
Norsk
Regnesentral
NORWEGIAN COMPUTING CENTER
(sfi)² Statistics for Innovation
Final Report
The Research Council of Norway
sfi
Centre for Research-based Innovation

Content

Foreword by Centre Director

(sfi)² — Statistics for Innovation is one of the Norwegian centres of excellence for research-based innovation, called Sentre for forskningsdrevet innovasjon (SFI) in Norwegian.

Professor Arnoldo Frigessi.
Photo: Lin Stensrud (NR).

These centres have been created by the Research Council of Norway with the aim to enhance the capability of the Norwegian economy to innovate by focusing on long-term research and forging alliances between research-intensive enterprises and prominent research groups.

The acronym (sfi)² comes from the identical abbreviations of Statistics for Innovation and Sentre for forskningsdrevet innovasjon, both as sfi, with the square power indicating the non-linear return of the investment!

Statistics for Innovation developed statistical methodologies, strategically necessary to achieve innovation goals in four key sectors of the Norwegian economy: petroleum, finance, marine and health.

(sfi)² operated from 2007 until 2014 as part of the Norsk Regnesentral (Norwegian Computing Center, NR) in Oslo, in partnership with the University of Oslo (UiO), the Norwegian University of Science and Technology (NTNU) and eleven partners: Biomolex, DNB, Gjensidige, Norsk Hydro, Institute of Marine Research, PubGene, Oslo University Hospital, Sencel Bioinformatics, Smerud Medical Research, Spermatech and Statoil. (sfi)² had an annual budget of about 40 million NOK (4.5 million Euro), including in-kind contributions. More than 100 scientists and professionals from universities, research institutes and industry have participated.

We hope that this booklet, containing the final report of (sfi)², will give the reader a good overview of our achievements — research results which have led to innovation and added value at our partners and beyond. We are proud of our eight (sfi)² years!

Arnoldo Frigessi
Director of (sfi)²

Foreword by Managing Director NR

The establishment of Norwegian Centres of Excellence has changed the research landscape in Norway. It started with Centres of Excellence in 2002, Centres for Research-based Innovation from 2007 and then Centres for Environment-friendly Energy Research in 2009. These are large, visible and long-term investments. The centres increase the cooperation between institutions in long-term relations. The competition is hard. For a research group it is necessary to carefully consider how to participate in a winning team. It is necessary to find good research challenges, strategic partners and alliances and improve its own competence.

Dr. Philos. Lars Holden.
Photo: Lin Stensrud (NR).

Statistics for Innovation, (sfi)², was one of the first 13 centres for research based innovation with application areas finance/ insurance, petroleum, marine and health.

This shows that statistics is an important discipline that is crucial for this wide range of application areas. This also shows that the participating partners together hold a high scientific level in statistics.

NR had a close co-operation with most of the partners in the centre before it started. But (sfi)² has made it possible to reach much closer cooperation between all the partners than we had earlier. In particular has the cooperation between NR and UiO been closer with seminars, seminar lunches and NR has been responsible for more PhD's and postdoc's than previously.

NR has also a closer cooperation with all the business partners than earlier. We have had more strategic discussions identifying important research topics than earlier. The long-term commitment has made it possible to continue the cooperation also during the financial crisis in 2009-2010 without reducing the activity.

(sfi)² has improved the legacy of statistics in Norway and in particular NR. It has been important in recruiting new staff and increased the motivation in the present staff. It has also made it easier to fund project outside the centre.

(sfi)² has given many different types of results. We have had some concrete scientific results that have been important for our partners. Some have already proven beneficial for our partners and others will become beneficial in the coming years. We have educated many PhD's that will prove their value and contribution to the Norwegian society in the coming years. For NR and all the other partners in the centre, (sfi)² has increased our activity.

Lars Holden
Managing director, Norwegian Computing Center

Summary

Science, industry and business possess the technology to collect, store, distribute and process huge amounts of data; valuable data, yet complex and incomplete. Statistics provides the methods for collecting, understanding and learning from data. Because knowledge is fragmented and data are incomplete, discoveries and decisions are uncertain. Statistics is the science to quantify how confident we are in our knowledge, to assess risk and to make decisions under uncertainty.

Statistics for Innovation developed statistical methodology which provided solutions to industrial innovation problems in all four sectors; petroleum, finance including insurance, marine and health. These four sectors share a common need for new statistical methods: (i) although much data from different sources are collected, it is often hard to detect the critical and informative portions; (ii) data are extremely complex, as we measure more in depth; (iii) the systems to be understood or predicted are per se more complex; (iv) "the easy part" of the analysis (in terms of objectives, data and methods) has already been done, and it is necessary to discover "second order effects" which are often those that can give competitive advantages; (v) and finally, evidence-based decision making is required more and more in society at large.

The Norwegian oil industry is still the key engine in the Norwegian economy, providing wealth and welfare. However, increasing exploration costs require reduced uncertainty in future decision-making, which can be achieved by a better exploitation of the multiple sources of data and geophysical knowledge. Banks, insurance companies and financial institutions require a thorough understanding of the combined effect of risk factors.

The fisheries and marine industry play a crucial role in the economy of coastal districts. There is a need to develop more accurate management tools of marine resources, based on vastly incomplete data and a better representation of the complexity of ecology, biology and the industry's own sampling schemes and operations. Genomic data play a fundamental role for the understanding of cancer and other diseases and the development of new treatments. The enormous amount of high-throughput genomic data requires new ways to protect discoveries against randomness.

These industrial challenges pointed to a series of methodological research lines: (1) inference based on multiple data sources, with varying resolutions, of different type and quality; (2) fully probabilistic inference, capturing and representing uncertainty; (3) efficient simulation approaches and algorithms for inference, scenario simulation and prediction.

Outstanding results

We have developed a rich portfolio of research projects, focusing on the innovative ideas of our business and industrial partners. We developed new statistical methodology (going under the technical names INLA, pair-copula constructions, weighted and monotone lasso, Bayesian network design, etc.), contributing to significant value creation.

With Gjensidige, we modelled car insurance claims by incorporating individual frailty effects, accounting for heterogeneity between policyholders, unlike the traditional approach used in the insurance industry. This is useful in policy pricing, marketing and portfolio optimisation. Similar models are also used for marketing strategies in order to take direct action towards customers that have a high churn probability, instead of un-stratified marketing.

With Norsk Hydro, we predicted electricity spot prices for various price areas in a very powerful way. We also forecast wind and solar production to improve spot price predictions both in the German and the Nordic electricity markets.

With Statoil, we created a tool to help decide the exploration drilling sequence in a prospect. This totally new approach allows to explore where one should collect data (drill) first to as rapidly as possible predict precisely the presence or not of hydrocarbons in the area. Also we have a fully operative tool for computing the probabilities for different lithology-fluid classes from seismic data and geological knowledge. PCube, as the software is called, is today a key tool in oil exploration and commercialised in various versions.

With the Oslo University Hospital (OUS) we developed a generic web-based system, providing statistical methodology and computing power to handle a variety of biological inquires on genomic datasets, useful as engine for purpose-tailored solutions, including therapy optimisation. Also with OUS, we patented a gene expression based molecular diagnostics test, to be used to aid in deciding whether to administer post-mastectomy radiotherapy in breast cancer. Although there are several prognostic tests on market today, this is the first test which can indicate response to radiotherapy after removal of breast in breast cancer patients.

With the Institute of Marine Research we estimated the numbers of fish caught in each age group, based on samples from boats and hauls within boat. Such estimates of catch-at-age are a critical input in most age structured stock assessment processes for commercial fish species throughout the world.

Dam - Røldal Suldal hydro power system. © Norsk Hydro.

The seismic vessel Ramform Atlas. Photo: Petroleum Geo-Services.

Photo: iStockphoto.com.

Photo: iStockphoto.com.

We developed methods for optimising sampling designs across various sampling programmes used for management advice on fish stocks. We created a tool for predicting the seasonally varying map of presence of juvenile fish in the Barents Sea. This is used to decide whether fishing areas should be temporarily closed for shrimp fishing in periods with too many juvenile fish within the area. Our method obtains much better estimates with the same sampling effort, or the same precision can be reached with much fewer samples.

In aquaculture, we developed models for the spread of parasites and viruses within marine fish farms. This allows investigating the effects of various management strategies and it can also be a useful tool for understanding transmission mechanisms and tracking transmission pathways.

With DNB, we developed a new approach, the pair-copula construction (PCC), which allows capturing stochastic dependence between assets, delivering precise quantiles of total risk, and corresponding needed reserves. PCC has been used to solve problems in finance and insurance, hydrology, internet data, genetics and environmental science, among others.

Also with DNB we have implemented a new method to render correlation matrices usable in risk assessment: If correlations are computed for pairs of financial assets, or when some of the correlations are manually adjusted (as in stress testing), the resulting correlation matrix might be illegal. We have developed a tool for finding the nearest legal correlation matrix in such cases.

With Pubgene, we created advanced but still easy to use text mining tools that allow users to find, organise and better make use of relevant information in large biomedical text repositories.

With Gjensidige we produced maps of Norway which predict the geographical landscape of risk of damages to insured buildings due to weather conditions under climate change scenarios. Our specialised calibration technique allows producing reliable downscaling of precipitation over the Norwegian mainland.

With Smerud we developed a method for joint analysis of clinical trials examining the effect of anti-TNF (tumor necrosis factor) drugs which are important for many autoimmune diseases.

Furthermore we have promoted a new research culture, at the interface between statistical science and long-term innovation. Through training of many students and young industrial researchers, we contributed to capacity building: a new generation of statisticians, ready to take key professional and academic positions.

Sammendrag

Vitenskap, næringsliv og forretningsvirksomhet innehar teknologi for å samle, lagre, fordele og prosessere store datamengder; verdifulle data, men allikevel komplekse og ufullstendige. Statistikk bidrar med metoder til å samle, forstå og lære fra data. Siden kunnskap er fragmentert og data er ufullstendige, gjør det oppdagelser og beslutninger usikre. Statistikk er vitenskapen for å kvantifisere hvor sikre vi er på vår kunnskap, for å måle risiko og ta beslutninger basert på usikkerhet.

"Statistics for Innovation" har utviklet statistiske metoder som gir løsninger på industrielle innovasjonsproblemer innen alle fire områder; petroleum, finans inkludert forsikring, marin og helse.

Disse fire områdene deler et felles behov for nye statistiske metoder: (i) selv om det samles inn mye data fra forskjellige kilder, er det ofte vanskelig å identifisere de kritiske og informative utvalgene; (ii) data er ekstremt komplekse når vi utfører grundigere målinger; (iii) systemene vi skal beskrive eller predikere er i seg selv mer komplekse; (iv) "den enkle delen" av analysen (ut fra mål, data og metoder) har allerede vært gjort, og det er nødvendig å oppdage "andre-ordens effekter" som ofte er de som kan gi konkurransefortrinn; (v) og til slutt, bevis-basert beslutningstaking kreves oftere i samfunnet generelt.

Den norske oljeindustrien er fremdeles hovedmotoren i norsk økonomi og sørger for velstand og velferd. Men økte letekostnader krever redusert usikkerhet i fremtidig beslutningstaking, som kan oppnås ved bedre utnyttelse av flere datakilder og geofysisk kunnskap.

Banker, forsikringsselskaper og finansinstitusjoner krever en grundig forståelse av kombinerte effekter av risikofaktorer. Fiskeriene og den marine industrien spiller en avgjørende rolle i økonomien i kystdistriktene. Det er et behov for å utvikle mer nøyaktige styringsverktøy for marine ressurser, basert på enorme mengder ufullstendige data og en bedre representasjon av kompleksiteten i økologi, biologi og industriens egne samplingsstrategier og operasjoner.

Genomiske data spiller en grunnleggende rolle for forståelsen av kreft og andre sykdommer og utviklingen av nye behandlinger.

Den enorme mengden gjennomstrømning av genomiske data krever nye måter å beskytte oppdagelser mot tilfeldigheter.

Disse utfordringene i næringslivet pekte på en rekke metodiske forskningsområder: (1) inferens, basert på mange datakilder, med varierende oppløsning eller forskjellig type og kvalitet; (2) full stokastisk inferens, som fanger opp og representerer usikkerhet; (3) effektive simuleringsmetoder og algoritmer for inferens, scenario-simulering og prediksjon.

Fremragende resultater

Vi har utviklet en stor portefølje av forskningsprosjekter som fokuserer på innovasjonsideene til våre næringslivspartnere. Vi har utviklet nye statistiske metoder (som går under de tekniske betegnelsene INLA, pair-copula constructions, weighted and monotone lasso, Bayesian network design osv.), som har bidratt til betydelig verdiskapning.

Sammen med Gjensidige har vi modellert bilforsikringskrav ved å inkludere variabler som tar høyde for heterogenitet mellom forsikringstakere, i motsetning til den tradisjonelle metoden som benyttes i forsikringsbransjen. Dette er nyttig i prising av forsikringspoliser, markedsføring og porteføljeoptimalisering. Lignende modeller brukes også i markedsføringsstrategier for å ta direkte kontakt med kunder som har stor sannsynlighet for å si opp polisen i stedet for å bruke ikke-stratifisert markedsføring.

Sammen med Norsk Hydro har vi predikert spotpriser for strøm i forskjellige prisområder. Vi har også laget prognoser for vind- og solproduksjon for å forbedre spotpris prediksjonene både i det tyske og det nordiske elektrisitetsmarkedet.

Sammen med Statoil har vi utviklet et verktøy som kan bidra til å bestemme hvilken sekvens man skal prøvebore i et felt. Denne helt nye metoden utforsker hvor man først bør samle data (bore) for raskest mulig å kunne avgjøre om det er hydrokarboner i området eller ikke.

Vi har også utviklet et verktøy til å beregne sannsynlighetene for forskjellige litologi-fluid klasser fra seismiske data og geologisk kunnskap, som er fullt operasjonelt. PCube, som verktøyet kalles, er i dag et nøkkelverktøy i oljeleting og det er kommersialisert i forskjellige versjoner.

Sammen med Oslo Universitetssykehus (OUS) har vi utviklet et generisk web-basert system som tilbyr statistiske metoder og regnekraft for å behandle en rekke biologiske problemstillinger om genomiske datasett. Dette er nyttig som motor i spesialbygde løsninger, inkludert optimalisering av terapi.

Sammen med OUS har vi også patentert en molekylær diagnosetest basert på genuttrykk, som kan benyttes som hjelp i beslutningen om anvendelse av postmastectomy strålebehandling ved brystkreft. Selv om det er mange prognosetester på markedet i dag er dette den første testen som kan indikere respons på strålebehandling etter brystfjerning for brystkreftpasienter.

Sammen med Havforskningsinstituttet har vi estimert antall fisk i hver aldersgruppe ved fangst basert på prøver fra båter og hal innen båter. Slike estimater av fangst-ved-alder er kritisk input i de fleste aldersbaserte bestandsmodeller for kommersielle fiskearter verden over.

Vi har utviklet metoder for å optimalisere samplingsdesign over forskjellige samplingsprogram som benyttes for råd om forvaltning av fiskearter.

Vi har også laget et verktøy for å predikere sesongvarierende kart over forekomst av fiskeyngel i Barentshavet. Dette brukes

Photo: iStockphoto.com.

til å ta beslutninger om midlertidig stengning av rekefiske i perioder når det er for mange fiskeyngel i samme område. Våre metoder oppnår bedre estimater basert på samme antall prøver, eller vi kan oppnå samme presisjon med langt færre prøver.

Innen fiskeoppdrett har vi utviklet modeller for spredning av parasitter og virus i oppdrettsanlegg. Dermed kan vi undersøke effekten av forskjellige driftsstrategier, og det kan også være et nyttig verktøy for å forstå overføringsmekanismer og spore overføringsbaner.

Sammen med DNB har vi utviklet en ny metode, par-copula konstruksjon (PCC), som beholder de stokastiske avhengighetene mellom ulike risikotyper og resulterer i presise kvantiler for totalrisiko og korresponderende behov for økonomisk kapital. PCC har blitt brukt til å løse problemer blant annet innen finans og forsikring, hydrologi, internett data, genetikk og klimaforskning.

Sammen med DNB har vi også implementert en ny metode for å omforme korrelasjonsmatriser slik at de kan benyttes i risikovurdering: hvis korrelasjonene er beregnet for aktivapar, eller hvis enkelte korrelasjoner justeres manuelt (f.eks. ved stresstesting), kan de resulterende korrelasjonsmatrisene være ugyldige. Vi har utviklet et verktøy for å finne den nærmeste lovlige korrelasjonsmatrisen.

Sammen med PubGene har vi utviklet avanserte, men likevel brukervennlige, tekstmining verktøy som lar brukeren finne, organisere og utnytte relevant informasjon i store biomedisinske tekstlagre.

Sammen med Gjensidige har vi generert kart over fastlands-Norge som predikerer et geografisk landskap med skaderisiko for forsikrede bygninger på grunn av værtyper i ulike klimaforandringsscenarioer. Vår spesialiserte kalibreringsteknikk kan produsere pålitelige nedskalerte nedbørskart over Norge.

Sammen med Smerud har vi utviklet metoder for samlet analyse av kliniske forsøk som undersøker effekten av anti-TNF (tumor necrosis factor) legemidler som er viktige for mange autoimmune sykdommer.

Videre har vi fremmet en ny forskningskultur, på grensen mellom statistikk og innovasjon på lang sikt. Gjennom utdanning av mange studenter og unge forskere har vi bidratt til kapasitetsbygging: en ny generasjon av statistikere, klare til å inneha nøkkelstillinger innen akademia eller næringslivet.

Basic facts about the Centre

Partners in (sfi)²

Host institution and research partner: Norwegian Computing Center (NR), *http:www.nr.no*

Contact person: Lars Holden, Managing Director

Research partner: University of Oslo (UiO), *http:www.uio.no/english/*

Contact person: Arne Bang Huseby, Associate Professor and Head of Department of Mathematics

Research partner: Norwegian University of Science and Technology (NTNU), *http:www.ntnu.edu*

Contact person: Håvard Rue, Professor, Department of Mathematical Sciences

Corporate partner: Statoil, *http:www.statoil.com*

Contact person: Rolf Helland, Manager Seismic Imaging and Interpretation

Corporate partner: DNB, *http:www.dnb.no/en*

Contact person: Roar Hoff, Executive Vice President, Head of Group Risk Analysis

Corporate partner: Gjensidige, *http:www.gjensidige.com*

Contact person: Martin Danielsen, Executive Vice President Product/ underwriting general insurance

Corporate partner: Norsk Hydro ASA, *http:www.hydro.com*

Contact person: Stefan Erath, Head of Market Analysis

Partner in marine: Institute of Marine Research, *http:www.imr.no/en*

Contact person: Jon Helge Vølstad, Research Group Manager

Partner in health: Oslo University Hospital (OUS), *http:www.ous-research.no/*

Contact person: Eivind Hovig, Professor

Corporate partner: PubGene AS, *http:www.pubgene.com*

Contact person: Tor-Kristian Jenssen, Head of Research and Development

Corporate partner: Sencel Bioinformatics AS, *http:www.sencel.com*

Contact person: Torbjørn Rognes, Chief Executive Officer

Corporate partner: Biomolex AS, *http:www.biomolex.com*

Contact person: Håvard Hildeng Hauge, Managing Director

Corporate partner: Smerud Medical Research AS, *http:www.smerud.com*

Contact person: Knut Terje Smerud, Chief Executive Officer

Corporate partner: Spermatech AS, *http:www.spermatech.com*

Contact person: Eirik Næss-Ulseth, Chief Executive Officer

Organisation

Board in 2014 (personal deputy in parenthesis)

Roar Hoff, DNB, chairman
Martin Danielsen, Gjensidige, vice chairman
Stefan Erath, (Rønnaug Sægrov Mysterud), Norsk Hydro
Rolf Helland, Statoil
Jon Helge Vølstad, Institute of Marine Research
Knut Terje Smerud, Smerud Medical Research
Eirik Næss-Ulseth, Spermatech
Håvard Hauge, Biomolex
Torbjørn Rognes, Sencel Bioinformatics
Tor-Kristian Jenssen, PubGene
Eivind Hovig, Oslo University Hospital
Arne Bang Huseby, University of Oslo
Håvard Rue, NTNU
Lars Holden, (Petter Abrahamsen), NR
André Teigland, (Petter Abrahamsen), NR
Observer: Terje Strand, Research Council of Norway

All partners are represented in the Board.

(sfi)²'s board in 2014. From left: Lars Holden, Martin Danielsen, Rønnaug Sægrov Mysterud, Håvard Hauge, André Teigland, Knut Terje Smerud, Arne Bang Huseby, Roar Hoff, Arnoldo Frigessi, and Terje Strand.

Legal organisation

(sfi)² is a legal part of NR. The Centre was organised as a part of the SAMBA department at Norwegian Computing Center, which has André Teigland as research director. The legal and administrative responsible is managing director Lars Holden, Norwegian Computing Center.

Scientific Advisory Committee (SAC)

From left: Professor Henry Wynn, London School of Economics, UK, chairman. Professor Sylvia Richardson, Imperial College, London, UK. Professor Paul Embrechts, ETH Zürich, Switzerland.

We regret the death of Professor George Casella, University of Florida, USA, who was a member of our SAC until 2012.

Director of the Centre

Arnoldo Frigessi, University of Oslo.

Arnoldo Frigessi is professor at the Department of Biostatistics, Institute of Basic Medical Sciences, Faculty of Medicine. He also has a 20% affiliation at NR.

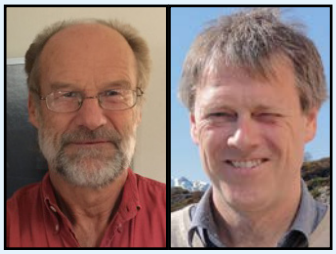
Key Innovators

From left: Kjersti Aas, NR, Finance ; Magne Aldrin, NR, Marine; Marit Holden, NR, Health ; Petter Abrahamsen, NR, Petroleum.

Key Scientists

From left: Ørnulf Borgan, Department of Mathematics, UiO; Nils Lid Hjort, Department of Mathematics, UiO; Ingrid K. Glad, Department of Mathematics, UiO: Geir O. Storvik, Department of Mathematics, UiO.

From left: Lars Holden, NR; Eivind Hovig, Oslo University Hospital; Knut Liestøl, Department of Informatics, UiO; Odd O. Aalen, Department of Biostatistics, UiO.



From left: Henning Omre, Department of Mathematical Sciences, NTNU; Håvard Rue, Department of Mathematical Sciences, NTNU

Effects of centre for the host institution and research partners

NR

(sfi)^2 has increased the activity in statistics at NR and among all the partners in the centre. It has strengthened the scientific level as a result of closer cooperation with universities, more time to science and closer cooperation with user partners that has made it possible to identify important scientific challenges that are important for the business partner. It has improved the recruitment to NR and to statistics in general and hence given significant results also outside the activity in the centre. About half of the PhD's and postdocs have been employed at NR in order to better understand the partners' problems and to have closer contact with the supervisors working on the same and similar problems.

NTNU

(sfi)^2 has contributed to increased research collaboration between NTNU, the Department of Mathematical Sciences, and the statistical society in Oslo. It has improved the network to some of the partners, especially Statoil. Participation to (sfi)^2 has strengthened the research competence, and through this competitiveness when applying and participating in new research projects. (sfi)^2 has resulted in three PhD-degrees. In particular it has contributed to the development of the R-INLA software for approximate Bayesian inference for latent Gaussian models, which would not have been developed without the support of (sfi)^2.

UiO

UiO took part in (sfi)^2 with the Department of Mathematics, the Department of Computer Science and the Department of Biostatistics, at the two faculties of Sciences and Medicine. The director of (sfi)^2 has been employed at UiO. Most of the PhD students affiliated to (sfi)^2 graduated at UiO. Our professors, associate professors, postdocs and students contributed significantly to the research projects during the eight years of the centre. This was one of the two centres of research based innovation with an UiO leader. Participation to (sfi)^2 has increased the contact between statistics at UiO and the industry, as also highlighted during 2013, which was the year UiO has devoted to innovation. Important research results have been published on leading scientific journals, and recruitment has benefit from the lively (sfi)^2 environment. The collaboration between UiO and NR has grown importantly, also leading to new joint ventures.

