

Modeller for økonomiske analyser av vedlikeholdsstrategier og vedlikeholdsbudsjetter



Modeller for økonomiske analyser av vedlikeholdsstrategier og vedlikeholdsbudsjetter

Harald Minken
Geir Dahl
Nils Gaute Voll

Transportøkonomisk institutt (TØI) har opphavsrett til hele rapporten og dens enkelte deler. Innholdet kan brukes som underlagsmateriale. Når rapporten siteres eller omtales, skal TØI oppgis som kilde med navn og rapportnummer. Rapporten kan ikke endres. Ved eventuell annen bruk må forhåndssamtykke fra TØI innhentes. For øvrig gjelder [åndsverklovens](#) bestemmelser.

Tittel: Modeller for økonomiske analyser av vedlikeholdsstrategier og vedlikeholdsbudsjetter

Forfattere: Harald Minken
Geir Dahl
Nils Gaute Voll

Dato: 12.2014

TØI rapport: 1380/2014

Sider 52

ISBN Elektronisk: 978-82-480-1594-9

ISSN 0808-1190

Finansieringskilde: Samferdselsdepartementet

Prosjekt: 3638 - Nytten av å øke vedlikeholdsbudsjettene i transportsektoren

Kvalitetsansvarlig: Kjell Werner Johansen

Emneord: Rehabilitering
Samfunnsøkonomi
Vedlikehold

Sammendrag:

Er det samfunnsøkonomisk lønnsomt å øke vedlikeholdsbudsjettene? Hvordan kan vedlikeholdet planlegges slik at midlene brukes optimalt? I denne rapporten har vi lagt til rette for et framtidig arbeid med disse spørsmålene ved å skissere en enkel og praktisk metode for å vurdere om budsjettet bør økes, ved å etablere de viktigste byggesteinene i en større og mer realistisk modell, og ved å gi framtidige arbeidere på feltet oversikt over litteraturen.

Title: Models for economic analysis of maintenance plans and maintenance budgets

Author(s): Harald Minken
Geir Dahl
Nils Gaute Voll

Date: 12.2014

TØI report: 1380/2014

Pages 52

ISBN Electronic: 978-82-480-1594-9

ISSN 0808-1190

Financed by: Ministry of Transport and Communications

Project: 3638 - Nytten av å øke vedlikeholdsbudsjettene i transportsektoren

Quality manager: Kjell Werner Johansen

Key words: Economic analysis
Maintenance
Rehabilitation

Summary:

From the point of view of social efficiency, ought we to increase maintenance budgets in the transport sector? How can we devise socially efficient maintenance and rehabilitation plans? In this report, we pave the way for future work on these problems by setting out a simple and practical method to assess if maintenance budgets ought to be increased, by establishing the foundations for a larger and more realistic model, and by providing future workers in the field with a survey of literature.

Language of report: Norwegian

Rapporten utgis kun i elektronisk utgave.

This report is available only in electronic version.

Transportøkonomisk Institutt
Gaustadalleen 21, 0349 Oslo
Telefon 22 57 38 00 - www.toi.no

Institute of Transport Economics
Gaustadalleen 21, 0349 Oslo, Norway
Telefon 22 57 38 00 - www.toi.no

Forord

I to prosjekter under Samferdselsdepartementets program for overordnet transportforskning (POT) har TØI i samarbeid med professor Geir Dahl, UiO, utviklet metoder for å finne optimale vedlikeholdsstrategier og beregne den samfunnsøkonomiske nytten av å øke vedlikeholdsbudsjettene i samferdselssektoren. Prosjektet "Vedlikeholdsstrategier" pågikk fra seinhøsten 2005 til våren 2008 og er dokumentert i TØI-rapport 957/2008. Oppfølgingsprosjektet "Nytten av å øke vedlikeholdsbudsjettene i transportsektoren", som dokumenteres her, startet høsten 2010. Det blei lagt på is høsten 2011 og i 2012 og 2013, men blei gjenopptatt i 2014. Som vi redegjør for i denne rapporten, har prosjektet ikke nådd alle sine mål, men sluttrapporten gir etter vårt syn gode muligheter for at samferdselsetatene skal kunne gjøre enkle analyser av nytten av å øke vedlikeholdsbudsjettene, samtidig som den legger grunnlaget for forskning på, og anvendelser av, mer avanserte modellverktøy.

Cand oecon Harald Minken har vært leder for prosjektet og har skrevet kapittel 1, 2 og 5. Professor Geir Dahl har skrevet kapittel 3, og forsker Nils Gaute Voll har skrevet kapittel 4. Assisterende direktør Kjell Werner Johansen har stått for kvalitetssikring av rapporten, og sekretær Trude Rømming har stått for den avsluttende tekstbehandlingen.

Oslo, desember 2014
Transportøkonomisk institutt

Gunnar Lindberg
direktør

Kjell Werner Johansen
avdelingsleder

Innhold

Sammendrag

1	Bakgrunn	1
1.1	Prosjektet "Vedlikeholdsstrategier"	1
1.2	Oppfølgingsprosjektet	2
1.3	Disposisjon av rapporten	3
2	Skissemessig vurdering av nytten av å øke vedlikeholdsbudsjettet	6
2.1	Optimal rehabiliteringstakt for et enkelt objekt	6
2.2	Mange objekter og et felles budsjett	8
2.3	Nytten av å øke budsjettet	9
3	Optimering: modell og metode	11
3.1	Kort beskrivelse	11
3.2	Formål med MOV	12
3.3	Hvorfor er dette vanskelig	12
3.4	Modellen	14
3.5	Kommentarer til modellen	15
3.6	Heltallig lineær optimering og MOV	16
3.7	Lagrange-relaksasjon	17
3.8	Heuristikk	18
3.9	Videreutvikling av algoritmene	19
4	Praktisk eksempel	20
4.1	Testsituasjonen	20
4.2	Optimal allokering av budsjettmidler	21
4.3	Konklusjon	24
5	Litteraturoversikt	25
5.1	Definisjoner og begrepsavklaringer	26
5.2	Klassifisering av litteraturen	27
5.3	Generell litteratur	28
5.4	Veg. Nedbrytning og effekt av tiltak	28
5.5	Veg. Optimal vedlikeholdsstrategi for ett objekt	33
5.6	Veg, mange objekter	34
5.7	Jernbane	37
5.8	Bru	38
5.9	Flyplass	38
5.10	Kommersielle vedlikeholdsstyringsystemer (PMS)	38
5.11	Konklusjon	39
A	Vedlegg A	
	DYNOPT – documentation	41

B Vedlegg B

Litteratur som det er vist til i kapittel 5

45

Sammendrag:

Modeller for økonomiske analyser av vedlikeholdsstrategier og vedlikeholdsbudsjetter

TØI rapport 1380/2014

Forfatter(e): Harald Minken, Geir Dahl, Nils Gaute Voll

Oslo 2014 52sider

Er det samfunnsøkonomisk lønnsomt å øke vedlikeholdsbudsjettene? Hvordan kan vedlikeholdet planlegges slik at midlene brukes optimalt? I denne rapporten har vi lagt til rette for et framtidig arbeid med disse spørsmålene ved å skissere en enkel og praktisk metode for å vurdere om budsjettet bør økes, ved å etablere de viktigste byggesteinene i en større og mer realistisk modell, og ved å gi framtidige forskere på feltet oversikt over litteraturen.

Dette er sluttrapporten i prosjektet "Nytten av å øke vedlikeholdsbudsjettene i transportsektoren", som var finansiert gjennom Samferdselsdepartementets POT-program og har pågått (med et lengre avbrudd) fra seinhøstes 2010 til 2014. Dette prosjektet var en oppfølging av et tidligere prosjekt, kalt "Vedlikeholdsstrategier", som pågikk fra høsten 2005 til våren 2008.

Oppgava i begge prosjektene var å finne den samfunnsøkonomisk beste måten å vedlikeholde en samling av likeartede objekter på over et antall år, gitt at det måtte skje innafor gitte årlige budsjetter. I "Vedlikeholdsstrategier" utviklet vi et dataverktøy for å finne slike optimale vedlikeholdsstrategier, og anvendte det på noen enkle eksempler fra vegsektoren. Målet var å kunne beregne nytten av å øke vedlikeholdsbudsjettene. Tankegangen var at for å kunne si noe om det, må vi gjøre antakelser om hvordan budsjettmidlene vil bli brukt. Og den antakelsen som vi så som den rimeligste, var at midlene ville bli brukt på beste måte, dvs. i tråd med en optimal vedlikeholdsstrategi. Først måtte vi altså ha verktøyet som fant optimale vedlikeholdsstrategier, og deretter ville eksperimenter med å bruke ulike budsjetter i verktøyet kunne gi oss nytten av en budsjettøkning.

Begge prosjekter befinner seg i et fagfelt på grensa mellom ingeniørkunnskap, operasjonsanalyse og økonomi. Å formulere en god vedlikeholds- og rehabiliteringsplan er vanskelig nok i seg sjøl. Det krever ingeniørkunnskaper om tilstandsmåling, nedbrytning av infrastrukturen og virkninger av tiltak. Men å finne en samfunnsøkonomisk optimal plan som skal kunne gjennomføres innafor gitte budsjetter, og undersøke om det er samfunnsøkonomisk lønnsomt å øke vedlikeholdsbudsjettene, krever som hovedregel langt mer kompliserte matematiske og operasjonsanalytiske verktøy enn det vi har vært vant til i samferdselssektoren i Norge. Som vist i Dahl og Minken (2008), er det eksistensen av årlige budsjettrestriksjoner som, sammen med blant annet at det finnes mange slags mulige tiltak å velge mellom, gjør problemet så komplisert. Likevel må vi ha med budsjettrestriksjoner i modellen om vi skal si noe fornuftig om å endre budsjettene. Uten budsjettrestriksjoner kan vi si noe om hvordan vedlikeholdspolitikken ideelt

burde innrettes, eller om hva som er det minste beløpet som må til for å oppnå en på forhånd gitt vedlikeholdsstandard, men det er dessverre ikke realistiske problemstillinger.

Kapittel 2 i denne rapporten går så langt som mulig i retning av å forenkle analyseverktøyet slik at en vurdering av om det er lønnsomt å øke vedlikeholdsbudsjettene skal kunne la seg gjennomføre med det kunnskapsgrunnlaget og det datagrunnlaget vi har nå.

Analyseverktøyet er ikke basert på detaljerte opplysninger om de enkelte objektene som skal vedlikeholdes, og pretenderer ikke å resultere i noen realistisk vedlikeholdsplan. Det som trengs, er snarere gjennomsnittlige sammenhenger for store klasser av objekter. Videre finnes det ikke noe bindende budsjett for det enkelte året, men bare en gjennomsnittlig årlig budsjettramme som skal overholdes i det lange løp. Det er en underliggende forutsetning at vedlikeholdsbudsjettene i utgangspunkt er tilstrekkelige til at et stabilt nivå på kvaliteten av infrastrukturen er mulig i det lange løp.

Verktøyet kan programmeres i EXCEL og gir umiddelbart en nyttekostnadsbrøk for tiltaket som skal analyseres, nemlig å øke gjennomsnittsbudsjettet med ei krone. I tillegg gir den optimal rehabiliteringsfrekvens for hver av klassene av objekter.

Den enkle modellen i kapittel 2 kan ikke erstatte en modell som tar begrensningene i de årlige budsjettene mer på alvor, som tar hensyn til at det finnes mange typer av tiltak, og som tar utgangspunkt i at objektene i utgangspunktet kan befinne seg i ulike tilstander. Bare en slik modell kan si noe om hva som er den beste strategien i et virkelig tilfelle, og bare i en slik modell er det mulig å studere slike ting som hvordan man på beste måte tar inn et vedlikeholdsetterslep.

Kapittel 3 og 4 rapporterer om arbeidet med et mer detaljert og realistisk verktøy. Det skal ikke bare kunne foreta en grov vurdering av lønnsomheten av å øke budsjettet, men også kunne brukes til prioritering av innsatsen til konkrete infrastrukture objekter, eller i alle fall til relativt homogene grupper av slike objekter. Videre kan det brukes til å angi hva slags tiltak som det lønner seg å bruke på det enkelte objekt, eller i alle fall på den enkelte gruppe av objekter.

Allerede i det første prosjektet etablerte vi et slikt verktøy. Men det var bare programmert for én tilstandsvariabel, dvs. infrastrukturens tilstand kunne ikke beskrives i flere dimensjoner. For veg betyr det for eksempel at vi kunne karakterisere vegene etter hvor ujamne de var i lengderetningen (såkalt IRI), men ikke på tvers (spordybde). Vi kunne altså ikke anvende all relevant kunnskap om vegens tilstand. I det nye prosjektet var en av oppgavene å rette på dette, slik at tilstanden kunne beskrives med en vektor med flere elementer. Vi ville også gjøre verktøyet mer generelt, slik at det kunne brukes ikke bare på vegdekker, men på hvilken som helst samling av noenlunde likeartede infrastrukture objekter. Kapittel 3 viser at vi har løst denne grunnleggende oppgava, og programmert et dataverktøy som kan løse et svært bredt spekter av slike problemer, og som er helt på høyden med de beste slike verktøy i utlandet. Men vi mangler ennå brukergrensesnittet som skal gjøre det enkelt og fleksibelt å benytte i praksis. Vi mangler derfor også erfaringer med verktøyet.

Kapittel 4 viser likevel at vi har gjennomført en rekke kjøring av programmet. Vi har sannsynligvis ikke funnet noen lovmessigheter om hva som utgjør en riktig vedlikeholdsstrategi eller noen tommelfingerregel for hvordan budsjettet skal tilpasses. Det var heller ikke å vente. Det er såpass mange former for avhengighet mellom tiltakene og såpass begrensede muligheter til å justere omfanget av tiltakene opp og ned i størrelse til at det lett kan bli store konsekvenser av små forandringer og omvendt. (Det som gjøres på ett objekt i år, bestemmer hvor mye som trenger gjøres i seinere år, og det innskrenker også mulighetene for omfattende tiltak på andre objekter i samme år. Bestemte krav til slutt-tilstanden bidrar sammen med budsjettet til dette.)

Dette har bidratt til beslutningen om å angripe spørsmålet om nytten av å endre budsjettet med enklere verktøy, slik vi gjør i kapittel 2. Om vi går for langt i forenkling, kan bare erfaringen gi svar på. Samtidig vil den mer realistiske modellen i kapittel 3 gi mulighet til optimering av virkelige vedlikeholds- eller rehabiliteringsstrategier, og kanskje gi en kontroll på regnestykkene med den enkle modellen.

Kapittel 5 er ment som en hjelp til de som forhåpentligvis vil arbeide videre med dette, enten det er analyser av den typen som er beskrevet i kapittel 2, eller det er ferdigstilling og bruk av verktøyet i kapittel 3 og 4. Det er en litteraturoversikt som både skal dekke den nyeste litteraturen, prøve å fastlegge noen definisjoner, og gi noen korte blikk på emnets historie. Vi veit at vi har utelatt mye, men sannsynligvis ikke noe vesentlig som en ikke vil kunne nøste opp ved hjelp av litteraturlistene i de arbeidene vi har tatt med.

Hovedsakelig er det den akademiske litteraturen som er dekket. Det er en ganske betydelig kløft mellom den akademiske litteraturen og litteraturen som beskriver kommersiell programvare eller praktisk utredningsarbeid på feltet, men vi tror at vegen fram er å ta i bruk den avanserte teorien. Noe mer utredningspreget stoff har vi likevel tatt med, sammen med det vi kjenner til av lærebøker.

Det er vårt håp at til tross for problemene vi har møtt i dette prosjektet, vil det bli arbeidet videre med modellene, brukergrensesnittene og de økonomiske analysene av vedlikeholdspolitikken. Gjennom å angi en enkel veg til analyser av budsjettet i kapittel 2, legge til rette for fullføring av arbeidet med den mer realistiske modellen i kapittel 3 og 4, og gi nye arbeidere på feltet mulighet til å skaffe seg en oversikt over fagfeltet og dets utvikling i kapittel 5, har vi gjort hva vi har kunnet så langt.

1. Bakgrunn

Dette er sluttrapporten i prosjektet "Nytten av å øke vedlikeholdsbudsjettene i transportsektoren", som var finansiert gjennom Samferdselsdepartementets POT-program og har pågått (med et lengre avbrudd) fra seinhøstes 2010 og ut år 2014. Dette prosjektet var en oppfølging av et tidligere prosjekt, kalt "Vedlikeholdsstrategier", som pågikk fra høsten 2005 til våren 2008.

1.1 Prosjektet "Vedlikeholdsstrategier"

I "Vedlikeholdsstrategier" stilte vi oss oppgava å finne den samfunnsøkonomisk beste måten å vedlikeholde en samling av likeartede objekter på over et antall år, gitt at det måtte skje innafor gitte årlige budsjetter. Vi utviklet et dataverktøy for å finne slike optimale vedlikeholdsstrategier, og brukte det på noen enkle eksempler fra vegsektoren. Målet var å kunne beregne nytten av å øke vedlikeholdsbudsjettene. Tankegangen var at for å kunne si noe om det, må vi gjøre antakelser om hvordan budsjettmidlene vil bli brukt. Og den antakelsen som vi så som den rimeligste, var at midlene ville bli brukt på beste måte, dvs. i tråd med en optimal vedlikeholdsstrategi. Først måtte vi altså ha verktøyet som fant optimale vedlikeholdsstrategier, og deretter ville eksperimentere med å bruke ulike budsjetter i verktøyet kunne gi oss nytten av en budsjettøkning.

Det er nokså sjølsagt at om vi bruker et vedlikeholdsplanleggingsprogram som ikke tar hensyn til at det finnes budsjetter, eller som tar hensyn til det på en feilaktig måte, vil vi aldri kunne si noe fornuftig om vedlikeholdsbudsjettene bør økes. Vi kan kanskje beregne kostnaden for å oppnå en på forhånd fastsatt standard på alle objekter, men vi vil ikke vite om denne standarden er den samfunnsøkonomisk mest fornuftige. Vi kan heller ikke vite hva som ville vært det samfunnsøkonomisk beste resultatet for et bestemt budsjettnivå dersom det var mulig å fordele midlene optimalt mellom objektene over flere år. Uten et budsjett som binder objektene sammen er optimalt vedlikehold et relativt enkelt problem på papiret (om vi ser bort fra mangelfulle data og usikre effektsamband), for strategien for hvert objekt kan legges opp uten å ta hensyn til konsekvensene for andre objekter. Men når vi tar hensyn til budsjettene, blir problemet straks svært komplisert, og krever svært avanserte modeller og løsningsalgoritmer.

Vår vitenskapelige artikkel fra det første prosjektet, Dahl og Minken (2008), er publisert i ett av de aller fremste operasjonsanalytiske tidsskriftene i verden. Problemet å beregne gevinsten ved å øke vedlikeholdsbudsjettene, slik vi formulerte det, var altså et godt stykke fra problemene som de vedlikeholdsansvarlige i etatene var vant med. Tilsvarende hyllevare fantes ikke, og det å ta i bruk vår modell slik den forelå, krevde relativt spesialisert kompetanse fra brukerne.

Nå gjorde vi en del i det første prosjektet for å formidle arbeidet vårt på en måte som kunne forstås av flere. Det viktigste er kapittel 3 i rapporten fra prosjektet, Minken, Dahl og Steinsland (2008). Vi hadde også en artikkel i et nordisk tidsskrift, Dahl og Minken (2007). Den genererte en forespørsel fra et dansk kommersielt selskap, men ingen henvendelser for øvrig.

TØI-rapporten av Minken, Dahl og Steinsland (2008) handler imidlertid om mer enn dataprogrammet vi utviklet. Den inneholder også en breiere utledning og drøfting av økonomiske prinsipper for planlegging av vedlikehold av samferdselsinfrastruktur. Levetider, eksistensverdi og kapitalverdi av infrastrukture objekter, og grunnprinsipper for nyttekostnadsanalyse av utbedringer og standardendring blei behandlet i kapittel 2. Begrepet vegkapital, sårbarhetsanalyse og effektivitetsmål for

vegnettverk er behandlet i kapittel 4. Spesielt vil vi minne om at vi i avsnitt 4.3 framsatte et forslag til å forbedre analysene av i hvilken grad vegbygging bidrar til økonomisk vekst. Forslaget går ut på å bruke et annet begrep om samferdselskapital i analysen, og se om det gir en bedre sammenheng med økonomisk vekst enn den som blei funnet den gangen. I stedet for anleggskostnaden (eventuelt nedskrevet) bør kapitalverdien beregnes prospektivt, som differansen mellom eksistensverdien og anleggskostnaden. Anleggskostnaden kan da justeres ned med en nyttekostnadsbrøk som sier hvor stor del av anleggskostnaden som faktisk har blitt til kapital.

1.2 Oppfølgingsprosjektet

Oppfølgingsprosjektet "Nytten av å øke vedlikeholdsbudsjettene i transportsektoren", som dokumenteres her, startet høsten 2010. Det blei lagt på is etter et snaut år, men blei gjenopptatt i 2014. Som vi redegjør for i denne rapporten, har prosjektet ikke nådd alle sine mål, men sluttrapporten gir etter vårt syn gode muligheter for at samferdselsetatene skal kunne gjøre enkle analyser av nytten av å øke vedlikeholdsbudsjettene, samtidig som den legger grunnlaget for forskning på, og anvendelser av, mer avanserte modellverktøy.

Prosjektet tok sikte på å anvende det verktøyet som var utviklet i det første prosjektet på flere transportområder. Verktøyet egner seg nemlig ikke bare for vedlikehold av vegdekker eller veger mer generelt, men, med høvelig omprogrammering, vil det kunne brukes på en hvilken som helst samling av likeartede objekter som skal vedlikeholdes eller rehabiliteres innafør rammene av felles årlige budsjetter. I prosjektet vi foreslo, ville vi omprogrammere det verktøyet som var utviklet, slik at brukeren sjøl kunne velge funksjonsformer som passer for en hvilken som helst aktuell anvendelse. Dette arbeidet var faktisk allerede påbegynt i etterkant av det forrige prosjektet.

Betrakt et sett av infrastrukture objekter som vil måtte dele på felles årlige vedlikeholdsbudsjetter. Objektene er gjenstand for nedbrytning som følge av trafikkbelastningen og andre forhold, og denne nedbrytningen avhenger også av hvor solid de er bygd i utgangspunktet. Objektene tilstand på et gitt tidspunkt kan måles ved en indikator som kan ha en eller flere dimensjoner. Nedbrytningstakten for objektene kan være forskjellig. Initialtilstandene kan være hva som helst mellom dårligst mulige og best mulige tilstand. Jo verre tilstand, jo større brukerkostnader. Myndighetene rår over en rekke mulige vedlikeholdstiltak med ulik intensitet og kostnad. En vedlikeholdsstrategi er en allokering av vedlikeholdstiltak til hvert av objektene i et visst antall år framover. Problemet er å finne en vedlikeholdsstrategi som minimerer summen av brukerkostnader og vedlikeholdskostnader, gitt budsjettbetingelsene og eventuelle krav til slutttilstanden for hvert av objektene.

I det opprinnelige prosjektet målte vi objektene tilstand med én og bare én tilstandsvariabel. En av oppgavene i det nye prosjektet var å omprogrammere slik at det var mulig å måle tilstanden i flere dimensjoner. Konkret kom vi etter hvert til at inntil fire dimensjoner kunne være aktuelle: For det første trenger vi minst én indikator som påvirker brukernes kostnader og opplevelse når objektet tas i bruk. I det første prosjektet var det vegdekkets jevnhet, målt ved IRI (international roughness indicator) eller en omforming av IRI. For det andre kunne det trenges en annen indikator for konstruksjonens styrke eller bæreevne. Den henger sammen med den første ved at bæreevnen svekkes av visse typer av trafikk (tungtrafikk), og at denne svekkelsen kan påvirke hvor fort brukernes kostnader og opplevelser forringes med tida. For det tredje vil tida siden siste fullstendige rehabilitering ("alderen") være en høvelig indikator for å fange opp faktorer som ikke er eksplisitt uttrykt på annen måte. Det er i virkeligheten ikke slik som vi antok i det første prosjektet, at en viss trafikkbelastning eller ett visst vedlikeholdstiltak vil ha samme virkning uansett hvor lenge det er siden en full rehabilitering. Og for det fjerde vil naturligvis trafikkvolumet, eller mer generelt bruksintensiteten, påvirke objektets tilstand og nedbrytningstakt.

Opgaven å legge programmet til rette for flere tilstandsvariable er løst. Se nærmere om det i kapittel 3. Videre var oppgava å legge til rette for ulike anvendelser - vegvedlikehold, vedlikehold på jernbanen, vedlikehold av bruer og tunneler. Det innebærer å fjerne alle sektorspesifikke trekk fra programmet. Dette er også gjort.

