Department of Computer Science

Delivering environmental information for decision support in dynamic conditions

Ville Kotovirta

Aalto University

DOCTORAL DISSERTATIONS

Aalto University publication series
DOCTORAL DISSERTATIONS 252/2016
VTT SCIENCE 143

Delivering environmental information for decision support in dynamic conditions

Ville Kotovirta

A doctoral dissertation completed for the degree of Doctor of Science (Technology) to be defended, with the permission of the Aalto University School of Science, at a public examination held at the lecture hall T2 of the school on 10 December 2016 at 12.

Aalto University
School of Science
Department of Computer Science

Supervising professor
Professor Juho Rousu, Aalto University, Finland

Thesis advisors
Professor Emeritus Olli Simula, Aalto University, Finland
Dr. Jaakko Hollmén, Aalto University, Finland

Preliminary examiners
Docent Seppo Kaitala, Finnish Environment Institute, Finland
Dr. Mika Sulkava, Natural Resources Institute Finland, Finland

Opponent
Professor Jukka Riekki, University of Oulu, Finland

Aalto University publication series
DOCTORAL DISSERTATIONS 252/2016

VTT SCIENCE 143

ISBN 978-952-60-7168-8 (printed)
ISBN 978-952-60-7167-1 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)
http://urn.fi/URN:ISBN:978-952-60-7167-1

ISBN 978-951-38-8480-2 (printed)
ISBN 978-951-38-8479-6 (pdf)
ISSN-L 2242-119X
ISSN 2242-119X (printed)
ISSN 2242-1203 (pdf)
http://urn.fi/URN:ISBN:978-951-38-8479-6

Cover art by Christina Lindberg

Unigrafia Oy
Helsinki 2016

Finland

NORDIC ECOLABEL
441 697
Printed matter

Aalto University

Aalto University, P.O. Box 11000, FI-00076 Aalto www.aalto.fi

Abstract

Author
Ville Kotovirta

Name of the doctoral dissertation
Delivering environmental information for decision support in dynamic conditions

Publisher School of Science

Unit Department of Computer Science

Series Aalto University publication series DOCTORAL DISSERTATIONS 252/2016

Field of research Computer and Information Science

Manuscript submitted 23 August 2016 **Date of the defence** 10 December 2016

Permission to publish granted (date) 13 October 2016 **Language** English

☐ **Monograph** ☒ **Article dissertation** ☐ **Essay dissertation**

Abstract

This thesis studied improvements in the timely delivery of relevant near real-time environmental information to support decision making in dynamic environments. Especially, the focus was on orchestration of the data processing tasks, presentation of the information to end-users, and data collection from the field where the users are operating. It was found that three system design principles can be used to improve the information delivery: 1) organising the synergies in data access and information processing as a component called Data operator, 2) including automatic analyses of the situation to support the interpretation of the information, and 3) harnessing of end-users and end-user devices as opportunistic and participatory sensors to collect data from the local conditions in order to complement other data sources.

The research was conducted by studying two application cases: ice navigation and water quality monitoring. In the ice navigation case, we developed a system to deliver in-situ and remote sensing data as well as forecasts by computational models about the meteorological, oceanographic and ice conditions to ice-going ships, a route optimisation method to support the decision making and information presentation, and a method for using ships and ship radars as a sensor network. In the water quality monitoring case, citizens were harnessed as observers of water turbidity and the algae situation in order to complement other data sources, and citizen observations were compared with expert observations.

Open data and open interfaces are important elements for accessing data, but they are not adequate to guarantee the optimal use of environmental data in near real-time applications. The whole processing chain from data sources to end-user awareness should be considered in order to take full advantage of the data. It is concluded that the three design principles are not limited to the application cases of this study, but are applicable to other domains of environmental monitoring as well, for example air quality, disaster and built environment monitoring. The amount of environmental data is growing exponentially, and new methods are needed to include these data in decision making in society.

Keywords environmental monitoring, environmental informatics, computer science, remote sensing, participatory sensing, ice navigation, route optimisation, water quality monitoring

ISBN (printed) 978-952-60-7168-8 **ISBN (pdf)** 978-952-60-7167-1

ISSN-L 1799-4934 **ISSN (printed)** 1799-4934 **ISSN (pdf)** 1799-4942

Location of publisher Helsinki **Location of printing** Helsinki **Year** 2016

Pages 137 **urn** http://urn.fi/URN:ISBN:978-952-60-7167-1

Aalto-yliopisto

Aalto-yliopisto, PL 11000, 00076 Aalto www.aalto.fi

Tiivistelmä

Tekijä
Ville Kotovirta

Väitöskirjan nimi
Ympäristötiedon välittäminen tukemaan päätöksentekoa muuttuvissa ympäristöolosuhteissa

Julkaisija Perustieteiden korkeakoulu

Yksikkö Tietotekniikan laitos

Sarja Aalto University publication series DOCTORAL DISSERTATIONS 252/2016

Tutkimusala Informaatiotekniikka

Käsikirjoituksen pvm 23.08.2016 **Väitöspäivä** 10.12.2016

Julkaisuluvan myöntämispäivä 13.10.2016 **Kieli** Englanti

☐ **Monografia** ☒ **Artikkeliväitöskirja** ☐ **Esseeväitöskirja**

Tiivistelmä

Tässä väitöskirjassa tutkittiin oikea-aikaisen ympäristötiedon välittämistä tukemaan päätöksentekoa muuttuvissa ympäristöolosuhteissa. Työssä keskityttiin koko prosessointiketjun tiedonkäsittelytehtävien hallintaan, tiedon esittämiseen loppukäyttäjille sekä tiedon keräämiseen paikallisista olosuhteista. Tutkimuksessa havaittiin, että soveltamalla kolmea periaatetta järjestelmän suunnittelussa tiedon välittämistä käyttäjälle voidaan parantaa: 1) organisoimalla synergiset tiedonkäsittelytehtävät komponentiksi nimeltä Dataoperaattori, 2) vähentämällä välitettävän ja esitettävän tiedon monimutkaisuutta automaattisella tilanteen analyysillä ja 3) valjastamalla loppukäyttäjät laitteineen keräämään muita tietolähteitä tukevaa tietoa paikallisista olosuhteista.

Työssä tutkittiin kahta sovellusta liittyen jäänavigointiin ja vedenlaadun seurantaan. Jäänavigointia varten toteutettiin järjestelmä välittämään satelliittikuvia, maanpäällisiä havaintoja, sekä sää-, jää- ja meritilanteen malliennusteita jäissä kulkeville laivoille, reitinoptimointimenetelmä helpottamaan moniulotteisen tiedon esittämistä käyttäjille ja menetelmä laivojen ja laivatutkien valjastamiseksi antureiksi keräämään tietoa jäätilanteesta. Vedenlaadun seurantaan kehitettiin järjestelmä kansalaisten vedenlaatuhavaintojen keräämiseksi järvi- ja merialueilta täydentämään muita tietolähteitä, ja kansalaisten tekemiä havaintoja verrattiin ammattilaisten havaintoihin.

Avoin data ja avoimet rajapinnat ovat tärkeitä elementtejä ympäristötiedon hyödyntämisessä, mutta tiedon optimaalinen käyttö päätöksenteossa muuttuvissa ympäristöolosuhteissa vaatii koko prosessointiketjun suunnittelua tietolähteiltä käyttäjälle esitettäväksi. Työssä esitetyt kolme periaatetta eivät rajoitu tutkittuihin sovelluksiin, vaan ovat sovellettavissa laajemmin ympäristömonitoroinnissa, esimerkiksi ilmanlaadun, luonnononnettomuuksien tai rakennetun ympäristön monitoroinnissa. Ympäristötiedon määrä kasvaa jatkuvasti ja uusia menetelmiä tarvitaan tiedon koko potentiaalin hyödyntämiseksi yhteiskunnassa.

Avainsanat ympäristömonitorointi, ympäristöinformatiikka, tietojenkäsittelytiede, kaukokartoitus, kansalaishavainnointi, jäänavigointi, reitinoptimointi, vedenlaadun monitorointi

ISBN (painettu) 978-952-60-7168-8 **ISBN (pdf)** 978-952-60-7167-1

ISSN-L 1799-4934 **ISSN (painettu)** 1799-4934 **ISSN (pdf)** 1799-4942

Julkaisupaikka Helsinki **Painopaikka** Helsinki **Vuosi** 2016

Sivumäärä 137 **urn** http://urn.fi/URN:ISBN:978-952-60-7167-1

Preface

That's one small step for mankind, one giant leap for me. This thesis wraps up the work I've been involved in for many years. The leading theme has been the delivery of relevant information to end-users operating in dynamic environments. The story began with ice navigation, which is crucial for Finland due to its northern location, and continued towards the monitoring of natural and urban environments, which has global interest due to climate change and urbanization.

A wise man once said that postgraduate students should not lead a team at the same time as they are striving to complete their studies. I dared to disagree and continued dreaming of a doctoral thesis while acting as a research team leader and a project manager. Looking back, and looking at the calendar, I have to admit that the wise man, Professor Tuomas Häme, was right. However, when one of the bigger projects ended, I had an opportunity to sit down, look back at the results and wrap them up as a book. I would like to thank my instructors Professor Olli Simula and Dr. Jaakko Hollmén and my supervisor Professor Juho Rousu for making this book a reality.

This journey would not have been possible without the contribution and support of many talented and passionate people. I was lucky to join the group led by Robin Berglund that developed information delivery for Finnish and Swedish icebreakers and helped to keep the Finnish and Swedish economies running even during the winter time. The systems developed for ice navigation with Jyrki Haajanen and Markus Laakso spun off to other projects. Together with Timo Toivanen, Markku Huttunen and Renne Tergujeff we pioneered mobile methods for collecting environmental observations from ordinary citizens. With Heikki Turtiainen and Dr. Tero Eklin we planned and led a whole research program to study environmental monitoring, and the concept of Data operator emerged during many fruitful discussions, especially with Veli-Pekka Luoma. Cooperation with Harri Hytönen led to a Data operator prototype and Docent Mauno Rönkkö was there to push results towards publications. I thank Dr. Jari Silander, Timo Pyhälahti, Yrjö Sucksdorff, Dr. Olli Saari, Dr. Tommy Jacobson, Dr. Eero Punkka, Matthieu Molinier, Heikki Pentikäinen, Markku Mikkola, Jukka Hemilä, Dr. Janne Saarela, Atso Haapaniemi, Antti Aalto, Janne Mikkonen, Dr. Markus Stocker, Okko Kauhanen, Professor Mikko Kolehmainen, Associate Professor Kostas Karatzas, Professor Venkatachalam Chandrasekar, Dr. Esko Juuso, Dr. Thomas Casey, Jenni Mansner, Teppo Veijonen, Simo Neuvonen, Ari Seinä, Lic.Sc.(Tech) Risto Jalonen, Dr. Lars

Axell, Professor Kaj Riska, Dr. Juha Karvonen, Dr. Rüdiger von Bock und Polach, Professor Pentti Kujala, Dr. Marko Järvinen, Matti Lindholm, and Dr. Kari Kallio.

I also acknowledge financial assistance from VTT Technical Research Centre of Finland, the Finnish Funding Agency for Innovation and the Academy of Finland, and cooperation with the Environmental Informatics group of the University of Eastern Finland, the Finnish Meteorological Institute, the Finnish Environment Institute, CLIC Innovation Ltd, Vaisala Oyj, HiQ Finland Oy, and the Measurement, Monitoring and Environmental Efficiency Assessment (MMEA) consortium.

Finally, I express my gratitude to pre-examiners Docent Seppo Kaitala and Dr. Mika Sulkava and to my opponent Professor Jukka Riekki.

Thank you dear friends for lifelong memories (and for all the memories that I no longer remember), for all the music and jamming, for keeping me running after the ball, and for never-ending discussions about life, the universe and everything.

Thank you my dear family for your love and support.

Ville Kotovirta

Espoo, 3 November 2016

Contents

List of Abbreviations

AIS	Automatic Identification System
BPEL	Business Processing Execution Language
CEP	Complex Event Processing
CI	Confidence Interval
EC2	Amazon Elastic Compute Cloud
ENVRI	Environmental Research Infrastructure
EPL	Event Processing Language
ESA	European Space Agency
ESB	Enterprise Service Bus
FMI	Finnish Meteorological Institute
GEO	Group on Earth Observations
GEOSS	Global Earth Observation System of Systems
GIS	Geographic Information System
GPS	Global Positioning System
HTTP	Hypertext Transfer Protocol
IBPlott	Icebreaker Plotter
IBNet	Icebreaker Network
ICE	Ice Conditions Equivalent
INSPIRE	Infrastructure for Spatial Information in the European Community
ISO/IEC	International Organization for Standardization / International Electrotechnical Commission
MMEA	Measurement, Monitoring and Environmental Efficiency Assessment research program
ODP	Open Distributed Processing

OGC	Open GIS Consortium
RDF	Resource Description Framework
SETI	Search for Extraterrestrial Intelligence
SMS	Short Message Service
SOA	Service-Oriented Architecture
SQL	Structured Query Language
SWE	Sensor Web Enablement
SYKE	Finnish Environment Institute
UI	User Interface
VGI	Volunteered Geographic Information
VSAT	Very-Small-Aperture Terminal
VTT	VTT Technical Research Centre of Finland
WSDL	Web Service Description Language
XML	Extensible Markup Language

List of Publications

This thesis is based on the following original publications which are referred to in the text as Publications **I**–**V**. The publications are reproduced with kind permission from the publishers.

I Kotovirta V., Kanniainen J., Veijonen T., Neuvonen S. 2006. Building an Application Framework for Monitoring the Environment. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2006, 31 July - 4 August 2006, Denver, Colorado, USA, pp. 2149-2152.

II Berglund R., Kotovirta V., Seinä A. 2007. A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea. Canadian Journal of Remote Sensing 33(5): 378-387.

III Kotovirta V., Jalonen R., Axell L., Riska K., Berglund R. 2009. A system for route optimization in ice-covered waters. Cold Regions Science and Technology 55(1): 52-62.

IV Kotovirta V., Karvonen J., von Bock und Polach R., Berglund R., Kujala P. 2011. Ships as a sensor network to observe ice field properties. Cold Regions Science and Technology 65(3): 359-371.

V Kotovirta V., Toivanen T., Järvinen M., Lindholm M., Kallio K. 2014. Participatory surface algal bloom monitoring in Finland in 2011–2013. Environmental Systems Research 3(24).

Author's Contribution

Publication I: Building an Application Framework for Monitoring the Environment

Kotovirta analysed the requirements, designed the framework architecture and participated in and led the development of the framework. Kotovirta participated in the design and developed the first versions of the Facade component.

Publication II: A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea

Kotovirta was the main developer of the IBPlott system (Icebreaker Plotter, a decision support tool for icebreakers). Kotovirta designed the Facade component in the data delivery architecture together with Berglund. Kotovirta participated in the development of the Façade component.

Publication III: A system for route optimization in ice-covered waters

Kotovirta designed the architecture, participated in the optimization framework development, integrated the ship transit model and the optimization framework to the ViewIce system (a decision support tool for merchant vessels) and developed the route visualisation component. Kotovirta performed the validation based on AIS (ship-borne Automatic Identification System) data.

Publication IV: Ships as a sensor network to observe ice field properties

Kotovirta composed the first idea of using ships as sensors together with Berglund. Kotovirta led the development, designed the architecture and integrated the components of ship transit model, ship radar image capture, image mosaicing and ice drift analysis into a prototype system.

Publication V: Participatory surface algal bloom monitoring in Finland in 2011–2013

Kotovirta composed the first ideas about developing a participatory sensing platform and led the development of the platform which was the basis of the LeväVahti application used for algae and water turbidity monitoring. Kotovirta participated in composing the first ideas of applying participatory sensing to algae monitoring in Finland. Kotovirta performed the data analysis of comparing observations by citizens and experts.

The thesis overview: In addition to Publications **I-V**, this thesis introduces previously unpublished ideas concerning extending the Facade model in Publications **I** and **II** towards the Data operator model. Kotovirta composed the first ideas of Data operator and led the discussions in the Measurement, Monitoring and Environmental Efficiency Assessment (MMEA) research program. Kotovirta led and participated in the architecture design and implementation of the Data operator prototype based on Enterprise Service Bus (ESB) architectural model.

The thesis manuscript has been proofread by Michael Bailey / Semantix Oy. I have personally examined and accepted/rejected the suggested modifications one by one. This has not altered the scientific content.

Guaranteed by

Ville Kotovirta

1. Introduction

From time to time we are surprised by the environment. We get caught in the rain, slip and slide on icy roads when we did not change winter tyres in time, we may be disappointed when going for a swim in an algae-choked lake, or have difficulties to breath when jogging unaware of pollen or bad air quality in the area. Not only citizens struggle with changing environmental conditions, but companies and businesses are also affected by the dynamics of the environment, and unpredictable events can cause harm and losses. On a larger scale, societies are sometimes surprised by environmental hazards such as storms, fires, landslides, earthquakes and tsunamis.

There are on-going activities that try to reduce the number of environmental surprises. The availability and interoperability of environmental data are increasing, as regional and global initiatives aim towards integration of environmental systems and opening up various data reserves. For example, Open Geospatial Consortium (OGC) aims at interoperable interfaces enabling sharing of environmental data. OGC Sensor Web Enablement (SWE) in particular concerns accessing and controlling various distributed sensors by standard interfaces (Botts et al., 2008). Global Earth Observation System of Systems, GEOSS, is a world-wide voluntary effort coordinated by the Group on Earth Observations (GEO) secretariat, aiming at global connectivity of already existing systems monitoring the environment and storing environmental data (GEO secretariat, 2010). Copernicus (previously known as Global Monitoring for Environment and Security, GMES) is the European Union contribution to GEOSS, and aims at achieving an autonomous Earth observation system consisting of remote (satellite) and in-situ sensors. The Data Observation Network for Earth (DataONE) is developing a system to support data discovery and access across diverse data centres distributed worldwide (Michener et al. 2012). Current trends are towards opening publicly produced data, including environmental data. For example, in Europe the INSPIRE directive (Infrastructure for Spatial Information in the European Community) obliges European public organizations to open up their environmental data sources for applications (European Parliament, Council, 2007). As a result of INSPIRE, in Finland, the Finnish Meteorological Institute (FMI) opened up an extensive amount of meteorological data, and the Finnish Environment Institute (Syke) followed with other environmental data such as hydrological data.

How is it possible that the environment still surprises us, even though we are surrounded by sensors and huge amounts of environmental data, data sources are opening up and open standards have been developed for sharing data, and we have mobile devices connected to the Internet in our pockets? The hypothesis behind the thesis is that open data and open interface standards are necessary elements, but they are not sufficient to take full advantage of environmental data. The whole environmental information delivery chain from data sources to end-user awareness should be considered and new system architectural models and design principles developed and applied for delivering the right information to the right place at the right time in the right format. The thesis studied two application cases, ice navigation and water quality monitoring. As a main contribution it was found that information delivery could be improved by three system design principles: 1) organising the synergies in data access and information processing as a component called Data operator, 2) including automatic analyses of the situation to minimise the amount of data delivered, to reduce the complexity of information presentation and to support interpretation of the information, and 3) harnessing of end-users to collect data from the local conditions in order to complement other data sources.

1.1 Research context, research questions and research objectives

This thesis examines improvements in near real-time environmental information delivery to support decision making of both professionals and laymen in dynamic environments. We focus on two application cases: 1) ice navigation, in which end-users are professionals making navigational decisions on board ships in ice-covered waters, and 2) water quality monitoring, in which end-users are citizens making decisions about recreational activities and authorities monitoring water quality. The thesis uses multidisciplinary approach in solving the research questions. It connects to the field of environmental monitoring (Artiola et al., 2004) and applies remote sensing and participatory sensing (Conrad and Hilchey, 2011) to ice monitoring, ice navigation and water quality monitoring (Karydis and Kitsiou 2013). The thesis focuses on near real-time aspects of environmental monitoring, not on collecting, using or analysing long-term time series. However, longer time series can be collected as a side product when e.g. participatory sensing methods are applied. The thesis does not contribute to environmental problem research, but aims at delivering data about the surrounding environmental conditions to support decision making related to users' activities and possible problems while operating in the environment. The thesis connects to the field of environmental informatics (Frew and Dozier, 2012) by finding ways to improve current state of the art in accessing the multitude of environmental data sources and delivering environmental information to users. However, it does not improve environmental data storage systems or access to long-term environmental time series, but concentrates on existing near real-time data sources and the delivery of relevant near real-time environmental information using modern methods of computer science. The general research question of the thesis is:

How could near real-time environmental information delivery be improved to reduce the amount of surprise and better support decision making of both professional and layman end-users operating in dynamic environments?

While considering the general research question we came up with more specific research questions addressing different aspects of the general question. The information delivery to end-users requires a multitude of processing tasks that need to be orchestrated in order to address the challenges of near real-time environmental information delivery. The first specific research question is:

Q1: How could the orchestration of information processing be implemented to address the challenges of near real-time environmental information delivery?

After the information processing, an important component of the information delivery chain is the information presentation to the user. All the time-varying observational and forecast information should be presented in an understandable and intuitive way to support the decision making. The second specific research question is:

Q2: How could the information be presented to support situation awareness and improve decision making?

End-users operating in the dynamic environment are the experts of the local conditions. Experienced users operating in the area can interpret the local situation and its evolution better than a generic model on a computer server, and even unexperienced users can collect relevant data to support the situational picture. In addition, the devices and vehicles used in the field can contain instrumentation that could be utilised for monitoring the local environment. The third specific research question is:

Q3: How could end-users be harnessed to collect additional data from the local environment to complement other data sources in the information processing?

The next task is to define research objectives that address the research questions. The research objectives are divided into a general research objective and three specific research objectives that are logical parts of the general objective. The general research objective addresses the general research question and it is:

To develop and study near real-time environmental information delivery architectures for the application cases of ice navigation considering professional end-users and water quality monitoring considering both professional and layman end-users.

The specific research objectives that address the specific research questions are:

O1: To develop orchestration of the information processing to tackle the challenges of near real-time information delivery.

O2: To develop the presentation of relevant information while minimising the complexity of multidimensional information in order to support the interpretation of the situation.

O3: To harness end-users and end-user devices as data collectors in order to complement other data sources.

