DANMARKS OG GRØNLANDS GEOLOGISKE UNDERSØGELSE RAPPORT 2011/140

Preliminary analyses of network for groundwater level monitoring

Anker L. Højberg & René K. Juhler

GEOLOGICAL SURVEY OF DENMARK AND GREENLAND
MINISTRY OF CLIMATE AND ENERGY

GEUS

DANMARKS OG GRØNLANDS GEOLOGISKE UNDERSØGELSE RAPPORT 2011/140

Preliminary analyses of network for groundwater level monitoring

Anker L. Højberg, & René K. Juhler

DE NATIONALE GEOLOGISKE UNDERSØGELSER FOR
DANMARK OG GRØNLAND, KLIMA- OG ENERGIMINISTERIET

G E U S

Table of Content

Foreword

The present work is part of the project “Udvikling af principper og metodikker til forbedring af DK-model” funded by the Danish National Monitoring and Assessment Programme for the Aquatic and Terrestrial Environment (NOVANA). The overall objective of the project has been minor model updates, model analyses and development of methods that can direct and support future model development.

The present report documents the study on the preliminary analyses of the network for groundwater monitoring.

Dansk Sammenfatning

Nærværende studie er en indledende analyse af det nationale overvågningsprogram for overvågning af den kvantitative status af den danske grundvandsressource. Formålet med projektet har været:

- At illustrere hvordan den nationale vandressource model (DK-modellen) kan anvendes til design af et fagligt funderet program for overvågning af den kvantitative grundvandsressource på national skala
- Identificere vidensbehov der skal afklares før der kan udformes et ensartet overvågningsnet på national skala.

Ifølge Vandrammedirektivet (VRD) (EC, 2000) skal der etableres et national overvågningsnet, der giver en sammenhængende og omfattende oversigt af den kvantitative status af grundvandsressourcen. Overvågningen finder sted ved monitering i grundvandsforekomster, hvor VRD giver mulighed for en gruppering af forekomsterne ud fra fysisk/kemiske forhold samt den aktuelle belastning af forekomsten. Gennem grupperingen kan antallet af stationer reduceres, så der kun overvåges i en eller enkelte forekomster indenfor en gruppe.

Det Nationale program for Overvågning af Vand og Natur (NOVANA) (Bijl et al., 2007) omfatter 121 indtag til overvågning af den kvantitative status. I Danmark er der udpeget 383 grundvandsforekomster og der sker således ikke en overvågning af samtlige forekomster. Det nuværende overvågningsnet er designet ud fra kriterier om en ensartet rumlig dækning samt anvendelse af eksisterende indtag med historiske data. Nettet tager derimod ikke højde for den gældende afgrænsning af grundvandsforekomsterne eller eventuelle grupperinger heraf.

I nærværende projekt er der skitseret en metode for anvendelse af DK-modellen til udformning af et fagligt baseret net for overvågning af den kvantitative status af grundvandsressourcen. Metoden er illustreret for Sjælland og tager udgangspunkt i modelsimuleringer til kvantificering af langtidspåvirkningerne af grundvandsstanden forårsaget af indvinding samt de fremtidige klimaændringer. En forudsætning for DK-modellens anvendelse er, at denne er i stand til at beskrive udviklingen i grundvandspotentialet forårsaget af ændret indvinding og/eller klima. Dette er indledningsvist testet ved anvendelse af historiske data for indvinding samt observerede tidsserier af grundvandsstanden for perioden 1990 – 2006, der indeholder såvel markante ændringer i indvindingen på Sjælland samt naturlige variationer af de klimatiske forhold med våde og tørre år. Metodens fokus er anvendelse af DK-modellen til forudsigelse af langtidseffekter, dvs. trends i udviklingen af grundvandspotentialet. Testen af modellen er derfor konstrueret på basis af en sammenligning mellem observerede og simulerede trends, hvor trenden er bestemt som hældningskoefficienten ved en linear regression af hhv. de observerede og simulerede tidsserier.

Til test af modellen er der anvendt pejletidsserier fra den nationale database Jupiter, hvor kun data markeret som repræsenterende et ”ro-vandspejl” er medtaget. Enkelte ”outlier” i en observeret tidsserie kan have markant betydning for estimering af hældningskoefficienten ved en lineær regression. For at undgå dette er der gennemført et kvalitetstjek af de observerede data, før der er lavet en sammenstilling med de simulerede tidsserier. Dette tjek er udført ved en statistisk test (Cook D).

Observerede og simulerede trends er sammenlignet både som et middel for hele Sjælland og for samtlige tidsserier medtaget i testen. Fittet mellem den observerede og simulerede trend for de enkelte tidsserier blev analyseret ved en cluster analyse, der guidede en efterfølgende gennemgang af de observerede tidsserier, hvor observerede tidsserier der var tydeligt påvirket af nærtstående indvindinger blev identificeret. Efter frasortering af ikke anvendelige tidsserier var der pejletidsserier fra i alt 1066 indtag til rådighed for den efterfølgende sammenligning med modellen.

Testen af modellen viste, at DK-modellen er i stand til at beskrive langtidsudviklingen af middel grundvandsstanden for Sjælland tilfredsstillende. Sammenligning af de enkelte observerede og simulerede tidsserier viser nogen spredning med størst afvigelse mellem data for tidsserier placeret i eller tæt ved indvindingsboringer. Modellens manglende evne til at simulere det indvindingsnære grundvandsniveau vurderes at være relateret til usikkerhed mht. hvilken situation disse observationer repræsenterer. Afhængig af de hydrogeologiske forhold samt den tid der medgår fra en pumpe slukkes til der foretages en pejling, vil pejledata fra indvindingsboringer afspejle et grundvandsniveau, der ligger mellem en fuld udvikling sænkningstragt og en upåvirket situation. Den øgede forskel mellem observeret og simuleret data for indvindingsboringer indikerer, at det ved kalibrering og validering af hydrologiske modeller kan være problematisk at tillægge pejletidsserier stammende fra indvindingsboringer for stor vægt,

Med modellen er det beregnet hvorledes den nuværende indvinding påvirker grundvandsstanden for Sjælland, ligesom grundvandspotentialet under et fremtidigt klima er beregnet ved anvendelse af IPCC klimascenarierne A1 og B1. Påvirkningerne af hhv. indvinding og et fremtidigt klima er opdelt i tre kategorier (lille, middel og stor påvirkning), der reflekterer den samlede effekt af indvinding/klimaændring samt de stedspecifikke hydrogeologiske forhold. Modelberegningerne er kombineret med den rumlige udbredelse af grundvandsforekomsterne samt de nationale overvågningsindtag for den kvantitative grundvandsressource på Sjælland. Denne kombination giver mulighed for en umiddelbar vurdering af hvorledes overvågningsfiltrene er fordelt mellem grundvandsforekomsterne, samt hvorledes filtrene er fordelt mellem områder med forventet lille, middel eller stor påvirkning.

Kombinationen af modelsimuleringer og data viser, at der er en del spredning mht. repræsentationen af overvågningsfiltre i grundvandsforekomsterne med en overvågning af 14 ud af i alt 63 regionale og dybde grundvandsforekomster. Af de 14 overvågede grundvandsforekomster er hovedparten repræsenteret med et enkelt filter, mens fem har to eller flere filtre. Sammenholdes filterplaceringen endvidere med den simulerede påvirkning fremgår det, at alle filtre ikke er placeret optimalt. Otte ud af de i alt 25 overvågningsfiltre på Sjælland, er placeret i områder hvor modellen forudsiger en lille påvirkning fra såvel indvindingen samt de anvendte klimascenarier. I fire af grundvandsforekomsterne er samtlige filtre placeret i områder med en forventet lille påvirkning. Mens det kan være relevant at have observationer i områder med en forventet lille påvirkning, bør dette ikke være de eneste områder, der overvåges indenfor en grundvandsforekomst.

For de øvre regionale grundvandsforekomster på Sjælland er der givet et eksempel på hvorledes modelsimuleringerne kan anvendes i forbindelse med gruppering af grundvandsforekomster samt identificering af optimale overvågningslokaliteter. Udnyttelse af grundvandsressourcen indenfor én grundvandsforekomst er generelt meget varierende i forskellige delområder af forekomsten. Én forekomst vil derfor ikke være karakteriseret ved én udnyttelsesgrad. I nærværende studie er følgende strategi anbefalet:

1. Gruppering af grundvandsforekomster baseret på den overordnede påvirkning af grundvandsforekomsten ved den eksisterende indvinding
2. Underopdeling af samtlige forekomster baseret på prædefinerede kategorier for påvirkningen (i dette studie er der anvendt tre kategorier: lille, middel og stor).
3. Indenfor hver gruppe af grundvandsforekomster betragtes hver kategori samlet på tværs af alle forekomsterne indenfor gruppen. De optimale overvågningslokaliteter indenfor én kategori, er de lokaliteter der giver mest information om variationen indenfor kategorien, dvs. de optimale lokaliteter er dem der optimerer variansen mellem observationerne. Denne analyse vil eksempelvis kunne gennemføres ved en PCA analyse til identificering af lokaliteter, hvor der forventes størst forskel (varians) mellem den hydrologiske respons, eller ved en cluster analyse til identificering af lokaliteter med sammenlignelig hydrologisk respons mht. påvirkning fra indvinding og klima.

Gennemførelsen af disse tre punkter kan baseres på modelsimuleringer med DK-modellen.

Udover en optimering af variansen mellem observationerne (pkt. 3 herover), skal følgende aspekter inddrages ved udpegningen af de optimale overvågningslokaliteter:

1. *Rumlig udbredelse af grundvandsforekomster*. Nogle grundvandsforekomster er opbygget af flere afgrænsede sandenheder, der er afgrænset fra hinanden. Dynamikken fra disse forekomster kan afspejle lokale forhold, hvorfra det ikke umiddelbart er muligt at ekstrapolerer viden til de øvrige forekomster i samme gruppe.
2. *Usikkerhed i modelberegningerne.* Hvor DK-modellen er i stand til simulering af udviklingen i potentialet kan modelberegningerne tillægges mere vægt, mens overvågningsbehovet er større hvor modelberegningerne viser sig mere usikre.
3. *Placering af eksisterende observationspunkter*. Til kalibrering af DK-modellen indgår alle observerede tidsserier, også observationer der ikke er del af det nationale overvågningsprogram. For en bedre rumlig dækning af observationer til kalibrering, bør områder med få eller ingen observationer bør prioriteres højest.
4. *Anvendelse af eksisterende indtag*. Ved inddragelse af eksisterende indtag skal der tages højde for følgende:
    a. Prioritering af indtag med lange tidsserier
    b. Teknisk udformning af boring og filter
    c. Administrative forhold, såsom tinglysning af boring

Udpegningen af de optimale overvågningslokaliteter skal således udformes som en multikriterieanalyse, der inddrager såvel de hydrogeologiske forhold, dvs. variationen i den hydrologiske respons fra indvinding/klima, samt tekniske og administrative forhold.

Gennem eksemplet for Sjælland er der identificeret et behov for opstilling af design kriterier for udformning af nettet så det kan tilgodese overvågningsbehovene. I nærværende studie blev den maksimale grundvandssænkning forårsaget af den nuværende indvinding samt en underopdeling i lav, middel og stor påvirkning anvendt som eksempel. Der eksisterer imidlertid alternative kriterier, såsom udnyttelsesgraden udtrykt som indvindingens størrelse i relation til grundvandsdannelsen, eller hvor stort volumen af grundvandsforekomsten der er påvirket. Endvidere er der behov for at klarlægge hvad der er kritisk og på basis heraf definerer meningsfulde kategorier for påvirkningen indenfor en enkelt grundvandsforekomst.

Udpegningen af grundvandsforekomster på Sjælland er baseret på en tidligere version af DK-modellen. Generelt er der god overensstemmelse mellem forekomsternes rumlige udbredelse og modellagene i DK-modellen, men der blev identificeret inkonsistens i nogle områder. På Fyn er grundvandsforekomsterne ligeledes udpeget på basis af en tidligere version af DK-modellen, men siden da er der sket en væsentlig opdatering af den geologiske model for Fyn. I Jylland har DK-modellen ikke været anvendt direkte til udpegning af grundvandsforekomster. Det vurderes derfor, at der for Fyn og Jylland vil være større inkonsistens mht. den rumlige udbredelse af grundvandsforekomsterne og modellagene i DK-modellen, hvorfor der vil være behov for en gennemgang af linket mellem grundvandsforekomster og DK-modellen.

På baggrund af erfaringerne fra nærværende studium baseret på data fra Sjælland, vurderes den nationale vandressource model at være velegnet til brug som støtteværktøj for evaluering samt re-design af den rumlige udformning af det nationale overvågningsprogram og gruppering af grundvandsforekomster. Forud for en detaljeret analyse af overvågningsnettet på nationalt niveau anbefales følgende trin imidlertid gennemført:

1. Definition af design kriterier, herunder specifikation af, hvordan disse kan anvendes til gruppering af grundvandsforekomster.
2. Alle indtag i NOVANA skal knyttes til en bestemt grundvandsforekomst og et model lag i DK-modellen.
3. Analyse af historiske tidsserier data til beregning af observerede tendenser, der kan sammenlignes med simulerede værdier for at teste DK-modellerne evne til simulering af tendenser mht. stigende/faldende grundvandsstand over en længere årrække. Det er væsentlige, at de observerede data ikke indeholder ”out-lier”, som vil resultere i en fejlagtig estimering af tendensen. Til dette formål er der opnået god erfaring i nærværende studium med anvendelse af den statistiske test Cooks D. Analyse af tidsserier giver endvidere en generel kvalitetssikring af data og vil dermed være til gavn for andre undersøgelser.

Ved gennemførelsen af den nationale analyse vil der endvidere være behov for videreudviklingen af en multikriterieanalyse til identificering af de optimale overvågningslokaliteter, herunder anvendelse af en cluster analyse eller PCA til karakterisering af den hydrologiske respons forårsaget af indvinding eller klimaændringer.

Introduction

The overall objective of the EU Water Framework Directive (WFD) (EU, 2000) is to achieve a good quantitative and qualitative status of all waters by 2015. To obtain this goal the EU member states are obliged to prevent further deterioration, protect and enhance the status of the water resources. The WFD thus prescribes both an improvement for waters already negatively affected by anthropogenic activities, as well as the protection with respect to future activities. Assessment of current status and detection of future trends in water bodies should be based on monitoring networks, designed to establish a coherent and comprehensive overview of the water status within each river basin district. Monitoring may be achieved by grouping of groundwater bodies, whereby monitoring can be reduced to monitoring in selected members of the group.

The Danish National Monitoring and Assessment Programme for the Aquatic and Terrestrial Environment (NOVANA) (Bijl et al., 2007) includes 121 screens for monitoring of the groundwater level. With a total of 383 groundwater bodies in Denmark not all groundwater bodies are monitored. The present network of monitoring screens is designed to provide a spatially even distribution with reuse of existing screens with historical data, and is not based on the configuration of groundwater bodies and possible grouping hereof.

A national water resource model has been developed for Denmark (www.vandmodel.dk). The model includes a continuous geological/hydrogeological model for the entire country, which has been updated in the period 2005 – 2009, based on detailed existing geological and hydrogeological models. The nationwide geological/hydrogeological interpretation provides a consistent basis for a three-dimensional delineation of the groundwater bodies, as well as link between groundwater bodies and location of monitoring screens.

It is planned to expand the Danish monitoring network on groundwater levels. For an optimal design of the expanded network there is a need for a comprehensive evaluation of the present network to assess its adequacy and identify optimal location for the future expansion of the network.

The present study is a preliminary study on how the national water resources model can be applied to evaluate the adequacy of the present monitoring network on groundwater levels and support future design of the monitoring network.

Project objectives

The project objectives are:

- To illustrate the possible use of the national model for the design of a monitoring network that is more technical based than the design of the present network.
- Identify requirements for a nationwide analysis of the monitoring programme.

While monitoring networks must be design to assess both the quantitative and qualitative status for all waters, the present study only considers monitoring of the quantitative status of the groundwater. In a monitoring design both the spatial and temporal resolution must be considered. However, all screens included in the national monitoring programme on the groundwater level is equipped by data logger and the temporal resolution is thus not considered.

The analysis is exemplified by using the DK-model and monitoring data for Sjælland

The National monitoring programme

The Danish National Monitoring and Assessment Programme for the Aquatic and Terrestrial Environment (NOVANA) includes 121 wells for monitoring of the development of the groundwater level. The spatial organisation of the monitoring network has primarily been designed to assure a uniform geographic distribution with monitoring in areas of different hydrological and land use characteristics, from which knowledge could be extrapolated to non-monitored areas. The national monitoring network was originally established in 1989, but has been adjusted several times to adapt new knowledge and changing need for monitoring (Jørgensen and Stockmarr, 2008). The spatial distribution of the present network for monitoring of the groundwater is provided in Figure 1.

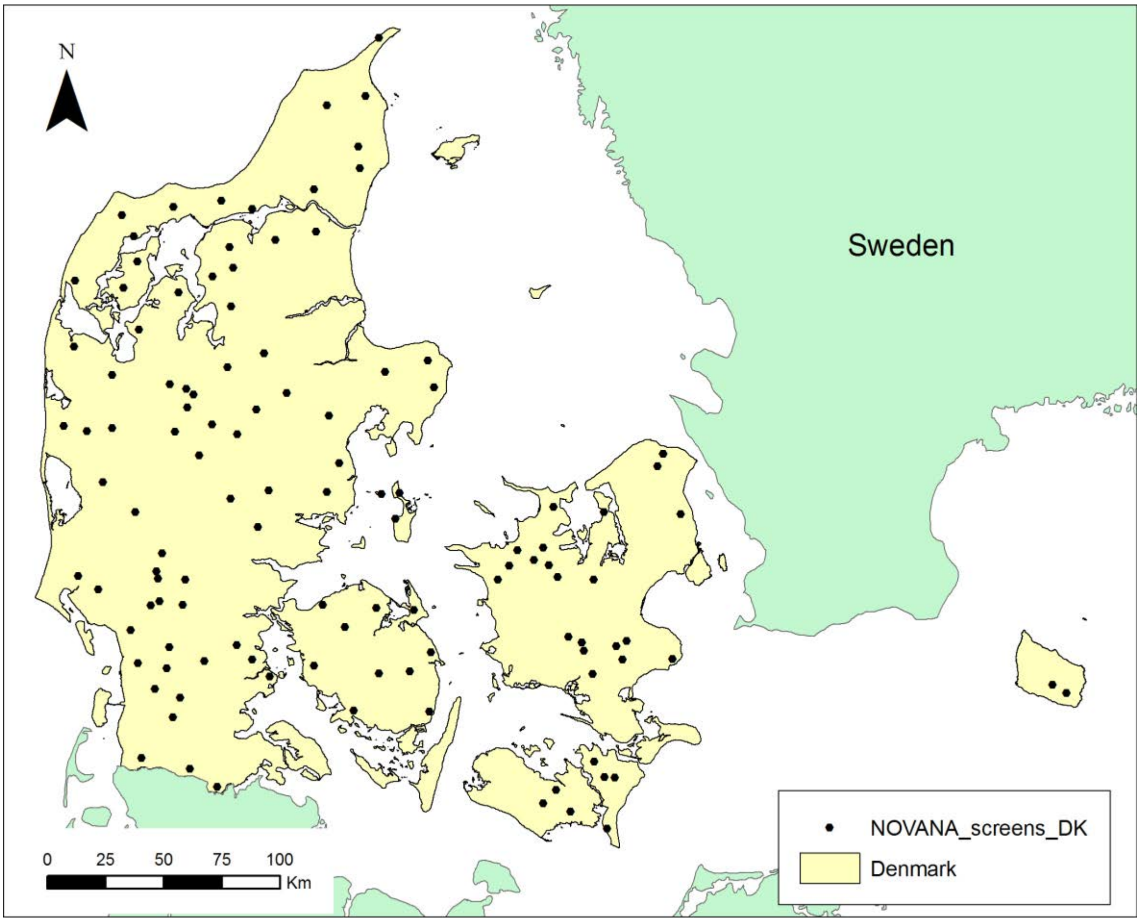

Figure 1 Location of screens in NOVANA for monitoring of the groundwater head

One of the most important objectives of the present monitoring network is to fulfil Denmark‟s obligations according to the WFD. Assessments on the quantitative status of the groundwater must be accomplished at the level of groundwater bodies. In Denmark groundwater bodies has been delineated centrally with focus on application of a standardised strategy and with the aim of achieving a homogeneous delineation with respect to size on numbers of groundwater bodies in the different regions of the country (Danish EPA, 2006).