Til gjengjeld må naturligvis brukeren da legge inn de egenskapene ved objektene, nedbrytningen og tiltakene som er aktuelle for hennes anvendelse. Det krever et ganske generelt og fleksibelt brukergrensesnitt. Det er på dette området prosjektet ikke har lyktes.¹

De neste arbeidsoppgavene går ut på å bruke det nye grensesnittet til å programmere opp anvendelser på veg, jernbane, bru og tunnel og bruke det til å gjennomføre utvalgte analyser på disse områdene. Det innebærer å spesifisere nedbrytingsfunksjoner, brukerkostnader og mulige tiltak og tiltakskostnader på hvert av disse anvendelsesområdene, og gjennomføre konkrete analyser på det grunnlaget. Spesielt ønsket vi å finne nyttekostnadsbrøken av en marginal budsjettøkning og optimal utforming av et løft for å ta inn vedlikeholdsetterslepet. I arbeidsplanen var det regnet med at vi skulle samarbeide nært med etatene om å spesifisere de nødvendige funksjonene, og at etatene skulle settes i stand til å gjøre egne analyser. Fram til sommeren holdt vi flere møter med vedlikeholdsfolk i vegvesenet og Jernbaneverket og forskere som arbeidet for dem. Men vi kom aldri så langt at vi klarte å få dem aktivt med i prosjektet. Det er mulig det ville vært annerledes om vi hadde hatt et generelt grensesnitt som de kunne være med på å spesifisere. Det er også mulig at vi hadde urealistiske forestillinger om hva vi kunne vente av dem.

Vi utarbeidet uansett et rammeverk for en spesifisering av de nødvendige funksjonene og kostnadsartene som vi mener vi være til god hjelp for et eventuelt framtidig arbeid med dette. En god del av dette brukte vi seinere i et mindre oppdrag for metodegruppa i NTP i slutten av 2011 - se Minken, Meyer, Veisten og Bai (2011). Denne rapporten er faktisk et langt stykke på veg et produkt av arbeidet i vedlikeholdsprosjektene. For et eventuelt framtidig arbeid med optimale vedlikeholdsstrategier, i regi av TØI eller etatene sjøl, viser vi spesielt til kapittel 2 og vedlegg 2 og 3 i den rapporten.

1.3 Disposisjon av rapporten

Kapittel 2

Da vi gjenopptok arbeidet i prosjektet i høst, sto det klart at dersom det skulle være noen mulighet på kort sikt for å nytteberegne økninger i vedlikeholdsbudsjettene, måtte vi gi en enkel oppskrift som gjorde det mulig for etatene å foreta beregningene sjøl. Nærmere bestemt måtte vi gi en oppskrift som samfunnsøkonomene i etatene kunne arbeide videre med og sjøl få vedlikeholdsekspertene i etatene til å levere input til. En slik oppskrift gir vi i kapittel 2. Den er ikke basert på detaljerte opplysninger om de enkelte objektene som skal vedlikeholdes, og pretenderer ikke å resultere i noen realistisk vedlikeholdsplan. Det som trengs til denne oppskrifta, er snarere gjennomsnittlige sammenhenger for store klasser av objekter. Videre finnes det ikke noe bindende budsjett for det enkelte året, men bare en gjennomsnittlig årlig budsjetttramme som skal overholdes i det lange løp. Det er en underliggende forutsetning at vedlikeholdsbudsjettene i utgangspunkt er tilstrekkelige til at et stabilt nivå på kvaliteten av infrastrukturen er mulig i det lange løp. Spørsmålet er da om det lønner seg å forbedre dette nivået ved å øke gjennomsnittsbudsjettet.

Modellen, som er en enkel videreføring av modellen i avsnitt 2.4 i Minken, Dahl og Steinsland (2008), kan programmeres i EXCEL og gir umiddelbart en nyttekostnadsbrøk for tiltaket som skal analyseres, nemlig å øke gjennomsnittsbudsjettet med ei krone. I tillegg gir den optimal rehabiliteringsfrekvens for hver av klassene av objekter. Arbeidet med de nødvendige kostnadene og funksjonssammenhengene kan bygge på systematikken som er bygget opp i vårt prosjekt og redegjort for i Minken, Meyer, Veisten og Bai (2011).

¹Et brukergrensesnitt som blei utviklet i 2011 har vi forkastet fordi for mye av begrepene og tankegangen var preget av vegdekkeanvendelsen. Medarbeideren hadde da sluttet. Da vi ikke fant noen andre som kunne ta over arbeidet, avbrøt vi prosjektet i september 2011. Da vi, i samråd med oppdragsgiver, tok opp igjen arbeidet i 2014, blei den nye medarbeideren langvarig sjuk, og arbeidet har derfor ikke kunnet fullføres til avtalt tid.

Kapittel 3

Den enkle modellen i kapittel 2 kan ikke erstatte en modell som tar begrensningene i de årlige budsjettene mer på alvor, som tar hensyn til at det finnes mange typer av tiltak, og som tar utgangspunkt i at objektene i utgangspunktet kan befinne seg i ulike tilstander. Bare en slik modell kan si noe om hva som er den beste strategien i et virkelig tilfelle, og bare i en slik modell er det mulig å studere slike ting som hvordan man på beste måte tar inn et vedlikeholdsetterslep.

I prosjektet har vi etablert en slik modell, men ikke fått utstyrt den med et godt brukergrensesnitt. Det er derfor ønskelig å fortsette arbeidet med denne modellen i en eller annen form. Å få den enkle modellen fra kapittel 2 på beina kan være nøkkelen til dette, for det vil kreve at etatene spesifiserer kostnader og funksjonssammenhenger som direkte kan tas i bruk i den mer avanserte modellen. Dermed løser vi ett av problemene som vi ikke løste i prosjektet vi rapporterer her, nemlig å få etatene til å engasjere seg mer direkte i problemstillingen å finne de samfunnsøkonomisk beste vedlikeholdsstrategiene og riktig nivå på vedlikeholdsbudsjettene.

I kapittel 3 gir vi derfor en forhåpentligvis forståelig framstilling av den mer avanserte modellen, i håp om å øke interessen for å utvikle den videre til et praktisk redskap og for å etablere videre forskning på dette området.

Problemstillingen presenteres først på en mest mulig generell måte, slik at man ser muligheter for anvendelser på ulike felt. Videre forklares den matematiske modellen som beskriver planleggingsproblemet. Modellen er en optimeringsmodell der optimeringen består i å minimere den samfunnsmessige total kostnaden over en viss tidshorisont. Denne minimeringen består i å finne en best mulig vedlikeholdsstrategi ved bruke avanserte algoritmer fra fagfeltet matematisk optimering. Presentasjonen er likevel skrevet slik at man ikke trenger bakgrunn i dette feltet for å komme inn i modell og metoder, i alle fall på et overordnet nivå. Et svært viktig trekk ved den matematiske metoden som benyttes er at den gjør det mulig å bedømme hvor god den løsningen man finner virkelig er. Dette skjer ved at det beregnes en nedre skranke på den optimale vedlikeholdskostnaden ved å utvide løsningsmengden i problemet. Ideen bak denne tilnærmingen blir forklart i nevnte kapittel uten å gå inn i detaljene i dette.

Kapittel 4

Et infrastrukturobjekt har ulike egenskaper som er av betydning for hvor godt det tilfredsstill de kravene som brukerne vil sette til det. Prinsipielt er det derfor naturligvis viktig å kunne operere med flere tilstandsvariable.

Imidlertid er det - i alle fall på vegsida - utviklet flere vanlig brukte sammensatte tilstandsindikatorer. Nedbrytningsfunksjonen og tiltakenes virkning vil da gjelde den sammensatte tilstandsvariabelen. De øvrige mulige tilstandsvariablene, som styrken til konstruksjonen og underbygningen, behøver likevel ikke bli helt utelatt, men kan fungere som variabler som påvirker parameterne i nedbrytningsfunksjonen og funksjonene som angir hvordan de ulike tiltakene virker. Det er fullt mulig å gjennomføre følsomhetsanalyser som kan illustrere betydningen av utviklingen i tilstandsvariablene som bare indirekte påvirker brukernes opplevelser og kostnader. Det er derfor ikke så urealistisk som en skulle tro å eksperimentere videre med modellen slik den blei utviklet i det første prosjektet. I kapittel 4 gjør vi nettopp det.

Vi tar utgangspunkt i å anvende programmet på et testeksempel i flere omganger for å illustrere generelt, uavhengig av testeksempelet vi ser på, hvordan programmet kan brukes som det foreligger. Via simuleringer vil vi forsøke å gi helt konkrete svar på spørsmål som naturlig stilles i forbindelse med allokering av budsjettmidler.

Kapittel 5

I prosjektet har vi gått gjennom en mengde vitenskapelig litteratur (pluss noe litt mindre vitenskapelig litteratur) om vedlikehold og rehabilitering. Spesielt har vi vært interessert i litteratur

som har et økonomisk perspektiv og som behandler tilsvarende problemer som vi har vært opptatt av. Noe av dette er gjennomgått i Minken, Meyer, Veisten og Bai (2011). Men vi tror det vil være nyttig for etatene om vi supplerer det med henvisninger til flere interessante artikler og rapporter. Dette er gjort i kapittel 5.

2. Skissemessig vurdering av nytten av å øke vedlikeholdsbudsjettet

Infrastruktur, som veger, jernbanelinjer, tunneler og flystriper, er ikke gjenstand for gjentatt hel utskifting og gjenanskaffelse. De dypere lag av slike konstruksjoner er i liten grad utsatt for nedbryting ved ordinær bruk, og trenger derfor ikke å skiftes ut. Hovedsakelig er det det øverste laget - vegdekket, skinnegangen - som brytes ned med tida og trenger å fornyes. Et nedslitt vegdekke eller en nedslitt skinnegang påfører brukerne ekstrakostnader i form av større slitasje på kjøretøy, lavere fart og større ulykkesrisiko. Dårlig ordinært vedlikehold og utsatt rehabilitering kan dessuten medføre at større deler av konstruksjonen bli skadet, hvilket igjen vil føre til enda mye større brukerkostnader.

Men rehabilitering er dyrt. For hyppig rehabilitering gir derfor unødige kostnader for samfunnet, mens for sjelden rehabilitering gir økte årlige brukerkostnader. Dessuten vil etaten som har ansvaret for infrastrukturen også påføres økte årlige kostnader, for eksempel i form av stadig hyppigere ikke-planlagt vedlikehold, reparasjoner og lapping. Derfor finnes en samfunnsøkonomisk optimal rehabiliteringstakt. Den skal vi nå finne i en enkel modell der det ikke er noen trafikkvekst. Deretter utvider vi perspektivet til mange objekter som skal rehabiliteres under et felles budsjett, som kan overskrides i det enkelte året, men som antas å være bindende på langt sikt.

2.1 Optimal rehabiliteringstakt for et enkelt objekt

Vi ser nå på en enkelt bit av infrastrukturen - det kan være en vegstrekning, ei bru eller et hvilket som helst annet slags infrastruktureobjekt. Da objektet blei bygd, påløp det en investeringskostnad C_0 . Seinere vil det trenges rehabilitering med en fast frekvens hvert n -te år. Første rehabilitering skjer altså n år etter bygging. Rehabiliteringskostnaden er en tiltakende funksjon av denne frekvensen, $R(n)$. Det er rimelig å anta $R(0) = 0$ og den deriverte $R'(n) > 0$. Om rehabiliteringen ikke skjer altfor sjelden, vil vi anta at objektet har uendelig levetid.¹

Brukerne vil ha en fast årlig nytte E av å bruke objektet på N . Men dessuten vil de også ha kostnader, og disse kostnadene avhenger av tida siden forrige rehabilitering, t . Vi skriver dem $C_B(t)$, der t går fra 0 til n og $C'_B(t) > 0$. Likeledes vil infrastruktureieren (etaten) ha årlige kostnader $C_E(t)$ på samme form. Så lenge vi ikke regner med at det finnes et budsjett, er det ikke essensielt å skille mellom brukerkostnader og etatskostnader, men vi vil få bruk for det seinere.²

Tabell 2.1 sammenfatter notasjonen vi skal bruke i dette kapitlet: Vi bruker kalkulasjonsrenta ρ til å neddiskontere nytten og kostnadene i en rehabiliteringsperiode til begynnelsen av perioden.

¹Levetida er den korteste av følgende to tidsrom: For det første den tida hvor objektet tjener sin opprinnelige hensikt. Den er over når det samfunnsmessige behovet for objektet opphører eller når folk får bedre måter å tilfredsstille dette behovet på. For det andre den tekniske levetida - tida inntil hele konstruksjonen er utslitt og må erstattes med en ny. Den økonomiske levetida er i vår definisjon den korteste av tida hvor et finns et behov for objektet og den tekniske levetida ved optimal rehabiliteringstakt.

²Dersom infrastruktureieren har driftskostnader som ikke er inkludert i det budsjettet vi seinere skal se på, og disse kostnadene øker med tida siden forrige rehabilitering, kan de åpenbart inkluderes i $C_B(t)$. For å markere det, kunne vi bruke den noe vagere betegnelsen brukskostnader.

Tabell 2.1: Notasjon

Investeringskostnad	C_0
Tidsavhengige årlige brukskostnader	$C_B(t)$
Tidsavhengige årlige etatskostnader	$C_E(t)$
Brukernytte	N
Antall år mellom hver rehabilitering	n
Rehabiliteringskostnad	$R(n)$
Kalkulasjonsrente	ρ

Samfunnets netto nytte av objektet over en hel rehabiliteringsperiode vil vi kalle $V(n)$. Vi har da:

$$V(n) = \int_0^n \left(N - (C_B(t) + C_E(t)) \right) e^{-\rho t} dt - R(n)e^{-\rho n} \quad (2.1)$$

Netto nytte $W(n)$ av et slikt infrastrukturobjekt med en uendelig levetid er investeringskostnaden pluss netto nytte av en uendelig kjede av gjentatte rehabiliteringer. Nyttten av den uendelige kjeden utgjør en geometrisk rekke med kvotient k , der k er diskonteringsfaktoren som omgjør $V(n)$ i hver av rehabiliteringsperiodene til nåverdien ved begynnelsen av perioden før. I en matematisk formelsamling vil en finne at summen av en uendelig geometrisk rekke er:

$$a + ak + ak^2 + \dots = \frac{a}{1 - k}$$

der a er et vilkårlig tall og kvotienten k må være mindre enn $|1|$ for at summen skal være endelig. I vårt tilfelle er a lik $V(n)$ og $k = e^{-\rho n}$. Derfor får vi at $W(n)$ blir:

$$\begin{aligned} W(n) &= -C_0 + \sum_{i=0}^{\infty} e^{-i\rho n} \int_0^n \left(N - (C_B(t) + C_E(t)) \right) e^{-\rho t} dt - \sum_{i=1}^{\infty} e^{-i\rho n} R(n) \\ &= -C_0 + (1 - e^{-\rho n})^{-1} \left(\int_0^n \left(N - (C_B(t) + C_E(t)) \right) e^{-\rho t} dt - R(n)e^{-\rho n} \right) \\ &= \left(\frac{N}{\rho} - C_0 \right) - (1 - e^{-\rho n})^{-1} \left(\int_0^n \left((C_B(t) + C_E(t)) \right) e^{-\rho t} dt + R(n)e^{-\rho n} \right) \end{aligned} \quad (2.2)$$

Legg merke til at den andre summen i første linje går fra $i = 1$, ikke $i = 0$. Det skyldes at rehabiliteringskostnaden først skal neddiskonteres til begynnelsen av perioden, før den neddiskonteres sammen med de årlige kostnadene til år null.

Den verdi av n som maksimerer $W(n)$ kaller vi den optimale rehabiliteringstakten, n^* . Vi finner n^* ved å derivere $W(n)$ og sette den deriverte til 0. Det gir etter litt regning følgende likning for å beregne n^* :

$$\frac{\rho}{1 - e^{-\rho n}} \left(\int_0^n (C_B(t) + C_E(t)) e^{-\rho n} dt + R(n) \right) = C_B(n) + C_E(n) + R'(n) \quad (2.3)$$

Her merker vi oss først at investeringskostnaden og den konstante delen av brukernytten ikke har noen innvirkning på optimal rehabilitering. Avskrivninger eller andre regler som tar utgangspunkt i investeringskostnaden, slik som å "bevare kapitalverdien", gir altså ikke noe fornuftig grunnlag for å bestemme vedlikeholds- og rehabiliteringspolitikken. Tolkningen av vilkåret (2.3) for optimal rehabilitering er for øvrig som følger:

Den første handlingen som spiller noen rolle, er den første rehabiliteringen. Det er også den siste handlingen som spiller noen rolle, for vi har forutsatt at alle rehabiliteringsperiodene er like lange. Venstresida i (2.3) er annuiteten³ av alle kostnader fra og med tidspunktet for den første

³Annuiteten til en utbetaling er et fast årlig beløp som er slik at utbetaling av hele beløpet nå er lik nåverdien av annuiteten utbetalt hvert år over en viss periode (her: en uendelig periode).

rehabiliteringen dersom vi velger å gjennomføre den etter n år, mens høyresida er kostnaden ved å utsette rehabiliteringen med ett år. Så lenge venstresida er større enn høyresida, er det dyrere i gjennomsnitt å rehabilitere etter n år enn å utsette det, og dersom venstresida er mindre enn høyresida, er det omvendt, og vi er egentlig allerede for seint ute.

2.2 Mange objekter og et felles budsjett

La oss nå anta at det finns J objekter, indeksert med j . Det betyr at objektene er nummerert fra 1 til J og at j kan betegne et hvilket som helst av objektene $1, 2, 3, \dots, J$. Alle objektene er i utgangspunkt brakt opp til den best mulige tilstand i år 0 ved hjelp av initielle investeringer $\mathbf{C}_0 = (C_{01}, \dots, C_{0j}, \dots, C_{0J})$. Hvert objekt har sin egen brukernytte og sin egen brukerkostnad og etatskostnad som funksjon av tida siden siste rehabilitering. Videre er også rehabiliteringskostnaden forskjellig for hvert av objektene. Som følge av alle kostnadsforskjellene kan de ha ulike rehabiliteringsperioder, sjøl om de startet første periode samtidig.

Netto nytte for objektene samlet er en funksjon $W(\mathbf{n})$ av vektoren $\mathbf{n} = (n_1, n_2, \dots, n_J)$ av rehabiliteringsperiodene. Ved en enkel utvidelse av likning (2.2) til tilfellet med mange objekter, har vi:

$$W(\mathbf{n}) = \sum_{j=1}^J \left[\left(\frac{N_j}{\rho} - C_{0j} \right) - \left(1 - e^{-\rho n_j} \right)^{-1} \cdot \left(\int_0^{n_j} (C_{Bj}(t) + C_{Ej}(t)) e^{-\rho t} dt + R_j(n_j) e^{-\rho n_j} \right) \right] \quad (2.4)$$

Vi har en felles budsjettbetingelse for alle objektene. Den gjelder budsjettet for hele den uendelige levetida samlet, eller med andre ord nåverdien av den gjennomsnittlige årlige budsjettutgiften. Kall dette gjennomsnittet B . Budsjettbetingelsen kan da skrives:

$$\sum_{j=1}^J \frac{1}{1 - e^{-\rho n_j}} \left(\int_0^{n_j} C_{Ej}(t) e^{-\rho t} dt + R_j(n_j) e^{-\rho n_j} \right) \leq \frac{B}{\rho} \quad (2.5)$$

La oss kalle skyggeprisen på budsjettrestriksjonen μ . Den vil være positiv. I tillegg til budsjettbetingelsen 2.5 vil det være L førsteordensbetingelser for optimum, en for hvert objekt. Det er i prinsipp nok til å regne ut alle de J optimale rehabiliteringsperiodene pluss skyggeprisen μ . For å gjøre det, danner vi Lagrangefunksjonen og setter de partiellderiverte av den til 0. For $j = 1, \dots, J$ gir det følgende førsteordensbetingelser:

$$G_j(n_j) - H_j(n_j) = \frac{\mu}{1 + \mu} \left[\frac{\rho}{1 - e^{-\rho n_j}} \int_0^{n_j} C_{Bj}(t) e^{-\rho t} dt - C_{Bj}(n_j) \right] \quad (2.6)$$

der

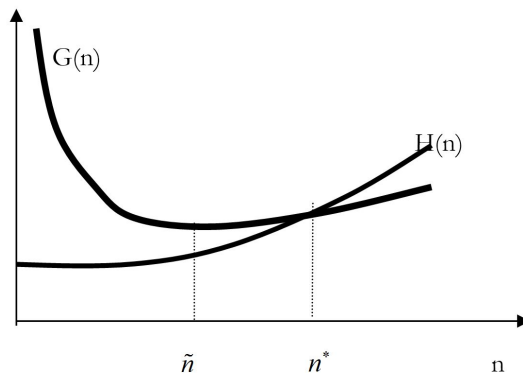
$$G_j(n_j) = \frac{\rho}{1 - e^{-\rho n_j}} \left[\int_0^{n_j} (C_{Bj}(t) + C_{Ej}(t)) e^{-\rho t} dt + R_j(n_j) \right]$$

$$H_j(n_j) = C_{Bj}(n_j) + C_{Ej}(n_j) + R'_j(n_j)$$

Det trekket ved (2.6) som en først kan merke seg, er at hver av førsteordensbetingelsene involverer én og bare én rehabiliteringsperiode n_j . Det betyr at de optimale periodene kan regnes ut én for én, uten å blande inn andre variable enn skyggeprisen. Det neste en kan merke seg, er at G_j har akkurat samme form som venstresida i (2.3), mens H_j har samme form som høyresida. Tolkningen er altså den samme, og det eneste budsjettbetingelsen gjør, er å stille krav om at hver G_j skal være *mindre* enn tilsvarende H_j , mens vi tidligere, når det ikke var noe budsjett å ta hensyn til, skulle ha at de var like.⁴ Figur 2.2, som er kopiert fra det tilsvarende diagrammet på side 8 i (Minken, Dahl, Steinsland, 2008), viser de to kurvene.⁵ Løsningen av denne modellen kan programmeres i

⁴Klammeparentesen i (2.6) vil være negativ, for brukerkostnadene i siste år før rehabilitering vil alltid være større enn annuiteten av brukerkostnadene i alle år.

⁵Formelen i (Minken, Dahl, Steinsland, 2008) som tilsvarende vår formel 2.6, lider dessverre av flere skrivefeil eller regnefeil. Ser vi bort fra det, vil det kanskje være nyttig å lese kapittel 2 der som en utdyping av vår tekst her.



Figur 2.1: TØI-rapport 957/2008

EXCEL. For hvert objekt j må en naturligvis først etablere formene på funksjonene som inngår, og det er her mesteparten av arbeidet ligger. Når det er gjort, kan man først velge en verdi for $\mu/(1 + \mu)$. Den må naturligvis være mellom 0 og 1, uten å komme helt opp til den øvre grensa. Da kan hver av førsteordensbetingelsene løses ved å bruke funksjonen "målsøking" i EXCEL, som fungerer slik at om man har n_j i en egen celle, kan man be EXCEL om å tilpasse verdien i den cella slik at et bestemt vilkår (her: førsteordensbetingelsen 2.6) blir oppfylt.

Men så må man tilpasse $\mu/(1 + \mu)$ slik at budsjettbetingelsen (2.5) blir oppfylt. Man må altså kontrollere om (2.5) er oppfylt eller ikke. Er den oppfylt med likhet, er løsningen funnet. Hvis den er oppfylt med ulikhet, må $\mu/(1 + \mu)$ senkes, og hvis betingelsen ikke er oppfylt, må $\mu/(1 + \mu)$ økes. Dette kan gjøres ved en systematisk framgangsmåte som heter det gyldne snitt (se golden section search på Wikipedia). Man må altså gå fram og tilbake mellom å løse førsteordensbetingelsene for gitt og søke etter den som gir likhet i (2.5).

Det er ikke sikkert at en løsning finnes. Det er mulig budsjettet er for stramt til å drive den samme rehabiliteringspolitikken i det uendelige uten at vedlikeholdskostnadene eller brukerkostnadene eksploderer. Det er også mulig at budsjettet er så romslig at budsjettet ikke blir oppbrukt i optimum, slik at $\mu = 0$. (Det er kanskje det man først bør anta om $\mu/(1 + \mu)$.) Men i alle andre tilfeller vil budsjettbetingelsen tvinge fram en langsommere rehabiliteringstakt enn det som ellers ville være optimalt.

