

**Risikovurdering i havbruk
ved bruk av
Bayesianske nett**

av

Kristian Johansen

Masteroppgave ved Institutt for informasjon- og medievitenskap

Universitetet i Bergen

Høst 2010 og Vår 2011



Forord

Før jeg avslutter denne masteroppgaven for godt, vil jeg gjerne takke dem som har hjulpet meg å gjennomføre denne. Først av alt vil jeg gjerne takke Bjørnar Tessem for at han har vært min veileder, og for råd og tips han har kommet med både under utviklingen av modeller og selve skrivingen av oppgaven.

Jeg vil også sende en stor takk til Fiskeridirektoratet for alle de midler jeg har fått tilgang til gjennom dem. De har gang på gang stilt seg til disposisjon for meg når jeg har trengt folk til intervjuer eller undersøkelser. Jeg har også fått låne et kontor av dem slik at jeg har kunnet oppholde meg i nærheten av dem gjennom oppgaven. Jeg vil også takke for at jeg har blitt invitert på konferanser og møter. Uten den hjelpen jeg har fått fra dem hadde denne oppgaven ikke vært mulig å gjennomføre. Jeg vil sende en ekstra takk til Kyst- og havbruksavdelingen, som jeg har hatt mest kontakt med.

Jeg vil takke Rådgivende Biologer AS for at de hadde mulighet til å være med på intervjuer og gi meg et nytt synspunkt på problemstillingene jeg har jobbet med. Hjelp fra dem har gjort at jeg har kunnet utvikle en bedre modell.

Jeg vil også takke min kjæreste, min familie og mine venner som har støttet meg mens jeg har skrevet denne oppgaven.

Bergen, Mai 2011

Kristian Johansen

Forord.....	3
1. Innledning	6
2. Bakgrunn.....	6
2.1 Akvakultur og fiskeoppdrett.....	6
2.1.1 Produksjonsyklus	7
2.1.2 Oppdrettsanlegg.....	8
2.1.3 Rømning av oppdrettslaks.....	9
2.2 Fiskeridirektoratet.....	10
2.2.1 Fiskeridirektoratet og oppdrettsnæringen i Norge.....	10
2.3 Dagens system for risikovurdering.....	11
2.3.1 Den operasjonelle risikovurderingen.....	12
2.3.1 Problemer med den operasjonelle risikovurderingen.....	12
2.3.3 Problemer med for lite data.....	13
2.4 Kunnskapsmodellering.....	14
2.5 Bayesianske Nett.....	15
2.5.1 Risk Assessments in Fisheries Surveillance.....	16
2.5.2 Crime risk factor analysis.....	17
2.5.3 Medical diagnosis.....	18
2.6 Sannsynlighetsteori og Risikoanalyse.....	18
3. Problemstilling.....	19
3.1 Motivasjon.....	19
3.2 Utforming av oppgaven.....	19
3.2.1 Bruk av kunnskapsmodellering.....	20
3.2.2 Bruk av bayesianske nett.....	20
3.2.3 Forskjeller fra arbeidet gjort med tradisjonelt fiske.....	21
3.2 Målet med oppgaven.....	21
3.2.1 Identifisering av de mest utsatte fasene i produksjonen.....	22
3.2.2 Rømning fra sjøanlegg.....	23
4. Metode.....	25
4.1 Design Science.....	25
4.2 Metoder for datainnsamling.....	25
4.3 Teknologier.....	26
4.4 Plan for gjennomføring.....	27
4.4.1 Bygging av modellen.....	27
4.4.2 Evaluering og revidering av modellen.....	29
5. Modelleringsprosessen.....	31
5.1 Kilder fra Fiskeridirektoratet.....	31
5.1.1 AkvaBest.....	31
5.1.2 NYTEK.....	32
5.1.3 Tidligere systemer for risikovurdering.....	32
5.1.4 Rømningsrapporter.....	32
5.2 Interaksjon med ekspertene.....	33
5.3 Gjennomføring.....	33
5.3.1 Første iterasjon.....	34
5.3.2 Andre iterasjon.....	38
5.2.3 Tredje iterasjon.....	41
5.4 Den endelige modellen.....	45
5.4.1 Målbare faktorer.....	45
5.4.2 Hjelpevariabler.....	49
5.4.4 Sammensetning av den endelige modellen.....	56

5.5 Diskretisering.....	58
5.6 Vekting.....	59
5.7 Validering av den endelige modellen.....	59
5.7.1 Validering av struktur og diskretisering av variabler.....	60
5.7.2 Validering av vekting.....	60
6. Bruk av modellen.....	60
6.1 Minimum- og maksimumsverdier.....	61
6.2 Bruk av modellen i et større system.....	61
6.3 Bruk av modellen for sammenligning av anlegg.....	62
6.3.1 Anlegg A.....	62
6.3.2 Anlegg B.....	65
6.3.3 Anlegg C.....	67
6.3.4 Sammenligning av de vurderte anleggene.....	69
7. Diskusjon.....	70
7.1 Kunnskapsmodelleringsprosessen.....	70
7.1.1 Oppbygning av prosessen.....	71
7.1.2 Interaksjon med ekspertene.....	71
7.1.3 Problemer med kunnskapsmodelleringsprosessen.....	72
7.2 Bruk av bayesianske nett for risikovurdering innen havbruk.....	73
7.2.1 Styrker ved bayesianske nett for risikovurdering innen havbruk.....	73
7.2.2 Svakheter ved bayesianske nett for risikovurdering innen havbruk.....	74
7.2.3 Andre implementasjoner av bayesianske nett.....	74
8. Konklusjon.....	75
9. Litteraturliste.....	77
10. Vedlegg:.....	78
10.1 Stor versjon av modellen fra iterasjon 1.....	78
10.2 Stor versjon av modellen fra iterasjon 2.....	79
10.3 Stor versjon av modellen fra iterasjon 3 (Endelig modell).....	80

1. Innledning

Fiskeindustrien er en viktig del av næringslivet. Fisk er en viktig eksportvare for Norge. I følge Statistisk Sentralbyrå ble det eksportert «Fisk, krepsdyr og bløtdyr og varer derav» for 17559 millioner kroner fra januar til april i 2011 (SSB, 2011). En stor og stadig voksende del av disse eksportvaren består av oppdrettsfisk avlet opp i havbruksanlegg. I 2007 var over halvparten av all eksportert fisk fra Norge oppdrettsfisk (SSB, 2007). Hovedsaklig dreier slik oppdrett seg om laks.

Denne delen av industrien er relativt ny i Norge sammenlignet med tradisjonelt fiske. Drift av havbruksanlegg begynte så smått i Norge på 70-tallet, og siden den gang har disse anleggene vokst kraftig i både produksjonsmengde og antall (SSB, 2007). Den kraftige økningen både i antall anlegg og størrelse på det enkelte anlegg fører med seg en del problemstillinger. Industrien har fått en del kritikk på grunn av at fisk som rømmer fra anleggene utgjør en alvorlig trussel for miljøet, og spesielt den norske villaksen.

Det ønskes å forbedre metodene som brukes for å vurdere risiko for rømming på anlegg som benyttes til oppdrett av laks. Målet med denne oppgaven er derfor å utvikle en modell av kunnskapen som finnes rundt dette. Denne modellen vil deretter brukes som et verktøy for å vurdere hvor høy risiko det er for at fisk rømmer fra et anlegg. Oppgaven vil utforske en utforming av en slik modell ved bruk av bayesianske nett.

Denne oppgaven vil først ta for seg bakgrunnskunnskap, og deretter presentere en detaljert beskrivelse av problemstillingen og presentere en plan for hvordan oppgaven vil ta for seg denne. Deretter vil det følge en rapportering av hvordan oppgaven ble utført og en presentasjon av den ferdige modellen. Til slutt vil det være en diskusjon av resultatene av oppgaven etterfulgt av en konklusjon.

2. Bakgrunn

Jeg vil nå gå gjennom en del bakgrunnsmateriale for de emner denne oppgaven vil omhandle. Dette bakgrunns materialet vil bli brukt for å stille en målrettet problemstilling for oppgaven og å legge en plan for hvordan disse målene kan nås på best mulig måte.

2.1 Akvakultur og fiskeoppdrett

Oppdrett av fisk er en type havbruk som kalles akvakultur. Akvakultur er definert som kultivering av organismer i vann. Akvakultur har blitt brukt som en kilde til mat siden 2000år f.Kr. (Rabanal,

1988). Oppdrett av fisk i åpent hav er derimot en nyere implementasjon av denne typen prosess, og nyere teknologier har gjort denne typen oppdrett effektiv og lønnsom ved at det trengs mindre arbeidere for å avle opp et større antall fisk (Rabanal, 1988). Norge har en lang kystlinje og mange steder langs denne er godt egnet til å drive slik oppdrett. Tall fra Fiskeridirektoratet viser at vi i Norge slutten av 2009 hadde 613354 tonn biomasse i produksjon (Fiskeridirektoratet, 2010b).

2.1.1 Produksjonsyklus

Dagens oppdrett av fisk har en bestemt produksjonsyklus, som vi ser gjengitt i figur 1. Denne syklusen består av tre faser: settefiskfasen, matfiskfasen og slaktfasen. Hver av disse fasene foregår på forskjellige steder. Før laksen i det hele tatt blir satt ut i havet avles den frem på land. Denne fasen kalles settefiskfasen. Laks blir avlet frem av befruktete egg i innendørs ferskvannsanlegg. Laksen kalles i denne fasen yngel og holdes i settefiskanlegget helt til den når en vekt på ca 100g. Når laksen når en vekt på 100g kalles den for smolt. Den tåler nå å stå i saltvann og er klar for å settes ut i havet.

Den neste fasen i produksjonsyklusen kalles matfiskfasen. Dette er den lengste fasen i produksjonen. I denne fasen settes fisken ut i de oppdrettsanleggene som finnes i havet. Laksen står i anlegget til den er 2,5 år, og veier da ca 5kg. I denne fasen vil flere større arbeidsoperasjoner bli utført på fisken, som for eksempel avlusning og flytting av fisk. Ved en avlusning helles et legemiddel ned i merden for å fjerne lakselus fra laksen. Dette kan ha konsekvenser for miljøet og er et svært omdiskutert tema, men er en nødvendighet for å sikre at laksen har god helse.

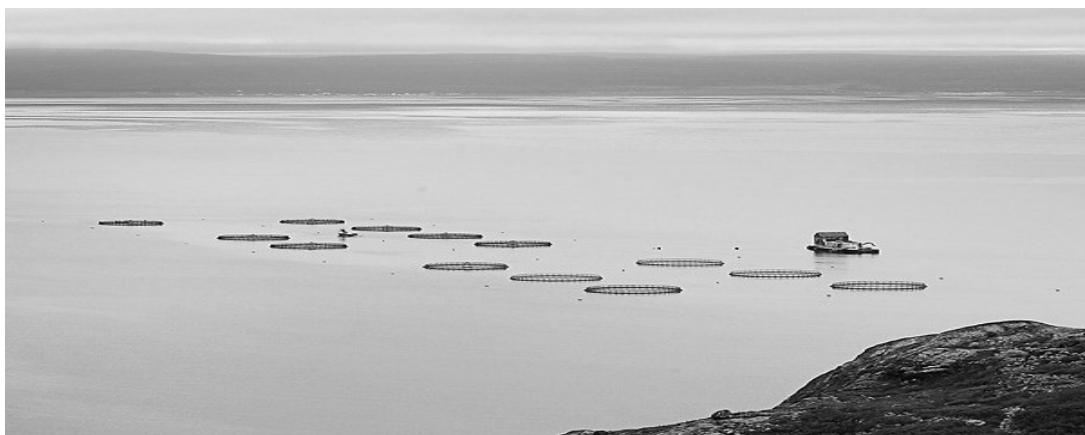
Den siste fasen er slakt av laksen. Laksen slaktes på spesialiserte anlegg som oftest finnes på land. Her blir fisken bedøvet og sløyet av maskiner. Deretter blir den ofte lagt på is og transportert videre for ytterligere prosessering.

Mellom disse fasene (og i tillegg i matfiskfasen dersom fisk skal flyttes) foregår det transport av fisk. For å gjennomføre denne transporten brukes det brønnbåter. En brønnbåt er en båt med et rom som holder fisken i saltvann gjennom transporten. Disse må ha stor kapasitet ettersom dagens oppdrettsanlegg ofte kan romme rundt en million laks. For å få fisken inn og ut av brønnbåten brukes det et pumpesystem som pumper fisken i den retningen som trengs.



Figur 1: Laksens produksjonsyklus

2.1.2 Oppdrettsanlegg



Bilde 1: Et typisk oppdrettsanlegg

Selve anlegget som fisken oppbevares i under matfiskfasen er det stedet hvor fisken tilbringer mest tid. Disse anleggene må planlegges nøye for at fisken skal være sikret, både med tanke på fiskens helse og rømmingssikkerhet. For at nye anlegg skal bli godkjent må de derfor oppfylle en del forhåndsbestemte krav.

Det første som er viktig med et anlegg er lokasjonen. Når et nytt anlegg skal etableres må lokasjonen først grundig undersøkes for å finne ut om den er akseptabel for fiskeoppdrett. En akseptabel lokasjon må ha tilstrekkelig gjennomstrømming av vann for at fisken skal få nok oksygen, samtidig som at naturkreftene ikke kan være for store. Dybde og temperatur er også viktig for fiskens helse. Det settes også krav om avstand mellom oppdrettsanlegg for å forhindre sykdommer og parasitter som kan transporteres mellom disse via strømmer i vannet eller andre smitteveier.

Et typisk oppdrettsanlegg består vanligvis av to typer hovedkomponenter: merder og en flåte for før til fisken, kontorer og lignende. En merde består av en not som holder fisken inne. Noten holdes oppe av en flytekrage. Begge disse delene av anlegget er fortløyd og loddet ned for å holde dem i den tiltenkte posisjonen. Flytekragen består vanligvis av plast eller stål. Stålanlegg er rigide mens plasthanlegg er mer fleksible, noe som ofte gjør dem mer egnet til å tåle sterke naturkrefter. Flytekragen har ofte gangveier slik at ansatte ved et anlegg lett kan komme seg innpå merden, og den er ofte dekket av et nett for å forhindre at rovfugler kommer til fisken.

Noten er nettet som holder fisken inne under havoverflaten. Dette nettet er strukket slik at det ligger som en tett pose under flytekragen. Noten blir holdt nede av lodd, men kan i noen tilfeller fremdeles bevege seg en del med strømmen. Nederst i noten er det ofte en hov som kalles en

dødfiskoppsamler. Fisk som dør i merden vil etter hvert synke til bunnen. Dødfiskoppsamleren brukes for å heise opp den døde fisken fra noten.

Flåten brukes til å oppbevare foret til fisken og kalles derfor ofte for en «fôrflåte». Fiskefôr er ofte laget av industrifisk (fisk som ikke spises av mennesker) og diverse næringstoffer, og er formet til pellets. Fisken blir foret ved at slike pellets blir pumpet gjennom rør fra fôrflåten og blåst ut over nøtene. Fôrflåten har også ofte kontorer og andre rom som brukes av de ansatte.

I tillegg til disse komponentene finnes det også en god del annet utstyr som brukes på oppdrettsanlegg. Eksempler på slikt ekstrautstyr er kraner for å heise opp nøter, høytrykksspylere og annet utstyr for vasking, utstyr for overvåkning og diverse annet utstyr for å forenkle eller forbedre driften av anlegget. Utenom dette trengs det også båter for å transportere folk og utstyr mellom fôrflåten og merdene. Det brukes også båter for å utføre mange av arbeidsoperasjonene som utføres ved anleggene.

2.1.3 Rømning av oppdrettslaks

Oppdrettsnæringen i Norge har fått kraftig kritikk fra flere instanser når det gjelder rømning av oppdrettslaks fra oppdrettsanleggene i havet. Rømning er et stort problem i næringen, da dette problemet ofte forekommer. Det eneste som står mellom oppdrettslaksen og det åpne havet er noten, og det er mange forhold som kan føre til at det kommer hull eller rifter og fisk kommer seg ut.

En stor fare med hull og rifter i nøter er at de ofte befinner seg under vannoverflaten, og derfor er svært lite synlige. Det kan derfor gå lang tid før disse blir oppdaget. Et lite hull kan derfor forårsake at mye fisk kommer seg ut. Slike hull oppstår oftest dersom det er noe som gnager på noten, dersom rovdyr prøver å få tak i fisken eller dersom noten kommer i kontakt med en båt. I tillegg til rifter og hull i nøter kan havari på annet utstyr også føre til rømninger.

Rømning av fisk kan forebygges ved at personellet på det enkelte anlegget følger forhåndsregler om sikkerhet og har gode rutiner både når det gjelder arbeidsoperasjoner som utføres på anlegget og daglig drift. Eksempler på slike gode rutiner er jevnlig undersøkelser av utstyr og målinger av forhold på anlegget. Det er også viktig at utstyret er tilpasset de miljømessige påkjennelser det kommer til å bli utsatt for under bruk.

Den største faren ved at oppdrettslaks rømmer er at den kan gå opp i elvene og formere seg med villaksen. Det er ikke fryktet at dette kan utrydde villaks i Norge, men på grunn av at det er en fare for at det genetiske mangfoldet som finnes naturlig hos villaksen kan bli forstyrret. Dette kan skje

på grunn av at oppdrettsfisken ofte er avlet fram med egenskaper som øker dens salgsverdi og derfor har en annen genetisk oppbygning enn villaksen .

Det finnes også andre grunner til å forhindre at laks rømmer fra anleggene. Selv om miljøfaktoren er den viktigste motivasjonen for å forhindre rømning av fisk fra slike anlegg, finnes det selvfølgelig også økonomiske grunner. Desto mindre fisk som rømmer fra anlegget, desto større vil den økonomiske gevinsten for oppdretter være.

2.2 Fiskeridirektoratet

Fiskeridirektoratet er Norges kontrollorgan for havbruk ¹. Direktoratet ble opprettet i år 1900, og har som formål å regulere fiskeri- og havbruksindustrien i Norge. Direktoratet eksisterer under Fiskeri- og kystdepartementet og har som mål «å sikre rammebetingelsene for en lønnsom og bærekraftig fiskeri- og havbruksnæring samt annet marint basert næringsliv». Fiskeridirektoratet har som hovedoppgaver å sette faglige premisser, forvalte bruk av maritime ressurser og å være et offentlig organ som samarbeider med næringen.

Fiskeridirektoratets fungere som en rådgivende støttespiller for Fiskeri- og kystdepartementet, og er med på å utvikle regelverkene og forskriftene som skal benyttes i fiske og havbruk. Disse reglene blir deretter fastsatt av Fiskeri- og kystdepartementet. Fiskeridirektoratet fungerer også som et kontrollorgan for industrien, og har ansvaret for å tilse at de lover og regler som blir satt blir etterfulgt. Direktoratet har et sentralt kontor i Bergen, men har også lokale avdelinger rundt om i landet.

2.2.1 Fiskeridirektoratet og oppdrettsnæringen i Norge

Fiskeridirektoratet er delt inn i flere avdelinger. Hver avdeling har hver sine oppgaver og ansvarsområder. Kyst- og havbruksavdelingen er den avdelingen som har ansvaret for akvakultur og fiskeoppdrett. De har ansvaret for å tilse at havbruk i Norge blir gjort på en måte som er økonomisk og miljømessig bærekraftig. Avdelingen er delt inn i et sentralt kontor som befinner seg ved Fiskeridirektoratets hovedkontor i Bergen og består av ledelsen for avdelingen, og regionkontorer fordelt utover landet.

Avdeling for kyst- og havbruk har ansvaret for tilsynet av oppdrettsanlegg, men på grunn av at næringen er såpass stor er det umulig å utføre tilsyn ved alle kontrollobjektene. De har derfor vedtatt å bruke risikoanalyse for å fordele ressursene de har tilgjengelig for kontrollering mest

¹ <http://fiskeridir.no/>

mulig effektivt. De bruker en todelt vurdering for å komme fram til risikoverdiene for de forskjellige kontrollobjektene. Denne typen vurdering blir foretatt årlig, og de resultater som vurderingen kommer frem til vil gjelde for hele året.

Vurderingen er delt inn i en sentral og en regional (også kalt: ”Strategisk risikovurdering” og ”Operasjonell Risikovurdering”). Den Strategiske risikovurderingen har som formål å finne ut hvilke områder, prosesser eller aspekter ved næringen som har høyest risiko og trenger fokus. Den strategiske risikovurderingen har de siste årene pekt ut rømning av oppdrettslaks som et av de områdene som trenger mest fokus. Denne vurderingen blir gjort på et sentralt nivå, det vil si at den sentrale risikovurderingen er en enkelt vurdering som gjelder for hele landet. På grunn av at rømning utgjør en såpass stor trussel mot villaksen i Norge, har dette lenge vært det temaet med høyest prioritering strategisk sett.

Den operasjonelle risikovurderingen er rettet mot risiko ved enkelte kontrollobjekter. Ved utførelsen av den operasjonelle risikovurderingen blir de faktorer som er funnet ved den strategiske risikovurderingen brukt for å gi de forskjellige kontrollobjektene en verdi som representerer deres relevante risiko for rømning. Denne vurderingen blir gjort på et regionalt nivå, som vil si at de enkelte regionkontorene foretar denne vurderingen for de kontrollobjektene de har ansvar for.

2.3 Dagens system for risikovurdering

Siden 1970-tallet da bruk av akvakultur for alvor begynte å bli brukt i Norge har oppdrettsanleggene blitt både flere og større. På grunn av at bruk av akvakultur i Norge har eksplodert så raskt har innsamling av data fra anleggene ennå ikke blitt prioritert. De data som finnes er som regel kun informasjon om spesifikasjoner, sertifiseringer og rømninger som har skjedd. Mange av forholdene på et anlegg ut over disse er derfor relativt ukjente for direktoratet. Unntaket for dette er observasjoner som har blitt gjort under inspeksjoner og lignende, og slik informasjon finnes ofte bare hos de enkelte inspektørene.

Mangel på informasjon om anleggene, rapportering og loggføring betyr at for å kunne observere et anlegg må inspektørene fysisk være tilstede på anlegget. Dette gjør det vanskelig og tidkrevende å bedømme om et oppdrettsanlegg følger de retningslinjer og standarder som er satt. Som alle andre kontrollorgan har også Fiskeridirektoratet begrensede ressurser, og det kan derfor være en utfordring å effektivt kontrollere alle havbruksanleggene i Norge. Siden direktoratet har for liten kapasitet til å måle alle kontrollobjektene er det derfor en naturlig løsning å bruke risikovurdering for å velge ut de objektene som trenger tilsynet mest.

2.3.1 Den operasjonelle risikovurderingen

Systemet som brukes av Fiskeridirektoratet for operasjonell risikovurdering ved havbruk i dag er et enkelt system der hvert av anleggene blir satt i en risiko-gruppe basert på en vurdering. Denne blir gjort under den operasjonelle risikovurderingen. Dette vil si at den kun blir gjennomført én gang i året. De regionale kontorene har muligheten til å gå inn og manuelt forandre denne vurderingen dersom det trengs.

Metoden som brukes har ingen rigid fast form og kan bli forandret fra år til år. Ved behov har det også blitt endret under selve risikovurderingen. For å utføre risikovurderingen blir det brukt et skjema som er utviklet av regionkontoret selv (i henhold til de faktorene som er funnet i den strategiske risikovurderingen) og inneholder en del variabler som representerer risikofaktorer eller grupper av disse.

Verdiene som kan settes inn i variablene er 0, 1 og 2, henholdsvis «Lav Risiko», «Middels Risiko» og «Høy Risiko». Noen av variablenes verdier, spesielt de som omhandler spesifikasjoner og lignende, beregnes ut fra data om lokasjonen mens andre blir subjektivt vurdert av de som utfører vurderingen. Når alle variablene i skjemaet er fastsatt, summeres risikoverdiene for alle variablene opp og utgjør til sammen den samlede risikoen for lokasjonen. Deretter blir lokasjonen klassifisert som «grønn», «gul» eller «rød» alt etter hvor høy den samlede verdien er.

Det ferdige skjemaet brukes så for å gjøre en vurdering der det blir bestemt hvilke lokasjoner som skal kontrolleres det neste året. Det planlegges også hvilke typer kontroller som skal gjøres på de forskjellige utvalgte anleggene. I tillegg til dette brukes skjemaet dersom det trengs flere kontroller utover året.

2.3.1 Problemer med den operasjonelle risikovurderingen.

Selv om de regionale kontorene manuelt kan oppdatere risikovurderingen gjennom året, kan det tenkes at dette systemet for risikovurdering er for statisk for sitt bruk. Det kan finnes forhold som fører til at risikoen ved et anlegg forandrer seg på et år, for eksempel ved fluktueringer i vær og strømmer eller utskifting av personell.

Noe som også kan bli et problem med denne typen vurdering er at mange av variablene blir vurdert ut fra subjektive meninger. Det er ønskelig at så mange av variablene som mulig blir vurdert ut fra målbare data for å unngå at det blir gjort feilaktige vurderinger. I dette systemet har det allikevel vært nødvendig å bruke subjektive vurderinger, siden det ikke finnes nok innsamlede data tilgjengelig.

Vurderingen har også noen metodiske svakheter. Et eksempel på dette er at i vurderingen blir en kostnadsfaktor direkte lagt til den samlede risikoverdien, det vil si at kostnadsfaktoren blir behandlet på samme måte som risikoer. Å behandle disse datatypene feil kan forårsake feil i beregningen av risiko. En vanlig måte for å beregne risiko er å multiplisere kostnaden med sannsynligheten for en hendelse. Det er derfor mulig at det vil oppnås et bedre resultat i vurderingen dersom konsekvensfaktorene blir multiplisert med den samlede risikoverdien som et eget ledd til slutt i prosessen.

Et annet aspekt ved systemet som kan forårsake unøyaktighet eller feil er metoden som blir brukt for å anslå risikoen. Siden alle faktorene er likevektede og det ikke tas hensyn til konsekvensen som er knyttet opp mot de forskjellige risiko-variablene, er det mulig at systemet returnerer ugyldige risikoer for lokasjonene. På grunn av dette er det også mulig for et kontrollobjekt å komme ut med lav risiko selv om en av variablene har en skyhøy verdi. For eksempel dersom et anlegg har en verdi i toppen av skalaen for én av faktorene, men har lav risiko på de andre vil den høye verdien allikevel ikke ha stor påvirkning på den sammenlagte risikoen. Dette eksempelet illustrerer en farlig situasjon som kan oppstå ved bruk av dagens system.

2.3.3 Problemer med for lite data

Som sagt tidligere har det også vært et problem for Fiskeridirektoratet at det ikke finnes gode kilder til data om forholdene på de enkelte anleggene. De dataene som finnes stammer som regel fra egenrapporteringer fra anleggene selv eller fra tidligere kontroller eller rapporter fra inspeksjoner gjort av direktoratet. Data fra egenrapporteringer er i tillegg vanskelig å verifisere, og dersom slik data blir brukt for å beregne en risikoverdi er det mulig at denne verdien blir upålitelig.

Et eksempel på en situasjon der feilrapportering kan forekomme er at et havbruksanlegg som har fisk ut over deres tillatte kapasitet ville fått bøter dersom de rapporterte den faktiske mengden, og de vil muligens derfor underrapportere mengden. Det vil alltid være en viss sannsynlighet for slik feilrapportering eller at data blir manipulert dersom slik egenrapportering blir brukt.

