



Nr B 2363
Oktober 2019

Dioxiner och PCBer i sik - statistiska analyser

Hannes Waldetoft

Författare: Hannes Waldetoft

Medel från: Svenska Insjöfiskarens Centralförbund och Stiftelsen Institutet för Vatten- och Luftvårdsforskning

Rapportnummer B 2363

ISBN 978-91-7883-106-7

Upplaga Finns endast som PDF-fil för egen utskrift

© **IVL Svenska Miljöinstitutet 2019**

IVL Svenska Miljöinstitutet AB, Box 210 60, 100 31 Stockholm

Tel 010-788 65 00 // www.ivl.se

Rapporten har granskats och godkänts i enlighet med IVL:s ledningssystem

Innehållsförteckning

Sammanfattning.....	4
Summary	6
Inledning	8
Frågeställningar	8
Dioxin och PCB	8
Kemi och användningsområden.....	8
Hälsorisker	9
Toxiska ekvivalenter.....	9
Reglering och kostrekommendationer	9
Tidigare framkomna samband	10
Data	10
Hur insamling av data påverkar slutsatser.....	10
Beroende variabel.....	12
Förklarande variabler.....	13
Metod.....	17
Länkfunktion	17
Modellspecifikation	18
Huvudeffekter	18
Interaktionseffekter	19
Tolkning av resultat av estimering.....	19
Resultat	20
Estimering av ursprunglig modell	20
Estimering av modell utan interaktioner	22
Fixa effekter	23
Fångstplats.....	26
Statistisk modellering av fetthalt.....	27
Diskussion	29
Referenser.....	32

Sammanfattning

IVL Svenska Miljöinstitutet bedriver mellan 2017 och 2019 projektet "Dioxiner i fet fisk – hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske" i syfte att klarlägga hur förekomst av dioxinlika ämnen i fet fisk från Vänern, Vättern och Östersjön varierar över tid och rum, med olika beredningsmetoder, olika delar av fisken och andra faktorer. En målsättning är att underlätta för den småskaliga fiskenäringen och dess associerade beredningsindustrier att förse konsumenter med livsmedel med lågt dioxin- och PCB-innehåll.

Denna delrapport syftar till att genom statistisk modellering undersöka hur olika morfometriska, tids- och rumsliga aspekter påverkar förekomsten av dioxinlika ämnen i sik (*Coregonus lavaretus*) från de tidigare nämnda vattnen. Andra frågeställningar och andra arter behandlas i separata rapporter. Med dioxinlika ämnen syftas här på ett sammanslaget mått av dioxiner, furaner och dioxinlika PCBer.

Resultaten visar på att det föreligger skillnader mellan de tre vattnen, där sik från Vänern generellt sett innehåller högre halter av dioxinlika ämnen än sik från Vättern och Bottniska viken. Skillnaderna mellan Vänern och Vättern visade sig huvudsakligen bero på skillnader i fetthalt, där den fetare Vänernsiken har högre halter. Den uppskattade skillnaden är att Vätternsiken i genomsnitt har 57% lägre fetthalt än Vänernsiken. Denna modellering kan dock inte svara på varför fetthalten i sik är högre i Vänern.

Sambandet som fanns mellan fetthalt och dioxinlika ämnen visade att en enhets ökning av fetthalten (en procentenhet) i genomsnitt leder till en ökning av halter av dioxinlika ämnen med 25%. Detta samband gäller för alla vatten inkluderade i undersökningen. Om en genomsnittlig sik i Vättern hade levt i Vänern beräknas det att den skulle ha 46% högre halter av dioxinlika ämnen, huvudsakligen på grund av att fetthalten generellt sett är högre i Vänernsiken än Vätternsiken. Modelleringen indikerade att sik från Bottniska viken har lägst halter av de tre vattnen. Det framkom att sik från Bottniska viken har i storleksordningen tre till fem gånger lägre halter av dioxinlika ämnen än sik från Vänern och Vättern. Denna skillnad beräknades under förutsättning att fetthalten är likadan mellan vattnen. I realiteten har dock sik från Bottniska viken lägre fetthalt än sik från Vänern (men inte Vättern). Därav är skillnaden i halter av dioxinlika ämnen mellan sik från Bottniska viken och Vänern sannolikt större än fem gånger.

Var inom ett vatten en sik fångas (dvs inom Vänern, Vättern eller Bottniska viken) tycks ha marginell betydelse för förekomsten av dioxinlika ämnen. Av större betydelse visades vara vilket av vattnen en sik fångas i. När Vänernsiken delades upp efter i vilken av Vänerns tre delbassänger (norra Värmlandssjön, södra Värmlandssjön samt Dalbosjön) den fångats i syntes inga indikationer på att någon av delbassängerna skulle vara mer fördelaktig att bedriva fisket i.

Ett samband mellan sikens längd och förekomst av dioxinlika ämnen hittades. Det visades på att korta sikar generellt sett har lägre halter än längre sik. Den uppskattade skillnaden är att den längre siken generellt sett har 22% högre halter än den korta (38 cm användes som gränsdragning för lång/kort).

Gällande säsongvariationer hittades inga indikationer på att någon säsong skulle vara mer fördelaktig att bedriva sikfiske på än någon annan.



Resultaten indikerar även att halterna av dioxinlika ämnen i sik minskar med tiden. Den skattade minskningen är ca 10% per år.

Överlag förstärktes uppfattningen att fetthalten har stor betydelse för halter av dioxinlika ämnen, men tydligt är att även andra faktorer, såväl morfometriska som tid- och rumsliga, spelar in.

Summary

Between 2017 and 2019, IVL Swedish Environmental Research Institute is running the project “Dioxins in oily fish – threats and potential for development of small scale professional coastal and lake fishing” with purpose of investigating how dioxin-like compounds in oily fish from Lake Vänern, Lake Vättern and the Baltic Sea varies with time and space, different cooking procedures, parts of the fish and other factors. A goal is to assist small-scale professional fishing and its associated industries to provide customers with foodstuffs low in dioxins and PCBs.

The purpose of this report is to by application of statistical modelling, examine how morphometric and spatiotemporal variation affect the prevalence of dioxin-like compounds in European whitefish (*Coregonus lavaretus*) in the waters mentioned above. Other research questions for other species are treated in separate reports.

The results indicate that differences with regards to dioxin-like compounds exist between the waters, where whitefish from Lake Vänern in general have higher levels than whitefish from Lake Vättern and the Gulf of Bothnia. Differences between Lake Vänern and Lake Vättern were shown to be a consequence of differences in fat content between the lakes, where whitefish from Lake Vänern have higher fat content. The estimated difference is that whitefish from Lake Vättern on average have 57% lower fat content compared to whitefish from Lake Vänern. Why the fat content is higher in Lake Vänern cannot be answered by the statistical modelling used in this report.

When increasing the fat content with one percentage, it is estimated that the content of dioxin-like compounds increases with 25%. This estimate applies to whitefish from all waters included in this study. If an average whitefish from Lake Vättern instead would have lived in Lake Vänern, it is estimated that it would have 46% higher levels of dioxin-like compounds, since the fat content is on average higher in Lake Vänern compared to Lake Vättern. The statistical modelling also indicated that whitefish from the Gulf of Bothnia have about three to five times lower levels of dioxin-like compounds than whitefish from Lake Vänern and Lake Vättern. This difference was calculated under the condition that the fat content in whitefish is constant between the waters. However, the fat content is in fact lower in whitefish from the Gulf of Bothnia than in whitefish from Lake Vänern (but not Lake Vättern). Due to this, the difference in levels of dioxin-like compounds between whitefish from the Gulf of Bothnia and Lake Vänern is likely to be larger than five times.

Where within a water a whitefish is caught (i.e. within Lake Vänern, Lake Vättern or the Gulf of Bothnia) seem to be of minor importance with regards to prevalence of dioxin-like compounds. More important is in which water a whitefish is caught. When whitefish from Lake Vänern were divided into from which of three sub-areas (in essence: north, south and southwest) it was caught, little variation in prevalence of dioxin-like compounds were seen between the areas. In other words, no specific area within Lake Vänern was found to be more beneficial to perform the fishing in.

An association between the length of a fish and dioxin-like compounds was found. Results indicate that longer fish on average have higher levels. The estimated difference is that longer whitefish on average have 22% higher levels of dioxin-like compounds (38 cm was used as cutoff value for long/short).

Regarding seasonal variations, no indications were found that any seasons is more beneficial to perform the fishing at.



The results also indicate that levels of dioxin-like compounds are reducing over time, with an estimated average reduction of 10% per year.

In general, the results enhance the belief that the fat content has a major influence on the prevalence of dioxin-like compounds, but evident is also that other factors, such as morphometric and spatiotemporal variations, are of importance.

