



Evaluering av flomvarslingsmodellverktøy

84
2014



R
A
P
P
O
R
T

Evaluering av flomvarslings modellverktøy

Rapport nr 84-2014

Evaluering av flomvarslingsmodellverktøy

Utgitt av: Norges vassdrags- og energidirektorat
Redaktør: Ingjerd Haddeland
Forfattere: Thomas Væringstad, Elin Langsholt og Thomas Skaugen

Trykk: NVEs hustrykkeri
Opplag: 25
Forsidefoto: Målestasjon i Sula. Foto: Ingjerd Haddeland, NVE
ISBN 978-82-410-1034-7

Sammendrag: Resultater fra 98 nedbørfelt simulert med HBV, ODM og DDD modellen er vurdert. DDD har noe bedre treff og færre falske alarmer enn HBV og ODM for alternativene der det ikke tillates en tidsforskyvning mellom observerte og simulerte flomtopper. Generelt scorer modellene bedre når man tillater en tidsforskjell i flomtoppen på et døgn, og i dette tilfellet er resultatene fra HBV litt bedre enn de to andre modellsystemene. Analysene indikerer også at det er rom for forbedringer når det gjelder modellenes evne til å treffe flomtoppene.

Emneord: Flomvarsling, hydrologisk modellering, evaluering

Norges vassdrags- og energidirektorat
Middelthunsgate 29
Postboks 5091 Majorstua
0301 OSLO

Telefon: 22 95 95 95
Telefaks: 22 95 90 00
Internett: www.nve.no

Desember 2014

Innhold

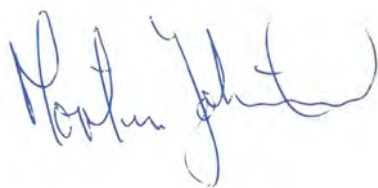
1	Innledning	6
2	Bakgrunn.....	6
3	Resultater	8
3.1	Nash-Sutcliffe og bias	8
3.2	Taylordiagram	9
3.3	Treff, falske alarmer og suksessindeks	10
4	Konklusjoner.....	13
	Referanser	15
	Vedlegg A: Figurer	16
	Vedlegg B: Tabeller	22
	Vedlegg C: Eksempler fra våre naboland	28

Forord

Flomvarslingens hydrologiske modellpark er utviklet og utvidet kraftig de siste 20 årene. For flomvarslingen er det flomsituasjonene som er de mest kritiske, og det er da man har størst behov for at modellene fungerer godt. Hittil er tradisjonelle kalibreringsmål som tar for seg hvordan modellen fungerer gjennom alle årets dager brukt for å evaluere modellene. Det har imidlertid de siste årene blitt et sterkere ønske om å bruke mer direkte mål på hvordan modellene fungerer i flomsituasjoner. Denne rapporten presenterer resultatene for en evaluering utført i 2014, og målet er å si noe om hvordan modellene fungerer i flomsituasjoner, gitt at observert nedbør og temperatur er brukt i de hydrologiske simuleringene. Planen er at tilsvarende analyser skal gjennomføres minst hvert annet år slik at eventuelle forbedringer synliggjøres.

Oslo, desember 2014

Morten Johnsrud
avdelingsdirektør



Hege Hisdal
seksjonssjef



Sammendrag

I denne rapporten presenteres et system som tallfester kvaliteten på modellene som inngår i NVEs flomvarslingssystem. Kvantifisering av resultatene er et ledd i å tallfeste hvor godt modellene treffer flomtoppene, samt å måle om modellene blir bedre over tid. Evalueringen er utført på grunnlag av hydrologiske modellsimuleringer basert på historiske tidsserier av nedbør og temperatur. Resultater fra 98 nedbørfelt simulert med HBV, ODM og DDD modellen er vurdert. For å evaluere hvordan modellene oppfører seg for de største vannføringene (flomtoppene), er det tatt i bruk kriterier som treffrate og falske alarmer. DDD har noe bedre treff og færre falske alarmer enn HBV og ODM for alternativene der det ikke tillates en tidsforskyvning mellom observerte og simulerte flomtopper. Generelt scorer modellene bedre når man tillater en tidsforskjell i flomtoppen på et døgn, og i dette tilfellet er resultatene fra HBV litt bedre enn de to andre modellsystemene. Analysene indikerer også at det er rom for forbedringer når det gjelder modellenes evne til å treffe flomtoppene.

1 Innledning

Modellsimuleringer av vannføring stemmer ikke alltid med observert vannføring, og evaluering av hydrologiske simuleringsresultater er nødvendig for å identifisere områder der simuleringene kan forbedres. Her presenteres et system som tallfester kvaliteten på modellene som inngår i NVEs flomvarslingssystem for hele Norge, men som også gir informasjon om hvert enkelt modellfelt. Evalueringen er utført på grunnlag av hydrologiske modellsimuleringer basert på historiske tidsserier av nedbør og temperatur. Fokuset er på de flommene som har blitt observert og/eller simulert i perioden 1961-2011. Modellsimuleringene er utført for nedbørfelt som inngår i NVEs operasjonelle flomvarslingssystem, der for tiden modellsimuleringer basert på HBV modellen (Hydrologiska Byråns Vattenbalansavdelning; Bergström, 1976; Sælthun, 1996), ODM (Open Distributed Model; basert på Beldring (2008)) og DDD (Distance Distribution Dynamics; Skaugen og Onof, 2014) inngår. Det er derfor tre sett med tidsserier som er evaluert for hvert nedbørfelt. Analysene er gjort for de nedbørfelt der resultater eksisterte våren 2014 for alle de tre modellsystemene; det vil si 98 nedbørfelt.

Det er mange årsaker til at simuleringsresultatene ikke er perfekte, blant annet vil inngangsdata (nedbør, temperatur), vannføringskurven og modellen i seg selv være usikkerhetskilder. Det kan være interessant å vite nøyaktig hva som forårsaker at simulert vannføring ikke alltid stemmer med observert vannføring. I denne rapporten er det sluttresultatet som er i hovedfokus, og det er derfor ikke gjort forsøk på å skille på faktorene som påvirker resultatet. Et viktig prinsipp for analysene presentert i denne rapporten er at resultatene for de ulike nedbørfeltene skal kunne sammenlignes. Det er derfor normaliserte resultater som sammenlignes, og for eksempel volumfeil blir angitt i fraksjon av årlig vannføring.

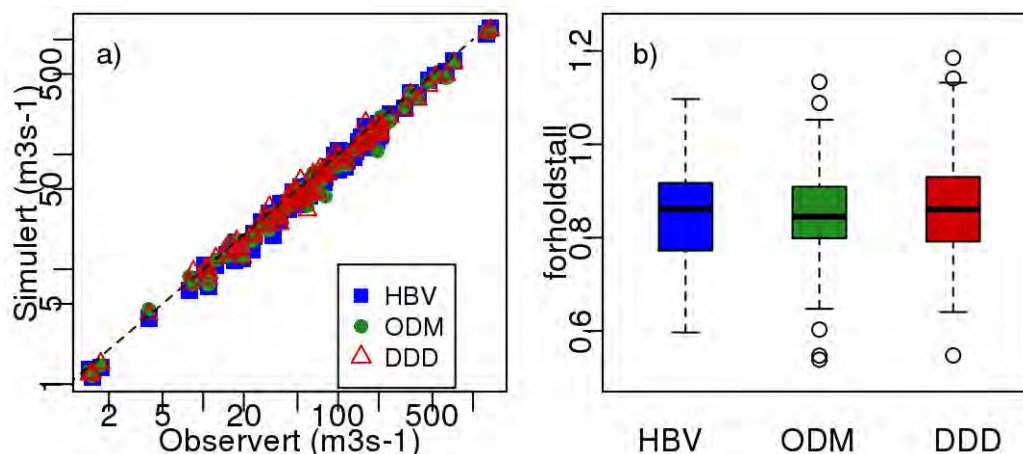
I denne rapporten blir ordet «modellsystem» brukt om henholdsvis HBV, ODM og DDD. Ordet «modell» blir brukt om et modelloppsett med et modellsystem for et nedbørfelt. Som nevnt over er det her modellresultater basert på historiske tidsserier (observasjoner) av nedbør og temperatur som er evaluert. Vannføringsprognosene er evaluert og presentert i Langsholt m fl. (2014).

2 Bakgrunn

Det finnes mange metoder for å evaluere hydrologiske modellers evne til å simulere vannføring. I NVEs flomvarsling brukes Nash-Sutcliffe (NS; Nash og Sutcliffe, 1970) og bias når HBV- og ODM-modellene kalibreres, mens DDD-modellene blir kalibrert kun på NS-verdier. I denne rapporten presenteres de tradisjonelle valideringskriteriene etter siste kalibreringsrunde (vinteren 2014), sammen med en generell evaluering av modellparken i form av Taylordiagram (Taylor, 2001). For å evaluere hvordan modellene oppfører seg for de største vannføringene (flomtoppene), er det også tatt i bruk kriterier som treffrate og falske alarmer. Teorien bak evalueringskriteriene blir presentert sammen med resultatene i kapittel 3. Hvilke 98 nedbørfelt som inngår i analysene, samt mer detaljerte resultater, finnes i vedlegg A og B. I beregningen av Nash-Sutcliffe, bias og

verdier som ligger til grunn for Taylordiagrammet brukes de observerte og simulerte vannføringsverdiene direkte, mens observerte og simulerte flomstørrelser blir brukt for å beregne treff, bom, falske alarmer og suksessindeks. Fokuset her er på de flommene som har blitt observert og/eller simulert i løpet av perioden 1961-2011.