Rapportene eller notater som stammer fra kontroller gjort av inspektører, eller andre lignende kilder inneholder mer pålitelige data. Problemet med disse er at de som regel ikke er strukturerte nok til å hente ut kvantifiserbar informasjon. Dette vil si at det er veldig vanskelig å generere data til en database eller et system for risikovurdering automatisk fra disse rapportene.

2.4 Kunnskapsmodellering

Kunnskapsmodellering er en prosess som brukes for å formalisere menneskelig kunnskap, og er i dag et vanlig verktøy for å samle inn kunnskap for bruk i intelligente systemer. Målet med prosessen er å samle inn informasjon om et gitt domene for å gjengi dette i en form som kan brukes av datamaskiner for å gjøre beregninger, eller for menneskelig bruk ved å gi en bedre forståelse av de forhold som finnes mellom faktorer i domenet. En kunnskapsmodell kan enten være vinklet mot et spesifikt problem, eller prøve å gjengi domenet som en helhet.

Personen som utfører denne prosessen kalles en kunnskapsmodellør (eng. «Knowledge engineer»). Denne personen kan være en ekspert i domenet selv, men det er ofte vanlig at kunnskapsmodelløren jobber med andre eksperter dersom dette ikke er tilfellet. Kunnskapsmodellørens jobb er å finne informasjonen som er tilgjengelig om domenet, hente denne inn og sette den sammen til modeller.

En kunnskapsmodellingsprosess der det skal utvikles en modell som for eksempel kan være et bayesiansk nett, kan beskrives som å bestå av tre hovedsteg: bygging, evaluering og revidering (Korb and Nicholson, 2010). I det første steget samles informasjonen om domenet inn. Informasjon blir samlet inn ved hjelp av intervjuer og samtaler med eksperter, data som er tilgjengelig og tidligere forskning gjort i domenet. Etter at informasjon og data er samlet inn bruker kunnskapsmodelløren denne for å bygge opp modellen. Kunnskapsmodellen i seg selv er et sett av alle faktorene, variablene og relasjonene mellom disse som er relevante for domenet. Dette settet skal gjenspeile de egenskaper domenet har i virkeligheten.

Etter at modellen er ferdigbygget skal den evalueres. I dette steget samarbeider ekspertene og kunnskapsmodelløren om å gå gjennom modellen for å finne ut hva som fungerer eller ikke. Denne prosessen kan gjøres ved at modellen testes med data for å se om resultatene samsvarer med den virkelige verden, eller den kan gjøres ved at modellen manuelt gjennomgås.

Etter evalueringen tar kunnskapsmodelløren med seg de tilbakemeldinger han/hun har fått og reviderer modellen. Den reviderte modellen kan deretter testes videre eller tas i bruk. Det er ikke uvanlig at disse tre stegene blir gjort i iterasjoner av bygging, evaluering og revidering. Det er heller ikke uvanlig at stegene glir inn i hverandre, slik at iterasjonene blir mer en uformell struktur enn en strengt etterfulgt plan.

Kunnskapsmodellering er godt egnet til intelligente systemer. På grunn av at kunnskapsmodellen er en representasjon av domenet kan denne brukes for å utføre undersøkelser ved å analysere data som finnes tilgjengelig om dette. På denne måten kan kompliserte domener ofte forenkles og gjøres mer forståelige.

2.5 Bayesianske Nett

Bayesianske nett (Korb and Nicholson, 2010) er et rammeverk som brukes for å modellere variabler og de kausale forholdene som finnes mellom disse. Disse nettene bruker sannsynlighetsteori for å beregne hvilken påvirkning de forskjellige variablene i modellen har på hverandre.

Bayesianske nett baserer seg på Bayes' teorem. Bayes' teorem brukes for å beregne sannsynligheten til en hypotese gitt en variabel mengde observerte faktorer. Dette teoremet kan brukes for å beregne sannsynligheter i et nettverk av relaterte risikofaktorer. Dette betyr i praksis at en kan finne ut hvor stor sannsynlighet det er for at en variabel har en tilstand basert på andre observerbare faktorer som inngår i modellen.

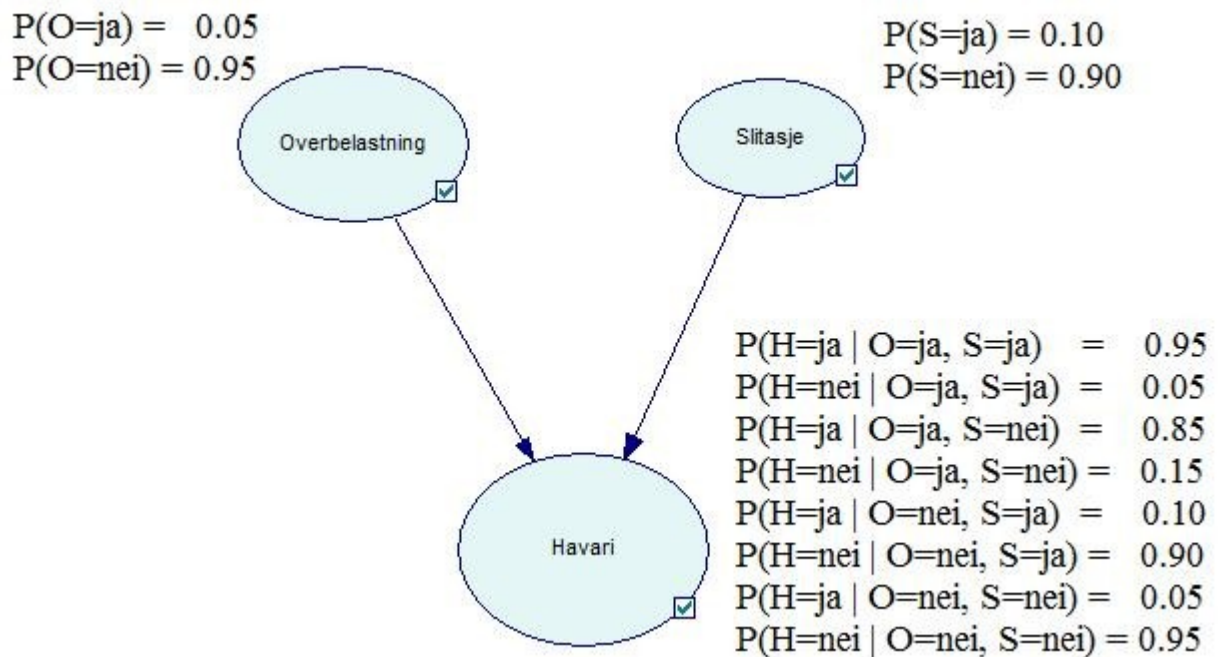
$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}.$$

Formel 1: Bayes' Teorem

Bayesianske nett har vist seg å være effektive for å beregne risiko i situasjoner som er preget av de samme utfordringene som preger oppdrettsnæringen (Pourret et al., 2008). Å organisere kunnskapen som finnes om et domene på den måten som gjøres i et slikt nett kan også være et nyttig referanseverk for ekspertene som jobber med dette.

Et enkelt eksempel på et bayesiansk nett ser vi i figur 2. Her har vi nodene «overbelastning» og «slitasje», som begge påvirker noden «havari». Her vises også de betingede sannsynlighetene hver av tilstandene til nodene. For «havari» ser vi at sannsynlighetene er avhengige av hvilke tilstand nodene som påvirker er i. Ved å bruke Bayes' teorem kan dermed sannsynlighetene for «havari» beregnes ut fra hvor sannsynlige tilstandene til de påvirkende nodene er, dersom tilstandene til disse nodene ikke er kjent.

Det finnes mange eksempler på hvordan bayesianske nett tidligere har blitt brukt for å modellere domener for å utføre risikoanalyser. Jeg vil nå gå gjennom et par av disse for å se på hvordan disse prosjektene har blitt utført, og hvilke resultater de har hatt.



Figur 2: Eksempel på et enkelt bayesiansk nett

2.5.1 Risk Assessments in Fisheries Surveillance

Et av de mest relevante eksemplene på bruk av Bayesianske nett i systemer for risikovurdering kommer fra et prosjekt utført av Fiskeridirektoratet i samarbeid med Professor Bjørnar Tessem ved Universitetet i Bergen (Tessem, 2009; Tessem et al. 2009). Ressursavdelingen ved Fiskeridirektoratet, som har ansvar for de ville ressursene som finnes i havet, jobber med kontrollering av den tradisjonelle fiskeindustrien. Et eksempel på et problem for denne avdelingen er at fiskebåter av og til dumper overflødig fisk dersom de har fylt kvoten. Dette skjer for eksempel ved makrellfiske. Hvis fiskebåten fisker opp annen fisk, dumpes ofte denne til fordel for den mer lønnsomme makrellen. Dette er ulovlig siden slik fisk ofte dør, og dette kan på sikt medføre skade på fiskebestanden i norske havområder.

For å motarbeide dette ville Ressursavdelingen bruke risikovurdering for å finne ut hvilke kontrollobjekter som burde vektlegges med kontroller og lignende. Denne jobben var vanskelig å gjøre manuelt og de vurderingene som ble gjort ble ofte for statiske til å kunne brukes effektivt. Det ble derfor vedtatt at denne risikovurderingen skulle gjøres automatisk, og bayesianske nett ble valgt som metode for å implementere dette systemet.

Tessem jobbet som kunnskapsmodellør i prosjektet. Ut fra intervjuer og samtaler med eksperter utformet de i samarbeid en kunnskapsmodell for domenet. Denne kunnskapsmodellen ble bygget

som et bayesiansk nettverk, og brukes for å beregne risikoer ved fiskebåter i Nordsjøen. Denne automatiske risikovurderingen er dynamisk og virker svært lovende.

Ved hjelp av denne håper Fiskeridirektoratet å lettere kunne opprettholde lovene og reglene for makrellfiske på en mer effektiv måte i forhold til de tilgjengelige ressursene.

2.5.2 Crime risk factor analysis

Roongrasamee Boondao ved Ubon Rajathanee University i Thailand har brukt bayesianske nett for å modellere kriminalitet i Bangkok, Thailand (Boondao, 2008). I dette arbeidet har han vært mest interessert i å se hvilke faktorer som spiller mest inn på én spesifikk kriminell handling, nemlig mord. Motivasjonen hans for dette var å finne ut hvilke faktorer myndighetene burde konsentrere seg om for å motarbeide kriminalitet i storbyen. I analysen som førte til den ferdige modellen brukte Boondao en fem-steps metode for å samle inn de data han trengte.

Denne metoden besto av disse stegene:

1. Identifisering av trekk og mønstre innen kriminalitet.
2. Etablere relasjoner mellom forskjellige faktorer.
3. Fastslå risikonivåer.
4. Gjenkjenning av trekk og mønstre fra data.
5. Forutsi risikofaktorer.

I steg 1-3 ble det brukt informasjon fra tidligere undersøkelser om kriminalitet og intervjuer med eksperter for å fastslå variabler, forhold mellom dem og vektning. Deretter ble det i steg 4 utført maskinlæring der tidligere data ble brukt for å revidere strukturen og sannsynlighetene i nettverket. I steg 5 ble det ferdige nettet brukt for å forutsi hvilke faktorer som hadde størst sannsynlighet for påvirkning på den kriminelle handlingen de var interessert i.

Etter å ha brukt bayesianske nett i dette arbeidet konkluderte Boondao med at det var et godt egnet verktøy for risikoanalyse. Ut fra de resultatene han oppnådde mente han at bayesianske nett i slike situasjoner kan hjelpe å oppnå en fleksibel løsning, som i tillegg er mulig å utvide dersom dette er ønskelig.

2.5.3 Medical diagnosis

Det finnes mange eksempler på prosjekter der bayesianske nett har blitt brukt for å utføre medisinsk diagnose. Diagnose er en prosess som går ut på å finne ut hvilke årsaker som er mest sannsynlig ut fra et sett med symptomer. Jeg har valgt ut et av disse siden diagnostisering og risikoanalyse kan sees på som relaterte disipliner.

Agnieszka Onisko fra Bialystok Technical University i Polen beskriver et system utviklet for diagnose av leversykdommer (Onisko, 2008). Dette prosjektet kalles HEPAR II og bruker et bayesiansk nett for å utføre diagnosen. Dette nettet ble utviklet ved bruk av både eksperter innen domenet og data. Modellen ble også validert ut fra klinisk data som var tilgjengelig for domenet.

Resultatene av å bruke et bayesiansk nett for en slik diagnose var i følge Onisko svært gode. I en test der systemets diagnoser ble testet mot menneskelige legers diagnose, ble det funnet at HEPAR II var dobbelt så treffsikker som de menneskelige legene i gjennomsnitt, og 40% bedre enn den beste av legene.

Onisko nevner at en negativ side med bayesianske nett er at det bruker mange numeriske parametre. Dersom data ikke finnes tilgjengelig for å beregne disse kan det bli vanskelig å parametrisere modellen for å få et fullstendig funksjonelt system.

2.6 Sannsynlighetsteori og Risikoanalyse

Sannsynlighetsteori er et felt innen matematikk. Sannsynligheten for en hendelse formaliseres ved at den blir tilegnet en sannsynlighetsvariabel, et tall mellom 0 og 1, som beskriver hvor sannsynlig det er at denne hendelsen opptrer.

Risiko er relatert til sannsynlighet. Risiko er et ambiguøst begrep som har mange definisjoner, men er vanligvis definert som sannsynligheten for at en eller annen hendelse kommer til å få et negativt utfall. I mange tilfeller økes også risikoen med kostnaden til det negative utfallet. Kostnad er her et generelt begrep på den skaden eller negative påvirkning det negative utfallet fører med seg. En vanlig måte å beregne risiko på er derfor å multiplisere sannsynligheten for at en hendelse skal opptre med denne negative hendelsens kostnad.

Risikoanalyse er et samlebegrep på metoder som brukes for å analysere risikoen for forskjellige handlinger eller konsekvenser av disse. Risikoanalyse kan brukes for å velge handlinger som har større sannsynlighet for å føre med seg gode konsekvenser.

Risikoanalyse brukes også av tilsynsorganer for å plukke ut hvilke objekter som er i en faresone, og

dermed trenger tilsyn. Dette blir gjort ved at alle tilsynsobjektene blir tilegnet en risikoverdi, som sier noe om hvor sannsynlig det er at disse har mangler, og trenger tilsyn. Som sett tidligere kan denne verdien være beregnet kun ut fra sannsynligheter eller de kan være et resultat av en mer komplisert utregning, som for eksempel tar for seg de beregnede kostnadene eventuelle uønskede hendelser eller negative utfall vil ha.

3. Problemstilling

Jeg vil nå presentere problemstillingene for denne oppgaven. Jeg vil først gå gjennom motivasjonen min for å utføre denne oppgaven og hvilken utforming denne oppgaven vil ha. Jeg vil deretter konkret beskrive hva målene med oppgaven er, ut fra det som tidligere er presentert.

3.1 Motivasjon

På grunnlag av de problemene Fiskeridirektoratet jobber med ved Avdeling for Kyst- og Havbruk ønsker de nå å utvikle en bedre metode for å beregne risikoen for rømning av oppdrettslaks fra sjøanlegg. De ønsker å ha en risikovurdering som tar hensyn til så mange av risikofaktorene som mulig og dermed gjøre det enklere å plassere ressurser hos de havbruksanleggene som trenger det mest.

De ønsker også at en slik risikovurdering skal kunne brukes så dynamisk som mulig. Dette vil si at en burde kunne få en oppdatert beregning dersom nye data blir tilgjengelige. Utviklingen av et slikt system vil også være nyttig for å identifisere faktorer som spiller inn på risikoen for at oppdrettsfisk rømmer, og kan være med på å gi en pekepinn for hva slags målinger som må gjøres og hva som må fokuseres på for å forbedre sikkerheten ved slike anlegg.

3.2 Utforming av oppgaven

Denne oppgaven vil være et forskningsprosjekt som følger metoden «Design Science» (Hevner et al., 2004). Formålet med denne vil være å utvikle en modell som representerer kunnskapen rundt rømning av oppdrettslaks. De mest utsatte fasene i produksjonszyklusen vil identifiseres, og modellen vil fokusere på disse.

Arbeidet med å utvikle en modell til bruk for risikovurdering av oppdrettsanlegg vil bestå av å utvikle en kunnskapsmodell for domenet. Denne kunnskapsmodellen vil være i form av et bayesiansk nett, som vil gi modellen den funksjonaliteten den trenger for å kunne utføre en

vurdering av risiko.

3.2.1 Bruk av kunnskapsmodellering

For å kunne utvikle en slik modell som jeg skal lage er det nødvendig å bruke metoden kunnskapsmodellering. I tillegg til å utvikle en modell som skal kunne brukes til risikovurdering ved Fiskeridirektoratet vil denne prosessen være nyttig for å identifisere faktorer som spiller inn på risikoen. Å identifisere disse faktorene er en viktig del av oppgaven.

Kravene for å utføre en kunnskapsmodelleringsprosess er at det må finnes eksperter på domenet som er tilgjengelige for kunnskapsmodelløren, eller at det finnes nok data tilgjengelig slik at prosessen kan utføres uten eksperter. Siden det er usikkert om nok data er tilgjengelig, har jeg sikret meg at eksperter kommer til å være tilgjengelig. Disse ekspertene skal være med på å hjelpe meg å sette meg inn i domenet og å evaluere og validere de modellene som jeg utvikler.

3.2.2 Bruk av bayesianske nett

Basert på de gode resultatene fra ressursavdelingens bruk av bayesianske nett for den samme type risikovurdering har Kyst- og havbruksavdelingen lagt frem et ønske om å utbedre sine systemer for risikovurdering ved bruk av samme teknologi. Det er også dokumentert at bayesianske nett har blitt brukt i lignende situasjoner, og gitt gode resultater (Onisko, 2008; Boondao, 2008; Tessem, 2009).

Som vi har sett tidligere i oppgaven finnes det mange eksempler på situasjoner hvor bayesianske nett har blitt brukt for å modellere kunnskap med et formål om å kunne vurdere enkelte faktorer som forskeren er interessert i. Fra eksemplene ser vi at en slik kunnskapsmodell har flere bruksområder, både for å vurdere hvor stor sannsynlighet det er for at en eller annen tilstand opptrer og å finne ut hvilke faktorer i modellen som har størst påvirkning på domenet i sin helhet.

Å utvikle en modell med disse egenskapene for bruk innen Kyst- og havbruksavdelingen vil gi avdelingen et verktøy som kan brukes til både å vurdere enkelte anleggs risiko, men også på sikt å gjøre en mer strategisk vurdering ved å finne ut hvilke situasjoner som har størst påvirkning på faren for rømning. Dette betyr at bayesianske nett er en fleksibel teknologi som supplerer avdelingen med et verktøy som er nyttig både i en tidlig form og dersom det utvikles videre.

Bayesianske nett bruker sannsynligheter for å vurdere risiko. Det vil si at dersom det ikke finnes mye data kan det brukes estimeringer i stede for direkte statistisk data. Denne egenskapen er svært nyttig for dette prosjektet da det generelt er lite data tilgjengelig for domenet. Dette gjelder spesielt i begynnelsen av utformingen av systemet. Etter hvert som data blir tilgjengelig er bayesianske nett

også et godt verktøy for å bruke maskinlæring for å oppdatere nettets sannsynligheter ved bruk av disse observerte dataene.

3.2.3 Forskjeller fra arbeidet gjort med tradisjonelt fiske

Det finnes mange likheter mellom problemene som ble møtt i ressursavdelingens arbeid (Tessem, 2009). Det er allikevel en god del forskjeller ved problemstillingene og risikofaktorene i Kyst- og havbruksavdelingen og de som ble møtt i Ressursavdelingen. Det er selvfølgelig en faglig forskjell på problemene, men hovedforskjellen ligger i mengden dokumentasjon og datakilder som er tilgjengelige.

Det finnes mindre registrert data for havbruksanlegg enn det gjør for fiskebåter, og de data som finnes for disse er ofte unøyaktige. Dette gjør at kunnskapsmodelleringen kanskje må angripes fra en annen vinkel for Kyst- og havbruksavdelingen enn det som ble gjort for ressursavdelingen. Det må legges mer vekt på å finne de relevante faktorene ved hjelp av intervjuer og samtaler enn å identifisere disse ved å bruke data som finnes fra før.

Når modellen ble utviklet for Ressursavdelingen hadde utviklerne et godt etablert system med kontrollbare data. Fiskebåtenes posisjoner, kvoter og fangster blir monitorert ved hjelp av fangstdagbøker, GPS, sluttseddel ved levering av fangst og lignende metoder. Dette gjør at en har en del synlige variabler som kan brukes i modellen for å beregne hvilke båter som har større risiko enn andre.

For Kyst- og havbruksavdelingen er det i dag et problem at mange av variablene ved et havbruksanlegg som utgjør risikofaktorer ikke er synlige. Med dette menes det at faktorene definitivt utgjør en risiko, men realistisk sett er svært vanskelige å observere. Dette gjelder spesielt faktorer som for eksempel menneskelige feil som kan føre til rømning.

Utviklingen av kunnskapsmodellen kan derfor ikke utføres med et grunnlag i de målte dataene som eksisterer hos Kyst- og havbruksavdelingen. Kunnskapsmodellen bør gjerne heller bygges opp ved å identifisere nye variabler som kan fylles med data fra observasjon av anlegg. Disse nye variablene vil være viktige for å komme fram til en nøyaktig risikovurdering.

3.2 Målet med oppgaven

Det første målet med oppgaven vil være å utvikle en modell som representerer kunnskapen som finnes om risikoer for rømning av oppdrettslaks. Modellen som jeg vil utvikle skal ha som hensikt å kunne brukes som et verktøy for Fiskeridirektoratet. Modellens hovedbruksområde skal være

risikoanalyse av oppdrettsanlegg og vil derfor være vinklet mot dette, men kunnskapsmodellen vil også si noe om det helhetlige bildet når det gjelder kausale forhold som finnes mellom faktorer ved oppdrettsanlegg.

Hensikten med denne kunnskapsmodellen er at den skal kunne brukes av Fiskeridirektoratet til å bygge et rammeverk som kan brukes for å beregne risikoer for hvert av anleggene.

Kunnskapsmodellen vil omfatte forskjellige faktorer som øker eller minsker risikoen for at fisk rømmer fra oppdrettsanlegg. Jeg vil bygge et system som først og fremst baserer seg på de tidligere systemene for vurdering av risiko som er laget av Fiskeridirektoratet, men vil også bruke mye tid på å identifisere flere variabler som kan legges til disse. Dette på grunn av at de tidligere systemene for risikovurdering kan bli sett på som mangelfulle.

Det andre målet med oppgaven er å identifisere de faktorene som spiller inn på risiko for rømning av laks ved sjøanlegg. Dette er et mål på grunn av at Fiskeridirektoratet ikke har mye data tilgjengelig. De planlegger å utvikle en database som skal inneholde innsamlede data fra oppdrettsanlegg, og trenger derfor å finne ut hvilke data som trengs. Å identifisere disse faktorene er en naturlig del av kunnskapsmodelleringsprosessen jeg skal utføre, og vil derfor gli inn i denne.

Produksjonsyklusen for oppdrettslaks er omfattende, og det er mange situasjoner der rømningsepisoder kan oppstå. Jeg vil derfor begrense denne oppgaven til den delen av produksjonen der det oppstår flest rømninger av fisk. Jeg vil på forhånd sette noen rammer for hvilke situasjoner jeg vil fokusere på.

For å finne disse fokusområdene vil jeg undersøke data for å se hvilke situasjoner som har størst sannsynlighet for at det forekommer rømning av laks. Jeg velger å prioritere på denne måten slik at den ferdige modellen skal være mest mulig nyttig for fremtidig bruk. Å modellere hele produksjonen ville blitt for omfattende for denne oppgaven.

3.2.1 Identifisering av de mest utsatte fasene i produksjonen

For å finne ut hvilke rammer jeg vil sette for kunnskapsmodellen har jeg et ønske om å finne ut hvor i produksjonsyklusen det oppstår flest rømninger. Jeg vil altså konsentrere systemet rundt de områdene som har størst frekvens, ikke de som nødvendigvis de som har høyest konsekvens. Dette er på grunn av at jeg vil utvikle et system som i hovedsak beregner risiko for at rømning oppstår, ikke hvilke følger den eventuelle rømningen har.

Det er på forhånd antatt at dette er i matfiskfasen, der fisken står i oppdrettsanlegg i sjø. Jeg vil allikevel samle inn data for å støtte denne antakelsen. Figur 3 er en grafisk representasjon av en

analyse av data som finnes i et dokument som Fiskeridirektoratet har satt sammen av data hentet fra rømningsrapporter. Dette dokumentet er ikke offentlig tilgjengelig. Rømningsrapportene er opprinnelig skrevet på papir, og har manuelt blitt satt sammen i et forsøk på å systematisere de data som er tilgjengelig om rømning.

I dokumentet finnes det også data om antall fisk som rømmer ved hver rømningsepisode, men jeg valgte å ikke ta hensyn til disse dataene i denne analysen. Dette er på grunn av at jeg som sagt i utgangspunktet er mer interessert i å finne ut hvilke situasjoner som har høyest sannsynlighet for rømning. Å finne ut hvilke situasjoner som har størst rømning i antall fisk vil dreie seg mer om å kartlegge hvilke situasjoner som påfører miljø og næring høyest kostnad.

3.2.2 Rømning fra sjøanlegg

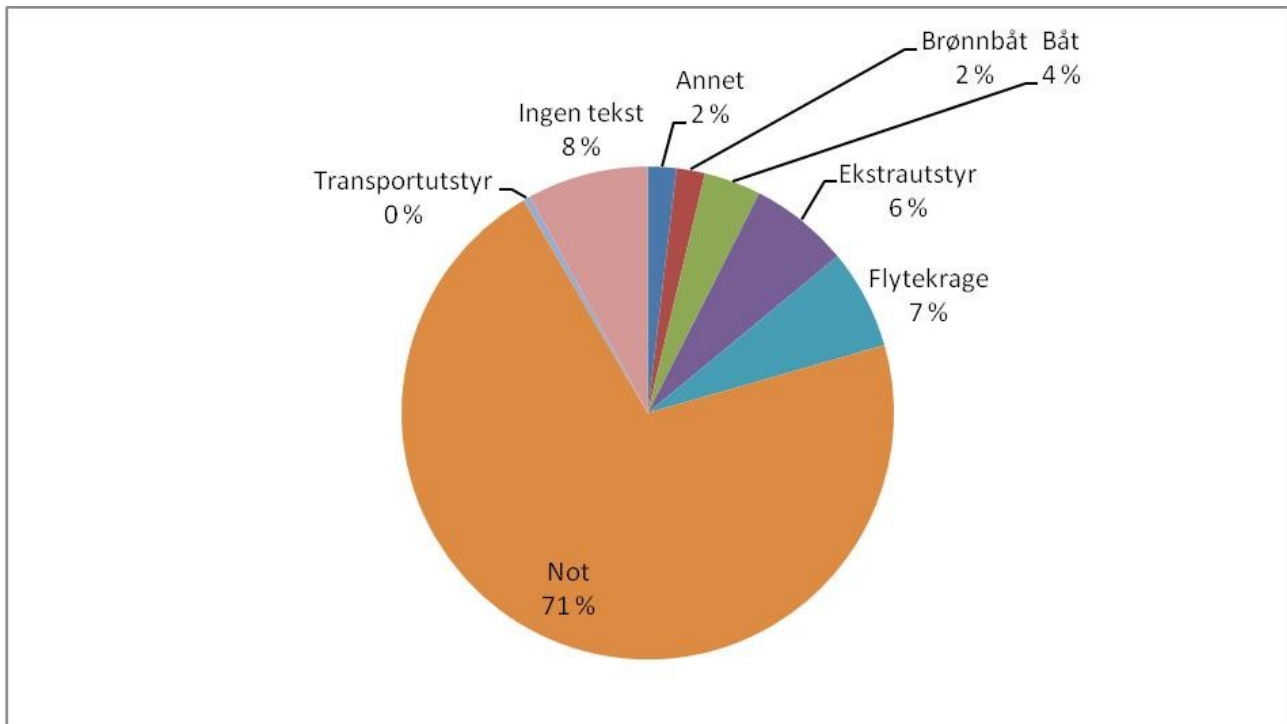
Det er på sjøanlegg de fleste rømninger skjer, når laksen er i matfiskfasen. Ut fra rømningsrapportene fra Fiskeridirektoratet vises det at 79.25% av all rømning som er registrert er fra sjøanleggene. Dette er naturlig siden det er i slike anlegg fisken oppholdes den største delen av tiden. Det er derfor viktig å legge merke til at det er mulig at risikoen for rømning faktisk er større ved andre faser i produksjonen. Det er allikevel her flest fisk blir rapportert rømt, og jeg vil derfor ikke ta hensyn til dette i mine prioriteringer.

