, Institut for Miljø- og Ressourceteknologi

⁵KU, Institut for Plante- og Miljøvidenskab

⁶DHI, Global kemikaliereregulering og registrering



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

Serietitel og nummer:	Teknisk rapport fra DCE - Nationalt Center for Miljø og Energi nr. 329
Kategori:	Forskningsbidrag
Titel:	VIDREUDVIKLING AF MODEL FOR MILJØFARLIGE FORURENENDE STOFFER
Undertitel:	Forbedring af statistik model og udvidelse af stofgruppe
Forfatter(e):	Peter B Sørensen ¹ Christian F Damgaard ¹ Anders Nielsen ² Hans Thodsen ¹ Henrik Tornbjerg ¹ Lærke Thorling ³ Poul B Løgstrup ⁴ Peter E Holm ⁵ Dorte Rasmussen ⁶
Institution(er):	¹ Aarhus Universitet, Institut for Ecoscience ² DTU AQUA, Institut for Akvatiske Ressourcer ³ GEUS, Afdeling for Geokemi ⁴ DTU, Institut for Miljø- og Ressourceteknologi ⁵ KU, Institut for Plante- og Miljøvidenskab ⁶ DHI, Global kemikaliregulering og registrering
Udgiver:	Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi ©
URL:	http://dce.au.dk
Udgivelsesår:	2025
Redaktion afsluttet:	Januar 2025
Faglig kommentering:	Martin M Larsen
Kvalitetssikring, DCE:	Iben Boutrup Kongsfelt
Sproglig kvalitetssikring:	Charlotte Kler
Ekstern kommentering:	Kommentarerne findes her:
Finansiel støtte:	Ingen ekstern finansiering
Bedes citeret:	Sørensen, P.B., Damgaard, C.F., Nielsen, A., Thodsen, H., Tornbjerg, H., Thorling, L., Løgstrup, P. B., Holm, P. E., Rasmussen, D., 2025 Videreudvikling af model for miljøfarlige stoffer i overfladevand. Forbedring af statistik model og udvidelse af stofgruppe. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 50 s. - Teknisk rapport nr. 329
	Gengivelse tilladt med tydelig kildeangivelse
Sammenfatning:	Resultater fra StatSurf vises for metallerne arsen, barium, bly, cadmium, chrom, kobber, nikkel, vanadium samt zink, mens kviksølv første bliver modelleret, når der foreligger en metode til at få biota prøver med i StatSurf. Simuleringen af metallerne i denne rapport er udført med en ny hurtig beregningsmetode, hvor alle ID15 oplande en koblet sammen, dog uden brug af X matricen.
Emneord:	Metalkoncentrationer, Vandløb, Modellering, Monitoringsdata
Foto forside:	Skjern Å, ved Skarrild, af Peter Borgen Sørensen
ISBN:	978-87-7156-922-3
ISSN (elektronisk):	2244-999X
Sideantal:	50

Indhold

Datablad	2
Forord	5
Sammenfatning	6
Summary	7
1 Indledning	8
2 Opsummering af modelprincip	9
2.1 Statistisk princip	9
2.2 Vandmodel	10
2.3 Retention	10
3 Landsdækkende resultater	12
3.1 Beskrivelse af modelresultat med nikkel som case	12
3.2 Hovedresultater fra de øvrige metaller	21
3.3 Potentiale for inkludering af forklarende variable	38
4 Konklusion og perspektiv	39
5 Referencer	41
6 Bilag 1	42
6.1 Introduction	42
6.2 Statistical framework	42
6.3 References	43
6.4 Supplements	43

Forord

Denne rapport beskriver resultatet af et projekt, hvis formål var at videreudvikle modellen "MetalStat" matematisk, så den statistiske styrke ved modellens forudsigelse forbedres. Rapporten afprøver de ny beregninger for metallerne: arsen, barium, bly, cadmium, chrom, kobber, nikkel, vanadium, zink og kviksølv i vandløb. Projektet er således udført for Miljøstyrelsen med henblik på at kunne understøtte vandplanlægningen med modellering af koncentrationen af miljøfarlige forurenende stoffer i overfladevand bl.a. i forhold til en vurdering af, om der er risiko for forekomst af koncentrationer over miljøkvalitetskrav.

Projektgruppen består af: Peter Borgen Sørensen, Christian Frølund Damgaard og Anders Nielsen, der primært har haft ansvar for den matematiske konstruktion af modellen, der er blevet diskuteret i en faglig arbejdsgruppe bestående af Lærke Thorling, Poul L. Bjerg, Peter Engelund Holm og Dorte Rasmussen. Desuden har Hans Todsén og Henrik Tornbjerg bistået med data for ID15 oplandene. Arbejdet har været fulgt af repræsentanter fra Miljøstyrelsen, primært bestående af Rasmus Dyrmosé Nørregaard og Kasper Lauge Thomassen. Martin Mørk Larsen har været fagfællebedømmer for Institut for Ecoscience. For DCE er rapporten kvalitetssikret af Iben Boutrup Kongsfelt. Projektet har været drøftet i denne gruppe. Miljøstyrelsen har haft et udkast af rapporten til kommentering.

Projektet er udført som et udviklingsprojekt, og den primære projektperiode var fra juni 2024 til december 2024 med afrapportering i 2024.

Det skal bemærkes, at i denne rapport anvendes "." som decimal separator.

Sammenfatning

Det har det haft stor prioritet i 2024 at få beregningerne i MetalStat kodet om til at anvende R pakken: R Template Model Builder (RTMB). Det har vist sig, at RTMB kan løse ligninger så hurtigt, at det både er muligt at udbygge modellen til en mere komplet statistisk model samtidig med, at beregningstiden er reduceret drastisk. I den nye og mere komplette version af modellen kobles den statistiske analyse af måledata direkte med brug af oplands-model for vandstrømningen mellem oplandene. Denne omkodning har gjort modellen generelt mere operationel for stoffer i vand-fasen, der bliver transporteret og omsat i overfladevand, og derfor ændres modellens navn fra MetalStat til StatSurf (**S**tatistical model for **S**urface waters).

StatSurf kobler ID15-oplande sammen under estimering af mediankoncentrationer for de enkelt oplande, hvilket betyder, at målinger i et vandløbssystem får mere generel betydning end bare at være et udtryk for tilstanden i det ID15-opland, der er målt. Dette princip kan derfor udnytte monitorings data langt mere effektivt og estimere geografiske forskelle selv uden brug af en X matrice. Det betyder ikke, at en X matrice er overflødiggjort, men at informationsniveauet af monitorings-data er øget.

Undersøgelsen af effekten af forskellige forklarende variable (X matrice) er forbedret i forhold til MetaStat, da estimering af koefficienterne til X matricen er blevet hurtigere. StatSurf kan derfor avendes i en systematisk undersøgelse af mulige X matricer af forskellige forskergrupper, uden at det behøver være gennem en server eller anden central beregningsenhed. Endvidere kan StatSurf evt. anvendes til beregninger direkte af Miljøstyrelsens ansatte, som derfor får nye muligheder for at drage nytte af modellens mange muligheder.

Tilbageholdelse af metaller i søer er inddraget i modellen ved udviklingen af en retentionsmodel, så søer og vandløb nu bliver modelleret i samme model, hvor effekten af en sø beregnes ved en hydraulisk kontakttid for det opland, hvori søen er placeret.

Denne rapport viser resultater fra StatSurf for metallerne arsen, barium, bly, cadmium, chrom, kobber, nikkel, vanadium samt zink, mens kviksølv første bliver modelleret, når der foreligger en metode til at få biota prøver med i StatSurf. Simuleringen af metallerne er i denne rapport er alene foretaget med brug af den statistiske model, uden brug af X matricen. Næste skridt vil naturligt være at bruge StatSurf til at identificere og evaluere en relevant X matrice for hvert metal for at undersøge mulige forklaringer for de målte koncentrationer. Der er også åbnet op for analyse af andre miljøfarlige forurenende stoffer end metaller.

Summary

It has been a high priority in 2024 to have the calculations in MetalStat re-coded in order to apply the R package: R Template Model Builder (RTMB). It has been shown that RTMB can solve equations so quickly that it is possible to expand the model into a more complete statistical model, while, at the same time, reducing the calculation time drastically. In the new and more complete version of the model, the statistical analysis of measurement data is directly linked to the use of an ID15-catchment model for the water flow between the ID15-catchments, which controls a mass balance for the metal. This re-coding has made the model generally more operational for substances transported in the water phase and, therefore, the name of the model has been changed from MetalStat to StatSurf (**Stat**istical model for **Surf**ace waters).

StatSurf connects ID15-catchments when estimating the median concentration for the individual ID15-catchments, which means that measurements in a watercourse system have more general impact than just being an expression of the local condition of the ID15-catchment that has been measured. This principle can therefore utilize monitoring data much more efficiently and estimate geographical differences even without the use of an X matrix. This does not mean that an X matrix has been made redundant, but rather that the level of information of monitoring data has improved.

The study of the effect of different explanatory variables (X matrix) has improved compared to MetalStat, as the estimation of the coefficients for the X matrix has become faster. StatSurf can therefore be used in a systematic investigation of possible X matrices by different research groups without having to go through a server or other central computational unit. Furthermore, StatSurf can be used for calculations directly by the Danish Environmental Protection Agency's employees, who then get new opportunities to take advantage of the model's many options.

Retention of metals in lakes has been included in the model in the development of a retention model, so that lakes and streams are now modelled in the same model, where the effect of a lake is calculated by a hydraulic contact time for the catchment area in which the lake is located.

This report shows results from StatSurf for the metals arsenic, barium, lead, cadmium, chromium, copper, nickel, vanadium and zinc, while mercury will only be modelled when there is a method for including biota samples in StatSurf. The simulation of the metals in this report is done using only the statistical model without the use of the X matrix. The next step will naturally be to use StatSurf to identify and evaluate a relevant X matrix for each metal in order to investigate possible explanations for the measured concentrations. Analysis of substances other than metals is also possible in the same way.

1 Indledning

En model benævnt MetalStat er udviklet i årene 2022 og 2023, hvor målte metalkoncentrationer fra overfladevand tolkes i en statistisk model med brug af oplysninger fra mindre vandløbsoplande på omkring 15 km² (ID15 oplande). Gennem dette udviklingsarbejde blev det klart, at MetalStat sætter store krav til håndtering af beregningsarbejdet pga. mange variable og ønsket om at bestemme den fælles fordeling af parametrene, så usikkerhederne i modellens forudsigelser kan beregnes. MetalStat har under udviklingsarbejdet hidtil håndteret denne udfordring ved brug af MCMC-metoder, hvor den matematiske struktur og metode blev undersøgt og testet i et fleksibelt udviklingsmiljø. Det har vist sig, at beregningstiden er en udfordring ved brug af MCMC metoder, så i 2024 er det primære formål at mindske beregningstiden med programmet Template Model Builder (TMB) (Kristen m.fl., 2016), som netop i år er blevet gjort anvendeligt gennem statistikprogrammet R (R Template Model Builder: RTMB). I øvrigt blev det anbefalingerne i Sørensen m.fl., (2024), at massebalancekorrektionen i den gamle version af MetalStat burde erstattes af en mere generel løsning af ligningerne, hvis det skulle vise sig regneteknisk muligt, hvilket derfor også har været formålet med udviklingsarbejdet i 2024. I denne rapport ændres modellens navn derfor fra MetalStat til StatSurf (Statistical model for SURface waters).

MetalStat er blevet udviklet på baggrund af overvågningsdata for de fem metaller bly, cadmium, kobber, zink og nikkel. I denne rapport bliver listen over metaller søgt udvidet til også at omfatte arsen, barium, chrom, vanadium og kviksølv i vandløb. Dog viste det sig, at de specielle forhold omkring biota betyder, at biota-data ikke kunne inkluderes i modellen i 2024, og at kviksølv derfor først kan modelleres i 2025 under forudsætning af, at en model for biotakontaminering bliver implementeret.

I denne rapport bruges følgende opdeling af ID15-oplandene:

- Målte ID15 oplande: Oplande, hvorfra der er registreret målinger fra en eller flere målestationer.
- Umålte ID15 oplande: Oplande, hvori der ikke har været registreret målinger.

Der anvendes målinger fra perioden 2011-2023, og et ID15-opland defineres som værende målt, hvis der på et eller andet tidspunkt i denne periode er indsamlet målinger i større eller mindre omfang.

Først opsummeres StatSurf med fokus på de ny elementer i forhold til MetalStat, svarende til en brug af massebalance mellem oplande i tolkning af målte koncentrationer samt inddragelse af retention i søer. Dernæst bruges landsdækkende resultater for nikkel til at beskrive modellens hovedresultater samt betydningen af at bruge massebalancen mellem målte værdier og at inddrage retention for søer. Landsresultater for 9 metaller vises derefter med en opsummering, der viser hvert metals potentiale for at blive modelleret på landsplan med brug af en X matrice i efterfølgende aktiviteter (2025).

2 Opsummering af modelprincip

I dette kapitel opsummeres de statistiske principper bag StatSurf som en udbygget version af MetalStat. Derefter beskrives brugen af massebalance mellem oplande i tolkning af målte koncentrationer med brug af en vandmodel, der beskriver vandstrømning mellem ID15-oplandene. Dernæst beskrives, hvordan StatSurf inddrager effekten af retentionen ved gennemstrømning af søer.

2.1 Statistisk princip

Modellen er beskrevet i detaljer i Bilag1, hvor der henvises til den matematiske beskrivelse. StatSurf er hierarkisk opbygget ligesom MetalStat på den måde at forstå, at modellen har en statistisk model, der tolker målingerne, og i denne model kan der indsættes en oplandsmodel, der søger at forklare de målte koncentrationsniveauer ud fra forholdene i ID15 oplandene. Oplandsmodellen vil typisk bestå af en liste af potentielle forklarende variable for hvert ID15 opland, (X matricen), hvis betydning estimeres i den statistiske model.

Den statistiske del af StatSurf er mere komplet end den statistiske del af MetalStat, da betydningen af vandtransporten mellem ID15 oplandene er taget med i StatSurf. Danmark er opdelt i omkring 3350 ID15 oplande. En mindre del af disse oplande er målte, da de indeholder målestationer. I MetalStat blev mediankoncentrationen estimeret i de målte oplande hver for sig og efterfølgende brugt til at forudsige koncentrationsniveauer i umålte oplande med støtte fra en X matrice. Dette princip er ændret i StatSurf, hvor mediankoncentrationer estimeres i både målte og umålte ID15-oplande. I estimererne indgår vandudveksling mellem ID15-oplandene, hvilket betyder, at estimatet af mediankoncentration i et umålt opland godt kan påvirkes af målinger i opstrøms og nedstrøms retning i andre oplande. De umålte ID15-oplande, der er helt isoleret fra målte ID15-oplande ved slet ikke at dele vand med disse, vil få en estimeret mediankoncentration svarende til en "landsforventning", der er baseret på alle målte ID15-oplande.

Ved at koble alle ID15 oplande sammen med deres indbyrdes vandudveksling opnås på denne måde en mere komplet statistisk model, hvor målinger i et vandløbssystem nu både anvendes til at give information om de specifikke målte ID15 oplande, men også til koncentrationsniveauer i umålte ID15-oplande opstrøms og nedstrøms det målte ID15-oplande.

Potentialet i den nye og mere komplette statistiske model vises i det følgende med resultater fra StatSurf uden brug af X matrice. Hovedforudsætningerne i denne del af modellen er:

- De anvendte ID15-oplande forudsættes repræsentative til at beskrive den landsdækkende rummelige variation i koncentrationsniveauer.
- Den relative forskel i koncentrationsniveau mellem ID15-oplande forudsættes at være log-normalfordelt med middelværdi 0.
- Målte koncentrationer inden for et ID15-opland forudsættes at være log-normal fordelte med samme spredning for alle målestationer. Dvs. at den relative variation forudsættes ens for de enkelte målte koncentrationer for alle ID15-oplande.

- Målinger under detektionsgrænsen forudsættes at følge samme log-normalfordeling som målingerne over detektionsgrænsen.
- Graden af tilbageholdelsen af et metal (retention) forudsættes at være ens for alle opløste fraktioner af det pågældende metal. Dvs. det forudsættes, at metallet ikke optræder på forskellige fraktioner i vandsøjlen på en måde, så en andel af vandfasens metal udfælder hurtigere end en anden andel, der f.eks. er adsorberet til partikler.
- Den relative landsdækkende forskel mellem år for alle ID15-oplandes mediankoncentrationer regnes for at være log-normalfordelte med middelværdi 0.
- Vandudvekslingen mellem ID15-oplandene samt vandtilførslen til de enkelte oplande fra hhv. overfladenært vand og grundvand regnes for konstante, uden at deres reelle usikkerhed inddrages i modellens usikkerhedsberegning.

2.2 Vandmodel

StatSurf anvender ligesom MetalStat input om vandtilførsel og udveksling mellem ID15-oplande fra DK-modellen (<https://denationalehydrologiske-model.dk>), men StatSurf anvender desuden output fra DK-modellen. Den estimerede grundvandstilførsel til ID15-oplandene har dog nogle store fiktive udsving mellem de enkelte ID15-oplande med store negative og positive tilførsler, hvilket skyldes, at vandløbsbunden og grundvanspejlet i DK modellen ikke er helt sammenfaldende pga. modelusikkerheder på meget lokal skala. Det giver store, men fiktive udsving i vandudveksling mellem vandløb og grundvand. På en lidt større skala over flere ID15-opland er der ikke sådanne fiktive udsving, så derfor blev vandtilførslen til de 3350 ID15-oplande udjævnet inden for 350 større aggregerede oplande på følgende måde:

1. Inden for hvert af de 350 aggregerede oplande blev grundvandstilførslen og overfladenær vandtilførsel summeret op over de ID15-oplande, som det aggregerede opland indeholdt. Denne sum blev divideret med det samlede areal af det aggregerede opland for at bestemme en areal specifik tilførsel af overfladenær vand samt grundvand.
2. Den arealspecifikke tilførsel af overfladenær vand samt grundvand blev multipliceret med arealet for hvert ID15-opland, der var indeholdt i det aggregerede opland, som estimerer for tilførelse af overfladenær vand og grundvand i hvert af disse oplande.

2.3 Retention

Den målte koncentration er en opløst fraktion bestemt på filtrerede prøver. Hvis der sker en flytning af metal fra opløst tilstand til bunden tilstand, vil det derfor i modellen indgå som en fjernelse af metal, der benævnes retention. I det omfang der sker en binding til partikler i suspension, der har en størrelse, der tilbageholdes af filterpapiret ved prøvefiltrering, vil metallet stadig befinde sig i vandsøjlen, selvom det regnes som bundet, men typisk vil en fjernelse betyde, at metallet bundfældes eller bindes til faste overflader og dermed kommer ud af vandsøjlen.

I MetalStat er retentionen modelleret som en første orden fjernelse i et såkaldt "stempelflow", hvor et vandvolumen i indløbet er forudsat at gennemløbe en opholdstid i oplandet for derefter at ankomme til udløbet. Det betyder, at alt vand i indløbet opholder sig i et bestemt tidsrum i oplandet, før det løber videre fra oplandet til næste ID15-opland eller ud i det marine miljø. Det vil være en god tilnærmelse til et virkeligt vandsystem, hvor et "stempelflow" simulerer forholdene i et vandløb, der gennemstrømmer et opland. Det er til gengæld en tvivlsom model for en sø, hvor der naturligt forgår en opblanding af søens vandvolumen. Under sådanne vilkår kan et vandvolumen, der ankommer til søens indløb, relativt hurtigt indfinde sig ved udløbet pga. søens interne opblanding. Der vil derfor være en stor forskel på, hvor lang tid et vandvolumen kan opholde sig i søen, før det løber ud i udløbet. Som model for dette forhold kan der anvendes en model for en "fuldt opblandet" beholder, hvor et indløb mixes momentant ud i hele søens vandvolumen. Fuldt opblandede forhold kan ses som værende komplementære til "stempel flow" modellen, og et virkeligt system vil altid være et sted mellem disse to yderligheder. Den reelle forskel mellem "stempelflow" og "fuld opblanding" er undersøgt nærmere i Bilag 1, supplement, hvor en Taylor analyse viser, at de to modeller næsten giver ens resultater, når opholdstiden i et opland er beskedent, mens forskellen bliver markant ved større opholdstider. Det vurderes derfor, at "fuld opblanding" er den bedste model for søer, hvor opholdstiden ofte er flere måneder, og da de to modeller er stort set ens for kortere opholdstider, dvs. for vandløb, så vælges fuldt opblandet som model i StatSurf for bedre at kunne beskrive forholdene i søerne.