For flomstørrelsene brukes to verdier: $1.0 \cdot Q_m$ og $1.2 \cdot Q_m$, der Q_m er middelflommen. Observert middelflom er gjennomsnittet av den største observasjonen hvert år i analyseperioden, mens simulert middelflom tilsvarer gjennomsnittet av den høyeste simulerte verdien hvert år. Det er bare dagene med observasjoner i analyseperioden (1961-2011) som er brukt for å regne ut størrelsen på observert og simulert middelflom. Forskjellen på observerte og simulerte middelflommer er illustrert i figur 1. I de aller fleste nedbørfelt er simulert middelflom lavere enn observert middelflom (figur 1 a). Medianverdien av forholdet mellom simulert og observert middelflom er litt over 80 prosent (figur 1b), og det er liten forskjell i resultatene mellom de tre modellsystemene. Et naturlig valg for å evaluere litt større flommer enn middelflommen ville være femårsflommen, som flomvarslingen bruker som kriterium for utsendelse av flomvarsel på oransje nivå (nivå 2 av 3). Imidlertid er bestemmelse av femårsflommen avhengig av valg av frekvensfordeling, og det finnes ikke offisielle størrelser for både simulert og observert femårsflom for alle 98 nedbørfelt. Det er her derfor valgt å bruke $1.2 \cdot Q_m$ som et alternativ til femårsflommen. Ved konsekvent bruk av $1.2 \cdot Q_m$ for både observerte og simulerte vannføringer er det dessuten entydig hva som sammenlignes.



Figur 1: a) Observert og simulert middelflom ($m^3 s^{-1}$, logaritmisk skala) for HBV, ODM og DDD for de 98 nedbørfeltene, og b) Boxplot som viser median, 25- og 75-persentilene (farget del av plottet), samt minimums- og maksimumsverdiene av forholdet mellom simulert og observert middelflom. «Outliers» i datasettet er markert med åpne sirkler.

3 Resultater

3.1 Nash-Sutcliffe og bias

Nash-Sutcliffe (NS) koeffisienten (Nash og Sutcliffe, 1970) blir beregnet på følgende måte:

$$NS = 1 - \frac{\sum(Q_o - Q_s)^2}{\sum(Q_o - \bar{Q}_o)^2} \quad (1)$$

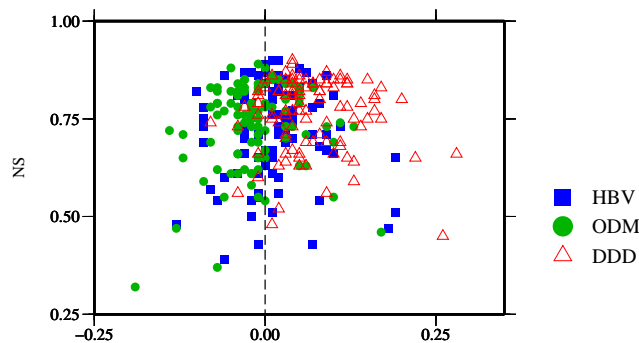
Nash-Sutcliffe koeffisienten kan variere fra $-\infty$ til 1. Dersom NS er 1, tilsvarer dette en perfekt tilpassing av modellert vannføring til observerte data. En NS verdi på 0 indikerer at modellprediksjonene er så nøyaktige som middelverdien av de observerte data. En NS verdi mindre enn null oppstår når gjennomsnittsverdien av observasjonene er en bedre indikator enn modellen, med andre ord når residualvariansen (telleren i ligning 1), er større enn variasjonen (nevneren).

Når det er forskjell i totalt volum observert og simulert vannføring, får man en *bias* i resultatene. Denne blir regnet ut på følgende måte:

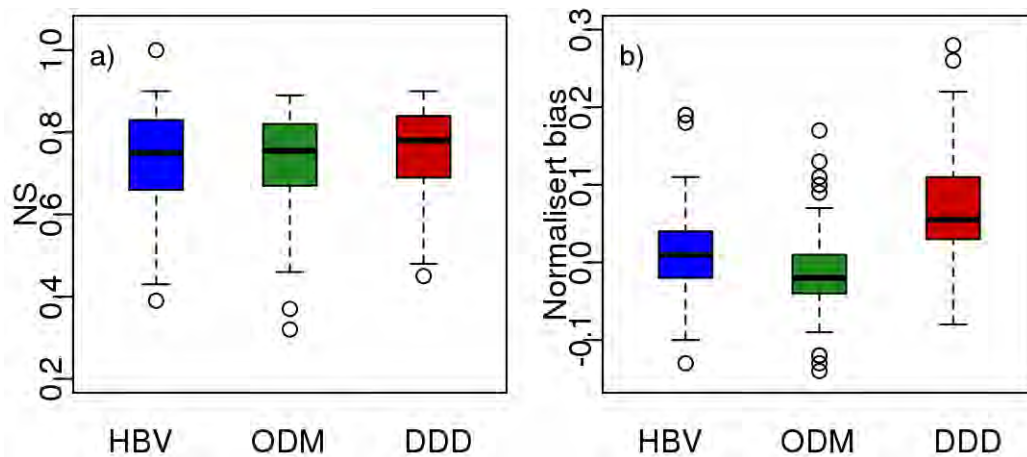
$$Bias = \sum_t (Q_o - Q_s) \quad (2)$$

der Q_o er observert vannføring (m^3s^{-1}) og Q_s er simulert vannføring (m^3s^{-1}). Målet er å ha en bias nærmest mulig 0.

Bias er i denne rapporten angitt som normalisert bias, det vil si forskjell i totalvolum dividert på total observert avrenning i nedbørfeltet i analyseperioden. Dette er gjort for at bias fra de ulike nedbørfelt skal kunne sammenlignes. Figur 2 og 3 viser resultatene på to ulike måter: NS verdier mot bias, og en statistisk presentasjon av NS verdiene og bias. Figur 2 indikerer at det er vanskelig å se en direkte sammenheng mellom NS og bias, selv om de aller største biasene (positive eller negative) er for nedbørfelt med relativt lave NS verdier. Figur 3a viser at totalt sett har DDD noe bedre NS verdier enn HBV og ODM, men dette kan ha sammenheng med at DDD som eneste modellsystem er kalibrert uten å ta hensyn til bias, noe som gjenspeiler seg i biastallene presentert i figur 3b.



Figur 2: NS-verdier (y-akse) og normalisert bias (x-akse; observert – simulert) for HBV, ODM og DDD.



Figur 3: a) Boxplot av NS-verdier og b) normalisert bias for HBV, ODM og DDD for de 98 nedbørfeltene. Positiv bias betyr at volum observert vannføring er større enn volum simulert vannføring.

3.2 Taylordiagram

Ved hjelp av et Taylordiagram (Taylor, 2001) kan man grafisk sammenstille tre ulike kriterier for hvor godt en simulert tidsserie stemmer med observasjoner. Samsvaret mellom de to tidsseriene er kvantifisert i form av deres korrelasjon, kvadratfeilen og amplituden av variasjonene (representert ved standardavviket). Et slikt diagram er spesielt nyttig nå man skal evaluere ulike aspekter ved komplekse modeller eller måle relativ godhet til mange ulike modeller (se for eksempel Covey m fl., 2000).

Taylordiagrammet viser standardavviket for simuleringene i forhold til observasjonene som avstand fra origo. Korrelasjonen mellom simulert og observert vannføring er gitt som cosinus til asimutvinkelen. Observasjonene er plottet ved (1.0), og dette punktet representerer altså ingen feil i standardavviket og null feil i korrelasjonen. Avstanden mellom punktet (1.0) og punktet som representerer simuleringensresultatene er proporsjonal med kvadratfeilen (RMSE). Gode simuleringer av amplituden av tidsserien sammenlignet med observasjonene er representert med et symbol nær den stiplede linja. God representasjon av fasen viser seg ved kort avstand mellom symbolet og 1 på x-aksen.

Gitt en simulert tidsserie f og en observert tidsserie r , er formlene for å beregne korrelasjonskoeffisienten (R), sentrert kvadratavvik (E^2) og standardavviket for de simulerte (σ_f) og observerte (σ_r) tidsseriene gitt som:

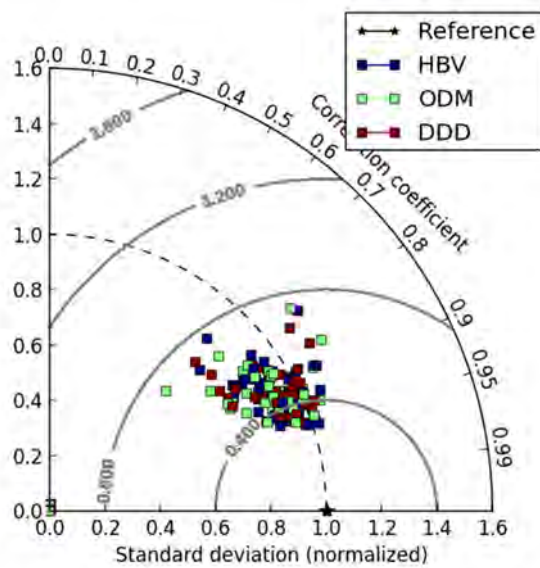
$$R = \frac{\frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})(r_n - \bar{r})}{\sigma_f \sigma_r} \quad (3)$$

$$E^2 = \frac{1}{N} \sum_{n=1}^N [(f_n - \bar{f}) - (r_n - \bar{r})]^2 \quad (4)$$

$$\sigma_f^2 = \frac{1}{N} \sum_{n=1}^N (f_n - \bar{f})^2 \quad (5)$$

$$\sigma_r^2 = \frac{1}{N} \sum_{n=1}^N (r_n - \bar{r})^2 \quad (6)$$

Figur 4 viser at det totalt sett, som for NS og bias, er vanskelig å rangere de ulike modellsystemene basert på den statistiske analysen som presenteres i Taylordiagrammet. Korrelasjonskoeffisientene, standardavviket og RMSE er forholdsvis like når man evaluerer de 98 modellfeltene under ett.



