

# **Klimaatmodellering**

Een blik op de toekomst

**Jasper Wijnands**

Student Bedrijfswiskunde & Informatica

**BWI Werkstuk**

Begeleider: Sandjai Bhulai

Afgerond op 9-maart-2007

## Management samenvatting

De probleemstelling die ik in mijn scriptie onderzoek is: hoe wordt met modellen onderzoek gedaan naar mogelijke klimaatveranderingen?

In deze scriptie zal ik verschillende klimaatmodellen in detail beschouwen. Deze klimaatmodellen zijn Global Climate Models (GCMs), Earth Models of Intermediate Complexity (EMICs) en weersvoorspellingsmodellen voor de korte termijn (High Resolution Limited Area Model) en de middellange tot lange termijn (Ensemble Prediction System).

Modellen die in de categorie GCM vallen, zijn over het algemeen geschikt voor voorspellingen van minder dan een uur tot enkele eeuwen in de toekomst. In deze modellen zijn veel klimatologische processen in detail uitgewerkt. Voor het draaien van deze modellen worden de modernste supercomputers gebruikt. De voorspellingstermijn loopt toch slechts tot enkele eeuwen door een beperkte rekencapaciteit. Om ook op een langere termijn voorspellingen te kunnen doen, zijn EMICs ontwikkeld. Bij deze modellen worden veel klimatologische processen in minder detail geprogrammeerd of zelfs weggelaten. De voorspellingshorizon van deze modellen loopt van enkele jaren tot 100.000 jaar. Voor de zeer korte termijn bestaan er gespecialiseerde weersvoorspellingsmodellen. Een voorbeeld hiervan is het Ensemble Prediction System (EPS). Het EPS valt in de categorie van GCMs en berekent 51 scenario's op basis van verschillende beginsituaties. De voorspellingstermijn loopt van 48 uur tot 10 dagen in de toekomst. Voor de voorspelling tot 48 uur gebruikt het KNMI het High Resolution Limited Area Model (HIRLAM). Een belangrijk aspect van dit model is dat niet de hele aarde gemodelleerd wordt, terwijl de resolutie juist wel hoog is.

Als eigen onderzoek heb ik de gemiddelde dagtemperatuur van morgen proberen te voorspellen met twee statistische modellen. Het betreft hier een regressiemodel en een tijdreeksmodel. De voorspellingen van deze modellen heb ik vergeleken met de voorspelling dat de gemiddelde dagtemperatuur van morgen gelijk is aan de gemiddelde dagtemperatuur van vandaag. Beide modellen slagen er niet in om significant betere voorspellingen te doen dan deze benchmark voorspelling.

Conclusie is dat klimaatmodellen voor zowel de korte termijn als de lange termijn erg complex zijn. In de meest nauwkeurige modellen wordt de gehele aarde driedimensionaal gemodelleerd, waarbij diverse klimatologische processen in detail worden gemodelleerd. Vervolgens wordt met deze modellen gesimuleerd wat het weer of klimaat zou kunnen worden in de toekomst.

Uit het eigen onderzoek kan ik concluderen dat complexe klimaatmodellen absoluut nodig zijn om nauwkeurige weersvoorspellingen te kunnen doen. Verder is één van de kerntaken van het KNMI om waarschuwingen af te geven bij extreme weerssituaties. Deze zal je met eenvoudige statistische modellen niet kunnen voorspellen, terwijl de complexe klimaatmodellen hier juist wel geschikt voor zijn.

## Executive summary

The formulation of the problem I will examine in my essay is: how are models used in research into possible climate changes?

In this essay, I will present a detailed view on several climate models. These climate models are Global Climate Models (GCMs), Earth Models of Intermediate Complexity (EMICs), and weather forecast models for the short term (High Resolution Limited Area Model) and the medium up to long term (Ensemble Prediction System).

Models in the category GCM are generally suitable for forecasts from less than an hour up to some centuries in the future. In these models many climatic processes are modelled in-depth. To run these models the most modern supercomputers are used. Nevertheless, the forecast period ranges only up to some centuries due to a restricted calculation capacity. To be able to forecast on a longer term EMICs have been developed. In these models a lot of the climatic processes are programmed in less detail or are omitted. The forecast horizon of these models ranges from some years up to 100,000 years.

For the very short term specialized weather forecast models are used. An example of this is the Ensemble Prediction System (EPS). The EPS belongs to the category of GCMs and calculates 51 scenarios based on different initial situations. The forecast period ranges from 48 hours up to 10 days in the future. For the forecast up to 48 hours the KNMI (Royal Dutch Meteorological Institute) uses the High Resolution Limited Area Model (HIRLAM). An important characteristic of this model is that only part of the earth is modelled. On the contrary, the resolution of the model is high.

I have done research to forecast the average day temperature of tomorrow with two statistical models. The two statistical models are a regression model and a time series model. I have compared the forecasts of these models to the forecast that the average day temperature of tomorrow is equal to the average day temperature of today. Both models do not perform significantly better than this benchmark forecast.

The conclusion is that climate models for both the short term and the long term are very complex. In the most accurate models the complete earth is modelled in three dimensions. Furthermore, several climatic processes are modelled in detail. Then, these models are simulated to predict the weather or climate in the future.

From the research I have done I can conclude that complex climate models are absolutely necessary to make accurate weather forecasts. Furthermore, one of the core functions of the KNMI is to warn for extreme weather conditions. It is not possible to predict these extreme conditions with simple statistical models, whereas the complex climate models are suited for this task.

# Inhoudsopgave

1	Inleiding .....	5
2	Het klimaatsysteem .....	7
3	Global Climate Models .....	10
3.1	Feedback van sneeuw- en ijsoppervlakken .....	11
3.2	Feedback van waterdamp en wolken .....	11
3.3	Feedback van micro-organismen in de oceanen .....	12
4	Earth Models of Intermediate Complexity .....	13
5	Weersvoorspellingen .....	16
5.1	Ensemble Prediction System .....	16
5.2	High Resolution Limited Area Model .....	21
6	Eigen onderzoek .....	23
6.1	Regressiemodel .....	23
6.2	Tijdreeksmodel .....	26
6.3	Resultaten .....	29
6.4	Conclusie eigen onderzoek .....	30
7	Conclusie .....	31
8	Literatuurlijst .....	32

# 1 Inleiding

De aarde: één van de meest complexe systemen die de mensheid ooit aanschouwd heeft en die we elke dag opnieuw aanschouwen. Aan de ene kant van het tijdsspectrum ligt het weer, dat van uur tot uur plotseling kan omslaan. Aan de andere kant, de lange termijn, is de aarde meer afhankelijk van astronomische krachten en zijn er zowel ijstijden als warmere periodes geweest. Het klimaat wordt zowel op de korte als lange termijn gekenmerkt door grilligheid.

Echter, het klimaat is in balans. Dat staat in ieder geval vast voor het verleden. Het klimaatsysteem is namelijk in staat om na veranderingen terug te keren naar haar oorspronkelijke staat. Neem als voorbeeld vulkaanuitbarstingen, waar as, zwaveldioxide en ook zware metalen de atmosfeer in worden geblazen. Het zelfreinigende proces van de atmosfeer werkt dan als volgt. In de atmosfeer ontstaan door een chemische reactie van ozon, waterdamp en ultraviolette straling OH radicalen (moleculen met een ongepaard elektron). Deze radicalen reageren onmiddellijk (binnen een seconde) met de schadelijke gassen, waardoor de concentratie schadelijke gassen langzaam wordt afgebroken.

In het heden is, terugkijkend naar het recente verleden, een opwarming van de aarde te zien sinds 1950. Het rapport van het IPCC (Intergovernmental Panel on Climate Change, opgericht door de Verenigde Naties en Wereld Meteorologische Organisatie) van 2 februari 2007 geeft aan dat het merendeel van de opwarming zeer waarschijnlijk veroorzaakt is door een ophoping van broeikasgassen. Met zeer waarschijnlijk wordt hier bedoeld dat de kans groter is dan 90 procent. De menselijke activiteiten die dit veroorzaken zijn voornamelijk industrialisering en ontbossing (Bruce e.a., 1996; Giambelluca en Henderson-Sellers, 1996; Watson e.a., 1996). Verder zorgt de groei van de wereldbevolking er voor dat deze activiteiten niet gemakkelijk tot stoppen gebracht kunnen worden.

De menselijke activiteiten zijn terug te zien in het opraken van fossiele brandstoffen. Verder zijn CO<sub>2</sub> niveaus in de atmosfeer ongekend hoog en nemen nog steeds toe. De ozonlaag is in de afgelopen eeuw dunner geworden. Een ander belangrijk punt is dat de chemische samenstelling van de atmosfeer verandert, waardoor de zelfreinigende capaciteit door oxidatie afneemt (Keeling e.a., 1976, 1995; WMO, 1994; Houghton e.a., 1996; MacKay e.a., 1997). De gevolgen van een dunnere ozonlaag zijn bijvoorbeeld dat meer ultraviolette straling het aardoppervlak bereikt. Verder neemt door de afnemende zelfreinigende capaciteit van de atmosfeer de luchtvervuiling toe.

Een hachelijk punt is dat de gevolgen van klimaatveranderingen pas laat zichtbaar zijn. Voorbeeld hiervan is het gat in de ozonlaag boven Antarctica, dat in 1984 ontdekt werd. Elk jaar rond september bereikt de ozonconcentratie zijn minimum. Boven een gebied zo groot als Afrika is dan de ozonconcentratie meer dan gehalveerd.

Dit gaat ook op voor toekomstige veranderingen van het klimaat. Op dat moment kan het al te laat zijn om de gevolgen te bestrijden.

Doordat industriële uitstoot wereldwijd voorlopig niet zal verminderen, zullen vragen over klimaatverandering alleen maar toenemen. Klimaatmodellen spelen een zeer belangrijke rol in het beantwoorden van deze vragen. Bovendien kunnen ze de noodzaak en het effect van wereldwijde milieuafspraken weergeven.

De probleemstelling van mijn scriptie is: hoe wordt met modellen onderzoek gedaan naar mogelijke klimaatveranderingen?

In deze scriptie zal ik bekijken hoe klimaatmodellen er uitzien. Verder zal ik kijken naar de aanpak voor korte termijn voorspellingen (weersvoorspellingen) en lange termijn voorspellingen (klimaatveranderingen tot 100.000 jaar).

Edward Lorenz kenmerkte de complexiteit van het klimaat met zijn uitspraak dat een vlinder in Brazilië maanden later een tornado in Texas tot gevolg zou kunnen hebben. Om te onderzoeken of complexe klimaatmodellen echt een toegevoegde waarde hebben, zal ik door praktijkonderzoek nagaan hoe deze modellen zich verhouden tot veel simpelere statistische voorspellingen.

## 2 Het klimaatsysteem

Om duidelijk te krijgen hoe het klimaatsysteem ongeveer werkt, volgt hier eerst de definitie van de United Nations Framework Convention on Climate Change (FCCC) uit 1992:

“‘Climate system’ means the totality of the atmosphere, hydrosphere, biosphere and geosphere and their interactions.”

Met de atmosfeer wordt de dampkring om de aarde aangeduid.

