The scientific case for eInfrastructure in Norway

The eInfrastructure Scientific Opportunities Panel

The scientific case for eInfrastructure in Norway

The eInfrastructure Scientific Opportunities Panel

Appointed by the Research Council of Norway

Galen Gisler (chair)	Physics of Geological Processes, UiO
Elena Celledoni	Mathematics, NTNU
Trygve Ulf Helgaker	Chemistry, UiO
Trond Iversen	met.no and Geosciences, UiO
Kjetill Sigurd Jakobsen	Biology, UiO
Colin Jones	Swedish Meteorological and Hydrological Institute
Anna Lipniacka	Physics, UiB
Arvid Lundervold	Biomedicine, UiB
Nils Reidar B. Olsen	Hydraulic and Environmental Engineering, NTNU
Koenraad De Smedt	Linguistics, UiB
Jacko Koster (secretary)	UNINETT Sigma
Gudmund Høst (advisor)	Research Council of Norway

Foreword

A modern and functional electronic infrastructure, eInfrastructure, is a prerequisite for a modern society. It includes electronic devices from the personal laptop to huge supercomputers interconnected with high-speed networks. In addition, efficient storage facilities for ever increasing collections of data as well as instruments, tools and services that enable effortless use of these resources are needed. In short, this is the eInfrastructure that is crucial for any highly developed society.

To elucidate the need for a further development of eInfrastructure in Norway for the next decade, the eVITA Programme Committee established the eInfrastructure Scientific Opportunities Panel, chaired by Dr Galen Gisler, University of Oslo, and consisting of eminent researchers from a variety of different disciplines. The Panel's efforts were supplemented through consultation with many researchers throughout Norway and abroad, from both academia and industry. The result of their efforts is the present report.

The eVITA Programme Committee is grateful to the Panel for the effort and dedication they invested in this report, and we are convinced that the report will be useful for the Research Council of Norway and decision makers in pointing out the opportunities for scientific progress that will be enabled by further sustained development of the national eInfrastructure.

Helge Holden

Chair

eVITA Programme Committee
October 2010

Executive Summary

Many Norwegian research groups use high-performance computing, fast data networks, archival storage, and associated services. To help these research groups maintain their internationally leading positions, the Research Council of Norway has invested in the development of the necessary infrastructure through the eVITA[1] programme, especially in the programs Notur, NorGrid, and NorStore. While it is important to obtain and use the best available hardware, no less important are the human resources for supporting and maintaining them and the institutes necessary for housing them. All this comprises a new kind of societal infrastructure, an electronic infrastructure or *eInfrastructure*. This eInfrastructure is not associated with individual projects or institutions, but exists at a national level. It is a *societal* infrastructure, an indispensable part of a "well-functioning research system"[2]. As research is vital to the future of any nation, so is eInfrastructure as vital as power lines and roads.

As an independent societal infrastructure, eInfrastructure requires its own stable and secure funding stream. This eInfrastructure is a *prerequisite* for research, so its funding must not be contingent upon research funding. On the contrary, those doing research and seeking funding need to be assured that the necessary eInfrastructure will be there, robust, stable, and well-maintained. Relative to present conditions, the funding stream for eInfrastructure in Norway needs to be boosted, fortified, and maintained at a sustainable level.

The present eInfrastructure initiative, running from 2005 to 2014, is managed by the eVITA Programme Committee, which reports to the Research Council. Looking beyond the end of this initiative, the Programme Committee appointed the eInfrastructure Scientific Opportunities Panel to assess the future growth of scientific needs for eInfrastructure in Norway. This report is authored by that Panel. The intention of this assessment is to furnish the eVITA Programme Committee and the Research Council with arguments from the research community for the further development of eInfrastructure. It might be expected that scientists, given the opportunity to speculate on future needs, would simply ask for more resources. But the more careful and realistic assessment attempted in this report considers concrete problems that require solutions; those solutions entail eInfrastructure.

This document covers areas that dovetail with the priorities of the Research Council[3], specifically: *climate and environment*, *energy*, *health and welfare*, *social challenges*, *industry*, and *quality of research*. Another Research Council priority — *innovation* — sprouts from the fertilisation of high-quality research with leading-edge technology.

Climate science and weather prediction require new algorithms to exploit the tightly coupled supercomputers that will be available. Norway's location in latitudes that are exceptionally dependent on climate change, and where weather patterns are difficult and complex, presents both unique challenges and opportunities in climate and weather research.

Norway's economy is built on the development and exploitation of sources of energy, and this is implicitly involved in many of the examples given in this report. The search for resources requires detailed and data-rich observations; the efficient extraction of resources requires intensive modelling and detailed monitoring; and the efficient use of energy requires extensive computational work on the physics and chemistry of combustion and turbulence and on the design of advanced new materials.

High-energy physics, traditionally an area of excellence in Norway, needs fast globally-connected networks and large storage systems as well as the middleware for processing experimental data on the grid. Solar physics, Earth sciences and fluid dynamics need tightly coupled supercomputers and advanced visualisation facilities for large-scale models. Chemists, material scientists, and atomic physicists do very detailed calculations on the microscale that result in just a few numbers of high precision; their needs are memory and computation intensive, but not data rich.

The life sciences have a rapidly growing need for eInfrastructure. Spanning from bioinformatics to neuroimaging, they are an extremely diverse community whose needs range from high performance desktop computing to grids and supercomputers; for them the storage of vast amounts of robust, secure, and easily accessible data is a *sine qua non*. The humanities (in particular, linguistics) and social scientists likewise need vast amounts of storage; they need systems for rich metadata and persistent identifiers, and their computational needs will increase as the benefits of indexing, analysis, and correlation become more apparent.

In sum, computational methods are in rapid evolution in all scientific areas, and demands for eInfrastructure will undoubtedly increase. Nevertheless, while the net message is indeed "more resources", the specific needs are diverse and nuanced. How much of that *more* is to be put into high performance computing, grid computing, software and services, networks, and storage, respectively, will be among the questions dealt with in this Panel's eInfrastructure Use Roadmap for in 2011.

Sammendrag

Mange norske forskningsgrupper er avhengige av tungregning, høyhastighetsnett, datalagring og beslektede tjenester. For at forskningsgruppene skal kunne opprettholde internasjonale ledende posisjoner innen sine fagområder har Norges forskningsråd investert i utviklingen av den nødvendige infrastruktur gjennom eVITA[1] programmet, spesielt i delprogrammene Notur, NorGrid og NorStore. Selv om det er viktig å skaffe best mulige datamaskiner, er det ikke mindre viktig å bidra med menneskelige ressurser for støtte og vedlikehold av utstyret, samt institusjoner der arbeidet kan utføres. Alt dette utgjør en ny samfunnsmessig infrastruktur, en elektronisk infrastruktur eller *eInfrastruktur*. Denne eInfrastruktur er ikke bare rettet mot enkeltprosjekter eller institusjoner, men tilbys på et nasjonalt nivå. Det er en *samfunnsmessig* infrastruktur, en uunnværlig del av "et velfungerende forskningssystem"[2]. På samme måte som forskning er viktig for landets fremtid har eInfrastruktur en rolle som kan sammenliknes med kraftlinjer og veier.

Som en uavhengig samfunnsmessig infrastruktur, trenger eInfrastruktur sin egen sikre og stabile finansiering. Denne eInfrastrukturen er en *forutsetning* for forskning, slik at finansieringen må være uavhengig av eventuelle midler fra prosjekter. Tvert i mot, forskere som søker midler må være sikre på at den nødvendige eInfrastruktur er til stede, robust, stabil og godt vedlikeholdt. I forhold til dagens situasjon må finansieringen av eInfrastruktur i Norge styrkes og holdes på et bærekraftig nivå.

Det nåværende eInfrastruktur initiativet, som løper fra 2005 til 2014, ledes av eVITA's programkomite, som rapporterer til Norges forskningsråd. For å beskrive forskningens fremtidige behov for eInfrastruktur i Norge etter denne perioden, oppnevnte programkomiteen et "eInfrastructure Scientific Opportunities Panel", som har skrevet denne rapporten. Hensikten er å gi eVITA's programkomite og Norges forskningsråd argumenter fra forskningsmiljøene for fremtidig utvikling av eInfrastruktur. Det kunne vært forventet at forskerne ville benytte en slik anledning til å be om mer ressurser. Imidlertid er det foretatt en nøktern og realistiske analyse av de faktiske behov som tar utgangspunkt i konkrete utfordringer i forskningen. Løsningen av disse problemene forutsetter en god eInfrastruktur.

Denne rapporten er strukturert etter faglige disipliner, men samtidig dekkes Forskningsrådets satsinger[3], herunder *energi og miljø*, *klima*, *helse og velferd*, *samfunnsutfordringer*, *næringsrelevant forskning* og *kvalitet i forskning*. *Innovasjon* — en annen av Forskningsrådets prioriterte områder — stimuleres gjennom samspillet mellom forskning av høy kvalitet og bruk av ledende informasjons- og kommunikasjonsteknologi.

Klimaforskning og værvarsling har behov for at nye algoritmer utvikles for fremtidens tett sammenkoplede superdatamaskiner for å bidra til å finne løsninger på klimarelaterte problemer. Norges nordlige beliggenhet gjør oss eksepsjonelt utsatte for klimaendringer, da værforholdene er vanskelige og komplekse. Dette gir unike utfordringer og muligheter for vær og klimaforskning.

Norges øknonomi er bygd på utvikling og utnyttelse av forskjellige typer energi, noe som vises i mange av eksemplene i denne rapporten. I leting etter ressurser er det behov for detaljerte observasjoner med store mengder data. En effektiv utvinnelse av ressursene vil i stor grad inkludere modellering og detaljert overvåkning. Norge har mange forskningsprosjekter som går ut på å forbedre energibruken, for eksempel innen forbrenningsteknikk og design av avanserte materialer.

Høyenergifysikk krever høyhastighetsnett på global skala, store lagringssystemer og spesialisert programvare for å muliggjøre distribuerte beregninger basert på eksperimentelle data. Store beregningsmodeller som brukes innen solfysikk, geofysikk og fluiddynamikk krever tett sammenkoblede regneanlegg og avanserte visualiseringsystemer for analyse av modellenes resultater. Detaljerte beregninger på liten skala innen kjemi, materialvitenskap og kjernefysikk gir ofte kun noen få tall som resultat, derfor er behovene innen disse fagområdene i hovedsak knyttet til økt regnekraft.

Biovitenskapelig og medisinsk forskning har sterkt voksende behov for eInfrastruktur. Samlet representerer de mange ulike forskningsfelt, spennende fra bioinformatikk til hjerneavbildning. Behovene for eInfrastruktur omfatter et bredt spekter fra kraftige PC'er til superdatamaskiner, og sentralt står behovet for robuste storskala lagringssystemer for tildels sensitive data. Humaniora (spesielt språkvitenskap) og samfunnsfagene har behov for kraftige lagringsressurser, systemer for metadata og gjenfinningsverktøy. Regnekraftbehovene innen disse fagområdene vil øke etterhvert som fordelene med å kunne indeksere, analysere og korrelere blir mer stadig mer åpenbare.

Beregningsmetoder er i rask utvikling innen alle forskningsområder og behovet for eInfrastruktur vil øke. Selv om denne rapportens hovedbudskap er "mer ressurser" er de spesifikke behovene mangfoldige og nyanserte. Hvor mye *mer* som trengs av tjenester knyttet til tungregneanlegg, grid-verktøy, programvare, høyhastighetsnett og lagring er sentrale spørsmål som vil bli behandlet i dette panelets veikart for eInfrastruktur i 2011.

---

[1] eVitenskap, Teori, og Anvendelser: eScience, Theory and Applications.

[2] From "Climate for Research", Report No. 30 to the Storting, 2008-2009, Ministry of Education and Research, p. 1.

[3] Ibid., p. 1.

About this Document

The eInfrastructure Scientific Opportunities Panel, appointed by the Research Council of Norway and reporting to the eVITA Programme Committee, is responsible for monitoring the development of scientific use of eInfrastructure. Our first assignment, from the Panel's *Terms of Reference* is:

> By June 1st, 2010, develop the scientific case for the eInfrastructure that can best serve Norwegian research groups and operational forecasting from 2015. eInfrastructure in the present context covers electronic resources such as large data collections, large-scale computing resources and high-speed networks, as well as the tools and services enabling efficient use of these resources.

This present document is intended to fulfil that assignment. Our job is not finished, however. We have another assignment yet to carry out:

> By June 1st, 2011, produce a first version of an eInfrastructure Use Roadmap. This Roadmap should cover current and new scientific areas for eInfrastructure use, taking into account the opportunities offered by existing and emerging large-scale international collaborations. The roadmap should also make international comparisons and, in collaboration with the NOTUR project leader, match applications against hardware architectures.

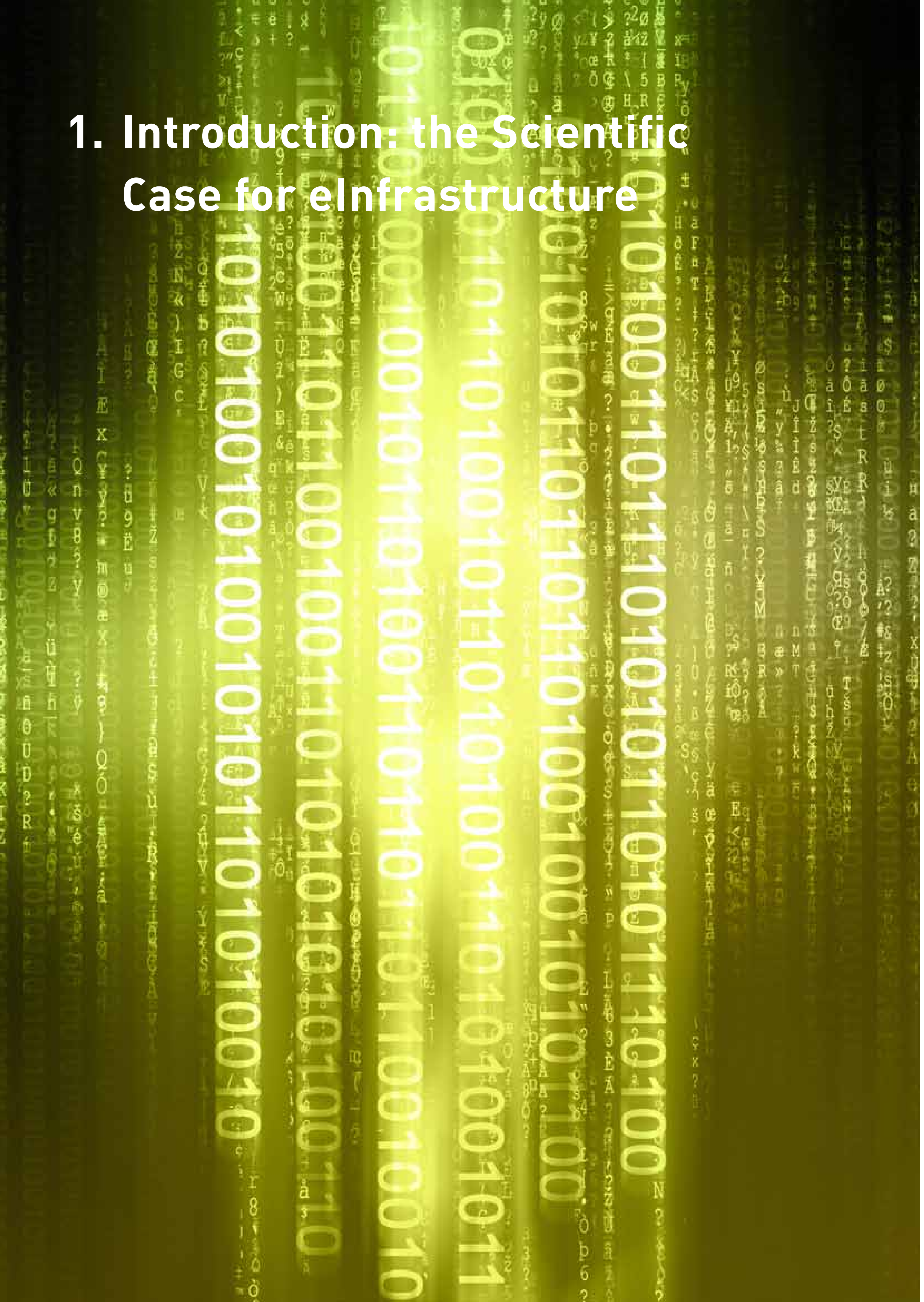
We will be working on that Roadmap over the course of the 2010-2011 academic year, noting that it is "a first version" and that changes will likely be necessary in subsequent years.

List of common abbreviations used in this document

CERN	European Organisation for Nuclear Research
ECMWF	European Centre for Medium-Range Weather Forecasts
EGI	European Grid Initiative
eIRG	eInfrastructure Reflection Group
ESFRI	European Strategic Forum on Research Infrastructures
FP7	The European Union's Seventh Framework Program
ICT	Information and Communication Technology
LHC	Large Hadron Collider
met.no	The Norwegian Meteorological Institute
NDGF	Nordic Data Grid Facility
RCN	Research Council of Norway
NILU	Norwegian Institute for Air Research
Notur	The Norwegian Metacentre for Computational Science
NorGrid	The Norwegian Grid Initiative
NorStore	The Norwegian Data Storage Infrastructure
NSD	Norwegian Social Science Data Services
NTNU	Norwegian University of Science and Technology, Trondheim
PRACE	Partnership for Advanced Computing in Europe
UiB	University of Bergen
UiO	University of Oslo
UiT	University of Tromsø

Table of Contents

1. Introduction: the Scientific Case for eInfrastructure



Science starts from intriguing observations that provoke thoughtful speculations; experiments or new observations then suggest theories and eventually laws and models. Observation, experiment, and theory are all now performed using an infrastructure consisting of fast computers, fast networks, and robustly accessible data storage. This infrastructure has also become indispensable for communities far beyond the natural sciences.

Computational science was born as the theorist's assistant, solving systems of equations that were intractable to analysis (§3). As the observer's assistant, it helps construct, analyse, and manipulate images and then digest large quantities of raw data into a form amenable for interpretation. As the experimenter's assistant, it helps design experiments and instruments, and it records and archives data. Computational science has become a bridge among theorists, observers, and experimenters by enabling "numerical experiments" that are impossible to do in the laboratory for reasons of size, duration, hazard, or ethics.

In systems that are critically sensitive to initial conditions and boundary data, like weather and climate (§4.4), theory, observations, and experiments are synthesised, using computers, into predictions that are constantly confronted by new observations. Immediate benefits of advanced basic research are conveyed to society through computers, and new types of science, like chaos and dynamical systems theory (§3), are born in the exchange.

The development of genome sequencing over the last few decades has led to explosive growth in the application of computational science to biology (§4.6), and spreading from there to epidemiology, crime prevention, and forensics. The vast amount of data stored in an individual's genome, and the need to compare and collate genomes of many individuals has pushed the development of novel strategies. Similar techniques and infrastructure have been used by the field of computational linguistics (§4.7), and cross-fertilisation has proven fruitful. Political science and the practice of law make use of data mining techniques, and the social sciences (§4.8) are alert to the possibilities that new ways of accessing data can offer. Areas of industrial applications (§4.9), important to Norway's future, benefit strongly from these new techniques.

The name *eInfrastructure*[4] has been given to the combination of (1) computers from laptops to the fastest supercomputers on the planet; (2) the networks that transfer data among them at fast speeds; and (3) the banks of storage and archival media that are accessible from these computers and networks; and (4) the instruments, tools, and services that enable researchers to make use of these resources (see §2 and §3 for a more complete discussion). This complex, assembled *ad hoc* from disparate parts over the last few decades, is now considered of vital importance in its own right, a component of the world's working as important to human endeavour as the airports, railways, roads, bridges, and cities that we take for granted.

It is for this reason that eInfrastructure is now being established and funded in a way that reflects its position as a fundamental societal infrastructure, independent of its use, as roads and railways are funded whether they transport food, sofas, or Members of Parliament.

The task of the eInfrastructure Scientific Opportunities Panel is to assess the anticipated needs of Norwegian research groups for eInfrastructure resources over the next decade, and to develop the scientific case for that eInfrastructure. This document, *The Scientific Case for eInfrastructure in Norway*, is the first result of that assessment. A subsequent document, whose first version is scheduled for June 2011, is intended to provide a Roadmap for the development of that eInfrastructure.

Fundamental versus Applied Research

The boundary between *Applied* and *Fundamental* science shifts over time (see §3 and §4.1-4.3 for examples of fundamental research that have become applications). In a sense, the distinction lies between *Problems* we must solve and *Questions* we are driven to ask. Public money readily funds research groups in Norway, as in other countries, to look for solutions to *Big Problems* like those in Table 1, because they are of concern to all. But we are unable to solve *Big Problems* unless we are bold enough to ask *Big Questions*, some of which are listed in Table 2. So research groups must also focus on these. Addressing *Big Problems* prepares students for the real work of tackling them and provides the public with advice they need. Addressing *Big Questions* sows seeds for future applications, provides the public with partial answers, and inspires students with the scope and vision to handle their work responsibly. International and interdisciplinary approaches and the eInfrastructure of sophisticated computers, fast networks, and reliable data storage are used to find solutions to these problems and answers to these questions.

A specific practical interdisciplinary example — volcanic ash from Eyjafjallajökull

Members of this panel were affected as were many others in Europe by the eruption of Eyjafjallajökull (Fig. 1) and the subsequent traffic chaos. Improved eInfrastructure will help to ameliorate our response to future similar events. In Table 3 we list questions that have been discussed by frustrated passengers in airport lobbies and by national newspapers, grouped by the disciplines that can address them with the aid of eInfrastructure.

Figure 1. The Eyjafjallajökull glacier and the volcanic ash plume, on 17 April 2010.

---

[4] Also known, principally in North America, as "cyberinfrastructure".

Specific needs for Norway

Among the activities reported on in the Scientific Areas (§4.1-§4.9) are several in which Norwegian researchers are currently among the world's top leaders. To maintain their leadership positions they will require access to the best eInfrastructure components available in the world. In addition, there are some particularly Norwegian requirements in certain of the scientific areas that call for special mention here.

In weather forecasting and climate predictions, the unique combination of steep mountains and deep fjords and Norway's position relative to the westerlies present a severe challenge, as well as the general difficulties of dealing with arctic weather patterns. These same factors also make it possible for Norway to realise greater benefits from using very high resolution numerical weather prediction models than in most other places. We thus have an interesting and unique testbed for the development of such advanced models, which will require the most advanced eInfrastructure available.

Fluid dynamics research has always been strong in Norway, partly because of our reliance on hydropower. With the increase of other energy sources, including wind, tidal, and solar power, understanding the effects of regulations and environmental impacts is increasingly important. Sedimentation in rivers is one factor that affects dams built for hydropower, and a specific Norwegian example is given in §4.3.

The official Norwegian language policy recognizes the importance of information technology solutions for language preservation and language use in society. Most expertise on Norwegian, as well as most of the appropriate resources and tools are located at Norwegian institutions; these will require good eInfrastructure for curation and processing tasks.

Natural Hazards: Predicting severe weather, §4.4 Predicting geohazards, §4.5 Building resilient communities, §4.8	**Energy and Resources:** Improving energy efficiency, §4.2, §4.3 Developing new materials, §4.3 Capturing and storing carbon, §4.5, §4.9
Disease: Expanding public health, §4.6 Fighting cancer, §4.6 Predicting/controlling epidemics, §4.6	**Environment and Climate:** Ensuring clean air and water, §4.4 Mitigating or adapting to climate change, §4.4, §4.9 Managing the space environment, §4.1
Human Populations: Ensuring adequate food for all, §4.6 Slowing population growth, §4.8	**War and Peace:** Resolving conflicts rationally, §4.8 Avoiding causes for war, §4.8

Table 1. Suggestive list of Big Problems, with sections of this document where they are addressed.

What are the fundamental constituents of matter? §4.1
How and when did the Universe begin — and what is its ultimate fate? §4.1
How did the Sun and the solar system come to be as they are? §4.1, §4.5
What processes shaped Earth's continents and oceans? §4.4, §4.5
How sensitive is Earth's climate system — are there "tipping points"? §4.4, §4.9

How, when, and where did life begin? §4.6
What are the mechanisms and triggers of biological evolution? §4.6
Can the mechanisms of ageing be controlled or stopped? §4.6
How common is life in the Universe? §4.1, §4.6

What is consciousness? §4.6
How did human intelligence evolve — and is it unique? §4.6

Can humans live together in peace and justice? §4.8
What is the carrying capacity of our planet? §4.4, §4.5, §4.9

Table 2. A suggestive list of Big Questions, with sections addressing them.

