

Model for søbelastning af kvælstof og fosfor

Fagligt notat fra DCE – Nationalt Center for Miljø og Energi

Dato: 28. Februar 2025 | 18



AARHUS
UNIVERSITET

DCE – NATIONALT CENTER FOR MILJØ OG ENERGI

Datablad

Fagligt notat fra DCE – Nationalt Center for Miljø og Energi

Kategori: Rådgivningsnotat

Titel: Model for søbelasting af kvælstof og fosfor

Forfatter(e): Peter Borgen Sørensen
Institution(er): Institut for Ecoscience

Faglig kommentering:
Kvalitetssikring, DCE:
Sproglig kvalitetssikring:

Hans Thodsen
Maj-Britt Bjergager
Charlotte Elisabeth Kler

Ekstern kommentering: Miljøstyrelsen. [Kommentarerne findes her.](#)

Rekvirent: Miljøstyrelsen, Styrelsen for Grøn Arealomlægning og Vandmiljø

Bedes citeret: Sørensen, B.P., 2025. Model for søbelasting af kvælstof og fosfor. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 24 s. – Fagligt notat nr. 2025|18

Gengivelse tilladt med tydelig kildeangivelse

Foto forside: Udløbet af Skals Å fra Fussing Sø (foto Peter Borgen Sørensen)

Sideantal (eksklusiv bilag): 24

Indhold

1	Indledning	4
2	Metode	6
2.1	Belastning med N og P	6
2.2	Belastning af vand (vandmodel)	7
2.3	Koncentration af P og N (stofmodel)	10
3	Illustrative beregninger	12
3.1	Hydraulisk responstid for 10 danske søer	12
3.2	Bryrup Langsø	13
4	Diskussion	20
5	Referencer	23
6	Bilag	24

1 Indledning

Som baggrund for belastningsberegningerne for intensivt målte sør giver nærværende notat en beskrivelse af en beregningsmodel for massebalancen for fosfor og kvælstof i de 10 mest undersøgte sører, som er afrapporteret i Johansson et. Al., (2024) (kap. 2.3). Den beskrevne metode er sammenholdt med en historisk metode (Windolf et al., 1996) i Jeppesen et al., (2025). Beregningsmodellen kan håndtere manglende viden om tilført masse af kvælstof (N) og fosfor (P) til sørerne. At denne viden mangler, skyldes dels, at der ikke er målt på alle vandløb ved indløb til en given sør, dels at sør kan have udveksling med grundvand gennem sør bunden, hvilket ikke kan måles med standardudstyr.

Selv opgørelsesmetoden følger beskrivelsen i Sørensen and Nielsen (2023). Formålet med metoden er at anvende målte vandvoluminer samt den totale koncentration af kvælstof og fosfor i både indløb og udløb fra sør til at beregne sørernes tilførsel og afgivelse af N og P. I de fleste tilfælde har en given sør et enkelt udløb, hvor der månedligt er målt vandvolumen og koncentration af N og P. Målingerne af vandvolumen og koncentration af N og P i indløbet i til sør er derimod normal ufuldkommen. Dette skyldes, at indløbet til sør oftest består af flere vandløb med et dominerende opmålt vandløb og en række mindre tilløb, hvor der ikke er foretaget målinger. Derudover kan umålt vand løbe til sør som tilstrømning af grundvand gennem sørens bund. Der kan også ske udsivning gennem bunden og dermed en skjult udledning af N og P, der skal lægges til udledningen via udløbet for at bestemme sørens samlede massebalance af N og P. Endelig er det muligt, at der i den samme sør både kan være ind- og udsivning gennem sør bunden. Pga. disse kombinationsmuligheder vil en massebalance for en sør altid være en tilnærmelse med en vis usikkerhed. Modellens udfordring er at minimere disse usikkerheder vel vidende, at de ikke kan elimineres helt.

Opgørelsesmetoden får input fra to modeller: (1) en vandmodel i form af volumenbalance af vand, der beregner det umålte volumen af vand, og som adskilles i to vandvolumener svarende til en grundvandsfødt tilstrømning og en overfladenær tilstrømning, og (2) en stofmodel, der beregner N- og P-koncentrationsniveauer for hver af de to ovenover beskrevne vandvolumener samt et bidrag fra punktkilder placeret i umålt opland. Disse to modeller kan beregne den umålte belastning af N og P i sør og identificere en evt. udledning af N og P gennem sør bunden. Modellerne forudsætter, at N- og P-koncentrationsniveauerne i det umålte vand kan beregnes ved brug af de koncentrationerne i det målte vand, idet modellernes parametre (modelkoefficienterne) bestemmes på baggrund af den målte andel af det tilstrømmende vand og målt N og P i indløb. Dette er en tilnærmelse og mulig usikkerhedskilde.

Vandmodellen kalibreres, så den målte udstrømning fra sør bliver en funktion af den målte indstrømning til sør. Den kalibrerede model kan efterfølgende beregne, hvor stort et indstrømmende, umålt vandvolumen, der mangler for fuldt at matche det målte volumen af vand i udløbet. Det beregnede umålte vandvolumen er opdelt i hhv. en grundvandsfødt og en overfladenær tilstrømning. Denne opdeling er valgt ud fra den betragtning, at koncentrationsniveauerne af næringsstoffer typisk ikke er ens for de to typer af vandvolumen.

En ændring i indstrømmet vandvolumen til et bestemt tidspunkt vil påvirke udløbet af vand fra søen over et stykke tid efter indløbstidspunktet, som en "baljeeffekt", hvor et udsving i indløbet ikke i samme omfang vil give udsving i udløbet, pga. den dæmpende effekt, som opstår ved at "baljens" overfladenniveau skal ændre sig, før udløbet fra "baljen" vil ændres. Denne forsinkende effekt er inkluderet i vandmodellen som et lineært reservoir, der understøtter modellen til at beregne umålt indstrømmet vandvolumen tidsmæssigt korrekt alene ud fra kendskab til udløbet.

P- og N-koncentrationen forudsiges ved hjælp af en statistisk model, der opdeler kilderne i tre komponenter: (1) punktkilder, (2) diffuse overfladenære kilder og (3) diffuse grundvandskilder. Punktkildernes indflydelse skal estimeres både for den målte del af oplandet og for den umålte del. Som input til modellen defineres et potentielt punktkildebidrag svarende til en opgørelse af massebidraget fra punktkilder i oplandet uden hensyntagen til den retention, der sker i oplandet fra det sted, punktkilden udledes, til målestationen nedstrøms eller til udløbet i søen. Denne retention betyder, at det potentielle punktkildebidrag vil være større end det reelle punktkildebidrag. Under antagelse af at det reelle punktkildebidrag er proportionalt med det potentielle bidrag, kalibreres det potentielle punktkildebidrag i modellen for N og P til at kunne estimere det reelle punktkildebidrag og dermed tage hensyn til retentionen i oplandet opstrøms.

I dette notat illustreres metoden for vandmodellens reservoarfaktor for 10 søer, mens en mere detaljeret illustrative gennemregning vises for N-belastning af Bryrup Langsø.

2 Metode

2.1 Belastning med N og P

Belastning af søen opdeles i hhv. målt og umålt belastning fra vandtilstrømning, en atmosfærisk belastning fra våd- og tørdeposition og et bidrag fra punktkilder:

$$Load_{in}|_{i,j} \approx Cin_{tot,mea}|_{i,j} \cdot Qin_{mea}|_{i,j} + Cin_{meabase}|_{i,j} \cdot Qin_{unmeabase}|_{j,i} + \\ Cin_{measurf}|_{i,j} \cdot Qin_{unmeasurf}|_{j,i} + \rho \cdot P_{unmea}|_{i,j} + At|_{i,j} \quad (1)$$

hvor

$Load_{in}|_{i,j}$: Belastning af søen, år nr. i og måned nr. j [kg/måned]

$Cin_{tot,mea}|_{i,j}$: Målt indløbskoncentration, år nr. i og måned nr. j [kg/m³]

$Qin_{mea}|_{i,j}$: Målt tilstrømmende vandvolumen, år nr. i og måned nr. j [m³/måned]

$Cin_{meabase}|_{i,j}$: Forudsagt indløbskoncentration i grundvand, år nr. i og måned nr. j [kg/m³]

$Qin_{unmeabase}|_{i,j}$: Umålt grundvandsflow, år nr. i og måned nr. j [m³/måned]

$Cin_{measurf}|_{i,j}$: Forudsagt indløbskoncentration i overfladenært vand, år nr. i og måned nr. j [kg/m³]

$Qin_{unmeasurf}|_{i,j}$: Umålt flow af overfladenært vand, år nr. i og måned nr. j [m³/måned]

ρ : Estimeret andel af potentielle punktkilder, der belaster søen [enhedsløs]

$P_{unmea}|_{i,j}$: Masse af N og P fra potentielle punktkilder i umålt del af oplandet, år nr. i og måned nr. j [kg]

$At|_{i,j}$: Atmosfærisk bidrag af N og P til søen, år nr. i og måned nr. j [kg]

Den totale koncentration af næringsstoffer ($Cin_{tot,mea}|_{i,j}$) er målt på målestasjonen ved indløbet sammen med det tilstrømmende vandvolumen ($Qin_{mea}|_{i,j}$). Koncentrationen i grundvand ($Cin_{meabase}|_{i,j}$) er beregnet ved brug af stofmodellen, hvor det forudsættes, at koncentrationsniveauet af N og P i umålt grundvand er identisk med koncentrationen i målte grundvandsføde tilløb. På samme måde er koncentrationsniveauet af N og P i umålt overfladenært vand ($Cin_{measurf}|_{i,j}$) i indløbet forudsat identisk med koncentrationen i målt overfladenært vand. Koncentrationsniveauet i hhv. målt grundvandsgenereret og målt overfladenær tilstrømning vil blive beregnet ved brug af stofmodellen som beskrevet i et af nedenstående afsnit. De umålte vandvolumener svarende til grundvand ($Qin_{unmeabase}|_{i,j}$) og overfladenært vand ($Qin_{unmeasurf}|_{i,j}$) beregnes ved brug af vandmodellen. De potentielle punktkilder ($P_{unmea}|_{i,j}$) beregnes ved at undersøge de oplande omkring søen, der formodes ikke at bidrage med vand til det målte indløb. Det atmosfæriske bidrag ($At|_{i,j}$) beregnes ved brug af nøgletal for deposition multipliceret med søens overfladeareal.

Udledt stof fra søen forudsiges som:

$$Load_{out}|_{i,j} \approx Cin_{meaout}|_{i,j} \cdot Q_{out}|_{j,i} \quad (2)$$

hvor $Cin_{meaout}|_{i,j}$ er den målte koncentration i udløbet [kg/m³]. I det omfang, der sker udsivning fra søen via bunden, forudsættes det, at koncentrationsniveaueret er ens overalt i søen.

2.2 Belastning af vand (vandmodel)

En volumenbalance af vand for søen giver:

$$\Delta V_{i,j} = Qin_{mea}|_{i,j} + Qin_{unmea}|_{i,j} + A \cdot (q_{pre}|_{i,j} - q_{eva}|_{i,j}) - Qout_{mea}|_{i,j} - Qout_{unmea}|_{i,j} \quad (3)$$

hvor

i, j : Indeks for hhv. år og måned.

$\Delta V_{i,j}$: Ændring i vandvolumen over en måned (m³/måned)

$Qin_{mea}|_{i,j}$: Målt volumen tilløbende vand over måned i for år j (m³/måned)

$Qin_{unmea}|_{i,j}$: Umålt volumen af tilstrømmende vand over måned i for år j (m³/måned)

$q_{pre}|_{i,j}$: Nedbør over måned i for år j (m³/måned)

$q_{eva}|_{i,j}$: Fordampning over måned i for år j (m³/måned)

A : Søareal (m²)

$Qout_{mea}|_{i,j}$: Målt vandvolumen ud af søen over måned i for år j (m³/måned)

$Qout_{unmea}|_{i,j}$: Umålt vandvolumen ud af søen (gennem søbund) over måned i for år j (m³/måned)

Ligning 3 kan omskrives til:

$$Qout_{mea}|_{i,j} = A \cdot (q_{pre}|_{i,j} - q_{eva}|_{i,j}) + Qin_{mea}|_{i,j} + Qin_{unmea}|_{i,j} - Qout_{unmea}|_{i,j} - \Delta V_{i,j} \quad (4)$$

hvor $Qout_{mea}|_{i,j}$ skal måles på station ved søens udløb, mens nedbøren enten kan findes fra en nedbørsmåler i nærheden af søen eller ud fra meteorologiske data. Fordampningen $q_{eva}|_{i,j}$ beregnes ud fra potentiel fordampning. Den umålte del af udløbet $Qout_{unmea}|_{i,j}$ vil typisk være udsivning fra søbunden, da en sø meget sjeldent har udløb. Den effektive vandudskiftning over en måned alene pga. nedbør og fordampning er beregnet som:

$$Qin_{vertical}|_{i,j} = A \cdot (q_{pre}|_{i,j} - q_{eva}|_{i,j}) \quad (5)$$

hvor $Qin_{vertical}|_{i,j}$ er netto tilført vandvolumen til søen pga. nedbør og for-dampning. Tilført vandvolumen fra nedbør, $q_{pre}|_{i,j}$ er estimeret fra meteorologiske data, mens den fordampede vandvolumen, $q_{eval}|_{i,j}$ er estimeret ved brug af en energibalancen for indstråling. Ligning 4 og 5 giver følgende:

$$Qout_{mea}|_{i,j} = Qin_{vertical}|_{i,j} + Qin_{mea}|_{i,j} + Qin_{unmea}|_{i,j} - Qout_{unmea}|_{i,j} - \Delta V_{i,j} \quad (6)$$

Anvendelse af ændringen af søens vandvolumen, $\Delta V_{i,j}$, som parameter i begreningen kræver måling af søens overfladekote, hvilket sjeldent er tilfældet. Der kan kompenseres for den manglende volumenbestemmelse ved at modellere den hydrauliske responstid for et indløb ved brug af lineære reservoarer, hvor de matematiske detaljer fremgår af Sørensen and Nielsen (2023). Resultatet er ligningen:

$$\begin{aligned} Qout_{mea}|_m &= \sum_0^l [F_t \cdot (Qin_{mea}|_{m-t} + Qin_{unmea}|_{m-t} + Qin_{vertical}|_{m-t})] - \\ Qout_{unmea}|_m \end{aligned} \quad (7)$$

Ligning 7 kræver beregning af den umålte indløbne vandmængde, hvilket gøres med følgende udtryk:

$$Qin_{unmea}|_m = \alpha_n \cdot Qin_{mea}|_m + \beta 1_i \cdot \gamma_j \quad (8)$$

hvor α_n er en proportionsfaktor mellem målt og umålt vand for den n 'te kombination af målestationer i indløbet til den specifikke sø. Ændringer i lokationerne, hvor der måles i indløb, kan ske ved, at målestationer flyttes, oprettes, eller nedlægges, og enhver af sådanne ændringer vil give en ny kombination af målestationer.

Parameteren $\beta 1_i$ estimerer det samlede årlige vandvolumen, som strømmer til søen, uden at der er nogen relation til årsvariation i det målt indløb. Dette volumen bestemmes for hvert år (i 'te år). Det vil således være at betragte som grundvandstilstrømningen til søen. Parameteren γ_j estimerer sæsonvariationen af grundvandstilstrømningen, hvilket giver følgende begrænsninger for γ_j :

$$\sum_{j=1}^{12} \gamma_j = 1 \quad (9)$$

I ligning 7 er F_t en lineær reservoarfaktor, som beskriver indflydelsen fra et indløb, der foregået t tidsskridt før tiden for udløbet, og t er dermed et indeks, der tæller antallet af måneder før udløbet, mens l er det antal måneder bagud i tiden, hvor indløbet kan have haft reel betydning for udløbet, og m er et indeks for hele rækken af tidsskridt (måneder) fra det første år, $i=1$, og den første måned $j=1$. Således gælder:

$$m = 12 \cdot (i - 1) + j \quad (10)$$

Ligningen for F_t er udledt i Sørensen and Nielsen (2023) og er følgende:

$$F_0 = 1 + K \cdot \left(1 - e^{-\frac{1}{K}}\right) \quad (11a)$$

$$F_t = K \cdot \left(1 - e^{-\frac{1}{K}}\right) \left(e^{-\frac{t-1}{K}} - e^{-\frac{t}{K}}\right), \quad t = 1, 2, 3, \dots \quad (11b)$$

I Ligning 11 a og b er K en reservoirkoefficient, hvor værdien $K=0$ betyder, at søen er så lille i forhold til gennemløbet af vand, at det kun er indløbet samme

måned som udløbet, der har betydning for udløbet. En større sø med forsinkelse mellem indløb og udløb, vil derimod have en Kværdi, der er markant større end 0. Dette illustreres for 10 sører i de illustrative beregninger senere i denne rapport.