Figur 4: Taylor diagram der hvert nedbørfelt er representert med en firkant; HBV, ODM og DDD.

3.3 Treff, falske alarmer og suksessindeks

Det mest interessante spørsmålet når man skal evaluere modeller brukt for flomvarslingsformål er «hvor godt treffer modellene flomtoppene»? Modellsimuleringer utført i forbindelse med flomvarsling kan evalueres med hensyn på om man treffer flommene. Dette kan beskrives ved følgende størrelser (Stanski m fl., 1989):

$$TR = \frac{T}{T + B} \quad (7)$$

$$FAR = \frac{FA}{T + FA} \quad (8)$$

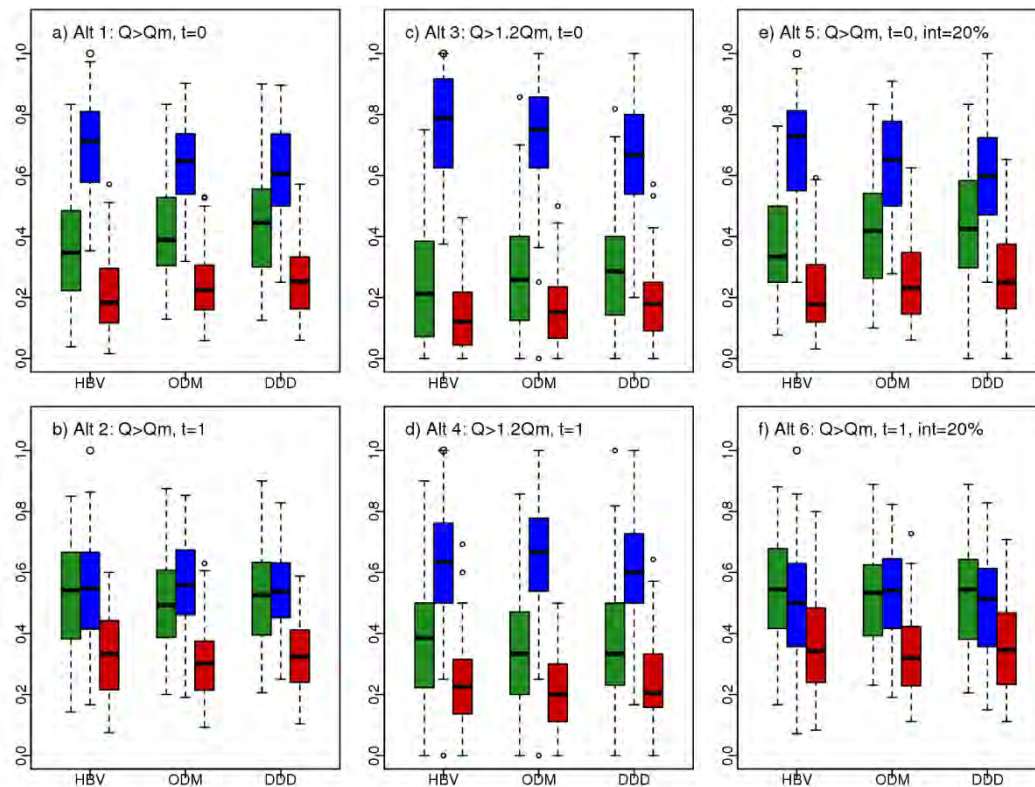
$$SI = \frac{T}{T + FA + B} \quad (9)$$

der T representerer antall treff, B antall bom og FA er antall falske alarmer. *Treffraten* (TR) beskriver forholdet mellom vellykkede simulerte hendelser og alle observerte hendelser. TR varierer fra 0 (dårligst) til 1 (best). Raten av *Falske alarmer* (FAR) beskriver forholdet mellom simulerte hendelser som ikke er observert og alle hendelser. *Suksessindeksen* (SI) er den mest beskrivende av disse fire målene, og varierer mellom 0 og 1 (best). Denne type evaluering blir blant annet brukt i Tsjekkia (Daňhelka og Vlasák, 2002) og i deler av USA (Corby og Lawrence, 2002), men de har noe ulik praktisering når det gjelder å definere hva et treff er. I USA bruker de også flere kategorier enn de som er nevnt ovenfor, blant annet for å framheve resultater innenfor ulike flomvarslingsnivåer (Corby og Lawrence, 2002).

En sammenstilling av alle modellens resultater for treffrate, falske alarmer og suksessindeks er vist i figur 5. For hvert nedbørfelt er bare de dagene i analyseperioden (1961-2011) der det eksisterer observert vannføring tatt med i beregningene. Det er i figur 5 brukt ulike alternativ for å definere «treff», se tabell 1. I alle alternativene sammenlignes simulerte og observerte vannføringer i forhold til døgnmiddelflommen (fraksjoner), alternativt $1.2 \cdot Q_m$. Bommer og falske alarmer er beregnet på tilsvarende måte som «treff», og til sammen er disse tre størrelsene grunnlaget for å beregne suksessindeksen. Perfekt score for treff og suksessindeks er 1, mens perfekt score for falsk alarm er 0. Se også resultatene i form av flere figurer og tabeller i vedlegg A og B.

Tabell 1: Alternative kriterier for å beregne treff.

Alt	Q	T (dager)	Intervall (prosent)	Forklaring
1	$Q > Q_m$	0	0	Simulert og observert vannføring overskrider Q_m på samme dag
2	$Q > Q_m$	1	0	Simulert vannføring overskrider Q_m samme dag eller med en dags forskyvning i forhold til når observert vannføring overstiger Q_m
3	$Q > 1.2 \cdot Q_m$	0	0	Simulert og observert vannføring overskrider $1.2 \cdot Q_m$ på samme dag
4	$Q > 1.2 \cdot Q_m$	1	0	Simulert vannføring overskrider $1.2 \cdot Q_m$ samme dag eller med en dags forskyvning i forhold til når observert vannføring overstiger Q_5
5	$Q > Q_m$	0	20	Simulert vannføring er innenfor ± 20 prosent av observert vannføring på samme dag, gitt at observert vannføring er større enn middelflommen
6	$Q > Q_m$	1	20	Simulert vannføring er innenfor ± 20 prosent av observert vannføring på samme dag eller med en dags forskyvning, gitt at observert vannføring er større enn middelflommen



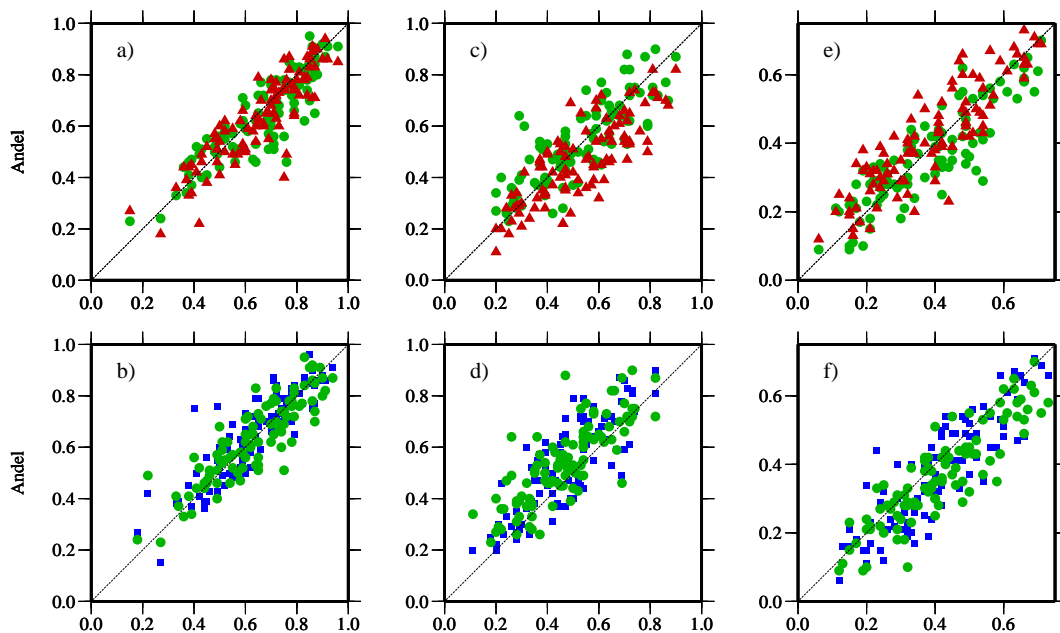
Figur 5: Treffrate (TR; grønn), rate av falske alarmer (FAR; blå) og suksessindeks (SI; rød) gitt de seks alternative kriteriene for "treff". Box-plottene viser TR, FAR og SI for alle nedbørfelt. Alle dager der det finnes observert vannføring i analyseperioden (1961-2011) er med i analysen, gitt at dagens vannføring er større enn gårsdagens vannføring.. Alternativene er forklart i tabell 1.

