

Het raadsel van de verdwenen hittegolven



Hoe het KNMI historische hittegolven
uit de boeken schrapte

Colofon

Auteurs

Frans Dijkstra
Jan Ruis
Rob de Vos
Marcel Crok

Uitgever

Wij zijn een ad hoc gevormde groep onderzoekers en wij brengen dit rapport op persoonlijke titel uit. Onze motivatie was om de onderste steen boven te krijgen wat betreft de temperatuurcorrecties die het KNMI had doorgevoerd in De Bilt.

Website:

Het rapport kan gedownload worden op

www.klimaatgek.nl

www.destaatvanhet-klimaat.nl

Financiering

Dit rapport is tot stand gekomen op eigen initiatief van de onderzoekers. In de zomer van 2018 hebben we via onze eigen kanalen (website en social media) een crowdfunding pagina opgezet. Dat leverde ongeveer 6500 euro op gedoneerd door 140 mensen. In een laat stadium hebben we ook enige funding ontvangen via de pas opgerichte Stichting Climate Intelligence.

Donaties

De kosten van het onderzoek en het schrijven van het rapport zijn aanzienlijk geweest. Donaties worden dan ook alsnog zeer op prijs gesteld.

Doneren kan, onder vermelding van 'Rapport Verdwenen Hittegolven', als storting op de rekening van:

M. Crok

IBAN NL66INGB0000520595

Contact

Rob de Vos: rob@klimaatgek.nl

Frans Dijkstra: fydijkstra1946@kpnmail.nl

Jan Ruis: janruis30@gmail.com

Marcel Crok: marcel.crok@gmail.com

Telefoon

Marcel Crok: (06) 16236275

Stichting Climate Intelligence

De Stichting Climate Intelligence heeft als doel het genereren van kennis over en inzicht in omvang, aard, oorzaak en gevolgen van klimaatverandering en het klimaatbeleid dat hierop betrekking heeft. Hiertoe probeert de stichting helder en transparant te communiceren naar het brede publiek wat er aan feiten beschikbaar is over klimaatverandering en klimaatbeleid en ook waar feiten overgaan in veronderstellingen, aannames en inschattingen.

Tevens voert en stimuleert de stichting een publiek debat hierover en financiert zij onderzoeksjournalistiek werk op dit terrein.



Voorwoord

Gedurende de lange droge zomer van 2018 werden Nederlanders in de media geconfronteerd met berichten over de forse toename van het aantal hittegolven in ons land sinds 1901.

Na een ‘fact check’ bleek dat die toename niet zozeer het gevolg was van klimaatverandering maar van een ‘correctie’ die het KNMI kort daarvoor had toegepast. De gemeten temperaturen tussen 1901 en 1951 van het station De Bilt waren namelijk per 1 januari 2016 fors bijgesteld. Daardoor verdwenen in één klap veel ‘oude’ hittegolven uit de statistieken.

Die correctie was uitkomst van een zogenaamde homogenisatie die het KNMI heeft toegepast op de 5 hoofdstations in ons land. Vanwege het ontbreken van parallelle meetdata in De Bilt heeft het KNMI de temperatuurreeks van Eelde (Groningen) als referentiestation voor de homogenisatie van De Bilt gebruikt.

Het KNMI heeft over de homogenisatie een technisch rapport uitgebracht (Brandsma 2016a). Het voor u liggende rapport beschrijft onze poging tot reconstructie van de homogenisatie van De Bilt zoals die in het KNMI-rapport is beschreven. Bovendien wordt uitgebreid aandacht besteed aan de vraag of de homogenisatie van De Bilt wel nodig was en of de keuze van het referentiestation Eelde de juiste is geweest.

We danken alle mensen die ons tot steun zijn geweest bij het maken van het rapport!

Maart 2019,

Frans Dijkstra
Jan Ruis
Marcel Crok
Rob de Vos

Inhoudsopgave

	Samenvatting	6
H1	Inleiding	10
H2	Waarom homogeniseren?	16
H3	Parallelmetingen op het KNMI-terrein	20
	Hutwisseling	20
	Locatiewisselingen in 1950 en 1951	22
	Conclusies	26
	Homogeniseren voor dummies	27
H4	De reconstructie van de homogenisatie van De Bilt	28
	Methoden	28
	Effect van de homogenisatie op de tropische, zomerse dagen en hittegolven en op ijsdagen	33
	Daggemiddelde temperatuur en dagelijkse minimumtemperaturen	34
	Enige klimatologische constateringn	35
	Conclusies	36
H5	Reconstructie van de homogenisatie op basis van de aanvullende informatie van het KNMI	37
	Hoor en wederhoor	37
	Driemaands gemiddelden	38
	Vergelijkingsperioden 56 maanden	39
	Maximale jaartemperatuur	40
	Conclusies	42
H6	De temperatuursprong en de correctie van het KNMI	43
	Duitse stations als referentie van de homogenisatie van De Bilt	44
	Methode	45
	Een nieuwe referentieset voor de homogenisatie van De Bilt	46
	Vergelijking Tx De Bilt met de referentieset	50
	Vergelijking Tx-zomer De Bilt met de referentieset	51
	Evaluatie homogenisatie De Bilt	52
	Conclusies	55
H7	Effect van de keuze vergelijkingsstation op de homogenisatie	56
	Werkwijze	57
	Resultaten	57
	Gemiddelde correctie Tx	58
	Effect van andere Nederlandse en Duitse stations	59
	Discussie	62
	Conclusies	63
	Literatuur	64
	Appendix	65
	Bijlage 1: Samenstelling Duits-Nederlands ensemble	65
	Bijlage 2: Detailgrafieken H4 en H5	69
	Bijlage 3: Vragen aan en antwoorden van het KNMI	79
	Bijlage 4: Smoothing for dummies	86

Samenvatting

1. **Het KNMI corrigeerde in 2016 haar oude temperatuurmetingen van vóór september 1951. Het deed dat vanwege een verandering van de meethut in 1950 en een verplaatsing van het meetstation in 1951. Op de jaarlijkse gemiddelde temperatuur hadden de correcties nauwelijks invloed. Warme tot zeer warme dagen werden echter drastisch naar beneden bijgesteld, tot wel 1,9 °C bij dagen boven de 28 °C. Als gevolg daarvan nam het aantal tropische dagen in de periode 1901-1951 in De Bilt af van 164 tot 76 en het aantal zomerse dagen van 931 tot 629. Van de 23 hittegolven die oorspronkelijk in de boeken stonden resteren er na de correcties (in jargon: de homogenisatie) nog maar 7. Als gevolg daarvan kon het KNMI in de warme zomer van 2018 claimen dat hittegolven nu veel vaker voorkomen dan een eeuw geleden.**
2. **Dit rapport beschrijft onze pogingen – een viertal onafhankelijke onderzoekers - om de homogenisatie van de temperatuurmetingen in De Bilt te reproduceren. Tevens is gekeken naar de noodzaak om überhaupt te homogeniseren.** De KNMI-correcties zijn tot stand gekomen na vergelijking van De Bilt met station Eelde (nabij Groningen). We hebben eveneens onderzocht tot welke correcties vergelijking met andere stations in Nederland en/of Duitsland zouden leiden.
3. **Hittegolven worden in Nederland alleen bepaald door de metingen in De Bilt. Er is sprake van een hittegolf als het op minimaal vijf dagen 25 °C of warmer is, waarvan op minimaal drie dagen gelijk of warmer dan 30 °C. In de ongecorrigeerde metingen vinden er meer hittegolven plaats in de periode 1901 t/m augustus 1951 (23 hittegolven) dan in de periode september 1951 t/m 2018 (19 hittegolven).** Hoewel die eerste periode korter duurt dan de tweede (51 tegen 67 jaar). Na de homogenisatie is die verhouding volledig omgedraaid: 7 hittegolven tegen 19 hittegolven. Brandsma, de onderzoeker van het KNMI die de homogenisatie heeft uitgevoerd schreef in vakblad Meteorologica in 2016: “Bekende problemen zoals de neerwaartse sprong in Tx [de maximumtemperatuur, red.] rond 1950 in De Bilt en het relatief hoge aantal hittegolven in de eerste helft van de 20ste (...) zijn hiermee verleden tijd.” Blijkbaar werd het hoge aantal hittegolven vóór 1950 als ‘een probleem’ gezien.
4. **Er is een statistische reden die de homogenisatie in De Bilt verdacht maakt. Het percentage tropische dagen in de periode vóór de wijziging (1906-1951) ten opzichte van erna (1951-2016) zakt in De Bilt veel verder dan bij de overige vier hoofdstations in Nederland (De Kooy, Beek, Eelde en Vlissingen.** Na homogenisatie bedraagt dat percentage bij de overige vier hoofdstations tussen de 35 en 39% terwijl het in De Bilt zakt naar 26%. Deze verschillen zijn niet genoemd door het KNMI.
5. **Normaal gesproken voeren onderzoekers bij de overgang naar een andere meetlocatie gedurende een periode parallelmetingen uit op zowel de oude als de nieuwe meetlocatie. Dat is bij de herlocatie op 27 augustus in De Bilt niet gebeurd. Historische documenten tonen aan dat er vanaf 1942 metingen zijn uitgevoerd met de Stevensonhut die vlakbij de Pagodehut was opgesteld. Het KNMI heeft merkwaardig genoeg geen gebruik gemaakt van die gegevens betreffende de overgang van de oude Pagodehut naar de nieuwe Stevensonhut in mei 1950.** Waarom die gegevens niet gebruikt zijn bij de homogenisatie is onduidelijk. Een rapport van Kramer et al uit 1954 suggereert dat het gemiddelde temperatuurverschil tussen de Pagodehut en de Stevensonhut gedurende een periode van negen maanden slechts 0,005°C bedroeg, met uiterste afwijkingen van +0,4°C tot -0,2°C. Wij hebben de parallelmetingen opgevraagd bij het KNMI maar niet gekregen. Gezien de claims die gedaan worden over het voorkomen van hittegolven na de homogenisatie is het van belang dat deze metingen alsnog zo spoedig mogelijk worden vrijgegeven.

6. **Het KNMI heeft (ook) parallelmetingen uitgevoerd van mei 2003 tot juni 2005. Die metingen vonden onder andere plaats op de oorspronkelijke meetlocatie (waar gemeten werd tot 27 augustus 1951) en de locatie waarheen de Stevensonhut in augustus 1951 werd verplaatst. De maximumtemperaturen in de zomer in de periode 2003-2005 op deze twee locaties verschillen niet significant van elkaar.** Dit suggereert dat de verplaatsing in 1951 geen groot effect gehad hoeft te hebben, terwijl het KNMI deze verplaatsing juist mede aangrijpt als noodzaak voor de homogenisatie. De locatiewisseling lijkt de grote correcties van de maximumtemperaturen tot 1,9°C dan ook niet te rechtvaardigen.
7. **Vanwege het vermeende gebrek aan parallelmetingen heeft het KNMI besloten om de dagwaarden van De Bilt te homogeniseren aan de hand van station Eelde dat 150 km ten noordoosten van De Bilt ligt. Daarbij is gebruik gemaakt van een statistische techniek, de *percentile matching method*. De keuze voor Eelde is saillant. Eerder, bij het tot stand komen van de Centrale Nederlandse Temperatuur (CNT), werd namelijk geconcludeerd dat Eelde daarvoor niet geschikt was.** Het KNMI-rapport waarin de homogenisatie wordt beschreven onderbouwt niet waarom alleen station Eelde geschikt was. Een nadeel van Eelde is dat metingen daar pas in 1946 begonnen. Daardoor konden slechts vier jaar aan data vóór de breuk (1946-1949) gebruikt worden en die werden vergeleken met vier jaar na de breuk (1952-1955). Uit ons onderzoek blijkt dat de methode gevoelig is voor de lengte van de periode.
8. **De methode verdeelt voor iedere maand de metingen onder in twintig percentielen. Met ‘slechts’ vier jaar aan data komt dat neer op ongeveer zes waarden per percentiel. Dat is weinig en leidt tot veel ruis in de resultaten. Het KNMI heeft daarom aansluitend middeling over verschillende maanden en percentielen toegepast om deze ruis te onderdrukken.** Uit ons onderzoek blijkt dat de uiteindelijke temperatuurcorrectie sterk afhankelijk is van de gekozen middeling. Het KNMI heeft echter niet laten zien hoe gevoelig het eindresultaat voor deze arbitraire keuzes is.
9. **Het lukte ons niet om op basis van de procedure zoals beschreven in het technische rapport van het KNMI de homogenisatie van het KNMI te reproduceren. Met name voor de maximumtemperaturen vinden wij sterk afwijkende resultaten.** Voor de jaargemiddelde maximumtemperaturen wijken de uitkomsten van het KNMI 0,05°C tot 0,12°C af van onze reconstructie en de fluctuaties per jaar zijn in de KNMI-homogenisatie veel groter dan in onze reconstructie. Met name jaren met warme zomers worden door het KNMI sterk naar beneden bijgesteld. Bijvoorbeeld voor 1947 – een jaar met oorspronkelijk vier hittegolven – komt het KNMI tot veel lagere maximumtemperaturen dan wij in onze reconstructie.
10. **Het KNMI heeft door de homogenisatie het aantal tropische dagen en hittegolven vóór 1951 sterk verminderd: van 164 tropische dagen naar 76, en van 23 naar 7 hittegolven. Bij onze reconstructie van de homogenisatie blijven er vóór 1951 echter 122 tropische dagen over en 13 hittegolven.** Dit betekent niet dat wij dit aantal als ‘het juiste’ beschouwen. Onze conclusie is wel dat de homogenisatie van het KNMI heeft geresulteerd in een sterke overcorrectie van de temperatuur op warme dagen.
11. **Na contact te hebben gezocht met het KNMI om de verschillen tussen hun en onze resultaten te kunnen verklaren bleek dat het KNMI acht maanden extra had gebruikt (de eerste acht maanden van 1950) om de beginperiode zo lang mogelijk te laten duren. En er is een extra middeling toegepast van de percentielen over drie aansluitende maanden. Aanvullend onderzoek maakte duidelijk dat de extra middeling weinig effect had maar gebruik van de extra acht maanden wel. Onze resultaten kwamen nu dichterbij die van het KNMI. Maar de grote verschillen bij de maximumtemperaturen blijven aanwezig. Wij houden na de homogenisatie aanzienlijk meer tropische dagen en hittegolven over. Wij constateren opnieuw dat de homogenisatie van het KNMI heeft geresulteerd in een sterke overcorrectie van de temperatuur op warme dagen. Het gebruik van de extra acht maanden**

is volgens ons aanvechtbaar omdat deze maanden (jan-aug) daardoor zwaarder tellen dan de overige maanden.

12. **Het feit dat schijnbaar arbitraire keuzen in de methodiek (percentielberekening, middeling, lengte van de vergelijkingsperiode) veel invloed hebben op de uitkomsten van de homogenisatie, laat zien dat deze methode niet geschikt is om dagwaarden te corrigeren met de pretentie van een nauwkeurigheid van 0,1 graad.** Dat is statistisch onverantwoord. Als je al homogeniseert met deze methode, dan zou je gehomogeniseerde dagtemperaturen tenminste moeten afronden op hele graden.
13. **Het klopt dat temperatuurmetingen in De Bilt een behoorlijke daling laten zien in de periode 1949 tot 1956. Het is verleidelijk om die afkoeling grotendeels toe te schrijven aan de verandering van de meethut en de verplaatsing. Uit ons onderzoek blijkt echter dat die afkoeling in alle stations in Nederland en ook Duitsland optreedt. Die afkoeling is dus vooral klimatologisch van aard.** Wij hebben door vergelijking met een ensemble van Nederlandse en Duitse stations wel geconcludeerd dat De Bilt inderdaad relatief iets warmer was vóór 1950 en dat enige correctie in de jaren vóór 1950 verdedigd kan worden. Wij concluderen dat de door het KNMI toegepaste correcties op de maximumtemperaturen gemiddeld van 1931-1949 wel kloppen maar in de jaren daarvóór tot overcorrectie leiden. Voor jaren met recordwarme zomers (zoals 1947) past het KNMI een overcorrectie toe van naar schatting 50%. Voor de maanden mei-september is de geschatte overcorrectie 210% en voor de drie warmste zomermaanden zelfs ruim 260% op basis van vergelijking met het ensemble.
14. **We hebben tenslotte onderzocht hoe gevoelig de homogenisatie van De Bilt is voor de keuze van het vergelijkingsstation. Dat effect blijkt groot te zijn.** Met Beek als vergelijkingsstation vinden wij voor De Bilt minder grote correcties voor de maximumtemperaturen en meer tropische dagen. Met Aachen als vergelijkingsstation vinden we voor De Bilt geen correctie voor de maximumtemperaturen en veel meer tropische dagen. Met geen enkel representatief vergelijkingsstation noch het door ons geselecteerde Duits-Nederlandse ensemble vinden wij zo weinig tropische dagen in De Bilt als bij de door het KNMI gebruikte homogenisatie aan de hand van Eelde. We concluderen dat de *percentile matching method*, zoals door het KNMI toegepast, afhankelijk is van arbitraire keuzen en daarom niet geschikt is om de klimatologische geschiedenis ingrijpend aan te passen.
15. **Het moge duidelijk zijn dat wij de homogenisatie van De Bilt zoals die is uitgevoerd met behulp van station Eelde onverdedigbaar vinden. De uitkomst is veel te afhankelijk van arbitraire keuzes wat betreft de stations waarmee je vergelijkt, de gebruikte statistische methode, de lengte van de vergelijkingsperiode en de gebruikte middeling.** Het KNMI heeft in het technische rapport over de homogenisatie bovendien nagelaten om zelf deze gevoeligheden bloot te leggen. Tevens heeft het KNMI geen enkele moeite gedaan om te laten zien dat de afkoeling in de periode 1949-1956 grotendeels klimatologisch van aard kan zijn omdat die zich bij alle homogene reeksen in Nederland en Duitsland voordoet. Dit rapport heeft laten zien dat een kleine correctie wellicht verdedigbaar is maar de homogenisatie van het KNMI heeft geleid tot flinke overcorrecties, vooral van warme dagen. Het beste zou zijn om de homogenisatie ongedaan te maken en met een breder opgezet team (waaronder wetenschappers van buiten het KNMI) een nieuwe start te maken. Het KNMI zou zich voorlopig moeten onthouden van claims over een vermeende toegenomen trend in hittegolven in Nederland.

1 Inleiding

De Koninklijke Nederlandse Academie van Wetenschappen (KNAW) stelde op 15 januari 2018 voor dat het herhalen van experimenten heel gewoon moet worden (KNAW 2018). In veel vakgebieden staat onderzoek regelmatig ter discussie vanwege de toegepaste methodiek. Soms ook dikken onderzoekers hun conclusies aan om hun publicatie beter in het oog te laten springen. De beste remedie tegen foute conclusies uit zulk onderzoek is dat wetenschappers geregeld elkaars onderzoek herhalen.

We hebben het advies van de KNAW opgevolgd en in dit rapport geprobeerd de door het KNMI in 2016 uitgevoerde homogenisatie van de temperatuur op station De Bilt te reproduceren. Deze homogenisatie vond plaats wegens een verandering van de meethut in 1950 en een verplaatsing in 1951 waarvan men veronderstelde dat deze invloed zouden hebben op de meetwaarden. De temperatuurregistraties die vóór deze instrumentele veranderingen plaatsvonden zouden dan gecorrigeerd moeten worden.

Het betreft de gemiddelde temperatuur (T_g), minimumtemperatuur (T_n) en maximumtemperatuur (T_x) per etmaal over de periode van 1 januari 1901 tot 1 september 1951. De in deze periode geregistreerde temperaturen zijn door de homogenisatie in 2016 bijgesteld. In onze reproductie van die homogenisatie hebben we ons met name gericht op de dagelijkse maximumtemperaturen (T_x) in De Bilt omdat die het sterkst zijn bijgesteld, maar ook omdat de maximumtemperaturen in de media de meeste aandacht hebben getrokken.

Homogenisatie wordt vaak toegepast bij verplaatsing van meetapparatuur. Bij verplaatsing veranderen de omgevingsomstandigheden en daarmee de resultaten van de metingen. Brandsma, de auteur van het aan de homogenisatie ten grondslag liggende technische rapport (Brandsma 2016a), beschrijft de noodzaak van homogeniseren zo: *“Homogeneous time series of daily T_n , T_x and T_{mean} are indispensable for climate change and variability studies.”*



Figuur 1 Ligging 5 hoofdstations

De homogenisatie van De Bilt was onderdeel van een homogenisatie van alle 5 hoofdstations. Behalve De Bilt zijn dat Groningen/Eelde, Maastricht, Den Helder/De Kooy en Vlissingen/Souburg.

Op alle vijf hoofdstations was sprake van een herlocatie van de meetapparatuur. Groningen werd in 1951 vliegveld Eelde, zo'n 10 km ten Z van de stad. Den Helder werd in 1972 vliegveld De Kooy, gelegen ten ZO van Den Helder. Maastricht werd in 1951 vliegveld Beek, ruim 9 km ten NNO van Maastricht, en Vlissingen werd van 1947 tot 1958 vliegveld Souburg en daarna weer Vlissingen. Deze vier stations zijn gehomogeniseerd op basis van parallelle metingen. Bij verplaatsing heeft men gedurende enige tijd zowel op de oude als de nieuwe locatie gemeten.

De Bilt was een geval apart. Op 17 mei 1950 werd de grote Pagodehut vervangen door een Stevensonhut. Op 16 september werd die hut enkele tientallen meters westwaarts verplaatst vanwege bouwwerkzaamheden. Op 27 augustus 1951 werd de hut zo'n 300m naar het zuiden verplaatst, naar een meetveldje. In het technische rapport over de homogenisatie schrijft Brandsma: *"For the relocation in De Bilt in 1951 no parallel observations have been made because the relocation was unforeseen and after the relocation the old location was disturbed due to building activities in the neighborhood of the former screen. Therefore, data of Eelde (before the relocation and screen change and thereafter) has been used for the homogenization of the series."* Een aantal jaren vóór de hutwisseling in 1950 zijn parallelmetingen verricht tussen Pagodehut en een Stevensonhut. De uitkomsten daarvan zijn merkwaardig genoeg niet bij de homogenisatie gebruikt (Brandsma 2016a).

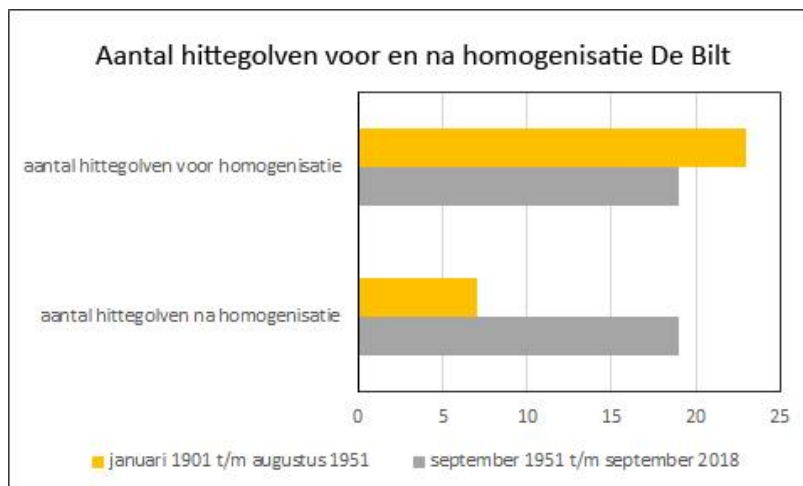
Voor de homogenisaties van de vijf stations maakt het KNMI gebruik van een statistische techniek genaamd *'percentile matching'*. Daarbij vergelijkt men de meetgegevens van de oude en de nieuwe locatie gedurende een overlappende periode. Vanwege het ontbreken van parallelle data bij De Bilt heeft het KNMI gekozen voor de data van station Eelde als referentie.

Wat is een hittegolf?

Van een hittegolf is sprake als in De Bilt in een periode van 5 dagen minstens 3 tropische dagen voorkomen en de overige dagen zomers zijn. Anders gezegd: vijf dagen lang moet de temperatuur in De Bilt boven de 25°C uitkomen, waarvan drie dagen boven de 30°C. Men spreekt van een 'landelijke hittegolf' als die zich in De Bilt voordoet. Indien op een andere plaats wordt voldaan aan de criteria kan men spreken over 'lokale hittegolven'. Deze worden niet officieel geregistreerd of geteld. Een hittegolf eindigt als op een dag de maximumtemperatuur van 25 °C niet meer bereikt wordt. Hittegolven kunnen dus langer duren dan 5 dagen.

Als na het einde van een hittegolf de maximumtemperatuur opnieuw boven de 25°C komt, begint de telling opnieuw. Een zomer kan daardoor meer dan één hittegolf tellen. De zomer van 2018 telde twee hittegolven die 6 en 9 dagen duurden. De zomer van 1947 telde volgens de oorspronkelijke metingen vier hittegolven van resp. 6, 6, 6 en 15 dagen.

De homogenisatie van De Bilt heeft geleid tot opvallend grote verschillen tussen de oude gemeten data en de nieuwe gehomogeniseerde data. Diverse media en internetsites berichtten al snel over deze opvallende grote correcties. De gevolgen van de homogenisatie werden nog beter zichtbaar in de warme zomer van het afgelopen jaar. Er werden in 2018 door het KNMI twee hittegolven geteld. Media berichtten dat hittegolven nu frequenter voor komen: van één keer in de twintig jaar een eeuw geleden tot elke twee tot drie jaar nu. De bron daarvan was het KNMI, dat zich baseerde op de gehomogeniseerde cijfers.



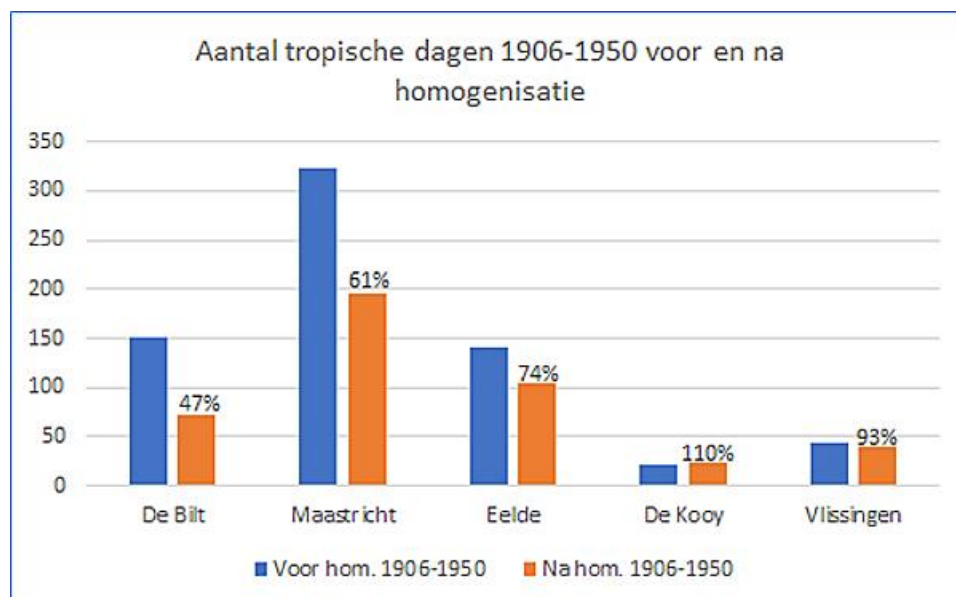
Figuur 2 Aantal hittegolven De Bilt voor en na homogenisatie

In figuur 2 is te zien wat het effect van de homogenisatie is op het aantal hittegolven van 1901 tot 1951. Vanaf 1 september 1951 zijn de temperaturen niet gehomogeniseerd.

Onderzoeker Brandsma van het KNMI die de homogenisatie heeft uitgevoerd schrijft:
"Bekende problemen zoals de neerwaartse sprong in Tx rond 1950 in De Bilt en het relatief hoge aantal hittegolven in de eerste helft van de 20ste eeuw (...) zijn hiermee verleden tijd."
 Meteorologica, 2016-4, pag. 4

Met "neerwaartse sprong" bedoelt Brandsma kennelijk een trendbreuk in het Tx-verloop als gevolg van de verplaatsing en verandering van meethut. Een van de auteurs van het voorliggende rapport, Frans Dijkstra, heeft al in 2017 kritiek geleverd op de homogenisatie van Brandsma.

Dijkstra (2017) schreef in Meteorologica: *"De homogenisatie van de temperatuurreeksen voor De Bilt leidt voor Tx tot onverwachte uitkomsten die niet in lijn zijn met de uitkomsten bij de andere stations en die onvoldoende verklaard kunnen worden uit de overgang van een pagodehut naar een Stevensonhut. De sterke vermindering van het aantal tropische dagen in de periode 1906-1950 in De Bilt en de veel lagere classificatie van de zomer van 1947 stemmen niet overeen met gehomogeniseerde gegevens van de andere stations. De temperatuursprong rond 1950 doet zich ook voor bij andere KNMI-stations en ook bij stations in Engeland en België."*

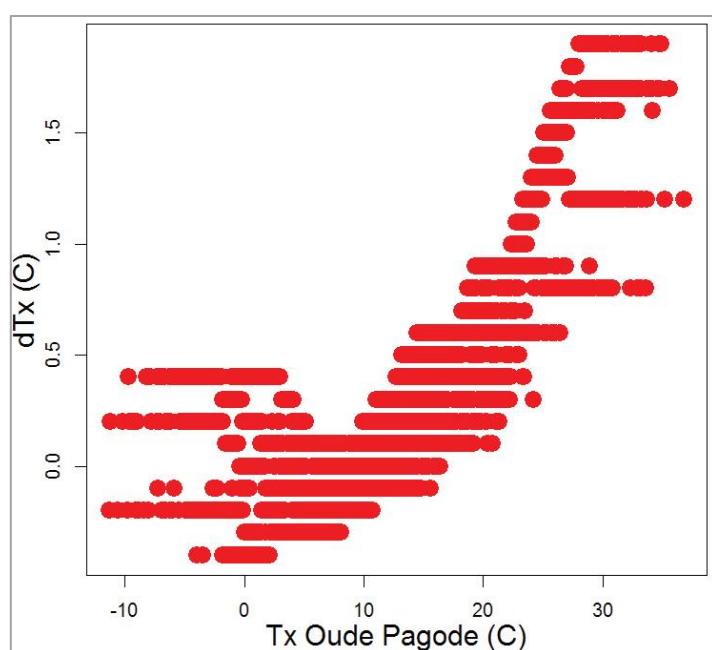


Figuur 3 Aantal tropische dagen op 5 hoofdstations voor en na homogenisatie

Ook in andere media is de homogenisatie omstreden (Rozendaal (2016,2017), Klimaatgek¹). Daarbij gaat het vaak om de meest opvallende gevolgen van de homogenisatie, namelijk de drastisch afname van zomerse en tropische dagen en hittegolven. Van de 931 zomerse dagen bleven er maar 629 over, en het aantal van 164 tropische dagen werd teruggebracht tot 76. Deze herschrijving van de meteorologische geschiedenis heeft tot gevolg dat er in deze periode door homogenisatie nog maar 7 hittegolven erkend worden, tegen 23 vóór de homogenisatie.

Afgezien van de vraag of de homogenisatie van De Bilt wel nodig was en of de gebruikte methode valide is, valt het op dat in De Bilt veel meer tropische dagen zijn geschrapt dan bij de andere KNMI-stations: in Maastricht, Eelde, De Kooy en Vlissingen worden na homogenisatie tussen 35,4 en 39,5% van de tropische dagen voor 1951 geteld, in De Bilt is dit slechts 26,4%. Het lijkt er sterk op dat De Bilt overgecorrigeerd is. In hoofdstuk 4 en 5 wordt daar aandacht aan besteed.

Oud-medewerker van het KNMI en emeritus hoogleraar Henk de Bruin (2016) schreef een kritisch artikel over de homogenisatie van De Bilt in Meteorologica. Uit dat artikel is onderstaande grafiek afkomstig die duidelijk de toegepaste correcties op de temperatuur weergeeft:



Figuur 4 Correctiewaarde dTx als functie van de gemeten Tx

Op de website Weerwoord heeft blogger 'Michiel'² specifiek gekeken hoe de homogenisatie uitpakt voor de maand september. In september zijn, net zoals de andere zomermaanden, alleen de warme dagen zwaar gecorrigeerd, maar de verdeling tussen correcties binnen september is zeer scheef. Figuur 5 toont de correctie per temperatuurinterval.

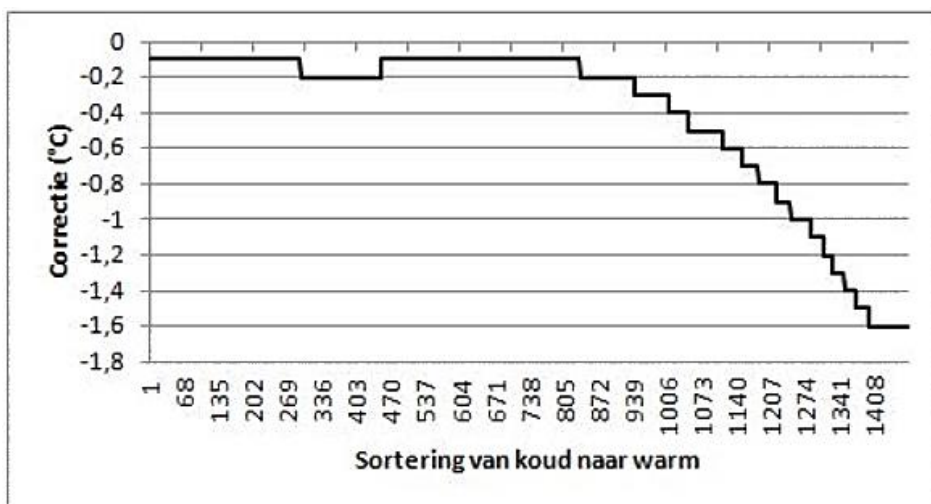
¹ <https://klimaatgek.nl/wordpress/2018/07/27/het-raadsel-van-de-verdwenen-hittegolven>

² <https://www.weerwoord.be/m/2227437>



Figuur 5 Correctie De Bilt per temperatuurinterval voor de maand september

In figuur 6 is de correctie weergegeven over alle septemberdagen in de gehomogeniseerde periode. Meer dan de helft (939) van alle septemberdagen (totaal 1470) heeft een correctie van 0,1 en soms 0,2 graden gehad. Hier zitten ook veel zonnige dagen tussen, zelfs volledig windstille dagen zoals 14 september 1925.



Figuur 6 Correctie De Bilt voor alle septemberdagen in de gehomogeniseerde periode

Ook is specifiek een aantal dagen met elkaar vergeleken om te laten zien hoe door het negeren van de invloed van de zon en de wind de toegepaste homogenisatietechniek inconsistent wordt:

Het weer op vrijdag 3 september 1909 te De Bilt:

Oude temperatuur 19,2 °C, nieuwe temperatuur 19,1 °C

Duur zonschijn: 10,7 uur

Gemiddelde windsnelheid 4,6m/s

Het weer op woensdag 29 oktober 1913 te De Bilt:

Oude temperatuur 19,1 °C, nieuwe temperatuur 18,3 °C

Duur zonschijn: 2,4 uur

Gemiddelde windsnelheid 6,7m/s

Michiel constateert: "De homogenisatie doet hier exact het tegenovergestelde van wat het zou moeten doen. Beide dagen hebben bijna dezelfde temperatuur, maar 29 oktober 1913, een dag met

een veel lagere zonnestand (vergelijkbaar met half februari), meer wind, veel minder zonneschijnduur en veel minder absolute zonneschijnduur, heeft toch een correctie gekregen die 8x hoger is dan de septemberdag die juist in aanmerking had moeten komen voor een dergelijke correctie. “

In het technische rapport over de homogenisatie schrijft Brandsma: *“For Tx large positive corrections (up to 1.9°C) are found in summer for the largest percentiles. This is mainly a result (of) a combination of the pagoda which was open at the bottom and affected by reflected sunlight and the enclosed location till 1951.”* Het is daarom opvallend dat de genoemde ‘dominante’ fysische factoren, namelijk zonlicht en ventilatie, niet gebruikt werden bij de homogenisatie: de homogenisatie is een louter statistische exercitie geworden. Het lijkt erop dat dat geresulteerd heeft in een overcorrectie van met name de hogere temperaturen.

Brandsma lijkt zich daarvoor te excuseren als hij schrijft: *“... het is onrealistisch om van een homogenisatiemethode te verwachten dat ze de allerhoogste extremen (zoals het aantal tropische dagen per jaar) perfect corrigeert.”*
Meteorologica, 2017-4, pag. 16

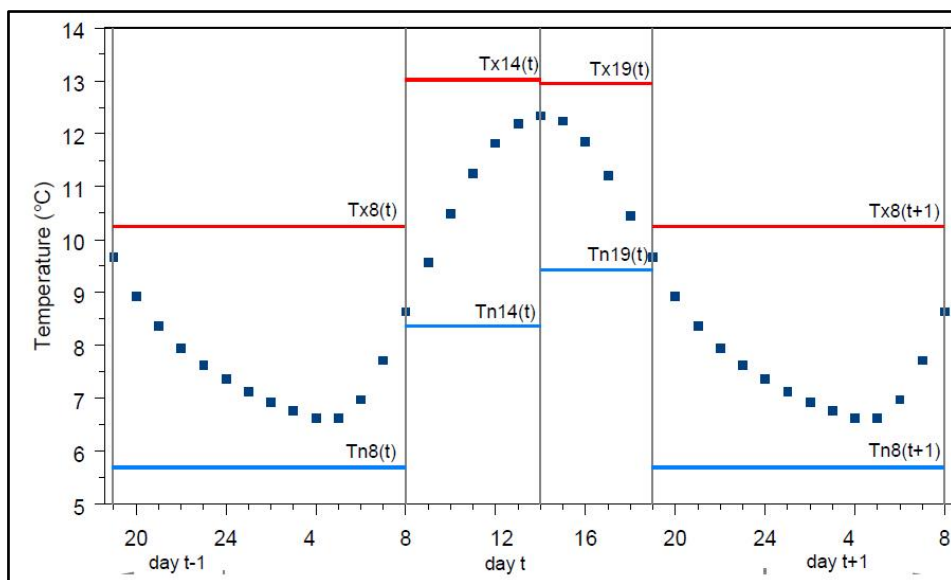
In het volgende hoofdstuk wordt onder andere ingegaan op de vraag of homogenisatie met behulp van een statistisch model, zonder daarbij de fysische factoren te betrekken, wel wetenschappelijk verantwoord is. En of homogeniseren vanuit meettechnische principes überhaupt wel mag.

