The influence of data quality on the detectability of sea-level height variations

K.I. van Onselen

NCG Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission

Delft, December 2000

Colophon

The influence of data quality on the detectability of sea-level height variations
K.I. van Onselen
Publications on Geodesy 49
ISBN 90 6132 273 1
ISSN 0165 1706

Publications on Geodesy is the continuation of Publications on Geodesy New Series

Published by: NCG Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission, Delft, The Netherlands
Printed by: Meinema Drukkerij, Delft, The Netherlands
Cover: Bi-linear pattern in sea-level rise estimated through trend values (in mm/yr) for 18 tide gauges in the North Sea area. * Indicate locations of tide gauges with corresponding trend values.

NCG Nederlandse Commissie voor Geodesie
P.O. Box 5030, 2600 GA Delft, The Netherlands
Tel.: +31 (0)15 278 28 19
Fax: +31 (0)15 278 17 75
E-mail: ncg@geo.tudelft.nl
Website: www.ncg.knaw.nl

The NCG Nederlandse Commissie voor Geodesie Netherlands Geodetic Commission is an institute of the Royal Netherlands Academy of Arts and Sciences (KNAW).

Contents

Summary

For low-lying areas like the Netherlands, an ever-increasing sea level can become a serious threat. This is especially true if sea level rise accelerates, e.g., due to greenhouse-gas induced warming. To anticipate potential troubles, it is important to have a good estimate of the expected behaviour of future sea levels. This requires an accurate description of the present-day sea level variation curve and of foreseeable changes in this "natural" pattern in the near future. These changes in the behaviour of future sea levels can be based, e.g., on models predicting global change, but this is beyond the scope of this thesis.

Much simplified, sea level rise over the last century could be described by a linear regression line. Accelerations of this "natural" pattern have occurred if the slope value of the regression line increases, or higher order regression coefficients are required to describe the sea-level rise curve. The better the natural sea level variation curve (as has occurred over the last century) can be determined, the easier it will be to detect a significant divergence from this curve. The objective of this thesis is to determine how well patterns in sea level height variations can be detected, given the limited quality of the data available.

The objective of this thesis requires long sea level height time series. Therefore, only tide gauge data has been used and altimetry sea level height series have not been considered. Tide gauges measure sea level heights relative to the tide gauge bench marks. Consequently, the resulting sea level height time series show both variations in absolute sea level and vertical movements of the tide gauge bench marks. By monitoring the height changes between the tide gauge bench marks and a stable reference height, these relative sea level heights can (in principle) be converted into absolute sea level changes. Unfortunately, locating a reference point which is truly stable over long time spans will be extremely difficult, if not impossible.

How well a specific sea level variation pattern can be detected depends on the sea level variations themselves, the quality of the tide gauge measurements and, if applicable, the quality of geodetic measurements used to connect the tide gauge bench marks in height. Based on existing literature, it has been tried to gain a clear understanding of these various aspects. Unfortunately, in the literature studied on processes which can influence relative sea-level heights, (almost) no mention is made of long-periodic processes (periods over 20 years), while analysis of tide gauge records shows that long-periodic fluctuations with significant amplitudes do occur in sea level height time series.

Sea level heights as used in this thesis are annual mean sea levels. The quality of these annual mean values not only depends on the quality of the tide gauge measurements, but also on the frequency of these measurements. Not only the quality of state-of-the-art techniques is important, but also of tide gauges and measuring frequencies which were used in the past. Since estimating long-term sea level variation curves requires long sea level height series, historical measurements have to be used as well. In chapter 3, an overview is given of the measuring precision and systematic errors and limitations characteristic for the six tide gauge systems commonly used during the last century. Based on information available for Dutch tide gauges, an estimate is given of how much the quality of annual mean sea levels deteriorates if mean values are based on, e.g., mean tide levels instead of on hourly measurements.

If data for a number of tide gauges is available, a common sea level variation curve, e.g., applying to the Dutch coast, can be estimated. Since tide gauge measurements are relative to the local tide gauge bench mark, any vertical movements of the tide gauges relative to one another will have introduced inconsistencies between the individual time series. These inconsistencies reduce the quality of a common sea level variation curve based on these tide gauge series.

As long as tide gauges experience only secular height movements relative to one another, the common oscillation pattern can still be discerned using techniques like SVD. However, the slope of the estimated common variation curve is determined by the rate of vertical movements of the individual tide gauges. If tide gauges undergo vertical movements which vary in rate and over time, the common oscillation pattern will be affected as well. By relating all sea level height series to the same reference frame (e.g., NAP) internal differences in relative sea level due to vertical movements of the tide gauge bench marks are removed from the data sets. Ideally, permanent monitoring of the tide gauge bench marks is applied.

Nowadays, this can be achieved by means of GPS. However, in the past height differences were usually based on spirit levelling.

In chapter 6, the quality of three geodetic techniques, i.e., GPS, gravimetry, and spirit levelling is described. In addition, limitations of these techniques when applied to monitoring height changes of tide gauge bench marks are discussed. Since changes in local gravity represent both variations in mass and changes in station height, gravimetry is not well suited for determining height differences. Uncertainties in height differences obtained by GPS can be reduced to within 1 cm. However, the quality of these measurements might be less in harbour areas (e.g., due to signal interference). GPS has the advantage that it allows for permanent monitoring over large distances, but measurements are only available for the last few decades. Spirit levelling can produce high precision height differences (over short distances), but is time consuming and prone to systematic errors (especially over long distances). However, levelled height differences are often the only type of height information available.

In the past, tide gauge bench marks have (hopefully) been connected to a local reference frame. Between some neighbouring local height datums, height differences have occasionally been obtained as well. However, only since the second European levelling network (UELN-73), the height difference between the continent of Europe and Scandinavia and Great Britain respectively is available. These height connections consist of only a single connection line and, consequently, errors in these height differences cannot be detected by testing. In chapter 8, an indirect method is introduced for connecting vertical datums, which results in dynamic height differences between the fundamental stations in the various height datum zones. An advantage of this method is that quality information (both precision and reliability) of the estimated height differences can de determined as well. Unfortunately, a high quality potential coefficient model is required. As a result, only if a new model (to be obtained from the planned GOCE mission) becomes available, height differences between datum zones could be derived with standard deviations of 1 cm.

The quality of sea level variation curves depends on the method used to estimate these curves. A number of data analysing techniques have been tested for their suitability for working with sea level height data. Sea level height time series have a number of specific characteristics, for instance non-stationarity, data quality which is not constant for the complete time series, and a wide range of periodic fluctuations with sometimes variable frequencies and amplitudes. As a result, most of the techniques examined do not work well when applied to sea level height data. It is found that the best techniques for smoothing sea level height series are moving average smoothing and Singular Spectrum Analysis, while estimates of future sea level heights should be based on either AR(I)MA modelling or regression.

To determine how well specific sea level variation patterns can be detected, experiments with a large variety of simulated sea level height time series have been performed. These simulated time series consist of the curve which needs to be detected (e.g., a linear trend), periodic fluctuations (based on actual tide gauge data) and simulated additional errors. This can either be inaccuracies introduced by the tide gauge equipment or the height measurements, or (uncorrected for) height variations between tide gauge bench marks. By applying regression to the simulated time series, it is examined whether or not the original sea level variation curve can be recovered. It should be noted that statistical significance of estimated regression coefficients is no guarantee that the "true" sea level variation curve is detected. For example, if linear regression is applied to a sea level series following a quadratic curve, the estimated trend value can still be statistically significant. For this reason, often trend estimates are shown as a function of an increasing number of observations. For the above mentioned example, estimated trend values will steadily increase with an increasing number of included observations. Only if the model (of a linear regression line) fits the data, and if enough observations are available, estimated trend values will stabilise around the trend value actually present in the data set.

First, experiments have been performed with sea level height data for a single tide gauge. In this case, the original data relative to the tide gauge bench mark can be used. If (based on external knowledge of the behaviour of the local sea level) long-periodic fluctuations could be eliminated from the data set, the detectability of a single linear regression line depends on the trend value and the noise level of the measurements. For sea level data with a trend of 1.5 mm/yr, even if a noise level of 5 cm applies, this trend can be detected if 35 observations are available. If a simulated time series contains long-periodic fluctuations based on data for tide gauge Den Helder, of the order of 90 years of observations are required before trend estimates stabilise around the actual trend value on which the data set is based. Therefore, it is concluded that long-periodic fluctuations are the main factor in determining the amount of data required to detect a linear trend in a sea level height time series.

In chapter 7, using six tide gauge data sets, a common sea level variation curve for the Dutch coast is estimated. In order to eliminate deviations from this common curve caused by height variations of the tide gauge bench marks relative to one another, all tide gauges have to be connected in height to the local reference system (NAP). Inaccuracies in the required height connections introduce inconsistencies between the time series. Since the actual height connection history for the tide gauges is unknown, a number of scenarios have been used to simulate height connection errors. Experiments show that the quality of the estimated common variation curve not only depends on the precision of the height measurements, but also on the time span between subsequent height connections. For higher levels of connection noise, it is more pronounced that the larger the time span between subsequent connections, the less dependable the estimated trend values will be. In order to detect future sea level rise accelerations, historical data has to be used as well. Experiments show that, if long periods have elapsed between historic height connections, the precision of future height connections is of almost no importance. Increasing the standard deviation of future height measurements from 5 mm to 2 cm, or increasing the time span between height connections from one to 10 years, hardly influences the results.

Finally, for the North Sea area, the quality of spatial variation patterns which can be derived based on trend values for 18 tide gauges, is examined. A spatial pattern in sea level height variations should be based on real differences in trend values for the various locations and not on variations resulting from measuring errors and height changes between tide gauge bench marks. Based on experiments with simulated time series, the following conclusions have been made. If height connections to a local reference frame are performed every 10 years, ranges of errors in trend estimates (as a function of latitude and longitude) are three times as large as results based on annual connection of heights. As a result of, e.g., post-glacial rebound, fundamental stations in the different datum zones can experience height changes relative to one another. If the individual time series (connected to the local datums) are not corrected for these relative vertical movements, this will result in large errors in the estimated spatial variation pattern. If height differences between vertical datum zones are based on results derived for European levelling networks, resulting errors in trend values (as a function of latitude and longitude) will be much larger than those caused by the post-glacial rebound movements (of the selected fundamental stations: Amsterdam, Newlyn, and Helsingborg) itself. This same holds for differences in vertical movements obtained by GPS measurements with a standard deviation of the order of 1 mm/yr.

Samenvatting

Zelfs al is de gemiddelde zeespiegelstijging klein, voor laaggelegen gebieden zoals Nederland kan dit op de langere termijn een flinke bedreiging vormen. Deze problemen verergeren als het tempo van zeespiegelstijging versnelt, bijvoorbeeld ten gevolge van het broeikaseffect. Om zo goed mogelijk op potentiële problemen in te kunnen springen, zijn goede voorspellingen van toekomstige zeehoogtes nodig. Hiervoor zijn twee dingen nodig, namelijk inzicht in het patroon dat zeehoogtes tot nu toe hebben gevolgd en te verwachten veranderingen ten aanzien van dit "natuurlijke" patroon. Deze veranderingen in het natuurlijke patroon kunnen worden gebaseerd op een GCM (global change model), maar dit is een onderwerp dat niet in dit proefschrift behandeld wordt.

In een verregaande versimpeling kan de zeespiegelstijging zoals die de afgelopen eeuw heeft plaatsgevonden beschreven worden door een eenvoudige, lineaire, trend. Versnellingen in zeespiegelstijging vinden plaats als de trendwaarde toeneemt, of een hogere orde regressiecoëfficiënt nodig is om het gedrag van de zeespiegel te beschrijven. Des te beter het "natuurlijke" patroon in zeespiegelstijging (over de afgelopen eeuw) beschreven kan worden, des te beter (en eerder) versnellingen ten opzichte van dit patroon gedetecteerd kunnen worden. Het doel van dit proefschrift is te bepalen hoe goed specifieke patronen in zeespiegelstijging gedecteerd kunnen worden, ondanks de beperkte kwaliteit van de beschikbare data.

Om de doelstelling van dit proefschrift te verwezenlijken zijn erg lange tijdreeksen met zeehoogtes nodig. Om die reden is alleen gebruik gemaakt van peilschaaldata en niet van zeehoogtes bepaald door satelliet radar altimetrie. Een peilschaal produceert zeehoogtes relatief ten opzichte van het peilmerk bij de peilschaal. Hierdoor bevat de verkregen meetreeks niet alleen veranderingen in absoluut zeeniveau, maar ook de verticale bewegingen van de peilschaal zelf. Door hoogteverschillen te registreren tussen het peilmerk en een stabiel referentiepunt kunnen de relatieve zeehoogtes die de peilschaal produceert omgerekend worden naar veranderingen in absoluut zeeniveau. Helaas is het in de praktijk erg moeilijk (zo niet onmogelijk) om een referentiepunt te vinden dat daadwerkelijk stabiel is over langere tijdsperiodes.

Hoe goed het patroon waarmee zeehoogtes variëren gedetecteerd kan worden hangt af van drie factoren: van de veranderingen in het zeeniveau zelf, van de kwaliteit van de peilschaalmetingen en eventueel nog van de kwaliteit van de geodetische metingen die gebruikt zijn om peilschalen in hoogte aan te sluiten aan een lokaal referentiestelsel. Uit literatuur die over deze onderwerpen verschenen is, is getracht inzicht te krijgen in deze verschillende factoren. Weinig informatie bleek voor handen te zijn over lang-periodieke fluctuaties (periodes langer dan 20 jaar) in zeehoogtes. Uit analyse van peilschaaldata blijkt echter dat deze langgolvige effecten wel degelijk aanwezig zijn.

De resultaten in dit proefschrift zijn gebaseerd op jaargemiddeldes. De kwaliteit van deze jaargemiddeldes hangt niet alleen af van de kwaliteit van de individuele metingen, maar ook van de frequentie waarmee de metingen zijn uitgevoerd. Omdat lange meetreeksen nodig zijn, wordt ook gebruik gemaakt van historische data. Daarom is niet alleen de kwaliteit die bereikt kan worden met het nieuwste type peilschalen interessant, maar is het ook belangrijk om de kwaliteit te kennen van de metingen die in het verleden gedaan zijn. Hoofdstuk 2 bevat een overzicht van de zes meest gebruikte types peilschalen, hun kwaliteit en hun specifieke eigenschappen en tekortkomingen. Op basis van informatie die voor Nederlandse peilschalen beschikbaar is, is in dit hoofdstuk ook beschreven hoe de relatie is tussen de kwaliteit van het jaargemiddelde en het aantal dagelijkse metingen.

Als voor meerdere peilschalen data beschikbaar is, dan kan dit gebruikt worden om een gezamenlijke curve in zeespiegelvariatie uit te rekenen, die bijvoorbeeld het gedrag van het zeeniveau langs de Nederlandse kust weer geeft. Peilschalen bepalen zeehoogtes ten opzichte van het lokale peilmerk. Als de peilschalen in hoogte ten opzichte van elkaar bewegen, dan ontstaan hierdoor afwijkingen tussen de verschillende meetreeksen. Dit beïnvloedt op zijn beurt de kwaliteit van de gezamenlijke curve die uit deze datasets bepaald wordt.

Zolang onderlinge hoogteveranderingen tussen peilschalen seculair van aard zijn, kunnen de gezamenlijke fluctuaties in zeeniveau nog steeds goed bepaald worden met behulp van een techniek als Singular Value Decomposition. Het is dan echter niet langer mogelijk om een goede schatting te maken van de trend in zeespiegelstijging langs de Nederlands kust. Door alle peilschalen regelmatig in hoogte te koppelen aan het nationale referentiestelsel (NAP), kunnen de relatieve veranderingen in gemeten zeehoogte ten gevolge van hoogteveranderingen van de peilschaal t.o.v. NAP uit de data geëlimineerd worden. Bij voorkeur vindt permanente registratie plaats van eventuele hoogteveranderingen van de peilschalen, wat echter pas sinds een aantal jaren mogelijk is. Tegenwoordig kan namelijk een permanente GPS ontvanger bij de peilschaal geplaatst worden. In het verleden konden hoogteverschillen alleen bepaald worden door middel van waterpassen.

Hoofdstuk 5 geeft een beschrijving van de kwaliteit van drie geodetische technieken, waterpassen, GPS, en gravimetrie, die gebruikt kunnen worden om hoogteveranderingen van peilschalen te meten. Alle drie deze technieken hebben zo hun specifieke tekortkomingen als ze gebruikt moeten worden om hoogtes van peilschalen te controleren. Gravimetrie is niet erg geschikt om hoogteverschillen te bepalen, aangezien een gemeten verschil in zwaartekracht zowel het gevolg kan zijn van een hoogteverandering als van een massaverandering. Als GPS gebruikt wordt, dan kunnen onzekerheden in de metingen teruggebracht worden tot waardes kleiner dan 1 cm. Helaas kan de kwaliteit van met GPS bepaalde hoogtes in havens vaak relatief slecht zijn, bijvoorbeeld ten gevolge van interferentie met signalen van communicatieapparatuur. Een voordeel van GPS is weer dat ook over lange afstanden hoogteverschillen min of meer continu bepaald kunnen worden. Helaas zijn GPS metingen pas sinds enkele tientallen jaren beschikbaar. Waterpasgegevens zijn beschikbaar sinds het begin van de 19e eeuw. Met behulp van waterpassen kunnen hoogteverschillen met een zeer goede precisie bepaald worden. Dit gaat echter alleen op bij korte afstanden. Over langere afstanden is waterpassen erg gevoelig voor systematische fouten.

In het verleden zijn peilschalen in hoogte gekoppeld aan het nationale referentiestelsel. Tussen sommige van deze nationale referentiesystemen zijn ook hoogteverschillen bepaald. Het hoogte-verschil tussen West Europa en Engeland of Scandinavië is pas voor het eerst bepaald in het tweede Europese waterpasnetwerk (UELN-73). Helaas bestaat zowel de connectie tussen Scandinavië en Europa als de aansluiting tussen Engeland en West Europa slechts uit één gewaterpast traject. Het gevolg hiervan is dat fouten in deze aansluitingen niet gevonden kunnen worden door toetsing. In hoofdstuk 7 wordt een indirecte methode beschreven om verticale datums in hoogte met elkaar te verbinden. Deze methode resulteert in verschillen in dynamische hoogte tussen fundamentele stations in de verschillende gebieden. Het voordeel van deze methode is dat ook de kwaliteit van de hoogteverschillen (zowel precisie als betrouwbaarheid) bepaald kan worden. Het nadeel van deze methode is dat om hoogteverschillen met een goede kwaliteit te krijgen ook een geopotentiaalmodel van zeer hoge kwaliteit nodig is. Pas als een nieuw model beschikbaar komt (te verkrijgen uit de geplande GOCE satelliet missie) kunnen hoogteverschillen tussen datumzones bepaald worden met standaardafwijkingen van ongeveer 1 cm.

De kwaliteit van een geschatte curve door een meetreeks van zeehoogtes hangt ook af van de techniek die gebruikt is om de curve te schatten. In dit proefschrift zij een aantal technieken onderzocht op hun toepasbaarheid bij het analyseren van zeespiegelfluctuaties. De gebruikte tijdreeksen met zeehoogtes hebben namelijk een aantal specifieke kenmerken. Om een paar voorbeelden te noemen: ze zij niet stationair, de kwaliteit van de metingen is niet constant, en ze bevatten een groot aantal periodieke fluctuaties die niet noodzakelijkerwijs een vaste frequentie en amplitude hebben. Door deze specifieke eigenschappen blijken de meeste van de onderzochte technieken niet goed toepasbaar op reeksen met zeeniveau's. De beste technieken om een meetreeks te smoothen blijken moving average smoothing en Singular Spectrum Analysis, terwijl extrapolatie van de meetreeks het beste kan gebeuren met behulp van AR(I)MA modellering of regressie.

Om te bepalen hoe goed specifieke curves gedecteerd kunnen worden, zijn voor dit onderzoek experimenten uitgevoerd met een groot aantal tijdreeksen met gesimuleerde zeehoogtes. Deze gesimuleerde data sets zijn opgebouwd uit de gesimuleerde curve (bijv. een lineaire trend), periodieke fluctuaties (gebaseerd op echte peilschaaldata) en toegevoegde fouten. Dit kunnen meetfouten zijn geïntroduceerd door de peilschalen zelf, fouten die voortkomen uit hoogteveranderingen van de peilschalen (waar niet voor gecorrigeerd is) of onnauwkeurigheden in de hoogteaansluitingen van de peilschalen. Gekeken wordt of door middel van regressie de originele variatiecurve uit de (gesimuleerde) metingen terug gevonden kan worden. Hierbij dient opgemerkt te worden dat als regressie coëfficiënten bepaald worden die statistisch significant zijn, dit geen garantie is dat de "echte" variatiecurve gevonden is. Ter illustratie, als lineaire regressie toegepast wordt op een kwadratische curve, dan is het goed mogelijk dat de geschatte trendwaarde statistisch significant is. Om deze reden worden in dit proefschrift trendschattingen vaak getoond als functie van het aantal waarnemingen waarop ze gebaseerd zijn. Voor het genoemde voorbeeld geldt dat geschatte trendwaardes gestaag zullen stijgen naarmate meerdere schattingen gebaseerd worden op meer waarnemingen. Slechts als het model van een lineaire trend past bij de data (en voldoende waarnemingen beschikbaar zijn) zullen trendschattingen zich stabiliseren rond de trendwaarde waarop de tijdreeks daadwerkelijk gebaseerd is.

In hoofdstuk 4 worden experimenten uitgevoerd met een enkele meetreeks. Als, in een hypothetisch geval, alle lang-periodieke fluctuaties uit de data geëlimineerd kunnen worden, dan is de detecteerbaarheid van een lineaire trend nog slechts afhankelijk van de trendwaarde zelf en het ruisniveau van de metingen. Zelfs als de metingen meetruis bevatten met een standaardafwijking van 5 cm, blijkt 35 jaar data voldoende om een trend van 1.5 mm/jaar te detecteren. Als de gesimuleerde tijdreeks periodieke fluctuaties bevat gebaseerd op data voor peilschaal Den Helder, dan blijkt dat minimaal 90 jaar data nodig is voordat de trendschattingen zich stabiliseren rond de trendwaarde waarop de meetreeks gebaseerd is. Het kan dan ook geconcludeerd worden dat de aanwezigheid van lang-periodieke fluctuaties in peilschaaldata de beperkende factor zijn voor het aantal jaar data dat nodig is voor een stabiele trendschatting.

In hoofdstuk 6 worden gezamenlijke variatiecurves in zeeniveau langs de Nederlandse kust geschat, gebaseerd op meetreeksen voor zes peilschalen. Om de invloed van onderlinge hoogteveranderingen van de peilschalen uit deze gezamenlijke curve te elimineren, worden alle peilschalen in hoogte aangesloten aan NAP. Onnauwkeurigheden in deze hoogteaansluitingen introduceren op hun beurt weer verschillen tussen de meetreeksen. Aangezien het erg moeilijk te achterhalen is wanneer hoogteaansluitingen van peilschalen daadwerkelijk hebben plaatsgevonden, en vooral wat de kwaliteit hiervan is, zijn hoogtefouten gesimuleerd op basis van een aantal scenario's. Uit experimenten met verschillende scenario's van hoogteaansluitingen blijkt dat de kwaliteit van de gezamenlijke curve niet alleen bepaald wordt door de precisie van de hoogteaansluitingen, maar ook door hun frequentie. Naarmate het ruisniveau van de hoogteaansluitingen groter wordt, geldt steeds meer dat langere periodes tussen de aansluitingen zorgen voor een schatting van de gezamenlijke trend die steeds slechter wordt.

Om toekomstige versnellingen in zeespiegelvariatie op te sporen moet ook historische data gebruikt worden. Uit experimenten met gesimuleerde data blijkt dat als er in het verleden vrij veel tijd heeft gezeten tussen opeenvolgende hoogteaansluitingen, de precisie van toekomstige hoogteaansluitingen (binnen zekere grenzen) nauwelijks meer van invloed is. Een verbetering in de standaardafwijking van toekomstige aansluitingen van 2 cm naar 5 mm, of permanent metingen in plaats van 10-jaarlijkse campagnes, blijkt nauwelijks invloed te hebben op de resultaten.

Als laatste is ook, voor het Noordzee gebied, gekeken hoe goed een ruimtelijk patroon in variaties in zeespiegelstijging bepaald kan worden op basis van data van 18 peilschalen. Zo'n ruimtelijk patroon zou gebaseerd moeten zijn op daadwerkelijke verschillen in trendwaardes voor de verschillende locaties, en niet het gevolg moeten zijn van meetfouten en van hoogteveranderingen tussen de peilschalen zelf. Gebaseerd op experimenten met gesimuleerde meetreeksen, zijn de volgende conclusies getrokken. Als peilschalen eens in de 10 jaar aangesloten zijn aan een lokaal referentiestelsel, dan zijn fouten in trendwaardes (als een functie van de locatie) drie keer zo groot als de fouten die voorvloeien uit jaarlijkse aansluiting. De fundamentele stations in de datumzones zullen ook in hoogte ten opzichte van elkaar bewegen, bijvoorbeeld als een gevolg van post-glacial rebound. Als de individuele meetreeksen (die in hoogte gekoppeld zijn aan de lokale datums) niet voor deze hoogteveranderingen worden gecorrigeerd, dan ontstaan hierdoor grote vertekeningen in het spatiele patroon dat door de trendwaardes geschat wordt. Als hoogteverschillen tussen de datumzones gebaseerd worden op gegevens uit de Europese waterpasnetwerken, dan ontstaan hierdoor fouten in geschatte trendwaardes (als functie van hun locatie) die veel groter zijn dan de fouten veroorzaakt door de verticale bewegingen van de fundamentele stations Amsterdam, Newlyn en Helsingborg. Dit geldt ook als hoogteveranderingen bepaald worden met behulp van GPS, terwijl deze metingen een standaardafwijking hebben van 1 mm/jaar. Kortom, zolang de meetprecisie van beschikbare methoden om de hoogteaansluiting te verrichten niet verbetert, is het beter om niet te corrigeren voor de hoogteveranderingen tussen de genoemde fundamentele stations.

Acknowledgements

Many people have helped me with the realisation of this PhD thesis; by supplying information I've needed on a number of subjects, by giving useful suggestions, and by making my stay and the Delft University more pleasant. Special thanks go to:

Roland Klees, for his support, the discussions we had and all his help in improving the earlier versions of this thesis.

My colleges at FMR who provided useful information and discussions on a wide range of subjects. In addition, they have provided an atmosphere which I found enjoyable to work in. Special thanks go to Lucy Mathers for her many suggestions on the use of the English language. Johannes Bouman has been my "room mate" for almost seven years. I'm grateful, not only for his company, but also for his help in solving many of the problems I've encountered while working on my thesis. I would also like to give special thanks to Wil Coops and Rene Reudink. Although they did not directly contribute to the scientific part of this thesis, they've given me a lot of moral and practical support.

Erik de Min, who has provided me with the software I required to estimate the variance-covariance matrix of geoid heights as used in the connection of vertical datums.

Radboud Koop, for the simulation of the "GOCE variance-covariance matrix", and for his moral support.

Phil Woodworth (from PSMSL) who has put a lot of effort in improving the quality of my thesis. Both by pointing out a number of errors in my use of the English language and by providing me with a wealth of additional information on a large number of subjects. I'm also grateful to Hans-Peter Plag (from Oslo University) for his remarks on aliasing effects caused by the use of annual mean values.

Last but not least, I am much obliged to my supervisor Martin van Gelderen. Without his ongoing moral and practical support, this PhD thesis might not have been realised.

Abbreviations

AP	Amsterdams peil
ARIMA	autoregressive-integrated moving average
ARMA	autoregressive-moving average
EGM96	Earth gravity model 1996
FG5	"free-fall absolute gravimeter"
FSO	fourteen-to-sixteen month oscillation; pole tide with changing period
GCM	Global Change Model
GEONZ97	geoid for the North Sea area 1997
GOCE	gravity field and steady-state ocean circulation explorer
GPS	global positioning system
IERS	international Earth rotation service
IGS	international GPS geodynamic service
ITRF	international terrestrial reference frame
JGM	joint gravity model
JILA	Joint Institute of Laboratory Astrophysics
MWG	modified Meissl/Wong&Gore truncation coefficients
NAP	normaal Amsterdams peil
NEREF	Netherlands reference frame
NN	Normal Null
OSU91A	Ohio State University gravity model 91A
PSMSL	permanent service for mean sea level
RLR	revised local reference
RMS	root mean square
SELF	sea level fluctuations: geophysical interpretation and environmental impact
SLR	satellite laser ranging
SPREP	South Pacific Regional Environment Programme
SSA	singular spectrum analysis
SVD	singular value decomposition
TGBM	tide gauge bench mark
TOPEX	ocean topography experiment
UELN	united European levelling network
VLBI	very long baseline interferometry
WGS-84	world geodetic system 1984
σ	standard deviation

Chapter 1

Introduction

1.1 Sea-level variations

In June 1999, the following news (translated from Dutch) was published on the information service of the Dutch television network (teletext):

> **London** Two islands in the Pacific have been swamped by rising sea levels as a result of the greenhouse effect. This has been reported by SPREP on an international climate change meeting in Bonn. The islands involved are Tebua Tarawa and Abunuea, both situated in Kiribati.

The above example is only one of the many "facts" that have been published in the last few years by our national news agency, but also by international news agencies on the subject of greenhouse-gas induced warming. Frequently, the news facts mentioned are not very accurate. For example, it is stated that the greenhouse effect is very dangerous to our society. However, this is a "natural" phenomenon without which the temperature on Earth would be thirty degrees lower (de Ronde, 1991). The real problem is *additional* greenhouse-gas induced warming of the Earth. Often, only a small part of an extensive report is used, for example to emphasis the writers opinions concerning the "dangers" of greenhouse-gas induced warming. As an example, the news quoted above gives the impression of a major disaster. However, the two islands mentioned are only very small islets (so-called motu) and, as explained by Pirazzoli (1990), all oceanic islands experience subsidence due to lithospheric cooling. Another example of known facts adapted to confirm one's opinion was taken from the web-page of the American Petroleum Institute:

> There is no credible evidence that sea level is rising worldwide as a result of human activities. However, changes do occur frequently from decade to decade, or by region. For example, waters around the Mississippi Delta have been slowly rising, but parts of Scandinavia have experienced a decline - not a rise - in sea level.

Although the interest in sea-level height variations (and especially in sea-level rise) seems to be a phenomenon of only the last few decades at most, this is certainly not true. Already as early as 1682 (in the Netherlands) and 1704 in Scandinavia a start has been made with recording variations in sea-level height, on a regularly basis (Mörner, 1979). The Permanent Service for Mean Sea Level (PSMSL) has monthly and annual mean sea-level values available from 1807 on (tide gauge Brest). The early tide gauges were often situated in harbour areas and sea-level height measurements were usually obtained in order to predict tidal movements.

From mean sea levels as recorded by the PSMSL, sea-level variation curves (like secular sea-level rise) can be determined over periods of up to around 150 years. An indication of sea-level variations on much longer time scales can be obtained from a wide variety of sources. As an example, in the Mediterranean area man-made structures (e.g., harbour constructions) built by the Roman Empire have been found, which are situated well below present-day sea level. The more precise it can be recovered when these constructions were built, and the better it can be estimated at what height above sea level they were constructed, the more accurate estimates of the secular sea-level rise over the elapsed period of time will be. Over even longer time spans, secular sea-level height variations can, for example, be estimated based on radiocarbon dating of remains of marine plants that lived in sediments of the inter-tidal belt; see Lambeck *et al.* (1998) for more details on the (limitations of) this method.

Examination of the different types of sea-level height data shows that sea-level height variation does not follow a simple linear secular rise or fall. Sea-level heights are influenced by a large number of periodic (or semi-periodic) processes operating on a wide range of time scales. As an example, figure 1.1 shows changes in eustatic sea level over the past 7000 years.

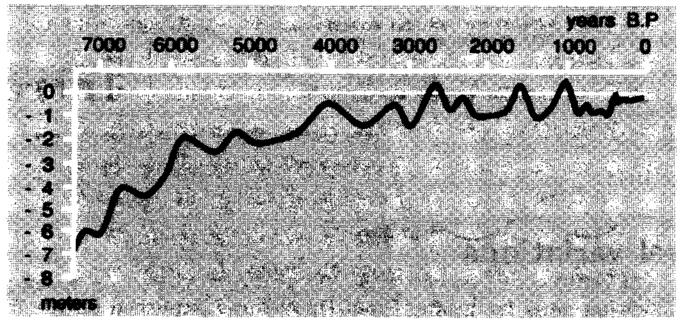

Fig. 1.1 Variations in eustatic sea-level height; reproduced from de Ronde (1991)

One of the main sources of periodic sea-level variations on geological time scales are the glacial cycles. According to Emery and Aubrey (1991), these cycles have time scales of 100 000 years with superimposed shorter periodic fluctuations. It is estimated that during the last interglacial period (125 000 years ago), sea level was two to six metres above the present level. During the last ice age, sea level lowered up to 120 metres below present-day levels. In addition to these large climatological cycles, smaller variations have occurred as well. Examples are the Medieval Warm Period (900 - 1250 AD) and the Little Ice Age (1300 - 1800 AD); see Varekamp and Thomas (1998). Figure 1.2 shows oscillations in mean winter temperatures that have occurred over the period between 800 and 1990 AD in the Netherlands. It can be concluded that substantial fluctuations in climate have already occurred, long before human society was able to influence its environment on a large scale.

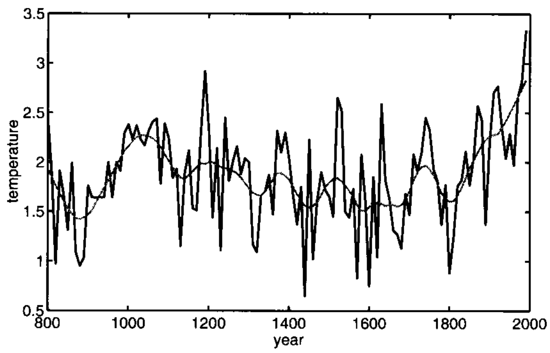

Fig. 1.2 Ten-yearly mean winter temperatures (in °C) as observed in the Netherlands (black), and moving averages determined over 150 years (grey). Reproduced from Können (1999).

Some problems with detecting sea-level variations

The second quotation given in the preceding, illustrates one of the major problems of the methods used to collect sea-level heights and variations in sea levels. Sea-level heights are determined relative to a specific location or area of land. For example, tide gauges monitor variations in sea-level height relative to the zero-point of the tide gauge. This zero-point is (usually) fixed to a specific location (the tide gauge bench mark) on the surface of the Earth. As another example, if the level of submergence beneath present-day sea level of historic buildings is used, this will not only give an indication of how much sea-level heights have increased, but also of the vertical movement of the buildings itself.

Most methods for determining sea-level variations will result in measurements of relative sea-level heights. The measurements contain the combined effect of variations in absolute sea-level heights and local vertical movements of the reference point (or area). Nowadays, methods like satellite altimetry are able to provide, more or less, variations in absolute sea levels, although small corrections for glacial isostatic adjustment are still required; see Peltier (1998). However, altimetry measurements are still obtained relative to the orbit of the altimeter satellite and are, therefore, dependent on coordinates determined for terrestrial reference stations. In addition, measurements are only available for the last few decades, and are (often) subject to drift of the altimeter.

The height of the coastline (used as reference for sea-level measurements) is influenced by a large number of processes; see, e.g., Emery and Aubrey (1991). Some of these processes, like local tectonic movements, will result in very localised effects. Other processes, for example post-glacial rebound, will influence larger areas. The processes influencing the height of the coastlines operate on a wide variety of time scales, varying from almost abrupt height changes (e.g., due to earthquakes), to changes on geological time scales (e.g., lithospheric cooling). Due to the fact that different locations experience different vertical movements, there is a wide range of values estimated for the variations in relative sea-level height obtained at different locations.

Relative sea-level height measurements contain both variations in absolute sea-level heights and vertical movements of the reference points (e.g., tide gauge bench marks) to which the sea-level measurements are related. Therefore, in order to estimate variations in global absolute sea-level height, the (local) height movements of the reference points should be removed from the measurements. This is not easy to achieve, since the heights of the reference points are influenced by a large number of processes. However, some researchers claim that "stable" areas can be found for which the land movements can be neglected. As an alternative, based on post-glacial rebound models (and depending on their quality), at least part of the vertical land movements could be removed from the relative sea-level height measurements.

It should be remarked that variations in absolute sea level are mainly of interest for scientific purposes. For many practical considerations variations in relative sea-level height will be more important. For example, for people living along the Dutch coast, it might be a relief that *absolute* global sea level increases with only 1 mm/yr. However, if due to local circumstances (e.g., significant subsidence) *relative* sea levels increase with 3 mm/yr, our present coastal defences (dunes and dikes) might no longer be sufficient in already a shorter period of time.

A second problem with detecting sea-level variation patterns is the time span of the measurements. If the observation period is not sufficient in relation to the longest (significant) fluctuations in the data set, this might lead to anomalous predictions. As a simple example, it is assumed that linear regression is applied to tide gauge data that contains a significant long-periodic signal (period is 60 years). As can be seen from figure 1.3, if a trend estimate is based on only the last 30 years of data, the derived value is much larger than the actual sea-level rise slope in the data.

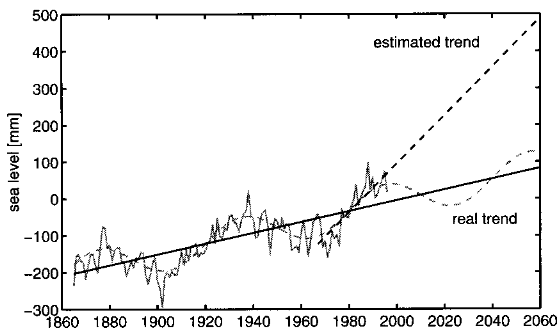

Fig. 1.3 Annual mean sea levels (grey), long-periodic fluctuations (dashed grey), estimated trend based on full 132 years of data (black), and estimated trend based on only 30 years of data (dashed black).

Figure 1.3 illustrates that the observation period should be sufficiently long in relation to the longest significant period present in the sea-level height data. But, as explained in the preceding, and can be seen from figure 1.1, sea-level heights contain periodic fluctuations with extremely long periods. Therefore, it will be impossible to define an observation period that includes at least one full cycle of the longest-possible periodic fluctuations. However, this is not necessary. Of importance is that the longest *significant* periodic fluctuations are included.

That only fluctuations that are significant need to be considered implies two different things. First of all, only if the amplitude of the resulting sea-level signal is significant in relation to all other signals influencing sea-level heights, should a process be taken into account. For example, Pitman (1979) estimates that plate movements have a maximum effect of 0.01 mm/yr on sea-level heights and this process can, therefore, be neglected. In addition, the periodicity of a process has only to be considered if its frequency is significant in relation to the period of time that is of interest to us. As a simple example, sea-level height fluctuations caused by the glacial cycle are very large. However, for our present-day problems with rising sea levels (and those of the next few generations) it is no relief that sea levels will fall again when the next Ice Age commences.

When estimating how many years of observations are required for a reliable estimate of sea-level variation curves, it has to be determined what kind of periodic phenomena have to be considered and which fluctuations have to be incorporated into the variation curve that is estimated. For the example of sea-level fluctuations caused by the glacial cycle, it will not be very useful to (try to) remove the resulting (small) trend from the sea-level height data. For periodic events on shorter time scales, it depends on the purpose of the sea-level research whether or not these fluctuations should be eliminated from the estimated sea-level variation curve. For example, if the overall trend in sea-level rise is required, sea-level height series should be used that contain at least one full period of the major periodic phenomena (of the order of 90 years of data are required; see chapter 4). On the other hand if the safety margins of dikes for the year 2020 have to be predicted, of importance is not the rise in mean sea level but the increase in magnitude of major storm surges.

Sea-level height time series as collected by individual tide gauges are influenced by a large number of processes. These processes cause sea-level height variations on a wide range of temporal and spatial scales. Some processes (like thermal expansion) will influence sea levels on a more or less global scale, while other processes (e.g., local atmospheric pressure variations) have a very localised effect. Influences on instantaneous sea-level heights may have a periodic nature (like tides and fluctuations in salinity), may be more or less secular (e.g., effect of post-glacial rebound), or can cause more or less random effects (high frequency signals like waves). On the subject of the various processes influencing sea-level heights, a wide range of literature has been published; see, for example, reviews by Emery and Aubrey (1991), Hamon and Godfrey (1980), and Fairbridge and Jelgersma (1990).

This thesis focuses on secular variations in sea-level heights. High-frequency oscillations will, in general, not hamper the detectability of, e.g., secular trends in sea-level rise. Very long-periodic fluctuations (with periods of at least a few decades), however, can have a significant influence on how well sea-level variation curves can be estimated. Unfortunately, literature on long-periodic fluctuations occurring in sea-level height time series is scarce.

Justification for sea-level research

According to Rietveld (1986), of the order of 20% of the surface of the Netherlands lies below mean sea level. Approximately 60% of the Netherlands is situated below high tide level (Pöttgens, 1991). These lands are protected by dunes and dikes. As explained by de Ronde (1991), the safety of the low-lying areas is especially threatened by storm surges and high river discharges. It is estimated (Xu, 1990) that if sea level would rise by 1.5 metres, the recurrence period of an extreme storm surge as has occurred in 1953, would decrease from 300 years to only 3 years.

Not only the Netherlands would be severely threatened by rising sea levels. A significant part of the world's population lives in low-lying areas. In addition, especially in developing countries, large amounts of people depend on the river deltas for their food production. In particular coral islands (the Maldives, Kiribati, Marshall Islands, etc.) are threatened by rising sea levels, since they often extend only several metres above sea level.

To mention a few figures, according to Zwick (1997), the Nile data (Egypt) represents only 2.3% of the area of the country, but 46% of the total cultivated surface is situated in this delta, and 50% of the population lives in this area. Gommes and du Guerney (1998) mention that most of the current largest urban concentrations are near the sea coasts. They also estimate that under a worst case sea-level rise scenario (global mean sea-level rise of 95 cm by the year 2100), between 0.3% (for Venezuela) and 100% (Kiribati and Marshall Islands) of the population would be affected by the rising water. The majority of the people that would be affected by this magnitude of sea-level increase live in China (72 million) and

Bangladesh (71 million). According to Fisk (1997), without measures adopted specifically to tackle rising sea levels, increased flooding will affect some 200 million people worldwide by the year 2080. Around 25% of the worlds coastal wetlands could be lost by this time due to sea-level rise alone.

A wide range of areas in the world is threatened by rising sea levels. Some of these areas (like the major part of the Netherlands) could in principle be protected by building (higher) dikes. For other areas, like coral islands, this is not really an option. Apart from the tremendous costs involved, a sea-wall around an atoll would kill the inner lagoon ecosystem. This would endanger the fishery, on which a large part of the population is dependent. In addition, they would lose the income from the tourist industry if the beaches disappear.

In order to take timely precautions, whether or not defences can be built or people have to be evacuated, it is important to know what magnitudes of relative sea-level rise are to be expected on which time scales. The general sea-level variation curve underlying sea-level height movements is of importance. Over longer time spans, higher order sea-level rise (e.g., quadratic or exponential increase) will have much more impact than linear rise. It is also of interest to determine whether or not observed sea-level rise accelerations are part of the "natural" pattern or caused by, e.g., greenhouse-gas induced warming. In the latter case, some protective measures might still be taken in order to reduce this greenhouse-gas induced warming in the near future. In addition, industrialised countries might be more willing to assist developing countries with coping with sea-level rise problems if these problems are caused by man-made climate changes instead of having "natural" origins.

It is difficult to distinguish between greenhouse-gas induced sea-level rise accelerations and "natural" variations in sea level. As explained in the preceding, long before human society was able to influence its environment on a large scale, significant variations in climate have occurred. Many authors have discussed that sea levels over the last 100 years have increased at a much higher rate than that which has occurred over the last few thousands of years. According to Warrick *et al.* (1996), sea-level rise over the last 2000 years has been of the order of a few tens of centimetres, while sea-level rise over the last 100 years is of the order of 1.0 to 2.5 mm/yr. Based on high waters at Liverpool since 1768, Woodworth (1999) concludes that the apparent high rates of sea-level rise observed in the twentieth century, are a result of an acceleration around the second half of the nineteenth century. On the other hand, based on dated salt mash peat sequences, Varekamp and Thomas (1998) claim that sea-level rise increase started already in the 17th century, with a major acceleration around 1800 AD (which again corresponds to the end of the Little Ice Age). In addition, substantially higher global average surface temperatures (relative to the mean value obtained between 1961 and 1990) have been determined for the period around 1940 (Nicholls *et al.*, 1996).

1.2 Objectives of this thesis

The study and interpretation of all aspects of sea-level height variations requires an interdisciplinary effort. Therefore, this thesis is necessarily limited to only a particular area of the problems involved. Central to this thesis is *the quality of the geodetic and tide gauge measurements.* The aim of the work performed can be described as follows.

> The objective of this thesis is to determine how well patterns in sea-level height variations can be detected, considering the fact that sea-level height time series used are affected by inconsistencies.
>
> Inconsistencies between time series are introduced by inaccuracies in the measurements involved. This can be measurement errors of the tide gauge equipment itself, or inaccuracies in the geodetic measurements that are used to connect the tide gauges in heights.

Methodology

Due to time considerations, information about the quality of the measurements (tide gauges, spirit levelling, GPS, and gravimetry) has been obtained from literature. No experiments have been performed to confirm the described values for the quality of these measurements. Only on the subject of connecting vertical datums, error propagation studies have been performed to obtain an indication of the quality with which datum connection parameters can be determined.

Conclusions derived concerning the detectability of sea-level variation patterns are all based on experiments with simulated data. The artificial sea-level height series used contain a specific (simulated) sea-level variation curve (e.g., a linear regression line) in combination with periodic fluctuations and simulated measuring errors. The periodic fluctuations are based on actual tide gauge data as provided by the PSMSL. The reason for using artificial sea-level height time series is that this allows the construction of data sets containing only specific phenomena, e.g., a time series with a specific trend and tide gauge measuring errors, but no inconsistencies introduced by connecting the tide gauge in height to an external reference system. For the simulated data sets the underlying sea-level variation curve is precisely known. Consequently, it can be examined how well this curve can still be detected if the time series contains all kind of measuring errors (e.g., introduced by the tide gauge itself, and the geodetic measurements used to connect tide gauges in height).

Further context and limitations

In this thesis, sea-level height variations are considered for a small area, the North Sea region. Of the wide variety of aspects of sea-level variations and measurements that influence the detectability of these variations, only the following components will be examined. Simulated data sets are used in which the sea-level heights follow simple curves that can be described by (a combination of) regression lines. The detectability of these sea-level variation curves is examined based on the statistical significance of the estimated regression parameters. In addition, it is examined how robust trend estimates are against a change in the number of observations on which these trend estimates are based. Furthermore, it is checked how much trend estimates change if different realisations of inconsistencies between the time series are applied. The following type of inconsistencies will be considered:

- Measuring errors of the tide gauge equipment.
- Inaccuracies in the geodetic measurements used to connect the tide gauges to a local reference frame.
- Vertical movements of local reference frames relative to one another.
- Inaccuracies in the height connections between the local reference frames.

As indicated in the preceding paragraphs, sea-level heights are usually measured relative to a specific reference on the Earth's surface. If a time series of sea-level heights is used to determine a sea-level variation curve, this curve will also be related to this specific height reference.

Tide gauges relate sea-level heights to the local tide gauge bench mark at specific time intervals. From the resulting sea-level height time series, a sea-level variation curve can be determined. If this curve can be described by, e.g., a (combination of) regression lines, based on the values estimated for the regression coefficients, future sea-level heights can be predicted. In this case, all sea-level values (sea-level variation curve and forecasted values) are given relative to the local tide gauge bench mark.

Sea level as measured by an individual tide gauge is influenced by a large number of local processes, e.g., by variations in sea-water density for a tide gauge located near the mouth of a large river. As a result, differences will occur between sea-level height data obtained by tide gauges situated along the same coastline. For this group of tide gauges, a common sea-level variation curve can be estimated that represents sea-level variation relative to this coastline. Since the individual sea-level height series are related to the different tide gauge bench marks, some kind of common reference surface has to be used as reference for the estimated common sea-level variation curve. This is shown in figure 1.4.

All sea-level height time series can be related to the same reference surface by determining the height differences between the local tide gauge bench marks and this reference. This height connection to a common (local) reference surface is in particular required if the tide gauge bench marks experience height changes relative to one another (e.g., due to local differences in subsidence rates). These relative height changes of the tide gauge bench marks introduce apparent differences in rates of sea-level rise between the sea-level height time series. By relating all sea-level height series to the same local reference surface, these apparent sea-level height variations are removed from the individual data sets.

If for a larger group of tide gauges, situated around the North Sea, measurements are available, a spatial pattern in sea-level variation curves can be determined. As a (hypothetical) example, it might be observed that rates of sea-level rise increases with latitude. Analogous to the preceding, all sea-level height time series have to be related to the same reference surface. Otherwise, if tide gauge bench marks

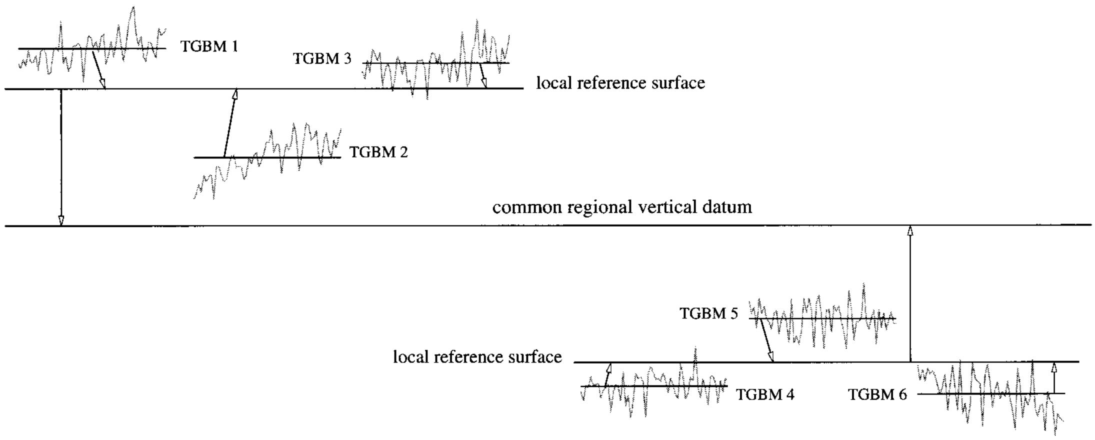

Fig. 1.4 Different reference surfaces to which sea-level height series can be related. TGBM indicates the local tide gauge bench marks. Arrows represent the height connections between the reference surfaces.

move relative to one another, anomalous differences are introduced into the individual sea-level variation curves.

Often, the height of the tide gauge bench marks have been determined relative to a local height system. In different regions, different reference surfaces have been used. Consequently, sea-level height series obtained along the various coastlines surrounding the North Sea, will have been connected to different local height systems. If these local reference surfaces experience vertical movements relative to one another (e.g., due to differences in uplift or subsidence rates), these movements are introduced as apparent sea-level height changes into the time series.

To determine a spatial sea-level variation pattern, the height differences between the local height systems have to be monitored; see figure 1.4. As a result, the spatial sea-level variation pattern is determined relative to a common reference height. As reference height simply one of the local height systems as used in the region can be selected.

The procedure described above is just one method for handling sea-level height data. Many different approaches could be envisioned. For example, nowadays it is often preferred to connect tide gauge bench marks into a global (geometric) reference system (e.g., ITRF92) by means of relative GPS measurements.

In principle, sea-level height time series can be related to any possible reference height. In this thesis it has been chosen to use sea-level heights relative to the local tide gauge bench marks if results for individual time series are examined. It is also possible to transfer these time series to sea-level variations relative to a local height system, or even relative to a global system like ITRF92. However, it has to be taken into account that all operations on sea-level height series in order to relate them to different reference surfaces, introduce uncertainties into the data sets. Whether the transformation is based on height measurements, gravity measurements, available transformation parameters, etc., these will always contain inaccuracies. Therefore, sea-level height time series relative to the local tide gauge bench mark (or even better, relative to the tide gauge zero), will be more accurate than those given relative to, e.g., ITRF92.

1.3 Outline

In order to realize the objectives of this thesis, literature has been studied and a wide range of simulation experiments have been performed. A description of the work performed and the results derived is divided over the subsequent chapters as follows. First, in chapter 2 the quality of the major types of tide gauge systems will be addressed. Inaccuracies in the tide gauge measurements will introduce inconsistencies between sea-level height time series. During the history of tide gauge operations, different types of tide gauges have been used with different error characteristics. Since not only the quality of the instantaneous sea-level height measurements is important, but the number of daily measurements as well, this chapter also discusses a number of methods for averaging instantaneous sea-level heights into annual mean values.

If tide gauges experience height changes relative to one another, and especially if the relative height movements are not constant over time, this hampers the detection of a common sea-level variation curve by means of, e.g., SVD. Furthermore, if spatial patterns in sea-level height variation are based on these data sets, the spatial variation pattern will not only represent variations in absolute sea-level rise, but vertical movements of the tide gauges relative to one another as well. These problems can be (partially) prevented by referring all sea-level heights to the same reference surface.

