

More than ten years of scientific progress and impact

CONTENTS

	More than ten years of scientific progress and impact	3
1.	What is e-Science? What is SeRC?	. 4
	Data and software are redefining science	. 4
	SeRC - home of e-Science.	. 4
	SeRC's three-pillar model	. 5
	Organisation - our MCPs	
	Or current six MCPs	
	Infrastructure	10
	Management	.12
	Diversity and gender balance	.12
	Education	.12
2.	Impact	14
	Affecting academia and society	.14
	KPIs - list of scientific achievements.	.14
	Impact case: GROMACS - cited once every 70 minutes	.18
	Impact case: High-performance CFD with Nek5000	.19
	Impact case: Knowledge transfer in deep learning	20
	Impact case: Learning "hard" ness	
	Impact case: Brain simulation	.22
	Impact case: Cancer screening with Al.	23
	Impact case: Immersing the public in the scientific process	24
	Impact case: Climate modelling and weather prediction	25
3.	The way forward	26
	e-Science is continuously redefined	26
	Our MCP foundation is spreading	.27

info@e-science.se, www.e-science.se

Text: © SeRC. The contents may be quoted provided the source is clearly stated.
Photo/rendering: © SeRC and its affiliations except 8 fusebulb/Shutterstock.com
9 Gorodenkoff/Shutterstock.com, Elena11/Shutterstock.com, SFIO CRACHO/
Shutterstock.com 12 Rido/Shutterstock.com 13 Lightspring/Shutterstock.com
20 Zapp2Photo/Shutterstock.com 21 Dmitry Kalinovsky/Shutterstock.com 22 Giovanni
Cancemi/Shutterstock.com 23 Alexxndr/Shutterstock.com 25 Michael Tjernström
26-27 Syda Productions/Shutterstock.com

Editing, form, layout, illustrations: Gunnar Linn, Linnkonsult www.linnkonsult.se Print: Åtta.45 Tryckeri AB, Järfälla, 2022

MORE THAN TEN YEARS OF SCIENTIFIC PROGRESS AND IMPACT

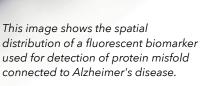
The Swedish e-Science Research Centre (SeRC) has been active for more than ten years. During this time, we have developed and tuned our operation and we are proud to present a strong and expanding portfolio of impact - on both scientific and societal levels.

We have managed to create and refine an efficient and highly competitive research environment with good gender balance and targeted education components for securing tomorrow's scientists within the area.

We also want to take the opportunity to highlight what we think is one main reason for the success of the centre, namely our flexible organisation into multidisciplinary collaboration programmes for seamless interaction with affiliated partners, tailor-suited to different applications, implemented throughout the entire operation, and powerful enough to meet the demands of the future.

Because science is evolving rapidly, with changing core technologies, increasing data volumes, and emerging application areas, e-Science must not only follow this evolution - but also stay ahead and be part of defining the paradigm shift.

We believe SeRC has the experience and capabilities to continue to define the state-of-the-art in e-Science, and are thrilled at this opportunity to showcase the amazing work carried out by the researchers in the centre.





Lean Meng/a

SeRC Director

Til Gild Otim Eiten

Erik Lindahl



Olivia Eriksson SeRC Coordinator

Dan Henningson

SeRC Co-Director

1. What is e-Science? What is SeRC?

This image and the one on page 5 show two defect orbitals for a point defect (divacancy) in silicon carbide (SiC). The green orbital is occupied in an excited state, and the purple orbital is occupied in the ground state. Si atoms are yellow, and C atoms are black. The use of computers and data in science is increasing fast. The concept of e-Science meets this evolution, and SeRC does the job.

DATA AND SOFTWARE ARE REDEFINING SCIENCE

How science is performed and understood is undergoing a disruptive change as the use of computers and data is becoming increasingly important in solving scientific problems. The result is e-Science, an important component in the complex ecosystem of rapid scientific progress.

e-Science is both the driving force and the main tool for the digital conception,

DEFINITION: e-SCIENCE

e-Science is the concept of computational method development, largescale data analysis and computational processing to gain scientific insight.

handling, and analysis of data, with human researchers in the loop to create qualitative insights and quantitative documentation from a data- and computational-driven analysis process.

e-Science is also the arena for convergence between high-performance computing (HPC) and data science. Formerly distinguishable, HPC is now producing data at an unprecedented scale and data science is now requiring computer capacity in par with the largest HPC applications. In addition, hardware and software for these two application areas are merging. This implies the convergence of methodology, developing in step for HPC and data science.

This metamorphosis of science into e-Science is constantly evolving with new scientific insights, applications and

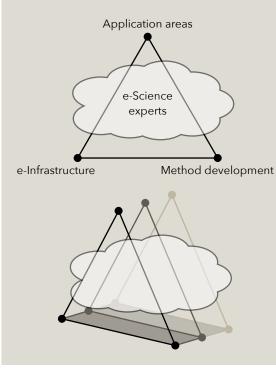
This evolving ecosystem is the home of SeRC, which continuously pushes the limits of what can be achieved."

methods being developed. This evolving ecosystem is the home of SeRC, which continuously pushes the limits of what can be achieved.

SeRC - HOME OF e-SCIENCE

SeRC is a research centre within the strategic research area (SRA) of e-Science, funded by the Swedish government Strategic Research Area Initiative and based on a collaboration between four universities: Kungliga Tekniska högskolan (KTH), Stockholms universitet (SU), Karolinska

SeRC'S THREE-PILLAR MODEL



The work in SeRC relies on three pillars of existing excellence in application areas, e-infrastructure, and method development.

Pillar 1 - application areas: SeRC's research is performed embedded within application areas of strategic relevance for our partner universities. This means that our application areas are selected with regard to both scientific and societal impact (see chapter 2). This impact also influences our organisation, since we have chosen a highly agile way to cluster our resources into so-called multidisciplinary collaboration programmes (see below), directly connected to the real-world challenges governing our research (see our impact cases in chapter 2).

Pillar 2 - e-infrastructure: To perform the necessary calculations, we need seamless access to world-class computational infrastructure.

Figure 1. Our three pillars are application areas, e-infrastructure, and method development. The two bottom pillars are mutual for different application areas. Conversely, since we are large cutting-edge users of these resources, we also need to be able to contribute to their development. For these reasons, SeRC includes the resources at a number of national infrastructure nodes (see pages 10-11) as native components.

Pillar 3 - method development: An integral part of SeRC is leading-edge research in method development, for instance in fields such as numerical analysis, visualisation, parallelisation, acceleration, and data science. The research groups contribute both with the latest research results on e-Science tools as well as a broad competence base in e-Science tool usage and methodology.

These three nodes are interconnected by human capital in the form of **e-Science experts**, which come in two flavours: embedded e-Science experts belonging to the application fields, and generic e-Science experts mainly connected to the computing e-infrastructure. The interaction between these experts and scientists in the core and application areas are of utmost importance for the success of any e-Science endeavour.

universitetThe first phase of SeRC was instrumen-
tal in establishing e-Science in Sweden and
in fostering the e-Science paradigm in many
different

Following the centre's continuing evolution, we are prepared to meet the high-end demands of the application areas."

research areas. Following the centre's continuing evolution, we are prepared to meet the high-end demands of the application areas and to deliver solutions to problems requiring advanced e-infrastructure and e-Science methodology, which we utilise both to create new knowledge within a number of different contexts and to define requirements for the further development of the e-infrastructure.