Inledning

Denna rapport är en del av projektet "Dioxiner i fet fisk – hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske" som genomförs av IVL Svenska Miljöinstitutet i samarbete med Sveriges Lantbruksuniversitetet och Statens Veterinärmedicinska Anstalt, med finansiering av IVL:s forskningsstiftelse och Svenska Insjöfiskarens Centralförbund. Projektet löper mellan 2017 och 2019 och det huvudsakliga syftet är att öka kunskaperna om hur dioxiner, furaner och PCBer varierar över tid och rum i fet fisk. Ambitionen är att kunna optimera småskaligt yrkesfiske vad gäller fångstens innehåll av dessa miljögifter. De vatten som är inkluderade i projektet är Vänern, Vättern och den svenska Östersjökusten. Här berörs bara sik, men projektet som helhet inkluderar ett flertal arter.

Frågeställningar

För tillfället begränsar förekomsten av dioxinlika ämnen i fet fisk från framför allt Vänern, Vättern och Östersjön möjligheten att utveckla svenskt småskaligt yrkesmässigt fiske (Karlsson, et al., 2017). För sik från Vänern och Vättern krävs att yrkesfiskare försäkras sig om att halterna av dioxinlika ämnen i fångsten inte överstiger EU:s gränsvärden innan den går till försäljning (Karlsson, 2016). Att analysera varje fångat parti är kostsamt och omöjliggör försäljning av färsk fisk. Därav finns det intresse av att rikta fisket mot exempelvis vissa platser, årstider eller individer med viss morfometri för att säkerställa att halterna är under gränsvärdena. I denna rapport har statistisk modellering använts som ett medel för att öka kunskapsläget och för att i längden kunna rikta denna kunskap mot ett kontrollfiskeprogram i Vänern och Vättern. De frågeställningar som därför har undersökts är:

- Hur påverkar morfometriska, tid- och rumsliga förändringar halter av dioxinlika ämnen i sik i Vänern, Vättern och längs den svenska Östersjökusten?

Morfometriska mått är kroppsliga egenskaper, här längd och vikt. Rumsliga variationer innebär variationer inom ett vatten men även mellan olika vatten. Förändringar i tid avser här årliga- och säsongsmässiga förändringar.

- Förändras de uppfattningar framkommit hittills i projektet?

Uppfattningar om vad som påverkar halterna av dioxinlika ämnen i sik har framkommit under tidigare studier och i projektet "Dioxiner i fet fisk – hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske". Det är av intresse att undersöka om dessa uppfattningar förstärks eller om något motsäggande framkommer.

Dioxin och PCB

Kemi och användningsområden

Dioxiner är en grupp ämnen där grundstrukturen är två bensenringar till vilka kloratomer i olika antal är bundna. Antalet kloratomer och hur de binder till bensenringarna ger upphov till olika

egenskaper. Dioxiner kategoriseras i två grupper: polyklorerade dibensofuraner (s.k. furaner) och polyklorerade dibenso-p-dioxiner, ofta benämnda som dioxiner. Dioxiner och furaner benämns sammanslaget som PCDD/F. Varje unik variant av dioxin, furan och PCB benämns som en *kongen*.

PCB, eller *polyklorerade bifenyler* består även de av två bensenringar till vilka kloratomer i olika antal och struktur är bundna. Vissa kongener har en plan struktur mellan de två bensenringarna och deras toxiska egenskaper liknar de hos PCDD/F. Dessa kongener kallas därför *dioxinlika PCBer* och övriga benämns som *icke-dioxinlika PCBer*. Sammantaget kallas i denna rapport dioxiner, furaner och dioxinlika PCBer för "dioxinlika ämnen".

De här ämnena har olika utsläppskällor och användningsområden. PCBer är syntetiskt framställda industrikemikalier som introducerades under slutet av 1920-talet. De användes bl.a. som tillsats i hydrauloljor och fogmassor, men har sedan 1970-talet fasats ut genom stegvisa förbud. PCDD/F däremot bildas oavsiktligt i vissa processer, t.ex. vid blekning av pappersmassa med elementärt klor, en process som förekom i svensk pappers- och massaindustri fram till 1990-talet (Hållén & Karlsson, 2018). PCDD/F bildas även oavsiktligt vid förbränning av avfall, och hamnar därför i biotoper via nederbörd. Utsläpp från massa- och pappersindustrin är lokala medan utsläpp från förbränning fördelas över större ytor.

Hälsorisker

Dioxinlika ämnen är persistenta och fettlösliga. Att de är persistenta innebär att de bryts ned till mindre farliga ämnen väldigt långsamt. I kombination med att de är fettlösliga leder detta till biomagnifikation, alltså att organismer högre upp i näringskedjan generellt sett har högre halter än organismer längre ned. Djurförsök har visat att dioxinlika ämnen har negativa effekter vad gäller immunförsvaret och reproduktionsförmågan. Höga doser kan påverka nervsystemet och hjärnans utveckling negativt (Cantillana & Aune, 2012).

Toxiska ekvivalenter

För att kvantifiera hur toxiskt ett prov är har World Health Organisation (WHO) utvecklat ett system där koncentrationen av varje kongen av dioxinlika ämnen i ett prov multipliceras med en toxisk ekvivalensfaktor (TEF) för att korrigera för att alla inte är lika toxiska. TEF-värdena är satta som värden mellan 0 och 1 där TEF=1 är värdet för den mest toxiska kongenen, 2378-TCDD. Kongener med ett lägre värde anses mindre toxiska. Användandet av TEF möjliggör att få ett unikt värde för toxiciteten av ett prov.

Reglering och kostrekommendationer

EU har satt gränsvärdena vad gäller dioxinlika ämnen för vilka ett överskridande innebär att fisken inte får saluföras. Halterna i sik från Vättern ligger generellt sett under gränsvärdena medan sik från Vätern ofta överskrider gränsvärdena (Karlsson, et al., 2017). Gränsvärdena för PCDD/F är 3,5 TEQ pg/g vv och 6,5 TEQ pg/g vv då dioxinlika PCBer är inräknade (European Commission, 2011). Våtvikt (vv) indikerar att den kemiska analysen gjorts på muskelprov från färsk eller fryst och tinad fisk.

Tidigare framkomna samband

Fetthalt har en stark samverkan med förekomst av dioxinlika ämnen (positiv korrelation). Tester med att använda fetthalt som proxy för dioxinlika ämnen har genomförts via en handhållen fetthaltsmätare, men mätaren visar stor mätosäkerhet vilket bidrar negativt till användandet av en sådan som en del av ett kontrollfiskeprogram för sik (Karlsson, et al., 2017).

Längs Bottenhavskusten i närhet av skogsindustrirecipienter har historiska utsläpp lagrade i sediment visat en lokal påverkan på halterna i fisk (Malmaeus, et al., 2012). Vissa PCB-kongener i abborre fanns i högre halter närmare utsläppskällan.

Vad gäller siken så har skillnader mellan vatten påträffats, där Vänern generellt sett tyckts ha högre halter jämfört med Vättern och Bottenhavet. En trolig förklaring är högre fetthalter i Vänern (Karlsson, et al., 2017).

Generellt sett minskar halterna av dioxinlika ämnen i miljön över tid, även om nedgången för dioxiner och furaner avstannat något (Karlsson & Malmaeus, 2014).

Data

De data som används till analysen har tagits fram av projektet "Dioxiner i fet fisk – hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske" men även av tidigare mindre delprojekt, och har samlats in mellan 2013 och 2019. I data ingår sik, strömming, gädda, abborre, lax och öring, men som nämnts ligger fokus här på sik, till stor del på grund av att den relativt stora datamängden för denna art möjliggör en mer komplex statistisk metod. För de andra arterna finns även andra frågeställningar, vilka behandlas separat. De individer som analyserats har fångats av IVL eller av yrkesfiskare i Vänern, Vättern eller Bottniska viken. Hel, ej urtagen fisk har levererats fryst till IVLs fisktoxikologiska enhet i Stockholm. Där har materialet preparerats och provberetts för att sedan skickas till ett ackrediterat laboratorium för analys. De laboratorier som anlitats är ALS (Prag, Tjeckien) och Eurofins (Hamburg, Tyskland), (Karlsson & Hällén, 2019).

Hur insamling av data påverkar slutsatser

Hur data har samlats in påverkar konstruktionen av en statistisk modell avsedd att finna samband mellan dioxinlika ämnen och andra variabler. Till exempel har de som bedrivit fisket själva rapporterat var de fångat fisken. De har alltså inte följt en given mall för var fisket ska bedrivas, vilket har lett till många fångstplatser med olika stora fångster per plats (för analysen av siken finns 38 olika fångstplatser). En fångstplats kan ses därför ses som slumpmässigt utvald ur en stor population av möjliga fångstplatser. För att kunna använda all information om fångstplatser är det därför lämpligt att inkludera fångstplats som en s.k. "random-effect" i modellen. Slutsatserna angående vilken påverkan olika fångstplatser har på förekomsten av dioxinlika ämnen kommer då att vara i termer av hur mycket variation fångstplatserna bidrar med och alltså inte i termer av halter för specifika fångstplatser. Anledningen till detta är att det är olämpligt att använda

dummy-variabler för att koda fångstplatserna. Variabeln skulle få 37 nivåer, vilket inte är optimalt. Ska slutsatser dras i termer av nivåer för olika fångstplatser bör informationen kodas om så att närliggande plaster slås ihop till en gemensam. Antalet dummy-variabler kan på så sätt minskas kraftigt.