2 Waarom homogeniseren?

Het KNMI kan een waarnemingsreeks homogeniseren als er veranderingen in waarnemingsomstandigheden hebben plaatsgevonden die een sprong of trendbreuk veroorzaken die niet klimatologisch van aard is. Te denken valt aan verplaatsingen van de stations en/of instrumenten, langzaam of abrupte veranderingen in de omgeving en veranderingen in instrumenten, meethutten en meetmethoden. De inhomogeniteiten kunnen volgens het KNMI zo groot worden dat een klimaatreeks zonder verdere correcties ongeschikt is voor klimaatonderzoek.³

Het KNMI vindt dat voor onderzoek naar klimaatverandering en variabiliteit het belangrijk is om deze inhomogeniteiten te corrigeren teneinde homogene datasets te verkrijgen. In de praktijk betekent dit dat men dat deel van de meetreeks dat inhomogeniteiten bevat homogeniseert. Die homogenisatie komt meestal neer op een wijziging van de meetgegevens met behulp van statistische technieken. De statistische techniek die toegepast is voor de homogenisatie van de dagelijkse temperaturen van De Bilt heet 'percentile matching' en wordt in hoofdstuk 4 nader toegelicht.

Een wisseling van meetinstrument zou geen inhomogeniteit mogen veroorzaken. De instrumenten dienen immers geijkt te zijn. Toch zijn er redenen om aan te nemen dat ijking van instrumenten niet altijd voldoende is. Sommige oudere thermometers waren op instrumenteel fysische grond niet nauwkeurig genoeg of hadden een langere reactietijd dan de moderne instrumenten. Ook de overgang van glazen vloeistofthermometers naar elektronische minimum-maximumthermometers leverde afwijkingen op. De elektronische meter geeft voor Tx lagere waarden aan, tot 0,5 °C.⁴



Figuur 7 Gemiddelde uurtemperaturen en gemiddelde waarden Tn en Tx De Bilt 1901-1970. Brandsma (2013)

Een bijzondere inhomogeniteit kan ontstaan door meetmethodiek. Tot en met 1932 werden Tn en Tx bepaald voor het interval van 8h tot 8h plaatselijke tijd, daarna tot 1971 van 19h tot 19h. Vanaf 1971 meet men van 0h - 0h UTC (gecoördineerde wereldtijd).

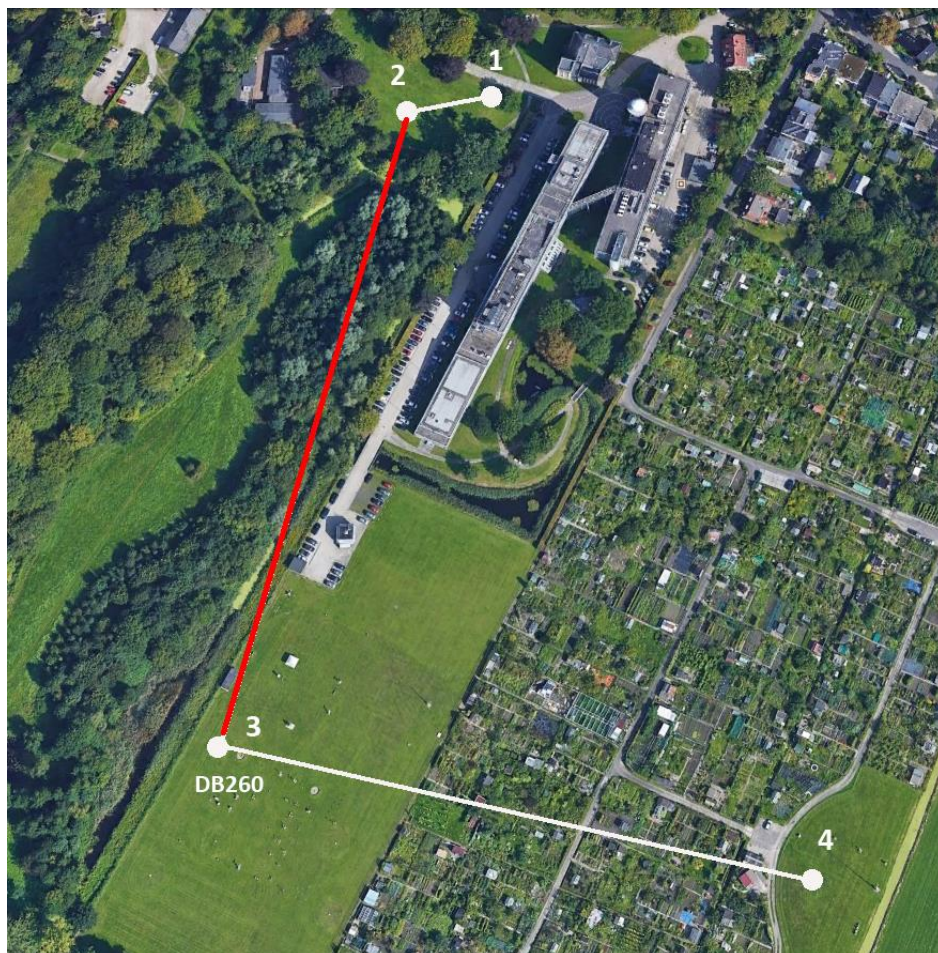
Deze verschillen beïnvloeden de gemiddelde waarden van Tn en Tx. Bovendien werden van 1946 tot 1971 twee waarnemingsmethodes naast elkaar gehanteerd, de synoptische (uurmetingen) en de klimatologische waarnemingen. Die laatste bestaan uit 3 metingen per etmaal van maximum- en minimumtemperatuur (om 8h, 14h en 19h) plus de papieren thermograafstroken. De metingen van

³ <https://www.knmi.nl/kennis-en-datacentrum/achtergrond/gehomogeniseerde-tijdreeksen-dagwaarden>

⁴ <https://judithcurry.com/2014/07/07/understanding-adjustments-to-temperature-data/>

1901 tot 1971 werden vanaf 2014 gestandaardiseerd met behulp van de klimatologische data zodat ze vergelijkbaar zijn geworden met de data vanaf 1971.

Omdat bij standaardisatie geen meetdata veranderd worden maar een correctie wordt toegepast op de wijze waarop T_n , T_x en T_g worden berekend wijkt standaardiseren strikt genomen af van homogeniseren vanwege bijvoorbeeld verandering van meetlocatie. Wij gaan in dit rapport voor wat betreft de ongehomogeniseerde etmaaltemperaturen uit van de gestandaardiseerde waarden van de hoofdstations.



Figuur 8 Verplaatsing van de meethut op 27 augustus 1951 (rode lijn), over een afstand van ongeveer 300m

Figuur 8 toont de verplaatsing van de thermometerhut in de loop van de jaren. Op 17 mei 1950 werd de grote pagodehut vervangen door een Stevensonhut (locatie 1). Vanwege bouwwerkzaamheden werd de nieuwe hut op 16 september 1950 tijdelijk enige tientallen meters westwaarts verplaatst (2). Op 27 augustus 1951 werd de hut over een afstand van ongeveer 300m zuidwaarts verplaatst naar een nieuw meetveld (3). In 2008 kreeg de hut zijn huidige locatie (4).

Aanleidingen om tot homogenisatie over te gaan zijn verandering van instrument, type meethut of meetlocatie. Elke verandering rondom een meetpunt en elke verplaatsing, hoe klein ook, levert immers andere fysische omstandigheden op en dus verschillende meetdata. Men kan dan parallelmetingen inplannen.

Lastiger wordt het als er inhomogeniteiten vermoed of ontdekt worden in oude meetreeksen. Soms zijn er aanwijzingen in metadata van een station die wijzen op een mogelijke inhomogeniteit. Vaak echter vermoedt men een inhomogeniteit op basis van het verloop van een meetreeks. Er kan dan sprake zijn van een opvallende trend of een sprong. Statistische technieken om een afwijkende trend en temperatuursprongen te ontdekken worden dan toegepast. Bij de samenstelling van de CNT (Centraal Nederlandse Temperatuur) werd de waarnemingsreeks van één station vergeleken met die van een ensemble van andere stations in Nederland. Lokale afwijkingen zullen dan in het oog springen.

Op het toepassen van homogenisatie van waarnemingsreeksen is de afgelopen jaren de nodige kritiek gekomen. Die kritiek is meestal van fundamentele aard en gebaseerd op de basisregels van meten, ijken en statistiek. Zo is temperatuurmeting op een bepaalde plaats conditioneel, dat wil zeggen afhankelijk van allerlei, al dan niet regelbare, parameters. Wat de relatie van het meetresultaat met de omstandigheden is, is vaak onbekend of wordt onvoldoende begrepen. Een voorbeeld daarvan is de invloed van het vochtgehalte van het bovenste deel van de bodem, kort gemaaid of langer gras et cetera.

Door die ontbrekende kennis wordt (bijvoorbeeld bij een verandering van locatie) de correctie niet fundamenteel aangepakt. Fundamenteel zou zijn om de invloed van de omstandigheden exact te kwantificeren op basis van fundamentele fysische of chemisch kennis. Omdat we dat niet kunnen wordt achteraf via statistische technieken homogenisatie toegepast.

Vanuit een fundamenteel principe van het meten van een grootheid is het toepassen van correcties op individuele data alleen toegestaan als per meetresultaat de fysische aard en de grootte van de correctie precies kunnen worden gegeven. Om hele datareeksen door een statistisch filter te halen om zo een veelvoud van veranderende condities (locatie, apparatuur, meethut, begroeiing, verstedelijking et cetera) te “corrigeren” is vanuit meettechnische principes niet goorloofd.

De Amerikaanse statisticus William Briggs⁵ reageerde uitgebreid op de homogenisatie van De Bilt naar aanleiding van een artikel van Frans Dijkstra in de Volkskrant⁶. De huidige temperatuurreeks van De Bilt is volgens Briggs een collage van een aantal afzonderlijke waarnemingsreeksen, waarvan de eerste twee zijn gehomogeniseerd:

“In the notes to the data it said in 1950 there was “relocation combined with a transition of the hut”. Know what that means? It means that the data before 1950 is not to be married to the data after that date. Every time you move a thermometer, or make adjustments to its workings, you start a new series. The old one dies, a new one begins. If you say the mixed marriage of splicing the disjoint series does not matter, you are making a judgment. Is it true? How can you prove it? It doesn’t seem true on its face. Significance tests are circular arguments here. After the marriage, you are left with unquantifiable uncertainty. This data had three other changes, all in the operation of the instrument, the last in 1993. This creates, so far, four time series now spliced together.”

⁵ <http://wmbriggs.com/post/14718/>

⁶ <https://www.volkskrant.nl/columns-opinie/het-publieke-weergeheugen-werkt-misleidend~b8f77b3e/>

Er is nog iets merkwaardigs aan de praktijk van homogeniseren. Als bijvoorbeeld een trendbreuk door homogenisatie wordt gladgestreken, werkt die homogenisatie door op alle aan de trendbreuk voorafgaande data. Voor het doel dat de wetenschapper voor ogen heeft, namelijk het creëren van een lange, homogene temperatuurreeks, maakt het echter niet uit of de reeks vóór of ná de breuk wordt gecorrigeerd. Dat dat laatste niet gebeurt heeft te maken met de overtuiging dat de huidige metingen correct zijn. Deze veronderstelling is echter maar beperkt houdbaar, namelijk tot de volgende homogenisatie vanwege verandering van meetinstrument, hut of meetlocatie. De Bruin zegt terecht hierover: *“Dit (homogeniseren) ondermijnt het eigen gezag, want wie zegt het grote publiek dat er momenteel wel goed wordt waargenomen en dat er later niet opnieuw moet worden geharmoniseerd?”*

Ook het homogeniseren op basis van de meetresultaten van een ander station (in dit geval Eelde) is vanuit meettechnische principes niet geoorloofd. IJken doe je op een standaard en niet op de resultaten van een ander station, omdat er dan geen sprake meer is van onafhankelijke meetgegevens.

We zouden het hierbij kunnen laten, homogeniseren moet je niet doen, maar de werkelijkheid is dat homogeniseren van weergegevens op grote schaal voorkomt. Het KNMI stelt: *“De temperatuurreksen van de vijf hoofdstations Den Helder, De Bilt, Groningen, Vlissingen en Maastricht zijn door stationsverplaatsingen inhomogeen. Voor De Bilt speelt daarnaast een verandering van een pagodehut naar een Stevensonhut een rol. De afzonderlijke reksen zijn daardoor over hun volle lengte, zonder correcties, meestal ongeschikt voor het doen van uitspraken over klimaatverandering en -variabiliteit.”* (Brandsma 2016b)

De homogenisatie door Brandsma roept nog meer vragen op. Hoe weet Brandsma dat de trendbreuk in het Tx signaal van De Bilt niet natuurlijk is? En hoe weet hij dat het aantal hittegolven in De Bilt te hoog was? Om dat soort uitspraken te kunnen doen moet De Bilt worden vergeleken met een norm, een referentie. Het technische rapport geeft daar geen uitsluitsel over.

Of er werkelijk sprake is van een inhomogeniteit in het signaal van De Bilt rond 1950, of er aantoonbaar teveel hittegolven in de eerste helft van de 20ste eeuw geteld zijn en of de homogenisatie van De Bilt correct is uitgevoerd zullen we in de volgende hoofdstukken proberen te achterhalen.

3 Parallelmetingen op het KNMI-terrein

Hutwisseling

Vanaf 1897 tot 1950 is de zogeheten Pagodehut gebruikt voor de temperatuurmetingen van station De Bilt. Het was een grote hut die aan de onderzijde open was, met een ventilerende dakconstructie. In figuur 9 is direct achter de Pagodehut een Stevensonhut te zien. Die is daar begin 1942 geplaatst, op 4m afstand van de Pagodehut.



Figuur 9 De Pagodehut met rechts achter de Stevensonhut

Van der Bijl (1952) noemt beide hutten en schrijft: *“Reeds spoedig opgemerkte verschillen gaven aanleiding tot het trekken van de conclusie, dat straling, afkomstig van het aardoppervlak, de nauwkeurigheid van de temperatuurmeting in de grote hut nadelig beïnvloedde. Zo wees bv. de thermometer in de grote hut op 22 april 1942 om 14 uur 3,0 °C hoger aan dan de thermometer in de Stevensonhut. Oorzaak was de sterke verhitting van het aangrenzende grasveld.”* Omdat Van der Bijl verder geen mededelingen doet over de metingen gaan we er van uit dat het genoemde voorbeeld exemplarisch is. Wel van belang is de constatering dat de Pagodehut blijkbaar gevoelig was voor warmtestraling vanaf de onderzijde, vanwege de open onderkant van de hut.

Brandsma (2011) vermeldt dat er parallelmetingen voorhanden zijn tussen Pagodehut en Stevensonhut: *“The change in screen type was accompanied by parallel measurements. We digitized and analyzed these data and found that the screen transition explains about half (0.37 °C) of the downward jump in summer (April-September) maximum temperature.”* Waarschijnlijk gaat het hier om dezelfde metingen als waarvan Van der Bijl verslag deed. Omdat deze gegevens voor ons onderzoek van belang kunnen zijn hebben we de data op 21 januari 2019 opgevraagd bij het KNMI. Helaas bleek het KNMI niet bereid om de gevraagde gegevens per ommegaande naar ons te sturen. Ze komen pas in maart 2019, na publicatie van een artikel van het KNMI in *Meteorologica*, beschikbaar, terwijl het KNMI (Brandsma 2011) wel al in 2011 gebruik heeft gemaakt van die parallele metingen. Hiervoor merkten we al op dat genoemde parallele metingen merkwaardigerwijs niet gebruikt zijn bij de homogenisatie.

Kramer et al (1954) verwijzen naar een periode van parallelmetingen tussen de Pagodehut en de nabij geplaatste Stevensonhut:

“In deze hut (bedoeld wordt de Pagodehut) en een normale Stevenson hut die er vlak bij geplaatst was, zijn gedurende een periode van enkele jaren dagelijks vergelijkende metingen verricht, die als resultaat hebben opgeleverd, dat beide hutten zeer goed overeenkomen ten aanzien van de gemiddelde over dagtemperatuur (gem. verschil over 9 maanden 0,005, uitersten +0,4, -0,2) en dat slechts de extremen bij de Stevenson-hut iets minder geprononceerd zijn. ”

Blijkbaar slaat dit op dezelfde parallele metingen die door Van der Bijl en Brandsma (2011) genoemd worden. De verschillen tussen de dagtemperaturen waren volgens Kramer zeer klein. Verschillen tussen de “extremen” (minimum en maximum-etmaaltemperaturen) waren bij de Stevensonhut iets minder geprononceerd.



Figuur 10 De proefopstelling van Kramer met links het vinydur hutje

Kramer et al (1954) hebben onderzoek gedaan naar parallelemetingen die tussen oktober 1945 en april 1947 gedurende een periode van 20 maanden zijn gedaan aan de grote Pagodehut en een nabij geplaatste klein kunststof hutje van vinydur. De opstelling daarvan is in figuur 10 te zien, links het bedoelde vinydurhutje. Dit hutje is kleiner dan de Stevensonhut en is daardoor en vanwege andere constructie- en materiaaleigenschappen niet een-op-een te vergelijken met de Stevensonhut.

De onderzoekers constateerden dat bij lagere temperaturen het verschil tussen Pagodehut en kleine hut iets groter was dan bij hogere temperaturen maar dat het verschil niet significant was. Dat is opvallend, omdat het KNMI stelt dat de grootste correcties voor T_x in de zomer nodig waren, bij de hoogste percentielen. Het technische rapport (Brandsma 2016a) stelt: *“For T_x large positive corrections (up to 1.9°C) are found in summer for the largest percentiles. This is mainly a result (of) a combination of the pagoda which was open at the bottom and affected by reflected sunlight and the enclosed location till 1951.”* Kramer et al hebben bovendien geconstateerd dat het verschil in de gemeten temperatuur tussen de Pagodehut en de vinydur hut vooral in het najaar en de winter

optrad, hoewel Kramer niet tijdens extreme zomertemperaturen heeft gemeten. Het KNMI gaat bij de homogenisatie van De Bilt echter uit van de grootste verschillen in de zomer.

Voorlopige conclusie is dat de verschillen tussen de metingen in de Pagodehut en de Stevensonhut waarschijnlijk klein waren. Een verschil van 0,005 °C tussen de gemiddelde dagwaarden over een periode van 9 maanden is verwaarloosbaar. Wel kunnen er verschillen zijn opgetreden in T_x en T_n tussen beide hutten. Vanwege de open onderzijde van de Pagodehut lijkt deze gevoelig te zijn geweest voor warmtestraling in de vroege morgen in combinatie met de lage windsnelheden in dit deel van de dag. Overdag zullen de verschillen waarschijnlijk gering zijn geweest vanwege de hogere windsnelheden (betere ventilatie in beide hutten).

De Pagodehut is in De Bilt in 2016 herbouwd. De resultaten van de metingen in deze nieuwe Pagodehut zijn op het moment van schrijven nog niet bekend. Wel heeft Brandsma⁷ in een presentatie laten zien dat op een dag in augustus 2017 een verschil is gevonden van 0,4 °C bij $T_x=24$ °C. Dat is veel lager dan de maximale correctie van 1,9 °C die het KNMI heeft toegepast op T_x .

Locatiewisselingen in 1950 en 1951

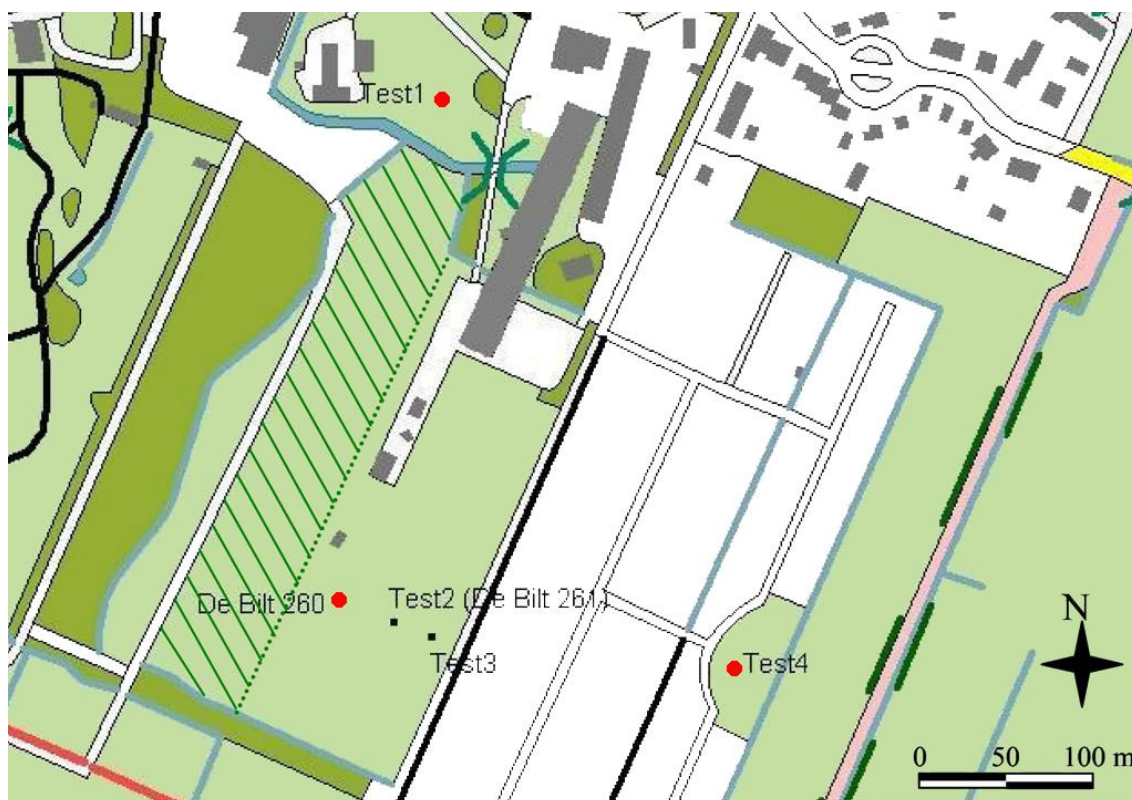
Na de officiële ingebruikname van de Stevensonhut op 17 mei 1950 werd de locatie van deze hut binnen 16 maanden tweemaal verplaatst. Bij de eerste locatiewijziging op 16 september 1950 werd de hut enkele tientallen meters westwaarts verplaatst. De nieuwe omstandigheden moeten veel geleken hebben op de oude: de omgeving waarin de hut dan staat is hetzelfde halfopen park als voorheen. Alleen de afstand tot de KNMI-gebouwen is wat groter geworden. We hebben geen gegevens gevonden over vergelijkingen tussen beide sites maar het lijkt aannemelijk dat deze verplaatsing geen grote gevolgen zal hebben gehad op de temperatuurmetingen.

Dat ligt minder voor de hand voor de tweede verplaatsing op 27 augustus 1951 naar een meetveld ongeveer 300m ten zuiden van de oude locatie. De afstand is relatief groot en de omgevingsomstandigheden zijn anders. Geen halfopen park met hier en daar een hoge boom maar een open meetveld zonder directe nabijheid van gebouwen. Wel is er ten Z en ten W van de nieuwe meetlocatie hoge begroeiing (bomen) aanwezig. De nieuwe locatie heeft van 27 augustus 1951 tot 25 september 2008 onder de codenaam DB260 dienst gedaan als meetlocatie voor de officiële temperatuurreeks van De Bilt.

Ook van deze verplaatsing zijn zoals al geschreven is geen parallelmetingen voorhanden. Wel zijn in de periode van mei 2003 t/m juni 2005 parallelmetingen verricht tussen een vijftal locaties op het KNMI-terrein. Daarbij waren behalve de 'nieuwe' locatie van DB260 vanaf 1951 ook de oude (tijdelijke) locatie van de periode 1950-1951 betrokken.

Een van de redenen om deze parallelmetingen te verrichten was “*....to explore the possibility of using present-day parallel measurements to correct for inhomogeneities, caused by changes in the surroundings and a relocation of the thermometer screen in 1951.*” (Brandsma 2011) In figuur 11 zijn de vijf locaties aangegeven.

⁷ Brandsma, T., 2018: presentatie op het symposium 'Old observation series in modern times', NVBM, De Bilt, 6-4-2018.



Figuur 11 Het testterrein voor de parallelmetingen tussen 2003-2005

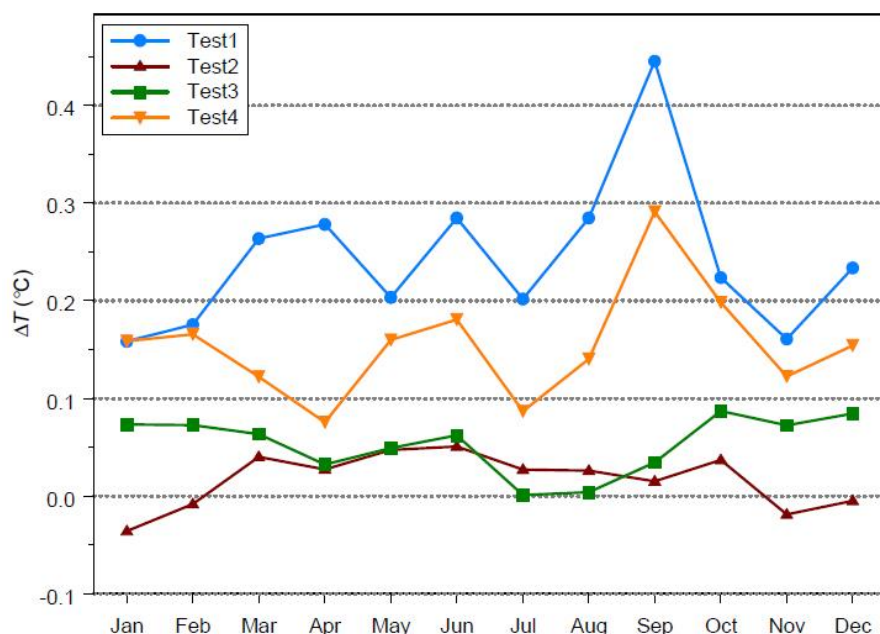
Test1 is de locatie waar de Stevensonhut heeft gestaan van 16 september 1950 tot 27 augustus 1951, nadat de Pagodehut wegens bouwwerkzaamheden moest verdwijnen. DB260 is de locatie waar de meetopstelling op 27 augustus 1951 naartoe is verplaatst. Test4 is de huidige locatie van de meetopstelling, in gebruik genomen op 25 september 2008. Test1 en DB260 zijn de locaties die voor de homogenisatie van de temperatuur in De Bilt van belang zijn.

Gedurende het experiment vond een aanzienlijke verandering plaats in het gebied dat in figuur 11 diagonaal groen gestreept is. Dit gebied was tot 1999 grasland geweest en vanaf die datum langzaam veranderd in ruderaal terrein met opschietende bosschages tot 3m hoog. In september en oktober 2004 werd dit ruderaal terrein veranderd in een park met waterpartijen en werden de struiken gerooid. Dit had tot gevolg dat er een inhomogeniteit ontstond in de meetreeks, het betreffende gebied was op slechts 12m van de meetlocatie DB260 gelegen.

Om de vergelijking tussen de locaties Test1 en DB260 niet te laten beïnvloeden door genoemde ingreep hebben we gebruik gemaakt van de gegevens die in een interim-rapport (Brandsma 2005) verschenen zijn en waarvan de testgegevens lopen van mei 2003 tot mei 2004. Bijkomend voordeel is dat in dit interim-rapport DB260 als referentielocatie gebruikt wordt. In het definitieve rapport (Brandsma 2011) is Test4 de referentielocatie geworden. Redenen daarvoor waren de genoemde inhomogeniteit in 2004 en het feit dat sinds 25 september 2008 Test4 de officiële meetlocatie van station De Bilt werd.

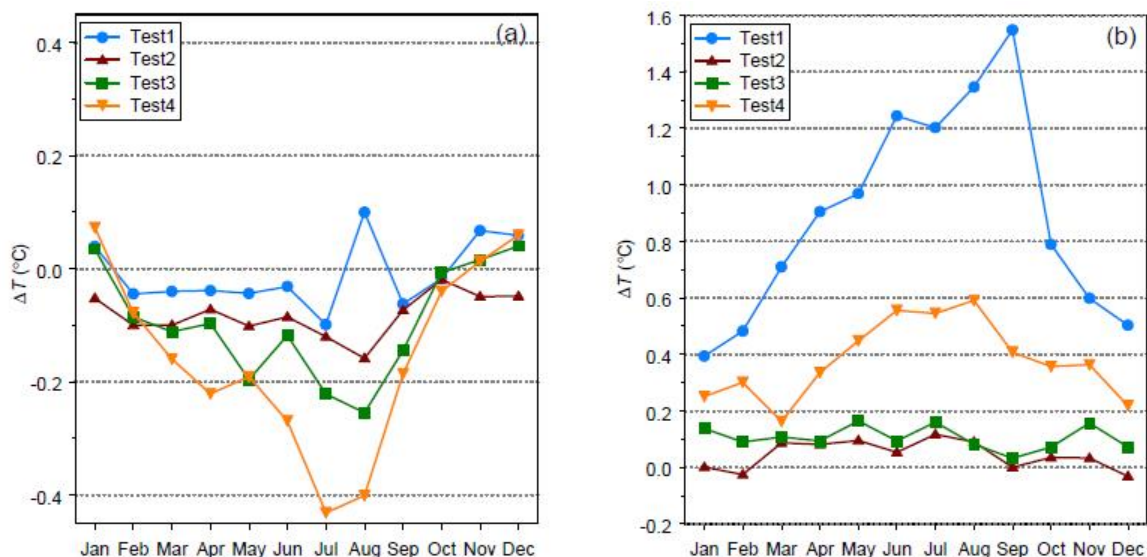
In figuur 12 is de gemiddelde dagelijkse temperatuur per maand aangegeven voor de 4 testlocaties ten opzichte van DB260. De afwijking van elk van de testlocaties is dus de afstand tussen de gemiddelde dagelijkse temperatuur van elke locatie ten opzichte van de 0.0 lijn. Duidelijk is te zien dat de waarden van de Test1-locatie aanmerkelijk hoger liggen dan van DB260. De cijfers van Test1 zijn ook hoger dan van de andere testlocaties. Wel valt op dat het verschil tussen Test1 en Test4 het geringst is. Test4 is de huidige (vanaf 25 september 2008) officiële meetlocatie van de temperatuur van station De Bilt.

De verschillen tussen de locaties kunnen te maken hebben met verschillen in windsnelheid, zonlicht, straling door nabijgelegen gebouwen et cetera. Met name van de windsnelheid is bekend dat het vaak een belangrijke oorzaak is van temperatuurafwijkingen, vooral overdag.



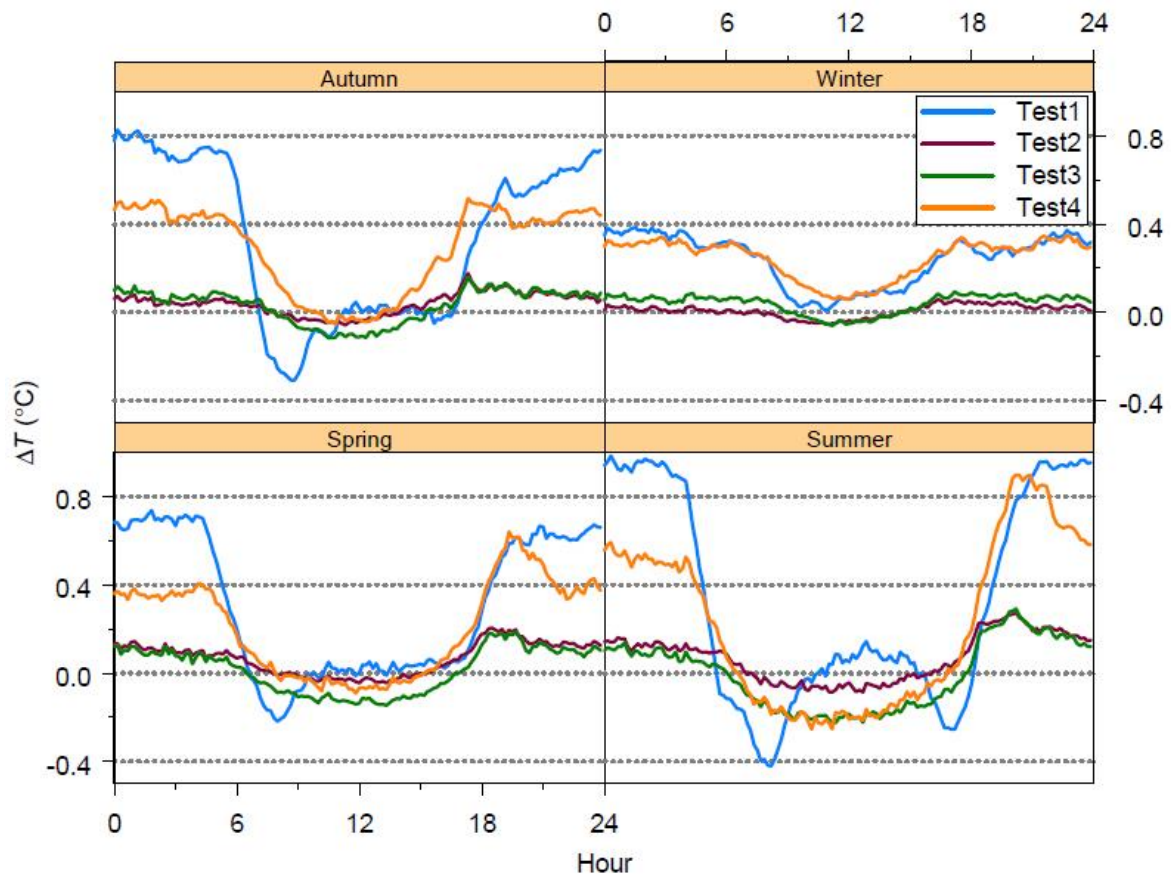
Figuur 12 Verskil tussen de gemiddelde dagelijkse temperatuur per maand van de 4 testlocaties ten opzichte van DB260

Bekijken we echter de gemiddelde dagelijkse maximum en minimum temperatuur per maand dan wordt het beeld heel anders:



Figuur 13 Verskil tussen de gemiddelde dagelijkse maximum (a) en minimum (b) temperatuur per maand van de 4 testlocaties ten opzichte van DB260

De gemiddelde T_x per maand van Test1 wijkt het minste af van die van DB260, terwijl de afwijkingen van de gemiddelde T_n van Test1 het grootst zijn. Om een verklaring te vinden voor die afwijkende minimum temperaturen op locatie Test1 heeft het KNMI de dagelijkse gang van de temperatuur ten opzichte van DB260 van de vier testlocaties in beeld gebracht:



Figuur 14 Verschillen tussen de 4 testlocaties en DB260 van de dagelijkse gang van de temperatuur

Te zien is dat de nachtelijke temperaturen van Test1 opvallend hoog zijn, met name in de zomer. Bovendien is te zien dat de temperatuur gedurende zonsopgang en zonsondergang fors daalt. Dat effect is het grootst in de zomer en het kleinst in de winter. Midden op de dag wijkt Tx van Test1 nauwelijks af van DB260.

De oorzaak van de hogere nachttemperaturen op Test1 moet niet gezocht worden in verschillen in windsnelheid, want de windsnelheid is 's nachts nagenoeg nul. Mogelijk is de kleine 'sky view' een oorzaak: de locatie is omringd door hoge bomen. Door die beperkte 'sky view' wordt de nachtelijke warmte-uitstraling naar boven gehinderd. Ook de warmtestraling van de nabijgelegen bebouwing kan een rol spelen.

De oorzaak van de dip bij zonsopgang en -ondergang wordt gezocht in het feit dat door de omringende begroeiing en bebouwing de thermometerhut op die momenten van de dag in de schaduw ligt. Waar de andere testlocaties dan verwarmd worden door de zon, daalt de temperatuur van de Test1-locatie doordat de warmte-uitstraling op die momenten groter is dan de opwarming door de zon.

Tijdens het warmste deel van de dag zijn de temperaturen op Test1 in de lente en herfst ongeveer gelijk aan die van DB260. In de zomer en winter zijn deze iets hoger dan op DB260 (ongeveer 0,1 °C).

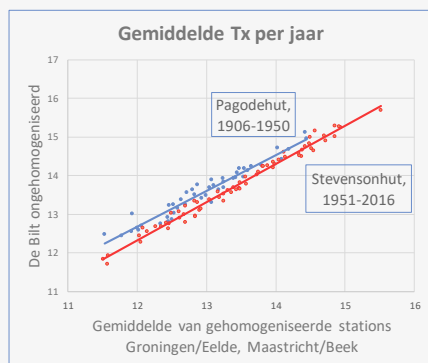
Conclusies

1. Het KNMI (Brandsma 2011) vond na het analyseren van parallelle gegevens dat de hutwisseling in 1950 ongeveer de helft ($0,37\text{ }^{\circ}\text{C}$) van de temperatuursprong rond 1950 verklaart. Wij kunnen dat niet verifiëren omdat de gegevens ons nog niet ter hand zijn gesteld.
2. Het KNMI heeft voor de homogenisatie om onbekende redenen geen gebruik gemaakt van de parallelle metingen tussen de Pagodehut en Stevensonhut.
3. Volgens Kramer et al (1954) is de gemiddelde afwijking van T_g tussen de Pagodehut en Stevensonhut over een periode van 9 maanden slechts $0,005\text{ }^{\circ}\text{C}$, met uiterste afwijkingen van $+0,4\text{ }^{\circ}\text{C}$ en $-0,2\text{ }^{\circ}\text{C}$. De extremen van de Stevensonhut zijn iets minder geprononceerd dan van de Pagodehut.
4. Kramer vond slechts geringe afwijkingen van de decadegemiddelden tussen Pagodehut en vinydur hut. Opvallend is dat het verschil in de zomermaanden op waarnemingstijd 14 uur ($\sim T_x$) tussen beide hutten gering is: de Pagodehut is gemiddeld $0,13\text{ }^{\circ}\text{C}$ warmer.
5. Het KNMI gaat uit van de grootste verschillen in de zomer en corrigeert dan T_x tot maximaal $1,9\text{ }^{\circ}\text{C}$ naar beneden. We concluderen voorzichtig dat de hutwisseling van Pagode naar Stevensonhut waarschijnlijk geen verschillen teweeg heeft gebracht die de sterke correctie van T_x tot $1,9\text{ }^{\circ}\text{C}$ rechtvaardigen.
6. Uit parallelle metingen (Brandsma 2008) blijkt dat gedurende het zomerhalfjaar de dagelijkse maximumtemperatuur op de Test1-locatie (de locatie tussen 16 september 1950 en 27 augustus 1951) niet significant afwijkt van die van DB260, terwijl het verschil van de gemiddelde dagelijkse minimumtemperatuur in het zomerhalfjaar $1,20\text{ }^{\circ}\text{C}$ bedroeg.
7. De verplaatsing van de meeteenheid van Test1 naar DB260 heeft nauwelijks effect gehad op T_x . Die locatiewisseling lijkt de grote correctie van T_x tot maximaal $1,9\text{ }^{\circ}\text{C}$ dan ook niet te rechtvaardigen.

