



UDVIKLING AF MODEL TIL UNDERSTØTTELSE AF VANDPLANLÆGNINGEN I FASTLÆGGELSEN AF KONCENTRATIONSNIVEAUER AF METALLER I DANSKE VANDLØB

Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 484

2022



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

UDVIKLING AF MODEL TIL UNDERSTØTTELSE AF VANDPLANLÆGNINGEN I FASTLÆGGELSEN AF KONCENTRATIONSNIVEAUER AF METALLER I DANSKE VANDLØB

Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi

nr. 484

2022

Peter Borgen Sørensen¹
Christian Frølund Damgaard¹
Poul Løgstrup Bjerg⁴
Hans Estrup Andersen¹
Peter Engelund Holm⁵
Henrik Tornberg¹
Jesper Leth Bak¹
Goswin Johann Heckrath³
Ane Kjeldgaard¹
Patrik Fauser²

¹ Aarhus Universitet, Institut for Ecoscience

² Aarhus Universitet, Institut for Miljøvidenskab

³ Aarhus Universitet, Institut for Agroøkologi

⁴ Danmarks Tekniske Universitet, Institut for Vand og Miljøteknologi

⁵ Københavns Universitet, Institut for Plante- og Miljøvidenskab



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

Serietitel og nummer:	Videnskabelig rapport fra DCE - Nationalt Center for Miljø og Energi nr. 484
Kategori:	Rådgivningsrapporter
Titel:	Udvikling af model til understøttelse af vandplanlægningen i fastlæggelsen af koncentrationsniveauer af metaller i danske vandløb
Forfattere:	Peter Borgen Sørensen ¹ , Christian Frølund Damgaard ¹ , Poul Løgstrup Bjerg ⁴ , Hans Estrup Andersen ¹ , Peter Engelund Holm ⁵ , Henrik Tornberg ¹ , Jesper Leth Bak ¹ , Goswin Johann Heckrath ³ , Ane Kjeldgaard ¹ , Patrik Fauser ²
Institution:	¹ Aarhus Universitet, Institut for Ecoscience, ² Aarhus Universitet, Institut for Miljøvidenskab, ³ Aarhus Universitet, Institut for Agroøkologi, ⁴ Danmarks Tekniske Universitet, Institut for Vand og Miljøteknologi, ⁵ Københavns Universitet, Institut for Plante- og Miljøvidenskab
Udgiver:	Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi ©
URL:	https://dce.au.dk
Udgivelsesår:	Februar 2022
Redaktion afsluttet:	Januar 2022
Faglig kommentering:	Jacob Carstensen og Martin Mørk Larsen
Kvalitetssikring, DCE:	Susanne Boutrup
Sproglig kvalitetssikring:	Charlotte Kler
Ekstern kommentering:	Miljøstyrelsens kommentarer findes her: https://dce2.au.dk/pub/komm/SR484_komm.pdf
Finansiell støtte:	Miljøstyrelsen har ydet finansiell støtte.
Bedes citeret:	Peter Borgen Sørensen, Christian Frølund Damgaard, Poul Løgstrup Bjerg, Hans Estrup Andersen, Peter Engelund Holm, Henrik Tornberg, Jesper Leth Bak, Goswin Johann Heckrath, Ane Kjeldgaard, Patrik Fauser. 2022. Udvikling af model til understøttelse af vandplanlægningen i fastlæggelsen af koncentrationsniveauer af metaller i danske vandløb. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 70 s. - Videnskabelig rapport nr. 484 https://dce2.au.dk/pub/SR484.pdf
	Gengivelse tilladt med tydelig kildeangivelse
Sammenfatning:	Dette projekt har til formål at udvikle en pilot-model til prædiktio n af forekomst af metaller i vandløb og har karakter af et pilotprojekt ved at udvikle en model, som kan afdække anvendeligheden af forskellige beskrivende data generelt til brug også for eventuelle fremtidige modelleringsprojekter for andre stofgrupper end metaller. Det bemærkes, at projektet er et første forsøg på at anvende NOVANA- data i kombination med beskrivende data til at opstille en landsdækkende model til estimering af miljøfarlige forurenende stoffkoncentrationer i vandløbsvand.
Emneord:	metaller, vandløb, koncentrationsniveau, model, MetalStat
Layout:	Grafisk Værksted, AU Silkeborg
Foto forside:	Beate Strandberg
ISBN:	978-87-7156-663-5
ISSN (elektronisk):	2244-9981
Sideantal:	70
Internetversion:	Rapporten er tilgængelig i elektronisk format (pdf) som https://dce2.au.dk/pub/SR484.pdf

Indhold

Forord	5
Sammenfatning	6
Summary	10
1 Indledning	14
2 Statistisk tilgang	16
3 Datagrundlag	17
4 Modelbeskrivelse	18
4.1 Princip	18
4.2 Oplandsmodel	19
4.3 Samlet statistisk model	23
4.4 Samlet oversigt over typer af model input og output	26
5 Mulige kilder til bly, cadmium og nikkel	30
6 Valg af X matrice	34
6.1 Belastningsvariable	34
6.2 Transportvariable	35
7 Modelresultater	38
7.1 Udvælgelse af metaller i forhold til modellerings- potentiale	38
7.2 Vurdering af modellens konvergens	38
7.3 Analyse af målte vandløbsstationer	41
7.4 Landsdækkende estimering	51
8 Konklusioner	58
8.1 Afgrænsning	58
8.2 Brug af MetalStat i tilrettelæggelsen af overvågningsprogrammet for metaller	58
8.3 Forslag til implementering af MetalStat	60
8.4 Evaluering af projektet med henblik på opmærksomheds-og læringspunkter til eventuelle kommende modelprojekter	61
9 Referencer	63
Bilag A, Statistiske grundbegreber	65
Bilag B. Oversigt over data aggregeret til ID 15-oplande	67
Bilag C, parameterestimering	68

Forord

Denne rapport beskriver resultatet af et projekt, hvis formål var at udvikle en pilot-model til prædiktion af forekomst af metaller i vandløb. Projektet er 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, om der er risiko for forekomst af koncentrationer over miljøkvalitetskravene

Projektgruppen består af: Peter Borgen Sørensen og Christian Frølund Damgaard har primært haft ansvar for den matematiske konstruktion af modellen. Hans Estrup har primært bidraget med datagrundlag og miljøkemisk ekspertise. Poul Løgstrup Berg, Peter Engelund Holm, Jesper Leth Bak og Goswin Johan Heckrath har primært bidraget med miljøkemisk forståelse til modellen samt forståelse af mulige kilder. Patrik Fauser har bidraget til forståelse af kilder, Henrik Tornbjerg og Jens Bøgestrand har understøttet datagrundlaget, og Ane Kjeldgaard har lavet rapportens GIS beregninger. Projektet har været fulgt af en gruppe med repræsentanter fra Miljøstyrelsen primært bestående af Maj-Brit Bjergager, Rasmus Dyrmos Nørregaard samt Berit Borksted, og fra DCE har projektet primært været varetaget af Susanne Boutrup. Projektet har været drøftet i denne gruppe. Projektet har desuden været fulgt af en styregruppe med repræsentanter for Miljøstyrelsen og miljøministeriets departement. Styregruppen er i løbet af projektperioden blevet orienteret om projektets fremdrift. Miljøstyrelsen har haft et udkast af rapporten til kommentering.

Projektet er udført som et udviklingsprojekt under rammekontrakten og var desuden understøttet af en tillægskontrakt. Den primære projektperiode var fra juni 2021 til december 2021.

Sammenfatning

Dette projekt har til formål at udvikle en pilot-model til prædiktion af forekomst af metaller i vandløb og har karakter af et pilotprojekt, der afdækker anvendeligheden af forskellige beskrivende data til brug for eventuelle fremtidige modelleringsprojekter for også andre grupper af miljøfarlige forurenende stoffer. Det bemærkes, at projektet er et første forsøg på at anvende målinger af metalkoncentrationer i vandløb fra NOVANA programmet i kombination med en række beskrivende variable med det formål at opstille en model til estimering af koncentrationer af miljøfarlige forurenende stoffer i vandløbsvand.

Som rumlig opløsning vælges ID 15 oplande, der er topografiske oplande på ca. 1500 ha. Disse oplande er allerede defineret ud fra topografi og koblet hydrologisk til hinanden for hele Danmark og vil derfor danne en naturlig basis for en rumlig model, hvor en rumlig analyse skal relateres til ID 15 oplande opstrøms udvalgte vandløbsvandområde. De etablerede vandløbsstationer under NOVANA er typisk placeret tæt ved udløbet fra et ID 15 opland. Nogle ID 15 oplande vil således føde vand til vandløb, der er målt for metaller i NOVANA, mens andre ID 15 oplande vil være uden for disse områder.

Der er valgt en data-drevet (statistisk) tilgang, hvor de eksisterende data fra NOVANA-programmet styrer opstilling og parametrisering af modellen. Modellen kan håndtere usikkerheden relateret til både prøvetagningen og den kemiske analyse samt usikkerheden ved modellens forudsigelser af koncentrationsniveauer for ikke overvågede vandløbsvandområder. Der er to hovedkilder til input:

- Målinger af metalkoncentrationer i vandfasen fra udvalgte vandløbsstationer under NOVANA opdelt på sæson og år.
- Oplysninger (uafhængige variable) om forhold i alle ID 15 oplande, der formodes at kunne have betydning for metalkoncentrationen i vandfasen, såsom f.eks. landbrugsdrift, jordbund og spildevandsudløb.

Modellen vil løbende kunne opdateres og dermed forbedres både i forhold til nye målinger og i forhold til nye oplysninger om ID 15 oplandene. Modellen har pga. sin statistiske opbygning fået navnet MetalStat.

Beregningsarbejdet for MetalStat kan inddeles i en fase, hvor modellens forudsigelser skal konvergere mod den mest realistiske sammenhæng mellem model og målinger, samt derefter en simuleringsfase, hvor modellen giver prædiktio-
ner jf. de forespørgsler, der måtte være interessante. Da der er mange variable i modellen, tager det lang tid for den at konvergere, men når det er overstået, vil modellen kunne give prædiktio-
ner momentant. Den lange tid, det tager modellen at konvergere, har besværliggjort udviklingsarbejdet i den relativt kort projektperiode, da enhver rettelse og ændring betyder, at modellen skal starte helt eller delvist forfra med at konvergere, og der er sket løbende justeringer som en naturlig del af udviklingsarbejdet. Modellen blev justeret i sin strukturligning op mod projektafslutning i december, og det har betydet, at kun de tre metaller bly, cadmium og nikkel er ved at være konvergeret, men endnu ikke helt er kon-

vergeret i forhold til denne nye tilstand ved projektets afslutning. En anden kørsel med bly, cadmium, nikkel, kobber og zink er også ved at konvergere, men vil kræve mindst en måneds ekstra regnetid, der vil forsætte ind i 2022, da det vil være en fordel at have en færdig konvergeret model som udgangspunkt for en evt. driftsfase, hvis en sådan besluttet. Grundet de ovenstående forhold, herunder den lange konvergenstid, blev den oprindelig liste af metaller reduceret til de fem metaller bly, cadmium, nikkel, kobber og zink i første omgang, og for at få beregninger så tæt på at konvergere som muligt blev omfanget i nærværende projekt yderligere fokuseret til bly, cadmium og nikkel. Det anbefales, at de øvrige metaller barium, chrom, kviksølv og arsen først inkluderes i modellen, når de første fem metaller er bragt i drift. Beregningerne med at konvergere fortsætter efter aflevering af denne rapport for at kunne understøtte en evt. kommende modelleringsaktivitet. De konkrete beregningsresultater i denne rapport må derfor ikke tages til indtægt for tolkning af metallers forekomst, da modellen i skrivende stund er ved at konvergere.

Projektet har haft fokus på vandkoncentrationen af metaller, da det er her, der er flest data, men andre matricer, såsom sediment, kunne godt inkluderes i modellen, når modellering af vandkoncentrationen er implementeret. Modelens multivariate struktur betyder, at målinger i forskellige miljømatricer kan understøtte hinanden med en samlet viden om et stofs forekomst.

Fire hovedspørgsmål blev opstillet fra projektets begyndelse, og er her nævnt i kursiv med en kort besvarelse:

1. *I hvilket omfang kan modellen forudsige koncentrationsniveauer for ikke overvågede vandløbsvandområder?* Ud fra de overvågede områder kan koncentrationsniveauerne forudsiges for hele Danmark som fordelinger, der beskriver usikkerheden.
2. *Med hvilken sikkerhed er prædiktionerne opnået ved en sådan modellering bestemt?* Sikkerheden for en prædiktion uden for overvågede områder vil afhænge af, hvor godt det er muligt at beskrive vigtige mekanismer i en landsdækkende oplandsmodel og derfor være meget stofafhængig. Oplandsmodellen vil til gengæld ikke have så stor betydning for prædiktionen inden for de overvågede områder, hvor usikkerheden primært skyldes målingen og dennes evne til at beskrive forholdende i vandløbet det pågældende sted. Begge steder er der en usikkerhed, men den er altså størst uden for de overvågede områder. Rapporten illustrerer, hvordan MetalStat kan kortlægge usikkerheden bag en forudsigelse af koncentrationsniveau uden for det overvågede område, så der ikke opstår en falsk sikkerhed i forhold til forudsagte værdier.
3. *En vurdering af hvorledes nærværende model med fordel kan indgå i tilrettelæggelsen af et fremtidigt overvågningsprogram for metaller i vandløb.* MetalStat kan forbedre evnen til at finde sammenhænge i overvågningsdata, der kan understøtte en bedre forståelse for metallernes skæbne og tilførsel til vandmiljøet. Denne slags analyser vil kunne indgå i tilrettelæggelsen af et fremtidigt overvågningsprogram. Det vil især være oplagt at prioritere et fremtidigt overvågningsprogram mod de ikke-monitorerede områder, som MetalStat forudsiger har største sandsynlighed for de højeste koncentrationsniveauer.

4. *En evaluering af projektet med henblik på opmærksomheds- og læringspunkter til eventuelle kommende modelprojekter.* Modeller for miljøfarlige forurenende stoffer er typisk meget komplekse, og forudsigelser fra disse vil være behæftet med en betydelig usikkerhed. Hvis det efterfølgende besluttet at anvende den udviklede model til beslutningsstøtte, skal der laves en version til driftsfasen, hvilket stiller specifikke krav, dels til en optimal drift og dels til en ekstra kvalitetssikring af selve kodningen, som i øvrigt kan understøttes ved at sammenligne resultater fra pilotmodellen med resultater fra driftsmodellen. MetalStat sætter nye standarder for, hvad der kan konkluderes på baggrund af overvågningsdata, og dermed også krav til de kompetencer, der kræves af dem, som skal anvende modellen. Det betyder, at der med fordel kan laves en faglig struktur bag modellen, som behandler resultater, og som sørger for drift og forbedringer. Den faglige struktur omkring modellen kan evt. være bygget op i fire faglige grupper, der virker i synergi:
- En *brugergruppe*, der dels forestår formidling af modellen til alle relevante i Miljøstyrelsen, dels formulerer de specifikationer, som modellen skal indfri.
 - En *miljøfaglig gruppe*, der løbende diskuterer modellens resultater i forhold til, om de virker miljøfagligt troværdige. Desuden skal denne gruppe sikre, at modellen bliver opdateret med ny viden om de processer, som modellen beskriver. Den miljøfaglige gruppe kan evt. udgive en tilstandsrapport, der konkluderer på baggrund af modelsimuleringer med brug af indrapporterede overvågningsdata.
 - En *driftsgruppe*, der forestår den regnetekniske kobling til datakilder samt dokumentation af modellens beregninger og sikrer, at modellen indfrier de ønskede specifikationer. Driftsgruppen skal sikre, at modellen leverer resultater svarende til forventningerne fra brugergruppen. De udtrukne gennemregninger vil kunne behandles med deskriptiv statistik som f.eks. histogrammer eller fraktilværdier vist for hvert ID 15 opland i GIS. Disse udtræk specificeres af brugergruppen.
 - En *udviklingsgruppe*, der kombinerer miljøkemisk faglig viden med matematisk og regneteknisk viden i samråd med den miljøfaglige gruppe og kommer med projektforslag til mulige udviklingsprojekter. Evaluerer og overvåger modellens resultater og laver forslag til forbedringer og justeringer i forhold til nye datakilder og forbedringer i modellens matematiske opbygning. Denne gruppe kan evt. også vurdere mulighederne for at udbrede modelleringen til andre stoffer og problemstillinger.

Da MetalStat er en multivariat model for metaller, vil viden om et metals koncentrationsniveau per automatik kunne bidrage med viden om et andet metals koncentrationsniveau på samme sted og til samme tid. På samme måde kan en inddragelse af en målt hjælpe-variable, som f.eks. litium, udnytte en mulig sammenhæng med koncentrationen af metal, til at understøtte en viden om niveauet for metalkoncentrationerne. Det bør overvejes, om sådanne hjælpe-variable skal med i den første driftsversion af MetalStat.

De aktuelle beregninger i dette projekt har afdækket nogle vigtige opmærksomhedspunkter, der bør adresseres under opstilling af en model til drift: X matricen bør evalueres grundigt på baggrund af resultatet fra dette pilot-projekt, når modellen er konvergeret, med henblik på at optimere den lineære oplandsmodel, især i forhold til de stærkt ulineære uafhængige variable såsom lerindhold og pH. Det synes muligt at reducere antallet af uafhængige variable, og i den forbindelse bør det overvejes, om andre skal testes, såsom areal med industri, by og ferskvand. Det bør undersøges, om det er muligt at fastholde betydningen af X matrices variable til kun at være positiv, da det vil modvirke estimering af negative koncentrationer og dannelsen af fiktive modsatrettede +/- effekter fra forskellige uafhængige variable. Datagrundlaget bør løbende opdateres til de nyeste NOVANA, efterhånden som disse bliver gjort tilgængelige.

MetalStat bruger estimeret tilførsel af grundvand i en ret simpel beskrivelse, hvor en bedre beskrivelse ville være baseret på en statistisk model for grundvandskoncentrationen, der udnytter viden om grundvandskoncentration fra oplandets boreprøver. På denne måde kan grundvandsovervågningen integreres i vandløbsovervågningen. På samme måde som beskrevet for grundvand kan andre matricer, som f.eks. sedimentbundet metal, kobles til MetalStat gennem en ny statistisk (latent) variabel for sedimentkoncentrationen, hvis der kan findes en troværdig funktion, der kobler adsorberet fase til opløst fase, evt.- baseret på ligevægtskemi eller første ordens udveksling mellem sediment og vandfase.

Summary

This project aims to develop a test model for predicting the presence of metals in running waters and is structured as a test project that uncovers the applicability of various descriptive data for use in any future modeling projects for other groups of substances as well. The project is a first attempt to use NOVANA data in combination with descriptive data to set up a model for estimating environmentally hazardous pollutant concentrations in running waters.

As spatial resolution, ID 15 catchments have been selected, which are topographic catchments of approx. 1500 ha each. These catchments are already defined based on topography and are connected hydrologically for the whole of Denmark. Therefore, they form a natural basis for a spatial modelling, where a spatial analysis must be related to upstream ID 15 catchments along the running water. The established stations under NOVANA are all typically located close to the outlet of an ID 15 catchment area. Thus, some ID 15 catchments will feed water to stations measured for metals in NOVANA, while other ID 15 catchments will be outside these areas.

A data-driven (statistical) approach has been used, where existing data from the NOVANA program controls the set-up and parameterization of the model. The model can handle both the uncertainty related to the sampling, the chemical analysis as well as the uncertainty in the model's predictions of concentration levels for unmonitored watercourses. There are two main sources of input:

- Measurements of metal concentrations in the water phase at selected stations under NOVANA, divided into season and year.
- Information (attributes) on conditions in all ID 15 catchments that have been assessed to have a potential effect on the metal concentration in the water phases, e.g. agricultural activities, soil type and wastewater outlets.

The model can be continuously updated with data and knowledge and can, thus, be improved both in relation to new measurements and in relation to the new information from the catchments that enter the ID 15 catchments. Due to its statistical structure, the model has been named MetalStat.

The calculation work for MetalStat can be divided into a phase, where the model predictions must converge towards the most realistic relation between model and measurements, and then a simulation phase, where the model yields result, cf. the queries that may be interesting. Since there are many variables in the model, it takes a long time for it to converge, but once this is completed, the model will be able to give results instantly. The long time it takes for the model to converge has made the development work more difficult in the relatively short project period, as any correction and change means that the model has to start converging from scratch, and ongoing adjustments have been made as a natural part of the development work. But for the project, this has meant that only the 3 metals lead, cadmium and nickel are about to converge by report submission, as the model was adjusted in its structural

equation towards project completion in December and therefore is not completely converged in relation to this new condition. Another run with lead, cadmium, nickel, copper and zinc is also converging, but will require at least one month of extra calculation time, which will continue into 2022, as it will be an advantage to have a finished converged mode as a starting point for a possible operating phase. So, after the initial list of metals was reduced to the 5 metals lead, cadmium, nickel, copper and zinc, and to get calculations as close to converging as possible, these 5 metals have further been reduced to lead, cadmium and nickel in this report. It is recommended not to include the other metals barium, chromium, mercury and arsenic in the model till after the first 5 metals have been put into operation. The convergence calculations continue after the submission of this report to support an upcoming modeling activity. The concrete calculation results in this report must, therefore, not be taken as a realistic interpretation of the occurrence of metals, as the model is still converging at the time of writing.

The project has focused on the water concentration of metals, as this is where most data are available, but other matrices, such as sediment, could be included in the model once the modeling of the water concentration has been implemented. The model's multivariate structure means that measurements in different environmental matrices can support each other with combined knowledge and the presence of a substance.

The following four main questions were posed from the beginning of the project:

1. *To what extent can the model predict concentration levels for uncontrolled watercourses?* Based on the monitored areas, the concentration levels can be predicted for the whole of Denmark as distributions that reflect the uncertainty.
2. *With what certainty are the predictions obtained by such modeling determined?* The certainty of a prediction outside monitored areas will depend on how well it is possible to describe important mechanisms in a nationwide catchment area model, and the uncertainty thus highly depends on the substance. The uncertainty will be greatest outside monitored areas, in which the uncertainty is primarily due to the analytical technique and the capability of the water sampling to reflect the conditions in the sampling station. In both places, there is uncertainty, but it is greatest outside the monitored areas. The report illustrates how MetalStat can map the uncertainty related to a prediction of a concentration level outside the monitored range so that a false certainty does not arise in relation to predicted values.
3. *An assessment of how the present model can be applied beneficially in the preparation of a future monitoring program for metals in running waters.* MetalStat can improve the ability to find correlations in monitoring data that can support a better understanding of the fate of the metals and their supply to the aquatic environment. This type of analysis could be part of the planning of a future monitoring program. It is also natural to prioritize a future monitoring program to have focus on the non-monitored areas that MetalStat predicts are most likely to have the highest concentration levels.
4. *An evaluation of the project to highlight attention and learning points for possible future model projects.* Models for environmentally hazardous pollutants are

typically very complex, and predictions from these are subject to considerable uncertainty. If it is subsequently decided to use the developed model for decision support, a version must be made for the operational phase, which places specific requirements partly for optimal operation and partly for extra quality assurance of the coding itself, which can otherwise be supported by comparing results from the pilot model with results from the operating model. MetalStat sets new standards for what can be concluded based on monitoring data and, thus, also requirements for the competencies required from those who will be using the model. This means that it is advantageous to create a professional structure behind the model, which processes results and seeks operation and improvements. The professional structure around the model could be built up in four groups that work in synergy:

- *A user group* that is responsible for disseminating the model to all relevant parties in the Danish Environmental Protection Agency and formulates the specifications that the model must meet.
- *An environmental science group* that continuously discusses the model's results in relation to whether they appear to be environmentally credible. In addition, this group must ensure that the model is updated with new knowledge about the processes that the model describes. The environmental scientist group may publish a condition report with conclusions based on model simulations, including monitoring data.
- *An operating group* that is responsible for the computational connection to data sources as well as documentation of the model's calculations and ensures that the model meets the desired specifications. The operating group must ensure that the model delivers results corresponding to the expectations of the user group. The extracted calculations can be processed with descriptive statistics such as histograms or fractional values shown for each ID 15 catchment in GIS. These extracts are specified by the user group.
- *A development group* that combines environmental chemical professional knowledge with mathematical and arithmetic knowledge in consultation with the environmental professional group and comes up with project proposals for possible development projects. Evaluates and monitors the model's results and makes suggestions for improvements and adjustments in relation to new data sources and improvements in the model's mathematical structure. This group could possibly also assess the possibilities of extending the modeling to other substances and issues.

As MetalStat is a multivariate model in terms of metals, knowledge of one metal's concentration level will automatically be able to contribute knowledge of another metal's concentration level in the same place and at the same time. In the same way, an auxiliary variable can be used if there is a connection between e.g., the lithium content and a metal, and if the lithium concentration is measured, then it will be possible to contribute knowledge about the metals, given the lithium auxiliary variables are included in the model in line with

the metals. It should be considered whether such auxiliary variables should be included in the first operating version of MetalStat.

The current calculations in this project have uncovered some important points that should be addressed when setting up a model for operation: (1) the X matrix should be thoroughly evaluated based on the results of this project when the model is converged to optimize the linear catchment model. Especially in relation to the strongly nonlinear attributes, e.g. the clay content and pH. Because it seems possible to eliminate some of the selected attributes to a smaller number, it should be considered whether others should be tested, e.g. areas of industry, city and fresh water. It should be investigated whether it is possible to maintain β to be only positive, as it will counteract the estimation of negative concentrations and the formation of fictitious opposite + / - effects from different attributes. The database should be continuously updated with the latest NOVANA data as they become available, which means that an operating model must expand its data base to also include NOVANA data from 2021.

MetalStat uses estimated groundwater supply in a rather simplified description, where a better description would be based on a statistical model of the groundwater concentration that utilizes knowledge of groundwater concentration from the catchment area's drilling samples. In this way, groundwater monitoring can be integrated into watercourse monitoring. In the same way as described for groundwater, other matrices, such as sediment-bound metal, are coupled to MetalStat through a new latent variable for sediment concentration, if a credible function can be found that links adsorbed phase to dissolved phase, possibly based on equilibrium chemistry or first-order exchange between sediment and water phase.

1 Indledning

Miljøstyrelsen gennemførte i 2020 et projekt med det formål ud fra eksisterende viden at afklare mulighederne for udvikling af matematiske modeller, der kan understøtte vandplanlægningen i fastlæggelsen af koncentrationsniveauer af miljøfarlige forurenende stoffer (MFS) i vandmiljøet med henblik på at understøttet risikovurdering, tilstandsvurdering og indsatsprogrammer for MFS i vandområder, for hvilke der ikke foreligger konkrete målinger (Sørensen et al., 2021). Dette arbejde konkluderede, at den samlede opgave med modellering af MFS i vandmiljøet er omfattende og kompleks, bl.a. fordi MFS omfatter mange forskelligartede grupper af stoffer. I projektrapporten (Sørensen et al., 2021) blev det derfor anbefalet at fokusere opstarten på modeludvikling på et eller flere af en række selvstændige delprojekter, som på baggrund af den eksisterende viden, herunder eksisterende modeller og tilgængelige data, er vurderet at være de mest oplagte til at påbegynde en proces hen mod modellering af MFS i overfladevand. Et af de i projektrapporten skitserede delprojekter fokuserer på forekomst af metaller i vandløb, mens de øvrige delprojekter omhandler pesticider og stoffer fra jordforureninger (eksempelvis kulbrinter og chlorerede opløsningsmidler). Miljøstyrelsen har vurderet, at den umiddelbart tilgængelige datamængde fra NOVANA-overvågningen er størst for metallerne, og da metallerne samtidig er en relevant stofgruppe ift. det overordnede formål, udgør dette delprojekt udgangspunktet for nærværende projekt.

Projektet har til formål inden for rammerne af eksisterende data at udvikle en pilot-model til prædiktion af forekomst af metaller i vandløb. Projektets primære leverance er således en model og en tilhørende rapport med en evaluering af projektet igennem en adressering og besvarelse af fire hovedspørgsmål:

1. I hvilket omfang kan modellen forudsige koncentrationsniveauer for ikke overvågede vandløbsvandområder. Herunder, men ikke begrænset til, en evaluering af datagrundlaget og identificering af eventuelle mangler, den rumlige opløsning (ID 15 oplandsmodellen), modelvalget, metodevalget samt muligheden for at forfine modellen ved et eventuelt fremtidigt projekt.
2. Med hvilken sikkerhed er prædiktionerne opnået ved en sådan modellering bestemt? Vurderingen heraf skal anvendes i overvejelserne om den videre anvendelse af resultaterne i tilstandsvurderingerne, herunder en vurdering af, om de prædikterede koncentrationsniveauer kan lægges til grund for tilstandsvurdering af vandløbsområder, dvs. ved en sammenholdelse af modellerede koncentrationer af konkrete stoffer med miljøkvalitetskrav jf. bekendtgørelse om fastlæggelse af miljømål for vandløb, søer, overgangsvande, kystvande og grundvand, Bek. nr. 1625 af 19/12/2017.
3. En vurdering af hvorledes nærværende model med fordel kan indgå i tilrettelæggelsen af et fremtidigt overvågningsprogram for metaller i vandløb. Eksempelvis ved udpegning af vandområder med risiko for overskridelser af miljøkvalitetskrav eller som udgangspunkt for en revision af den geografiske udbredelse af vandløbsvandområderne i henhold til tilstandsvurderingerne for så vidt angår MFS.
4. En evaluering af projektet med henblik på opmærksomheds-og læringspunkter til eventuelle kommende modelprojekter.