2.3 Nyten av å øke budsjettet

Vi ønsker å finne verdien av å øke det gjennomsnittlige årlige budsjettet med en krone. En setning i matematikken som heter oppfyllingssetningen, kan brukes her. Den sier her at økningen i samfunnets netto nytte ved optimal rehabiliteringspolitikk, $W(\mathbf{n}^*)$, som følge av en marginal forbedring av det langsiktige budsjettet B/ρ , er lik μ^* , altså den optimale skyggeprisen. Det vi ønsker å finne, er imidlertid ikke virkningen av å endre B/ρ , men av å endre B . Regnestykket blir:

$$\frac{\partial W^*}{\partial B} = \frac{\partial W^*}{\partial (B/\rho)} \cdot \frac{\partial (B/\rho)}{\partial B} = \frac{\partial W^*}{\partial (B/\rho)} \cdot \frac{1}{\rho} = \mu^* \frac{1}{\rho} \quad (2.7)$$

Siden slutten av 1990-tallet har vi i Norge brukt netto nytte per budsjettkrone (NNB) som lønnsomhetsmål. Prosjektet å øke det gjennomsnittlige budsjettnivået i alle framtidssår med én krone, koster det offentlige $1/\rho$ kroner og gir en netto nytte på $\mu^* \cdot \rho^{-1} - \rho^{-1}$. Netto nytte per budsjettkrone i dette tilfellet blir

$$NNB = \frac{\mu^* \frac{1}{\rho} - \frac{1}{\rho}}{\frac{1}{\rho}} = \mu^* - 1 \quad (2.8)$$

Det vil altså lønne seg samfunnsøkonomisk å gjennomføre en varig gjennomsnittlig budsjettøkning på én krone dersom $\mu^* > 1$, eller med andre ord dersom $\mu/(1 + \mu) > 1/2$.

3. Optimering: modell og metode

Dette kapitlet omhandler optimeringsmodellen i prosjektet. Vi redegjør for begreper og selve problemstillingen, samt gir en innføring i metodene som benyttes i programmet. I noen grad peker vi også på ideer for videre arbeid med modell og metoder.

3.1 Kort beskrivelse

MOV er et dataprogram som kan brukes til strategisk planlegging av vedlikehold i en gitt infrastruktur, samt tilhørende analyse av betydningen av rammevilkår som f.eks. budsjetter. Programmet er av svært generell natur og vil dermed kunne anvendes på slike problemstillinger innen flere sektorer, som f.eks. vei eller jernbane.

Programsystemet er basert på avanserte matematiske metoder fra fagfeltet matematisk optimering. Vi kombinerer ulike idéer og teknikker fra områder som diskret optimering og dynamisk optimering (optimal kontroll). Det er en underliggende matematisk modell, altså en matematisk beskrivelse av planleggingsproblemet. Vi vil her redegjøre for denne modellen samt algoritmene som benyttes. Vi vil dermed presentere potensialet for bruk av systemet, men også få fram antagelser og begrensninger som er tilstede.

Modellen omhandler en infrastruktur, et system av objekter (f.eks. deler i et veinettverk), der hvert objekt er utsatt for en slitasjeprosess som gradvis medfører vedlikeholdsbehov. Vedlikehold kan utføres på ulike objekter, til ulike tider, og av ulikt omfang. Man opererer med en viss tidsperiode som planleggingshorisont. Starttilstander for objektene må spesifiseres. Videre antar vi at man beskriver ønsket tilstand i systemet ved avslutning av planperioden. Begrepet "tilstand" her kan svare til noe ganske komplisert, alt avhengig av anvendelsen. I modellen er dette håndtert på svært generelt vis, noe som nettopp gir den fleksibiliteten man ønsker for et slikt verktøy.

Hvert element i systemet er utsatt for en nedbrytning (slitasje) som må beskrives matematisk. For eksempel vil en veistrekning slites ut fra forventet årlig trafikk. Videre kan vedlikehold utføres på objektene, og mulighetene her må spesifiseres. Dette betyr at ved hvert tidsskritt i modellen (f.eks. hvert år) og for hvert objekt i systemet, står man overfor et valg mellom ulike vedlikeholdstiltak. Man kan f.eks. utføre et omfattende vedlikehold, et noe enklere vedlikehold eller unnlate vedlikehold fullstendig. Merk at vedlikeholdstiltakene for de enkelte objektene er uavhengige av hverandre. Det kan vurderes om koplinger mellom tiltakene kan innføres i en framtidig utgave av MOV, men dette er en komplisert utvidelse å håndtere algoritmisk.

Det er modellen og optimeringen som skal bestemme hvilke valg av vedlikeholdstiltak som er best ut fra helheten, altså ved å se alle objektene, budsjettet, og hele planperioden under ett. Dette er en svært komplisert problemstilling generelt, og derfor må matematiske metoder og effektive beregninger på datamaskiner benyttes for å finne gode løsninger. Det kan i noen enkle tilfelle være mulig å finne optimale, eller ganske gode, vedlikeholdsstrategier ved manuelle metoder. Men straks det er variasjon mellom objekter, i slitasje og mulig vedlikehold, og antall objekter og tidsskritt øker, så er den matematiske modellen og algoritmene der eneste mulighet for å finne gode strategier.

Programsystemet MOV er skrevet i språket MATLAB, som er et system for tekniske beregninger og matematikk. Selve algoritmene er utviklet basert på generelle prinsipper, men disse er skreddersydd

for den modellen vi har her og med tanke på fleksibilitet i anvendelsesområder.

Resten av kapitlet gir en innføring i den matematiske modellen og metodene som benyttes. Intenjonen er å presentere dette slik at man kan lese dette uten nevneverdig kjennskap til fagområdet matematisk optimering. En kort omtale av sentrale programmoduler blir også gitt. Vi vil også diskutere muligheter og begrensinger ved modellen, samt kort omtale aktuelle utvidelser.

En langt mer utførlig beskrivelse av modell og metoder, samt en anvendelse på veinnettverk, finnes i artikkelen av Dahl og Minken (2008) fra tidsskriftet *Computers & Operations Research*. Videre, i selve MATLAB koden, er det skrevet en viss dokumentasjon av datastrukturer og hovedskrittene i de respektive rutinene.

3.2 Formål med MOV

Hovedformålet med MOV er å kunne gjøre god strategisk planlegging av vedlikehold i infrastruktur-systemer.

Programmet håndterer en svært generell problemstilling. Fordelen med generaliteten er at man har flere mulige anvendelsesfelt, f.eks.

- infrastruktur av veier i et avgrenset geografisk område
- jernbane: vogner, skinnegang eller annet utstyr
- andre områder for transport (f.eks. gods)

MOV er implementert i programsystemet MATLAB, som er avansert matematisk programvare for tekniske beregninger.

MOV består av programmoduler for å spesifisere en konkret problemstilling i henhold til en matematisk modell, samt algoritmer for beregning av optimal vedlikeholdsstrategi. Resultatene rapporteres og presenteres til bruker. Programmet gir typisk løsninger som er nær optimale ut fra den spesifiserte problemstillingen. Naturligvis vil det alltid være usikkerhet knyttet til data som man putter inn i modellen. Men for gitte data ønsker man en løsning som helst er optimal eller nesten-optimal. Det er dette MOV er egnet for.

3.3 Hvorfor er dette vanskelig

Man kan kanskje tro at vedlikeholdsplanlegging ikke er så vanskelig: man setter opp en vedlikeholdssyklus for hvert objekt og samordner disse litt slik at budsjettet holder.

For noen problemer kan det være nokså greit, men generelt er det langt mer komplisert.

En problemstilling som dukker opp i MOV kan best forklares via et grunnleggende kombinatorisk problem som vi nå diskuterer; det kalles gjerne *ryggsekkproblemet*. Dette problemet dukker opp som et spesialtilfelle av MOV, dersom input data er satt på passende vis. Faktisk kan vi si at MOV problemet er en generalisering av ryggsekkproblemet der dynamikk er introdusert.

La n være et positivt heltall, og anta at a_1, a_2, \dots, a_n og c_1, c_2, \dots, c_n er gitte positive tall. Videre la b være et gitt positivt tall. Vi sier at en delmengde S av $\{1, 2, \dots, n\}$ er *tillatt* dersom $\sum_{i \in S} a_i \leq b$. Her angir S hvilke tall vi velger ut: vi tar med a_i for hver $i \in S$. Dette utvalget er tillatt dersom summen av tallene ikke overskrider det gitte tallet b . Vi har nå følgende problem:

- **Ryggsekkproblemet (RS):** Finn en tillatt delmengde S som maksimerer $\sum_{i \in S} c_i$ over alle tillatte delmengder.

Eksempel. La $n = 4$, $a_1 = 2$, $a_2 = 4$, $a_3 = 4$, $a_4 = 6$. Videre, la $b = 7$. Da er f.eks. $S = \{1, 2\}$ tillatt fordi $a_1 + a_2 = 6 \leq 7$, mens $S = \{3, 4\}$ ikke er tillatt fordi $a_3 + a_4 = 10 > 7$. I dette tilfelle, siden $n = 4$ er så liten, kan vi bestemme alle tillatte mengder, som er

$$\emptyset, \{1\}, \{2\}, \{3\}, \{4\}, \{1, 2\}, \{1, 3\}.$$

Hvis $c = (c_1, c_2, c_3, c_4) = (1, 3, 4, 3)$ (altså $c_1 = 1$, $c_2 = 3$ etc.), så får vi følgende liste med tillatte mengder S og tilhørende verdier på $\sum_{i \in S} c_i$:

S	\emptyset	$\{1\}$	$\{2\}$	$\{3\}$	$\{4\}$	$\{1, 2\}$	$\{1, 3\}$
$\sum_{i \in S} c_i$	0	1	3	4	3	4	5

Så vi ser at $S = \{1, 3\}$ er (entydig) optimal løsning.

Problem RS kalles ryggsekkproblemet fordi man kan tolke det som å finne en optimal pakking av en ryggsekk: man velger blant n gjenstander, der hver har en viss vekt og en viss nytteverdi. Videre vil man bære høyest vekten b , og RS problemet er da å velge ut de gjenstandene som maksimerer totalnytteten uten å overskride vektbegrensingen b .

Som sagt, hvis n er liten, kan man greit løse RS ved metoden over, en opplisting av alle muligheter. Men for stor n er situasjonen annerledes. F.eks. la $a, c \in \mathbb{R}^{34}$ være gitt ved

$$a = (3, 1, 11, 6, 6, 2, 12, 3, 6, 7, 11, 12, 4, 1, 4, 21, 6, 7, 8, 16, 5, 2, 5, 6, 3, 5, 2, 2, 8, 5, 1, 2, 1, 3),$$

$$c = (8, 4, 1, 8, 8, 4, 11, 4, 7, 5, 3, 15, 1, 3, 1, 17, 5, 17, 2, 11, 2, 5, 5, 1, 7, 9, 10, 13, 5, 3, 2, 6, 2, 2),$$

og la $b = 24$. En naturlig første metode er da å gå gjennom alle delmengde S og for hver sjekke om den er tillatt, dvs. om $\sum_{i \in S} a_i \leq b$, og dessuten beregne $\sum_{i \in S} c_i$ slik at kan sammenlikne og finne en optimal løsning. Men antall delmengder her er

$$2^{34} = 17.179.869.184$$

så det er altså over 17 milliarder løsninger å sjekke (og for hver skal to summer beregnes). Problemet er at antall delmengder av en mengde med n elementer er 2^n , som vokser eksponentielt i n . Hvis dataene (tallene) har en kjent struktur, eller hvis c har en "pen" struktur, kan det tenkes at man kan unngå å se på alle delmengder. Også i det generelle tilfelle kan man gjøre mer fornuftige søk enn metoden vi nettopp omtalte.

Likevel, RS er et kjent vanskelig beregningsproblem, det er et såkalt *NP-hardt* optimeringsproblem. Dette er et sentralt begrep i fagfeltet algoritmer og kompleksitet, og innebærer i praksis at ingen effektiv algoritme er kjent, til tross for iherdige forsøk på å finne dette. De fleste tror heller ikke det er mulig å finne en effektiv algoritme, selv om dette ennå ikke er bevist matematisk.

Svært forenklet kan man si at en algoritme er en *polynom-algoritme* dersom antall regneoperasjoner vokser som et polynom når man ser det som funksjon av størrelsen på input. I RS er input vektorene a og c , begge av lengde n , samt tallet b . Størrelsen på input er antall binære bits som kreves for å lagre input tallene.

For et *NP-hardt* problem er ingen polynomalgoritme funnet, og hvis man skulle finne en slik,, vil alle andre problemer i klassen *NP*, en stor klasse av "vanskelige" problemer, også løses i polynom regnetid. Fra en mer praktisk synsvinkel regnes RS blant de "enkleste" *NP*-harde problemene, slik at hvis b ikke er for stor, vil man gjerne kunne løse det ved hjelp av dynamisk programmering.

For å vende tilbake til MOV modellen kan vi litt forenklet si at vi i hver tidsperiode har et RS liknende problem: velg ut de prosjekter som skal realiseres slik at summen av vedlikeholdskostnadene (a_i -ene) ikke overskrider budsjettet b for dette tidsskrittet. Her kan c_i oppfattes som den

samfunnsmessige nytten av å realisere prosjekt *i*. MOV er likevel langt mer komplisert fordi det er av dynamisk karakter: man har slike RS problemer i hvert tidsskritt, kanskje 30 perioder, men alle disse RS-problemene er linket sammen ved at man kan skyve vedlikehold ut i tid fordi det kan være mer lønnsomt.

Det er denne avhengigheten av beslutninger om vedlikehold over tid som skaper den store utfordringen. Det er lagt ned mye arbeid i å finne skreddersydde matematiske metoder for å løse disse optimeringsproblemene.

3.4 Modellen

Vi presenterer nå modellen, og bruker terminologi som skal være nokså uavhengig av konkrete anvendelser.

- **System.** Vi betrakter et *system* med et antall *objekter*. Disse objektene kan være, f.eks., tog eller veisegmenter. Hvert objekt er utsatt for en form for nedbrytning eller slitasje, og det er dette som gir behov for vedlikehold. Nedbrytning må spesifiseres i input til MOV. Det er ulike måter å gjøre dette på.
- **Beslutningene.** Bruker av MOV ønsker å finne en sekvens av best mulige beslutninger om vedlikehold for hvert objekt over hele tidperioden. Man kan tenke seg ulike former for vedlikehold for hvert objekt, slik som et lettere vedlikehold eller et mer omfattende vedlikehold. Som regel er det også ønskelig å ha som alternativ at intet vedlikehold foretas. Det kan også være snakk om ulike typer vedlikehold. For hvert objekt må man derfor spesifisere en mengde med ulike *mulige* vedlikeholdstiltak. I tillegg må man gi tilknyttede kostnader, og vi opererer med to typer kostnader, som nå beskrives.
- **Kostnader.** Det er to kostnadstyper:
 - Type 1: Normale kostnader (brukerkostnad)
 - Type 2: Vedlikeholdskostnad

Med normale kostnader mener vi kostnader som er der til daglig og er knyttet til bruken av objektet. Denne kostnaden kan avhenge av “kvaliteten” på objektet (se under) og av hvor mye objektet er i bruk, f.eks. trafikkvolum for en veistrekning. Kostnaden må også spesifiseres av bruker. Kostnader av Type 2, vedlikeholdskostnader, er kostnader som er valgfrie og som det tas beslutninger om å iverksette. Bruker spesifiserer for hvert objekt en liste over mulige vedlikeholdstiltak, og tilhørende kostnader. Et tiltak som består i å ikke gjøre noe vedlikehold, med kostnad 0, er da mulig (og vanlig) å spesifisere. Hvis man holder kostnader av Type 2 lave (lite vedlikehold) vil kostnadene av Type 1 gjerne bli større (med realistiske data), og det å finne en best mulig balanse her er det optimeringen i MOV skal finne ut av.

- **Tilstander og dynamikk.** MOV er en dynamisk modell der en viss tids-horisont splittes i et antall intervaller. Antall tidsskritt velges av bruker, og kan f.eks. være slik at ett tidsintervall svarer til ett år. I hvert tidsskritt utføres vedlikehold og objektene er utsatt for nedbrytning. Angående dynamikken, altså utviklingen over tid, antar vi at status for hvert objekt kan sammenfattes ved en *tilstand*. Vi opererer altså med en mengde av mulige tilstander, og denne mengden kan variere fra objekt til objekt. Dette gir et *tilstandsrom* for hvert objekt. Dynamikken er slik at nedbrytningen av et objekt i tidsintervallet gir en (muligens) ny tilstand (som representerer lavere “kvalitet”). Videre vil vedlikehold endre tilstanden, og denne endringen bestemmes av valg av vedlikeholdstiltak. Vi tenker at vedlikeholdet skjer umiddelbart i starten av tidsintervallet, mens nedbrytningen gjelder for hele tidsintervallet. For hvert objekt må det angis en starttilstand som gjelder ved oppstart av planperioden. Videre angir man en slutttilstand som objektet skal være i ved slutten av planleggingsperioden.

- **Om tilstandsrommet.** Vi antar bare at tilstandsrommet er endelig, for hvert objekt, uten å bygge inn mer struktur eller antagelser her. Dette gir en enorm fleksibilitet, og gjør at MOV kan ha en rekke anvendelsesmuligheter. Det bør sies at regnetiden for optimeringen avhenger av størrelsen (antall elementer) på tilstandsrommet, mer om dette senere. I prinsippet kan kontinuerlige tilstandsrom benyttes i slike modeller, men da er det langt vanskeligere, nærmest umulig, å finne optimale løsninger, unntatt i visse tilfelle der nedbrytning og vedlikeholdsfunksjoner er av enkel og pen form. Et endelig tilstandsrom kan også fint fange opp at man måler tilstander via ulike parametre, slik som f.eks. er tilfelle for veistandard. F.eks. hvis det er to parametre, med henholdsvis 2 og 3 mulige verdier, så får vi totalt 6 kombinasjoner, så vårt tilstandsrom har størrelse 6. Det er klart at denne konstruksjonen kan gi gigantisk stort tilstandsrom, desom man har en del ulike slike parametre. En enkel erfaring med MOV, med dagens algoritmer og implementasjon, er at man kan håndtere opp i mot tilstandsrom av størrelse ca 1000 (for hver objekt). Regnetiden er avhengig av antall objekter og andre forhold, så dette er bare en enkel anbefaling.

Merk at dynamikken beskrives ved å spesifisere mulige overganger mellom tilstander. Jo mer komplisert tilstandsrommet er, jo mer data må angis og tilstandsovergangene blir jo også mange, og dette må også beskrives, selv om dette kan gjøres indirekte via oppskrifter som kan programmeres. Alt dette må altså bruker spesifisere slik at Matlab får oversatt dette i MOV til det konkrete optimeringsproblem som skal løses.

- **Budsjett.** Bruker må spesifisere budsjettet, brutt ned på de enkelte tidsintervallene. Det er et (absolutt) krav i modellen at budsjettet overholdes i hver periode. Dette kravet kopler vedlikeholdsproblemene for de ulike objektene sammen.
- **Optimeringsproblemet.** Optimeringsproblemet er å bestemme hvilket vedlikehold man skal benytte for hvert objekt og hver tidsperiode, og gjøre dette slik man minimerer en viss kostnadsfunksjon. Denne funksjonen summerer opp alle kostnader av Type 1 og Type 2 over alle objekter og tidsintervaller, basert på valgt vedlikeholdsstrategi.

3.5 Kommentarer til modellen

Problemstilling og modellen i MOV er beskrevet over, uten at vi gikk inn i alle detaljer. Modellen er svært generell, men har også sine begrensninger. Tilstandsrommet er diskret (endelig), ikke kontinuerlig. Dette innebærer at man ikke nøyaktig kan beskrive f.eks nedbrytning som løsning av en differensiallikning. Noen stor innvending er likevel ikke dette, forutsatt at man lar antall tilstander være stort nok; da vil forskjellen mellom en kontinuerlig og approksimativ diskret løsning være liten. Vi ser på et dynamisk system der tilstanden ved tid $t + 1$ avhenger av tilstanden ved tid t . Dette er en *første-ordens* differenslikning som forøvrig er ganske generell: dynamikken ved tid t kan avhenge av t slik at systemet kan være ikke-homogent og gjerne ikke-lineært.

Denne "første-ordens" egenskapen er viktig og blir utnyttet av algoritmene i MOV. Faktisk er det denne egenskapen som utnyttes slik at man får subproblemer som kan løses effektivt som *korteste vei problemer i en asyklisk graf* (spesialtilfelle av Bellman's algoritme). Denne algoritmen har faktisk lineær regnetid som funksjon av størrelsen på input.

Det kan forekomme anvendelser der nedbrytningen også avhenger av tiden siden siste vedlikehold. Dette passer ikke inn i nåværende MOV modell. Imidlertid, med en viss "pris" kan man håndtere denne svakheten ved å introdusere en ekstra parameter som angir tid siden siste vedlikehold. I så fall må tilstandsrommet utvides ut fra dette (slik vi har beskrevet før). Resultatet er en førsteordens modell som da passer inn i MOV oppsettet. Denne konstruksjonen vil raskt gi stort tilstandsrom, og er trolig bare gjennomførbart dersom man har få tilstander forøvrig.

Til slutt her kan vi kommentere spesialtilfellet der man bare har ét enkelt objekt. Da kan optimeringsproblemet løses ved hjelp av dynamisk programmering, evt. korteste vei i en asyklisk graf (se over). Dette er en enkel algoritme. Merk likevel at problemet ikke er trivielt i betydningen

at man alltid kan finne en eksplisitt løsning analytisk. Dette er fordi vi ønsker å håndtere helt generelle kostnadsdata. Det finnes analyser der man finner analytiske optimale vedlikeholdsstrategier for ét objekt, men da ligger strenge antagelser på kostnadsfunksjoner til grunn.

3.6 Heltallig lineær optimering og MOV

Heltallig lineær optimering er lineære optimeringsproblemer der variablene er heltallige. I vanlig lineær optimering er variablene kontinuerlige. Vi betegner heltallig lineær optimering med ILP (Integer Linear Programming). Slike problemer dukker opp i en rekke anvendelser, bl.a. der man skal ta diskrete beslutninger, altså velge mellom et fåtall alternativer. Generelt er ILP beregningsmessig vanskelig, NP-hardt, men mye forskning har gitt en lang rekke gode metoder som er anvendbare i praksis, uten at dette nødvendigvis er enkelt. Mye av denne utviklingen har kommet i de siste 30-40 årene, og etterhvert har denne delen av optimering vokst fram til et svært aktivt og spennende felt.

Generelt kan ILP skrives på følgende form

$$\begin{aligned} \min \quad & c^T x \\ \text{forutsatt at} \quad & Ax \leq b \\ & 0 \leq x \leq \mathbf{1} \\ & x \text{ er heltallig.} \end{aligned} \tag{3.1}$$

Her betegner O og $\mathbf{1}$ nullvektoren og vektoren med bare enere (av passende dimension), mens $c \in \mathbb{R}^n$ og $b \in \mathbb{R}^m$ er gitte vektorer. A er en reell $m \times n$ matrise, slik at $Ax \leq b$ er et system av m lineære ulikheter i variabelen $x = (x_1, x_2, \dots, x_n)$. I problemet over er x en kolonnevektor, men vi identifiserer kolonnevektorer med tilhørende n -tupler. De siste begrensingene innebærer at $x_i \in \{0, 1\}$, men skrives gjerne slik av praktiske grunner. Kostnadsfunksjonen (objektivfunksjonen) er lineær, nemlig $f(x) = c^T x = \sum_{j=1}^n c_j x_j$.

Hvis vi fjerner kravet om at x er heltallig, får vi et vanlig LP (lineær optimering) problem, som kalles LP *relaksasjonen* til ILP. En rekke praktiske metoder for å løse ILP tar utgangspunkt i å løse LP relaksasjonen, og går videre med f.eks. variabel-fiksering etc. I noen tilfelle vil man få en optimal løsning av LP relaksasjonen som (tilfeldigvis) er heltallig, og da er det klart at denne også må være optimal i ILP problemet. Det er en rik teori om hvilke matriser og systemer av lineære ulikheter som har denne egenskapen (uavhengig av c).

MOV problemet kan skrives som et ILP problem, nemlig som

$$\begin{aligned} \min \quad & c^T x \\ \text{forutsatt at} \quad & Ax \leq b \\ & Cx \leq d \\ & 0 \leq x \leq \mathbf{1} \\ & x \text{ er heltallig.} \end{aligned} \tag{3.2}$$

Her bruker vi ny notasjon og har skrevet de lineære ulikhetene splittet i to delsystemer $Ax \leq b$ og $Cx \leq d$. En nærmere forklaring av MOV er som følger:

- variabelen x har et objekt for hver mulige beslutning, en $(0, 1)$ -variabel (pga de to siste begrensingene) som angir om det tilhørende vedlikeholdstiltaket iverksettes ($x_j = 1$) eller ikke ($x_j = 0$). Antall slike variable er

$$\text{NCOMP} \cdot \text{NTIME} \cdot \text{NSTATE}$$

der NCOMP, NTIME and NSTATE er antall objekter, antall tidsskritt og antall tilstander (for hvert objekt).