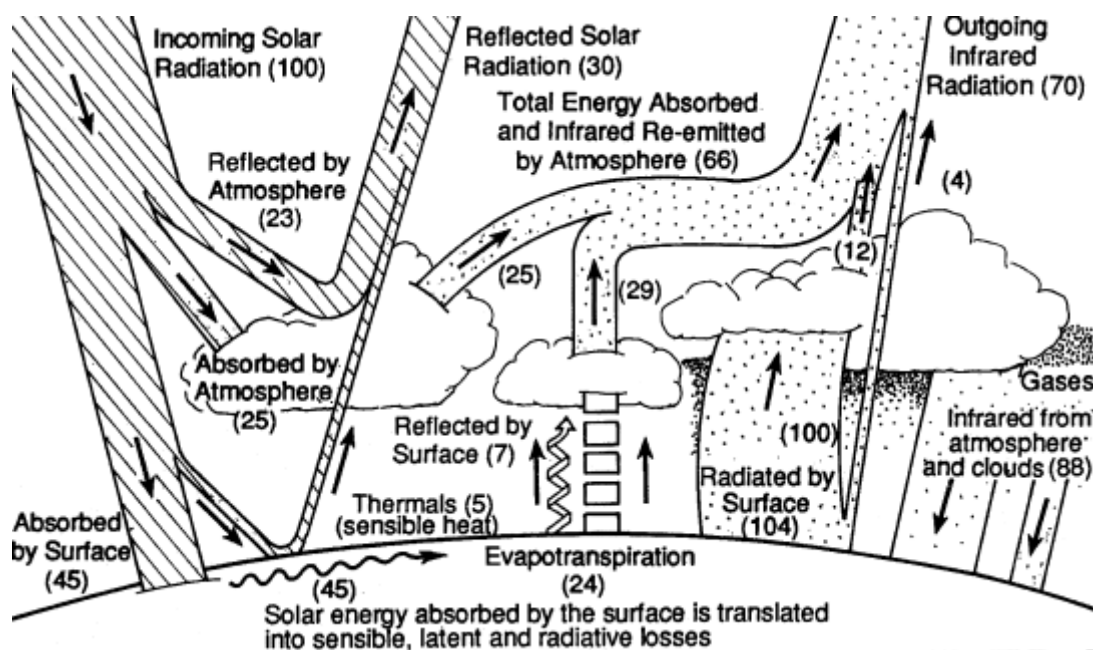
De hydrosfeer is het geheel van water op, onder en boven het oppervlak van een planeet. Dit omvat dus alle oceanen, zeeën, rivieren, ijs, sneeuw en grondwater.

De biosfeer bevat alle aardse organismen, dus zowel ter land, ter zee als in de lucht.

De geosfeer bevat de vorm en verschuivingen van het aardoppervlak.

Het belangrijkste in deze definitie zijn de interacties tussen de atmosfeer, hydrosfeer, biosfeer en geosfeer.

Ter verduidelijking zal ik nu in meer detail op het klimaatsysteem in gaan. Het startpunt van leven op aarde is de zon. De zon speelt een zeer grote rol in het klimaatsysteem. Aan de hand van figuur 1 zal ik het energiebeheer van de aarde uitleggen.



Figuur 1. Schematische weergave van het energiebeheer van de aarde (getallen tussen haakjes zijn procenten)

Het energiebeheer van de aarde begint bij binnenkomende straling van de zon (100 procent).

Bij een klimaatsysteem dat in balans is, zal de totaal uitgestraalde energie overeenkomen met de binnenkomende energie. Zo niet, dan zal de aarde opwarmen of afkoelen.

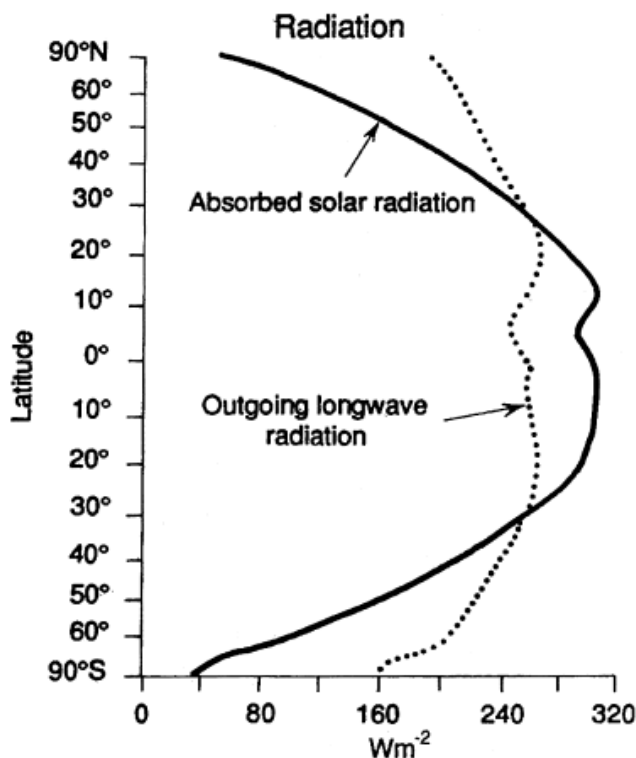
Van de binnenkomende straling van de zon, wordt ongeveer 23 procent gereflecteerd door de atmosfeer en 7 procent door het aardoppervlak (bijvoorbeeld door grote ijs- en sneeuwoppervlakken). Verder wordt 25 procent geabsorbeerd door de atmosfeer en de overige 45 procent van de energie wordt opgenomen door het aardoppervlak. Van de binnenkomende straling wordt dus 70 procent door onze planeet geabsorbeerd.

Deze warmte verspreidt zich vervolgens over de aarde. Een deel van de energie komt door verdamping van water van het aardoppervlak terug in de atmosfeer. Het aardoppervlak geeft ook direct infrarode straling af. Op deze manier wordt ook energie terug de atmosfeer ingebracht. Een deel daarvan wordt geabsorbeerd door de atmosfeer, een groot deel wordt teruggekaatst naar de aarde door wolken en broeikasgassen en een klein deel verdwijnt direct terug het heelal in (4 procent van de oorspronkelijke binnengekomen energie van de zon). De atmosfeer geeft vervolgens de resterende energie af, zodat de energiebalans (als het klimaatsysteem in evenwicht is) na verloop van tijd neutraal is.

De lucht in de atmosfeer bestaat voor het grootste deel uit stikstof ( $N_2$ ) en zuurstof ( $O_2$ ). Deze chemische verbindingen houden echter geen infrarode straling vast. Hoewel broeikasgassen maar een klein gedeelte van de atmosfeer vormen, spelen ze wel een zeer grote rol in het energiebeheersysteem van de aarde.

Het energiebeheer van de aarde is een dynamisch proces. De hoeveelheid energie die de zon uitstraalt, varieert over meerdere jaren gemeten. Dit is een natuurlijke oorzaak waar het klimaatsysteem zich aan moet aanpassen. Aan de andere kant is de hoeveelheid broeikasgassen in de atmosfeer voor een deel de oorzaak van menselijke activiteiten. Ook hieraan zal het klimaatsysteem zich moeten aanpassen om te zorgen dat er evenveel energie de aarde binnenkomt als verlaat.

Een tweede punt dat belangrijk in het klimaatsysteem is, is het verschil in binnenkomende zonneenergie op verschillende geografische breedtes. Op de noordpool (90 graden noorderbreedte) of zuidpool (90 graden zuiderbreedte) komt namelijk veel minder energie door zonnestraling binnen dan op de evenaar. Door de verspreiding van energie in de aarde en atmosfeer, wordt richting de polen meer energie afgegeven dan geabsorbeerd. Rond de evenaar wordt juist meer energie geabsorbeerd dan afgegeven (zie figuur 2).



Figuur 2. Weergave van het energiebudget van de aarde op verschillende geografische breedtes



Een andere oorzaak van energie-onbalans is de rotatie van de aarde om haar as. Door deze rotatie ligt telkens een ander gebied in de maximale warmte van de zon. Behalve een onbalans op de geografische breedte van de aarde, is er dus ook in de geografische lengterichting een energie-onbalans.

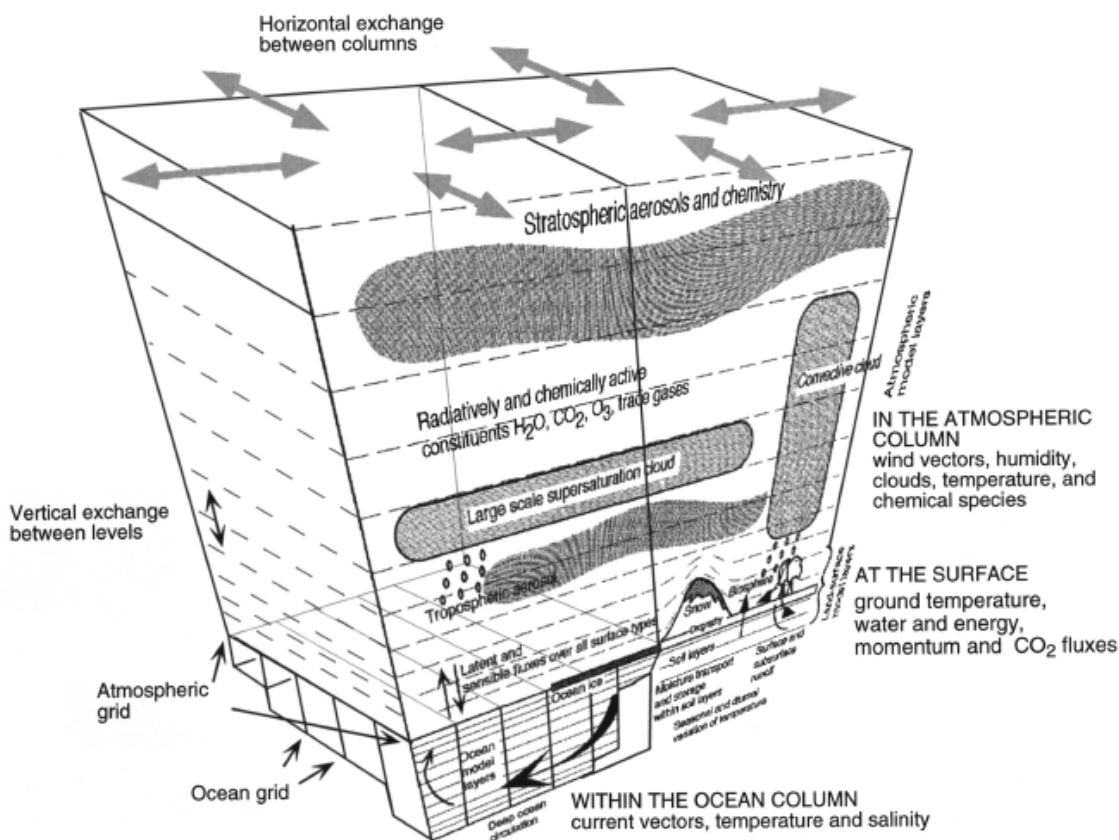
Doordat het klimaatsysteem deze onbalans probeert te herstellen, ontstaat een diversiteit aan klimatologische processen. Voorbeelden hiervan zijn stromingen in de atmosfeer of stromingen in de diepere lagen van de oceanen.

Globale energiebalans, energie-onbalans op verschillende geografische breedtes en rotatie van de aarde zijn de drie belangrijkste kenmerken van het klimaatsysteem. Helaas is met alleen deze drie processen het klimaatsysteem nog niet (nauwkeurig) te modelleren. Dit komt doordat dit een te grove vereenvoudiging van de werkelijkheid is.

### 3 Global Climate Models

Global Climate Models (GCMs) horen tot de meest gedetailleerde klimaatmodellen. In deze modellen komt kennis samen van diverse vakgebieden, waaronder natuurkunde, scheikunde, biologie, economie en sociale wetenschappen. Naast de belangrijkste kenmerken van het klimaatproces, zoals in het vorige hoofdstuk is besproken, worden nog veel meer processen gemodelleerd. Het is echter niet mogelijk om alle kennis die we op dit moment over het klimaat hebben, in een Global Climate Model weer te geven. Zo een model zou zelfs op de snelste supercomputer niet meer te draaien zijn.

Om in de categorie GCM te vallen, moet een klimaatmodel in ieder geval volledig driedimensionaal zijn en een model van de atmosfeer en van de oceanen bevatten, die aan elkaar gekoppeld zijn. Om het klimaat met behulp van numerieke methoden te kunnen simuleren, wordt de aarde onderverdeeld in kleinere driedimensionale gebieden. Voor de oceanen zijn deze gebieden over het algemeen kleiner dan voor de atmosfeer (zie figuur 3). Dit komt doordat stromingen in de oceaan in een veel kleiner gebied kunnen bestaan dan dat in de atmosfeer het geval is. Draaikolken in de oceaan (die van belang zijn voor klimaatmodellering) hebben namelijk een doorsnede van 10 tot 50 kilometer. In de atmosfeer hebben deze gebieden een doorsnede van ongeveer 1000 kilometer.