The quantitative status of groundwater bodies should consider not only the direct impact extraction may have on the groundwater levels but also secondary effects on surface water systems, as described Annex V table 2.1.2 in the WFD, and provided in Table 1.

Elements	Good status
Groundwater level	The level of groundwater in the groundwater body is such that the available groundwater resource is not exceeded by the long-term annual average rate of abstraction. Accordingly, the level of groundwater is not subject to anthropogenic alterations such as would result in: — failure to achieve the environmental objectives specified under Article 4 for associated surface waters, — any significant diminution in the status of such waters, — any significant damage to terrestrial ecosystems which depend directly on the groundwater body, and alterations to flow direction resulting from level changes may occur temporarily, or continuously in a spatially limited area, but such reversals do not cause saltwater or other intrusion, and do not indicate a sustained and clearly identified anthropogenically induced trend in flow direction likely to result in such intrusions.

Table 1 Table 2.1.2 from Annex V in the WFD on good quantitative status of the groundwater

In Denmark, several groundwater aquifers may be arranged vertically at different levels with or without hydrological contact. Commonly the aquifers close to terrain are in hydrological contact to streams, while the deep aquifers do not interact with the surface water system. Since both the direct impact by extraction in the groundwater system as well as effects on the surface water system must be considered in the evaluation of the quantitative status, the overall concept in the delineation of groundwater bodies has been to separate the groundwater system according to their expected interaction with the surface water system. This has led to the definition of three categories of groundwater bodies (Danish EPA, 2006):

- Surface near groundwater bodies. Defined in areas where sand extents to terrain. The groundwater bodies are in hydrological contact with surface water system, are often small non-continuous aquifers that are delineated on the basis of topographical divides.
- Regional groundwater bodies. Are similarly in hydrological contact with the streams but forms generally larger continuous aquifers of regional extent.
- Deep groundwater bodies. Covers the deep laying aquifers without hydrological contact to the surface water system.

The initial delineation of the groundwater bodies was carried out for the entire country by a consultancy and later revised by the regional water authorities. The official version of the physical extent and status of the groundwater bodies are hosted centrally at the common

GIS-centre in Aalborg, from which a version was received for the present study in September 2010.

In the next planning period of the national monitoring programme running from 2011 to 2015, the design of the monitoring network will be evaluated. Expansion of the programme with inclusion of new screens is similarly planned with focus on establishing new surface near monitoring wells for gaining more knowledge on the surface near processes and stream-aquifer interaction.

Study area

The analyses described in the present report are exemplified by using the DK-model for Sjælland. This sub model was selected due to the large number of groundwater observation wells located on Sjælland.

Geological and hydrogeological model

The geological model for Sjælland is composed of a Quaternary sequence of alternating clay and sand deposits overlying Pre-quaternary clay, marl and chalk/limestone. The Quaternary deposits has been conceptualised as clay from terrain to the Pre-quaternary surface, with sand lenses in four vertical levels. Surface near geology (toplayer) was adopted from the national map on surface near geology, which was simplified into three classes: clay, sand and peat/others and assumed to represent the top 3 m geology. Pre-quaternary clay and marl only exist in the western part of Sjælland, while chalk and limestone forms the deepest aquifer used for drinking water abstraction in the entire model domain. The conceptual geological model is shown in Figure 2, while the thicknesses of the sand lenses at the four levels are shown in Figure 3. For the chalk/limestone aquifer a constant thickness of 50 m was assumed in the entire area.

In the numerical model the geological layers are used as computational layers, with the exception of the three topmost layers (toplayer, kl1 and ks1) which were combined to one layer. The numerical model thus consists of 10 vertical layers.

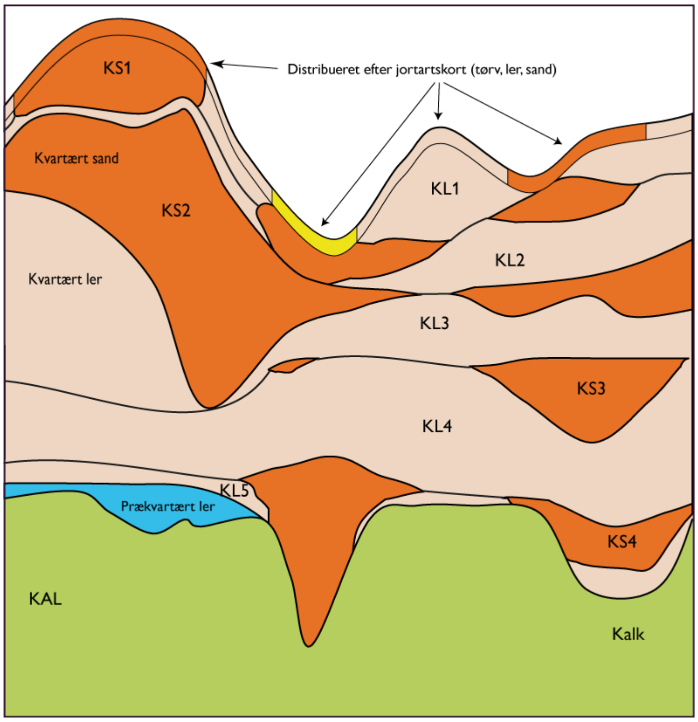

Figure 2. Conceptual geological model for the study area Sjælland. KS: quaternary sand (sand 1 to sand 4); KL: quaternary clay; KAL: Chalk/Limestone.

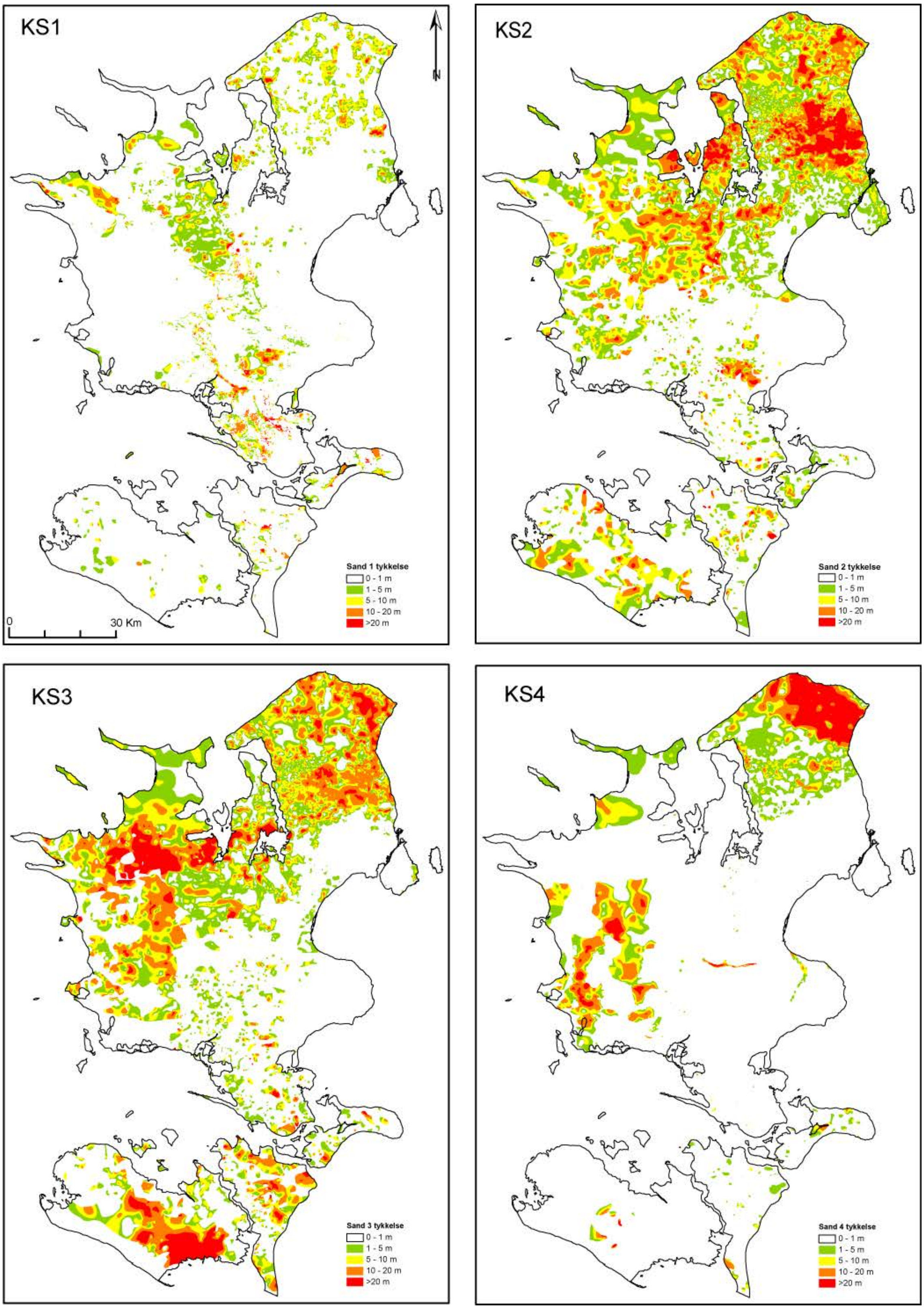

Figure 3. Thickness of the sand lenses in the four vertical levels, Sand 1 (KS1) to Sand 4 (KS4).

The topmost level of sand lenses forms discontinuous aquifers. Extraction is relative sparse from this layer and dominated by single wells managed by private well owners. The second and third level of sand lenses form regional aquifers and are used for groundwater abstraction in the western and northern part of Sjælland, Figure 4. The lower sand lenses (sand 4) represents deep quaternary sand deposits, predominantly associated with valleys cut into the pre-quaternary surface. Aquifers in the lower sand are thus discontinuous, and only exploited locally. The lower chalk aquifer constitutes the main aquifer in the eastern part of Sjælland, where the thickness of the quaternary sequence is limited. Most of the water supply for the greater Copenhagen area is based on the chalk aquifer. The location of the extraction wells in the chalk and the three lower sand aquifers are shown in Figure 4.

The grey shaded areas in the figure represent areas where the sand aquifers Sand 2 - Sand 4 do not exist. For the Chalk/Limestone aquifer the shaded area is the extent of the Pre-quaternary clay on top of the Chalk/Limestone aquifer, where it is found at greater depth and less exploited.

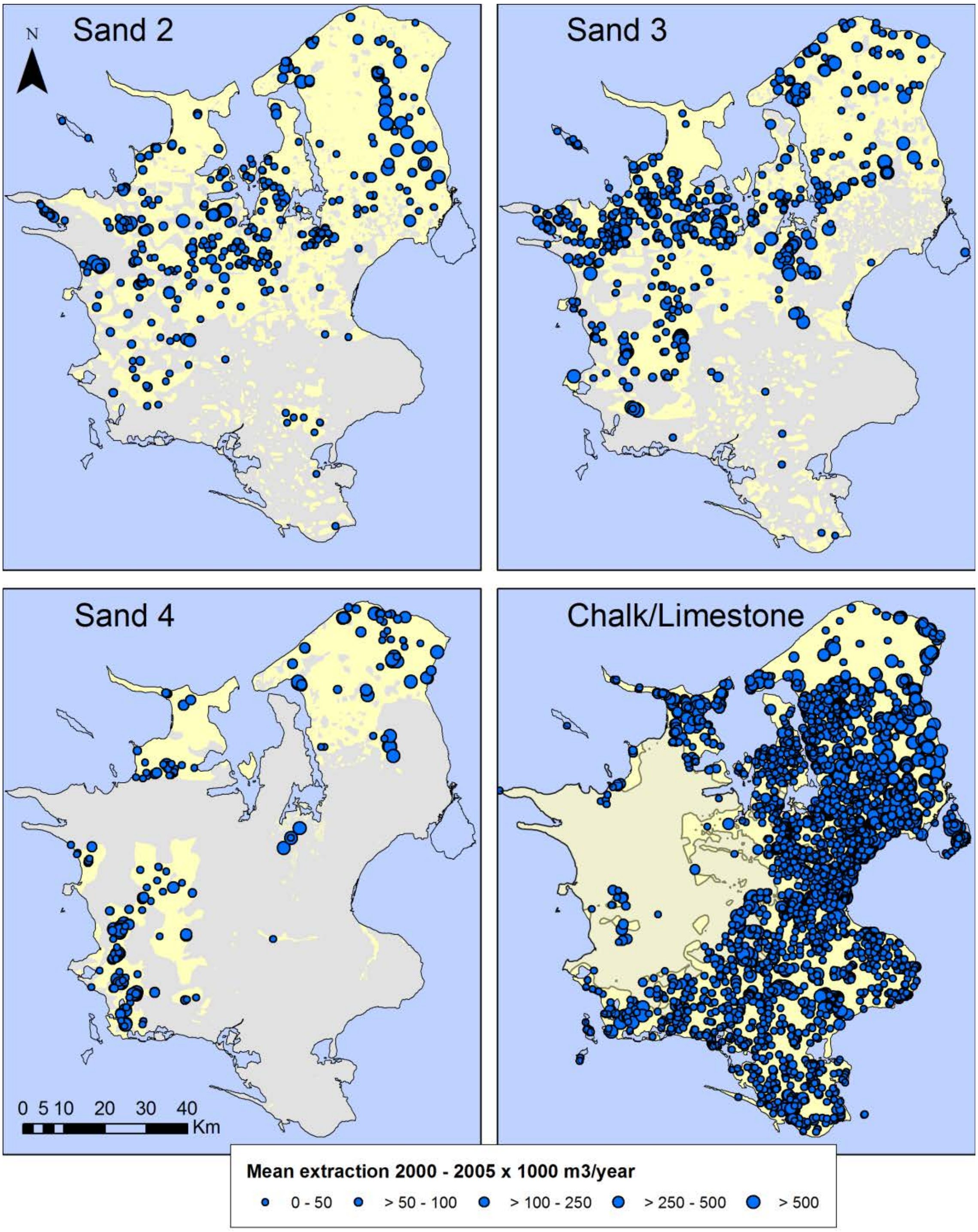

Figure 4. Extraction in the sand aquifers in level 2 - 3 and the chalk/limestone aquifer.

The temporal development of the water abstracted on Sjælland is shown in Figure 5. A decline in extracted levels is observed during the 1990‟ties, which is attributed major efforts in water savings. From 2005 another significant reduction is observed. This reduction is, however, coincidence with a structural reform in Denmark, with formation of new institutes responsible for water administration at the regional level and reorganisation of the obligations among the regional and local water authorities. The drop in extraction levels from 2005 is thus ascribed incomplete reporting of extraction to the national database JUPITER, and not a further reduction in extraction.

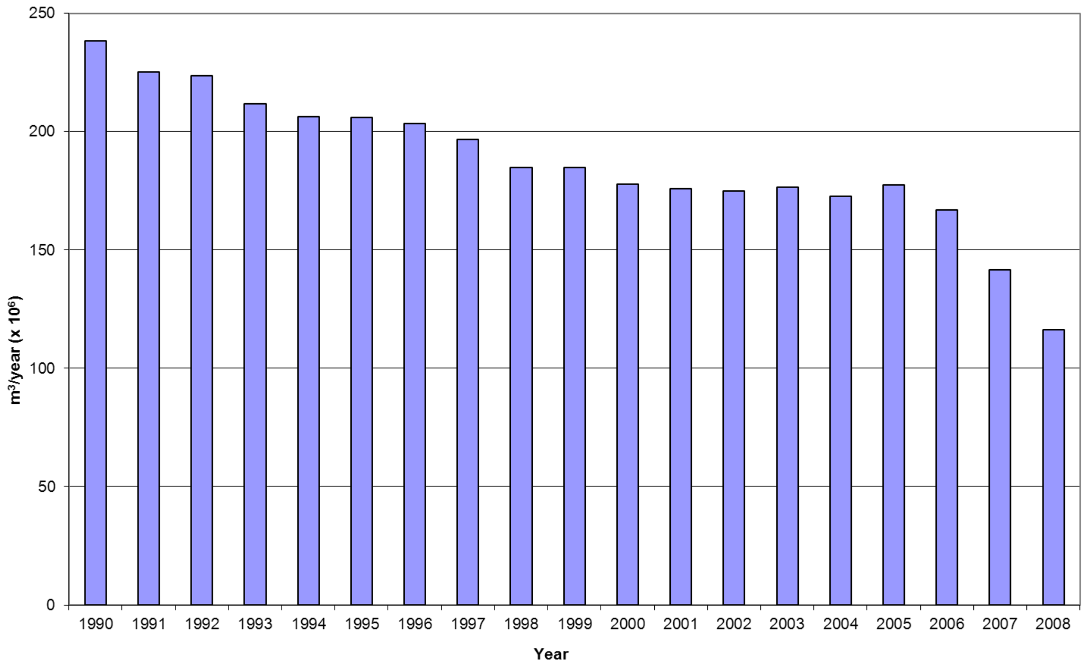

Figure 5. Temporal development of extraction on Sjælland.

Groundwater bodies and NOVANA screens on Sjælland

A total of 69 groundwater bodies have been delineated on Sjælland. Six groundwater bodies are defined as surface near groundwater bodies, while the majority (52) belongs to the category of regional groundwater bodies, and the remaining 11 are designated as deep groundwater bodies, Figure 6. While the surface near and deep groundwater bodies are associated to only one aquifer and one model layer in the vertical direction, the regional groundwater bodies often spans the two sand aquifers “Sand 2” and “Sand 3”.