- Begrensningene $Ax \leq b$ er et system av lineære ulikheter som sikrer at vektoren x representerer en sekvens av beslutninger for hvert objekt med angitt starttilstand og gitt slutttilstand, og slik at tilstander i ethvert tidsskritt er knyttet til forrige tilstand via dynamikken angitt over. Dynamikken er altså beskrevet ved hjelp av lineære ulikheter. Dette er mulig pga. endelig tilstandsrom.
- De resterende begrensningene $Cx \leq d$ representerer at summen av vedlikeholdskostnader i hver tidsperiode er høyst lik budsjettet for den perioden.
- Kostnadsfunksjonen, som skal minimeres, er $f(x) = c^T x$ som er summen av alle kostnader av Type 1 og Type 2 (se over) som følger av beslutningsvektoren x .

Det ble vist i artikkelen av Dahl og Minken (2008) at dette optimeringsproblemet er *NP*-hardt. Så, derfor kan man ikke regne med at det finnes effektive algoritmer for problemet, i en viss teoretisk forstand, slik at den løser *alle* slike problemer effektivt. Så, målet er derfor å finne metoder som fungerer best mulig for praktiske problemer man er interessert i å løse.

3.7 Lagrange-relaksasjon

VI algoritmen for å løse MOV benyttes en slagkraftig teknikk som kalles *Lagrange-relaksasjon*. I (3.2) relakseres budsjettkravene $Cx \leq d$ ved at disse fjernes fra begrensningene og inkluderes på et vis i kostnadsfunksjonen. Dette gir følgende såkalte *Lagrange-subproblem*

$$\begin{aligned} \min \quad & c^T x + \lambda^T (Cx - d) \\ \text{forutsatt at} \quad & Ax \leq b \\ & 0 \leq x \leq \mathbf{1} \\ & x \text{ er heltallig.} \end{aligned}$$

Vi betegner dette problemet med $Lsub(\lambda)$. Her er λ en ikke-negative vektor med såkalte Lagrange-multiplikatorer (duale variable). Altså er fjerningen av begrensningene $Cx \leq d$ kompensert på et vis ved å addere straffeledet $\lambda^T (Cx - d)$ i kostfunksjonen. Dette representerer en straff, altså høyere kostnad, hvis en ulikhet (eller flere) fra $Cx \leq d$ brytes.

La $v(LSub(\lambda))$ være optimale verdi i $Lsub(\lambda)$ og la $v(ILP)$ være optimal verdi i det opprinnelige problemet ILP. Da er det lett å se at

$$v(Lsub(\lambda)) \leq v(ILP).$$

Dermed har vi en nedre skranke på den optimal verdien i det opprinnelige problemet. Dette er av stor betydning fordi man da har en mulighet til å måle kvaliteten på en tillat løsning x i ILP ved å sammenlikne med den nedre skranken. Hvor god denne skranken er, vil avhenge av valget av multiplikatorvektoren λ .

Hovedpoenget med Lagrange-subproblemet er at det er lettere å løse enn ILP problemet, og som sagt at det er en approksimasjon. Ved at budsjettkravene er fjernet er også koplingen mellom de ulike objektene fjernet, slik at Lagrange-subproblemet kan løses for hvert objekt separat. Problemet $Lsub(\lambda)$ dekomponeres altså i uavhengige subproblemer, ett for hvert objekt. Dette er en betydelig forenkling, og et viktig poeng er at slike subproblem kan løses svært effektivt som korteste-vei-problemer i en enkel asyklisk graf.

Videre i algoritmen modifiseres λ på en iterative måte ved å lage en sekvens av λ -er og for hver av disse løses $Lsub(\lambda)$. Det er generelle prinsipper for oppdatering av λ som benyttes, og som baserer seg på beregningene så langt i algoritmen. Kort fortalt er oppdateringen beslektet med "steepest descent" metoder fra ikke-lineær optimering for deriverbare funksjoner, men siden

funksjonen her ikke er deriverbare, brukes et beslektet begrep til derivert (nemlig subgradienter). Denne modifikasjonen kalles *subgradientmetoden*, og er en sentral metode innen ikke-deriverbar optimering, se Nemhauser og Wolsey (1988).

På denne måten løser vi altså problemet

$$\max\{v(Lsub) : \lambda \geq O\}.$$

Dette kalles gjerne det *duale Lagrange problemet*, og hvis den optimale verdien av dette betegnes med v^* er det klart at

$$v(ILP) \geq v^*.$$

Her er $v(ILP)$ den optimale verdien vi prøver å finne, mens v^* er den beste nedre skranken vi finner fra Lagrangeteknikken.

I neste avsnitt ser vi å hvordan vi finner *øvre* skranker på $v(ILP)$. Med både øvre og nedre skranker kan man si noe absolutt om hvor god løsning man har funnet.

3.8 Heuristikk

De løsningene man får fra subproblemet $Lsub(\lambda)$ kan være gode og nesten-optimale, men, dessverre vil de kunne bryte noen av de begrensningene som er relaksert. Altstå kan budsjettet være overskredet for noen av tidsperiodene.

Vår erfaring, basert på en lang rekke testkjøringer, er at løsninger fra $Lsub(\lambda)$ når λ har konvertert typisk bare bryter budsjettet så vidt; overskridelsene er gjerne små. Dette innebærer at man kan se hvordan en liten justering av budsjettet ville gi en tilnærmet optimal løsning. Denne informasjonen er nyttig i planleggingen. Vi bør si at dette er en erfaring ut fra visse typer tester, som riktignok er forsøkt satt med en rimelig realisme i dataene. Det er imidlertid mulig å lage eksempler der løsningen ikke vil være så god. Dette er noe man må forvente fordi problemet som nevnt er *NP-hardt*.

Selv om en løsning som bryter budsjettet litt kan være svært interessant, har vi også laget metoder som gir en tillatt *primalløsning*, altså en løsning som er tillatt i ILP. Denne løsningen oppfyller budsjettet (og de andre begrensningene i ILP), men den er ikke nødvendigvis optimal. Dette er en heuristikk, dvs. en metode basert på fornuftige prinsipper og som dessuten er rask. Ofte gir den også meget gode løsninger. Heuristikken tar som input en dualløsning fra $Lsub(\lambda)$ og velger en tilfeldig rekkefølge på de forslåtte vedlikeholdstiltakene. Deretter behandles tiltakene i denne rekkefølgen og velges inntil budsjettet ikke tillater noen flere tiltak. Dermed er budsjettet overholdt. Merk at den trukne rekkefølgen som brukes her har betydning for hvilken løsning man får. Dermed kan to testkjøringer, med samme input data, ende opp med to forskjellige primalløsninger. Det er altså en viss form for randomisering i primalalgoritmen.

Merk at det kan hende at heuristikken ikke finner en tillatt løsning; dette kan skje hvis det gjenværende budsjettet er svært lavt (etter at andre tiltak er valgt), og slutttilstanden for objektet er satt slik at "tilstrekkelig" vedlikehold er nødvendig. Dette er en uunngåelig problemstilling så lenge vi spesifiserer både start- og slutttilstand for objektene. Man kan enkelt modifisere modellen samt algoritmen slik at slutttilstanden er fri, men at man da bruker en restverdi som inkluderes i modellen. (Det er ikke opplagt hvordan man setter en fornuftig restverdi, og dette er grunnen til at vi krever spesifisering av slutttilstand.)

Det er mulig å lage andre heuristikker som kan gi gode løsninger. Erfaringene med heuristikken over er svært gode, men når modellen kommer i bruk for en konkret setting kan alternative metoder eventuelt utvikles. Dette er en fornuftig angrepsmåte fordi en heuristikk bør skreddersys til den konkrete anvendelsen og de typiske trekk man har i kostnadsfunksjoner etc. Vi har enkelte generelle idéer til forbedring av heuristikken.

3.9 Videreutvikling av algoritmene

Det ligger mye arbeid bak algoritmene som er implementert i MOV. Imidlertid kan det være behov for å videreutvikle disse algoritmene av forskjellige grunner:

- Det kan være interessante planeksempler der gapet mellom øvre og nedre skranke er såpass stort at man er usikker på kvaliteten av løsningen som programmet finner.
- Regnetiden kan bli uforholdsmessig stor, spesielt der tilstandsrommet er stort.

Vi nevner nå noen idéer til videreutvikling av metodikken som vi mener det er interessant å se nærmere på:

1. *Nedre skranke*: Hovedmetodikken for å lage nedre skranke er basert på Lagrange-relaksasjon der budsjettkravene relakseres. Dette virker som et velfungerende prinsipp, men derimot mener vi at metodene for å finne beste nedre skranke kan forbedres. Det har vist seg at subgradientmetoden er følsom for parametre som påvirker skritt lengde i hver iterasjon. Her bør man finne ut hva som er gode standardvalg, og eventuelt se på dynamisk justering av disse parametrene underveis i algoritmen. Videre er stabiliseringen av subgradientretningene viktig, og kan trolig forbedres. Her finnes også generell teori man kan støtte seg til.
2. *Heuristikk/Øvre skranke*: Heuristikken bør forbedres. Dagens utgave er basert på enkle idéer og er laget for å være rask. Man kan tillate seg å bruke noe mer regnetid, og i stedet la være å kalle heuristikken i hver subgradientiterasjon. Det er en fleksibilitet her som man kan utforske nærmere. Nye heuristikker kan bruke andre prinsipper for utvelgelse av vedlikeholdsprosjekter i hvert tidsskritt. Stadig er det fornuftig å basere seg på dualløsningen, som altså kan overskride budsjettet. Dessuten kan det være en god idé å bringe inn dynamikken bedre i heuristikken: valgene man har gjort ved tidspunkt t kan påvirke valgene ved neste tidspunkt. For eksempel, et prosjekt som ble valgt i dualløsningen, men falt utenfor i primalløsningene ved tid t , kan få en økt preferanse ved tid $t + 1$.
3. *Ny overordnet idé*: En ny, og ganske annerledes metode vil være å se på ILP formuleringen av problemet og lage en algoritme basert på å løse LP relaksasjoner av problemet kombinert med variabel-fiksering. Med en rask LP-solver som f.eks. OPL-CPLEX vil man kunne løse initiell LP relaksasjon raskt, og bruke den optimal løsningen til å fiksure (låse) noen variable. Deretter modifiseres problemet ut fra disse valgene og man løser den nye LP relaksasjonen. Så fikseres igjen noen variable, og man fortsetter slik. Dette vil også være en rask metode, som virker meget rimelig å se nærmere på, spesielt fordi tidligere teoretiske resultater indikerer at LP relaksasjonene kan være meget gode for dette problemet.

4. Praktisk eksempel

Som vist i kapittel 3 og dokumentert i vedlegg A, har vi etablert et system som raskt kan finne tilnærmet optimale vedlikeholdsstrategier for en gitt vektor \mathbf{B} av årlige (eller periodevise) budsjetter. Det gir mulighet til å foreta eksperimenter ved å variere \mathbf{B} på en systematisk måte. Den enkleste formen for slike eksperimenter er følsomhetsanalyser: Hva skjer om vi øker budsjettene med en fast prosent? Hva skjer om vi øker det i de første årene på bekostning av en reduksjon i årene etterpå, eller om vi utsetter vedlikeholdsinnsatsen noen år? Hva skjer om kostnadene for visse typer av tiltak øker, eller vi introduserer mer slitesterke materialer? Vil det lønne seg å bygge mer solid?

Simulering er en litt mer avansert form for slike eksperimenter. Det er en måte å ta hensyn til usikkerhet på. Usikkerheten kan knytte seg til målefeil ved registrering av tilstanden eller virkningene av tiltak, eller en mer grunnleggende usikkerhet om nedbrytningen eller virkningene. Eller det kan være budsjettbevilgningene i framtida som er usikre. Hvis en kan si noe om sannsynlighetsfordelingen til de usikre parametrene, kan en da bruke simulering til å finne sannsynligheten for den framtidige utviklingen av nøkkelvariable, som den gjennomsnittlige tilstanden, antall objekter som befinner seg i en uakseptabel tilstand, det samlede vedlikeholdsetterslepet osv. En kan også undersøke hvilke budsjettøkninger som skal til for å oppnå gitte mål for gjennomsnittstilstanden med en viss sannsynlighet. En kan tenke seg at slike eksperimenter blir programmert slik at de kan gjennomføres med få tastetrykk.

I dette kapitlet skal vi ta for oss et eksempel på mulige eksperimenter.

4.1 Testsituasjonen

I eksperimentet tenker vi oss at programmet skal brukes til å finne den fordelingen av budsjettmidlene over tid som gir de laveste samfunnsøkonomiske kostnadene.

MOV krever et visst arbeid med å spesifisere input til modellen. Her skal vi anta at det dreier seg om 25 objekter, alle med gitt initialtilstand. Vi antar videre at det kan finnes 40 ulike tilstander som et objekt kan være i. Analyseperioden består av 30 perioder (år). Nedbrytningen av objektene fra én periode til den neste som funksjon av trafikkbelastningen, er slik som beskrevet i Dahl og Minken (2008). Alle objekter følger samme lov for nedbrytningen. Det samme gjelder kostnaden per kilometer for brukerne ved de ulike tilstandene og mengden av mulige vedlikeholdstiltak som står til disposisjon for infrastrukturholderen. Men vi antar at objektene har ulik størrelse, slik at både brukerkostnaden og tiltakskostnadene varierer med størrelsen på objektet.

I Dahl og Minken (2008) har alle årlige budsjetter verdien 125. I vårt tilfelle skal vi gjennomføre simulering for å finne den best mulige fordelingen av budsjettmidlene over år. Vi antar derfor at de årlige budsjettene kan anta verdier mellom 115 og 125. Om vi nå forutsetter at alle verdier i dette området er like sannsynlige, kan vi trekke 30 tilfeldige tall i dette området, ett for hvert år. For hver trekning av en slik tillatt budsjettvektor $\mathbf{B} = \{B_1, B_2, \dots, B_{30}\}$, anvender vi en algoritme på MOV for å beregne optimale samfunnsøkonomiske kostnader. Dette gjentar vi mange ganger (totalt 1000 vektorer i tre ulike omganger), og skaffer oss på den måten et materiale til å bedømme hvordan den optimale fordelingen av budsjettene ser ut.

I testeksempelet er kun budsjettet mulig å endre, alle andre egenskaper ligger fast.

Et slikt testeksempel gjenspeiler den virkeligheten som en bruker av programmet vil møte. Antall objekter, deres beskaffenhet, hvordan de kan vedlikeholdes, hvordan de nedbrytes og andre egenskaper vil ved de fleste anvendelser være gitt av empiri. Det eneste som normalt vil være mulig å endre på er budsjetteringen.

Utenom dette normalt ilfellet, der vi bare endrer budsjettet, finnes det også andre interessante anvendelser, nemlig analyse av endring i objektenes egenskaper, hvilke tiltak som er tilgjengelige, og hvor de virker. Se nærmere i Minken, Dahl og Steinsland(2008) om dette.

4.2 Optimal allokering av budsjettmidler

Når MOV kjøres på et gitt sett av input-verdier rapporteres flere ulike egenskaper ved den løsningen som MOV velger, som øvre og nedre skranke for optimal løsning, foreslått løsning, vedlikeholdskostnad for foreslått løsning for hver periode, total kostnad for foreslått løsning for hver periode, hvilke objekter som skal vedlikeholdes i foreslått løsning, i hvilken rekkefølge og hvor mye vedlikehold som skal gjennomføres osv. Dette har vi ikke gått nærmere inn på her.

Et firma eller en etat vil normalt ønske å gjøre den totale kostnaden både ved bruk og vedlikehold av objektene minst mulig. I dette avsnittet undersøker vi hvordan det kan gjøres, gitt at vi kan disponere en fast mengde budsjettmidler fritt over alle tidsperiodene.

Problemet som MOV løser for gitte årlige budsjetter, er krevende. Det samme må sies om problemer som krever at MOV skal kjøres gjentatte ganger med sikte på en best mulig fordeling av budsjettmidlene over årene. Vi har ingen garanti for at den total kostnaden som MOV rapporterer endrer seg gradvis når budsjettene endrer seg, tvert i mot vil vi forvente at ved små endringer i budsjett kan vi få diskontinuerte hopp i rapportert total kostnad. Årsaken til dette er at algoritmen kan velge en annen korteste veg når MOV benytter Bellman's algoritme. En stor mengde standardmetoder for minimering av funksjoner, som bygger på kontinuitet, kan dermed ikke benyttes i vår leting etter et budsjett som gir minste total kostnad. Å lete etter løsninger manuelt er utelukket grunnet det enorme antallet muligheter funksjonen har for variasjon, så vi må benytte algoritmer til å lete for oss. Videre vil også heuristikken for rapportering av beste funnet løsning i MOV vanskeliggjøre direkte bruk av algoritmer, ettersom vi får små variasjoner i rapportert beste løsning fra gang til gang. Vi kan her velge å enten gjøre heuristikken bedre eller å slå den av, og vi har valgt det siste.

Heldigvis finnes det flere algoritmer som ikke er avhengige av kontinuitet for å kunne anvendes. En av de meste kjente er Nelder-Mead algoritmen, forkortet NM-algoritmen, se Nelder og Mead (1965), og den vil vi benytte her.

La $\mathbf{MOV}(\mathbf{B})$ være problemet som løses i kapittel 3 med parametre som i vår testsituasjon skissert i kapittel 4.1 der $\mathbf{B} = \{B_1, B_2, \dots, B_{30}\}$ betegner en vektor av budsjetter over 30 år. For å kunne anvende NM-algoritmen innfører vi et ledd som straffer avvik, både opp og ned, fra fastsatt totalbudsjett¹. Problemet vi skal løse blir da,

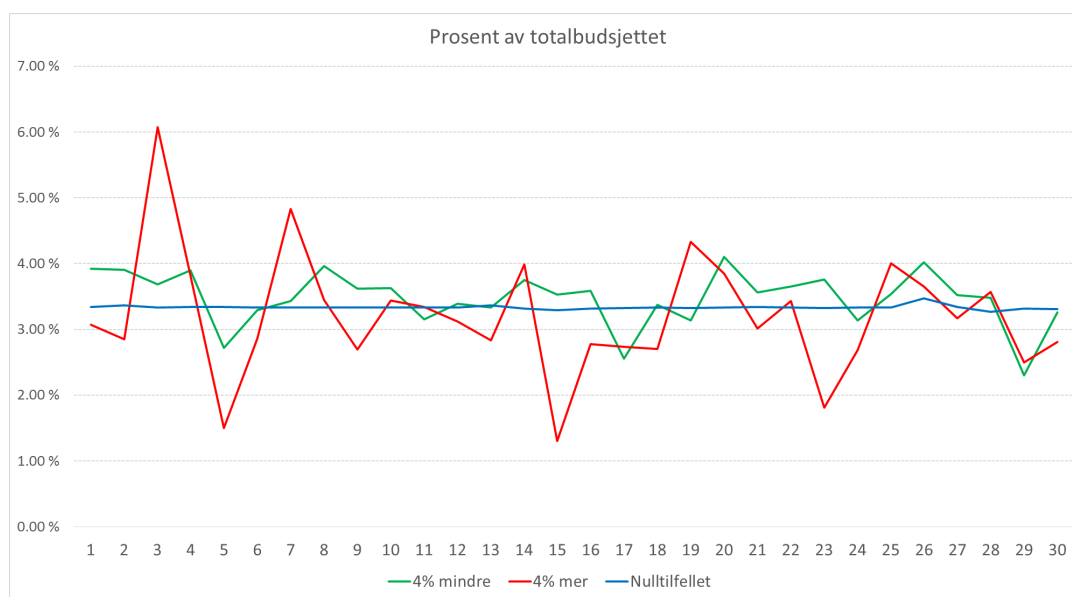
$$\min \mathbf{MOV}(\mathbf{B}) + f \cdot \left(K - \sum_{i=1}^{30} B_i \right)^2 \quad (4.1)$$

hvor K er den totale budsjettkosten vi ønsker å begrense oss til over alle periodene. Vi har også tatt med en faktor f som bestemmer styrken på hvor mye avvik straffes med. Vi har satt $f = 0.1$ for at straffen for avvik fra gitt verdi ikke skal få dominere funksjonen i ligningen 4.1.

¹Her kunne vi selvsagt tenkt oss at vi heller straffet avvik over totalbudsjett og premierte avvik under totalbudsjett men det har vi valgt å se bort fra her i og med at det er fort gjort å påvirke optimaliseringen for mye ved en slik premiering

Når vi anvender NM-algoritmen på funksjonen i 4.1 får vi for et budsjett som inn-vektor en ut-vektor som er et nytt budsjett hvor algoritmen har søkt å minimere samfunnsøkonomisk kostnad og avvik fra totalt tilgjengelig budsjett ved å omfordele tilgjengelige budsjettmidler mellom de ulike årene.

Ulempen med NM-algoritmen er at den ikke bruker all informasjonen vi måtte ha om funksjonen som skal minimeres. Hvis vi har gitt en funksjon med mange lokale minima vil den ofte konvergere mot det første og beste minimaet den finner. Figur 4.1 viser fordelingen av budsjettmidler (ut-



Figur 4.1: Resultat av tre kjøring av NM-algoritmen på problem 4.1 med ulikt totalbudsjett. x -aksen er år fra 1 til 30, y -aksen er andelen av det totale budsjettet over hele perioden som foreslås brukt.

vektoren) for tre kjøring av NM-algoritmen, hver med ulikt totalbudsjett som inn-verdi. De ulike kurvene viser resultatet etter 800 iterasjoner. Figuren viser hvor stor andel (i prosent) av totalbudsjettet som av NM-algoritmen er foreslått benyttet i hver budsjettperiode.

I det som kalles nulltilfellet er K fordelt likt over alle periodene før NM-algoritmen begynner. For de andre tilfellene er det gitt en økning(rød linje) og reduksjon(grønn linje) på 4% i totalbudsjettet K som deretter er fordelt likt på alle periodene før NM-algoritmen satt i gang. Vi observerer at de ulike budsjettene som legges inn i NM-algoritmen resulterer i tre ganske ulike svar hva størrelsen på budsjettet per år angår. Rapportert totalkostnad varierer imidlertid veldig lite imellom de ulike tilfellene (0.36% mellom høyeste og laveste totalkostnad). NM-algoritmen går tydelig i første og beste minima den finner og blir værende der.

Et stort antall kjøring av NM-algoritmen har blitt gjennomført på ulike budsjetter for å få en ide om hvordan funksjonen i ligning 4.1 ser ut. Konklusjonen er at funksjonen er ikke-konveks og særdeles irregulær, selv på et rimelig oversiktelig testeksempel som vårt. Ved en kjøring av NM-algoritmen på 100 ulike budsjetter ble det for eksempel funnet 70 ulike lokale minima. Det er ingen grunn til å anta at kompleksiteten denne funksjonen viser vi gå ned ved praktiske anvendelser av MOV.

Vi henvises derfor til å gjøre søk etter gode lokale minima blandt et stort antall mulige, og fremgangsmåten vi har valgt er etter mye testing et stegvis tilfeldig søk etter minima ved hjelp av NM-algoritmen. Fremgangsmåten er som følger:

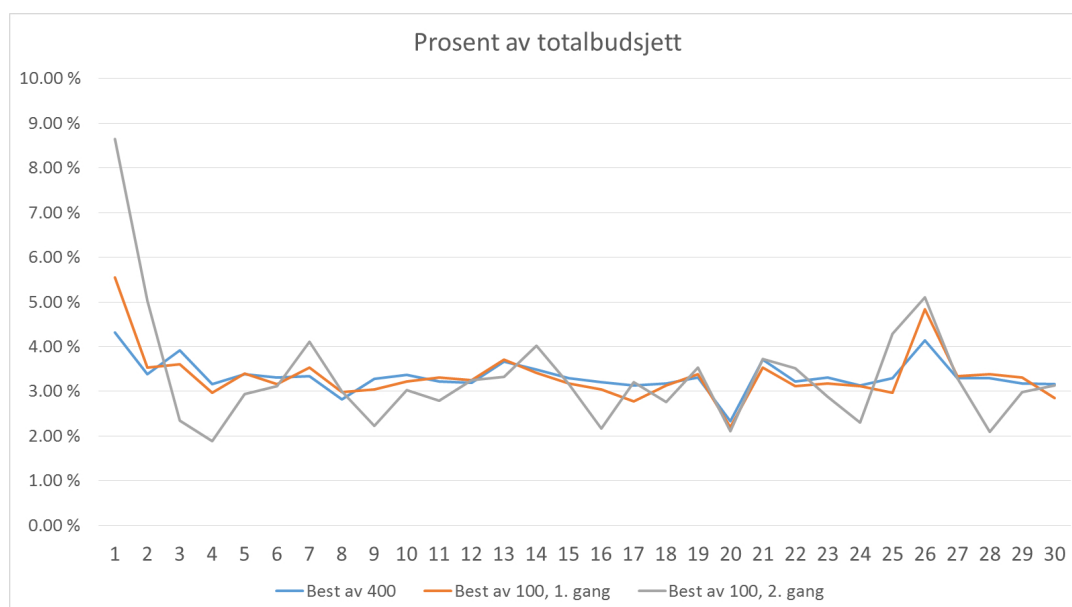
- Vi genererer 400 tilfeldige punkt på overflaten til en kule i \mathbb{R}^{30} og legger dette til nullbudsjettet

i Figur 4.1. Deretter kjøres NM-algoritmen, med 200 iterasjoner, på disse budsjettene. Vi kaller svaret fra denne kjøringen for T_0 .