Materials science, §4.2:
Can the transparency of aircraft windows be maintained under the impact of volcanic ash?

Fluid dynamics, §4.3:
What size and number density of ash particles can be tolerated by jet engines, internal combustion engines, or propellers?

Atmospheric science, §4.4:
Can we do better at
- measuring the densities and particle sizes of volcanic ash when it is in the air?
- predicting the propagation and fallout of volcanic ash?

What are the short- and long-term climate consequences of massive ash eruptions?

Earth science, §4.5:
What processes lead to the fragmentation of magma and massive ash eruptions?
Can the characteristics of ash be predicted from observations of eruptions?

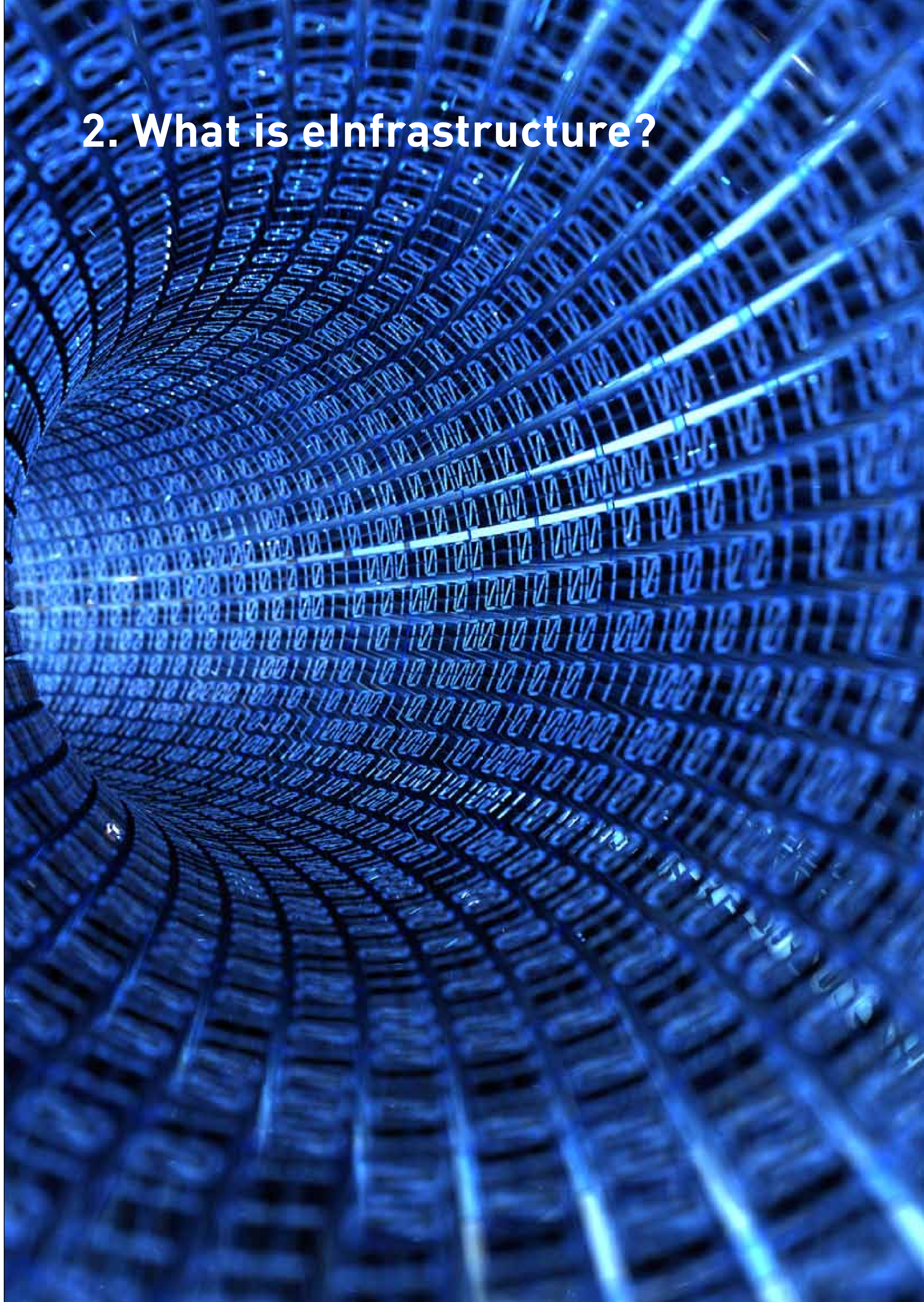
Life science, §4.6:
What characteristics of volcanic ash particles are hazardous to life?

Linguistics, §4.7:
Can different attitudes to the event be detected in media coverage in different countries?
Is language itself affected, through the coining of new words?

Social science, §4.8:
Can we do more of our work without the use of air travel?

Table 3. Practical questions about volcanic ash from Eyjafjallajökull.

2. What is eInfrastructure?



As viewed by the EU's FP[5] Capacities Programme, *eScience* is an innovative way of conducting scientific research by the creation of a new environment for academic and industrial research in which virtual communities share, federate, and exploit the collective power of scientific facilities. The eInfrastructure Reflection Group (eIRG)[6] defines *eInfrastructure* as "this new research environment in which all researchers — whether working in the context of their home institutions or in national or multinational scientific initiatives — have shared access to unique or distributed scientific facilities (including data, instruments, computing and communications), regardless of their type and location in the world."

The information and communication technologies (ICT) enable eInfrastructure as a new kind of societal infrastructure that includes associated resources, tools, and services. Technologies belonging to eInfrastructure include computer facilities and peripherals, high-performance and high-capacity networks, grids and collaborative environments, support for software development and life cycle management, tools to manage and share resources, data and on-line content, and applications to process and present research activity. These components support the needs of *all* researchers, even those who do not make extensive use of ICT. Using eInfrastructure, researchers share access to large data collections, advanced tools for data analysis, large-scale computing resources, and high-performance visualisation. Significant improvements in the productivity of multi-institutional collaborations result from better eInfrastructure.

We think of eInfrastructure as the integrating mechanism, the *glue* between regions and different scientific disciplines. New opportunities for researchers arise from remote access to computing services, new instrumentation, and virtual organisations. New scientific communities arise from new eInfrastructures; researchers working in different fields but on similar challenges attain new levels of collaboration and new ways of sharing data, using sophisticated new simulation tools and virtual environments. This highlights the importance of providing eInfrastructure as a service, rather than continuing with a product- or technology-oriented approach. The eIRG promotes this approach to ensure the continued ability of eInfrastructure to act as an innovation engine and accelerate the transition of leading-edge, research-focused ICT applications into solutions that benefit society as a whole. The European Strategic Forum on Research Infrastructures (ESFRI)[7], which supports policy-making on research infrastructures in Europe, emphasises that eInfrastructure is a critical factor in *all* research infrastructures. Norway is involved in several projects proposed in ESFRI's Road-map for Research Infrastructures[8]. The important enabling role of eInfrastructure is realised by the European Council[9] and eInfrastructure is a central pillar of EUs research and innovation policy.

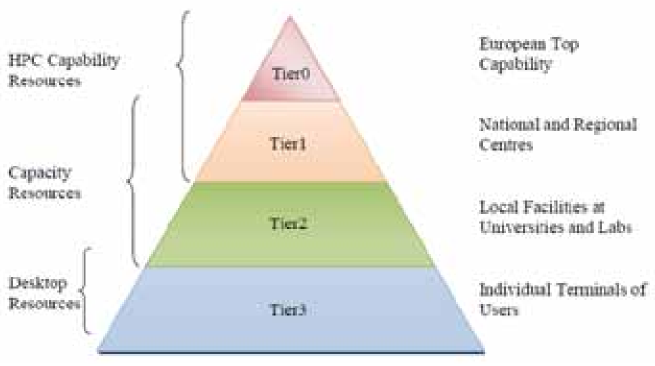


Figure 2. The European high-performance computing pyramid. The top level presently consists of systems in the Petascale range.

High performance computing services

Modern scientific research requires massive computing resources: the natural sciences have long-established needs and patterns of use in this area, while the humanities and social science communities are emerging as users (see §3). High performance computing is essential for both basic and applied research, and is the core component of eInfrastructure.

A sustainable infrastructure for high-performance computing consists of systems of differing architectures supporting different classes of algorithmic processes in a cost-efficient manner. A specific vision for the overall infrastructure is the pyramid (Fig. 2) whose top represents capability computing and lower levels are capacity computing down to researcher workstations. In *capability* computing (or supercomputing) sufficient power and internal bandwidth is available to solve latency-bound or communication-intensive problems; scalability and overall application performance are of vital importance. Capability computing facilities are large, expensive, and unique (nationally or internationally) in their own kind at the time of installation. Examples of disciplines for which capability computing is essential are weather forecasting, climatology, fluid and plasma dynamics, combustion and nuclear fusion. Capacity computing includes a larger number of smaller and less expensive high-performance systems for simulations with more modest computational requirements, for which scalability is not so critical.

Distributed computing services: grids and clouds

Another component of global eInfrastructure is the *grid*, an environment for distributed computing and sharing distributed data that allows new methods of collaborative research. The grid uses many distinct computer systems, from capability systems down to workstations, and data sources with transparent interoperability between institutional and national environments Grid infrastructures can be truly enormous, measured by the aggregated capacities of the available compute and storage systems and the number of researchers that make use of them. With grids and modern high-capacity communication networks, pan-European virtual research communities achieve very broad impacts.

In *cloud computing* shared computers, storage, software, and information are provided through the internet on demand, like a public utility. Clouds typically consist of services delivered through large data centres and often appear as single points of access for all consumers' computing needs. A number of commercial offerings have emerged in recent years. Customers typically rent usage from a third-party provider; they consume resources as a service and pay for the resources that they use. Cloud computing has considerable long-term potential for use in large-scale science.

Storage and curation services for scientific data

A knowledge society is built primarily on information, not merely on hardware and software. The preservation and free transmission of knowledge requires good access to digital information. Modern society demands advanced levels of data curation[10]. Scientific data collections are not merely stored or archived, but are subject to frequent revision and enhancement. Curation adds value to digital information to enhance its contemporary and future use, by ensuring that data are valid and available for discovery and reuse. Readily available scholarly and academic information resources enable and facilitate cross-institutional, cross-border, and cross-disciplinary science.

Modern research infrastructures include the storage and curation of scientific data. Digital content across all disciplines must be part of research infrastructure. The rapid explosion of digital information that must be saved is a significant challenge. The production and use of unprecedented quantities of complex scientific data, including research input gained by means of observation, experience, or experiment, and research outputs like reports and publications, will place stringent demands on the infrastructure. These data will need to be stored, maintained, published, and made openly accessible, but the data must also be validated so that researchers can trust them. Trust in data is enhanced by the curators: qualified domain specialists who deal with the issues of permanency, provenance, authenticity, integrity, interoperability, and quality of the primary data and associated metadata, but also with security, ownership, confidentiality and privacy issues.

Core services, software, and personnel

Beyond the physical and electronic resources themselves, eInfrastructure requires services to ensure that the overall infrastructure provides an ever-advancing, agile, robust, and user-friendly production environment. These services are provided by *people*, and therefore the eInfrastructure must include a long-term commitment to maintaining and continuously improving the skills and competence of its human resource base. Research, development, support, training, community development, outreach, and education are all fundamental aspects, and lead to strength, flexibility and the capability to respond to new areas of science. Such services include:

- Technical support for installation and maintenance of the equipment
- User support for helping researchers use the infrastructure components, including advice on appropriate algorithms for appropriate architectures
- Security infrastructure for authentication and authorisation
- Research staff for developing new algorithms and new architectures and making the best out of both
- Staff that maintain contacts with counterparts in other countries.

Present organisation of eInfrastructure

The Norwegian national infrastructure, funded in part by the Research Council of Norway, presently consists of a high-performance computing project (Notur II), a project for scientific data services (NorStore), and a national grid initiative (NorGrid). This distributed infrastructure has been built by a consortium of partners that consists of the universities that host and operate the resources (UiB, UiO, UiT, NTNU) and a coordinating partner, UNINETT Sigma[11]. UNINETT provides the national network for research and education and provides services for identity management. The Norwegian Social Science Data Services[12] assists researchers involved in empirical research in gathering and analysing data, and with issues of methodology, privacy, and research ethics. NSD has extensive experience in documentation of data. NSD's data holdings provide information about the human society at different levels.

The Nordic Data Grid Facility[13] coordinates and supports a production grid infrastructure leveraging the national infrastructures from Denmark, Finland, Norway and Sweden. NDGF does not own or operate its own large-scale eInfrastructure resources for computation and storage, but builds on the national infrastructures built by the national organisations. NDGF coordinates a distributed centre as part of the Nordic contribution to the worldwide collaboration on analysing data from the Large Hadron Collider located at CERN. The national research networks of the five Nordic countries collaborate in NORDUnet[14], which provides further connection to their European counterparts.

Scientific endeavour recognises no borders, but addresses questions on a global scale. Some challenges are globally relevant, like climate and energy; some are globally distributed, like health; and others, like high energy physics, are simply too large to be undertaken by any single national community. Important steps toward providing pan-European services have been taken by the European Commission through ESFRI and eIRG. Major European projects are constructing general-purpose eInfrastructure service layers, avoiding the fragmentation of national policies. Projects with Norwegian participation include the European Grid Initiative[15] and the Partnership for Advanced Computing in Europe[16].

Linking national eInfrastructures into a powerful global network will dramatically change the way research is conducted and irrevocably alter the landscape of science, providing research opportunities for many future researchers.

---

[5] The European Commission's Seventh Framework Program (FP7) defines the preparatory, implementation, and construction phases of research infrastructures.

[6] eIRG: www.e-irg.eu. eIRG's mission "is to pave the way towards a general-purpose European eInfrastructure", from the eIRG Roadmap (2010): www.e-irg.eu/images/stories/eirg_roadmap_2010_layout_final.pdf.

[7] ESFRI: cordis.europa.eu/esfri.

[8] ESFRI projects that Norway is involved in include (first three hosted or coordinated in Norway): CESSDA: The Council of European Social Science Data Archives. www.cessda.org; ECCSEL: European Carbon Dioxide Capture and Storage Laboratory Infrastructure. www.eccsel.org; SIOS: Svalbard Integrated Arctic Earth Observing System. www.forskningsradet.no/sios; EISCAT_3D: The European Incoherent Scatter radar system. www.eiscat3d.se; ESS: The European Spallation Source is the next generation neutron research facility. www.ess-neutrons.eu; CLARIN: The Common Language Resources and Technology Infrastructure. www.clarin.eu; ELIXIR: The European Life Sciences Infrastructure for Biological Information. www.elixir-europe.org.

[9] The future of ICT research, innovation and infrastructures: Adoption of Council Conclusions, Council of the European Union, 2009, RECH 425, COMPET 487, TELECOM 248.

[10] Data curation refers to the policy and practice regarding the creation, management, and long-term care of data.

[11] sigma.uninett.no

[12] NSD: www.nsd.uib.no

[13] NDGF: www.ndgf.org

[14] NORDUnet: www.nordu.net

[15] EGI: www.egi.eu

[16] PRACE: www.prace-project.eu

3 Mathematics, Statistics and Information Science

In March 2003 the Société Nautique de Genève became the first European team to win the America's Cup. Swiss designers teamed up with numerical analysts at École Polytechnique Fédérale de Lausanne and produced accurate computer approximations of a sailing boat and its interaction with winds and water, reproducing more than 400 different boat configurations, and solving problems with more than 160 million unknowns.

Figure 3. The sailing boat Alinghi.

A combination of advanced physical models and numerical methods, and a powerful supercomputer, made it possible to reproduce all scenarios the boat Alinghi (Fig. 3) could possibly incur during a race. "Indeed, a precise prediction of the transition location and turbulence development on, e.g., boat appendages is crucial to obtain an accurate estimate of the forces acting on the different boat components.[17]" Similar simulation challenges have been addressed by teams of Norwegian mathematicians and engineers in the field of ship design, offshore installations, wind energy, and computational fluid dynamics in general (see also §4.3 and §4.9).

Mathematics, Statistics and Information Science as part of the eInfrastructure

Mathematics is the common language of science. As such mathematics is fundamental to eScience and an essential component of eInfrastucture for science. *Mathematical models* encode the essence of scientific problems, formulating them in simplified and rigorous form. The solution is then searched for by *mathematical methods*, by studying the existence and features of the solutions, or approximated using *numerical algorithms*.

The simulation of natural processes using modern computer tools began in the 1940s with the groundbreaking work of the mathematician John von Neumann, leading through the years to enormous progress in all branches of science and technology. Two major mathematical contributions to this development deserve to be mentioned: the discovery of the Fast Fourier Transform (FFT) algorithm[18], and the development of efficient techniques for the computer solution of large linear system of equations and eigenvalue problems. Both these subjects have their roots in the pre-computer era[19] but are still under active investigation. The FFTW[20], developed at the beginning of the millennium, is a library of different FFT algorithms each designed for a particular hardware architecture to achieve the highest computational performance. Such features are important in distributed computing (see §2) and are the new trend in software design. Mathematical research addresses also the group theoretical fundamentals of FFTs in the search for new horizons and new areas of application for this algorithm.

A recognised property of science is that applications benefit from mathematical rigour, and new mathematics is in turn inspired by new scientific ideas. Similarly in eInfrastructure, hardware and algorithms make progress hand-in-hand. The most advanced computational facilities inspire advanced mathematical algorithms and those in turn lead to new ideas in hardware.

Simulation technology and modern computing have a firm mathematical foundation and are an integrated part of the scientific method. Huge steps forward can be expected in science when important new mathematical algorithms, as the examples mentioned above, are discovered and made available as new scientific tools.

Improvements in computer speed follow Moore's law, and double every 18 months, but theoretical insights and algorithmic advances have resulted in analogous and sometimes even more dramatic increase in the speed of simulation processes. The case of partial differential equations is illustrated in Fig. 4.[21] Another example is given by Monte Carlo simulation of spin systems in nanoscience; over a 25-year period, form 1970 to 1995, the relative improvement due to Moore's law amounts to a speedup of three orders of magnitude, while additional seven orders of magnitude should be attributed to improved algorithms[22].

---

[17] mathicse.epfl.ch/cmcs/NewResearch/americascup.php3. Quote from research description of the CFD simulations for the Alinghi team.
[18] Cooley, JW, Tukey JW, 1965. An algorithm for the machine calculation of complex Fourier series, *Math. Comput.* **19**, 297.
[19] Gauss developed an FFT algorithm in the nineteenth century to speed up celestial mechanics calculations. See: MT Heideman, DH Johnson, CS Burrus, Gauss and the history of the FFT, *IEEE ASSP Magazine*, October 1984.
[20] Frigo, M and Johnson, SG, 2005. The design and implementation of FFTW, *Proceedings of the IEEE* **93**: 216–231.
[21] A Science-Based Case for Large-Scale Simulation, US Dept of Energy, volume 1, 2003.
[22] A Science-Based Case for Large-Scale Simulation, US Dept of Energy, volume 2, 2004.

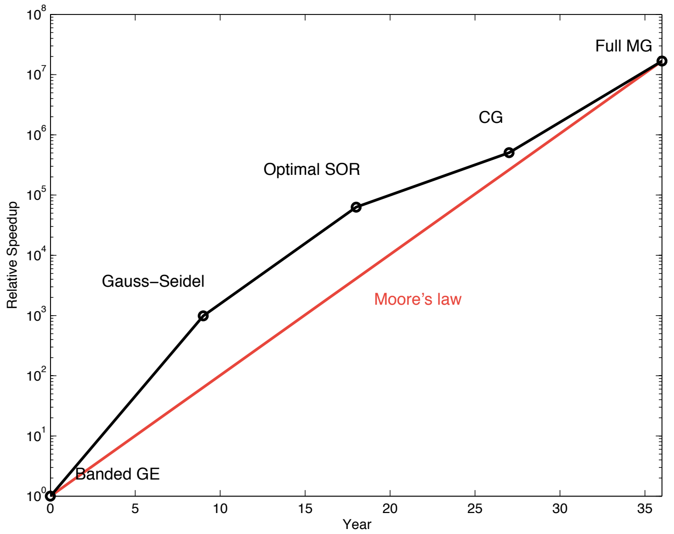


Figure 4. The solution of the electrostatic potential equation on a uniform cubic grid of n x n x n cells. The relative gains of some solution algorithms for this problem and Moore's Law for improvement of processing rates over the same period (illustrated for the case where n = 64)[21].

Mathematical algorithms are an integrated component of eInfrastructure. Consider the web search. Improved search algorithms are the key to Google's success, for example. The PageRank algorithm gives a measure of the relevance of a webpage to a particular user search, not just by means of its content, but also by how many other pages link to it. This apparently simple strategy gave Google its leading status among web search engines (see also §4.7). In mathematical terms this approach can be described as a clever choice of metric to describe the objective function of an optimisation problem. The web stores a good portion of human knowledge; it saves us trips to libraries; it helps us find people; it gives us the latest news form the world. This anarchic container would be useless without effective algorithms for retrieving information. The invention of the World Wide Web (§4.1) is one of the major cultural revolutions of the twentieth century, but its usefulness is the result of timely improvements of mathematical search algorithms to cope with the web's constant expansion. In Norway, this technology, guaranteed the success of the company FAST[23]. Emerging challenges in this field are problems of data security and privacy with implications for social sciences and law.

Mathematical tools for eInfrastructure: some emerging trends and scientific challenges

Reliability of simulations

Mathematics plays a substantial role in the quality control of simulations, by producing estimates of the error committed by the uncertainty of the models and by the approximations used. Analysis of simulation methods proves their reliability as predictive tools, and, most importantly, describes their limitations.

NASA's Mars Climate Orbiter mission failed in 1999 because of a lack of complete end-to-end verification of the navigation computer models, among other factors. A mistake in the computation of the corrections of the spacecraft's trajectory as it approached Mars led to the loss of the spacecraft itself. The control of similar mechanical systems and the accurate reproduction of their geometry is highly important to the Norwegian offshore industry and in particular for the analysis of risk and reliability of marine structures and operations[24]. Precious expertise gained by engineers through years of research is part of the national heritage of the Norwegian oil endeavour.

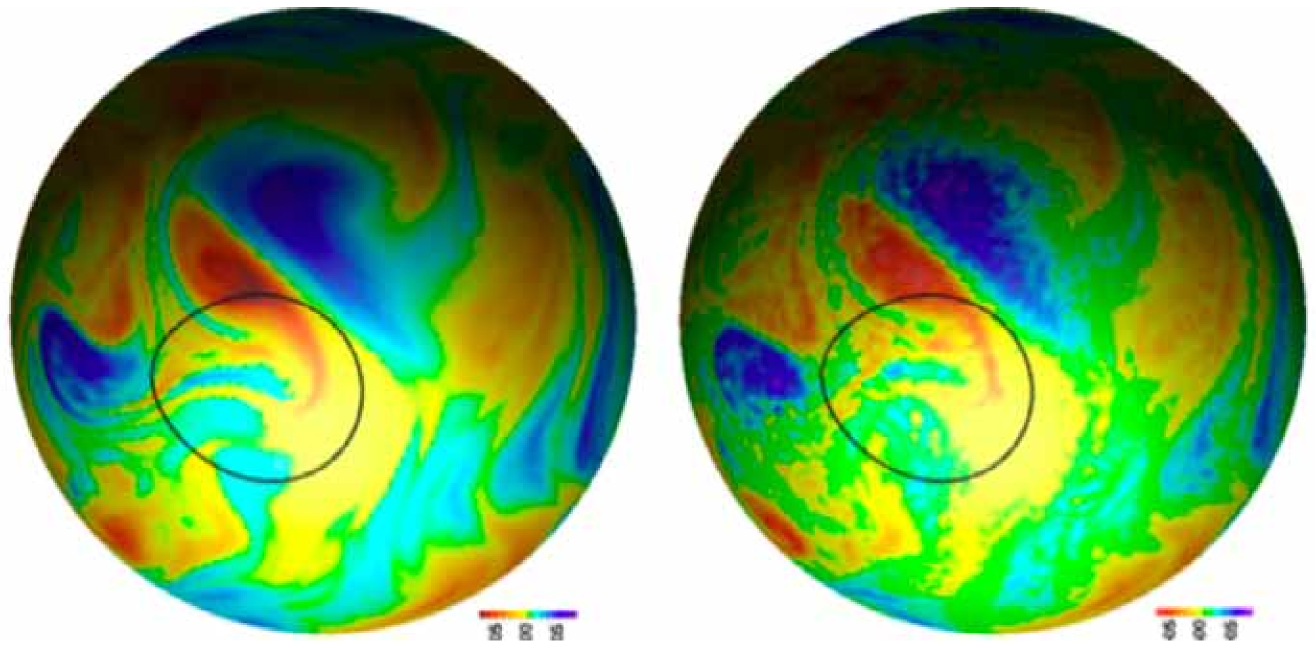


Figure 5. Simulation of potential vorticity with two different numerical methods.