Fjernelsen af et stof fra vandfasen betragtes som en proces, der foregår i vandsøjlen, og den hydrauliske opholdstid i et opland kan ses som virkningstiden for fjernelsesprocessen. Den hydrauliske opholdstid kan estimeres som: V/Q , hvor V er vandvolumen i oplandet, og Q er det vand volumen, der pr tidsrum løber ud af oplandet. Som alternativ til opholdstiden kan kontakttiden anvendes svarende til A/Q , hvor A er det våde areal i oplandet.

Hvis fjernelsesprocessen primært forgår i vandsøjlen, er opholdstiden at foretrække frem for kontakttiden, men hvis fjernelsesprocessen derimod forgår på kontaktflader mellem vandsøjlen og sediment/planter, eller hvis sedimentationen ud af vandsøjlen er så langsom, at den ikke kan tilendebringes under opholdstiden, kan kontakttiden være mere velegnet til at beskrive tilbageholdelsen. Variablen hydraulisk opholdstid kan ikke skelne mellem en lavvanded sø og en dyb sø, hvis de har samme vandvolumen og samme vandgennemstrømning, da de vil have samme opholdstid. Modsat gælder for kontakttiden, hvor den lavvandede sø vil have større hydraulisk kontakttid end den dybe sø, og dermed vil den lavvandede sø have en større retention, end den dybde sø under disse vilkår.

De beregninger, der er brugt i denne rapport, har anvendt hydraulisk kontakttid. Om det er hydraulisk opholdstid eller kontakttid, der anvendes i StatSurf, afhænger dog alene af inputtet til StatSurf, så et andet input, som er baseret på hydraulisk opholdstid, vil betyde, at StatSurf bruger opholdstiden i stedet.

3 Landsdækkende resultater

I dette kapitel vises landsdækkende resultater for arsen, barium, bly, cadmium, chrom, kobber, nikkel, vanadium og zink. Først beskrives et resultat fra StatSurf detaljeret med brug af nikkel som eksempel, dernæst vises overordnede resultater for de øvrige metaller.

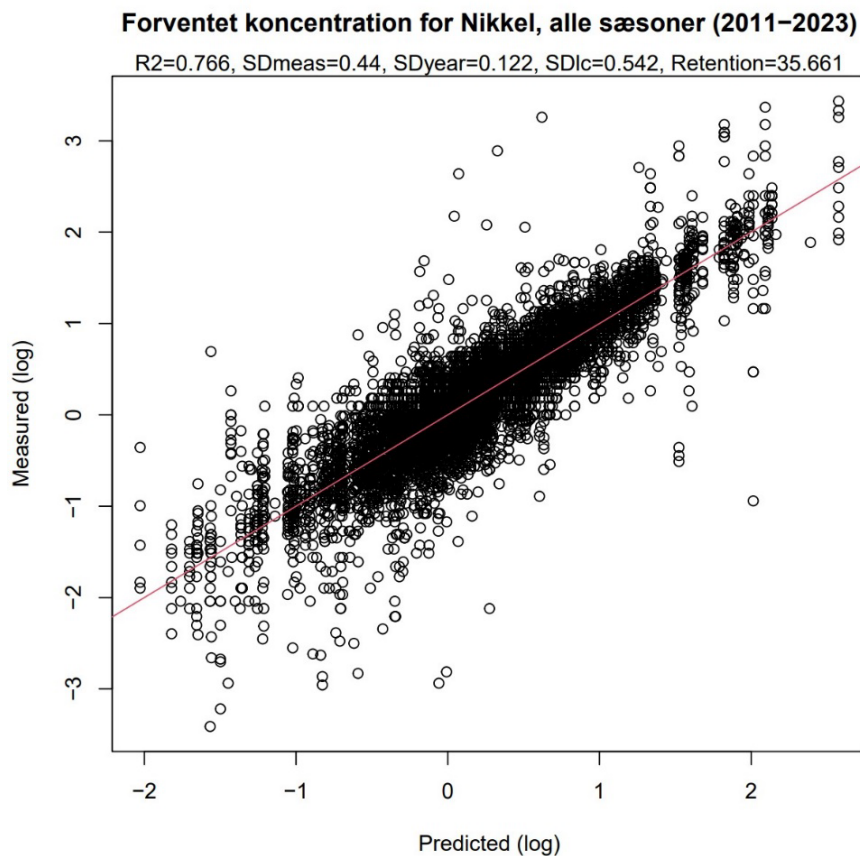
3.1 Beskrivelse af modelresultat med nikkel som case

I dette afsnit illustreres outputtet med brug af nikkel data i perioden fra og med 2011 til og med år 2023. Der kan være en stor forskel mellem sæsoner, så det er muligt at udvælge, hvilke måneder det ønskes at inddrage i modellen, men antallet af målinger mindskes jo korte tidsperiode, der vælges. I denne rapport vælges det at inddrage målinger fra hele året i en samlet analyse.

Modellens generelle resultater ses i Figur 3.1, hvor de forudsagte mediankoncentrationer for de målte ID15-oplande afbildedes på x-aksen sammen med de enkelte målinger på y-aksen. Hvis de estimerede medianer var en eksakt gengivelse af målingerne, ville alle datapunkterne ligge på $y=x$, svarende til den røde linje, og den lineære regression koefficient (R) ville have værdien $R^2=1$. Der ses en tydelig sammenhæng mellem estimerede medianer og måling med en R^2 på 0.77 (vist på Figur 3.1 som $R^2=0.766$). Dette plot er vigtigt som en første evaluering af modellen, især for at undersøge i hvor høj grad ID15-oplandene kan beskrive landsdækkende forskelle i koncentrationsniveauer for de målte ID15-oplande. Men det er vigtigt at gøre klart, at Figur 3.1 ikke kan tages som et billede for modellens evne til at forudsige koncentrationsniveauer i de ID15-oplande, som er umålte. På Figur 3.1 er værdien vist for en række nøglevariable i linjen over plottet udover værdien for R^2 . Standard afvigelsen for målte log koncentrationer (SD_{meas}) beskriver variationen mellem gentagende målinger i samme ID15-opland. En specifik tolkning af denne standardafvigelse er følgende: Den relative forskel mellem målte koncentrationer er $\exp(0.44)=1.55$, hvilket betyder, at 68% af målingerne ligger inden for $\mu/1.55$ og $\mu \cdot 1.55$, hvor μ er mediankoncentrationen. To andre standardafvigelser er ligeledes angivet på Figur 3.1 som henholdsvis standardafvigelsen mellem år ($SD_{year}=0.123$) og mellem ID15-oplande ($SD_{lc}=0.542$). Det er væsentligt at konstatere, at standardafvigelsen mellem ID15-oplande er den største, svarende til den variation i koncentrationsniveauer, der potentielt kan beskrives med forklarende variable (X matrice). Derfor er SD_{lc} et vigtigt mål til at vurdere betydningen af en forklarende variabel i test af kommende X matricer.

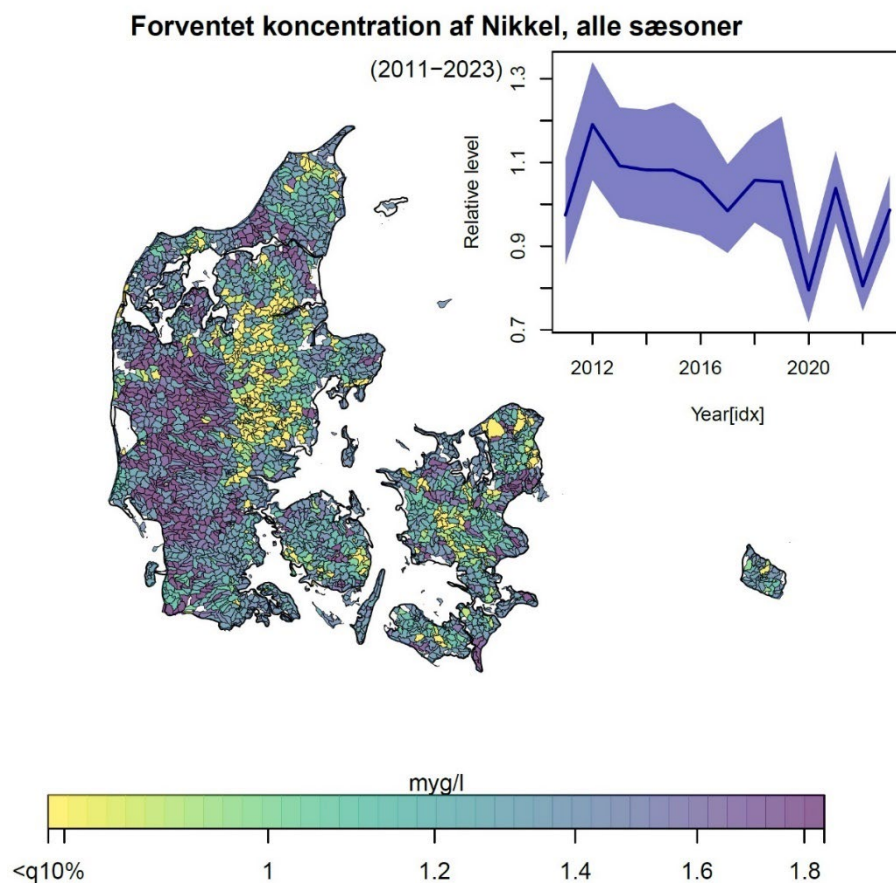
Den sidste variabel i listen over variable over plottet på Figur 3.1 er retentionskoefficienten ($Retention=35.661$) for fjernelse af nikkel fra vandfasen. Denne variabel kan tolkes ved at anvende relationen for tilbageholdelse i et opland: $1-1/(1+Retention \cdot Th)$ (Bilag1, Supplement, Eq. 12), hvor Th er den beregnede hydraulisk kontakttid i hvert ID15-opland. Ved at indsætte medianværdien for Th i: $1-1/(1+Retention \cdot Th_{median})$ fremkommer en karakteristisk medianfjernelse fra vandfasen ved passage af et ID15-opland. For nikkel bliver medianfjernelse fra vandfasen ved gennemløb af et ID15-opland med en retentionskoefficient på 36 således 3.4%. Denne medianfjernelse dækker selvfølgelig over en stor variation i retention mellem de enkelte oplande, hvor oplande, der indeholder søområder, har langt den største fjernelse på op mod 100%. For metaller skyldes en fjernelse typisk en adsorption, udfældning, eller optag i biota, mens for organiske forurenende miljøfarlige stoffer kan det også skyldes nedbrydningsprocesser og fordampning.

Figur 3.1. Sammenligning mellem modellens forudsigtelse (prediction) og faktiske målinger i samme opland (Measured), hvor værdierne er på log skala.



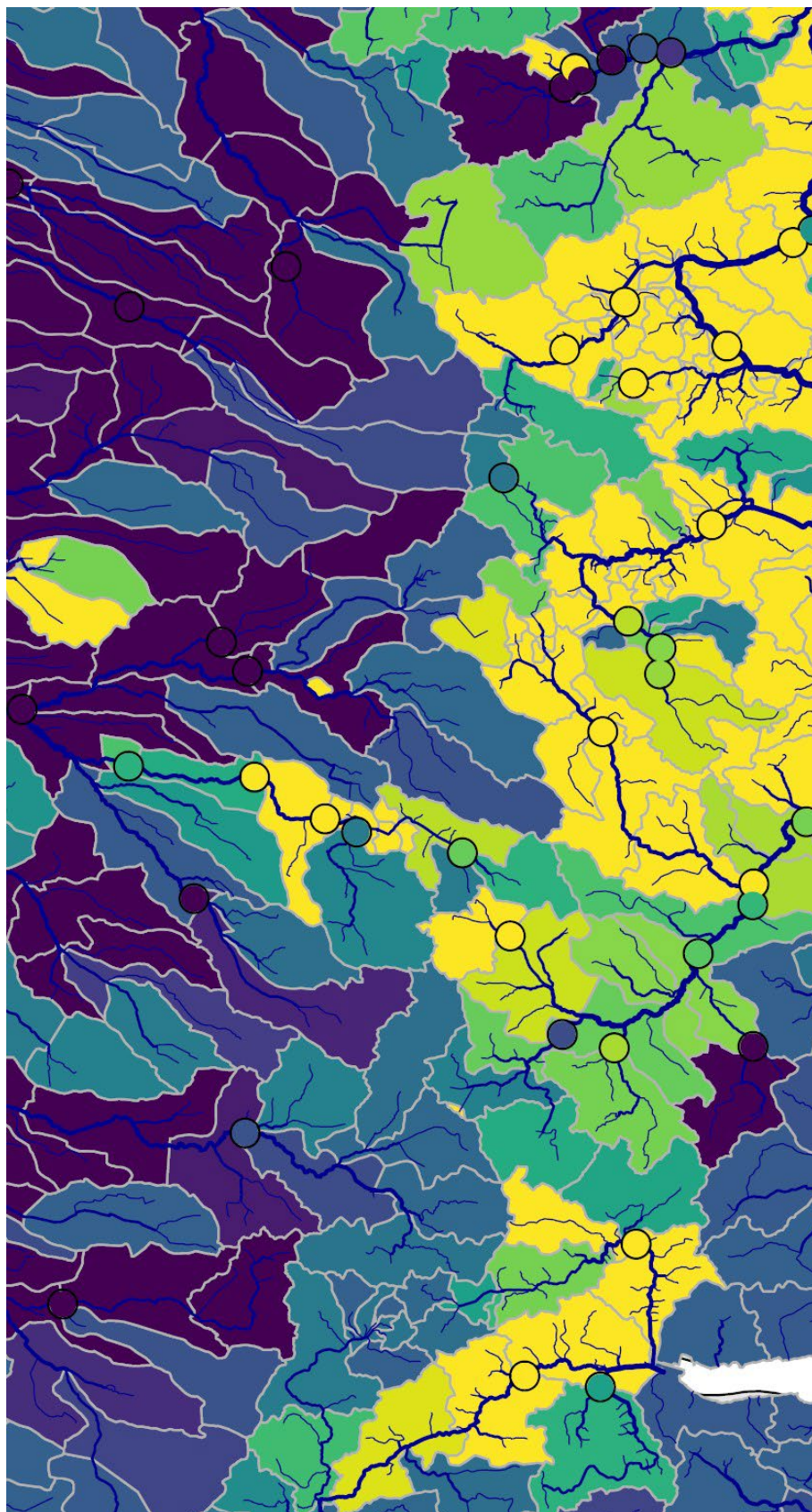
De viste resultater i denne rapport anvender udelukkende målte koncentrationer samt vandudveksling mellem oplande som input samtidig med, at der skelnes mellem år for at undersøge for generelle tidstrends. Det betyder, at der ikke tages hensyn til forklarende variable (X matrice), som varierer mellem forskellige oplande. Modelresultaterne vises på et kort over alle oplande, hvor farven afhænger af værdier for den estimerede mediankoncentration, dels med en graf, som viser den landsdækkende relative forskel mellem år med angivelse af 95 % konfidensinterval (Figur 3.2). Der ses en svag tendens til en fald i nikkelkoncentrationen gennem årene, med et siksak mønster mellem 2020, 2021, 2022, og 2023. Den slags udsving kan have en reel forklaring som f.eks. forskelle mellem våde og tørre år. Det kan dog også være artefakter, som indikerer, at tidstrenden skal tages med forbehold, da sådanne forskelle fra det ene år til det andet evt. blot kan være udtryk for forskellige strategier for udvælgelse af målestationer fra det ene år til det andet år. Derfor bør en tidstrend betragtes over en længere årrække. Effekten mellem år kan estimeres mellem målestationer, også selvom hver station ikke er målt for hvert eneste år, men hvis der er en generel forskel på kriterierne for, hvordan stationer er udvalgt fra det ene år til den næste, vil det blive tolket af modellen som en effekt mellem år. Effekter mellem år er ikke kun taget med i modellen for at undersøge for tidstrends, men også for at eliminere mulige forskelle mellem år, både artefakter og reelle, i at skabe bias i modellens øvrige estimater.

Figur 3.2. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for nikkel med brug af målinger hele året.

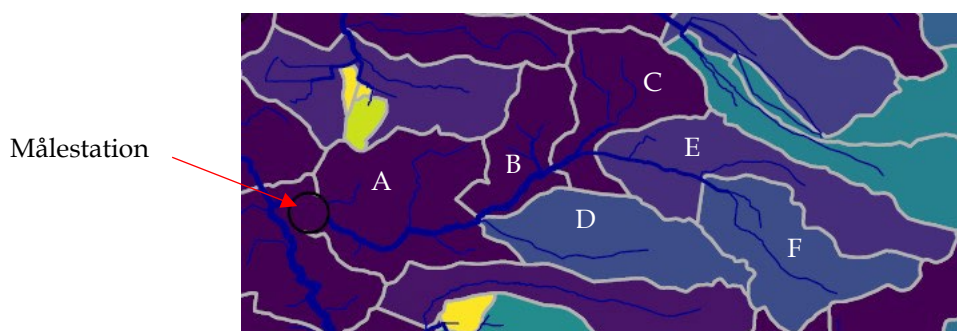


Det er tydeligt, at Figur 3.2 viser nogle markante geografiske forskelle i de estimerede mediankoncentrationer, især omkring israndslinjen fra den sidste istid. Dette kan undre, set i lyset af at hovedparten af ID15-oplandene slet ikke har målinger og derfor alene er estimeret på baggrund af målte oplande uden brug af X matrice, der potentielt kan adskille forskelle i karakteristika mellem ID15-oplande. Grunden til, at der alligevel er tydelige geografiske forskelle, skyldes, at StatSurf har vandudvekslingen med i estimatet for mediankoncentrationerne. Dette er illustreret i Figur 3.3, hvor der er zoomet ind på grænsen mellem hhv. Skjern Å og Gudenåens oplande. Farvekoden er den samme, som er anvendt på Figur 3.2, og desuden er målestationerne vist som cirkler, hvor middelværdien af de målte værdier i hver station angiver cirkelns farve efter samme farvekode, som er anvendt for oplandene. Derudover er vandløbskortet lagt ind over med en strektykkelse for hver vandløbsstrækning, der afhænger af vandløbets hierarki målt efter Strahler hierarkiet (Horton, 1945), der adskiller vandløbsstrækninger hierarkisk efter det største antal forgreninger i vandløbsnetværket ud til oplandsgrænse for hovedoplandet, der løber til strækningen. De fleste målestationer ligger omkring udløbet fra et ID15-opland, og der ses en klar tendens til, at farven af en cirkel matcher farven af det opland, som har stationens placering som udløb. Dette sammenfald er selvfølgelig ikke overraskende, da modellen anvender data både til at bestemme middelværdi i cirklerne og til at estimere mediankoncentrationer i ID15-oplande. Dette mønster er en visuel validering af at modellen opfører sig som forventet.

Figur 3.3. Udsnit af Figur 3.2. Omkring midten af figuren hvis området ved Tinnets kraft, med Skjern Å's opland mod venstre og Gudenåen mod højre. Cirklene er placeret ved målestationer og farven svarer til gennemsnitslig målt koncentration.