Inaccuracies in the measurements required to connect the tide gauge bench marks in height introduce inconsistencies between the time series. In chapter 5, geodetic measuring techniques will be described that can be used (and have been used in the past) to connect tide gauge bench marks to a local reference frame. In chapter 7, an indirect method for deriving height differences between different local height datums is described. Based on error propagation studies, a quality assessment is made for datum connection parameters obtained using the proposed method.

The accuracy of derived sea-level variation patterns not only depends on the quality of the data, but also on the data analysing techniques used to determine these patterns. In chapter 3, a selection of data analysing techniques will be discussed, which can be applied to monitor sea-level height variations or predict future values.

In chapters 4, 6, and 8, based on simulated data sets, the detectability of sea-level variation patterns will be examined. For an *individual* tide gauge, variations in sea level relative to the tide gauge bench mark will be considered in chapter 4. Of interest is how well different variation curves can be distinguished and whether or not it is possible to determine the onset year of sea-level rise accelerations. A *common* sea-level variation curve for a group of tide gauges will be determined in chapter 6. For this purpose, all sea-level height series used are connected to a local reference frame. It will be examined how much the detectability of a common variation curve for this group of tide gauges is influenced by the inaccuracies in the height measurements between the tide gauge bench marks and the local reference frame. *Spatial variations patterns* for a group of tide gauges situated in the North Sea area are examined in chapter 8. The individual time series used in this chapter are not only influenced by height connections between tide gauge bench marks and the local reference frames, but by height changes between the different vertical datums as well.

The major conclusions of this thesis will be summarised in chapter 9. In addition, based on knowledge gained with the work performed, recommendations for future research and new to be established sea-level monitoring systems will be given.

Chapter 2

Tide gauge measurements

2.1 Introduction

Sea level is conventionally monitored using tide gauges, which relate variations in sea-level height to a local tide gauge bench mark. Over the years a number of different tide gauge systems have been developed, with varying measuring precision and recording principles. In section 2.2 error characteristics will be discussed for the six major tide gauge systems that have been used in the past or are still in use today, i.e., tide poles, tide poles with float, stilling-well tide gauges, acoustic reflection tide gauges, pressure tide gauges and open-sea pressure tide gauges.

The quality of sea-level height time series and, consequently, the quality of estimated patterns in sea-level variation, not only depends on the measuring accuracy of the tide gauge systems, but on the measuring frequency as well. In addition, since usually only some kind of mean values (hourly, daily, monthly, or even yearly values) are available, the method used to form these mean values also influences the quality of sea-level height time series. These effects of sampling rate and averaging method will be discussed in section 2.3.

Finally, since tide gauges measure sea-level height relative to a local tide gauge bench mark, height differences between the tide gauge bench marks have to be determined, in order to relate time series for different tide gauges. Connection of tide gauge bench marks to a local reference frame will be discussed in chapter 5.

2.2 Error characteristics of tide gauge instruments

Local sea-level height variations have been measured for thousands of years using a variety of measuring techniques. At first, sea-level measuring systems often consisted of a vertically mounted pole, relative to which instantaneous sea-level height could be determined by an observer. Since reading sea-level height relative to a pole is difficult in the presence of waves, the introduction of a float that moves vertically in a well significantly improved the accuracy of the measurements. This system was further improved by automating the recording of the sea-level data, e.g., by means of a pen driven by the float across a chart, which in turn is mounted on a circular drum. According to Pugh (1987), self-recording gauges began operating in the beginning of the nineteenth century. Mechanical recording has a number of disadvantages as well, therefore, digital recording leads to a further improvement in the accuracy of the measurements.

Besides improvements on the "traditional" stilling-well with float system, a number of other sea-level measuring devices, based on other measuring principles have been developed during the last century. Examples are tide gauges that measure pressure at a fixed point below the sea surface and tide gauges that measure the acoustic reflection time between a fixed point and the instantaneous sea surface.

For planning of required measurement durations, for sites where new tide gauges have to be installed, only the accuracy of state-of-the-art tide gauge systems is important. However, error characteristics of older sea-level measuring systems are relevant as well since often older sea-level records have to be included in order to estimate a specific phenomenon in sea-level height variations. For some sites, tide gauge data have been recorded for more than a century, and, to determine meaningful results from these data, variations in data quality over the length of the time series (e.g., due to changes in applied measuring techniques) have to be taken into account. Error characteristics and limitations for the six major techniques, i.e., tide poles, tide poles with float, stilling-well tide gauges, pressure tide gauges, acoustic reflection tide gauges, and open-sea pressure tide gauges, will be described in the following sections.

2.2.1 Tide poles

One of the first systems for monitoring sea-level variations consisted of a so-called tide pole, vertically mounted on a site assumed representative for the area of interest. At regular intervals the height of the instantaneous sea level relative to this pole can be read by an observer. The gauge zero should be connected to a permanent bench mark on shore, the so-called tide gauge bench mark.

The precision of sea-level measurements from tide pole readings is determined by random reading errors that, according to Montag (1970), depend on the state of the sea surface, the quality of the markers, the distance between observer and tide pole, light conditions, and the constitution of the observer. From experiments carried out under different weather conditions, Montag (1970) concluded that the average error due to these reading errors amounts to less than 2 cm. Pugh (1987) is slightly less optimistic, he states a reading precision of 2 cm but only for calm weather conditions; in the presence of waves this precision deteriorates. He claims, based on experiments, that in the presence of waves with a height of 1.5 metres, experienced observers are able to read the tide pole with a precision of 5 cm.

Apart from random reading errors, the accuracy of tide pole measurements is also determined by the presence of systematic errors and blunders. Blunders can originate from gross errors either in the readings of the sea-level heights or in the registration of these heights. Systematic errors can result

1. from the construction of the pole itself
   e.g., deviations in the scale of the pole, or problems with the connection between the individual one-metre sections of the pole,
2. from environmental conditions
   e.g., due to different illumination during day and night, or
3. they are caused by the operator
   e.g., operator has a tendency to read slightly too high values.

Furthermore, sea level in the presence of waves is often determined as the average between the crest and trough of the waves. As a result, waves give rise to additional systematic errors, due to two different mechanisms. First of all, as explained by Pugh (1987), for an observer viewing the pole at an angle, the apparent trough may actually be the crest of an intermediate wave. This wave is obscuring the true trough at the pole, resulting in an averaged level that is too high. In addition, since water waves differ from a true sine oscillation, averaging between crest and trough values will give a systematic error. The value of this error depends on the state of the sea surface, but on average will be of the order of 1 mm; see Montag (1970).

Although tide poles have a large number of limitations they have some advantages as well. For example, due to the small amount of technology involved they are relative cheap and easy to install and operate. Consequently, they provide an easy method to check readings of other tide gauge systems, e.g, to check the datum imposed for pressure tide gauge measurements (see section 2.2.5).

2.2.2 Tide pole with float

As indicated by Pugh (1987), the problem of reading the tide pole (especially in the presence of waves) can be minimised by fitting a transparent tube alongside the tide pole. This tube is connected to the sea by means of a narrow inlet tube, which prevents waves from entering the tube connected to the tide pole, see figure 2.1. As a result, the height of the sea level inside the tube can easily be read relative to the tide pole.

How well waves are attenuated depends on the so-called time constant of the specific tide pole system. According to Smithson (1997), this time constant is determined by the characteristics of the inlet tube and stilling tube and is a function of the gravitational acceleration, the kinematic viscosity of sea water, the diameter and length of the inlet tube, and the internal diameter of the stilling tube.

A tide pole with a float in a stilling tube has a lot of error characteristics and limitations in common with stilling-well tide gauges. Since stilling-well tide gauges are one of the more popular tide gauges in use

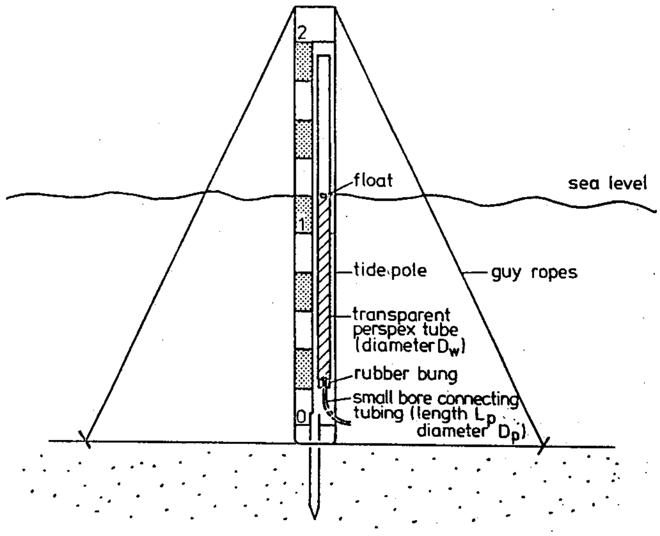

Fig. 2.1 Tide pole with float in stilling tube; reproduced from Smithson (1997)

nowadays, these characteristics will be described in detail in the next section. In this section, problems related to using a float and a stilling tube will only be discussed briefly.

The accuracy of tide pole measurements fitted with a float in a stilling tube is determined by mechanical problems and the deviation between the water level in the tube and the open sea level outside. Mechanical problems are, e.g., friction of the float within the stilling tube and deviations in the construction of the pole itself as discussed in section 2.2.1.

Deviations between the water level inside and outside the stilling tube depend, e.g., on the state of the sea surface in connection with the size of the inlet and stilling tube. Waves, which cause reading inaccuracies if "normal" tide poles are used, are damped out by the filtering properties of the inlet and stilling-tube construction. How well these waves are attenuated depends on the size of these tubes. Unfortunately, the filtering characteristics of the inlet and stilling-tube construction might change over time as a result of environmental conditions (like silting up by sediments and marine growth).

Other factors contributing to a deviation between levels inside and outside the stilling tube are differences in salinity and temperature between the "open sea" water and the water inside the tube and pressure deviations due to the inlet tube being placed in a tidal stream. An additional problem with salinity and temperature differences in the stilling tube is that their effects are extremely difficult to estimate if dye is mixed into the water in order to increase the visibility of the level of water in the tube and/or some substances are added to the water to prevent freezing.

2.2.3 Stilling-well tide gauges

The two previous tide gauge systems have the major disadvantage that tide gauge readings have to be performed by a human observer. Apart from the fact that this facilitates the occurrence of blunders this makes them unsuitable for long term measurement campaigns with a high measuring frequency. A higher measuring frequency allows elimination of, e.g., high frequency disturbances, and results in more precise (e.g., hourly) mean values.

Since the mid-nineteenth century self-recording gauges have been in operation. They often consist of a so-called stilling well (a tube or other enclosed area, connected to the open sea by an orifice or inlet pipe) with a float. This float is connected to the recording system, which in the past was a mechanical device, whereas newer tide gauges use digital recording. Figure 2.2 shows an example of a stilling-well tide gauge with orifice and mechanical recording on a drum.

In order to prevent aliasing and to allow for the use of a float, high frequency movements of the water

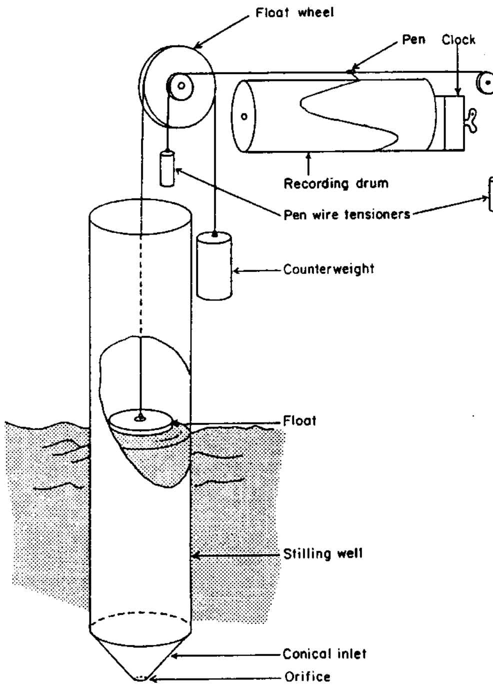

Fig. 2.2 Stilling-well tide gauge with mechanical recording on a drum; reproduced from IOC (1985)

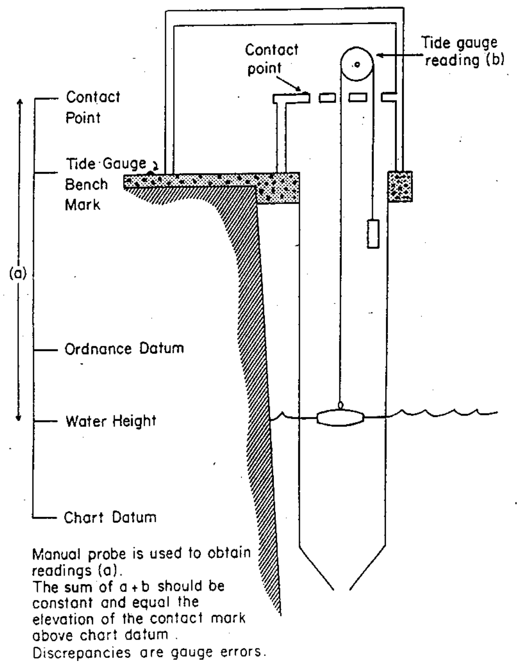

Fig. 2.3 Datums involved with stilling-well tide gauge; reproduced from Smithson (1997)

surface have to be filtered out mechanically. The stilling well provides a protection of the float system against environmental conditions (e.g., from the effect of wind), whereas the relatively small diameter of the inlet provides a mechanical filter of high frequency movements like waves. Damping characteristics of the stilling well are, mainly, determined by the ratio between the diameter of the inlet (orifice) and the diameter of the stilling well itself.

Damping effect of the stilling-well system is determined by the flow coefficient through the inlet. This flow coefficient has a value between 0 (inlet is completely closed) and 1 (ideal flow conditions); for more details see Sager and Matthäus (1970). They also give detailed relations between water level response in the stilling well and constant, linear, and periodic variations respectively, of the "open" sea water level.

The accuracy of sea-level variations based on stilling-well tide gauges is determined by the stilling well itself (deviations between the water level inside the stilling well and the "open" sea), by mechanical problems of the float system, and by problems with the recording of data. If digital recording is used to record samples of sea-level height at discrete periods in time, the sampling period has to be adapted to the highest significant frequency in order to prevent aliasing.

Mechanical recording has the advantage that it allows for a continuous monitoring of sea-level height variations, and thus the problem of aliasing is averted. However, mechanical recording of sea-level heights yields inaccuracies in recorded sea level heights as well. Main problems, as indicated by e.g., Lennon (1970), Van der Made (1987), and Xu (1990), are problems with the time annotation (e.g., due to a deviation in the rotational speed of the drum, or even worse, due to variations in this rotation), deformation of the drum, and deformation of the paper (e.g., due to variations in humidity). In addition, as indicated by Pugh (1987), reading of the produced charts is a tedious procedure that is prone to errors. He estimates that due to the width of the chart trace, precision of sea-level height charts is of the order of 2 cm for levels and 2 minutes in time. According to Woodworth (2000), an advantage of charts was that, even if the record was noisy, e.g., as a results of seiches or a passing ship, the experienced person doing the analysis of the data would be able to pick out the smooth tide line through the noise.

Mechanical problems with the float system include movements of the float and wire, friction and backslash in float and counterpoise suspension, variable tension in the suspension wire of the float, and variations in the length of the float suspension (again based on Lennon (1970), Van der Made (1987), and Xu (1990)). A number of these mechanical problems can be checked against by the so-called "van de Casteele test". In this test, the distance between the contact point and water level is measured manually (e.g., with a steel tape) and compared to the recorded sea-level height, see figure 2.3. Plots of the difference between these two distances through time (the so-called "van de Casteele diagrams") give an indication of probable mechanical errors in the tide gauge system. For an "ideally" operating tide gauge the diagram shows a straight line, while different deviations from this straight line correspond to different types of mechanical problems. For more details the interested reader is referred to, e.g., Smithson (1997).

The stilling-well structure itself also limits the accuracy of sea-level heights. As indicated in the preceding section, mechanical filtering by means of a relative small inlet not only attenuates high frequency signals it also causes a time delay between sea-level variations in the "open" sea and resulting water level variations in the stilling well. In addition, filtering characteristics change if the diameter of the inlet changes, e.g., as a result of siltation, marine growth and accumulation of weed or trash. Tide gauges situated in some sites will be more prone to clogging up than tide gauges situated in other sites. Besides, as investigated by Cross (1968), the flow through an orifice inlet is asymmetrical, resulting in a Bernoulli effect of the water level in the well if significant surging outside the well occurs. Cross (1968) estimates that this can cause an error of 14 cm if waves are present with heights between 1.5 and 3 metres.

Deviations in density (in temperature and salinity) between the sea water and the water inside the stilling well causes different water levels inside and outside the well. As described by Van der Made (1987) if salinity of the "open" sea-water changes (e.g., due to variations in runoff of a nearby river), the effect on the density of the water outside the stilling well will be much larger than the effect on the water inside the well. As an example, Van der Made (1987) shows that for a stilling well with an inlet tube 4 metres below the water level and a density of 1010 kg/m^3 inside the well, a density increase outside the well of 20 kg/m^3, would yield a deviation between the level inside and outside the well of 4 cm.

Especially stilling-well tide gauges at estuary sites can show large deviations between the water level inside and outside the stilling well due to variations in salinity and temperature through the tidal cycle. As explained by, e.g., Pugh (1987), with rising tide the density of estuary water increases. Consequently, at high tide the density of the "open" sea water is higher than inside the stilling well, where the density is an average of the "open" sea density during the filling up of the stilling well. As a result, the water level in the stilling well can be significantly higher than outside the stilling well. In extreme cases, like tide gauges in the river Mersey (U.K.) where the tidal range is of the order of 10 metres, Lennon (1970) estimates that at spring high water the level inside the well can be 6 cm higher than outside the well. Pugh (1987) even gives a difference in water level of 12 cm for a tidal range of 10 metres.

Deviations between the water level inside the stilling well and the "open" sea level can also be caused by differences in pressure. If the stilling well is situated in a tidal stream, the stilling-well structure itself causes pressure disturbances by (partly) blocking the flow. In addition, other obstacles in the vicinity of the stilling well can cause pressure differences. The result is a draw-down of the water level, which gives (analogous to differences in density) systematic errors in recorded sea-level heights.

According to Diamante *et al.* (1987), the standard deviation of a single measurement of sea-level height made by a stilling-well tide gauge will not be better than about 5 cm. However, according to Woodworth (2000), this estimate had a political aganda to get float gauges replaced by next generation acoustic systems in the USA.

The precision of sea-level heights can be improved by forming hourly mean values, since this reduces the influence of high frequency (nearly) random errors as introduced by waves. Hamon and Godfrey (1980) estimate that an adequately maintained tide gauge situated in a suitable location can yield hourly mean values with a precision of the order of 1 cm, whereas daily mean values will have a standard deviation of the order of 1 mm. A suitable location is (probably) a site that is the least possible influenced by local processes, so the tide gauge is, e.g., not installed near a (large) river mouth.

In section 2.3 different methods will be discussed that can be used to form longer-term average values like monthly and yearly means. The precision of these mean values will be significantly better than that of the single samples of sea-level heights because the effect of most random influences (like the effect of waves) will be strongly reduced. However, a number of systematic errors (like draw-down by pressure differences and relative high water levels in the stilling well during high tides due to differences in density) also occur in sea-level measurements. These errors cannot be reduced by forming monthly or yearly averages of sea level. As explained by Diamante *et al.* (1987) special attention has to be paid to these systematic effects since, if not taken into account by other methods, they will introduce long-period effects that are almost identical to the actual trend in sea-level height.

One possible way of dealing with systematic effects is to combine sea-level readings for a number of tide gauges in a region. However, it should be noted that some of these systematic effects (like hydrodynamic draw-down) will influence all tide gauges (of the same type) in the same manner and, consequently, it is not possible to completely eliminate these effects by simple averaging tide gauge data over a large region. Combining time series for a number of tide gauges in a region does not necessarily yield a mean time series with an improved accuracy. Although taking an area average will reduce some of the systematic errors present in the individual time series, some new errors will be introduced as well. For example, the accuracy of the combined time series is limited by inaccuracies in the determined height differences between the tide gauge benchmarks of the tide gauges under consideration.

2.2.4 Acoustic reflection tide gauges

According to, e.g., Diamante *et al.* (1987), systematic errors inherent to stilling-well tide gauges can be largely reduced if the float is replaced by a remote sensor that does not require physical contact with the actual water surface. For this type of sensor it is no longer necessary to attenuate the high-frequency fluctuations, which makes mechanical filtering by means of a narrow orifice or inlet pipe superfluous.

Figure 2.4, gives a schematic drawing of a state-of-the-art acoustic reflection tide gauge. Sea-level height is measured by means of the time taken by an acoustic pulse to travel between the acoustic sensor, the instantaneous sea surface and back. From this two-way travel time and the height of the acoustic sensor relative to the tide gauge bench mark the instantaneous sea-level height can be determined.

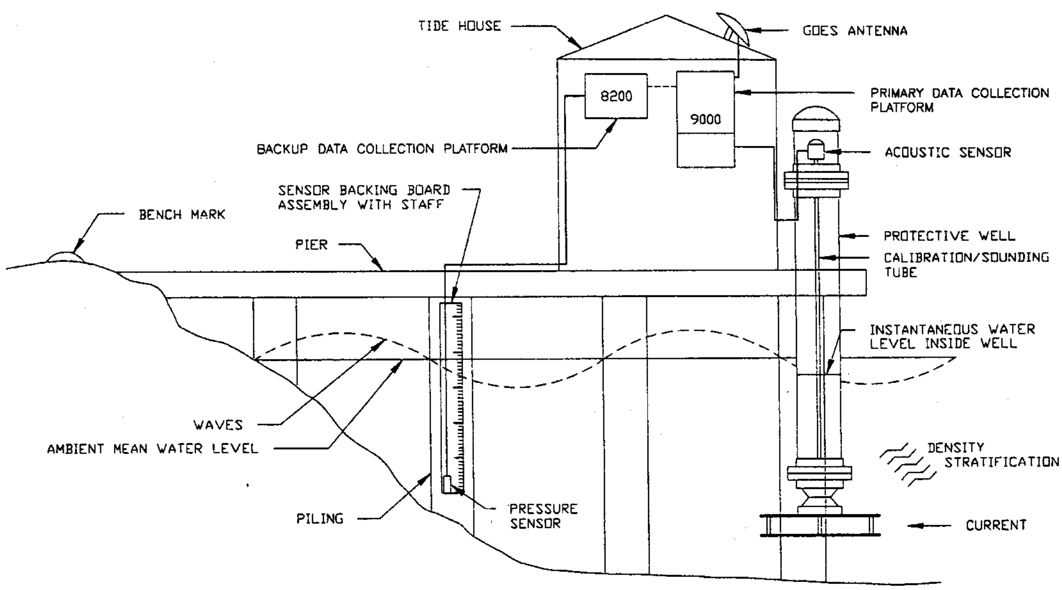

Fig. 2.4 Reflection tide gauge system; reproduced from Martin *et al.* (1996)

The two-way travel time (t_r) not only depends on the distance between the acoustic sensor and the sea surface (l_i) but also on the velocity of the acoustic pulse in the sounding tube (C_a), i.e.,

$$t_r = \frac{2l_i}{C_a}. \tag{2.1}$$

For dry air at a temperature of 10° and one hPa pressure, timing must be accurate within $6{\cdot}10^{-5}$ seconds, in order to detect a 1 cm change in sea-level height.

Velocity of sound in air depends on temperature, pressure and humidity, therefore, in order to accurately determine variations in sea-level height, variations in these parameters have to be taken into account. The new acoustic reflection tide gauge system as described by Martin *et al.* (1996) consists of a self-calibrating acoustic sensor that compensates for variations in C_a (due to temperature changes) in the sounding tube.

In theory, it is not necessary to attenuate high-frequency movements. However, in practice, present day acoustic measuring equipment still requires some kind of protective well in order to limit power consumption and procure accurate enough results; see Diamante *et al.* (1987). An open protective well might be sufficient. This has the advantage that sea-level conditions in the well will be more resemblant to the open sea (less sensitive to, e.g., density build-ups, etc.). In addition, this type of well is less prone to siltation and does not have a strong filtering effect. On the other hand, for a state-of-the-art station configuration Martin *et al.* (1996) still propose the use of a protective well with an orifice opening. However, they use a 15 cm diameter protective well with a 5-cm diameter orifice, whereas for stilling-well gauges Pugh (1987) recommends an orifice to well diameter ratio of 0.1.

According to Martin *et al.* (1996), sea-level heights are determined as average values over 6-minute intervals based on instantaneous measurements with a sampling rate of 1/s. Averages are determined over 3-minute periods, while sample outliers exceeding three times the standard deviation corresponding to this 3 minute interval are removed from the data. The resulting new averages provide one sea-level height measurement for every 6 minutes with a resolution of $\pm$ 1 mm. Smithson (1997) is slightly less optimistic; he estimates that a resolution of only about 3 mm can be achieved by averaging measurements with a sampling rate of 1/s over periods of 3 minutes.

2.2.5 Subsurface pressure tide gauges

As another alternative for the stilling-well tide gauge, nowadays often subsurface pressure tide gauges are used. These tide gauges measure pressure at a fixed point somewhere below the sea surface. This pressure can be converted to sea-level height using the relationship between measured pressure (P) and depth (D) as given by, e.g, Pugh (1987):

$$P = P_{A_w} + \rho_w g D, \tag{2.2}$$

in which P_{A_w} is the atmospheric pressure acting on the sea surface, ρ_w is the mean density in the water column, and g is the gravitational acceleration.

As an example of a subsurface pressure tide gauge system, figure 2.5 shows a schematic drawing of a bubbler gauge. The cylinder is open at the bottom so that water can flow in. A steady flow of compressed air or other gas is let into the connecting tube and can bubble out through an orifice (the pressure-point). As explained by Pugh (1987), for low rates of gas escape, gas pressure equals water pressure (P). This pressure is transmitted along the tube and recorded by the recording system, see Pugh (1987) for a description of recording systems.

Equation 2.2 requires atmospheric pressure measured at the instantaneous sea surface. Since this pressure cannot be obtained in practice, atmospheric pressure is measured by a sensor at a height (h_a) above the sea surface. According to Carrera *et al.* (1996) the relation between atmospheric pressure (P_A) at height h_a and atmospheric pressure at the instantaneous sea surface (P_{A_w}) is

$$P_{A_w} = P_A g h_a \rho_a, \tag{2.3}$$

in which ρ_a is the density of the air between the sensor and the sea surface. According to Woodworth (2000), usually a differential pressure transducer is used, i.e., a transducer which substracts atmospheric pressure from the sub-surface pressure. For this type of transducer, air pressure subtraction is automatic.

From equation (2.2) and equation (2.3) it is clear that the precision of sea-level heights determined by subsurface pressure tide gauges depend on the precision of the measurements of atmospheric pressure and "water" pressure, the precision of the value used for the gravitational acceleration, the precision of the values used for the water and air density respectively, and on the precision of the determined height between the pressure point in the cylinder and the atmospheric pressure sensor.

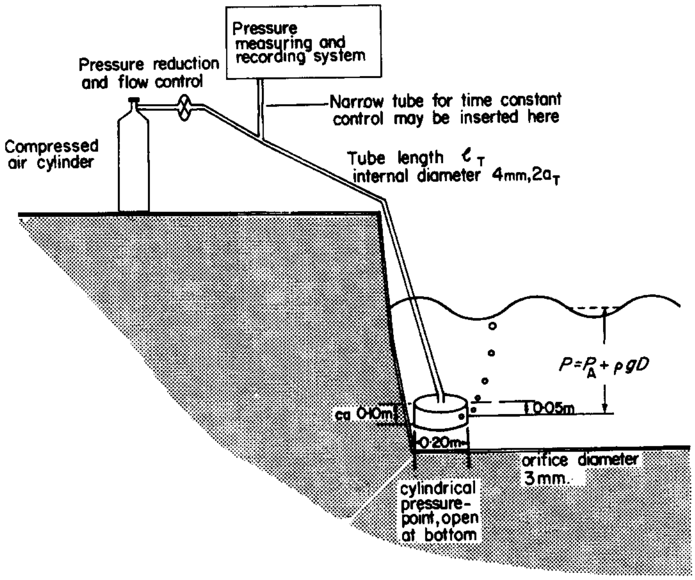

Fig. 2.5 Subsurface pressure tide gauge system; reproduced from Pugh (1987)

Carrera *et al.* (1996) give a comprehensive description of these error sources contributing to the overall precision of determined sea-level heights. Many of these error sources themselves depend on measurements of a number of other parameters, e.g., water density is a function of salinity, water temperature and pressure. According to Carrera *et al.* (1996), of the six parameters involved, uncertainties in "water" pressure, water density and gravity contribute the most to the resulting uncertainty in sea-level height, whereas the contribution of uncertainties in atmospheric pressure, atmospheric density and height between the pressure point in the cylinder and the atmospheric pressure sensor are relatively small.

Pugh (1987) estimates that the pneumatic system in a subsurface pressure tide gauge with a connecting tubes up to 200 metres, can produce water head equivalents (i.e., the depth of water that would produce a specific pressure) with a precision of 1 cm. Besides random errors, the accuracy of determined sea-level heights is also influenced by the occurrence of systematic errors. As described by Xu (1990), systematic errors are introduced by the non-linear relationship between water depth and pressure, hydrostatic pressure ranges between 0.1019 and 0.1034 kg/cm^2. Furthermore, large errors occur when water is forced into the connection tube by waves, for more details see Pugh (1987).

One of the major advantages of subsurface pressure tide gauges is that they are relatively convenient to use and they can be operated under difficult environmental conditions. One of their major drawbacks is that it is often difficult to relate the zero-height point of the system to the land based tide gauge bench mark. This is due to (different) biases and drifts inherent to the air pressure sensor and the "water" pressure sensor. This problem is addressed in detail by Woodworth *et al.* (1996), who estimate that subsurface pressure tide gauge data is often related to the land datum with a precision of about 2 cm.

Woodworth *et al.* (1996) describe a method that seems to be able to provide datum control with a precision of the order of only 1 mm. This method is based on an additional pressure point situated at a known height approximately near mean sea level. A description of this method is beyond the scope of this chapter and interested readers are referred to Woodworth *et al.* (1996).

2.2.6 Open-sea pressure gauges

The tide gauge instruments as described in the preceding sections can only be used to measure sea-level variations relative to a tide gauge bench mark on land, or on an off-shore platform. Using so-called open-sea pressure gauges it is also possible to measure sea-level heights at sites far away from the coast.

Usually, these tide gauge systems consist of a pressure sensor, placed on the ocean bottom, which is able to measure and record the pressure of the overlying water column. Depending on the water depth in which the tide gauge has to operate, different systems have been developed. In principle there are two categories of open-sea pressure gauges: deep-sea pressure gauges and pressure gauges for use on continental shelves.

According to Pugh (1987), in relative shallow water the tide gauge is often attached to a ground line, which simplifies the recovering of the gauge. In shallow water it is also possible to transmit recorded data to, e.g., a surface buoy, either through a cable or acoustically. This has the major advantage that the sea-level data can be collected without removing the tide gauge from its location on the bottom of the sea. Consequently, data becomes available almost real-time and tide gauges can continue operating on the same location (at least for as long as its power supply lasts).

For deep-sea operations, pressure tide gauges are usually built into a protective framework that free-falls to the bottom of the sea. These pressure gauges can operate in depths of over 4000 metres, with a resolution of 0.01 metres; see Pugh (1987). Since, presently, no methods have been developed to transmit the data while the pressure tide gauge is situated at the bottom of the sea, after a certain amount of time (usually 1 year), the tide gauge has to be recovered. This has the disadvantage that sea-level data is only available after recovery of the tide gauge, and undisturbed time series are only available over relative short periods of time. After recovery, the tide gauge can be lowered back to the bottom of the sea, but it will always be in a (slightly) different location and the measurements series is discontinued for the time period needed for the recovery and replacement of the tide gauge.

A new deep-sea pressure tide gauge called "MYRTLE" (Multi Year Return Tide Level Equipment) can, partly, overcome these problems. Figure 2.6 shows a schematic drawing of this tide gauge system, which consists of four data logger capsules. After one year of operation the first capsule is released and floats to the ocean surface where it can be recovered. After two years of recording the second capsule is released, etc. To prevent loss of data, each capsule contains data for the full measurement series as performed up to the release of the capsule. After four years of operation, the whole framework is released and can be recovered.

Major limitations for the accuracy of open-sea pressure system are the accuracy of the pressure sensor itself, drift of the pressure transducer zero, and settlement or lift of the instrument relative to the bottom of the sea. Pugh (1987) estimates that the framework will settle into the sediments with a velocity of a few cm/month. The accuracy of the pressure transducer is, amongst others, affected by its sensitivity to temperature changes. Pugh (1987) estimates that, in depths of (at most) 200 metres drift of the transducer zero can be reduced to a few cm/month, while in depths of around 4000 metres drift can be significantly larger. In addition, according to Banaszek (1985), the pressure sensor, and especially the framework in which the tide gauge is mounted, can seriously distort the velocity field and, consequently, pressure detected by the sensor deviates from the hydrostatic pressure that is related to depth. Spencer and Vassie (1997) state that the main problem with pressure sensors at great depths is "creep", i.e., a monotonic (although not linear) drift of the pressure measured by the sensor.

Variations in density through the water column above the pressure transducer should be taken into account as well, e.g., by calibration of the instrument based on measurements through the water column. Furthermore, pressure as measured by the transducer not only depends on the amount of water overlying the transducer but also on the atmospheric pressure acting on the sea surface. Rae (1976) estimates that pressure variations due to atmospheric pressure can be of the order of 50 millibars. However, since sea-level response to variations in atmospheric pressure is almost inverse barometric (at least for "open" seas), the total pressure at the sea bottom will not be significantly affected. On the other hand, according to Rae (1976), density variations (e.g., due to internal waves) can cause variations in total bottom pressure of about a few millimetres.

2.3 Sampling rate and averaging method of tide gauge readings

In the preceding section, for the six major tide gauge systems, error sources have been described. However, predictions of sea-level variation curves are usually based on some kind of average values (e.g., monthly

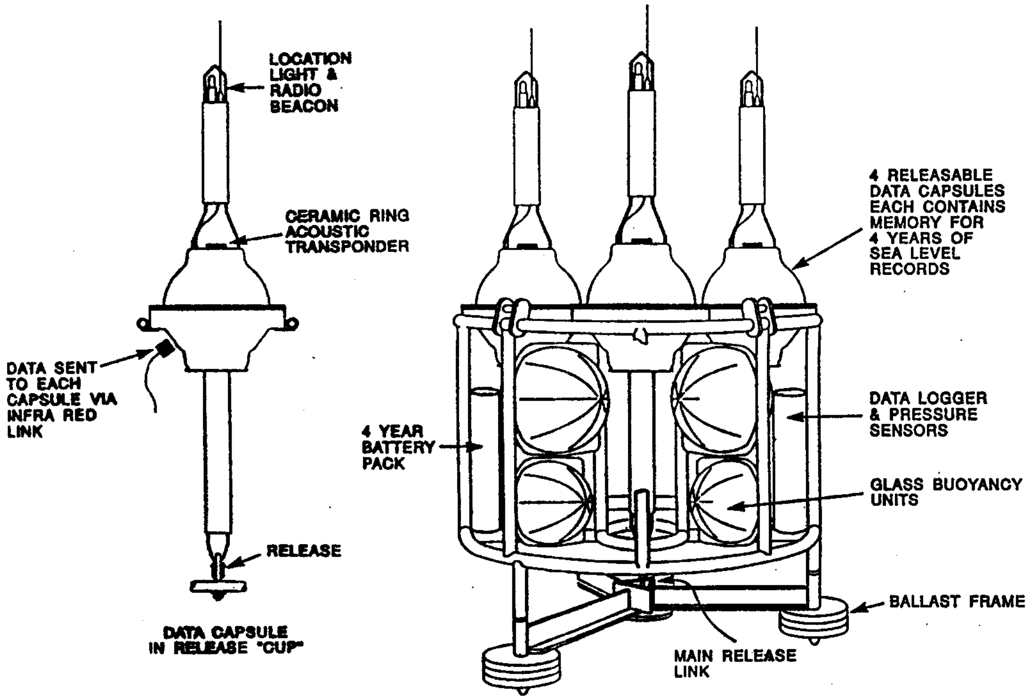

Fig. 2.6 Deep see pressure tide gauge system capable of 4 years of recording; reproduced from Smithson (1997)

or yearly sea-level heights) instead of on instantaneous measurements of sea-level height. Consequently, the accuracy of these predictions is not only determined by the accuracy of the individual tide gauge measurements but also by the sampling rate of these measurements and the method applied to these individual measurements to form mean values.

The reason for basing evaluations of sea-level variations on averaged sea-level heights is simply that records of instantaneous sea-level values are, in general, not available. In the past, when observations of sea-level height were simply written down, only monthly mean values (or even yearly mean values) were recorded for long-term keeping. Even after the introduction of mechanical recording (e.g., on a paper chart mounted on a drum) usually only monthly mean values were stored. For the more recent past, sometimes hourly mean values are available, often obtained by digitising the paper charts at hourly intervals. Although state-of-the-art tide gauge systems usually work with a relatively high sampling rate, again, only average values are stored. As an example, acoustic reflection tide gauges measure once every second, but only 3-minute average values are stored.

An advantage of mean sea-level heights is that, depending on the method used to form the averages and the time span over which the average is taken, high frequency signals (like waves) and periodic effects (like tides) are, partly, removed from the data. Different averaging methods, which have been developed through the years, will be described in the following sections. In complexity these methods range from simply taking the arithmetic mean of only two measurements a day, to low-pass filtering based on 24 hourly values a day.

A limitation of historical sea-level data based on manual readings by an observer is that these data often have a relatively low sampling rate, i.e., mean values are based on only a few readings a day. In addition, the oldest data is often based on irregular sampling over the tidal cycle, e.g., a tide pole was only read during daytime (at high and low water). Since this limited number of sea-level height samples outlines the averaging method that can be used, its effects will be described in the following sections where the different averaging methods will be introduced.

Finally, it should be noted that the accuracy of mean sea-level values is also influenced by how well outliers and systematic errors in the individual sea-level measurements have been corrected for. As described in detail by Pugh (1987), prior to forming the average values, individual measurements should be checked for reading errors and corrected for timing errors. Timing errors, i.e., measurements are not recorded at the exact time they are supposed to be recorded, might for example result from variations in the rotational speed of the drum on which the paper chart is mounted. In addition, it might be necessary to use interpolated values to fill gaps in the data series. Since this is beyond the scope of this chapter, for a description of methods that can be used to check the recorded measurements, the interested reader is referred to, e.g., Pugh (1987).

Depending on the averaging method used, periodic effects (with a period that is short relative to the time span over which the measurements are averaged) are, partly, removed from the resulting mean values. However, the influences of low-frequency tides (like the lunar nodal tide with a period of 18.6 year) will more or less remain in monthly and even in yearly mean sea-level heights. In addition, mean monthly and annual sea levels are influenced by the occurrence of storm surges.

2.3.1 Low-pass filtering of hourly values

The most advanced method of forming mean (monthly or annual) sea-level heights is low-pass filtering of hourly (mean) sea-level heights to obtain smoothed daily mean sea levels. By taking the arithmetic mean of these smoothed daily sea levels over a period of a month or a year, respectively monthly or annual mean sea-level heights can be derived.

Different low-pass filters, requiring a different number of hourly (mean) sea-level values have been developed, the most widely used is the so-called Doodson X_0 filter, which uses 39 hourly values. As described in, e.g., IOC (1985) the smoothed daily value ($X_F(t)$) is derived by applying the filter F_m to the 39 hourly values in the symmetric window around time t:

$$X_F(t) = \frac{1}{30} F_0 X(t) + \frac{1}{30} \sum_{m=1}^{19} F_m \left[X(t+m) + X(t-m)\right] \tag{2.4}$$

in which the filter elements are defined by

m	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
F_m	0	2	1	1	2	0	1	1	0	2	0	1	1	0	1	0	0	1	0	1

This Doodson low-pass filter is designed to better eliminate the influence of diurnal and semi-diurnal tidal constituents from the resulting monthly and annual mean sea-level heights. How well these filters succeed in eliminating tides, depends on the characteristics of the local tidal regime, i.e., the relative contribution of the different tidal constituents. Rossiter (1958) gives examples of attenuation factors for a number of tidal constituents for two different tidal regimes.

Besides the 39 hour filter, Doodson has also developed filters using even more hourly values. Examples are filters using 72, or even 168 hourly values; for filter coefficients see Pugh (1987). According to Pugh (1987), monthly means based on these longer filters differ insignificantly from those based on the filter using 39 hourly values. In addition, the filter based on 39 hourly values can be applied to time series containing occasional data gaps. The longer the filter length, the larger the amount of data that is lost on either side of the gap.

2.3.2 Arithmetic mean of hourly values

Instead of using low-pass filters to smooth the hourly values in a first step, often the hourly values themselves are used directly to form monthly or annual mean sea-level heights. These mean values are simply determined as the arithmetic mean over all (mean) hourly sea-level values in a month or a year respectively. According to Rossiter (1958), this method is the most widely used because it involves less computational effort and results in mean values close to those based on low-pass filtering. Pugh (1987) gives the maximum (remaining) contributions of a number of tidal constituents to monthly mean sea levels.

From 1971 onwards in the Netherlands mean (monthly and annual) sea-level heights have been determined as arithmetic means of hourly sea-level heights; see Van der Hoek Ostende and Van Malde (1989). At first these hourly values were derived by digitising mechanically produced sea-level charts at hourly intervals.

Starting in 1987, mechanical recording on charts has been replaced by digital recording of 10-minute mean values. The resulting hourly values are rounded off to the cm-level, the arithmetic means over all hourly values in a year are rounded to the mm-level.

Figure 2.7, shows for 10 tide gauges in the Netherlands for which more than 100 years of tide gauge data is available, which method of forming annual mean sea-level heights has been applied through the years. This figure clearly shows that, the further back in time, the smaller the number of daily observations used to determine annual mean sea-level heights. In the following sections, in descending order of number of daily observations, these historical methods of forming annual mean sea-levels will be described.

2.3.3 Arithmetic mean of 3-hourly values

Prior to using hourly measurements, annual mean sea-level heights in the Netherlands were based on 8 measurements/day. As described by Van der Hoek Ostende and Van Malde (1989), sea-level values at 2, 5, 8, 11, 14, 17, 20, and 23 o'clock were read from the paper charts and rounded off at cm-level. From these 3-hourly values annual mean sea level was simply determined as the arithmetic mean over all values in one year, rounded to 1 mm.

As described by Van der Hoek Ostende and Van Malde (1989) measurements performed before 1961 usually refer to Amsterdam time, later measurements are usually made with reference to MET (Middle European Time), the time difference between the two systems being 40 minutes. However, according to Van der Hoek Ostende and Van Malde (1989) the effect of this time-shift on the determined annual mean values is negligible.

Prior to using 3-hourly measurements, annual mean sea-level heights were (usually) based on only 4 daily measurements, see section 2.3.4. As can be seen from figure 2.7, the moment when the change was made between these two sampling rates varies widely for the various stations. As an example, for the station Den Helder already around 1885 mean sea levels were based on 8 observations/day, whereas only from 1936 onwards this method was used for all tide gauge stations along the Dutch coast.

Experiments based on hourly mean values for the period between 1971 and 1986, as performed by Van der Hoek Ostende and Van Malde (1989), showed that the difference between annual mean sea-level heights based on the 3-hourly values and annual mean sea-level heights based on hourly values is relatively small. The mean value for the difference between these two time series of annual values ranges between -0.4 and +0.6 mm for the various stations, whereas the standard deviations range between 0.3 and 1.0 mm. According to Van der Hoek Ostende and Van Malde (1989), the stations under consideration are situated in largely varying tidal regimes, which can be assumed representative for the Dutch coastal zone. Therefore, they conclude that using 3-hourly values instead of hourly values will hardly introduce errors for other tide gauge stations as well.

2.3.4 Arithmetic mean of 6-hourly values

From figure 2.7 it can be seen that after mechanical recording became available, for most stations the switch was made from annual mean sea levels based on (daytime) tide levels to annual mean sea-level heights based on 6-hourly values. Analogous to the method as applied to the 3-hourly measurements, sea-level heights at 2:00, 8:00, 14:00, and 20:00 o'clock Amsterdam time were read from the paper chart and the arithmetic mean of all values in one year (rounded at the mm-level) was taken to form annual mean sea-level heights.

Van der Hoek Ostende and Van Malde (1989) have compared, for the period between 1971 and 1986, annual mean sea levels based on 6-hourly values with those based on hourly values. For a number of stations the difference between the two resulting time series is rather large (of the order of 1 cm). These are all stations situated in a rather special tidal regime, where a short-duration increase in sea-level height is seen at low tide, which is related to a relative long duration of low tide. Consequently, relative large deviations are introduced for these stations if annual mean sea level is based on 6-hourly values instead of on hourly values.

Also for stations situated in more "normal" tidal regimes there are significant differences between annual mean values based on 6-hourly values and those based on hourly values. For these stations, the

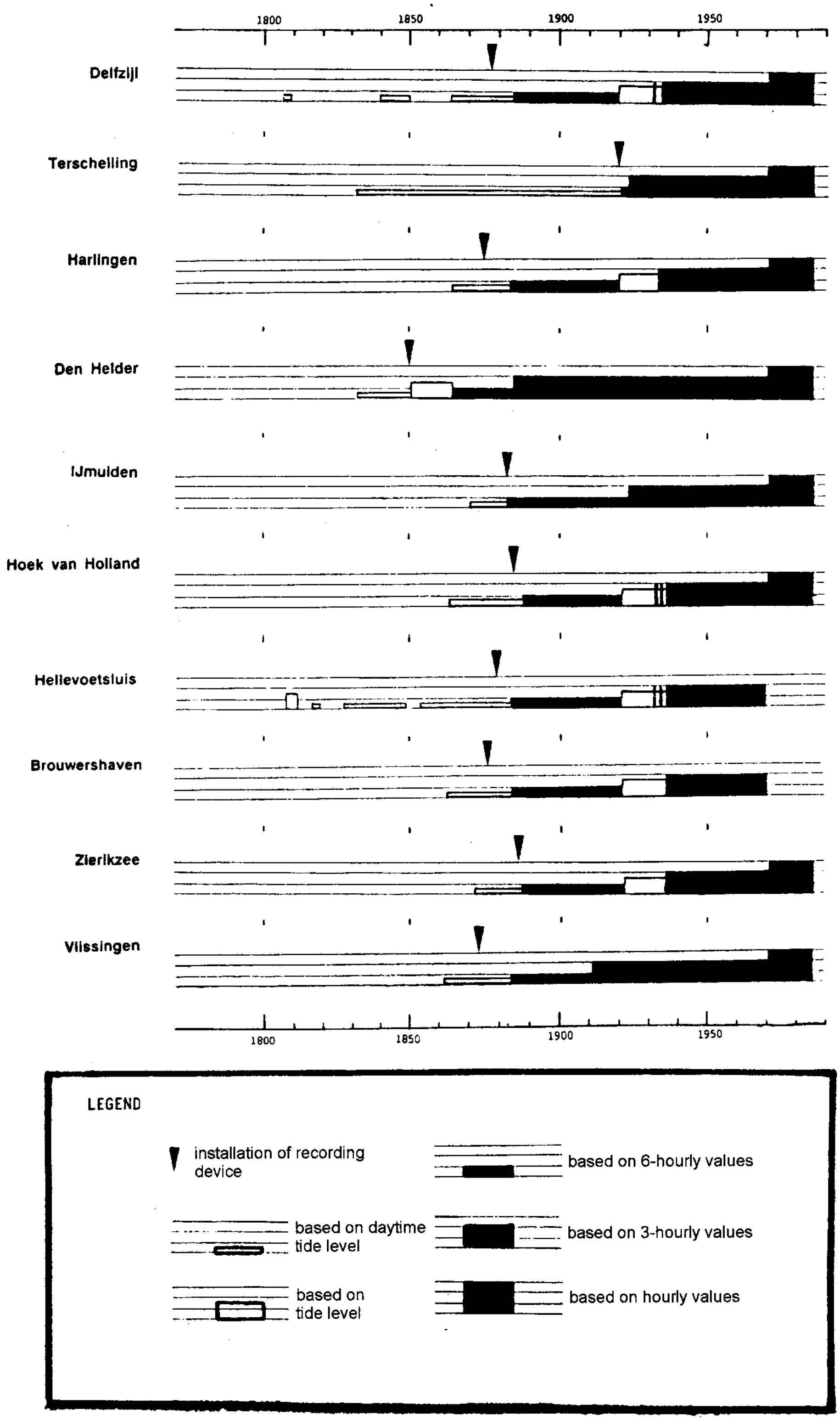

Fig. 2.7 Historical overview of averaging methods applied to tide gauge measurements to form mean annual sea-level height; reproduced from Van der Hoek Ostende and Van Malde (1989)

mean value for the difference between the two time series of annual values range between -3 and +6 mm, whereas the standard deviations range between 1.2 and 2.8 mm; see Van der Hoek Ostende and Van Malde (1989).

2.3.5 Mean sea-level heights determined from mean tide level

For a number of stations, between 1920 and 1935, yearly mean sea-level heights were based on mean tide levels; see figure 2.7. Since hourly values were, in principle, available this was probably done in order to simplify computations. Contrary to the period before the installation of mechanical recording devices (see next section), these tidal mean values were not only based on daytime measurements of high and low tide level but on nighttime values as well. These high and low tide values were simply read from the paper charts with a resolution of 1 cm; see Van der Hoek Ostende and Van Malde (1989) for more details. Until 1971, in addition to hourly sea-level values, these high and low water values have been recorded.

By taking the arithmetic mean of resp. all high tide values and all low tide values in one year, annual mean high tide and low tide values were determined and rounded to the nearest mm. By averaging these two mean tide values the annual mean tide value was found. From these annual mean tide levels, annual mean sea levels were determined by applying a specific shift. According to Van der Hoek Ostende and Van Malde (1989) this was based on the, false, assumption that with a high enough accuracy the difference between annual mean sea level and annual mean tide level for a specific station could be assumed constant.

According to Pugh (1987), the difference between mean sea level and mean tide level is a result of shallow-water tidal harmonics. However, variations in these two levels are highly correlated. As explained by Pugh (1987), for every combination of tidal constituents (e.g., M_2 and M_4, or K_1 and K_2), the systematic difference between the mean sea level and mean tidal level can be computed. The resulting shifts are dependent of the location of the tide gauge.

For the stations under consideration, shifts to determine mean sea level from mean tide levels range between -15.5 and +18 cm. Unfortunately, although the values used for these corrections are known, it is (as yet) unknown how these values have been derived; see Van der Hoek Ostende and Van Malde (1989) for more details.

For the period between 1971 and 1986, Van der Hoek Ostende and Van Malde (1989) compared annual mean sea levels based on mean tide values with annual mean sea levels based on hourly values of sea-level height. For the various stations they found a wide range in differences between these two time series of annual mean values. The mean value for the differences ranging between -25 and + 21 cm, and the standard deviations ranging between 0.4 and 1.5 cm.

In addition, experiments by a number of authors (see Van Malde (1986)) have shown that the tidal regime along a major part of the Dutch and German coasts is changing, i.e., the mean difference between high and low tide increases. Consequently, estimating mean sea-level height by applying a shift to mean tide level can yield an (increasingly) large systematic error. As a result of this phenomenon and the large standard deviations found between time series based on mean tide values and hourly sea-level values, Van der Hoek Ostende and Van Malde (1989) conclude that annual mean sea-level values based on mean tide values are rather inaccurate.

2.3.6 Mean sea-level heights determined from mean daytime tide level

As described by Van der Hoek Ostende and Van Malde (1989), in the Netherlands, before the installation of mechanical recording devices only high and low water level were measured. Measurements were made during daytime (between 6:00 and 18:00 Amsterdam time), with a 1-cm resolution. All values of high water level in one year were combined in a mean daytime high-tide level. Mean daytime low-tide level was accordingly determined as the arithmetic mean of all daytime low tide values in a year. Both mean tidal values were determined with a resolution of 1-cm. Next, annual mean daytime tide level was calculated as the arithmetic mean of mean daytime high-tide level and mean daytime low-tide level, and rounded off to an integer number of cm. This rounding off was performed as follows: if the integer part of the mean value was an even number it was rounded down, an uneven value was rounded up.

From these annual mean daytime tide levels the annual mean sea levels were determined by applying a shift. The various stations have shifts that are, usually, constant over the whole period of time for which this method was applied. However, for some stations the value used for the shift has been changed at a certain time. For the 10 stations under consideration, values for the shifts range between -17 and +18 cm. Analogous to the applied shifts to correct mean tide levels, it has not been possible to discern how these values have been determined; see Van der Hoek Ostende and Van Malde (1989).