Den endelige statistiske model ser ud som følger:

$$Q_{out,mea}|_m \approx \sum_0^l [F_t \cdot ((1 + \alpha_n) \cdot Q_{in,mea}|_m + \beta 1_t \cdot \gamma_t)] - \beta 2_i, \quad \beta 1_i, \beta 2_i \geq 0 \wedge [\beta 1_i > 0 \rightarrow \beta 2_i = 0] \quad (12)$$

Hvor den lineære reservoarfaktor F_t er estimeret ved brug af ligning 11 a-b. Ligning 12 vil sammen med Ligning 9 og 11 a-b blive anvendt til at give et Maximum Likelihood-estimat af modellens parametre. Parameteren $\beta 1_i$ beskriver den del af indløbet, der ikke afhænger af målt indløb i tilstrømmende vandløb, og tolkes som en tilstrømning af grundvand. Parameteren $\beta 2_i$ beskriver den del af udløbet, der ikke er målt, svarende til en mulig udsivning af vand fra søen gennem søbunden. Denne udsivning er forudsat konstant gennem året, men forskellig fra år til år. Det er nødvendigt at opstille nogle betingelser for de to parametre $\beta 1_i$ og $\beta 2_i$. Det skyldes, at $\beta 1_i$ og $\beta 2_i$ kan kompensere for hinanden, så en tilvækst af værdien for $\beta 1_i$ kan udlignes ved en tilsvarende vækst i værdien for $\beta 2_i$. Denne effekt svarer til, at ekstra tilført vand kan udlignes ved ekstra volumen umålt udløb, men det er ikke muligt simultant at bestemme et umålt indløb og udløb, der ikke afhænger af det målte vandvolumen, i samme Maximum Likelihood-estimat. Vandmodellen antager derfor, at den måned i løbet af et år, hvor der er mindst målt vandindstrømning, er helt domineret af grundvandsgenereret vand, og denne måned betegnes referencemåned for det specifikke år. Hvert år har dermed en referencemåned, der kan være forskellig fra år til år betegnet $j_{ref,i}$. Det giver følgende relation for den grundvandsgenererede og målte tilstrømning:

$$Q_{in,meabase}|_{i,j} \approx Q_{in,mea}|_{j_{ref,i}} \quad (13)$$

Ligning 13 forudsætter, at grundvandsfødt (base-flow) målt vand er konstant over året, men forskellig fra år til år. Den målte tilførte vandmængde, der ikke relateres til grundvandsfødt vand, henregnes til overfladenært vand:

$$Q_{in,measurface}|_{i,j} \approx Q_{in,mea}|_{i,j} - Q_{in,mea}|_{j_{ref,i}} \quad (14)$$

Det umålte tilførte grundvandsgenererede vandvolumen beregnes som et bidrag fra hhv. det umålte vand i tilløbene og grundvandstilstrømning til søen:

$$Q_{in,unmeabase}|_{j,i} \approx \alpha_n * Q_{in,meabase}|_{i,j} + \beta 1_i \cdot \gamma_j \quad (15)$$

Ligning 15 forudsætter, at andelen af grundvandsfødt vand (base-flow) i den umålte tilstrømning er den samme som andelen i den målte. I ligning 15, ses det, at umålt grundvandsfødt vand kan variere både mellem år og hen over året mellem månederne. Da den overfladenære tilstrømning forudsættes at udvise store variationer fra måned til måned som respons på aktuelle hydrologiske forhold, og da disse variationer forudsættes ens inden for søens op-

land, forudsættes det, at den umålte overfladenære tilstrømning af vand korrelerer med den målte tilstrømning. Den umålte overfladenære vandtilstrømning er derfor beregnet som:

$$Qin_{unmeasurface}|_{j,i} \approx \alpha_n * Qin_{measurface}|_{i,j} \quad (16)$$

Det samlede volumen vand tilført søen med base-flow over en måned er beregnet som:

$$Qin_{base}|_{j,i} \approx Qin_{meabase}|_{i,j} + Qin_{unmeabase}|_{i,j} \quad (17)$$

Det samlede volumen vand tilført søen fra overfladenær tilstrømning over en måned er beregnet som:

$$Qin_{surface}|_{j,i} \approx (1 + \alpha_n) * Qin_{measurface}|_{i,j} \quad (18)$$

mens den samlede mængde udløbende vand er givet som:

$$Qout|_{j,i} = \beta 2_i + Qout_{mea}|_{i,j} \quad (19)$$

2.3 Koncentration af P og N (stofmodel)

Massebalance for stof (både N og P) i det målte vand i indløbet er

$$\begin{aligned} Qin_{tot,mea}|_{i,j} \cdot Qin_{mea}|_{i,j} &= Qin_{meabase}|_{i,j} \cdot Qin_{meabase}|_{i,j} + Qin_{measurf}|_{i,j} \cdot \\ &Qin_{measurf}|_{i,j} + P_{effmea}|_{i,j} \end{aligned} \quad (20)$$

hvor $Cin_{tot,mea}|_{i,j}$ er totalkoncentrationen af stof i den målte del af indløbet svarende til $Qin_{mea}|_{i,j}$, mens $Cin_{meabase}|_{i,j}$ og $Cin_{measurf}|_{i,j}$ er totalkoncentrationen i hhv. det målte grundvandsfødte og målte overfladenære vandvolumen. Det effektive bidrag fra punktkilder i det målte indløb er taget med som $P_{effmea}|_{i,j}$ og er defineret som den masse, der ankommer i søens målte indløb, og som er et resultat af punktkildebelastningen i det målte opland. Ligning 20 kan omskrives til følgende udtryk med brug af den beskrivende dimensionsløse variable $X_{i,j}$:

$$\frac{Qin_{tot,mea}|_{i,j}}{Qin_{mea}|_{i,j}} = \frac{Cin_{meabase}|_{i,j} + (Cin_{measurf}|_{i,j} - Qin_{meabase}|_{i,j}) \cdot X_{i,j} + P_{effmea}|_{i,j}}{Qin_{mea}|_{i,j}}, \text{ hvor } X_{i,j} = \frac{(Cin_{measurf}|_{i,j} - Qin_{meabase}|_{i,j}) \cdot X_{i,j}}{Qin_{mea}|_{i,j}} \quad (21)$$

Kvantificeringen af en given punktkildes reelle bidrag til belastningen vil afhænge af både punktkildens styrke, hvor udledningen finder sted og retentionen i oplandet. Derfor er der defineret en parameter ρ svarende til den andel af masse fra punktkilder i det målte opland, der optræder i søens indløb

$$P_{effmea}|_{i,j} = \rho \cdot P_{mea}|_{i,j} \quad (22)$$

hvor $P_{mea}|_{i,j}$ er den samlede punktkildestyrke i det målte opland, og ρ vil blive bestemt med den statistiske model. Ud fra ligning 21 og 22 kan følgende statistiske model nu opstilles:

$$Y_{i,j} = \theta_0 + \theta_1 \cdot m + (\gamma_1 i + \gamma_2 j) \cdot X_{i,j} + \rho \cdot \frac{P_{mea}|_{i,j}}{Qin_{mea}|_{i,j}} \quad (23)$$

hvor koncentrationen af stof i grundvandsfødt målt indløb er forudsat beskrevet som en lineær relation i tiden med ligningen:

$$Cin_{meabase}|_{i,j} \approx \theta_0 + \theta_1 \cdot m \quad (24)$$

Argumentet er, at grundvandsfødt vand er flere år om at dannes og dermed tilskrives en træg ændring over tid. Parametrene ρ , θ_0 , θ_1 , γ_1_i og γ_2_j bestemmes alle med brug af den statistiske model (ligning 23).

Den proportionalitetsfaktor, der multipliceres på $X_{i,j}$ er forudsat at bestå af et bidrag, der varierer over året, og en forskel mellem år: $\gamma_1_i + \gamma_2_j$. Der kan være en tidslig udvikling over årene, der ikke er ens for hhv. grundvandsfødt og overfladenært vand. Hvis der f.eks. sker et ændret N- og P-forbrug på markfladen, vil dette først have en effekt på overfladenært vand, da grundvandsfødt vand har en større forsinkelse. Denne forskel beskrives med parameteren γ_1_i . Det formodes desuden, at rodzonens optag af N og P kan have en vis betydning for koncentrationsniveauet i det overfladenære vand, og parameteren γ_2_j kan derfor tolkes som en mulig sæsonvariation pga. en årstidsafhængig betydning af planteoptaget.

Ved at sammenligne ligning 21 og 23 ses det, at ledet $(Cin_{measurf}|_{i,j} - Cin_{meabase}|_{i,j})$ i ligning 21 svarer til ledet $(\gamma_1_i + \gamma_2_j)$ i ligning 23, så koncentrationen i det målte tilløbende overfladenære vand kan estimeres ved:

$$Cin_{measurf}|_{i,j} \approx Cin_{meabase}|_{i,j} + \gamma_1_i + \gamma_2_j \quad (25)$$

Det følger af ovenstående, at koncentrationen i det overfladenære vand er mere dynamisk i forhold til både sæsonen og fra et år til et andet end koncentrationen i grundvandet.

Som det ses i ligning 23, er punktkilderne inkluderet som en del af den afhængige variabel Y. Bidraget fra disse punktkilder er ikke direkte målbart og må derfor estimeres som et bidrag opstrøms i oplandet til en given målestation.

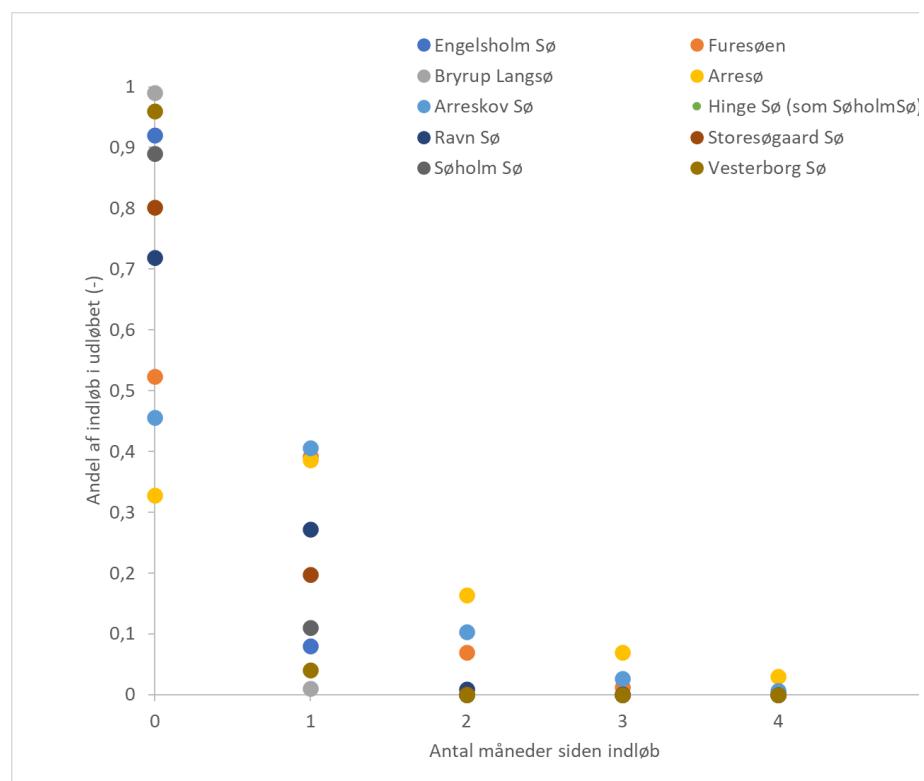
De forventede koncentrationsniveauer fra ligningerne 24 og 25 er ikke målbare størrelser, der kan valideres. Således kan forudsigtelserne fra ligningerne 24 og 25 evt. justeres eller erstattes ved hjælp af ekspertvurdering, hvis værdierne forekommer urealistiske, og der findes lokal viden om korrekte værdiintervaller.

3 Illustrative beregninger

3.1 Hydraulisk responstid for 10 danske søer

Indløb af vand i en ø på et bestemt tidspunkt vil påvirke udløb fra øen over et stykke tid efter indløbstidspunktet. Denne forsinkende effekt i hydraulisk respons mellem indløb og udløb er med i vandmodellen som et lineært reservoir (ligning 11 a-b), og svarer ikke til vandets hydrauliske opholdstid, selvom de to begreber er beslægtede. Den forsinkende effekt, når der anvendes månedsvandføring, er ubetydelig for mindre øer med stor hydraulisk belastning (stor vandtilstrømning) som f.eks. Bryrup Langsø. For større øer kan der være flere måneders forsinkelse på et udløb. Den hydrauliske forsinkelse er vist for 10 danske øer på figur 2.1 og kan forklares med følgende eksempel: For Furesøen, måned nr. 0, er værdien 0,52 (orange punkt på y-aksen), hvilket indikerer, at indløb af 1 m^3 vand til øen giver anledning til et udløb af $0,52 \text{ m}^3$ vand den samme måned. De $0,52 \text{ m}^3$, der løber ud pga. et indløb på 1 m^3 , er ikke det samme vand, som løb ind, men kan forstås som vand, der ikke ville være løbet ud, hvis der ikke var tilført 1 m^3 . Måneden efter (måned nr. 1) giver indløbet af 1 m^3 i indløb månedens før anledning til udløb af ca. $0,4 \text{ m}^3$ vand fra Furesøen, og måneden efter igen kommer der stadig et lille "skvulp" ud af øen svarende til knap $0,1 \text{ m}^3$. Det fremgår af figur 2.1, at de hydraulisk hårdt belastede (lille opholdstid) øer, såsom Bryrup Langsø, Vesterborg Sø, Engelsholm Sø, Søholm Sø, og Hinge Sø, alle har en hurtig responsid, hvor et volumen i indløb stort set giver samme volumen i udløbet for den samme måned, mens de øvrige øer har en betydelig forsinkelse med Arresø som øen med størst forsinkelse. Arresø adskiller sig fra de andre øer ved, at der løber mere vand ud i månedens efter indløbet, end den måned indløbet foregår i. Dette er en konsekvens af, at en del af en given måneds indløb foregår sidst på månedens, og at dette vand kun i begrænset omfang giver anledning til en påvirkning af udløbet i samme måned som indløbet. Denne egenskab, at et indløb i en given måned har størst betydning for udløbet månedens efter, er fælles for meget store øer, hvor forsinkelsen er så stor, at et indløb primært bidrager til udløbet de efterfølgende måneder. Denne forsinkelse mellem indløb og udløb anvendes til at placere et volumen af umålt vand mest korrekt i forhold til, hvilken måned et vandvolumen er strømmet ind i øen for at relaterne det rigtige vandvolumen til de målte indløbskoncentrationer for N og P på det pågældende tidspunkt. Det har størst betydning at inddrage denne forsinkelse, hvor koncentrationsniveauet varierer mellem måneder, og især hvis der er en korrelation mellem vandføring og koncentrationsniveau.

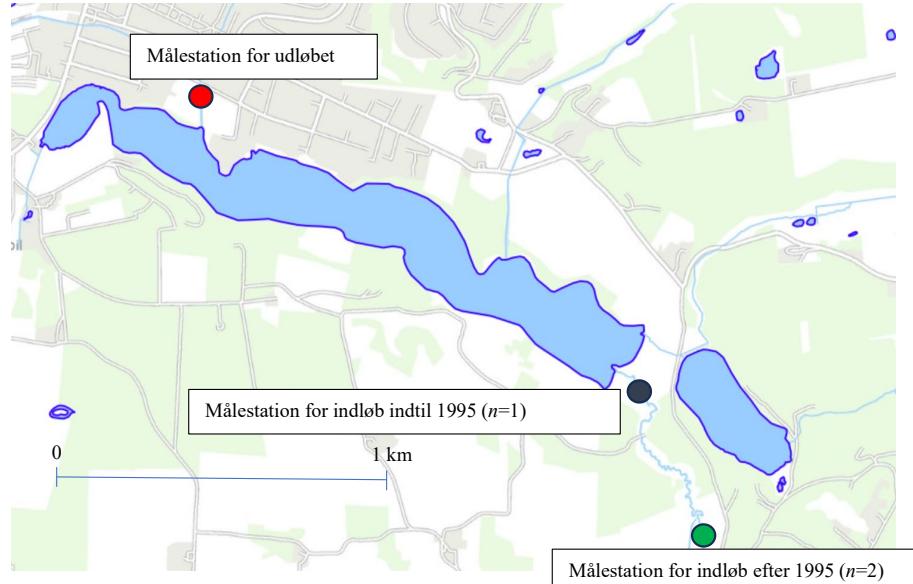
Figur 2.1 Hydraulisk forsinkelse for sørerne. Bemærk, at Hinge Sø og Søholm Sø følger hinanden.



3.2 Bryrup Langsø

Bryrup Langsø er en mellemstor lavvandet sø, der dækker et areal på 37 ha og med en gennemsnitlig dybde på 4,6 m. Indløbsmålestasjonen er placeret i hovedindløbet, men stationen blev flyttet mellem år 1995 og 1996. Andelen af målt vand kan have ændret sig pga. denne flytning, hvorfor der regnes med to kombinationer af målestationer i indløbet ($n = 2$). Desuden ses i figur 2.2, at der er tre mindre vandløb, der løber ud i søen uden at være opmålt, hvilket yderligere indikerer, at modellen skal identificere en vis andel af umålt vand i indløb til søen.

Figur 2.2. Oversigt over Bryrup Langsø med angivelse af målestationernes placering.



Data indsamles på målestationerne som en del af Danmarks Nationale Overvågningsprogram (NOVANA), Svendsen og Norup. (2005a) og Svendsen m.fl. (2005b), og data i casestudiet er hentet fra dette program. Tidserien dækker perioden fra 1990 til og med 2021 med månedsværdier (Sørensen and Nielsen, 2023).

Vandmodel

Estimater for koefficienterne i ligning 11 a-b og 12 er givet for Bryrup Langesø i tabel 2.1. Det fremgår af tabellen, at β_{1_i} , volumenet af estimeret tilstrømmende grundvand, er estimeret for hvert år, og at der årligt strømmer mellem 0,1 og 3,66 millioner m^3 målt grundvand ind i søen, mens β_{2_i} , udsivning af vand gennem søbunden, som konsekvens af disse vandmængder, er estimeret til 0. Søen modtager dermed en del vand, som ikke afhænger af den aktuelle tilførsel af vand i det målte tilløb, hvilket tyder på en stor grundvandstilstrømning til søen. Den generelle sæsonvariation for denne grundvandstilstrømning ses i γ_j , der angiver andelen af grundvand, der er tilstrømmet i måned nr. j . F.eks. er der den første måned ($j=1$) og det første år ($i=1$) ifølge modellen strømmet $\beta_{1_1} \cdot \gamma_1 = 90000 \cdot 0,073 = 7047 m^3$ grundvand til søen. Det ses i tabel 1, at α_1 og α_2 reelt har identiske værdier, svarende til 0,2, hvilket betyder, at stationsflytningen fra år 1995 til år 1996 vist på figur 2.2 ikke har haft målbar betydning for andelen af umålt vand i tilløbet. Dette underbygges også af, at flytningen er af en relativt beskeden distance. Den estimerede værdi for α_n på 0,20 indikerer, at der løber 20 % mere vand til søen, end der måles på målestationen. Det virker troværdigt ud fra, at der tre mindre og umålte tilløb til søen, jf. figur 2.2.

Tabel 2.1. Koefficienterne i ligning 11 og 12, estimeret for Bryrup Langsø ($R^2=0,95$). Indeks 1-32 svarer til indeks for år, mens indeks 1-12 svarer til indeks for måned nr. og indeks 1-2 svarer til indeks for forskellig placering af målestation for indløbet.