Figuur 3. Schematische weergave van de aarde in een GCM

Geologische kenmerken van de aarde worden in het model verwerkt, zoals de vorm van de oceanen, de verschillende dieptes van de oceanen op elke plaats, de hoogteverschillen op het aardoppervlak. In deze schematische weergave van de aarde worden allerlei verschillende klimatologische processen gemodelleerd. Elk proces heeft een bepaald deelgebied waar het

zich kan afspelen. In een GCM wordt vervolgens de toestand in elk gebied gesimuleerd per tijdstap, waarbij interactie tussen de verschillende gebieden kan plaatsvinden. Via differentiaalvergelijkingen worden veel globale processen in de atmosfeer en oceanen beschreven, zoals de beweging van energie, de beweging van waterdamp in de atmosfeer en de beweging van zout in de oceanen. Deze differentiaalvergelijkingen worden gebruikt om de toestand van deze processen voor elk gebied voor de volgende tijdstap te berekenen. Bij de simulatie wordt in ieder geval gezorgd dat natuurkundige wetten als globale energiebalans (zoals eerder beschreven) en behoud van massa niet geschonden worden.

Naast de geleidelijke veranderingen zoals stroming van warmte, kan ook door een interne of externe kracht het klimaat beïnvloed worden. Vulkanische uitbarstingen en verhoging van CO<sub>2</sub> niveaus door menselijke activiteiten zijn voorbeelden van interne veranderingen. Variatie in stralingsenergie van de zon is een externe kracht. In een GCM zullen dus ook plotselinge veranderingen meegenomen moeten worden.

Verder vormt feedback een belangrijk kenmerk van GCMs. Er vindt feedback plaats als de output van een proces andere processen beïnvloed, waardoor de uiteindelijke output anders is. Er kan sprake zijn van een positieve of een negatieve feedback. Bij een positieve feedback wordt het oorspronkelijke effect versterkt. Bij een negatieve feedback wordt het oorspronkelijke effect afgezwakt. Met het modelleren van de feedback worden heel duidelijk interacties tussen de atmosfeer, hydrosfeer, biosfeer en geosfeer meegenomen. Om de complexiteit van een GCM duidelijk te maken, zal ik een aantal van deze feedback processen beschrijven.

### 3.1 Feedback van sneeuw- en ijsoppervlakken

Stel dat de CO<sub>2</sub> concentratie in de atmosfeer toeneemt of dat de stralingsenergie van de zon toeneemt. Deze veranderingen hebben beiden een stijging van de globale oppervlakte temperatuur tot gevolg. Door hogere oppervlakte temperaturen zal echter een deel van het ijs en de sneeuw dat op de aarde aanwezig is smelten. Bijvoorbeeld een deel van de ijsvlakten op de noord- of zuidpool. Hierdoor wordt de totale oppervlakte van de cryosfeer kleiner. De cryosfeer wordt gevormd door de gebieden waar water voorkomt in de vorm van sneeuw, permafrost, pakijs of gletsjers en behoort tot de hydrosfeer. Deze ijsoppervlakken spelen een belangrijke rol in het energiebeheersysteem van de aarde. Het weerkaatst namelijk zonnestraling, zodat in totaal minder warmte door de aarde wordt opgenomen. Als de totale grootte van ijsoppervlakken afneemt door smelting, zal dus meer zonneenergie door de aarde geabsorbeerd worden. Hierdoor zal de globale oppervlakte temperatuur verder stijgen. De feedback van sneeuw- en ijsoppervlakken is in dit geval dus een positieve feedback. Het oorspronkelijke effect van een stijgende globale oppervlakte temperatuur wordt namelijk versterkt.

### 3.2 Feedback van waterdamp en wolken

Stel dat de CO<sub>2</sub> concentratie in de atmosfeer toeneemt of dat de stralingsenergie van de zon toeneemt. Deze veranderingen hebben beiden een stijging van de globale oppervlakte temperatuur tot gevolg. Hierdoor zal meer water verdampen en de hoeveelheid waterdamp in de atmosfeer neemt toe. Waterdamp heeft de eigenschap om infrarode straling die de aarde uitstraalt vast te houden. Deze straling zou anders ontsnapt zijn naar het heelal. Het effect is dus een verdere opwarming van de aarde, oftewel een positieve feedback.

Als door de toegenomen waterdamp het aantal wolken toeneemt, ontstaat weer een ander effect. Wolken hebben twee verschillende effecten in het energiebeheersysteem van de aarde. Het eerste effect is dat inkomende stralingsenergie van de zon wordt weerkaatst. Dit is dus een koelend effect. Het tweede effect is dat infrarode straling van de aarde wordt vastgehouden. Dit heeft een opwarmend effect. Welk van deze twee effecten overheerst, is afhankelijk van de hoogte van de wolk. Bij wolken hoog in de atmosfeer overheerst het opwarmende effect. Wolken laag in de atmosfeer hebben netto juist een koelend effect. Of de feedback van wolken in dit scenario positief of negatief is, hangt er dus vanaf op welke hoogte de wolken (voornamelijk) zullen ontstaan. Het ontstaan en modelleren van (de hoogte en dikte van) wolken is een zeer moeilijk probleem waar volop onderzoek naar wordt gedaan.

### 3.3 Feedback van micro-organismen in de oceanen

Stel dat de CO<sub>2</sub> concentratie in de atmosfeer toeneemt. Deze verandering heeft een stijging van de globale oppervlakte temperatuur tot gevolg. Van alle extra CO<sub>2</sub> in de atmosfeer wordt een groot deel opgelost in de oceanen en verdwijnt op deze manier uit de atmosfeer. Door de grotere hoeveelheid CO<sub>2</sub> die in de oceanen is opgelost, worden de oceanen zuurder. Hierdoor wordt het voor (micro-)organismen met kalkskeletten veel moeilijker om te overleven. Dit kan invloed hebben op andere processen binnen de oceaan. Aan de andere kant zijn sommige micro-organismen in staat om CO<sub>2</sub> om te zetten in zuurstof. Door hogere CO<sub>2</sub> niveaus zou het aantal van deze micro-organismen kunnen toenemen, wat een negatieve feedback zou betekenen.

Over de precieze gevolgen van dit feedback mechanisme is nog niet genoeg bekend en ik neem daarom aan dat dit in de meeste Global Climate Models niet is verwerkt.

Global Climate Models worden gebruikt om het klimaat van minder dan een uur tot enkele eeuwen in de toekomst te voorspellen. Met de rekencapaciteit van de huidige supercomputers is het niet mogelijk om met GCMs voorspellingen te doen verder in de toekomst zonder het aantal details te verminderen.

## 4 Earth Models of Intermediate Complexity

Om toch onderzoek te kunnen doen naar de ontwikkeling van het klimaat op de langere termijn, is een nieuw type klimaatmodel ontwikkeld. Deze categorie van klimaatmodellen heet Earth Models of Intermediate Complexity (EMIC).

In deze modellen worden sommige kenmerken van de GCMs vereenvoudigd. Een voorbeeld daarvan is het constant veronderstellen van de oppervlakteweerkaatsing van zonneenergie op verschillende geografische breedtes. Deze feedback van ijs- en sneeuwoppervlakken (zoals hierboven beschreven) die gemodelleerd is in GCMs, wordt daar als erg belangrijk gezien. Vanwege de beperkt beschikbare reken capaciteit wordt dit effect niet gemodelleerd in veel Earth Models of Intermediate Complexity.

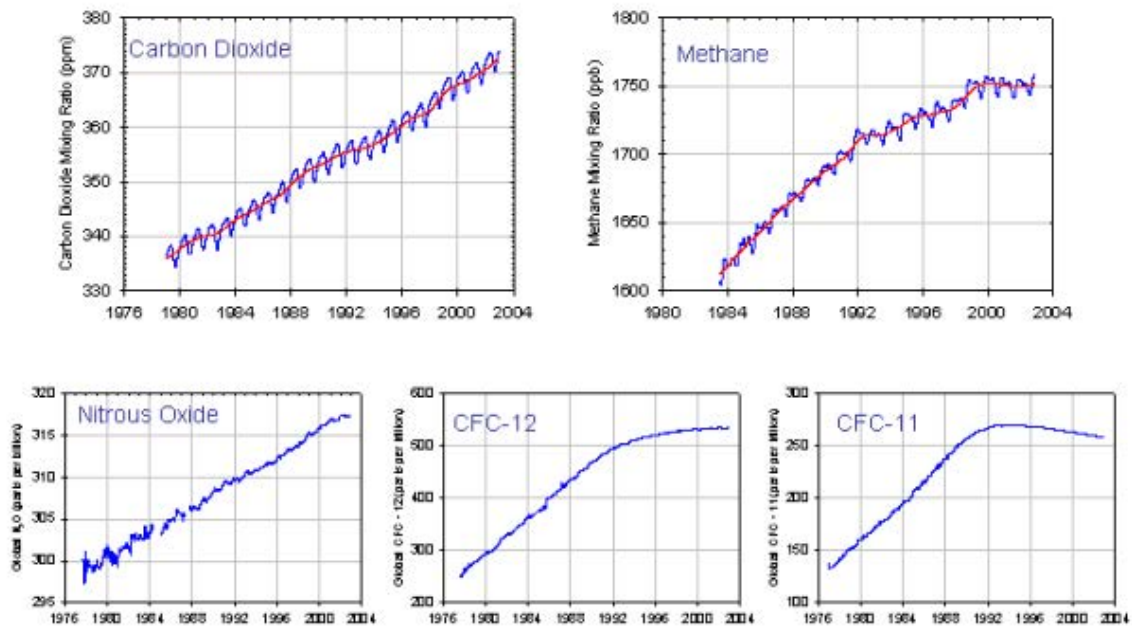
Sommige processen spelen op de korte termijn een belangrijke rol in het verloop van het klimaat, maar zijn voor de langere termijn minder belangrijk. Een voorbeeld hiervan is de modellering van wolken. Om nog nauwkeuriger voorspellingen voor de korte termijn te kunnen doen, is dit een belangrijk verbeterpunt voor Global Climate Models. Voor de langere termijn is de precieze modellering minder van belang. In EMICs wordt wolkenvorming soms simpel afhankelijk gesteld van de oppervlaktetemperatuur.

Sommige EMICs zijn slechts tweedimensionaal. Waar Global Climate Models de hele aarde modelleren (driedimensionaal), wordt bij deze EMIC bijvoorbeeld de geografische lengte als dimensie weggelaten. Wat overblijft is een model met de geografische breedte (noodzakelijk om de energie-onbalans op verschillende geografische breedtes te kunnen modelleren) en de hoogte.

De tijdsschaal van GCMs en EMICs verschilt dus per definitie sterk. Waar Global Climate Models voorspellingen doen over het klimaat voor minder dan een uur in de toekomst tot ongeveer 200 jaar, begint de voorspellende capaciteit van een EMIC pas bij enkele jaren. Met een EMIC kunnen voorspellingen voor maximaal 100.000 jaar in de toekomst worden gedaan. Deze tijdsschalen zijn indicatief voor de twee modelcategorieën. Er zijn ook voorbeelden bekend waar Global Climate Models ingezet worden voor langere termijn voorspellingen en EMICs voor kortere termijnen dan aangegeven. Voor de meerderheid van de modellen uit deze categorieën klopt deze tijdsschaal.

Earth Models of Intermediate Complexity zijn in het verleden bijvoorbeeld gebruikt om de gevolgen van de maatregelen uit het Kyoto protocol te voorspellen. Met dit verdrag zijn industrielanden overeengekomen om de uitstoot van broeikasgassen met gemiddeld 5% te verminderen ten opzichte van het niveau in 1990. De periode waar naar gekeken wordt is 2008–2012. Het gaat hier om de gassen koolstofdioxide ( $\text{CO}_2$ ), methaan ( $\text{CH}_4$ ), distikstofoxide ( $\text{N}_2\text{O}$ ) en bepaalde fluorverbindingen zoals CFK's, HFK's, PFK's en  $\text{SF}_6$ . Deze gassen dragen in verschillende mate bij aan het broeikaseffect. Om dit te kunnen vergelijken is het  $\text{CO}_2$ -equivalent ingesteld. Methaan heeft bijvoorbeeld een  $\text{CO}_2$ -equivalent van 21. Dat houdt in dat 1 kilo methaan over een periode van 100 jaar 21 maal zoveel aan het broeikaseffect bijdraagt als 1 kilo koolstofdioxide.  $\text{SF}_6$  heeft het hoogste  $\text{CO}_2$ -equivalent, namelijk 23.900.