ORGANISATION - OUR MCPS

SeRC is based on the concept of a number of research **communities** (see figure 2), eight today, constituting the veritable arenas where our progress is being made. The communities represent the centre's current different research areas, continuously under

institutet (KI) and Linköpings universitet (LiU). It was founded in 2010 from the Swedish Government Bill on Research Policy. In this bill, a total of 24 different strategic research areas were defined, of which e-Science is one.

Designed as a virtual centre built up around the different projects and collaboration of approximately 50 principal investigators connected to the four academic players above, SeRC connects the higher end of e-Science activities in a broad spectrum, with demanding applications regarding both simulation and analysis.



Associate Professor of Climate Modelling at the Department of Meteorology, Stockholm University

From where did you come to SeRC?

I was introduced to SeRC by a colleague who had been, and still is, involved. e-Science plays an important role in climate research; we produce, analyse and use large amounts of data, and rely heavily on efficient e-infrastructure, so it is natural for our community to be part of SeRC.

What do you think is the major advantage with SeRC's MCP setup?

The multidisciplinary MCP structure allows for combining frontline expertise between core data scientists and specialists in a range of applications, and for finding common grounds, and problems and solutions, with other applications.

What has SeRC meant for your professional development?

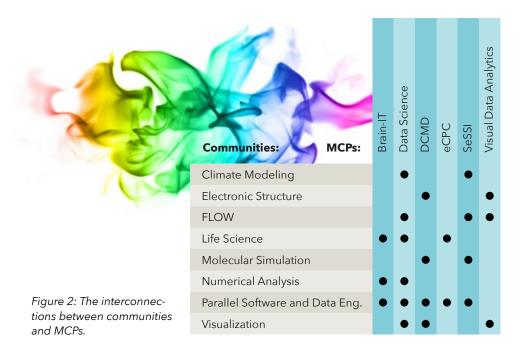
SeRC has given me a chance to explore e-Science applications that I would otherwise not have been able to invest time in. For instance, with seed funding from SeRC I have both initiated a project applying machine-learning methods to separate aerosol effects from meteorological influence on clouds, and contributed to securing a Europe-wide Horizon 2020 grant for developing next-generation climate models. re-evaluation to match the needs governing the e-Science domain, and each one of our affiliated researchers originally stems from one of these.

The current communities focus on core e-Science methodology, like Numerical Analysis, Visualization, or Parallel Software and Data Engineering, whereas other focus on some area of application of this methodology, like Molecular Simulation, Electronic Structure, FLOW (computational fluid dynamics), Climate Modeling, or Life Science.

SeRC has a focus on added value and bringing together strong research teams. To achieve this, the work is organised into **multidisciplinary collaboration programmes** (MCPs, see fact box and figure 2). The process of implementing the MCPs meant a significant upshift in utilising our human resources, aligning them in novel ways with the aim of boosting creativity and progress. With the MCPs in place, SeRC benefits from a common efficient

With the MCPs in place, SeRC benefits from a common efficient setup applying to all research areas and projects, influencing both work procedures and approaches."

setup applying to all research areas and projects, influencing both work procedures and approaches. The importance of this restructuring of our operation should not be underestimated (see more about this in chapter 3).



MULTIDISCIPLINARY COLLABORATION PROGRAMMES (MCPs)

The MCPs are designed to constitute the working units in which our three-pillar model is directly imposed on different research topics. The MCPs are the way we fill our three-pillar model with activity. In each MCP, there are a number of collaborating projects, where each more application-oriented project must clearly define how collaboration with a more methoddevelopment-oriented project is achieved, and vice versa. An important aspect within each MCP is the involvement of key e-Science experts. This way, substantial collaboration has been achieved between method-development researchers and researchers from application areas, directly contributing to the success of SeRC. The six current MCPs include **Brain-IT**, **Data-driven Computational Materials Design (DCMD)**, e-Science for Cancer **Prevention and Control (eCPC)**, SeRC **Data Science, SeRC Exascale Simulation Software Initiative (SeSSI) and Visual Data Analytics in e-Science applications** (see pages 8–9).

Since the MCPs represent a number of different ways to view the e-Science field, they are not mutually exclusive; there is some overlap between them – which is highly desired.

Read more about our organisation on e-science.se/org

Read more about our projects on e-science.se/projects

This image shows a partition of electronic density in molecular subgroups using Voronoi segmentation.

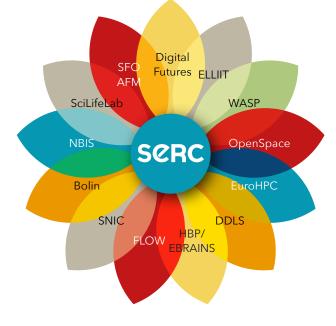
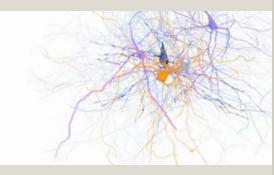


Figure 3: A 2020 snapshot of SeRCs affiliations, which are organisations "owning" the application areas in which we implement our MCP-based way of working and in which the responsible management are clearly familiar with SeRC's role and importance. Please note that the "petals" are exchangeable over time, depending on which projects we are engaged in.

CENTRE ORGANISATION

Board: Morten Dæhlen, Chairman, UiO, Anders Ynnerman, Vice-Chairman, LiU, Gunilla Svensson, SU, Juni Palmgren, KI, Jan Gulliksen, KTH.
Director: Dan Henningson, KTH.
Co-director: Erik Lindahl, SU/KTH.
Coordinator: Olivia Eriksson, KTH.
Management: Director, co-director, coordinator and a varying subgroup of the SeRC Pls.
Advisory Group: An international group of highly respected researchers from industry and academia. MCP BRAIN-IT

Coordinator: Jeanette Hellgren Kotaleski



Neuroscience research is currently producing data at an enormous pace and volume, measuring and quantifying structure, function and dysfunction of the brain at many levels. Neuroinformatics tools and computational approaches are needed to make sense of this vast amount of experimental data and integrate the resulting findings into a computational theory.

In MCP Brain-IT we develop, enhance and use e-Science approaches, including multi-scale modelling and scalable computer simulations, to facilitate bridging molecular-, cellular-, synaptic-, network-, and system-level perspectives in studies on brain plasticity, memory and learning, conducted together with computational neuroscientists, brain theoreticians, cognitive neuroscientists, and psychologists.

This SeRC research contributes to the understanding of learning and memory formation in the brain. This will have impact for brain-inspired Al and for the treatment of brain disorders, the latter because synaptic plasticity is impaired in both neurodegenerative and neuropsychiatric disorders.