Koefficient	Estimat	Enhed	Koefficient	Estimat	Enhed	Koefficient	Estimat	Enhed
β_{1_1}	900000	$m^3/år$	$\beta_{1_{28}}$	1700000	$m^3/år$	$\beta_{2_{23}}$	0	$m^3/måned$
β_{1_2}	1040000	$m^3/år$	$\beta_{1_{29}}$	1600000	$m^3/år$	$\beta_{2_{24}}$	0	$m^3/måned$
β_{1_3}	900000	$m^3/år$	$\beta_{1_{30}}$	2120000	$m^3/år$	$\beta_{2_{25}}$	0	$m^3/måned$
β_{1_4}	860000	$m^3/år$	$\beta_{1_{31}}$	2440000	$m^3/år$	$\beta_{2_{26}}$	0	$m^3/måned$
β_{1_5}	1660000	$m^3/år$	$\beta_{1_{32}}$	1720000	$m^3/år$	$\beta_{2_{27}}$	0	$m^3/måned$
β_{1_6}	3660000	$m^3/år$	β_{2_1}	0	$m^3/måned$	$\beta_{2_{28}}$	0	$m^3/måned$
β_{1_7}	920000	$m^3/år$	β_{2_2}	0	$m^3/måned$	$\beta_{2_{29}}$	0	$m^3/måned$
β_{1_8}	560000	$m^3/år$	β_{2_3}	0	$m^3/måned$	$\beta_{2_{30}}$	0	$m^3/måned$
β_{1_9}	1180000	$m^3/år$	β_{2_4}	0	$m^3/måned$	$\beta_{2_{31}}$	0	$m^3/måned$
$\beta_{1_{10}}$	720000	$m^3/år$	β_{2_5}	0	$m^3/måned$	$\beta_{2_{32}}$	0	$m^3/måned$
$\beta_{1_{11}}$	1620000	$m^3/år$	β_{2_6}	0	$m^3/måned$	γ_1	0,073	$1/måned$
$\beta_{1_{12}}$	480000	$m^3/år$	β_{2_7}	0	$m^3/måned$	γ_2	0,098	$1/måned$
$\beta_{1_{13}}$	1360000	$m^3/år$	β_{2_8}	0	$m^3/måned$	γ_3	0,134	$1/måned$
$\beta_{1_{14}}$	1280000	$m^3/år$	β_{2_9}	0	$m^3/måned$	γ_4	0,110	$1/måned$
$\beta_{1_{15}}$	1360000	$m^3/år$	$\beta_{2_{10}}$	0	$m^3/måned$	γ_5	0,110	$1/måned$
$\beta_{1_{16}}$	1700000	$m^3/år$	$\beta_{2_{11}}$	0	$m^3/måned$	γ_6	0,098	$1/måned$
$\beta_{1_{17}}$	980000	$m^3/år$	$\beta_{2_{12}}$	0	$m^3/måned$	γ_7	0,073	$1/måned$
$\beta_{1_{18}}$	540000	$m^3/år$	$\beta_{2_{13}}$	0	$m^3/måned$	γ_8	0,073	$1/måned$
$\beta_{1_{19}}$	920000	$m^3/år$	$\beta_{2_{14}}$	0	$m^3/måned$	γ_9	0,061	$1/måned$
$\beta_{1_{20}}$	620000	$m^3/år$	$\beta_{2_{15}}$	0	$m^3/måned$	γ_{10}	0,049	$1/måned$
$\beta_{1_{21}}$	100000	$m^3/år$	$\beta_{2_{16}}$	0	$m^3/måned$	γ_{11}	0,061	$1/måned$
$\beta_{1_{22}}$	760000	$m^3/år$	$\beta_{2_{17}}$	0	$m^3/måned$	γ_{12}	0,061	$1/måned$
$\beta_{1_{23}}$	1060000	$m^3/år$	$\beta_{2_{18}}$	0	$m^3/måned$	α_1	0,20	-
$\beta_{1_{24}}$	1280000	$m^3/år$	$\beta_{2_{19}}$	0	$m^3/måned$	α_2	0,21	-
$\beta_{1_{25}}$	880000	$m^3/år$	$\beta_{2_{20}}$	0	$m^3/måned$	K	0,01	-
$\beta_{1_{26}}$	1360000	$m^3/år$	$\beta_{1_{21}}$	0	$m^3/måned$			

$\beta_{1_{27}}$	1800000	m ³ /år	$\beta_{1_{22}}$	0	m ³ /måned
------------------	---------	--------------------	------------------	---	-----------------------

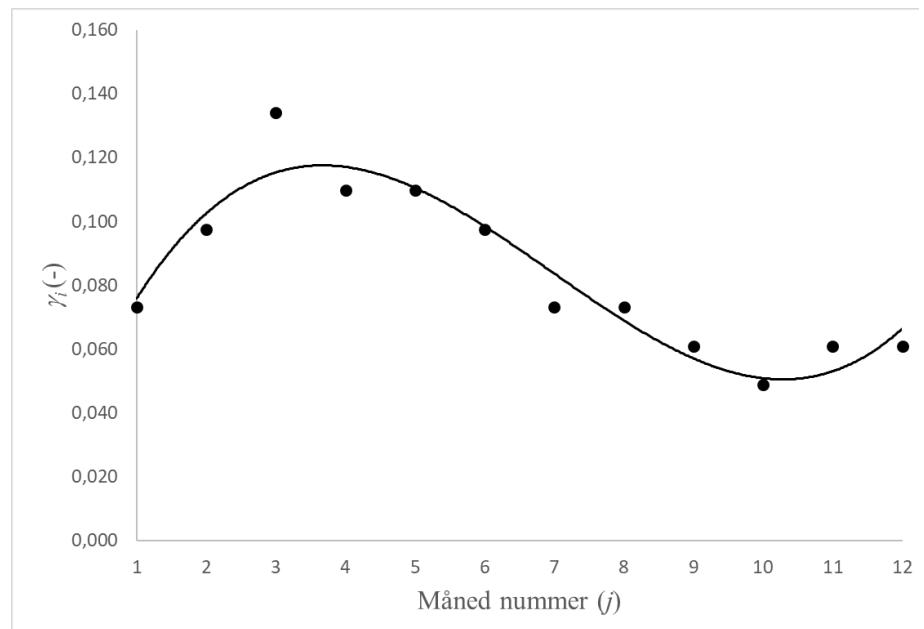
Modellen tolker β_{1_i} som en grundvandstilstrømning til søen, hvilket er en tolkning, som er svær at validere, da det ikke er muligt at måle denne vandmængde. Det er dog muligt at foretage en indirekte validering ved at korrelerede estimerede β_{1_i} -værdier med nedbørsmængden i søens opland. En sådan korrelation er vist i tabel 2.2, hvor værdien for β_{1_i} er søgt beskrevet ved brug af nødbørsmængden (mm/år) for hhv. det samme år som det år, hvor β_{1_i} blev estimeret ($Prec_i$), året før ($Prec_{i-1}$) og året før igen ($Prec_{i-2}$). Tabel 2.2. er en sekventiel ANOVA, hvilket betyder, at nedbør samme år inddrages først som forklarende variabel med en forklaret kvadratsum på $91 \cdot 10^6$, hvilket er beskedent sammenlignet med en uforklaret kvadratsum (residualer) på $75714 \cdot 10^6$. Derefter indføres nedbør året før som forklarende variabel alene for den variation, som nedbøren samme år ikke kan forklare. Her forklares en kvadratsum på $17899 \cdot 10^6$, hvilket er langt mere end nedbøren samme år. Til sidst afprøves nedbøren to år før for at se, om denne variabel kan forklare noget af den resterende kvadratsum. Her ses igen en meget mindre forklaringskraft med værdien $416 \cdot 10^6$. Af disse tre mulige bidrag for hvert sit år er det derfor kun nedbøren året før, som er signifikant med en p værdi på 0,02 (Pr=0,02). Tabel 2.2 viser således, at β_{1_i} -estimatet alene korrelerer med nedbøren året før ($Prec_{i-1}$), mens der ikke er nogen sikker korrelation med nedbør hhv. samme år ($Prec_i$) eller nedbøren to år før ($Prec_{i-2}$). Dette indikerer, at β_{1_i} udtrykker en grundvandstilstrømning, som er påvirket af nedsivning året før, men som er upåvirket af mere nylig falden nedbør.

Tabel 2.2. ANOVA-tabel, der viser og tekster betydningen for β_i af årsnedbør hhv. samme år ($Prec_i$), året før ($Prec_{i-1}$) og to år før ($Prec_{i-2}$) det år, hvor β_{1_i} er estimeret. Tabellen er sequentiel anova som forklaret i teksten.

Respons variabel: β_{1_i}					
Beskrevende variable	Kvadrat Sum	Df	F	Pr (>F)	
$Prec_i$	$91 \cdot 10^6$	1	0,031	0,86	
$Prec_{i-1}$	$17899 \cdot 10^6$	1	6,147	0,02	
$Prec_{i-2}$	$416 \cdot 10^6$	1	0,143	0,71	
Residualer	$75714 \cdot 10^6$	26			

Grundvandstilstrømmingens variation hen over året er modelleret ved brug af koefficienten γ_j , og den årlige variation for Bryrup Langsø af koefficienten γ_j ses på figur 2.3. Figuren viser en sæsonvariation, hvor grundvandstilstrømmingen stiger gennem vinteren, hvor grundvandsmagasinet fyldes op, mens den falder gennem sommerhalvåret, hvor grundvandsmagasinet tømmes.

Figur 2.3. Variation over året i estimeret umålt grundvandstilførelse til søen (figur fra Sørensen og Nielsen, 2023).



Modellen skal forudsige den umålte belastning. Mængden af umålt vand i forhold til målt vand er opgjort i tabel 2.1 Det umålte udløb ($Q_{out,unmea}|_{i,j}$) er estimeret til nul for hele perioden.

Tabel 2.3. Den estimerede månedlige gennemsnitlige tilførsel af vand eksklusiv fordampning og nedbør direkte i søen ($Q_{in_{vertical}|i,j}$).

Kombination af målestation (n)	Middeltilførsel (1000 m ³ /måned)			
	Reservoirtyp	Målt	Umålt	Total
1	Grundvandsfødt	200	129	330
	Overfladenært	291	58	350
	Total	492	188	679
2	Grundvandsfødt	214	152	366
	Overfladenært	225	47	273
	Total	439	199	638

Tabel 2.3 viser, at det umålte vandvolumen udgør omkring 1/3 af indløbet, og at den umålte del domineres af grundvandstilført vand.

Kvælstof-model

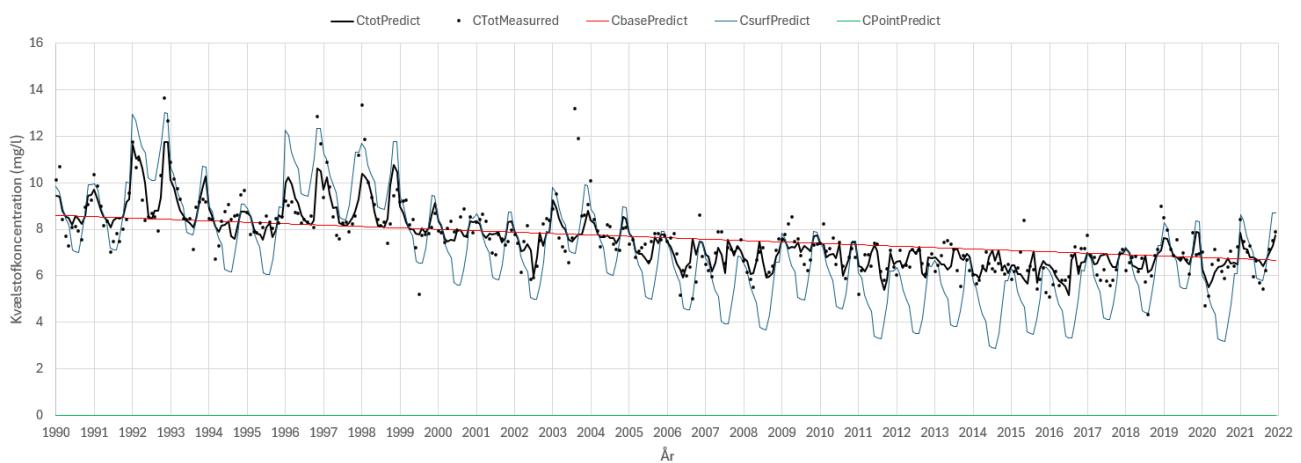
Den målte N-koncentration i indløbet er anvendt sammen med den målte vandmængde i ligning 23 og under forudsætning af, at det målte base-flow svarer til referenceflowet for det pågældende år (se tabel 2.4). Punktkildefaktoren ρ blev estimeret til en værdi tæt på nul, hvorfor punktkilder er negligeret i modellen. Dette er i overensstemmelse med, at opgørelsen over punktkilder i oplandet ikke har påvist betydelige punktkilder (Sørensen and Nielsen, 2023).

Tabel 2.4. Estimerede parametre for ligning 23, $R^2=0,70$.

Parameter	Estimat	Enhed	Parameter	Estimat	Enhed	Parameter	Estimat	Enhed
γ_{1_1}	0	mg/l	$\gamma_{1_{17}}$	-1,51	mg/l	γ_{2_1}	1,24	mg/l
γ_{1_2}	0,15	mg/l	$\gamma_{1_{18}}$	-2,05	mg/l	γ_{2_2}	1,00	mg/l
γ_{1_3}	3,21	mg/l	$\gamma_{1_{19}}$	-2,25	mg/l	γ_{2_3}	0,29	mg/l
γ_{1_4}	0,95	mg/l	$\gamma_{1_{20}}$	-0,88	mg/l	γ_{2_4}	-0,12	mg/l
γ_{1_5}	-0,59	mg/l	$\gamma_{1_{21}}$	-1,21	mg/l	γ_{2_5}	-0,39	mg/l
γ_{1_6}	-0,66	mg/l	$\gamma_{1_{22}}$	-2,44	mg/l	γ_{2_6}	-1,47	mg/l
γ_{1_7}	2,79	mg/l	$\gamma_{1_{23}}$	-2,17	mg/l	γ_{2_7}	-1,54	mg/l
γ_{1_8}	1,82	mg/l	$\gamma_{1_{24}}$	-1,81	mg/l	γ_{2_8}	-1,56	mg/l
γ_{1_9}	2,33	mg/l	$\gamma_{1_{25}}$	-2,68	mg/l	γ_{2_9}	-0,91	mg/l
$\gamma_{1_{10}}$	0,07	mg/l	$\gamma_{1_{26}}$	-2,02	mg/l	$\gamma_{2_{10}}$	0,05	mg/l
$\gamma_{1_{11}}$	-0,81	mg/l	$\gamma_{1_{27}}$	-2,12	mg/l	$\gamma_{2_{11}}$	1,37	mg/l
$\gamma_{1_{12}}$	-0,52	mg/l	$\gamma_{1_{28}}$	-1,26	mg/l	$\gamma_{2_{12}}$	1,36	mg/l
$\gamma_{1_{13}}$	-1,30	mg/l	$\gamma_{1_{29}}$	-0,92	mg/l	θ_0	8,61	mg/l
$\gamma_{1_{14}}$	0,76	mg/l	$\gamma_{1_{30}}$	0,18	mg/l	θ_1	-0,0050	mg/(l·t)
$\gamma_{1_{15}}$	-0,13	mg/l	$\gamma_{1_{31}}$	-2,02	mg/l	ρ	0	-
$\gamma_{1_{16}}$	-1,10	mg/l	$\gamma_{1_{32}}$	0,66	mg/l			

Tidsserier med modellerede koncentrationer er vist i figur 2.4 sammen med de målte værdier. Den tykke sorte kurve er modelberegnet totalkoncentration, som kan sammenlignes med målingerne (punkterne). Ekstreme målte værdier synes dårligt beskrevet af modellen. Den tynde blå kurve er den modelberegnede koncentration i det overfladenære vand, mens den røde kurve er den modelberegnede koncentration i base-flow. Koncentrationen i det overfladenære vand udviser variation over året med høje værdier i vinterperioden og de laveste værdier om sommeren. Dette stemmer fint overens med, at det overfladenære vand er i forbindelse med rodzonen og derfor er påvirket af udvaskning og planteoptag. Forskellen i koncentration i det overfladenære vand mellem vinter og sommer er ca. 3 mg/l. Koncentrationen i base-flow forudsættes at falde lineært gennem perioden, fordi det forudsættes, at

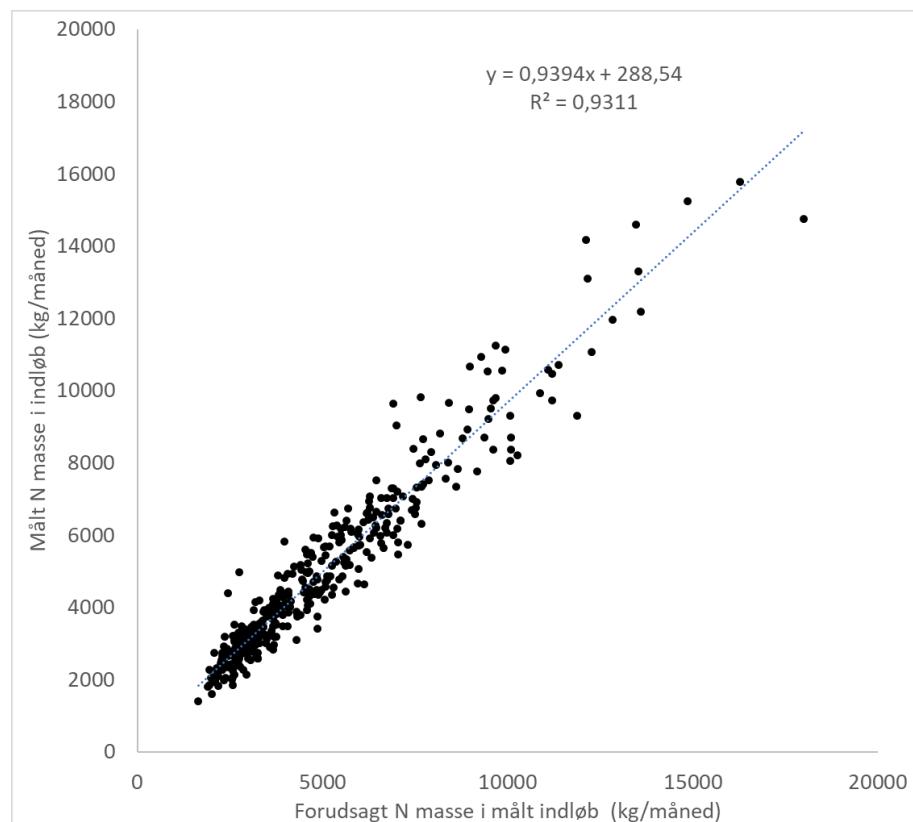
grundvandskoncentrationen ikke ændrer sig dramatisk fra år til år, men at der her er tale om mere langsigtede ændringer, der godt kan beskrives som en lineær ændring over hele perioden. Da sommervandføringen domineres af grundvandsfødt vand, bør de målte koncentrationer om sommeren (mellem gridlinjerne) være tæt på den røde linje på figur 2.4. Dette kan også observeres på figuren, dog med en tendens til at sommer-koncentrationerne ligger under den røde kurve en del år i perioden 2006-2018. Koncentrationen i det overfladenære vand synes at falde væsentligt gennem perioden, hvilket stemmer overens med, at netop det overfladenære vand responderer tydeligt på tiltag i landbruget. Der ses desuden en tendens til, at koncentrationsniveauet øges lidt i det overfladenære vand sidst i perioden. Det er især estimatet af det koncentrationsniveau i base-flow, der er kritisk for massebalancen, fordi det meste umålte vand er grundvandsfødt jf. tabel 2.3.



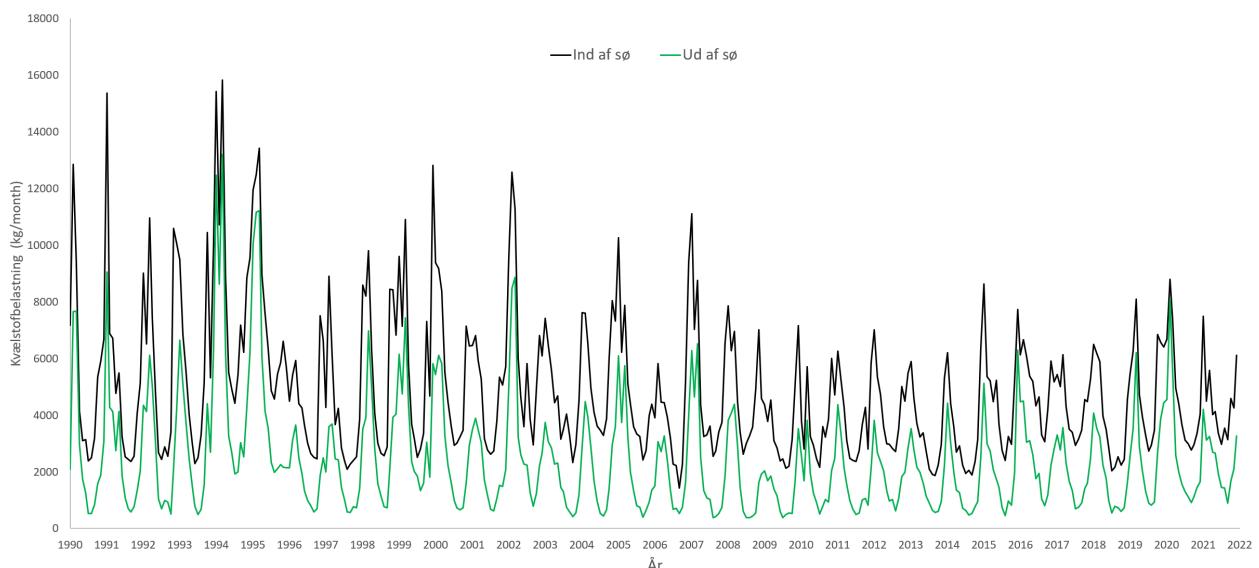
Figur 2.4. Beregnede (sort kurve) samt målte (punkter) koncentrationsniveauer af N i indløbet. Den blå kurve er det beregnede koncentrationsniveau i det overfladenære vand, mens den røde kurve er beregnet koncentration i base-flow. De lodrette gridlinjer svarer til hele år. (Figur fra Sørensen og Nielsen, 2023).

