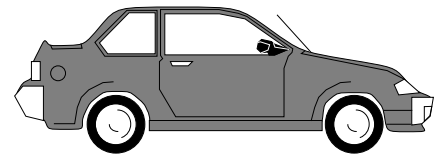
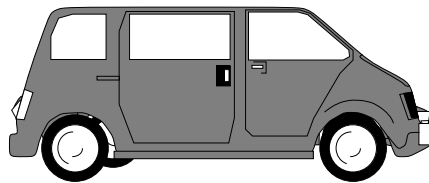


# Ny metodikk for trafikkberegninger på grunnlag av telling

## Kalibrert for kortidstillinger i Oslo



STAT/07/97

Magne Aldrin  
Turid Follestad

Oslo  
oktober 1997

**Tittel/Title:** Ny metodikk for trafikkberegninger på grunnlag av tellinger - Kalibrert for korttidstillinger i Oslo

**Dato/Date:** oktober

**År/Year:** 1997

**Notat nr:**

Note no: STAT/07/97

**Forfatter/Author:** Magne Aldrin, Turid Follestad

**Sammendrag/Abstract:**

Rapporten beskriver en ny metode for å estimere trafikkmengden på en vei der det er utført tellinger. Trafikkmengden estimeres for et år som helhet, angitt ved årsdøgntrafikk (ÅDT), og for hver enkelt time i året. Om det finnes telldata hvor bilene er inndelt i klasser (biltype eller lengdeklasser), kan beregningene gjøres for hver klasse av biler. Metoden gir også anslag for usikkerheten i ÅDT-estimatene. Videre gir metoden grunnlag for å estimere trafikkarbeid i et geografisk område med et tilhørende usikkerhetsanslag. Dette kan brukes til å optimalisere telleopplegget ved å avveie i hvilken grad større veier bør telles i lengre tidsperioder enn mindre veier, og hvilke ukedager og når på døgnet det bør telles. Metoden er kalibrert for korttidstillinger i Oslo. Metodens anvendelses område kan utvides til langtidstillinger og til andre områder ved rekalkulering av endel sentrale parametre.

Arbeidet er utført på oppdrag fra Statens Forurensningstilsyn. Trafikkdataene som arbeidet bygger på er levert av Statens Vegvesen Oslo, Vegtrafikksentralen.

**Emneord/Indexing terms:** Trafikkberegninger, ÅDT-beregninger, Redusert rang regresjon

**Målgruppe/Target group:**

**Tilgjengelighet/Availability:** Åpen

**Prosjektdata/Project data:**

**Prosjektnr/Project no:**

**Antall sider/No of pages:**

**Satsningsfelt:**

Research field:

---

## Innhold

<b>Del 1: Overordnet beskrivelse</b> .....	1
1 Innledning .....	1
2 Bakgrunn.....	2
3 Ny metode for trafikkberegninger basert på kombinasjon av basiskurver.....	3
3.1 Beregning av trafikkmengde ut fra tellinger .....	4
3.2 Usikkerhet i ÅDT-estimat .....	12
3.3 Beregning av trafikkarbeid med usikkerhet .....	12
4 Sammenligning av presisjon til ny og nåværende metode.....	13
<b>Del 2: Detaljert framstilling 15</b> .....	15
5 Beskrivelse av metoden basert på kombinasjon av basiskurver .....	15
5.1 Etablering av basiskurver.....	15
5.2 ÅDT-beregninger basert på et gitt antall basiskurver .....	16
5.3 Valg av antall basiskurver .....	18
5.4 Usikkerhet i ÅDT-estimat .....	21
5.5 Beregning av trafikkarbeid i et område, med usikkerhetsanslag .....	23
5.6 Omregning mellom inndeling av biler etter biltype og etter lengde .....	25
6 Datagrunnlag.....	25
6.1 Oversikt over data .....	25
6.2 Prosedyre for korrigering av data.....	28
7 Beregninger basert på et simuleringseksperiment .....	30
7.1 Prediksjon basert på basiskurver.....	31
7.2 Etablering av ligninger for valg av antall basiskurver .....	32
7.3 Etablering av ligninger for beregning av usikkerhet.....	33
7.4 Beregning av korrelasjoner .....	33
7.5 Sammenligning med nåværende metodikk basert på variasjonskurver .....	34
8 Mulige forbedringer .....	38
Referanser: .....	40
Appendix A Trekking av talletidspunkter.....	41

# Del 1: Overordnet beskrivelse

Rapporten er delt i to hoveddeler. Del 1 (kapittel 1 - kapittel 4) gir en oversikt over bruk av og egenskaper med den nye metoden for trafikkberegninger. Del 2 (kapittel 5 - kapittel 8) beskrives der nye metoden mer grundig. Denne beskrivelsen er ment å være såpass detaljert at den kan danne grunnlag for en implementering av metoden, og vi har derfor funnet det fornuftig å gi en relativt detaljert og teknisk framstilling. Del 2 inneholder også en oversikt over datagrunnlaget, og hvordan feil i dataene er korrigert.

## 1 Innledning

På veinettet i Norge foregår det en rekke trafikktellinger. For endel av tellepunktene er tellingene i prinsippet kontinuerlige (men kan ha manglende data), for andre telles det kun i korte perioder. Ut fra trafikktellingene gjøres det ulike typer beregninger. En av de viktigste er ÅDT (trafikk pr. døgn i gjennomsnitt over et år) på det enkelte tellepunkt. For å beregne ÅDT brukes pr. i dag de målte dataene sammen med et sett såkalte variasjonskurver, som angir trafikkvariasjon over året.

På oppdrag fra Statens Forurensningstilsyn har Norsk Regnesentral (NR) i et par år vært engasjert for å utarbeide en ny og bedre metode for trafikkberegninger (Aldrin 1995b), som på sikt vil kunne erstatte metoden basert på variasjonskurver. Vi har utarbeidet en helhetlig metodikk for trafikkberegninger i Norge som kan

- estimere trafikk på timesbasis et helt år
- beregne ÅDT med usikkerhetanslag som avhenger av veistørrelse, antall tellesteder og tellemønster
- beregne trafikkarbeid med usikkerhetsanslag
- optimalisere telleopplegget før tellinger gjøres
- utføre alle beregninger for biler i 5 lengdeklasser
- gi en forståelse av trafikkmønsteret
- brukes til å detektere feil i datagrunnlag

Så langt er metodikken kalibrert på trafikkdata for Oslo og med sikte på kortidstelling (mindre enn 2 uker). Vi antar den har et noe større anvendelsesområde (f.eks. storbyer, og tellinger opp til 2 måneder), men for at metoden skal brukes generelt må den rekalibreres på representative data fra hele Norge, og for både korte og lange tellinger. Metoden er god nok til å tas i bruk nå, men enkelte detaljer bør på sikt forbedres. Videre bør metoden utvides til å kunne beregne dimensjonerende time, og mer generelt en frekvensfordeling av trafikkmengde pr. time. Viktige forbedringer på kort sikt omfatter dermed

- implementering
- rekalkulering av metoden på representative data fra Norge
- kalibrering for generelt telleomfang, dvs. også for lengre telleperioder
- egne basiskurver for totalt antall biler og for tunge biler
- endel ulike forbedringer av detaljer i eksisterende metodikk
- utvidelser til å omfatte bl.a. beregning av dimensjonerende time

Disse omhandles litt mer detaljert i kapittel 8.

Før vi går videre vil vi presisere hva vi mener med enkelte begreper: a) Med usikkerhet mener vi her estimeringsusikkerhet som skyldes at vi ikke har observert trafikken hele året, mens usikkerhet som skyldes målefeil (feilregistreringer) vil komme i tillegg. b) En veilenke er en retningsbestemt forbindelse mellom to knutepunkt, og to parallelle veilenker i motsatt retning utgjør et veilenkepar. TrafikkteLLinger kan kun foretas i et punkt på veilenka, men vi anser det som representativt for trafikken på hele veilenka. Vi vil derfor ofte bruke betegnelsen veilenke i stedet for punkt. c) I dette arbeidet vil minste tidsenhet for tellinger og estimerer være en time, og med begrepet antall talletimer vil vi mene antall timer med tellinger.

## 2 Bakgrunn

Et overvåkningssystem for lokal forurensning er etablert i Oslo. Ideelt sett har systemet behov for opplysninger på timesbasis for alle veilenker. Imidlertid finnes det kontinuerlige (maskinelle) tellinger kun for et fåtall av veiene i Oslo (nivå 1 punkter, i denne studien 16 veier i 2 retninger, dvs 32 veilenker). For de fleste veilenker må man derfor bruke beregnede tall basert på tellinger i begrensede tidsperioder. For noen veier (nivå 2 punkter) gjennomføres maskinelle tellinger 2-4 uker pr. år. Det foreligger planer om å utføre tellinger for de resterende veilenkene (1000-2000 stykker), men for de aller fleste av disse vil man kun ha mulighet til å gjennomføre tellinger i korte tidsperioder, fra noen få timer til noen få døgn. Slike tellinger kan gjøres manuelt, men det er mulig at disse etterhvert kan erstattes av videotellinger.

Overvåkningssystemet for lokal luftforurensning i Oslo har behov for trafikkdata fordelt på biltype. Manuelle tellinger vil også typisk være inndelt etter biltype. For nivå 1 punktene finnes det imidlertid tellinger fordelt på fem lengdeklasser. Disse er nummerert fra 1 til 5 med følgende inndeling målt i meter: [0-5.5], [5.6-7.6], [7.7-12.4], [12.5-15.9] og [16.0 og lengre]. Det foreliggende arbeidet er basert på en slik inndeling etter lengde. I tillegg er det behov for en omregningsmetode mellom inndeling i lengde og biltype.

Hovedprinsippene i dagens metode for beregning av ÅDT er beskrevet i "Håndbok 146 Trafikkberegninger" utgitt av Statens Vegvesen, kapittel B2 og B5. Dagens metode baserer seg på bruk av variasjonskurver for trafikk over et år, en uke og et

døgn, samt en eventuell justering for spesielle dager. I Aldrin og Follestad (1993) er det beregna nye variasjonskurver for Oslo. For Oslo finnes det tre variasjonskurver for år, tre for uke, fem for hverdagsdøgn og en for helgedøgn. En begrensende faktor ved denne metoden er at det blant disse må velges kun *en* variasjonskurve for år, *en* for uke og *en* for døgn. Kurvene kan velges manuelt eller ved tilpasning til data. Videre gjelder disse kurvene for summen av alle bilklasser, som er dominert av små biler. Disse variasjonskurvene vil derfor være brukbare for lengdeklasse 0-5.5m, men passer ikke så godt for de større lengdeklassene hvor bilene som følger et annet kjøremønster.

### **3 Ny metode for trafikkberegninger basert på kombinasjon av basiskurver**

Den nye metoden tar i bruk såkalte basiskurver, som har en lignende fortolkning som variasjonskurver. Men mens den nåværende metoden bruker *en* kurve for år, *en* for uke og *en* for døgn, er den nye metoden basert på kombinasjon av flere kurver. Det har den fordelen at kompleksiteten av beregningene kan tilpasses til hvor mye data som er tilgjengelig. Hvis det er få telldata tilgjengelig brukes en enkel modell, mens en mer detaljert modell brukes hvis man har mange telletimer. Videre er modellen tilpasset hver lengdeklasse (men dette kan i prinsippet også gjøres for variasjonskurve-metoden).

Metoden er basert på at det foreligger mange telletimer for endel veilenker, og få telletimer for andre veilenker. Hovedideen i metoden er følgende: Først brukes dataene fra nivå 1 punktene til å etablere et sett med basiskurver over en tidsperiode, for eksempel fra starten av 1994 til slutten av 1996. Tidsenhet er en time. Den første basiskurven er en gjennomsnittskurve for alle nivå 1 punkter. Den er essensielt satt sammen av variasjonskurver for år, uke og døgn og tar også hensyn til en tidstrend og til spesielle dager. De neste basiskurvene er å betrakte som justeringer av den første kurven, og kurvene er ordna med synkende grad av viktighet.

Trafikken på en enkelt veilenke kan betraktes som en kombinasjon av slike basiskurver pluss et feilledd. Jo flere basiskurver som inngår, jo mindre blir feilleddet. Men for å bruke mange basiskurver trengs mye data. For en veilenke med tellinger noen få timer vil det lønne seg å bruke få kurver, kanskje en, to eller tre. For veilenker med tellinger noen uker (nivå 2 punkter) vil det lønne seg å bruke noen flere kurver. Antall kurver som bør brukes kan modelleres som en funksjon av antall tilgjengelige telletimer, av veistørrelse og av hvilken lengdeklasse det er snakk om.

ÅDT-estimatene vil selvsagt være usikre, og usikkerheten avhenger av antall telletimer, veistørrelse og lengdeklasse. De enkelte ÅDT-estimatene kan slås sammen til et estimat for trafikkarbeid i et geografisk område. Også usikkerheten i disse estimatene vil avhenge av antall telletimer, veistørrelse og lengdeklasse, og en

kan benytte dette til å optimalisere telleopplegget ved å avveie hvor mye det bør telles på store veier og hvor mye på små veier.

Metoden kan deles i følgende tre hoveddeler:

- Trafikkberegninger (kapittel 3.1), som består av a) etablering av basiskurver (detaljer i kapittel 5.1), b) ÅDT-beregninger basert på et gitt antall basiskurver (detaljer i kapittel 5.2) og c) usikkerhet i ÅDT-estimater (detaljer i kapittel 5.2)
- Usikkerhet i ÅDT-estimat (kapittel 3.2, detaljer i kapittel 5.4)
- Beregning av trafikkarbeid i et område, med usikkerhetsanslag (kapittel 3.3, detaljer i kapittel 5.5)

Dessuten er det behov for omregningsregler mellom inndeling av biler etter biltype og etter lengde. Dette har ikke vært et tema i dette prosjektet, men kapittel 5.6 og Aldrin (1995 b) beskriver et opplegg for å etablere slike omregningsregler.

### 3.1 Beregning av trafikkmengde ut fra tellinger

Når det foreligger tellinger på en veilenke beregnes antall biler pr. time ved å kombinere  $k$  basiskurver. Basiskurvene har en gitt verdi for hver time  $t$ , og de  $k$  verdiene i time  $t$  angis som  $b_{1t}$  til  $b_{kt}$ . Antall biler beregnes v.hj. a. formelen

$$\text{estimert antall biler i time } t = c \cdot \exp(\alpha_1 b_{1t} + \alpha_2 b_{2t} + \dots + \alpha_k b_{kt}), \quad (1)$$

hvor koeffisientene  $c$  og  $\alpha$ -ene beregnes ut fra de tilgjengelige tellingene. ÅDT beregnes ved å ta sum av timesestimaterne over et år og dele på antall døgn. YDT og HDT beregnes tilsvarende ved å ta med henholdsvis yrkesdøgn og helgedøgn i summen. Dette gjøres separat for hver lengdeklasse. Total ÅDT beregnes ved å summere over lengdeklassene. Uten å ha nevnt det eksplisitt, har vi antatt at alle tellinger er gjort når det er normal trafikk, og ikke når veien er stengt f.eks. p.g.a. veiarbeid. Om veien har vært stengt i kjente perioder kan timesestimaterne for disse periodene settes til 0.

Basiskurvene er beregna på forhånd ut fra nivå 1 punkter. De bør oppdateres jevnlig, f.eks. en gang i året. Basiskurvene er ordna med minkende grad av viktighet. Beregning av basiskurver og bruk av dem er nærmere beskrevet i kapittel 5.1 og kapittel 5.3.

Figur 1 viser de to første basiskurvene for hver av lengdeklassene. Hver basiskurve er her dekomponert i fire deler: i) trend, ii) sesongvariasjon, iii) spesielle dager (slik de falt i 1994) og iv) ukedag- og døgn-variasjon. Basiskurven framkommer altså som summen av disse. Kurvene er i log-skala. Da kan en differanse på for eksempel 0.1 tolkes som omtrent 10% forskjell i original skala, mens en differanse på 0.20 kan tolkes som 20% forskjell osv. Dette er fordi en differanse  $d$  i logskala blir  $e^d$  i original skala, som er tilnærma lik  $100d\%$ .

La oss betrakte dekomposisjonen av den første basiskurven for lengdeklasse 0-5.5m. Trendkurven viser en svak økning på omtrent 3% pr. år. Sesongkurven viser at trafikken typisk er lav midt på sommeren og i vintermånedene. Kurven for spesielle dager viser at det typisk er lavere trafikk i jula, påska og omkring helligdage i mai. Døgnvariasjonen er tydelig, med lav trafikk om natta. Morgen- og ettermiddags-rushet fra mandag til fredag sees også tydelig. Døgnvariasjonen er ganske lik fra mandag til fredag, og fra lørdag til søndag. Noe forskjell er det dog, og det er viktig å modellere inn all systematisk variasjon som kan tilskrives eksakt informasjon slik som ukedag.

Den andre basiskurven er ikke så tolkbar, men kan anses som en justering av den første. Om den andre basiskurven *legges til* (dvs.  $\alpha_2$  er positiv) den første vil det blant annet gi mindre trafikk i timene før morgenrushet og til en viss grad under morgenrushet, mens trafikken om ettermiddagen vil øke. Om en i stedet *trekker fra* (dvs.  $\alpha_2$  er negativ) den andre basiskurven, vil effekten være motsatt. Den tredje basiskurven utgjør en ytterligere justering osv.

Kurvene for de andre og større lengdeklassene har mange fellestrekk med kurvene for lengdeklasse 1. En vesentlig forskjell er at det ikke sees noe morgen- og ettermiddagsrush i første basiskurve, noe som er svært rimelig. Et annet trekk er at trafikkveksten (trenden i første basiskurve) er større for de tyngre lengdeklassene enn for lengdeklasse1.