1.2 Outline of the thesis

The thesis consists of an overview and five publications. Section 2 defines a reference architecture for near real-time environmental information delivery and discusses the challenges of developing such a system, Section 3 summarizes the implementation of the specific research objectives, Section 4 discusses the findings in relation to research questions and Section 5 presents the summary and conclusions. The details of the implementations are given in Publications **I-V**. Publication **I** implements the first specific research objective and depicts the initial information delivery orchestration architecture with the initial concept of Facade. Publication **II** implements the second specific research objective for the ice navigation application case and describes the data delivery architecture as well as information presentation, Publication **III** contributes to the second specific research objective and considers minimising the amount of data used in the presentation by ice routing, i.e. route optimisation for ice navigation. Publications **IV** and **V** implement the third specific research objective by developing an opportunistic data collection method for ice navigation and a participatory data collection approach for water quality monitoring. In addition to Publications **I-V**, the overview of the thesis presents the Data operator model as a new result. Data operator extends the original concept of Facade (Publication **I**) from a single application point of view towards multiple applications and takes into account data collection from users.

2. Near real-time environmental information delivery

In this section we define the concept of near real-time environmental information delivery and present a reference architecture that defines key elements and their relations in the domain at a high level of abstraction. We also discuss the challenges of developing such systems and describe the contributions of the thesis in relation to the reference architecture.

2.1 Conceptual view

The thesis connects mainly to the fields of environmental monitoring and environmental informatics. Environmental monitoring was defined by Artiola et al. (2004) as follows: “Environmental monitoring is the observation and study of the environment. In scientific terms, we wish to collect data from which we can derive knowledge”. Environmental knowledge is needed by society in conjunction with different processes and at multiple levels, when the activities are dependent on the environmental conditions or are likely to affect the environment on a local or global scale. Environmental knowledge is relevant to many businesses, such as energy production, production of goods and services, traffic, logistics, navigation, disaster mitigation, forestry and mining of raw materials, and also to the daily activities of citizens.

Environmental monitoring is originating from the need to understand the state and progress of the environment and the effect of human activities in the environment. As environmental problems have become worse and sensor technology has become more advanced and available, more and more environmental data have been produced by environmental monitoring. The amount of data has raised a need of collecting, storing, managing and analysing the data, and the field of environmental informatics has emerged (Hilty and Page, 1995). More recently, Frew and Dozier (2012) defined environmental informatics as “the application of data science to environmental problems.” By data science they mean collection, management, exploitation, communication and preservation of environmental information about the state of Earth’s biosphere (and associated spheres) consisting of large complex multidimensional datasets.

Artiola et al. (2004) refer to “the staircase of knowing” by Roots (1997), which includes steps of observation and measurement, data, information, knowledge, understanding and wisdom. Observation and measurement lead to data through verification, data lead to information through selection and testing, information leads to knowledge through organization and interpretation, knowledge leads to understanding through comprehension and integration, and understanding leads to wisdom through judgment. The contribution of the thesis mainly concentrates on the first steps of the staircase of knowing, i.e. observation and measurement, data and information. We harness end-users and end-user devices as sensors to complement other existing data sources and consider how data are processed into information. The thesis also contributes to the latter steps of knowledge and understanding, as we consider how the information is presented to the user to increase the knowledge and understanding of the situation and improve decision making. There are many definitions for data, information and knowledge (e.g. Zims, 2007). In this thesis, by "data" we refer to data that have potential value to end-users but have not yet been processed into a usable form, and by "information" we refer to data that have been processed into a usable form and have value for decision making.

We concentrate on delivering near real-time information about the surrounding environment for decision making in a dynamic environment. Delivery architectures integrate with existing near real-time environmental data sources and process data into relevant information (Figure 1). The data sources include for example measurement data collected by in-situ and remote sensors, interpolations or forecasts calculated by computational models, analyses compiled by human experts and data produced by citizen scientists. The data sources can be built specifically for the application or they can be general sources such as satellite systems and weather models providing data via open interfaces for a multitude of applications. The information delivery chain reduces the amount of data and increases the amount of information, i.e. usefulness or value of the data from the end-user's point of view (the scales and forms of the graphs in Figure 1 concerning the data amount and data value are only illustrative).

By the term near real-time we refer to systems that deliver data to end-users in seconds, minutes or hours after the actual measurement or model run is available from the data sources. In other words, we concentrate on time scales which Artiola et al. (2004) defined as daily (>24 hours), hourly (>60 minutes) and instantaneous (<1 second). This differs from environmental monitoring applications that collect historical time series of some parameters for long-term analysis, in which the latest measurements are not as important as the earlier measurements of a long time series. These time scales were referred to by Artiola et al. (2004) as seasonal (>4 months), annual (>1 year), generation-lifetime (20-100 years) and geologic (> 10,000 years). In this thesis we are discussing applications in which the value of the produced data for the user is decreasing as a function of time in minutes or hours, as the environmental conditions evolve.

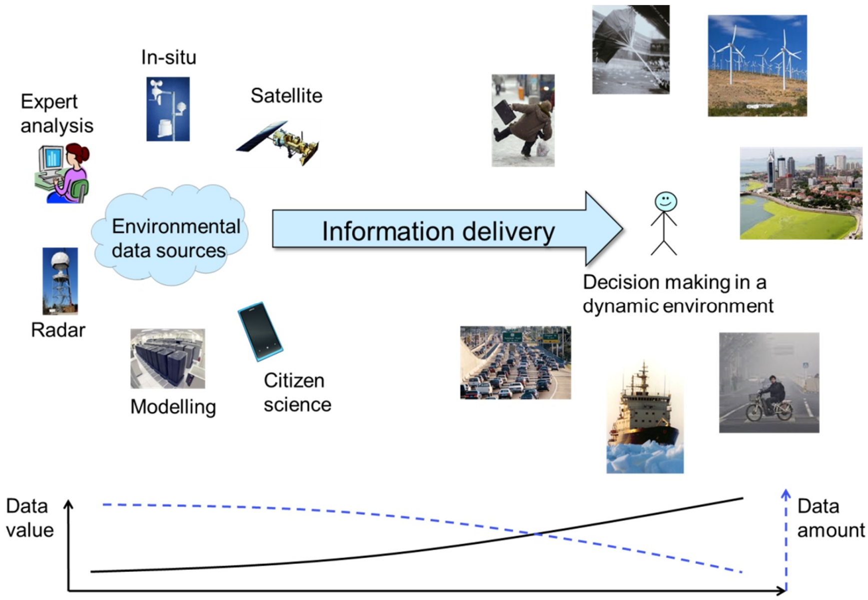

Figure 1. A conceptual view of near real-time environmental information delivery. Data coming from sensors, models, human experts and citizen scientists are processed and personalized to end-users so that the data volume is minimized and the data value, i.e. the amount of information, is maximized.

Example applications conforming with the conceptual view presented in Figure 1 include a system for ice navigation support that delivers satellite data, model data and in-situ observations to ice-going ships (Publications **II**, **III**, **IV**); a participatory algae monitoring system that complements remote sensing data and expert measurements with citizen observations (Publication **V**); an air quality monitoring system that collects air quality measurements from distributed sensor networks and generates visualizations as well as sends alerts to citizens about the changing air quality (Lim et al., 2012); a tsunami early warning system that collects seismic observations, buoy observations and tide gauge observations about water level from various areas to produce customized warning messages for delivery via different channels such as the web, TV broadcasting, SMS and e-mail (Wächter et al., 2012); a system for tornado monitoring and forecasting that collects weather radar and meteorological observations, as well as weather forecasts, and produces near real-time visualizations of the developing situation for weather services and emergency response personnel (Plale et al., 2006); and a near real-time water quality management system that monitors flow and salinity of surface water deliveries and seasonal wetland drainage and soil salinity in surface soils (Quinn et al., 2010).

2.2 Reference architecture

We present a reference architecture that defines the key elements of a near real-time environmental delivery system and their relations at a high level of abstraction. Different specific architectures for near real-time environmental information processing and delivery have been presented before: Cao (2009) proposed use of the Enterprise Service Bus (ESB) architectural model to handle the data collection from distributed sensors developed by a variety of manufacturers relying on various standards; Motwani et al. (2010) discussed use of the ESB architectural model and complex event processing (CEP), the so called event-driven SOA, as a means of achieving extensibility and interoperability in a heterogeneous environmental sensor network; Usländer et al. (2010) proposed a service-oriented architecture called the Sensor Service Architecture (SensorSA) based on OGC SWE and aiming at integration of various data sources, whether in-situ, airborne or spaceborne sensors, or data storages; Lim et al. (2012) proposed an architecture applying ESB and CEP to make an air quality monitoring system more interoperable, scalable and stable; Lee et al. (2010) discussed the use of cloud computing to deal with varying computing resource needs due to changing environmental conditions; Suakanto et al. (2012) proposed an architecture using cloud computing in sensor data processing for disaster early warning; Li and Wu (2011) proposed a service-oriented architecture for configuring data processing chains automatically in cases in which the requested data do not yet exist. The data processing integrates sensors, automatic processing and manual processing to fulfil the user's complex tasks in near real-time.

The ENVRI reference model environmental science research infrastructure presented by Chen et al. (2013b) aims to model the "archetypical" environmental research infrastructure. The model identifies five subsystems including data acquisition that collects raw data from sensor arrays, various instruments, or human observers, and brings the measurements (data streams) into the system; data curation that facilitates quality control and preservation of scientific data; data access that enables discovery and retrieval of data housed in data resources managed by a data curation subsystem; data processing that aggregates the data from various resources and provides computational capabilities and capacities for conducting data analysis and scientific experiments; and community support to manage, control and track user activities and help users to conduct their roles in communities.

The ENVRI model defines relevant elements and terminology for modelling environmental data systems. However, in this work it is important to emphasize the near real-time aspect of the information delivery, the role of end-users in producing additional data to complement other data sources, and the need for orchestrating the data processing efficiently and robustly. Figure 2 provides a reference architecture at a high level of abstraction that zooms in to the information delivery arrow presented in the conceptual view of Figure 1. The arrow is divided into two separate arrows pointing in opposite directions describing the direction of information and data flows from the end-user's point of view. One of the arrows describes the part of the system that delivers the processed information to users, while the other describes the data collection from users to complement other data sources in information processing. In the information delivery chain the relevant data sources are interfaced and relevant data are retrieved to be processed into relevant information from the end-user's point of view. Information processing may include many processing steps such as rectification and calibration (of satellite images), data quality control, compression (lossless or lossy), filtering, data fusion, machine learning, pattern recognition, modelling and forecasting. Finally, the information is delivered to end-users and presented in a useful way, for example graphically or textually, to support decision making. The data interfacing may include or precede a process of agreeing about interfaces, data formats, data prices and the terms of data usage. In the data collection chain, end-users and end-user devices are provided with interfaces to enable participatory and opportunistic data collection, and the data are processed (for example data quality control, filtering, compression) and published as a data source to the information delivery system or other systems that might be interested in the data. The whole processing chain of delivering information and collecting data is controlled by the orchestrator component that manages the workflow of processing and communication between the processing components. The orchestrator can be a script gluing the processing components together, or it can be a more advanced software that controls the processing execution, handles the exceptions, warns of possible problems with the processing, communicates with human operators and end-users about the status of processing and guarantees as robust information delivery as possible. All processing components can be distributed and they can contain manual tasks or be totally autonomous.

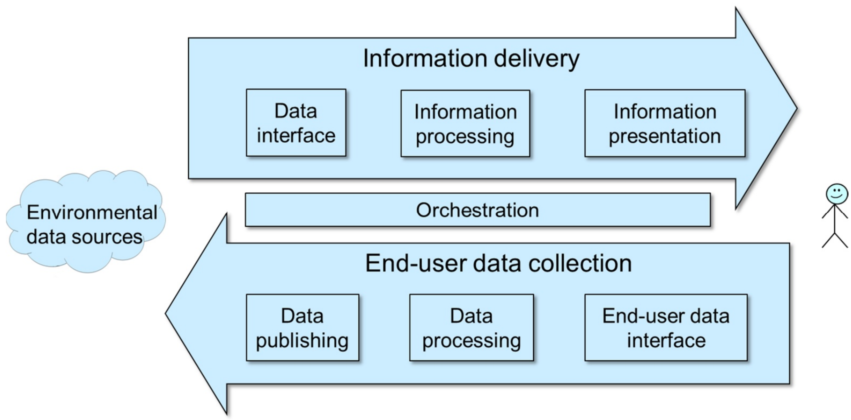

Figure 2. The reference architecture of near real-time environmental information delivery at a high level of abstraction. The main components of the information delivery chain include a data interface for accessing and retrieving data, information processing and information presentation. Data collection from end-users includes an interface to collect data from end-users, data processing and publishing data as a data source available to the same or possibly other information delivery systems. The orchestration controls the processing and communication between the components of both information delivery and data collection.

2.3 Challenges of near real-time environmental information delivery

In this section we analyse general challenges of near real-time environmental information delivery that guide the design choices of the implementations. These are based on requirements for distributed near real-time environmental monitoring systems presented by Kotovirta et al. (2013), but they are repeated here for convenience and elaborated. Related work has been carried out by Chen et al. (2013a), who analysed requirements for environmental science research infrastructures.

We categorised the challenges by using four viewpoints that capture different aspects of the whole system under consideration. The viewpoints are from the Open Distributed Processing (ODP) framework, an international standard for distributed system specification published by ISO/IEC (ISO/IEC 10746-1, 1998). The ODP framework uses five viewpoints to specify particular concerns of the whole system: the Enterprise Viewpoint, which concerns the organisational situation in which business is to take place, the Information Viewpoint, which concerns modelling of the shared information manipulated within the system of interest, the Computational Viewpoint, which concerns the design of the analytical, modelling and simulation processes and applications provided by the system, the Engineering Viewpoint, which tackles the problems of diversity in infrastructure provision, and the Technology Viewpoint, which concerns real-world constraints applied to the existing computing platforms on which the computational processes must be executed. We use the first four viewpoints, and do not here consider the technological challenges which are part of the implementation. Other viewpoints have also been suggested for modelling environmental information systems. Rönkkö et al. (2013) presented a method for classifying environmental monitoring systems and recognised three relevant viewpoints: application domain, functionality and architecture. Chen et al. (2013b) chose the science viewpoint, information viewpoint and computational viewpoint for modelling environmental science research infrastructures.

2.3.1 Challenges from the enterprise viewpoint

Integration of heterogeneous distributed systems. Environmental data are produced by many sensors and sensor networks, including in-situ sensors (fixed, ad hoc, wired, wireless), people using their mobile phones (i.e. citizen science and community-based environmental monitoring, Conrad and Hilchey, 2011), and remote-sensing sensors such as radars and satellites. In addition, computer models produce forecasts based on the measurements, and human experts and researchers provide analyses based on the numerical data produced by sensors and numerical models. Before the information is presented to end-users by the end-user applications, it is collected and processed by various distributed processing components. All in all, an environmental monitoring system needs to handle the integration of distributed heterogeneous systems managed by different organizations for retrieving the data and processing and delivering the information.

Extensibility. During the operation of the system, new data sources or processing services may become available that provide additional relevant information for the end-users. The monitored environment may behave unpredictably, and data input and output as well as data processing needs may grow as a result of a sudden event in the environment. The system should be extensible so that new data and processing can be included. In addition, extensibility is required when new end-users and end-user applications are introduced to the system.

Cost efficiency. In principle the value gained from a monitoring system defines how much the system can cost. Some critical applications must be operating reliably with strict data security regardless of the costs, but in many cases costs restrict the architectural choices of the system and thus define its overall performance and usability. The system should be economically sustainable, so that at least in the long run the benefits of the system are greater than the costs. The use of cost-efficient solutions and re-usable components in the architecture can decrease the overall costs, and still provide good enough performance to meet the requirements set by the application domain. In addition to the development costs, the running costs, such as data processing, data transmission, data storage, and maintenance costs, are also relevant.

2.3.2 Challenges from the information viewpoint

Relevance of data delivery. The data delivered to the end-user should contain only the relevant information, i.e. the data are personalized to the user needs, and it should be presented in a way that supports adequate comprehension of the situation. Any useless data will burden the data communication channel, cause additional costs, require additional time from the user to check the data, overload the user with too much information and impair user understanding of the situation. This may imply a push type of service that monitors the environment on behalf of the user and notifies the user of any relevant changes in the environment. The system can even be proactive, delivering relevant information that the user was even not aware of (Rönkkö et al., 2012). The amount of data delivered may depend on the environmental situation. The data processing chain might be operating fluently, but still the user does not receive any data until something relevant occurs in the environment and data delivery is initiated. In order to deliver relevant information, the system should know the user needs, the user role, and the user context including e.g. the device and software that the user is using, the location of the user, and the data transmission line bandwidth.

Data quality / reliability. Data quality is crucial, especially when important decisions are made on the basis of the delivered information. The quality of the information produced by the data processing chain is dependent on the data quality not only of the raw data sources, but also of the computational components of the processing chain. Williams et al. (2011) discussed the importance of describing and exchanging uncertainty information when developing loosely coupled, interoperable environmental monitoring systems. Quality control is required to ensure that the input data for a processing component is of the expected quality. Any critical deviations, for instance due to malfunctioning measurement devices, statistical deviations due to noise, or computational results exceeding the expected value ranges, are to be detected and reported.

2.3.3 Challenges from the computational viewpoint

Timely data delivery. The system should deliver information in time, i.e. the resources reserved for data processing, data storage and data transmission should be adequate for the time requirements of the application area and the end-users. In general, the value of the latest observations and forecasts in dynamic environmental conditions decrease as a function of time, although the criticality of timely delivery varies between applications. For example, forecasted dispersion of pollen in an area for the next day can be reported with some delay, but information about a toxic release related to a chemical accident in a nearby factory should be delivered as quickly as possible. In ice navigation the situation may be static for days, or change in hours or even minutes in harsh weather conditions. However, even if the situation is stable, information about the observed and predicted stability should be delivered without delay in order for operators to be able to make the right decisions in the field.

Scalability. When new data sources, processing and end-users are introduced in the system more computing power, data storage and data transmission capacity is required. The system should therefore be scalable to handle the increased amount of data and processing. The scalability goes both ways – the resources should be scaled down if the amount of input data, processing or the number of end-users is decreased.

2.3.4 Challenges from the engineering viewpoint

Configurability. New uses of the system may be achieved, not by adding new data sources or processing components to the system, but by re-configuring the system's existing components. Furthermore, the system's performance may be adjusted or optimized through configuration – of data computing resources, data storage capacity, end-user roles, or security policy, for example.

Operational reliability. The system should be successful in delivering the information in every case, as a failure to deliver critical information may be costly, even life-threatening. In addition, there should not be false alarms as these will reduce user trust in the system. For example, ships navigating in ice are dependent on updated information in order to choose safe and cost-efficient routes. Allergic people must remain aware of surrounding air quality in order to take their medicine in time, and people living close to a factory where a chemical release has just taken place must be alerted about the accident and advised to take cover.

Data security. The more valuable the information produced by the system is, the more important is the data security of the data processing chain. The key principles of data security must be included, i.e. confidentiality, integrity and availability. Only authorized parties should be able to access the data and the processing components; the data should be protected from any malicious party trying to steal or alter them. When important decisions are made on the basis of the data or computed events, users should be guaranteed that the data are unchanged and that no system is pretending to be a reliable data source in order to mislead the user or cause harm. As the data processing chain consists of loosely coupled processing components managed by different organizations, service requests can come from cross security domains, and therefore the propagation of identity across domains is important.

2.4 Contributions of the thesis in relation to the reference architecture

The thesis studied the specific research objectives focusing on two application cases, 1) ice navigation, and 2) participatory water quality monitoring. Figure 3 maps the contributions of Publications **I-V** and the new results presented in the thesis overview (**O**) to the reference architecture. Publication **I** concentrates on the preliminary ideas of solving the orchestration using Business Processing Execution Language (BPEL) and introduces the concept of Facade to coordinate synergies in the data processing. Publication **II** introduces the ice navigation case and shows how the multitude of information sources is presented to users at sea. Publication **III** considers minimising the amount of data transferred and presented by applying route optimisation that integrates ice model output with a ship transit model. Publication **IV** harnesses ships and ship radars as an opportunistic data collection sensor network, while publication **V** harnesses citizens as participatory sensors of water quality. The thesis overview (**O**) extends the orchestration towards the Enterprise Service Bus (ESB) architectural model, and elaborates the Facade model towards the Data operator model to serve multiple applications.

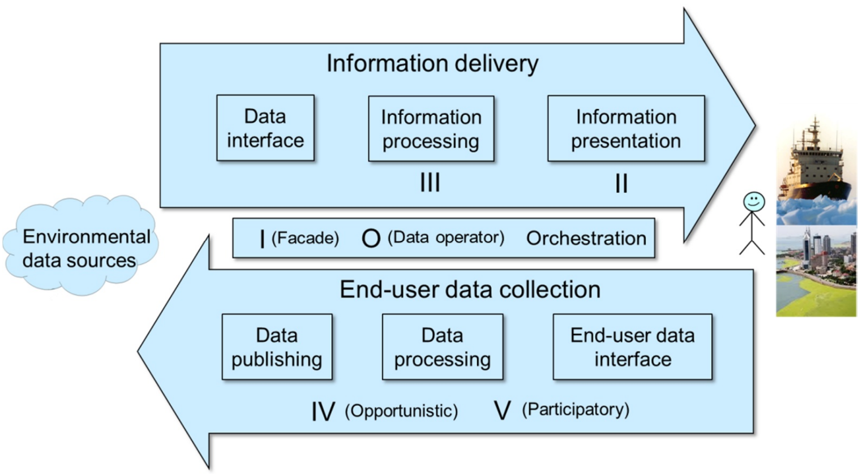

Figure 3. The research concentrated on improvements in the information delivery architectures of two application cases, ice navigation and water quality monitoring. The main focus of contributions of each Publication (**I-V**) and the new results presented in the thesis overview (**O**) are mapped to the reference architecture.

2.4.1 Application cases

2.4.1.1 Ice navigation

Information about prevailing environmental conditions is essential for navigation in ice-covered sea areas. Information is needed for cost-efficient and safe route planning, to find the easiest routes, to avoid icebergs and other obstacles and to reduce the probability of getting stuck or ending up in dangerous areas. The ice field can be static or move, mainly due to the driving forces of wind and sea currents. When a wind stress field becomes convergent the weakest part of the ice cover can be crushed, building up heaps of broken ice blocks above and below the water line (Wadhams, 2002). Such a deformation feature is called a pressure ridge and it can reach a thickness of tens of meters, being thus impenetrable even for the strongest of icebreakers. In addition, the convergent ice field causes compression in the ice that hampers ship transit. Ice navigation knowhow is especially relevant in heavily trafficked sea areas where ice plays a major role in winter time, such as the Gulf of St. Lawrence in Canada and the Baltic Sea in Europe. Moreover, global climate change is opening up new geographically shorter international shipping routes linking the Atlantic and Pacific Oceans by the Northern Sea Route or the Northwest Passage (Smith and Stephenson, 2013).