Figur 3 viser hvordan rømningsepisodene er fordelt på de ulike komponentene ved et sjøanlegg. Det vises tydelig fra figur 3 at det er nøtene som er mest utsatt. Etter nøtene er det flytekrage som har størst frekvens når det gjelder rømning. Dette er de to største komponentene ved oppdrettsanlegget. Vi ser også at feil i sammenheng med båttrafikk og ekstrautstyr utgjør en god del av tilfellene. Det må også nevnes at de forskjellige kategoriene ikke er eksklusive. For eksempel kan båter komme for nærme anlegget og skade noten med propellen. Denne hendelsen vil da kunne gå inn under to kategorier.

I tillegg til at rømningsrapportene viser at det er størst sannsynlighet for at rømning oppstår i sammenheng med feil eller uønskede hendelser ved not, flytekrage, båttrafikk og ekstrautstyr, er ekspertene også enige i at dette er de viktigste komponentene å sikre. Ekspertene legger til at det ofte kan være et samspill mellom disse komponentene. For eksempel er båttrafikk ikke et problem før det oppstår kontakt mellom båt og not/flytekrage.

På grunnlag av denne analysen og innspill fra ekspertene vil jeg derfor se på kategoriene not, flytekrage og båt som viktigste kategoriene. Det er verdt å notere seg at det er også en god del av disse rapporterte hendelsene som ikke har noen beskrivende tekst på hvilken komponent det er som

feilen har oppstått i.



Figur 3: Rømning av oppdrettslaks fra sjøanlegg

Som nevnt tidligere er rømningsrapportene til Fiskeridirektoratet egenrapporteringer fra oppdretterne selv. Det tas derfor forbehold om kvaliteten på den informasjonen som brukes for å finne disse tallene. Det er flere kilder til feil ved denne typen rapporter. Det er ikke sikkert at oppdretter har fullstendig informasjon om hva som har skjedd eller hvor mange fisk som har rømt. Det er også mulig at oppdretter med vilje manipulerer tall for å unngå bøter eller lignende.

Det fantes også mange rapporter som ikke ble kategorisert etter komponent, da denne informasjonen ikke var inkludert i rapporten. I slike tilfeller ble rapporten tatt ut av statistikken. En siste kilde til feil ved denne type innmelding er at det også er mulig at det er en god del rømminger som ikke blir rapportert i det hele tatt.

Som et resultat av denne undersøkelsen vil jeg fokusere oppgaven min på å modellere risikobildet rundt komponenten not. Jeg vil inkludere arbeidsoperasjoner som blir utført på denne komponenten og båttrafikk rundt denne, både som følge av drift og annen kommersiell trafikk. I og med at noten er festet til flytekragen vil denne også bli tatt med i modellen. Modellen skal inneholde alle de feil og uønskede hendelser som kan oppstå i sammenheng med disse komponentene, og hvilke faktorer som fører til disse.

4. Metode

Jeg vil nå gå gjennom metoden jeg vil bruke for å oppnå målet med oppgaven. Jeg vil se på arbeidets overordnede metode som er design science, hvilke metoder jeg vil bruke for datainnsamling og interaksjon med eksperter på området, og til slutt presentere en plan for hvordan oppgaven skal utføres.

4.1 Design Science

Oppgaven følger en metode som kalles «Design Science». Denne metoden er definert av Hevner et al. (Hevner et al., 2004), og er en resultatorientert metode for forskning der målet er å utvikle en artefakt. Artefakten er typisk et stykke software, en modell, en prototype eller lignende, og skal kunne representere en mulig løsning på et problemområde. For at forskningen skal være meningsfull må artefakten enten presentere en løsning på et problem som ikke tidligere er løst, eller presentere en bedre løsning enn de som allerede finnes.

Forskningsprosessen for et prosjekt som bruker design science vil være arbeidet med å utvikle artefakten. Denne prosessen kan gjerne bestå av iterasjoner av utvikling og evaluering. Ved å bruke iterasjoner blir problemområdet gradvis dekket av artefakten, som til slutt gjennomgår en endelig evaluering for å tilse at den representerer en fornuftig og tilstrekkelig løsning på problemet.

Det gir mening å bruke design science som en metode for dette prosjektet. Det rammeverket for forskning som design science gir et prosjekt passer godt med problemstillingen og de mål som allerede er satt for denne oppgaven. Artefakten som skal utvikles i denne oppgaven er en modell, og denne modellen blir utviklet som en løsning på et spesifikt problem, nemlig risikovurdering for rømning av oppdrettslaks.

4.2 Metoder for datainnsamling

I samsvar med de funn som tidligere har blitt gjort i prosjekter som har brukt kunnskapsmodellering skal jeg bruke intervjuer av eksperter som en kilde til data (Pourret et al., 2008). Jeg vil bruke uformelle eller semistrukturerte intervjuer og samtaler med ekspertene for å hente inn informasjon. Jeg velger å forholde meg til ekspertene på denne måten på grunn av at hverken kunnskapsmodelløren eller ekspertene på forhånd vet nøyaktig hvordan modellen kommer til å formes, og jeg tror at diskusjoner der tema som omhandler modellen blir gjennomgått vil være mest nyttig.

Ekspertene som skal intervjues i min oppgave jobber daglig ved Fiskeridirektoratet, ved Kyst- og

havbruksavdelingen. Jeg vil gjennom arbeidet ha kontakt med flere av de ansatte på avdelingen, både med inspektører som ofte er ute på inspeksjoner ved oppdrettsanlegg, og de som jobber administrativt. Det vil være inspektørene jeg kommer til å forholde meg mest til når det gjelder intervjuer, siden disse ofte vil sitte med detaljkunnskap om forhold på anleggene.

Med å bruke uformelle intervjuer håper jeg å ikke hindre framgangen i prosessen ved å sette en for rigid form på samtalene jeg har med ekspertene. Intervjuene vil heller ha temaer som blir diskutert, med spørsmål jeg har forberedt, og jeg vil lede samtalen i den retningen jeg mener fører til det mest produktive resultatet. Resultatene fra intervjuene vil påvirke hvordan jeg tolker annen data jeg har tilgjengelig og hvordan disse blir satt sammen i modellen.

Jeg har også snakket med Kyst- og havbruksavdelingen om å delta på møter de skal ha hvor relevant informasjon blir diskutert. Dette vil være en god kilde til informasjon som kan benyttes i kunnskapsmodelleringsprosessen.

Jeg kommer også til å bruke data fra dokumenter lagt fram av Fiskeridirektoratet. Disse dokumentene vil først og fremst bli brukt for å etablere de faktorer som finnes for domenet, og hvilke sammenhenger som finnes mellom disse. Denne typen data vil bli brukt som bakgrunnstoff for intervjuene med ekspertene.

Modellen vil med disse typene datakilder være forankret i både ekspertenes kunnskap om domenet, og dokumentasjon som finnes. Dette er i tråd med metoden for kunnskapsmodellering, og kan gi et solid grunnlag for en valid modell.

4.3 Teknologier

Jeg vil bruke et grafisk grensesnitt for å utvikle kunnskapsmodellen for denne oppgaven. Dette grensesnittet er et program kalt GeNIe². GeNIe er et grensesnitt for SMILE, som er et system utviklet for å modellere bayesianske nett. SMILE har også støtte for å gjøre beregninger med slike nettverk. GeNIe gir brukeren muligheten til å utvikle bayesianske nett med et UML-lignende grensesnitt.

GeNIe og SMILE er utviklet av Decision Systems Laboratory, som er en forskningsgruppe fra Universitetet i Pittsburgh. Decision Systems Laboratory forbeholder seg retten til distribusjon av programvaren, men både GeNIe og SMILE er gratis og kan fritt benyttes til det brukeren måtte ønske. Dette gjelder både kommersiell og privat bruk.

GeNIe og SMILE er valgt på grunn av at de har tidligere blitt brukt for å utvikle lignende nettverk,

² <http://genie.sis.pitt.edu/>

og har i disse tilfellene tilført mye funksjonalitet (Tessem, 2009; Onisko, 2008). GeNIe gir muligheten for å rask utvikling av enkle bayesianske nett, men har også støtte for for eksempel maskinlæring. Å velge GeNIe vil gi meg muligheter til å benytte seg av den funksjonaliteten jeg skulle komme til å trenge gjennom arbeidet. Annen argumentasjon for å bruke disse teknologiene er at denne oppgaven ikke har noen økonomisk støtte, og gratis programvare derfor er foretrukket.

4.4 Plan for gjennomføring

Prosessen med å utvikle modellen vil som sagt bestå av iterasjoner av bygging, evaluering og revidering. Dette betyr at hele prosessen vil starte med innsamling av informasjon. Denne informasjonen vil bli samlet inn ved hjelp av de metoder jeg har diskutert tidligere. Når jeg har samlet inn tilstrekkelig med informasjon om domenet vil jeg kunne lage utkast for en modell som deretter kan bli vurdert av ekspertene. Denne vil deretter bli revidert ut fra de synspunkter ekspertene har kommet med.

Grunnen til at modelleringsprosessen ble utført som en iterativ prosess er at det må etableres en felles forståelse av domenet mellom ekspertene og kunnskapsmodelløren. Denne felles forståelsen er noe som oppstår over tid og som et resultat av å jobbe med domenet. Det er derfor gunstig med en iterativ prosess. En slik prosess tillater ekspertene og kunnskapsmodelløren å tilnærme seg en enighet i et miljø der det ikke er en katastrofe dersom det blir gjort feil. Dersom det oppstår uenigheter mellom partene vil disse vanligvis bli oppdaget i evalueringen og kan forbedres i neste iterasjon.

For hver iterasjon vil modellen komme nærmere en akseptabel representasjon av domenet. Til slutt vil jeg sitte igjen med en modell som både jeg og ekspertene mener er en tilstrekkelig representasjon. Denne modellen vil gå gjennom en endelig validering for å sikre at den faktisk er en gyldig representasjon av domenet. Jeg vil nå gå gjennom iterasjonenes planlagte oppbygning gjennom hele arbeidet.

4.4.1 Bygging av modellen

Dette steget er preget av innsamling og vurdering av data. Identifisering av risikofaktorer og relasjonene mellom disse er de sentrale målene med dette steget. Det første som må gjøres er å identifisere hva som mangler i modellen. Dette kan gjøres ved å analysere innsamlet informasjon, gjennom diskusjoner med eksperter eller ved å studere evalueringer av modeller fra tidligere iterasjoner.

Den første iterasjonen i en slik prosess vil være spesiell i og med at kunnskapsmodelløren gjerne ikke har stor kjennskap til domenet. Det vil derfor være viktig for kunnskapsmodelløren å tilegne seg den mest grunnleggende kunnskapen om domenet først for så å bygge videre på denne i senere iterasjoner (Onisko, 2008).

Etter hvert som jeg får et bilde av hvordan domenet er satt sammen, kan jeg begynne å utvikle en formell representasjon av dette. Denne formelle representasjonen består av et sett av faktorer som spiller inn på risikoen for at oppdrettslaks rømmer, og et sett av relasjoner som beskriver hvilke faktorer som påvirker hverandre. Det er også viktig å etablere hva slags påvirkning som finnes her. Den formelle representasjonen blir utviklet i grensesnittet GeNIE som er beskrevet tidligere i oppgaven. En grafisk representasjon av modellen vil gjøre modellens innhold lett å forholde seg til både for meg selv og for ekspertene jeg skal jobbe med.

For å oppnå en oversiktlig struktur i modellen ønsker jeg å dele inn risikofaktorene i grupper. Jeg håper at dette vil gjøre modellen mer forståelig for både ekspertene og meg selv. Jeg håper også at disse gruppene kan brukes som submodeller. På grunn av at det finnes lite tidligere innsamlede data som kan brukes for beregning av modellen, vil det muligens være enklere å innføre modellen til bruk delvis, det vil si en submodell om gangen. Gruppene vil bestå av relaterte faktorer som deler nok egenskaper til å si at en gruppering vil være logisk.

Etter at faktorene og relasjonene mellom disse er etablert, må relasjonene bli vektet før nettet kan brukes for å vurdere risiko. Dette kan komme til å bli en stor utfordring i dette tilfellet. Det er som sagt tidligere ikke samlet inn mye data fra havbruksanlegg, og at de innsamlede dataene stammer fra egenrapporteringer. Det burde likevel være gode muligheter for å samle inn data fra kilder som ikke blir brukt i dag. Mange av de miljømessige faktorene kan kanskje måles ved hjelp av for eksempel værmeldinger eller rapporter om rovdyrbestand. De tekniske faktorene kan tenkes å kunne måles i samarbeid med leverandører (som i Norge kan pålegges å oppgi informasjon om leveranser til havbruksanlegg, i følge ekspertene jeg har snakket med).

Dersom det er mulig å samle nok data vil jeg prøve å bruke maskinlæring for å parametrisere nettet. Jeg ser på dette som usannsynlig basert samtaler jeg har hatt med Kyst- og havbruksavdelingen om mengden data de har tilgjengelig. Alternativet til dette er å bruke ekspertenes estimeringer av sannsynlighetene i modellen. Det finnes mange metoder for å gjøre dette (Korb and Nicholson, 2010), men disse er ofte vanskelige å få nøyaktige uten tilgang på data (Onisko, 2008). Jeg vil allikevel bruke disse metodene for å vekte modellen, om ikke annet for å kunne demonstrere modellens tiltenkte bruk. Vektingen av modellen vil bli gjort i de siste iterasjonene, når strukturen i modellen er tilnærmet ferdigutviklet.

En del av prosessen med å legge vekt til relasjonene mellom faktorene vil være å finne fornuftige diskretiseringer av faktorenes mulige verdier. Det er vanlig å bruke diskrete verdier for variabler i bayesianske nett for å simplifisere bruken av disse. I dette tilfellet vil det sannsynligvis ofte dreie seg om mulige tilstander av sannsynligheter for at et eller annet faremoment er til stede. Slike sannsynligheter kan for eksempel være lav, middels og høy, eller med flere eller færre diskrete verdier. Hvor mange verdier som trengs for en relasjon vil avhenge av hvor gradvis påvirkning den påvirkende faktoren har på den påvirkede.

Samtidig som at modellen skal brukes for å vurdere risiko, er det også et mål at denne skal være med på å identifisere de faktorer som er relevante. Når faktorer blir identifisert vil det derfor vektlegges om det er realistisk å finne data som sier noe om faktorens tilstand. Dersom faktorer blir identifisert, men det ikke finnes noen måte å måle faktorens tilstand, kan disse ikke være med som separate faktorer i modellen. Disse må da kanskje heller modelleres under mer generelle faktorer, avhengig av hva slags type faktor det er snakk om.

Når jeg snakker om faktorer i denne teksten tenker jeg på individuelle, observerbare tilstander eller hendelser ved et anlegg. Dette kan være helt konkrete faktorer som for eksempel temperatur eller strømstyrke, men også mer abstrakte faktorer som fiskehelse. At det oppstår gnag mellom to komponenter kan også være en faktor.

Et eksempel på en slik type tilstand kan være miljøforhold. Faktorer som representerer miljøforhold kan være for eksempel strømstyrke ved et anlegg. Det finnes i dag registreringer om hva slags værforhold det finnes på et anlegg, men disse er ofte mangelfulle og blir per i dag ikke utnyttet til sitt fulle av Fiskeridirektoratet for å gjøre en fullstendig risikoanalyse. I slike tilfeller vil jeg se på hvordan data som er tilgjengelig bedre kan brukes, samtidig som at jeg vil diskutere med ekspertene om hvordan det er mulig å anskaffe data av bedre kvalitet.

Oppbygningen av modellen vil ha tre aspekter: Faktorene og grupperinger av disse, strukturen av relasjoner som finnes mellom faktorene og vektingen av relasjonene mellom faktorene. Denne oppgaven vil som forklart tidligere fokusere mest de to første aspektene, faktorene og relasjonene. Dette på grunn av usikkerheten når det gjelder vektingen av faktorene.

4.4.2 Evaluering og revidering av modellen

Evaluering av artefakten er et viktig aspekt av design science. Hevner et al spesifiserer at “The utility, quality, and efficacy of a design artifact must be rigorously demonstrated via well-executed evaluation methods” (Hevner et al., 2004). Dette vil derfor bli gjort i hver iterasjon, samtidig som

en mer grundig validering vil bli gjort av den ferdige modellen.

Målet med denne evalueringen i iterasjonene er å gjennomgå modellen for å finne ut hva som fungerer og hva som trenger revidering. Denne evalueringen skal i mitt arbeid være uformell og bestå av at modellen diskuteres med ekspertene. I disse diskusjonene vil styrker og svakheter med modellen bli diskutert. Mindre justeringer vil bli gjort på modellen som revideringsteget av iterasjonen. Dersom det finnes større mangler, som overflødige faktorer eller relasjoner, eller faktorer eller relasjoner som mangler, vil disse bli tatt hensyn til i neste iterasjon.

Når jeg har en ferdig versjon av modellen som ser ut som om den kan virke i praksis vil jeg utføre en endelig validering av denne. Denne valideringen vil bli gjort mer formell enn de evalueringene som blir gjort mellom iterasjonene. Det er flere sider av modellen som må valideres. Først vil jeg validere strukturen i modellen, deretter vil jeg validere diskretiseringen av variablene og dersom det lar seg gjøre vil jeg til slutt validere vektingen. Jeg vil dele modellen inn slik og gjøre tester hver for seg for å finne ut om alle sider ved modellen er korrekte. For å gjøre disse testene vil jeg bruke ekspertene på Fiskeridirektoratet.

Når jeg snakker om strukturen i modellen mener jeg relasjonene mellom variablene. Dette er det aspektet av modellen som kommer til å stå ferdig først, og det er derfor naturlig å begynne med å validere denne. Det er også et svært viktig aspekt, siden resten av modellen bygger på strukturen. Måten jeg vil validere strukturen i modellen er ved å intervju et fåtall eksperter, vise dem en liste med variablene i modellen og deres relasjoner, og deretter be dem å svare på om de relasjonene jeg har funnet faktisk eksisterer i virkeligheten. Jeg vil etterpå sammenligne svarene fra de forskjellige ekspertene og se om det er samsvar mellom meningene deres. Ut fra opplysningene jeg får fra disse intervjuene vil jeg kunne si noe om validiteten til strukturen i modellen min.

Valideringen av diskretiseringene vil skje på samme måte. Forskjellen her vil være at valideringen vil gjøres mens jeg etablerer diskretiseringen. Jeg vil gjøre intervjuer der jeg ber ekspertene sette de mulige tilstandene for variabler i tilfellene der dette trengs. Jeg vil være med på dette for å passe på at de diskretiseringene som blir gjort er hensiktsmessige i forhold til modellen. Jeg vil gjøre dette uavhengig med forskjellige eksperter for å kunne sammenligne forskjellige meninger etterpå. Dersom det er stor uenighet om kategoriseringer vil jeg holde nye møter, gjerne med flere eksperter til stede der disse diskuteres.

Valideringen av vektingen i modellen vil være vanskelig å få nøyaktig på grunn av at det sannsynligvis finnes for lite data til å teste modellen i sin helhet. Jeg håper allikevel at det kan finnes noen statistikker eller data som kan brukes for å gjøre en estimering på om vektene er gode

nok eller ikke.

5. Modelleringsprosessen

Jeg vil nå gå gjennom hvordan modelleringen ble utført i praksis. Jeg vil først gå gjennom de kilder som jeg har funnet og brukt som et grunnlag for modellen jeg har utviklet. Deretter vil jeg gå gjennom interaksjoner jeg har hatt med eksperter. Til slutt vil jeg se på utviklingsprosessen, gå gjennom hver iterasjon av arbeidet og se på hvordan modellen har modnet for hver av disse.

Til slutt vil jeg på den endelige modellen og se på valideringen av denne. Modellen vil være hovedresultatet av denne oppgaven, og jeg vil derfor gå gjennom denne i detalj.

5.1 Kilder fra Fiskeridirektoratet

Før jeg begynte modelleringen måtte jeg ha tilgang på informasjon og data om domenet. Det ble gjort klart på forhånd at det ikke fantes mye innsamlede data fra oppdrettsanleggene tilgjengelig. Det fantes derimot en god del dokumenter som omhandlet generell sikkerhet ved anlegg og hvilke uønskede hendelser som kan oppstå ved disse. Faktorer som spiller inn på risikoer har derfor i stor grad blitt identifisert i dokumenter som disse. Disse dokumentene er for det meste veiledninger gitt ut for oppdretterne selv, undersøkelser og lignende informasjon. Noen av disse kildene er ikke offentliggjort, og vil derfor ikke bli sitert i oppgaven.

5.1.1 AkvaBest

En av kildene jeg har brukt for å identifisere faktorer som spiller inn på risikoen for rømning er en erfaringsbase som kalles AkvaBest (Fiskeridirektoratet, 2010a). Dokumentene som finnes i denne erfaringsbasen er resultatet av et prosjekt utført av Fiskeridirektoratet for å identifisere fareområder tilknyttet rømning av oppdrettslaks. AkvaBest er hovedsaklig en veiledning for både næringen og myndighetene, og beskriver retningslinjer for hvordan en kan ivareta sikkerheten ved et anlegg.

I AkvaBest finnes det mye informasjon om hva som kan gå galt på de forskjellige delene av et anlegg og hvordan uønskede hendelser eller risikofylte tilstander kan unngås. I disse dokumentene finnes også bilder som beskriver de aktuelle hendelsene. AkvaBest går gjennom hele produksjonsklusen og inkluderer transport og montering av utstyr. Mange av de uønskede hendelsene som blir beskrevet i AkvaBest er forårsaket av menneskelige feil.

5.1.2 NYTEK

NYTEK er en standard for kvalitet og dimensjonering på utstyr som finnes på et anlegg (Fiskeridirektoratet, 2004). Alle anlegg i Norge skal etter forskriftene fra Fiskeridirektoratet være sertifisert etter denne standarden, eller ha et dugelighetsbevis som tilsier at de er i god nok stand til å brukes.

NYTEK setter også standarden for hvordan lokaliteten til et anlegg skal klassifiseres.

Klassifiseringen av et anlegg er et mål på de miljøkrefter som finnes ved dets lokalitet. Utstyret på et anlegg blir valgt med denne klassifiseringen som et grunnlag, for å tilse at det skal tåle de krefter som kommer til å virke på det i sjøen. Mange forhold tas i betraktning for å utføre lokalitetsklassifiseringen, men bølgehøyde og strømstyrke er de to viktigste faktorene.

Dette dokumentet inneholder mange eksempler på hvilke uønskede hendelser som kan opptre på et anlegg i sammenheng med utstyr. Det blir også forklart hvilke konsekvenser disse uønskede hendelsene kan ha, og hvordan mindre feil kan være med på å forårsake rømninger. Dette dokumentet beskriver også en del om hvordan miljøkrefter påvirker utstyret.

5.1.3 Tidligere systemer for risikovurdering

Jeg har tidligere i oppgaven nevnt dagens system for risikovurdering. Dette systemet inneholder en del faktorer som kan brukes som et utgangspunkt for den nye modellen. Dette dokumentet er et enkelt skjema for risikovurdering, og er utviklet av Kyst- og havbruksavdelingen til bruk i den årlige operasjonelle risikovurderingen de utfører. Faktorene som finnes i dette dokumentet er svært generelle forhold fra anlegg.

5.1.4 Rømningsrapporter

Jeg har tidligere i oppgaven brukt data fra rømningsrapporter for å identifisere hvilken fase i produksjonsyklusen det oppstår flest rømninger. Rømningsrapporter blir levert inn til Fiskeridirektoratet av oppdretterne dersom rømningsepisoder oppstår. Fiskeridirektoratet har satt sammen data fra slike rømninger for å få et bedre bilde av hvilke typer anlegg, komponenter og situasjoner det oppstår rømninger.

Matfiskfasen ble tidligere i oppgaven identifisert som den fasen som hadde størst antall rømning. Videre kan data fra rømningsrapporter være med på å si noe om hvilke komponenter eller situasjoner innenfor denne fasen som medfører størst sannsynlighet for rømning. Det tas allikevel forbehold om kvaliteten på denne dataen, siden rømningsrapporter sendt inn til Fiskeridirektoratet

er egenrapportert.

5.2 Interaksjon med ekspertene.

Kyst- og havbruksavdelingen ved Fiskeridirektoratet har stilt seg til disposisjon når det gjelder samtaler og intervjuer med eksperter på områder som har med rømning å gjøre. Jeg har hatt kontakt med flere av de ansatte ved avdelingen. Jeg har hatt mange samtaler med de som jobber direkte med tilsyn, og som ofte har arbeidsoppgaver som inkluderer inspeksjon av sjøanlegg. I tillegg til disse har jeg også hatt samtaler med ansatte ved Fiskeridirektoratet som jobber sentralt med mer strategiske problemstillinger, men som har viktige synspunkter på problemet.

Interaksjonen med ekspertene har vært svært uformell. Det vil si at jeg ikke har brukt noe som helst struktur for å intervjuer eller samle inn informasjon, utover å ha generelle planer over hvilke faktorer som skal diskuteres og spørsmål jeg har til ekspertene. Dette er gjort med vilje for ikke å legge bånd på den naturlige kunnskapsutvekslingen som bør være til stede i kunnskapsmodelleringen.

Formatet på møtene har vanligvis vært diskusjoner mellom meg selv og en til to eksperter, men det har vært tilfeller der jeg har samlet flere eksperter for å få flere synspunkter på noe dersom jeg har trengt dette. Jeg har også deltatt på en del møter mellom ekspertene der jeg selv ikke har vært aktiv, men der relevant informasjon har blitt diskutert dem imellom.