Formålet med den statistiske model er at beregne N-belastningen for den ikke-målte del af det tilførte vandvolumen, mens den målte belastning kan anvendes til at teste modellens evne til at forudsige belastningen ved at sammenligne den modelberegnede ”målte” belastning med den faktisk målte. En sådan test er vist på figur 2.5.

Figur 2.5. Sammenligning af modelbereget og målt belastning, alene for den målte andel af indløbet vand (figur fra Sørensen og Nielsen, 2023).



Den endelige opgjorte månedlige kvælstofbelastning, som udgøres af den målte belastning og den beregnede umålt belastning med brug af modellens forudsigelser for det umålte land er vist på figur 2.6. I figuren er den målte tilførsel af kvælstof vist sammen med den målte kvælstoftransport ud af søen. I Bryrup Langsø ses en betydelig tilbageholdelse af kvælstof svarende til gennemsnitligt 2500 kg N pr. måned.



Figur 2.6. Den samlede N-belastning ind og ud af Bryrup Langsø, hvor målinger er huludfyldt med modelberegninger (figur fra Sørensen og Nielsen, 2023).

4 Diskussion

Formålet med udviklingen af modellen er at forudsige transporten af kvælstof (N) og fosfor (P) ind og ud af en sø med ufuldstændigt opmålte vandvolumener, og hvor den ikke-målte belastning skal forudsiges ved brug af viden om den målte andel af vandet. I de fleste tilfælde vil en given sø kun have et enkelt udløb, hvorfor udstrømningen har tendens til at være bedre beskrevet end tilstrømningen. Søer kan dog have udsivning gennem bunden, hvilket vil bidrage til tab af N og P på samme måde som udløbet ved overfladen. Indløbet til søen kan bestå af flere vandløb, hvoraf ofte kun en del er målte. Der vil ofte være målt på et hovedtilløb, mens mindre vandløb ikke måles. Det er klart, at der er usikkerheder relateret til de målte vandvolumener og næringsstofkoncentrationsniveauer, men i dette notat er alene fokuseret på at minimere den usikkerhed, der introduceres fra den ikke-målte andel af belastningen. Kvalitetssikring af de overvågningsdata, der bruges som datainput, er afgørende for belastningsberegningen. Den velkendte generelle modelleringsbetingelse "garbage in is garbage out" vil også gælde for belastningsforudsigelserne: Hvis datainputtet til belastningsforudsigelserne ikke er af tilstrækkelig kvalitet, vil det medføre betydelig usikkerhed.

Vandmodellen udtrykker udløbet vandvolumen som funktion af målt indløbet volumen. Det kan grafisk ses som en funktion med målt udløb på y-aksen og målt indløb på x-aksen. I denne model tolkes skæringspunktet på y-aksen som den umålte netto-mængde vand, der tilføres søen under fiktive vilkår, hvor der ikke kommer noget målt vand ind i søen, altså som en nettvandudveksling gennem sør bunden eller fra umålte vandkilder, der løber til søen uafhængigt af målt tilløb. Hvis skæringen estimeres som negativ, tolkes det som tab af vand på grund af udsivning gennem sør bunden. Under ideelle forhold, hvor alt tilløbet vand er målt, vil der være en hældning på kurven mellem indløbet vand og udløbet vand på værdien én. I det omfang denne hældning er større end én (der løber mere vand ud af søen, end der er målt at løbe ind), tolkes det som, at der findes en andel af umål vand i den reelle indstrømmende vandmængde. Dette er baseret på det rationale, at hvis der f.eks. måles to m^3 vand ved udløbet af søen for hver målt én m^3 ved indløbet (hældning lig med to), så er der én m^3 umål vand i tilløbet for hver m^3 målt vandvolumen i indløbet. Sæsonudsvinget i base-flow antages uafhængigt af målestationskombinationen. Denne antagelse kan dog anfægtes, hvis det umålte vandvolumen er stort. I sådanne tilfælde vil det imidlertid altid være en udfordring at foretage valide belastningsforudsigelser. En ændring i placering af målestationerne er i modellen antaget at finde sted i begyndelsen af et år, hvilket typisk er tilfældet i overvågningsprogrammet. Den statistiske model for vandbalance kan dog ændres, så den omfatter ændringer i målingen fra en hvilken som helst måned og et hvilket som helst år.

Modellen antager, at den mindste månedlige tilstrømning af vand i et år alene består af base-flow, hvorfor denne måned bruges som referencemåned for bestemmelse af det målte grundvandsfødte vandvolumen det pågældende år. Derfor kan våde perioder over sommeren, som sædvanligvis er perioden med lavest vandføring, påvirke modellens estimatorer af base-flow om sommeren, hvilket kan føre til overestimering af base-flow i referencemåneden. Desuden vil denne målte andel af base-flow variere over året, hvilket modellen ikke tager højde for i den nuværende version.

Akkumuleringen af vand i søen er inkluderet i modellen som en hukommelseseffekt i en lineær reservoirmodel, der inkluderer de seneste måneders vandininput for at modellere den aktuelle måneds output. Udstrømningen fra søen antages således at være lineært afhængig af vandvolumenet i søen. Dette er typisk gyldigt for mindre udsving i ind- og udløbende vandmængder, da større ændringer i forholdet mellem indgående vandmængde og udgående vandmængde måneden efter vil have en tendens til at være ikke-lineære.

N- og P-koncentrationen forudsiges ved hjælp af en statistisk model, der opdeler kilderne i tre komponenter: (1) punktkilder, (2) grundvandsføde kilder og (3) overfladenære kilder. Punktkildernes styrke skal identificeres både i oplandet til det målte tilløb og i søens opland, og modellen estimerer punktkildens reelle bidrag. Punktkildebiddinget kan enten beskrives som et potentiel eller reelt bidrag. Et potentiel bidrag svarer til den opgjorte kildestyrke (massebidrag) i oplandet fra punktkilder uden hensyntagen til retentionen i oplandet fra det sted, punktkilden udleder til vandløbet og nedstrøms til udløbet i søen. Det reelle punktkildebidding er det bidrag, punktkilder reelt giver til belastningen af søen under hensyntagen til retentionen. En opgørelse af punktkilder i søens opland kan bruges til at beregne det potentielle punktkildebidding, mens det reelle bidrag er svært at prædiktere, da det også kræver en meget detaljeret viden om retentionen. Modellen søger at løse dette dilemma ved statistik at estimere det effektive punktkildebidding gennem fitning af massebalancen for indløbet til det potentielle punktkildebidding (Sørensen og Nielsen, 2023). Denne fitning udføres ud fra den betragtning, at det reelle punktkildebidding, modsat mere diffuse bidrag, ikke følger vandflowet, og derfor ikke kan estimeres i indløbet sammen med koncentrationsbidraget for hhv. grundvandsfødt vand og overfladenært vand. En vis andel af punktkildebiddinget kan dog med denne metode fejlagtigt blive modelleret som en del af mere diffuse kilder, især hvis punktkildene er placeret på længere afstand opstrøms søen, og hvis der er andre søer opstrøms, der påvirker koncentrationsniveauet i det tilstrømmende vand. Det vurderes derfor, at modellen er bedst til at håndtere punktkilder for søer, i hvis tilstrømning der ikke findes andre søer opstrøms. I tilfælde af systemer, hvor vand fra én sø strømmer til andre søer nedstrøms, kan en mere omfattende model for transport og om sætning af N og P være bedre. I det omfang de potentielle punktkilder fremstår små i forhold til den samlede masse af stof, vil der dog ikke være nogen større gevinst i at forfine beskrivelsen af punktkilder.

Adskillelsen af vandafstrømningen i henholdsvis grundvandsfødt vand og overfladenært vand er en grov kategorisering, der bruges til at beskrive den store variation i vandtransportruter. Det specifikke tilfælde, hvor koncentrationsniveauet i det overfladenære vand falder i vækstsæsonen, indikerer dog realisme bag modellen. Forudsætningen om en lineær relation mellem koncentrationsniveauet af N og P og tid bygger på en formodning om, at koncentrationsudviklingen i grundvandet er relativt langsom og kun i lille omfang afhængig af forskelle i udvaskningen fra rodzonern fra år til år. Det vil dog være muligt at indføre en mere fleksibel funktionel sammenhæng, hvilket kan overvejes i fremtidige estimerter.

Det forudsættes i modellen, at næringsstofkoncentrationen i det umålte grundvand svarer til koncentrationen i det målte grundvand i tilløbet. Dette kan være kritisk for især søer med en betydelig tilstrømning af grundvand direkte gennem sør bunden, hvis dette umålte grundvand har en anden koncentration end det målte vand i tilløbet. Derfor bør massebalancen tages med forbehold for søer med stor umålt grundvandstilstrømning.

Modellen for vand og næringsstoffer bør løbende evalueres og evt. justeres i tæt synergি med modellering af vand, N og P i andre landsdækkende modeller. De landsdækkende modeller kan adskille forskellige kilder til N og P og således tage hensyn til forskellige forhold mellem målte og umålte oplande. Landsdækkende modeller er dog behæftet med forholdsvis stor usikkerhed, når de bruges lokalt for de enkelte sører. Det er således et dilemma, om det er bedre at ekstrapolere lokale målinger til umålte oplande uden hensyn til forskellige karakteristika mellem målte og umålte oplande, eller om landsdækkende modeller skal bruges til at tage hensyn til disse forskelle, selvom disse modeller er meget usikre på lokal skala og derfor risikerer at indføre en fiktiv forskel. Således foretager GEUS i 2024 en revision af den landsdækkende QNP-model, hvor modellen for tilstrømmet vand til hvert ID15-opland forbedres. Dette kan forbedre muligheden for at kunne beregne både grundvandsgeneret og overfladenært tilstrømmet vandvolumen for de enkelte ID15-oplande som delvis erstatning for den lokale vandmodel, som er beskrevet i dette notat. Når de opdaterede resultater fra dette arbejde er færdigt, bør det derfor undersøges, om nogle af de mest kritiske forudsætninger i vandmodellen kan eliminieres med virkning også for usikkerheden i stofmodellen.

5 Referencer

Jeppesen, E., Sørensen, P.B., Johansson, L.S. Søndergaard,M., Lauridsen, T., L., Nielsen, A., Mejlhede, P. , 2025: Recovery of lakes from eutrophication: changes in nitrogen retention capacity and the role of nitrogen legacy in 10 Danish lakes studied over 30 years. *Hydrobiologia* 852, 377–387 (2025). <https://doi.org/10.1007/s10750-024-05478-6>

Johansson, L.S., Søndergaard, M. & Sørensen, P.B., 2024: Sører 2022. NO-VANA. Aarhus Universitet, DCE – Nationalt Center for Miljø og Energi, 96 s. - Videnskabelig rapport nr. 591

Svendsen, L.M. & Norup, B. (eds.) 2005a: NOVANA. Nationwide Monitoring and Assessment Programme for the Aquatic and Terrestrial Environments. Programme Description – Part 1. National Environmental Research Institute, Denmark. 53 pp. – NERI Technical Report No. 532.

Svendsen, L.M., Bijl, L. van der, Boutrup, S. & Norup, B. (eds.) 2005b: NO-VANA. National Monitoring and Assessment Programme for the Aquatic and Terrestrial Environments. Programme Description – Part 2. National Environmental Research Institute, Denmark. 138 pp. – NERI Technical Report No. 537

Sørensen og Nielsen, 2023, Statistical model concept to quantify input and output of water, nitrogen and phosphorus for lakes with partly gauged watersheds, Arxiv, Cornell University: <https://doi.org/10.48550/arXiv.2312.08120>

Windolf, J., E. Jeppesen, J. P. Jensen & P. Kristensen, 1996: Modelling the seasonal variation in in-lake concentration and retention of nitrogen in 16 shallow lakes: a four-year mass balance study. *Biogeochemistry* 33: 25–44. <https://doi.org/10.1007/BF00000968>

6 Bilag

Følgende bilag svarer til dokumentation på platformen Arxiv,
(December 2023):

[arXiv:2312.08120v1](https://arxiv.org/abs/2312.08120v1)

Statistical model concept to quantify input and output of water, nitrogen and phosphorus for lakes with partly gauged watersheds

Peter Borgen Sørensen^{*1}, Anders Nielsen²

¹*Department of Ecosystem Science, Faculty of Technical Sciences, Aarhus University, C.F. Møllers Allé, Building 1120, 8000 Aarhus, Denmark.*

²*WateriTech, Døjsøvej 1, 8660 Skanderborg, Denmark*

[^{*}pbs@ecos.au.dk](mailto:pbs@ecos.au.dk)

Abstract

Valid mass load predictions of nutrients, in particular nitrogen (N) and phosphorus (P), are needed for the limnological understanding of single lake ecosystems as well as larger river/lake ecosystems. The mass of N and P that enters a lake will determine the ecological state of the lake, and the mass release from the lake will determine the ecological state of downstream ecosystems. Hence, establishing sound quantifications of the external load is crucial and e.g. contributes to the foundation of assessments of necessary management interventions to improve or preserve the ecological integrity of lakes. The external load of N and P is an integral of several pathways, each having different contributions to the total mass load. Around the world, balances of N and P have been derived for decades to support both lake water quality monitoring and research, but it can be difficult and, thus, costly to make detailed and sufficiently covering measurement campaigns in all tributaries (surface as well as groundwater) in the watershed of the N and P load including seasonality and temporal change from year to year. Thus, load prediction is facing challenge of uncertainty due to unmeasured loads, which can be a consequence of limited resources available for the water flow recordings and water concentration measurements in inlets around the lake, or simply due to invisible water flow taking place through the lake bottom. The lake outlet will typically take place in one single river, so the outlet recording seems easier to measure than inlets, however, the outlet may also have unmeasured parts in cases where water is leaching out through the lake bottom. In this paper, we propose a method that applies incomplete data sets (incomplete in the sense of temporal frequency and percentage of gauged watershed) to generate time series that predict the N and P loads entering and leaving the lake.

Key words: Lakes, ecological modelling, nitrogen, phosphorus, load predictions

Introduction

The estimation of the nutrient load to lakes of nitrogen (N) and phosphorus (P) is a fundamental task needed for developing ecological models of lake processes. The uncertainty of the load estimates directly influences the uncertainty of the lake models, and therefore the predictions of the load need to be refined as much as possible in order to ensure valid lake modeling.

Specifically, the purpose of this method is to deliver optimal use of existing data to quantify N and P concentrations of the inflowing water as well as the water volumes entering and leaving the lake to predict the load of N and P. The major challenge is how to assess the unmeasured fraction of water, N and P.

The method has been developed as a part of the Danish National Monitoring Program (NOVANA), Svendsen et al., 2005a and 2005b and the data in the case study was obtained from this program.

Methods

The principle is based on two linear model assumptions: (1) A mass balance of water, where the unmeasured fraction of water is predicted and separated into a baseflow and a surface-near flow, respectively; (2) A mass balance of N and P, where the measured concentration level is decomposed into a baseflow water concentration, a surface-near water concentration and a contribution from point sources. The predictions from these two linear models are used together with point source estimate, to predict the unmeasured load of N and P into the lake and downstream from the lake.

Water model

A mass balance of water during one month for the lake is as follows:

$$\Delta V_{i,j} = Q_{in,mea}|_{i,j} + Q_{in,unmea}|_{i,j} + A \cdot (q_{pre}|_{i,j} - q_{eva}|_{i,j}) - Q_{out,mea}|_{i,j} - Q_{out,unmea}|_{i,j} \quad (1)$$

Where

i,j : indexes of respectively year and month

$\Delta V_{i,j}$: Change in water volume in the lake for one month ($m^3/month$)

$Q_{in,mea}|_{i,j}$: Measured amount of inflow water for one month ($m^3/month$)

$Q_{in,unmea}|_{i,j}$: Unmeasured amount of inflow water for one month ($m^3/month$)

$q_{pre}|_{i,j}$: Precipitation for one month ($m/month$)

$q_{eva}|_{i,j}$: Evaporation for one month ($m/month$)

A : Lake area (m^2)

$Q_{out,mea}|_{i,j}$: Measured water outflow in the outlet of the lake for one month ($m^3/month$)

$Q_{out,unmea}|_{i,j}$: Unmeasured water outflow from the lake (leaching from bottom) for one month ($m^3/month$)

Eq. 1 can be rearranged as:

$$Q_{out,mea}|_{i,j} = A \cdot (q_{pre}|_{i,j} - q_{eva}|_{i,j}) + Q_{in,mea}|_{i,j} + Q_{in,unmea}|_{i,j} - Q_{out,unmea}|_{i,j} - \Delta V_{i,j} \quad (2)$$

The $Qout_{meal}|_{i,j}$ can be measured using a flow recorder at the outlet of the lake and the precipitation can be measured at a nearby station or estimated using methodological data. Evaporation can be measured using the energy balance of incoming energy balance. The $Qout_{unmeal}|_{i,j}$ is the unmeasured outflow that can take place as net leaching from the lake bottom. The volume of water exchange due to precipitation and evaporation is estimated as

$$Qin_{vertical}|_{i,j} = A \left(q_{pre}|_{i,j} - q_{eva}|_{i,j} \right) \quad (3)$$

Where $Qin_{vertical}|_{i,j}$ is the net volume of water supply into the lake for one month due to precipitation and evaporation. The $q_{pre}|_{i,j}$ variable is estimated using meteorological data, and $q_{eva}|_{i,j}$ is estimated using an energy balance by using methodological data from the energy radiation.

$$Qout_{meal}|_{i,j} = Qin_{vertical}|_{i,j} + Qin_{meal}|_{i,j} + Qin_{unmeal}|_{i,j} - Qout_{unmeal}|_{i,j} - \Delta V_{i,j} \quad (4)$$

There are two challenges displayed in Eq. 3. The volume change of water in the lake is rarely measured in Denmark and, thus, in praxis is unknown, the following linear reservoir model for the right-hand side of Eq. 4 is defined:

$$Qout_{meal}|_m = \sum_0^l [F_t \cdot (Qin_{meal}|_{m-t} + Qin_{unmeal}|_{m-t} + Qin_{vertical}|_{m-t})] - Qout_{unmeal}|_m \quad (5)$$

Where t is an index that counts time steps as months, and m is an index counting time steps from $i=1$ and $j=1$ as

$$m = 12 \cdot (i - 1) + j \quad (6)$$

and the linear reservoir weighting factor F_t is defined as the fraction of the inflow at month m that sequentially will contribute to the outflow the same month ($t=0$), the next month ($t=1$), the month after next month ($t=2$), etc. The index l is the number of months that are considered to have additional outflow because of an inflow at month m . Thus, the F_l factor will sum up to

$$\sum_0^l F_t = 1 \quad (7)$$

The linear reservoir model is derived based on a box model having continued input of a mass unit during a time unit, and where the output from the box is proportional to what is still left in the box, see Figure 1.