Today's computer power is insufficient to master the grand challenges of our time. Our ability to predict hurricanes and extreme weather depends on improving mathematical models (§4.4). During hurricanes Gustav and Ike in 2008, different prediction codes gave conflicting estimates of the extent of storm surge, and the consensus results were not good enough to warn the authorities of potential flooding. In October 1987, a strong local hurricane surprised people in southern England. It is believed[25] that the failure to predict this storm was due to nonconservation of potential vorticity in the numerical weather forecasting methods (Fig. 5). These problems underscore an emerging need for new and better models encoding the physics of natural processes, and sound approximation strategies based on analytical insight and clever mathematical algorithms.

The use of appropriate computational tools to attack particularly hard scientific computing challenges has proven essential to the successful outcome of the resulting simulations. Symplectic integrators are of crucial importance in celestial mechanics simulations over long times. A new integrator for the solar system, incorporating new symplectic and structure preserving methods, contributed to the 2004 Geophysical Time Scale and was responsible for realigning epochal boundaries by millions of years[26].

Complex models and probabilistic modelling

Computers have changed the way mathematical research is carried out. The discovery of solitons[27] as solutions of certain nonlinear partial differential equations arose from computational evidence. Another example is the calculation[28] in 2007 of the representations of the group E_8, an important milestone in the theory of Lie groups, structures invented in the nineteenth century by the Norwegian Sophus Lie.

The increased availability of computational power has led to the use of more complex mathematical models in computer simulations, but these require a detailed understanding of the mathematical theory behind them. Without suitable mathematical algorithms it is impossible to exploit the available computational resources optimally. As the hardware constantly evolves so also must the search for appropriate software be pursued in turn.

With the increasing complexity of models comes an increasing complexity among model results. In fields like climate and weather forecast, biology, medicine and exploitation of natural resources, it is necessary to understand and *quantify the uncertainties* inherent in the models and the extent to which simulation outputs actually relate to real-world processes.

In many cases the primary objective of numerical simulation of time evolving systems is the prediction of coarse-grained (macroscopic) dynamics. Numerous algorithmic approaches have been introduced to extract effective, lower-dimensional, macroscopic dynamics (multiscale models and methods, see §4.3, §4.4, §4.6 and §4.9). The model problems may be either deterministic or stochastic and the study of the transition from the microscopic to the macroscopic description is the crucial challenge to address (see for example §4.4 on weather forecasting and §4.9 on reservoir modelling).

At the Centre of Excellence for Ships and Ocean Structures, during recent years a lot of attention has been devoted to the development of simulation based methods for statistical prediction of extreme loads and responses in complex systems. Such methods can be computationally very challenging in terms of the requirements on sheer computational power. For many such problems a direct approach is simply impossible with present day computers, even supercomputers. Consequently, the research has focused on developing new techniques for reducing the demands on the processors. The key ingredient is a reformulation of the deterministic models in a stochastic setting. This work has resulted in very efficient and powerful statistical methods to reduce the computational cost of large scale simulation. Many problems have thus become accessible with current computational facilities and the desired predictions have successfully been produced. This research has very important practical implications for example in offshore installations and ship design, and its further success and development will depend on the availability of good eInfrastructure.

Algorithms for analysis and storage of data

Computer power offers also new mathematical challenges in the field of data analysis, as often large amounts of data, are produced by computer simulations. The use of advanced and clever *statistical methods* has brought significant advance in numerous fields like computational biology (§4.6), signal separation, image analysis and pattern recognition, with important applications in medical sciences, but also in social sciences and humanities. As eInfrastructure becomes more important to fields like biology, medical sciences and linguistics (§4.6, 4.7), special needs and new challenges in the field of data processing and pattern recognition spontaneously emerge from these sciences.

A fresh look at an old problem can lead to progress. Estimating the changing average temperature on our planet is a very challenging problem of reconstructing information from data. The difficulty arises because the temperature data are from measurement points scattered in an uneven fashion on the globe. Brute force computations on even the most powerful supercomputers lead to unsatisfactory results. A recent breakthrough has been obtained by using modern methods of spatial statistics, relying on the clever use of sparse matrix computations. The results[29] give precise estimates of the statistical error due to the uncertainty of the measurements (Fig. 6).

---

[23] Fast Search and Transfer, bought by Microsoft. www.microsoft.com/enterprisesearch/en/us/fast-customer.aspx

[24] T. Moan *et al.*, 1981. The Alexander L. Kielland Accident, NOU 11: 1981, Universitetsforlaget, Oslo.

[25] TD Ringler, J Thuburn, JB Klemp, WC Skamarock, A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured C-grids. J. Computational Physics, 229, 2009.

[26] J Laskar, P Robutel, F Joutel, Gastineau, ACM Correia, and B Levrard, 2004. A long-term numerical solution for the insolation quantities of the Earth, *Astron. Astrophys.* **428**, 261-285.

[27] Zabusky, N. J.; Kruskal, M. D. 1965. Interaction of solitons in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**: 240–243.

[28] aimath.org/E8/

[29] F Lindgrenn, J Lindström, H Rue, An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach, technical report.

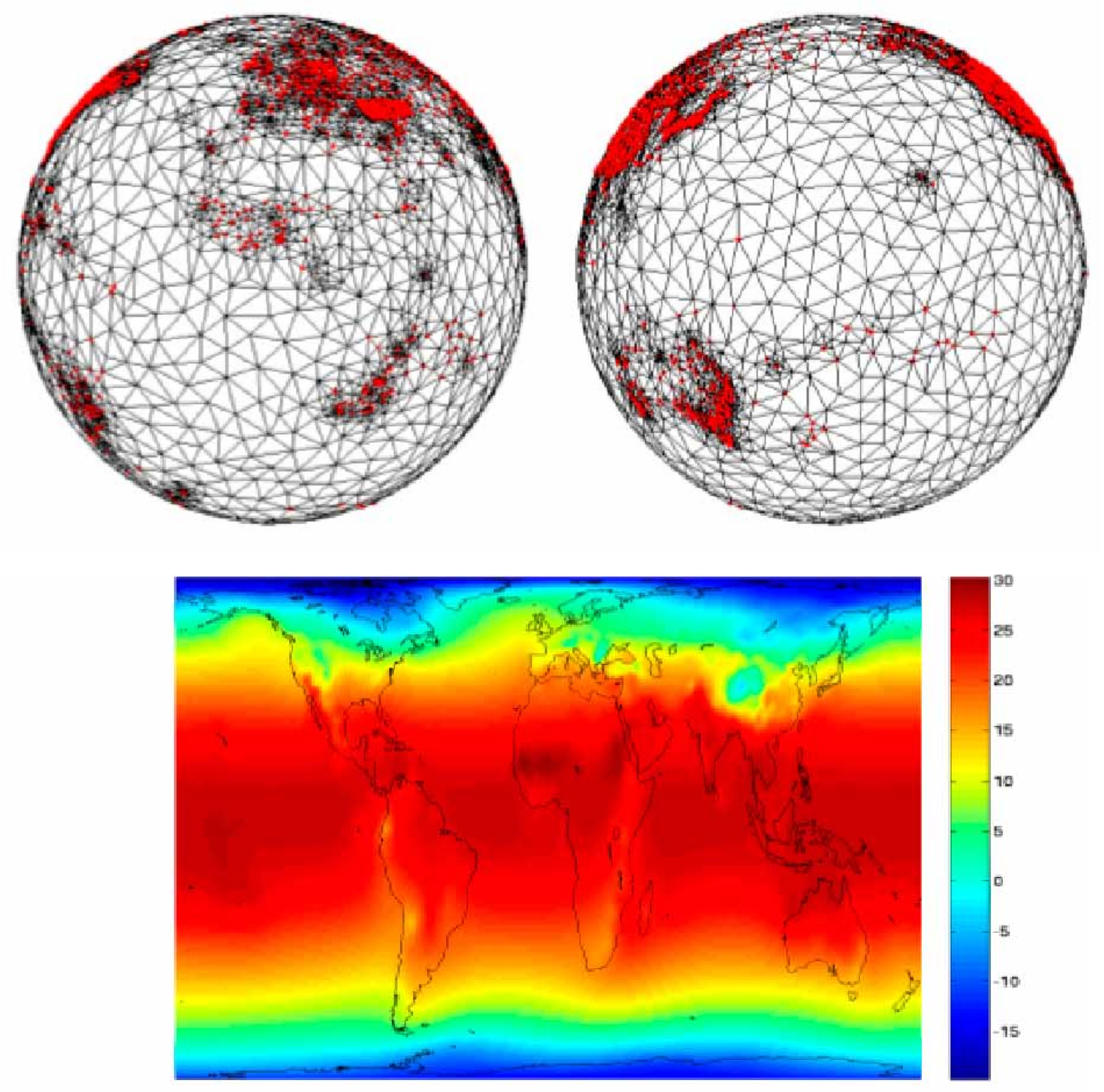


Figure 6. Scattered measurements of temperature (above) and global reconstruction of average temperature for 1980 (below).

Heterogeneous computing and GPUs

Heterogeneous computing[30] refers to the use of heterogeneous processing cores with different compute characteristics to maximize performance. In particular, the combination of multi-core CPUs and many-core accelerators (graphic processing units, GPUs) offers unprecedented floating-point processing power and energy efficiency. Hardware accelerators, like GPUs, are special-purpose hardware designed to maximize the performance of a restricted set of operations, and will hence deliver a much higher number of floating point operations per second or per watt than traditional CPUs. Such heterogeneous systems have become increasingly popular in recent years and they require tailored software components and specially developed numerical techniques. A 2D grid fits naturally with the rendering model built into a GPU and many computations naturally map into grids: matrix algebra, image processing, numerical discretizations of PDEs. A simple example is multiplying each value in the stream of data by a constant (increasing the brightness of an image). GPUs have already been tested for general purpose computing in various applications in computational science. The most recent Top500 list[31], which ranks the most powerful supercomputers in the world, illustrates how GPU technology is at the forefront: the fastest system worldwide in theoretical performance is the Chinese Nebulae, based on GPU technology with 2.98 petaflops. This trend is expected to continue and there seems to be an increasing community of researchers believing that future high-performance computing will be heterogeneous (see §2, §4.2, §4.3, §4.6). This means increased challenges on the programming side, but also increased rewards at the performance side.

Applied and Fundamental Research to develop eInfrastructure

The contribution of mathematics, statistics and information science to the solution of the grand challenges of our time is seen in the multidisciplinary context of eScience, where mathematicians interact with other scientists to exploit, develop and redefine eInfrastructure. The interpretation of results produced by computer simulations for a particular application is undertaken as a collaborative effort between mathematicians and other scientists. In this process important physical features of the solutions are identified, accuracy of the simulations is assessed, and undesired artefacts from the approximation processes are eliminated. The verification procedure consists of comparing computer simulations with experimental data. Verification of advanced computer tools is also accomplished when events that occur in reality are *accurately* reproduced at the computer, as for instance tsunamis and extreme weather events. Accuracy is a well defined concept in Mathematics and can be measured; the mere reproduction of a physical event by an apparently faithful computer-generated image is not true verification. The process of rendering visually large amounts of correctly calculated and certified numerical computer output, is an important new sub-discipline of simulation technology and is known as *visualisation*.

Various branches of mathematics contribute in their own ways to eInfrastructure: differential geometry, harmonic analysis, optimisation theory, dynamical systems, partial differential equations and numerical analysis. Expertise will continue to be demanded from very different branches of mathematics for simulation technology and eScience.

The importance of mathematics in eScience extends beyond direct interaction; sometimes abstract mathematical theories find unexpected areas of application. Certain mathematical constructions are universal and have long lasting effects on science and technology. The theory of Lie transformation groups developed in the nineteenth century is an example. Lie groups gained importance in physics and other sciences throughout the years and will continue to play a fundamental role in the future. This field has just begun influencing computational science and, for example, it is still not clear how to reproduce exact symmetry transformations on computers.

Addressing the foundations of computational mathematics is a safe investment for the future of eInfrastructure. Deeper understanding of the foundations is likely to bring us genuinely new and revolutionary interpretations of eInfrastructure, new technology and new algorithms. To safeguard spontaneously emerging connections of mathematics to eScience, the European Commission has recently published the INFRA-2010-1.1.28 Call for Research Infrastructures, addressing Infrastructures for *Mathematics and its interfaces with science, technology and society at large*. We hope such auspicious initiatives will be of inspiration to the Norwegian authorities.

Conclusions

A variety of the latest high performance computing facilities and the newest technologies within eInfrastructure are required to enable and inspire modern research in mathematics, statistics and information science. For example research on general purpose computing on GPUs has attracted a lot of attention in recent years. The hardware in this field is developing very quickly, and adequate funding schemes are necessary to be able to cope with this development.

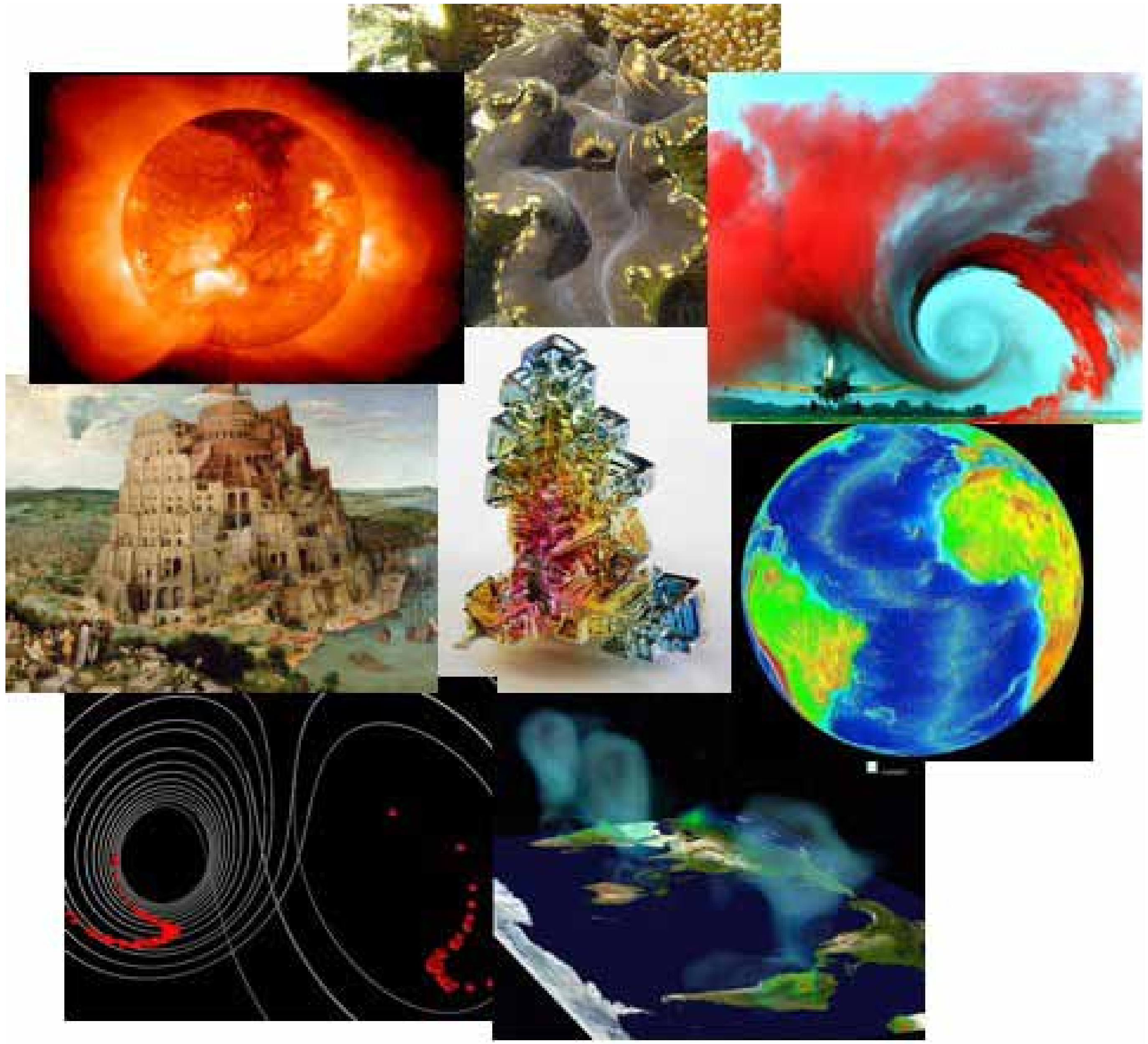
Mathematicians, statisticians and computer scientists at all Norwegian universities and research institutions will continue to actively contribute to eScience in various ways. The new resources and algorithms will allow for the solution of more complex mathematical problems and equations with impact on other sciences. We will witness significant breakthroughs in uncertainty quantification of complex systems with important practical implications in oil and offshore engineering, climate, biology, medicine and other fields. New algorithms, and better ways of exploiting eInfrastructure will emerge from the synergetic interaction of mathematics and statistics with information science and a variety of applications.

---

[30] Computational Science and Engineering: Challenges and Opportunities. Contributions Towards a Centre for Computational Science and Engineering at NTNU and SINTEF. September 2010.

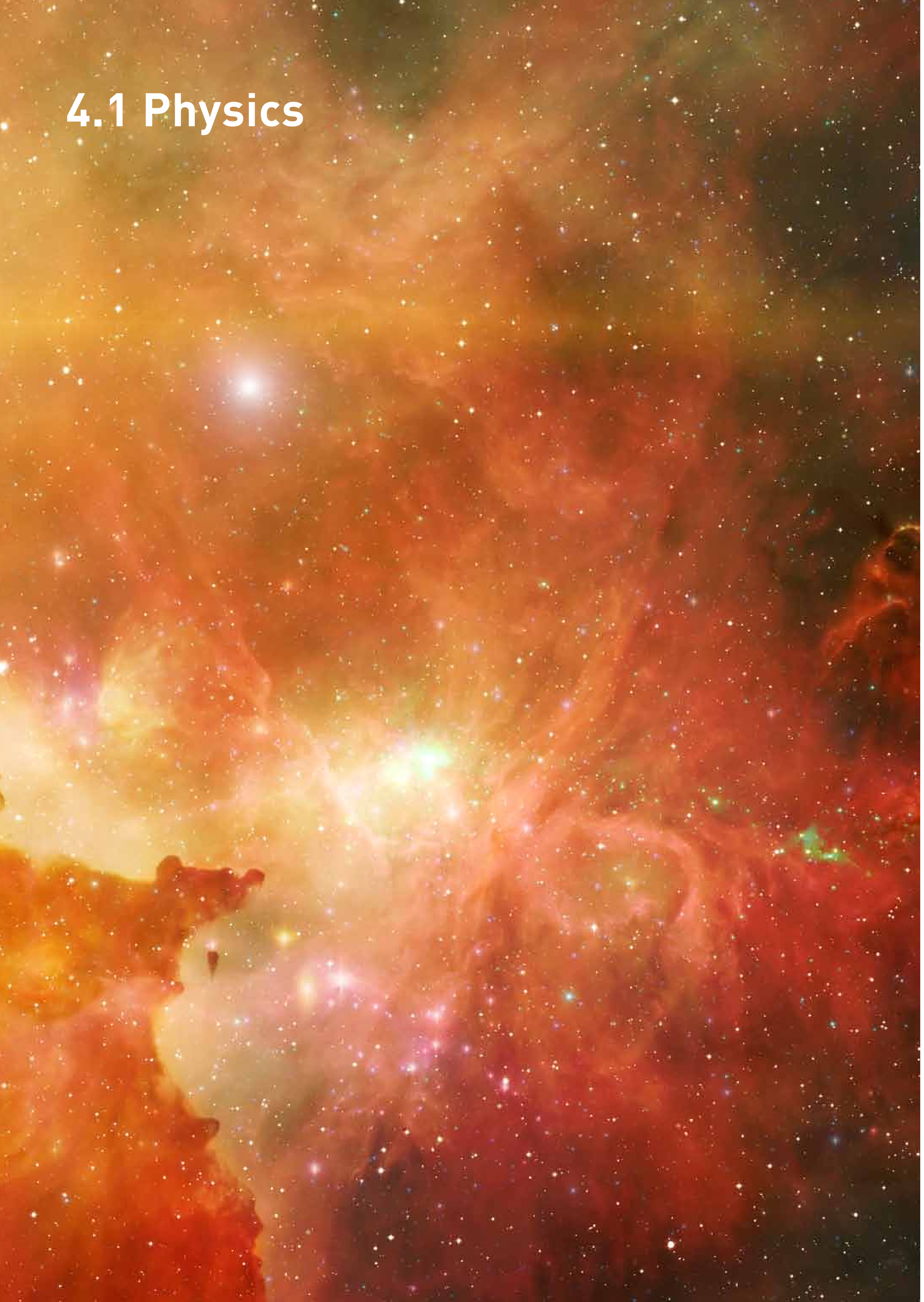
[31] University of Mannheim, University of Tennessee, and NERSC/LBNL. Top 500 supercomputer sites. http://www.top500.org/

4. Scientific Areas



In the following subsections, we describe specific areas of scientific endeavour, attempt to illustrate the breadth of current research in Norway concerning use of eInfrastructure, and their anticipated needs over the next decade. We highlight their socio-economic relevance, scientific challenges, and expectations for the future.

4.1 Physics



High energy particle and nuclear physicists developed Monte Carlo methods to describe the interactions of various types of radiation with matter. Such codes are now common in other disciplines. Monte Carlo modelling of medical radiation offers significant improvements in health care. In telecommunications, Monte Carlo methods are used to design efficient networks, and in finance, it is used to reduce the uncertainties of transactions and to improve forecasts needed by insurance companies.

Summary

The laws of physics are the basis for understanding the behaviour of matter at all levels of complexity, from sub-nuclear particles to proteins, from cosmic radiation to the formation of clouds, the origin of galaxies, and the evolution of the Universe. Understanding the nature and interaction of matter is the basis of technology. It is hard to overestimate the societal consequences of basic physics discoveries. *"One day sir, you may tax it"* was the famous reply of Michael Faraday to William Gladstone, British Chancellor of the Exchequer, when asked of the practical value of electricity.

High performance computing is vital in many branches of physics. Physicists perform massive data analyses, large data transfers, and simulations. All aspects of eInfrastructure are important, and physics is a test bench for many developments in processors, storage, and networking. Large and costly international laboratories and facilities collect and transfer data all over the world, where they are stored and analysed in hundreds of places by people collaborating with each other. This led, at CERN[32], to the birth of the World Wide Web and later to grid computing.

Solar physics

The Sun, our nearest star, is the source of the energy that makes life possible on Earth. Understanding how it works in detail is beyond the reach of our calculations. It generates energy through nuclear-fusion reactions in its deep interior. The heat from those reactions moves outward from the core, driving the strong convection currents coursing through the outermost layers. These currents interact with the dynamo that produces the Sun's magnetic field, creating the solar atmospheric phenomena: sunspots, flares and prominences, coronal mass ejections, and the solar wind.

Studying these phenomena requires substantial computer resources. Presently the best that can be done is to simulate a rather small portion of the atmosphere: a cube 20 thousand kilometres on a side at a resolution of 20 kilometres. Solving the full magnetohydrodynamic equations with radiative heat transport in this cube requires Petaflop-scale resources. Astrophysicists at UiO currently use Notur resources and the Pleiades computer at the NASA Ames facility in California to perform these calculations. Their use of these machines is about 16 million core-hours per year, and they generate simulation data at the rate of 9 Terabytes per hour of simulated solar time. Analysis of the simulation data itself requires supercomputer time; an interactive immersive visualisation system is needed to fully exploit the results.

These computations suffer from having to specify artificial boundary conditions for the magnetic field produced by the solar dynamo. Calculations of this kind cannot provide any information about how changes in the interior affect surface activity; massive parameter studies only investigate sensitivity to boundary conditions. Coupling solar atmospheric simulations to solar dynamo calculations running simultaneously might just be possible on Exaflop machines.