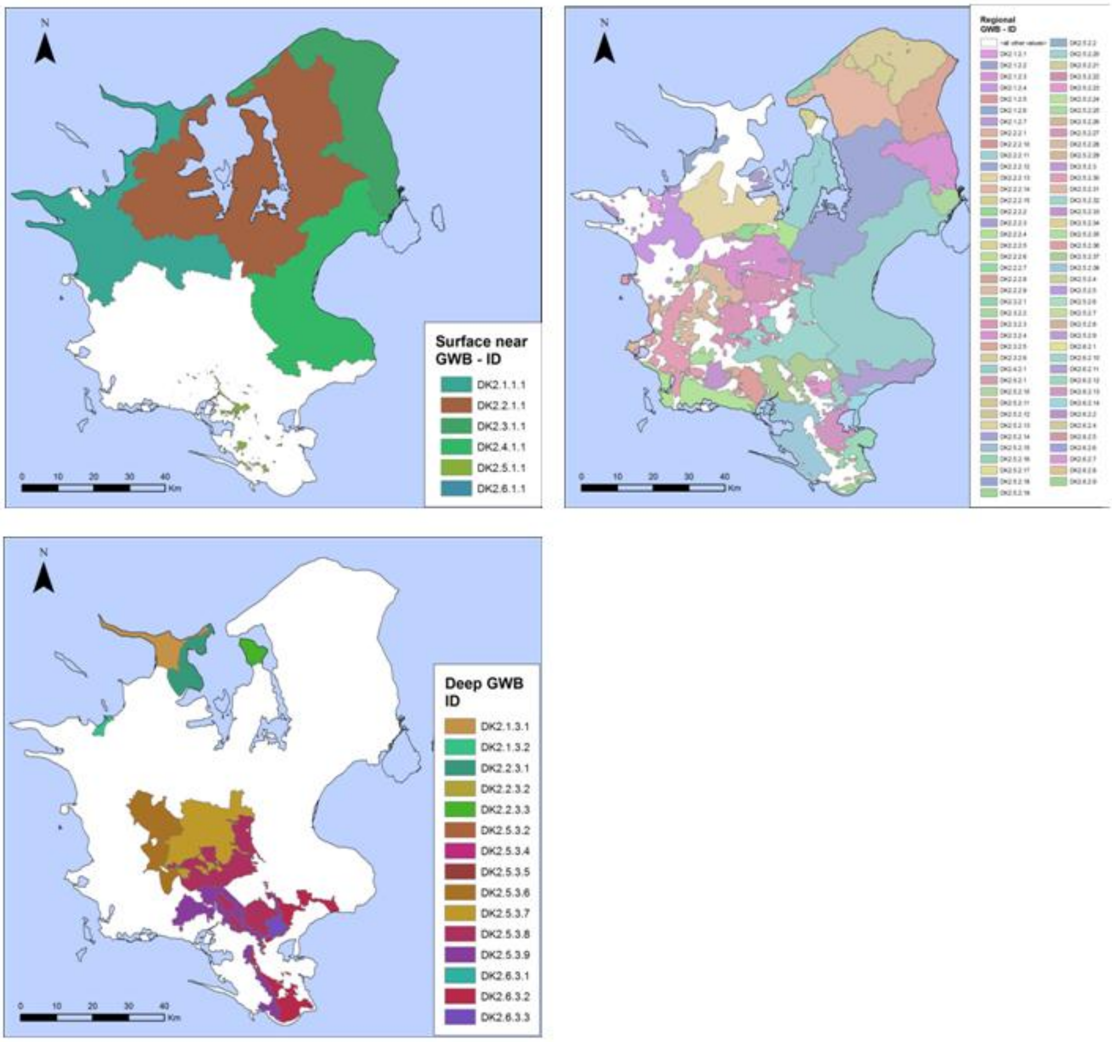

Figure 6 Groundwater bodies on Sjælland in the three levels

The groundwater bodies have been associated to the model layers in the DK-model by the regional water authorities and this link between groundwater bodies and model layers have been adapted in the present study.

The national monitoring programme (Bijl et al., 2007) includes 25 screens on Sjælland for monitoring of the groundwater level. The spatial locations of the screens are shown in Figure 7.

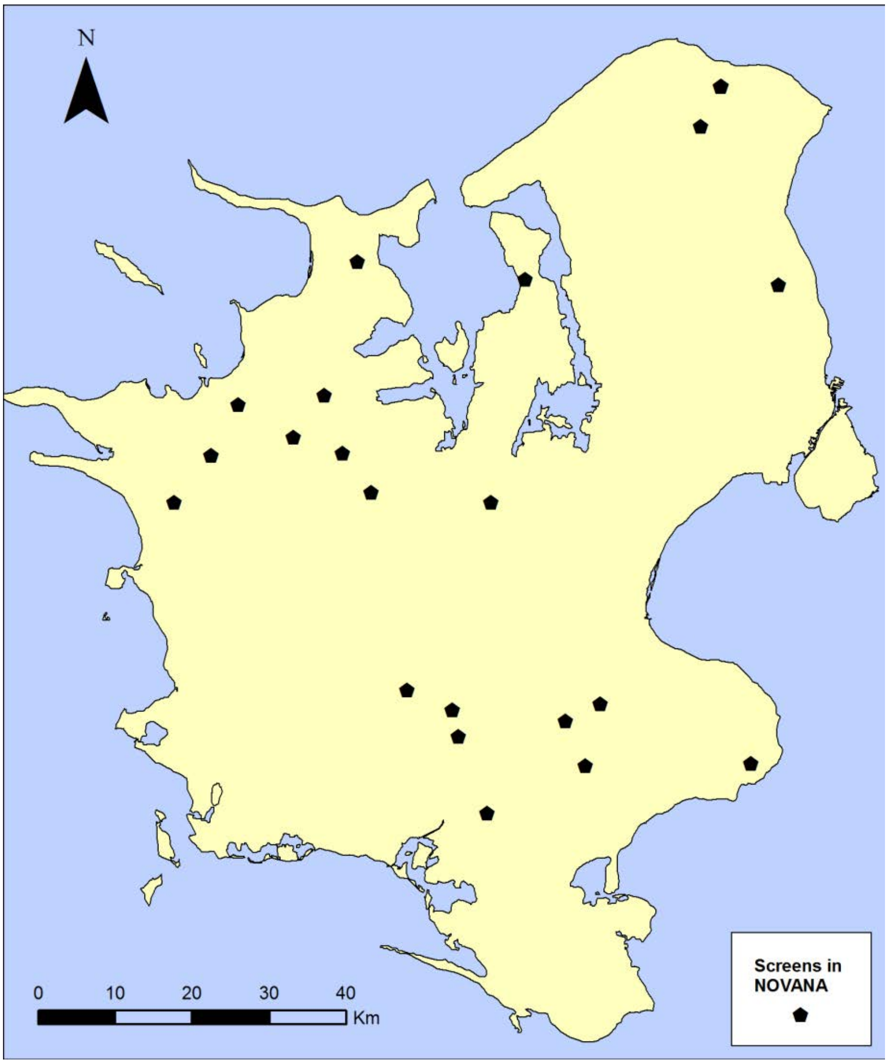

Figure 7 Screens in the national monitoring programme for monitoring the groundwater level on Sjælland

Each screen in the monitoring programme has been associated to a groundwater body by the regional water authorities and Table 2 lists the screens and associated groundwater bodies for Sjælland. Most of the screens included in NOVANA have been associated with a groundwater body, but five screens on Sjælland are missing a groundwater body. In three cases a groundwater body could be estimated based on the spatial correlation between the screens and the delineation of the groundwater bodies (marked by * in Table 2). For the remaining two screens, no groundwater body was defined.

The Groundwater bodies have been numbered by four digits, where the third digit defines whether the groundwater body is 1) surface near, 2) regional or 3) deep. For some screens in Table 2 there is a mismatch between the numbering and the classification ("Class"), which is similarly received from the regional water authorities. Furthermore, there is some inconsistency in the classification where screens belonging to the same groundwater body have been classified as Regional and Deep, respectively.

The national water resource model has been used as the basis for the initial three-dimensional delineation of the groundwater bodies. In the present study, the vertical arrangement of the screens in model layers has been updated using the most current version of the model, column “model layer” in Table 2. In the proceeding part, the screens will be displayed in accordance with the model layers in the table.

Table 2 Description of NOVANA screens on Sjælland used to monitor the groundwater level. * marks screens for which a groundwater body was estimated in the present study

ID	Model layer	GWB	Class
216.625_5	3	?	Regional
216.625_1	8	?	Regional
204.397_2	1	2.1.2.4	Regional
203.90_1	5	2.1.2.4	Deep
204.397_1	5	2.1.2.4	Deep
191.102_1	10	2.1.3.1	Deep
192.11B_1	10	2.2.2.11	Regional
197.166_1	5	2.2.2.13	Regional
197.334_1	5	2.2.2.13	Regional
197.476_2	5	2.2.2.13	Regional
205.336_1	5	2.2.2.13	Deep
205.342_1	3	2.2.2.2	Deep
206.1237_1	3	2.2.2.6	Regional
182.402_1	5	2.3.1.1	Terrain
182.319_1	10	2.3.2.1	Regional
194.129B_1	10	2.3.2.1	Regional
182.317_1	7	2.3.2.3	Regional
187.1057_1	5	2.3.2.6	Regional
217.163_1	10	2.4.2.1*	Regional
217.206_1	10	2.4.2.1*	Regional
216.272_1	10	2.5.2.32	Regional
216.529_1	10	2.5.2.32	Regional
217.474_1	10	2.5.2.32	Regional
221.278_1	10	2.5.2.37*	Regional
218.343_1	10	2.6.2.11	Regional

Methods

The overall methodology applied in the present study is use of the national water resources model to identify the long term impact by the current extraction and expected future changes and the spatial variations in such impacts. The use of the model for such analyses requires that the model is able to reproduce the long terms trends caused by extractions and climatic changes with sufficient accuracy. The first step in the analyses has thus been a test of the models ability to simulate the trend in historical data. As the amount of extraction has changed significantly from the early nineties to present, Figure 5, the use of historical data will provide a useful test on the models ability to simulate impact by extraction. The models ability to predict the effected of climatic change cannot be tested directly, but the simulation period (1990 – 2006) covers natural climatic variations.

Historical data used for model the evaluation consists of time series on groundwater heads from the national database JUPITER. Prior to use all time series were subject to a quality check focussing at identifying outliers that may affect trends in the data series. Time series identified as reliable were used to test the model, and finally the different model simulations for identifying effects of extraction and climatic changes were carried out. The different steps in the analyses can be summaries as:

1. Statistical test on observed data to identify outliers that may affect the overall trend in the time series.
2. Computation of observed and simulated trends by linear regression for all screens included in the analysis.
3. Evaluation of the models ability to reproduce observed trends.
4. Simulation of impacts by the present extraction under the current climate
5. Simulation of effect of expected future climatic change

The methods applied in the analysis are outlined in the following sections.

Analyses and selection of observed time series

The time series included in the analysis where extracted from the national database JUPITER. The database contains data from all screens, i.e. screens only used for monitoring as well as screen used for groundwater extraction. Observations in extraction wells measured during extraction may not provide a realistic level of the groundwater head in the aquifer, as the screen introduces a pressure loss due to resistance, which will result in head level lower than the surrounding aquifer. Custom for observations in extraction wells are that the pumps must be switch off for several hours prior to observation. As the time without pumping prior to observation varies and the extent to which a groundwater head will recover within a given time varies depending on the amount of water extracted and the hydrogeological characteristics of the aquifer, the observations in an extraction well will display a groundwater level somewhere between a fully recovered aquifer and a fully developed cone of depression. Data from extraction wells may thus describe conditions different from those described by the numerical models, and be considered more uncertain when used for model calibration and validation. The JUPITER database contains, however, a large number of observations from extraction wells, and it is desirable to analysis to which extend these data can be used for modelling purposes. Data from extraction wells were therefore included in the present analysis to explore the extent to which data from extraction wells deviates from other data. JUPITER provides the possibility of marking whether observations represent a situation with or without pumping. In the initial data analysis all observations were included. This resulted in the inclusion of many time series with large variations in the observation reflecting both conditions with and without pumping. It was therefore decided only to include data from the extraction wells if they were marked as representing a situation with no pumping. By doing so, too many data was excluded, as the marking of data in JUPITER is not consequent. For some time series, including time series from screens in NOVANA, the marking was only made for the more recent data, thus only parts of the time series were included in the analysis. For small scale detailed modelling it is possible to manually inspect all time series and evaluate whether they are valid for model use. This is not an option for nationwide assessment, e.g. application in the national water resources model or the assessments in the present study. For the inclusion of more data in future nationwide assessments, the evaluation at local scale must be documented centrally.

Observed time series are in the present study used to evaluate the DK-models ability to simulate trends in groundwater heads. To accomplish this, each observed time series must have a simulated counterpart, representing the groundwater head at the same location, and only observed time series from screens with x, y and z-coordinates were included in the analysis. Further requirements set for the inclusion of time series are that they contain recent data and have sufficient data to evaluate trends and yearly variations. The following criteria were formulated for the selection of time series: 1) only data between 1. January 1990 and 31 December 2008 are included 2) the time series must include data after 1. January 2000, 3) the time series must include 20 or more observations in total, and 4) in average there must be a minimum of 2 observations per year in active years.

Based on the criteria above for including a time series, a total of 1168 time series were selected. As the observations will be used to evaluate the DK-models ability to estimate trends, it must be assured that trends estimated from observation data is not erroneous due to outliers in a time series. A number of statistically based methods for identification of possible outliers have been developed, e.g. a commonly used „rule of thumb" that identifies possible outliers as points whose standardized residual is greater than 3.3 (corresponding to the .001 alpha level). In general, there are two types of widely used methods for assessing outliers: 1) statistics that assess the overall impact of an observation on the regression result (e.g. Cooks distance D) and 2) statistics that assess the specific impact of an observation on the regression coefficients (e.g. DFBETA). In the present study simple outlier statistics such as the alpha level 0.01 and Studentized residuals was tested and found less suited for the identification of outliers in the monitoring data series. Aiming for a more robust statistic for identification of a possible outlier the Cooks distance was evaluated. In common terms Cook's distance measures the effect of deleting a given observation in a dataset used for regression. It is important to note that data points identified as possible outliers may not be real outliers, i.e. outliers should only be removed if there is reason to believe that these observations are caused by errors or events that is not belonging to the system under study (e.g. malfunctions in instrumentation or high influence from water works and pumping). Non critical elimination of possible outliers may lead to models that are not a proper description of the system under study, i.e. possible outliers may hold valuable information. For this reason the time series containing suspected outliers should be inspected manually.

SAS 9.2 was used for the data analysis. Cook's D measures the change in the parameter estimates caused by deleting each observation and can be expressed as

$$D_i = [1/(p\, s^2)]\, (\mathbf{b} - \mathbf{b}_{(i)})'\, (\mathbf{X}'\mathbf{X})\, (\mathbf{b} - \mathbf{b}_{(i)})$$

for linear models, where $\mathbf{b}_{(i)}$ is the vector of parameter estimates obtained after deleting the *i*th observation.

Description of procedure: (*Steps in the data analysis)*

1. Import of data
2. Computation of descriptive statistics for each screen, such as start and end of monitoring period, number of observations and range of depths
3. Calculation of mean groundwater level for each screen on month, quarter and yearly basis
4. Simple linear regression and Cook D statistics using prog reg
5. Manual evaluation of possible outliers
6. Simple linear regression using prog reg on observed and simulated data

Based on the CookD statistics 158 time series were identified, where the slope in a linear regression analysis is affected by a single value. For each of these time series a graph was constructed highlighting the value identified in the CookD analysis, as exemplified in Figure 8 for a CookD analysis based on monthly averaged data. The time series where manually checked by visual inspection. Since the CookD test is carried out on averaged data (month, quarterly or yearly mean) values identified in the analysis may consists of one or more observations. If the value consisted of a single observation, the observation was assumed to be erroneous and removed from the time series. Where the values consisted of several observations points the entire time series were removed from the dataset. This was in general observed for low groundwater head levels in screens used for abstraction, indicating that the measurements was affected by the abstraction, as illustrated in Figure 9.

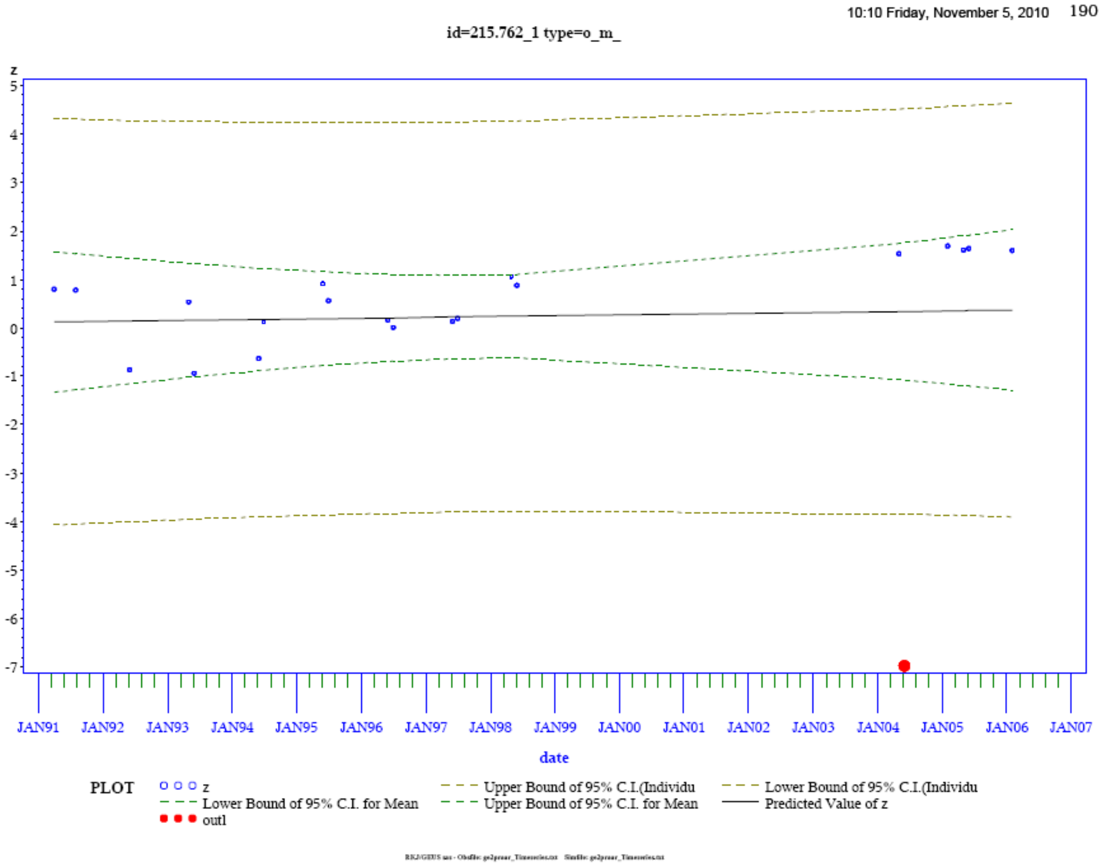

Figure 8 A data series from a monitoring screen (DGU nr. 215.762). Regression line and monthly averages are shown (▪) Upper and lower bonds (95%) are shown as dotted curves (individual and means). A single point at May 2004 is identified as a possible outlier

Figure 9 Observed time series where more than one observation was identified as an outlier by the CookD test

As described in the following section, the comparison of observed and simulated trend in the groundwater level resulted in the identification of additional time series that was discarded, primarily due to significant impact from extractions. After deselecting, a total of

1064 observed time series were included in the analysis. A few screens contained a very high number of observations due to the presence of automatic logger equipment. However, in general, the time series had less than 300 points and the distribution of data densities in the screens is shown in Figure 10. The temporal variation in number of active screens and observation data points in shown in Figure 11 from which it is seen that the number of screens decreases markedly from 2005. Due to establishment of automatic data logger the number of observations increases from 2003, and decreases again from 2005 with the reduction of the number of active screens, and probably due to a lower collection frequency in the automatic data loggers.