The data sources that support ice navigation are heterogeneous, including satellite sensors (e.g. Pettersson et al., 2000; Pedersen and Saldo, 2005; Vainio et al., 2000), weather observations and forecast, ice charts (e.g. Karvonen et al., 2003) and ice model data (e.g. Axell, 2005; Mårtensson, 2012). The ships are moving out of reach of the main telecommunication networks and rely on satellite communication links. Maritime VSAT (Very-Small-Aperture Terminal) systems using geosynchronous satellites can reach a data range of up to 16 Mbit/s, but only the Iridium satellite system can reach polar regions with data rates from 2.4 kbit/s to 128 kbit/s (depending on the solution and service level). Therefore, the information delivery architecture should adjust the amount of transferred data to suit the communication capabilities. Thick client architectural models (i.e. client software providing rich functionality independent of the central server) are relevant, as the communication link, whether satellite or terrestrial, can be affected by breaks more often than fixed landline connections.

2.4.1.2 Water quality monitoring

Fresh water is essential for agriculture, industry and human existence, and water quality is a function of land processes that generate pollution and thus is an indicator of overall environmental health. The management of water resources and controlling of societal processes need updated information about water quality of seas, lakes and rivers, and has global application possibilities. The data sources include satellite imagery (Palmer et al., 2015), automatic in-situ measurements, human measurements and laboratory analysis of water samples (Karydis and Kitsiou, 2013). Citizen science has also been recognised as a promising way of collecting additional information (Pyhälahti et al., 2015; Publication **V**).

In the participatory water quality monitoring case, the remote sensing data and expert observations were complemented with water turbidity and algae observations originating from citizens in Finland and in the Baltic Sea area. Algal mass occurrences, in particular cyanobacterial surface blooms, are one of the most distinguishing effects of eutrophication in lakes and the coastal waters of the Baltic Sea (Solimini et al., 2006). Therefore, the frequency and intensity of cyanobacterial blooms are used to assess the ecological status of surface water bodies in Europe under the European Water Framework Directive (Carvalho et al., 2013). However, algal bloom occurrences in water bodies vary greatly in terms of both space and time, which requires frequent monitoring of the algae situation. Remote sensing by satellites can provide high temporal and spatial resolution bloom information of sea areas (e.g. Reinart and Kutser, 2006). However, satellite methods using visual or near infrared channels require clear skies and therefore cannot be used every day, and continuous data series cannot be made. In addition, monitoring of small water bodies requires satellite images with a good spatial resolution (<30 m), and at present, such images are not operationally available on a daily basis. Therefore, on-ground visual observations by trained experts are an important method in algae monitoring, and e.g. in Finland the method has been used since 1998 (Rapala et al. 2012). However, there are insufficient resources to cover all lakes and times, and thus not all bloom events can be detected. In this study we applied citizen science to increase both temporal and spatial coverage of surface bloom visual observations in lakes and coastal regions of the Baltic Sea.

3. Implementing the research objectives

In this section a description is presented of how the specific research objectives were approached. Two application cases, ice navigation and water quality monitoring, were considered in the study. Ice navigation involves professional users on board ships navigating in ice-covered areas, and water quality monitoring involves both professional users and laymen situated near water bodies.

3.1 O1: Orchestration of information processing and delivery

Implementation of the first specific research objective, *O1: To develop orchestration of the information processing to tackle the challenges of near real-time information delivery*, is described here. First, we developed an orchestration architecture using the Service-Oriented Architecture (SOA) model and Business Processing Execution Language (BPEL) and developed the Facade component for implementing synergies in the processing tasks. Then we developed an architecture for event processing based on the Enterprise Service Bus (ESB) model and extended the Facade model towards the Data operator model.

3.1.1 Orchestration with BPEL

Solving the challenge of integrating heterogeneous distributed systems suggested the use of the SOA model, as this is ideal for integrating heterogeneous systems (Erl, 2005). However, the basic SOA model does not provide the processing chain orchestration as such, and the architecture applied BPEL for the task. A BPEL script defines the execution workflow and a workflow engine takes care of the execution. We developed an information processing and delivery architecture based on the Service-Oriented Architecture (SOA) model and Business Process Execution Language (BPEL) (Publication **I**). The work originated from the ice navigation case but was influenced by other applications in environmental monitoring applying remote sensing data in the domains of disaster monitoring, forestry, forest fire monitoring, traffic monitoring, disposal site monitoring, season monitoring for tourism, air quality monitoring and water quality monitoring.

To solve the challenges of extensibility and configurability we implemented the data processing tasks as web services, described their interfaces using Web Service Description Language (WSDL) and scripted the sequence of processing using BPEL. The architecture is presented in Figure 4, and it includes a module library that contains all the necessary processing modules defined as proxy web services and a workflow engine that executes workflow descriptions defined with a workflow editor. Executable workflow descriptions are published as web services and they can be accessed and initiated by other applications, or they can be part of other workflow descriptions, thus allowing a hierarchical composition of workflows. In the architecture, there are two example applications calling the workflow web services, an user interface (UI) that enables manual initiation of workflow execution, and the Facade subsystem that operates according to end-user profiles that define what kind of data end-users require. End-users define and store profiles using a dedicated profile UI. As the workflow engine executes the workflow, different proxy web services can run on the same or separate web server instances that can be deployed on the same or separate machines, as illustrated in the bottom of Figure 4. In the illustration, the example workflow contains calls to processing tasks and a data storage service for storing and retrieving data.

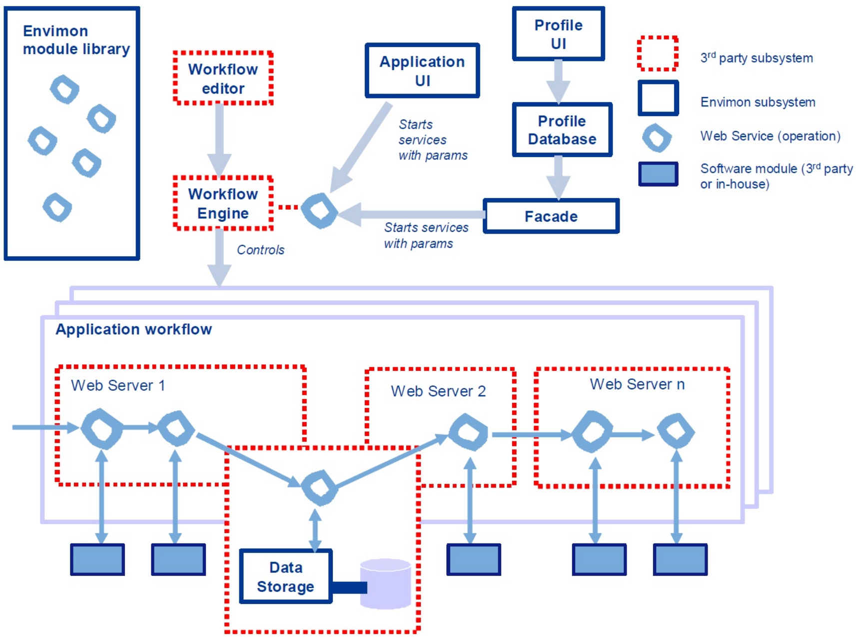

Figure 4. BPEL orchestration architecture including the Facade component (from Publication **I**). Processing tasks are implemented as web services that are orchestrated by the workflow descriptions defined with a workflow editor and executed by a workflow engine. Facade initiates needed workflows according to end-user profiles that define end-user data needs.

3.1.2 Facade component

In order to solve the challenges of timely data delivery and relevance of data delivery, we needed to consider the synergies in the data processing tasks to minimise the amount of processing and processing time, and to include end-user specific needs in the data processing. When considering the whole processing chain from input to delivery it appears that instead of a single workflow there should be a multitude of workflow definitions for a single application. One workflow would be quite complex if it should take care of all the possibilities for tailoring different products for different data users. For instance, there could be a workflow process that takes care of bulk data production, and another that tailors and filters bulk data files according to different user needs. If there is more than one workflow definition taking care of all the processing needed by a single application, a higher level of control is required. This new level of control could in principal be implemented as a top-level workflow description that controls the other lower-level workflows, but depending on the complexity of the control logic it could be programmed using other programming languages than workflow languages, or at least additional control modules could be needed as part of the control workflow.

We implemented a top-level orchestrator component that we called Facade (Kotovirta, 2003, Publication **I**, Publication **II**). Facade retrieves relevant data products and processes and delivers relevant information to end-users based on end-user profiles that define user-specific needs. Figure 4 shows the role of Facade in the BPEL orchestration architecture. It acts as a top-level control for a workflow engine and a multitude of workflows, and calls the workflows to do the actual processing jobs defined by user profiles. Users manage their data service orders using a profile UI, and Facade ensures that the final products are delivered to them. Thus, Facade is an active component that takes responsibility for the information delivery orchestration so that relevant files are available according to the order without delays after the data become available in the data sources. Facade is triggered by events generated by new available data, e.g. when a satellite has flown over or a forecast model is run.

Facade implements the synergies in the processing tasks so that the amount of processing and thus the processing time are minimised. Data retrieval and other general processing tasks are performed first when new data are available, then further processing generates more specific results according to end-user needs (Figure 5). For example, users may be operating in different sea areas and require different parts of the very same satellite data, or users may be using different data communication channels with varying bandwidths and require the same data in different resolutions. A large satellite image could be retrieved, rectified and calibrated only once, but then cropped, filtered and resampled many times according to user-specific needs. When a new user need appears the system should first check what is already available and process only the necessary steps.

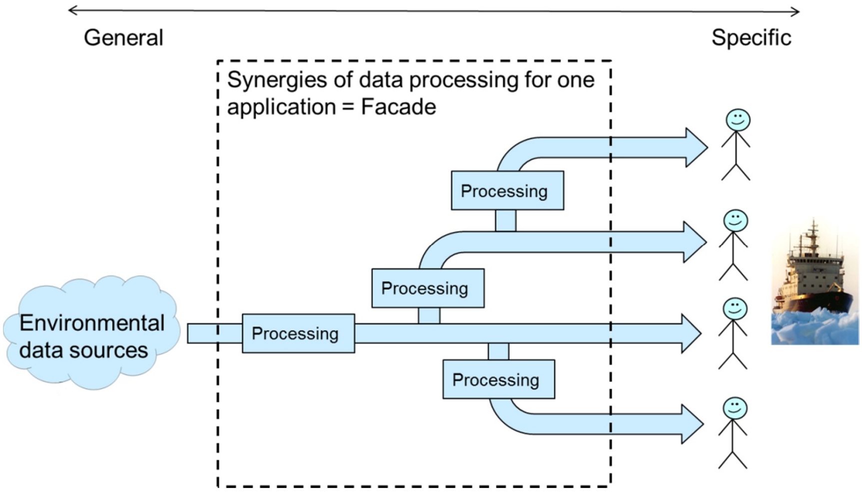

Figure 5. A schematic view of the role of Facade in the environmental information processing. Synergies of processing and delivery to users with different specific needs are implemented only once in a component called Facade. For example in the case of ice navigation, a radar satellite image is retrieved and pre-processed only once and then cropped and delivered according to different end-user needs.

3.1.3 Orchestration with ESB

Whereas BPEL orchestration provides a solution for processing intensive tasks, we were also considering the processing of data flows in which a number of events need to be processed efficiently. An event can be considered as a measurement produced by a sensor, a forecasted value produced by a computational model, or a feature calculated from the data, e.g. a parameter exceeding a threshold value, or a more complex phenomenon detected in the environment. The challenge of delivering relevant information also suggests the use of event processing. Instead of large regular data amounts the user could be provided with short notifications about static conditions and with more comprehensive data amounts when critical changes occur in the environment. Event-based architecture can help in implementing a push type of service that monitors the environment on behalf of the user and notifies the user of any relevant changes in the environment, or even a proactive system delivering relevant information that the user was even not aware of (Rönkkö et al., 2012).

The need for event-based processing directed the research towards an environmental information delivery architecture in which the orchestration is implemented using the event-based SOA on Enterprise Service Bus (ESB) (Kotovirta et al., 2013). The Enterprise Service Bus (ESB) is a component that mediates messages between various distributed systems by transforming and routing the messages (Chappell, 2004). It extends the traditional client-server model and promotes asynchronous message oriented design for communication between systems. We extended the architecture with a complex event processing (CEP) component which can combine events from multiple sources and infer patterns of more complex events. CEP consists of a real-time event correlation mechanism controlled by rules encoded with an event pattern language (Luckham, 2002).

3.1.3.1 Prototype system for near real-time environmental information delivery

We implemented a prototype system based on the event-based SOA and ESB architectural model for near real-time environmental information delivery for serving multiple applications (Figure 6). The first version of the prototype was described by Kotovirta et al. (2013), and here we describe the extended version as a new contribution. Extensions include an additional cloud storage system, a semantic database and a data catalogue service, but we describe also the main parts of the prototype for convenience. Kotovirta led and participated in the development of the architecture design as well as the prototype system implementation, and participated in the development of selected parts of the prototype.

The prototype implements proxy web services for accessing data sources and processing components, and implements the process chain orchestration by using the WSO2 Enterprise Service Bus open source implementation. The prototype was installed on Linux nodes that run on the commercially available Amazon Elastic Compute Cloud (Amazon EC2), and it used noSQL Microsoft Azure and Amazon S3 cloud storages for sensor data storing, PostGIS SQL database for geographical data storing, and Profium semantic storage for storing data in RDF format (Resource Description Framework). The open source data portal software CKAN was used to implement a data catalogue for static data sources and for near real-time data streams that the prototype processed.

The complex event processing was implemented using an open source event stream processing and event correlation engine called Esper, developed by EsperTech. Esper provides a specific language for expressing event conditions, called Event Processing Language (EPL). The anomalies or events that the users are interested in are transformed into EPL sentences and subscribed to Esper. In the prototype, detected events about quality problems or interesting anomalies can be delivered to end-users using a notification service that uses e-mail, HTTP and SMS protocols for sending notification messages. Quality control was implemented using a proprietary quality control software library that includes rudimentary statistical tests and algorithms for data quality purposes (e.g. average, standard deviation, minimum, maximum, histogram, linear regression, fast Fourier transform). The scalability of computations was improved by integrating the open source distributed real-time computation system, called Storm, to distribute the computation of quality control, complex event processing, and generic computation service.

The data processing chain's data security is highly dependent on the level of data security of separated services. However, as the Enterprise Service Bus (ESB) acts as a central point of communication it can improve the overall data security by controlling the traffic between the services. The data security service component WSO2 Identity server was included in the prototype architecture to take care of message confidentiality and integrity, identity and authentication, authorization and privacy, access policies, and federation of identities.

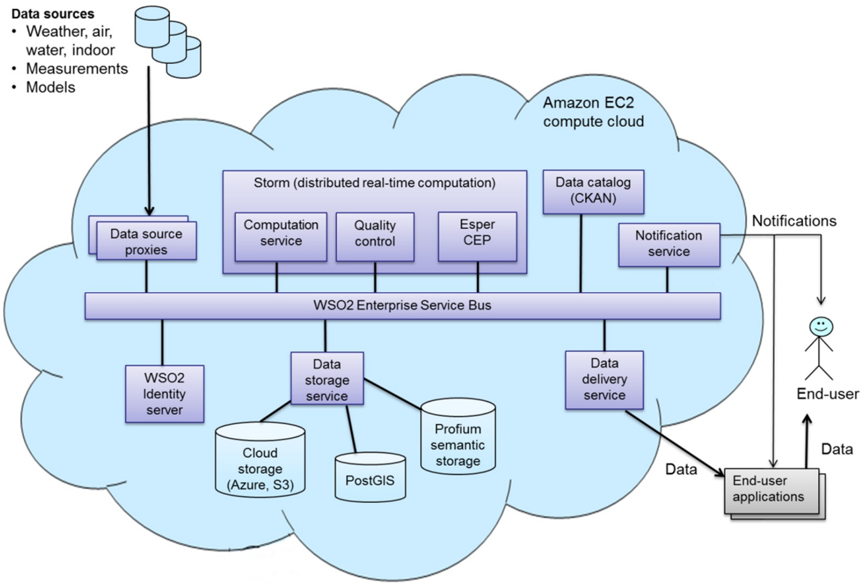

Figure 6. The environmental information delivery system prototype architecture based on Enterprise Service Bus (ESB) architectural model, Esper Complex Event Processing (CEP), distributed computing platform Storm, Amazon EC2 cloud computing platform and Amazon and Windows Azure cloud storages.

3.1.4 Data operator

In this section we present the concept of Data operator that extends the Facade model towards multiple applications. The discussion and definition of Data Operator is a new contribution in the thesis overview. The data delivery architecture prototype by Kotovirta et al. (2013) can be considered as a Data Operator prototype, as it was designed to serve multiple environmental monitoring applications accessing similar data sources and requiring similar processing. Various data sources were connected, including a weather station provided by Vaisala Ltd., different sensor networks measuring indoor air quality, temperature and building energy consumption, a pollen forecast model provided by the Finnish Meteorological Institute (FMI), a weather forecast model provided by FMI, the water level forecast model provided by the Finnish Environment Institute (SYKE), and participatory sensing observations provided by VTT Technical Research Centre of Finland's EnviObserver platform (Kotovirta et al. 2012). The data, processed information and notifications were delivered to various applications, including an application assisting power network management in fault detection, an application delivering environmental data to citizens, an agricultural application helping farmers in decision making, e.g. related to pesticide spraying, a facility management system presenting indoor comfort indices to facility managers and a storm path prediction application alerting citizens about being in the path of a storm system (Stocker et al., 2015).

While serving a multitude of applications, we realised that similar processing tasks were needed repeatedly in the information processing. The applications needed to access the same data sources, control the quality of those data, store historical data for later use, and generate notifications based on the same data but with differing notification rules. As in the case of Facade there was a need to implement synergies in the data processing only once (the challenge of cost-efficiency) and to actively process general data on behalf of end-user applications based on specific end-user needs and deliver the information to end-users in good time (thus meeting the challenges of timely delivery and relevance of data delivery). However, in contrast to the case of Facade, this time the system was serving a multitude of applications instead of multiple end-users of a single application. The whole prototype system was taking the role of Facade for multiple applications, and it was named Data operator. Kotovirta composed the first ideas of Data operator and led the discussions for defining the role of Data operator in environmental monitoring in the Measurement, Monitoring and Environmental Efficiency Assessment (MMEA) research program. In this thesis the term Data operator model is used to describe the design principle of implementing the synergies of data processing as a component called Data operator. The prototype implementation by Kotovirta et al. (2013) is referred to as the Data operator prototype. Figure 7 illustrates the role of Data operator in the information delivery architecture serving multiple applications.

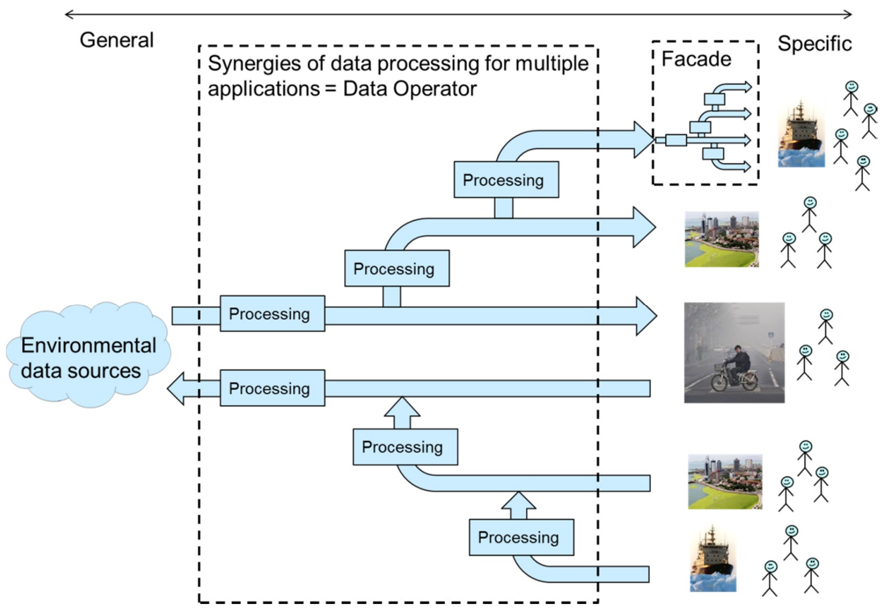

Figure 7. A schematic view of the role of Data operator in environmental information processing and delivery. Synergies in information delivery and data collection between multiple applications are implemented as a separate component called Data operator. For example, access to weather data sources could be implemented only once, instead of each application (ice navigation, algae monitoring, air quality monitoring) doing that separately. In addition, user observations about weather conditions from each application could be available to all applications via Data operator. The role of Facade in serving a single application is also depicted in the figure to help compare the roles of Facade and Data operator.

3.2 O2: Information presentation

One important component of the information delivery chain is the information presentation to the user. All the time-varying observational and forecast information should be presented in an understandable and intuitive way to support the decision making. In this section we use the ice navigation case as an example, and describe the accomplishment of the second specific research objective, *O2: To develop the presentation of relevant information while minimising the complexity of multidimensional information in order to support the interpretation of the situation*.

The ice navigation case needs advanced solutions for information presentation, as a multitude of data sources are available for seafarers in ice-covered waters. For example, near real-time delivery of both optical and radar satellite images to ice-going ships has been made for several years in different ice-covered sea areas such as the Arctic (Pettersson et al., 2000), the Antarctic (Danduran et al., 1997; Pedersen and Saldo, 2005) and the Baltic Sea (Vainio et al., 2000). In addition to satellite images, other data such as weather observations and forecasts, water level observations and forecasts, ice forecasts including level ice thickness and ice drift produced by computational models, ice thickness charts produced by automatic algorithms using radar satellite images (Karvonen et al., 2003) and ice charts made by human experts are also delivered to users on board (Publication **II**).

We developed an end-user application called Icebreaker Plotter (IBPlott) that presents ice navigation-related information to the end-user, as well as the traffic information coming from the Icebreaker Network system (IBNet) used to coordinate and communicate the daily assistance activities of Finnish, Swedish and Estonian icebreakers. IBPlott allows the user to choose which data layers are visible on the screen and enables panning and zooming as well as setting the time of the view (Figure 8). The symbols representing environmental data, such as meteorological, oceanographic and ice parameters, are filtered according to zoom level in order to avoid too many symbols on the screen, which might be confusing to the user. IBPlott presents the latest optical and radar satellite images that show observation of the ice field from a larger area. A method was developed to compare satellite images from different times or different instruments. One satellite image can be drawn only in small window at the mouse location, and when the mouse is moved, the user experiences a feeling of seeing through the topmost image, and can compare features in both images.