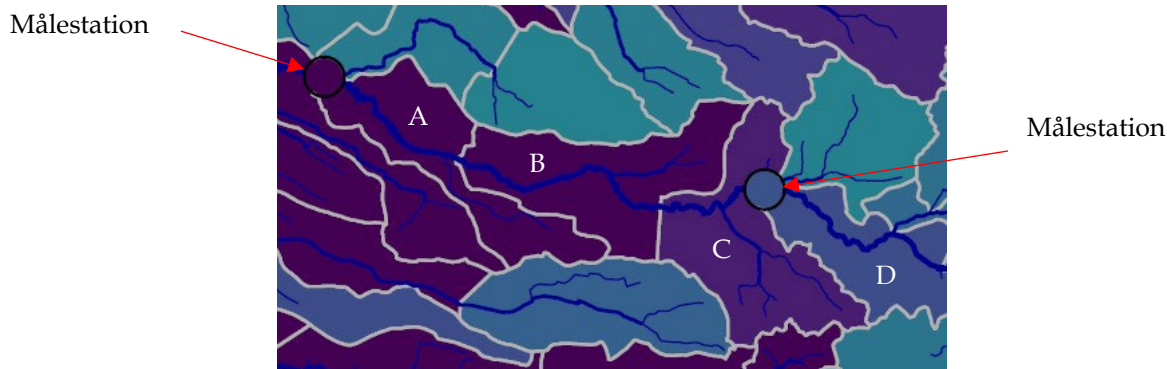


Den tydelige opdeling, der ses på Figur 3.3 hen over israndslinjen, undersøges nærmere i Figur 3.4, hvor ID15-oplandene for hele Fjederholt Å kan ses. Fjederholdt Å er tilløb til Rind Å, som igen er en del af Skjern Å oplandet. ID15-oplandene i Fjederholt Å oplandet er navngivet med bogstaverne A til F. Opland A er det nederste opland for Fjederholt Å, som også er det ID15-opland, der har en målestation, hvor der er målt høje nikkel koncentrationer, svarende til den mørke farve af cirklen. Dvs. at ID15-opland A får estimeret en høj mediankoncentration, for at modellen bedst muligt kan estimere de målte værdier i udløbet fra oplandet. Meget af det vand, som løber ud af opland A, kommer fra opland B, da det følger hovedløbet af Fjederholt Å. Det betyder, at den estimerede mediankoncentration fra opland B også skal være høj, fordi det stort set er det samme vand, som der måles på i opland A, og det samme gælder for opland C. Opland E er dog så langt væk fra opland A, at vandvolumet fra dette opland er beskedent sammenlignet med udløbet fra opland A. Det betyder, at udløbet fra opland E godt kan have en lavere mediankoncentration samtidigt med, at koncentrationen er høj i opland A. Derfor vil modellen forvente en mediankoncentration i opland E, som er tættere på, hvad der kan forventes på landsplan, uden hensyn til lokale forhold, og denne tendens forstærkes med det faldende vandvolumen op til opland F. Både for opland F og D gælder det, at det vandvolumen, der strømmer herfra til opland A, er så beskedent, at modellen ikke estimerer mediankoncentration op efter den højere koncentration i ID15-opland A. StatSurf er konservativ på den måde, at der kun afviges fra forventning på landsplan, hvis der er målinger som nødvendiggør dette.



Figur 3.4. Farvelægning efter forventet mediankoncentration for nikkel i Fjederholt Å (skala som Figur 3.2) med angivelse af målestation, hvor der er indsamlet analyser, samt hvilke ID15-oplande som Fjederholt Å afvander.

Et andet eksempel på, hvordan umålte ID15-oplande kan påvirkes af målte ID15-oplande, vises på Figur 3.5 med en serie af 4 forbundne ID15-oplande langs Omme Å. I udløbet fra opland A ligger en målestation, der har målt høje koncentrationer, mens der opstrøms i udløbet fra opland D er målt lavere koncentrationer. De to umålte oplande (B og C), som ligger mellem de to målte (A og D), skal både opfylde betingelsen om, at deres input nedstrøms til opland A skal indeholde en høj koncentration, mens deres input opstrøms fra opland D har en lavere koncentration. Det betyder, at estimatet for disse to umålte oplande lægger sig imellem de to målte med en stigende koncentration fra opland C til B. Det tvinger mediankoncentrationen i ID15-opland B til at være noget mindre end i ID15-opland A, selv om disse to oplande langt hen ad vejen deles om det samme vandvolumen.

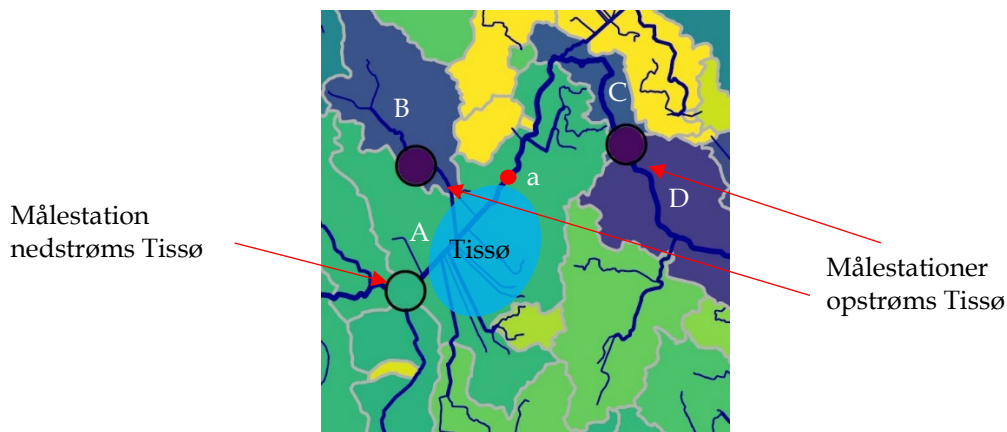


Figur 3.5. Farvelægning efter forventet mediankoncentration for nikkel i en oplandsserie langs Omme Å (skala som Figur 3.2) med angivelse af to målestationer, hvor der er indsamlet anlaysr.

Retentionen har desuden en indflydelse på estimatet af mediankoncentration for umålte oplande, hvor ID15-oplande med relativ høj retention vil blive estimeret til at have en lavere mediankoncentration sammenlignet med ID15-opland med relativ lav retention.

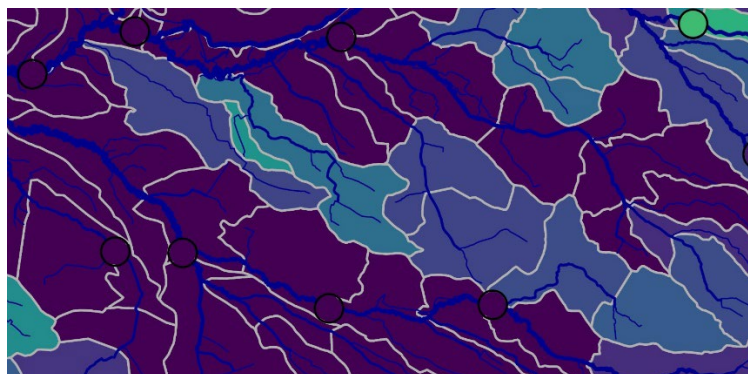
StatSurf inddrager retention i ID15-oplande, hvor metaller tilbageholdes med en faktor, som estimeres i modellen. Som beskrevet i teksten til Figur 3.1, så fjernes typisk 3,5 % af nikkel i et opland, men dette er median af alle ID15-oplande, der varierer voldsomt alt efter ferskvandarealet og vandflowet i de enkelte ID15-oplande - og dermed kontakttiden i hvert opland. På Figur 3.6 ses ID15-oplandene omkring Tissø, der er indikeret på figuren ved en blå ellipse. Her ses det, at der på målestationen fra ID15-opland A er målt relativt lave koncentrationer, vist med den grønne farve i cirklen, og ID15-opland A er derfor tillagt en tilsvarende lav mediankoncentration, selvom to målestationer tæt på indløbet til oplandet viser høje koncentrationsniveauer med de mørkt farvede cirkler. Hvis ikke der var placeret en sø i opland A, ville modellen ikke have været i stand til at estimere en lav koncentration i opland A, samtidig med at der var høj koncentration i oplandene B og C. Det skyldes, at massebalancen mellem oplanden uden en sø ikke ville have tiladt en stor masse af indstrømmende nikkel fra oplanden B og C, samtidig med at koncentrationen for opland A var lav, da en tilsvarende stor masse så også skulle strømme ud af opland A. Desuden ses det, at estimatet for opland C, der ligger mellem de målte oplande A og D, er tæt på koncentrationsniveauet i opland D, med omtrent samme mørke farve, selvom opland A har et markant lavere koncentrationsniveau. Dette spring i koncentrationsniveau er muligt at beskrive med StatSurf pga. retentionsmodellen, da masse derved kan tilbageholdes i søer og dermed give koncentrationsfald mellem indløb og udløb af søer vist på Figur 3.6.

I øvrigt illustrerer Figur 3.6, hvordan StatSurf er en oplandsmodel, der tillægger hele ID15-oplandet den samme mediankoncentration. Således er det oplagt, at indløbet helt frem til Tissø inden opland A vil have en høj mediankoncentration, svarende omtrent til opland C, men da Tissø ikke dækker hele arealet i opland A, betyder det, at Halleby Å, der løber ind i søen fra opland C (markeret ved a), må have en høj mediankoncentration, svarende til ID15-opland C, helt hen til indløbet af Tissø inde midt i opland A. Især i oplande, der indeholder søer, kan der således være sådanne store ikke-modellerede koncentrationspring inde i oplandene mellem indløb og udløb til og fra søer.

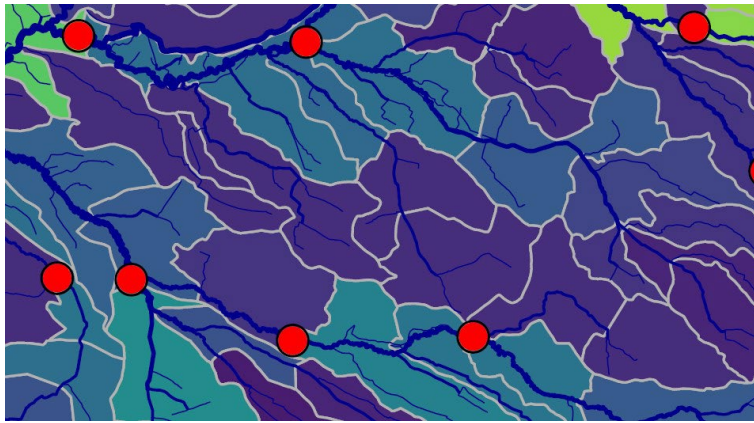


Figur 3.6. Udpluk af ID15-oplandet med Tissø (opland A), samt de nærmeste ID15-oplande ved indløbet til søen (oplande B, C, og D). Cirklerne angiver målestationer og farveskale følger Figur 3.2.

På Figur 3.7. ses et udpluk af ID15-oplande fra Skjern Å oplandet, hvor Omme Å løber nederst på figuren, mens Karstof Å afgrænser mod nord. Modellen forudsiger, at mediankoncentrationen er lavere (lysere farver) mellem Omme Å og Karstof Å, og det kunne antyde, at vandet i de mindre vandløb mellem hovedløbene Omme Å og Karstof Å har en mere beskeden mediankoncentration. Men det kan også være udtryk for, at modellen i disse ID15-oplande mangler oplysninger, og derfor tillægges disse oplande en mediankoncentration, der er mere tæt på forventningen på landsplan. Dette kan undersøges nærmere ved at udnytte, at StatSurf har en fuld usikkerhedsvurdering af alle estimater, hvilket betyder, at modellen estimerer usikkerheden på estimaterne af mediankoncentrationen i hvert ID15-opland, og usikkerheden på estimaterne er vist på Figur 3.8, hvor målestationerne vises som røde cirkler. På Figur 3.8 er høj estimeret usikkerhed vist i mørke lilla farver, mens lav estimeret usikkerhed vises som lysere farver. Det ses, at netop ID15-oplandene mellem Omme Å og Karstof Å har højere usikkerhed end ID15-oplandene, der gennemstrømmes af hovedløbene af de to åer, da disse er målte. Så konklusionen er, at de lavere mediankoncentrationer mellem Omme Å og Karstof Å kan skyldes en større usikkerhed af estimaterne. Igen vil gyldige forklarende variable kunne hjælpe med til at give et mere kvalificeret estimat af disse oplande mellem Omme Å og Karstof Å. Alternativt kan der indsamles flere målinger, som beskrevet neden for.

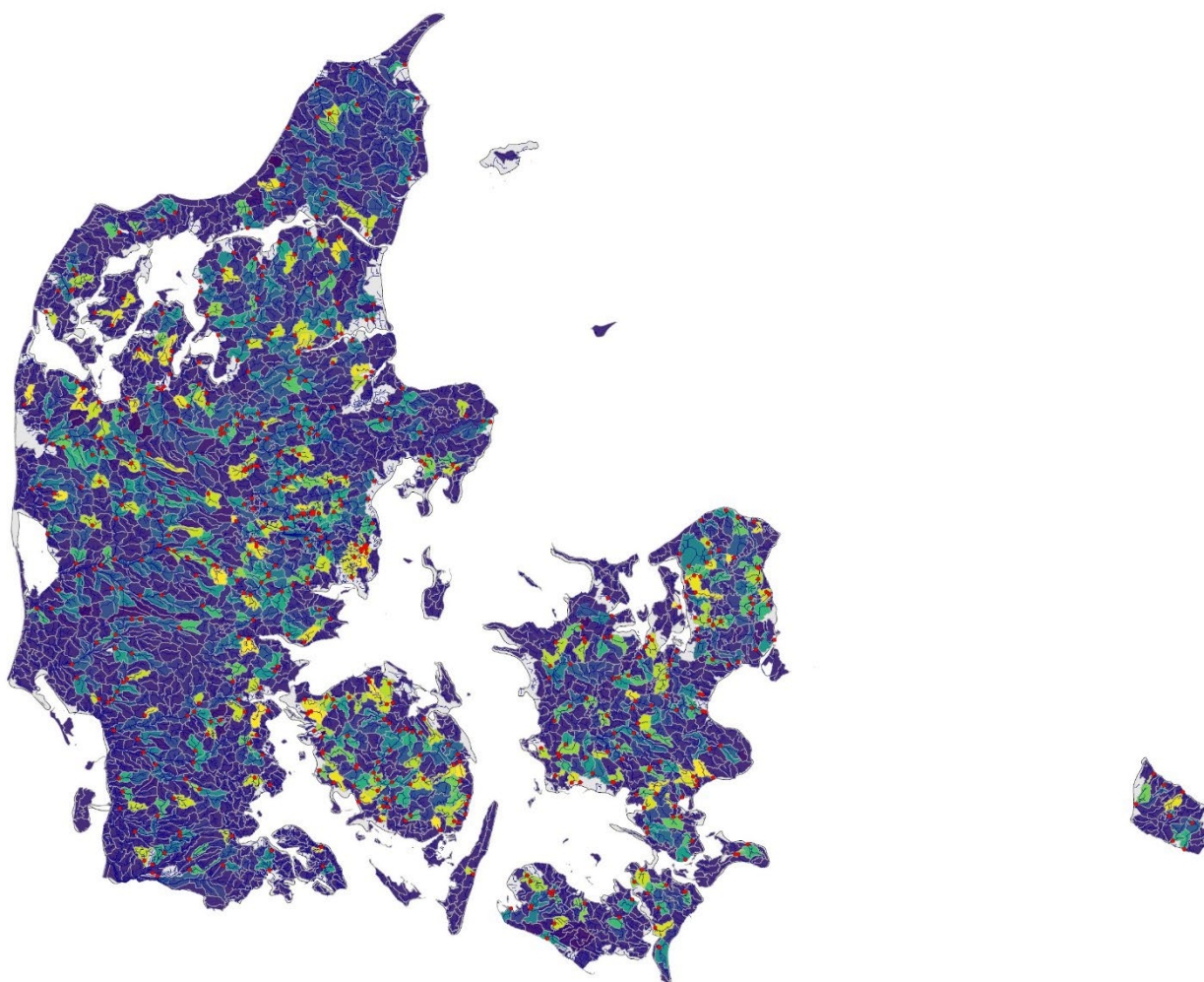


Figur 3.7. Udpluk af ID15-oplande fra Skjern Å oplandet mellem Omme Å i syd og Karstoff Å i nord med farve skala som Figur 3.2



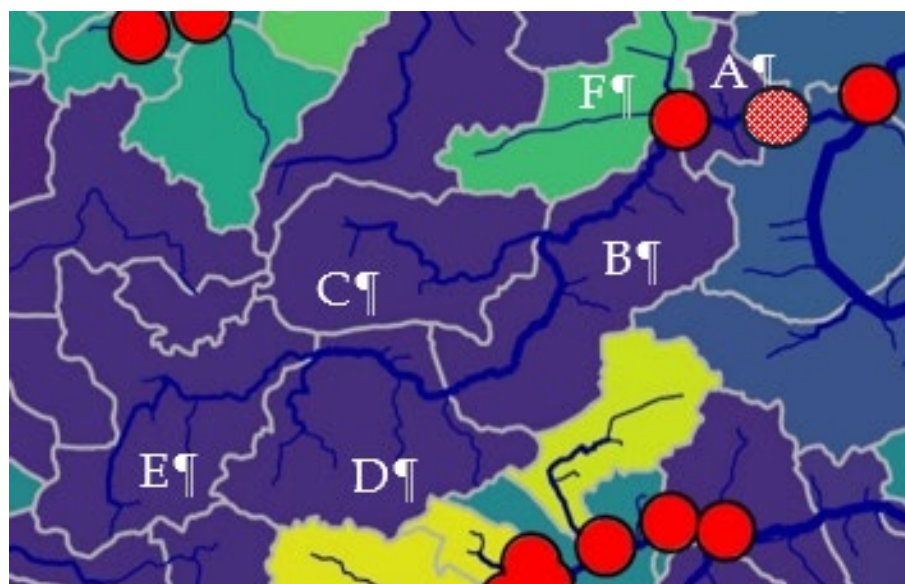
Figur 3.8. Farvekode efter graden af usikkerhed ved estimaterne af mediankoncentration for samme område som vist på Figur 3.7. Mørke farver er større usikkerhed, mens lyse farve er mindre usikkerhed. De røde prikker viser placering af målestationer.

En landsdækkende kortlægning af usikkerheden på estimater af mediankoncentrationen kan give input til en kostoptimal planlægning af monitoringsaktiviteter. På Figur 3.9 ses usikkerheden på estimatet af mediankoncentrationen på landsplan. En optimal planlægning af monitoringsaktiviteter vil fokusere på de mørke områder og gå nedstrøms inden for et mørkt område for at måle på så meget vand som muligt. Dette illustreres i Figur 3.10, der viser Tange Ås ID15-oplande fra udspring til udløb i Gudenåen ved Tangeværket. Opland A grænser op til udløbet i Gudenåen og afvander således fra oplandene B, C, D, E og F. Der er målt for udløbet fra opland F, hvilket betyder, at usikkerhed for mediankoncentrationen fra dette opland med indløb til opland A er lavere end for de øvrige oplande omkring Tange Å. Ved at måle i udløbet fra opland A, vil det være muligt at mindske usikkerheden langs hovedløbet af Tange Å gennem opland B og D. Opland E vil nok ligge så langt opstrøms, at der stadig vil være en større usikkerhed på mediankoncentrationen, selvom der måles i opland A, men en strategi kan være at måle i udløbet fra opland A (mønstret rødcirkel på Figur 3.10) og efterfølgende opdatere modellen med disse målinger for derefter at beslutte, om der skal indsamles flere målinger fra Tange Ås opland.



Figur 3.9. Relativ usikkerhed på estimer af mediankoncentration for ID15-oplande, mørke farver er højere usikkerhed, mens lyse farver er lavere usikkerhed. De røde prikker viser placering af målestationer.

Figur 3.10. Relativ usikkerhed, som udpluk fra Figur 3.10 af Tange Ås opland (ID15-oplande A, B, C, D, E, F). Den mønstrede røde cirkel angiver den optimale nye målestation, hvorfra målinger kan nedbringe usikkerheden for Tange Ås opland



Ovenstående analyse har fokus på mediankoncentrationen i hvert opland, hvilket er et af hovedresultaterne fra modellen, men modellen kan også estimere andre nyttige variable. For eksempel kan den fittede model anvendes til at estimere massen af nikkel ved at multiplicere vandvolumen som strømmer fra ID15-oplandene med de estimerede mediankoncentrationer, hvilket vises på Figur 3.11. Det ses klart af figuren, at den transporterede masse af nikkel domineres af de store Vestjyske åer, da det er her, de højeste koncentrationer er til stede, samtidig med at vandløbene har stor vandføring. Således estimeres Skjern Å til at udlede omkring 15 kg nikkel per døgn til Ringkøbing Fjord. Det er muligt at estimere den samlede masse af nikkel, der per dag vaskes ud til marine områder, svarende til 371 kg/d, og den samlede fjernelse af nikkel fra overfladevandet på vejen til det marine miljø kan estimeres til 16 kg/d.

Figur 3.11. Transporteret masse af nikkel ud af ID15-oplandene (g/d). De røde cirkler angiver målestationer.