Read more: e-science.se/mcp-brain-it

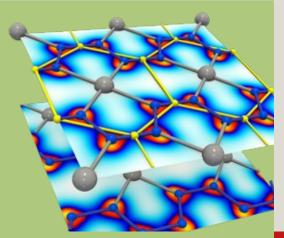
MCP DATA-DRIVEN COMPUTATIONAL **MATERIALS DESIGN (DCMD)**

Coordinator: Igor Abrikosov

New innovative materials are a crucial part of many emerging technologies. The integration of materials science with data-driven methods stands to be disruptive for the field of materials design. MCP DCMD coordinates efforts in this new and exciting research direction.

Our projects bring together condensed-matter theory and modelling, quantum chemistry, high-throughput computation, computer science, statistics, and visualisation to explore the adoption of semantic web techniques, machine learning, data mining, visual intelligent decision-support systems, and other methods.

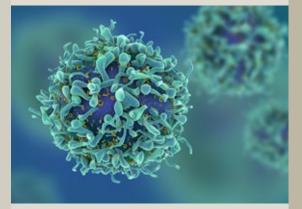
We expect these activities to result in the discovery of new materials, phenomena, and novel insights into the interrelations of materials properties. An overarching aim is to facilitate the design of materials for sustainability.



Read more: e-science.se/mcp-dcmd

MCP e-SCIENCE FOR CANCER PREVENTION AND CONTROL (eCPC) Coordinators: Mark Clements and Keith Humphreys





It is increasingly important to bring advanced computational methods into the medical sciences. SeRC early identified that e-Science could make a strong contribution to the field of cancer screening, cancer diagnostics, and cancer treatment.

Within MCP eCPC, a range of e-Science experts (mathematics, statistics, bioinformatics, computer science) work in integration with molecular scientists, epidemiologists and clinicians to develop a modularised e-Science framework for personalised screening and treatment.

Our top achievements include key contributions to the development and commercialisation of the Stockholm3 test for prostate cancer diagnostics as well as the development of an AIbased diagnosis and grading tool, an evaluation framework, and a large number of sophisticated simulation tools and algorithms.

Read more: e-science.se/mcp-ecpc

MCP DATA SCIENCE

Coordinator: Hedvig Kjellström

In addition to the ongoing simulation-oriented activity in e-Science, we are seeing a rapid increase in methods making use of real-life data. This is the field of work for MCP Data Science.

For some applications, for instance climate and flow modelling, the challenge is to extract knowledge from enormous amounts of data, such as learning approximation models of highly complex physical systems to enable faster prediction.

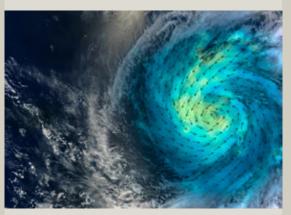
Often, datasets are so large that they cannot be stored completely in memory, but have to be analysed sequentially, which poses challenges in terms of hardware use and machine-learning methodology. In other applications, for instance medical diagnostics, data are instead often very sparse and incomplete, and the challenge is instead to "fill in the gaps" and make the most use of the few existing data.



Read more: e-science.se/mcp-data-science

MCP SeRC EXASCALE SIMULATION SOFTWARE INITIATIVE (SeSSI)

Coordinator: Berk Hess



The goal of MCP SeSSI is to improve performance towards the exascale of key scientific codes. Exascale computers will enable much larger computations than are currently possible, but most scientific codes are limited in scaling to 10,000-100,000 cores. Going beyond this requires deep and specialised knowledge about algorithms, communication patterns, and new hardware.

In SeSSI, we address this with a multi-disciplinary team focussing on a small set of scientific codes, including members from life science, mechanics, climate research, mathematics, and computer science.

This successful approach has led to very highly cited new versions of scientific codes, significant co-funding from national and international programmes, a joint "strong research environment", and several high-impact application publications.

Read more: e-science.se/mcp-sessi

MCP VISUAL DATA ANALYTICS Coordinator: Ingrid Hotz

Scientific and technical research relies on effective analysis of increasingly large and complex data. To avoid bottlenecks in the scientific process, it is important that the efforts spent on generating data is matched with the effort spent on data analysis, which, in turn, requires the development of technology and methodology.

The goal of MCP Visual Data Analytics is to develop visual analytics environments tailored to large-scale, complex, and dynamic data enabling interactive multi-scale analysis for knowledge discovery.

Fundamental concepts are interaction, filtering, and explorations on multiple levels of detail to support scientific reasoning, as well as scalability of the approaches in terms of complexity and data size, and integrated solutions for dynamic data in a coherent way.



Read more: e-science.se/mcp-vda



Professor of Computer Science at the Division of Robotics, Perception and Learning, KTH

From where did you come to SeRC?

I became engaged in SeRC in 2018, when SeRC expanded into the data-science area in addition to its original core in e-Science. Many different research projects in SeRC now use data-driven methods, and they complement the more traditional simulation-based e-Science approaches.

What do you think is the major advantage with SeRC's MCP setup?

The MCPs present an excellent way for researchers from different disciplines to meet and collaborate on concrete and relevant research problems. SeRC was highly functional even before the MCP design, but the setup really helps us to interact and engage. Research is about constantly building new knowledge through interaction with others, and the MCPs facilitate this.

What has SeRC meant for your professional development?

I have learnt a lot, and still learn a lot, about natural-science and life-science applications of machine learning and data science, and it is a fantastic multi-disciplinary environment. I don't know of any other research environment quite like this, and I'm proud to be part of it.

INFRASTRUCTURE

SeRC has an active part in four different hardware infrastructures, which we both utilise and help develop for the best possible national capabilities in the e-Science domain.

The Center for High Performance Computing at KTH (PDC) operates

leading-edge, high-performance computers on a national level. PDC offers easily accessible computational resources that primarily cater to the needs of Swedish academic research and education. To a large extent, the centre is funded by the Swedish Research Council via Swedish National Infrastructure for Computing (SNIC). PDC takes part in major international projects to develop high-performance computing for the future and stay a leading national resource in parallel computing. The centre also houses the SeRC Open Space (or SeRC room) - an area for informal interactions between researchers and e-Science experts - following the original SeRC idea of bringing researchers from various scientific domains, computer scientists, and high-performance computing



(HPC) practitioners together to share their expertise and assist and inspire each other's research. The PDC director is part of SeRC's management group.

The National Supercomputer Centre (NSC) provides leading-edge high-performance computing resources and support to users throughout Sweden and hosts Sweden's new large-scale resource for AI/ML, which is an NVIDIA SuperPOD funded by the Knut and Alice Wallenberg Foundation. The centre is an independent organisation within Linköpings universitet, to a large extent funded by the Swedish Research Council via SNIC. The NSC director is part of SeRC's management group. Visualization Center C is a research

and science centre in Norrköping, Sweden, conducting a unique mix of leading visualisation research and public outreach activities. The centre hosts Sweden's leading ennvironments for visualisaiton including VR and media labs, interactive exhibitions and a range of immersive environments such as a 3D stereoscopic fulldome theatre. The centre is a result of a close collaboration between the Municipality of Norrköping, Linköping university, Norrköping Science Park and Interactive Institute Swedish ICT. The SeRC-funded project OpenSpace is making heavy use of the dome theatre for astrovisualisation for research and public outreach. The centre is also the coordinator of the national project WISDOME connecting five science centres in Sweden, with funding from the Knut and Alice Wallenberg Foundation.