Projektet har karakter af et pilotprojekt og har blandt andet til formål at af-dække anvendeligheden af forskellige beskrivende data (eksempelvis nedbør, udbragte mængder af husdyrgødning og jordbundstyper) til brug for eventuelle fremtidige modelleringsprojekter for yderligere stofgrupper. Det bemærkes, at projektet er et første forsøg på at anvende NOVANA-data i kombination med beskrivende data til at opstille en egentlig model til estimering af MFS-koncentrationer i vandløbsvand. Projektet kan derfor uanset udfaldet indgå i det videre forløb med modellering af MFS i overfladevand, som minimum ved at afklare, om modellering af koncentrationsniveauer af metaller i ikke-overvågede vandløbsvandområder, for så vidt angår MFS, er mulig. Det er således vigtigt, at projektet hele tiden har det forvaltningsunderstøttende perspektiv for øje, dvs. et overordnet formål om at kunne understøtte vandplanlægningen med hensyn til risikovurdering, tilstandsvurdering og indsatsprogram for MFS.

Modeludviklingen må nødvendigvis tage udgangspunkt i de metaller, der foreligger god viden om i forhold til forekomst i vandmiljøet fra især NOVANA-overvågningen, hvilket omfatter følgende metaller:

- Barium
- Bly
- Cadmium
- Chrom
- Kobber
- Nikkel
- Zink
- Kviksølv
- Arsen.

I denne gruppe indgår således både kobber og zink, der begge i projektrapporten for projektet i 2020 blev anbefalet som fokusstoffer med klare kilder til overfladevand fra dyrkede overflader (Sørensen et al., 2021).

Som rumlig opløsning vælges ID 15 oplande, der er topografiske oplande på ca. 1500 ha (Troldborg et al., 2015). Disse oplande er allerede defineret ud fra topografi og koblet hydrologisk til hinanden for hele Danmark og vil derfor danne en naturlig basis for en rumlig model, hvor en rumlig analyse skal relateres til ID 15 oplande opstrøms udvalgte vandløbsvandområder. De etablerede vandløbsstationer under NOVANA er alle placeret ved udløbet fra et ID 15 opland. Nogle ID 15 oplande vil således føde vand til vandløb, hvor der er målt for metaller i NOVANA, mens andre ID 15 oplande vil være uden for disse områder. Som nævnt oven for udgør projektets endelige produkt en model og en tilhørende rapport. Modeldelen af leverancen omfatter estimer af metalkoncentrationer for de ID 15 oplande, der ligger uden for de målte vandområder, samt angivelse af beregningsmetode og mulig dokumentation for kodning i det omfang, det evt. viser sig nyttigt. Det er således ikke muligt inden for den berammede tid at udvikle egentligt beslutningsstøttende software til brug i administrationen, men projektets resultater kan danne grundlag for senere udvikling af sådan software.

2 Statistisk tilgang

Der er valgt en data-drevet (statistisk) tilgang, hvor de eksisterende data fra NOVANA-programmet styrer opstilling og parametrisering af modellen. Måleprogrammet under NOVANA betragtes som en stor observationsmodel af virkeligheden (målemodel), der ”modellerer” vand-koncentrationerne på de målte stationer. De indsamlede prøver fra en målestation repræsenterer derfor kun fejlbehæftede proxyer værdier for stationens ’sande’ vandkoncentration, der aldrig kan måles direkte og derfor kaldes en latent variabel, fordi den sande værdi altid vil være ukendt. Observationsmodellen kobles sammen med en teoretisk model (strukturmodel inklusiv oplandsmodel), der beskriver logiske og kausale sammenhænge i oplandene, så den samlede statistiske model er en blanding mellem en observationsmodel for dataindsamling og en teoretisk model, der beskriver MFS-koncentrationer ud fra en række parametre. En statistisk model bestemmer variationsområderne for værdisætningen til alle parameterestimer, der indgår i modellen, og har derfor usikkerhedsanalysen som et indbygget resultat. Den statistiske tilgang kan desuden håndtere målinger under detektionsgrænsen og manglende målinger for enkelt-metaller i prøver, hvor andre metaller er målt. Da statistiske metoder er styrende for modellen, vil den i rapporten blive benævnt MetalStat.

Det statistiske princip kan forstås som en trinvis tilgang, dog med den bemærkning, at modellen i sin endelige form vil integrere alle estimer i en samlet estimering. Med udgangspunkt i denne trinvis tilgang defineres to modeller:

1. Observationsmodel, som tager hensyn til analyseusikkerhed, inkl. detektionsgrænser samt usikkerheden ved prøvetagning. Opgaven er at estimere latente variable som udtryk for den sande, men ukendte, værdi af metalkoncentrationerne i vand med brug af aktuelle måledata.
2. Strukturmodel, der søger at forklare forskelle i de latente vandkoncentrationer mellem vandløbsstationer og gennem tid.

I en statistisk model er variable ikke beskrevet som en talværdi, men ved en sandsynlighedsfordeling, hvor variansen beskriver usikkerheden af tilgængelig viden. Denne fordeling gør det muligt at finde en vilkårlig fraktilværdi for en vandkoncentration fra et målt opland.

Et koncentrationsniveau tilskrives en række effekter (årsager), der inddeles efter, hvad der vides om dem:

- Deterministiske effekter (fixed effects) mellem vandløbsstationer. Disse effekter kan henføres til en årstidsafhængighed eller effekter fra kendte forhold i de enkelte ID 15 oplande, beskrevet gennem en oplandsmodel, så som landbrugsaktivitet i oplandet, spildevandstilledning etc.
- Stokastiske effekter (random effects) mellem både vandløbsstationer og år. De stokastiske effekter kortlægger, i hvilket omfang de deterministiske effekter er dækkende for alle de mulige effekter, der styrer variationer i vandkoncentrationer.

3 Datagrundlag

3.1.1 Målte koncentrationer

Som datagrundlag for de statistiske analyser anvendes kvalitetssikrede data fra måling af miljøfarlige forurenende stoffer (MFS) i vandløb ved stationer, der indgik i NOVANA i perioden 2010-2019.

For bly, cadmium og nikkel er der 2298 enkeltprøver fordelt på 174 stationer, der opfylder det opstillede krav om, at der er målt for mindst en af disse tre metaller. Tilsvarende for de fem metaller bly, cadmium, nikkel, kobber, zink er der i alt 2358 enkeltprøver fordelt på 182 stationer, i hvilke mindst en af disse fem metaller er målt. Et overblik over de målinger, der er fortaget, kan ses i Tabel 4.1.

Målingerne er fordelt over 10 år, og et år er inddelt i de fire sæsoner: Vinter (december, januar, februar), forår (marts, april, maj), sommer (juni, juli, august) og efterår (september, oktober, november).

Tabel 4.1. Samlet overblik over antallet af målinger fordelt på metallerne målt under NOVANA.

Metal	Antal målte prøver	Antal prøver med koncentration under detektionsgrænsen
Arsen	535	1
Barium	934	0
Bly	2249	589
Cadmium	2296	219
Chrom	990	1
Kobber	1309	3
Kviksølv	774	550
Nikkel	2247	5
Zink	1127	7

3.1.2 Landsdækkende GIS data

Til at estimere koncentrationsniveauer i strukturmodellen kræves uafhængige variable (X matrice), der giver input til prædiktionerne med landsdækkende data. Disse data kan i princippet tage mange former og f.eks. være resultater fra andre modeller. MetalStat er bygget op omkring data for ID 15 oplande, som er defineret i den nationale kvælstofmodel (Troldborg et al., 2015). Disse oplande dækker hele Danmark og er hver ca. 15 km² store. GIS data fra disse oplande er brugt som X matrice. Det samlede datasæt, der har været til rådighed, fremgår af Bilag B og danner datagrundlag til Kapitel 6. Det skal dog påpeges, at dette datasæt er under løbende udbygning. Denne udvælgelse skal ses som en løbende fremtidig aktivitet mod en stadig forbedret modellering, hvilket illustreres i denne rapport.

4 Modelbeskrivelse

Centrale statistiske grundbegreber er forklaret i Bilag A som Boks 1, 2 og 3. MetalStat er bygget op om estimering af den fælles posteriori sandsynlighedsfordeling, som afsøges med en Metropolis-Hastings algoritme. Denne algoritme anvender iterationer og er derfor ret regnetung, så modellen vil typisk skulle bruge fuld CPU-tid i en længere periode for at konvergere, men når det ønskes, kan resultater fra modellen derefter genereres relativt hurtigt på baggrund af de sidste få tusinde gennemregninger. Så brugeren af modellen vil ikke bemærke, at Metropolis-Hastings algoritmen hele tiden beregner. Hvis der inddrages nye måledata fra f.eks. NOVANA eller andre data for de enkelte oplande, vil det påvirke likelihood-funktionen, der kobler modellen til målingerne (Bilag A, Boks 1) og dermed resultaterne. Der vil således være behov for, at modellen løbende får tid til at konvergere, indtil modellen ikke længere kan forbedre sin overensstemmelse med de målte koncentrationer ved flere gennemregninger med Metropolis-Hastings algoritmen. På denne måde vil der ske en løbende ”oplæring” af modellen, efterhånden som nye data bliver tilgængelige.

4.1 Princip

Målte koncentrationer af miljøfarlige forurenende stoffer i miljøet vil typisk variere log-normalfordelt (Sørensen and Johansson, 2020). Derfor vælges en log-normal model, der forudsætter, at $\log(C_L)$ er normalfordelt. Intuitivt giver en log-normalfordeling god mening for koncentrationsværdier, dels fordi $C_L \geq 0$, og fordi $C_L \rightarrow 0 \Rightarrow \log(C_L) \rightarrow -\infty$. En strukturligning relaterer den latente koncentration (C_L) til en koncentration (C_{opl}), som bliver estimeret i hvert ID 15 opland i Danmark. For de oplande, hvor der er målinger, vil C_L og C_{opl} derved blive fittet mod hinanden i denne strukturligning ved at bruge likelihood funktionen (Bilag A, Boks 1). Det log-normale princip betyder, at det er bedst at relatere C_L og C_{opl} til hinanden i strukturligningen gennem deres forhold:

$$F_s \cdot F_i \cdot F_n = \frac{C_L}{C_{opl}} \quad (1)$$

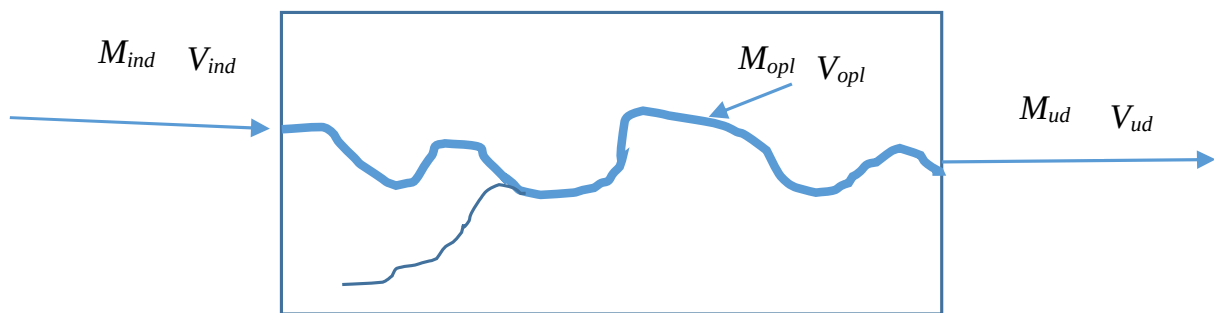
hvor $F_{s=1}$ er defineret som et fast forhold (intercept), der har en værdi for hver af de fire sæsoner ($s = 1, 2, 3, \text{eller } 4$), svarende til hhv. vinter, forår, sommer og efterår. F_i er defineret som påvirkning af forholdet mellem oplande (stationer), og F_n er påvirkning af forholdet mellem år. Ved at tage logaritmen på begge sider af Ligning 1 fremkommer:

$$\log(C_L) = \log(C_{opl}) + \log(F_s) + \log(F_i) + \log(F_n) \quad (2)$$

Det ses fra Ligning 2, at i den log-normale model er forholdene i Ligning 1 udtrykt som additive størrelser. Så hvis f.eks. modellen estimerer $\log(F_2) = 0,5$, betyder en sæson effekt, der giver $\text{EXP}(0,5)=1,64$, forskydning af forholdet mellem C_L og C_{opl} , for sæson 2. Hvis $\log(F_2) = 0$, giver det forholdet $\text{exp}(0)=1$, svarende til at der ikke er nogen relativ forskel for sæson 2.

4.2 Oplandsmodel

Oplandsmodellen skal tilvejebringe oplysninger om ID 15 oplande, der kan forklare tilførslen af metaller til vandløbet samt bruge de afstrømmede vandmængder til at estimere det resulterende koncentrationsniveau. Sådanne modeller kan tage forskellige former, som beskrevet i Sørensen et al., (2021), hvor der især skelnes mellem relativt komplekse deterministiske modeller og mere simple empiriske modeller. Da formålet med MetalStat er at give landsdækkende estimer af koncentrationsniveauer, vælges en relativ simpel empirisk tilgang, da det vurderes, at en mere kompleks deterministisk tilgang vil være for datakrævende og regnetung. Men oplandsmodellen i MetalStat kan løbende forbedres og evt. gøres mere kompleks, hvis det ønskes i fremtiden. For hvert ID 15 opland opstilles en stationær massebalance, hvor følgende masser og vandmængder defineres som transporteret mængde over tidsrummet Δt :



Hvor

M_{ind} : Masse af metal, der løber ind i oplandet fra opstrøms oplande [g / m³]

V_{ind} : Vandvolumen ind i oplandet fra opstrøms oplande [m³]

M_{opt} : Masse af metal i det vand, der tilføres vandløbet fra selve oplandet [g / m³]

V_{opt} : Vandvolumen, der løber til vandløbet fra selve oplandet [m³]. Dette vand dækker over mindre vandløb, overfladisk afstrømning, nedbør og kilder samt tilførsel via dræn og grundvand direkte til vandløbet. Det dækker også over antropogen tilførsel, f.eks. via renseanlæg og overløbsbygværker.

M_{ud} : Masse af metal i vandet, der løber ud af oplandet [g / m³]

V_{ud} : Vandvolumen, der løber ud af oplandet [m³]

Masse af stof, der passerer et tværsnit af vandløbet, er beskrevet ved:

$$M|_{\Delta t} = \int_0^{\Delta t} \int_A c \cdot v \, dA \, dt \quad (3)$$

Hvor A er vandløbets tværsnitsareal, v og c hhv. er den lokale vandhastighed (m / s) og metalkoncentration (µg / l), hvor lokal henfører til, at værdien varie-

rer både med tid og placering i tværsnittet. Inden for tidsrummet Δt forudsættes det, at mængden af vand og stof er uafhængige, dvs. at den lokale koncentration af stof ikke korrelerer med den lokale vandhastighed. Mere specifikt forudsættes følgende:

$$M|_{\Delta t} \approx C \cdot V \quad (4)$$

Middelkoncentrationen C udtrykker koncentrationen i hele tværsnittet over tidsrummet Δt , og V er det volumen vand, der transporteres over tværsnittet og over tidsrummet Δt . De målte koncentrationer er indsamlet som punktmålinger (stikprøver). Det betyder, at det faktisk er c og ikke C , der er målt, hvor MetalStat estimerer C som latente variable med brug af de målte c værdier.

Da oplandsmodellen er stationær, forudsættes det at:

$$\frac{\partial M|_{\Delta t}}{\partial t} \approx 0 \quad (5)$$

Forudsætningen om stationære forhold afhænger især af tidsrummet Δt , da der typisk vil være udsving over korte tidsrum både i vandmængder og vandkoncentration. Hvis derimod Δt er en længere periode, f.eks. et helt år, vil en del af fluktuationerne over tiden udlignes. I det omfang oplandsmodellens forudsætning om stationær tilstand ikke holder i forhold til måleresultater, vil MetalStat tillægge den deterministiske effekt "sæson" og den tilfældige effekt "år" en betydning.

Hvis der sker en tilbageholdelse af metallet i oplandet, forudsættes det at ske efter et første ordens udtryk:

$$\left. \frac{dM}{dt} \right|_{udfældning} = -\rho \cdot M \quad (6)$$

ρ er en tilbageholdelsesrate [1/s]. ρ regnes for en konstant i modellen både tidsmæssigt og i forhold til placering i landet. Løsning af Ligning 6 for den samlede tilførte masse $M_{ind} + M_{opl}$ giver følgende ligning over et tidsrum svarende til den hydrauliske opholdstid i oplandet ($T_{h,opl}$):

$$M_{ud} = (M_{ind} + M_{opl}) \cdot e^{-\rho \cdot T_{h,opl}} \quad (7)$$

Det forudsætter en lineær model svarende til:

$$M_{opl} = \beta \cdot X_{opl} \quad (8)$$

X_{opl} er værdien for en uafhængig variabel i oplandet, der har betydning for tilføjeisen af metal.

Det bemærkes, at β er forskellig for hvert metal, og for hvert metal er β en vektor med et antal elementer, der svarer til antallet af uafhængige variable i X_{opl} , og X_{opl} er en række fra en matrix (X matrixen) med så mange rækker, som der er oplande, og med et antal søjler, som svarer til en konstant arealmæssig tilføjelse, et antal kontinuerte uafhængige variable, samt et antal niveauer for de kategoriske uafhængige variable minus en. Ligning 8 har således følgende form udtrykt ved de enkelte koefficienter:

$$M_{opl} = \beta_0 \cdot A_{opl} + \sum_{k=1}^Z \beta_k \cdot X_{opl,j} \quad (9)$$

A_{opl} er oplandets areal, svarende til $X_{opl,0}$, og Z er antallet af uafhængige variable, og hvor β_0 svarer til en tilførsel af metal per areal opland, som kan anses for at være konstant over hele landet.

Vandbalancen over oplandet giver

$$V_{ud} = V_{ind} + V_{opl} \quad (10)$$

Udløbskoncentrationen fra oplandet beregnes med brug af Ligningerne 7 - 10:

$$C_{ud} = \frac{M_{ind} + \beta \cdot X_{opl}}{V_{ind} + V_{opl}} \cdot e^{-\rho \cdot T_{h,opl}} \quad (11)$$

Hvis der er flere indløb til et opland, kan ligning 11 beregnes med input fra hvert opland, der løber ind:

$$C_{ud} = \frac{\sum_{i=1}^N [M_{ind,i}] + \beta \cdot X_{opl}}{\sum_{i=1}^N [V_{ind,i}] + V_{opl}} \cdot e^{-\rho \cdot T_{h,opl}} \quad (12)$$

N er antallet af vandløb, der løber ind i oplandet, og $M_{ind,i}$ og $V_{ind,i}$ er hhv. masse metal og volumen vand for det i 'te vandløb, der løber ind i oplandet.

Følgende notation bruges for to oplande i serie, hvor opland 1 løber til opland 2:



V_1 og V_2 er vandvolumen, der tilføres vandløbet fra hhv. opland 1 og 2; T_{h1} , T_{h2} er den hydrauliske opholdstid i hhv. opland 1 og 2; X_1 , og X_2 er uafhængige variable for hhv. opland 1 og 2; C_1 og C_2 er metalkoncentrationen for udløb af hhv. opland 1 og 2.

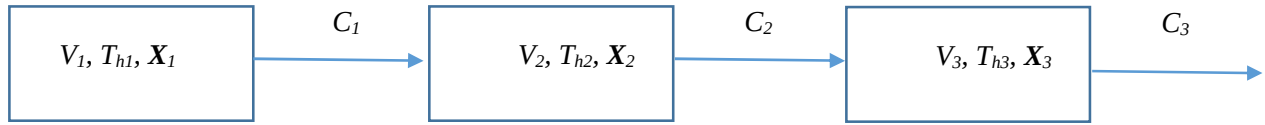
For opland 1 bruges ligning 12, hvor $M_{ind}=0$:

$$C_1 = \beta \cdot \frac{X_1}{V_1} \cdot e^{-\rho \cdot T_{h1}} \quad (13)$$

Koncentrationen for udløbet af opland 2 bestemmes med brug af ligning 12 og 13, hvor $V_1 = V_{ind}$ og $V_2 = V_{opl}$:

$$C_2 = \frac{\beta}{V_1 + V_2} \cdot [X_1 \cdot e^{-\rho \cdot (T_{h1} + T_{h2})} + X_2 \cdot e^{-\rho \cdot T_{h2}}] \quad (14)$$

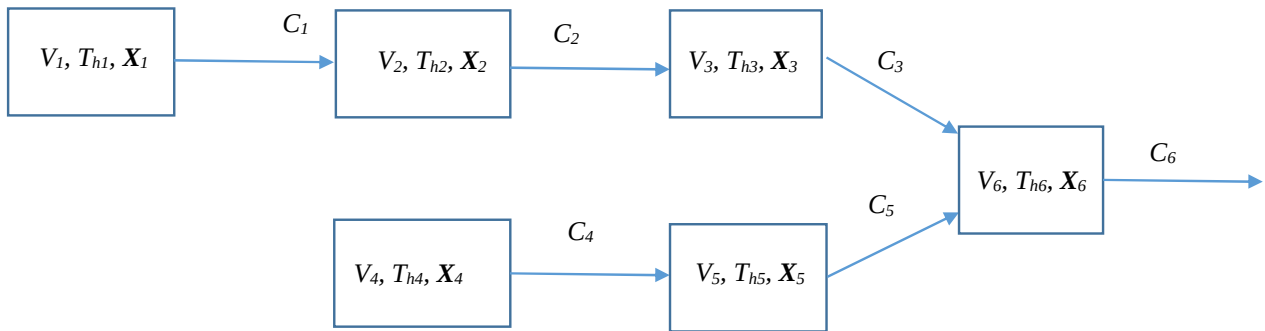
Nu udvides princippet til tre oplande:



Ligningerne 12 og 14 giver:

$$C_3 = \frac{\beta}{V_1+V_2+V_3} \cdot [X_1 \cdot e^{-\rho \cdot (Th_1+Th_2+Th_3)} + X_2 \cdot e^{-\rho \cdot (Th_2+Th_3)} + X_3 \cdot e^{-\rho \cdot (Th_3)}] \quad (15)$$

Princippet kan også udvides til sammenløb af to serier af oplande:



Hvor princippet ved Ligning 12, 14 og 15 giver:

$$C_6 = \frac{\beta}{V_1+V_2+V_3+V_4+V_5+V_6} \cdot [X_1 \cdot e^{-\rho \cdot (Th_1+Th_2+Th_3+Th_6)} + X_2 \cdot e^{-\rho \cdot (Th_2+Th_3+Th_6)} + X_3 \cdot e^{-\rho \cdot (Th_3+Th_6)} + X_4 \cdot e^{-\rho \cdot (Th_4+Th_5+Th_6)} + X_5 \cdot e^{-\rho \cdot (Th_5+Th_6)} + X_6 \cdot e^{-\rho \cdot Th_6}] \quad (16)$$

Denne massebalance kan bruges i en statistisk model, der vil estimere β og ρ . Det antal ubekendte parametre, som modellen skal estimere for hvert metal, er altså en parameter for hver kontinuert attribut, for hvert niveau i kategoriske uafhængige variable minus en, et intercept samt ρ .

Ligning 16 kan generaliseres til følgende form:

$$C_{opt,i} = \frac{1}{\{\sum V\}_i} \cdot \sum_{k=0}^Z [\beta_k \cdot \sum_{j=1}^M [X_{j,k} \cdot \Phi_{i,j} \cdot e^{-\rho \cdot Th_{i,j}}]] \quad (17)$$

Hvor $\Phi_{i,j}$ er en funktion som beskrevet neden for med ligningerne (18a og 18b), og $\Theta_{i,j}$, $Th_{i,j}$, og $\{\sum V\}_i$ er faste konstanter, mens β_k og ρ er de parametre, som modellen skal estimere. M er samlet antal oplande. Funktionen $\Phi_{i,j}$ har følgende udtryk:

$$\Phi_{i,j} = 1, \quad \text{hvis opland } j \text{ løber til } i, \text{ eller } i = j \quad (18a)$$

$$\Phi_{i,j} = 0, \quad Th_{i,j} = 0 \text{ hvis opland } j \text{ ikke løber til } i \quad (18b)$$

$Th_{i,j}$ afhænger af, hvilken vej vandet har løbet fra opland j til opland i for de kombinationer af oplande, hvor der faktisk løber vand fra opland j til opland

i. Dette princip danner et system af ligninger, der har mere eller mindre fælles information fra de uafhængige variable.

4.2.1 Hydraulisk opholdstid

Den hydrauliske opholdstid i hvert opland ($T_{h,opl}$) er bestemt ved:

$$T_{h,opl} = \frac{\Delta t \cdot V_{tot,opl}}{\{\Sigma V\}_i} \quad (19)$$

$V_{tot,opl}$ er det samlede vandvolumen, der er akkumuleret i oplandet. Da det samlede vandvolumen ikke er kendt i hvert opland, bruges ferskvandsarealet ($A_{fersk,opl}$), som er kendt for hvert opland, som en approximation. Det betyder, at der forudsættes en konstant middeldybde (d) for ferskvand i alle oplande:

$$T_{h,opl} = \frac{\Delta t \cdot A_{fersk,opl} \cdot d}{\{\Sigma V\}_i} \quad (20)$$

Ligningerne 20 og 4 kan omskrives til:

$$M_{opl} = M'_{opl} \cdot e^{-\rho' \cdot \frac{A_{fersk,opl}}{\{\Sigma V\}_i}} \quad (21)$$

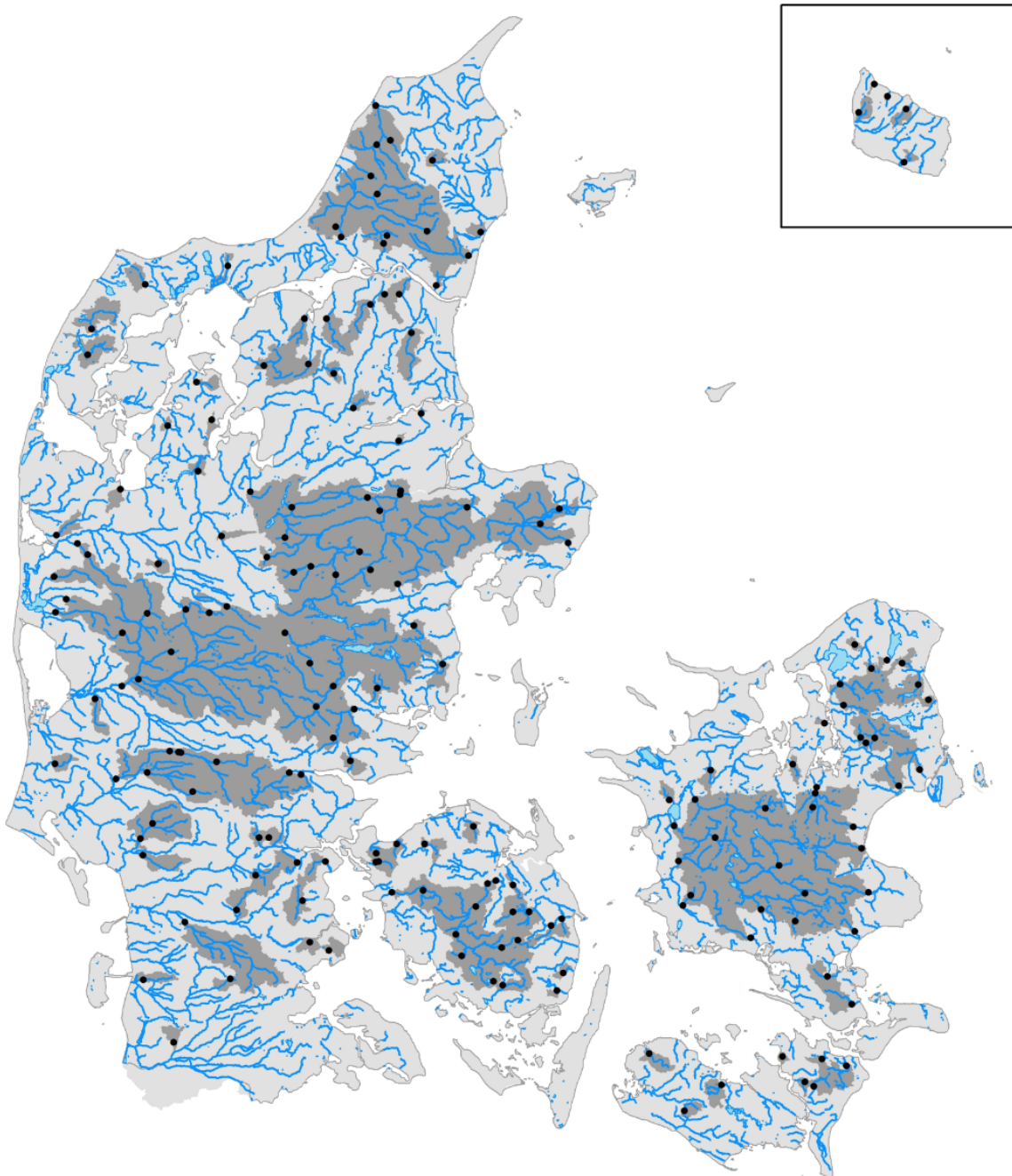
hvor

$$\rho' = \Delta t \cdot d \cdot \rho \quad (22)$$

Så ved at bruge $\frac{A_{fersk,opl}}{\{\Sigma V\}_i}$ som udtryk for $T_{h,opl}$ vil modellen estimere parameteren ρ' i Ligning 22 som udtryk for en eventuel tilbageholdelse.