Oslo University Hospital

(sfi)^2 has contributed to increased activity in genomic research at the Oslo University Hospital (OUS), and has strengthened the hospital's research profile. New statistical methods have been developed, including methods for three-dimensional analysis of genome data. Moreover, analysis of integrative genomics at OUS has been improved. (sfi)^2 has also resulted in several PhD-degrees. In particular, (sfi)^2 has contributed substantially to the development of the Genomic HyperBrowser in the project GeneStat, which is a generic web-based system, providing statistical methodology and computational solutions to handle a variety of biological inquires on genomic datasets. It is a de facto flagship for Norwegian academic bioinformatics, and it is increasingly being used both nationally and internationally.

Photo: iStockphoto.com.

Effects of centre for the company partners, public partners and society at large

Review of what is considered the most important effects

(sfi)² has developed a rich portfolio of research projects, focusing on the innovative ideas of our business and industrial partners. Many of our projects have achieved important results that contribute to value creation. However, it is important to realise that it takes time for research results to become operational. Some of the results have been implemented by the partners and are in daily use, while other results still need more testing and evaluation by the partners. Even though (sfi)² is concluded, the collaborations and most projects continue, providing the opportunity for many of the innovative results to be implemented and used.

This section presents feedback from our partners. First, we discuss some general comments we have received from many of the partners. Then we present more specific feedback from each of the partners

How did we experience (sfi)²?

Feedback from the partners

(sfi)² has organised all projects in terms of innovation aims, and not method-wise, in order to maintain a constant focus on delivering solutions for the partners. An essential point for innovation is to follow the interests of the partners, in the sense that the projects should be rooted at the partners. Useful results for the partners increase the probability of their implementation.

Being part of (sfi)² has given many of the partners the opportunity to pursue interesting ideas that would not necessarily result in an immediate product or service, but have a larger long-term potential. Brainstorming sessions, which included researchers working both in academia and in contracted research as well as people working in the industry, were held before selecting innovation aims and resulted in a multitude of interesting ideas. Many of them could not be pursued in (sfi)², but hopefully some of them can be pursued in the future.

A major benefit of (sfi)² for the partners has been the access to competent personnel and research institutions. Some of the partners have had projects and cooperation with NR for many years before (sfi)². The centre has strengthened these relationships, expanded the network between NR and the partners, and simultaneously contributed to an increased knowledge base with the partners.

"(sfi)² has been able to keep key science personnel during the whole period. This continuity has strongly contributed to the success." – Board meeting

DNB

DNB has participated in the project TotalRisk that has developed methodology that can be used in the model for quantifying the total risk for the company. The results are among other used in DNBs capital requirement assessments, which involve quantification of extreme events or tail risk. A particular challenge is related to aggregation of different risk types while preserving the tail dependency. This is challenging theoretically as well as computationally when the available time (for simulations) is short due to operational causes.

Many of the theoretical and methodological results have been implemented and are operational. The change in the total risk model has influenced DNBs evaluation of capital requirements and risk profile. Financial capital is the basis for capital allocation; hence the innovative results have also contributed to a more precise risk-based credit pricing.

(sfi)² has provided solutions to problems that probably could not be solved internally. Also, through the process of making the results operational, DNB has increased their knowledge base within risk modelling. It is of great value that a network between DNB and the research environment in (sfi)² has been established. This is now used when discussing both smaller and larger problems related to the focus areas of (sfi)².

In addition, a DNB co-worker has been associated with (sfi)² fulltime for a year. Through this period he received valuable professional progress which is relevant in his current position.

Photo: iStockphoto.com.

"DNB finds it important that strong Norwegian research environments exist, with whom the bank can cooperate to better understand and quantify risk. (sfi)² has obviously contributed to this, which will give spillover effects for the entire Norwegian financial community." – Roar Hoff, DNB

Hydro

Fundamental power market models are used to analyse and forecast market developments. While such models are considered as the best approach in the long-term, short- and medium-term analysis demands statistical models to cover short-term variations as well as price range scenarios. Such scenarios are of vital importance in Hydro Energy's operational activity.

(sfi)^2 has given Hydro Energi AS a better basis to develop and implement price models in the short- and medium-term. The centre represented a wide range of methods, thematic areas and statisticians which led to a business advantage for solving challenging price modelling issues. This was a valuable supplementation to Hydro's work within the energy markets, as their strength is on the fundamental side.

Many of the subprojects within Elprice ended up in models used by Hydro Energi AS on a daily basis for portfolio optimisation and price modelling in the short- and medium-term. Hydro Energy's total annual revenue was more than 6.5 bill NOK in 2013. Improvement of pricing models is of crucial importance for the company.

Powerstation. Photo: Terje S. Knudsen, © Norsk Hydro.

"Through the work in (sfi)^2, we have seen the value of combining statistical expertise with eager market analysts; price models have developed, for the benefit of good water resource optimization."
– Rønnaug Sægrov Mysterud, Norsk Hydro

Gjensidige

Gjensidige has been involved in two of the main projects; ClimateInsure and CustomerLife.

Using climate scenarios and different climate models when predicting future damages, has contributed to an increased focus and consciousness-raising on these problems. The main result from (sfi)^2 has been a strengthened knowledge base in Gjensidige, in addition to focusing more on long-term thinking despite that insurance pricing is more short-term based. More specifically there has been an increased focus on

- surface water problems, particularly related to the water mains in the municipalities;
- building new houses in areas that are exposed to flooding, landslides or avalanches.

This area of research is currently not very applicable directly in products for Gjensidige in a short-term view. But in a long-term view they have increased knowledge of climate change.

The worst customers in an insurance company are causing a disproportional large amount of the claims. Thus, it is both just and profitable to develop good methods to locate this group. Some of the methods developed in (sfi)^2 has been transferred to Gjensidige and are currently in use. The potential added value in using such methods is large savings in this group (several million NOK). Using these types of methods will result in both increased premium income and reduced claim payments, and hence improved competitiveness for Gjensidige.

Other methods developed in the (sfi)^2 has not been implemented, but are tested with possibility to adjust or replace current methods. Gjensidige has received important contributions and good advices when developing new fraud solutions.

Photo: iStockphoto.com.

Statoil

Statoil is committed to accommodating the world's energy needs in a responsible manner, applying technology and creating innovative business solutions. This is addressed in an exploration programme with a strong focus on exploration activities using the latest techniques and technologies to ensure long-term production growth and value creation.

Within FindOil, (sfi)^2 has developed a new methodology and software (PCube) for use in oil exploration. The approach is patented, and the software is in regular use. The added value from using this tool is difficult to estimate, since it is a part of a larger framework. However, the total cost in oil exploration is very large, so small improvements imply very large savings. Other methodology developed in (sfi)^2 to help planning exploration campaigns, based on Bayesian network, has potential for improved decision making.

(sfi)^2 has strengthened the collaboration between Statoil, NR and NTNU. Statoil personnel involved in (sfi)^2 projects has increased their knowledge base in statistics and stochastic models. A large creative process, which has resulted in many new ideas, has been of great value. This will be further pursued by Statoil and NR.

Statoil is also in the process of trying to commercialise some of the methods we have been working on within (sfi)^2. This will be of benefit for oil exploration on the Norwegian Continental Shelf at large.

Photo: iStockphoto.com.

Institute of Marine Research

A major goal for Institute of Marine Research (IMR) is to provide research-based advice to Norwegian authorities on aquaculture and the ecosystems of the Barents Sea, the Norwegian Sea, the North Sea and the Norwegian coastal zone. The aim of research and management advice provided by IMR is to ensure that Norway's marine resources are managed and harvested in a sustainable way.

Estimates of catch composition and abundance indices of fish in numbers-at-age serve as critical inputs in many stock assessments. The project StatMarine has developed methods to obtain precise estimates and especially be able to quantify the associated uncertainty. Historically, the resulting estimates have usually been reported without any measure of uncertainty.

In particular, (sfi)^2 has supported major advances in the development of statistical methods for:

1. Estimating catch in numbers-at-age and associated uncertainty.
2. Estimating abundance indices in numbers-at-age from trawl and acoustic surveys with associated uncertainty.
3. Development of novel statistical assessment methods that can account for the propagation of errors input data from 1) and 2) to the stock assessments results that are used for quota advice.

The advances of methods developed in (sfi)^2 provide a framework for determining the adequate monitoring efforts of commercial fisheries and fish populations to support stock assessments, including cost-effective sampling and subsampling schemes for the estimation of age-composition of fish. Due to the very large costs of conducting sample surveys of fisheries and fish stocks, this will help IMR in the prioritisation of research to achieve major objective.

Photo: iStockphoto.com.

Smerud Medical Research

Smerud Medical Research has been involved in ComplexClin, and the results of that project have already attracted some interest from external parties, and from existing biotech clients. Smerud Medical Research has also been able to exploit the results from this project in their marketing and positioning towards other, new client opportunities.

In terms of potential added value, Smerud Medical Research believes that the data generated through ComplexClin can be further utilised commercially; partly by implementing similar methods in other drug development programmes, and partly through using ComplexClin as a starting brick in a huger set of examples which can help justify a change in regulatory thinking. However, this has a 10-years perspective, and requires a lot of supportive documentation and similar projects.

Another benefit from participation has been that Smerud Medical Research has had an opportunity to promote job vacancies to a group of very qualified statisticians; especially the graduates. One is hired through this route, and they hope to attract at least one more statistician from the (sfi)^2 society.

Most valuable, has been the general experience obtained from being a partner in a huger R&D consortium. Indirectly, the experience has been used to become an expert company in terms of raising public R&D grants into pan-European drug development consortia, and as such, Smerud Medical Research claims that (sfi)^2 enabled them to create, develop and make a commercial success out of the new processes and services. Through this work, they believe they have contributed significantly to attract EU-grants to the Norwegian clinical trial sector, beneficial both for smaller, commercial businesses, and also for the public hospitals and clinics that have an interest in drug development projects.

Lastly, Smerud Medical Research has on-going discussions with relevant partners in (sfi)^2, with whom they see opportunities to continue business collaboration to a mutual interest.

Photo: iStockphoto.com.

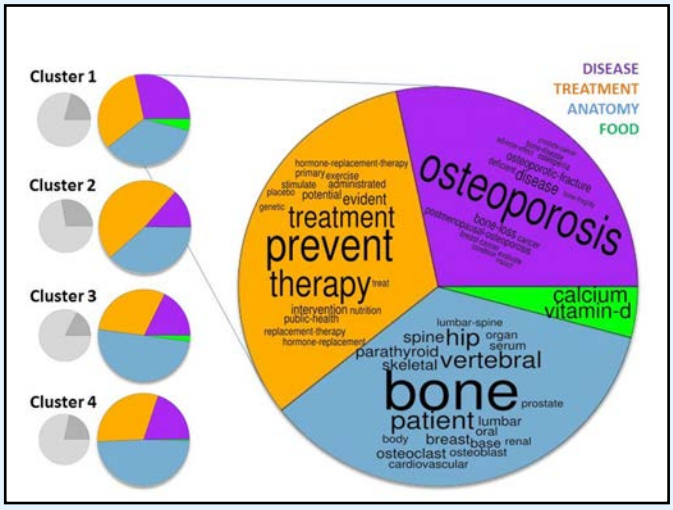
"From the general experience in being a partner in (sfi)^2, we have been encouraged to initiate similar consortia within the biotech sector itself, and mostly on a pan-European level." – Knut Terje Smerud, Smerud Medical Research

Biomolex, PubGene, Sencel Bioinformatics, Spermatech

Biomolex, PubGene, Sencel Bioinformatics and Spermatech have all been participating in the project BioInfStat. For Biomolex and PubGene new methods have been developed that are implemented and in use. These methods are mainly routines and additional modules that are parts of a larger product, which makes it difficult to quantify the added value. However, the participation in (sfi)^2 has resulted in important and necessary competitive advantages.

(sfi)^2 has strengthened the collaboration between these companies and NR. In addition there are on-going discussions between several partners and NR for future project collaboration.

Participation in (sfi)^2 has increased the companies' knowledge base in statistics and contributed to a better understanding of how statistical methods can be used in future product development.



Visualisation of multifocus labels for cluster 1 of the "bone fracture" search results.

Catch-at-age

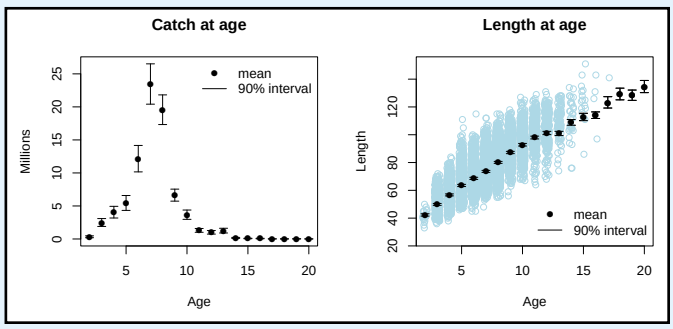
The Institute of Marine Research (IMR) needs to produce accurate estimates of the numbers of fish caught in the Norwegian commercial fish reservoirs, stratified by species and age group, on the bases of samples of caught fish from boats and hauls within boat. Such estimates of catch-at-age are a critical input in most age structured stock assessment processes for commercial fish species throughout the world.

The sampling design has also changed: previously, most samples were taken randomly from boats in ports, while more recently most samples were from a fixed "reference fleet", composed of a rather small number of boats. Bayesian hierarchical models are particularly appropriate when different type of data need to be jointly analysed, resorting to assumptions of conditional independence of data given the underlying truth. A Bayesian hierarchical model was designed for catch-at-age data of the type age and length, length-stratified ages, and length only, without using other expensive age determination techniques.

For example, by measuring the individual rings of the otoliths in the ear of a fish, it is possible to estimate its growth and age. Effects of age reading errors can be estimated. In addition, a model component allows the estimation of discards when data are available. Although this model component has not been used yet, IMR expects that the recent discard ban in the EU will put more pressure on Norway to document total catch including discard in the future. An important component of the model is the inclusion of random effects to account for positive correlation in both fish size and age within the sampling units. The model is complex and the estimation is based on Markov Chain Monte Carlo (MCMC).

A user-friendly implementation allows fishery researchers to use the algorithm on various species. The model is used to provide input data for the assessment for Northeast Arctic cod, haddock, and saithe, and North Sea Cod, haddock, and saithe, and is implemented and operationalised for North Sea Herring, Norwegian spring spawning herring, blue whiting, mackerel, and coastal cod. It is being implemented for Sebastes marinus, Sebastes mentella, and Greenland halibut, and the goal is to extend it to all species assessed by age structured models.

The model is described in several papers, where we also illustrate the advantages of our Bayesian approach in terms of appropriate modelling of precision and accuracy of the description of the uncertainty of the estimates.



Estimated catch-at-age and length-at-age with 90% confidence intervals for cod.

Fish farm. Photo: iStockphoto.com

Bycatch

We have developed a new way to predict the (seasonally varying) map of presence of juvenile fish in shrimp fishery basins of the Barents Sea, based on sample data. Again, a Bayesian approach was preferred, because it is based on a single probabilistic model jointly for all data, and the prior model allows borrowing strength between regions poor of data and regions data rich, in a coherent way.

The method outperforms current approaches, so that with the same sampling effort much more precise estimates can be obtained; alternatively, the same precision can be reached with fewer samples, reducing costs. The estimation of juvenile cod is used to decide whether fishing areas in the Barents Sea should be temporarily closed for shrimp fishing, to reduce involuntary by-catch which could spoil the population. We are working on a further improvement of the space-time Bayesian model, which we believe will turn very important for the Norwegian Directorate of Fisheries, where the responsibility of such management decisions reside.

Spread of infectious diseases

A further task has been to develop generic modelling tools for exploring effects of different intervention strategies on disease outbreaks in aquaculture industry. Bayesian spatio-temporal model has been developed for the spread of infectious salmon anaemia and other parasites and viruses in Norwegian salmon farming, based on traditional epidemiological and genetic data characterising the disease agent.

The purpose is to compare effects of various management strategies, to inform decisions, through the understanding of transmission mechanisms and pathways. Even if pathways for the spread of pathogens between aquaculture stocks depend on both pathogen and host characteristics, the context is very much the same from a modelling point of view, which allowed us to develop a generic framework. Farms are located at fixed coordinates, seaway distances between farms play an important role, networks representing commercial trade or local contact are relevant, and characteristics of the farms (water temperature, biomass, number of fish) are important.

Our models are used to investigate different intervention strategies to reduce the number of disease outbreaks, using simulations of disease spread under various scenarios. Interventions such as progressed culling of infected farm stocks, neighbourhood surveillance and culling, or vaccination, have been explored.

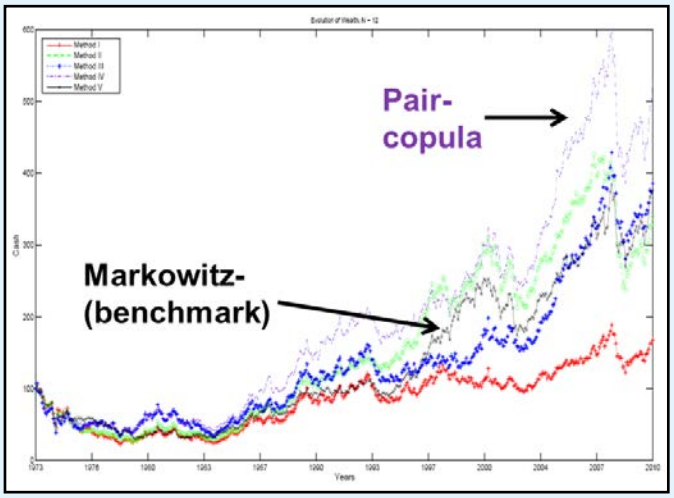
Risk management

Since the global financial crisis in 2008 it is more important than ever for financial institutions to quantify and evaluate their risk exposures. At the same time, the financial business has become more complex due to new financial instruments, global market information quickly available, and new risk structures, requiring advanced statistical tools. Further, pressure from shareholders and international changing regulatory environment, have made it more important for banks and insurance companies to understand and evaluate their risk exposures.

The Basel II international capital framework for banking has led to more use of statistical methods, and the Solvency II capital adequacy regime for insurance companies introduces a need for new methods to quantify risk over time, beyond traditional actuarial models. Risk management confronts us with heavy-tailed and non-Gaussian distributions, rapid changes and complex dependencies. This forces us to look beyond standard statistical models to develop a more sophisticated methodology.

Pair-copula constructions

With our partner DNB, we developed the new pair-copula construction (PCC) approach, which allows capturing stochastic dependence between assets, delivering more precise quantiles of total risk, and corresponding needed reserves. Since our first paper, presenting the inferential basis of PCC, PCCs have been extensively used in finance, econometrics, insurance, hydrology, forestry, pattern recognition, communication, neuroscience, medical imaging, genetics and environmental science, among others. The PCC is based on an expansion of any continuous multivariate density into a product of many pair-copula, each thus depending on just two variables. These variables are transformations of the original variables, in particular they are conditional distributions. The PCC model and the corresponding estimation algorithm are hierarchical in nature, so that the local choices of pair-copula can be freely made, while the transformations of the original variables which are necessary in the hierarchy are always available when needed, as they depend on copula chosen previously. Together with many other researchers worldwide, we contributed extensive theory and methods for PCC. The impact of PCC is still in fiery, with many industries testing it out, as we do with DNB currently.



The accumulation of wealth if 100 $ is invested in a portfolio consisting of 12 American industry indices in 1973 and then rebalanced every month using 5 different strategies. The strategy that is based on the pair-copula construction outperforms its competitors. Source: University of Queensland Business School.

Transformation of correlation matrices

With DNB we have implemented two new methods to render correlation matrices usable in risk assessment: if correlations are computed separately for pairs of financial assets (as happens when they are observed over different time spans or time scales), or when some of the correlations are manually adjusted (as in stress testing), the resulting correlation matrix might be illegal, and cannot be used further in simulations, prediction and analysis.

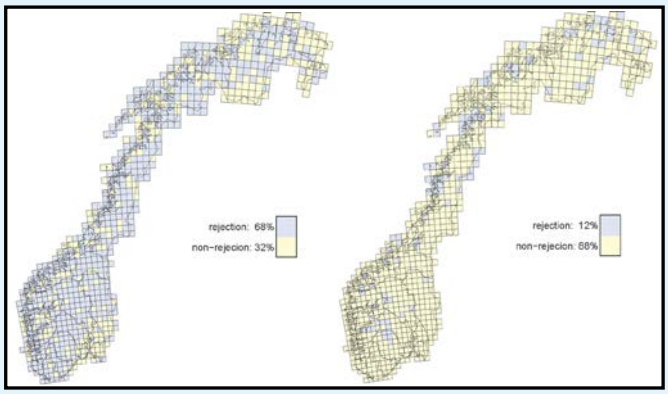
We have developed a tool for finding a legal correlation matrix which is near to the data in some way. We assume a pseudo-likelihood model, with Gaussian or beta components, of the empirical correlations and perform maximum pseudo-likelihood estimation. In the Bayesian version, we used certain natural priors. The method performs well compared to other shrinking approaches, and is now in use at DNB.

Climate change

The prediction of the impact of climate change on many economic sectors is urgent, because of the implications this has on industrial adaptation strategies and on mitigation policies. Equipping such predictions with realistic level of uncertainty is very important.

We studied the effect of climate change for the insurance industry in Norway. Gjensidige provided a unique data set, the first (and probably still the only one) public data containing all private buildings they insured in Norway over a time period of twenty years and all the claims filed in this period due to damages produced by weather events. Regressing these data with meteorological historical logs, allows learning the effect of weather on this type of claims. Using climate scenarios for the future, one can then predict future claim distributions.

There were several difficulties that needed to be overcome, related to the geographical scale of the analysis, the insufficient precision of regional models used for downscaling, and the complex interaction of the weather variables, precipitation, temperature, draining and snow. We developed a Bayesian Poisson Hurdle model with Ising spatially smoothed weather variable selection to model the historical relations between claims and weather. We analysed the quality of downscaling methods and prepared a new statistical calibration of dynamical downscaling. The production of risk maps for the insurance is of interest for future market developments and might also lead to new contract types.



Discrepancies between downscaled reanalysis precipitation and gridded observations (autumn). Spatially independent two-sample statistical tests reported as the fraction of grid cells rejected for uncalibrated (left panel) and calibrated (right panel) reanalysis out of sample test data.

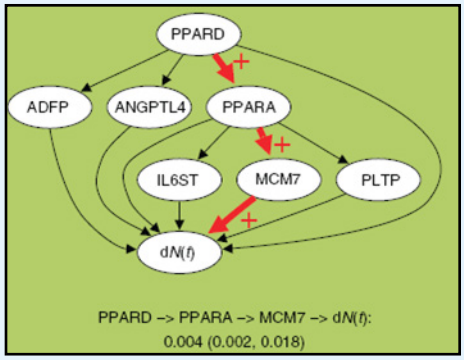
Breast cancer

Breast cancer is the most common cancer in females. Radiotherapy is known to prevent loco-regional recurrence (LRR) of cancer and to favour disease free survival in high-risk breast cancer patients. It is the standard treatment after breast conserving surgery. Postmastectomy radiotherapy (PMRT) is also recommended for patients estimated to have a high risk of LRR. PMRT is associated with significant side effects. There is an urgent need to determine strong biomarkers which help in taking better decision about PMRT, avoiding over-treatment.

It is now understood that breast cancer, like most other cancers, is a complex disease with many subtypes which can be characterised from a genetic point of view. Similarly, one can speculate that some of the about 25000 human genes, in particular their expression which describes their level of gene activity, play a role in the susceptibility to radiotherapy.

We identified seven genes associated with LRR in a group of high-risk breast cancer patients, which can be used to predict the usefulness of PMRT. Using a well described Danish cohort, we elaborated a two-step version of the lasso approach for Cox models, to find interactions between gene expressions and PMRT, which impacts time to local recurrence. The seven genes were further validated on an independent cohort. The score we derived from the expression levels of these seven genes has been patented and further studied.

We hope it can be developed further into a clinical diagnostic tool which can be used by doctors to personalise therapy. This project is quite typical for the so called $p>n$ setting, as there were about n=200 women in the cohort, randomised to PMRT or not, and p=25000 genes to select between



Dynamic path model, with direct and indirect effects, for breast cancer.