- De tre budsjettene, $B_1(T_0)$, $B_2(T_0)$ og $B_3(T_0)$ som gav lavest totalkostnad i T_0 velges og fra hvert av dem genereres 100 nye budsjetter som som NM-algoritmen kjøres på. Svaret fra denne kjøringen kaller vi for T_1 .
- De tre budsjettene, $B_1(T_1)$, $B_2(T_1)$ og $B_3(T_1)$ som gav lavest totalkostnad i T_1 velges og fra hvert av dem genereres 100 nye budsjetter som som NM-algoritmen kjøres på. Svaret fra denne kjøringen kaller vi for T_2 .
- Det budsjettet, $B_1(T_2)$ som gav lavest totalkostnad i T_2 rapporteres som beste budsjett.

Denne fremgangsmåten er konseptuelt enkel og kan etter all sannsynlighet også anvendes når MOV blir brukt på realistiske tilfeller. Men den er ganske tidkrevende. Totalt tar det rundt tre og et halvt døgn i ren beregningstid å gjennomføre disse beregningene på vårt testdatasett og det totale antall ganger MOV algoritmen kalles er i underkant av en halv million. Det er rimelig å anta at beregningstiden vil øke raskt med et større tilstandsrom og flere objekter.

For å vise at metoden er brukbar ble den anvendt på vårt testeksempel. Resultatet av det



Figur 4.2: Andel av totale budsjettmidler foreslått brukt per år.

kan sees i Figur 4.2.

Som vi ser er det tydeligste trekket at budsjettet i første periode trekkes kraftig opp. Dette betyr i vårt testeksempel at det er hensiktsmessig å ta igjen mye av vedlikeholdsetterslepet tidlig i perioden. Dette er som ventet². Besparelsen på totalkosten er liten. Den laveste totalkostnaden rapportert er 1.61% bedre enn nulltilfellet, hvor totalbudsjettet deles likt på alle periodene. Men ettersom totalkosten er stor sammenlignet med totalbudsjettet, utgjør besparelsen i beste tilfelle sett opp mot nulltilfellet hele 5.2% av totalbudsjettet. Selv om totalbudsjettet er det samme, sparer det derfor i vårt testeksempel samfunnet for kostnaden av drøye halvannen periodes vedlikehold.

²Dersom vi hadde brukt neddiskontering, slik at periodene fikk mindre vekt i regnestykket jo lengre ut i tid de kom, ville det trolig ha vært enda tydeligere. Men selv uten neddiskontering gir det åpenbart gevinst å komme så tidlig som mulig inn på et godt spor.

4.3 Konklusjon

Problemsystemet MOV vil raskt kunne finne en nær-optimal vedlikeholdstrategi for gitte årlige budsjetter. Det gir gode muligheter til å studere hva som skjer når budsjettene varieres. Vi har vist hvordan MOV kan brukes til å optimere både vedlikeholdsbudsjettets størrelse og tidsutvikling. Metoden som skisseres i dette kapitlet er generell og kan anvendes på det samme sett av tilfeller som MOV kan anvendes på. Tilgjengelig regnekraft setter dog større begrensninger ved en slik optimalisering av budsjettet enn en anvendelse av MOV alene, da NM-algoritmen må kjøre MOV mange ganger i sekvens. Ved en anvendelse på et praktisk tilfelle er det å anbefale at det settes av midler til implementering av en annen algoritme en NM-algoritmen for søk etter minima, da det finnes metoder som er mer effektive, men ble for kompliserte å implementere her.

5. Litteraturoversikt

Prosjektet som denne rapporten rapporterer om, befinner seg i et fagfelt på grensa mellom ingeniørkunnskap, operasjonsanalyse og økonomi. Hovedsiktemålet i prosjektet er å utvikle redskaper for økonomisk analyse av vedlikeholdsstrategier, spesielt spørsmålet om det er samfunnsøkonomisk lønnsomt å øke vedlikeholdsbudsjettene. Men dette krever som hovedregel langt mer kompliserte matematiske og operasjonsanalytiske verktøy enn det vi har vært vant til i samferdselssektoren i Norge. Det krever også ingeniørkunnskaper om tilstandsmåling, nedbrytning av infrastrukturen og virkninger av tiltak.

Kapittel 2 i denne rapporten går så langt som mulig i retning av å forenkle analyse-verktøyet slik at en vurdering av om det er lønnsomt å øke vedlikeholdsbudsjettene skal kunne la seg gjennomføre med det kunnskapsgrunnlaget og det datagrunnlaget vi har nå. Kapittel 3 og 4 rapporterer om arbeidet med et mer detaljert og realistisk verktøy, som ikke bare skal kunne foreta en grov vurdering av lønnsomheten av å øke budsjettet, men også kunne brukes til utarbeiding av vedlikeholdsstrategier, dvs. til prioritering av innsatsen til konkrete infrastrukture objekter, eller i alle fall til relativt homogene grupper av slike objekter, og til å angi hva slags tiltak som det lønner seg å bruke på det enkelte objekt, eller i alle fall på den enkelte gruppe av objekter. Kapitlene viser at vi har løst den grunnleggende oppgaven, nemlig å programmere et dataverktøy som kan løse et svært bredt spekter av slike problemer, og som er helt på høyden med de beste slike verktøy i utlandet. Men vi mangler ennå brukergrensesnittet som skal gjøre det enkelt og fleksibelt å benytte i praksis. Vi mangler derfor også en del erfaringer med verktøyet.

Kapittel 5 er ment som en hjelp til de som forhåpentligvis vil arbeide videre med dette, enten det er analyser av den typen som er beskrevet i kapittel 2, eller det er ferdigstilling og bruk av verktøyet i kapittel 3 og 4. Det er en litteraturoversikt som både skal dekke den nyeste litteraturen, prøve å fastlegge noen definisjoner, og gi noen korte blikk på emnets historie. Vi har ikke tilstrekkelig oversikt til å kunne si at vår oversikt har med all den beste litteraturen, samtidig som vi har unngått all den dårlige. Snarere blir det opp til de som vil arbeide med dette å sortere klinten fra hveten. Vi veit at vi har utelatt mye, men forhåpentligvis ikke noe vesentlig som en ikke vil kunne nøste opp ved hjelp av litteraturlistene i de arbeidene vi har tatt med.

Hovedsakelig er det den akademiske litteraturen som er dekket. Det er en ganske betydelig kløft mellom den akademiske litteraturen og litteraturen som beskriver kommersiell programvare eller praktisk utredningsarbeid på feltet, men vi tror at vegen fram er å ta i bruk den avanserte teorien. Noe mer utredningspreget stoff har vi likevel tatt med, sammen med det vi kjenner til av lærebøker.

Litteraturen vi har sett på, er listet opp i vedlegg B. Det er altså dette vedlegget, ikke litteraturlista i hovedteksten, som må brukes om en vil se nærmere på noe av litteraturen som er nevnt i kapittel 5. (Litteraturlista dekker referansene i kapittel 1-4.) Ikke alle vil ha tilgang til de akademiske tidsskriftene der mye av litteraturen står, men ved å bruke for eksempel Google Scholar vil det i mange tilfeller være mulig å oppspore fritt tilgjengelige versjoner av artiklene. Om ikke det går, kan artikler bestilles gjennom et bibliotek.

Vi starter med å definere de viktigste begrepene i avnitt 5.1. Deretter definerer vi hva slags litteratur vi har tatt for oss, og skisserer vi hvordan denne litteraturen kan inndeles (avsnitt 5.2). Selve litteraturgjennomgangen faller i følgende deler:

Vi begynner med et avsnitt (5.3) om litteratur av mer generell karakter, dvs. lærebøker og framstillinger som ikke er knyttet til noen bestemt form for infrastruktur. Deretter behandler vi veg i tre store avsnitt - (5.4) om nedbrytning av vegdekkets tilstand og effekten av tiltak, (5.5) om optimering av vedlikehold eller rehabilitering av et enkelt objekt, og (5.6) om optimering av vedlikehold av mange objekter eller et nettverk som helhet. Så følger fire kortere avsnitt om henholdsvis jernbane (5.7), bru (5.8) og flyplasser (5.9), og til slutt et kort avsnitt om kommersielle vedlikeholdsprogrammer (5.10). Avsnitt 5.11 inneholder en kort oppsummering.

5.1 Definisjoner og begrepsavklaringer

Vår begrepsbruk samsvarer ikke helt med etatenes, men følger Minken m.fl. (2011). Grunnen til forskjellene, som er nærmere drøftet i Minken m.fl. (2011), er at vi ønsker definisjoner som ikke er etatsspesifikke, og at vårt hovedfokus er den økonomiske betydningen av vedlikehold og rehabilitering. Derfor er det hensiktsmessig å skille mellom aktiviteter som gjentas rutinemessig hvert år eller hyppigere, og aktiviteter som man kan velge å gjøre i det enkelte året eller utsette til et seinere år.

Drift: Driftsoppgaver er oppgaver som må gjentas hyppig (minst en gang hvert budsjettår), og som styres av faste regler for hva som skal gjøres, samt hvor ofte eller hva som utløser dem.

Vedlikehold: Vedlikehold er tiltak som foretas sjeldnere enn en gang hvert budsjettår, og som ikke er fullstendig regulert av faste regler. Det finnes et valg mellom ulike tiltak (inkludert ingen tiltak) i det enkelte år. Tilstanden som skal oppnås gjennom disse tiltakene, kan den enten være bestemt av faste regler eller være gjenstand for økonomiske avveininger.

Rehabilitering: Vedlikehold som gjensker konstruksjonens opprinnelige tilstand fullt ut, eller idet minste gjensker tilstanden til de viktigste av konstruksjonens komponenter fullt ut.

Reinvestering: Reinvestering er å vrake et objekt og erstatte det med et nytt av samme slag. Når deler av objektet har uendelig levetid, brukes begrepet reinvestering også om de mest fullstendige formene for rehabilitering.

Oppgradering eller fornyelse: En reinvestering eller rehabilitering med elementer av nyinvestering.

Reparasjon: Tiltak som utføres på objekter som har sluttet å fungere etter hensikten, med sikte på å gjensker funksjonaliteten. Jernbaneverket skiller mellom forebyggende og korrektivt vedlikehold. Det er det samme som skillet mellom planlagt og ikke planlagt vedlikehold. Reparasjoner er da en form for korrektivt eller ikke-planlagt vedlikehold.

Tilstand: Objektets tilstand er det eller de forhold som bestemmer hva det koster i vid forstand for brukerne å bruke objektet (enten det er i form av anstrengelser eller tidsbruk eller pengeutlegg), og hva slags framtidige kostnader eieren av objektet vil måtte påta seg for å holde det i funksjonsdyktig stand. Tilstanden i det enkelte år (eller i en annen periode) måles med en eller flere tilstandsvariable.

Nedbrytning: Uten tiltak av den typen vi har beskrevet ovenfor, vil objektets tilstand brytes ned med tida. Nedbrytningen skyldes typisk miljø- eller klimafaktorer, bruksintensiteten (i vårt tilfelle trafikkmengdene eller mengdene av spesielle typer av trafikk, som tungtrafikk), bruksmåten (for eksempel fart, last, piggdekkbruk), hvor solid konstruksjonen er bygd (spesielt styrken i underbygningen, materialene i konstruksjonen), samt det vi kan kalle "tidas tann", som omfatter alle slags faktorer som vi ikke har tatt hensyn til på annen måte.

Nedbrytningen uttrykkes i **nedbrytningsfunksjoner**, som gir tilstandsvariablene i én periode

som funksjon av tilstanden i forrige periode (eller i prinsipp hele objektets historiske liv), de nedbrytingsfaktorene det utsettes for i inneværende periode, og de tiltak som settes inn i inneværende periode.

Nedbrytningen er vanskelig å spå nøyaktig om. Man kan derfor enten anta at den skjer i henhold til en deterministisk prosess eller en stokastisk prosess. Når det gjelder veger, antas det ofte at den faktoren som betyr mest for usikkerheten, er utviklingen av bruksintensiteten, dvs. trafikkutviklingen (Chootinan m.fl. 2006). Et alternativ til å se tilstandsutviklingen som en stokastisk prosess, er å foreta hyppige inspeksjoner av tilstanden og revidere vedlikeholdsplanen i tråd med det. Det er åpenbart den beste løsningen når det gjelder å lage en årlig vedlikeholdsplan, men om man ønsker å skaffe seg oversikt over de langsiktige implikasjonene av nedbrytningen og vedlikeholdet, er dette ikke en farbar veg. De stokastiske modellene krever mye data som må anslås subjektivt. Til prinsippavgjørelser om den langsiktige utviklingen av vedlikeholdsbudsjettene og hvordan de skal disponeres, kan derfor en deterministisk modell være det beste.

Kostnadsarter: Kostnadene kan deles i to: tilstandskostnader og tiltakskostnader. Tilstandskostnadene kan igjen deles i to: Kostnader ved regulær drift og kostnaden ved avbrudd eller feil. Slik sett finns det tre kostnadsarter. Det er to parter som påføres kostnadene, vi kan kalle dem henholdsvis brukerne og etaten. Brukerne har kostnader ved regulær drift som avhenger av objektets tilstand. Det kan f.eks. dreie seg om tidskostnader, ulykkeskostnader og slitasje på kjøretøyet, dersom det er et vegobjekt det dreier seg om. Tilsvarende har etaten kostnader ved regulær drift, nemlig driftskostnader. Ved avbrudd og feil oppstår kostnader for brukerne ved de alternativene som de må bruke i stedet, og kostnaden for etaten ved å etablere slike alternativer og ved å drive det nødvendige ikke-planlagte vedlikeholdet. Tiltakskostnadene består av etatens kostnader til å gjennomføre det planlagte tiltaket og brukernes kostnader i den tida objektet ikke er tilgjengelig på grunn av verdlikehold. Se nærmere om kostnadsarter i Minken m.fl. (2011).

Det er mulig og ønskelig å inkludere de forventede kostnadene ved avbrudd og feil i de regulære drifts- og brukerkostnadene ved å anta at sannsynligheten for avbrudd og feil øker når objektets tilstand forverrer seg. Disse forventede kostnadene utgjør da en del av stigningen i brukerkostnadene og etatskostnadene som følger med nedbrytningen.

Om det ikke foreligger noe budsjett, blir det også uten betydning hvem som bærer de ulike kostnadene, så vi kan redusere kostnadsartene til bare to: tiltakskostnader og tilstandskostnader. Men straks vi har med et budsjett å gjøre, blir det vesentlig å skille etatens kostnader fra brukerkostnadene. I motsetning til mange forfattere tror vi at det er viktig i den forbindelsen å innse at både etaten og brukerne faktisk har kostnader ved ikke-planlagt vedlikehold, og at brukerne også har kostnader ved planlagt vedlikehold. I noen sammenhenger kan det dessuten være viktig å skulle mellom brukerkostnader og eksterne kostnader og mellom sluttbrukere og operatører. Derimot kan vel etatens ordinære driftskostnader som regel antas å være konstante, slik at de kan utelates fra planleggingssystemet.

De fleste planleggingssystemer bruker en langt grovere inndeling av kostnadsartene enn den vi har skissert her. Vi tror det gjør systemene mindre brukbare i praksis.

5.2 Klassifisering av litteraturen

Det vi ser på i dette kapitlet, er vitenskapelig litteratur om optimale strategier for vedlikehold og rehabilitering av infrastruktur i transportsektoren. Med vitenskapelig litteratur mener vi artikler i vitenskapelige tidsskrifter og bøker. Med optimale strategier mener vi beste måte å drive vedlikeholdet (eller rehabiliteringen) på i et langsiktig perspektiv. Det vil enten si maksimering av samfunnsøkonomisk overskudd eller minimering av de samfunnsøkonomiske kostnadene.

Litteraturen kan deles inn etter disse skillelinjene:

1. Hvordan er klassen av objekter som modellen eller artikkelen tar for seg, avgrenset? (Helt

generell, visse typer av transportmidler, visse typer av komponenter)

2. Er det ett eller flere objekter som skal vedlikeholdes?
3. Er det én eller flere perioder?
4. Finnes det budsjettbetingelser?
5. Er objektenes tilstand beskrevet med én eller flere tilstandsvariable?
6. Er tilstanden beskrevet med kontinuerlige eller diskrete variable?
7. Er nedbrytningen kontinuerlig eller diskret?
8. Er nedbrytningen deterministisk eller stokastisk?
9. Kan vedlikeholdet drives med ett eller flere slags tiltak?
10. Er virkemiddelbruken kontinuerlig (hvert virkemiddel kan ha ulike intensiteter) eller diskret?
11. Er virkningen av virkemidlene deterministisk eller stokastisk?
12. Hvilke kostnadsarter finns? (Hovedkategoriene er her brukerkostnader og infrastruktureiers kostnader. Brukerkostnadene kan deles i kostnader ved ordinær bruk og kostnader i forbindelse med tiltak. Infrastruktureiers kostnader kan deles inn i planlagte tiltak og avbøtende tiltak.)
13. Er det mulig å inspisere objektenes tilstand i hver periode?

Vi skal ikke bruke denne klassifiseringen til å til å gi en fullstendig karakteristikk av alle artikler vi omtaler. Det viktigste for oss blir om optimeringsmodellen som er brukt gjelder ett eller flere objekter, og om det finnes et budsjett eller ikke. Det vil likevel bli klart at spørsmålet om nedbrytningen er stokastisk eller deterministisk vil ha stor betydning.

5.3 Generell litteratur

Bortsett fra litteraturen om vedlikehold av bestemte slags objekter, som vegdekker, maskiner eller bygninger, finnes det også en mer abstrakt litteratur som vil være nyttig uansett hva det er som skal vedlikeholdes, bare det har en økonomisk verdi og er gjenstand for nedbrytning over tid. Dekker og Wildeman (1997), Nicolai (2008) og Nicolai og Dekker (2006) er glimrende eksempler på slik litteratur.

Lerfald m.fl. (2008) er en oversikt over gjeldende standarder, krav og regler for vedlikehold av all slags infrastruktur, ikke bare samferdsel. Den inneholder også en gjennomgang av tilstanden til infrastrukturen, tilstandsmålinger og mulige tiltak. Andersson m.fl. (2010) har som mål å etablere samfunnsøkonomiske analysemodeller for vedlikehold på samferdselsområdet i Sverige, og likner i så måte mye på vårt prosjekt.

Aurstad (2011) er en lærebok i drift og vedlikehold av veger.

5.4 Veg. Nedbrytning og effekt av tiltak

5.4.1 Funksjonell og strukturell tilstand

Vegens tilstand kan sees fra to sider: fra trafikantenes side og fra vegbyggernes side. Vegens *funksjonelle tilstand*, eller kjørbareheten, er hvor grei og behagelig den er å bruke for trafikantene. Hvis vegen er ugrei og ubehagelig, antas det å gi opphav til målbare ekstrakostnader i form av langsommere fart, økte kjørekostnader og flere ulykker. Vegens *strukturelle tilstand*, eller bæreevnen, er hvor motstandsdyktig den er mot å bli slitt ned av vær, vind, tungtrafikk og andre påkjenninger.

Den funksjonelle tilstanden blir målt med indikatorer på hvor ujamn den er på tvers (spordybde) og i lengderetningen (International Roughness Index, IRI), samt eventuelt omfanget av sprekker, humper og hull, svake kanter osv. Det vanligste i våre dager er å bruke IRI. Det finns også ulike indikatorer på den samlede kjørbareheten eller kjøreopplevelsen, og det finnes standard formler for å regne om fra den ene indikatoren til den andre.¹

Den strukturelle tilstanden for en ny veg blir uttrykt ved det såkalte *strukturtalet*, som er en vektet sum av tjukkelsen på de ulike lagene som vegen er oppbygd av. Med årene vil bæreevnen svekkes, slik at tjukkelsen på lagene ikke lenger gir det rette bildet av hvor solid vegen er. Den effektive bæreevnen på et gitt tidspunkt måles da ved *det effektive strukturtalet*, som er strukturtalet for den nybygde vegen med vektene på hvert av lagene justert ned.²

5.4.2 Et viktig eksperiment

Utgangspunktet for det aller meste av forskningen om nedbrytningen av vegen er et stort eksperiment som blei gjennomført av amerikanske myndigheter i slutten av 50-åra. På testbaner med ulike slags konstruksjon og vegdekke blei biler kjørt rundt og rundt i lengre tid, og slitasten på vegen og reduksjonen i bæreevnen blei målt. Det resulterte blant annet i en formel for levetida til et vegdekke som funksjon av trafikkvolumet og hvor solid vegen var bygget (HRB 1962).³ Oversatt fra et kjørbarehetsmål (der et høyt tall er bra) til et ujamnhetsmål (der et lavt tall er bra), ser formelen slik ut:

$$s_t = s_0 + (\bar{s} - s_0) \left(\frac{n_t}{N} \right)^b \quad (5.1)$$

Her er s_t vegdekkets ujamnhetsmål på tidspunkt t , s_0 er tilstanden når vegdekket er nytt eller så godt som nytt, og \bar{s} er tilstanden når den har blitt så dårlig at hele vegdekket må fornyes. n_t er den kumulative belastningen som vegen er utsatt for fra den var ny og til tidspunkt t . Belastningen er målt i såkalte standardaksler, fordi det er akseltrykket (ikke bilens vekt) som betyr noe for hvilke belastninger vegen utsettes for. b er en faktor som bestemmer forma på funksjonen som knytter ujamnhetsmålet til den kumulative belastningen.

Vi ser av formelen at når den kumulative belastningen har nådd punktet N , vil s_t være lik \bar{s} . N er altså det kumulative antall standardaksler som vegen tåler før den må reasfalteres. Slik formelen framstår, er det bare trafikken, eller nærmere bestemt antall standardaksler siden vegdekket sist blei fornyet, som bestemmer den funksjonelle tilstanden. Men det er bare delvis sant, for andre faktorer er også i sving, og de kan tenkes å gjemme seg bak variablene N og b .

Forskerne bak HRB (1962) antok at N avhang av strukturtalet D (altså grovt sett hvor tjukt vegdekket var) og av trafikkbelastningen, målt i standardaksler og justert for forekomsten av doble aksler. La L_1 være det kumulative akseltrykket fra antallet standardaksler som har passert siden vegen var ny, og L_2 være en dummyvariabel som teller 1 hvis en aksel er enkel og 2 om den er dobbel, og la A_0 , A_1 , A_2 og A_3 være koeffisienter som skal estimeres. Formelen de estimerte på, var formulert slik:

$$N = A_0(D + 1)^{A_1}(L_1 + L_2)^{-A_2}L_2^{A_3} \quad (5.2)$$

¹Siden IRI måler ujamnhetsmål, er et lavt tall bra, mens det er omvendt for indikatorene som måler kjørbarehet. Legg også merke til at i land som bruker inches og miles, blir IRI et helt annet tall enn i land som bruker millimeter og meter, eller meter og kilometer.

²Vegen består av mange lag. En første inndeling er i overbygning og underbygning. Overbygningen består av et vegdekke og et vegfundament. Vegdekket består av et slitelag og et bindelag. Fundamentet består av et bærelag, et forsterkningslag og eventuelt et filterlag. Det er tjukkelsen på vegdekket som betyr mest for det effektive strukturtalet. Det effektive strukturtalet måles som regel ved å slippe et lodd ned på vegen og måle hvor mange centimeter vegoverflaten blir presset ned. Hvis vegen ikke blir brukt til annet enn det den er konstruert for, vil det tilnærmet være bare det effektive strukturtalet for *vegdekket* som endrer seg med tida, mens styrken av underbygningen er konstant.