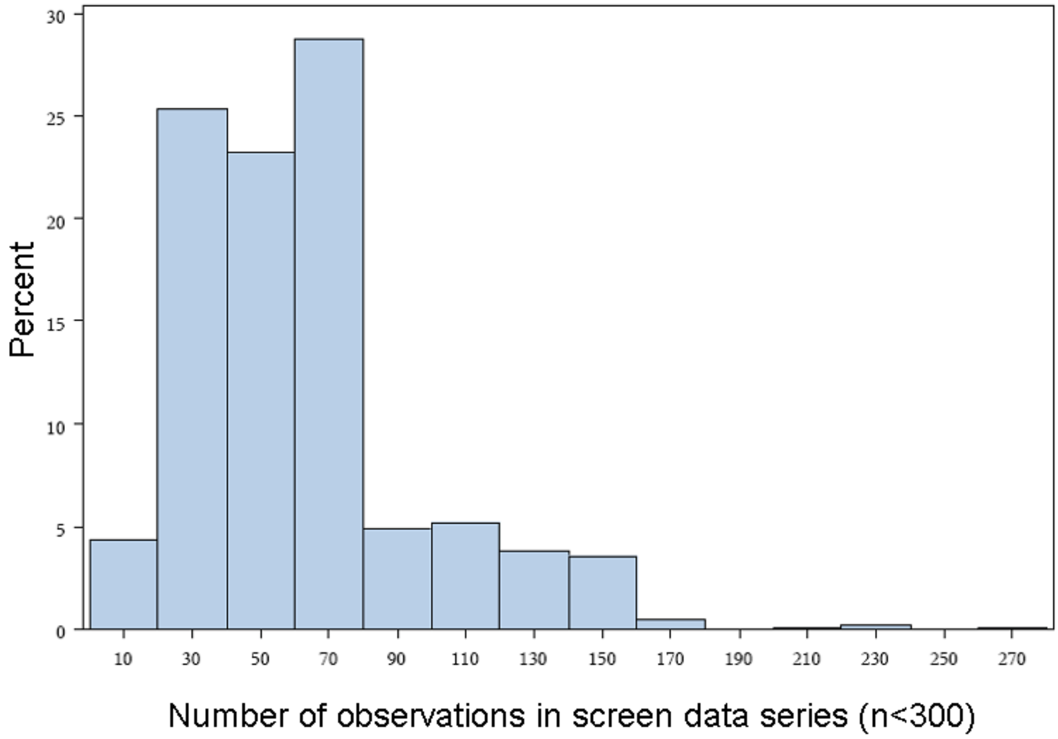

Figure 10 Data density in screens with less than 300 data points

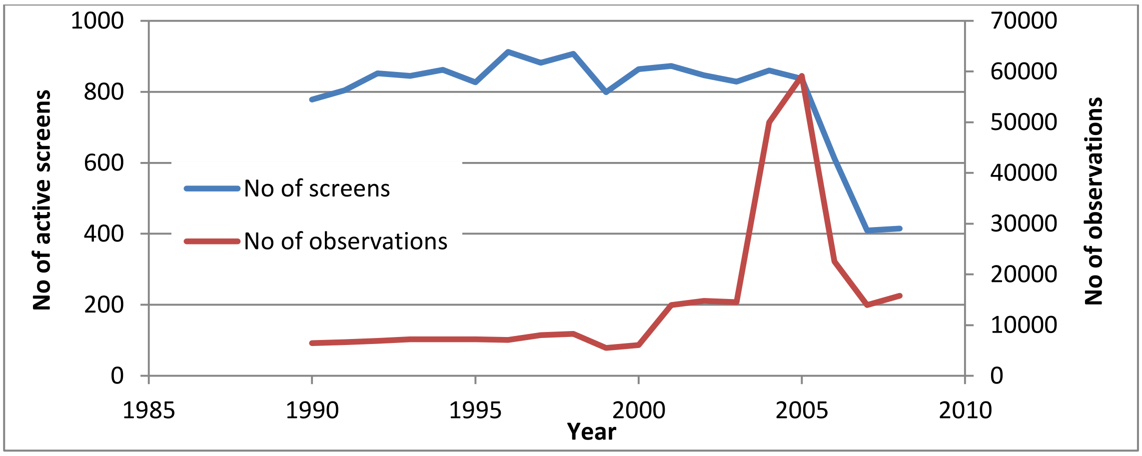

Figure 11 temporal variation in active screens and observation points

The spatial location of the selected screens is shown in Figure 12 which also includes the screens in the national monitoring programme. From the figure it is noticed that four wells (five screens) are included in the monitoring programme, but not selected in the present analysis. With the removal of data not marked as representing a situation without pumping four of these screens do only contain data for one to two years, while the last screen has been deselected due to suspicious data, with numerous data points displaying the exact same value.

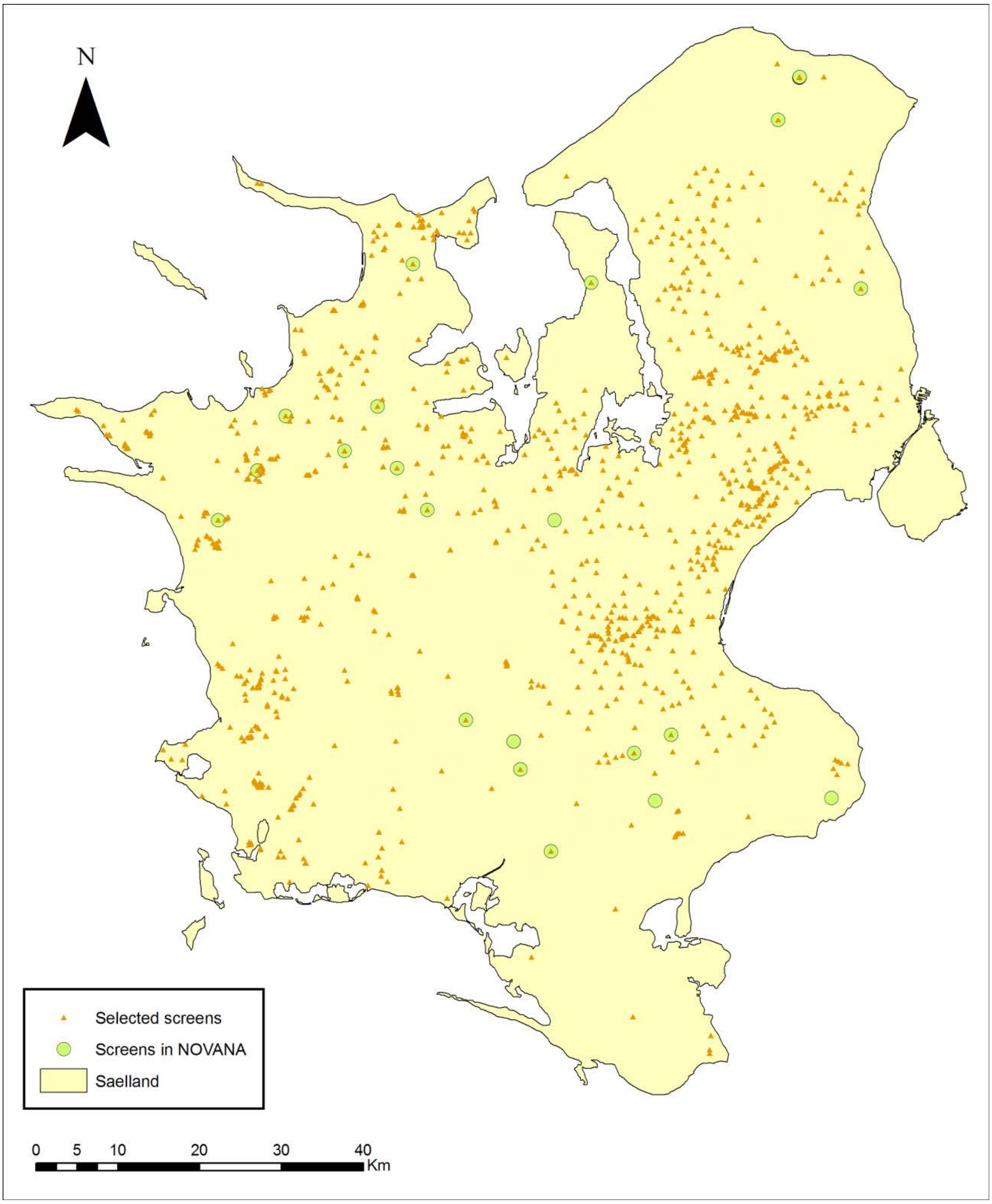

Figure 12. Location of screens with selected time series, and screen included in the national monitoring programme (NOVANA).

Test on simulated long terms trends

The DK-models ability to simulate long terms trends has been tested by comparing both the overall observed and simulated development in the groundwater levels and by comparing the observed and simulate development for each individual screen. The applied methods are described in the present section while the results are presented in the result section.

The overall development of the groundwater level on Sjælland have been analysed in previous studies. HUR (2005) employed time series on groundwater levels from 1100 screens to analyse the development from 1989 to 2003, and found a statistical significant trend with an increase in the groundwater level in the studied period, which was attribute a general decrease in groundwater abstraction. In a following study (Christensen and Sonnenborg, 2006), a previous version of the DK-model was used to evaluate, whether the analysis based on observation data was representative for the entire Sjælland. In this analysis, the development in the groundwater level was computed in two ways: 1) using only time series at location corresponding to the observations, and 2) using simulated head development for the entire model. It was found that the trend in the groundwater level was overestimated if only the locations with observations were included. This indicates that observations close to abstractions wells, with the highest response to changes in abstractions, were overrepresented in the analysis by HUR.

In the present study, the method applied by HUR is slightly modified for evaluation of the overall trend, as is briefly summarised here. The analysis is based on the period 1993 – 2005. The starting year 1993 is selected to allow for a “warm up” period for the model, while the end year 2005 is selected as the number of active screens is halved from 2005 to 2008, Figure 11. A mean groundwater head is computed for the entire period, as well as a mean yearly head for each year in the period. The groundwater heads are “normalised” by subtracting the period mean from the year mean. By this, a new time series for each screen is constructed displaying the yearly deviation from the mean for the entire period. Finally, one aggregated time series is constructed by averaging across all screens. In contradiction to the HUR analysis, the yearly mean values have been weighted according to the number of active screens in the actual year to account for the difference in number of active from year to year.

The second test on the models ability to simulate trends in the groundwater head development is carried out by comparing trends in observed and simulated head for the individual screens. The analysis is carried out by simple linear regression on observed and simulated time series, and comparing the slopes of the two regressions. The test period is similarly 1993 – 2005. Not all observed time series extends the entire period, and the observation frequency varies greatly among the time series as well as within a single time series. In order to compare the regression based on the two data sets, data is paired, i.e. simulated values are only included at dates for which observation exists. The result from the analysis is shown in Figure 13.

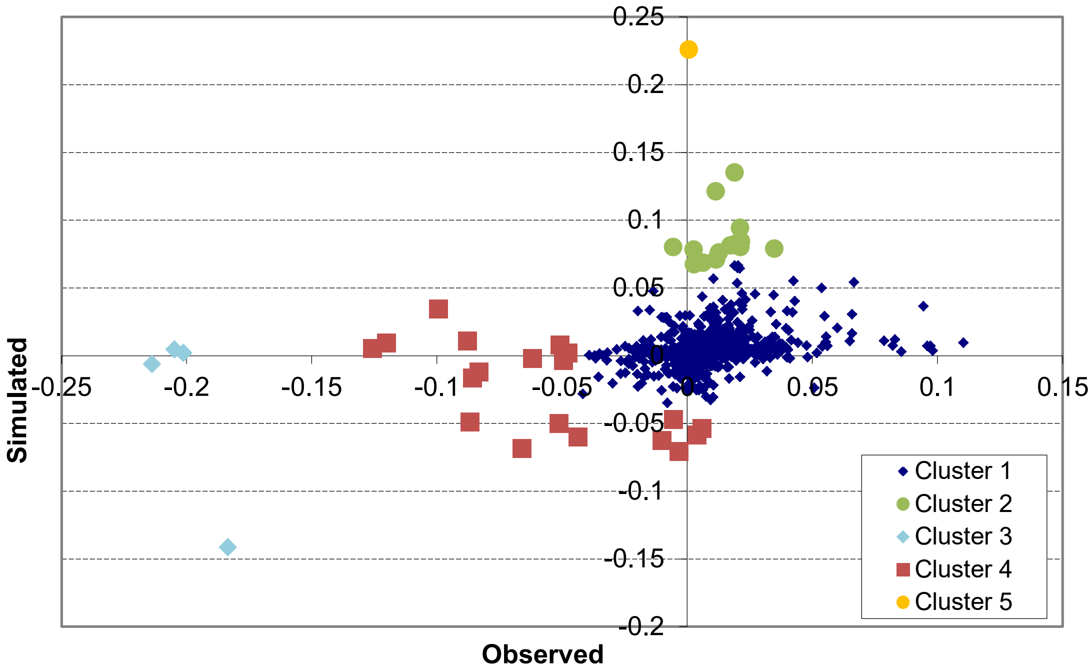

Figure 13. Simulated vs. observed slopes from linear regression of time series for individual screens.

The dataset was analysed by a clustering analysis on the estimated trends (i.e. the slope of the regression line for each screen). In this analysis variables used for clustering was the slope estimated using observed and simulated data. In the ideal world the trend analysis would result in identical results for each screen irrespectively from the choice of data series used (observed or simulated data series). However, in the real world situation and use of actual monitoring data there will be differences between observed and simulated data series. This will cause screen slope estimates that differs when trend analysis from observed time series are compared to simulated series. For example, in some screens a steeper slope is estimated using observed data than when using simulated observation series. As a first step towards identifying screens where the models may be optimised or monitoring improved a cluster analysis was made comparing slopes based on simulated and observed data for each screen. Ordinary significance tests such as analysis of variance *F* test are not valid for testing differences between clusters, and testing of clusters against the null hypothesis that objects are assigned randomly to clusters are also not suitable for determining optimal number of clusters. Thus, the optimal number of clusters was determined by inspecting the development of several statistical characteristics in relation to cluster number (flexible beta distance, R^2, RMS standard derivate, and semi partial R^2). Based on this analysis a set of five clusters was identified as suitable for the actual cluster analysis. The outcome of the five cluster analysis was manually inspected. Cluster 2 and 5 are all extraction wells and both observed and simulated heads are expected to be associated with large uncertainties and removed from the dataset. Cluster 3 consists of four wells, of which three have a significant negative slope in the observed data but no simulated trend. These three time series originated from screens in extraction wells, and a plot of the time series revealed a significant and sudden drop in the head level at the end of the time series, as illustrated in Figure 14. The drop cannot be justified by changes in the amount of water extracted, and are thus likely to reflect a water level during pumping, and was removed from the time series.

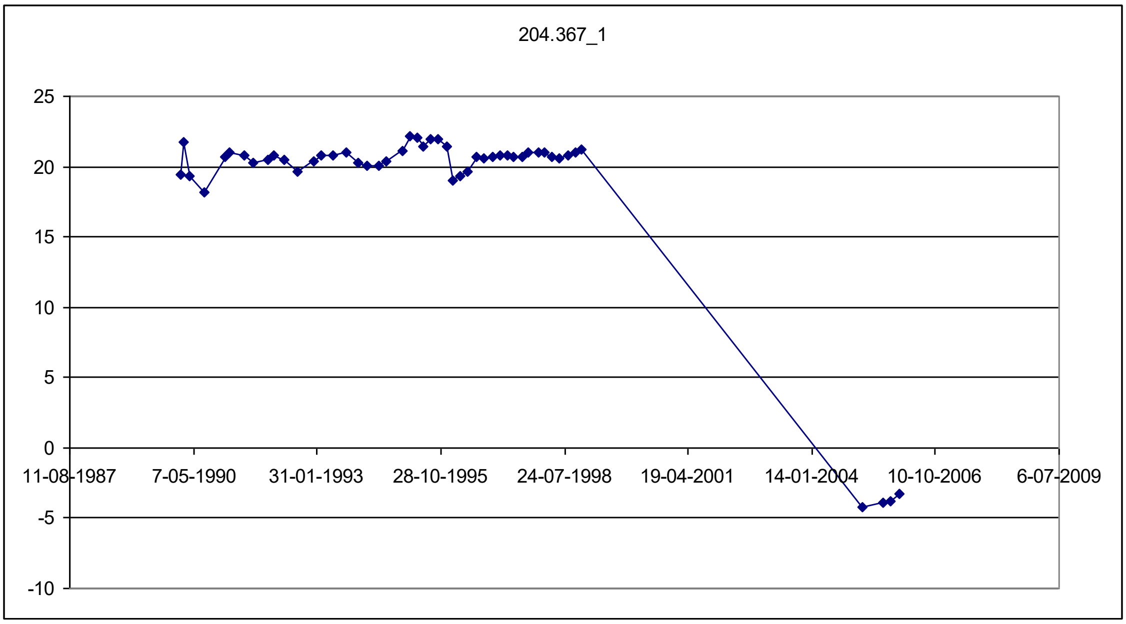

Figure 14. Time series with sudden drop, indicating the observations reflect the groundwater level during pumping.

Time series identified in cluster 4 did not show any distinct pattern. Outliers were identified and removed in four of the time series in cluster 4. Another seven time series in cluster 4 were removed completely due to suspicious data points. Some of these time series have large variation between minimum and maximum levels, Figure 15 (A), indicating that the observations are affected by pumping. Other time series show a distinct minimum or maximum level, indicating either: 1) misplacement of a datalogger with groundwater levels outside the dataloggers measuring interval 2) malfunction of a datalogger, or 3) erroneous measurements or reporting from manually observation. An example is provided in Figure 15 (B), where a maximum value of exactly 8.62 (more than 10 m above the mean level) is observed from 1996 to 1998, after which a new maximum of exactly -1 m is observed.

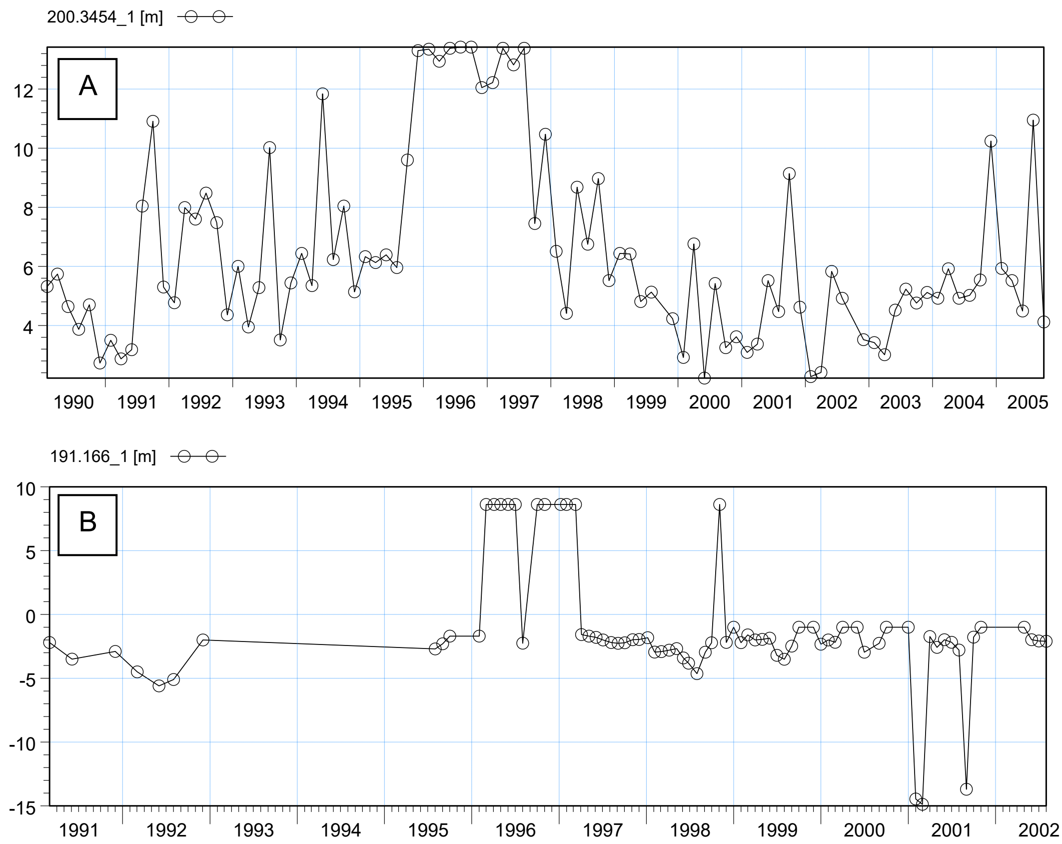

Figure 15. Examples of groundwater head time series in cluster 4, which have been removed from the dataset.