4.3 Samlet statistisk model

På Figur 4.1 ses de vandløbsstationer, hvor der er foretaget målinger i NO-VANA i perioden 2010-2019. Det mørkegrå område afstrømmer til mindst en målestation.



Figur 4.1. Geografisk placering af de 174 vandløbsstationer, hvor der er målt for enten bly, cadmium eller nikkel, med angivelse (mørkegrå) af de områder af Danmark, der afvander til mindst en målestation.

De målte metalkoncentrationer, $y_{i,n,s,z,m}$, for metal m i et opland i i år n og sæson (årstid) s blev modelleret ved hjælp af en Bayesiansk hierarkisk observationsmodel. For hver kombination af opland, år og sæson kan der være flere enkeltmålinger, der beskrives med indeks z . Sæson bestemmes ud fra den typiske inddeling i vinter (dec., jan. og feb.), forår (mar., apr. og maj), sommer (jun., jul. og aug.) og efterår (sep., okt. og nov.). De logtransformerede metalkoncentrationer antages at være multinormalfordelte.

$$\log(y_{i,n,s,z,m}) \sim MN(lc_{i,s}, \Sigma_{m\grave{a}le}) \quad (23)$$

hvor $lC_{i,s}$ er latente variable af dimension m af de estimerede metalkoncentrationer $\log(y_{i,n,s,m})$ på en logskala, og $\Sigma_{m\grave{a}le}$ er den estimerede kovariansmatrice for de forskellige metaller af dimension $m*m$. Som det kan ses af ligning 23, har vi valgt at aggregere målinger fra forskellige år i den samme latente variabel. Dette er gjort under antagelsen af, at den systematiske variation i metalkoncentrationerne fra år til år er begrænset og for ikke at få for mange latente variable i modellen. I stedet for er effekten af år blevet modelleret som en stokastisk effekt i den strukturelle ligning (se neden for).

Ikke alle metaller har været målt ved alle stikprøvetagninger, og disse ikke-målte værdier og målinger under detektionsgrænsen er blevet estimeret ved hver Metropolis-Hastings iteration med den prior begrænsning, at alle manglende værdier er mindre end den højst målte koncentration, og at værdier under detektionsgrænsen er mindre end detektionsgrænsen. Motivationen for denne huludfyldningsprocedure er i) ikke arbitrært at fastlægge målinger under detektionsgrænsen til f.eks. halvdelen af detektionsgrænsen, ii) at bestemme usikkerheden for manglende målinger og målinger under detektionsgrænsen, og iii) at udnytte informationen i den estimerede måle kovariansmatrice som en del af huludfyldningsproceduren. Alternativt kunne man udarbejde en censureret statistisk model, men denne fremgangsmåde er relativt kompliceret med multidimensionale data, og vælges ikke da den vil udfordre beregningstiden med ekstra variable.

De estimerede metalkoncentrationer blev modelleret i en strukturel model med brug af ligning 2, som også blev antaget at være multinormal fordelt,

$$lC_{i,s} \sim MN(\log(C_{opl}(\beta, \rho')) + \alpha Z_s + \sigma_1 \gamma_i + \sigma_2 \gamma_n, \Sigma_{res}) \quad (24)$$

hvor Z_s er en designmatrice, som modellerer effekter af årstid (intercept arbitrært sat til vinter), $C_{opl,i}(\beta, \rho')$ er bestemt af ligningerne 17 og 22, som er relateret til oplandet (se oplandsmodel), γ_i er stokastiske effekter mellem vandløbsstationer, γ_n er stokastiske effekter mellem år. De år, hvorfra der er stikprøver, som bestemmer $lC_{i,s}$, indgår ved en vægtet værdi, der summer op til en for hver latent variabel. Parametrene α , β , ρ' , σ_1 , og σ_2 bliver estimeret. Dimensionen af middelværdien er m . Σ_{res} er den residuelle strukturelle kovariansmatrice med dimension $m*m$. Bemærk at ved at tage log til middelværdien i ligning 24 (se ligning 1 og 2) antages en stærk prior på de mulige middelværdier, da kun positive koncentrationer er tilladte.

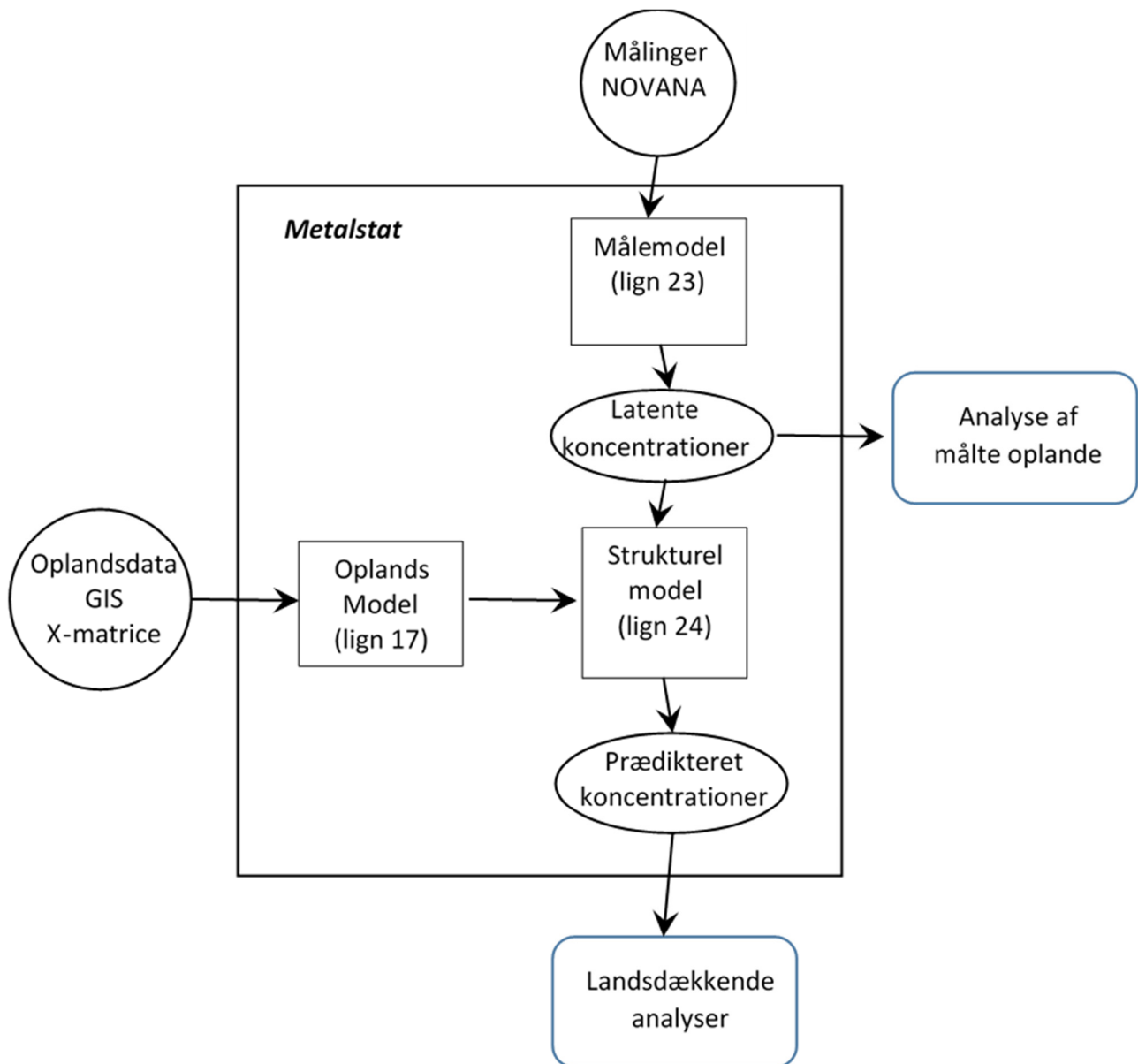
Den fælles posteriori sandsynlighedsfordeling af de latente variable ($lC_{i,s}$), koefficienterne, de tilfældige effekter og de to kovariansmatricer blev alle estimeret ved hjælp af Metropolis-Hastings algoritmen (Boks 5). Nye værdier af kovariansmatricen blev trukket ved hjælp af Wishart matrixfordeling, $\Sigma_{u+1} \sim WMD\left(10000, \frac{\Sigma_u}{10000}\right)$, hvor u er iterationsnummer og de øvrige værdier fra normalfordelinger med passende standardafvigelser af skridt størrelsen.

Den lineære relation mellem tilført masse og X (Ligning 8) kan for nogle kombinationer af β beregne negative masser i et omfang, der giver negative koncentrationer. Det håndteres ved, at en værdi for β forkastes i Metropolis-Hastings algoritmen, hvis den er årsag en negative værdi af C_{opl} .

I Bayesiansk statistik bliver antagelsen om normalfordelte stokastiske effekter i praksis udført ved at give de tilfældige effekter en stærk normalfordelt prior, som sikrer, at de stokastiske effekter er standard normalfordelte, $\gamma_{j,m} \sim N(0,1)$.

De øvrige estimerede parametre antages at have uniforme prior sandsynlighedsfordelinger i deres respektive domæne.

En oversigt over den samlede modelstruktur er vist i Figur 4.2. Målinger fra NOVANA programmet indgår i en likelihoodfunktion i målemodellen til at estimere median koncentrationer som latente variable. Denne latente variabel kan både anvendes til at analysere koncentrationsniveauer i de målte oplande og indgå i en strukturel model sammen med GIS data for hele landet til at prædikerer koncentrations niveauer for hele landet.



Figur 4.2. Samlet modelstruktur, hvor målinger fra NOVANA programmet indgår i en målemodel, der kan analysere koncentrationsniveauerne i de målte oplande, og som også kan indgå sammen med oplandsmodellen i en strukturel model til at prædikerer koncentrationer for hele landet.

4.4 Samlet oversigt over typer af model input og output

Modellen kan give flere forskellige typer relevante output, som det fremgår af Figur 4.3 og forklares i det følgende.

Der er to hovedkilder til input til MetalStat:

- Målinger af metalkoncentrationer i vandfasen ved faste vandløbsstationer opdelt på sæson og år.
- Oplysning om forhold i oplandene, der formodes at kunne have betydning for metalkoncentrationen i vandfasen. Dette definerer de oplandsrelaterede deterministiske effekter, som indføres i modellen.

Disse to kilder til input samles i likelihood-funktionen, og som output kommer:

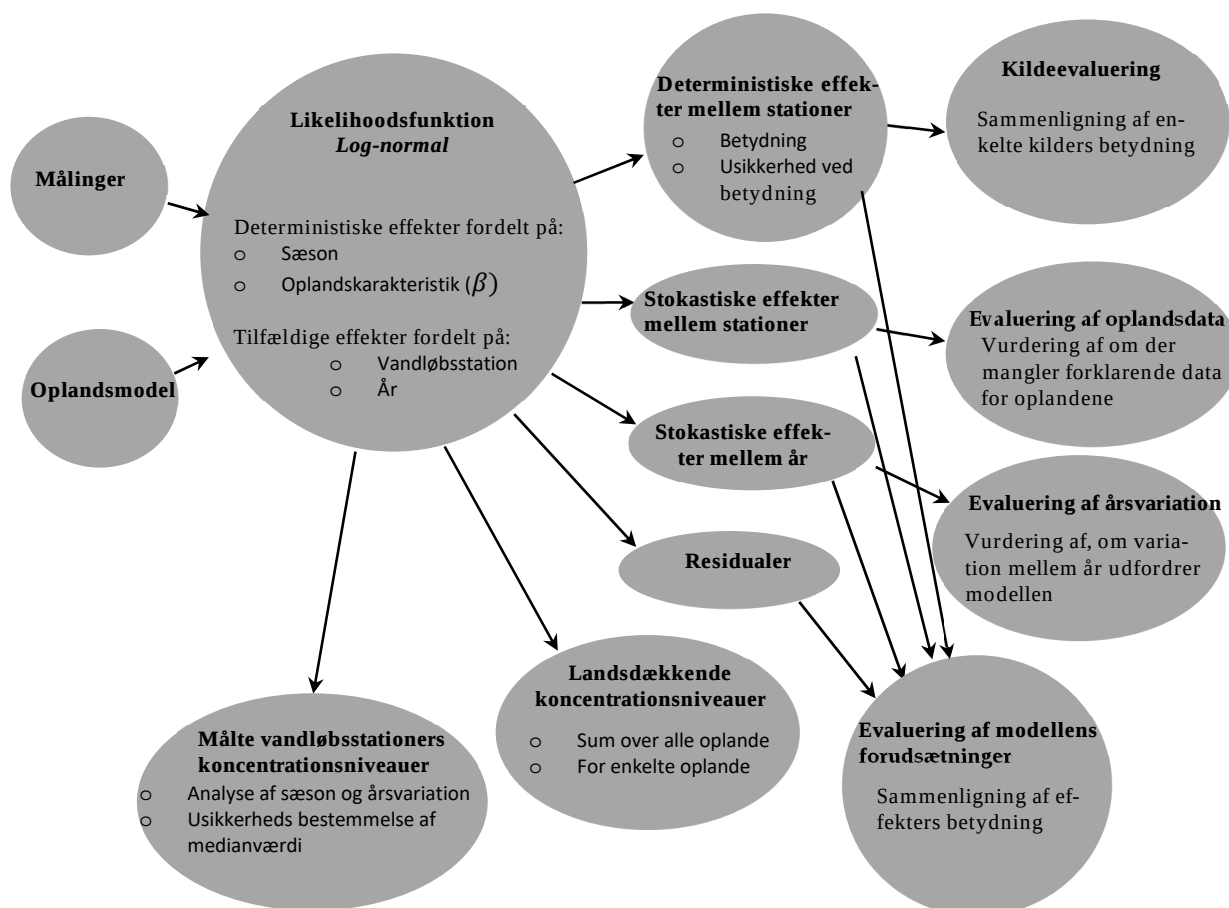
- *Betydning af deterministiske effekter.* De deterministiske effekter, som er taget med i MetalStat, er kun formodede effekter baseret på faglig viden om kilder transport og skæbne. MetalStat vil tillægge disse effekter en større eller mindre reel betydning i forhold til:
 - Oplandsrelaterede effekter: Hvilke oplandsrelaterede forhold har større / mindre betydning?
 - Sæsonrelaterede effekter: Hvilke sæsoner (vinter, forår, sommer, efterår) har større / mindre betydning?
- *Betydning af stokastiske effekter.* I det omfang der er forskelle mellem vandløbsstationer, som endnu ikke kan identificeres gennem deterministiske effekter, regnes de som værende stokastiske effekter. På denne måde estimerer modellen stokastiske effekter mellem år, så hvis enkelte år er særlige, eller der er en tidsmæssig udvikling fra år til år, vil modellen estimere dette som stokastiske effekter. Det er derfor klart, at de stokastiske effekter er meget centrale som redskab til at vurdere, om modellen i tilstrækkeligt omfang evner at identificere og bruge viden til at beskrive metalkoncentrationerne.
- *Landsdækkende koncentrationsniveauer.* MetalStat vil kunne prædiktere fraktilværdien for koncentrationsniveauerne i alle ID 15 oplande, enten som et samlet resultat for alle oplande eller for hvert opland.
- *Målte koncentrationsniveauer.* For et opland, der indeholder en målt vandløbsstation, vil modellen typisk kunne estimere fraktilværdier, der er mindre usikre end for et opland uden en målt vandløbsstation. Her vil modellen kunne levere et mere præcist estimat, end det hidtil har været muligt, og med et usikkerhedsinterval for bestemmelse af koncentrationen. Det vil også være muligt for modellen af undersøge effekter af år og sæson, der er specifikke på den målte station.

For de målte vandløbsstationer kan MetalStat opdele den samlede variation i koncentrationsniveau i hhv. deterministiske og stokastiske effekter mellem vandløbsstationer samt residual variation, hvor de deterministiske effekter mellem vandløbsstationer afspejler modellens evne til at bruge kendt viden om de enkelte oplande til at forudsige koncentrationsniveauet. Residualer fremstår i modellen som en variation i de målte koncentrationsværdier uden nogen form for systematik og kan skyldes mange forhold, f.eks. lokale regn hændelser op til indsamlingstidspunktet, en særlig kraftig og kortvarig tilstrømning fra et bestemt drænrør opstrøms indsamlingsstedet eller analysekemiske årsager. Årsagen til residualer kan enten være styret af effekter, der

ikke reelt lader sig beskrive med nogen realistisk model, mens andre potentielt vil kunne beskrives med en mere kompleks model, f.eks. med en model, der i højere grad inddrager kortvarige dynamiske forhold. Med udgangspunkt i de tre definerede typer af effekter kan modellens resultater og forudsætninger evalueres på følgende områder:

- *Kildeevaluering*, hvor de deterministiske effekter tildeles en vis grad af betydning for hver af de kilder til metal, som modellen har med som uafhængig variabel. Hvis nogle kilder estimeres til at være ubetydelige, kan det overvejes at negligere dem i modellen.
- *Evaluerings af oplandsmodel*, hvor forholdet mellem deterministiske og stokastiske effekter mellem vandløbsstationer vil indikere, i hvilken grad oplandsmodellen er i stand til at forklare forskellen mellem vandløbsstationer. En granskning og kortlægning af de målte vandløbsstationer, der har det største bidrag til stokastiske effekter, vil give en rettesnor for, hvad der kan tænkes at være årsag til den ukendte del af effekter mellem vandløbsstationer, som ikke kan beskrives med oplandsmodellen. På lignende måde kan det evt. pege på, at oplandsmodellen bør gøres tidsdynamisk, hvis der er betydelige stokastiske effekter mellem år.
- *Evaluerings af modelforudsætninger generelt*, hvor residualvariationen sammenholdes med den variation, som kan beskrives med hhv. deterministiske og stokastiske effekter mellem vandløbsstationer og mellem år. I det omfang residualvariationen er betydelig, vil det give en rettesnor for potentialet i yderligere at forbedre modellens opbygning.

Det er vigtigt løbende at evaluere og opdatere modellen, så den afspejler den nyeste viden både i forhold til at opdatere måledata, efterhånden som de udbygges med nye målinger, og i forhold til at evaluere oplandsdata samt modelforudsætninger baseret på de målte vandløbsstationer.



Figur 4.3. Samlet modellerings-struktur med angivelse af de forskellige typer output, hvoraf analysen for hhv. deterministiske og stokastiske effekter kan bruges til at evaluere modellens evne til at forudsige koncentrationsniveauet.

5 Mulige kilder til bly, cadmium og nikkel

Viden om kilderne til metaller er vigtig, dels når de uafhængige variable (X matricen) til oplandsmodellen skal defineres, og dels når den betydning, som MetalStat tillægger dem, skal vurderes. Vurderingen af spredningen og forureningen med metaller vanskeliggøres dog af variationen i baggrundskoncentrationer af metallerne i jord, underjord, akvifermaterialer og sediment (Sørensen et al., 2021).

I dette kapitel gennemgås kort centrale kilder for især bly, cadmium og nikkel som illustration for den viden, som bør inddrages for en kommende model, der dækker flere metaller. I kapitel 6 bliver denne viden udmøntet ved valg af datakilder til X matricen i Ligning 9.

Kilderne kan opdeles i hhv. diffuse kilder, der har en vis arealmæssig udstrækning langs vandløbet, og punktkilder, der tilfører stof til vandløbet i enkelte punkter. Mange mindre punktkilder kan samlet set opfattes som diffuse kilder. Således vil drænudledning, der typisk foregår fra mange dræn, blive betragtet som en diffus kilde. Både diffuse kilder og punktkilder kan være udvaskningsbetingede, hvilket er illustreret i Tabel 5.1:

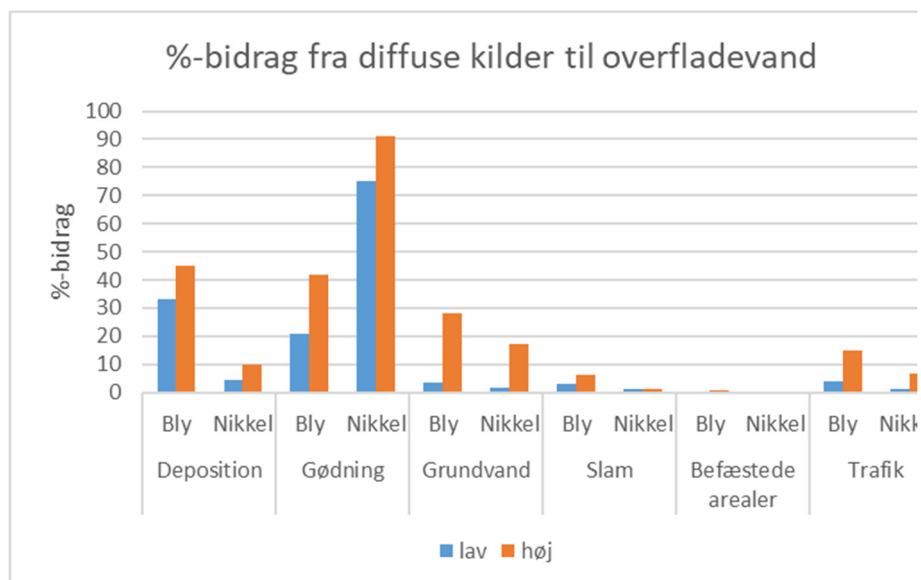
Tabel 5.1. Kilder fordelt på hovedtyper

Diffus	Punktkilde
• Dræning • Grundvand • Overfladeafstrømning fra ikke befæstede areal • Atmosfærisk tilførsel • Spredt bebyggelse	• Lokal udsivning fra forurenede jord/grundvand (grunde/lossepladser) • Afstrømning fra befæstede arealer • Udløb fra renseanlæg

Der er gjort en række erfaringer fra Miljøstyrelsens projekt om diffuse kilder til de 23 hovedvandoplande (DHI, 2020). Der er for de enkelte stoffer stor forskel på bidraget fra forskellige diffuse kilder til hovedvandoplandene. De absolutte og relative bidrag af stoffer fra de forskellige diffuse kilder varierer mellem hovedvandoplandene, hvilket skyldes forskelle mellem deres areal, arealanvendelse, vandføringer og de gennemsnitlige grundvandskoncentrationer af stoffer.

For metallerne nikkel og bly gælder det, at deposition, gødning og grundvand bidrager med >90% af det diffuse bidrag til overfladevand, alle med typisk mindstebidrag på 10 % (DHI, 2020) (se Figur 5.1, der viser de relative bidrag af bly og nikkel til overfladevand for de diffuse kilder). På grund af det begrænsede bidrag fra de øvrige diffuse kilder sammenholdt med det omfattende arbejde, der ligger i dataindsamling og beregning af de diffuse kildebidrag, har vi i projektets modeludvikling valgt at kun tage bidrag fra atmosfærisk deposition, dræning og overfladeafstrømning fra ikke-befæstede arealer, hvilke indeholder bidrag fra gødning, slam og grundvand, med i de efterfølgende betragtninger. Bidraget fra grundvand vil indeholde en fraktion af nedsivningen fra befæstede og ikke-befæstede arealer.

Figur 5.1. Bidrag fra atmosfærisk deposition, udbringning af slam og gødning og grundvandsindstrømning for bly og nikkel. Værdier angiver laveste og højeste bidrag til hovedvandoplandene, fra DHI (2020).



Atmosfærisk våd- og tørdeposition varierer ikke voldsomt geografisk i Danmark. Vind, nedbørsmængde, solindstråling og lufttemperatur påvirker afsætningen af luftforureningskomponenter (ELLERMANN ET AL., 2019). Den samlede deposition af tungmetaller (summen af tør- og våddeposition) kan estimeres ud fra målinger af våddeposition samt ved beregning af tørdeposition ud fra målingerne af atmosfærens indhold af partikelbundne tungmetaller. I NOVANA har våddepositionen og den atmosfæriske koncentration af partikelbundne tungmetaller været målt siden hhv. 1990 og begyndelsen af 1980'erne. I 2010 blev antallet af målestationer reduceret, så der siden 2010 kun har været målt luftkoncentrationer af tungmetaller ved Anholt og ved Risø, hvor der tidligere blev målt luftkoncentrationer af tungmetaller på fem målestationer. Mængden af nedbør, antallet af byger, nedbørsintensiteten, samt i hvilket omfang transport af luftmasser med høje koncentrationer af partikler falder sammen med regnhændelser, påvirker variationen i våddepositionen (ELLERMANN ET AL., 2021).

Lokal udsivning fra forurenede jord / grundvand, typisk fra historiske forureninger, vil kunne relateres til kortlagte forurenede grunde på vidensniveau 2, hvor hver grund specifikt udvælges i det omfang, den vurderes at være en mulig kilde til bly, cadmium eller nikkel. Der vil dog være mange lokale forureninger, hvor den specifikke afstand til vandløbet inden for ID 15 oplandet kan være afgørende. En mulighed kunne være at fokusere på de mest vandløbsnære forureninger.

Afstrømning fra befæstede arealer samt udløb fra renseanlæg styres dels af den vandmængde, der udledes, og koncentrationsniveauet. Her kunne data fra PULS-databasen finde anvendelse (<https://miljoportal.dk/nyheder/2020/nyt-puls-system-er-i-luften/>).

5.1.1 Tilførsel af metaller via grundvand til overfladevand

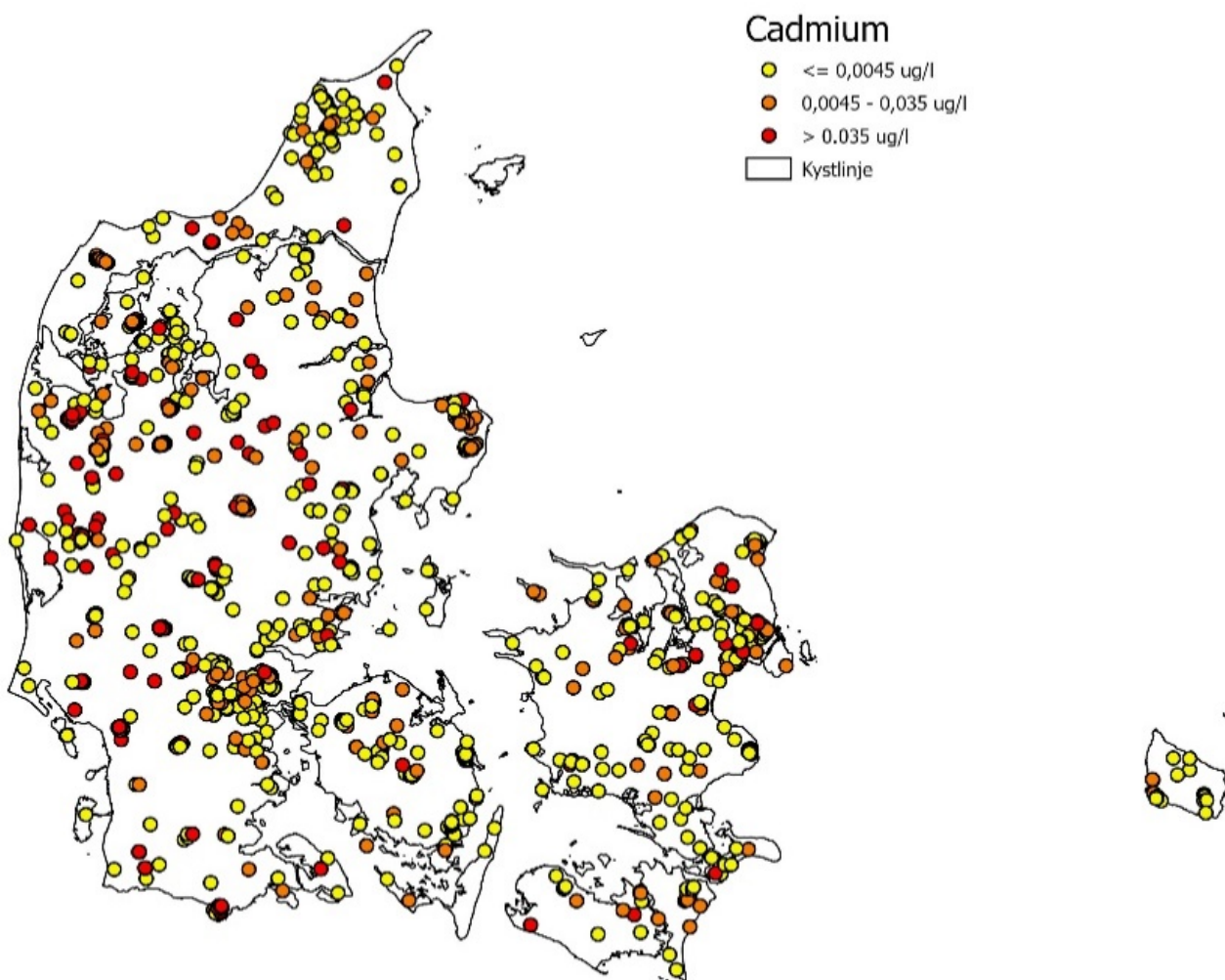
I denne rapport vil MetalStat således alene estimere en mulig grundvandstilførsel gennem den mængde grundvand, der strømmer til vandløbet uden hensynstagen til forskelle i koncentrationsværdier. Derfor vil tilførsel af metaller fra grundvand få i det følgende en mere uddybet beskrivelse, især med henblik på at beskrive mulighederne for senere at inddrage grundvand som en udvidelse af MetalStat.