Homogeniseren voor dummies

Dummy 1. Vaste correcties voor alle dagen voor de inhomogeniteit.

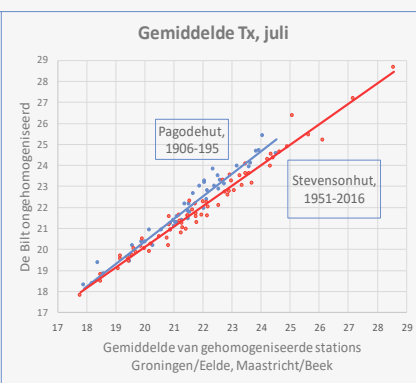
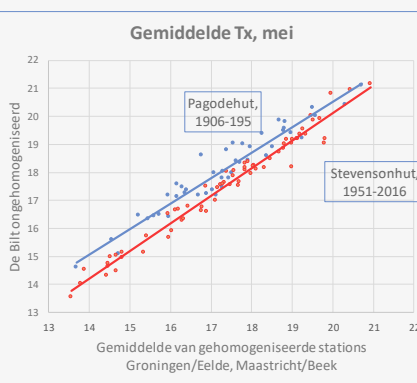
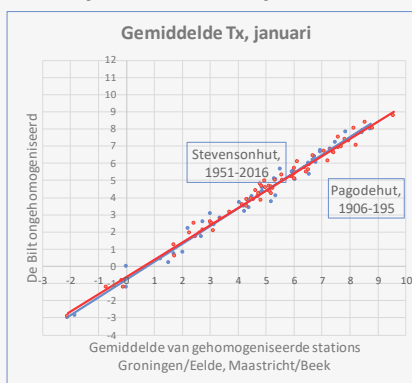
Omdat station De Bilt verdacht wordt niet homogeen te zijn, zetten we de gemiddelde temperatuur per jaar in De Bilt uit tegen de gemiddelde temperatuur van stations die gehomogeniseerd zijn.



De rode lijn geeft de metingen van De Bilt in de Stevensonhut, de blauwe lijn in de pagodehut. De blauwe lijn ligt ongeveer 0,3°C hoger dan de rode lijn. Dus was het in de pagodehut gemiddeld iets te warm. Hiervoor kan gecorrigeerd worden door van alle Tx-waarden 0,3°C af te trekken.

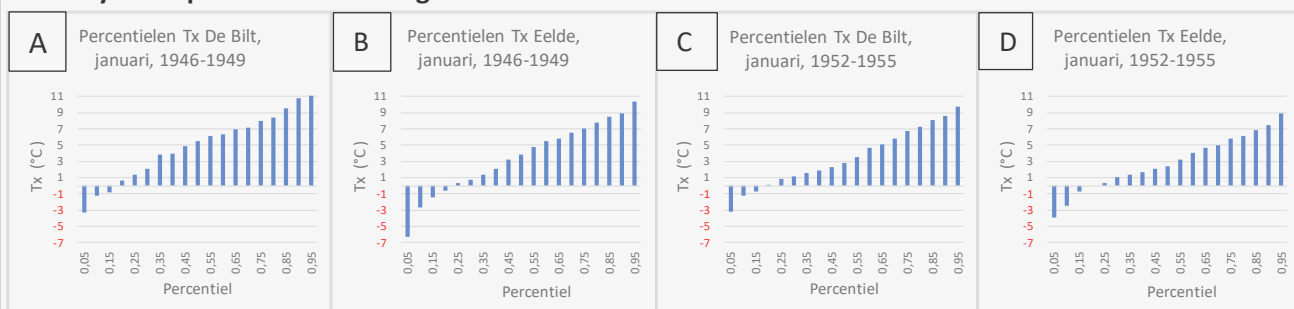
Maar het hangt ook van het seizoen af. Daarom herhalen we dit in **Dummy 2** voor alle maanden van het jaar. Hier worden alleen januari, mei en juli getoond.

Dummy 2. Correcties per maand.



In januari is het in de pagodehut niet warmer dan in de Stevensonhut, maar in mei is het in de pagodehut 0,5-0,8°C warmer. In warme julimaanden is het in de pagodehut ongeveer 0,8°C te warm, maar in koude julimaanden maakt het niet uit. Dat brengt ons op **Dummy 3**. Je moet het niet alleen per maand bekijken, maar ook onderscheid maken tussen koude en warme dagen: sorteert de dagelijkse temperaturen in twee vergelijkingsperioden van de koudste dag tot de warmste dag. De koudste 5% vallen onder het 5%-percentiel, de daarop volgende 5% vallen onder het 10% percentiel, enzovoorts. In **Dummy 3** wordt dit getoond voor de maand januari.

Dummy 3. De percentiel matching methode.



De reeks voor De Bilt heeft een inhomogeniteit rond 1950, de reeks voor Eelde is homogeen. De ruwe correcties voor De Bilt berekenen we per maand en per percentiel als volgt:

A-B = verschil tussen De Bilt en Eelde voor 1950.

C-D = verschil tussen De Bilt en Eelde na 1950.

(A-B) - (C-D) is het verschil tussen beide verschillen, waarmee per percentiel kan worden gecorrigeerd voor de verandering in De Bilt. Dit doen we voor alle 12 maanden. De ruwe correcties vertonen grote verschillen tussen aangrenzende maanden en aangrenzende percentielen. Daarom worden de correcties afgevlakt met een smoothing-filter.

4 De reconstructie van de homogenisatie van De Bilt

In dit hoofdstuk beschrijven we onze reconstructie van de homogenisatie die door het KNMI is toegepast op de etmaaltemperaturen van De Bilt. Die homogenisatie betreft T_g , T_n en T_x van 1 januari 1901 tot 1 september 1951. Het KNMI heeft de temperatuurreeksen van De Bilt gecorrigeerd voor de trendbreuk die in 1950/1951 zou zijn opgetreden. Een aantal detailgrafieken zijn te vinden in bijlage 2 (zie Appendix).

In het technische rapport van de homogenisatie schrijft het KNMI (Brandsma 2016a): *“For the investigated stations, it is known that there have been major relocations that caused inhomogeneities. In addition, for station De Bilt the relocation was accompanied by a major change in thermometer screen.”* Zowel de hutwisseling als de herlocatie in 1950 en 1951 zijn beschreven in hoofdstuk 1 en 3.

Omdat er volgens het KNMI geen parallelle metingen beschikbaar waren heeft men volgens het technische rapport als referentiereeks de data van Eelde gebruikt voor de periode 1946-1949 (vóór de veranderingen in De Bilt) en 1952-1955 (na de veranderingen in De Bilt). Eelde is ook gebruikt om de reeks van Groningen te homogeniseren bij de verplaatsing in 1951 vanuit de stad Groningen naar vliegveld Eelde, 10 km ten zuiden van de stad. We hebben gebruik gemaakt van de niet gehomogeniseerde maar wel gestandaardiseerde reeks (Brandsma 2013) van De Bilt.⁸ Met deze twee reeksen beschikken wij over hetzelfde basismateriaal als het KNMI.

Methoden

Het KNMI heeft de temperatuurreeksen gehomogeniseerd met de *percentiel matching methode*. Daarbij worden voor en na de transitie van een meetpunt gedurende een reeks van jaren parallelle metingen verricht, op grond waarvan voor elke maand percentielen in stappen van 0,05 van de gemeten temperaturen worden bepaald. Door de metingen per percentiel per maand vóór en na de transitie met elkaar te vergelijken kan een correctie worden berekend, waarmee de metingen van vóór de transitie kunnen worden aangepast.

Alle metingen tussen twee percentielen worden verminderd met de correctie die is berekend voor het dichtstbijzijnde hogere percentiel. In een periode van 4 jaar zijn er 113 - 124 dagen voor dezelfde maand. Dat betekent dat per percentiel in de vergelijkingsperioden gemiddeld ongeveer 6 metingen beschikbaar zijn. Dit leidt tot veel ruis in de data, doordat grote verschillen kunnen optreden tussen aangrenzende maanden en aangrenzende percentielen. Daarom heeft men de reeksen afgevlakt met een Loess smoother, waarvoor de span op 0,6 is gesteld op grond van *“a mixture of cross-validation and expert opinion.”*

Voor het berekenen van de percentielen verwijst het KNMI naar Hyndman en Fan (1996). Niet vermeld wordt welke methode er is gebruikt van de diverse mogelijkheden die door deze auteurs zijn genoemd. Volgens Wikipedia (2018) zijn er 5 hoofdmethoden om percentielen te bepalen, met een aantal subvarianten. In Excel zijn vanaf 2013 twee percentiel-functies beschikbaar: `percentielen.exc` en `percentielen.inc`. Het verschil zit in het wel of niet meetellen van de mediaan bij de beide helften van de populatie. Inclusieve percentielen liggen dicht bij de mediaan dan exclusieve percentielen en dit verschil neemt toe naarmate een getal verder van de mediaan af ligt. Wij hebben de inclusieve en de exclusieve methode gebruikt en het effect daarvan onderzocht op de resultaten van de

⁸ https://cdn.knmi.nl/knmi/map/page/klimatologie/gegevens/daggegevens/temp_260.txt

homogenisatie. Het KNMI-rapport vermeldt niet welke software gebruikt werd voor de Loess smoothing. Wij hebben de Loess utility voor Excel gebruikt volgens Peltier.⁹

De span van de smoothing kan worden uitgedrukt als de fractie van het aantal meetpunten dat wordt meegerekend in de smoothing. Er is geen objectieve maat voor de te gebruiken span. Gebruikt men een hoge span, dan wordt de ruis in de data sterk afgevlakt, maar verdwijnen fluctuaties op kortere termijn, die wel relevant kunnen zijn voor wat men wil laten zien. Gebruikt men een lage span, dan kan de ruis in de data zodanig overheersen dat men geen trend meer ziet.

Bij een span van 1,0 worden alle punten meegerekend. Bij korte reeksen moet een hoge span worden genomen, omdat anders de ruis domineert. Bij lange reeksen is een hoge span niet nodig. Bij een tijdreeks over meer dan 100 jaar (zie bijvoorbeeld figuur 53 en 54 in bijlage 2) is het voldoende 20 of hoogstens 30 punten mee te rekenen (de span is dan minder dan 0,2 of 0,3) om de ruis af te vlakken, maar toch relevante fluctuaties in de trend te blijven zien.

Bij de percentiel matching methode heeft men te maken met kleine reeksen: 19 punten voor de percentielen en 12 punten voor de maanden. De keuze van de span is daarbij van grote invloed op het resultaat. Wij hebben verschillende waarden van de span gebruikt en het effect van de span onderzocht op de resultaten van de homogenisatie. Bij de Loess utility in Excel is het minimaal aantal punten dat in de span kan worden meegerekend 5. Bij 12 maanden is de minimale span dus 0,417. Voor 19 percentielen is de minimale span 0,27.

Het technische rapport van het KNMI (Brandsma 2016a) vermeldt over de uiteindelijke correcties: *"The 5 and 95 percentiles are taken as the upper and lower limit, respectively. It was found that for the given number of days these are reasonable limits. For the final correction, temperature values more extreme than the limits are corrected using corrections of the nearest limit, so beyond the limits the correction remains constant."* Wij interpretern dit als volgt: temperaturen t/m het 5^e percentiel krijgen de correctie van het 5^e percentiel, temperaturen boven het 5^e t/m het 10^e percentiel krijgen de correctie van het 10^e percentiel, etc., temperaturen boven het 90^{ste} percentiel en boven het 95^{ste} percentiel krijgen de correctie van het 95^{ste} percentiel. Deze interpretatie wordt bevestigd door het feit, dat de correcties in augustusmaanden 1,9 graad bedragen voor alle temperaturen boven 28 graden. Dat is in de augustusmaanden van 1946-1949 het 90^{ste} percentiel (zie tabel 1).

We hebben in Excel een VBA-programma¹⁰ geschreven dat voor iedere dagtemperatuur van 1901 t/m 1950 eerst op grond van de percentielen voor De Bilt van 1946-1949 bepaalt onder welk percentiel van de betreffende maand die temperatuur valt, waarna de daarbij behorende correctie wordt toegepast. Een alternatief kan zijn, dat daarvoor de percentielen van 1952-1955 worden gebruikt. Wij hebben het verschil tussen beide methoden onderzocht. Voor de tropische dagen blijkt dit niets uit te maken, aangezien die allemaal in beide perioden in de twee hoogste percentielen vallen.

⁹ <https://peltiertech.com/loess-smoothing-in-excel/>

¹⁰ VBA = Visual Basic for Applications, de standaard programmeertaal voor macro's in Excel

Resultaten en discussie

Percentielen TX De Bilt 1946-1949 per maand													Percentielen TX Eelde 1946-1949 per maand												
01	02	03	04	05	06	07	08	09	10	11	12		01	02	03	04	05	06	07	08	09	10	11	12	
5	-3,3	-4,7	0,8	9,2	11,2	15,2	16,6	17,2	16,1	7,2	3,3	-3,0	5	-6,4	-6,8	-0,8	8,0	10,4	14,4	15,8	17,2	14,8	6,1	2,7	-2,9
10	-1,3	-3,1	1,9	10,0	12,7	16,1	17,4	18,4	16,3	8,0	4,8	-0,2	10	-2,7	-5,0	-0,4	9,2	11,5	14,8	16,4	17,4	15,7	7,6	4,0	-0,7
15	-0,9	-2,0	2,7	11,0	13,6	16,4	17,9	19,1	17,1	9,2	5,6	2,2	15	-1,5	-4,2	0,4	9,8	12,6	15,5	17,0	18,0	16,5	8,3	5,0	1,1
20	0,6	-1,0	3,1	11,4	14,5	17,0	18,2	19,4	17,2	10,9	6,1	2,5	20	-0,6	-3,0	1,2	10,2	13,5	15,9	17,4	18,5	16,9	10,3	5,9	2,1
25	1,4	0,1	4,2	12,0	15,3	17,3	19,2	19,6	17,9	12,1	6,9	3,0	25	0,3	-1,9	2,0	11,0	14,6	16,1	18,4	18,9	17,4	11,6	6,6	3,0
30	2,1	1,1	5,2	12,2	16,1	18,0	20,2	20,0	18,2	13,0	7,6	3,8	30	0,8	0,9	2,9	11,3	15,0	16,5	18,9	19,3	17,9	12,0	6,9	3,5
35	3,8	2,7	6,5	12,8	16,5	18,3	20,3	20,2	18,5	13,3	7,9	4,3	35	1,4	2,8	4,0	11,6	15,4	16,9	19,6	19,7	18,3	12,7	7,6	4,0
40	4,0	4,2	7,5	13,3	16,9	18,6	20,9	20,7	19,2	14,1	8,2	4,8	40	2,1	3,5	5,6	11,9	16,0	17,3	19,9	19,9	18,6	13,1	7,9	4,3
45	4,9	5,3	8,9	13,9	17,6	19,1	21,2	21,1	19,7	15,0	8,7	5,3	45	3,3	4,5	6,7	12,5	16,4	17,7	20,3	20,2	18,9	13,6	8,2	4,8
50	5,5	6,4	9,6	14,3	18,3	19,5	22,2	21,8	20,1	15,4	9,0	6,0	50	3,8	6,0	7,6	13,4	16,9	18,2	21,1	21,0	19,6	14,5	8,5	5,2
55	6,1	7,4	11,0	15,3	19,2	20,0	23,2	22,0	20,5	15,7	9,3	6,3	55	4,8	6,6	8,5	14,1	17,5	18,5	21,3	21,6	19,8	15,0	8,8	5,7
60	6,3	7,8	12,1	16,0	19,3	20,2	23,7	23,2	21,2	16,2	10,0	6,8	60	5,5	7,1	9,7	14,4	17,9	19,1	22,0	21,9	20,1	15,3	9,2	6,2
65	7,0	8,3	12,3	16,4	20,0	20,8	24,5	24,1	22,7	16,3	10,3	7,5	65	5,8	7,5	10,5	15,2	18,6	19,5	22,4	22,4	20,8	15,8	9,6	6,9
70	7,2	8,7	13,0	16,9	20,5	21,2	25,6	24,7	23,3	17,0	11,1	8,0	70	6,5	7,8	10,8	15,6	19,5	19,8	23,0	23,0	21,8	16,2	10,2	7,5
75	8,0	9,0	13,5	17,3	21,4	22,1	26,9	25,3	24,0	17,5	11,9	8,3	75	7,1	8,4	11,7	16,2	20,2	20,3	24,9	23,8	22,6	16,6	10,8	8,2
80	8,4	9,4	14,8	19,1	22,3	23,2	28,2	26,0	24,9	18,0	12,2	9,2	80	7,8	8,8	12,5	17,4	21,5	21,5	25,7	24,0	23,6	17,4	11,3	8,6
85	9,6	10,0	15,3	20,9	23,9	25,6	29,0	26,7	25,3	18,9	12,5	9,9	85	8,5	9,2	12,9	19,1	22,0	23,2	26,8	24,9	24,0	17,9	11,7	8,9
90	10,8	10,5	16,6	21,8	25,3	28,8	30,3	28,0	26,1	19,3	13,2	10,6	90	8,9	9,6	14,7	20,9	23,5	27,6	27,7	26,7	25,2	18,6	12,3	9,8
95	11,2	12,3	17,8	24,1	27,2	31,2	32,0	29,6	28,2	20,2	14,1	11,3	95	10,4	10,8	15,8	23,2	25,8	29,6	31,1	27,6	26,7	19,7	13,0	10,9

Percentielen TX De Bilt 1952-1955 per maand													Percentielen TX Eelde 1952-1955 per maand												
01	02	03	04	05	06	07	08	09	10	11	12		01	02	03	04	05	06	07	08	09	10	11	12	
5	-3,2	-0,8	1,7	8,3	10,4	13,7	15,9	16,6	12,6	8,6	0,8	0,3	5	-4,0	-3,4	0,2	7,3	9,5	12,8	15,9	16,1	12,4	7,2	0,1	-0,2
10	-1,2	-0,2	2,5	9,2	11,4	14,7	16,5	17,3	13,8	9,8	2,5	1,7	10	-2,5	-2,2	1,3	8,4	10,5	13,8	16,5	17,1	13,0	9,6	2,7	0,9
15	-0,7	0,1	3,7	10,0	12,3	15,4	17,1	18,2	14,7	11,0	4,7	3,1	15	-0,7	-1,3	2,0	8,9	11,5	14,5	16,7	17,5	13,8	10,3	3,7	1,8
20	0,1	0,9	4,5	10,2	12,6	15,9	17,5	18,8	15,4	11,8	5,5	3,7	20	0,0	-1,0	2,8	9,9	12,5	15,5	17,3	18,0	14,8	10,8	5,1	2,5
25	0,8	1,2	4,9	10,6	13,5	17,1	18,1	19,2	16,0	12,2	6,5	4,4	25	0,3	-0,6	3,8	10,0	12,7	15,9	17,5	18,6	15,6	11,2	6,7	3,0
30	1,2	1,9	5,9	11,0	14,0	17,3	18,4	19,5	16,2	12,8	7,5	4,9	30	1,1	-0,1	4,3	10,2	13,0	16,3	18,0	18,8	15,7	11,8	7,0	3,8
35	1,6	2,3	6,4	11,2	14,5	17,5	18,7	19,9	16,3	13,0	8,2	5,8	35	1,4	0,6	5,2	10,6	13,5	17,1	18,4	19,2	15,9	12,2	7,6	4,8
40	1,9	2,9	6,8	11,5	15,3	18,1	19,2	20,1	17,0	13,5	8,5	6,2	40	1,7	1,4	5,6	10,8	14,1	17,8	18,7	19,7	16,4	12,6	7,9	5,1
45	2,3	3,7	7,8	11,8	15,6	18,5	19,6	20,3	17,1	14,1	9,2	6,8	45	2,1	2,0	6,2	11,0	14,9	18,3	18,8	20,0	16,6	12,9	8,5	5,4
50	2,8	4,5	8,3	12,4	16,2	18,9	20,2	20,7	17,3	14,2	9,5	7,2	50	2,4	2,7	6,7	11,3	15,4	18,8	19,3	20,5	16,8	13,4	8,9	6,0
55	3,6	4,8	8,8	13,0	16,9	19,2	20,4	21,5	17,7	14,3	9,8	7,6	55	3,2	3,0	7,2	11,8	16,4	19,1	19,5	20,9	17,0	13,9	9,0	6,3
60	4,7	5,2	9,4	13,4	17,4	20,1	20,9	22,0	17,9	14,5	10,1	8,1	60	4,1	3,8	7,8	12,1	17,7	19,8	19,9	21,7	17,4	14,3	9,6	6,9
65	5,1	5,9	10,7	13,7	18,4	20,5	21,0	22,4	18,5	14,9	10,3	8,4	65	4,6	4,3	8,8	13,2	18,3	20,2	20,6	22,0	17,7	14,4	9,9	7,7
70	5,9	6,2	11,2	14,3	19,6	21,4	22,1	22,8	19,1	15,2	10,7	8,9	70	5,0	5,6	10,0	13,7	19,0	20,8	21,0	22,6	18,2	14,8	10,5	8,1
75	6,8	6,9	12,2	15,2	20,2	22,1	22,6	23,2	19,5	15,9	11,7	9,2	75	5,8	5,8	10,8	14,3	19,5	21,5	22,3	23,2	19,4	15,0	11,0	8,9
80	7,3	7,4	12,8	15,9	20,5	23,2	23,3	23,8	20,2	16,4	12,1	9,8	80	6,1	6,6	11,7	15,4	20,5	22,6	22,9	23,9	20,0	15,3	11,4	9,2
85	8,1	8,0	13,5	18,4	22,2	23,7	24,4	24,5	21,0	17,1	12,3	10,3	85	6,8	7,7	11,9	16,7	21,7	23,7	23,8	24,6	21,3	16,2	11,7	9,5
90	8,7	8,5	15,0	20,2	25,3	25,1	26,1	25,8	22,2	17,7	13,1	10,6	90	7,5	8,6	13,3	19,2	22,8	24,6	26,6	25,6	22,1	16,8	12,0	10,2
95	9,7	9,8	16,2	21,5	27,3	27,3	27,9	27,2	24,2	18,4	14,3	12,4	95	9,0	9,7	14,6	20,6	25,5	26,9	27,6	26,7	23,6	17,5	13,3	11,5

Tabel 1 Percentielen van de maximale etmaaltemperaturen per maand in de vergelijkingsperioden 1946-1949 en 1952-1955 in Eelde en De Bilt.

Tabel 1 laat de temperaturen per maand en per percentiel zien voor de twee stations in de twee vergelijkingsperioden. We hebben daaruit de correcties berekend zoals Brandsma beschrijft.

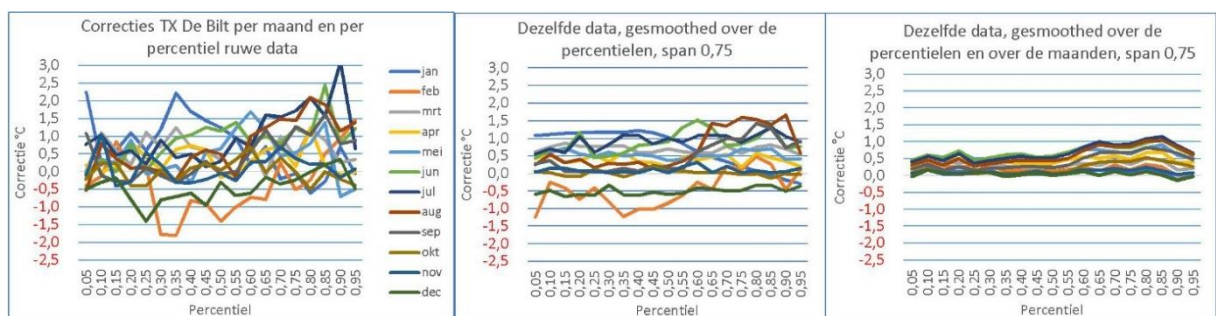
- 1) Voor elke maand is het verschil berekend tussen de overeenkomstige percentielen in De Bilt en Eelde in de twee perioden en vervolgens is het verschil tussen die twee verschillen berekend. Dat levert de ruwe correcties op die in tabel 2 zijn weergegeven. De belangrijkste conclusie die uit tabel 2 kan worden getrokken is, dat de ruis in deze data zeer groot is. Dat is niet verwonderlijk, omdat de basis van de correcties bestaat uit data uit twee zeer verschillende meteorologische perioden. Van 1946-1949 was er één extreem warme zomer (1947), één warme zomer (1948) en één extreem koude winter (1947). Van 1952-1955 waren er alleen koele zomers, drie zachte winters en één koude winter (1954). Tropische dagen kwamen van 1952-1955 weinig voor en geen enkel percentiel heeft in die periode in De Bilt of in Eelde een $T_x > 30^\circ\text{C}$. Bovendien liggen De Bilt en Eelde 147 km van elkaar. Door de gegevens van die stations in die perioden met elkaar te vergelijken kan het niet anders, dan dat men veel ruis genereert. Het is de vraag of daaruit betekenisvolle conclusies kunnen worden getrokken.

Correctie TX De Bilt per maand per percentiel: $\Delta T = \Delta T1 - \Delta T2$												
$\Delta T1$: temperatuurverschil De Bilt - Eelde voor de relocatie												
$\Delta T2$: temperatuurverschil De Bilt - Eelde na de relocatie												
	01	02	03	04	05	06	07	08	09	10	11	12
5	2,3	-0,4	0,0	0,2	-0,1	-0,1	0,8	-0,5	1,1	-0,2	-0,1	-0,5
10	0,2	-0,1	1,1	0,0	0,3	0,4	1,1	0,8	-0,1	0,3	1,1	-0,3
15	0,6	0,9	0,6	0,1	0,2	0,0	0,5	0,4	-0,3	0,2	-0,4	-0,2
20	1,1	0,1	0,2	0,9	0,9	0,8	0,6	0,1	-0,3	-0,4	-0,3	-0,8
25	0,6	0,2	1,1	0,4	-0,1	0,0	0,2	0,1	0,1	-0,4	0,6	-1,4
30	1,2	-1,8	0,7	0,1	0,1	0,5	0,9	-0,0	-0,1	0,1	0,2	-0,8
35	2,2	-1,8	1,3	0,6	0,2	0,9	0,4	-0,2	-0,3	-0,3	-0,3	-0,7
40	1,7	-0,8	0,7	0,7	-0,3	1,0	0,5	0,4	-0,1	0,1	-0,3	-0,6
45	1,5	-0,9	0,6	0,6	0,5	1,3	0,1	0,6	0,3	0,2	-0,2	-1,0
50	1,3	-1,4	0,4	-0,2	0,7	1,2	0,3	0,5	-0,1	0,1	-0,0	-0,3
55	1,0	-1,0	0,9	-0,0	1,2	1,4	1,0	-0,2	0,0	0,4	-0,2	-0,7
60	0,2	-0,7	0,8	0,3	1,7	0,9	0,7	1,0	0,6	0,7	0,3	-0,6
65	0,7	-0,8	-0,1	0,6	1,2	1,0	1,6	1,3	1,2	-0,0	0,3	-0,2
70	-0,2	0,3	1,0	0,7	0,4	0,8	1,6	1,5	0,6	0,4	0,7	-0,4
75	-0,1	-0,5	0,4	0,2	0,5	1,2	1,7	1,5	1,3	-0,0	0,4	-0,2
80	-0,6	-0,2	1,2	1,2	0,8	1,0	2,1	2,1	1,0	-0,5	0,2	-0,0
85	-0,3	0,5	0,9	0,2	1,4	2,4	1,6	1,9	1,6	0,0	0,2	0,2
90	0,7	0,9	0,3	-0,1	-0,7	0,8	3,1	1,2	0,8	-0,2	-0,2	0,3
95	-0,1	1,5	0,3	-0,0	-0,5	1,2	0,7	1,4	0,9	-0,4	0,1	-0,5

Tabel 2 Correcties per maand en per percentiel als gevolg van vergelijking van de percentielen per maand in de vergelijkingsperioden.

- 2) Om tot bruikbare correcties per maand en per percentiel te komen is een sterke afvlakking van de verschillen tussen aangrenzende percentielen en maanden nodig. Daarvoor moet men in tabel 2 eerst een smoothing uitvoeren over de kolommen (de percentielen), en vervolgens over de rijen (de maanden), of andersom. Figuur 15 laat het effect van de smoothing zien. De linker grafiek is de grafische weergave van tabel 2, de middelste grafiek laat het resultaat zien van de smoothing over de percentielen, de rechter grafiek laat het resultaat zien na smoothing over de percentielen en over de maanden. Bij de smoothing over de maanden heeft het KNMI vóór januari de gegevens van december en november gekopieerd, en na december de gegevens voor januari en februari. Alleen op die manier kan worden voorkomen, dat deze maanden onbedoeld worden beïnvloed door het feit dat ze aan het begin of het einde van de reeks staan. Dat lijkt ons een juiste methode, omdat de smoothing aan het begin en het eind van een reeks minder nauwkeurig wordt. Wij hebben om deze reden 5 maanden toegevoegd aan het begin en het einde van een jaar, zodat ook zinvol gesmoothed kan worden over een periode van 10 maanden. De span is dan $10/12 = 0,83$.

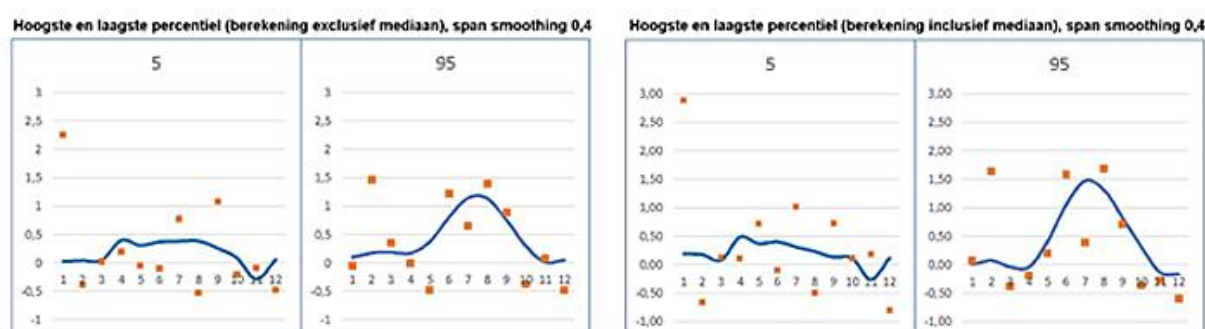
3)



Figuur 15 Correctiefactoren per maand en per percentiel. Links de data zonder smoothing, midden de data na smoothing over de percentielen, rechts de data na smoothing over de percentielen en over de maanden.

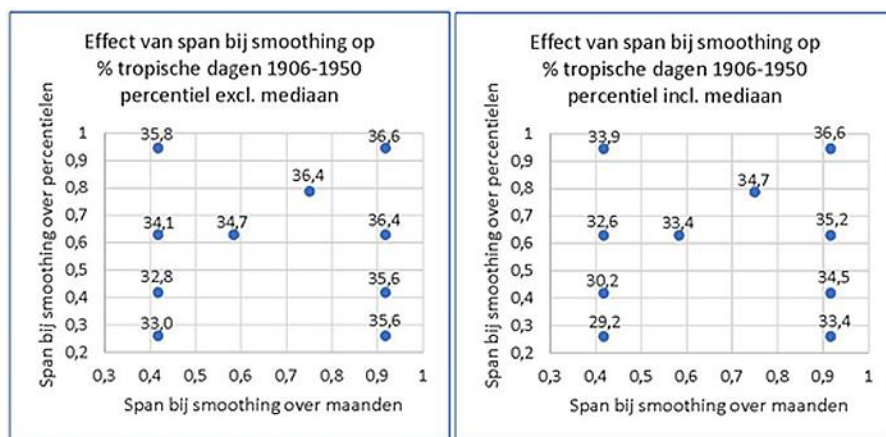
De correctiefactoren voor de 12 maanden zijn per percentiel in beeld gebracht in de figuren 49 t/m 51 van bijlage 2. Daarbij zijn drie varianten gebruikt van de span van de twee smoothings: 0,4, 0,6 en 0,9. Deze percentielen zijn exclusief de mediaan. Uit deze figuren blijkt dat de span een grote invloed heeft op de omvang van de correcties. Het is dus mogelijk om door variatie van de span het resultaat te sturen in een gewenste richting. Het is moeilijk om op objectieve gronden een keuze te maken. Het KNMI schrijft dat men ‘*expert opinion*’ gebruikt. Wat die opinie inhoudt maakt men niet duidelijk, de keuze voor de span wordt niet gemotiveerd. Figuur 16 in het technisch KNMI-rapport (Brandsma, 2016a), hier weergegeven als figuur 52 in bijlage 2 komt het meest overeen met onze figuur 50 van bijlage 2 maar vertoont minder fluctuaties. Waarom het KNMI niet is uitgekomen op ongeveer dezelfde figuren als onze figuren 50 of 51 in bijlage 2 is niet duidelijk.

Wij hebben de figuren 49 t/m 51 in bijlage 2 ook berekend met percentielen inclusief de mediaan. Figuur 16 geeft correcties voor de hoogste en de laagste percentielen. Door te rekenen inclusief de mediaan worden de correcties groter. Bij inclusieve percentielen worden de extremen dichter naar de mediaan getrokken.



Figuur 16 Effect van verschillende methoden om percentielen te berekenen op de correcties voor het hoogste en laagste percentiel van Tx in De Bilt. Links de percentielen berekend exclusief de mediaan, rechts de percentielen berekend inclusief de mediaan.

We hebben het effect van de span op de correcties diepgaander onderzocht. Voor een aantal varianten van de span voor de twee smoothings hebben we de Tx-reeks van De Bilt gecorrigeerd, de tropische dagen geteld en het percentage daarvan berekend dat vóór 1951 valt. Dat is gedaan met percentielen die zijn berekend exclusief de mediaan en inclusief de mediaan. Figuur 17 laat het resultaat zien. Hoe lager de span van de smoothing, hoe minder tropische dagen voor 1951 geteld worden. Bij percentielen inclusief de mediaan zijn de verschillen het grootst. Het laagste dat we gevonden hebben is 29,2% (de laagste span in beide smoothings bij percentielen inclusief de mediaan). Het getal 26%, waar de gehomogeniseerde reeks van het KNMI toe leidt, hebben we niet kunnen reproduceren.



Figuur 17 Effect van span van de smoothing op percentage tropische dagen vóór 1950. Percentielberekening exclusief mediaan (links) en inclusief mediaan (rechts).

Wij zien geen reden waarom De Bilt met het aantal tropische dagen van 1906-1951 sterk zou moeten afwijken van de andere Nederlandse hoofdstations. Daarom constateren we dat bij onze werkwijze voor De Bilt een span van 0,75 optimaal is met gebruik van percentielen exclusief de mediaan¹¹. Daarmee keert de zomer van 1947 terug in de top 3 van grote zomers van de 20^{ste} eeuw en houdt die van 1911 zijn gedeelde 4^e-6^e plaats.

Effect van de homogenisatie op de tropische, zomerse dagen en hittegolven en op ijsdagen

In de figuren 53 en 54 van bijlage 2 zijn de aantallen tropische dagen resp. de zomerse dagen van 1901-2018 weergegeven volgens drie reeksen: de ongehomogeniseerde reeks van De Bilt, de homogenisatie van het KNMI en onze reconstructie van de homogenisatie. In figuur 55 van bijlage 2 zijn de hittegolven weergegeven volgens de drie reeksen. In alle gevallen is de conclusie dat door de homogenisatie van het KNMI de zomerse en tropische dagen voor 1951 te ver zijn teruggebracht. Ook de door het KNMI geclaimde reductie van het aantal hittegolven voor 1951 houdt geen stand: volgens het KNMI zouden slechts 24% van de hittegolven voor 1951 hebben plaatsgevonden, tegen 37% volgens onze herziening.

Op 21 augustus 2018 meldden de media dat volgens het KNMI het recordaantal zomerse dagen (51) op die dag zou zijn overschreden. Volgens onze herziening van de gehomogeniseerde reeks stond toen het recordaantal zomerse dagen nog steeds op naam van 1947: 54 en zonder homogenisatie 64. Uiteindelijk kwam 2018 op 13 oktober uit op 55 zomerse dagen, waarmee het record van 1947 volgens onze reconstructie van de homogenisatie nipt is overschreden, maar nog altijd 9 dagen onder de werkelijke metingen blijft.