Tilsammen har vi beregna 8 basiskurver, slik at antall basiskurver  $k$  i ligning (1) vil være mellom 1 og 8. Men antall basiskurver som bør brukes avhenger av antall telletimer som tilgjengelig, siden koeffisientene i ligninga bestemmes fra telldataene. Om det er talt få timer velges  $k$  liten, om det er talt mange timer velges  $k$  stor. Vi har laget en beregningsregel som velger  $k$  ut fra telleomfang, antatt trafikkmengde og lengdeklasse. Denne beregningsregelen er beskrevet nærmere i kapittel 5.4.

Telleomfang omfatter både antall telletimer og hvilke tidspunkter, og karakteriseres ut fra hvor mange timer som er talt i ulike tidsperioder som angitt i Tabell 1. Timene i et døgn er nummerert fra 1 til 24, slik at time 1 er fra kl. 0.00 til kl. 1.00 osv., slik at med timene 8-9 menes tidsperioden fra kl. 7.00 til kl. 9.00.



## Tabell 1 Karakterisering av telldata.

---

hverdag 8-9

hverdag 10-15

hverdag 16-17

hverdag 7,18-24

hverdag 1-6

lørdag 10-24

lørdag 1-9

søndag 10-24

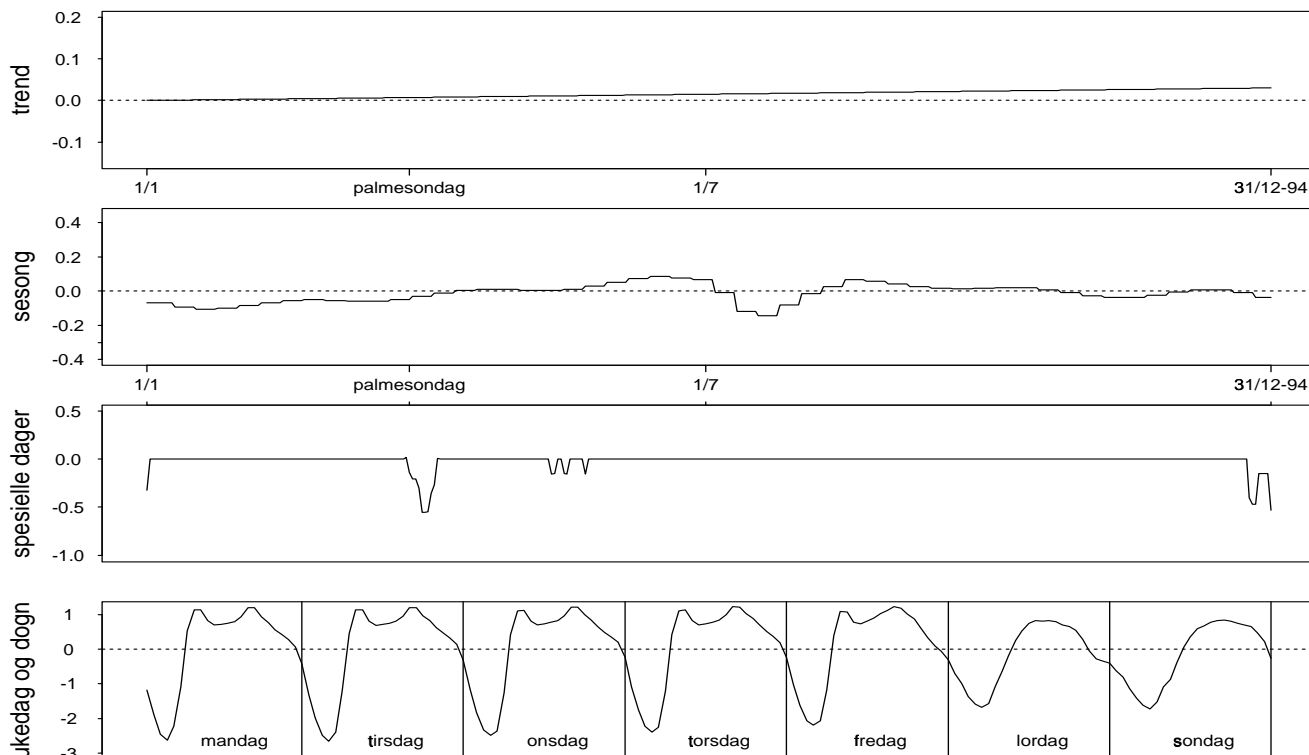
søndag 1-9

---

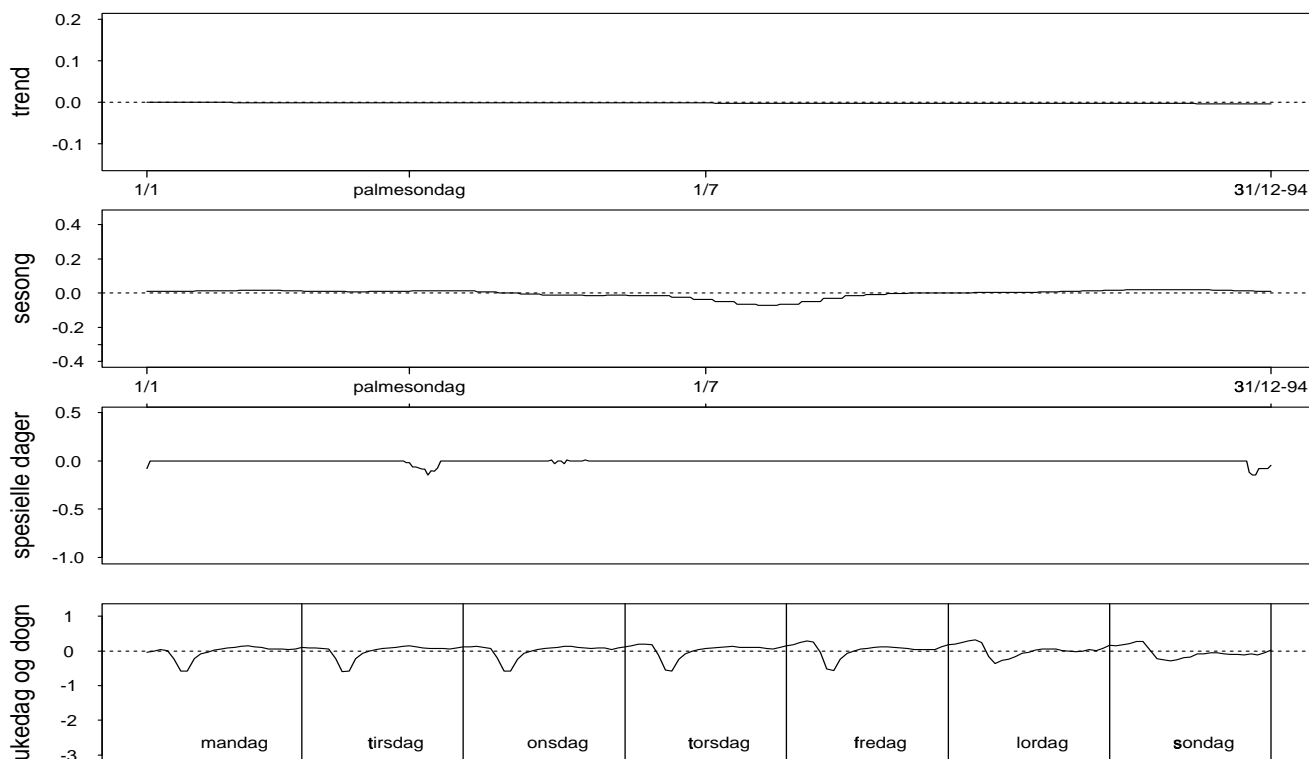
**Figur 1** 2 dekomponerte basiskurver for hver av lengdeklassene.  
a)

## 2 dekomponerte basiskurver for lengde 0-5.5 m

1. basiskurve, dekomponert



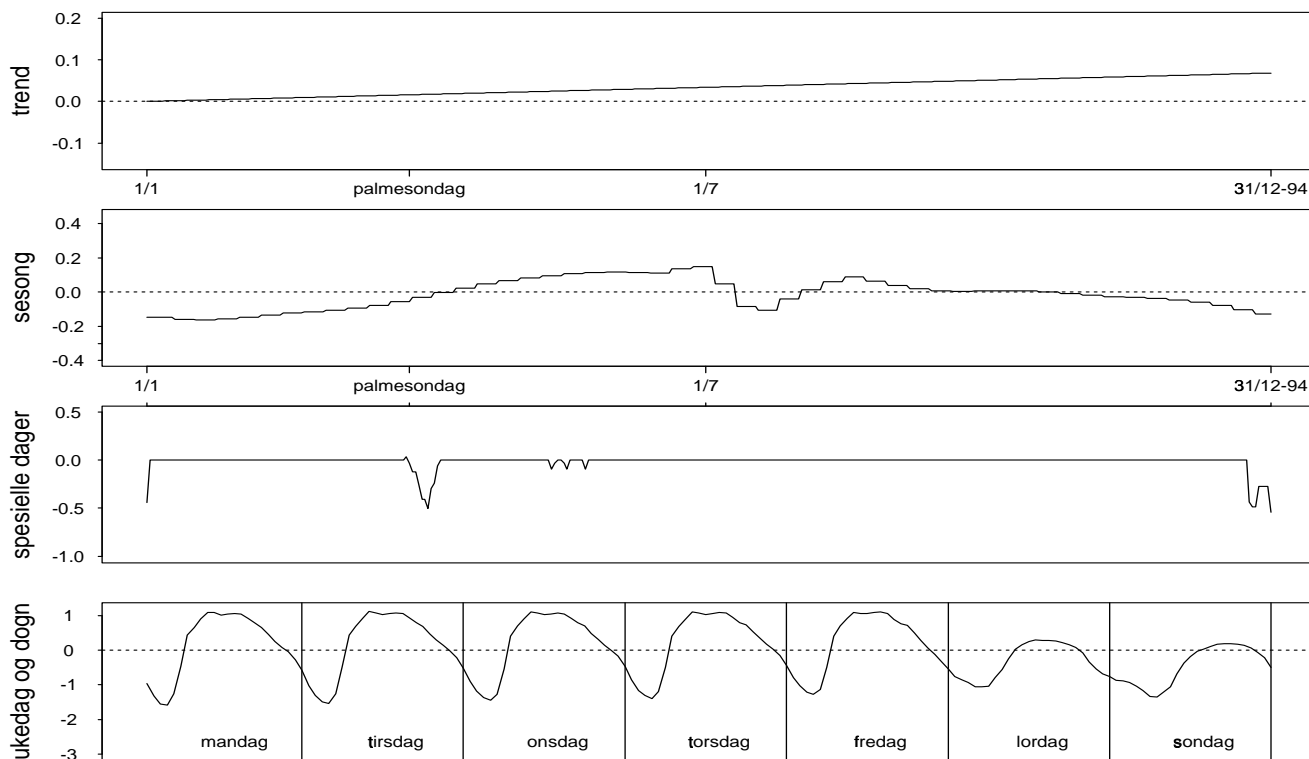
2. basiskurve, dekomponert



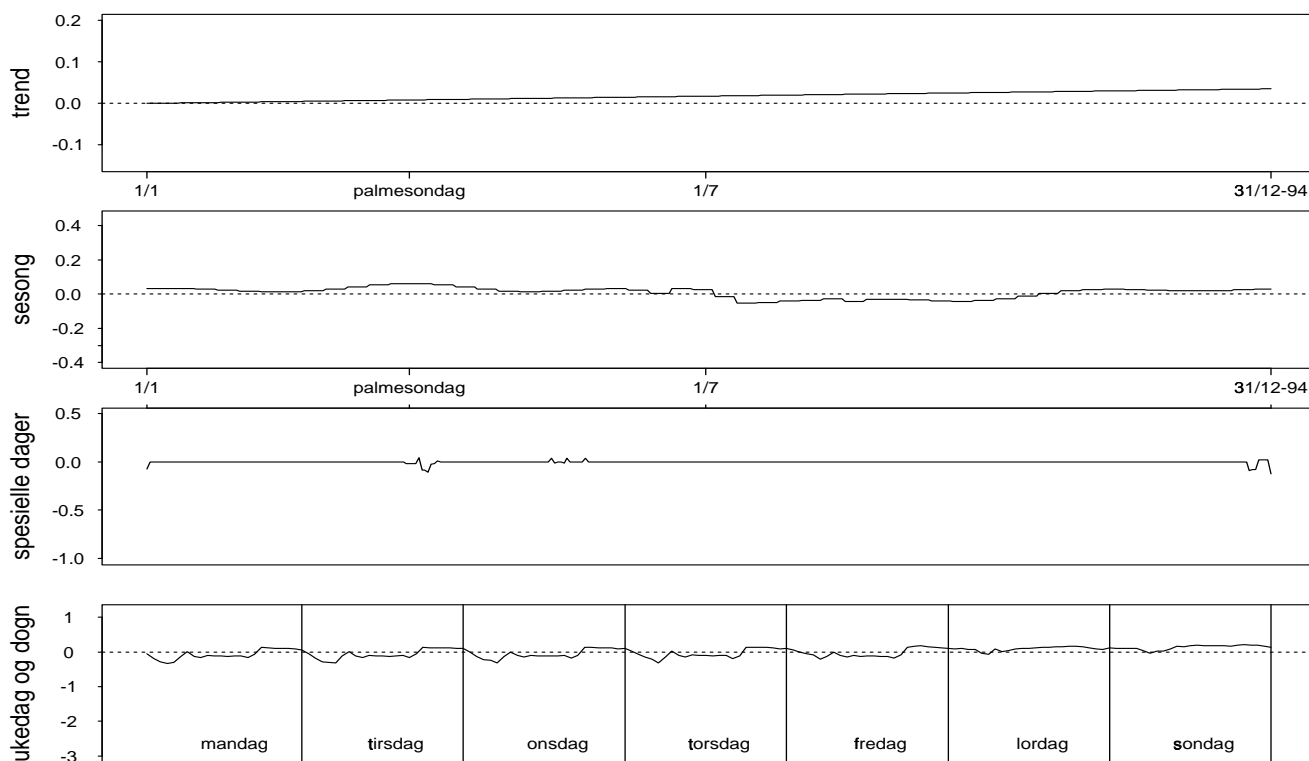
b)