Jeg fikk også muligheten til å være med på en revisjon på et oppdrettsanlegg. Dette var en naturlig del av prosessen med å tilegne seg kunnskap om domenet. Å se oppbygningen av et anlegg og hvordan dette fungerte i praksis gjorde det enklere å se sammenhenger mellom faktorer som jeg tidligere bare hadde lest om.

I tillegg til ekspertene ved Fiskeridirektoratet har jeg også ved et par anledninger vært i kontakt med utenforstående eksperter. Dette gjelder ansatte ved Rådgivende Biologer i Bergen. Rådgivende Biologer er et firma fra Bergen som driver med arbeid innen oppdrettsnæring. De utfører de fleste typer undersøkelser som gjøres ved oppdrettsanlegg, inkludert konsekvensanalyser. Jeg tok kontakt med Rådgivende Biologer for å få nye synspunkter på domenet, og for å få en bredere kunnskap om problemet.

5.3 Gjennomføring

Modelleringsprosessen har blitt utført som en iterativ prosess. Iterasjonene har bestått av to trinn: Vurdering av forhold/utvikling av modellen og evaluering av modellen. Her følger en gjennomgang av denne prosessen, delt inn i de respektive iterasjonene.

5.3.1 Første iterasjon

I begynnelsen av den første iterasjonen dreide mye av arbeidet seg om å tilegne seg informasjon om domenet. Mye tid ble brukt på å sette seg inn i oppdrettslaksens produksjon. Dette foregikk gjennom samtaler med ekspertene og kildene jeg hadde tilgjengelig.

5.3.1.1 Analyse av rømningsrapporter

Det første som ble gjort var å analysere rømningsrapportene fra Fiskeridirektoratet for å finne ut hvilke faktorer tilknyttet sjøanlegg som hadde flest rømninger. Det ble her funnet at komponenten «not» på sjøanlegg hadde flest rømninger, og faktorer fra denne komponenten ble derfor videre undersøkt. Dette har blitt diskutert tidligere i oppgaven.

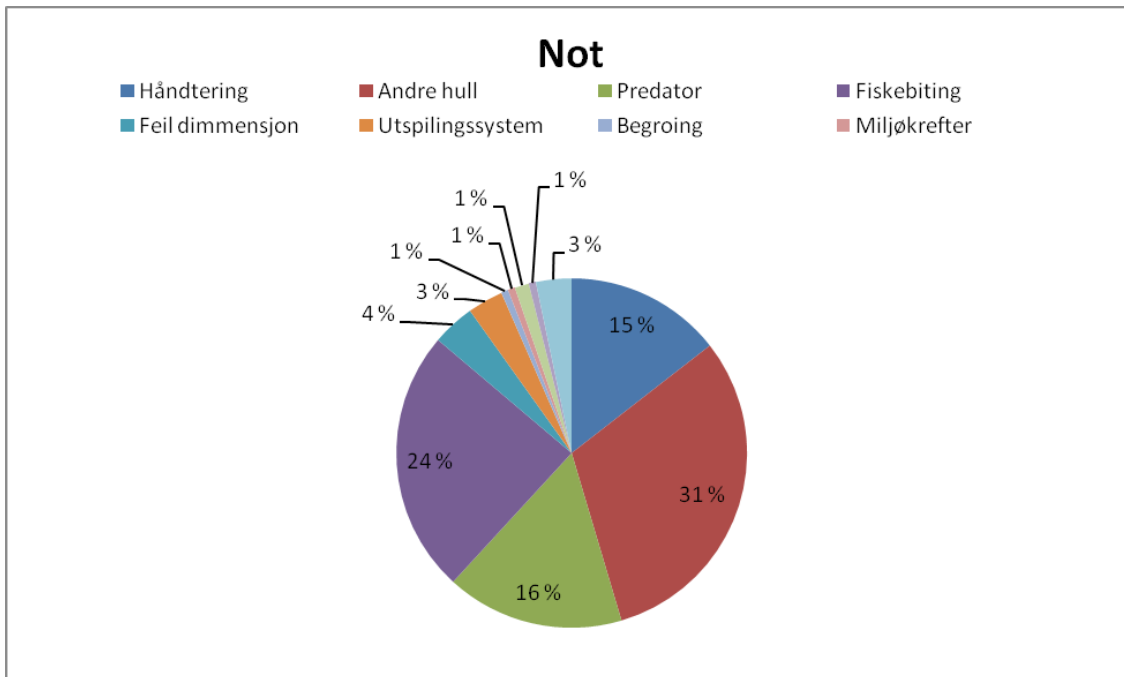
Videre ble rømningsrapportene analysert for å finne ut hvilke uønskede hendelser eller prosesser som er årsaken til rømninger. Siden det er mest fokus på komponenten «not» ble kun de uønskede hendelsene for denne komponenten analysert.

Figur 4 viser en analyse av de faktorer som påvirker rømning for komponenten «not». Her kan vi tydelig se at fire faktorer har høyest frekvens av rømning. Disse faktorene er «Andre Hull», «Fiskebiting», «Predatorer» og «Håndtering». Disse faktorenes høye frekvens kan tyde på at de er kritiske faktorer som har stor innvirkning på om det rømmer fisk fra anleggene.

For å sette tallene fra figur 4 i perspektiv kan vi se på tabell 1. Denne viser hvor mange prosent av rømning totalt (fordelt på alle faser i produksjonsyklusen) disse fire uønskede hendelsene utgjør i følge rømningsrapportene. Som vi kan se utgjør summen av disse fire ca 48.5% av totalen, som vil si at nesten halvparten av alle tilfeller der det rømmer laks stammer fra disse årsakene. Det er derfor trygt å si at noten vil være et viktig fokuspunkt for modellen, og de fire uønskede hendelsene som har blitt avdekket vil være viktige faktorer å undersøke for å forhindre rømning.

Andre Hull	17.4%
Fiskebiting	13.7%
Predatorer	9.26%
Håndtering	8.14%
Sum	~48.5%

Tabell 1: Faktorer for rømning ved nøter fordelt på total rømning



Figur 4: Faktorer som fører til rømning for nøter

Dataene fra rømningsrapportene var allikevel ganske utydelige. Kategorien «Andre Hull» representerer flere forskjellige faktorer som har ført til hull i not og dermed rømning, men gir ikke mye informasjon ut over det. Dette betyr at det kanskje kan finnes mange mindre eller uidentifiserte årsaker som samlet utgjør en stor brøkdel av faktorene som fører til rømning.

Det har også blitt reist spørsmål om hvorfor kategoriene ”Andre hull”, ”Fiskebiting” og ”Predator” har såpass høy frekvens. Det har blitt spekulert i om det er fordi disse kategoriene ikke tilsier at oppdretter har gjort en direkte feil. At predatorer gnager hull på nøtene er utenfor oppdretternes kontroll, og det er derfor godt mulig at noen velger å rapportere disse uønskede hendelsene som årsaken til rømning slik at de selv ikke kan bli straffet.

Selv om disse dataene kanskje kan sies å være av tvilsom kvalitet, stammer denne fra faktiske rømningsrapporter, og ved å identifisere disse faktorene fikk jeg et utgangspunkt for å konstruere begynnende modeller.

5.3.1.2 Faktorer hentet fra NYTEK og AkvaBest

Jeg hadde også et ønske om å forankre modellen jeg skulle lage ved å ta i bruk relevant data som allerede fantes om anleggene. Etter diskusjon med ekspertene fikk jeg vite at hvert anlegg har en lokalitetsklassifisering. Denne klassifiseringen er spesifisert i NYTEK (Fiskeridirektoratet, 2004), og er et mål på hvor store naturkreftene er på lokaliteten til anlegget. Denne klassifiseringen blir så brukt for å beregne styrken som trengs i utstyret for at det skal klare seg mot naturkreftene.

For å klassifisere miljøforholdene på et anlegg er det vanlig at strøm og bølgehøyde blir målt i en måned, og de målte dataene blir brukt for å gi anlegget en klassifisering. Det kan også bli brukt andre metoder for å måle dette, for eksempel lengre målinger av strøm og bølger eller ekstrapolering av eldre målinger. Slike undersøkelser blir ikke utført av Fiskeridirektoratet, men av utenforstående firmaer. Dette gjør at slike klassifiseringer i noen tilfeller ikke er helt pålitelige. Ekspertene har gitt uttrykk for at det er sannsynlig at det finnes anlegg som er underdimensjonerte på grunn av dette.

Som nevnt tidligere redegjør NYTEK for en del uønskede hendelser som kan oppstå med utstyr i sammenheng med både montering, aktivt bruk og normal drift. Det beskrives også hvilke konsekvenser disse uønskede hendelsene kan ha. Et eksempel på slike forhold som blir beskrevet er for eksempel den uønskede hendelsen «Feil fortøyning» som fører til den mer generelle kategorien «Mulig kontakt med andre deler ved skjevbelastning». Denne fører igjen til «Gnag og slitasje mot andre deler og utstyr» som til slutt kan føre til «Havari» eller «Hull i not».

AkvaBest (Fiskeridirektoratet, 2010a) inneholder mye informasjon om hva som påvirker risikoen for rømning. Mange faktorer er til stede i både AkvaBest og NYTEK, men i AkvaBest er det mer fokus på hva som fører til den uønskede hendelsen. De fleste av faktorene som påvirker risikoen for rømning ved et anlegg har ifølge AkvaBest et grunnlag i menneskelige feil.

Det kan sies at AkvaBest og NYTEK komplementerer hverandre. AkvaBest sier noe om hvordan uønskede hendelser oppstår, mens NYTEK redegjør for hvilke konsekvenser disse har og hvilke komponenter som blir påvirket. Ved å hente faktorer fra begge disse var det derfor mulig for meg å opprette begynnende modeller.

5.3.1.3 Bygging av modell i første iterasjon

De tidlige modellene jeg lagde var som sagt preget av de første tre kildene jeg først hadde tilgjengelig: rømningsrapporter, AkvaBest og NYTEK. Jeg bygget opp den første modellen ved å bruke de uønskede hendelsene som faktorer i modellen. Disse uønskede faktorene hadde selv faktorer som påvirket dem, men disse var som sagt som regel menneskelige feil. Jeg valgte å ikke ta hensyn til disse, siden jeg ikke var sikker på hvordan jeg skulle håndtere dem. Dette var akseptabelt siden målet med den første modellen var ikke å utvikle en perfekt modell av domenet, men å legge et grunnlag for videre utvikling.

AkvaBest og NYTEK inneholder en stor mengde uønskede hendelser som kan brukes som faktorer. De første utkastene av modellen jeg lagde ble derfor alt for store og uhåndterlige. Jeg fant tidlig ut

at jeg måtte begynne å slå sammen uønskede hendelser der dette var mulig. Eksempler på dette er uønskede hendelser som har mye til felles, som «Manglende driftsrutiner», «Feil prosedyre/opplæring» og «Gale prosedyrer ved skifting av not, levering, tørking/desinfisering.». Disse kunne samles under «Antatte dårlige rutiner på anlegg».

På tross av at jeg samlet mange av faktorene i kategorier, var den første modellen preget av veldig spesifikke hendelser i sammenheng med utstyr. Et eksempel på dette er «Ujevn heving av not», som sier noe om at det er en sannsynlighet for at noten blir hevet feil under en arbeidsoperasjon. Dette kan igjen føre til skader eller slitasje på noten, som kan være en påvirkende faktor for rømning.

Etter samtaler med ekspertene hadde jeg et inntrykk av at disse faktorene alene ikke var nok. I tillegg til de uønskede hendelsene ville jeg også modellere de miljømessige faktorene som for eksempel biomasse og helse i fiskebestanden. Jeg ville også ta hensyn til miljøkreftene som påvirket anlegget i forhold til den lokalitetsklassifiseringen det hadde.

5.3.1.4 Evaluering og revidering av modell i første iterasjon

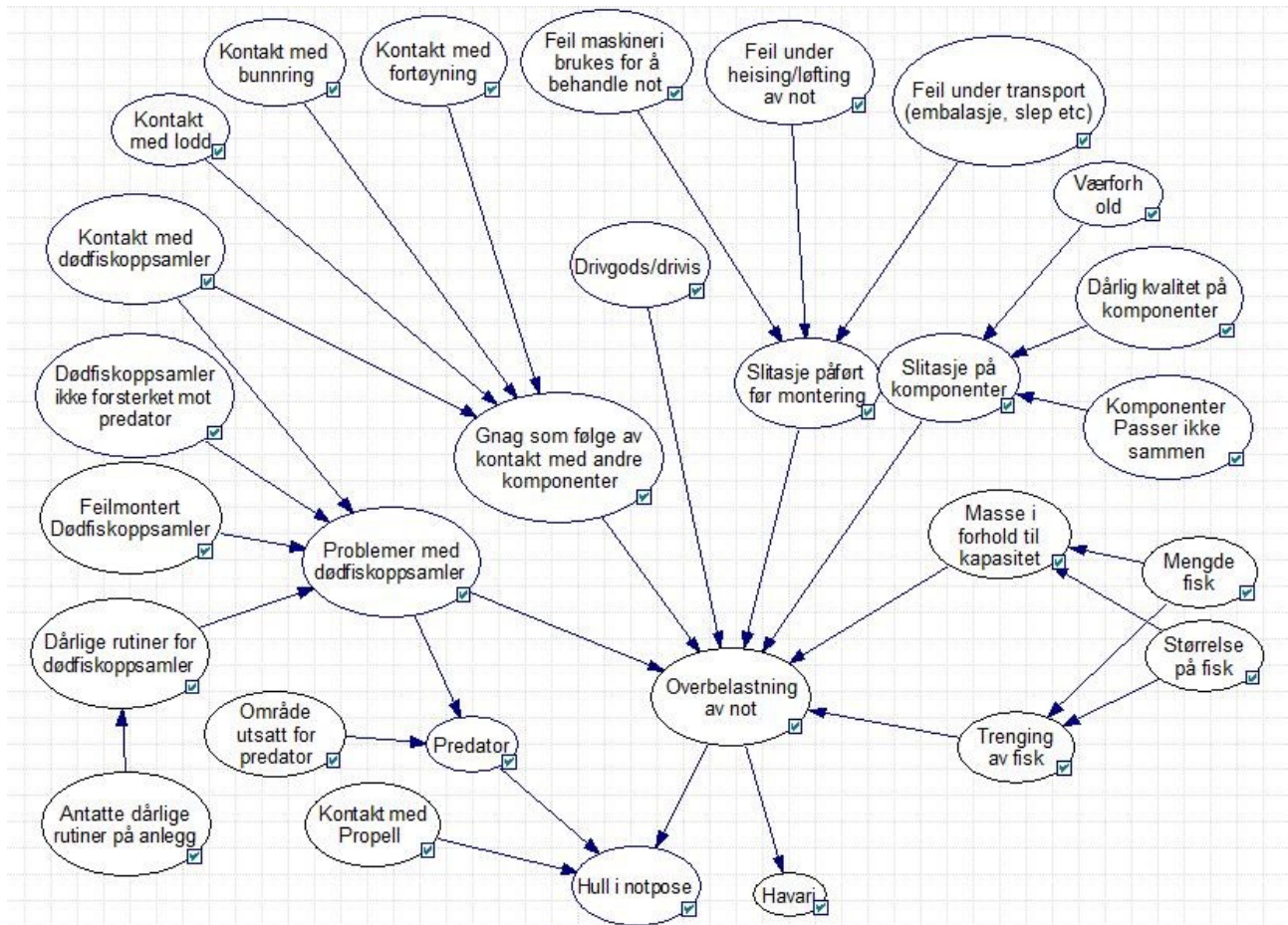
For å evaluere den første modellen brukte jeg først av alt resultatene fra analysen av rømningsrapportene. Denne viser hvilke faktorer som oftest er årsaken til at rømninger oppstår. Dersom en modell skal være god må den kunne representere disse årsakene. Jeg satte det derfor som et krav at den første modellen i hvert fall skulle kunne representere de årsakene som hadde størst frekvens.

Når ekspertene ble konsultert om relasjonene og variablene i modellen fikk jeg generelt sett positive tilbakemeldinger. Den første modellen var selvfølgelig ikke en representativ modell av domenet, men fungerte bra som et utgangspunkt og en måte å jobbe seg mot en mer fullstendig forståelse av domenet. I og med at modellen var et førsteutkast hadde den flere svake sider.

Modellen var ganske uoversiktlig, som vi ser i figur 5. Den hadde ikke en modulær oppbygning som jeg hadde planlagt før jeg begynte med arbeidet, og det var for enkelte faktorer vanskelig å se sammenhengen mellom faktoren og dens konsekvenser. Det var også et par uønskede hendelser som jeg hadde misforstått, og dermed var plassert feil. Dette var feil som oppsto på grunn av at jeg ikke ennå hadde nok kunnskap om domenet.

Modellen manglet faktorer der en kunne fastsette tilstander hvis en hadde håndfaste data. Et eksempel på dette er «Ujevn heving av not», som er en uønsket hendelse som kan opptre ved tidspress eller feil bruk av maskineri. Denne variabelen hadde potensielt stor påvirkning på om det

rømmer fisk eller ikke, men det ville være veldig vanskelig å finne data på om denne hendelsen oppsto eller ikke. Jeg ville utvide modellen min til å inkludere variabler som kunne måles, og var årsaker til at slike hendelser opptrer.



Figur 5: Modellen fra første iterasjon (større bilde finnes i vedlegg 10.1)

5.3.2 Andre iterasjon.

Andre iterasjon av kunnskapsmodelleringen var preget av å identifisere de faktorer som var årsaker til de uønskede hendelsene som ble identifisert i første iterasjon. I tillegg til dette var det ønskelig å innføre mer struktur i modellen, og bytte ut de spesifikke hendelsene med mer generelle kategorier av hendelser dersom dette var mulig. Jeg ville med disse kategoriene prøve å gjøre modellen mer modulær, som beskrevet i planen for oppgaven. I tillegg til disse hovedmålene for iterasjonen var det selvfølgelig et mål i seg selv å tilegne seg mer kunnskap om domenet slik at denne iterasjonens modell ble en mer nøyaktig formell representasjon av domenet.

5.3.2.1 Bygging av modellen i andre iterasjon.

Problemet med modellen fra første iterasjon var en mangel på variabler som kunne brukes som punkter for input. På grunn av dette bestemte jeg meg for å se hvilke felles faktorer som påvirket verdiene for de ulike variablene. Som tidligere nevnt er de fleste uønskede hendelsene på en eller annen måte et resultat av menneskelige feil. Disse menneskelige feilene gjaldt feil avgjørelser gjort av ledelse, feil utførte rutiner, feil bruk av utstyr og så videre.

For å håndtere disse menneskelige feilene ville jeg først se om de kunne deles inn i kategorier. Ved å dele de menneskelige feilene inn i typer, kunne jeg lettere holde oversikten på hvilke typer feil som økte sannsynligheten for hvilke uønskede hendelser. Kategorier av menneskelige feil som ble funnet var «Uerfarent Personell», «Manglende teknisk utstyr», «For lite personell» og «Problemer med drift/avgjørelser». Disse kategoriene ble lagt til som påvirkende faktorer for de uønskede hendelsene som allerede var lagt til i modellen. En av tankene bak disse var at de skulle være målbare, og at deres tilstander skulle brukes for å beregne sannsynlighetene for de andre faktorene i modellen.

For å innføre mer struktur i modellen og forsøke å gjøre den mer modulær bestemte jeg meg også for å klassifisere de skadene som kunne bli påført utstyret. Jeg bestemte meg for å utvide modellen til å modellere i hvilken fase av arbeidet skaden ble påført. Kategoriene jeg fant for disse fasene var «Slitasje påført før montering», «Slitasje påført under montering», «Slitasje påført under arbeid» og «Slitasje påført over tid». Alle faktorene i modellen kunne sies å påvirke minst en av disse kategoriene.

5.3.2.2 Evaluering og revidering av modell i andre iterasjon.

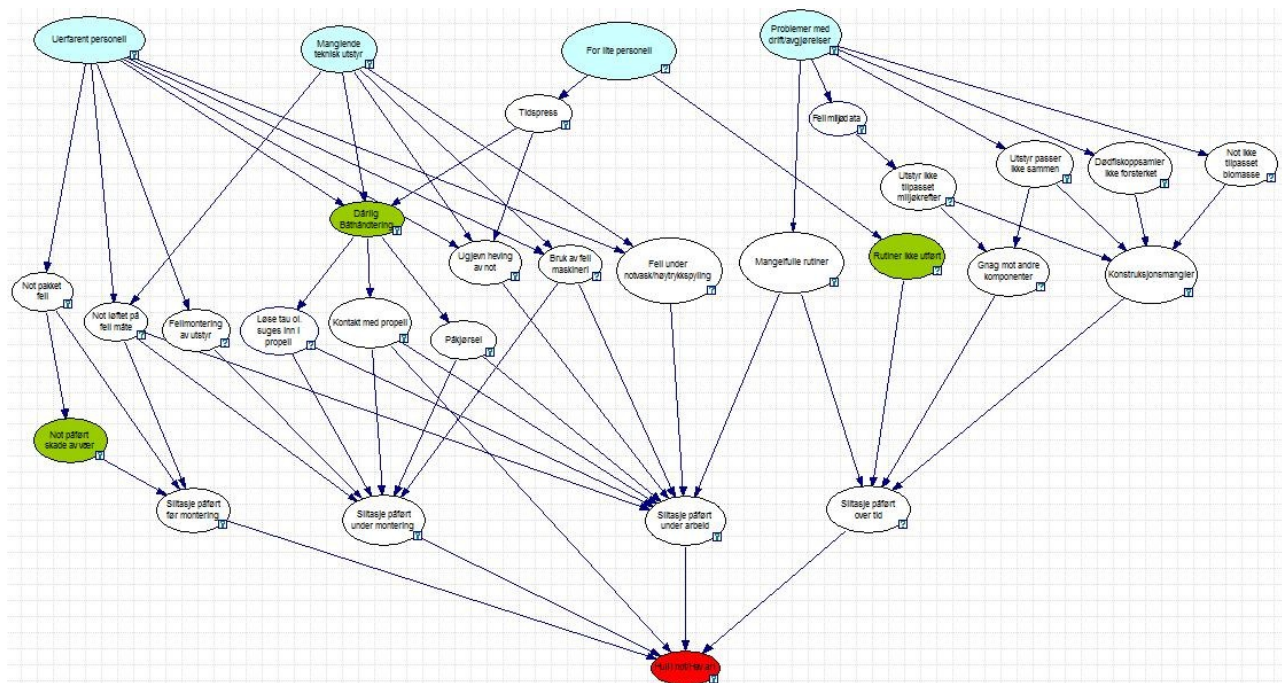
Figur 6 viser modellen fra andre iterasjon. Modellen fra denne iterasjonen ble godt mottatt av ekspertene. De mente at modellen representerte et godt bilde av de faktorene som påvirket domenet og hvilke relasjoner som fantes mellom disse. Faktorene i modellen var velkjente for dem siden disse var tatt ut fra dokumenter de allerede hadde god kjennskap til, noe som gjorde modellen lett for dem å akseptere.

Selv om denne modellen var et godt utgangspunkt for videre arbeid forsto jeg etter flere samtaler med ekspertene at denne modellen ikke kunne brukes slik den var. Dette var på grunn av at de faktorene jeg hadde brukt var vanskelige å måle og hendelsene som ble brukt fremdeles var for spesifikke. Dette gjaldt kategoriene «Uerfarent Personell», «Manglende teknisk utstyr», «For lite personell» og «Problemer med drift/avgjørelser». Disse kategoriene var helt klart viktige faktorer

og måtte bli med i modellen, men mange av dem var ifølge ekspertene vanskelige å finne målbare data for.

Ekspertene kunne også bekrefte at strukturen i modellen var representativ for domenet, men det ble stilt spørsmål til om en slik inndeling av modellen ville være fordelaktig. Argumenter i mot en inndeling av uønskede hendelser på denne måten var at det ikke hadde mye å si i hvilken fase av arbeidet feilen oppsto. Det var ønskelig å opprette en struktur i modellen som heller ga informasjon om mer distinkte kategorier for feil, som for eksempel «Feil båthåndtering» eller «Utstyr ikke tilpasset miljøforhold».

Dette ledet til ideen med å opprette hjelpevariabler som hver for seg oppsummerte en større og mer generell gruppe av uønskede hendelser som kan føre til rømming. Det ville være nyttigere for å dele modellen inn i slike hjelpevariabler for forhold til å dele feil inn i hvilken fase av arbeidet feilen oppsto, dette på grunn av at hjelpevariablene i seg selv sier noe om hvilke grupper av uønskede hendelser som påvirker den totale risikoen. En slik strukturering av modellen vil også gjøre modellen modulær, som var et av delmålene med modelleringen. Dette kom til å bli en stor forandring av modellen, og ble ikke gjort i andre iterasjon, men tatt med som input til tredje iterasjon.



Figur 6: Modellen fra andre iterasjon (større bilde finnes i vedlegg 10.2)

5.2.3 Tredje iterasjon

Etter å ha lagd de første modellene og diskutert disse med ekspertene hadde jeg bedre innsikt i hvilke faktorer som var viktige i domenet, og hvordan disse fungerte sammen. Mellom andre og tredje iterasjon var jeg også med på et anlegg og fikk se hvordan utstyret faktisk så ut, og hva som ble vektlagt på en inspeksjon. Dette var svært lærerikt og hjalp meg å sette meg inn i domenet. Jeg deltok også på en del møter og en større konferanse som omhandlet risikovurdering for oppdrettsanlegg. Observasjoner gjort ved disse anledningene gjorde til at jeg hadde nye ideer om hvordan modellen burde struktureres.

I denne iterasjonen bestemte jeg meg derfor for å designe en ny modell fra bunnen av. Tanken med den nye modellen var å utvikle den med tanke på struktur først av alt, slik at den ble mer funksjonell enn modellene fra de to første iterasjonene. Selv om jeg brukte mange av de samme relasjonene fra de første modellene ble strukturen i den tredje modellen bygget fra grunnen av. For å bygge en god struktur ble modellen delt inn i tre «lag»: et øverste lag med noder beregnet for input, et nederste lag beregnet for output og et midterste lag hvor beregninger av sannsynligheter og risikoer ble gjort.

5.2.3.1 Lagvis inndeling av modellen

Et tilbakevendende problem med modellene fra de to første iterasjonene var at det ikke fantes noen klare punkter for input til modellen. Dette var fordi faktorene enten var for generelle som for eksempel «Uerfarent personell», eller for spesifikke som «Ujevn heving av not». Jeg ønsket derfor å begynne utviklingen av den nye modellen med å identifisere direkte målbare forhold som finnes ved anlegg.

Jeg hadde altså et godt bilde på hvilke uønskede hendelser som opptrådte på anleggene og kunne føre til rømming, men det manglet informasjon om hva som forårsaket disse. Eksempelet som har blitt brukt tidligere: «Ujevn heving av not» kan brukes for å nøyere forklare dette. Ujevn heving av en not fører til økt risiko for at det oppstår skade på noten og at fisk kan komme seg ut. I og for seg er ikke hendelsen så interessant for modellen, men dersom det kunne finnes og måles data for hva som forårsaker «Ujevn heving av not», kunne dette ha blitt brukt for å lage en mye mer utfyllende modell.