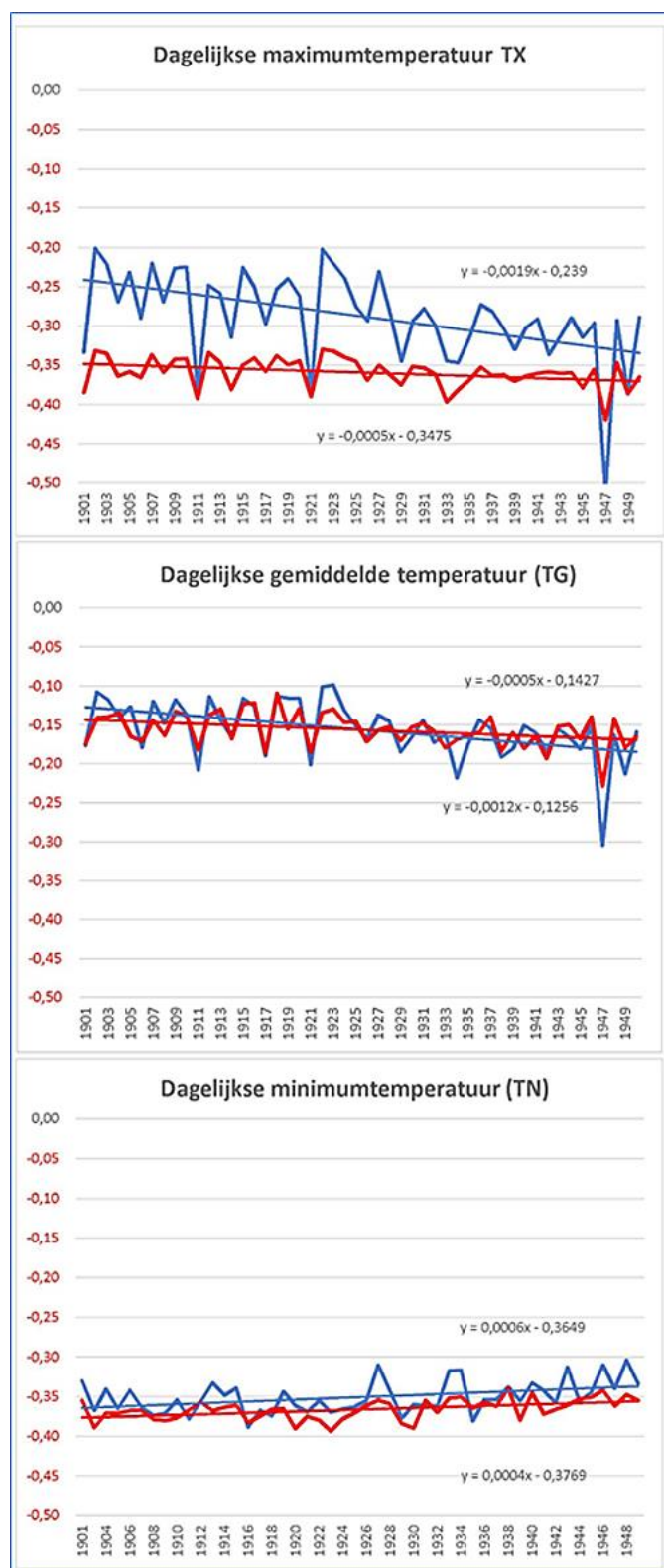
Het vermoeden van Dijkstra (2017) dat bij de homogenisatie in De Bilt sprake is van overcorrectie wordt door deze reconstructie bevestigd. Het effect van de pagodehut, die tot 1951 werd gebruikt, en het effect van de verplaatsing in 1951 zijn waarschijnlijk overschat.

Figuur 56 in bijlage 2 laat het effect zien van de drie verschillende homogenisaties op het aantal ijsdagen. De trendlijnen zijn in deze figuur berekend met een span van 11 punten uit 118 (0,09), om een mogelijke relatie met de zonnecyclus zichtbaar te maken. De verschillen zijn op het eerste gezicht niet groot. Het KNMI telt vóór 1951 3 ijsdagen minder dan zonder homogenisatie. Volgens onze herziening worden vóór 1951 50 ijsdagen meer geteld. Dit is 9% meer dan het KNMI vindt.

¹¹ Bij percentielen inclusief de mediaan worden de extremen naar de mediaan toe getrokken. Wij zien geen reden dat dit gewenst zou zijn, en kiezen daarom voor percentielen exclusief de mediaan.

Daggemiddelde temperatuur en dagelijkse minimumtemperaturen

Figuur 18 geeft de gemiddelde correcties per jaar voor Tx, Tg en Tn als gevolg van de homogenisatie van het KNMI en onze reconstructie. Op de Y-as is het verschil uitgezet tussen de ongehomogeniseerde, maar wel gestandaardiseerde data van De Bilt en de homogenisatie. De negatieve getallen betekenen dat de homogenisatie de oorspronkelijke waarden heeft verlaagd.



Figuur 18 Gemiddelde correcties per jaar als gevolg van de homogenisatie van Brandsma (blauw) en onze reconstructie van die homogenisatie (rood).

Bij Tx valt op dat onze reconstructie leidt tot gemiddeld grotere correcties dan de homogenisatie van het KNMI. Het verschil loopt van 1901-1950 af van 0,12 tot 0,05 graden. Maar voor 1947 komt het KNMI tot veel grotere correcties. Dat geldt in mindere mate ook voor 1911 en 1921. In warme zomers leidt de methode van het KNMI kennelijk tot extra grote correcties die niet uit onze reconstructie blijken. De fluctuaties in de correcties zijn bij het KNMI veel groter dan bij onze reconstructie. Waarom 1947 er bij het KNMI zo sterk uitspringt kunnen we niet verklaren. Ook het KNMI heeft daar desgevraagd geen verklaring voor gegeven.¹²

Bij Tg en Tn zijn de verschillen tussen de homogenisatie van het KNMI en onze reconstructie klein, gemiddeld over de periode 1901-1950 een paar honderdste graden. Wel zijn de fluctuaties bij het KNMI wat groter dan bij ons. Bij Tg valt vooral 1947 sterk op doordat de homogenisatie van het KNMI voor dat jaar een veel hogere correctie oplevert. Dit hangt samen met de nog veel grotere correctie bij Tx.

De trend in de correcties in de periode 1901-1950 is bij het KNMI voor Tx sterk dalend en bij Tn licht stijgend. Voor Tg is de trend licht dalend. De sterke daling van de correcties bij Tx is in onze reconstructie niet aanwezig.

Enige klimatologische constatering:

Uit de figuren 53 en 54 in bijlage 2 kan worden geconcludeerd, dat het aantal zomerse dagen ($T_x \geq 25$) in De Bilt sneller stijgt dan het aantal tropische dagen ($T_x \geq 30$). In het aantal tropische dagen zijn golfbewegingen te zien, terwijl het aantal zomerse dagen een tamelijk constante stijging vertoont vanaf 1960. Voor 1950 was het tamelijk constant en tussen 1950 en 1960 trad een duidelijke daling van het aantal zomerse dagen op.

Beweringen die vaak in de media worden gedaan (ook door KNMI en PBL) dat het aantal zeer hete dagen in de toekomst sterker stijgt dan gemiddeld worden door de historische waarnemingen niet ondersteund. In werkelijkheid is de maximale jaartemperatuur sinds 1947 niet gestegen. Alleen dankzij de homogenisatie kon het KNMI schrijven dat het op 27-7-2018 0,1 graad warmer werd dan op 27-6-1947 (35,7 versus 35,6). De werkelijk gemeten temperatuur op 27-6-1947 was 36,8. Volgens onze herziene homogenisatie heeft 1947 met 36,2 nog steeds het absolute record voor De Bilt, dat dus al 71 jaar staat.

¹² Correspondentie tussen F.Y. Dijkstra en A.M.G. Klein Tank, januari 2017

Conclusies

1. De uitkomst van de homogenisatie met de percentiel matching methode, zoals door het KNMI (Brandsma 2016a) uitgevoerd, is sterk afhankelijk van de gebruikte span bij de smoothing en van de manier waarop de percentielen worden berekend. Zonder smoothing is de ruis in de data zo groot, dat geen conclusies per percentiel en per maand zijn te trekken.
2. Dit herhalingsonderzoek bevestigt niet dat bij homogenisatie het aantal tropische dagen en het aantal hittegolven voor 1951 in De Bilt zo sterk daalt als het KNMI claimt. Met de door ons aanbevolen span van de smoothing blijven de data van De Bilt in overeenstemming met die in andere Nederlandse stations.
3. De statistische onzekerheden in de percentiel matching methode bij de door het KNMI gebruikte datareeksen van twee ver van elkaar verwijderde vergelijkingsstations zijn dermate groot dat het aanbeveling verdient andere vergelijkingsstations bij de homogenisatie te betrekken.
4. Zolang er geen algemeen geaccepteerde methode is gevonden om tot gehomogeniseerde daggegevens voor de Bilt te komen dient het KNMI zich in de media te onthouden van uitspraken over klimatologische records die in strijd zijn met de oude, niet gehomogeniseerde reeksen. Het is beter geen homogenisatie te hebben dan een verkeerde homogenisatie.

5 **Reconstructie van de homogenisatie op basis van de aanvullende informatie van het KNMI**

Het vorige hoofdstuk laat zien dat reconstructie van de homogenisatie op basis van de informatie uit het Technisch Rapport van het KNMI (Brandsma 2016a) niet tot dezelfde resultaten leidt. De jaargemiddelden van de gemiddelde en de minimumtemperaturen bleken vrij nauwkeurig reproduceerbaar te zijn, maar bij de maximumtemperaturen is er in de jaargemiddelden een substantieel verschil, dat niet in de tijd constant is en voor de jaren met hete zomers het grootst is. Met het aantal tropische dagen in de periode 1901-1950 wijkt onze reconstructie van de homogenisatie aanzienlijk af van die van het KNMI. Dat leidde bij ons tot het vermoeden, dat door het KNMI nog andere bewerkingen van de data zijn toegepast, dan in het Technisch Rapport (Brandsma 2016a) zijn vermeld.

Hoor en wederhoor

We hebben geprobeerd hoor en wederhoor toe te passen en hebben het KNMI vroegtijdig in kennis gesteld van onze belangrijkste bevindingen. Ons voorstel om daarover een gesprek te hebben werd door het KNMI afgewezen. Wel wilde het KNMI schriftelijke vragen beantwoorden. In bijlage 3 staan onze vragen van 25 oktober 2018 en de op 9 november 2018 ontvangen antwoorden. Voor dit rapport zijn de volgende punten relevant:

- (1) Het KNMI blijkt een defaultmethode in de R-software te hebben gebruikt voor het bepalen van de percentielen. In hoeverre die methode afwijkt van de standaardmethoden van Excel (exclusief of inclusief mediaan) kunnen wij niet beoordelen. Het KNMI schrijft dat er geen reden was om een andere keuze te maken.
Wij hebben vastgesteld dat de keuze van de methode voor het bepalen van de percentielen invloed heeft op het aantal tropische dagen dat men vindt. Er kan dus wel degelijk een reden zijn om af te wijken van de defaultmethode van de R-software.
- (2) Op welke wijze de span van de smoothing is bepaald legt het KNMI in het technisch rapport niet uit, ook niet hoe men daarbij '*expert opinion*' heeft toegepast. Wel stelt het KNMI dat het effect van een kleine verandering van de span '*op het oog*' klein is.
Wij hebben vastgesteld dat het effect van de span op het aantal tropische dagen groot is en betreuren het, dat het KNMI hier niet verder op in is gegaan.
- (3) In het antwoord op onze derde vraag stelt het KNMI dat er geen aanvullende aanpassingen zijn gedaan op de data, die kunnen verklaren waarom wij niet dezelfde resultaten vinden. Vervolgens meldt het KNMI dat niet de data van 1946-1949 zijn vergeleken met die van 1952-1955, zoals in het rapport staat, maar de data van januari 1946- augustus 1950 met die van januari 1952 – augustus 1956. Dit is gedaan om zoveel mogelijk data van vóór de overgang op de Stevensonhut mee te nemen. Bovendien zijn de percentielen voor elke maand gebaseerd op lopende gemiddelden van drie maanden.
- (4) Over de statistische fout in de berekende correcties meldt het KNMI dat die moeilijk is te bepalen, maar dat men rekening houdt met een fout van een paar tiende graad op de grootste correcties.
Naar onze mening is deze opvatting niet onderbouwd. Wij vinden door variatie van percentielberekening, de span en de vergelijkingsstations veel grotere verschillen ten opzichte van de homogenisatie van het KNMI dan een paar tiende graad. We vinden bovendien de grootste correcties voor andere percentielen (0,80 – 0,85) dan het KNMI (0,95).

Het KNMI gaat nergens in op de gevolgen die de gemaakte keuzen hebben voor de uitkomsten van de homogenisatie. Het is ons dus onbekend of het KNMI de gevoeligheid van de methode voor die keuzen heeft onderzocht. Het meest opvallend zijn twee dingen die we hierna zullen onderzoeken:

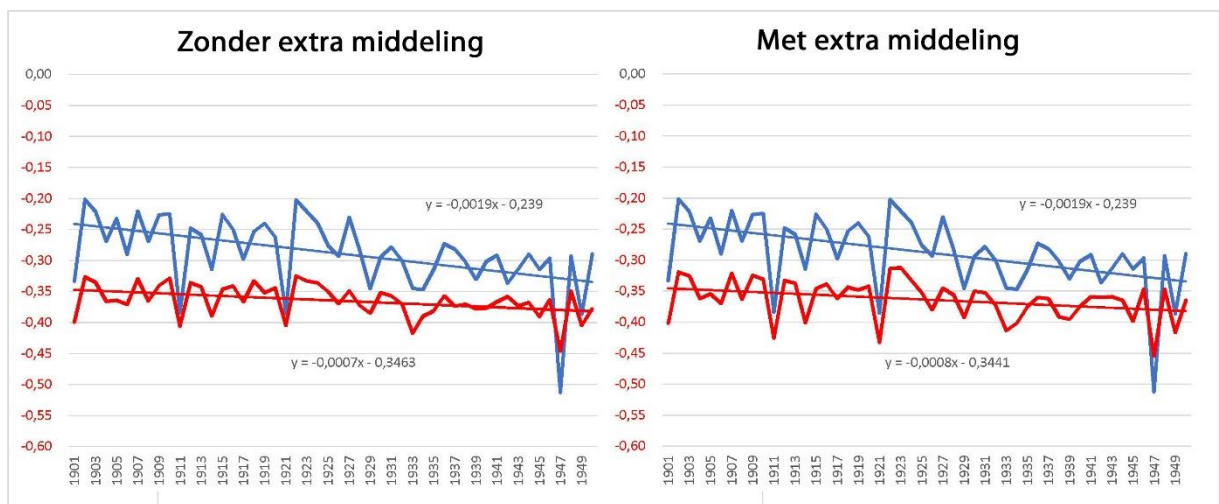
- percentielen gebaseerd op lopende driemaands gemiddelden en
- vergelijkingsperioden van 56 maanden in plaats van 48 maanden.

Driemaands gemiddelden

Door de percentielen te baseren op lopende gemiddelden van drie maanden wordt een derde afvlakking geïntroduceerd: eerst een driemaands gemiddelde over de percentielen, dan een smoothing over de verschillen tussen de percentielen in de oude en de nieuwe situatie, en vervolgens een smoothing van die uitkomsten over de maanden.

Bij een span van 0,6 berekent men een smoothing over de maanden op basis van 7 punten ($7/12 = \text{ca } 0,6$). Door uit te gaan van driemaands gemiddelden neemt men aan weerszijden van die 7 punten een punt extra, waardoor men in feite 9 punten meerekent. Ongeveer hetzelfde effect is te bereiken door bij de smoothing over de maanden uit te gaan van een span van 0,75.

We hebben geprobeerd de homogenisatie te reconstrueren op basis van lopende gemiddelden van de percentielen over drie maanden. Figuur 19 laat het effect zien op de jaargemiddelden van de correcties van Tx.



Figuur 19 Gemiddelde correctie Tx per jaar, 1901-1950, volgens de homogenisatie van het KNMI (blauw) en volgens onze reconstructie (rood). Links: de percentielen in onze reconstructie gebaseerd op percentielen per maand. Rechts: de percentielen in onze reconstructie gebaseerd op driemaands gemiddelden.

Het effect op de jaargemiddelden is gering, gemiddeld zijn de jaargemiddelden na de extra middeling $0,02\text{ }^{\circ}\text{C}$ hoger, maar in de jaren met warme zomers gemiddeld $0,04\text{ }^{\circ}\text{C}$ lager. Het zijn slechts kleine verschillen, maar toch lijken ze symptomatisch voor de keuzen die het KNMI heeft gemaakt in zijn homogenisatie: jaren met veel warme dagen worden extra naar beneden gecorrigeerd.

Figuur 57 in bijlage 2 laat de correcties per maand en percentiel zien, berekend met een lopend driemaands gemiddelde van exclusieve percentielen en een span van 0,6 bij de smoothings. Deze figuur is vergelijkbaar met figuur 50 in bijlage 2. Het belangrijkste verschil tussen figuur 57 en figuur 50 is de spreiding van de meetpunten: die liggen door de extra middeling over 3 maanden, dicht bij de gesmoothde lijnen.

In tabel 3 wordt het effect van de verschillende smoothings op de aantallen tropische dagen getoond. De middeling van de percentielen over drie maanden levert voor de periode 1906-1950 vier

extra tropische dagen op, bij een span van 0,6. Instellen van de span op 0,75 levert nogmaals 4 extra tropische dagen op. Het geringe aantal tropische dagen in de homogenisatie van het KNMI kan dus niet verklaard worden uit de middeling van de percentielen over 3 maanden. Gezien het geringe effect van deze extra middeling zien we hier verder van af, en gebruiken een span van 0,75.

Volgens KNMI		Reconstructie homogenisatie volgens dit rapport		
Originele metingen	Gehomogeniseerd	Span 0,6	Span 0,75	Span 0,6, extra middeling
164	76	114	122	118

Tabel 3 Aantal tropische dagen 1901-1950 volgens originele metingen en 4 varianten homogenisatie

Vergelijkingsperiode 56 maanden

Zoals hiervoor gemeld, heeft het KNMI in afwijking van wat in het technisch rapport staat, niet een periode van 48 maanden gebruikt voor de overlappende periode, maar van 56 maanden, 1 januari 1946 t/m augustus 1950, om zoveel mogelijk metingen van voor de transitie van de meetapparatuur mee te nemen. Dit heeft een aantal bezwaren:

- (1) Hierdoor gaat de vergelijking niet over een heel aantal jaren. De maanden januari-augustus wegen daardoor zwaarder dan de maanden september-december.
- (2) Er wordt in de periode 1952-1956 een extra strenge winter meegerekend en een extra koele zomer. Wat dat voor gevolgen heeft voor de matching van de percentielen is niet eenvoudig te voorspellen. Uit diverse door ons uitgevoerde reconstructies met andere vergelijkingsstations in Nederland en Duitsland, en andere vergelijkingsperiodes (4 jaar, of 15 jaar) blijken onverwachte en soms grote effecten op te treden (zie hoofdstuk 6 en 7).
- (3) Het bezwaar dat de twee vergelijkingsperiodes meteorologisch onvergelijkbaar zijn wordt door de toevoeging van 8 maanden aan beide perioden niet opgeheven:
 - 1946-1950 telde 1 extreem strenge winter, 1 extreem warme zomer en 1 warme zomer, en daar komen door toevoeging van 8 maanden 1 normale zomer en winter bij,
 - 1952-1956 telt na toevoeging van 8 maanden 1 koude winter, 1 strenge winter en alleen koele zomers.
 Die vergelijkingsperiodes blijven onvergelijkbaar. Bovendien sloeg omstreeks 1950 de AMO-index (Atlantic Multidecadal Oscillation) in de zomermaanden om, wat de inleiding was tot een lange periode met veel slechte zomers die tot 1975 aanhield.

Wij hebben de reconstructie van de homogenisatie herhaald met toevoeging van de door het KNMI genoemde 8 maanden. We hebben de homogenisatie uitgevoerd met exclusieve en inclusieve percentielen en een span van 0,75 voor de beide smoothings.

Figuur 58 in bijlage 2 laat de jaargemiddelden van de correcties van T_x , T_g en T_n zien met gebruikmaking van exclusieve en inclusieve percentielen. Bij vergelijking met figuur 18 in hoofdstuk 4 valt het op, dat onze reconstructie voor T_x niet meer tot grotere correcties leidt dan de homogenisatie van het KNMI. Het verschil dat we in hoofdstuk 4 vonden kan dus blijkbaar verklaard worden uit de toevoeging van 8 maanden aan de vergelijkingsperiode.

Toch blijven er grote verschillen tussen onze reconstructie en de homogenisatie van het KNMI. De grote fluctuaties van de correcties per jaar, die de KNMI-homogenisatie laat zien, vinden wij niet. Dit geldt in het bijzonder voor de grote correcties in jaren met hete zomers (1911, 1921 en 1947) die het KNMI vindt. Voor T_n blijkt er opnieuw weinig verschil tussen onze reconstructie en de homogenisatie van het KNMI. Wel zijn de fluctuaties per jaar bij de KNMI-homogenisatie wat groter dan bij onze reconstructie. T_g neemt een middenpositie in, wat logisch lijkt. Onze reconstructies met exclusieve percentielen liggen dicht bij de KNMI-homogenisatie dan die met inclusieve percentielen.

We kunnen constateren, dat door toevoeging van de 8 extra maanden in 1950 en 1956 onze reconstructie dichter bij de homogenisatie van het KNMI komt, maar dat de verschillen in jaren met hete zomers blijven bestaan. Dat blijkt ook uit de aantallen tropische dagen, zoals tabel 4 laat zien.

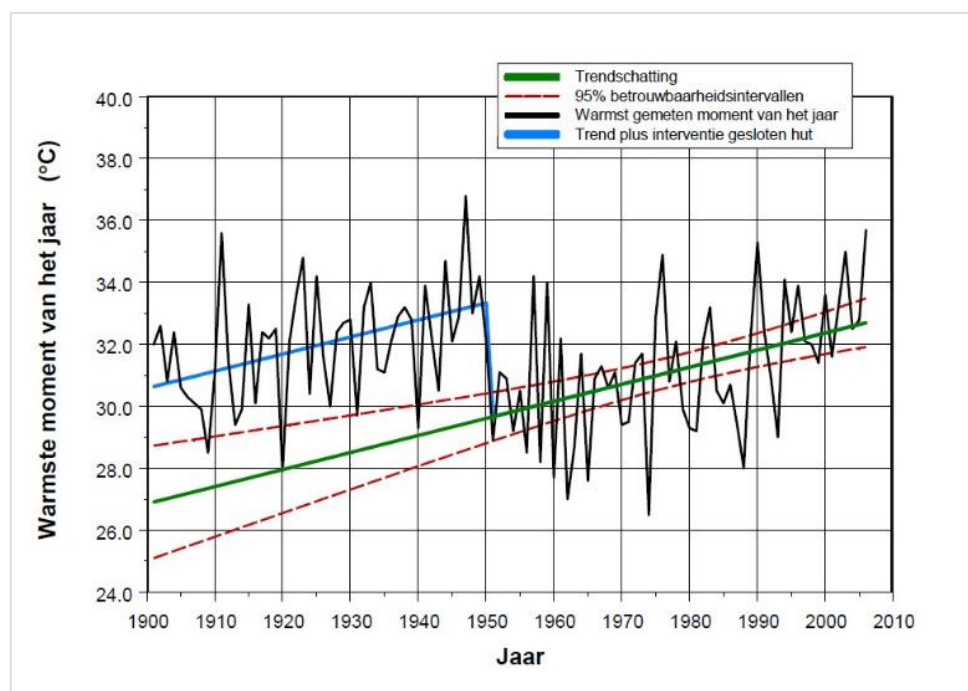
Volgens KNMI		Reconstructie homogenisatie volgens dit rapport			
Origineel	Homogenisatie	Span 0,6 excl.	Span 0,75 excl.	Span 0,60 incl	Span 0,75 incl
164	76	104	113	102	109

Tabel 4 Aantal tropische dagen 1901-1950 volgens originele metingen en 4 varianten homogenisatie met vergelijkingsperioden van 56 maanden

In alle varianten van onze reconstructie vinden wij veel meer tropische dagen vóór 1950 dan het KNMI: onze uitkomsten variëren van 102-122, tegen 76 volgens het KNMI.

Brandsma (2017) stelt in zijn weerwoord op Dijkstra (2017) dat het “*onrealistisch is om van een homogenisatiemethode te verwachten dat ze de allerhoogste extremen (zoals het aantal tropische dagen per jaar) perfect corrigeert.*” Dat kunnen we onderschrijven, maar in dat licht is het verbazingwekkend dat het KNMI en het PBL in de zomer van 2018 juist veel in de publiciteit zijn getreden met de sterke toename van hittegolven, tropische en zomerse dagen in de laatste 50 jaar, vergeleken met de periode daarvoor, meestal zonder te vermelden dat de cijfers van voor 1950 gecorrigeerd zijn. Aangezien het KNMI zelf toegeeft, dat het onrealistisch is om te verwachten dat deze correcties perfect zijn, zou op dit punt grote terughoudendheid verwacht mogen worden.

Maximale jaartemperatuur



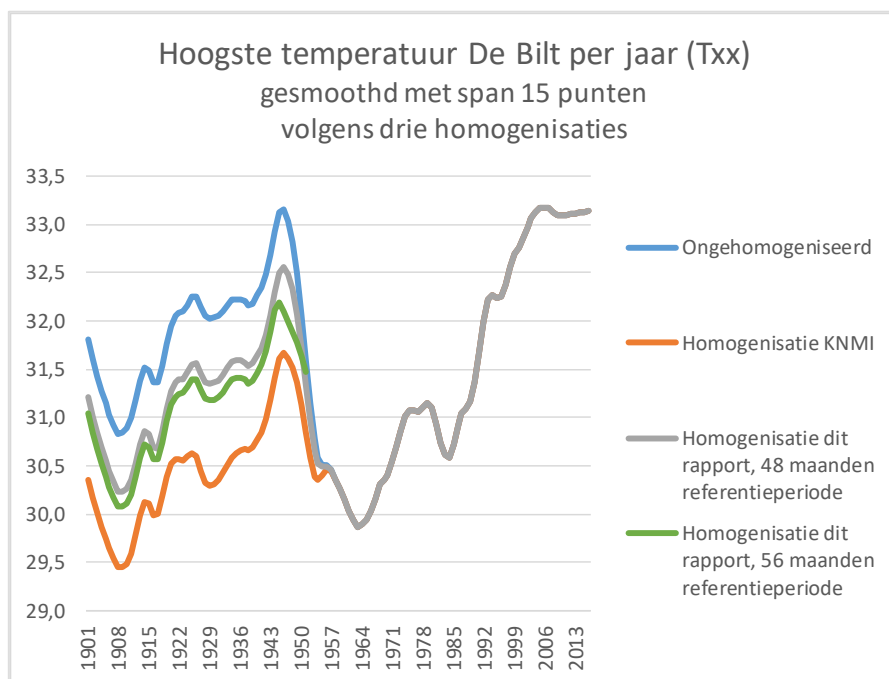
Figuur 20 Ontwikkeling van de maximale jaartemperatuur volgens Visser (2007).

De ontwikkeling van de maximale temperatuur in Nederland (c.q. De Bilt) is iconisch geworden voor de hittegolven van de toekomst. Sommige auteurs (o.a. Mureau c.s., 2013) hebben de vraag gesteld ‘*wanneer bereiken we in Nederland de 40 graden?*’ Visser (2007) deed een poging de trend in de maximale jaartemperaturen van 1901-1951 te corrigeren voor de (vermeende) te hoge metingen in de pagodehut in De Bilt. Zie figuur 20.

Door de hele temperatuursprong rond 1950 toe te schrijven aan de pagodehut, kwam Visser tot een correctie van 3,5 °C voor de maximale jaartemperaturen in de periode 1901-1950. Hij ging er daarbij impliciet van uit, dat in Txx een lineaire trend zit over een periode van meer dan 100 jaar. Uit onze

analyses volgt dat dit niet juist is: in werkelijkheid is er sprake van een golfbeweging. De temperatuursprong rond 1950 kan slechts ten dele worden toegeschreven aan instrumentele veranderingen en is grotendeels klimatologisch van aard, zoals in hoofdstuk 6 wordt aangetoond.

Figuur 21 laat het effect zien van de verschillende homogenisaties op de temperatuursprong. Zonder homogenisatie is er een sprong van 3 °C. Door de homogenisatie van het KNMI wordt de sprong gehalveerd tot 1,5 °C. Met onze reconstructies van de homogenisatie is de temperatuursprong 2-2,3 °C.



Figuur 21 Trend in de hoogste jaartemperatuur in de Bilt volgens originele metingen KNMI en drie homogenisaties.

Conclusies

1. De door het KNMI toegepaste middeling van de percentielen over drie maanden heeft weinig invloed op de uitkomst van de homogenisatie: gemiddeld komt er een 0,02 °C hogere Tx uit, maar in jaren met hete zomers wordt Tx gemiddeld 0,04 °C lager.
2. Het verlengen van de vergelijkingsperiode van 48 tot 56 maanden heeft een duidelijke invloed op de uitkomst van de homogenisatie. Tx wordt daardoor gemiddeld 0,1 °C hoger, Tg iets minder en op Tn heeft het geen invloed.
3. Onze reconstructie van de homogenisatie komt door het verlengen van de vergelijkingsperiode van 48 tot 56 maanden dicht bij de homogenisatie van het KNMI, maar de verschillen tussen de jaren blijven bij de KNMI-homogenisatie groter dan bij onze reconstructie, in het bijzonder in jaren met hete zomers.
4. Het feit dat onze reconstructie dicht bij de homogenisatie van het KNMI komt betekent niet dat de vergelijkingsperiode van 56 maanden beter zou zijn dan 48 maanden. De eerder genoemde bezwaren tegen de vergelijkingsperiode van 56 maanden blijven van kracht.
5. Het aantal van 76 tropische dagen van 1901-1950 waar het KNMI op uitkomt wordt door onze reconstructie niet bevestigd. In alle varianten van de reconstructie komen wij op 102-122 tropische dagen. Het is evident, dat daardoor ook meer hittegolven vóór 1950 berekend worden.
6. Het feit, dat schijnbaar arbitraire keuzen in de methodiek (percentielberekening, smoothing, lengte van de vergelijkingsperiode) veel invloed hebben op de uitkomsten van de homogenisatie, laat zien dat deze methode niet geschikt is om daggegevens te corrigeren met de pretentie van een nauwkeurigheid van 0,1 graad. Dat is statistisch onverantwoord.
7. Gehomogeniseerde dagtemperaturen dienen te worden afgerond op tenminste hele graden. Dan zou het KNMI van 1901-1950 ook 104 tropische dagen tellen, omdat 29,5 °C statistisch niet te onderscheiden is van de tropische waarde 30 °C. Het aantal hittegolven vóór 1950 zou dan niet 7 zijn maar 10 tot 12.

6 De temperatuursprong en de correctie van het KNMI

Brandsma (2016b) schrijft over een "*.. neerwaartse sprong in Tx rond 1950 in De Bilt..*", waarmee hij kennelijk doelt op een trendbreuk in Tx als gevolg van de verandering van meethut en verplaatsing (hier verder als "instrumentele breuk" betiteld).

Zoals we verder zullen zien doet zich een neerwaartse sprong rond 1950 voor in een groot deel van NW-Europa tussen 1949 en 1956. Die sprong is van klimatologische aard. De veronderstelde instrumentele breuk in 1950/51 valt dus binnen deze klimatologische neergang. Als we corrigeren voor de gemeenschappelijke klimatologische sprong dan moet er een trendbreuk detecteerbaar zijn in De Bilt die andere stations niet vertonen.

Een trendbreuk of inhomogeniteit in de temperatuur wordt gedetecteerd door vergelijking met een referentiereeks van naburige stations. Essentieel bij het detecteren van een breuk is de gebruikte referentiereeks. Gebruikt men een andere set van referentiestations dan zal het resultaat, zoals het tijdstip en de grootte van de breuk, waarschijnlijk anders uitpakken. De referentiereeks bestaat vaak uit het gemiddelde van de waarden van homogene nabijgelegen stations. Sprongen in een tijdreeks worden gedetecteerd op grond van het verschil met de homogene referentiereeks als de trend van de verschilreeks (het residu) een breuk vertoont op een bepaald tijdstip. De grootte van de breuk is het verschil tussen de gemiddelde waarden van het residu voor en na de breuk.

Van der Schrier et al (2011a) gebruiken voor de samenstelling van de CNT (Centraal Nederlandse Temperatuur) een referentiereeks van een aantal langjarige Nederlandse temperatuurreksen om te bepalen of er in de temperatuurreksen van De Bilt sprake is van inhomogeniteiten. De homogeniteitstesten tonen breuken rond 1918-1920, 1950, 1968-1969 en 1976. Voor de breuk rond 1918-1920 is voor de samenstelling van de CNT niet gecorrigeerd en evenmin voor de breuk rond 1976; die breuken waren niet significant.

Over de breuk rond 1950 schrijven Van der Schrier et al in Meteorologica (2011b): "*..de breuk rond 1950 is duidelijk verwant aan een reeks aanpassingen die plaatsvonden in het begin van de 50-er jaren. Op 17 mei 1950 werd de Pagode vervangen door een Stevenson hut. Op 16 september 1950 is de opstelling in westelijke richting verplaatst, om op 27 augustus 1951 300 m naar het zuiden te verhuizen.*" In hetzelfde artikel lezen we dat niet alleen gecorrigeerd werd voor de breuk rond 1950 maar ook voor een verlaging van de Stevensonhut in 1961 en de vervanging van de Stevensonhut door een schotelhutje in 1993. Wij konden niet achterhalen of en welke referentiestations daarbij werden gebruikt. Dat is overigens een euvel van veel KNMI-rapporten: wat men precies heeft gedaan wordt niet vermeld. De normale praktijk is, dat men parallelle metingen doet op dezelfde locatie met het oude instrument en het nieuwe instrument. Daar zijn geen referentiestations bij nodig.

In een eerder rapport van Van Ulden (2009) lezen we dat er twee breuken zijn gedetecteerd, niet vier, te weten in 1950 en 1976. Daarbij werden Den Helder, Groningen/Eelde, Oudenbosch en Maastricht/Beek gebruikt als referentiestations. De grootte van de trendbreuk in 1950 werd bepaald als het verschil van de gemiddelden tussen 1935-1949 en 1951-1965.

In dit hoofdstuk gaat het over de breuk rond 1950. Het KNMI (Brandsma 2016a) heeft voor de correctie van deze breuk slechts één reeks gebruikt als referentie: Eelde. Afgezien van het feit dat het gebruik van slechts één station als referentie statistisch gezien problematisch kan zijn is dat nog om een andere reden opvallend. In het Scientific Report (Van Ulden 2009) wordt gesteld dat Eelde te sterk afwijkt van de meer centraal gelegen stations en daarom niet werd meegenomen in de CNT. Waarom het KNMI desondanks Eelde toch als referentie nam voor de homogenisatie van De Bilt is onduidelijk.

Duitse stations als referentie voor de homogenisatie van De Bilt

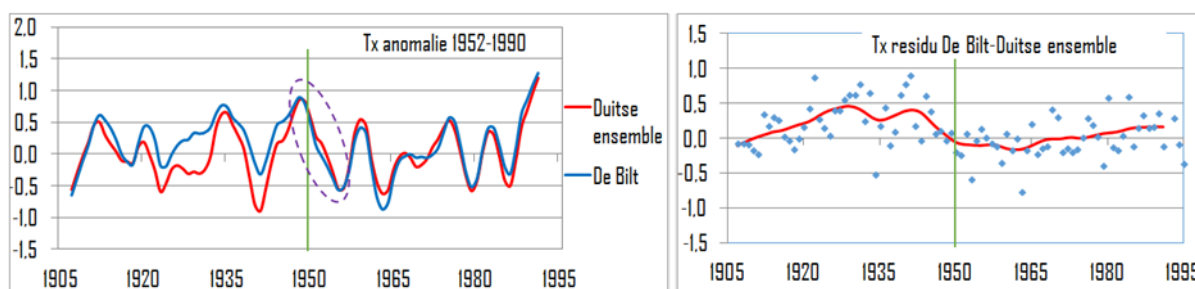
Voor het KNMI is de aanleiding om de dagelijkse temperaturen van De Bilt tussen 1901 en 1951 te corrigeren een breuk in de reeks als gevolg van hutwisseling en locatieverandering. Om te onderzoeken of er rond 1950 inderdaad sprake is van een instrumentele breuk hebben we een referentieset van meetstations samengesteld die geen correctie of homogenisatie hebben ondergaan. Om te vermijden dat we daarvoor de al of niet gehomogeniseerde stations in Nederland gebruiken (die bovendien onderling correleren omdat bij homogenisatie vaak dezelfde referentiesets worden gebruikt), hebben wij gekozen voor een volledig onafhankelijke referentieset die we het *Duits ensemble* noemen. Aan die eis voldeden zeven meetstations in Duitsland:

1. Aachen:	1930-2010
2. Münster:	1901-1991 (1945 en 1946 afwezig)
3. Solingen:	1936-2002
4. Geisenheim:	1935-2015
5. Jena Sternw:	1901-2015
6. Lindenberg:	1907-2015
7. Potsdam:	1901-2015

Tabel 5 De 7 stations van het Duitse ensemble

De stations zijn geselecteerd uit de verzameling van 1246 Duitse stations in de ECAD databank¹³. Veel stations vielen af vanwege het ontbreken van teveel data (waarschijnlijk als gevolg van oorlogshandelingen), veel andere reeksen beginnen pas na 1950. De gebruikte etmaalreeksen van Tx zijn non-blended en voor zover we hebben kunnen nagaan ongehomogeniseerd. In deze studie zijn alleen de jaarlijks gemiddelden van de dagelijkse maximumtemperaturen gebruikt voor analyse. De 2 ontbrekende jaren 1945 en 1946 voor Münster werden opgevuld met die van het meest nabijgelegen station Winterswijk met een kleine correctie. De methode waarmee het Duitse ensemble werd samengesteld is beschreven in Bijlage 1.

In Figuur 22 (links) vergelijken we de jaargemiddelde Tx van de ongehomogeniseerde data van De Bilt met Tx van het Duitse ensemble voor de periode 1907-1991. De Tx reeksen zijn uitgezet als anomalie ten opzichte van het gemiddelde van 1952-1990 en vervolgens gladgestreken met een Loess filter:

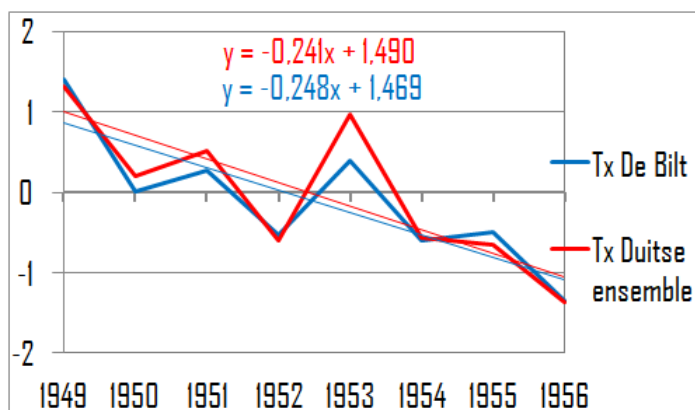


Figuur 22 Links: Tx (°C anomalie van de jaargemiddelde Tx in het interval 1952-1990), Loess-gefilterd met span 8 jaar. De streepjes-ellips markeert de klimatologische neergang en de groene verticale lijn de instrumentele veranderingen rond 1950. Rechts: het verschil (residu) tussen de jaarlijkse Tx in De Bilt en de jaarlijkse Tx van het Duitse ensemble (°C anomalie 1952-1990). De rode lijn is een Loess filter met span 15 jaar door de punten.