Genomic HyperBrowser

Sequence-level genomic data consists of pieces of information, each positioned on its specific locations along the DNA. For example, genes are entities of variable length, localised on the chromosomes. For any given position of the DNA, there exists multiple information regarding its physical structure, its function in nuclear organisation, its role in disease, and so on. Thus, any given nucleotide position of the DNA may be annotated with many, up to thousands, associations.

This creates an enormous matrix of information, where each of the 3.2 billion columns is a position of the DNA and each of the thousands of rows represents an annotation track. It is an enormous task to start analysing this giant matrix of tracks, though this is necessary to unravel underlying structure. Here we can find the answers to the organisation of biological systems.

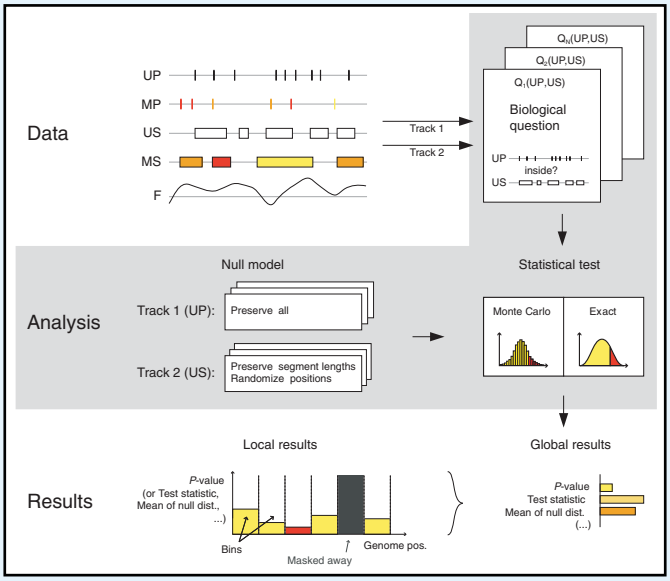
Do the tracks mimic each other? Where in the genome do they behave surprisingly incoherently? Where do strong associations appear? Which tracks fall out from a common pattern in specific areas of the genome?

We developed a generic methodology and implemented the Genomic HyperBrowser (GHB), to handle many types of biological inquiries in real time. The software is open source and widely extensible. The architecture, the software logic, the need for high performance computing and the implementation technology, all call for the development of new statistical methodology, based on pre-calculation of sufficient statistical summaries, progressive approximations which compromise between computational time and precision of results in a controlled manner.

The GHB is constructed on an abstract mathematical representation of the tracks. With the help of stochastic geometry on the line, we translate genomic elements into geometrical objects. Biological hypotheses of interest are then translated into relations between (marked or unmarked) points, segments and functions. Thus we created an abstract representation of genomic investigations, which allowed the construction of a generic tool.

The questions of interest are expressed in terms of stochastic properties of random processes, and inference is reduced to the computation of probabilities of rare events, describing the deviation from null-model behaviour. The GHB is up and running on the cluster of computers of the University of Oslo and can be accessed freely at http://hyperbrowser.uio.no/.

Interesting issues related to sequential Monte Carlo testing have been resolved. There are hundreds of registered GHB users worldwide who perform new analysis and hopefully discover new biology.



Flow diagram of the mathematics of genomic tracks.

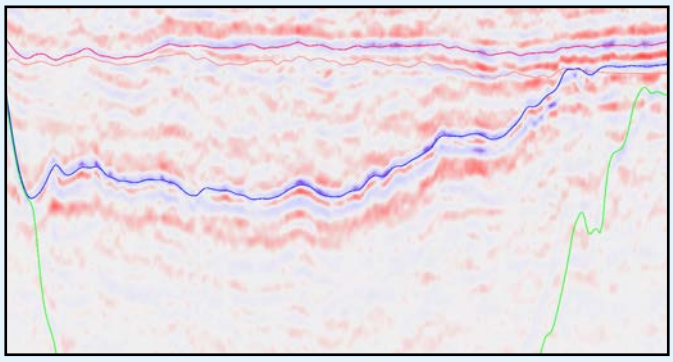
PCube

One of the problems considered with Statoil was lithology and fluid prediction, where the goal was to find the probability of oil and gas at a location given seismic data. Amplitude versus offset (AVO) is the standard, simple geometric procedure that describes how seismic signals are reflected and allows determining the fluid content, the porosity, the density etc., which all give indications in the presence of hydrocarbon.

AVO data was mainly used qualitatively, concluding positively or negatively on such presence with no further quantifications. This traditional approach was radically improved by our method, which can be considered as one of the first quantitative approaches using seismic data. We compute the probabilities for different lithology-fluid classes merging seismic data with geological knowledge. The method combines stochastic rock physics relations between the elastic parameters and the different lithology-fluid classes, with the results from a fast Bayesian seismic simultaneous inversion from seismic data to elastic parameters.

The methodological development provides a consistent framework for integrating knowledge generated by the different geo-disciplines. We incorporated simple rules for vertical stacking: due to the densities of fluids, there will never be water above the oil in a reservoir, whereas there may be gas, and so on. We developed and patented a method that introduces these constraints locally without exploding the computational time.

We further extended the model to additional type of data. Our method (called PCube) was implemented in C++ and the software tool received a graphical user friendly interface, so that the number of users at Statoil increased rapidly. Two commercial actors have also been involved since 2012. Fraunhofer, through Sharp Reflections, have been given the rights to create software based on the earlier version of PCube for open commercial sale, whereas Blueback Reservoir AS (now Schlumberger) developed a Petrel-plugin based on the more sophisticated version, and intended only for internal Statoil use.



Seismic horizons detected by PCube.

Where to drill?

When searching for hydrocarbons, obtaining a direct measurement in a location means drilling a well, which costs between 20 and 50 million USD. In addition there is a need for indirect measurements between wells, for example seismic geo-imaging of the rock structures. The two most common indirect data types are obtained by probing the earth with sound waves (seismic), and electromagnetic waves (CSEM). There is also considerable qualitative knowledge of the geological formations which can be formalised into spatial statistical prior distributions.

Our idea was to represent the possible drilling locations in an area (called prospect) as nodes of a directed acyclic graph (DAG). Each node has a discrete sample space, for example {dry, gas, oil}. Oil and gas require the presence of each of a source rock, a reservoir, a migration and a trap for the hydrocarbons, where each of these events has a distinct spatial dependency between neighbouring prospects. The number of nodes in a prospect ranges from 10 to 50.

As causal geological mechanisms drive the topology of the network, Statoil geologists and geo-physicists have worked on the elicitation of the graph topology and of the needed conditional probabilities in the DAG. This means that in this application and contrary to the previous ones, the networks are available, but there are no data yet acquired on the nodes.

How to design a sampling strategy which minimises the number of wells to be drilled in a prospect, while allowing the conclusion that the prospect does not have hydrocarbons with high probability? Which is the optimal sequence of few, say 5, wells to be drilled to maximise information?

Several papers have been written on the theme, ranging from methodology to algorithmically challenges and the actual application. After a re-analysis of historical prospects, we run a prospective case. Statoil will run the approach in parallel with their other operational procedures, in order to compare and integrate recommendations. A commercial version of the approach is also on its way while a PhD student is currently working on new theoretical aspects.

The seismic vessel Ramform Vanguard. Photo: Petroleum Geo-Services.

Research

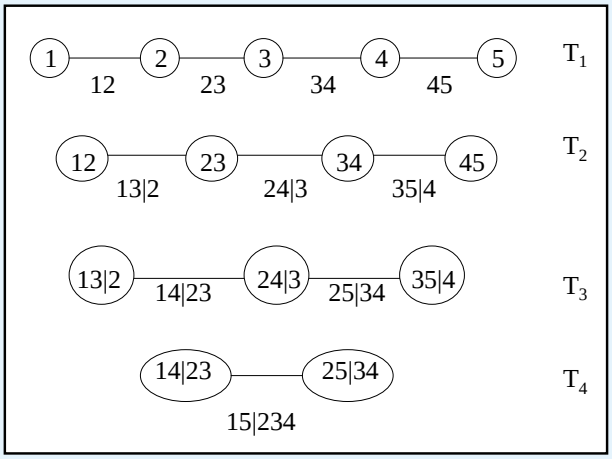
Research achievements

Methodology developed at (sfi)² was clearly motivated by our industrial challenges: during the eight years, we have developed new statistical methodology and published more than 220 papers, which are receiving attention from the scientific community. This section describes the key results achieved in each of the research projects

ComplexDepend – Statistics for complex stochastic dependence

The aim has been to produce scientific results which underpin the innovation strategy of (sfi)² by providing new statistical methods and computational tools to current and future innovation projects.

Key areas where we contributed original methods include multivariate copula models, experimental design in Bayesian networks, stochastic models for infectious disease epidemics with known contact networks, Bayesian computing with INLA, sequential Monte Carlo methods and testing, high dimensional regressions and regularisation, statistical causality for processes over time, statistical model downscaling, hierarchical Bayesian models, Bayesian and non-Bayesian methods for data integration, probabilistic forecasting and model comparison. In this global and highly competitive world, in order to produce outstanding innovation, to be best and first, one needs to build on its own competitive edges, among which unique methodological developments take a key role.



Pair-copula construction.

Elprice – Electricity price sensitivity

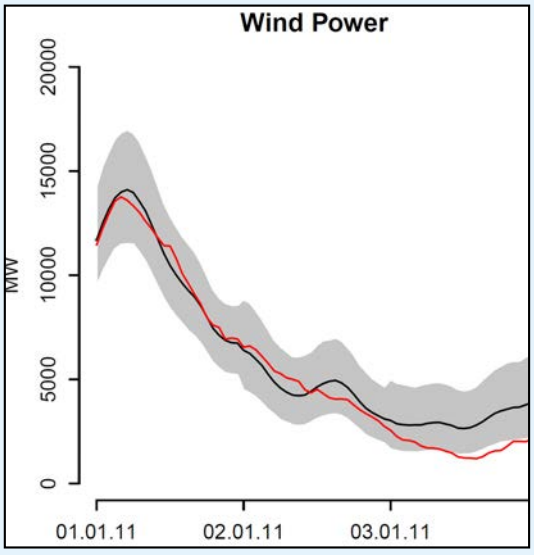
On the 9th of June 2014, solar power for the first time comprised more than half of the German power consumption. The impact of unregulated, renewable (solar and wind) production is increasing in the German and Nordic electricity markets, and the traded volumes are large (349 TWh of electricity was traded on the Nord Pool Spot's Elspot market in 2013). This is just one example of how the liberalised electricity markets are constantly evolving. That means that every substantial improvement in models will contribute to better risk or profit margins for the producers or consumers.

Elprice has developed new statistical methodologies for understanding, decomposing and forecasting electricity markets. The Nordic electricity price is influenced by local, regional and global drivers. Water reservoir levels is a local driver, the German coal price is a regional driver, while the oil price is a global driver. Elprice has developed models for a causal understanding of the interplay between such price drivers in the short and long run. In the very short run, like the next hours or days, forecast models are essential. To this end, Elprice has developed forecasting models for sun and wind power production, based on weather forecasts. The state-of-the-art for forecasting electricity prices is to forecast the demand and, conditional on this, forecast the price. This standard procedure partly ignores the manner in which spot prices are settled. Elprice has therefore built models for forecasting the key ingredients in the formation of the next day's price, like the flow of electricity between two markets and the way the producers and consumers of electricity bid in these markets.

With these new methods, combined with application of state-of-the-art methods, Elprice has given a better risk assessment and improved profit margins through

- improved models for modelling long-term risk in the electricity markets;
- integration of forecasting models for renewable production into electricity price forecasting models in a coherent manner;
- implementation of the first statistical methodology for forecasting of electricity prices where all key ingredients are included.

Power lines, Frøystul. © Norsk Hydro.



Predicted wind power production (black) with confidence interval (grey), and actual wind power production (red).

ClimateInsure – Climate change and insurance industry

The insurance industry is highly exposed to risks caused by weather related events. In a warming world there is evidence for changes in the spatial distribution, frequencies, and intensities of both ordinary and more catastrophic weather events. Turning weather pattern projections into robust and reliable loss estimates at a local scale is imperative for the successful modelling of impacts of climate change to the insurance sector.

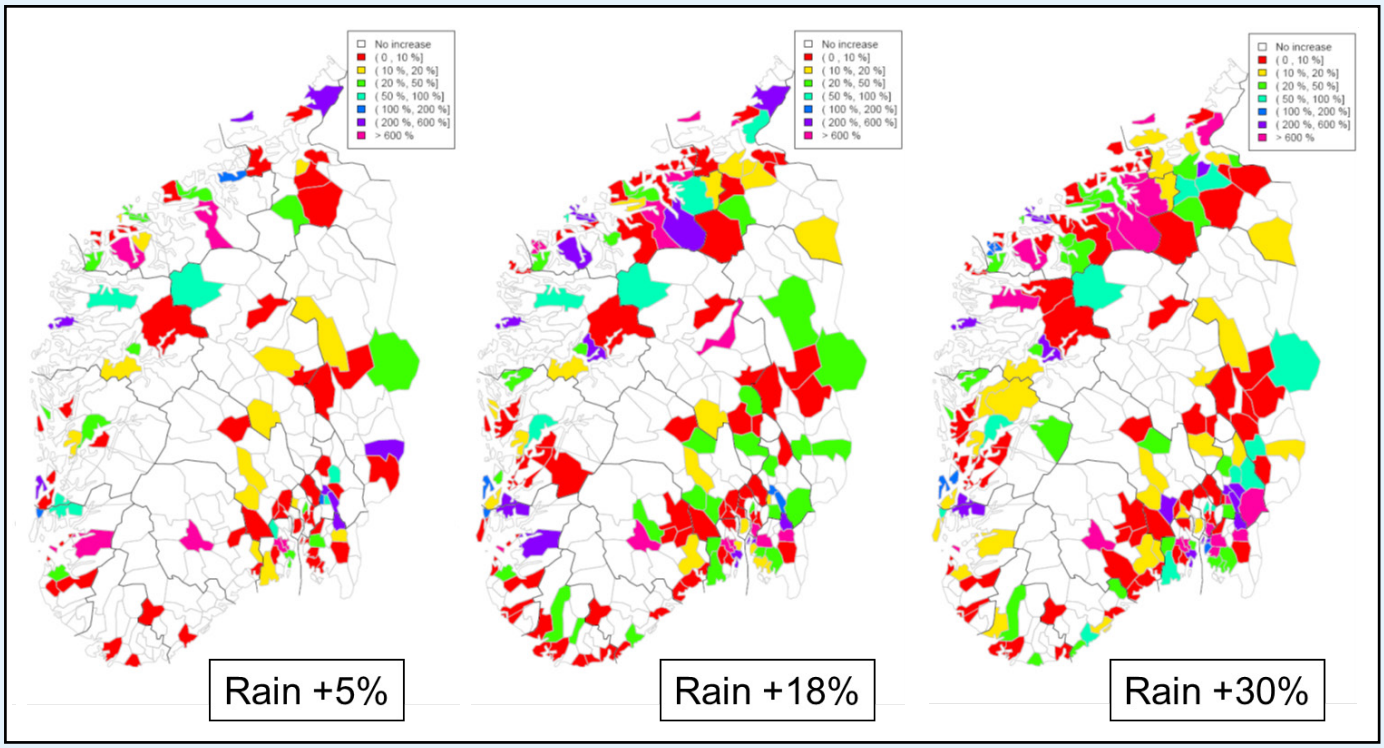
ClimateInsure has contributed to the comprehensive financial and safeguarding challenges of mitigation and adaptation. In order to understand patterns of risk in the light of climate change, over time and space, it is necessary to explore the relation between weather/climate events and incurred losses analysing historical data. A Bayesian Poisson hurdle statistical model that explains number of residential building water claims in terms of various meteorological and hydrological variables under a spatially smooth variable selection was established at municipality level.

The model depicts changes in claim frequency either from certain synthetic changes in reference weather patterns or from more sensible future weather projections obtained from regionally downscaled global climate model scenarios. ClimateInsure combined full distributional as well as local effect statistical tests for evaluating the latter and applied them to past ERA-40 reanalysis model precipitation, with no hope of climate models performing better than a reanalysis.

Inconsistencies in downscaled products in general leave demand for a full distributional calibration. ClimateInsure experimented with the Doksum Shift function for bias correcting a reanalysis. Further investigation is currently ongoing with the aim of extending the calibration to climate model data. With such a calibration in place, projections of future claim frequencies can be derived from feeding climate scenarios into our Bayesian Poisson hurdle model.

Based on the above methods, ClimateInsure

- identified same day precipitation intensity as the most important meteorological variable for the generation of water losses;
- estimated quite dramatic changes in the risks of precipitation-related damages to residential buildings as a result of plausible climate scenarios, with an increase in the mean number of losses of more than twice the percentage increase in precipitation even for very moderate scenarios;
- demonstrated that past ERA-40 reanalysis precipitation suffers from serious discrepancies as compared with observations;
- established a 25x25 km^2 grid based bias correction for ERA-40 reanalysis precipitation.



Percentage increase of the 95-percentile of predicted yearly number of water losses under a low (+ 5%), medium (+ 18%) and high (+ 30%) future precipitation scenario relative to a reference year (2001).

CustomerLife – Statistics for modelling customer life in an insurance company

The insurance business is based on having fundamental knowledge and behavioural understanding about the customers as a group, but also on the individual level. CustomerLife has developed statistical methodology to help Gjensidige (or any other non-life insurance company) to better understand their customers. By understanding how customers behave, the company can improve the pricing of their products, target their campaigns more efficiently and avoid intrusive interaction with customers. The main projects within CustomerLife have been churn prediction, individual motor insurance modelling and soft fraud detection.

In Norway, it has become quite easy for a customer to change insurance company if a competitor presents a better offer. Attracting new customers is often more expensive than actions to retain existing customers. Being able to identify customers with a high risk of leaving is therefore valuable. Instead of more or less random target groups, the company can target campaigns towards customers that are likely to leave. Each month, our statistical model predicts the probability for each single policyholder that he or she will leave the company the next month, i.e. that all his or her policies are cancelled. We have also modelled the time until a policyholder leaves the company for corporate motor insurance, where each policyholder, in this case a company, has several cars insured. Here we have used survival analysis with an individual frailty effect for each company, a methodology commonly used in medicine.

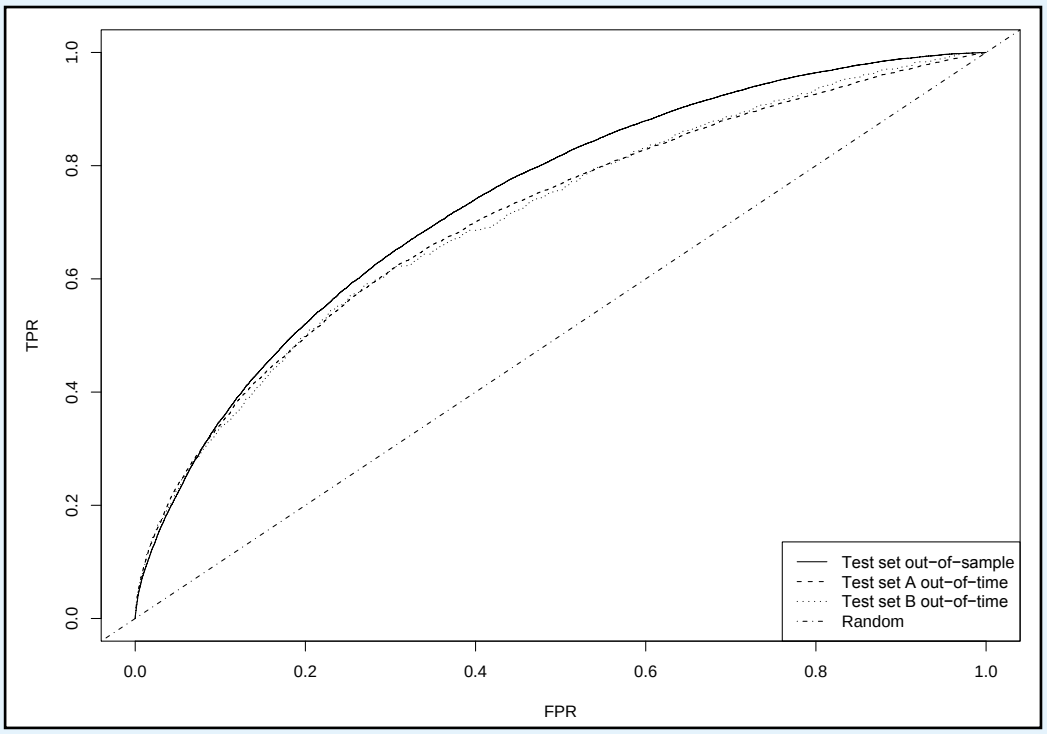
When pricing motor insurance products, the premium is set based on characteristics of the policyholders and their vehicles. Two policyholders that have the same age, live in the same city, drive the same car and the same distance pay in principle the same premium. The premium reflects the risk of accidents, and a higher premium means a higher accident risk. One policyholder may however, be a worse driver (or more unlucky) than the other driver and have more accidents than what his or her premium is supposed to cover. We have modelled this individual behaviour that cannot be explained by the given characteristics.

This makes it possible for the company to identify the policyholders that are worse or better drivers than previously accounted for and adjust their premiums accordingly. We modelled the number of claims for each policyholder using a mixed Poisson regression model where explanatory variables describing the policyholder and the vehicle were included, and used several years of data for each policyholder. For each policyholder, an individual effect was estimated. We followed a number of policyholders forward in time, and were able to predict which policyholders that would have more claims than the average policyholder in their group also in the future.

Insurance fraud leads to higher costs for the insurance company and also for the non-fraudulent policyholders who have to pay an increased premium. Soft fraud occurs when a valid claim is exaggerated or when a policyholder attempts to alter their policy after the event that lead to the claim had occurred. Investigating claims require resources from the company, and it is undesirable that innocent policyholders are treated as potentially fraudulent. The aim is therefore to reduce the number of claims that needs to be investigated by correctly identifying more claims as fraud. The method used is a combination of several statistical models, instead of using just one model which is more common. The result is, for every claim, a probability that the claim is a valid claim. Our method was shown to outperform the current method by having a smaller expected loss.

The project CustomerLife has

- developed new methodology that better identifies fraudulent claims;
- developed a model that identifies high risk policyholders in car insurance that can be used for better ranking and pricing of policyholders;
- identified important factors to identify policyholders that have a high risk of leaving the company.



Prediction performance of our model that identifies customers likely to leave the company, compared to random classification of policyholders. Our model is far better than random classification.

TotalRisk – Statistics for modelling the risk of financial institutions

Though extreme events in the financial world manifest themselves through high gains and especially high losses, very often it is the co-movement of underlying instruments in times of crises that triggers such events; the so–called perfect storm scenario. Hence, the modelling of the joint occurrence of extremes ought to be of prime concern when modelling financial risk.

The objective of TotalRisk has been to renew the tools used for risk modelling in financial institutions, producing more reliable and useful estimates of risk. Complex dependency and interactions are key aspects. Dependence structures are often non-linear, requiring new statistical methods, such as copula-based approaches.

Copulas have been used extensively for modelling multivariate financial return data. However, when different pairs of risk factors in a portfolio have very different tail dependence, the standard copula theory will not allow for the construction of an appropriate model. In TotalRisk we have developed a new method for constructing stochastic multivariate models from smaller components denoted pair-copula constructions (PCCs). The corresponding paper has already been cited 475 times.

To determine the ability of a financial institution to deal with an economic crisis, stress-testing is often used. One stress scenario is increased correlations between assets or risk types. Manually adjusting a correlation matrix often makes it statistically invalid. In TotalRisk we have developed a novel Bayesian method for rehabilitation of illegal correlation matrices.

In risk management applications one usually is interested in quantiles far out in the tail of the probability distributions. For such problems, importance sampling (IS) often provides an efficient means of generating estimates with high accuracy. In TotalRisk we have developed a new IS technique, specially designed for estimating the risk of real-world credit portfolios.