På liknande sätt har det inte planerats i förväg under vilka årstider fisket har bedrivits, av vilket en konsekvens blivit att inga sikar fångats i Vänern under sommartid, p.g.a att sikfiske sällan bedrivs under denna årstid. Slutsatser angående en årstidseffekt i Vänern bör därför dras med stor försiktighet.

Teorin bakom en möjlig säsongsvariation kommer från att den skulle hänga ihop med fiskens lek. Problemet är att siken består av flertalet underarter som leker vid olika tidpunkter, och det finns inte rapporterat till vilken underart en individ tillhör. En möjlig säsongsvariation kan därför vara olika för olika underarter. Det komplicerar tolkningen av en eventuellt signifikant säsongseffekt.

Tillgängliga data har även ett stort bortfall för vissa variabler som skulle kunna vara intressanta, vilket gör att de inte kan inkluderas i en statistisk modell. Till exempel kan relationen mellan dioxinlika ämnen och fångstdjup vara av intresse, men fångstdjup har rapporterats för ett för litet antal fiskar för att den informationen ska vara lämplig att inkludera i en modell. Även information om en individs ålder saknas, men kommer tillkomma i ett senare skede av projektet.

Utöver detta är data obalanserad. Det innebär att fördelningen av antalet fiskar (observationer) för olika variabler eller kombinationer av variabler är ojämn. Det leder till att samband som kanske finns mellan någon variabel och halten av dioxinlika ämnen eventuellt inte kan upptäckas. Exempelvis har det på vissa fångstplatser fångats 20 eller fler individer, medan för många andra, har endast en fisk fångats. En jämnare fördelning är att föredra.

Ett annat sätt på vilket datainsamling och analys av muskelprover påverkar slutsatser är att vissa observationer är s.k. samlingsprov. Det innebär att muskelvävnad från flera fiskar från samma fångst homogeniserats innan analys. De koncentrationer och morfometriska mått som finns för dessa prov är därför medelvärden. Att analysera flera individer i ett samlat prov har gjorts för att reducera kostnader. Majoriteten av proverna i datasetet är dock analyserade som individprover. Hur de här samlingsproven fördelar sig vad gäller antalet individer ses i Tabell 1.

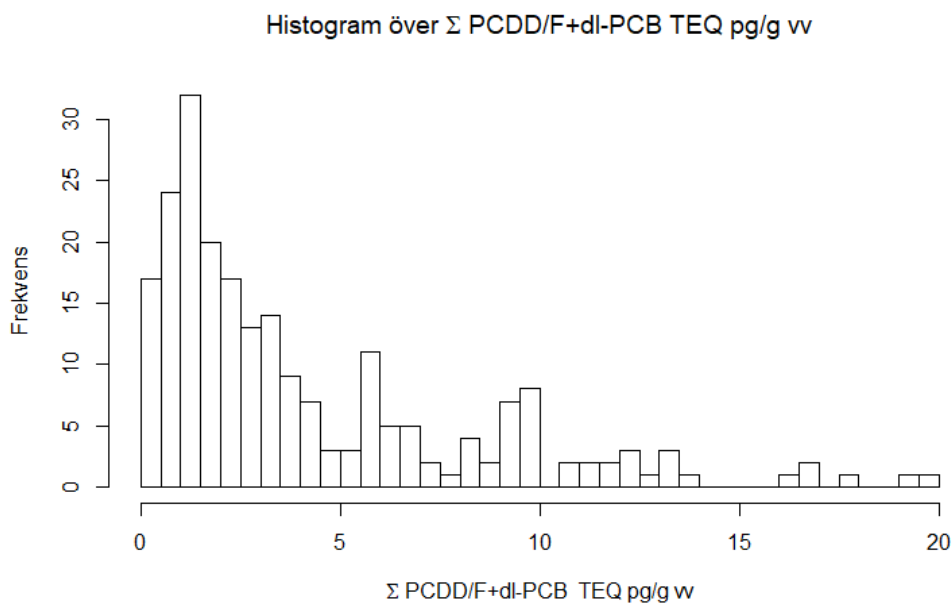
Tabell 1. Frekvenstabell över antal individer per samlingsprov.

Antal individer	Antal observationer
1	204
2	12
3	8
4	7
5	4
6	3

Genom att, för vissa observationer, ha tillgång till medelvärden istället för information om varje individ minskar mängden tillgänglig information, vilket får en påverkan på de tester som används för att dra slutsatser om enskilda variabler. Hypotestesten kommer att förkastas mer sällan jämfört med om varje individ hade analyserats. I statistiska termer benämns det som att testen är konservativa. I och med att ett samlingsprov i viss mån innehåller information om flera individer kommer varje observation i estimeringen av en modell ges en vikt som är lika stor som antalet individer i motsvarande samlingsprov. De får då större inflytande på estimaten än de prover som gjorts på endast en individ. Överlag är andelen samlingsprov få vilket leder till en uppfattning att de inte utgör några större problem vid hypotestest av parameterskattningar.

Beroende variabel

Den här rapporten fokuserar endast på det sammanslagna mått som anger förekomsten av dioxiner, furaner och dioxinlika PCB i TEQ pg/g vv, och kommer därför att använda det måttet som beroende variabel i den statistiska modelleringen. Anledningen är att det är detta mått som innehåller mest information om dioxinlika ämnen. Måttet är summan av koncentrationer av de 17 PCDD/F-kongenerna och de 12 PCB-kongenerna viktade med respektive TEF-värde. Figur 1 visar fördelningen av dioxinlika ämnen i siken från Vänern, Vättern och Bottniska viken. Antalet observationer är 224. En stor andel värden ligger under EU:s gränsvärde, men ett flertal överskridande värden syns.



Figur 1. Histogram över dioxinlika ämnen angivna i toxiska ekvivalenter pg/g våtvik (vv) i sik från Vänern, Vättern och Bottniska viken. Gränsvärdet för saluföring ligger på 6.5 TEQ pg/g vv.

Förklarande variabler

De förklarande variabler som finns tillgängliga att inkludera i en modell visas i Tabell 2. Variablerna *År*, *Säsong* och *Längd.binär* finns inte med i originaldata utan är skapade utifrån information om fångstdatum och fiskens längd. Den binära längdvariabeln antar värdet 0 om en fisk är kortare än 38 cm och 1 om den är längre. I analysen är *längd.binär=0* referensnivå. Som en del av ett kontrollfiskeprogram har det föreslagits att tillåta försäljning av sik kortare än 43 cm. Intentionen var att ha detta som gräns vid kodning av den binära variabeln, men väldigt få individer i data är längre än 43 cm. För att få en jämnare fördelning av "långa" och "korta" sikar, och därmed en bättre statistisk analys, sattes gränsen istället till 38 cm. Att inkludera information om en individs ålder är av intresse att ha med som förklarande variabel, men denna information finns inte tillgänglig vid skrivandet av denna rapport. Åldersbestämning av den sik som finns tillgänglig pågår under 2019.

Alla laboratorieanalyser av sik har gjorts av ett och samma laboratorium (ALS), så det finns ingen anledning att inkludera någon "Lab"- variabel i en modell. Vidare anses "År" här vara en kontinuerlig variabel.

Tabell 2. Möjliga förklarande variabler att inkludera i en modell.

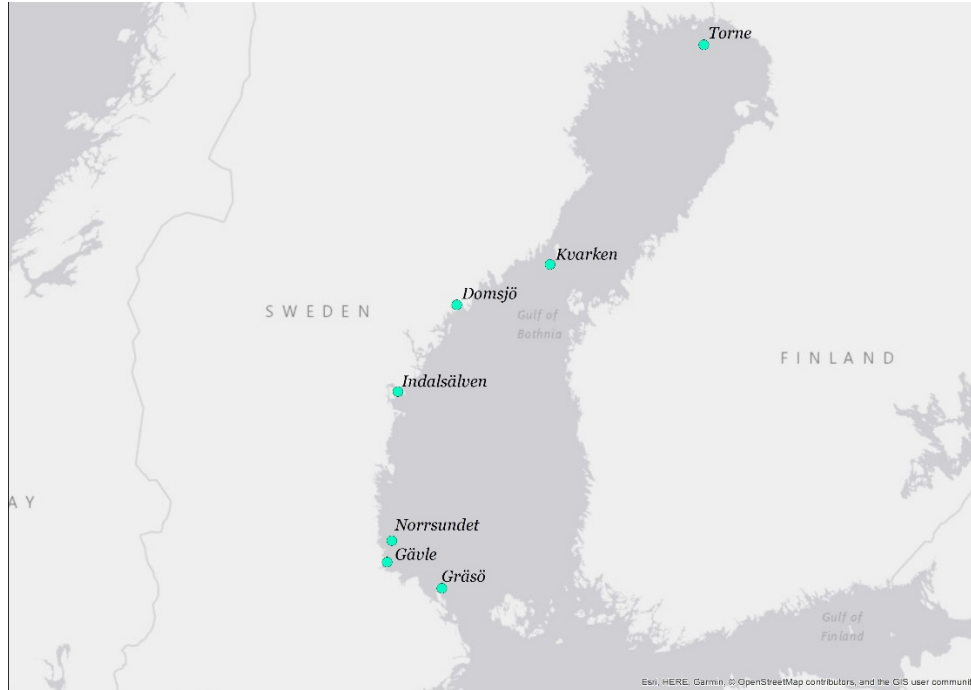
Variabelnamn	Andel bortfall (%)	Kategorisk/kontinuerlig	Antal nivåer
Laboratorium	0,0	Kat	1
Vatten	3,0	Kat	3
Fångstplats	0,0	Kat	38
Fångstdatum	4,2	-	-
Säsong	4,2	Kat	4
År	4,2	Kont	-
Preparationsdatum	11	-	-
Fångstkoordinater	21	Kont	-
Antal individer	0,0	Kat	6
Kön	12	Kat	2
Längd.binär	1,3	Kat	2
Längd	1,3	Kont	-
Vikt(hel)	11	Kont	-
Vikt(somatisk)	0,42	Kont	-
Levervikt	19	Kont	-
Gonad	25	Kont	-
Konditionsfaktor(hel)	11	Kont	-
Konditionsfaktor(somatisk)	1,7	Kont	-
Lerversomatiskt index	8,8	Kont	-
Gonadsomatiskt index	14	Kont	-
Fetthalt	0,84	Kont	-
Fiskare	37	Kat	14
Fiskemetod	53	Kat	4
Fångstdjup	57	Kont	-
Ålder(otolit)	70	Kont	
Ålder(fjäll)	78	Kont	-