In figuur 4 is te zien hoe de concentratie van broeikasgassen in de atmosfeer zich de afgelopen 20–25 jaar heeft ontwikkeld.



Figuur 4. Trends in hoeveelheid broeikasgassen<sup>1</sup>

De broeikasgassen in figuur 4 dragen bij tot 97% van het totale broeikaseffect. Voor methaan, CFC-12 en CFC-11 (beide CFK's, oftewel chloorfluorkoolstofverbindingen) is te zien dat de stijgende trend tot stilstand is gebracht. Voor CFC-11 is zelfs een lichte daling te zien. Voor koolstofdioxide en distikstofoxide is de trend echter nog steeds stijgend. De reductiepercentages die in het Kyoto protocol zijn afgesproken, verschillen van land tot land. Luxemburg moet zijn uitstoot bijvoorbeeld met 28% verminderen terwijl Portugal zijn uitstoot met 27% mag laten stijgen. Het reductiepercentage voor Nederland is 6%.

Het Kyoto protocol is in 1997 opgesteld en het verdrag zou in werking treden als de parlementen van 55 landen, die samen 55 procent van de uitstoot van broeikasgassen veroorzaken, het hadden bekrachtigd. Op 23 oktober 2004 werd het verdrag ook in Rusland goedgekeurd. Hierdoor was de eis van 55 landen die samen 55 procent van de uitstoot van broeikasgassen veroorzaken, gehaald. Op 16 februari 2005 (90 dagen na de ondertekening door Rusland) is het Kyoto protocol officieel in werking getreden

Om afspraken te maken voor het Kyoto protocol was het natuurlijk erg belangrijk om te kunnen inschatten wat het effect zou zijn van een maatregel. Hierbij was men vooral geïnteresseerd in de klimaatveranderingen op de lange termijn, oftewel langer dan enkele eeuwen. Dus de tijdschaal was langer dan wat met een GCM kon worden gemodelleerd. De ratificatie van het verdrag door lidstaten was een proces dat enkele jaren duurde. In die jaren werd met klimaatmodellen geprobeerd het effect van de maatregelen te bepalen. T.M.L. Wigley gebruikte in 1998 een EMIC met een sterk vereenvoudigde representatie van het klimaat. Hierdoor was er rekencapaciteit over om verschillende scenario's te simuleren. Bij zo een scenario werden aannames gemaakt over verschillende factoren die het klimaat beïnvloeden. Bijvoorbeeld de concentratie van CO<sub>2</sub> in de atmosfeer. Allereerst werd deze concentratie geschat, waarbij aangenomen werd dat de maatregelen uit het Kyoto protocol niet waren genomen. Daarnaast werden er verschillende scenario's voor deze concentratie opgesteld naar aanleiding van de maatregelen uit het Kyoto protocol. Zo werd dit gedaan voor

<sup>1</sup> Afbeelding afkomstig van [http://nl.wikipedia.org/wiki/Afbeelding:Trends\\_gassen\\_ozonlaag.PNG](http://nl.wikipedia.org/wiki/Afbeelding:Trends_gassen_ozonlaag.PNG).

verschillende factoren die het klimaat beïnvloeden en die een afhankelijkheid hadden met maatregelen uit het Kyoto protocol. Hierdoor ontstonden verschillende scenario's voor het hele klimaatsysteem. Deze werden met het EMIC gesimuleerd. Zo werd beter duidelijk wat het effect was van de maatregelen uit het Kyoto protocol. De conclusie van dit onderzoek was dat als alle landen zich aan het Kyoto protocol zouden houden de veranderingen in het klimaat beperkt zouden zijn.

Op de dag van inwerkingtreding van het Kyoto protocol in 2005 waren wetenschappers dan ook al van mening dat de doelstelling onvoldoende was. Sindsdien wordt er gesproken over het Post-Kyoto protocol.

## 5 Weersvoorspellingen

Waar EMICs bedoeld zijn om het klimaat op de langere termijn te voorspellen, bestaan er ook veel verschillende modellen waarmee het klimaat op de korte termijn voorspeld kan worden. In dit geval hebben we het over het weer.

Het KNMI is de belangrijkste instelling in Nederland die zich bezighoudt met het voorspellen van het weer. Het doel hiervan is het ondersteunen van de veiligheid, economie en een duurzaam milieu. Om de veiligheid te ondersteunen geeft het KNMI bijvoorbeeld waarschuwingen uit bij gevaarlijke weersituaties.

Voor langetermijntoontwikkelingen verricht het KNMI onderzoek naar de veranderingen in het klimaat. Het beschikbaar stellen van bij het KNMI aanwezige kennis, data en informatie is een kernactiviteit.

Voor weersvoorspellingen hanteert het KNMI de begrippen korte, middellange en lange termijn. Korte termijn betekent hier 0 tot en met 48 uur in de toekomst. Middellange termijn geeft de periode aan van 2 tot 5 dagen vooruit en lange termijn de periode van 6 tot 10 dagen in de toekomst. Voor de middellange en de lange termijn gebruikt het KNMI voorspellingen van het European Centre for Medium-Range Weather Forecasts (ECMWF). Voor de korte termijn genereert het KNMI voorspellingen met het HIRLAM model.

### 5.1 Ensemble Prediction System

De modellen die het ECMWF gebruikt om het weer te voorspellen, vallen in de categorie Global Climate Models. Zoals eerder besproken worden GCMs gebruikt om het klimaat van minder dan een uur tot enkele eeuwen in de toekomst te voorspellen. De modellen van het ECMWF vallen uiteraard in het korte termijn gedeelte van dit spectrum. De voorspellingen die het ECMWF afgeeft die het meest ver in de toekomst liggen, zijn voorspellingen tot 21 dagen.

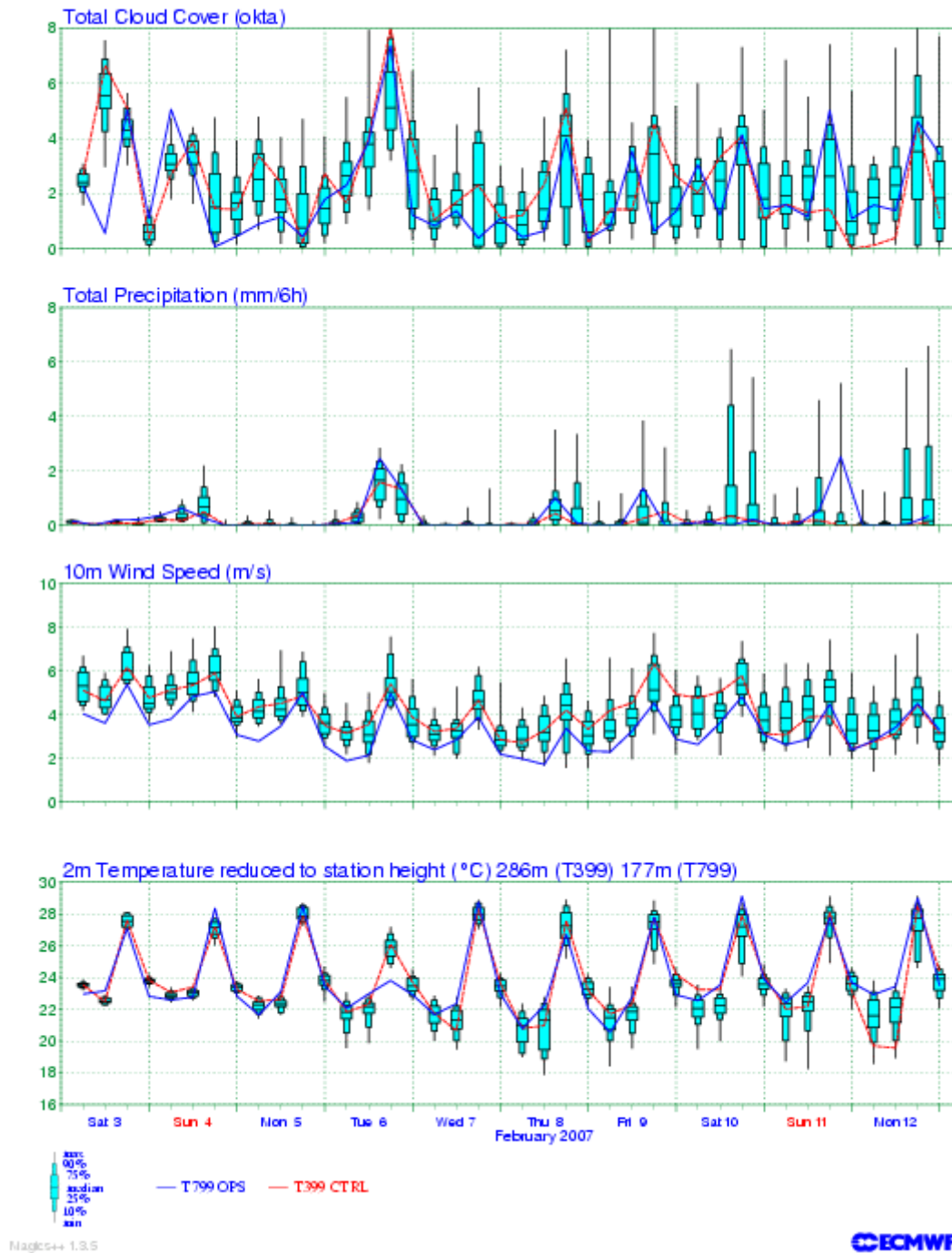
De belangrijkste voorspellingen van het ECMWF worden gemaakt met de volgende twee voorspellingsmodellen. Allereerst de voorspelling tot tien dagen in de toekomst op een resolutie van ongeveer 25 kilometer. Deze resolutie geeft de grootte van de vlakken aan waarin de aarde is opgedeeld. De gehele aardbol is in dit model dus opgedeeld in vlakken van ongeveer 25 bij 25 kilometer. Dit is de hoogste resolutie die het ECMWF in haar modellen gebruikt. Ondanks deze hoge resolutie en toenemende nauwkeurigheid van de voorspellingen, blijft er altijd een bepaalde onzekerheid in de voorspellingen. Deze onzekerheid is niet geheel op te lossen door de resolutie nog verder te verhogen.

Zoals in de inleiding al is aangegeven met het vlindereffect, is het weer ook wel te beschrijven met chaostheorie. Het vlindereffect van Edward Lorenz geeft aan dat een vlinder in Brazilië maanden later een tornado in Texas tot gevolg zou kunnen hebben. Dit komt terug bij het modelleren van het klimaat. De beginsituatie, oftewel de stand van het klimaat op dit moment, is namelijk niet met zekerheid nauwkeurig vast te leggen. Hoe langer de voorspellingstermijn wordt, hoe meer deze fouten in de begintoestand zorgen voor afwijkingen ten opzichte van de werkelijkheid.