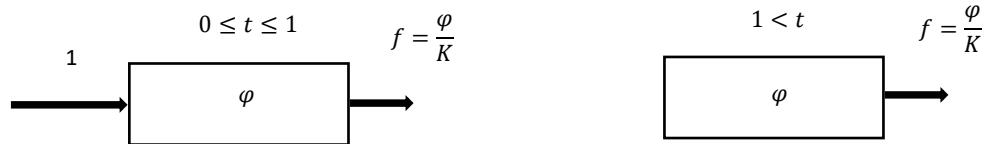


Figure 1. A linear reservoir model. To the left, the mass unit is introduced into the box continuously during a time unit. On the right, the condition for $t>1$ is shown, where the mass is released from the box proportionally to the remaining mass. Parameter K is a reservoir constant that controls the release rate and, thus, the time scale of the response for the reservoir.

$$\frac{d\varphi}{dt} = 1 - \frac{\varphi}{K}, \quad \text{for } 0 \leq t \leq 1 \quad (7)$$

$$\frac{d\varphi}{dt} = -\frac{\varphi}{K}, \quad \text{for } 1 < t \quad (8)$$

The initial condition for Eq 6 is $\varphi = 0$ for $t=0$, and the solution of Eq. 6, 7 is:

$$\varphi = K \cdot \left(1 - e^{-\frac{t}{K}}\right), \text{ for } 0 \leq t \leq 1 \quad (9)$$

$$\varphi = \varphi_{t=1} \cdot e^{-\frac{t-1}{K}}, \text{ for } 1 < t \quad (10)$$

The flux of mass out of the box (f) can now be calculated using the relationship $f = \frac{\varphi}{K}$ as

$$f = \left(1 - e^{-\frac{t}{K}}\right), \text{ for } 0 \leq t \leq 1 \quad (11)$$

$$f = \frac{\varphi_{t=1}}{K} \cdot e^{-\frac{t-1}{K}}, \text{ for } 1 < t \quad (12)$$

The mass leaving the box during the initial time unit ($0 \leq t \leq 1$) can be calculated using Eq 11 as

$$F_0 = \int_0^1 f dt = 1 + K \cdot \left(e^{-\frac{1}{K}} - 1\right) \quad (13)$$

The mass that leaves the box after the initial time unit during the period from t to $t+1$ can be derived using Eq 12 and 11 for $t>1$ as

$$F_t = \int_t^{t+1} f d\omega = K \cdot \left(1 - e^{-\frac{1}{K}}\right) \left(e^{-\frac{t-1}{K}} - e^{-\frac{t}{K}}\right) \quad (14)$$

Where F_t is the fraction of the inflow at step $m-t$ that is contributing to the outlet at step m . The principle is shown in Figure 2 for $l=10$ and two K values (0.5 and 2). Thus, if a lake is estimated to have $K=0.5$, then the major part of inflowing water will also result in an increased outflow in the same month, however, a fraction of 0.38 will lead to an increase the month after, and a small fraction will lead to increased outflow two months after. For $K=2$, the delay of the response to an input is strong, and Figure 2 shows that the increase in outflow is largest the month following the month when inflow took place. Thus, the K parameter characterizes the hydraulic property of the lake, where a lake having a large inflow of water and a small surface area will tend to have small values of K , while larger lakes having larger surface limited inflow will tend to have larger K values. The physical property of the outlet will also play a role, where a wide outlet will tend to keep the lake volume constant and therefore result in a small value of K in contrast to a narrower outlet.

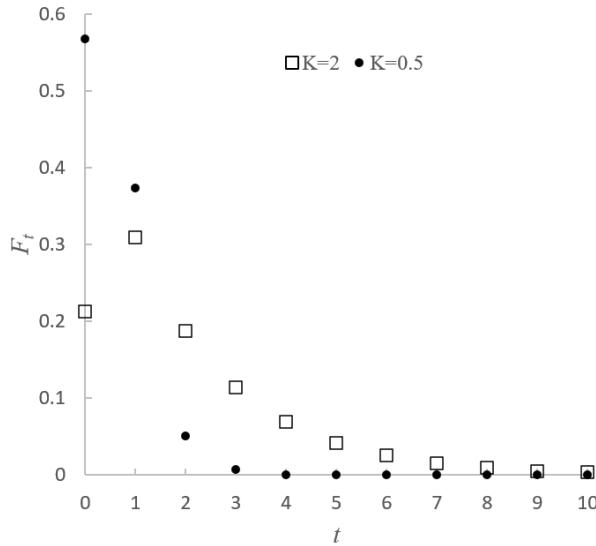


Figure 2. Numerical example for $l=10$ and the K values 0.5 and 2.

The unmeasured part of the water volume needs to be estimated in Eq. 5, and for this purpose the following linear model is used:

$$Qin_{unmeal}|_m = \alpha_n \cdot Qin_{meal}|_m + \beta 1_i \cdot \gamma_j \quad (15)$$

Where α_n is the slope between the measured and unmeasured flow for the n^{th} combination of measuring station location. In some cases, the locations of the water flow stations are changed and a slope needs to be estimated for each new location in order to take into account that this may influence the fraction of unmeasured water. The parameter $\beta 1_i$ estimates a yearly water volume entering the lake during the year i without having any relation to $Qin_{meal}|_m$. This could be ground water surplus to the lake through the bottom or from aquifers adjacent to the lake perimeter. The parameter γ_j describes the seasonal distribution of this surplus water volume due to seasonality of the ground water table. Thus, the following constrain applies to γ_j :

$$\sum_{j=1}^{12} \gamma_j = 1 \quad (16)$$

The unmeasured outflow of water is assumed to be seepage through the lake bottom, as lakes rarely have more than one outlet at the surface. This seepage is assumed to be constant during the year, but may differ between years. The argument is that the seepage will tend to be rather constant due to the relatively fixed water table in the lake, however, during a longer period of several years it may change. So, the unmeasured outflow is estimated as a constant each year:

$$Qout_{unmeal}|_m \approx \beta 2_i, \beta 2_i \geq 0 \quad (17)$$

The sign $\beta 2_i$ is restricted to conditions where unmeasured water is leaving the lake. Eqs. 5, 15 and 17 are combined in the statistical model for the measured outflow as:

$$Qout_{meal}|_m \approx \sum_0^l [F_t \cdot ((1 + \alpha_n) \cdot Qin_{meal}|_m + \beta 1_i \cdot \gamma_j)] - \beta 2_i, \beta 1_i, \beta 2_i \geq 0 \wedge [\beta 1_i > 0 \rightarrow \beta 2_i = 0] \quad (18)$$

Where the linear reservoir factor F_t is estimated using Eq 13 and 14. Eq. 18 is used in a Maximum Likelihood estimate of the parameters. However, for practical fitting there is a challenge using Eq. 18

without including the constraints for $\beta 1_i$ and $\beta 2_i$. This is because the value of respectively $\beta 1_i$ and $\beta 2_i$ can compensate for each other such that an increase in the $\beta 1_i$ value can be compensated by a similar decrease in the $\beta 2_i$ value. The interpretation of this phenomenon is that if constant unmeasured water flow into and out of the lake exist simultaneously, then this is hidden for the mass balance governing the model.

The model assumes that the month having the least measured inflow of water is the month where the contribution from the surface-near water flow is negligible as an inflow contributor, this month is denoted the reference month for the specific year. Thus, every year has a reference month that may be different from year to year, denoted j_{ref_i} . The measured baseflow as a fraction of the inflow for each month is estimated based on the reference month:

$$Qin_{meabase}|_{i,j} \approx Qin_{mea}|_{j_{ref_i}} \quad (19)$$

The measured surface-near water flow is now predicted as:

$$Qin_{measurface}|_{i,j} \approx Qin_{mea}|_{i,j} - Qin_{meabase}|_{i,j} \quad (20)$$

The unmeasured base-flow entering the lake is predicted as:

$$Qin_{unmeabase}|_{j,i} \approx \alpha_n \cdot Qin_{meabase}|_{i,j} + \beta 1_i \cdot \gamma_j \quad (21)$$

In Eq. 21, it is assumed that the volume of water entering the lake independently of the measured inlet volume is a baseflow. The argument is that unmeasured surface-near inflow correlates close to the variation in measured inflow because the surface-near water tends to induce large variations in the inflowing water volume, being a short term respond to rainfall that often takes place in the entire large catchment and, thus, is effective also on unmeasured areas. The unmeasured surface-near flow entering the lake is thus predicted as:

$$Qin_{unmeasurface}|_{j,i} \approx \alpha_n \cdot Qin_{measurface}|_{i,j} \quad (22)$$

The total baseflow entering the lake can now be predicted as

$$Qin_{base}|_{j,i} \approx Qin_{meabase}|_{i,j} + Qin_{unmeabase}|_{i,j} \quad (23)$$

The total flow of water from surface-near sources is predicted as

$$Qin_{surface}|_{j,i} \approx (1 + \alpha_n) \cdot Qin_{measurface}|_{i,j} \quad (24)$$

Eq. 15 assumes that the fraction of baseflow in unmeasured inflow is similar to the fraction of base flow in the measured inflow. The total waterflow out of the lake is predicted as

$$Qout|_{j,i} = Qin_{surface}|_{j,i} + Qin_{base}|_{j,i} \quad (25)$$

N and P model

The load of N and P is divided in two categories: a measured part and an unmeasured part. The measured part is predicted as the measured concentration level in the measured water in the inflow, and the unmeasured part is predicted using a model as described in the following. A mass balance is established for the measured transport of N and P that divides the total measured load into three transport routes: (1) $Qin_{measurface}|_{i,j}$; (2) $Qin_{meabase}|_{i,j}$; (3) N and P entering from point sources in the measured part of the lake watershed ($P_{effmea}|_{i,j}$) as

$$\begin{aligned} Load_{i,j} &\approx Cin_{tot,mea}|_{i,j} \cdot Qin_{mea}|_{i,j} + Cin_{meabase}|_{i,j} \cdot Qin_{meabase}|_{i,j} + Cin_{unmeasurf}|_{i,j} \cdot \\ &Qin_{unmeasurf}|_{i,j} + P_{unmea}|_{i,j} \end{aligned} \quad (26)$$

The unmeasured concentrations are assumed equal to the measured concentration for the base flow and surface-near flow, respectively:

$$Cin_{unmeabase}|_{i,j} \approx Cin_{meabase}|_{i,j}, \quad Cin_{unmeasurf}|_{i,j} \approx Cin_{measurf}|_{i,j} \quad (27)$$

Where $Cin_{meabase}|_{i,j}$ is the base-flow concentration in the measured inlets and $Cin_{measurf}|_{i,j}$ is the surface-near concentration in the measured inlet water. The measured concentration will obviously only be a total bulk concentration of the mixed inlet water between base-flow and surface-near water. Hence, the base-flow and surface-near flow components need to be estimated in the total inlet concentration of the mixed water.

The total mass balance provides:

$$\begin{aligned} Cin_{tot,mea}|_{i,j} \cdot Qin_{mea}|_{i,j} &= Cin_{meabase}|_{i,j} \cdot Qin_{meabase}|_{i,j} + Cin_{measurf}|_{i,j} \cdot Qin_{measurf}|_{i,j} + \\ P_{effmea}|_{i,j} \end{aligned} \quad (28)$$

where $Cin_{tot,mea}|_{i,j}$ is the total bulk concentration measured in the inlet to the lake in the same water volume that was measured as $Qin_{mea}|_{i,j}$, the concentration in respectively base-flow and surface-near flow is $Cin_{meabase}|_{i,j}$ and $Cin_{measurf}|_{i,j}$. The effective contribution from point sources in the inlet is added as $P_{effmea}|_{i,j}$ and defined as the contribution form point sources that have discharged to the measured inlet.

It is not possible to measure the concentration levels separately in the baseflow and in the surface-near flow, as the water is mixed up in the streams. However, these concentration levels can be predicted using parameters estimated by a statistical model as follows, where Eq 29 is rearranged to form:

$$Cin_{tot,mea}|_{i,j} = Cin_{meabase}|_{i,j} \cdot \frac{Qin_{meabase}|_{i,j}}{Qin_{mea}|_{i,j}} + Cin_{measurf}|_{i,j} \cdot \frac{Qin_{measurf}|_{i,j}}{Qin_{mea}|_{i,j}} + \frac{P_{effmea}|_{i,j}}{Qin_{mea}|_{i,j}} \quad (29)$$

And this equation can be rewritten as

$$\begin{aligned} Cin_{totmea}|_{i,j} &= Cin_{meabase}|_{i,j} \cdot (1 - X_{i,j}) + Cin_{measurf}|_{i,j} \cdot X_{i,j} + \frac{P_{effmea}|_{i,j}}{Qin_{mea}|_{i,j}} \quad \text{where } X_{i,j} = \\ \frac{Q_{measurf}|_{i,j}}{Qin_{mea}|_{i,j}} \end{aligned} \quad (30)$$

And further rewritten to the form:

$$Cin_{totmea}|_{i,j} = Cin_{meabase}|_{i,j} + (Cin_{measurf}|_{i,j} - Cin_{meabase}|_{i,j}) \cdot X_{i,j} + \frac{P_{effmea}|_{i,j}}{Qin_{mea}|_{i,j}} \quad (31)$$

The quantification of effective point sources' contribution to the load by measured inlet water will typically be uncertain, both due to the quantification of the point source strength, where the discharge takes place, and the retention from the position of the point sources' discharge to the point of measurement at the lake inlet. Thus, a parameter ρ is defined as an effective point source factor as

$$P_{effmea}|_{i,j} = \rho \cdot P_{mea}|_{i,j} \quad (32)$$

Under the condition that the local point source strength is known the Eq. 15 has the following form, where the depended variable is:

$$Y_{i,j} = \theta o + \theta 1 \cdot t + (\gamma 1_i + \gamma 2_j) \cdot X_{i,j} + \rho \cdot \frac{P_{meal}|_{i,j}}{Q_{in,meal}|_{i,j}} \quad (33)$$

Where t is time since the first measurement, and ρ is the measured impact from point sources.

Comparing Eq. 31 and 33, it is seen that the measured baseflow concentration can be predicted as

$$C_{in,meabase}|_{i,j} \approx \theta o + \theta 1 \cdot t \quad (34)$$

And the measured surface-near fraction can be predicted as

$$C_{in,measurf}|_{i,j} \approx C_{in,meabase}|_{i,j} + \gamma 1_i + \gamma 2_j \quad (35)$$

Consequently, the base-flow concentration is assumed to change gradually with time as a linear and second order relation. The argument is that the base-flow concentration is related to ground water that typically takes several years to form and, thus, is attributed to a sluggish change during time. The surface-near concentration is more dynamic, both in relation to the season and from one year to another year.

As seen in Eq. 31, the point sources are included as part of the dependent variable Y. The contribution from these point sources is not directly measurable and must therefore be estimated as a contribution upstream in the catchment to a given measuring station.

The predicted concentration levels from Eq. 25 and 26 are latent in the way that they are not measurable quantities that can be validated. Thus, the predictions from Eq. 25 and 26 may be adjusted or replaced using expert judgement, if the values seem unrealistic and local knowledge about proper value intervals exists.

Load prediction

The predicted water volumes and concentration levels are multiplied to predict the mass load in and out the lake:

$$\begin{aligned} Load_{in}|_{i,j} &\approx C_{in,tot,mea}|_{i,j} \cdot Q_{in,mea}|_{i,j} + C_{in,meabase}|_{i,j} \cdot Q_{in,unmeabase}|_{j,i} + C_{in,measurf}|_{i,j} \cdot \\ &Q_{in,unmeasurf}|_{j,i} + \rho \cdot P_{unmea}|_{i,j} + At|_{i,j} \end{aligned} \quad (36)$$

Where $P_{unmea}|_{i,j}$ is the contribution from unmeasured point sources and $At|_{i,j}$ is the atmospheric deposition on the lake surface. The load out of the lake is estimated as

$$Load_{out}|_{i,j} \approx C_{out,mea}|_{i,j} \cdot Q_{out}|_{j,i} \quad (37)$$

Where $C_{out,mea}|_{i,j}$ is the measured concentration in the outlet.

Results

Case study Bryrup Langsø

The Danish Lake Bryrup Langsø was selected as an illustrative case. The lake is a small to middle sized shallow lake covering an area of 37 ha and with a mean depth of 4.6 m. The inlet measuring stations are located along the major inlet, but the station was relocated between 1995 and 1996, so the fraction of measured water may have changed as a consequence. Hence, the model concept introduced in this paper distinguishes between these two inlet stations ($n=2$). Three smaller untagged streams enter the lake, as seen in Figure 1, so this indicates that the model should identify a fraction unmeasured water flow into the lake.

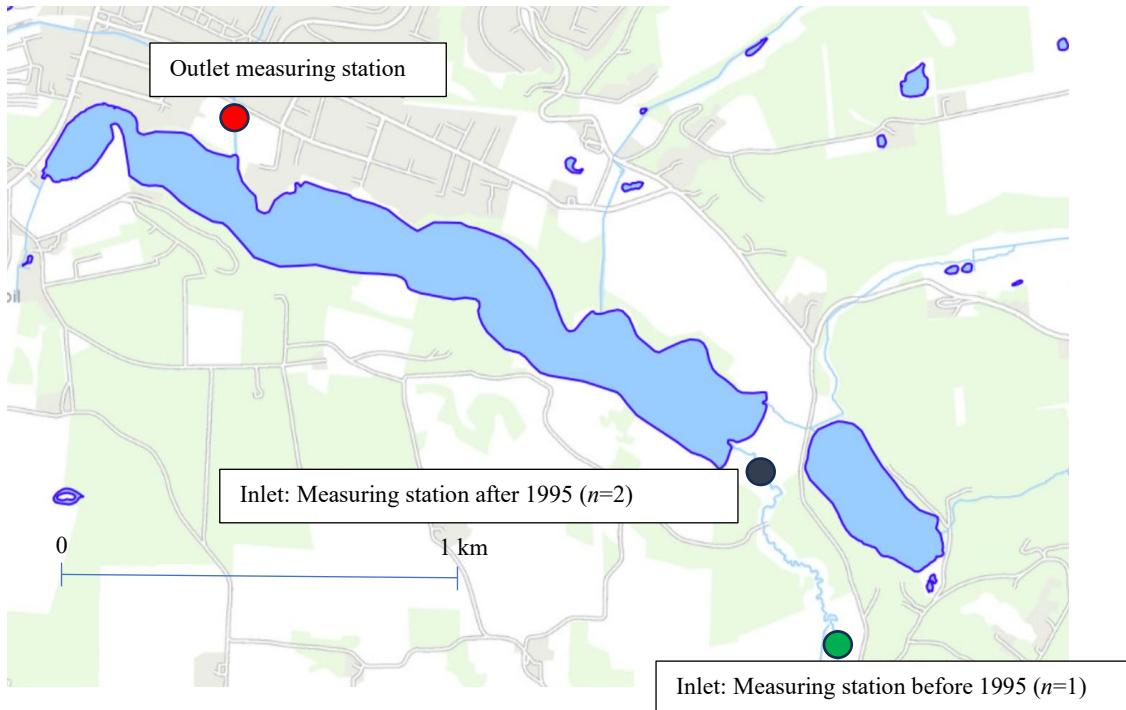


Figure 1. The overview of Bryrup Langsø case study. The lake is shaped as a rift formed by the ice age. The inlet stations are located at the main stream entering at the eastern end, and the outlet is located at the western end.

The data are collected at the stations as a part of the Danish National Monitoring Program (NOVANA), Svendsen et al., 2005a and 2005b, and the data in the case study are obtained from this program. The water flow is predicted using local calibrated correlations between water flow and depth. This results in continuous water flow prediction that is integrated to predict the water volume passing the cross section during one month. The total concentrations of N and P have been measured during the years, typically one or two bulk water samples per month. The measured concentration levels have been linear interpolated, and the mean value is multiplied with the paired water volume to predict the actual mass of transported matter for one month. Monthly data was collected in the period 1990-2020, as listed in Table A1, see appendix.