Sik har fångats på 38 olika platser, fördelat mellan Vänern, Vättern och Bottniska viken. Hur många individer som fångats vid hur många platser visas i Tabell 3.

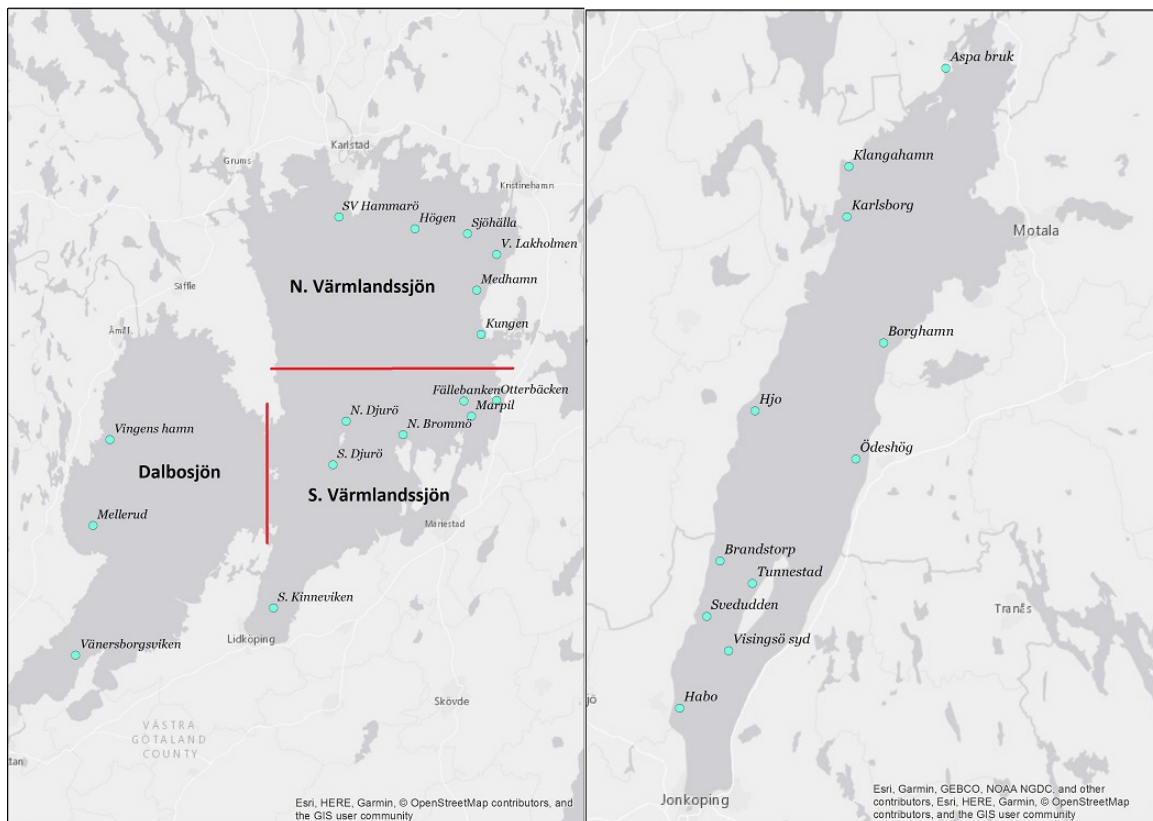
Tabell 3. Frekvenstabell över antal sikar per fångstplats.

Antal individer	Antal fångstplatser
1	12
2	5
3	3
4	4
5	2
6	2
7	2
9	1
10	1
11	1
14	1
17	1
18	1
24	1
38	1

För att få en uppfattning av hur fångstplatserna fördelar sig inom de berörda vattnen visas dessa i Figur 2 och Figur 3. Uppdelningen av Vänern i dess delbassänger ligger till grund för Figur 6.



Figur 2. Fångstplatser i Bottniska viken.



Figur 3. Fångstplatser i Vänern och Vättern. För Vänern har de delbassängerna Norra Värmlandssjön, Södra Värmlandssjön och Dalbosjön markerats.

Metod

Den statistiska metod som använts för att analysera samband mellan dioxinlika ämnen och ett antal av de förklarande variablerna i Tabell 2 är en *generaliserad linjär mixed modell* (Generalized Linear Mixed Model). Metoden kan ses som en utbyggnad i två steg av en vanlig multipel regressionsmodell. Ett av stegen är att tillåta att den beroende variabeln har en annan fördelning än normalfördelad (som är ett av antagandena i en regressionsmodell). Det är det generaliserande steget. Den beroende variabeln antas här vara gamma-fördelad. Observationer från en gamma-fördelning kan bara anta positiva värden, och variansen ökar med ökande värden, något som är lämpligt för analys av dioxinlika komponenter, där spridningen förväntas vara större för större uppmätta TEQ-värden och värdena kan inte vara negativa. Det andra steget är att inkludera fler stokastiska (slumpmässiga) komponenter i modellen. Vanligtvis anses feltermen i en regressionsmodell vara en normalfördelad slumpvariabel, och det är den enda slumpmässiga komponenten. Här anses även varje fångstplats vara en realisation av en normalfördelning, på grund av att det stora antalet fångstplatser och att de är att anse som slumpmässigt valda. Fångstplats kan därför benämnas som en *random-effect*. Variablerna i en vanlig regressionsmodell benämns som *fixed-effects*. Här finns alltså båda typerna med, därav benämningen "mixed".

Även en modell med fetthalt som beroende variabel anpassades. Anledningen är att, som framkommit i tidigare stycken, *Säsong*- och *Vatten* kan ha variation i fetthalt som bakomliggande orsak. Ett kompletterande sätt att undersöka detta är att konstruera en modell med fetthalt som beroende variabel. Modellen med fetthalt som förklarande variabel är en *generaliserad linjär modell*.

Länkfunktion

När den beroende variabeln inte längre anses normalfördelad måste en länkfunktion användas för att linjärisera modellen. Den enklaste varianten av en mixed modell har en fixed-effect och en mixed-effect och den beroende variabeln antas fortfarande vara normalfördelad. I matematiska termer ser den ut som:

$$y = \beta_0 + U_j + \beta_1 X + \varepsilon$$

Där feltermen, ε , antas normalfördelad med medelvärde 0 och varians σ^2 . Det skrivs i statistiska termer som $\varepsilon \sim N(0, \sigma^2)$. Här är U en random-effect som även den är normalfördelad, fast med variansen τ^2 . Indexet j skulle här ha 38 nivåer (en för varje fångstplats) och alla nivåer ses som realisationer av en normalfördelning med väntevärde noll och varians τ^2 .

När y inte längre anses normalfördelad modelleras istället väntevärdet av den beroende variabeln $E(y)$, och detta väntevärde är föremål för en länkfunktion $g(\cdot)$. På grund av att väntevärden modelleras inkluderas inte feltermen i modellen. Modellen skrivs nu som:

$$g(E(y)) = \beta_0 + U_j + \beta_1 X \quad (1)$$

Här anses alltså toxiciteten av dioxinlika ämnen, y , vara gammafördelad och fångstplatserna, U_j , anses vara realisationer från en $N \sim (0, \tau^2)$ fördelning. Länkfunktionen som kommer användas för alla estimeringar är den logaritmiska, $g(\cdot) = \ln(\cdot)$. De vanligaste alternativen gällande länkfunktion när en gammafördelning antas är: logaritmisk, invers, och identitet. Identitetsfunktionen applicerar ingen transformation. Residualanalys av modellen i ekvation (2) specificerad med de tre alternativen av länkfunktioner visade på att den logaritmiska är mest lämplig.