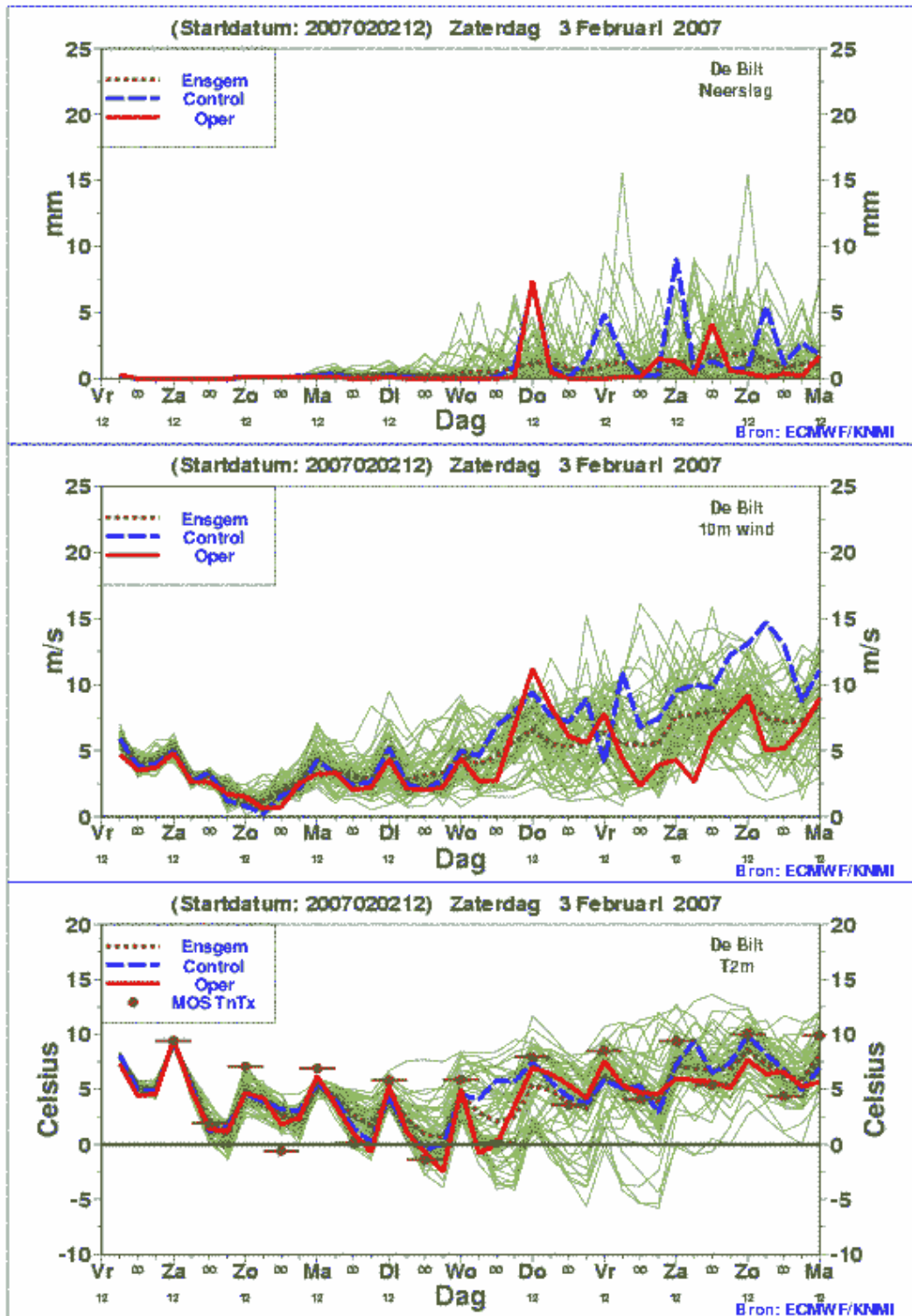
Daarom bepaalt het ECMWF de voorspelling tot tien dagen in de toekomst ook nog op een andere manier. Het Ensemble Prediction System (EPS) berekent 51 weersvoorspellingen op een resolutie van 50 kilometer. Al deze 51 voorspellingen hebben een andere begintoestand als input van het klimaatmodel. De begintoestand van het eerste voorspellingsmodel is hierbij een beetje verstoord om de effecten van onzekerheden in die begintoestand te simuleren. Zie figuur 5 voor de output van het Ensemble Prediction System. Hier zijn de resultaten weergegeven door voor elk tijdstip de 10%, 25%, 50%, 75%, 90% kwantielen te berekenen en de minimale en maximale waarde. De rode lijn (T399) is de verwachting op basis van het EPS zonder verstoring in de begintoestand. De blauwe lijn (T799) is de voorspelling van het eerste voorspellingsmodel met een hoge resolutie van 25 kilometer. Doordat een globaal klimaatmodel gebruikt wordt, is het mogelijk om voor elke plaats in de wereld een 10-daagse voorspelling zoals in figuur 5 af te geven.

EPS Meteogram  
 Caguas (98m) 18.2°N 66.15°W  
 Deterministic Forecast and EPS Distribution Saturday 3 February 2007 00 UTC



Figuur 5: EPS output van 3 februari voor de plaats Caguas (Puerto Rico)

In plaats van de weergave van kwantielen is het ook mogelijk om alle 51 voorspellingen weer te geven. Door de 51 voorspellingen over elkaar te leggen, wordt een band weergegeven. Deze band is gelijk een maatstaf voor de nauwkeurigheid van de voorspellingen. Deze uiteindelijke voorspelling wordt door het KNMI gebruikt als belangrijkste voorspelling voor de middellange en lange termijn.



Figuur 6. 10-daagse verwachting KNMI van 3 februari op basis van EPS

In figuur 6 zijn de 51 voorspellingen van het Ensemble Prediction System als groene lijnen te zien. Het is duidelijk te zien hoe kleine veranderingen in de begintoestand op langere termijn resulteren in een geheel andere voorspelling.

Een ander verschil tussen de EPS voorspellingen en één gemiddelde voorspelling heeft betrekking op de economische waarde van de voorspelling. De economische waarde van de EPS voorspellingen is veel hoger dan van één gemiddelde voorspelling. Met de EPS

voorspellingen is bijvoorbeeld een risicoanalyse mogelijk voor bedrijven met weersafhankelijke activiteiten. Waar de gemiddelde / verwachte temperatuur of neerslag slechts één voorspelling is, kan met de 51 mogelijke scenario's worden geanalyseerd wat er allemaal kan gebeuren in de toekomst.

Dit kan bijvoorbeeld toegepast worden in de Formule 1, waar de pitstopstrategie en bandenkeuze beter bepaald kunnen worden met deze 51 scenario's (hiervoor is een kortere termijn voorspelling nodig dan deze 10-daagse voorspelling). In het ideale geval is het team voorbereid als het weer plotseling omslaat en het begint te regenen. Met de voorspellingen van het EPS heeft het team dan van tevoren besloten hoe te handelen als deze situatie zich voordoet.

Een ander voorbeeld is het waarschuwen voor gevaarlijke weersomstandigheden door het KNMI. Sommige van de 51 scenario's kunnen bijvoorbeeld aangeven dat er een storm op komst is. Het KNMI kan dan aangeven dat dit een risico is voor die betreffende dag. Dit is erg nuttig voor bedrijven in de watersportsector.

In totaal zijn 18 landen lid van het ECMWF. Het gaat hier om België, Denemarken, Duitsland, Spanje, Frankrijk, Griekenland, Ierland, Italië, Luxemburg, Nederland, Noorwegen, Oostenrijk, Portugal, Zwitserland, Finland, Zweden, Turkije en Groot-Brittannië. Verder heeft het ECMWF een samenwerkingsverband met nog eens 10 landen, namelijk Tsjechie, Estland, IJsland, Kroatië, Litouwen, Hongarije, Marokko, Roemenië, Servië en Slovenië.

Door dit grote samenwerkingsverband is het mogelijk om flinke uitgaven aan supercomputers te doen. Deze zijn essentieel voor het berekenen van een nauwkeurige weersvoorspelling. Hoe hoger de rekenkundige kracht van de computer, des te meer details er in het klimaatmodel meegenomen kunnen worden. Het ECMWF heeft een High Performance Computing Facility, dat bestaat uit twee identieke, maar onafhankelijke, IBM Cluster 1600 supercomputer systemen. Elk van de systemen bestaat uit ongeveer 2250 IBM Power5+ processors. De kloksnelheid per processor is 1.9 GHz.



Figuur 7. Eén van de twee IBM Power5+ Cluster 1600 systemen van het ECMWF

Hoewel dit erg veel lijkt, is deze rekenkundige capaciteit absoluut vereist. Het opstellen van de verwachting van het EPS met deze twee supercomputer systemen duurt namelijk ongeveer een dag.

Hierdoor is deze weersvoorspelling per definitie niet geschikt voor een periode van 24 uur in de toekomst. Als de resultaten van deze berekeningen bekend zijn, is de periode van 24 uur namelijk al om.

## 5.2 High Resolution Limited Area Model

Voor de korte termijn gebruikt het KNMI daarom een ander model. Dit model berekent de weersvoorspelling tot 48 uur in de toekomst. Het model heet HIRLAM, wat staat voor High Resolution Limited Area Model. Het HIRLAM systeem is ontwikkeld door een samenwerkingsverband van Denemarken, Frankrijk, Finland, Ierland, IJsland, Noorwegen, Spanje, Zweden en Nederland (KNMI).

Voor de weersverwachting van vijf of meer dagen vooruit moeten gegevens uit de hele wereld in de berekeningen worden verwerkt. Ontwikkelingen op duizenden kilometers afstand kunnen enkele dagen later van invloed zijn op het weer in Europa. Voor de weersverwachtingen tot hooguit 36 uur vooruit is het bijvoorbeeld voldoende gegevens uit een gebied tot een afstand van 1300 kilometer in de berekeningen mee te nemen. Dus, om het weer voor de korte termijn te voorspellen, heeft het KNMI geen globaal klimaatmodel nodig. HIRLAM valt dus niet in de categorie van Global Climate Models.

De naam HIRLAM (High Resolution Limited Area Model) geeft aan dat een gelimiteerde oppervlakte wordt gemodelleerd, maar wel met een hoge resolutie. De huidige resolutie van het HIRLAM model is 22 bij 22 kilometer. Dit is een hogere resolutie dan de modellen van het ECMWF, terwijl toch minder rekencapaciteit vereist is. Doordat niet het klimaat over de hele wereld berekend wordt, is de inzameltijd van de gegevens ook een stuk korter. Om de berekeningen van HIRLAM uit te kunnen voeren, heeft het KNMI begin van deze eeuw de beschikking gekregen over een nieuwe supercomputer. Het gaat hier om de Sun Fire 15 K. Deze supercomputer bestaat uit 106 UltraSPARC III Cu processors, die allemaal draaien op 1.2 GHz.

Deze supercomputer heeft duidelijk lagere specificaties dan de supercomputer systemen van het ECMWF. Niettemin behoort de supercomputer van het KNMI tot de snelste van Nederland. Recentelijk heeft het KNMI de beschikking gekregen over een nieuwe supercomputer, namelijk een Altix 3700 van Silicon Graphics. Deze supercomputer heeft 240 processoren die op 1.6 GHz draaien. Op dit moment hoort deze supercomputer tot de top 10 van snelste systemen in Nederland.

Het HIRLAM model is in principe ook met een moderne desktop pc te draaien. De runtime van het model is dan echter zo lang dat de voorspellingsperiode al voorbij is als de resultaten uiteindelijk beschikbaar zijn. Hier is dus duidelijk te zien waarom een supercomputer noodzakelijk is.

Naast het HIRLAM heeft het KNMI ook de beschikking over een ander model, namelijk het XHRLAM. Dit model is in principe hetzelfde als het HIRLAM, maar heeft een hogere resolutie. De resolutie van XHRLAM is 11 kilometer. Om dit model door te kunnen rekenen, wordt een kleiner gebied gemodelleerd dan bij HIRLAM. Hierdoor is de voorspellingshorizon van XHRLAM maximaal 24 uur in plaats van 48. Het voordeel van de hogere resolutie is dat extremen in de wind of neerslag beter behouden blijven. Dit is van groot belang voor de taak van het KNMI om weerswaarschuwingen af te geven. In figuur 8 zijn de karakteristieken van HIRLAM en XHRLAM weergegeven.

	HIRLAM	XHRLAM
resolutie	0.2°	0.1°
aantal roosterpunten	400*330	300*300
verwachtingstermijn	48 uur	24 uur
aantal verw. per dag	4	8
aflevertijd	+3 uur	+2 uur

Figuur 8. Configuratie van het HIRLAM en XHRLAM model

Een resolutie van 0.2 graden komt overeen met oppervlakken van 22 bij 22 kilometer en 0.1 graad met 11 bij 11 kilometer. De aflevertijd is het aantal uren na het waarneemtijdstip dat de voorspelling beschikbaar is. Bij het XHRLAM wordt ongeveer de helft van de aflevertijd besteed aan het verzamelen van de waarnemingen en de andere helft is runtime van het model.