Tungmetaller findes i jordens geologiske udgangsmateriale og frigives herfra ved kemiske, fysiske og biologiske nedbrydningsprocesser. Stofferne er således en naturlig del af de terrestriske og akvatiske økosystemer, herunder grundvand. Menneskets aktiviteter har desuden resulteret i, at tungmetaller er frigivet til miljøet, hvorfor mange tungmetaller i dag forekommer i miljøet i koncentrationer, der ligger over baggrundskoncentrationen (Nilsson et al., 2019).

Den følgende karakteristik af de tungmetaller, der omhandles i nærværende rapport, har taget udgangspunkt i beskrivelsen fra Mortensen et al. (2021). Bly forekommer naturligt i sulfider, sulfater og karbonater. Desuden kan det indgå i feldspat og glimmermineraller. Bly binder sig til lermineraller, organisk stof og jern- og mangan oxid/hydroxider og har almindeligvis en lav opløselighed. Cadmium optræder som bikomponent i for eksempel sulfidholdige mineraller. Koncentrationen i grundvand er typisk meget lav. Cadmium er mest mobilt i grundvand under oxiderede forhold, hvis pH er under 8. Mobiliteten øges med faldende pH. Cadmium kan binde sig til organisk stof, lermineraller og udfælder med jern- og mangan-hydroxider. Nikkel findes blandt andet i sulfidholdige mineraller. Mobiliteten i grundvand afhænger af pH og redox-forhold, og ved neutrale pH forhold kan mobiliteten nedsættes ved sorption. Kobber er indbygget i for eksempel pyroxener, amfiboler og biotit. Mobiliteten i grundvand afhænger af pH og redox-forhold og er størst under sure og oxiderede forhold. Kobber fælder ud med jern- og manganforbindelser. Zink findes i en række almindelige mineraller som pyroxener, amfiboler og glimmermineraller. Det binder sig til jern- og manganforbindelser, lermineraller og organisk stof. Mobiliteten i grundvand er som for cadmium størst under sure forhold.

Den geografiske variation i grundvandets indhold af tungmetaller kan beskrives på grundlag af udtræk fra den nationale boringsdatabase Jupiter. Til brug for vandområdeplanerne for tredje planperiode 2021-2027 har GEUS sammenstillet data for bl.a. tungmetaller for 2050 grundvandsforekomster for perioden 2013-2019. For hvert indtag, der er tilknyttet en grundvandsforekomst, er der for hvert stof beregnet en stofs specifik middelværdi af de årlige middelværdier på baggrund af kvalitetssikrede rådata (Mortensen et al., 2021). Datasættet indeholder 1964 observationer for bly, 1827 observationer for cadmium, 9161 observationer for nikkel, 2246 observationer for kobber og 2558 observationer for zink. Figur 5.2. eksemplificerer den geografiske fordeling af de målte grundvandskoncentrationer af cadmium.

I den nuværende version af modellen MetalStat bruges en estimeret tilførsel af grundvand (X_{grundv}) som attribut for en mulig grundvandsbåret tilførsel af metal til overfladevand i interaktion med hhv. pH, ler og organisk stof i jordmatrixen. Som ovenstående eksempel viser, er dette en forsimplet beskrivelse, da koncentrationsniveauet i grundvandet synes at variere geografisk. En bedre beskrivelse, som på sigt kan indarbejdes i modellen, vil være baseret på et estimat på grundvandskoncentrationen (C_{grundv}), hvor massen af metal, der tilføres overfladevand, er givet som $\alpha \cdot C_{grundv} \cdot X_{grundv}$, hvor α udtrykker en evt. tilbageholdelse (retention), når grundvandet strømmer til vandløbet, og X_{grundv} er estimeret grundvandbidrag. For hvert ID 15 opland kan C_{grundv} ses som en latent variabel, der kan estimeres med brug af de grundvandsprøver, der er indrapporteret for det pågældende opland, på den samme måde som vandkoncentrationen i vandløbet ses som en latent variabel, der er målt i NOVANA. Derved kan grundvandsovervågningen integreres i vandløbsovervågningen.



Figur 5.2. Koncentrationsniveauer for cadmium i grundvand. Data fra Mortensen et al. (2021).

Nilsson et al. (2019) har sammenstillet måledata fra grundvand og vandløb, hvilket blandet andet illustrerer de betydeligt lavere koncentrationsniveauer i vandløb sammenlignet med grundvand for bly og cadmium, Tabel 5.3.

Tabel 5.3. Tungmetalkoncentrationer i grundvand og vandløb. Andel af indtag med fund og mediankoncentrationer i grundvandsovervågningen (80 indtag) og vandløbsvand (21 stationer) i kontrolovervågningen (Thorling et al., 2016; Thodsen et al., 2016). MKK: miljøkvalitetskrav. Tabellen er modificeret efter Nilsson et al. (2019).

	Grundvand (2015)			Vandløb (2011-2015)		
	Antal indtag med konc. > kravværdi/antal indtag	Hyppigheden af indtag med konc. > kravværdi	Mediankonc. (ug/l)	Antal fund > MKK/antal prøver	Hyppighed af fund > MKK	Mediankonc. (ug/l)
Bly	6/ 80	8%	0,22	27/304**	8.9%**	0,029
Cadmium	3/80	4%	0,08	0/316	0%	0,005
Kobber	3/80	4%	1,65	58/172	34%*	1,05
Nikkel	28/134***	21%***	1,50***	46/304**	15%*, **	1,0
Zink	11/80	14%	1,65	54/74	73%*	7,75

*Korrigeret for baggrunds niveau

**Vurderet i forhold til miljøkvalitetskrav gældende i Bek. Nr. 1625, 2017

***Data fra 2013 (Thorling et al., 2015)

6 Valg af X matrice

En viden om metallernes kilder skal implementeres i modellen som uafhængige variable, der kan bestemmes svarende til X matricen i Ligning 8. Modellen forudsætter, at der er en lineær sammenhæng mellem en uafhængig variabel og tilførslen af masse metal, og skønnes dette ikke til at være tilfældet, kan variablene eventuelt transformeres for bedre at opnå linearitet. Hvis der er mulig interaktion mellem to uafhængige variable fra en faglig vurdering, laves en ny variabel ved multiplikation af de disse to uafhængige variable.

I det følgende listes de valgte uafhængige variable, der opdeles efter hhv. belastnings-variable og transport-variable. Med belastningsvariable menes uafhængige variable, der betyder noget for, hvor store stofpuljer der kan forekomme i oplandet, mens transport-variable beskriver, hvor let disse stofpuljer kan blive transporteret til vandløbet. Der er en interaktion imellem hhv. belastnings- og transportvariablene, da stof kun kan transporteres til vandløbet i det omfang, der er en stofpulje i oplandet, mens en stofpulje i oplandet kun kan transporteres til vandløbet i det omfang, der er en transportvej. Det er dog kun nogle interaktioner, der giver mening, hvilket kan illustreres ved, at belastningen fra et renseanlæg ikke afhænger af nogen transport-variable, da spildevandet udledes direkte til vandløbet, mens betydningen af stof fra udbragt gylle på landbrugsjord kan afhænge af tilstedeværelsen af makroporer og dræn.

Hvis en uafhængig variabel ikke har reel betydning for modellering, vil det vise sig ved, at MetalStat tillægger den attribut en værdi omkring nul. På den måde er MetalStat en fleksibel model, hvor nye uafhængige variable kan testes for deres betydning og løbende inddrages for derved at forbedre modellen mod mere præcise prædiktioner.

Forurenede grunde indgår ikke som variabel i denne version af MetalStat, da der ikke har været ressourcer i projektet til at inddrage data fra disse fordelt på ID 15 oplande. Det bør overvejes til fremtidige udbyggede versioner.

Reference til Id i parentes for uafhængige variable referer til Tabel over datakilder i Bilag B.

6.1 Belastningsvariable

Følgende belastningsvariable er valgt til at beskrive belastning af stof:

- Spildevandsmængde. Her bruges vandmængde fra renseanlæg (X_{rens}) direkte (Id 21)
- Regnvandsbetingede udløb. Her bruges vandmængde fra regnvandsbetingede udløb (X_{regn}) direkte (Id 20)
- Udbragt gylle fra grise. Her bruges udbragt mængde N fra svinegylle fordelt på dyreenheder (X_{svin}) direkte (Id 9)
- Intensivt landbrug. Her bruges arealet med intensivt landbrug (X_{land}) direkte (Id 8)

- Grundvands-indløb (eller tilstrømning) til vandløbet. Her bruges andel af grundvand (f_{grund}) (Id 19) samt det tilførte vandvolumen til oplandet: $X_{grundv} = f_{grund} \cdot V_{opl}$. Dette er en forsimpning i forhold til at koncentrationniveauet i grundvandet varierer, men det vil ikke føre til en overestimulering af grundvandsbidraget, men nærmere modsat, da noget af variationen i grundvandsstilførelsen pga. forskellige koncentrationer niveauer delvist vil tilskrives stokastiske effekter.

6.2 Transportvariable

Der er valgt følgende transportvariable, der beskriver, i hvilken grad stof i form af metalmængde, der er tilstede i en pulje i oplandet, kan blive transporteret til vandløbet.

- Drænastrømning. Her bruges drænvandføring (D) (Id 11) og oplands areal (A_{opl}) til at estimere volumen af vand tilført i oplandet fra dræn:

$$X_{dræn} = \frac{D \cdot A_{opl}}{1000} [\text{m}^3]$$

- Makroporer. Her bruges areal med makroporer direkte (Id 13) : X_{makro} [ha]. Dette svarer til et areal med makroporer i høj risikoklasse (4 + 5), og massetilførslen til vandløbet pga. makroporefaciliteret transport forudsættes derfor proportionalt med dette areal.
- Nettonedbør. Her bruges nettonedbør (N_{net}) (Id 17) og oplands areal (A_{opl}) til at estimere volumen af vand tilført i oplandet fra nettonedbør:

$$X_{ned} = \frac{N_{net} \cdot A_{opl}}{1000}$$

hvor X_{ned} har enheden m^3

- pH. $M_{H^+} \propto \int_A 10^{-pH} da$ pH tillægges betydning, idet metaller forekomst i vandfasen kan afhænge af, at elementarladninger (forsuring) gør metalatomer mobile. Denne betydning kan dække over komplekse sammenhænge, hvor der vælges en simpel empirisk model, hvor det forudsættes, at den masse metal, der evt. mobiliseres til vandfasen i et opland, er proportional med antal mol H^+ i oplandet. Desuden forudsættes, at følgende attribut er proportional med antal mol H^+ i et opland:

$$M_{H^+} \propto \int_A 10^{-pH} da$$

- Dette integral løses numerisk med brug af konstant grid størrelse for estimeret pH (250 X 250 m):

$$M_{H^+} \propto \int_A 10^{-pH} da \propto X_{pH} = \sum_{opl} 10^{-pH}$$

- Da et ID 15 opland typisk er omkring 15 km^2 , betyder det, at der typisk summeres op over ca. 240 kvadrater på hver 250 X 250 m, hvilket vurderes at give en tilstrækkelig detaljeret beskrivelse af effekter fra pH. pH værdien estimeres som to uafhængige variable ud fra middelværdi i hhv. 0-30 cm dybde (Id 6) og 1-2 m dybde (Id 5) for at beskrive både den overfladenære og dybere pH.
- Kulstofmængde. Det forudsættes, at massen af metal, der frigives til vandløbet, kan være påvirket enten positivt eller negativt af massen af kulstof i

oplandet, som i modellen er estimeret som middel kulstof-indhold gange oplandsareal. Her bruges organisk kulstof i procent samt oplandsareal:

$$X_{org} = \frac{f_{org}}{100} \cdot A_{opt}$$

- Kulstofmængden estimeres både i a- og b-horisonten (hhv. Id 4 og 5).
- Lermængde. Massen af metal, der transporteres med ler, forudsættes at afhænge af ler-indholdet i jorden, enten negativt ved sorption til jordmatricen eller positivt ved mobilisering med lerkolloider. Denne afhængighed forudsættes at være proportional med massen af ler, der er til stede i jorden i oplandet, og denne masse forudsættes at være proportional med andelen af ler i jorden ganget med oplandsarealet:

$$X_{ler} = \frac{f_{ler}}{100} \cdot A_{opt}$$

- hvor f_{ler} er andel (%) af ler i oplandets topjord (Id 2).

6.2.1 Interaktioner

I det omfang en transport-variable har betydning for, hvor meget en belastnings-variable kan påvirke koncentrationsniveauet i vandløbet, vil der være en interaktion imellem disse, men da hver interaktion skal beskrives som en ny variabel, er det vigtigt kun at forudsætte interaktioner, hvor det skønnes mest relevant for at begrænse antallet af variable i modellen. I Tabel 6.1. ses en oversigt over uafhængige variable og forudsatte interaktioner mellem disse. Det ses i tabellen, at der er 14 forskellige typer af variable samt 13 interaktioner mellem disse. Det giver i alt 27 uafhængige variable, der bliver brugt til estimering i modellen. Der er opstillet en mulig interaktion mellem tilstedeværelsen af intensivt landbrug (X_{land}) og hhv. overfladenært ler ($X_{land}: X_{lera}$), kulstof ($X_{land}: X_{Ca}$), estimeret drænflow ($X_{land}: X_{dræn}$), makroporer flow ($X_{land}: X_{makro}$) samt nettonedbør ($X_{land}: X_{ned}$). På lignende måde er undersøgt interaktion mellem historisk udbragt gylle fra grise (X_{svin}) og overfladenært ler ($X_{svin}: X_{lera}$) og kulstof ($X_{svin}: X_{Ca}$), estimeret drænflow ($X_{svin}: X_{dræn}$), makroporeflow ($X_{svin}: X_{makro}$) samt nettonedbør ($X_{svin}: X_{ned}$).

Indløbet af grundvand (X_{grundv}) undersøges for transport-variable, der er dybere, i et forsøg på at forbedre geokemisk relevans: Interaktion mellem H^+ for 1-2 m ($X_{grundv}: X_{Hb}$), mængden af ler i b-horisonten ($X_{grundv}: X_{lerb}$) samt kulstof i b-horisonten ($X_{grundv}: X_{Cb}$).

Tabel 6.1. Tabel over de anvendte uafhængige variable og deres interaktioner

Variable X1, X2,..., X14, svarer til notationen i modeludskriften.			Belastnings-variable (interaktioner vist med :)			
			Intensivt landbrug X_{land}	Gylle fra smågrise X_{svin}	Regnvands bet. udløb X_{regn}	Renseanlæg X_{rens}
						Grundvands inflow X_{grundv}
Transport-variable	H ⁺ for 0-13 cm	X_{Ha}				
	H ⁺ for 1-2 m	X_{Hb}				$X_{grundv}: X_{Hb}$
	Ler a-horisont	X_{lera}	$X_{land}: X_{lera}$	$X_{svin}: X_{lera}$		
	Ler b-horisont	X_{lerb}				$X_{grundv}: X_{lerb}$
	C a-horisont	X_{Ca}	$X_{land}: X_{Ca}$	$X_{svin}: X_{Ca}$		
	C b-horisont	X_{Cb}				$X_{grundv}: X_{Cb}$
	Drænflow	$X_{dræn}$	$X_{land}: X_{dræn}$	$X_{svin}: X_{dræn}$		
	Makroporeflow	X_{makro}	$X_{land}: X_{makro}$	$X_{svin}: X_{makro}$		
	Netto nedbør	X_{ned}	$X_{land}: X_{ned}$	$X_{svin}: X_{ned}$		

De valgte uafhængige variable dækker i forskellig grad de mulige antropogene kilder, som vist i Tabel 6.2. Dræning er dækket ind af drænet areal samt interaktionen mellem drænet areal og hhv. areal intensivt landbrug og antal dyreenheder udbragt svinegylle. Desuden kan både arealet med makroporer og nettonedbøren ses som mulige bidrag for dræntransporteret stof, også i interaktion mellem henholdsvis intensivt landbrugsareal og areal med makroporer. Betydning af grundvandsindstrømning er søgt beskrevet med den vandmængde, som strømmer til fra grundvandet i interaktion med de dybeste horisonter for henholdsvis pH, ler og organisk stof. Overfladeafstrømning fra åbent land er søgt beskrevet ved areal med intensivt landbrug og udbragte dyreenheder svinegylle i interaktion med hhv. pH, ler og organisk stof i de øverste horisonter samt makroporeflow og nettonedbør. Atmosfærisk tilførsel samt spredt bebyggelse regnes som en generel baggrundsbelastning. Lokal udsivning fra forurenede jord dækkes ikke med de valgte uafhængige variable. Afstrømning fra befæstede arealer beskrives af vandmængden fra regnvandsbetingede udløb. Renseanlæg dækkes af spildevandsmængden.

Tabel 6.2. Uafhængige variable fordelt på de antropogene Kilder.

Diffuse kilder	Punktkilder
- Dræning ($X_{dræn}$, X_{land}: $X_{dræn}$, X_{svin}: $X_{dræn}$, X_{makro}, X_{land}: X_{makro}, X_{svin}: X_{makro}, X_{ned}, X_{land}: X_{ned}, X_{svin}: X_{ned}) - Grundvand (X_{grundv}, X_{grundv}: X_{Hb}, X_{grundv}: X_{lerb}, X_{grundv}: X_{Cb}) - Overfladeafstrømning (åbent land) (X_{land}, X_{land}: X_{lera}, X_{land}: X_{Ca}, X_{land}: $X_{dræn}$, X_{land}: X_{makro}, X_{land}: X_{ned}, X_{svin}, X_{svin}: X_{lera},	- Lokal udsivning fra forurenede jord (Ikke medtaget) - Afstrømning fra befæstede arealer (X_{regn}) - Udløb fra rensesanlæg (X_{rens})

7 Modelresultater

De resultater, som vises i denne rapport, er kun anvendelige for at beskrive brugen af modellen og bør ikke indgå i nogen form for beslutningsstøtte. Resultatafsnittet er medtaget for at vise, hvordan MetalStat både kan analysere resultaterne fra de målte vandløbsstationer, forudsige koncentrationer i alle ID 15 oplande i Danmark og understøtte arbejdet med at opnå bedre kausal forståelse for metalleres forekomst.

7.1 Udvalgelse af metaller i forhold til modellerings-potentiale

MetalStat kan sættes op til alle relevante metaller, og hvis der er mangelfulde data til rådighed for et metal, vil det vise sig som en stor usikkerhed i estimerterne omkring dette metal. Den samlede liste af metaller indeholder barium, bly, cadmium, chrom, kobber, nikkel, zink, kviksølv og arsen. Af disse er bly, cadmium, kobber, nikkel og zink udvalgt til modeludvikling i to faser, hvor første fase omhandler bly, cadmium og nikkel, og anden fase indeholder alle de udvalgte metaller. Grunden til, at der er valgt at udvikle i to faser med hhv. tre og fem metaller, er et ønske om at optimere beregningsarbejdet, da modellen med tre metaller er tættere på at være konvergeret, og derfor bedre kan nå at kunne danne baggrund for denne afrapportering.

7.2 Vurdering af modellens konvergens

Modelpakken Mathematica (<https://www.wolfram.com/mathematica/>) er brugt som udviklingsmiljø, og herfra kan der under brug af Metropolis-Hastings algoritmen løbende trækkes mange forskellige oplysninger, hvoraf en af de vigtigste er deviance (afvigelse), der afspejler uoverensstemmelsen mellem modellens estimerede koncentrationsværdier sammenholdt med målinger. Så længe modellen konvergerer (se evt. Boks 3), vil uoverensstemmelsen løbende blive mindsket, og denne periode er først slut, når en yderligere reduktion i uoverensstemmelsen ikke synes mulig, det vil sige når forskellen mellem model og måldata ikke længere udviser en nedadgående tendens.

Konvergeringsperioden er især lang, når modellen startes op med helt nye data, da alle variable så skal justeres på plads. For at minimere den nødvendige tid til opstart af modellen er det derfor en fordel at lade modellen konvergere trinvist efter følgende princip:

1. Den latente variabel for mediankoncentration på alle stationer med målinger skal konvergere først uden at inddrage deterministiske og stokastiske effekter.
2. Deterministiske effekter inddrages med brug af de latente variable uden hensyntagen til stokastiske effekter.
3. Med udgangspunkt i de konvergerede resultatet punkt 2 sættes både de stokastiske og deterministiske effekter til at konvergere.
4. Den samlede model sættes til at konvergere.

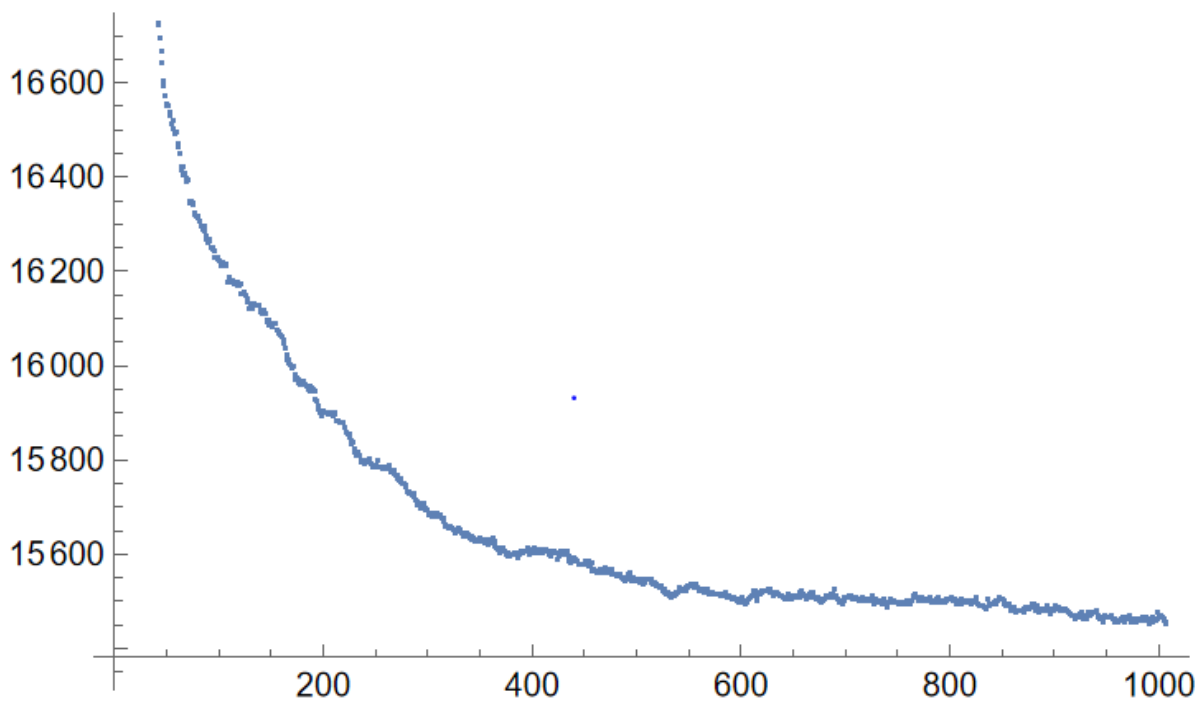
Den trinvis opstart er begrundet i, at den latente variabel primært er styret af målingerne, så det kan være en fordel at få et godt estimat på den latente variabel først. Dette inkluderer også huludfyldning med og estimerer for koncentrationer under detektionsgrænsen. Når der foreligger gode estimerer for den latente variabel, vil det kunne styre Metropolis-Hastings algoritmens iterationer af deterministiske effekter til bedst muligt at beskrive den latente variabel. For at øge regnehastigheden på dette trin, holdes de stokastiske effekter ude. Når både den latente variabel og de deterministiske effekter er konvergeret (trin 1 og 2), inkluderes de stokastiske effekter i trin 3. Til sidst inkluderes den fulde model i trin 4, og modellens resultater må først anvendes, når dette sidste trin viser, at alt er konvergeret på plads.

Da de sidste justeringer i modellens strukturligning foregik i slutningen af projektperioden, er konvergeringsperioden ikke tilendebragt for trin 2 og 3 for de tre metaller bly, cadmium og nikkel. Trin 1 er dog fuldført, hvilket betyder, at den latente variabel er i tæt overensstemmelse med medianværdier estimeret med en simpel metode.

Forskellen mellem trin 2 og 3 illustreres i figurerne 7.1 og 7.2, hvor Figur 7.1 viser Metropolis-Hastings algoritmens iterationer på trin 2, hvor det kun er deterministiske effekter, der konvergeres på plads, mens Figur 7.2 viser iteration af både deterministisk og stokastiske effekter. Beregningerne i figurer 7.1 og 7.2 er sat i gang samtidig, så forskellen i antallet af gennemregninger, som er afbildet på x-akserne, viser forskellen i regnetid. Begge figurer er resultat af ca. 14 dages regnetid. Ved at sammenligne antallet af gennemregninger, ses således, at trin 2 regnes ca. 10 gange hurtigere end trin 3. På Figur 7.1 ses også en asymptotisk tendens til, at uoverensstemmelsen konvergerer mod en konstant værdi, mens det endnu ikke ses tydeligt på Figur 7.2. Der ses også en forskel mellem Figur 7.1 og 7.2 på y-aksen, hvor uoverensstemmelsen for Figur 7.1 er større end for Figur 7.2, hvilket er en naturlig konsekvens af, at indførelsen af stokastiske effekter gør modellen i stand til at nærme sig målingerne yderligere og dermed udvise en mindre uoverensstemmelse. Når iterationen i Figur 7.1 er konvergeret, kan de estimerede deterministiske effekter overføres til iterationen i Figur 7.2 for derved at hjælpe med til at konvergere.

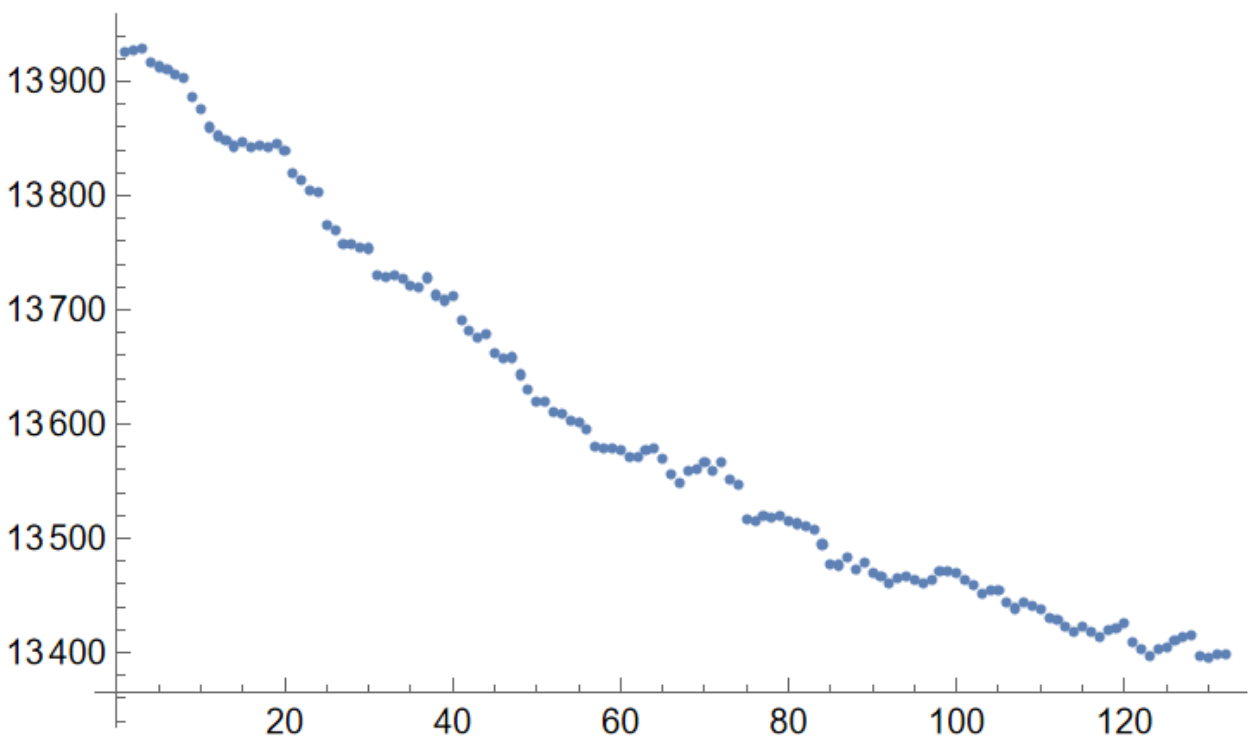
Det forventes, at iterationen af de tre metaller på trin 2 er tilendebragt i januar 2022. Derved kan iterationen på trin 3 opdateres med gode estimerer for de deterministiske effekter, og det forventes, at trin 3 er færdig i marts 2022. Derefter skal der bruges et par måneder på at sikre, at trin 4 er tilendebragt, og dermed at alle effekter har fået deres korrekte betydning. Iteration med bly, cadmium, nikkel, kobber og zink er også sat i gang, og det forventes at få en konvergeret model for alle fem metaller vil kunne være på plads omkring august 2022.