It should be noted when comparing sea-level data for various tide gauges, that although for most tide gauges (as mentioned in these sections) the switch to 6-hourly measurements was made around the same time (somewhere around 1885) for station Terschelling annual mean sea levels were still based on annual mean daytime tide level until around 1920, see figure 2.7.

For the period between 1923 (resp. 1936) and 1960, Van der Hoek Ostende and Van Malde (1989) compared, for 7 stations along the Dutch coast, annual mean sea-level heights based on mean annual daytime tide level with those based on mean annual tide level. They found differences between the mean values of these two time series of up to 6.4 mm (for station Den Helder). However, as explained by Van der Hoek Ostende and Van Malde (1989), this difference in mean value does not necessarily have a large influence on the accuracy of the mean daytime tide level method. Since both time series of annual mean sea level are obtained by applying a correction to the mean tide levels, the difference between the two series can be minimised by changing the shift as applied to the mean annual daytime tide values. However, standard deviations for the time series based on mean daytime tide level are between 10 and 30% higher than standard deviations for the series based on mean tide level; both relative to annual mean sea-level heights based on 3-hourly measurements.

As described by Van der Hoek Ostende and Van Malde (1989) in addition to inaccuracies introduced by using only daytime samples, less precise annual mean sea levels are obtained because annual mean daytime high tide and low tide levels are rounded off to an integer number of centimetres. Compared to annual mean sea-level heights based on mean high tide and low tide values, which are rounded to the nearest mm, this rounding off already introduces a standard deviation of about 3 mm.

2.4 Conclusions and recommendations

In the preceding sections error characteristics of a number of tide gauge systems have been discussed. It is clear that tide gauges without automatic recording are unsuitable for high quality monitoring of sea-level height variations over a longer period of time, due to the limited measuring frequency and the susceptibility to reading and recording errors. State-of-the-art tide gauge systems are equipped with a mechanical or digital recording device, which allows continuous measurements or measurements with a relative high sampling rate. These recording devices introduce some errors of their own, e.g., errors due to mechanical problems (like friction) or due to aliasing. In addition, the quality of the measured sea-level heights is influenced by the occurrence of systematic errors. Tide gauges based on different techniques, i.e., based on a float in a stilling well, acoustic travel times, or a comparison between water pressure and atmospheric pressure, are susceptible to different systematic errors.

For a new site, a suitable tide gauge system should be selected based on characteristics of the specific location (e.g., easy to reach for maintenance), environmental conditions (e.g., the occurrence of large currents), expected measuring precision, and available budget. However, for analysing existing tide gauge data, it should be remembered that often less optimal tide gauge systems have been used. For example, in the Netherlands tide gauge data from the 19th century is usually based on manual recording.

Since often for longer periods of time only monthly or annual sea-level heights are available, the method used to form these average values should be considered as well. Especially for relatively old tide gauge data, averaging methods have been used that yield biased mean values. For example, in the Netherlands only from 1935 onwards for all tide gauges mean values were based on 3-hourly or hourly values. Prior to this date, often only 6-hourly values were used or mean sea-level heights were even based on mean tide levels.

If differences in tide gauge systems or averaging methods have occurred over the time span of the measurements, this will influence estimates based on the time series. Therefore, when evaluating long series of sea-level values, any available information concerning how the mean sea-level heights have been obtained, should be taken into account. Often it will not be possible to adapt the measured values in order to improve the consistency of the time series. However, at least some care can be taken when examining results based on a time series for which it is known that major differences in measurement accuracy might have been introduced by changes in tide gauge equipment or averaging methods.

Chapter 3

Techniques for analysing sea-level data

3.1 Introduction

This chapter discusses a number of analysing techniques that can be used to estimate patterns in time series of sea-level heights. These methods range in simplicity from simply fitting a straight line through all data points to the more complicated AR(I)MA models and filtering based on wavelets. Data analysing techniques can be used for two different purposes. First of all, they can be used to monitor sea-level heights. The objective is to reveal a pattern in past sea-level height variations and to examine how this pattern changes over time. Secondly, data analysing techniques can be used to forecast future sea-level heights, based on a predicted pattern in available sea-level data.

From monthly and annual mean sea-level values, as supplied by the Permanent Service for Mean Sea Level (PSMSL), the major part of the effects of the diurnal and semi-diurnal tides have disappeared. However, these time series still contain a wide range of periodic fluctuations. By smoothing the data sets, the underlying decadal variation pattern can be revealed. How much smoothing is required depends on the purpose for which the resulting data set is required.

Predictions of future sea-level heights require insight in expected changes in the present-day sea-level rise curve for the near future. How future sea levels are influenced by, e.g., greenhouse-gas induced warming, is difficult to predict. This not only requires insight in expected warming of global temperatures (e.g., based on a GCM), but knowledge on how this increase in temperature relates to sea-level height changes as well. The relation between a change in global temperature and sea-level rise at a specific location is complicated and requires knowledge of, for instance, thermal expansion and changes in ice load; see e.g., Warrick *et al.* (1996) for more details.

Accurate predictions of future sea levels should not only be based on expected sea-level height variations as a result of anthropogenic activities, but estimates of "natural" changes in sea-level heights have to be included as well. Estimates of "natural" sea-level height variations for the near future could be obtained by forecasting sea-level values based on existing (long) sea-level height time series.

Different purposes of forecasting future sea levels require different degrees of detail. For a general indication of the behaviour of mean sea levels over the years, a simple estimate of the trend value of sea-level rise can be used. The long-periodic cyclic pattern can be added to the forecasted values to give some extra information. However, a trend estimate (or estimate of a higher-order sea-level variation curve) is not sufficient for assessing whether or not current dikes will still be sufficient protection against flooding of our low-lying areas in 20 or 30 years time. This kind of predictions should not only forecast the expected trends and cyclic patterns, but on top of these the extremes that can occur due to shorter periodic fluctuations like spring tides and storm surges.

Modelling future sea level height changes based on models predicting global change is beyond the scope of this thesis. Therefore, in this chapter it is assumed that we are only interested in predicting future sea levels to gain insight in what can be expected based on the patterns in sea level that have occurred up to now. Consequently, forecasting techniques should be able to predict a mean trend in sea-level rise, possibly in addition with expected long-periodic fluctuations. A number of techniques that could be used for this purpose, will be described in section 3.3.

If more than one time series is available, it should be verified whether or not a common curve is sufficient to describe the sea-level variations for this group of tide gauges. Methods to examine the correspondence between different time series and to estimate a common curve will be explained in section 3.4. In addition, fitting of a spatial sea-level variation pattern to the group of tide gauges will be described.

Characteristics of sea-level height time series As compared to data sets usually employed in literature on analysing time series, sea-level height data has some specific characteristics. Time series as used by, e.g., Box and Jenkins (1976), usually consist of only a moderate number of samples. In addition, they have a homogeneous measuring accuracy and contain a limited number of periodic effects, for which the (constant) period is often known in advance.

Sea-level time series often cover very long periods of time. For example, along the Dutch coast, for nine tide gauges sea-level time series containing over 100 years of data are available. Although long time series have the advantage that they allow the detection of a pattern emerged in noise and in long-term periodic fluctuations, they have a disadvantage as well. The longer the time series, the larger the chance that circumstances have changed over time, e.g., changes in the environment near the tide gauge or in the tide gauge equipment used. As a result, the precision of the sea-level height measurements might vary and even the underlying sea-level variation curve might have been modified.

Contrary to, e.g., time series containing monthly sales data with only a very limited number of periodic effects, sea-level height series contain a wide variety of periodic effects. Some of these fluctuations have constant periods that are (more or less) known, like astronomical tides and rotational tides. In addition, sea-level variations are generated by processes that cause more or less periodic variations in sea-level height, but for which the period is not constant over time.

In addition to periodic fluctuations tide gauge time series are also contaminated by noise. As a result of changes in tide gauge systems and frequency of measurements, the amount of random noise will not be constant over the time series. Apart from changes in the precision of mean sea-level values, sea-level height series can be influenced by a number of systematic effects. Some effects have a slowly increasing impact, like silting up of channels in the neighbourhood of the gauge. Others, like relocation of the tide gauge or harbour reconstructions may cause discontinuities in the sea-level height time series.

3.2 Smoothing of tide gauge data

A large number of analysing techniques are available for smoothing of time series. Only a limited number of techniques will be discussed in this chapter. They can be divided into the following categories:

- Techniques based on averaging.
  A smoothed value is determined as the weighted average of a (specific) number of measurements surrounding or preceding the sample under consideration. Examples are: moving average smoothing and exponential smoothing.
- Techniques based on singular value decomposition of groups of measurements.
  Reconstructions based on the first (few) singular values, representing the signal common to a group of time series can be used as smoothed signal. The method of Singular Spectrum Analysis allows this technique to be applied to individual time series.
- Techniques based on processing in the spectral domain.
  Examples are Fourier and Wavelet analysis.

The various smoothing techniques will be assessed on how well they are able to separate the underlying decadal pattern from higher frequency fluctuations and noise. However, since the "true" decadal pattern is not known, it is impossible to indicate which technique predicts this pattern best. Furthermore, using, e.g., the sum of squared errors between the original and smoothed data will not give an indication of how well a technique works either. The less the data is smoothed, the smaller this sum of errors will be.

Since a real measure indicating the precision of the techniques is difficult to find, results based on the various techniques will be compared. The robustness of the techniques is important as well. The term robustness is used as an indication of how much predictions change if a (small) change in the model is made. A suitable technique should not only be able to distinguish the decadal pattern, but this pattern should be robust against small changes in the model as well.

Most of the methods as will be discussed in the following sections will be applied to annual mean sea levels for tide gauge Den Helder. On a few occasions mean yearly values for tide gauge Vlissingen will be used.

3.2.1 Techniques for pre-treating individual tide gauge data

Pre-treating of data might be necessary in case significant discontinuities occur in the data set. For the tide gauge data under consideration, no gaps occur within the time series. Therefore, no correction for missing data is needed. However, due to its specific characteristics (e.g., the possibility that the tide gauge has been relocated) other corrections to the data set might be needed. Two specific pre-treating methods will be mentioned.

"Normalising of data" Monthly mean sea-level heights are related to calendar months, which range in period between 28 and 31 days. As explained in detail by Cartwright (1983), this non-uniform sampling of averages and irregular intervals between them can cause blurring of spectral information. This might cause a problem when using Fourier-type techniques to examine the data.

Monthly mean sea levels can be normalised by applying weights depending on the number of days in the corresponding month. In the time domain, the difference between normalised and original values is not very large. For example, for tide gauge Den Helder the difference between original annual mean sea levels and annual mean sea-level heights based on normalised monthly values is of the order of a few mm.

Omitting deviating parts In extreme cases, e.g., if large jumps occur, visual inspection of the data itself can reveal that part of the data strongly deviates from the general pattern present. The visibility of this deviation might be enhanced by detrending the time series; see figure 3.1. In the left-hand side figure (for station Den Helder), apart from periodic fluctuations, no clear pattern emerges. The right-hand side figure (tide gauge Vlissingen) shows a jump in the data around 1885, whereas in the remaining part of this series a trend can still be discerned. This is a clear indication that the overall trend estimated for this time series does not fit the remaining part of the data very well.

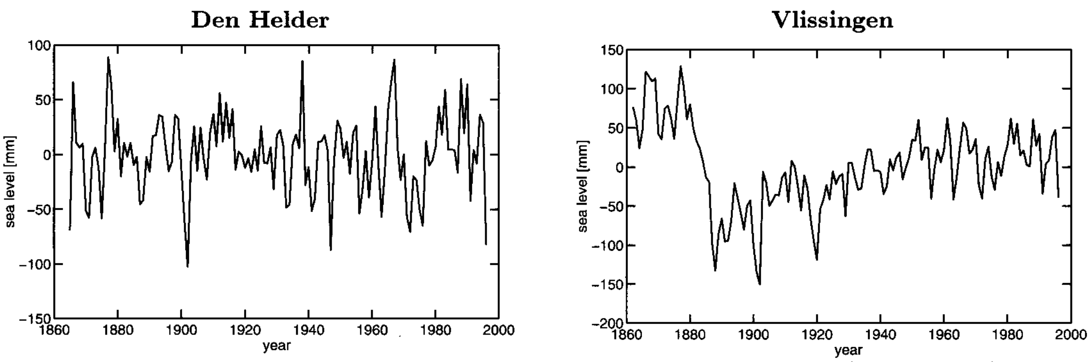

Fig. 3.1 Detrended annual mean sea-level heights for resp. tide gauge Den Helder (left-hand side) and tide gauge Vlissingen (right-hand side)

For an extreme case, like the sea-level height series for tide gauge Vlissingen, jumps in the data set can easily be identified by visual inspection of the data itself. By omitting the deviating part of the data set, a more consistent set of sea-level heights, allowing more reliable predictions, can be obtained. If less extreme jumps are present, in a later stage, deviations in the data might be revealed because, e.g., AR(I)MA modelling shows difficulties with data containing (small) jumps.

As a simple example to illustrate the necessity of removing deviating observations in the first part of the time series, linear regression is applied to annual mean sea-level values for tide gauge Vlissingen. Based on the complete data set as provided by the PSMSL, i.e., values between 1862 and 1997, a sea-level rise of 1.2 mm/yr is estimated. The left-hand side of figure 3.2 shows the annual mean sea levels, the estimated trend, and the expected sea-level height for the year 2050 based on this trend.

From figure 3.2, a relative sudden increase in measured sea-level heights around 1885 is clearly visible. Omitting this first part of data and using records starting only in 1887 (which still implies 110 years of data), yields a significantly larger estimated yearly sea-level rise. By omitting the first 25 years of data the estimated trend increases with 1 mm/yr, i.e., from 1.2 to 2.2 mm/yr. As a result, if this estimated trend is used to predict the expected increase in sea-level height for the year 2050 (relative to 1997) a twice as large value is found, i.e., 114 mm instead of only 63 mm.

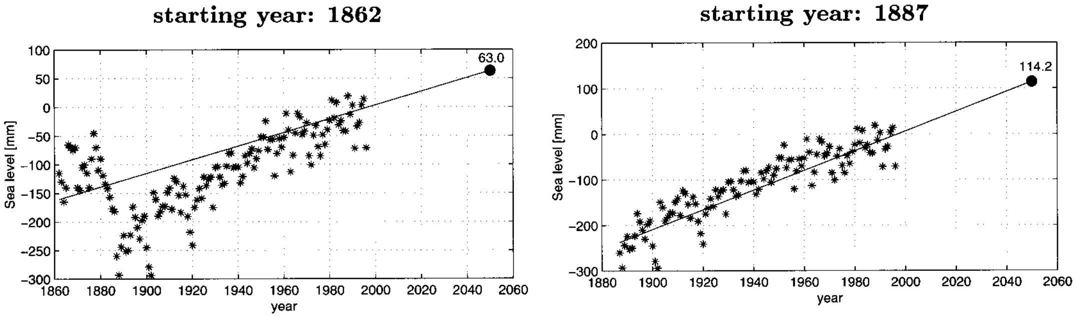

Fig. 3.2 Tide gauge Vlissingen: annual mean sea-level values, estimated trend, and predicted sea level for the year 2050; relative to the year 1997.

If only a few measurements in the time series contain strongly deviating values, these samples can be found by outlier tests. However, for data sets as those obtained by tide gauge Vlissingen, a significant part of the time series contains sea-level heights that do not correspond to the remainder of the data set, e.g., because the tide gauge has been relocated. In principle, this type of discontinuities in the time series can be revealed by testing whether or not the proposed model fits the data. For example, if a linear regression line is fitted to the data, significant discontinuities occur if the hypothesis $H_0 : E\{y_i\} = a + bx_i$ is rejected in favour of the alternative hypothesis $H_1 : E\{y_i\} \neq a + bx_i$. The test statistic

$$t = \frac{\hat{\sigma}^2}{\sigma_0^2} \tag{3.1}$$

can be used, which should follow a $F_{1-\alpha}(N-1, \infty)$ distribution (under H_0), with N the number of observations, $\hat{\sigma}$ the a-posteriori standard deviation, and σ_0 the a-priori standard deviation. Unfortunatelly, for sea-level height series it is very difficult to make a realistic estimate of this a-priori standard deviation. Depending on the value used, the model $y_i = a + bx_i$ is either accepted or rejected.

As an alternative method, two trend values are estimated. One value ($\hat{b}'$) for the complete time series, and one value ($\hat{b}$) based only on the last part of the data set (the observations which have been obtained well after a possible relocation of the tide gauge). It is assumed that discontinuities have occurred in the beginning of the time series if the difference in trend ($\hat{\Delta b} = \hat{b}' - \hat{b}$) is significant. In this case the hypothesis $H_0 : E\{\hat{\Delta b}\} = 0$ is rejected in favour of $H_1 : E\{\hat{\Delta b}\} \neq 0$. This can be tested by:

$$t = \frac{\hat{\Delta b}}{\hat{\sigma}_{\hat{\Delta b}}} = \frac{\hat{\Delta b}}{\sqrt{\hat{\sigma}^2_{\hat{b}'} + \hat{\sigma}^2_{\hat{b}}}} \tag{3.2}$$

which should (under H_0) follow a Student's distribution with $N-4$ degrees of freedom.

For tide gauge Vlissingen, the discontinuities in the beginning of the time series are obvious from the plot of annual mean sea levels. Consequently, they can easily be removed from the data set. For other time series inspection of the sea-level values does not reveal discontinuities, although in some cases (e.g., Delfzijl as discussed in section 3.4) discontinuities have occurred. Consequently, before sea-level data is used to determine a (common) variation curve, the data sets should be checked (e.g., by testing) for discontinuities. However, the purpose of this thesis is not to derive optimal results with actual tide gauge data but to examine the influence of measuring errors on the detectability of sea-level variation curves. In chapters 4, 6, and 8 simulated data sets are used, and no tests for discontinuities have to be applied.

3.2.2 Moving average smoothing

Every element in the time series is replaced by a (weighted) average of surrounding elements. The number of surrounding elements is defined by the width of the smoothing window. The size of this window might be taken equal to the largest period of (major) period fluctuations present in the data. This would suggest a window size of 19 years so that it contains a full period of the lunar nodal tide (period 18.6 years). As an alternative, the amplitude spectrum following from a Fourier transform of the data can give an indication of the maximum period of fluctuations in the time series.

In a so-called rectangular window, all elements have equal weights. The triangular window assigns greater weight to the observation being smoothed in the centre of the window. This window requires an uneven number of samples within the window. Weights are obtained by convolving 2 rectangular windows followed by scaling in order to realize unitary total weight. Different windows show a different behaviour in the time and frequency domain. In principle, a window could be designed with specific properties (in time or space domain) but this is beyond the scope of this chapter.

Figure 3.3 shows annual mean sea levels (tide gauge Den Helder), smoothed by a 19 year moving average filter. Results based on a rectangular window show a large similarity with those calculated using a triangular window. Next it is tested how robust smoothed values are against a small change in window size. Figure 3.4 shows moving average smoothed sea levels for three rectangular window sizes. From this figure it is clear that a change in window size from 17 to 19 or even 21 years, hardly changes the resulting smoothed values. Same results are obtained if weights corresponding to triangular windows are used.

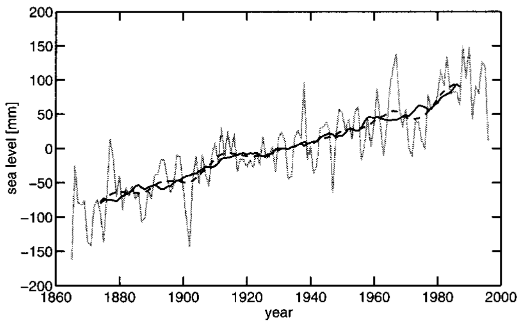

Fig. 3.3 Annual mean sea levels (grey) and moving average smoothed series (window 19 years); rectangular (black) or triangular window (dashed).

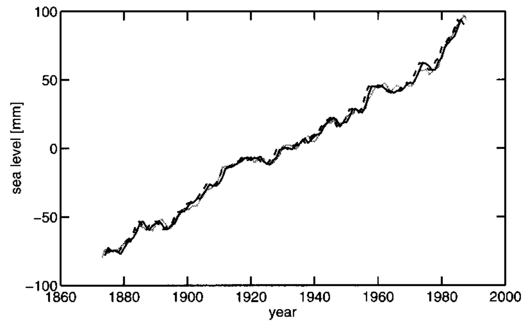

Fig. 3.4 Moving average smoothed annual mean sea levels; rectangular windows: 17, 19, and 21 years.

It can be concluded that moving average smoothing yields smoothed values that, based on visual inspection, seems to reveal the decadal pattern in the data rather well. Different weighting within the windows leads to slightly different smoothed values, but the differences are not very large. Smoothed values do not change significantly if the window size is changed a little.

3.2.3 Exponential smoothing

Contrary to moving average smoothing, exponential smoothing bases the smoothed value only on preceding observations. The most recent observation gets the largest weight, older observations receive exponentially decreasing weights. In its simplest form, so-called single exponential smoothing, the smoothed value at a time t is determined from a combination of the observed value at this time (X_t) and the smoothed value of the preceding observation (S_{t-1}):

$$S_t = \alpha X_t + (1 - \alpha) S_{t-1} \tag{3.3}$$

A problem with exponential smoothing is to find an appropriate value for the smoothing parameter α and an initial smoothed value. According to Makridakis and Wheelwright (1978), the first smoothed value (S_1) could simply be taken equal to the first observed value in the time series. Any deviations caused by this initial value will cancel out after a certain period of time.

Figure 3.5 compares original and exponentially smoothed sea-level data. It is clear that the smaller the value of α, the smoother the resulting series. However, this figure also shows that the smaller the value of α, the longer the smoothed values are influenced by the selected initial value S_1. Since the first value in the time series has a large negative value, the first smoothed values are of lower level than the time series itself would suggest. This "lagging" behind of the smoothed values dissolves rather quickly for larger values of α, but for $\alpha = 0.05$ it continues almost throughout the whole time series.

It can be concluded that exponential smoothing is probably not well suited for smoothing sea-level data. This is due to a number of characteristics of this method. First of all, it is difficult to indicate what value

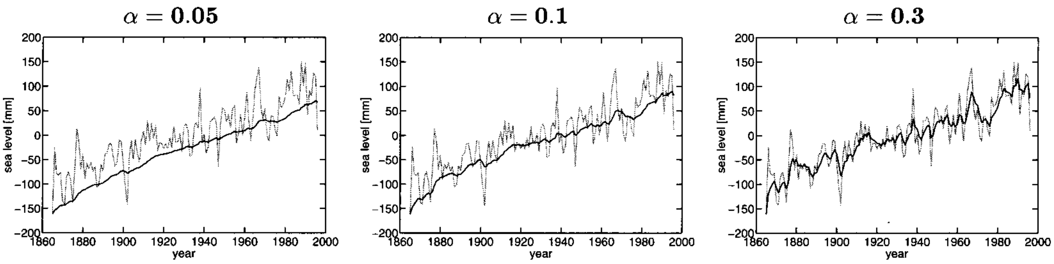

Fig. 3.5 Tide gauge Den Helder. Annual mean sea-level heights (grey) and exponentially smoothed series for 3 different values of the smoothing parameter α.

should be used for the smoothing parameter α. In addition, the smoother the required values, the longer the time series is influenced by the selected initial smoothed value. As a result, the level of the smoothed time series seems to be significantly below that suggested by the sea-level heights themselves.

3.2.4 Singular spectrum analysis

Singular value decomposition can be applied to a data matrix containing time series for a number of locations. The first singular vector (referred to as first principal component) contains the largest possible part of coherent variability present in the data. Given this first principal component, the second component contains the largest possible resulting coherent variability, etc.

One of the disadvantages of singular value decomposition is that it can only be applied to a group of time series. As an extension of this method, Heinen (1992) has developed the so called SSA method (singular spectrum analysis). In this method, each time series is divided into M sub-series with length $N - M + 1$, in which N is the total number of samples in the series, and M is the size of the window. Following Heinen (1992), an example of how the data matrix (X) is formed, based on one time series is:

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_M \\ x_2 & x_3 & \cdots & x_{M+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{N-M+1} & x_{N-M+2} & \cdots & x_N \end{bmatrix} \tag{3.4}$$

As a result, SSA can also be applied to a single time series. How well the characteristics of the time series (trend and periodic components) are separated into different principal components depends on the size of the window (M). This window size should be a compromise between on the one hand having one principal component representing both trend and period fluctuations (small window size), and on the other hand having the trend being spread out over a large number of components (large window size).

This is one of the difficulties with the SSA method. The size of the window influences the principal components, and should be selected separately for every individual time series. There is no fixed relation between the number of samples in the series and an appropriate window size. In practice, the window size has to be selected by comparing results based on varying window sizes.

Figure 3.6 shows the first 12 principal components for tide gauge Den Helder, based on a window size of 20 years (which yields, according to Heinen (1992), the best results). A window size of 20 years implies that the signal can be reconstructed from 20 independent principal components. From figure 3.6 it can be seen that the first principal component follows more or less a linear trend. The second and third component represent relatively long-periodic fluctuations, the next components higher periodic fluctuations, etc. Later components represent mainly noise and have not been depicted.

In figure 3.7 reconstructions based on specific (combinations of) principal components are compared to the actual annual mean sea-level heights. Reconstructed signal in the first figure is based on only the first principal component. The second figure shows a reconstruction based on the first two principal components, etc. As is to be expected, the more principal components are added, the better the reconstructed signal resembles the original data. Based on the required smoothness of the resulting pattern, the number of principal components used to reconstruct the smoothed signal should be selected.

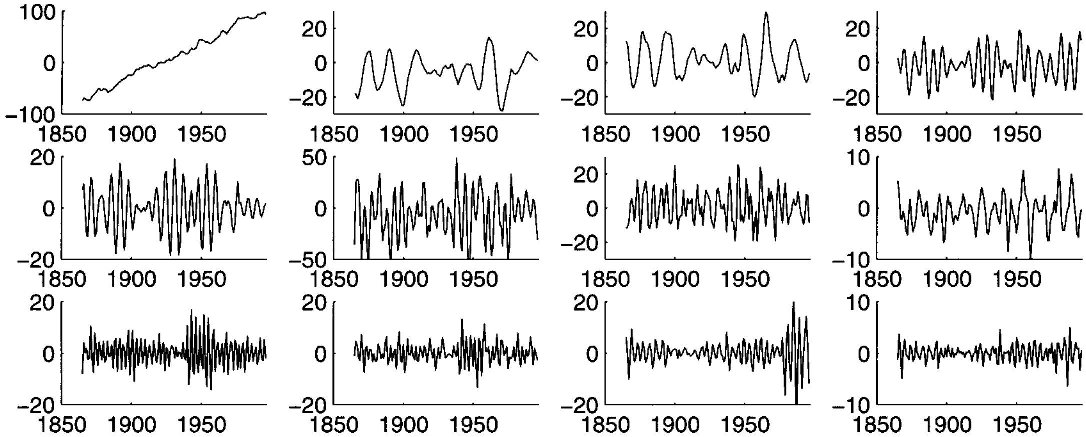

Fig. 3.6 First 12 singular vectors following from SSA (window 20 years) for tide gauge Den Helder.

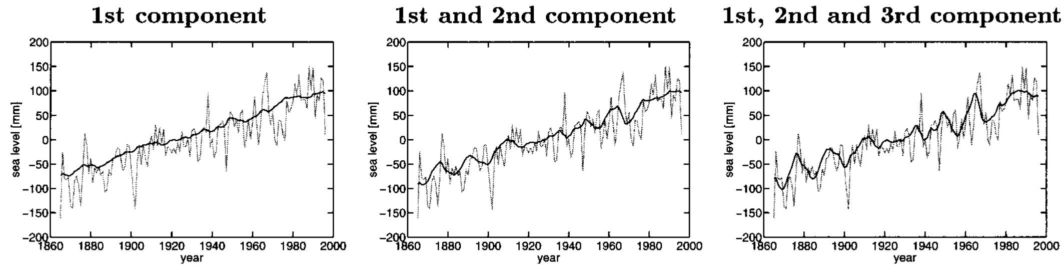

Fig. 3.7 Singular value decomposition for tide gauge Den Helder. Gray: original tide gauge data, black: reconstructions based on (combination of) principal components.

In the preceding, a window size of 20 years has been used. To check whether or not a change in the selected window size significantly influences the results, the first component has been determined based on a number of window sizes. Figure 3.8 shows that a small change in window size, i.e., to either 19 or 21 years, hardly influences the first component. Changing the window size with 5 years, leads to more significant deviations in the first component. However, as can be seen from the right-hand side of figure 3.8, the first principal components for window sizes of 15, 20 and 25 years still show a strong resemblance.

It can be concluded that Singular Spectrum Analysis yields smoothed values that seem to resemble the pattern in the time series. A small change in window size does not significantly change the smoothed values. Even larger changes in window size do not cause substantial changes in smoothed values, e.g., a change in window size from 20 to 15 years causes deviations within 15 mm.

3.2.5 Spectrum analysis

The last category of techniques does not consider the time series themselves, but operates in the spectral domain. A simple example is Fourier analysis in which sine and cosine functions at different frequencies are correlated with the observed time series. The resulting Fourier coefficients can be depicted in an amplitude spectrum, showing the relative significance of periodicity at the various frequencies.

As an example, figure 3.9 shows the amplitude spectrum for tide gauge Den Helder based on annual mean values. Detrended sea level time series have been used, because, otherwise the spectrum would be dominated by an extreme high amplitude at very low frequencies. From figure 3.9 it can be seen that the higher frequencies (up to about period of 5 years/cycle) have relatively little power. Furthermore, significant periodic effects with periods of about 15 and 18 years can be seen.

When examining figure 3.9, it should be kept in mind that this figure is based on annual mean values.

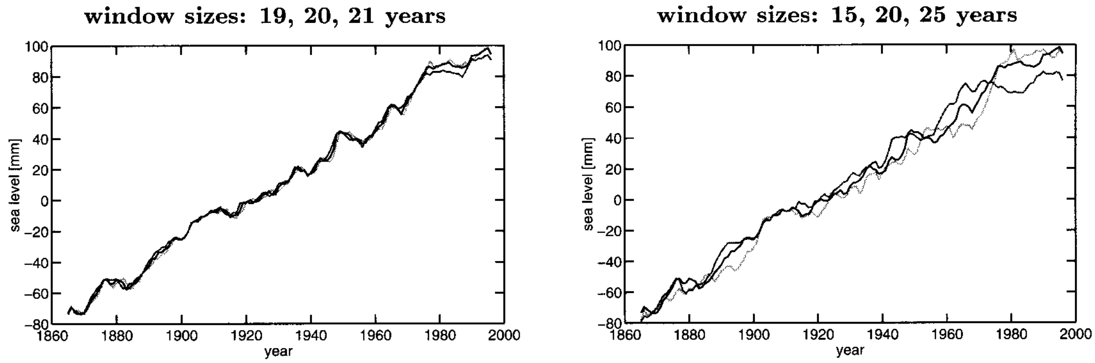

Fig. 3.8 First principal components for annual mean sea-level values for tide gauge Den Helder, based on different window sizes.

Consequently, aliasing can occur if periodic fluctuations with periods up to two years are present in sea level heights, which are not filtered out by forming annual mean sea levels. Using f_s for the sampling frequency and f_0 for the frequency of the periodic oscillation involved, aliasing occurs if $f_s < 2f_0$. In this case, the frequency of the oscillation (f_0) aliases at frequency ($f_s - f_0$). Especially oscillations with periods close to one year alias in relatively longer periods.

Plag (1997) has discussed that the period of the pole tide is not constant in time but changes spatially coherent with time. He identifies three jumps in periods and renames the pole tide "fourteen-to-sixteen month oscillation" (FSO). According to Plag (2000), the most significant oscillations with periods between five and eight years, as depicted in figure 3.9, are basically due to aliasing of the different phases of the FSO as a result of using annual mean sea levels instead of monthly values.

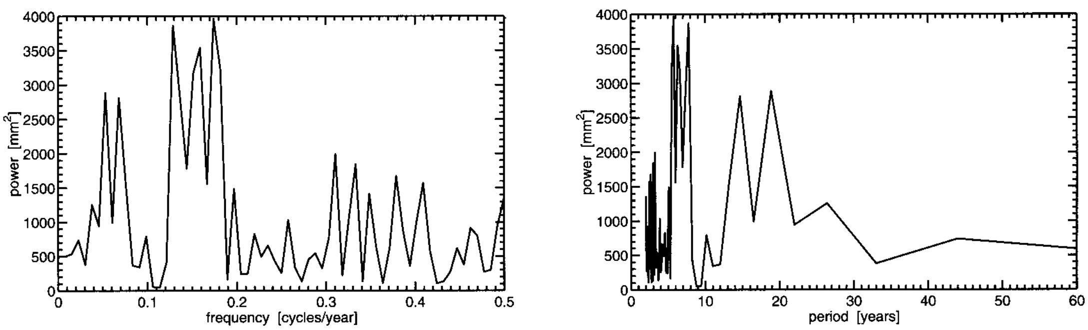

Fig. 3.9 Power spectrum versus frequency (left-hand side) or period (right-hand side) based on annual mean sea levels for tide gauge Den Helder. A best straight-line fit is removed from the data and a Hanning window is applied.

Spectrum analysis can also be used to smooth time series. Low-pass filtering in the frequency domain allows all oscillations with a period smaller than a specific cut-off period to be removed from the signal. However, this method leads to strongly deviating smoothed values if applied to original tide gauge data. This is due to the trend in the sea-level heights. Therefore, in order to obtain more likely results, sea-level heights have to be detrended prior to low-pass filtering. As a result, all figures shown in this section present sea-level variations relative to the linear trend in the data, and are indicated by "detrended sea levels".

An additional problem is the selection of an appropriate low-pass filter. Experiments have been performed with Butterworth and Chebyshev type I filters; see Oppenheim *et al.* (1983). Type I indicates a Chebyshev filter for which the stopband response is maximally flat, and a specific amount of ripple is allowed in the passband. Figure 3.10 shows the effect on the smoothed values of a change in the order of these low-pass filters. No results are shown for the 15th order Chebyshev filter since for this high-order filter numerical problems arise. All Chebyshev filters used are based on 5 dB ripple in the passband.

From figure 3.10 it can be seen that the higher the order of the filter, the smoother the results. In general, results based on Chebyshev low-pass filtering are smoother than those based on Butterworth filtering. From figure 3.10 it is also clear that filtered sea-level heights deviate at most 15 mm from the linear trend in the time series.

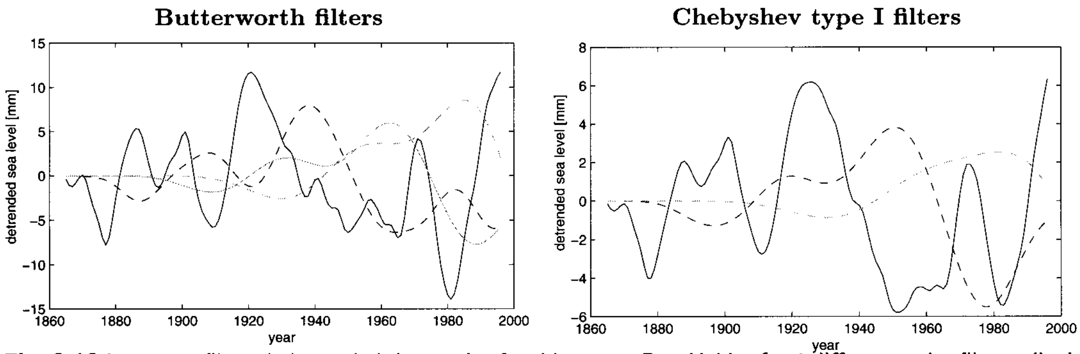

Fig. 3.10 Low-pass filtered *detrended* time series for tide gauge Den Helder for 4 different order filters all with cut-off period 20 years; black: 2nd order, dashed black: 5th order, grey: 10th order, and dashed grey: 15th order.

All low-pass filtered values as shown in figure 3.10 have been derived based on a cut-off period of 20 years. Figure 3.11 shows low-pass filtered, detrended, sea-level heights based on five different cut-off periods. As is to be expected, the higher the cut-off period, the smoother the resulting values. This figure shows that with increasing cut-off period, the pattern in the smoothed values oscillates at a lower frequency. Consequently, the position of the the maximum values shifts towards later years.

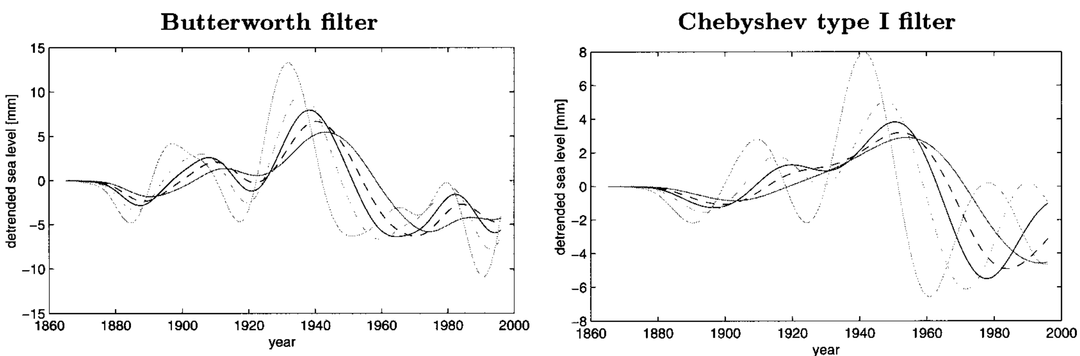

Fig. 3.11 Low-pass filtered *detrended* time series for tide gauge Den Helder with 5th order filter for 5 different cut-off periods; grey: 15 years, dashed grey: 18 years, black: 20 years, dashed black: 22 years, dark grey: 25 years.

It can be concluded that Fourier analysis can be used to deduce at which frequencies periodic fluctuations are present within the data. However, Fourier analysis assumes these fluctuations to be present throughout the time series, with a constant amplitude and frequency. Sea-level data contains periodic fluctuations that are not present (at the same frequency) during the whole measuring period. Fourier analysis will only show that these fluctuations are present, but can give no insight in when they occur.

Low-pass filtering can be used to remove higher frequency oscillations from the data. However, this method can only be used on detrended data. The filtered values deviate only a little from the linear trend in the data. In addition, how these smoothed values fluctuate around the linear trend line is highly influenced by the choice of low-pass filter. It can be concluded that Fourier analysis is not a very suitable tool for smoothing sea-level height time series, but can be used to give an indication of the periodic fluctuations significantly present in the data set.

3.2.6 Smoothing based on wavelets

Some of the processes generating sea-level variations show significant fluctuations in amplitude and period. For example, although variations in atmospheric pressure show periodic patterns, for specific periods of time they may be absent or completely changed. Variations in frequency content require a technique that can not only give an indication of the frequencies present in the data, but also of when these frequencies occur. This information can be obtained from a number of techniques, among which wavelet analysis. For a treatment of the theory of wavelet analysis, see, e.g., Strang and Nguyen (1996).

Wavelet analysis can be explained as follows. By correlating a time series with sine and cosine functions at different frequencies, periodicities at the corresponding frequencies can be detected. The sine and cosine functions used are constant in period and amplitude throughout time. Instead of sine and cosine functions, wavelet analysis correlates the time series against wavelets. A wavelet being a waveform of limited duration with an average value of zero. The wavelets can be scaled, i.e., spread out over a larger period of time, or shifted in time. This implies that the maximum values of the wavelet occur at a different moment. The time series is now correlated against different scaled versions of the wavelet, to which an increasing time shift is applied.

Wavelets can be used to analyse data in various ways, but in this chapter wavelet analysis will only be used to decompose a time series in constituting data sets of decreasing smoothness. For the theory involved, see Strang and Nguyen (1996). Decomposition can be described as splitting the signal in a high-scale, low frequency part (the approximation) and a low-scale, high-frequency part (the details). Next the low frequency part is again split into an approximation and a detail part. This is repeated until the desired smoothing level is reached.

Figure 3.12 shows a level 3 decomposition of annual mean sea levels. A specific wavelet, coiflet3, has been used; see Daubechies (1992) for information on the construction of this type of wavelet. Level 3 decomposition consists of a level 3 approximation signal, and three detail signals. The original time series can be reconstructed by adding these four signals. The approximation signal shows a trend in combination with occasional occurrence of (long-periodic) fluctuations. The detail 3 signal shows no trend, but only relatively long-periodic fluctuations. The level 2 detail contains fluctuations with relatively higher frequencies, whereas the shortest period events show up in the detail 1 signal.

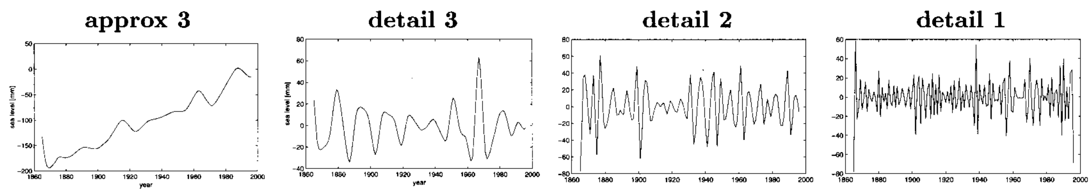

Fig. 3.12 Signal decomposition at level 3, based on coiflet3. Annual sea-level heights for tide gauge Den Helder.

It seems that wavelet analysis is an adequate technique to separate the underlying pattern in a time series from higher periodic fluctuations and noise. However, the results derived are highly dependent on choices made during the wavelet analysis. First of all, since wavelet analysis is based on correlating the time series with wavelets on various scales to which various time shifts have been applied, results are influenced by the choice of wavelet. Based on the characteristics of the time series involved, a wavelet should be selected. However, for sea-level height data no specific choice of wavelet is evident.

In figure 3.13 the level 3 approximation based on coiflet3 is compared with level 3 approximations based on two other types of wavelets. The two other wavelets seem equally suitable for representing the sea-level height series. Although there seems to be a reasonable similarity between the results in some parts, major differences occur as well. For example, level 3 approximation based on coiflet3 shows relatively few periodic fluctuations, whereas results based on symlets3 are less smooth.

Results are also influenced by the selected level of decomposition. The higher the level of decomposition, the smoother the approximated signal; see figure 3.14 in which approximated signals are shown for three increasing levels of decomposition. It should be noted that an approximation corresponding to a lower level of decomposition can be reconstructed by adding detail signals to a higher level approximation.

The achieved decomposition is determined by the selection of a wavelet family, the smoothness of the

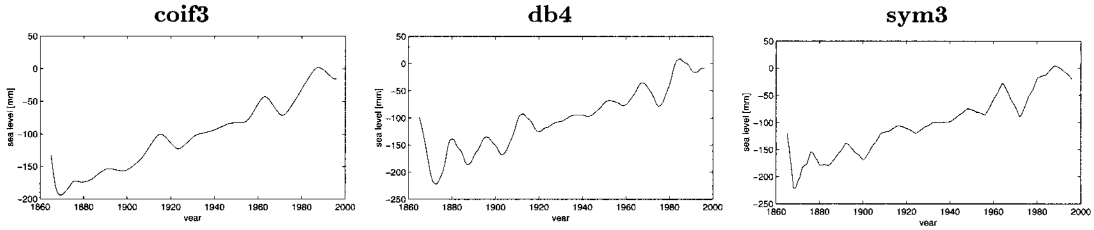

Fig. 3.13 Level 3 approximation (decomposition at level 3) based on coiflets, symlets or Daubechies wavelets.

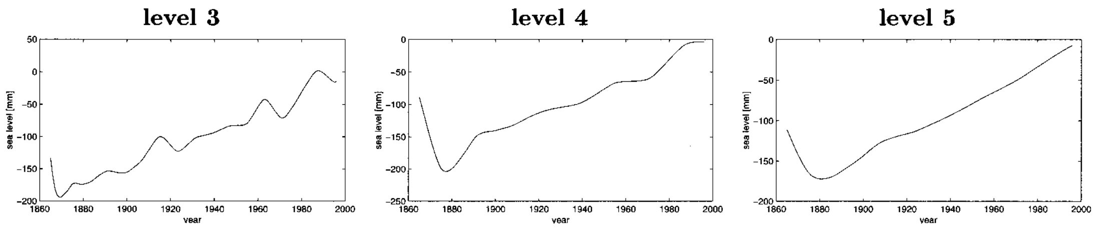

Fig. 3.14 Wavelet decomposition (approximated signals) based on coiflet3 for 3 different decomposition levels.

wavelet (e.g., selecting db2, or db8 wavelets), and the level used for the decomposition. A small change in one of these choices, e.g., decomposition at level 4 instead of level 3, can lead to significant changes in the resulting decomposition. On the other hand, results are still similar enough that more than one choice of a combination of wavelet plus decomposition level seems suitable for representing a trend underlying the data. Consequently, it might be rather difficult to, objectively, decide which wavelet and decomposition should be used, while this choice has a significant influence on the results.

An additional problem is that the wavelet and decomposition level required to yield a sufficient smooth approximation (with enough resemblance to the original data) depends on the number of samples in the time series. For example, if instead of annual sea-level heights mean monthly sea-level values are used, level 3 decomposition does not produce smooth results. A much higher level of decomposition, e.g., level 8 decomposition, is required to produce relatively smooth results.

3.3 Prediction of future sea levels

As indicated in section 3.1, an accurate prediction of future sea-level heights requires information on expected changes from the present-day situation, e.g., based on a GCM. This is beyond the scope of this thesis, therefore, in this section, only extrapolations of sea level height time series will be considered. A wide variety of analysing techniques is available for forecasting future values based on an existing time series. Only a limited number of these techniques will be examined in this chapter. Depending on the fundamentals on which these techniques are based, they can be divided into three categories:

- Fitting of a specific curve to the measurements.
  All techniques considered are regression techniques. A specific order of polynomial is assumed. Coefficients are estimated that represent a curve which yields the best least-squares fit to the data available. With these coefficients, the regression model can be used to estimate values for every future moment in time.
- Techniques based on averaging measurements.
  Measurements have equal weights (moving average methods), or exponentially decreasing weights (exponential smoothing methods). Future values are based on:
  - smoothed average values
    examples: moving average prediction and single exponential smoothing
  - smoothed average values with (smoothed) trends
    examples: linear moving average prediction, Brown's one-parameter exponential smoothing and Holt's two-parameter exponential smoothing

  - smoothed average values with smoothed trend and seasonality
    example: Winter's linear and seasonal exponential smoothing
- Fitting of a curve of *averaged* measurements to the data.
  AR(I)MA modelling will be discussed, which is based on selecting a specific combination of autoregressive values of the signal at previous times and moving average values of white noise at present and preceding times. The coefficients for these autoregressive and moving average parameters are estimated as those which yield a model that best fits the data available. Based on the selected number of autoregressive and moving average parameters, and the values estimated for their coefficients, future values for the time series can be forecasted.

Analysing techniques will be assessed based on how well they are able to forecast the decadal pattern in sea-level variation. Since the "true" future sea-level variation pattern is not known, the methods will be judged based on the following characteristics. Only part of the available data is used to estimate the model parameters. For the remaining period, future values are predicted and compared to the actual data. Techniques are judged on how well the predicted values follow the actual data. It is also tested how much forecasted values vary if the amount of sea level data on which the model estimation is based is changed.

The various forecasting techniques will be applied to both original and smoothed annual mean sea-level heights for tide gauge Den Helder. As smoothed data results from either moving average smoothing or singular value decomposition will be used.

3.3.1 Linear regression

Linear regression implies that a line $\hat{y}_i = a + bx_i$ is fitted through the data. The regression coefficients a (intercept) and b (slope) have to be selected as to minimise the squared difference between the estimated trend line and the original data series. With these two regression coefficients, future data points can be estimated for every moment in time.

An advantage of linear regression is that it is simple to apply. On the other hand, linear regression is very sensitive to (large) outliers in the data. Consequently, in this section, linear regression will only be applied to smoothed sea-level height series.

First, linear regression is used to fit a trend line trough moving average smoothed annual mean sea levels. Results are shown in figure 3.15, for both a rectangular window and a triangular smoothing window. If the fitted trends are used to predict a sea-level height value for the year 2050, a value of resp. 59.9 mm (rectangular window) or 59.5 mm (triangular window) is found.

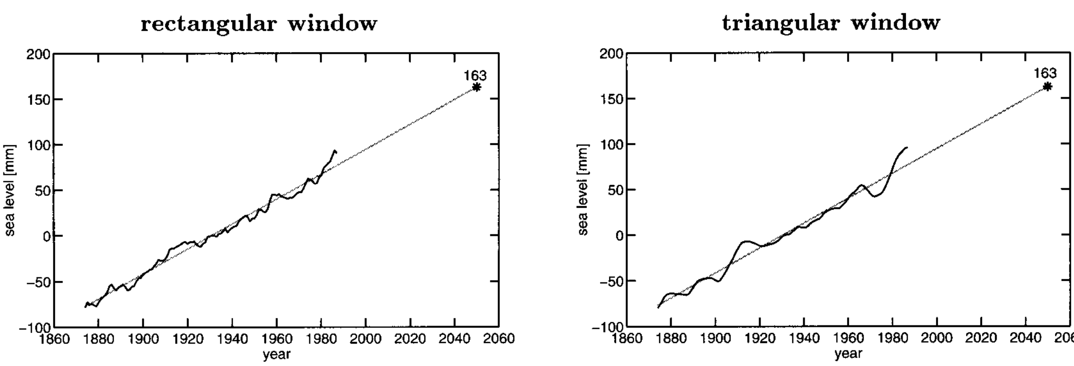

Fig. 3.15 Moving average smoothed values (window: 19 years), linear trend fitted through these smoothed values (grey), and predicted sea level for 2050.

From figure 3.15 it can be concluded that for predicted values for the year 2050 it makes no difference whether the rectangular or the triangular window is used to smooth the original time series. This is confirmed by table 3.1, in which regression coefficients and predicted future sea levels have been summarised for a number of smoothing techniques. This table also shows that a small change in window size, e.g., to 17 or 21 years, hardly influences the values determined for the intercept and slope of the fitted trend line, and, consequently the values predicted for the sea-level height in the year 2050.

To test how much results are influenced by the number of elements in the time series, the first 30 years of data have been omitted prior to moving average smoothing the data. As can be seen from table 3.1, this has only a very small influence on the trend fitted through these smoothed values. However, the standard deviations with which this trend can be determined slightly deteriorate.

Smoothing method	intercept in mm	slope mm/yr	predicted value for 2050	smoothed value in 1996
no smoothing	-2726 ± 3.1	1.41 ± 0.08	168.8 mm	
moving average				
rectangular window 19 yrs	-2637 ± 0.5	1.37 ± 0.02	163.0 mm	90.3 mm
triangular window 19 yrs	-2627 ± 0.7	1.36 ± 0.02	162.7 mm	95.9 mm
rectangular window 17 yrs	-2651 ± 0.6	1.37 ± 0.02	164.1 mm	93.8 mm
triangular window 17 yrs	-2647 ± 0.8	1.37 ± 0.02	164.1 mm	98.2 mm
rectangular window 21 yrs	-2624 ± 0.5	1.36 ± 0.01	162.0 mm	85.4 mm
triangular window 21 yrs	-2610 ± 0.6	1.35 ± 0.02	161.4 mm	93.4 mm
moving average first 30 yr omitted				
rectangular window 19 yrs	-2557 ± 0.7	1.33 ± 0.03	159.2 mm	90.3 mm
triangular window 19 yrs	-2554 ± 0.9	1.32 ± 0.04	159.3 mm	95.9 mm
SSA				
window size 18 yrs	-2569 ± 0.5	1.34 ± 0.01	171.8 mm	96.2 mm
window size 20 yrs	-2545 ± 0.5	1.32 ± 0.01	171.3 mm	94.2 mm
window size 22 yrs	-2487 ± 0.4	1.30 ± 0.01	167.9 mm	87.0 mm

Table 3.1 Linear regression coefficients and predicted sea-level values based on original and smoothed data for tide gauge Den Helder.

Next, a linear trend is fitted through the first principal component derived with SSA. Results, for three different window sizes, are shown in figure 3.16 and table 3.1. It can be concluded that a, small, change in window size does not have a significant influence on the predicted value for the year 2050.

Comparing the results based on SSA with those derived for moving average smoothing shows that trend estimates are slightly smaller (but only of the order of 0.05 mm/yr), and values estimated for the intercepts are somewhat larger. As a result, predicted sea levels for 2050 based on SSA smoothed data are larger (of the order of 5 to 10 mm) than results based on moving average smoothed data. It can be concluded that results based on these two smoothing techniques (and using different windows) show a large similarity. From table 3.1 it can also be seen that smoothing the data prior to regression yields a significant increase in precision with which the regression coefficients can be determined.

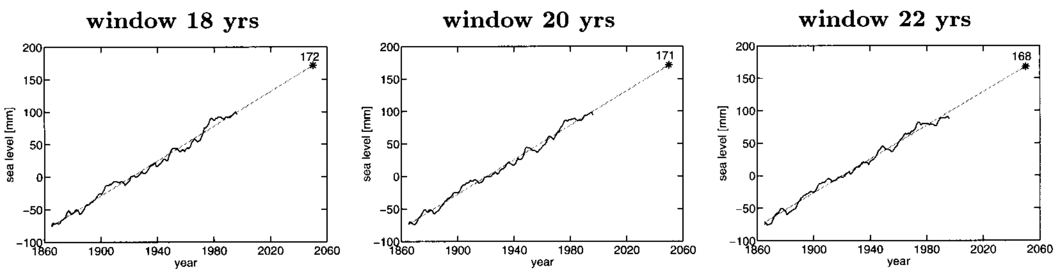

Fig. 3.16 First component based on SSA, linear trend fitted through these smoothed values (grey), and predicted sea level for 2050.