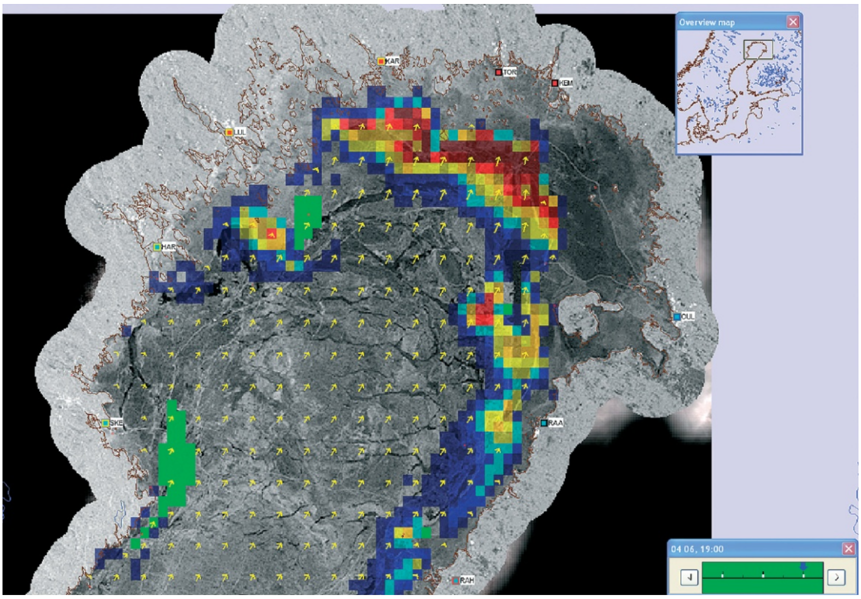

Figure 8. Example visualisation of ice information in the end-user application IBPlott. In the image, ice drift and compression forecasts calculated by an ice model are presented on top of a Radarsat image. The red colour indicates that ice is accumulating in the northern parts of the Bay of Bothnia (approx. 65N 23E) and causing compression in the area. Green colour indicates divergence of the ice field near the coast of Sweden. The user can view the situation at different times using the time browser tool in the lower right corner. The overview map in the upper right corner indicates the geographic location. (From Publication **II**)

3.2.1 Route optimisation for reducing complexity

We studied how the complexity of the presentation of multidimensional information could be reduced by an automatic analysis of the situation. If an algorithm could calculate the most relevant features in the environment that affect the decision making it would be sufficient to present only these features to the user. In the case of ice navigation, the information is mainly used for planning optimal routes through the ice field, and thus we aimed at calculating the best route choices automatically in order to minimise the amount of data presented and maximise the amount of information to support rapid comprehension of the situation. By minimising the data presented we can also minimise the data delivered over a low-bandwidth data communication channel to the ships, if the analysis is done on the server side ashore and not on the client side on board.

We developed a route optimization method for ships navigating in ice covered areas that combines ice modelling, ship transit modelling and optimization methods and aims at an efficient implementation for rapid results in an operative environment (Publication **III**). A computational ice model (Wilhelmsson, 2002; Axell, 2005) estimates the ice field properties, such as equivalent ridged ice thickness, level ice thickness and ice concentration (Figure 9). The equivalent ridged ice thickness means the average thickness of ice that is equivalent to the effect of the density and size distribution of ice ridges (heaps of ice blocks) on the ship performance. Level ice thickness means the thickness of the undeformed ice and it can reach 0.6-0.8 m during normal winters in the Baltic Sea, where multiyear ice does not form, and up to 3 m in the polar regions where multiyear ice exists (Wadhams, 2002). Ice concentration describes the area of the water surface covered by ice as a percentage of the whole area of interest. We assume that at <70% ice concentration a ship can avoid all ice and achieve its open water speed, at >95% ice concentration the ice field is solid and a ship reaches its ice speed (i.e. the maximum speed it can reach due to ice), and between 70% and 95% ice concentration the ship speed is a linear combination of the open water speed and the ice speed.

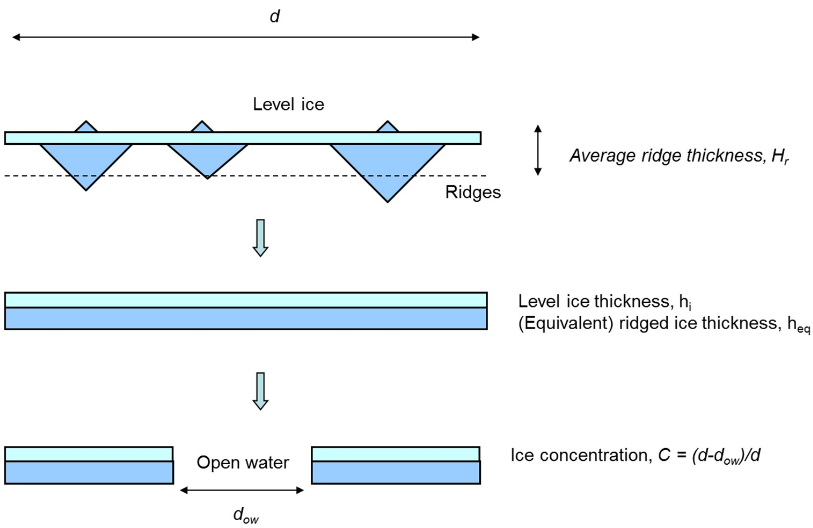

Figure 9. Illustration of the ice field properties used in the route optimisation. Level ice thickness is the thickness of unbroken solid ice, equivalent ridged ice thickness estimates the effect of ice ridges of various sizes in the ice field, and ice concentration gives the proportion of ice in relation to open water. (Illustration from Publication **III**)

The ship transit model estimates the ship speed in given ice conditions. We calculated the ice resistance caused by level ice and ridged ice separately using formulae given by Riska et al. (1997) and the effect of ice concentration using formulae given in Publication **III**. The net thrust produced by the propellers of a vessel were calculated using formulae by Kämäräinen (1986). The ship speed in ice is determined by the speed that makes ice resistance and net thrust equal. The ship speed is used in the cost function to calculate the travel time of a given route, and the optimisation process alters an initial route guess to find alternatives with lower cost, i.e. shorter travel time. Three optimization methods were tested: 1) Powell's method, which is a standard method in non-derivative optimization (Powell, 1964), 2) the simplex method (also known as the Polytope algorithm) by Nelder and Mead (1965), which is a reliable and simple optimization method, and 3) simulated annealing (Kirkpatrick et al., 1983), which is suitable for difficult optimization problems. Although simulated annealing is more capable of finding its way out of local minima, Powell's method provided good enough results faster and was therefore selected to be used in the prototype implementation.

One purpose of applying ship route optimization was to present multidimensional data to end-users in an easily understandable way. The optimization is started by giving an initial guess and as the optimization continues an animation about the progress is presented. The optimization ends as the ending criteria are reached or the user stops the optimization. The user can stop the process at any time, and the best result so far is presented. It is also possible to give more than just one initial guess and all the optimized result routes are drawn on the screen. This enables easier comparison of different initial route alternatives and their optimized results. We developed a simple method to study the variability of the optimal route. The method takes random samples of the routes that are within prescribed bounds of the optimum, and draws these routes together with the optimum (Figure 10). Without any further analysis this presentation gives visual information about the sensitivity of the optimal route. If all the result routes close to the optimum go close to some point, there is a narrow corridor presented in the visual presentation. This suggests that one should not divert too much from the route near that point. On the other hand, if there are route segments going relatively far from the optimum route, there is probably a wider corridor where the user can more freely choose his route according to other preferences.

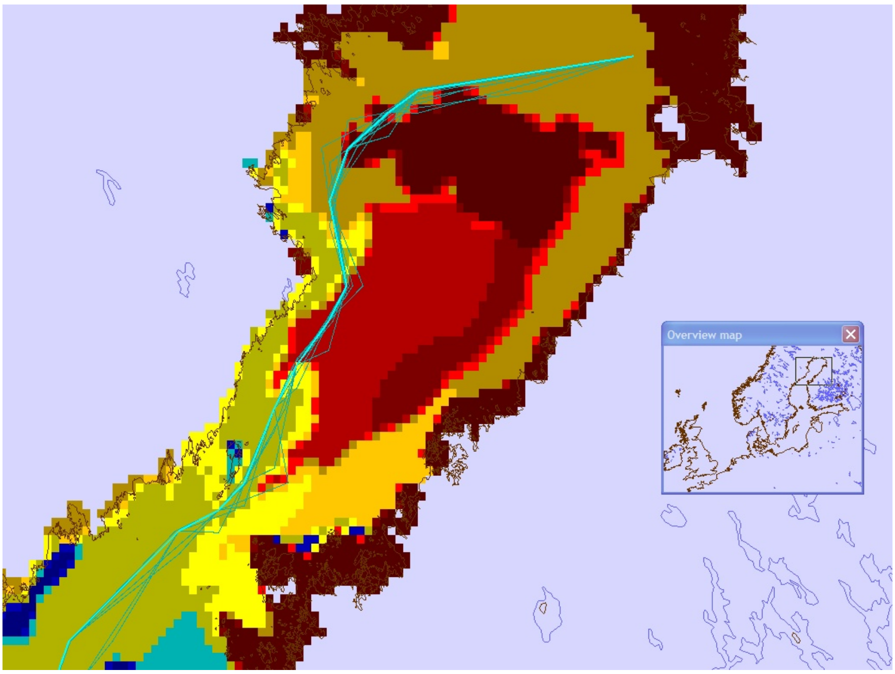

Figure 10. Example presentation of the optimized route through the ice field (from Publication **III**). One model parameter, the level ice thickness, is visualised on the map with colours indicating the ice thickness (blue indicates open water, yellowish colours ice and reddish colours thicker ice. A random sample of routes within bounds of 5% longer transit times are drawn together with the optimal route, which provides visual information about the sensitivity of the optimum.

3.3 O3: Harnessing end-users for data collection

In this section we describe the accomplishment of the third specific research objective, *O3: To harness end-users and end-user devices as data collectors in order to complement other data sources*. First, we take a look at citizen science and participatory sensing, and then discuss the contributions of opportunistic and participatory approaches.

3.3.1 Citizen science and participatory sensing

Citizen science and participatory sensing provide an emerging data source for environmental monitoring and also support near real-time applications. Citizen science involves citizens in producing scientifically meaningful observations or analyses (Haklay, 2012). Citizen science has been pioneered by the SETI@Home initiative (the Search for Extraterrestrial Intelligence), which utilized the idle time of millions of participants' computers in analysing distant radio signals in the search for extra-terrestrial life. Later Rosetta@home harnessed citizen computers to determine the 3-dimensional shapes of proteins in the search for cures for some major human diseases. Rosetta@home realised that people could also provide brain power for the project instead of just providing their computers, and developed a computer game FoldIt that puts players to compete in protein folding, with good results (e.g. Eiben et al., 2012). In other examples citizens have been harnessed to classify galaxies from deep space images (GalaxyZoo project), assign land classes using remote sensing images (Fritz et al., 2012; Hu and Wu, 2012), detect tumour cells in pathological breast cancer images (Candido dos Reis et al., 2015), and analyse remote sensing imagery to provide assessments of damage caused by earthquakes and other disasters (Barrington et al., 2011).

As a sub-field of citizen science, citizens have been harnessed for collecting information about the environment. Citizens have been involved in environmental monitoring, to some extent, for over a hundred years - for example the Christmas Bird Watch, started by ornithologists of North America, has been ongoing since 1900 (Haklay, 2012). Recent advances in information and communication technology (ICT) and increased awareness of the status of the environment, in particular global climate change, have activated people even more to participate in environmental monitoring and enabled efficient ways to collect, store and share the data (Burke et al., 2006; Conrad and Hilchey, 2011).

There are many terms that describe citizen contribution to environmental monitoring and each term has its own subtle nuance and viewpoint. For example, Volunteered Geographic Information (VGI) means that individuals using GPS devices collect geographic information for mapping the environment (Goodchild, 2007) and OpenStreetMap is a successful example of a large scale VGI activity. However, two main paradigms can be detected for citizen contributions. Users can provide observations actively themselves, which is called participatory sensing, or enable their mobile devices to collect data automatically, which is called opportunistic sensing (Lane et al., 2008).

Methods for opportunistic and participatory citizen sensing have been presented for various applications and in various forms. Eiman (2010) presented a method for mapping noise levels in a city area using mobile phones; Mednis et al. (2011) developed a system for road irregularity detection using data from accelerometers of mobile phones; Ginsberg et al. (2009) studied how search engine query data could be used in the recognition of influenza epidemics; Wang et al. (2013) studied how images produced by people using social media tools can be used to observe the state of natural world; Olmanson et al. (2008) used citizen Secchi depth measurements as an in-situ data source for satellite image calibration; Sunyoung et al. (2011) developed a mobile application (Creek Watch) for citizens to monitor waterways (amount of water, rate of flow and amount of litter); Lowry and Fienen (2013) presented a method for stream stage monitoring in which citizen passers-by make observations using fixed measuring devices; Toivanen et al. (2013) presented a method for observing Secchi depth and turbidity using an inexpensive measurement device and a mobile phone camera; Leeuw and Boss (2014) developed the HydroColor mobile application which estimates the concentration of total suspended matter and the backscattering coefficient from water reflectance measured by a mobile phone camera; Cao and Thompson (2014) used smartphones as sun photometers for the remote sensing of atmospheric optical depth; Molinier et al. (2014) presented a method for observation of tree parameters using a mobile phone and remote sensing data.

Architectures and platforms for more generic participatory sensing have been suggested, for example A Scalable Architecture for Global Sensing and Monitoring G-Sense (Perez et al., 2010), Platform for Remote Sensing using Smartphones (Das et al., 2010), the personal environmental impact report (Mun et al., 2009) and the EnviObserver participatory sensing system (Kotovirta et al., 2012).

3.3.2 Opportunistic data collection: ships as ice sensors

For the ice navigation case we developed a method to harness the ship radars and ships themselves as a sensor network to collect data from the field that could help other users moving in or approaching the area (Publication **IV**). The system collects marine radar images and ship performance observations, forms mosaics of images coming from multiple radars, calculates ice drift from successive radar images, analyses trafficability in different sea areas, and delivers processed images and ice drift information to end-users (Figure 11).

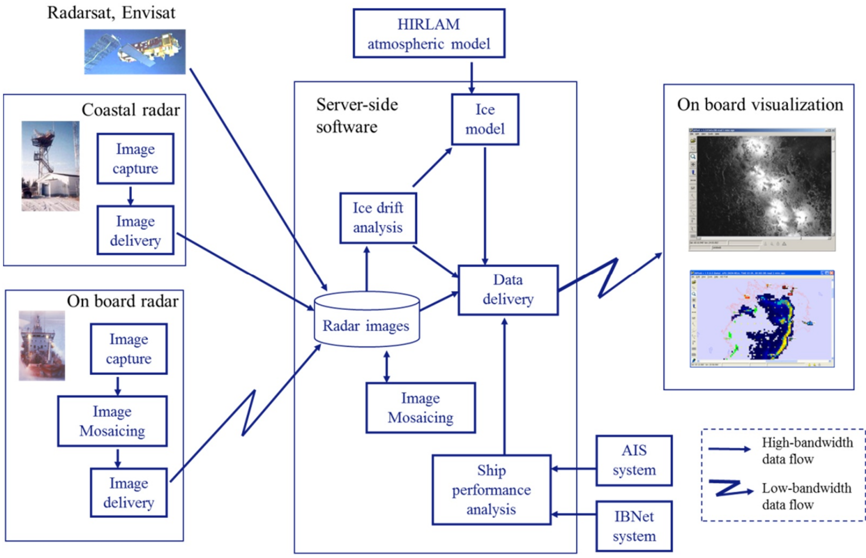

Figure 11. The conceptual architecture of ships as a sensor network system. Image mosaics are formed from radar images collected from coastal and on-board radars, ice drift is analysed from successive radar images, and trafficability of the ice field is estimated on the basis of ship speed and ship specification data coming from the AIS (Automatic Identification System) and IBNet (Icebreaker network) systems.

Trafficability estimation using ships as sensors is based on the finding that theoretical speeds of different types of commercial ships correlate as a function of the level ice thickness (Riska et al., 1997). We assume that the maximum speed a ship can achieve in the ice can be used to derive a one-dimensional indicator of the ice field properties that we call the ice conditions equivalent value (ICE). All the phenomena such as level ice, ridges and ice compression are taken into account as thicker ice. The ICE value calculated by one ship can be used to determine the speed of another ship in the same area. This can be done if we have the ICE-v curves of both ships, i.e. functions that determine the relationships of ship speeds (v) and the ICE value. ICE-v curves can be determined by ship transit modelling or by statistical methods. We determined ICE–v curves statistically using the AIS (Automatic Identification System) data, IBNet (Icebreaker Network) system data and ice chart data. A dataset was collected that includes the ship speed, time, location from the AIS system, level ice thickness and ice concentration (in the ship's location in space and time) from ice charts, and information about the closest port and icebreaker activities from the IBNet system. Locations in which the ships would probably not use full engine power were removed by selecting only the data points in which the ships were not assisted by an icebreaker and were not close to any port. The statistical analysis is presented in Publication **IV**.

Satellite radar images cover large areas and are useful in navigation. However, when the ice field is moving it would be useful to get an updated view of the situation. Ships carry radars that are used in tactical navigation to observe the ice conditions near the ships. Coastal and ship-borne radar images can be used to complement satellite images in limited areas, and image mosaics formed from shipborne radar images from many ships can cover a larger area. We prototyped a system that collects radar images from ships, forms image mosaics of single images from different locations and delivers mosaics to users. In the prototype we had only one image capture device available, but demonstrated the image data collection, image mosaicing and data delivery architecture. Figure 12 presents an example image mosaic formed from images taken on board the icebreaker Otso in the Gulf of Bothnia. Multiple radar images with specific time intervals are combined in the same image from two parallel tracks as Otso first moved south and later travelled north again. Thus, the image mosaic shows a larger area than a single radar image would. Different shades of grey illustrate the strength of the backscattered radar signal and reveal features in ice.

In addition to forming image mosaics, we also considered determining ice drift from sequential images. Ice compression caused by a moving ice field is a major factor in estimating the trafficability. Different methods have been used to observe ice drift, such as buoys (Heil et al., 2000, Inoue et al., 2009, Rampal et al., 2008), marine coastal radars (Sun et al., 2004, Mahoney et al., 2007), camera images (Leisti et al., 2009), and successive satellite images (Karvonen et al., 2007, Gutierrez and Long, 2003). We used radar images collected from ships and coastal radars and applied two approaches: visual estimation and numerical computation. To estimate the ice drift visually, the user can view an animation of successive radar images from a stationary ship or coastal radar, or compare marine radar image mosaics with each other and with satellite radar images. We developed an experimental colour-coded image mosaic by rendering different individual radar images from different times to different RGB (Red, Green, Blue) channels. As a result, static parts in the image are rendered in black and white, whereas moving parts appear in colours. As the time difference between the images and the projection of the image are known, ice drift can be estimated by calculating the distance between red, green and blue versions of an ice feature in the same image. We also calculated the ice drift numerically from sequential radar images by using the method developed for satellite radar images (Karvonen et al., 2007). The details of the numerical method are given in Publication **IV**.

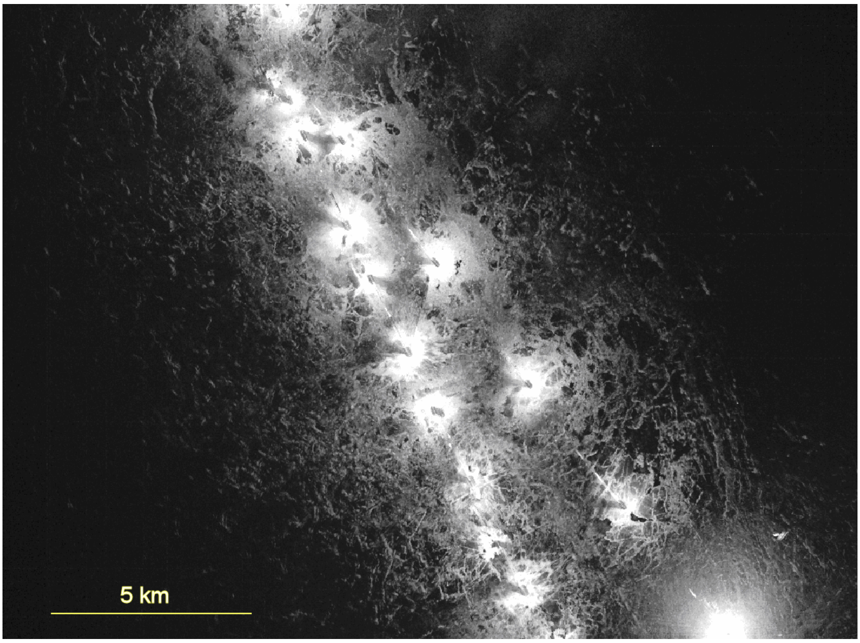

Figure 12. An example image mosaic formed from radar images captured on board the icebreaker Otso in the Gulf of Bothnia. Two tracks are merged in the image as Otso first moved south and later travelled north again using an almost parallel route. The bright spot in the lower right-hand corner is an (overexposed) artefact created by the mosaicing algorithm when Otso was stationary for some hours. (From Publication **IV**)

3.3.3 Participatory data collection: citizens as water quality sensors

To study the use of citizen end-users as participatory data collectors we developed a participatory sensing platform called EnviObserver (Kotovirta et al., 2012), and applied the platform to air quality monitoring, plant disease monitoring, storm monitoring and also to water transparency, water turbidity (Toivanen et al., 2013) and algae monitoring (Publication **V**). The water turbidity monitoring utilised a simple measurement device which consists of a container and two measurement tags at different depths inside the container. The container is filled with water and the user takes a picture looking inside the container through a hole in the lid. The water transparency and water turbidity can be determined by a two-phase algorithm that first searches for the measurement tags in the image, and then determines water quality values based on RGB values of the tag pixels. The test users were recruited by the Finnish Environment Institute and included people who collect water quality samples professionally, people from water monitoring companies and water protection associations, as well as private citizens.