## 2 dekomponerte basiskurver for lengde 5.6-7.6 m

1. basiskurve, dekomponert



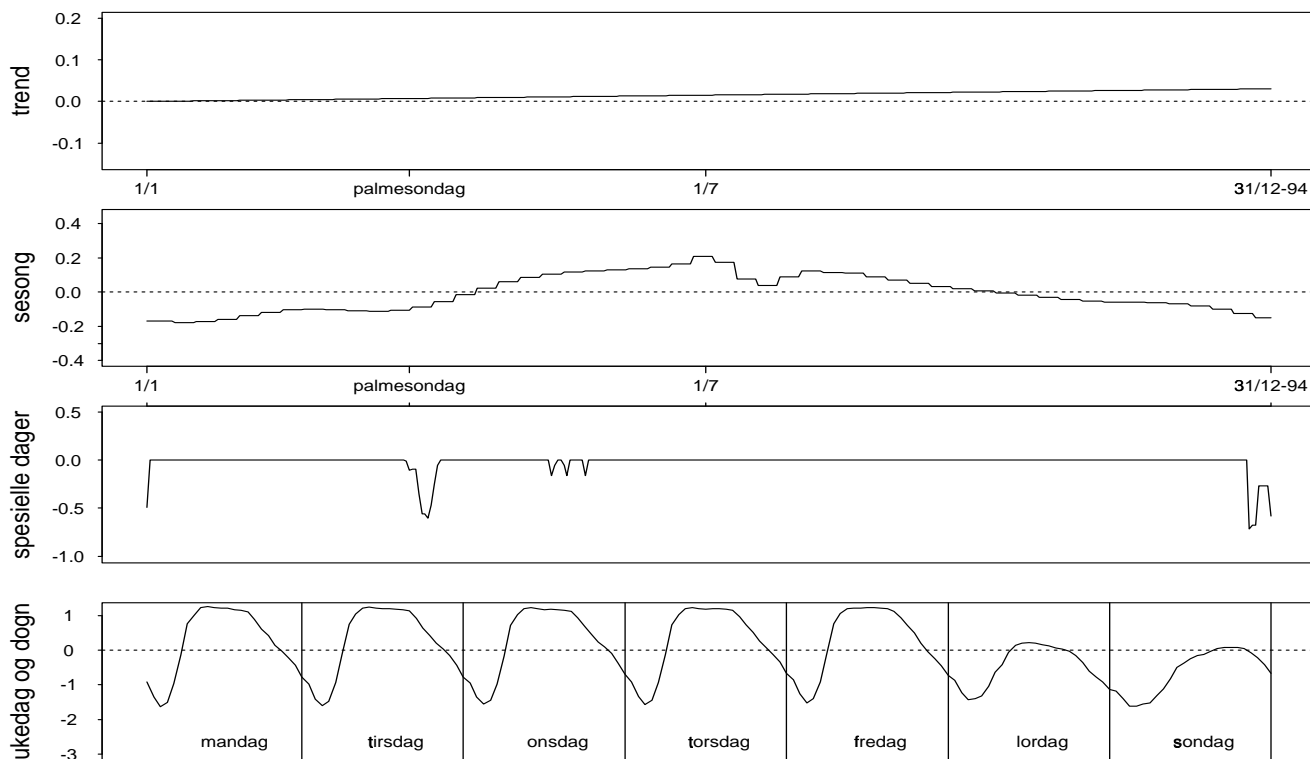
2. basiskurve, dekomponert



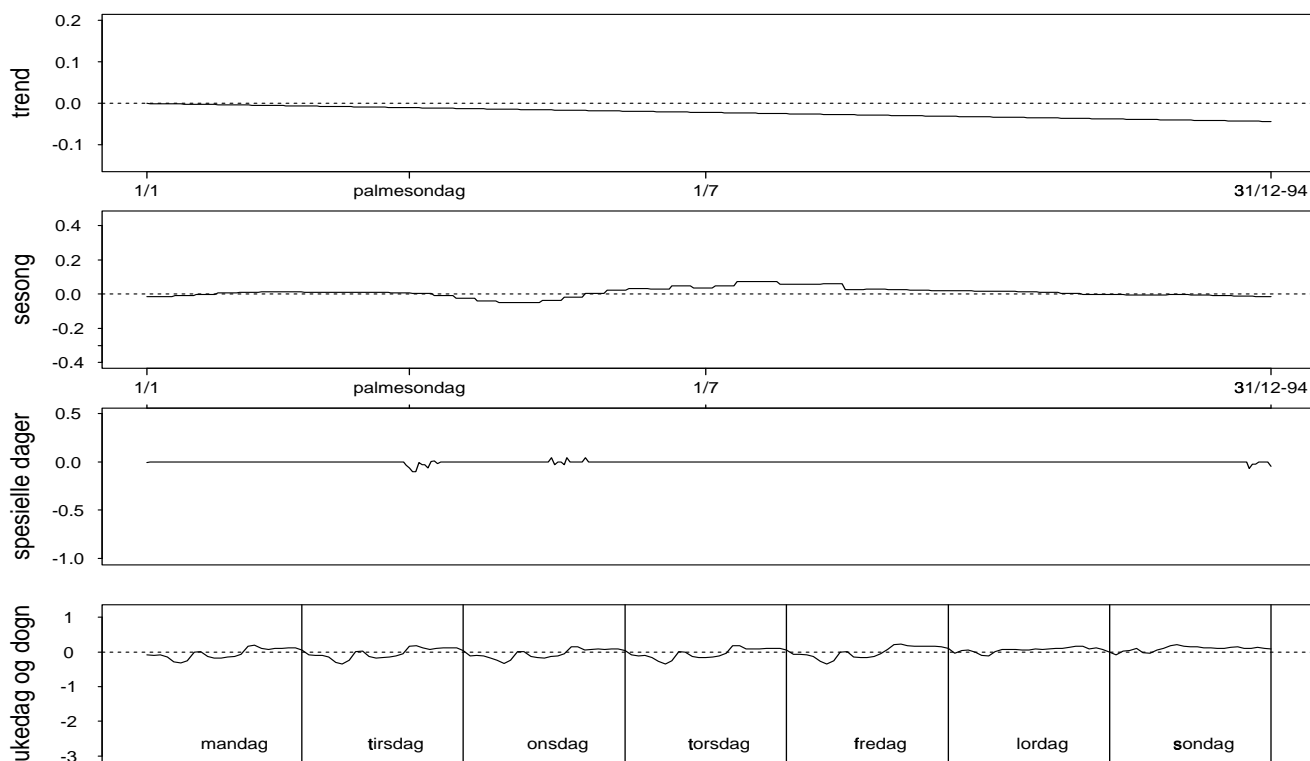
c)

## 2 dekomponerte basiskurver for lengde 7.7-12.4 m

### 1. basiskurve, dekomponert



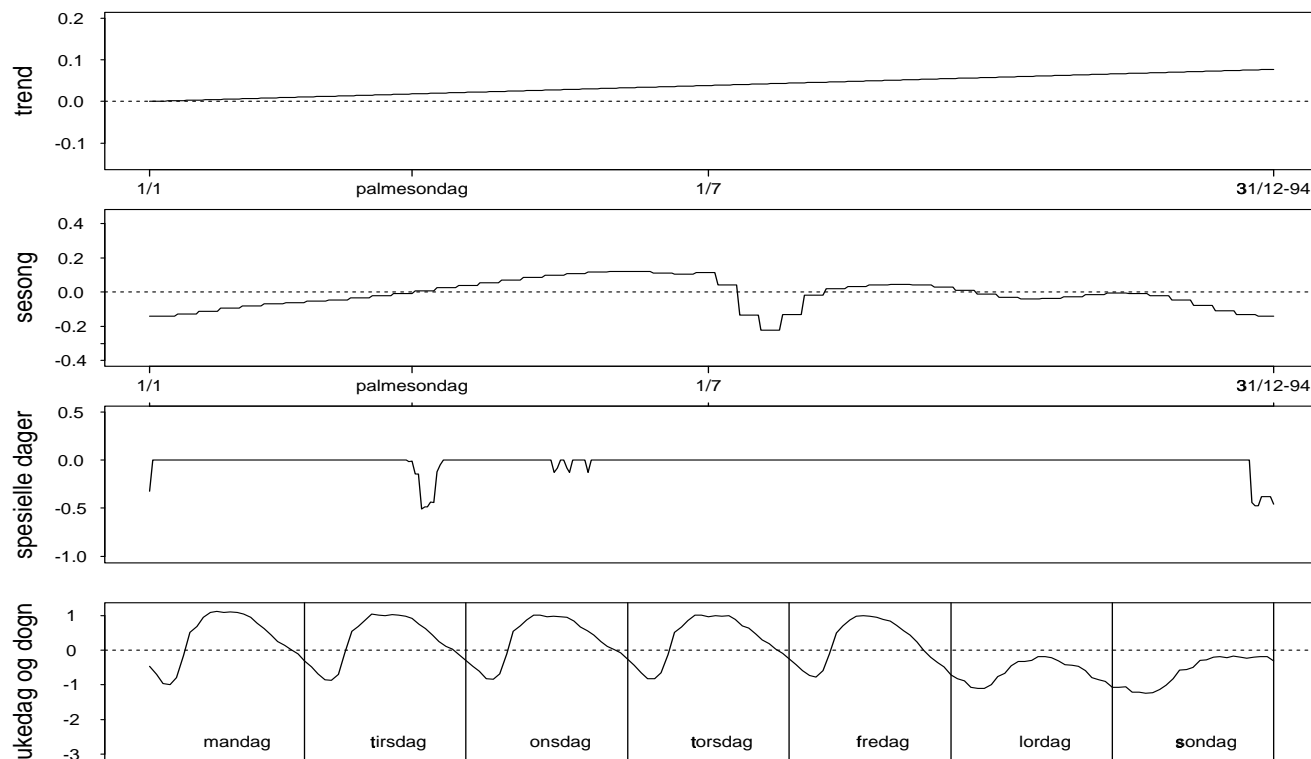
### 2. basiskurve, dekomponert



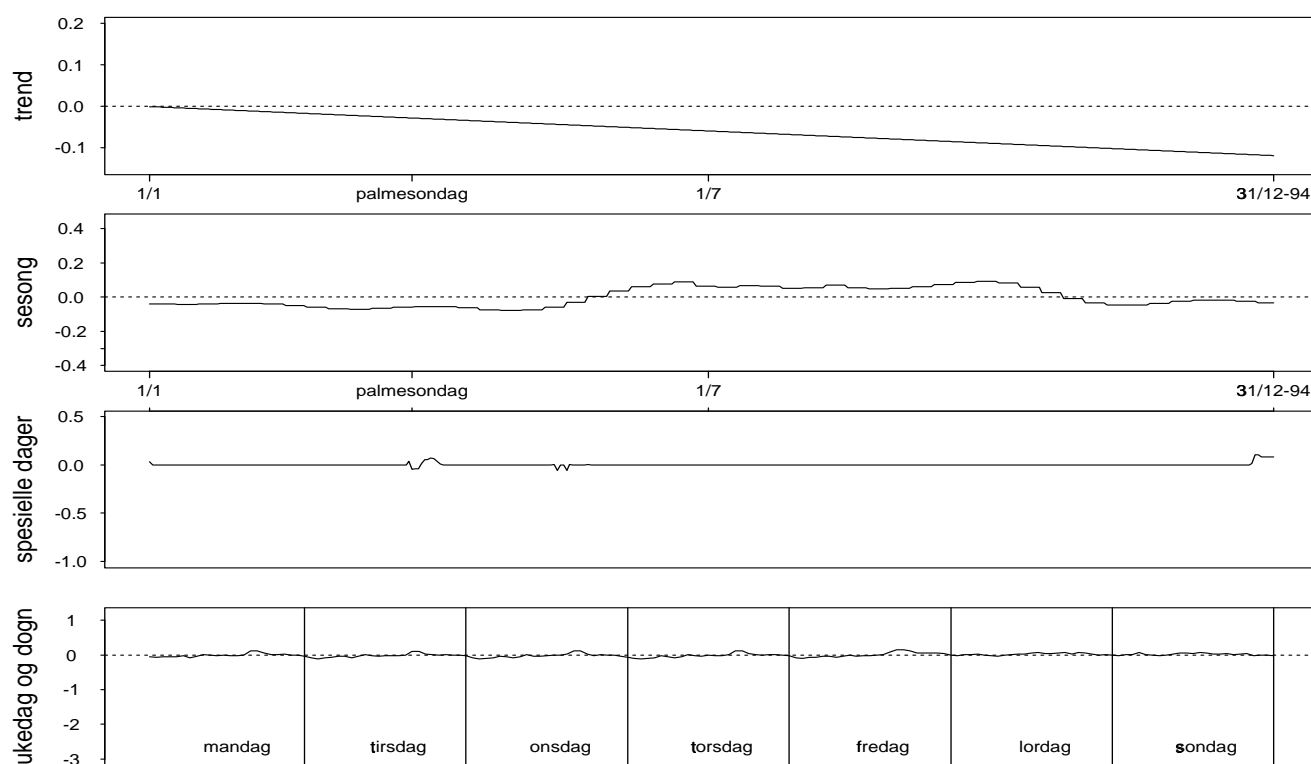
d)

## 2 dekomponerte basiskurver for lengde 12.5-15.9 m

### 1. basiskurve, dekomponert



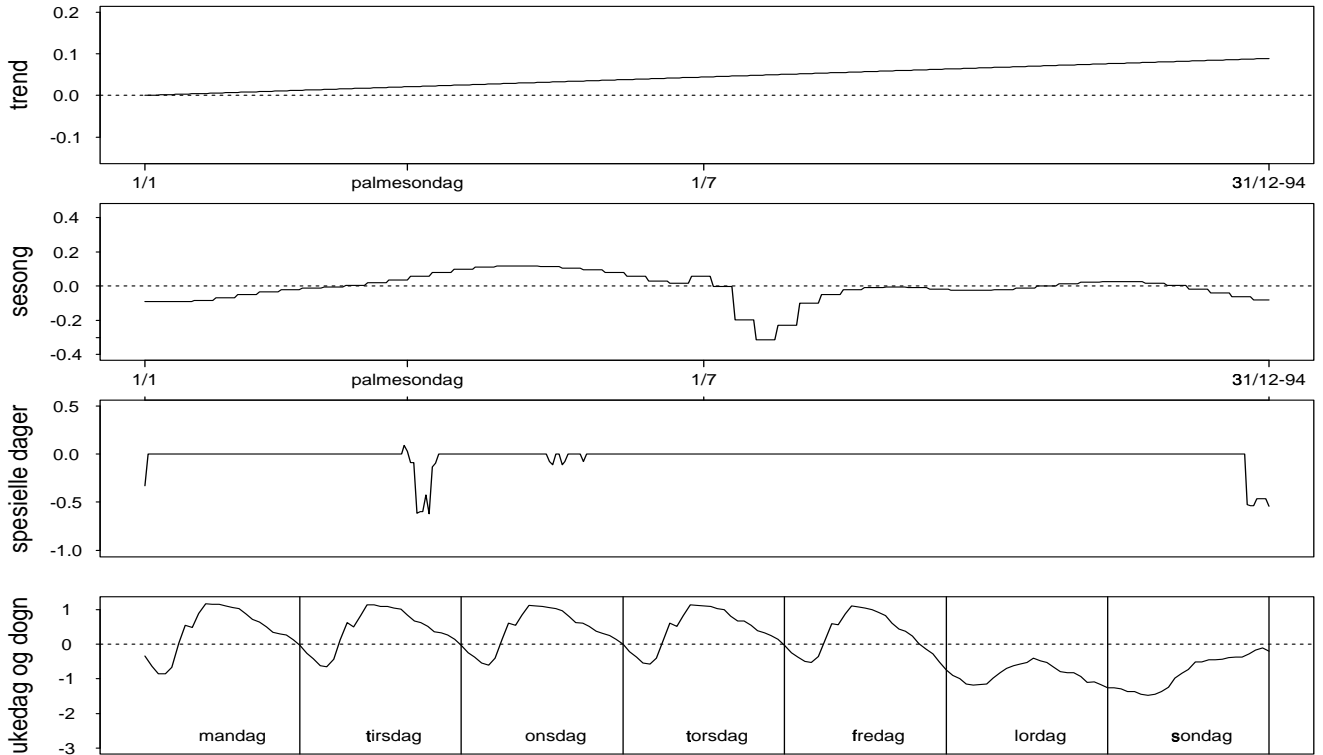
### 2. basiskurve, dekomponert



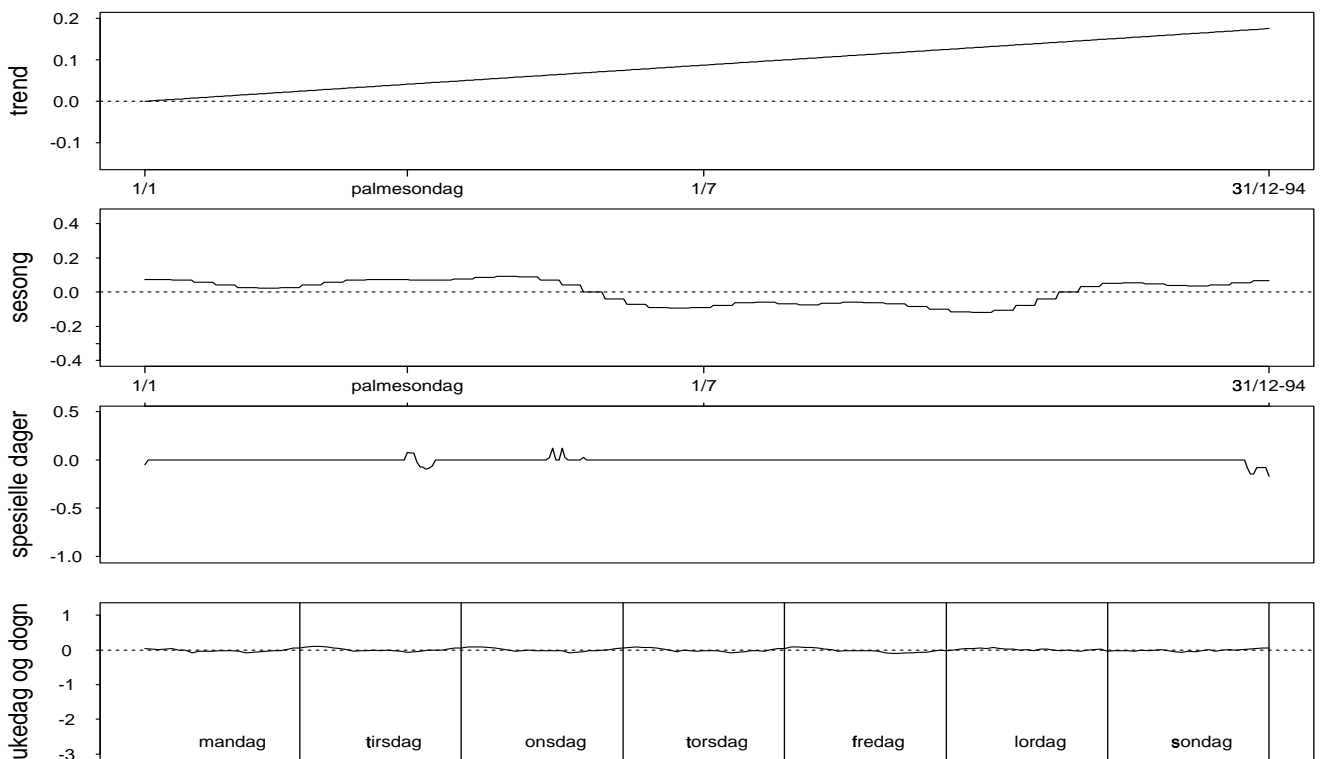
e)

## 2 dekomponerte basiskurver for lengde 16.0 m og lengre

### 1. basiskurve, dekomponert



### 2. basiskurve, dekomponert



### 3.2 Usikkerhet i ÅDT-estimat

ÅDT-estimatene vil være usikre og det er viktig å tallfeste denne usikkerheten. Usikkerheten avhenger av lengdeklasse, antatt trafikkmengde og tilgjengelige tell-data. Vi bruker et grovt, foreløpig, estimat til å anslå trafikkmengde i den aktuelle lengdeklasse, kalt  $\hat{A}DT(0)$ . Videre beregner vi antall talte timer innen 9 bestemte tidsperioder. Ut fra dette definerer vi variablene  $z_1$  til  $z_{10}$  angitt i Tabell 2.

