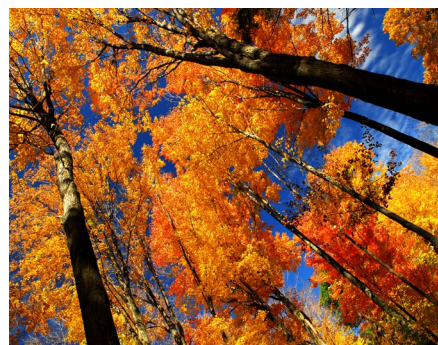
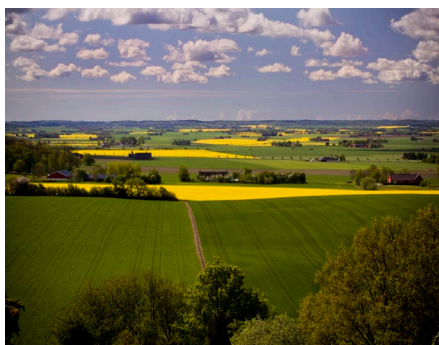


Homogenisering av månadsmedeltemperatur 1860–2021

Magnus Joelsson



Pärmbild.

Bilderna föreställer en vy över gröna och gula fält under en blå himmel med stackmoln, ett antal hus i snöstorm och brandgula löv i trädkronor sedd underifrån, med blå himmel och slöjmoln i bakgrunden.

Klimatologi nr 59

Homogenisering av månadsmedeltemperatur 1860–2021

Magnus Joelsson

Granskad och Godkänd av: Cristoffer Wittskog

Funktion: Nätansvarig för det meteorologiska observationsnätet

Denna sida är avsiktligt blank

Sammanfattning

I denna rapport presenteras sammanställningen och homogeniseringen av 456 tidsserier för månadsmedeltemperatur. Tidsserierna innehåller data från totalt 836 väderstationer i det svenska stationsnätverket mellan åren 1860 och 2021. Tidsserierna är sammankopplade för att kunna använda så mycket data som möjligt och använda så litet uppskattad (interpolerad) data som möjligt. Sammankopplingarna utförs med en nyutvecklad automatisk metod vilken beskrivs i rapporten. Homogeniseringen och interpoleringen utförs med en vidareutveckling av homogeniseringverktöget HOMER (Bart).

Sammankopplingen ger en täckningsgrad av data för alla tidsserier mellan åren 1860 och 2021 på 44 %. Motsvarande täckningsgrad för icke-sammankopplade tidsserier längre än 15 år är 30 %. 257 av de 456 sammankopplade tidsserier består av över 60 års data (motsvarande två normalperioder). 173 av de icke-sammankopplade tidsserier består av över 60 års data.

För 22 av tidsserierna detekteras inga homogenitetsbrott och de kan därför betraktas som homogena. I övrigt detekteras 1 788 homogenitetsbrott vilket motsvarar en mediantid på ett homogenitetsbrott per 17 år av observationer. Vanligast förekommande är tre homogenitetsbrott i en och samma tidsserie. 37 % av homogenitetsbrotten har stöd i metadata eller sammanfaller med kopplingspunkter. 37 % av datapunkterna har en negativ korrektion (observationerna har en högre temperatur än det homogeniserade värdet), 24 % positiv korrektion. Resterande 40 % är inte korrigerade i det homogeniserade datasetet.

Summary

In this report, the merging and homogenisation of 456 time series of monthly mean temperature, are presented. The time series consists of data from in total 836 weather stations in the Swedish station network between the years 1860 and 2021. The time series are merged in order to use more of the available data and to restrict the interpolation of data. The merging is performed with a newly developed automatic tool, which is described here. The homogenisation and interpolation of data is performed with a development of the homogenisation tool HOMER (Bart).

The merging of the time series between the years 1860 and 2021 gives a data coverage over of 44 %. Corresponding data coverage for the raw time series with more data than 15 years is 30 %. 257 of the 456 merged time series consist of over 60 years of data (which corresponds to two normal periods). 173 of the raw times series consist of over 60 years of data.

For 22 of the time series no homogeneity breaks are detected. These time series can be considered to be homogeneous. For the other series 1 788 homogeneity breaks are detected which corresponds to 17 years of observations per break. Most common is three homogeneity break in a series. 37 % of the homogeneity breaks is supported in meta data or coincide with merging points. 37 % of the data is negatively corrected (observations has higher temperature than the homogenised value), 24 % positively corrected. The remaining 40 % is not corrected.

Denna sida är avsiktligt blank

Innehållsförteckning

1	BAKGRUND	9
2	SYFTE	10
3	METODIK	10
3.1	Stationskopplingar	10
3.1.1	Principer	10
3.1.2	Automatisk metod	11
3.1.2.1	Gruppering av stationer	11
3.1.2.2	Sammankoppling av stationsgrupper	12
3.2	Homogenisering	13
3.2.1	HOMER	13
3.2.1.1	Detektionsfunktioner	13
3.2.1.2	ANOVA-korrektionsmodell	14
3.2.1.3	Praktiskt användning av HOMER	14
3.2.1.4	Bart	15
3.2.1.5	Metadata	15
3.2.2	Korrigerig av gradvisa inhomogeniteter	16
4	RESULTAT	16
4.1	Stationskopplingar	16
4.2	Homogenisering	16
4.2.1	Homogenitetsbrott	16
4.2.2	Korrektion	17
4.2.3	Korrektion av gradvisa inhomogeniteter	17
4.2.4	Exempel: Jokkmokk	17
5	DISKUSSION	18
6	SLUTSATSER	19
7	APPENDIX	19
7.1	ANOVA-korrektionsmetoden, minimalt exempel	19
8	REFERENSER	20
9	FIGURER OCH TABELLER	22

Denna sida är avsiktligt blank

1 Bakgrund

Klimatnormal definieras som väder sett över 30 år (WMO, 2017). Exempelvis definieras den klimatologiska medeltemperaturen för en viss punkt som medeltemperaturen för den punkten under 30 år. För att fastställa förändringar i den klimatologiska medeltemperaturen krävs alltså mer än 60 års data. Alla undersökningar av klimatets förändringar behövs således långa tidsserier av meteorologiska observationer. Väderstationer som är aktiva på dessa tidsskalor (10–100 år) har ofta varit med om olika förändringar som potentiellt har påverkat observationerna: mätinstrument eller annan utrustning har bytts ut eller degenererat, observatörer bytts ut eller stationen flyttats. Även mätmetoder och rutiner har förfinas och förändras genom åren, exempelvis tidpunkten för dagliga observationer). Vidare kan också omgivningarna genomgå förändringar exempelvis i form av växtlighet eller byggnader. En plats som vid mätseriens start befunnits i utkanten av en stad kan efterhand finnas mitt i stadskärnan. Venema et al. (2012) uppskattar att en tidsserie ifrån en typisk västerländsk mätstation har ett homogenitetsbrott vart 15:e till vart 20:e år.

I vissa fall har dylika förändringar noterats i inspektionsprotokoll för väderstationerna, men det är rimligt att misstänka att många förändringar inte finns dokumenterade. En stor del av de inspektionsprotokoll som finns upprättade är dock ännu inte digitaliserade och är därför svårtillgängliga.

1950 gav Conrad och Pollack följande definition av “homogenitet” för klimatologiska tidsserier:

A numerical series representing variations of a climatological element is called “homogeneous” if the variations are caused only by variations of weather and climate.

Ungefär: “en numerisk serie som representerar variationer av en klimatologisk variabel kallas ‘homogen’ om variationerna endast orsakas av väder och klimat”.

För att göra observationer i en icke-homogen tidsserie jämförbara med varandra behöver tidsserien homogeniseras. Homogenisering innebär att en tidsserie justeras så att den motsvarar en mätserie uppmätt under samma förutsättningar under hela perioden. Ofta efterliknas förutsättningarna vid den senaste observationen, då detta förenklar fortsatt förlängning av tidsserien med nya observationer. Om exempelvis en förflyttning av en mätstation resulterar i konstant högre uppmätta temperaturer, kan mätvärdena i tidsserien fram till flytten justeras upp i motsvarande grad.

För att lösa problemet med bristande dokumentation över observationsserier (så kallad “metadata”) och för att ta fram kvantitativa korrektionstermer har ett antal statistiska homogeniseringsverktyg utvecklats (Ribeiro et al. 2016). Ett vanligt tillvägagångssätt är att jämföra en tidsserie (kandidatserie) med en referensserie för att hitta homogenitetsbrott. Referensserien kan antingen vara en annan tidsserie som antas beskriva samma klimatologi (exempelvis från en närbelägen väderstation) och som parvis jämförs med kandidatserien eller en sammanfogad serie (exempelvis en genomsnittlig serie från ett antal närbelägna väderstationer). HOMER (HOMogenizaton softwarE in R) är ett verktyg som kombinerar en parvis detektionsfunktion med en sammanfogad detektionsfunktion (Mestre et al. 2013).

Dagliga temperaturobservationer har utförts i Sverige sedan 1722 (Bergström och Moberg, 2002) då mätserien i Uppsala startade på initiativ av bland annat Anders Celsius. 1754 startades mätserien i Stockholm (Moberg et al. 2002), vilket är den längsta pågående mätserien utan betydande avbrott i Sverige. 1858–1860 installerades 24 mätstationer av Kungliga Vetenskapsakademi, vilka snart kom att förvaltas av SMHI:s föregångare Statens Meteorologiska Centralanstalt. Stationsnätets omfattning ökade därefter de kommande 100 åren; 1962 var över 400 stationer aktiva. Sedan dess har antalet stationer successivt åter minskat till 251 stationer 2021. Under dessa år har

stationer upphört eller behövt flyttas. 1995 inleddes en omfattande automatisering av stationsnätet (Alexandersson, 2002).

Observationerna registreras för närvarande i Meteorologisk Observationsdatabas för Realtid och Arkiv (MORA). Ibland har, vid en flytt eller automation, en ursprunglig station och en ersättande station registreras som två separata stationer i MORA. Särskilt vid automation har det också ofta funnits en period med parallella mätningar. För att studera förändringar i klimatet krävs långa tidsserier; det är då fördelaktigt att sammankoppla dessa stationers tidsserier. I SMHI:s månatliga redovisning "Väder och vatten" har enligt tradition 100 delvis sammankopplade tidsserier använts för att beskriva Sveriges klimat. Det homogeniserade dataset som fram till 2022 utgjort underlaget för SMHI:s klimatindikator för månadsmedeltemperatur omfattade 35 kopplade tidsserier.

Även tidsserier från historiskt icke-besläktade stationer kan sammankopplas om de två tidsserierna beskriver två klimatologiskt likvärdiga platser. Kopplingar av data från olika stationer har fördelarna att de långa tidsserierna blir fler relativt data från enskilda stationer och att data från tidsserier som är för korta för att homogeniseras separat (i HOMER tidsserier med mindre än 15 års data) kan användas. Å andra sidan förkastas en del data från överlappande tidsserier som sammankopplas och närbelägna parallella tidsserier används i homogeniseringsprocessen. De sammankopplade serierna riskerar också att ha homogenitetsbrott vid kopplingarna.

Inför den i denna rapport beskrivna homogeniseringen av månadsmedeltemperatur i Sverige har en sammankoppling av tidsserier utförts, där alla tillgängliga månadsmedeltemperaturuppgifter ingått som ingångsdata. Då detta är ett mycket omfattande arbete har en automatisk metod för sammankopplingar av tidsserier utvecklats.

Det sammankopplade datasetet har sedan homogeniserats med en helautomatiserad versionen av HOMER (Joelsson, 2022a). Homogeniseringen av månadsmedeltemperatur i Sverige 1860–2021 finns även beskriven i Joelsson et al. (2022b).

2 Syfte

Projektet syftar till att sammankoppla och homogenisera långa tidsserier av månadsmedeltemperaturer från 1860 för att beskriva temperaturutvecklingen i Sverige över den tid då reguljära observationer representativa för hela Sveriges klimat utförts.

3 Metodik

I följande avsnitt beskrivs metodiken för sammankopplingen (avsnitt 3.1) och homogeniseringen (avsnitt 3.2) av tidsserier.

3.1 Stationskopplingar

3.1.1 Principer

Sammankopplingen av tidsserier måste beakta följande utgångspunkter:

- Tidsserier som kopplas samman ska representera samma klimat.
- Den sammankopplade tidsserien ska vara så heltäckande som möjligt.
- Den sammankopplade tidsserien ska vara så homogen som möjligt.
- Ny data ska kontinuerligt kunna läggas till.

Från dessa utgångspunkter kan ett antal principer uppställas för sammankopplandet av tidsserier från olika stationer:

- i. Tidsserier med god täckning föredras framför dataserier med sämre täckning.
- ii. Stationer måste vara tillräckligt närbelägna.

- iii. Tidsserierna ska inte vara svagt korrelerade, i den mån korrelation kan testas.
- iv. Tidsserier från ö-, kust- och inlandsstation ska inte kopplas med varandra.
- v. Tidsserier som innehåller de nyaste observationerna föredras.
- vi. Antalet kopplingar bör hållas lågt för bättre homogenitet. Exempelvis kan i de fall där täta kopplingar fram och tillbaka mellan tidsserier skulle ge maximal täckning i den sammanlagda tidsserien, färre kopplingar föredras.
- vii. För sammankoppling av tidsserier i ett dataset som ska homogeniseras kan ytterligare ett krav ställas på stationskopplingarna: tidsserier med ett långt överlapp kan användas i homogeniseringsprocessen och bör därför inte sammankopplas.

Med utgångspunkt i dessa principer har en automatisk kopplingsprocedur utvecklats.

Alla månadsmedeltemperaturer baserade på 2-meterstemperaturobservationer från SMHI:s stationsnät, digitaliserade och registrerade i MORA, används som ingångsdata till kopplingsprocessen.