The algae monitoring was based on visual estimation of the amount of algae in the water – the very same method that is used by the trained expert observers making national algae observations in Finland. Two applications were developed for receiving both mobile observations from ad-hoc locations and observations from stationary observation sites defined by citizens: the mobile phone application Levävahti (Algae Watch) and the web-based lake information system Järviwiki (Lake wiki, Rapala et al., 2012). Ordinary citizen observers were recruited by publishing invitations in local and national news services and water-related events and exhibitions at the beginning of the summer season in Finland. The architecture of the data collection system is depicted in Figure 13.

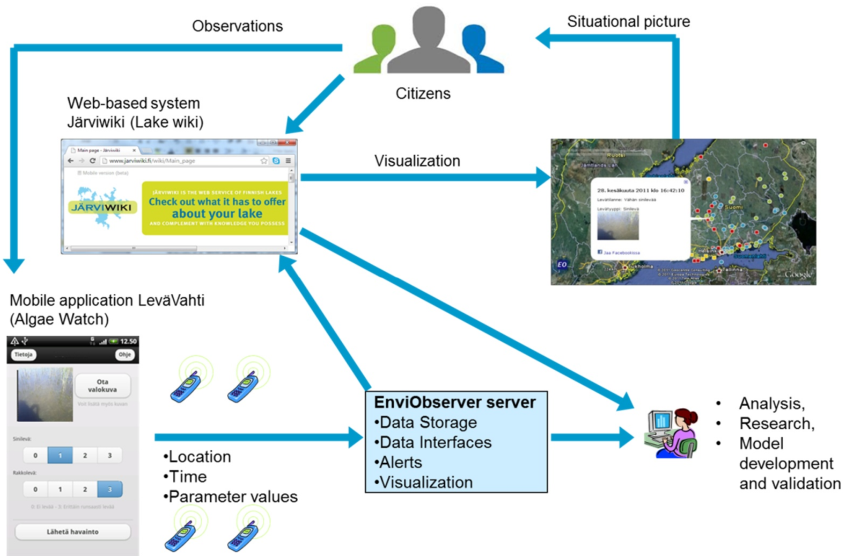

Figure 13. Overall architecture of the participatory algae monitoring system. Two participatory sensing systems were used to collect citizen algae observations: the mobile phone application Levävahti (Algae Watch) and the collaborative web service Järviwiki (Lake wiki). The LeväVahti observations are stored in the EnviObserver server, Lake wiki observations are stored in the Lake wiki system, and all observations are visualised in a web presentation in real-time. The observations are used together with expert observations and satellite-based algae products to estimate the overall algae situation. (From Publication **V**)

Concerning algae observations, a total of *4572* trained expert observations and *872* citizen observations, of which *269* were made using the mobile phone application, were received in the summer of 2011. In the summer of 2012, *4427* expert observations and *319* citizen observations (*156* mobiles) were received, and in the summer of 2013, *4150* expert observations and *465* citizen observations (*134* mobiles) were performed. We compared the citizen observations with the expert observations in order to determine the reliability of the citizen observations. We couldn't directly compare the citizen and expert observations, because they were not made in the same location or at the same time. The mobile phone citizen observations were made wherever citizens were moving and whenever they had time and will to make observations, whereas expert observations were done regularly in pre-determined locations. We calculated weekly averages of both citizen and expert observations and calculated the correlations between the averages. In order to estimate the accuracy of the correlations we used the bootstrapping method, in which the observed data are used as an approximating distribution of the real distribution. We resampled random samples from each week's dataset and constructed a sampled observation dataset for each week that was equal in size to the original dataset. We then calculated new sampled averages for each week, and used them in calculating the correlations between citizen and expert observations. As a result, we obtained distributions for weekly averages and also for the correlations of the weekly averages, and from the distributions we calculated the confidence intervals both for the weekly averages and the correlations of averages. The correlations between the weekly averages of citizen and expert observations were: *0.72, 95%* confidence interval (CI) [*0.53 0.86*]; *0.65, 95%* CI [*0.35 0.86*]; and *0.56, 95%* CI [*0.29 0.76*] for the years 2011, 2012 and 2013. Figure 14 shows example results from the year 2011, other results are given in Publication **V**.

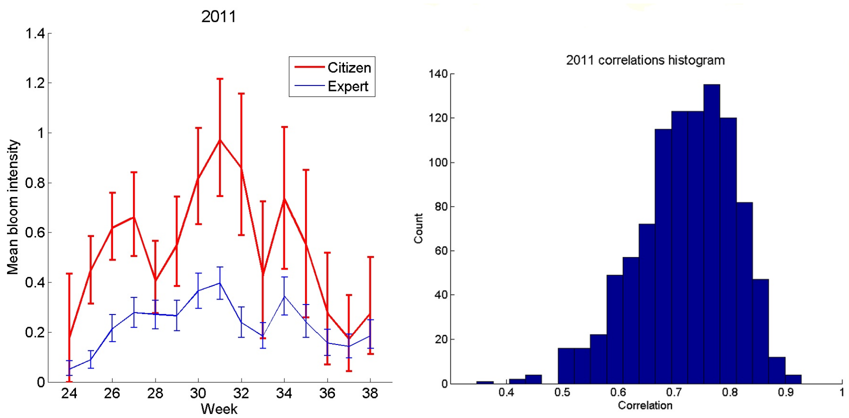

Figure 14. Weekly bloom intensity averages of citizen and expert observations and 95% confidence intervals in the summer of 2011 (on the left), and the histogram of the correlations of the expert and citizen observations produced using the bootstrapping method (on the right).

4. Discussion

In this section we discuss the contribution of the thesis in relation to the research questions. We organise the discussion around the specific research questions, namely *Q1: How could the orchestration of information processing be implemented to address the challenges of near real-time environmental information delivery?, Q2: How could the information be presented to support situation awareness and improve decision making?, and Q3: How could end-users be harnessed to collect additional data from the local environment to complement other data sources in the information processing?*

4.1 Q1: Orchestration

4.1.1 Addressing the challenges

The challenges of near real-time environmental information delivery guided our work of implementing the orchestration of information processing, delivery and data collection from end-users. The challenge of integrating heterogeneous distributed systems suggested the use of the Service-Oriented Architecture (SOA) paradigm, in which the data sources and processing components are loosely coupled and access each other via well-defined interfaces. Workflow control is required to orchestrate the information processing, and for this purpose we developed architectures based on Business Process Execution Language (BPEL) and Enterprise Service Bus (ESB) technologies.

Loose coupling helps in solving the challenge of extensibility and configurability, as new sources or processing units can be added when requirements are extended. It can also improve operational reliability, as components can be substituted with other components providing similar functionality. For example, if some input data source fails to deliver data, the source could be substituted with another source providing similar data. On the other hand, if a processing module ceases to function, the system could use an alternative module to finalize the calculations in time. The central workflow system controls and tracks the message delivery between the services and can detect anomalies in the processing and react accordingly. For example, if some processing component fails to complete its task in time, the central control can trigger other service calls that try to solve the problem, use substitutes, alert human operators and notify data end-users about delays or quality problems. However, loose coupling requires well-defined interfaces of the service components and proper testing so that components behave according to the specification, work as a part of a workflow and can be substituted with components providing similar functions.

The challenges of operational reliability and scalability suggested the use of cloud services for computing and data storage. Cloud computing is foreseen to provide the 5th utility, after water, electricity, gas and telephony, for meeting the everyday needs of the general community (Buyya et al., 2009). Lee et al. (2010) demonstrated that cloud computing resources are sufficiently elastic to deal with the unpredictable loads of real-world sensing applications. Cloud services also help in tackling the challenge of cost-efficiency. The required data processing, data storage, networking and uptime availability can be gained without the need to invest in and maintain the computing hardware and required expertise. Resources can be adjusted to meet the changing needs for data input, storage and processing, and thus keep the costs at an optimal level. The data can be automatically replicated to different server farms in different geographical locations, depending on the cloud service provider, which improves the reliability of data storing.

The challenge of relevance of data delivery guided us towards event-based architecture, anomaly detection and delivering data only when relevant changes are detected. The Data operator prototype aims at observing the environment and data sources on behalf of the end-user applications and notifying relevant changes in the data or in the environment. In a static situation it might be that no data are delivered at all, but the data flows are triggered when relevant phenomena are detected. For example, a decline in the forecasted water quality, or high winds caused by an approaching storm centre, are examples of triggering events. A failure in a data source or missing data values in a data stream are examples of quality anomalies that the user or system operator might be interested in. In the prototype we used only simple algorithms for anomaly and quality problem detection, such as detecting threshold values of parameters, but in future work more advanced algorithms might be used. For example, algorithms for near real-time anomaly detection of sensor stream data have been developed by Yao et al. (2010) and by Hill and Minsker (2010). However, there are limits to what an algorithm can detect and in many cases only human users can interpret the situation. There is also a question of responsibilities when decisions are made. Was a bad decision due to data quality problems, information delivery problems, a programming error in the algorithm interpreting the situation, a logical error in the algorithm, or human error in interpreting the situation?

4.1.2 Data operator

The challenges of cost-efficiency and timely delivery led the development towards a Facade component that firstly wraps the synergies in the data processing tasks serving multiple end-user needs, and secondly is an active top level orchestrator that fulfils end-user information needs in a suitable time frame. Data operator was defined as an extension of Facade towards serving multiple applications in wrapping synergies in processing. When a single application is built by a sole developer (or cooperating developers), it is possible to design the architecture so that the processing is performed efficiently, minimising redundant tasks. However, when multiple applications are developed by multiple developers the coordination is more challenging and even though all developers would benefit from a common implementation it would be more difficult to achieve. We succeeded in developing a Data operator prototype as we were involved in a larger research program that developed multiple applications (Measurement, Monitoring and Environmental Efficiency Assessment, MMEA), and a common solution was possible because application developers were communicating about mutual problems.

4.1.2.1 Fragmentation problem

The Data operator model is relevant when applications are using the fragmented environmental data sources. Instead of implementing the data access many times by different applications, the applications should coordinate their efforts and implement the access only once. Despite open standards and open data trends, fragmentation of data sources is the reality which applications face. This has been noticed by many authors. For example, Goward (2007) wrote: “With all this wealth of land observations it is increasingly difficult to understand why these remotely sensed data are not a more pervasive element of data systems that help inform human societies about the state and dynamics of our home planet.” After nearly ten years the situation has improved, but still interoperability of environmental data is a big issue. Giuliani et al. (2011) stated that although administrations and governments are recognizing that geospatial data are an important component of an information infrastructure (such as e-government), a huge amount of geospatial data is stored in different places by different organizations, and the vast majority of these data are not being used as effectively as they should be. Ceccato et al. (2014) wrote that many barriers remain in terms of data, services, practice and policy that need to be overcome if climate and environmental information are going to play a significant part in reducing climate-related risks. They listed barriers including lack of access to relevant local and globally accessible data and lack of policies for data sharing as well as technological constraints to knowledge and data sharing. Yue et al. (2015) stated that with countless different sensors generating heterogeneous data varying in quality, precision, type, format, scale, granularity, structure and semantics, data integration, sharing, reuse, and project collaboration across disciplines have become a major bottleneck limiting the use of the data.

There are more and more environmental data available produced by various in-situ sensors, remote sensors, computational models making predictions, human specialists, and citizens using their mobile devices. The data are produced by private and public stakeholders globally. We can estimate that the amount of environmental data is growing at the same exponential rate as the amount of data overall. However, at the same time there is a growing amount of data that are not used to their full potential. The problem is that the data are distributed throughout various systems in various organisations in various countries. Instead of being interoperable the data are fragmented inside these various systems and cultures, and cannot be used seamlessly in applications. We can distinguish between different types of fragmentation:

- Fragmentation in space and time. Different sensors and models have different spatial and temporal resolutions, and the data cannot be combined without data synchronization, such as interpolation, extrapolation and advanced data fusion algorithms.
- Fragmentation in technology. Different sensor systems and computational models are based on different technologies and on different principles, and are using different data models and formats to store and exchange the data.
- Fragmentation in semantics. Different organisations, different countries, and different disciplines have different traditions and legacies and therefore different semantics to describe the very same data parameters and measurements.
- Fragmentation in data quality. Some systems consider data quality more seriously than others. For example, official weather observations might be carefully calibrated and the measurement methods and locations are carefully designed, whereas a participatory weather sensing system might use layman observers and a variety of devices with varying data quality and without any guarantees of calibration.
- Fragmentation in data politics. Different organisations have differing views on how the data should be shared, or should it be shared at all. For example, companies may keep the data only for their own purposes and business reasons.

4.1.2.2 Data operator to mitigate fragmentation

As described in the introduction, there are ongoing activities promoting development of standards for sharing data (e.g. OGC) and opening at least publicly produced environmental data (e.g. INSPIRE). There are also efforts to develop new technologies for increasing the interoperability of systems producing environmental data. Bröring et al. (2011) considered the integration of various sensor systems on-the-fly with minimal human intervention. Nativi et al. (2013b) presented the GEO model initiative, which is a generic concept for increasing access to environmental models and their outputs and to facilitate greater interaction between models, resulting in webs of interacting models. Nativi et al. (2013a) discussed a brokering approach for promoting multi-disciplinary interoperability, i.e. interoperability among diverse disciplinary and domain systems such as those of meteorology, biodiversity, climate, seismology, etc. Yue et al. (2015) envisioned the emergence of intelligent GIServices that facilitate information discovery and integration over the network and automate the assembly of GIServices to provide value-added products helping users to perceive environmental surroundings.

Although these efforts provide architectures, technology and standards for interoperability, they do not discuss how systems are encouraged to adopt these technologies and standards. In many cases the need for interoperability is not taken into account when systems are designed and built, because there is no need for interoperability at the time. If a system is designed to serve purposes of one company only there is no need to design it to be interoperable, even though the data would be valuable to other systems and applications later. Therefore, standards for interoperability, open data trend and goodwill are not sufficient to solve the fragmentation problem and get environmental data into use.

The Data operator model was developed because the fragmentation problem exists – we wanted to solve the access to fragmented data sources only once for all the applications. One interesting question for future research is to find out how Data operator helps mitigate the fragmentation problem. Fragmentation increases the costs of accessing data originating from various sources. As Data operator considers the synergies in data and processing needs of various applications, it thus reduces the costs of implementing and running the applications. This benefits the applications that can as a result use more data, which again benefits data providers who can sell more data to applications. Data operators can be considered as two-sided platforms which benefit both data providers and data users. Two-sided platforms provide infrastructure and rules that facilitate the two groups' transactions (Eisenmann et al., 2006). The two groups are attracted to each other by a phenomenon that economists call the network effect. This opens up a business potential for companies implementing the Data operator model.

In the future, there could be Data operator companies that make profit by mediating data processing between data providers and data users. They will serve applications with environmental information and provide data owners with a channel to sell their data to new applications, and thus they create a motivation to get the data into use from fragmented systems. However, at the initial phase of developing a two-sided platform a common problem called the 'chicken-and-egg' situation may prevent the platform from emerging. In a chicken-and-egg situation there are too few participants on each side of the platform, which inhibits the positive feedback of the network effects. The inter-dependence of demand on both sides requires a decision on which side to develop first (Evans and Schmalensee, 2010). Future research is needed to solve the 'chicken-and-egg' problem of Data operator companies.

4.2 Q2: Information presentation

Information presentation is an important part of an information delivery architecture and the challenge of delivering relevant data. The whole architecture processes relevant information from the end-user point of view, but the presentation ensures that the information is delivered, understood and applied in the decision making. Visualisation is an efficient way to present the multitude of data, as humans are good at visual perception, image analysis and recognising patterns, colour, movement and shape. However, without careful design the human visual system is easily overloaded and the information is not delivered. Visualisation of environmental data has been studied earlier. For example, Rink et al. (2014) developed methods to visualise environmental data that are used as an input to environmental models in order to find inconsistencies and artefacts in the data before model runs, Lundblad et al. (2011) considered how visualisation of weather data can help in verifying weather forecasts and analysing the situation, Nelson et al. (2011) developed methods to visualise time-dependent multidimensional water quality data, and Multimäki et al. (2015) studied the visualisation of rain, humidity and pollen concentrations on the same map presentation.

We applied a method used widely in the field of GIS (Geographic Information System) and visualised different information sources as layers on a map that the user can switch on and off, thus choosing the data sources he wants to focus on. However, as our users are operating in a dynamic environment the time dimension of the data becomes an important factor. All the data objects have geographical coordinates and time coordinates, and as the user browses in time and space those objects whose coordinates match the view settings are drawn. When the user browses forward or backward in time an animation can be shown on the screen. However, one challenge is to cope with the different time resolutions of different data sources. For example, satellite images can be taken on consecutive days, while an ice model forecast uses three-hour time steps. The user must have experience and understand the time differences in the presentation in order to interpret the situation correctly.

We wanted to go further than just to visualise various information sources in different layers. We considered the use of automatic analysis to interpret the situation on behalf of the user, and instead of presenting multidimensional information to reduce the complexity and give optimal route choices to support the route planning. This could be considered as an ultimate data compression, as instead of all data only route points are delivered, presented and used in the decision making. The ship route optimization through the ice field is a difficult problem, as ice field modelling and ship transit modelling are both difficult tasks by themselves. Previous studies have considered mainly ship route optimisation in open waters (e.g. Benjamin et al., 2006, Witt and Dunbabin, 2008, Kosmas and Vlachos, 2012), but work has also been done on ship route optimisation in ice-covered waters. Nam et al. (2013) developed an ice navigation method for the Arctic areas using Dijkstra's algorithm, Choi et al. (2013) developed a method based on the genetic algorithm and tested it on a static ice model output in the Arctic region, and Guinness et al. (2014) applied A* algorithm to route optimisation taking into account ice conditions and available icebreaker assistance.

One key question concerning application of automatic analysis of the situation is whether the end-users are willing to trust the results calculated by the system. Because of the uncertainties in ice modelling and ship transit modelling there are many uncertainties in the optimisation results, and more experienced captains probably want to plan routes by themselves using the latest satellite data, weather forecasts and ice model outputs. However, finding an optimal route using multidimensional data is not an easy task for humans, although suitable for computers. Therefore, automatic analyses could give additional value for decision making. For example, in cases when the ice model has predicted the development of the ice field correctly, even an experienced user could trust the routes calculated by the computer and take them into account in route planning. The usability of the optimised routes also depends on the situation that the user faces. When ships are moving in areas where the route choices are limited or they need to follow strict orders by icebreakers, independent navigation is not possible. On the other hand, route optimization might be useful in areas where there are many route alternatives or no icebreaking assistance is available. In addition, the route optimization could be used by icebreakers to estimate safe routes for ships they are assisting.

Improvements in the current implementation should be considered to improve the usability. We used a visual presentation of the variability of the optimum, but in the future a numerical calculation of the uncertainties could be presented. The three parameters that were used to describe the ice field, namely the level ice thickness, the ridged ice thickness and the ice concentration, describe the mean effect of the varying ice thickness to the ship's transit. In the future, we could consider using more detailed description of the ice field, for instance a probability density function (pdf) of the ice thickness and calculate pdf's of ship speeds and transit times instead of plain mean values. For instance, the probability of getting stuck could be a relevant variable in comparing route alternatives. However, presenting pdf's to end-users increases again the complexity of the presentation and requires that the users are familiar with the concept of probability.

Ice thickness is not the only factor affecting the ship speed. The movement of the ice field, i.e. ice drift, should also be included in the optimization, as it might cause compression in ice which makes the ice field more difficult to penetrate, or decompression that might open up new leads (widened cracks in the ice caused by divergence in the ice field, Wadhams, 2002). Ships try to follow leads and also tracks made by other vessels, but modelling these accurately is a challenge for ice models. However, observational data could be used to update ice model outputs concerning leads and tracks. In addition to satellite data, one source of information would be other vessels moving in the area, their tracks and performance in the ice, images captured by ship radars and observations made by humans on board. This leads us in the next section to discuss harnessing end-users as data collectors.

4.3 Q3: Harnessing end-users

We demonstrated two different approaches for harnessing end-users for collecting information from the local conditions, an opportunistic approach using ships and ship radars as sensors and a participatory approach enabling citizen end-users to make observations actively. In the opportunistic approach, the ships operate normally in the ice and the observations are collected as a side product without affecting the actual operations. In the participatory approach, citizens are required to actively observe water quality and deliver their observations with a mobile phone or a web application.

4.3.1 Ships as sensors

The demonstration showed how ships could be used as sensors of the trafficability of the ice field, and ship radar images can be used as an additional source of information about the local conditions. The use of vehicles and vehicle instrumentation as opportunistic sensors has been suggested previously. Marine observations from ships have a long tradition, but lately technology has enabled more efficient measurement by taking advantage of ships as mobile platforms for automatic water quality and meteorological data collection (Petersen, 2014). In addition to marine observation, vehicles have been used as sensor platforms on land. Haberlandt and Sester (2010) suggested and explored theoretically a method for using car windscreen wipers as rain detectors, and Perttunen et al. (2011) presented a method to determine road surface conditions by using accelerometer and GPS readings from a mobile phone attached to a rack in the windshield. Varanka et al. (2008) presented a method to determine road slipperiness based on information of heavy vehicle front and driven axle speeds, engine speed and torque, and Tergujeff et al. (2014) applied the method for a fleet of heavy vehicles to determine road slipperiness in a wider geographical area. Bröring et al. (2015) presented the enviroCar platform for collecting automobile data for traffic monitoring. Guo et al. (2016) presented a method to use standard cars equipped with conventional low-cost sensors as data collectors to support more accurate lane mapping for navigation purposes, e.g. for enhancing autonomous driving.

In the demonstration of ships as sensors we had only one image capture device on board a single ship, and thus we could not test mosaicing of images coming from multiple ships at the same time. However, we formed mosaics using images captured from multiple tracks of a single ship, which demonstrates the idea of using images from multiple ships. Despite the promising results we have not yet applied the method to a fleet of ships (e.g. icebreakers). Image mosaics combining images from many ships would cover a larger area and update the information in the satellite images. They would consist of images from the same time but different spatial locations, and ice drift calculation could be performed by comparing radar images taken from different ships operating in the same area, with each other and with satellite images. In addition, radar images taken from different locations, but looking at the same feature in the ice field, could reveal additional information that facilitates the interpretation of the ice field. Radar images are not an alternative to satellite images but they can be utilized in data fusion with these. The Sentinel-I mission of the European Space Agency (ESA) consists of two satellites and improves the imaging frequency compared to previous single radar satellite systems. However, when the ice field is moving and the situation changes rapidly, additional image data from local conditions could be useful. Even a single image capture device installed in a carefully selected radar location could give relevant information about the movements of the ice field to estimate ice compression (converging ice) in the area.