Når først modellen er konvergeret, er det en langt mindre regnetung opgave at opdatere modellen med ny viden såsom et nyt års måldata fra NOVANA.



Antal gennemregninger

Figur 7.1. Her ses, hvordan MetalStat konvergerer for metallerne bly, cadmium og nikkel på trin 2, hvor de stokastiske effekter ikke indgår. På y-aksen vises uoverensstemmelsen mellem model og målinger (deviation), mens antallet af gennemregninger vises på x-aksen.



Antal gennemregninger

Figur 7.2. Her ses, hvordan MetalStat konvergerer for metallerne bly, cadmium og nikkel på trin 3, hvor både de deterministiske og stokastiske effekter ikke indgår. På y-aksen vises uoverensstemmelsen mellem model og målinger, mens antallet af gennemregninger vises på x-aksen.

I det følgende vises kun de foreløbige resultater for de tre metaller bly, cadmium og nikkel, da disse er bedst konvergeret og derfor mest troværdige i deres estimater.

7.3 Analyse af målte vandløbsstationer

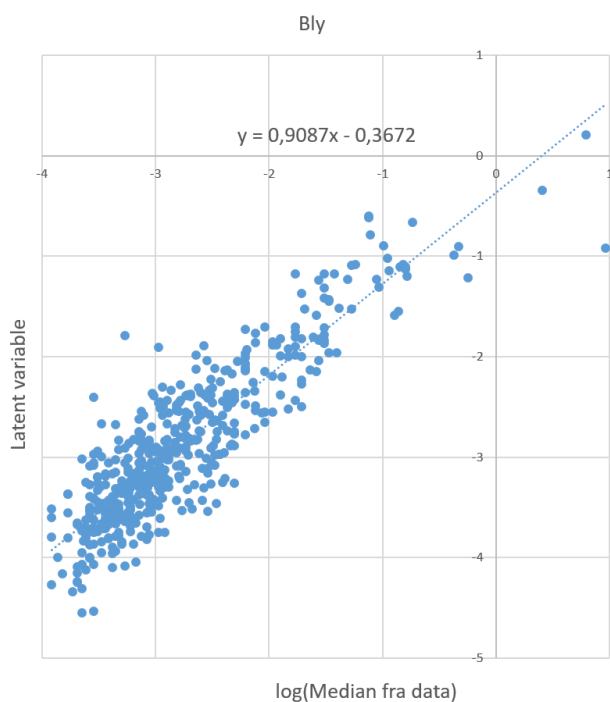
MetalStat indeholder en statistisk model for målingerne med median-koncentrationen som latent variabel, og derfor vil denne model kunne forbedre den hidtidige praksis med at tolke måledata. Da modellen stort set er konvergeret i forhold til den latente variabel (trin 1), vil nedenstående analyse i hhv. kapitel 7.3.1 og 7.3.2 være at betragte som et korrekt output.

Modellen vil naturligvis være mere præcis for målte områder end for umålte. De stokastiske effekter vil give en rettesnor for, hvor meget det kan bidrage med ny viden at øge antallet af målte vandløbsstationer, eller om der er potentiale i at forbedre oplandsmodellen til bedre at kunne identificere deterministiske effekter.

7.3.1 Sammenligning mellem MetalStat og en simpel rangordningsmodel

Når komplekse modeller testes, bør første skridt være at lave en generel test mod simple sammenhænge, som modellen bør kunne eftervise. I den forbindelse er det oplagt at undersøge, om modellens latente koncentrationer rent faktisk afspejler median-værdier for målte koncentrationer. Dette gøres ved at sammenligne simple estimater for mediankoncentrationer mod estimater af en latent variabel fra modellen. MetalStat estimerer middelværdien for log-normalfordelte koncentrationer på hver vandløbsstation og for hver sæson, hvilket svarer til logaritmen til median-koncentrationerne. Det er derfor muligt at finde mediankoncentrationer i måledata med en simpel rangordningsmetode, der rangordner de enkelte målinger, og tage værdien i midten af denne rang. I tilfælde af at denne værdi ligger under detektionsgrænsen, er estimatet dog udelukket. MetalStat kan godt finde median-værdier, der ligger under detektionsgrænsen, men da disse ikke kan estimeres med den simple rangordningsmetode, udgår de ved denne sammenligning med MetalStat.

På Figur 7.3 er medianestimer fra MetalStat sammenlignet med rangordningsmetoden for bly, og der ses en tydelig lineær relation mellem disse. Hvis MetalStat var helt sammenfaldende med rangordningsmetoden, ville der være en tæt lineær sammenhæng langs $y=x$, altså med hældningen én og skæring gennem nul. Der ses en tendens til en hældning på én, da den fittede værdi er 0,91, og skæringen ligger ret tæt på (0,0), men dog med en afvigelse på 0,37, der fremstår signifikant ($\text{Pr}<0,001$). Med en standardafvigelse (Std. error) for hældningen på 0,023 ses det, at hældningen på 0,91 ligger ca. fire standardafvigelser under én, og således er hældningen signifikant mindre end én ($\text{Pr}<0,001$). Det betyder, at der er en lille forskel mellem MetalStat og rangordningsmetoden.



```

Output

Call:
lm(formula = MeasModelMedian ~ DataMedian, data = Dataset)

Residuals:
    Min       1Q   Median       3Q      Max
-1.42558 -0.25382 -0.00783  0.25465  1.55055

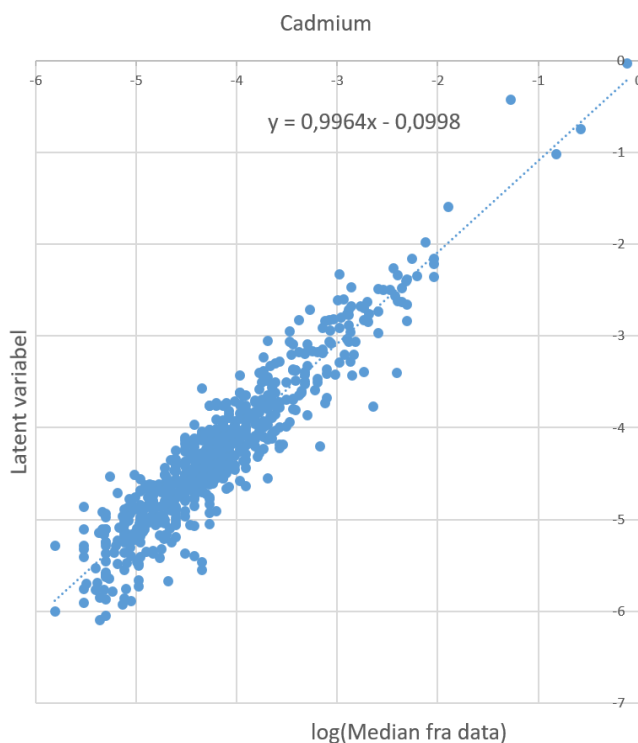
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.36721    0.06267   -5.86 0.00000000855 ***
DataMedian   0.90865    0.02188   41.54 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.366 on 487 degrees of freedom
Multiple R-squared:  0.7799, Adjusted R-squared:  0.7794
F-statistic: 1725 on 1 and 487 DF, p-value: < 2.2e-16

```

Figur 7.3. Sammenhæng mellem den estimerede latente variabel for log(median) (y-akse) med log(median) (x-akse) bestemt for bly med simpel rangordning, der hvor der er tilstrækkeligt med målinger over detektionsgrænsen.

Analysen for cadmium er vist på Figur 7.4. Her ses en hældning tæt på en (0,996) og en skæring tæt på (0,0), som ikke er signifikant (pr=0.12).



```

Output

Call:
lm(formula = MeasModelMedian ~ DataMedian, data = Dataset)

Residuals:
    Min       1Q   Median       3Q      Max
-1.12695 -0.17818  0.01169  0.19742  0.93882

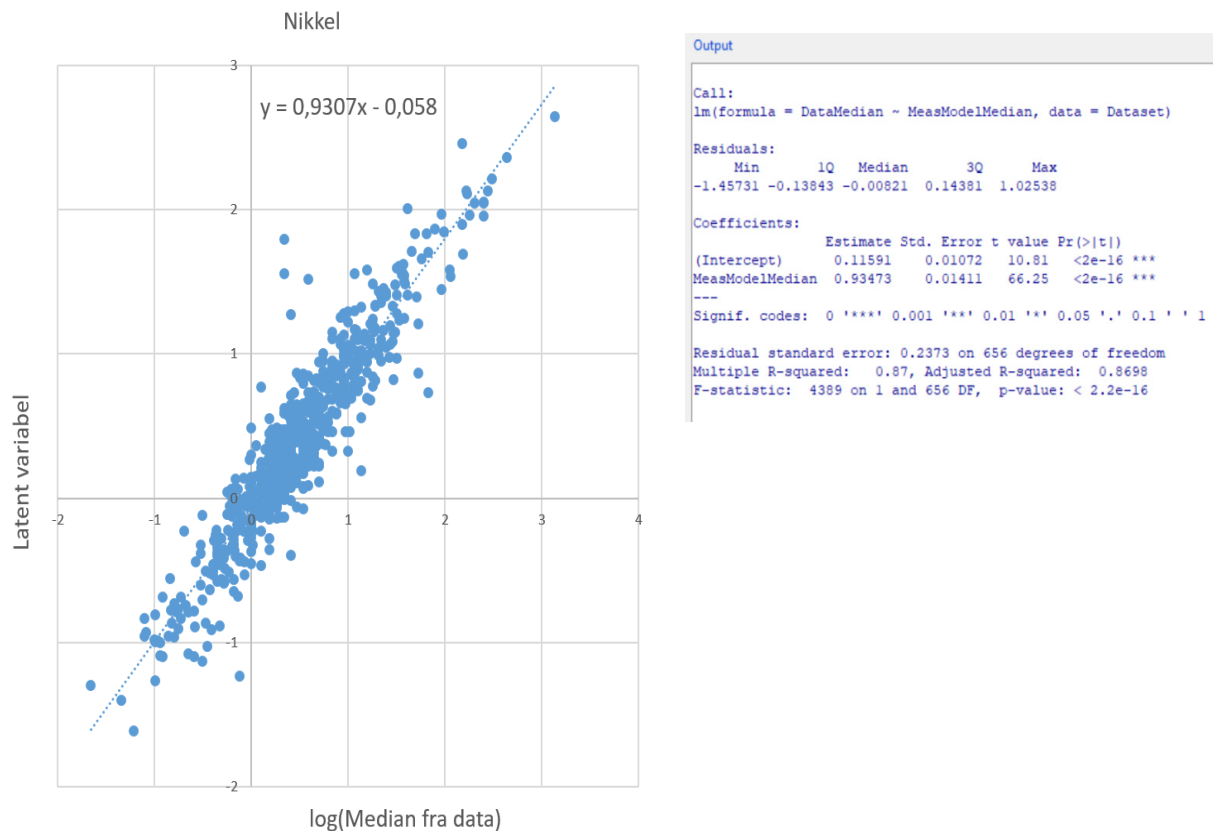
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.09980    0.06448   -1.548  0.122
DataMedian   0.99641    0.01514   65.794 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2979 on 614 degrees of freedom
Multiple R-squared:  0.8758, Adjusted R-squared:  0.8756
F-statistic: 4329 on 1 and 614 DF, p-value: < 2.2e-16

```

Figur 7.4. Sammenhæng mellem den estimerede latente variabel for log(median) (y-akse) med log(median) (x-akse) bestemt for cadmium med simpel rangordning, der hvor der er målinger nok over detektionsgrænsen.

Analysen for nikkel er vist på Figur 7.5. Her ses en relativt tæt sammenhæng med en hældning tæt på en (0,93) og en skæring tæt på (0,0). Både hældning og skæring er dog signifikant ($Pr < 0,001$) forskellig fra hhv. en hældning på en og en skæring gennem (0,0).



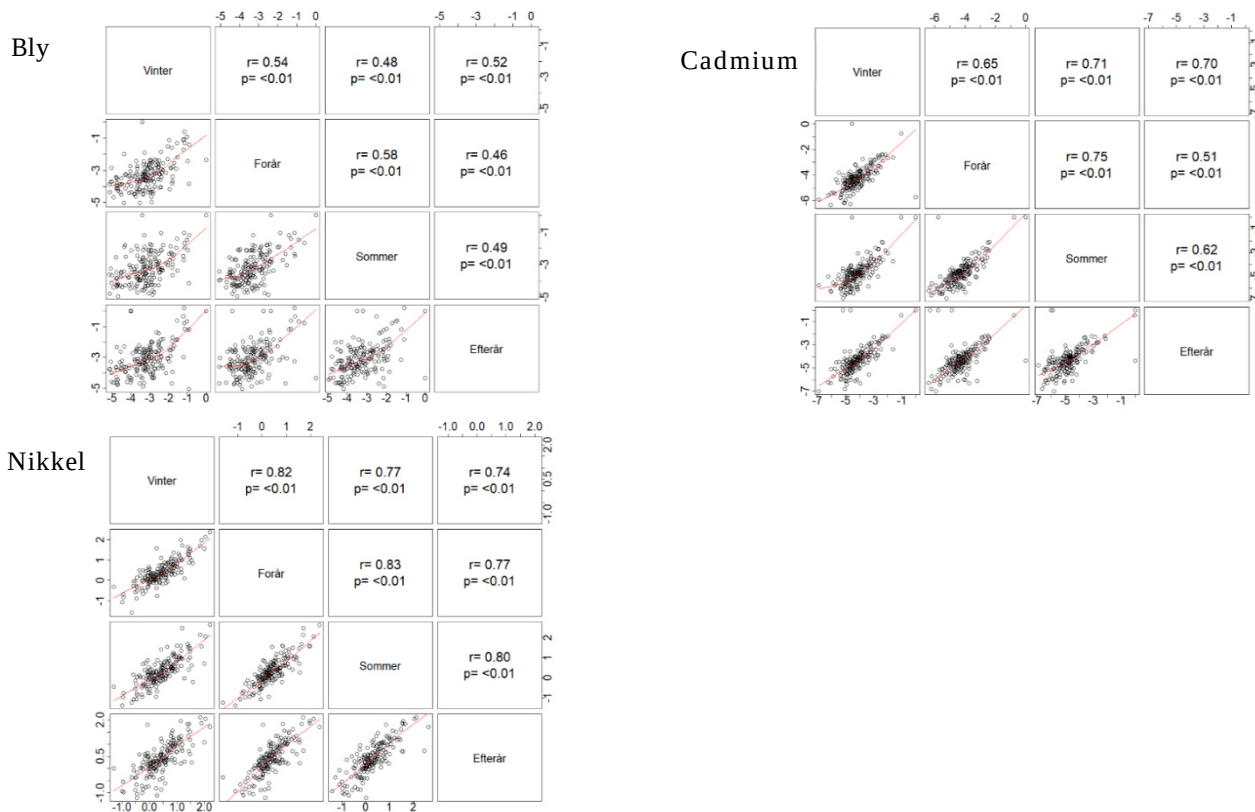
Figur 7.5. Sammenhæng mellem den estimerede latente variabel for log(median) (y-akse) med log(median) (x-akse) bestemt for nikkel med simpel rangordning, der hvor der er målinger nok over detektionsgrænsen.

Der er som vist på figurerne 7.3-7.5 en tæt sammenhæng mellem MetalStat estimerterne for de latente median-koncentrationer og et uafhængigt estimat af samme medianværdier med brug af en rangordningsmetode. Der er dog fundet en mindre afvigelse for bly og nikkel, men det kan lige såvel være pga. usikkerhed for estimatet med brug af en simpel rangordningsmetode som metoden bag MetalStat.

7.3.2 Analyse for forskelle mellem sæsoner

I Figur 7.6 er de fire sæsoners estimerter for de tre metaller afbilledet mod hinanden i en matrice for hver målt vandløbsstation og for hver parvise kombination af fire sæsoner. Korrelationskoefficienten vises i samme matrice i øverste højre trekant i hver matrice. Hvis der er en klar forskel mellem vandløbsstationer, bør det typisk medføre, at en vandløbsstation med et højt koncentrationniveau for en sæson også vil have en tendens til at have en høj koncentration for de andre sæsoner, altså at koncentrationerne ligger langs en $y=x$ linje i Figur 7.6. For bly ses relativt begrænset korrelation (r omkring 0,5). Dette tyder på, at forekomsten af bly ikke har dominerende stedspecifik effekter, der klart kan adskille vandløbsstationer fra hinanden, svarende til overstående forklaring. En tilsvarende analyse er i Figur 7.6 vist for cadmium, hvor der ses en noget mere udtalt korrelation sammenlignet med bly. For nikkel ses på Figur 7.6 den

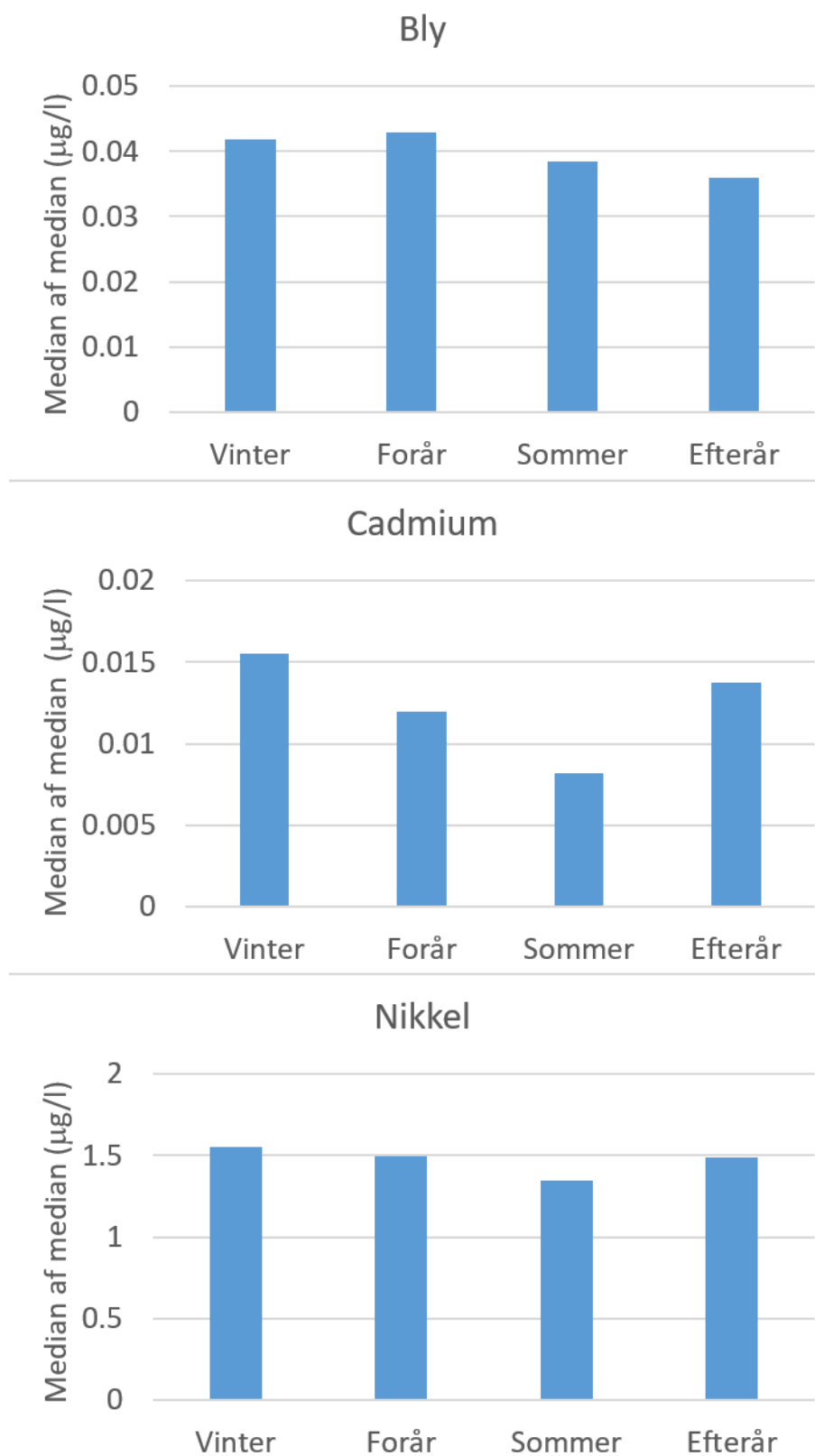
stærkeste korrelation mellem sæsoner sammenlignet både med bly og cadmium. Så dette tyder på, at bly ikke følger stærke mønstre, der kan beskrives som forskelle mellem vandløbsstationer, mens det i mere udtalt grad er tilfældet for cadmium og endnu mere for nikkel.



Figur 7.6. De fire sæsoners estimater for de 3 metaller for hver målt vandløbsstation er afbilledet for alle parvise kombinationer af to sæsoner i en matrice. Korrelationskoefficienten vises i samme matrice i øverste højre trekant af hver matrice, hvor p værdi svarer til test af H_0 hypotesen, at koncentrationsniveauet for to sæsoner er uafhængige af hinanden.

På Figur 7.7 ses median-værdierne af de vandløbsspecifikke medianværdier for hver målt vandløbsstation opdelt mellem hver sæson. Hvis der er forskel mellem sæson i medianen af median-værdierne, vil det kunne afsløre sæsonafhængige kilder. I Figur 7.7. ses det, at medianen for medianen af cadmium er mindst om sommeren, hvilket kunne pege på, at det er den overfladenære afstrømning i det åbne land, der typisk er størst i vinterhalvåret, som bringer mest cadmium til vandløbet. Hvis cadmium var bragt til vandløbet med f.eks. by-spildevand og grundvand, ville det således ikke vise den sammenhæng, som ses i Figur 7.6., da by-spildevand og grundvand typisk har størst andel af vandløbets vandføring om sommeren.

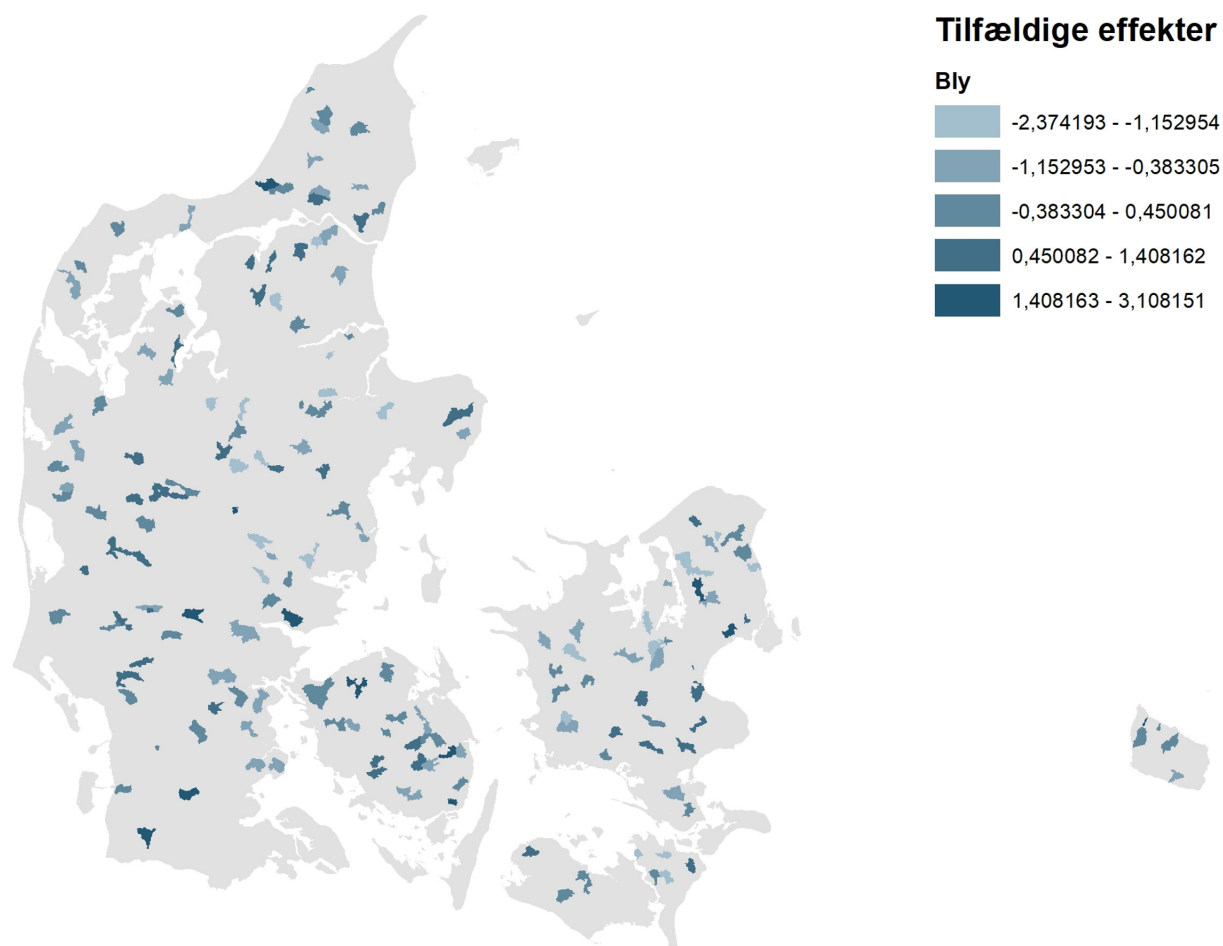
Figur 7.7. Median af median for målte vandløbsstationer fordelt mellem sæson.