Binnen de streepjes-ellips (figuur 22 links) is het interval te zien van de klimatologische neergang. Vóór 1950 is Tx in De Bilt relatief hoger dan het Duitse ensemble. Daarna loopt De Bilt in de pas met het Duitse ensemble wat erop wijst dat het ensemble representatief is. De klimatologische neergang

¹³ <https://www.ecad.eu/>

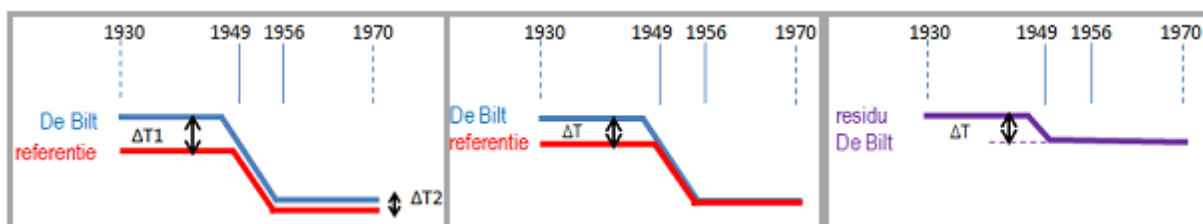
tussen 1949 en 1956 verloopt in De Bilt niet anders dan bij het Duitse ensemble ondanks de door het KNMI veronderstelde instrumentele breuk rond 1950:



Figuur 23 Tx (°C anomalie van 1952-1990) in het interval van de klimatologische neergang (1949-1956). De Bilt en het Duitse ensemble tonen kwalitatief en kwantitatief dezelfde neergang.

De neergang doet zich niet alleen voor in De Bilt maar ook bij Duitse en Nederlandse stations. De hutwisseling in 1950 en de herlocatie in 1951 veroorzaken geen detecteerbare trendbreuk binnen dit interval: de Tx van de referentieset en De Bilt dalen met identieke hellingshoek. De conclusie is dat de "neerwaartse sprong" rond 1950 in een groot gebied zichtbaar is in temperatuurreksen en dus klimatologisch van aard is. Het effect van de instrumentele veranderingen is dus binnen deze klimatologische neergang niet detecteerbaar, maar figuur 22-rechts toont dat Tx in De Bilt in de jaren vóór 1950 (1920-1945) relatief hoger is. Deze trendbreuk zou veroorzaakt kunnen zijn door de genoemde instrumentele veranderingen. Waarom de trendbreuk in Tx niet precies rond 1950 plaatsvond hebben wij niet verder onderzocht maar verderop zullen we zien dat deze instrumentele breuk zich wél voordoet rond 1950 in Tx van de zomermaanden. In de zomermaanden is de klimatologische neergang overigens het markantst (zie figuur 32).

Methode



Figuur 24 Schematische representatie van de gevolgde methode. Links: Tx in De Bilt (zwart) tussen 1930 en 1970. Tussen 1949 en 1956 is de klimatologische neergang (zie figuur 23). Rood: Tx van het vergelijkingsstation/referentie. $\Delta T1$ = Tx referentie min Tx De Bilt over het vergelijkingsinterval 1930-1949 en $\Delta T2$ idem over 1952-1970. Midden: beide Tx reeksen uitgezet als anomalie van 1952-1970.. De Bilt is relatief warmer vóór 1950 ($\Delta T = \Delta T1 - \Delta T2$). Rechts: verschil tussen de Tx reeksen van het middelste plaatje. Dit is het residu zoals getoond in figuur 22-rechts. ΔT is de geschatte benodigde neerwaartse correctie voor de instrumentele breuk in De Bilt.

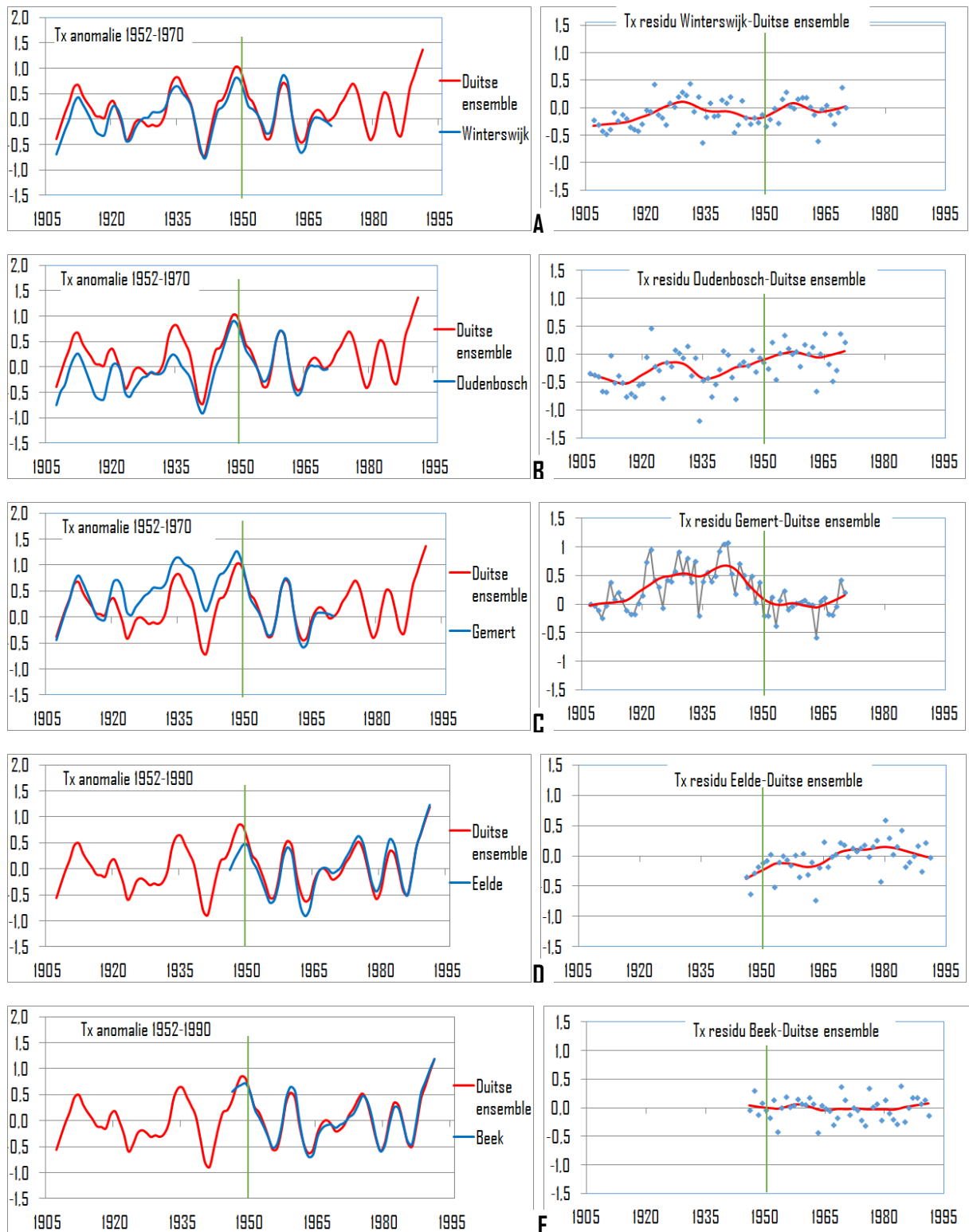
In figuur 24 is de gevolgde methode zeer schematisch weergegeven. We willen een schatting maken van de benodigde Tx correctie in De Bilt (ΔT) en die vergelijken met de Tx correctie van het KNMI.

De schatting van de benodigde Tx correctie vindt hier plaats door vergelijking van Tx De Bilt met Tx van een representatieve referentieset.

Een nieuwe referentieset voor de homogenisatie van De Bilt

Een aantal Duitse stations ligt wel wat ver van De Bilt en dat kan invloed hebben op de resultaten. Daarom gaan we met onverdachte Nederlandse stations en dichtbijgelegen Duitse stations een geschiktere referentieset samenstellen. Hiervoor werd al gerefereerd aan het feit dat het gebruik van slechts één referentiestation voor de homogenisatie van De Bilt uit statistisch oogpunt problematisch kan zijn. Het gebruik van één station wordt normaal gesproken beperkt tot een nabijgelegen station en uitgevoerd op basis van parallelmetingen gedurende een bepaalde periode. Parallelmetingen rond 1950 zijn bij De Bilt niet voorhanden en de temperatuurreeks van het nabijgelegen station Soesterberg begint pas in 1951.

Behalve Eelde had het KNMI bij de keuze van één of meer referentiestations ook andere (ongecorrigeerde) stations kunnen betrekken, te weten Winterswijk, Oudenbosch, Gemert, Beek en Eelde. We hebben deze stations vergeleken met het Duitse ensemble om te beoordelen welke van deze stations in aanmerking komen om te worden opgenomen in de nieuwe referentieset.



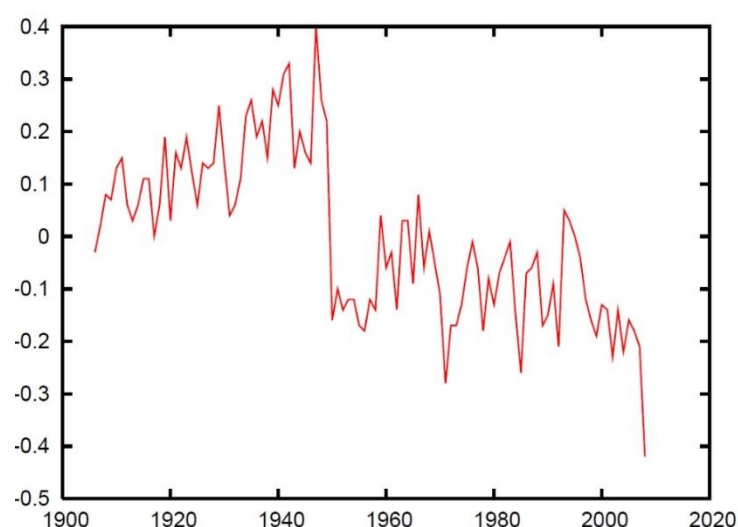
Figuur 25 Figuren links: Tx ($^{\circ}\text{C}$ anomalie van 1952-1990) van vergelijkingsstations en het Duitse ensemble, Loess-gefilterd met span 8 jaar. Figuren rechts: het verschil tussen de jaarlijkse Tx van de vergelijkingsstations en de jaarlijkse Tx van het Duitse ensemble. De rode lijn is een Loess filter met span 15 jaar door de punten.

Figuur 25 (linker grafieken) toont dat na 1950 de Tx'en van alle stations goed in de pas lopen met Tx van het Duitse ensemble. Oudenbosch (figuur 25B) en Gemert (figuur 25C) vertonen daarentegen grote systematische trendafwijkingen van het Duitse ensemble in de periode vóór 1950. De rechter grafieken in figuur 25 tonen de residuen waarmee trendafwijkingen zichtbaar worden gemaakt. Winterswijk (figuur 25A) vertoont geringe trendafwijkingen van het Duitse ensemble, afgezien van de neergang vóór 1920. In het Scientific Report (Van Ulden 2009) staat dat er kleine trendbreuken zijn in Winterswijk in 1940, 1950 en 1960 als gevolg van instrumentveranderingen. In figuur 25A-rechts zijn die trendbreuken slechts klein, maar kunnen wel relevant zijn als alleen de kleine 4-jaars intervallen rond 1950 (1946-1949 en 1952-1955) voor homogenisatie worden beschouwd. Voor grotere vergelijkingsintervallen zoals die in dit hoofdstuk worden gebruikt (1931-1949 en 1952-1970) is het verschil met het Duitse ensemble marginaal ($0,08^{\circ}\text{C}$).

Oudenbosch vertoont veel meer opwarming dan de Duitse referentiereeks ($0,36^{\circ}\text{C}$ van 1907-1970). Hoe deze opwarming lokaal door natuurlijke factoren kan zijn ontstaan is vooralsnog onbekend. Het KNMI heeft in Oudenbosch alleen een trendbreuk rond 1966 en 1984 gedetecteerd als gevolg van verplaatsing, waarbij Rotterdam, Deelen, Gemert en de gehomogeniseerde reeks van De Bilt als referentie werden gebruikt.

Gemert (figuur 25C) toont een opwarmende trend van 1907 tot 1940 waarna een sterke daling inzet. Metadata van Gemert vermelden een renovatie van het station in 1949. Het KNMI heeft een trendbreuk rond 1950 gedetecteerd met de referenties Oudenbosch, De Bilt en Winterswijk.

Ter vergelijking hier de trendbreuk van Gemert zoals gepubliceerd door Van der Schrier et al in *Climate of the Past* en *Meteorologica* (2011):



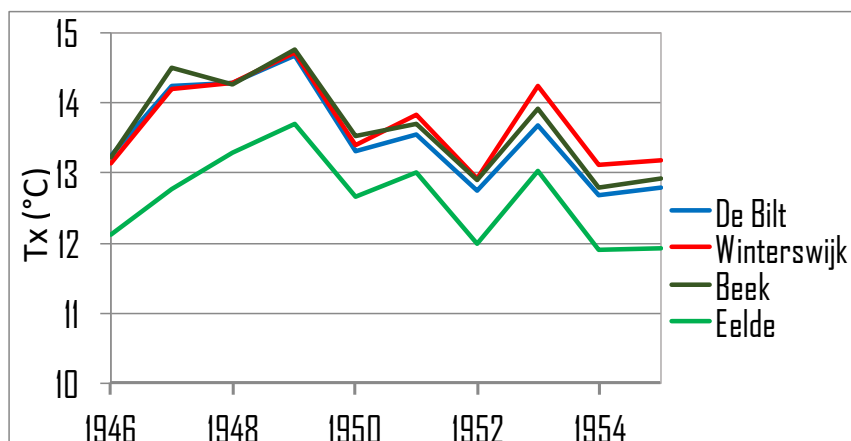
Figuur 26 Trendbreuk in de jaarlijks gemiddelde Tx van Gemert in vergelijking met een referentie, volgens Van der Schrier c.s. Vergelijk met Figuur 25C-rechts.

In figuur 26 zien we een zeer scherpe inhomogeniteit in 1950 die in figuur 25C-rechts afwezig is. In figuur 25C-rechts is de piek rond 1940, niet 1950, bovendien daalt Tx geleidelijk van 1940-1952 en niet plotseling verticaal zoals in figuur 26. Het KNMI wijt de oplopende trend van 1907-1949 aan de geleidelijke vegetatiegroei. Het patroon in figuur 25C-rechts bevestigt deze interpretatie niet.

Op basis van de voorgaande resultaten vallen Oudenbosch (zie ook bijlage 1) en Gemert af om te worden opgenomen in de nieuwe referentieset van nabije stations. Het KNMI heeft gekozen voor Eelde als het enige referentiestation. Waarom voor Eelde is gekozen om De Bilt te homogeniseren

wordt in het technisch rapport niet vermeld, en evenmin waarom niet méér referentiestations zijn gekozen zoals gebruikelijk is. (Zie ook bijlage 3 voor een antwoord van het KNMI op deze vraag t.a.v. Beek).

Het residu van Eelde (figuur 25D-rechts) toont de tekortkoming van Eelde als enig referentiestation: de meetreeks begint pas in 1946 waardoor er maar 4 jaren vóór 1950 beschikbaar zijn. De periode van 1946 t/m 1949 speelt een belangrijke rol in de homogenisatie van De Bilt. We zien ook dat deze 4 jaren een relatief lage Tx vertonen. Dat doet vermoeden dat de homogenisatie van De Bilt, waarbij de temperatuurreksen van Eelde van 1946-1949 en 1952-1955 als referentie zijn gebruikt, niet representatief is. Het residu van Beek (Figuur 25E-rechts) toont dat Beek veel beter in de pas loopt met het Duitse ensemble dan Eelde (figuur 25D-rechts).



Figuur 27 Vergelijking van Tx (°C) van De Bilt, Winterswijk, Beek en Eelde in het interval 1946-1955.

In figuur 27 zijn voor de periode 1945-1955 behalve de Tx van De Bilt en Eelde ook die van Beek en Winterswijk weergegeven. Op basis van deze grafiek ligt een keuze voor Beek als referentiestation meer voor de hand dan van Eelde. Ook Winterswijk doet dan mee. Dat heeft als voordeel dat deze reeks al veel eerder begint.

Op basis van deze resultaten is een referentieset samengesteld uit Winterswijk, Beek en Eelde en de 3 meest nabijgelegen stations uit het Duits ensemble:

1	Winterswijk ¹⁴	1907 - 1970
2	Beek	1946 - 1991
3	Eelde	1946 - 1991
4	Aachen	1930 - 1991
5	Solingen	1936 - 1991
6	Münster	1907 - 1991

Tabel 6 De 3 Nederlandse en de 3 Duitse stations van de referentieset en de gebruikte jaren van Tx.

¹⁴ november 1944 afwezig, opgevuld met Münster november 1944 met kleine correctie

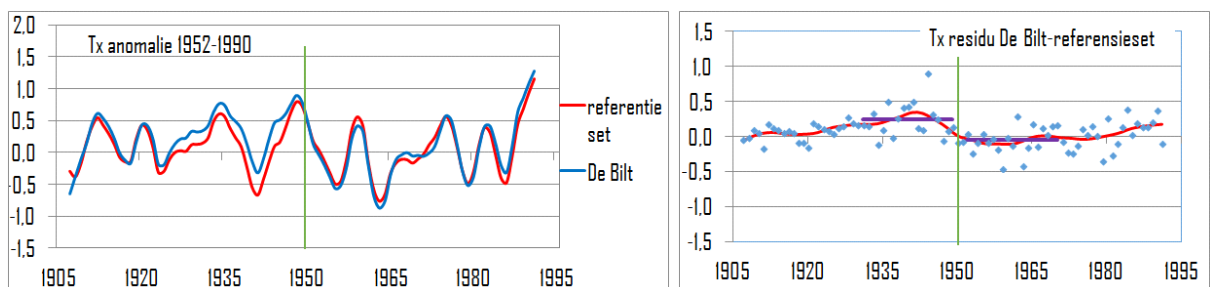


Figuur 28 De ligging van de 6 referentiestations.

De afstand van Eelde tot De Bilt is vergelijkbaar met de afstanden van de overige stations in de referentieset. Het argument dat deze stations te ver noordelijk of te ver zuidelijk liggen voor een analyse vervalt hierdoor omdat het KNMI Eelde gebruikte als referentie. Omdat de reeks van Münster in 1991 eindigt laten we de reeksen die verder doorlopen ook eindigen in 1991. Voor de analyse van de instrumentele breuk rond 1950 is een langere reeks ook niet relevant.

Deze stations zijn de basis voor onze nieuwe referentieset, hierna 'referentieset' genaamd. Op basis van deze referentieset van 6 stations hebben we een ensemble van de gemiddelde anomalie van de jaarlijks gemiddelde Tx bepaald. In Bijlage 1 is de methode beschreven om van deze stations, die verschillende perioden beslaan, een referentieset te maken voor de periode 1907-1991. We claimen overigens niet dat deze referentieset optimaal is voor vergelijking met De Bilt; uiteraard zullen er meteorologische verschillen zijn.

Vergelijking Tx De Bilt met de referentieset.



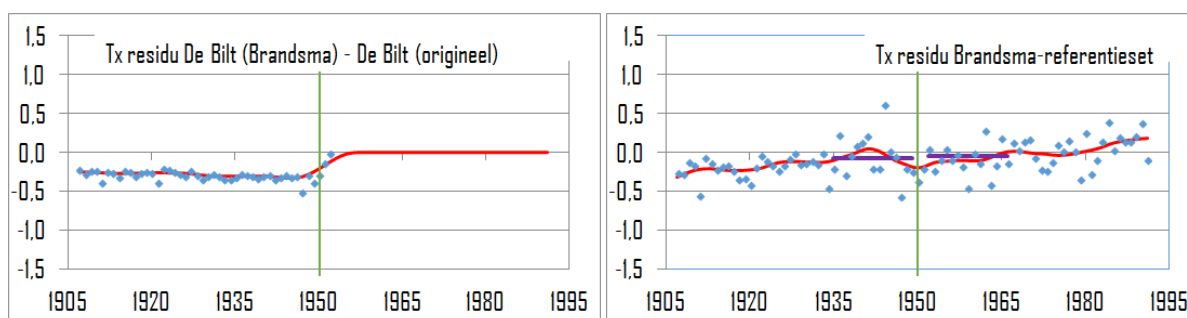
Figuur 29 Links: Tx (°C anomalie van 1952-1990) van De Bilt en de referentieset, Loess-gefilterd met span 8 jaar. Rechts: verschil tussen de jaarlijkse Tx van De Bilt en de jaarlijkse Tx van de referentieset. De rode lijn is een Loess filter met span 15 jaar door de punten. De paarse lijnen stellen de gemiddelde residuen voor in de intervallen 1931-1949 en 1952-1970; het niveauverschil is $0,29 \pm^{15}0,13$ °C.

¹⁵ Standaardfout van het verschil tussen twee gemiddelden:

$$\sigma_{M_1 - M_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Figuur 29-links laat zien dat na 1950 de data van De Bilt goed in de pas lopen met de referentieset, een sterke aanwijzing dat de referentieset representatief is. Het verschil tussen Tx De Bilt en Tx referentieset betreft alleen het interval 1925-1950. De afwijking van Tx De Bilt ten opzichte van de referentieset is te zien in het residu (figuur 29-rechts). Over de oorzaak van het feit dat er geen positief residu van 1906-1925 is kunnen we alleen speculeren. De instrumentele breuk eind jaren '40 is $-0,29 \pm 0,13$ °C. (NB: in het interval 1931-1970, dat gebruikt werd voor de schatting voor de grootte van de instrumentele breuk is de referentieset het gemiddelde van alle 6 stations). De hypothese is dat dit de benodigde correctie is voor die instrumentele breuk.

De vraag is daarom of de breukcorrectie die het KNMI op Tx De Bilt heeft toegepast voldoende is om de bovengenoemde instrumentele breuk ($-0,29 \pm 0,13$ °C) te compenseren. Onderstaande grafiek toont de feitelijk toegepaste correctie en vergelijken we het resultaat van deze correctie met de referentieset:

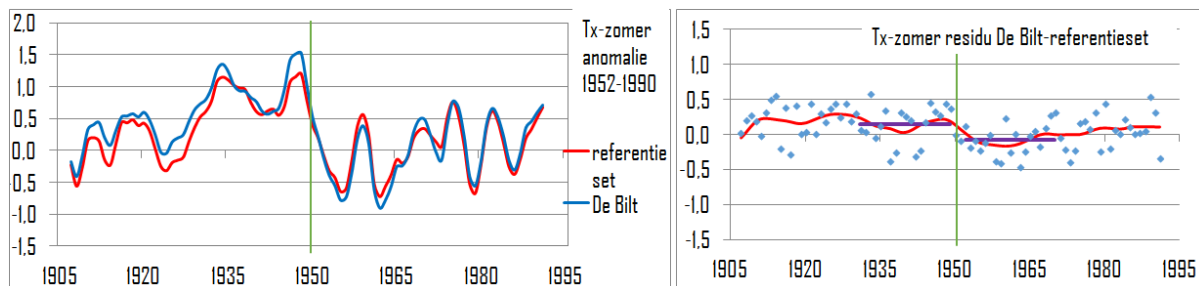


Figuur 30 Links: de breukcorrectie (verschil tussen de jaarlijkse Tx van de door het KNMI gehomogeniseerde De Bilt en de ongehogeniseerde De Bilt). De breukcorrectie is variabel en bedraagt $-0,32 \pm 0,01$ °C in het interval 1931-1949. Rechts: resultaat van de breukcorrectie op het verschil tussen Tx van de door het KNMI gehomogeniseerde De Bilt en Tx van de referentieset. De rode lijn is een Loess filter met span 15 jaar door de punten. De paarse lijnen stellen de gemiddelde residuen voor in de intervallen 1931-1949 en 1952-1970; het niveauverschil is $-0,03 \pm 0,08$ °C.

Figuur 30-links toont de correcties die het KNMI heeft toegepast op Tx in het interval 1901-1951. De correctie is variabel; in het interval 1931-1949 is de door het KNMI toegepaste correctie gemiddeld $-0,32 \pm 0,013$ °C. Figuur 30-rechts toont het resultaat van deze correctie op Tx De Bilt (vergelijk met figuur 22-rechts). De instrumentele breuk is vrijwel volledig gecompenseerd ($-0,03 \pm 0,08$ °C). De homogenisatie heeft niet geleid tot onder- of overcorrectie, althans wat de jaargemiddelden van Tx in het interval 1931-1970 betreft. Tussen 1910-1949 en 1952-1970 is het verschil $-0,13 \pm 0,05$ °C. Dat duidt wel op een overcorrectie.

Vergelijking Tx-zomer De Bilt met de referentieset

De invloed van de homogenisatie van de temperatuur van De Bilt is het grootst in de zomermaanden. Het technische rapport (Brandsma 2016a) stelt: “For Tx large positive corrections (up to 1.9°C) are found in summer for the largest percentiles”. Het is daarom van belang om de Tx- zomer van De Bilt te vergelijken met die van de referentieset. Tx-zomer is de gemiddelde maximale dagtemperatuur in de zomermaanden mei t/m september.

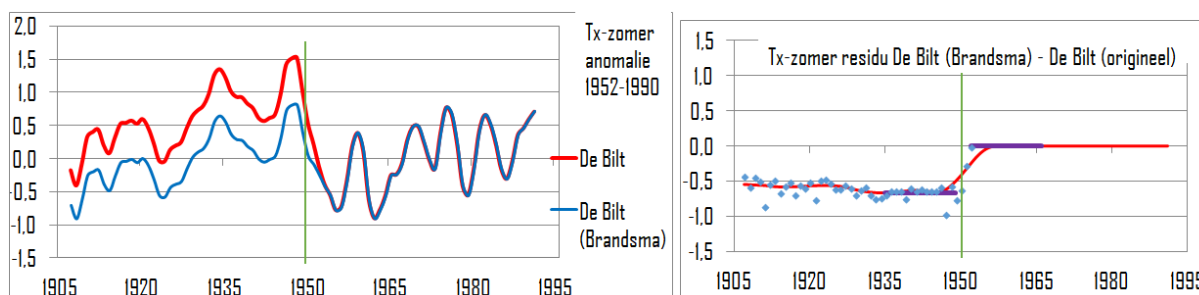


Figuur 31 Links: Tx-zomer ($^{\circ}\text{C}$ anomalie van 1952-1990) van De Bilt en de referentieset, Loess-gefilterd met span 8 jaar. Rechts: verschil tussen Tx-zomer van De Bilt en de referentieset. De paarse lijnen stellen de gemiddelde residuen voor in de intervallen 1931-1949 en 1952-1970; het niveauverschil is $0,22 \pm 0,09$ $^{\circ}\text{C}$.

Figuur 31 toont de resultaten van deze vergelijking. Door Tx alleen van de zomermaanden te nemen verandert het beeld enigszins: de klimatologische neergang tussen 1949 en 1956 is veel markanter. Ook is er nu een duidelijke sprong tussen vóór en ná de 1950-markering (figuur 31-rechts). Dat bevestigt dat er een breuk is rond 1950 van $0,22 \pm 0,09$ $^{\circ}\text{C}$, waarschijnlijk als gevolg van de instrumentele veranderingen. Maar de grootte van deze breuk is wel afhankelijk van het gekozen interval gezien de variabiliteit in het residu.

Evaluatie homogenisatie De Bilt

Zoals eerder aangegeven zijn we benieuwd of we de breukcorrectie van het KNMI kunnen reproduceren. Figuur 32-links toont de anomalie van Tx-zomer van De Bilt van de niet gehomogeniseerde reeks (rood) en de gehomogeniseerde reeks (blauw):



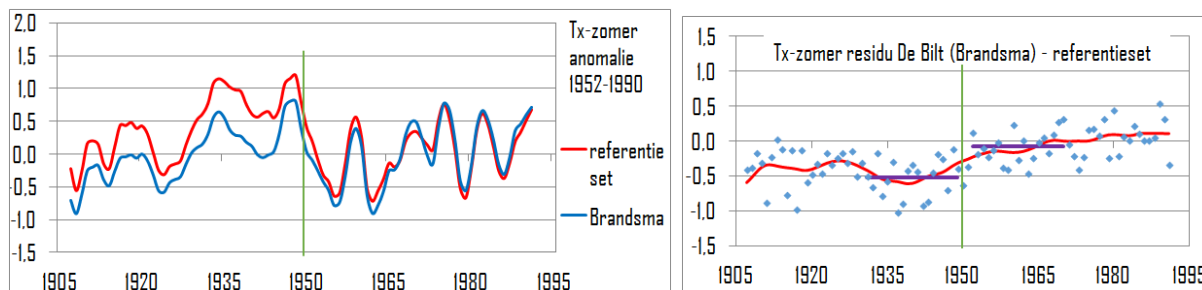
Figuur 32 Links: Tx-zomer ($^{\circ}\text{C}$ anomalie van 1952-1990) van De Bilt en de gehomogeniseerde Tx De Bilt, Loess-gefilterd met span 8 jaar. Rechts: verschil tussen Tx-zomer van de door het KNMI gehomogeniseerde De Bilt en de ongehomogeniseerde De Bilt. De paarse lijnen stellen de gemiddelde residuen voor in de intervallen 1931-1949 en 1952-1970; het niveauverschil is $0,67 \pm 0,02$ $^{\circ}\text{C}$.

Tx-zomer van de gehomogeniseerde reeks van De Bilt is in het interval 1952-1991 identiek aan dat van de ongehomogeniseerde Tx-zomer van De Bilt; de homogenisatie-correctie was alleen toegepast op het interval 1901- september 1951. Duidelijk is dat homogenisatie de Tx-zomer in dat interval sterk heeft verlaagd, de grootte van die verlaging is goed te zien in het residu van figuur 25-rechts.

Volgens de verwachting is elke waarde van Tx-zomer in het interval 1901-1951 door de homogenisatie verlaagd. Waar het om gaat is het verschil tussen vóór en ná 1950. De breukcorrectie rond 1950 is $-0,67 \pm 0,02$ $^{\circ}\text{C}$ (figuur 32-rechts). Anders gezegd: de originele Tx-zomer in De Bilt vóór 1950 is door homogenisatie met gemiddeld $0,67$ $^{\circ}\text{C}$ verlaagd. Die verlaging zou de benodigde correctie zijn voor de instrumentele breuk. Wat ook opvalt is dat 1947, oorspronkelijk een recordzomer, met 1 $^{\circ}\text{C}$ is verlaagd en 1911, eveneens een warme zomer, met $0,86$ $^{\circ}\text{C}$. Waarom de door het KNMI gevolgde homogenisatieprocedure recordwarme zomers relatief veel meer verlaagt

dan andere jaren is ons niet duidelijk geworden. Deze temperatuur-afhankelijke correctie wordt in ieder geval niet gesteund door figuur 31-rechts.

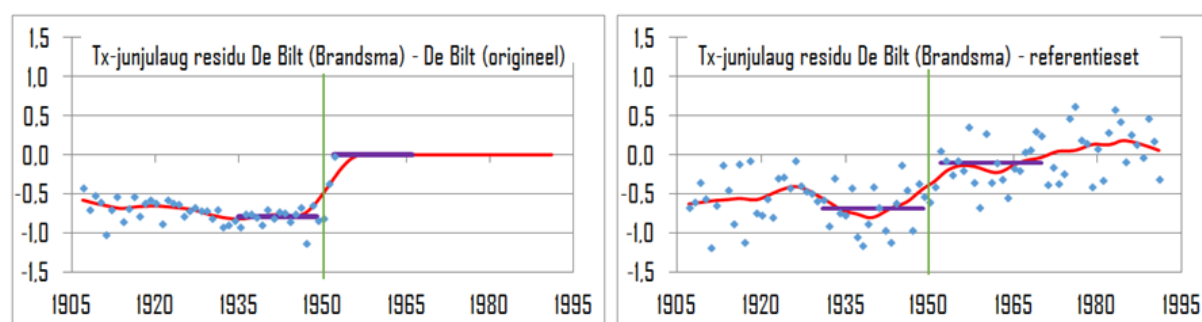
Om de grootte van de breukcorrectie te evalueren vergeleken we de gehomogeniseerde Tx-zomer met onze referentieset:



Figuur 33 Links: Tx-zomer ($^{\circ}\text{C}$ anomalie van 1952-1990) van de gehomogeniseerde De Bilt en van de referentieset, Loess-gefilterd met span 8 jaar. Rechts: verschil tussen Tx-zomer van de gehomogeniseerde De Bilt en de referentieset. De rode lijn is een Loess filter met span 15 jaar door de punten. Het niveauverschil tussen de paarse horizontale lijnen is $0,45 \pm 0,08$ $^{\circ}\text{C}$.

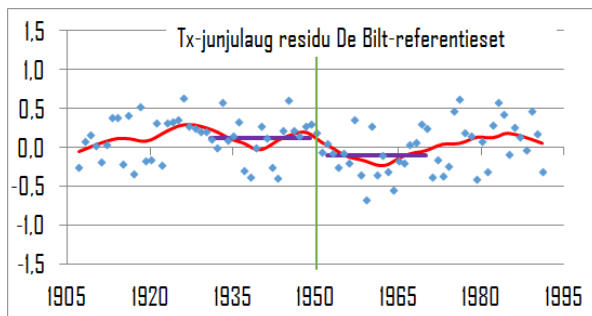
Figuur 33-links toont dat de breukcorrectie tot gevolg heeft dat Tx-zomer in het interval 1907-1950 een stuk lager is dan Tx-zomer van referentieset. In figuur 33-rechts toont de grootte van de breukcorrectie in vergelijking met de referentieset; deze bedraagt $-0,45 \pm 0,08$ $^{\circ}\text{C}$. Dat wijst op een *flinke overcorrectie* want het verschil had nul moeten zijn. De benodigde correctie voor de instrumentele trendbreuk was namelijk $-0,22 \pm 0,09$ $^{\circ}\text{C}$ (zie figuur 31). De breukcorrectie van het KNMI is echter $-0,67^{\circ}\text{C}$ ($-0,22^{\circ}\text{C}$ $-0,45^{\circ}\text{C}$, zie ook figuur 32).

De geschatte grootte van de overcorrectie is wel afhankelijk van de gekozen lengte van de intervallen vóór en ná 1950 zoals te zien is in figuur 32-rechts: de neerwaartse correctie neemt geleidelijk af van 1950 naar 1907. Bovendien blijkt de grootte van de overcorrectie afhankelijk te zijn van de temperatuur: beperken we de zomermaanden tot de warmste maanden juni, juli en augustus dan wordt de overcorrectie nog groter:



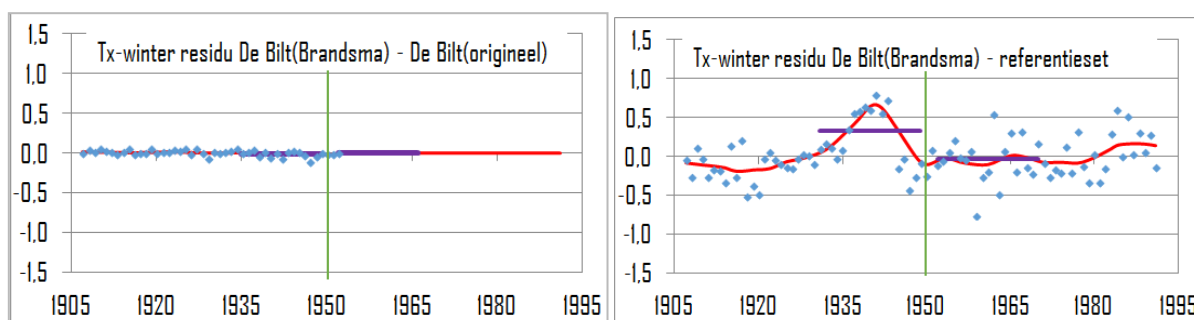
Figuur 34 Links: de door het KNMI toegepaste breukcorrectie op de gemiddelde Tx van juni, juli en augustus. De toegepaste correctie bedraagt gemiddeld $-0,8 \pm 0,03$ $^{\circ}\text{C}$ in het interval 1931-1970. Rechts: gevolg van deze breukcorrectie. Het niveauverschil tussen de paarse horizontale lijnen is $0,58 \pm 0,15$ $^{\circ}\text{C}$.

Zoals figuur 34 toont is de gemiddelde overcorrectie nu groter: $0,58 \pm 0,15$ $^{\circ}\text{C}$. De benodigde correctie was echter niet $-0,8$ $^{\circ}\text{C}$ maar $-0,22$ $^{\circ}\text{C}$:



Figuur 35 Verschil tussen de gemiddelde Tx van juni, juli en augustus van De Bilt en de referentie set. Het niveauverschil tussen de paarse lijnen is $0,22 \pm 0,09$ °C.

In de wintermaanden (november t/m maart) daarentegen is de door het KNMI toegepaste correctie op Tx vrijwel nihil, waardoor de wintermaanden een ondercorrectie vertonen:



Figuur 36 Links: de door het KNMI toegepaste correctie op Tx-winter ($-0,01 \pm 0,01$ °C). Rechts: verschil tussen Tx-winter van de gehomogeniseerde De Bilt en de referentieset. Het niveauverschil tussen de paarse horizontale lijnen is $0,37 \pm 0,14$ °C.

Figuur 36-rechts toont dat de ondercorrectie alleen geldt voor het interval 1930-1944. Zoals in figuren 22 en 29 ook al werd getoond lijkt de instrumentele breuk voor wat betreft Tx en Tx-winter niet rond 1950 maar rond 1945 te liggen, dit in tegenstelling tot Tx-zomer waar de instrumentele breuk wel rond 1950 ligt (zie figuren 33 en 34). Wat hiervan de reden is hebben wij niet verder onderzocht.