Water model

The coefficients in the linear model of the water balance model (Eq. 4) are estimated in Table 1.

coefficient	Estimate	unit	coefficient	Estimate	unit	coefficient	Estimate	unit
β_{1_1}	900000	m^3	$\beta_{1_{28}}$	1700000	m^3	$\beta_{1_{23}}$	0	$m^3/month$
β_{1_2}	1040000	m^3	$\beta_{1_{29}}$	1600000	m^3	$\beta_{1_{24}}$	0	$m^3/month$
β_{1_3}	900000	m^3	$\beta_{1_{30}}$	2120000	m^3	$\beta_{1_{25}}$	0	$m^3/month$
β_{1_4}	860000	m^3	$\beta_{1_{31}}$	2440000	m^3	$\beta_{1_{26}}$	0	$m^3/month$
β_{1_5}	1660000	m^3	$\beta_{1_{32}}$	1720000	m^3	$\beta_{1_{27}}$	0	$m^3/month$
β_{1_6}	3660000	m^3	β_{2_1}	0	$m^3/month$	$\beta_{1_{28}}$	0	$m^3/month$
β_{1_7}	920000	m^3	β_{2_2}	0	$m^3/month$	$\beta_{1_{29}}$	0	$m^3/month$
β_{1_8}	560000	m^3	β_{2_3}	0	$m^3/month$	$\beta_{1_{30}}$	0	$m^3/month$
β_{1_9}	1180000	m^3	β_{2_4}	0	$m^3/month$	$\beta_{1_{31}}$	0	$m^3/month$
$\beta_{1_{10}}$	720000	m^3	β_{2_5}	0	$m^3/month$	$\beta_{1_{32}}$	0	$m^3/month$
$\beta_{1_{11}}$	1620000	m^3	β_{2_6}	0	$m^3/month$	γ_1	0.073	1/month
$\beta_{1_{12}}$	480000	m^3	β_{2_7}	0	$m^3/month$	γ_2	0.098	1/month
$\beta_{1_{13}}$	1360000	m^3	β_{2_8}	0	$m^3/month$	γ_3	0.134	1/month
$\beta_{1_{14}}$	1280000	m^3	β_{2_9}	0	$m^3/month$	γ_4	0.110	1/month

$\beta_{1_{15}}$	1360000	m^3	$\beta_{2_{10}}$	0	$m^3/month$	γ_5	0.110	1/month
$\beta_{1_{16}}$	1700000	m^3	$\beta_{2_{11}}$	0	$m^3/month$	γ_6	0.098	1/month
$\beta_{1_{17}}$	980000	m^3	$\beta_{2_{12}}$	0	$m^3/month$	γ_7	0.073	1/month
$\beta_{1_{18}}$	540000	m^3	$\beta_{2_{13}}$	0	$m^3/month$	γ_8	0.073	1/month
$\beta_{1_{19}}$	920000	m^3	$\beta_{2_{14}}$	0	$m^3/month$	γ_9	0.061	1/month
$\beta_{1_{20}}$	620000	m^3	$\beta_{2_{15}}$	0	$m^3/month$	γ_{10}	0.049	1/month
$\beta_{1_{21}}$	100000	m^3	$\beta_{2_{16}}$	0	$m^3/month$	γ_{11}	0.061	1/month
$\beta_{1_{22}}$	760000	m^3	$\beta_{2_{17}}$	0	$m^3/month$	γ_{12}	0.061	1/month
$\beta_{1_{23}}$	1060000	m^3	$\beta_{2_{18}}$	0	$m^3/month$	α_1	0.20	-
$\beta_{1_{24}}$	1280000	m^3	$\beta_{2_{19}}$	0	$m^3/month$	α_2	0.21	-
$\beta_{1_{25}}$	880000	m^3	$\beta_{2_{20}}$	0	$m^3/month$	$K:$	0.01	-
$\beta_{1_{26}}$	1360000	m^3	$\beta_{1_{21}}$	0	$m^3/month$			
$\beta_{1_{27}}$	1800000	m^3	$\beta_{1_{22}}$	0	$m^3/month$			

Table 1. Estimated values for the coefficients of the water balance model (Eq. 4) for the lake Bryrup Langsø. $R^2=0.95$.

The β_{1_i} values are total volume of water entering the lake during one year, independent of the variation of the measured inflow. The γ_j values distribute β_{1_i} between the months. The water model interprets β_{1_i} as a base flow, so the value should have correlation to the precipitation the year before, and this is investigated for Bryrup Langsø in Table 2, where β_{1_i} is used as a response variable for the linear combination of the yearly precipitation (mm/year), respectively the same year ($Prec_i$), the former year ($Prec_{i-1}$) and the year before the former year ($Prec_{i-2}$). If the intercept is related to the base-flow into the lake, then the intercept in January should be correlated to the precipitation the year up to January (year number $i-1$) and not the year that starts in January (year number i). This is seen to be the case in Table 2, where only the precipitation the year before the intercept is seen to be significant, and the year before the former year is not significant ($Prec_{i-2}$).

Response variable: β_{1_i}					
Prediction variables	Sum square	Df	F	Pr (>F)	
$Prec_i$	$91 \cdot 10^6$	1	0.031	0.86	
$Prec_{i-1}$	$17899 \cdot 10^6$	1	6.147	0.02	
$Prec_{i-2}$	$416 \cdot 10^6$	1	0.143	0.71	
Residuals	$75714 \cdot 10^6$	26			

Table 2. ANOVA table testing the influence of different years' precipitation on the response variable β_i

The intercept variation between months (γ_i) is shown during the year in Figure 1. Bryrup Langsø is in a climatic zone having a seasonality, where the evaporation from the soil surface and vegetation is high during the summer period, such that the formation of surface-near water is limited in this part of the season. While the winter period is characterized by having much more precipitation than evaporation. Thus, the variation in Figure 1 is expected, where the baseflow to the lake is increasing during the winter period build up, having a top value in spring and followed by a decrease during the summer period.

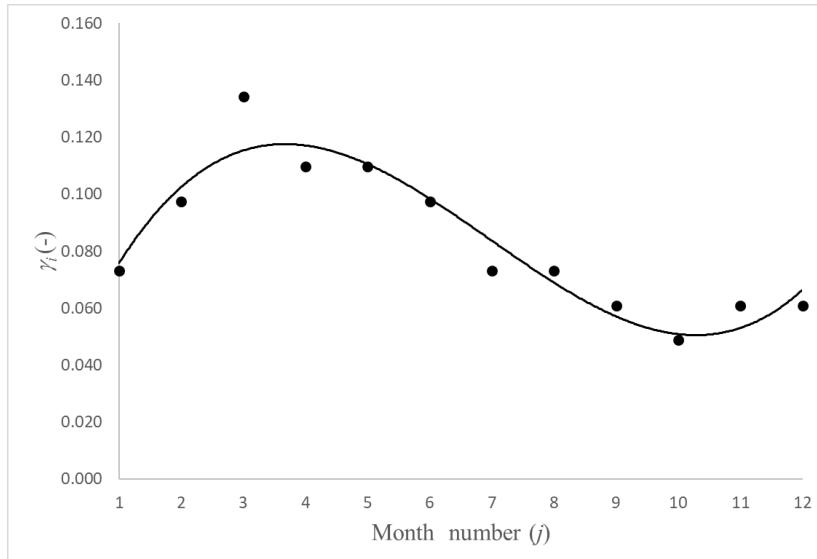


Figure 1. The variation in incept (γ_i) during the year.

The slopes (α) in Table 1 are estimated for two different gauge locations corresponding to Table A1 in appendix A, where it is seen that station location “d” is changed to “e” from year 1995 to 1996. This is because the location of the inlet station was moved, so the fraction of respectively measured and unmeasured water in the inlet may also have changed. However, the slope value of $\alpha_1 = 0.20$ is nearly similar to the value of $\alpha_2 = 0.21$, so the change of station location does not indicate a considerable change in the fraction of unmeasured water. It is also seen that the value $K=0.01$ discloses a limited hydraulic memory effect from the former month, which, again, shows a high hydraulic load of water. The water volume inflow during one month divided by lake area is typically 1-3 m, indicating a high hydraulic load level on the lake and, thus, a highly limited hydraulically memory from the former month.

The general water volume budget is disclosed in Table 3. The unmeasured outflow ($Q_{out,unmea}|_{i,j}$) is zero for the entire data set, so the lake does not have net release of water at any time, however, it is impossible to identify conditions where a simultaneous inflow and out flow exist through the lake bottom. The unmeasured base-flow can either come from wells outside the lake or form wells directly flowing into the lake from the lake bottom.

Station location (n)	Mean monthly volume inflowing water (1000 m ³ /month)			
	Reservoir type	Measured	Unmeasured	Total
1	Base flow	200	129	330
	Surface-near flow	291	58	350
	Total	492	188	679
2	Base flow	214	152	366
	Surface-near flow	225	47	273
	Total	439	199	638

Table 3. The mean monthly inflow exclusive $Q_{in,vertical}|_{i,j}$ is divided into respectively base-flow and surface-near flow and measured and unmeasured volumes. The inlet measuring station was relocated at the end of 1995, so $n=1$ covers the measurement before 1996 (1990-1995), and $n=2$ includes the measurements from year 1996 (1996-2021).

N Model

The N concentration in the inlet is analyzed using equations 24-26. The coefficient table for Eq. 23 is shown in Table 3. The point source factor ρ was estimated to be negative and non-significantly

different from zero. Thus, the model could not detect any influence to the inlet due to point sources. In Denmark, the point sources for N are typically small compared to the leaching from soil, so a non-significant contribution from point sources is not surprising. This coincides with the comparison between the point source strength in Table A1 for measured catchment that has a level in the interval 50-100 kg/month, with the load from measured water in Figure 5, where the levels are much higher, at 1500-12000 kg/month.

coefficient	Estimate	unit	coefficient	Estimate	unit	coefficient	Estimate	unit
γ_{1_1}	0	mg/l	$\gamma_{1_{17}}$	-1.51	mg/l	γ_{2_1}	1.24	mg/l
γ_{1_2}	0.15	mg/l	$\gamma_{1_{18}}$	-2.05	mg/l	γ_{2_2}	1.00	mg/l
γ_{1_3}	3.21	mg/l	$\gamma_{1_{19}}$	-2.25	mg/l	γ_{2_3}	0.29	mg/l
γ_{1_4}	0.95	mg/l	$\gamma_{1_{20}}$	-0.88	mg/l	γ_{2_4}	-0.12	mg/l
γ_{1_5}	-0.59	mg/l	$\gamma_{1_{21}}$	-1.21	mg/l	γ_{2_5}	-0.39	mg/l
γ_{1_6}	-0.66	mg/l	$\gamma_{1_{22}}$	-2.44	mg/l	γ_{2_6}	-1.47	mg/l
γ_{1_7}	2.79	mg/l	$\gamma_{1_{23}}$	-2.17	mg/l	γ_{2_7}	-1.54	mg/l
γ_{1_8}	1.82	mg/l	$\gamma_{1_{24}}$	-1.81	mg/l	γ_{2_8}	-1.56	mg/l
γ_{1_9}	2.33	mg/l	$\gamma_{1_{25}}$	-2.68	mg/l	γ_{2_9}	-0.91	mg/l
$\gamma_{1_{10}}$	0.07	mg/l	$\gamma_{1_{26}}$	-2.02	mg/l	$\gamma_{2_{10}}$	0.05	mg/l
$\gamma_{1_{11}}$	-0.81	mg/l	$\gamma_{1_{27}}$	-2.12	mg/l	$\gamma_{2_{11}}$	1.37	mg/l
$\gamma_{1_{12}}$	-0.52	mg/l	$\gamma_{1_{28}}$	-1.26	mg/l	$\gamma_{2_{12}}$	1.36	mg/l
$\gamma_{1_{13}}$	-1.30	mg/l	$\gamma_{1_{29}}$	-0.92	mg/l	θ_0	8.61	mg/l
$\gamma_{1_{14}}$	0.76	mg/l	$\gamma_{1_{30}}$	0.18	mg/l	θ_1	-0.0050	mg/(l·t)
$\gamma_{1_{15}}$	-0.13	mg/l	$\gamma_{1_{31}}$	-2.02	mg/l	ρ	0	-
$\gamma_{1_{16}}$	-1.10	mg/l	$\gamma_{1_{32}}$	0.66	mg/l			

Table 4. Estimated parameters for Eq. (24), $R^2=0.70$.

The time series of predicted concentrations is shown in Figure 2 together with the measured values subtracted point source estimates. Some extreme values, especially some higher levels, seem poorly described by the model.

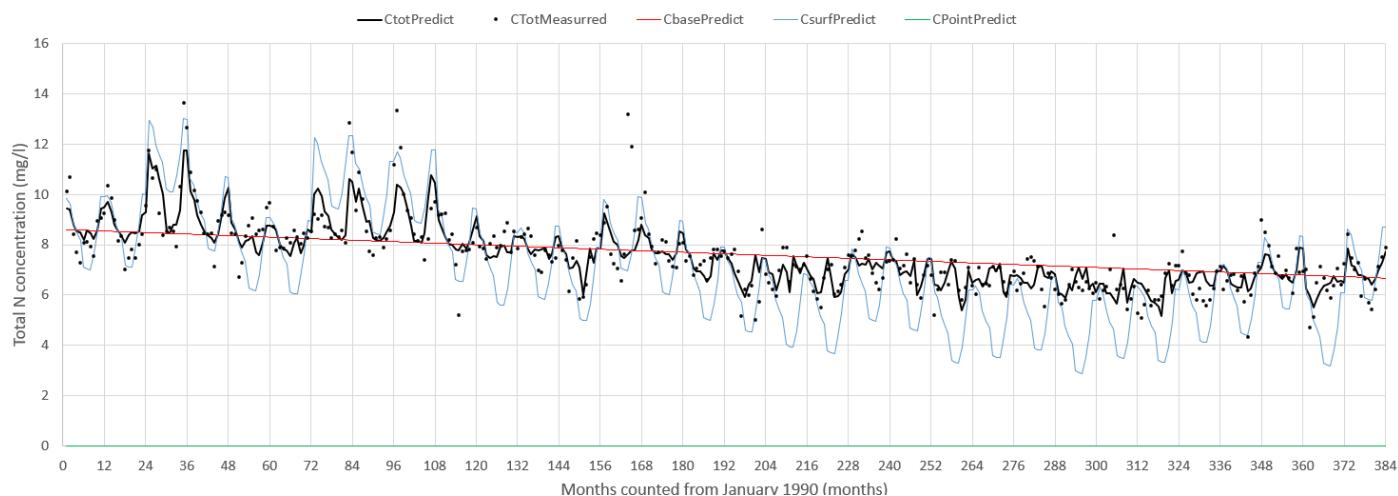


Figure 2. A time series of predicted (solid black curve) and measured inlet concentrations (circles) for N. The predicted surface water concentration is shown as a blue line, and the base-flow concentration as a red line.

The values from Table 3 are used in Eq. 25 and 26 to predict respectively $Cin_{meabase}|_{i,j}$ and $Cin_{measurf}|_{i,j}$. The seasonal relation is shown in Figure 2 for 1990. $Cin_{meabase}|_{i,j}$ does not change much during one single year, so the dashed curve of $Cin_{meabase}|_{i,j}$ is nearly constant in Figure 2. The N concentration for $Cin_{measurf}|_{i,j}$ is likely to show variations during the year due to the effects of the growing season, having increased plant uptake in the spring and summer season leading to a drop in the $Cin_{measurf}|_{i,j}$ value.

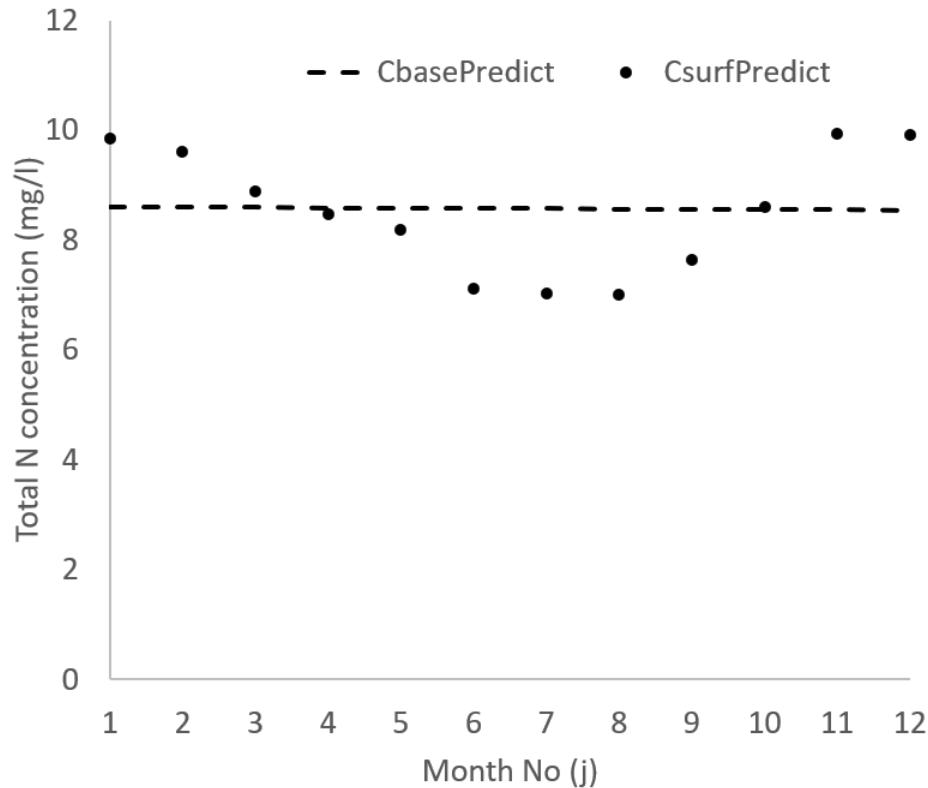


Figure 3. The variation of predicted concentration values during 1990, where CbasePredict is the predicted base water concentration, and CsrfPredict is the predicted surface-near water concentration.

The long-term variation in concentration levels is shown in Figure 3 for the winter flow concentration (January, $j=1$). The base-flow concentration shows a weak decrease during the period, while the surface-near concentration primarily shows a rapid decrease in the period 2000-2005 from a relatively high concentration level to a lower level.