Using ships as sensors of the trafficability has been used by icebreakers operating in the Baltic Sea. Ship captains have already earlier used the speed of a certain ship as an indicator of the ice conditions in that area. They estimate the ice thickness based on their know-how about the performance of various ships in ice. Using the AIS (Automatic Identification System), the speeds of all ships can be accessed, which improves the feasibility of the method. We used a simple algorithm to determine ICE-v curves that can be used to estimate the speed for a ship given the ice conditions equivalent (ICE), or the ice conditions equivalent given the ship speed. However, the determination of ICE-v curves using the data available contains major sources of uncertainties. Since the AIS does not include engine parameters we have to assume that the ships are using full (or nearly full) thrust when moving in ice. In reality, they do not necessarily use full thrust all the time or they can be slowing down for other reasons than thicker ice. In the future, there should be access to non-static propulsion system parameters in order to determine the thrust delivered by the propellers. This would enable more accurate estimations of the resistance caused by the ice field and improve the statistical determination of ICE–v curves. Another source of error is the information about the prevailing ice conditions. Ice charts and ice models are the best view to the situation in any given point, but they contain spatial and temporal uncertainties. In addition, they do not include tracks made by other vessels or leads that the ships may follow and achieve higher speeds than one would expect from the ice thickness given by an ice chart.

4.3.2 Citizens as sensors

We demonstrated the use of citizens as observers of the algae situation. Based on our analysis, the systematic positive correlations between the expert and citizen algae observations in successive years suggest that citizen observers can extend the current observation network spatially and temporally and provide additional information that supports the algae monitoring. The usefulness of citizen observations was recognised earlier by e.g. Delaney et al. (2008), who concluded that with proper training, citizens can provide reliable aid in collecting knowledge about both native and invasive crabs. A study by Gallo and Waitt (2011) concluded that citizen scientists are able to detect and report invasive plants in their local areas, and that the data can be used by professional scientists. D'Hondt et al. (2012) created noise maps based on citizen observations, and concluded that they are comparable to official simulation-based noise maps.

We wanted to validate the method in realistic conditions, i.e. to collect observations from volunteers who were moving close to waters and wanted to report algae observations. The differences in location and timing of expert and citizen observations did not enable direct comparisons and caused uncertainties in the results, as we used averaging of both space and time. When looking at the graph in Figure 14 we notice that the citizen observations are biased towards higher bloom intensity values than the expert observations. It appeared that citizens made fewer 'no algae detected' observations than the experts. Experts were instructed to observe being there are algae or not, but citizens observed whenever they found it useful. Citizens tend to make observations more often when algae are visible and omit 'no algae' reports. In the future, to study the quality of citizen observations in comparison to expert observations in more detail, a special campaign should be organised to ensure that citizens and experts observe the same algae situation in the same region at the same time.

There are still open issues related to participatory sensing, such as how to motivate the citizens to participate and continue making observations, how to ensure the quality of the data produced, how to preserve the data privacy and enable trust between observers and the system, and how to enable continuous activities instead of project-based campaigns. Different frameworks have been suggested to motivate citizen observers. Reddy et al. (2010a) presented a recruitment framework for identifying potential participants for data collections, and Juong-Sik and Hoh (2010) discussed an incentive mechanism for stimulating participatory sensing applications. Micro-payments as an incentive mechanism were explored by Reddy et al. (2010b). Gamification has been studied as a method for motivation for example by Han et al. (2011) and Bowser et al. (2014).

Haklay et al. (2010) and Comber et al. (2013) discussed the importance of data quality when using VGI (Volunteered Geographic Information). In many cases, the quality of data varies, it is not documented, it fails to follow scientific principles of sampling design, and its coverage is incomplete (Goodchild and Li, 2012). Kalantari et al. (2014) discussed the common lack of metadata associated with data, and suggested that creation of metadata can provide better understanding and appreciation of data quality. See et al. (2013) compared the quality of remote sensing image analysis performed by experts and non-experts, and discussed the relevance of training material in cases where non-experts had difficulties.

In citizen science systems there is a risk of faulty input due to human errors or even service misuse. The risk for misuse is higher when no registration to the service is required. We did not require any registration for the mobile part of the system, which makes adoption of the system easier, but, at the same time, makes misuse of the system more likely and hampers the analysis of user activity, motivation and data quality. In addition, when the observations are based on human senses the measurements are not of uniform quality, but vary according to individual skill, sensitivity and mood. Tasks should be designed to cope with these kinds of error sources. Alabri et al. (2010) described a framework combining the data quality control and trust metrics to enhance the reliability of citizen science data. Kuan et al. (2010) and Yang et al. (2011) proposed reputation management systems to evaluate the trustworthiness of gathered data. Sunyoung et al. (2011) suggested that creating a successful citizen sensing application requires designing the application together with various stakeholders and ensuring that the gathered data can be put to use.

One relevant issue about harnessing citizens as data collectors is the continuity of the activities. There are examples of successful projects in which citizens have provided additional value, but the activity has ceased when the project has ended. In the future, new mechanisms, standards and cooperation between public and private stakeholders are needed to enable new markets for citizen contributions. Commercial interests for companies to develop citizen science systems and use citizen data in their business would enable sustainable activities and make citizens a continuous data source to complement other sources. In the future, Data operator companies could implement the synergies in citizen data collection and data exchange between systems, find new value for citizen data and promote the emergence of new participatory sensing markets.

5. Summary and conclusions

Timely delivery of relevant environmental information is crucial for decision making in dynamic environments. This thesis developed improvements in delivering near real-time information about the surrounding environment to support decision making in dynamic environmental conditions. We focused on two application cases, ice navigation and water quality monitoring, and formed three specific research objectives to find answers to three specific research questions. We wanted to improve the orchestration of the information processing in order to tackle the challenges of near real-time environmental information delivery, to improve the information presentation to end-users, and to harness end-users as collectors of additional data about the local conditions to support other information sources.

To study the research questions we developed information processing and delivery orchestration architectures using the Service-Oriented Architecture (SOA) model, Business Process Execution Language (BPEL) and Enterprise Service Bus (ESB) model. We developed an orchestrator component called Facade to manage the synergies in the data processing from the point of view of one application serving a multitude of end-users with different needs, and extended the Facade model towards the Data operator model that implements the synergies of data processing of a multitude of applications. We considered how the information presentation could be improved by using automatic analyses to reduce the complexity of multidimensional data and applied route optimisation to the ice navigation case. Finally, we applied opportunistic methods for harnessing ships and ship radars to collect additional data to support ice navigation, and participatory methods to harness citizens to make observations about water quality to complement expert observations and satellite data.

As a conclusion, we answer the main research question: *How could near real-time environmental information delivery be improved to reduce the amount of surprise and better support decision making of both professional and layman end-users operating in dynamic environments?* We found that three system design principles can be used. By organising the synergies in the data processing according to the Data operator model, applications can access the fragmented data sources more efficiently, save costs in the implementation and running of the system, and improve the robustness of the information delivery. Information presentation is a relevant part of the information delivery chain and by using a computer to analyse the situation on behalf of the user we can reduce the amount of data delivered, reduce the complexity of information presentation and support users in decision making. Harnessing end-users as data collectors can provide additional information that complements other data sources and supports the situational picture. Figure 15 presents an updated version of Figure 1 in which we have added the three system design principles to the conceptual view: Data operator to implement the synergies in accessing and processing data, automatic analyses to reduce the amount of data delivered and support the information interpretation by reducing the complexity of information presentation, and data collection from end-users to complement other data sources.

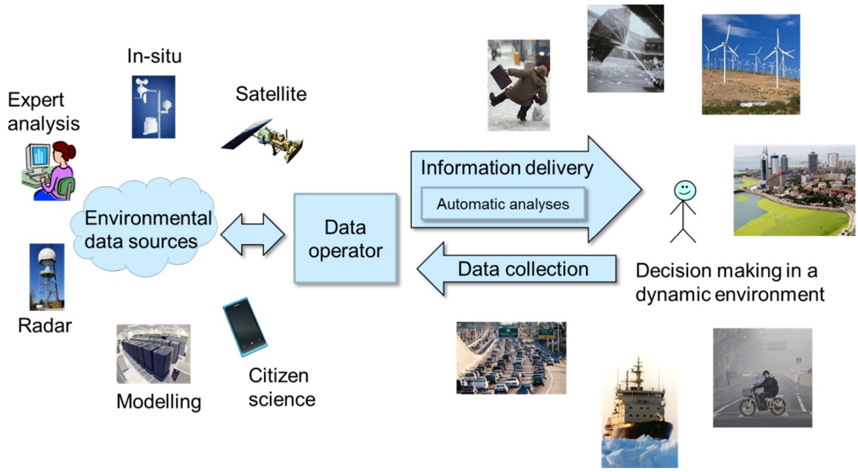

Figure 15. Answering the main research question. Near real-time environmental information delivery can be improved by 1) Data operator to implement the synergies in the data processing and data access, 2) automatic analyses to reduce the amount of data delivered and to support the information interpretation, and 3) harnessing end-users as collectors of data about local conditions.

We focused on two application cases and are limited to those in our studies. However, as a final conclusion we suggest that the design principles could be applied to other applications coping with near real-time environmental data such as road traffic, weather, air quality, disaster and built environment monitoring. In addition, the methods could be extended to environmental monitoring and environmental informatics applications that consider historical time series, long-term analysis and longer time scales (seasonal, annual, etc.), for example to tackle air pollution, deforestation and climate change.

Open data and open interfaces are important elements for accessing data sources, but they are not adequate to guarantee the optimal use of ever-increasing amounts of environmental data. The whole processing chain from data sources to end-user awareness should be considered in order to take full advantage of the data. The amount of data is increasing due to technological advances and the global interest in environmental issues. Humanity is producing more and more data about the environment, either by intentional environmental monitoring or unintentionally by various activities and processes in society. In addition to traditional environmental monitoring, environmental data are produced by people moving in the environment using mobile devices equipped with sensors, by people using social media tools to take pictures and videos, make sound recordings and write texts about environmental issues, and by human-driven and autonomous vehicles that measure the environmental conditions for navigational purposes. New methods are needed to improve the use of these data in decision making in society.

References

Alabri A, Hunter J. 2010. Enhancing the quality and trust of citizen science data. In Proceedings of 6th IEEE International Conference on e-Science, eScience 2010, Brisbane, Australia, 7-10 December 2010, pp. 81-88.

Artiola J., Pepper I., Brusseau M. 2004. Environmental Monitoring and Characterization, Elsevier Academic Press, Amsterdam, The Netherlands, 2004.

Axell L. 2005. Extended ice forecast modelling for the Baltic Sea. In: Dahlin, H., Flemming, N.C., Marchand, P., and Petersson, S.E. (eds.), European Operational Oceanography: Present and Future, Proceedings of the Fourth International Conference on EuroGOOS, 6-9 June 2005, Brest, France, pp. 531-534.

Barrington L., Ghosh S., Greene M., Har-Noy S., Berger J., Gill S., Lin A., Huyck C. 2012. Crowdsourcing earthquake damage assessment using remote sensing imagery. Annals Of Geophysics 54(6): 680-687.

Benjamin M., Curcio J., Leonard J. 2006. Navigation of Unmanned Marine Vehicles in Accordance with the Rules of the Road. In Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 15-19 May 2006, Orlando, Florida, pp. 3581-3587.

Botts M., Percivall G., Reed C., Davidson J. 2008. OGC® Sensor Web Enablement: Overview and High Level Architecture. GeoSensor Networks, Lecture Notes in Computer Science 4540: 175-190.

Bowser A., Hansen D., Preece J., He Y., Boston C., Hammock J. 2014. Gamifying citizen science: A study of two user groups. In Proceedings of the ACM Conference on Computer Supported Cooperative Work (CSCW), 14-18 March 2014, Vancouver, Canada, pp. 137-140.

Bröring A., Maue P., Janowicz K., Nuest D., Malewski C. 2011. Semantically-Enabled Sensor Plug & Play for the Sensor Web. Sensors 11(8): 7568-7605.

Bröring A., Remke A., Stasch C., Autermann C., Rieke M., Möllers J. 2015. enviroCar: A Citizen Science Platform for Analyzing and Mapping Crowd-Sourced Car Sensor Data. Transactions in GIS 19(3): 362-376.

Burke J., Estrin D., Hansen M., Parker A., Ramanathan N., Reddy S., Srivastava M. 2006. Participatory Sensing. In Proceedings of the International Workshop on World-Sensor-Web (WSW'2006), 31 October - 3 November 2006, Boulder, Colorado, USA, pp. 117-134.

Buyya R., Chee Shin Yeo, Venugopal S., Broberg J., Brandic I. 2009. Cloud computing and emerging IT platforms: vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems 25(6): 599-616.

Candido dos Reis F., Lynn S., Raza A., Eccles D., Hanby A., Provenzano E., Caldas C., Howat W., McDuffus L., Liu B., Daley F., Coulson P., Vyas R.,. Harris L., Owens J., Carton A., McQuillan J., Paterson A,, Hirji Z., Christie S., Holmes A., Schmidt M., Garcia-Closas M., Easton D., Bolla M., Wang Q., Benitez J., Milne R., Mannermaa A., Couch F., Devilee P., Tollenaar R., Seynaeve C., Cox A., Cross S., Blows F., Sanders J., de Groot R., Figueroa J., Sherman M., Hooning M., Brenner H., Holleczek B., Stegmaier C., Lintott C., Pharoah P. 2015. Crowdsourcing the General Public for Large Scale Molecular Pathology Studies in Cancer, EBioMedicine 2(7): 681-689.

Cao J. 2009. Flexible data collection over distributed sensors on Enterprise Service Bus. In Proceedings of International Conference on Frontier of Computing Science and Technology, 17-19 December 2009, Shanghai, pp. 226-230.

Cao T, Thompson J. 2014. Remote Sensing of Atmospheric Optical Depth Using a Smartphone Sun Photometer. PLoS ONE 9(1): e84119.

Carvalho L., Poikane S., Lyche Solheim A., Phillips G., Borics G., Catalan J., De Hoyos C., Drakare S., Dudley B., Järvinen M., Laplace-Treyture C., Maileht K., McDonald C., Mischke U., Moe J., Morabito G., Nõges P., Nõges T., Ott I., Pasztaleniec A., Skjelbred B., Thackeray S. 2013. Strength and uncertainty of lake phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704: 127-140.

Ceccato P., Fernandes K., Ruiz D., Allis E. 2014. Climate and environmental monitoring for decision making. Earth Perspectives 2014 1(16).

Chappell D. 2004. Enterprise Service Bus. O'Reilly Media Inc.

Chen Y., Hardisty A., Preece A., Martin P., Atkinson M., Zhao Z., Magagna B., Schentz H., Legré Y. 2013a. Analysis of Common Requirements for Environmental Science Research Infrastructures. In Proceedings of the International Symposium on Grids and Clouds (ISGC), 17-22 March 2013, Taipei, Taiwan.

Chen Y., Martin P., Magagna B., Schentz H., Zhao Z., Hardisty A., Preece A., Atkinson M., Huber R., Legré Y. 2013b. A Common Reference Model for Environmental Science Research Infrastructures. In: Page B., Fleischer A., Göbel J., Wohlgemuth V., (eds.), 27th International Conference on Environmental Informatics for Environmental Protection, Sustainable Development and Risk Management, 2-4 September 2013, Hamburg, Germany, pp. 665-673.

Choi M., Chung H., Yamaguchi H., De Silva L. 2013. Application of genetic algorithm to ship route optimization in ice navigation. In Proceedings of the 22nd International Conference on Port and Ocean Engineering under Arctic Conditions, POAC'13, June 9-13, 2013, Espoo, Finland.

Comber A., See L., Fritz S., Van der Velde M., Perger C., Foody G. 2013. Using control data to determine the reliability of volunteered geographic information about land cover, International Journal of Applied Earth Observation and Geoinformation 23: 37-48.

Conrad C., Hilchey K. 2011. A review of citizen science and community-based environmental monitoring: issues and opportunities. Environmental Monitoring and Assessment 176(1): 273-291.

D'Hondt E, Stevens M, Jacobs A. 2013. Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. Pervasive Mobile Computing 9(5): 681-694.

Danduran P., Mouchot M., Garello R., Fleury D., Thepaut E. 1997. First realtime use of RADARSAT SAR imagery for ship navigation in Antarctica. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) '97, 3-8 August, 1997, Singapore, pp. 1678-1680.

Das T., Mohan P., Padmanabhan V., Ramjee R., Sharma A. 2010. PRISM: Platform for Remote Sensing using Smartphones. In Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys '10), 15-18 June 2010, San Francisco, California, USA, pp. 63-76.

Delaney D., Sperling C., Adams C., Leung B. 2008. Marine invasive species: validation of citizen science and implications for national monitoring networks. Biological Invasions 10(1).

Eiben C., Siegel J., Bale J., Cooper S., Khatib F., Shen B., Players Foldit, Stoddard B., Popovic, Z., Baker D. 2012. Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. Nature Biotechnology, Feb 2012, 30(2):190-192.

Eiman K. 2010. NoiseSPY: a real-time mobile phone platform for urban noise monitoring and mapping. Mobile Networks and Applications 15(4):562-574.

Eisenmann T., Parker G., Van Alstyne M. 2006. Strategies for two-sided markets. Harvard Business Review 84(10): 92-101.

Erl T. 2005. Service-oriented architecture: concepts, technology, and design. Prentice Hall Professional Technical Reference.

European Parliament, Council. 2007. Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), Brussels, Belgium.

Evans D., Schmalensee R. 2010. Failure to Launch: Critical Mass in Platform Businesses. Review of Network Economics 9(4).

Frew J., Dozier J. 2012. Environmental Informatics. Annual Review of Environment and Resources 37: 449-472.

Fritz S., McCallum I., Schill C., Perger C., See L., Schepaschenko D., van der Velde M., Kraxner F. and Obersteiner M. 2012. Geo-Wiki: An online platform for improving global land cover. Environmental Modelling and Software 31(2012): 110-123.

Gallo T., Waitt D. 2011. Creating a successful citizen science model to detect and report invasive species. Bioscience 61(6): 459-465.

GEO secretariat. 2010. Report on progress. GEO Secretariat, Geneva, Switzerland.

Goodchild M. 2007. Citizens as sensors: the world of volunteered geography. GeoJournal, August 2007, 69(4): 211-221.

Goodchild M., Li L. 2012. Assuring the quality of volunteered geographic information. Spatial Statistics 1: 110-120.

Ginsberg J., Mohebbi M., Patel R., Brammer L., Smolinski M., Brilliant L. 2009. Detecting influenza epidemics using search engine query data. Nature 457(7232): 1012-1014.

Giuliani G., Ray N., Schwarzer S., De Bono A., Peduzzi P., Dao H., Van Woerden J., Witt R., Beniston M., Lehmann A. 2011. Sharing environmental data through GEOSS. International Journal of Applied Geospatial Research 2(1).

Goward S. 2007. Land remote sensing in the 21st Century. Geotechnologies in service to human societies. GeoFocus (Editorial) 7:1-4.

Guinness R.E., Saarimäki J., Ruotsalainen L., Kuusniemi H., Goerlandt F., Montewka J., Berglund R., Kotovirta V. 2014. A method for ice-aware maritime route optimization. In Proceedings of IEEE/ION Position, Location and Navigation Symposium (PLANS) 2014, 5-8 May 2014, Monterey, California, USA, pp. 1371-1378.

Guo C., Kidono K. Meguro J., Kojima Y., Ogawa M., Naito T. 2016. A Low-Cost Solution for Automatic Lane-Level Map Generation Using Conventional In-Car Sensors. IEEE Transactions on Intelligent Transportation Systems, PP(99) (Early Access).

Gutierrez S., Long D.G. 2003. Application of optical flow and scale space methods to sea–ice motion in Antarctica. Proceedings of SPIE - The International Society for Optical Engineering 5155: 249-256.

Haberlandt U., Sester M. 2010. Areal rainfall estimation using moving cars as rain gauges: A modelling study. Hydrology and Earth System Sciences 14: 1139-1151.

Haklay M. 2012. Citizen Science and Volunteered Geographic Information – overview and typology of participation. In: Sui DZ, Elwood S, Goodchild MF (eds.) Volun-

teered Geographic Information, Public Participation and Crowdsourced Production of Geographic Knowledge, Springer, Berlin, pp. 105-122.
Haklay M., Basiouka S., Antoniou V., Ather A. 2010 How Many Volunteers Does it Take to Map an Area Well? The Validity of Linus' Law to Volunteered Geographic Information. Cartographic Journal 47(4): 315-322.
Han K., Graham E., Vassallo D., Estrin D. 2011. Enhancing Motivation in a Mobile Participatory Sensing Project through Gaming. In 2011 IEEE International Conference on Privacy, Security, Risk, and Trust, 9-11 October 2011, Boston, Massachusetts, USA, pp. 1443-1448.
Heil P., Fowler C.W., Maslanik J.A., Emery W.J., Allison I. 2000. A comparison of East Antarctic sea–ice motion derived using drifting buoys and remote sensing. In Proceedings of International Symposium on Sea Ice and its Interactions with the Ocean, 18–23 June 2000, Fairbanks, Alaska, USA, pp. 139-144.
Hilty L., Page B. 1995. Environmental Informatics as a New Discipline of Applied Computer Science. In: Avouris N., Page B. (eds.), Environmental Informatics: Methodology and Applications of Environmental Information Processing. Springer Netherlands, pp. 1-11.
Inoue J., Enomoto T., Miyoshi T., Yamane S. 2009. Impact of observations from Arctic drifting buoys on the reanalysis of surface fields. Geophysical Research Letters 36(8): L08501.
Hill D., Minsker B. 2010. Anomaly detection in streaming environmental sensor data: A data-driven modeling approach. Environmental Modelling & Software 25(9): 1014-1022.
Hu Z., Wu W. 2012. A satellite data portal developed for crowdsourcing data analysis and interpretation. In Proceedings of 2012 IEEE 8th International Conference on E-Science (e-Science), 8-12 October 2012, Chicago, USA, pp. 1-8.
ISO/IEC 10746-1. 1998. Information technology Open Distributed Processing Reference Model: Overview, ISO/IEC standard.
Juong-Sik L., Hoh B. 2010. Dynamic pricing incentive for participatory sensing. Pervasive and Mobile Computing 6(6): 693-708.
Kalantari M., Rajabifard A., Olfat H., Williamson I. 2014. Geospatial Metadata 2.0 – An approach for Volunteered Geographic Information. Computers, Environment and Urban Systems 48: 35-48.
Karvonen J., Similä M., Heiler I. 2003. Ice thickness estimation using SAR data and ice thickness history. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2003, 21-25 July 2003, Tolouse, France, 1:74-76.
Karvonen J., Simila M., Lehtiranta J. 2007. SAR-based estimation of the Baltic Sea ice motion. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2007, 23-28 July 2007, Barcelona, Spain, pp. 2605-2608.
Karydis M., Kitsiou D. 2013. Marine water quality monitoring: A review. Marine Pollution Bulletin 77: 23-36.
Kirkpatrick S., Gelatt Jr. C., Vecchi M. 1983. Optimization by Simulated Annealing. Science 220(4598): 671-680.
Kotovirta V., Berglund R., Haajanen J., Mansner J., Tergujeff R. 2003. Delivering Near Real-Time Met-Ice-Ocean Observation and Forecast Data – the IWICOS Façade. In: Building the European Capacity in Operational Oceanography, Proceedings of the Third International Conference on EuroGOOS, 3–6 December 2002, Athens, Greece, pp. 639-641.
Kotovirta V., Hytönen H., Aalto A. 2013. Extensible Architecture for Distributed Near Real-Time Environmental Monitoring. Measurement, Monitoring and Environmental Assessment (MMEA) research program, Research report D1.1.7,

Cleen Ltd, 2013, Helsinki. Available in https://portal.cleen.fi/_layouts/IWXmlPublications.aspx?Source=MMEA&FileId=1673 (accessed Sept 2016)

Kotovirta V., Toivanen T., Tergujeff R., Huttunen M. 2012. Participatory Sensing in Environmental Monitoring – Experiences. Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing, 4-6 July 2012, Palermo, Italy, pp. 155-162.