Oslo astrophysicists are also involved in observational projects using satellites and ground based telescopes. They are partners in the Japanese Hinode[33] satellite, in orbit since 2006, and host the European data centre for that satellite. They will likely do the same for a NASA satellite to be launched in the next few years, so data handling and archival storage are also priorities.

High Energy Particle and Nuclear Physics, and Astroparticle Physics

High energy particle and nuclear physics, and astroparticle physics seek to answer basic questions about the nature, origin, and ultimate structure of matter in the Universe. Tools of the trade are accelerator systems tens of kilometres long producing high energy beams to probe matter at microscopic scales; and hundred-megapixel, eight-storey-high detectors, filled with electronics and processors, measuring with microscopic precision particles produced in high energy collisions.

In Norway, large research groups at UiB and UiO are active in the ATLAS[34] and ALICE[35] experiments at the Large Hadron Collider[36] at CERN and are members of international collaborations, numbering thousands of researchers in hundreds of laboratories all over the world. The management and distribution of data are critical tasks facing the collaborations. These data, which are proprietary to the collaborations, are a treasure chest of unprecedented discoveries, and are replicated to scientific computing centres around the world for analysis by all members of the collaborations.

The LHC will study collisions of proton and heavy ion beams, ultimately of energy seven times higher than has been achieved in a particle accelerator. When the colliding beams of protons are made more intense, the interesting data rate will be increased, making pattern recognition, analysis and simulation of each event more compute intensive. Fig. 7 illustrates the tasks of finding traces of interesting particles in the collisions at relatively low detector occupancy.

The ability of LHC physicists, including Norwegians, to analyse the data in an efficient and competitive way depends on the capabilities of the underlying eInfrastructure: the hardware, software, networking and very importantly, the professional services to move and process the data. The internationally connected infrastructure must provide services to all members of the collaborations; and it must cope with frequent updates of proprietary scientific software and related auxiliaries. This is not possible without qualified professional staff able to work in a diverse international environment.

---

[32] CERN is the European Organisation for Nuclear Research, see public.web.cern.ch/public

[33] hinode.nao.ac.jp/index_e.shtml

[34] atlas.ch

[35] aliceinfo.cern.ch/Collaboration/index.html

[36] lhc.web.cern.ch/lhc/. The LHC is a circular particle accelerator 27 km in circumference. Phase 1 began at CERN in March 2010, and it is already setting world records for the highest-energy manmade particle collisions.

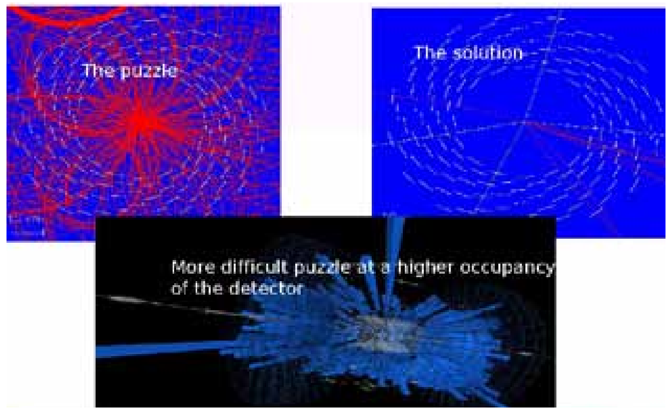


Figure 7. Illustration of the pattern recognition challenge in LHC detectors. Top: finding high energy tracks in Compact Muon Solenoid detector central tracker; bottom: occupancy of ATLAS detector with 2009 beam-splash events.

High energy physics leads the world in eInfrastructure use. Computer simulations are used for the design of detectors and accelerators, the electronics, data collection, and data filtering systems, and for the analysis and interpretation of the results. The data handling problem is immense: to every byte of collected data corresponds nearly a Megabyte of simulated data. So too are the problems of collaboration: thousands of scientists spread worldwide need the best tools for working together.

High energy collisions recreate conditions that were prevalent picoseconds after the Big Bang. Only around 4% of the Universe is built of known matter[37] and how this came to be is unknown. On the way to making the matter we know, the Universe went through the state of a *quark-gluon plasma*: a primordial soup of quasi-free nucleonic constituents, with yet unknown properties. In heavy ion collisions at the LHC this state will be studied, and particles from the earliest times of our Universe will be recreated. Among the products may be particles of the dark matter that makes up a quarter of the Universe's mass, and is responsible for the formation and stability of galaxies.

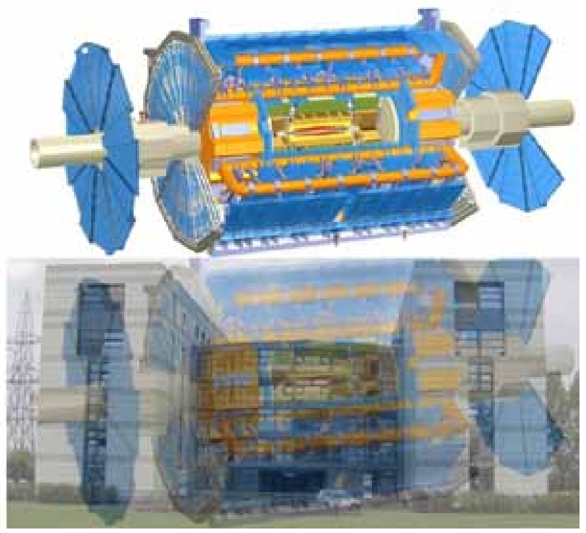


Figure 8. The ATLAS detector and its real size "shadow" on the building hosting ATLAS and CMS[38] experiment offices.

To capture rare events in high-energy collisions the beams must have very high intensity, with a collision frequency so high that one collision takes place before the particles produced in the previous one have left the detector. The data rate is equivalent to over 1 Petabyte *per second* if all event data could be recorded, but this is far beyond the capabilities of current technology. The collaborations have implemented a fast layered trigger system designed to select interesting rare events from the huge background of well understood events. The trigger system uses real-time information to identify the most interesting few events out of the 40 million beam crossings that occur every second in the detector. There are three trigger levels, the first based in detector electronics and the other two in software on a computer cluster nearby. For the ATLAS detector (Fig. 8) a few hundred events per second remain to be stored for further analysis, equivalent to a few Petabytes each year. Individuals and groups within the collaboration write their own codes to perform further analysis, for testing particular physical models or searching for new particles.

Grid computing is extensively used for event reconstruction, allowing the parallel use of university and laboratory computers throughout the world for the compute-intensive task of reducing large quantities of raw data into a form suitable for analysis. Grid middleware provides the ability to manage and access all the collaboration's distributed resources. The software for these tasks has been under development for many years by a large number of people, and refinement continues. Norway plays a leading role in development of grid middleware.

The LHC collaborations have detailed plans of eInfrastructure needs until 2013, but extrapolation to later years is an educated guess (Fig. 9). Processor needs grow slowly after 2013 until the Super LHC era, likely to start in 2020. Addressable memory per core, presently 4GB, will need to increase with the larger event sizes. Large data transfer rates and the geographical distribution of users places high requirements on the links. At present the distributed centres typically operate with 20Gb/s LAN, and 10Gb/s interconnect between the centres. A breakthrough in interconnect technology, if it happens, might lead to a decrease in required storage space.

The Super LHC era will bring a step-like increase in Processor needs and memory-per-core, because of the increased time needed to simulate events matching experimental data. The number of particles per event will increase fourfold, leading to much longer time spent on pattern recognition and interaction simulation. The interesting data volume to be simulated will increase by a similar factor, leading to further increases in processor and storage requirements.

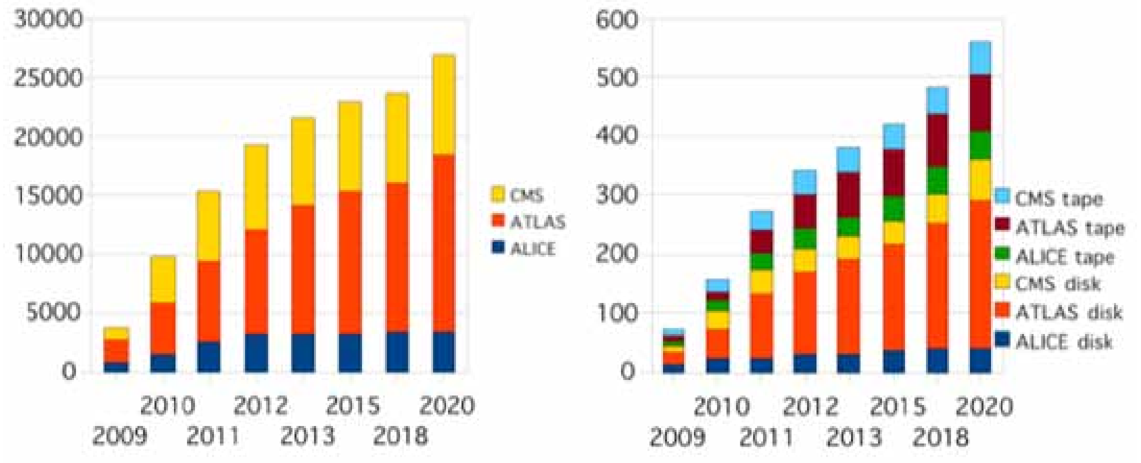


Figure 9. Left, estimated processor requirements for LHC in numbers of 8-core Intel Xeon E5430 processors. Right, projected storage in Petabytes.

Astroparticle physics studies various forms of ultra high energy radiation from the Universe: protons, neutrinos, gamma rays, and antimatter particles, to understand the nature of high energy phenomena in the Universe. Examples of these phenomena are various types of supernovae explosions, activities at centres of galaxies often related to the presence of black holes, and possibly the accumulation and annihilation of dark matter particles leading to high energy radiation. A further goal is to understand the formation of galaxies and large scale structures by simulating the time evolution of dark matter clusters and filaments on the scale of billions of years. The simulations are compared to maps of the tiny variations in the temperature of the microwave background radiation, measured presently by the NASA WMAP[39] satellite. Better accuracy is expected from the European PLANCK[40] satellite. Patterns on various angular scales give important clues to the origins and the overall structure of the Universe. The scientific questions of astrophysics, astroparticle and high energy physics are closely related. Similar detector devices are used and there is similar approach to the use of eInfrastructure. A powerful eInfrastructure facilitates movement of staff and expertise between these disciplines.

Atomic and Low Energy Nuclear Physics

In quantum mechanics, high-performance computing pioneers like Hartree[41] developed both theory and methods for the self-consistent calculations of atomic energies using the day's top computers. Demands have increased since with the development of femto- and even attosecond laser pulses, that can now resolve motion of electrons within atoms. New phenomena and new questions prompt theoretical studies of time-dependent quantum systems, through solving the fundamental time-dependent Schrödinger equation, a linear differential equation with 3*N* space variables (for *N* particles) and one time variable. One of the largest direct calculations to date, involving nearly 23 000 cores on the Cray XT4 in Edinburgh, was for a helium atom in which both electrons are removed simultaneously by absorption of photons from a strong laser beam. Physicists at UiO and UiB are active in computational studies of laser light interactions with atomic systems.

Another application being actively pursued by Norwegian physicists is the theoretical and experimental study of quantum dots. They exhibit macroscopic manifestations of quantum mechanical phenomena and have applications in areas as diverse as biomedical imaging, quantum computing, photovoltaics, and light-emitting devices. These applications require machines with large on-core memory and ultrafast communications between cores, like the Cray XT4 at UiB.

Calculations of the internal structure of atomic nuclei involves solving many-body problems using iterative matrix techniques where the matrix sizes can be of order many Gigabytes. Physicists at UiO carry out these calculations with American colleagues on the fastest computer in the world, the Jaguar Cray XT5 machine at Oak Ridge National Laboratory in Tennessee. Smaller versions of these calculations can be done on the Cray XT4 at UiB. The same techniques are used to understand BoseEinstein condensates, which are fragile experimental states that are used to investigate problems in fundamental physics.

Requirements for eInfrastructure

For solar physics, and for atomic and low-energy nuclear physics, the tightly coupled nature of the problems require highly integrated supercomputers with fast interconnects and large memory. The more loosely connected problems of high-energy and astroparticle physics requires grids with high bandwidth and high data storage.

Expectations from 2015

The boosting of LHC beam intensities after 2020 will cause a stepwise increase in the requirements in all aspects of eInfrastructure. In other areas, the increase is likely to be more gradual, but physics will probably still lead other areas of science in use and demand for eInfrastructure.

---

[37] *Baryonic* matter, of which stars and planets and everything we know is made, constitutes only about 4% of the mass of the Universe, according to recent studies. The remainder is detectable so far only by its gravitational effects, and comes in two forms, *dark matter* (about 25%) and *dark energy* (about 70%).

[38] cms.web.cern.ch/cms/index.html

[39] map.gsfc.nasa.gov

[40] www.sciops.esa.int/index.php?project=PLANCK

[41] Hartree, D. R., 1957. *The calculation of Atomic Structures*. New York: Wiley & Sons.

4.2 Chemistry and Materials

In 1990, Uzi Landman and co-workers studied the workings of atomic force microscopy (AFM) by carrying out moleculardynamics studies of a tip of nickel atoms hitting a sheet of gold atoms[42]. In doing so, they discovered that gold atoms jumped up to contact the approaching nickel tip; as the nickel tip was subsequently withdrawn, a nanometer-sized golden wire was created (Fig. 10). This result amazed the discoverers but was confirmed experimentally with AFM a few years later. This is an example of the discovery of new phenomena that cannot be predicted in a simple manner from the physical laws that operate on the atomic scale. Even though the laws that govern the atoms and their motions are known, these laws sometimes lead to surprising behaviour that can be revealed only through numerical simulations.

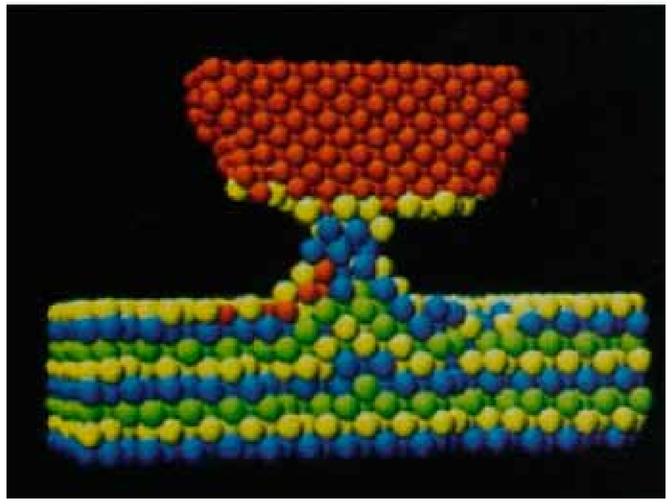


Figure 10. Molecular dynamics simulation of an atomic force microscope. Before the hard nickel tip touches the soft gold surface, gold atoms jump up to cover it; as the tip is withdrawn, a nanowire of gold atoms is formed. Nickel atoms are coloured red; gold atoms from consecutive layers are coloured yellow, blue, green, and so on.

Scientific Challenges

Computational chemistry and materials science began with the development of quantum mechanics in the 1920s, when it was realised that molecules and materials consist of charged particles in motion, governed by the laws of quantum mechanics. Paul Dirac wrote[43] "The underlying physical laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known and the difficulty is only that the exact application of these laws leads to equations that are much too complicated to be soluble." Dirac could not foresee the spectacular emergence of the electronic computer in the second half of the twentieth century, which made a numerical attack on the many-body problem possible. First-principles simulations of chemical processes have become commonplace and are nowadays performed even by non-specialists in support of experimental activities and measurements. Today, about 40% of all articles in the Journal of the American Chemical Society are supported by computation; likewise, at the Department of Chemistry at UiO, about one-third of its staff have written articles supported by computation. This is an remarkable development for a science usually perceived as experimental and empirical in nature. Today, computation not only plays an important role in the interpretation and prediction of experimental observations — it is increasingly viewed as an alternative to experiment. In 1998, the Nobel Prize in Chemistry was shared by two of the founding fathers of electronic-structure simulations: Walter Kohn "for his development of density-functional theory" and John Pople "for his development of computational methods in quantum chemistry", in recognition of the enormous impact that such simulations have had on chemistry and materials science[44].

Computational chemistry owes its importance to the development of accurate electronic-structure models and their implementation in software packages, widely distributed to the community of chemists and materials scientists. The continued development of codes such as the Dalton[45] quantum chemistry code enables simulations of a broader range of physical phenomena on more realistic systems, using ever improving computer platforms. Twenty years ago quantum chemistry was applied to systems containing a handful of atoms; it is nowadays applied to hundreds of atoms, following the development of new electronic-structure models and new numerical techniques. Systems with thousands of atoms can be simulated by using methods that describe parts of the system at a high first-principles level, while the surroundings are described at a lower semi-empirical level. Important problems such as defects in solids and enzymatic reactions can be studied using multiscale simulations with a method hierarchy spanning a broad range of spatial and temporal domains.

While simulations are being performed on larger systems over longer time spans, experiments are being performed with ever higher spatial and temporal resolutions — for example, femtosecond laser pulses are now used to probe reaction mechanisms of simple molecular systems. The computational challenge is then to provide sufficient accuracy to match that of experiments. Today, computational chemistry is capable of achieving accuracies comparable with or surpassing those of measured reaction enthalpies and atomisation energies, but it cannot yet routinely deliver accuracies needed in many spectroscopic studies. Reaction barriers and reaction rates are also demanding, requiring a better description of the motion of electrons in molecules and solids. Indeed, besides the development towards larger systems, the development towards higher accuracy constitutes the grand challenge of computational chemistry and materials science over the coming decade.

The situation regarding computation and simulations becomes more complex in nanoscience. The objects of study, whose building blocks are nanotubes, quantum dots, clusters and nanoparticles, are structures of size up to 100 nm. These have been studied since the 1980s, following the development of a number of new experimental tools such as scanning tunnelling and atomic force microscopes. New emergent phenomena have been discovered, such as the gold nanowire production cited at the beginning of this section. Simulations have proven essential in unraveling the physics of new phenomena such as giant magnetoresistance[46], which is now used in commercial hard disks. In the future, simulations of nanoscience will continue to play an important

---

[42] Reproduced from U Landman, WD Luedtke, NA Burnham, RJ Colton, 1990. Atomistic mechanisms and dynamics of adhesion, nanoindentation, and fracture, *Science*, **248**, 454-461.

[43] P. A. M. Dirac,1929. *Proc. R. Soc. London Ser. A* **123**, 714.

[44] nobelprize.org/nobel_prizes/chemistry/laureates/1998/

[45] www.kjemi.uio.no/software/dalton/dalton.html. Dalton is developed by Norwegian scientists and others.

[46] P. Grünberg, R. Schreiber, Y. Pang, M. B. Brodsky, and H. Sowers (1986). Layered magnetic structures: evidence for antiferromagnetic coupling of Fe layers across Cr interlayers" *Physical Review Letters* **57** (19): 2442–2445.

role in unraveling the secrets of nanostructures. The report from a 2002 US Department of Energy Workshop on "Theory and Modelling in Nanoscience"[47] concluded "investment in the national nanoscience initiative will pay greater scientific dividends if it is accelerated by a new investment in theory, modelling and simulation", warning that the "absence of quantitative models that describe newly observed phenomena increasingly limits progress in the field." The fundamental challenges and opportunities in simulations are in the areas of complex nanostructures and nano-interfaces and the growth of such structures. Components that need improving include the electronic-structure methods, methods for classical and non-classical molecular dynamics, and Monte Carlo methods. Simulations of nano-sized systems need to deal with widely different length and time scales and to treat systems involving a variety of interactions that have traditionally been treated by different methods and techniques.

Requirements for infrastructure

New developments in computer technology will make it possible to perform simulations in ten years time that are unthinkable today, provided we are able to utilise in an efficient manner the power of massively parallel computers. Recent benchmarks on Argonne's Blue Gene/P[48], demonstrate scaling of the Dalton code to over 20 000 cores. However, to take full advantage of tomorrow's technology in all aspects of simulation, existing codes must be upgraded and rewritten, sometimes from scratch, for efficient calculations on thousands of processors. Such work is already underway in a number of research groups and will revolutionise simulations in chemistry and materials science over the next decade. In view of such developments, ready access to massively parallel computers is essential for Norwegian theoretical and computational chemists to stay at the vanguard of their science.

Quantum-mechanical simulations are computationally demanding, but have low requirements on data storage[49] for input and output. Graphical processing units or GPUs are becoming widely used and may accelerate central computational tasks by an order of magnitude. Memory requirements vary; many simulations can be carried out using a few gigabytes of memory, while some require orders of magnitude more; terabytes of memory are being used in a few recent applications. Some methods and algorithmic developments depend on fast interconnects with large bandwidths and shared-memory architectures; the need for such architectures will therefore increase. On the whole, however, quantum mechanical simulations are flexible in that they can be adapted to a wide variety of computers[50].

A critical factor for all simulations in chemistry and materials science is high computer stability. Many production applications require a week or more of computing time, even when fully parallelised on clusters. Typically, such calculations cannot be restarted without considerable loss. Maintenance and optimisation of production codes on specific platforms is important work that cannot easily be undertaken within small research groups. Such work is best performed in support groups within the national supercomputing framework.

Expectations from 2015

Over the next five years, computational chemistry will shift towards large parallel computers. Codes are being constantly improved with faster algorithms — for example, the current experimental version of Dalton calculates electronic energies and molecular forces one to two orders of magnitude faster than the currently released version. Such improvements never reduce the needs for computing power; the increased efficiency is used instead to improve the quality of the simulations. Even with vastly improved codes and computers, production runs will typically stretch over one or two weeks, using thousands of processors — only in this manner will it be possible to carry out research at the forefront of chemistry and materials science.

---

[47] www.cs.odu.edu/~keyes/scales/reports/nano_2002.pdf

[48] With 294 192 PowerPC 450 850 MHz processors, designed to run continuously at 1 Petaflop.

[49] Some applications can gain from abundant scratch storage that is removed upon the completion of the run.

[50] The Centre for Theoretical and Computational Chemistry, a Centre of Excellence funded by the RCN, used 18.000.000 core hours on Stallo, 1.500.000 hours on Titan, and 250.000 hours on Hexagon and Njord in 2009.

4.3 Fluid Dynamics

Computational Fluid Dynamics (CFD) is used in a large number of scientific areas including turbulence modelling, complex physics, chemistry, geosciences, medicine, engineering and meteorology. Examples include oil spills in the ocean, spreading of volcanic particles in the atmosphere, study of the flow of blood and spinal fluid in the human body, computation of wave forces on ships, three-phase flow in oil pipelines and separators, air flow around cars, sailboats, airplanes, and wind turbines, oil/air/water flow and temperature in car engines, ventilation in buildings, etc. The current chapter describes some areas in which state-of-the-art research in fluid mechanics is being conducted at NTNU, and their requirements for eInfrastructure in the future.

Numerical modelling of particles in turbulent flow

Almost all natural and industrial fluid flows are turbulent, characterised by an unsteady and irregular fluid motion. Turbulent flows of liquids and gases are diffusive, dissipative, intrinsically threedimensional and non-linear. They are nevertheless governed by the Navier-Stokes equations. Suspensions of solid particles in gases and liquids are of practical concern, as in the transport of solid particles or aerosols in the atmosphere, of microorganisms in the ocean, or of sediment in rivers.

Understanding the role of particles in industrial fluid flow can improve energy efficiencies of machinery and help to reduce emission of harmful pollutants in combustion engines. Sediment particles in river flows affect artificial structures such as dams and bridges. Learning about the mechanics of sediment transport leads to better appreciation of the environmental impacts of such structures and helps prevent accidents.

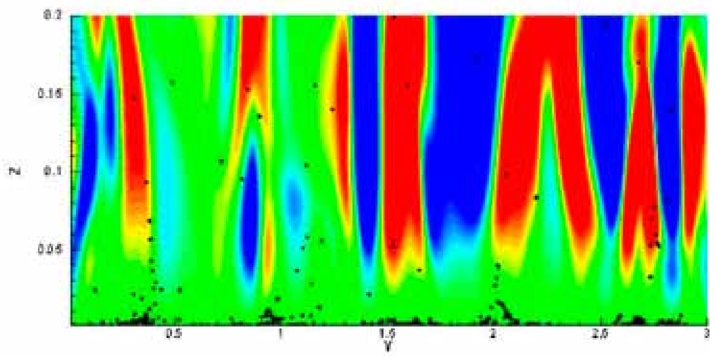


Figure 11. A cross-sectional instantaneous view of near-wall flow, with tiny spherical particles (black dots) embedded in a turbulent flow field. The wall is at the bottom and the primary flow is directed out of the paper. The colours reflect turbulent variations in the velocity field. The particles cluster in certain areas, with implications for energy loss in mixed flows.