## 6 Eigen onderzoek

Ik zal nu een eigen onderzoek presenteren dat gericht is op het klimaat op de korte termijn (het weer). Vanwege het ontbreken van een supercomputer kan ik geen eigen onderzoek doen naar klimaatveranderingen op de langere termijn.

Zoals eerder besproken worden voor weersvoorspellingen van meer dan 48 uur tot 10 dagen in de toekomst zeer complexe Global Climate Models gebruikt. Hoewel voor een termijn korter dan 48 uur geen globaal klimaat model gebruikt wordt, is het HIRLAM ook zeer complex. Wat ik in dit onderzoek wil vaststellen, is of het gebruik van deze zeer complexe modellen noodzakelijk is om een goede voorspelling te kunnen doen. Een geheel andere aanpak is bijvoorbeeld een statistisch model (regressiemodel) waarbij verschillende metingen als windsnelheid en neerslag als verklarende variabelen worden gebruikt.

Eén van de kerntaken van het KNMI is om te zorgen dat dit soort metingen publiekelijk beschikbaar zijn. Voor dit onderzoek zal ik daarom gebruik maken van meetgegevens van 1 januari 1991 tot en met 24 januari 2007 van weerstation De Bilt. Deze data zal ik gebruiken om de gemiddelde dagtemperatuur voor elke dag in de periode 2 januari 2001 tot en met 24 januari 2007 te voorspellen.

Ik heb twee statistische modellen opgesteld om het weer te voorspellen. Het eerste model is een regressiemodel dat met diverse weerkundige metingen van vandaag een voorspelling voor morgen probeert te doen. Het tweede model is een tijdreeksmodel dat op basis van een tijdreeks van 10 jaar terug tot vandaag een voorspelling voor morgen afgeeft. De modellen die ik heb opgesteld, zijn gemaakt om de gemiddelde dagtemperatuur van morgen te voorspellen. Deze waarden zijn te vergelijken met de werkelijke gemiddelde temperatuur op de verschillende dagen, omdat deze ook in de meetgegevens voorkomen.

Om zo een model te vergelijken met het HIRLAM van het KNMI, is het het meest wenselijk om de voorspellingen van het KNMI voor de gemiddelde temperatuur van morgen te hebben van 2 januari 2001 tot en met 24 januari 2007. Helaas worden deze gegevens niet publiekelijk beschikbaar gesteld. Daarom zal ik de modellen testen door de voorspellingen te vergelijken met een nulvoorspelling. Deze voorspelling is dat de gemiddelde dagtemperatuur van morgen gelijk is aan de gemiddelde dagtemperatuur van vandaag.

Ik zal bekijken of de twee statistische modellen die ik heb opgesteld beter presteren dan deze nulvoorspelling. Verder wordt een voorspelling die niet meer dan 1 graad Celsius afwijkt van de werkelijke temperatuur als goed gezien. Het is mogelijk om het percentage van de voorspellingen te berekenen dat in deze categorie valt. Zo kan de absolute performance van een model bepaald worden.

### 6.1 Regressiemodel

Voor het regressiemodel gebruik ik als verklarende variabelen metingen van het KNMI. In figuur 9 is het begin van de datafile te zien met metingen voor weerstation De Bilt van 1 januari 2001 tot en met 24 januari 2007.

STN	YYYYMMDD	FG	FHX	FX	TG	TN	TX	SQ	SP	DR	RH	PG	VVN	NG	UG
260	20010101	58	80	150	34	10	63	0	0	78	95	9966	30	8	91
260	20010102	60	70	150	87	56	112	27	34	9	3	9918	75	6	83
260	20010103	53	70	130	67	45	99	52	66	39	40	10009	58	6	86
260	20010104	62	80	140	78	58	91	0	0	51	49	9949	30	8	89
260	20010105	59	100	200	76	62	107	0	0	100	189	9874	33	8	92
260	20010106	40	60	100	59	41	76	19	24	29	21	9983	38	5	92
260	20010107	44	50	110	52	36	68	0	0	3	3	10080	60	6	91

Figuur 9. Diverse metingen van weerstation De Bilt

De afkortingen boven de kolommen hebben de volgende betekenis:

STN = stationsnummer (260=De Bilt)

YYYYMMDD = datum (YYYY=jaar MM=maand DD=dag)

FG = etmaalgemiddelde windsnelheid in 0.1 m/s

FHX = hoogste uurgemiddelde windsnelheid in 0.1 m/s

FX = hoogste windstoot in 0.1 m/s

TG = etmaalgemiddelde temperatuur in 0.1 graden Celsius

TN = minimum temperatuur in 0.1 graden Celsius

TX = maximum temperatuur in 0.1 graden Celsius

SQ = zonnenschijnduur in 0.1 uur (-1 indien minder dan 0.05 uur)

SP = percentage van de langst mogelijke zonnenschijnduur

DR = duur van de neerslag in 0.1 uur

RH = etmaalsom van de neerslag in 0.1 mm (-1 indien minder dan 0,05 mm)

PG = etmaalgemiddelde luchtdruk in 0,1 hPa

VVN = minimum opgetreden zicht (0=minder dan 100m, 1=100-200m, 2=200-300m, ..., 49=4900-500m, 50=5-6km, 56=6-7km, 57=7-8km, ..., 79=29-30km, 80=30-35km, 81=35-40km, ..., 89=meer dan 70km)

NG = bedekkingsgraad van de bovenlucht in octa's

UG = etmaalgemiddelde relatieve vochtigheid in procenten

De temperatuurmetingen zijn gedaan op een hoogte van 1.5 m.

De bedekkingsgraad van de hemel door wolken wordt ieder uur bepaald en weergegeven in achtsten (octa's). De waarde voor NG geeft het gemiddelde van de 24 uurlijkse waarnemingen. De betekenis van de waarden is als volgt.

0	Onbewolkt
1	Vrijwel onbewolkt
2	Licht bewolkt
3	Half bewolkt
4	Half bewolkt
5	Half tot zwaar bewolkt
6	Zwaar bewolkt
7	Vrijwel geheel bewolkt
8	Geheel bewolkt
9	Bovenlucht onzichtbaar

Het minimale zicht wordt op bemande stations elk uur visueel waargenomen door herkenning van markante punten in de omgeving van een meetstation, waarvan de afstand tot de waarnemer bekend is. Op automatische stations worden hiervoor zichtmeters gebruikt.

Neerslag omvat regen, motregen, onderkoelde regen, sneeuw, hagel, ijsregen, ijsnaalden, neerslag vanuit mist, dauw, enz. De hoeveelheid wordt uitgedrukt in millimeters regenwater.



10 Millimeter neerslag op 1 vierkante meter komt overeen met 10 liter regenwater per vierkante meter. De duur is de totale tijd (eventueel onderbroken) waarin sprake is van meetbare neerslag.

Wind heeft betrekking op de horizontale verplaatsing van lucht op een hoogte van 10 m boven het maaiveld. De gemiddelde snelheid is het gemiddelde van de 24 uurgemiddelde waarden in het etmaal. De hoogste van deze 24 uurgemiddelde waarden wordt apart vermeld onder: maximale uurgemiddelde snelheid. De gemeten hoogste uitschieter van de windsnelheid wordt vermeld als hoogste windstoot.

Ik heb besloten om de categorie VVN te interpreteren als een ratio-scaled variabele en dus niet als een nominale variabele. De waarden voor deze variabele zijn namelijk al redelijk ratio-scaled. Zie bijvoorbeeld de sprong van 50 voor 5-6 kilometer naar 56 voor 6-7 kilometer. Hetzelfde heb ik besloten voor de bedekkingsgraad NG, want een hogere waarde voor NG geeft een hogere bedekkingsgraad aan.

Als verklarende variabelen in het regressiemodel gebruik ik alle variabelen zoals hierboven besproken, behalve het stationsnummer (STN) en de datum (YYYYMMDD). Met de waarden voor deze variabelen op dag  $t$  probeer ik de gemiddelde dagtemperatuur van dag  $t + 1$  te voorspellen ( $TG_{t+1}$ ).

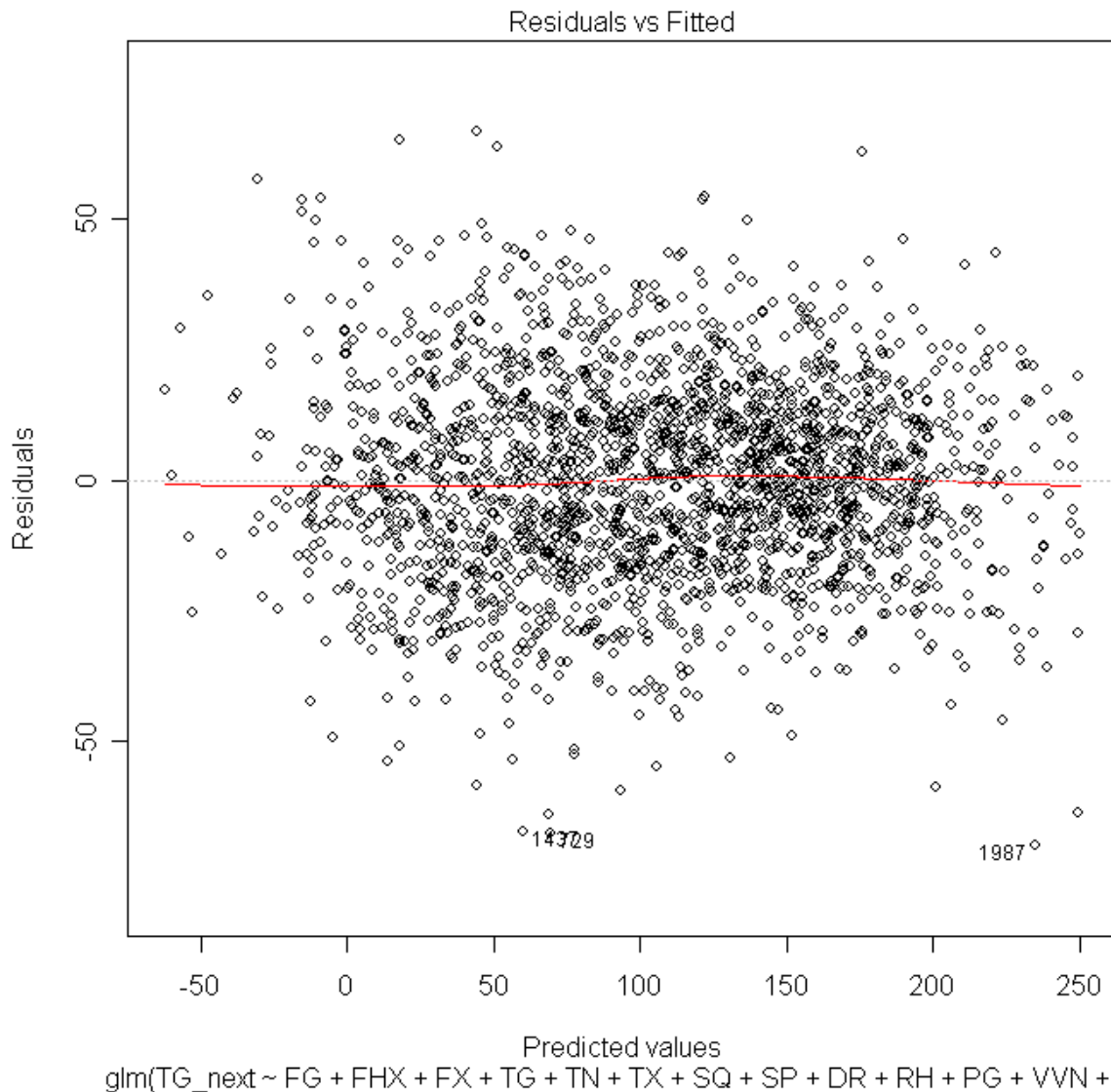
Ik heb dit regressiemodel als generalized linear model in het softwarepakket R gemodelleerd. De regressie wordt uitgevoerd op alle verklarende variabelen plus de interacties tussen elk tweetal verklarende variabelen. De waarden van deze verklarende variabelen voor dag  $i$  komen terug in de vector  $x_i$  in onderstaand model.