Modellspecifikation

Modellen som specificerats i ekvation (1) utvecklas här med de variabler och interaktioner mellan variabler som anses intressanta och möjliga (vissa har för stort bortfall och är därför inte lämpliga att inkludera i modellen) att undersöka. Interaktionstermer inkluderas på basis av teoretiska resonemang. Modellen presenteras först, därefter förklaras dess komponenter. De ingående förklarande variablerna och intressanta interaktioner är:

$$\ln(E(y)) = \beta_0 + U_j + \beta_{1l}Vatten + \beta_2Fett + \beta_3Längd.binär + \beta_4År + \beta_5CF + \beta_{6k}Säsong + \beta_{7l}Fett * Vatten + \beta_8Fett * Längd.binär + \beta_{9k}Fett * Säsong \quad (2)$$

$$l = 1,2 \quad k = 1,2,3 \quad \text{och} \quad U_j \sim N(0, \tau^2)$$

Där U_j är den komponent av random-effecten som är associerad med fångstplats j och y är summan av dioxiner, furaner och dioxinlika PCB:er angivet i TEQ pg/g vv.

Huvudeffekter

Kodningen och förklaringen av de fixa effekter som är inkluderade är:

- *Vatten* har tre nivåer: Väneren, Vättern och Bottniska viken. I originaldata finns observationer för Bottenhavet och Bottenviken, men på grund av få observationer för dessa områden har de slagits ihop till "Bottniska viken".
- *Fett* anger fetthalten i procent hos en sik. I analysen har den skalats om för att ha medelvärde noll och standardavvikelse lika med ett.
- *Längd.binär* antar värdet noll om en sik är kortare än 38 cm och 1 om längre än 38 cm.
- *År* anses kontinuerlig, med 2015=1, 2016=2, etc. Variabeln är skalad om till medelvärde noll och standardavvikelse lika med ett.

- CF är konditionsfaktorn, beräknad som $vikt/längd^3$. Variabeln är skalad om till medelvärde noll och standardavvikelse lika med ett. Konditionsfaktorn kan beräknas på hel eller somatisk vikt, i detta fall avses den somatiska.
- *Säsong* har fyra nivåer, en för varje årstid.

De kontinuerliga variablerna har skalats om för att estimeringsproceduren för en generaliserad linjär mixed modell kräver det. Från variablerna i Tabell 2 har högerledet i ekvation (2) konstruerats. Målet är att inkludera både variabler som har låg andel bortfall och att få med variabler som beskriver morfometri, tid- och rumsliga aspekter. Även fetthalt har inkluderats då den förväntas ha stor påverkan på halterna av dioxinlika ämnen i sik. Några alternativ av morfometriska variabler är funktioner av varandra, tex konditionsfaktorn som är en funktion av längd och vikt. Därav behöver inte både längd, vikt, och konditionsfaktor inkluderas. Korrelationen mellan längd och vikt(somatisk) är 0,8. Det är därför lämpligt att bara inkludera den ena. Den tillgängliga spatiala informationen som inte inkluderats är fångstens koordinater. Denna information sammanfaller med fångstplatsens namn, som använts istället.

Interaktionseffekter

De interaktioner som finns med i ekvation (2) förklaras och motiveras nedan:

Interaktion mellan längd och fetthalt

Längd fungerar här som en proxy för ålder (yngre fiskar är generellt sett kortare). Mindre fiskar är generellt sett inte toppredatorer, utan äter i större utsträckning plankton som har lägre halter än fisk som är byte för större fisk. Därav kan mindre fiskar ha en lägre ökning av dioxinlika ämnen för samma ökning i fetthalt som en större fisk. Ett liknande argument ligger till grund för det beslut som tagits om att endast tillåta försäljning av strömming kortare än 17 cm.

Interaktion mellan fetthalt och "Vatten"

Det finns en misstanke om att relationen mellan fetthalt och dioxinlika ämnen varierar mellan de olika vattnen. Det är av intresse att undersöka om olika vatten har olika magnituder av sambandet mellan fetthalt och dioxinlika ämnen. Samma fetthaltsökning kan innebära olika ökning av halter av dioxinlika ämnen för olika vatten.

Interaktion mellan fetthalt och säsong

Fetthalt har visats ha starkt samband med förekomst av dioxinlika ämnen, och när fetthalt och säsong endast är inkluderade som huvudeffekter, kan en säsongseffekt vara svår att upptäcka då den egentligen beror på att fetthalten varierar säsongsvis. I en linjär regression tolkas en säsongseffekt exempelvis som att "givet fetthalten (och andra förklarande variabler) så estimeras den genomsnittliga skillnaden i dioxinlika ämnen vara $\hat{\beta}_x$ för vinter jämfört med höst" (fiktivt exempel), men på grund av att fett varierar med säsong inkluderas interaktionstermen.

Tolkning av resultat av estimering

Här ges en kort beskrivning av hur de tabeller som visar resultat från olika estimeringar tolkas.

Alla nivåer som är satta till referensnivå är inkluderade i interceptet, β_0 . Anledningen är att modellen inte går att estimeras annars. Referensnivåerna är; *Vatten=Vättern*, *Säsong=höst* och *Längd.binär=kort*. Övriga nivåer på kategorivariabler tolkas som en jämförelse mot referensnivån.

Det mest relevanta är om en parameter är signifikant och i så fall om motsvarande parameterskattning är negativ eller positiv. Tecknet anger sambandets riktning. I de fall där en kategorivariabel skattats (Vatten, säsong) och nivåerna är minst tre görs en vidare analys av signifikansnivåer. Den proceduren kommer vara en multipel jämförelse som korrigerar för att signifikansnivåerna ändras när flertalet parvisa test görs. Signifikansnivån som kommer användas genomgående är 5%. Det innebär att nollhypotesen i ett hypotestest förkastas om p-värdet är lägre än 0,05.

Estimerad påverkan från kontinuerliga förklarande variabler på den beroende variabeln görs lämpligast i relativa termer, på grund av att den logaritmiska länkfunktionen används. Dessutom måste det tas i beaktning att de kontinuerliga variablerna är skalade till medelvärde noll och standardavvikelse lika med ett. Tolkningen görs som att en enhets ökning av X i genomsnitt leder till en procentuell förändring av Y på $100 * e^{\hat{\beta}_x / sd(X)}$, där $\hat{\beta}_x$ är parameterskattningen för variabel X och $sd(X)$ är standardavvikelsen av observationerna för denna, innan den skalades om. Standardavvikelseerna för de kontinuerliga variablerna på originalskala visas i Tabell 4. För kategorivariabler korrigeras det inte för någon standardavvikelse, men samma ekvation används. Tolkningen görs däremot inte för en enhets ökning utan i termer av att referensnivån för variabeln skiftar till en annan nivå.

Tabell 4. Standardavvikelse för kontinuerliga variabler i modell (2).

Variabel	Standardavvikelse
Fett	2,43
År	1,01
CF	0,14

Resultat

Här redovisas resultat från estimering av ekvation (2) och varianter av denna. Även resultat från en modell med fetthalt som beroende variabel presenteras.

Estimering av ursprunglig modell

Estimering av modellen i ekvation (2) presenteras i Tabell 5.

Tabell 5. Resultat av estimering av modell (2). Formella tester är Wald t-tester. Kontinuerliga variabler är skalade till medelvärde noll och standardavvikelse lika med ett.

	Estimat & signifikans
Intercept	1,05***
Vatten=Bottniska.viken	-1,84***
Vatten=Vättern	-0,29
Fett	0,32
Längd.binär=lång	0,19**
År	-0,11**
CF	-0,16**
Säsong=vår	0,24*
Säsong=sommar	0,96**
Säsong=vinter	0,35*
Vatten=Bottniska.viken*Fetthalt	0,31
Vatten=Vättern*Fetthalt	0,17
Fetthalt*Längd.binär=lång	0,06
Fetthalt* Säsong=vår	0,17
Fetthalt* Säsong=sommar	0,45
Fetthalt* Säsong=vinter	0,15
Antal obs.	224
Antal grupper: Fångstplats	38
Varians: Fångstplats(intercept)	0,13
Varians: Residual	0,32
***p<0,001, **p<0,01, *p<0,05	

Fokus vid denna estimering är att kunna minska modellen genom att ta bort variabler och interaktioner som inte är signifikanta (som alltså inte kan bidra till att förklara variationer i dioxinlika ämnen i sik). Noterbart är att inga interaktioner är signifikanta. Det beror rimligtvis på att interaktionerna kan vara överflödiga när en logaritmisk länkfunktion används. Om den logaritmiska länkfunktionen linjäriserar sambanden är det rimligt att interaktionerna inte blir signifikanta i estimeringen.

Att interaktionerna inte är signifikanta, tillsammans med argumentationen som belyser att de rimligtvis är överflödiga för denna generaliserade modell, gör att de tas bort ur modellen. Den huvudeffekt som inte är signifikant är fetthalt. Valet faller ändå på att låta den vara kvar på grund av att den är med i alla interaktioner, vilket kan leda till att estimeringen huvudeffekten blir osäker.