Particles in industrial flow

Particles embedded in a fluid are usually modelled as spheres. The motions of non-spherical particles[51] are by far more intriguing with the further complexity of rotational motions. Tiny particles tend to cluster in some preferred areas of a turbulent flow field, as seen in Fig. 11, depending on particle mass and shape. Feedback from the particles on the flow field depends on the relative concentration. To study this, simulations that are larger than those done today are required. Computers with greater speed and storage capacity are needed for increased particle loading and Reynolds number[52].

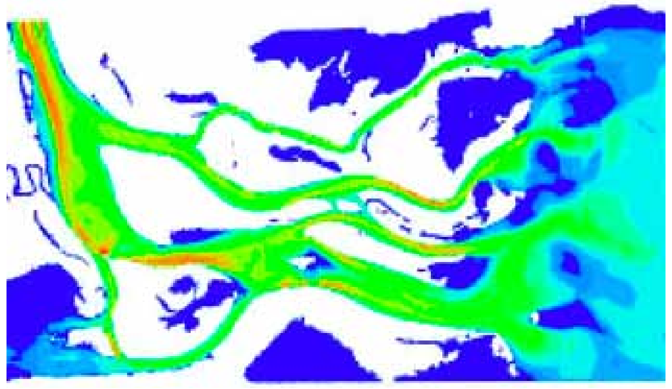


Figure 12. Plan view of computed velocities in the delta of Lake Øyern, the largest freshwater delta in Northern Europe. Red colours represent high velocities and blue colours low velocities. The water flows from top left towards the right.

Modelling sediment transport in deltas

Hydropower regulations can cause geometrical changes in the deltas of the reservoirs. Sediment particles are deposited at the mouths of rivers, increasing the bed levels. When the deposits get too high, the river seeks an alternative path. A complex system of channels, a delta, is formed. The pattern depends on the sediment properties, the water discharge rate, and the water level in the lake or ocean. Secondary currents and recirculation are further complications. Deep scour holes can occur[53], and vegetation can affect the sediment deposition[54]. Numerical models are used to calculate the environmental implications of such changes over a few weeks time (Fig. 12). These computations are currently at the limit of the capabilities of existing computer clusters in Norway. It will be very useful to model much longer times, up to 20-30 years. Similar techniques are also useful for geological reservoir modelling, and petroleum companies are currently investing considerable research efforts in this area. However, such computations will require many orders of magnitude more computing power than available today.

Combustion and turbulence

Most of the world's energy is produced by the combustion of biomass, coal, natural gas, and oil. A considerable fraction of air pollution and man-made greenhouse gases come from this source. The challenges are to use these fuels more effectively and reduce the polluting emissions. This requires more detailed knowledge about the combustion processes.

Tremendous progress in combustion science and technology in the past two decades has been made possible by advances in computational resources. We can now simulate the details of some simplified combustion phenomena. Engine manufacturing industries routinely employ tailored combustion codes for their development work, based on limiting simplifications. But simple descriptions of chemical reactions result in unreliable predictions of emission levels. Improving the models with greater fidelity under realistic combustion conditions is absolutely crucial to guide the manufacture of efficient and clean fuels. See also §4.9 for the industrially relevant autoignition problem.

---

[51] For example, man-made carbon nanotubes, which could be hazardous upon inhalation.

[52] The Reynolds number measures the relation between inertial and viscous forces in a fluid. At high Reynolds number (when inertial forces dominate) flows are likely to be turbulent.

[53] Eilertsen, R., Olsen, N. R. B., Ruether, N. and Zinke, P. 2008. River bed changes in the Lake Øyern delta distributaries over a three-year period (2004 - 2007) as revealed by interferometric multibeam, International Geological Congress (IGC2008) Oslo, Norway.

[54] Zinke, P. and Olsen, N. R. B. 2009. Numerical modelling of water and sediment flow in a delta with natural vegetation, 33rd IAHR Congress, Vancouver, Canada.

Gaseous combustion

The physical and chemical laws of gaseous combustion are reasonably well known. These include the equations of motion, mixing, heat and mass transfer, chemical reactions, and the interaction between chemistry and gas flow. In principle, the coupled set of equations can be solved with a computer and suitable numerical tools. The complexity of turbulent flames allows direct numerical simulation (DNS) to be done only for very simple cases[55]. The results of such numerical experiments can be used to supplement experimental data in the development of simple combustor designs.

The length and time scales in turbulent flames (Fig. 13) span many orders of magnitude. A combustor can have a length of 1 meter and an overall mixing time of 1 second, while the reactions occur on scales of 1 micrometer and 1 nanosecond. The computational requirements of turbulent combusting flow are orders of magnitude larger than for a non-reacting flow. The required computing power for DNS calculations is extreme and requires a cluster with one million cores or more to model complex geometries.

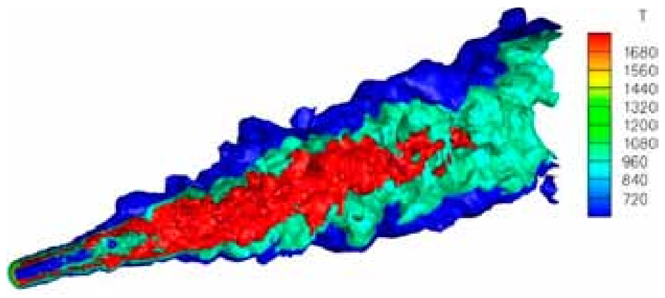


Figure 13. A temperature field from a Large Eddy Simulation of a turbulent flame, showing details of the large-scale eddying motion of the flow. Acoustic waves and stability problems can be diagnosed with simulations like this.

Multi-scale methods in Computational Fluid Dynamics

Multi-scale modelling will play a key role in the future development of computational fluid dynamics[56]. For example, the interface between two fluids, or between a fluid and a solid, is strongly affected by surface molecular effects. A multi-scale test case is the impact of a droplet on a flat plate, which has applications in ink-jet printing, spray cooling, pesticide spraying, erosion processes due to rain, and others (Fig. 14). A macroscopic model computing the main flow field using the Navier-Stokes equation and a microscopic model for the surface forces are coupled. There is a movable contact line, namely the gas-liquid interface in contact with the solid surface. High Weber numbers[57] increase the complexity of the flow, because the droplets break up and splash. It is possible to compute the process in 2D using current computers, but a 3D computation will require substantially more cores. A system with 100 000 cores should be sufficient for a 3D solution.

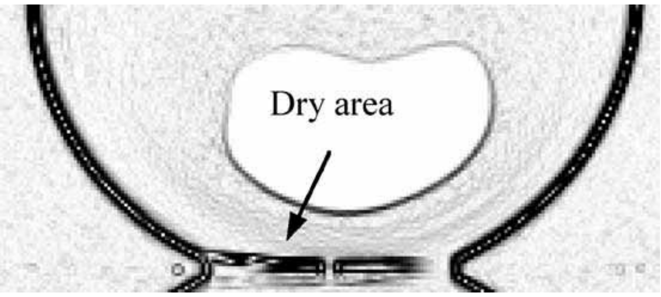


Figure 14. The contact line dynamics for moderate Weber number[58] in the impact of a droplet on a plate.

Scientific challenges and solutions

When a fluid dynamics code is used on a parallel machine, the work is divided among the different processors. Whether a shared memory model or a distributed memory model is used, fast connections are necessary to transport data among the parts of the grid. For very large numbers of cores achieving good scaling is a challenge. CFD programs on current clusters may not scale well above 1000-10 000 cores, but increased bus capacity on future systems will help. It may also be possible to develop numerical algorithms where the amount of data transfer between the cores is reduced.

Requirements for eInfrastructure

The main eInfrastructure requirement for fluid mechanics is raw computational power. A simulation of detailed turbulent fluid flow always improves its accuracy with more computing power. Progress in this science is therefore highly dependent on the available resources. Computational fluid dynamics programs mainly use distributed memory systems with very fast interconnects. A system with between 100 000 and 1 million cores would be required to address problems in turbulence, combustion, and particle transport. Good computer stability and uptime are highly desirable.

Expectations from 2015

It is expected that the numbers of cores in future clusters will continue to increase. The need for computing power in fluid dynamics is effectively unlimited. With more available power, we can compute with finer grids and resolve flow fields and turbulence in greater detail. The accuracy will increase, and it becomes possible to compute new phenomena that had not been possible earlier. The numerical algorithms used in CFD will mature. Unstructured polyhedral grids, finite-volume methods and multi-grid solvers will be more widely used, and open source CFD programs like OpenFOAM may become more prevalent.

---

[55] Gruber, A. (2006) Direct numerical simulation of turbulent combustion near solid surfaces. Doctoral thesis 2006:14, Norwegian University of Science and Technology, Department of Energy and Process Engineering, Trondheim.

[56] Müller, B. and Walker, C. work in preparation.

[57] The Weber number measures the ratio between inertial forces and surface tension. At high Weber numbers, when inertial forces dominate, a flow is likely to disintegrate into a spray.

[58] S. Sikalo, C. Tropea, E.N. Ganic, 2005. Dynamic wetting angle of a spreading droplet. *Experimental Thermal and Fluid Science* **29**, 795-802.

4.4 Climate Science and Weather Prediction

In March 2010, the European Centre for Medium-Range Weather Forecasts[59] reported that their primary measure of forecast skill[60] surpassed the 10-day mark, for the first time ever. This landmark achievement highlights the enormous strides in reliability of numerical weather predictions since the first computerised forecasts were made in the early 1950s. This advance is the direct result of the interplay between improved scientific understanding and efficient use of computational resources. On average, the weather is now predicted 8 days into the future with the same skill as for 5 days in 1980. Forecast reliability is also predicted, along with probabilities for the occurrence of various weather events. Similar progress has also been achieved for predictions 2-3 days ahead, with stronger skill requirements, particularly concerning accurate prediction of extreme events.

Of supreme importance for future planning and policy making is the understanding of how our climate will evolve in time. The climate system involves complex interactions of the atmosphere, the oceans, the cryosphere[61], the biosphere, and the land surface (Fig. 15). Computer-based climate simulations are the only viable approach to understanding the climate system. These models also constitute the ultimate test of our understanding of processes determining the statistical behaviour of weather. Non-linear and unstable processes eventually lead to loss of weather predictability because inaccurate representation of small-scaled processes propagates to larger scales. Scale interactions thus constitute a major source of error in climate projections. Tropical convection, which plays a key role in driving global atmospheric circulation, is an important example.

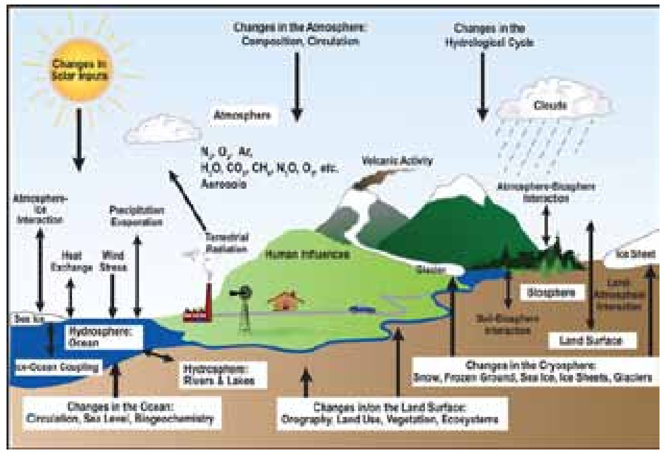


Figure 15. The climate system includes all processes interacting among the subsystems of the fluid earth, including atmosphere, ocean, ice, the biosphere, and the land surface.

Predictions and projections

Weather prediction was the first non-military scientific problem addressed in computational science in the 1940s. Vilhelm Bjerknes, 50 years before, had already developed its scientific basis by combining the equations of fluid dynamics and thermodynamics for the atmosphere and the oceans. If these equations could be solved, future development from a given known state could be calculated. But the equations cannot be solved analytically, so computers are necessary.

Early on, meteorologists used computers for increased understanding. Formal limits for weather predictability were established by Lorenz[62], who defined *chaos* as critical sensitivity of a time-dependent process to its initial condition: a tiny initial error grows over a finite time until predictions become worthless. There are two distinct types of predictions: *Predictions of Type* 1 are typically weather *predictions* based on accurate knowledge of the initial state; and are skilful if they provide more information than climate statistics. *Predictions of Type* 2 are typically climate *projections*, which estimate changes in the occurrence of all possible weather statistics resulting from a change in external forcing[63].

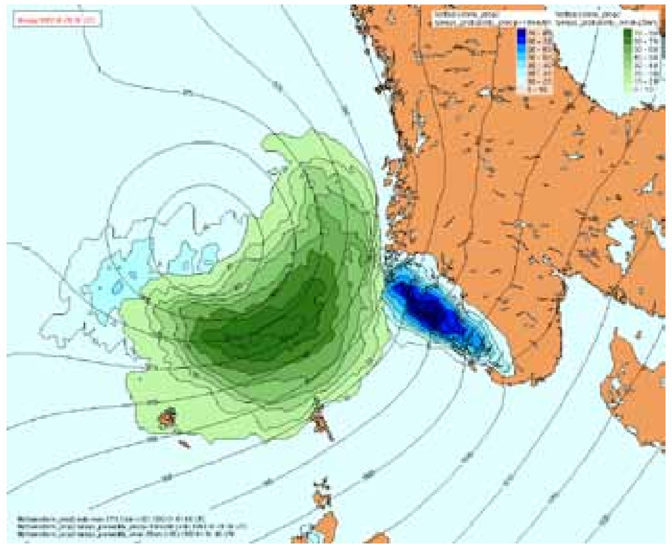


Figure 16. Probability of wind speed exceeding 25 m/s (green) and precipitation exceeding 10mm over 6 hours (blue) predicted 36h ahead of the time when a hurricane force cyclone hit NW Norway 1st January 1992.

A useful weather prediction for a given location consists of three major elements: (1) a consensus prediction that contains all possible predictable features; (2) information about the uncertainty of the consensus; and (3) expected probabilities of particular events (see Fig. 16, for example). *Ensemble predictions* consisting of several individual predictions for the same time are developed to accomplish this, starting from slightly different initial data or with different algorithms or physics included. Ensembles typically contain around 50 members. Single deterministic predictions with greater detail are also used, but miss uncertainties and probabilities.

Ensemble predictions for 10 days have been produced operationally at ECMWF for more than 15 years, with monthly, seasonal,

---

[59] www.ecmwf.int/. ECMWF is an intergovernmental organisation supported by 31 states. See their press release "Landmark in forecast performance", www.ecmwf.int/publications/cms/get/ecmwfnews/1268389540174.

[60] The correlation coefficient between forecast and observed 500 hPa height anomalies being greater than 60%.

[61] The cryosphere consists of ice sheets, glaciers, sea ice, and permafrost.

[62] E. N. Lorenz, 1963. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*. Vol.**20**, 130—141.

[63] External forcing is typically ignored in weather forecasting, while the initial state is irrelevant for climate projections. Intermediate cases, considering both initial conditions and external forcing, are seasonal forecasts and decadal predictions.

and annual forecasts increasingly appearing. The Norwegian Meteorological Institute (met.no) has run a high-resolution limited-area ensemble system for the 2-3 day range since 2005. Ensembles are also used in climate projections, which are time-dependent responses to scenarios that change the driving conditions of the climate system[64]. A single model gives one estimate of the climate system's response to a change in forcing, but an ensemble of models can estimate the likelihood of changes in the frequency distribution of all relevant climate parameters.

Models for climate projections

Earth System Models (ESMs) aim to integrate all important interacting processes across the climate system (Fig. 15), while traditional Climate Models (GCM for global, RCM for regional) include only the *physical* interactions among the subsystems, with chemistry and bio-geo-chemistry prescribed. Such models formed the basis of the Fourth IPCC Assessment report, delivering global climate projections for a variety of greenhouse gas emission scenarios at a range of future time scales (Fig. 17). ESMs are considerably more complex than GCMs, as they include a wealth of possible feedbacks among the physical, chemical and biological components of the climate system, rendering them computationally more demanding than GCMs. For full participation in the international CMIP5[65] experiment, a model needs to produce approximately 10 years of simulation per day of calculation. Reduction of model complexity and resolution is necessary to achieve this with present supercomputers.

While ESMs are necessary for climate projections of a century or longer, GCMs with fewer interacting processes can be run with higher spatial resolution for predictions up to a few decades. High-resolution models provide improved information on climate variability associated with high-impact weather events as well as regional effects of topography and ground surface heterogeneity.

Small-scale processes in the climate system, like cumulus convection and cloud-radiation interactions, have significant impact on large-scale circulation. With a horizontal grid spacing of ~100 km, present-day GCMs cannot resolve cloud formation, cumulus convection, and a number of other important processes. These unresolved processes are represented in a highly simplified manner; the statistical effects of such processes within a model grid box are controlled by, and communicated back onto, the resolved large-scale variables, which the model formally predicts. Cumulus convection, which is important in driving global circulation, has proven to be very difficult to represent in this way. Simulations capable of resolving convection thus promise significant improvements in our ability to simulate the global climate system. Decreasing the grid spacing of a GCM from the present-day 100 km scale to the 1 km scale necessary for resolving convection would increase the computational cost by a factor of 1 million[67]. This bottleneck has major consequences for the best use of computational resources in the foreseeable future.

The Norwegian Earth System Model (NorESM), developed by the Norsk Klimasenter[68] in collaboration with the National Centre for Atmospheric Research[69] in the USA, will be Norway's contribution to the next IPCC assessment. Forced at the boundaries by GCMs, Regional Climate Models are run at considerably higher spatial resolution, to provide detailed local climate information consistent with the large scale climate simulated by the driving GCM. RCMs are run in Norway at met.no, the Bjerknes Centre for Climate Research, and the Institute for Marine Research. These models are used to investigate potential impacts of future climate change and to develop adaptation strategies for Norway.

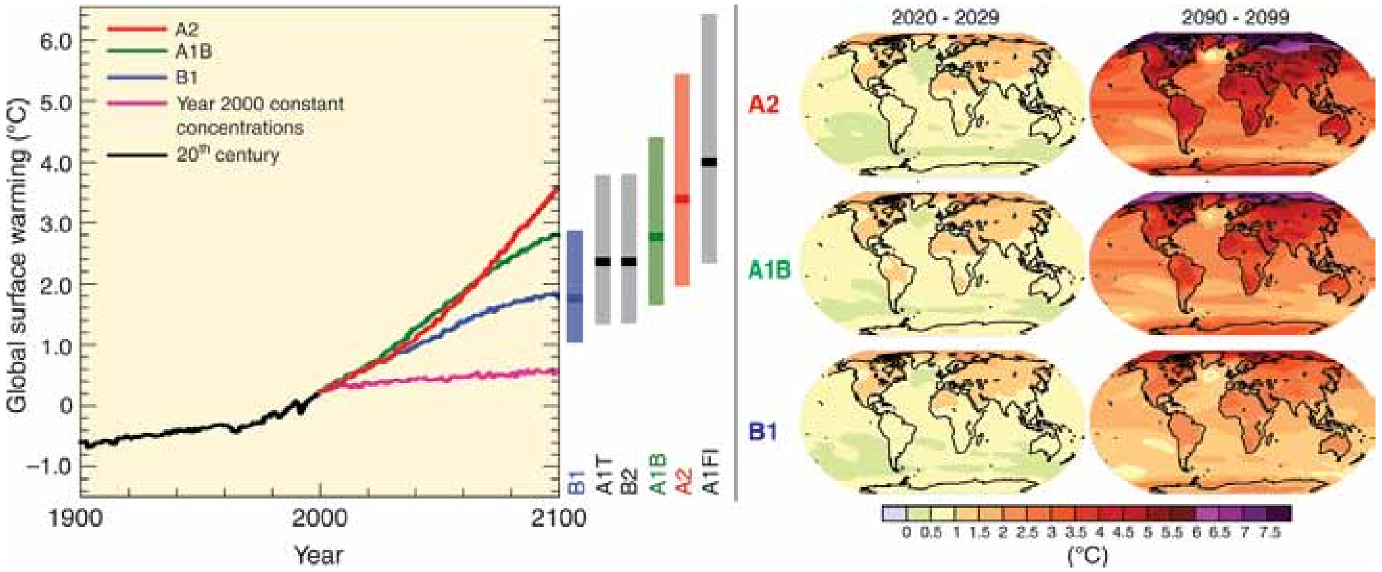


Figure 17[66]. Left: Solid lines are multi-model global averages of surface warming for different scenarios as continuations of twentieth century simulations, with best estimate and likely range. Right: Projected surface temperature changes for the early and late twentyfirst century.

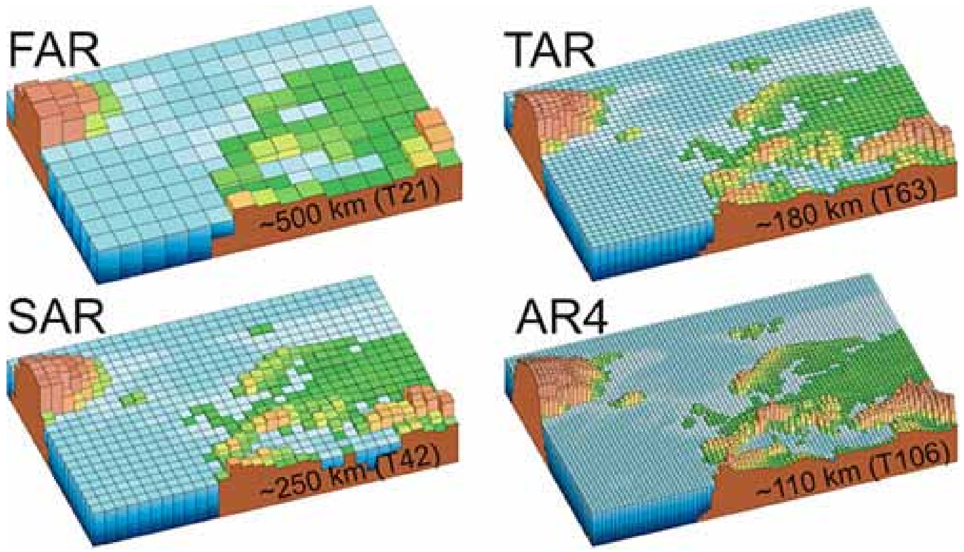


Figure 18. Improving resolution greatly enhances the ability of a model to predict local processes that can affect global circulation.

Prediction models

Models designed for weather prediction share many basic similarities with ESMs, but with reduced complexity and higher spatial resolution. Global forecast models are run in a few centres, with ECMWF as the most important source for Norway. For short-range atmospheric forecasting Norway relies on the HIRLAM[70] model and the UK Met Office's Unified Model[71]; for ocean forecasting, on the ROMS[72] model coupled to the met.no sea-ice model. Coupling ocean and atmospheric models may significantly improve forecasts since deep Arctic waters are occasionally mixed to the surface by strong winds. Accurate characterisation of the initial state is achieved through advanced data-assimilation systems, which can be as computationally demanding as the forecast model.

Models for air quality

Through NILU[73] and met.no, Norway has a central role in European work on air quality modelling. Models have been developed for problems ranging from acid rain and eutrophication to photo-oxidants and airborne particulates (including the volcanic ash from the recent Eyjafjallajökull eruption). The models run off-line with meteorological data imported from weather forecast models. Air quality forecasting is operationalised through the MACC[74] project coordinated by ECMWF. Output from national models are fed into integrated assessment models which estimate optimal plans for emission reductions to achieve politically defined targets, such as the European emission protocols. There is a growing interest to couple air quality models with climate models to develop coordinated emission policies to address future mitigation options for air quality and climate warming.

Requirements for eInfrastructure

To improve weather predictions and climate simulations, drastically increased model resolution, increased model complexity, and a larger number of ensemble integrations are required. The international climate community is presently pursuing development options in all three directions. Task-level parallelism on capacity machines yields more ensembles, while algorithms are being developed that will allow efficient use of future Exascale computers (i.e. with ~10^{18} floating point operations per second) for cloud-resolving climate change projections and weather predictions. Given the enormous societal benefits resulting from accurate estimates of regional climate change and weather, the costs involved in developing such computational systems are in fact relatively small[75].

Expectations from 2015

The trend in supercomputing points towards a rapid increase in the number of processors on a given machine while the speeds of individual processors are not likely to increase greatly. A number of difficult bottlenecks exist in weather and climate models preventing full exploitation of such future large-scale parallel systems. These bottlenecks arise from frequent input/output and tight coupling between adjacent model grid cells leading to the need for frequent inter-processor communication. Present GCMs can run relatively efficiently on supercomputers up to about 20 000 cores, beyond which inter-processor communication leads to a saturation of model speedup. Achieving the factor of a million improvement necessary for convection-resolving GCMs requires a major overhaul in the formulation of these models.