$$\Omega: \begin{cases} Y_i \sim N(\mu_i, \sigma^2) \\ \eta_i = x_i^T \beta \\ \eta_i = \mu_i \\ \text{for } i = 1, \dots, 2211 \end{cases}$$

Vervolgens heb ik met de functie `step()` de effecten zonder significante invloed uit het model verwijderd. De `step()` functie werkt op basis van het Akaike criterium, dat ook rekening houdt met het aantal parameters in het model. Het overgebleven generalized linear model verklaart de temperatuur van de volgende dag op basis van alle afzonderlijke verklarende variabelen (zoals hierboven gedefinieerd) plus de volgende interacties:

FG:FHX	FHX:TG	FX:TG	TG:TN	TN:TX	TX:SQ	SQ:RH	SP:DR	DR:RH	RH:VVN	PG:VVN	VVN:NG
FG:FX	FHX:TN	FX:TN	TG:SQ	TN:DR	TX:SP	SQ:PG	SP:RH	DR:PG		PG:NG	VVN:UG
FG:TN	FHX:SQ	FX:SP	TG:SP	TN:NG	TX:DR	SQ:NG	SP:VVN	DR:UG			
FG:TX	FHX:SP	FX:DR	TG:DR		TX:RH	SQ:UG					
FG:SQ	FHX:RH	FX:RH	TG:RH		TX:PG						
FG:SP	FHX:NG	FX:PG	TG:PG								
FG:RH		FX:VVN									
FG:PG		FX:UG									
FG:VVN											
FG:NG											
FG:UG											

De voorspellingen van dit model heb ik afgezet tegen de afwijking van de werkelijke temperatuur. Het resultaat is te zien in figuur 10.



Figuur 10. Plot van de voorspellingen tegenover afwijking van de werkelijke temperatuur

Zowel de residuen op de y-as als de voorspellingen op de x-as hebben als eenheid 0,1 graden Celsius. De voorspellingen lopen uiteen van -6,2 graden Celsius tot 25 graden Celsius. De afwijkingen lopen uiteen van -6,7 tot +7 graden Celsius.

De (denkbeeldige) band met residuen tussen -10 en 10 geeft de voorspellingen aan die minder dan 1 graad Celsius van de werkelijke gemiddelde dagtemperatuur liggen.

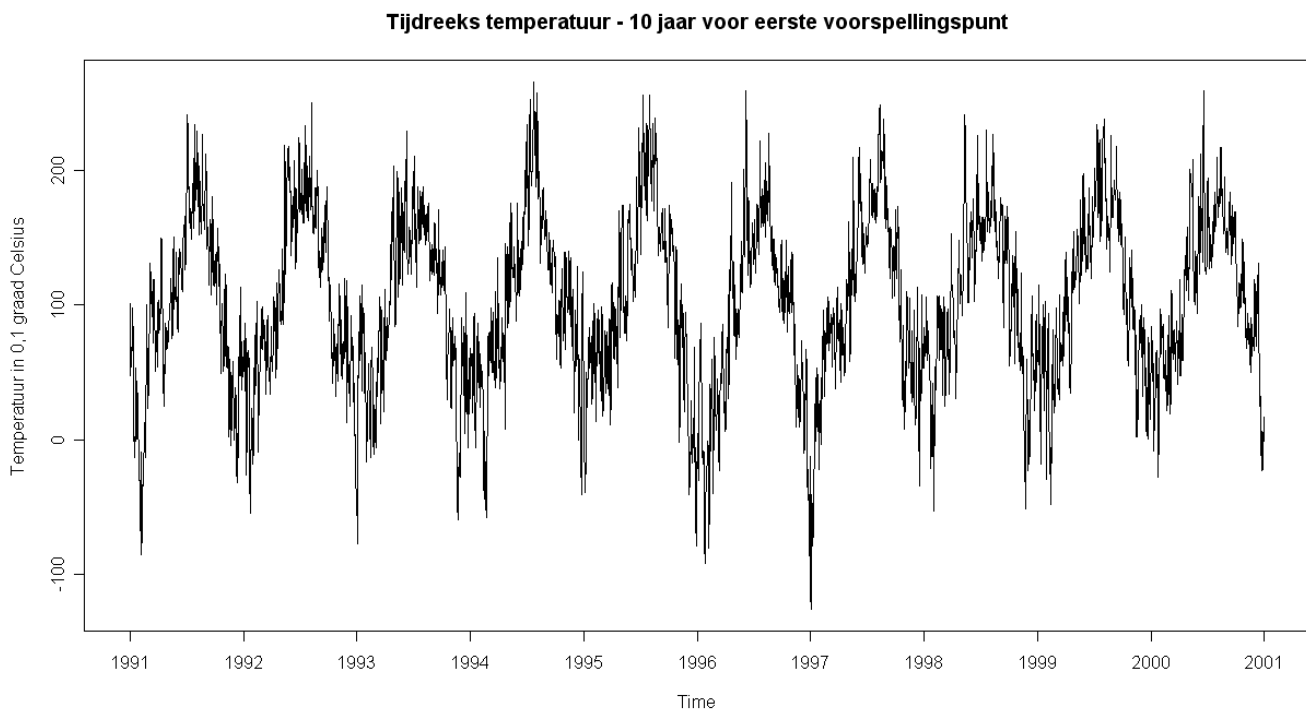
Een gedetailleerd overzicht van de resultaten zal ik geven na het bespreken van het tijdreeksmodel.

## 6.2 Tijdreeksmodel

Voor het tijdreeksmodel gebruik ik data over de gemiddelde dagtemperatuur in De Bilt uit de periode 1 januari 1991 tot en met 24 januari 2007. Voor elk voorspellingspunt in de periode van 1 januari 2001 tot en met 24 januari 2007 maak ik een tijdreeksmodel op basis van de gemiddelde dagtemperatuur van 0 tot 10 jaar terug. De gemiddelde dagtemperatuur van 29

februari in schrikkeljaren laat ik buiten beschouwing om makkelijker de seizoenscomponent te kunnen schatten. De voorspelling voor 1 maart 2004 zal ik daarom buiten beschouwing laten.

Als voorbeeld zal ik de procedure beschrijven om een voorspelling te maken voor de temperatuur van 2 januari 2001. Allereerst maak ik een tijdreeks van de gemiddelde dagtemperatuur van 2 januari 1991 tot en met 1 januari 2001. Deze tijdreeks is te zien in figuur 11.



Figuur 11. Tijdreeks gemiddelde dagtemperatuur van 2 januari 1991 tot en met 1 januari 2001

Er is duidelijk een seizoenscomponent aanwezig in deze tijdreeks. Een goed model voor deze tijdreeks is daarom:

$$X_{j,d} = m_j + s_d + Y_{j,d} \text{ voor } j = 1991, d = 2, \dots, 365 \text{ en } j = 1992, \dots, 2000, d = 1, \dots, 365 \text{ en } j = 2001, d = 1$$

Hierin is  $X_{j,d}$  de gemiddelde dagtemperatuur op jaar  $j$  en dag  $d$ ,  $m_j$  de trend component voor jaar  $j$ ,  $s_d$  de seizoenscomponent voor dag  $d$  en  $Y_{j,d}$  een stationaire tijdreeks.

Om de trend component  $m_j$  te schatten, wordt de gemiddelde temperatuur voor elk jaar bepaald volgens de formule:

$$\hat{m}_j = \frac{1}{365} \left( \left( \sum_{d=2}^{365} X_{j,d} \right) + X_{j+1,1} \right)$$

Deze formule is afhankelijk van de datum waarvoor de voorspelling gedaan wordt.

Vervolgens wordt de seizoenscomponent geschat door:

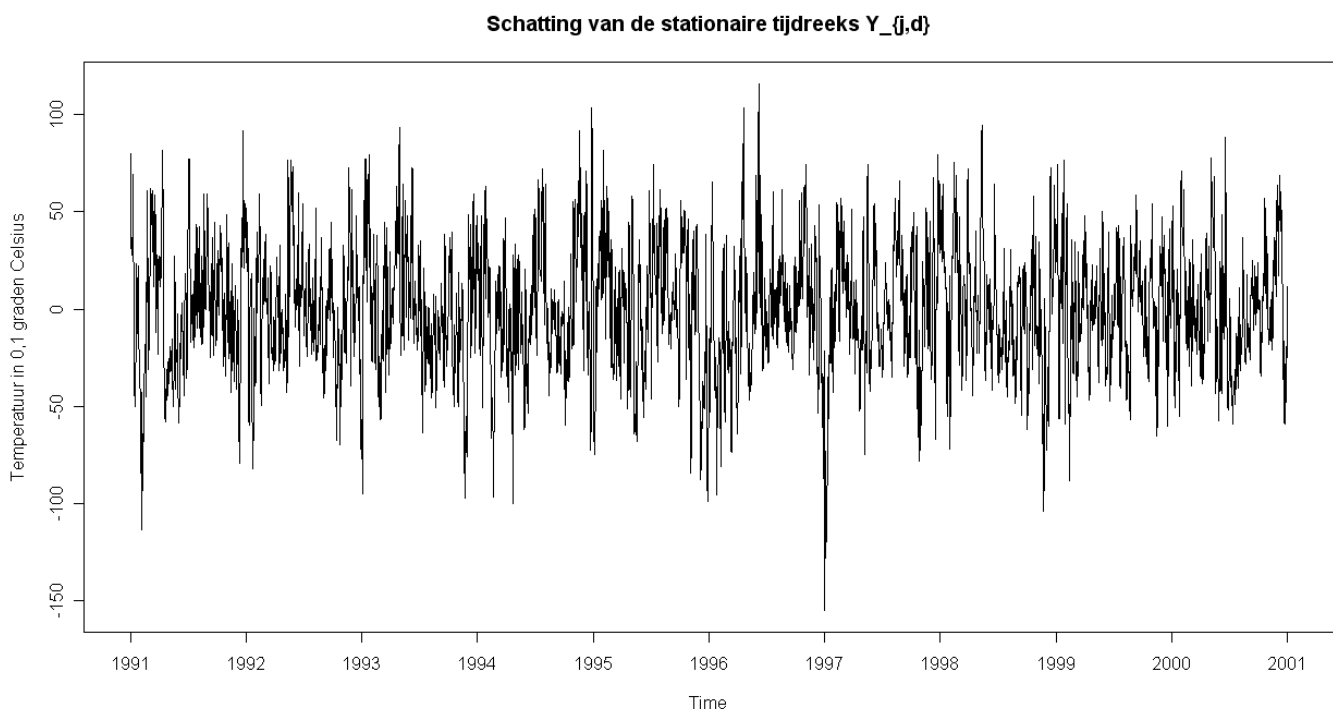
$$\hat{s}_d = \frac{1}{10} \sum_{j=1991}^{2000} X_{j,d} \text{ voor } d = 2, \dots, 365 \text{ en}$$

$$\hat{s}_d = \frac{1}{10} \sum_{j=1992}^{2001} X_{j,d} \text{ voor } d = 1$$

Een schatting voor de stationaire tijdreeks  $Y_{j,d}$  is vervolgens

$$\hat{Y}_{j,d} = X_{j,d} - \hat{m}_j - \hat{s}_d \text{ voor } j = 1991, d = 2, \dots, 365 \text{ en } j = 1992, \dots, 2000, d = 1, \dots, 365 \text{ en } j = 2001, d = 1$$

De schatting van de tijdreeks  $Y_{j,d}$  is te zien in figuur 12.



Figuur 12. Schatting van de tijdreeks  $Y_{j,d}$

Om deze overgebleven stationaire tijdreeks te modelleren, heb ik de tijdreeks onderzocht met behulp van de autocorrelatie functie en partiële autocorrelatie functie in R. Voor de verschillende periodes die ik onderzocht heb, vertoont de autocorrelatie functie geen cutoff punt. Daarom heb ik besloten om voor elke voorspelling deze tijdreeks te modelleren als een autoregressief proces. De orde van het autoregressieve proces wordt door R bepaald op basis van het Akaike criterium.