Simulated impact by extraction

Developments in groundwater heads are affected by both the actual extraction and the current climate. As both quantities vary over time, the reason for an observed development in the groundwater head level cannot be concluded on the basis of observation data alone. Model simulations have thus been carried out to aid in separation and quantification of the effect of the two components.

For monitoring purposes it is relevant to evaluate the impact of the actual level of extraction on the groundwater system. The effect of variations in the current climate was filtered out by constructing a standard climatic dataset resembling the present mean conditions. This was achieved by recycling the climatic data for 2004, which was found best to resemble the month mean precipitation for the period 1990 – 2005. Alternatively, a mean climatic year could be constructed by daily averaging over the entire period, i.e. 1st January is a mean of 1st January for all years, and so on. Recycling of a mean year is however preferred in general, as this preserves the day to day variation, which is smeared out by daily averaging. From 1990 to 2008 significant variations are observed in the amount of groundwater extracted, Figure 5. As discussed previously, the decrease in extraction observed from 2005 is believed to be a result of incomplete reporting to the national database, and not due to an actual decrease in the water extracted. For this reason, the mean extraction in the period 2000 – 2005 was chosen as representative for the actual amount of groundwater extracted.

A model simulation was carried out with standard climate and mean extraction as described above. A groundwater head resembling no extraction was applied as initial conditions, and the model simulation was run for 30 years, which corresponds to the typical timeframe for which extraction permits for drinking water purposes are issued. Two aspects may be critical when the impact of extraction on groundwater head is evaluated, namely the maximum drawdown and the temporal development. The maximum drawdown is critical for various reasons, such as risk for subsidence and oxidation of natural chemical components with subsequent release of toxic elements to the water phase. Similar important is, however, the evaluation of whether stationary condition will develop, or the groundwater level will continue to drop under the current conditions. The simulation was thus analysed in two ways:

1. The maximum drawdown caused by the actual extraction was determined as difference in groundwater head between the initial conditions (no extraction) and the groundwater head level after 30 years of pumping.
2. The head drop from 10 to 30 years after pumping start was computed. By this, it was assumed that the aquifer may be vulnerable to a continuously decrease in the groundwater head, if a new stationary situation was not reached after 10 years after pump start.

Simulated of impact by climatic changes

Areas most sensible to the expected future climatic changes are identified by comparing the development in the groundwater head applying current and future climatic conditions. As future climatic conditions, climate in the period 2071 – 2100 as predicted by the IPCC scenarios A2 and B2 were applied (IPCC, 2000), which represent scenarios characterised by relatively high and moderately atmospheric CO_2 released, respectively. The A2 and B2 scenarios have been used as standard scenarios in the past half-decade, but new scenarios have recently been developed by IPCC. The new scenarios are, however, not available for hydrological modelling yet, and the A2 and B2 scenarios have therefore been applied for illustration purposes.

The climatic scenarios were generated using data from the general circulation model (GCM) from Hadley Centre HadAM3H, which were downscaled dynamically using the regional climate model (RCM) HIRHAM developed by the Danish Meteorological Institute (Christensen et al., 1996, 1998). (http://prudence.dmi.dk/). The climatic data was bias corrected by the delta change method, where delta change factors were computed based on monthly mean values for the reference period 1961 – 1990. The climatic scenarios and bias correction is described in detail by Roosmalen et al. (2007).

In all simulations a constant extraction was applied corresponding to the mean extraction in the period 2000 – 2005. The model was run for a period of 17 years (1990 – 2006) with initial conditions corresponding to a groundwater head subjective to mean extraction 2000 – 2005 and the current climate. Simulated groundwater heads were averaged for the last 10 years and differences between the current climate and scenario A2 and B2 was computed. With the applied initial heads the initial condition is close to the expected end situation, and a warm up period of six years was assumed adequate for the two climatic scenarios. The results therefore represent a dynamic equilibrium under the different climatic conditions and current extraction levels, and differences among the simulations with different climatic input thus resemble long terms differences due to different climatic conditions.

Discussion and results

Simulated trends

The result of the analyses on the models ability to simulate the overall trend in the groundwater level for Sjælland is shown in Figure 16, comparing the observed and simulated normalised heads. From the figure it is seen that the model is able to reproduce the year to year variation well. The overall development in the groundwater head for the period 1993 – 2005 is slightly underestimated by the model, as illustrated by the linear trend lines. The model predicts a mean increase of 0.0471 m/year in the period compared to an increase of 0.0582 m/year based on observation data. The difference in the linear trends for observed and simulated data is dominated by the discrepancies in the very dry years (1996 and 1997) and to a smaller extent by the wet year 2002. Although the model does not capture the yearly variation for the most dry/wet years, it reproduces the overall development satisfactorily.

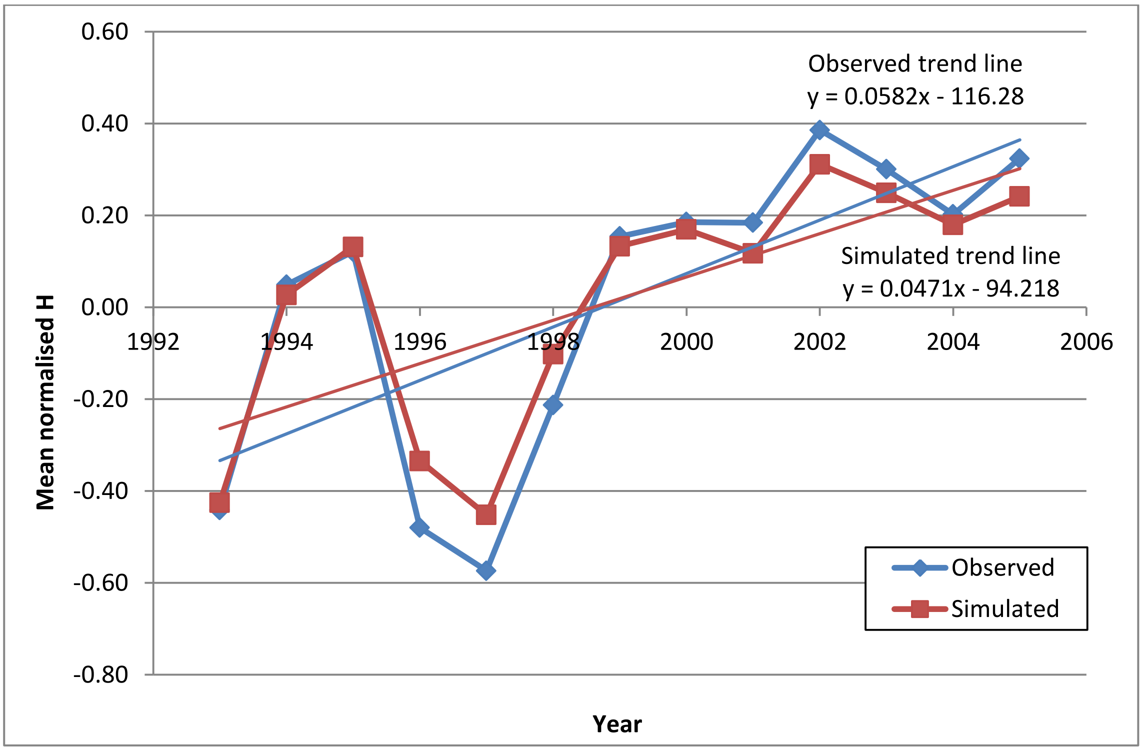

Figure 16. Observed and simulated mean normalised head for the 1175 screens included in the study.

The trend in the groundwater head varies greatly for the individual screens with both positive and negative developments in the groundwater level, as illustrated in Figure 17, displaying the slopes computed from the observed time series at the individual screens. From the figure it is also observed that approximately 45 % of all time series have an absolut slope less than the mean slope for all screens.

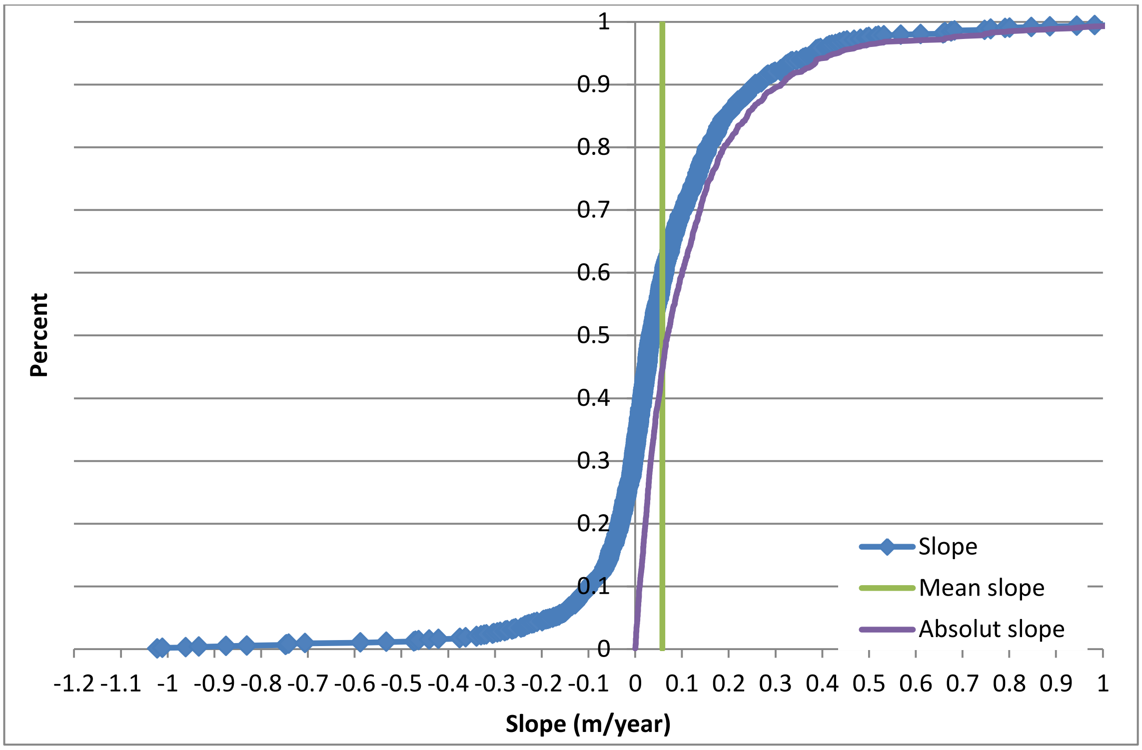

Figure 17. Slopes in linear regression computed from observed time series

Comparison of observed and simulated trends is shown in Figure 18. The results indicate that the dataset may be further divided into different cluster for additional manual interpretation, but this was not done in the present study.

The trend line in Figure 18 is based on the entire dataset. With a slope of 0.39 it is seen that the model in general underestimates the observed trend. Deviation from the 1 : 1 line is primarily caused by the screens with the highest observed increase in groundwater level (largest positive observed slope), where the model simulates a much smaller increase. Similarly, for a smaller group of screens a significant decrease in the groundwater level is observed (negative slope), while the model does not predict any significant development (slope close to zero). The majority of the data are, however, clustered around the 1 : 1 line.

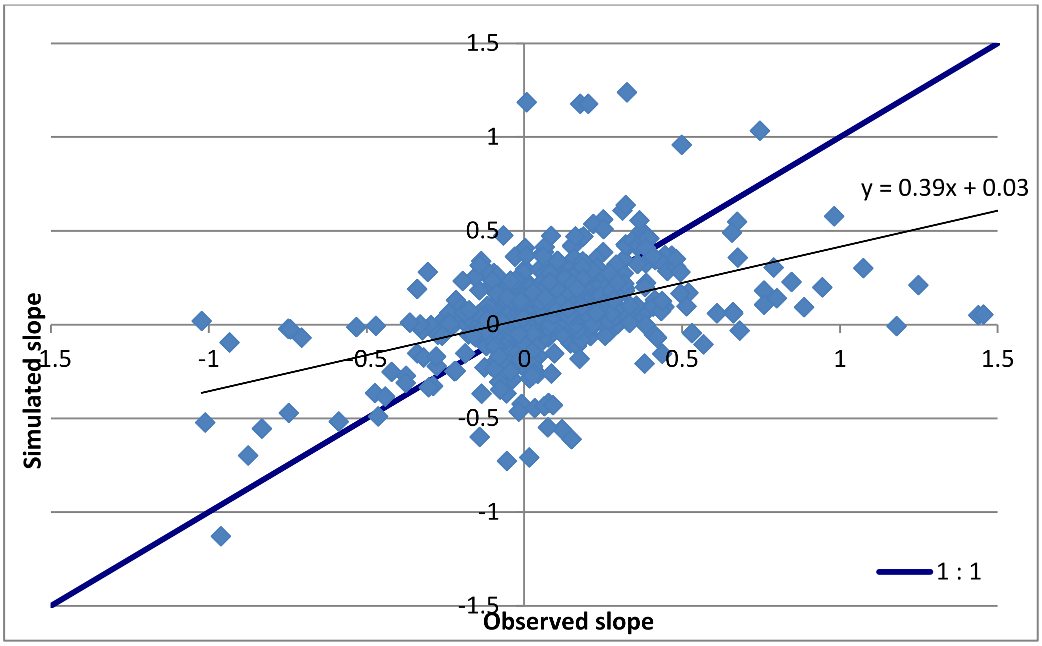

Figure 18. Simulated versus observed slope for all screens.

Using only data located more than half grid cell from an extraction well a few outliers with large deviation from the 1:1 line is still observed, Figure 19, but the datapoints are in general closer to the 1:1 line with a slope of 0.61. It is especially noticed that the aforementioned datapoints close to the x-axis are due to observations in extractions wells.

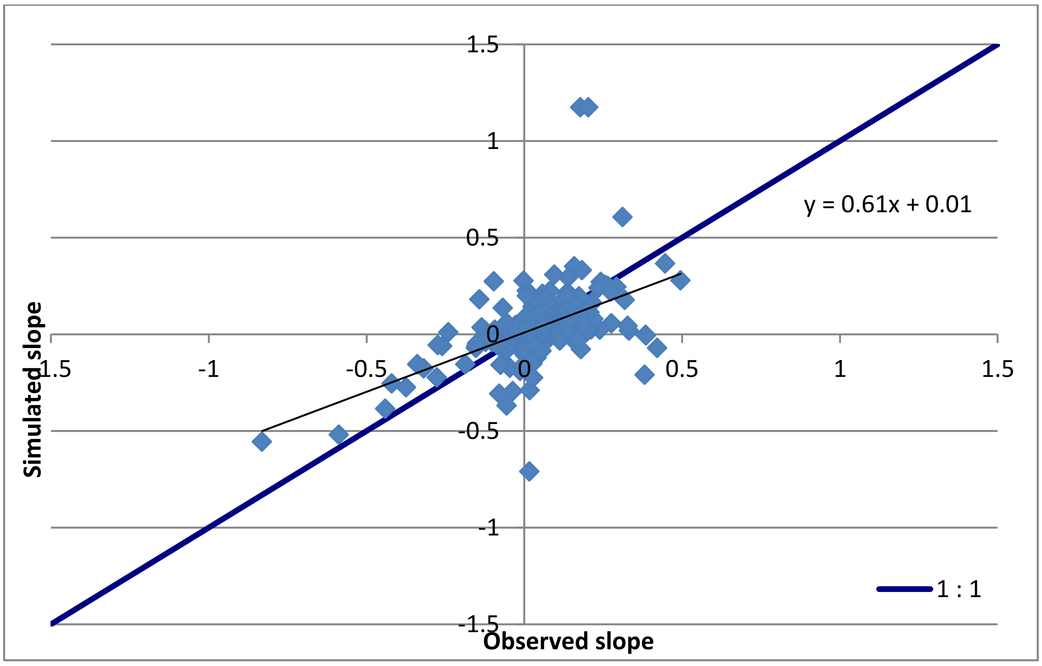

Figure 19. Simulated vs observed slope for screens located more than half a grid cell from an extraction well.

Three situations exist with respect to observed and simulated trends: 1) the observed and simulated trends are consistence (both either positive or negative), 2) the simulated trend is positive, while a negative trend is observed (false positive trend), and 3) the simulated trend is negative while the observed is positive (false negative trend). The critical conditions arise where the model simulate the reverse of the observed trend. Screens with simulated false positive or negative trends are shown in Figure 20.

Computation of the trends in the individual screens do not include an analysis on whether the trend is statistical significant. However, the computation of trends is sensitive to the start and end conditions. Due to seasonal variations different results may be obtained if the start and end conditions are at seasonally high or low groundwater level. This is in particular problematic where very low trends are estimated, and data points with an observed trend with an absolute value less than 0.6 cm/year (one tenth of the overall observed trend for the entire period) is not included in Figure 20.

A false positive or negative trend is simulated for 313 screens (slope > ± 0.6 cm/year), of which 229 is located closer than 250 m from an extraction well (half a cell size), see Figure 20. The majority of the screens where the model simulates the reverse trend are thus associated to extraction. Of the remaining 84 screens, 63 have a slope less than the mean slope of +/- 6 cm/year.