Figur 5 indikerer at det er relativt liten forskjell i resultatene mellom de ulike modellsystemene. Medianverdiene (se også tabell 2) viser at DDD har noe bedre treff og suksessindeks enn HBV og ODM for alternativene der det ikke tillates en tidsforskyvning mellom observerte og simulerte topper (alternativene 1, 3 og 5). Generelt scorer modellene bedre når man tillater en tidsforskyvning i flomtoppen på et døgn, noe som til en viss grad skyldes forskyvningen som eksisterer mellom nedbørdøgn og vannføringsdøgn (se Engeland, 2014). Det er sannsynligvis også et resultat av noe sen avrenningsrespons i enkelte modeller (Engeland, 2014). Et noe overraskende resultat er at det er så mange falske alarmer i resultatene, noe som bidrar sterkt til lave suksessindekser. Dette kan også til en viss grad ha sammenheng med forskyvningen mellom nedbørdøgn og vannføringsdøgn og en litt for treg avrenningsrespons, men nærmere undersøkelser kreves før man kan konkludere.

Figur 6 viser at det stort sett er slik at dersom et nedbørfelt gir et dårlig eller godt resultat for TR, FAR eller SI i et av modellsystemene, så er det samme tilfelle i de to andre modellsystemene. Figurene i vedlegg A viser at det er en tendens til at treffraten er noe dårligere på Vestlandet enn i de andre regionene i Norge, - sannsynligvis et resultat av forholdsvis små nedbørfelt og stor variasjon i nedbør som gjør at man der kan regne med større utfordringer med hensyn på modellenes drivende data (nedbør og temperatur).

Tabell 2: Medianverdiene fra figur 5. Beste resultat for hvert alternativ og kriterium er markert med uthevet skrift.

Alt	TR			FAR			SI		
	HBV	ODM	DDD	HBV	ODM	DDD	HBV	ODM	DDD
1	0.347	0.389	0.446	0.711	0.648	0.606	0.183	0.225	0.252
2	0.542	0.493	0.523	0.548	0.559	0.537	0.333	0.301	0.324
3	0.212	0.258	0.285	0.788	0.750	0.667	0.121	0.151	0.179
4	0.387	0.333	0.333	0.634	0.667	0.600	0.226	0.200	0.205
5	0.333	0.419	0.424	0.729	0.651	0.598	0.178	0.231	0.250
6	0.544	0.533	0.544	0.500	0.541	0.514	0.343	0.320	0.346



Figur 6: a) TR, b) FAR og c) SI for HBV (x-akse) plottet mot treffrate for ODM (grønn) og DDD (rød), og c) TR, d) FAR og e) SI for DDD (x-akse) plottet mot treffrate for HBV (blå) og ODM (grønn).

4 Konklusjoner

Hydrologiske modeller for 98 nedbørfelt basert på HBV, ODM og DDD modellene er vurdert. Median Nash Sutcliffe verdi for de 98 nedbørfeltene er i underkant av 0.8 for alle tre modellsystemene, noe som regnes som et godt resultat. Det er ikke funnet store forskjeller mellom modellsystemene med hensyn på korrelasjon, kvadratfeil og variasjon i forhold til observerte tidsserier. DDD, i motsetning til HBV og ODM, er ikke kalibrert med tanke på bias. Dette resulterer i noe høyere NS verdi, mens også større bias, for DDD i forhold til HBV og ODM.

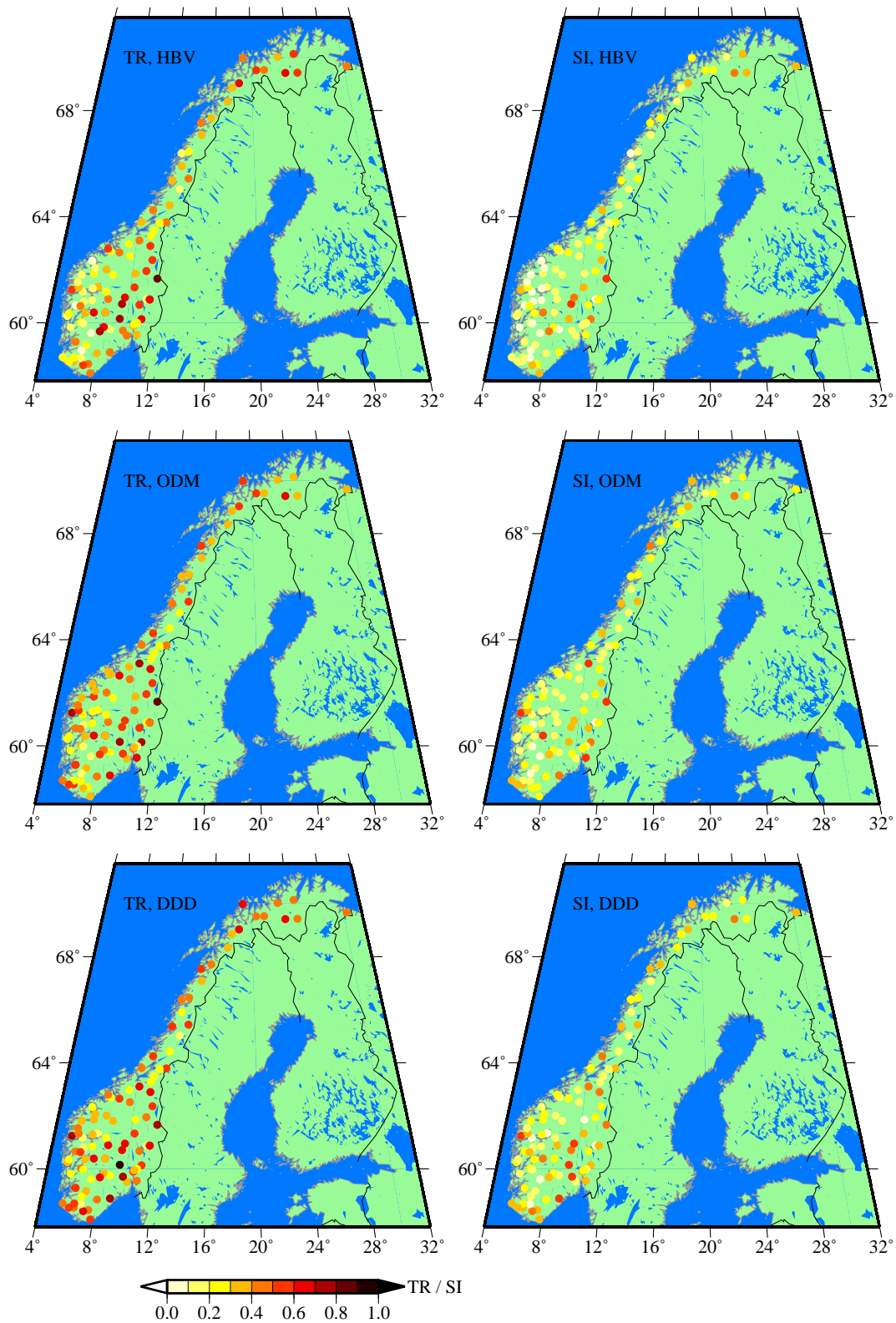
Sammenligninger av observerte og simulerte flomtopper for perioden med observasjoner for hvert av de 98 nedbørfeltene i perioden 1961-2011 er analysert, viser at DDD har noe bedre treff og suksessindeks enn HBV og ODM for alternativene der det ikke tillates en tidsforskyvning mellom observerte og simulerte flomtopper. Generelt scorer modellene bedre når man tillater en tidsforskjell i flomtoppen på et døgn, og i dette tilfellet er resultatene fra HBV litt bedre enn de to andre modellsystemene. Analysene indikerer at det er rom for forbedringer når det gjelder å treffe flomtoppene, og et stort antall falske alarmer bidrar sterkt til lav suksessindeks.

Det anbefales kvartalsvis gjennomgang og eventuelt oppdatering av modellene dersom nye vannføringskurver eller nye inngangsdata gjør at forrige kalibrering ikke er gyldig. Analyser og rapportering tilsvarende denne rapporten skal utarbeides minst hvert annet år slik at eventuelle forbedringer synliggjøres.