For å oppnå målene mine med modellen i tredje iterasjon bestemte jeg meg for å organisere strukturen i den nye modellen i «lag». Jeg hadde det nederste laget som besto av de variablene jeg var interessert i å finne ut verdien for, nemlig «Hull i not» og «Havari». Jeg hadde også en god begynnelse på det midterste laget som besto av uønskede hendelser som kunne føre til rømming. Det

var denne typen informasjon det fantes mest av i de kildene jeg hadde fått tilgang til hos Fiskeridirektoratet. Det jeg manglet i de tidligere modellene var det øverste laget, som skulle være laget der input blir lagt inn. Denne inputen skulle bestå av direkte observerbare forhold fra anleggene.

5.2.3.2 Identifisering av målbare variabler

Jeg vurderte flere måter for å finne ut hvilke variabler som var målbare for domenet. Jeg hadde allerede et overblikk over aktuelle forhold, men bestemte meg tidlig for at jeg måtte konferere med ekspertene for å finne ut hvilke av disse som i det hele tatt var nødvendige og/eller mulige å inkludere. Jeg ønsket å få flere synspunkter på dette, gjerne fra flere eksperter. Jeg vurderte å intervju ekspertene en og en, men bestemte meg for at det var mer produktivt å samle flere for å diskutere sammen.

Jeg ville allikevel ha et utgangspunkt for denne diskusjonen, og bestemte meg for å bruke et dokument som ekspertene allerede var kjent med. På et av møtene jeg var med på mellom jeg lagde de første og den endelige modellen ble det utviklet et enkelt skjema for operasjonell risikovurdering. Dette er et enkelt skjema som kan brukes til for å beregne en risiko for et anlegg ved å legge inn enkel informasjon, og er en videreutvikling av Kyst- og havbruksavdelingens tidligere systemer for risikovurdering.

Dette dokumentet har åtte variabler som kan gis verdier. Disse er «Eksponering», «NYTEK-status», «Anleggstype», «Kompetanse», «IK-status», «Siste Kontroll», «Tips» og «Egen vurdering». Disse verdiene er veldig generelle, men gir et overblikk over de overordnede aspektene som påvirker risikoen ved et anlegg. Jeg valgte derfor ut «Eksponering», «NYTEK-status», «Kompetanse» og «IK-status» fra dette dokumentet, og la til «Utstyr» og «Eksterne parter» som variabler jeg selv mente var relevante.

Disse seks kategoriene mente jeg til sammen vil romme alle variabler som påvirker risikoen ved et anlegg. Etersom det finnes slike klare inndelinger av variablene vurderte jeg om det burde lages submodeller for hver av disse, som hver for seg tar for seg sin enkelte kategori av målbare variabler. Jeg valgte å ikke gjøre dette, da det finnes mange uønskede hendelser som blir påvirket av faktorer fra flere av kategoriene.

Jeg ville bruke disse generelle kategoriene for å få ekspertene til å tenke på hvilke målbare faktorer som kan passe under disse. Dette ville jeg gjøre i en samling av eksperter, for å få dem til å diskutere sammen og komme frem til en felles enighet om hvilke variabler de synes er aktuelle. Ved

å gjøre dette på en slik måte kan jeg også få en slags validering av de målbare variablene. Ingen målbare variabler ville bli tatt med dersom ekspertene ikke var enige om at de burde være det.

5.2.3.3 Bygging av modell i den tredje iterasjonen.

Identifiseringen av de målbare variablene ga gode resultater. Sammen med ekspertene avdekket jeg en rekke forhold som kunne tenkes å være målbare ved et anlegg. Når jeg sier at de kan tenkes å være målbare vil det si at det per i dag gjerne ikke finnes noen måte å måle dem på, men at de allikevel antas å være observerbare.

En viktig erfaring jeg gjorde meg når jeg lagde de første modellene mine var at det ikke nyttet å lage for spesifikke variabler. Dersom modellen inneholder for mange spesifikke uønskede hendelser vil den bli uhåndterlig og vanskelig å jobbe med, både for utvikler og de som senere skal bruke modellen. Etter at jeg hadde identifisert de målbare variablene ville jeg derfor opprette hjelpevariabler som var generaliseringer av de mer spesifikke uønskede hendelser som kan oppstå ved et anlegg.

Hjelpevariablene skulle utgjøre det midterste laget av modellen. Jeg ville bruke de faktorene og relasjonene jeg allerede hadde til å lage disse hjelpevariablene. Dette gjaldt spesielt de grupperingene av uønskede hendelser jeg allerede hadde funnet. Siden et tidligere problem hadde vært at disse var for spesifikke, ville jeg gjøre dem enda mer generelle. Jeg ville også at hver av hjelpevariablene skulle si noe om en spesifikk del av risikoen, som i seg selv er interessant. Dette på grunn av jeg ønsket å ha en modulær tilnærming til problemet.

Jeg har altså hentet mange av hendelsene fra de tidligere modellen jeg har laget, men trukket disse sammen i kategorier. Eksempler på slike sammentrekninger av uønskede hendelser er for eksempel at «Ujevn heving av not» og «Feil under notvask/høytrykksspyling» har sammen med andre lignende hendelser blitt trukket sammen til «Feil bruk av utstyr».

Dette har blitt gjort på en slik måte at det ikke er noe tap av informasjon. Kategoriene skal være disjunkte og det vil derfor være ekvivalent å bruke kategoriene som hjelpevariabler istedenfor de enkelte uønskede hendelsene. Alle de uønskede hendelsene jeg tidligere har brukt passer inn under en av de nye hjelpevariablene.

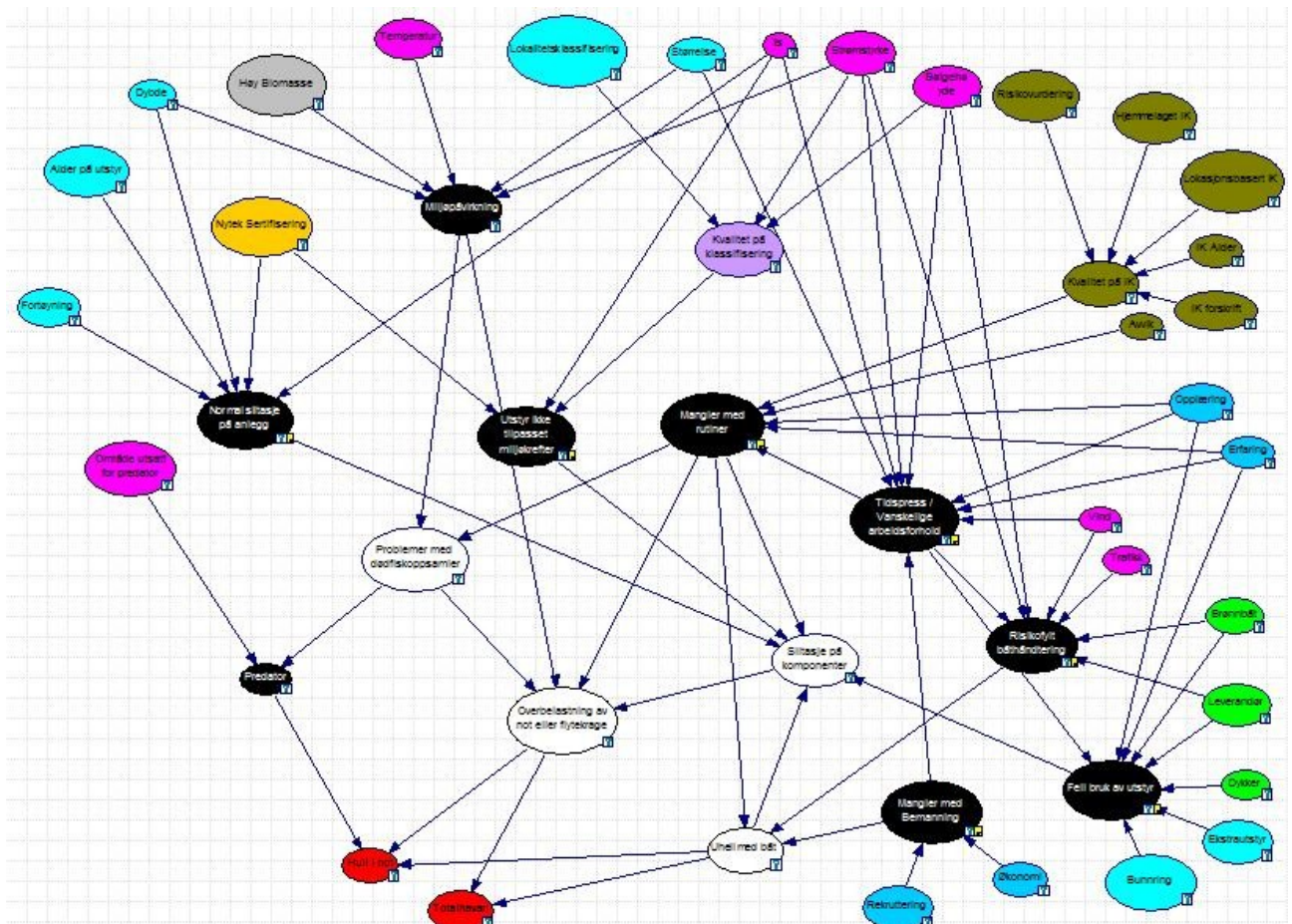
Ved å bruke disse hjelpevariablene tar jeg et bevisst valg om at modellen ikke skal ta hensyn til sannsynligheten for at en enkelt spesifikk uønsket hendelse oppstår. Dette gjør jeg fordi det ikke er dette som er hovedpoenget med modellen. Poenget med modellen er å beregne en overordnet risiko for rømning av fisk, ikke å modellere hvilken spesifikk hendelse som fører til dette. Å bruke

hjelpevariablene på denne måten vil tilføre modellen funksjonalitet ved at de kan si noe om interessante deler av domenet, som separat har mening i en risikovurdering. På denne måten vil også modellen bli modulær.

5.2.3.4 Evaluering og revidering av modell i den tredje iterasjonen.

Modellen fra den tredje iterasjonen var ifølge ekspertene en god representasjon av domenet. Denne modellen var mer komplett enn modellen fra de to første iterasjonene, som et resultat av måten den ble bygd opp. Ved å identifisere alle tenkelige målbare faktorer og bruke dem som et utgangspunkt ble denne modellen også mulig å bruke for beregning.

Valideringen av modellen ble som i de andre iterasjonene gjort ved å diskutere med ekspertene. Mange av disse diskusjonene forgikk mens modellen ble utviklet, og resulterte for det meste i små justeringer av modellen, men ingen større strukturelle forandringer. Disse justeringene dreide seg om små forandringer på et par relasjoner eller faktorer. Etter noen runder med validering og små forandringer oppnådde modellen et nivå hvor jeg valgte å bruke den som en endelig modell.



Figur 7: Modellen fra iterasjon 3 (større bilde finnes i vedlegg 10.3)

5.4 Den endelige modellen

Modellen jeg satt igjen med fra tredje iterasjon så ut til å kunne representere domenet på en god måte. Den oppfylte både målet om at den skulle kunne brukes for å vurdere risiko for et anlegg, og den identifiserer en mengde målbare faktorer som kan brukes for å gjøre dette. På grunn av at modellen er modulær kan den i tillegg brukes for å finne ut hvordan forskjellige deler av risikobildet ved anlegget påvirker den totale risikoen.

På grunn av at modellen tilsynelatende oppfyller disse målene ville jeg derfor gå videre til neste fase i kunnskapsmodelleringsprosessen. I denne fasen gikk den endelige modellen også gjennom en endelig validering. Denne valideringen var mer formell enn de som har blitt utført i iterasjonene. Jeg vil nå gi en detaljert beskrivelse av den endelige modellen og valideringen av denne.

5.4.1 Målbare faktorer

De målbare faktorene som ble identifisert ble brukt som et grunnlag for den endelige modellen. Modellen er bygd opp rundt disse og det er derfor naturlig å presentere disse. Det er seks forskjellige kategorier av målbare faktorer. Jeg vil nå gi en kort forklaring på disse kategoriene og de enkelte faktorene disse inneholder. Jeg vil gi en kort begrunnelse for hvorfor de er viktige og hva de innebærer.

Eksposering:

Faktorene som finnes under denne kategorien sier noe om hvordan de utenforstående forholdene er rundt et anlegg. Med utenforstående forhold menes det her forhold som vær, trafikk, rovdyrbestand og så videre. Kort sagt all informasjon om miljøet og andre forhold rundt anlegget.

Vind: Denne faktoren er et mål på hvor sterk vinden er på et anlegg. Vind gjør båtåndtering og arbeidsrutiner vanskeligere å utføre på en trygg måte.

Bølgehøyde: Et mål på hvor høye bølger som er målt ved et anlegg. Bølger gjør båtåndtering og arbeidsrutiner vanskeligere å utføre på en trygg måte

Strømstyrke: Mål på hvor sterk strøm det er ved et anlegg. Strømstyrken påvirker fiskens miljø i tillegg til at båtåndtering og arbeidsrutiner blir vanskeligere å utføre på en trygg måte.

Is: Denne faktoren beskriver hvor stor sannsynlig det er for at det finnes unormalt mye drivis og eller tilfrysing i vannet rundt anlegget. Dersom det finnes unormalt mye is ved et anlegg i en periode øker sannsynligheten for at anlegget blir utsatt for ekstra slitasje.

Trafikk: Et mål på hvor mye trafikk det er rundt et anlegg. Dersom et anlegg ligger nær en havn

eller i et ellers trafikkert område vil det være større sannsynlighet for påkjørsel fra båt.

Område utsatt for Predator: En faktor som tilsier om et anlegg holder til i et område som er preget av høy rovdyrbestand. Enkelte rovdyr kan bite/rive hull i noten for å komme seg inn til fisken.

Temperatur: Dette er et mål på om temperaturen i anlegget er gunstig eller ikke. En ugunstig temperatur vil påvirke fiskens helse. Dette er på grunn av at oksygenbehovet hos fisk forandrer seg med temperaturen i vannet.

NYTEK-status: Denne kategorien inneholder bare en faktor: «NYTEK-sertifisering». NYTEK er en standard for utstyr som brukes til oppdrett. Alt utstyr som brukes i dag må enten ha en NYTEK-sertifisering eller et dugelighetsbevis som tilsier at det er i god nok stand til å brukes i drift. Denne variabelen representerer om alt, noe eller ingenting av utstyret på et anlegg er sertifisert i henhold til NYTEK-standarden

Utstyr: Denne kategorien beskriver de utstyrsmessige forholdene ved et anlegg. Dette er forhold som klassifisering, utstyrstyper, ekstrautstyr, alder og så videre. Dette er viktige variabler på grunn av at mange av de uønskede hendelsene som oppstår er forårsaket av feil ved utstyr.

Alder: Denne variabelen er et mål på alderen på utstyret. Gammelt utstyr vil ikke være like sikkert som nytt utstyr og kan derfor utgjøre en signifikant risiko for havari og rømming.

Lokalitetsklassifisering: Denne faktoren sier noe hva slags klassifisering anlegget har i forhold til miljøkrefter. Denne klassifiseringen blir brukt for å beregne anleggets dimensjonering.

Kvalitet på klassifisering: Denne faktoren skal si noe om den antatte kvaliteten på lokalitetsklassifiseringen. Ettersom målingene som ligger til grunnlag for denne klassifiseringen kan bli sett på som mangelfulle og at beregningene ikke blir gjort av Fiskeridirektoratet selv kan det være fornuftig å si noe om hvor pålitelig denne er. Denne variabelen har blitt tatt med på grunn av at det kan være at anlegget blir utsatt for større miljølaster enn det er beregnet på. Denne faktoren er ikke direkte målbar, men blir beregnet ut fra lokalitetsklassifiseringen og de faktiske observerte miljøforholdene på anlegget.

Dybde: Dette er et mål på hvor dyp noten på et anlegg er. Desto dypere noten er desto vanskeligere blir den å håndtere og overvåke. Ved bruk av dype nøter vil det også være vanskelig å dykke, noe som hever risikoen ytterligere. Dype nøter har derimot en positiv påvirkning på fiskens miljø, på

grunn av at fisken da får mer rom og bedre tilgang på oksygenrikt vann.

Merder: Denne faktoren sier noe om hvor mange merder det er på et anlegg. Dersom et anlegg har mange merder er det større sannsynlighet for at disse blir lagt opp på siden av hverandre, noe som kan forverre fiskens miljø.

Fortøyning: Dette er et mål på kvaliteten på fortøyning ved et anlegg. Dersom det har blitt brukt feil type fortøyning eller denne er montert feil vil det være større sannsynlighet for slitasje og skade.

Ekstraustyr: Denne faktoren sier noe om hvor mye ekstraustyr det er på et anlegg. Desto mer ekstraustyr det finnes desto større er sannsynligheten for at det blir brukt feil, spesielt i kombinasjon med uerfarent eller for lite personell.

Bunnring: Bunnring blir brukt i bunnen av noten for å holde noten på rett plass. Disse blir av og til brukt feil, og dersom et anlegg har denne vil sannsynligheten øke for slitasje på komponenter.

For mye biomasse: Alle anlegg har en grense for hvor mye biomasse de kan holde i merdene. Biomasse er et mål på kilo levende fisk. Dersom denne grensen blir overgått økes sannsynligheten for at fiskens miljø blir dårligere.

Eksterne parter: Denne kategorien består av de risikoer som forårsakes av utenforstående tjenesteytere som benyttes ved et anlegg. Dette er for eksempel leverandører, dykkere, brønnbåter og andre tjenesteytere som har tilknytning til anleggene. Risikoen vil også øke for disse dersom det byttes tjenesteytere ofte. Dette er på grunn av at det er større sannsynlighet for at uhell opptrer dersom tjenesteyter ikke er kjent på anlegget. Risikoen vil også øke dersom tjenesteyter ikke har tilstrekkelig kunnskap jobben de utfører.

Brønnbåt: Et mål på hvor stor risiko brønnbåttrafikken på et anlegg utgjør. Dette målet bør kanskje være en sammensetning av hvor mye brønnbåttrafikk det finnes på anlegget, om anlegget bruker samme båt/firma/kaptein og hvor godt tilrettelagt anlegget er for brønnbåttrafikk.

Dykker: Denne variabelen sier noe om hvor sannsynlig det er at feil oppstår i sammenheng med dykking ved arbeidsoperasjoner, vedlikehold eller lignende. Dette kan være arbeidsoperasjoner som for eksempel vasking av not. Dersom det er uerfarne dykkere som utfører disse oppgavene er det større risiko for at feil blir gjort.

Leverandør: Denne variabelen sier noe om hvor stor sannsynlighet det er for at trafikk forårsaket av leveranser påfører skade på not eller anlegg generelt.

Kompetanse: Denne kategorien inneholder faktorer som gir et innblikk i hvordan kompetansenivået er på et anlegg. Med kompetanse menes det her den samlede erfaringen, kunnskapen på anlegget, og om det er tilstrekkelighet med arbeidskraft tilgjengelig.

Opplæring: Denne faktoren tilsier om opplæringen ved et anlegg er tilstrekkelig. Et anlegg som ikke gir de ansatte tilstrekkelig opplæring vil kunne få problemer med at utstyr ikke blir brukt på korrekt måte eller at rutiner ikke blir utført riktig.

Erfaring: Dette er et mål på den samlede erfaringen som finnes på et anlegg. Det samlede antall år arbeiderne har jobbet i næringen fordelt på antall arbeidere. Dersom det bare er nyutdannede arbeidere på et anlegg vil det være større sannsynlighet for at feil oppstår enn om arbeiderne er erfarne.

Økonomi: Dette er et mål på de økonomiske midlene et anlegg har til rådighet. Dersom det er mangel på økonomiske midler vil det kunne forekomme en nedprioritering av bemanning. Et anlegg med dårlig økonomi vil derfor ha større risiko for uønskede hendelser som følger av at de er underbemannet.

Rekruttering: Et mål på hvor gode rekrutteringsmuligheter anlegget har. Det kan for eksempel være tenkelig at det er lettere å finne nye ansatte i nærheten av byer og tett bebygde områder enn i avsidesliggende strøk. Dersom det er vanskelig å finne folk til å jobbe på et anlegg vil det være større sannsynlighet for at anlegget blir underbemannet. Dette fører til en høyere risiko for anlegget.

IK-status: Denne kategorien inneholder variabler som dreier seg om kvaliteten på rutiner og internkontroll ved et anlegg. De fleste av faktorene i denne kategorien brukes kun for å vurdere en generell kvalitet på internkontrollen ved anlegget.

IK forskrift: Denne faktoren tilsier om anleggets system for internkontroll følger forskriftene fra Fiskeridirektoratet.

Hjemmelaget: Denne faktoren tilsier om systemet for internkontroll er laget av oppdretter eller selskap. Dersom et system er hentet utenfra vil det være dårligere enn et som er spesifikt laget for deres eget anlegg.

Lokasjonbasert: Denne faktoren tilsier om systemet er laget spesifikt for det enkelte anlegget eller ikke. Selskaper har ofte like systemer for flere anlegg, noe som kan gjøre at kvaliteten på disse går ned på grunn av at anlegg ofte er forskjellige.

Alder: Denne faktoren sier noe om hvor lenge det er siden systemet for internkontroll blir revidert. Dersom det skjer forandringer på anlegget og systemet for internkontroll ikke oppdateres er det sannsynlig at kvaliteten på systemet går ned.

Risikovurdering: Denne faktoren sier noe om kvaliteten på den interne risikovurderingen ved anlegget. Dersom denne er dårlig er det mer sannsynlig at skadelige forhold ikke oppdages. Kvaliteten på internkontrollen går derfor ned dersom dette er tilfellet.

Kvalitet på IK: Denne faktoren er et mål på den generelle kvaliteten på internkontrollen ved et anlegg. Dersom kvaliteten på internkontrollen er dårlig vil det være større sannsynlighet for at det oppstår mangler ved rutinene på anlegget.

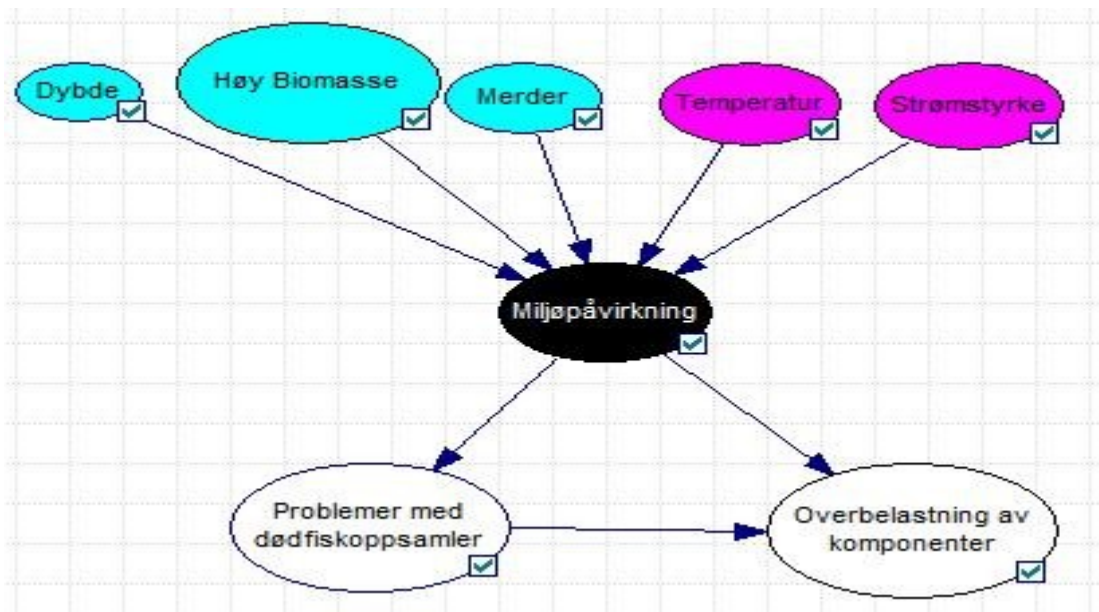
Avvik: Denne faktoren er et mål på hvor mange alvorlige avvik anlegget har. Hvis et anlegg har mange alvorlige avvik kan dette tyde på at det er mangler ved rutinene på anlegget.

5.4.2 Hjelpevariabler.

De direkte målbare faktorene er som sagt det første laget i modellen. Disse faktorene påvirker så det neste laget i modellen, som er et lag som består av hjelpevariabler. Disse hjelpevariablene er kategorier av hendelser eller faktorer som har nok felles egenskaper til at de kan plasseres under samme node i modellen. Dette midterste laget består kun av hendelser som ikke nødvendigvis garanterer utslipp/havari, men som påvirker risikoen for dette. Samlet vil de verdiene disse hjelpevariablene gir kunne brukes for å beregne den totale risikoverdien for et anlegg.

Jeg vil nå gå gjennom hjelpevariablene og gi detaljerte beskrivelser av dem.

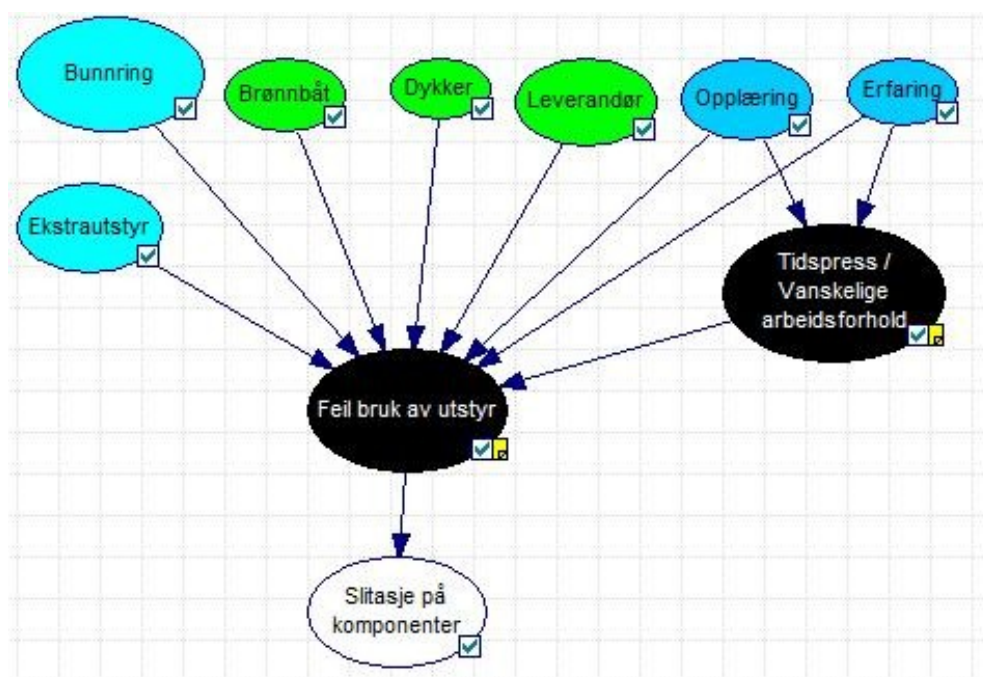
Miljøpåvirkning: Denne hjelpevariabelen representerer en kategori av faktorer eller hendelser som spiller inn på miljøet i vannet rundt og i merden. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 8. Med miljø menes det her faktorer som påvirker de biologiske forholdene og fiskens levedyktighet. Eksempler på disse forholdene er oksygen-nivå i vannet og gjennomstrømning av vann gjennom merden. Dersom miljøet i vannet er dårlig vil det være større sannsynlighet for diverse uønskede hendelser, for eksempel fiskedød. Disse uønskede hendelsene kan føre til belastning eller gnag på utstyret, som igjen kan føre til overbelastning, hull eller rømning.