Masse (g/d) for Nikkel alle sæsoner (2011–2023)

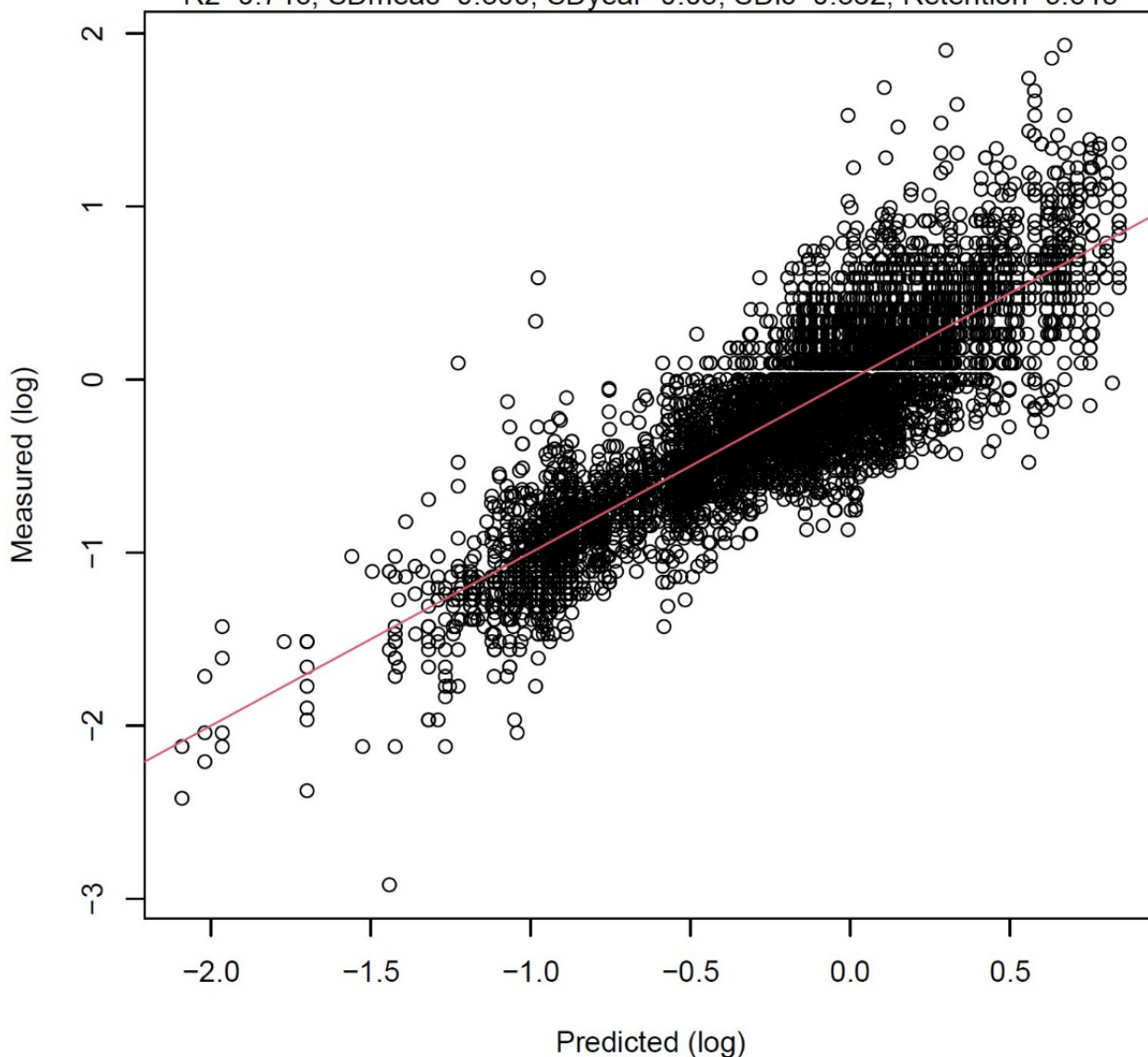


3.2 Hovedresultater fra de øvrige metaller

Modellens generelle resultater for Arsen ses på Figur 3.12, svarende til Figur 3.1 for nikkel, hvor de forudsagte mediankoncentrationer for de målte ID15-oplande afbildes på x-aksen sammen med de samhörørende målinger på y-aksen. Dette plot viser, at variationen mellem oplande typisk dominerer over variationen mellem enkelte målinger fra samme opland, da $SD_{lc}(0.382)$ er større end $SD_{meas}(0.309)$, mens retentionen på omkring 10 (9.648) er noget mindre end retentionen for nikkel og giver en medianfjernelse i oplandene på omkring 1 %.

Forventet koncentration for Arsen, alle sæsoner (2011–2023)

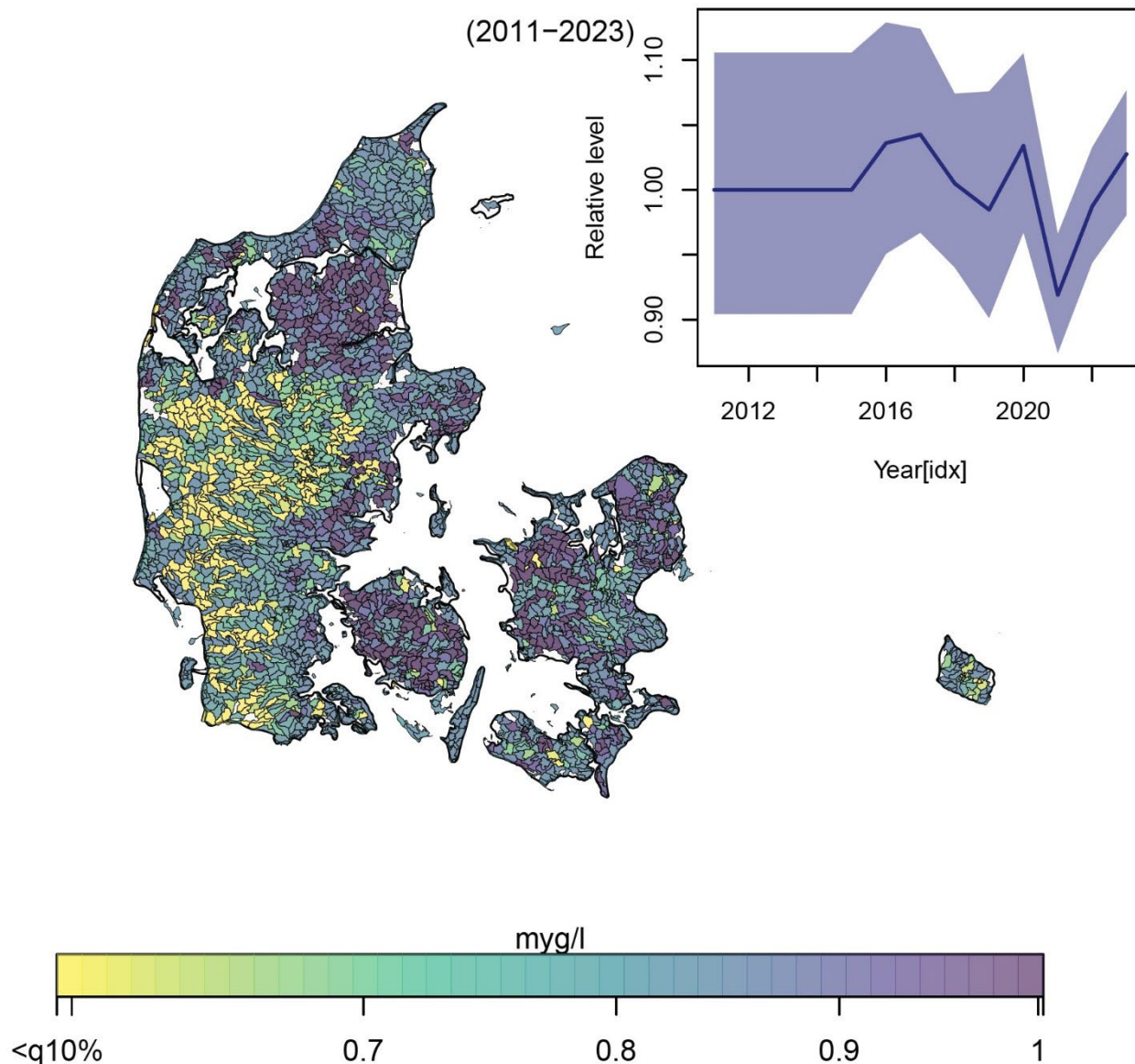
$R^2=0.746$, $SD_{meas}=0.309$, $SD_{year}=0.05$, $SD_{lc}=0.382$, $Retention=9.648$



Figur 3.12. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for arsen med brug af målinger fra alle sæsoner (hele året).

På Figur 3.13 ses forventet mediankoncentration i ID15-oplandene for arsen. Modsat for nikkel viser arsen den omvendte betydning af israndslinjen, med den laveste koncentration i det vestlige og sydlige Jylland.

Forventet koncentration af Arsen, alle sæsoner

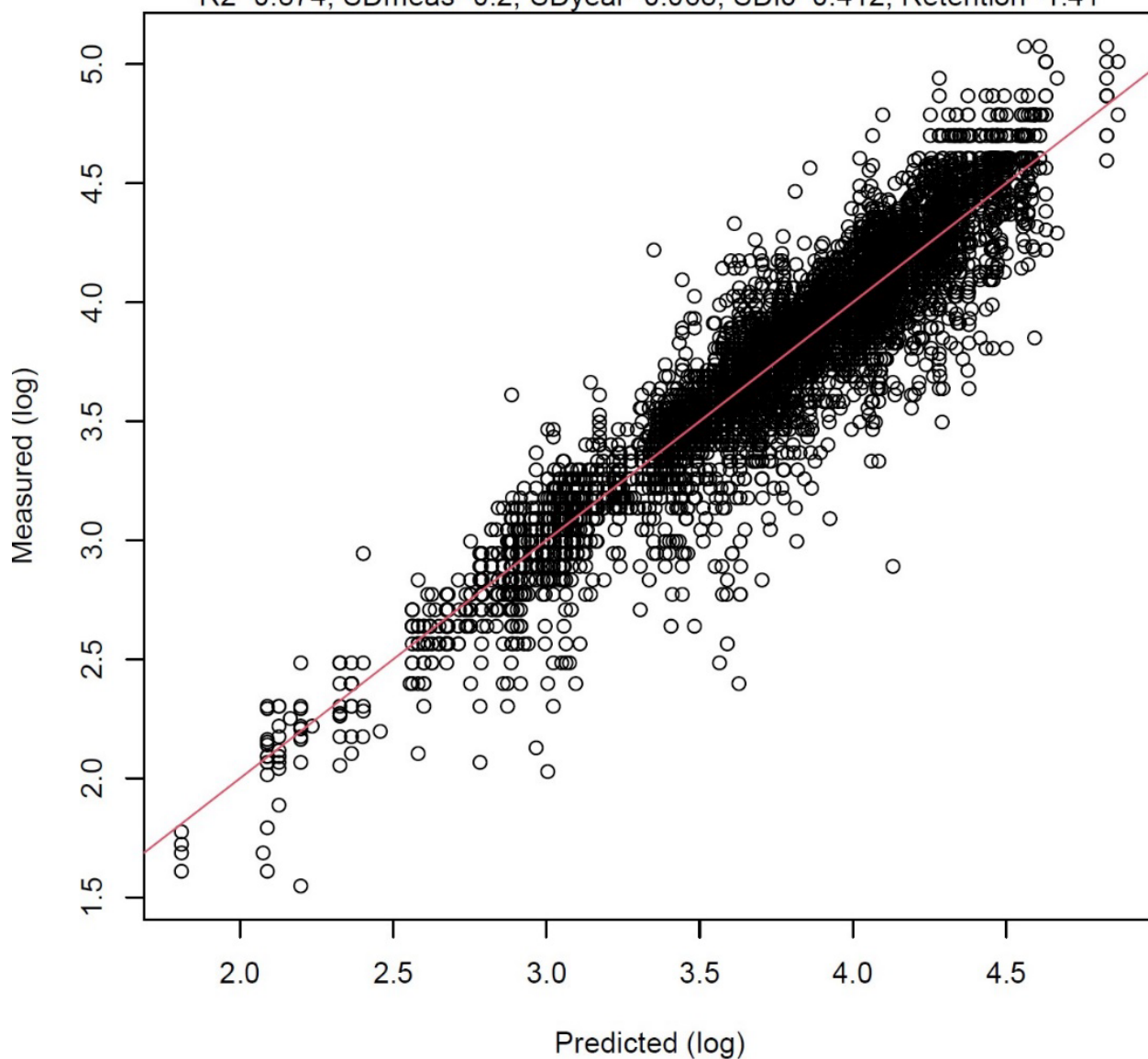


Figur 3.13. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for arsen med brug af målinger fra alle sæsoner (hele året).

På Figur 3.14 ses de generelle resultater for barium, variationen mellem oplande (SDlc) estimeres til at være dominerende over variationen mellem målinger på i samme opland (SDmeas). Dette resultat, sammenholdt med at alle målinger er målt over detektionsgrænsen, giver en høj R^2 værdi på 0.874. Retentionen estimeres som værende meget lav med værdien 1 (1.41) sammenlignet med både nikkel og arsen. Denne lave retention giver en medianfjernelse i et ID15-opland på omkring 0.1 %.

Forventet koncentration for Barium, alle sæsoner (2011–2023)

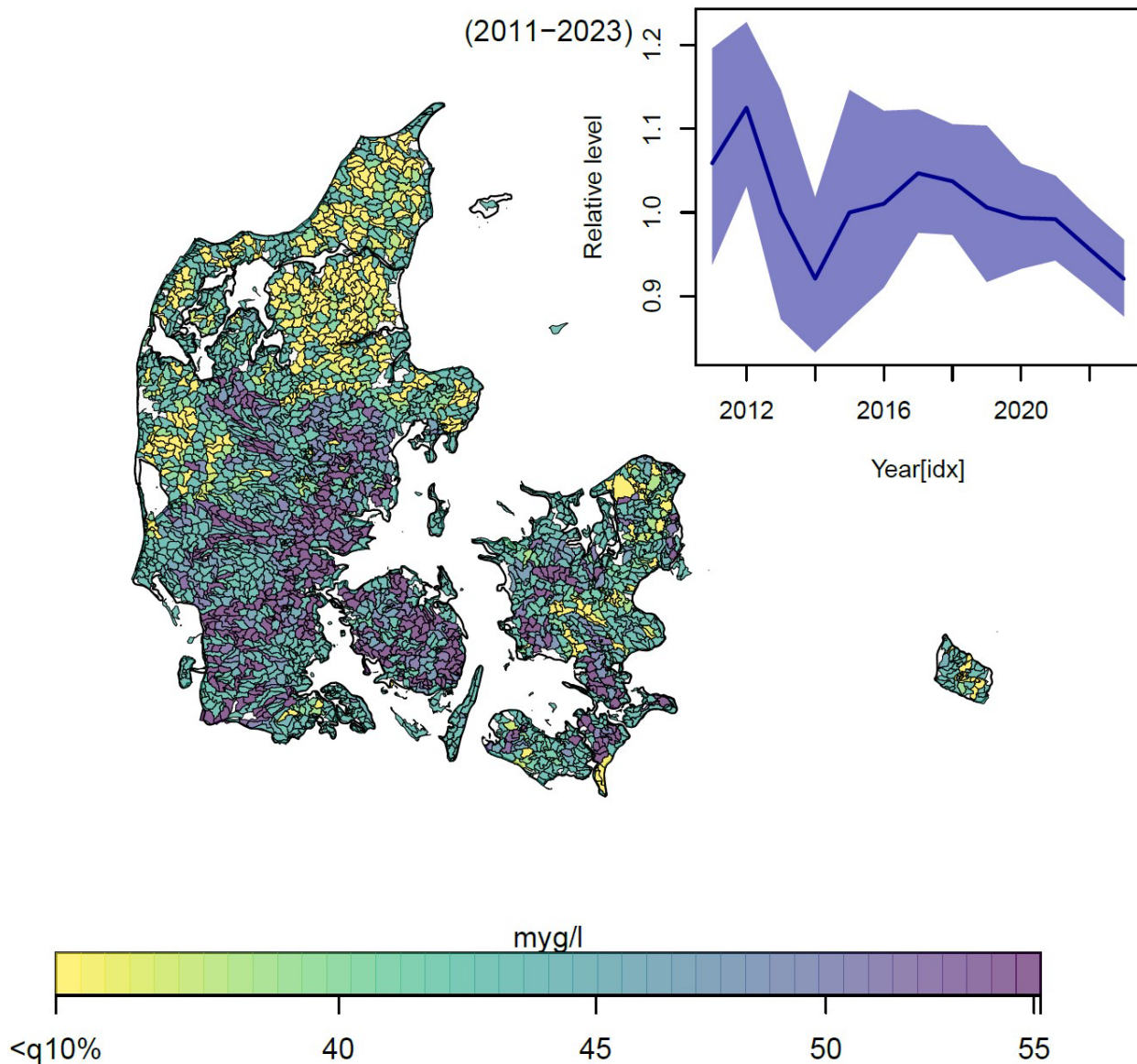
$R^2=0.874$, $SD_{meas}=0.2$, $SD_{year}=0.068$, $SD_{lc}=0.412$, $Retention=1.41$



Figur 3.14. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for barium med brug af målinger fra alle sæsoner (hele året).

Forventet mediankoncentration af barium for ID15-oplandene er vist på Figur 3.15, hvor det især er den sydlige halvdel af Jylland, samt Fyn, der udviser en forhøjet koncentration. De tilfældige effekter af år antyder en beskedent faldende tendens fra omkring 1.1 ned til cirka 0.9 i relativ forskel.

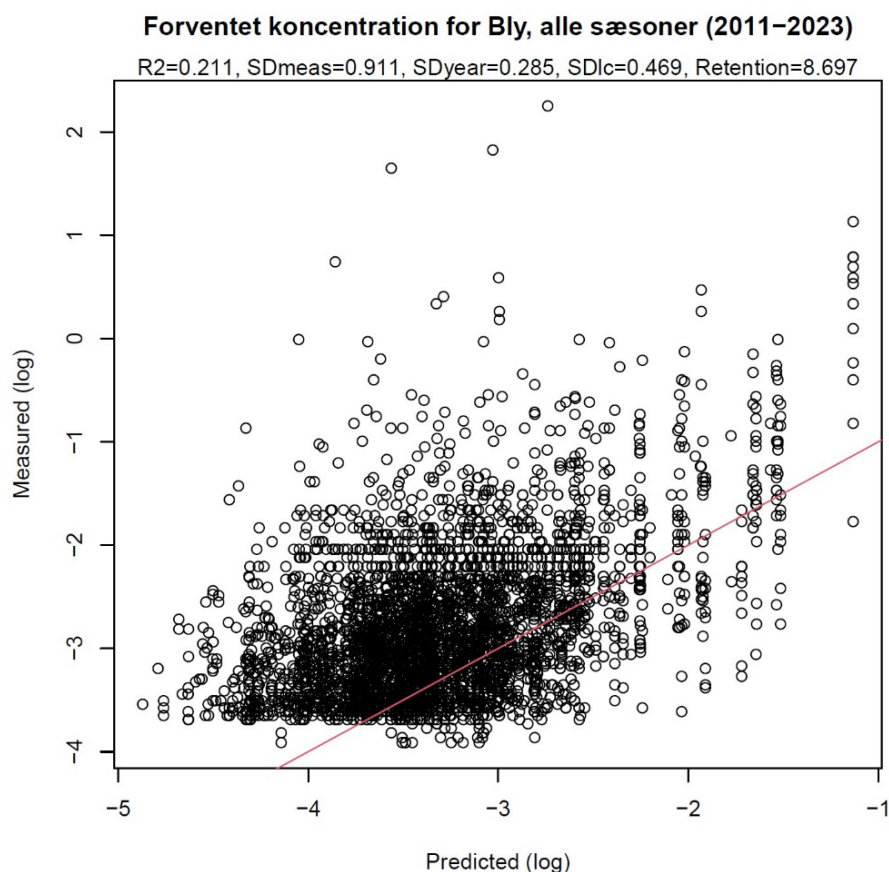
Forventet koncentration af Barium, alle sæsoner



Figur 3.15. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for barium med brug af målinger fra alle sæsoner (hele året).