As a last test, it is checked how much forecasted values change if the year in which the forecasts are started is changed. As an example, experiments have been performed using the first principal component (window: 20 years) for tide gauge Den Helder. If linear trends are determined based on data up to resp.

1967 or 1976, results hardly differ from those based on the complete time series up to 1996. The difference between the two trend lines is very small and they follow the principal component rather well. Therefore, it can be can be concluded that linear regression, applied to smoothed tide gauge data, seems robust against a change in starting year for the forecasts. The years 1967 and 1976 have been selected because, for a number of techniques discussed in the following sections, this yields the largest deviations between the forecasts.

It can be concluded that forecasts based on linear regression applied to smoothed sea-level data are rather consistent. Moving average smoothing and SSA modelling yield forecasted sea-level heights for the year 2050 at most 1 cm apart. A small change in window (size or weights) hardly changes the forecasted value either. In addition, if applied to smoothed data, linear regression seems to be robust against a change in starting year for the forecasts.

3.3.2 Other forms of regression

Instead of a linear relation between observed sea-level heights and time, other relationships might exist. Using regression analysis, the parameters defining these relationships can be estimated. However, a value for these parameters can be determined, whether or not the relationship is actually present in the data. Therefore, beforehand some insight is needed in the kind of relation between sea-level heights and time.

A simple test to check if a specific relationship exists is to plot the residuals between actual data and predictions based on the relationship. If a specific relationship is consistent with the data, the residuals should not show a systematic pattern. In figure 3.17 residuals are shown for a linear trend fitted through smoothed data following from SSA. They show some periodic behaviour, but no remaining higher order pattern (e.g., a parabolic curve).

Another indication of how well the estimated pattern fits the data is the so-called determination coefficient (r^2). As explained by Makridakis and Wheelwright (1978), this is the ratio between the total variation explained by the regression line and the total variation in the original data. The closer to 1 the value for this coefficient, the better the regression line explains the variance in the data. As an example, for the linear trend line fitted through the first principal component (window 20 years) for tide gauge Den Helder, the value found for $r^2 = 0.990$, indicating a good fit.

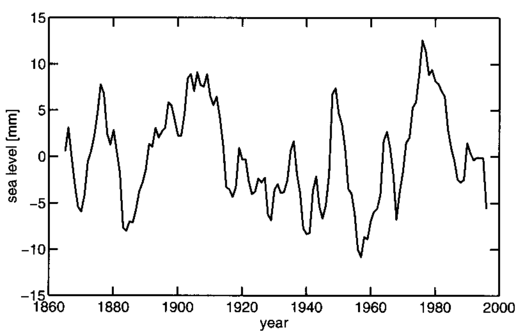

Fig. 3.17 Difference between first component based on SSA (window 20 years) and linear trend fitted through these smoothed values.

Experiments show that if an exponential fit is applied instead of a linear one, this leads to very unrealistic results. If instead of linear regression, a second, third or fourth order polynomial is fit through the smoothed data, results are almost indistinguishable from those based on a first order polynomial. As an example, it is assumed that a quadratic curve should only be used to describe sea-level variation if the estimated quadratic regression coefficient ($\hat{c}$) is significant in relation to its standard deviation ($\hat{\sigma}_{\hat{c}}$); see chapter 4. In this case, the hypothesis that sea-level variation is linear ($H_0 : E\{\hat{c}\} = 0$) can be tested against the hypothesis that the variation is quadratic ($H_1 : E\{\hat{c}\} \neq 0$). The test statistic

$$t = \frac{\hat{c}}{\hat{\sigma}_{\hat{c}}} \tag{3.5}$$

can be used, which should follow a Student's distribution with $N-2$ degrees of freedom (under H_0) and N the number of observations. With a confidence level of 95% and $N = 132$, $\hat{t}_{N-2} \approx 1.98$. If quadratic regression is applied to SSA smoothed data for tide gauge Den Helder, it is found that $\hat{c}/\hat{\sigma}_{\hat{c}} = 1.638$, which implies that the quadratic term is not statistically significant.

It can be concluded that, at least for the data set used in this section, simple linear regression seems to give satisfactory results, which do not seem to improve with the inclusion of higher order polynomial terms. An exponential fit instead of a normal fit deteriorates the results.

3.3.3 Moving average prediction

Whereas moving average smoothing assigns the value averaged over the window to the central sample in the window, moving average prediction allocates this value to the first sample following the samples in the window. Usually, weights are not applied and every sample in the window contributes the same to the resulting moving average value. For a time series (X) consisting of N samples, the value at a specific time $t+1$ (with $t \leq N$) is forecasted as:

$$\hat{X}_{t+1} = \frac{1}{M} \sum_{i=t-M+1}^{t} X_i \tag{3.6}$$

in which M is the number of samples within the window. For moving average prediction, a window size should be selected that minimises the squared difference between the forecasted value and the actual value of these time series; see Makridakis and Hibon (1984). For annual mean sea-level heights for tide gauge Den Helder this yields a window size of 17 years.

After a suitable value for the window size is chosen, future values for the time series (at times k after the last known value N) can be predicted using

$$\tilde{X}_{N+k} = \frac{1}{M} \sum_{i=N-M+k}^{N-1+k} X_i \tag{3.7}$$

replacing X_i by $\tilde{X}_i$ in the right-hand side of equation (3.7) if the subscript exceeds N.

As an example, the left-hand side of figure 3.18 shows estimated values based on moving average prediction (window 17 years). The extrapolated values are based on only the first 112 years of data. Compared to the actual values for the period between 1977 and 1997, predicted values are systematically low. The further away from 1977, the larger the deviation between the predicted and actual values. If this method is used to make predictions for an even larger amount of time in advance, results are even worse; see right-hand side of figure 3.18. This is one of the disadvantages of using moving averages for predicting future values. If the time series contains a trend, moving average prediction systematically underestimates the actual values.

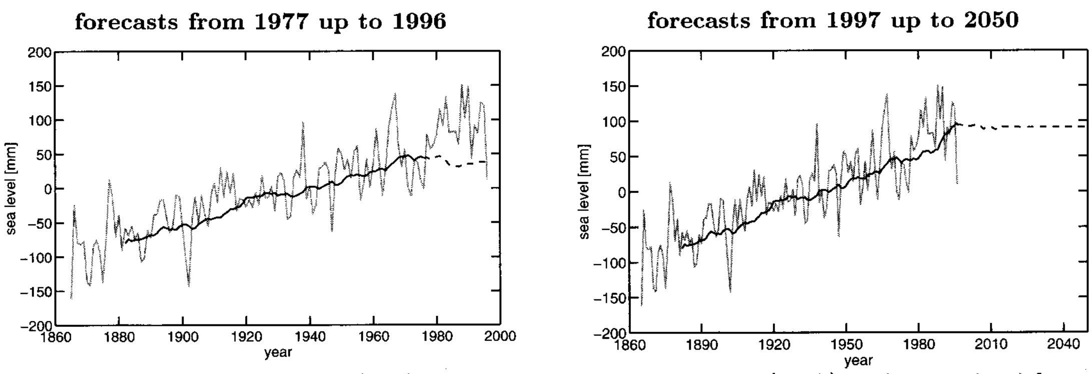

Fig. 3.18 Annual mean sea levels (grey), moving average predicted values (black), and extrapolated future values (dashed); window size: 17 years.

3.3.4 Linear moving average prediction

If a trend is present in the data, moving average predictions lag behind the actual measured values. As described by Makridakis and Wheelwright (1978), this can be corrected for by taking moving averages of these moving average values, since this second moving average gives an estimate of how much the original moving averages lag behind the actual values. Following Makridakis and Hibon (1984), the method of linear moving average prediction is described by a set of equations. The first equation gives the simple moving average at time t:

$$S_t' = \frac{1}{M} \sum_{i=t-M+1}^{t} X_i \tag{3.8}$$

with M the number of samples within the window. Next, the moving average of equation (3.8) is taken:

$$S_t'' = \frac{1}{M} \sum_{i=t-M+1}^{t} S_i' \tag{3.9}$$

The intercept (a_t) and slope (b_t) for time t can be determined from these moving averages as:

$$\begin{aligned} a_t &= 2S_t' - S_t'' \\ b_t &= \frac{2}{M-1}(S_t' - S_t''). \end{aligned} \tag{3.10}$$

Using these coefficients, the predicted value for time $t+1$ is found as:

$$\hat{X}_{t+1} = a_t + b_t \tag{3.11}$$

and future values (at times k after the last known value N) can be forecasted as:

$$\tilde{X}_{N+k} = \hat{a}_N + k \cdot \hat{b}_N \tag{3.12}$$

Figure 3.19 shows annual mean sea-level values for tide gauge Den Helder and moving average smoothed values for the first 112 years of data. In addition, forecasted sea-level heights starting from 1977 onwards are given. Comparison with the left-hand side of figure 3.18 shows that predictions based on linear moving averages are significantly better than those based on simple moving averages. Instead of flattening out, predicted values show a trend.

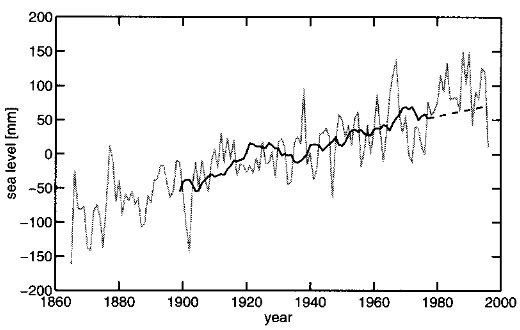

Fig. 3.19 Gray: annual mean sea levels, black: linear moving average values (up to 1976), dashed: extrapolated future values. Window: 17 yrs.

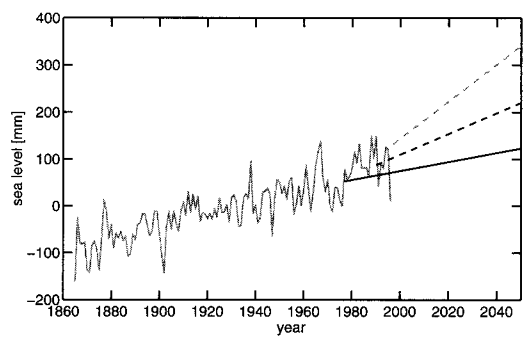

Fig. 3.20 Annual mean sea levels (grey) and linear moving average values; starting in 1977 (black), in 1990 (black dashed), or in 1997 (grey dashed).

The results shown in figure 3.19 look rather promising. However, forecasts can drastically change if a different starting year for the forecasts is used, i.e., slope and intercept values are used based on a different number of observations. As an example, figure 3.20 shows extrapolated future annual sea-level heights based on three different starting years for the predictions. It can be seen that, in this case, the further in time the predictions are started, the larger the value of the trend and intercept.

The above example shows that linear moving average prediction does not yield reliable estimates. This can be explained as follows. Forecasts are based on linear regression beyond the last observation available, using values for the slope and intercept as determined for this last observation. In order for this method to give reliable results, the values for intercept and slope should not vary to much over time. Unfortunately, even for smoothed sea-level heights, a large variation in these regression parameters is found. As a result, forecasts based on linear moving average prediction are highly influenced by the selection of the year in which the forecasts are started.

3.3.5 Single exponential smoothing

Exponential smoothing methods assign exponentially decreasing weights to older measurements. In this section, results for the least complex method, called single exponential smoothing, will be discussed. Following, e.g., Makridakis and Wheelwright (1978), for $t \leq N$, with N the total number of observations in the time series, the predicted value at time $t+1$ can be determined as:

$$\hat{X}_{t+1} = \alpha X_t + (1-\alpha)\hat{X}_t \tag{3.13}$$

in which X_t is the observed value, $\hat{X}_t$ the predicted value at time t, and α is the smoothing parameter that needs to be estimated.

For each individual time series, a value for α is determined as that value that yields the smallest mean squared difference between the predicted and observed values. The first predicted value ($\hat{X}_1$) is simply taken equal to the first observed value in the time series. For six tide gauge data sets along the Dutch coast, values for α are estimated ranging between 0.24 and 0.63.

As described by Makridakis and Hibon (1984), future values for the time series (at times k after the last known value N) are forecasted as:

$$\tilde{X}_{N+k} = \hat{\alpha} X_N + (1-\hat{\alpha})\tilde{X}_{N+k-1} \tag{3.14}$$

Single exponential smoothing has been applied to the first 102 years of data for tide gauge Den Helder. Results are shown in the left-hand side of figure 3.21. It can clearly be seen that although the smoothed values, prior to 1976, follow the actual values rather well, the predicted values deviate strongly. The further away from the year 1976, the flatter the line of predicted values becomes. In addition, if predictions are started in 1968, a completely different set of predicted values is found.

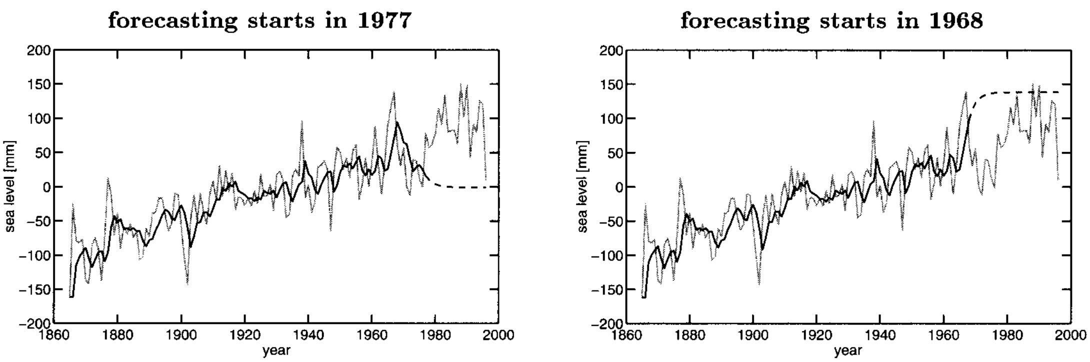

Fig. 3.21 Tide gauge Den Helder; grey: original mean annual sea levels, black: exponential smoothed values, dashed: forecasted values. Forecasting starts in 1977 ($\alpha = 0.34$), or in 1968 ($\alpha = 0.37$).

The above shows that single exponential smoothing is not suited for forecasting sea-level heights. As can be seen from equation (3.14), forecasted values are all based on the last observation. As a result, single exponential smoothing, analogous to moving average prediction, is only suited for stationary signals, i.e., time series that do not contain a trend.

3.3.6 Linear exponential smoothing

Two exponential smoothing methods will be discussed that have been developed for processing time series containing a trend. The first method, Brown's one-parameter exponential smoothing, is based on double averaging the measurements. In the second method, Holt's two-parameter linear exponential smoothing, the trend is included as an additional value that needs to be smoothed.

Brown's one-parameter linear exponential smoothing

Analogous to linear moving average prediction, the lag between the smoothed and actual values is estimated based on the difference between single and double smoothed observations. Following Makridakis and Wheelwright (1978), for times $t \leq N$, with N the total number of observations in the time series, the first smoothed values are determined from

$$S'_t = \alpha X_t + (1-\alpha)S'_{t-1} \tag{3.15}$$

and the double exponential smoothed values from

$$S''_t = \alpha S'_t + (1-\alpha)S''_{t-1} \tag{3.16}$$

From these two smoothed series, the intercept (a_t) and slope (b_t) can be estimated as:

$$\begin{aligned} a_t &= 2S'_t - S''_t \\ b_t &= \frac{\alpha}{1-\alpha}(S'_t - S''_t) \end{aligned} \tag{3.17}$$

Based on these coefficients, values for time $t+1$ can be predicted as:

$$\hat{X}_{t+1} = a_t + b_t \tag{3.18}$$

Analogous to single exponential smoothing, an appropriate value for α is determined as that value that minimises the squared difference between the predicted and actual observations. For the six tide gauges along the Dutch coast, estimated values for α vary between 0.05 and 0.23. They are significantly smaller than those determined for single exponential smoothing, implying that a relatively larger amount of preceding observations will contribute substantially to the prediction.

Based on the intercept and slope values determined for the last observation in the time series (N), future values (at times k after the last known value N) are forecasted as:

$$\tilde{X}_{N+k} = \hat{a}_N + k \cdot \hat{b}_N \tag{3.19}$$

If Brown's method is used to forecast annual sea levels for tide gauge Den Helder, results are highly influenced by the year selected to start the predictions. As can be seen from figure 3.22, forecasted values based on two different starting years deviate strongly. This is caused by the fact that the intercept and slope of the forecasted values are taken equal to the intercept and slope corresponding to the last observation used. Since the time series contain large periodic fluctuations, a different staring year for the forecasts will yield a different value for the slope and intercept. Smoothing of the time series removes the short-periodic fluctuations from the data sets. However, the long-periodic phenomenon will still be present. Consequently, smoothing of the time series prior to analysis does not yield more reliable forecasts.

Holt's two-parameter linear exponential smoothing

Instead of "double" smoothing the observations, the trend is included as an additional value that is smoothed. As a result, this method is more flexible, but requires an additional smoothing parameter for this trend value to be estimated based on an iterative process. Unfortunately, predictions of future values are, again, based on the value of the slope determined for the last observation. Consequently, results will not be very reliable if this slope varies significantly over time, e.g., because the time series considered contains large periodic fluctuations. Therefore, results based on this method will not be shown.

3.3.7 Linear and seasonal exponential smoothing

An adaptation to linear exponential smoothing that can handle periodic data is Winter's linear and seasonal exponential smoothing. This method is described in detail by Makridakis and Wheelwright

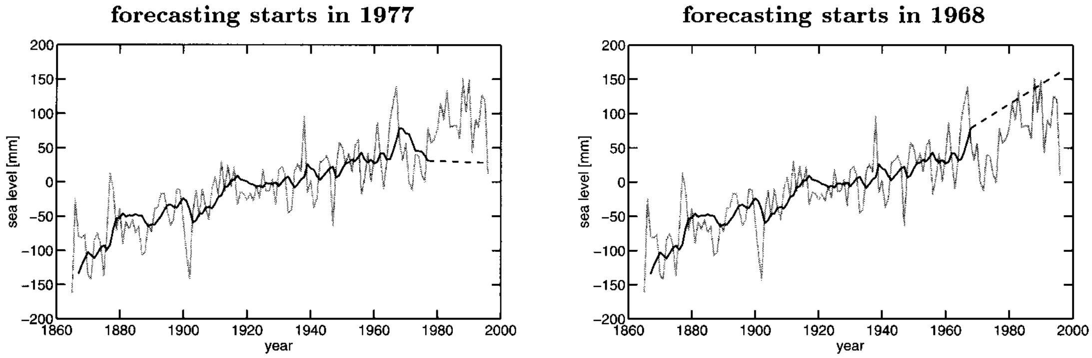

Fig. 3.22 Tide gauge Den Helder; grey: original mean annual sea levels, black: Brown's one-parameter linear exponential smoothed values ($\alpha = 0.1$), dashed: forecasted values. Forecasting starts in 1977 or in 1968.

(1978), and basically involves three different "smoothing equations", one to smooth the seasonality, one to smooth the trend, and one to smooth the observations. First the length of seasonality (L) has to be determined. In case of mean sea-level height time series, a period of 19 years could be used since this includes one full cycle of the lunar nodal tide. The smoothed observations (S_t) can be determined from:

$$S_t = \alpha \frac{X_t}{P_{t-L}} + (1-\alpha)(S_{t-1} + T_{t-1}) \tag{3.20}$$

This equation contains the trend determined for the previous observation (T_{t-1}). This trend is smoothed with a separate parameter β:

$$T_t = \beta(S_t - S_{t-1}) + (1-\beta)T_{t-1} \tag{3.21}$$

In addition, equation (3.20) contains the seasonal adjustment factor (P_{t-L}) for the corresponding data point in the previous period. To smooth randomness, this seasonal adjustment factor is smoothed with a third parameter γ:

$$P_t = \gamma \frac{X_t}{S_t} + (1-\gamma)P_{t-L} \tag{3.22}$$

For times $t \leq N$, with N the total number of observations, sea-level heights are predicted with:

$$\hat{X}_{t+1} = (S_t + T_t)P_{t-L+1} \tag{3.23}$$

Future values for the time series (at times k after the last known value N) can be forecasted as:

$$\tilde{X}_{N+k} = (\hat{S}_N + k \cdot \hat{T}_N)\hat{P}_{N-L+k} \tag{3.24}$$

One of the problems related to Winter's method is to find appropriate values for the three smoothing parameters and initial values for the seasonal adjustment factors. That combination of values for α, β, and γ should be used that minimises the total squared difference between the predicted and observed values. For tide gauge Den Helder the values $\alpha = 0.08$, $\beta = 0.02$, and $\gamma = 0.01$ are found.

A relative large number of observations is available. Therefore, it can be assumed that problems caused by the selection of initial values will have cancelled out long before the forecasting is started. As a result, the following simple method can be applied to estimate initial values for the first 19 (the period assumed for seasonality) seasonal adjustment factors. A value 1 is assumed for all P_{t-L}; substitution of this value in equations (3.20), (3.21), and (3.22) yields the 19 initial values for S_t, T_t, and P_t respectively.

For tide gauge Den Helder, small values are determined for the smoothing parameters, implying that a very large amount of previous values contribute to the smoothed observation, trend, and seasonal adjustment factor. This can be seen from figure 3.23 in which Winter's linear and seasonal exponentially smoothed values are shown. The predicted values follow the observations rather well, the curve for the predicted values being very smooth, i.e., most periodic fluctuations have been smoothed out. Forecasted sea levels up to 2050 are shown in the right-hand side of figure 3.23. These forecasted values look rather consistent with the available observed data.

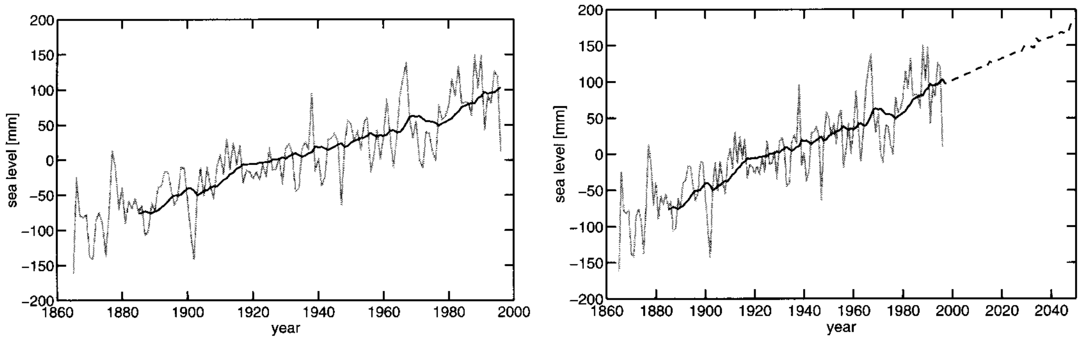

Fig. 3.23 Tide gauge Den Helder; grey: original annual mean sea levels, black: Winter's linear and seasonal exponential smoothed values ($\alpha = 0.08$, $\beta = 0.02$, $\gamma = 0.01$), dashed: forecasted values.

Results in figure 3.23 have been derived using a period of 19 years (> one full cycle of the lunar nodal tide) for the length of seasonality. Experiments with other lengths of seasonality show that Winter's linear and seasonal exponential smoothing seems robust in relation to changes in period of seasonality. Unfortunately, a change in starting year for the forecasts can cause a major change in forecasted values.

As an example, in figure 3.24 results are shown based on two different starting years for the forecasting. This figure shows that for both starting years the forecasts seem to follow the actual data rather well. However, there is a significant difference between the 2 sets of forecasts. This is caused by the fact that Winter's method bases the forecasted values on the smoothed value and smoothed trend value determined for the last observation. Despite smoothing of the different parameters, these values still show major fluctuations with time.

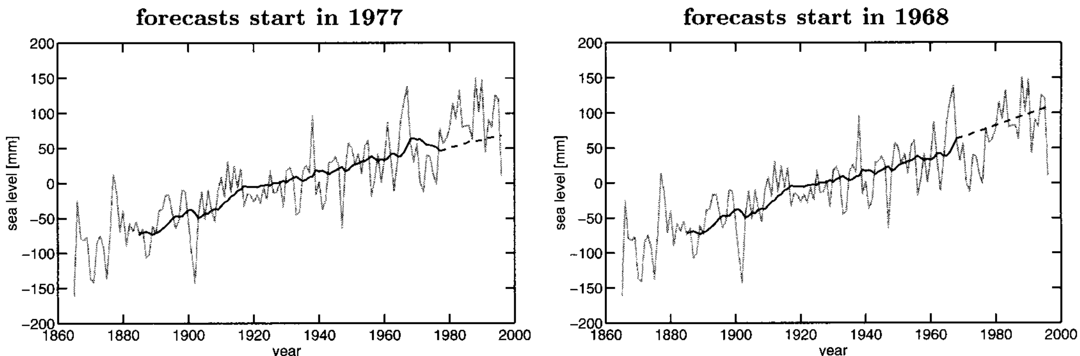

Fig. 3.24 Gray: original data, black: Winter's linear and seasonal exponential smoothing predictions, dashed: forecasted values. Forecasting starts in 1977 (α 0.09, β 0.02, γ 0.02) or in 1968 (α 0.09, β 0.02, γ 0.01).

3.3.8 AR(I)MA modelling

AR(I)MA is the acronym for autoregressive (integrated) moving average, a technique described in detail by Box and Jenkins (1976). AR(I)MA modelling is based on a specific number of autoregressive and moving average terms that are fitted to the data in order to yield minimal total squared differences between modelled values and actual data. Based on the values determined for the ARMA coefficients, future values can be forecasted. For a description of the different types of models and the mathematics involved, see Box and Jenkins (1976). In this section, only some characteristics of applying AR(I)MA modelling to sea-level data will be described and some results based on annual mean sea-level heights for tide gauge Den Helder will be shown.

For a stationary signal, it is assumed that the signal at time t can be written as a combination of an autoregressive process and a moving average process:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \cdots - b_q \varepsilon_{t-q} \qquad (3.25)$$

The autoregressive process consists of a combination of values of the signal at previous moments in time:

$$y_t = a_1 y_{t-1} + a_2 y_{t-2} + \cdots + a_p y_{t-p} + \varepsilon_t \tag{3.26}$$

in which ε_t is a white noise process, i.e., a sequence of uncorrelated random variables with zero mean and constant variance. The moving average process is a combination of this white noise at time t and at preceding times:

$$y_t = \varepsilon_t - b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} - \cdots - b_q \varepsilon_{t-q} \tag{3.27}$$

It is assumed that p autoregressive and q moving average coefficients are sufficient to describe the signal.

As described by Box and Jenkins (1976), a stationary input series implies that it has a constant mean, variance and autocorrelation through time. For non-seasonal, non-stationary signals this can be achieved by differencing the input signal. This yields the so-called autoregressive integrated moving average (ARIMA) process. In the ARIMA process, the observations y_t are substituted by $z_t = \nabla^d y_t$. In which d is the number of times the signal has to be differenced in order to achieve stationarity and ∇ is the differencing operator $\nabla y_t = y_t - y_{t-1}$.

If the signal contains significant periodic components they have to be removed by seasonal differencing the signal over these periods. Although this can be easily done for a time series that contains only a few periodic components, this becomes difficult if the number of periodic fluctuations increases, especially if the period and amplitude are not really constant but change with time.

Identification To forecast future values based on an AR(I)MA process, first it has to be determined how often the time series has to be differenced (indicated by a parameter d) and how many autoregressive and moving average coefficients are needed. This is the most difficult part of AR(I)MA modelling.

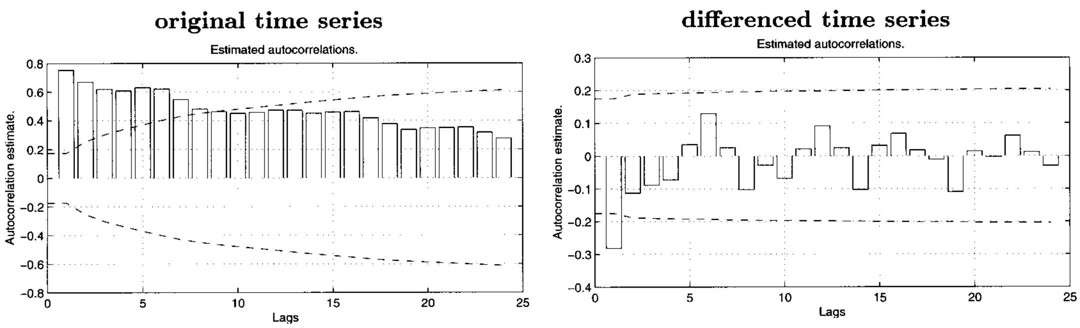

Fig. 3.25 Autocorrelation functions for annual mean sea levels. Dashed lines: 2σ of correlations.

A tool to determine the parameter d is the autocorrelation function. As described by Makridakis and Wheelwright (1978), for a stationary time series the autocorrelations drop to zero after the second or third time lag, while for non-stationary signals they are still significant after larger lags of time. This is demonstrated by the left-hand side of figure 3.25 in which the autocorrelation function for annual mean sea levels of tide gauge Den Helder is shown. From the right-hand side in figure 3.25, showing autocorrelations for the differenced time series, it is clear that after lag 1 autocorrelations are below the dashed lines (corresponding to twice their standard deviation). Therefore, it can be concluded that differencing the time series once is sufficient, i.e., $d = 1$.

Next, it has to be decided how many autoregressive (p) and moving average (q) parameters are needed to build the model. This decision is based on the autocorrelations and, so-called, partial autocorrelations of the time series. In general it holds that an autoregressive process of order p has p partial autocorrelations significantly different from zero; see Makridakis and Wheelwright (1978) for more details. Figure 3.26 shows the partial autocorrelation functions for the first difference of the time series. Partial autocorrelations differ significantly from zero for the first 4 lags.

According to Box and Jenkins (1976), for an ARMA process with p autoregressive and q moving average parameters, while $p < q$, the autocorrelation function should be a mixture of exponentials and damped sine waves after the first $q - p$ lags. On the other hand, if $p > q$, the partial autocorrelation function should be dominated by a mixture of exponentials and damped sine waves after the first $p - q$ lags. Based on the plots of the correlations and autocorrelations as shown in figures 3.25 and 3.26, it seems difficult to decide on the number of required autoregressive and moving average parameters.

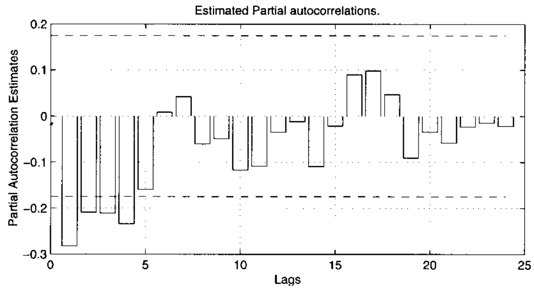

Fig. 3.26 Partial autocorrelation function for differenced time series for tide gauge Den Helder.

Estimation and verification; original time series The data set is represented by a number of ARMA models, with varying numbers of autoregressive and moving average parameters. Models considered are constituted from between 0 and 6 moving average parameters, while the number of autoregressive parameters also ranges between 0 and 6. For a number of the ARMA models, parameters could not be estimated because these models are not stationary and/or invertible. Based on the following test, it will be decided which of the (remaining) ARMA models fits the data best.

First, it is checked whether or not additional parameters in the ARMA models have significant values as compared to the error with which they can be estimated. It is found that the following ARMA(p,q) models have values that are at least twice their standard deviation: (1,0), (2,0), (3,0), (4,0), (0,1) and (0,2). Also acceptable might be, (5,0), (0,3), (1,1), (2,2), (4,1), (4,2), (5,1), and (6,1), which have values about equal to their standard deviations.

As a second test, the autocorrelations of the residuals are examined. If the selected models fit the data, the residuals should not contain any pattern. Consequently, autocorrelations for the residuals should be within the 95% confidence level for white noise. For most tested models, the major part of the autocorrelations for the first 24 lags are within this confidence level. Exceptions are: model (1,0) for lag 2, model (2,0) for lag 3, model (3,0) for lag 4, model (0,1) for lag 1, and model (0,2) for lag 8.

Next, a portmanteau lack of fit test is applied to the autocorrelations. With this test inadequacies in the estimated model can be detected by looking at a group of autocorrelations instead of the individual values; see Box and Jenkins (1976). It is found that, based on the autocorrelations of the residuals, all models are accepted except for models (2,2), (4,1), and (4,2).

How well models fit the data can also be compared based on the sum of squares of the residuals. The better the model fits the data, the smaller this sum of squares. It is found that the sum of squares has the same order of magnitude for all tested models, except for model (4,2). Relative high values are determined for the autoregressive models (1,0), (2,0), (3,0), and mixed models (2,2), and (4,1). However, the difference in sum of squares is not sufficiently large to indicate which model fits the data best.

As a final test, cumulative periodograms have been determined. These diagrams can be a useful tool to inspect the residuals for significant periodic effects in the data. If the selected ARMA models fit the data, the residuals should be random. As explained by Box and Jenkins (1976), the power spectrum of white noise has a constant value and, consequently, the cumulative spectrum of white noise plotted against frequency shows a straight line. For most models under consideration, the cumulative periodograms are within the 95% confidence interval corresponding to white noise. Small deviations, but still within the 90% confidence interval are found for models (0,1) and (2,0).

Based on the tests described above, it can be concluded that a single model that outperforms the others cannot be identified. A model that seems to fit the data is (4,0). Acceptable might also be: (5,0), (0,3), (1,1), (5,1) and (6,1). Some of these models contain a rather large number of parameters. However, as stated in the principle of parsimony (Box and Jenkins (1976)), the smallest possible number of parameters should be used to describe the data. Therefore, it is concluded that it is difficult to determine the required number of autoregressive and moving average parameters to describe the original time series for tide gauge Den Helder. Neither plots of the autocorrelation and partial autocorrelation functions, nor tests conducted on specific ARMA models give enough information to decide which model fits this data set best.

Estimation and verification; first principal component Next, it is tried to find a suitable ARMA model for a smoothed time series, i.e., for the first principal component of tide gauge Den Helder. To create a stationary data set the first difference of this time series is used. ARMA models are estimated that are constituted from between 0 and 6 autoregressive and between 0 and 6 moving average parameters.

Same tests are performed on the different ARMA models as described in the preceding. It is found that the portmanteau lack of fit test is rejected for all models under consideration. Therefore, it can be concluded that none of the tested models can be used to accurately model the differenced first principal component for tide gauge Den Helder.

Estimation and verification; seasonal differenced first principal component A possible cause for the rejection of the portmanteau test is the presence of significant periodic fluctuations in the data. Based on autocorrelation plots of the residuals for the various tested ARMA models, it is clear that at lag 20 the autocorrelation has by far the largest value. Since this value is negative, this is probably an indication of a significant periodic fluctuation with a period of around 40 years.

Next, seasonal differencing over a period of 40 years is applied to the data (the first principal component for tide gauge Den Helder). Autocorrelations and partial autocorrelations for this seasonal differenced data set are shown in figure 3.27. From this figure it can be concluded that all correlations and partial autocorrelations are within the 95% confidence interval corresponding to white noise, with the exception for time lag 1. Therefore, additional seasonal differencing is not necessary.

Based on figure 3.27, it might be concluded that the differenced and seasonal differenced first principal component for tide gauge Den Helder can be modelled by either a pure autoregressive or a pure moving average model, both with only one parameter. In the following, this assumption will be tested.

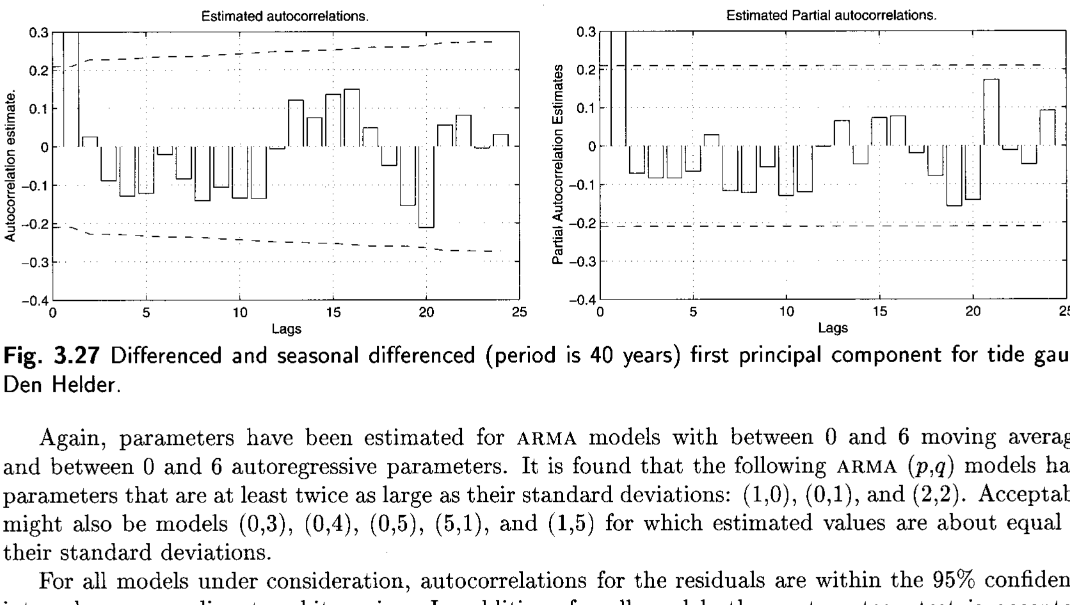

Fig. 3.27 Differenced and seasonal differenced (period is 40 years) first principal component for tide gauge Den Helder.

Again, parameters have been estimated for ARMA models with between 0 and 6 moving average, and between 0 and 6 autoregressive parameters. It is found that the following ARMA (p,q) models have parameters that are at least twice as large as their standard deviations: (1,0), (0,1), and (2,2). Acceptable might also be models (0,3), (0,4), (0,5), (5,1), and (1,5) for which estimated values are about equal to their standard deviations.

For all models under consideration, autocorrelations for the residuals are within the 95% confidence interval corresponding to white noise. In addition, for all models the portmanteau test is accepted. Finally, cumulative periodograms for the residuals reveal that all values remain in the 95% confidence interval corresponding to white noise. Therefore, it can be concluded that no remaining pattern is present within the residuals and all models seem to fit the data. The sum of squares of the residuals does not give an indication which model fits the data best, since it is of the same order of magnitude for all models under consideration.

Based on the test results described above, models that seem to fit the data best are (2,2), (0,1), and (1,0). While acceptable might be models (0,3), (0,4), (0,5), (5,1), and (1,5). However, with the concept of parsimony in mind, the best choice to model the data seems to be either a pure moving average (0,1) or a pure autoregressive model (1,0) with only 1 parameter.

Forecasting Based on the autocorrelation and partial autocorrelation plots, the significance of the estimated parameters, and with the concept of parsimony in mind, the first principal component for tide gauge Den Helder should be modelled with either a (1,0) or a (0,1) ARMA model. However, before applying this model, the data has to be differenced. In addition, seasonal differencing with a period of 40 years is needed.

For a pure autoregressive (1,0) model, future values of the time series are forecasted as:

$$\hat{z}_{t+k+1} = a_1 \hat{z}_{t+k} + \hat{\mu}_z \tag{3.28}$$

while for (0,1) pure moving average model values are simply forecasted as:

$$\hat{z}_{t+k+1} = \hat{\mu}_z \tag{3.29}$$

In these equations $\hat{\mu}_z$ is the estimate of the mean value in the time series.

To test how well the (1,0) and (0,1) model are able to forecast future values, only part of the data is used to estimate the model parameters. Based on these parameters, values for the remaining period can be predicted and compared to the actual values. Modelled and forecasted values are shown in figure 3.28, for two different starting years of the forecasting. From this figure it can be concluded that there is hardly any difference between results based on the pure autoregressive and results based on the pure moving average model. This is not surprising, since the two forecasting equations are almost equal. Figure 3.28 also shows that predictions starting in 1968 seem to fit the actual data better than those starting in 1977.

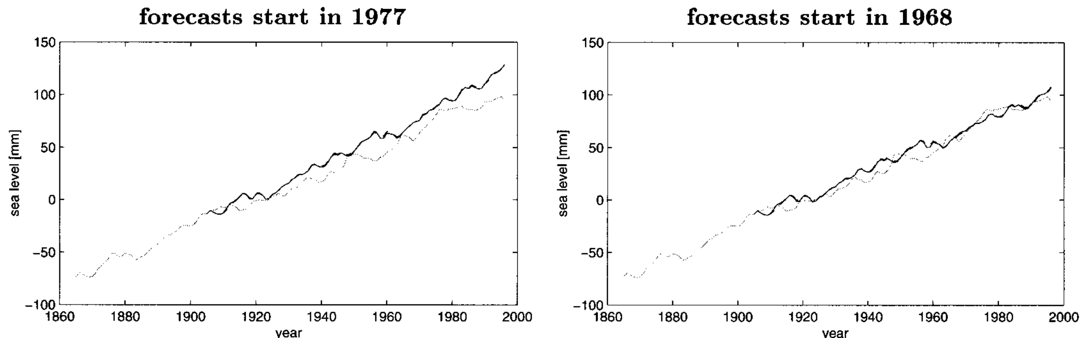

Fig. 3.28 First principal component for tide gauge Den Helder (grey), predicted and forecasted values based on Arma modelling. Black: autoregressive model ($a_1 = 0.3$), dashed: moving average model ($b_1 = 0.2$).

Next, both the autoregressive and moving average model are used to forecast future sea-level heights up to the year 2050; values are given in table 3.2. Three different data sets have been used, i.e., data available up to 1996, up to 1976, and up to 1967. There is a wide range in forecasted values for the year 2050. Predicted sea-level increases relative to the modelled values for the year 1996 are given as well. Although these values show more similarity, the variation between different estimates is still large.

period	model (p,q)	forecasted value for 2050	modelled value for 1996	increase
until 1996	(1,0)	183.4 mm	117.0 mm	66.4 mm
	(0,1)	182.4 mm	116.8 mm	65.5 mm
until 1976	(1,0)	210.5 mm	128.7 mm	81.8 mm
	(0,1)	208.2 mm	129.0 mm	79.2 mm
until 1967	(1,0)	156.0 mm	106.6 mm	49.4 mm
	(0,1)	159.4 mm	107.6 mm	51.7 mm

Table 3.2 Forecasted sea-level height for 2050, modelled level for 1996, and increase in 2050 relative to 1996. Differenced and seasonal differenced (period: 40 years) first principal component for tide gauge Den Helder

Conclusions It can be concluded that AR(I)MA modelling is not well suited to forecast future sea-level heights, if applied to original annual mean sea levels. Due to the specific characteristics of these sea-level height time series, it is very difficult to determine the number of autoregressive and moving average parameters needed for an accurate description of the data.

On the other hand, if applied to smoothed sea-level heights, AR(I)MA modelling seems to be able to forecast future sea-level values rather well. Results based on a pure (1,0) autoregressive model, or on a pure (0,1) moving average model are very similar. However, some dependence on the year in which the forecasts are started is found.

Since either a (1,0) or a (0,1) model is used to describe the data, forecasted values for the differenced and seasonal differenced time series are either based on the mean value or on the mean value in combination with the previous forecasted value. As a result, forecasts for the smoothed sea-level heights are mainly determined by the linear trend (due to differencing the series) and fluctuations with a period of 40 years (due to the seasonal differencing).

3.4 Sea-level data for a group of tide gauges

If more time series are available, one can either choose to determine a common sea-level variation curve for this group of tide gauges, or try to estimate a spatial pattern in sea-level height variation. Which of these two methods is the most appropriate depends, a.o., on the number of available time series, their quality (length, occurrence of data gaps, noise level), and on the correlation between the different time series.

If only a small number of time series are available, and especially if their quality is not very good, it will not be feasible to estimate a reliable spatial variation pattern. If a large enough amount of time series is available, it should be checked whether one common sea-level variation curve is sufficient to describe sea-level variation in the area, or a spatial pattern in sea-level rise has to be determined.

Even if the underlying sea-level variation curve is the same for all time series considered, these data sets will not be the same. Variations between the time series can be the result of, a.o.,

- measuring errors introduced by the tide gauges themselves
- height changes of the tide gauges that are not corrected for
- inaccuracies in height connections between tide gauges
- local sea-level variations in the area of individual tide gauges

All experiments performed in the remainder of this chapter are based on actual relative sea-level height data. This implies that height changes of the tide gauges themselves are included in the sea-level height series. Results for tide gauges that are connected in height will only be shown in chapters 6 and 8.

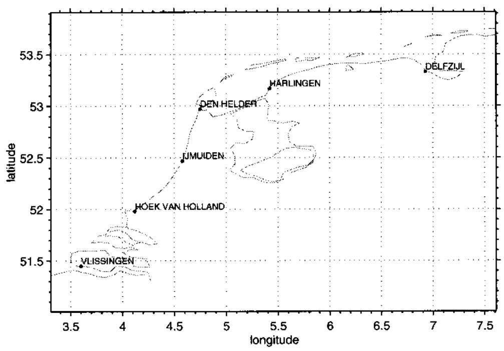

Fig. 3.29 Six stations in the Netherlands for which the PSMSL holds more than 100 years of data.

tide gauge	data from	trend mm/yr	data from	trend mm/yr
Delfzijl	1865	1.57	1886	1.84
Harlingen	1865	1.32	1886	1.23
Den Helder	1865	1.40	1886	1.42
IJmuiden	1872	1.45	1886	2.03
Hoek van Holland	1864	2.34	1886	2.30
Vlissingen	1862	1.19	1886	2.13
mean value		1.55		1.82

Fig. 3.30 Estimated trends and mean trend value for six tide gauges along the Dutch coast.

If for six tide gauges along the Dutch coast (see figure 3.29) trend estimates are determined, it seems that no common sea-level variation curve is present. As can be seen from table 3.30, trend estimates (for the common period 1886 - 1997) vary between 1.23 and 2.30 mm/yr.

If we look at the correlations between the time series, see left-hand side of figure 3.31 we see that the correlation coefficients between the six time series are at least 0.84, while most correlations are of the order of 0.9. If 10 stations in the North Sea area are added to the group of tide gauges, correlations decrease significantly, between some tide gauges even negative correlations are found; see right-hand side of figure 3.31. However, even for this group of 16 tide gauges, most correlations are of the order of at least 0.6. Therefore, it can be concluded that significant correlation is present between sea-level height series for a number of tide gauges in the North Sea area. In appendix A, some information concerning the stations used is given.

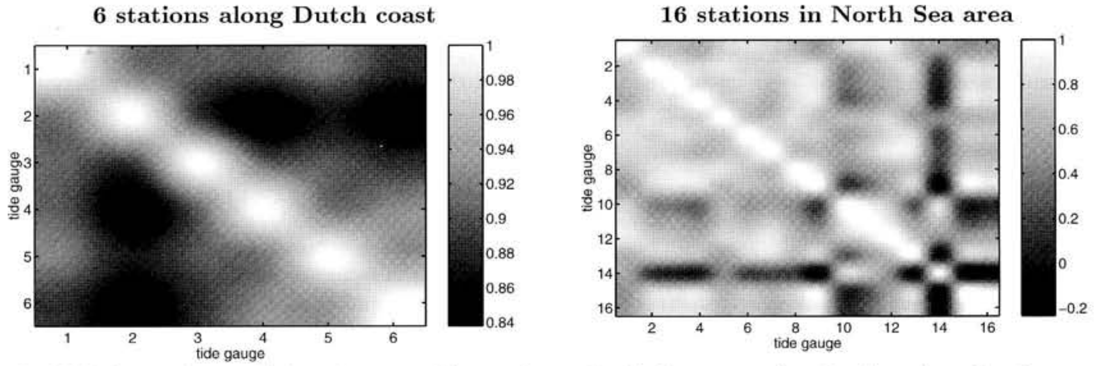

Fig. 3.31 Correlation coefficients between different (normalised) tide gauge series. Stations along Dutch coast: data from 1886 - 1997. Stations in North Sea area: data from 1898 - 1981.

Singular Value Decomposition (SVD) As a second method to examine the similarity between sea-level height series, SVD will be used. For this purpose, a data matrix (A) is created in which every column contains a sea-level height series for an individual tide gauge. The SVD of this data matrix yields:

$$A = USV^T \tag{3.30}$$

The U matrix contains as many singular vectors as there are time series in the data set. These singular vectors (which are orthogonal) contain different variation curves from which all individual time series can be reconstructed. The S matrix contains the singular values, which represent the power of the singular vectors relative to one another. Finally, the components in the V matrix indicate how strongly the different singular vectors contribute to the different time series.

If the time series in a data set consist solely of random noise, all singular values will be similar in size. On the other hand, if all time series are exactly the same, the first singular value will be very large (it contains all the energy present in the data set: $\sqrt{n\sum_{i=1}^{N} x_i}$ with n the number of time series and N the number of samples), while all other singular values are zero. Therefore, it is assumed that if the first singular value is much larger than the others, there is a large similarity between the time series. In general it can be stated that the first singular vector following from SVD analysis represents that part of the signal that is common to all time series in the data set.

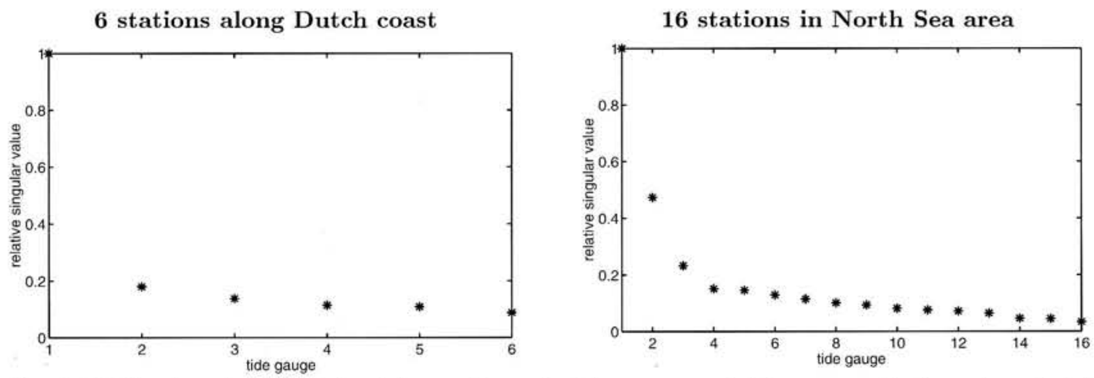

Fig. 3.32 Singular values (scaled relative to first value) for a group of tide gauges. Stations along Dutch coast: data from 1886 - 1997. Stations in North Sea area: data from 1898 - 1981. Normalised data.

The left-hand side of figure 3.32 shows the singular values following from SVD for the group of six tide gauges along the Dutch coast. The first singular values is roughly five times as large as the other singular values. Therefore, it can be concluded that the curve represented by the first singular vector is rather strong in all six time series. For the group of 16 tide gauges, see right-hand side of figure 3.32, the first singular values is only (approximately) twice as large as the second singular value. Figure 3.33 shows the six singular vectors (multiplied by their singular values) for the group of six tide gauges. The first singular vector clearly contains a trend in combination with a (long) periodic pattern.

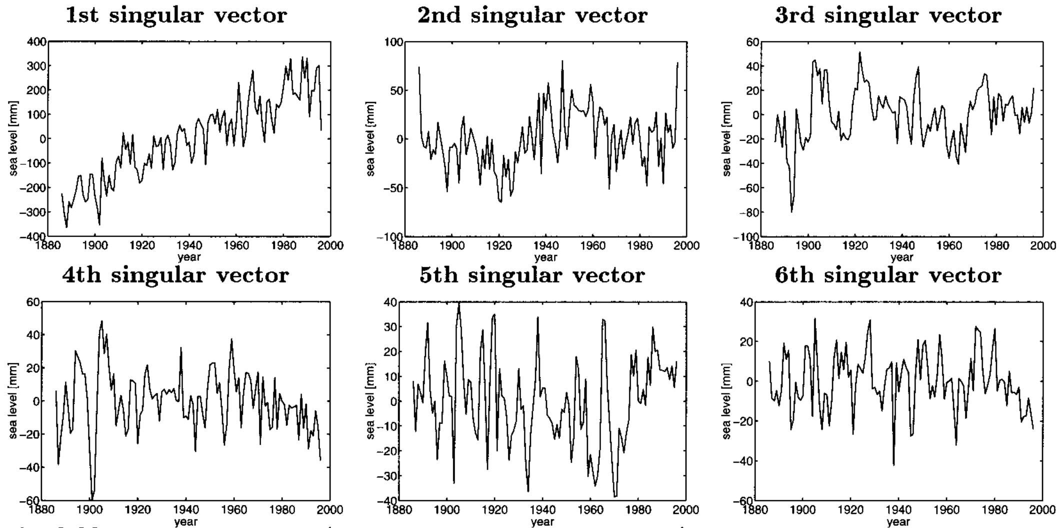

Fig. 3.33 Reconstructed signal (singular vector times singular value) based on group of six stations along the Dutch coast. (Note the differences in scale).