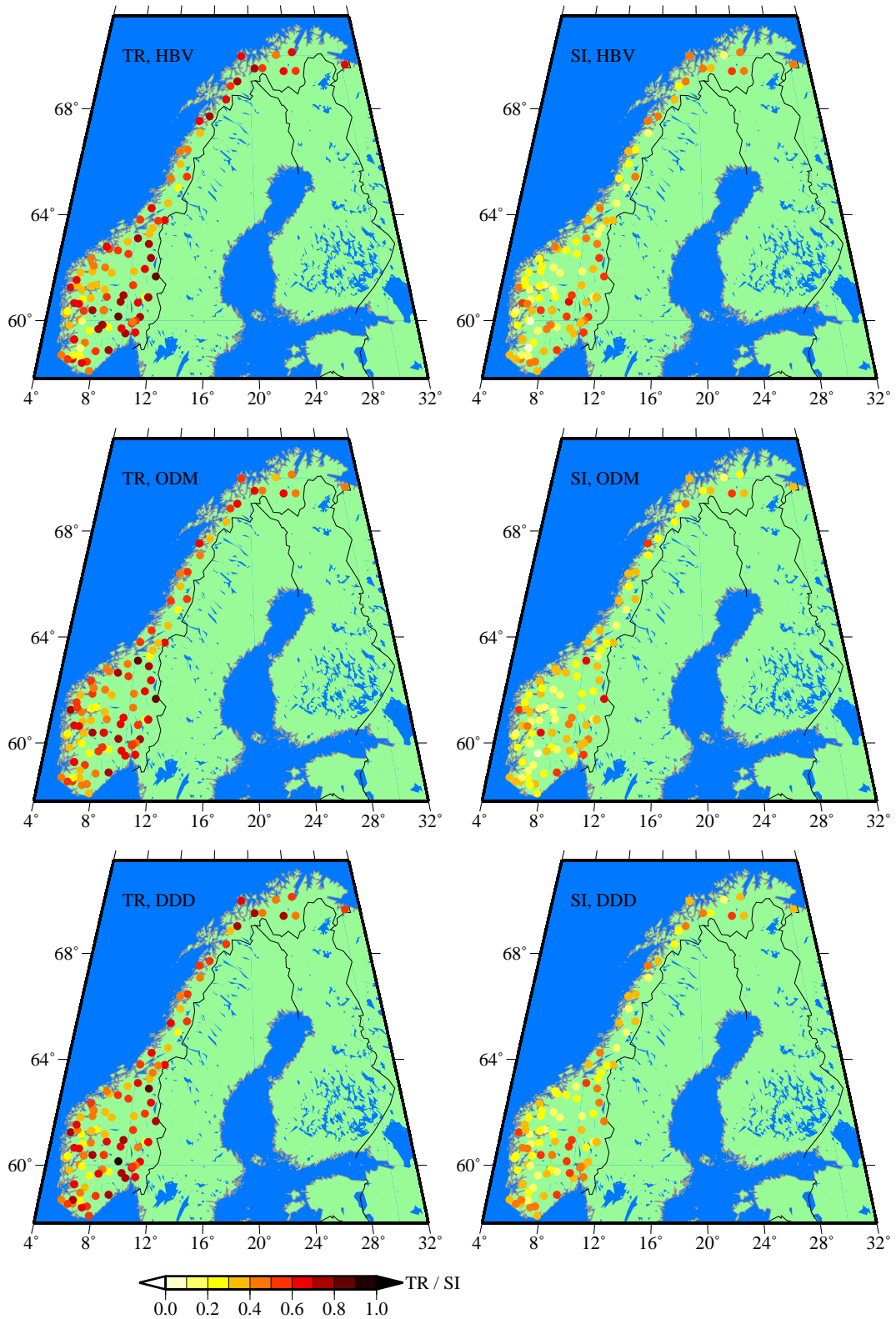
Referanser

- Beldring, S. (2008) Distributed Element Water System, *NVE report 4-2008*, ISBN 978-82-410-0670-8, Norges vassdrags- og energidirektorat, Oslo, Norge (51 s).
- Bergström, S. (1976) Development and application of a conceptual runoff model for Scandinavian catchments, *SMHI Report RHO 7*, Norrköping, Sverige, 134 s.
- Corby, R. J., og W.E. Lawrence (2002) A categorical flood forecast verification system for Southern Region RFC river forecasts, *NOAA Technical Memorandum NWS SR-212*, Fort Worth, USA. Tilgjengelig her:
<http://www.srh.noaa.gov/ssd/techmemo/sr220.pdf>
- Covey, C., K.M. AchutaRao, S.J. Lambert og K.E. Taylor (2000): Intercomparison of Present and Future Climates Simulated by Coupled Ocean-Atmosphere GCMs. PCMDI Report No 66. Program for Climate Model Diagnosis and Intercomparison, Lawrence Livermore National Laboratory, University of California, Livermore, CA.
- Daňhelka, J., og T. Vlasák (2002) Evaluation of real-time flood forecasts in the Czech Republic 2002-2012. Tilgjengelig her:
http://www.chmi.cz/files/portal/docs/poboc/CB/pruvodce/vyhodnoceni_en.html
- Engeland, K. (2014): Evalueringer av flomvarslingsmodeller. Analyser av flomhendelser der den hydrologiske modellen kan forbedres (NVE-rapport under arbeid).
- Langsholt m fl. (2014): Forbedring av flomvarslingens verktøy. Rapport fra FoU-prosjektet 81057, 2011-2014 (NVE-rapport under arbeid).
- Nash, J. E., og J. V. Sutcliffe (1970) River flow forecasting through conceptual models part I — A discussion of principles, *Journal of Hydrology*, 10 (3), 282–290
- Olsson, J., og G. Lindström (2008) Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *Journal of Hydrology*, 350, 14-24.
- Skaugen T. and C. Onof, 2014. A rainfall runoff model parameterized from GIS and runoff data. *Hydrol. Process.* 28, 4529-4542, DOI:10.1002/hyp.9968.
- Stanski, H. R., L.J. Wilson og W.R. Burrows (1989) Survey of common verification methods in meteorology, WMO, *World Weather Watch Report No. 8* (TD No. 358).
- Sælthun, N.R. (1996) The "nordic" HBV model: description and documentation of the model version developed for the project Climate Change and Energy Production, *NVE Publication no. 7-1996*, ISBN 82-410-0273-4, Norges vassdrags- og energidirektorat, Oslo, Norge, 30 s.
- Taylor, K.E. (2001) Summarizing multiple aspects of model performance in a single diagram, *Journal of Geophysical Research-Atmospheres*, 106(D7), 7183-7192, doi:10.1029/2000JD900719