Visualisation Studio at KTH (VIC) is a resource for students and researchers. The studio houses many different technologies to view and interact with large amounts of data and its different applications, including haptics, eye-tracking, and various stereoscopic technologies. The studio is funded by a grant from the Knut and Alice Wallenberg's foundation, together with Norrköping's Visualization Center C and Linköping University.



Professor of Theoretical Physics at the Department of Physics, Chemistry and Biology, Linköping University

From where did you come to SeRC?

I'm one of the co-applicants for SeRC, so I was involved in the original proposal for the centre. I was very interested in the visionary idea of the SeRC centre, and also in "moving up" to e-Science, which is a much broader scope than materials modelling which was my scope at the time.

What do you think is the major advantage with SeRC's MCP setup?

The restructuring of SeRC into the MCP-based organisation and workflow meant as giant a leap forward for me as coming to SeRC had been in the first place. Everything fell into place: even better contact with the core technologies, even better collaboration and understanding between us people in the system.

What has SeRC meant for your professional development?

The true "gold medal" for a researcher is the knowledge that one has managed to create something that passes extremely high demands, like those applied by Nature magazine. I am very humbled by the fact that it would not have been possible without the highly functional environment of SeRC that I am deeply submerged into.

MANAGEMENT

The strategic, tactical, and operational governance of SeRC is characterised by active, engaged and committed management – for two main reasons: partly because the agile functionality of the centre requires a swift-footed leadership with substantial understanding of the great span of different contexts in which SeRC is active; and partly because our human capital is our best resource, which means that we are prepared to go to considerable lengths to keep our operation running smoothly, effectively and efficiently.

SeRC combines our own recruitment of principal investigators with the connection of excellent researchers already identified and recruited by our affiliated universities, in proximal and relevant areas. This makes us able to put extremely high trust in the people we recruit, which in turn has been instrumental for having all researchers engage in and contribute to jointly evolve SeRC's strategic direction.

SeRC is defined by the researchers participating in the centre, and evolves with them. We take particular pride in identifying and supporting junior researchers to help them build successful careers and contribute to our joint environment. SeRC aims to be an environment open and welcoming to all, where everyone feels safe and comfortable in sharing opinions and participating in the scientific discourse.

This all, in resonance with the manageable size of the centre where all colleagues are within reach, adds attachment and engagement to our staff. In addition to our own management resources, we use external consultants for developing the organisation as well as keeping up the "team spirit".

DIVERSITY AND GENDER BALANCE

The continuous improvement of diversity and gender balance on all levels in SeRC is one of the major goals of SeRC, in particular since computational topics have frequently had a severe deficit in female researchers. To address this goal, SeRC has introduced several actions ranging from activities to support and motivate researchers of all backgrounds to establish a more open and welcoming structural organisation.

- To improve diversity, particularily regarding gender, this has been included as a dimension in the development and evaluation of the MCPs.
- To attract female PIs to SeRC, newly employed assistant professors at the participating universities have been strategically invited to serve as PIs for new projects.
- To improve awareness of diversity challenges, a session at our annual meetings is typically devoted to the topic equal opportunities.
- To survey the current status and diversity in SeRC groups, a questionnaire on gender aspects was distributed and evaluated among the participants of the 2019 annual meeting.
- To attract the next generation of talented scientists to SeRC, the 2019 master students were invited to the same year's annual meeting with a balance of female and male participants. Some of these have been hired as PhD students.

These initiatives are accompanied with regular evaluations of the current situation.

Although the establishment of sustainable changes is a slow process, there is evidence of improved diversity, in particular that the proportion of female researchers has increased over SeRC's timespan.

EDUCATION

To ensure a wide scope of courses, SeRC collaborates with the EuroCC National Competence Center Sweden (ENCCS) in offering specialised workshops focussing on education. SeRC also collaborates with eSSENCE, a strategic collaborative research programme in e-Science between three Swedish universities (Uppsala University, Lund University and Umeå University), providing so-called SeSE courses (see fact box).

In collaboration with our academic departments and SeSE, SeRC offers PhD-level courses to educate the next generation of scientists in fields where the role of e-Science is either emerging or has already become indispensable. Since its inception, we followed the philosophy that SeRC courses should be offered on topics that were not available at the Swedish universities.

Up to 2019, SeRC were involved in two types of courses. The core courses focused on topics common to every application domain of e-Science such as scientific computing, data management, and data visualisation. These courses were open to a broad community of students. The specialised courses were focussed on the e-Science aspects of each specific field, for instance climate modelling. Thus, SeRC trained students in both basic and advanced topics depending on the needs of the PhD students not only in Sweden but in some cases also from outside Sweden. In general, core courses were offered more frequently than the specialised courses.

Over last few years, many of e-Science courses have been adopted as master-level or PhD-level courses at the universities. In particular, several core courses became redundant. Therefore, we have reorganised

SeSE

The graduate school Swedish e-Science Education (SeSE) provides education in fields where the use of e-Science is emerging and where such education can have an immense impact on the research as well as advanced training for students in fields that are already



computer-intensive. SeSE is jointly run by SeRC and eSSENCE.

The course curriculum is tailored to meet a broad set of prerequisites and will foster collaborations between Swedish researchers, utilising e-Science tools and methods and also making it possible for inter-disciplinary collaborations on a Nordic level.

One the most popular courses is the Introduction to High Performance Computing at PDC which annually attracts 30-50 students from Sweden and other Nordic countries.

Since 2015 over 550 students have taken e-Science courses, of which around 27% were female. About 10% of our students came from universities which are not part of the SeSE collaboration.

sese.nu

the course structure. The creation of the MCP structure (see pages 6–9) provided the perfect framework for this change. We now align the research activities within SeRC with our teaching efforts. Therefore, each MCP will offer at least one course per year within their area of expertise.

This setup ensures that e-Science education remains consistent with the research horizon in various fields, and results in a clearer identity for the courses we provide compared with what can be found elsewhere in academia.

Most of our PhD courses are available at master level, together with the specific master-level courses offered.

e-SCIENCE COURSES e-science.se/education Our human capital is our best resource. SeRC portait: INGRID HOTZ

Professor of Scientific Visualisation at the Division of Media and Information Technology, Linköping University

From where did you come to SeRC?

Before becoming part of SeRC, I led a group for comparative visualisation at the Zuse Institute in Berlin and then spend some time as the visualisation group leader at the German Aerospace Center (DLR).

What do you think is the major advantage with SeRC's MCP setup?

In contrast to typical interdisciplinary projects, the MCP design enforces us to build projects that go beyond one-to-one collaboration between core and application area. Everybody has to rethink their own work in general terms, which often leads to unexpected new ideas and generates synergies between different application areas that wouldn't naturally collaborate.

What has SeRC meant for your professional development?

SeRC made it very easy for me to settle in Sweden and find interesting collaboration partners. It became my scientific home. The long-term perspective provided by SeRC is a clear advantage. The fact that the work later is actually used in the applications is a rewarding experience that I rarely had before.