The fact that the first singular vector contains a clear trend and the others have (almost) no trend is no guarantee that a common trend is present for the group of tide gauges. This is confirmed by experiments with two simulated data sets. Both sets contain detrended values for the six Dutch tide gauges with additional simulated trends. For one data set, all time series contain a trend of 1.5 mm/yr; the second data set has time series with trend values ranging between 1.0 and 2.0 mm/yr. After SVD, for both data sets it is found that only the first singular vector holds a trend.

The differences in trend between the time series are not visible from the singular vectors or values themselves. However, as indicated in the preceding, the elements in the V matrix give a measure of how strongly the singular vectors are represented in the individual time series. Therefore, different values for these elements lead to a different scaling of the first singular vector and, consequently, to different trend values. As a result, the values for the elements in the V matrix corresponding to the first singular vector can be used as as indication of the similarity of the overall trend values in the individual time series.

Table 3.3 gives the elements in the V matrix, which relate the first singular vector to the six different time series. This table clearly shows that although significant differences are present, no clear pattern can be discerned in these values. For comparison, for the simulated data set in which all time series contain the same trend, the values for these V elements are more or less the same size. For the simulated data set with increasing trend values (1.0 mm/yr for Delfzijl up to 2.0 mm/yr for tide gauge Vlissingen), a pattern with increasing values for the elements of the V matrix is clearly visible.

data set	Delfzijl	Harlingen	Den Helder	IJmuiden	Hoek van Holland	Vlissingen
original data	0.41	0.31	0.33	0.44	0.48	0.44
simulated data						
same trend	0.42	0.42	0.41	0.41	0.40	0.39
increasing trends	0.32	0.36	0.39	0.43	0.46	0.48

Table 3.3 Factors relating the first singular vector following from SVD (for a group of six tide gauges along the Dutch coast) to the individual time series.

Detection of discontinuities in time series The method of singular value decomposition can also be used as a tool to detect if one of the time series contains a set of deviation observations. As already indicating in the preceding, tide gauges Vlissingen and IJmuiden show strong discontinuities in the beginning of these time series (sea-level heights show a more or less abrupt change in height). Visual inspection of the other four tide gauge series does not reveal major deviations.

Figure 3.34 shows reconstructed signals (singular vectors times singular values) based on different singular vectors. The group of tide gauges for which SVD is performed contains only four tide gauges. Although data for tide gauges Vlissingen and IJmuiden has not been included, reconstructions based on the second and third singular value show a deviation in the first part of the time series.

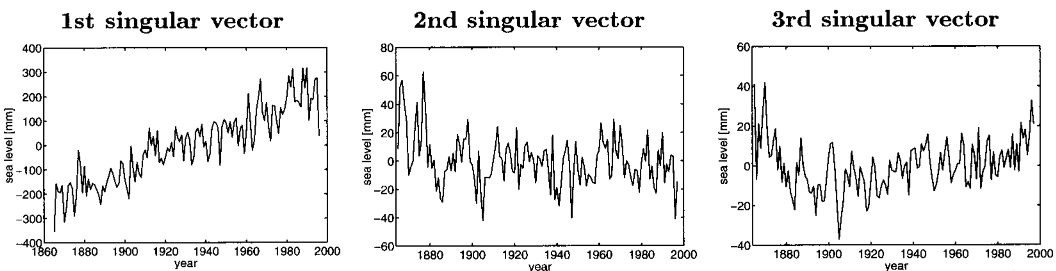

Fig. 3.34 Reconstructed signals (singular vector times singular value) based on group of 4 stations (Delfzijl, Harlingen, Den Helder, Hoek van Holland) along the Dutch coast. Data from 1865 to 1996.

The deviation in the first part of the reconstructed signals, as shown in figure 3.34, is not a common feature to all four time series, because, in this case it would show up in the reconstruction based on the first singular vector. If reconstructions are based on a combination of the second and third singular vector, deviations in the first part of the data set are enhanced for tide gauge Delfzijl, while they almost completely disappear for the other tide gauges; see figure 3.35. Therefore, it is concluded that tide gauge Delfzijl has produced the time series with deviating observations for the first part of the data set.

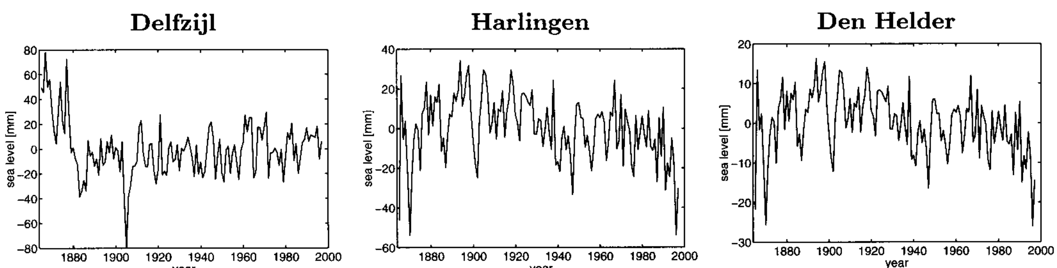

Fig. 3.35 Reconstructed signals based on both 2nd and 3rd singular vectors for different stations. Data set consists of 4 tide gauges: Delfzijl, Harlingen, Den Helder, Hoek van Holland. Data from 1865 to 1996.

Comparison between data sets In figure 3.33, the first singular vector (multiplied by its singular value) as derived for a group of six tide gauges along the Dutch coast was shown. Next, this data set is extended with sea-level height series for 10 additional tide gauges in the North Sea area. This requires a reduction of the length of the time series; see appendix A. The left-hand side of figure 3.36 compares the first singular vectors based on these two data sets. A large similarity between the two singular vectors, both in trend and periodic fluctuations, is clearly visible.

It can be argued that the large similarity between the first singular vectors, shown in the left-hand side of figure 3.36, is a results of the overlap between the two data sets and not of a common sea-level variation pattern in the area. To check whether this is true or not, a third data set is created, consisting of the 10 tide gauges that are not situated in the Netherlands. In the right-hand side of figure 3.36, the first singular vector based on this data set is compared to that based on the six tide gauges along the Dutch coast. The similarity between the periodic pattern in both vectors is still strong, but the trends seem to deviate somewhat.

In figure 3.36, reconstructions based on only the first singular vector are compared. For the data set with the six tide gauges along the Dutch coast, all higher singular values are small as compared to the

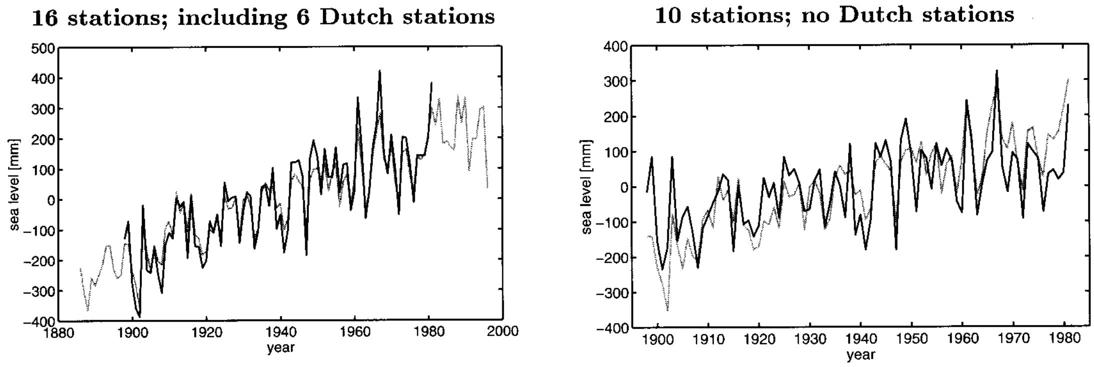

Fig. 3.36 Reconstructed signal (first singular vector times first singular value) based on a group of 6 six stations along the Dutch coast (grey) or a group of (16 or 10) stations in North Sea area (black). Normalised data.

first singular value. However, for both the data set with 16 and 10 stations, the second singular value is significant as compared to the first singular value. In addition, for these two data sets, the second singular vector shows a clear trend. Therefore, it might be more realistic to compare the results for the "Dutch" data set with reconstructions for the "North Sea" data sets based on both the first and the second singular vector. Results are shown in figure 3.37.

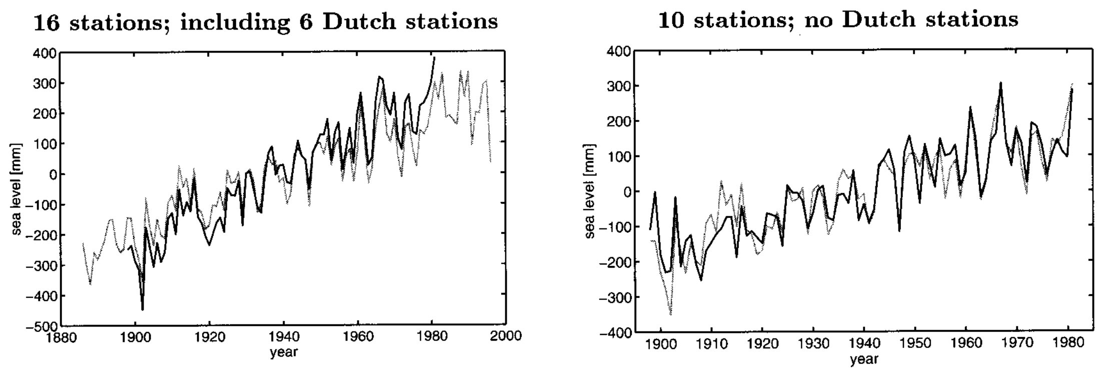

Fig. 3.37 Reconstructed signals. Gray: group of 6 six Dutch stations (1st singular vector times 1st singular value), black: group of (16 or 10) stations in North Sea area (1st and 2nd singular vector times 1st + 2nd singular value). Normalised data.

The data sets used above do not contain sea-level height data for the United Kingdom. In order to include tide gauges from the U.K., the time span considered for the calculations has to be reduced to the common data period from 1898 to 1965; see appendix A. Figure 3.38 shows correlations between the two tide gauges in the U.K. and the 16 stations in the North Sea area as used in the preceding. With a few exceptions, most correlations are between 0.5 and 0.8.

Figure 3.39 compares reconstructions (first singular vector times first singular value) based on the data sets with and without the two tide gauges in the U.K. It is clear that adding these two tide gauges to the data set does not really change the first singular vector, i.e., the sea-level variation curve common to the group of tide gauges.

Estimation of a common variation curve It is assumed that the variation curve consists of (a combination of) regression lines. Consequently, the relation between the sea-level height data and the common regression coefficients can be written as:

$$\hat{Y} = A\hat{x} \tag{3.31}$$

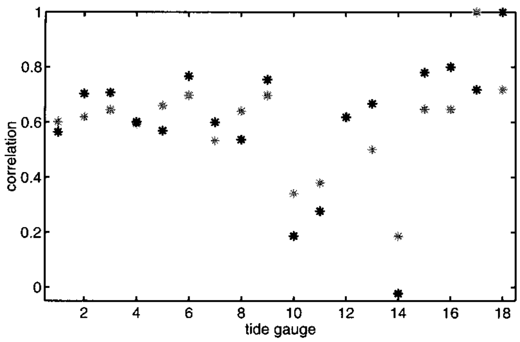

Fig. 3.38 Correlations between Aberdeen II (grey) or North Shields (black) and 17 other stations in the North Sea area.

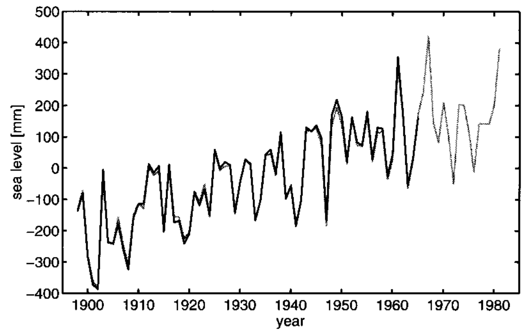

Fig. 3.39 Reconstructed signals. Gray: based on 16 stations (data: 1898 - 1981), black: data added for Aberdeen II and North Shields (data: 1898 - 1965).

in which Y is a vector containing the sea-level height measurements for all tide gauges, A is the design matrix containing, a.o., (functions of) the years in which observations are made, and x is the vector with (unknown) regression coefficients. These regression coefficients can be determined as the weighted least-squares estimate of x:

$$\hat{x} = (A^T W A)^{-1} A^T W Y \tag{3.32}$$

in which W is used to indicate the weight matrix. If sea-level height measurements are uncorrelated (no correlations within the time series and no correlations between the different time series), this weight matrix is given by:

$$W = \begin{pmatrix} \sigma^2_{y^1} & & & 0 \\ & \sigma^2_{y^2} & & \\ & & \ddots & \\ 0 & & & \sigma^2_{y^n} \end{pmatrix}^{-1} \tag{3.33}$$

in which $\sigma^2_{y^i}$ are the variances applying to the individual time series and n is the number of time series in the data set.

Estimation of spatial variation patterns in sea-level heights In principle, different attributes of the sea-level series can be used as input for a spatial pattern. For example, a spatial pattern of trend values could be determined, or a pattern can be derived for the elements in the V matrix that relate the first singular vector to the different time series.

In general, the relation between station coordinates and, e.g., trend values can be written as:

$$\hat{z} = A\hat{p} \tag{3.34}$$

with z the quantities through which the spatial pattern is estimated (e.g., trend values as determined for the individual time series), A the design matrix containing, a.o., (functions of) tide gauge coordinates, and p a vector containing the (unknown) coefficients that form the spatial pattern. These coefficients can be determined as the weighted least-squares estimate:

$$\hat{p} = (A^T W A)^{-1} A^T W z \tag{3.35}$$

in which W is the weight matrix of the quantities z (e.g., trend values). Assuming these quantities as obtained for the individual tide gauges to be uncorrelated, the weight matrix is given by equation (3.33) if $\sigma^2_{y^i}$ are replaced by the variances $\sigma^2_{z_i}$. If a spatial variation pattern is estimated for trend values, for $\sigma^2_{z_i}$ the a-posteriori variances of the trend values estimated for the individual time series can be used.

Depending on the number of stations through which the spatial pattern has to be determined, a large number of relations between station coordinates and, e.g., trend values, could be examined. Unfortunately, only for a very limited number of stations, sea-level height series with sufficient quality are available, and they are not well distributed over the North Sea area. Consequently, only relative simple spatial patterns can be, realistically derived.

Based on experiments with a variety of (2 dimensional) regression models it is found that models based on more than six regression parameters do not provide realistic results. These regression models result in extreme large variations in trend values; e.g., at the borders of the region under consideration, trend values of the order of 8 mm/yr are found. Consequently, only the following models will be considered:

linear:	$z = a + b_1x + b_2y$
bi-linear	$z = a + b_1x + b_2y + c_1xy$
quadratic without xy:	$z = a + b_1x + b_2y + c_2x^2 + c_3y^2$
quadratic:	$z = a + b_1x + b_2y + c_1xy + c_2x^2 + c_3y^2$

using x and y for the longitude and latitude of the tide gauges respectively.

Figure 3.40 shows resulting spatial patterns, estimated through trend values for the 18 sea-level height series in the North Sea region. All four patterns show a decrease in trend values towards the North-east. This is in agreement with post-glacial rebound models, which estimate maximum rates of uplift in Fennoscandia. In order to examine how well the various models fit the data sets, in table 3.4 the RMS of the fit between actual trend values (as determined at the tide gauge locations) and those following from the various spatial models are given. The RMS of this data fit is much smaller if a quadratic model is estimated instead of a (pure) linear one. However, difference with the two mixed models is not very large. Consequently, the bi-linear model is probably the best model to represent the trend values.

weighted trend values				trend values have uniform weight			
linear	bi-linear	quadratic $-xy$	quadratic	linear	bi-linear	quadratic $-xy$	quadratic
0.444	0.402	0.413	0.393	0.446	0.406	0.416	0.395

Table 3.4 RMS of data fit (in mm/yr) between actual trend values corresponding to tide gauges and values following from the different models. To estimate variation patterns, trend values are weighted (by their standard deviations) or have uniform weight.

If instead of weighted values (based on the a-posteriori standard deviation of the individual trend estimates) all trend values have equal weight, the RMS of the data fit (for the four spatial patterns) does not really increase. However, as can be seen from figure 3.41, some differences occur in the spatial patterns estimated through the trend values. The larger differences occur for the two spatial models containing the cross-term xy.

3.5 Conclusions and recommendations

In this chapter, a number of analysing techniques for smoothing sea-level data and predicting future sea-level heights have been examined. In addition, methods have been discussed to check the similarity between different sea-level data sets. Furthermore, the estimation of a common sea-level variation curve or a spatial variation pattern for a group of tide gauges is described. Based on derived results in this chapter, the following conclusions can be made.

Concerning smoothing of time series A good measure to describe the quality of the smoothing techniques is difficult to obtain. The less the time series is smoothed, the smaller the error between the smoothed and actual data will be. However, for a reliable smoothing method it should hold that a (small) change in the model (e.g., in the choice of base functions or in the value used for a smoothing parameter) should not cause major changes in smoothed values.

In general it is found that for the more complex techniques, results are extremely influenced by the choice of the model parameters. Unfortunately, no obvious choice (following from external information or based on the time series themselves) of these parameters is available.

From the techniques under consideration in this chapter, only moving average smoothing and Singular Spectrum Analysis seem more or less robust against a small change in their model parameters. Therefore, it is concluded that these methods are best suited to distinguish between a decadal pattern in sea-level variation and higher periodic fluctuations and noise. Techniques operating in the spectral domain should only be used to determine at which frequencies, significant, fluctuations are present in the data.

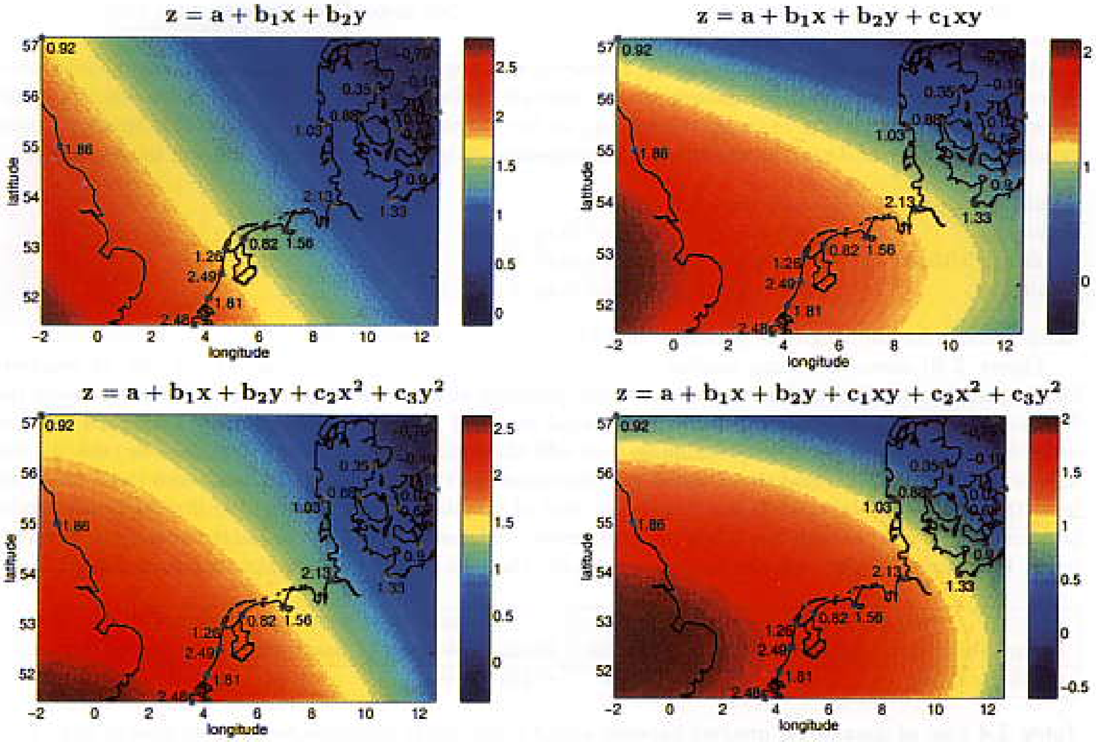

Fig. 3.40 Different spatial patterns estimated through (weighted) trend values (in mm/yr) for 18 tide gauges in the North Sea area. * indicate locations of tide gauges with corresponding trend values.

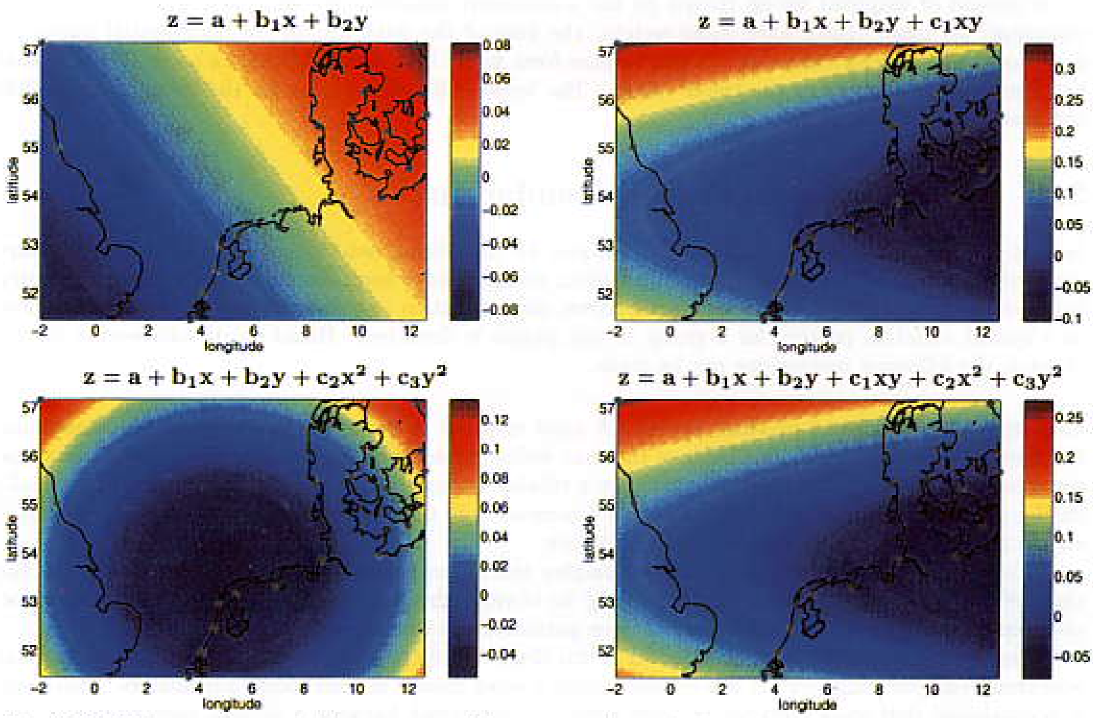

Fig. 3.41 Difference (in mm/yr) between spatial patterns estimated through weighted trend values and those based on uniform weights. * indicate locations of tide gauges.

Concerning predicting future sea levels In general, sea-level height series contain a trend in combination with very long-periodic fluctuations. The majority of the techniques (as discussed in this chapter) are not well suited to handle this kind of data. Often, forecasted values are based on the (slope and intercept determined for) the last available observation. Due to the long-periodic fluctuations, forecasted values are significantly influenced by the choice of the year in which the forecasting is started.

Of the techniques considered in this chapter, linear regression and AR(I)MA modelling appear to be best suited for forecasting future sea-level heights. For linear regression, reliable estimates require long time series. Moreover, they should not contain (large) parts with deviating observations. The method of AR(I)MA modelling can only, reliably, be used on smoothed sea-level heights. Linear regression only yields a linear trend for future sea levels, whereas with AR(I)MA modelling a linear trend in combination with a decadal pattern is found. However, due to this decadal pattern, forecasts based on AR(I)MA modelling are more susceptible to a change in starting year for the forecasting.

Concerning detection of specific sea-level variation curves Using the statistical significance of regression parameters, the distinction can be made between various sea-level variation curves, e.g., between linear and quadratic sea-level rise. As another example, based on the significance of the difference in trend, it can be decided whether one slope value throughout the time series is sufficient to describe the sea-level variations or a combination of two regression lines with different trend values has to be used. The significance of determined regression parameters can also be used to estimate the onset year of sea-level rise accelerations.

Concerning spatial variations Quantities relating different types of base functions (regression lines, singular vectors, etc.) can be used as input to estimate spatial patterns in sea-level variations. However, the quality of available sea-level data in in the North Sea area is not very good. Only for a limited number of tide gauges long time series are available. Of these, some have large data gaps. In addition, suitable time series are not well distributed over the North Sea area. As a result, at most, quadratic variation patterns can be estimated.

Chapter 4

Detectability of curves in relative sea level

4.1 Introduction

The main focus of this thesis is the role of geodetic measuring techniques on the detectability of sea-level variation patterns. To investigate these effects, experiments will be conducted on artificial data sets. These artificial data sets will partly be based on actual sea-level height data for a number of tide gauges, and partly on patterns simulated based on known information concerning sea-level height variations.

The reason for using artificial data sets is the following. Based on a well considered choice of analysing techniques and amount of tide gauge data to be used, various patterns can be "detected" in actual sea-level data. As an example, in section 3.2.1 it has been shown that if simple linear regression is applied to tide gauge data without prior smoothing, significantly different trends can be found if the amount of data available is changed.

By using artificial data, it is possible to build a number of data sets containing specific phenomena, e.g., a time series with a linear trend in combination with measuring noise, or a data set consisting not only of this linear trend but also of (long-term) periodic fluctuations. For these artificial data sets it can be examined what kind of sea-level variation patterns can still be, reliably, discerned due to the limited accuracy of the measurements, possible systematic errors, limited knowledge of the processes causing the sea-level variations, limitations of the analysing techniques, etc.

To exemplify the artificial data sets and the cases being investigated, in this chapter detectability of variation curves in individual sea-level height series will be considered. It will be assumed that sea-level height data is available for an individual tide gauge situated at an "ideal" location. Sea level height for a tide gauge at an "ideal" location is only determined by processes influencing regional sea level and not by all kind of site dependent processes like variations in river runoff.

For this "ideally located" tide gauge the detectability of a pattern in relative sea level will be investigated. This implies that only variations in sea-level height relative to the local tide gauge bench mark will be considered, changes in height of the tide gauge itself will not be taken into account. Since, in this chapter, only patterns in individual sea-level series are of interest, considering only relative sea levels is sufficient. In order to compare sea-level changes for a group of tide gauges, height variations of the tide gauges themselves have to be taken into account. This can be achieved by connecting all tide gauge bench marks to a local reference frame, as will be discussed in chapter 5.

Based on artificial data sets, it will be examined whether or not it is possible to detect specific sea-level variation curves, and how long the time series need to be for a reliable estimate of these patterns. A large number of simulated data sets, containing different patterns, periodic fluctuations and noise functions will be introduced. The general structure of the simulated time series will be explained in section 4.3. Prior to this, the methods used to estimate (combinations of) regression curves will be explained in section 4.2.

An acceleration in sea-level rise should not be very difficult to detect if the onset time of the acceleration is known. Unfortunately, this is not a very realistic assumption. For example, greenhouse-gas induced warming and a resulting increase in global temperature are expected to cause (or have already caused) an increase in ocean temperatures. As a result, sea levels are expected to rise, e.g., due to thermal expansion of the ocean water and increased melting of ice from glaciers and ice sheets. When this phenomenon is expected to start (or has already started) to affect local sea levels is difficult to assess. Furthermore, due to the so-called buffer effect of the oceans, a number of scientists expect that there will not be a straight relationship between an increase in global air temperature and a resulting increase in ocean water temperature. They expect that at first, ocean temperatures might hardly change, possibly followed by a relative large change in ocean temperatures in later years. Therefore, in order to detect an increase in sea-level rise, e.g., due to greenhouse-gas induced warming, it should be possible to, reliably, distinguish this effect from the "normal" pattern in sea-level variation.

For detecting (changes in) linear trends, the following data sets will be considered. First, in section 4.4, "ideal" measurement series will be used. These are time series from which all periodic fluctuations have been removed based on modelling of processes influencing sea-level heights. Next, in section 4.5 (simulated) short-periodic fluctuations will be added to the data sets. Finally, in section 4.6, the influence of long-periodic fluctuations on the detectability of patterns in relative sea-level height will be examined.

4.2 Estimating curves in individual time series

Based on experiments with a large variety of analysing techniques in chapter 3, it has been found that regression seems the most suitable technique for determining variation curves in sea-level height data. For more complex techniques, results can be significantly influenced by the choice of model parameters. Unfortunately, for these methods there is often no clear indication which model parameters are best suited for the time series under consideration, and new parameters have to be determined for different data sets. Regression is a relatively simple technique and determination of regression coefficients is straightforward.

All experiments as conducted in chapters 4, 6 and 8 are performed on time series with annual mean sea levels. As has been indicated in chapter 3, the use of annual mean sea levels instead of monthly values causes aliasing. However, aliasing mostly occurs in the higher frequencies, for example between 4 and 8 years. Since this thesis focuses on secular sea level height variations, this aliasing should not cause significant problems. In figure 4.1, trend values estimated for tide gauge Den Helder are plotted as a function of the starting year for the regression. This figure demonstrates that trend estimates do not change significantly if monthly mean values are used instead of annual ones.

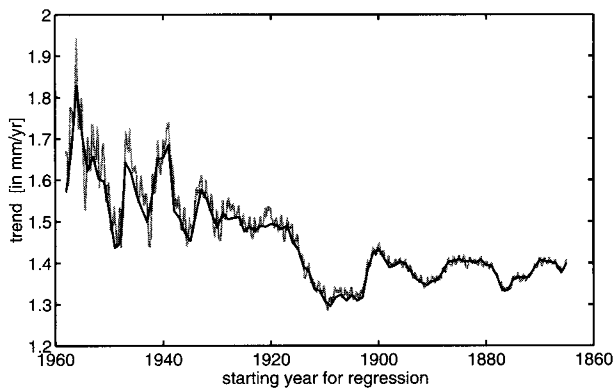

Fig. 4.1 Estimated trends versus the starting year for the regression. Annual mean values (black) or monthly mean values (grey) for tide gauge Den Helder up to the year 1997 are used. The left-most estimate is based on (40 years of) data between 1958 and 1997, the right-most estimate is based on (132 years of) sea level heights between 1865 and 1997.

How well a specific variation curve fits the time series will be tested based on plots of trend estimates as a function of the number of observations included in the regression, based on plots of differences between the time series and the fitted curve, and based on the statistical significance of the estimated regression coefficients. It should be remarked that in order to really check the validity of the assumed variation curve, the reliability of the applied statistical tests has to be considered as well. This is, however, beyond the scope of this thesis.

4.2.1 Secular sea-level increase

If regression is applied to sea-level data the following should be kept in mind. For most time series it is possible to determine regression coefficients that are significant in relation to their standard deviation. However, this is no guarantee that the regression model based on these parameters represents the sea-level variation curve in which we are interested. As an example, if linear regression is applied to data for tide gauge Vlissingen, a trend of 1.2 mm/yr is found. This trend value is significant (with 95% confidence level). However, the first 25 years of data strongly deviate from the remainder of the data set and, consequently, the regression model based on a trend of 1.2 mm/yr represents the sea-level variation curve not very well; see figure 3.2 in chapter 3.

A second example, for tide gauge Den Helder, is shown in figure 4.2. Based on visual inspection of the annual mean sea levels, one linear trend seems sufficient to describe the sea-level variation curve. It could be argued (see Emery and Aubrey (1991)), that for a reliable trend estimate of the order of 20 years of data is sufficient, since this includes at least one full circle of the lunar nodal tide, which is one of the main contributors to sea-level variations on longer time scales. In figure 4.2, linear regression lines are plotted for windows containing 20 years of sea-level data. A large variation in estimated slopes for the different windows is clearly visible.

Figure 4.2 shows that using a different segment of data yields a different slope value for the sea-level variation curve. This is confirmed by figure 4.3, in which estimated trend values are given determined for moving windows containing 20 years of sea-level heights. From figure 4.3 it can be seen that trend estimates vary between -1 mm/yr (window centred around the year 1970) and +7 mm/yr (window centred around 1980).

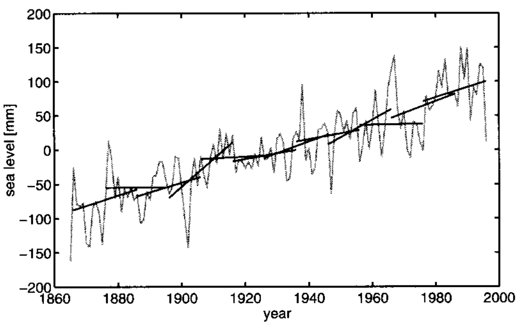

Fig. 4.2 Annual mean values (grey) and linear regression lines over windows of 20 years (black).

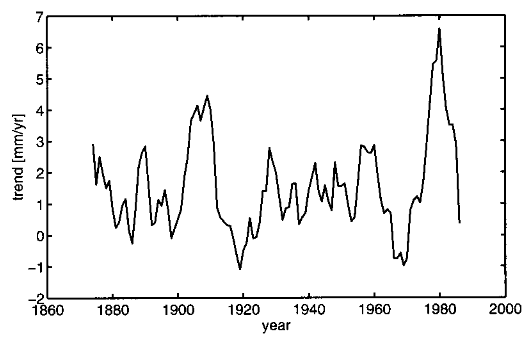

Fig. 4.3 Trends determined in moving windows of 20 years. Plotted versus central year in window.

Using $\hat{b}$ for the value estimated for the trend, and $\hat{\sigma}_{\hat{b}}$ for its a-posteriori standard deviation, the significance of an estimated trend value is given by:

$$t = \frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} \tag{4.1}$$

This significance can be used to test the hypothesis that no significant trend is present ($H_0 : E\{\hat{b}\} = 0$), against the hypothesis that there is a trend in the data ($H_1 : E\{\hat{b}\} \neq 0$). The test statistic in equation (4.1) should follow a Student's distribution with $N-2$ degrees of freedom (under H_0); with N the total number of observations. The trend can be estimated with a-posteriori standard deviation:

$$\hat{\sigma}_{\hat{b}} = \sqrt{\frac{\sum \varepsilon_i^2}{N-2}} \cdot \frac{1}{\sqrt{\sum (x_i - \bar{x})^2}} \tag{4.2}$$

in which $\bar{x}$ is the mean value of the years under consideration and ε_i are the differences between the actual values and the values predicted by the linear regression model.

If the differences ε follow a normal distribution with zero mean and a specific standard deviation σ_ε, equation (4.2) can be rewritten as:

$$\hat{\sigma}_{\hat{b}} = \sqrt{\frac{N}{N-2}} \cdot \frac{\hat{\sigma}_{\hat{\varepsilon}}}{\sqrt{\sum (x_i - i\bar{x})^2}} \tag{4.3}$$

Furthermore, for a vector x containing subsequent years of observation, it can be derived that

$$\sum_{i=1}^{N}(x_i - \bar{x})^2 = \frac{N(N^2-1)}{12} \tag{4.4}$$

Substitution of equation (4.4) into equation (4.3) shows that the standard deviation of the trend is determined by the level of the measurement errors and the number of observations used as follows:

$$\hat{\sigma}_{\hat{b}} = \hat{\sigma}_{\hat{\varepsilon}} \cdot \sqrt{\frac{12}{(N-2)(N^2-1)}} \tag{4.5}$$

Based on equation (4.5), it can be concluded that, for a window of 20 years, a trend estimate is significant if $\hat{b} > 0.07\ \hat{\sigma}_{\hat{\varepsilon}}$. It might be argued that for the small window size of 20 years, the majority of the slope values as shown in figure 4.3 might not be statistically significant. This is confirmed by the left-most side of figure 4.4. In this figure, those slope estimates that are significant (with 95% confidence) in relation to their a-posteriori standard deviation are denoted by $*$. Only a limited number of slope values are statistically significant. However, these few significant trend estimates still show major variations, i.e, they vary between 2 and 7 mm/yr.

Figure 4.4 also shows that an increase in window size over which slope values are determined, yields relatively more trend estimates that are significant in relation to their a-posteriori standard deviations. In addition, the larger the window size, the smaller the variations in trend estimates. However, even for windows containing 40 years of observations, trend estimates vary between 0.7 and 2.3 mm/yr.

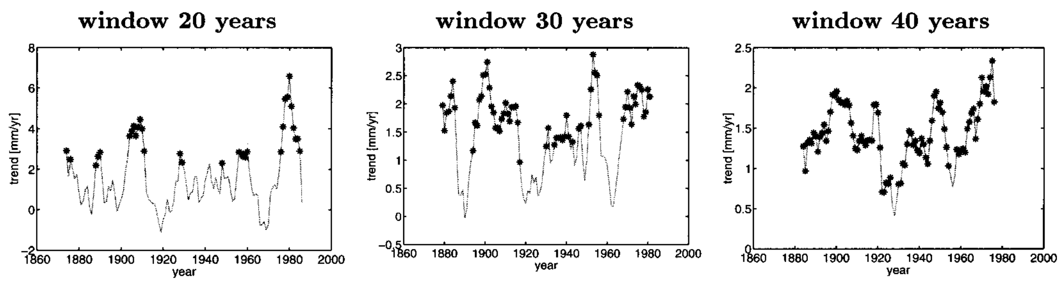

Fig. 4.4 Estimated trends (grey) and trends that are significant relative to their a-posteriori standard deviation ($*$); determined over moving windows with different sizes.

4.2.2 Accelerations in sea-level rise

Instead of a steady sea-level increase throughout the complete time series, an acceleration in sea-level rise might have occurred, e.g., due to increased industrial activity or greenhouse-gas induced warming. As a simple example, it is assumed that the sea-level variation curve can be represented by two separate linear regression lines, with a higher trend value applying to the second part of the data set, i.e.,

$$\begin{aligned} y_i &= a_1 + b_1 x_i + \varepsilon_i \qquad \text{for } i = 1, \ldots, n_1 \\ y_i &= a_2 + b_2 x_i + \varepsilon_i \qquad \text{for } i = n_1 + 1, \ldots, N \end{aligned} \tag{4.6}$$

Besides testing the significance of the two trend estimates ($\hat{b}_1, \hat{b}_2$) themselves, it can be tested whether or not the difference in trend ($\hat{b}_2 - \hat{b}_1$) is significant. An increase in trend has occurred if the hypothesis ($H_0 : E\{\hat{b}_2 - \hat{b}_1\} = 0$) is rejected in favour of the hypothesis ($H_1 : E\{\hat{b}_2 - \hat{b}_1\} > 0$). As test value

$$t = \frac{\hat{b}_2 - \hat{b}_1}{\sqrt{\hat{\sigma}^2_{\hat{b}_1} + \hat{\sigma}^2_{\hat{b}_2}}} \tag{4.7}$$

can be used, which should follow a Student's distribution with $N - 4$ degrees of freedom (under H_0).

If the differences between the actual values and the values predicted by the regression model follow a normal distribution with zero mean and standard deviation σ_ε, analogous to equation (4.3), the a-posteriori variances for the trend estimates can be estimated from:

$$\hat{\sigma}_{\hat{b}_1}^2 = \hat{\sigma}_{\hat{\varepsilon}}^2 \cdot \frac{12}{(n_1-2)(n_1^2-1)} \tag{4.8}$$

and

$$\hat{\sigma}_{\hat{b}_2}^2 = \hat{\sigma}_{\hat{\varepsilon}}^2 \cdot \frac{12}{(n_2-2)(n_2^2-1)} \tag{4.9}$$

in which $n_2 = N - n_1$ has been used.

In case of a change in trend value at the beginning of the time series only a relatively small part of the measurements will contain the low slope values. Consequently, the total number of observations (N) will be relatively large as compared to the number of deviating samples (n_1). In this case, the test statistic for testing whether or not a change in trend has occurred can be approximated by:

$$t \approx \frac{\hat{b}_2 - \hat{b}_1}{\hat{\sigma}_{\hat{\varepsilon}}} \sqrt{\frac{(n_1-2)(n_1^2-1)}{12}} \tag{4.10}$$

On the other hand, if the change in trend occurs towards the end of the time series, the number of samples to which the first slope value applies (n_1) will be large as compared to the number of samples based on the second trend value (n_2). Consequently, the expression to test the significance of the change in trend can be simplified to:

$$t \approx \frac{\hat{b}_2 - \hat{b}_1}{\hat{\sigma}_{\hat{\varepsilon}}} \sqrt{\frac{(n_2-2)(n_2^2-1)}{12}} \tag{4.11}$$

Using 1976 as year to divide the time series, Heinen and Hoogkamer (1993) found, for six tide gauges along the Dutch coast, no increase in trend value that was statistically significant (with 95% confidence level). However, figure 4.5 (for tide gauge Den Helder) shows estimated significances for the difference in trend values based on division years of the time series ranging between 1868 and 1994. For a division of the time series in either 1969, 1970, or 1971, the increase in trend value is statistically significant.

Figure 4.6 shows annual mean sea-levels in combination with linear regression lines determined for the data set from 1865 up to 1971, and for the data set from 1971 up to 1996. This figure shows a large jump between the two trend lines. Although the difference in trend values is significant, the combination of the two regression lines does not seem a good representation for the sea-level variation curve.

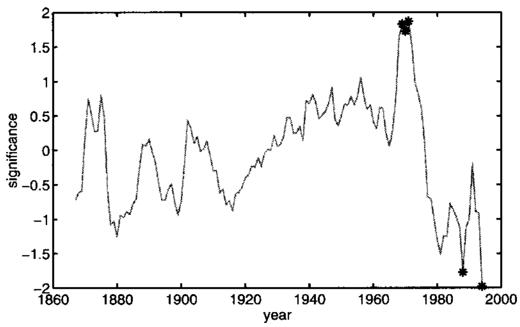

Fig. 4.5 Significance of difference in trend between first and second part data set, versus date used to split the time series.

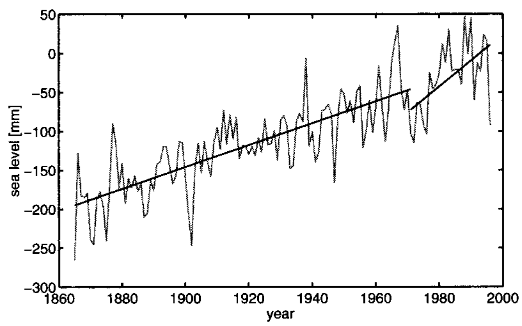

Fig. 4.6 Annual mean sea levels for Den Helder (grey), linear regression lines for 1st and 2nd part data set (black).

A jump in sea-level height between the two regression curves applying to the first and second part of the data set can be prevented by adding the restriction that the two regression curves should be joined. As a result, equation (4.6) changes to:

$$\begin{aligned} y_i &= a_1 + b_1 x_i + \varepsilon_i && \text{for } i = 1, \ldots, n_1 \\ y_i &= a_1 + b_1 x_{n_1} + b_2(x_i - x_{n_1}) + \varepsilon_i && \text{for } i = n_1 + 1, \ldots, N \end{aligned} \tag{4.12}$$

Unfortunately, due to this restriction, relations for the standard deviations of the trends become rather complex. As an example, the standard deviation of the trend value for the first part of the data set not only depends on the number of observations in this part of the time series, but on the number of samples in the second part of the data set as well. Equations are rather complex and, therefore, not shown.

Using equation (4.12), no statistically significant increase in trend value is found towards the end of the time series for one of the six tide gauges along the Dutch coast. For tide gauges Vlissingen and IJmuiden, increases in slope value at the beginning of the time series are significant. However, inspection of the annual mean sea levels (see e.g., figure 3.2 in chapter 3 for tide gauge Vlissingen) shows that a more or less abrupt change in sea-level height has occurred at the beginning of these time series. An increase in trend does not comply with these data sets.

4.2.3 Quadratic sea-level rise

As a second example to describe sea-level rise acceleration, quadratic regression is applied to the second part of the time series. In this case, the following relations hold between sea-level height (y) and year of observation (x).

$$\begin{aligned} y_i &= a_1 + b_1 x_i + \varepsilon_i && \text{for } i = 1, \ldots, n_1 \\ y_i &= a_2 + b_2 x_i + c x_i^2 + \varepsilon_i && \text{for } i = n_1 + 1, \ldots, N \end{aligned} \tag{4.13}$$

in which c is the quadratic regression term. With the additional constraint that the two regression lines have to join, this set of equations changes to:

$$\begin{aligned} y_i &= a_1 + b_1 x_i + \varepsilon_i && \text{for } i = 1, \ldots, n_1 \\ y_i &= a_1 + b_1 x_{n_1} + b_2 (x_i - x_{n_1}) + c(x_i - x_{n_1})^2 + \varepsilon_i && \text{for } i = n_1 + 1, \ldots, N \end{aligned} \tag{4.14}$$

Sea level heights can be assumed to follow a quadratic pattern instead of a linear rise if the variance of the measuring errors based on the linear curve is significantly larger than the variance based on a quadratic curve. This implies that the hypotheses $H_0 : E\{\hat{\sigma_l^2} - \hat{\sigma_q^2}\} = 0$ is rejected in favour of the alternative hypotheses $H_1 : E\{\hat{\sigma_l^2} - \hat{\sigma_q^2}\} > 0$. In these relations $\hat{\sigma_l^2}$ is the a-posteriori variance of the measuring errors (ε_i^l) based on the linear model, which is given by:

$$\hat{\sigma_l^2} = \frac{\sum_{i=1}^{N} \varepsilon_i^l}{(N-2)} \tag{4.15}$$

and $\hat{\sigma_q^2}$ is the a-posteriori variance of the measuring errors (ε_i^q) based on the quadratic model, which can be estimated as:

$$\hat{\sigma_q^2} = \frac{\sum_{i=1}^{N} \varepsilon_i^q}{(N-3)} \tag{4.16}$$

In these equations N is the number of observations.

To test the above mentioned hypothesis, the following test statistic can be used:

$$t = \frac{\hat{\sigma_l^2}}{\hat{\sigma_d^2}} \tag{4.17}$$

in which $\hat{\sigma_d^2}$ is the a-posteriori variance of the difference between the errors based on the linear model (ε_i^l) and the errors based on the quadratic model (ε_i^q). This variance can be estimated as:

$$\hat{\sigma_d^2} = \frac{\sum_{i=1}^{N} (\varepsilon_i^l - \varepsilon_i^q)^2}{(N-2)-(N-3)} = \sum_{i=1}^{N} (\varepsilon_i^l)^2 - \sum_{i=1}^{N} (\varepsilon_i^q)^2 \tag{4.18}$$

If the H_0 hypothesis applies, the test statistic in equation (4.17) should follow a $F_{1-\alpha}(N-2, 1)$ distribution.

Although, strictly speaking, the above equations should be used to distinguish between linear and quadratic sea-level rise, for reasons of simplicity, the following method will be applied. It is assumed that a quadratic variation curve should only be used to describe sea-level rise if the estimated quadratic regression coefficient ($\hat{c}$) is significant in relation to its standard deviation ($\hat{\sigma}_{\hat{c}}$). Consequently, the hypothesis that sea-level rise is linear ($H_0 : E\{\hat{c}\} = 0$) is tested versus the hypothesis that the increase is quadratic ($H_1 : E\{\hat{c}\} > 0$). The following test statistic is used:

$$t = \frac{\hat{c}}{\hat{\sigma}_{\hat{c}}} \tag{4.19}$$

which should follow a Student's distribution with $N-2$ degrees of freedom (under H_0). This method can be justified by the fact that if the estimated quadratic regression coefficient is not statistically significant, estimated values for the intercept and slope will not differ much from those based on linear regression.

If the two regression curves follow equation (4.13), the variances of the regression coefficients for the second part of the time series are independent from the number of data points in the first part of the data set. It can be derived, see appendix B.1, that the a-posteriori standard deviation for the quadratic regression coefficient depends on the differences (ε_i) between the actual values and those predicted by the quadratic regression curve (represented by the a-posteriori standard deviation $\hat{\sigma}_{\hat{\varepsilon}}$) and the number of available observations in the second part of the time series, n_2, as follows:

$$\hat{\sigma}_c = \hat{\sigma}_{\hat{\varepsilon}} \cdot \sqrt{\frac{180}{(n_2-2)(n_2^4 - 5n_2^2 + 4)}} \tag{4.20}$$

With the restriction that the two regression curves have to join, a simple analytical expression for the variance of the quadratic regression coefficient cannot be derived. This variance not only depends on the number of samples in the second part of the time series, but on the number of observations in the first part of the data set as well. The resulting expression for the variance of the quadratic regression coefficient contains a large number of elements and cannot easily be simplified.

4.2.4 Onset of sea-level rise acceleration

Although an acceleration in sea-level rise might be expected (from inspection of sea-level data or external information), the actual onset year will usually not be known. In order to fit a combination of regression curves to a sea-level height data set, an estimation of the onset year of the sea-level rise acceleration has to be made. The following method will be used to estimate the onset year of sea-level rise acceleration.

First, a window will be selected in which the onset year should be found. If sufficient external information is available concerning the sea-level variation curve, this window might be small, otherwise, a large window has to be used. Next, using the first year in the window, the data set is divided into two time series for which regression coefficients are determined based on equation (4.12) (linear rise 2nd part data set) or equation (4.14) (quadratic rise 2nd part data set). This procedure is repeated for every possible division year of the time series. For all combinations of regression curves, the significance of the regression parameters is determined. It is assumed that the highest value for the significance of the difference in trend (resp. the significance of the quadratic regression coefficient) corresponds to the most likely value for the onset year of the sea-level rise acceleration.

As a test of the estimated onset year of the sea-level rise acceleration, significances of the other regression coefficients are examined as well. For example, the onset year could also be estimated as that year which yields the highest significance of the trend determined for the first part of the time series. Estimated onset years of sea-level rise accelerations based on the significance of resp. $\hat{b}_1$, $\hat{b}_2$, or $\hat{b}_2 - \hat{b}_1$, should not show a large variation.

4.3 General structure of simulated data sets

All experiments as performed in the remainder of this chapter use simulated time series. Time series are constructed containing a signal part (the curve that needs to be detected), a noise part (measuring noise and irregularities in the beginning of the time series), and possibly periodic fluctuations. All data sets are based on annual mean sea-level heights, with values available for the years 1865 up to 1996. Simulated time series used to examine the detectability of accelerated sea-level increase due to greenhouse-gas induced warming, contain annual mean values up to the year 2100.

4.3.1 Simulated patterns in sea-level height time series

A wide range of (combinations of) regression curves could be envisioned for describing sea-level height variations over long periods of time for different locations. Of these, only the following simple examples will be examined.

- one specific linear trend throughout the time series
- transition to a higher trend value in the beginning of the time series
  Change can occur suddenly or there might be a "transition period" of a few years.
- transition to a higher trend value at the end of the time series
  Change can occur suddenly or there might be a "transition period" of a few years.
- change from linear to quadratic increase at the end of the time series

For the most simple data sets, one trend value applying throughout the time series will be used. According to the 1995 report of the Intergovernmental Panel on Climate Change (Warrick *et al.*, 1996), over the past 100 years, sea level has increased with 10 to 25 cm. Therefore, to simulated a data set containing the same linear trend throughout the time series, slope values of either 1.0, 1.5 , 2.0 or 2.5 mm/yr will be used.

Although over the past 100 years sea-level rise seems to be of the order of 1.0 to 2.5 mm/yr, historical data reveals that this is a much higher rate than that which has occurred over the last thousands of years. On average, sea-level rise over the past 2000 years has probably been of the order of a few tens of centimetres. The relatively higher values in sea-level increase, as determined for the last century, are probably due to increased industrial activity leading to increased global temperatures.

The actual sea-level variation curve for the past few centuries is not known. Therefore, it is simply chosen to use a combination of two linear trend lines (with the second trend value larger than the first) to describe sea-level rise acceleration resulting from increased industrial activity. Processes contributing to an increased sea-level height variation over the past century have been described by Warrick *et al.* (1996). They give a median estimate for their combined influence over the last 100 year of 8 cm. Consequently, curves used to simulate the onset of increased industrial activity will start with a 0.8 mm/yr lower slope value, i.e., with trends of resp. 0.2, 0.7, 1.2, or 1.7 mm/yr.

Although it is certain that the rate of sea-level increase during the last 100 years is relatively high as compared to the average value over the last 2000 years, it is uncertain when this acceleration in sea-level rise has started. Even though much research has been conducted on long sea-level height series (with over 100 years of data), it has not been possible to prove a sea-level rise acceleration over the last century. Based on high water levels at Liverpool since 1768, Woodworth (1999) suggests an acceleration in sea-level rise around mid or late nineteenth century.