Figur 8: Miljøpåvirkning med assosierte variabler og relasjoner

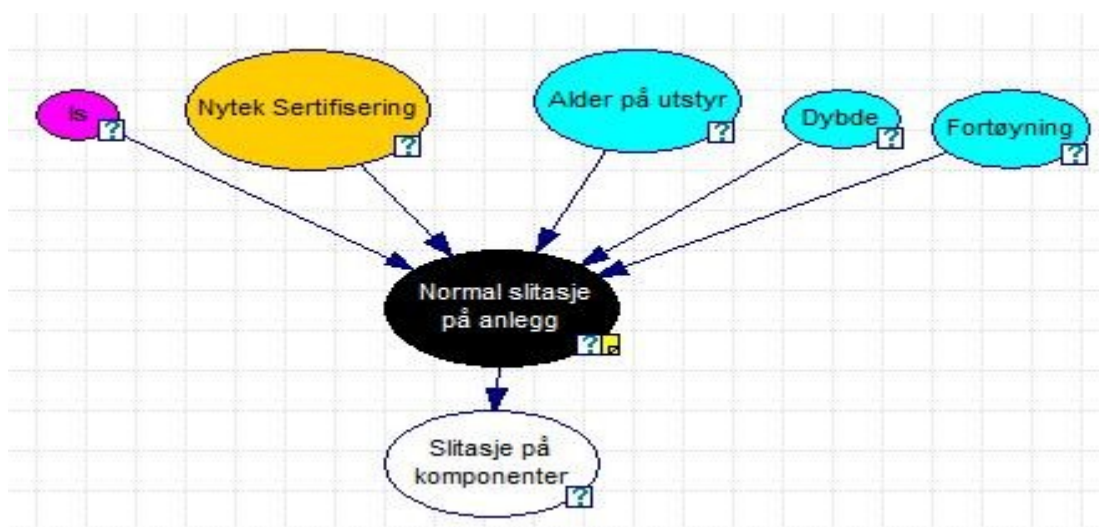
Feil bruk av utstyr: Denne variabelen representerer sannsynligheten for at utstyret som er til stede ved et anlegg blir brukt på feil måte. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 9. Dersom dette skjer økes sannsynligheten for overbelastning, hull og rømning.

Sannsynligheten for at utstyr blir brukt på feil måte øker dersom det finnes mye ekstrautstyr på anlegget, dersom det er problemer med opplæring eller erfaring på anlegget eller dersom det finnes risiko som et følge av bruk av utenforstående tjenesteytere som for eksempel dykkere, brønnbåter eller leverandører.



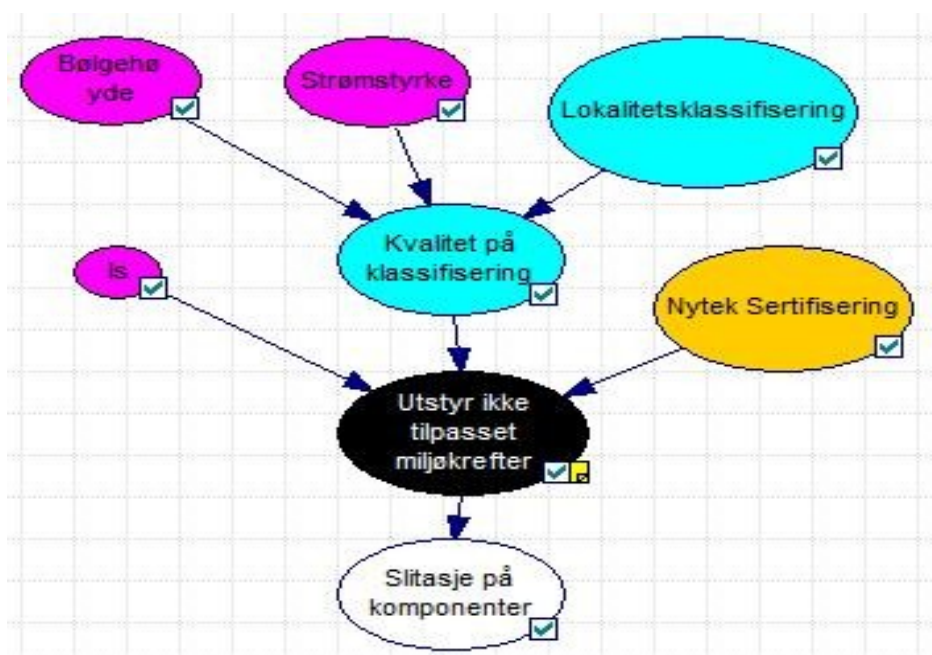
Figur 9: Feil bruk av utstyr med assosierte variabler og relasjoner

Normal slitasje på anlegg: Denne variabelen representerer den antatte normale slitasjen på et anlegg. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 10. Med normal slitasje menes det her slitasje som følge av normale driftsforhold, uten uønskede hendelser eller slitasje fra feil i driften. En viktig faktor for denne variabelen er alderen på anlegget. Et gammelt anlegg vil ha mye større slitasje enn et nytt. Andre faktorer som spiller inn er dimensjonene på anlegget (dybde, størrelse på merd osv.) og kvaliteten på fortøyningen. Dersom anlegget i perioder har mer is enn normalt vil dette også føre til at utstyret slites fortere. Ekspertene har også uttrykt at det er sannsynlig at utstyr uten NYTEK-sertifisering vil slites fortere. Dersom et anlegg er veldig slitt vil det ikke tåle påkjennelser fra andre faktorer, og øker derfor sannsynligheten for at rømninger skjer.



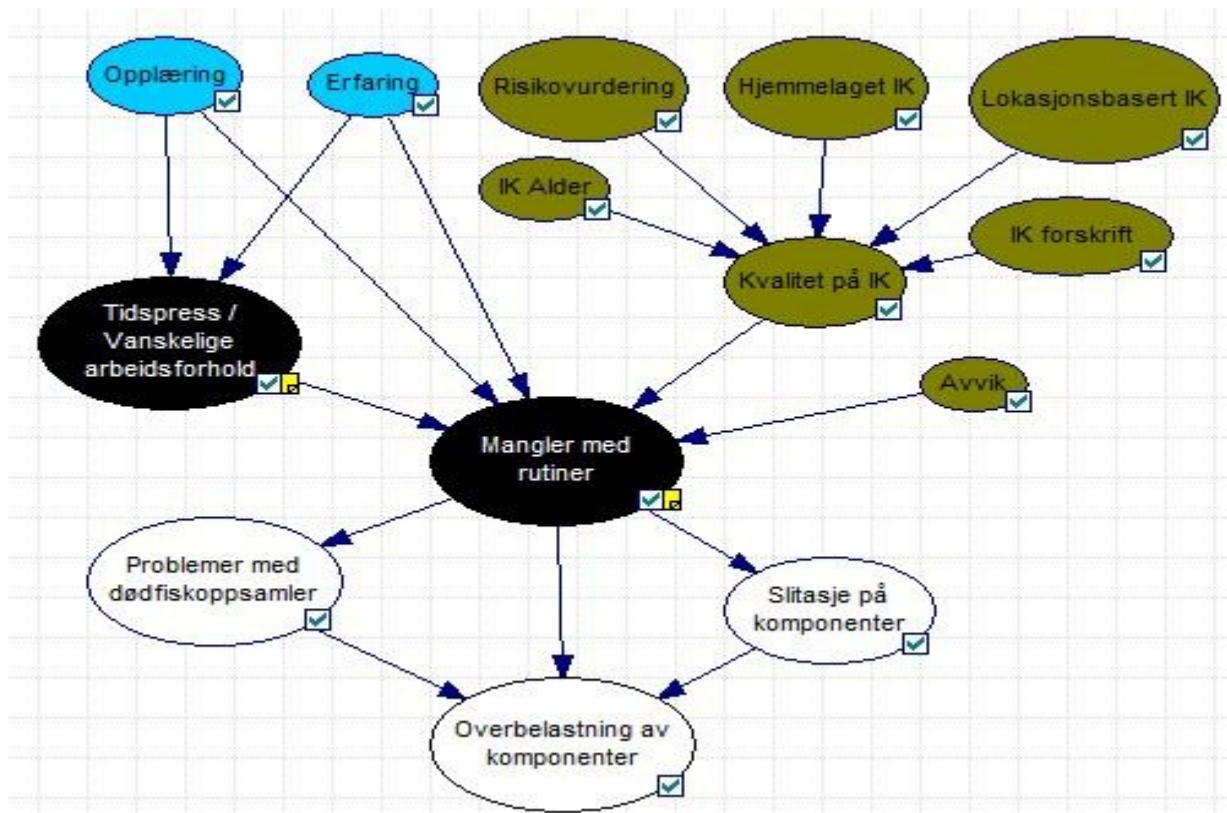
Figur 10: Normal slitasje på anlegg med assosierte variabler og relasjoner

Utstyr ikke tilpasset miljøkrefter: Denne variabelen representerer hvor godt anlegget er tilpasset de naturkrefter som opptrer i området rundt anlegget. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 11. Hvert anlegg har en klassifisering som tilsier hvor sterke naturkrefter anlegget tåler. Denne klassifiseringen er som sagt tidligere beregnet ut fra begrensede målinger gjort ved anlegget. Dersom det måles naturkrefter ved anlegget som overgår anleggets klassifisering vil dette utgjøre en stor risiko for rømning. Andre faktorer som spiller inn på denne variabelen er utstyrets sertifisering og unormalt mye is på anlegget.



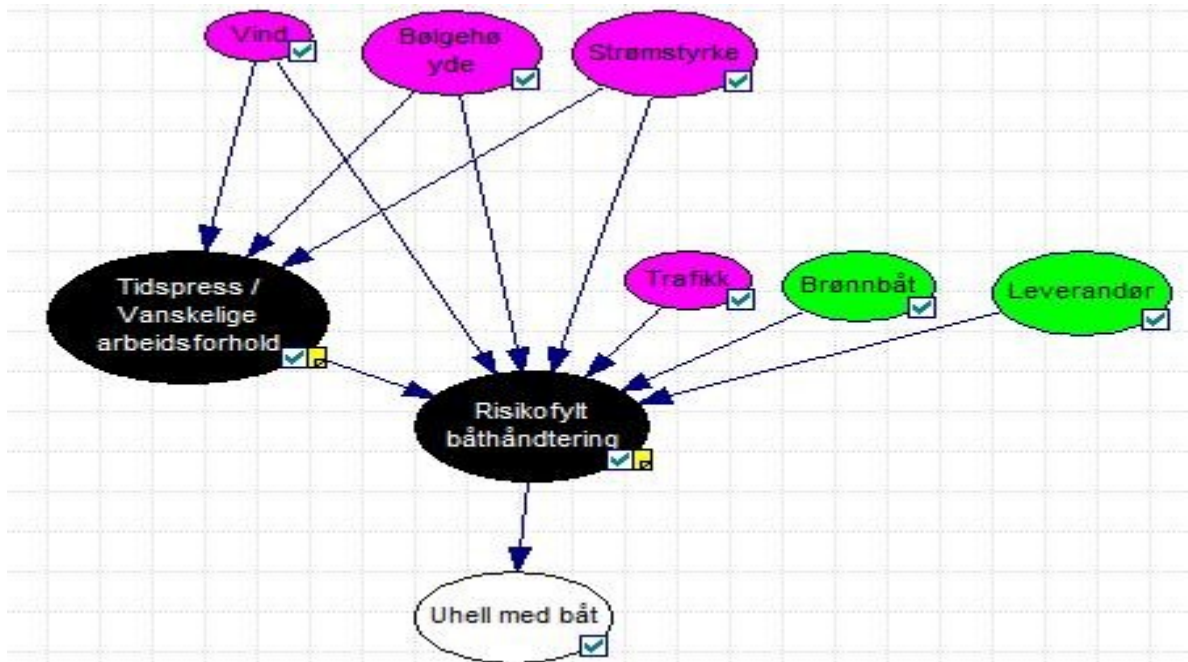
Figur 11: Utstyr ikke tilpasset miljøkrefter med assosierte variabler og relasjoner

Mangler ved rutiner: Denne variabelen sier noe om hvorvidt det er sannsynlig at det finnes mangler ved rutinene på et anlegg. Hjelpesvariabelens assosierte variabler og deres relasjoner vises i figur 12. Med mangler menes det her at de rutinene som finnes på anlegget enten er for dårlige eller at de blir utført på en ikke tilstrekkelig måte. Det er flere observerbare forhold på et anlegg som kan peke på at det er mangler ved rutiner, for eksempel at anlegget har dårlig internkontroll, at de har ansatte med lite erfaring i næringen eller at de har dårlig opplæring. Mangler ved rutiner kan føre til at usikre forhold ved anlegget ikke blir tatt hånd om eller oppdaget, og kan derfor utgjøre en stor risiko for anlegget.



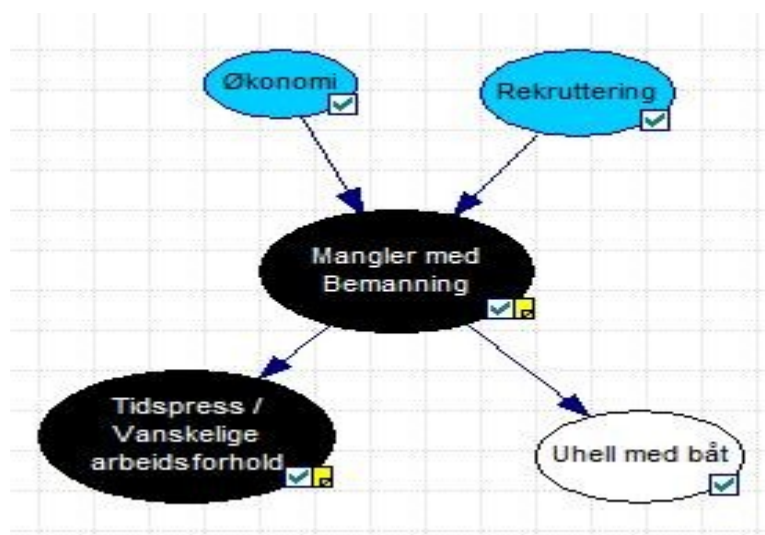
Figur 12: Mangler ved rutiner med assosierte variabler og relasjoner

Risikofylt båtåndtering: Denne variabelen representerer hvor stor sannsynlighet det er for at det opptrer risikofylt båtåndtering ved anlegget. Hjelpesvariabelens assosierte variabler og deres relasjoner vises i figur 13. Med risikofylt båtåndtering menes det her båttrafikk, enten i sammenheng med anlegget eller annen trafikk i nærheten, som har en sannsynlighet for å gjøre skade på anlegget. Båter kan skade både not og flytekrage ved et anlegg, enten ved kontakt med propell eller ved feil fortøyning. Dersom et anlegg ligger i et område med mye kommersiell eller privat trafikk vil det være større sannsynlighet for ulykker. Det er også større sannsynlighet for at uhell skjer dersom anlegget benytter seg av utenforstående tjenesteytere som ikke er kjent ved anlegget. Det er også større sannsynlighet for ulykker dersom det byttes tjenesteytere ofte. Bølgehøyde, vind og strømstyrke er faktorer som gjør båtåndtering vanskeligere, og spiller derfor også inn på denne variabelen.



Figur 13: Risikofylt båthåndtering med assosierte variabler og relasjoner

Mangler ved bemanning: Denne variabelen representerer sannsynligheten for at et anlegg har problemer med bemanning. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 14. Dette kan for eksempel opptre dersom økonomien til anlegget er dårlig eller hvis anlegget ligger i et område med lite tilgang til arbeidskraft på grunn av at rekruttering er vanskelig. Dersom et anlegg har problemer med bemanning kan det bety at anlegget kan bli preget av tidspress.

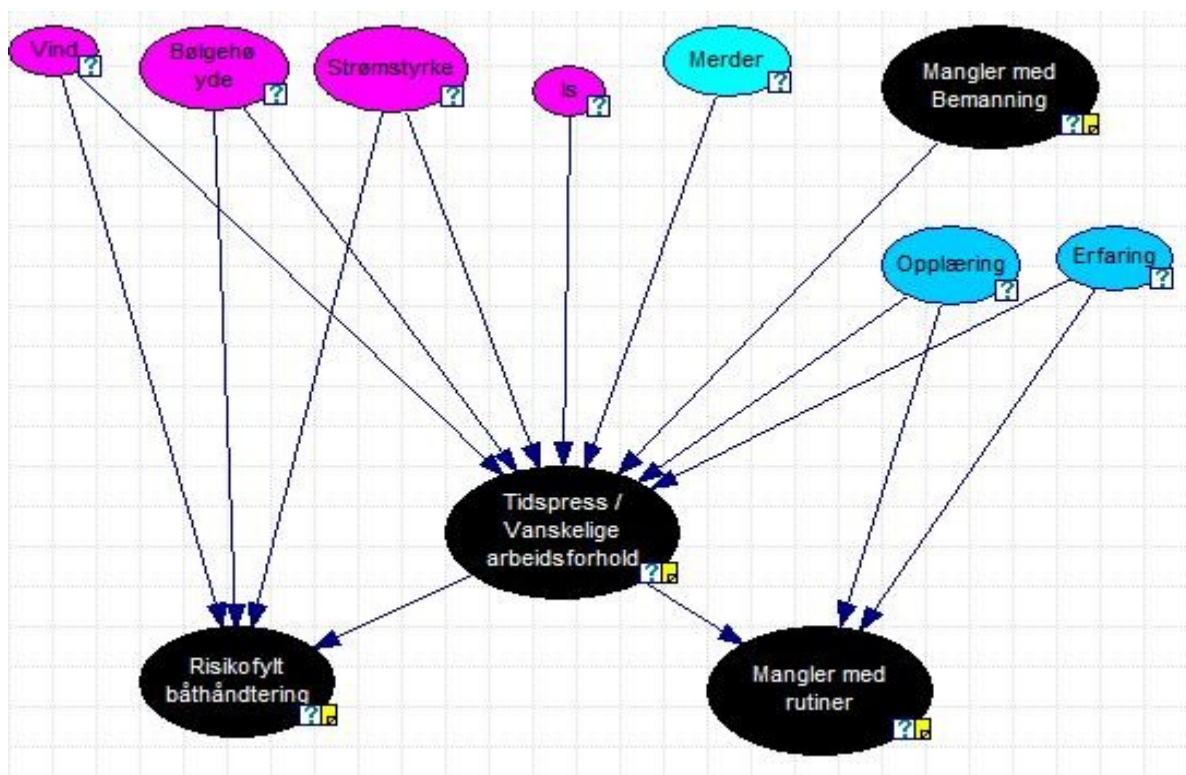


Figur 14: Mangler med bemanning med assosierte variabler og relasjoner

Tidspress/Vanskelige arbeidsforhold: Denne variabelen representerer sannsynligheten for at det

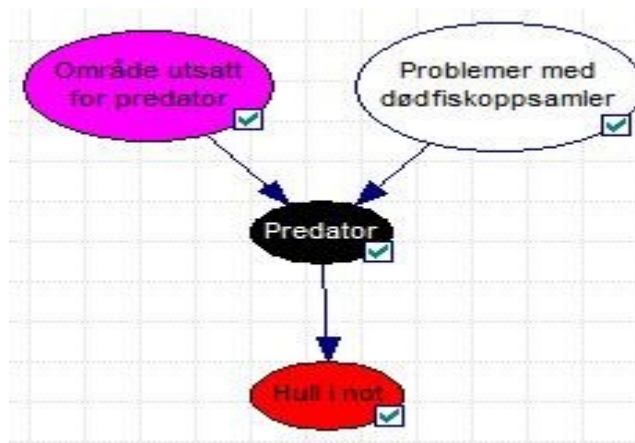
oppstår tidspress på et anlegg eller det på en eller annen måte er vanskelige arbeidsforhold på anlegget. Med dette menes det her at de ansatte ikke har muligheten til å fullføre alle de oppgavene de har på en tilstrekkelig måte på grunn av for liten tid, arbeidskraft eller værforhold.

Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 15. Sannsynligheten for tidspress går opp med antall merder på anlegget, dersom det er problemer med bemanning, opplæring eller tidspress, eller dersom naturkreftene på anlegget er så store at de kan vanskeliggjøre arbeidsoppgaver. Dersom det er tidspress på et anlegg kan det forårsake problemer med driften på anlegget. Det kan for eksempel føre til at det tas snarveier ved bruk av utstyr, rutiner eller båtåndtering. Dersom slike arbeidsoppgaver ikke blir utført på riktig måte kan det igjen føre til at det skjer uønskede hendelser som rømming eller havari.



Figur 15: Tidspress med assosierte variabler og relasjoner

Predator: Denne variabelen sier noe om hvor stor sannsynlighet det er for at anlegget har problemer med predatorer. Hjelpevariabelens assosierte variabler og deres relasjoner vises i figur 16. Med predatorer menes det her dyr som kan lage hull på noten for å komme seg inn til fisken. Sannsynligheten for predatorer øker dersom anlegget befinner seg i et utsatt område, eller dersom anlegget har problemer med dødfiskoppsamleren. Dersom det er problemer med dødfiskoppsamleren kan det samle seg død fisk nederst i noten, noe som tiltrekker seg rovdyr.

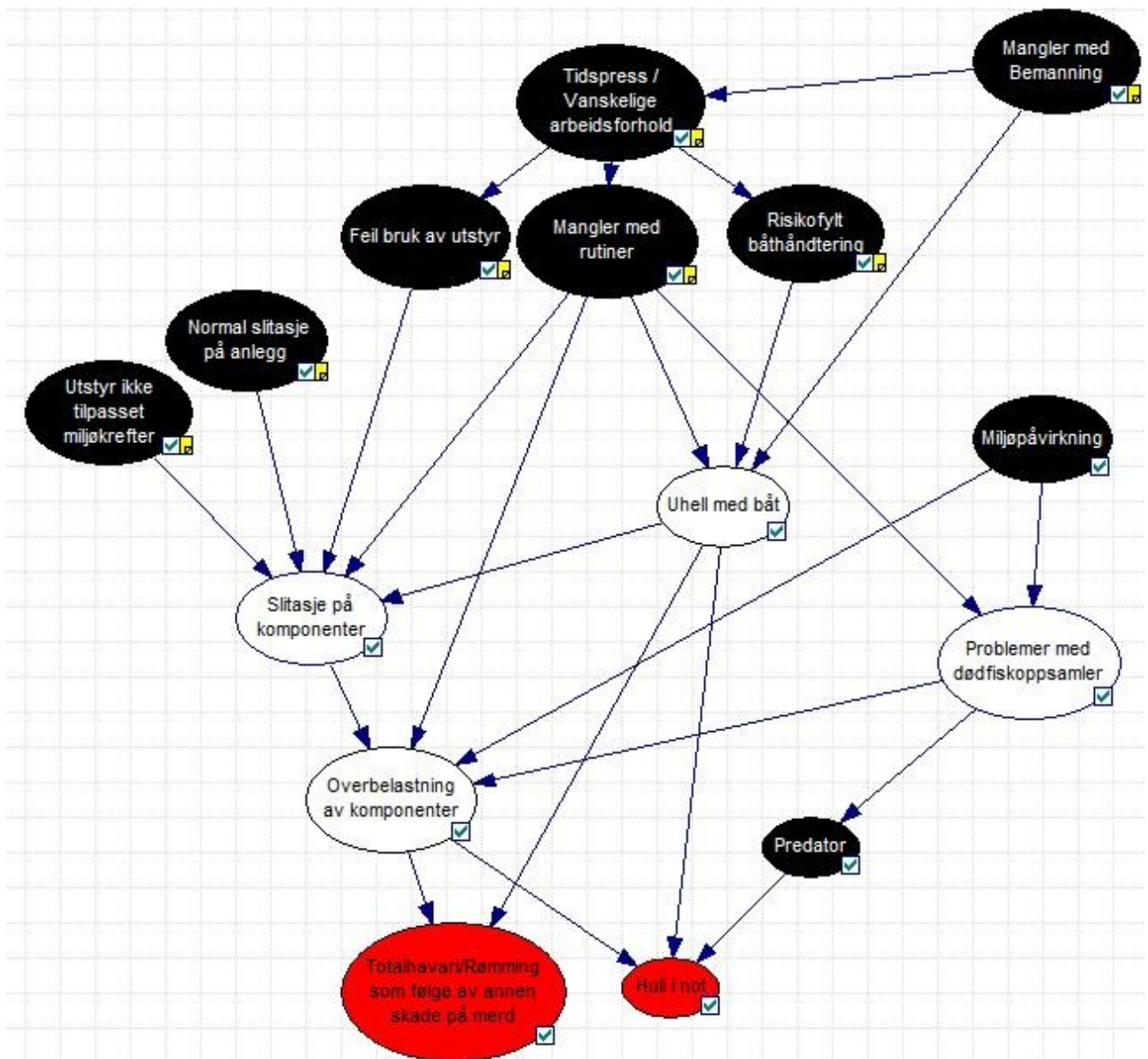


Figur 16: Predator med assosierte variabler og relasjoner

5.4.4 Sammensetning av den endelige modellen

Når jeg har koblet de målbare variablene opp mot de generelle hjelpevariablene har jeg videre sett på hvordan hjelpevariablene i seg selv leder fram til en eller flere av de uønskede hendelsene som fører til direkte rømningsepisoder. Dette gjelder de to hendelsene i modellen som representerer en faktisk rømning er «Totalhavari/Rømning som følge av annen skade på merd» og «Hull i Not». «Hull i Not» representerer en situasjon der det har kommet et hull eller en rift i noten, mens «Totalhavari/rømning som følge av annen skade på merd» representerer rømning som følge av totalhavari eller andre skader på merden som en helhet. Disse to hendelsene utgjør sammen den totale risikoverdien for et anlegg.

Mellom hjelpevariablene og de to resultatnodene har jeg lagt inn fire noder, som har som funksjon å knytte resten av modellen sammen. Relasjonene mellom disse nodene, hjelpevariablene og de nodene som representerer faktiske rømninger er vist i figur 17. De fire nodene som knytter hjelpevariablene sammen med rømningsnodene er i seg selv ikke så interessante, i og med at deres funksjon kun er å knytte sammen modellen til en enhet. Jeg vil nå kort gi en forklaring på nodene som knytter modellen sammen.



Figur 17: Sammensetning av hjelpevariabler

Problemer med dødfiskoppsamlere: Denne verdien representerer hvor stor sannsynlighet det er for at det kan oppstå problemer med dødfiskoppsamleren på anlegget. Dette kan for eksempel skje dersom rutiner ikke blir fulgt eller dersom mye fisk dør på grunn av at miljøet er dårlig.

Uhell med båt: Denne verdien representerer sannsynligheten for at det skjer faktiske uhell i sammenheng med båttrafikk ved anlegget. Denne verdien vil gå opp dersom det er stor sannsynlighet for risikofyllt båthåndtering. Dersom slike uhell skjer kan det påføre slitasje på utstyr i tillegg til at det kan i noen tilfeller føre direkte til rømning.