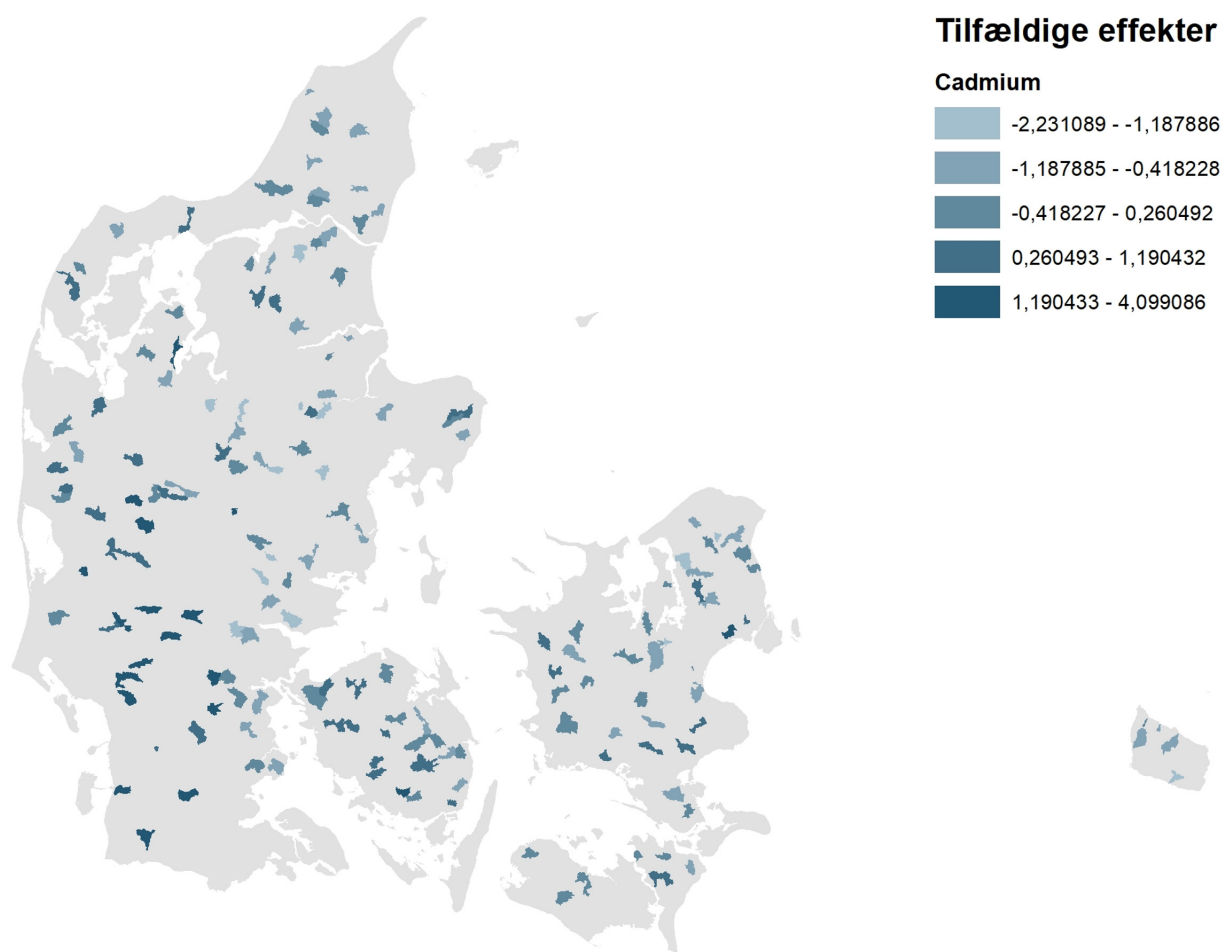
7.3.3 Analyse for stokastiske effekter mellem vandløbsstationer

De stokastiske effekter afslører, om der er forskelle i metalkoncentrationer, der ikke kan beskrives med oplandsmodellen. Det er derfor meget vigtigt at undersøge modellen for stokastiske effekter, og dette er det en stor hjælp at

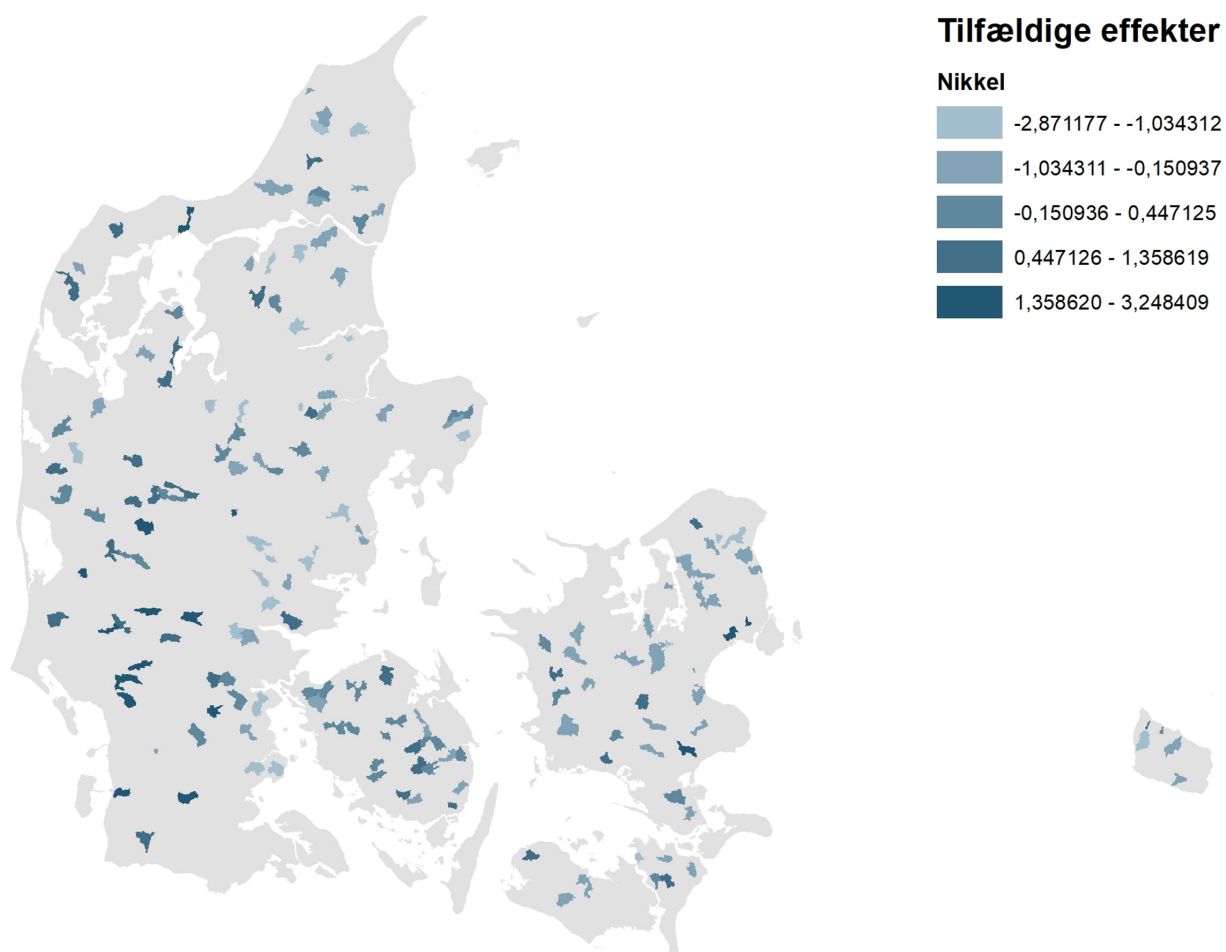
afbillede de estimerede stokastiske effekter i GIS for at se, om der er et geografisk mønster, se figurerne 7.8-10. Det ses, at de stokastiske effekter for bly ligger spredt over hele landet, mens niveauet for cadmium og nikkel synes at have en tendens til at være størst i Vestjylland. Da de stokastiske effekter viser forskelle, som oplandsmodellen ikke kan beskrive, vil en nærmere granskning af de specifikke vandløbsstationer med høje værdier kunne indkredse kritiske effekter, som modellen ikke har med, og som derfor bør inddrages i en forbedret version.



Figur 7.8. Stokastiske effekter mellem vandløbsstationer for bly svarende til $\log(F_i)$ for bly (Ligning 2).



Figur 7.9. Stokastiske effekter mellem vandløbsstationer for cadmium svarende til $\log(F_i)$ for cadmium (Ligning 2).



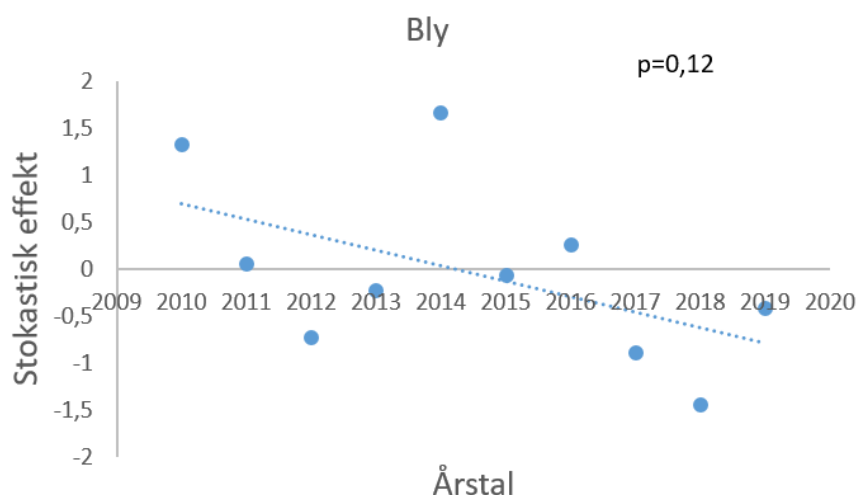
Figur 7.10. Stokastiske effekter mellem vandløbsstationer for nikkel svarende til $\log(F_i)$ for nikkel (Ligning 2).

7.3.4 Analyse af stokastiske effekter mellem år

Hvis der er en generel tendens til, at median-koncentrationer i et bestemt år ligger på et andet niveau end i andre år, vil det kunne ses som stokastisk effekt mellem år. Oplandsmodellen regner stationært for hele den målte periode, så alle effekter, der udtrykker en tidstrend fra år til år, vil bidrage til de stokastiske effekter mellem år.

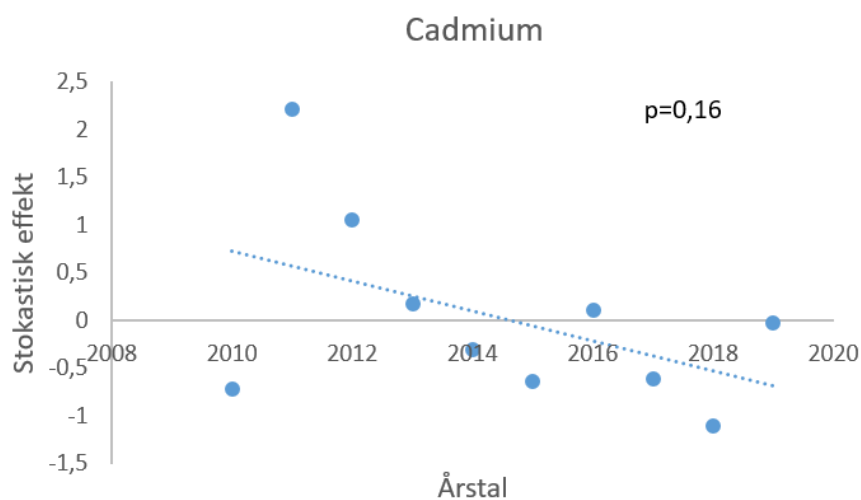
På Figur 7.11 ses de stokastiske effekter mellem år for bly, hvor der ses et vist fald. En simpel lineær model for det tidsmæssige forløb af de stokastiske effekter mellem år viser dog, at faldet ikke synes signifikant ($p=0,12$), men hvis tendensen fortsætter de næste få år, kan den faldende koncentration mellem år måske blive signifikant.

Figur 7.11. Stokastiske effekter mellem år for bly ($\log(F_n)$ i Ligning 2), hvor p -værdien tester H_0 hypotesen, at der ingen lineær tidstrende er.



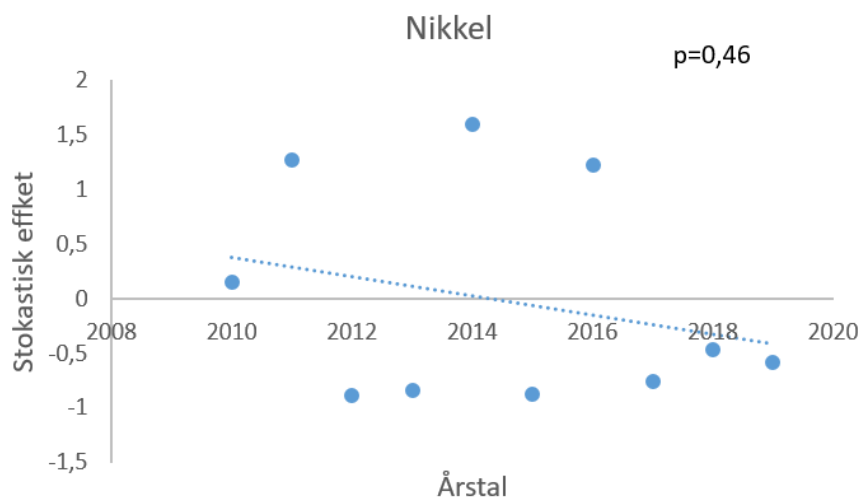
En lignende analyse er gennemført for cadmium (Figur 7.12), og her ses også et fald i median koncentration, der dog heller ikke er signifikant ($p=0,16$). I øvrigt ses det i Figur 7.12, at år 2011 havde en markant større og positiv stokastisk effekt sammenlignet med de andre år, mens år 2010 omvendt havde en negativ effekt, så nærmere granskning af forskellen mellem år 2010 og 2011 synes relevant i bestræbelse på at forbedre forståelsen af vigtige transportveje for cadmium.

Figur 7.12. Stokastiske effekter for cadmium mellem år ($\log(F_n)$ i Ligning 2), hvor p -værdien tester H_0 hypotesen, at der ingen lineær tidstrende er.



Nikkel er analyseret for stokastiske effekter mellem år (Figur 7.13), og her synes der ikke at være nogen tydelig tidstrend på samme måde som indikeret for bly og cadmium.

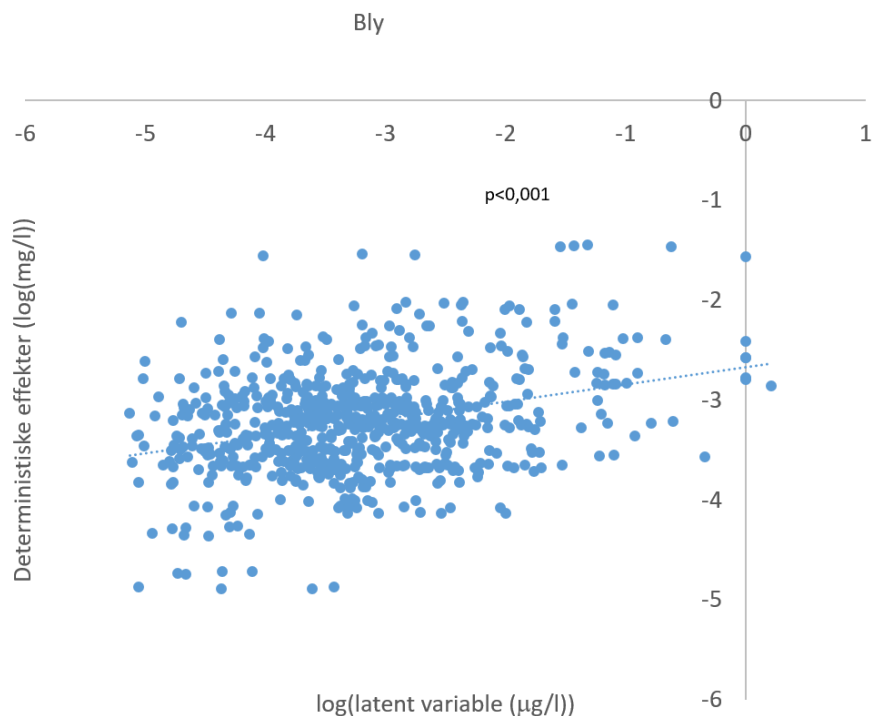
Figur 7.13. Stokastiske effekter mellem år for nikkel ($\log(F_n)$ i Ligning 2), hvor p -værdien tester H_0 hypotesen, at der ingen lineær tidstrende er.



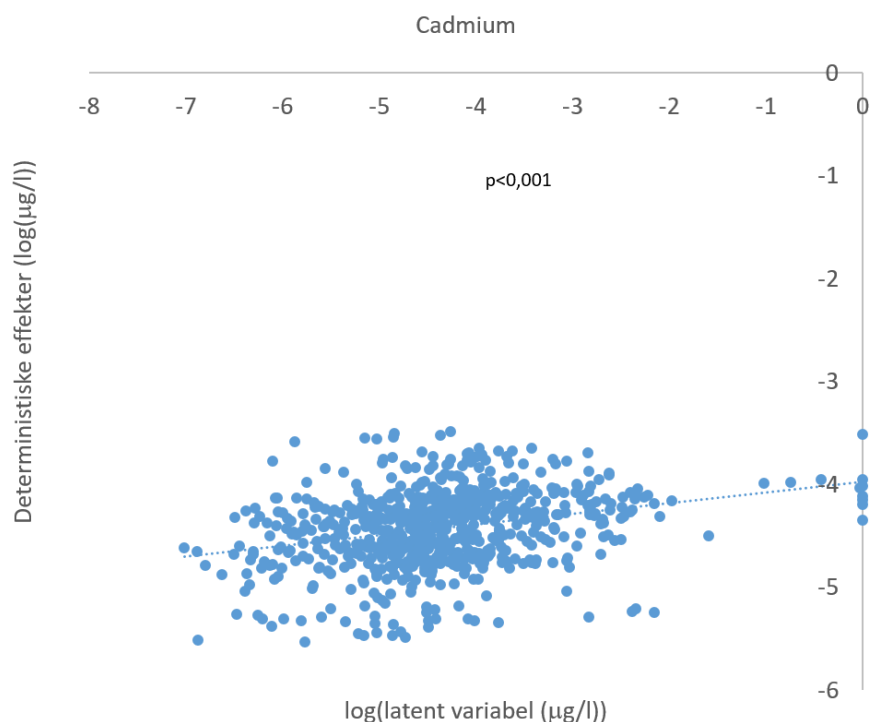
7.3.5 Analyse af oplandsmodellen

Modellens evne til at adskille forskelle i metalkoncentrationer geografisk og derudfra udpege områder med høje koncentrationsniveauer fra områder med lave koncentrationer afhænger af, hvor godt de deterministiske effekter kan prædiktere den latente mediankoncentration. Det ses tydeligt af figurerne 7.14-7.16, at modellens prædiktioner med den nuværende parametrisering viser, at der er en svag tendens til en vis prædiktions for bly. Da modellen ikke er konvergeret, kan denne manglende prædiktions dog endnu ikke tages som udtryk for, hvor godt oplandsmodellen kan beskrive en reel variation i koncentrationsniveauerne.

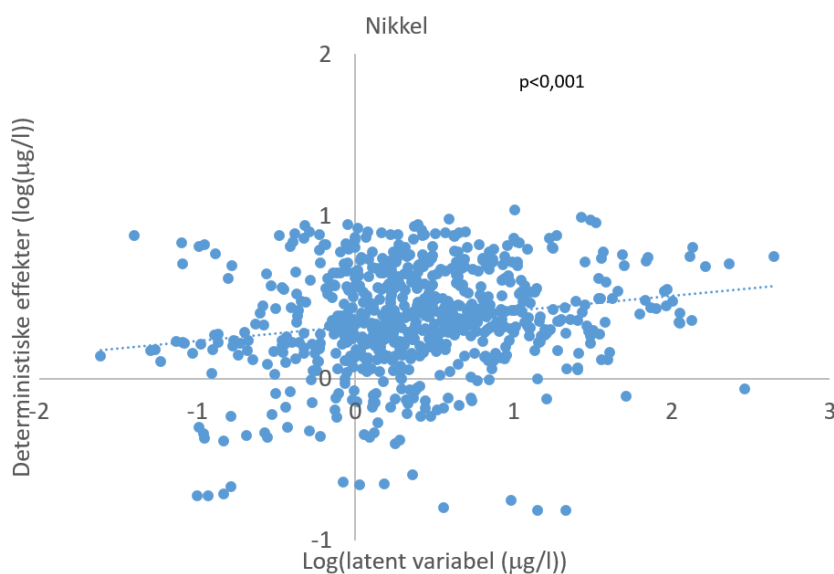
Figur 7.14. Prædiktions af latent mediankoncentration for bly mod prædiktions fra deterministiske modelkomponenter. p -værdien tester H_0 hypotesen, at hældning er 0 for det fittede stiplede linje.



Figur 7.15. Prædiktion af latent mediankoncentration for cadmium mod prædiktionen fra deterministiske modelkomponenter. p -værdien tester H_0 hypotesen, at hældning er 0 for det fittede stiplede linje.



Figur 7.16. Prædiktion af latent mediankoncentration for nikkel mod prædiktionen fra deterministiske modelkomponenter. p -værdien tester H_0 hypotesen, at hældning er 0 for det fittede stiplede linje.



7.4 Landsdækkende estimering

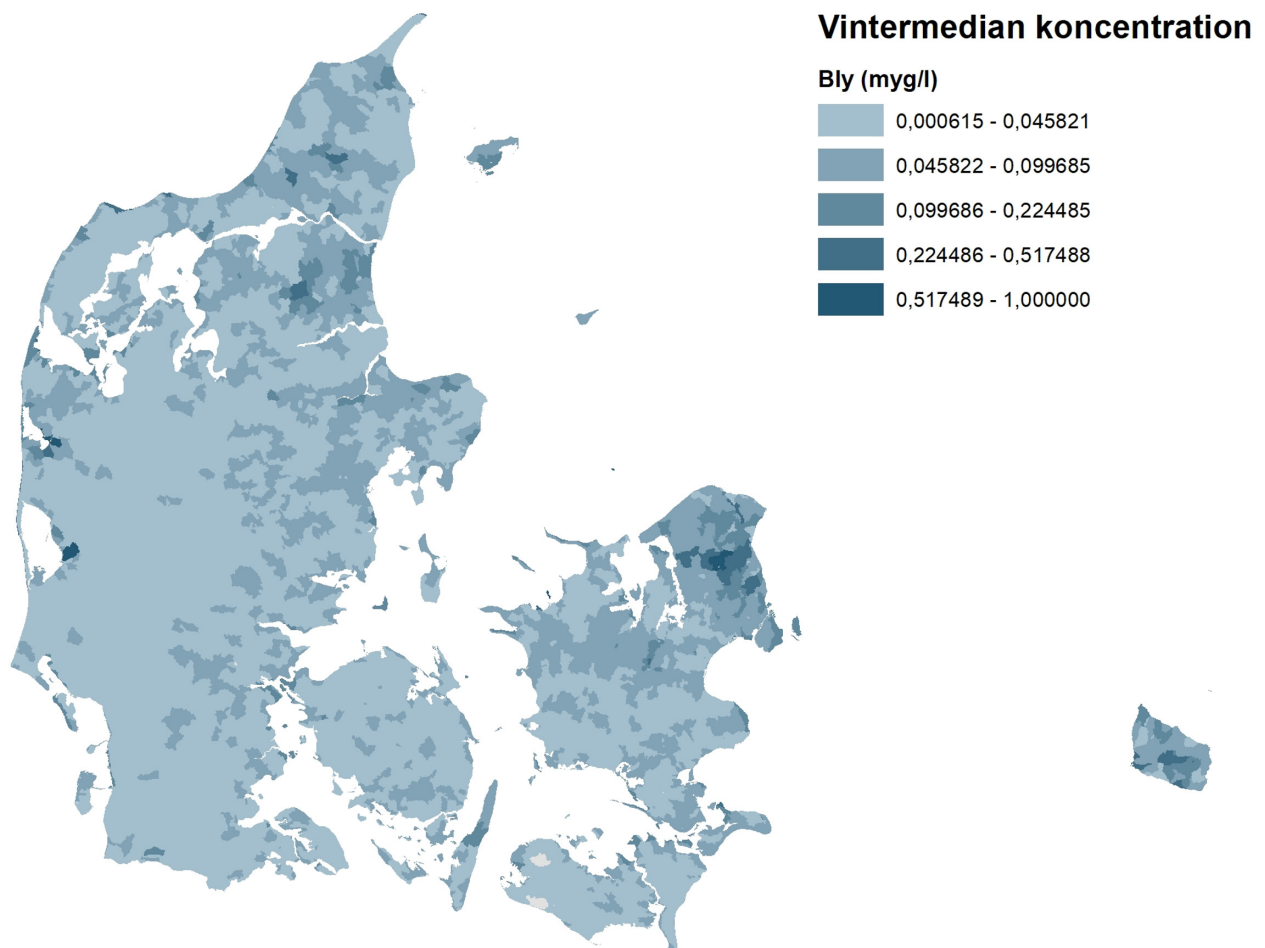
7.4.1 Geografisk fordeling af estimater

MetalStat kan estimere median-koncentrationsniveauer med brug af de deterministiske effekter for alle ID 15 oplande. På Figur 7.17-7.19 ses således median-estimer for vintersæsonen for hhv. bly, cadmium og nikkel. For bly og cadmium ses tydelige regionale forskelle, mens nikkel er mere jævnt fordelt over landet. Det er vigtigt at undersøge rimeligheden i de forskelle, der afspejles med sådanne kortlægninger, før de tages til indtægt som gyldige, og i den forbindelse

at det vurderes, om de uafhængige variable, der giver de regionale forskelle, med rimelighed kan siges at reflektere en reel kausalitet.

Der vil også kunne beregnes kort for de andre tre årstider, og det er vigtigt at påpege, at der er stor usikkerhed ved estimatet for de enkelte oplande, hvilket bliver illustreret med et enkelt opland på Figur 7.20.

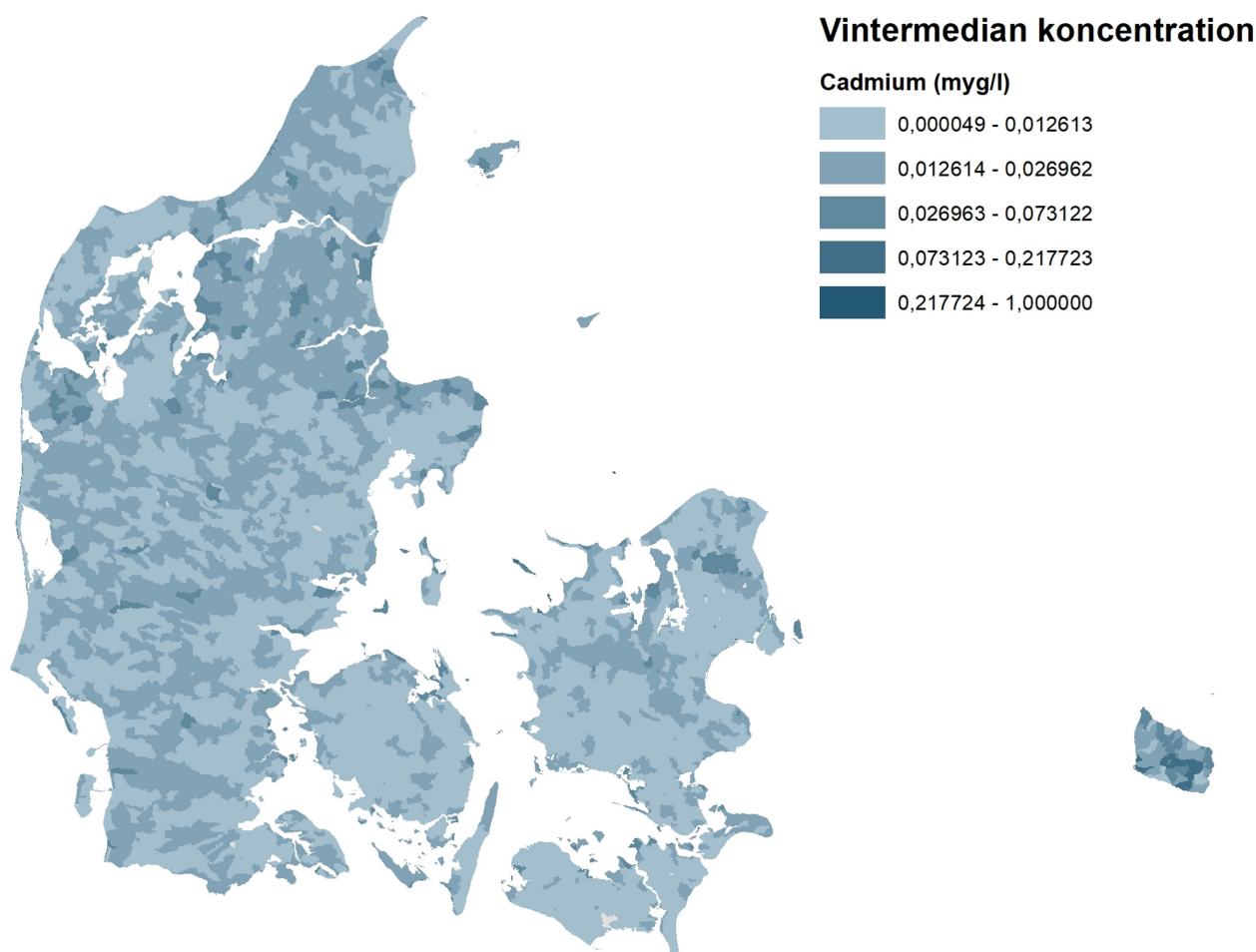
De estimerede median-koncentrationer for bly ses på Figur 7.17, hvor der synes at være en gradient over israndslinjen ned gennem Jylland med de laveste koncentrationer mod vest. Som første skridt i en analyse af dette resultat er de stokastiske effekters geografiske fordeling, som vises på Figur 7.8, vigtige, da de kan afsløre, om denne gradient er et artefakt pga. stokastiske effekter. På Figur 7.8 ses ikke en tydelig tendens til, at de stokastiske effekter er større eller mindre over israndslinjen, hvilket underbygger, at der rent faktisk er lavere blykoncentration i det vestlige Jylland.



Figur 7.17. Prædiktion af mediankoncentration for bly og fordelt på alle ID 15 oplande for vintersæsonen.

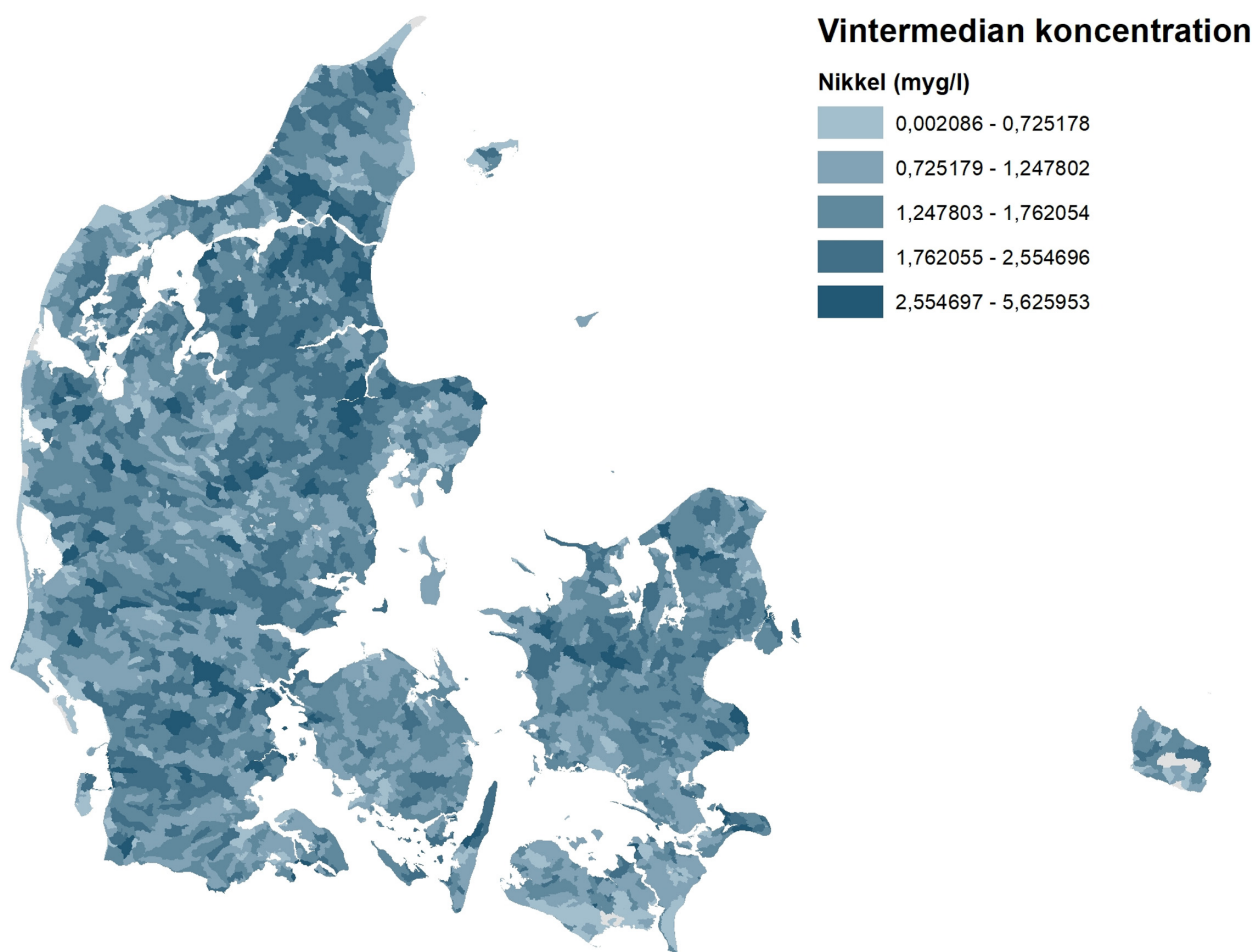
Betragtes cadmium (Figur 7.18), ses det, at koncentrationen af cadmium har en tendens til at være højere i det nordlige Jylland og nordvestlige Sjælland sammenlignet med Fyn og Vestjylland. Hvis denne forskel sammenholdes med den geografiske udbredelse af de stokastiske effekter for cadmium (Figur 7.9), ses det, at de stokastiske effekter for cadmium giver større bidrag til koncentrationsniveauet i det vestlige Midtjylland sammenlignet med resten af

landet, hvilket betyder, at Nordjylland og Nordvestsjælland ikke nødvendigvis er de eneste potentielle højniveau-områder, da den deterministiske model underestimerer koncentrationerne i Vestjylland.