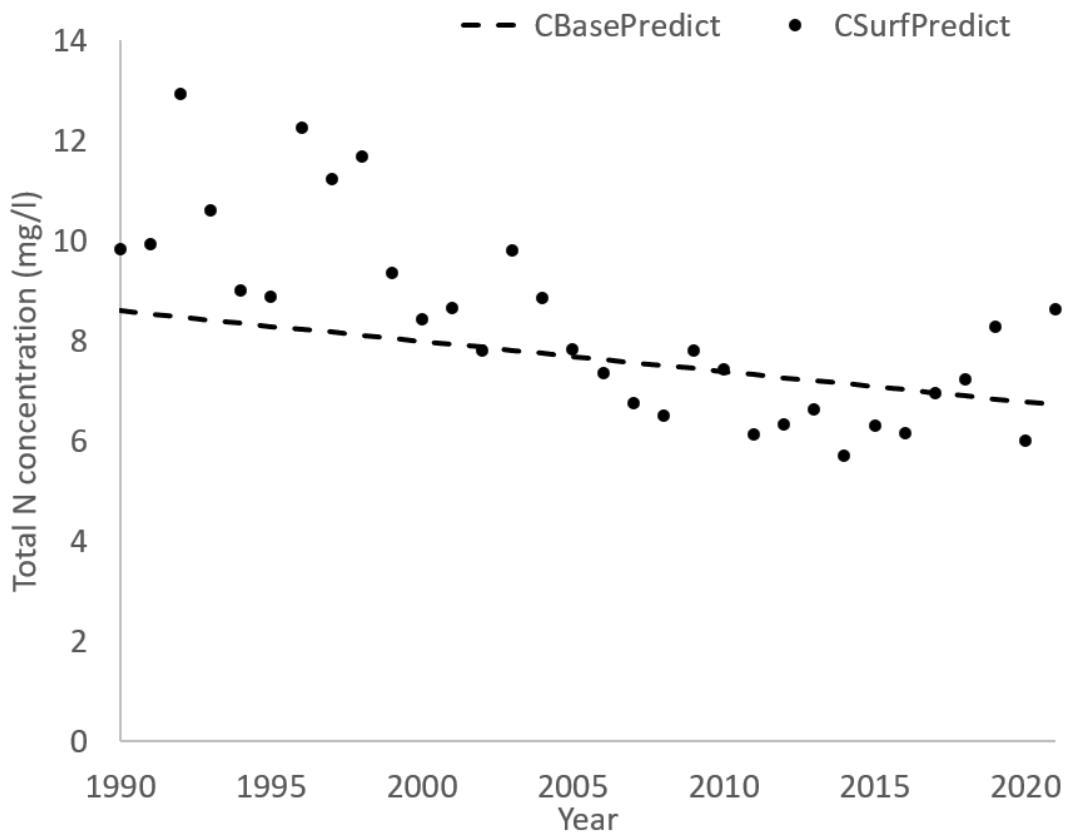


Figure 4. Long-term time trend of the N concentration (January every year), where CbasePredict is the predicted base water concentration and CsurfPredict is the predicted surface-near water concentration.

The purpose of the statistical model is primarily to predict the diffuse source load of N due to unmeasured fractions of water bodies, however, the application of the model can be evaluated by considering the capability of the model to predict the load for the measured fraction of the water bodies. Such an evaluation is shown in Figure 5.

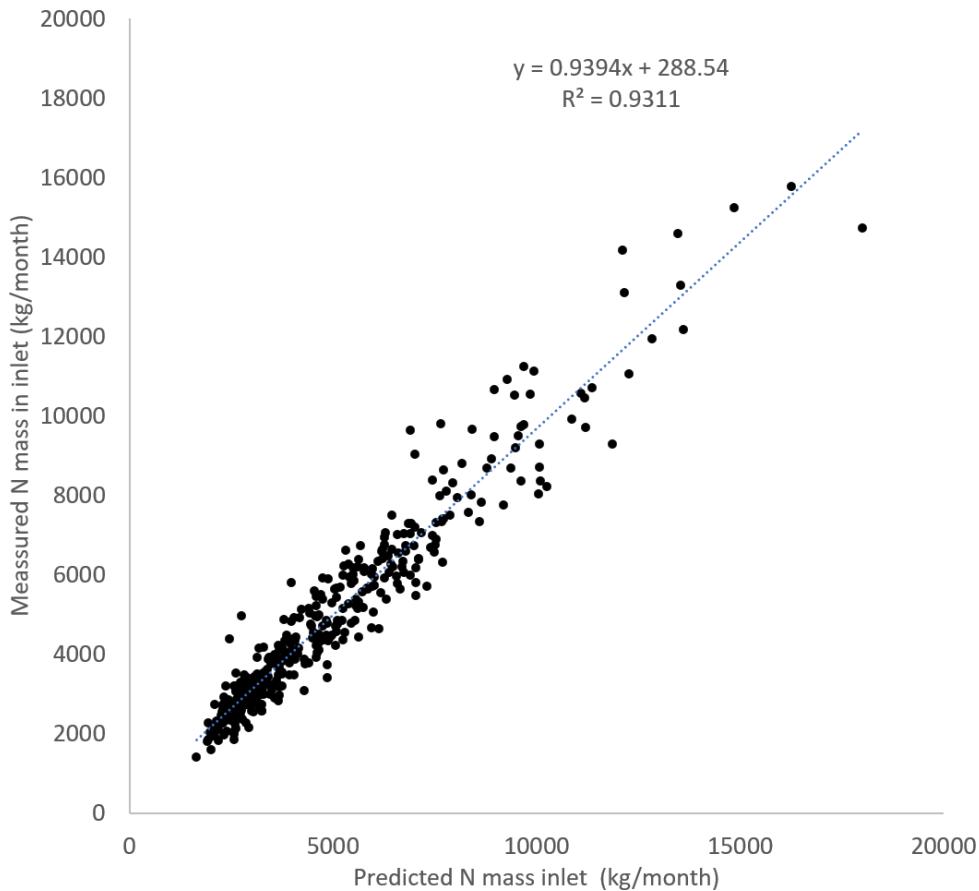


Figure 5. Comparing the predicted diffuse load of N with the measured diffuse load.

The total load in and out of the lake is shown in Figure 5. The final result for N is the predicted input and output of N for Bryrup Langsø, as shown in Figure 6, where the measured load is added to the predicted unmeasured load for each month in the time series. The inlet load is seen to be substantially larger than the outlet, disclosing a retention of N in the lake as expected. On average, the lake is removes 2500 kg N from the fresh water ecosystem every month.

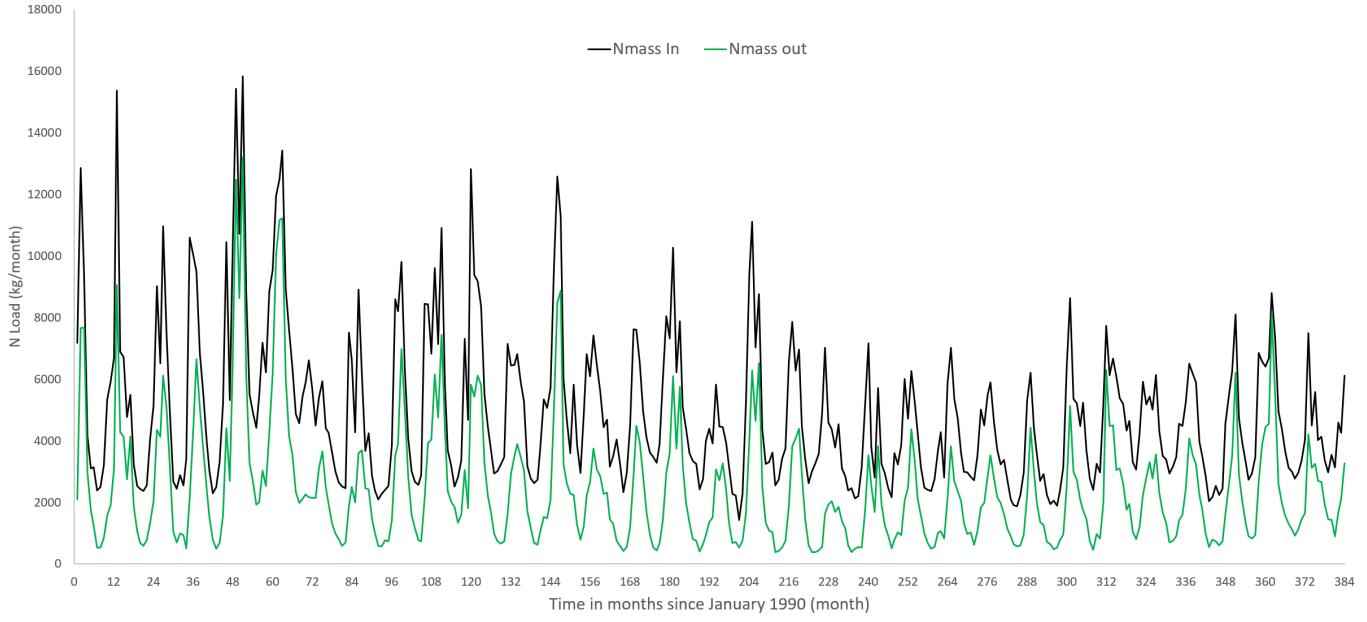


Figure 6. The total predicted load of N in and out of the Bryrup Langsø.

P Model

The P concentration in the inlet is analyzed using the equations 24-26. The coefficient table for Eq. 24 is shown in Table 5. The point source factor in Table 5 is estimated to be 0.1056, which indicates some influence due to point sources.

coefficient	Estimate	unit	coefficient	Estimate	unit	coefficient	Estimate	unit
γ_{1_1}	0	mg/l	$\gamma_{1_{17}}$	0.004	mg/l	γ_{2_1}	0.081	mg/l
γ_{1_2}	-0.039	mg/l	$\gamma_{1_{18}}$	-0.016	mg/l	γ_{2_2}	0.078	mg/l
γ_{1_3}	-0.024	mg/l	$\gamma_{1_{19}}$	-0.065	mg/l	γ_{2_3}	0.068	mg/l
γ_{1_4}	-0.021	mg/l	$\gamma_{1_{20}}$	-0.072	mg/l	γ_{2_4}	0.040	mg/l
γ_{1_5}	-0.044	mg/l	$\gamma_{1_{21}}$	-0.039	mg/l	γ_{2_5}	0.054	mg/l
γ_{1_6}	-0.093	mg/l	$\gamma_{1_{22}}$	-0.035	mg/l	γ_{2_6}	0.085	mg/l
γ_{1_7}	-0.023	mg/l	$\gamma_{1_{23}}$	-0.068	mg/l	γ_{2_7}	0.103	mg/l
γ_{1_8}	-0.095	mg/l	$\gamma_{1_{24}}$	-0.080	mg/l	γ_{2_8}	0.130	mg/l
γ_{1_9}	-0.007	mg/l	$\gamma_{1_{25}}$	-0.069	mg/l	γ_{2_9}	0.109	mg/l
$\gamma_{1_{10}}$	-0.053	mg/l	$\gamma_{1_{26}}$	-0.049	mg/l	$\gamma_{2_{10}}$	0.099	mg/l
$\gamma_{1_{11}}$	-0.074	mg/l	$\gamma_{1_{27}}$	-0.064	mg/l	$\gamma_{2_{11}}$	0.080	mg/l
$\gamma_{1_{12}}$	-0.003	mg/l	$\gamma_{1_{28}}$	-0.065	mg/l	$\gamma_{2_{12}}$	0.074	mg/l
$\gamma_{1_{13}}$	-0.060	mg/l	$\gamma_{1_{29}}$	-0.079	mg/l	θ_o	0.09647	mg/l
$\gamma_{1_{14}}$	-0.088	mg/l	$\gamma_{1_{30}}$	-0.058	mg/l	θ_1	-0.00011	mg/(lt)
$\gamma_{1_{15}}$	-0.042	mg/l	$\gamma_{1_{31}}$	-0.037	mg/l	ρ	0.016	-
$\gamma_{1_{16}}$	-0.049	mg/l	$\gamma_{1_{32}}$	-0.068				

Table 5. Estimated parameters for Eq. (23), $R^2=0.49$.

The time series of predicted concentration is show in Figure 7 together with the measured values subtracted point source estimates. Some extreme values, especially some higher levels, seem poorly described by the model.

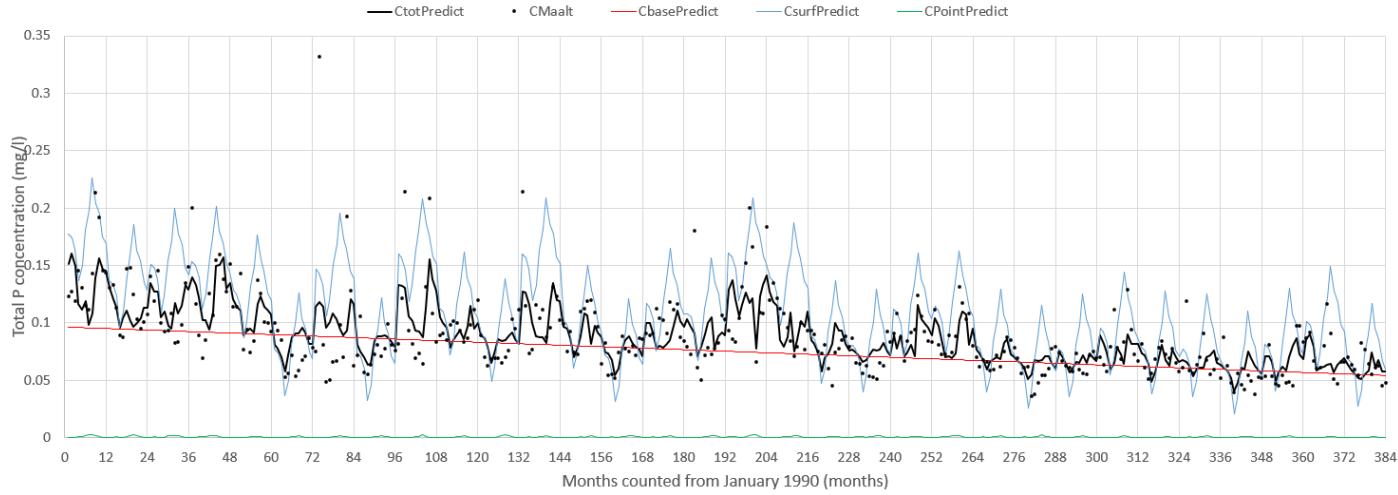


Figure 7. A time series of predicted (solid black curve) and measured inlet concentrations (circles) of P. The predicted surface water concentration is shown as a blue line and the base-flow concentration as a red line.

The values from Table 5 are used in Eqs. 24 and 25 to predict respectively $Cin_{meabase}|_{i,j}$ and $Cin_{measurf}|_{i,j}$. The seasonal relation is shown in Figure 2 for 1990. $Cin_{meabase}|_{i,j}$ does not change much during one single year, so the dashed curve of $Cin_{meabase}|_{i,j}$ is nearly constant in Figure 2. The P concentration for $Cin_{measurf}|_{i,j}$ is not likely to show variations during the year, as the N concentration in Figure 3 indicates another fate and transport mechanism for P compared to N, where P is more likely to be governed by adsorption kinetic and not so much by plant uptake as N.

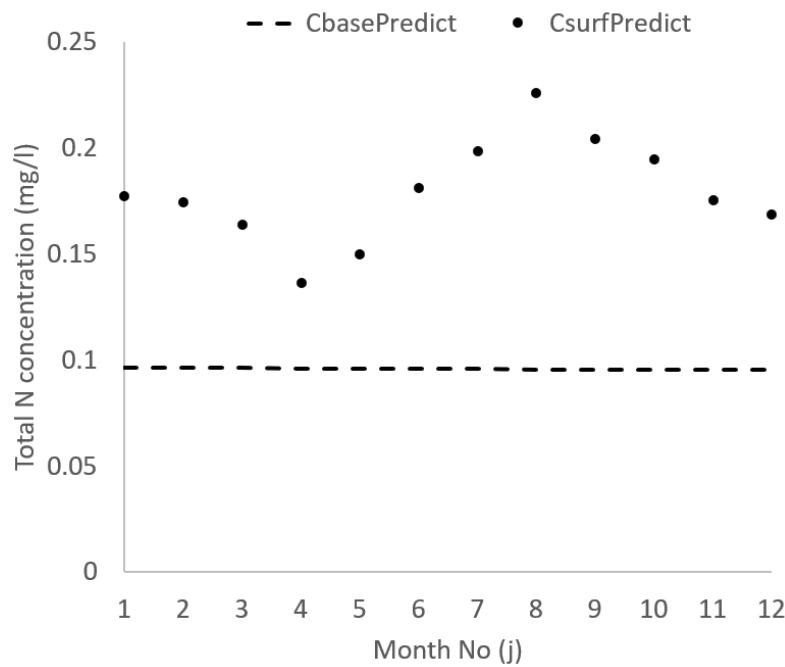


Figure 8. The variation of predicted concentration values during year 1990, where CbasePredict is the predicted base water concentration and CsrfPredict is the predicted surface-near water concentration.

The long-term variation in concentration levels is shown for P in Figure 9 for the winter flow concentration (January, $j=1$). The surface concentration of P is more likely to show a continuous drop during the entire period compared to the similar graph for N (Figure 4).

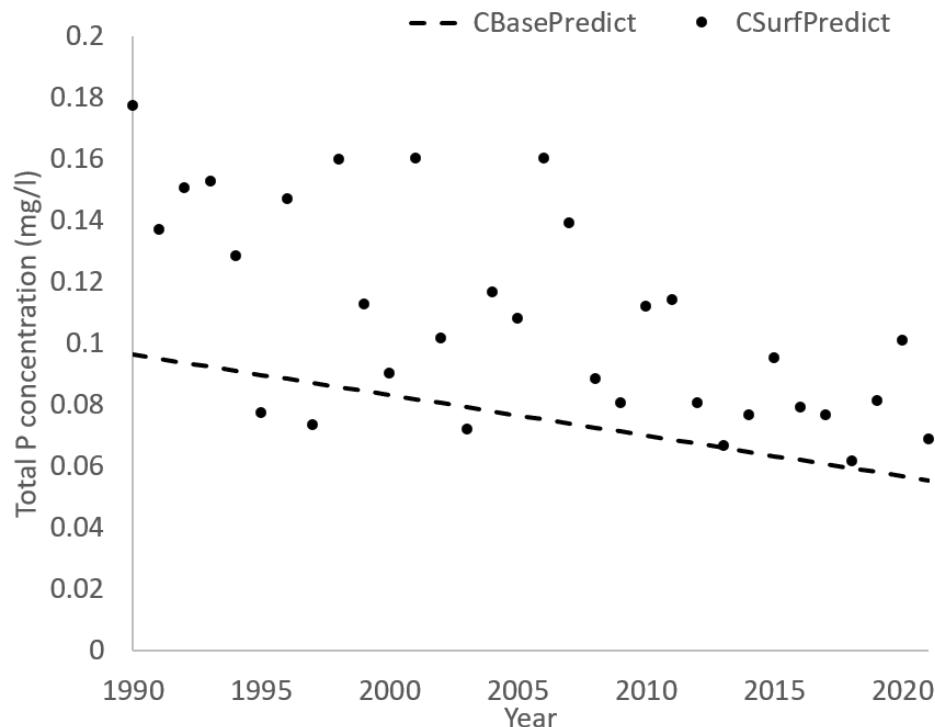


Figure 9. Long-term time trend of the N concentration (January), where CbasePredict is the predicted base water concentration and CsrfPredict is the predicted surface-near water concentration.

The load of P from the measured part of the lake catchment is compared with the measured load in Figure 10.