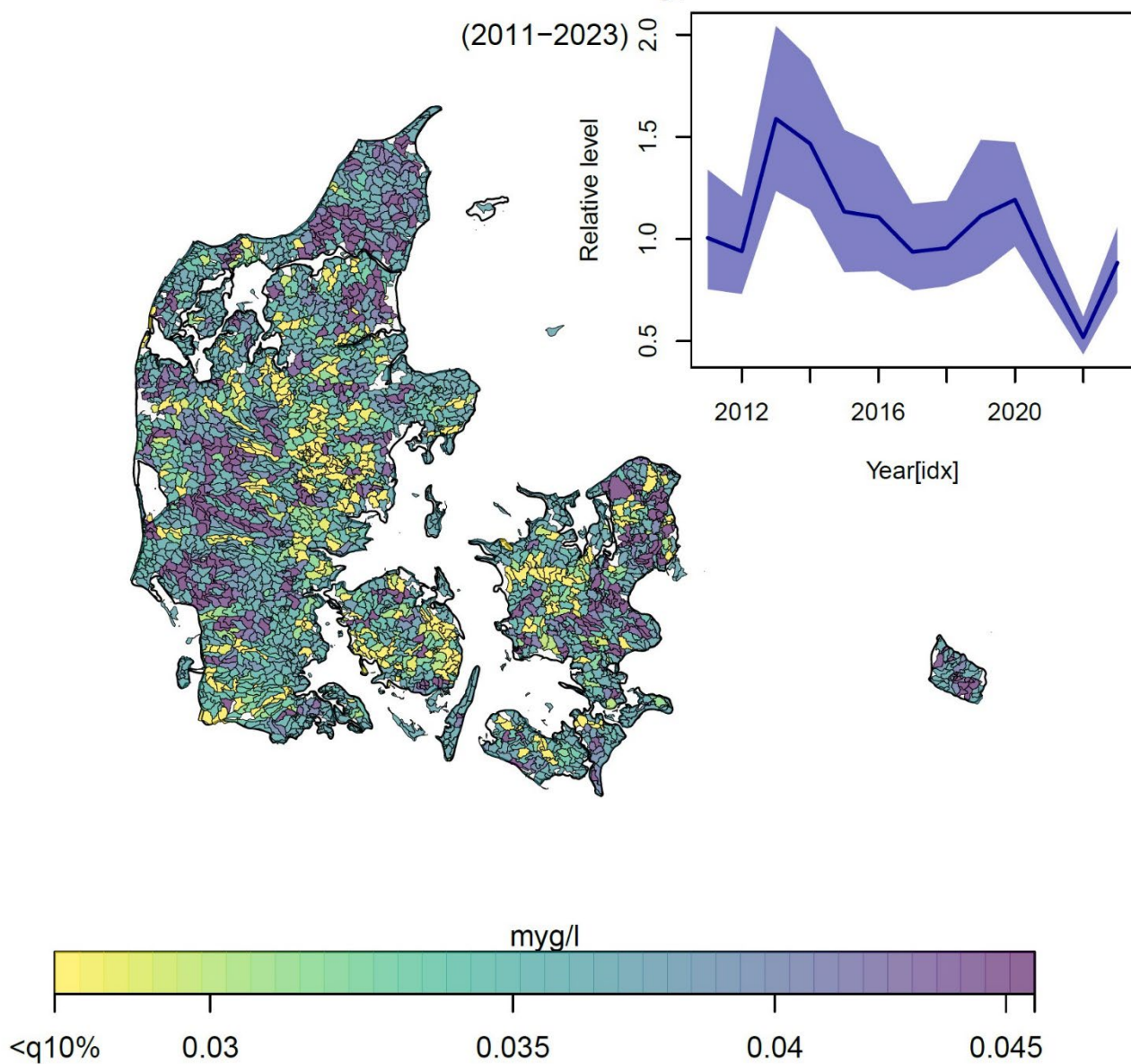
Det generelle resultat for bly ses i Figur 3.16. Bly afviger fra de oven for omtalte metaller ved at punkterne på Figur 3.16 er langt mere spredte. Den spredte punktsky skyldes to faktorer: for det første ligger en del af målingerne under detektionsgrænsen, hvilket i dette plot betyder, at målingen er afsat på y-aksen ved detektionsgrænsen. Det betyder, at punktskyen på y-aksen er trunkeket omkring detektionsgrænsen, med enkelte målinger, som blev analyseret med en lavere detektionsgrænse end flertallet af de andre målinger. Modellen har ingen detektionsgrænse i sin forudsigelse og kan derfor fint forudsige en koncentration til at være under detektionsgrænsen. Det giver en forskel mellem model og målinger for målinger under detektionsgrænsen, som ikke kan tolkes som en reel uoverensstemmelse med model og måling, og værdier for R^2 bliver meningsløse. En anden forklaring, som kan begrunde den spredte punktsky, er, at en historisk kontaminering fra bly i benzin, der stoppede for en del år siden, betyder, at metallet generelt set har haft tid til at fordele sig ud i alle ID15-oplande uden nogle betydende menneskeskabte eller mineralogiske kilder af nyere dato, der kan skabe forskelle mellem de enkelte ID15-oplande. Dette kan igen forklare, at variationen af de enkelte målte koncentrationer i samme opland (SDmeas) er dominerende over variationen mellem oplandene (SDlc). Retentionen for bly på 9 (8.697) svarende til retentionen for arsen.

Figur 3.16. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for bly med brug af målinger fra alle sæsoner (hele året).



På Figur 3.17 ses den forventede mediankoncentration fordelt på ID15-oplandene for bly. Forskellene mellem oplande ses at være beskedne, svarende til at Figur 3.16 viser, at variationen mellem oplande domineres af variationen inden for oplande. De relative tilfældige effekter af år udviser en svag tidstrend med store udsving i intervallet 0.5-1.5.

Forventet koncentration af Bly, alle sæsoner

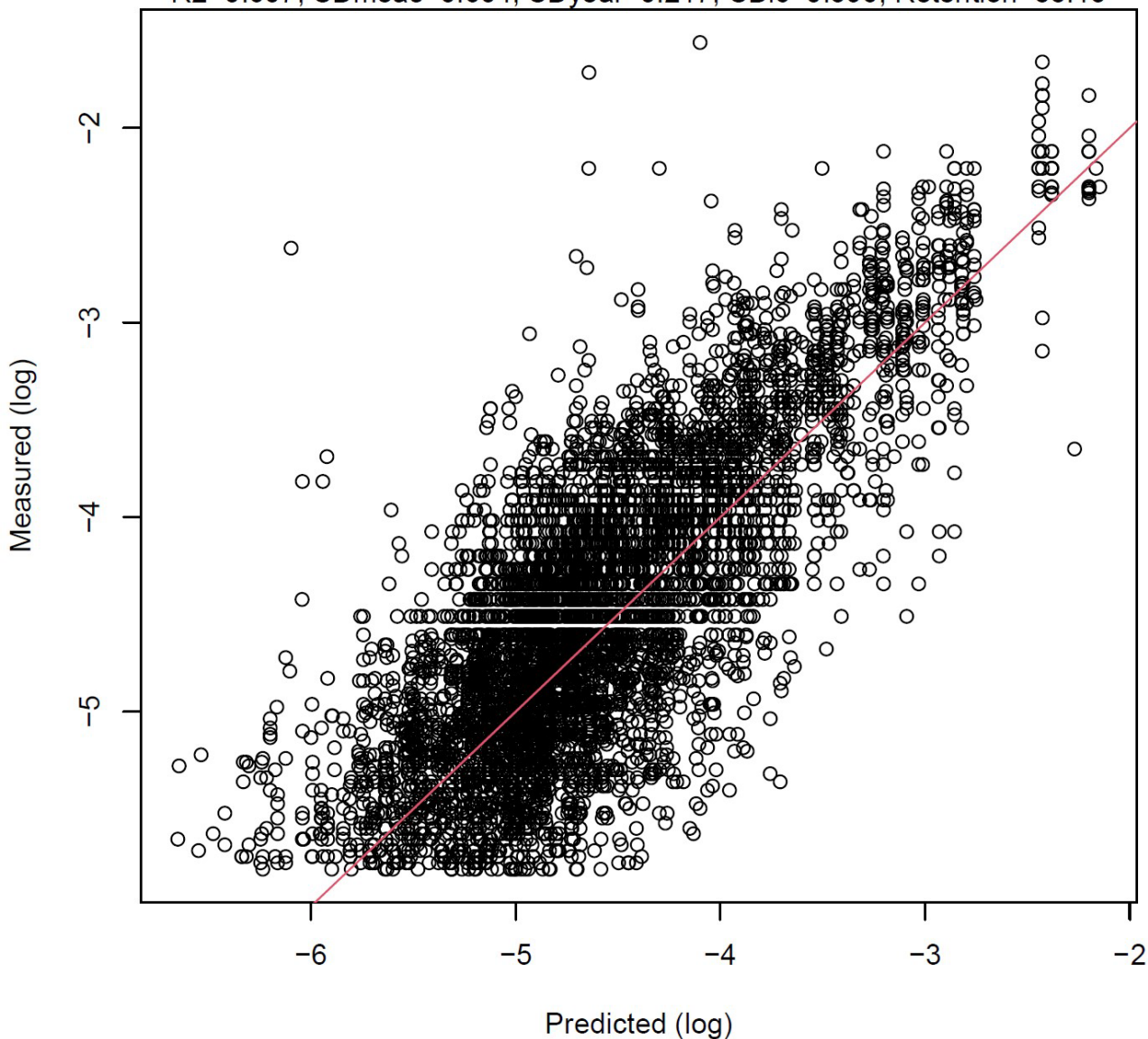


Figur 3.17. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for bly med brug af målinger fra alle sæsoner (hele året).

De generelle resultater for cadmium vises i Figur 3.18, hvor der ses en del spredning af punktskyen, som er trunkeket neden for detektionsgrænsen. Variationen mellem målinger inden for oplandene (SDmeas) har omtrent samme værdi som variationen mellem oplande (SDlc). Cadmium adskiller sig fra de andre metaller ved at have en markant større retention med en værdi på 88 (88.15), hvilket giver en medianfjernelse på omkring 8.2%, og bare middelstore søer vil stort set fjerne cadmium fra vandfasen uanset indløbskoncentrationen.

Forventet koncentration for Cadmium, alle sæsoner (2011–2023)

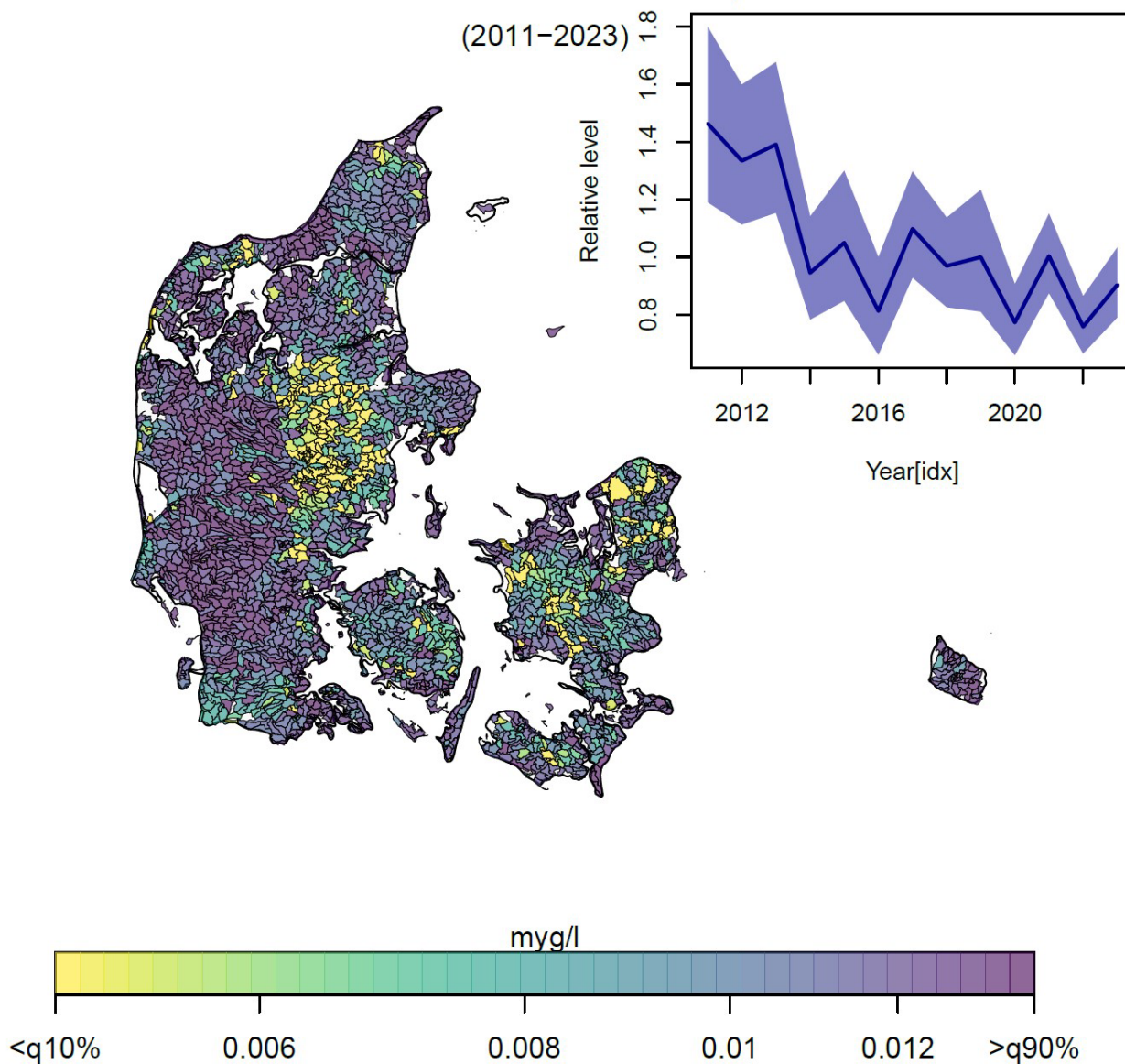
$R^2=0.607$, $SD_{meas}=0.604$, $SD_{year}=0.217$, $SD_{lc}=0.596$, $Retention=88.15$



Figur 3.18. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for cadmium med brug af målinger fra alle sæsoner (hele året).

Den forventede mediane koncentration i ID15-oplandene for cadmium ses afbildet i Figur 3.19, med et stort sammenhængende område med de højeste koncentrationer i Vestjylland og et stykke ned i Sønderjylland. Den tilfældige effekt af år udviser en faldende tendens med relativ størrelse på 1.4 til 0.8, svarende til et fald på omkring 60 % i koncentrationsniveau over hele perioden

Forventet koncentration af Cadmium, alle sæsoner

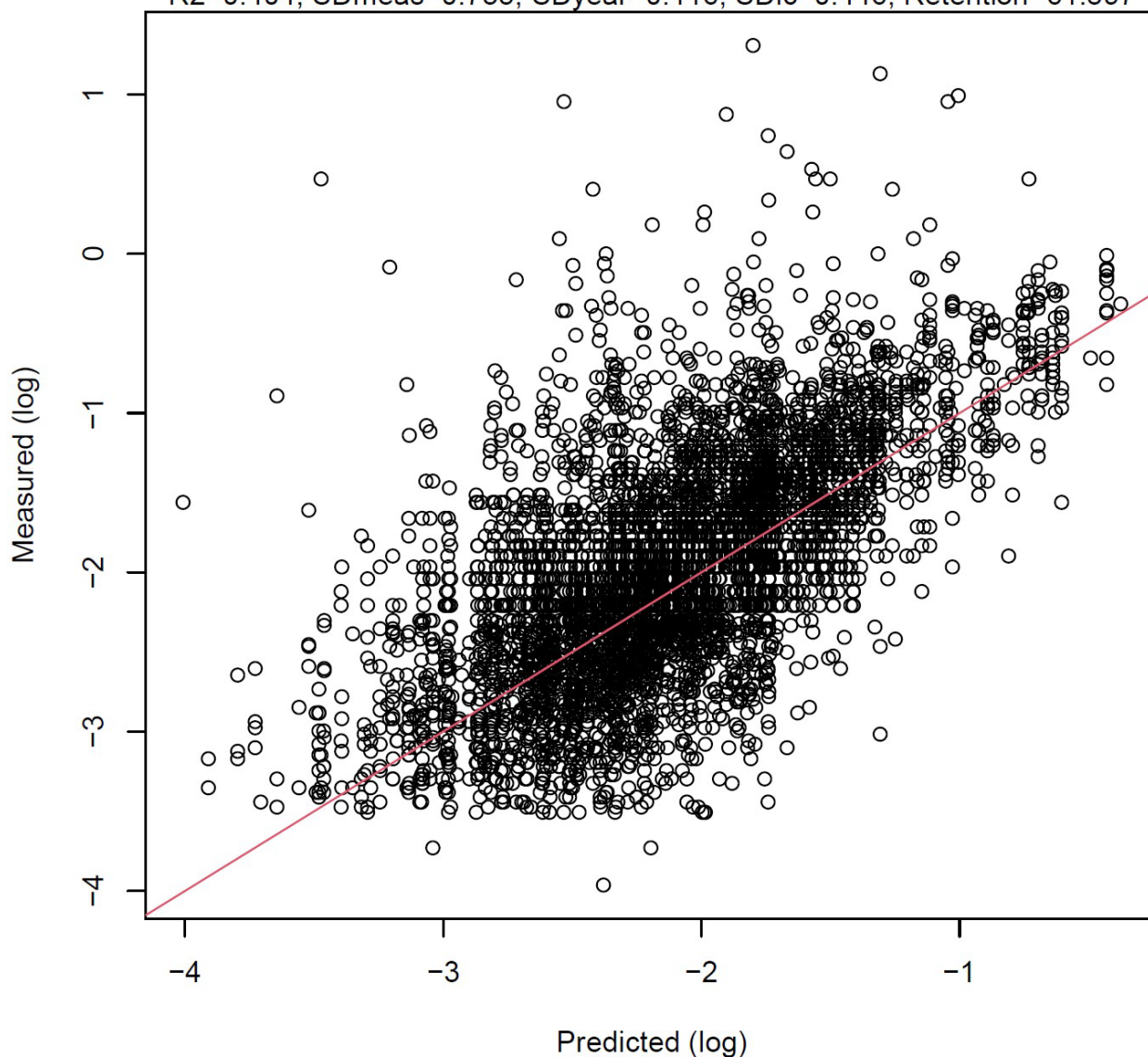


Figur 3.19. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for cadmium med brug af målinger fra alle sæsoner (hele året).

De generelle resultater for chrom vises på Figur 3.20, en relativt spredt punktsky med en variation inden for oplandene (SDmeas) på 0.735, der er noget større end variationen mellem oplandene (SDlc), der har værdien 0.446. Retentionen på 62 (61.897) ligger omkring midt i intervallet blandt metallerne.