Figur 7.18. Prædiktion af mediankoncentration for cadmium og fordelt på alle ID 15 oplande for vintersæsonen.

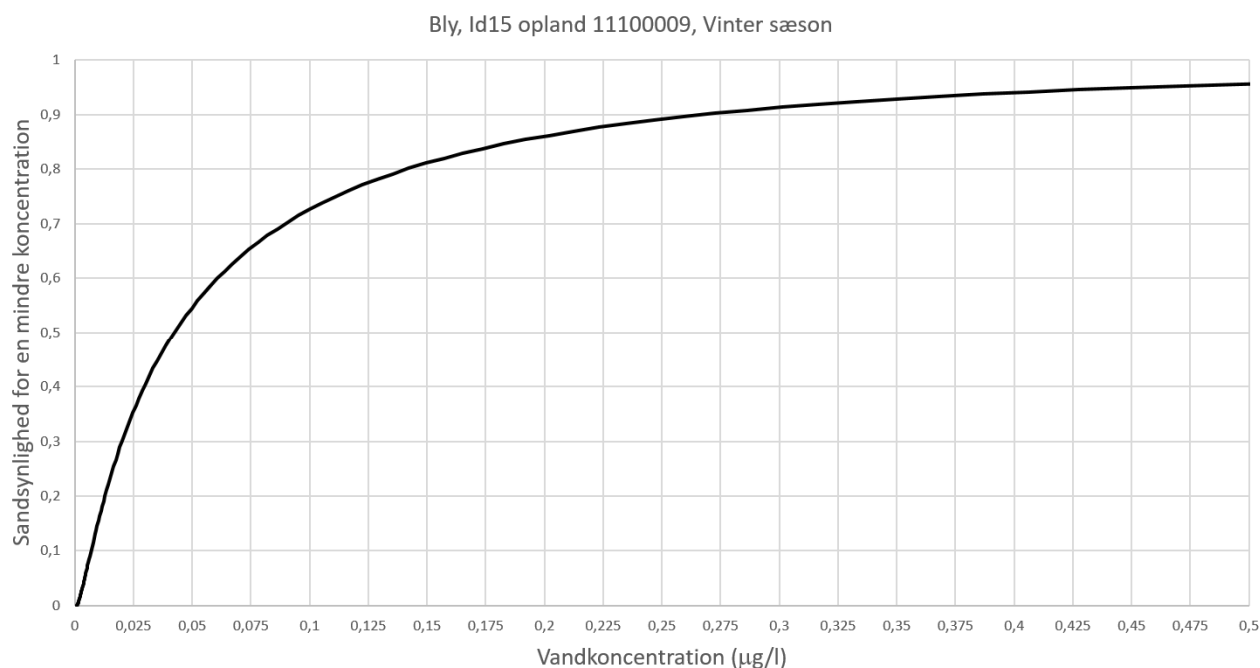
For nikkel ses der i Figur 7.19 ikke nogen forskel mellem hhv. det østlige og vestlige Jylland, men på Figur 7.10 ses en klar tendens til høje værdier af stokastiske effekter i det vestlige Jylland. Derfor kan denne manglende forskel mellem det østlige og vestlige Jylland være udtryk for, at oplandsmodellen ikke helt får alle faktorer med, der har betydning for at kunne adskille nikkelkoncentrationen i hhv. det østlige og vestlige Jylland.



Figur 7.19. Prædiktion af mediankoncentration for nikkel og fordelt på alle ID 15 oplande for vintersæsonen.

7.4.1.1 For enkelte oplande

Usikkerhed ved estimering af koncentrationstværdier kan plottes for alle oplande. I Figur 7.20 ses således prædiktion af koncentrationen i det umålte ID 15 opland 11100009 vist som en akkumuleret sandsynlighed, hvor der således er en 90 % sandsynlighed for, at koncentrationen er under 0.27 µg/l.



Figur 7.20. Den kumulerede sandsynlighedsfordeling af koncentrationsestimatet for bly og for ID 15 opland 11100009 (vinter sæson).

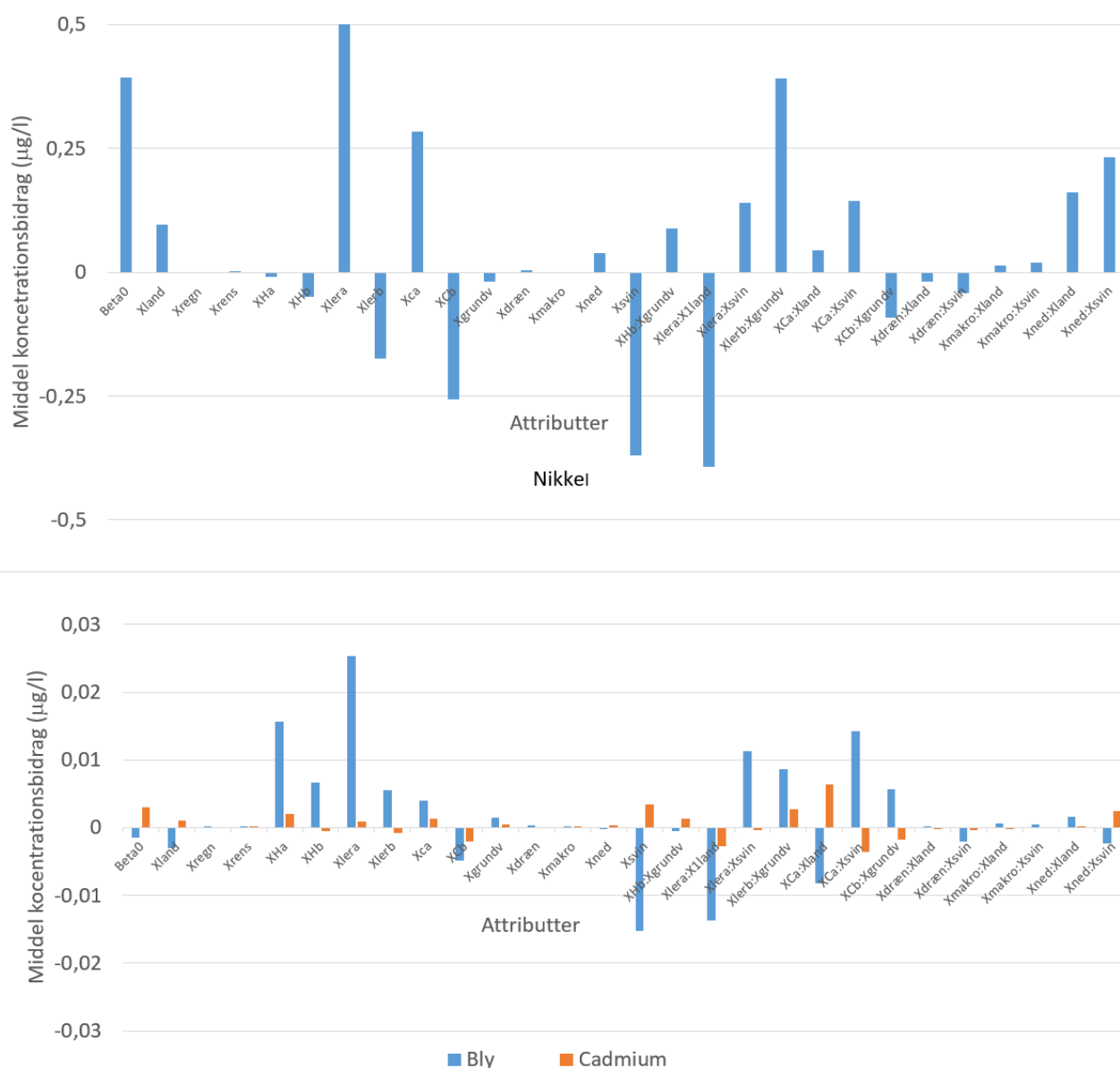
7.4.2 Parameterestimering

Resultatet fra parameterestimering er vist i Bilag C. I denne ses median for hver parameter samt usikkerheden udtrykt som hhv. nedre 2,5 % fraktil samt øvre 97,5 % fraktil. Desuden vises en p -værdi svarende til sandsynligheden for, at parameteren kan være positiv, så hvis $p=1$, betyder det, at parameteren er sikkert bestemt som værende positiv, og derfor at den pågældende effekt er signifikant, mens værdien -1 betyder, at den pågældende effekt påvirker koncentrationen negativt. Hvis p -værdien er mellem -1 og 1 , betyder det, at den pågældende effekt ikke er entydig og dermed ikke stærk signifikant, og ud fra typiske konventioner vil en p -værdi i intervallet $-0,975 - 0,975$ betragtes som værende ikke signifikant. Det er vigtigt at påpege, at en signifikant parameter ikke behøver at være betydningsfuld, men omvendt vil en ikke-signifikant parameter sjældent blive betragtet som betydningsfuld. Om en parameter fremkommer med en værdi, der er reel betydningsfuld, kan ikke direkte ses i Bilag C, men kræver en nærmere analyse, som illustreret i næste afsnit.

7.4.3 Betydning af oplandskarakteristika

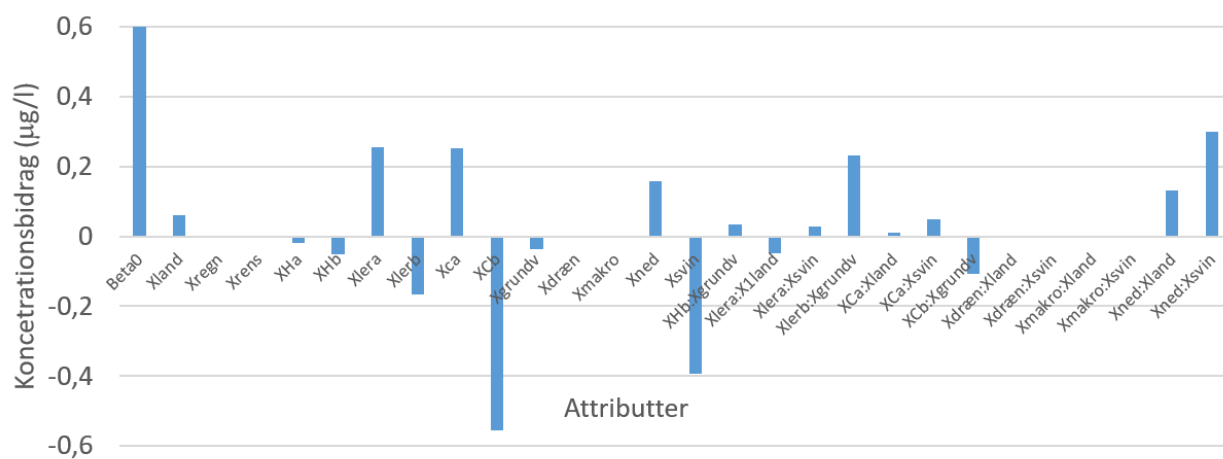
For hvert opland kan de estimerede parametre (Bilag C) bruges til at beregne betydningen for prædikterede koncentrationer for de forskellige uafhængige variable ved at bruge ligningerne 1 og 8. Ved på den måde at bestemme den estimerede betydning for alle oplande kan en gennemsnitlige betydning bestemmes for hele landet som et generelt udtryk for den pågældende uafhængige variabels estimerede betydning for koncentrationsniveauet af metallet, se Figur 7.21. De uafhængige variable er beskrevet i kap. 6.2. Denne analyse skal dog anvendes med forsigtighed, da $+$ / $-$ effekter kan forstyrre en tolkning. Således ses der for nikkel en stor positiv betydning fra kulstof i a-horisonten (Xca), men en tilsvarende negativ betydning for kulstof i b-horisonten (Xcb), hvilket giver grund til at genoverveje X matricen og/eller udstikke nogle begrænsninger for β . En relevant begrænsning vil være at fastholde $\beta > 0$,

svarende til at oplandsmodellen estimerer mulige bidrag af metal i interaktion mellem transportvariable.



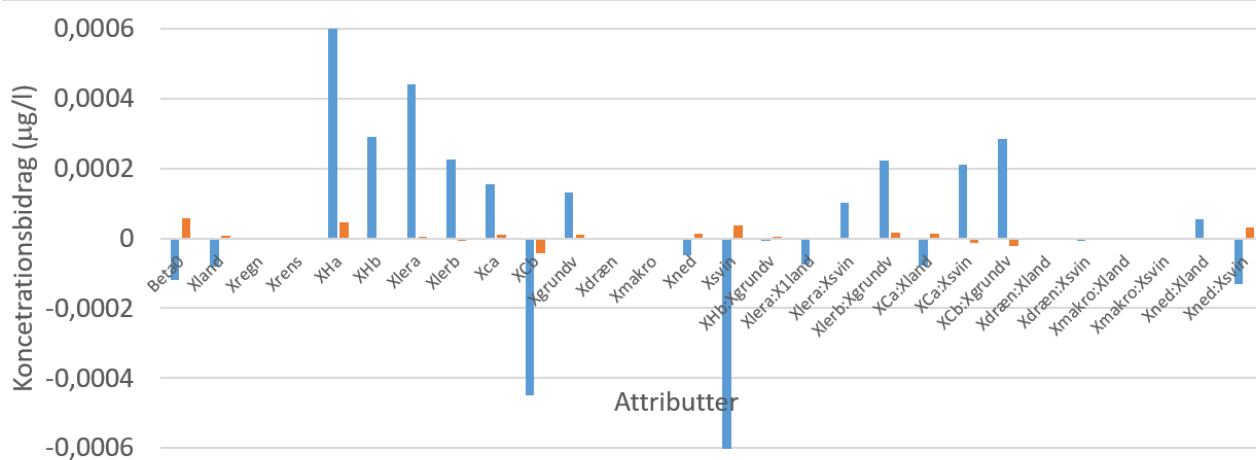
Figur 7.21. Estimeret koncentrationsbidrag for nikkel øverst samt bly og cadmium nederst fra MetalStats uafhængige variable som middel for hele Danmark. Beta0 svarer til Ligning 9.

Det er også muligt at se estimeret betydning i forhold til de enkelte oplande og bruge dem som rettesnor for, hvad der kan have betydning for koncentrationsniveauet i oplandet netop der, som vist på Figur 7.21 for ID 15 oplandet 11100001.



Attributter

■ Nikkel



Attributter

■ Bly ■ Cadmium

Figur 7.22. Estimeret koncentrationsbidrag for nikkel øverst samt bly og cadmium nederst fra MetalStats uafhængige variable som middel for ID 15 opland 11100001. Beta0 svarer til Ligning 9.

8 Konklusioner

8.1 Afgrænsning

Beregningsarbejdet med at få modellen til at konvergere har vist sig at være stort pga. de mange variable og er ikke tilendebragt ved afslutningen af nærværende projekt, fordi strukturligningen blev modificeret hen mod slutningen af projektperioden. Hver eneste manglende måling og målinger under detektionsgrænsen er ukendte for modellen og bliver derfor estimeret som en variabel, hvilket bidrager til et stort regnearbejde. Det betyder, at udviklingsarbejdet besværliggøres, da enhver rettelse og ændring betyder, at modellen skal starte helt eller delvist forfra med at konvergere. Til gengæld håndterer modellen manglende målinger på en troværdig måde, både i forhold til målinger under detektionsgrænsen og helt manglende målinger for enkelte metaller, hvor andre metaller er målte. Det vil ikke være det samme problem under drift, når først modellen er konvergeret. Men det har for projektet betydet, at kun de tre metaller bly, cadmium og nikkel er tættest på at være konvergeret, da modellen blev justeret i sin strukturligning i december, hvorved meget af den foregående kalibrering gik tabt. Så den oprindelige liste af metaller blev afkortet til de fem metaller bly, cadmium, nikkel, kobber og zink i første omgang, og for at få beregninger så tæt på at konvergere som muligt, er disse fem metaller derefter yderligere reduceret til bly, cadmium og nikkel for afrapportering i denne rapport. Dog er konvergering af modellen med alle fem metaller fortsat efter projektets afslutning, og resultatet heraf vil kunne anvendes som opstart i en eventuel drift-model, der derved ikke får samme behov for at konvergere. Det anbefales, at de øvrige metaller barium, chrom, kviksølv og arsen først inkluderes i modellen, når de første fem metaller er bragt i drift.

Projektet har haft fokus på vandkoncentrationen af metaller, da det er her, der er flest data, men andre matricer, såsom sediment, kunne godt inkluderes i modellen, når modellering af vandkoncentrationen er implementeret.

8.2 Brug af MetalStat i tilrettelæggelsen af overvågningsprogrammet for metaller

Modellen, der er udviklet i dette projekt, er en prototype, som tester, om en statistisk model, hvori der indgår en oplandsmodel, kan bruges til bedre at tolke overvågningsdata og derved gøre det muligt at prædiktere koncentrationniveauer for de ikke-målte områder. De konkrete beregningsresultater i denne rapport kan ikke bruges som grundlag for beslutninger vedr. de metaller, der er afbildet, og dette gælder især, da modellen i skrivende stund fortsat er ved at konvergere. Denne rapport viser dog, at brug af modellen er lovende, og det åbner op for et nyt paradigme, hvor modeller bruges mere aktivt i synergi med overvågningsaktiviteter for MFS. Modeller af den type, som er udviklet i dette projekt, kan bidrage med:

- *Forbedret og mere præcis tolkning af overvågningsdata.* En model kan forbedre evnen til at finde sammenhænge i overvågningsdata, der kan understøtte en bedre forståelse for metallernes skæbne og tilførsel til vandmiljøet. Dette eksemplificeres i rapporten ved, at bly udviste en mere diffus kontaminering mellem målte vandløbsstationer end cadmium og nikkel, hvilket understøtter en formodning om, at forekomsten af bly i højere grad end cadmium og nikkel er styret af en gammel historisk kontaminering fra bly i benzin og derfor har spredt sig mere jævnt mellem alle vandløbsstationer. Som andet eksempel viste den stokastiske effekt for år sig som en tidstrend for bly og cadmium, der dog endnu ikke er signifikant.
- *Forudsigelse af koncentrationsniveauer uden for overvågede områder.* Ud fra de målte vandløbsstationer i overvågningsprogrammet kan koncentrationsniveauerne estimeres uden for overvågede områder. Troværdigheden af denne estimering vil dog afhænge af, hvor godt det er muligt at beskrive vigtige mekanismer i en landsdækkende oplandsmodel og derfor være meget stofafhængig. Rapporten illustrerer, hvordan en statistisk model er nødvendig, da den kan kortlægge usikkerheden bag estimerer uden for overvågede områder, så der ikke opstår en falsk sikkerhed i forhold til, at værdier får for stor troværdighed.
- *Evaluerings af mulige kilder til forekomsten af metaller.* De kilder til forekomst af metaller, som er med i modellen, får tillagt en betydning, der viser deres bidrag til koncentrationsniveauerne både for de enkelte oplande og for hele landet. Det er dog vigtigt at påpege, at denne betydning kun er retningssigende og bør følges op af en miljøkemisk granskning, da den korrelation, som modellen finder, altid bør vurderes med miljøkemisk ekspertise, før den godtages som reel.
- *Tilrettelæggelse af fremtidigt overvågningsprogram for metaller i vandløb.* Det vil være oplagt at prioritere et fremtidigt overvågningsprogram mod de ikke-moniterede områder, som ifølge MetalStat vil have størst sandsynlighed for at være de områder, som har de højeste koncentrationsniveauer. Desuden vil den forbedrede forståelse for koncentrationsniveauerne i overvågede områder give ny viden til planlægning af et fremtidigt overvågningsprogram. F.eks. viser MetalStat, at tilfældige effekter for cadmium mellem målte vandløbsstationer er højest i vestjyske vandløb, hvilket antyder, at der kan være hot spots med lokale høje koncentrationer.
- *Evaluerings af den eksisterende forståelse af metaller forekomst.* Da modellen er statistisk, åbner det op for en afsøgning af kilder til usikkerhed:
 - Hvor godt kan vi forudsige et koncentrationsniveau, både i de overvågede områder og uden for disse?
 - Er der nogle mønstre i vores manglende forståelse, der kan give rettesnor for nye undersøgelser / overvågning?
 - Hvor stor en betydning kan tillægges de forskellige kilder, der er med i oplandsmodellen, og hvor meget kan disse forklares af den samlede variation i koncentrationsniveau?

Et udviklingsmiljø for modeludvikling er nødvendigt, så længe selve princippet bag modellen er under testning. Derfor beskrives MetalStat i denne rapport ud fra en prototype, der er sat op i et udviklingsmiljø, hvor kodningen ikke er rettet mod en egentlig drift, men til testning af principper og funktionalitet.

8.3 Forslag til implementering af MetalStat

Brugen af modeller som MetalStat i samspil med en overvågning af MFS er vigtig at se som et kompetenceløft, da modeller som MetalStat sætter nye standarder for, hvad der kan konkluderes på baggrund af overvågningsdata, og dermed også krav til kompetence hos den bruger, som skal anvende modellerne. Det betyder, at der bør laves en faglig struktur bag modellerne, som behandler resultater, og som sørger for drift og forbedring af modellerne. En model bør i denne sammenhæng ikke opfattes som en fast færdigudviklet størrelse, men som en del af en læringsproces, hvor ny viden løbende vil indgå i et stadigt bedre vidensgrundlag, og hvor modellen løbende udvikles til bedst muligt at understøtte nye aktiviteter, der kan forbedre vidensgrundlaget. Den faglige struktur omkring modellering kunne være bygget op i fire faglige grupper:

- *En brugergruppe*, der dels forestår formidling af modellerne til alle relevante i Miljøstyrelsen og dels definerer de specifikationer, som modellerne skal indfri. Opgaven er at implementere brugen af modellerne i sagsbehandling på den mest optimale måde.
- *En miljøfaglig gruppe*, der evt. årligt diskuterer modellernes resultater i forhold til, om de virker troværdige, og i forhold til, om de kan bidrage med ny viden. Desuden skal denne gruppe sikre, at modellerne bliver opdateret med ny viden omkring de processer, som modellerne beskriver. Modellernes resultater vil løbende blive evalueret i forhold til relevant faglig viden som en del af en kvalitetssikring. Dette er nødvendigt, da modellerne er så komplekse, at de ikke bare reproducerer simple kontrollerbare sammenhænge, men også mere komplekse mønstre. Hvis modellerne f.eks. tillægger en attribut stor betydning for koncentrationsniveauet af et metal, bør denne betydning kun blive taget til indtægt som reel, hvis den også kan begrundes med miljøfaglig viden eller viden om mulige kilder. Den miljøfaglige gruppe skal således være mellemlid mellem modellernes output og brugergruppen samt løbende være i dialog med udviklingsgruppen. Den miljøfaglige gruppe kan evt. udgive en tilstandsrapport, der konkluderer på baggrund af modelsimuleringer med brug af opdaterede overvågningsdata.
- *En driftsgruppe*, der forestår den regnetekniske kobling til datakilder samt dokumentation af modellernes beregninger og sikrer, at modellen indfrier de definerede specifikationer. Da modellerne er bygget op om en løbende iteration, vil de hele tiden være regneaktive. Så der kræves en konstant adgang til en CPU. Princippet er, at computeren konstant vil gennemregne modellerne, og hvis der ønskes et resultat fra modellerne, hentes de sidste få tusinde gennemregninger, der vil afspejle statistikken omkring målingerne samt usikkerheden på estimerede parametre. De udtrukne gennemregninger vil kunne behandles med deskriptiv statistik som f.eks. histogrammer eller fraktilværdier vist for hvert ID 15 opland i GIS. Disse udtræk specificeres af brugergruppen. Når nye data tilgår modellerne, f.eks. i form af et nyt års overvågningsresultater, vil deres variable justere sig på

plads svarende til en tilstand af ny viden fra de nye data. Driftsgruppen vil tjekke, om modellerne er konvergeret, før der igen kan udtages resultater fra dem. Typisk vil den periode, der skal bruges at konvergere, være få uger, da der kun vil fremkomme mindre justeringer i modellernes samlede opsætning. Løbende monitoring og optimering af modellernes X matricer som disse understøtter oplandsbeskrivelsen på bedst mulig måde.

- *En udviklingsgruppe*, der kombinerer miljøkemisk faglig viden med matematisk og regneteknisk viden i samråd med den miljøfaglige gruppe og kommer med projektforslag til mulige udviklingsprojekter. Gruppen evaluerer og overvåger modellernes resultater og laver forslag til forbedringer og justeringer i forhold til nye datakilder og fordringer i modellernes matematiske opbygning. Denne gruppe kan også vurdere mulighederne for at udbrede modelleringen til nye stoffer og problemstillinger.

8.4 Evaluering af projektet med henblik på opmærksomheds-og læringspunkter til eventuelle kommende modelprojekter

8.4.1 MetalStat specifikt

Da MetalStat er en multivariat model mht. metaller, vil viden om et metals koncentrationsniveau implicit kunne bidrage med viden om et andet metals koncentrationsniveau på samme sted og til samme tid. På samme måde kan en hjælpevariabel bruges, hvis der er en sammenhæng mellem f.eks. litiumindholdet og et eller flere metaller, og hvis litiumkoncentrationen er målt, så vil det kunne bidrage med viden om metallerne, hvis litium eller en anden hjælpevariabel inddrages i observationsmodellen på linje med metallerne. Det bør overvejes, om sådanne hjælpevariabler skal med i den første driftversion af MetalStat.

X matricen bør evalueres grundigt med henblik på at optimere oplandsmodellen, især i forhold til de stærkt ulineære uafhængige variable som f.eks. lerindhold, pH. Valget af ID15 oplande som reference er gjort som et kompromis mellem ønsket om en landsdækkende model, der er realistisk at anvende i forhold til tilgængelige data og regneteknisk på den ene side, og ønsket om at kunne beskrive lokale forhold troværdigt på den anden side. Dette kan være en udfordring over for stærkt ulineære variable, der kan variere stærkt inden for et ID15 opland og derfor være dårligt beskrevet gennem en enkelt X værdi for hvert IDF 15 opland. En sådan udfordring kan dog ofte løses ved at opstille en model i GIS, der integrerer en troværdig middelværdi indenfor ID15 oplandet. Denne mulighed bør undersøges især for pH i en evt. driftversion af MetalStat. Det synes muligt at eliminere nogle af de valgte uafhængige variable til et mindre antal, og i den forbindelse bør det overvejes, om andre skal testes, såsom areal i oplandet med industri, by og ferskvand. Det bør undersøges, om det er muligt at fastholde β til kun at være positiv, da det vil modvirke estimering af negative koncentrationer og dannelsen af fiktive modsatrettede +/- effekter fra forskellige uafhængige variable. Datagrundlaget bør løbende opdateres med de nyeste NOVANA-data, efterhånden som disse bliver gjort tilgængelige. Før en sådan opdatering af data foretages, anbefales det at bruge de nyeste data til validering af MetalStat, så man løbende kan følge med i, hvordan modellen forbedrer sine prædiktive evner.