Na het modelleren van de tijdreeks  $X_{j,d}$  op bovenstaande manier is het mogelijk om een voorspelling te geven voor 2 januari 2001. Met R voorspel je het volgende punt van de stationaire tijdreeks, in dit geval  $Y_{2001,2}$ . Vervolgens wordt daarbij opgeteld de seizoenscomponent, in dit geval  $s_2$  voor 2 januari. Daarna wordt ook de schatting van de trend component voor het laatste jaar hierbij opgeteld. Het verkregen getal is de voorspelling van de gemiddelde dagtemperatuur voor 2 januari 2001.

Om alle voorspellingen volgens deze procedure te genereren heb ik volgende R code geschreven:

```

predictions<-numeric(2210)
for(i in 1:2210){
  temperature_matrix <- matrix(temperature[(i+1):(3650+i)],2),
  byrow=TRUE, ncol=365)
  mhat<-apply(temperature_matrix,1,mean)
  shat<-apply(temperature_matrix-mhat,2,mean)
  Yhat<- temperature_matrix-outer(mhat,shat,"+")
  remaining.ts<-
  ts(c(t(Yhat)[,1],t(Yhat)[,2],t(Yhat)[,3],t(Yhat)[,4],t(Yhat)[,5],t(Yh
  at)[,6],t(Yhat)[,7],t(Yhat)[,8],t(Yhat)[,9],t(Yhat)[,10]))
  ar_remaining<-ar(remaining.ts)
  predictions[i]<-
  predict(ar_remaining,se.fit=FALSE)[1]+shat[1]+mhat[10]
}

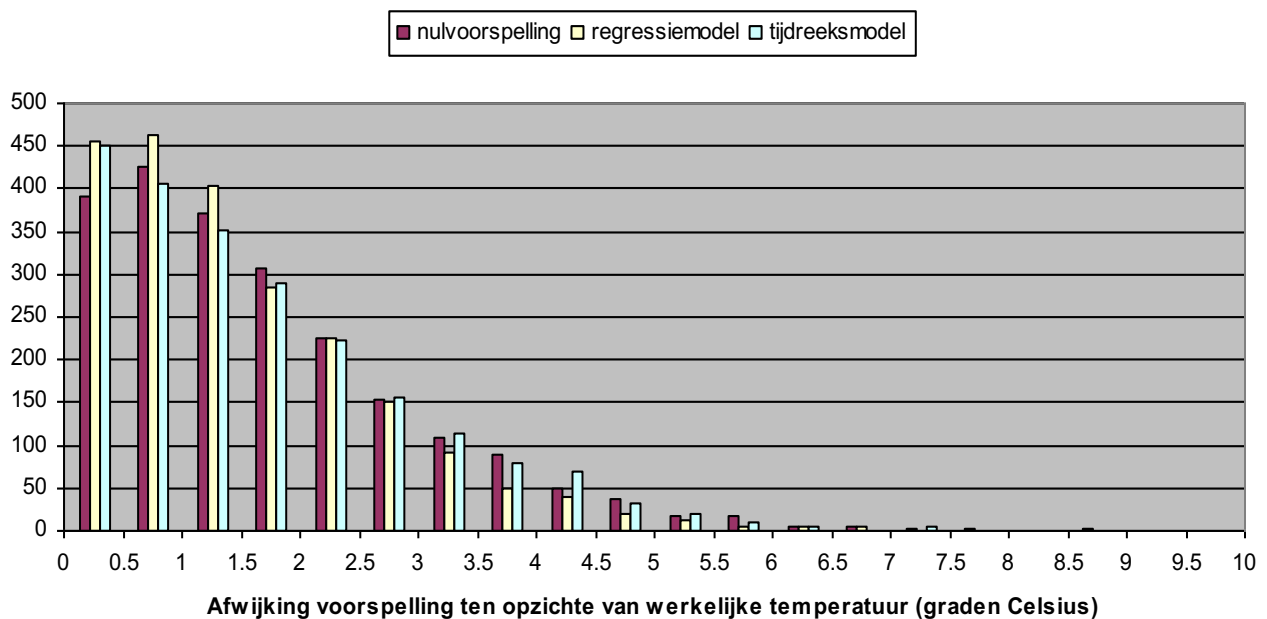
```

Na het draaien van deze code bevat de vector 'predictions' de voorspellingen voor de gemiddelde dagtemperatuur.

### 6.3 Resultaten

Na het verkrijgen van de voorspellingen van beide modellen is het zaak om de nauwkeurigheid van de voorspellingen te bepalen. Hiervoor heb ik de absolute afwijkingen van elk van de voorspellingen voor de periode 2 januari 2001 tot en met 24 januari 2007 berekend ten opzichte van de werkelijke gemiddelde dagtemperaturen. Van deze afwijkingen heb ik vervolgens een histogram gemaakt (zie figuur 13).

#### Histogram



Figuur 13 Histogram

In dit histogram zijn de afwijkingen te zien van zowel de nulvoorspelling, het regressiemodel als het tijdreeksmodel. In de bucket met een absolute afwijking tot een halve graad Celsius presteren zowel het regressiemodel als het tijdreeksmodel beter dan de nulvoorspelling. In de bucket met een absolute afwijking van een halve graad tot één graad Celsius presteert het regressiemodel beter dan de nulvoorspelling, maar het tijdreeksmodel slechter. Over het algemeen zijn de verschillen van beide modellen met de nulvoorspelling niet erg groot.

Voor zowel de nulvoorspelling, het regressiemodel als het tijdreeksmodel heb ik de gemiddelde absolute afwijking berekend. Deze zijn te zien in onderstaande tabel

	<b>Gemiddelde afwijking</b>
nulvoorspelling	1,64
regressiemodel	1,47
tijdreeksmodel	1,64

Ook hier zijn de minimale verschillen te zien. Het tijdreeksmodel heeft zelfs dezelfde gemiddelde absolute afwijking als de nulvoorspelling.

Als absolute performance heb ik van tevoren gedefinieerd het percentage voorspellingen dat niet meer dan 1 graad Celsius afwijkt van de werkelijke temperatuur. De resultaten zijn te zien in onderstaande tabel

	<b>Afwijking kleiner dan 1 graad (%)</b>
nulvoorspelling	36,9%
regressiemodel	41,6%
tijdreeksmodel	38,8%

Beide statistische modellen presteren dus beter dan de nulvoorspelling. Echter, het verschil in performance is erg klein. Een model waarbij in maar 40% van de gevallen de voorspelling maximaal 1 graad afwijkt van de werkelijke temperatuur is niet echt bruikbaar.

## 6.4 Conclusie eigen onderzoek

Dit onderzoek heeft aangetoond dat een relatief eenvoudig regressiemodel of tijdreeks niet in staat is om bruikbare voorspellingen af te leveren. De performance is nauwelijks beter dan de performance van de nulvoorspelling, die als voorspelling gewoon de temperatuur van vandaag geeft. De conclusie is daarom dat de zeer complexe modellen van het KNMI wel degelijk nodig zijn om goede voorspellingen te kunnen doen

Verder is één van de kerntaken van het KNMI om waarschuwingen af te geven bij extreme weersituaties. Om extreme weersituaties te voorspellen, zijn complexe klimaatmodellen zeer geschikt. Deze situaties kun je met een statistisch model niet goed zien aankomen.

## 7 Conclusie

Klimaatmodellen gericht op de korte termijn of de lange termijn zijn beiden erg complex. Er bestaan verschillende categorieën modellen, die gericht zijn op een verschillende voorspellingshorizon. Global Climate Models en Earth Models of Intermediate Complexity bestrijken samen een voorspellingshorizon van enkele uren tot 100.000 jaar in de toekomst. Global Climate Models zijn de meest complexe modellen en worden over het algemeen gebruikt tot ongeveer enkele eeuwen in de toekomst. De hele aarde wordt driedimensionaal gemodelleerd en veel klimatologische processen worden geïmplementeerd. Vervolgens wordt gesimuleerd hoe de toekomst er uit ziet door de klimatologische processen in te laten werken op de begintoestand. In EMICs zijn veel klimatologische processen vereenvoudigd of zelfs weggelaten. Hierdoor kan er een langere termijn mee gesimuleerd worden.

De ontwikkeling en nauwkeurigheid van klimaatmodellen loopt hand in hand met de stijging van de beschikbare reken capaciteit. In de meest complexe Global Climate Models zit nog niet alle kennis verwerkt die we hebben over klimatologische processen. Een model dat al deze kennis in zich heeft, zou op de snelste supercomputer niet te draaien zijn. Daarom zal in de toekomst de nauwkeurigheid van klimaatmodellen toenemen.

De beschikbare reken capaciteit bepaalt ook de voorspellingshorizon van de verschillende modellen. Als er een oneindige reken capaciteit zou zijn, zouden EMICs zelfs niet eens nodig zijn. Ook voor de voorspellingen verder dan 100.000 jaar in de toekomst zouden dan Global Climate Models gebruikt kunnen worden. Op dit moment is de toepasbaarheid van GCMs beperkt tot enkele eeuwen in de toekomst.

Aan de korte termijn van het spectrum van klimaatvoorspellingen liggen de weersvoorspellingen. Hiervoor zijn enkele gespecialiseerde modellen gemaakt. Het Ensemble Prediction System en het HIRLAM worden door het KNMI gebruikt voor weersvoorspellingen in Nederland. Vooral het Ensemble Prediction System is inventief doordat hier uitgegaan wordt van verschillende (verstoorde) beginsituaties. Hierdoor wordt de onnauwkeurigheid van de voorspellingen duidelijker weergegeven.

Door een steeds grotere reken capaciteit van supercomputers, zal de nauwkeurigheid van de weersvoorspellingen in de toekomst nog flink toenemen. Dit zal het dagelijkse leven van mensen veraangenamen en veiliger maken, doordat extreme weersituaties beter voorspeld kunnen worden en de mens hiervoor gewaarschuwd wordt.

Uit het eigen onderzoek kan ik concluderen dat complexe klimaatmodellen absoluut nodig zijn om nauwkeurige weersvoorspellingen te kunnen doen. Vooral voor het afgeven van waarschuwingen bij extreme weersituaties. Deze zal je met eenvoudige statistische modellen niet kunnen voorspellen, terwijl de complexe klimaatmodellen hier juist wel geschikt voor zijn.

## 8 Literatuurlijst

- Beckman, K. (1991), *De milieubeweging (7) De grote vervuiler*
- Damsté, J.S. (2006), *Wat doen de micro-organismen in de oceanen?* in oktobernummer 2006 van tijdschrift Quest
- ECMWF, <http://www.ecmwf.int/products/forecasts/d/epsmain>
- Intergovernmental Panel on Climate Change (2007), *Climate Change 2007: The Physical Science Basis*
- KNMI, [http://www.knmi.nl/over\\_het\\_knmi/](http://www.knmi.nl/over_het_knmi/)
- KNMI, [http://www.knmi.nl/waarschuwingen\\_en\\_verwachtingen/extra/pluim.html](http://www.knmi.nl/waarschuwingen_en_verwachtingen/extra/pluim.html)
- McGuffie, K. en A. Henderson-Sellers (1997), *A Climate Modelling Primer*
- McGuffie, K. en A. Henderson-Sellers (2001), *Forty Years of Numerical Climate Modelling*
- Ministerie van VROM, <http://www.vrom.nl/pagina.html?id=22062#b22066>
- Schneider (1992), *Coupled Climate System Modelling*
- Tijm, S. en B.W. Schreur (2001), *HIRLAM in meer detail*
- Wigley (1998), *The Kyoto Protocol: CO<sub>2</sub>, CH<sub>4</sub> and climate implications*
- Wikipedia (geen auteur bekend):
- <http://nl.wikipedia.org/wiki/Cryosfeer>
- <http://nl.wikipedia.org/wiki/Kyoto-protocol>
- [http://nl.wikipedia.org/wiki/Radicaal\\_%28scheikunde%29](http://nl.wikipedia.org/wiki/Radicaal_%28scheikunde%29)
- <http://nl.wikipedia.org/wiki/Vlindereffect>