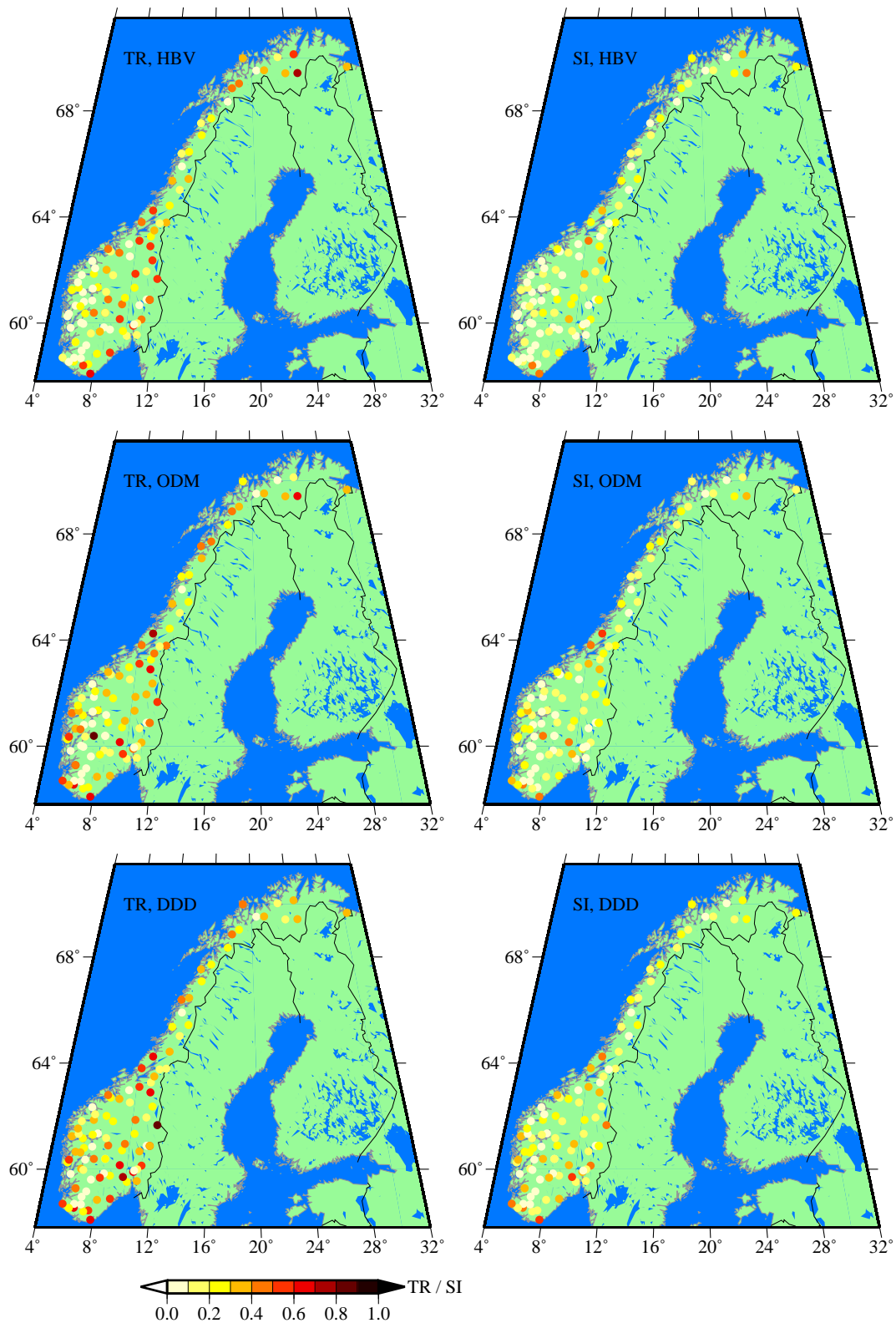
Vedlegg A: Figurer



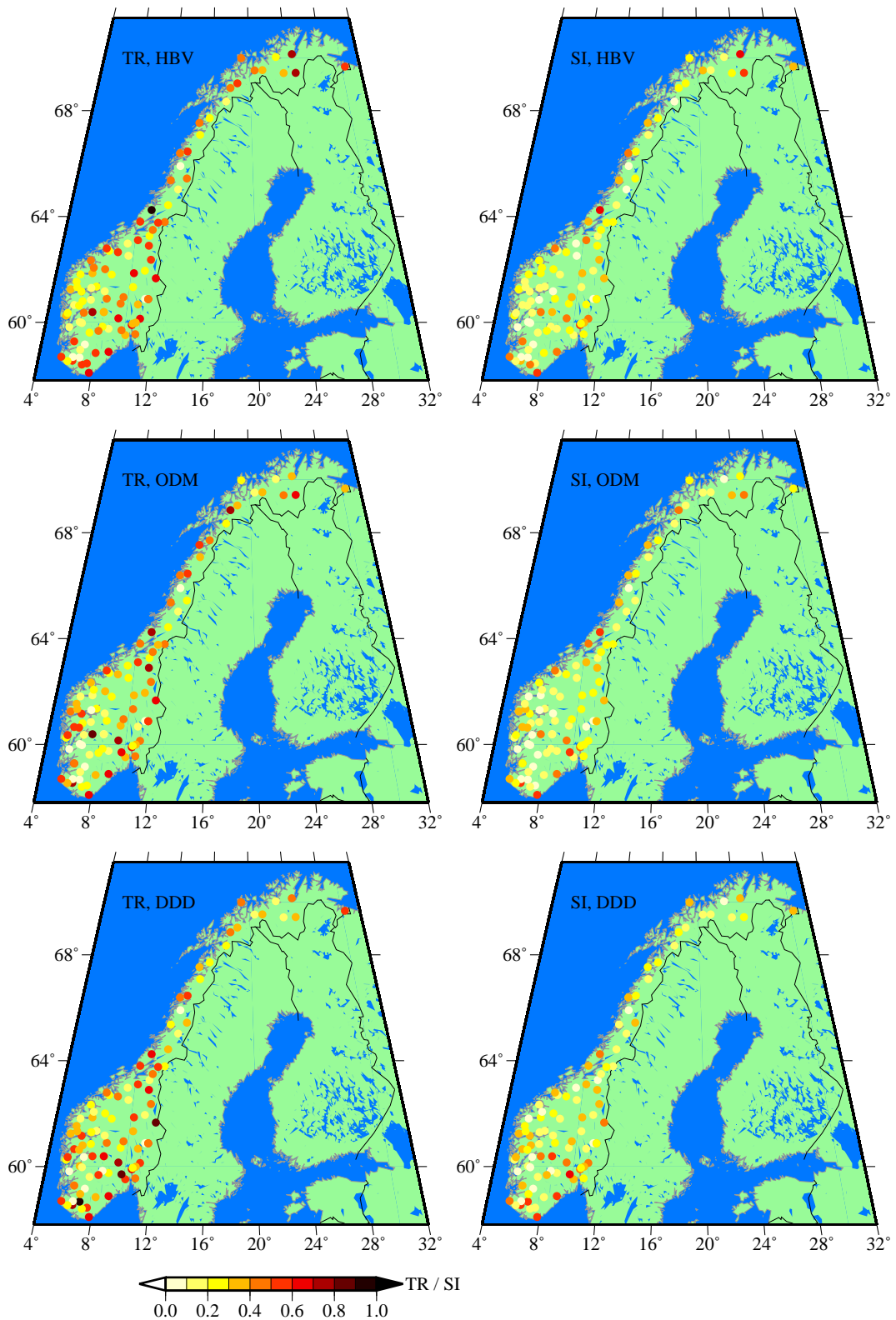
Figur A1: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 1.



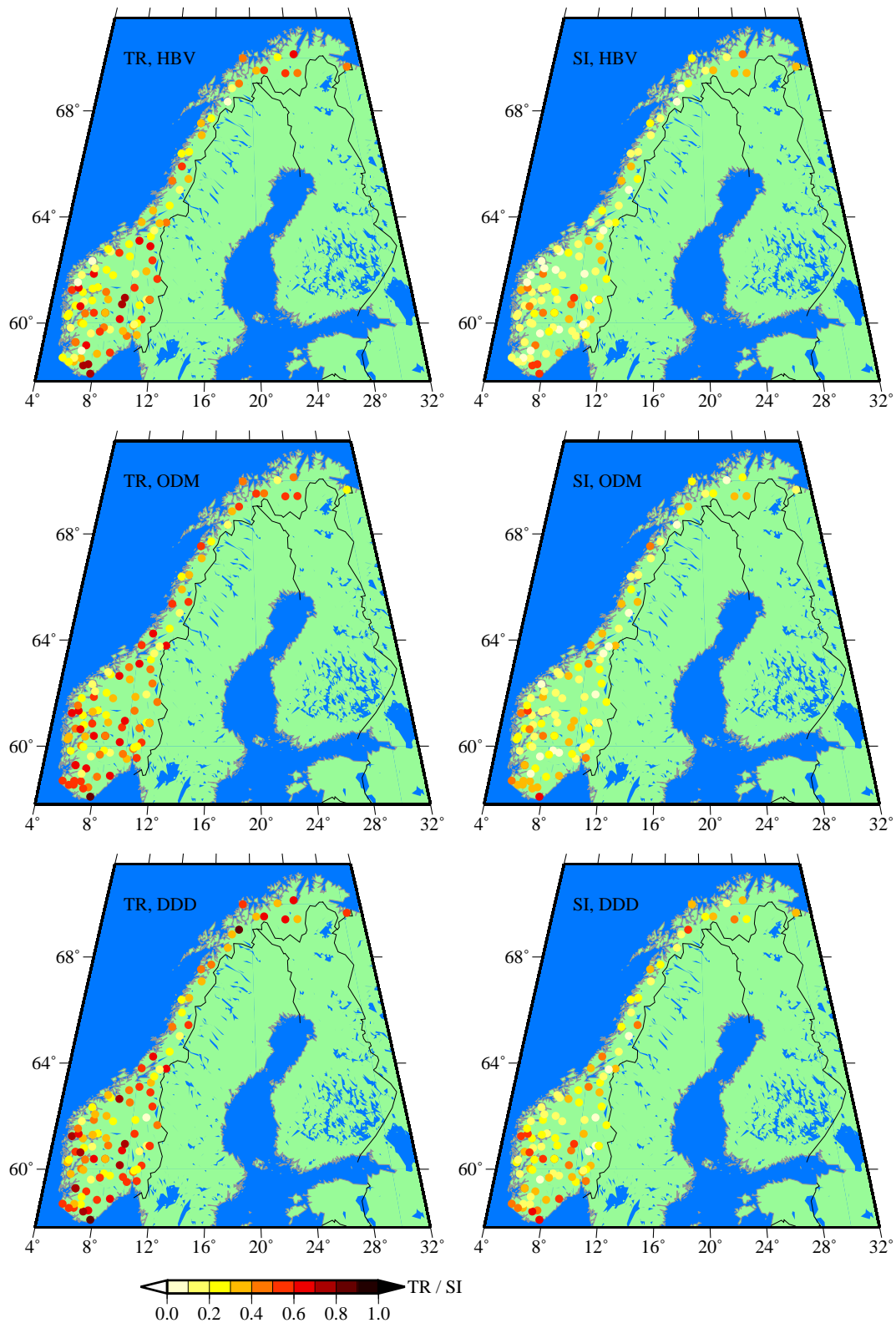
Figur A2: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 2.



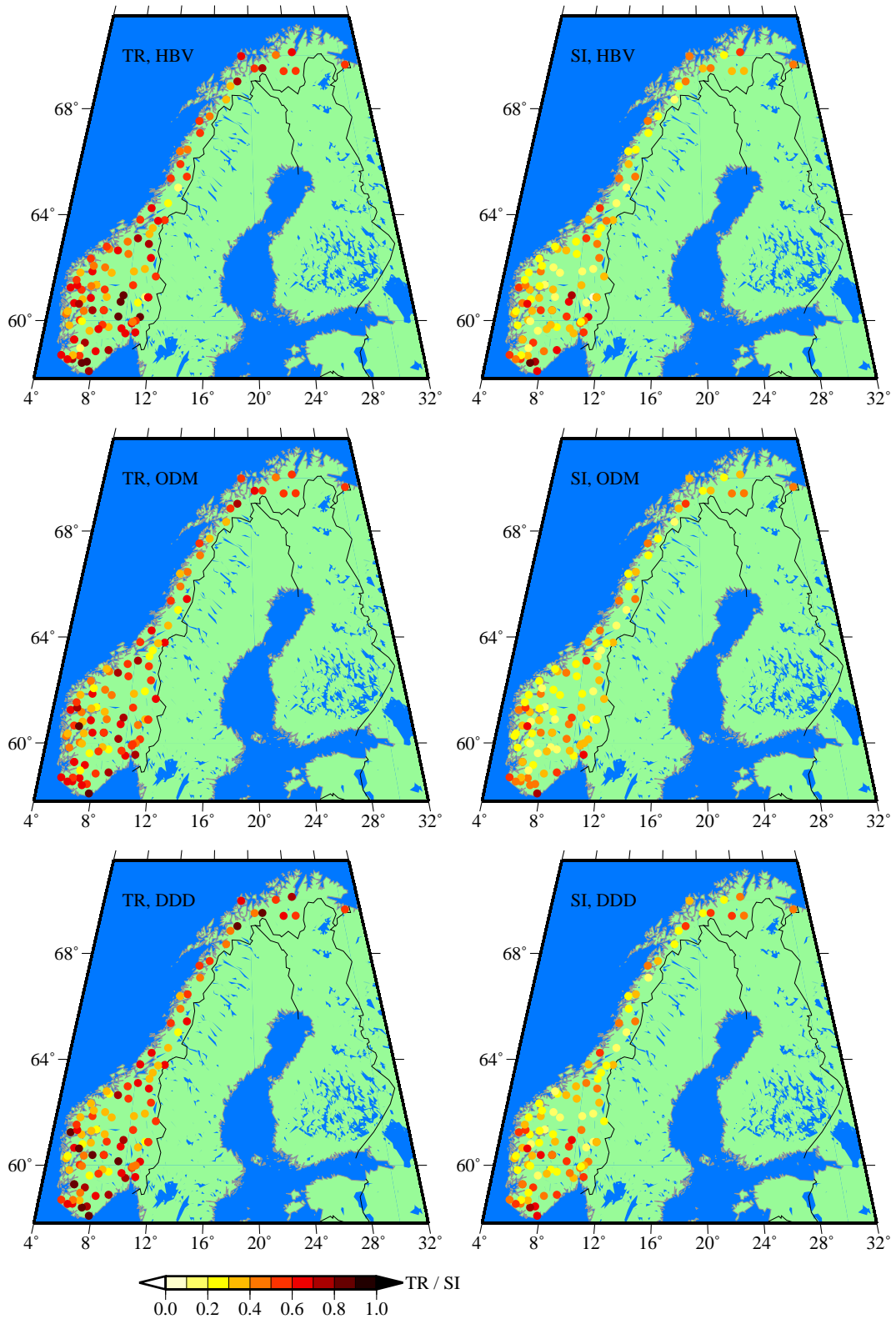
Figur A3: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 3.