Estimering av modell utan interaktioner

Resultat från estimering av modellen i ekvation (2) utan interaktioner visas här. Modellen ser då ut som:

$$\ln(E(y)) = \beta_0 + U_j + \beta_{1l}Vatten + \beta_2Fett + \beta_3Längd.binär + \beta_4År + \beta_5CFS + \beta_{6k}Säsong \quad (3)$$

$$l = 1,2 \quad k = 1,2,3 \quad \text{och} \quad U_j \sim N(0, \tau^2)$$

De huvudsakliga resultaten vad gäller svaren på frågeställningarna görs från denna estimering i kombination med resultaten från analys av hur fetthalten påverkar andra variabler. De fixa effekterna förklaras först, därefter "random effect" (Fångstplats). Resultaten syns i Tabell 6.

Tabell 6. Resultat av estimering av modell (5) utan interaktionstermer. Formella tester är Wald t-tester. Kontinuerliga variabler är skalade till medelvärde noll och standardavvikelse lika med ett.

	Estimat & signifikans
Intercept	1,07***
Vatten=Bottniska.viken	-1,79***
Vatten=Vättern	-0,33
Fett	0,54**
Längd.binär=lång	0,20**
År	-0,11**
CF	-0,13**
Säsong=vår	0,19
Säsong=sommar	0,73***
Säsong=vinter	0,29
Antal obs.	224
Antal grupper: Fångstplats	38
Varians: Fångstplats(intercept)	0,14
Varians: Residual	0,32
***p<0,001, **p<0,01, *p<0,05	

Fixa effekter

Jämfört med tidigare estimering är fetthalt nu signifikant. Estimatet är positivt vilket indikerar att ökande fetthalt leder till ökande halter av dioxinlika ämnen. Det överensstämmer med tidigare uppfattning om fetthaltens påverkan. Det estimerade sambandet är att om fetthalten ökar med en enhet så ökar halten av dioxinlika ämnen i genomsnitt med $e^{0.54/2.43} = 25\%$.

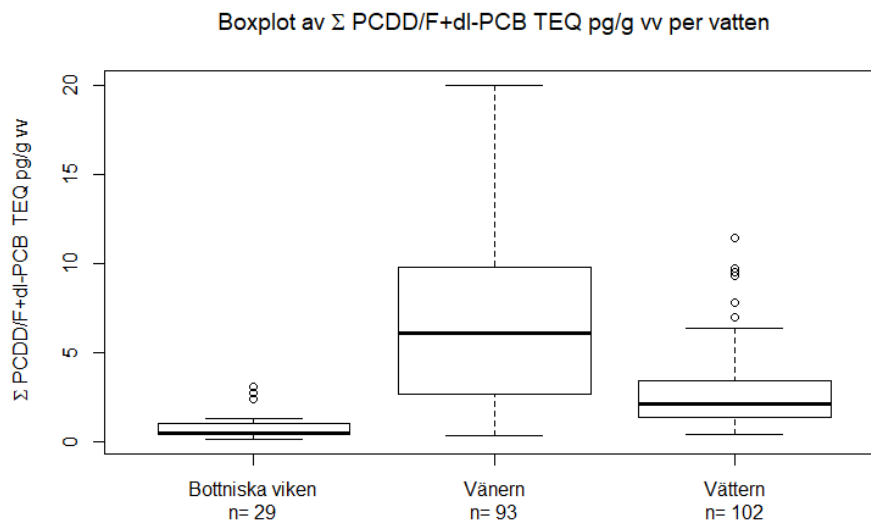
Längd är signifikant, med positivt tecken, vilket indikerar att individer längre än 38 cm tenderar att ha högre halter av dioxinlika ämnen. Detta var den väntade riktningen. Den estimerade skillnaden är att sik längre än 38 cm i genomsnitt har $e^{0.20} = 22\%$ högre halter än de korta.

Även "År" är signifikant, men med negativt tecken, vilket tyder på en generell minskning av halterna med tiden. Den tidigare uppfattningen är att halterna minskar med tiden. Estimeringen indikerar att för varje år minskar halterna med i genomsnitt $1 - e^{-0.11/1.01} = 10\%$.

Konditionsfaktorn, CF, är signifikant med negativt tecken, vilket tyder på att individer som har låg vikt i förhållande till sin längd har högre halter. Konditionsfaktorn samvarierar med fetthalten (korrelation=0,5) och som nämnts görs tolkningen av den signifikanta "CF" utifrån att fetthalten hålls konstant. För två sikar som endast kan särskiljas på grund av olika konditionsfaktor, tycks konditionsfaktorn ha en signifikant påverkan på halter av dioxinlika ämnen. En praktisk implementering av detta blir komplicerad då fetthalten är okänd. Vad gäller morfometri anses det mer lämpligt att förlita sig på skillnader i längd, än på skillnader i konditionsfaktor, men sambandet mellan konditionsfaktorn och fetthalt kommer undersökas närmare i stycket "Statistisk modellering av fetthalt". Tidigare uppfattningar om konditionsfaktorn är inte tydliga i det fall modellen kontrollerar för skillnader i fetthalt.

Skillnader mellan olika vatten

Hur halterna varierar mellan olika vatten analyseras här med multipla jämförelsetester. Tukey's korrigerad används. Detaljerad information om multipla jämförelser finns i ett urval av statistisk litteratur, exempelvis (Montgomery, 2017). De ursprungliga signifikanserna för olika vatten i Tabell 5 görs mot referensnivån som är Vättern. Tukey's test används för att testa alla kombinationer av vatten och för att hålla signifikansnivån på 5-%, tabell 7 visar resultatet. Nollhypotesen, H_0 , är att ingen skillnad mellan två olika vatten finns, mothypotesen, H_1 , är att skillnad finns. Testet kompletteras med en boxplot över halter per vatten (Figur 4).



Figur 4. Boxplot av dioxinlika ämnen uppdelat på vattendrag. "n" indikerar antalet observationer per vatten.

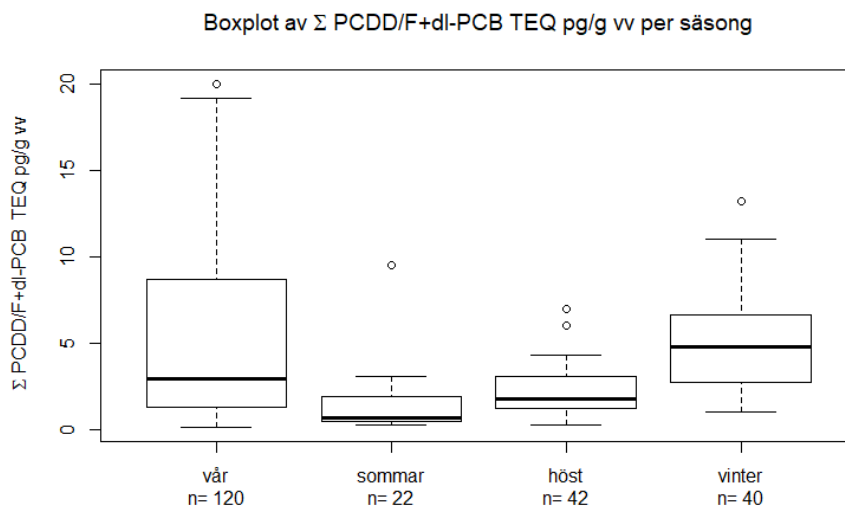
Tabell 7. Jämförelse mellan olika vatten. Tukey's multipla jämförelse används. 5% signifikansnivå.

H_0	Estimat	Standardfel	p-värde
Bottniska viken-Vänern=0	-1,79	0,29	<0,001
Vättern-Vänern=0	-0,33	0,22	0,31
Vättern-Bottniska viken=0	1,47	0,30	<0,001

Testerna indikerar att Bottniska viken har generellt sett lägre halter av dioxinlika ämnen än Vänern och Vättern. Ingen skillnad hittades mellan Vänern och Vättern, trots att tidigare uppfattningar är att en sådan skillnad ska föreligga. Anledningen kan vara att det är skillnader i fetthalt mellan dessa vatten som ligger bakom skillnaderna i dioxinlika ämnen. Denna modell, som kontrollerar för skillnader i fetthalt, kan därför inte påvisa en skillnad i dioxinlika ämnen mellan Vänern och Vättern. För att utreda frågan om det är fetthalt som styr skillnader mellan olika vatten avseende förekomst av dioxinlika ämnen estimeras i en modell med fetthalt som beroende variabel (stycke "Statistisk modellering av fetthalt"). Estimaterna här indikerar att Bottniska viken har ca 3 till 5 ggr lägre halter än Vänern och Vättern. Notera att denna skillnad är oberoende av skillnader i fetthalt. Tas skillnader i fetthalt i beaktning är skillnaden förmodligen än större.

Skillnader mellan olika säsonger

Variationen mellan olika årstider testas också med hjälp av Tukey's multipla jämförelse. Anledningen är samma som för "Vatten"-variabeln. Resultatet visas i Tabell 8. De formella testerna kompletteras med en boxplot över halter av dioxinlika komponenter fördelat på de olika årstiderna (Figur 5).



Figur 5. Boxplot av dioxinlika ämnen uppdelat per säsong. "n" indikerar antalet observationer per årstid.