T	Breuk volgens Duitse ensemble	Breuk volgens referentieset	Breukcorrectie KNMI
Tx	$-0,37 \pm 0,11$	$-0,29 \pm 0,13$	$-0,32 \pm 0,01$
Tx-zomer	$-0,40 \pm 0,14$	$-0,22 \pm 0,09$	$-0,67 \pm 0,02$
Tx-jul/jun/aug	$-0,49 \pm 0,14$	$-0,22 \pm 0,09$	$-0,80 \pm 0,03$
Tx-winter	$-0,34 \pm 0,13$	$-0,37 \pm 0,14$	$-0,01 \pm 0,01$

Tabel 7 Kolom 2 en 3: grootte van de instrumentele breuk in Tx (°C) in De Bilt in het interval 1931-1970 volgens vergelijking met het Duitse ensemble en met de referentie set. Kolom 3: de door het KNMI toegepaste correctie voor de instrumentele breuk.

Op basis van deze resultaten in tabel 7 kunnen we concluderen dat de homogenisatie door het KNMI een aanzienlijke overcorrectie heeft opgeleverd in de zomermaanden en een ondercorrectie in de wintermaanden.

Conclusies

1. De trendbreuk in Tx in De Bilt, die volgens het KNMI veroorzaakt zou zijn door instrumentele veranderingen rond 1950, valt in het interval van een klimatologische neergang die zich voordeed in een groter deel van NW Europa tussen 1949 en 1956. Door de verschilreeksen te nemen van De Bilt met representatieve ensembles van andere stations werd bevestigd dat er zich in De Bilt inderdaad een residuele trendbreuk in Tx voordeed: Tx in De Bilt is vóór 1950 relatief te hoog. Hierbij werden twee vergelijkingsensembles gebruikt, het ene bestaande uit 7 Duitse stations (D ensemble) en het andere uit 3 nabijgelegen Duitse stations en 3 geschikte Nederlandse stations (D-N ensemble).
2. Voor de homogenisatie van De Bilt heeft het KNMI slechts één meetstation gebruikt: Eelde, waarvan maar vier jaren vóór 1950 beschikbaar waren die bovendien een relatief lage Tx vertonen. Daardoor zijn er twijfels of de homogenisatie met uitsluitend Eelde als referentie wel representatief is.
3. Om de door het KNMI toegepaste neerwaartse correcties op Tx De Bilt vóór 1950 te evalueren hebben we in deze studie een ruwe schatting gemaakt van de benodigde breukcorrectie. Die schatting kwam tot stand door vergelijking van de ongehomogeniseerde en gehomogeniseerde reeksen van De Bilt met het D-N ensemble. Daarbij werden de vergelijkingsintervallen 1931-1949 en 1952-1990 gebruikt. Het D-N ensemble is representatief omdat deze het temperatuurverloop in De Bilt na 1952 in grote mate reproduceert.
4. Voor de jaargemiddelde Tx bleek de benodigde breukcorrectie niet significant te verschillen van de door het KNMI toegepaste Tx correctie. De KNMI Tx correctie voor de wintermaanden (november t/m maart) is nihil waardoor deze een sterke ondercorrectie vertoont in vergelijking met het D-N ensemble. De KNMI Tx correctie voor de zomermaanden (mei t/m september) is daarentegen weer veel te groot: er is sprake van een overcorrectie van ruim 200%. Voor de drie warmste zomermaanden (juni, juli, augustus) loopt de overcorrectie zelfs op tot ruim 260%. Daarbij is ook gevonden dat de jaargemiddelde Tx in record warme jaren (1947 en 1911) door het KNMI relatief sterker naar beneden werden bijgesteld, hetgeen niet wordt gesteund door onze vergelijking met het referentie-ensemble

7 Effect van de keuze vergelijkingsstation op de homogenisatie

Het KNMI (Brandsma 2016a) heeft gekozen voor Eelde als vergelijkingsstation voor de homogenisatie van de dagelijkse temperaturen van De Bilt. Voor Eelde is een homogene reeks beschikbaar vanaf 1 januari 1946, maar dit station ligt 147 km ten noordoosten van De Bilt. Andere stations met metingen die teruggaan tot 1946 of eerder, en die dichterbij De Bilt liggen zijn: Gemert (65 km), Oudenbosch (75 km) en Winterswijk (105 km), maar die reeksen zijn door instrumentele veranderingen niet homogeen. Toch zijn deze stations wel gebruikt bij de constructie van de CNT-reeks (Van der Schrier c.s., 2011). Vliegveld Beek ligt iets dichterbij De Bilt dan Eelde (138 km). Van Beek is ook een homogene reeks beschikbaar vanaf 1 januari 1946.

Het KNMI motiveert haar keuze voor Eelde als volgt (KNMI, 2018): *“Eelde is beter geschikt als parallel station dan Beek.... Dit heeft te maken met de afwijkende hoogteligging van Beek (114 m boven Nap) en de directe nabijheid van het Maasdal. Daarnaast is voor de overheersende windrichting in Nederland (ZW-W) de afstand van De Bilt en Eelde tot grote wateroppervlakken vergelijkbaar.”*

Het KNMI heeft blijkbaar niet onderzocht hoe groot het effect van het hoogteverschil en de nabijheid van het Maasdal is. Het is waar dat De Bilt en Eelde bij de overheersende ZW-wind beter vergelijkbaar zijn, de stations liggen bij ZW-wind althans in elkaars verlengde. Maar doel van de homogenisatie van het KNMI was mede om het aantal hittegolven vóór 1950 te verminderen (Brandsma 2016a). Tijdens hittegolven is de overheersende windrichting niet ZW maar vaak ZO, en dan zou Beek beter geschikt kunnen zijn als vergelijkingsstation. De vraag moet dus gesteld worden wat het effect is van de keuze van geschikte alternatieve vergelijkingsstations.

Aangezien we hier alleen de referentieperiode 1946-1955 gebruiken voor de homogenisatie van De Bilt kunnen we vergelijkingsstations gebruiken die in ieder geval binnen deze periode geen homogenisaties hebben ondergaan. In H6 zagen we dat er vijf Nederlandse stations beschikbaar zijn met een meetreeks die tenminste de periode 1946-1955 omvat en geen homogenisaties hebben ondergaan, namelijk Oudenbosch, Gemert, Winterswijk, Eelde en Beek.

Uit vergelijking met het Duitse ensemble bleek in H6 dat Gemert een grote trendbreuk vertoont rond 1950, een breuk die eveneens door het KNMI gerapporteerd is. Het KNMI signaleerde voor Oudenbosch trendbreuken in 1966 en 1984 als gevolg van herlocatie. Oudenbosch zou dus in principe wel in aanmerking kunnen komen als vergelijkingsstation omdat we ons beperken tot 1946-1955. In H6 bleek echter dat Oudenbosch grote trendafwijkingen vertoont in vergelijking met het Duitse ensemble. Om die reden werd Oudenbosch niet opgenomen in de referentieset (het Duits-Nederlandse ensemble in H6). Als we in H6 figuur 25-rechts het Tx residu van Oudenbosch echter vergelijken met dat van Eelde tussen 1946 en 1955 dan wijken deze stations binnen deze periode weinig van elkaar af: het verschil tussen het gemiddelde 1946-1949 en het gemiddelde 1952-1955 is 0,16°C. Voor Eelde is dit 0,22°C. Het verschil is dus zeer klein. Omdat bovendien het KNMI voor Oudenbosch geen breuk heeft gedetecteerd rond 1950 kan voor ons doel hier Oudenbosch wel meegenomen worden als vergelijkingsstation.

Voor Winterswijk heeft het KNMI kleine inhomogeniteiten gevonden in 1940, 1950 en 1960. In figuur 25A-rechts van H6 zien we dat het verloop tussen 1946 en 1955 relatief groot is (0,26°C). In H4 werd geconcludeerd dat de verschillen wel relevant zijn als alleen de kleine 4-jaars intervallen rond 1950 worden beschouwd; voor grotere vergelijkingsintervallen, zoals die in H6 worden gebruikt (1931-1949 en 1952-1970) is het verschil met het Duitse ensemble marginaal (0,08°C). Hoewel Winterswijk niet als het meest representatieve vergelijkingsstation geldt voegen we hem hier wel toe aan de lijst van vergelijkingsstations om te zien of dit station sterk afwijkende resultaten geeft ten opzichte van de andere vergelijkingsstations.

Tenslotte nemen we ook het Duits-Nederlandse ensemble als referentie mee. Dit ensemble bevat de bovengenoemde stations Eelde, Beek en Winterswijk ook de meest nabijgelegen Duitse stations Aachen, Solingen en Münster. Deze drie Duitse stations liggen niet veel verder van De Bilt dan Eelde. Bij hittegolven komt de wind meestal uit Duitsland, zodat de nabijgelegen Duitse stations wel relevant kunnen zijn voor Brandsma's probleemstelling.

Werkwijze

We hebben het effect van de keuze van het vergelijkingsstation onderzocht door de percentiel matching methode toe te passen, zoals beschreven in hoofdstuk 4, met parallelle data voor alle stations waarvoor in de periode 1946-1955 daggegevens beschikbaar zijn. Dat zijn behalve de bovengenoemde Nederlandse stations Eelde, Beek, Oudenbosch en Winterswijk ook het meest nabijgelegen Duitse station Aachen en het Duits-Nederlandse ensemble.

Voor het vergelijken van het effect van de keuze van alternatieve vergelijkingsstations gebruiken we twee criteria:

- (1) het aantal tropische dagen in de periode 1906-1950 als percentage van het aantal tropische dagen in de periode 1906-2016
- (2) de gemiddelde correctie die Tx ondergaat als gevolg van de homogenisatie

Resultaten

Uit tabel 8 blijkt dat de keuze van het vergelijkingsstation van grote invloed is op het aantal tropische dagen dat men vindt na homogenisatie van De Bilt: bij een span van 0,6 (zoals Brandsma zegt te gebruiken) varieert het van 31,6% (Winterswijk) tot 42,2% (Aachen). Die percentages liggen ver boven de uitkomst van de homogenisatie van het KNMI met Eelde als vergelijkingsstation (26,4%). Maar daarbij is in feite een span van 0,75 gebruikt door een extra voortschrijdende middeling over driemaandelijke perioden.

Met een span van 0,75 vinden we percentages tropische dagen in De Bilt variërend van 31,6% (Winterswijk) tot 42,7% (Aachen). Met geen enkel alternatief vergelijkingsstation vinden wij van 1906-1950 zo weinig tropische dagen als het KNMI vindt met Eelde als vergelijkingsstation (26,4%).

		tropische dagen 1906-1950 / 1906-2016		gemiddelde correctie Tx 1906-1950	
Vergelijkingsstation(s)	Afstand tot De Bilt (km)	Span 0,6	Span 0,75	Span 0,6	Span 0,75
Eelde (Brandsma)	147	26,4%		-0,29	
Eelde (reconstructie Brandsma)	147	34,7%	36,4%	-0,37	-0,37
Beek	138	38,0%	36,6%	-0,07	-0,06
Winterswijk	105	31,6%	31,6%	-0,39	-0,42
Oudenbosch	75	37,0%	37,0%	-0,30	-0,31
Aachen	159	42,2%	42,7%	-0,00	0,00
Duits-Nederlands ensemble		35,8%	36,4%	-0,12	-0,14

Tabel 8 Effect van representatieve vergelijkingsstations en span van de smoothing op homogenisatie Tx De Bilt.

De keuze voor een span van 0,6 of 0,75 heeft in een aantal gevallen vrijwel geen invloed op het aantal tropische dagen, in een aantal andere gevallen is het verschil 1 tot 2 procentpunten, maar voor sommige stations in tegengestelde richtingen: het is niet altijd zo dat een hogere span tot meer

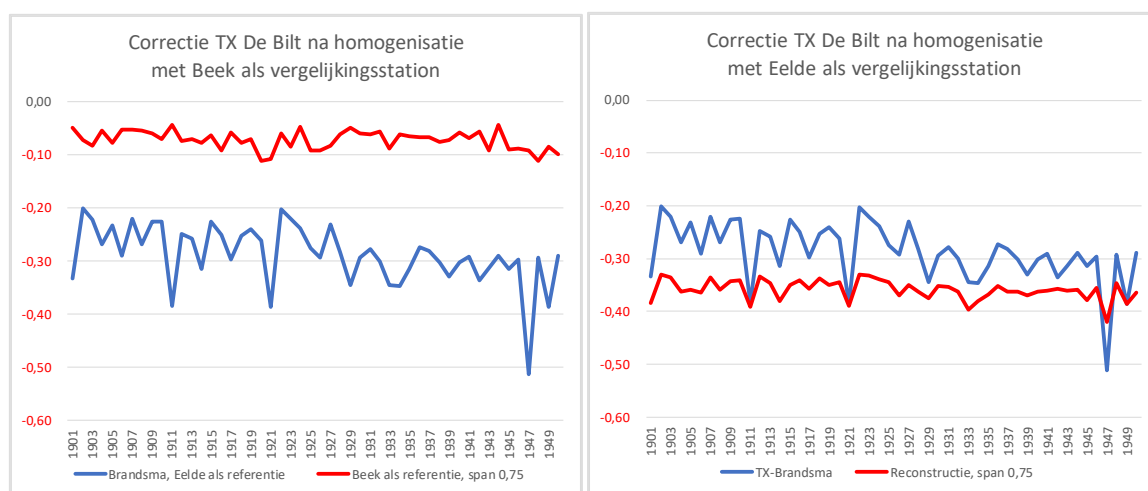
tropische dagen leidt.

Kijken we naar de uitkomst van het Duits-Nederlandse ensemble dan is het percentage tropische dagen vrijwel gelijk aan dat van de Nederlandse stations, met uitzondering van Winterswijk. Zoals eerder genoemd is Winterswijk een minder representatief vergelijkingsstation gezien de kleine trendbreuk rond 1950. Het percentage tropische dagen voor de overige Nederlandse stations, inclusief de Eelde reconstructie en het Duits-Nederlandse ensemble, is tamelijk constant rond de 35% - 36%.

Dit stemt overeen met wat Dijkstra (2017) vond voor de gehomogeniseerde data van de KNMI-hoofdstations.

Gemiddelde correctie Tx

De gemiddelde correctie van Tx door de homogenisatie varieert van $-0,39\text{ }^{\circ}\text{C}$ tot $0,00\text{ }^{\circ}\text{C}$, afhankelijk van het onderzochte vergelijkingsstation zoals vermeld in tabel 8. Dit lijkt geen groot verschil, maar bij de hoogste percentielen lopen de grootste correcties op tot meer dan $-1\text{ }^{\circ}\text{C}$. Brandsma vindt zelfs correcties van $-1,9\text{ }^{\circ}\text{C}$ voor de hoogste percentielen in augustus.



Figuur 37 Effect van de keuze voor Eelde of Beek als vergelijkingsstation op correctie Tx voor De Bilt

Figuur 37 laat zien hoe de homogenisatie uitwerkt voor de tijdreeks van de jaargemiddelden van Tx. De blauwe lijnen in de beide grafieken geven de correcties van de homogenisatie van het KNMI met Eelde als vergelijkingsstation. De rode lijnen geven onze reconstructies van de homogenisatie met Eelde dan wel Beek als vergelijkingsstation. Enige zaken vallen hierin op:

- (1) Onze reconstructie met Eelde als vergelijkingsstation komt gemiddeld lager uit dan die van het KNMI ($-0,37$ versus $-0,29$ bij het KNMI), terwijl onze reconstructie wel tot meer tropische dagen leidt dan de homogenisatie van het KNMI (tabel 8).
- (2) De variatie tussen de jaren is bij het KNMI veel groter dan in onze reconstructie, met als grote uitschieter het jaar 1947, waar het KNMI de gemiddelde Tx tweemaal zo sterk corrigeert als in onze reconstructie.
- (3) Onze homogenisatie met Beek als vergelijkingsstation komt gemiddeld hoger uit dan die van het KNMI ($-0,07$ versus $-0,29$ bij het KNMI).
- (4) De variatie tussen de jaren is bij onze homogenisatie met Beek als vergelijkingsstation kleiner dan met Eelde als vergelijkingsstation. Het jaar 1947 springt er bij homogenisatie met Beek niet meer uit.

Effect van andere Nederlandse en Duitse stations.

Wij hebben ook het effect onderzocht van minder representatieve en niet-representatieve vergelijkingsstations met als doel om de uitkomsten die de *percentiel matching methode* met deze stations oplevert te evalueren.

Gemert is niet representatief gezien de grote trendbreuk rond 1950 (zie figuur 25 in H6). De verder gelegen Duitse stations zijn vanwege de afstand mogelijk minder representatief om te gebruiken voor de homogenisatie van De Bilt. Tabel 9 vat de resultaten met deze vergelijkingsstations samen.

		tropische dagen 1906-1950 / 1906-2016		gemiddelde correctie Tx 1906-1950	
Vergelijkingsstation(s)	Afstand tot De Bilt (km)	Span 0,6	Span 0,75	Span 0,6	Span 0,75
Gemert	65	44,8%	42,8%	0,13	0,12
Geisenheim	343	38,5%	39,3%	-0,07	-0,06
Lindenberg	606	27,7%	29,0%	-0,34	-0,35
Potzdam	542	33,9%	36,4%	-0,20	-0,21
Jena	467	36,4%	33,9%	-0,19	-0,18
Ensemble Duitse stations		33,9%	33,9%	-0,12	-0,14

Tabel 9 Effect van andere vergelijkingsstations en span van de smoothing op homogenisatie Tx De Bilt.

In tabel 9 zien we dat Gemert tot een positieve correctie van Tx De Bilt leidt en een hoog percentage tropische dagen. Lindenberg leidt tot het laagste percentage. De overige vergelijkingsstations leveren percentages tropische dagen die iets boven of onder de in tabel 8 genoemde 35-36% liggen. Het Duitse ensemble levert dezelfde Tx correctie op als het Duits-Nederlandse ensemble in tabel 8 en een iets lager percentage tropische dagen.

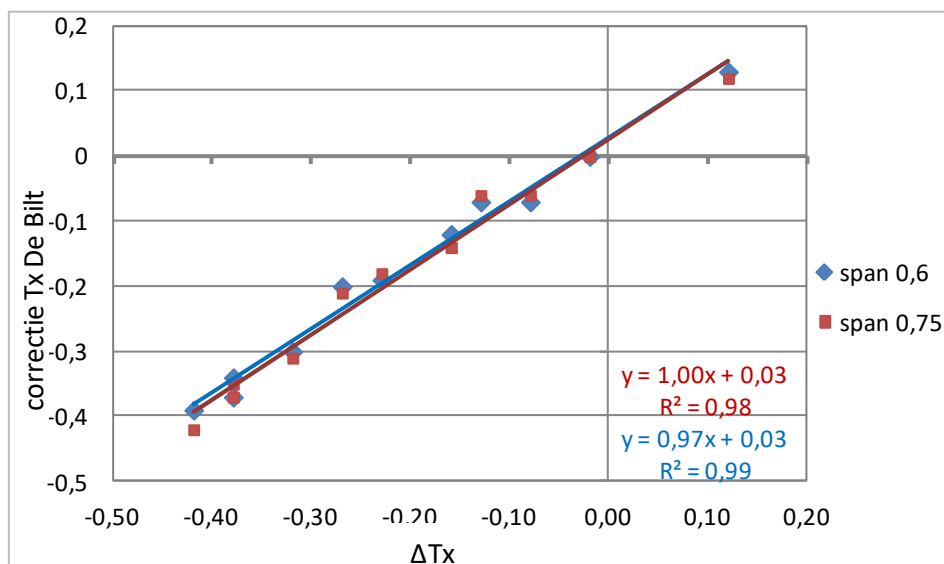
De resultaten in tabel 9 zijn slechts illustratief en hoogstwaarschijnlijk niet representatief als homogenisatie van De Bilt, maar kunnen wel gebruikt worden om de resultaten van de *percentiel matching methode* met station Eelde te evalueren. Wij onderzochten of er een verband bestaat tussen de gemiddelde Tx correctie (genoemd in tabel 8 en 9) en de gemiddelde ΔTx per vergelijkingsstation in de twee vergelijkingsperioden van telkens 4 jaar. De jaargemiddelde ΔTx = temperatuurverschil De Bilt tussen [1946-1949] en [1952-1955] minus temperatuurverschil Eelde tussen [1946-1949] en [1952-1955]. Dit is hetzelfde als $\Delta Tx = \Delta T1 - \Delta T2$ in tabel 4 van H4 waar de formule gebruikt werd voor de percentielen per maand.

In onderstaande tabel is de Tx-correctie, zoals vermeld in tabel 8 en tabel 9, gezet naast de jaargemiddelde ΔTx :

vergelijkings-station	ΔT_x	correctie De Bilt Tx 1906-1950	
		span 0,6	span 0,75
Gemert	0,12	0,13	0,12
Aachen	-0,02	0	0
Beek	-0,08	-0,07	-0,06
Geisenheim	-0,13	-0,07	-0,06
Eelde	-0,38	-0,37	-0,37
Lindenberg	-0,38	-0,34	-0,35
Winterswijk	-0,42	-0,39	-0,42
Oudenbosch	-0,32	-0,3	-0,31
Potzdam	-0,27	-0,2	-0,21
Jena	-0,23	-0,19	-0,18
D-ensemble	-0,16	-0,12	-0,14
DN-ensemble	-0,16	-0,12	-0,14

Tabel 10 Combinatie van gegevens uit tabel 8 en 9 en de jaargemiddelde ΔT_x

Als we de gemiddelde Tx correctie De Bilt uitzetten tegen ΔT_x krijgen we onderstaande grafiek:



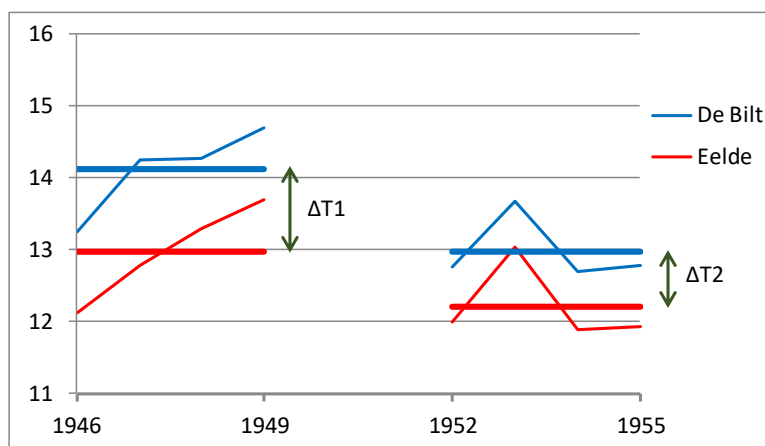
Figuur 38 Relatie tussen jaargemiddelde ΔT_x (°C) en de gemiddelde Tx correctie De Bilt voor alle vergelijkingsstations (exclusieve percentielen)

De vergelijking van de trendlijnen is bij benadering $y = x$. De gebruikte span maakt geen noemenswaardig verschil. De totale variatie in de Tx-correctie van De Bilt wordt voor 98-99% verklaard door ΔT_x .

Dat is een onverwacht resultaat want in de *percentiel matching methode* wordt ΔT_x berekend per maand per percentiel (tabel 2). De *percentiel matching methode* voegt dus blijkbaar niets significant toe aan wat niet al uit de jaargemiddelde ΔT_x blijkt (althans als we ons beperken tot jaargemiddelden). Kennelijk wordt de ΔT_x over de maanden, percentielen en jaren zo sterk uitgemiddeld dat het resultaat gelijk is aan de jaargemiddelde ΔT_x . Een indicatie daarvoor bleek al uit figuur 15-rechts van H4: de smoothing leidt tot een sterke afvlakking over de percentielen en maanden.

Uit figuur 38 blijkt ook dat naarmate ΔT_x sterker negatief is (dus naarmate Tx van het vergelijkingsstation in het interval 1946-1949 relatief lager is in vergelijking tot De Bilt) dat dan de

neerwaartse correctie van Tx in De Bilt in de periode 1906-1950 groter is. Deze situatie is weergegeven in figuur 39 waarin $\Delta T1$ groter is dan $\Delta T2$:

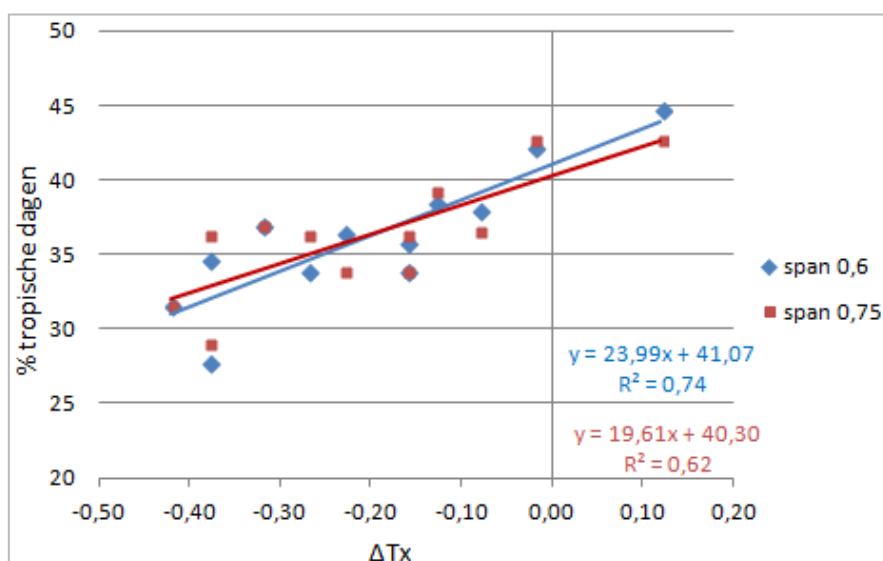


Figuur 39 Schematische weergave van $\Delta T1$ en $\Delta T2$ tussen Tx. De vette lijnen zijn de gemiddelde Tx'en in de intervallen 1946-1949 en 1952-1955 en de dunne lijnen de originele Tx'en.

In figuur 39 zijn $\Delta T1$ en $\Delta T2$ negatief want Tx Eelde is lager dan Tx De Bilt. In het bovenstaande voorbeeld is Tx De Bilt 1946-1949 = 14,12 °C en 1952-1955 = 12,98 °C. Eelde 1946-1949 = 12,98 °C en 1952-1955 = 12,21 °C. $\Delta Tx = \Delta T1 - \Delta T2 = -1,14$ °C + 0,76 °C = -0,38 °C. De ΔTx is een goede voorspeller van de Tx correctie op basis van Eelde (zie tabel 10: -0,37°C). De formule voor de schatting van de Tx correctie is dus:

$$Tx \text{ correctie} = \Delta Tx = \Delta T1 - \Delta T2 \quad (1)$$

Eelde en Winterswijk zijn dus vergelijkingsstations die tot een grote Tx-correctie in De Bilt leiden. Als het vergelijkingsstation een klein ΔTx -verschil met De Bilt vertoont dan is de Tx correctie klein. Met andere woorden: als voor de homogenisatie van De Bilt een ander vergelijkingsstation wordt gebruikt dan Eelde, bijvoorbeeld Beek of Aachen die een klein ΔTx -verschil met De Bilt vertonen, dan is de Tx-correctie in De Bilt ook klein. Tx in De Bilt in de periode 1906-1950 is dan hoger en er refteren dan waarschijnlijk ook meer tropische dagen:



Figuur 40 Relatie tussen jaargemiddelde ΔTx (°C) en het percentage tropische dagen De Bilt voor alle vergelijkingsstations (exclusieve percentielen)

In figuur 40 zien we dat het percentage tropische dagen eveneens afhankelijk is van het ΔTx -verschil, hoewel de relatie wat minder sterk is.

Als we ons beperken tot de zomermaanden dan verwachten we volgens figuur 32-rechts van H6 een hogere Tx-correctie. Op basis van bovenstaande formule 1, toegepast op Tx-zomer, voorspellen we een Tx correctie van $-0,69^{\circ}\text{C}$. De homogenisatie van het KNMI levert een Tx-correctie van $-0,62^{\circ}\text{C}$ en onze herhomogenisatie levert een Tx-correctie van $-0,64^{\circ}\text{C}$. Dat komt dus in de buurt van de voorspelling. Uit figuur 32 van H6 bleek dat de Tx-correctie van het KNMI een aanzienlijke overcorrectie oplevert in vergelijking met de referentieset (en dat geldt dan ook voor onze herhomogenisatie). Dat ligt aan Eelde als vergelijkingsstation omdat Eelde een relatief lage Tx vertoont zoals ook al af te leiden viel uit figuur 25D-rechts van H6. Zou Aachen of Beek als vergelijkingsstation zijn gebruikt dan zou Tx-zomer hoogstwaarschijnlijk geen overcorrectie opleveren.

De *percentiel matching methode* met Eelde als vergelijkingsstation is nog om een andere reden discutabel omdat de analyse zich noodzakelijkerwijs beperkt tot 4-jaarsintervallen. Data van Eelde starten pas vanaf 1946. Maar Tx varieert sterk over de jaren en dus ook ΔTx . Hier het resultaat als het Duits-Nederlandse ensemble als referentie wordt gebruikt:

Vergelijkingsinterval Duits-Ned. ensemble	ΔTx	verwachting	
		Tx-correctie	% tropische dagen
1946-1955	$-0,16^{\circ}\text{C}$	$-0,13^{\circ}\text{C}$	37,2%
1941-1960	$-0,38^{\circ}\text{C}$	$-0,35^{\circ}\text{C}$	32,9%
1936-1965	$-0,37^{\circ}\text{C}$	$-0,34^{\circ}\text{C}$	33,0%
1931-1970	$-0,30^{\circ}\text{C}$	$-0,27^{\circ}\text{C}$	34,4%

Tabel 11 Invloed van het vergelijkingsinterval op ΔTx van het Duits-Nederlands ensemble en de verwachte Tx-correctie en % tropische dagen volgens figuur 38 en 40 (span 0,75).

Volgens figuur 38 zou dan ook de Tx correctie en volgens figuur 40 ook het percentage tropische dagen variëren afhankelijk van het gekozen vergelijkingsinterval. Dat maakt de KNMI-procedure extra arbitrair.

Discussie

Uit tabel 8 blijkt dat de gemiddelde correctie van Tx in De Bilt uiteenloopt van $-0,39^{\circ}\text{C}$ tot $0,0^{\circ}\text{C}$, afhankelijk van het vergelijkingsstation. Stations die een grote jaargemiddelde ΔTx opleveren leiden tot de grootste neerwaartse Tx correcties in De Bilt in de periode 1906-1950.

De uitkomst van de *percentiel matching methode* is, wat de jaargemiddelden betreft, vrijwel gelijk aan het Tx verschil (ΔTx) tussen De Bilt en een vergelijkingsstation in het vergelijkingsinterval. Kennelijk worden de percentielen over de maanden en jaren (zie tabel 1 en 2) sterk uitgemiddeld. Het gebruik van Eelde of Winterswijk als vergelijkingsstation leidt daardoor tot grote Tx correcties voor De Bilt, het gebruik van Beek en Aachen leidt daarentegen tot veel kleinere correcties. Met andere woorden: omdat geen enkel vergelijkingsstation goed met De Bilt is te vergelijken, is de uitkomst van de percentiel matching methode geheel afhankelijk van de keuze van het vergelijkingsstation, en die keuze kan, gelet op tabel 8 en figuur 37 alleen maar arbitrair zijn.

Een beperking van vergelijkingsstation Eelde is dat er maar 4 jaar vóór 1950 beschikbaar zijn als vergelijkingsinterval. Gezien de variabiliteit van Tx van jaar tot jaar en daarmee ook de ΔTx over verschillende vergelijkingsintervallen kan de homogenisatie met Eelde als niet representatief worden beschouwd.

Een andere beperking van vergelijkingsstation Eelde is dat het leidt tot een grote overcorrectie van Tx in de zomermaanden in De Bilt. Dat komt omdat de Tx-zomer in Eelde relatief koel is in vergelijking met andere vergelijkingsstations. Andere vergelijkingsstations zoals Beek, Aachen en het Duits-Nederlandse ensemble leiden niet tot overcorrectie.

Bij deze conclusies past nog een voorbehoud: zoals Michiel opmerkt is de *percentiel matching methode* een puur statistische methode die geen rekening houdt met verschillende meteorologische omstandigheden.¹⁶ Het opwarmende effect van de pagode is vooral te verwachten op dagen met hoge zonnestand en weinig wind. De *percentiel matching methode* houdt daar geen rekening mee, maar past op bewolkte, winderige, maar relatief warme dagen dezelfde correctie toe als op zonnige dagen met weinig wind met dezelfde temperatuur. Het is twijfelachtig of dit wetenschappelijk verantwoord is.

Conclusies

1. Het resultaat van de percentiel matching methode is sterk afhankelijk van de keuze van het vergelijkingsstation en van de overheersende windrichting.
2. Voor een realistische correctie van het aantal tropische dagen en hittegolven vóór 1951 in De Bilt is Beek waarschijnlijk geschikter als vergelijkingsstation dan Eelde, omdat tijdens hete zomerperioden de heersende windrichting niet ZW is maar ZO.
3. Een 'beste' vergelijkingsstation dat in alle maanden en bij alle windrichtingen beter is dan de andere bestaat niet. De correctie voor De Bilt hangt af van de temperatuurontwikkeling bij een vergelijkingsstation tussen 1946 en 1955 en die ontwikkeling varieert sterk per vergelijkingsstation.
4. Een bezwaar van Eelde als vergelijkingsstation is dat Eelde voor 1950 relatief koel was in vergelijking met de meeste andere stations. Dit verschil leidt voor De Bilt tot grotere correcties dan bij de keuze van een ander vergelijkingsstation.
5. Een ander bezwaar van Eelde is dat er maar 4 jaar vóór 1950 beschikbaar zijn terwijl blijkt dat vergelijkingsintervallen van verschillende grootten tot aanmerkelijk verschillende homogenisatie resultaten leiden.
6. De uitkomst van de percentiel matching methode op de Tx correctie van De Bilt in de jaren vóór 1950 is, wat de jaargemiddelden betreft, bijna exact gelijk aan het relatieve temperatuurverschil tussen het vergelijkingsstation en De Bilt over 1946-1949.
7. Als Eelde voor de homogenisatie wordt gebruikt dan is de neerwaartse correctie in De Bilt vóór 1950 groot, leidt het tot overcorrectie in de zomermaanden en minder resterende tropische dagen dan als men een ander vergelijkingsstation gebruikt zoals Beek, Aachen of het Duits-Nederlandse ensemble.

¹⁶ <https://www.weerwoord.be/m/2227437>

9 Literatuur

Brandsma, T., 2005: Parallel air temperature measurements at the KNMI-terrain in De Bilt (the Netherlands) May 2003 – April 2005. KNMI-publicatie 207, HISKLIM 7

Brandsma, T., 2011: Parallel air temperature measurements at the KNMI observatory in De Bilt (the Netherlands) May 2003 – June 2005. KNMI Scientific Report, WR 2011-01

Brandsma, T., 2016a: Homogenization of daily temperature data of the five principal stations in the Netherlands (version 1.0). KNMI Technical Report TR-356.

Brandsma, T., 2016b: Homogenisatie van dagelijkse temperaturen van de KNMI hoofdstations. *Meteorologica* 25(2), 4-8.

Brandsma T., Jilderda R. and Sluijter R., 2013: Standardization of data and methods for calculating daily Tmean, Tn and Tx in the Netherlands for the 1901-1970 period.

Bruin, H. de, 2016: De Oude Pagode en een kortebroekenheld. *Meteorologica* 26(3), 26-27.

Dijkstra, F., 2017: Commentaar op de homogenisatie van dagelijkse temperaturen van de KNMI hoofdstations. *Meteorologica* 26(4), 12-15.

Hyndman, R. J. and Y. Fan, 1996: Sample quantiles in statistical packages, *American Statistician* 50, 361–365.

KNAW, 2018: Replication studies. Improving Reproducibility in the Empirical Sciences. Advisory Report KNAW, 15-1-2018.

Kramer, C., Post, J.J. en Woudenberg, J.P.M., 1954: Nauwkeurigheid en betrouwbaarheid van temperatuur- en vochtigheidsbepalingen in buitenlucht met behulp van kwikthermometers. Mededelingen en verhandelingen KNMI, nr. 60.

Mureau, R., W. van den Berg, W. Hazeleger en E. Min, 2013: Vaker hoge maxima? *Meteorologica* 22(3), 26-29

Rozendaal, S., 2016: Creatief met cijfers: KNMI kan geen thermometer lezen. KNMI goochelt de tropische zomer van 1947 van 1 naar 6 in de toptien. Elsevier, 27 oktober 2016.

Rozendaal, S., 2017: De hete pagode van De Bilt. Elsevier, 6 mei 2017

Schrier, G. van der, Ulden, A. van, Oldenborgh, G.J. van, 2011a: The construction of a Central Netherlands temperature. *Clim. Past*, 7, 527–542, 2011.

Schrier, G. van der, Ulden, A. van, Oldenborgh, G.J. van, 2011b: De Centraal Nederland Temperatuurreeks. *Meteorologica* 20(4), 10-15

Ulden van, A., Van Oldenborgh, G.J., Van der Schrier, G., 2009: The Construction of a Central Netherlands Temperature. KNMI Scientific Report WR 2009-03.

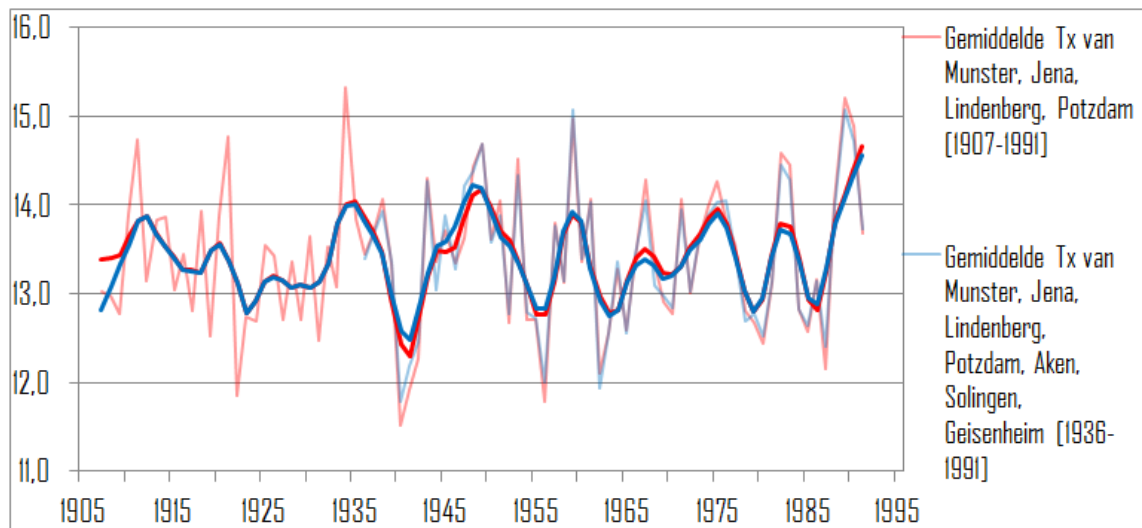
Visser, H., 2007: Kans op extreem warme dagen in Nederland. Milieu en Natuur Planbureau, rapport 550032010/2007.