Slitasje på komponenter: Denne verdien representerer den totale sannsynlige slitasjen på anlegget. Alle forhold som har en sannsynlighet for å slite på komponentene ved et anlegg vil påvirke denne

verdien.

Overbelastning av komponenter: Denne verdien representerer sannsynligheten for at slitasje kan komme til å bli til en overbelastning. Dersom dette skjer vil det enten oppstå hull i noten, eller en annen skade på anlegget som fører til rømning.

Forholdet mellom «Slitasje på komponenter», «Overbelastning av komponenter», «Hull i not» og «Totalhavari» kan virke komplisert. Tanken bak strukturen her er at slitasje på komponenter aldri vil forårsake hull i noten eller annen skade før det faktisk skjer en overbelastning. Dersom et anlegg har høy sannsynlighet for slitasje på komponenter vil selvfølgelig risikoen for at hull kan oppstå gå opp, men dette skjer gjennom at risikoen for overbelastning går opp. Det er også flere faktorer som kan forårsake en overbelastning direkte, uten å gå via slitasje. Et eksempel på dette er miljøpåvirkning, som kan føre til massedød, som igjen har en liten sannsynlighet for å overbelaste noten.

5.5 Diskretisering

Når modellens struktur var ferdig måtte jeg finne de mulige tilstandene til de forskjellige variablene i modellen. I de fleste tilfellene hadde variablene et fåtall tydelige tilstander og var derfor rimelig enkle å etablere. Et eksempel på en slik variabel er «For høy biomasse». Denne har bare to tilstander, enten en akseptabel biomasse eller en for høy biomasse.

I noen tilfeller var det snakk om kontinuerlige variabler. Et eksempel på en slik variabel er «Vind», som måles i meter i sekundet. For denne typen variabler er ofte den letteste måten å håndtere dem på å utføre en diskretisering (Korb and Nicholson, 2010). For variabler av denne typen i modellen min var diskretiseringen en forholdsvis kort prosess. For å ikke gjøre modellen for komplisert ville jeg ha så få tilstander for hver faktor som mulig. Samtidig var det viktig å ikke miste informasjon ved å ha for få. Jeg brukte den innsikten jeg har fått i domenet for å ta en avgjørelse for hver av faktorene. Vanligvis var det nok med tre tilstander for de kontinuerlige variablene.

Diskretiseringen av hjelpevariablene med unntak av «Normal slitasje på komponenter» og «Predator» ble gjort ved å bruke lav, middels, og høy risiko. Dette er et valg jeg har tatt selv, på grunn av at disse variablene representerer hvor høy sannsynlighet det er for at noe er galt. Ved å bruke de diskretiseringene jeg har gjort kan jeg si noe om hvor sannsynlig det er at den enkelte faktoren er innenfor en eller annen risikogruppe.

5.6 Vekting

Vekting av et bayesiansk nettverk gjøres ved at de betingede sannsynlighetene (heretter bare kalt sannsynlighetene) for hver av de mulige tilstandene til en variabel gis en verdi. Disse sannsynlighetene kan finnes via tre kilder. Innsamlede data fra domenet kan brukes for å sette sannsynligheter, eksperter i domenet kan sette sannsynlighetene basert på kunnskap eller litteratur fra domenet kan brukes for å hente inn sannsynligheter. I dette tilfellet finnes det lite innsamlet data. For å sette sannsynligheter som kan brukes i nettverket må det derfor brukes enten litteratur fra domenet eller eksperter for å finne disse (Korb and Nicholson, 2010).

Siden min modell består av hjelpenoder som er generaliseringer av uønskede hendelser, er det vanskelig å finne eksakte sannsynligheter for disse. Både eksperter og de kildene jeg har tilgang til står i enighet om at det finnes kausale forhold mellom inputnodene og hjelpenodene, men på grunn av at disse er generelle hadde de problemer med å vurdere konkret hvor stor denne påvirkningen er.

På grunn av dette har uthenting av vektene foregått på denne måten: Jeg har satt en standard påvirkning som alle inputnodene har på hjelpenodene. Denne standard påvirkninger har jeg så økt eller minsket dersom jeg har inntrykk av at faktoren er mer viktig eller mindre viktig. Inntrykket jeg har av faktorenes viktighet stammer fra samtaler med ekspertene og litteraturen jeg har gått gjennom.

På denne måten har jeg kunne etablere sannsynligheter for alle variablene. Disse sannsynlighetene vil ikke være nøyaktige. Dette er akseptabelt på grunn av at målet med å vekte modellen var å ha en funksjonell modell der sannsynlighetene hadde et grunnlag i virkeligheten, ikke å utvikle en modell som uten modifiseringer kunne brukes til å vurdere risiko. For å utvikle en modell som kan brukes i industrien vil det sannsynligvis kreves større mengder data enn det som finnes tilgjengelig i dag.

5.7 Validering av den endelige modellen

Det siste steget i kunnskapsmodelleringsprosessen er å utføre en endelig validering av modellen. For å gjøre dette hadde jeg planlagt å dele modellen inn i tre aspekter som skal valideres separat. Disse er modellens struktur, diskretiseringen av variablene og vektingen av modellen. De to første aspektene ble gjort etter planen, men vektingen viste seg å være vanskelig å validere.

For å utføre disse valideringene brukte jeg ekspertene jeg tidligere har jobbet med, og fikk i tillegg hjelp av eksperter utenfra Fiskeridirektoratet som ikke før hadde sett modellen. Å få nye synspunkter fra eksperter som ser på næringen fra en annen vinkel styrket resultatet av den endelige

valideringen.

5.7.1 Validering av struktur og diskretisering av variabler

Formålet med valideringen av modellens struktur var å sikre at ingen av relasjonene i modellen ikke hadde et grunnlag i virkeligheten. Jeg gikk derfor gjennom alle faktorene og deres relasjoner til andre faktorer systematisk med ekspertene (både ekspertene fra Fiskeridirektoratet og de utenforstående ekspertene). Jeg gjorde dette med en til to eksperter om gangen, og hadde totalt tre gjennomganger. Jeg ba dem under disse gjennomgangene å bekrefte hver av relasjonene. Etter dette sammenlignet jeg resultatene for å se om bekræftelsene fra de forskjellige ekspertene stemte overens.

Det viste seg at alle ekspertene var enige om som godt som alle relasjonene. Unntaket var at en av ekspertene stilte spørsmål til et fåtall av relasjonene. På grunn av at det kun var én av ekspertene som stilte spørsmål til relasjonene valgte jeg å se bort fra at disse var ubegrunnede. De andre ekspertene bekreftet disse relasjonene.

Valideringen av diskretiseringene ble gjort på samme måte, men ble kun gjennomgått av to eksperter. Dette var på grunn av at de fleste tilstandene for variablene var relativt selvforklarende. Jeg mener derfor at en rigid gjennomgang av disse ikke var like nødvendig som gjennomgangen av strukturen. Ekspertene som gikk gjennom diskretiseringene var enige i de valgene jeg hadde tatt i de tilfellene der kontinuerlige variabler skulle diskretiseres.

5.7.2 Validering av vekting

På grunn av at vektningen ble gjort kun ved å estimere forskjeller i påvirkningskraft fra faktorene, vil denne ikke være nøyaktig. Validering av denne vil derfor ikke være nødvendig. Dersom det hadde vært data tilgjengelig kunne en ha brukt maskinlæring for å forbedre vektene, og fått et mer nøyaktig sett sannsynligheter. En kunne også ha brukt data for å validere modellen.

Vektene i modellen min er ikke tiltenkt å være representative for domenet, men er estimeringer som er satt på grunn av at modellen skal kunne brukes for å demonstrere bruk. Å validere denne vil derfor være meningsløst.

6. Bruk av modellen

Jeg vil nå gå gjennom et par kjøreeksampler for modellen og se på hvordan den kan brukes for å vurdere risiko for oppdrettsanlegg og sammenligne risikoverdier fra flere anlegg. Dette vil gi et

grundigere innblikk i modellen og hvordan bruk av denne er tiltenkt. Vektene i modellen er estimeringer som jeg selv har gjort på grunnlag av informasjon fra ekspertene og de kildene jeg har tilgjengelige. Resultatene fra disse kjøreeksemplene vil derfor ikke være representative for faktiske anlegg, men ment for å illustrere bruken av modellen min.

6.1 Minimum- og maksimumsverdier.

Modellens formål er å sette en risikoverdi på et anlegg ut fra observerbare forhold som finnes på denne. Det finnes en minimum og en maksimumsverdi for disse risikoverdiene for min modell. Minimumsverdien er den minste risikoen et anlegg kan ha i modellen. Denne verdien oppnås dersom alle forholdene på et anlegg ligger til rette for at det ikke skal oppstå uønskede hendelser. Maksimumsverdien oppnås dersom alle forholdene ligger til rette for at uønskede hendelser kommer til å oppstå ved anlegget.

	Minimum	Maksimum	Differanse
Hull i Not	15.72%	29.22%	13.5%
Totalhavari/annen rømning	12.08%	18.61%	6.53%

Tabell 2: Minimums og maksimumsverdier for modellen

Tabell 2 viser hva minimum og maksimumsverdiene for nodene «Hull i Not» og «Totalhavari» er. Her ser vi også hva differansen mellom disse verdiene er. Disse verdiene representerer de verdier av risikoer som anlegg kan bli tilegnet.

Et sentralt poeng her er at disse tallene ikke representerer en sannsynlighet for at feil skal oppstå ved et anlegg, men er en risikoverdi som kun har som formål å si noe om hvor stor risiko det finnes ved anlegg relativt til minimumet og maksimumet.

Dette gjelder også til en viss grad hjelpevariablene. Disse har tre mulige tilstander, lav, medium og høy risiko, som gjenspeiler hvor sannsynlig det er at anlegget ligger i disse. Med de sannsynlighetene jeg har lagt til i modellen vil det alltid være en risiko for at et anlegg ligger i enten medium eller høy risikogruppe. Det vil derfor med mine sannsynligheter alltid være rundt 5% sannsynlighet for at anlegget ligger i gruppen for medium risiko, og rundt 5% sannsynlighet for at det ligger i gruppen for høy risiko.

6.2 Bruk av modellen i et større system.

Modellen min er implementert ved bruk av det grafiske modelleringsverktøyet GeNIe. Dette

verktøyet gir brukeren muligheten til å sette beviser for noder i modellen. Med beviser menes her at vi har informasjon om at en node har en bestemt tilstand eller verdi. Dette betyr at en kan legge inn den informasjonen vi vet om inputnodene i modellen, for å deretter vurdere hvilke tilstander som er mest sannsynlige for de noden vi ikke har informasjon om.

Ved å legge inn den informasjonen vi har tilgjengelig om et anlegg inn i modellen, vil vi kunne bruke dette verktøyet for å automatisk beregne hvor sannsynlige de forskjellige tilstandene i resten av modellen er. Det er disse sannsynlighetene som til slutt aggregeres i de endelige risikoverdiene som sier noe om den totale risikoen ved anlegget.

I et fullt utviklet system ville denne prosessen fungert på samme måte, men det ville sannsynligvis bli brukt et mer brukervennlig verktøy. GeNIe er et verktøy for utvikling av bayesianske nett, og å bruke dette verktøyet for å sette inn informasjon for hvert av anleggene ville vært upraktisk. Det er mer sannsynlig at mange av disse dataene ville bli direkte lest inn i systemet fra en database dersom systemet ble tatt i bruk.

6.3 Bruk av modellen for sammenligning av anlegg.

Jeg vil nå gå gjennom et kjøreeksempel der modellen brukes for å vurdere tre forskjellige oppdrettsanlegg. I og med at data ikke finnes tilgjengelig for faktiske oppdrettsanlegg vil anleggene i disse eksemplene være oppdiktete. Disse blir hver for seg tildelt risikoverdier, som til slutt blir sammenlignet.

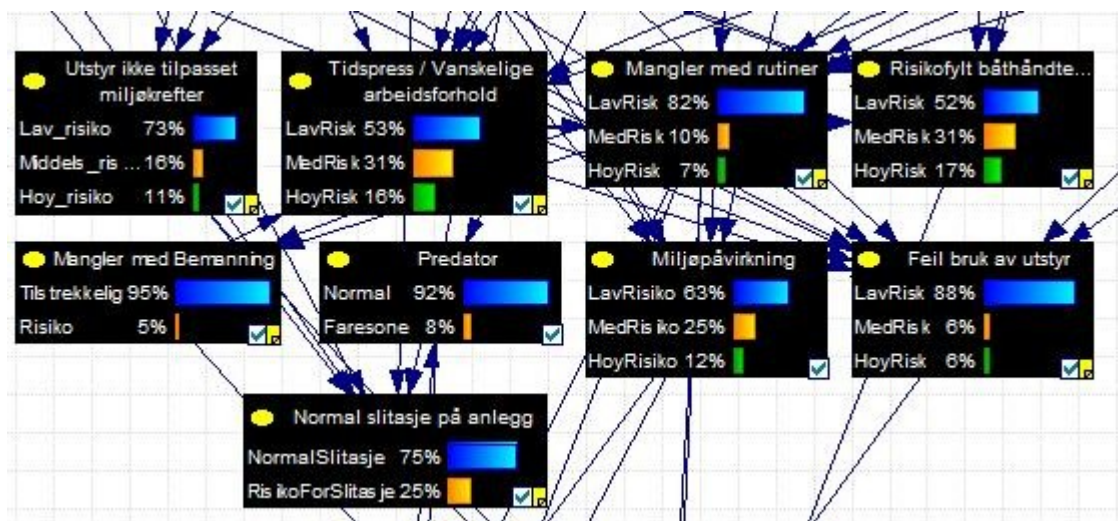
For å gjøre dette vil jeg bruke GeNIe for å sette inn beviser for alle inputnodene i modellen, for hver av anleggene. Jeg vil deretter bruke GeNIe for å beregne de resterende hjelpevariablene, og de totale risikoverdiene. Ved å gjøre dette får hvert av anleggene en samlet risikoverdi, samtidig som at en kan se på hjelpevariablenes verdier for å finne ut detaljer om risikoverdien.

6.3.1 Anlegg A

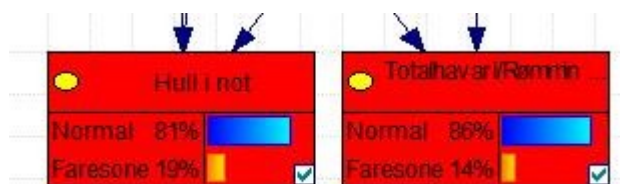
Anlegg A er et tiltenkt anlegg som er i topp stand og som kun har problemer med sterke naturkrefter og ugunstig temperatur. De observerbare tilstandene som finnes ved dette anlegget blir interpretert og modellens inputnoder settes til de respektive verdiene. Verdiene for hele anlegget vises her i tabell 3.

Faktor:	Tilstand
Alder på utstyr:	Nytt
Dybde	Grunn
Høy biomasse	Akseptabel biomasse
Lokalitetsklassifisering	Klasse 5
Fortøyning	Godkjent
Ekstraustyr	Ikke risiko
Merder	Få merder
Bunnring	Ikke bunnring
Is	Unormalt
Strømstyrke	Sterk
Vind	Sterk
Temperatur	Ugunstig
Bølgehøyde	Høy
Trafikk	Trafikkert
Område utsatt for predator	Ikke utsatt
NYTEK sertifisering	Sertifisert
Risikovurdering	God
Hjemmelaget IK	Hjemmelaget
Lokasjonbasert IK	Lokasjonbasert
Alder IK	Brukbar
IK forskrift	Forskrift
Avvik	Akseptabelt
Opplæring	God
Erfaring	God
Rekruttering	Lett
Økonomi	God
Brønnbåt	Ikke risiko
Dykker	Ikke risiko
Leverandør	Ikke risiko

Tabell 3: Distribusjon av tilstander i modellen for anlegg A



Figur 18: Hjelpesvariablenes verdier for anlegg A



Figur 19: Total risiko for anlegg A

Fra figur 18, som viser hjelpevariablenes verdier for anlegg A, kan vi se at det er fire variabler som har fått tydelig høyere sannsynlighet for å være i enten en medium- eller høyrisiko gruppe. Dette er variablene «Tidspress/Vanskelige arbeidsforhold», «Risikofylt båthåndtering», «Miljøpåvirkning» og «Normal slitasje på anlegg». At disse verdiene vises er logisk med tanke på hvilke forhold som finnes på anlegget. Sterke naturkrefter er med på å øke farene med båthåndtering, samtidig som at det er fare for at det blir vanskelig å utføre arbeidsoperasjoner. Utstyr vil også påføres større slitasje. Årsaken for miljøpåvirkningens høye verdi kan finnes i temperaturforholdene. I og med at det var ugunstige temperaturforhold på anlegget vil det være større sannsynlighet for at flere fisk dør.

Figur 19 viser de totale risikoene for anlegg A. Som vi ser er disse for «Hull i not» 19% og for «Totalhavari» 14%. Disse verdiene er avrundinger. I tabell 4 ser vi den faktiske sannsynligheten fra modellen. Tabell 4 viser også risikoverdien, som er modellens verdi trukket fra de minimale verdiene. Den viser også avstand fra maksimumsverdien. Risikoverdien og avstanden fra maksimumsverdien sier noe om hvordan anlegget ligger i forhold til det spekteret som modellen kan måle. Det er disse verdiene som er mest interessante, og som kommer til å bli sammenlignet

med verdiene fra de andre anleggene.

Anlegg A	Verdi i modell	Risikoverdi	Avstand fra maksimumsverdi
Hull i not	18.74%	3.02	10.48
Totalhavari	14.33%	2.25	4.28

Tabell 4: Risikoverdier for anlegg A

6.3.2 Anlegg B

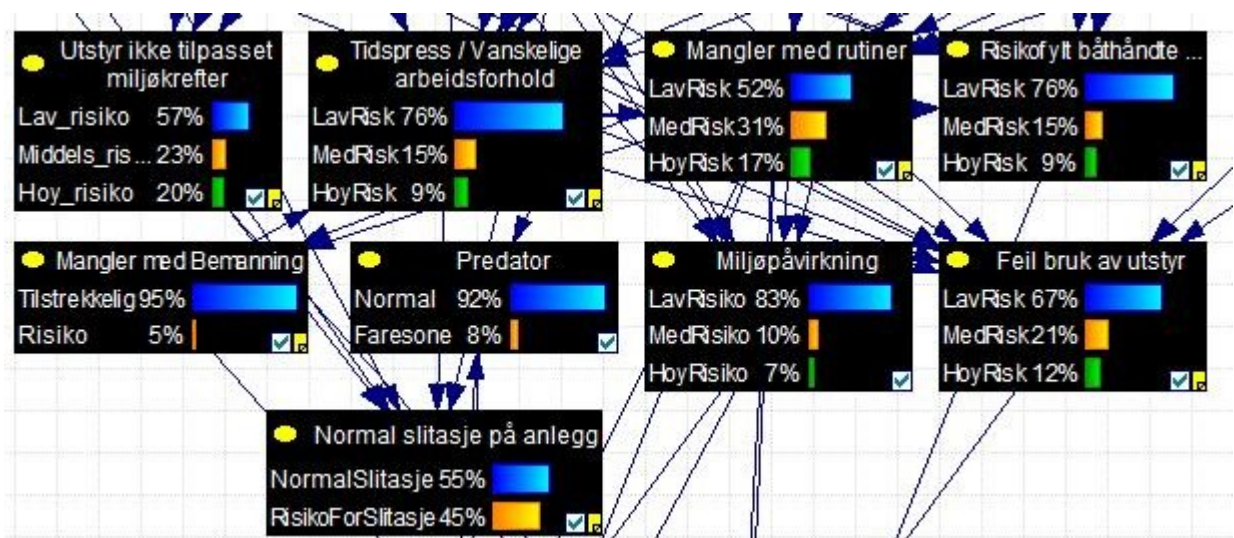
Anlegg B er et anlegg som har middels sterke, men ikke unormale miljøkrefter. De negative forholdene ved dette anlegget er at mye av utstyret er dårlig. Utstyret på anlegget har også kun dugelighetsbevis. Anlegget har problemer med kvaliteten på internkontrollen og har hatt mange avvik tidligere. Det er allikevel god opplæring på anlegget, det brukes utenforstående tjenesteytere som er kjent på anlegget, og de ansatte har lang erfaring innen næringen. Tabell 5 viser distribusjonen av tilstander i modellen for anlegg B.

Faktor:	Tilstand:
Alder på utstyr:	Gammelt
Dybde	Dyp
Høy biomasse	Akseptabel
Lokalitetsklassifisering	Klasse 2
Fortøyning	Manglende
Ekstraustyr	Risiko
Merder	Få merder
Bunnring	Bunnring
Is	Normal
Strømstyrke	Middels
Vind	Middels
Temperatur	Gunstig
Bølgehøyde	Middels
Trafikk	Ikke trafikkert
Område utsatt for predator	Lite utsatt
NYTEK sertifisering	Dugelighetsbevis.
Risikovurdering	Manglende
Hjemmelaget IK	Ikke hjemmelaget

Lokasjonbasert IK	Ikke lokasjonbasert
Alder IK	Utdatert
IK forskrift	Ikke forskrift
Avvik	Mange
Opplæring	God
Erfaring	God
Rekruttering	Lett
Økonomi	God
Brønnbåt	Ikke risiko
Dykker	Ikke risiko
Leverandør	Ikke risiko

Tabell 5: Distribusjon av tilstander i modellen for anlegg B

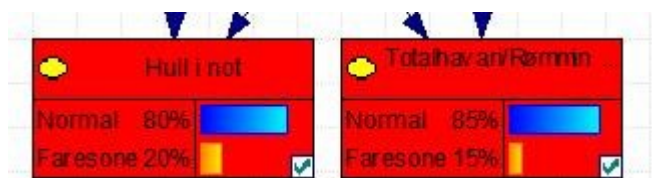
Som vi ser ut fra figur 20 er det her en del høye verdier for de forskjellige risikogruppene. «Normal slitasje på anlegg» har den høyeste verdien. Dette kommer av at anlegget er gammelt og utstyrmessig har mange dårlige egenskaper. «Mangler med rutiner» har også høye verdier som et resultat av at det kan være dårlig kvalitet på internkontrollen ved anlegget. Sannsynligheten for at internkontrollen er dårlig er representert i modellen av noden «Kvalitet på IK». Denne vises her i figur 21. Figur 21 viser også at det er sannsynlig at kvaliteten på klassifiseringen av anlegget er dårlig, noe som gjør at «Utstyr ikke tilpasset miljøkrefter» får forhøyede verdier. Anleggets totale verdier vises i figur 22 og i tabell 6.



Figur 18: Hjelpevariablenes verdier for anlegg B



Figur 19: Kvalitet på klassifisering og IK for anlegg B



Figur 20: De totale risikoverdiene for anlegg B

Anlegg B	Verdi i modell	Risikoverdi	Avstand fra maksimumsverdi
Hull i not	19.79%	4.07	9.43
Totalhavari	14.69%	2.61	3.92

Tabell 6: Risikoverdier for anlegg B

6.3.3 Anlegg C

Anlegg C skal representere et anlegg som har godt over gjennomsnittet med feil og mangler.

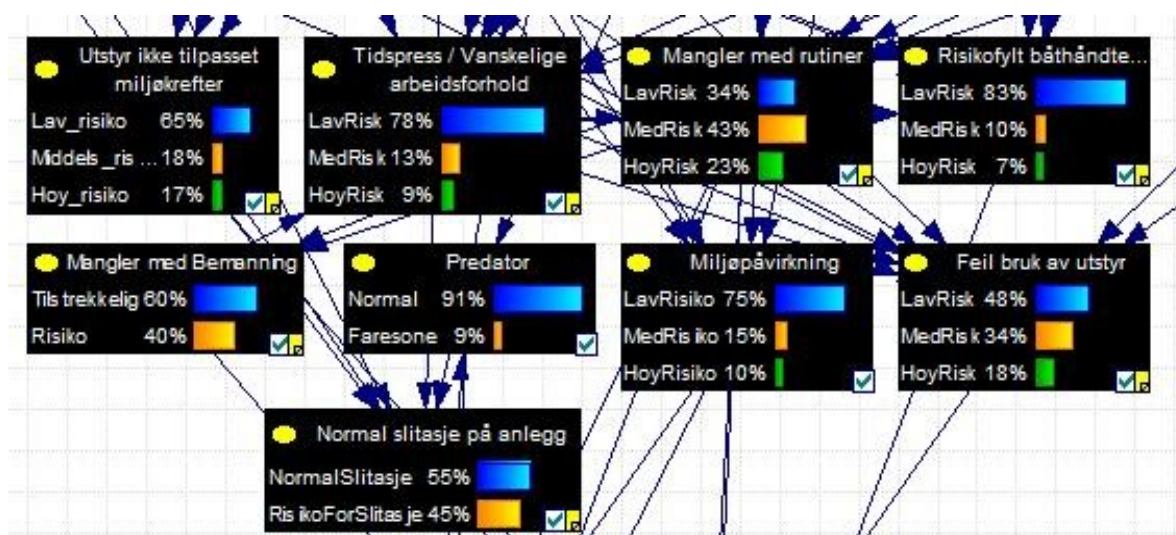
Anlegget vil allikevel ha lave naturkrefter som påvirker. Anlegget vil derfor ha dårlige tilstander for alle faktorene utenom de som sier noe om naturkreftene.

Faktor:	Tilstand:
Alder på utstyr:	Gammelt
Dybde	Dyp
Høy biomasse	Akseptabel
Lokalitetsklassifisering	Klasse 1
Fortøyning	Manglende
Ekstraustyr	Risiko

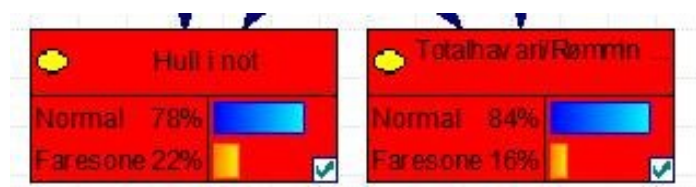
Merder	Få merder
Bunnring	Bunnring
Is	Normal
Strømstyrke	Lav
Vind	Svak
Temperatur	Gunstig
Bølgehøyde	Lav
Trafikk	Ikke trafikkert
Område utsatt for predator	Ikke utsatt
NYTEK sertifisering	Dugelighetsbevis
Risikovurdering	Manglende
Hjemmelaget IK	Ikke hjemmelaget
Lokasjonbasert IK	Ikke lokasjonbasert
Alder IK	Utdatert
IK forskrift	Ikke forskrift
Avvik	Mange
Opplæring	Dårlig
Erfaring	Dårlig
Rekruttering	Vanskelig
Økonomi	Dårlig
Brønnbåt	Risiko
Dykker	Risiko
Leverandør	Risiko

Tabell 7: Distribusjon av tilstander i modellen for anlegg C

Ut fra hjelpevariablene som er vist i figur 23 ser vi at anlegg C har høye risikoer for alle variablene, utenom «Risikofylt båthåndtering», «Utstyr ikke tilpasset miljø» og «Predator». Disse har ikke høye verdier på grunn av at de er avhengige av miljøet rundt anlegget, som i anlegg C sitt tilfelle ikke hadde noen problemer assosiert med seg. I figur 24 og tabell 8 ser vi de totale risikoene for anlegg C.