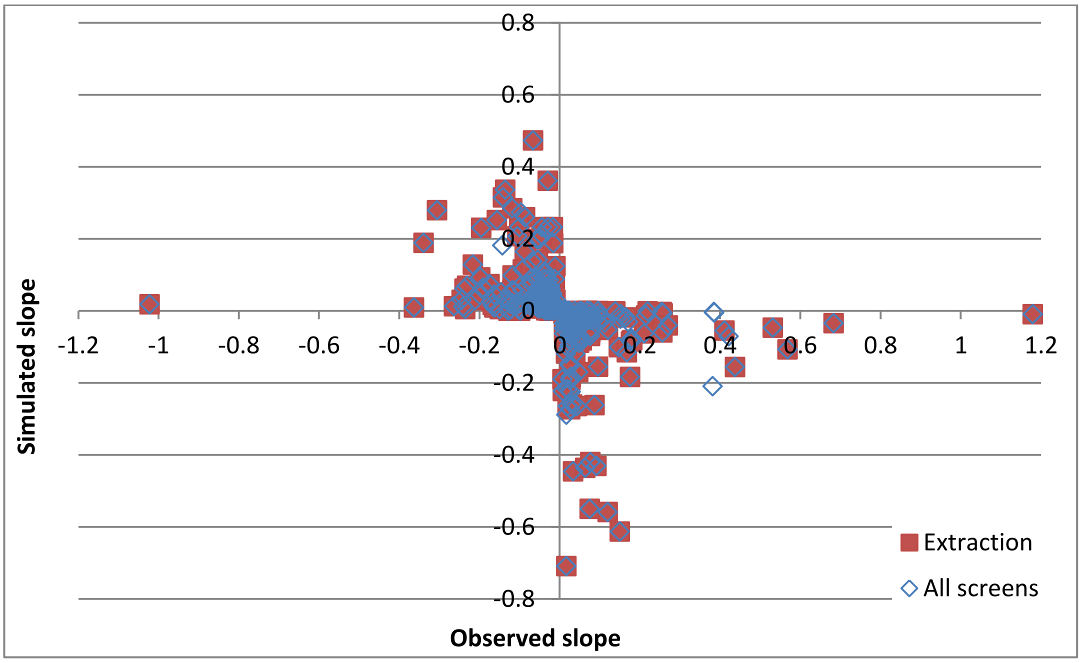

Figure 20. Data points where the model simulates false positive or false negative trends.

In Figure 21 observed and simulated slopes are grouped by the model layers. From this it is seen that the two groups with most deviation between observed and simulated development primarily is associated to the chalk/limestone aquifer.

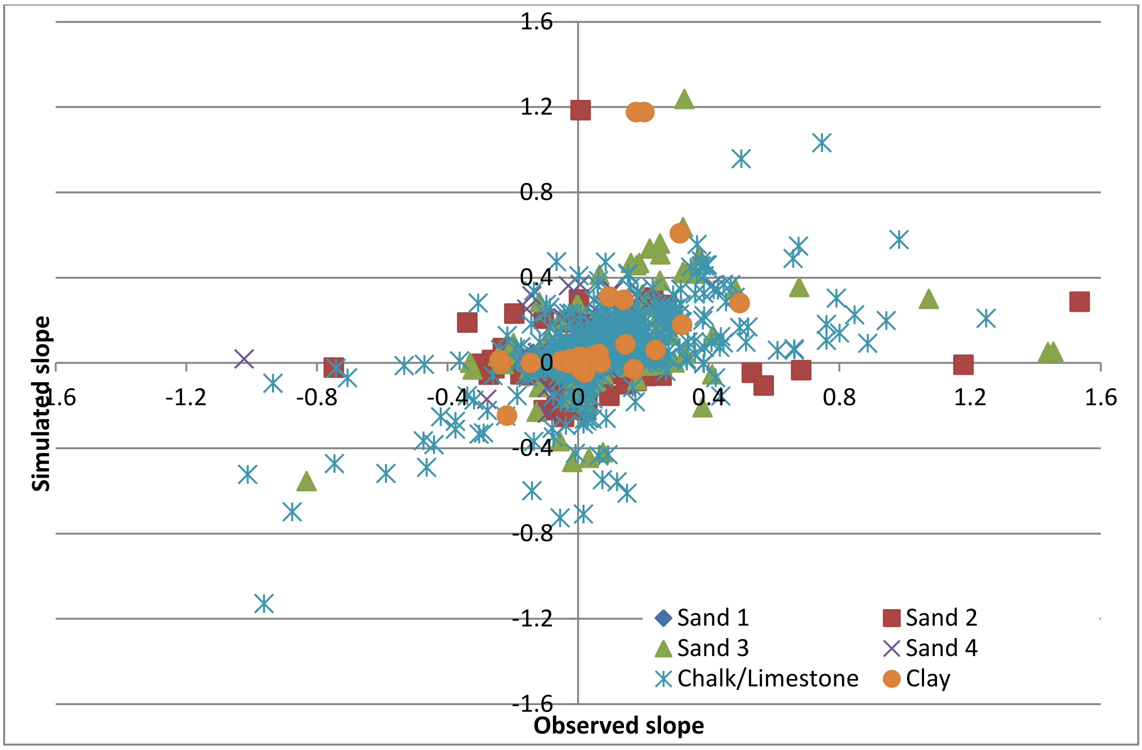

Figure 21. Simulated versus observed slope for individual model layers.

Figure 22 and Figure 23 displays the spatial distribution of observed versus simulated trends for the individual screens in the three most important sand aquifers and the lower Chalk/Limestone aquifer. Screens in which the absolute slope of the head development is less than on tenth of the overall development for the entire period (6 cm/year) are marked in the figures below, indicating that results from these may be less reliable.

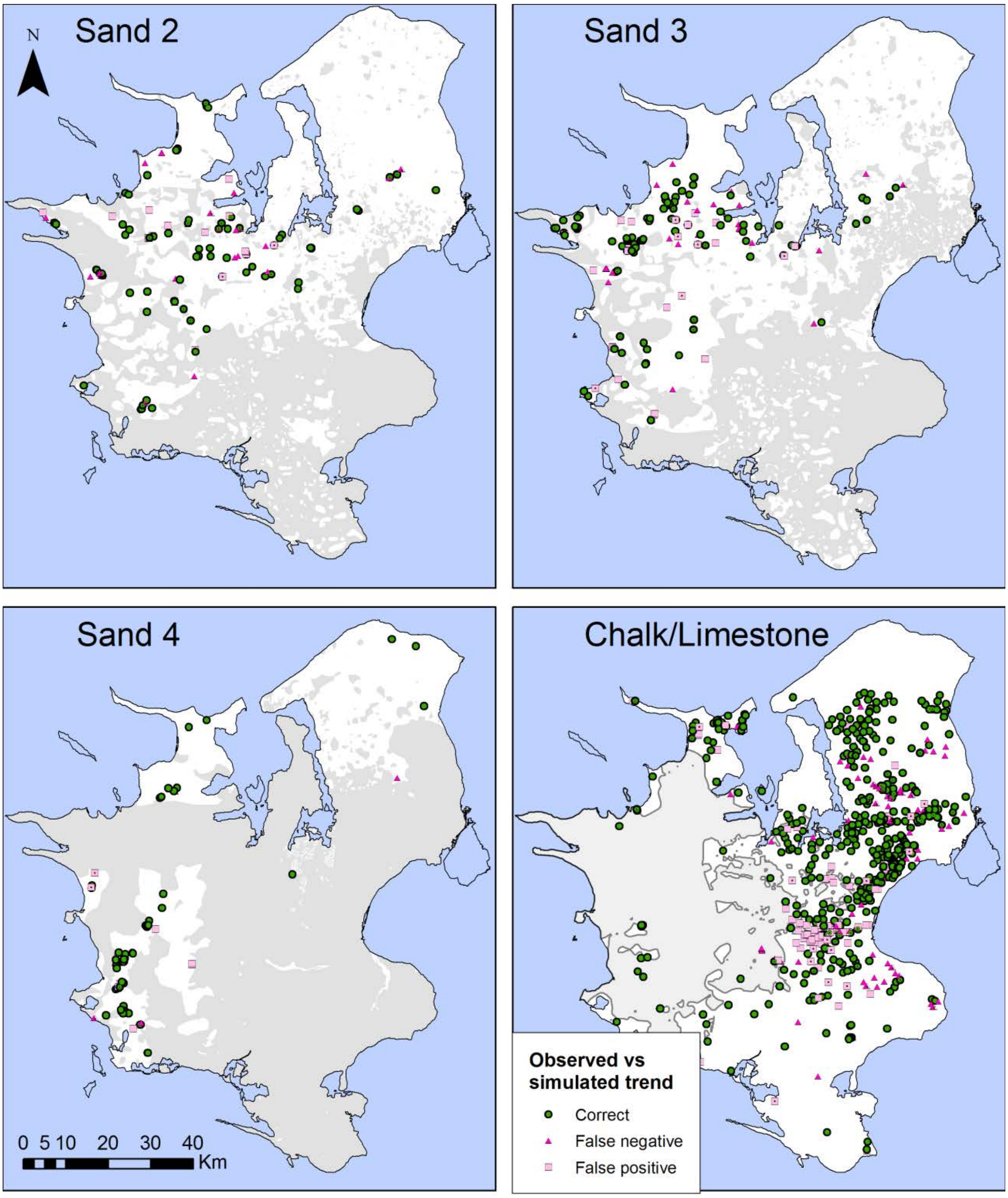

Figure 22 Observed versus simulated trends in individual screens, screens in which a trend of less than 0.6 cm per year is estimated are indicated by a dot in the symbol.

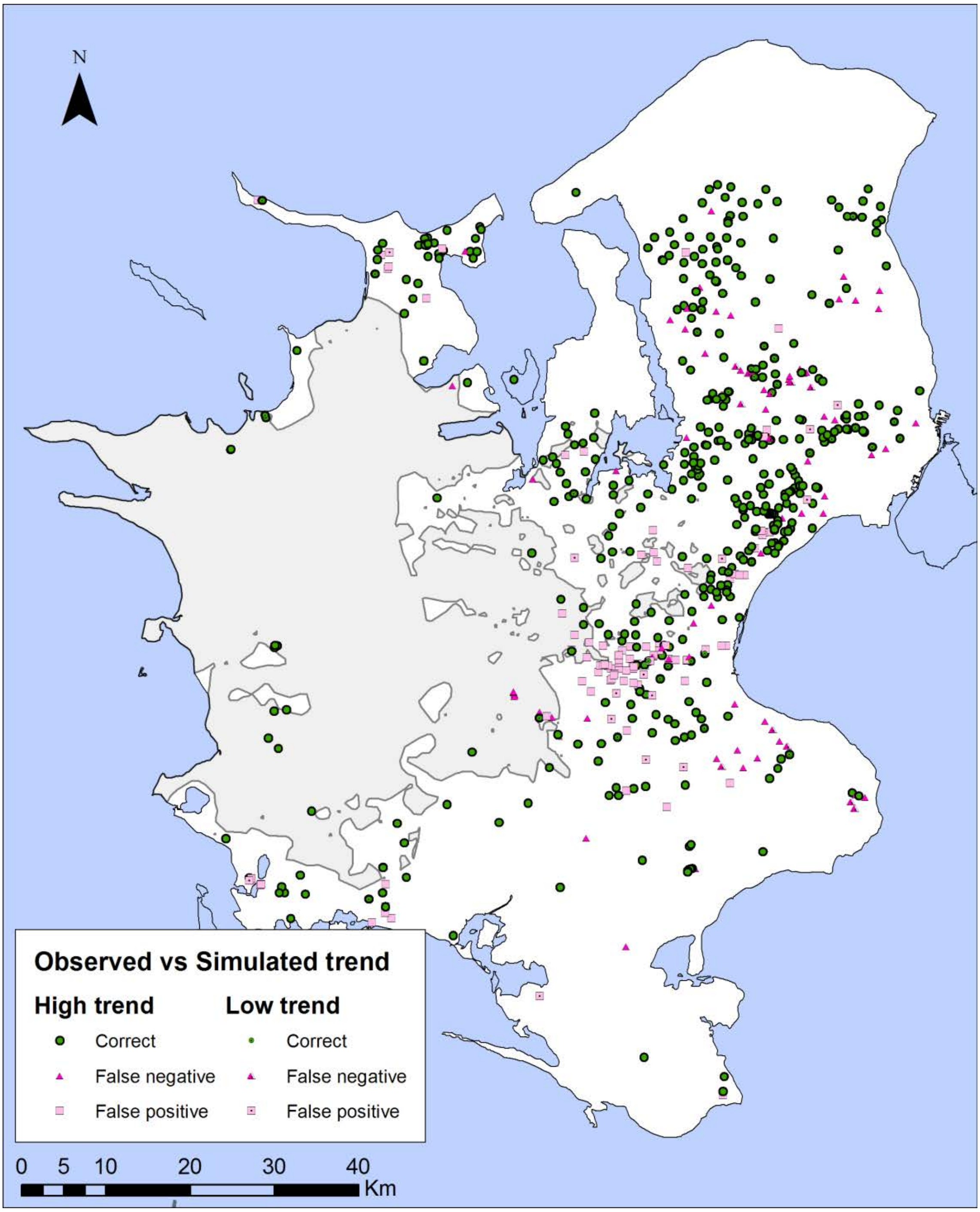

Figure 23 Observed versus simulated trends in individual screens in the Chalk/Limestone aquifer. Screens in which a trend of less than 0.6 cm per year is estimated are indicated by a dot in the symbol

In large parts of the area the model is able to predict the sign of the trend, i.e. whether the historical data on climate and extraction results in an increase or decrease in the groundwater level. However, discrepancies between observed and simulated trends are similarly observed. Several reasons may cause the simulation of a reverse trend, among which the most important are expected to be 1) errors or missing details in the hydrogeological model, 2) neglecting of heterogeneity in the hydrogeological parameters, and 3) uneven distribution of extraction rates among screens in a well field, (in the model it is assumed that the extraction is equally distributed among all active screens in a well field).

For the three sandy aquifers, the screens in which the reverse trend is simulated are relatively scattered indicating that the discrepancies are caused by local features such as 1) and 3) above. In the Chalk/Limestone aquifer the false negative and false positive tend to group, which is particularly observed in the central-eastern part of Sjælland around Køge. Most of the screens with a simulated false positive trend in this area are associated to well fields.

In summary, the analysis illustrates that the observed trends are associated with large variations that includes both negative and positive trends. The model is able to reproduce the sign of the trends for the majority of the screens. Discrepancies between observed and simulated trends are predominantly associated to screens close to extractions, and may be caused by an insufficient hydrogeological model or erroneous distribution of extractions among screens in a the well fields. For screens located more than half a grid cell size (> 250 m) from an extraction, errors in the simulated trends are predominantly associated to small trends.

Overall, the model is expected to be able to predict the development in the groundwater heads with sufficient accuracy to be used in the further analysis of the monitoring network.

Impact by extraction

The simulated impact from groundwater extraction on groundwater head development is illustrated in Figure 24 and Figure 25. Initial condition in the simulation is no extraction, and the model is run for 30 years with mean extraction for the period 2000 – 2005. Figure 24 displays the total decrease in groundwater head after 30 years pumping, i.e. the maximum drawdown introduced by pumping. Decrease in the last 20 years (from 10 to 30 years after start of pumping), Figure 25, indicates that it takes a very long time for a new dynamic equilibrium to established, and there is a risk for groundwater mining with continuously decrease in the groundwater level. Figure 24 and Figure 25 includes the delineation of the regional and deep groundwater bodies, superimposed on the model layer to which they have been associated by the regional water authority.

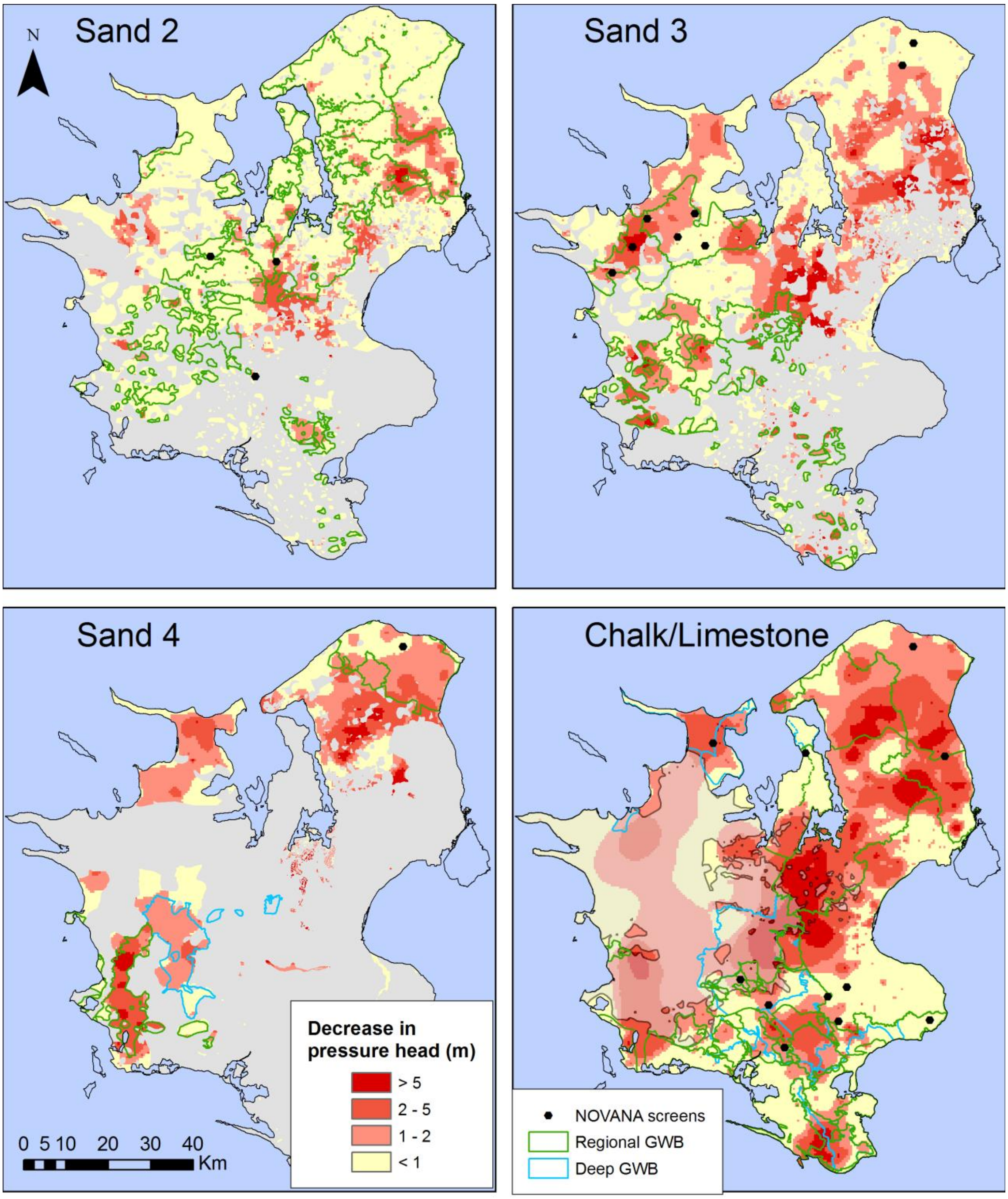

Figure 24. Simulated decrease in groundwater level by 30 years extraction.