**Tabell 2** Forklaringsvariable i modeller for usikkerhet i ÅDT-estimer.

---

$z_1 = \hat{A}DT(0) + 0.1$
$z_2 = 0.1 + \text{antall telletimer på hverdager i timene 8-9}$
$z_3 = 0.1 + \text{antall telletimer på hverdager i timene 10-15}$
$z_4 = 0.1 + \text{antall telletimer på hverdager i timene 16-17}$
$z_5 = 0.1 + \text{antall telletimer på hverdager i timene 7 samt 18-24}$
$z_6 = 0.1 + \text{antall telletimer på hverdager i timene 1-6}$
$z_7 = 0.1 + \text{antall telletimer på lørdager i timene 10-24}$
$z_8 = 0.1 + \text{antall telletimer på lørdager i timene 1-9}$
$z_9 = 0.1 + \text{antall telletimer på søndager i timene 10-24}$
$z_{10} = 0.1 + \text{antall telletimer på søndager i timene 1-9}$

---

For hver lengdeklasse beregnes standardavviket til ÅDT-estimatet ved formelen

$$\text{standardavvik} = \sqrt{c \cdot z_1^{\gamma_1} \cdot \dots \cdot z_{10}^{\gamma_{10}}} \quad . \quad (2)$$

Koeffisientene  $c$  og  $\gamma$ -ene avhenger av lengdeklasse. Tallverdier for disse er angitt i Tabell 9 i kapittel 5.4. I det kapittelet finnes det flere detaljer og et eksempel.

### 3.3 Beregning av trafikkarbeid med usikkerhet

I tillegg til trafikk på et enkelt punkt eller en enkelt veilenke kan man være interessert i et samlemål for totaltrafikk, f. eks. i et geografisk område. Dette kalles trafikkarbeid, og en hensiktsmessig definisjon av trafikkarbeid vil her være en veid sum av all trafikk, hvor vektene øker med størrelsen på bilene og med lengde på veilenkene. Videre bør de ulike lengdeklassene vektet med tanke på at større biler forurenses mer enn mindre biler. Anta at det i et område er  $n$  veilenker med lengder  $v_1, v_2, \dots, v_n$ , med tellinger for hver av disse. La  $w_1$  til  $w_5$  betegne vektene for de 5 lengdeklassene, og la  $\hat{A}DT_{\bar{i}}$  betegne et estimat for ÅDT for lengdeklasse  $l$  på veilenke  $i$ . Trafikkarbeid kan da estimeres med

$$\text{estimert trafikkarbeid} = \sum_{i=1}^n v_i \left( \sum_{l=1}^5 w_l \cdot \text{ÅDT}_{li} \right) \quad (3)$$

Usikkerheten til dette estimatet for trafikkarbeid avhenger av usikkerheten til de enkelte ÅDT-estimatene, og av korrelasjonen mellom dem. Kapittel 5.5 viser hvordan usikkerheten til det estimerte trafikkarbeid kan beregnes.

#### 4 Sammenligning av presisjon til ny og nåværende metode

Presisjonen til den nåværende variasjonskurvemethoden så vel som den nye metoden er beregna v.h.j.a. en simuleringsstudie. Denne studien er basert på data fra nivå 1 punktene, og tar ikke i bruk virkelige tellinger fra veilenker med få telletimer. Men veilenker med kontinuerlige tellinger kan spille rollen som veilenker med få telletimer, ved at man bare tar i bruk tellingene fra svært korte tidsperioder. Tabell 3 viser hvor store feil i ÅDT vi kan forvente for de to metodene på grunnlag av tellinger i omtrent 1 1/2 døgn. Vi oppgir her midlere absolutt feil (som tilsvarer ca. 1.25 standardavvik i en normalfordeling). For lengdeklasse 1 gir variasjonskurvemethoden typisk en feil på 9.0%, mens den nye metoden basert på kombinasjon av basiskurver gir en feil på 7.2%. Gevinsten ved å bruke den nye metoden skyldes i stor grad dens evne til å justere kompleksiteten i modellen etter hvor mange telldata som er tilgjengelig. Den nye metoden er også best for de andre lengdeklassene, og for disse skyldes gevinsten i større grad at den nye metoden er tilpasset hver enkelt av lengdeklassene. For de store lengdeklassene er feilprosenten stor uansett metode, da det er færre biler i disse klassene, men ved beregning av trafikkarbeid i et område vil feilene i stor grad utligne hverandre.

**Tabell 3** Gjennomsnittlig absolutt feil i ÅDT-estimer i % av sann ÅDT.

Lengdeklasse	Variasjonskurver	Kombinasjon av basiskurver
1	9.0	7.2
2	22.0	14.2
3	25.8	15.1
4	38.0	27.6
5	46.2	33.2

En nedgang fra 9.0% til 7.2% kan synes som en liten gevinst ved å ta i bruk en mer kompleks metode. Men i et eksempel i kapittel 7.5 viser vi at for å minske usikkerheten fra 9.0% til 7.2% trengs det omtrent 6 ganger så mange telletimer,



ved bruk av en og samme metode. Dermed kan man uten tap av presisjon minske antall telletimer til  $1/6$  ved å gå over fra å bruke den tradisjonelle variasjonskurve-metoden til å bruke metoden basert på kombinasjoner av basiskurver.

## Del 2: Detaljert framstilling

Denne delen er tenkt kun til bruk for spesielt interesserte, blant annet ved implementering. Den er detaljert og teknisk skrevet og bruker en mer presis notasjon en Del 1.

### 5 Beskrivelse av metoden basert på kombinasjon av basiskurver

#### 5.1 Etablering av basiskurver

For hver veilenke modelleres antall biler pr. time innen hver lengdeklasse som en funksjon av tidspunktet. Trafikken dekomponeres i nivå, trend, sesongvariasjon over året, spesielle dager, variasjon over ukedagene og variasjon over døgnet. Modellen og metoden (reduert rang regresjon) er i hovedtrekk de samme som er beskrevet i Aldrin (1995 a). Den viktigste forskjellen er at tidsenhet nå er time, slik at en også må inkludere døgnvariasjon. Videre er helligdager som faller på hverdager i utgangspunktet modellert som en søndag. Nivået for disse dagene er deretter korrigert ved hjelp av noen av variablene for spesielle dager, mens variasjonen over døgnet vil være som for en søndag.

Modellen er en regresjonsmodell hvor responsvariablene er logaritmen til (antall biler +1). Konstanten 1 legges til fordi antall biler pr. time kan være 0 for lengdeklassene over 5.5m. Responsvariablene modelleres som en lineær funksjon av en rekke forklaringsvariable som er angitt i Tabell 4. Modellen er dermed multiplikativ i original skala.

La  $t$  betegne en enkelt time, la  $y_t$  betegne logaritmen til (antall biler + 1) på en veilenke ved tid  $t$ , og la  $\mathbf{x}_t$  betegne de tilhørende forklaringsvariablene. En redusert rang regresjonsmodell (se f. eks. Davies og Tso 1982) med rang  $k$  kan for hver veilenke skrives

$$y_t = \alpha_0 + \alpha_1 \beta_1^T \mathbf{x}_t + \dots + \alpha_k \beta_k^T \mathbf{x}_t + \varepsilon_t \quad (4)$$

hvor  $\alpha$ -ene er skalarer knytta til hver respons (veilenke), mens  $\beta$ -ene er vektorer felles for alle responser. Videre er  $\varepsilon_t$  er et feilledd, mens  $T$  betyr transponert. Ved å definere  $i$ -te basiskurve som  $\mathbf{b}_{it} = \beta_i^T \mathbf{x}_t$  kan dette også skrives

$$y_t = \alpha_0 + \alpha_1 \mathbf{b}_{1t} + \dots + \alpha_k \mathbf{b}_{kt} + \varepsilon_t. \quad (5)$$

**Tabell 4** Forklaringsvariable.

- 
- Lineær trend
  - Sesongvariasjon over året, som består av 17 forklaringsvariable: 6 par sinus- og cosinus-funksjoner med periode 1, 1/2, ... , 1/6 år, 1 sinusfunksjon med periode 18 uker med virkning fra og med uke 25 og til og med uke 32, og 2 par sinus- og cosinus-funksjoner med periode 8 uker og 4 uker med virkning i same periode.
  - Spesielle dager, 17 indikatorvariable: 1) ekstra fridag på hverdag, 2) inneklemt dag, 3) julaften, 4) 1. og 2. juledag, 5) romjul hverdag, 6) romjul helg, 7) nyttårsaften, 8) nyttårsdag, 9) lørdag før palmesøndag, 10) palmesøndag, 11) mandag og tirsdag i påskeuka, 12) onsdag i påskeuka, 13) skjærtorsdag og langfredag, 14) påskeaften, 15) 1. påskedag, 16) 2. påskedag, 17) tirsdag etter påske.
  - Ukedagvariasjon og døgnvariasjon.  $7 \times 24 = 168$  indikatorvariable, en for hver time for hver enkelt av ukedagene.
- 

Denne modellen estimeres ut fra de 32 veilenkene i datasettet av nivå 1 punkter. Estimeringa skjer ved først å estimere nivået til hver responsvariabel ved å sentrere disse, og deretter utføre en redusert rang regresjon. Basiskurvene er ordna med minkende grad av viktighet. Den første basiskurven er en gjennomsnittskurve for alle veilenker i datasettet, mens de neste basiskurvene er å betrakte som justeringer av den første kurven.

Basiskurvene er vist i Figur 1 i Del 1 og diskutert der. I tillegg vil vi kommentere at for spesielle dager er utslagene er mindre enn vist i Aldrin (1995 a). Det skyldes at helligdager som faller på hverdager her betraktes som en søndag, før de korrigeres ved hjelp av indikatorvariable for spesielle dager.

## 5.2 ÅDT-beregninger basert på et gitt antall basiskurver

I kapittel 5.1 modellerte vi logaritmen til (antall biler pr. time +1) som en lineærkombinasjon av flere basiskurver, samt et konstantledd som representerer nivået. Basiskurvene er felles for alle veilenker, men hvordan de vektlegges via  $\alpha$ -koeffisientene er ulik fra veilenke til veilenke. For en veilenke med tellinger noen relativt få timer, beregnes nye  $\alpha$ -koeffisienter på grunnlag av telldataene, mens basiskurvene er som før. Dermed kan trafikkmengden estimeres for alle timer i en tidsperiode, f. eks. over et år.

Hvor mange basiskurver som bør inngå i lineærkombinasjonen vil avhenge av hvor mye telldata som er tilgjengelig til å bestemme koeffisientene i lineærkombinasjonen. Videre vil optimalt antall basiskurver kunne variere med lengdeklasser.

La  $b_{1t}$ ,  $b_{2t}$ ,  $b_{3t}$  osv. angi basiskurvene. De ulike mulighetene for opptil 2 samt generelt  $k$  basiskurver er angitt i Table 5, hvor  $\hat{\alpha}$  over  $\alpha$ -ene betyr at de er estimert fra de tilgjengelige målingene.

**Table 5** Mulige lineærkombinasjoner.

---

0 basiskurver med estimerte koeffisienter, men 1 med fast koeffisient	$\hat{y}_t(0) = \hat{\alpha}_0 + 1 b_{1t}$
1 basiskurve med estimert koeffisient	$\hat{y}_t(1) = \hat{\alpha}_0 + \hat{\alpha}_1 b_{1t}$
2 basiskurver med estimerte koeffisienter	$\hat{y}_t(2) = \hat{\alpha}_0 + \hat{\alpha}_1 b_{1t} + \hat{\alpha}_2 b_{2t}$
$k$ basiskurver med estimerte koeffisienter	$\hat{y}_t(k) = \hat{\alpha}_0 + \hat{\alpha}_1 b_{1t} + \dots + \hat{\alpha}_k b_{kt}$

---

I den første ligninga er  $\alpha_1$  fastsatt til 1, fordi basiskurvene er normert slik at den mest typiske verdien for  $\alpha_1$  er 1. Sjøl om den faktisk tar i bruk en basiskurve, nummererer vi denne ligninga med  $k=0$ .

La nå  $a_t$  betegne antall biler i time  $t$ . Ved å transformere modellen (5) og droppe indeksen  $j$  får vi at forventninga til  $(a_t + 1)$  er gitt ved:

$$E(a_t + 1) = E(\exp(\varepsilon_t)) \cdot \exp(\alpha_0) \cdot \exp(\alpha_1 b_{1t} + \dots + \alpha_k b_{kt}). \quad (6)$$

Minste kvadraters estimater  $\hat{\alpha}$  finnes ved hjelp av lineær regresjon på de tilgjengelige målingene med basiskurvene som forklaringsvariable og logaritmen til (antall biler pr. time +1) som respons. Et rett fram estimat for  $a_t$  vil ut fra dette være

$$\exp(\hat{y}_t(k)) - 1 = [\exp(\hat{\alpha}_0) \cdot \exp(\hat{\alpha}_1 b_{1t} + \dots + \hat{\alpha}_k b_{kt})] - 1, \quad (7)$$

men dette har to svakheter. For det første vil dette typisk gi underestimering av  $a_t$ , fordi  $\exp(\varepsilon_t)$  har en forventning som er større enn 1 selv om  $\varepsilon_t$  har forventning 0. For det andre kan vi få et negativt estimat for antall biler pr. time!

Vi definerer derfor en ny alternativ modell hvor forventninga til  $a_t$  er gitt ved

$$E(a_t) = c \cdot \exp(\tilde{\alpha}_1 b_{1t} + \dots + \tilde{\alpha}_k b_{kt}), \quad (8)$$

hvor  $\tilde{\alpha}$  over  $\alpha$ -ene angir at disse generelt kan være noe ulike de fra modellen (6). For data hvor  $a_t$  typisk er langt fra 0 vil modellene (6) og (8) være tilnærma lik hverandre, men hvis  $a_t$  er 0 eller nær 0 er de to modellene ulike. Vi estimerer nå antall biler pr. time med

$$\text{estimert antall biler i time } t = \hat{a}_t = \hat{c} \cdot \exp(\hat{\alpha}_1 b_{1t} + \dots + \hat{\alpha}_k b_{kt}), \quad (9)$$

hvor

$$\hat{c} = (1/n) \sum (a_t / \exp(\hat{\alpha}_1 b_{1t} + \dots + \hat{\alpha}_k b_{kt})), \quad (10)$$

hvor  $n$  er antall målinger og summen tas over de  $n$  tidspunkt det finnes målinger for. Dette vil alltid gi positive estimater, som i tillegg er omtrentlig forventningsrette. Men som estimater for  $\alpha$ -ene har vi i (9) for enkelhets skyld brukt minste kvadraters estimatene  $\hat{\alpha}$  som er konsistente for modell (6), men ikke for modell (8). Det er derfor trolig mulig å forbedre estimatene i (9), noe som kan være viktig for veilenker med lite trafikk. En mulighet er å bruke en generalisert lineær modell med Poisson- eller gamma-fordeling og logaritmisk link-funksjon, men andre ikke-lineære teknikker kan også brukes. Vi har ikke hatt tid til å prøve ut slike forbedringer.

Nå finnes det målinger  $a_t$  for noen timer, og estimater  $\hat{a}_t$  for alle andre timer. Dermed estimeres ÅDT som summen av  $a_t$ -ene og  $\hat{a}_t$ -ene over et år delt på antall døgn i året. Vi lar heretter ÅDT( $k$ ) betegne et ÅDT-estimat basert på  $k$  basiskurver.

Ved få data, eller om dataene inneholder noen ekstreme (og kanskje gale verdier), kan ÅDT-estimatene bli ustabile for  $k \geq 1$ . Vi har derfor modifisert prosedyren over på to måter:

i) Når det gjelder  $\alpha_1$  er det kun meningsfylt med positive verdier. For  $k=0$  er dette automatisk oppfylt, siden  $\alpha_1$  da er fastsatt til å være 1. For  $k=1$  settes  $\hat{\alpha}_1$  til 0 hvis regresjonsestimatet er negativt. Tilsvarende modifikasjon er ikke gjort for  $k \geq 2$ , da dette er mer komplisert.

ii) ÅDT( $k$ ) justeres i forhold til ÅDT(0) slik: Hvis ÅDT( $k$ )  $> 3 \cdot \text{ÅDT}(0)$  settes ÅDT( $k$ ) =  $3 \cdot \text{ÅDT}(0)$  og hvis ÅDT( $k$ )  $< (1/3) \cdot \text{ÅDT}(0)$  settes ÅDT( $k$ ) =  $(1/3) \cdot \text{ÅDT}(0)$ .

### 5.3 Valg av antall basiskurver

Det gjelder å velge antall basiskurver  $k$  slik at trafikken modelleres mest mulig nøyaktig, dvs.  $k$  må være stor nok. Samtidig må overtilpasning unngås, dvs.  $k$  må ikke være for stor i forhold til mengde av telldata vi har.

La  $e_k = \text{ÅDT} - \text{ÅDT}(k)$  være feilen i et ÅDT-estimat basert på  $k$  basiskurver. Vi ønsker å minimere denne feilen. Det gjøres ved først å modellere den forventede kvadratiske feil (betegnes med  $E(e_k^2)$ ) som en funksjon av lengdeklasse, antatt trafikkmengde, tilgjengelige telldata og av antall basiskurver  $k$ . Deretter velges den  $k$  som gir minst forventa kvadratisk feil. For en gitt  $k \leq 8$  har vi modellert  $E(e_k^2)$  som en multiplikativ funksjon på formen

$$E(e_k^2) = c_{lk} \cdot z_1^{\gamma_{1k}} \cdot \dots \cdot z_{10}^{\gamma_{10k}} \quad \begin{array}{l} l = 1, \dots, 5 \\ k = 0, \dots, 8 \end{array} \quad (11)$$

hvor forklaringsvariablene (z-ene) er angitt i Tabell 2. Verdien av  $z_1$  er forskjellig for hver lengdeklasse, de øvrige z-ene er like, uavhengige av lengdeklasse. Videre er  $c_l$  en separat konstant for hver lengdeklasse. Alle øvrige koeffisienter ( $\gamma$ -ene) er felles for alle lengdeklasser, for å begrense antall modeller. Koeffisientene er angitt i Tabell 6, v.h.j.a. beregninger som vi beskriver i kapittel 7.