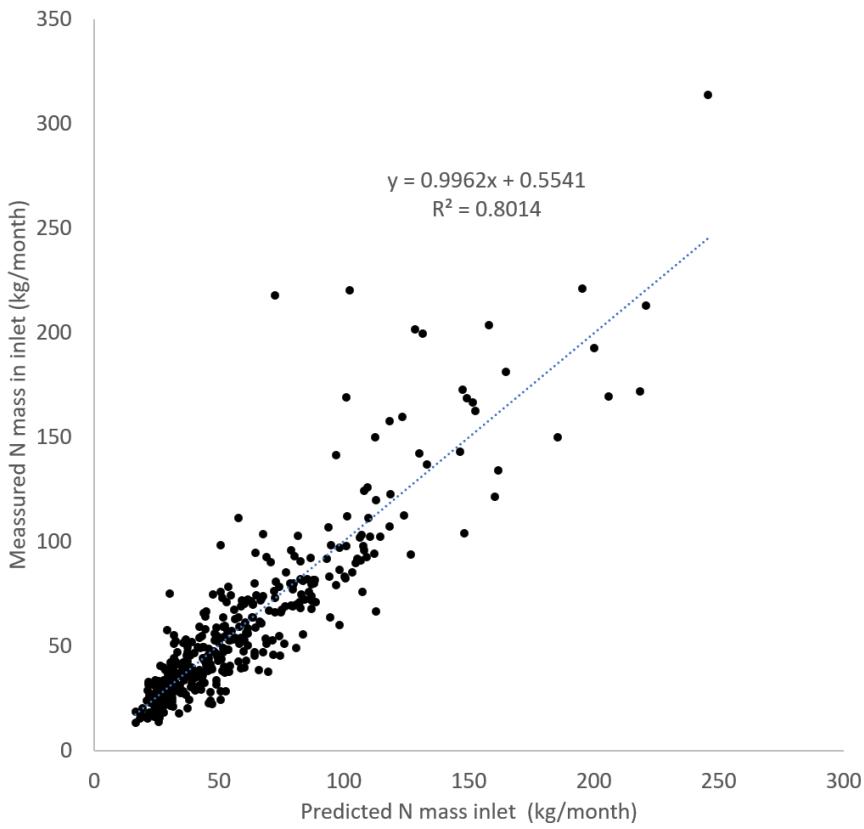


Figure 10. Comparison of the predicted load of P with the measured diffuse load.

The total load in and out of the lake is shown in Figure 11. The retention of P in the lake is not so pronounced as for N. The average retention is 14 kg/month.

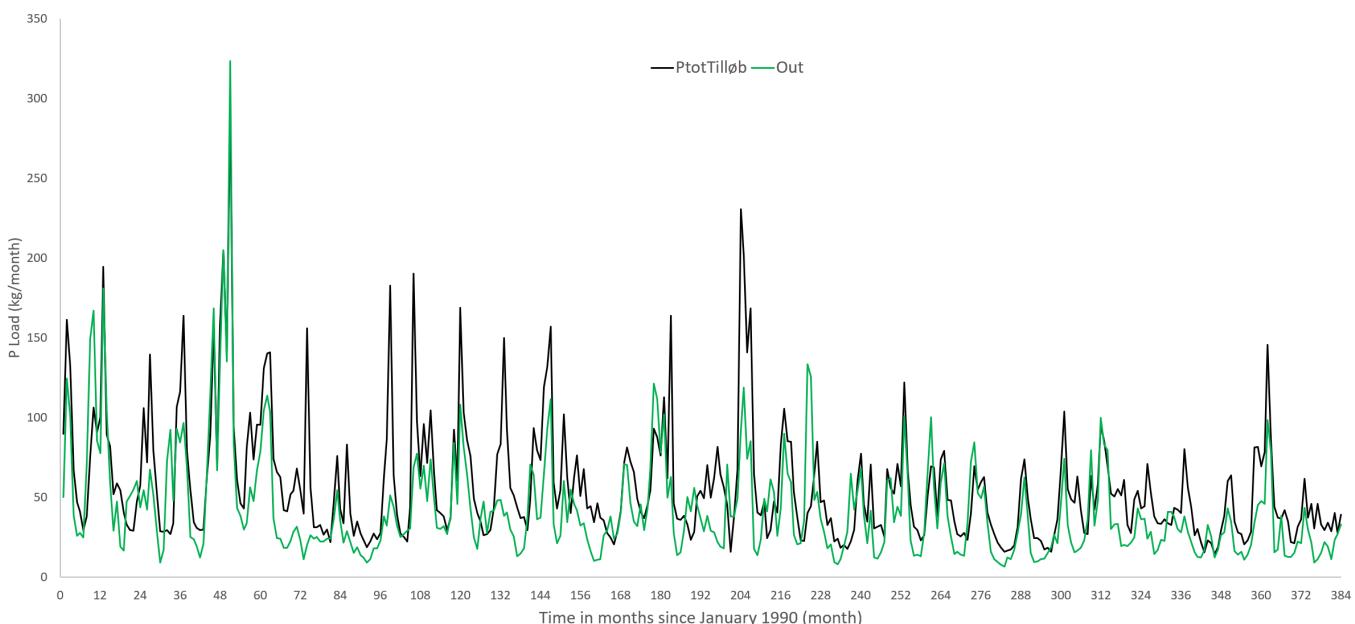


Figure 11. The total load of P in and out of the Bryrup Langsø.

Discussion

The purpose of the development of the presented statistical model concept is to predict the load of N and P in and out of a lake, where only a fraction of the inflow and outflow is measured, and the unmeasured fraction thus needs to be predicted. In most cases, the lake will only have one single outlet, so the outflow tends to be better described than the inflow. However, lakes may have seepage from the bottom that will contribute to the release of N and P in the same way as the outlet at the surface. The inlet to the lake may consist of several rivers and streams, out of which only a fraction of the water body is measured. There may be a main tributary entering the lake, while smaller streams are unmeasured. Obviously, there are uncertainties related to the measured load of both water and nutrient concentration levels that will alter the load prediction also for the measured fraction of the water bodies. However, in this method the focus has solely been allocated to predicting the unmeasured fraction of the load. Thus, proper quality assurance of the monitoring programs that are used as data input is critical for the load calculation. The well-known general modelling condition of “garbage in is garbage out” will also apply for the load predictions in case the data input to the load predictions is not of sufficient quality and, thus, introduces significant uncertainty.

The model assumes that the minimum monthly measured inlet flow of water during a year is equivalent to the base-flow. Thus, the model assumes a seasonality of the discharge, where there is a dry season and/or a season where the evaporation from the soil and plant surfaces evaporates all the precipitation. Better models could be applied for estimating the base-flow for the reference month. Extreme rain events during summer may lead to overestimation of the base-flow during the reference month.

Conceptually, the model relies on a linear mass balance model relating measured outflow to measured inflow. The model coefficients are used to predict the unmeasured part separated into a baseflow and a surface-near flow. This separation is beneficial from the point of view that the concentration levels of N and P not necessarily are similar. The intercept of the linear model can be interpreted as the volume of water that runs out of the lake when no measured water enters the lake. Thus, the estimated intercept predicts the water exchange through the lake bottom or from unmeasured wells entering the lake, and a negative value estimate indicates loss of water due to seepage through the lake bottom. The slope between inflow and outflow is assumed to reflect the fraction of unmeasured water in the inflowing rivers and streams. This is based on the rationale that if e.g. 2 m³ water is measured at the outlet of the lake for every measured 1 m³ at the inlet (slope equal 2), then there is 1 m³ unmeasured water in the inflow for every 1 m³ of measured water volume. For the baseflow, the assumption is that the change in intercept between months is equivalent to a similar change in measured baseflow. This is combined with the assumption that the smallest measured monthly inlet flow of water is entirely baseflow.

The accumulation of water is included as a memory effect in a linear model that includes last month's water input to model the actual month's output. Thus, the water accumulation impact on water outflow is assumed to be linear between the incoming volume and the resulting outgoing volume the month after. This is valid for smaller changes in water volume, as the larger changes in the relation between incoming water volume and outgoing water volume the month after tends to be nonlinear.

Any replacement of the measuring stations is assumed to take place at the start of a year. However, the statistical model for the water balance could be modified to include changes in measuring starting at any month and year.

The seasonality in base-flow is assumed to be independent of the measurement station combination, and this assumption could be challenged in case a large amount of unmeasured water exists. However,

in case the amount of unmeasured water is large, it will always be a challenge to make valid load prediction.

The N and P concentrations are predicted using a statistical model that decomposes the sources of nutrient into three components: (1) point sources; (2) surface-near sources; (3) base-flow source. The point source's strength must be identified both in the catchment of the measured inflow and in the catchment of the lake, and the model will estimate the effective load of the point source strength. However, some fraction of the point source contribution may be modelled as part of the baseflow, if the point sources are located at longer distance upstream from the lake, and especially if there are other lakes upstream that influence the concentration level of the inflowing water. The model works best for lakes that have inflow from streams and rivers not having other lakes upstream. In case of lake systems where water from one lake flows into other lakes downstream, a more comprehensive model for the load in and out each lake may be better. Such a model could apply the model in this paper for the upper catchment lakes using the principle in the paper.

The separation of the water body into respectively baseflow and surface-near flow is a rough categorization used to include the large variation in water transport routes. However, the specific case where the surface-near concentration level makes a decrease during the growing season indicates some realism behind the model. The surface-near sources of N and P are considered to come from the upper soil layer or from run off and drain systems in the catchment and will contribute to the surface-near flow as well.

Lake load of N and P was modelled using similar monitoring data from NOVANA by Windolf et al., (1996). Their approach was different from the approach in the paper in relation to:

- The lake water volume needed to be determined for each month to solve the mass balance.
- The mass balance of water was more aggregated without distinction between surface-near and base-flow water, which can have the consequence that unmeasured surface-near water may be hidden in seepage or ground water inflow through the lake bottom in the mass balance equation.
- The inlet concentration was not decomposed in surface-near water concentration and base-flow concentration.
- The unmeasured base-flow concentration needed to be estimated using additional information about ground water concentration levels.

In some circumstances, it may be critical that the method in this paper uses the predicted base-flow concentration to predict the unmeasured base-flow concentration in cases where point sources may be hidden as base-flow concentration due to upstream lakes. In such cases, the method can be adjusted by replacing the predicted base-flow concentration with independent measurements of ground water concentrations, as done by Jørgen et al., 1996.

Reference

Svendsen, L.M. & Norup, B. (eds.) 2005a: NOVANA. Nationwide Monitoring and Assessment Programme for the Aquatic and Terrestrial Environments. Programme Description – Part 1. National Environmental Research Institute, Denmark. 53 pp. – NERI Technical Report No. 532.

Svendsen, L.M., Bijl, L. van der, Boutrup, S. & Norup, B. (eds.) 2005b: NOVANA. National Monitoring and Assessment Programme for the Aquatic and Terrestrial Environments. Programme Description – Part 2. National Environmental Research Institute, Denmark. 138 pp. – NERI Technical Report No. 537.

Windolf, J., E. Jeppesen, J. P. Jensen & P. Kristensen, 1996. Modelling the seasonal variation in in-lake concentration and retention of nitrogen in 16 shallow lakes: a four-year mass balance study. Biogeochemistry 33: 25–44.

Appendix A

Table A1

The table is also available as supplementary material.

Year	i	j	$Q_{out,meal _{i,j}}$	Station location	$Q_{in,meal _{i,j}}$	$jref_t$	$Q_{in,vertical _{i,j}}$	N				P			
								$CNin_{totmeal _{i,j}}$	$CNout_{totmeal _{i,j}}$	$P_{meal _{i,j}}$	$P_{unmeal _{i,j}}$	$CPin_{totmeal _{i,j}}$	$CPout_{totmeal _{i,j}}$	$P_{meal _{i,j}}$	$P_{unmeal _{i,j}}$
				m ³ /month			mg/l			Kg/month	mg/l	mg/l		Kg/month	mg/l
1990	1	1	541037	1	545773	7	46169	10.109	3.871	121	79.4	0.1231	0.0932	19.3	12.9
1990	1	2	1328141	1	958569	7	47883	10.672	5.767	145	80.1	0.1276	0.0938	25.7	13.0
1990	1	3	1127606	1	811620	7	10843	8.430	6.800	165	77.3	0.1189	0.0897	30.2	12.6
1990	1	4	492480	1	344308	7	-6186	7.716	6.024	119	71.6	0.1457	0.0789	18.6	11.7
1990	1	5	374976	1	244786	7	-26419	7.290	4.661	107	68.6	0.1309	0.0698	16.4	11.6
1990	1	6	375840	1	239664	7	3130	8.070	3.323	135	66.9	0.1145	0.0736	25.7	11.2
1990	1	7	291946	1	179378	7	-23811	8.098	1.826	139	67.3	0.1115	0.0856	26.6	11.1
1990	1	8	294624	1	193992	7	-1379	7.912	1.796	174	68.0	0.1431	0.2409	36.2	11.2
1990	1	9	520992	1	290387	7	47547	7.539	1.609	155	68.0	0.2133	0.2865	29.3	10.9
1990	1	10	739238	1	459224	7	33350	8.948	2.124	135	69.8	0.1915	0.2258	22.8	11.2
1990	1	11	593568	1	491495	7	18706	9.063	3.177	107	72.4	0.1456	0.1446	15.2	11.5
1990	1	12	666922	1	550819	7	25898	9.265	4.615	114	73.9	0.1438	0.1168	17.6	11.6
1991	2	1	1473120	1	1192849	8	35362	10.347	6.139	85	74.6	0.1308	0.1228	12.2	20.9
1991	2	2	613993	1	508217	8	14570	9.862	6.955	109	75.4	0.1331	0.1717	18.7	21.0
1991	2	3	703616	1	507748	8	9204	8.999	5.871	127	72.6	0.1131	0.0848	23.2	20.6
1991	2	4	536803	1	380042	8	14644	8.145	5.120	83	66.8	0.0895	0.0546	11.8	19.7
1991	2	5	767629	1	445694	8	-18333	8.330	5.379	73	63.9	0.0874	0.0617	9.5	19.6
1991	2	6	452304	1	267244	8	3838	7.007	4.100	102	62.1	0.1470	0.0426	18.5	19.2
1991	2	7	344174	1	198789	8	-26009	7.478	3.116	106	62.5	0.1480	0.0492	19.5	19.1
1991	2	8	287928	1	181390	8	-16060	7.815	2.423	142	63.2	0.1248	0.1654	29.4	19.2
1991	2	9	287194	1	192924	8	1155	7.462	2.047	121	63.2	0.1034	0.1778	22.7	18.9
1991	2	10	381672	1	212856	8	16433	7.990	2.035	100	65.0	0.0947	0.1442	16.1	19.1
1991	2	11	565574	1	335956	8	37784	8.432	2.376	71	67.6	0.1010	0.1066	8.3	19.5
1991	2	12	595408	1	397871	8	26345	9.539	3.431	78	69.2	0.1070	0.0736	10.4	19.6
1992	3	1	894586	1	596288	7	24556	11.765	4.861	95	48.4	0.1408	0.0610	12.4	12.0
1992	3	2	688288	1	443764	7	16917	10.660	5.993	119	49.7	0.1187	0.0614	18.9	12.1
1992	3	3	976009	1	748829	7	29177	10.975	6.256	138	45.2	0.1459	0.0691	23.4	11.5
1992	3	4	818294	1	581042	7	16843	9.243	6.187	93	36.1	0.1001	0.0631	12.0	10.1
1992	3	5	570767	1	386877	7	-25674	8.377	5.502	83	31.4	0.0923	0.0519	9.7	9.9
1992	3	6	242352	1	183332	7	-49709	8.515	4.328	111	28.6	0.0930	0.0375	18.8	9.2
1992	3	7	227128	1	175485	7	-19936	8.675	3.030	115	29.3	0.1082	0.0760	19.7	9.1
1992	3	8	366673	1	221493	7	31375	8.519	2.682	151	30.4	0.0829	0.1999	29.6	9.2
1992	3	9	362102	1	209606	7	-3056	7.941	2.576	131	30.4	0.0838	0.2556	22.9	8.8
1992	3	10	315783	1	245744	7	20047	10.304	1.632	110	33.2	0.0981	0.1524	16.3	9.2
1992	3	11	825552	1	630509	7	50417	13.628	2.569	81	37.4	0.1348	0.1129	8.5	9.8
1992	3	12	833518	1	636757	7	19563	12.648	5.086	88	39.8	0.1490	0.1015	10.7	9.9
1993	4	1	1005757	1	691131	6	48106	10.866	6.614	105	48.7	0.2005	0.0963	12.6	12.2
1993	4	2	716910	1	502181	6	11589	10.171	7.089	130	49.9	0.1162	0.0964	19.1	12.4
1993	4	3	571184	1	390190	6	-633	9.758	6.525	149	45.5	0.0892	0.0445	23.6	11.7
1993	4	4	417624	1	287758	6	-14085	9.270	5.892	104	36.3	0.0697	0.0570	12.2	10.3
1993	4	5	324421	1	211479	6	-28282	8.467	4.631	92	31.5	0.0849	0.0610	9.9	10.1

2019	30	12	963146	2	638716	7	29661	6.941	4.611	39	12.3	0.0830	0.0495	4.7	1.8
2020	31	1	944102	2	649373	9	33648	7.025	4.814	41	12.8	0.0870	0.0488	4.7	1.8
2020	31	2	1767958	2	1192155	9	73743	4.725	4.563	76	15.8	0.0919	0.0557	9.1	2.2
2020	31	3	1367473	2	812148	9	2981	5.120	4.053	100	17.9	0.0664	0.0578	12.1	2.5
2020	31	4	701215	2	400996	9	-14905	6.477	3.687	41	12.7	0.0601	0.0222	4.7	1.8
2020	31	5	587604	2	300906	9	-20867	7.124	3.372	28	11.7	0.0611	0.0295	3.1	1.7
2020	31	6	508823	2	249716	9	-9427	6.670	3.080	74	15.7	0.0805	0.0759	8.9	2.2
2020	31	7	494290	2	246977	9	-1714	6.192	2.621	80	16.2	0.1165	0.0272	9.6	2.3
2020	31	8	463982	2	239958	9	-1751	5.887	2.400	134	20.9	0.0911	0.0280	16.5	2.9
2020	31	9	427740	2	222009	9	-8347	6.358	2.121	99	17.9	0.0511	0.0302	12.1	2.5
2020	31	10	483024	2	249480	9	27463	7.059	2.297	65	14.8	0.0467	0.0321	7.7	2.1
2020	31	11	572292	2	296205	9	17215	6.410	2.535	22	11.1	0.0656	0.0394	2.3	1.6
2020	31	12	535572	2	347927	9	31413	7.229	3.079	32	12.0	0.0661	0.0397	3.6	1.7
2021	32	1	994408	2	654066	6	27947	8.434	4.224	41	12.2	0.0705	0.0439	5.2	1.7
2021	32	2	665684	2	370317	6	6297	7.474	4.685	76	15.0	0.0638	0.0448	10.0	2.1
2021	32	3	736993	2	467631	6	13340	7.033	4.409	101	17.1	0.0593	0.0297	13.5	2.3
2021	32	4	584913	2	310770	6	-13974	7.294	4.600	41	12.1	0.0543	0.0158	5.1	1.7
2021	32	5	630061	2	382618	6	24109	5.952	4.228	28	11.1	0.0828	0.0178	3.4	1.6
2021	32	6	517628	2	271264	6	-21650	6.624	3.713	75	14.9	0.0765	0.0296	9.8	2.0
2021	32	7	484386	2	295042	6	-2236	5.688	3.008	81	15.4	0.0646	0.0456	10.6	2.1
2021	32	8	654433	2	390872	6	9092	5.439	2.198	136	19.9	0.0549	0.0295	18.3	2.7
2021	32	9	413749	2	314772	6	14011	6.225	2.170	101	17.0	0.0614	0.0273	13.4	2.3
2021	32	10	616283	2	463188	6	27127	7.115	2.709	65	14.1	0.0627	0.0384	8.5	1.9
2021	32	11	628364	2	387810	6	19041	7.515	3.359	22	10.5	0.0451	0.0432	2.4	1.5
2021	32	12	802914	2	563289	6	29400	7.903	4.069	32	11.4	0.0475	0.0414	3.9	1.6