Forventet koncentration for Chrom, alle sæsoner (2011–2023)

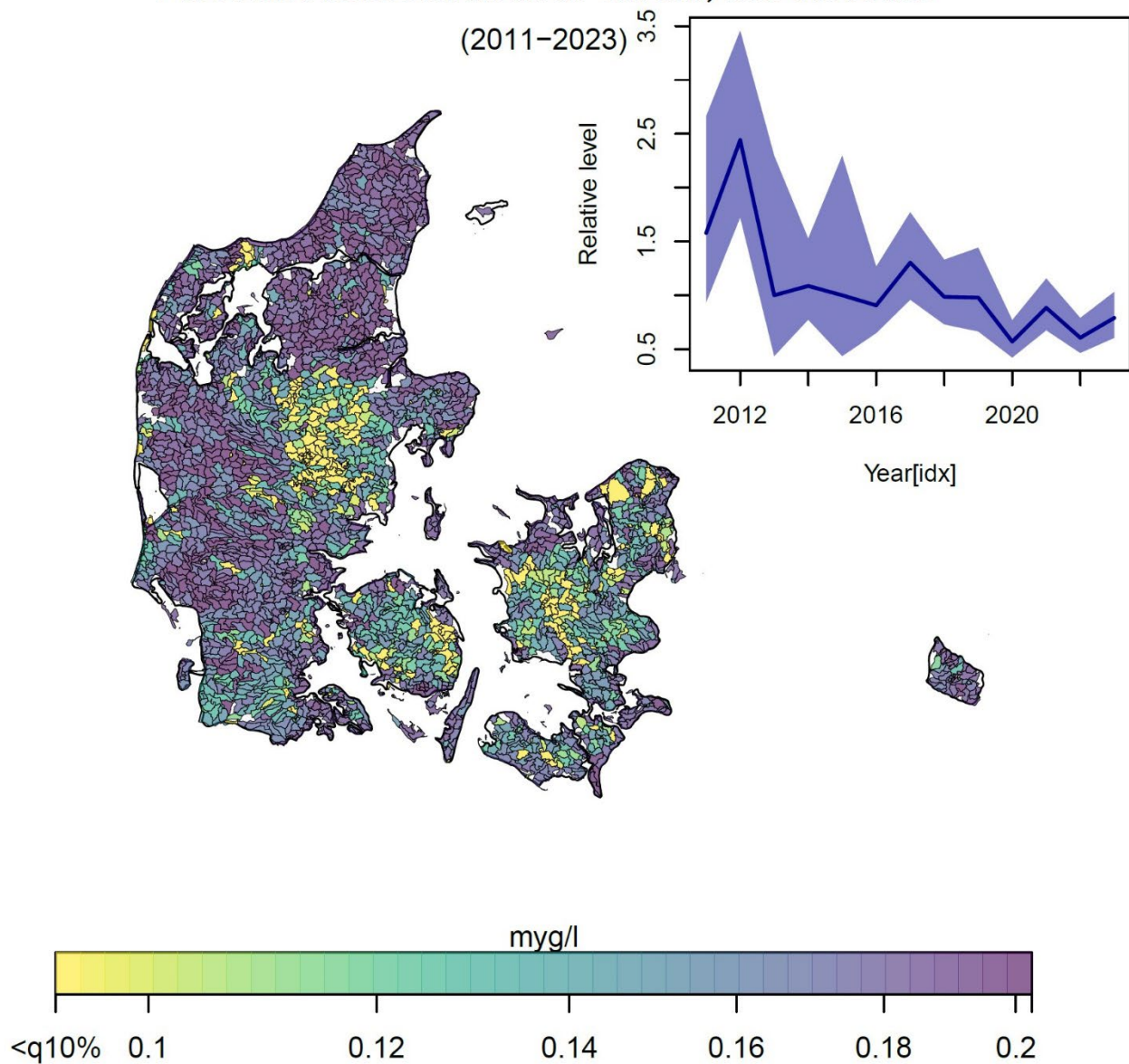
R²=0.404, SDmeas=0.735, SDyear=0.416, SDlc=0.446, Retention=61.897



Figur 3.20. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for chrom med brug af målinger fra alle sæsoner (hele året).

Den forventede mediankoncentration i ID15-oplandene for chrom ses afbildet i Figur 3.22, med de højeste koncentrationer i Vestjylland og Himmerland. Den tilfældige effekt af år udviser en faldende tendens med relativ størrelse fra omkring 2 ned til omkring 0.6 svarende til et fald på omkring 70 % i koncentrationsniveau over hele perioden

Forventet koncentration af Chrom, alle sæsoner

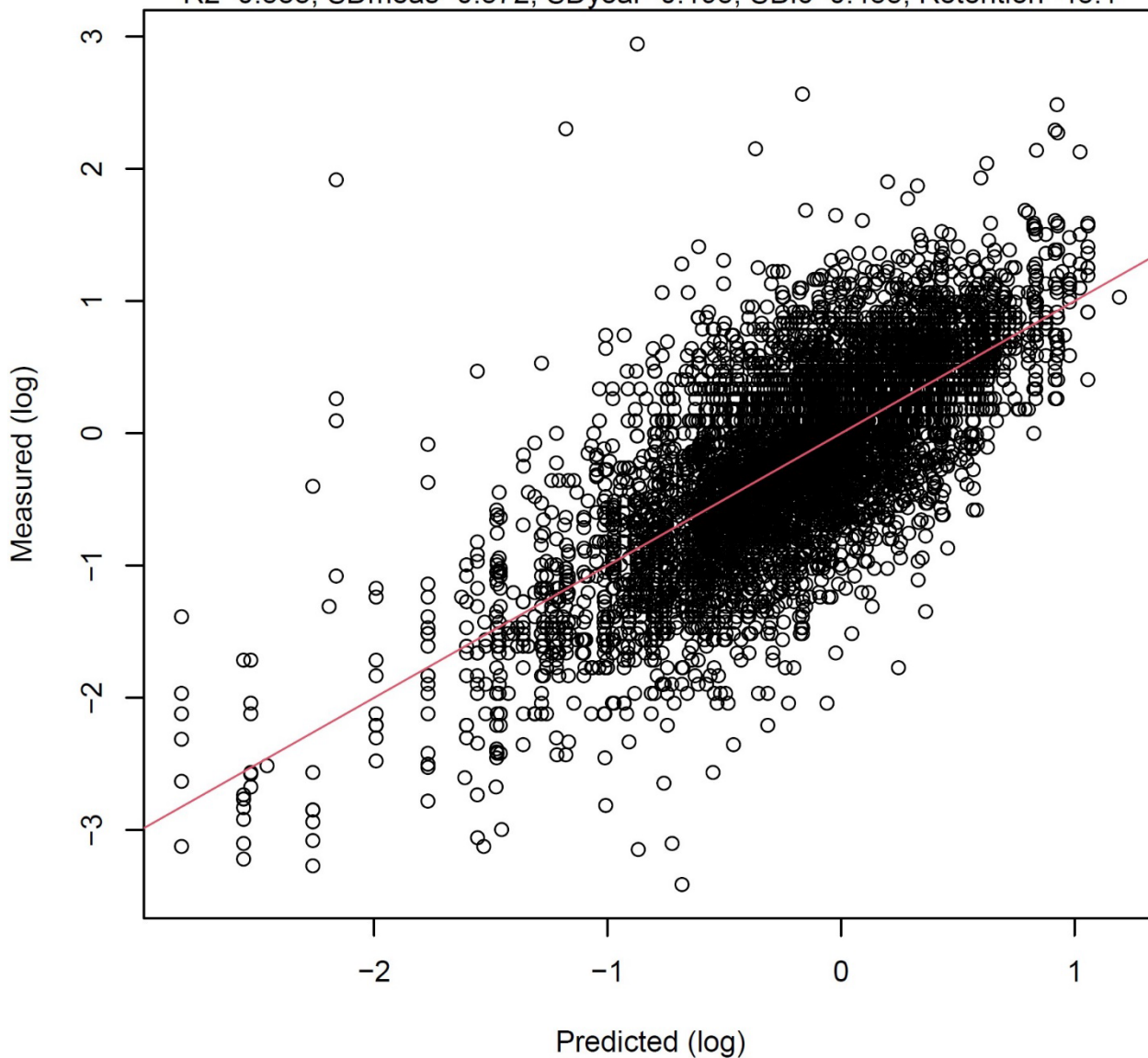


Figur 3.21. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for chrom med brug af målinger fra alle sæsoner (hele året).

De generelle resultater for kobber vises på Figur 3.22, hvor der ses en punktsky med en relativt stor variation inden for oplandene (SDmeas) på 0.572, der svarer til variationen mellem oplandene (SDlc) med værdien 0.455. Retentionen på 52 (52.089) ligger midt i intervallet blandt metallerne.

Forventet koncentration for Kobber, alle sæsoner (2011–2023)

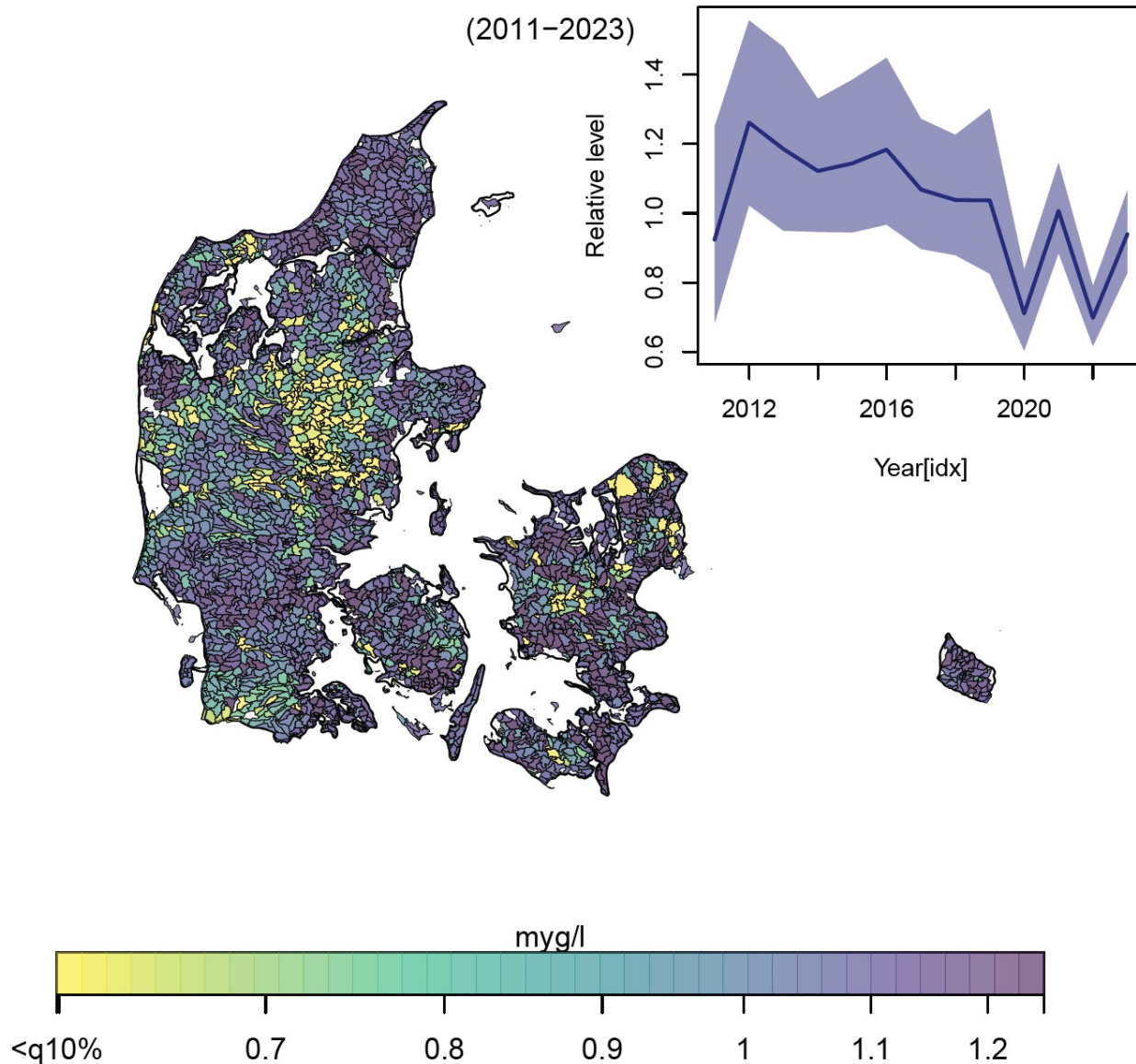
R²=0.533, SDmeas=0.572, SDyear=0.195, SDlc=0.455, Retention=48.1



Figur 3.22. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for kobber med brug af målinger fra alle sæsoner (hele året).

Den forventede mediankoncentration i ID15-oplandene for kobber ses afbildet i Figur 3.23, med en relativt spredt variation med en tendens til at Gudenå's opland har lavere koncentrationer end resten af landet. Den tilfældige effekt af år udviser en svag faldende tendens med relativ størrelse på omkring 1.2 ned til omkring 0.8, svarende til et fald på omkring 30 % i koncentrationsniveau over hele perioden

Forventet koncentration af Kobber, alle sæsoner

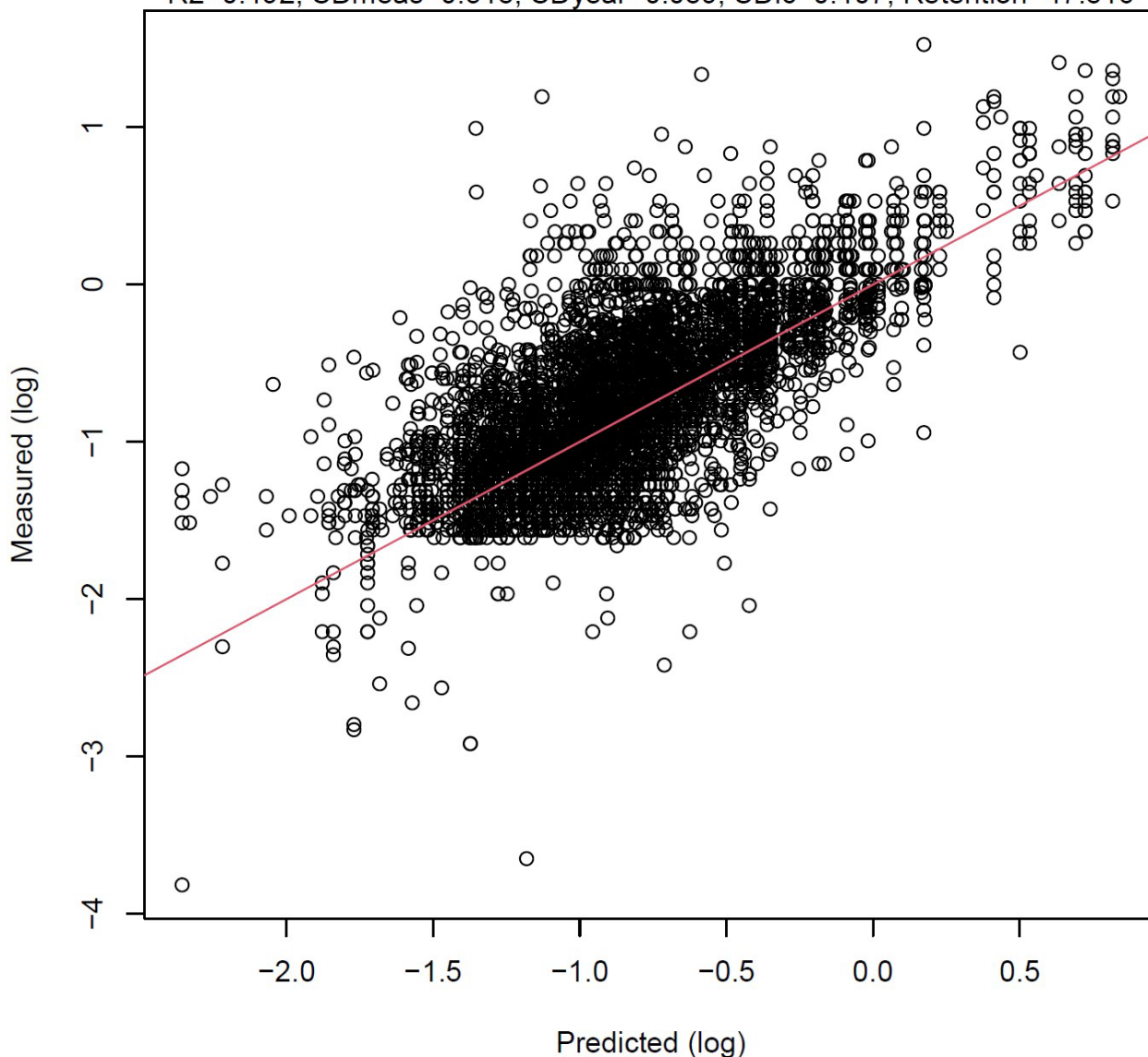


Figur 3.23. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for kobber med brug af målinger fra alle sæsoner (hele året).

De generelle resultater for vanadium vises i Figur 3.24, hvor der ses en punktsky med den største variation inden for oplandene (SDmeas) på 0.518, i forhold til variationen mellem oplandene (SDlc), med værdien 0.407. Retentionen på 48 (47.816) ligger midt i intervallet blandt metallerne. Der ses også en tydelig trunkering nederst i punktskyen i forhold til de målte værdier pga. et betydeligt antal målinger under en detektionsgrænsen.

Forventet koncentration for Vanadium, alle sæsoner (2011–2023)

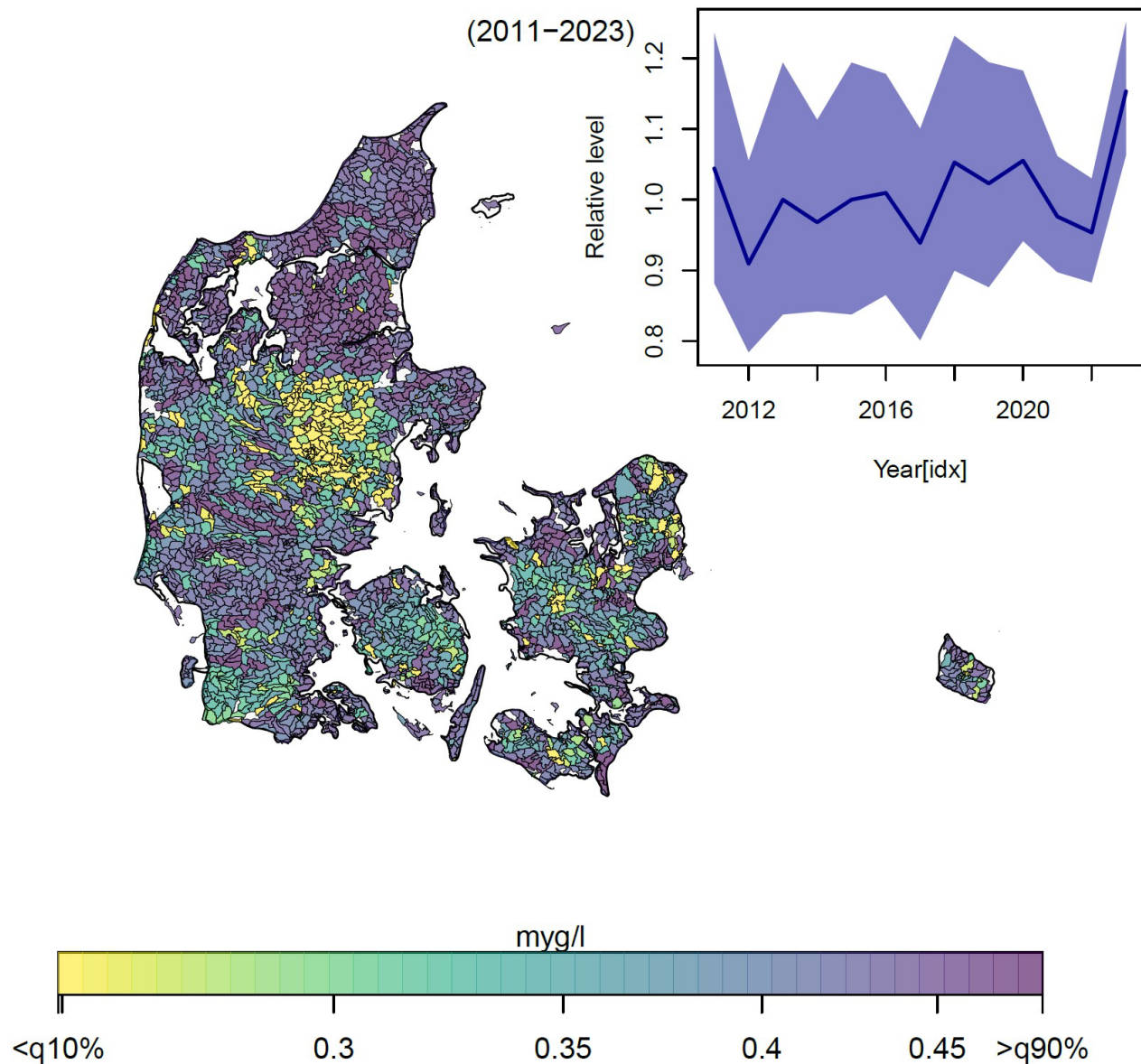
$R^2=0.492$, $SD_{meas}=0.518$, $SD_{year}=0.089$, $SD_{lc}=0.407$, $Retention=47.816$



Figur 3.24. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for vanadium med brug af målinger fra alle sæsoner (hele året).