Kosmas O., Vlachos D. 2012. Simulated annealing for optimal ship routing. Computers & Operations Research 39(2012): 576-581.

Kuan LH, Salil SK, Wen H. 2010. Are you contributing trustworthy data?: the case for a reputation system in participatory sensing. In Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems (MSWIM '10), October 17-21 2010, Bodrum, Turkey, pp. 14-22.

Kämäräinen J. 1986. Icebreaking tanker transit simulation model. In International Offshore and Navigation Conference and Exhibition, Polartech '86, 27-30 October 1986, Espoo, Finland, Vol 2, pp. 1138-1159.

Lane N., Eisenman S., Musolesi M., Miluzzo E., Campbell A. 2008. Urban sensing systems: opportunistic or participatory? In Proceedings of the 9th workshop on Mobile computing systems and applications, 25–26 Feb 2008, California, USA, pp. 11-16.

Lee K., Murray D., Hughes D., Joosen W. 2010. Extending sensor networks into the Cloud using Amazon Web Services. In IEEE International Conference on Networked Embedded Systems for Enterprise Applications (NESEA), 25-26 November 2010, Suzhou, China, pp. 1-7.

Leeuw T., Boss E. 2014. Crowdsourcing water quality data using the iPhone camera. Poster presented at the 2014 Ocean Science Meeting, 22 February - 3 March 2014, Honolulu, Hawaii. Available online http://www.eposters.net/pdfs/crowdsourcing-water-quality-data-using-the-iphone-camera.pdf. (Accessed Sept 2016.)

Leisti H., Riska K., Heiler I., Eriksson P., Haapala J. 2009. A method for observing compression in sea ice fields using IceCam. Cold Regions Science and Technology 59: 65-77.

Li H., Wu B. 2011. A Service-Oriented Architecture for Proactive Geospatial Information Services. Future Internet 3(4): 298-318.

Lim S.B., Yoon K., Eo Y.D. 2012. Ubiquitous air quality monitoring system with service oriented architecture middleware. Journal of Convergence Information Technology 7(6): 193-201.

Lowry C., Fienen M. 2013. CrowdHydrology: crowdsourcing hydrologic data and engaging citizen scientists. GroundWater 51(1):151-156.

Luckham D. 2002. The Power of Events: An Introduction to Complex Event Processing in Distributed Enterprise Systems. Addison-Wesley Professional.

Lundblad P., Lofving H., Elovsson A., Johansson J. 2011. Exploratory Visualization for Weather Data Verification. In Proceedings of 15th International Conference on Information Visualisation, 13-15 July 2011, London, United Kingdom, pp. 306-313.

Mahoney A., Eicken H., Shapiro L. 2007. How fast is landfast sea ice? A study of the attachment and detachment of nearshore ice at Barrow, Alaska. Cold Regions Science and Technology 47: 233-255.

Mednis A., Strazdins G., Zviedris R., Kanonirs G. 2011. Real time pothole detection using Android smartphones with accelerometers. In Proceedings of the 7th

IEEE International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS'11), 27–29 June 2011, Barcelona, Spain, pp. 1-6.

Michener W., Allard S., Budden A. 2012. Participatory design of DataONE—enabling cyberinfrastructure for the biological and environmental sciences. Ecological Informatics 11(2012):5-15.

Molinier M., Häme T., Toivanen T., Andersson K., Mutanen T. 2014. Relasphone - Mobile phone and interactive applications to collect ground reference biomass data for satellite image analysis. In Proceedings of the Geoscience and Remote Sensing Symposium (IGARSS) 2014, 13-18 July 2014, Quebec, Canada, pp. 836-839.

Motwani R., Motwani M., Harris F., Dascalu, S. 2010. Towards a scalable and interoperable global environmental sensor network using service oriented architecture. In Proceedings of the 2010 Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP 2010), 7-10 December 2010, Brisbane, Australia, pp. 151-156.

Multimäki S., Mäkilä A., Ahonen-Rainio P. 2015. Combining Two Datasets into a Single Map Animation. In Proceedings of the 1st ICA European Symposium on Cartography, EuroCarto 2015, 10-12 November 2015, Vienna, Austria, pp. 41-42.

Mun M., Reddy S., Shilton K., Yau N., Burke J., Estrin D., Hansen M., Howard E., West R., Boda P. 2009. PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In MobiSys '09 Proceedings of the 7th international conference on Mobile systems, applications, and services, 22–25 June 2009, Kraków, Poland, pp. 55-68.

Mårtensson S., Meier H., Pemberton P., Haapala J. 2012. Ridged sea ice characteristics in the Arctic from a coupled multicategory sea ice model. Journal of Geophysical Research 117:C00D15.

Nam J., Park I. Lee H., Kwon M., Choi K., Seo Y. 2013. Simulation of optimal arctic routes using a numerical sea ice model based on an ice-coupled ocean circulation method. International Journal of Naval Architecture and Ocean Engineering 5(2): 210-226.

Nativi S., Craglia M., Pearlman J. 2013a. Earth Science Infrastructures Interoperability: The Brokering Approach. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 6(3): 1118-1129.

Nativi S., Mazzetti P., Geller G. 2013b. Environmental model access and interoperability: The GEO Model Web initiative. Environmental Modelling & Software 39(2013): 214-228.

Nelder J., Mead R. 1965. A Simplex Method for Function Minimization. Computer Journal 7:308-313.

Nelson J., Williams G., Obregon O, Buahin C. 2011. Dynamic Multidimensional Visualization for Water Quality Data in Rivers. In Proceedings of World Environmental and Water Resources Congress, 22-26 May 2011, Palm Springs, California, USA, pp. 4811-4819.

Olmanson L., Bauer M., Brezonik P. 2008. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. Remote Sensing of Environment 112: 4086-4097.

Palmer S., Kutser T., Hunter P. 2015. Remote sensing of inland waters: Challenges, progress and future directions. Remote Sensing of Environment 157: 1-8.

Pedersen L., Saldo R. 2005. Experience with near real time distribution of envisat asar data to end-users. European Space Agency, (Special Publication) ESA SP 572: 733-738.

Perez A., Labrador M., Barbeau J. 2010. G-sense: a scalable architecture for global sensing and monitoring. IEEE Network 24(4):57-64.

Perttunen M., Mazhelis O., Cong F., Kauppila M., Leppänen T., Kantola J., Collin J., Pirttikangas S., Haverinen J., Ristaniemi T., Riekki J. 2011. Distributed Road Surface Condition Monitoring Using Mobile Phones. In Proceedings of International Conference on Ubiquitous Intelligence and Computing (UIC 2011), 2-4 September 2011, Banff, Canada, pp. 64-78.

Petersen W. 2014. FerryBox systems: State-of-the-art in Europe and future development. Journal of Marine Systems 140(2014): 4-12.

Pettersson L., Sandven S., Dalen O., Melentyev V., Babich N. 2000. Satellite radar ice monitoring for ice navigation of a tanker convoy in the Kara Sea. In Proceedings of the International Offshore and Polar Engineering Conference, 27 May - 2 June 2000, Seattle, USA, pp. 691-696.

Plale B., Gannon D., Brotzge J., Droegemeier K., Kurose J., McLaughlin D., Wilhelmson R., Graves S., Ramamurthy M., Clark R., Yalda S., Reed D., Joseph E., Chandrasekar V. 2006. CASA and LEAD: Adaptive Cyberinfrastructure for Real-Time Multiscale Weather Forecasting. Computer 39(11): 56-64.

Powell M. 1964. An efficient method for finding the minimum of a function of several variables without calculating derivatives. Computer Journal 7(2): 155-162.

Pyhälahti T., Toivanen T., Kallio K., Järvinen M., Molinier M., Koponen S., Kotovirta V., Peng C., Anttila S., Laanen M., Lindholm M. 2015. Advances in Combining Optical Citizen Observations on Water Quality With Satellite Observations as Part of an Environmental Monitoring System. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2015, 26-31 July 2015, Milan, Italy, pp. 5395-5398.

Quinn N., Ortega R., Rahilly P., Royer C. 2010. Use of environmental sensors and sensor networks to develop water and salinity budgets for seasonal wetland real-time water quality management. Environmental Modelling & Software 25(9):1045-1058.

Rampal P., Weiss J., Marsan D., Lindsay R., Stern H. 2008. Scaling properties of sea ice deformation from buoy dispersion analysis. Journal of Geophysical Research: Oceans 113(3).

Rapala J., Kilponen J., Järvinen M., Lahti K. 2012. Finland: Guidelines for monitoring of cyanobacteria and their toxins. In: Chorus I. (ed.), Current approaches to cyanotoxin risk assessment, risk management and regulations in different countries. Federal Environment Agency (Umweltbundesamt), Germany, pp. 54-62.

Reddy S., Estrin D., Srivastava M. 2010a. Recruitment framework for participatory sensing data collections. In Proceedings of the 8th International Conference on Pervasive Computing, 17-20 May 2010, Helsinki, Finland, pp. 138-155.

Reddy S., Estrin D., Hansen M., Srivastava M. 2010b. Examining Micro-Payments for Participatory Sensing Data Collections. In Proceedings of the 12th ACM International Conference on Ubiquitous Computing (Ubicomp '10), 26-29 September 2010, Copenhagen, Dennmark, pp. 33-36.

Reinart A., Kutser T. 2006. Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. Remote Sensing of Environment 102(1-2): 74-85.

Rink K., Bilke L., Kolditz O. 2014. Visualisation strategies for environmental modelling data. Environmental Earth Sciences 72(10):3857-3868.

Riska K., Wilhelmson M., Englund K., Leiviskä T. 1997. Performance of merchant vessels in ice in the Baltic. Winter Navigation Research Board, Report 52, December 1997, Espoo, Finland.

Roots E. 1997. Inclusion of different knowledge systems in research. In: Manseau M. (ed.), Terra Borealis. Traditional and Western Scientific Environmental

Knowledge. Workshop Proceedings, 10-11 September 1997, Northwest River, Labrador, Canada, pp. 42–49.

Rönkkö M., Kotovirta V., Karatzas K., Bastin L., Stocker M., Kolehmainen M. 2012. Proactive Environmental Systems: the Next Generation of Environmental Monitoring. In: Seppelt, R., Voinov, A.A., Lange, S., Bankamp, D. (eds.), iEMSs 2012 International Congress on Environmental Modelling and Software, 1-5 July 2012, Leipzig, Germany, pp. 1-8.

Rönkkö M., Kotovirta V., Kolehmaine, M. 2013. Classifying environmental monitoring systems. In: Environmental Software Systems. Fostering Information Sharing: 10th IFIP WG 5.11 International Symposium ISESS 2013, 9-11 October 2013, Neusiedl am See, Austria, pp. 533-542.

See L, Comber A, Salk C, Fritz S, van der Velde M, Perger C., Schill C., McCallum I., Kraxner F., Obserstein M. 2013. Comparing the Quality of Crowdsourced Data Contributed by Expert and Non-Experts. PLoS ONE 8(7): e69958.

Smith L., Stephenson S. 2013. New Trans-Arctic shipping routes navigable by midcentury. In Proceedings of the National Academy of Sciences of the United States of America 110(13): E1191-E1195.

Solimini A., Cadroso A., Heiskanen A. 2006. Indicators and methods for the ecological status assessment under the Water Framework Directive. Office for Official Publications of the European Communities, Luxembourg. ISBN 92-79-02646-1.

Stocker M., Kauhanen O., Hiirsalmi M., Saarela J., Rossi P., Rönkkö M., Hytönen H., Kotovirta V., Kolehmainen M. 2015. A software system for the discovery of situations involving drivers in storms. In Proceedings of Environmental Software Systems. Infrastructures, Services and Applications: 11th IFIP WG International Symposium (ISESS 2015), 25-27 March 2015, Melbourne, Australia, pp. 226-234.

Suakanto S., Supangkat S., Suhardi, Saragih R., Nugroho T., Nugraha I. 2012. Environmental and disaster sensing using cloud computing infrastructure. In Proceedings of the 2012 International Conference on Cloud Computing and Social Networking (ICCCSN), 26-27 April 2012, Bandung, West Java, Indonesia, pp. 1-6.

Sun H., Shen Y., Qiu D. 2004. Measuring sea ice drift via cross-correlation of radar ice images. Journal of Hydrodynamics 16(6): 716-720.

Sunyoung K., Robson C., Zimmerman T., Pierce J., Haber E. 2011. Creek Watch: Pairing Usefulness and Usability for Successful Citizen Science. In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI '11), 7-12 May 2011, Vancouver, Canada, pp. 2125-2134.

Tergujeff R., Silvonen P., Pilli-Sihvola E., Hautala R. 2014. Dynamic road slipperiness detection based on in-vehicle sensors. In Proceedings of the 10th ITS European Congress, 16-19 June 2014, Helsinki, Finland, 9 pp.

Toivanen T., Koponen S., Kotovirta V., Molinier M., Chengyuan P. 2013. Water quality analysis using an inexpensive device and a mobile phone. Environmental Systems Research 2(9).

Usländer T., Jacques P., Simonis I., Watson K. 2010. Designing environmental software applications based upon an open sensor service architecture. Environmental Modelling & Software 25(9):977-987.

Vainio J., Similä M., Grönvall H. 2000. Operational use of RADARSAT SAR data as aid to winter navigation in the baltic sea. Canadian Journal of Remote Sensing 26(4): 314-317.

Varanka M., Erkkilä K., Noponen K., Mäkelä K., Seppänen T. 2008. Automatic road slipperiness detection of heavy duty vehicles. In 11th International Symposium

on Wireless Personal Multimedia Communications (WPMC 2008), 8-11 September, Saariselkä, Finland, 2008, 5 pp.
Wadhams P. 2002. Ice in the ocean. Gordon and Breach Science.
Wang J., Korayem M., Crandall D. 2013. Observing the natural world with Flickr. 2013. In IEEE International Conference on Computer Vision Workshops, 1-8 December 2013, Sydney, Australia, pp. 452-459.
Wilhelmsson T. 2002. Parallellization of the HIROMB ocean model, Licentiate Thesis, Royal Institute of Technology, Department of Numerical and Computer Science, Stockholm, Sweden.
Williams M., Cornford D., Bastin L., Jones R., Parker S. 2011. Automatic processing, quality assurance and serving of real-time weather data. Computers & Geosciences 37(3): 351-362.
Witt J. Dunbabin M. 2008. Go with the Flow: Optimal AUV Path Planning in Coastal Environments. In Proceedings of the Australasian Conference on Robotics and Automation, 3-5 December 2008, Canberra, Australia, pp. 1-9.
Wächter J., Babeyko A., Fleischer J., Häner R., Hammitzsch M., Kloth A., Lendholt M. 2012. Development of tsunami early warning systems and future challenges. Natural Hazards and Earth System Sciences 12(6):1923-1935.
Yang H., Zhang J., Roe P. 2011. Using Reputation Management in Participatory Sensing for Data Classification. Procedia Computer Science 5:190-197.
Yao Y., Sharma A., Golubchik L., Govindan R. 2010. Online anomaly detection for sensor systems: A simple and efficient approach. Performance Evaluation 67(11):1059-1075.
Yue P., Baumann P., Bugbee K., Jiang L. 2015. Towards intelligent GIServices. Earth Science Informatics 8(3): 463-481.
Zins C. 2007. Conceptual approaches for defining data, information, and knowledge. Journal of the American Society for Information Science and Technology 58(4): 479-493.

I

Publication I

Kotovirta V., Kanniainen J., Veijonen T., Neuvonen S. 2006. Building an Application Framework for Monitoring the Environment. In Proceedings of International Geoscience and Remote Sensing Symposium (IGARSS) 2006, 31 July - 4 August 2006, Denver, Colorado, USA, pp. 2149-2152.

Building an application framework for monitoring the environment

Ville Kotovirta, Jenni Kanniainen, Teppo Veijonen, Simo Neuvonen
VTT Technical Research Centre of Finland
P.O. Box 1000, FI-02044 VTT, Finland

***Abstract*—In the Envimon project we are building prototype information systems for diverse real-world environment monitoring applications. The applications are built onto a common software framework, which will be designed and implemented in the project. The framework can be used for developing diverse applications in a straightforward and cost-efficient way. It enables employing diverse data sources, pre-processing, analyzing and preparing the data into proper products, and delivering them at the right time, in the right form and via the appropriate channel to different end users that use diverse terminal equipment. The core of the framework consists of workflow management. A workflow description written in a workflow language represents the data process used in an application, and the workflow engine manages the execution of the process. The workflow tasks are implemented as web services, and the framework architecture utilizes XML-based standards like WSDL (Web Service Description Language) and BPEL (Business Process Execution Language). The feasibility of the concept introduced by the software framework is verified by implementing several environment monitoring systems for the real-world applications in the areas of disaster monitoring, forestry, forest fire, maritime, traffic monitoring, disposal site monitoring, and season monitoring for tourism.**

Keywords—earth observation; spatial data; framework; environment monitoring; workflow; BPEL; web services; web service orchestration

I. Introduction

There are more and more sensors and sensor networks available around and above us and new applications that combine earth observation (EO) data with in-situ measurements are developed to monitor changes in our environment. In the Envimon project we build prototype information systems for diverse real-world environment monitoring applications in the areas of disaster monitoring, forestry, forest fire, maritime, traffic monitoring, disposal site monitoring, and season monitoring for tourism. The applications have common requirements for top-level functionality as they collect data coming from a diversity of sensors, preprocess and analyze the data, and deliver final products to other systems, users and decision makers. Some of the applications operate in near real-time fashion, and the processing they involve must be carried out as fast and reliably as possible. To achieve speed in the processing we need to reduce human related tasks whenever it is possible, and make the processing chain as automatic as possible. Some of the analysis steps still require manual work and are not possible to be carried out without a human expert, but some of the more rudimentary processing phases can be automated for achieving the near real-time requirement around the clock. One of the goals of Envimon is to build a prototype of a common framework called EOFrame that facilitates the development of automated EO data processing chains, and onto which the applications and services can be built. The framework architecture utilizes web services, an XML-based workflow language BPEL (Business Process Execution Language, [1]) and workflow engines for developing and running automatic data processing chains.

Web service and workflow technologies have been applied in earth observation earlier as well. The Multiple Application Support Service (MASS) environment aimed at prototyping solutions for an open service-oriented and distributed environment for service users and service providers of EO, meteorological and GIS data [2]. The project's results are utilized in the Service Support Environment (SSE) developed for the Ground Segment Department at ESA-ESRIN [3]. SSE facilitates service chaining, i.e. the creation of new services from a horizontal set of basic services supplied by multiple service providers. A user of the SSE portal can place an order for a data product he is interested in, and the data is delivered either synchronously when the service returns a result immediately or asynchronously if the processing takes time, and the product must be delivered afterwards. The core of the SSE Portal is the workflow engine that executes business processes based on BPEL. Our approach has similarities to SSE, but as SSE facilitates existing services to be accessible by a variety of users, our framework facilitates application and service development by utilizing web service and BPEL techniques. Applications developed with EOFrame can in principle be registered as services within SSE. An example of a framework approach for processing earth observation data in a distributed manner is described in [4]. The authors present a framework for building flexible and scalable processing pipelines that include preprocessing, processing and automated result analysis as independent modules. The framework gives the flexibility to add and remove modules on the fly, as well as re-use existing code. The approach utilizes Java RMI and dedicated software for managing the workflow. Our approach differs in the sense that instead of proprietary interfaces and workflow engines we use standard components like web services for component interfaces, and BPEL and 3rd party workflow engines for workflow control.

In this paper we describe the requirements of the framework, the system architecture of our design, and analyze the results of using the framework in building the applications.

II. REQUIREMENTS FOR A COMMON FRAMEWORK

The Envimon applications process earth observation data and some of them use the same data sources from the same satellite instruments. Therefore, already in the early stages of the project it was clear that the application processing chains share some common functionality that should be implemented only once as a common framework for all of the applications.

Common EO data sources require common preprocessing steps such as geometric and radiometric rectification, geo-coding of satellite images and mosaicing of images into composite images. Near real-time applications need automated data retrieval so that new data is available for the processing whenever they are available. There are different kinds of processing that the applications use for analyzing the preprocessed files. After the analysis phase the data are delivered to systems or users that will do some additional processing or take actions based on the data. The delivery part provides the users with the processing results tailored to their specific needs, for example what data format they accept, what are the relevant parameters they want to retrieve, and what spatial and temporal areas they are interested in.

In automatic chains, the input, preprocessing, analysis and delivery must be controlled by an intelligent manager mechanism that notices stuck processing services, recovers from processing failures, collects log information about processing status, and informs administrators about problems it cannot solve by itself. The processing chain from data input to delivery consists of separate processing modules and a workflow control mechanism is required for enabling description and execution of application-specific data processing workflow. A workflow control calls data processing modules to accomplish tasks defined in a workflow specification. The workflow execution can be done parallelly, so that service tasks can be executed on separate remote computers, and the workflow control handles the communication between distributed processes. Some processing modules may support multiprocessing, but this is out of scope of workflow control's requirements. The workflow may contain conditional branches, i.e. the execution of a task will take place if certain conditions are fulfilled. For example, if the result of a pre-analysis is interesting, and can be automatically tested, the workflow execution might call an optional data input source to retrieve additional data for more accurate post-analysis. These requirements comply with common workflow patterns that are discussed in more detail e.g. in [5].