Figur A4: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 4.



Figur A5: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 5.



Figur A6: Treffrate (TR) og suksessindeks (SI) for 98 nedbørfelt, HBV, ODM og DDD, alternativ 6.

Vedlegg B: Tabeller

I tabell B1 og B2 presenteres antall treff, bom og falske alarmer for henholdsvis alternativ 1 og 2.

Tabell B1: Alternativ 1

FeltNr	Treff	HBV		Treff	ODM		Treff	DDD	
		Bom	Falsk alarm		Bom	Falsk alarm		Bom	Falsk alarm
2.11	15	13	16	15	13	17	14	14	21
2.145	16	13	24	13	16	24	15	14	18
2.265	7	6	14	7	6	17	6	7	16
2.268	7	20	32	8	19	28	9	18	25
2.279	21	11	20	24	8	24	18	14	21
2.291	7	13	31	8	12	29	6	14	23
2.32	7	14	17	7	14	17	5	16	31
2.323	3	21	39	5	19	31	3	21	26
2.604	22	11	13	21	12	11	17	16	12
2.614	16	10	29	12	14	22	16	10	22
2.634	6	6	11	7	5	7	5	7	11
3.22	8	15	13	15	8	7	11	12	11
6.1	9	24	25	6	27	23	9	24	18
8.2	9	21	27	14	16	20	15	15	16
8.6	7	17	17	8	16	15	8	16	16
12.171	22	9	12	16	15	13	21	10	13
12.178	14	6	19	15	5	20	18	2	13
12.192	7	9	26	5	11	28	5	11	20
12.193	8	11	22	11	8	18	14	5	7
12.215	4	6	14	4	6	15	6	4	6
12.7	19	10	16	16	13	21	20	9	10
15.79	16	5	13	16	5	19	10	11	13
16.132	8	17	16	8	17	17	8	17	10
16.193	4	20	21	5	19	12	5	19	14
16.66	17	10	30	13	14	36	10	17	21
16.75	23	6	20	16	13	20	19	10	23
18.1	8	9	13	11	6	8	12	5	4
20.2	13	27	31	16	24	27	18	22	26
21.47	2	23	30	4	21	29	5	20	29
22.16	19	23	26	12	30	18	19	23	20
22.22	8	10	6	7	11	6	9	9	9
24.9	22	15	19	14	23	22	23	14	20
25.24	4	14	16	7	11	34	6	12	17
25.32	1	7	11	2	6	8	1	7	8

26.2	7	14	18	12	9	12	10	11	10
26.26	2	19	23	4	17	22	5	16	18
27.16	2	8	5	2	8	8	5	5	9
27.24	7	25	28	16	16	18	16	16	11
28.7	9	27	24	16	20	12	15	21	10
35.16	5	8	9	8	5	5	9	4	3
41.1	10	14	35	13	11	33	14	10	32
42.2	5	33	26	10	28	27	10	28	28
48.1	8	25	34	12	21	25	15	18	18
48.5	5	26	40	4	27	37	7	24	35
50.1	17	30	37	17	30	39	18	29	30
50.13	15	8	12	15	8	7	16	7	11
55.4	7	32	28	11	28	19	13	26	15
55.5	4	14	9	4	14	10	4	14	8
62.1	5	29	26	5	29	33	6	28	22
62.18	1	14	11	6	9	11	6	9	8
62.5	18	20	23	17	21	25	19	19	18
72.5	4	30	36	10	24	57	8	26	23
73.27	5	15	17	5	15	15	5	15	10
75.23	8	16	26	6	18	34	4	20	19
77.3	4	24	24	6	22	23	9	19	15
79.3	3	13	16	4	12	8	2	14	11
82.4	17	16	18	25	8	17	24	9	11
83.2	14	24	27	19	19	22	17	21	16
84.11	4	21	39	12	13	30	10	15	28
86.12	2	16	14	7	11	10	6	12	7
88.4	11	29	30	20	20	28	15	25	28
97.1	1	25	36	10	16	28	7	19	26
98.4	6	26	30	11	21	26	11	21	21
104.23	5	18	21	7	16	13	5	18	10
105.1	14	12	21	10	16	19	12	14	27
109.42	13	15	37	17	11	30	16	12	32
109.9	7	19	40	12	14	36	11	15	30
112.8	7	17	25	8	16	25	7	17	26
122.11	11	10	19	13	8	23	14	7	14
122.9	11	13	20	18	6	10	16	8	9
123.31	8	28	37	9	27	18	10	26	21
124.2	7	24	34	9	22	28	8	23	25
127.11	12	13	16	12	13	18	14	11	20
127.13	4	16	17	4	16	10	5	15	17
133.7	9	15	20	10	14	15	11	13	16
138.1	13	15	14	14	14	11	16	12	10
139.15	7	32	27	10	29	34	7	32	20
139.35	10	19	37	8	21	39	8	21	40

148.2	7	12	12	9	10	11	11	8	16
151.15	16	20	18	19	17	27	20	16	20
152.4	10	20	23	10	20	28	7	23	28
156.1	7	20	30	10	17	23	13	14	17
157.3	5	32	27	13	24	15	16	21	18
162.3	7	16	28	9	14	22	9	14	34
165.6	13	17	20	16	14	9	15	15	11
168.2	4	7	9	4	7	4	5	6	5
174.3	4	9	16	5	8	12	5	8	11
191.2	9	17	23	10	16	23	10	16	23
196.35	18	12	21	15	15	15	20	10	22
200.4	13	15	16	15	13	21	17	11	21
206.3	9	9	19	9	9	39	8	10	16
208.3	8	9	10	6	11	9	7	10	11
212.1	18	8	13	16	10	11	17	9	10
212.49	5	11	24	5	11	21	7	9	27
223.2	16	14	16	11	19	20	14	16	25
234.13	12	12	9	8	16	8	11	13	14
247.3	16	17	14	12	21	14	16	17	16
311.46	20	4	11	20	4	14	18	6	20

Tabell B2: Alternativ 2

FeltNr	Treff	HBV			ODM			DDD	
		Bom	Falsk alarm	Treff	Bom	Falsk alarm	Treff	Bom	Falsk alarm
2.11	20	8	12	17	11	15	18	10	17
2.145	17	12	22	13	16	22	15	14	14
2.265	8	5	10	8	5	15	7	6	12
2.268	9	18	29	9	18	26	9	18	23
2.279	23	9	12	24	8	20	20	12	20
2.291	8	12	30	8	12	28	7	13	22
2.32	7	14	17	9	12	17	8	13	29
2.323	7	17	31	7	17	27	5	19	24
2.604	24	9	11	21	12	10	19	14	11
2.614	19	7	28	16	10	19	18	8	21
2.634	8	4	9	7	5	6	5	7	11
3.22	16	7	7	16	7	4	15	8	9
6.1	18	15	12	16	17	14	16	17	11
8.2	17	13	15	17	13	15	19	11	12
8.6	11	13	12	12	12	10	14	10	10
12.171	24	7	9	19	12	11	21	10	10
12.178	17	3	13	15	5	19	18	2	11
12.192	12	4	18	10	6	20	12	4	16
12.193	13	6	16	13	6	15	15	4	7
12.215	5	5	13	4	6	13	6	4	6
12.7	21	8	12	19	10	18	21	8	9
15.79	16	5	8	17	4	14	12	9	11
16.132	12	13	12	15	10	15	14	11	8
16.193	7	17	17	6	18	11	7	17	12
16.66	21	6	25	15	12	31	12	15	21
16.75	23	6	18	17	12	19	19	10	21
18.1	12	5	6	12	5	5	12	5	4
20.2	22	18	19	16	24	27	20	20	24
21.47	6	19	26	5	20	29	6	19	29
22.16	24	18	17	18	24	15	24	18	16
22.22	8	10	5	7	11	6	10	8	9
24.9	23	14	15	15	22	19	25	12	17
25.24	8	10	15	9	9	32	6	12	15
25.32	2	6	9	3	5	7	2	6	7
26.2	13	8	10	12	9	12	10	11	10
26.26	3	18	19	7	14	19	9	12	14
27.16	2	8	4	2	8	8	7	3	7
27.24	16	16	16	18	14	16	18	14	9
28.7	17	19	12	17	19	12	16	20	10
35.16	8	5	4	8	5	5	9	4	3
41.1	16	8	29	15	9	30	16	8	29