Based on the above-mentioned methods and on careful application of state-of-the-art methods, TotalRisk has developed the following innovations already operational at DNB:

- a tool for transforming a non-positive definite correlation matrix to a legal one;
- a tool for importance sampling of the tail of the credit risk distribution;
- parallelisation of simulations using the Open MP technology;
- the copula methodology for computing total economic capital.

	NS	IS	RE	HF	NGB	NB	IB	LB	USD	EUR	JPY	GBP	CI
NS	1	0	0.33	0	0.02	−0.05	−0.04	0.01	−0.02	0.02	0	0.01	0.01
IS	0.77	1	0.38	0	0.02	−0.05	−0.03	0.01	−0.02	0.02	0	0.01	0.01
RE	0.32	0.24	1	0.03	0.01	−0.06	0	0.01	−0.01	0.01	0	0.01	0.02
HF	0.06	0.06	−0.16	1	0	0.01	0.01	0	0.01	−0.01	0	0	0.03
NGB	−0.13	−0.02	0.02	0.16	1	−0.01	0	−0.01	−0.01	0.01	0	0	0.04
NB	0.2	0.2	0.25	0.3	0.68	1	0	0.01	0.01	0	0	0	−0.14
IB	0.2	0.2	−0.03	0.37	0.29	0.64	1	−0.01	0.01	−0.01	−0.01	0	−0.06
LB	−0.05	−0.04	−0.11	−0.28	−0.54	−0.85	−0.59	1	0.01	−0.01	0	0	0.03
USD	0.14	0.07	0	−0.26	−0.06	−0.16	−0.23	0.06	1	−0.01	0	0	−0.05
EUR	−0.08	−0.13	0.01	0.18	0.16	0.21	0.14	−0.2	0.27	1	0	0	0.05
JPY	−0.04	−0.03	0.04	0.08	0	0.04	0.1	−0.05	−0.54	−0.2	1	0	0.01
GBP	0	−0.02	−0.06	0.02	0.04	0.04	−0.01	0.02	0.46	0.47	−0.36	1	0.02
CI	−0.9	−0.9	−0.9	0	0	0	0	0	0	0	0	0	1

Correlation matrix; lower triangular is the illegal correlation matrix and upper triangular is the correction made to transform it to a legal one.

StatMarine – Statistics for management of Norwegian marine resources

Forming scientifically based management advice of marine resources involves applying time series of estimates based on data originating from several monitoring programmes to models that combine the data inputs and estimates, and predicts effects of harvesting on age/length structured stock abundance. Fundamental inputs to models are time series of estimates characterising the commercial fisheries, e.g. total removals and age- and size structure, and estimates of stock status independent of the fisheries, typically from research surveys. Monitoring programmes are prohibitively costly, and cost efficient methods for both sampling and estimation are therefore critical to ensure proper allocation of sampling effort and risk management of quota advice.

In this project the main focus has been on estimating removals by commercial fisheries, namely estimation of catch-at-age and bycatch of juvenile fish in shrimp fisheries.

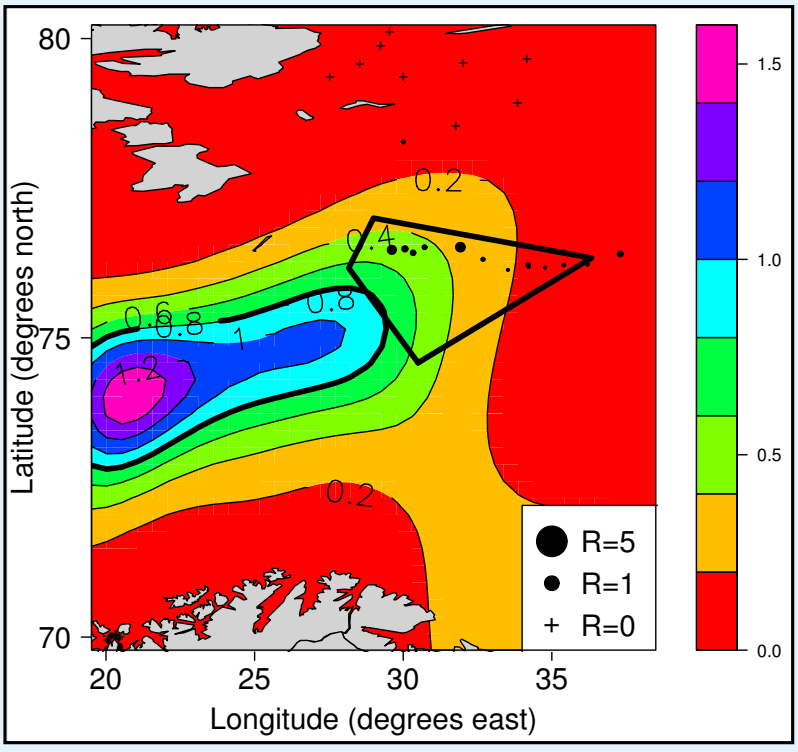
StatMarine has developed a framework model that includes a model that efficiently estimates catch-at-age applying state-of-the-art statistical methodology by combining data from multiple monitoring programmes and different types of sample data recognising the complex hierarchical cluster data. The approach is generic and is implemented as an R-library with a graphical user interface that controls in- and out-put of data.

It is currently implemented for 11 different commercially important fish species at the Institute of Marine Research and is currently being implemented for other species and is being considered used by other institutions for other species.

The model is complex and the estimation is based on MCMC and has been extensively evaluated through simulations and thorough comparisons with classical sampling theoretical approaches for sample survey data. Based on this framework we have developed methods that optimise the sampling strategy by appropriate spread of sampling effort measured through precision of selected parameters constrained by a given cost, or alternatively find the cost for a required level of precision.

We have also developed an innovative way to predict the (seasonally varying) map of presence of juvenile fish, based on sample data. The method outperforms current approaches, so that with the same sampling effort much more precise estimates can be obtained; alternatively, the same precision can be reached with fewer samples, reducing costs. This is used to decide whether fishing areas in the Barents Sea should be temporarily closed for shrimp fishing in periods with too many juvenile fish within the area.

The summarised effort on developing methods and implementation for estimation of input used in stock assessment models for quota advice has enabled to take quantification of risk and cost efficiency a step further: We are now in a position to develop a framework that provides realistic quantification of uncertainty in parameters used for management advice which is linked with uncertainty due to sampling. More specifically this means that uncertainty in quota advice can be controlled, and associated cost monitoring programmes may be established. This research has and will have huge interest and focus in the years to come, to achieve the goal of realistic risk quantification and control of management of marine resources.



Contour plots of the estimated bycatch ratio. Circle areas are proportional to the bycatch ratios. Samples with zero bycatch rates are marked with a " + " symbol.

Infect – Modelling spread of infectious diseases in fish farming

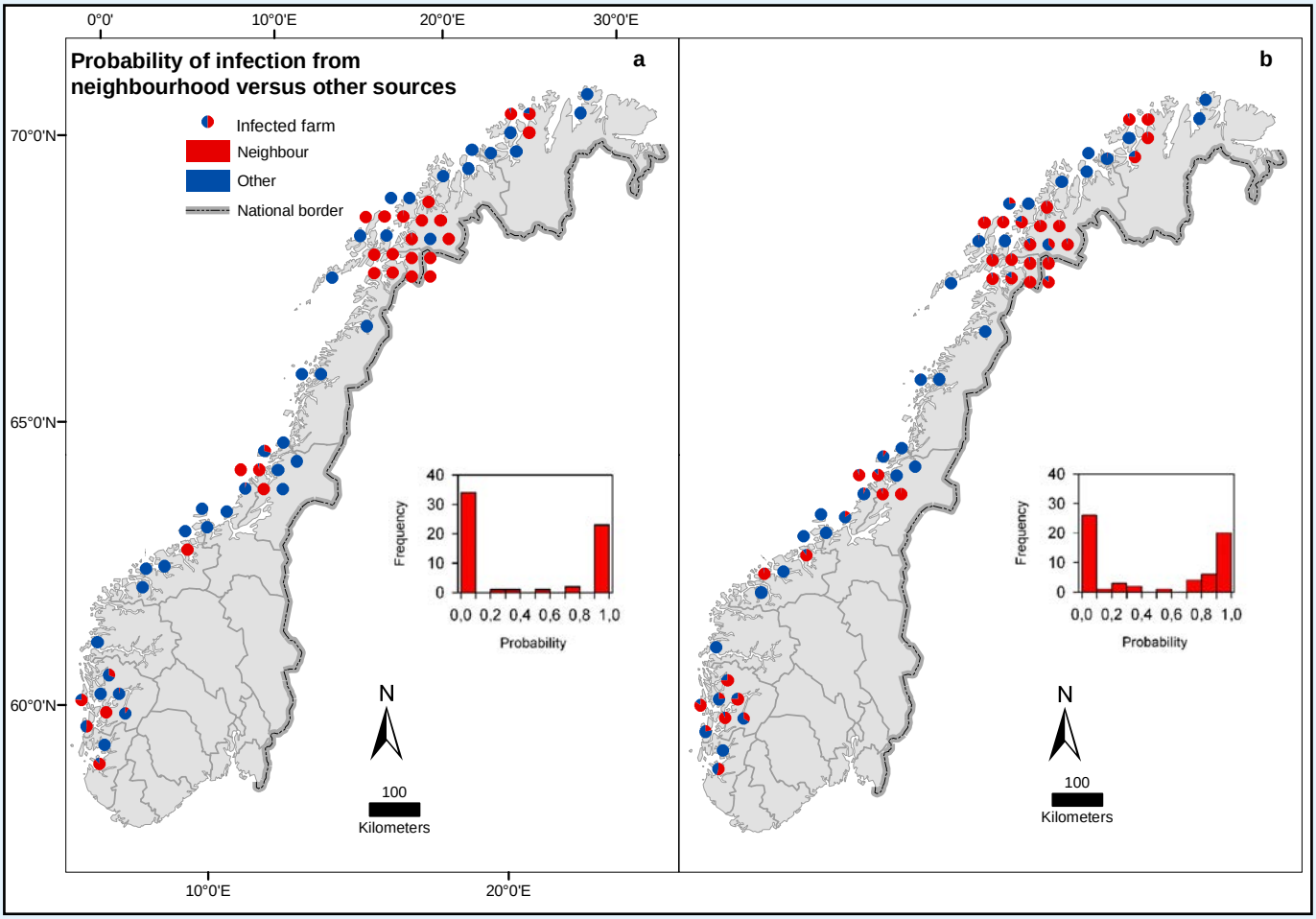
Infectious diseases, including parasites, constitute a constant threat to the Norwegian fish farming industry with major economic implications, in addition to being a problem for fish welfare and the environment.

Infect has developed stochastic models for the spread of various infectious diseases within and between marine fish farms, where all Norwegian salmon farms are modelled simultaneously. The set of salmon farms is regarded as a network where the degree of contact between each pair of farms depends on their seaway distance.

Separate models are developed for salmon lice (a parasite), and the two virus diseases pancreas disease (PD) and infectious salmon anaemia (ISA). The latter model takes into account genetic information on the ISA virus.

Based on these models, Infect has

- provided new insight into the dynamics of the various diseases, especially regarding the importance of water contact;
- provided general tools for investigating the effects of new disease-preventing management strategies;
- shown that introducing a culling strategy may reduce the number of PD outbreaks by 60%.



Probability of infection from neighbourhood versus other sources.

Infect – Human epidemiology

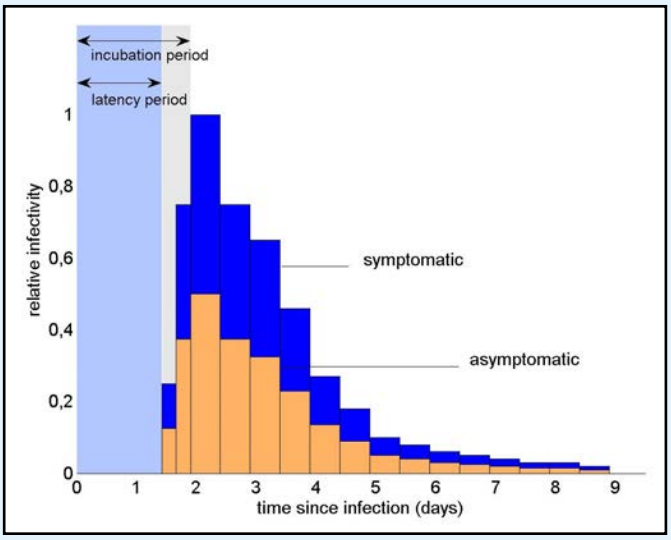
Influenza is an acute viral infection that spreads easily through droplets or by contaminated hands. WHO estimates that approximately 5-10% of the adult population and 20-30% of children get infected each year. Influenza has a significant economic impact associated with medical treatment (direct cost) and lost productivity (indirect cost).

The importance of presenteeism, i.e. employees attending work despite an illness, is increasingly being recognised as a cost driver: sick people have reduced work capacity, and for transmissible diseases like influenza, sick people increase the risk for co-workers to become infected.

In this project, we reviewed the literature on influenza transmission in the workplace and effects of sick leave behaviour on influenza transmission. We performed a survey among all 31 member countries in the European Influenza Surveillance Network (EISN) on national recommendations for influenza particularly targeting the work force and opinions on their effectiveness. We developed a dynamical model to explore the health consequences of reduced influenza presenteeism. To inform the model, we conducted a survey on current influenza sick leave practice in Norway. The model was combined with a probabilistic health-economic analysis to assist in design and decision-making of public health interventions in Norway.

Some key results from Infect are:

- there is a clear gap in the literature regarding knowledge about influenza transmission in the workplace and the effect on transmission of social distancing measures, e.g. sick leaves, targeting the work force;
- approximately half of the European countries in EISN have implemented recommendations for reducing influenza presenteeism in pandemic preparedness plans. There is a consensus opinion among European policy-makers that lowering presenteeism is effective to prevent spread of influenza;
- the survey among a convenience sample of Norwegian employees indicates that approximately 70% of people with influenza symptoms take sickness absence;
- the dynamic influenza model show that for seasonal influenza interventions to reduce presenteeism may reduce the attack rate by 0-23%, while for pandemic influenza interventions may reduce the attack rate by 0-12%. The largest relative effects were obtained assuming low transmissibility;
- generally, Positive Net Health Benefits (NHBs) were associated with early withdrawal (0.5-1 day after symptom onset); increasing adherence was only found cost-effective assuming low reproductive numbers.



Infectivity profile assumed in the dynamic model for influenza.

ComplexClin – Statistics for complex design of clinical studies

Rheumatoid Arthritis is a chronic disorder, and treatment includes drugs to help control disease and limit joint damage, such as the relatively inexpensive disease modifying anti-rheumatic drugs (DMARDs) and the more expensive biological anti-TNF (tumor necrosis factor) drugs. The latter drugs have been tested in clinical trials either alone or in combination with DMARDs, against another anti-TNF drug, just DMARDs, placebo or placebo and DMARDs combined.

ComplexClin has developed a mixed treatment comparison regression model for ranking the biological anti-TNF drugs with respect to treatment effect. We have conducted a literature research finding eligible clinical trials containing various comparisons of 9 different anti-TNF drugs, DMARDs and placebo combinations, resulting in a complex structural network of comparisons.

Our Bayesian MTC model enabled a comparison and ranking of the anti-TNF drugs with respect to their ACR50 (American College of Rheumatology) effect. Our model took into consideration relevant factors such as duration of RA disease prior to study start and drug doses given during the trials.

Based on this method ComplexClin

- quantified the impact of duration of RA disease prior to study start and drug doses given during the trials on treatment effect for each anti-TNF drug and all over;
- ranked the different anti-TNF drugs when given alone and when given jointly with MTX.

Hands deformed from Rheumatoid Arthritis. Photo: iStockphoto.com.

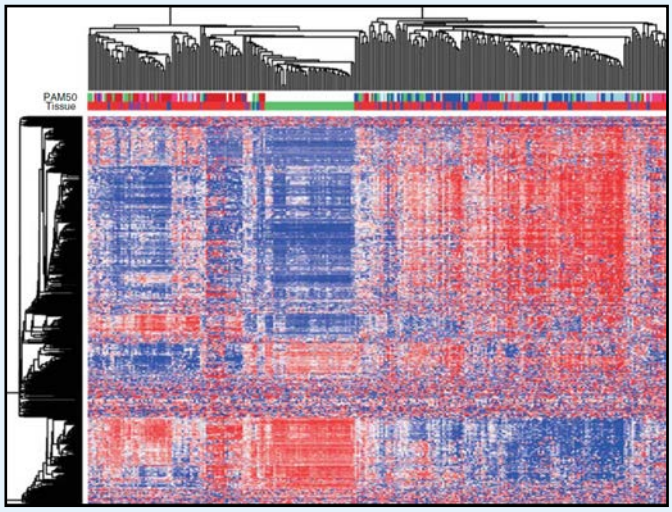
GeneStat – Statistics for genomic research

New molecular technologies allow analysing the function of genes in patients with an inheritable disease, like cancer or schizophrenia, in order to find the genetic signatures which allow early prevention, prediction of the efficacy of a therapy, production of new diagnostic tests and the development of new drugs.

GeneStat has developed new statistical methodologies which allow to extract knowledge from such huge amount of complex genomic data, collected to disclosure unknown molecular mechanisms of disease and health. Finding the biomarkers which carry true information about mechanisms, is challenged by the high risk of false positives: there are some 25000 genes, each could be a biomarker or any combinations of them — there are explosively many. Typically a cohort may include some hundreds of patients, very few compared to the number of possible biomarkers, and statistical methods have the task to identify the significant ones, minimising the risk for false positive and negative discoveries.

GeneStat developed classes of new methods to select the few genes which are most likely to be truly relevant; they include variations of Lasso penalised regression (for generalised linear or monotone non-linear models), bias-free preselection, multiple testing corrections and efficient Monte Carlo tests, and a generic mathematical construction of the comparison of genomic tracks in one and three dimensions (using stochastic geometry).

Based on these new methods and on a careful application of state-of-the-art methods, GeneStat



Hierarchical clustering of the methylation level of the 500 most variable gene regions.

- identified 7 genes, which allow predicting the efficacy of postmastectomy radiotherapy in breast cancer and which lead to the patenting of a new diagnostic test with Inven2;
- developed and implemented the Genomic Hyperbrowser, an open source tool to study and compare genomic tracks, which is used by hundreds of user worldwide and has led to important discoveries;
- discovered and validated several new gene signatures, which help understand molecular mechanisms in osteoporosis, psychiatric disorders, cervical, breast and colorectal cancers.

BioInfStat – Statistics for bioinformatics

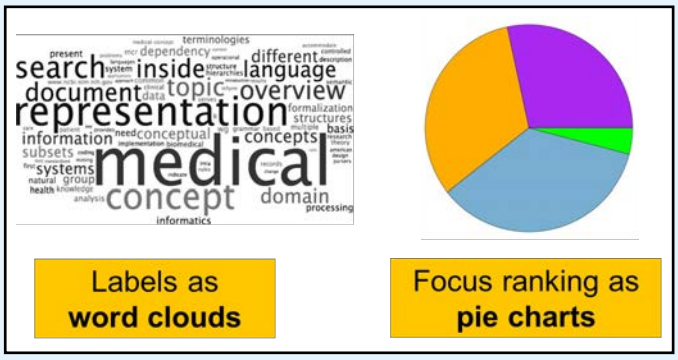
The overall goal has been to provide users with text mining tools that enable them to easily get an overview of the relevance of documents returned from a biomedical search and efficiently single out the useful information.

Document collections resulting from searches in the PubMed literature are often so large that some organisation of the returned information is necessary. Clustering is an efficient tool for organising search results. To help the user to decide how to continue the search for relevant documents, the content of each cluster can be characterised by a set of representative key words or cluster labels.

Through the project such solutions, based on adapted state-of-the-art methodology, have been developed and integrated into the CoreMine system, and as a result users can now make a keyword based search in PubMed, cluster the returned results and receive clusters accompanied by textual labels and summaries describing the contents. This is also integrated with the network concept of CoreMine, enabling visualisation of cluster relationships in a network.

Through our work we have seen that what is the best clustering, cluster labels or cluster summaries, very often depend on who you are asking. Different users will have different needs and interests, e.g. a biomedical researcher may see things differently from a GP. We have therefore developed a novel concept, called multi-focus, which give users the possibility to get an overview of the contents from multiple viewpoints.

The multi-focus visualisations give an overview of clusters along axes that the general labels are not able to convey. The approach is generic and should be applicable to any biomedical (or other) domain with any selection of foci where appropriate focus vocabularies can be established.



Visualisation through familiar concepts.

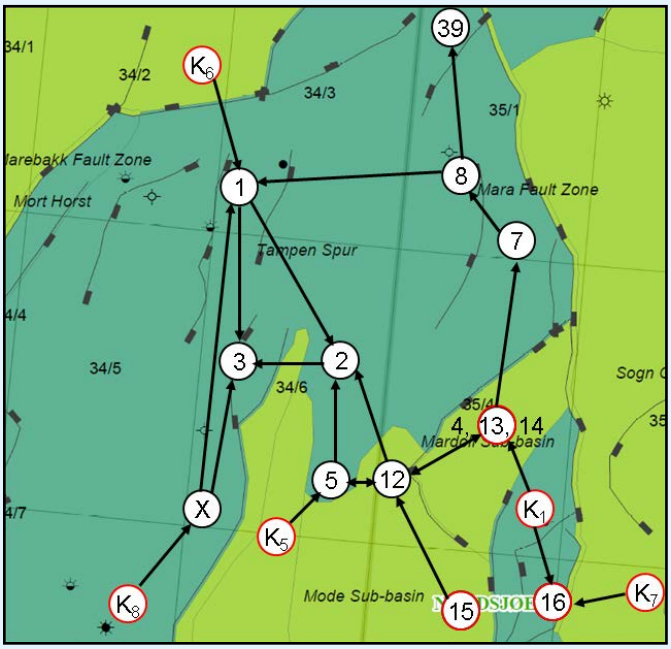
FindOil – Statistics for oil and gas exploration

The project name nicely summarises the goal of this project: To find oil, or more precisely, to find oil more efficiently. We have attacked this problem from two angles: By looking at the correlation between possible prospects (candidates for exploration drilling), and by looking at the data at a single prospect. In both cases, we use statistical models to obtain a new and better understanding of the problem.

Currently, oil companies tend to ignore the correlation between prospects when setting up drilling plans. Even after a new well is drilled, the update of other prospects is handled in an ad hoc manner. However, for an optimal exploration programme, the value of information at each well should be considered, together with the local discovery probability. We have shown that by setting up a full correlation model between prospects, based on existing geological information, the drilling sequence can in general be extended to give more discoveries without drilling more dry wells. The challenge for this model is to get a good user interface, so that geologists can use it.

At the prospect level, seismic data is the most important data source before drilling. We have set up a full consistent stochastic model for rock type, fluid and seismic data, and used it to find the probability for rock and fluid given seismic data. The resulting stochastic model is a spatial model incorporating rock and fluid continuity, as well as the physical model for linking rock and fluid to seismic data. This cannot be solved explicitly, so we have to rely on some approximations, the main being that we go from 3D to 1D, and solve only one vertical trace at a time. This is not as bad as it may sound, due to very continuous data.

We have developed a systematic trial and error approach for rock types and fluids that reduces the error in the approximations still needed for a vertical trace, and gives impressive results. This approach has been patented. Several implementations of this exist, under the name PCube. One version is generally available in commercial software, but does not use the patent. Statoil has one commercial implementation of the patented approach, while even the research version has a graphical user interface and is being used regularly.



Example of prospect network.

Awards

In addition to many of our research results being presented to international conferences as invited lectures, we highlight the following awards, given to (sfi)^2 staff.

- Sverdrup prize:

  Young scientists: Sara Martino 2009, Ida Scheel 2011, Ingrid Hobæk Haff and Kjetil Røysland 2013.

  Main: Tore Schweder 2011, Nils Lid Hjort 2013.

- Election of members in academies of science during the centre period:

  The Norwegian Academy of Science and Letters (DNV): Ørnulf Borgan, Arnoldo Frigessi, Knut Liestøl, and Håvard Rue.

  The Norwegian Academy of Technological Sciences (NTVA): Arnoldo Frigessi, and Lars Holden.