2. Impact

So, how does SeRC perform? How relevant are the centre's results for the surrounding world?

AFFECTING ACADEMIA AND SOCIETY

The major performance indicator for a research centre is its impact. For academic research, two main aspects of impact are relevant. On the one hand, we have the **scientific impact**, meaning the effect of our work on other academic and scientific work in terms of citations and similar accepted indicators. On the other hand, we have what we call **societal impact**, meaning the extent to which the scientific results are forwarded into society at large, and industrial development in particular.

Both aspects are of major importance for SeRC. A large part of our work also concerns method development, which is just as relevant for academic research as for industrial research, not to mention non-research applications. All these aspects contribute critically to the broad impact of SeRC, ranging from fundamental research to education, outreach and commercial usage of the results.

KPIS - LIST OF SCIENTIFIC ACHIEVEMENTS

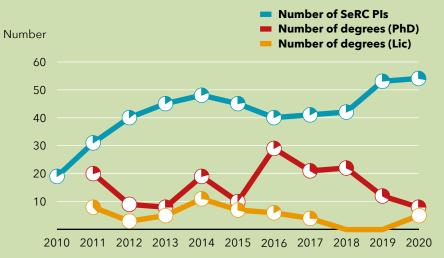


Figure 4: The number of SeRC-financed **principal investigators** per year since the start of the centre as well as the numbers of **academic degrees** achieved by SeRC-affiliated personnel (corresponding to highly qualified personnel moving on from SeRC to other organisations, transferring knowledge and experience). For each node, the gender balance is indicated as a small pie chart (filled part denotes proportion women).

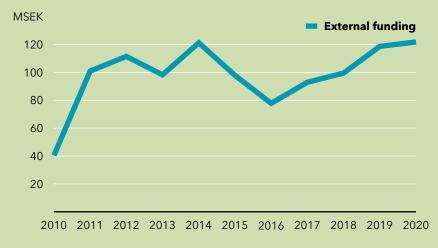


Figure 5: The external funding has risen from 40.4 MSEK 2010 to 121.8 MSEK 2020.

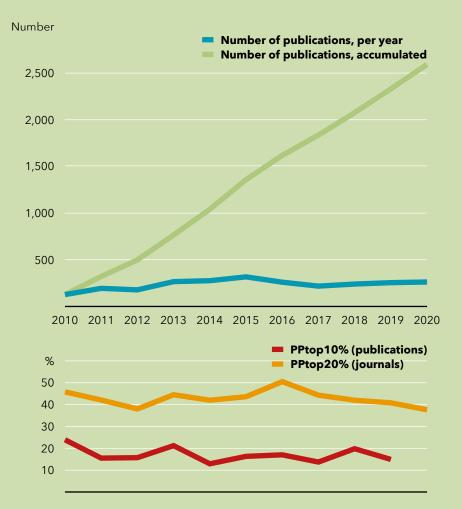


Figure 6: SeRC research has excellent scientific impact. The number of (peer-reviewed) **publications** since the start of the centre, both per year and accumulated, along with **distribution percentages.** PPtop10% indicates the proportion of publications in the 10% most frequently cited publications in a WoS category, compared with publications in the same category and year. Self-citations are excluded. Only articles and reviews are considered. PPtop20% indicates the proportion of publications in the 20% most frequently cited journals in a WoS category, compared with publications and year. Only articles and reviews are considered. The fluctuations are due to a few extremely high-impact papers.



Professor of Biostatics at the Department of Medical Epidemiology and Biostatistics, Karolinska Institutet

From where did you come to SeRC?

I was one of the founding members of SeRC and contributed to the writing of the first proposal. My own niche is quantitative modelling in the medical sciences. SeRC has contributed to my "dream come through" by offering the possibility to work in a group of scientists with extremely high skills in mathematics, statistics and computer science focusing on solving real-world problems.

What do you think is the major advantage with SeRC's MCP setup?

It provides a unique multidisciplinary perspective, incorporating computer science in its broadest sense to a number of applied science domains.

What has SeRC meant for your professional development?

I have thoroughly enjoyed the SeRC environment with junior and senior scientists from several universities and faculties coming together and slowly over the years forming an identity around e-Science. The "brainwashing" events regularly organised by the SeRC leadership have been highly motivating and interesting.

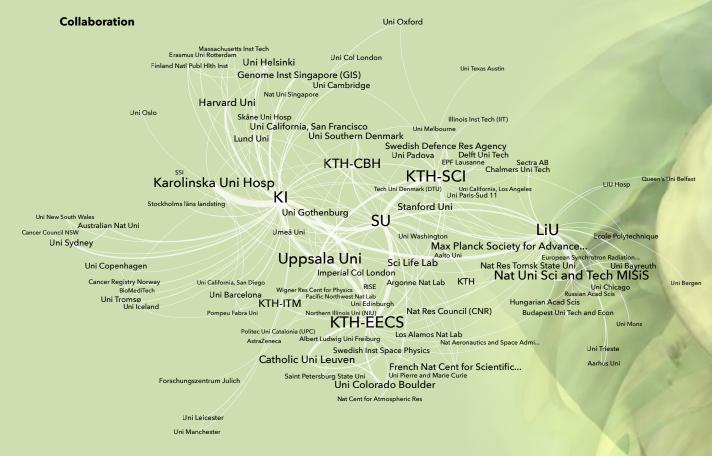
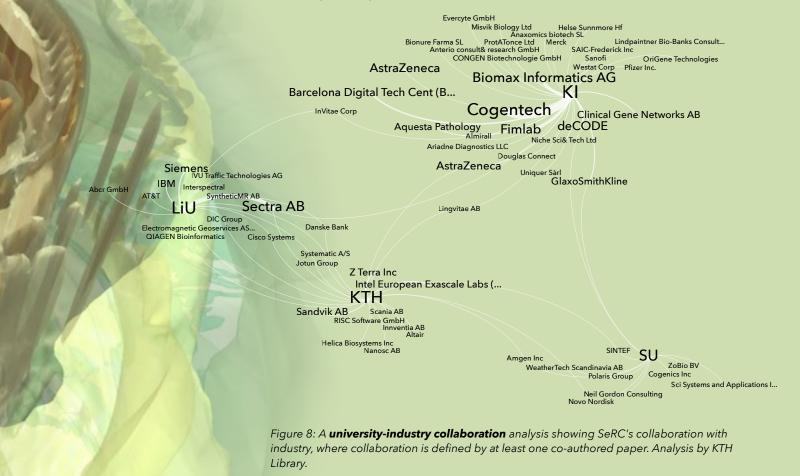


Figure 7: A **collaboration** analysis showing how scientific work has been cited between different players, meaning both scientific and societal impact. The figure shows organisations (nodes) with at least eight papers and linked by at least three co-authored papers. Publications with more than 25 organisations are excluded. The size of the nodes indicates the number of papers of each organisation, and the size of the lines shows to what extent the organisation collaborates. Analysis by KTH Library.