---

[64] For example, changed greenhouse gas emissions lead to a change in the response of the atmosphere to heat radiation.

[65] CMIP5 is the Coupled Model Intercomparison Project Phase 5, intended to support the next assessment report of the IPCC, the Intergovernmental Panel on Climate Change, www.ipcc.ch.

[66] Caption text is truncated from the IPCC version of this figure. The full caption, as required by the IPCC, is given on the figure attributions page at the end of this document.

[67] Decreasing the size of horizontal grid cells by a factor n increases the number of cells by a factor n^2, with the vertical resolution fixed. For numerical stability, the time step must then be reduced by a factor n also, so the total cost goes up by a factor n^3. Better vertical resolution and greater complexity add further to the cost.

[68] Norsk klimasenter, www.norclim.no, was formed by the Norwegian Meteorological Institute (met.no) and the Bjerknes Centre, with important contributions from the Nansen Environmental and Remote Sensing Centre (NERSC), the Norwegian Institute for Air Research (NILU), the Norwegian Polar Institute, UiB, and UiO.

[69] www.ncar.ucar.edu

[70] High Resolution Limited Area Model, hirlam.org, a research cooperation of European meteorological institutes.

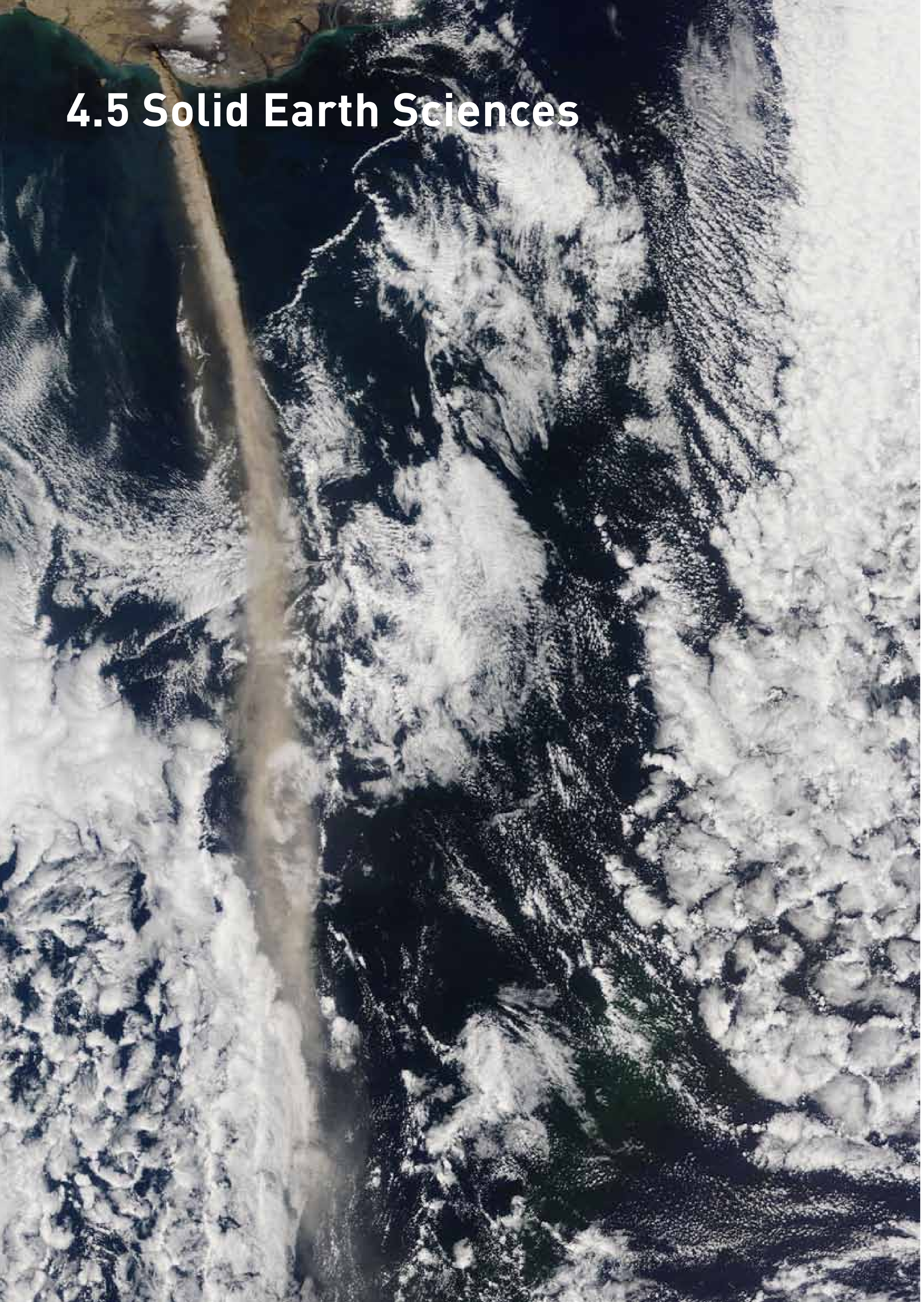
[71] www.metoffice.gov.uk/science/creating/daysahead/nwp/um.html

[72] Regional Ocean Modelling System, www.myroms.org, an open source code.

[73] NILU is the Norwegian Institute for Air Research, www.nilu.no.

[74] Monitoring Atmospheric Composition and Climate, www.gmes-atmosphere.eu, an atmospheric services program funded by FP7.

[75] As an example the next-generation climate supercomputer procurement at the US Oak Ridge National Laboratory was recently awarded to Cray Inc. at a cost of $47 million.



4.5 Solid Earth Sciences

Plate tectonics and earthquakes

The crust of the Earth is made up of a number of tectonic plates that move about, at times colliding with one another, then rifting apart again. These motions, slow though they are, lead to events that are violent and hazardous, including earthquakes and volcanic eruptions. The plate motions are probably connected to patterns of convection deep within Earth's mantle (Fig. 19), but our understanding of cause and effect is limited at best. Tectonic plates sometimes slide steadily against each other without much friction. This is called aseismic creep. In other instances friction halts the sliding for a while until the built-up stress breaks the sticking and slip occurs violently, in an earthquake. Conceptually, this is simple, but sticking and slipping are microscopic processes, while fault zones are thousands of kilometres in length. This multi-scale resolution problem is beyond the reach of present day computers.

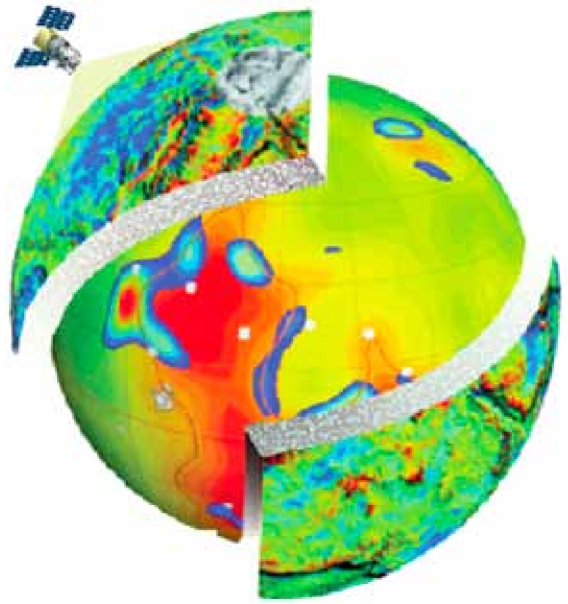


Figure 19. Sketch of components of a Whole-Earth model, including hot spots at the core-mantle boundary and features of the crust.

Laboratory studies of rocks under stress use acoustic and piezo-electric sensors to determine the stick-slip conditions. These experiments generate Terabytes of data from initiation until breakdown, and the samples involved are only a few centimetres across. Reliably understanding breakdown precursors for such samples will require analysis of many such experiments. The next step, already beginning in at least one well-known fault zone[76], is to repeat such analyses on rocks already subject to tectonic stresses *in situ*, and observe their states during aseismic creep and before earthquakes of different sizes. Data from such observations will range well into the Petabyte scale for a single fault system.

Volcanism

A complementary effect of the tectonic motion of plates is volcanism. When plates collide, sometimes one is subducted underneath another. Material in the subducted plate is pushed deep down into the mantle, where some of it melts and rises buoyantly into the crust of the overriding plate forming a magma chamber that feeds chains of volcanos. This occurs all around the Pacific Ocean, in the "Ring of Fire" where most of Earth's volcanos and violent earthquakes occur (Fig. 20).

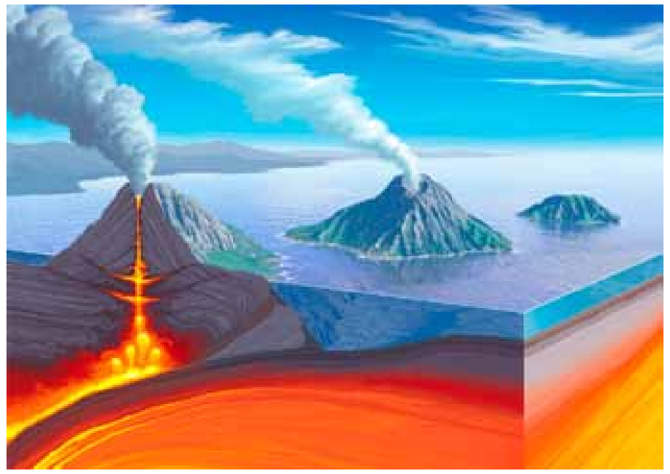


Figure 20. An artist's view of the formation of a chain of volcanos at a subduction zone.

Forecasting the eruption of active volcanos becomes significantly better as instrumentation improves and local observatories gain experience. Nevertheless, evacuations are sometimes ordered too early, with significant disruption to lives and economies, and sometimes not early enough, resulting in loss of life. Seismic data, gas emission, thermal data from satellites or *in situ*, borehole measurements, and surface deformations all contribute to the data that are analysed and compared with previous activity. Expanding the coverage from known active and dangerous volcanos to dormant volcanos that may erupt sometime in the future makes further demands on both data storage and processing. The eruption process itself is almost overwhelmingly complex, including phenomena such as supersonic jetting, fragmentation of rock, ejection of volcanic bombs, pyroclastic flows, production of dust and aerosols, and emission of gas. Performing a full computer simulation of an eruption is presently impossible. Parts of such eruptions have been simulated by scientists in the US[77], and work is beginning on such simulations in Norway[78]. Two-dimensional full scale simulations will be possible on Petascale machines; three-dimensional simulations will require Exascale.

---

[76] The SAFOD project (San Andreas Fault Observatory at Depth) earthquake.usgs.gov/research/parkfield/safod_pbo.php has drilled deep holes into an aseismic portion of California's famous San Andreas Fault. The data are made available to geologists worldwide, including some in Norway.

[77] Ogden, D. E., Glatzmaier, G. A., and Wohletz, K. H., 2008, Effects of vent overpressure on buoyant eruption columns: Implications for plume stability, *Earth and Planetary Science Letters*, v. **268**, p. 283-292.

[78] Gisler, G, 2009. Simulations of the explosive eruption of superheated fluids through deformable media. *Marine and Petroleum Geology Journal, Special Issue* Volume **26**, pp1888-1895.

Large igneous provinces and whole-Earth modelling

Not all of Earth's volcanos occur at plate boundaries. The Icelandic volcanic fields, whose recent Eyjafjallajökull eruption severely disrupted the lives of local farmers and the European travel industry, arises from the spreading centre of a mid-ocean ridge. Large Igneous Provinces (LIP), associated with plumes of hot material ascending through Earth's mantle, have produced some of the most violent activity the Earth has ever seen. The greatest mass extinction in the history of life on Earth may have been provoked by the outbreak of one such LIP, the Siberian Traps, 250 million years ago at the end of the Permian period[79]. The Yellowstone "super-volcano" is a manifestation of another LIP. Both mantle plumes and mid-ocean ridge spreading are part of the mantle convection system that drives the motion of the tectonic plates.

Computer models of plate motion through geologic time that include mantle plumes and ocean ridges have been produced by Norwegian researchers[80]. Local models of subduction-induced volcanism exist, but are still too primitive to predict the localisation of outbreaks, and there is yet no good model of how LIPs are formed and evolve. Whole-Earth models, including magnetic field generation, mantle plumes and mantle convection, plate motion, subduction, earthquakes, and volcanism, are well beyond the capability of any machines now available. But important strides in this direction can be made using multi-Petaflop machines.

Requirements for eInfrastructure

The multi-scale demands of problems in the solid Earth sciences demand very large and very tightly coupled computer systems. Groups at UiO working on these problems presently use millions of core-hours on the Norwegian national computing infrastructure, and the growth will increase dramatically as both whole-Earth modelling and the modelling of explosive volcanism mature. Exascale machines will be needed for adequate performance of such models. A crucial need for Earth sciences is the development of high-quality interactive visualisation facilities to aid in the development and understanding of detailed three-dimensional models.

Expectations from 2015

With the current rate of progress, an Exascale machine should be available somewhere in Europe within the next ten years. With Norwegian expertise and interest in solid-Earth natural sciences, including whole-Earth modelling, large igneous provinces, and explosive volcanism, we think a good case can be made for siting the first one in Norway.

---

[79] Svensen, H., Planke, S., Polozov, A., Schmidbauer, N., Corfu, F., Podladschikov, Y., and Jamtveit, B. (2009) Siberian gas venting and the end-Permian environmental crisis. *Earth and Planetary Science Letters*, **277**, 490-500.

[80] Torsvik, T.H., Müller, R.D., Van der Voo, R., Steinberger, B., and Gaina, C. 2008. Global plate motion frames: toward a unified model, *Reviews of Geophysics*, **41**, RG3004.

4.6 Life Sciences

The life sciences (biology and medicine and related interdisciplinary fields[81]) involve the study of living organisms from nanobes and viruses to man and even extraterrestrial life. The socio-economic relevance is high because these fields involve the understanding and treatment of disease, the ageing population, environmental threats, and the cost of health care, for example. The field is much broader than physics or chemistry, but not as mature in the use of mathematical theory and numerical modelling. The recent sprouting of new subfields like bioinformatics, biomedical computing, systems biology, computational medicine, biomedical imaging, and computational neuroscience illustrates the explosive growth in the usage and need of eInfrastructure in the Life Sciences.

Bioinformatics

Bioinformatics is a large, ill-defined field, associated with the algorithmic, analytical and computational challenges in genome research. But it also includes computational science related to biological macromolecules such as proteins and nucleic acids. The application of computers and quantitative methods to generate, explore and confirm biological hypotheses using high-throughput molecular and sequencing data is the core of bioinformatics. It is both data intensive and compute intensive, with heavy use of statistics to build models, make inferences and perform validation. Eventually nearly all biologists will use bioinformatics methods through well-adapted programs and databases for high-throughput sequencing, but expert support may be needed.

The emergence of high-throughput technologies has vastly increased the amount and range of data obtainable on cells, organisms, and populations. Bioinformatics databases have grown rapidly, and represent a vital source of information for molecular biology and medicine. This has facilitated studies of biological processes but it also requires new ways of organising, analysing, and interpreting the data[82]. Bioinformatics has its own terminology and methodology, dedicated publishing channels and conferences, and courses of study at universities. Its research agenda relies heavily on current technology in molecular biology and biotechnology; in return, biological research is strongly influenced by advances in bioinformatics. Most molecular biological and genomic research facilities now include divisions or teams devoted to bioinformatics; often bioinformatics and molecular biology faculties are co-located.

Biomedical computing

Biomedicine, or preclinical medicine, combines medicine with the fundamental biosciences and is concerned with the theory, knowledge, and research behind clinical medicine. Biomedical research aims at a deeper, molecular understanding of the mechanisms underlying disease as well as new diagnostic and therapeutic practices. This lays the foundation of modern medicine and medical practice. Biomedical computing is becoming increasingly important for the measurement of bio-signals, bio-structure, and bio-function; for analysing, storing and retrieving recorded data; and for modelling and understanding. Recognising the importance of computing, the National Institute of Health[83] (USA) has incorporated National Centres for Biomedical Computing in their Roadmap plan[84].

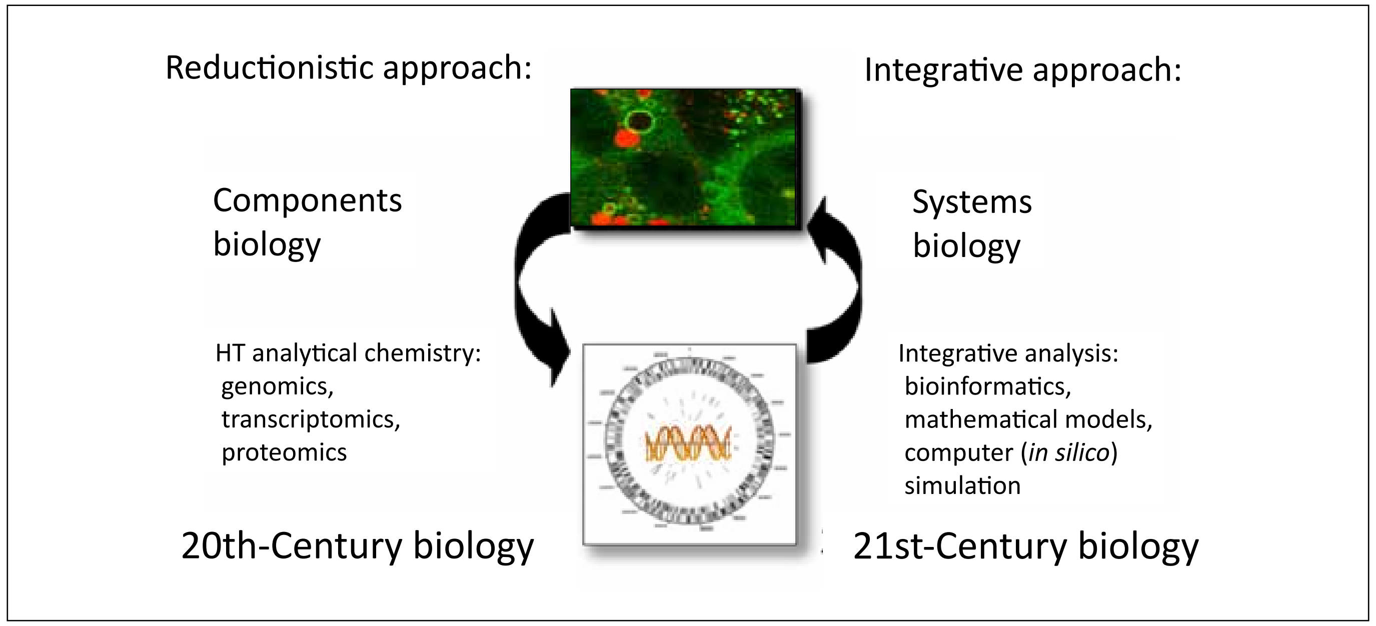


Figure 21. Systems biology as a shift from a reductionist to an integrative approach[85].

---

[81] Just a few of the related fields are: molecular biology, ecology, genetics, neuroscience, embryology, oncology, cardiology, computational biology, mathematical physiology, biomechanics, biophysics, systems biology, …

[82] "Biology in the twenty-first century is being transformed from a purely lab-based science to an information science as well" www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html

[83] www.nih.gov

[84] See www.ncbcs.org. The mission of the NCBCs is to create software and tools to help the biomedical community to analyse, model, and share data on human health and disease.

[85] Adapted from Palsson, BO, 2006. *Systems biology: Properties of reconstructed networks*. Cambridge Univ. Press.

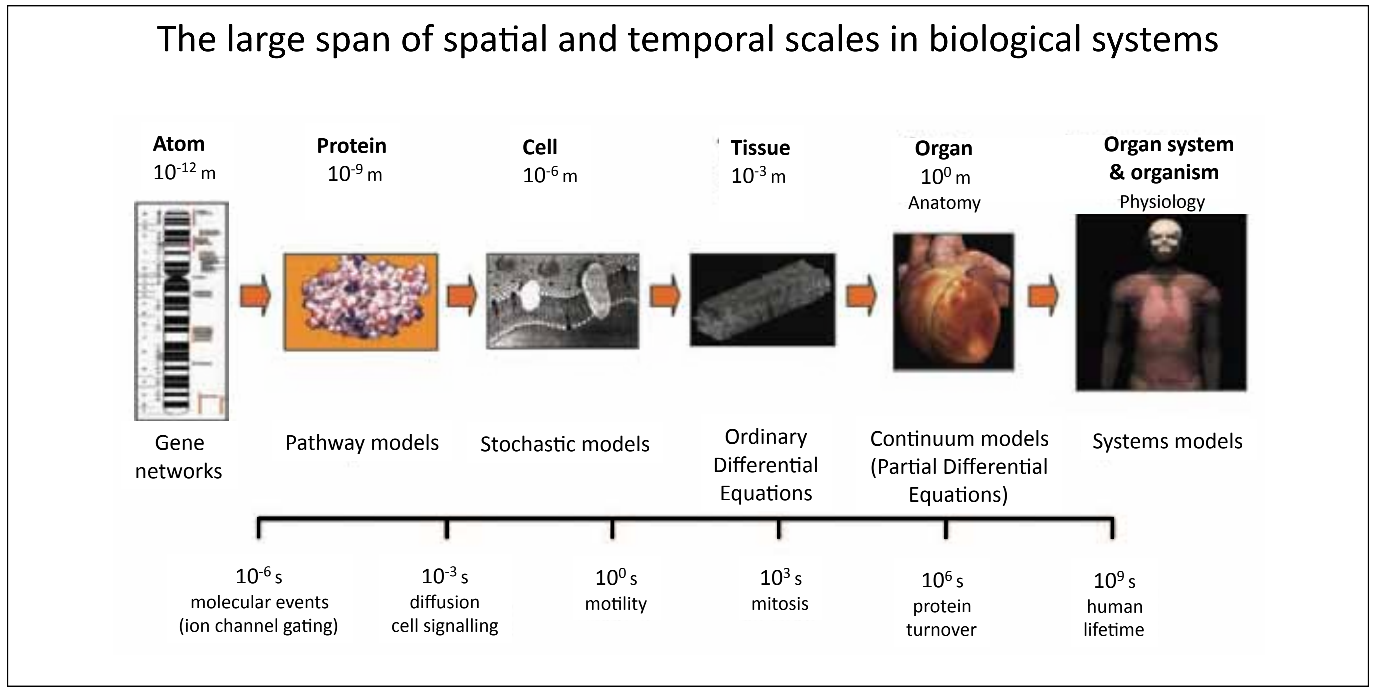


Figure 22. The large span of spatial and temporal scales in biological systems[89].

Systems biology

The field of systems biology focuses on interactions among biological systems, and reflects a paradigm shift in molecular and cellular biology (Fig. 21). The aim is the modelling and discovery of *emergent properties*, properties of a system that arise from collective interactions and are understood using e.g. dynamical systems theory. Systems biology obtains and analyses complex data from multiple experimental sources. Accordingly, it splits into a large number of subdisciplines with names like *genomics*[86] and *interferomics*[87]. Wikipedia lists no fewer than eleven of these *-omics*.

The enormous range of temporal and spatial scales involved in systems biology (Fig. 22) requires experimental and computational methods that span those scales competently. The methodology of systems biology includes the use of robots and automated sensors to aid in experimentation and data acquisition. Mechanistic (*in silico*) computational models, data mining techniques, and very large online databases put stringent demands on available eInfrastructure. Examples of new cooperative developments include the BioModels Database and the Systems Biology Markup Language[88].

Computational medicine

Computational medicine is a new multidisciplinary field employing mathematics, physics, information technology, and biomedical engineering to understand disease mechanisms and improve diagnosis, prediction, and treatment of disease. Useful clinical information for improving health is obtained by collecting and analysing heterogeneous data from personalised genetics profiles, bio-signals, advanced imaging and other phenotypic information. The field of computational medicine is a high priority of the French National Institute for Research in Computer Science and Control[90], and several universities and research institutions have established Computational Medicine units[91].

Biomedical imaging

During the last decade, improvements in spatial resolution, temporal resolution, and the ability to acquire multichannel or multispectral images from organ or tissue samples, whether *in vivo* or *ex vivo*, have increased the sophistication of biomedical imaging. The amount of imaging data from a biological experiment or a medical exam can be tens to hundreds of megabytes, where relevant tissue or cellular information is distributed in both time

---

[86] The study of organismal DNA sequences, including intra-organismal cell-specific variation (Wikipedia).