³Levetida til vegdekket er tidsrommet før det må fornyes fullstendig. Dette punktet må defineres subjektivt som dårligste akseptable tilstand, eller estimeres som den samfunnsøkonomisk beste avveiningen mellom kostnaden for en rehabilitering og mer ordinære vedlikeholds- og brukskostnader. Implisitt er formelen for levetida av vegdekket ikke bare en funksjon av den kumulative trafikkbelastningen og vegens styrke, men også av det punktet som utløser reinvestering.

Her er N , D , L_1 og L_2 gitte data, og vi ser at N er antatt å avhenge av styrketallet og trafikken målt i standardaksler, og eventuelt andre faktorer (som været) gjennom A_0 .

Forskerne formulerte også sammenhengen mellom b og styrketallet D og trafikk-belastningen $L_1 + L_2$ med en helt tilsvarende likning som likningen for N , og estimerte koeffisientene i den.

De tre likningene kan ikke utgjøre et konsistent system. \bar{s} må være eksogent gitt i likning (5.1). Hvilken N som skal til for at vegen skal bli så dårlig som \bar{s} , kan estimeres uten å bruke den formelen, forutsatt at det finnes data om den kumulative belastningen på veger som har blitt så dårlige som \bar{s} . Data om andre tilstander er irrelevant. Når N er anslått, kan den første formelen brukes til å estimere b . Likning (5.2) kan bruke N , sammen med styrketallet og sammensetningen av trafikken på enkle og doble aksler, til å bestemme blant annet b , dersom b kan identifiseres med A_2 . Men b er allerede estimert fra den første likningen! Og fra den tredje, som vi ikke har skrevet opp!

5.4.3 Re-estimering og nye eksperimenter

Uansett, dette eksperimentet og estimeringen av disse likningene blei brukt i lang tid til å utarbeide standarder for vegbyggingen. Vi veit nå at vegene som blei bygd etter denne standarden, ikke holdt så lenge som den estimerte funksjonen skulle tilsi. Small og Winston (1988) og Small m.fl. (1989) påviste flere metodefeil i estimeringen. De så bort fra den første formelen og reestimerte den andre med mer moderne estimeringsmetoder. Den reviderte formelen, med A_2 på 3,6, blei utgangspunkt for det meste av seinere arbeider om vegslitasje, spesielt beregninger av vegtrafikkens marginale eksterne kostnader. Det er potensen 3,6 som viser hvor mye større skade et tungt kjøretøy gjør enn et lett. Med en slik høy potens blir kostnaden som lette kjøretøyer påfører vegen, nesten helt uten betydning i forhold til kostnaden som tunge kjøretøyer forårsaker. Seinere har Madanat m.fl. (2002) brukt enda bedre økonometriske metoder på det samme materialet. Samtidig har det omsider også vært gjennomført nye empiriske undersøkelser i større skala, bl.a. i det såkalte Long Term Pavement Performance-prosjektet (LTPP blant venner), se (FHA 2013).

Naturligvis er nedbrytningen av en veg ikke en prosess som ter seg på samme måte overalt, uavhengig av vegtype, konstruksjon, klima osv. Det vil derfor alltid være behov for å estimere nedbrytningen i konkrete tilfeller. Man behøver slett ikke anta at nedbrytningen har samme matematiske form som i (5.1) eller (5.2), og mange ulike estimeringsmetoder er i bruk. Det vil imidlertid ofte være et problem at historiske data er ufullstendige, inneholder feil, og at tiltak på vegen ofte ikke er registrert, slik at det er vanskelig å vite om en bedring av tilstanden skyldes en tilfeldig feil eller et tiltak. Flere forfattere, blant annet Ben-Akiva og Ramaswamy (1993), Ben-Akiva m.fl. (1993), Gao m.fl. (2011) har foreslått løsninger på dette problemet.

Dye (2010) har utviklet et enhetlig system for å karakterisere ytelsesnivået (level of service) for elementene i det amerikanske hovedvegssystemet. Hensikten er blant annet å påvise behov for oppgradering og vedlikehold, sette mål for forbedringer etc. Flora m.fl. (2010) har videreutviklet det vanlige målet på vegens strukturelle styrke, og Bryce m.fl. (2012, 2013) har utviklet et mål på den strukturelle styrken på nettverksnivå.

5.4.4 Nordiske arbeider

I norsk sammenheng omhandler Sund (2004, 2005), Saba m.fl. (2006) og Evensen (2006) testing, vurdering og videreutvikling av modeller for nedbrytning på norske veger, mens Öberg (2001), Wågberg (2001) og Jacobson og Wågberg (2007) er eksempler på svenske arbeider på området. Saba m.fl. kommer for øvrig fram til den nedstemmende konklusjonen at et ikke finnes gode nedbrytningsmodeller. Lerfald (2010) gjelder bæreevnen og dens betydning for vedlikeholdspolitikken.

Larsen m.fl. (2002) dokumenterer EU-prosjektet PAV-ECO, der alle sider av vedlikeholdsplanlegging og nyttekostnadsanalyse av vedlikehold i 15 europeiske land er kartlagt, og beste praksis anbefalt.

Det har også nylig vært gjennomført et nordisk prosjekt på dette området, der det er utarbeidet en nordisk nedbrytningsmodell, kalt 3xP Nordic, med parametere estimert på grunnlag av teststrekninger i tre land (NordFoU 2006, Christiansen m.fl. 2010, Skar m.fl. (2012)). Brukerhåndboka er Baltzer og Mollerup (2013). Dette har gitt et nytt grunnlag for å formulere nedbrytningfunksjoner, både til bruk for valg av materialer og vedlikeholdsopplegg for enkeltstrekninger (project level) og til bruk i optimering av strategier for vegdekkevedlikehold (network level). Lang og Grivet (2013) har bygd på NordFoU-PPM og vedlikeholdsstyringssystemet HDM-4 og formulert formler både for nedbrytningen i flere dimensjoner og for virkningene av tiltak.

5.4.5 En mye brukt forenkling

En eksplisitt eller implisitt antakelse i det meste av litteraturen er at normene for vegbygging og akseltrykkbestemmelsene er satt slik at de nederste delene av vegen, vegfundamentet og underbygningen, aldri utsettes for nedbrytning.⁴ Endringer i det effektive strukturtallet vil da utelukkende være endringer i *vegdekkets* bæreevne. Da vil det være en sammenheng mellom vegens funksjonelle tilstand og dens strukturelle tilstand. Sprekker, hull etc., som innebærer en dårlig funksjonell tilstand, vil gjør at trafikken og været (særlig nedbør og vekslinger mellom frost og mildvær) får større negativ virkning på bæreevnen til vegdekket. Og omvendt: Svekket bæreevne gir utslag i sprekkdannelse og spor som reduserer den funksjonelle tilstanden.

Det betyr at det ikke er urimelig å bruke funksjonell tilstand som hovedmålet på vegens tilstand under normale omstendigheter. Dette er standard i litteraturen.

5.4.6 Stokastisk eller deterministisk nedbrytning?

Nedbrytningen kan modelleres som en deterministisk eller en stokastisk prosess. I utgangspunkt er det såpass mange ukjente eller usikre faktorer som påvirker den, at det virker mest naturlig å anta at den skjer stokastisk. Den vanligste stokastiske modellen er en markov beslutningsprosess (MDP). Men to ting taler for å nøye seg med å estimere forventningsverdien. Det ene er at det krever mer data om man også skal si noe fornuftig om avvikene fra forventet utvikling. Ng m.fl. (2011) snakker for eksempel om at "the curse of dimensionality" gjør det vanskelig å løse storskala MDP-problemer. Derfor er stokastiske modeller mye vanligere når modellen bare omfatter ett objekt enn når den omfatter et helt nettverk. Det andre er at det blir vanskeligere å bedømme om vedlikeholdsbudsjettene lar seg holde på langt sikt dersom vi er usikre på hvordan infrastrukturens tilstand vil være.⁵ Siden det er mer interessant å introdusere budsjettrestriksjoner når vi har mange objekter, bidrar også det til at stokastiske modeller brukes mest i enkle sammenhenger med ett eneste objekt.

Et annet moment i valget mellom stokastisk og deterministisk modell er dette: Jo flere typer av virkemidler og tiltak man antar, og jo flere dimensjoner tilstandsrommet har, jo flere parametre blir det som skal estimeres, og dette trekker i retning av å anta deterministisk nedbrytning. Hvis man dessuten har mange objekter og skal anta ulik nedbrytning for hvert av objektene, blir det egentlig en illusjon at man tar hensyn til usikkerheten ved subjektivt å anslå alle disse parametrene. Om det ikke tar for lang tid å kjøre modellen, vil det alltid være mulig å håndtere usikkerheten på en annen måte, nemlig ved simulering eller følsomhetsanalyse i etterkant av modellkjøringen.

Den vanligste stokastiske modellen er som sagt en markovprosess. Det innebærer at både tida og vegens tilstand er diskrete variable. Sannsynligheten for at tilstanden s_t på tidspunkt t skal være S , der S er én av et endelig antall mulige tilstander, avhenger av tilstanden i perioden før, s_{t-1} , men ikke av tidligere tilstander, som s_{t-2} , s_{t-3} osv. Styrken til en slik modell er at den diskrete oppdelingen av tida i år samsvarer med hva slags data vi har, i og med at tilstanden som

⁴Dette kan også kreve at rene driftsoppgaver som rensking av stikkrenner m.m. gjennomføres punktlig.

⁵Dersom den usikre tilstandsutviklingen er skeivt fordelt, vil vedlikeholdsplanen som er utarbeidet med forventet nedbrytning og gitte budsjetter, enten kreve mer penger enn budsjettet tilsier, eller medføre underforbruk, alt etter som det er høyre- eller venstreskeivhet. Dette kan unngås ved å bruke en annen verdi enn forventningsverdien i den deterministiske beregningen. Men hvilken? For å svare på det må vi kjenne fordelingen.

oftest bare blir registrert en gang i året. Svakheten er at det ikke er rimelig å anta at forhistorien ikke har noe å si. Alt annet likt vil en og samme tilstand på en veg som sist blei rehabilitert for 15 år siden, slå annerledes ut til neste år enn om vegen var rehabilitert for 2 år sida. I tillegg kommer at å estimere sannsynlighetsfordelingen i seg sjøl krever mer data enn om vi bare estimerer et punktanslag.

5.4.7 Estimering med usikre data

For å ta hensyn til at modeller basert på MDP er følsomme for feil og hull i data, utviklet Madanat and Ben-Akiva (1994) to versjoner av en latent Markov beslutningsprosess (LMDP), én der infrastrukturens tilstand blei inspisert med faste mellomrom, og en der hyppighetene på inspeksjonene blei valgt på en kostnadseffektiv måte. Ben-Akiva og Gopinath (1995) er en videreutvikling av dette. Madanat m.fl. (1995) dreier seg om hvordan MDP-sannsynlighetene skal beregnes. Mishalani og Madanat (2002) oppsummerer denne litteraturen.

Durango og Madanat (2002) bruker MDP og metoden "adaptiv kontroll". De finner sannsynlighetsfordelingen for nedbrytningen under mange ulike forutsetninger, tillegger hver av løsningene subjektive sannsynligheter og oppdaterer disse sannsynlighetene når det kommer ny informasjon. Durango-Cohen og Madanat (2008) kombinerer den samme tilnærmingen til usikkerhet om nedbrytningen med muligheten for at inspeksjonen av tilstanden kan ha feil. De trekker konklusjoner om hvilke type usikkerhet som er viktigst.

Smilowitz and Madanat (2000) behandler konstruksjon av MDP-sannsynligheter for problemer som involverer mange objekter når mangelfulle data gjør estimering umulig. Guillaumot m.fl. (2003) kombinerer deres tilnærming med inspeksjoner og "adaptiv kontroll".

Ng m.fl. (2011) formulerer to modeller med usikkerhet basert på heltallsprogrammering, altså ikke MDP. De beskriver usikkerheten ved hjelp av et intervall rundt en utgangsverdi for hver av de usikre parametrene. Deres formulering gjør det mulig å sammenlikne deterministiske og stokastiske modeller og beregne forskjellen på vedlikeholdskostnadene når nedbrytningen eller virkningen av tiltaket er usikker og når den er sikker, samt å anslå verdien av å innhente mer presis informasjon. Modellene gjelder et enkelt objekt. En bibetingelse er at objektet skal ha en viss minste kjørlighet til enhver tid. Modellen kan naturligvis løses for mange objekter som kan sammenliknes med hverandre ved utarbeiding av en felles plan, og det ser ut til at dette er den bruken som forfatterne tenker seg. Uansett, hvis man er mest opptatt av å sikre et minstenivå på den funksjonelle tilstanden, samt å finne ut hva redusert usikkerhet kan spare av midler for å oppnå dette, er dette en aktuell modell.

Hva slags data trenger vi til å estimere nedbrytningen? Til å svare på det er Chu og Durango-Cohen (2008b) en viktig kilde. De framholder at tilstanden ikke kan fanges opp på en god måte med et eneste tall, dvs. tilstandsrommet bør være flerdimensjonalt. Videre framhever de at en bør bruke paneldata, dvs. kombinere tidseriedata om enkeltobjekter med tverrsnittsdata om mange ulike objekter. På den måten fanger en opp virkninger som ikke kan identifiseres med bare tidsserie-data eller bare tverrsnittsdata, og en kan identifisere særtrekk for det enkelte objektet. Paneldata er derfor den eneste måten å få gode estimater på når det ikke bare forekommer nedbrytning, men også vedlikehold. Vedlikehold skjer jo ikke samtidig på alle objekter, og det er heller ingen tilfeldig målefeil. Som allerede Ben-Akiva og Ramashwami (1993) var inne på, må vedlikehold og nedbrytning altså estimeres samtidig. Chu og Durango-Cohen (2000a) foreslår å bruke intervensjonsanalyse sammen med metoden i Chu og Durango-Cohen (2008a) for å finne virkningen av vedlikeholdet.

Endelig oppsummerer de erfaringer som viser at vegdekker generelt ikke har markovegenskapen. Markovegenskapen vil si at tilstanden i tidligere perioder, som $t - 1$, $t - 2$ osv., er uten betydning for nedbrytningen i periode $t + 1$, Bare tilstanden i periode t har betydning. De viser så hvordan en skal sette opp og estimere en modell som tar hensyn til alt dette. Dette er jo alvorlig for alle modeller basert på MDP.

5.4.8 Effektmodeller

Vi har sagt lite om modellering av hvordan tiltak kan forbedre tilstanden, men en viktig referanse på det området er Ouyang og Madanat (2004), som bygger på Paterson (1990). Chu og Chen (2012) drøfter virkningen av tiltak på en mer abstrakt måte. Chu og Durango-Cohen (2008b), som vi viste til i forrige avsnitt, vil naturligvis gi viktige retningslinjer dersom en skal estimere virkningene. Ellers finnes det litt her og der i litteraturen om effekten av konkrete tiltak.

5.5 Veg. Optimal vedlikeholdsstrategi for ett objekt

5.5.1 Sammenheng med investeringspolitikken

Formålet med planlegging og styring av vedlikeholdet av samferdselsinfrastrukturobjekter er å sikre optimalt vedlikehold og rehabilitering over objektet levetid, samt optimal levetid for objektet (dvs. optimal reinvesteringstakt). Som vi har vært inne på, krever det at objektet blir bygget tilstrekkelig solid, slik at ikke levetida blir for kort. Men det skal heller ikke bygges *for* solid. Det riktige styrketallet er punktet der en liten økning av styrken koster like mye ekstra i byggekostnader som nåverdien av de sparte vedlikeholds- og rehabiliteringskostnader og de sparte brukerkostnadene som den økte styrken gir opphav til (Small m.fl. 1989).

5.5.2 Problemet

Forutsatt at investeringen gir objektet riktig styrke, er en optimal vedlikeholdsstrategi et spørsmål om å minimere de samlede kostnadene over levetida. De samlede kostnadene er da tiltakskostnadene pluss brukerkostnadene.⁶ Vi trenger altså å vite hvordan vegen blir dårligere for brukerne etter hvert som tida går uten større vedlikehold (helt rutinemessig vedlikehold, samt driftsoppgaver som må gjøres uansett, kan vi se bort fra), og vi trenger å vite hvilke ekstrakostnader det fører til for brukerne.

Alternativt kan problemet formuleres som minimering av vegholderens kostnader over levetida, gitt at et visst minstenivå eller gjennomsnittsnivå for objektets tilstand blir oppnådd. Brukernes kostnader er da implisitt tatt hensyn til gjennom denne bibetingelsen.

Det finns lesverdige litteraturoversikter i Durango-Cohen og Madanat (2008) og mange av de andre arbeidene vi skal nevne nedenfor, kanskje spesielt de som har Madanat som medforfatter. Problemet blei først formulert som et deterministisk problem med kontinuerlig tid og kontinuerlige tilstander. Eksempler på dette:

Friesz og Fernandez (1979) var de første som formulerte problemet som et optimalt kontrollteori-problem. Fernandez og Friesz (1981) og Markow og Balta (1985) behandler optimal reinvestering- eller rehabiliteringsstrategi. I motsetning til resten av denne litteraturen, som bruker én tilstandsvariabel, bruker Markow og Balta to, en for funksjonell og en for strukturell tilstand. Tsunokawa og Schofer (1994) brukte en glatt kurve for å beskrive tilstandsforbedringen ved et tiltak, siden de ikke kunne takle sprangvise endringer i tilstanden.

Li og Madanat (2002) og Ouyang og Madanat (2006) er svært sentrale referanser på dette feltet. Grunnen er at de beviser enkle regler for hvordan rehabilitering av vegdekker skal drives om det ikke finnes noen budsjettbegrensninger eller andre kompliserende hensyn. Regelen er å rehabiliterer hver gang objektet når en viss minstestandard, og derfra foreta en fullstendig rehabilitering, slik at vegdekket blir så godt som nytt. I følge Chu og Chen (2012) er dette en vanlig regel å følge, men da med minstestandarden fastsatt på grunnlag av ingeniørskjønn, ikke på grunnlag av maksimering av samfunnsøkonomisk overskudd, slik som litteraturen vi ser på.

Både Li og Madanat (2002) og Ouyang og Madanat (2006) løser problemet analytisk, og viser at usikkerhet om nedbrytningen ikke har store konsekvenser for løsningen.

⁶Vi ser bort fra eksterne miljøkostnader.

Forutsetningene i Li og Madanat (2002) er kontinuerlig tid og tilstand, deterministisk nedbrytning, ingen endring av trafikken eller miljøet over tid, og at både nedbrytningen og virkningen av tiltak har markovegenskapen, dvs. bare tilstanden i inneværende periode har betydning for tilstandsutviklingen i neste periode. Det finnes bare ett slags tiltak, men det kan brukes med en intensitet som kan velges på en kontinuerlig skala. Ethvert brudd på disse forutsetningene er altså en mulighet for at den optimale løsningen kan oppføre seg betydelig annerledes.

I disse og andre litt eldre artikler finnes det bare én type av tiltak, men den kan ofte brukes med ulik intensitet. I Rashid og Tsunokawa (2012) er det imidlertid hele tre typer av tiltak, som alle skal brukes optimalt. Enkle og billige tiltak kan ikke bare forbedre kjøreforholdene betydelig på kort sikt, men også forsinke nedbrytningen. Dermed blir det sjeldnere nødvendig med full rehabilitering. Slike tiltak kan også være et alternativ til full fornyelse av det øvre laget av dekket (Labi og Sinha 2003, 2005). Dette er ikke bare tatt med i modellen til Rashid og Tsunokawa, men også i Gu m.fl. (2012).

Lamprey m.fl. (2008) vurderer bare én type tiltak, men det er ikke reasfaltering som de fleste andre, men nettopp forebyggende vedlikehold i form av enkle tiltak som kan utsette reasfalteringen. På sett og vis er premisset at Li og Madanat tar feil - det er ingen grunn til å vente med tiltak til objektet har nådd dårligste tilstand. Artikkelen inneholder klare og gode definisjoner av vanlige begreper i vedlikeholdslitteraturen.

Lee og Madanat (2014) ser konstruksjonen av vegen og det seinere vedlikeholdet (fra nå til evigheten) i sammenheng. De understreker at trafikk og miljø varierer fra sted til sted, og at utgangspunktet derfor må være det enkelte objektet, ikke et abstrakt gjennomsnittsobjekt (bottom-up, ikke top-down). Det gjelder sjøl om det endelige formålet er en samlet plan. Derfor handler deres artikkel om optimal konstruksjon og vedlikehold av ett enkelt objekt.

Vi merket oss i avsnitt 5.5.1 at Small m.fl. (1989) hadde utledet en regel for riktig avveining mellom ressursene som legges ned i konstruksjonen, dvs. hvilken styrke vegen skal bygges med, og ressursene som brukes på vedlikehold. Men vi merket oss også i avsnitt 5.4.2 og 5.4.3 at forutsetningen for arbeidene til både AASHO (dvs. HRB 1962), Small m.fl. (1989) og Madanat m.fl. (2002) var at tilstanden som utløste rehabilitering var gitt på forhånd, ikke bestemt optimalt. Det er dette Lee og Madanat retter på. Men de har også med et lettere vedlikeholdstiltak som bremser nedbrytningen. Nedbrytningen i modellen deres er for øvrig kontinuerlig, og avhenger ikke bare av tilstanden i forrige periode, men også av trafikkmengden og den strukturelle tilstanden. Tida er diskret.

5.6 Veg, mange objekter

5.6.1 Oversikter

Gao (2011) og Gao and Zhang (2011) er en doktorgradsavhandling og en rapport, begge to svært like. Begge inneholder gode generelle oversikter over hele feltet, fra nedbrytingsfunksjoner og stokastiske tilstandsmodeller til optimering av samfunnsøkonomisk nytte av vedlikehold av mange objekter under en budsjettrestriksjon. Litteraturlista er på åtte sider, med mange referanser som ikke er tatt med i vår rapport. Modellen som blir behandlet, har det særpreget i forhold til andre modeller at budsjettet er usikkert. Vedlikeholdet skal altså planlegges under usikkerhet om hva bevilgende myndigheter finner på i framtida.

5.6.2 Uten budsjetter

Golabi m.fl. (1982) antok diskret tid og diskrete tilstander, og var den første til å bruke markov beslutningsprosesser (MDP) til å modellere nedbrytningen av objektene. Resultatet er ikke bestemte vedlikeholdsplaner for hvert objekt, men hvor store andeler av all infrastruktur i en bestemt tilstand som skal få et bestemt tiltak.

Carnahan (1988) er også en modell med diskret tid og tilstander, mens Durango-Cohen og Sarutipand (2007) har kontinuerlige tilstander og diskret tid. Et viktig verk i denne litteraturen er Ouyang og Madanat (2004), som formulerer og løser problemet med å finne den vedlikeholdspolitikken som minimerer den samfunnsøkonomiske kostnaden ved å vedlikeholde et sett av infrastrukture objekter over et endelig tidsrom, gitt at det finnes flere ulike slags vedlikeholds- og rehabiliteringstiltak, og gitt realistisk nedbrytningsfunksjoner og funksjoner for virkningen av tiltakene. De bruker blandet heltallsoptimering til å løse problemet. I Dahl og Minken (2008) har vi både brukt funksjonsformene for nedbrytning og ideer om virkningen av tiltak derfra. Spesielt viktig er at modellen tar hensyn til at det finnes nedbrytning sjøl i den aller beste tilstand, ellers vil objektene aldri komme ut av den tilstanden.

Seyedshohadaie m.fl. (2010) handler om å løse optimeringsproblemet "i verste fall" på en systematisk måte (conditional value at risk) som egner seg godt sammen med optimering. Resultatet er en optimal vedlikeholdsstrategi som er robust innafor visse grenser for usikkerheten. En modellvariant kan visstnok også ta hensyn til budsjettrestriksjoner på kort sikt (men ikke i det lange løp). Denne modellen er altså på grensa til å blir tatt med i neste avsnitt.

5.6.3 Med budsjetter

Dahl og Minken (2008), Karabakal m.fl. (1994) og Yeo m.fl (2013)

Dahl and Minken (2008) og Karabakal m.fl. (1994) er svært like. I begge artikler er modellen en diskret, heltalls optimeringsmodell med deterministisk nedbrytning og virkning av tiltakene, og tiltakene på de ulike objektene må finansieres innafor en budsjettskranke. Løsningsmåten er også i begge tilfeller i hovedsak den samme.

Vi kjente ikke til artikkelen til Karabakal m.fl. da vi skreiv vår artikkel, og blei først oppmerksom på den i 2014. Grunnen er at artikkelen bare var offentliggjort i en rapportserie fra et universitet i USA, at de har vært aktive på andre felter og ikke har noen andre arbeider om transport, og at artikkelen deres heller ikke blei sitert i noe vitenskapelig tidsskrift før i 2009.