Since no information concerning the actual onset of increased sea-level rise is available, each of the following three scenarios will be examined:

- in 1895 an abrupt increase in slope with 0.8 mm/yr has occurred
- over a transition period of 10 years, starting in 1890, a linear increase in slope value has occurred until trend is: 1.0, 1.5, 2.0, or 2.5 mm/yr
- over a transition period of 30 years, starting in 1880, a linear increase in slope value has occurred until trend is: 1.0, 1.5, 2.0, or 2.5 mm/yr

It is expected that, due to greenhouse-gas induced warming, an acceleration in sea-level rise may occur in the near future, or might already be taking place. According to Warrick *et al.* (1996), depending on specifics of the climate models used, predictions for global sea-level height increase in the year 2100 (relative to 1990) range between 20 and and 86 cm, with a best estimate of 49 cm.

Since it is not (yet) known when this acceleration in sea-level rise has started, or will start in the near future, as onset date the year 1991 will be used. To simulate the effect of greenhouse-gas induced warming on future sea-level heights, one of the following scenarios will be assumed:

- in 1991 an abrupt increase in slope occurs, with trend values of 1.8, 4.5, or 7.8 mm/yr
- over a transition period of 10 or 30 years, starting in 1991, a linear increase in slope occurs
  Final trend value is chosen that yields sea-level height increase for the year 2100 (relative to 1990) of resp. 20, 49, or 86 cm. In left-hand side of figure 4.7, curves for 49 cm higher sea level are shown.

- starting in 1991, sea-level heights change quadratically
  A linear trend component of 2, 3, or 4 mm/yr is assumed respectively. An accompanying quadratic regression coefficient is used that yields a sea-level height for the year 2100 that is 49 cm higher than that in the year 1990. Resulting curves are shown in the right-hand side of figure 4.7.

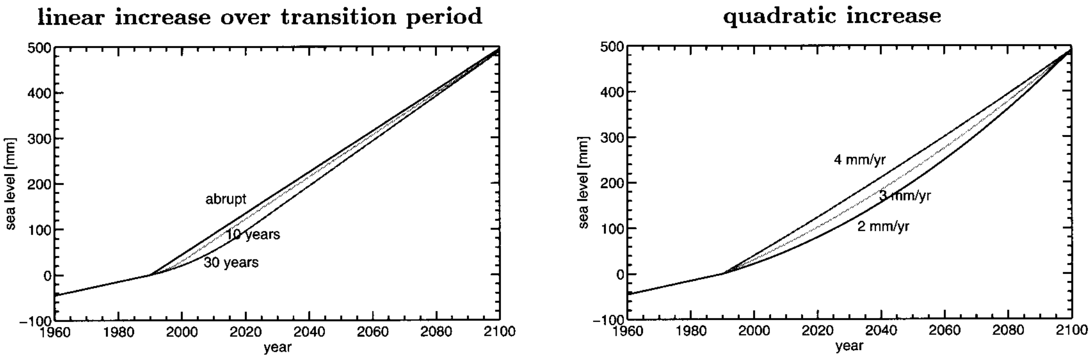

Fig. 4.7 Simulated sea-level variation patterns. First, a trend of 1.5 mm/yr; from 1991 accelerated sea-level rise. Left: abrupt change in trend to 4.5 mm/yr, or linear increase over transition period. Right: quadratic increase with linear components of 2, 3, and 4 mm/yr.

4.3.2 Simulated noise in sea-level height time series

The purpose of this chapter is to examine how well sea-level variation curves can be detected if the measurements are affected by noise and/or periodic fluctuations present in the data. To simulate noise for the annual mean sea-level heights, one (or a combination) of the following scenarios is used:

- same quality of measurements throughout the time series
  It is assumed that measurements are only affected by normally distributed random errors. The same noise level applies throughout the time series.
- different measuring systems have been used
  It is still assumed that measurements are only affected by random errors, but different standard deviations apply to different parts of the time series.
- tide gauge has been relocated
  Relocation of the tide gauge will cause a systematic error consisting of a jump in height throughout the remainder of the time series.

For the more simple noise functions it is assumed that the same measuring quality applies throughout the time series. Noise is simulated by using normally distributed random values with zero mean and a standard deviation of 0.5, 1, 5, or 10 mm.

Over the years different tide gauge systems have been used and yearly sea-level heights have been based on varying number of daily measurements. As a result, the assumption of a uniform measuring precision might not be very realistic. As an alternative, noise functions can be simulated based on the known history for specific tide gauges.

For tide gauge Den Helder, until 1885 sea levels have been observed 4 times daily, until 1970 8 observations/day have been used and since 1970 yearly sea levels have been based on hourly observations. Experiments conducted by Van der Hoek Ostende and Van Malde (1989) show how much the precision of sea-level heights deteriorates, when using a smaller number of daily observations; see section 2.3. Based on these experiments, measurement noise is simulated by using normally distributed random values with a standard deviation of 3 mm until 1885, σ 1.5 mm until 1970, and from 1970 on a precision of 1 mm. An example of a resulting noise function is shown in the left-hand side of figure 4.8.

For tide gauge Den Helder, from the start of the period for which sea-level heights are available (through the PSMSL), annual sea levels have been based on at least 4 observations/day. For a number of other tide gauges along the Dutch coast, there have been periods in which annual sea-level heights have been based on mean tide levels. As a result, precisions of these annual sea levels will be much lower. For example, for tide gauge Hoek van Holland the history of observations has been: until around 1890 mean daytime tide levels, until 1920 4 observations/day, until 1935 mean tide levels, until 1970 8 observations/day, since 1970 hourly values.

Based on the history for tide gauge Hoek van Holland, a second noise function, with a relatively high noise level, has been created. This noise function is simulated by using normally distributed random values with a precision of 20 mm until 1890, a standard deviation of 3 mm until 1920, σ 15 mm for the period until 1935, a precision of 1.5 mm until 1970, and finally, from 1970 on a noise level of 1 mm. An example of this noise function is shown in the right-hand side of figure 4.8.

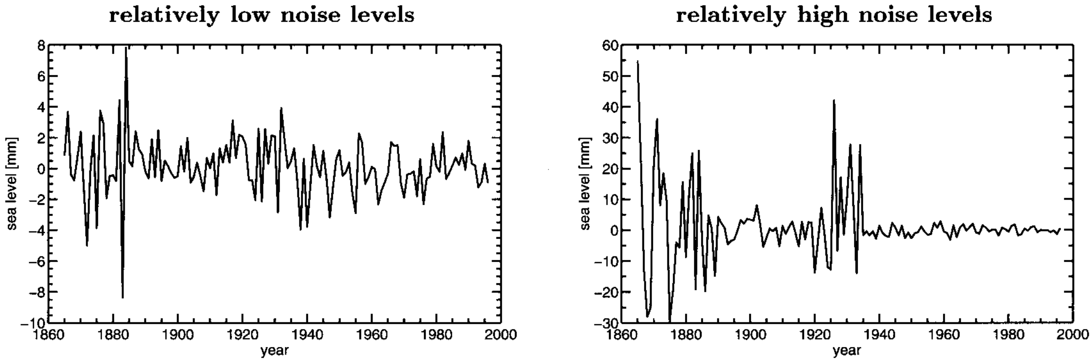

Fig. 4.8 Left: noise function based on tide gauge Den Helder (yielding a relatively low noise level). Right: noise function based on tide gauge Hoek van Holland (yielding a relatively high noise level).

If a tide gauge is relocated, this will often result in a jump in sea-level height. Visual inspection of actual tide gauge series reveals (small) jumps in some of the data sets. For example, in the sea-level height series for tide gauge Vlissingen, around 1885 a, rather sudden, decrease in sea level of the order of 10 cm is found. This effect is simulated by increasing sea-level heights by 10 cm up to 1890. The year 1890 is used since this corresponds to the transition from using mean daytime tide levels to 4 observations/day in the simulated noise function. It is quite possible that at this time also the tide gauge system has been changed, e.g., from a tide pole to a stilling-well tide gauge with mechanical recording.

4.3.3 Simulated periodic fluctuations in sea-level height time series

Periodic fluctuations could be modelled based on knowledge of the processes underlying these fluctuations. However, although a lot of information is known concerning processes influencing sea-level heights on relatively short time scales, i.e., up to periods of two years, not much is known concerning longer periodic processes. As can be seen from figure 3.9, actual tide gauge data contains a lot of periodic fluctuations with periods longer than 2 to 10 years.

Since simulating periodic fluctuations based on knowledge of processes underlying the fluctuations seems very difficult, the following strategy has been used. Using SSA (see section 3.2.4), sea-level data for an actual tide gauge is divided into a number of principal components. By using only specific principal components, a signal can be constructed with a required frequency contents.

To simulate periodic fluctuations, SSA has been applied to yearly sea-level heights for tide gauge Den Helder. A window size of 20 years has been used, implying that the total signal can be reconstructed based on 20 principal components. Long-periodic fluctuations are simulated based on principal component 2 up to 7, in combination with the detrended reconstruction for the first principal component. Resulting periodic fluctuations are shown in the middle graph in figure 4.9.

Short-periodic fluctuations are simulated by reconstructing a signal based on principal components: 9, 13, 15, 16, 18, 19, and 20. These components have been selected because they contain (almost) only information with periods up to 5 years. Resulting signal is shown in the left-most side of figure 4.9.

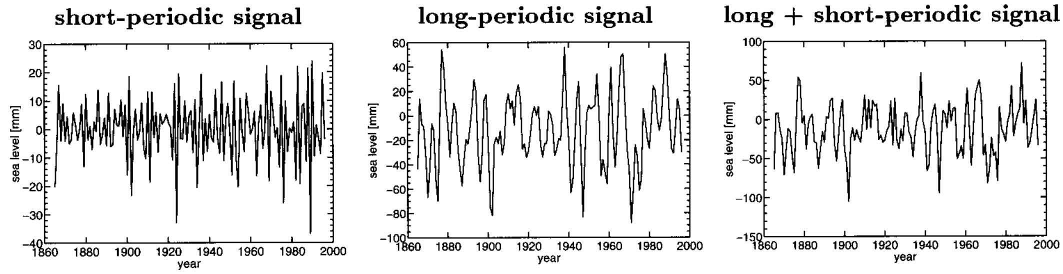

Fig. 4.9 Periodic fluctuations simulated from data for tide gauge Den Helder. Reconstructions based on specific singular values following from SSA analysis have been used.

4.4 "Ideal" measurement series

First, an "ideal" measurement series is assumed, i.e., all periodic fluctuations have been eliminated from the data set based on modelling of the processes causing these fluctuations. The remaining time series consists only of a linear trend. Of course, it is not very realistic to assume that all periodic fluctuations can be removed from the sea-level height series. Our current knowledge of the processes underlying these fluctuations is simply not sufficient. However, these "ideal" time series can be used to indicate what is maximally attainable from sea-level height data. Furthermore, they can be used as a first step in building more realistic and, therefore, more complicated simulated sea-level height series.

4.4.1 Same linear trend throughout the time series

If a time series consists of a linear trend in combination with random measuring noise, a simple relation between the time needed to detect this trend and the standard deviation of the measuring noise can be derived as follows. In section 4.2, it has been derived that the hypothesis that no significant trend is present ($H_0 : E\{\hat{b}\} = 0$) can be tested against the hypothesis that there is a trend in the data ($H_1 : E\{\hat{b}\} \neq 0$), by means of the test statistic

$$t = \frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} = \frac{\hat{b}}{\hat{\sigma}_{\hat{\varepsilon}}} \cdot \sqrt{\frac{(N-2)(N^2-1)}{12}} \tag{4.21}$$

with $\hat{b}$ the value estimated for the trend, $\hat{\sigma}_{\hat{b}}$ its a-posteriori standard deviation, $\hat{\sigma}_{\hat{\varepsilon}}$ the a-posteriori standard deviation of the measurement errors, and N the number of yearly observations. The test statistic in equation (4.21) should follow a Student's distribution with $N-2$ degrees of freedom (under H_0). Consequently, with a 95% confidence level, a trend b can be detected if the number of observations is at least

$$\sqrt{(N-2)(N^2-1)} \geq 2\sqrt{3}\frac{\hat{\sigma}_{\hat{\varepsilon}}}{\hat{b}} \cdot t_{0.05}(N-2) \tag{4.22}$$

with $t_{0.05}(N-2) \approx 1.7$ for 18 or more observations.

The smaller the trend that has to be detected, the larger the number of observations that are needed. In addition, if the precision of the measurements improves, the required number of observations decreases. Approximately (for relative large values of N), the following relation holds:

$$\frac{N_2}{N_1} \approx \left(\frac{\sigma_{\varepsilon_2}}{\sigma_{\varepsilon_1}}\right)^{\frac{2}{3}} = \left(\frac{b_1}{b_2}\right)^{\frac{2}{3}} \tag{4.23}$$

Therefore, in order to detect a trend that is half as large, approximately 1.6 times the number of observations are needed. In figure 4.10 the relation is plotted between trend values and the number of observations required to detect this trend with a 95% confidence level.

The relation between required number of observation and measuring noise only holds if the same noise level applies to all observations. In figure 4.8 two noise functions have been shown that have been simulated based on the known history for tide gauges Den Helder (relatively low noise level) and Hoek van Holland (higher noise level). These noise functions are added to linear sea-level rise functions with trend values between 1.0 and 2.5 mm/yr. Next, linear regression is applied to (parts of) these data sets.

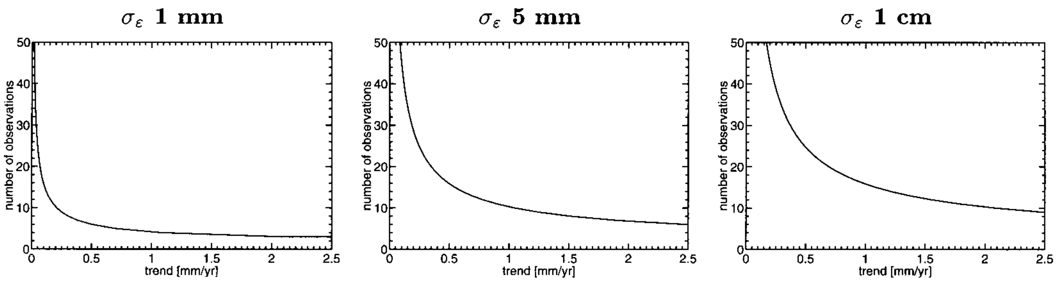

Fig. 4.10 Number of observations required to detect a trend with 95% confidence level. 3 levels of measuring noise.

In figure 4.11, for data containing a trend of 1.5 mm/yr, estimated trends are plotted versus the number of observations used. Trend estimates are shown for 100 different realizations of measuring noise. The left-most estimate is based on only the last 10 years of observations, the next point on the last 11 years of data, etc. The right most trend estimates use the full data set of 132 years of observations.

For the relative low noise level, figure 4.11 shows that if of the order of 15 to 20 observations have been included, trend estimates (based on different realizations of measuring noise) all stabilise around the actual slope value in the time series. If the relative high noise function (based on tide gauge Hoek van Holland) applies, using around 70 years of data causes a somewhat deviating value in estimated trend. This is due to the relatively large noise level for the period between 1920 and 1935.

Figures corresponding to the three other trend values will not be shown, since, apart from a difference in level, the pattern in estimated trends is the same. This is not surprisingly since all four simulated signals consist of a linear trend with deviations from this trend based on the same noise functions.

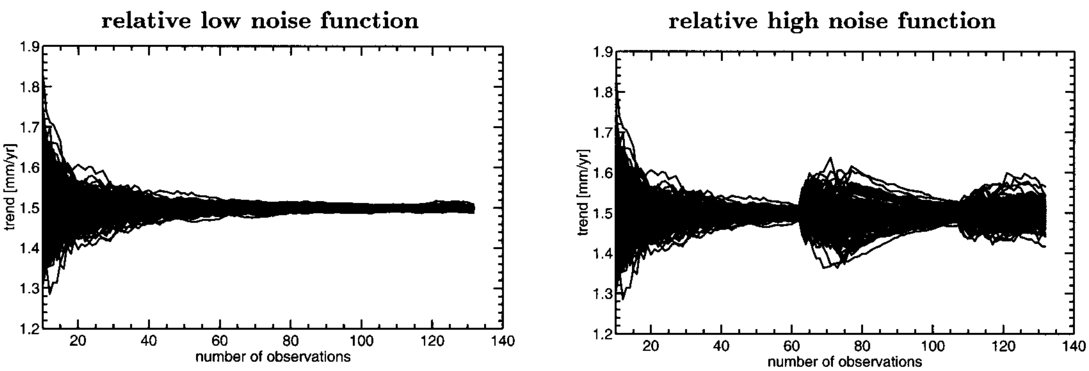

Fig. 4.11 Estimated trend versus number of observations (left-most estimate based on data for 1987-1996, right-most estimate on data from 1865 up to 1996). Based on a trend of 1.5 mm/yr and (100 realizations of) measuring noise. Left: tide gauge Den Helder (low noise level), right: Hoek van Holland (high noise level).

Relocation of the tide gauge As explained in section 4.3.2, this effect is simulated by increasing sea-level heights by 10 cm up to 1890. From simulated time series, even if the large noise level applies, the sudden decrease in sea-level height is quite visible from a plot of the data itself. As indicated in section 3.2.1, this kind of deviating parts of data should be removed prior to analysing the sea-level heights.

If trends are estimated based on the complete time series, results are highly influenced by the deviating first 25 years of data. Slope values estimated are resp. 0.3, 0.8, 1.3, and 1.8 mm/yr. The deviation between the estimated trends and the actual trends in the time series is independent of the actual trend values. This deviation is solely determined by the size of the jump in sea level (Δsl), the number of observations in the beginning of the time series containing the increased sea levels (n), and the total number of observations (N). It can be derived (see appendix B.2) that the deviation in trend, Δb_j, is related to these variables as:

$$\Delta b_j = 6\Delta\text{sl}\frac{n(n-N)}{N(N^2-1)} \tag{4.24}$$

In section 3.2.1 it has been shown that the discontinuities in the sea-level height series introduced by the relocation of the tide gauge could be detected by testing. However, for the simulated data sets used, the jump in sea level is clearly visible from plots of the annual mean sea levels. Therefore, no further tests have been performed.

4.4.2 Transition to higher trend in the beginning of the time series

In this section, it will be examined whether or not it is possible, based on the limited accuracy of sea-level height time series, to prove an acceleration in sea-level height in the beginning of this time series. As described in section 4.3.1, it is assumed that a 0.8 mm/yr lower slope value applies to the first part of the data set, with 1895 as onset year of the sea-level rise acceleration.

If linear regression is applied to time series containing two different rates of sea-level increase, trend values are estimated that are somewhat smaller than those that apply to the last part of the data set, i.e, values of resp. 0.9, 1.4, 1.9, and 2.4 mm/yr are determined. The deviation between these estimated values based on the complete time series and the slope value that applies to the last part of the time series is indicated by Δb_t. This deviation in trend (Δb_t) is independent from the slope value that applies to the last part of the time series. It can be derived, see appendix B.3, that Δb_t only depends on the difference in trend ($\bar{b} = b_2 - b_1$) in the two parts of the data set, the number of samples in the beginning of the time series containing the lower trend value(n_1), and the total number of samples (N) as:

$$\Delta b_t = \frac{n_1\bar{b}}{N(N^2-1)}\left(2n_1^2 + 9n_1 - 9N - 3n_1N + 1\right) \tag{4.25}$$

For large values of N, and n_1 small relative to N, this expression can be approximated by:

$$\Delta b_t \approx -3\bar{b}\left(\frac{n_1}{N}\right)^2 \tag{4.26}$$

As an example, in case $N = 130$ and $n_1 = 30$, this approximation introduces an error of 8%.

Of interest is whether or not it is possible, based only on available time series, to prove that a change in trend has occurred. If two different trend values apply, sea-level heights are given by equation (4.12). An increase in trend has occurred if the hypothesis $H_0 : E\{\hat{b}_2 - \hat{b}_1\} = 0$ is rejected in favour of the hypothesis $H_1 : E\{\hat{b}_2 - \hat{b}_1\} > 0$. As test value the significance of the estimated difference in trend can be used, i.e.,

$$t = \frac{\hat{b}_2 - \hat{b}_1}{\sqrt{\hat{\sigma}^2_{\hat{b}_1} + \hat{\sigma}^2_{\hat{b}_2}}} \tag{4.27}$$

This test statistic should follow a Student's distribution with $N-4$ degrees of freedom. As explained in section 4.2, no simple analytical expression can be derived for this test statistic. However, without the restriction that the two regression curves should be joined (sea levels following equation (4.6)), the test statistic can be simplified to the relation given in equation (4.10). Therefore, if the total number of observations $N = 132$ and the difference in trend value $b_2 - b_1 = 0.8$ mm/yr, with a confidence level of 95% the difference in trend cannot be detected if

$$\sqrt{(n_1 - 2)(n_1^2 - 1)} < 7.1 \cdot \hat{\sigma}_{\hat{\varepsilon}} \tag{4.28}$$

Consequently, for $\hat{\sigma}_{\hat{\varepsilon}}$ 1 mm, at least 5 measurements containing the lower trend value are required (i.e., $n_1 \geq 5$). If this standard deviation is 1 cm, at least 18 observations are needed.

In figure 4.12 estimated trends are plotted versus the number of observations used, for time series in which the major part of the data sets contain a trend of 1.5 mm/yr. Trend estimates are shown for 100 different realizations of measuring noise, using three different noise levels. This figure clearly shows that the larger the noise level, the larger the spread in estimated trend values based on different realizations of measuring noise. In addition, if more observations are included, the spread in trend estimates reduces.

If the quality of sea-level height measurements is very good (a standard deviation of 1 mm applies), the change in trend after about 100 years of observations have been included is clearly visible. Also a fair indication can be obtained of the onset time of this change in trend. If annual mean sea levels have a measuring precision of 0.5 cm, the lowering in trend estimates is still visible, but a lot more observations containing these lower slope values are required to really discern this change in trend. For the highest noise level (σ 1 cm) lower trend values are (almost) impossible to discern.

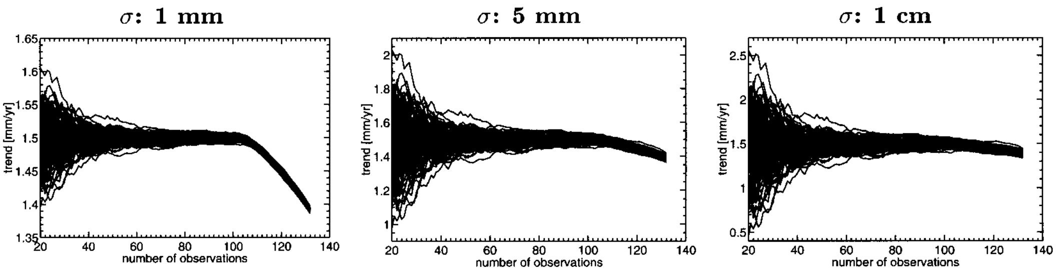

Fig. 4.12 Estimated trend versus the number of observations used (left-most estimate based on data for 1977-1996, right-most estimate on data from 1865 up to 1996). First 30 years: trend of 0.7 mm/yr; in 1895 abrupt change to 1.5 mm/yr. For 100 realizations of measuring noise (3 noise levels).

Next, it is assumed that the onset time of the increased sea-level rise is unknown. For one specific realization of measuring noise added to a time series containing a trend of 1.5 mm/yr, it will be tried to estimate this onset year, based on the significance of the regression parameters. Using different years to divide the time series, the time series is split in two parts, for which separate trend values are estimated. It is assumed that the division year that yields the highest values for the significance of the regression parameters should be used as estimate for the onset year of the sea-level rise acceleration; see section 4.2 for more details on this method. For a division of the time series ranging between 1891 and 1900, table 4.1 shows estimated trend values, their standard deviation and significance.

For a noise level of 1 mm, based on the significance of the first trend value $(\hat{b_1}/\hat{\sigma}_{\hat{b_1}})$ the year 1897 is selected to divide the time series. If the significance of the second trend value $(\hat{b_2}/\hat{\sigma}_{\hat{b_2}})$ is used as an indicator, the year 1894 would be selected. Finally, highest value for the significance of the difference in trend $(\hat{b_2} - \hat{b_1}/(\hat{\sigma}^2_{\hat{b_2}} + \hat{\sigma}^2_{\hat{b_1}})^{\frac{1}{2}})$ is derived if the year 1896 is used to divide the time series. These values are close to the actual year in which the trend acceleration occurs, i.e., 1895. The values estimated for the trends are also close to their actual values.

If a larger measuring noise level applies (0.5 or 1 cm) it holds, for both parts of the data set, that the longer the time series the higher the significance of the estimated trend. For the significance of the difference in trend, a maximum value is found within the range of years as shown in table 4.1. However, values for this significance vary widely, and other (even higher) "maximum" values are for example found for the year 1882 (for σ 5 mm) and 1887 (for σ 1 cm). Therefore, it can be concluded that the method of determining the onset year of the sea-level rise acceleration based on the maximum significance of the difference in trend, only works in case of a very good quality of measurements.

If, instead of a constant noise level, the noise function simulated based on the history for tide gauge Den Helder (relatively low noise level function; see figure 4.8) is used, trend estimates are derived as shown in the left-most graph in figure 4.13. The change in slope value is clearly visible, and a reasonable estimate of the onset time of the acceleration can be made. This is confirmed by table 4.2, based on the significance of resp. the second trend value and the difference in trend onset years 1895 and 1899 are found. If the relatively high noise function applies, change in trend is no longer visible. Neither from trend estimates or from significances of trend values.

The two right-hand side graphs in figure 4.13 show estimated trends based on data sets in which increase in trend value has occurred over transition periods of 10 or 30 years. Both for a measuring precision of 1 mm. From both graphs the change in slope value is clearly visible. However, comparison with results in figure 4.12 yields that the larger the transition period, the more difficult it is to estimate the onset time

	standard deviation 1 mm					standard deviation 5 mm					standard deviation 1 cm				
year	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$
1891	0.53	18.3	1.49	453.2	32.9	0.45	5.1	1.50	102.2	11.8	0.47	3.2	1.52	65.6	7.0
1892	0.57	23.8	1.49	472.3	38.1	0.49	6.1	1.51	101.1	12.2	0.52	3.7	1.53	64.9	*7.1*
1893	0.60	31.1	1.49	484.4	45.0	0.54	6.8	1.51	100.9	12.2	0.57	4.3	1.53	64.1	7.0
1894	0.64	40.6	1.49	*489.1*	53.3	0.58	7.7	1.51	99.9	12.2	0.62	4.8	1.53	63.5	7.0
1895	0.67	51.0	1.50	487.0	61.0	0.62	8.7	1.51	98.5	*12.4*	0.66	5.4	1.53	62.5	7.0
1896	0.70	55.6	1.50	482.8	*61.2*	0.65	9.0	1.52	99.1	11.7	0.70	6.0	1.53	61.6	7.0
1897	0.73	*57.0*	1.50	473.0	58.0	0.68	10.0	1.52	97.8	11.9	0.74	6.5	1.54	60.9	6.8
1898	0.76	54.8	1.50	460.6	52.1	0.71	10.4	1.52	98.4	11.4	0.78	7.1	1.54	60.2	6.8
1899	0.79	50.5	1.51	452.0	45.1	0.75	11.1	1.53	97.1	11.3	0.81	7.5	1.54	60.0	6.6
1900	0.81	48.5	1.51	437.1	40.7	0.78	11.9	1.53	95.8	11.2	0.83	8.1	1.54	59.1	6.7

Table 4.1 Estimated trends and their significances; t_{b_1}, t_{b_2}, and $t_{b_2-b_1}$ indicate resp. the test statistic (significance) of the 1st trend value, the 2nd trend value and the difference in trend. For various division years of the time series. Up to 1895 trend of 0.7 mm/yr, in 1895 abrupt increases to 1.5 mm/yr. Constant noise level, σ: 1 mm, 5 mm, or 1 cm.

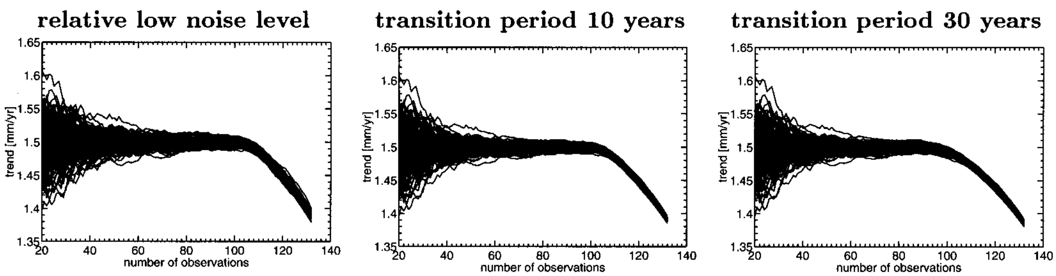

Fig. 4.13 Estimated trend versus the number of observations used (left-most estimate based on data for 1977-1996, right-most estimate on data from 1865 up to 1996). First, trend of 0.7 mm/yr; in 1895 abrupt change to 1.5 mm/yr (with relative low noise level), or a transition period applies (with constant noise level; σ 1 mm). For 100 realizations of measuring noise.

of the sea-level rise acceleration. From significances of trend estimates (see table 4.2), onset time can still be determined, although, in case of a transition period of 30 years estimates are off by a few years.

4.4.3 Transition to higher trend at the end of the time series

It is assumed that sea-level rise acceleration in the second part of the time series can be represented by a linear trend with a higher slope value. The two parts of the time series can, again, be described by equation (4.12). The test statistic for testing whether or not the increase in trend value is significant is given by equation (4.27). Analogous to the preceding section, a simple analytical expression for this test statistic can be derived if sea-level heights follow equation (4.6) instead of equation (4.12). As derived in section 4.2, without the restriction that the two regression curves should be joined, the test value for the significance of the difference in trend $(b_2 - b_1)$ can be approximated by:

$$t \approx \frac{\hat{b}_2 - \hat{b}_1}{\hat{\sigma}_{\hat{\varepsilon}}} \sqrt{\frac{(n_2 - 2)(n_2^2 - 1)}{12}} \tag{4.29}$$

in which n_2 $(= N - n_1)$ is the number of observations containing the increased sea-level rise, $\hat{b}_1$ and $\hat{b}_2$ are the estimated values for the trend in respectively the first and second part of the data set, $\hat{\sigma}_{\hat{\varepsilon}}$ is the a-posteriori standard deviation of the measuring errors, N is the total number of observations, and n_1 is the number of observations to which the first trend value applies.

The test statistic in equation (4.29) should follow a Student's distribution with $N - 4$ degrees of freedom (under H_0). It can be used to test whether or not the hypothesis $H_0 : E\{\hat{b}_2 - \hat{b}_1\} = 0$ is rejected

	relative low noise level					transition period 10 years					transition period 30 years				
year	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$	$\hat{b}_1$	t_{b_1}	$\hat{b}_2$	t_{b_2}	$t_{b_2-b_1}$
1891	0.54	8.7	1.48	317.4	15.2	0.54	19.1	1.48	445.3	33.6	0.58	20.3	1.47	378.7	31.2
1892	0.58	10.4	1.48	319.2	16.4	0.57	24.7	1.49	462.0	38.8	0.61	24.5	1.48	393.0	33.9
1893	0.61	12.0	1.48	325.8	17.2	0.61	32.0	1.49	473.0	45.4	0.65	29.4	1.48	405.3	37.1
1894	0.64	14.0	1.49	325.0	18.2	0.65	40.8	1.49	478.4	52.6	0.68	35.0	1.48	415.1	40.5
1895	0.68	15.7	1.49	*327.6*	18.8	0.68	50.9	1.50	476.3	*59.7*	0.71	42.2	1.48	418.2	45.0
1896	0.71	17.6	1.49	325.7	19.5	0.71	53.1	1.50	*479.3*	57.5	0.74	44.2	1.49	433.3	43.8
1897	0.73	19.4	1.50	320.9	20.0	0.74	*55.7*	1.50	472.3	56.0	0.77	48.6	1.49	437.6	45.0
1898	0.76	21.1	1.50	315.0	20.2	0.77	55.3	1.50	462.4	51.9	0.79	52.0	1.49	439.9	*45.1*
1899	0.79	22.6	1.50	308.2	*20.3*	0.79	52.1	1.51	442.4	45.9	0.81	51.8	1.50	*450.5*	42.4
1900	0.81	24.0	1.50	300.8	20.1	0.82	50.7	1.51	426.7	42.0	0.84	*54.0*	1.50	448.2	41.7

Table 4.2 Estimated trends and their significances; t_{b_1}, t_{b_2}, and $t_{b_2-b_1}$ indicate resp. the test statistic (significance) of the 1st trend value, the 2nd trend value and the difference in trend. For various division years of the time series. Up to 1895 trend of 0.7 mm/yr, in 1895 this trend value either changes abruptly to 1.5 mm/yr (with relative low noise level), or a transition period applies (with constant noise level σ 1 mm).

in favour of the alternative hypothesis ($H_1 : E\{\hat{b}_2 - \hat{b}_1\} > 0$). For a relatively large number of observations (e.g., of the order of at least 60 years of data), with a confidence level of 95%, an increase in trend value $\hat{\Delta b} = \hat{b}_2 - \hat{b}_1$ cannot be detected if for the number of samples with the higher trend value (n_2) holds:

$$\sqrt{(n_2 - 2)(n_2^2 - 1)} < 5.7 \cdot \frac{\hat{\sigma}_{\hat{\varepsilon}}}{\hat{\Delta b}} \tag{4.30}$$

Whether or not an acceleration in sea-level height can be detected is mainly determined by the number of observations containing the increased sea-level heights, the increase in trend and the noise level of the measurements. This is confirmed by results presented in figures 4.14 and 4.15. In these figures, estimated trends are plotted versus the year up to which observations are assumed available. The left-most estimates in the graphs are only based on observations up to 1920, whereas for the right-most estimates observations up to 2100 or 2040 have been used.

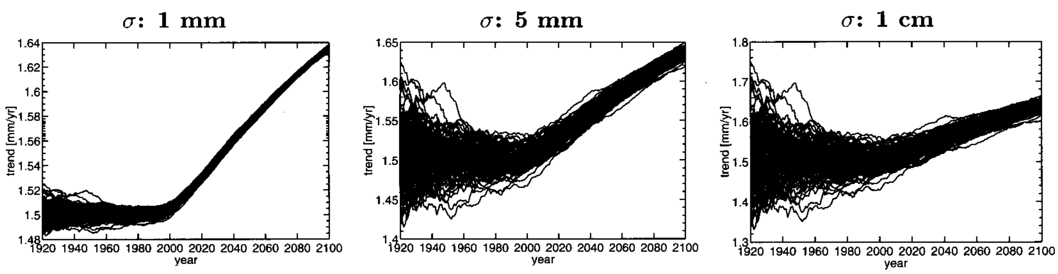

Fig. 4.14 Estimated trend versus the date up to which observations have been used. First, trend of 1.5 mm/yr; in 1991 this trend abruptly changes to 1.8 mm/yr. For 100 realizations of measuring noise (3 noise levels).

Estimated trends in figure 4.14 are based on a time series in which sea-level rise experiences only a small acceleration, i.e., from 1.5 to 1.8 mm/yr. This figure clearly demonstrates that if the quality of the measurements is very good (σ: 1 mm) this acceleration already becomes visible if a small amount of future observations becomes available. For larger noise levels, curves of estimated trend values show larger fluctuations. As a result, more future observations are required to detect the sea-level rise acceleration.

If greenhouse-gas induced warming leads to a larger difference in trend (e.g., increase to 4.5 mm/yr), the influence of measuring noise on the required number of observations to detect this acceleration is much less profound. As can be seen from figure 4.15, although the spread in estimated trend values increases with increasing noise level, the sea-level rise acceleration is still clearly visible after a limited amount of future observations has been included.

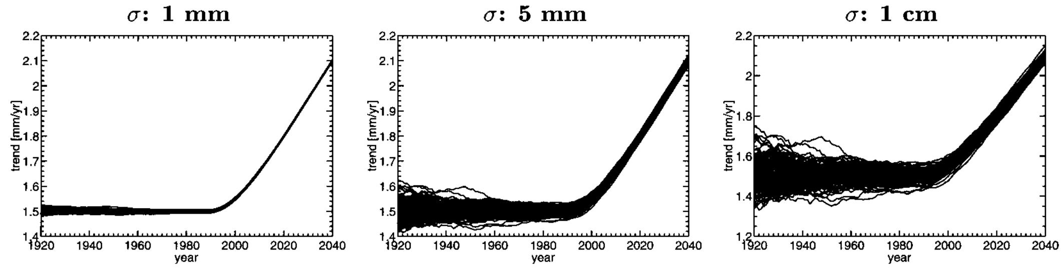

Fig. 4.15 Estimated trend versus the date up to which observations have been used. First, trend of 1.5 mm/yr; in 1991 trend abruptly changes to 4.5 mm/yr. For 100 realizations of measuring noise (3 noise levels).

4.4.4 Transition to quadratic increase at the end of the time series

It is of major importance to determine whether or not sea-level heights in the next decades increase following a linear or a higher order (e.g., quadratic or exponential) pattern. On longer time scales, the effects of a higher order increase are much more pronounced than sea-level rise corresponding to a linear pattern. Consequently, the less years of observations needed to distinguish between various patterns of sea-level rise the better.

For a time series consisting solely of a variation curve in combination with normally distributed random noise, it can be derived how many years of observations are required to distinguish between a linear rise and quadratic sea-level increase. The following parameters will be used; see also figure 4.16. After a period of Δx years, both sea-level variation curves yield an increase in sea-level height of Δsl. After a period of X_r years ($X_r > \Delta x$) will have passed, the first curve yields a sea-level increase Δsl_1, the second curve an increase in sea-level height that is β times larger, i.e., $\Delta sl_2 = \beta \Delta sl_1$.

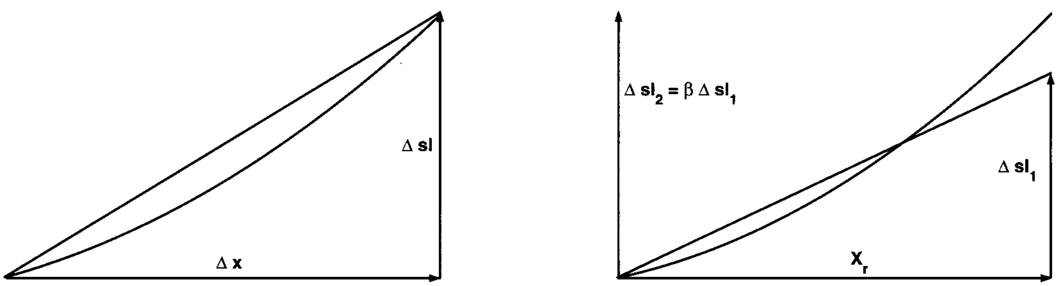

Fig. 4.16 Two sea-level rise acceleration curves. After Δx years, both curves yield the same (Δsl) sea-level rise. After X_r years, increase resulting from second curve (Δsl_2) is β times increase from first curve (Δsl_1).

In section 4.2 it has been explained that sea-level increase follows a quadratic pattern instead of a linear one, if the hypothesis $H_0 : E\{\hat{c}\} = 0$ is rejected in favour of the alternative hypothesis ($H_1 : E\{\hat{c}\} > 0$). The test statistic

$$t = \frac{\hat{c}}{\hat{\sigma}_{\hat{c}}} \tag{4.31}$$

can be used, which should follow (under H_0) a Student's distribution with $N - 2$ degrees of freedom. It can be derived, see appendix B.4, that this test statistic can be determined as:

$$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}} = \frac{\beta - 1}{X_r - N} \frac{b_l}{\hat{\sigma}_{\hat{\varepsilon}}} \sqrt{\frac{(N-2)(N^4 - 5N^2 + 4)}{180}} \tag{4.32}$$

In this equation $\hat{\sigma}_{\hat{\varepsilon}}$ is the differences between the actual values and those predicted by the quadratic model, and b_l is the linear regression coefficient.

Based on equation (4.32) it can be determined how many years of observations are required to detect a quadratic sea-level increase that would yield, e.g., after 50 years, a rise in sea level that would be 1.5 times as large as that based on the linear increase (with slope value b_l). It can be concluded that if the factor β is large, only a relatively small amount of data is required to distinguish between linear and quadratic sea-level rise. In addition, the level of measurement errors is of major importance for the number of years required to distinguish between linear and quadratic sea-level rise.

As an example, based on a measurement precision of 1 cm, it takes 20 years to distinguish between a linear trend of 1.5 mm/yr (yielding an increased sea level of only 7.5 cm in 50 years of time), and a quadratic sea-level increase resulting in a sea-level height increase of 30 cm in 50 years time. For measurements with a standard deviation of 1 mm, only 10 years of observations are needed to distinguish between these two sea-level variation patterns.

In the above it has been shown how many years of observations are, in theory, required to distinguish between linear and quadratic sea-level rise acceleration. However, it is not really known when greenhouse-gas induced warming has started (or will start) to effect sea-level heights. Consequently, after the onset of sea-level rise acceleration, first a number of observations will be required to detect that an acceleration has occurred before different acceleration patterns can be distinguished.

An indication whether or not a linear curve applies, can also be obtained by inspecting deviations between the estimated linear curve and the actual annual sea-level heights. As can be seen from figure 4.17, these errors have distinctly different patterns if the simulated time series contains two linear trends or a linear trend followed by quadratic increase. If the trend value changes abruptly to the higher value (see left-most graph in figure 4.17), errors are more or less random. If the trend value increases gradually (over a transition period of 30 years), the errors seem to follow a slowly increasing pattern. From the right-most graph, it can clearly be seen that if a linear pattern is fitted to data containing quadratic increase, this leads to large errors following a specific, parabolic, pattern.

All time series used for figure 4.17 contain measuring noise with a standard deviation of 1 cm. The year 1990 has been used to divide the time series. If different years are used to divide the time series and/or other values are used for the transition period or the linear component in case of quadratic sea-level rise, result are similar to those shown in figure 4.17.

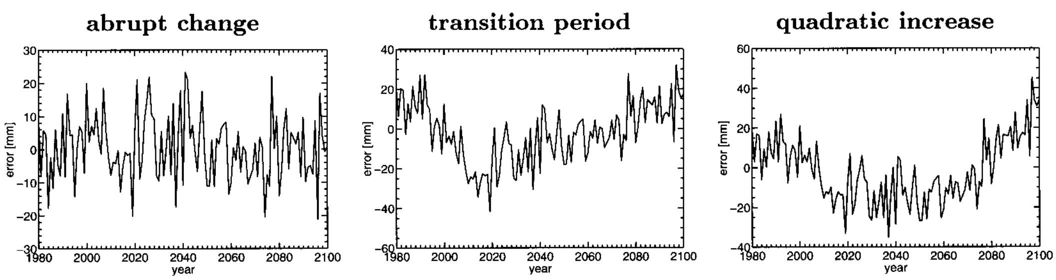

Fig. 4.17 Error between actual values and estimated linear pattern; division year: 1990. After 1990 increased linear rise (abrupt change or transition period) or quadratic increase (linear component: 3 mm/yr). Constant noise level of 1 cm.

Analogous to preceding sections, for simulated time series containing quadratic sea-level rise it will be tried whether or not the onset time of the sea-level rise acceleration can be determined based on the significance of the regression coefficients. First, for one specific realization of measuring noise, the time series are divided into two parts to which linear regression is applied separately. From the significance of the estimated trend values and the difference in trend between the two parts, no indication can be obtained of which year should be used to divide the time series.

Next, it is tried to fit a quadratic pattern to the second part of the time series. Table 4.3 gives estimated onset years of sea-level rise acceleration, based on the maximum value derived for the significance of one of the three regression coefficients, i.e., significance of the slope value for the first part of the time series, significance of the linear coefficient for the second part of the time series and significance of the quadratic coefficient. The notation > 2000 (resp. < 1980) is used to indicate that the larger the first (second) part of the time series, the higher the significance of the regression coefficient.

Results presented in table 4.3 demonstrate that the larger the linear component on which quadratic sea-level rise is based, the better the onset year can be estimated. In addition, if the quality of the measurements is very good, the onset year is estimated rather well for all types of quadratic increase under consideration.

test statistic	linear component 2 mm/yr			linear component 3 mm/yr			linear component 4 mm/yr		
	σ 1 mm	σ 5 mm	σ 1 cm	σ 1 mm	σ 5 mm	σ 1 cm	σ 1 mm	σ 5 mm	σ 1 cm
$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	1991	> 2000	> 2000	1990	1993	1999	1990	1991	1996
$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	1991	> 2000	> 2000	1990	1995	1992	1990	1992	1990
$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$	< 1980	< 1980	< 1980	1990	< 1980	< 1980	1990	< 1980	< 1980

Table 4.3 Estimated onset year of sea-level rise acceleration based on significance of trend in 1st and 2nd part of the data set and the quadratic regression term. First, trend of 1.5 mm/yr; from 1991 on, quadratic sea-level rise (linear components 2, 3, or 4 mm/yr). Constant noise level, σ: 1 mm, 5 mm, or 1 cm.

After determining the quadratic regression parameters, errors between the resulting quadratic sea-level rise and actual values should again be checked for systematic patterns. If the residuals appear strongly patterned, a different higher order sea-level increase pattern (exponential, cubic, etc.) applies to the data. In this chapter, only quadratic increase is considered, and therefore, no further experiments are conducted.

4.5 Time series containing short-periodic fluctuations

In this section, time series will be considered that contain periodic fluctuation with relatively high frequencies, i.e., periods up to, say, 5 years. These short-periodic fluctuations have been simulated based on actual tide gauge data; see section 4.3.3. No additional measuring noise is added to the simulated time series since the highest principal components have been used to simulate the periodic signal. In principle, these components contain the very high frequency oscillations and measuring noise.

4.5.1 Same linear trend throughout the time series

If the simulated time series are used to estimate a linear trend, results as shown in figure 4.18 are derived. This figure shows the estimated trend as a function of the number of observations used for the time series based on a trend of 1.5 mm/yr. Figures corresponding to the three other trend values will not be shown, since, apart from a difference in level, the pattern in the estimated trends is the same.

From figure 4.18 it can be seen that for a time series consisting solely of a linear trend in combination with short-periodic fluctuations, the value estimated for the trend stabilises after around 30 years of observations have been included. To really be able to detect the trend in the measurements series, after the trend value has stabilised a number of additional observations are needed. Consequently, it might be concluded that of the order of 45 to 60 years of observations are required to detect a trend in a time series containing high frequency fluctuations.

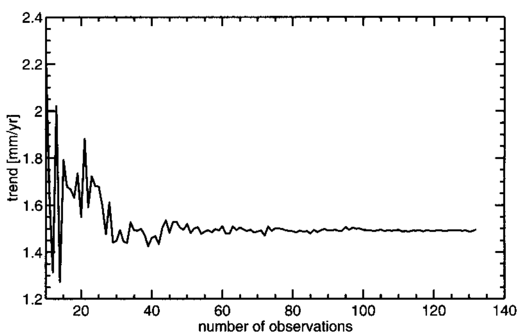

Fig. 4.18 Estimated trend versus the number of observations used (left-most estimate based on data for 1987-1996, right-most estimate on data from 1865 up to 1996). Simulated signal based on a trend of 1.5 mm/yr and relatively short-periodic fluctuations.

Due to the fact that the time series consist of a linear trend in combination with the same pattern of periodic fluctuations, the curves for the estimated trends follow the same pattern for all four trend values. In addition, the same standard deviations apply to the values estimated for the four trends. However, for different trend values, this same standard deviations has different implications concerning the number of observations required to detect this trend. An estimated trend is significant if the hypothesis $H_0 : E\{\hat{b}\} = 0$ is rejected in favor of the alternative hypothesis $(H_1 : E\{\hat{b}\} \neq 0)$. This can be tested with the statistic

$$\frac{\hat{b}}{\hat{\sigma}_{\hat{b}}} \sim t_\alpha(N-2) \tag{4.33}$$

in which $t_\alpha(N-2)$ is a Student's distribution with $N-2$ degrees of freedom, and $\hat{\sigma}_{\hat{b}}$ is the a-posteriori standard deviation for the trend. This implies (for a 95% confidence level) at least 13 measurements for a trend of 2.5 mm/yr, 15 for a trend of 2.0 mm/yr, 17 measurements are required for a trend of 1.5 mm/yr, and a trend of 1.0 mm/yr requires at least 21 measurements.

The values presented above do not contradict the conclusion that of the order of 50 observations are required to detect a linear trend in a time series containing short-periodic fluctuations. They only imply that the trend value estimated based on 25 measurements is significant. A significant trend value does not guarantee that this estimate represents the "actual" sea-level rise. After 25 measurements have been included, trend estimates may still vary significantly, not as a result of inaccuracies in the data, but due to periodic fluctuations.

As described in section 3.2 time series containing measuring noise and high frequency oscillations can be smoothed prior to analysing the data. In this section it was found that probably the best techniques to smooth sea-level height series are moving average smoothing and SSA (Singular Spectrum Analysis). These techniques will be applied in order to try to improve the estimates of the trend and shorten the time required to detect the trends in the time series. It should be remarked that smoothing introduces (additional) correlation between the smoothed values. In theory, these correlations should be incorporated in the regression model.

Moving average smoothing Since the simulated periodic signal contains mainly energy in periods up to around five years, moving average smoothing has been applied to the data with window sizes of 3, 5, or 7 years. Both rectangular and triangular windows have been applied; see section 3.2.2 for more details. The longer the smoothing window to be applied, the more data should be available. Consequently, the shortest window yielding acceptable results should be used.

If prior to the regression analysis, moving average smoothing is applied to a simulated time series containing short-periodic fluctuations and a linear trend of 1.5 mm/yr, depending on the number of observations used, trends are estimated as shown in figure 4.19. From this figure it is clear that the larger the window size used, the smoother the curve of the estimated values. However, comparison with the results derived without smoothing the data, see figure 4.18, reveals that smoothing does not (or hardly) reduces the number of observations required to detect the trend.

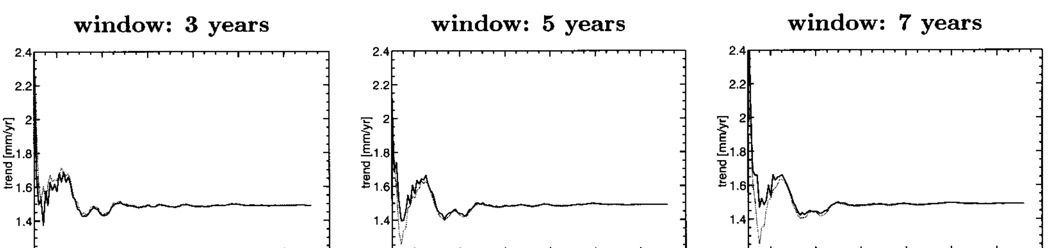

Fig. 4.19 Estimated trend versus the number of observations (left-most estimate based on data for 1987-1996, right-most estimate on data from 1865 up to 1996). Simulated signal based on a trend of 1.5 mm/yr and short-periodic fluctuations. Moving average smoothing (3 windows) is applied. grey: triangular, black: rectangular windows.

Singular Spectrum Analysis As described in section 3.2.4, the method of SSA requires the selection of a window size. After some experimenting with various window sizes it has been found that relatively small window sizes of the order of 6 to 8 years yield the best results. At least, in this case, were the simulated time series contains fluctuations with maximum periods of (approximately) five years. Results for estimated trends based on time series smoothed using SSA with window sizes of 6, 7, or 8 years are shown in figure 4.20. For each plot, reconstructions based on only the first principal component has been used as smoothed signal.

Comparison of the estimated trends as presented in figure 4.20 with those corresponding to the non-smoothed data (figure 4.18) shows that the required number of observations can be somewhat reduced by applying SSA to the time series. However, if estimates are only based on 20 to 60 years of observations, there is a significant variation in estimated value based on different window sizes.

In addition, the values around which the slope estimates stabilise are systematically higher than the actual trend present in the data. Experiments with time series containing different linear trends show that the larger the trend value, the larger the systematic error between estimated and actual trend value. The systematic error in the value around which the estimated trend stabilises is found to be around 2.5% of the actual trend value. This tendency of SSA to yield relative high values for the smoothed observations has already been discussed in section 3.3.1.

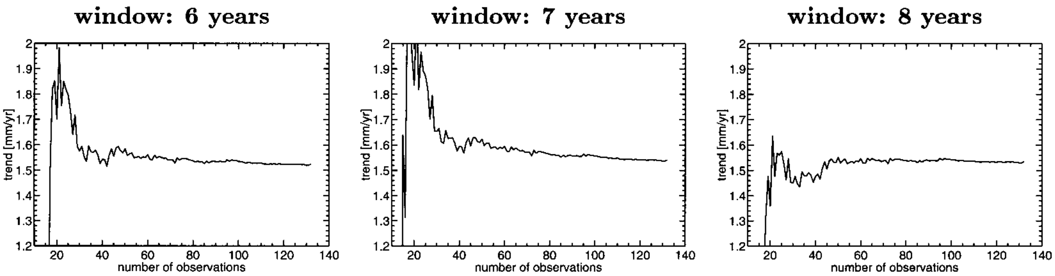

Fig. 4.20 Estimated trend versus the number of observations used (left-most estimate based on data for 1987-1996, right-most estimate on data from 1865 up to 1996). Simulated signal based on a trend of 1.5 mm/yr and relatively short-periodic fluctuations. Signal is smoothed using SSA with 3 different window sizes.