  The Royal Norwegian Society of Sciences and Letters (DKNVS): Håvard Rue.

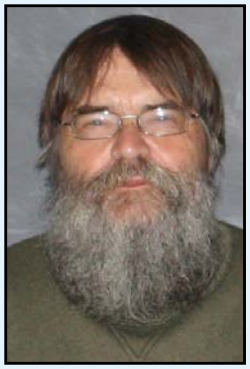
- American Statistical Association, Section on Bayesian Statistical Science, Student Paper Competition 2014: Reinaldo Marques, "Particle Move-Reweighting Strategies for Online Inference".

International cooperation

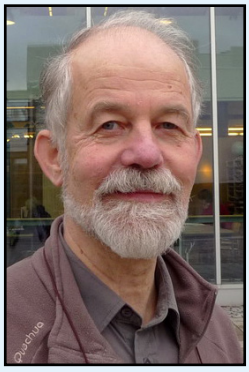
International appointments at (sfi)²

(sfi)² has appointed 4 adjunct professors:

Professor Peter Guttorp (University of Washington, Seattle, USA) was appointed in a 20% position at (sfi)² and NR, starting from 2009, as guest professor. Guttorp has contributed to various projects, in particular Climate-Insure and ComplexDepend.

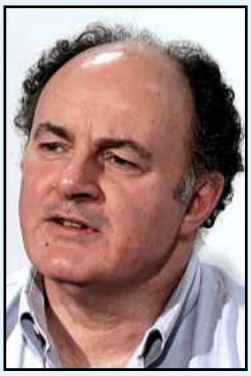


Professor Emeritus Elja Arjas (Helsinki University, Finland) was appointed in a 20% position at (sfi)², starting in 2013, as guest professor at the Department of Biostatistics, UiO. Arjas has contributed to various projects, in particular GeneStat, ComplexDepend and StatMarine.



Professor Sylvia Richardson (Imperial College, London, UK) was appointed in a 20% position at (sfi)², starting late 2009, as guest professor. Richardson has contributed to various projects, in particular GeneStat and Complex-Depend.

Associate Professor Gianpaolo Scalia Tomba (University of Rome II, Italy) ws appointed in a 20% position at (sfi)², starting in 2010. Scalia Tomba has contributed to various projects, in particular Infect, ComplexClin, Elprice, CustomerLife and ComplexDepend.



(sfi)² and STOR-i are international partners

As a key part of (sfi)² internationalisation we have developed a major strategic academic partnership with the Statistics and Operational Research Doctoral Training Centre (STOR-i) at Lancaster University. This key cross-disciplinary centre has been selected for its international research reputations and for the degree of alignment of their agendas with that of (sfi)² 's.

The STOR-i Doctoral Training Centre, a joint venture between the Departments of Mathematics and Statistics and Management Science at Lancaster University, offers a four-year PhD programme in Statistics and Operational Research (STOR) developed and delivered with industrial partners.

The first year of the programme comprises taught courses, projects and group activities to give a robust grounding in STOR, develop key research skills and identify students' research agenda. On successful completion of the first year students will progress to a PhD in years 2-4. This is when he/she will encounter real-life commercial challenges, develop leading-edge research and make a real impact on major industrial and scientific applications. Among the industrial partners of STOR-i there are Shell, Unilever, the UK met office, several pharma-ceutical companies (see full list: http://www.stor-i.lancs.ac.uk/partner_links).

The collaboration between STOR-i and (sfi)² has led to three PhD students who are currently co-supervised, several reciprocal visits of staff and students, and the organisation of joint workshops. Frigessi is a member of the scientific committee of STOR-i.

MASTMO – MSc programme in MAthematical and STatistical MOdelling

(sfi)² is a partner in the NTNU lead international project MASTMO (https://wiki.math.ntnu.no/mastmo).

MASTMO - MSc Programme in Mathematical and Statistical Modelling is bilateral Masterprogrammes, funded by NORAD (NOMA-2007/10048). The courses are held at Hawassa University (HU) in Ethiopia. MASTMO is based on and linked to the Industrial Mathematics programme run by the Department of Mathematical Sciences, NTNU.

The MSc Programme has a distinct applied profile with focus on a set of core applications. At the completion of the project Hawassa University shall be capable of managing the MASTMO MSc programme by itself. Qualification of HU faculty, implementation of the programme at HU and graduation of the first three cohorts of MASTMO MSc candidates, are the major activities of the programme.

(sfi)² staff has participated to the teaching and to the super-vision of students. We are also partner of a follow-up project, funded by Norhed, for a new period of the MASTMO School, integrated by a new PhD programme.

Main workshops and conferences

Abel Symposium: Statistical analysis for high-dimensional data, May 5-9, 2014, Lofoten

The topic of the Abel symposium was statistical and machine learning methodologies specifically developed for inference in situations where the number of parameters *p* is very much larger than the sample size *n*, with a particular reference to genomic applications.

Speakers presented new theories and methods, challenging applications and computational solutions. Specific themes were, among others, variable selection and screening, penalised regression, sparsity, thresholding, low dimensional structure, computational challenges, non-convex situations, learning graphical models, sparse covariance and precision matrices, semi and non-parametric formulations, multiple testing, classification, factor models, clustering, preselection.

The 14th Annual Winter School: Big Data and Modern Statistics, January 19-24, 2014, Geilo

The winter school, organised by SINTEF and (sfi)^2, focused on important innovative statistical theories, methods and computational tools that are being developed to handle more complex and high-dimensional data and subtle scientific questions.

It presented challenging statistical problems that require intensive computational approaches, structural model learning, causal inference, graphical models, functional data analysis, tree structured data, spatial data, point patterns and shape analysis, and more, both in a frequentist and Bayesian perspective. Applications spanned from genomics to computational anatomy, from ice with air bubbles images to geology, from human movement recognition to neurobiology, and from networks to personalise therapy.

Photo: iStockphoto.com.

Workshop on climate change effects in the Nordic industry, economy and society, June 3-4, 2013, Oslo

This workshop, organised by Norwegian Computing Center and the University of Oslo, focused on the prediction of effects of climate change in various branches of the Nordic economy and society. The Nordic countries have world class science communities in both statistics and climate research. Industry at large and society are among the world leaders in the analysis of the effect of climate change on their enterprises, activity and organisation.

The workshop was an occasion for stakeholders, statisticians and climate scientists to review the current understanding of climate change effects in key areas of the Nordic societies, and to start designing the research road map needed to reach a more precise understanding of future scenarios, with realistic measures of uncertainty, thus empowering business, technology and political decision makers with precise assessment of risk.

Photo: photos.com.

GEOSTATS – Ninth International Geostatistics Congress, June 11-15, 2012, Oslo

200 geostatisticians from all over the world participated on this 5 day conference in Oslo arranged by Norwegian Computing Center, University of Oslo and the Norwegian University of Science and Technology.

Topics were the latest developments in geostatistics and spatial statistics, both theoretical and applied, as well as practical issues in the many applications of statistics to natural phenomena. The International Geostatistics Congress has been an important arena for sharing and discussing the development of geostatistics in all these years.

The Second Workshop on Bayesian Inference for Latent Gaussian Models with Applications, May 30–June 1, 2012, Trondheim

A hundred participants joint the 3 day workshop, sponsored by (sfi)^2. Latent Gaussian models have numerous applications, for example in the social sciences, spatial and spatio-temporal data, epidemiology, climate modelling, and image processing. This workshop brought together researchers who develop methodology and fit datasets in this broad model class.

Trondheim. Photo: iStockphoto.com.

InFER2011 (Inference For Epidemic-related Risk), March 28- April 1, 2011, Warwick, UK

InFER2011 was a 5 day conference, supported by (sfi)^2, on the use of statistical inference in epidemic control and risk management. It drew together all those interested in analysing epidemic data, and using the results of those analyses to improve disease control. Through the use of statistical inference, the conference focused on bridging the gap between theoretical quantitative work on epidemics, and the data obtained from the field during disease outbreaks.

The topics included: Statistical inference in epidemics; Risk management; Guidance and policy; Novel applications.

Network modeling — Methods and applications in biology, medicine and sociology, May 7-8, 2012, Oslo

70 participants joint the 2 day workshop, organised by the Centre for Biostatistical Modelling in the Medical Sciences (BMMS) and (sfi)^2. Networks are everywhere, on the Internet, in social settings, in infectious disease, in biological systems.

During the last few years the theory of networks has emerged as a discipline in its own right with a well-developed theory and with very broad applications in many areas. The aim of this workshop was to present and discuss recent advances in this new and exciting discipline.

The Norwegian Academy of Science and Letters.
Photo: Anne-Marie Astad.

Modern Statistics for Climate Research: An international workshop, February 1-2, 2010, Oslo

50 participants from all five Nordic countries, Germany, United Kingdom and United States participated at the 2 days workshop. The aim of the workshop, organised and funded by (sfi)^2, Centre for Biostatistical Modelling in the Medical Sciences (BMMS) and the Norwegian Academy of Science and Letters, was to put together a group of central statisticians worldwide, developing advanced statistical methodology for climate research and adaptation.

The workshop was an important occasion to understand state-of-the-art science and future trends of a new discipline which might have the name statistical climatology, and will hopefully help to recruit young talents to this area of science of growing importance.

Photo: iStockphoto.com.

3rd Vine Copula Workshop, December 15-16, 2009, Oslo

In the past years, we have been leading in the development of a new general method to construct stochastic multivariate models and perform inference: the pair-copula construction (PCC). This builds on the probabilistic concept of Vine copula.

This was the third workshop on this specialised topic. The first two workshops were in Delft. Among the 20 participants and speakers, there were statisticians from Delft, München, British Columbia (Canada) and UK.

The Computing Center, Kristen Nygaards hus.

International Workshop on Causal Modelling, September 21-23, 2009, Oslo

This workshop was organised by the University of Oslo and (sfi)^2, and sponsored by the Research Council of Norway. Causal inference has become a major topic within the field of statistics with very important implications for fields like epidemiology, clinical medicine and social science.

Recent developments and breakthroughs have given a much more systematic view of how statistical analysis can contribute to a causal understanding. Nevertheless, a number of important issues remain.

In particular, this workshop focused on counterfactual causality; causality and time analysed by stochastic processes, including martingales; Granger causality, local dependence and dynamic path analysis. Results from the project "Dynamic modelling and causality", funded by the Research Council of Norway, were also presented.

Oslo. Photo: iStockphoto.com.

Insuring Future Climate Change — Preparing and acting today: A European conference for the insurance sector, November 3-4, 2008, Oslo

The conference was attended by 130 researchers, actuarians and staff from many Europeans insurance companies. It was part of our programme Climate Change and the Insurance Industry (CCII), which was funded by an EU Marie Curie Industry-Academia Partnerships and Pathways (IAPP) grant.

This conference brought together researchers, stakeholders and decision-makers from insurance, science and politics in order to stimulate discussion and dialog. The aim was to promote pro-active actions for reducing risks before serious impacts of climate change are upon us. We put together an exciting programme with speakers from Germany, France, Austria, the Netherlands, UK, Italy, Sweden and Norway.

The sessions covered a variety of topics such as climate risk communication, management and estimation for the insurance sector. Case studies from different regions in Europe covering the most central types of climate extremes such as storms, flooding, sea-level rise and snow-blizzards were presented.

Photo: iStockphoto.com.

ECMOR XI, European Conference on Mathematics of Oil Recovery, September 8-11, 2008, Bergen

(sfi)^2 with NR was co-organiser of this conference with 170 participants of 16 countries, who held 88 presentations. The ECMOR Conferences are a forum for presenting new results in mathematical research and development on the exploitation of oil and gas reservoirs.

An important aspect of the conferences is the interplay between theory and practice, where computations play a central role. Continuing high oil prices have ensured high research activity in recent years. We trust that this research effort has produced results that will contribute towards better recovery solutions. Lars Holden was co-chairman of the scientific committee.

Bergen. Photo: iStockphoto.com.

Training of researchers

(sfi)² had more than 100 collaborating scientists, from our research, public and enterprise partners, covering different roles in our research teams, from hands-on implementers to discussion partners, from theoretical developers to innovators. (sfi)² participated to the UiO and NTNU training programmes and funded many PhD students and postdocs.

We funded the first period of junior positions at NR and UiO which are permanent and will therefore continue also after after the conclusion of (sfi)². This is very important to create a long lasting effect, beyond the life-time of our centre. We were supporting master students associated to our centre, by funding their travel to workshops and conferences. NR offers 4-5 summer jobs annually, and some were working on (sfi)² projects.

PhD supervision was organised in teams, with competence on the methodological and applied side. Most junior staff is travelling to conferences and gives presentations of their results also to the applied community to which their work appeals.

All our students were recruited based on international calls. We have had students and postdocs from many different countries, from the US to Ethiopia, from Italy to China.

The principal researchers at (sfi)² have maintained their teaching obligations, and they were able to take themes from (sfi)² into various master courses. As part of the graduate school in biostatistics, we have organised and funded several courses in statistical genomics.

Here we present some feedback from three researchers who completed their PhD within the centre:

Linn Cecilie Bergersen

Already during my master thesis, my supervisor Ingrid Glad invited me along to small workshops and interesting seminars at (sfi)². This maintained my interest of a PhD at (sfi)². When I started my PhD, I got to work with extremely talented researchers and my PhD encompassed the entire range of elements which I believe a PhD in statistics should; statistical theory, applied statistics, real data sets, computational aspects and knowledge about the field of application.

I would like to emphasise the value that (sfi)² brought to me by encouraging me to attend conferences and to present my work; to make me confident speaking and presenting in front of an audience, and to meet other PhD students and researchers from all over the world working with similar things.

Through (sfi)² I established contact with Sylvia Richardson, and I got the chance to work closely with her during a 3 months research stay in Paris. Today I am working in a Norwegian software as a service company, Cxense ASA, within the field of big data.

I knew nothing about the media industry from before I started in Cxense, and my time as a PhD student at (sfi)² taught me nothing about business directly, but everything about being a researcher, communicating and presenting statistics to anyone, working in teams, being critical when that is needed, asking the right questions, working through difficult problems and being structured. All which are skills I've found extremely valuable entering business and a field I had limited knowledge about before I started.

Ingrid Hobæk Haff

After working as a research scientist at the Norwegian Computing Center for a few years, I wanted to do a PhD in order to improve my researcher skills and competence. (sfi)² offered a unique opportunity to do a PhD that was both methodological and applied, within themes that I found very interesting.

The years at (sfi)² were instructive and exciting, with a varied work environment, among others due to visits from universities abroad. I also got the opportunity to spend a semester at l'Université Catholique de Louvain in Belgium. After finishing my PhD, I returned to my position as researcher at the Norwegian Computing Center.

Ingrid Hobæk Haff with the PhD evaluation committee, from left: Professor Christian Genest, Department of Mathematics and Statistics, McGill University; Ingrid Hobæk Haff, NR; Associate Professor Ida Scheel, Department of Mathematics, UiO; Professor Paul Embrechts, Department of Mathematics, ETH Zürich.

Gabriele Martinelli

I chose to join a doctoral programme in partnership with the (sfi)², since I wanted to be exposed to a centre where I could appreciate different uses of applied statistical methodology, and learn much more than in a usual statistics department.

I believe that my PhD experience was very positive, thanks to my supervisors, to the interaction with NTNU and with NR, and to the interactions with my colleagues, with whom we had several opportunities of sharing knowledge and ideas.

Currently, I am working with applied statistics and decision support problems in a large oil&gas service company, Schlumberger.

Employment of PhD-candidates (2015)

Employment of PhD-candidates (number)							
By centre company	By other companies	By public organisations	By universities	By research institutes	Outside Norway	Other	Total
-	8	1	6	5	-	-	**20**

Communication/Popular dissemination of knowledge

(sfi)^2 was present in the Norwegian society at large through visibility in the media and target events. Together with NR, we presented statistics to the junior public at Forskningstorget; we presented the results obtained in our centre in several workshops and public events, including an evening at Litteraturhuset in Oslo, in occasion of the international year of statistics.

An article celebrating (sfi)^2 is published on a peer reviewed journal, see Frigessi, Holden and Teigland (2015)[1] .

1. Frigessi A., Holden L., and Teigland A. (2015). (sfi)^2 Statistics for Innovation – The experience of the Oslo centre in industrial statistics, Appl. Stochastic Models Bus. Ind., 31, 82–93, with discussion.

Future prospects

NR has about 40 persons within statistics that work with contract research within statistics funded by the Research Council, EU, other kinds of research funding and private and public clients. This activity has been important for building up the competence that made it possible to get the SFI.

This contract research activity has continued during the lifetime of the centre and continues after the centre is closed. Some of the activity has been moved between the centre and the contract research part during the lifetime of (sfi)^2, depending on the matureness of the research and patent considerations.

NR has had a portfolio of projects with Statoil, DNB, Gjensidige, Hydro and Institute of Marine Research that continues after (sfi)^2 has ended.

NR will also continue the research cooperation with UiO and Oslo University Hospital with a portfolio of new projects. Professor Håvard Rue, NTNU, will continue to work part-time at NR. NR has also been awarded a new centre for research based innovation, *Big Insight*. Some of our activity in finance and health will continue as part of the new centre.

Financing through the life of the centre

Summary sheet for the main categories of partners (MNOK)

Contributor	Cash	In-kind	Total
Host		29,4	**29,4**
Research partners	4,0	110,3	**114,3**
Companies	15,7	49,1	**64,7**
Public partners	3,2	44,7	**47,9**
RCN	80,0		**80,0**
Sum	**102,9**	**232,5**	**336,3**

Distribution of resources (MNOK)

Type of activity	
Research projects	327,0
Common centre activities	3,0
Administration	6,3
Sum	**336,3**

Results

Key figures

	Total
Scientific/scholary publications (peer reviewed)	223
Dissemination measures for users	25 workshops, 21 courses, 4 conferences, 2 career days. Weekly (sfi)^2 lunches, informal gathering of staff and students with a lecture on relevant topics. Weekly Tuesday seminar in statistics, at the Department of Mathematics, UiO. Monthly Thursday seminar in biostatistics, at the Department of Biostatistics, UiO.
Dissemination measures for the general public	Forskningstorget in Oslo, yearly from 2007 to 2014. Open NTVA-meeting in Oslo: "Statistics for Innovation - (sfi)^2: Fra gener til elektrisitetspriser, statistics in action" (April 2012). Celebrating the International Year of Statistics in Oslo: "Fra tall til innsikt" (November 2013).
Number of new/improved methods/ models/prototypes finalised	14
Number of new/improved products/ processes/services finalised	28
PhD-degrees completed	20

Conclusions

Patents

- Kolbjørnsen, Odd; Buland, Arild; Hauge, Ragnar. "Method of Modelling A Subterranean Region of the Earth". 2008. Patent number/License number: GB2463242.
- Mohammed, H; Tramm, T; Myhre, S; Alsner, J; Børresen-Dale, AL; Sørlie, T; Overgaard, J; Frigessi, A. "A gene signature associated with efficacy post-mastectomy radiotherapy in breast cancer". Application number 61607316, EFS ID 12237804, 6 March 2012. INVEN-32535/US-1/PRO Patent Application by UiO in collaboration with University of Aarhus.
- Inven2 as, UiO; OUS: "New gene index, composed of 7 genes, the expression of which allows to predict the efficacy of radiotherapy". A patent has been filed and confirmed.
- Sandve, GK; Gundersen, S; Rydbeck, H; Glad, IK; Holden, L; Holden, M; Liestøl, K; Clancy, T; Frigessi, A; Hovig, E. "Genomic Hyperbrowser", DOFI 14.10.2011, Invent2 as, Invention 11083.
- Sørlie, T; Frigessi, A; Børresen-Dale, A-L; Myhre, S; Mohammed, H; Overgaard, J; Alsner, J; Tramm, T. "A gene signature associated with efficacy of postmastectomy radiotherapy in breast cancer". Provisional Patent number: INVEN-32535/US-1/PRO. Filed March 6, 2012. The provisional patent application for the gene signature has been filed by Inven2 AS on behalf of OUS and UiO (U.S. Provisional Patent Application Serial No. 61/607,316).

NR had a close cooperation with most of the partners before the centre started. We knew what to expect from each other. Hence, we have been able to focus on the important problems of finding the good innovation challenges, scientific solutions and recruitment. We have had no disagreements on patents or any conflicts during the lifetime of the centre. Some of the critical parts are:

Valuable innovation: Focus on the development of new statistical methodology that is really necessary to achieve successful and valuable innovation aims. Developed methods and tools should eventually be built inside innovative products and services. At (sfi)^2, this is the case for several of our results, while others are not yet part of operational pipelines as they are still evaluated, compared and optimised with our partners.

Culture: Establish a new research culture, at the interface between statistical science and long-term innovation attractive for people from academia, industry, contract research and PhD students. Here seminars, lunches and many good dialogs have contributed to the success of (sfi)^2.

People: Recruitment of new master and PhD students is critical for all partners. (sfi)^2 has trained about 20 PhD students and 10 master students that all have a promising career in front of them in a very important area for the industry and the society at large. A good leadership of the centre that makes all the partners contribute in a constructive manner is very important.

High scientific level: This is the most important part of all research. More than 220 papers that have received much attention and many important results prove that this has been a focus for the centre.

We believe that (sfi)^2 has been a great success. This had not been possible without the generous funding from the Research Council of Norway and all the partners and people contributing to the centre in their various roles.

From the science festival Forskningstorget on Karl Johans gate. Photo: NR.