Tabell 8. Jämförelse mellan olika årstider. Tukey's multipla jämförelse används.

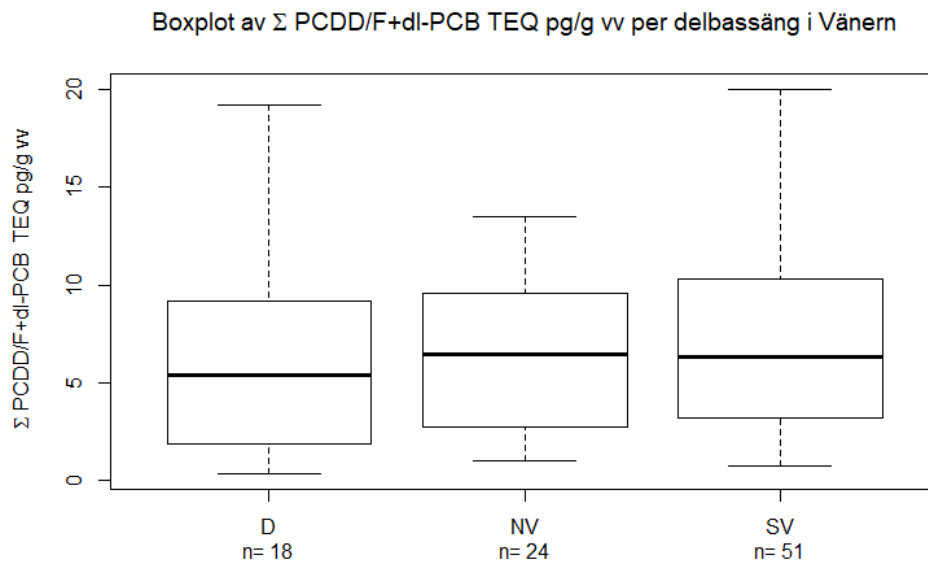
H_0	Estimat	Standardfel	p-värde
Vår-höst=0	0,19	0,11	0,28
Sommar-höst=0	0,73	0,20	0,0019
Vinter-höst=0	0,29	0,16	0,21
Sommar-vår=0	0,54	0,21	0,053
Vinter-vår=0	0,10	0,16	0,91
Vinter-sommar=0	-0,44	0,25	0,27

Testerna hittar endast en signifikant skillnad mellan höst och sommar, vilket innebär att det inte finns några signifikanta skillnader mellan några andra kombinationer av årstider. Tilläggas bör att boxploten visar på att halterna är lägst sommartid, men tittar man på den estimerade differensen mellan sommar och höst så är denna positiv, vilket skulle innebära att sommaren har generellt sett högre halter av dioxinlika ämnen. Förklaringen ligger i att modellen kontrollerar för skillnader mellan vatten och fångstplatser, vilket innebär att den tar i beaktning att inga sikar fångats under sommaren i Väneren, och på grund av höga halter i just Väneren kan modellen kompensera för detta. Den estimerade nivån sommartid kan därför vara högre än den som syns i boxploten. Många observationer sommartid kommer dock från samma fångstplats och det är få observationer för denna årstid. Den signifikanta skillnaden mellan sommar och höst bör därför inte tas som sann. Vid vilken årstid sikfisket bedrivs tycks alltså ha marginell påverkan på halter av dioxinlika ämnen, men även här kan skillnader i fetthalt vara det som påverkar skillnader mellan årstider. Det utreds i stycket "Statistisk modellering av fetthalt".

Fångstplats

Informationen om var, inom ett vatten, en sik fångats har inkluderats som en "random effect". Det beskrevs tidigare, men upprepas här, att konsekvensen blir att estimeringen sker i termer av hur mycket variation den sammanlagda informationen om fångstplatser bidrar med, till skillnad från att få ut en estimerad nivå av dioxinlika ämnen per fångstplats. Formella test av varianser testar nollhypotesen att variansen är lika med noll, och mothypotesen att den är större än noll. Om testet visar att variansen är noll (eller väldigt nära noll) behöver inte "Fångstplats" inkluderas i modellen, annars bör den behållas. Det tests som används här är ett *likelihood-ratio test*. När det används här visar det att "Fångstplats" bör behållas i modellen, vilket betyder att fångstplatsinformationen bidrar till att förklara variationer i dioxinlika ämnen i sik (testet presenteras inte). Det bör dock nämnas att testet har vissa tveksamheter som går att läsa mer om i (Waldetoft, 2019).

Ett annat angreppssätt för att få en uppfattning om hur fångstplatser är associerade med halter av dioxinlika ämnen är att jämföra den estimerade variansen för fångstplats och den återstående variansen som inte kan förklaras av någon av modellens komponenter. Den estimerade variansen som fångstplatsinformationen kan förklara är 0,14 och den varians som varken de fixa effekterna eller fångstplatsinformationen kan förklara är 0,32. Den totala variansen som inte kan förklaras av fångstplats eller de fixa effekterna är därför $0,14 + 0,32 = 0,46$. Det innebär att det har estimerats att $0,14/0,46 = 30\%$ av variationen som de fixa effekterna inte kan förklara, förklaras av fångstplatsinformationen. Det är svårt att säga om det är mycket eller lite, men resonemanget kan kompletteras med en boxplot över dioxinlika ämnen i siken i Vänern uppdelat per delbassäng Figur 6. Fångstplatserna har grupperats ihop till tre delbassänger: Norra Värmlandssjön, Södra Värmlandssjön och Dalbosjön (Figur 3). Spridningen och medianvärdena är väldigt lika för de tre delbassängerna, vilket ger uppfattningen att olika fångstplatser inom ett vatten inte är av stor betydelse för förekomst av dioxinlika ämnen i sik. Dessutom kan standardavvikelsen för variansskattningen jämföras mot de estimat som avser skillnader mellan vatten. Standardavvikelsen för variansskattningen är $\sqrt{0,14} = 0,37$, vilket kan ses som den estimerade genomsnittliga avvikelsen från medelvärdet med avseende på halter av dioxinlika ämnen för fångstplatser inom ett vatten. Jämförs detta mot tex skillnaden mellan Vänern och Bottniska viken på 1,47 så är den betydligt större. Detta resonemang förstärker uppfattningen att val av vatten är av större vikt än val av fångstplats inom ett vatten.



Figur 6. Boxplot över dioxinlika ämnen per delbassäng i Vänern. D=Dalbosjön, NV=norra Värmlandssjön, SV= södra Värmlandssjön.

Statistisk modellering av fetthalt

En modell med fetthalt som beroende variabel estimeras för att undersöka fetthaltens samverkan med olika variabler. *Vatten*, *Säsong*, *Längd.binär*, *CF* och *År* är förklarande variabler. Fångstplatsinformationen kommer inte inkluderas. Frågorna kring fångstplats anses tillräckligt besvarade från tidigare estimeringar. När fångstplatsinformationen inte inkluderas finns ingen random-effect med. Därför är modellen en *generaliserad linjär modell*. Fetthalten antas som tidigare vara gamma-fördelad och länkfunktionen är logaritmisk. I estimeringen behöver kontinuerliga variabler inte skalas om. Modellen ser ut som:

$$\ln(E(y)) = \beta_0 + \beta_{1l}Vatten + \beta_{2k}Säsong + \beta_3CF + \beta_4År + \beta_5Längd.binär \quad (4)$$

$$l = 1,2 \quad k = 1,2,3$$

Där y är en siks fetthalt i procent.

Modellen har efter estimering utvärderats och den passar data utan inkludering av fler förklarande variabler. Resultatet från de hypotestester som ingår i estimeringen anses därför trovärdiga. Dessa resultat visas i Tabell 9.

Tabell 9. Resultat av estimering av modell (4). Modellen är en generaliserad linjär modell med fetthalt som beroende variabel.

	Estimat & signifikans
Intercept	0,24
Vatten=Bottniska.viken	-0,60**
Vatten=Vättern	-0,88***
Säsong=vår	-0,18
Säsong=sommar	-0,04
Säsong=vinter	-0,02
CF	1,45***
År	-0,07*
Längd.binär	0,04
Antal obs.	224
***p<0,001, **p<0,01, *p<0,05	

Säsongsvariationer tycks inte ha någon påverkan på fetthalten. Inte heller tyckts det finnas skillnad i fetthalt mellan långa och korta individer. Det tycks även att fetthalten i sik minskar med tiden. En modell utan de icke-signifikanta variablerna estimeras nu för att få bättre träffsäkerhet i de signifikanta variablerna. Resultatet syns i Tabell 10.

Tabell 10. Estimering av modell (4) utan icke-signifikanta variabler.

	Estimat & signifikans
Intercept	0,19
Vatten=Bottniska.viken	-0,53***
Vatten=Vättern	-0,84***
CF	1,38***
År	-0,06
Antal obs.	222
***p<0,001, **p<0,01, *p<0,05	

Konditionsfaktorn tycks ha en påverkan på fetthalten. Gällande konditionsfaktorn kan alltså sägas att den har en samvariation med fetthalt. Hög konditionsfaktor är associerat med hög fetthalt, vilket inte är förvånande.