3.1.2 Automatisk metod

Den automatiska stationssammankopplingen sker i två steg:

1. Gruppering av stationer som möjligen kan sammankopplas
2. Sammankoppling av grupperna till en sammankopplad tidsserie

I avsnitt 3.1.2.1 och 3.1.2.2 förklaras de två stegen var för sig.

När ett dataset av sammankopplade stationer sammanställts prövas datasetet i HOMER. Om någon av tidsserierna inte kan användas i HOMER på grund av att det inte finns tillräckligt många referensserier, tas kopplingsprocessen om från början. De av HOMER utsorterade huvudstationerna finns då med på en lista med olämpliga huvudstationer, dessa kan då inte bli huvudstationer i den nya sammankopplingsrundan.

3.1.2.1 Gruppering av stationer

Processen för grupperingen av stationer är åskådliggjord i figur 1.

Samtliga ingående tidsserier listas enligt antalet datapunkter:

1. En potentiell huvudstation (station **A**) väljs genom att gå igenom de stationer som ännu inte blivit grupperade och som inte finns med på en lista över olämpliga huvudstationer. Serierna tas i ordning utifrån antalet datapunkter; serier med flest datapunkter undersöks först, serien med minst antal datapunkter undersöks sist.
2. Andra ännu icke-grupperade stationer sorteras efter distansen mellan dem och station **A**. Dessa stationer utgör potentiella stödstationer.
3. De potentiella stödstationerna som ännu inte testats med station **A** undersöks en och en tills en (endast en) lämplig stödstation hittas. Alla par av stationer som redan testats finns registrerade i en kontrolltabell. Stödstationen (station **B**) godkänns om:
 - a. Station **A** och station **B** båda är ö-stationer eller båda är inlandsstationer.
 - b. Station **A** och station **B** båda är längre från kusten eller båda är närmare kusten än bredden på kustremsan eller stationerna är närmare varandra än den mest kustnära av dem är nära kusten.
 - c. Distansen mellan station **A** och station **B** är mindre eller lika med tröskelvärdet (25 km vilket motsvarar distansen mellan Holmön och Holmögdadd).
 - d. Det inte finns någon möjlig huvudstation som är närmare station **B**. Om lämpligheten mellan station **B** och en annan tänkbar huvudstation (**C**) undersökts i ett tidigare skede i proceduren och funnits icke-lämpliga är station **C** inte längre möjlig som huvudstation för station **B** och hindrar alltså inte längre att station **B** grupperas med station **A**, även om station **C** ligger närmare station **B** än station **A**.

- e. Överlappet mellan stationerna inte är längre än tröskelvärdet (15 års data vilket är den kortaste tillåta tidsserien i HOMER). Om station **B** finns på listan över olämpliga huvudstationer kan överlappet vara längre än tröskelvärdet 15 år.
 - f. Det finns komplimenterande data i station **B** (data vid tidpunkter som inte finns i station **A**)
 - g. Korrelationen mellan de två tidsserierna inte är lägre än tröskelvärdet (0.9)
 - h. (Om överlappet är för litet för att räkna ut korrelationen) Korrelationen mellan de två tidsserierna och en tredje tidsserie **D** (den närmsta stationen med en tidsserie som har tillräckligt många gemensamma datapunkter med både huvudserien och stödserien) inte skiljer sig mer åt än tröskelvärdet. (Om inget av korrelationstesten går att genomföra kan tidsserierna fortfarande kopplas ihop.)
4. Klara station **B** testen i punkt 3. grupperas den med station **A** och ses inte längre som grupperingsbar. Data för station **A** och **B** slås preliminärt ihop. Den preliminärt sammanslagna tidsserien används bara i grupperingssyfte, inte som slutgiltig tidsserie. Station **A**:s rad i kontrolltabellen nollställs. Klarar station **B** inte testen i punkt 3., ses station **A** inte längre som möjlig huvudstation för station **B**. Utfallet i testet i punkt 3. registreras i kontrolltabellen.
 5. Punkt 1. till 4. repeteras till dess att inga nya stationspar testas.
- Processen punkt 1. till 5. ger ett antal grupper av stationer.

3.1.2.2 Sammankoppling av stationsgrupper

Stationsgrupperna kopplas ihop enligt följande (se även figur 2):

1. Huvudstationerna gås igenom en och en. Stationerna tas i ordning som de grupperades alltså huvudserier med flest datapunkter undersöks först, huvudserien med minst antal datapunkter tittas på sist.
2. Alla stödstationer som sorterats bort i tidigare kopplingar läggs till som potentiella stödstationer. Testet beskrivet i punkt 3 i stycket ovan görs för alla nya potentiella stödstationer.
3. Alla potentiella kopplingspunkter för alla tidsserier i gruppen hittas (streckade linjer i figur 2). Alla tidpunkter som är sista tidpunkten i en nivå listas. En nivå är data i följd där sporadiska uppehåll i datatäckning är upp till tre år. (Nivå 3 i figur 2 illustrerar ett uppehåll i station **A** på över tre år.)
4. Den mest lämpliga stationen för varje potentiell nivå (tiden mellan två potentiella kopplingspunkter eller slut-/startpunkten för hela gruppen) undersöks. Först undersöks det senaste nivån (nivå 1 i figur 2) och sist den äldsta nivån (nivå 10 i figur 2).
 - a. Stationen som används för föregående nivå (alltså kronologisk kommande nivå) noteras (exempelvis station **A** i nivå 6, figur 2).
 - b. Den av stationerna i gruppen som har mest data för nivån noteras (exempelvis station **B** i nivå 7, figur 2).
 - c. Om station **B** inte har mer än 10 % bättre datatäckningsgrad (alltså andelen befintlig data i den aktuella tidsperioden) än station **A** för nivån väljs station **A** för nivån (som i exemplet i figur 2), annars väljs station **B** (jämför motsvarande situation, nivå 9 i exemplet i figur 2).
5. De station som inte upptar någon nivå görs tillgängliga för andra stationsgrupper att kopplas till (station **D** i figur 2).
6. Huvudstationen i varje grupp byts från den station med flest datapunkter till den station som har data i den senaste nivån (station **A** i figur 2).

3.2 Homogenisering

Datasetet med sammankopplade tidsserier homogeniseras med HOMER med det automatiska skriptet Bart (version 0.17.10). Som en del av homogeniseringsproceduren i HOMER utfylls de luckor som finns i datasetet med hjälp av korrektionsmetoden ANOVA.

3.2.1 HOMER

HOMER (Mestre et al. 2013) är en produkt av projektet COST Action ES0601 (HOME) som pågick mellan 2007 och 2011 (Venema et al. 2012). Projektet syftade till att delvis utvärdera befintliga homogeniseringsmetoder och delvis till att konstruera ett nytt homogeniseringsverktyg utifrån resultaten.

HOMER är en relativ homogeniseringsmetod vilket innebär att varje tidsserie jämförs med andra liknande tidsserier. En absolut homogeniseringsmetod tittar på förändringar inom tidsserien. Svårigheten att skilja på klimatologiska och icke-klimatologiska förändringar gör absoluta metoder svårare att använda (Venema et al. 2020). HOMER räknas också som en "maximum likelihood"-metod, eftersom den i likhet med exempelvis ACMANT (Domonkos, 2020) bygger på den statistiska metoden "maximum likelihood" för att hitta homogenitetsbrott. Detta i motsats till metoder som använder hypotestest, exempelvis SNHT (Alexandersson 1986) och MASH (Szentimrey, 2011).

HOMER inkluderar ett antal funktioner för att detektera homogenitetsbrott och ANOVA-korrektionsmodellen (Causinus och Mestre, 2004) för att korrigera tidsserierna med hjälp av de detekterade homogenitetsbrotten samt en funktion för att hitta den optimala månaden för ett homogenitetsbrott, då homogenitetsbrott detekteras årsvis.

3.2.1.1 Detektionsfunktioner

HOMER inkluderar tre funktioner för att detektera homogenitetsbrott. Tidsserien för vilken homogenitetsbrott ska detekteras kallas kandidatserie, de tidsserier som används för att hitta homogenitetsbrott kallas referensserier. Referensserier är antingen tidsserier som är välkorrelerade med kandidatserien eller tidsserier vars stationer är närbelägna kandidatstationen.

1. **Parvis detektion:** En tidsserie som består av differensen mellan kandidatserien och en referensserie delas upp i ett antal segment genom att minimera den sammanlagda interna variansen (vilket motsvarar standardavvikelsen i kvadrat) inom alla segment enligt homogeniseringsverktyget PRODIGE (Causinus och Mestre, 2004). Antal segment bestäms genom ett "penalized likelihood criteria" MBIC (Modified Bayesian Information Criterion, Zhang och Siegmund, 2007). Det går inte att avgöra till vilken av de två serierna de homogenitetsbrott som upptäcks på detta sätt hör till. Därför provas varje kandidatserie mot ett antal referensserier; hör homogenitetsbrottet till kandidatserien uppträder brottet även i skillnaden med andra referensserier. Den här parvisa detektionsmetoden kan användas både på års- och årstidsmedel. Parvis detektion benämns som "pairwise detection" i litteraturen. I figur 3 visas ett exempel på den parvisa detektionsfunktionen.
2. **Sammanfogad detektion:** ANOVA-korrektionsmodellen konstruerar en "idealisk" tidsserie för korrektion och datautfyllnad. För närmare beskrivning se nedan. Denna idealiska tidsserie används också för att upptäcka homogenitetsbrott på samma sätt som i punkt 1, alltså att differensen mellan kandidatserien och den idealiska tidsserien delas upp i segment. Sammanfogad detektion benämns som "joint detection" i litteraturen.
3. **Årscykel detektion:** ACMANT (Domonkos, 2020) är ett närbesläktat homogeniseringsverktyg där det ingår en funktion för att hitta inhomogeniteter i årscykelns amplitud. Årscykel detektion kallas "ACMANT" eller "ACMANT-style detection" i litteraturen.

3.2.1.2 ANOVA-korrektionsmodell

HOMER inkluderar ANOVA-korrektionsmodellen (Causinus och Mestre, 2002) för korrektion och datautfyllnad. Metoden är beskriven i detalj i Domonkos (2022). Grundantagandet är att värdet för en klimatvariabel (x , exempelvis temperatur) kan formuleras som summan av en klimatterm (μ , gemensam för stationer $s_j \forall j \in [1, p]$ inom ett delnätverk av stationer som delar klimat, men variabel i tid $t_i \forall i \in [1, n]$), en stationsterm (v , specifik för varje station inom delnätverket, men konstant i tid mellan två homogenitetsbrott) samt en restterm (ε , vilket inkluderar väder, mätfel och diverse brus):

Ekvation 1

$$x_{i,j} = \mu_i + v_j + \varepsilon$$

Ett ekvationssystem kan då ställas upp med ekvationerna för medelvärden, dels medelvärden för varje tidssteg (här månad) över alla stationer i delnätverket och dels medelvärden för varje station över tidsserien (för varje period mellan två homogenitetsbrott):

Ekvation 2

$$\frac{1}{p} \sum_{j=1}^p x_{i,j} = \mu_i + \frac{1}{p} \sum_{j=1}^p v_j$$

Ekvation 3

$$\frac{1}{n} \sum_{i=1}^n x_{i,j} = \frac{1}{n} \sum_{i=1}^n \mu_i + v_j$$

Vidare kan summan av alla klimattermer antas vara noll; klimattermen beskriver således klimatologisk avvikelse. Ekvationssystemet har då $n + p + 1$ ekvationer med $n + p$ obekanta. När termerna μ_i och v_j på detta sätt är bestämda kan $x_{i,j}$ approximeras.

Ett minimalt exempel för att illustrera ANOVA-korrektionsmodellen finns i appendix 6.1.

3.2.1.3 Praktiskt användning av HOMER

De rutinen för användningen av HOMER som utvecklats inom det norska Meteorologisk institutet (Elin Lundestad, persolog kommunikation) beskrivs i figur 4:

1. Homogenitetsbrott detekteras för de råa serierna (med dataluckor) med den parvisa detektionsmetoderna (årsvis och säsongvis) samt den sammanfogade detektiosmetoden.
2. De råa serierna homogeniseras med hjälp av de godkända homogenitetsbrotten, se nedan.
3. Homogenitetsbrott detekteras med årscykelmetoden på de homogena tidsserierna.
4. De råa serierna homogeniseras med hjälp av de godkända homogenitetsbrotten.
5. Homogenitetsbrott detekteras för de homogena serierna (med dataluckor) med den parvisa detektionsmetoderna (årsvis och säsongvis) samt den sammanfogade detektiosmetoden.
6. Punkt 2 och 3 upprepas.
7. De optimala månaderna för varje homogenitetsbrott bestäms.
8. En slutgiltig homogenisering utförs.

HOMER presenterar resultaten av detektionsfunktionerna under punkterna 1, 3 och 5 och ger användaren möjlighet att godkänna eller underkänna de föreslagna homogenitetsbrott i varje tidsserie. Även kriterier för att tolka resultaten har utvecklats och kan översättas till ett poängsystem, se figur 5:

1. Varje homogenitetsbrott detekterade av parvis-årsvis-, sammanfogad- och årscykeldetektionsfunktionerna tilldelas tre poäng.
2. Homogenitetsbrott detekterade av parvis-säsongvis-detektionsfunktionen tilldelas en poäng.
3. Poäng tillhörande året innan eller året efter ett potentiellt homogenitetsbrott kan överföras till detta med ett avdrag på tre poäng.
4. Om ett potentiellt homogenitetsbrott kommer upp i tolv eller mer poäng godkänns brottet. Kommer ett potentiellt homogenitetsbrott upp i nio eller mer poäng men mindre än tolv uppsöks stöd för brottet i metadata (exempelvis dokumentation från mätstationerna). Hittas stöd i metadata godkänns brottet, om inte antecknas brottet för senare omgångar.
5. Två brott i samma tidsserie kan inte vara närmare i tid än två år.