I hovedtrekk er det slik at koeffisientene  $\gamma_2$  til  $\gamma_{10}$  er negative. Dermed minsker usikkerheten når antall telletimer øker, noe som er rimelig. Utgangspunktet for å lage slike modeller er at ved få telletimer lønner det seg å bruke liten  $k$ , mens ved mange telletimer vil en stor  $k$  gi mindre forventet kvadratisk feil. Ligningene bør gjenspeile dette. Vi vil presentere to eksempler på bruk av ligningene for lengdeklasse 1.

#### *Eksempel 1*

Anta først at vi for en veilenke har talt en hverdag fra og med time 7 til og med time 24, og at estimatet  $\hat{A}DT(0)$  er 20000 biler pr.døgn. Vi beregner  $E(e_k^2)$  v.h.j.a. ligning (11) og Tabell 6. Øverste linje i Tabell 7 viser kvadratrota til  $E(e_k^2)$  for ulike  $k$ , og vi ser at kvadratrota og dermed også selve  $E(e_k^2)$  er minst for  $k=0$ . Om vi øker antall telletimer slik at vi teller et helt døgn en hverdag og et helt døgn en lørdag, og  $\hat{A}DT(0)$  fortsatt er 20000, får vi resultatene som er gjengitt i andre linje i tabellen. Nå er  $E(e_k^2)$  minst for  $k=3$ . Om vi øker antall telletimer ytterligere til en hel uke, får vi resultatene som er gjengitt i tredje linje i tabellen, og nå er  $E(e_k^2)$  minst for  $k=8$ .

#### *Eksempel 2*

Neste eksempel tar utgangspunkt i den samme tellesituasjon som forrige eksempel, dvs. vi har talt en hverdag fra og med time 7 til og med time 24, og at estimatet  $\hat{A}DT(0)$  er 20000 biler pr.døgn. Deretter fem-dobler vi antall telletimer slik at vi teller fem hverdager fra og med time 7 til og med time 24. Til slutt fem-dobler vi nok en gang antall telletimer. Resultatet er gjengitt i Tabell 8, og viser at det lønner seg å ha  $k$  større enn 0 først når vi har talt 25 hverdager fra og med time 7 til og med time 24. Om det kun telles på dagtid hverdager, vil det altså i praksis alltid være tilstrekkelig med  $k=0$ .

**Tabell 6** Regresjonskoeffisienter i modeller for valg av k for k=0, ... , 8.

Variabelnavn	Antall basiskurver				
	0	1	2	3	4
c lengdeklasse 1	414.2858	320.6233	136.6247	77.3192	76.3266
c lengdeklasse 2	63.4686	55.1062	36.9768	33.8253	36.3280
c lengdeklasse 3	41.7237	35.8355	20.3968	23.1586	34.0798
c lengdeklasse 4	33.7893	28.8915	25.8615	26.7891	29.2201
c lengdeklasse 5	47.0015	36.6512	43.4738	50.3551	52.5095
$\gamma_1$ , ÅDT-prediksjon med k=0	1.1210	1.0819	1.1669	1.2159	1.2322
$\gamma_2$ , hverdag 8-9	-0.0735	-0.0578	-0.0482	-0.1104	-0.1016
$\gamma_3$ , hverdag 10-15	-0.1754	-0.1618	-0.1482	-0.1566	-0.1700
$\gamma_4$ , hverdag 16-17	-0.0903	-0.0672	-0.0818	-0.0890	-0.0941
$\gamma_5$ , hverdag 7,18-24	-0.0652	-0.0854	-0.1130	-0.1172	-0.1338
$\gamma_6$ , hverdag 1-6	0.0467	-0.0357	-0.0537	-0.0592	-0.0846
$\gamma_7$ , lørdag 10-24	0.0128	-0.0411	-0.0590	-0.0600	-0.0755
$\gamma_8$ , lørdag 1-9	0.0593	-0.0108	-0.0270	-0.0281	-0.0535
$\gamma_9$ , søndag 10-24	0.0048	-0.0354	-0.0624	-0.0725	-0.0997
$\gamma_{10}$ , søndag 1-9	0.1021	0.0292	0.0138	0.0103	-0.0166

Variabelnavn	Antall basiskurver			
	5	6	7	8
c lengdeklasse 1	85.0867	78.8017	76.5876	80.7399
c lengdeklasse 2	56.1138	96.2448	99.2034	107.8489
c lengdeklasse 3	42.9184	42.8810	43.1507	44.1507
c lengdeklasse 4	35.1893	34.6784	36.6200	37.5988
c lengdeklasse 5	57.5363	59.2817	61.0577	64.2172
$\gamma_1$ , ÅDT-prediksjon med k=0	1.2364	1.2641	1.2795	1.2947
$\gamma_2$ , hverdag 8-9	-0.1053	-0.1051	-0.1095	-0.1170
$\gamma_3$ , hverdag 10-15	-0.1762	-0.1781	-0.1822	-0.1856
$\gamma_4$ , hverdag 16-17	-0.1087	-0.1097	-0.1179	-0.1265
$\gamma_5$ , hverdag 7,18-24	-0.1497	-0.1568	-0.1621	-0.1763
$\gamma_6$ , hverdag 1-6	-0.1089	-0.1241	-0.1329	-0.1379
$\gamma_7$ , lørdag 10-24	-0.0873	-0.0959	-0.1087	-0.1187
$\gamma_8$ , lørdag 1-9	-0.0704	-0.0771	-0.0836	-0.0964
$\gamma_9$ , søndag 10-24	-0.1073	-0.1172	-0.1253	-0.1319
$\gamma_{10}$ , søndag 1-9	-0.0336	-0.0399	-0.0548	-0.0652

**Tabell 7** Kvadratrot til  $E(e_k^2)$  for 3 tellesituasjoner, første eksempel. Minimumsverdi for hver tellesituasjon er utheva.

Tellesituasjon	Antall basiskurver								
	0	1	2	3	4	5	6	7	8
1 hverdag 7-24	<b>3010</b>	3162	3431	3246	3897	4455	5135	5710	6492
2 døgn (hverdag+lørdag)	3911	2587	2493	<b>2321</b>	2402	2441	2629	2740	2920
1 uke	3753	1828	1544	1298	1133	1032	1048	1010	<b>1003</b>

**Tabell 8** Kvadratrot til  $E(e_k^2)$  for 3 tellesituasjoner, andre eksempel. Minimumsverdi for hver tellesituasjon er utheva.

Tellesituasjon	Antall basiskurver								
	0	1	2	3	4	5	6	7	8
1 hverdag 7-24	<b>3010</b>	3162	3431	3246	3897	4455	5135	5710	6492
5 hverdager 7-24	<b>2184</b>	2353	2514	2230	2622	2902	3319	3627	4015
25 hverdager 7-24	1579	1745	1837	<b>1526</b>	1756	1882	2135	2292	2470

#### 5.4 Usikkerhet i ÅDT-estimat

La nå  $\hat{A}DT(d)$  være et  $\hat{A}DT$ -estimat basert på et antall basiskurver som er valgt ut fra data som angitt i kapittel 5.3, og la nå  $e_d = \hat{A}DT - \hat{A}DT(d)$  være tilsvarende feil. Vi ønsker å anslå usikkerheten til  $\hat{A}DT$ -estimatet, og gjør det ved å modellere  $E(e_d^2)$  som en funksjon av lengdeklasse, antatt trafikkmengde og tilgjengelige telldata. Vi kunne i prinsippet brukt modellene i kapittel 5.3, men disse har flere svakheter. De tar ikke hensyn til at valg av  $k$  medfører en ekstra usikkerhet. Videre er en felles modell for alle lengdeklassene noe restriktiv. Vi har derfor laget modeller for  $E(e_d^2)$ , som tar vare på usikkerheten forbundet med valg av  $k$ :

$$E(e_d^2) = c_l \cdot z_1^{Y_{11}} \cdot \dots \cdot z_{10}^{Y_{10l}} \quad l = 1, \dots, 5. \quad (12)$$

Disse modellene er på lignende form som (11), men det er separate modeller for hver lengdeklasse, og ingen avhengighet av  $k$ .

Forventa kvadratisk feil er summen av varians og kvadratet av forventningsskjevheten. Forventningsskjevheten er relativt liten (se kapittel 7.5), slik at anslaget på  $E(e_d^2)$  også er et anslag for variansen, og kvadratrotta av dette er et anslag på standardavviket.



Koeffisientene er beregna i kapittel 7.3 og angitt i Tabell 9. Tabellen viser at usikkerheten minker når antall telletimer øker, videre at tellinger på hverdager bidrar til å minske usikkerheten mer enn tellinger i helger. Det betyr imidlertid ikke at man bare bør telle på hverdager, se eksemplet nedafor. Tilsvarende gir tellinger på hverdager 10-15 og 7,18-24 større gevinst en tellinger i rushtida. Merk også at tellinger natt til lørdag eller søndag faktisk kan øke usikkerheten for de største lengdeklassene.

**Tabell 9** Regresjonskoeffisienter i modeller for usikkerhet

Variabelnavn	Lengdeklasse				
	1	2	3	4	5
c	2.8450	168.7742	48.9633	19.9283	68.2979
$\gamma_1$ , ÅDT-prediksjon med k=0	1.4227	1.6580	1.0440	1.0120	0.9305
$\gamma_2$ , hverdag 8-9	-0.0675	-0.0674	-0.0721	-0.0374	-0.0638
$\gamma_3$ , hverdag 10-15	-0.1074	-0.1785	-0.2369	-0.1567	-0.1762
$\gamma_4$ , hverdag 16-17	-0.0658	-0.0700	-0.0571	-0.0915	-0.0765
$\gamma_5$ , hverdag 7,18-24	-0.0980	-0.0506	-0.0803	-0.0562	-0.0607
$\gamma_6$ , hverdag 1-6	-0.0487	0.0050	-0.0042	-0.0149	-0.0201
$\gamma_7$ , lørdag 10-24	-0.0485	-0.0284	-0.0107	0.0026	0.0304
$\gamma_8$ , lørdag 1-9	-0.0132	-0.0109	-0.0139	0.0246	0.0334
$\gamma_9$ , søndag 10-24	-0.0657	0.0056	0.0143	-0.0145	-0.0280
$\gamma_{10}$ , søndag 1-9	-0.0038	-0.0038	0.0354	0.0618	0.1139

### Eksempel

La oss ta som et eksempel ta for oss to ulike tellesituasjoner for lengdeklasse 1. Anta det for en veilenke er talt seks timer hverdag 10-15 og to timer hverdag 16-17, og at estimatet ÅDT(0) er 48000 biler pr.døgn. Ligning (12), Tabell 2 og Tabell 9 gir da

$$E(e_d^2) = \frac{2.8450 \cdot 48000 \cdot 0.1^{1.4227} \cdot 0.1^{-0.0675} \cdot 6.1^{-0.1074} \cdot 2.1^{-0.0658}}{0.1^{-0.0980} \cdot 0.1^{-0.0487} \cdot 0.1^{-0.0485} \cdot 0.1^{-0.0132} \cdot 0.1^{-0.0657} \cdot 0.1^{-0.0038}} \quad (13)$$

slik at kvadratota til  $E(e_d^2)$  blir 4753. Anta så at man alternativt fordeler de åtte talte timene anderledes, med tre timer hverdag 10-15, to timer hverdag 16-17, en time hverdag 7,18-24 og to timer søndag 10-24, mens ÅDT(0) fortsatt er 48000 biler pr.døgn. Da blir kvadratota til  $E(e_d^2)$  blir 3966. Usikkerheten blir altså betydelig mindre ved å spre tellingene.

## 5.5 Beregning av trafikkarbeid i et område, med usikkerhetsanslag

Kapittel 5.4 tok for seg usikkerhet i et ÅDT-estimat for en enkelt veilenke og en enkelt lengdeklasse. Imidlertid vil det være ønskelig å kunne anslå usikkerheten i det totale trafikkarbeidet i et geografisk område. En hensiktsmessig definisjon av trafikkarbeid vil her være en veid sum av all trafikk, hvor vektene øker med størrelsen på bilene og med lengde på veilenkene. Videre bør de ulike lengdeklassene vektas med tanke på at større biler forurenses mer enn mindre biler. Anta at det i et område er  $n$  veilenker med lengder  $v_1, v_2, \dots, v_n$ , med tellinger for hver av disse. La  $w_1$  til  $w_5$  betegne vektene for de 5 lengdeklassene, og la  $\text{ÅDT}_{li}$  betegne sann ÅDT (i motsetning til i kapittel 3.3 hvor det sto for et estimat) for lengdeklasse  $l$  på veilenke  $i$ . En hensiktsmessig definisjon av trafikkarbeid er da

$$\text{trafikkarbeid} = \sum_{i=1}^n v_i \left( \sum_{l=1}^5 w_l \cdot \text{ÅDT}_{li} \right). \quad (14)$$

Dette estimeres ved å erstatte sann  $\text{ÅDT}_{li}$  med et estimatet  $\text{ÅDT}(d)_{li}$ , hvor antall basiskurver er valgt ut fra data etter den angitte prosedyren og altså kan variere avhengig av lengdeklasse og tellesituasjon.

Usikkerheten til dette estimatet avhenger av usikkerheten til det enkelte estimat  $\text{ÅDT}(d)_{li}$ , men også av korrelasjon mellom ulike estimater. Når det gjelder estimater som er basert på tellinger på ulike tidspunkt, er det ingen grunn til å tro at disse er korrelert. Når estimatene derimot er basert på samtidige (eller nesten samtidige) tellinger, kan det gi estimater som er positivt korrelert, særlig når veilenkene er i nærheten av hverandre. Estimaterne for ulike lengdeklasser på samme veilenke vil være basert på tellinger som er gjort samtidig, og vil typisk være positivt korrelert. Disse korrelasjonene er beregna som forklart i kapittel 7.4 og angitt i Tabell 11. La  $E(e_{li}^2)$  betegne forventa kvadratisk feil for det enkelte estimat  $\text{ÅDT}(d)_{li}$  beregna ut fra modellene kapittel 5.4, og la  $\rho_{ll'}$  betegne korrelasjonen mellom estimater for lengdeklassene  $l$  og  $l'$  på samme veilenke gitt av Tabell 11. Usikkerheten (målt som forventa kvadratisk feil) i estimatet for trafikkarbeid er da gitt ved

$$\sum_{i=1}^n v_i^2 \left( \sum_{l=1}^5 w_l^2 \cdot E(e_{li}^2) + 2 \sum_{l,l'} w_l \cdot w_{l'} \cdot \sqrt{E(e_{li}^2)} \cdot \sqrt{E(e_{l'i}^2)} \cdot \rho_{ll'} \right). \quad (15)$$

hvor den siste summen tas over alle 10 par av lengdeklasser  $l$  og  $l'$ . Den siste summen innafor parentesen er korreksjonsledd som tar hensyn til at estimatene for de ulike lengdeklassene er basert på samtidige tellinger og derfor er korrelerte.

### Eksempel

La oss ta et eksempel med to veilenker, og kun lengdeklasse 1 og 5. De nødvendige størrelsene er gitt i Tabell 10.

**Tabell 10** Data for eksempel.

---

$v_1 = 1$	$v_2 = 2$
$w_1 = 1$	$w_5 = 3$
$\text{ÅDT}_{11} = 20000$	$\sqrt{E(e_{11}^2)} = 2000$
$\text{ÅDT}_{51} = 2000$	$\sqrt{E(e_{51}^2)} = 200$
$\text{ÅDT}_{21} = 10000$	$\sqrt{E(e_{21}^2)} = 1000$
$\text{ÅDT}_{52} = 1000$	$\sqrt{E(e_{52}^2)} = 100$
$\rho_{15} = 0.03$	

---

Dermed blir

$$\text{trafikkarbeid} = 1 \cdot 1 \cdot 20000 + 1 \cdot 3 \cdot 2000 + 2 \cdot 1 \cdot 10000 + 2 \cdot 3 \cdot 1000 = 52000 ,$$

mens forventa kvadratisk feil i dette estimatet blir

$$\begin{aligned} & 1^2(1^2 \cdot 2000^2 + 3^2 \cdot 200^2 + 2 \cdot 1 \cdot 3 \cdot 2000 \cdot 200 \cdot 0.03) \\ & + 2^2(1^2 \cdot 1000^2 + 3^2 \cdot 100^2 + 2 \cdot 1 \cdot 3 \cdot 1000 \cdot 100 \cdot 0.03) \\ & = 8864000 \end{aligned} ,$$

Kvadratrotta til dette er 2977, og utgjør et standardavvik for det estimerte trafikkarbeidet.

Usikkerhetsanslagene til det enkelte ÅDT-estimat avhenger av estimatet ÅDT(0) som er basert på tellinger. Men om dette erstattes med et foreløpig estimat basert på skjønn, kan man uten å gjøre tellinger grovt angi usikkerheten i både ÅDT-estimatene og estimatet for trafikkarbeidet for et gitt telleopplegg. Dette kan brukes til

å designe et telleopplegg hvor det telles mer på store veier enn på mindre, slik at usikkerheten i trafikkarbeidet blir minst mulig med færrest mulig telletimer.