[87] The study of organismal, tissue, or cell level transcript correcting factors (Wikipedia).

[88] www.ebi.ac.uk/biomodels-main/. European Bioinformatics Institute, European Molecular Biology Laboratory. sbml.org. SBML is an international effort to develop a format for models of biological processes.

[89] Adapted from: C. Dollery and R. Kitney, *Systems biology: A vision for engineering and medicine*, Technical report, The Academy of Medical Sciences and The Royal Academy of Engineering, London, UK, Feb. 2007.

[90] INRIA www.inria.fr/inria/strategie/priorites

[91] E.g. Johns Hopkins University - www.icm.jhu.edu; University of Oulu - www.computationalmedicine.fi; Chinese Academy of Sciences - english.ia.cas.cn; Uppsala University - www.medsci.uu.se/klinfarm/compmed.

and space. Frequently, the researcher or clinician will inspect and describe these images qualitatively, and extract quantitative data as to size, shape, change, or motion, and physiological or molecular parameters such as flow, perfusion, diffusion, and permeability.

Important research areas include multimodal image registration[93]; image segmentation; object tracking; and conversion of three-dimensional image data into accurate geometry and computational models[94], each with a broad range of applications. The field of bioimaging uses new technologies within high performance computing, grid[95], and cloud computing. Computational neuroimaging[96] is a subdiscipline that typically uses functional magnetic resonance imaging (fMRI) to quantitatively investigate the relationship between the brain and behaviour (Fig. 23). The emerging field of imaging genetics[97], correlating image features with associated gene profiles in genome-wide association studies (GWAS), is also both data-intensive and compute-intensive.

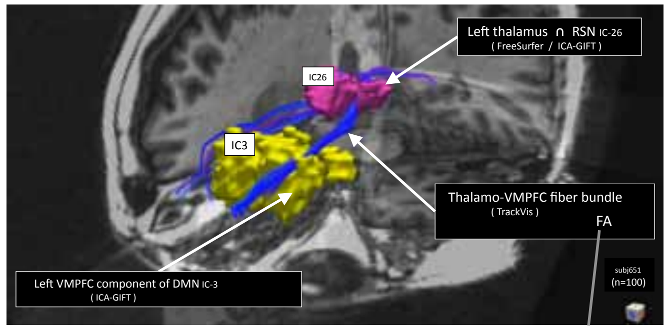


Figure 23. fMRI image of the brain of a healthy volunteer in a study of cognitive ageing, obtained automatically with freely available image processing and analysis tools[92].

Computational neuroscience

Understanding the nervous system is the mandate of the neurosciences. Computational neuroscience attempts to model everything from subcellular processes within a single neuron to sensory processing, memory, networks, and neural correlates of consciousness. Because of the tremendous complexity of neural circuits this field requires staggering amounts of information and computational power. In Norway the Kavli Institute for Systems Neuroscience and the Centre for the Biology of Memory at NTNU[98] study how spatial location and spatial memory are handled in the brain, and the Computational Neuroscience group at the Norwegian University of Life Sciences[99] study brain activity and large-scale models of neuronal networks.

A consortium of academic and commercial institutions in the UK universities is creating a virtual laboratory for neurophysiology in the CARMEN project[100], which archives data from experimental procedures and the software that were used to process the data. Collaborators can share tools, methods, and algorithms, and run analyses on common computer resources. The Blue Brain project[101] is a first attempt to perform detailed simulations of the mammalian brain to understand its function and dysfunction.

Novel neuromorphic computing architectures

In January 2010, a Brussels Workshop on Future and Emerging Technologies (FET) presented a proposal to design, build, and operate a large scale Neuromorphic Computation Facility (NCF) for basic research and application development[102]. This will be a configurable physical model of high-fidelity brain microcircuits, with a design based on the massively parallel architecture of the brain, rather than on von Neumann architecture. This conceptually different computer, with 10^{11} neurons and 10^{15} dynamic synapses, will explore the dynamics of complex brain-like systems.

A related FET-Open project, Brain-i-Nets: Novel Brain-Inspired Learning Paradigms for Large-Scale Neuronal Networks[103], will produce a set of rules for synaptic plasticity and network reorganisation to describe the adaptive processes occurring in the living brain during learning, and to port these rules into neuromorphic hardware. The US Defence Advanced Research Projects Agency has funded IBM and university partners for Phase 1 of the SyNAPSE[104] initiative. This program hopes to create new electronics hardware and architecture that can understand, adapt, and respond to informative environments in ways that mimic biological brains. IBM's BlueMatter software for neuroscience modelling performs cat-brain-scale simulations on the Blue Gene architecture (Fig. 24).

Scientific challenges

The great variety among the life sciences presents a challenge in building an efficient Norwegian infrastructure for the diverse needs of biologists and medical scientists. Local facilities should be strengthened, and communication among different institutions must be enhanced. The educational challenges of multidisciplinarity, mutual understanding, and the cultural barriers between the exact sciences and biology and medicine remain a problem[105]. Bioimaging must deal with visualising and exploring large-scale, multidimensional data. Finally, in many experimental situations the dimensionality of the data can be orders of magnitude larger than the number of available samples so that feature selection and model selection are difficult. The newer mathematics of data mining and machine learning, like compressed sensing and random matrix theory, are needed.

---

[92] Courtesy of Martin Ystad, Tom Eichele, Erlend Hodneland, and Judit Haasz. See also: Lundervold A, 2010. *Nonlinear Biomedical Physics*, 4 (Suppl 1):S9.
[93] Slomka PJ, Baum RP, 2009. *Eur J Nucl Med Mol Imaging*, **36** Suppl 1:S44–S55.
[94] Young PG, Beresford-West TBH, Coward SRL, Notarberardino B, Walker B, Abdul-Aziz A, 2008. *Philos Transact A Math Phys Eng Sci*, **366**(1878):3155–3173.
[95] Keator DB, Wei D, Gadde S, et al., 2010. *Front Neuroinformatics*, **3**(30); Ruiz M, 2009. *Meth Mol Biol* **569**:219-238.
[96] Deco G, Jirsa VK, Robinson PA, Breakspear M, Friston K, 2008. *PLoS Comput Biol*, **4**(8):e1000092.
[97] Bigos KL, Weinberger DR, 2010, Imaging genetics-days of future past. *Neuroimage*, in press.
[98] commonweb.ntnu.no/cbm/ The Centre for the Biology of Memory is a Centre of Excellence funded by RCN.
[99] compneuro.umb.no
[100] Code Analysis, Repository & Modelling for E-Neuroscience, www.carmen.org.uk, funded by the Engineering and Physical Sciences Research Council of the UK.
[101] bluebrain.epfl.ch, of the Brain and Mind Institute of the Ecole Polytechnique in Lausanne, Switzerland
[102] cordis.europa.eu/fp7/ict/fet-proactive/docs/flagship-ie-jan10-07_en.pdf, not yet funded.
[103] brain-i-nets.kip.uni-heidelberg.de, recently funded by the EU.
[104] Systems of Neuromorphic Adaptive Plastic Scalable Electronics, www.darpa.mil/dso/thrusts/bio/biologically/synapse/index.htm
[105] National Research Council of the National Academies (Ed): Mathematics and Twentyfirst Century Biology. The National Academies Press, Washington D.C. 2005, books.nap.edu/catalog.php?record_id=11315

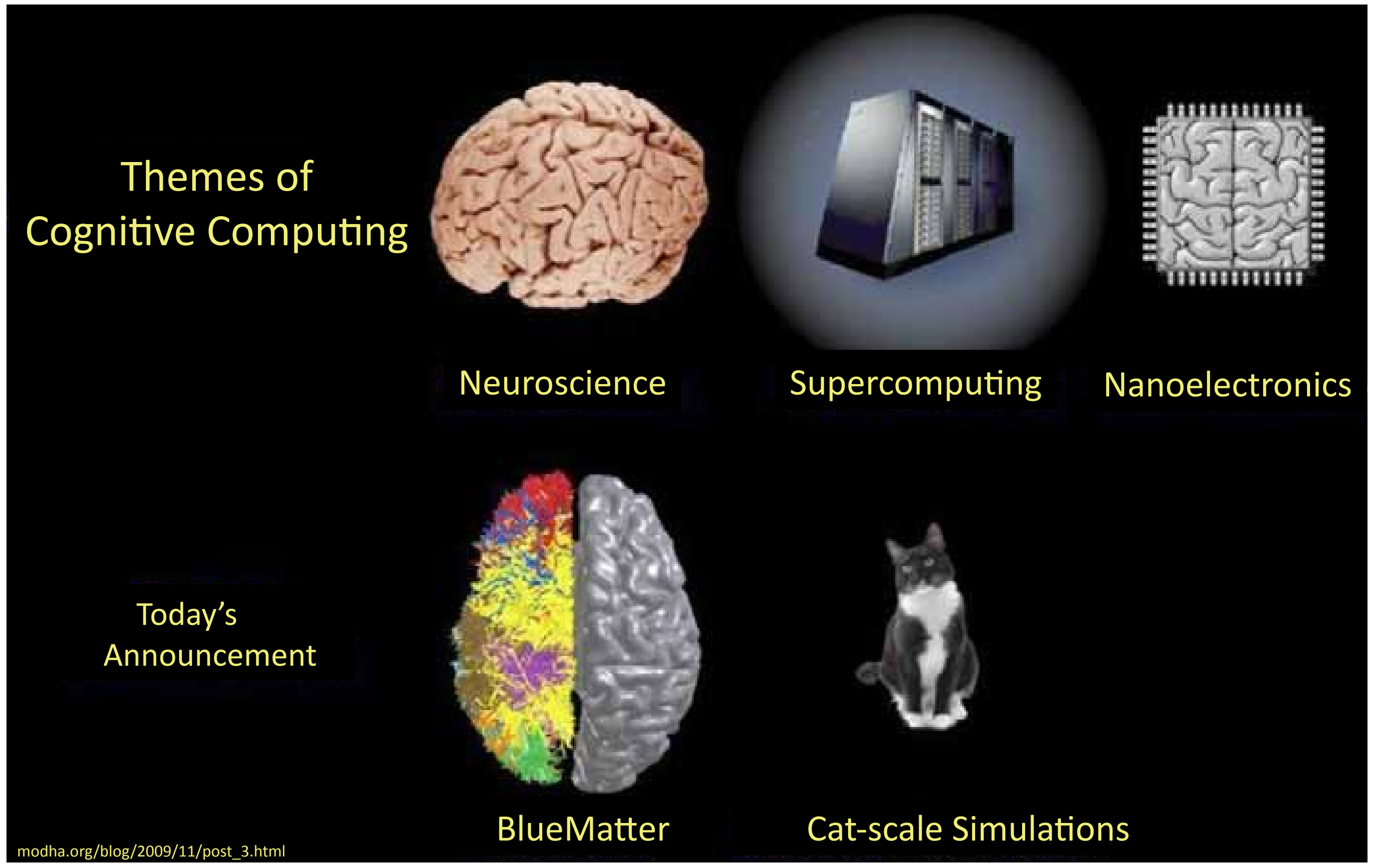


Figure 24. The announcement of the DARPA SyNAPSE award to IBM and the BlueMatter simulation platform[106].

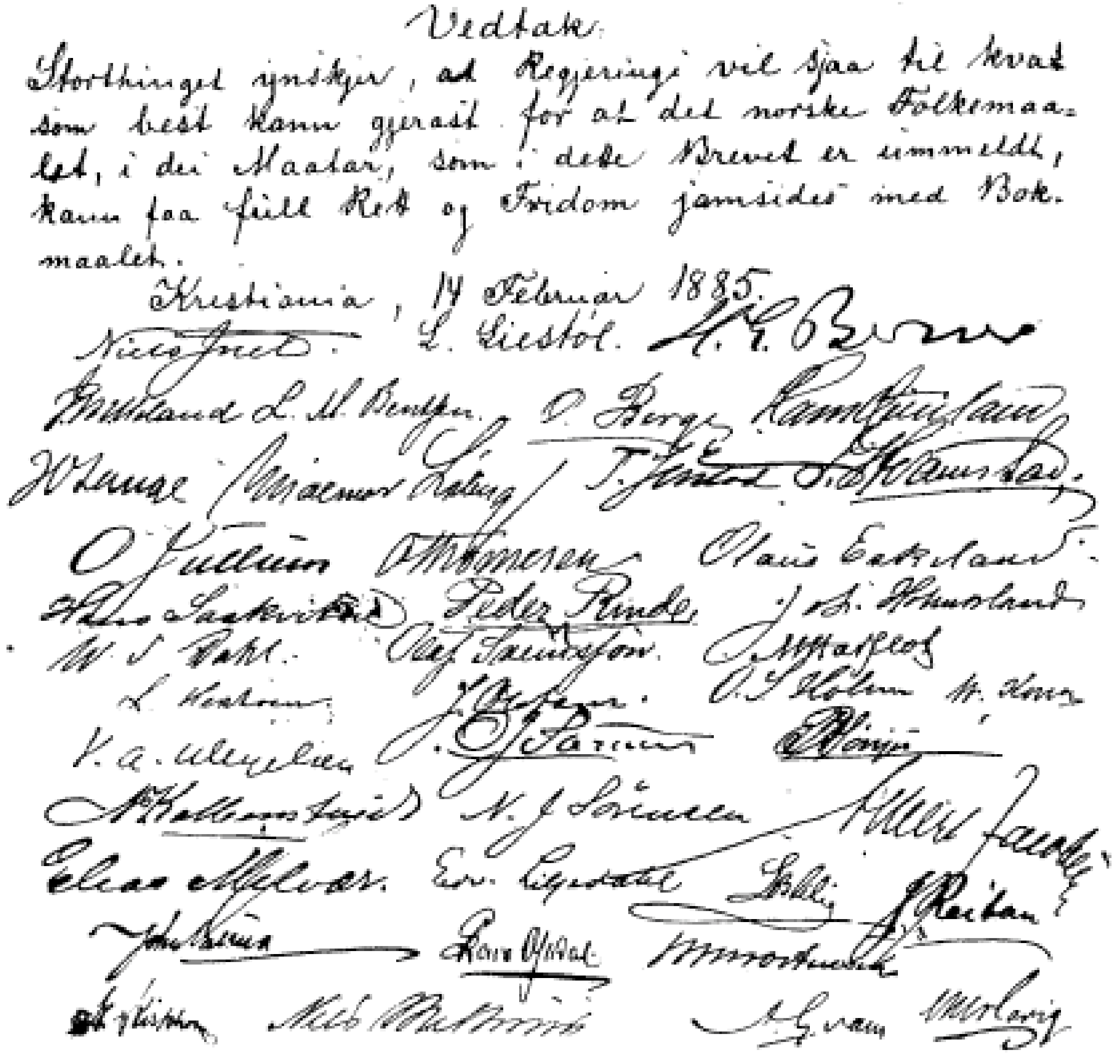
Expectations from 2015

During the next decade the life sciences will become increasingly exact and quantitative[107], and possibly more demanding of eInfrastructure resources than chemistry and physics are. Research in the life sciences is already generating data amounts comparable to nuclear physics, and the volume of complex data will grow rapidly. Researchers and clinicians need to acquire the knowledge and tools to effectively use future eInfrastructure and successfully apply the lessons learned. The field of genomics has already exploded through second generation sequencing technology. This shifts focus towards enormous efforts in functional and comparative genomics generating huge amounts of data within this particular field. Completely new sets of tools for the analysis of genome data will be needed. It is also to be expected that genomics will have an impact within the entire life science field — and outside the life sciences in psychology, anthropology, archaeology, and humanistic sciences as well.

---

[106] modha.org
[107] Cohen JE, 2004. Mathematics is biology's next microscope, only better; biology is mathematics' next physics, only better. *PLoS Biol*, **2**(12):e439.

4.7 Language, Cognition, and Information



Vedtak.

Storthinget ynskjer, at Regjeringi vil sjaa til kvat som best kann gjerast for at det norske Folkemaalet, i dei Maatar, som i dette Brevet er ummeldt, kann faa full Rett og Fridom jamsides med Bokmaalet.

Kristiania, 14 Februar 1885.

L. Liestøl.

Peder Rinde

The interdisciplinary field of language and information is not yet a big player in HPC, but computers are playing a steadily more prominent role in linguistics, modern textual scholarship, history, information and communication sciences, and other areas of the humanities and social sciences. Computational requirements vary with the multiple roles that language plays, as instrument of communication and as carrier of cultural content and historical knowledge.

Overview and current status

Language processing addresses two major challenges of theoretical and practical relevance. On the one hand, language is a complex, uniquely human capacity which is intimately linked to logical thinking, knowledge representation, problem solving, self-reflection, consciousness, emotion, creativity, and learning. For about half a century, *computational linguistics* has used computer modelling to explore this capacity, while its applied branch, *language and speech technology*, has developed practical applications such as machine translation, writing tools, information search, etc.

On the other hand, language is a major carrier of cultural content and historical knowledge. Much of this is endangered. Not only is a language dying out every two weeks, but many collections of language and multimedia materials are in acute danger of being lost. Audio and video recordings are rapidly decaying, old digital formats are no longer supported, materials are insufficiently documented and their makers have left the scientific community. Only a large scale digitisation effort can preserve these materials and make them accessible for future generations.

Powerful language models promise deeper insight into the unique human cognitive capacity, better information access in our multilingual society, and a content analysis of vast archives of cultural and historical materials. The effort needed to achieve these challenges is extensive and requires a well-organised computational infrastructure and a better integration of skills from the Humanities and Social Sciences with those of Computer Science and Information Science.

Language has a twofold complexity. First, the same meaning can be expressed in multiple ways. A keyword-based search for *"Supercomputer with 10 000 CPUs"* is unlikely to retrieve texts containing *"Ten thousand node HPC cluster."* A deeper semantic analysis is needed. Second, language is ambiguous, as users know from misunderstandings and jokes, like this from Groucho Marx: *"One morning, I shot an elephant in my pyjamas. How he got in my pyjamas, I don't know."* Ordinary sentences can have thousands of syntactic analyses, with consequences for computational requirements.

Spoken language presents additional challenges. Models of speech recognition involve advanced signal processing as well as linguistic and cognitive research. Differences in voice, tempo, dialect, and other speaker characteristics are difficulties. Multimedia resources will require extensive digitisation and new approaches to annotation and interpretation to make them searchable. In speech processing, dictation systems are popular in domains where keyboard use is awkward or too slow, such as medical consultations, heavy industries, environments requiring gloves, and so on. Text-to-speech systems are necessary for the blind and useful for a wide range of public and private information services using auditory channels.

Rule-based approaches from the 1960s were followed by neurally inspired models and an explosive development of machine learning and hybrid approaches from the 1990s onward. Statistical language models can be applied at all levels of analysis, from the disambiguation of speech sounds to the selection of the most likely interpretation for a whole sentence. However, these models are computationally intensive and require large, high-quality data sets for training and testing.

Data mining and classification techniques are used in different areas of linguistics. In dialectology, recent models use quantifiable measures such as Levenshtein distances[109] in clustering algorithms for dialects. Authorship attribution involves memory-based learning, shallow parsing, and genetic optimisation. Text sciences investigate texts in their various editions, publications, and interpretations, as works of literary, cultural, philosophical, religious, or historical importance. Many literary texts have a long history of editions, while historical materials (such as rune inscriptions) can be considered at multiple levels of normalisation and interpretation.

Modern computational tools can provide the researcher with many parameters to dynamically control the rendering of a text for particular purposes. The interactive dynamic presentation of the Wittgenstein Archive in Bergen[110] provides an illustration. An immense amount of manual work has been spent encoding the text in forms suitable for exploration. Currently there is still a lack of good computational tools to support this work.

Socio-economic relevance

In our society faced with an information deluge, language processing is of great importance. Providers of web search and information retrieval services, such as Google and Comperio[111] offer a range of language-enabled services, including translation and summarisation of web pages, searching for web pages with similar content, and complex customised data mining to structure the information found. Applications such as NewsExplorer[112] and NewsBlaster[113] attempt to classify, analyse, and cluster multiple news stories, producing summaries, timelines, and relevant named entities. As more people use social networks on the web, mass opinion estimation based on natural language processing and sentiment analysis, and assembling interest groups across different languages, form attractive new areas of research. Computer Assisted Language Learning[114] and tools for people with language disorders or handicaps, are also of high social relevance. User profiling for targeted advertisements is already a commercial application.

---

[109] The minimum number of edits required to transform one string into another; see en.wikipedia.org/wiki/Levenshtein_distance

[110] WAB, wab.aksis.uib.no/transform/wab.php?modus=opsjoner

[111] Google: www.google.no, Comperio: comperio.no

[112] emm.newsexplorer.eu/NewsExplorer/home/en/latest.html. NewsExplorer is a project of the Joint Research Centre of the European Commission.

[113] newsblaster.cs.columbia.edu. NewsBlaster is a project of the Natural Language Processing Group at Columbia University.

[114] CALL, en.wikipedia.org/wiki/Computer-assisted_language_learning

Current scientific challenges and expectations from 2015

It is expected that language, cognition, and information technology will fuse with multimodal input and output technologies, like speech, sensing and tactile input, immersive 3D visualisation, advanced robotics, and combinations. Information content will be increasingly mined and mapped along thematic, social, geographic, and time lines. Categorisation and aggregation tools will be used on different scales, ranging from international trend forecasting and mass opinion mining to searching for needles in information haystacks. With new modelling techniques applied to increasing amounts of data, the need for computer power will remain beyond the capacity of modern desktops.

Models of language and cognition deal with the problem of finding patterns in badly structured data. A challenge is to develop annotated language materials of sufficient size, with sufficient detail and precision in annotation, and with as few access restrictions as possible. Such materials are very labour-intensive and expensive to develop. Individual researchers, projects, groups, or institutions may possess valuable materials, but rarely have the financial resources to maintain a corpus infrastructure for external access and reuse. Cooperation and coordination on a larger scale is needed. Through CLARIN[115], the humanities are promoting a European centre structure in which Norway would contribute machines, data, tools, and web services.

Data repositories with adequate metadata and tools for cataloguing, filtering, data sharing, and interactive collaboration will become important. In the next decade there will be an enormous increase in available digitised language and multimedia data. The present digitisation effort at the National Library of Norway[116], along with other archives and libraries, combined with the increasing production of language materials destined for digital media, will produce a very large and very diverse mass of texts and recordings which must be properly managed. Cooperation on standards and technical formats at the national and international level is needed to streamline access to language resources.

In the medium to long term, curation will be one of the main challenges, including such diverse aspects as managing data provenance, quality assurance, access protocols, migration to new formats and platforms, and upkeep of standards for encoding and metadata. In an eInfrastructure for language, there will be a demand for tools and services, not merely for access, but for enriching the materials, producing secondary resources such as lexicons and word-nets, and developing translation equivalents and collocations. Using techniques that derive new information from existing resources, the underlying knowledge can be viewed in a different way.

History, archaeology, history of art, ethnography, and social sciences deal with not only with written sources, but with experiential data like stories and videos and a wide variety of artefacts and other objects that need to be mapped in time and geography. This calls for new types of databases allowing data mining using dimensions of space, time, thematic, and theoretical parameters. Already, humanities scholars use GPS tagging and overlays of multiple views as in Google Earth or GIS systems (see §4.8). Stories and events are being mapped to the places in which they occur. Visualisation and interactiveness of data will be part of the eInfrastructure, and augmented reality is beginning to play a useful role.

Requirements for eInfrastructure

Language models are characterised by data-intensive processing. Within four years, language models for analysis and interpretation, as well as applications such as machine translation, will need to be trained on corpora which are 500 times bigger than today's. Grammars, production systems, training materials, application contexts or other parameters in the model evolve continuously. Seemingly simple changes will often have unexpected global effects, so that automated validation, regression testing, progress reporting, and error detection are necessary to keep track of the evolution of a model.

Recent experience in Norwegian-to-English machine translation[117] shows that demand for compute-intensive tasks increases sharply as models become more complex. Exploring thousands of candidate translations per input sentence, an end-to-end translation of a 5000-sentence corpus took 155 computing hours; when more vocabulary was added, this number went up to 246 computing hours. Development cycles of the next generation of language models will require clusters which remain about a factor 100 bigger than desktops in the years to come.

The largest word collections for Norwegian[118] do not themselves present a problem for storage, but the annotation of these collections is complex and may take up hundreds of times more space than the source. Computational services needed for processing and searching these data are significant. End users need full-time access through online search and filtering systems, with fast turnaround times. Most research projects in computational linguistics develop their own software, but the infrastructure should be to provide an inventory of existing resources tools, web services and pipelines[119].