3.2.1.4 Bart

Bart (Joelsson et al. 2022a) är en automatisering av HOMER enligt avsnitt 4.2.1.3. Bart har konstruerats för stor flexibilitet vad gäller HOMER:s användningsrutiner och bedömningskriterierna för homogenitetsbrott. Inom detta projekt har endast skönlägesinställningarna, beskrivna i avsnitt 4.2.1.3. använts. Kärnfunktionerna i HOMER är i princip orörda i Bart-algoritmen. En homogenisering med Bart motsvarar alltså en homogenisering med HOMER om ovan nämnda rutiner och kriterier används. Två tillägg har dock gjorts, båda gäller valet av referensserier:

1. En hybridfunktion (hybrid mellan en funktion som väljer referensstationer baserat på korrelation mellan kandidatstationen och referensstationen och en funktion som väljer referensstationer baserat på det geografiska avståndet mellan kandidatstationen och referensstationen) är utvecklad.
2. Ett tröskelvärde för det maximala antalet stationer finns implementerat för att begränsa beräkningstiden.

Ett problem med att använda korrelationskriteriet för att välja referensstationer är att överlappet mellan två serier kan inverka på korrelationskoefficienten. Ett problem med att använda geografisk närhet mellan två stationer är att dessa två stationer i princip skulle kunna beskriva skilda klimat trots att de är närbelägna. Ett problem med båda dessa metoder är att de inte säkerställer datatäckningen över tid. Den nya hybridfunktionen kategoriserar alla potentiella referensstationer enligt korrelationen, inom varje kategori sorteras referenserna enligt hur närbelägna de är. Funktionen säkerställer att minst tre referensserier har 95 % gemensam täckning med kandidatserien under hela dess täckningsperiod.

Bart finns beskriven i detalj i bilaga 1: Bartguide.

3.2.1.5 Metadata

Som inmatning till Bart ges en lista på potentiella homogenitetsbrott, så kallade "metadatabrott". I detta projekt listas tre typer metadatabrott:

- i. Tidpunkter för sammankopplingar enligt avsnitt 3.1.
- ii. Tidpunkter för "hårda" sammankopplingar, alltså där två stationer sammankopplats i den digitala databasen MORA.

- iii. Tidpunkter för flyttar, observatörsbyten, utrustningsbyten och andra potentiella homogenitetsbrott som nämns i inspektionsprotokoll och liknande dokumentation ("metadata"). Hittills har en liten andel av all tillgänglig metadata digitaliserats.

Potentiella homogenitetsbrott behandlas i Bart som ett homogenitetsbrott detekterat av någon av detektionsfunktionerna, vilket i praktiken motsvarar användningen av metadata i det traditionella utvärderingen av potentiella homogenitetsbrott, se avsnitt 3.2.1.3. Samtliga metadatabrott är listade i tabell 1.1 i bilaga 2: Tidsserier.

3.2.2 Korrigering av gradvisa inhomogeniteter

Gradvisa förändringar kan ofta representeras av ett antal stegfunktioner snarare än linjära trendkorrektioner eftersom gradvisa förändringar ofta inte är linjära i tid (Venema et al., 2020). Homogeniseringen från 2014 innehöll linjära trendkorrektioner för åtta stationer: Lund, Uppsala, Stockholm, Svenska Högarna, Falun, Härnösand, Frösön och Umeå. Stockholm har två separata perioder av trendkorrigeringar. Detta ger en möjlighet att jämföra dessa linjära trendkorrigeringar med motsvarande stegfunktionskorrigeringar.

4 Resultat

Resultaten av homogeniseringen av månadsmedeltemperatur i Sverige 2022 är beskriven här. Resultaten av de automatiska stationskopplingarna (se avsnitt 4.1) är presenterade i avsnitt 5.1 och homogeniseringen av det sammankopplade nätverket finns presenterad i avsnitt 5.2.

4.1 Stationskopplingar

Det sammankopplade datasetet består av data från 836 stationer i 456 sammankopplade tidsserier. I tabell 1 finns datatäckningen och antalet saknade data listade för det ursprungliga och det sammankopplade datasetet samt den del av det ursprungliga datasetet som kan användas av HOMER (tidsserier med minst 15 års data). Det sammankopplade datasetet inkluderar något fler data och har nästan hälften av saknad data jämfört med det ursprungliga datasetet med minst 15 års data per tidsserie. Datatäckningen för det sammankopplade datasetet är 44 %, motsvarande för de serier med minst 15 år är 30 %. Hur datatäckningen varierar i tid är beskrivet i figur 6.

Fördelningen av tidsseriernas längd är förskjutna mot längre stationer i det sammankopplade datasetet, se figur 7. Exempelvis är antalet tidsserier med minst 60 års data (längden av två normalperioder) i det sammankopplade datasetet 257 jämfört med 173 i det ursprungliga datasetet.

Samtliga stationskopplingar är presenterade i bilaga 3: Stationskopplingar i figurerna 1.1–1.456 samt i tabell 1.1. Stationskopplingarna är jämförda med stationskopplingarna i 2014 års homogeniserade dataset i figur 2.1–2.21.

4.2 Homogenisering

4.2.1 Homogenitetsbrott

Under homogeniseringsprocessen av de 456 sammankopplade tidsserierna hittades totalt 1 788 homogenitetsbrott. För 22 tidsserier hittas inga homogenitetsbrott, för 90 tidsserier hittas tre homogenitetsbrott, vilket är det vanligast förekommande antal homogenitetsbrott. Antalet homogenitetsbrott per tidsserie finns presenterat som histogram i figur 8.

Den genomsnittliga tiden mellan två homogenitetsbrott per tidserie presenteras som histogram i figur 9. Mediantiden för den genomsnittliga tiden mellan två homogenitetsbrott sett över alla tidsserier är 17 år. För att se hur homogenitetsbrotten är fördelade i tid är antalet brott per station och decennium presenteras som ett histogram i figur 10.

37 % av de godkända homogenitetsbrotten har stöd i metadata eller i sammanfaller med kopplingspunkter. 34 % av de föreslagna homogenitetsbrotten i metadata motsvarar ett godkänt brott.

Referensnätverket för varje tidsserie finns beskriven i bilaga 4: Referenstidsserier. Samtliga godkända homogenitetsbrott är listade i tabell 1.2 i bilaga 2.

4.2.2 Korrektion

60 % av datasetet är korrigerad av homogeniseringen, se figur 11. 42 % av data är korrigerad med mindre än 0,5 °C, 3 % av data är korrigerade med 1 °C eller mer. 37 % av datapunkterna har en negativ korrektion (observationerna har en högre temperatur än det homogeniserade värdet), 24 % positiv korrektion. Korrektionen är störst både vad gäller det genomsnittliga värdet och det genomsnittliga absolutbeloppet för normalperioden 1871–1900: –0,2 °C respektive 0,3 °C. Se figur 10. Korrektionen för 1991–2020 är mindre än en tiondels grad för både det genomsnittliga värdet och det genomsnittliga absolutbeloppet. Korrektionen är i genomsnitt störst för vårmånaderna och minst för höstmånaderna. Se figur 11. Skillnaden är liten (mindre än en tiondels grad) men signifikant.

4.2.3 Korrektion av gradvisa inhomogeniteter

Jämförelse mellan 2014 och 2022 års homogeniserade dataset för de åtta stationer med linjära trendkorrektioner (Lund, Uppsala, Stockholm, Svenska Högarna, Falun, Härnösand, Frösön och Umeå) illustreras i figur 3.1–3.8 i bilaga 2.

I sju av åtta tidsserier (alla utom Härnösand) finns stegfunktionskorrektionen i 2022 års homogeniserade dataset under perioden för den linjära trendkorrektionen i 2014 års homogeniserade dataset. I sex av dessa sju fall (alla utom Umeå) är tecken för nettokorrektionen (alltså skillnaden mellan korrektionerna vid början och slutet på perioden) för perioden samma i 2022 och 2014 års homogeniserade dataset. I sex av de sju fallen (alla utom Falun-Lugnet) är nettokorrektionen under perioden större i 2014 års dataset än i 2022 års dataset.

4.2.4 Exempel: Jokkmokk

Regelbundna temperaturmätningar startade i Jokkmokk november 1860 på Kungliga Vetenskapsakademins initiativ och är därmed en av ett 30-tal stationer i MORA som startade innan 1861. Månadsmedeltemperaturserien ifrån Jokkmokk (med stationsnumret 169880) är MORA:s åttonde mest omfattande serie i fråga om antalet månadsvärden (1 873 stycken) med en cirka en procent saknade värden mellan 1860 och 2018 då stationen lades ner. Jokkmokk har därför ingått i de 35 stationer som tidigare utgjort det homogeniserade nätverket och därmed också ingått som underlag för klimatindikatorn Temperatur. Temperaturserien från Jokkmokk fick viss uppmärksamhet i ett nyhetsinslag i SVT (2020) eftersom temperaturen där beräknas ha ökat mest av SMHI:s stationsplatser sedan 1860.

Tidsserien ifrån stationen 169880 Jokkmokk innehåller inga kopplingar ifrån den automatiska kopplingsprocessen utan betraktas i MORA som en lång sammanhängande serie. Stationen har dock flyttat cirka 13 gånger. Dessutom finns instrumentbyten och observatörsbyten registrerade, dock är anteckningarna förmodligen inte helt kompletta.

Fem homogenitetsbrott detekteras i Jokkmokk (1958, 1975, 1978, 1992 och 2000). Tre av dessa brott finns registrerade i metadata (1958, 1975 och 1978). En flytt finns registrerad 1957-08-01. Här finns också en kommentar om att temperaturburen ska ersättas samt att temperaturburen "kommer stå på en altan". Även 1975 finns en flytt registrerad till en 11 m lägre belägen plats än det tidigare registrerade läget. År 2000 flyttar stationen till en 6 m lägre belägen plats än det tidigare läget.

Mellan 1860 och 1957 justeras temperaturen ned 0,71 °C på årsnivå (märk dock att varje månad justeras med olika korrektionstermer). Mellan 1958 och 1974 justeras temperaturen ned 0,97 °C. Efter flytten runt 1958 mäts således något högre temperaturer än den föregående perioden, möjligtvis inverkade en icke-optimal placering av temperaturburen. Mellan 1975 och 1978 justeras temperaturen ner 0,80 °C. Efter flytten 1975 mäts något lägre temperaturer än

föregående period, förändringen i altitud förklarar inte denna förändring. Mellan 1979 och 1992 justeras temperaturen ned 0,38 °C, mellan 1993 och 1999 0,26 °C. Efter 2000 är stationen ojusterad. Efter flytten 2000 mäts således något lägre temperaturer än föregående period, förändringen i altitud förklarar inte heller denna förändring.

Fram till den sista 30-årsperioden (1989–2018) ökade årsmedeltemperaturen från den första 30-årsperioden (1860–1889) med 1,8 °C. I det råa datasetet ökar temperaturen med 1,2 °C. Störst är förändringen på vintern (december, januari, februari) med 2,7 °C, minst på sommaren (juni, juli, augusti) med 0,6 °C. Även justeringen (skillnaden mellan rådata och den homogeniserade värdena) är mellan den första och sista 30-årsperioderna störst på vintern med 0,9 °C och minst på sommaren med 0,3 °C.

5 Diskussion

Vilken temperatur som skulle ha observerats vid en historisk tidpunkt på platsen för en nu aktiv station med en modern metod är för alltid höljt i dunkel. Homogenisering syftar till att med god sannolikhet närma sig detta idealiska värde. Det är omöjligt att veta huruvida homogenisering är korrekt, särskilt då metadata är ofullständig. Det har dessutom visat sig att olika homogeniseringsmetoder ger olika resultat (Joelsson et al. 2022a, 2022b, Venema et al. 2012).

Homogenisering är ett asymmetriskt problem: ett värde kan vara sant i mening att det är korrekt, men det kan också vara sant i meningen att det är observerat. Om exempelvis ett korrekt värde för en tidpunkt i en mätserie är 10 °C och det observerade värdet är 11 °C är felet alltså 1 °C. Om en homogenisering ger att värdet är 9 °C är felet alltså lika stort oavsett om det observerade värdet eller det homogeniserade värdet används. Dock är det observerade värdet sant i mening att det är observerat och bör i detta fall föredras framför det felaktiga homogeniserade värdet. Detta resonemang får till följd att homogenisering inte endast ska syfta till att minska felet utan minska felet med minsta möjliga åverkan på rådata.

Antalet brott som detekteras av en homogeniseringsmetod har i en studie med syntetisk data och kända homogenitetsbrott visat sig till viss del bero på känsligheten i metoden (Joelsson et al. 2022a); en känsligare metod detekterar fler brott än en mer okänslig metod, både verkliga och falska brott. Fler detekterade (och korregerade) brott ger generellt ett lägre medelfel (brott lönar sig). Det är dock viktigt att väga denna minskning i fel med åverkan på rådata då skillnaden mellan det observerade och det homogeniserade värdet ofrånkomligen ökar med antalet brott. När verklig data homogeniseras är felet okänt vilket uppmanar till viss försiktighet.

Den genomsnittliga tiden mellan två homogenitetsbrott i den aktuella homogeniseringsstudien är 17 år. Detta motsvarar det typiska värdet för en västerländsk mätstation som är ett homogenitetsbrott var 15:e till 20:e år (Venema et al. 2012). Detta ger stöd för lämplig nivå av känsligheten i den aktuella metoden.

Den nuvarande metoden innehåller inga gradvisa korrektionstermer, på inrådan av WMO (Venema et al. 2020). 2014 års homogenisering innehöll nio linjära trendkorrektioner fördelade på åtta stationer. Nettokorrektionen under dessa perioder har i de flesta fall samma tecken i båda homogeniseringsdataseten, men är mindre i det aktuella datasetet. För mellanliggande värden är i många fall skillnaden större än nettokorrektionen. Den bör poängteras här att det är omöjligt att avgöra vilken av homogeniseringarna som är mest korrekt.