**Tabell 11** Korrelasjon mellom ÅDT-estimer for ulike lengdeklasser.

Lengdeklasse	1	2	3	4	5
1	1.00	0.01	0.10	0.08	0.03
2	0.01	1.00	0.51	0.07	0.03
3	0.10	0.51	1.00	0.52	0.48
4	0.08	0.07	0.52	1.00	0.65
5	0.03	0.03	0.48	0.65	1.00

## 5.6 Omregning mellom inndeling av biler etter biltype og etter lengde

I tellingene ved nivå 1 punktene er bilene inndelt i 5 lengdeklasser, og alle beregninger i denne rapporten er gjort for disse lengdeklassene. Imidlertid vil manuelle tellinger gruppere bilene etter biltype. Likeledes vil luftovervåkningssystemet for Oslo, hvor beregningene er tenkt brukt, trenge biltypeklassifiserte tall som inndata. Det er derfor nødvendig å finne en omregningsregel fra lengdekategorier til biltypekategorier og vice versa. Når en slik regel er funnet, kan tellingene gjøres med biltypeklassifisering, som deretter transformeres til lengdeklassifiserte tellinger. De nødvendige beregninger gjøres så for de ulike lengdeklassene, og sluttsvarene transformeres deretter til biltypeklassifisering som overvåkningssystemet kan bruke.

Det er i Aldrin (1995 b) foreslått et opplegg for å finne en omregningsregel mellom lengdeklasser og biltypekategorier. Ved å gjennomføre manuelle tellinger på nivå 1 punktene kan en beregne sammenhengen mellom de maskinelle og de manuelle tellingene. Denne sammenhengen vil antagelig variere med ukedag, tid på døgnet og hvilken type vei det er snakk om. Det enkleste vil være å se bort fra denne variasjonen, men dette kan modelleres relativt enkelt. Det vil trolig gi en langt mer nøyaktig omregningsregel, og hindre systematiske feilberegninger.

## 6 Datagrunnlag

### 6.1 Oversikt over data

De 5 lengdeklassene er i meter: 0-5.5, 5.6-7.6, 7.7-12.4, 12.5-15.9 og 16.0 og lengre. Den nåværende analysen er basert på tellinger på timesbasis fra starten av 1994 til høsten 1996. 16 veier med tellinger i to retninger er med. Disse er listet i Tabell 12 med anslag på ÅDT for hver enkelt lengdeklasse og hvert år, sammen

med hvor stor andel av hvert enkelt år det finnes pålitelige data for. Vålerengatunnelen var med i det opprinnelige datasettet, men er ikke tatt med i analysen her fordi den fikk en ekstrem økning i trafikken omkring åpninga av Ekebergtunnelen, og dermed er atypisk.

Dataene er levert av Statens Vegvesen Oslo, Vegtrafikksentralen, på et såkalt RTD-format. I RTD-formatet er antall biler registrert pr. felt, og ned til 5 minutters tidsintervaller. Vi har aggregert rådataene til trafikk pr. veiretning og pr. time. Deretter er dataene korrigert etter en prosedyre angitt i kapittel 6.2. Alle beregninger er utført på dette korrigerede datasettet. Dataene har ikke på forhånd gått gjennom kontrollrutiner som Vegdirektoratets KRK-system. Om så var tilfelle, ville nok vårt arbeide med korrigering vært enklere, men det er usikkert om KRK-systemet ville avdekket tilstrekkelig med feil.

Vi vil kommentere en detalj i Tabell 12 som kan virke litt pussig. Vestgående løp i Vaterlandtunnelen har svært liten trafikk i lengdeklassene 2 til 5. Dette synes rart, men siden tallene er stabile fra år til år har vi valgt å bruke dem.

**Tabell 12** Veier i datasettet, med årlige ÅDT-anslag for hver enkelt lengdeklasse, og andel av tidsperioden det finnes data for. Tall basert på tellinger minst 30% av året er utheva. (forts. neste side)

Vei	Retning	ÅDT								
		Lengdekl.1			Lengdekl.2			Lengdekl.3		
		94	95	96	94	95	96	94	95	96
RV 4, Tr.h.vn., Ammerud	Fra sentrum	12855	<b>12615</b>	<b>12349</b>	750	<b>715</b>	<b>702</b>	398	<b>379</b>	<b>369</b>
RV 4, Tr.h.vn., Ammerud	Mot sentrum	13152	<b>12812</b>	<b>12730</b>	662	<b>636</b>	<b>640</b>	431	<b>409</b>	<b>406</b>
RV 190, Ekeberg tunnelen	Mot sentrum	-	28616	<b>26687</b>	-	2610	<b>2335</b>	-	1612	<b>1336</b>
RV 190, Ekeberg tunnelen	Fra sentrum	-	26143	<b>24039</b>	-	1668	<b>1448</b>	-	841	<b>704</b>
E18, Fiskvollbukta	Fra Moss	9623	<b>8919</b>	<b>8841</b>	458	<b>424</b>	<b>394</b>	347	<b>321</b>	<b>325</b>
E18, Fiskvollbukta	Fra Oslo	11666	<b>11110</b>	<b>11204</b>	427	<b>399</b>	<b>386</b>	391	<b>361</b>	<b>358</b>
E18, Frognerstranda	Mot Drammen	<b>31764</b>	<b>31624</b>	<b>32480</b>	<b>2701</b>	<b>3235</b>	<b>3818</b>	<b>808</b>	<b>751</b>	<b>826</b>
E18, Frognerstranda	Mot Oslo	<b>28836</b>	<b>27955</b>	<b>28729</b>	<b>4880</b>	<b>6173</b>	<b>6879</b>	<b>928</b>	<b>924</b>	<b>931</b>
RV 150, Ringvn., Gaustad	Mot Smestad	18655	<b>18294</b>	<b>18250</b>	1472	<b>1599</b>	<b>1685</b>	372	<b>382</b>	<b>376</b>
RV 150, Ringvn., Gaustad	Mot Tåsen	19715	<b>19401</b>	<b>19625</b>	1594	<b>1790</b>	<b>1719</b>	426	<b>435</b>	<b>423</b>
RV 150, Granfosstunnelen	Nordgående løp	<b>9028</b>	<b>9901</b>	<b>10168</b>	<b>207</b>	<b>235</b>	<b>252</b>	<b>174</b>	<b>219</b>	<b>225</b>
RV 150, Granfosstunnelen	Sydgående løp	<b>9409</b>	<b>10459</b>	<b>10936</b>	<b>291</b>	<b>343</b>	<b>357</b>	<b>226</b>	<b>279</b>	<b>287</b>
E6, Karihaugen	Fra Oslo	<b>12936</b>	<b>13148</b>	<b>13528</b>	<b>774</b>	<b>952</b>	<b>1034</b>	<b>558</b>	<b>536</b>	<b>544</b>
E6, Karihaugen	Fra Hamar	<b>13077</b>	<b>12894</b>	<b>13255</b>	<b>449</b>	<b>733</b>	<b>873</b>	<b>272</b>	<b>483</b>	<b>582</b>
RV 150, Løren	Fra Økern	28879	<b>28319</b>	<b>28936</b>	1209	<b>1224</b>	<b>1267</b>	733	<b>741</b>	<b>754</b>
RV 150, Løren	Fra Sinsen	26920	<b>26516</b>	<b>27303</b>	1198	<b>1209</b>	<b>1259</b>	731	<b>721</b>	<b>727</b>
E6, Ringvn., Manglerud	Mot Sinsen	24561	<b>24002</b>	<b>22895</b>	1028	<b>1002</b>	<b>929</b>	727	<b>718</b>	<b>680</b>
E6, Ringvn., Manglerud	Mot Ryen	23072	<b>22723</b>	<b>22541</b>	961	<b>978</b>	<b>967</b>	739	<b>734</b>	<b>712</b>
E18, Maritim	Fra Oslo	38261	<b>38178</b>	<b>39932</b>	3623	<b>3560</b>	<b>3559</b>	748	<b>779</b>	<b>798</b>
E18, Maritim	Mot Oslo	39275	<b>39438</b>	<b>41580</b>	2840	<b>2730</b>	<b>2744</b>	1092	<b>1095</b>	<b>1166</b>
E6, Mortensrud	Mot Oslo	14949	15013	<b>15072</b>	949	974	<b>1009</b>	582	588	<b>597</b>
E6, Mortensrud	Mot Stockholm	15296	15321	<b>15445</b>	835	870	<b>914</b>	507	515	<b>522</b>
RV 150, Ringvn., Nydalen	Fra Sinsen	<b>17074</b>	<b>20101</b>	<b>21465</b>	<b>806</b>	<b>981</b>	<b>1002</b>	<b>482</b>	<b>565</b>	<b>574</b>
RV 150, Ringvn., Nydalen	Fra Tåsen	<b>17369</b>	<b>20537</b>	<b>21883</b>	<b>755</b>	<b>896</b>	<b>970</b>	<b>454</b>	<b>534</b>	<b>542</b>
E18, Fjellinja	Vestgående løp	<b>27757</b>	<b>31869</b>	<b>37038</b>	<b>1510</b>	<b>1668</b>	<b>1880</b>	<b>844</b>	<b>861</b>	<b>915</b>
E18, Fjellinja	Østgående løp	<b>30326</b>	<b>34397</b>	<b>39181</b>	<b>968</b>	<b>1096</b>	<b>1262</b>	<b>720</b>	<b>781</b>	<b>828</b>
RV 162, Vaterlandstu.	Vestgående løp	<b>6823</b>	<b>7272</b>	7727	<b>14.9</b>	<b>15.2</b>	13.4	<b>63.8</b>	<b>65.6</b>	47.6
RV 162, Vaterlandstu.	Østgående løp	<b>8217</b>	<b>8497</b>	8839	<b>72.2</b>	<b>73.7</b>	70.5	<b>132</b>	<b>139</b>	118
RV 161, Vøyen bru	Fra Sandakervn	<b>10653</b>	<b>10439</b>	<b>10295</b>	<b>300</b>	<b>279</b>	<b>276</b>	<b>122</b>	<b>121</b>	<b>116</b>
RV 161, Vøyen bru	Fra Maridalsvn	<b>11088</b>	<b>10925</b>	<b>10895</b>	<b>262</b>	<b>253</b>	<b>253</b>	<b>121</b>	<b>119</b>	<b>112</b>
RV 168, Wergelandsvn.	Mot Parkveien	<b>8307</b>	<b>8390</b>	<b>8231</b>	<b>139</b>	<b>119</b>	<b>111</b>	<b>57</b>	<b>53.5</b>	<b>55.9</b>
RV 168, Wergelandsvn.	Mot sentrum	<b>5629</b>	<b>5801</b>	<b>5676</b>	<b>84.7</b>	<b>83.8</b>	<b>77.4</b>	<b>38.6</b>	<b>39.3</b>	<b>36.1</b>

Vei	Retning	ÅDT						Prosentandel observert		
		Lengdekl.4			Lengdekl.5			94	95	96
		94	95	96	94	95	96			
RV 4, Tr.h.vn., Ammerud	Fra sentrum	68.1	<b>68.7</b>	<b>67.3</b>	85.7	<b>86.1</b>	<b>86.1</b>	0	<b>51.8</b>	<b>69.3</b>
RV 4, Tr.h.vn., Ammerud	Mot sentrum	75.2	<b>76.1</b>	<b>74.9</b>	90.8	<b>90.4</b>	<b>93.6</b>	0	<b>52.1</b>	<b>69.4</b>
RV 190, Ekebergtunnelen	Mot sentrum	-	758	<b>625</b>	-	631	<b>593</b>	0	1.4	<b>33.7</b>
RV 190, Ekebergtunnelen	Fra sentrum	-	574	<b>498</b>	-	656	<b>581</b>	0	1.4	<b>33.8</b>
E18, Fiskvollbukta	Fra Moss	180	<b>179</b>	<b>194</b>	216	<b>225</b>	<b>256</b>	0	<b>49</b>	<b>61.2</b>
E18, Fiskvollbukta	Fra Oslo	191	<b>179</b>	<b>188</b>	263	<b>262</b>	<b>297</b>	0	<b>50.7</b>	<b>61.4</b>
E18, Frognerstranda	Mot Drammen	<b>358</b>	<b>385</b>	<b>422</b>	<b>446</b>	<b>384</b>	<b>399</b>	<b>37.8</b>	<b>79.1</b>	<b>54.9</b>
E18, Frognerstranda	Mot Oslo	<b>416</b>	<b>431</b>	<b>465</b>	<b>660</b>	<b>605</b>	<b>525</b>	<b>37.7</b>	<b>79.2</b>	<b>55</b>
RV 150, Ringvn., Gaustad	Mot Smestad	104	<b>122</b>	<b>117</b>	95.2	<b>113</b>	<b>100</b>	23.7	<b>95.9</b>	<b>68.3</b>
RV 150, Ringvn., Gaustad	Mot Tåsen	74.1	<b>84.6</b>	<b>79.1</b>	92.6	<b>98.6</b>	<b>83.4</b>	23.7	<b>95.9</b>	<b>68.4</b>
RV 150, Granfosstunnelen	Nordgående løp	<b>67.5</b>	<b>74.5</b>	<b>76.5</b>	<b>20.7</b>	<b>34.7</b>	<b>35.5</b>	<b>91.6</b>	<b>62.3</b>	<b>47.2</b>
RV 150, Granfosstunnelen	Sydgående løp	<b>74.8</b>	<b>82.8</b>	<b>87.5</b>	<b>45.5</b>	<b>69.1</b>	<b>65.4</b>	<b>91.2</b>	<b>65</b>	<b>46.8</b>
E6, Karihaugen	Fra Oslo	<b>209</b>	<b>263</b>	<b>303</b>	<b>345</b>	<b>436</b>	<b>524</b>	<b>82.3</b>	<b>91.2</b>	<b>62.4</b>
E6, Karihaugen	Fra Hamar	<b>93.5</b>	<b>207</b>	<b>289</b>	<b>139</b>	<b>362</b>	<b>599</b>	<b>73.3</b>	<b>45.1</b>	<b>62.4</b>
RV 150, Løren	Fra Økern	125	<b>135</b>	<b>139</b>	180	<b>189</b>	<b>180</b>	0	<b>98.7</b>	<b>67.6</b>
RV 150, Løren	Fra Sinsen	127	<b>132</b>	<b>134</b>	161	<b>167</b>	<b>163</b>	0	<b>98.5</b>	<b>67.7</b>
E6, Ringvn., Manglerud	Mot Sinsen	159	<b>162</b>	<b>128</b>	308	<b>297</b>	<b>216</b>	0	<b>88</b>	<b>64.1</b>
E6, Ringvn., Manglerud	Mot Ryen	141	<b>141</b>	<b>114</b>	266	<b>256</b>	<b>197</b>	0	<b>87.7</b>	<b>63.9</b>
E18, Maritim	Fra Oslo	434	<b>471</b>	<b>514</b>	416	<b>442</b>	<b>485</b>	25.5	<b>97.3</b>	<b>57.5</b>
E18, Maritim	Mot Oslo	439	<b>470</b>	<b>504</b>	494	<b>536</b>	<b>609</b>	25.5	<b>97.2</b>	<b>57.4</b>
E6, Mortensrud	Mot Oslo	166	167	<b>174</b>	318	323	<b>333</b>	0	12.6	<b>54.9</b>
E6, Mortensrud	Mot Stockholm	202	205	<b>211</b>	231	236	<b>240</b>	0	12.6	<b>54.9</b>
RV 150, Ringvn., Nydalen	Fra Sinsen	<b>56.6</b>	<b>95.3</b>	<b>104</b>	<b>90.5</b>	<b>153</b>	<b>157</b>	<b>95.4</b>	<b>97.3</b>	<b>67.4</b>
RV 150, Ringvn., Nydalen	Fra Tåsen	<b>52.7</b>	<b>87.9</b>	<b>99.2</b>	<b>66.2</b>	<b>119</b>	<b>119</b>	<b>95.3</b>	<b>97.4</b>	<b>67.4</b>
E18, Fjellinja	Vestgående løp	<b>688</b>	<b>711</b>	<b>777</b>	<b>164</b>	<b>169</b>	<b>191</b>	<b>86.7</b>	<b>74.3</b>	<b>37</b>
E18, Fjellinja	Østgående løp	<b>683</b>	<b>725</b>	<b>796</b>	<b>101</b>	<b>124</b>	<b>131</b>	<b>86.6</b>	<b>74.9</b>	<b>37.6</b>
RV 162, Vaterlandstu.	Vestgående løp	<b>0.941</b>	<b>1.25</b>	0.743	<b>0.165</b>	<b>0.166</b>	0.113	<b>85.9</b>	<b>87.6</b>	22.7
RV 162, Vaterlandstu.	Østgående løp	<b>17.9</b>	<b>17.1</b>	17.2	<b>26.3</b>	<b>24.4</b>	21.2	<b>87.5</b>	<b>87.6</b>	22.7
RV 161, Vøyen bru	Fra Sandakervn	<b>14.8</b>	<b>17.6</b>	<b>12.8</b>	<b>11.6</b>	<b>9.5</b>	<b>5.71</b>	<b>53.4</b>	<b>93.6</b>	<b>56.6</b>
RV 161, Vøyen bru	Fra Maridalsvn	<b>5.36</b>	<b>7.03</b>	4.41	<b>11.1</b>	<b>9.22</b>	<b>5.38</b>	<b>53.4</b>	<b>93.5</b>	<b>56.5</b>
RV 168, Wergelandsvn.	Mot Parkveien	<b>3.3</b>	<b>3.02</b>	<b>2.8</b>	<b>5.82</b>	<b>2.42</b>	<b>0.527</b>	<b>31.4</b>	<b>90.8</b>	<b>54.8</b>
RV 168, Wergelandsvn.	Mot sentrum	<b>4.46</b>	<b>3.94</b>	<b>3.38</b>	<b>13.7</b>	<b>1.99</b>	<b>0.46</b>	<b>31.4</b>	<b>92.5</b>	<b>54.8</b>