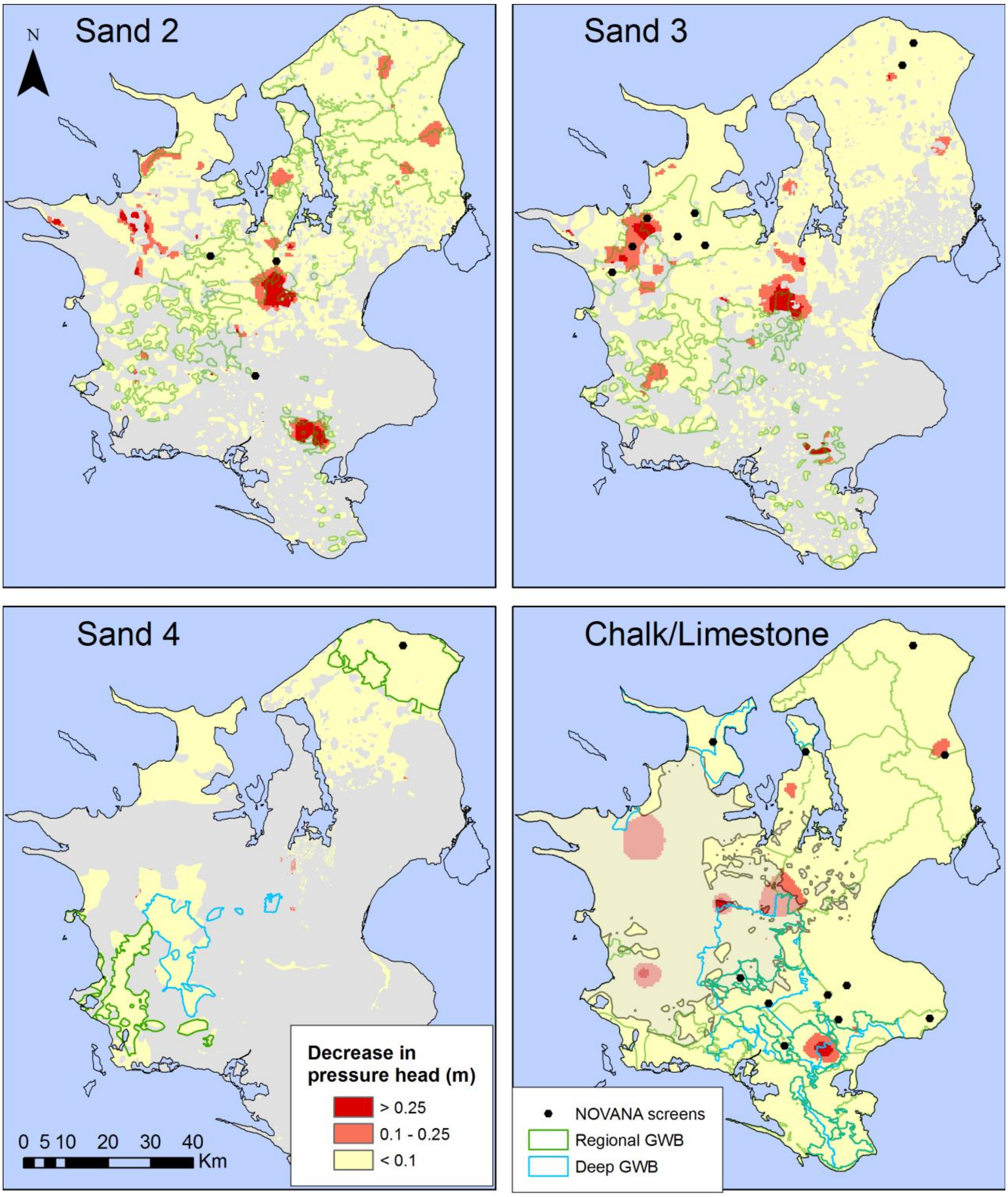

Figure 25. Decrease in groundwater level from 10 to 30 years from start of pumping.

Impact by climatic change

Differences in simulated groundwater heads using current climatic conditions and the A2 and B2 climatic scenario are displayed in Figure 26 and Figure 27, respectively. Noticeable from the figures are that the two future climatic scenarios displays the same pattern, but with the most impact by the B2 scenario. It is also noted that all values are positive, meaning that the groundwater level is expected only to increase. The overall pattern of the impact by climatic change is comparable to a previous study on the effect of climatic changes on Sjælland (Sonnenborg et al., 2006), although the previous study also predicted areas with decreasing groundwater levels in the western part of Sjælland. The two studies are based on two different versions of the DK-model, and the primary reason to the differences between the two studies is expected to be due to the recent update of the geological model in the DK-model, but also due to differences in conditions such as simulation period and abstraction data. The objective of a present study under KFT (Koordineringsenhed for forskning i klimatilpasning) is to simulate the impact climatic change has on the groundwater levels nationwide using a new climatic scenario A1B. The study includes also an assessment on the uncertainty on such simulations. The simulations on the impact by climatic change in the present study should thus only be considered as illustration on how such simulation can help in the analysis of the monitoring network.

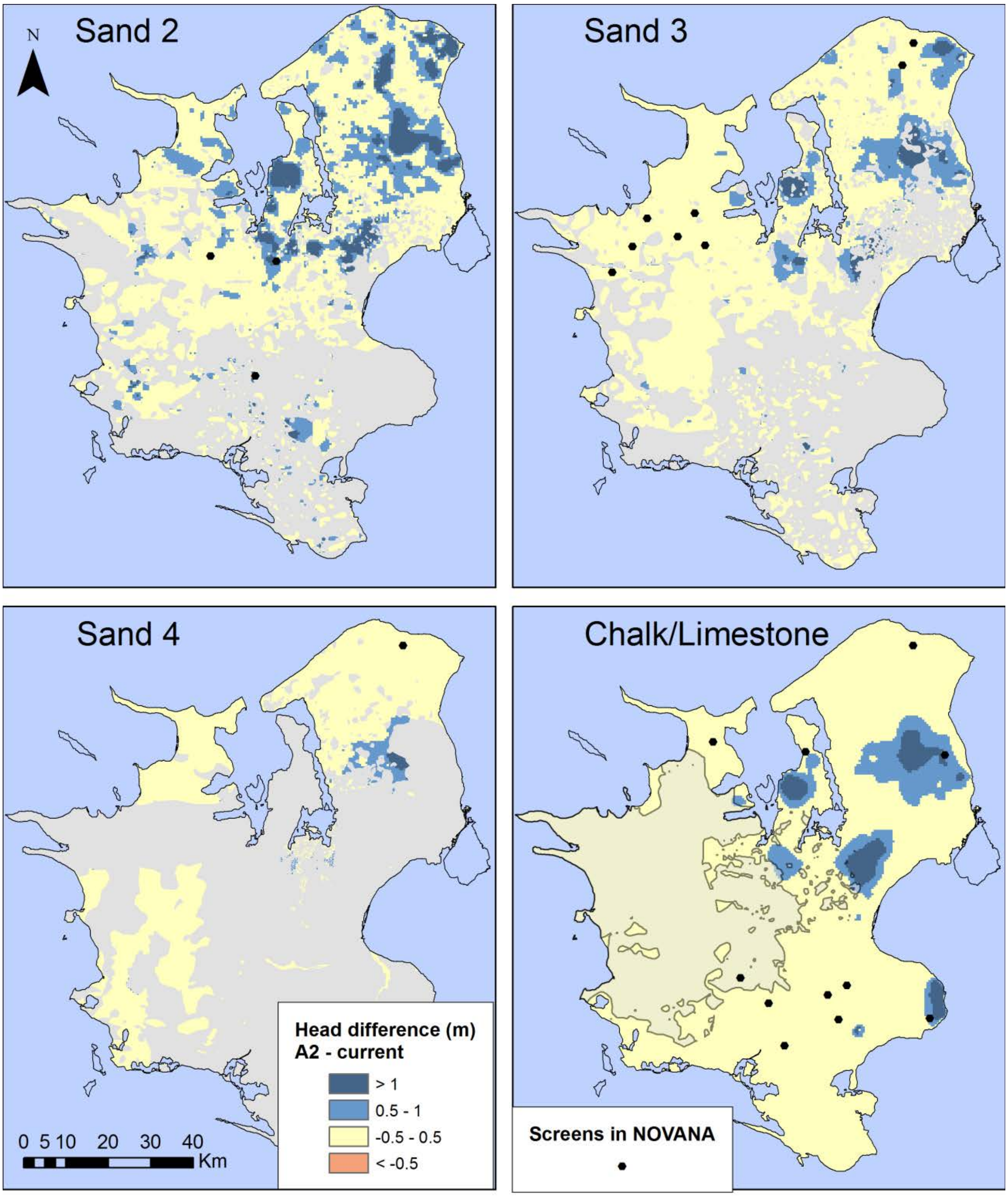

Figure 26 Difference in drawdown using climate scenario A2 compared to current climate.

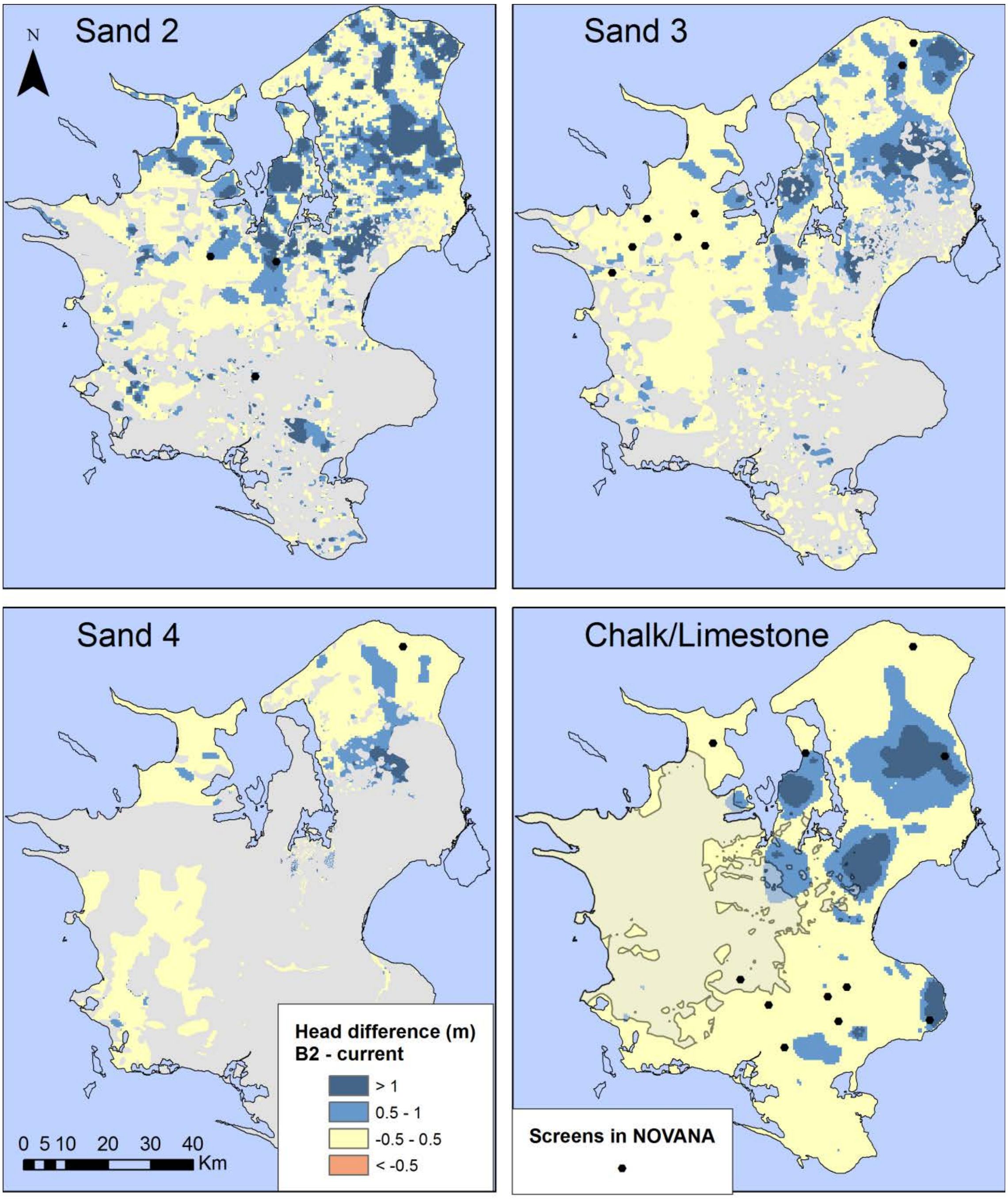

Figure 27 Difference in drawdown using climate scenario B2 compared to current climate.

Evaluation of the monitoring network

A first step in the evaluation of the present monitoring programme can be carried out by combining the spatial distribution of the screens included in the monitoring network with the simulated impact by extraction under the current climate, Figure 24 and Figure 25, and simulated effect due to climatic change, Figure 26 and Figure 27. This comparison is summarised in Table 3 where the expected impact from extraction and climatic change for each screen is categorised into three classes according to Table 4.

Table 3 Expected impact by extraction and climate for monitoring screens on Sjælland, ordered by groundwater body.

ID	Model layer	Drawdown Total	Drawdown > 10 years	Impact by climate	GWB
216.625_5	3	Medium	Low	Medium	?
216.625_1	8	Medium	Low	Medium	?
204.397_2	1				2.1.2.4
203.90_1	5	Medium	Low	Low	2.1.2.4
204.397_1	5	High	Medium	Low	2.1.2.4
191.102_1	10	Medium	Low	Low	2.1.3.1
192.11B_1	10	Low	Low	Low	2.2.2.11
197.166_1	5	Medium	Medium	Low	2.2.2.13
197.334_1	5	Medium	Low	Low	2.2.2.13
197.476_2	5	Low	Low	Low	2.2.2.13
205.336_1	5	Low	Low	Low	2.2.2.13
205.342_1	3	Low	Low	Medium	2.2.2.2
206.1237_1	3	Low	Low	Medium	2.2.2.6
182.402_1	5	Low	Low	Low	2.3.1.1
182.319_1	10	Medium	Low	Low	2.3.2.1
194.129B_1	10	Medium	Low	High	2.3.2.1
182.317_1	7	Low	Low	Low	2.3.2.3
187.1057_1	5	Medium	Low	Medium	2.3.2.6
217.163_1	10	Low	Low	Low	2.4.2.1*
217.206_1	10	Low	Low	Low	2.4.2.1*
216.272_1	10	Medium	Low	Low	2.5.2.32
216.529_1	10	Medium	Low	Low	2.5.2.32
217.474_1	10	Low	Low	Low	2.5.2.32
221.278_1	10	Medium	Low	Low	2.5.2.37*
218.343_1	10	Low	Low	High	2.6.2.11

Table 4 Classification of Low, Medium and High impact based on head level change caused by extraction or climatic change

	Drawdown in m due to pumping Total	Drawdown in m due to pumping > 10 years after start of pumping	Head change in m due to climatic change
Low	< 1	< 0.1	< ± 0.5
Medium	1 – 5	0.1 – 0.25	± 0.5 - ± 1.0
High	> 5	> 0.25	> ± 1.0

The summary in Table 3 provides some basic insight into how the different screens are expected to respond to the impact from extraction and climatic change. This is valuable information in the analysis of the monitoring data and redesign of the monitoring network. For example, optimal conditions for monitoring effects of future climatic changes are locations with low impact from extraction and high sensitivity towards climatic change. On the other hand, monitoring impact from extraction and the possible long term groundwater mining must be based on screens located in areas indicated in the simulations as affected by extraction. Table 3 shows that the monitoring programme includes screens for both examples. However, it is also obvious from Table 3 that not all screens are located optimal. Eight screens (one third of all screens) are located in areas where the effect by extraction and climate are both expected to be low, and in four of the groundwater bodies all screens are expected to have low sensitivity towards both pressures. While it may be relevant to have some screens located in such areas, they should not comprise the majority and not be the only screens within a single groundwater body.

From Table 3 it is also evident that only 14 of the 63 regional and deep groundwater bodies on Sjælland are monitored, most with only one screen, but five groundwater bodies have two or more screens.

According to the WFD groundwater bodies may be grouped for monitoring purposes. This means that all groundwater bodies do not need to be monitored, but it must be assured that the monitored location provides sufficient information on the remaining groundwater bodies in the group. In other words, a relation must be established between groundwater bodies within a single group that allows information (monitoring) in one groundwater body to be transferred to the other members of the group. Such relations have not been defined at present.

Central in the strategy for grouping groundwater bodies should be the expected impact from extraction. Two measures have been defined in the previous sections: 1) total drawdown in groundwater head due to extraction and 2) long term decrease in the groundwater level indicating groundwater mining. Groundwater bodies in which a dynamic equilibrium is not reached fails to meet a good quantitative status, and can as such not be grouped. Comparison between groundwater bodies with significant areas identified as converging slowly to a dynamic equilibrium (Figure 25) and the assignment of a poor status by the water authorities agrees well, with the exception of two groundwater bodies (DK2.5.2.23 and DK2.6.2.10) located in Sand 2 and 3 in the southern part of Sjælland. Here the model simulation indicates slowly or failing convergence to stationarity for the entire groundwater bodies while they have been defined as being in good quantitative status.

Grouping is relevant to consider for groundwater bodies in which a new dynamic equilibrium is expected to develop. A possible criterion for grouping can be based on a measure of the simulated drawdown caused by extraction. Since the groundwater bodies have not been delineated on the basis of impacts from extraction they are characterised by being affected by extraction to varying degree in different areas within a single groundwater body. A simple average of the total drawdown over the entire area of the groundwater body will indicate the degree to which the groundwater body is exploited and may thus be helpful information for the characterisation of a groundwater body having a good or poor status. For grouping a simple average is, however, not an adequate measure, as groundwater bodies with high impact in only a small percentage of the area, will be classified as being less affected by extraction than a groundwater body with medium impact in the entire area. A more useful measure may thus be an average only in areas where some impact is observed.

For illustration purposes, the average drawdown for groundwater bodies located in Sand 3 is shown in Figure 28 computed on the basis of the entire groundwater body (Mean_all) and areas in which the drawdown is greater than a threshold value of 1 m (Mean_gt1). The figure also illustrates groundwater bodies being classified by the regional water authorities as having a good or poor status. It is noted that two groundwater bodies do not have a value for Mean_gt1, because the simulated drawdown was less than 1 m in the entire areas for these groundwater bodies. From Figure 28 it is seen that the average drawdown computed on the basis of the entire groundwater body (Mean_all) and the assigned quantitative status agrees well. Groundwater bodies with a Mean_all of approximately 1 m or less have been classified as being in good status, while groundwater bodies with larger drawdown are classified as being in poor status. When the average drawdown is computed only for areas with a drawdown greater than 1 m (Mean_gt1) the groundwater bodies group a little different, as groundwater body DK2.2.2.13 (poor status) has a Mean_gt1 similar to the groundwater bodies classified as being in good status.