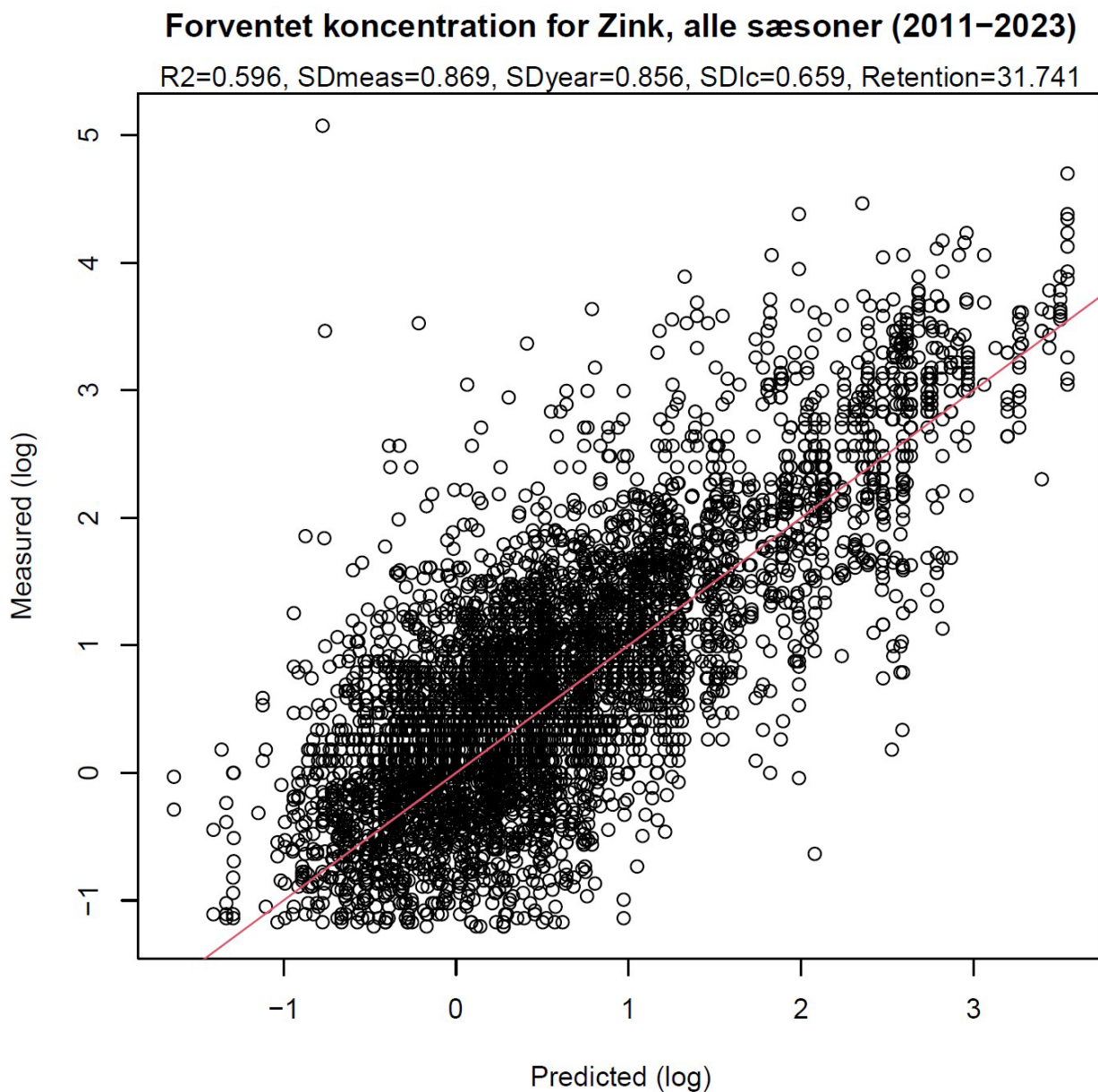
Den forventede mediankoncentration i ID15-oplandene for vanadium vises på Figur 3.25, med de højeste koncentrationer i Vensyssel og Himmerland, mens relativt lave koncentrationer forventes i Gudenåens opland. Den tilfældige effekt af år udviser ingen tendens gennem perioden.

Forventet koncentration af Vanadium, alle sæsoner



Figur 3.25. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for vanadium med brug af målinger fra alle sæsoner (hele året).

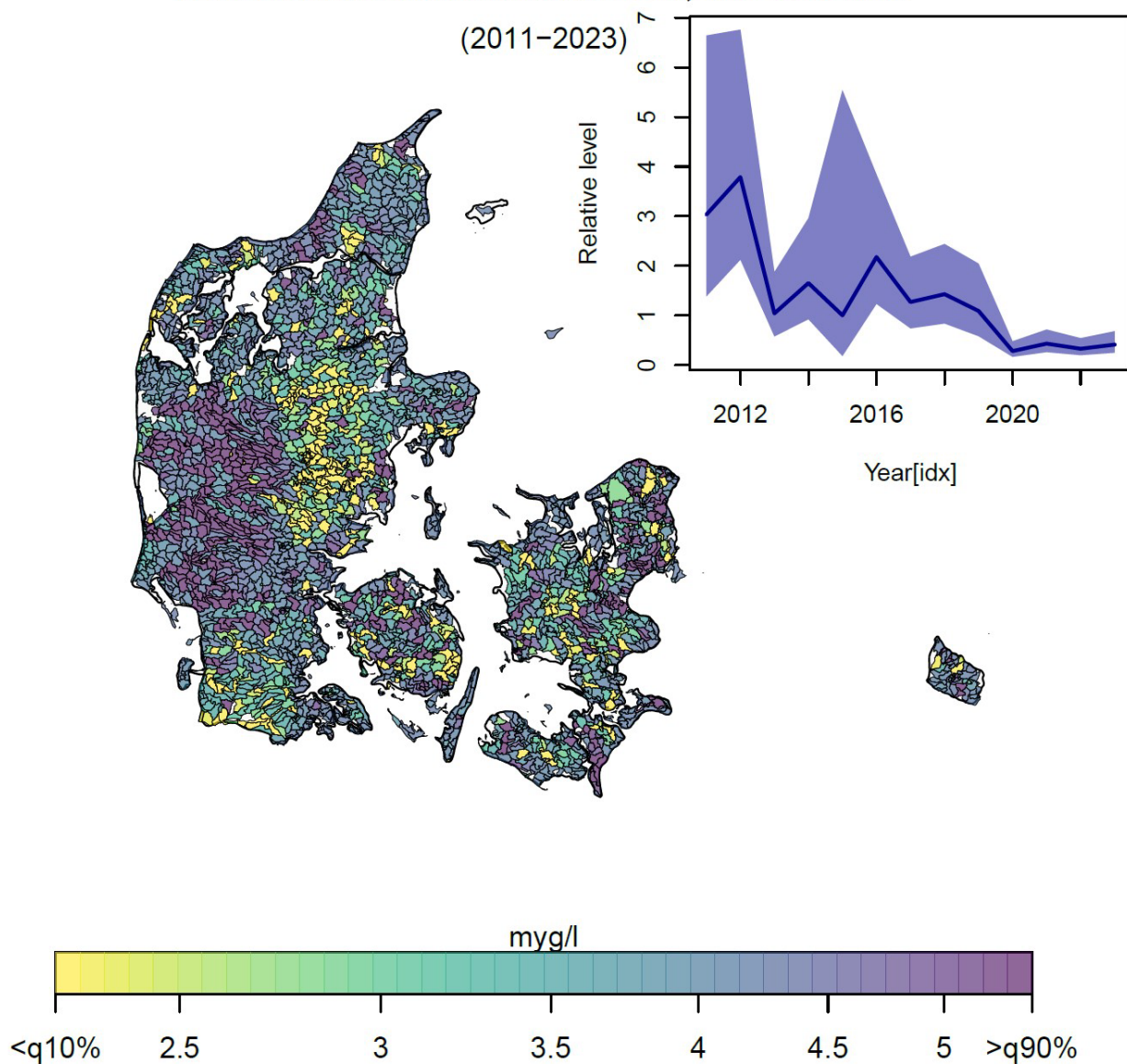
De generelle resultater for zink vises på Figur 3.26, hvor der ses en punktsky med en stor variation både inden for oplandene (SDmeas) på 0.869 og i forhold til variationen mellem år (SDyear) med værdien 0.856, og der er også en betydelig variation mellem oplandene (SDlc) på 0.659. Denne høje værdi for SDyear skyldes et fald i koncentrationen, som ses på Figur 3.26. Retentionen på 32 (31.741) ligger midt i intervallet blandt metallerne.



Figur 3.26. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for zink med brug af målinger fra alle sæsoner (hele året).

Den forventede mediankoncentration i ID15-oplandene for zink vises på Figur 3.27, med de højeste koncentrationer i Vestjylland, Himmerland og Nordøstsjælland. Den tilfældige effekt af år udviser en kraftig faldende tendens gennem perioden med et relativt fald fra ca 3-4 til ca 0,5, svarende til 80-90 % reduktion. Med en så voldsom tidstrend kunne det overvejes at udbygge modellen med en tidsfunktion, der kan beskrive det faldende forløb, som ses på Figur 3.27.

Forventet koncentration af Zink, alle sæsoner



Figur 3.27. Landsdækkende oversigt over forventet mediankoncentration ved udløb af hvert ID15-opland for zink med brug af målinger fra alle sæsoner (hele året).

3.3 Potentiale for inkludering af forklarende variable

Ved at inkludere forklarende variable (X matrice) i modellen vil nogle af disse forskelle mellem ID15-oplande blive forklaret og vil således potentielt kunne levere en adskillelse af ID15-oplandes mediankoncentrationer i intervallet fra 1/2 til 2 gange en landsmedian for alle metallerne. Reelt vil en X matrice dog aldrig være ideel, og der vil derfor altid være en rest variation, som bidrager til SDlc, og X matrixens betydning skal måles ved at se, hvor stor del af SDlc, der kan forklares med X matricen.

Selvom en X matrice i runde tal kan beskrive den samme relative variation mellem ID15-oplande for alle metaller, betyder det ikke, at en X matrice er lige relevant for alle metaller. Det kan ses ved at betragte den tilfældige variation mellem målinger foretaget i samme opland (SDmeas) med variationen mellem oplande (SDlc), hvor der er stor forskel mellem metallerne. For bly er SDmeas (0.91) således betydeligt større end SDlc (0.47), hvilket betyder, at det ikke er muligt at beskrive andet end en mindre del af den målte variation. Dette er i modsætning til barium, hvor SDmeas (0.20) er noget mindre end SDlc (0.48), så her er det potentielt muligt at eliminere den målte uforklarede variation. Forholdet mellem SDlc og SDmeas angiver, hvor meget variation en X matrice potentielt kan beskrive sammenlignet med variationen mellem enkelte målinger i samme ID15-opland. En høj værdi viser således høj relevans for at afsøge mulige X matricer. Dette forhold viser, at en X matrice er mindst relevant at undersøge for bly (0.51) og mest relevant for barium (2.06).

Tabel 3.1. Oversigt variation i og mellem ID15-oplande

Metal	SDmeas	SDlc	SDlc/SDmeas
Arsen	0.31	0.38	1.24
Barium	0.20	0.41	2.06
Bly	0.91	0.47	0.51
Cadmium	0.60	0.60	0.99
Chrom	0.74	0.45	0.61
Kobber	0.57	0.46	0.80
Nikkel	0.44	0.54	1.23
Vanadium	0.52	0.41	0.79
Zink	0.87	0.66	0.76

4 Konklusion og perspektiv

Som anbefalet af Sørensen et al. (2024) er MetalStat som en del nærværende projekt blevet omkodet i RTMB og har dermed fået en hurtigere beregning af de statistiske variable. Konklusionen på dette arbejde er, at der nu ligger en meget effektiv model med mere generel anvendelighed for andre stoffer end metaller. Modellen vil i fremtidige rapporter blive benævnt StatSurf. I denne rapport er StatSurf fittet til monitoringsdata for ni metaller med visning af generelle resultater, som dels kvantificerer, hvor godt ID15-oplande kan bruges til at adskille forskellige koncentrationsniveauer, og dels kortlægge mediankoncentrationer geografisk for alle ID15-oplande.

StatSurf kobler ID15-oplandene sammen under estimering af medianer for de enkelte oplande, hvilket betyder, at målinger i et vandløbssystem får mere generel betydning end bare at være et udtryk for tilstanden i det punkt, der er målt. Dette princip kan derfor udnytte monitorings data langt mere effektivt og estimere geografiske forskelle uden brug af en X matrice. Det betyder ikke, at en X matrice er overflødiggjort, men at informationsniveauet af monitoringsdata er øget.

Output fra MetaStat er genberegnet for de fem metaller (cadmium, bly, kobber, nikkel, zink) med brug af StatSurf uden brug af X matricen med opdaterede måldata til og med 2023. Denne gruppe af metaller er blevet udvidet med arsen, barium, chrom og vanadium. En model for metaller i biota, såsom kviksølv, blev ikke som planlagt færdigudviklet i rapportperioden pga. en forsinkelse i udarbejdelsen af korrelationen mellem oplande for vandanalyser, der ikke levede tilstrækkelig med projekttid til at færdiggøre en beskrivelse af rummelig korrelation, som tager hensyn til dels, at biota kan flytte sig aktivt rundt i oplandet, og dels, at optaget og frigivelsen mellem biota og vandfasen har en lang tidsfaktor sammenlignet med transportprocesser i vandfasen.

Desuden bliver effekter af søer inddraget i modellen ved udviklingen af en retentionsmodel, så søer og vandløb nu bliver modelleret i samme model, hvor effekten af en sø beregnes ved en hydraulisk kontakttid for det opland, hvori søen er placeret.

Undersøgelsen af effekten af forskellige forklarende variable (X matrice) er også betydeligt forbedret i StatSurf, da estimering af koefficienterne til X matricen er blevet hurtigere. Så selvom StatSurf bruger en X matrice på samme måde som MetalStat, så er det blevet regneteknisk lettere at teste og udvikle X matricen, da estimering af koefficienterne er hurtigere. Det er således muligt med en standard bærbar computer og uden reel forsinkelse pga. af beregningstid at undersøge, om en karakteristisk variable i ID15 oplandene, f.eks. grundtilstrømning, spildevands udledning, eller landbrugsareal, har betydning for forudsigelsen af målte vandkoncentrationer. Det har været prioriteret i 2024 at gøre StatSurf operationel, så en X matrice kan undersøges og anvendes uden væsentlig forsinkelse pga. beregningstid.

StatSurf er en relativt kort kode i R, som beregner hurtigt nok på en almindelig computer, og kræver blot basale regnetekniske kompetencer med at bruge software pakker i R. Dette åbner op for en helt anden implementering af modellen, end det var tilfældet for MetalStat. Dels vil StatSurf kunne bruges i en systematisk undersøgelse af mulige X matricer af forskellige forskergrupper, uden at det behøver være gennem en server eller anden central enhed. Endvidere kan StatSurf anvendes til beregninger direkte af Miljøstyrelsens ansatte, som derfor får nye muligheder for at drage nytte af modellens mange muligheder.

I MetalStat er grundvand taget med i X matricen med brug af en særskilt og relativ simpel regressionsmodel (Sørensen mfl., 2024), der udnytter målinger fra boreprøver til at estimere en generel grundvandskoncentration for de grundvandsmagasiner, der føder vand til ID15-oplandene. Der ligger et stort potentiale i også at erstatte denne regressionsmodel med en mere generel model i RTMB, der kobler grundvands målinger til estimer af en generel koncentration i grundvandsmagasinerne. Dels vil det betyde, at usikkerheden i grundvands-estimerne får reel indflydelse på usikkerhed i estimerne i StatSurf, og dels vil det give en bedre udnyttelse af grundvandsdata.

5 Referencer

Horton, R. E. (1945), "Erosional development of streams and their drainage basins: hydro-physical approach to quantitative morphology", *Geological Society of America Bulletin*, 56 (3): 275-370, doi:10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2, S2CID 129509551

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell. 2016. TMB: Automatic Differentiation and Laplace Approximation. *Journal of Statistical Software* 70:1 - 21.

Sørensen, P.B., Damgaard, C.F., Bak, J.L., Andersen, H.E., Voutchkova, D., Thorling, L, Bjerg, P.L., Holm, P.E., 2024, Videreudvikling og klargøring af Me-talStat for landsdækkende beregninger. Aarhus Universitet, DCE - Nationalt Center for Miljø og Energi, 68 s. - Videnskabelig rapport nr. 606.

6 Bilag

Dette bilag indeholder metodebeskrivelsen for StatSurf, som vil blive publiceret i videnskabeligt tidsskrift vinteren 2024/25.

Statistical model for predicting contaminant concentrations in fresh water from monitoring data

Peter B Sorensen, Anders Nielsen, Peter E Holm, Poul B Løgstrup, mfl., Christian F Damgaard,

Introduction

The pollution by contaminants may cause ecotoxicological effects that have the potential to harm ecosystem structure and function and have detrimental effects on biodiversity (REF). To monitor the possible detrimental effects of contaminants, the concentration of selected contaminants is measured in Danish ecosystems, including many organic and inorganic contaminants in one or more of the three freshwater compartments: water phase, sediment, and biota.

The task of making country wide ecological risk assessment of many different contaminants is challenging and resource constraints hinders optimal monitoring at the national level. In Denmark, 3351 catchment areas are set up in such a manner that water will run out from an area into the next area at a single outlet point (REF). The concentration of contaminants is sampled at a fraction of these catchments areas typically at the outlet points. The sampling is irregular, but typically, the sampling is performed each month across several years. Consequently, to obtain the maximal information of the resources invested in the monitoring activities and provide credible ecological risk assessments at the national level statistical and predictive modelling is required as an integrated part of the monitoring activity.

Traditionally, the concentration of contaminants has been predicted by modelling the flow of water and mass balances of substance without linking a statistical model to specific sampling sites with measured concentrations of contaminants (REF). However, such a prediction method precludes state-of-the-art statistical analyses of the measured concentrations, and the different sources of uncertainty cannot be quantified. Here, we will instead focus on the concentration measurements and make predictions by using the possible upstream sources as explaining variables in a hierarchical statistical model consisting of a measurement model and a process model.

The hierarchical model consists of latent variables that model the concentration distribution at each of the 3351 catchment areas. These latent variables are conditioned by the measured concentrations at the sampling sites using a measurement model, and a process model where the expected water flow between different areas is modelled and information on possible upstream explanatory variables (e.g. degree of farming, soil type, etc.) of each area is analyzed in a regression model.

Importantly, the covariation between different catchment areas is modelled explicitly using the information on the expected water flow between the areas and their respective properties. For example, if a sampling site is downstream from another sampling site, then a positive covariation of the measurements at the two sites is expected, and if the water at the downstream sampling site is mixed with the influx of water from another stream, then the positive covariation is expected to be relatively reduced. Furthermore, if two sampling sites are downstream from areas with similar properties, e.g. degree of farming, then these sampling sites are expected to have positive covariation. Since, the covariation among areas is modelled in a realistic manner by modelling the waterflow among areas, the resulting predictions of the concentrations of contaminants in the different catchment areas and their uncertainties are estimated without bias.

Statistical framework

Measurement model

The measured concentrations are left-censored at a known detection limit (D) and assumed to follow a truncated log-normal distribution, and in the case a measurement is below the detection limit it is modelled by the cumulative log-normal distribution. Furthermore, the concentrations are assumed to vary from year-to-year at the national level. The unknown but true mean values of the log-transformed contaminant concentrations at catchment area i are modelled with

latent variables lC_i , and the measured concentrations at the sampling sites are indexed as $z_{i,y,j}$, where y is the sampling year, and j are replicates from the same area and year. Consequently, the likelihood function of the measurement equation is

$$L(z_{i,y,j}) = \begin{cases} \phi(\log(z_{i,y,j}), lC_i + \delta_y, \sigma_0^2) & , z_{i,y,j} \geq D \\ \Phi(\log(D), lC_i + \delta_y, \sigma_0^2) & , z_{i,y,j} < D \end{cases} \quad (1),$$

where $\phi(x, \mu, \sigma_0^2)$ and $\Phi(x, \mu, \sigma_0^2)$ are density and cumulative density respectively for the normal distribution with mean μ and variance σ_0^2 evaluated in x , and δ_y are random effects of year with density function $\phi(\delta_y, 0, \sigma_y^2)$.