This image shows a visualisation of the storm Isabel over Florida using feature level-sets. The colours of the surfaces represents the severity of the weather conditions concerning high wind speed and high precipitation. The red surface highlights the area with the worst conditions. One can observe that the weather conditions are less severe in the eye of the hurricane.

University-industry collaboration



SERC IMPACT CASES

On the following pages, we will take a look at a few prominent examples of SeRC's impact. Note that this impact is of both **scientific** and **societal** nature, indicating the relevance of the centre to both the global academic community and society at large where we find the challenges that drive our research.

GROMACS - cited once every 70 minutes

This SeRC-developed supercomputer tool provides the necessary foundation for understanding the dynamics of molecular life science.

Performed within MCP SeSSI (see page 9)

Modern technology and research have transformed life science into a molecular field of research. However, understanding biology on the molecular level requires tools that can capture dynamics. SeRC's GROMACS tool enables the use of computers as virtual microscopes to study how molecules inside cells and viruses move and interact. The cross-disciplinary research area, called computational biophysics, involves extensive research in physics, computer science and biological applications.

But even traditional supercomputers do not provide sufficient performance to provide the necessary timescale resolution for the calculations. The SeRC GROMACS team was among the first in the world to invent new methods that could use graphics processor unit (GPU) accelerators to speed up simulations by an order-of-magnitude. To go with this, faster algorithms and new methods based on statistical mechanics were developed.

GROMACS has contributed to the establishment of a new generation of computational microscopes that are used by researchers all over the world. A key reason for this impact was the decision to make the source code available as open access. According to the **Intersect 360** market update at the

Stanford 2018 HPC Conference, GROMACS is the single most used of all generally available high-performance computing (HPC) applications in the world.

This has had additional major societal impact for the vendors developing the CPU and GPU hardware that form the core of supercomputers. In particular **NVIDIA** and **Intel** have long collaborated with the SeRC team to further improve this impact. The scientific impact of GROMACS is among the largest of any research ever performed in Sweden; the first paper has been cited over 12,000 times, the work introducing GPU accelerated simulations was the most cited Swedish scientific publication in 2014, and the 23rd most cited of all scientific publication citing this work roughly once every 70 minutes.

By far the most important impact of the GROMACS research is that the work has enabled thousands of researchers worldwide to increasingly use simulations for their fields of data-driven life science.

> Read a more comprehensive text on: e-science.se/ic-gromacs

SCRC IMPACT CASE

High-performance CFD with Nek5000

SeRC is investing in the continuous development of the open-source simulation code Nek5000 that is able to perform largest-scale turbulence simulations using the most powerful supercomputers.

Performed within MCP SeSSI (see page 9)

Turbulence and fluid mechanics are present everywhere from nano-sized laboratories to climate and astronomy. Simulations have gained a dominating influence on all research and development in the field, but extreme simulation resolution or sophisticated modelling is required to cover the extreme range of active length and time scales. This calls for the continuous development of large software packages.

SeRC's approach is focused around the open-source code Nek5000, a general-purpose Computational Fluid Dynamics (CFD) code that may be coupled to any other multi-physics framework, optimisation methods, and novel applications in machine learning and data science. Nek5000 has been developed since the mid 1980s and is a flavour of the well-known finite-element method.

Our groups within SeRC have used Nek5000 for more than 15 years in more than 25 PhD projects, and have performed some of the largest simulations on the system. The focus has been on algorithmic developments and methods to adapt the computational mesh during the simulation to account for where most grid points are needed, an approach called adaptive mesh refinement. Equally relevant is our development of a new version of the Nek5000 kernels to take advantage of the latest generation of computer architetures -GPU-based accelerators - as will become operational within **EuroHPC1**.

Fully resolved high-fidelity simulations as provided by SeRC's additions to Nek5000 have now reached the level of confidence to compete with the largest experimental facilities, at the same time providing more detailed data for advanced data analysis. Since such extremely large computations require more than 10 billion grid points, turbulence simulations can be viewed as one of the prime creators of large data sets, which in turn help improving engineering design or understanding, for instance, climate science, flow control, or biological flows. Overall, computational fluid dynamics and Nek5000 in particular are fine proof of the raison d'être and applicability of e-Science.

With Nek5000 as the backbone, SeRC has become an active partner in several EU projects and centres of excellence. We also had a number of collaborations with industries where Nek5000 was used as the method of choice. Examples that were directly driven from concrete industrial interest include novel topology optimisation of heat sinks to maximise natural convection, or, as shown in the image, simulations of an oscillating wing in a Saab-initiated project aiming to understand how changes in turbulent flow along a wing surface affect the control of an aircraft.

Read a more comprehensive text on: e-science.se/ic-nek5000

1 The European High Performance Computing Joint Undertaking (EuroHPC JU) is a joint initiative between the EU, European countries and private partners to develop a World Class Supercomputing Ecosystem in Europe. SCRC IMPACT CASE

Knowledge transfer in deep learning

Deep learning outperforms all earlier types of machine learning, especially with its capability - discovered by a SeRC researcher - to transfer knowledge from one task to another.

Performed within MCP Data Science (see page 9)

What was previously mere science fiction, such as self-driving cars and phones that let you log in just showing your face, is now reality. Machine learning and its progression deep learning are key components, containing techniques that process data for learning to understand gathered information combined with large (artificial) neural networks trained with, in turn, large volumes of data points. Deep learning has transformed the artificial intelligence (AI) and data science fields.

But does every specific application of deep learning require its own build-up of "knowledge" within the neural network? A SeRC researcher has found that it's not necessarily so.

In our original research, an AI model was first trained with a simple visual recognition task to assign an image to one of 1,000 available categories, covering a wide range of general visual classes such as man-made objects, animals, plants, and scenes. We observed the surprising result that the AI model also learned to solve other visual recognition tasks, such as finding smiling faces or specific activities in the images. The model could even retrieve other instances of a query object (such as the Statue of Liberty) across a large dataset of images.

The insight that what the model learns in one task can be extremely useful for other tasks operating on similar sensory information has proven to be extremely valuable in computer vision (automatic understanding of visual data). In fact, the original research paper constituted a key turning point for the entire field of computer vision, convincing prominent researchers to accept deep learning as the go-to method for image analysis. This is not least stated by the magazine **Nature** in May, 2015²:

"This success has brought about a revolution in computer vision; ConvNets are now the dominant approach for almost all recognition and detection tasks and approach human performance on some tasks."

This piece of research is listed among the top 100 most influential deep-learning works between 2012 and 2017³ based on the number of citations.

2 pubmed.ncbi.nlm.nih.gov/26017442

3 github.com/terryum/awesome-deep-learning-papers

Read a more comprehensive text on: e-science.se/ic-deeplearning

Learning "hard"ness

In this project, SeRC is developing a machine-learning approach to tailor-explore the big digitalised data of materials and accelerate the search of advanced coatings serving both the robustness and accuracy of manufacturing cutting tools.

The metal-cutting industry faces a quick shift of its market because of altered manufacturing demands and novel societal regulations. To operate in a time of such rapid changes, **Sandvik Coromant** and **Seco Tools** have adopted an approach based on digitalisation and smart materials design, which can be done by machine-learning algorithms exploring a digitalised materials database.