Ytterligare en öppen fråga är hur många stationer som ska ingå i ett homogeniserat dataset. Idealiskt används många tidsserier; detta ökar robustheten i homogeniseringen då varje serie kan jämföras med fler klimatologiskt lika kandidatserier. Fler tidsserier kan dock i realiteten innebära fler korta tidsserier och tidsserier av lägre kvalitet. Sammankoppling av tidsserier råder till viss del bot på andelen saknade data, men inför å andra sidan inhomogeniteter i datasetet. En annan fördel med ett stort antal tidsserier är att det förmodligen ökar robustheten och representativiteten i regionala klimatanalyser. Ett mindre dataset, kan visserligen ha bättre kvalitet och täckningsgrad, men ger varje tidsserie färre klimatologiskt lika tidsserier att jämföras med. I det aktuella datasetet ingår så stor del av det tillgängliga datasetet som möjligt.

En möjlighet vid sammankopplingen var att använda de stationsföljningar som ligger till grund för datasetet "Månadens väder och vatten" i tillägg till de automatiska kopplingarna. Detta skulle förenkla jämförelser mellan dessa två dataset. För att sammankopplingen ska vara konsekvent har dock hela nätverket sammankopplats med den automatiska metoden, men tröskelvärden har hämtats från "Månadens väder och vatten"-datasetet för att göra de två dataseten så lika som möjligt.

Den aktuella homogeniseringen tar ingen hänsyn till hur många aktiva referensstationer som finns vid en tidpunkt vid utvärderingen av homogenitetsbrott. Eventuellt är detta en bidragande anledning till att homogenitetsbrotten inte fördelar sig jämnt i tid för de aktiva tidsserierna. Utvärderas homogenitetsbrott relativt hur många aktiva referenser som finns tillgängliga skulle detta kunna leda till mindre robust homogenisering då brott skulle accepteras med mindre stöd i områden och tidsperioder med mindre datatäckningsgrad. Utvärderas homogenitetsbrott däremot med absoluta antal brottssignaler försvårar det detektionen av homogenitetsbrott i dessa områden och under dessa perioder. Det är tänkbart att en relativ utvärdering kommer att testas vid framtida homogeniseringar.

Detektionen av årscykelshomogenitetsbrott är i den aktuella homogeniseringen inte skild från inhomogeniteter i medelvärde. Ett homogenitetsbrott endast i årscykeln kommer alltså inte kunna detekteras. Brott på års- och säsongsnivå är i den aktuella homogeniseringen inte särskilda. Den beprövade traditionella metoden har istället använts. Särskiljning av olika typer av homogenitetsbrott kommer eventuellt testas vid kommande homogeniseringar.

Homogeniseringen ger i det aktuella fallet inte någon signifikant skillnad på den övergripande temperaturtrenden för Sverige sett över hela perioden (Joelsson et al. 2022b), dock indikerar linjära passningar av temperaturerna i det homogeniserade datasetet en något snabbare uppvärmning än motsvarande för det råa datasetet. Detta är en följd av att homogeniseringen korrigerar fler värden negativt än positivt, historiska observationer antas alltså vara för "varma" snarare än för "kalla". Detta är vad som förväntas då en suboptimal mätmetod (exempelvis dåligt placerad termometer eller otillräckligt strålningsskydd) tenderar att mäta för höga temperaturer snarare än för låga och då mätmetoderna antas ha förbättrats över tid.

6 Slutsatser

Data från 836 tidsserier för månadsmedeltemperatur är automatisk sammankopplade till 456 tidsserier med en ny automatisk metod. Antalet tidsserier med mer än 60 års data ökar i och med sammankopplingarna från 173 till 257 stycken. Antalet värden som behöver interpoleras in en motsvarande homogenisering utan sammankopplingar minskas från 70 % till 55 %.

Det sammankopplade datasetet homogeniseras med en automatiserad version av HOMER, baserad på en traditionell procedur. HOMER hittar och korrigerar totalt 1 788 homogenitetsbrott, vilket motsvarar i genomsnitt cirka 17 år mellan två homogenitetsbrott och att 60 % av datasetet korrigeras varav 37 % negativt (mätningarna för "varma") och 24 % positivt (mätningarna för "kalla"). Tre procent av datasetet korrigeras med 1 °C eller mer. Tjugotvå serier anses homogena.

Saknade data fylls ut med hjälp av ANOVA-korrektionsmetoden.

De linjära trendkorrektionerna i 2014 års data återskapas till viss del med stegfunktionskorrektioner i 2022 års data, dock är nettokorrektionerna överlag mindre.

7 Appendix

7.1 ANOVA-korrektionsmetoden, minimalt exempel

För att illustrera modellen kan ett minimalt hypotetiskt exempel ställas upp. Låt matrisen $\mathbf{X}$ vara månadsmedeltemperaturen för tre månader $t_i \forall i \in \{1, 2, 3\}$, för tre stationer $s_j \forall j \in \{1, 2, 3\}$. Raderna i matrisen motsvarar stationerna och kolonnerna i matrisen motsvarar tiden.

$$\mathbf{X} = \begin{bmatrix} -0,1 & 0,9 & 1,9 \\ 0,0 & 1,0 & 2,0 \\ 0,1 & 1,1 & 2,1 \end{bmatrix}$$

Genomsnittet för varje station över hela perioden beräknas enligt ekvation 2:

$$\frac{(-0,1 + 0,9 + 1,9)}{3} = 0,9 = \frac{\mu_1 + \mu_2 + \mu_3}{3} + v_1$$

$$\frac{(0,0 + 1,0 + 2,0)}{3} = 1,0 = \frac{\mu_1 + \mu_2 + \mu_3}{3} + v_2$$

$$\frac{(0,1 + 1,1 + 2,1)}{3} = 1,1 = \frac{\mu_1 + \mu_2 + \mu_3}{3} + v_3$$

Genomsnittet för varje månad över nätverket beräknas enligt ekvation 3:

$$\frac{(-0,1 + 0 + 0,1)}{3} = 0 = \mu_1 + \frac{v_1 + v_2 + v_3}{3}$$

$$\frac{(0,9 + 1,0 + 1,1)}{3} = 1 = \mu_2 + \frac{v_1 + v_2 + v_3}{3}$$

$$\frac{(1,9 + 2,0 + 2,1)}{3} = 2 = \mu_3 + \frac{v_1 + v_2 + v_3}{3}$$

Slutligen:

$$\mu_1 + \mu_2 + \mu_3 = 0$$

Ekvationsystemet kan exempelvis lösas enligt ersättningsprincipen:

$$v_1 = 0,9$$

$$v_2 = 1,0$$

$$v_3 = 1,1$$

$$\mu_1 = -\left(\frac{0,9 + 1,0 + 1,1}{3}\right) = -1$$

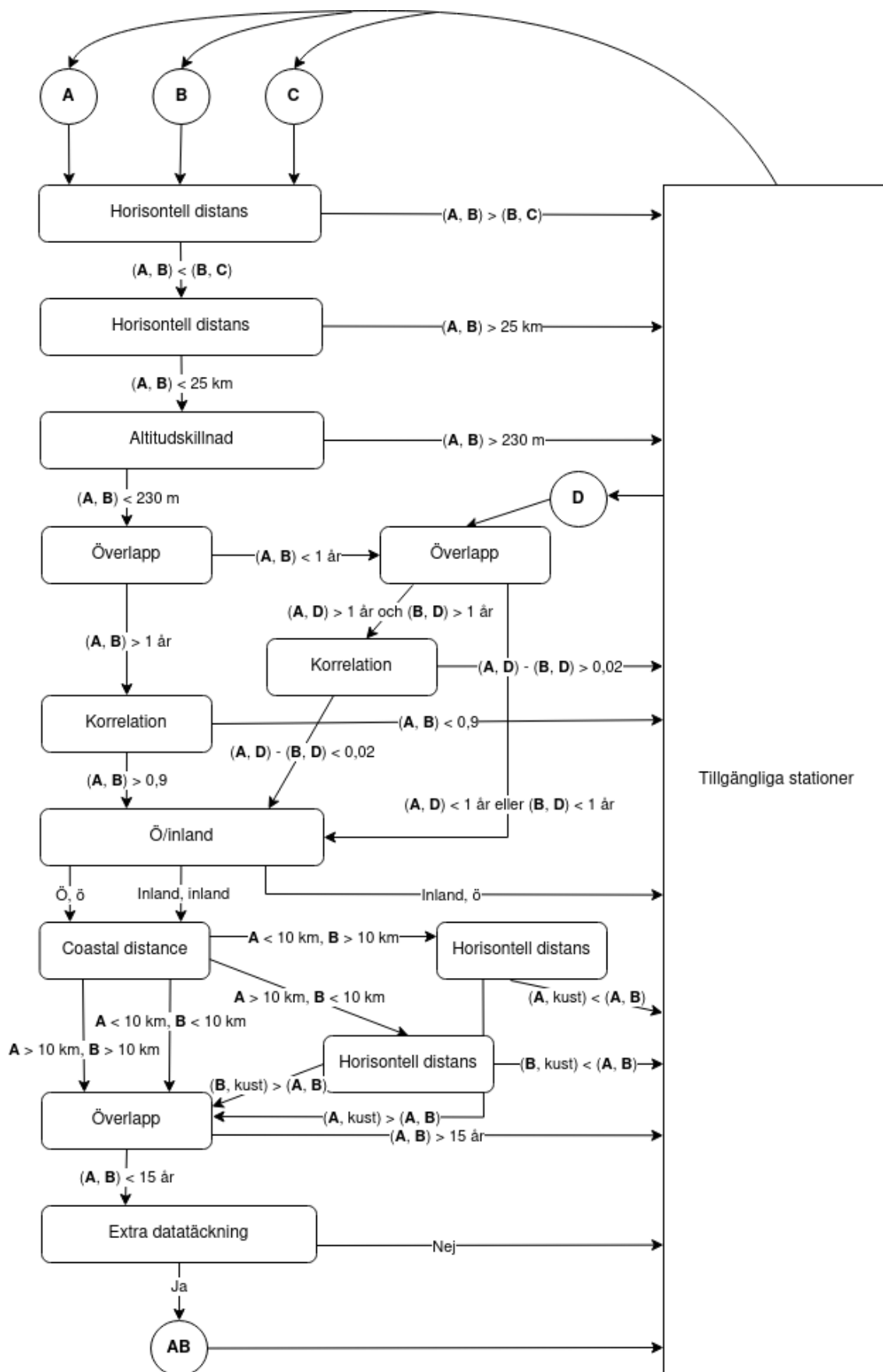
$$\mu_2 = 1 - \left(\frac{0,9 + 1,0 + 1,1}{3}\right) = 0$$

$$\mu_3 = 2 - \left(\frac{0,9 + 1,0 + 1,1}{3}\right) = 1$$

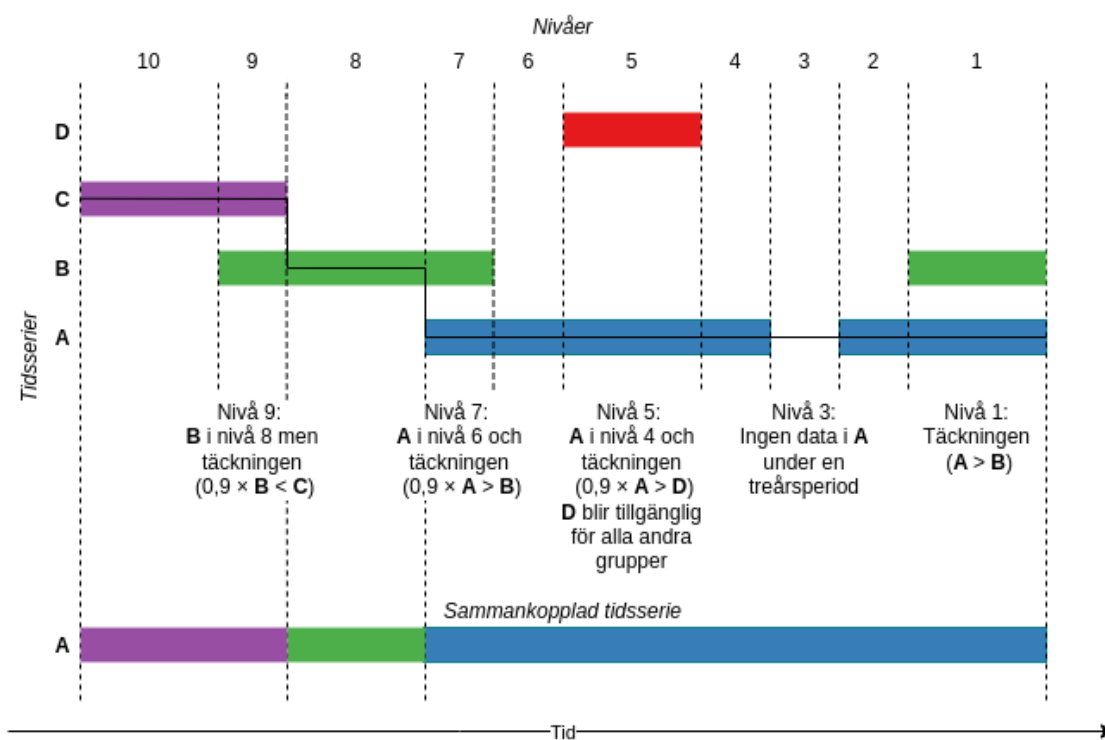
8 Referenser

- World Meteorological Organization, 2017. WMO guidelines on the calculation of climate normals.
- Venema, V.K., Mestre, O., Aguilar, E., Auer, I., Guijarro, J.A., Domonkos, P., Vertacnik, G., Szentimrey, T., Stepanek, P., Zahradnicek, P. och Viarre, J., 2012. Benchmarking homogenization algorithms for monthly data. *Climate of the Past*, 8(1), s. 89–115.
- Conrad, L. och Pollack, L.W., 1950. *Methods in Climatology*, Harvard University Press, Cambridge, Mass., s. 222..