Karabakal m.fl. framhever at stokastisk nedbrytning ved hjelp av markovprosesser går dårlig sammen med budsjettbetingelser. De som bruker markovprosesser, vil derfor i stedet velge å minimere kostnaden ved å oppnå en bestemt minstilstand for alle objekter. Men dermed vil årlige utgifter kunne variere urealistisk mye fra år til år.

En annen artikkel med likhetstrekk med Dahl og Minken (2008) er Yeo m.fl. (2013). Først beregnes samfunnsøkonomisk beste løsning for hvert objekt uten budsjettet, og så brukes en mønstergjenfinningsalgoritme eller en evolusjons-algoritme for å finne løsninger innenfor budsjettet.

5.6.4 Mot realistisk størrelse

Det er begrensninger på hvor store problemer som kan løses med metodene til Dahl og Minken eller Karabakal m.fl. Med størrelsen mener vi da antall objekter, tilstandsrommets størrelse, antall tidsperioder osv. Gao og Zhang (2012) har foreslått en metode som kan utvide grensene for hvor store problemer som kan løses. Det vil da være mulig å utvikle nesten-optimale løsninger for vegnettverk av realistisk størrelse. Vi har ikke sett nærmere på dette i vårt prosjekt, men det kan forhåpentligvis gjøres i andre sammenhenger.

Sathaye and Madanat (2011) benytter resultater i Li og Madanat (2002) og Ouyang og Madanat (2006) til å formulere den optimale løsningen for hvert av objektene på analytisk form. Det vil si at de gir slipp på å kunne anvende mange ulike vedlikeholdstiltak, og formulerer problemet i stedet som å finne det optimale nivået som utløser full rehabilitering. Det blir da forholdsvis enkelt å utvikle en algoritme som å sørge for at løsningen holder seg innafor budsjettet. Sathaye and Madanat (2012) videreutvikler metoden.

5.6.5 Modeller for fordeling av ressurser på mer overordnet nivå

Chan m.fl. (1994, 2003) foreslår å kombinere nyttekostnadsanalyse og mål for å allokere budsjettet på delområdet og distrikter. Tsunokawa og Van Diep (2008) kritiserte de subjektive elementene i disse allokeringmodellene, og foreslo en framgangsmåte basert på samfunnsøkonomisk nytte som målfunksjon. Oppgava er altså å drive virksomhet på et sett av delområder (eller vedlikeholde et sett av objekter) med mål å maksimere samfunnsøkonomisk nytte under et felles budsjett. Det kreves at vi for en hver allokering av budsjettet kan regne ut hva hvert av underområdene kan få ut av disse midlene, eller med andre ord, at vi kan formulere nytten på hvert underområde som en funksjon av budsjetttildelingen. Optimeringsproblemet vil da være ganske enkelt.

Det er sterke likhetstrekk mellom Tsunokawas og Van Dieps metode og den vi har gjort greie for i denne rapportens kapittel 2. Begge deler kan karakteriseres som "ovenfra og ned", i den forstand at de egner seg til å fordele et budsjett på den best mulige måten, eller til å si noe om nytten av å øke budsjettet, men blir litt for enkle når det gjelder å utarbeide en detaljert handlingsplan på hvert av underområdene.

5.6.6 Usikkerhet

Modellen i Chootinan m.fl. (2006) tar hensyn til usikkerhet, og bruker en målfunksjon som enten minimerer forventede vedlikeholdskostnadene gitt et mål på gjennomsnittstilstanden i nettverket, eller maksimerer forventningen til tilstanden for gitt budsjett. Imidlertid sies metoden å kunne brukes på andre målfunksjoner også. Deres viktigste bidrag er å vise hvilke konsekvenser usikkerheten har. De regner både med usikkerhet om hvilken av flere mulige nedbrytningsmodeller som er rett, og usikkerhet om trafikk, vær og andre forhold som påvirker nedbrytningen. De bruker stokastisk simulering til å finne sannsynlighetsfordelingen for objektene framtidige tilstand, og en genetisk algoritme til å velge en optimal flerårig vedlikeholdsplan for hele nettverket. Resultatet er at de kan vise at en reit deterministisk modell sannsynligvis vil undervurdere hvor dyrt det er å oppnå et visst tilstandsnivå i nettverket som helhet, og overvurdere tilstanden som kan oppnås med et gitt budsjett.

Medury og Madanat (2013) viser hvordan MDP-basert nedbrytning faktisk kan brukes også til å lage detaljerte vedlikeholdsstrategier for et helt nettverk (dvs. en "bottom-up"-modell). Strategiene holder seg innenfor budsjettet på kort sikt. På lengre sikt kan forventede kostnader beregnes.

Ohlmann og Bean (2009) fremmer en algoritme for å ta hensyn til budsjetter når det skal gjøres tiltak på en samling objekter som utvikler seg stokastisk (i følge en ikke-homogen MDP).

Durango-Cohen og Tadepalli (2006) handler om å benytte data fra overvåkingsutstyr til å planlegge vedlikehold. Metoden består av to deler, en del der data blir brukt til å forutsi framtidig tilstand, og en del som bruker dette til å optimere vedlikeholdet. State-of-the-art sies å være markovmodeller (markov beslutningsprosesser) som optimeres under budsjettbetingelser.

5.6.7 Forenkling

Wang m.fl. (2011) er et matematisk program for å gruppere likeartede og/eller nærbeliggende vegsegmenter sammen på en måte som kan spare vedlikeholdskostnader. En kan tenke seg at vedlikeholdet i hver gruppe skal settes ut på anbud, og oppdraget vil gå ut på å bruke samme tiltak på alle segmenter i en gruppe. Hvis segmentene har svært ulik tilstand, vil det innebære at noen segmenter blir behandlet med et tiltak som er unødvendig dyrt. Hvis segmentene blir for små og oppstykkede, vil administrative kostnader og kostnader ved å flytte utstyret omkring, bli unødig store. Denne problemstillingen er logisk sett en forundersøkelse før programmet for å finne optimale vedlikeholdsstrategier for hele samlingen av segmenter. Det er også grunn til å merke seg at problemstillingen må forutsettes løst på forhånd dersom vi skal trekke konklusjoner om lønnsomheten av å øke eller redusere budsjettet.

5.6.8 Integrering

Det er en trend å integrere planleggingen av vedlikeholdet for helt ulike komponenter i vegsystemet. Sinha and Fwa (1989) handler om å bygge et integrert styringssystem for dekker, bruer osv.

En annen form for integrering oppstår når trafikken på de ulike objektene avhenger av tilstanden på alle objektene. Det kan også oppstå køproblemer. Trafikken på lenkene i nettverket som skal vedlikeholdes kan da eventuelt beregnes ved å finne brukerlikevekten i nettverket. Det kan være aktuelt i byområder. Brukerlikevekten kan være følsom for endringer i vegkvaliteten som følge av nedbrytning og vedlikeholdstiltak. Vedlikehold i et nettverk der trafikken avhenger av kjørbarheten på hver av lenkene kan formuleres som et problem med to nivåer (bilevel program), se Ng m.fl. (2009). Chu og Chen (2012) er en liknende tottrinnsmodell, der oppgaven er å fastlegge den tilstanden som utløser rehabilitering under hensyn til at rehabiliteringen utløser atferdsendringer hos brukerne av nettverket.

Flere andre studier tar også hensyn til etterspørselseffektene av vedlikeholdet, for eksempel Durango-Cohen og Sarutipand (1997).

5.7 Jernbane

Det internasjonale standardverket i jernbanedrift er (så langt vi skjønner) Esveld (2001). Olsson og Veiseth (2011) er en kort generell innføring i jernbanedrift på norsk. Zoeteman and Esveld (2004) oppsummerer beste praksis når det gjelder vedlikeholdsplanlegging. Daniels (2008) er en gjennomgang av litteraturen om vedlikeholdskostnader.

I 2001 blei det etablert et langsiktig prosjekt i det norske Jernbaneverket med sikte på mer systematisk vedlikeholdsstyring og vedlikeholdsarbeid. Blant annet skulle det lages en ny lærebok og utvikles bedre databaserte beslutningsstøtteverktøy. Norddal m.fl. (2012) gir en oversikt over dette arbeidet, mens Rausand og Vatn (2008) og Podofillini m.fl. (2006) behandler teorien bak. Svee (2011) er håndboka som skulle lages.

I Sverige har det vært arbeidet mye med å estimere marginale eksterne kostnader ved jernbanetrafikk. Det ligger litt på sida av vårt emne og er ikke tatt med her. Men vi nevner Öberg og Andersson (2009), som er en gjennomgang av kostnadene som nedbrytningen av sporet gir opphav til.

Ferreira og Murray (1997) er en oversikt over nedbrytning, vedlikeholdsstyring og optimering fra et australsk perspektiv. Simson m.fl. (2000) redegjør for et vedlikeholdsplanleggingsystem med samfunnsøkonomisk innretning. Det gjelder ett objekt, ingen budsjettbegrensning, men en nedre grense for funksjonell tilstand, og med mange alternative tiltak. De tar hensyn til at det både finns planlagt og ikke-planlagt vedlikehold.

Lake m.fl. (2000, 2002) tar spesielt for seg utskifting av sviller. Higgins (1998) og Higgins m.fl. (1999) er en operasjonsanalytisk modell for å minimere summen av forsinkelser forårsaket av vedlikehold og forsinkelser forårsaket av manglende vedlikehold.

Melching og Liebman (1988) handler om avhengighet mellom vedlikeholdsprosjekter, i et nettverk der noen lenker må være brukbare for at det skal ha noen hensikt å vedlikeholde andre lenker. Jovanovic (2004) redegjør for et generisk verktøy som tillater brukerne å definere segmenter, velge parametre til nedbrytingsfunksjonene og stille kvalitetskrav.

Peng m.fl. (2011) etablerer og løser en stor optimeringsmodell for vedlikeholdsplanlegging på kort sikt, der det gjelder å minimere summen av reisekostnadene for vedlikeholdsarbeiderne (eller kostnaden ved flytting av maskiner og utstyr) og kostnaden for forstyrrelser i trafikken. Modellen tar hensyn til når de enkelte strekningene er tilgjengelige for vedlikehold, hvilke tiltak som gjensidig utelukker hverandre, og hvilke som er avhengige av tiltak andre steder for å få effekt.

Zhang m.fl. (2013) behandler et tilsvarende kortsiktig planleggingsproblem med vekt på at ikke-planlagt vedlikehold er mye dyrere enn planlagt. Furuya og Madanat (2012), derimot, omhandler en modell for å finne optimale vedlikeholdsstrategier på lengre sikt. Hva mer er: Den tar fullt hensyn til ulikhetene mellom objektene og eventuelle avhengighetsforhold mellom dem når det gjelder vedlikeholdet.

Sjøl om de fleste artiklene her ikke handler om vedlikeholdsstrategier på langt sikt, viser de alle noe av det som særpreger optimalt vedlikehold på jernbane i forhold til vedlikehold på veg. Det finnes jo få omkjøringsmuligheter i nettet, slik at noen strekninger blir avgjørende viktige for at systemet som helhet skal fungere. Samtidig blir det vanskeligere å kombinere vedlikehold med å opprettholde trafikken. I tillegg er toleransen for ulykker og uhell svært liten og kravene til pålitelighet store. Alt dette innebærer at andre mål har forrang for samfunnsøkonomi, eller sagt på en mer operasjonsanalytisk måte: Maksimeringen av samfunnsøkonomisk overskudd må foregå under mange ulike bibetingelser (i tillegg til budsjettbetingelsen).

En annen implikasjon av avhengighetene og de strenge standardkravene er at tilstanden til objektet eller komponenten enten er brukbar eller ikke brukbar, slik at teori for sannsynligheten for at en komponent skal svikte blir viktigere enn teori der objektet kan ha mange mer eller mindre akseptable tilstander. Zhao m.fl. (2006) er et eksempel.

5.8 Bru

Nedbrytningen av bruer handler i stor grad om miljøets påvirkning på betongen. Det avgjørende er også her sikkerheten, dvs. sannsynligheten for at en viktig del av strukturen skal svikte. Så lenge sannsynligheten for at det skal skje er ubetydelig, er brua brukbar. Og så lenge den er brukbar, er brukerkostnadene konstante. Abu-Tair m.fl. (2002) behandler nedbrytningen av betongen i bruer som en stokastisk prosess, og forventet levetid som følge av det. Verma m.fl. (2014), som er tilfeldig valgt av oss som ikke har nærmere kjennskap til emnet, gir en brei oversikt over nyere litteratur på området.

Bolukbasi m.fl. (2004, 2006) vurderer nedbrytingsfunksjoner og prognoser for framtidig tilstand for mange typer av komponenter i bruer, både før og etter rekonstruksjon. Andre liknende arbeider er Lee (2007). Lee m.fl. (2007) bruker en avansert algoritme for å supplere mangelfulle opplysninger om resultatene av tidligere inspeksjoner, med sikte på å lage bedre nedbrytningsfunksjoner.

Al-Subhi m.fl. 1990 er et tidlig eksempel på en modell som bruker heltallsprogrammering til å optimere vedlikeholdet av mange bruer under en budsjettrestriksjon. De mulige tiltakene er nybygging, rehabilitering og større vedlikeholdstiltak. En annen modell for optimering av bruvedlikehold finnes i Bocchini og Frangopol (2011).

Ryall (2001) er en håndbok i bruvedlikehold. En omfattende gjennomgang av alle typer av nedbrytningsmodeller, tiltak og optimeringsverktøy på området er gjort i doktorgradsarbeidet til Robelin (2006). Professor Madanat var med i bedømmelseskomiteen hans.

5.9 Flyplass

Gendreau og Soriano (1998) er en nokså gammel gjennomgang av vedlikeholdsstyringssystemer for rullebanedekket på flyplasser. Lange (2010) er den nye norske håndboka på området.

5.10 Kommersielle vedlikeholdsstyringssystemer (PMS)

Her noterer vi bare noen referanser til de store kommersielle vedlikeholdsstyringssystemene. Ingen av dem har oss bekjent skaffet seg gode optimeringsmoduler ennå, men de kan sikkert brukes som

utgangspunkt for optimering med ett eller annet av programmene vi har nevnt.

FHWA (2002) gjør greie for programmet HERS-ST. (Det finnes nyere versjoner.) Kerali m.fl. (2006) gjør greie for HDM-4, som sannsynligvis er det mest utbredte. Mannisto og Tapio (1994) dokumenterer det finske styringssystemet HIPS for vedlikehold av vegdekker og fremmer forslag om integrering av styringssystemet for vegdekker og bruer til et felles system, IMS.

En kort og instruktiv innføring i utviklingen av datasystemer for planlegging av vedlikehold er kapittel 1 og 2 i Chootinan m.fl. (2006). Framstillingen gjelder vedlikehold av vegdekker. Oppgava er å legge opp vedlikeholdet av et stort antall objekter over flere år *på en måte som sikrer funksjonaliteten* uten å bruke for mye penger. Det var på området vegdekker at slike datasystemer først blei utviklet. Men framstillingen har overføringsverdi til andre typer av objekter.

Zoeteman and Esveld (2004) gir en oversikt over alle de store programmene som er i bruk. SVV (2007) gir en kortfattet oversikt over programmer som er vurdert for norske forhold.

Vedlikeholdsstyringssystemet for bruer som er i bruk i Norge, Pontis, er dokumentert i Golabi og Shepherd (1997).

5.11 Konklusjon

Gjennomgangen, spesielt av litteraturen om vegdekker, har tydelig vist at det har skjedd mye på disse fagfeltene på de siste ti årene. Det er nå trolig realistisk å finne optimale vedlikeholdsstrategier for nettverk med mye flere objekter enn tidligere. Optimeringsmodulene i kommersiell programvare har inntil det siste ikke vært i stand til å ta hensyn til budsjettbegrensninger på riktig måte, men vi nærmer oss kanskje et punkt der dette endrer seg. Da vil økonomiske analyser av vedlikehold kunne gjennomføres i mye større utstrekning, og på langt mer detaljert nivå, enn i dag.

Men enn så lenge vil slike analyser kreve betydelige kunnskaper i matematikk, operasjonsanalyse, programmering og økonomi av de som skal utføre dem. Forhåpentligvis vil det være mulig å bygge opp denne kunnskapen i norske miljøer.

Gjennomgangen har også bekreftet at det ennå ikke er veldig mange modeller som tar hensyn til budsjettbegrensninger på en riktig måte. Vårt eget arbeid fra 2008 og 2011 står seg fremdeles godt, men vi har også funnet ideer i litteraturen som kan brukes til å forbedre det. Det finnes miljøer både i USA og Asia som tydeligvis arbeider med samme slags modeller som vi, og som kanskje arbeider kontinuerlig med det, der vi bare har drevet sporadisk.

Av litteraturen om vedlikehold på jernbanen har vi etter hvert skjönt at modeller på det området nødvendigvis må inkludere flere bibetingelser enn på veg. Der finnes det litt mindre av økonomiske avveininger og litt mer av absolutt krav, for eksempel til sikkerhet og pålitelighet. Noe tilsvarende gjelder bru.

I alle modeller vil det være usikkerhet om nedbrytningen og virkningen av tiltakene, og om de empiriske dataene som er samlet inn om tilstanden. Det vil også blant annet være usikkerhet om de framtidige budsjettene. I noen tilfeller kan det vises at dette ikke påvirker den optimale løsningen vesentlig. I andre tilfeller lar det seg gjøre å finne ut hvor usikker den er. Flere forhold taler for at det er bedre å vurdere usikkerheten i etterkant, og oppdatere modellen når nye data foreligger, enn å bygge usikkerheten inn i modellen.

Referanser

- Ahuja R.K., T. L. Magnanti, J. B. Orlin. *Network flows: theory, algorithms, and applications*. Englewood Cliffs, New Jersey: Prentice-Hall, 1993.
- Bertsekas D. P. *Dynamic Programming and Optimal Control, Vol. I*. Belmont, Massachusetts: Athena Scientific, 2000.
- Dahl G., H. Minken. Towards sound methods of Cost-Benefit Analysis of road network rehabilitation strategies. *Nordic Road and Transport Research*, no 2 2007.
- Dahl G., H. Minken. *Methods based on discrete optimization for finding road network rehabilitation strategies*, Computers and Operations Research 35, (2008), 2193–2208.
- Labi S., K. C. Sinha. Life-Cycle Evaluation of Flexible Pavement Preventive Maintenance. *ASCE Journal of Transportation Engineering* 2005; 131 (10): 744–751.
- Li Y., S. Madanat. A steady-state solution for the optimal pavement resurfacing problem. *Transportation Research Part A* 2002; 36: 525–535.
- Minken H., G. Dahl, C. Steinsland. Samfunnsøkonomisk vurdering av vedlikeholdsstrategier, oppgradering og standardutforming i vegnettet. TØI-rapport 957/2008.
- Minken H., S.F. Meyer, K. Veisten, Y. Bai. Samfunnsøkonomisk analyse av vedlikehold - hva trengs i etatene? TØI-rapport 1185/2011
- Mosteller, F. A k -Sample Slippage Test For an Extreme Population *The Annals of Mathematical Statistics*, 1948, Vol 19, No 1: 58–65.
- Nelder, John A., R. Mead. *A simplex method for function minimization*. Computer Journal 7, (1965): 308-313.
- Nemhauser G.L., L. Wolsey. *Integer Programming and Combinatorial Optimization*. New York: Wiley, 1988.

A. Vedlegg A

DYNOPT – documentation

Dette avsnittet er skrevet på engelsk; delvis fordi koden og kommentarer i den er på engelsk.

In this chapter the main structure and some details of DYNOPT are described. The emphasis will be on main variables, and the interaction and main purpose of the various functions that constitute DYNOPT. For all details we refer to the code itself.

Main structure

The program is written in MATLAB and consists of several MATLAB functions, each in a separate file.

The overall structure is specified in the main program `Main.m` where three functions are called (invoked):

<code>InputData.m:</code>	specifies the input data to the problem
<code>Optimize.m:</code>	performs the optimization and solves the problem
<code>OutputResults.m:</code>	presents the solution found and various results

Each of these functions call other functions, and these again call some functions. This will be explained later.

The program sends output to the MATLAB command window. This includes information on which of the main functions that is active and also reports on the progress of optimization. Finally a brief report on the solution is given. The active function is indicated like this: ... Optimize

All other information should be self-explained by giving text, name of variables etc, and such information is given without the leading dots. In addition, the session is written to a datafile called `datafile.txt`: this includes the input data and the optimization results.

Global variables

The program contains a number of global variables, i.e., variables that may be accessed and altered in all those functions in which they are declared as global variables. These variables contain information about the state transition graphs, and the size of the system under consideration. Using global variables may not be the best programming “ethics”, but it is certainly useful. In the code the name of global variables is always written using capital letters.

NTIME	Number of time steps
NSTATE	Number of states
NCOMP	Number of components
NINVEST	Vector: Number of investment levels for each component
NEDGES	Vector: Number of edges (state transitions) for each component
GRAPH_I	Vector: GRAPH_I(cp,k) is initial vertex (state) of edge number k for component cp
GRAPH_J	Vector: GRAPH_J(cp,k) is initial vertex (state) of edge number k for component cp
GRAPH_S	Vector: GRAPH_S(cp,k) is (type 1) cost of edge number k for component cp
GRAPH_Sinvest	Vector: GRAPH_Sinvest(cp,k) is rehabilitation cost (type 2 cost) of edge number k for component cp
INITSTATE	Vector: INITSTATE(cp) is the initial state for component cp
TERMSTATE	Vector: TERMSTATE(cp) is the terminal state for component cp
BUDGET	Vector: BUDGET(t) is the budget in time period t
OPTGAP	optimization parameter: optimality gap stopping criterion
MAXITER	optimization parameter: maximum number of Lagrange iterations

The remaining part of this chapter describes the various MATLAB functions of DYNOPT in more detail. MATLAB functions are written like this: `ReachableStates`.

Input and output functions

There are some functions for communication between DYNOPT and the user of the program. Here they are:

InputData:

- Input: none
- Returns: none
- Purpose: Sets up the problem to be studied. Calls, depending on a variable `pr` (set in this function), either `TestCase1`, `TestCase2` or `TestCase3`. Calls `ReachableStates`. Writes the problem data to file `datafile.txt`.

ReachableStates:

- Input: `cp` (component)
- Returns: 1 if `TERMSTATE` was modified, 0 otherwise.
- Purpose: Determines if the given terminal state of comp. `cp` (`TERMSTATE(cp)`) can be reached. If not it finds the “nearest” state that can be reached, and redefines `TERMSTATE` accordingly.

TestCase1:

- Input: none
- Returns: none
- Purpose: Sets up a very small test example with only 2 states and 2 components.

TestCase2:

- Input: none
- Returns: none

- Purpose: Sets up a “reasonable” test example with specified size (NSTATE etc.), random graph, specified number of rehabilitation levels, degrading and costs. Here the state space is assumed “linear”.

TestCase3:

- Input: none
- Returns: none
- Purpose: Sets up a random state graph and random costs; not realistic.

OutputResults:

- Input: none
- Returns: none
- Purpose: Prints all the results of the computations to the file `results.txt`.

Optimization functions

Optimize:

- Input: none
- Returns: none
- Purpose: Main function for the optimization algorithm. Runs several iterations in which lower and upper bounds on the optimal value are computed, and gradually improved using Lagrangian relaxation and the subgradient method. The budget constraints are relaxed and the Bellman Ford algorithm is used repeatedly, for each component. Calls the functions `BellmanFord`, `Heuristic1`, `BellmanFord`, `ConstrainedBellmanFord`. A central variable is `X_PATH` which represents the current solution: it is a matrix with a row for each component `cp`, and this row contains the edges in the path for `cp`. A path complete describes the development the states and therefore the rehabilitation strategy for `cp`. A parameter is `MaxIterations` which is the maximum number of Lagrangian iterations (updates of the dual variables). It seems that around 50 is a good value, but a higher value may decrease the optimality gap.

BellmanFord:

- Input: `cp` (component), `Lambda` (Lagrangian cost vector)
- Returns: `Cost(i,t)` (optimal cost to state `i`, time `t`), `Epath` (edges in the shortest path found)
- Purpose: computes the shortest path in the “dynamic graph” obtained from the state graph. Uses the Bellman Ford algorithm (dynamic programming).

SubgradientMethod:

- Input: `Lambda`, `Surplus`, `InvestCost`, `it`
- Returns:
- Purpose: `Lambda` is the dual variable (associated with the relaxed budget constraints, one for each time period), and the subgradient method iteratively adjusts this vector to a near-optimal dual solution. In the adjustment the `Surplus` vector is used, saying how much the current relaxed solution (computed by Bellmann-Ford) violates the budget constraint. If large violation, the corresponding `Lambda` is reduced. If small violation: `Lambda` is increased. There is a step length parameter which is geometrically decreased in each iteration (ρ^{it}), where `it` is iteration count.