4.5.2 Transition to higher trend in the beginning of the time series

Analogous to preceding sections, it is tried whether or not it is possible to detect a change in trend value at the beginning of the time series, for a trend embedded in short-periodic fluctuations. Linear regression is applied to three time series, one with an abrupt change in slope value, and two with a gradual increase in trend over a transition period of 10 or 30 years. Resulting trends estimates are shown in figure 4.21. In this figure, the change in trend after around 100 years of observations have been included is clearly visible. Whether or not the change in slope value has occurred abruptly or over a transition period does not yield significantly different results.

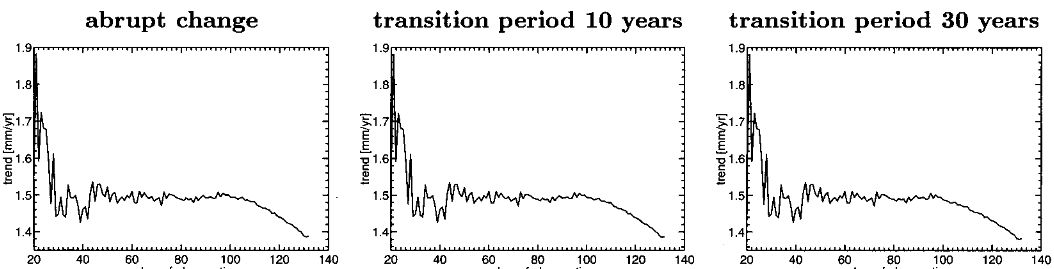

Fig. 4.21 Estimated trend versus number of observations (left-most estimate based on data for 1977-1996, right-most estimate on data for 1865-1996). First, trend of 0.7 mm/yr; increasing to 1.5 mm/yr: abruptly (in 1895), over transition period of 10 (starting 1890) or 30 years (starting 1880). Short-periodic fluctuations.

From figure 4.21, a reasonable indication can be obtained of the onset year of the sea-level rise acceleration. In section 4.2, it has been explained that this onset year can be estimated as the year that yields maximum values for the significance of the regression parameters. In section 4.4.2 it has been shown that this method only works if the quality of the measurements is very good (σ 1 mm).

For time series containing both short-periodic fluctuations and measuring noise based on data for tide gauge Den Helder it is found that the more observations the first resp. second part of the data set contain, the higher the significance of the trend value estimated for this part. Based on the significance of the difference in trend onset years 1898 (abrupt change) or 1899 (transition period) are found. These estimates of the onset year of the sea-level rise acceleration are off by a few years (1895 was used to simulate the time series). However, if these years are used to divide the time series, trends are estimated close to their actual values of 0.8 and 1.5 mm/yr. For the first part of the data set a slope value of 0.82 or 0.86 mm/yr is found, for the second part of the data set a trend of 1.48 or 1.50 mm/yr is estimated. Therefore, the method of determining the onset year of sea-level rise acceleration based on the maximum value for the significance of the regression parameters works rather well for a time series with a transition to a higher slope value in the beginning of a data set containing short-periodic fluctuations.

4.5.3 Transition to higher trend at the end of the time series

As shown in section 4.5.1, the value estimated for a trend in a time series containing short-periodic fluctuations stabilises after around 30 years of data have been included. In order to detect a change in estimated trend value, a number of "stable" years are needed prior to the change in trend. Consequently, if a change in trend occurs after around 45 to 60 years of observations have been obtained, it should be possible to detect this change in trend from the trend estimates.

In figure 4.22, estimated trends are plotted versus the year up to which observations have been included. The complete time series, starting in 1865 has been used. Consequently, if a change in trend is assumed to start in 1990, over 100 years of data containing the "old" slope value are available, and the change in trend should be visible. This is confirmed by figure 4.22, even the small change in trend from 1.5 mm/yr to 1.8 mm/yr is clearly visible in the trend estimates.

From figure 4.22 it can also be seen that the larger the increase in trend, the sooner the sea-level rise acceleration can be detected from the estimated trends. Already around the year 2010 the increased trend estimates are clearly visible if the slope value increases to 4.5 mm/yr. If the slope value increases only to 1.8 mm/yr, a lot more observations containing the "new" trend value will be required in order to detect this acceleration. This should not be a real problem, since, the smaller the increase in trend value, the smaller its impact on sea-level heights over time spans of 30 or 40 years.

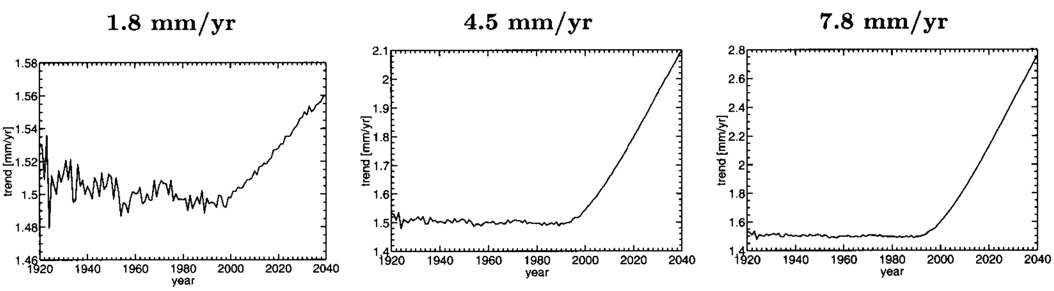

Fig. 4.22 Estimated trend versus the date up to which observations have been used. First, trend of 1.5 mm/yr; in 1991 abrupt change to 1.8, 4.5, or 7.8 mm/yrs. Short-periodic fluctuations.

Next, for different years to divide the time series, separate values for the trend and its standard deviation and significance are determined for the first and second part of the time series. Years used to split the data set range between 1980 and 2010. If the increase in trend value is only small, i.e., from 1.5 mm/yr to 1.8 mm/yr, the longer the first part of the time series, the higher the significance of its slope value. Based on the significance of the trend estimate in the second part of the data set and the significance of the difference in trend onset years of resp. 1990 and 1991 are determined.

If the change in trend is relatively large (increase to 4.5 or even 7.8 mm/yr), based on the significance of the slope value estimated for the second part of the time series and the significance of the difference in trend, the year 1990 is selected to split the data set. Based on the significance of the trend value for the first part of the time series resp. the year 1994 (increase to 4.5 mm/yr) or 1992 (increase to 7.8 mm/yr) would be selected. Based on the maximum value for the significance of the regression parameters, sea-level rise acceleration towards the end of a time series containing short-periodic fluctuations can be determined rather well.

4.5.4 Transition to quadratic increase at the end of the time series

Based on visual inspection of the annual mean sea-level heights themselves or plots of estimated trend values, linear increase (possibly with a gradual transition period) and quadratic sea-level rise at the end of the time series cannot be distinguished. Therefore, time series are divided in two parts, division years varying between 1980 and 2000. Linear regression is applied to the first part, and either linear or quadratic regression to the second part of the time series. As described in section 4.2, the onset year of the sea-level rise acceleration is estimated as the year that yields the maximum values for the regression parameters.

In table 4.4, results are summarised for estimated onset years based on the significance of the various regression parameters. This table shows that only for the time series based on the largest linear component, the significance of regression coefficients can give an indication of the onset year of the sea-level rise acceleration. For the other two data sets, the significance of the estimated regression coefficients increases with either increasing the length of the first part of the data set (indicated by > 2000) or with increasing the second part of the data set (indicated by < 1980).

Table 4.4 also shows that, for this time series, the same onset years are estimated based on linear and quadratic regression for the second part of the data set. For the significance of the quadratic regression term it holds that the larger the second part of the time series, the larger its significance.

For all three sea-level rise acceleration curves, the value determined for the quadratic regression coefficient is significant, for division years of the time series ranging between 1980 and 2000 (1995 for the curve with a linear component of 4 mm/yr). However, based on (the quadratic) regression coefficients, onset year of the sea-level rise acceleration cannot be determined. If, as a result, an arbitrarily selected year is used, estimated regression coefficients can deviate significantly from their actual values. As an example, for the time series based on linear increase of 4 mm/yr, trend estimates for the first part of the data set range between 1.42 and 1.60 mm/yr if division years between 1980 and 2000 are used.

linear component	linear regression			quadratic regression		
	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$
2 mm/yr	> 2000	> 2000	> 2000	> 2000	> 2000	< 1980
3 mm/yr	> 2000	> 2000	> 2000	> 2000	1999	< 1980
4 mm/yr	1998	1991	1993	1998	1991	< 1980

Table 4.4 Estimated onset year of sea-level rise acceleration based on significance of regression coefficients. To the 2nd part of the data set either linear or quadratic regression is applied. First, trend 1.5 mm/yr; after 1990, quadratic sea-level rise (linear components 2, 3, or 4 mm/yr). Short-periodic fluctuations.

Next it is tried whether or not results can be improved by smoothing the time series prior to analysing the data. Analogous to preceding sections, the first principal component following from SSA with window sizes varying between 4 and 8 years have been used. Estimated onset years of sea-level rise accelerations based on the significance of the various regression coefficients are presented in table 4.5. For the simulated data set based on the linear component of 2 mm/yr, only the significance of the trend values estimated for the first or the second part of the time series gives an indication of the year that should be used to divide the time series. For the two data sets based on larger linear components, significance of the quadratic regression term can be used to estimate the onset year of sea-level rise acceleration as well.

From the results shown in table 4.5 it can be concluded that smoothing the time series by means of SSA improves the detectability of the onset year of sea-level rise acceleration. For the time series based on the two larger linear components, estimates of the onset year based on significances of different regression coefficients and based on different window sizes do not vary much. However, they are systematically low, ranging between 1984 and 1989, while the actual onset year is 1991. Still, estimated regression parameters resemble their actual values reasonably well. Using again the example of estimated trends for the first part of the data set, for a time series based on a linear component of 4 mm/yr, values are estimated between 1.43 and 1.46 mm/yr (using the division years as indicated in table 4.5).

window size	linear component 2 mm/yr			linear component 3 mm/yr			linear component 4 mm/yr		
	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$
4 years	1997	1993	< 1980	1989	1989	< 1980	1989	1988	< 1980
5 years	1994	1990	< 1980	1988	1989	1985	1988	1988	1984
6 years	1988	1990	< 1980	1988	1987	1986	1988	1987	1986
7 years	1989	1990	< 1980	1988	1987	1986	1988	1987	1986
8 years	1989	1988	< 1980	1987	1986	1985	1987	1986	1985

Table 4.5 Estimated onset year of sea-level rise acceleration based on significance of (quadratic) regression coefficients. First, trend of 1.5 mm/yr; after 1990, quadratic sea-level rise (linear components 2, 3, or 4 mm/yr). Short-periodic fluctuations. Signal is smoothed using SSA with different window sizes.

4.6 Time series containing short- and long-periodic fluctuations

In this section, signals will be examined that consist of simulated trend values in combination with both short- and long-periodic fluctuations. These periodic fluctuations have been simulated based on actual tide gauge data; see section 4.3.3. Time series containing only long-periodic fluctuations will not be considered separately, since estimated trends are more or less the same with or without adding short-periodic fluctuations to a data set containing long-periodic fluctuations. No additional measuring noise will be added since the higher principal components are included in these simulated time series.

4.6.1 Same linear trend throughout the time series

For a time series containing (significant) fluctuations with periods mainly up to five years, trend estimates stabilise after around 30 years of observations have been included. In figure 4.23 trend estimated are shown based on a time series containing both short and long-periodic fluctuations. From thsi figure it can be seen that values estimated for the trend stabilise only after around 80 to 90 years of observations have been included. The pattern in estimated trend values is mainly determined by the long-term periodic fluctuations, if the short-periodic fluctuations are omitted from the simulated time series, the curves as shown in figure 4.23 hardly change. Only estimated trends based on a slope of 1.5 mm/yr are given, because, estimated trend curves based on the three other slope values are the same (except for the general trend level).

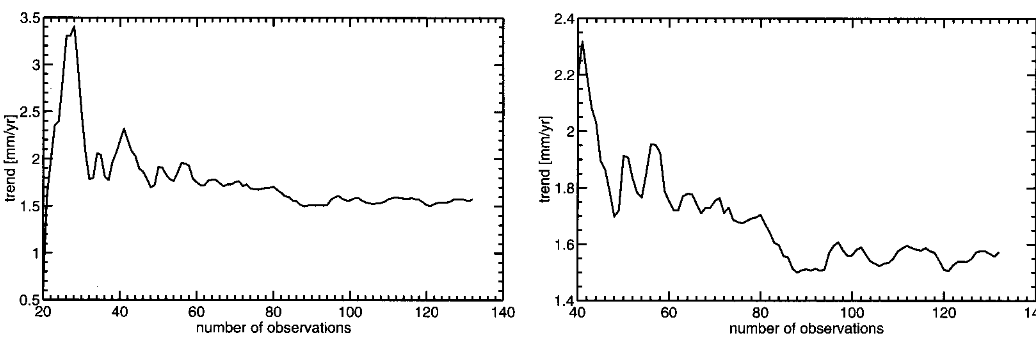

Fig. 4.23 Estimated trend versus the number of observations used. Simulated signal based on a trend of 1.5 mm/yr, short- and long-periodic fluctuations. Right-most estimate is based on data between 1865 and 1996, left-most estimate is based on data for the period 1977-1996 (left panel), or 1957-1996 (right panel).

Smoothing of the simulated time series prior to analyzing the data does not reduce the number of required observations needed for a stable (correct) estimate of the trend. For example, if SSA is used to smooth the data, depending on the window size used, curves of estimated trends become smoother, but the number of required observations is not really reduced. In addition, as already reported in preceding sections, SSA smoothed time series yield trend estimates that are systematically high.

Examining plots of estimated trends based on various window sizes, see figure 4.24, shows that the size of the systematic error not only depends on the trend in the data (as shown in section 4.5.1) but also on the window size used. Figure 4.24 shows that the larger the window size used, the larger the bias in estimated trends.

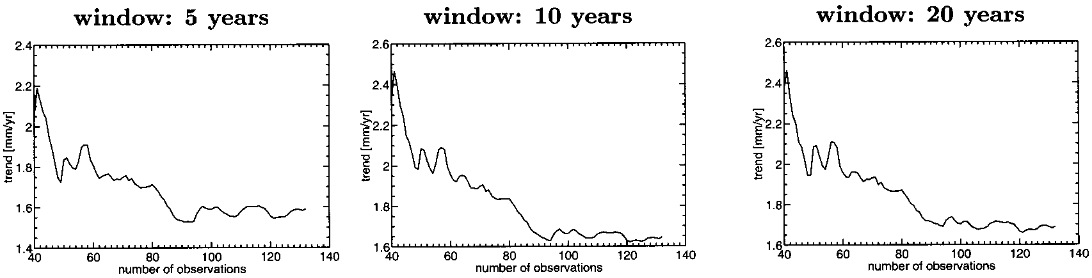

Fig. 4.24 Estimated trend versus the number of observations used (left-most estimate based on data for 1957-1996, right-most estimate on data from 1865 up to 1996). Simulated signal contains trend of 1.5 mm/yr, short- and long-periodic fluctuations. Signal is smoothed using SSA with different window sizes.

If the tide gauge has been relocated during the observation period, an example of a simulated data set is shown in the left-hand side of figure 4.25. It has been assumed that the observed pattern of short- and long-periodic fluctuations has not been influenced by the relocation of the tide gauge. Visual inspection of the sea-level height series shows that the first few observations deviate from the major part of the data set. However, the actual jump in sea level is not very clear. As a result, although it is clear that the first samples contain deviating values, the onset of the change in sea-level height is difficult to determine.

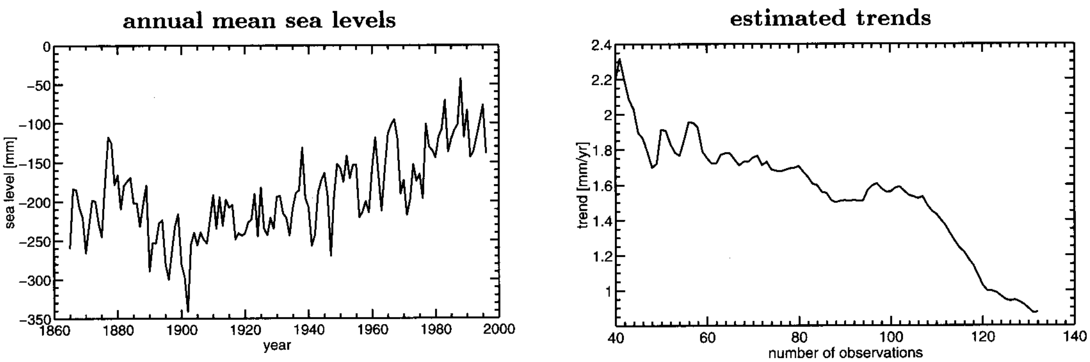

Fig. 4.25 Simulated signal and estimated trend. Based on a trend of 1.5 mm/yr, short- and long-periodic fluctuations. In 1890 a decrease (by 10 cm) in sea-level height has occurred.

The right-hand side of figure 4.25 shows estimated trends as a function of the number of observations used. After around 100 to 110 years of observations have been included, values estimated for the trend decrease. However, estimated trends stabilise only after around 90 years of observations have been used. Therefore, it is difficult to really separate the change in estimated trend due to the jump in sea level from variations in trend caused by the (long) periodic fluctuations.

More observations, obtained after the jump has occurred, might be needed to really detect the change in sea-level height from the estimated slope values. However, inspection of both the tide gauge data itself and the estimated trend values should provide a clear indication that "something has happened", although the actual identification of (the onset of) the jump might be difficult.

4.6.2 Transition to higher trend in the beginning of the time series

Visual inspection of plots showing annual mean sea-level values for time series with a 0.8 mm/yr lower trend value applying to the first 30 years of data does not reveal that the first few observations deviate from the major part of the time series. From plots of estimated trends (see figure 4.26), the change in trend value at the beginning of the time series is not really visible either.

Compared to the results based on the same trend value throughout the time series (see figure 4.23) estimated trend values are slightly lower after around 100 years of observations have been included. However, based on only the trend estimates as shown in figure 4.26, the decrease in estimated trends after around 100 years of observations have been included cannot be distinguished from the variation in estimated trend values due to the long-periodic fluctuations in the time series.

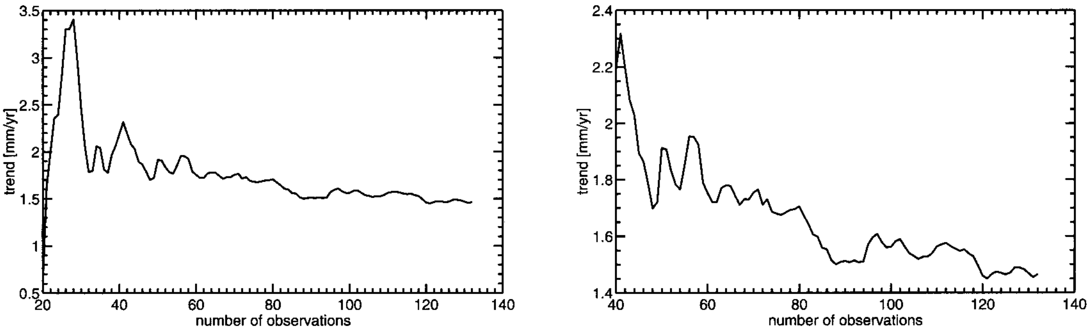

Fig. 4.26 Estimated trend versus the number of observations used. First, trend of 0.7 mm/yr; in 1895 (in figure: 102 observations) abrupt increase to 1.5 mm/yr. Short- and long-periodic fluctuations. Right-most estimate is based on data between 1865 and 1996, left-most estimate is based on data for the period 1977-1996 (left panel), or 1957-1996 (right panel).

From plots of errors between the estimated linear pattern and the actual sea-level heights the change in trend is not really visible. Splitting the data set in two time series for which separate trend values are determined does not reveal the change in trend either. Using division years ranging between 1867 and 1910, for both the first and second part of the data set it holds that the longer the time series, the smaller the standard deviation. Consequently, the significance of the difference in trend fluctuates widely, with a "maximum" value for the year 1901. However, if the first part of the time series contains observations until at least 1918, an even larger value for the significance of difference in trend is obtained.

Next, experiments have been performed in order to improve the detectability of the change in trend. First, time series is smoothed using SSA with various window sizes. Neither visual inspection of the first principal component or trends estimated for this component reveal the change in trend at the beginning of the time series.

Plots of detrended first principal components show that, if relatively large window sizes have been applied (of the order of at least 15 years) the first part of the smoothed observations deviate somewhat from the majority of the observations. As an example, for a window size of 15 years, the first 30 years of observations show a decreasing sea-level height, whereas in the remainder of the data set sea levels more or less fluctuate around zero. Unfortunately, this effect is highly dependent on the window size used, and cannot be assumed reliable enough to really indicate that a change in trend value has occurred.

As a final method to improve the detectability of the change in trend, trends have been estimated over moving windows containing only a limited number of observations. It is found that curves of trend values are highly influenced by the window size used. Even for larger windows, trends estimated over these windows fluctuate widely. If same windows are applied to time series with a trend of 1.5 mm/yr throughout the data set, comparison of the results reveals that for small window sizes the patterns are more or less the same. For larger window sizes (e.g., larger than 20 years), a deviation in estimated trends is clearly visible for the first part of the data sets. The larger the window size used, the more significant this difference in estimated trends. Unfortunately, based on only the results for the data set containing two different trend values, reliable identification of the change in trend is not really possible.

4.6.3 Transition to higher trend at the end of the time series

In figure 4.27 estimated trends are plotted as a function of the year up to which observations are included. For the larger increases in trend, this change in trend value is clearly visible from plots of the estimated slope values. However, since the estimated trend values show significant fluctuations, it is difficult to really determine the onset of the greenhouse-gas induced sea-level rise acceleration.

For the increase in trend with only 0.3 mm/yr the change in trend is not clearly visible from the estimated slope values. However, as can be seen from (the left-hand side of) figure 4.27, from the estimated trends it is clear that "something has happened" in the second half of the time series. At first, trend estimates show large oscillations. These oscillations decrease and fluctuate around a value of (approximately) 1.5 mm/yr. After around 1980 or 1990, estimated trend values show a slow, but regular increase.

For the data set with the increase in trend value to 1.8 mm/yr, a lot of measurements containing the higher trend value will be needed to detect this change in trend. Fortunatelly, even on a time scale of 50 to 100 years, the effect of a 0.3 mm/yr rise in sea level is relatively small.

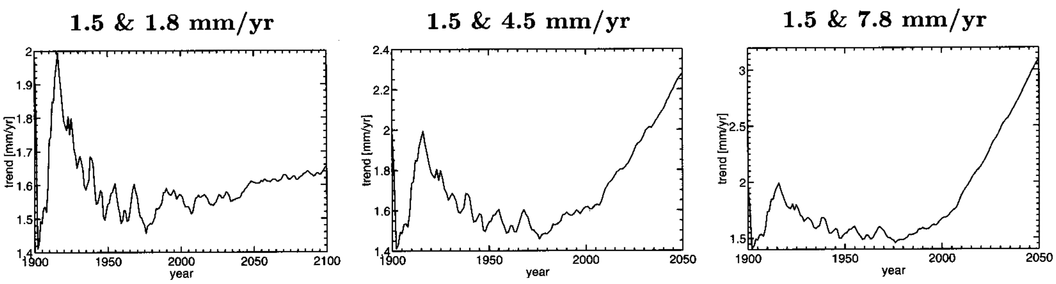

Fig. 4.27 Estimated trend versus the date up to which observations have been used. First, trend of 1.5 mm/yr; in 1991 abrupt change to 1.8, 4.5, or 7.8 mm/yr. Short- and long-periodic fluctuations.

Next, data sets are split into two separate series, using years ranging between 1980 and 2000 to divide the time series. The onset year of the sea-level rise acceleration is estimated based on the significance of the regression parameters; see section 4.2 for more details. Results for estimating the onset year are given in table 4.6. Only for the largest increase in trend value (from 1.5 to 7.8 mm/yr) it is really possible to determine the onset year of the increased sea-level rise. If the slope value increases to 4.5 mm/yr, onset years of either 1983 (based on the significance of the trend in the second part data set) or 1992 (significance difference in trend) are found. Using these years to divide the time series trends are estimated of resp. 1.4 and 1.6 mm/yr for the first part and 4.4 and 4.6 mm/yr for the second part of the data set.

Subsequently, it is tried whether or not smoothing of the time series can improve the estimates of the onset year and trend values. As an example, the first principal component resulting from SSA with various window sizes is used. From table 4.6, it can be concluded that smoothing of the time series improves the detectability of the onset year of sea-level rise acceleration based on the significance of regression parameters. However, SSA smoothing of data prior to regression introduces some difficulties as well.

The onset year determined from SSA smoothed sea-level heights seams highly dependent on the window size used. The larger this window size, the earlier the date that is indicated as most likely onset year. In addition, for different acceleration rates, onset years close to the actual value (1991) are found for different window sizes. For example, in case of a slope increase to 7.8 mm/yr, best estimate of onset years is based on original data or SSA smoothed data using a very small window size.

The above experiments have all been based on the full simulated time series with observations up to the year 2100. Of interest is how many years of observations after the onset of the increased sea-level rise (the year 1990) are required in order to detect this acceleration. Therefore, experiments have been repeated with smaller and smaller time series. Results for a number of data sets are given in table 4.7.

From table 4.7 it can be concluded that although the significance of the difference in trend decreases, same onset years and trend values are estimated if the number of observations is reduced with up to 40 years. If sea-level heights are only available up to the year 2040, based on the significance of regression parameters, the onset year of the sea-level rise acceleration can no longer be estimated. If, as a result,

window size	increase to 1.8 mm/yr			increase to 4.5 mm/yr			increase to 7.8 mm/yr		
	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$
none	> 2000	< 1980	> 2000	> 2000	1983	1992	> 2000	1990	1991
8 years	> 2000	1983	> 2000	> 2000	1988	1990	1991	1988	1988
10 years	> 2000	1983	1999	1996	1987	1989	1989	1987	1987
12 years	> 2000	1984	1997	1992	1987	1987	1987	1986	1986
15 years	> 2000	1982	1993	1989	1986	1985	1985	1985	1984
18 years	> 2000	< 1980	< 1980	1986	1984	1984	1984	1984	1983
20 years	> 2000	< 1980	< 1980	1985	1983	1983	1984	1983	1983

Table 4.6 Estimated onset year of sea-level rise acceleration based on significance of trend in 1st and 2nd part of the data and difference in trend. First, trend 1.5 mm/yr; in 1991 abrupt change to 1.8, 4.5, or 7.8 mm/yr. Short- and long-periodic fluctuations. Original data and signals smoothed using SSA (different window sizes).

an arbitrarily selected year is used to divide the time series, estimated regression parameters can differ significantly from their actual values. Therefore, it can be concluded that data up to 2050 or 2060 is needed in order to really detect a sea-level rise acceleration (starting in 1990) with an increased trend value of 4.5 mm/yr.

time series	based on	year	first part			second part			
			$\hat{b}_1$	$\hat{\sigma}_{\hat{b}_1}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\hat{b}_2$	$\hat{\sigma}_{\hat{b}_2}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{\sqrt{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})}}$
data up to 2100	1st trend	1992	1.56	0.018	*88.1*	4.41	0.019	237.7	111.2
	2nd trend	1987	1.47	0.018	84.1	4.31	0.016	*272.2*	120.0
	difference	1987	1.47	0.018	84.1	4.31	0.016	272.2	*120.0*
data up to 2080	1st trend	1992	1.58	0.019	*84.9*	4.37	0.030	147.5	80.0
	2nd trend	1985	1.47	0.018	80.1	4.20	0.023	*181.6*	92.5
	difference	1985	1.47	0.018	80.1	4.20	0.023	181.6	*92.5*
data up to 2060	1st trend	1993	1.61	0.019	*83.0*	4.50	0.060	75.2	46.0
	2nd trend	1983	1.48	0.019	76.3	4.16	0.043	*96.6*	56.8
	difference	1984	1.49	0.019	77.1	4.19	0.043	96.6	*56.8*
data up to 2040	1st trend	1994	1.69	0.021	*79.9*	4.41	0.129	34.3	20.9
	2nd trend	< 1980							
	difference	< 1980							

Table 4.7 Selected year for dividing the time series based on significance of resp. 1st trend value, 2nd trend, and difference in trend. In addition, based on these division years, trends their standard deviation and significance are given. Different amounts of data, SSA smoothing, window size: 12 years.

Next, instead of an abrupt change in trend, a gradual increase in trend over a specific transition period is applied. In figure 4.28 estimated trends are plotted based on transition periods of 10 and 30 years, both periods starting in 1991. Comparing the trend estimates with those presented in figure 4.27 leads to the following conclusions. In case of a very small increase in trend value the pattern in estimated trends is (virtually) the same for the abrupt change in trend and for the two transition periods. For the two higher increases in trend value it holds that the larger the transition period, the more observations will be needed to detect the sea-level rise acceleration.

For the above examples it has been assumed that the same trend value applies up to the onset of the greenhouse-gas induced sea-level rise accelerations. Subsequently, it will be examined if deviations in the beginning of the time series influence the detectability of a change in trend at the end of the data set. Three different sea-level height series have been simulated. One with the "pre-industrial" sea-level increase rate of only 0.7 mm/yr. For the second data set it is assumed that the tide gauge has been

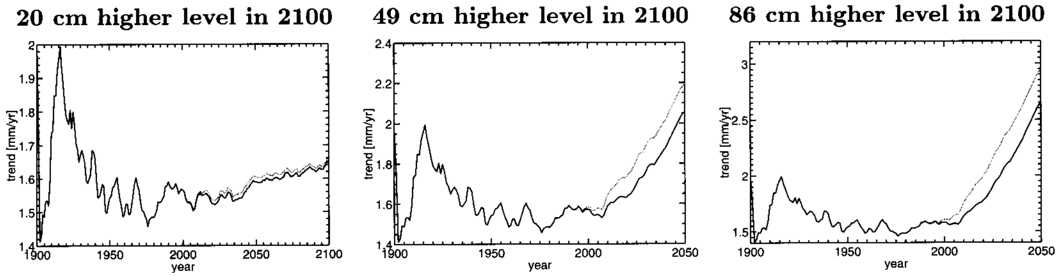

Fig. 4.28 Estimated trend versus the date up to which observations have been used. First, trend of 1.5 mm/yr; in 1991, over transition period of 10 (grey) or 30 years (black), linear increase in slope. Short- and long-periodic fluctuations.

relocated in 1890. The third simulated time series is a combination of the first two.

Visual inspection of the simulated time series with the starting trend of 0.7 mm/yr shows that from the observations themselves it cannot be deduced that the first 30 years of data are based on a smaller trend value. In addition, it is clear that the rate of sea-level rise in the last part of the data set is higher than in the first part, but the onset of this acceleration cannot easily be determined. For the data sets containing the jump in sea-level height, the greenhouse-gas induced sea-level rise acceleration with 3 mm/yr is no longer visible from the time series themselves.

If linear regression is applied to the time series, without prior removal of the deviating first observations, trends are estimated as plotted in figure 4.29. This figure shows that if the tide gauge has not been relocated (left-most graph in figure 4.29), the increase in slope value starting around 1990 is clearly visible. However, if the tide gauge has been relocated in the beginning of the time series, sea-level rise acceleration at the end of the time series can no longer be detected. The increase in estimated trend due to the greenhouse-gas induced sea-level rise acceleration cannot be distinguished from the "recovering" of the trend estimates from the relocation of the tide gauge.

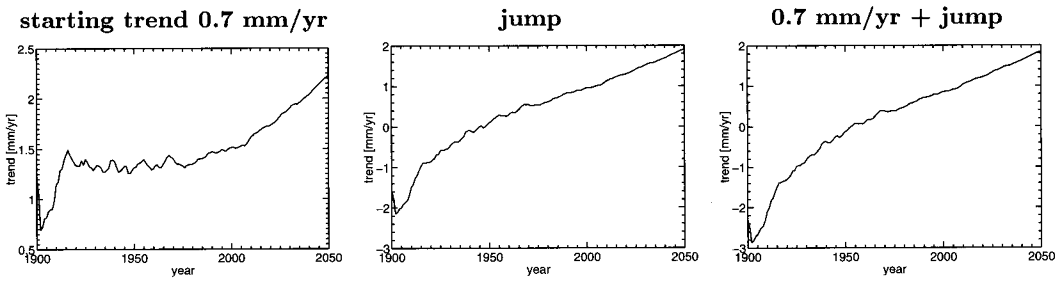

Fig. 4.29 Estimated trend versus the date up to which observations have been used. First part data: up to 1895 trend 0.7 mm/yr, or 10 cm jump in 1890, or both low trend value + jump in sea-level height. After starting period, trend of 1.5 mm/yr; in 1991 abrupt change to 4.5 mm/yr. Short- and long-periodic fluctuations.

Fortunately, for a time series containing a 10 cm jump in sea-level height, plots of the data itself already give a fair indication that the first 30 or so observations deviate from the majority of the data. Therefore, a simple solution is to omit these deviating observations from the data set prior to applying linear regression. It is not evident how many observations should be omitted. This is not of major importance as long as sufficient data is removed and the remaining time series is kept long enough for the trend values to (more or less) stabilise before the onset of the sea-level rise acceleration. This is confirmed by figure 4.30 in which trend estimates are given for time series in which observations have been omitted until resp. 1890, 1895, and 1900. Although the pattern in estimated slopes is different in the beginning, the sea-level rise acceleration at the end of the time series is clearly visible from all three graphs.

4.6.4 Transition to quadratic increase at the end of the time series

It is not possible to distinguish between quadratic sea-level rise and linear sea-level rise acceleration over a transition period, based on visual inspection of either plots of the data itself or estimated trend

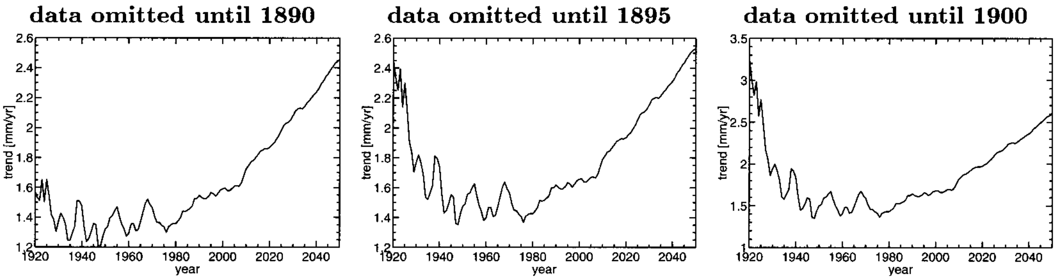

Fig. 4.30 Estimated trend versus the date up to which observations have been used. Trend 0.7 mm/yr, 10 cm jump in 1890, in 1895 abrupt increase to 1.5 mm/yr, in 1991 to 4.5 mm/yr. Short- and long-periodic fluctuations. For regression first few observations have been omitted.

values. In figure 4.31 errors are plotted between actual simulated time series (containing quadratic sea-level rise after 1990) and estimated linear patterns. Compared to results for time series without periodic fluctuations (figure 4.17), errors are significantly larger. As a result, only for the time series based on the smallest linear component (which has, therefore, the strongest quadratic component), the parabolic pattern is really visible.

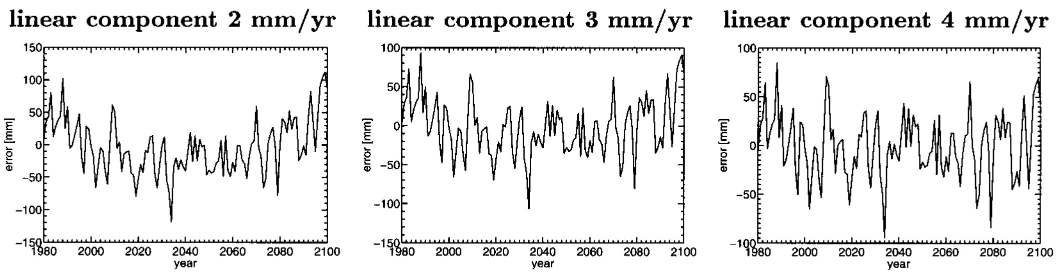

Fig. 4.31 Error between actual values and estimated linear pattern; division year 1990. After 1990 quadratic increase (linear components 2, 3, or 4 mm/yr). Short- and long-periodic fluctuations.

Next, it is tried whether or not it is still possible to determine the onset year of the (quadratic) sea-level rise based on the significance of the regression parameters. Estimated onset years based on linear regression are given in table 4.8, while results based on quadratic regression are shown in table 4.9. Either original data or time series smoothed using SSA have been used. Based on these tables, the following conclusions can be made.

- If linear regression is applied, it is only really possible to determine the onset year of the acceleration for the time series based on a linear component of 4 mm/yr.
- If quadratic regression is applied, based on significance of trend and quadratic coefficient for second part of the time series, onset year can be estimated for the time series with the smallest linear component.
- Based on significance of all three regression coefficients, onset year of sea-level rise acceleration can be determined if quadratic regression is applied to time series with linear component of 3 mm/yr.
- For the time series with a linear component of 4 mm/yr, based on the significance of the quadratic regression coefficient, onset year cannot be determined. Based on the significance of trends for first and second part of the data set, same onset year is determined based on linear regression and based on quadratic regression.

4.7 Conclusions and recommendations

Based on experiments performed in the preceding sections, with a variety of simulated sea-level height series, the following conclusions and recommendations can be made.

window size	linear component 2 mm/yr			linear component 3 mm/yr			linear component 4 mm/yr		
	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{b}_2-\hat{b}_1}{(\hat{\sigma}^2_{\hat{b}_2}+\hat{\sigma}^2_{\hat{b}_1})^{\frac{1}{2}}}$
none	> 2000	> 2000	> 2000	> 2000	1990	> 2000	> 2000	1988	1997
8 years	> 2000	> 2000	> 2000	> 2000	1998	> 2000	> 2000	1991	1994
10 years	> 2000	> 2000	> 2000	> 2000	1998	> 2000	1999	1991	1993
12 years	> 2000	> 2000	> 2000	> 2000	> 2000	> 2000	1996	1990	1991
13 years	> 2000	> 2000	> 2000	> 2000	> 2000	> 2000	1994	1990	1990
14 years	> 2000	> 2000	> 2000	1998	> 2000	> 2000	1993	1989	1990
15 years	> 2000	> 2000	> 2000	1997	> 2000	> 2000	1992	1989	1990
16 years	> 2000	> 2000	> 2000	1997	> 2000	> 2000	1993	1989	1989
17 years	> 2000	> 2000	> 2000	1997	> 2000	> 2000	1991	1988	1989
18 years	> 2000	> 2000	> 2000	1997	> 2000	> 2000	1991	1988	1988
20 years	> 2000	> 2000	> 2000	1997	> 2000	1999	1989	1987	1987

Table 4.8 Estimated onset year of sea-level rise acceleration based on significance of regression coefficients. First, trend 1.5 mm/yr; after 1990 quadratic increase. Short- and long-periodic fluctuations. Original data and signals smoothed using SSA (different window sizes). Linear regression applied to 2nd part data set.

window size	linear component 2 mm/yr			linear component 3 mm/yr			linear component 4 mm/yr		
	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$	$\frac{\hat{b}_1}{\hat{\sigma}_{\hat{b}_1}}$	$\frac{\hat{b}_2}{\hat{\sigma}_{\hat{b}_2}}$	$\frac{\hat{c}}{\hat{\sigma}_{\hat{c}}}$
none	> 2000	> 2000	< 1980	> 2000	> 2000	< 1980	> 2000	> 2000	< 1980
8 years	> 2000	> 2000	< 1980	> 2000	1996	< 1980	> 2000	1991	< 1980
10 years	> 2000	> 2000	< 1980	> 2000	1994	< 1980	> 2000	1990	< 1980
12 years	> 2000	1997	< 1980	1998	1992	< 1980	1997	1989	< 1980
13 years	> 2000	1996	1983	1997	1991	< 1980	1996	1989	< 1980
14 years	> 2000	1996	1983	1996	1991	1982	1994	1988	< 1980
15 years	> 2000	1994	1982	1995	1990	1982	1993	1987	< 1980
16 years	> 2000	1995	1981	1994	1989	1981	1992	1987	< 1980
17 years	> 2000	1995	< 1980	1993	1988	< 1980	1991	1986	< 1980
18 years	> 2000	1995	< 1980	1992	1988	< 1980	1990	1985	< 1980
20 years	> 2000	1995	< 1980	1991	1986	< 1980	1988	1983	< 1980

Table 4.9 Estimated onset year of sea-level rise acceleration based on significance of regression coefficients. First, trend 1.5 mm/yr; after 1990 quadratic increase. Short- and long-periodic fluctuations. Original data and signals smoothed using SSA (different window sizes). Quadratic regression applied to 2nd part data set.

Concerning the occurrence of periodic fluctuations Actual tide gauge data contains fluctuations with a wide range of frequencies. It is found that long-periodic fluctuations are the main factor in determining the amount of data required to detect a linear trend in a sea-level height series. As an example, if periodic information is simulated based on data for tide gauge Den Helder, of the order of 90 years of observations are required before trend estimates stabilise around their actual value. As a result:

- Long time series are required, even for detecting one single trend throughout the time series.
- If, for a time series containing data from 1865 up to 1996, the first 30 years of data are based on a 0.8 mm/yr lower slope value, this change in trend cannot be detected.
- If, after 125 years of data has been obtained, in 1990 an increase in trend occurs, the amount of (future) observations required to detect this acceleration depends on the magnitude of this increase in trend. Without prior smoothing of the data, observations up to 2100 are not sufficient to detect a small trend change from 1.5 to 1.8 mm/yr. If trend increases from 1.5 to 4.5 mm/yr (smoothed) observations up to around 2050 are required.

Concerning smoothing of time series Different techniques can be applied to smooth sea-level height series and, therefore, reduce the influence of unmodelled periodic fluctuations. From experiments with smoothing by means of SSA it can be concluded that:

- Smoothing of sea-level height data prior to linear regression will not significantly reduce the number of observations required for the trend estimates to stabilise around their actual value.
- Smoothing enhances the feasibility to determine the onset year of sea-level rise acceleration based on the significance of regression coefficients.
- SSA smoothing of data yields trend estimates that are systematically high. The bias between estimated and actual trend depends both on the window size used and the actual trend in the data.

Concerning distinction between linear and quadratic sea-level rise On longer time scales, effects of higher order sea-level rise curves are more pronounced than expected sea-level increase following from an increased linear curve. Based on experiments with simulated time series following a quadratic sea-level increase, the following conclusions can be made:

- For time series based on a small linear component, if linear regression is applied, onset year of sea-level rise acceleration cannot be determined based on the significance of regression parameters. However, (for smoothed data sets) based on the significance of the regression coefficients, the onset year can be estimated if quadratic regression is applied.
- For quadratic sea-level rise with a large linear component (and, therefore, a small quadratic component), the longer the second part of the data set, the larger the significance of the quadratic regression term. Therefore, this significance cannot be used to estimate the onset year of the sea-level rise acceleration. Based on the significance of the estimated linear coefficients, the same onset year is determined as for linear regression applied to this data set.

Concerning deviating observations As a result of a change in recording equipment or relocation of the tide gauge itself, the first part of the time series might deviate from the pattern in the remainder of the data set. Based on the complete time series, a trend value with good accuracy and significance can be found even if of the order of 25 years of deviating observations are present at the beginning of the time series. However, this value deviates from the trend representing the major part of the data set. Therefore, it is important that these deviating observations are removed from the time series prior to analysing the data. Based on simulated time series it was found that:

- The deviation in estimated trend as a result from a jump in sea-level height is only determined by the total number of observations, the number of observations containing the deviating values, and the size of the jump, and not by the trend value itself.
- If a 10 cm jump has occurred in 1890, a sea-level increase in 1990 from 1.5 to 4.5 mm/yr can no longer be discerned. Neither from plots of estimated trend values nor from significances of regression parameters.
- Although actual identification of the event is difficult, from plots of annual mean sea levels it is clearly visible that the first part of the data set contains deviating values. Fortunately, as long as sufficient data is removed (and remaining series is kept long enough) the selection of the date up to which observations are to be omitted does not significantly influence the results.

Depending on the actual trend in the time series and the level of measuring noise, only a limited number of observations are required to determine trend estimates that are statistically significant. By smoothing the data this number of observations can be reduced. However, finding a trend that is significant is no guarantee that this trend represents the "true" sea-level height variation curve. Departures from this actual variation curve can result, for example, from deviating observations in the beginning of the time series, long-periodic fluctuations or the occurrence of higher order increase patterns.

Chapter 5

Connecting tide gauges to a local height system

5.1 Introduction

Tide gauges measure sea-level height variations relative to a local tide gauge bench mark. Consequently, any height changes of these bench marks will be incorporated into the sea-level height series. In chapter 4, the detectability of specific variation curves for an individual tide gauge have been investigated. Of interest were variations in relative sea level, i.e., sea-level height movements relative to the local tide gauge bench mark. Therefore, height changes of the tide gauge itself did not need to be considered.

In chapter 6, the detectability of a common sea-level variation curve for a group of tide gauge, situated along the Dutch coast, will be examined. Any local vertical movements of the tide gauge bench marks will have introduced inconsistencies between the individual time series. In order to determine a common sea-level variation curve, these inconsistencies have to be removed. This can be achieved by connecting the tide gauge bench marks in height.

In the past, usually no direct height connections were made between tide gauges. However, often tide gauges bench marks have been connected to the local height system. As a result, a common variation curve in sea-level heights relative to this local height system can be determined. However, inaccuracies in the height connections will influence the quality of the sea-level height series relative to the local reference frame. Therefore, in this chapter the height connection to the local reference frame will be considered.

On the subject of connecting tide gauge bench marks a large variety of (workshop) reports and articles have appeared over the last few decades. A substantial amount of research has been conducted and is still in progress on the use of different geodetic techniques for establishing height connections between tide gauges and local, regional and global height networks. It is beyond the scope of this thesis to actually investigate the various measuring techniques and to try to obtain an ideal measuring approach for connecting tide gauges. Instead, existing literature on this subject is reviewed in order to obtain insight in attainable quality of height connections between tide gauges and a local reference frame. This, in order to qualify the effect of the height connection on the quality of sea-level height series as obtained by local tide gauges.

Before addressing the subject of connecting tide gauge bench marks to a local reference frame, it should be remarked that the quality of a sea-level height series relative to a height datum is influenced by the height connection between the "zero" of the tide gauge and the tide gauge bench mark as well. The tide gauge system most commonly used during the first part of tide gauge operations, and still in use in a lot of countries today, is the stilling-well tide gauge. This type of tide gauge has been described in section 2.2.3. As can be seen from figure 2.3, the "internal" height connection consists of two parts, i.e., the relation between the gauge zero and the contact point and the height difference between this contact point and the tide gauge bench mark.

In manuals on sea-level height measurements, e.g., IOC (1985), it is often remarked that these height differences should be checked on a regular basis in order to detect errors in calibration and, among others, variations in the length of the float suspension. According to Dillingh (1998), if a systematic error is found while checking the gauge zero, the difference in height is prorated over the measurements up to the time of the last check of the tide gauge zero.

No indication is given of the quality of the actual height differences determined between gauge zero and tide gauge bench mark. As a result, although the measuring error should be almost negligible small due to averaging over an enormous amount of measurements, usually a standard deviation of 1 cm is assumed for annual mean sea-level heights relative to the tide gauge bench mark.

On the subject of relating the zero-height point of a subsurface pressure tide gauge to the local tide gauge bench mark research has been conducted by Woodworth *et al.* (1996). They conclude that the precision of this connection will, in general, be of the order of only 2 cm. A new method, proposed by Woodworth *et al.* (1996), should be able to provide the height between the zero of the tide gauge and the tide gauge bench mark with a precision of the order of only 1 mm.

Traditionally, tide gauge bench marks have been connected to the local height datum by spirit levelling. In most countries in the North Sea area, starting somewhere in the second half of the 19th century, (re)levellings and subsequent adjustments of the national height networks have been performed every few decades. Nowadays, it is often proposed (e.g., by Carter *et al.* (1989)) to interconnect tide gauges using space geodetic techniques, possibly in combination with gravity measurements.

One of the main advantages of using GPS to connect tide gauge bench marks as compared to spirit levelling is the, relatively, low costs and high speed with which the connections can be performed. Using GPS, heights of a number of tide gauge bench marks can be interconnected in such short time intervals that it is probably safe to assume that the area is not influenced by, for example, subsidence. Levelling, on the other hand, is very time consuming and tide gauge bench marks might subside during the completion of the whole levelling network.

Although GPS is the proposed method of choice (e.g., by Carter *et al.* (1989)) for connecting tide gauge bench marks, it is not advisable to focus only on GPS and disregard levelling, mainly because of the following two reasons. First of all, GPS is a relatively new technique that can only be used to improve the height connection of recent and future sea-level height series. For the comparison of the many old tide gauge records available, results obtained by various (re)levellings and adjustments of local levelling networks have to be considered. Secondly, space geodetic techniques and spirit levelling result in fundamentally different heights.

Different scientists have different views on the subject of which type of height connection, orthometric or geometric, should be preferred for sea-level studies based on tide gauge data. Therefore, prior to describing the various geodetic measuring techniques, in section 5.2, advantages and problems related to the different height systems will be addressed.

In section 5.3, for a number of geodetic techniques that can be used for the vertical connection of tide gauge bench marks, an overview will be given of the expected precision of the determined heights and systematic errors that can occur. In section 5.4 the selection of observation sites will be briefly addressed. Of interest are maximum distances that are allowed between stations and problems originating from using stations at specific locations, e.g., near the coast. Finally, in section 5.5, required sampling of measurements will be discussed.

Before addressing the topics as described in the preceding, it is useful to consider what accuracy we would like to achieve with the height connection. The simplest answer is of course, the more accurate the height connections, the better. In one of the latest workshops (see Carter (1994)) held on the subject of connecting tide gauge bench marks it has been concluded that over 5 year intervals, vertical motions should be available with a standard deviation of 1 to 2 millimetres/year. Over time intervals of a few decades, this precision should be of the order of 0.3 to 0.5 mm per year. However, preventing (or correcting for) systematic errors in the height connection is probably of even more importance than obtaining a sufficient measuring precision.

In the following, the expression orthometric height will be used to indicate height differences obtained by spirit levelling (in combination with gravimetry). In reality, a number of countries in the North Sea area (like France and Germany) have adopted local height systems based on normal heights. In addition, different orthometric heights (based on different methods to estimate mean gravity) are in use, or have been used in the past. However, in principle, all these different types of heights can be converted into (one specific type of) orthometric heights. Therefore, only orthometric heights will be mentioned even when it would be more correct to distinguish different types of heights.

5.2 Problems related to the different "height" systems

The different geodetic techniques as will be described in this chapter, can be divided into three categories, based on the type of "heights" they produce. Space geodetic techniques yield geometric height differences. Based on levelling (in combination with gravimetry), heights relative to an equipotential surface can be obtained. Finally, gravity measurements provide information concerning the combined effect of changes in mass and height.

Space geodetic techniques On of the advantages of using space geodetic techniques for connecting tide gauge bench marks is that all bench marks heights can be related to the same global geodetic reference frame. As a further advantage this facilitates the comparison between tide gauge data and results obtained from satellite altimetry.

Space geodetic techniques yield geometric height differences while some applications of tide gauge data might require a connection in orthometric height. The question whether or not a connection in geometric height is sufficient for sea-level variation studies has been the subject of a lot of debate. For this thesis it has been preferred to connect tide gauges in orthometric height, a choice that will be clarified in section 5.2.1. Orthometric heights can be obtained from geometric heights if the geoid height at the station is known. An in-depth treatment of the accuracy and problems related to geoid heights is beyond the scope of this chapter; see, e.g., de Min (1996). However, in section 5.2.3 some specific problems related to sea-level variation studies will be discussed.

Spirit levelling Levelling in combination with gravimetry results in potential differences between stations. These potential differences can be converted into heights by dividing them by a gravity value. Since levelling only yields height differences, a specific equipotential surface should be used as a reference for the levelling network. Often heights are related to (some kind of approximation of) the geoid.

In different countries, over the years different height systems have been developed by applying a specific approach to the selection of the gravity values to be used. Examples are: dynamic heights, (different realisations of) orthometric heights, and normal heights. Since so many different approaches exist, often introducing errors due to specific approximations used, the UELN levelling network is based on original geopotential differences relative to NAP instead of converting them into height values.

In theory, levelled "heights" are relative to the geoid. Since, this is a reference surface that cannot be used directly, in most countries some kind of approximation has been introduced. At a specific tide gauge (or worse, a group of tide gauges) sea-level height variations have been measured over a certain period of time. The mean sea level derived from these observations is used as an approximation to the geoid.

Due to various processes (like changes in salinity, wind forcing, etc.) mean sea level is in reality not an equipotential surface. Consequently, tide gauges at different locations will yield different reference surfaces. As a result, constant offsets exist between the reference surfaces (vertical datums) as used in different countries. In order to compare tide gauge data for tide gauges located in different height datum zones, the height differences between the various reference surfaces have to be determined. A procedure for vertical datum connection will be described in chapter 7. In this section, only the connection of tide gauges to a local height datum will be considered.