42.2	12	26	18	15	23	21	15	23	23
48.1	15	18	23	13	20	22	16	17	15
48.5	6	25	37	7	24	35	9	22	34
50.1	24	23	30	18	29	36	20	27	26
50.13	17	6	10	17	6	4	17	6	11
55.4	14	25	19	13	26	17	15	24	13
55.5	5	13	8	5	13	9	5	13	7
62.1	9	25	20	11	23	27	7	27	22
62.18	10	5	2	9	6	9	9	6	6
62.5	23	15	17	23	15	18	23	15	14
72.5	13	21	28	13	21	51	15	19	21
73.27	6	14	16	6	14	13	6	14	9
75.23	8	16	23	6	18	31	6	18	17
77.3	10	18	17	7	21	21	11	17	14
79.3	5	11	12	7	9	6	4	12	10
82.4	21	12	14	25	8	17	25	8	11
83.2	15	23	26	20	18	21	17	21	15
84.11	15	10	26	14	11	25	15	10	22
86.12	7	11	9	7	11	10	6	12	7
88.4	15	25	27	20	20	26	16	24	25
97.1	12	14	23	15	11	20	14	12	17
98.4	13	19	19	15	17	22	14	18	17
104.23	12	11	11	11	12	7	8	15	7
105.1	18	8	18	12	14	14	12	14	25
109.42	16	12	32	20	8	26	16	12	30
109.9	13	13	33	14	12	31	15	11	29
112.8	9	15	21	10	14	21	8	16	25
122.11	16	5	14	15	6	21	17	4	14
122.9	18	6	13	20	4	9	16	8	8
123.31	11	25	32	10	26	17	13	23	20
124.2	11	20	29	12	19	26	14	17	21
127.11	15	10	13	16	9	16	15	10	20
127.13	10	10	11	6	14	8	8	12	12
133.7	13	11	14	13	11	13	12	12	16
138.1	17	11	8	15	13	10	17	11	9
139.15	10	29	23	10	29	32	9	30	19
139.35	11	18	34	11	18	34	11	18	37
148.2	9	10	12	11	8	11	13	6	15
151.15	20	16	14	21	15	26	20	16	20
152.4	11	19	21	11	19	24	10	20	25
156.1	13	14	22	15	12	14	14	13	15
157.3	17	20	18	18	19	10	17	20	15
162.3	9	14	26	11	12	20	10	13	33
165.6	18	12	12	19	11	7	16	14	9

168.2	8	3	7	4	7	4	6	5	5
174.3	8	5	13	5	8	12	7	6	11
191.2	13	13	22	14	12	21	10	16	22
196.35	22	8	19	18	12	14	21	9	20
200.4	19	9	13	16	12	19	18	10	19
206.3	14	4	13	12	6	33	13	5	12
208.3	10	7	10	8	9	9	7	10	9
212.1	18	8	8	18	8	8	20	6	8
212.49	7	9	21	6	10	19	7	9	27
223.2	18	12	14	13	17	16	17	13	20
234.13	13	11	8	10	14	7	11	13	11
247.3	20	13	10	15	18	12	18	15	15
311.46	20	4	10	21	3	11	18	6	18

Vedlegg C: Eksempler fra våre naboland

I Sverige (Göran Lindström, e-post nov 2011) brukes toppfeilen som et mål på eventuell systematisk over- eller underestimering:

$$Tf = \frac{Q_{ms} - Q_{mo}}{Q_{mo}} \quad (10)$$

der Q_{ms} er middelflommen for simulert vannføring, og Q_{mo} er middelflommen av observert vannføring. Iblant brukes også middelveien av absoluttfeilen som et mål:

$$MAE = \frac{1}{n} \frac{\sum (|Q_s - Q_o|)}{\overline{Q_o}} \quad (11)$$

der n = antall dager og Q_o -bar er middelveien av observasjonene. For vårfloppen beregnes ofte som volumfeilen i snøsmeltingsperioden. SMHI utfører også noen sannsynlighetsprognoser, se Olsson og Lindström (2008).

I Finland (Juha Markkula, epost 12.10.2011) har de et automatisk system som med jevne mellomrom beregner valideringsscore for prognosene, og lager valideringsgrafer som blir gjort tilgjengelig på websidene deres.

De viktigste prognosevalideringskriteriene for vannføring i elver i Finland er:

- Feil i flomtoppen (peak discharge error): Median av absoluttfeilen av maksimalvannføringen.
- Feil i tidspunkt for flomtoppen: Median av absoluttfeilen i tidspunktet for flomtoppen.

De viktigste prognosevalideringskriteriene for tilsig til vann i Finland er:

- Gjennomsnittlig absoluttfeil (mean absolute error)
- Relativ feil
- Nash-Sutcliffe

Denne rapportserien utgis av Norges vassdrags- og energidirektorat (NVE)

F.o.m 51. utgivelse i 2014.

- Nr. 51 Forslag til nytt vektsystem i modellen for kostnadsnormer
- Nr. 52 Jord- og sørpeskred i Sør-Norge mai 2013. Monica Sund
- Nr. 53 Årsrapport for utførte sikrings- og miljøtiltak for 2013
- Nr. 54 Naturfareprosjekt DP. 1 Naturskadestrategi
- Nr. 55 Naturfareprosjektet DP.6 Kvikkleire. Effekt av progressiv...
- Nr. 56 Naturfareprosjektet DP.6 Kvikkleire. Effekt av progressiv...
- Nr. 57 Naturfareprosjektet DP.6 Kvikkleire. Sikkerhet ifm utbygging..
- Nr. 58 Naturfareprosjektet DP.6 Kvikkleire. Sikkerhet ifm utbygging...
- Nr. 59 Naturfareprosjektet DP.6 Kvikkleire. Likestilling mellom bruk...
- Nr. 60 Skredfarekartlegging i Høyanger kommune
- Nr. 61 Flaumsonkart Delprosjekt Førde. Orvedal og Peereboom
- Nr. 62 Naturfareprosjektet Dp. 5 Flom og vann på avveie.
- Nr. 63 Naturfareprosjektet DP. 3.2 Datasamordning Ministudie av...
- Nr. 64 Naturfareprosjektet. Delprosjekt 2. Beredskap og krisehåndterin
- Nr. 65 Grønne tak og styrtregn. Effekten av ekstensive tak med...
- Nr. 66 Norges vannbalanse i TWh basert på HBV-modeller.
- Nr. 67 Effekt av lagringstid på prøve kvalitet. Haakensen / NIFS.
- Nr. 68 Effect of storage time on sample quality. Haakensen / NIFS.
- Nr. 69 Flomsonkart. Delprosjekt Fagernes. Ahmed Reza Naserzadeh..
- Nr. 70 Status høsten 2014 - resultater og veien videre. Haakensen/NIF
- Nr. 71 Aktive vannføringsstasjoner i Norge, Lars Evan Pettersson.
- Nr. 72 Smarte målere (AMS) og feedback. VasaaETT og Heidi Kvalvåg.
- Nr. 73 Filefjell og Anestølen. Evaluering av måledata for snø,..
- Nr. 74 Avbrottsstatistikk 2013. Astrid Ånestad.
- Nr. 75 Energibruk i undervisningsbygg. Langseth og Multiconsult m.fl.
- Nr. 76 Naturfareprosjektet: Delprosjekt 2. Beredskap og krisehåndterin
- Nr. 77 Naturfareprosjektet: Delprosjekt 6. Kvikkleire. Haakensen.
- Nr. 78 Status og prognoser for kraftsystemet 2014.Synnøve Lill Pauler
- Nr. 79 Snøskredvarslingen. Evaluering av vinteren 2014. NIFS. Barfod.
- Nr. 80 Norwegian Avalanche Warning Service. Program Review. NIFS. Gr
- Nr. 81 Oppsummeringsrapporter ifm høring 2-2014. Mi Lagergren.
- Nr. 82 Oppsummeringsrapporter ifm høring 2-2014. Mi Lagergren.
- Nr. 83 Inventory of glacier-related hazardous events in Norway.Jacksor
- Nr. 84 Evaluering av flomvarslingas modellverktøy. Ingjerd Haddeland.
- Nr. 85 Kartlegging av oppvarmingsutstyr i husholdningene. Magnussen.
- Nr. 86 Elsertifikat Årsrapport 2013
- Nr. 87 Naturfareprosjektet: Droneteknologi. Haakensen/NIFS.
- Nr. 88 Naturfareprosjektet: Delprosjekt 6. Kvikkleire. Haakensen/NIFS.
- Nr. 89 Klimaendring og damsikkerhet: En pilotstudie.. D.Lawrence.
- Nr. 90 Rapport Troms. Graziella Devoli/ NIFS.

Nr. 91

Rapport Kvikkleire. Graziella Devoli /NIFS

Nr. 92

Ground improvement for marginally stable slopes. NIFS.



Norges
vassdrags- og
energidirektorat

Norges vassdrags- og energidirektorat

Middelthunsgate 29
Postboks 5091 Majorstuen
0301 Oslo

Telefon: 09575
Internett: www.nve.no