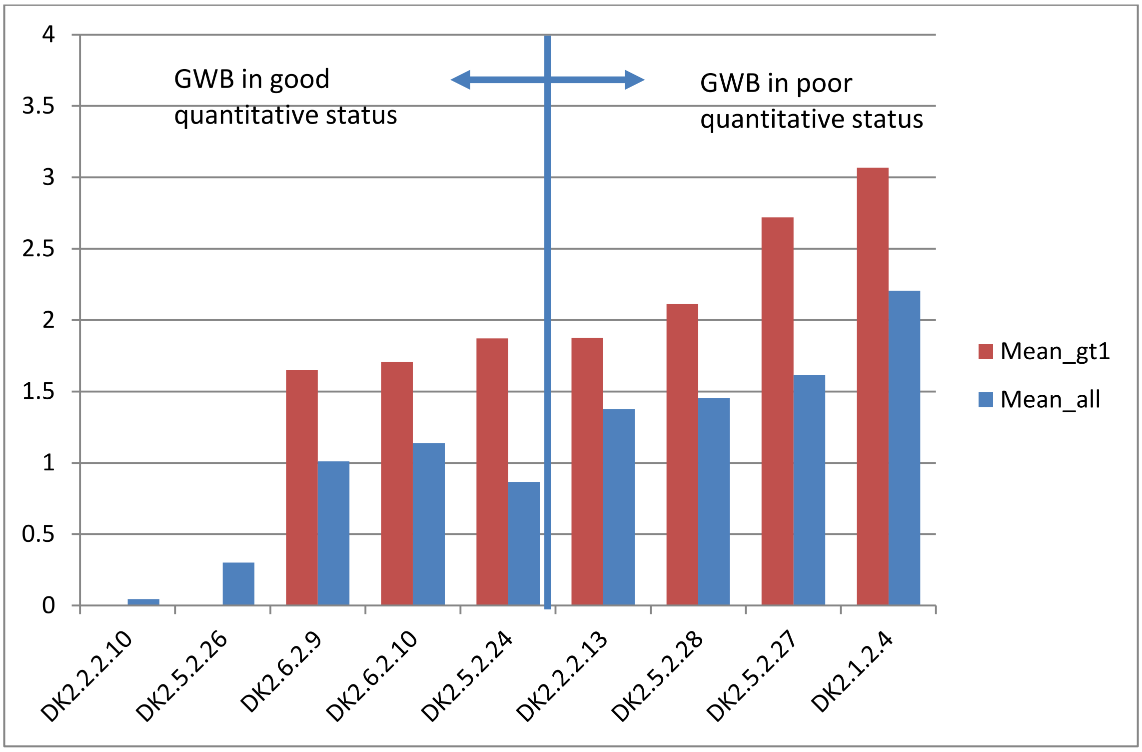

Figure 28 Average drawdown in groundwater bodies in Sand 3 computed as an average for the entire groundwater body (Mean_all) and as an average for areas where the drawdown is greater than 1 m (Mean_gt1)

Assuming that all groundwater bodies in Sand 3 were classified as being in good quantitative status Mean_gt1 can be used to group the groundwater bodies. For illustration purposes we define four groups, Figure 29: 1) Drawdown less than 1 m in the entire groundwater body, i.e. Mean_gt1 = 0; green in Figure 29, 2) 1 m < Mean_gt1 < 2 m; blue in Figure 29, 3) 2 m < Mean_gt1 < 3 m; yellow in Figure 29 and 4) Mean_gt1 > 3 m; black in Figure 29. Subjective evaluation of the grouping by visual inspection of Figure 29 confirms that such grouping appears meaningful. However, the intervals used for grouping the groundwater bodies are arbitrary and only used for illustration purposes.

In Figure 29 some inconsistencies are observed between the extent of Sand 3 and the groundwater bodies, where the groundwater bodies are defined in areas where Sand 3 does not exist (grey shaded areas) this is probably due to the fact that the groundwater bodies were delineated on the basis of an earlier version of the geological model than the one included in the version of the DK-model used in the present study. Furthermore it is noted that no groundwater body has been defined in the north-eastern part of Sjælland, although Sand 3 exists in most of this area and is exploited as shown in Figure 4. Some screens in NOVANA are located in the north-eastern part of Sjælland, but they have been associated to groundwater bodies in deeper layers.

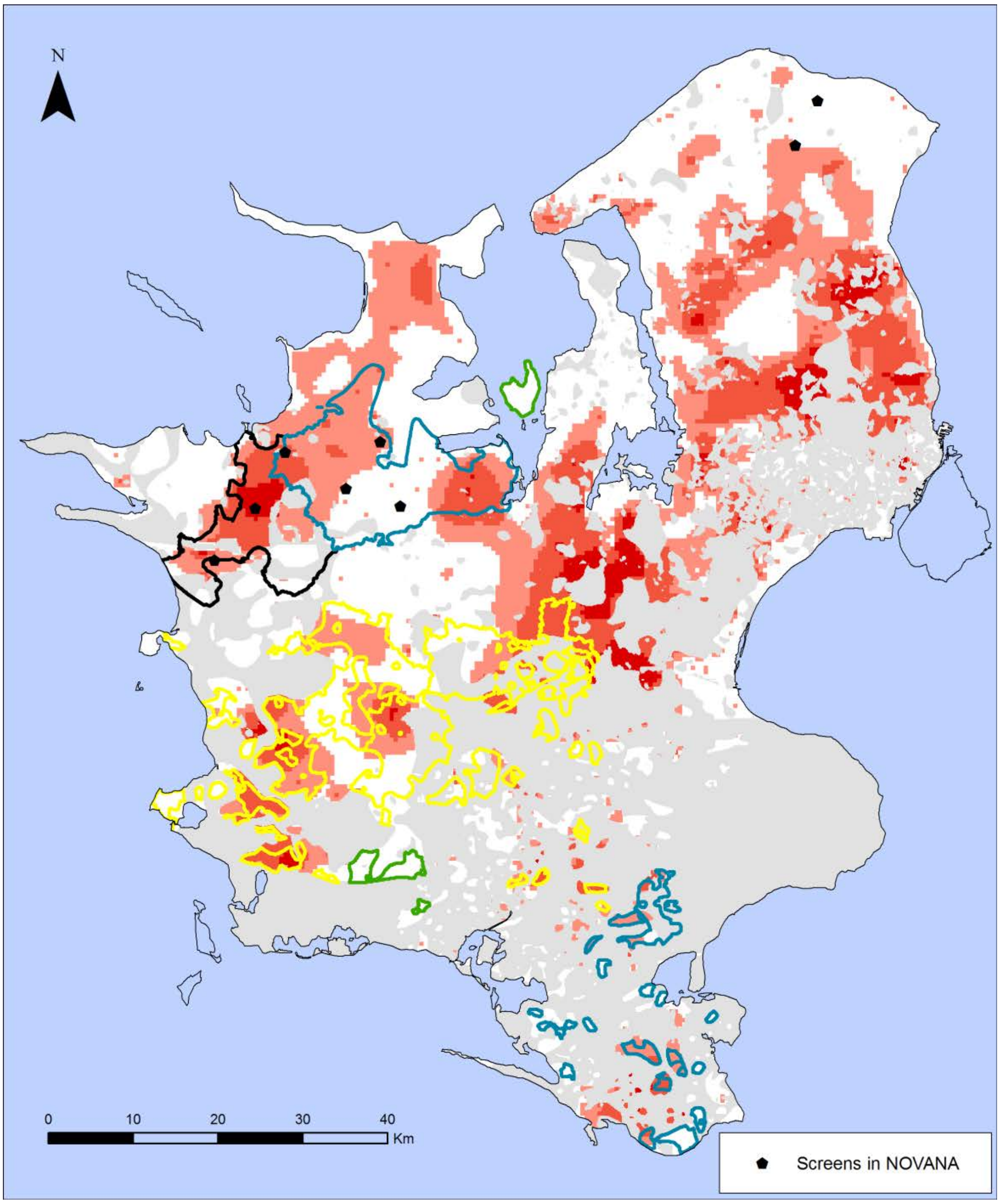

Figure 29 Example of groundwater body grouping in Sand 3

Following a grouping of the groundwater bodies the best strategy for monitoring must be defined, which include the number and spatial location of the monitoring screens. The WFD prescribe that the monitoring network must provide a coherent and comprehensive overview of the status, but specific design criteria are not provided, which are therefore to be defined at the national scale. Such design criteria have not been established in NOVANA.

In the design phase it must be evaluated whether a large spatial coverage with one or few observation points in many groundwater bodies are superior to a network design based on more intensive monitoring in few groundwater bodies, from which knowledge can be extrapolated to other areas. Common approaches for evaluation and design of a monitoring network is to analysis the variations in the data, and design the network so as to optimise the description of the variability, i.e. include screens in the network that best covers the variations in the aquifers (Gangopadhyay et al., 2001; Winter et al., 2000). With the use of the DK-model, and its expected ability to simulate the impact by extraction and climate acceptable, the model can be used to target the monitoring network in terms of identifying locations expected to display different aspects, such as high or low impact from climatic change and/or extraction.

Four screens are included in NOVANA for groundwater body DK2.2.2.13 (blue groundwater body in north-western Sjælland in Figure 29). Based on the model simulation it is expected that the four monitoring screens provide supplementary information, i.e. they are not all located in areas where the impact is expected to be similar. The configuration of screens in DK2.2.2.13 thus provides a comprehensive overview of the status in the groundwater body, which can be extrapolated to the other groundwater bodies. Such configuration may thus be superior to one screen in each of the blue groundwater bodies.

Establishment of design criteria is required before a detailed analysis and re-design of the network can be accomplished. In the example herein, the simulated total head change due to extraction is used as design criteria. This criterion may be combined with a criterion on optimising the variation among areas displaying identical overall hydrological response. Following the approach with different categories of impact (low, medium and high) the dynamics in each of the category may be analysed by the model for all groundwater bodies and sampling points representing one category selected across the groundwater bodies belonging to one group so as to optimise the variation within the category. Monitoring of one category may thus be composed of monitoring points in different groundwater bodies. Alternative criteria may similarly be chosen, such as water extracted expressed as a percentage of the groundwater infiltrated to the groundwater body or the area affected by extraction. Furthermore, meaningful intervals used for grouping must be identified allowing the groundwater bodies to be grouped on the basis of the expected impact, e.g. low, medium and high impact. In a previous study indicators have been used to assess the exploitable groundwater resource at the national level (Henriksen et al., 2008). However, these indicators were developed at a larger scale and cannot be directly transferred to the scale of groundwater bodies.

In addition to the design criteria, selection of monitoring locations, i.e. which groundwater body to monitor within a group of groundwater bodies, needs to take some basic aspects into consideration:

1. *Spatial extent of the groundwater bodies*. The group of four groundwater bodies indicated by blue in Figure 29 consists of one larger groundwater body and three

groundwater bodies composed of small isolated sand units. Monitoring data from the isolated sand units may display local features that may not be easily extrapolated.

2. *Uncertainty in simulated effect*. The evaluation on the models ability to simulate the trend in the development of the groundwater heads can be used as supplementary information for selecting the monitoring points. Where the model proves to simulate the trend from historical data correctly, confidence can be given to the model with respect to its ability to simulate the effect of variations in climate and extraction strategy, and more uncertain locations can be given higher priorities for monitoring.
3. *Location of existing observation points.* Calibration and validation of the national water resources model are based on all observation data from existing screens, and not limited to data from the monitoring programme. The density of monitoring data varies spatially, with low density in some areas limiting the possibility of testing the national water resources model. When new locations are included in NOVANA, areas with low density should be prioritised for improvement of the spatial coverage
4. *Use of existing screens*. If existing screens are adopted by the national monitoring programme, the following issues must be considered in the selection
    a. *Existing time series.* The detection of trends in time series requires that data exists for several years. Screens with long time series of good quality are thus preferred to screens with limited data
    b. *Technical aspects*. Before a screen can be adopted into the national monitoring programme the technical installations must comply with the standards defined in NOVANA
    c. *Administrative aspects*. Arrangement must be made with the local landowner to secure a future existence and access of the screens

Summary and recommendation

The present study illustrates an approach by which the national water resource model can support in the evaluation and possible re-design of the national monitoring programme on the quantitative status. The approached is based on model simulations used to evaluate the development in the groundwater head due to extraction and climatic changes. Results from the model simulations form the basis for identifying areas suitable for monitoring changes in the groundwater head due to the expected climatic changes. For monitoring of the general development in the groundwater head, groundwater bodies may be grouped. In the present study, it is illustrated how the model simulations can be used to aid in such grouping.

A prerequisite for model use is that the model is able to reproduce the observed development in the groundwater level. A test on the DK-models ability to reproduce the long term trends in the development of the groundwater head illustrated that the model is able to reproduce the overall development satisfactorily. Comparing the observed and simulated trends in the individual screens did, however, show spatial variation with respect to the adequacy by which the model mimicked the observed trends. The latter test does thus provide information on how the uncertainty in the modelled trends varies spatially.

Basic information on the adequacy of the current monitoring network for the study area (Sjælland) was achieved by combing the spatial location of the monitoring screens with the three dimensional extent of the groundwater bodies and simulated impact by climatic changes and current extraction. This simple analyses immediately revealed: 1) monitoring screens exists in only 14 of 69 groundwater bodies, 2) one-third of the screens where located in areas where the simulated impact by both climatic change and current extraction was low and 3) in four groundwater bodies, all monitoring screens where located in areas where the simulated effect by both climatic change and extraction was low.

An example of a possible grouping of the groundwater bodies is suggested. The impact by extraction is computed as the difference in heads for two simulations representing a virgin situation (no extraction) and the present level of extraction, respectively. A simple average drawdown is computed for areas where the impact by extraction is above a threshold (> 1 m) for the individual groundwater bodies and used as the offset for grouping. The grouping is thus based on the expected (simulated) impact that the present extraction has on the groundwater level in the groundwater bodies. The method thus takes into account both the actual pressure (extraction level) and the hydrogeological characteristics within the individual groundwater bodies.

Analysis of the adequacy of the monitoring network and possible re-design requires the establishment of a design criterion. Such criterion has not been developed so far, but should be defined within the national monitoring programme. Based on the example presented for Sjælland, the following overall strategy is recommended:

1. Grouping of groundwater bodies based on the overall impact caused by the present extraction

2. Subdividing of the groundwater bodies in predefined categories of impact (in the present study three categories were used: low, medium and high)
3. Within each group of groundwater bodies the individual categories are analysed across all groundwater bodies. Optimum locations of monitoring points are locations that provide the most information on the variation in the specific category for a group of groundwater bodies.

Several basic aspects must additionally be considered for identification of the optimal monitoring location within a group of groundwater bodies, which includes: 1) evaluation on whether a sampling point will display local features or can be extrapolated to other groundwater bodies within a group, 2) model uncertainty where less priority can be given to areas where the DK-model captures the observed trends satisfactorily, 3) the spatial coverage of monitoring screens, also screens not included in NOVANA and 4) evaluation on whether existing screens can be adopted.

The use of the DK-model was illustrated for Sjælland. Definition of the groundwater bodies, their spatial extent and link between groundwater bodies and monitoring screens in the national monitoring programme was adapted from the regional water authority. In the study some inconsistencies where found with respect to the numbering of the groundwater bodies and there classification of regional or deep. The groundwater bodies were originally delineated on the basis of the geological model included in the DK-model and the groundwater bodies were associated to the layer of the geological model. In the present study some inconsistencies were found between the spatial extent of the groundwater bodies and the geological model. Continuously development of the geological, without revising the extent of the groundwater bodies, is likely to explain some of the inconsistency, although it does not appear likely to explain all.

The request for a coherent and comprehensive monitoring programme by the WFD impose a need for detailed characterisation of the groundwater bodies and estimation of the expected impact from extraction, before grouping of the groundwater bodies can be accomplished. It is therefore crucial that there exists a reliable link between the groundwater bodies, screens in the national monitoring programme and layers in the national water resources model. For Sjælland the groundwater bodies were delineated on the basis of the national water resources model, and good correlation between the groundwater bodies and the model were generally found. An earlier version of the DK-model was similarly used for the initial delineation of groundwater bodies on Fyn, but this model has been subject to significant revision since, and the link between groundwater bodies, model layers and NOVANA screens are likely to need a revision. The DK-model was not used in the delineation of groundwater bodies in Jylland and the groundwater bodies have only been associated to model layers for parts of Jylland.

Based on the experiences from the present study based on data from Sjælland, it is concluded that the quality of the national water resource model is expected to be adequate for use as a supporting tool in the spatial design of the national monitoring programme and the grouping of groundwater bodies. The study illustrated that useful information can be obtained by combined information on spatial location of monitoring screens and groundwater bodies combined with simulated impact by climatic change and extraction. The same information was used to illustrate a simple approach for grouping of the groundwater bodies, based on the expected impact. Although consistency between groundwater bodies, monitoring wells and hydrogeological layers in the DK-model was generally found for Sjælland, some inconsistency was found, which may similarly be expected for the remaining part of the country.

The following steps are recommended prior to a detailed analysis of the monitoring network at the national scale:

1. Definition of design criterion, including a specification on how this criterion can be used to group the groundwater bodies.
2. All screens in NOVANA must be associated to a specific groundwater body and a model layer in the DK-model.
3. Analysis of historical time series data for computation of observed trends that can be compared to simulated values to test the DK-models ability to simulated trends. Essential is that the observed data do not include outliers that will result in an erroneous estimate of a trend. For this purpose good experience was obtained in the present study by using the Cook"s D measure. Analysis of times series will provide a general quality check to data and thus be beneficial to other studies.

References

Bijl, L. van der, Boutrup, S. & Jensen, P.N. (red.) (2007) NOVANA. Det nationale program for overvågning af vandmiljøet og naturen. Programbeskrivelse 2007-09 – del 2. Danmarks Miljøundersøgelser, Aarhus Universitet. 120 s. – Faglig rapport fra DMU nr. 615. http://www.dmu.dk/Pub/FR615.pdf.

Christensen BSB & Sonnenborg TO (2006). Grundvandsstandens udvikling på Sjælland 1989-2001, Projekt udført for Vandplan Sjælland, GEUS rapport 2006/59

Christensen, J.H., O.B. Christensen, P. Lopez, E. van Meijgaard, and M. Botzet. (1996) The HIRHAM4 regional atmospheric climate model. SC Rep. 96-104. Danish Meteorological Inst., Copenhagen.

Christensen, O.B., J.H. Christensen, B. Machenhauer, and M. Botzet. (1998) Very high-resolution regional climate simulations over Scandinavia: Present climate. J. Clim. 11(12):3204–3229.

Danish EPA (2006) Revision af udpegningen af grundvandsforekomster i Danmark. arbejdsrapport fra Miljøstyrelsen Nr. 6 2007.

EC (2000) Directive 2000/60/EC of the European Parliament and of the Council of October 23 2000 establishing a framework for community action in the field of water policy. Official Journal of the European Communities, L327/1-L327/72, 22.12.2000.

Gangopadhyay S, Gupta AD and Nachabe MH (2001) Evaluation of ground water monitoring network by principal component analysis. Ground Water, 39 (2), 181-191

Henriksen HJ, Troldborg L, Højberg AL and Refsgaard JC (2008) Assessment of exploitable groundwater resources of Denmark by use of ensemble resource indicators and a numerical groundwater–surface water model. Journal of Hydrology, 348: 224 – 240. DOI: 10.1016/j.jhydrol.2007.09.056

Højberg AL, Refsgaard JC, van Geer F, Jørgensen LF and Zsuffa I (2007) Use of Models to Support the Monitoring Requirements in the Water Framework Directive, Water Resources Management. http://dx.doi.org/10.1007/s11269-006-9119-y

IPCC. (2000) Special report on emissions scenarios (SRES): A special report of working group III of the intergovernmental panel on climate change. Cambridge Univ. Press, Cambridge, UK.

Jørgensen LF, Refsgaard JC and Højberg AL (2007) The inadequacy of monitoring without modelling support, J. Environ. Monit., 9(9), 931-942

Jørgensen LF & Stockmarr J (2009) Groundwater monitoring in Denmark: characteristics, perspectives and comparison with other countries. Hydrogeology Journal, 17: 827–842. DOI 10.1007/s10040-008-0398-7

Sonnenborg TO, Christensen BSB, van Roosmalen L and Henriksen HJ (2006) Klimaændringers betydning for vandkredsløbet i Danmark. GEUS rapport 2006/22, København.

van Roosmalen L, Christensen BSB and Sonnenborg TO (2007) Regional Differences in Climate Change Impacts on Groundwater and Stream Discharge in Denmark. Vadose Zone Journal, 6 (3): 554 – 571. DOI 10.2136/vzj2006.0093

Winter TC, Mallory SE, Allen TR and Rosenberry DO (2000) The use of principal component analysis for interpreting ground water hydrographs. Ground Water, 38 (2), 234-246