Appendix 1

Statement of accounts for the complete period of centre financing

Funding (in MNOK)

Funding	RCN	NR	UiO	NTNU	Statoil	DNB	Gjensidige	Hydro	Smerud	Biomolex	Pub-Gene	Spermatech	Sencel	OUS	IMR	Total
FindOil	9,7	0,3	3,0	4,8	13,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	30,8
TotalRisk	8,0	0,0	2,8	0,0	0,0	10,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	20,9
CustomerLife	6,2	0,0	2,5	0,0	0,0	0,0	5,7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	14,4
StatMarine	9,6	2,4	5,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	12,1	29,1
Elprice	5,5	0,1	1,1	0,0	0,0	0,0	0,0	11,8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	18,5
ClimateInsure	4,9	0,0	3,0	2,7	0,0	0,0	0,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	11,5
ComplexClin	1,8	0,2	1,3	0,0	0,0	0,0	0,0	0,0	3,2	0,0	0,0	0,0	0,0	0,0	0,0	6,4
BioInfStat	3,2	0,1	1,5	0,0	0,0	0,0	0,0	0,0	0,0	5,7	9,3	1,4	3,8	0,0	0,0	25,0
BioInfStat FR	5,8	3,9	7,5	7,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	34,8	0,0	59,5
Infect	1,8	0,0	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,3
GeneStat+PPS	3,9	1,1	24,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	29,1
ComplexDepend	12,7	5,7	33,1	10,1	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	61,8
Administration	0,0	6,1	0,4	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	6,5
Costs	6,9	9,6	4,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	20,5
Sum	**80,0**	**29,4**	**89,7**	**25,1**	**13,0**	**10,1**	**6,9**	**11,8**	**3,2**	**5,7**	**9,3**	**1,4**	**3,8**	**34,8**	**12,1**	**336,3**

Cost (in MNOK)

Funding	NR	UiO	NTNU	Statoil	DNB	Gjensidige	Hydro	Smerud	Biomolex	Pub-Gene	Spermatech	Sencel	OUS	IMR	Total
FindOil	10,8	6,2	4,8	9,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	30,8
TotalRisk	8,2	5,7	0,0	0,0	6,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	20,9
CustomerLife	8,5	4,2	0,0	0,0	0,0	1,7	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	14,4
StatMarine	11,9	8,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	8,9	29,1
Elprice	8,8	1,9	0,0	0,0	0,0	0,0	7,8	0,0	0,0	0,0	0,0	0,0	0,0	0,0	18,5
ClimateInsure	2,4	5,7	2,7	0,0	0,0	0,9	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	11,7
ComplexClin	1,9	1,3	0,0	0,0	0,0	0,0	0,0	3,2	0,0	0,0	0,0	0,0	0,0	0,0	6,4
BioInfStat	4,2	1,5	0,0	0,0	0,0	0,0	0,0	0,0	5,7	8,3	1,4	3,8	0,0	0,0	25,0
BioInfStat FR	5,4	11,8	7,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	34,8	0,0	59,5
Infect	0,0	2,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	2,3
GeneStat+PPS	1,4	27,6	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	29,1
ComplexDepend	12,3	36,1	10,1	0,0	0,0	0,3	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	58,8
Administration	6,1	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	6,1
Costs	19,7	4,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	23,7
Sum	**101,7**	**116,7**	**25,1**	**9,0**	**6,9**	**2,9**	**7,8**	**3,2**	**5,7**	**8,3**	**1,4**	**3,8**	**34,8**	**8,9**	**336,3**

Appendix 2

List of Postdocs, Candidates for PhD and MSc degrees during the full period of the centre

Postdoctoral researchers with financial support from the Centre budget

Name	M/F	Nationality	Years/period in the centre
Egil Ferkingstad	M	Norway	2007-2009
Sara Martino	F	Italy	2008-2009
Ida Scheel	F	Norway	2008-2012
Alex Lenkoski	M	USA	2012-2014
Xiangping Hu	M	China	2013-2015

Postdoctoral researchers working on projects in the centre with financial support from other sources

Name	M/F	Nationality	Source of funding	Years/period in the centre
Hiroko Solvang	F	Japan	OUS	2007-2010
Kjetil Røysland	M	Norway	UiO	2007-2013
Glenn Lawyer	M	USA	UiO/NFR	2008-2009
Andrew Brown	M	UK	UiO	2008-2010
Geir Kjetil Sandve	M	Norway	UiO	2008-2012
Ståle Nygård	M	Norway	OUS/UiO	2009-2013
Finn Lindgren	M	Sweden	NTNU	2010-2012
Daniel Simpson	M	New Zealand	NTNU	2013-2014
Kukatharmini Tharmaratnam	F	Sri Lanka	UiO	2011-2014
Valeria Vitelli	F	Italy	UiO	2013-2015

PhD candidates who have completed with financial support from the centre budget

Name	M/F	Nationality	Scientific area	Years/period in the centre	Thesis title	Main thesis advisor
Steffen Grønneberg	M	Norway	Finance	2007-2011	Some applications of stochastic process techniques to statistics	Kjersti Aas, Nils Lid Hjort
Marion Haugen	F	Norway	Health, Finance	2007-2010	Using frailty models to account for heterogeneity and associations in cancer epidemiology and insurance	Tron A. Moger
Ingrid Hobæk Haff	F	Norway	Finance	2008-2012	Pair-copula constructions – an inferential perspective	Arnoldo Frigessi, Kjersti Aas, Ørnulf Borgan
Linn Cecilie Bergersen	F	Norway	Method, Health	2009-2013	Guiding the Lasso: Regression in High Dimensions	Ingrid K. Glad, Nils Lid Hjort
Gabriele Martinelli	M	Italy	Petroleum	2009-2012	Petroleum prospect exploration using Bayesian Networks	Jo Eidsvik, Ragnar Hauge, Håkon Tjelmeland
Halfdan Rydbeck	M	Sweden	Health	2010-2012	Integrative epigenome analysis	Eivind Hovig
Øystein Sørensen	M	Norway	Health	2011-2014	Rank and Measurement Error Modeling for Robust Genomic Inference	Magne Thoresen, Arnoldo Frigessi

PhD candidates who have completed with other financial support, but associated with the centre

Name	M/F	Nationality	Years/period in the centre	Thesis title	Main thesis advisor
Sara Martino	F	Italy	2007-2007	Approximate Bayesian Inference for Latent Gaussian Models	Håvard Rue, Håkon Tjelmeland
Ståle Nygård	M	Norway	2007-2008	Statistical modeling of complex time-dependent data in genomics and ecology	Ørnulf Borgan, Arnoldo Frigessi, Magne Aldrin
Nils Haavardsson	M	Norway	2007-2008	Hydrocarbon product optimization in multi-reservoir fields - tools for enhanced value chain analysis	Arne Bang Huseby
Hege Bøvelstad	F	Norway	2007-2009	Survival prediction from high-dimensional genomic data	Ørnulf Borgan, Arnoldo Frigessi, Knut Liestøl
Yiting Xue	F	China	2007-2010	Mathematical and statistical models of infectious diseases and their applications in health economic evaluation	Birgitte F. de Blasio
Hayat Mohammed	F	Ethiopia	2007-2011	Statistical tools for high-dimensional genomic data	Arnoldo Frigessi
Xi Zhao	F	China	2007-2010	Systems Biology and Genomics of Breast Cancer through mining multi-level molecular profiles	Ole Christian Lingjærde, Anne-Lise Børresen-Dale
Gudmund H. Hermansen	M	Norway	2008-2013	Model selection and Bayesian nonparametrics for time series and non-standard regression models	Nils Lid Hjort, Arnoldo Frigessi
Sveinung Gundersen	M	Norway	2009-2013	Representation and integrated analysis of heterogeneous genomic datasets	Eivind Hovig
Thiago G. Martins	M	Brazil	2011-2014	Bayesian analysis of latent Gaussian models	Håvard Rue
Kristoffer H. Hellton	M	Norway	2011-2014	On high-dimensional principal component analysis in genomics: consistency and robustness	Magne Thoresen
Marissa LeBlanc	F	Canada	2008-2014 Defence in 2015	Correcting for sample overlap in cross-trait analysis of genome-wide association studies	Bettina Kulle Andreassen, Arnoldo Frigessi
Geir-Arne Fuglstad	M	Norway	2011-2015 Defence in 2015	Modelling Spatial Non-stationarity	Håvard Rue
Tonje Gulbrandsen Lien	F	Norway	2011-2015 Defence in 2015	Data integration in high dimensional genomics	Ingrid K. Glad
Reinaldo A. G. Marques	M	Brazil	2011-2014 Defence in 2015	Sequential inference in dynamic hierarchical models	Geir Storvik
Jonas Paulsen	M	Norway	2011-2015 Defence in 2015	Inferential analysis of genomic 3D organization	Eivind Hovig
Olav Nikolai Breivik	M	Norway	2012-2015 Defence in 2015	Applications of Bayesian hierarchical models	Geir Storvik
Christian Rohrbeck	M	Germany	2013-2016 Defence in 2016	Bayesian Spatial Monotonic Multivariate Regression with an application in climate change for the insurance industry	Arnoldo Frigessi, Jonathan Tawn

PhD students with financial support from the centre budget who still are in the process of finishing studies

Name	M/F	Nationality	Scientific area	Years/period in the centre	Thesis topic	Main thesis advisor
Marie Lilleborge	F	Norway	Petroleum	2012-2016	Efficient Information gathering in discrete Bayesian Networks	Ragnar Hauge
Martin Jullum	M	Norway	Petroleum, Method	2012-2015	A new approach to the focused information criterion and an approximate Gaussian procedure applied to geophysical inversion	Nils Lid Hjort

Appendix 3
List of Publications

Submitted papers

Eikvil, L; Jenssen, TK; Holden, M. Multi-focus cluster labeling. Submitted, December 2014.

Ferkingstad, E; Løland, A. Coping with area price risk in electricity markets: Forecasting Contracts for Difference in the Nordic power market. Submitted, June 2014.

Hobæk Haff, I; Aas, K; Frigessi, A; Graziani, VL. Structure learning in BBNs using regular vines. Submitted, March 2015.

Lenkoski, A; Hobæk Haff, I; Neef, LR; Løland, A. Calibrated Probabilities and the Investigation of Soft Fraud in Automobile Insurance Claims. Submitted, December 2013.

Lilleborge, M; Hauge, R; Eidsvik, J. Information gathering in Bayesian Networks with an application to petroleum prospecting. Submitted, February 2015.

Marques, R; Storvik, G. Particle move-reweighting strategies for online inference. Submitted, 2013.

Sørensen, Ø; Frigessi, A; Thoresen, M. Covariate selection in high-dimensional generalized linear models with measurement error. Submitted, July 2014.

Sørensen, Ø; Vitelli, V; Frigessi, A; Arjas, E. Bayesian inference from rank data. Submitted, May 2014.

Scheel, I; Frigessi, A; Hammer, H; Storvik, B. A Bayesian strategy for ranking customers by individual unobserved estimated risk factors. Submitted, 2014.

Papers (published and accepted)

Aalen, OO. Armitage lecture 2010: Understanding treatment effects: the value of integrating longitudinal data and survival analysis. Statistics in Medicine, 2012; Vol. 31: 1903–1917.

Aalen, OO. Survival analysis, longitudinal analysis and causal inference. Lifetime Data Analysis, 2010; Vol. 16(1): 1-2.

Aalen, OO. Understanding disease processes. Statistics in Medicine, 2010; Vol. 29(11): 1159-1160.

Aalen, OO; Frigessi, A. What can Statistics Contribute to a Causal Understanding? Scandinavian Journal of Statistics, 2007: Vol. 34(1): 155-168.

Aalen, OO; Røysland, K; Gran, JM; Ledergerber, B. Causality, mediation and time: a dynamic viewpoint. Journal of the Royal Statistical Society: Series A (Statistics in Society), 2012; Vol. 175(4): 831–861.

Aanes, S; Engen, S; Sæther, BE; Aanes, R. Estimation of fish stock dynamical parameters from catch-at-age data and indices of abundance: can natural and fishing mortality be separated? Canadian Journal of Fisheries and Aquatic Sciences, 2007; Vol. 64: 1130-1142.

Aanes, S; Nedreaas, K; Ulvatn, S. Estimation of total retained catch based on frequency of fishing trips, inspections at sea, transhipment, and VMS data. ICES Journal of Marine Science, 2011; Vol. 68(8): 1598-1605.

Aanes, S; Vølstad, JH. Efficient statistical estimators and sampling strategies for estimating the age-composition of fish. Canadian Journal of Fisheries and Aquatic Sciences. Published on the web 03 March 2015. doi: 10.1139/cjfas-2014-0408.

Aas, Kjersti; Berg, Daniel. Modelling dependence between financial returns using pair-copula constructions. In DEPENDENCE MODELING: Handbook on Vine Copulae, D. Kurowicka and Harry Joe (eds.), World Scientific Publishing Co., February 2011.

Aas, Kjersti; Czado, Claudia; Frigessi, Arnoldo; Bakken, Henrik. Pair-copula constructions of multiple dependence. Insurance: Mathematics and Economics, 2009; Vol. 44 (2): 182-198.

Aas, K; Puccetti, G. Bounds on total economic capital: the DNB case study. Extremes, 2014; doi: 10.1007/s10687-014-0202-0.

Aldrin, M; Lyngstad, TM; Kristoffersen, AB; Storvik, B; Borgan, Ø; Jansen, PA. Modelling the spread of infectious salmon anaemia (ISA) among salmon farms based on seaway distances between farms and genetic relationships between infectious salmon anaemia virus isolates. Journal of the Royal Society Interface, 2011; Vol. 8(62): 1346-1356.

Aldrin, M; Mortensen, B; Storvik, G; Nedreaas, K; Aglen, A; Aanes, S. Improving management decisions by predicting fish bycatch in the Barents Sea shrimp fishery. ICES Journal of Marine Science, 2012; Vol. 69 (1): 64-74.

Aldrin, M; Raastad, R; Tvete, IF; Berild, D; Frigessi, A; Leegaard, TM; Monnet, DL; Walberg, M; Müller, F. Antibiotic resistance in hospitals: a ward-specific random effect model in a low antibiotic consumption environment. Statistics in Medicine, 2013; Vol. 32(8): 1407-1418.

Aldrin, M; Storvik, B; Frigessi, A; Viljugrein, H; Jansen, PA. A stochastic model for the assessment of the transmission pathways of heart and skeleton muscle inflammation, pancreas disease and infectious salmon anaemia in marine fish farms in Norway. Preventive Veterinary Medicine, 2010; Vol. 93: 51-61.

Aldrin, M; Storvik, B; Kristoffersen, AB; Jansen, PA. Space-time modelling of the spread of salmon lice between and within Norwegian marine salmon farms. PLoS ONE, 2013; Vol. 8(5): e64039, doi:10.1371/journal.pone.0064039.

Ambrosi, A; Glad, IK; Pellin, D; Cattoglio, C; Mavilio, F; Di Serio, C; Frigessi, A. Estimated Comparative Integration Hotspots Identify Different Behaviors of Retroviral Gene Transfer Vectors. PLoS Comput Biol, 2011; Vol. 7(12): e1002292.

Antoniadis, A; Glad, IK; Mohammed, H. Local comparison of empirical distributions via nonparametric regression. Journal of Statistical Computation and Simulation, 2014; doi: 10.1080/00949655.2014.929133.

Baghishani, H; Rue, H; Mohammad-zadeh, M. On a hybrid data cloning method and its application in generalized linear mixed models. Statistics and Computing, 2012; Vol. 22(2): 597–613.

Beguin, J; Martino, S; Rue, H; Cumming, SG. Hierarchical analysis of spatially autocorrelated ecological data using integrated nested Laplace approximation. Methods in Ecology and Evolution, 2012; Vol. 3: 921–929.

Berg, D. Copula goodness-of-fit testing: An overview and power comparison. European Journal of Finance, 2009; Vol. 15(7/8): 675-701.

Berg, D; Aas, K. Models for construction of multivariate dependence. European Journal of Finance, 2009; Vol. 15(7/8): 639-665.

Berg, D; Quessy, JF. Local Power Analyses of Goodness-of-fit Tests for Copulas. Scandinavian Journal of Statistics, 2009; Vol. 36(3): 389-412.

Bergersen, LC; Glad, IK; Lyng, H. Weighted Lasso with Data Integration. Statistical Applications in Genetics and Molecular Biology, 2011; Vol. 10(1). doi: 10.2202/1544-6115.1703.

Bergersen, LC; Tharmaratnam, K; Glad, IK. Monotone Splines Lasso. Computational Statistics & Data Analysis, 2014; Vol. 77: 336-351.

Bhattacharjya, D; Eidsvik, J; Mukerji, T. The value of information in portfolio problems with dependent projects. Decision Analysis, 2013; Vol. 10(4): 341-351.

Bivand, RS; Gómez-Rubio, V; Rue, H. Approximate Bayesian inference for spatial econometrics models. Spatial Statistics, 2014; Vol. 9: 146-165.

Bivand, RS; Gómez-Rubio, V; Rue, H. Some spatial statistical extensions to R-INLA. In press, Journal of Statistical Software.

Bjørkvoll, EM; Grøtan, V; Aanes, S; Sæther, BE; Engen, S; Aanes, R. Stochastic Population Dynamics and Life-History Variation in Marine Fish Species. American Naturalist, 2012; Vol. 180(3): 372-387.

Blangiardo, M; Cameletti, M; Baio, G; Rue, H. Spatial and spatio-temporal models with R-INLA. Spatial and Spatio-temporal Epidemiology, 2013; Vol. 4: 33-49.

Bolin, D; Guttorp, P; Januzzi, A; Jones, D; Novak, M; Podschwit, H; Richardson, L; Särkkä, A; Sowder, CD; Zimmerman, A. Statistical prediction of global sea level from global temperature. In press, Statistica Sinica. doi: 10.5705/ss.2013.222w.

Bolin, D; Lindgren, F. Spatial models generated by nested stochastic partial differential equations, with an application to global ozone mapping. Annals of Applied Statistics, 2011; Vol. 5(1): 523-550.

Bray, F; Haugen, M; Moger, TA; Tretli, S; Aalen, OO; Grotmol, T. Age-incidence curves of nasopharyngeal carcinoma worldwide: bimodality in low-risk populations and aetiologic implications. Cancer Epidemiol Biomarkers Prev., 2008; Vol. 17(9): 2356-2365.

Brechmann, EC; Czado, C; Aas, K. Truncated regular vines in high dimensions with application to financial data. Canadian Journal of Statistics, 2012; Vol. 40(1): 68-85.

Brown, AA; Jensen, J; Nikolova, YS; Djurovic, S; Agartz, I; Server, A; Ferrell, RE; Manuck, SB; Mattingsdal, M; Melle, I; Hariri, AR; Frigessi, A; Andreassen, OA. Genetic variants affecting the neural processing of human facial expressions: evidence using a genome-wide functional imaging approach. Translational Psychiatry, 2012; Vol. 2: e143, doi: 10.1038/tp.2012.67.

Buland, A; Kolbjørnsen, O. Bayesian inversion of CSEM and magnetotelluric data. Geophysics, 2012; Vol. 77(1): E33-E42.

Buland, A; Kolbjørnsen, O; Carter, AJ. Bayesian Dix inversion. Geophysics, 2011; Vol. 76(2): R15-R22.

Buland, A; Kolbjørnsen, O; Hauge, R; Skjæveland, Ø; Duffaut, K. Bayesian lithology and fluid prediction from seismic prestack data. Geophysics, 2008; Vol. 73(3): C13-C21.

Bøhn, SK; Myhrstad, MC; Thoresen, M; Holden, M; Karlsen, A; Tunheim, SH; Erlund, I; Svendsen, M; Seljeflot, I; Moskaug, JØ; Laake, P; Arnesen, H; Tonstad, S; Collins, AR; Drevon, CA; Blomhoff, R. Blood cell gene expression associated with cellular stress defence is modulated by antioxidant-rich food in a randomised controlled clinical trial of male smokers. BMC Medicine, 2010; Vol. 8:54.

Bøhn, SK; Russnes, KM; Sakhi, AK; Thoresen, M; Holden, M; Moskaug, JØ, Myhrstad, MC; Olstad, OK; Smeland, S; Blomhoff, R. Stress associated gene expression in blood cells is related to outcome in radiotherapy treated head and neck cancer patients. BMC Cancer, 2012; Vol. 12:426. doi:10.1186/1471-2407-12-426.

Bøvelstad, HM; Borgan, Ø. Assessment of evaluation criteria for survival prediction from genomic data. Biometrical Journal, 2011; Vol. 53(2): 202-216.

Bøvelstad, HM; Nygård, S; Borgan, Ø. Survival prediction from clinico-genomic models — a comparative study. BMC Bioinformatics, 2009: Vol. 10:413.

Bøvelstad, HM; Nygård, S; Størvold, HL; Aldrin, M; Borgan, Ø; Frigessi, A; Lingjærde, OC. Predicting survival from microarray data — a comparative study. Bioinformatics, 2007; Vol. 23(16): 2080-2087.

Cameletti, M; Lindgren, F; Simpson, D; Rue, H. Spatio-temporal modeling of particulate matter concentration through the SPDE approach. AStA Advances in Statistical Analysis, 2013; Vol. 97(2): 109-131.

Cheng, Y; Lenkoski, A. Hierarchical Gaussian graphical models: Beyond reversible jump. Electronic Journal of Statistics, 2012; Vol. 6: 2309-2331.

Claeskens, G; Hjort, NL. Minimizing average risk in regression models. Econometric Theory, 2008; Vol. 24: 493-527.

Cooke, R; Joe, H; Aas, K. Vines Arise. In DEPENDENCE MODELING: Handbook on Vine Copulae, D. Kurowicka and Harry Joe (eds.), World Scientific Publishing Co., Scheduled Fall 2010.

Craigmile, PF; Guttorp, P; Lund, R; Smith, RL; Thorne, PW; Arndt, D. Warm streaks in the U.S. temperature record: What are the chances? Journal of Geophysical Research – Atmospheres, 2014; Vol. 119(10): 5757-5766.

Dakovic, R; Czado, C; Berg, D. Bankruptcy prediction in Norway: a comparison study. Applied Economic Letters, 2010; Vol. 17(17): 1739-1746.

de Blasio, BF; Iversen, BG; Scalia Tomba, G. Effect of Vaccines and Antivirals during the Major 2009 A(H1N1) Pandemic Wave in Norway — And the Influence of Vaccination Timing. PLoS ONE, 2012; Vol. 7(1): e30018. doi: 10.1371/journal.pone.0030018.

de Blasio, BF; Kasymbekova, K; Flem, E. Dynamic model of rotavirus transmission and the impact of rotavirus vaccination in Kyrgyzstan, Vaccine, 2010; Vol. 28(50): 7923-7932.

de Blasio, BF; Rae Neilson, A; Klemp, M; Skjeldestad, FE. Modeling the impact of screening policy and screening compliance on incidence and mortality of cervical cancer in the post-HPV vaccination era. Journal of Public Health, 2012; Vol. 34(4): 539-547.

de Blasio, BF; Seierstad, TG; Aalen, OO. Frailty effects in networks: comparison and identification of individual heterogeneity versus preferential attachment in evolving networks. Journal of the Royal Statistical Society, Series C (Applied Statistics), 2011; Vol. 60(2): 239- 259.

Disanto, G; Sandve, GK; Berlanga-Taylor, AJ; Morahan, JM; Dobson, R; et al. Genomic Regions Associated with Multiple Sclerosis Are Active in B Cells. PLoS ONE, 2012; Vol. 7 (3): e32281. doi:10.1371/journal.pone.0032281.

Disanto, G; Sandve, GK; Berlanga-Taylor, AJ; Ragnedda, G; Morahan, JM; Watson, CT; Giovannoni, G; Ebers, GC; Ramagopalan, SV. Vitamin D receptor binding, chromatin states and association with multiple sclerosis. Human Molecular Genetics, 2012; Vol. 21(16): 3575-3586.

Disanto, G; Sandve, GK; Ricigliano, VA; Pakpoor, J; Berlanga-Taylor, AJ; Handel, AE; Kuhle, J; Holden, L; Watson, CT; Giovannoni, G; Handunnetthi, L; Ramagopalan, SV. DNase hypersensitive sites and association with multiple sclerosis. Human Molecular Genetics, 2014; Vol. 23(4): 942-948.

Di Serio, C; Pellin, D; Ambrosi, A; Glad, IK; Frigessi, A. Biostatistics Meets Bioinformatics in Integrating Information from Highdimensional Heterogeneous Genomic Data: Two Examples from Rare Genetic Diseases and Infectious Diseases. In: Computational Intelligence Methods for Bioinformatics and Biostatistics, Lecture Notes in Computer Science, 2012; Vol. 7548: 10-20.

Dyrrdal, AV; Lenkoski, A; Thorarinsdottir, TL; Stordal, F. Bayesian hierarchical modeling of extreme hourly precipitation in Norway. Environmetrics, 2014; doi: 10.1002/env.2301.

Eidsvik, J; Ellefmo, SL. The Value of Information in Mineral Exploration Within a Multi-Gaussian Framework. Mathematical Geosciences, 2013; Vol. 45(7): 777-798.

Feldmann, K; Scheuerer, M; Thorarinsdottir, TL. Spatial postprocessing of ensemble forecasts for temperature using nonhomogeneous Gaussian regression. In press, Monthly Weather Review.

Ferkingstad, E; Frigessi, A; Lyng, H. Indirect genomic effects on survival from gene expression data. Genome Biology, 2008; Vol. 9:R58.

Ferkingstad, E; Frigessi, A; Rue, H; Thorleifsson, G; Kong, A. Unsupervised empirical Bayesian multiple testing with external covariates. Annals of Applied Statistics, 2008; Vol. 2(2): 714-735.

Ferkingstad, E; Holden, L; Sandve, GK. Monte Carlo null models for genomic data. In press, Statistical Science.

Ferkingstad, Egil; Løland, Anders; Wilhelmsen, Mathilde. Causal modeling and inference for electricity markets. Energy Economics, 2011; Vol. 33(3): 404-412.

Fleischer, T; Frigessi, A; Johnson, KC; Edvardsen, H; Touleimat, N; Klajic, J; Riis, MLH; Haakensen, VD; Wärnberg, F; Naume, B; Helland, Å; Børresen-Dale, AL; Tost, J; Christensen, BC; Kristensen, VN. Genome-wide DNA methylation profiles in progression to in situ and invasive carcinoma of the breast with impact on gene transcription and prognosis. Genome Biology, 2014; Vol. 15(435).

Fong, Y; Rue, H; Wakefield, J. Bayesian inference for generalized linear mixed models. Biostatistics, 2010; Vol. 11(3): 397-412.

Friederichs, P; Thorarinsdottir, TL. Forecast verification for extreme value distributions with an application to probabilistic peak wind prediction. Environmetrics, 2012; Vol. 23(7): 579-594.