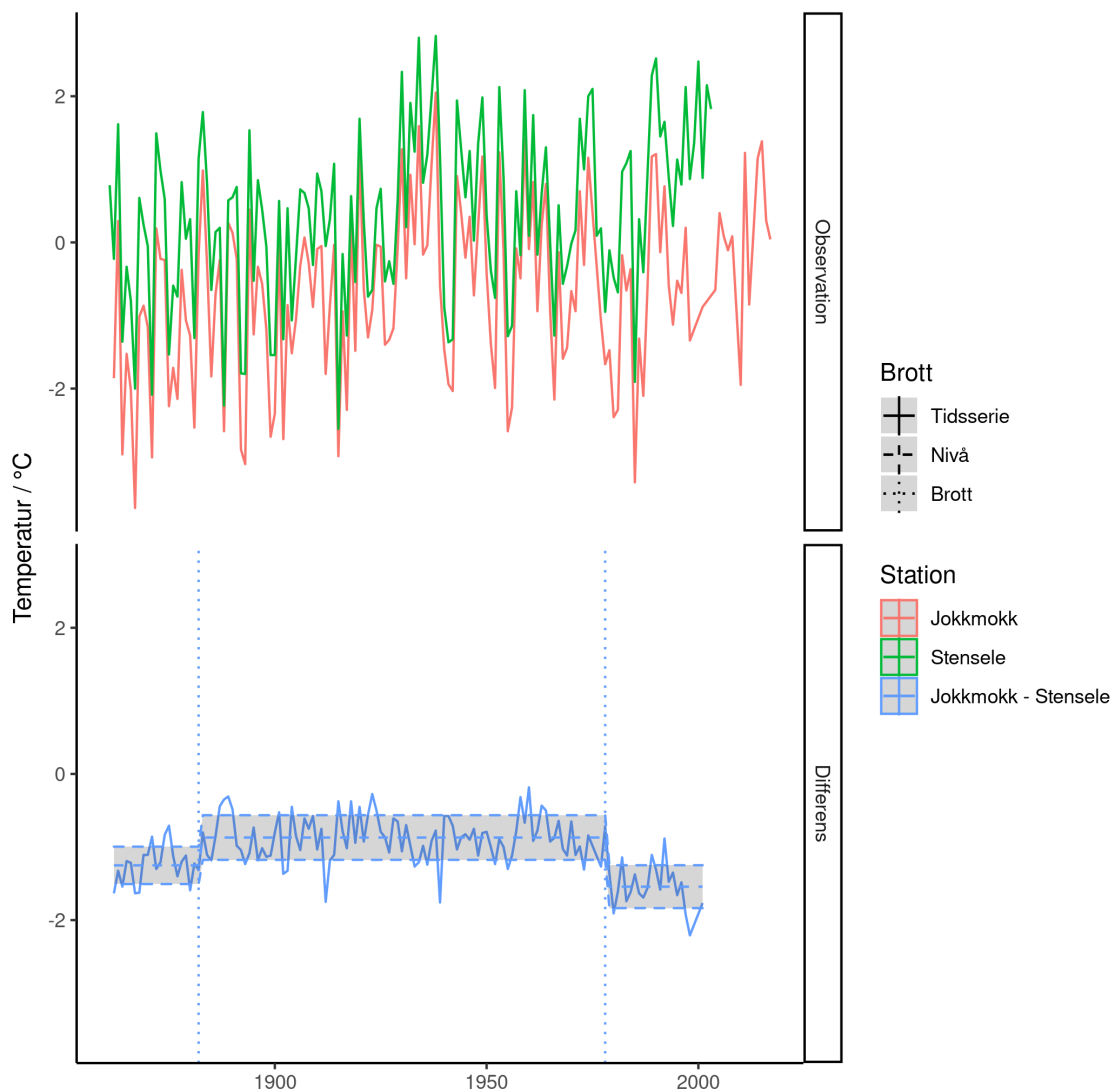
- Ribeiro, S., Caineta, J. och Costa, A. C., 2016. Review and discussion of homogenisation methods for climate data. *Physics and Chemistry of the Earth, Parts A/B/C*, 94, s. 167–179.
- Mestre, O., Domonkos, P., Picard, F., Auer, I., Robin, S., Lebarbier, E., Böhm, R., Aguilar, E., Guijarro, J., Vertachnik, G. och Klancar, M., 2013. HOMER: a homogenization software—methods and applications. *Idöjárás*, 117(1), s. 47–67.
- Venema, V., Trewin, B., Wang, X., Szentimrey, T., Lakatos, M., Aguilar, E., ... och Rasul, G., 2020. Guidelines on Homogenization, 2020 Edition. World Meteorological Organization.
- Alexandersson, H., 1986. A homogeneity test applied to precipitation data. *Journal of climatology*, 6(6), s. 661–675.
- Joelsson, L. M. T., Sturm, C., Södling, J., Engström, E., och Kjellström, E. (2022a): Automation and evaluation of the interactive homogenization tool HOMER. *International Journal of Climatology*, 42(5), 2861–2880. <https://doi.org/10.1002/joc.7394>
- Joelsson, L. M. T., Engström, E. och Kjellström, E. (2022b). Homogenisation of Swedish mean monthly temperature series 1860–2021. *International Journal of Climatology*.
- Szentimrey, T., 2011: Manual of homogenization software MASHv3.03. Hungarian Meteorological Service
- Bergström, H. och Moberg, A., 2002. Daily air temperature and pressure series for Uppsala (1722–1998). In *Improved Understanding of Past Climatic Variability from Early Daily European Instrumental Sources* (s. 213–252). Springer, Dordrecht.
- Moberg, A., Bergström, H., Krigsman, J.R. och Svanered, O., 2002. Daily air temperature and pressure series for Stockholm (1756–1998). In *Improved understanding of past climatic variability from early daily European instrumental sources* (s. 171–212). Springer, Dordrecht.
- Alexandersson, H., 2002. Temperatur och nederbörd i Sverige 1860–2001.
- Caussinus, H. och Mestre, O., 2004. Detection and correction of artificial shifts in climate series. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 53(3), s. 405–425.
- Domonkos P., 2020. ACMANTv4: Scientific content and operation of the software. <https://github.com/dpeterfree/ACMANT/tree/ACMANTv4.4>. (2022-09-15)
- Sveriges Television (SvT), Unik beräkning: Sverige värms upp dubbelt så snabbt som övriga jordklotet, 2020. <https://www.svt.se/nyheter/inrikes/unik-berakning-sverige-varms-upp-dubbelt-sa-snabbt-som-ovriga-jordklotet>. (2022-08-22)



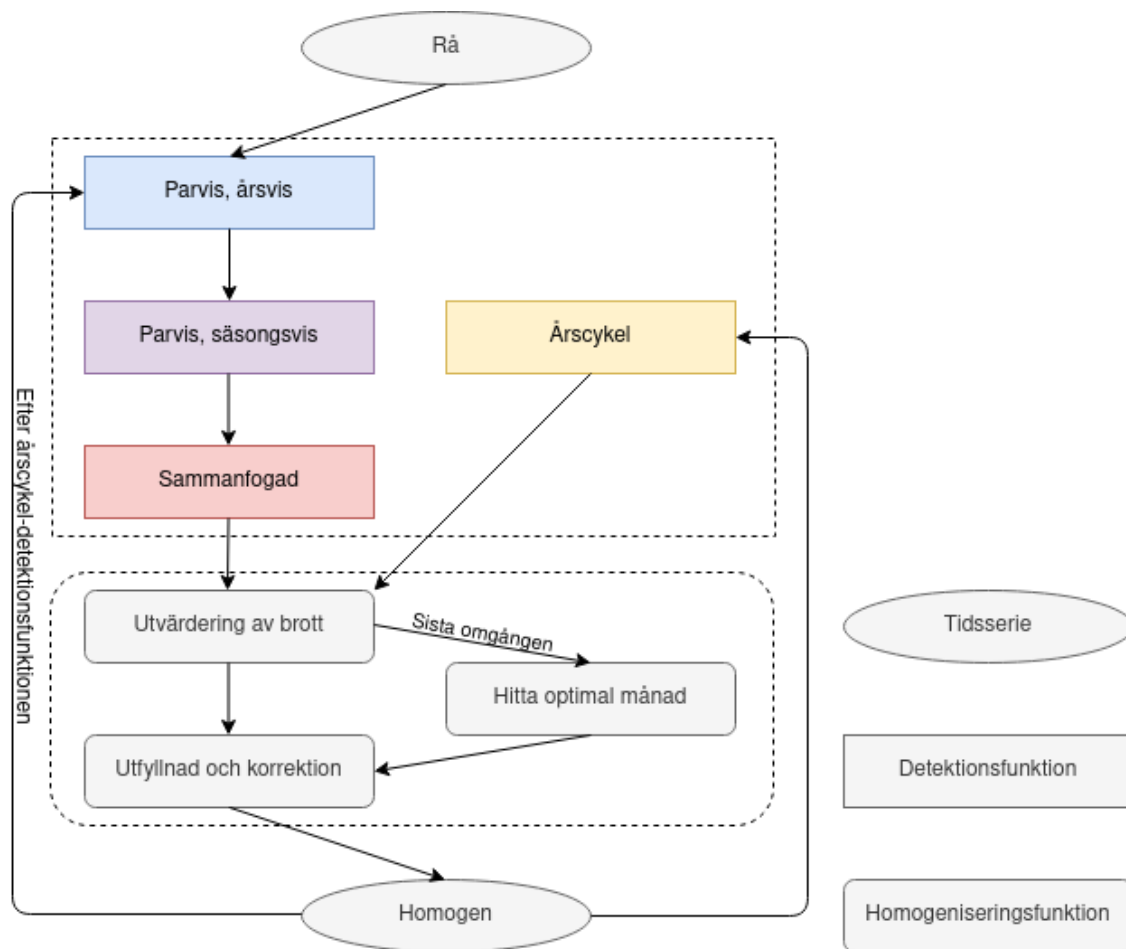
Figur 1: Flödesschema över grupperingsprocessen i den automatisk kopplingsmetoden



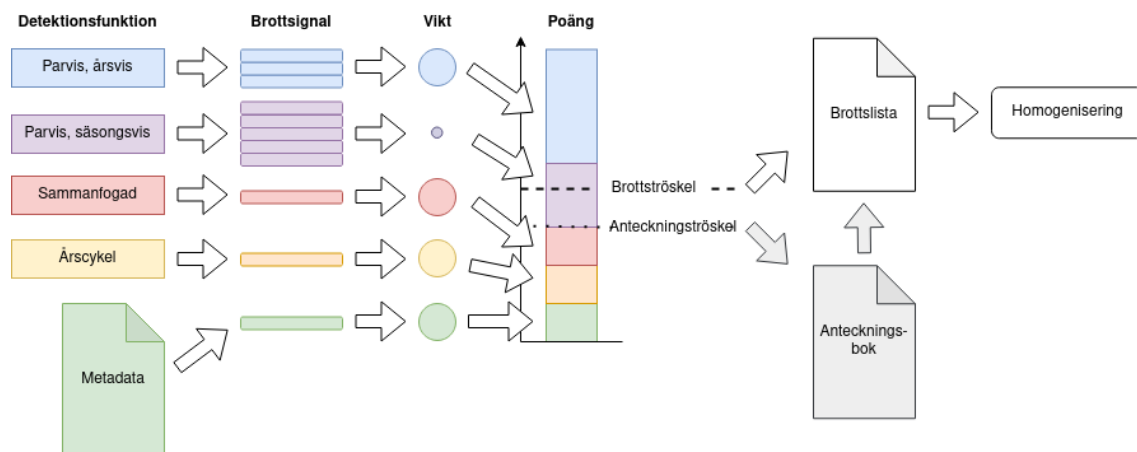
Figur 2: Illustration av ett fiktivt exempel av en sammankoppling av stationsgruppen ABCD



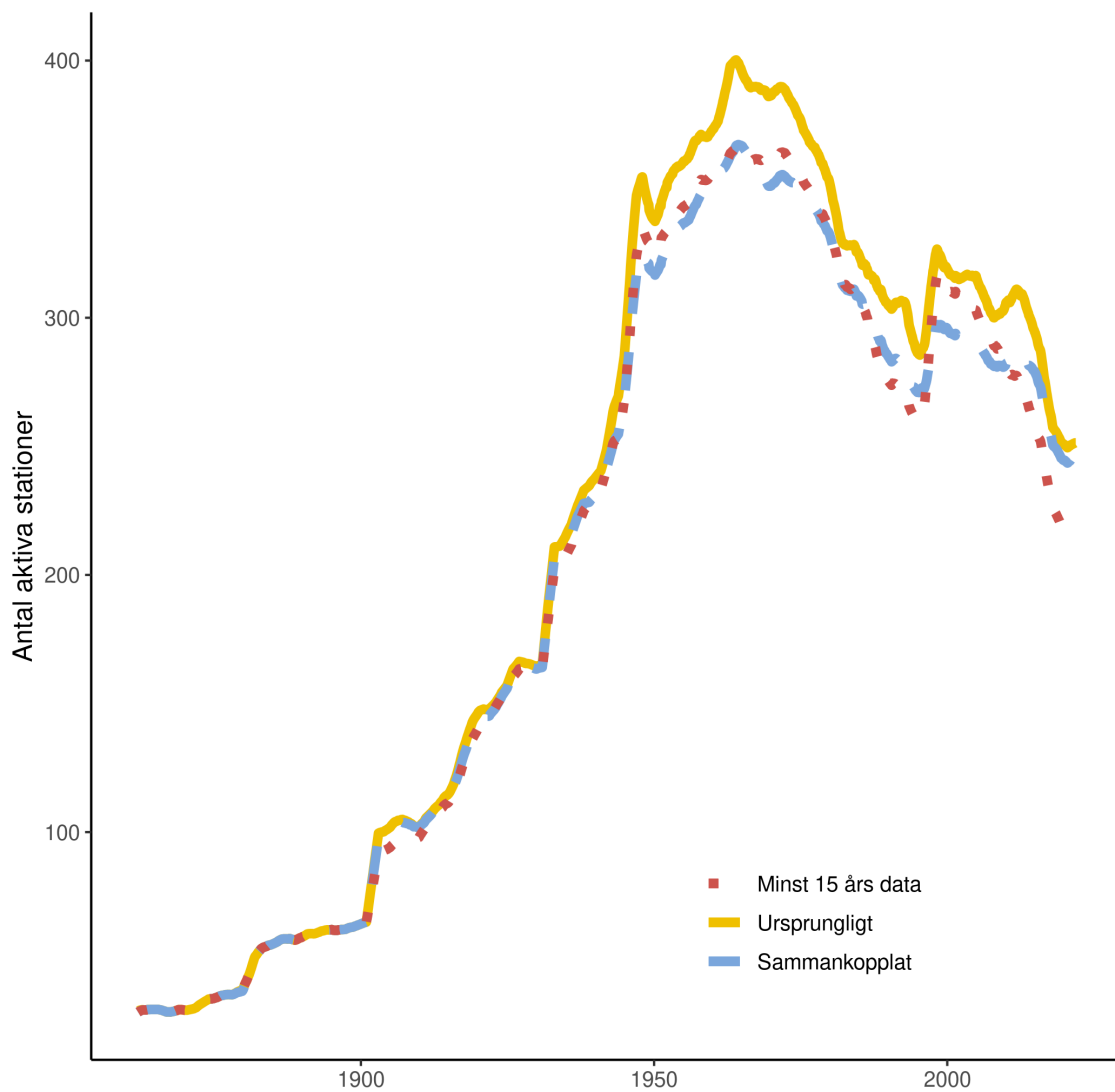
Figur 3: Exempel på parvis detektion av homogenitetsbrott mellan Stensele och Jokkmokk, den övre panelen visar de observerade årsmedeltemperaturerna för Stensele (grön) respektive Jokkmokk (röd), den nedre panelen visar skillnaden mellan de två serierna (heldragen linje), de detekterade homogenitetsbrotten (prickade linjer) samt medeltemperaturen och standaravvikelsens övre och nedre gräns för varje nivå (streckade linjer). De detekterade brotten utplaceras så standaravvikelesen (grått fält) minimeras.