## 6.2 Prosedyre for korrigerings av data

De maskinelle målingene vil typisk inneholde endel observasjoner hvor måleinstrumentet ikke har fungert skikkelig og altså viser feil. I andre tilfeller kan veien ha vært avstengt p.g.a. asfaltering o.l. Da er de registrerte dataene egentlig ikke feil, men de er ikke representative for normalsituasjonen. Slike observasjoner må derfor lukes ut før dataene brukes som grunnlag for endelige beregninger. Videre kan dataene være forskjøvet i tid, slik at tidspunktet må korrigeres. Slike feil og mangler i dataene kan være av svært ulik karakter, og det er vanskelig å finne noen automatisk prosedyre for å finne og korrigere feil. I datasettet for nivå 1 punktene har vi fulgt en prosedyre som i stor grad er basert på skjønn, og i noen grad på en statistisk modell. Den bruker først og fremst dataene for totalt antall biler pr. time (summen av alle lengdeklassene), og i noen grad data for lengdeklassen 0-5.5m. Dataene for de andre lengdeklassene brukes ikke, da det ville blitt svært tidkrevende. Dog er det nok ønskelig at prosedyren utvides, slik at hver enkelt lengdeklasse blir kontrollert. Prosedyren består av følgende fire punkter:

- i) Alle observasjoner hvor det for lengdeklassen 0-5.5m er registrert 0 biler pr. time er fjerna. For andre lengdeklasser er 0 biler pr. time ingen umulighet, men for den minste og vanligste lengdeklassen vil det i praksis aldri passere så få biler pr. time i en normalsituasjon.
- ii) Deretter er ulike størrelser avleda fra total trafikk (summen av alle lengdeklasser) plotta mot tid. Slike størrelser kan for eksempel være trafikken i 8. time (kl. 7-8) delt på trafikken i 16. time (kl 15-16). Dataene for ulike atypiske perioder kan korrigeres ved at de forskyves i tid, eventuelt sløyfes de.
- iii) For noen av veiene er det registrert svært lav trafikk noen netter. Det skyldes antakelig veiarbeid. For å luke ut de fleste av disse har vi sløyfa alle målinger hvor totalt antall biler pr. time er lavere enn en nedre verdi. Denne minimumsverdien er fastsatt subjektivt for hver enkelt vei, og er angitt i Tabell 13. Dette punktet i prosedyren kunne forøvrig gjerne vært gjennomført samtidig med, eller i stedet for, punkt i).
- iv) Dataene er nå av en såpass brukbar kvalitet at det er meningsfylt å tilpasse en statistisk modell, som i første omgang brukes til å luke ut flere feil. Utgangspunktet er igjen totalt antall biler pr. time. En regresjonsmodell (se kapittel 5.1) tilpasses ved hjelp av redusert rang regresjon med 15 basiskurver. Deretter beregnes residualene, som er avvik mellom den systematiske del av modellen og målingene. Standardavviket til residualene beregnes for hver veilenke, og alle observasjoner hvor residualen er større enn 5 standardavvik sløyfes. Verdien 5 er bestemt subjektivt ut fra tidsplott av residualene. Dette er en grov prosedyre, som for eksempel ikke tar hensyn til at standardavviket kan avhenge av tid på døgnet.

Dataene som er igjen etter denne prosedyren, danner grunnlaget for analysen som er gjengitt tidligere. Et eksempel på bruk av prosedyren finnes i Aldrin (1995 b).



**Tabell 13** Nedre grense for totalt antall biler pr. time i det korrigerede datasettet.

---

<b>Veinavn</b>	<b>Nedre verdi</b>
RV 4, Tr.h.vn, Ammerud	11
RV 190, Ekeberg tunnelen	21
E18, Fiskvollbukta	6
E18, Frognerstranda	21
RV 150, Ringveien, Gaustad	6
RV 150, Granfosstunnelen	6
E6, Karihaugen	6
RV 150, Ringveien, Løren	21
E6, Ringveien, Manglerud	11
E18, Maritim	21
E6, Mortensrud	6
RV 150, Ringveien, Nydalen	6
E18, Fjellinja	21
RV 162, Vaterlandstunnelen	21
RV 161, Vøyen bru	6
RV 168, Wergelandsveien	11

---

## 7 Beregninger basert på et simuleringseksperiment

Vi betrakter en enkelt veilenke, eller et veilenkepar, da de to lenkene i et veilenkepar vanligvis telles samtidig. Vi er interessert i anslag på antall biler på timesbasis gjennom et helt år, og over året som helhet målt som ÅDT. Det vil typisk foreligge tellinger i en relativt kort tidsperiode for dette veilenkeparet. Videre har vi bortimot kontinuerlige tellinger for et mindre antall veilenker (nivå 1 punktene).

Vi har utført et eksperiment som simulerer denne situasjonen, basert på virkelige trafikkdata fra nivå 1 punktene. Detaljene i simuleringseksperimentet beskrives i de neste delkapitlene, og kapittel 7.1 illustrerer hovedideen. Hensikten med simuleringseksperimentet er å besvare følgende spørsmål:

- Hvor mange basiskurver skal en ta i bruk for et gitt tellemønster og et gitt antall telletimer?
- Hvor stor usikkerhet blir det i estimater for ÅDT og trafikkarbeid?
- Hvordan bør tellemønsteret være?

- Hvor god er den nye metoden basert på kombinasjon av basiskurver i forhold til den nåværende metoden basert på variasjonskurver?

## 7.1 Prediksjon basert på basiskurver

På forhånd er ÅDT beregna for hver veilenke, og hvert enkelt år. Det er gjort på grunnlag av alle tilgjengelige data og er gjengitt i Tabell 12. De ÅDT-beregninger som er basert på tellinger minst 30% av året vil være relativt sikre, og vil i simuleringstudien spille rollen som sann ÅDT. Dette er altså fasiten.

Deretter utføres første del av simuleringseksperimentet slik:

- To veilenkepar velges ut. Disse skal illustrere veilenker med få telletimer. Ut fra de resterende 28 veilenkene beregnes 8 basiskurver som angitt i kapittel 5.1.
- Deretter utføres følgende eksperiment for hver av de fire veilenkene: For et gitt år (1994, 1995 eller 1996) med minst 30% data velges det ut et lite antall telletidspunkter innafor en 2-ukersperiode. Telletidspunktene velges tilfeldig ut fra en prosedyre beskrevet i kapittel Appendix A. Bare tidspunkter det finnes virkelige tellinger for kan bli valgt. Tellingene som virkelig er registrert på disse tidspunktene spiller rollen som de fåtallige telletimene en har til rådighet. Ut fra disse telletimene beregnes for hver av lengdeklassene ÅDT-estimer med fra 0 til 8 basiskurver etter metoden i kapittel 5.2. ÅDT-estimatene kan sammenlignes med "sann" ÅDT, som er fasiten. For hvert år trekkes det 500 ulike sett av telldata, som vil kunne overlappe hverandre i noe grad. Ved å variere de valgte telletidspunktene innafor samme år, vil ÅDT-estimatene også variere, mens "sann" ÅDT er fast. Samme prosedyre gjennomføres for hvert enkelt år og for hver av de fire veilenkene, dermed vil også "sann" ÅDT variere. Alle trekninger av telletidspunkt er uavhengige av hverandre.
- Deretter er det to nye veilenkepar som velges ut til å illustrere veilenker med få telletimer. Dette gjentas slik at hver veilenke en gang har spilt rollen som en veilenke med få målinger.

Eksperimentet er utført for år med minst 30% data, og det gir 74 sanne ÅDT-verdier som skal predikeres for hver lengdeklasse (se Tabell 12). Hver av disse skal igjen predikeres ut fra de 500 genererte sett med tellinger. Dette gir  $500 \times 74 = 37000$  ulike tellesituasjoner. For hver slik tellesituasjon har vi for hver lengdeklasse 9 ulike prediksjoner basert på 0 til 8 basiskurver, samt en sann ÅDT, slik at prediksjonsfeil kan beregnes. Videre er telletidspunktene for hver tellesituasjon også registrert.

Hvert sett av telletidspunkter er generert etter en prosedyre beskrevet i Appendix A. Hensikten med prosedyren er å generere mange sett med realistiske telletidspunkter innafor en kort tidsperiode (maksimalt to uker). Telletidspunktene kan karakteriseres med ukedag og time på døgnet, og Tabell 14 gir en oversikt hvordan telletidspunktene er fordelt.

**Tabell 14** Statistikk for antall timer talt innen ulike tidsintervaller for de simulerte tellesituasjonene.

Tidspunkt	Minimum	25% percentil	Median	75% percentil	Maksimum	Gjennomsnitt
hverdag 8-9	0	2	3	5	18	3.6
hverdag 10-15	0	3	7	12	49	8.4
hverdag 16-17	0	2	3	5	18	3.6
hverdag 7,18-24	0	0	1	8	61	5.4
hverdag 1-6	0	0	1	6	42	4.1
lørdag 10-24	0	0	0	10	30	5.1
lørdag 1-9	0	0	0	0	18	0.9
søndag 10-24	0	0	0	9	30	4.9
søndag 1-9	0	0	0	0	18	0.9
Totalt						36.9

## 7.2 Etablering av ligninger for valg av antall basiskurver

For hver lengdeklasse og for hver  $k$  (antall basiskurver) mellom 0 og 8 har vi 37000 prediksjonsfeil  $e_k = \text{ÅDT} - \text{ÅDT}(k)$  etter eksperimentet beskrevet i kapittel 7.1. Ut fra dette lager vi modeller for forventet kvadratisk prediksjonsfeil  $E(e_k^2)$  som funksjon av lengdeklasse, antatt trafikkmengde, telletidspunkt og antall basiskurver  $k$ , slik som angitt i ligning (11). Koeffisientene estimeres via følgende modell på logaritmiske data

$$\log(e_k^2 + 1) = \gamma_{0,lk} + \gamma_{1k} \log(z_1) + \dots + \gamma_{10k} \log(z_{10}) + \varepsilon \quad \begin{matrix} l = 1, \dots, 5 \\ k = 0, \dots, 8 \end{matrix} \quad (16)$$

der forklaringsvariablene er presentert i Tabell 2.

Koeffisientene i denne modellen estimeres ved minste kvadraters metode. Deretter transformeres modellen til formen i ligning (11). Fordi feilen  $e_k$  teoretisk sett kan være 0, er et 1-tall lagt til på logaritmisk form. Som i kapittel 5.2 ser vi bort fra dette ved tilbaketransformering. I motsetning til hva som var tilfelle i kapittel 5.2, har dette få negative konsekvenser, fordi feilene stort sett er langt større enn 1. Når det gjelder konstanten  $c_{lk}$  i ligning (11) er denne her estimert ved  $\hat{c}_{lk} = \exp(\hat{\gamma}_{0,lk} + (1/2)\hat{\sigma}^2)$ , hvor  $\hat{\gamma}_{0,lk}$  er minste kvadraters estimat for  $\gamma_{0,lk}$  og  $\hat{\sigma}^2$  er det ordinære estimat for variansen til  $\varepsilon$  i (16). Dette uttrykket baserer seg på at  $e_k^2$  er lognormal, noe som ikke er tilfelle. Likevel har vi valgt å bruke denne formen da dette ga svakt bedre resultater enn det alternative estimatet som vi har brukt i ligning (10) i kapittel 5.2 og i kapittel 7.3. Merk at hensikten med disse ligningene er å velge  $k$ , og ikke først og fremst å estimere korrekt nivå på  $E(e_k^2)$ . Resultatene av disse beregningene er presentert Tabell 6 i kapittel 5.3.

### 7.3 Etablering av ligninger for beregning av usikkerhet

Vi har nå etablert en metode for å beregne antall basiskurver  $k$  som funksjon av bl.a. tilgjengelige telletidspunkt. For å gjennomføre en rettførdig test av hvordan metoden virker i praksis gjør vi følgende utvidelse av simuleringseksperimentet: To veilenkepar velges ut til å illustrere veilenker med få målinger, og koeffisientene i Tabell 6 reestimeres på grunnlag av dataene fra de resterende 28 veilenkene (dvs. et litt mindre datasett). For hver enkelt av de simulerte tellesituasjonene for de fire valgte veilenkene velges nå  $k$  mellom 0 og 8 for hver enkelt lengdeklasse v.h.j.a. ligning (11) med reestimerte koeffisienter. Dette gjentas slik at hver veilenke en gang har spilt rollen som en veilenke med få målinger. Det blir altså valgt ut en  $k$  for hver av de 37000 tellesituasjonene for hver lengdeklasse. Etter dette står vi igjen med et datasett med *en* prediksjon for hver av de 37000 tellesituasjonene for hver lengdeklasse, sammen med tilhørende sann ÅDT og telletidspunkt.

Vi kan dermed beregne *en* prediksjonfeil  $e_d$  for hver av de 37000 tellesituasjonene for hver lengdeklasse. Ut fra dette har vi estimert koeffisientene i ligning (12), og disse gjengitt i Tabell 9 i kapittel 5.4. Tilsvarende som i kapittel 7.2 har estimeringa foregått via en modell for  $\log(e_d^2 + 1)$ , og vi har igjen sett bort fra 1-tallet ved tilbaketransformering. Konstanten  $c_l$  er estimert ved

$$\hat{c}_l = \frac{1}{37000} \sum_{i=1}^{37000} \frac{e_i^2}{z_{1i}^{\hat{\gamma}_{1l}} \cdot \dots \cdot z_{10i}^{\hat{\gamma}_{10l}}} \quad l = 1, \dots, 5, \quad (17)$$

hvor tellesituasjonene er nummerert med indeksen  $i$ . Dette tilsvarer ligning (10) i kapittel 5.2

### 7.4 Beregning av korrelasjoner

I kapittel 5.5 hevdet vi at ÅDT-estimer basert på samtidige tellinger kan være positivt korrelerte, og vi presenterte i Tabell 11 en korrelasjonsmatrise mellom ÅDT-estimer for de fem lengdeklassene basert på samtidige tellinger. I kapittel 7.3 etablerte vi et datasett med *en* prediksjonsfeil for hver av de 37000 tellesituasjonene for hver lengdeklasse. For hver tellesituasjon er tellingene for de fem lengdeklassene gjort samtidig. La  $e_{li}$  betegne feil for  $i$ -te tellesituasjon for lengdeklasse  $l$ . For hver tellesituasjon kan vi så beregne forventet kvadratisk feil  $E(e_{li}^2)$  ut fra ligning (12) og Tabell 9. Fordi forventningsskjevheten er svært liten (se kapittel 7.5) er dette også et estimat for variansen. Dermed kan korrelasjonen mellom ÅDT-estimatene for lengdeklasse  $l$  og lengdeklasse  $l'$  beregnes som

$$\rho_{ll'} = \frac{1}{37000} \sum_{i=1}^{37000} \frac{e_{li} \cdot e_{l'i}}{\sqrt{E(e_{li}^2)} \cdot \sqrt{E(e_{l'i}^2)}}. \quad (18)$$

Vi har undersøkt om korrelasjonene varierer med tellemønsteret slik som usikkerheten gjør det, men så er ikke tilfelle.

En kan også tenke seg at samtidige tellinger på de to veilenkene i et veilenkepar vil gi positivt korrelerte ÅDT-estimer. I simuleringeksperimentet vi hittil har presentert har disse tellingene ikke vært samtidige. For å undersøke dette spørsmålet har vi derfor utført et nytt, tilsvarende simuleringseksperiment hvor trekkingsmekanismen for telletidspunkt er modifisert slik at tellingene på de to veilenkene i samme par er samtidige. 100 tellesituasjoner er simulert pr. veilenke. De aktuelle korrelasjonene er mellom ÅDT-estimatene for de 5 lengdeklassene i den ene veilenka og de 5 lengdeklassene i den andre veilenka. De 25 estimerte korrelasjonene ligger alle mellom -0.09 og 0.01, og er dermed ubetydelige. Dermed fører samtidige tellinger på de to veilenkene i et veilenkepar ikke til korrelasjoner i ÅDT-estimatene.

Ved tellinger i kryss vil en også typisk ha samtidige tellinger. Vi har ikke undersøkt dette, da vi ikke har noen kryss i datasettet vårt. Men siden trafikk inn i krysset langs en vei vil gå ut av krysset en annen vei, er det rimelig å anta at det vil være visse positive korrelasjoner mellom disse estimatene.