10 Appendix

BIJLAGE 1: Samenstelling Duits-Nederlands ensemble

Voor het samenstellen van een ensemble van de zeven Duitse stations moeten we rekening houden met het aantal jaren die stations gemeenschappelijk hebben. We kunnen een gemiddelde nemen van 4 stations voor het interval 1907-1991 (Münster, Jena, Lindenberg en Potsdam) en 3 extra stations vanaf 1936 (Aken, Solingen en Geisenheim). Het KNMI doet dat ook bij het samenstellen van de CNT, waarbij van 1901 tot 1950 het gemiddelde van 4 "gecorrigeerde" stations (CNT4) werd gekoppeld aan het gemiddelde van 6 stations (CNT6). Er werd "een kleine correctie" (helaas ongespecificeerd) toegepast voor de overgang van 4 naar 6 referentiestations.

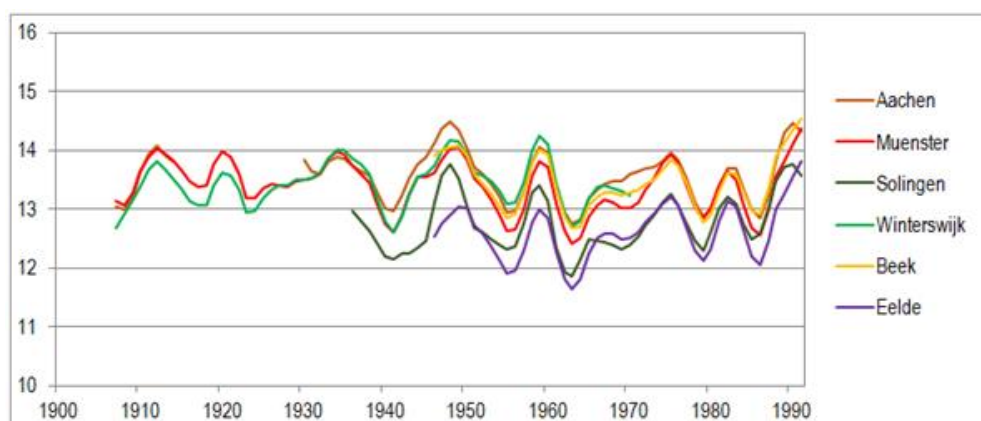
Wij hebben deze methode van het KNMI nagevolgd. In figuur 28 zijn de beide Tx reeksen uitgezet over de periode 1907-1991: rood: 4 stations; blauw dezelfde 4 stations plus 3 extra stations:



Figuur 41 Tx (°C) van 2 subsets bestaande uit de gemiddelden van 4 stations (rood) en van 7 stations (blauw). De dikke lijnen zijn verkregen met een Loess filter met span 8 jaren.

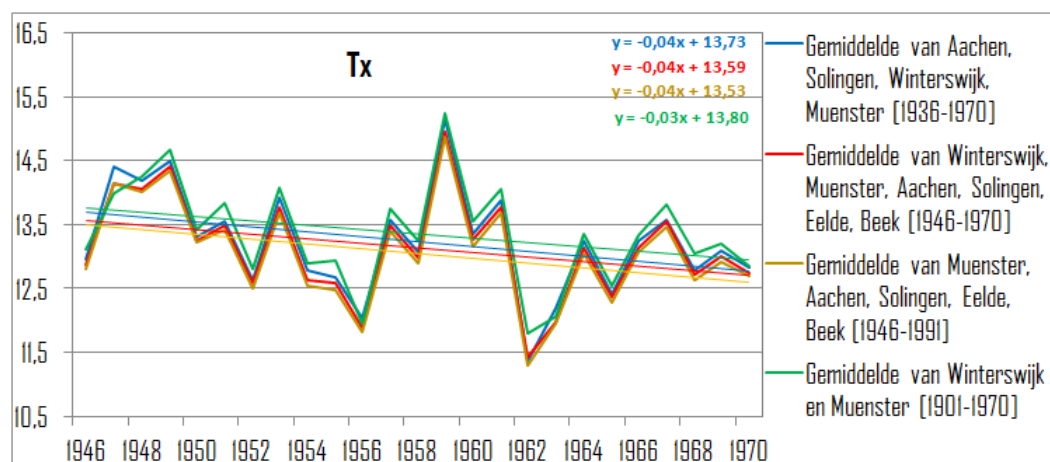
De richtingscoëfficiënt van Tx in het gemeenschappelijk interval 1936-1991 is exact hetzelfde voor het gemiddelde van de 4 stations en het gemiddelde van de 6 stations. Op de gemiddelde Tx van de 4 stations (rood) werd daarom alleen een kleine positieve correctie toegepast (0,049°C) zodat de gemiddelden van de twee reeksen in de kalibratieperiode 1936-1991 gelijk zijn. Vervolgens werden de twee reeksen (rood: 1907-1935 en blauw: 1936-1991) aan elkaar gekoppeld tot één reeks: het Duitse ensemble.

Samenstelling Duits-Nederlandse ensemble



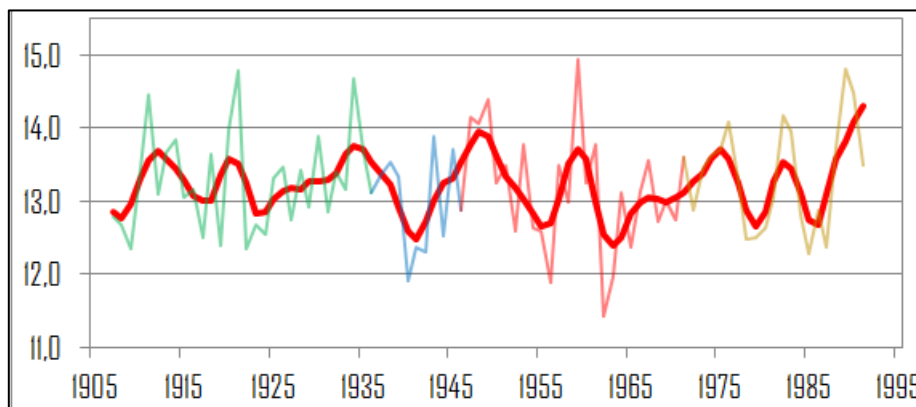
Figuur 42 T_x (°C) van de zes stations die gebruikt werden in het samenstellen van de referentieset, Loess-gefilterd met span 8 jaar.

Om de verschillende perioden van de stations aan elkaar te koppelen gebruiken we het interval dat alle 6 stations gemeenschappelijk hebben: 1946-1970 (de kalibratieperiode). We gebruiken het gemiddelde van alle 6 stations in deze periode om correcties toe te passen op drie subsets, te weten 1: Winterswijk en Münster (1907-1970), 2: Aachen, Solingen, Winterswijk, Münster (1936-1970), 3: Münster, Aachen, Solingen, Eelde, Beek (1946-1991):



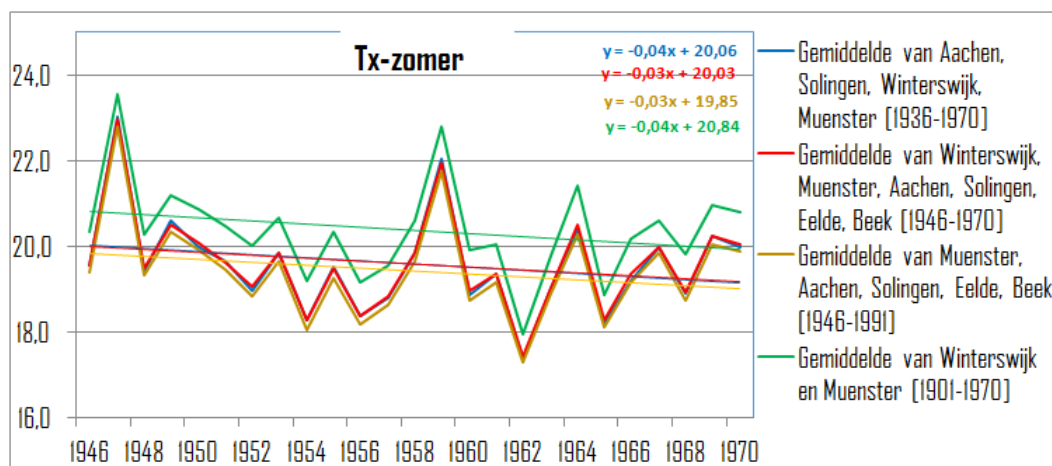
Figuur 43 Gemiddelde T_x (°C) van vier ensembles.

Basis voor de correcties is de rode lijn in figuur 16 met het gemiddelde van alle stations (de hoofdataset). De richtingscoëfficiënten van de 3 subsets verschillen niet significant van die van de hoofdataset ($p > 0,9$). We corrigeren daarom alleen voor het verschil van de gemiddelde T_x van de 3 subsets met die van de hoofdataset. De blauwe lijn (subset 2) ligt gemiddeld $0,102^\circ\text{C}$ boven de rode lijn, de gele lijn (subset 3) ligt gemiddeld $0,076^\circ\text{C}$ onder de rode lijn en de groene lijn (subset 1) ligt gemiddeld $0,23^\circ\text{C}$ boven de rode lijn. Met deze correcties kunnen we nu de 4 reeksen aan elkaar koppelen. Subset 1 (groen) van 1907-1935; subset 2 (blauw) van 1936-1945; subset 3 van 1946-1970; subset 4 (geel) van 1971-1991:



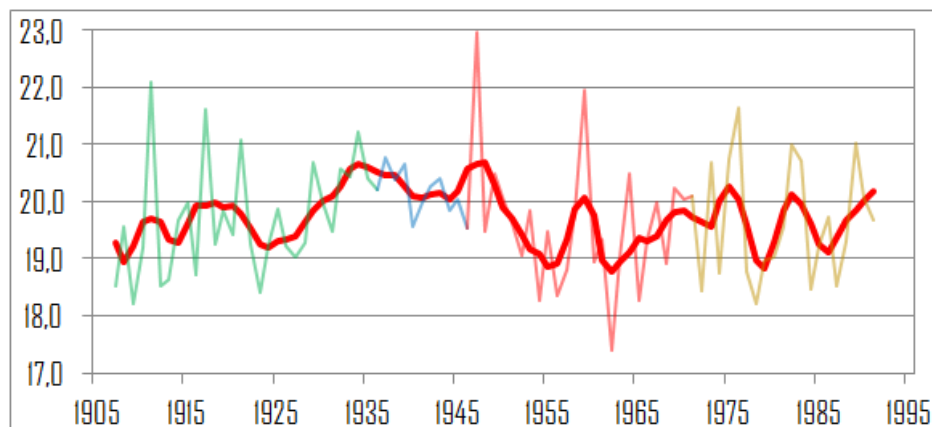
Figuur 44 Gemiddelde Tx (°C) van vier subsets samengevoegd tot een Duits-Nederlandse ensemble. De dikke rode lijn is verkregen met een Loess filter met span 8 jaar.

Voor de constructie van Tx-zomer werd dezelfde methode gevolgd als voor Tx:



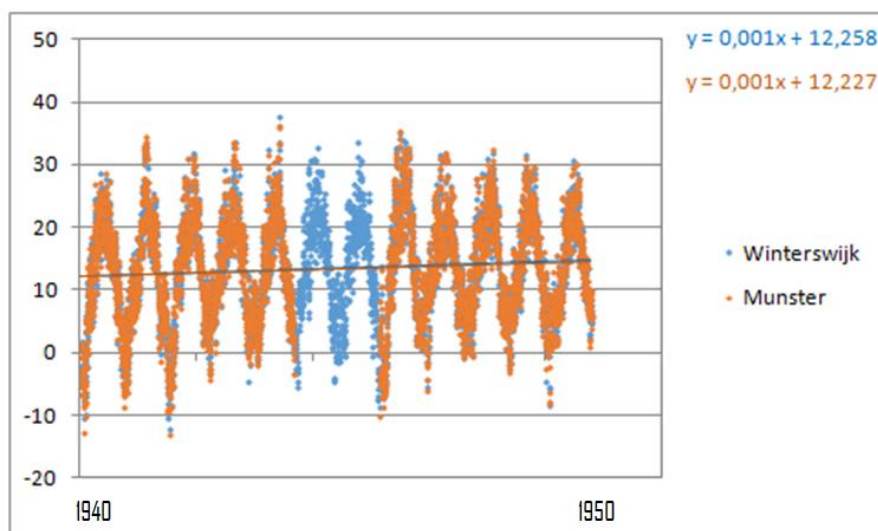
Figuur 45 Gemiddelde Tx-zomer (°C) van vier ensembles.

Aangezien de hellingshoeken van de 4 subsets in het interval 1946-1970 niet significant verschillen corrigeren we alleen voor het verschil van de gemiddelde Tx-zomer van de 3 subsets met die van de hoofdset op dezelfde wijze zoals hierboven beschreven voor Tx. Dit is het resultaat:



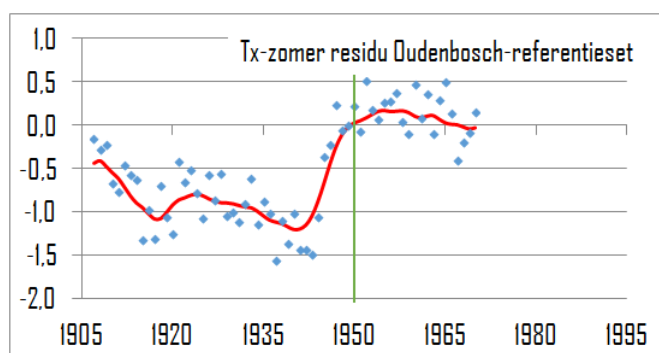
Figuur 46 Gemiddelde Tx-zomer (°C) van vier subsets samengevoegd tot een Duits-Nederlandse ensemble. De dikke rode lijn is verkregen met een Loess filter met span 8 jaar.

Voor Münster ontbreken de jaren 1945 en 1946. Voor het opvullen van deze jaren werden de gegevens van het dichtstbijzijnde station Winterswijk gebruikt. In figuur 19 staan de dagelijkse Tx-waarden van Münster en Winterswijk in het gebruikte kalibratie interval 1940-1950. De Tx-waarden van Münster werden met een kleine correctie voor het niveauverschil van de trendlijnen ($-0,031^{\circ}\text{C}$) gekopieerd naar Winterswijk.



Figuur 47 Dagelijkse Tx-waarden ($^{\circ}\text{C}$) van Winterswijk (blauw) en Münster (bruin) tussen 1940 en 1950.

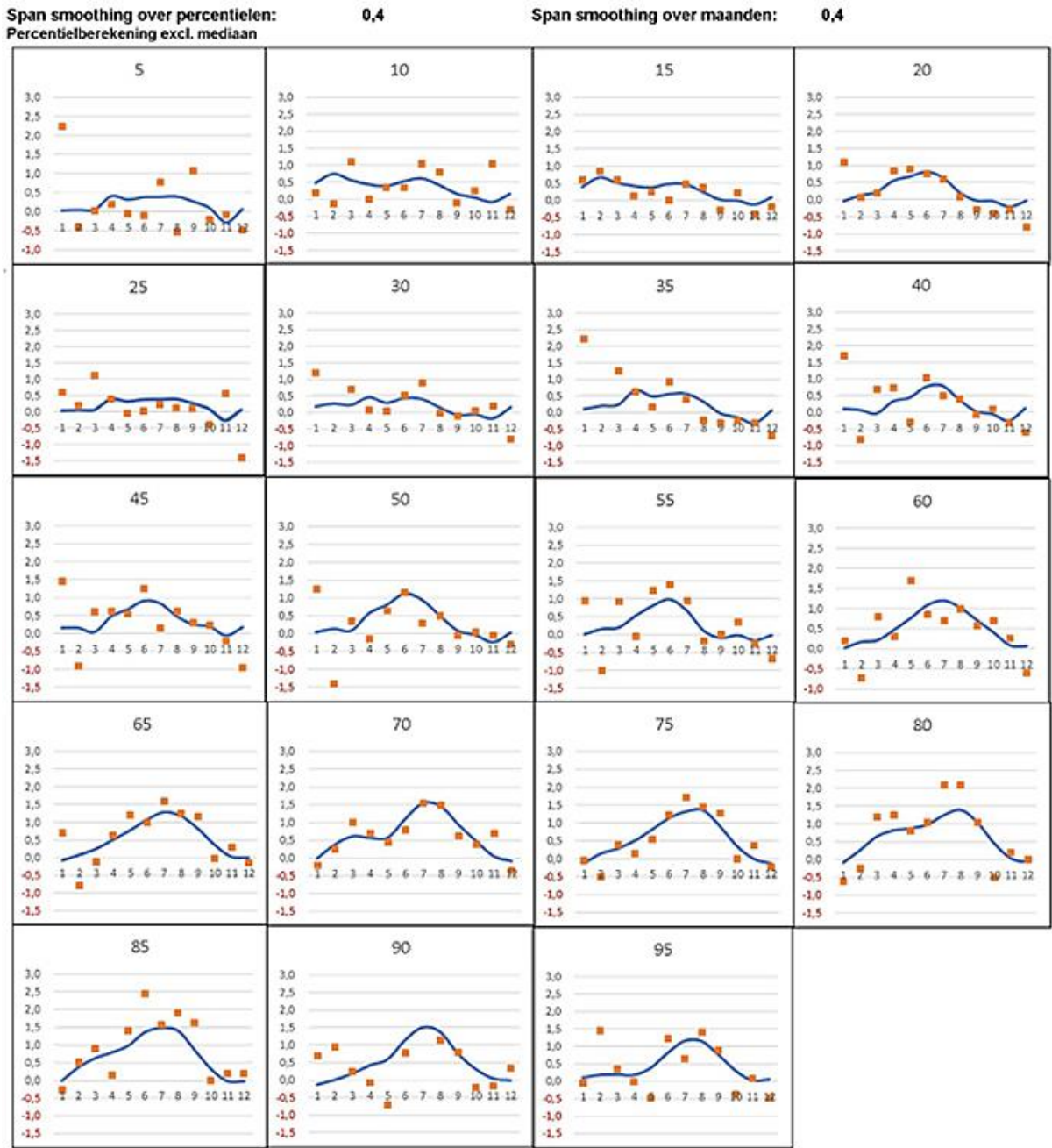
Oudenbosch werd uitgesloten van de referentieset mede vanwege de merkwaardige trendbreuk die Tx-zomer in Oudenbosch vertoont:



Figuur 48 Verschil tussen Tx-zomer van Oudenbosch en de referentieset. Oudenbosch vertoont een grote trendbreuk van meer dan 1°C .

Het KNMI vermeldt deze opvallende trendbreuk niet. Oudenbosch is notabene gebruikt in de constructie van de CNT en is ook gebruikt als referentie bij diverse homogenisaties waardoor deze stations en de CNT zijn "vervuild". Oudenbosch is in de jaren vóór 1950 ruim een graad lager dan de referentieset.

BIJLAGE 2: Detailgrafieken H4 en H5

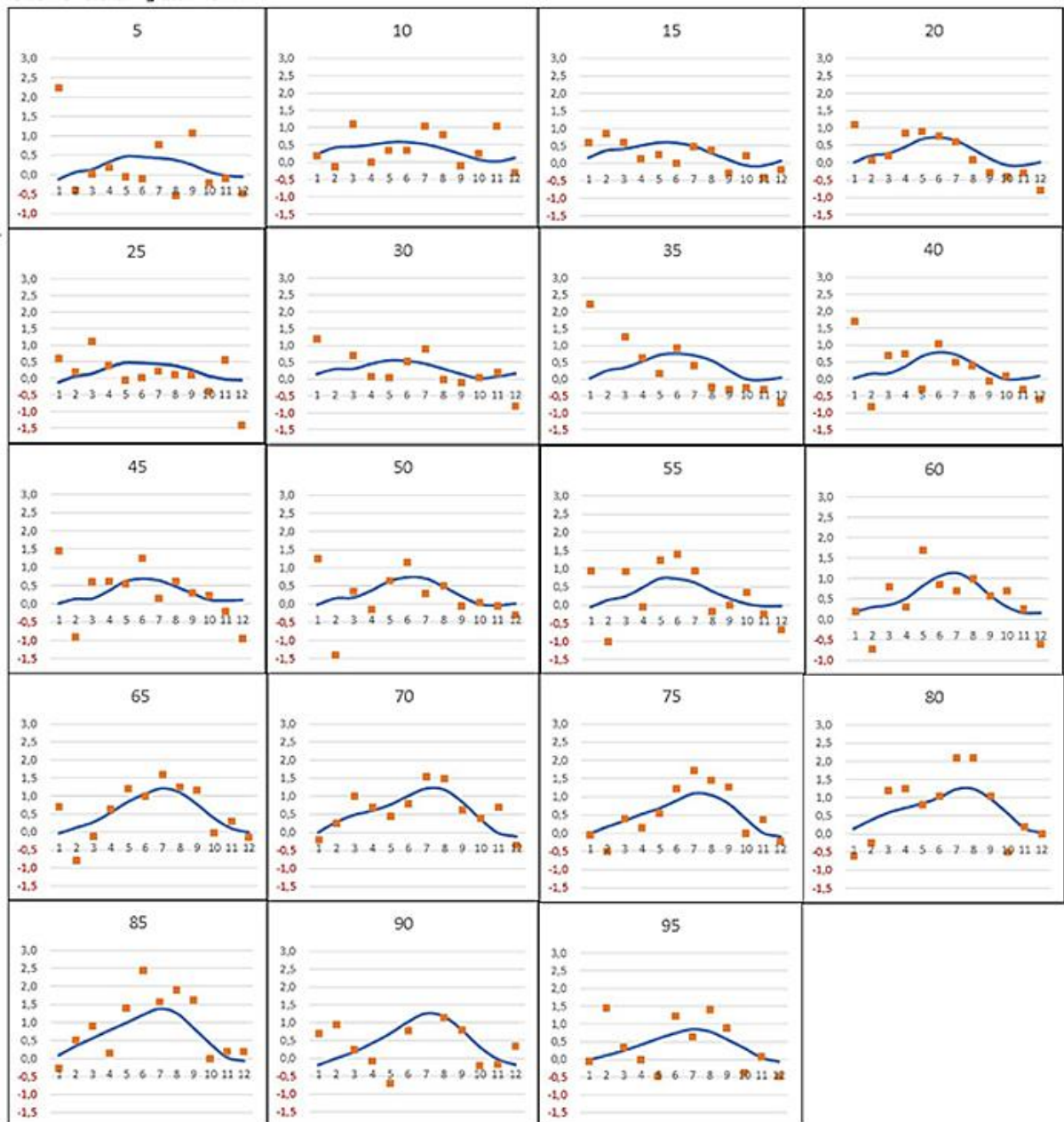


Figuur 49 Correcties Tx De Bilt, per maand en per percentiel, lage smoothing.

Span smoothing over percentielen:
Percentielberekening excl. mediaan

0,6

Span smoothing over maanden: 0,6



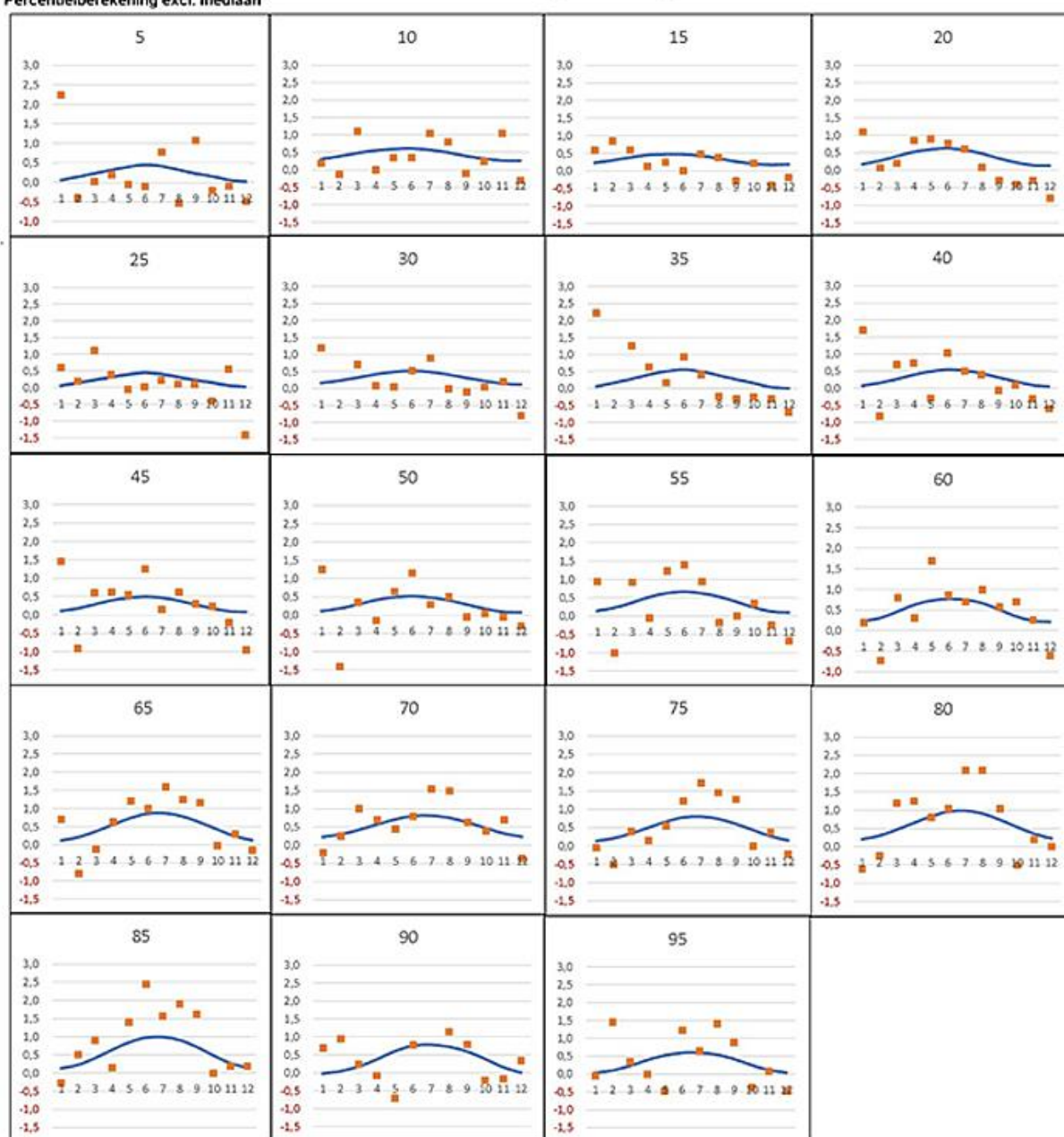
Figuur 50 Correcties Tx De Bilt, per maand en per percentiel, matige smoothing.

Span smoothing over percentielen:
Percentielberekening excl. mediaan

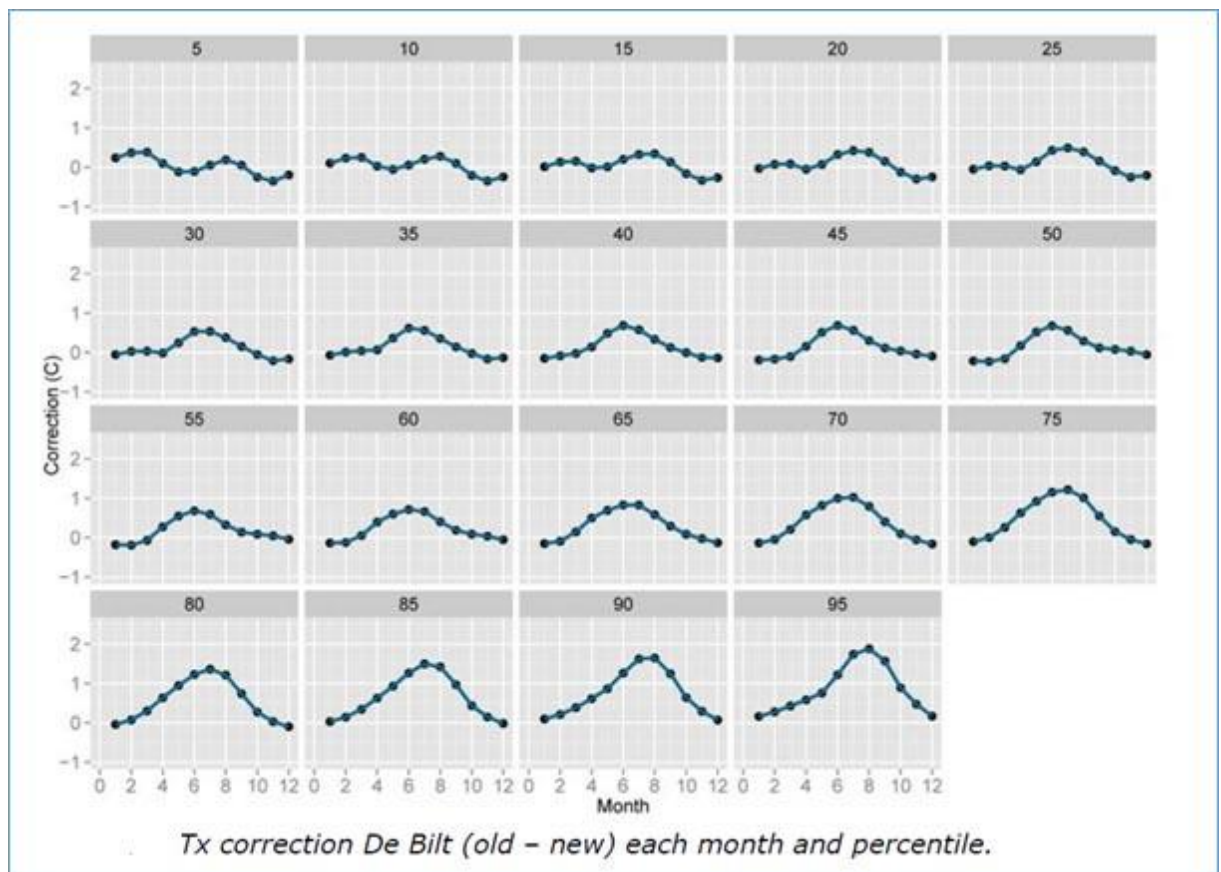
0,9

Span smoothing over maanden:

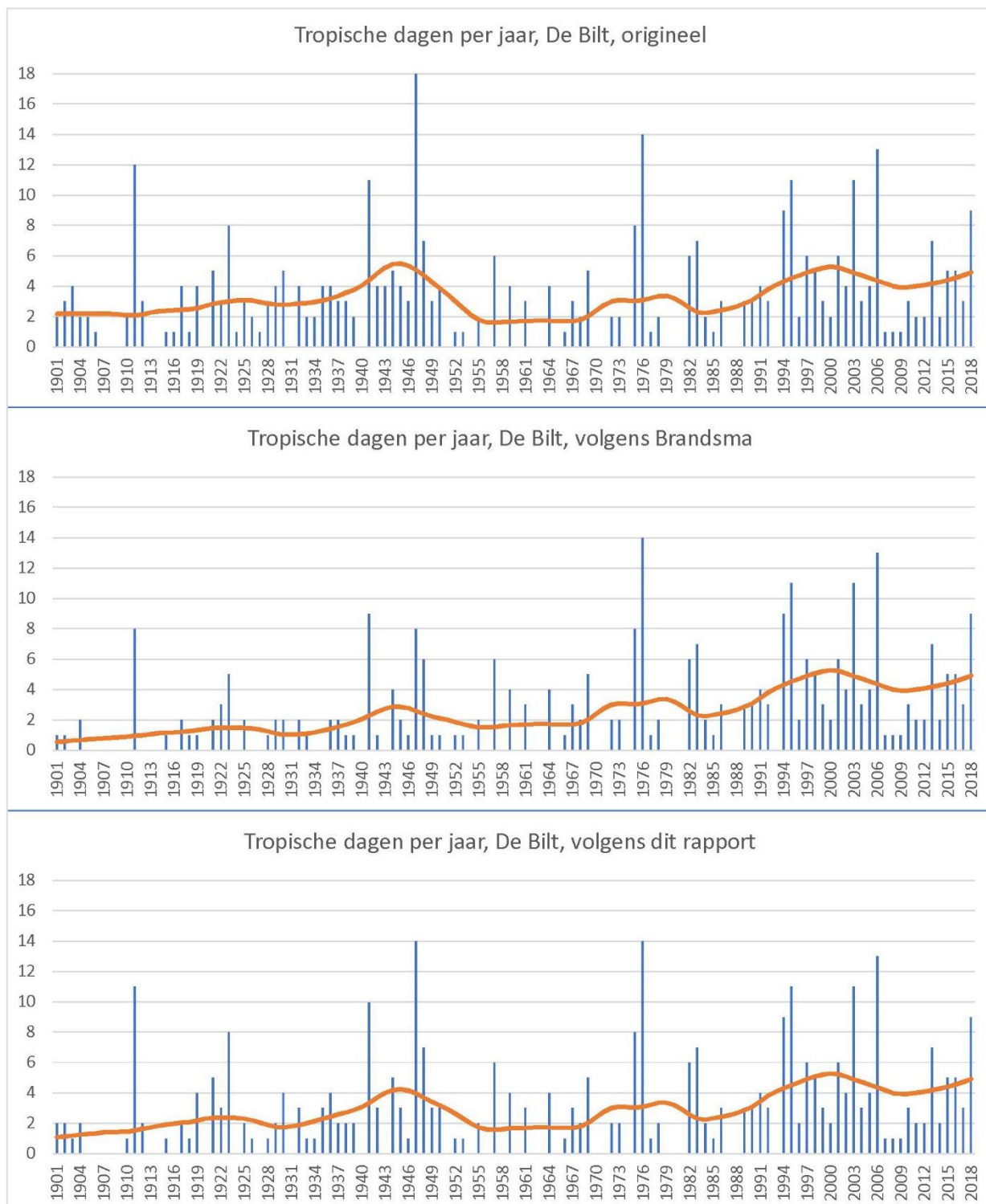
0,9



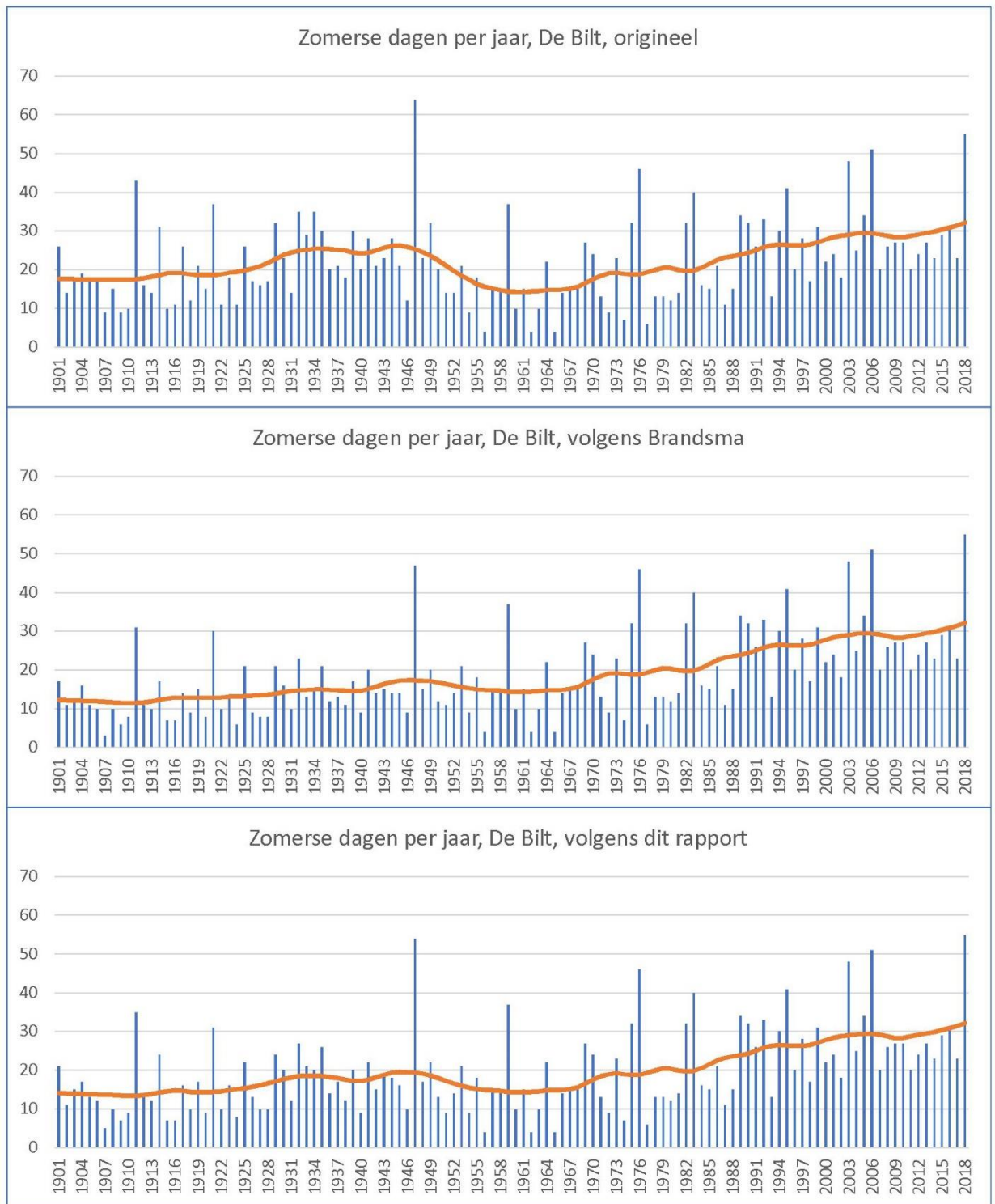
Figuur 51 Correcties Tx De Bilt, per maand en per percentiel, hoge smoothing.