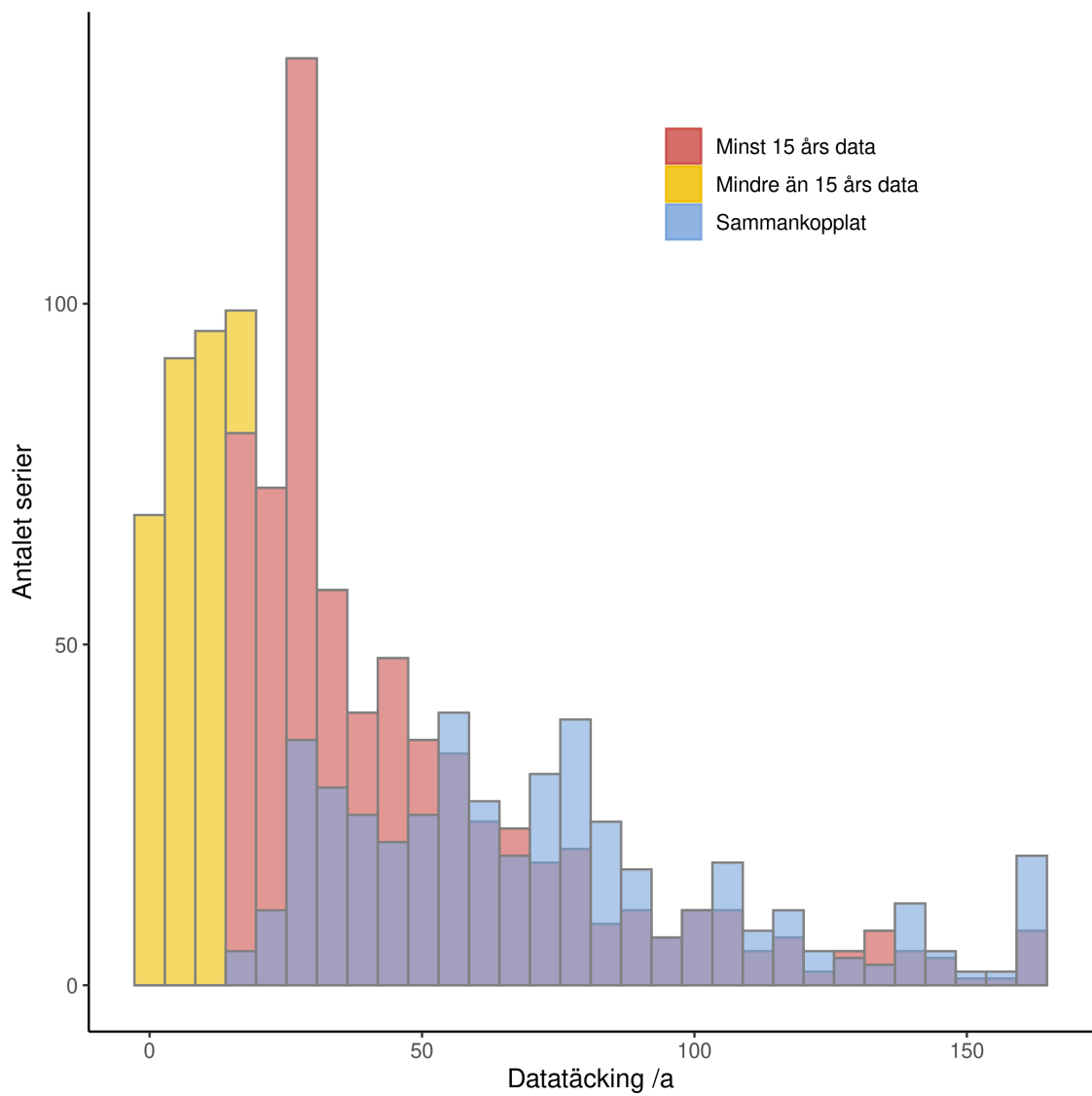
Figur 4: Flödesdiagram, rutin för användningen av HOMER/Bart



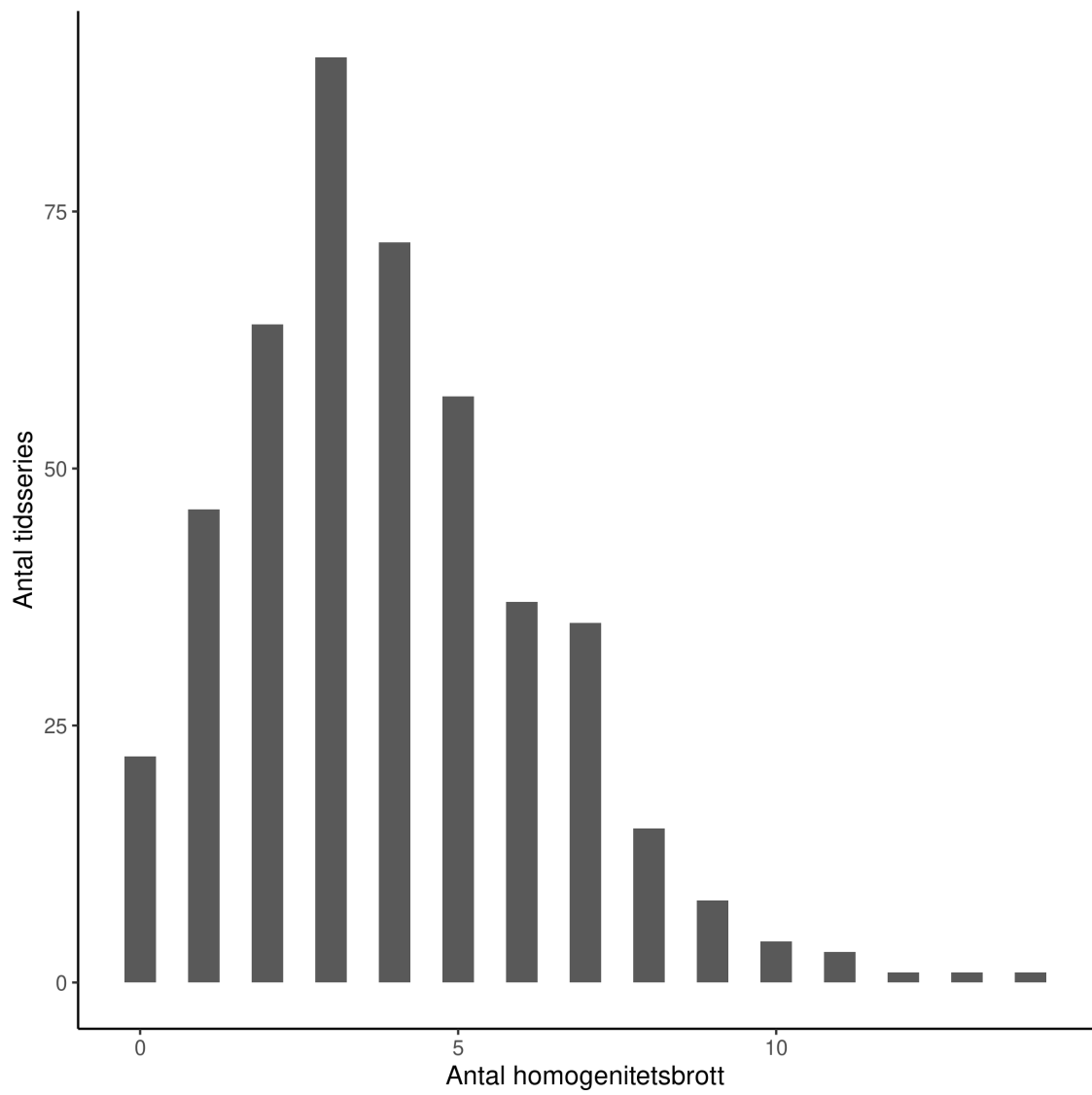
Figur 5: Utvärdering av brott i HOMER/Bart



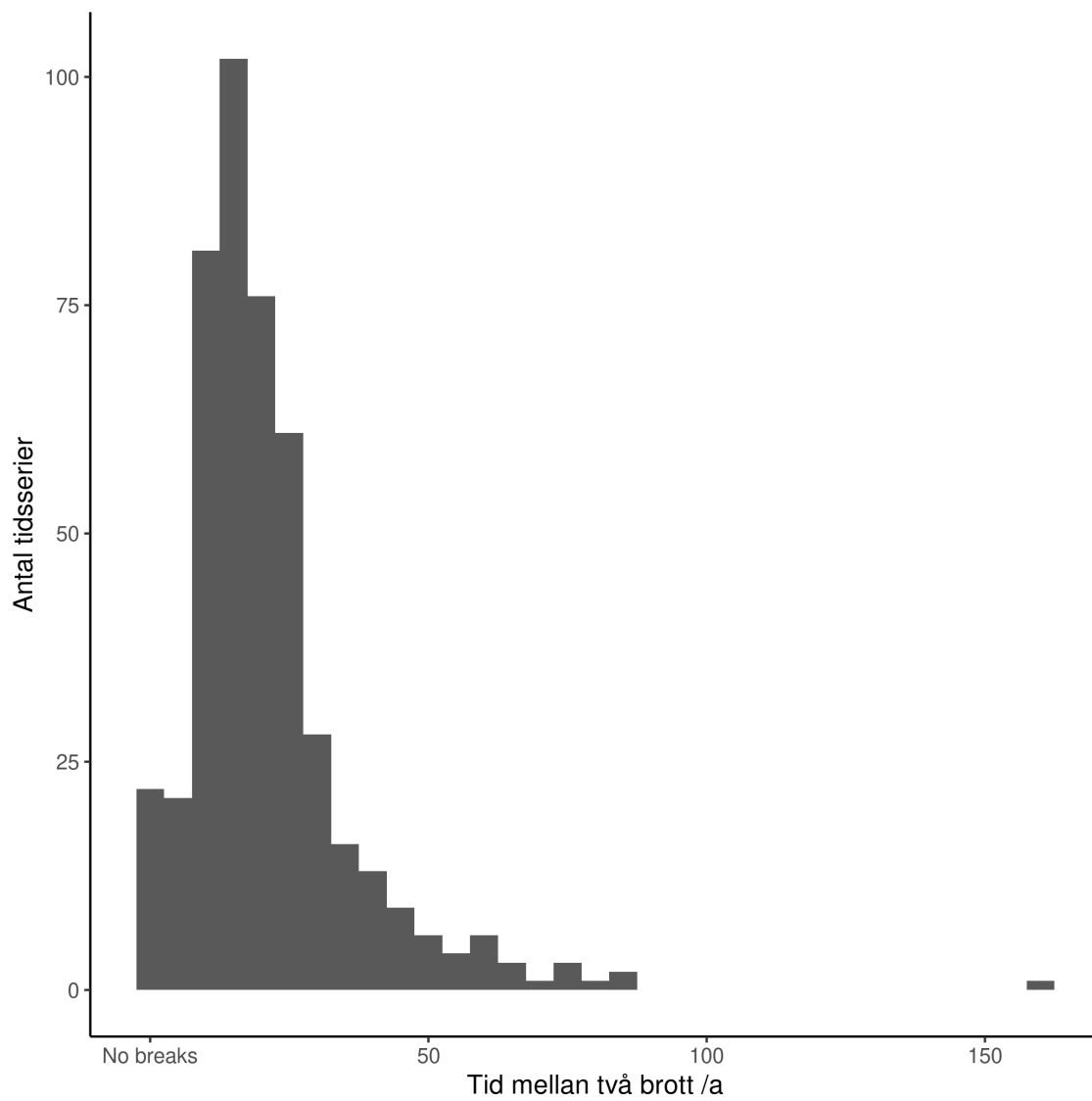
Figur 6: Antalet aktiva stationer med digitaliserade tidsserier 1860–2021 för det ursprungliga datasetet (gul heldragen linje), ursprungliga tidsserier med mer än 15 års data (röd prickad linje) samt det sammankopplade datasetet (blå streckad linje)