Frigessi, A; Løland, A; Pievatolo, A; Ruggeri, F. Statistical rehabilitation of improper correlation matrices. Quantitative Finance, 2011; Vol. 11(7): 1081-1090.

Frühwirth-Schnatter, S; Frühwirth, R; Held, L; Rue, H. Improved auxiliary mixture sampling for hierarchical models of non-gaussian data. Statistics and Computing, 2009; Vol. 19(4): 479-492.

Fuglstad, GA; Simpson, DP; Lindgren, FK; Rue, H. Exploring a new class of non-stationary spatial Gaussian random fields with varying local anisotropy. In press, Statistica Sinica. doi: 10.5705/ss.2013.106w.

Gómez-Rubio, V; Bivand, RS; Rue, H. Spatial models using Laplace approximation methods. Handbook of Regional Science, 2014; 1401-1417.

Grønneberg, S. The Copula Information Criterion and its implications for the Maximum Pseudo Likelihood Estimator. In DEPENDENCE MODELING: Handbook on Vine Copulae, D. Kurowicka and H. Joe (eds.), World Scientific Publishing Co., 2011.

Grønneberg, S; Hjort, NL. The Copula Information Criteria. Scandinavian Journal of Statistics, 2014; Vol. 41(2): 436-459.

Grønneberg, S; Hjort, NL. On the errors committed by sequences of estimator functionals. Mathematical Methods of Statistics, 2011; Vol. 20(4): 327- 346.

Gundersen, S; Kalas, M; Abul, O; Frigessi, A; Hovig, E; Sandve, GK. Identifying elemental genomic track types and representing them uniformly. BMC Bioinformatics, 2011; Vol. 12(1): 494-515.

Günther, CC; Tvete, IF; Aas, K; Hagen, JA; Kvifte, L; Borgan, Ø. Predicting Future Claims Among High Risk Policyholders Using Random Effects. D. Silvestrov and A. Martin-Löf (eds.): Modern Problems in Insurance Mathematics, EAA Series, Springer 2014; 171-185.

Günther, CC; Tvete, IF; Aas, K; Sandnes, GI; Borgan, Ø. Modelling and predicting customer churn from an insurance company. Scandinavian Actuarial Journal, 2014; Vol. 2014(1): 58-71.

Guttorp, P. Statistics and Climate. Annual Review of Statistics and Its Application, 2014; Vol. 1: 87-101.

Guttorp, P. Climate statistics and public policy. Statistics, Politics and Policy, 2012; Vol. 3(1). doi: 10.1515/2151-7509.1055.

Guttorp, P. The role of statisticians in international science policy. Environmetrics, 2011; Vol. 22(7): 817-825.

Guttorp, P; Craigmile, PF. Space-time modelling of trends in temperature series. Journal of Time Series Analysis, 2011; Vol. 32(4): 378-395.

Guttorp, P; Januzzi, A; Novak, M; Podschwit, H; Richardson, L; Sowder, CD; Zimmerman, A; Bolin, D; Särkkä, A. Assessing the uncertainty in projecting local mean sea level from global temperature. Journal of Applied Meteorology and Climatology, 2014; Vol. 53(9): 2163-2170.

Guttorp, P; Kim, TY. Uncertainty in Ranking the Hottest Years of U.S. Surface Temperatures. Journal of Climate, 2013; Vol. 26(17): 6323-6328.

Guttorp, P; Thorarinsdottir, TL. Bayesian Inference for Non-Markovian Point Processes. In: Advances and Challenges in Space-time Modelling of Natural Events, Lecture Notes in Statistics 2012; Vol. 207: 79-102.

Guttorp, P; Thorarinsdottir, TL. What Happened to Discrete Chaos, the Quenouille Process, and the Sharp Markov Property? Some History of Stochastic Point Processes. International Statistical Review, 2012; Vol. 80 (2): 253-268.

Guttorp, Peter; Xu, Jia. Climate change, trends in extremes, and model assessment for a long temperature time series from Sweden. Environmetrics, 2011; Vol. 22(3): 456-463.

Halle, CL; Andersen, EKF; Lando, M; Aarnes, EK; Hasvold, G; Holden, M; Syljuåsen, RG.; Sundfør, K; Kristensen, GSB; Holm, R; Malinen, E; Lyng, H. Hypoxia-induced gene expression in chemoradioresistant cervical cancer revealed by dynamic contrast-enhanced MRI. Cancer Research, 2012; Vol. 72 (20), 5285-5295.

Halle, CL; Lando, M; Svendsrud, DH; Clancy, T; Holden, M; Sundfør, K; Kristensen, GSB; Holm, R; Lyng, H. Membranous Expression of Ectodomain Isoforms of the Epidermal Growth Factor Receptor Predicts Outcome after Chemoradiotherapy of Lymph Node-Negative Cervical Cancer. Clinical Cancer Research, 2011; Vol. 17(16): 5501-5512.

Hammer, H; Kolbjørnsen, O; Tjelmeland, H; Buland, A. Lithology and fluid prediction from prestack seismic data using a Bayesian model with Markov process prior. Geophysical Prospecting, 2012; Vol. 60: 500–515.

Handel, AE; Sandve, GK; Disanto, G; Berlanga-Taylor, AJ; Gallone, G; Hanwell, H; Drabløs, F; Giovannoni, G; Ebers, GC; Ramagopalan, SV. Vitamin D receptor ChIP-seq in primary CD4+ cells: relationship to serum 25-hydroxyvitamin D levels and autoimmune disease. BMC Medicine, 2013; Vol. 11:163. doi:10.1186/1741-7015-11-163.

Handel, AE; Sandve, GK; Disanto, G; Handunnetthi, L; Giovannoni, G; Ramagopalan, SV. Integrating multiple oestrogen receptor alpha ChIP studies: overlap with disease susceptibility regions; DNase I hypersensitivity peaks and gene expression. BMC Medical Genomics, 2013; Vol. 6:45. doi:10.1186/1755-8794-6-45.

Hansen, LV; Thorarinsdottir, TL. A note on moving average models for Gaussian random fields. Statistics & Probability Letters, 2013; Vol. 83(3): 850-855.

Hansen, L; Thorarinsdottir, TL; Ovcharov, E; Gneiting, T; Richards, DP. Gaussian random particles with flexible Hausdorff dimension. In press, Advances in Applied Probability.

Haug, O; Dimakos, XK; Vårdal, JF; Aldrin, M. Climate change and its impact on building water damage. 38th ASTIN Colloquium, Manchester, UK, July 13-16, 2008.

Haug, O; Dimakos, XK; Vårdal, JF; Aldrin, M; Meze-Hausken, E. Future building water loss projections posed by climate change. Scandinavian Actuarial Journal, 2011; Vol. 2011(1): 1-20.

Haugen, M; Bray, F; Grotmol, T; Tretli, S; Aalen, OO; Moger, TA. Frailty modeling of bimodal age-incidence curves of nasopharyngeal carcinoma in low-risk populations. Biostatistics, 2009; Vol. 10(3): 501-514.

Haugen, M; Moger, TA. Frailty modelling of time-to-lapse of single policies for customers holding multiple car contracts. Scandinavian Actuarial Journal, 2014; doi: 10.1080/03461238.2014.963886.

Hinne, M; Lenkoski, A; Heskes, T; van Gerven, M. Efficient sampling of Gaussian graphical models using conditional Bayes factors. In press, Stat.

Hirst, D; Storvik, GO; Rognebakke, H; Aldrin, M; Aanes, S; Vølstad, JH. A Bayesian modelling framework for the estimation of catch-at-age of commercially harvested fish species. Canadian Journal of Fisheries and Aquatic Sciences, 2012; Vol. 69(12): 2064-2076.

Hjort, NL. Approximating Data. Journal of the Korean Statistical Society, 2008; Vol. 37/3,: 221-225.

Hjort, NL; McKeague, IW; Van Keilegom, I. Extending the Scope of Empirical Likelihood. Annals of Statistics, 2009; Vol. 37(3): 1079-1111.

Hjort, NL; Walker, SG. Quantile pyramids for Bayesian nonparametrics. Annals of Statistics, 2009; Vol. 37(1): 105-131.

Hobæk Haff, I. Comparison of estimators for pair-copula constructions. Journal of Multivariate Analysis, 2012; Vol. 110: 91-105.

Hobæk Haff, I. Parameter estimation for pair-copula constructions. Bernoulli, 2013; Vol. 19(2):462-491.

Hobæk Haff, I; Aas, K; Frigessi, A. On the simplified pair-copula construction — Simply useful or too simplistic? Journal of Multivariate Analysis, 2010; Vol. 101(5): 1296-1310.

Hobæk Haff, I; Frigessi, A; Maraun, D. How well do regional climate models simulate the spatial dependence of precipitation? An application of pair-copula constructions. Accepted for publication in Journal of Geophysical Research: Atmospheres.

Hobæk Haff, I; Segers, J. Nonparametric estimation of pair-copula constructions with the empirical pair-copula. Computational Statistics & Data Analysis, 2014; doi: 10.1016/j.csda.2014.10.020.

Holden, H; Holden, L. Optimal rebalancing of portfolios with transaction costs. Stochastics: An International Journal of Probability and Stochastic Processes, 2013; Vol. 85(3): 371-394.

Holden, H; Holden, L; Holden, S. Contract adjustment under uncertainty. Journal of Economic Dynamics & Control, 2010; Vol. 34: 657-680.

Holden, L; Hauge, R; Holden, M. Adaptive independent Metropolis–Hastings. Annals of Applied Probability, 2009; Vol. 19(1): 395-413.

Holden, L; Løland, A; Lindqvist, O. Valuation of long term, flexible gas contracts. Journal of Derivatives, 2011; Vol. 18(3): 75-85.

Holden, M; Deng, S; Wojnowski, L; Kulle, B. GSEA-SNP: applying gene set enrichment analysis to SNP data from genome-wide association studies. Bioinformatics, 2008; Vol. 24(23): 2784-2785.

Husberg, C; Nygård, S; Finsen, AV; Damås, JK; Frigessi, A; Øie, E; Waehre, A; Gullestad, L; Aukrust, P; Yndestad, A; Christensen, G. Cytokine expression profiling of the myocardium reveals a role for CX3CL1 (fractalkine) in heart failure. Journal of Molecular and Cellular Cardiology, 2008; Vol. 45(2): 261-269.

Huseby AB, Haavardsson NF. Multi-reservoir production optimization. European Journal of Operational Research, 2009; Vol. 199(1): 236-251.

Jansen, PA; Kristoffersen, AB; Viljugrein, H; Jimenez, DD; Stien, A; Aldrin, M. Sea lice as a density-dependent constraint to salmonid farming. Proceedings of the Royal Society of London. Biological Sciences, 2012; Vol. 27(1737): 2330-2338.

Jemtland, R; Holden, M; Reppe, S; Olstad, OK; Reinholt, FP; Gautvik, VT; Refvem, H; Frigessi, A; Houston, B; Gautvik, KM. Molecular disease map of bone characterizing the postmenopausal osteoporosis phenotype. Journal of Bone and Mineral Research, 2011; Vol. 26(8): 1793-1801.

Katz, RW; Craigmile, PF; Guttorp, P; Haran, M; Sansó, B; Stein, ML. Uncertainty analysis in climate change assessments. Nature Climate Change, 2013; Vol. 3: 769-771.

Kjønsberg, H; Hauge, R; Kolbjørnsen, O; Buland, A. Seismic pre-drill prospect risk and reservoir quality assessment. Geophysics, 2010; Vol. 75(2): O9-O19.

Kolbjørnsen, O; Hauge, R; Drange-Espeland, M; Buland, A. Model-based fluid factor for controlled source electromagnetic data. Geophysics, 2012; Vol. 77(1): E21-E31.

Kresse, SH; Rydbeck, H; Skårn, M; Namløs, HM; Barragan-Polania, AH; et al. Integrative Analysis Reveals Relationships of Genetic and Epigenetic Alterations in Osteosarcoma. PLoS ONE, 2012; Vol. 7 (11): e48262. doi: 10.1371/journal.pone.0048262.

Kristensen, P; Aalen, OO. Understanding mechanisms: opening the "black box" in observational studies. Scandinavian Journal of Work, Environment & Health, 2013; Vol. 39(2): 121-124.

Kristensen, VN; Lingjærde, OC; Russnes, HEG; Vollan, HKM; Frigessi, A; Børresen-Dale, AL. Principles and methods of integrative genomic analyses in cancer. Nature Reviews Cancer, 2014; Vol. 14(5): 299-313.

Kuijjer, ML; Rydbeck, H; Kresse, SH; Buddingh, EP; Lid, AB; Roelofs, H; Bürger, H; Myklebost, O; Hogendoorn, PCW; Meza-Zepeda, LA; Cleton-Jansen, AM. Identification of osteosarcoma driver genes by integrative analysis of copy number and gene expression data. Genes Chromosomes Cancer, 2012; Vol. 51: 696–706.

Kulle, B; Frigessi, A; Edvardsen, H; Kristensen, V; Wojnowski, L. Accounting for haplotype phase uncertainty in LD estimation, Genet Epidemiol. 2008 February; 32(2): 168-78. Erratum in: Genet Epidemiol., 2008; Vol. 32(6): 586-587.

Lando, M; Holden, M; Bergersen, LC; Svendsrud, DH; Stokke, T; Sundfør, K; Glad, IK; Kristensen, GB; Lyng, H. Gene Dosage, Expression, and Ontology Analysis Identifies Driver Genes in the Carcinogenesis and Chemoradio-resistance of Cervical Cancer. PLoS Genetics, 2009; Vol. 5(11): e1000719. doi: 10.1371/journal.pgen.1000719.

Lando, M; Wilting, SM; Snipstad, K; Clancy, T; Bierkens, M; Aarnes, EK; Holden, M; Stokke, T; Sundfør, K; Holm, R; Kristensen, GSB; Steenbergen, RDM; Lyng, H. Identification of eight candidate target genes of the recurrent 3p12p14 loss in cervical cancer by integrative genomic profiling. Journal of Pathology, 2013; Vol. 230(1): 59-69.

Lavine, JS; Bjørnstad, ON; de Blasio, BF; Storsaeter, J. Short-lived immunity against pertussis, age-specific routes of transmission, and the utility of a teenage booster vaccine. Vaccine, 2012; Vol. 30(3), 544-551.

Lawyer, G; Ferkingstad, E; Nesvåg, R; Varnäs, K; Agartz, I. Local and covariate-modulated false discovery rates applied in neuroimaging. NeuroImage, 2009; Vol. 47(1): 213-219.

Lawyer, G; Nesvag, R; Varnas, K; Frigessi, A; Agartz, I. Investigating possible subtypes of schizophrenia patients and controls based on brain cortical thickness. Psychiatry Research, 2008; Vol. 164(3): 254-264.

LeBlanc, M; Kulle, B; Sundet, KS; Agartz, I; Melle, I; Djurovic, S; Frigessi, A; Andreassen, OA. Genome-wide study identifies PTPRO and WDR72 and FOXQ1-SUMO1P1 interaction associated with neurocognitive function. Journal of Psychiatric Research, 2012; Vol. 46(2): 271-278.

Lenkoski, A. A direct sampler for G-Wishart variates. Stat, 2013; Vol. 2(1): 119-128.

Lenkoski, A; Eicher, TS; Raftery, AE. Two-Stage Bayesian Model Averaging in the Endogenous Variable Model. Econometric Reviews, 2014; Vol. 33 (1-4): 122-151.

Lerch, S; Thorarinsdottir, TL. Comparison of nonhomogeneous regression models for probabilistic wind speed forecasting. Tellus A, 2013; Vol. 65: 21206.

Li, Y; Brown, P; Rue, H; al-Maini, M; Fortin, P. Spatial modelling of lupus incidence over 40 years with changes in census areas. Journal of the Royal Statistical Society: Series C (Applied Statistics), 2012; Vol. 61: 99–115.

Lindgren, F; Martins, T; Rue, H; Simpson, D. Discussion on "Spatial prediction in the presence of positional error". Environmetrics, 2011; Vol. 22(2): 127.

Lindgren, F; Rue, H. Bayesian spatial modelling with R-INLA. In press, Journal of Statistical Software.

Lindgren, F; Rue, H; Lindström, J. An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. Journal of the Royal Statistical Society, Series B (Statistical Methodology), 2011; Vol. 73(4): 423-498.

Lindgren, G; Lindgren, F. Stochastic Asymmetry Properties of 3D Gauss-Lagrange Ocean Waves with Directional Spreading. Stochastic Models, 2011; Vol. 27(3): 490-520.

Løland, A; Dimakos, XK. Modelling Nord Pool's NO1 area price. Journal of Energy Markets, 2010; Vol. 3(1): 73-92.

Løland, A; Ferkingstad, E; Wilhelmsen, M. Forecasting transmission congestion. Journal of Energy Markets, 2012; Vol. 5(3): 65-83.

Løland, A; Huseby, RB; Hjort, NL; Frigessi, A. Statistical corrections of invalid correlation matrices. Scandinavian Journal of Statistics, 2013; Vol. 40(4): 807-824.

Ma, Y; Guttorp, P. Estimating daily mean temperature from synoptic climate observations. International Journal of Climatology, 2013; Vol. 33(5): 1264-1269.

Marques, R; Storvik, G. Reweighting Schemes Based on Particle Methods. Ettore Lanzarone and Francesca Ieva (eds.): The Contribution of Young Researchers to Bayesian Statistics, Springer Proceedings in Mathematics & Statistics, 2014; Vol. 63: 73-76.

Martinelli, G; Eidsvik, J. Dynamic exploration designs for graphical models using clustering with applications to petroleum exploration. Knowledge-Based Systems, 2014; Vol. 58: 113-126.

Martinelli, G; Eidsvik, J; Hauge, R. Dynamic decision making for graphical models applied to oil exploration. European Journal of Operational Research, 2013; Vol. 230(3): 688-702.

Martinelli, G; Eidsvik, J; Hauge, R; Førland, MD. Bayesian networks for prospect analysis in the North Sea. The American Association of Petroleum Geologists, 2011; Vol. 95(8): 1423-1442.

Martinelli, G; Eidsvik, J; Hokstad, K; Hauge, R. Strategies for petroleum exploration on the basis of Bayesian Networks: a case study. SPE journal, 2014; Vol. 19(4): 564-575.

Martinelli, G; Eidsvik, J; Sinding-Larsen, R; Rekstad, S; Mukerji, T. Building Bayesian networks from basin-modelling scenarios for improved geological decision making. Petroleum Geoscience, 2013; Vol. 19(3): 289-304.

Martino, S; Aas, K; Lindqvist, O; Neef, LR; Rue, H. Estimating Stochastic Volatility Models Using Integrated Nested Laplace Approximations. The European Journal of Finance, 2011; Vol. 17(7): 487-503.

Martino, S; Akerkar, R; Rue, H. Approximate Bayesian Inference for Survival Models. Scandinavian Journal of Statistics, 2011; Vol. 38(3): 514-528.

Martino, S; Rue, H. Case Studies in Bayesian Computation using INLA in "Complex data modeling and computationally intensive statistical methods". Secchi P. and Mantovan P. (editors), Contributions to Statistics 2010, pp 99-114.

Martins, TG; Rue, H. Extending Integrated Nested Laplace Approximation to a Class of Near-Gaussian Latent Models. Scandinavian Journal of Statistics, 2014; doi: 10.1111/sjos.12073.

Martins, TG; Simpson, DP; Illian, J; Rue, H; Geirsson, O. Discussion of "Beyond mean regression". Statistical modelling, 2013; Vol. 13(4): 355-361.

Martins, TG; Simpson, DP; Lindgren, FK; Rue, H. Bayesian computing with INLA: New features. Computational Statistics & Data Analysis, 2013; Vol. 67: 68-83.

Moger, TA; Haugen, M; Yip, BHK; Gjessing, HK; Borgan, Ø. A hierarchical frailty model applied to two-generation melanoma data. Lifetime Data Analysis, 2011; Vol. 17(3): 445-460.

Molyneux, SD; Waterhouse, PD; Shelton, D; Shao, YW; Watling, CM; Tang, QL; Harris, IS; Dickson, BC; Tharmapalan, P; Sandve, GK; Zhang, X; Bailey, SD; Berman, H; Wunder, JS; Iszvak, Z; Lupien, M; Mak, TW; Khokha, R. Human somatic cell mutagenesis creates genetically tractable sarcomas. Nature Genetics, 2014; Vol. 46(9): 964-972.

Moxnes, JF; de Blasio, BF; Leegaard, TM; Moen, AEF. Methicillin-Resistant Staphylococcus aureus (MRSA) Is Increasing in Norway: A Time Series Analysis of Reported MRSA and Methicillin-Sensitive S. aureus Cases, 1997–2010. PloS ONE, 2013; Vol. 8(8): e70499. doi: 10.1371/journal.pone.0070499.

Myhre, S; Mohammed, H; Tramm, T; Alsner, J; Finak, G; Park, M; Overgaard, J; Børresen-Dale, AL; Frigessi, A; Sørlie, T. In Silico Ascription of Gene Expression Differences to Tumor and Stromal Cells in a Model to Study Impact on Breast Cancer Outcome. PLoS ONE, 2010; Vol. 5(11): e14002. doi: 10.1371/journal.pone.0014002.

Möller, A; Lenkoski, A; Thorarinsdottir, TL. Multivariate probabilistic forecasting using Bayesian model averaging and copulas. Quarterly Journal of the Royal Meteorological Society, 2013; Vol. 139(673): 982-991.

Nair, KN; Kolbjørnsen, O; Skorstad, A. Seismic inversion and its applications in reservoir characterization. First Break, 2012; Vol. 30(3): 83-86.

Nesvåg, R; Frigessi, A; Jönsson, EG; Agartz, I. Effects of alcohol consumption and antipsychotic medication on brain morphology in schizophrenia. Schizophrenia Research, 2007; Vol. 90(1-3): 52-61.

Nesvåg, R; Saetre, P; Lawyer, G; Jönsson, EG; Agartz, I. The relationship between symptom severity and regional cortical and grey matter volumes in schizophrenia. Progress in Neuro-Psychopharmacology and Biological Psychiatry, 2009; Vol. 33(3): 482-490.

Nordskag, J; Kjøsnes, Ø; Nguyen, A; Hokstad, K. CSEM exploration in the Barents Sea: Joint CSEM & seismic interpretation. SEG Technical Program Expanded Abstracts, 2013; pp. 810-814. doi: 10.1190/segam2013-0883.1.

Nygaard, V; Liu, F; Holden, M; Kuo, WP; Trimarchi, J; Ohno-Machado, L; Cepko, CL; Frigessi, A; Glad, IK; van de Wiel, M; Hovig, E; Lyng, H. Validation of oligo-arrays for quantitative exploration of the transcriptome. BMC Genomics, 2008; Vol. 9:258. doi:10.1186/1471-2164-9-258.

Nygård, S; Borgan, Ø; Lingjærde, OC; Størvold, HL. Partial least squares Cox regression for genome-wide data. Lifetime Data Analysis, 2008; Vol. 14(2): 179-195.

Omair, A; Holden, M; Lie, BA; Reikerås, O; Brox, JI. Treatment outcome of chronic low back pain and radiographic lumbar disc degeneration are associated with inflammatory and matrix degrading gene variants: a prospective genetic association study. BMC Musculoskeletal Disorders, 2013; Vol. 14:105. doi:10.1186/1471-2474-14-105.

Orskaug, Elisabeth; Scheel, Ida; Frigessi, Arnoldo; Guttorp, Peter; Haugen, Jan Erik; Tveito, Ole Einar; Haug, Ola. Evaluation of a dynamic downscaling of precipitation over the Norwegian mainland. Tellus A, 2011; Vol. 63(4): 746-756.

Otnaess, MK; Djurovic, S; Rimol, LM; Kulle, B; Kähler, AK; Jönsson, EG; Agartz, I; Sundet, K; Hall, H; Timm, S; Hansen, T; Callicott, JH; Melle, I; Werge, T; Andreassen, OA. Evidence for a possible association of neurotrophin receptor (NTRK-3) gene polymorphisms with hippocampal function and schizophrenia. Neurobiology of Disease, 2009; Vol. 34(3): 518-524.

Paul, M; Riebler, A; Bachmann, LM; Rue, H; Held, L. Bayesian bivariate meta-analysis of diagnostic test studies using integrated nested Laplace approximations. Statistics in Medicine, 2010; Vol. 29(12): 1325-1339.