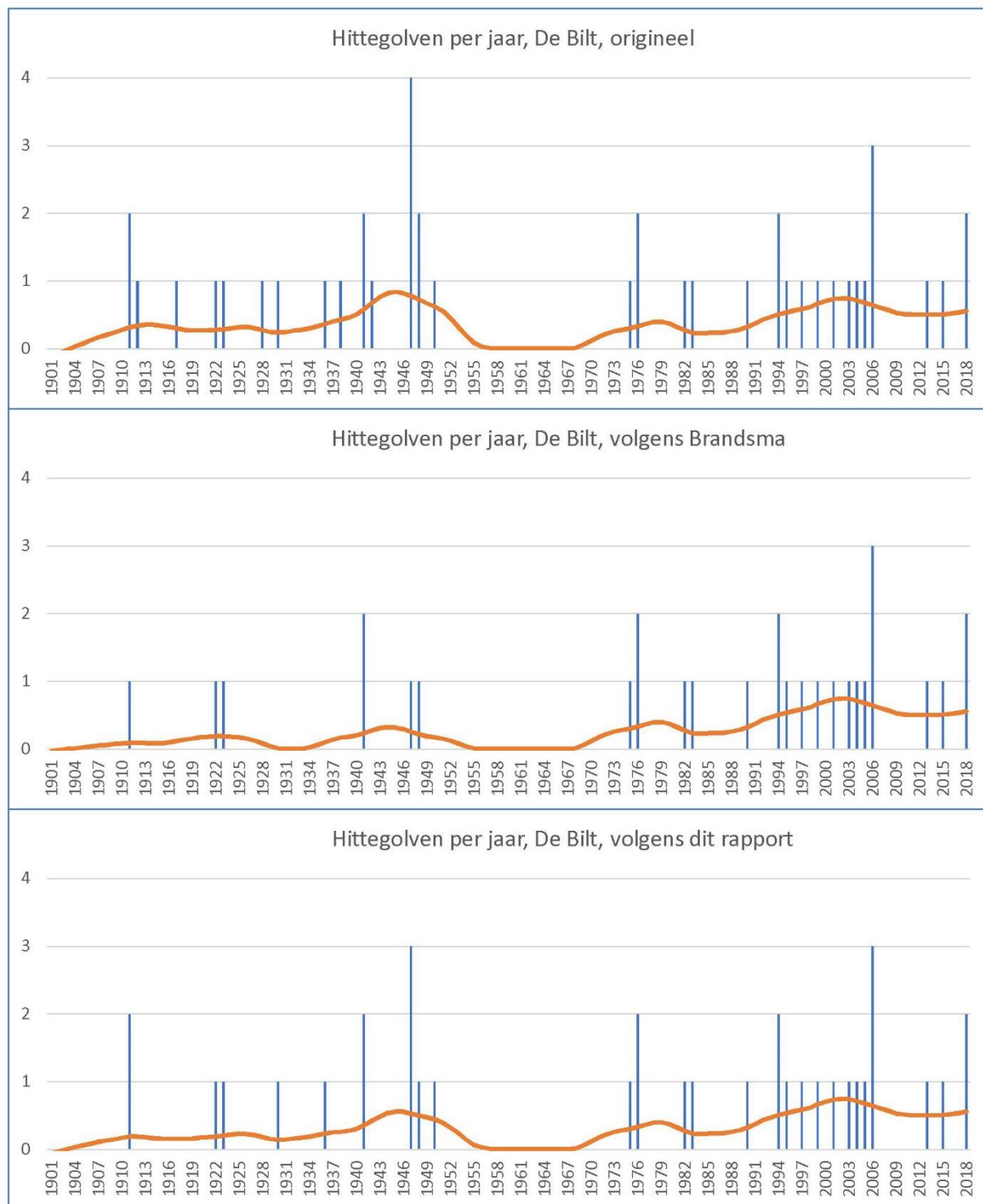
Figuur 52 Correcties Tx De Bilt, per maand en per percentiel volgens Brandsma (2016a), vergelijkbaar met figuur 47.



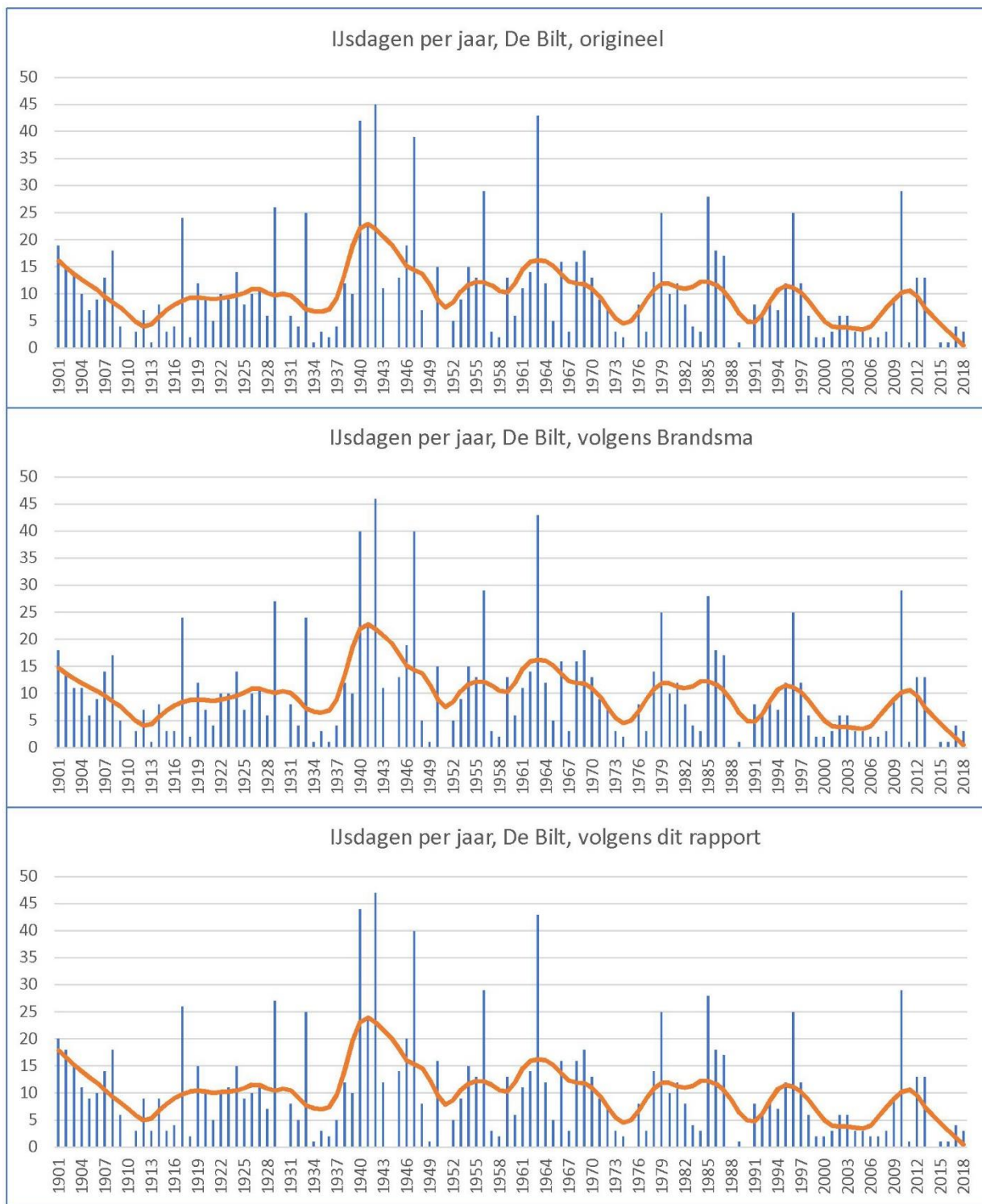
Figuur 53 Aantal tropische dagen per jaar in De Bilt volgens ongehomogeniseerde data, homogenisatie volgens Brandsma en herziene homogenisatie volgens dit rapport. De vloeiende lijn is een Loess-smoothing met een span van 20 punten uit 118 (0,17). De homogenisatie volgens dit rapport is evenals die van Brandsma gebaseerd op de data van Eelde (1946-1949 resp. 1952-1955) als vergelijkingsstation. Gebruikte span bij de smoothings 0,75.



Figuur 54 Aantal zomerse dagen per jaar in De Bilt volgens ongehomogeniseerde data, homogenisatie volgens Brandsma en herziene homogenisatie volgens dit rapport. De vloeiende lijn is een Loess-smoothing met een span van 20 punten uit 118 (0,17). De homogenisatie volgens dit rapport is evenals die van Brandsma gebaseerd op de data van Eelde (1946-1949 resp. 1952-1955) als vergelijkingsstation. Gebruikte span bij de smoothings 0,75.



Figuur 55 Aantal hittegolven per jaar in De Bilt volgens ongehomogeniseerde data, homogenisatie volgens Brandsma en herziene homogenisatie volgens dit rapport. De vloeiende lijn is een Loess-smoothing met een span van 20 punten uit 118 (0,17). De homogenisatie volgens dit rapport is evenals die van Brandsma gebaseerd op de data van Eelde (1946-1949 resp. 1952-1955) als vergelijkingsstation. Gebruikte span bij de smoothings 0,75.

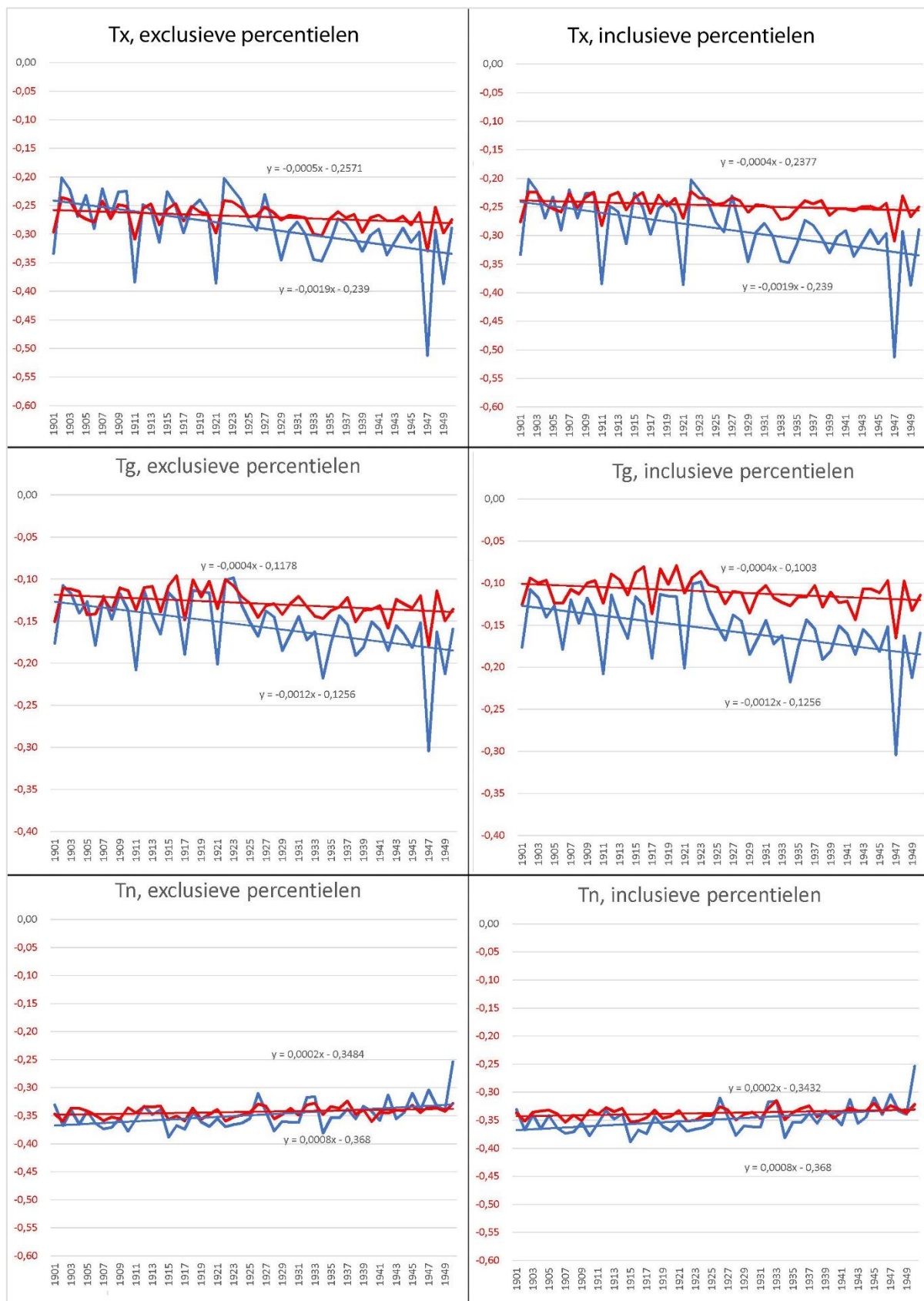


Figuur 56 Aantal ijsdagen dagen per jaar in De Bilt volgens ongehomogeniseerde data, homogenisatie volgens Brandsma en herziene homogenisatie volgens dit rapport. De vloeiende lijn is een Loess-smoothing met een span van 11 punten uit 118 (0,17) om een mogelijke relatie met de zonnecyclus zichtbaar te maken. De homogenisatie volgens dit rapport is evenals die van Brandsma gebaseerd op de data van Eelde (1946-1949 resp. 1952-1955) als vergelijkingsstation. Gebruikte span bij de smoothings 0,75.

TX De Bilt met Eelde als vergelijking, extra middeling over 3 lopende maanden in het begin
 Span smoothing over percentielen 0,60 Span smoothing over maanden: 0,58
 Percentielberekening excl. mediaan



Figuur 57 Correcties Tx De Bilt, per maand en per percentiel, met extra middeling van de percentielen over 3 lopende maanden. Vergelijkbaar met detailfiguur 47.



Figuur 58 Gemiddelde correcties van Tx, Tg en Tn per jaar volgens de homogenisatie van het KNMI (blauw) en volgens onze reconstructie (rood), vergelijkingsperiode van 56 maanden. Links: onze reconstructie gebaseerd op percentielen exclusief de mediaan, rechts: onze reconstructie gebaseerd op percentielen inclusief de mediaan.

BIJLAGE 3: Vragen aan en antwoorden van KNMI

Email naar Theo Brandsma van het KNMI

25 oktober 2018

Geachte heer Brandsma

Ondergetekenden zijn bezig met het reproduceren van de homogenisatie van de dagelijkse temperatuur-metingen in De Bilt van 1 januari 1901 tot 1 september 1951 zoals die door het KNMI is uitgevoerd.

Wij hebben de homogenisatie van Tg en Tx in De Bilt trachten te reproduceren aan de hand van de methode, zoals door Brandsma beschreven in het technisch rapport TR-356. De uitkomsten van die reproductie wijken op een aantal punten sterk af van die van het KNMI. Die afwijkingen kunnen het gevolg zijn van een te summiere beschrijving in het technische rapport van de gevolgde methode. Dat is dan ook de reden om onderstaande vragen te formuleren:

Voor het berekenen van de percentielen wordt verwezen naar Hyndman & Fan (1966), maar in die publicatie staan diverse methoden beschreven.

Bij het berekenen van de ruwe correcties (per percentiel en per maand: het verschil van de verschillen tussen Eelde en De Bilt, voor en na de relocatie in De Bilt) ontstaan data met grote verschillen tussen aangrenzende maanden en aangrenzende percentielen. Dat lost het KNMI op door tweemaal te smoothen, eerst over de percentielen en daarna over de maanden. Er wordt gekozen voor een span van 0,6 zonder deze te motiveren. Vermoedelijk wordt dezelfde span gebruikt bij het smoothen over 19 percentielen als over 12 maanden. Niet vermeld wordt wat het effect is van andere keuzen van de span. Wij hebben diverse combinaties van de span geprobeerd, en geconstateerd dat het uiteindelijke resultaat daar tamelijk sterk door wordt beïnvloed.

Uiteindelijk leidt dit tot correcties per maand en per percentiel. Voor Tx in De Bilt vindt het KNMI in het hoogste percentiel in de zomermaanden correcties van 1,9 °C. Onze herhaling van de homogenisatie leidt bij gebruik van dezelfde span als bij het KNMI tot aanzienlijk lagere correcties: 0,9-1,3 °C in de drie hoogste percentielen. Ook opmerkelijk is dat de hoogste correcties die wij vinden niet in de hoogste percentielen zijn, maar in het 85-percentiel.

Ook tussen onze homogenisatie en de door het KNMI gehomogeniseerde gemiddelde temperatuur (Tg) vinden wij verschillen. Zie figuur 1. Over het algemeen zijn het geen grote verschillen, behalve in 1911, 1921, 1934, 1947 en 1949. Vooral in 1947 corrigeert het KNMI Tg veel sterker dan uit onze reconstructie volgt. Bovendien is de trend in de correcties in de periode 1901-1951 bij het KNMI sterker dalend dan onze homogenisatie.

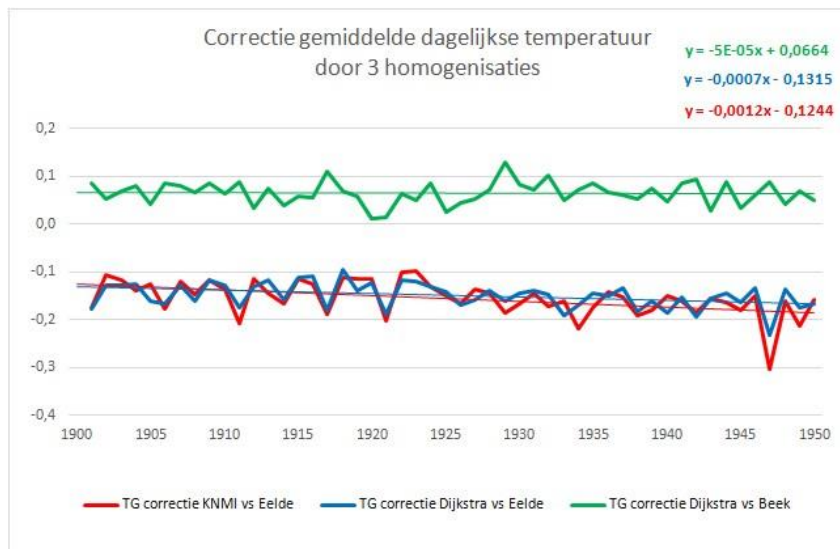
Het KNMI heeft besloten om de data voor De Bilt te homogeniseren met Eelde als referentiestation. Waarom voor Eelde is gekozen is in het technisch rapport niet verantwoord. Wij hebben dezelfde homogenisatietechniek ook toegepast met als referentiestation Beek, om te zien of deze keuze van invloed is op de uitkomsten. Dat was inderdaad het geval: de resultaten bij het gebruik van Beek wijken sterk af van die van Eelde. Zie figuur 1, 2 en 3.

Dit alles leidt tot een aantal vragen waar we graag een antwoord op zouden willen krijgen:

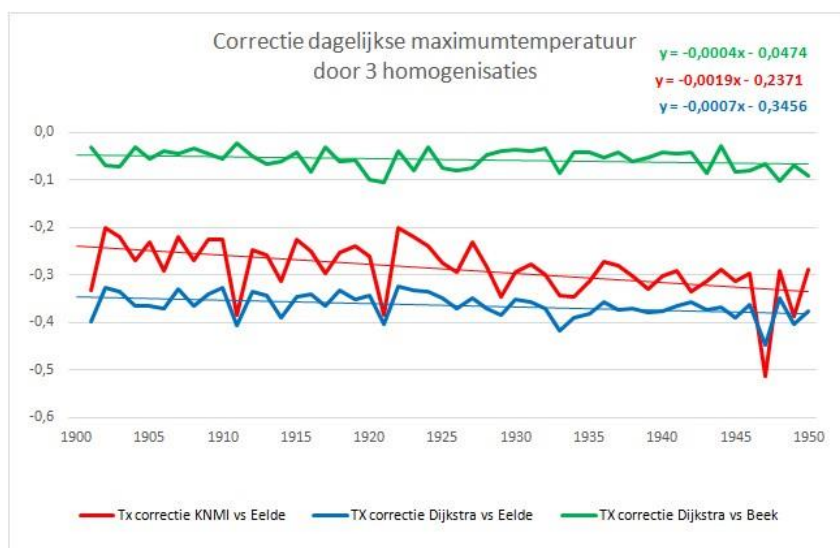
- 1) Welke methode van Hyndman & Fan is gebruikt om percentielen te berekenen, en waarom is voor die methode gekozen?
- 2) Waarop berust de keuze voor een span van 0,6? Is onderzocht wat het effect op de homogenisatie is bij andere keuzen van de span? Zo ja, wat kwam daaruit? Zo ja, waarom zijn deze resultaten niet gepubliceerd in TR-356?
- 3) Welke aanvullende correcties heeft het KNMI nog meer toegepast dan die omschreven worden in het technische rapport? Kunnen die verklaren waarom wij lagere correcties vinden in het hoogste percentiel dan het KNMI?
- 4) Zijn er andere stations dan Eelde gebruikt als vergelijkingsstation voor de homogenisatie van de data van De Bilt? Zo ja, welke en wat kwam daaruit? Waarom is dan toch voor Eelde gekozen?
- 5) Wat is de statistische fout in de berekende correcties?
- 6) Er zijn vele veranderingen geweest in instrumentarium, meetschema's en oprukkende bebouwing in de afgelopen eeuw. Waarom is nu juist de opstellingsverandering (pagode in de beschutte tuin -> Stevenson naar achterterrein) in 1950/1951 geselecteerd om een homogenisatie door te voeren in de reeks van de Bilt? Hoe (met welke statistische test) heeft het KNMI aangetoond dat deze overgang in 1950/1951 wel gehomogeniseerd moet worden en andere wijzigingen in het verleden niet?
- 7) Waarom is de toegepaste wetenschappelijke methode niet gepubliceerd in een peer-reviewed tijdschrift en alleen in een TR-rapport?
- 8) Waarom heeft het KNMI na de homogenisatie de ongecorrigeerde meetdata offline gehaald? Zouden niet, vanuit het oogpunt van transparantie, zowel de ongecorrigeerde als de gecorrigeerde data online moeten blijven staan? Is het KNMI bereid de ongecorrigeerde data weer online te zetten?

We danken u bij voorbaat voor uw medewerking.

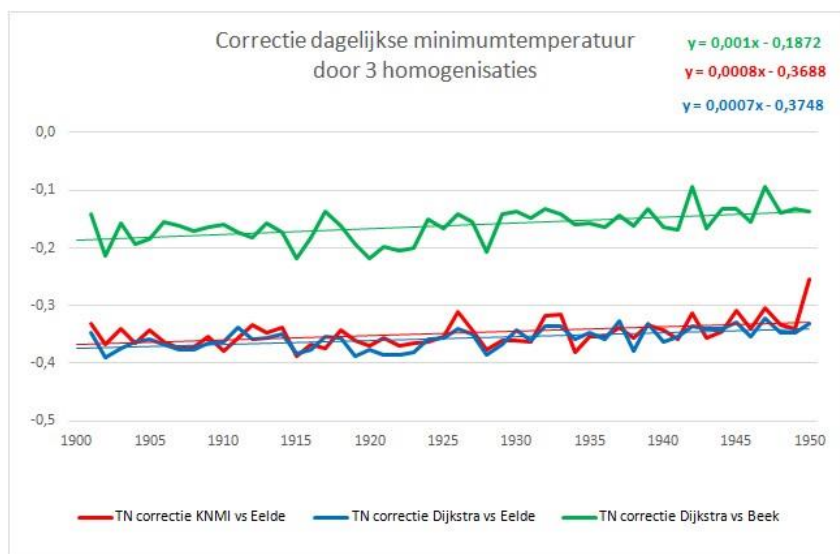
Frans Dijkstra, Rob de Vos, Marcel Crok, Jan Ruis, Henk de Bruin



Figuur 59 Correctie Tg De Bilt door 3 homogenisaties



Figuur 60 Correctie Tx De Bilt door 3 homogenisaties



Figuur 61 Correctie Tn De Bilt door 3 homogenisaties

De Bilt, 9 november 2018

Antwoorden op de vragen van Crok et al. (zoals verwoord in de mail van 25 oktober 2018)

1) *Welke methode van Hyndman & Fan is gebruikt om percentielen te berekenen, en waarom is voor die methode gekozen?*

We gebruiken R-software voor onze berekeningen. In R gebruiken we de functie `quantile` voor het berekenen van kwantielen. De documentatie zegt het volgende over deze functie:

“Quantile returns estimates of underlying distribution quantiles based on one or two order statistics from the supplied elements in `x` at probabilities in `probs`. One of the nine quantile algorithms discussed in Hyndman and Fan (1996), selected by type, is employed.

All sample quantiles are defined as weighted averages of consecutive order statistics. Sample quantiles of type `i` are defined by:

$$Q[i](p) = (1 - \gamma) x[j] + \gamma x[j+1],$$

where $1 \leq i \leq 9$, $(j-m)/n \leq p < (j-m+1)/n$, $x[j]$ is the j th order statistic, n is the sample size, the value of γ is a function of $j = \text{floor}(np + m)$ and $g = np + m - j$, and m is a constant determined by the sample quantile type.”

De tekst hierboven noemt negen algoritmen/methoden. Hier hebben we type 7 gebruikt met:

“ $m = 1-p$. $p[k] = (k - 1) / (n - 1)$. In this case, $p[k] = \text{mode}[F(x[k])]$.”

Dit is de default methode in R. Er was geen reden om daarvan af te wijken.

2) *Waarop berust de keuze voor een span van 0,6? Is onderzocht wat het effect op de homogenisatie is bij andere keuzen van de span? Zo ja, wat kwam daaruit? Zo ja, waarom zijn deze resultaten niet gepubliceerd in TR-356?*

Onderaan p.20 in TR-356 staat:

The smoothing is done using a so-called loess smoother (Cleveland, 1979). Using a mixture of cross-validation and expert opinion we determined a span of 0.6 (family = Gaussian and degree=2).

De keuze van de span is dus op basis van cross validation (generalized cross validation in dit geval). Voor afzonderlijke kwantielen en maanden geeft dat een iets andere span. We hebben gekozen om met één span te werken. De span van 0.6 bleek daarbij een goed compromis. Op het oog is het effect van een kleine verandering in span klein. We hadden daarom geen reden om het effect van de span op de gehele homogenisatie verder te onderzoeken.

3) *Welke aanvullende correcties heeft het KNMI nog meer toegepast dan die omschreven worden in het technische rapport? Kunnen die verklaren waarom wij lagere correcties vinden in het hoogste percentiel dan het KNMI?*

Het KNMI heeft geen aanvullende aanpassingen gedaan. Uitgangspunt zijn de gestandaardiseerde dagelijkse temperaturen van de vijf hoofstations zoals beschreven in TR-340.

Het rapport bevat voor De Bilt niet voldoende detail om de resultaten precies te reproduceren. Ten eerste, in een poging om voor De Bilt zoveel mogelijk data mee te nemen, hebben we voor sommige

maanden ook gebruik gemaakt van data in 1950 (t/m augustus, het moment voor de breuk door de overgang pagode/stevenson). Om voor en na de breuk intervallen van gelijke lengte te krijgen is de data na de breuk waar nodig aangevuld met data uit 1956. Ten tweede, vanwege de grote verschillen tussen aangrenzende maanden (zie figuur 4 uit TR-356) hebben we de berekening van de kwantielen voor elke maand steeds gebaseerd op lopende drie maanden. Dus voor de berekening van de januari kwantielen gebruiken we DJF, voor februari, JFM, etc. Dit zorgt voor een conservatieve schatting van de aanpassingen. Voor het tussenjaar (tussen de overgang naar Stevensonhut en de verplaatsing naar een open locatie 300 m naar het zuiden) is gebruik gemaakt van de helft van de aanpassing.

Onderstaande tabel geeft aan welke data voor welke maand gebruikt is. Daar waar maanden weggelaten zijn (omdat ze in het inhomogene interval vallen) is dat aangegeven.

Month	Before	After
jan	1946-50 (omit D 50)	1952-56 (omit D 56)
feb	1946-50	1952-56
mar	1946-50	1952-56
apr	1946-50	1952-56
may	1946-50	1952-56
jun	1946-50	1952-56
jul	1946-50	1952-56
aug	1946-50 (omit S 50)	1952-56 (omit S 56)
sep	1946-50 (omit SN 50)	1952-56 (omit SN 56)
oct	1946-49	1952-55
nov	1946-49	1952-55
dec	1946-50 (omit ND 50)	1952-56 (omit ND 56)

Achteraf gezien was het verstandig geweest deze detaillering als appendix op te nemen het rapport. Hopelijk lukt het met bovenstaande gegevens wel om de resultaten te reproduceren.

4) *Zijn er andere stations dan Eelde gebruikt als vergelijkingsstation voor de homogenisatie van de data van De Bilt? Zo ja, welke en wat kwam daaruit? Waarom is dan toch voor Eelde gekozen?*

Alleen Eelde is gebruikt als vergelijkingsstation voor de correctie van De Bilt. De twee kuststations Den Helder/De Kooy en Vlissingen vallen af door hun ligging dicht bij zee. Hierdoor zijn de stations op dagbasis slecht vergelijkbaar met De Bilt. Van de overige twee stations is Eelde beter geschikt als parallel station dan Beek (net als Eelde is er met Beek een overlap van 1946-1956). Dit heeft te maken met de afwijkende hoogteligging van Beek (114 m boven nap) en de directe nabijheid van het Maasdal. Daarnaast is voor de overheersende windrichting in Nederland (ZW-W) de afstand van De Bilt en Eelde tot grote wateroppervlakken vergelijkbaar.

Ook wat het aantal hittegolven betreft komt Eelde sterk overeen met De Bilt. In de gehomogeniseerde reeksen tellen beide stations 24 hittegolven t/m 2015. Eelde telt 9 hittegolven in de periode t/m 1950 tegenover 7 in De Bilt. Na 1950 telt Eelde 15 hittegolven tegenover 17 in De Bilt.

5) *Wat is de statistische fout in de berekende correcties?*

In de literatuur is weinig bekend over statistische onnauwkeurigheid bij aanpassingen in het kader van homogenisatie van reeksen van dagwaarden. Dit is een moeilijk probleem vanwege het gebruik van korte reeksen met autocorrelatie en de correlatie tussen de reeksen onderling. Recent is

wetenschappelijk onderzoek gestart naar het kwantificeren van onnauwkeurigheid (o.a. INDECIS (<http://www.indecis.eu/>)).

Op dit moment houden we rekening met een fout van een paar tiende graden op de grootste correcties.

6) *Er zijn vele veranderingen geweest in instrumentarium, meetschema's en oprukkende bebouwing in de afgelopen eeuw. Waarom is nu juist de opstellingsverandering (pagode in de beschutte tuin -> Stevenson naar achterterrein) in 1950/1951 geselecteerd om een homogenisatie door te voeren in de reeks van de Bilt? Hoe (met welke statistische test) heeft het KNMI aangetoond dat deze overgang in 1950/1951 wel gehomogeniseerd moet worden en andere wijzigingen in het verleden niet?*

Voorafgaand aan de homogenisatie is voor de vijf hoofdstations in 2013 een standaardisatie van de dagtemperaturen doorgevoerd beschreven in TR-340. Door de standaardisatie zijn de procedures om dag minima en maxima te berekenen gestandaardiseerd en wordt er gebruik gemaakt van uniforme databronnen.

Verschillen tussen de verschillende versies van de Stevensonhut en de moderne KNMI schotelhut zijn eerder onderzocht door het KNMI (Brandsma en Van der Meulen, 2008). Deze verschillen zijn bijna een orde kleiner dan die nu onderzochte verschillen.

Voor De Bilt is het effect van verstedelijking onderzocht (Brandsma et al., 2003). Ook dit effect is verwaarloosbaar op de totale temperatuurtrend in De Bilt.

Van De Bilt is vooral de inhomogeniteit rond 1950 in de maximumtemperatuur in de zomer bekend. Deze is makkelijk aan te tonen met willekeurige statistische test. Visser (2007) bijvoorbeeld analyseerde de breuk in Txx en vond een sprong van 3.7 ± 1.4 °C (95%-betrouwbaarheidsinterval). De grootte van de breuk is overigens wel afhankelijk van het model dat je fit voor en na de breuk. Het KNMI komt met een ander model op correcties die ongeveer de helft kleiner zijn. De breuk is ook duidelijk zichtbaar in de in de gemiddelde maandtemperaturen (zie bijv. Schrier et al., 2011).

Zoals in TR-356 aangegeven, zijn er nog verbeteringen mogelijk in de huidige homogenisatie. We verwachten echter niet dat dat zal leiden tot grote veranderingen in de belangrijkste temperatuurtrends.

7) *Waarom is de toegepaste wetenschappelijke methode niet gepubliceerd in een peer-reviewed tijdschrift en alleen in een TR-rapport?*

Voor publicatie in een peer-reviewed tijdschrift is het nodig dat onderzoek vernieuwend is. Dat is hier niet het geval. De percentile-matching methode is al eerder gepubliceerd en toegepast.

8) *Waarom heeft het KNMI na de homogenisatie de ongecorrigeerde meetdata offline gehaald? Zouden niet, vanuit het oogpunt van transparantie, zowel de ongecorrigeerde als de gecorrigeerde data online moeten blijven staan? Is het KNMI bereid de ongecorrigeerde data weer online te zetten?*

Gemeten waarden blijven naast de gehomogeniseerde beschikbaar en zijn uitgangspunt voor een eventuele volgende (verbeterde) homogenisatieslag. Via <https://www.knmi.nl/nederland-nu/klimatologie/gemeten-reeksen> zijn zowel de gehomogeniseerde reeksen als de gemeten reeksen beschikbaar. Ieder die dat wil kan gebruik maken van een of beide soorten reeksen. Het KNMI zelf gebruikt voor de voorlichting de gehomogeniseerde reeksen. Voor de relevante temperatuurindicatoren geven deze reeksen hetzelfde consistente beeld.

Literatuur

Brandsma, T., Homogenization of daily temperature data of the five principal stations in the Netherlands (version 1.0). KNMI Technical Report: TR-356, pp43, 2016.

Brandsma, T., R. Jilderda and R. Sluijter. Standardization of data and methods for calculating daily Tmean, Tn and Tx in the Netherlands for the 1901-1970 period. KNMI, Technical Report, TR-340, 45p., 2013.

Brandsma, T., G.P. Können and H.R.A. Wessels. Empirical estimation of the effect of urban heat advection on the temperature series of De Bilt (The Netherlands). *Int. J. Climatology*, 23, 829-845, 2003.

Brandsma, T. and J.P. van der Meulen. Thermometer Screen Intercomparison in De Bilt (the Netherlands), Part II: Description and modeling of mean temperature differences and extremes. *Int. J. Climatology*, 28, 389-400, 2008. doi:10.1002/joc.1524.

Schrier, G. van der, A.P. van Ulden and G.J. van Oldenborgh. The construction of a Central Netherlands temperature. *Climate of the Past*, 7, 527-542, 2011.

Visser, H. Kans op extreem warme dagen in Nederland: Een analyse van historische data, modelvoorspellingen en consequenties voor de volksgezondheid. MNP Rapport 550032010/2007, 2007.

BIJLAGE 4: Smoothing voor dummies

Smoothing is een statistische techniek waarmee fluctuaties in een reeks metingen worden afgevlakt, zodat de onderliggende trend zichtbaar wordt gemaakt.

De eenvoudigste manier om een trend zichtbaar te maken is de **regressielijn**. Dat is een rechte lijn die het best aansluit bij een reeks metingen. De vergelijking van zo'n regressielijn is een lineaire expressie van de vorm $y = ax + b$. De mate waarin de meetpunten gespreid zijn rond de lijn wordt uitgedrukt in de regressiecoëfficiënt. Meestal wordt het kwadraat daarvan vermeld: r^2 .

Vaak is een onderliggende trend niet lineair, maar is wel een gekromde lijn te vinden die beter aansluit bij een reeks metingen. De vergelijking van die lijn kan kwadratisch zijn, of derdegraads, of hoger. Men kan dan spreken van een **polynoomfunctie**. Andere varianten van gekromde lijnen kunnen een logaritmisch of exponentieel verband weergeven.

Als er geen wiskundig verband is binnen een meetreeks kan een fluctuerende trend zichtbaar worden gemaakt met twee technieken: **voortschrijdend gemiddelde** of **smoothing**.

Het **voortschrijdend gemiddelde** is een lopend gemiddelde van een aantal metingen. Bij elk meetpunt wordt de gemiddelde waarde genomen van die meting en een aantal aan het meetpunt voorafgaande waarden en een even groot aantal er op volgende waarden. Het nadeel van een voortschrijdend gemiddelde is dat het niet aan het begin en het einde van de reeks kan worden berekend, omdat er geen metingen voor en na het einde van de reeks zijn. De mate waarin fluctuaties in de reeks worden afgevlakt hangt af van het aantal punten dat in het voortschrijdend gemiddelde wordt meegerekend.

Voor **smoothing** bestaan diverse statistische methoden. De **loess-smoothing** maakt gebruik van een soort regressie rondom elk meetpunt (nadere uitleg is niet voor dummies). Het resultaat lijkt op een voortschrijdend gemiddelde, maar met een vloeiender curve van begin tot eind van de meetreeks. Net als bij het voortschrijdend gemiddelde hangt de mate waarin een fluctuaties in de reeks worden afgevlakt af van het aantal punten dat wordt meegerekend. Dat wordt de **span** genoemd. Het hangt van het statistische pakket af, in welke eenheid de span wordt uitgedrukt, meestal het aantal punten dat wordt meegerekend (de span is dan een heel getal), of de fractie van het aantal meegerekende punten van het totaal aantal punten (de span ligt dan tussen 0 en 1). In dit rapport wordt de span uitgedrukt in fracties van 0 tot 1.

Voorbeelden van smoothing.

De hierna volgende grafieken geven het aantal tropische dagen per jaar in De Bilt, volgens de ongehomogeniseerde reeks met verschillende trendlijnen volgens de hierboven beschreven technieken.