”År” är inte längre signifikant, varför det nu inte verkar som om fetthalten skulle minska med tiden, även om variabeln var signifikant i estimeringen av modell (4). Nämnas bör dock att p-värdet ligger nära gränsen för signifikans.

En multipel jämförelse genomförs för att tydligare utreda skillnader mellan Vätern, Vättern och Bottniska viken med avseende på fetthalt i sik. Resultatet visas i Tabell 11.

Tabell 11. Jämförelse mellan olika vatten med avseende på fetthalt. Tukey’s multipla jämförelse används.

H_0	Estimat	Standardfel	p-värde
Bottniska viken-Vätern=0	-0,53	0,13	<0,001
Vättern-Vätern=0	-0,84	0,09	<0,001
Vättern-Bottniska viken=0	-0,31	0,13	0,057

Skillnaden mellan Bottniska viken och Vättern, med avseende på fetthalt är inte signifikant, medan de andra kombinationerna av vatten är signifikanta, vilket innebär att analysen visar på att Vätern har högst fetthalt. Den estimerade skillnaden är att siken från Vättern har $1 - e^{-0.84} = 57\%$ lägre fetthalt än siken från Vätern. Det är alltså högst troligt att det finns en signifikant skillnad gällande dioxinlika ämnen mellan Vätern och Vättern, men att den i själva verket beror på en skillnad i fetthalt.

Denna del avslutas med ett kort räkneexempel med avseende på skillnaderna i dioxinlika ämnen mellan sik från Vätern och Vättern: mediansiken från Vättern har en fetthalt på 1.3% och en medianhalt dioxinlika ämnen på 2,11 TEQ pg/g vv. Skulle denna sik istället lev i Vätern skulle den förväntas ha en fetthalt på $1,3 * \frac{1}{1-0,57} = 3.0\%$. Den förväntade halten av dioxinlika ämnen blir då $2,11 * 1,249^{(3,0-1,3)} = 3,1$, vilket motsvarar en ökning på 46%. Siffran 1,249 är det estimerade sambandet mellan dioxinlika ämnen och fetthalt från sidan 23. Det uppskattas alltså från dessa data att halten av dioxinlika ämnen för en ”normal” sik är ungefär 46% högre i Vätern än Vättern.

Diskussion

I detta stycke diskuteras resultaten utifrån frågeställningarna som var:

- Hur påverkar morfometriska, tid- och rumsliga förändringar halter av dioxinlika ämnen i sik i Vätern, Vättern och längs den svenska Östersjökusten?
- Förstärks de uppfattningar framkommit hittills i projektet?

Sikens morfometri tycks ha en påverkan på förekomsten av dioxinlika ämnen. Den statistiska modelleringen indikerade att sikar under 38 cm generellt sett har lägre halter av dioxinlika ämnen än sikar över 38 cm. En plausibel förklaring är att den mindre siken generellt är yngre och därmed

under en kortare tid exponerats för dioxinlika ämnen. Liknande samband mellan längd och halter har funnits förekomma i strömming (Karlsson & Hållén, 2019).

Konditionsfaktorn visade sig signifikant med negativt tecken, vilket skulle betyda att sikar med låg vikt i förhållande till sin längd har generellt sett högre halter av dioxinlika ämnen. Utifrån ett kontrollfiskeprogram anses det likväl rimligast att inte ta konditionsfaktorn i beaktning. Anledningen är att konditionsfaktorn var signifikant givet att fetthalt och längd hålls konstanta. I mer praktiska termer betyder det att för två lika långa sikar med samma fetthalt, tycks skillnader i konditionsfaktor ha en betydelse. När sikar fångas är dock fetthalten okänd, så resultatet från estimeringen går implementera rent praktisk. Alternativet är att utnyttja att fetthalt har positiv samvariation med både dioxinlika ämnen och konditionsfaktorn. Samvariationen mellan dioxinlika ämnen och konditionsfaktor är däremot för svag för att vara en bra proxy (korrelationen mellan konditionsfaktor och dioxinlika ämnen i TEQ är 0,34).

Olika säsonger/årstider visade sig ha små skillnader i halter av dioxinlika ämnen. En signifikant skillnad hittades mellan sommar och höst men en argumentation som belyser att det saknas observationer under sommarn från Vänern och att få observationer finns under sommaren gör att slutsatser kring denna signifikans bör dras med försiktighet. Utifrån en tidigare uppfattning om att fetthalten kan variera säsongsvist estimerades en modell med fetthalt som beroende variabel. Resultaten visade inte på något samband mellan säsong och fetthalt. Att rikta sikfisket mot någon/några särskild(a) säsonger anses inte efter denna modellering relevant med avseende på fångstens halter av dioxinlika ämnen. Däremot skulle resultaten kring säsongsvariationer kunna ändras om det var känt till vilken underart en individ tillhör, då olika underarter av sik leker olika tider på året och därav avger rommen (innehållande dioxinlika ämnen) vid olika tidpunkter.

Gällande olika vatten gav modelleringen en indikation att halterna är lägst i Bottniska viken. De generellt sett högre halter av dioxinlika ämnen som påvisats i Vänern jämfört med Vättern (Karlsson, et al., 2017), förklaras efter modelleringen i denna rapport av skillnader i fetthalt mellan de två vattnen, där Vänern har högre fetthalter. Modellerna som konstruerats här kan inte svara på varför fetthalterna skiljer sig mellan olika vatten, utan indikerar bara att skillnader finns.

Gällande olika fångstplatser inom ett vatten, indikerade estimeringen att det i termer av ett kontrollfiskeprogram tycks det spela störst roll i vilket vatten fisket bedrivs, inte vid vilken plats inom vattnet.

De resultat som framkommit angående morfometriska, tids- och rumsliga aspekter begränsades delvis av hur datainsamlingen skett. Vad gäller statistisk styrka, alltså förmågan hitta skillnader som faktisk finns, vore det fördelaktigt om data var mer balanserad. Det skulle innebära att fördelningen av observationer för varje kategori av förklarande variabel skulle vara jämnare. Till exempel varierar antalet observationer per säsong kraftigt, och för Vänern finns inga observationer under sommaren. Dessutom är antalet individer fångade vid varje fångstplats väldigt varierande, vilket gör det svårare att avgöra fångstplatsens betydelse. Positivt är dock att fångstplatserna är jämnt utspridda i respektive vatten.

Gällande ekonomiska aspekter är det av intresse att få ut så mycket information som möjligt för en given budget. Görs datainsamlingen med den statistiska analysen i åtanke kan samma statistiska styrka fås med ett lägre antal observationer och därmed lägre kostnader. Det bör dock nämnas att möjligheterna för detta, i projektet "Dioxiner i fet fisk – hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske", är begränsade. Fisk skickas in av lokala yrkesfiskare och om data ska bli mer balanserad krävs att fiskarna får instruktioner på förhand om var och när de ska bedriva fisket. Rimligheten i detta får diskuteras utanför denna rapport.



En avslutande kommentar är att fler sikar analyseras kontinuerligt och kommer läggas till de data som ligger till grund för denna rapport. Data kommer även att kompletteras med individernas ålder, vilket är en variabel av intresse. Därav kan resultat och slutsatser i denna rapport komma att uppdateras framöver.

Referenser

- Cantillana, T. & Aune, M., 2012. *Dioxin-och PCB-halter i fisk och andra livsmedel 2000-2011*, u.o.: Livsmedelsverket.
- European Commission, 2011. Commission Regulation (EU) No 1259/2011 of 2 December 2011 amending Regulation (EC) No 1881/2006 as regards maximum levels for dioxins, dioxin-like PCBs and non dioxin-like PCBs in foodstuffs. *Off J Eur Union*, Issue Off J Eur Union 320, pp. 18-23.
- Hållén, J. & Karlsson, M., 2018. *Dioxiner i sediment från Vänern och Vättern*, u.o.: IVL-rapport B-2310.
- Karlsson, M., 2016. *Kontrollprogram för sikfisket i Vänern och Vättern*, u.o.: IVL.
- Karlsson, M. et al., 2017. *Dioxiner i fet fisk- hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske*, s.l.: IVL rapport B2310.
- Karlsson, M. & Hållén, J., 2019. *Dioxiner i fet fisk - Hot och utvecklingsmöjligheter för svenskt småskaligt kust- och insjöfiske. Årsrapport 2018*, Stockholm: IVL pm:2019-05-17.
- Karlsson, M. & Malmaeus, M., 2014. *Optimerat utnyttjande av lax och strömming - förstudie med förslag till provtagningsprogram*, Stockholm: IVL rapport B 2211.
- Malmaeus, M., Karlsson, M. & Rahmberg, M., 2012. *Bottensedimentens roll för dioxinsituationen i industrirecipienter*, u.o.: IVL-rapport B2053.
- Montgomery, D. C., 2017. *Design and analysis of experiments*. 9:th red. Singapore: John Wiley & Sons.
- Waldetoft, H., 2019. *Dioxin and PCB content in European Whitefish - a generalized linear mixed model*, Stockholm: Masters thesis in statistics, Stockholm University.