An additional problem that has to be considered in relation to levelling networks is the occurrence of secular height variations and how they have been treated in the definition of the local reference frame. As an example, to obtain the UELN80 adjustment, levelling data in the Fennoscandian region has been reduced to the common epoch 1960. Unfortunately, different reduction models have been used in the four Nordic countries; see e.g., Ehrnsperger *et al.* (1982) for more details on the reduction methods used.

Changes from one datum to another complicate the interpretation and comparison of tide gauge data. Consequently, sea-level height accelerations, e,g, due to greenhouse-gas induced warming will be more difficult to detect. This is not only a problem in the Nordic countries, but also in other countries where a change in reference surface has been introduced between different adjustments of the levelling networks.

Gravimetry Both variations in height and variations in mass cause changes in local gravity. Therefore, if the variation in mass is known, gravity measurements can be used to monitor vertical movement of tide gauge bench marks. Consequently, in order to use gravity measurements for determining height changes, some kind of assumption has to be made concerning corresponding changes in mass.

Since actual vertical movements consist of both a change in orthometric height and a change in geoid height, in order to determine variations in orthometric height the change in geoid height has to be known as well. However, according to Warrick *et al.* (1996), in studies of sea-level changes on time scales of less than a century, the change in geoid height is usually one order of magnitude smaller than the corresponding change in station height, and can (often) be neglected.

The relation between estimated change in height and measured change in gravity is determined by the selected gravity gradient factor. The vertical gravity gradient is a combination of the free-air gravity gradient (due to displacement of the station) and subsurface density field variations (due to influx of mass). Consequently, a change in height corresponds to a change in gravity somewhere between the following two extremes:

- No change in mass has occurred near the observation point. In this case the gravity gradient equals the free-air gravity gradient, which can be approximated by -0.3086 mGal/m.
- The change in height is due to influx of mass, in an extreme case by an addition of a Bouguer plate with density ρ (in g/cm^3). The gravity gradient can be estimated as (-0.3086 + 0.0419 ρ) mGal/m. Using a standard crustal density $\rho = 2.67$ g/cm^3 this becomes -0.1967 mGal/m.

However, based on theoretical considerations Jachens (1978) concludes that the range of values the gravity gradient can assume should be much larger, because variations in subsurface density can not only subdue the free-air gravity gradient but also enhance or even dominate this normal value. Jachens (1978) mentions an example where a certain amount of groundwater can be extracted before significant subsidence starts to occur.

Based on the above it can be concluded that great care has to be taken if variations in gravity have to be converted into variations in station height. Derived vertical movements contain a large uncertainty due to uncertainties in the value used for the vertical gravity gradient. On the other hand, gravity measurements can be a useful tool to discern between observed sea-level height changes that are due to actual variations in absolute sea level and those due to changes in height of the tide gauge itself. Except in case of extreme subsurface density variations, the effect of subsidence of the tide gauge will be much larger than the effect of an additional layer of water. Directly on top of the layer of water (and depending on its density) the corresponding gravity gradient is only about 0.04 mGal/m.

5.2.1 Orthometric heights versus geometric heights

Often it is difficult to place the GPS antenna right on top of the tide gauge bench mark, because of instrumental problems, interference with other signals or multipath problems. In this case, the height connection between tide gauge and (as nearby as possible) GPS site, is achieved by spirit levelling. The justification for this method is not easily found in the articles and reports on this subject, but the reasoning is probably as follows. First of all, only changes in height of the tide gauge during the sea-level variation observation period and height changes of the tide gauges relative to one another are important. In addition, because the distance between tide gauge and GPS site is kept as short as (practically) possible, it is assumed safe to mix GPS derived heights with levelled height differences. This is probably true, at least as no, significant, change in geoid height occurs between the tide gauge and GPS station.

To monitor seasonal and interannual variations in individual tide gauges or groups of tide gauges and to connect tide gauge data with satellite altimetry missions with durations of a limited number of years, a geometric connection of tide gauge heights might be sufficient. However, for studying long term variations in sea-level height and possible changes in the "natural" patterns due to, e.g., greenhouse-gas induced warming, very long tide gauge records are required.

As long as the height of a tide gauge shows, on average, only a secular variation, to detect an acceleration in relative sea-level increase for an individual tide gauge, no height information is required. If, however, sudden changes in height have occurred or the rate of vertical movement of the tide gauge changes over the years, information concerning the height of the tide gauge is required for a reliable interpretation of the sea-level variation pattern.

In most countries in the North Sea area, starting somewhere in the second half of the 19th century, tide gauge bench marks have been connected to the national height system by means of spirit levelling. Therefore, for most tide gauges in this area, at least some information concerning their orthometric height relative to the local datum is known, or can at least be derived from, e.g., available normal heights. For tide gauges partaking in various levellings, an indication can be given of the height variation of the tide gauge bench mark over the years.

Prior to the era of space geodetic techniques, geometric heights were not usually determined and the only height data obtainable are orthometric heights relative to the local height datum. Since accurate geoid heights corresponding to historical levellings are not available, to convert these orthometric heights into geometric heights, present day geoid information has to be used. This requires a stationary geoid for time spans of 100 to 150 years. This is (at the centimetre to sub-centimetre level) not a valid assumption. As an (extreme) example, Ekman (1989) estimates a maximum uplift of the geoid of 0.7 mm/yr due to post-glacial rebound.

A final argument in the discussion whether orthometric or geometric heights should be used is the purpose to which the obtained sea-level height series will be applied. As indicated in the preceding, for correcting relative sea-level height series for local height changes, geometric heights can be used. The resulting sea-level heights relative to the global geocentric reference frame can be used to estimate patterns and variations in patterns of absolute sea-level heights.

For investigating oceanographic effects of, e.g., greenhouse-gas induced warming, variations in sea-surface topography are of more importance than variations in absolute sea level. As explained by, e.g., Mitchum (1994), oceanographers use the geoid as reference surface for their hydrodynamic equations.

5.2.2 Limitations of heights derived in the past

Although, in theory, orthometric heights derived at a number of occasions are available for at least some tide gauges in the North Sea area, there are limitations to their usefulness. A large number of shortcomings of, especially the older, orthometric heights could be listed, but only a few of the more evident ones will be mentioned. First of all, in the past, levellings have often been carried out without additional gravity measurements. Consequently, only height increments have been obtained instead of "true" orthometric height differences. However, for areas like the Netherlands, with very small horizontal gravity gradients, this does not cause problems at the millimetre-level.

A second problem is the accuracy of the older levellings (especially the ones performed in the 19th century), which might not be compatible with those performed over the last few decades. Although internal precision of historical levelling networks is usually very good, questions can be raised concerning the reliability of the height information.

Only orthometric height connections relative to the local height datum have been performed. Scant information is available concerning height differences between the various height systems in use in the area. Even for local height connections, the number of occasions on which tide gauge bench mark heights have been determined is very limited. A major restriction to the usefulness of historic height differences is that usually no additional information concerning this height change is available. It is unknown whether a change in height found between two adjustments is due to:

1. a secular change in height, e.g., due to subsidence or post-glacial uplift
2. an abrupt change in height e.g., due to construction activities in the area of the bench marks
3. relocation of the tide gauge and tide gauge bench mark
4. a change in the reference surface of the local height datum
5. a change in levelling and/or adjustment procedure

Although the number of limitations to the usefulness of orthometric heights obtained in the past is endless, this is the only height information available. Converting them into geometric heights will make them even less reliable. Therefore, any orthometric height information available should be used, but with a clear picture in mind of their limited reliability.

5.2.3 Problems related to the geoid

As described in the preceding, space geodetic techniques provide a fast, and relatively cheap, method for determining station heights. The resulting geometric heights can be converted into orthometric height if the geoid height of the station is known. There are a number of possibilities to obtain the geoid height at a station, but most obvious are using a local geoid model available for the specific region or calculating the geoid height separately for each station, based on a geopotential coefficient model, preferably in combination with terrestrial gravity measurements in an area surrounding the station.

A substantial amount of geoid models are available, providing an estimate for the local geoid, the geoid in a specific region, or even a global geoid. In general it can be stated that the quality of the geoid heights provided by these models depends (apart from the quality of the global model) on the quality of the topography and local surface gravity data included. As a result, geoid models will not have a uniform precision, but local differences in quality occur depending on the local gravity information available. As an example, for the global OSU91A model Rapp and Balasubramania (1992) estimate a precision of 26 cm for ocean areas, a standard deviation between 38 and 56 cm in land areas with surface gravity data (depending on the quality of the terrestrial gravity measurements), and a precision of 200 cm for land areas where no terrestrial gravity information is available.

de Bruijne *et al.* (1997) have derived a geoid for the North Sea area (the GEONZ97 model), based on a global geopotential model (EGM96) in combination with local gravity data and additional measurements obtained by GPS, levelling and altimetry. Resulting geoid heights are shown in figure 5.1. Estimated standard deviation of the residuals between the GEONZ97 model and the data used (both NEREF points and TOPEX data) is 3.2 cm.

Figure 5.2 shows, for the overlap area of the Netherlands, differences between the GEONZ97 geoid and the so-called "De Min geoid" (see de Min (1996) for more details). This figure demonstrates that although both geoid models provide geoid heights with estimated standard deviations of a few centimetres, significant differences between the two models exist, especially in areas near the coast and close to the German border.

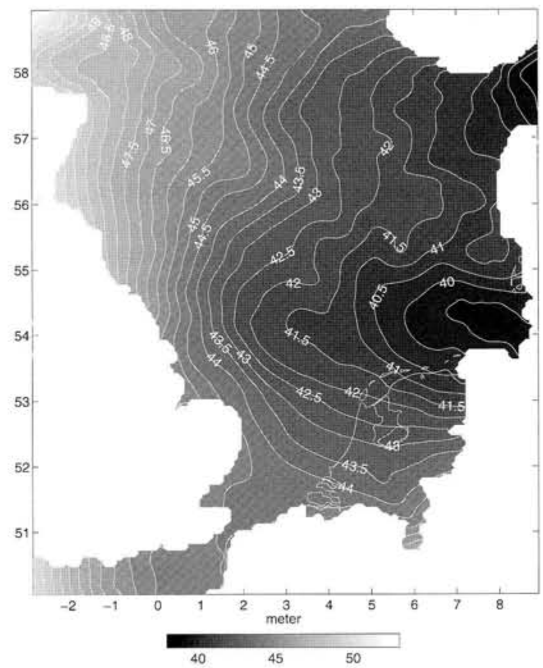

Fig. 5.1 Preliminary geoid for the North Sea area (GEONZ97); reproduced from Haagmans *et al.* (1998).

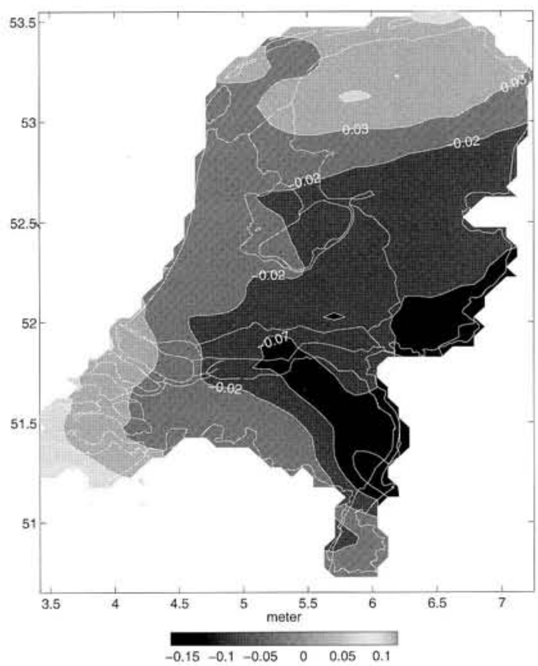

Fig. 5.2 Difference between preliminary geoid (GEONZ97) corrected to NAP and the "De Min geoid" for the Netherlands; reproduced from de Bruijne *et al.* (1997).

On the subject of geoid computation, problems related to the "geoid concept" and the feasibility of combining GPS measurements with geoid heights to obtain orthometric heights, a large amount of literature has been published over the years. In this section, only a few items of interest for the detectability of sea-level variations from groups of tide gauges will be mentioned.

The potential of the geoid is not known. Consequently, there is no unambiguous relation between the geoid and its reference ellipsoid. In addition geoid undulations are affected by the (unknown) differences between the centre of mass of the reference ellipsoid and the centre of mass of the earth.

The geoid is not a stationary surface but shows periodic movements and secular variations. For example, due to post-glacial rebound, in the northern part of the Gulf of Bothnia the geoid rise can amount to 0.7 mm/yr; see e.g., Ekman (1989). As a result of solid earth tides, the crust shows vertical deformations with an amplitude of around 30 cm, the variations of the equipotential surface are similar but its amplitude is about twice as large. A number of processes, resulting from greenhouse-gas induced warming, will effect the geoid as well. It can be expected that melting of glaciers and small ice caps and reduction of the major ice sheets will induce local changes in geoid height.

Another problem is that, due to the sun and moon moving in orbits close to the equator, a part of the tides does not vary periodically. Its average value is latitude dependent and is called the permanent tide. As described by Ekman (1993) depending on how this permanent tide is handled, different geoids result. Unfortunately, in different countries, the permanent tide has been treated in different ways. As explained by Ekman (1989), it is important that for all measurement systems applied (levelling, gravimetry and space geodetic systems) the permanent tide is handled the same way. As an example, Ekman (1989) estimates an error of 9 cm if tidally corrected GPS heights are converted into orthometric heights using the mean geoid instead of a non-tidal geoid. For an in-depth discussion on problems related to using terrestrial gravity anomalies, see Heck (1990).

Except for a few stations situated well inside large datum zones, terrestrial gravity information obtained by different countries and/or institutes will have to be used. In addition, for stations in coastal zones, gravity information for the ocean area will be required. Combining gravity data obtained from different sources can lead to some complications in the geoid height determination. First of all, between different countries, the quality and data density of the surface gravity data may show large variations. Secondly, even if gravity data sets are of uniform quality, they will have been reduced to different reference surfaces, i.e., to the specific equipotential surface used in the various countries. In a worse case, the reference surface used is not even an equipotential surface, e.g., because it is based on mean sea level for a group of tide gauges.

5.3 Error characteristics of measuring techniques

In this section, some specific characteristics of the measuring techniques will be described, and an indication of their expected measuring precision will be given. Precision estimates refer to the "height" usually associated with the measuring technique under consideration, i.e., geometric heights for space geodetic techniques and gravity accelerations for gravity measurements. It will not be tried to propagate these values into standard deviations for differences in orthometric height.

The resulting accuracy of station heights is not only determined by the formal error budget, but by systematic errors as well. Therefore, an indication will be given of systematic errors hampering the geodetic measurements. These errors are divided into two groups, i.e., systematic errors inherent to the measuring process itself and systematic errors caused by the "outside world". With this last group all kind of geophysical signals (e.g., solid earth tides), oceanographic signals (e.g., ocean load tides), atmospheric signals (e.g., atmospheric load tides) and local disturbances (e.g., variation in ground-water level) are meant. Some problems, specifically related to the selection of the station location (e.g., near the coast) will be discussed in section 5.4.

The geodetic measuring techniques considered in this chapter are: spirit levelling, GPS, and absolute and relative gravimetry. In general it can be stated that levelling is a time consuming and, therefore, expensive technique. However, levelling more or less supplies orthometric heights, without the additional errors introduced by geoid models or assumptions concerning mass changes associated with measured gravity changes. Furthermore, heights derived more than a few decades ago have always been based on levelling. In addition, even when space geodetic techniques are used for the major part of the height connection, it will often be necessary to connect the space geodetic stations to the actual tide gauge bench mark by means of levelling.

GPS yields fast and relative cheap station height differences. Furthermore, it is now possible to place permanent GPS stations at or near the tide gauge bench mark, thus providing continuous monitoring of vertical movements of the tide gauge. VLBI yields precise baseline estimates over very large distances. However, since it receives signals in the microwave range, it suffers from the same problems concerning, e.g., tropospheric path delay, as GPS. SLR on the other hand, operates in the optical range and has, therefore, different signal propagation characteristics. In addition, SLR provides coordinate estimates in a reference frame that, in theory, coincides with the centre of mass of the earth.

It is often stated (e.g., by Carter *et al.* (1989)) that gravity measurements can provide an independent method to monitor changes in height of tide gauge bench marks. Unfortunately, gravimeters sense the combination of height and mass changes, and mass changes are difficult to estimate. Consequently, it might be difficult to convert gravity variations into height changes. However, gravity observations can be used to distinguish between relative sea-level rise due to increase in absolute sea level and a rise due to subsidence of the tide gauge; see section 5.2.

Usually, determined heights refer to a specific epoch. This is only valid if during the time interval over which the underlying observations have been made, no (significant) changes in height have occurred. Since the earth is non-rigid, no matter how short the observation period, some vertical motion will always take place. Following Mather (1974) these motions can either be included (averaged out) in the definition of the heights, or they have to be modelled and eliminated from the observations prior to estimating the heights. Subsequently, estimated height changes between epochs are the result of secular variations (like subsidence) or abrupt changes (like earthquakes).

5.3.1 Levelling

Traditionally, stations have been connected in height by means of spirit levelling. Starting sometime in the second half of the 19th century, major parts of the countries in the North Sea area have been covered by levelling measurements. Relevellings and, subsequent, adjustments of first order networks have been performed every few decades. Starting in 1955, combined adjustments for a number of levelling networks in Europe have been derived. The first United European Levelling Network, UELN-53, was followed by UELN-73 and UELN-95, in which more levelling data covering a larger area has been included.

For a single connection in height of a tide gauge bench mark into the local reference system, one levelling is sufficient. However, the more often the height connection is performed, the better height variations of the tide gauge bench mark can be modelled, and future changes can be predicted. Unfortunately, in subsequent first order levellings often (slightly) different stations have been occupied. This can be a result of a change in policy (the decision that stations should be part of a primary levelling network) or simply due to the limited lifetime of the bench marks. Rietveld (1986) estimates that on a yearly base, between 2% and 5% of the total number of bench marks in the Netherlands disappears. As a result, the first order levellings conducted over the last century have only a small amount of bench marks in common.

As indicated in the preceding, (first order) levelling is very time consuming. As a result, the time required to complete the whole levelling network may be almost equal to the interval between subsequent relevellings of the network. As explained by Barbarella and Radicioni (1990), if the area under consideration is subject to significant vertical movements this may lead to problems with closing the levelling loops. Furthermore, the time interval between the relevellings will be different for the various stations. This complicates the interpretation of height differences following from subsequent adjustments of the levelling networks.

A European levelling network like UELN is based on levelling measurements obtained in completely different epochs, covering areas with significant vertical movements (e.g., the Scandinavian region). Consequently, in order to combine these measurements, velocity estimates of the stations should be considered as well; see e.g., Augath (1991) for more details. Unfortunately, different reduction models have bee used in the Scandinavian countries to compensate for post-glacial land uplift, while in most other countries no reduction has been applied at all.

Precision of levelled height differences In general, it is assumed that the standard deviation of levelled height differences can be described by:

$$\sigma_{\Delta H}(\mathrm{mm}) = \alpha\sqrt{S} \tag{5.1}$$

with α the unit standard deviation and S the levelling distance in kilometres. For, recently measured, first order levelling networks, a value $\alpha \approx 0.6$ should be feasible. This is a measure of the internal precision of the levelling networks. Even for the oldest primary levelling this internal precision is very good. As an example, Murre (1985) quotes a value of 0.75 mm $\sqrt{\text{km}}$. However, due to a large number of systematic influences, the accuracy of heights derived in the older networks might not be very good.

Levelling is influenced by a large number of (unmodelled) systematic effects, their influence increasing with increasing levelling distances. For levelling connections performed over only short distances, the square root law is probably adequate to estimate the quality of levelled height differences. For larger distances, the influence of systematic errors becomes more and more pronounced, and, as a worst case scenario, the standard deviation of levelled height differences might follow:

$$\sigma_{\Delta H}(\text{mm}) = \alpha S \tag{5.2}$$

see Vaníček and Krakiwsky (1986) for more details. Consequently, errors in estimated heights, especially those based on older networks, might be of the order of many tens of centimetres.

Internal precisions of (most) levelling networks are known. For example, for the first-order levelling networks in the Netherlands, the following values for the internal precision of the complete network (a-posteriori standard deviations) have been estimated; see Murre (1985) for more details on the networks and adjustment techniques applied.

- 1st primary levelling, measurements obtained between 1875 and 1885, $\hat{\sigma} = 0.75$ mm$\sqrt{\text{km}}$
- 2nd primary levelling, measurements obtained between 1926 and 1940, $\hat{\sigma} = 0.72$ mm$\sqrt{\text{km}}$ (original adjustment), or $\hat{\sigma} = 0.86$ mm$\sqrt{\text{km}}$ (based on new adjustment, conducted by Murre (1985))
- 3rd primary levelling, measurements obtained between 1950 and 1959, $\hat{\sigma} = 1.1$ mm$\sqrt{\text{km}}$ (original adjustment), or $\hat{\sigma} = 1.08$ mm$\sqrt{\text{km}}$ (based on new adjustment, conducted by Murre (1985))
- 4th primary levelling, measurements obtained between 1965 and 1978, $\hat{\sigma} = 1.00$ mm$\sqrt{\text{km}}$
- 5th primary levelling, measurements obtained between 1996 and 1999, $\hat{\sigma} = 0.72$ mm$\sqrt{\text{km}}$ for the optical levellings, and $\hat{\sigma} = 0.2$ mm $+ 0.25$ mm $\cdot$ km for the hydrostatic levellings (Brand, 2000).

Due to a number of reasons, it is very difficult to convert internal precisions of (first-order) levelling networks into estimates for the precision of determined heights of tide gauge bench marks. First of all, tide gauge bench marks will not often have been included in the first order levelling networks. Hopefully, they have been part of the second order networks. However, it is difficult to verify (especially for the older networks), whether the actual tide gauge bench mark has been included in the levelling network, or a bench mark in the area has been used. In addition, the accuracy of determined heights of tide gauge bench marks is not only determined by the internal precisions of the levelling networks, but influenced by a large number of systematic errors as well.

Systematic effects in levelling measurements The reliability of derived height differences, especially over long distances, is limited by the occurrence of systematic errors. In the past, a lot of research has been conducted on systematic errors in levelling measurements and procedures and methods that can be applied to minimise their effect. It will not be tried to give an in-depth description of all errors and possible solutions to reduce their influence, only an indication of possible error sources will be given.

In general it can be stated that systematic errors accumulate either with levelling distance or with elevation difference along the levelled path. Some of these errors, especially those due to, e.g., geophysical processes, will be time dependent. Although the influence of (some of these) systematic errors is nowadays largely reduced, by specific observing procedures or modelling, they will certainly limit the reliability of derived heights from the earlier levellings.

Comprehensive descriptions of a wide range of systematic errors inherent to the technique of spirit levelling are, e.g., given by Richards (1985) and Kukkamäki (1980). Examples of this type of errors are: vertical movements of the rod or levelling instrument, incorrect scale of the rod, and refraction. These systematic errors can introduce significant differences between subsequent first order levellings.

As an example, levelling refraction (in a temperate climate) can be of the order of 0.05 - 0.10 mm for every metre of elevation difference. The refraction effect in levelling is proportional to the elevation difference and the vertical temperature gradient, and proportional to the square of the sight length. Variations in mean sight length between different first order levellings can yield apparent relative elevation changes if no correction for refraction is applied. Nowadays effect of refraction can be reduced, e.g., by modelling based on measurements of temperature and temperature gradient. Unfortunately, in older levellings no refraction correction has been applied, and often no temperature information is available.

A second group of systematic errors is not so much the result of the measuring process itself but is caused by interfering processes in the area of the stations or processes influencing the reference surface. Some examples of (processes causing) systematic errors in levelled height differences are:

- Movements of the underground bench marks defining the national height datum.
  As described by, e.g., Rietveld (1986), since there is no stable rock on the surface of the Netherlands, underground bench marks defining NAP throughout the country have been founded in Pleistocene layers. Research (see, e.g., Lorenz *et al.* (1991)) indicates that these Pleistocene layers are not completely stable, and regional differences in movements of the underground bench marks have occurred.
- Tidal variations of the equipotential surface.
- Variations in equipotential surface due to isostatic recovery (e.g., post-glacial land uplift).
  With a maximum of 0.7 mm/yr in the northern part of the Gulf of Bothnia.
- Vertical movement of the earth's crust due to isostatic recovery.
  Depending on its location, a station can experience an uplift (in regions formerly covered by ice), or subsidence (in peripheral regions). As explained by Ashkenazi *et al.* (1990) this can result not only in a time dependent error but in a differential error as well for stations with different displacement rates. Combining measurements to which different reduction methods have been applied, leads to additional errors.
- Localised movements (in the area) of the bench mark
  As described by Cunietti *et al.* (1986), if the bench mark is superficially placed in sedimentary soils, it will be subject to lifting during rainy periods and settling during dry periods, resulting in a more or less periodic movement. Bench marks situated in sand will show only small vertical movements, while this amplitude may be significant for bench marks established in clay.
- Solid earth tide
  According to Ashkenazi *et al.* (1990), maximum error due to these astronomic effects is 0.1 mm/km.
- Ocean loading and deviation of the vertical due to attraction of ocean tides
  Since this effects measurements in coastal regions, this is important for the connection of tide gauge bench marks. Tilts can be of the order of 1 mm/km.

5.3.2 GPS

GPS is a fast, and relatively cheap alternative for connecting tide gauge bench marks in height. In addition, permanent occupation by a GPS receiver allows for continuous monitoring of height changes of the tide gauge bench mark. Continuous monitoring is only possible if the GPS antenna can be placed at the tide gauge bench mark itself. Otherwise, the rate of height determinations is determined by the frequency with which the levelling connection between tide gauge bench mark and GPS station is performed.

Levelling for the last part of the height connection between tide gauge bench mark and reference surface will be required if the location of the tide gauge bench mark does not allow the GPS antenna to be placed at this site. The tide gauge bench mark might be situated close to tall buildings, blocking part of the sky, or there might be too much obstacles causing multipath. In addition, as found, e.g., by Zerbini *et al.* (1996) for the SELF project, in harbour areas (where a major part of tide gauges is situated) GPS signals will suffer interference with signals stemming from communication transmitters and radar equipment.

Over the years, different procedures have been developed to monitor tide gauge bench mark heights by means of GPS. At first, it was proposed to use GPS only for height connections over distances of up to a few hundreds of kilometres. Observations were made in specific campaigns, in which selected stations were occupied for a few days, and campaigns were repeated every few months or years. This local network of GPS stations was connected into a global geocentric reference frame, by a height connection to the nearest VLBI or SLR station. See, for example, Carter *et al.* (1989) for more details on this subject.

Nowadays, it is often preferred to have permanently operating GPS receivers at (or near) specific tide gauges. Station coordinates can be determined in a so-called "fiducial station" solution, in which the network is solved while holding fixed the coordinates of at least three fiducial stations (e.g., VLBI, SLR, or GPS stations that are part of an ITRF solution). Following, e.g., Blewitt (1994), it is also possible to apply a "fiducial-free" method, i.e., no coordinates need to be fixed to previously (externally) determined values.

As an example, in the SELF project (see Zerbini *et al.* (1996)), sites in the area under consideration that are part of ITRF92 have been used as fiducial station. The selection of the reference system by means of using (a specific solution for the coordinates of) fiducial stations is of major importance for the accuracy of the derived height differences. As explained by Ashkenazi *et al.* (1993), selecting a different subset of fiducial stations will yield a slightly different reference frame. In addition, the fiducial stations themselves might be subject to vertical deformations. According to Zerbini *et al.* (1996), ITRF92 provides station positions to within 2 cm and velocities to within 3 mm/yr.

Precision of GPS based height differences Most commonly, as a measure to quantify the precision of GPS measurements, baseline length repeatability is used. However, height determination is of poorer quality than horizontal positioning, due, e.g., to the satellite configuration relative to the location of the receivers. An additional problem is the strong correlation between the vertical component and parameters (for clock, troposphere and bias) that need to be estimated; see, e.g., Springer *et al.* (1994) for more details. According to Bock (1996) standard deviations of height estimates are of the order of 3 times as large as standard deviations for horizontal positions.

Most estimates for the precision of GPS based height differences are, more or less, distance independent. On baselines of the order of 50 km, Ronde, de *et al.* (1993) found standard deviations for vertical positions of around 15 mm. According to Springer *et al.* (1994), repeatability of baseline heights (daily solutions) is of the order of 10 mm (8.9 mm + 1.1 ppb). From a comparison between GPS and SLR results, Springer *et al.* (1994) find an agreement of 12.5 mm for the derived height components.

Zerbini *et al.* (1996) state that absolute geocentric heights (ITRF92 coordinates) can be determined with a precision of around 20 mm, while using the approach developed for SELF, uncertainties in heights can be reduced to within only 10 mm. In general, it is assumed (see, e.g., Blewitt (1994)) that over distances of up to thousands of kilometres, based on continuous GPS occupation, global station heights can be monitored with a standard deviation of the order of 1-2 cm. On regional scales, a height precision of around 1 cm should be attainable. However, according to Ashkenazi *et al.* (1993), sub-centimetre precision for vertical movements of tide gauge bench marks will only be possible if a pure GPS based global reference frame (like IGS92) is used.

Systematic effects in GPS measurements GPS derived heights are influenced by a large number of, systematic, effects. Station velocities can further be influenced by deviations in satellite distribution and systematic errors between two different GPS campaigns. Systematic errors inherent to using GPS to determine height differences can be divided into a number of categories:

- receiver related errors
  According to Chrzanowski *et al.* (1990) the majority of these errors are independent of the baseline length. As a result, their relative effect will be larger for short baselines. Examples are: multipath, interference with "electronic" noise (e.g., signals transmitted by communication devices and radar equipment), and signal propagation delay inside the receiver; see, e.g., Seeber (1993).
  Another example is the variation in height between the mechanical centre of the antenna and its electrical phase centres for the L_1 and L_2 signals. Blewitt (1994) estimates a systematic error that can be larger than 1 cm, in case different antenna types are used and no calibration is applied. In addition, for very long baselines, the variations in phase centre no longer cancel out and, again, calibration is required.
- satellite related errors
  Mainly errors due to uncertainties in satellite location and in configuration of satellite relative to receiver. Examples are: orbital errors, signal propagation delay inside the satellite, and geometric distribution of satellites. For example, due to the specific location of the tide gauge bench mark, part of the sky might be blocked, resulting in a less favourable satellite configuration.
- signal propagation errors
  Refraction in ionosphere and troposphere can cause significant lengthening of the signal path. The major part of signal propagation errors can, for instance, be eliminated by combining measurements at L_1 and L_2 frequency (ionospheric effect), meteorological models possibly in combination with local measurements of pressure and temperature (dry troposphere), and additional measurements with water vapour radiometers (wet troposphere).
  Residuals of signal propagation delays after correction, e.g., due to atmospheric variations not described by the models used, cause errors in estimated station positions. Additional errors are

caused by inadequacies in the so-called mapping function, i.e., the relation between atmospheric delay estimated for actual measurements at a specific elevation angle, and zenith atmospheric delay predicted by the model.
As explained by Chrzanowski *et al.* (1990), estimated height difference between stations is mainly influenced by the relative refraction error (part of the atmospheric delay that is uncommon for both stations). This effect only depends on the elevation angle of incoming signals and not on the baseline length.

- errors related to reference frame
A reference frame is introduced by selecting a specific solution for the fiducial station coordinates (e.g., ITRF93), or by using a specific set of satellite ephemeris (usually given in WGS-84).
If very precise orbits are available, a fiducial-free solution of the GPS network can be determined. In this case coordinates of only one point will be kept fixed. The fiducial station approach is required if available orbits are not accurate enough. Possible inaccuracies in the fiducial station coordinates may cause distortions in the network. As explained by Ashkenazi *et al.* (1993), slightly different reference frames are defined if a different set of fiducial stations (for which coordinates are determined in a specific global reference frame solution) is selected.
An additional problem is movement of the fiducial stations due to plate motions. Usually, plate motion is reduced to a standard epoch using plate motion models.

In addition to the systematic errors described above, station heights can also be influenced by a number of physical processes (e.g., periodic vertical deformations), which affect the GPS measurements and interfere with the phenomenon we want to determine (long term variations and abrupt changes in station height). As explained by, e.g., Baker *et al.* (1995) (relatively short) periodic deformations like tidal loading deformations are important because they can alias into the longer period deformations that are of interest.

- solid earth tide
Causes vertical deformations of the earth's crust with an amplitude of up to 40 centimetres. The time average is not zero and depends on the latitude of the station (the so-called permanent tide). Usually, GPS measurements are corrected for the solid earth tides based on earth tide models determined from seismology. According to Baker *et al.* (1995), these type of models yield vertical displacements with uncertainties of about 2 millimetres (mainly due to lateral variations in the structure of the earth and inelasticity at tidal frequencies).
- ocean tide loading
The effect of ocean tide loading depends on the proximity of the station to ocean areas and the tidal range in these oceans. For specific locations, the range of vertical deformation can be of the order of 10 centimetres. The effects of ocean tide loading should (for the most part) cancel out in daily averages of station positions. However, as explained by Blewitt (1994), in reality solutions are formed for a deviating time period. Blewitt (1994) gives the example of 30-hour data windows, in which residual vertical displacements of the order of 1 centimetre remain.
Ocean tidal loading is often ignored, although for sub-centimetre precision tidal loading corrections are even required for stations located 500 kilometres inland.
- atmospheric pressure loading
At high latitudes this can cause vertical deformations with a range of a few centimetres. For stations near the coast, part of the effect of variations in atmospheric pressure are compensated by changes in sea-level height. However, the relation between variation in atmospheric pressure and sea level (in shallow water, near the coast) is not well established; see, e.g., Chelton and Enfield (1986). This leads to uncertainties in modelling the effects of atmospheric pressure variations for coastal stations.

5.3.3 Gravimetry

Gravity measurements can provide an independent method to distinguish between relative sea-level rise due to subsidence of the tide gauge and an actual increase in absolute sea level. As explained by Zerbini *et al.* (1996), the advantage of absolute gravity measurements is that they do not require a terrestrial reference system and are, therefore, not influenced by the selected realization of the reference frame. However, since gravimeters are sensitive to the combined effect of mass and height changes, and variations in mass are difficult to model, it is difficult to (reliably) convert measured gravity changes into variations in station height. This problem has been addressed in section 5.2, in this section only the accuracy of the gravity measurements themselves will be considered.

In principle two different types of gravity measurements are available, relative gravity measurements and absolute gravity measurements. There is a fundamental difference between these two techniques. Absolute gravity measurements require precise timing and distance measurement. For relative gravity measurements, only one of the two quantities, time or distance, has to be observed with great precision. Differences in this quantity between two gravity stations can be transformed into a relative gravity difference between these sites. For more details on this subject, see Torge (1989).

In the past, relative gravimeters provided measurements with a significantly higher accuracy than absolute instruments. Nowadays, the internal precision of absolute gravimeters is (at least) as good. Relative gravimeters can measure gravity changes over long periods of time and can even be used for continuous monitoring. Measurements of absolute gravity are usually obtained over relative short periods of time (one or two days).

Often, the best approach considered for monitoring height changes of bench marks is continuous measurements with a superconducting gravimeter in combination with repeated measurements with an absolute gravimeter; see for example Richter *et al.* (1993). This yields, on the one hand, a relation between the absolute gravity measurements in time, and on the other hand a restraint on possible drift of the relative gravity measurements.

Precision of gravity measurements Estimates for precision of absolute gravity measurements are of the order of 1 to 2 μGal. Based on results obtained for the SELF project, Zerbini *et al.* (1996) cite an average precision of 1.5 μGal for absolute gravity measurements, excluding outliers.

For a continuously operating superconducting gravimeter, measurements at the 0.1 μGal level should be feasible, although it should be kept in mind that permanent measurements are subject to instrument drift. Still, after an initial drift, Virtanen and Kääriäinen (1995) have found measurements with a superconducting gravimeter to stabilise at the (less than) 0.1 μGal/month level.

The above mentioned values are internal instrument precisions. Measurements are hampered by all kind of systematic influences as well. These systematic errors can (to some extent) be detected by comparing measurements with different (types of) instruments. According to Lambert (1994), systematic errors of the order of 6 to 10 μGal can be present in various types of absolute gravimeters. In addition, due to maintenance work, offsets (of the order of 10 μGal) can be introduced in the measurement series. However, for the new FG5 instrument it is claimed that no systematic errors beyond the 2 μGal level will occur. This complies with experiments performed by Lambert (1994) who found measurements made with several FG5 instruments and a JILA instrument to agree at the 2 μGal level.

Systematic effects in gravity measurements Of main interest are variations in gravity with time. Consequently, systematic errors in gravity measurements that yield a constant offset are of less interest than systematic errors due, e.g., to limited stability of the measuring system. As an example, measured variations in gravity due to instrumental drift are difficult to distinguish from gravity variations caused by actual vertical movements of the station.

Following Torge (1989) and Richter *et al.* (1993), examples of systematic effects in absolute gravity measurements (with an free-fall type of instrument) are: uncertainties in length or time standard, variations in the optical path length, and phase shifts in the electronic time measurements system.

Systematic effects present in relative gravity measurements are largely dependent on the type of gravimeter used, i.e., a spring gravimeter or a superconducting gravimeter. For spring gravimeters, errors are introduced by the construction of the gravimeter itself (e.g., uncertainties in zero position and reading scale, unstable voltage, etc.), external magnetic fields and variations in atmospheric pressure and temperature. More details are given by, e.g., Torge (1989) and Scherneck (1986).

In addition, gravity measurements are influenced by all kind of (local) geophysical, geochemical and oceanographic processes that cause gravity variations. This is further complicated by the fact that local gravity measurements not only depend on local phenomena but can be influenced by gravity changes in the far field as well. Examples of processes yielding gravity signals interfering with the phenomena we wish to observe, i.e., (secular) height changes of tide gauge bench marks, are:

- Gravity signals created by the observation site itself
  As explained by Harnisch (1993), often gravimeters are installed on observation pillars, which introduce a disturbing gravity field of their own. Consequently, moving the gravimeter from the centre to the edge of the pillar can yield gravity changes of a few μGal.
- Variations in ground-water level
  Variations in ground-water level and soil moisture can cause gravity variations of the order of a few tens of microgal. Unfortunately, (the effect of) soil moisture and ground-water variation is difficult to model, e.g., from measurements at a monitoring well close to the gravity station.
- Microseismic noise
  Gravity stations near the coast are hampered by microseismic noise caused by wave activity.
- Atmospheric pressure variations
  Movements of atmospheric mass influences gravity measurements due to the Newtonian attraction of the variation in air density on the one hand, and a loading effect of the earth's crust on the other hand. These two processes effect gravity in opposite directions.
  According to Harnisch (1993), the total effect of both attraction and loading can be described by a linear coefficient between -0.2 and -0.4 μgal/hPa. In extreme cases, this can yield a gravity variation of the order of 18 μGal (unfortunately, with uncertainties of the same order).
  Uncertainties in this simple model relating local atmospheric pressure measurements to gravity changes are:
  - For different sites and instruments, different regression coefficients may apply.
  - Local gravity is not only influenced by variations in local pressure, but by variations in regional distributions of air masses as well.
  - For stations situated near the coast, the influence of atmospheric pressure loading is reduced by variations in sea-level height, complicating the relation between local pressure and gravity.
- (Unmodelled) tidal effects
  Global and regional tidal models are used to eliminate tidal effects from the measurements. As indicated by Carter *et al.* (1989), for stations situated near the coastline, vertical attraction of nearby tidal ocean masses can yield significant gravity changes (of the order of several μGal). This is complicated by the fact that for complex coastal structures, tides are difficult to model.
- Polar motion
  As explained by, e.g., Ekman (1989), polar motion causes variations in gravity due to variations in the centrifugal force of the earth. According to Lambert (1994), this effect can be of the order of 13 μGal but can largely be removed based on pole position data as supplied by the International Earth Rotation Service.

5.4 Selection of observation sites

In this section, some remarks will be made concerning the selection of observation sites. Two aspects will be (briefly) discussed, i.e, spatial sampling of the measurements (maximum distances that are allowed between stations) and observation sites that are less suitable for specific techniques. Temporal sampling and time span of measurements will be discussed in section 5.5.

At first it should be remarked that for connecting tide gauges to a local (or global) reference frame, and for monitoring vertical movements of tide gauge bench marks, a spatial sampling and selection of specific sites can be used that yields optimal results. However, if velocities and movements of tide gauge bench marks over the last few decades (or even last century) have to be determined, sites should be used that have been occupied in previous campaigns as well.

For monitoring station velocities on time scales of a few years GPS measurements can be performed on the same sites as used in previous satellite geodetic campaigns. Often, these sites will still be accessible, but it should be checked whether or not environmental conditions have, significantly, changed. For example, due to construction activities part of the sky might be blocked, multipath may have increased considerably, or interference with other electronic sources has increased.

If station velocities, or velocity values for a specific area, on time scales of a few decades or more have to be considered, sites have to be used that have been part of national (first order) levellings. As indicated in section 5.3.1, this might prove difficult because lifetime of bench marks is limited and first order levellings as conducted over the last century do not have a lot of sites in common.

Spatial sampling of levelling observations Levelling can provide accurate height differences over short distances. However, this technique is sensitive to systematic errors, which can increase rapidly with increasing levelling distances. In the past, long distance levellings have been used to connect stations to the national height datum. Nowadays, serious doubts have arisen concerning the reliability of these connections and height changes based on (re)levellings over these long distances.

For first order levelling networks, often a maximum sight length of 50 metres is prescribed. Unfortunately, older levellings might be based on longer sight lengths, or even worse, sight lengths that are not approximately uniform throughout the network. If different mean sight length have been used for the various first order levellings, this can lead to apparent height changes if refraction has not been corrected for. Strange (1980) gives an example of an apparent uplift of 15.5 cm between a levelling with mean sight length of 61 metres in 1955 and a relevelling with mean sight length of 29 metre in 1965.

Spatial sampling of gps stations In order to determine height velocities, same sites have to be occupied in subsequent GPS campaigns. In addition, for the fiducial station approach, a number of stations for which the coordinates are known in a global reference frame, have to be included in the network. Some care has to be taken with the selection of fiducial stations, since, using a different set of fiducial stations (as available from a specific realization of a global reference frame) leads to a slightly different local reference system. In addition, comparing station heights determined in GPS campaigns based on different (realizations of) global reference systems might lead to apparent height changes as well.

As described in section 5.3.2, when the technique of GPS was first applied to connecting tide gauge bench marks, it was only used for connections up to at most a few hundreds of kilometres. Larger scale height connections were based on either SLR or VLBI measurements. Over the last years the precision of long distance GPS measurements (distances over 1000 km) has improved significantly. Consequently, Carter (1994) claims that GPS can now be used for tide gauge bench mark connection on a global scale, without a specific need of including SLR or VLBI measurements. However, questions can be raised concerning the long-term stability of a network based solely on GPS measurements.

Although GPS can be used for height connections over large distances, precision of measurements deteriorates slightly with distance. This is due to, for instance, signal propagation errors, which in particular hamper the resolution of ambiguities. As an example, Bosworth (1994) estimates a precision of a few millimetres for distances up to around 200 km, for distances up to 500 km he estimates the precision to decrease to 3 to 5 mm. According to Bosworth (1994), over distances between 500 and 1000 km height differences can probably still be determined with sub-centimetre accuracy, but for larger distances precisions will be of the order of (a few) centimetres. This is contradicted by Springer *et al.* (1994), who claim that, based on combined orbits of all IGS analysis centres, baselines up to 3000 kilometres can be determined without loss of accuracy.

Site selection in relation to the stability of bench marks In order to determine reliable height changes between measuring campaigns, stability of the bench marks used is of major importance. Movements obtained for the tide gauge bench marks should represent (secular) vertical movement in the region of the bench mark and not specific localised (abrupt or short-periodic) movements of the monument itself due to, e.g., shortcomings in its construction.

In general local movements of bench marks consist of two major components: seasonal height variations and a random walk component; see Langbein and Johnson (1997) for more details. They estimate that random walk noise can cause time-correlated errors of the order of 1 or 2 $\mathrm{mm}/\sqrt{\mathrm{yr}}$, and is probably the result of small disturbances of the bench marks, caused by weather effects interacting on the surface soils, which accumulate with time. According to Langbein and Johnson (1997), it is difficult to separate the (secular) vertical movements under consideration from the random-walk noise.

Following, e.g., Vaníček and Krakiwsky (1986), examples of mechanisms that can influence the stability of geodetic monuments are:

- Direct disturbance (or even destruction) of the monument
- Instabilities resulting from construction shortcomings
  As a result bench mark may be subject to thermal expansion and contraction or corrosion
- Variations in moisture content of the soil
  This results in volume changes of the underground of the bench marks that, in turn, yields vertical displacements. The amount of displacement depends on the type of soil (its compressibility)

underneath the monument. In general, bench marks established in peat or clay will show larger movements than those placed in a layer of sand or on rock.

- Expansion and contraction of the soil or rock due to temperature changes
  Vaníček and Krakiwsky (1986) give estimates of displacements between 0.1 and 0.25 mm per levelled section between bench marks that are keyed to bedrock, as a result of variations between day and night temperatures.
- Erosion of soil on which the monument is established by wind and/or water.

Instability effects of geodetic monuments can be minimised by both design and selection of favourable locations. Ideally, tide gauge bench marks should be placed on a stable rock surface. Unfortunately, in coastal zones bedrock is not often found on the surface. As an example, according to Rietveld (1986) there is no stable rock on the surface of the Netherlands and half of its surface is covered by clay and peat sediments with thicknesses up to 20 metres.

Site selection in relation to groundwater Variations in soil moisture influence the stability of geodetic monuments. In addition, ground-water variations have a direct (significant) influence on gravity measurements. This effect can be minimised by avoiding sites situated on soil with a high porosity. Preferably gravimeters should be placed on crystalline bedrock, although, as described by Lambert (1994), even on bedrock fluctuations in water table can occur.

Problems with sites near the coast Unfortunately, there is a wide range of problems associated with measurements at coastal sites. Reliability of determined height differences are influenced by noise from wave activity, attraction of tidal ocean masses, tidal loading, uncertainties in atmospheric pressure loading models, and differences in amount of water vapour over land and over water. In principle, all geodetic measuring techniques as considered in this chapter are influenced by (some of) these phenomena, but some more than others.

Wave activity creates microseismic noise, which can seriously decrease the accuracy of gravity measurements. The effect of wave activity will only be significant near the coast. Consequently, IOC (1994) advises (absolute) gravity stations to be placed at least 1 to 10 kilometres inland.

Tidal loading directly influences station heights with periods ranging from 8 hours up to years. In addition ocean loading tilts of around 1 mm/km can be created. When comparing measurements obtained at different measuring epochs, the effects of tidal loading have to be accounted for. Corrections are based on ocean tide models in combination with earth response models. Sites as far inland as 1000 km are still influenced by ocean loading effects. As explained by Vaníček and Krakiwsky (1986), for these large distances from the coast, the quality of existing earth response models and ocean tide models is sufficient to estimate vertical displacements resulting from ocean loading.

For stations within, e.g., 5 kilometres from the coast, major uncertainties arise in predicted vertical displacements. These uncertainties mainly stem from inaccuracies in the ocean tide models themselves. The nearer to the coast the observation station, the more important a, high resolution, tide model becomes. Unfortunately, local tidal regime is highly influenced by the coastal structure (shape of the coastline and ocean bottom) and, consequently, difficult to describe.

In addition to ocean tide loading, gravity measurements are also influenced by attraction from tidal ocean mass. This tidal attraction is only significant for gravity stations situated near (e.g., within a few km) the coast. Carter *et al.* (1989) estimate that vertical attraction of tidal ocean layers can cause gravity changes of several μGal for stations within 1 kilometre from the coast.

Analogous to ocean tide loading, atmospheric pressure loading has a direct influence on station heights. Based on local measurements of atmospheric pressure (part of) this effect can be modelled. For stations in coastal regions, the influence of variations in atmospheric pressure is reduced by variations in sea level. For a stations on a (small) island, an increase in atmospheric pressure will yield a redistribution of the water mass and, consequently, hardly any pressure loading.

For stations in coastal zones, the relation between atmospheric pressure and vertical displacement is difficult to estimate, mainly due to the following two reasons. First of all, only for open ocean areas there is a direct relation between variations in atmospheric pressure and variations in sea level, the so-called inverse barometer law. As has been explained by, e.g., Chelton and Enfield (1986), this inverse barometer law is not valid in coastal waters. A second problem is due to the fact that how much the influence of atmospheric pressure is reduced not only depends on the distance of the station from the coast, but also on the geometry of the coastline.

Space geodetic measurements have to be corrected for atmospheric delays caused by refraction in the ionosphere and troposphere. For GPS and VLBI the so-called wet tropospheric delay is the most difficult signal propagation delay to correct for, even though this yields relatively the smallest delays. Usually, the wet tropospheric delay in zenith direction in estimated based on a mapping function in combination with local measurements of temperature, pressure, and water vapour content. With a mapping function this zenith delay is converted into atmospheric delay for the elevation angle used.

Most mapping functions assume azimuthal homogeneity and, as a result, tropospheric delay is only elevation dependent. Azimuthal asymmetries in the refractive index (also called atmospheric gradients) can cause systematic errors in the determined height differences. For VLBI some results have been published concerning the influence of atmospheric gradients on the measurements. It was found that the influences of atmospheric gradients increases with decreasing elevation angle, especially for measurements obtained below 7° elevation; see Chen and Herring (1997) for more details.

Since GPS measurements are usually obtained with a 15° cutoff, MacMillan and Ma (1997) conclude that GPS measurements will be less sensitive to atmospheric gradients. Based on VLBI simulations they have found that for a 15° cutoff, deviations in vertical site estimates are half as large as those based on a 7° cutoff.

Most results have been published for VLBI stations situated inland. For GPS stations situated near the coast, part of the received signals has travelled over sea, and part over land. The water vapour content over sea might differ significantly from that over land. Consequently, if measurements to the various satellites are corrected based on the zenith tropospheric delay and applying the same mapping function for all measurements, systematic errors can be introduced in the derived height differences.

Sites with significant signal interference One additional source of noise that can reduce the quality of (space geodetic) measurements is interference of electronic signals. At some locations, large quantities of electronic noise can be present that interferes with GPS signals. As an example, Zerbini *et al.* (1996) mentions harbour areas were interference from, e.g., communication transmitters can be expected.

5.5 Required sampling and time span of measurements

In this section, some remarks will be made concerning the time span of measurements and the separation in time between subsequent measuring campaigns. The different geodetic techniques have different characteristics concerning temporal sampling.

Completion of a first order levelling network requires a large amount of time, measurements have started somewhere in the second half of the 19th century, and subsequent campaigns are separated in time by a significant number of years (decades). As an example, in the Netherlands levelling campaigns for the first order network have been performed between 1875 and 1885, between 1926 and 1939, in the period 1950 to 1959, between 1965 and 1980, and, finally the fifth primary levelling from 1996 to 1999.

GPS measurements are only available for the last two decades. At first measurements were obtained in specific measurement campaigns of relative short duration, nowadays continuous height monitoring is feasible. Absolute gravity measurements are usually obtained over short periods of time (up to around 2 days), while relative gravity measurements can, in principle, be used for continuous monitoring.

Since the gravitational field of the earth is not time independent, the geoid will slightly change with time. Consequently, in order to combine measurements obtained at different times, kinematic effects have to be taken into account. As an example, in order to form a united European levelling network, the levelling networks as available in the various countries should be reduced to one common reference epoch.

Space geodetic measurements are made in a global terrestrial reference system that is realized through a set of station coordinates obtained at a specific epoch, for example ITRF92. Due to plate motion, the fiducial stations move in time as well. Based on plate motion models, fiducial coordinates can be reduced to the observation epoch.

Great care should be taken with reducing obtained heights to a specific reference epoch. Measurements on which heights are based are obtained over a specific time interval. As explained by Vaníček *et al.*

(1987), only if the stations are not subject to significant height changes during this observation period, can a reference epoch be safely defined. Any vertical movements of the station that occur during the measurement campaign will influence the accuracy of the derived heights.

For GPS measurements vertical height variations of the stations during the measurement campaign should not cause major problems. Measurements can be obtained in relative short time intervals. Deformations that occur during the measurements are mainly of a periodic nature and can (for a large part) be corrected for.

Levelling is very time consuming and the completion of a first order levelling network will usually require several years. For large networks, the time required to perform the measurements might be of the same order as the time interval between subsequent levellings. This makes the selection of a reference epoch for heights determined in these levellings rather arbitrary.

Ideally, levelled height connections, for example between a GPS site situated inland and the tide gauge bench mark are performed within the smallest possible amounts of time. If this height connection is repeated regularly, it should be kept in mind that the occurrence of some types of systematic errors depends on the specific time of day or season. Examples are, deformation of the rod as a result of heating by the sun, shifts in readings due to variations in illumination of the rod by direct sunlight, and refraction. If systematic effects differ significantly between levellings, this can lead to apparent height changes of the tide gauge bench mark. Holdahl (1978) mentions the example of differences in refraction errors yielding apparent height changes for levellings performed in distinctly different seasons.