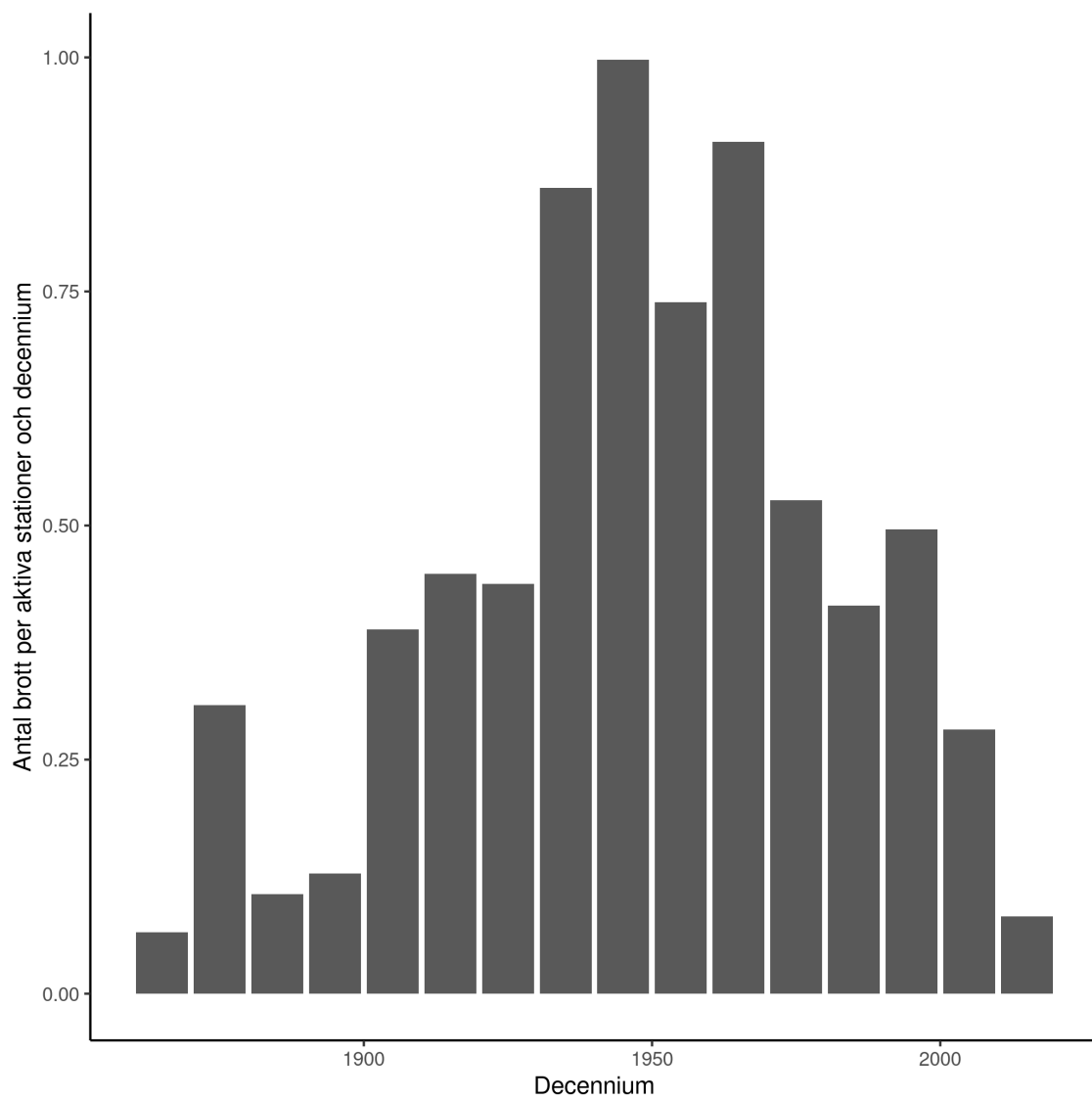
Figur 7: Histogram över tidsseriernas längd i det ursprungliga datasetet (gult och rött), där de röda staplarna beskriver den del av datasetet som kan användas i HOMER samt den sammankopplade datasetet (blå)



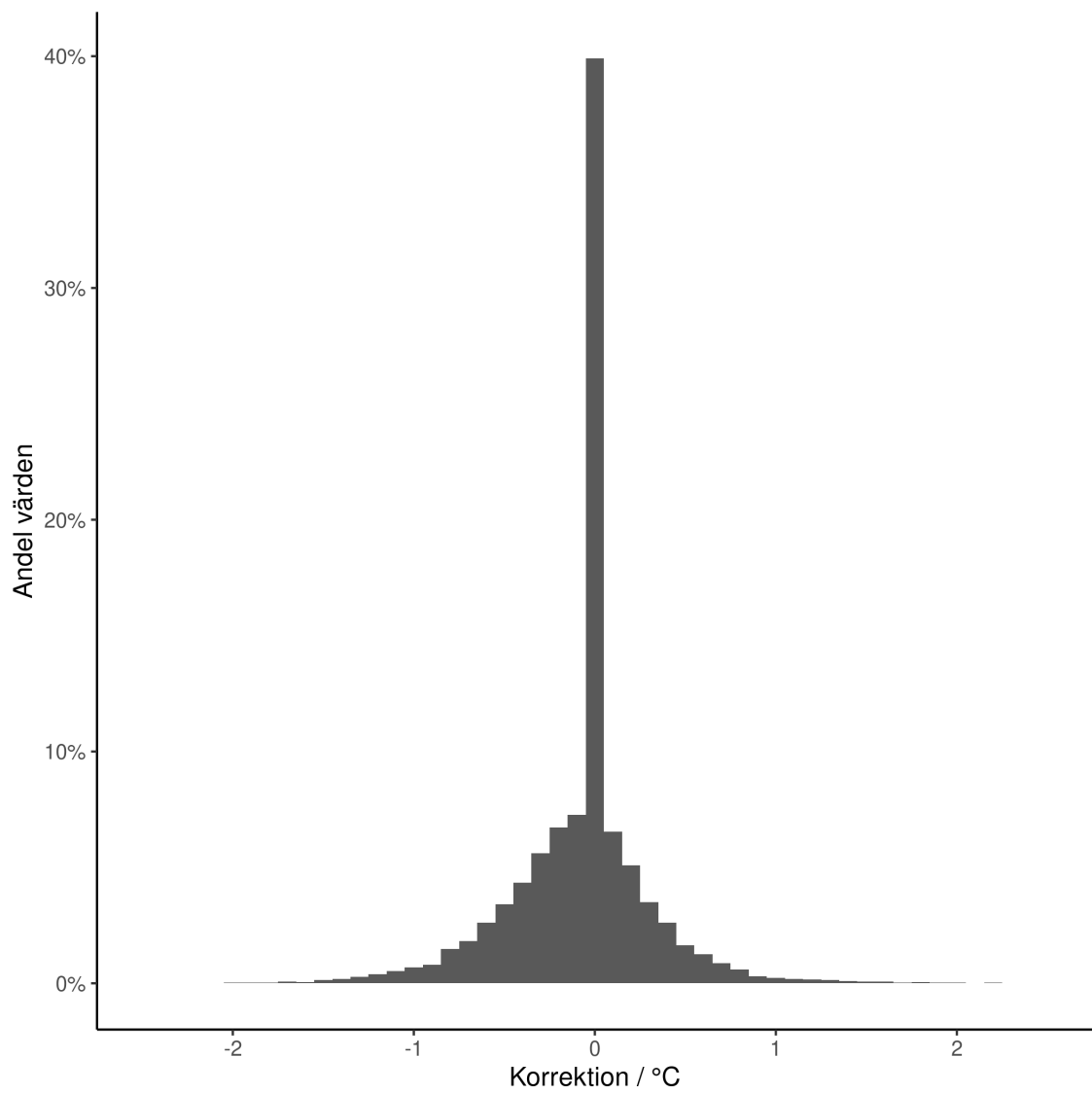
Figur 8: Histogram över antal homogenitetsbrott per tidsserie



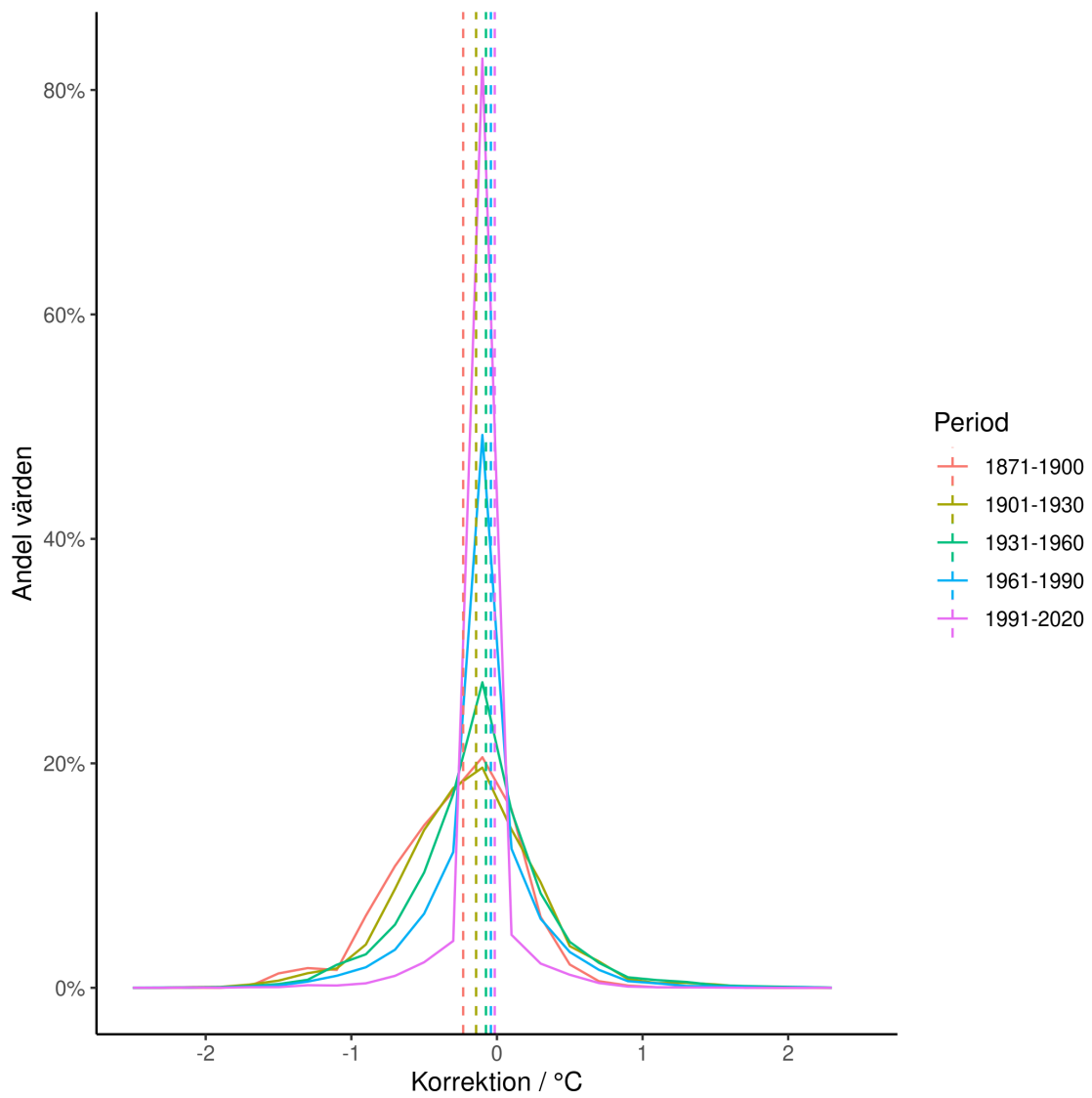
Figur 9: Histogram över den genomsnittliga tiden mellan två brott per tidsserie



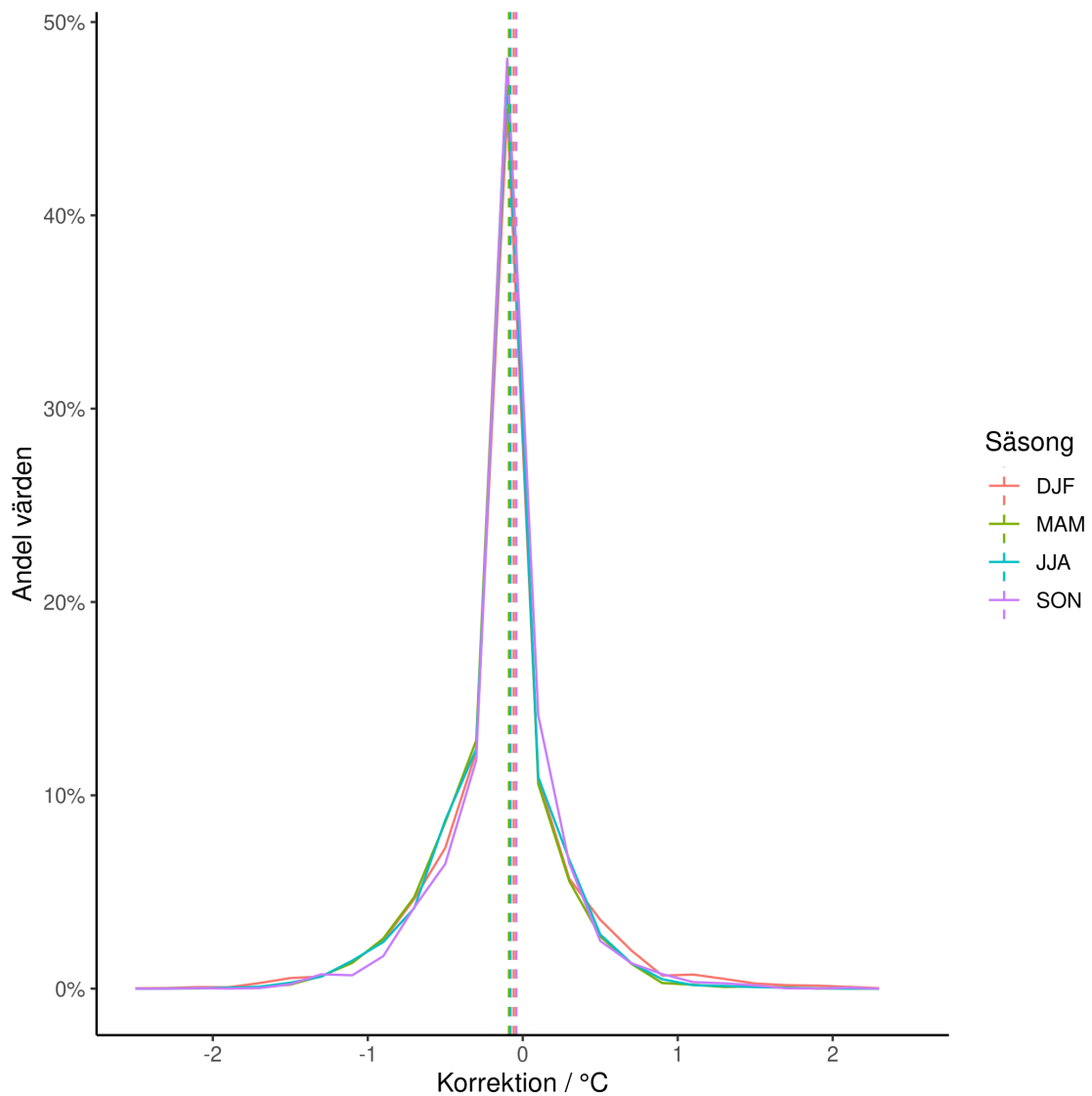
Figur 10:Antalet homogenitetsbrott per aktiva stationer och decennium för varje decennium 1860–1869 till 2010–2019