Network of catchments

Within Denmark, a number of non-overlapping catchment are set up A1, A2, ..., AI (I = 3351), in such a manner that water will run out from an area into the next area at a single outlet point (REF). The average size of the areas is 13 km². Each area deliver water to 0 or 1 other area, and each area receive water from 0 or 1, 2, or more other areas. This information is collected in the connection matrix:

$$\Psi_{i,j} = \begin{cases} 1, & \text{if } A_i \rightarrow A_j \\ 0, & \text{otherwise} \end{cases} \quad (2).$$

These connections defines a hierarchy of levels L1, L2, ..., Ln (n = 42), where level 1 does not receive water from any other area, and Li receives water from an area of level Li-1. The highest level is 42, but 90% of areas are at level 6 or lower.

Water flow model

Using the information on how water flows between the catchment areas and the properties of the areas the relationship between the concentration in different areas may be expressed as,

$$lC_i = \begin{cases} \log(q_i + \sum_{k=1}^K \beta_k X_{k,i}) - \log(Q_i) + \log(\alpha) - \log(1 + \theta Th_i) + \varepsilon_i & , \text{if } A_i \text{ is in L1} \\ \log\left(\frac{1}{\alpha} \Psi'(\mathbf{IC} \odot \mathbf{Q}) + q_i + \sum_{k=1}^K \beta_k X_{k,i}\right) - \log(Q_i) + \log(\alpha) - \log(1 + \theta Th_i) + \varepsilon_i & , \text{if } A_i \text{ is in L2 or higher} \end{cases} \quad (3),$$

where q_i is the volume of water generated in area i , Q_i is the water volume running out from area i , $X_{k,i}$ is the level of explanatory variable k at area i , β_k are the corresponding regression coefficients, α is the unexplained background contaminant concentration entering all catchments, Th_i is the mean hydraulic retention time of area i , θ is the retention parameter, \mathbf{IC} and \mathbf{Q} are vectors of length I, and ε_i models the process residual variation where $\varepsilon_i \sim N(0, \sigma_\varepsilon^2)$. The form of Eq. 3 is derived in Supplements as Eq. 36.

References

Horton, R. E. (1945), "Erosional development of streams and their drainage basins: hydro-physical approach to quantitative morphology", Geological Society of America Bulletin, 56 (3): 275-370, doi:10.1130/0016-7606(1945)56[275:EDOSAT]2.0.CO;2, S2CID 129509551

Kristensen, K., A. Nielsen, C. W. Berg, H. Skaug, and B. M. Bell. 2016. TMB: Automatic Differentiation and Laplace Approximation. Journal of Statistical Software 70:1 - 21.

Supplements

Model for water and contaminant exchange between catchments (g function)

Mass of metal passing through a cross section of a stream or river assuming one dimensional flow is described by:

$$M|_{\Delta t} = \int_0^{\Delta t} \int_A cv \, d\tau \, dt \quad (5),$$

where A is the cross-sectional area of the stream or river, and c and v is the contaminant concentration ($\mu\text{g/l}$) and local one dimensional water velocity in direction of flow (m/s) respectively. Both c and v vary with time and location in a fixed cross-section which may be important to include in transport models locally around sources of contaminants due to potential correlation between concentration and water velocity. However, on a larger scale of the catchment size it is assumed that the concentration level of water and concentration level is independent, i.e. the local concentration of matter does not correlate with the local water velocity. More specifically, the following is assumed:

$$M|_{\Delta t} \approx \bar{C}_{\Delta t, A} Q \quad (6),$$

where the mean concentration in time and over the cross section is defined as:

$$\bar{C}_{\Delta t, A} = \frac{1}{\Delta t} \int_0^{\Delta t} \bar{C}_A dt \quad (7)$$

The volume of water passing the cross section during Δt is defined as

$$Q = \int_0^{\Delta t} \int_A d\tau \quad (8),$$

In Eq. 6, the mean concentration over the cross section is defined as

$$\bar{C}_A = \int_A c d\tau \quad (9),$$

The sampling is assumed to be representative in the way that the expected concentration is equivalent to the mean concentration as defined in Eq. 7. This is typically valid if there are no strong sources of contaminant close upstream to the sampling station and if there are no strong temporal changes in concentration level in form of strong pulse type sources. The concentration C expresses the concentration throughout the cross section over the period Δt , and Q is the volume of water transported over the cross section and over the period Δt . The measured concentrations have been collected as point measurements in time.

As the catchment area model is assumed stationary:

$$\frac{\partial M|_{\Delta t}}{\partial t} \approx 0 \quad (10)$$

The assumption of stationary conditions depends in particular on the period Δt , as there will typically be fluctuations over short periods of time both in water volumes and water concentration. If, on the other hand, Δt is a longer period then over time some of the fluctuations will be equalized. In the model the Δt value is considered as a season (winter, spring summer, autumn).

The exchange of contaminants between catchments needs to include retention (removal) of contaminant in the catchment. For organic contaminants this retention is both due to degradation and to accumulation into unmeasured compartments. For inorganic contaminants like metals the retention is caused by transfer to unmeasured compartments. The modelling of retention is based on a first order removal process that must be combined with a catchment water mixing model. In the catchment the water is distributed between running water (streams/rivers) and steady water (lakes) having different conditions for mixing. In case the ID15 catchment only includes running water then the mixing is limited, while for catchments that includes lake the degree of mixing will be higher. The question is to identify an "effective" mixing for a catchment to cover this heterogeneity in mixing regimes. Two mixing extremes are identified in the following to identify the boundaries for realistic effective mixing:

- Perfect mixed conditions, where the water body in the catchment is assumed to be a homogeneous mixture for the contaminants. The concentration in the outlet is assumed equal to the concentration everywhere inside the catchment. A mass balance for a perfectly mixed box, having first order removal is:

$$\frac{d(C_{out}V)}{dt} = M_{in} - QC_{out} - V\theta C_{out} \quad (11)$$

Steady state assumption:

$$\frac{d(C_{out}V)}{dt} \equiv 0 \rightarrow C_{out} = \frac{M_{in}}{Q+V\theta} = \frac{C_{in}}{1+\theta\frac{V}{Q}}, \text{ using } M_{in} = QC_{in} \quad (12)$$

- Plug flow conditions, where the water is assumed to run through the catchment without any mixing. Thus, the water is assumed to enter the catchment as an independent “parcels” of water having no interaction with each other until the parcels exists the catchment as discharge.

A first order removal model for a parcel of water running along a transect in the catchment is:

$$\left. \frac{d(C)}{dt} \right|_{tranc} = \theta \cdot C \quad (13)$$

The travel time needed to travel from inlet to outlet of the catchment is V/Q , so Eq. 17 can be integrated from $t=0$ to $t=V/Q$ yielding:

$$C_{out} = C_{in} e^{-\theta \frac{V}{Q}} \quad (14)$$

The two approaches reflected in Eqs. 12 and 14 can be compared by comparing the two functions:

$$f1 = \frac{1}{1+x} \quad (15a)$$

$$f2 = e^{-x} \quad (15b)$$

The Taylor expansion of the two functions yields:

$$T_{f1} = 1 - x + x^2 - x^3 + \dots \quad (16a)$$

$$T_{f2} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \quad (16b)$$

The two Taylor expansions (Eqs. 16a and b) are identical for the zero and first order term and thus the two model will be closely similar for “small” values of x . For “large” values of x both Eqs 15a and b will approach zero, predicting a retention that removes all contaminants ($C_{out}=0$), so the models are also identical for “large” values of x . Thus, there may be an intermedia region for x values between “small” and “large” where the two models differ in the prediction and this region is disclosed in the numeric analysis in Figure 1. The two models are seen to be closely similar for values close to one (small x values) as predicted by the Taylor expansions and similarity is seen also for values close to zero (larger x values). However especially in the lower to middle range there exist a deviation between the two models where the complete mixed box model is predicting higher values (lesser retention) than the plug flow model. Real systems will tend to be between the two models where the completely mixed box model is more likely to simulate lakes and the plug flow closer to simulate streams. For ID15 catchments having no lakes the travel time from inflow to out flow is typically short so, in this case, the two models are likely to reproduce each other, while for lakes where the retention time is much higher there may be a risk for the plug flow model to overpredict retention. Based on this reasoning, the completely mixed model (Eq. 15) is thus selected as the model for retention.

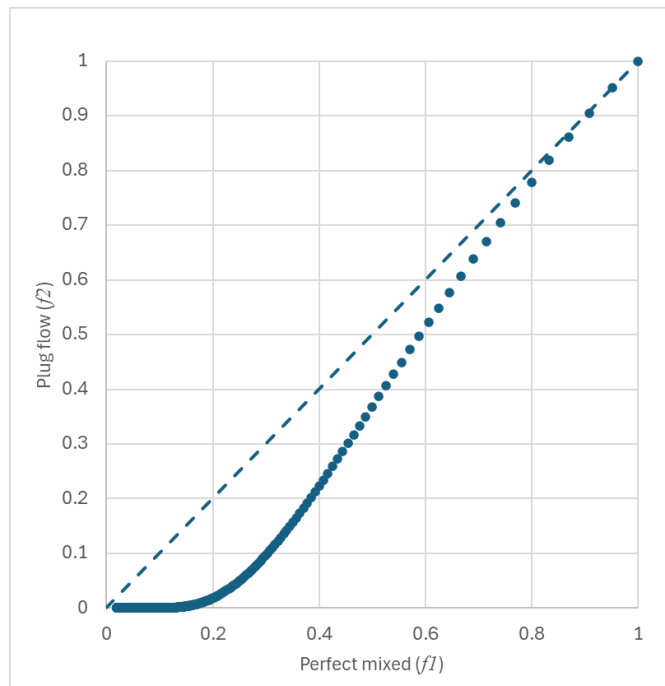


Figure 1. Numerical comparison of respectively plug flow (x axis) and completely mixed box (y axis) as described by the Eqs. 15 a and b. The interval of x values is 0-50 having constant steps ($dx=0.05$).

The first order steady state removal of the contaminant in the catchment area, under perfect mixed condition is derived using Eq. 12:

$$m = M_{in} - M_{out} = M_{in} - \frac{M_{in}}{1 + \theta \frac{V}{Q}} = M_{in} \cdot \left(1 - \frac{1}{1 + \theta \frac{V}{Q}} \right) \quad (17)$$

where m is the mass of removed contaminant in the catchment.

Illustrative model for two catchments

The mass flow between two catchments is illustrated for two catchments where no other catchments discharge to the upstream catchment. The system is illustrated in Figure 2.

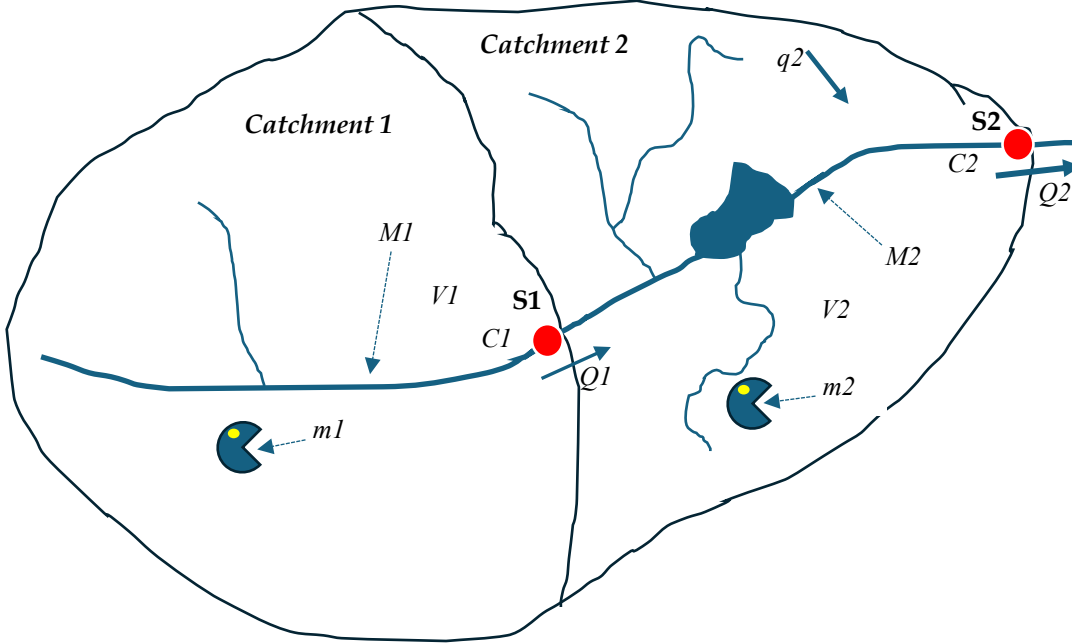


Figure 2. Illustration of the mass flow between two catchments, where water is discharged from Catchment 1 to catchment 2. The measuring stations are assumed to be located at the outlet of the catchments. Signature: M : the source of contaminants in the catchments, m : the mass that is retained/degraded in the catchments, V : free surface water volume (running waters and lakes) in the catchments, C : expected concentration in the catchments.

Catchment 1

The mass balance of Catchment 1 is

$$C1Q1 = M1 - m1 \quad (18),$$

where $C1$ is water concentration at the outlet from Catchment 1, $Q1$ is water volume running from Catchment 1 during the time period Δt , and $M1$ is the mass transport of contaminant to the surface waters in Catchment 1. The mass transport $M1$ is assumed to be a result of different sources in the Catchment. Every source is described in the catchment by a X value, where one source could be groundwater surplus to the water discharge, another one could be wastewater from facilities in the catchment or agricultural area that potentially could contribute with mass from agricultural activity, thus forming a vector X . All those factors are potential contributions, with unknown real impact on the mass balance for the catchment, so the X vector is weighted with a β vector having a coefficient for each source. The value of β is estimated using the statical model including uncertainty estimates. So, the mass transport to the water in Catchment 1 is:

$$M1 = \beta X_1, \quad \beta_i \geq 0 \quad (19),$$

where all elements in vector β are positive and the values will be estimated by the model, however, this estimation has a challenge as illustrated in the following. In case the retention is small ($m \approx 0$) the log concentration for Catchment 1 is:

$$\log(C1) = \log(\beta X_1) - \log(Q1), \quad m1 = 0 \quad (20)$$

If the task is to estimate the values of β using measurements for $C1$ then it is an open problem due to the following rule for $s > 0$:

$$\log(\beta X_1) = \log\left(s^{\frac{1}{s}} \beta X_1\right) = \log(s) + \log(\beta' X_1) \quad (21)$$

So, both β and β' can fit the measurement equally well and the value $\log(s)$ for all $s > 0$ will generate a floating intercept of the statistical model that will make it impossible to find a unique value of the parameter vector β . Thus, the estimation of intercept and effects of e.g. years in the statistical model of IC will interfere with the estimation of a β vector. To solve this challenge first element of X_1 ($X_{1,0}$) is defined to be the volume of water generated in Catchment 1 ($Q1$) and the related parameter $\beta_{1,0}$ is fixed to be one. Thus, only one value α satisfies:

$$\log(\beta X_1) = \log(\alpha) + \log(Q1 + \sum_{k=1}^P \beta_k X_{1,k}) \quad (22)$$

In other words: It is not possible to adjust the value of α and obtain the same results for $\log(\beta \cdot X_1)$ by adjusting the value of $\beta_{k \geq 1}$ if $Q1$ is non-zero for at least a fraction of the catchments, and this is obviously the case. In this equation the value of β is fixed having the coefficient value one for $Q1$ ($\beta_0=1$) and the value of α has an interpretation as follows: Let consider the special case where there are no sources other than $Q1$ in the X_1 matrix and no retention ($\theta=0$) then:

$$\sum_{k=1}^P \beta_k X_{1,k} = 0 \Rightarrow \log(C1) = \log(\alpha) + \log(Q1) - \log(Q1) = \log(\alpha) \Rightarrow \alpha = C1 \quad (23)$$

Thus the value of α is the background concentration of the water entering the catchment that cannot be explained by sources listed the X matrix.

The equation for mass source within Catchment 1 is thus:

$$M1 = \alpha(Q1 + \sum_{k=1}^P \beta_k X_{1,k}) \quad (24)$$

From Eq. 17:

$$m1 = M1 \left(1 - \frac{1}{1 + \theta \frac{V1}{Q1}}\right) \quad (25)$$

The mean hydraulic retention time of the catchment water body is defined as

$$Th_1 = \frac{V1}{Q1} \quad (26)$$

Eqs. 18, 24, 25 and 26:

$$C1 = \frac{\alpha(Q1 + \sum_{k=1}^P \beta_k X_{1,k})}{Q1 \cdot (1 + \theta Th_1)} \quad (27)$$

The concentration CI in log form:

$$IC1 = \log(C1) = \log(\alpha) + \log(Q1 + \sum_{l=1}^P \beta_l X_{1,l}) - \log(Q1) - \log(1 + \theta Th_1) \quad (28)$$

For catchment 2

The discharge of water from Catchment 2 is a sum of inflow from Catchment 1 and surplus of water in catchment 2 ($q2$):

$$Q2 = Q1 + q2 \quad (29)$$

The mass balance of contaminants for Catchment 2 is:

$$C2Q2 = C1Q1 + M2 - m2 \quad (30)$$

Using the same form as Eq. 25 for Catchment 2:

$$M2 = \alpha(q2 + \sum_{k=1}^P \beta_k X_{2,k}) \quad (31)$$

It is noticed that the surplus of water (q_2) is used in Eq. 31 instead of the outflow from Catchment 1 (Q_1) in Eq. 27, this difference is because the Catchment 1 is an upper catchment without inflow from any other catchments and thus the discharge from this catchment is equivalent with the surplus. The mass balance assumes that:

$$q_2 \geq 0 \quad (32)$$

The Eq. 32 is valid for nearly all catchments under the Danish conditions and smaller negative values is renumbered to zero having only small effect on the results.

Combining with an equation having same form as Eq. 25, but extended to include the mass coming from Catchment 1:

$$m_2 = (C_1 \cdot Q_1 + M_2) \cdot \left(1 - \frac{1}{1 + \theta T h_2}\right) \quad (33)$$

Combining Eqs. 30, 31 and 33:

$$C_2 = \frac{(C_1 Q_1 + \alpha \cdot (q_2 + \sum_{k=1}^P \beta_k X_{2,k})) \cdot \left(1 - \frac{1}{1 + \theta T h_2}\right)}{Q_2 (1 + \theta T h_2)} \quad (34)$$

The equation of the latent variable IC_2 is the log of Eq. 34:

$$IC_2 = \log \left(e^{(IC_1)} Q_1 + \alpha \cdot (q_2 + \sum_{k=1}^P \beta_k X_{2,k}) \right) - \log(Q_2) - \log(1 + \theta T h_2) \quad (35),$$

where the CI variable is substituted by ICI to make a model in the log space that can support log-normal distributed concentration measurements.

General model for all catchments

Event though the Eq. 35 is only derived for the downstream catchment of two catchments it is possible to generalize the equation to cover any catchment. This generalization is based on the following argument: In Eq. 35 the term “ $e^{(lC1)} \cdot Q1$ ” is quantifying the mass transport into catchment 2 from Catchment 1. In this simple case the Catchment 1 was an upper catchment without any inflow from other catchments, however, the term “ $e^{(lC1)} \cdot Q1$ ” is not depending on what ever there exist an upstream catchment for Catchment 1 or not. Some catchments have inflow from several upstream catchments where the inflow is a sum of all catchments discharging into the downstream catchment. General equations valid for all catchments are defined in the following.

$$lC_i = \log \left(\sum_{j=1}^l \left[\Psi_{i,j} e^{(lC_{j,s,y})} Q_{j,y} \right] + \alpha \cdot \left(q_i + \sum_{l=1}^p \beta_l X_{i,l} \right) \right) - \log(Q_i) - \log(1 + \theta Th_i) \quad (36)$$

Where $\Psi_{i,j}$ is the connection matrix (Eq. 2).

The mean hydraulic retention time is defined as:

$$Th_{i,s,y} = \frac{V_{i,s,y}}{Q_{i,s,y}} \quad (37)$$

VIDEREUDVIKLING AF MODEL FOR MILJØFARLIGE STOFFER I OVERFLADEVAND

Forbedring af statistik model og udvidelse af stofgruppe

Resultater fra StatSurf vises for metallerne arsen, barium, bly, cadmium, chrom, kobber, nikkel, vanadium samt zink, mens kviksølv første bliver modelleret, når der foreligger en metode til at få biota prøver med i StatSurf. Simuleringen af metallerne i denne rapport er udført med en ny hurtig beregningsmetode, hvor alle ID15 oplande en koblet sammen, dog uden brug af X matricen.

ISBN: 978-87-7156-922-3

ISSN: 2244-999X