SCRC IMPACT

> SeRC is engaged in a common research programme with the companies, with the goal to reduce the time it takes to develop novel materials from the current 10-15 years to 5-6 years. The SeRC approach implements machine-learning strategies in search for hard materials, building on quantum-mechanics calculations with density-functional methodology, developing guiding concepts to classify the full materials dataset (big data) into regularised subsets in which machine learning can be utilised to learn hardness.

> We use large-scale computer simulations as well as novel calculation and data-management methodologies together with new data infrastructure. With an extensive database utilised by machine learning, we can also handle the challenge that hardness

is not the only property required for an efficient cutting tool. Proposed materials are synthesised and carefully characterised in state-of-the-art experiments. Moreover, in collaboration with the companies, we provide industrial verification of their cutting-tool performance.

The work has generated both significant insights and new and upgraded products, beneficial in end-user cutting applications. A synergy is that the increased competitiveness of the partaking industry Sandvik Cormant and Seco Tools has societal relevance in acting on globalisation and changing demographics.

On the general level, the main impact of our research is a paradigm shift within materials development. We are turning predictive theoretical search into a natural first step of the design process. The project yields scientific breakthroughs and overcomes multiple technological barriers to enable innovation and strengthen the competitiveness of Swedish industry.

Read a more comprehensive text on: e-science.se/ic-hardness Performed within MCP DMCD (see page 8) SCRC IMPACT CASE

Brain simulation

Unlocking the mysteries of the brain poses a lot of challenges that can only be faced by coordinated cross-disciplinary efforts including both experimental and computational science.

Performed within MCP Brain-IT (see page 8)

The brain is probably the most complex object in the known universe, not least considering its amazing learning and memory capabilities. Theoretical and computational approaches offer an integrative framework for connecting existing knowledge about the brain with new experimental data.

SeRC's work in the field is aimed at enabling such a framework with the help of efficient computing on computer clusters, methods for computer simulations and mathematical modelling on several scales. We also work in close collaboration with AstraZeneca on method development. A complementary part to this neuroscience research is an attempt to derive brain-like AI that builds on the developed cognitive brain theories. This effort is envisaged to contribute to the next-generation intelligent systems, better suited to work alongside humans.

late. This model will serve as a simulation platform to study complex phenomena at the cellular and network levels of the brain that are difficult to measure using real experiments. We have also predicted how vulnerable different brain regions are by constructing brain models capturing aspects of resting-state fMRI (functional magnetic resonance imaging) data, providing an opportunity to study the effects of many brain diseases and a model framework for computational studies of

putational studies of therapeutic brain-tissue stimulation. Furthermore, we develop theories and computational models of working memory,

a centrepiece of human

cognition. As a result, our

SeRC has made a number of noteworthy contributions. Together with experimental collaborators we reconstructed and simulated, as the first in the world, a nearly full-scale cellular-level model of the mouse striatum, a brain structure involved in decision making and motor control. Importantly, the whole model-building pipeline for this effort is released as open source and now adapted by **EBRAINS**⁴, so that other labs can reproduce the model when additional data accumumodel predictions about memory information processing have stimulated a paradigm shift in the field by inspiring a new model-based approach to analysis of brain recordings in cognitive tasks. A rapidly growing number of experimental publications exploiting our proposed method of single-trial analysis of oscillatory brain activity bears testimony to the SeRC scientific impact.

Read a more comprehensive text on: e-science.se/ic-brainsim

4 EBRAINS (**ebrains.eu**) is a pan-European research infrastructure resulting from the European flagship project The Human Brain Project (**humanbrainproject.eu**).

SERC IMPACT CASE

Cancer screening with AI

The Stockholm3 prostate-cancer test uses e-Science to boost efficiency, both in terms of effectiveness and costs, far beyond what traditional tests can provide. SeRC is one of the developers of the test.

For the past 25 years, most prostate cancer testing in Sweden has started with a prostate-specific antigen (PSA) test. However, the PSA test is known to have high rates of false positive results, which lead to unnecessary biopsies, and false negative results, which may lead to delayed detection of potentially treatable disease.

To reduce the harms and costs from PSA testing, SeRC and other researchers at **Karolinska Institutet (KI)** have developed a new prostate-cancer test (the **Stockholm3** test) which was designed in the STHLM3 and STHLM3-MRI diagnostic trials with the help of AI and machine learning for the prediction algorithm. The SeRC team provided input to the trial design, the study analysis, and the development of a microsimulation framework to evaluate the cost-effectiveness of the Stockholm3 test in which hundreds of cores can be used to simulate millions of individuals.

Around 40,000 Stockholm3 tests are currently being performed annually, and the entire county of Stavanger in Norway has abandoned PSA testing in favour of Stockholm3 testing. International validation studies are ongoing in Germany, the Netherlands, Denmark, Norway, Finland, the UK, and the US.

The baseline results of the research were published in 2015, demonstrating a reduction in the number of prostate biopsies by 32% while detecting the same number of advanced prostate cancers. In 2021, the SeRC team reported on (a) the effectiveness of magnetic resonance imaging (MRI) for prostate cancer diagnostics in the prestigious New England Journal of Medicine and (b) how the Stockholm3 test can be effectively combined with MRI. The integration of e-Science in complex clinical trial design has already had large local impact, and the medical e-Science component (including simulation) is developing into an international strength at KI.

We expect the results from these trials will have major impact on prostate cancer screening and diagnostics both in Sweden and internationally in the years to come. As an indication of the success of the approach, the Stockholm3 group was awarded an **EIT Innovators Award** in 2017 "for developing a product or service with a high potential for societal and economic impact".

Read a more comprehensive text on: e-science.se/ic-cancerscreening Performed within MCP eCPC (see page 8)

SCRC IMPACT CASE

Immersing the public in the scientific process

SeRC's combination of hardware, methods and experts is the perfect foundation for creating new ways of presenting science to the scientists of the future. OpenSpace is a great example of this.

The presence of role models can increase the interest of children and adolescents to become interested in the fields of science, technology, engineering, and mathematics to address the challenges of the 21st century. In many cases, practicing experts can serve as these role models, but it can be challenging for these subject matter experts to present to a wider public.

For this reason, we have in SeRC been developing several software platforms that bridge the gap between data-driven research and public outreach and enable scientists to perform presentations more efficiently and effectively.

One example is **OpenSpace**⁵, primarily targeting the astronomy domain, and with partners such as the **American Museum of Natural History** in New York and with co-funding from **NASA.** Our research aims at capturing the spirit of space travel through interactive visualisations, based on the development of a software platform that can serve both as a tool for novel visualisation research and a tool for domain scientists to present their data. The project develops new techniques to show complex data sets in new, easier understood ways, sparking the interest of discovery and joy in the next generation.

Performed within

(see page 9)

MCP Visual Data Analytics

The combination of self-driven discoveries and the ability to interact with subject matter experts generates impact far beyond traditional approaches to science communication. In 2019 alone, the collaboration partners organised about 75 live public presentations using the OpenSpace software that reached about 415,000 members of the general public, many of which are children.