Besides computing, storage and search capacity, data curation is especially important for language and text sciences. Language data comes from a wide range of disparate sources, and can be annotated in various ways by other actors. Metadata including adequate information about data provenance is extremely important. Trusted identification of individuals, organisations, and services are important to give credit to authors and sources of data creation, and to provide authorisation based on licenses and rights. This identification needs to be persistent so that it will be possible to determine who created the data and how they were created decades afterwards.

Most humanities research groups lack the personnel to use eInfrastructure effectively. Due to lack of experience and support, many researchers simply do not know what can be done. Capacities for working with infrastructures must therefore be built up in departments of languages and digital humanities, implying long-term funding and good institutional strategies for bridging gaps among scientists from previously isolated disciplines, and between scientists and technicians. Transnational cooperation, as is currently undertaken in the CLARA[120] project, is also required.

---

[115] Common Language Resources and Technology Infrastructure, www.clarin.eu. CLARIN is a project under the ESFRI roadmap.
[116] nb.no
[117] The LOGON project, www.emmtee.net/index.php?page=1&lang=en
[118] Norsk aviskorpus, avis.uib.no and NoWac www.hf.uio.no/tekstlab/nowac.html, both about 700 million words and growing.
[119] Example pipeline tools are UIMA, Unstructured Information Management Architecture, developed by IBM; see en.wikipedia.org/wiki/UIMA and GATE, General Architecture for Text Engineering, an open source tool originating in the UK, see gate.ac.uk/overview.html
[120] Common Language Resources and their Applications, a Marie Curie initiative of the European Commission, see clara.uib.no.

4.8 Social Sciences

Geographical Information Systems

The future development of Geographical Information Systems[121] will lead to intensive use of computers, data storage, and networks, especially among communities, such as social scientists, who are not presently heavy users of eInfrastructure. Examples of primitive, but widely used, GIS are Google Earth with its multiple layers of data, and satellite navigation systems that include information about service stations, points of interest, and traffic or weather warnings. Sociological studies of changes in land-use patterns, surveys for exploitable resources, tracking migrational patterns, and developing adaptation strategies for climate change will all involve the integration of eInfrastructure with GIS[122].

Datasets used up to now in GIS can be large, but not huge, on the order of tens of Gbytes. Modern dynamic scientific queries will require the collation of hundreds to thousands of such datasets, resolved in time as well, leading eventually to Petabyte real-time searches. Active GIS systems will have data added to them constantly, from new satellite observations, or new data obtained on the ground, and these must be co-registered and collated with earlier data. More challenging is the concept of adding *experiential data* to GIS, such as stories, anecdotes, audio and video recordings, vastly expanding the size of databases for the use of researchers in the social sciences. Licensing rights and privacy issues for data requires care in handling and curation. Better user support for social scientists who are not expert computer users is urgently needed to help manage these issues.

Presently, social science researchers using GIS rely on cloud computing services provided commercially. Privacy, licensing, and reliability issues lead to a requirement for more secure systems for data storage and curation, and computational needs will expand greatly in the coming decades.

Economics

The science of economics seeks to understand how decisions taken by individuals, firms, and policy makers interact to produce the aggregate performance of an economy. With millions of agents, the overall complexity is enormous. In practice, economists rely on simple and stylised models in which different markets and types of decisions can be isolated, using statistical data or reported behaviour from small samples of surveyed agents.

Administrative registers have recently emerged as a promising new data source. These record the actual behaviour of all workers and firms over extended periods of time, and incorporate the complexity and heterogeneity of their interactions. If the registers are complete, these interactions add up to the aggregate economy and can be used to build a more realistic model.

Administrative data are used to identify differences in behaviour and preference and to estimate the consequences of unidentified differences. With millions of observations, stochastic modelling of individual decisions is a computational challenge. Single estimations can take tens of thousands of core-hours. In many cases repeated estimations and bootstrapping techniques increase the computational demands by orders of magnitude. Storage and memory requirements are not large, but the data frequently involve licensing and privacy issues.

Computational evolutionary finance

Evolutionary finance applies Darwin's principle of natural selection to study trading behaviour and asset prices in financial markets. In this perspective, a financial market can be seen as a selection mechanism that transfers wealth to traders who are well adapted to the environment from traders who are less well adapted. The trading strategies of wealthy traders determine the prices of financial assets, first, because those strategies are backed by more wealth, and second, because the strategies of wealthy traders are more likely to be copied by other traders. Wealth in financial markets is therefore the counterpart to fitness in biological systems.

Natural selection in financial markets produces rational behaviour in the aggregate. The rationality assumptions that form the basis of the standard theory of financial markets can be seen as a proxy for this outcome. In many problems of interest, this proxy is excellent, in others not. Evolutionary finance aims to provide answers to questions on which standard theory is silent (for example how markets become efficient) or questions that cannot be analyzed with the standard tools.

Computational evolutionary finance employs techniques from computer science to carry out controlled experiments in market design or regulation, for example. The effects of new regulatory measures can be assessed before they are implemented. A typical model in this field consists of a detailed description of the market microstructure, and a large number of individual agents who make portfolio and trading decisions within that microstructure. The agents' trading strategies are represented as computer programs, and the model is solved by subjecting these programs to natural selection until the aggregate price process has converged to a stationary process. Data from the model can then be analysed and compared with data from other market microstructures from other experimental treatments.

Access to High Performance Computing resources is a necessary prerequisite for carrying out such experiments. Typically, one would like to use many CPUs (500) for a relatively short time (12 hours) to generate data from one experiment. Currently, the workload is embarrassingly parallel, and the bottleneck is CPU cache performance. Storage needs are in the order of 10 GB per experiment. In the future, it is expected that computational evolutionary finance will benefit from parallel processing with a moderate amount of interprocess communication.

Requirements for eInfrastructure

The main requirements for social science applications are high bandwidth and an efficient data storage system with good protocols for data curation and security.

Expectations from 2015

As experiential data begins to be encoded into GIS systems, the complexity of managing and accessing the data will increase explosively. As the effects of climate change begin to be felt, the societal needs for adaptation, and the requirements of greater efficiencies and green energy sources will put demands on all aspects of eInfrastructure.

---

[121] GIS are systems in which data of various sorts are geo-referenced, or tied to a location on the surface of the Earth.
[122] S. Wang, 2010. A CyberGIS framework for the synthesis of cyberinfrastructure, GIS, and spatial analysis, Annals of the Association of American Geographers, **100**, 1-23.

4.9 Industrial Applications

Production from oil and gas reservoirs off the Norwegian shore remains of great importance to Norway's economy. Exploitation of these limited resources becomes more difficult as the fields are depleted. Economic benefits can be gained by studying the migration and trapping of hydrocarbons in source rock through using new algorithms and substantial computational resources.

The delivery of waste carbon gases into reservoirs for storage also involves migration and trapping. Reactions between the gases and the host rock may cause energy and volume changes with consequences for the stability and longevity of the storage. The reactions take place on the pore scale, micrometres to millimetres in size, while the reservoirs are order kilometres in size. Full high-resolution modelling of systems like these is not yet feasible, so sub-grid modelling of pore-scale processes is used. Significant advances in computer hardware over the next decade will help.

Pore-scale sequestration of carbon dioxide

Large scale reservoir models are totally reliant on the knowledge of the constitutive physical parameters like relative permeability and capillary pressure curves for reservoir rocks. Traditionally, these parameters are measured experimentally on core samples, but experiments of this type are time-consuming and challenging to perform. As a consequence, experimental data are often sparse, and single data points represent large volumes in the reservoir. This leads to an over-simplification of the reservoir model and causes large uncertainties in reservoir simulations. Furthermore, these experimental data give no specific information regarding the processes acting on the pore scale, particularly the fluid behaviour in the pore system.

An alternative approach is therefore to simulate the behaviour of the fluids confined in the pore system by applying advanced computer modelling (Fig. 25). Computer-aided analysis provide much additional information to laboratory measurements and enriches already existing experimental data. In the particular case of carbon dioxide storage, computer simulations can give fast and reliable information about the quality of a potential reservoir rock with respect to the trapping and storage.

To best capture the complexity of the pore space and the confined fluids, multiphase simulations can be used on images of the pore space. This can be achieved with the lattice-Boltzmann method. The method is also suited to investigate different fluid properties such as viscosity, surface tension and density in addition to variation in flow rate. These computer simulations are in most cases very computationally intensive and require massive parallelization to be executed on high-performance computing clusters in order to handle the sample sizes that are required to accurately capture multiphase flow properties of the rock.

Seismic imaging, modelling and inversion

Seismic imaging is an important method for mapping the upper part of Earth's crust. For exploration for hydrocarbons and optimising production from existing oil and gas fields it is vital. New hydrocarbon reservoirs are increasingly hard to find and there is an increasing demand for new seismic imaging technology. The next generation of seismic models will no longer be based on the arrival time of individual phases but on the complete information available in recorded waveforms, so methods for modelling and inversion must be improved. Seismic imaging has made significant advancements in this direction in recent years with the use of adjoint methods in waveform tomography. Computational resources with fast inter-node communication are key ingredients for the application of these imaging techniques in Norway.

Existing codes to calculate full waveform propagation in high resolution 3D-Earth models are computationally expensive but provide a unique opportunity to benchmark complicated regional Earth models produced by the Norwegian geodynamical community. Predicted effects for the individual models can be compared with waveforms (seismograms) observed over the whole frequency band between earth tides and microseismicity.

Seismic images of geological formations are created using reflected sound waves generated by an artificial source at the surface. There is a nonlinear relationship between mechanical rock properties and measurements at the surface, and an inverse problem must be solved to obtain estimates of the rock parameters. The solution of this highly nonlinear inverse problem involves numerical simulations of a large number of experiments and is computationally demanding. The size of the inverse problems to be solved is limited by computational resources. Running these inversions as standard tools in the Norwegian earth sciences community will require greater availability of substantial computing resources.

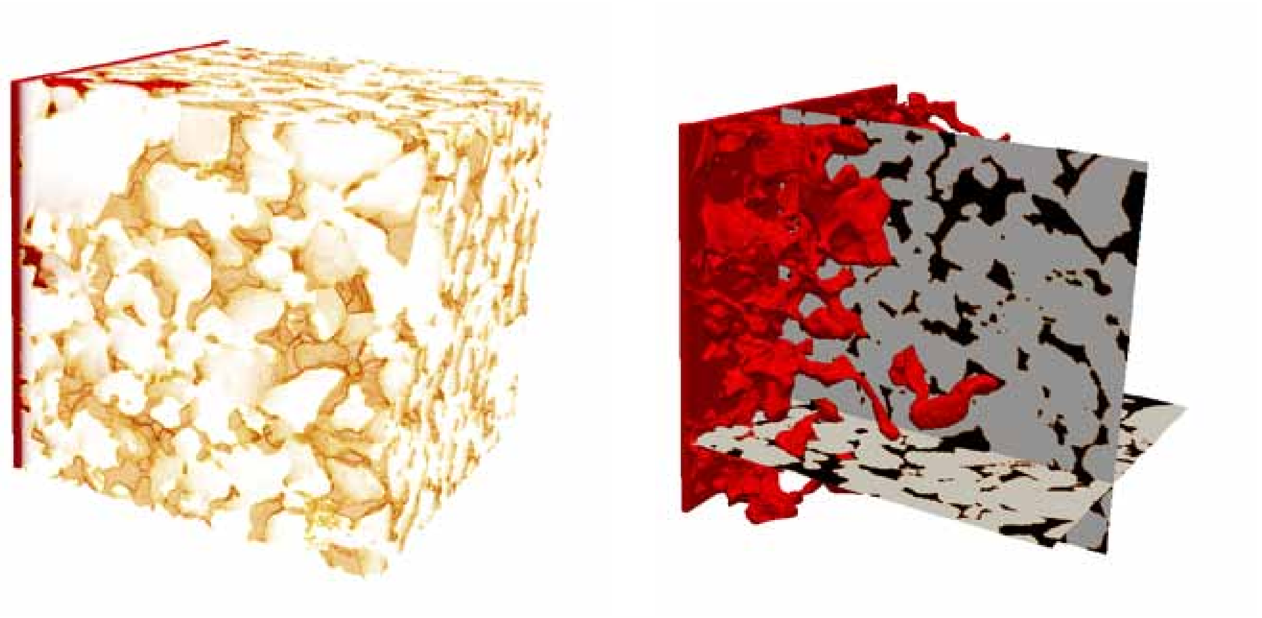
Computationally there are two main challenges; the size of the grid of the numerical discretization and the number of simulations required to model a full seismic survey. The number of grid nodes is determined by the wavelengths of the waves and the size of the computational domain. For a realistic simulation the computational grids need between 1 and 100 GB of storage. Large simulations thus require either out-of-core solutions or the use of domain decomposition. With today's standard hardware a single simulation can be performed in a few hours on a single processor. However, a realistic seismic survey may require several hundred thousand simulations. These simulations are independent and run in parallel on a cluster with several thousand processors.

Geothermal Energy

Shallow geothermal energy is a well-known technology widely used in Norway through ground-to-air heat pumps, but the extractable energy is small compared to that available from warm rocks at depths greater than 300 metres. There are presently no deep geothermal energy plants in Norway; a pilot plant begun at the Rikshospitalet University Hospital in Oslo was halted before completion because of unforeseen problems associated with the local geology. Many countries around the world exploit deep geothermal energy, but the challenge in Norway is that the gradient of heat increase with depth is rather low. Despite this, the potential is so vast that this source will undoubtedly be exploited over the next century. New modelling and exploration techniques using eInfrastructure resources will help.

Wind Energy

Wind energy is an easily exploitable energy source, and its use is growing rapidly. Wind energy research centres have been formed or expanded in many countries for enhanced exploitation. Optimal design requires coupling aerodynamical and structural models, which is computationally intensive. Models of wind systems from the regional to the microscale, including full time dependence, face challenges of tight coupling and massive parallelism. Recording both observational and geophysical structural data to monitor performance requires substantial storage capacity. Finally, the sporadic and highly variable energy source from wind farms needs coupling to energy storage systems (mainly water reservoirs) in better models of the electric power grid.



Fugure 25. Simulation of injection of a non-wetting fluid, e.g. gas, (red) into a micro-CT image of sandstone already initially filled with water. The first picture shows the entire sample while the second picture shows the distribution of fluids inside the pore space. The total size of the sample is 1.8 x 1.8 x 1.8 mm^3.

Autoignition in non-premixed combustion

Autoignition is important in many energy production systems, such as diesel engines and gas turbines. Until the early 1990s it was thought that diesel ignition is purely driven by the chemistry and unaffected by turbulence. It has subsequently been recognised that turbulence may affect the ignition time and flame development significantly. Careful modelling of the interactions between flow and chemistry will give deeper knowledge of how the fluid mechanics affects autoignition and assist the design of the efficient low-polluting combustion systems.

An important aspect of the interaction between turbulence and chemistry involves the location of ignition spots when re-lighting gas turbines. The combustion chamber should be designed to provide efficient mixing conditions enabling easy re-lighting if the turbine blows out. Three-dimensional Direct Numerical Simulations (DNS) have shown that there is a correlation between the location of the ignition spot (where the temperature is highest) and locations of low turbulent mixing rates[123]. Current computations are being done on a cluster with around 100 processors. If a cluster with around 100 000 processors were available, the Reynolds number could be increased significantly without loss of resolution. It would also be possible to study how the combustion points spread in the domain. This would be very important when designing auto engines with low emission rates and small fuel consumption. Also, it would be useful for the design of systems to capture carbon dioxide, to reduce the emission of greenhouse gases.

These are only a few examples of many that could be gleaned from industrial applications, including increased efficiency in processes ranging from transportation to refining and manufacturing.

Requirements for eInfrastructure

Secure data storage is important for resource extraction, geothermal energy, and carbon sequestration problems. Also needed is the capability for multi-scale computing on a large, tightly coupled system in the Petascale range. For the autoignition problem and energy efficiency in many other industrial applications, multi-scale computing will again require Petascale to Exascale systems. A full solution of the seismic inverse problem is possible, but only through massive use of computing power. Typically clusters with fast storage systems of the order of hundreds of TB and computing power of tens of Tflops are necessary.

Expectations from 2015

The availability of substantial computational resources will enable scientists in industrial applications to solve forward problems efficiently and thus contribute to the development of good models for the inverse solution methods require in seismic imaging. The design of reusable and interoperable components will facilitate the application of methods used in seismic imaging to a wider variety of geophysical inverse problems. Other industrial applications will benefit from similar approaches.

With increasing demands for greater energy efficiency and more effective use of resources, ever-improving computational infrastructures will become an increasingly vital requirement for Norwegian industry.

---

[123] Løvås, T., Lowe, A., Cant, R.S. and Mastorakos, E., Three-dimensional direct numerical simulations of autoignition in turbulent non-premixed flows with simple and complex chemistry, FEDSM2006-98108, ASME Joint U.S.-European Fluids Engineering Summer Meeting, Miami, FL, USA, (2006).

5. Concluding remarks

This survey of the present uses and anticipated future needs for the scientific use of and needs for eInfrastructure in Norway is necessarily incomplete. Our panel has solicited input from researchers who presently use components of the infrastructure, through personal contacts and by invitation to seminars. Our first seminar was hosted by the RCN on 14 December 2009, in conjunction with our panel's kickoff meeting. We held closed panel meetings in January and February, prior to another open seminar on 19 March 2010. Following that, we held two more closed panel meetings as we assembled and edited the contributions we had received into this present document.

We wish to acknowledge all those who contributed to those seminars, and those who have helped us in other ways, including the many one-on-one meetings we have each held with interested researchers[124]. Our time has been limited, however, and this document is too short to include all that we have learned. We have undoubtedly left out much that could have been said. Nevertheless, we feel we have assembled a fair, if incomplete, representation of the present state and future needs for eInfrastructure in Norway.

We will start from the present document as we assemble the Roadmap for the Future Development of eInfrastructure in Norway over the course of the coming year. Armed with this report, we will solicit more input, and then write a document which will aim to be more specific as to priorities for future development, guided by the articulated missions of the RCN and the Ministry of Education and Research. We will include comparisons with existing and future planned eInfrastructures in other countries, and highlight areas where interaction and cooperation are desirable and beneficial. Transnational eInfrastructure planning, especially regarding efforts in Europe such as eIRG and PRACE, will figure strongly in our deliberations over the next year.

---

[124] Specifically, we wish to acknowledge help and contributions from numerous individuals, including, from the University of Oslo: Mats Carlsson, Hans Eide, Morten Hjorth-Jensen, Janne Bondi Johannessen, Karen O'Brien, Stephan Oepen, Lars Oftedal, Christian-Emil Ore, Farid Ould-Saada, Ravidran Ponniah, Alex L. Read, Lynn Rosentrater, Bjørn H. Samset, Gunnar Wollan; from the Norwegian University of Science and Technology: Helge I. Andersson, Børge Arntsen, Ivar Ståle Ertesvåg, Helge Holden, Terese Løvås, Bernhard Müller, Arvid Næss, Håvard Rue, Einar Rønquist; from the University of Bergen: Inge Jonassen, Raymond Nepstad; from the University of Tromsø: Bénédicte Ferré, Kenneth Ruud; from Simula Research Laboratory: Xing Cai, Hans-Petter Langtangen; Kristin Bakken (The National Library of Norway), Idar Barstad (Uni Research), Laurent Bertino (Nansen Environmental and Remote Sensing Center), Dag Bjørge (Meteorological Institute), Paul Gibbon (Jülich Supercomputing Centre), Mats Hamrud (European Centre for Medium-Range Weather Forecasts), Håvard Helstrup (Bergen University College), Lars Holden (Norwegian Computing Center), Terje Lensberg (Norwegian School of Economics and Business Administration), Knut Andreas Lie (SINTEF), Hans Ekkehard Plesser (Norwegian University of Life Sciences), Thomas Ramstad (Numerical Rocks), Knut Røed (Ragnar Frisch Centre for Economic Research), Johannes Schweitzer (NORSAR), Oxana Smirnova (Lund University), Dieter van Uytvanck (Max Planck Institute for Psycholinguistics), Ingrid Helen Garmann Østensen (Norwegian Institute of Public Health); and we apologise to those we may have left off this list.

Figure Attributions

Figure 1. Wikipedia Commons, photograph by Bjarki Sigursveinsson.

Figure 2. Thomas Eickermann, Forschungszentrum Juelich, Germany (PRACE).

Figure 3. www.sailinganarchy.com/fringe/2006/images/alinghi%20new%20main.jpg

Figure 4. A Science-Based Case for Large-Scale Simulation, US Dept of Energy, volume 1, 2003.

Figure 5. TD Ringler, J Thuburn, JB Klemp, WC Skamarock, 2009, A unified approach to energy conservation and potential vorticity dynamics for arbitrarily-structured C-grids. J. Computational Physics, 229.

Figure 6. F Lindgrenn, J Lindström, H Rue, An explicit link between Gaussian fields and Gaussian Markov random fields: The SPDE approach, technical report.

Figure 7. Image combination by Anna Lipniacka from CERN archives.

Figure 8. Image combination by Anna Lipniacka from CERN archives.

Figure 9. Graph constructed by Anna Lipniacka.

Figure 10. Reproduced from U Landman, WD Luedtke, NA Burnham, RJ Colton, 1990. Atomistic mechanisms and dynamics of adhesion, nanoindentation, and fracture, *Science*, **248**, 454-461.

Figure 11. Lihao Zhao and Helge I. Andersson, Department of Energy and Process Engineering, NTNU.

Figure 12. Zinke, P., Olsen, N. R. B., Bogen, J. and Rüther, N. (2010) "3D modelling of the flow distribution in the delta of Lake Øyern, Norway", Hydrology Research, Vol, 41, No. 2, pp. 92-103.

Figure 13. Balram Panjwani, PhD candidate, Department of Energy and Process Engineering, NTNU.

Figure 14. S. Sikalo, C. Tropea, E.N. Ganic, 2005. Dynamic wetting angle of a spreading droplet. *Experimental Thermal and Fluid Science* **29**, 795-802.

Figure 15. Earth System Model Schematic from IPCC 2007 AR4: *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the 4th Assessment Report of the Intergovernmental Panel on Climate Change* [Core writing team, Pachauri R.K. and Reisinger A. (eds)]. IPCC, Geneva, Switzerland, 104pp.

Figure 16. From calculations by the Norwegian Meteorological Institute.

Figure 17. Derived from: IPCC 2007 AR4, cited above. The full caption legend, as required by the IPCC is: "Left panel: Solid lines are multi-model global averages of surface warming (relative to 1980-1999) for the SRES scenarios A2, A1B and B1 shown as continuations of the 20th century simulations. Bars in the middle of the figure indicate the best estimate (solid line within each bar) and the likely range assessed for the six SRES marker scenarios at 2090-2099 relative to 1980-1999. Right panels: Projected surface temperature changes for the early and late 21st century relative to the period 1980-1999. The panels show the multiGCM average projections for the A2 (top), A1B (middle) and B1 (bottom) SRES scenarios averaged over decades 2020-2029 (left) and 2090-2099 (right)." The acronyms and symbols in this caption are defined in the IPCC report.

Figure 18. Image from IPCC 2007 AR4, cited above.

Figure 19. Trond Torsvik, UiO.

Figure 20. Artist John Kaufmann: jek2004.com/Island-Arc.jpg.

Figure 21. Adapted from Palsson, BO, 2006. *Systems biology: Properties of reconstructed networks*. Cambridge University Press.

Figure 22. Adapted from C. Dollery and R. Kitney, *Systems biology: A vision for engineering and medicine*, Technical report, The Academy of Medical Sciences and The Royal Academy of Engineering, London, UK, Feb. 2007.

Figure 23. Courtesy of Martin Ystad, Tom Eichele, Erlend Hodneland, and Judit Haasz. See also: Lundervold A, 2010. On consciousness, resting state fMRI, and neurodynamics. *Nonlinear Biomedical Physics*, 4 (Suppl 1):S9.

Figure 24. Adapted from http://modha.org/blog/2009/11/post_3.html

Figure 25. Thomas Ramstad, Numerical Rocks.

The report can be ordered at:
The Research Council of Norway
P.O.Box 2700 St. Hanshaugen
NO-0131 OSLO

Telephone: +47 22 03 70 00
Telefax: +47 22 03 70 01
post@forskningsradet.no
www.forskningsradet.no

ISBN 978-82-12-02831-9 (printed version)
ISBN 978-82-12-02832-6 (pdf)