Heuristic1:

- Input: X, SCost
- Returns: X1 totalSCost] = Heuristic1(X,SCost)
- Purpose: Adjusts input solution X into a new solution X1 which is feasible, i.e., satisfies the budget constraints. The components are randomly ordered, and for each cp, in this order, the best path (strategy sequence) is chosen using Bellman-Ford and where only edges satisfying the “rest budget constraint” are permitted. Note: there is a potential problem in that there may (towards the end) be no feasible path to the selected terminal state (without violating the rest budget constraint). If this happens, an error message is printed to screen and the program stops.

Vedlegg B: Litteratur som det er vist til i kapittel 5

OBS. Litteraturen som det er vist til i kapittel 1-4 er ikke tatt med her, men finnes i litteraturlista bakerst i hovedteksten.

Abu-Tair, A.I., C. McParland, J.F. Lyness and A. Nadjai (2002) Predictive models of deterioration rates of concrete bridges using The Factor Method based on historic inspection data. Paper 262, Proceedings of the 9th Durability of Building Materials and components (DBMC 02), Brisbane, Australia.

Al-Subhi, K.M., D.W. Johnston and F. Farid (1990) Resource-Constrained Capital Budgeting Model for Bridge Maintenance, Rehabilitation and Replacement. *Transportation Research Record* **1268**.

Andersson, M., J. Nyström, K. Odolinski, L. Wieweg och Å Wikberg (2010) En strategi för utveckling av en samhällsekonomisk analysmodell for drift, underhåll och reinvestering av transportinfrastruktur. VTI 2010-12-10.

Aurstad, J. (2011) (red.) Lærebok, drift og vedlikehold av veger. Statens vegvesen Vegdirektoratet. VD rapport nr. 53.

Baltzer, S. and M. Mollerup (2013) 3xP Nordic User's Guide. NordFoU-PPM.

Ben-Akiva, M., F. Humplick, S. Madanat and R. Ramaswami (1993) Infrastructure management under uncertainty: Latent performance approach. *Journal of Transportation Engineering* **119**(1), 43-58.

Ben-Akiva, M. and R. Ramaswamy (1993) An approach for predicting latent infrastructure facility deterioration. *Transportation Science* **27**(2), 134-153.

Ben-Akiva, M. and D. Gopinath (1995) Modeling infrastructure performance and user costs. *Journal of Infrastructure Systems* **1**(19), 33-43.

Bocchini, P. and D.M. Frangopol (2011) A probabilistic computational framework for bridge network optimal maintenance scheduling. *Reliability Engineering and System Safety* **96**, 332-349.

Bolukbasi, M., J. Mohammadi and D. Arditi (2004) Estimating the future condition of highway bridge components using national bridge inventory data. *Practice Periodical on Structural Design and Construction* **9**(1).

Bolukbasi, M.M., D. Arditi and J. Mohammadi (2006) Deterioration of reconstructed bridge decks. *Structure and Infrastructure Engineering: Maintenance, Management, Life-Cycle Design and Performance* **2**(1), 23-31.

Bryce, J.M., G.W. Flintsch, S.W. Katicha, B.K. Diefenderfer (2012) Developing a Network-Level Structural Capacity Index for Asphalt Pavements. *Journal of Transport Engineering* **139**(2), 123-129.

- Bryce, J.M., G.W. Flintsch, S.W. Katicha, B.K. Diefenderfer (2013) Developing a Network-Level Structural Capacity Index for Structural Evaluation of Pavements. Virginia Center of Transportation Innovation and Research. Final Report VCTIR 13-R9.
- Carnahan, J. (1988) Analytical framework for optimizing pavement maintenance. *Journal of Transportation Engineering* **114**(3), 307-322.
- Chan, W.T., T.F. Fta and C.Y. Tan (1994) Road maintenance planning using genetic algorithms. I: Formulation. *Journal of Transportation Engineering* **120**(5), 693-709.
- Chan, W.T., T.F. Fta and C.Y. Tan (2003) Optimal fund-allocation analysis for multi-district highway agencies. *Journal of Infrastructure Systems* **9**(4), 167-175.
- Chootinan, P., A. Chen, M. Horrocks and D. Bolling (2006) A multi-year pavement maintenance program using a stochastic simulation-based genetic algorithm approach. *Transportation Research A* **40**(9), 725-743.
- Christiansen, A.S., M.L. Holst and S. Baltzer (2010) Development of performance measures, modeling and calibration. NordFOU-PPM, Report no. 2.
- Chu, C.-Y. and P.L. Durango-Cohen (2008a) Incorporating Maintenance Effectiveness in the Estimation of Dynamic Infrastructure Performance Models. *Computer-Aided Civil and Infrastructure Engineering* **23**, 174-188.
- Chu, C.-Y. and P.L. Durango-Cohen (2008b) Estimation of dynamic performance models for transportation structure using panel data. *Transportation Research B* **42**, 57-81.
- Chu, J.C. and Y.-J. Chen (2012) Optimal threshold-based network level transportation infrastructure life-cycle management with heterogeneous maintenance actions. *Transportation Research B* **46**, 1123-1143.
- Dahl, G. and H. Minken (2008) Methods based on discrete optimization for finding road network rehabilitation strategies. *Computers and Operations Research* **35**(7), 2193-2208.
- Daniels, L.E. (2008) Track maintenance costs on rail transit property. TCRP Web-only document 43.
- Dekker, R. and R.E. Wildeman (1997) A Review of Multi-component Maintenance Models with Economic Dependence. *Mathematical Methods of Operations Research* **45**, 411-435.
- Durango, P.L. and S. Madanat (2002) Optimal maintenance and repair policies in infrastructure management under uncertain facility deterioration rates: an adaptive control approach. *Transportation Research A* **36**, 763-778.
- Durango-Cohen, P. and P. Sarutipand (2007) Capturing interdependencies and heterogeneity in the management of transportation infrastructure systems. *Journal of Infrastructure Systems* **13**(2), 115-123.
- Durango-Cohen, P.L. and N. Tadeballi (2006) Using advanced inspection policies to support investments in maintenance and repair of transportation infrastructure facilities. *Journal of Transportation Engineering*, **January 2006**, 60-68.

- Durango-Cohen, P.L. and S. Madanat (2008) Optimization of inspection and maintenance decisions for infrastructure facilities under performance model uncertainty: A quasi-Bayes approach. *Transportation Research A* **42**, 1074-1085.
- Dye Management Group, Inc. (2010) Development of Levels of Service for the Interstate Highway System. NCHRP Report 677, Transportation Research Board.
- Esveld, C. (2001) Modern railway track. Second edition. MRT-Productions.
- Evensen, R. (2006) Vegdekker. Grunnlag for innretting av vegdekkevedlikehold. *Sandvika: ViaNova Plan og Trafikk AS*.
- Federal Highway Administration (2002) HERS-ST 2.0. Highway Economic Requirement System – State Version. Overview..
- Federal Highway Administration (2013) About LTTP.
<http://www.fhwa.dot.gov/research/tfhrp/programs/infrastructure/pavements/ltpp/index.cfm>
- Fernandez, E. and T. Friesz (1981) Influence of demand-quality interrelationships on optimal policies for stage construction of transport facilities. *Transportation Science* **15**(1), 16-31.
- Ferreira, L. and M.H. Murray (1997) Modelling railtrack deterioration and maintenance: current practices and future needs. *Transport Reviews* **17**(3), 207-221.
- Flora, W.F., G.P. Ong and K.C. Sinha (2010) Development of a Structural Index as an Integral Part of the Overall Pavement Quality in the INDOT PMS. JTRP Technical Report 2010/11. Purdue e-pubs.
- Friesz, T. and E. Fernandez (1979) A model of optimal transport maintenance with demand responsiveness. *Transportation Research B* **13**(4), 317-339.
- Furuya, A. and S. Madanat (2012) Accounting for Network Effects in Railway Asset Management. *Journal of Transport Engineering* **139**(1), 92-100.
- Gao, L. (2011) Optimal Infrastructure Maintenance Scheduling Problem under Budget Uncertainty. PhD dissertation, University of Texas at Austin.
- Gao, L. and Z. Zhang (2011) Optimal Infrastructure Maintenance Scheduling Problem under Budget Uncertainty. Southwest Region University Transportation Center, Report No. SWUTC/11/161028-1,
- Gao, L. and Z. Zhang (2012) Approximation approach to problem of large-scale pavement maintenance and rehabilitation. *Transportation Research Record* **2304**.
- Gao, L., J.P. Aguiar-Moya and Z. Zhang (2011) Performance Modeling of Infrastructure Condition Data with Maintenance Intervention. *Transportation Research Record* **2225**, 109-116.
- Gendreau, M. and P. Soriano (1998) Airport pavement management systems: An appraisal of existing methodologies. *Transportation Research A* **32**(3), 197-214.
- Golabi, K., R. Kulkarny and G. Way (1982) A statewide pavement management system. *Interfaces* **12**(6), 5-21.

- Golabi, K. and R. Shepherd (1997) Pontis: A system for maintenance optimization and improvement of US bridge network. *Interfaces* **27**(1).
- Gu, W., Y. Ouyang and S. Madanat (2012) Joint optimization of pavement maintenance and resurfacing planning. *Transportation Research B* **46**, 511-519.
- Guillaumot, V., P. Durango-Cohen og S. Madanat (2003) Adaptive optimisation of infrastructure maintenance and inspection decisions under performance model uncertainty. *Journal of Infrastructure Systems* **9**(4), 133-139.
- Higgins, A. (1998) Scheduling of railway track maintenance and crews. *Journal of the Operational Research Society* **49**(10), 1026-1033.
- Higgins, A., L. Ferreira and M. Lake (1999) Scheduling rail track maintenance to minimize overall delays. Proceedings of the 14th international symposium on transportation and traffic theory, Jerusalem, Israel.
- Highway Research Board (HRB) (1962) The AASHO Road Test. Report 5: Pavement Research, Special report 61E.
- Jacobson, T. og L.-G. Wågberg (2007) Utveckling och uppgradering av prognosmodell för beläggnings slitage från dubbade däck samt en kunskapsöversikt över inverkan faktorer. VTI notat 7-2007.
- Jovanovic, S. (2004) Railway track quality assessment and related decision making. In: 2004 IEEE International Conference on Systems, Man and Cybernetics.
- Karabakal, N., J.C. Bean and J.R. Lohmann (1994) Scheduling pavement maintenance with deterministic deterioration and budget constraints. Technical Report 94-18, College of Engineering, University of Michigan.
- Kerali, H.G.R., J.B. Odoki and E.E. Stannard (2006) Overview of HDM-4. The Highway Development and Management Series, Volume 1. Association mondiale de la route (AIPCR) / World Road Association (PIARC), Paris.
- Labi, S. and K.C. Sinha (2003) The effectiveness of maintenance and its impact on capital expenditures. Paper 208, Joint Transportation Research Program, Purdue University.
- Labi, S. and K.C. Sinha (2005) Life-cycle evaluation of flexible pavement preventive maintenance. *Journal of Transportation Engineering* **131**(10), 744-751.
- Lake, M.R., L. Ferreira and M.H. Murray (2000) Cost-Benefit Analysis of Sleeper Replacement Strategies: Simulation Model. *Transportation Research Record* **1713**, 15-21.
- Lake, M.R., L. Ferreira and M.H. Murray (2002) Using Simulation to Evaluate Rail Sleeper Replacement Alternatives. *Transportation Research Record* **1785**, 58-63.
- Lamprey, G., S. Labi and Z. Li (2008) Decision support for optimal scheduling of highway pavement preventive maintenance within resurfacing cycle. *Decision Support Systems* **46**, 376-387.
- Lang, J. og N. Grivet (2013) Kalibrering av HDM-4 – Nedbrytningsmodeller og effektmodeller. WSP Sverige AB.

- Lange, G. (2010) Håndbok for vedlikehold og rehabilitering av flyplassdekker. SINTEF, ATI, Forsvarsbygg. ID-nummer AV-H-U007.
- Larsen, H.J.E., G. Hildebrand and R.A. McDonald (2002) Economic evaluation of pavement maintenance. Report 114 2002, Danish Road Institute.
- Lee, J. (2007) A methodology for developing bridge condition rating models based on limited inspection records. PhD Dissertation, Griffith University, Australia.
- Lee, J., K. Le, Y.-C. Loo, M. Blumenstein and H. Guan (2007) ANN-based bridge condition rating model using limited structural inspection records. 10th Conference on Science and Technology, Ho Chi Minh City University of Technology.
- Lee, J. and S. Madanat (2014) Joint optimization of pavement design, resurfacing and maintenance strategies with history-dependent deterioration models. *Transportation Research B* **68**, 141-153.
- Lerfald, BO (2010) Et mer funksjonelt vegnett i øst-vestretning i Jämtland-Trøndelag. Faktorer som har betydning for bæreevnen. Rapport SBF IN A10003, SINTEF Byggforsk.
- Lerfald, B.O., S. Bruaset, L.S. Hafskjold, I.-L. Solberg, K. Holmøy, K.L. Moe, K. Heilemann, D. Berthelsen og I. Hoff (2008) SIP – Future Rehabilitation Strategies for Physical Infrastructure. Status for infrastruktur i Norge. Rapport SBF IN A08016, SINTEF Byggforsk.
- Li, Y. and S. Madanat (2002) A steady state solution for the optimal pavement resurfacing problem. *Transportation Research A* **36**(6), 525-535.
- Madanat, S. og M. Ben-Akiva (1994) Optimal inspection and repair policies for infrastructure facilities. *Transportation Science* **28**(1), 55-61.
- Madanat, S., R. Mishalani and W.W. Ibrahim (1995) Estimation of infrastructure transition probabilities from condition rating data. *Journal of Infrastructure Systems* **1**(2), 120-125.
- Madanat, S., J.A. Prozzi and M. Han (2002) Effect of Performance Model Accuracy on Optimal Pavement Design. University of California Transportation Center.
- Mannisto, V. and R. Tapio (1994) Infrastructure management system: Case study of the Finnish Road Administration. *Transportation Research Record* **1455**.
- Markow, M. and W. Balta (1985) Optimal rehabilitation frequencies for highway pavements. *Transportation Research Record* **1035**, 31-43.
- Martland, C.D., S. McNeil, D. Acharya, R. Mishalani, J. Eshelby (1990) Applications of expert systems in railroad maintenance: Scheduling rail relays. *Transportation Research A* **24**(1), 39-52.
- Medury, A. and S. Madanat (2013) Incorporating network considerations into pavement management systems: A case for approximate dynamic programming. *Transportation Research C* **33**, 134-150.
- Melching, C.S. and J.S. Liebman (1988) Allocating railroad maintenance funds by solving binary knapsack problems with precedence constraints. *Transportation Research B* **22**(3), 181-194.

- Minken, H., S.F. Meyer, K. Veisten og Yu Bai (2011) Samfunnsøkonomisk analyse av vedlikehold – hva trengs i etatene? TØI-rapport 1185/2011
- Mishalani, R. and S. Madanat (2002) Computation of infrastructure transition probabilities using stochastic duration models. *Journal of Infrastructure Systems* **8**(4), 139-148.
- Ng, M.W., D. Lin and S.T. Waller (2009) Optimal long-term infrastructure maintenance planning accounting for traffic dynamics. *Computer-Aided Civil and Infrastructure Engineering* **24**(7), 459-469.
- Ng, M.W., Z. Zhang and S.T. Waller (2011) The price of uncertainty in pavement infrastructure management planning: An integer programming approach. *Transportation Research C*, **19**(6), 1326-1338.
- Nicolai, R.P. (2008) Maintenance Models for Systems subject to Measurable Deterioration. PhD Thesis, Erasmus Universiteit Rotterdam.
- Nicolai, R.P. and R. Dekkert (2006) Optimal Maintenance of Multi-Component Systems: A review. Chapter 11 in: Kobbacy and Murthy (eds.) *Complex System Maintenance Handbook*. Springer Verlag.
- Norddal, P.K., H. Svee and H.K. Endresen (2012) RAMS implementation for railways. Proceedings of the 21st International Congress on Maintenance and Asset Management.
- NordFou (2006) Pavement Performance Models. Network level analysis. Development of performance measures, modelling and calibration. <http://www.nordfou.org/projekter/nedbrydningsmodellFlexBefaest.ast>
- Ohlmann, J.W. and J.C. Bean (2009) Resource-constrained management of heterogeneous assets with stochastic deterioration. *European Journal of Operational Research* **199**(1), 198-208.
- Olsson, N. og M. Veiseth (2011) Jernbanetrafikk. Tapir akademisk forlag.
- Ouyang, Y. and S. Madanat (2004) Optimal scheduling of rehabilitation activities for multiple pavement facilities: Exact and approximate solutions. *Transportation Research A* **38**(5), 347-365.
- Ouyang, Y. and S. Madanat (2006) An analytic solution for the finite-horizon pavement resurfacing planning problem. *Transportation Research B* **40**(9), 767-778.
- Paterson, W.D.O. (1990) Quantifying the effectiveness of pavement maintenance and rehabilitation. Proceedings at the 6th REAAA Conference, Kuala Lumpur, Malaysia.
- Peng, F., S. Kang, X. Li, Y. Ouyang, K. Somani and D. Acharya (2011) A Heuristic Approach to the Railroad Track Maintenance Scheduling Problem. *Computer-Aided Civil and Infrastructure Engineering* **26**(2), 129-145.
- Podofillini, L., E. Zio and J. Vatn (2006) Risk-informed optimization of railway tracks inspection and maintenance procedures. *Reliability Engineering and Systems Safety* **91**(1), 20-35.

- Rausand, M. and J. Vatn (2008) Reliability Centred Maintenance. Chapter 4 in: Kobbacy and Murthy (eds.) *Complex Systems Maintenance Handbook*, Springer Verlag.
- Rashid, M.M. and K. Tsunokawa (2012) Trend curve optimal control model for optimizing pavement maintenance strategies consisting of various treatments. *Computer-Aided Civil and Infrastructure Engineering* **27**(3), 155-169.
- Robelin, C.A. (2006) Facility-Level and System-Level Stochastic Optimization of Bridge Maintenance Policies for Markovian Management Systems. PhD Dissertation, University of California Transportation Center.
- Ryall, M.J. (2001) Bridge management. Butterworth Heinemann, Oxford.
- Saba, G., A. Huvstig, G. Hildebrand, E. Sund, R. Evensen, H. Sigursteinsson and J. Elsander (2006) Performance prediction models for flexible pavements: a state-of-the-art report. Teknologirapport nr. 2477, Teknologivdelingen, Statens vegvesen, Vegdirektoratet, Oslo.
- Sathaye, N. and S. Madanat (2011) A bottom-up solution for the multi-facility optimal resurfacing problem. *Transportation Research B* **45**(7), 1004-1017.
- Sathaye, N. and S. Madanat (2012) A bottom-up optimal pavement resurfacing solution approach for large scale networks. *Transportation Research B* **46**, 520-528.
- Seyedshohadaie, S.R., I. Damjanovic and S. Butenko (2010) Risk-based maintenance and rehabilitation decisions for transportation infrastructure networks. *Transportation Research A* **44**, 236-248.
- Simson, S.A., L. Ferreira and M.H. Murray (2000) Rail track maintenance planning: An assessment model. Transportation Research Record 1713.
- Sinha, K.C and T.F. Fwa (1989) On the total highway management. *Transportation Research Record* **1229**.
- Skar, A., M.L. Holst, J. Larsen, S. Baltzer and M. Mollerup (2012) Validation of performance models. Report no. 1 from the NordFoU-PPM 2 project.
- Small, K. A. and C. Winston (1988) Optimal Highway durability. *American Economic Review* **78**(3), 560-569.
- Small, K. A., Winston, C., and Evans, C. A. (1989). Road Work; A new highway pricing and investment policy. The Brookings Institution.
- Smilowitz og Madanat (2000) Optimal inspection, maintenance and rehabilitation policies for networks of infrastructure facilities under measurement and forecasting uncertainty. *Computer Aided Civil Infrastructure Engineering* **15**(1), 5-13.
- Statens vegvesen, teknologiavdelingen (2007) Vurdering av EDB-system for beregning av nedbrytning av veg. Rapport nr. 2520.
- Sund, E.K. (2004) Strategiske analyser av dekkevedlikehold på riksvegnettet, SINTEF rapport STF22 A04333.
- Sund, E.K. (2005) Vegkapital delprosjekt VEG. Tilstand/tilstandsutviklingsmodeller. Delprosjekt A5: Effekt av vedlikeholdstiltak. SINTEF-notat.

- Svee, H. (2011) Håndbok for vedlikehold. Jernbaneverket. Dokument-ID: STY-601058
- Tsunokawa, K. and J.L. Schofer (1994) Trend curve optimal control model for highway pavement maintenance: case study and evaluation. *Transportation Research A* **28**(2), 151-166.
- Tsunokawa, Koji and Dinh Van Hiep (2008) A unified optimization procedure for road asset management. 6th ICPT, Sapporo, Japan July 2008.
- Verma, S.K., S.S. Badauria and S. Ahktar (2014) Probabilistic evaluation of service life for reinforced concrete structures. Chinese Journal of Engineering Article ID 648438, <http://dx.doi.org/10.1155/2014/648438>.
- Wang, I.L., Y.C.J. Tsai and F. Lee (2011) A network flow model for clustering segments and minimizing total maintenance and rehabilitation costs. *Computers and Industrial Engineering* **60**(4), 593-601.
- Wågberg, L.-G. (2001) Utveckling av nedbrytningsmodeller – Sprickinitiering och sprickpropagering. VTI Meddelande 916. Statens väg- och transportforskningsinstitut (VTI), Linköping.
- Yeo, H., Y. Yoon and S. Madanat (2013) Algorithms for bottom-up maintenance optimization for heterogeneous infrastructure systems. *Structure and Infrastructure Engineering*, **9**(4), 317-328.
- Zhang, T., J. Andrews and R. Wang (2013) Optimal Scheduling of Track Maintenance on a Railway Network. *Quality and Reliability Engineering International* **29**(2), 285-297.
- Zhao, J., A.H.C. Chan, C. Roberts and A.B. Stirling (2006) Assessing the economic life of rail using a stochastic analysis of failures. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* **220**, 103-111.
- Zoeteman A. and C. Esveld (2004) State of the art in railway maintenance management: Planning systems and their application in Europe. *2004 IEEE International Conference on Systems, Man and Cybernetics* **5**, 4165-4170.
- Öberg, J. and E. Andersson (2009) Determining the deterioration cost for railway tracks. *Journal of Rail and Rapid Transit*, **223**(2), 121-129.
- Öberg, G. (red.) (2001) Statligt belagda vägar – tillståndet på vägytan och i vägkroppen, effekter och kostnader. VTI Notat 44-2001, Statens väg- och transportforskningsinstitut (VTI), Linköping.

Transportøkonomisk institutt (TØI)

Stiftelsen Norsk senter for samferdselsforskning

TØI er et anvendt forskningsinstitutt, som mottar basisbevilgning fra Norges forskningsråd og gjennomfører forsknings- og utredningsoppdrag for næringsliv og offentlige etater. TØI ble opprettet i 1964 og er organisert som uavhengig stiftelse.

TØI utvikler og formidler kunnskap om samferdsel med vitenskapelig kvalitet og praktisk anvendelse. Instituttet har et tverrfaglig miljø med rundt 70 høyt spesialiserte forskere.

Instituttet utgir tidsskriftet Samferdsel med 10 nummer i året og driver også forskningsformidling gjennom TØI-rapporter, artikler i vitenskapelige tidsskrifter, samt innlegg og intervjuer i media. TØI-rapportene er gratis tilgjengelige på instituttets hjemmeside www.toi.no.

TØI er partner i CIENS Forskningscenter for miljø og samfunn, lokalisert i Forskningsparken nær Universitetet i Oslo (se www.ciens.no). Instituttet deltar aktivt i internasjonalt forsknings-samarbeid, med særlig vekt på EUs rammeprogrammer.

TØI dekker alle transportmidler og temaområder innen samferdsel, inkludert trafiksikkerhet, kollektivtransport, klima og miljø, reiseliv, reisevaner og reiseetterspørsel, arealplanlegging, offentlige beslutningsprosesser, næringslivets transport og generell transportøkonomi.

Transportøkonomisk institutt krever opphavsrett til egne arbeider og legger vekt på å opptre uavhengig av oppdragsgiverne i alle faglige analyser og vurderinger.

Besøks- og postadresse:

Transportøkonomisk institutt
Gautstadalléen 21
NO-0349 Oslo

22 57 38 00
toi@toi.no
www.toi.no