## **7.5 Sammenligning med nåværende metodikk basert på variasjonskurver**

I simuleringstudien har vi for hver lengdeklasse estimert ÅDT for 37000 tellesituasjoner ved hjelp av den nye metoden basert på kombinasjon av basiskurver. I tillegg har vi laget estimer ved hjelp av to varianter av metoden basert på variasjonskurver (dagens metode), som skiller seg ved hvordan kurvene er valgt. I den ene varianten brukes de samme kurvene hele tida, nemlig den "midterste" av variasjonskurvene som er presentert i Aldrin og Follestad (1993). I den andre varianten velges års- og ukes-kurvene ved å sammenligne trafikk hverdag i forhold til helg, mens døgnkurven for hverdag velges ved å sammenligne morgenrush med ettermiddagsrush, slik det er anbefalt i Aldrin og Follestad (1993). Om det ikke er tilstrekkelig telletimer til å gjøre disse beregningene, er den "midterste" kurven valgt. For hver metode er det så beregna gjennomsnittlig absolutt feil innen hver lengdeklasse, hvor gjennomsnittet er tatt over alle 37000 tellesituasjoner, og hvor absolutt feil er definert som absoluttverdien til  $(\text{sann } \text{ÅDT} - \text{estimert } \text{ÅDT}) / (\text{sann } \text{ÅDT})$ . Resultatet er gitt i øverste venstre del av Tabell 15, som er den samme som Tabell 3. Hvor stor feil de ulike metodene gir varierer naturlig nok med antall telletimer de er basert på, og midtre og nederste del av Tabell 15 viser tilsvarende tall for henholdsvis få og mange telletimer. Her er få telletimer definert som at det er talt maksimalt 8 timer på hverdager fra og med 7-ende til og med 24-de time, mens mange telletimer er definert som at det er talt minst 40 slike timer. Dette utgjør henholdsvis 10% og 6% av de 37000 tellesituasjonene. En ser at den nye metoden er langt bedre enn begge de to variantene av metoden basert på varia-

sjonskurver. Dette gjelder for alle lengdeklasser og ved både få og mange telletimer.

Et vesentlig trekk ved basiskurvemetoden er muligheten til å velge antall basiskurver ut fra den gitte tellesituasjon. For å undersøke potensialet for forbedringer av denne del av metoden har vi beregna gjennomsnittlig absolutt feil ved andre valg av  $k$ . En mulighet er å alltid bruke samme  $k$ , uavhengig av tellesituasjonen, og tabellen viser hva dette ville gitt for  $k = 0, 1, 2$  eller  $3$ . Tabell 15 viser at det å velge  $k$  ut fra data er bedre enn å bruke en fast  $k$  hele tida. Til slutt har vi for hver enkelt tellesituasjon valgt den  $k$  som gir minst prediksjonsfeil (“ideell”  $k$ ). Dette er det ikke mulig å gjøre i praksis, men det angir en nedre grense for hvor godt det er mulig å gjøre det ved bedre prosedyrer for å velge  $k$ . Vi ser av Tabell 15 at usikkerheten ved  $k$  valgt ut fra data er mye større enn ved “ideell”  $k$ , noe som kan bety at forbedringer er mulig.

**Tabell 15** Gjennomsnittlig absolutt feil i ÅDT-estimer i % av sann ÅDT.

Lengdeklasse	Variasjonskurver		Kombinasjon av basiskurver						
	kurver valgt ut fra data	faste kurver	$k$ valgt ut fra data	$k = 0$	$k = 1$	$k = 2$	$k = 3$	“ideell” $k$	
	Alle tellesituasjoner								
1	9.0	10.3	7.2	11.1	8.4	8.3	7.7	2.6	
2	22.0	21.1	14.2	15.1	15.2	18.0	19.1	6.8	
3	25.8	24.2	15.1	16.6	17.0	17.7	19.9	7.0	
4	38.0	37.0	27.6	28.8	29.4	32.5	34.3	15.6	
5	46.2	45.6	33.2	35.9	35.3	43.5	49.7	20.4	
	Tellesituasjoner med antall timer hverdag 7-24 $\leq 8$								
1	12.1	13.0	9.3	15.5	11.8	12.0	12.7	3.5	
2	25.1	25.5	18.4	19.5	19.7	24.9	28.0	8.3	
3	28.6	28.7	21.2	22.5	23.9	26.8	32.9	9.5	
4	48.1	48.4	34.5	36.8	34.4	41.0	44.8	18.4	
5	63.5	63.8	40.5	45.2	40.6	54.0	62.9	25.1	
	Tellesituasjoner med antall timer hverdag 7-24 $\geq 40$								
1	6.9	7.7	5.4	8.2	6.4	6.2	5.3	2.1	
2	18.7	17.2	11.8	12.4	12.0	12.7	13.2	5.9	
3	23.0	20.9	11.8	13.5	12.4	11.7	13.0	5.6	
4	31.5	30.1	24.3	25.1	25.9	27.2	27.4	15.2	
5	35.0	33.6	27.5	29.0	30.0	35.3	38.2	17.6	

For lengdeklasse 1 ga den beste variasjonskurvemetoden 9.0% feil i gjennomsnitt over alle simulerte tellesituasjoner, mens den nye basiskurvemetoden ga 7.2% feil. Dette kan synes som en liten gevinst ved å ta i bruk en mer kompleks metode. Men en annen måte å bedømme hvor stor forskjell det er mellom disse to usikkerhetstillene er å beregne hvor mange flere telletimer det trengs for å minske usikkerheten fra 9.0% til 7.2%. I mange vanlige statistiske beregninger vil usikkerheten minke med  $1/\sqrt{n}$ , der n er antall telletimer. I vår situasjon vil imidlertid usikkerheten synke atskillig langsommere av to grunner. For det første vil tellingene nær hverandre i tid være høyt positivt korrelert. For det andre vil det uansett antall telletimer være en forskjell mellom vår forenklete modell og virkeligheten. Dette siste gjelder i enda større grad for variasjonskurvemetoden enn for basiskurvemetoden, siden variasjonskurvemetoden er mye mindre fleksibel. Når det gjelder metoden basert på basiskurver kan vi bruke ligning (12) og Tabell 9 for å beregne hvor fort usikkerheten synker. La oss ta for oss en veilenke i lengdeklasse 1 med  $\text{ÅDT}=48.000$  biler pr. døgn. Om vi har talt 10 timer fordelt på 2 timer på hverdag 8-9, 3 timer på hverdag 10-15 og 2 timer på hverdag 16-17 gir det en usikkerhet på 9.0% (målt som variasjonskoeffisient som er proporsjonal med forventet absolutt feil). Om vi 6-dobler antall telletimer blir usikkerheten 7.2%. På denne måten kan forskjellen mellom 9.0% og 7.2% ansees som svært stor.

I kapittel 5.4 viste vi i et eksempel at det lønner seg å spre tellingene slik at man f.eks. både teller hverdager og helger. Men ved manuelle tellinger kan det være vanskelig å få talt helger eller om natta. For den delmengden av de 37000 tellesituasjonene hvor det kun er talt på hverdager mellom time 7 og time 24 (10% av tellesituasjonene) har vi derfor beregnet Tabell 16 som tilsvarer Tabell 15. For lengdeklassene 2 til 5 er den nye basiskurvemetoden bedre enn variasjonskurvemetoden. Men for lengdeklasse 1 og med få telletimer er den nye og (den ene varianten av) den gamle metoden jevn gode. Likeledes ser vi at det da stort sett lønner seg å velge  $k=0$ , noe vi også viste i et eksempel i kapittel 5.3. Dette peker hen mot følgende modifikasjon av basiskurvemetoden: Hver basiskurve kan dekomponeres i deler som har å gjøre med årstid, spesielle dager, ukedag og tid på døgnet. Hvis det finnes tellinger for både hverdag og helg gir dette informasjon også om årstidsvariasjonen og om spesielle dager (se Aldrin og Follestad 1993), og det er rimelig å bruke nåværende variant av basiskurvemetoden. Om det derimot kun er observert på hverdager har man liten informasjon om variasjon mellom hverdag og helg og over året. Det kan da være rimelig å bruke en felles basiskurve ( $k=0$ ) for variasjon over år og døgnvariasjonen i helgene, men kombinere den komponent av basiskurvene som har å gjøre med døgnvariasjon på hverdager (jfr. dekomposisjon i Figur 1).

**Tabell 16** Gjennomsnittlig absolutt feil i ÅDT-estimer i % av sann ÅDT. Kun tellesituasjoner med tellinger hverdag 7-24.

Lengdeklasse	Variasjonskurver		Kombinasjon av basiskurver						
	kurver valgt ut fra data	faste kurver	$k$ valgt ut fra data	$k = 0$	$k = 1$	$k = 2$	$k = 3$	“ideell” $k$	
	Alle tellesituasjoner med tellinger kun hverdag 7-24								
1	9.1	10.4	8.6	8.5	14.1	16.0	16.5	3.6	
2	22.4	20.7	15.5	15.5	27.7	43.2	45.3	9.0	
3	27.6	24.9	14.5	14.5	34.2	43.2	42.7	7.8	
4	37.9	35.7	29.4	29.4	47.3	62.6	66.9	18.4	
5	43.8	41.5	37.8	37.8	53.1	83.0	93.9	26.4	
	Tellesituasjoner med tellinger kun hverdag 7-24, antall timer $\leq 8$								
1	10.3	11.2	10.4	10.4	32.5	37.8	48.3	6.8	
2	23.2	22.6	20.0	20.0	39.8	75.1	95.3	14.6	
3	27.5	26.9	18.7	18.7	56.6	78.1	86.3	13.8	
4	38.5	37.3	36.5	36.5	63.8	107.7	118.3	25.3	
5	53.7	53.3	53.4	53.3	71.3	144.4	170.8	42.4	
	Tellesituasjoner med tellinger kun hverdag 7-24, antall timer $\geq 40$								
1	6.3	6.9	5.5	5.6	6.3	6.2	5.3	2.0	
2	18.6	16.1	11.0	11.0	14.1	17.2	16.4	5.7	
3	23.8	20.8	10.6	10.6	13.6	14.1	13.5	4.6	
4	34.2	31.8	22.3	22.3	28.8	30.0	30.7	14.5	
5	35.1	32.7	24.6	24.5	32.1	41.0	44.4	15.9	

Til slutt har vi regna ut gjennomsnittlig feil (altså uten å ta absoluttverdien) for hver lengdeklasse for den nye metoden. Disse er gjengitt i Tabell 17. Vi ser at gjennomsnittlig feil er svært liten i forhold til gjennomsnittlig *absolutt* feil. Vi konkluderer med at en eventuell forventningsskjevhet er uten betydning i forhold til den generelle usikkerheten. Videre er da estimer for forventa kvadratisk feil i praksis også estimer for variansen.



**Tabell 17** Gjennomsnittlig feil i % av sann ÅDT.

Lengdeklasse	Kombinasjon av basiskurver, $k$ valgt ut fra data
1	-0.7
2	-1.1
3	0.0
4	-3.0
5	-2.7

## 8 Mulige forbedringer

Det vil alltid være rom for forbedring av en metode, og på sikt kan man tenke seg videreutviklinger. Aktuelle forbedringer av metodikken beskrevet i denne rapporten inkluderer følgende.

- *Implementering.*
- *Rekalibrering på representative data fra Norge*, f. eks. for tre typer veier: storby, småby, land.
- *Rekalibrering for et generelt tellemønster*, dvs. også for lenge telleperioder. Når antall talte timer utgjør en stor andel av et år må det tas hensyn til at antall timer i et år er endelig.
- *Flere veier i datasettet, spesielt noen litt mindre veier.* Nivå 1 punktene består av hovedveier, og danner i grunnlag for beregning av basiskurvene og for simuleringstudien. I det foreliggende arbeidet er det antatt at nivå 1 punktene til en viss grad er representative for mindre veier. Dette er ikke helt urimelig, da hovedtrekkene i variasjonen over et døgn, en ukedag og et år er felles for alle veier. Likevel vil resultatene i større grad kunne generaliseres om mindre veier inkluderes i datasettet.
- *Kalibrering for totalt antall biler* i tillegg til 5 lengdeklasser, fordi noen steder telles bare totalt antall. Evt også en egen kurve for tunge biler, dvs. summen av de fire største lengdeklassene.
- *Kombinasjon kun av noen komponenter av basiskurvene.* I kapittel 7.5 diskuterte vi en variant av basiskurvemetoden hvor kun en kombinerte noen komponenter av basiskurvene, og ikke hele. Dette er spesielt aktuelt hvis man kun har trafikktegninger på hverdager, hvor det kan lønne seg å kun kombinere den komponent av basiskurvene som har å gjøre med døgnvariasjon på hverdager
- *Bedre metoder for å kombinere basiskurvene.* I kapittel 5.2 brukte vi en multiplikativ modell (8) for å kombinere basiskurvene, men gikk via en additiv modell på logaritmisk skala for å estimere hvor stor vekt hver basiskurve skulle ha.

Dette er ikke helt tilfredstillende for veilenker med liten trafikk, og kan forbedres. En mulighet er å bruke såkalte generaliserte lineære modeller med Poisson- eller gamma-fordeling og logaritmisk link-funksjon. En annen mulighet er å estimere den modellen ved veid minste kvadraters metode, som i dette tilfelle vil kreve ikke-lineær optimering.

- *Myk nedvekting (v.h.j.a. ridge regresjon) av basiskurvene* i stedet for et fast antall. Dette kan generelt gi mer presise trafikkberegninger og bedre usikkerhetsberegninger. Konkret betyr dette at alle basiskurver får koeffisienter ulik 0, men disse koeffisientene "krympes" mot 0 hvis det er få talletimer. Graden av krymping spiller samme rolle som antall komponenter i den nåværende versjon av metoden.
- *Bedre metoder for valg av antall basiskurver.* Valg av antall basiskurver er vesentlig, og i kapittel 7.5 viste vi at det kan være et potensiale for forbedringer. I kapittel 5.3 er modell (12) estimert slik at koeffisientene  $\gamma_{2k}$  til  $\gamma_{10k}$  er felles for alle lengdeklasser, og det er mulig dette er for restriktivt. Derimot kan det tenkes at andre typer restriksjoner kan være gunstig, f. eks. at  $\gamma_{2k}$  til  $\gamma_{10k}$  pålegges å synke med økende  $k$ . En annen forbedringsmulighet kan være å estimere koeffisientene i (12) ved ikke-lineære metoder, jfr. forrige punkt. Dette gjelder også for modell (11) for estimering av usikkerhet.
- *Bruk av basiskurvemetoden i kombinasjon med andre trafikkmodeller som tar hensyn til veinettet.* Basiskurvemetoden tar ikke i bruk informasjon om hvordan veier henger sammen, f. eks. at antall biler inn i et kryss må være det samme som biler ut av et kryss.
- *Estimering av frekvensfordeling av trafikkmengde pr. time for alle timene i et år.* Ut fra dette skal det kunne beregnes f. eks. dimensjonerende time.
- *Usikkerhetsberegninger for estimerer for YDT og HDT.*

### **Referanser:**

Aldrin, M. og Follestad, T. (1993), "Nye variasjonskurver for Oslo med statistisk vurdering av telleopplegg", PROSAM rapport nr. 30.

Aldrin, M. (1995 a), "A statistical approach to the modelling of daily car traffic," *Traffic Engineering + Control*, **36**, 489-493.

Aldrin, M. (1995 b), "Metoder for å beregne trafikkdata til overvåkingssystemet for luftforurensning og støy i Oslo", NR-notat, STAT/13/95.

Davies, P.T. og Tso, M.K.-S. (1982), "Procedures for reduced-rank regression", *Applied Statistics*, **31**, s. 244-255.

## Appendix A Trekking av telletidspunkter

Det telles i 1 eller 2 påfølgende uker. Juleuka og påskeuka er holdt utafør, og det blir heller ikke talt på helligdager som 17.mai og Kr.himmelfartsdag.

### I. Ukenummer

Først trekkes et ukenummer uniformt mellom 1 og 50 som angir den første uka det skal telles i.

### II. Hverdager.

Før trekking av telletidspunktene, genereres en tilfeldig rekkefølge på hverdagene i hver av de to aktuelle ukene. Når det deretter trekkes ut n dager i ei uke, brukes de første n dagene i den tilfeldige ordna rekka av dager.

Timene på hverdagene deles i fem timegrupper:

- 1) Morgen - timene 7-8
- 2) Formiddag - timene 9-15
- 3) Ettermiddag - timene 16-17
- 4) Kveld, og tidlig morgen - timene 18-24 og 7
- 5) Natt - timene 1-6

For hver av de fem timegruppene trekkes først antall uker, med 80% sannsynlighet for 1 uke og 20% sannsynlighet for 2 uker. For hver uke trekkes deretter antall hverdager, og videre trekkes det for hver av disse dagene antall timer og hvilke timer det skal telles i.

Antall timer og hvilke timer det skal telles i trekkes uniformt. Men ved trekkingen av antall dager, er det gitt ulik sannsynlighet for å trekke 0,1,2,3,4 eller 5 dager. For timegruppe 1,2 og 3 er sannsynligheten  $1/9$  for 0,  $3/9$  for 1,  $2/9$  for 2 og  $1/9$  for 3,4 eller 5 dager med tellinger. For timegruppe 4 og 5 (kveld og natt) er sannsynligheten  $9/18$  for 0,  $4/18$  for 1,  $2/18$  for 2 og  $1/18$  for 3, 4 eller 5 dager. For disse timegruppene er det altså 50% sannsynlighet for ikke å trekke noen timer.

### III.Lørdager og søndager.

Telletidspunktene for lørdager og søndager trekkes hver for seg, men etter samme prosedyre.

For disse to dagene blir døgnet delt i to timegrupper:

- 1) Dag - timene 10-24
- 2) Natt - timene 1-9

For hver timegruppe trekkes først antall uker, som i dette tilfellet er lik antall dager. For helgedagene settes det, i motsetning til for hverdagene, en positiv sannsynlighet for å trekke 0 uker, og det trekkes 0, 1 eller 2 uker med sannsynlighet på henholdsvis 50%, 35% og 15% for timegruppe 1 (dag) og 85%, 10% og 5% for timegruppe 2 (natt).

Deretter trekkes antall timer og hvilke timer det skal telles i for hver av de to timegruppene. Dette gjøres, som for hverdagene, uniformt.

Dersom punktene I.-III. resulterer i færre enn to timer med tellinger, forkastes trekkinga og en ny trekking foretas. Videre forkastes trekkinga, og en ny foretas, hvis det ikke finnes virkelige tellinger for alle de utrukne timene.