OpenSpace has been featured in several events in Sweden. One example is the annual **Brilliant Minds** event that draws 750 of the most influential people in the world, at which SeRC's Anders Ynnerman was an invited speaker in 2019. The software has been distributed to some of most prestigious planetariums in the world, such as the **Adler Planetarium** in Chicago or the **Morrison Planetarium** at the California Academy of Sciences in San Francisco. Other partners include the **Iziko Planetarium** in Cape Town and the **Natural History Museum** in Vienna.

Read a more comprehensive text on: e-science.se/ic-openspace

5 openspaceproject.com

SCRC IMPACT CASE

Climate modelling and weather prediction

SeRC research makes us understand the processes leading to Arctic climate change amplification and how to predict weather in Arctic conditions.

The Arctic is experiencing about twice as much warming than the global average, but the reasons for this is not quite understood. And although our climate models are very helpful for the global climate change, they are not very good at representing the current climate in the Arctic. For future scenarios, the Arctic is the region where the uncertainty is the largest. The need for reliable prediction of weather and environmental variables such as sea-ice extent and thickness is increasing and of high societal relevance as the activities in the region is increasing with decreasing sea-ice area.

Whereas global climate models solve the necessary flow equations on a horizontal grid with boxes stacked in the vertical, SeRC has contributed to the development of a special kind of single-column model (SCM) that selects a single vertical stack of boxes from the bottom of the ocean to the top of the atmosphere at a location such that it can be studied in detail while the larger-scale flow is kept under control.

One of the novelties here is in the coupling between the atmospheric column and the ocean column into an atmosphere-ocean single-column model

6 ec-earth.org7 ecmwf.int8 mosaic-expedition.org

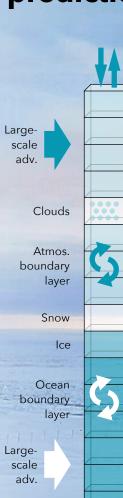
9 applicate-h2020.eu

(AOSCM) taken out of the global Earth system model **EC-Earth**⁶ that is developed in a European consortium where Sweden is a very active partner. The tool can be seamlessly integrated to also mimic the use for global forecasts on the medium-to-seasonal range at the **European Centre for Medium-range Weather Forecasts (ECMWF)**⁷. The new concept allows a more detailed comparison with observations and requires much less computational resources.

The concept is affecting climate-model development at large; at the latest **Multidisciplinary Drifting Observatory for the Study of Arctic Climate (MOSAiC)**[®] workshop, the modelling team concluded that AOSCM will be one of the primary modelling tools to be used to improve

regional and global models. Also, the AOSCM is listed in the Horizon 2020 database as one of the main results of the EU project **APPLICATE**⁹.

Read a more comprehensive text on: e-science.se/ic-climate



Performed within MCP SeSSI (see page 9)

25

3. The way forward

e-Science must continue to address needs in new application areas.

e-SCIENCE IS CONTINUOUSLY REDEFINED

The goal for e-Science is to enable scientific discovery by using the tools that modern computer technology and digitally available information offer. e-Science has become an essential and pervasive methodology in many domains and at the same time it is constantly expanding to new areas and paving ways for new research efforts in virtually all academic disciplines.

The future need for refinement and expansion of the e-Science paradigm as an interface to the rapidly changing underpinning tools and information sources is underlined by the on-going convergence of HPC and data science, combined with the ever-increasing need for data analysis.

For the strategic development of SeRC, we therefore need to spearhead the inclusion of new methods, data and workflows under the e-Science umbrella and thus continue to support data-driven human scientific insight. As an example of this, two main trends require attention:

Firstly, the core technologies around us, constituting the landscape in which e-Science navigates and operates, is changing. New components are constantly introduced, data volumes are increasing, and the era of machine learning is here. e-Science needs to embrace the potential offered by inclusion of intelligence-driven science and in this contribute to the paradigm shift for scientific discovery offered by the combination of human and machine intelligence.

Secondly, the number and types of

applications of e-Science are changing just as rapidly. We must address needs in new application areas and be swift-footed to include new areas in which the introduction of e-Science methodology can lead to break-through science.

OUR MCP FOUNDATION IS SPREADING

The MCP setup is a highly valuable foundation for the future of SeRC in particular and e-Science endeavours in general. From the vantage points of our MCPs, we are continuously scanning our surroundings – near and far – for possible novel applications, with steadily increasing knowledge and experience to meet new demands in new environments.

Bearing in mind that the number and nature of our current MCPs are not necessarily the same as tomorrow's, we realise that the MCP concept is a very flexible tool for bringing together the right minds, the right tools, and the right methods for any application. Through our MCPs, we make "expeditions" into application areas, not only for passive exploration but also for active development, not only into "secure" areas but also into worlds we haven't seen before.

The MCP concept of involvning diverse reserach groups from both method development and application areas to focus on a joint larger scientific goal has been a great catalyst of new cross-disciplinary collaborations in SeRC.

Other initiatives in Sweden are also applying concepts similar to our MCPs.

For instance, the Swedish Wallenberg Al, Autonomous Systems and Software Program (WASP)¹⁰, the Datadriven life science (DDLS)¹¹ strategic initiative from Knut and Alice Wallenberg Foundation (KAW) and the SciLifeLab¹² research community programmes. The success of the SeRC MCPs has served as proof of concept and inspired management and senior researchers in these initiatives to strive for multi-disciplinarity and shown a way to the next generation of successful research environments.

When plans for a strategic international collaboration initiative emerged between the Saõ Paolo and Stockholm regions, spearheaded by Karolinska Institutet (KI) and the Royal Institute of Technology (KTH) within the areas of life science and engineering, the use of the MCP concept was a natural choice. Identifying MCPs where not only researchers from method-development and application areas collaborate, but where these researchers come from different countries, will be a new challenge with great benefits. Having such a clear focus in future internationalisation will ensure that the impact will be long-term.

In short, we have only seen the beginning of what the flexible MCP concept can deliver.

10 wasp-sweden.org 11 scilifelab.se/data-driven 12 scilifelab.se



Professor of Fluid Mechanics at the Department of Mechanics, KTH

From where did you come to SeRC?

I was involved in the original centre proposal around 2009, contributing some minor parts on CFD simulations. At the time, I worked with a new supercomputer project at KTH and SU, so were were right in the topic.

What do you think is the major advantage with SeRC's MCP setup?

Large projects involving different areas, like the MCP setup, are very interesting, potentially yielding more "disruptive" output. Also, the setup has clearly strengthened the connections between scientists and infrastructure.

What has SeRC meant for your professional development?

SeRC has been instrumental in my development, both from the actual "funding" perspective, allowing me to focus on the e-Science work, and in providing a relevant network in Sweden. The continuity in the centre, and the developments we can experience as a team together, is unique compared to other centres or teams that I have been part of. I also really enjoy the strategy meetings where we can all actually come up with relevant input for the development of the centre. This image shows how molecular dynamics simulations with GROMACS enable SeRC researchers to model how small drugs (red) bind to the nicotinic acetylcholine receptor (orange/blue) in lipid membranes to change signalling in our nervous system, which has important potential applications to treat for instance depression or abuse disorders.



e-science.se