Figur 21: Hjelpesvariablene for anlegg C



Figur 22: De totale risikoene for anlegg C

Anlegg C	Verdi i modell	Risikoverdi	Avstand fra maksimumsverdi
Hull i not	21.56%	5.84	7.66
Totalhavari	15.99%	3.91	2.62

Tabell 8: Risikoverdier for anlegg C

6.3.4 Sammenligning av de vurderte anleggene

Etter at vurderingen av de forskjellige er gjort kan risikoverdiene hentes ut og sammenlignes. Det er på denne måten en kan bruke modellen for å finne ut hvilke anlegg som trenger tilsyn. Å gjøre dette kan for eksempel være nyttig for Kyst- og havbruksavdelingen ved Fiskeridirektoratet. Dersom de hadde hatt tilgang på de data som modellen krever kunne et system utført en slik vurdering og sammenlignet alle anleggene. Dette kunne ha blitt gjort hver gang data fra anleggene ble forandret eller ny data ble tilgjengelig. På denne måten kunne risikovurderingen ha blitt gjort på en dynamisk måte.

Selve sammenligningen gjøres ved å sammenligne risikoverdiene. Tabell 9 viser disse verdiene for

anlegg A, B og C som er brukt i kjøreeksempelen. Vi kan tydelig se at Anlegg C har de høyeste verdiene. I en reell vurderings situasjon ville det ha vært mange flere kontrollobjekter. Et eksempel på en måte å bruke disse dataene på er at en kunne sortert ut alle anleggene som hadde en risikoverdi over et visst nivå. Disse kunne ha blitt sett på som de som burde fokuseres på med tanke på kontroller og lignende. I vårt tilfelle kunne terskelen for fokusering for eksempel vært 4. Vi hadde da hatt to anlegg innenfor denne grensen: Anlegg B og C.

	Risiko for «Hull i not»	Risiko for «Totalhavari/annen rømning»
Anlegg A	3.02	2.25
Anlegg B	4.07	2.61
Anlegg C	5.84	3.91

Tabell 9: Sammenligning av risikoverdier for Anlegg A, B og C

7. Diskusjon

Risikovurdering innen havbruk har trolig ikke blitt systematisert på samme måte som jeg har gjort i denne oppgaven tidligere. Ved å samle kunnskapen som finnes rundt risiko for rømning av oppdrettslaks og utvikle et bayesiansk nett som modellerer dette har jeg vært med på å starte en prosess som til slutt kan ende i en sikrere og mer dynamisk måte å vurdere risiko for havbruksanlegg. Jeg vil nå diskutere de erfaringer jeg har gjort meg ved utføre kunnskapsmodelleringsprosessen.

7.1 Kunnskapsmodelleringsprosessen

Å utvikle en kunnskapsmodell ved bruk av bayesianske nett er en kreativ prosess. Etter kunnskapsmodelløren har samlet inn informasjon om domenet må denne settes sammen på en måte som både gjør modellen nyttig og som tar vare på all den relevante informasjonen som finnes om domenet. Å oppnå dette er en av grunnene til at jeg valgte å utføre modelleringsprosessen på den måten jeg gjorde. Ved å utvikle modellen i iterasjoner, der hver iterasjon resulterte i en prototype, ble det lett for meg å jobbe meg mot en fullstendig forståelse av domenet i samarbeid med ekspertene.

Jeg vil nå gå gjennom og diskutere forskjellige aspekter ved den prosessen jeg har vært gjennom. Jeg vil se på hvilke styrker og svakheter en slik prosess har ført med seg, hva som ble gjort på en god måte og hva som kunne ha blitt gjort bedre.

7.1.1 Oppbygning av prosessen

Oppbygning av selve prosessen har gått etter planen og har fungert bra. I og med at jeg startet fra et utgangspunkt der jeg ikke hadde noe kunnskap om rømning av oppdrettslaks var det nødvendig for meg å bruke mye tid i de første iterasjonene på å tilegne meg kunnskap om dette. Siden metoden kunnskapsmodellering ikke stiller særlige krav til resultatet av iterasjonene fikk jeg muligheten til å bruke tiden på å konsentrere meg om å lære meg domenet.

Etter hvert som jeg hadde nok kunnskap om domenet til at jeg selv kunne finne relasjoner mellom de forskjellige forholdene i domenet, kunne jeg begynne å konsentrere meg om kvaliteten til modellen.

Jeg vil si at en av hovedgrunnene til at denne prosessen fungerte bra var at jeg fikk muligheten til å bruke mye tid i de første iterasjonene på å sette meg inn i domenet. Metoden legger også opp til at kunnskapsmodelløren skal få hjelp av ekspertene til å tilegne seg kunnskap om domenet, noe som også gjør det svært mye lettere å sette seg inn i domenet.

Etter hvert som jeg beveget meg framover i prosessen brukte jeg mer og mer tid på selve modellen, og mindre tid på å sette meg inn i detaljer om domenet.

7.1.2 Interaksjon med ekspertene.

Jeg hadde god kontakt med ekspertene. Disse var tilgjengelige for meg i prosessen fra starten av. Samtaler med dem dreide seg i begynnelsen mest som sagt om å få et overblikk over domenet, og senere å skaffe seg en mer detaljert oversikt.

Ekspertene var svært behjelpelige med å utvikle modellen, men allikevel var det en teknologisk sperre mellom meg og dem. Det virket ofte som om de hadde problemer med å forstå hvordan modellen min skulle brukes, og hva nytten av denne ville være. På grunn av at det var vanskelig å forstå hvilke type relasjoner og faktorer modellen skulle inneholde, var det derfor også vanskelig for dem å gi konkrete svar på enkelte spørsmål jeg hadde.

Jeg fant ut at det var en fordel å unngå tekniske forklaringer på hvordan modellen fungerte, og heller stille spørsmål på en slik måte at jeg fant den informasjonen jeg trengte. For eksempel når modellen skulle evalueres var det bedre å vise en liste over relasjonene og faktorene enn å vise

modellen i sin helhet. Det kan hende at dette var på grunn av at en liste over de kausale forholdene i modellen var nærmere deres måte å tenke på enn å se et nettverk av disse. Det var lettere å forholde seg til en og en, enn alle samtidig.

Når det gjaldt sannsynligheter for de forskjellige variablene i modellene var det vanskelig hente ut informasjon ved hjelp av ekspertene. Dette var mye fordi at i den endelige modellen der vektene skulle legges til besto av generelle faktorer. For ekspertene, som ikke vanligvis jobber med sannsynligheter, var det vanskelig å gi et konkret svar på hvordan en målbar faktor påvirket en samling av hendelser.

For eksempel var det klart for dem at faktoren «Vind» gjorde at arbeid ble vanskeligere på et anlegg og derfor hadde en direkte innvirkning på faktoren «Tidspress / Vanskelige arbeidsforhold». Det var derimot vanskelig for dem å si noe konkret om hvor stor denne påvirkningen var. Dette var tilfellet med mange av faktorene. I noen tilfeller kunne de si noe om at en faktor hadde større påvirkning enn noen andre, eller at en bestemt faktor hadde liten eller stor påvirkning.

7.1.3 Problemer med kunnskapsmodelleringsprosessen

Det har gjennom min prosess vært få problemer. Det er allikevel et problem som må nevnes. Det har gjennom arbeidet mitt vært en teknisk barriere mellom meg og ekspertene jeg har jobbet med. Med dette mener jeg at de teknologier og verktøy jeg har brukt for å utvikle kunnskapsmodellene har vært ukjent og kan kanskje ha vært vanskelig å forstå for ekspertene.

På grunn av dette har det ofte vært forvirringer og misforståelser om hvordan data skal brukes i modellen, og dermed hvilke type data jeg trenger å vite noe om. Disse forvirringene har av og til gjort arbeid mindre effektivt på grunn av kommunikasjonsproblemer.

Disse problemene stammet også litt fra at ekspertene ikke var helt sikre på hva slags system jeg skulle utvikle. Mange, spesielt i begynnelsen, trodde jeg skulle være med på å utvikle en database til bruk i risikovurderingen. Dette stemte forsåvidt delvis, siden arbeidet jeg gjorde skulle være med på å identifisere hvilke data som skulle være med i en slik database. Det var allikevel ikke dette som var hovedmålet med min oppgave.

Når jeg utviklet modeller virket det ikke alltid som om at ekspertene så poenget med disse, noe som er naturlig dersom de ikke viste at jeg skulle utføre en kunnskapsmodellering. Jeg prøvde å løse dette problemet ved å gi en kort presentasjon på hvordan bayesianske nett fungerte, og hvorfor en modellering av kunnskapen er nødvendig. Dette gjorde etter hvert at det ble en bedre forståelse for hva jeg jobbet med, men den tekniske barrieren forsvant aldri helt.

For å unngått dette problemet fra starten av burde jeg ha brukt mer tid på å forklare målet mitt med oppgaven for ekspertene før jeg begynte med selve kunnskapsmodelleringen. På denne måten hadde de hatt en bedre forståelse for hva slags data jeg trengte og hvordan de kunne hjelpe meg. Jeg ville også brukt mer tid på å forklare de teknologiene jeg brukte fra starten av. Å gjøre dette kunne ha gjort prosessen mer effektiv.

7.2 Bruk av bayesianske nett for risikovurdering innen havbruk

Det fantes flere gode grunner for at denne modelleringen skulle gjøres i form av et bayesiansk nettverk. Bayesianske nett har i andre tilfeller med suksess blitt brukt for å vurdere risiko eller å løse lignende problemer (Tessem, 2009; Boondao, 2008; Onisko, 2008). Et av disse er et lignende prosjekt ved en annen avdeling innen Fiskeridirektoratet, hvor resultatene hadde vært gode.

Det var allikevel en negativ side ved å bruke bayesianske nett i denne oppgaven. Mange av metodene for utvikling og finjustering av et bayesiansk nettverk er avhengig av at det finnes innsamlede data tilgjengelig. For domenet jeg har jobbet med har dette ikke vært tilgjengelig. Samlinger av data er ikke et krav for å bruke bayesianske nett, siden dette også kan bli gjort ved bruk av eksperter og litteratur, men spesielt vekting av nettet kan bli vanskelig dersom data ikke finnes.

Jeg vil nå gå gjennom hva som har fungert med å bruke bayesianske nett, hva som ikke har fungert og hvordan en kan sammenligne arbeidet jeg har gjort med tidligere applikasjoner av denne teknologien.

7.2.1 Styrker ved bayesianske nett for risikovurdering innen havbruk.

Utviklingen av et bayesiansk nett har vært delt inn i tre deler. Først var det viktigst å identifisere de faktorer som skulle inkluderes, og opprette strukturen mellom dem. Denne strukturen består av kausale forhold mellom faktorene, som vil si at de påvirker hverandre på en eller annen måte. Det finnes mange forskjellige typer interaksjon mellom faktorer i slike nettverk. For eksempel kan en faktor ha positiv innvirkning på en annen, eller negativ innvirkning.

De fleste av forholdene mellom faktorene i min modell var enkle. Enten hadde en faktor positiv innvirkning, eller så hadde de en negativ innvirkning. I et par tilfeller ville også kombinasjonen av enkelte tilstander i forskjellige faktorer føre til en ekstra påvirkning.

Bayesianske nett har vist seg å være svært fleksible når det gjelder å modellere forskjellige slike kausale forhold (Tessem, 2009). Dette er på grunn av at disse forholdenes egenskaper vises igjen i

sannsynlighetene til en variabel, som lett kan håndteres. Gjennom hele arbeidet mitt hadde jeg ingen problemer med å modellere de forholdene jeg trengte med denne teknologien.

Bruken av programvaren GeNIe har også gjort denne prosessen enklere. Denne programvaren har flere nyttige verktøy som har gjort arbeidet med å utvikle modellen videre. Ved å bruke denne teknologien er det også lettere å eksportere modellen for videre bruk, for eksempel i et større system.

De neste to delene av utviklingen av det bayesianske nettet har vært å konfigurere faktorene og variablene. Dette er delt inn i vekting og diskretisering. Diskretisering av kontinuerlige variabler blir gjort for å forenkle nettverket. Det er mulig å bruke kontinuerlige variabler i bayesianske nett, men dette er ofte ikke nødvendig.

Proessen med å diskretisere variablene i modellen har vært ganske lett. Dette er mest på grunn av at den inneholder flere faktorer som har få tilstander. De kontinuerlige variablene som finnes har blitt delt inn i to eller tre variabler. At det finnes få mulige tilstander for faktorene i nettet er positivt, siden dette hjelper å holde kompleksiteten nede.

7.2.2 Svakheter ved bayesianske nett for risikovurdering innen havbruk.

Den største svakheten med å bruke et bayesiansk nett for risikovurdering for rømning av oppdrettslaks er at vektingen ikke er generert fra data. De metodene som finnes for å etablere vektorer for et bayesiansk nett som gir best resultater er ofte avhengige av data og kan ikke brukes i mitt tilfelle.

Det finnes også metoder for å bruke eksperter til å finne disse vektene. Jeg valgte å ikke gjøre dette i mitt arbeid hovedsaklig på grunn av at ekspertene var nølende med å sette konkrete tall på forhold mellom faktorer i modellen. Dersom jeg hadde brukt ekspertene til å gjøre dette ville modellen allikevel måttet finjusteres for å være nøyaktig.

Dersom det senere kommer data tilgjengelig kan disse vektene utbedres, noe som vil føre til at modellen kan testes og evalueres i sin helhet, og deretter kanskje tas i bruk av Fiskeridirektoratet som en del av et system for vurdering av risiko.

7.2.3 Andre implementasjoner av bayesianske nett.

Ut fra de resultatene jeg har sett av arbeidet jeg har gjort, finnes det mange likhetstrekk med andre lignende implementasjoner av bayesianske nett. I arbeidet gjort av Tessem for Ressursavdelingen ved Fiskeridirektoratet kan vi se dette (Tessem, 2009). Også i dette arbeidet ble det funnet at

eksperter hadde vanskelig for å forholde seg til konkrete sannsynligheter. Ekspertene ga som i mitt arbeid bedre respons når de ble bedt om å gi tilbakemeldinger om en faktors relative viktighet. Dette problemet er også nevnt av Boondao (Boondao, 2008), som brukte en 7-steps prosess kun for uthenting av sannsynligheter.

Tessem satte også på samme måte som meg selv vekt på modellens relasjoner, ut fra de utsagn ekspertene har kommet med angående avhengigheter og faktorerers viktighet. Han presiserer også at for at modellen skal være nyttig for bruk, bør det utføres maskinlæring- og teknikker for datainnsamling for å justere vektene. Dette er den samme argumentasjonen jeg har brukt for vektingen av min modell.

Nytten med å begynne kunnskapsmodelleringsprosessen med å sette seg inn i domenet som skal modelleres er også dokumentert i tidligere forskning. Onisko argumenterer for dette, og presiserer at spesifikt for modellering innen medisin er det anbefalt at kunnskapsmodelløren er kjent med terminologi, variabler og interaksjoner mellom disse (Onisko, 2008). Det kan også nevnes at hun her anbefaler interaksjon med eksperter som ikke direkte er med i kunnskapsmodelleringen. Jeg har fulgt begge disse anbefalingene for min prosess, og kan bekrefte at dette har hjulpet meg å oppnå gode resultater.

8. Konklusjon

Dette arbeidets mål har vært todelt. Det første målet var å utvikle et bayesiansk nett som også var en kunnskapsmodell som kan brukes til å utføre en risikovurdering for rømning av oppdrettslaks fra oppdrettsanlegg til havs. Dette målet var motivert av et ønske fra Fiskeridirektoratet om å utbedre deres systemer for risikovurdering innen havbruk. Det andre målet med oppgaven var å identifisere målbare faktorer som påvirket denne risikoverdien. Dette var motivert av at Fiskeridirektoratet ønsket å opprette en database med relevant informasjon, og trengte å vite hva de burde ha med i denne.

I oppgaven min har jeg oppnådd begge av disse målene. Gjennom å utføre en kunnskapsmodelleringsprosess, har jeg i iterasjoner kommet nærmere og nærmere en nøyaktig formalisering av den kunnskapen som finnes rundt oppdrettsanlegg og rømning av laks fra disse. Gjennom denne prosessen har jeg identifisert de faktorer som kan være målbare som påvirker risiko for rømning. Jeg har deretter brukt disse for å opprette et bayesiansk nett som kan brukes for å vurdere en total risiko for et anlegg når det gjelder rømning av oppdrettslaks.

Selv om de sannsynlighetene som brukes for å regne ut risikoverdien for anlegg per i dag ikke er

tilstrekkelig nøyaktige for å brukes av Fiskeridirektoratet, er modellen et godt eksempel på hvordan et bayesiansk nett kan brukes for å løse denne oppgaven. Jeg har også stor tro på at dersom det utføres en finjustering av sannsynlighetene som brukes i denne modellen vil denne med tiden være brukbar som et reellt verktøy for Fiskeridirektoratet. Dette avhenger selvfølgelig mye av hvilke data de har tilgjengelige i fremtiden.

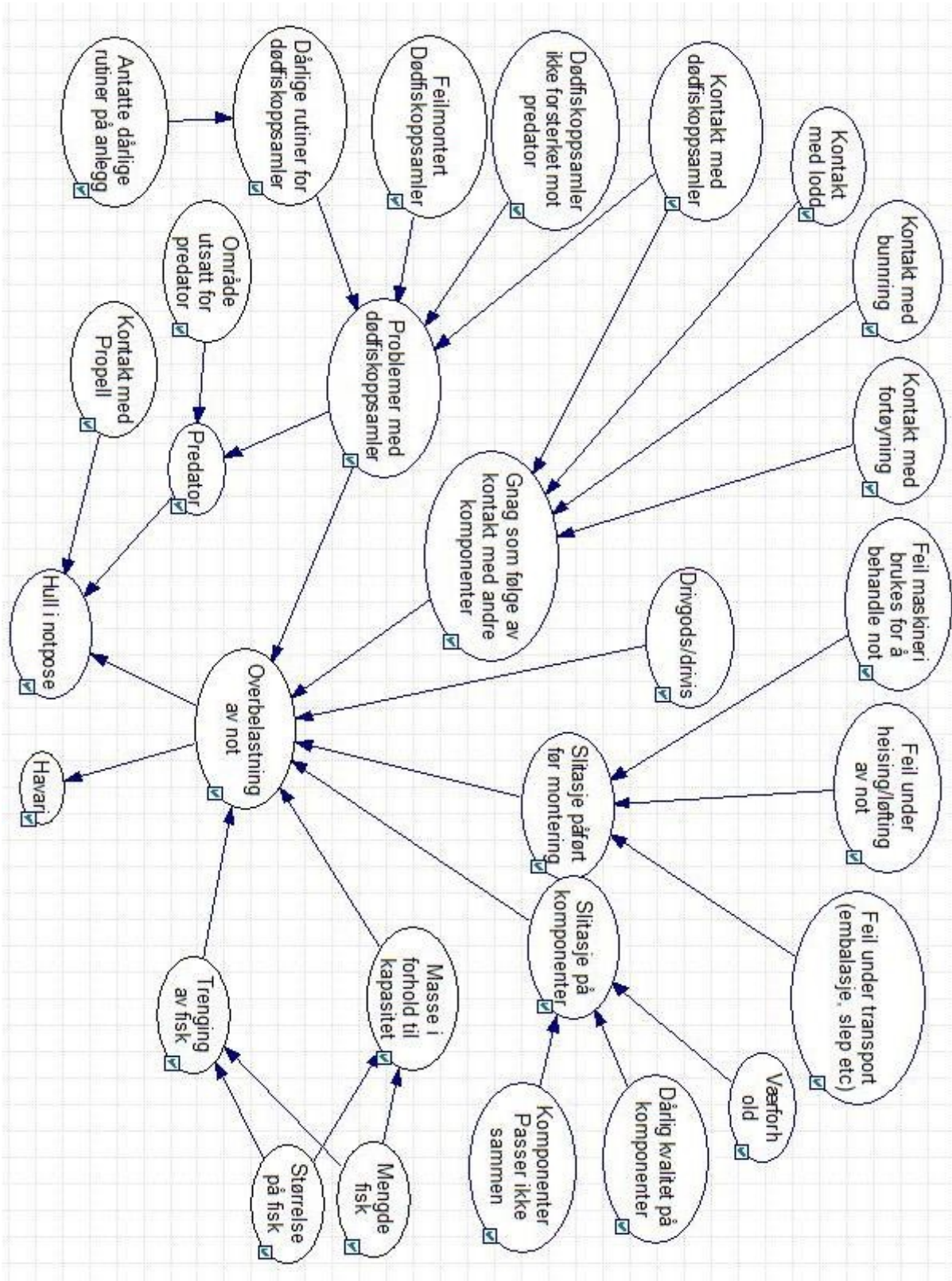
Proessen i med å oppnå disse målene har vært lang, og det kan nevnes både positive og negative sider med denne. Jeg fornøyd med strukturen i prosessen, og mener dette er en av grunnene til at modellen har den kvaliteten den har. Dersom jeg skulle utført en slik prosess igjen ville jeg brukt mer tid i begynnelsen på å informere ekspertene om hvordan resultatet var tiltenkt, og hva slags informasjon jeg trengte fra dem. Disse vanskelighetene er allikevel typiske for utviklingen av bayesianske nett, og alt i alt er jeg fornøyd.

9. Litteraturliste

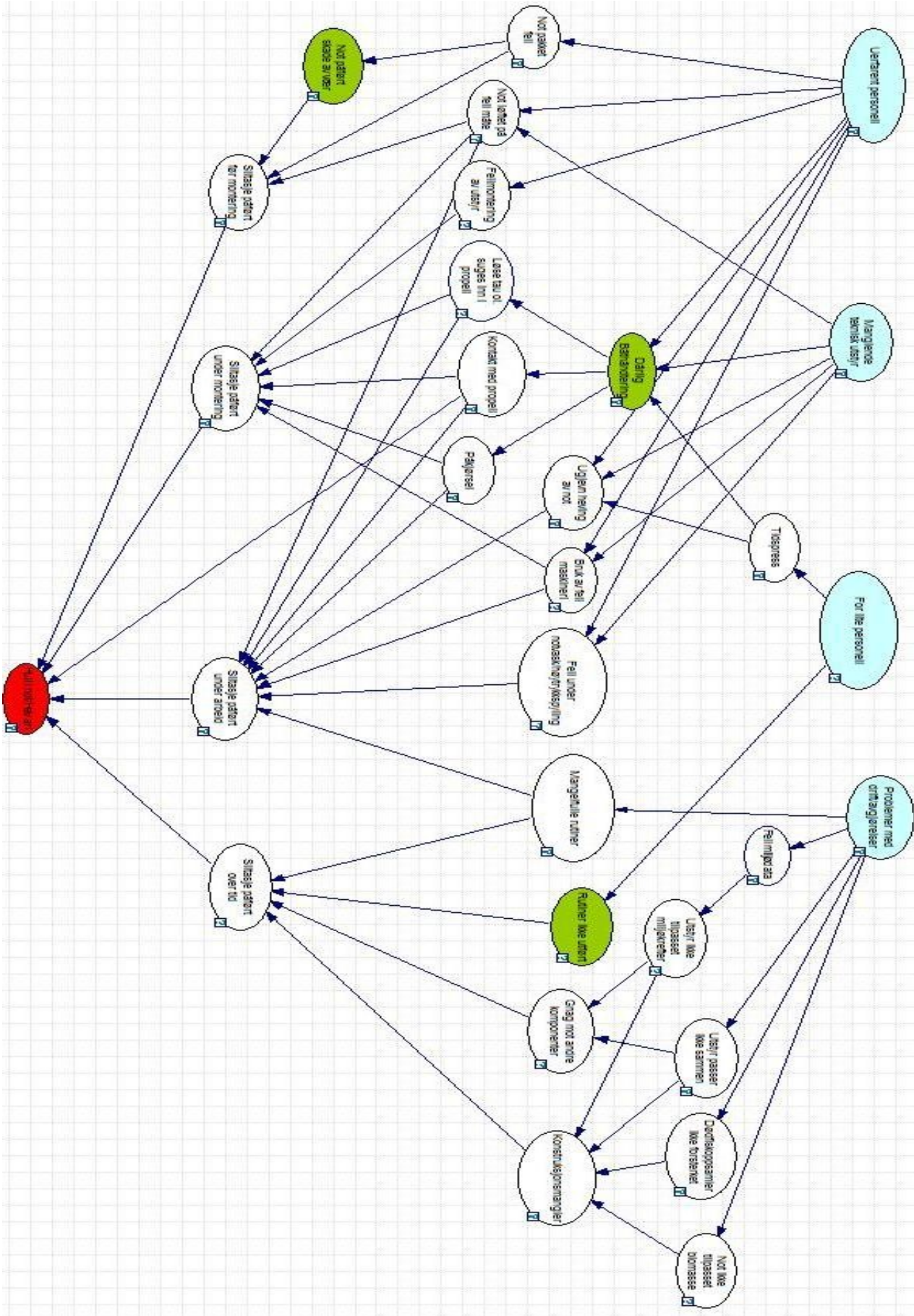
- BOONDAO, R. 2008. Crime Risk Factors Analysis. *In: O. POURRET, P. N. A. B. M. (ed.) Bayesian Networks: A Practical Guide to Applications.* John Wiley & Sons, Ltd.
- FISKERIDIREKTORATET 2004. NYTEK -Orientering om NYTEK.
- FISKERIDIREKTORATET 2010a. AkvaBest -Erfaringsbase for rømning.
- FISKERIDIREKTORATET 2010b. Laks, Regnbueørret og Ørret -Matfiskproduksjon.
- HEVNER, A. R., MARCH S.T. , J PARK, & RAM, S. 2004. Design science in information systems research. *Mis Quarterly*, 28, 75-105.
- KORB, K. B. & NICHOLSON, A. E. 2010. *Bayesian artificial intelligence*, Boca Raton, FL, CRC Press.
- ONISKO, A. 2008. Medical Diagnosis. *In: O. POURRET, P. N. A. B. M. (ed.) Bayesian Networks: A Practical Guide to Applications.* John Wiley and Sons, Ltd.
- POURRET, O., NAÏM, P. & MARCOT, B. 2008. *Bayesian networks : a practical guide to applications*, Chichester, England ; Hoboken, NJ, John Wiley.
- RABANAL, H. R. 1988. History of Aquaculture. FAO.
- SSB 2007. Primærnæringene: Fra landbruk til havbruk. Statistisk Sentralbyrå.
- SSB 2011. Eksport av varer uten skip og oljeplattformer. April og januar - april 2010 og 2011 Foreløpige tall. Millioner kroner (Rettet 16.05.2011 kl. 12.55). Statistisk Sentralbyrå.
- TESSEM, B. 2009. Risk Assessments in Fisheries Surveillance: A Case Study in Knowledge Engineering with Bayesian Networks.
- TESSEM, B. KAYSER, J. JOHE, P. & ULVATN, S. 2009. Developing Decision Support with Bayesian Networks in Fisheries Surveillance. *The Open Information Systems Journal*, 26-35.

10. Vedlegg:

10.1 Stor versjon av modellen fra iterasjon 1



10.2 Stor versjon av modellen fra iterasjon 2



10.3 Stor versjon av modellen fra iterasjon 3 (Endelig modell)