MetalStat bruger estimeret tilførsel af grundvand (X_{grundv}) som attribut for mulig grundvandsbåret tilførsel af metal i interaktion mellem hhv. ler, pH og organisk stof i jordmatrice, men det er en meget simpel beskrivelse, da koncentrationsniveauet i grundvandet naturligvis vil variere. Alene for bly, cadmium og nikkel kan det ses, at koncentrationsniveauet er relativt højt for grundvand sammenlignet med niveauet i vandløb. En bedre beskrivelse kunne være et estimat af grundvandskoncentrationen (C_{grundv}), hvor massen af metal tilført er givet som $\alpha \cdot C_{grundv} \cdot X_{grundv}$, hvor α udtrykker en evt. tilbageholdelse, når grundvandet strømmer til vandløbet. For hvert ID 15 opland kan C_{grundv} ses som en latent variabel, der kan estimeres med brug af de grundvandsprøver, der er indrapporteret for det pågældende opland, på samme måde som MetalStat betragter vandkoncentrationen i vandløbet som en latent variabel, der er målt i NOVANA-programmet. På denne måde kan grundvandsovervågningen integreres direkte ind i vandløbsovervågningen.

På samme måde som beskrevet for grundvand kan andre matricer, som f.eks. sedimentbundet metal, kobles til MetalStat gennem en ny latent variabel for sedimentkoncentrationen, hvis der kan findes en troværdig funktion, der kobler adsorberet fase til opløst fase, evt. baseret på ligevægtskemi eller første ordens udveksling mellem sediment og vandfase.

8.4.2 Modeller for MFS generelt

Da udfordringen med modellering for metaller og miljøfarlige forurenende stoffer i det hele taget er betydelig, bør der løbende arbejdes på at forbedre modelleringen i tæt synergi med det datagrundlag, som kan understøtte modellerne i løbende samspil med overvågningens dataindsamling.

Det er en fordel, at modeller for MFS bygges op efter statistiske principper, hvis deres formål er at estimere koncentrationsniveauer og den usikkerhed, der forbundet med estimerne. Da koncentrationsniveauet for mange MFS opfører sig log-normalfordelt, vil de principper, der anvendes i MetalStat, have potentiale til at kunne bruges til mange andre stoffer.

Resultaterne fra MetalStat viser også, at en model ikke kun kan bidrage med prædiktioner for ikke-målte områder, men også forbedre vidensgrundlaget for de målte områder på en måde, der ikke lader sig afsløre ved mere deskriptive metoder. En øget brug af modeller betyder derfor, at den viden, som kan erhverves fra brug af overvågningsdata, udvides og skifter karakter.

En evt. implementering af MetalStat kan danne rettesnor for, hvordan andre modeller for andre MFS skal implementeres, da en del af udfordringerne vil være sammenfaldende. Oplandsmodellen, som er udviklet til MetalStat ud fra især en oplandsmodel for næringsstoffer suppleret med forskellige andre uafhængige variable, er relevant til at teste over for andre MFS. Således er byspildevand og regnvandsbetingede udløb beskrevet, hvilket er relevant for f.eks. PAH'r, mens oplandsmodellens uafhængige variable for landbrugsdrift kunne være relevant for f.eks. pesticider.

ID 15 oplandene kan via GIS kobles til ønskede arealmæssige klassificeringer, såsom overfladevandområder og vandløbsområder

9 Referencer

Adhikari, K., Kheir, R.B., Greve, M.B., Bocher, P.K., Malone, B.P., Minasny, B., McBratney, A.B. and Greve, M.H. 2013. High-Resolution 3-D Mapping of Soil Texture in Denmark. *Soil Sci. Soc. Am. J.* 77, 860-876. doi: DOI 10.2136/sssaj2012.0275

Adhikari, K., Kheir, R.B., Greve, M.B., Bocher, P.K., Malone, B.P., Minasny, B., McBratney, A.B. and Greve, M.H. 2014. Mapping soil pH and bulk density at multiple soil depths in Denmark. *GlobalSoilMap. GlobalSoilMap: Basis of the global spatial soil information system.* ed. / Dominique Arrouays; Neil McKenzie; Jon Hempel; Anne Richers de Forges; Alex McBratney. CRC Press, 2014. p. 155-160.

Andersen, H. E. & Heckrath, G. (redaktører). 2020. Fosforkortlægning af dyrkningsjord og vandområder i Danmark. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 340 s. - Videnskabelig rapport nr. 397 <http://dce2.au.dk/pub/SR397.pdf>

Bak, J.L., Jensen, J. & Larsen, M.M. 2015. Belysning af kobber- og zinkindholdet i jord. Indhold og udvikling i kvadratnettet og måling på udvalgte brugstyper. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 72 s. - Videnskabelig rapport fra DCE - Nationalt Center for Miljø og Energi nr. 159 <http://dce2.au.dk/pub/SR159.pdf>

DHI, 2020. Kvantificering af tilførsel af miljøfarlige forurenende stoffer fra diffuse kilder til vandmiljøet. Rapport for Miljøstyrelsen, september 2020.

Ellermann, T., Bossi, R., Nygaard, J., Christensen, J., Løfstrøm, P., Monies, C., Geels, C., Nilesen, I. E., & Poulsen, M. B., 2019: Atmosfærisk deposition 2018. NOVANA. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi. 84s. – Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi nr. 351. <http://dce2.au.dk/pub/SR351.pdf>

Ellermann, T., Bossi, R., Nygaard, J., Christensen, J., Løfstrøm, P., Monies, C., Geels, C., Nilesen, I. E., & Poulsen, M. B., 2021: Atmosfærisk deposition 2019. NOVANA. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi. 90s. – Videnskabelig rapport fra DCE – Nationalt Center for Miljø og Energi nr. 415. <http://dce2.au.dk/pub/SR415.pdf>

Kotlar, A., de Jong van Lier, Q., Andersen, H.E., Nørgaard, T., Iversen, B.V. 2020. Quantification of macropore flow in Danish soils using near-saturated hydraulic properties. *Geoderma*, 375, 114479.

Larsen, S.E., Windolf, J. og Kronvang, B. 2015. Validering af fosformodellen. Notat fra DCE, AU, 16 s

Levin, G. 2019. Basemap03. Technical documentation of the method for elaboration of a land-use and land-cover map for Denmark. Aarhus University, DCE – Danish Centre for Environment and Energy, 86 pp. Technical Report No. 159

Mortensen, M.H., Ernstsén, V., Voutchkova, D & Thorling, L. 2021. Dokumentationsrapport. Udvikling af metode til vurdering af grundvandsforekomsters kemiske tilstand for udvalgte uorganiske sporstoffer og salte. Danmarks og Grønlands Geologiske Undersøgelse Rapport 2021 / 19.

Motarjemi, S., A.B. Møller, F. Plauborg, and B.V. Iversen. 2019. Predicting tile drainage discharge using machine learning algorithms (in prep.).

Møller, A.B., A. Beucher, B. V. Iversen, and M.H. Greve. 2018. Predicting artificially drained areas by means of a selective model ensemble. *Geoderma* J320: 30–42. <https://www.sciencedirect.com/science/article/pii/S0016706117318116?via=ihl> (verified 19 June 2018).

Nilsson, B., Kronvang, B., van't Veen, S., Troldborg, L., Thorling, L., Boutrup, S., Mørk Larsen, M., Rasmussen, J., Hinsby, K. & Kazmierczak, J., 2019. Vurdering af grundvandets kemiske påvirkning på vandløb og kystvande. Vol 1(2): Hovedrapport. Danmarks og Grønlands Geologiske Undersøgelse Rapport 2019 / 2.

Sørensen, P.B., Andersen, H.E., Fauser, P., Bjerg, P.L., Bro, R., Holm, P.E. & Abrahamsen 2021 Muligheder for modellering af miljøfarlige forurenende stoffer i overfladevand. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 60 s. - Videnskabelig rapport nr. 414. <http://dce2.au.dk/pub/SR414.pdf>

Sørensen, P.B. & Johansson, L.S. 2020. Vurdering af kampagnemåling af miljøfarlige forurenende stoffer i vandløb og søer. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 46 s. - Teknisk rapport nr. 167. <http://dce2.au.dk/pub/TR167.pdf>

Thodsen, H., Windolf, J., Rasmussen, J., Bøgestrand, J., Larsen, S.E., Tornbjerg, H., Ovesen, N.B., Kjeldgaard, A. & Wiberg-Larsen, P. (2016): Vandløb 2015. NOVANA. Aarhus Universitet, DCE – GEUS 125 Nationalt Center for Miljø og Energi, 68 s. - Videnskabelig rapport fra DCE - Nationalt Center for Miljø og Energi nr. 206

Thorling, L., Brusch, W., Ernstsén, V., Hansen, B., Laier, T., Larsen, F., B., Mielby, S og Sørensen, B. L. (2015): Grundvand. Status og udvikling 1989 – 2013. Teknisk rapport, GEUS 2015.

Thorling, L., Hansen, B., Larsen, C.L., Larsen, F., Mielby, S., Johnsen, A.R., & Troldborg, L. (2016): Grundvand. Status og udvikling 1989 – 2015. Teknisk rapport, GEUS 2016.

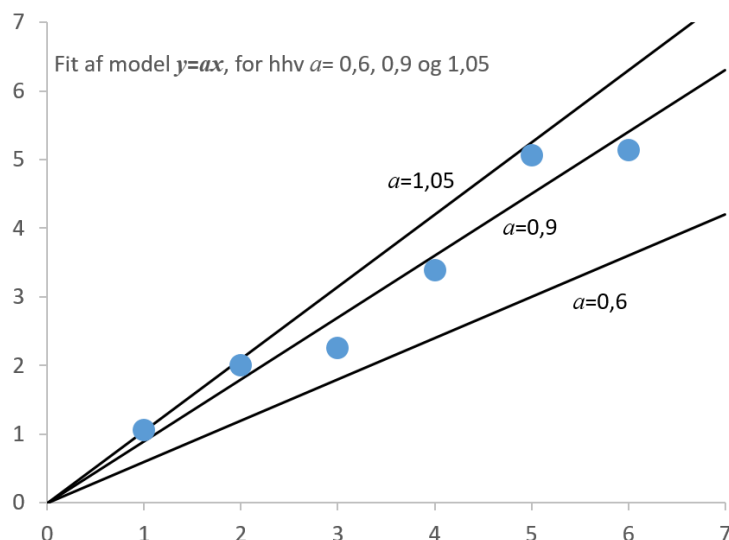
Troldborg, L., Tornbjerg, H., Blicher-Mathiesen, G., Kronvang, B., Thodsen, H., Ernstsén, V., 2015, National kvælstofmodel. Oplandsmodel til belastning og virkemidler. Metode rapport. De Nationale Geologiske Undersøgelser for Danmark og Grønland, 2015.

Bilag A, Statistiske grundbegreber

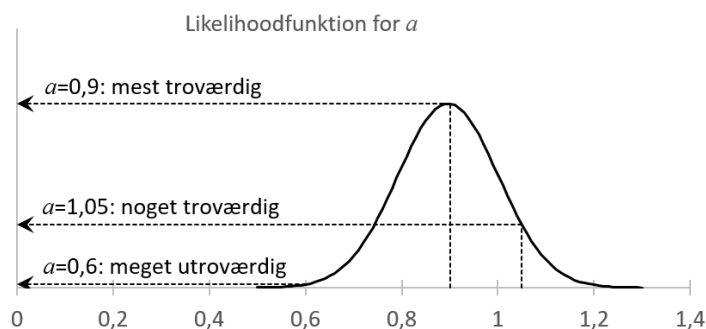
Boks 1: Kobling af model til målinger

En statistisk model kan bestemmes gennem en likelihood-funktion, der beregner, hvor troværdigt modellen kan forudsige de målte værdier. Følgende eksempel illustrerer princippet.

En simpel model $y=ax$ vælges til at beskrive seks målte sammenhænge mellem x og y , hvor værdien af parameteren a er ukendt. Neden for ses tre forskellige a værdier med hver sin linje:



Intuitivt vurderes $a=0,9$ til at være mest troværdig, mens $a=0,6$ er mindst troværdig, men en likelihood-funktion kan kvantificere graden af troværdighed. Hvis målingerne forudsættes at variere over og under linjen bestemt med a efter en standard normalfordeling ($N(0,1)$), så kan troværdigheden af en vilkårlig a værdi bestemmes:



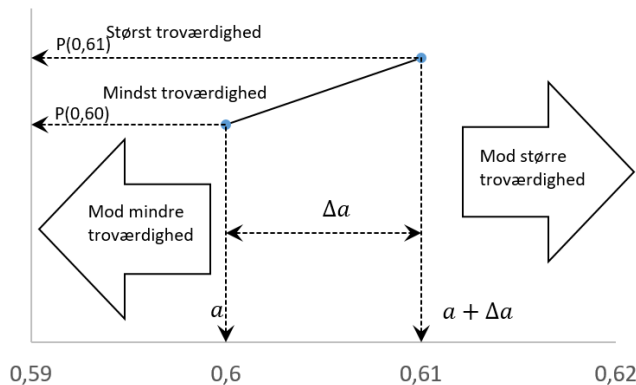
Nu kan troværdigheden (likelihood) beregnes for hver af de tre a værdier, og det ses, at $a=0,6$ er meget utroværdig og kan lades helt ude af betragtning, $a=1,05$ har en vis troværdighed, mens $a=0,9$ er mest troværdig (maksimum likelihood estimatet). Likelihood-funktionen indikerer altså usikkerhedsintervallet for a som en naturlig del af at bestemme a , og det er netop de statistiske modelleres styrke, at de på den måde bestemmer modelparameterens usikkerhed. Hvis man øger antallet af de seks målpunkter, vil likelihood-funktionen typisk blive indsnævret usikkerheds(konfidens)-intervallet for a .

I realiteten er modellen for metalkoncentrationer langt mere kompleks med flere hundrede parametre og flere metaller på en gang, men princippet med at bruge en likelihood-funktion til at finde fordelingen af en parameter følger med dette simple eksempel.

Boks 3: Metropolis-Hastings algoritmen

Princip. Metropolis-Hastings algoritmen er en metode til at estimere parametrene i en model. Modellen anvender den posteriori sandsynlighedsfordeling (Boks 2) for modellens parametre i en Metropolis-Hasting algoritme til at bestemme parameterens fordeling (forventet værdi og spredning rundt om forventningen). Grundprincippet illustreres i nedenstående eksempel med afsæt i Boks 1. Troværdige værdier for parameteren a er ukendte, og i komplekse modeller kan de ikke beregnes direkte ud fra likelihood-funktionen, fordi det er kun muligt at beregne en troværdighed, når a værdien først er valgt. Dette løses ved at "søge" efter troværdige a værdier med følgende metodik:

Da vi ikke kender de mest troværdige værdier for a , sætter vi bare værdien til $a=0,6$ ud fra rent gætværk, og vi anvender den posteriori sandsynlighedsfordeling, her illustreret med likelihood-funktion i Boks 1, til at beregne troværdigheden for denne værdi ($P(0,60)$). Derefter trækker vi en ny værdi fra en kandidat fordeling, f.eks. 0,61, og beregner troværdigheden for 0,61 ($P(=,61)$), så nu har vi to beregninger af troværdighed som illustreret i figuren:



Sandsynligheden for at vælge $a=0,61$ sættes til $P(0,61)/(P(0,61)+P(0,6))$, og ellers beholdes værdien $a=0,60$. Man sætter computeren til at lave et tilfældigt valg med brug af disse sandsynligheder. Denne algoritme vil med størst sandsynlighed vælge den mest troværdige værdi ($a=,61$), da $P(0,61)>P(0,60)$, så hvis algoritmen gennemløbes mange gange, vil der være en tendens til at finde de mest troværdige værdier ved gå "op ad bakke" på likelihood-funktionen, men nogle gange vil algoritmen også gå "ned af bakke", hvilket kan virke dumt, men det er smart, da algoritmen derved kortlægger usikkerheden på a , svarende til hele likelihood-funktionen i Boks 1.

Behov for at Konvergere (engelsk: Burn in) I eksemplet oven for er værdien $a=0,60$ brugt i Metropolis-Hastings algoritmen, og ifølge likelihood-funktionen i Boks 1 svarer det til "meget utroværdigt", hvilket derfor ikke virker optimalt, det er dog et vilkår, når beregningerne startes med algoritmen, at de mest troværdige intervaller for parametrene ikke kendes. Disse intervaller skal algoritmen selv søge sig frem til. Derfor starter en Metropolis-Hastings algoritme altid med at konvergere, hvor der søges efter de troværdige intervaller for parametrene, hvilket på engelsk typisk benævnes Burn in. Så længe modellen ikke er færdig konvergeret, skal modellens resultater ikke betragtes som rigtige.

Bilag B. Oversigt over data aggregeret til ID 15-oplande

Id	Beskrivelse	Tema	Periode	Enhed	Reference
1	Oplandsstørrelse			hektar	
2	Gns. lerindhold i opland (data for 0-30 cm)	30,4 m raster		%	Adhikari et al. (2013)
3	Gns. lerindhold i opland (data for 100-200 cm)	30,4 m raster		%	Adhikari et al. (2013)
4	Gns. kulstofindhold i opland (data for 0-30 cm)	30,4 m raster		%	Adhikari et al. (2013)
5	Gns. kulstofindhold i opland (data for 100-200 cm)	30,4 m raster		%	Adhikari et al. (2013)
6	Gns. pH i opland (data for 0-30 cm)	30,4 m raster		Antal protoner	Adhikari et al. (2014)
7	Gns. pH i opland (data for 100-200 cm)	30,4 m raster		Antal protoner	Adhikari et al. (2013)
8	Landbrug opdelt på intensivt og ekstensivt	Basemap	2011, 2016, 2018	hektar	Levin (2019)
9	Kvælstof tilført med svinøgødnings (alt. DE svin)			kg N (DE)	Bak et al. (2015)
10	Drænet areal	Drained_v3		hektar	Møller et al. (2018)
11	Vandføring, dræn, oplands-gns. (kun for drænedde arealer)	250 m raster	1990-2017	mm/år	Motarjemi et al. (2019)
12	Vandføring via makroporer, dræn, oplands-gns. (kun hvor makroporer)	250 m raster	1990-2017	mm/år	Motarjemi et al. (2019)
13	Areal med makroporer i høj risikoklasse (4 + 5)	250 m raster		hektar	Kotlar et al. (2020)
14	Jordtab fra mark til vandløb ved erosion	10 m raster		ton	Andersen & Heckrath (2020)
15	Vandføring, vandløb		Gns. 2010-2019	1000 m³/år	Overfladevandsdatabasen ODA https://odaforalle.au.dk
16	Nedbør		Gns. 2010-2019	mm/år	DMI
17	Nettonedbør		Gns. 2010-2019	mm/år	Egne beregninger på grundlag af DMI's data
18	Vanding		Gns. 2010-2019	mm/år	Jupiter databasen (https://www.geus.dk/produkter-tydelser-og-faciliteter/data-og-kort/national-boringsdatabase-jupiter)
19	Grundvands-andel af samlet vandføring	BFI		fraktion	Larsen et al. (2015)
20	Regnvandsbetingede udløb, vandmængde		Gns. 2010-2019	1000 m³/år	Miljøstyrelsen - Fagdatacenter for punktkilder
21	Renseanlæg, vandmængder		Gns. 2010-2019	1000 m³/år	Miljøstyrelsen - Fagdatacenter for punktkilder
22	Renseanlæg, tungmetaller		Gns. 2010-2019	kg/år	Miljøstyrelsen - Fagdatacenter for punktkilder

Bilag C, parameterestimering

2,5 %, 50 %, 97,5%: Fraktilværdier

P (+): Estimeret sandsynlighed for at parameterværdien er positiv

Parameter	metal	2.5%	50%	97.50%	p (+)
Intercept (winter)	Bly	-2,68967	-2,68485	-2,67786	0
Spring	Bly	0,007439	0,014042	0,019877	1
Sommer	Bly	0,001853	0,013469	0,022822	1
Autumn	Bly	-0,01182	-0,00349	0,004259	0,201835
Beta0	Bly	-0,00997	-0,00551	-0,00253	0
<i>X_{land}</i>	Bly	-0,03383	-0,0246	-0,0198	0
Xregn	Bly	-0,00059	0,006609	0,010518	0,93578
Xrens	Bly	0,045401	0,049755	0,053813	1
XHa	Bly	0,016254	0,021766	0,024986	1
XHb	Bly	0,041857	0,044872	0,057679	1
Xlera	Bly	0,008931	0,010592	0,011663	1
Xlerb	Bly	0,00026	0,001779	0,003764	1
Xca	Bly	0,001892	0,006443	0,015862	1
XCb	Bly	-0,02104	-0,0167	-0,00808	0
Xgrundv	Bly	0,042709	0,04798	0,054741	1
Xdræn	Bly	0,006661	0,017734	0,022184	1
Xmakro	Bly	-0,0001	0,003745	0,01028	0,963303
Xned	Bly	-0,01077	-0,00563	-0,00046	0,018349
Xsvin	Bly	-0,06578	-0,06002	-0,05159	0
XHb:Xgrundv	Bly	-0,01203	-0,00668	-0,00199	0,009174
Xlera:X1land	Bly	-0,00378	-0,0036	-0,00316	0
Xlera:Xsvin	Bly	0,001487	0,001698	0,001998	1
Xlerb:Xgrundv	Bly	0,005165	0,008268	0,010889	1
XCa:Xland	Bly	-0,00971	-0,00803	-0,00489	0
XCa:Xsvin	Bly	0,005707	0,006936	0,008591	1
XCb:Xgrundv	Bly	0,038384	0,050349	0,056389	1
Xdræn:Xland	Bly	-0,00233	0,001392	0,011381	0,651376
Xdræn:Xsvin	Bly	-0,04898	-0,03777	-0,02087	0
Xmakro:Xland	Bly	0,040689	0,045526	0,049095	1
Xmakro:Xsvin	Bly	0,012473	0,017646	0,021382	1
Xned:Xland	Bly	0,016344	0,019397	0,025122	1
Xned:Xsvin	Bly	-0,01721	-0,01407	-0,00617	0
gamma	Bly	-0,09635	-0,08894	-0,08393	0
sd random site	Bly	0,001	0,001	0,001	1
sd random year	Bly	0,001	0,001	0,001	1
sigma structural	Bly-Bly	0,742929	0,787862	0,83547	1
sigma measure	Bly-Bly	0,637417	0,637417	0,637417	1
sigma structural	Bly-Cadmium	0,276948	0,313543	0,380771	1
sigma measure	Bly-Cadmium	0,123276	0,123276	0,123276	1
sigma structural	Bly-Nikkel	0,19818	0,216426	0,263967	1
sigma measure	Bly-Nikkel	0,038036	0,038036	0,038036	1

Parameter	metal	2.5%	50%	97.50%	p (+)
Intercept (winter)	Cadmium	-4,17341	-4,16505	-4,14941	0
Spring	Cadmium	0,005237	0,013473	0,017261	1
Sommer	Cadmium	-0,04811	-0,04373	-0,03692	0
Autumn	Cadmium	0,024582	0,0382	0,043914	1
bo	Cadmium	0,038812	0,051944	0,056226	1
Xland	Cadmium	0,03356	0,038205	0,045918	1
Xregn	Cadmium	-0,02088	-0,01704	-0,0086	0
Xrens	Cadmium	-0,00108	0,011334	0,015561	0,963303
XHa	Cadmium	0,008516	0,012517	0,021089	1
XHb	Cadmium	-0,02055	-0,01659	-0,01026	0
Xlera	Cadmium	0,000553	0,001559	0,002881	1
Xlerb	Cadmium	-0,00229	-0,00111	0,000182	0,027523
Xca	Cadmium	0,001958	0,009046	0,014816	1
XCb	Cadmium	-0,04001	-0,03046	-0,02093	0
Xgrundv	Cadmium	0,059353	0,064382	0,0695	1
Xdræn	Cadmium	-0,00703	-0,00101	0,004296	0,302752
Xmakro	Cadmium	0,003195	0,008691	0,017887	1
Xned	Cadmium	0,024623	0,029531	0,032712	1
Xsvin	Cadmium	0,05188	0,060826	0,065548	1
XHb:Xgrundv	Cadmium	0,073746	0,078247	0,082317	1
Xlera:X1land	Cadmium	-0,00323	-0,00321	-0,00222	0
Xlera:Xsvin	Cadmium	-0,00042	-0,00021	0,00035	0,302752
Xlerb:Xgrundv	Cadmium	0,009276	0,011242	0,014672	1
XCa:Xland	Cadmium	0,022158	0,027288	0,029439	1
XCa:Xsvin	Cadmium	-0,00867	-0,00758	-0,00569	0
XCb:Xgrundv	Cadmium	-0,08281	-0,07176	-0,05317	0
Xdræn:Xland	Cadmium	-0,02885	-0,02067	-0,0162	0
Xdræn:Xsvin	Cadmium	-0,04005	-0,03027	-0,01874	0
Xmakro:Xland	Cadmium	-0,07586	-0,07253	-0,06493	0
Xmakro:Xsvin	Cadmium	-0,02039	-0,01272	-0,00382	0
Xned:Xland	Cadmium	0,003216	0,005822	0,01027	1
Xned:Xsvin	Cadmium	0,055865	0,064589	0,071112	1
gamma	Cadmium	-0,07192	-0,06536	-0,05611	0
sd random site	Cadmium	0,001	0,001	0,001	1
sd random year	Cadmium	0,001	0,001	0,001	1
sigma structural	Cadmium-Bly	0,276948	0,313543	0,380771	1
sigma measure	Cadmium-Bly	0,123276	0,123276	0,123276	1
sigma structural	Cadmium-Cadmium	0,799806	0,846599	0,955161	1
sigma measure	Cadmium-Cadmium	0,301862	0,301862	0,301862	1
sigma structural	Cadmium-Nikkel	0,335955	0,369047	0,465479	1
sigma measure	Cadmium-Nikkel	0,03575	0,03575	0,03575	1

Parameter	metal	2.5%	50%	97.50%	p(+)
Intercept (winter)	Nikkel	0,451379	0,458369	0,465386	1
Spring	Nikkel	0,008572	0,020443	0,025032	1
Sommer	Nikkel	0,052566	0,062725	0,06807	1
Autumn	Nikkel	0,000472	0,006272	0,015358	0,981651
bo	Nikkel	0,060483	0,066163	0,076111	1
Xland	Nikkel	0,02747	0,034099	0,044013	1
Xregn	Nikkel	-0,02526	-0,01861	-0,00696	0
Xrens	Nikkel	0,035294	0,040161	0,04515	1
XHa	Nikkel	-0,00365	-0,00054	0,000803	0,293578
XHb	Nikkel	-0,02058	-0,01445	-0,00863	0
Xlera	Nikkel	0,008127	0,011452	0,013484	1
Xlerb	Nikkel	-0,00386	-0,00241	-0,00059	0
Xca	Nikkel	0,015887	0,01959	0,025803	1
XCb	Nikkel	-0,0428	-0,03823	-0,03455	0
Xgrundv	Nikkel	-0,02817	-0,02589	-0,02072	0
Xdræn	Nikkel	0,006513	0,013312	0,018097	1
Xmakro	Nikkel	-0,00639	-0,00291	0,003194	0,137615
Xned	Nikkel	0,031737	0,034523	0,042988	1
Xsvin	Nikkel	-0,06918	-0,06306	-0,05651	0
XHb:Xgrundv	Nikkel	0,043929	0,049261	0,053419	1
Xlera:X1land	Nikkel	-0,00493	-0,00445	-0,00406	0
Xlera:Xsvin	Nikkel	0,000775	0,000903	0,001205	1
Xlerb:Xgrundv	Nikkel	0,014124	0,016075	0,022149	1
XCa:Xland	Nikkel	-0,00123	0,001912	0,00502	0,743119
XCa:Xsvin	Nikkel	0,001261	0,003038	0,004211	1
XCb:Xgrundv	Nikkel	-0,04044	-0,03554	-0,02886	0
Xdræn:Xland	Nikkel	-0,03417	-0,02669	-0,01737	0
Xdræn:Xsvin	Nikkel	-0,04732	-0,03406	-0,02698	0
Xmakro:Xland	Nikkel	0,030305	0,036878	0,041369	1
Xmakro:Xsvin	Nikkel	0,028242	0,033924	0,039119	1
Xned:Xland	Nikkel	0,07863	0,085279	0,089373	1
Xned:Xsvin	Nikkel	0,049521	0,059498	0,063192	1
gamma	Nikkel	-0,15639	-0,15051	-0,14652	0
sd random site	Nikkel	0,001	0,001	0,001	1
sd random year	Nikkel	0,001	0,001	0,001	1
sigma structural	Nikkel-Bly	0,19818	0,216426	0,263967	1
sigma measure	Nikkel-Bly	0,038036	0,038036	0,038036	1
sigma structural	Nikkel-Cadmium	0,335955	0,369047	0,465479	1
sigma measure	Nikkel-Cadmium	0,03575	0,03575	0,03575	1
sigma structural	Nikkel-Nikkel	0,398547	0,420885	0,504663	1
sigma measure	Nikkel-Nikkel	0,140304	0,140304	0,140304	1

UDVIKLING AF MODEL TIL UNDERSTØTTELSE AF VANDPLANLÆGNINGEN I FASTLÆGGELSEN AF KONCENTRATIONSNIVEAUER AF METALLER I DANSKE VANDLØB

Dette projekt har til formål at udvikle en pilot-model til prædiktion af forekomst af metaller i vandløb og har karakter af et pilotprojekt ved at udvikle en model, som kan afdække anvendeligheden af forskellige beskrivende data generelt til brug også for eventuelle fremtidige modeleringsprojekter for andre stofgrupper end metaller. Det bemærkes, at projektet er et første forsøg på at anvende NOVANA- data i kombination med beskrivende data til at opstille en landsdækkende model til estimering af miljøfarlige forurenende stoffkoncentrationer i vandløbsvand.