Paulsen, J; Lien, T; Sandve, GK; Gundersen, S; Trengereid, K; Glad, I; Borgan, Ø; Hovig, E. Handling realistic assumptions in hypothesis testing of 3D co-localization of genomic elements. Nucleic Acids Research, 2013; Vol. 41(10): 5164-5174.

Paulsen, J; Rødland, EA; Holden, L; Holden, M; Hovig, E. A statistical model of ChIA-PET data for accurate detection of chromatin 3D interactions. Nucleic Acids Research, 2014; Vol. 42(18): e143. doi: 10.1093/nar/gku738.

Paulsen, J; Sandve, GK; Gundersen, S; Lien, TG; Trengereid, K; Hovig, E. HiBrowse: Multi-purpose statistical analysis of genome-wide chromatin 3D organization. Bioinformatics, 2014; Vol. 30(11): 1620-1622.

Pitzer, VE; Atkins, KE; de Blasio, BF; Van Effelterre, T; Atchison, CJ; et al. Direct and Indirect Effects of Rotavirus Vaccination: Comparing Predictions from Transmission Dynamic Models. PLoS ONE, 2012; Vol. 7(8): e42320. doi: 10.1371/journal.pone.0042320.

Reitan, T; Aas, K. A New Robust Importance Sampling Method for measuring VaR and ES allocations for Credit Portfolios. Journal of Credit Risk, 2010/2011; Vol 6(4): 1-37.

Reppe, S; Refvem, H; Gautvik, VT; Olstad, OK; Høvring, PI; Reinholt, FP; Holden, M; Frigessi, A; Jemtland, R; Gautvik, KM. Eight genes are highly associated with BMD variation in post-menopausal Caucasian women. Bone, 2010; Vol. 46(3): 604-612.

Riebler, A; Held, L; Rue, H. Estimation and extrapolation of time trends in registry data — Borrowing strength from related populations. Annals of Applied Statistics, 2012; Vol. 6(1): 304-333.

Riebler, A; Held, L; Rue, H; Bopp, M. Gender-specific differences and the impact of family integration on time trends in age-stratified Swiss suicide rates. Journal of the Royal Statistical Society: Series A (Statistics in Society), 2012; Vol. 175: 473–490.

Roos, M; Martins, TG; Held, L; Rue, H. Sensitivity Analysis for Bayesian Hierarchical Models. Bayesian Analysis, 2014; doi: 10.1214/14-BA909.

Rue, H; Martino, S. Approximate Bayesian inference for hierarchical Gaussian Markov random fields models. Journal of Statistical Planning and Inference (Special Issue: Bayesian Inference for Stochastic Processes.), 2007; Vol. 137(10): 3177-3192.

Rue, H; Martino, S; Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. Journal of the Royal Statistical Society Series B (Statistical Methodolgy), 2009; Vol. 71(2): 319-392.

Rye, M; Sandve, GK; Daub, CO; Kawaji, H; Carninci, P; Forrest, AR; Drabløs, F; FANTOM consortium. Chromatin states reveal functional associations for globally defined transcription start sites in four human cell lines. BMC Genomics, 2014; Vol. 15:120. doi: 10.1186/1471-2164-15-120.

Røysland, K. A martingale approach to continuous-time marginal structural models. Bernoulli, 2011; Vol. 17(3): 895-915.

Røysland, K; Gran, JM; Ledergerber, B; von Wyl, V; Young, J; Aalen, OO. Analyzing direct and indirect effects of treatment using dynamic path analysis applied to data from the Swiss HIV Cohort Study. Statistics in Medicine, 2011; Vol. 30(24): 2947-2958.

Sandve, GK; Ferkingstad, E; Nygård, S. Sequential Monte Carlo multiple testing. Bioinformatics, 2011; Vol. 27(23): 3235-3241.

Sandve, GK; Gundersen, S; Johansen, M; Glad, IK; Gunathasan, K; Holden, L; Holden, M; Liestøl, K; Nygård, S; Nygaard, V; Paulsen, J; Rydbeck, H; Trengereid, K; Clancy, T; Drabløs, F; Ferkingstad, E; Kalas, M; Lien, T; Rye, MB; Frigessi, A; Hovig, E. The Genomic HyperBrowser: an analysis web server for genome-scale data. Nucleic Acids Research, 2013; Vol. 41(W1): W133-W141.

Sandve, GK; Gundersen, S; Rydbeck, H; Glad, IK; Holden, L; Holden, M; Liestøl, K; Clancy, T; Drablos, F; Ferkingstad, E; Johansen, M; Nygaard, V; Tøstesen, E; Frigessi, A; Hovig, E. The differential disease regulome. BMC Genomics, 2011; Vol. 12:353. doi:10.1186/1471-2164-12-353.

Sandve, GK; Gundersen, S; Rydbeck, H; Glad, IK; Holden, L; Holden, M; Liestøl, K; Clancy, T; Ferkingstad, E; Johansen, M; Nygaard, V; Tøstesen, E; Frigessi, A; Hovig, E. The Genomic HyperBrowser: inferential genomics at the sequence level. Genome Biology, 2010; Vol. 11(12): R121. doi: 10.1186/gb-2010-11-12-r121.

Sandve, GK; Nekrutenko, A; Taylor, J; Hovig, E. Ten Simple Rules for Reproducible Computational Research. PLoS Computational Biology, 2013; Vol. 9(10) : e1003285. doi:10.1371/journal.pcbi.1003285.

Scheel, I; Aldrin, M; Frigessi, A; Jansen, PA. A stochastic model for infectious salmon anaemia (ISA) in Atlantic salmon farming. Journal of the Royal Society Interface, 2007; Vol. 4: 699-706.

Scheel, I; Ferkingstad, E; Frigessi, A; Haug, O; Hinnerichsen, M; Meze-Hausken, E. A Bayesian hierarchical model with spatial variable selection: the effect of weather on insurance claims. Journal of the Royal Statistical Society: Series C (Applied Statistics), 2013; Vol. 62(1): 85-100.

Scheel, I; Green, PJ; Rougier, JC. A Graphical Diagnostic for Identifying Influential Model Choices in Bayesian Hierarchical Models. Scandinavian Journal of Statistics, 2011; Vol. 38(3): 529-550.

Scheel, I; Hinnerichsen, M. The Impact of Climate Change on Precipitation-related Insurance Risk: A Study of the Effect of Future Scenarios on Residential Buildings in Norway. The Geneva Papers, 2012; Vol. 37: 365–376.

Schefzik, R; Thorarinsdottir, TL; Gneiting, Tilmann. Uncertainty quantification in complex simulation models using ensemble copula coupling. Statistical Science, 2013; Vol. 28(4): 616-640.

Schweder, T; Hjort, NL. Confidence Distribution, the Frequentist Distribution Estimator of a Parameter: A Review Discussions. International Statistical Review, 2013; Vol. 81(1): 56-68.

Solvang, H; Lingjærde, O; Frigessi, A; Børresen-Dale, A; Kristensen, V. Linear and non-linear dependencies between copy number aberrations and mRNA expression reveal distinct molecular pathways in breast cancer. BMC Bioinformatics, 2011; Vol. 12:197. doi:10.1186/1471-2105-12-197.

Stenevik, E; Vølstad, JH; Høines, Å; Aanes, S; Óskarsson, G; Jacobsen, J; Tangen, Ø. Precision in estimates of density and biomass of Norwegian spring spawning herring based on acoustic surveys. In press, Marine Biology Research.

Storvik, B; Storvik, G, Fjørtoft, R. On the combination of correlated images using meta-Gaussian distributions. IEEE Transactions on Geoscience and Remote Sensing, 2009; Vol. 47(7): 2372-2379.

Storvik, G. On the flexibility of Metropolis-Hastings acceptance probabilities in auxiliary variable proposal generation. Scandinavian Journal of Statistics, 2011; Vol. 38(2): 342-358.

Sørbye, SH; Hindberg, K; Olsen, LR; Rue, H. Bayesian multiscale feature detection of log-spectral densities. Computational Statistics & Data Analysis; 2009; Vol. 53(11): 3746-3754.

Sørbye, SH; Rue, H. Scaling intrinsic Gaussian Markov random field priors in spatial modelling. Spatial Statistics, 2013; Vol. 8: 39-51.

Sørensen, Ø; Frigessi, A; Thoresen, M. Measurement error in Lasso: impact and likelihood bias correction. In press, Statistica Sinica. doi: 10.5705/ss.2013.180.

Tharmaratnam, K; Claeskens, G. A comparison of robust versions of the AIC based on M-, S- and MM-estimators. Statistics: A Journal of Theoretical and Applied Statistics, 2013; Vol. 47(1): 216-235.

Thon, K; Rue, H; Skrøvseth, SO; Godtliebsen, F. Bayesian multiscale analysis of images modeled as Gaussian Markov random fields. Computational Statistics & Data Analysis, 2012; Vol. 56(1): 49-61.

Thorarinsdottir, TL. Calibration diagnostics for point process models via the probability integral transform. Stat, 2013; Vol. 2(1): 150-158.

Thorarinsdottir, TL; Gneiting, T; Gissibl, N. Using Proper Divergence Functions to Evaluate Climate Models. SIAM/ASA Journal of Uncertainty Quantification, 2013; Vol. 1(1): 522-534.

Thorarinsdottir, TL; Løland, A. Comments on: Space-time wind speed forecasting for improved power system dispatch. TEST, 2014; Vol. 23(1): 32-33.

Thorarinsdottir, TL; Scheuerer, M; Feldmann, K. Statistical post-processing of ensemble forecasts. Promet, 2012; Vol. 37(3/4): 43-52.

Thorarinsdottir, TL; Scheuerer, M; Heinz, C. Assessing the calibration of high-dimensional ensemble forecasts using rank histograms. Journal of Computational and Graphical Statistics, 2015; doi: 10.1080/10618600.2014.977447.

Thorarinsdottir, TL; Sillmann, J; Benestad, R. Studying Statistical Methodology in Climate Research. EOS, Transactions American Geophysical Union, 2014; Vol. 95(15): 129.

Tramm, T; Mohammed, H; Myhre, S; Kyndi, M; Alsner, J; Børresen-Dale, AL; Sørlie, T; Frigessi, A; Overgaard, J. Development and validation of a gene profile predicting benefit of post-mastectomy radiotherapy in patients with high-risk breast cancer: A study of gene expression in the DBCG82bc cohort. Clinical Cancer Research, 2014; Vol. 20(20): 5272-5280.

Tvete, IF; Olsen, IC; Fagerland, M; Meland, N; Aldrin, M; Smerud, KT; Holden, L. An approach to combining parallel and cross-over trials with and without run-in periods using individual patient data. Clinical Trials, 2012; Vol. 9 (2): 164-175.

Tøstesen, E; Sandve, GK; Liu, F; Hovig, E. Segmentation of DNA sequences into twostate regions and melting fork regions. J Phys Condens Matter, 2009; Vol. 21(3):034109. doi: 10.1088/0953-8984/21/3/034109.

van de Wiel, MA; Leday, GGR; Pardo, L.; Rue, H; Van Der Vaart, AW; van Wieringen, WN. Bayesian analysis of RNA sequencing data by estimating multiple shrinkage priors. Biostatistics, 2013; Vol. 14(1): 113-128.

Villani, A; Frigessi, A; Liljeros, F; Nordvik, MK; de Blasio, BF. A Characterization of Internet Dating Network Structures among Nordic Men Who Have Sex with Men. PLoS ONE, 2012; Vol. 7(7): e39717. doi: 10.1371/journal.pone.0039717.

Vøllestad, LA; Hirst, D; L'Abee-Lund, JH; Armstrong, JD; MacLean, JC; Youngson, AF; Stenseth, NC. Divergent trends in anadromous salmonid populations in Norwegian and Scottish rivers. Proceedings of the Royal Society B, 2009; Vol. 276: 1021-1027.

Watson, CT; Disanto, G; Sandve, GK; Breden, F; Giovannoni, G; et al. Age-Associated Hyper-Methylated Regions in the Human Brain Overlap with Bivalent Chromatin Domains. PLoS ONE, 2012; Vol. 7(9): e43840. doi: 10.1371/journal.pone.0043840.

Wilhelmsen, M; Dimakos, XK; Husebø, TA; Fiskaaen, M. Bayesian modelling of credit risk using Integrated Nested Laplace Approximations. In: Rethinking Risk Management and Reporting: Measurement, Uncertainty, Bayesian Analysis and Expert Judgement, Risk Books, London, 2010.

Xue, Y; Kristiansen, IS; de Blasio, BF. Dynamic modelling of costs and health consequences of school closure during an influenza pandemic. BMC Public Health, 2012; Vol. 12:962. doi:10.1186/1471-2458-12-962.

Xue, Y; Kristiansen, IS; de Blasio, BF. Modeling the cost of influenza: the impact of missing costs of unreported complications and sick leave. BMC Public Health, 2010; Vol. 10:724. doi: 10.1186/1471-2458-10-724.

Young, LJ; Piegorsch, WW; Guttorp, P. In memory of George Casella. Environmetrics, 2013; Vol. 24(5): 279-280.

Yue, YR; Simpson, D; Lindgren, F; Rue, H. Bayesian Adaptive Smoothing Splines Using Stochastic Differential Equations. Bayesian Analysis, 2014; Vol. 9(2): 397-424.

Zechmeister, I; de Blasio, BF; Garnett, G. HPV-vaccination for the prevention of cervical cancer in Austria: a model based long-term prognosis of cancer epidemiology. Journal of Public Health, 2010; Vol. 18: 3–13.

Zhao, X; Rødland, EA; Sørlie, T; Naume, B; Langerød, A; Frigessi, A; Kristensen, V; Børresen-Dale, AL; Lingjærde, OC. Combining Gene Signatures Improves Prediction of Breast Cancer Survival. PLoS ONE, 2011; Vol. 6(3): e17845. doi: 10.1371/journal.pone.0017845.

Technical reports

Aas, K. Validering av modell for finansielle indikatorer. NR-Note, SAMBA/25/09, 2009.

Aas, K. Sammenligning av metoder for risikoaggregering. NR-Note, SAMBA/06/11, 2011.

Aas, K; Hobæk Haff, I. DNB Total Risk Model Version 4: Technical report. NR-Note, SAMBA/04/12, 2012.

Aas, K; Holden, M. Importance Sampling from DNB's Credit Risk Model. NR-Note, SAMBA/31/12, 2012.

Aas, K; Low, RKL. Portfolio optimization using pair-copula constructions. NR-note SAMBA/62/12, 2012.

Almendral Vazquez, A; Kolbjørnsen, O. EM inversion software, USER MANUAL, NR-Note, SAND/11/08, 2008.

Dimakos, XK; Løland, A. Modelling the NO1 area price. NR-Note, SAMBA/44/08, 2008.

Dimakos, XK; Storvik, B. Priselastisitet i forsikring, NR-Note SAMBA/36/07, 2007.

Dimakos, XK; Storvik, B; Vårdal, J. Kundelojalitet i PVK/BIL NL. NR-Note, SAMBA/51/08, 2009.

Eikvil, L; Holden, M. Word sense disambiguation and co-occurrence ranking, Applied to PubMed abstracts. NR-Note, SAMBA/46/09, 2009.

Eikvil, L; Holden, M. Efficient search for biomedical information. NR-Note, SAMBA/54/11, 2011.

Ferkingstad, E; Holden, L; Sandve, GK. Monte Carlo null models in ecology. Norwegian Computing Center, NR-Note, SAMBA/20/13, 2013.

Ferkingstad, E; Wilhelmsen, M; Løland, A. Classification model for the net capacity utilisation. NR-Note, SAMBA/61/10, 2010.

Günther, CC; Tvete, IF; Aas, K; Borgan, Ø. Modellering og prediksjon av kundeavgang. NR-Note, SAMBA 10/10, 2010.

Haug, O; Dimakos, XK; Vårdal, JF; Aldrin, M. Fremtidig utvikling i vannskader som følge av klimaendringer. NR-Note, SAMBA/12/08, 2008.

Hobæk Haff, I. Parameter estimation for pair-copula constructions, NR-Note, SAMBA/36/10, 2010.

Hobæk Haff, I; Vårdal, J; Aas, K. Modellering av finansielle indikatorer. NR-Note, SAMBA/48/08, 2008.

Hobæk Haff, I; Wilhelmsen, M; Dimakos, XK: Uførhet. NR-Note, SAMBA/01/08, 2008.

Holden, L; Haug, O. A Multidimensional Mixture Model for Unsupervised Tail Estimation, NR-Note SAMBA/09/09, 2009.

Holden, M; Aas, K. Parallellisering av Totalrisiko-programmet. NR-Note, SAMBA/43/10, 2010.

Holden, M; Aas, K. Parallellisering av Totalrisiko-programmet ved hjelp av GPU. NR-Note SAMBA13/11, 2011.

Holden, M; Aas, K. Parallellisering av simulering fra vines. NR-Note, SAMBA/18/11, 2011.

Holden, M; Glad, IK; Hauge, HH; Hovig, E; Liestøl, K. Image restoration and analysis of biomolecular radioactivity images II. NR-Note, SAMBA/11/12, 2012.

Holden, M; Glad, IK; Hauge, HH; Liestøl, K. Image restoration and analysis of biomolecular radioactivity images, NR-Note, SAMBA/27/08, 2008.

Kolbjørnsen, O; Almendral Vazquez, A; Buland, A. Bayesian common midpoint inversion for controlled source electromagnetics, NR-Note, SAND/12/08, 2008.

Kolbjørnsen, O; Hauge, R; Buland, A. Including facies geometry in lithology and fluid prediction. NR-Note, SAND/12/2011, 2011.

Løland, A. Justering til gyldig korrelasjonsmatrise. NR-Note, SAMBA/11/08, 2008.

Løland, A; Lindqvist, O. Valuation of commodity-based swing options: a survey. NR-Note, SAMBA/38/08, 2008.

Orskaug, E: Storskader. NR-Note, SAMBA/59/10, 2010.

Orskaug, E; Haug, O. Skadeprediksjoner basert på ECHAM4 klimamodelldata. NR-Note, SAMBA/29/09, 2009.

Orskaug, E; Scheel, I, Frigessi, A; Guttorp, P; Haugen, JE; Tveito, OE; Haug, O. Supplemental material to: Evaluation of a dynamic downscaling of precipitation over the Norwegian mainland. NR-Note, SAMBA/50/10, 2010.

Rognebakke, H; Hirst, D. Catch-at-age — Version 2.0: New features and validation of program. NR-Note, SAMBA/57/11, 2011.

Rognebakke, H; Hirst, D; Storvik, GO. Catch-at-age — Version 2.0: Technical Report. NR-Note, SAMBA/44/11, 2011.

Rognebakke, H; Hirst, D; Storvik, GO; Aldrin, M. Catch-at-age for multiple stocks: Modelling Skrei and Coastal Cod simultaneously. NR-Note, SAMBA/46/11, 2011.

Scheel, I; Hinnerichsen, M. The Impact of Climate Change on Insurance Risk: A Study of the Effect of Climate Change Scenarios in Norway. NR-Note, SAMBA/52/10, 2010.

Stien, M; Kolbjørnsen, O. D-vine Creation of Non-Gaussian Random Fields. NR-Note, SAND/07/08, 2008.

Storvik, B; Løland, A. Negative prices in the electricity market. NR-Note, SAMBA/08/10, 2010.

Tvete, IF; Løland, A. Oil prices and exchange rates: A survey. NR-Note, SAMBA/47/07, 2007.

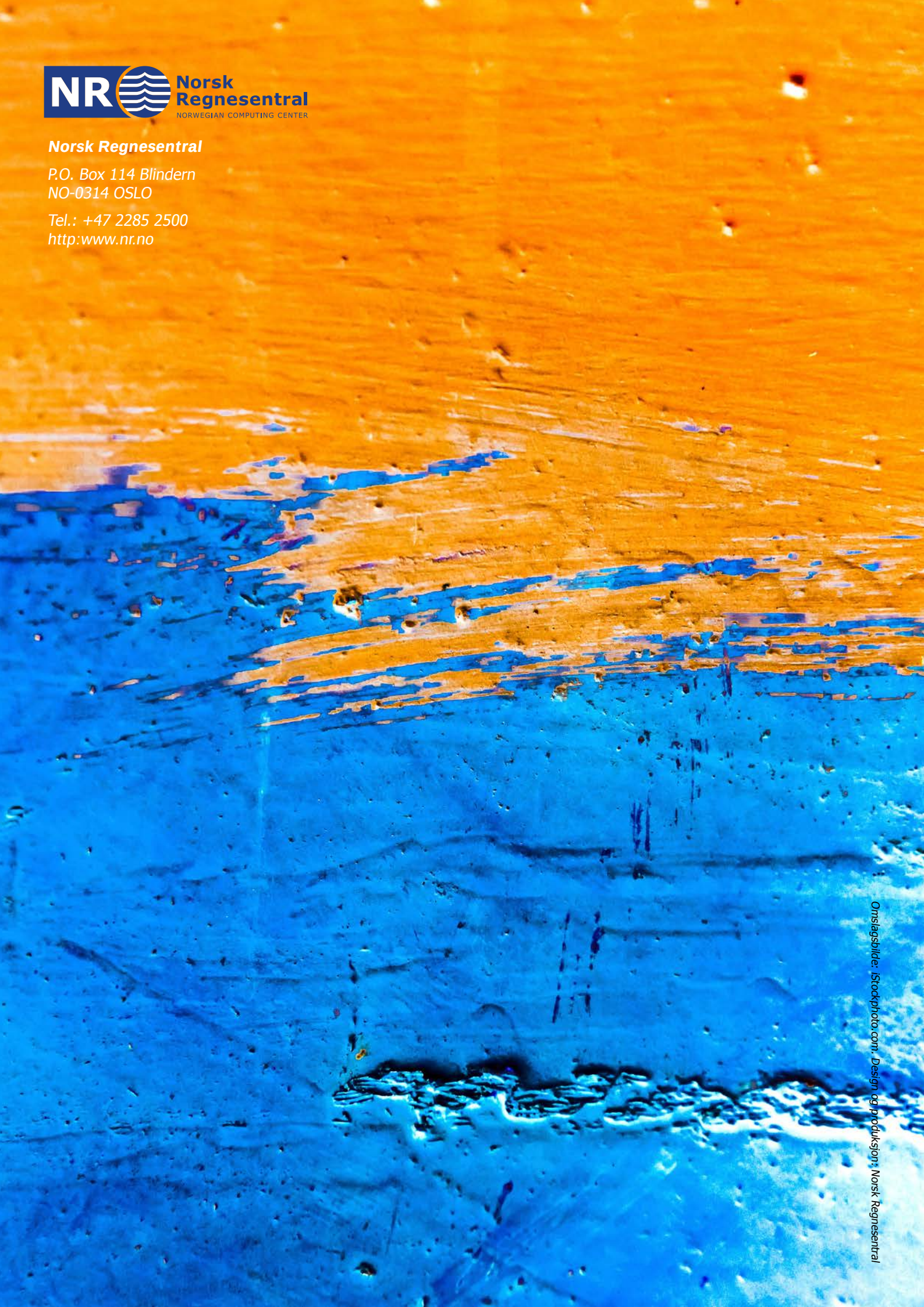
Vårdal, JF; Haug, O; Dimakos, XK; Aldrin, M. Fremtidig utvikling i vannskader som følge av klimaendringer — resultater. NR-Note, SAMBA/13/08, 2008.

Wilhelmsen, M; Dimakos, XK. En studie av sammenhengen mellom makroøkonomi og forsikring. NR-Note, SAMBA/41/07, 2007.

Books

Abrahamsen, P; Kolbjørnsen, O; Hauge, R (eds). Geostatistics Oslo 2012. Springer. (ISBN: 978-94-007-4152-2). pp 573. 2012.

Claeskens, G; Hjort, NL. Model Selection and Model Averaging. Cambridge University Press, (ISBN: 978-05-218-5225-8) 2008.



Norsk Regnesentral

P.O. Box 114 Blindern
NO-0314 OSLO

Tel.: +47 2285 2500
http:www.nr.no

Omslagsbilde: iStockphoto.com. Design og produksjon: Norsk Regnesentral