Figur 11: Histogram över korrektion



Figur 12: Fördelningen av korrektioner (heldragen linje) och dess genomsnitt (streckade linjer) för olika normalperioder: 1871–1900 (röd), 1901–1930 (gul), 1931–1960 (grön), 1960–1991 (blå) samt 1991–2020 (lila)



Figur 13: Fördelningen av korrektioner (heldragen linje) och dess genomsnitt (streckade linjer) för olika säsonger: december-januari-februari (röd), mars-april-maj (gul), juni-juli-augusti (blå) samt september-oktober-november (lila)

Tabell 1: Antalet tidsserier, data och saknade data det ursprungliga datasetet, alla tidsserier över 15 år samt det sammankopplade datasetet

Dataset	Antal tidsserier	Antal data	Antal saknade data
Ursprungligt	961	417 972	1 450 212
Minst 15 års data	686	394 161	939 423
Sammankopplat	456	395 674	490 790

SMHIs publiceringar

SMHI publicerar sju rapportserier. Tre av dessa, R-serierna är avsedda för internationella läsare och skrivs oftast på Engelska. I de övriga serierna används oftast Svenska men även Engelska.

Seriernas namn	Publiceras sedan
RMK (Report Meteorology and Climatology)	1974
RH (Report Hydrology)	1990
RO (Report Oceanography)	1986
METEOROLOGI	1985
HYDROLOGI	1985
OCEANOGRAFI	1985
KLIMATOLOGI	2009

I serien KLIMATOLOGI har tidigare utgivits:

1. Lotta Andersson, Julie Wilk, Phil Graham, Michele Warburton (University KwaZulu Natal) (2009)
Local Assessment of Vulnerability to Climate Change Impacts on Water Resources in the Upper Thukela River Basin, South Africa – Recommendations for Adaptation
2. Gunn Persson, Markku Rummukainen (2010)
Klimatförändringarnas effekter på svenskt miljömålsarbete
3. Jonas Olsson, Joel Dahné, Jonas German, Bo Westergren, Mathias von Scherling, Lena Kjellson, Fredrik Ohls, Alf Olsson (2010)
En studie av framtida flödesbelastning på Stockholms huvudavloppssystem
4. Markku Rummukainen, Daniel J. A. Johansson, Christian Azar, Joakim Langner, Ralf Doescher, Henrik Smith (2011)
Uppdatering av den vetenskapliga grunden för klimatarbetet. En översyn av natur-vetenskapliga aspekter
5. Sten Bergström (2012)
Framtidens havsnivåer i ett hundraårsperspektiv – kunskapssammanställning 2012
6. Jonas Olsson och Kean Foster (2013)
Extrem korttidsnederbörd i klimatprojektioner för Sverige
7. FNs klimatpanel – Sammanfattning för beslutsfattare. Effekter, anpassning och sårbarhet. Bidrag från arbetsgrupp 2 (WG 2) till den femte utvärderingen (AR 5) från Intergovernmental Panel on Climate Change, IPCC (2014)
8. Att begränsa klimatförändringar. FNs klimatpanel – Sammanfattning för beslutsfattare. Bidrag från arbetsgrupp 3 (WG 3) till den femte utvärderingen (AR 5) från Intergovernmental Panel on Climate Change (2015)
9. Erik Kjellström SMHI. Reino Abrahamsson, Pelle Boberg. Eva Jernbäcker Naturvårdsverket. Marie Karlberg, Julien Morel
Energimyndigheten och Åsa Sjöström SMHI (2014)
Uppdatering av det klimatvetenskapliga kunskapsläget
10. Risker och konsekvenser för samhället av förändrat klimat – en kunskapsöversikt (2014)
11. Gunn Persson (2015)
Vägledning för användande av klimatscenarier
12. Lotta Andersson, Anna Bohman, Lisa van Well, Anna Jonsson, Gunn Persson och Johanna Farelus (2015) Underlag till kontrollstation 2015 för anpassning till ett förändrat klimat
13. Gunn Persson (2015)
Sveriges klimat 1860-2014. Underlag till Dricksvattenutredningen

14. Anna Eklund, Jenny Axén Mårtensson, Sten Bergström, Emil Björck, Joel Dahné, Lena Lindström, Daniel Nordborg, Jonas Olsson, Lennart Simonsson och Elin Sjökvist (2015) Sveriges framtida klimat. Underlag till Dricksvattensutredningen
15. Elin Sjökvist, Jenny Axén Mårtensson, Joel Dahné, Nina Köplin, Emil Björck, Linda Nylén, Gitte Berglöv, Johanna Tengdelius Brunell, Daniel Nordborg, Kristoffer Hallberg, Johan Södling, Steve Berggreen-Clausen (2015) Klimatscenarier för Sverige - Bearbetning av RCP-scenarier för meteorologiska och hydrologiska effektstudier
16. Elin Sjökvist, Gunn Persson, Jenny Axén Mårtensson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Linda Nylén, Alexandra Ohlsson och Håkan Persson (2015) Framtidsklimat i Dalarnas län – enligt RCP-scenarier
17. Linda Nylén, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Värmlands län – enligt RCP-scenarier
18. Gunn Persson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Örebro län – enligt RCP-scenarier
19. Alexandra Ohlsson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Västmanlands län – enligt RCP-scenarier
20. Elin Sjökvist, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson (2015) Framtidsklimat i Uppsala län – enligt RCP-scenarier
21. Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Stockholms län – enligt RCP-scenarier
22. Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Södermanlands län – enligt RCP-scenarier
23. Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Östergötlands län – enligt RCP-scenarier
24. Gitte Berglöv, Magnus Asp, Steve Berggreen-Clausen, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Västra Götalands län – enligt RCP-scenarier
25. Alexandra Ohlsson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Jönköpings län – enligt RCP-scenarier
26. Gunn Persson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015) Framtidsklimat i Kalmar län – enligt RCP-scenarier

27. Gitte Berglöv, Magnus Asp, Steve Berggreen-Clausen, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Kronobergs län – Enligt RCP-scenarier
28. Gunn Persson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Hallands län – enligt RCP-scenarier
29. Alexandra Ohlsson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Skåne län – enligt RCP-scenarier
30. Alexandra Ohlsson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Anna Johnell, Jenny Axén Mårtensson, Linda Nylén, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Blekinge län – enligt RCP-scenarier.
31. Gunn Persson, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Gotlands län – enligt RCP-scenarier
32. Gitte Berglöv, Magnus Asp, Steve Berggreen-Clausen, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Norrbottens län – enligt RCP-scenarier
33. Gitte Berglöv, Magnus Asp, Steve Berggreen-Clausen, Emil Björck, Jenny Axén Mårtensson, Linda Nylén, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Västerbottens län – enligt RCP-scenarier
34. Linda Nylén, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Jämtlands län – enligt RCP-scenarier
35. Linda Nylén, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Västernorrlands län – enligt RCP-scenarier
36. Linda Nylén, Magnus Asp, Steve Berggreen-Clausen, Gitte Berglöv, Emil Björck, Jenny Axén Mårtensson, Alexandra Ohlsson, Håkan Persson, Elin Sjökvist (2015)
Framtidsklimat i Gävleborgs län – enligt RCP-scenarier
37. Jonas Olsson, Weine Josefsson (red.) (2015) Skyfallsuppdraget - ett regeringsuppdrag till SMHI
38. Gunn Persson, Linda Nylén, Steve Berggreen-Clausen, Peter Berg, David Rayner och Elin Sjökvist (2015)
Från utsläppsscenarioer till lokal nederbörd och översvämningsrisker
39. Anna Eklund, Jenny Axén Mårtensson, Sten Bergström och Elin Sjökvist (2015)
Framtidens vattentillgång i Mälaren, Göta älv, Bolmen, Vombsjön och Gavleån. Underlag till Dricksvattenutredningen
40. Anna Bohman (Centrum för klimatpolitisk forskning, CSPR) vid Linköpings universitet, Lotta Andersson, SMHI och CSPR, Linköpings universitet samt Åsa Sjöström, SMHI. (2016)
Förslag till en metod för uppföljning av det nationella klimatanpassningsarbetet. Redovisning av ett regeringsuppdrag December 2016

41. (2017)
Karttjänst för framtida medelvattenstånd
längs Sveriges kust
42. Anna Eklund, Linda Tofeldt, Johanna
Tengdelius-Brunell, Anna Johnell, Jonas
German, Elin Sjökvist, Maria Rasmusson,
Elinor Andersson (2017)
Vattennivåer, tappningar, vattentemperaturer
och is i Vättern
Beräkningar för dagens och framtidens
klimatförhållanden
43. Anna Eklund, Anna Johnell, Linda Tofeldt,
Johanna Tengdelius-Brunell, Maria
Andersson, Cajsa-Lisa Ivarsson, Jonas
German, Elin Sjökvist och Elinor Andersson
(2017)
Vattennivåer, tappningar, vattentemperaturer
och is i Hjälmaren Beräkningar för dagens
och framtidens klimatförhållanden
44. Anna Eklund, Linda Tofeldt, Anna Johnell,
Maria Andersson, Johanna Tengdelius-
Brunell, Jonas German, Elin Sjökvist, Maria
Rasmusson, Ulrika Harbman, Elinor
Andersson (2017)
Vattennivåer, tappningar, vattentemperatur
och is i Vänern Beräkningar för dagens och
framtidens klimatförhållanden
45. Sofie Schöld, Cajsa-Lisa Ivarsson, Signild
Nerheim och Johan Södling (2017)
Beräkning av högsta vattenstånd längs
Sveriges kust
46. Katarina Stensen, Johanna Tengdelius-
Brunell, Elin Sjökvist, Elinor Andersson,
Anna Eklund (2017)
Vattentemperaturer och is i Mälaren
Beräkningar för dagens och framtidens
klimatförhållanden
47. Jonas Olsson, Peter Berg, Lennart Wern,
Johan Södling, Lennart Simonsson,
Wei Yang, Anna Eronn (2017)
Extremregn i nuvarande och framtida klimat
– Analyser av observationer och
framtidsscenarier.
48. Signild Nerheim, Sofie Schöld, Gunn
Persson och Åsa Sjöström (2017)
Framtida havsnivåer i Sverige
49. Anna Eklund, Katarina Stensen, Ghasem
Alavi, Karin Jacobsson, Diala Abdoush
(2018)
Sveriges stora sjöar idag och i framtiden.
Klimatets påverkan på Vänern, Vättern,
Mälaren och Hjälmaren.
Kunskapssammanställning januari 2018
50. Gunn Persson, Christina Wikberger,
Jorge Amorim (2018)
Klimatanpassa städer med grönska
51. Katarina Losjö, Lennart Wern, Johan
Södling (2019)
Uppföljning av riktlinjer för bestämning
av dimensionerande flöden
52. Sjökvist, Elin (2019)
Sommaren 2018 – en glimt av framtiden?
53. Översättning av Summary for
Policymakers (2019)
FN:s klimatpanel – Sammanfattning för
beslutsfattare
Global uppvärmning på 1,5°C
54. Karin Hjerpe, Therése Sjöberg, Karin
Lundgren Kownacki, Lotta Andersson,
Åsa Sjöström (2020)
Myndigheters arbete med
klimatanpassning 2019
55. Therése Sjöberg, Karin Hjerpe, Karin
Lundgren Kownacki, Lotta Andersson
(2020)
Kommunernas arbete med
klimatanpassning 2019 - Analys av
statusrapportering till SMHI
56. Klimatförändringar och biologisk
mångfald – Slutsatser från IPCC och
IPBES i ett svenskt perspektiv (2020)
57. FN:s klimatpanel IPCC –
Sammanfattning för beslutsfattare (2020)
Specialrapport om Klimatförändringar
och marken
58. FN:s klimatpanel IPCC –
Sammanfattning för beslutsfattare (2020)
Specialrapport om Havet och kryosfären
i ett förändrat klimat

Denna sida är avsiktligt blank



Sveriges meteorologiska och hydrologiska institut
601 76 NORRKÖPING
Tel 011-495 80 00 Fax 011-495 80 01

ISSN 1654-2258