A data storage service is required for storing and querying for data files in different processing stages. For example, whenever an application starts searching for new data or downloading a large file it might check whether the new data file is already downloaded and preprocessed by the application itself or by some other application. Also, an application may need a result file from the latest processing round in the analysis phase of the current round for comparison purposes.

The required system is depicted in Fig. 1. The framework has a common pool of processing modules that retrieve different data from different data sources, preprocess and analyze data, and deliver them to different users. Processing modules constitute an operative processing chain defined by an application developer. A workflow control manages the processing chain execution, and a data storage stores and retrieves data products during the processing.

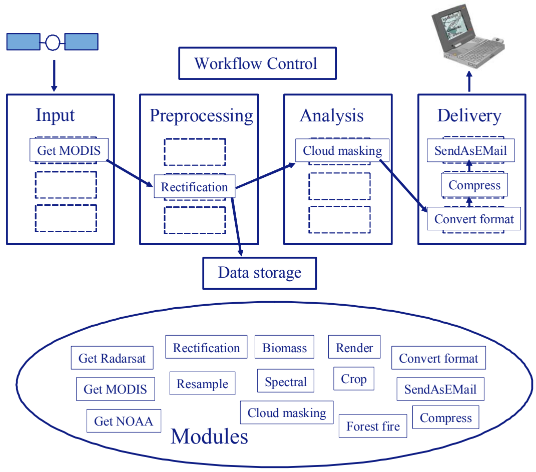

Figure 1. The required framework for environment monitoring applications.

The purpose of a framework is to make implementing target applications more efficient. In order to understand how the framework can be utilized in application development EOFrame's relationship to an application and different user roles are illustrated in Fig 2.

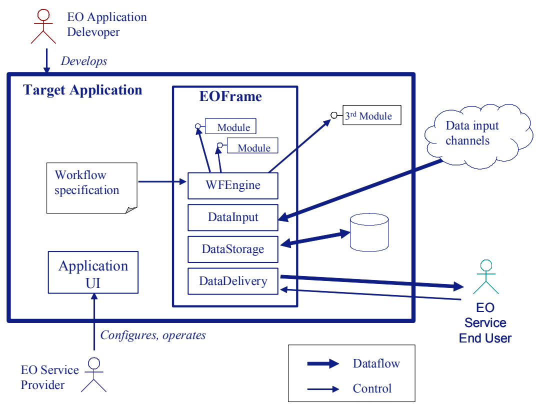

Figure 2. The relationships between EOFrame, an application, and different user roles.

An EO Application Developer implements the target application by coding the application logic and the user interface. The developer uses EOFrame components for data retrieval, processing and delivery. With the workflow engine the developer specifies a workflow of the application specific processing that may include calls to individual processing modules and predefined workflow processes provided by

EOFrame Module Library, or to some in-house or 3rd party processing modules. Before modules can be called by the workflow engine, they must be interfaced by implementing an interface, specified by EOFrame. An EO Service Provider configures and operates the target application to serve its customers. A service provider is not expected to be aware of EOFrame's existence in the application system at all. An EO Service End User orders and retrieves data from the application by way of the service provider. Without having to know anything about EOFrame or even the target application, the user may be unknowingly using EOFrame's data delivery service. End user is concerned about data quality, reliability and accuracy of the service, and not about how the system is built.

III. FRAMEWORK ARCHITECTURE

The requirement for a manageable workflow control and a possibility to add new processing modules to the processing chains suggested that we should look into workflow languages and engines when designing the overall architecture of the framework. Workflow management systems have been around since the late 1970s, but every now and then they are rediscovered in a marketing wave such as web services choreography as the latest [6]. SOA (Service-Oriented Architectures) with web service orchestration is currently among the hottest topics within IT [7]. BPEL seems to be becoming a de facto standard for workflow languages. There are many commercial and open source workflow engines available for BPEL, like Service Orchestrator from OpenStorm, Composite Application Builder from Vergil, Oracle's BPEL Process Manager, Process eXecution Engine (PXE) from FiveSight, ActiveBPEL from ActiveBPEL LLC, and Twister from Smartcomps. We realized that the field is not mature yet, but the engines are still developing and many bugs are waiting for fixes. We chose to use Oracle's BPEL Process Manager for editing and running BPEL processes, as it was freely available and appeared to be the most robust of the tested engines. However, in the beginning there were some problems in running even a simple BPEL process, but the problems disappeared as new versions became available. The framework architecture is not dependent on a specific workflow engine, so it is possible to compare different engines later on.

The architecture is depicted in Fig 3. It contains a workflow engine that executes workflow descriptions defined with a workflow editor. Processing modules are plugged in by defining a web service interface for them, and so they can be located anywhere in the Web. Envimon module library contains the WSDL (Web Service Description Language, [8]) definitions of all the processing modules needed by Envimon applications. A workflow engine takes care of calling the services defined in workflow descriptions. A workflow description that is published in the workflow engine can be accessed as a web service from any other application, like from an Envimon application for instance. A workflow process can also be part of some other workflow description allowing a hierarchical composition of workflows. When considering the whole data processing chain from input to delivery it appears that instead of a single workflow there could be a multitude of workflow definitions for a single application. One workflow would be too complex if it should take care of all the possibilities for tailoring different products for different data users. For instance, there could be a workflow process that takes care of bulk data production, and another that tailors and filters bulk data files according to different user needs. If there is more than one workflow definition taking care of all the processing needed by a single application a higher level of control is required. This new level of control can in principal be implemented as a top-level workflow description that controls the other lower-level workflows, or it can be an external piece of software that calls the workflow processes when required.

The delivery requirement is implemented as a separate software component called Façade that, based on the user profile definitions, tailors and delivers analyzed products to end-users. Façade has access to user profile data, and knows what kinds of data products are wanted and how they should be processed. Façade calls the workflow processes on a workflow engine to do the actual processing job, and in this sense Façade acts as a top-level control for a workflow engine and a multitude of workflows. Users manage their data service orders via Profile UI, and Façade takes care that the final products are delivered to them as defined in the profile database. The façade concept as part of data processing chains was originally introduced in [9].

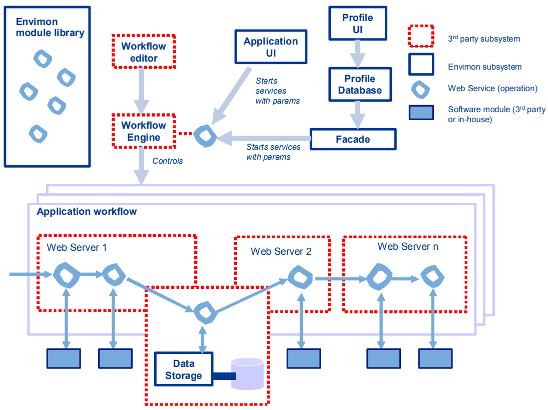

Figure 3. The EOFrame architecture.

Data storage component is used to store processed spatio-temporal data for later use. It does not necessarily store the files themselves, but the metadata describing the details of a file and information about from where it can be downloaded. Data storage implements its storing and querying services as web services, and hence the storing and retrieving of files can be included in the workflow. One of the challenges of processing large files is to transfer data between processing modules located on separate machines. The framework contains web services for publishing data files on a web or ftp server, and downloading files from a given location using HTTP and FTP. These services can be included in the workflow if the processing is distributed between separate machines and data files must be transferred.

processing environment, and how the robustness and reliability of the automatic data processing chain are ensured. Autonomic Computing ([10]) is emerging as a new approach to the design of complex distributed computer systems, and its relevance to automatic EO data processing chains should be studied.

Although the project concentrates on earth observation, the concept is more general in nature. The architecture could be utilized in diverse applications, whose key characteristics include workflow control, data storage, data intensive processing, distribution of processing, subscriber profile and domain specific processing modules. Examples of prospect application areas are: industry (procurement, production monitoring), retail and banking (customer relationship management and marketing), telecommunications (network analysis and management), health care (medical image analysis and prognoses), business management (business intelligence) and science (bioinformatics, space research).

IV. Applying the Framework in Application Development

The feasibility of the concept introduced by the software framework is verified by implementing several environment monitoring systems for the diverse real-world applications. However, it is difficult to develop a framework at the same time as the applications that are based on it are developed. Finding common functionality requires that you already have some functionality implemented in the applications. On the other hand, applications have their own deadlines and requirements that must be followed and cannot wait a framework to be ready to provide some of the functionality. Therefore, we had to implement applications parallelly with the framework development, and the framework was included in the application development not until in the late stages. Some applications could not utilize the framework, as they did not have a need for near real-time processing after all.

One of the applications with a near real-time requirement is the maritime application in which MODIS images are delivered to ships navigating in ice-infested sea areas. MODIS data can be considered as one of the data sources that help merchant vessels to navigate in ice and icebreakers to coordinate their assistance services. The data gets old very quickly, especially if the ice field is under movement, and they must be delivered to the ships as fast as possible. Also, the size of the data files delivered must be minimized because of the low-bandwidth data communication link to the ships. File sizes are reduced by tailoring data products, i.e. by cropping and resampling the images for different sea areas. The maritime application will be implemented using EOFrame. First, the data processing functionality required for processing MODIS data was implemented as separate software modules written in C. Then the modules were implemented as web services, and a BPEL description was written for enabling the workflow. Next we will include the maritime service in Façade's profile database, which enables users to order MODIS images tailored for their areas of interest. After doing this, the workflow of the application service should be manageable and inclusion of new processing in the workflow should be easier. For instance, an automatic cloud masking module to detect useless images is under consideration.

V. Conclusions

We are developing a prototype of a framework that facilitates building of automated and distributed data processing chains for large data volumes. The system is under construction, and more research and development work is still required before the prototype can really help application development. In the implementation we apply BPEL for describing workflows and workflow engines for executing them. We can imagine that in the future we see more and more individual and chained EO and GIS services available in the Web. When different processing tasks are distributed among different service providers and organizations a whole lot of new research challenges emerge. For instance how to control billing, quality of processing, and data security in a distributed

Acknowledgment

Envimon is a research project funded by Tekes (Finnish Funding Agency for Technology and Innovation) and run by VTT Information Technology. The partners are Stora Enso Inc., PIEneering Inc., Savcor Indufor Inc., City of Helsinki/City Planning Department (KSV), Helsinki Metropolitan Area Council (YTV), Finnish Radiation and Nuclear Safety Authority (STUK), National Research Institute for Earth Science and Disaster Prevention (NIED, Japan).

References

[1] T. Andrews, F. Curbera, H. Dholakia, et.al, Business Process Execution Language for Web Services, Version 1.1, 2003, http://www-106.ibm.com/developerworks/webservices/library/ws-bpel/

[2] Y. Coene, C. Heylen and P.G. Marchetti. An Environment Supporting Earth Observation and GIS Service Orchestration: The Mass Project, IEEE Geoscience and Remote Sensing Newsletter, Issue 129, ISSN 0161-7869, page 19-22, December 2003.

[3] Y. Coene, C. Bawin, Service Support Environment Architecture, Model and Standards, white paper, December 2004. http://earth.esa.int/rtd/Documents/SSE_Whitepaper_2.pdf

[4] P. Votava, R. Nemani, C. Bowker,A. Michaelis, A. Neuschwander, J. Coughlan, Distributed application framework for Earth Science data processing, International Geoscience and Remote Sensing Symposium (IGARSS), v 2, 2002, pp. 717-719

[5] W.M.P. van der Aalst, A.H.M. ter Hofstede, B. Kiepuszewski, and A.P. Barros, Workflow Patterns., Distributed and Parallel Databases, 14(3), pp. 5-51, July 2003.

[6] M. zur Muehlen, Workflow & BPM Research web site, 2004, http://www.workflow-research.de/

[7] J. Rosenberg, A. Mateos. Autonomic SOA Web Services - Achieving Fully "Business-Conscious" IT Systems. Web Services Journal. 2005. http://webservices.sys-con.com/read/136205_p.htm

[8] W3C: web services activity. 2006. http://www.w3.org/2002/ws/

[9] V. Kotovirta, R. Berglund., J. Haajanen, J. Mansner., R. Tergujeff, Delivering Near Real-Time Met-Ice-Ocean Observation and Forecast Data – the IWICOS Façade, Proceedings of 3rd EuroGOOS conference, 2003.

[10] R. Sterritt. Autonomic Computing. Innovations in Systems and Software Engineering, Vol. 1, No. 1, Springer, ISSN 1614-5046, pp. 79-88, 2005

Publication II

II

Berglund R., Kotovirta V., Seinä A. 2007. A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea. Canadian Journal of Remote Sensing 33(5): 378-387.

Can. J. Remote Sensing, Vol. 33, No. 5, pp. 378–387, 2007

Technical note / Note technique

A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea

Robin Berglund, Ville Kotovirta, and Ari Seinä

Abstract. The two most heavily navigated waterways in the world, where seasonal sea ice plays an important role in navigation, are the Gulf of St. Lawrence in Canada and the Baltic Sea in Europe. Spaceborne synthetic aperture radar (SAR) has been used operationally by the icebreakers in the Baltic Sea since 1992. Today, both RADARSAT and Envisat advanced synthetic aperture radar (ASAR) data are delivered to end-users. The captain and mates on Finnish and Swedish icebreakers interpret the images using a work station called IBPlott developed by the VTT Technical Research Centre of Finland. The geographical information system (GIS) type of work station combines and displays all available relevant information required for making routing and ship assistance decisions. The information available includes satellite images, ice charts, positions and destinations of the ships moving in the area, visualization of current speed and tracks (from the Automatic Identification System (AIS)), ice forecasts, water level, and weather forecasts and the most recent weather observations. In this paper we present the operational system in use for making SAR images available to the end-users in near real time in the Baltic Sea, combining the satellite images with other information products. The end-user experiences of the system and future development ideas are presented briefly.

Résumé. Les deux voies maritimes les plus empruntées au monde et dans lesquelles la glace saisonnière joue un rôle important pour la navigation sont le Golfe du fleuve Saint-Laurent au Canada et la Mer Baltique en Europe. L'imagerie spatiale radar à synthèse d'ouverture (RSO) est utilisée de façon opérationnelle pour la mer Baltique depuis 1992. De nos jours, des données RADARSAT et Envisat RSOA (« radar à synthèse d'ouverture avancée ») sont fournies aux utilisateurs sur le terrain. A bord des brise-glaces finlandais et suédois, le capitaine et l'équipage interprètent les images en utilisant la station de travail IBPlott, développée par VTT. Ce logiciel de type SIG combine et affiche toutes les informations nécessaires aux décisions concernant le routage maritime ou l'assistance à d'autres bateaux. L'information disponible inclut des images satellite, des cartes de glaces, les positions et destinations des bateaux navigant dans la zone, la visualisation de la vitesse et de la direction actuelles (par le système AIS (« Automatic Indentification System »), le niveau d'eau, des prévisions de conditions de glace et météorologiques, ainsi que les plus récentes observations. Dans cet article, nous présentons un système opérationnel qui rend les images RSO disponibles aux utilisateurs dans la Mer Baltique quasiment en temps-réel, conbinant les images satellite avec d'autres informations. Les expériences du système par les utilisateurs terrain et les idées pour de futurs développements sont brièvement exposées.

Introduction

The two most heavily navigated waterways in the world, where seasonal sea ice plays an important role in navigation, are the Gulf of St. Lawrence in Canada and the Baltic Sea in Europe. In the Baltic Sea the marine traffic has increased by 34% during the last 10 years, and the trend is expected to continue. The total cargo turnover in the Baltic Sea in 2003 was about 731 million tonnes, and it is expected to grow up to 1.2 billion tonnes by 2020 (Swedish Maritime Administration, 2006). Some 40% of marine transportation takes place during winter (from December to April). According to conservative predictions, growth of oil transportation from Russia to western Europe is expected to reach 200–250 million tonnes by 2015.

Parts of the Baltic Sea surface freeze every winter. The maximum annual ice extent occurs between January and March, when ice covers 52 000 – 420 000 km^2, with an average of 218 000 km^2. The ice season lasts from a few weeks in the south up to 7 months in the north. The ice occurs as fast ice that exists in coastal archipelago areas and drift ice that has a dynamic nature, being forced by winds and currents. Ice ridges

Received 22 December 2006. Accepted 4 July 2007. Published on the *Canadian Journal of Remote Sensing* Web site at http://pubs.nrc-cnrc.gc.ca/cjrs on 19 October 2007.

R. Berglund[1] and V. Kotovirta. VTT Technical Research Centre of Finland, PO Box 1000, FI-02044 VTT, Finland.

A. Seinä. Finnish Institute of Marine Research, PO Box 2, FI-00561 Helsinki, Finland.

[1]Corresponding author (e-mail: robin.berglund@vtt.fi).

and brash-ice barriers are the most significant obstructions to navigation in the Baltic Sea. Powerful, ice-strengthened vessels can break through thick level ice, but they are not capable of navigating through ridges and heavy brash-ice barriers without icebreaker assistance. Ice dynamics creates high pressure in the ice field and can be dangerous to vessels and cause time delays of up to several days. Under normal conditions the sailing time from the ice edge to the northernmost ports of the Baltic Sea is 1 day (400 nautical miles), but under severe conditions it can extend to 1 week.

Winter navigation is made possible by the use of icebreakers and ice-strengthened vessels and by restricting navigation. The national maritime authorities restrict navigation by requiring a certain minimum ice class and size of the vessel for it to be assisted. These restrictions are determined per harbor based on the current and forecasted ice conditions. Also, up to half of the smaller harbors may be closed for the winter. The icebreakers need detailed ice information for route planning. The smoothness of traffic has been possible owing to better ice monitoring, where use of earth observation (EO) data has become more important. Considerable savings in ice navigation could be made by optimizing the use of satellite-based operational ice monitoring (Grönvall and Seinä, 2002; Seinä et al., 2006). Spaceborne synthetic aperture radar (SAR) has been used operationally by the icebreakers in the Baltic Sea since 1992 (Vainio et al., 2000). The first trials used European Remote Sensing satellite 1 (ERS-1) data, and since 1998 Finland and Sweden have jointly acquired 100–150 RADARSAT scenes per ice season for ice monitoring and icebreaker operations.

Near-real-time delivery of satellite images to ice-going ships has been done for several years in different ice-covered sea areas like the Arctic (Pettersson et al., 2000; Smirnov, 2005), the Antarctic (Danduran et al., 1997; Toudal Pedersen and Saldo, 2004), and the Baltic Sea (Håkansson et al., 1995; Vainio et al., 2000). The client systems using satellite images on board range from simple image display applications suitable for images that have a coordinate grid "burned" into the image to tailored geographical information system (GIS) clients that have advanced overlay capabilities (e.g., the IceNav system developed by a Canadian company called Enfotec). One possibility is to include weather and ice information in the on-board electronic chart display and information system (ECDIS) (Chang, 2001; Diarbakerly et al., 2002).

National ice services of the Baltic Sea are responsible for collecting, analyzing, and distributing sea ice data as stated in Helsinki Commission (HELCOM) recommendation 25/7. At the moment there are nine ice services in the Baltic Sea publishing products with different quality levels. In Finland, ice service responsibilities are defined in the Act for the Finnish Institute of Marine Research and annual corollaries of the Ministry of Traffic and Communication. The service is financed mostly by government and partly by users. For a survey of other ice information providers, see Sandven et al. (2005) and Smirnov (2005).

In this paper we describe the ice navigation support system developed for icebreakers operating in the Baltic Sea and show how satellite data are utilized together with other observational and forecast data in ice navigation. The system is used operationally during the winter. It represents a GIS system that provides the user with satellite images and other observational and forecast data and is used as a navigation support system, although it is not ECDIS compatible. The system consists of an on-board end-user system called IBPlott developed by the VTT Technical Research Centre of Finland and data processing and delivery systems on the server side ashore developed and run by the Finnish Institute of Marine Research (FIMR) and VTT.

Requirements for an on-board ice navigation support system for icebreakers in the Baltic

Ships in the Baltic Sea have to contact the icebreakers when approaching the ice edge to obtain advice on waypoints to follow. For their operational and tactical navigation and assistance planning, the icebreakers require information about prevailing and forecasted ice conditions and about the traffic situation. It is also important that the information be presented near real time in an intuitive and easy-to-use way, which sets requirements for the system.

The general requirements for an on-board ice navigation support system for icebreakers operating in the Baltic Sea are presented in this section. The requirements are based on discussions and interviews with end-users, some of which were presented earlier by Seinä et al. (2006).

End-user requirements for ice information

This chapter presents the end-user requirements for ice information, i.e., information that describes the prevailing and predicted ice conditions. Ice information is needed to plan safe and cost-efficient routes through the ice field. Information about cracks, leads, ridged ice fields, ice concentration, and ice compression helps captains avoid areas of harsh ice conditions. If the icebreakers have good knowledge of the ice conditions, how the conditions are likely to evolve, and how individual ships perform in ice, there is enough information to define the waypoints in such a way that both the risks for the ship to be damaged and the total time the ships have to wait for icebreaker assistance are minimized.

Users require both optical and radar satellite images. SAR images are especially useful because their resolution ranges from 100 to 400 m and they are daylight independent and almost weather independent. Present satellite sensors are sensitive to the properties of the ice field surface and lack direct information about the ice thickness. The new satellite CryoSat-2, to be launched in 2009, will try to measure the ice thickness based on radar altimetry, i.e., measuring the freeboard of the ice floating on water (Drinkwater and Rebhan, 2003). The spatial and temporal resolution, however, will be limited, and therefore the products are probably not suitable for near-real-time operations. Therefore, satellite images are not enough to determine the properties of the ice field.

Observations about the ice thickness, which is compiled in the ice charts, are needed to help interpret the information in the satellite images. The input data of ice charts consist of ground truth, visual and digital airborne data, and spaceborne data from various satellites. The quality of the charts is highly dependent on the use and availability of EO data.

Accurate ice drift estimates require information about sea currents. In shallow basins like the Baltic Sea, water-level variations cause sea currents that affect ice motion, and therefore water-level observations are required. Water-level variations are caused mainly by atmospheric pressure and wind, as the amplitude of the tide in the Baltic Sea is negligible.

Satellite images and other observational data are not enough for route planning purposes, as they describe how the situation used to be, and plans are about the future. Although the systems are tuned to deliver the data to the end-users as fast as possible, observations are still always in the past. When the ice field is changing slowly, observations can be used for nowcasting, but observational data become outdated in a matter of hours when the ice field is moving. Therefore, proper forecasts are needed to support route planning. Forecasts cover a period from days to weeks ahead, but for users at sea a period of 24–54 h is mostly suitable.

Weather forecasts are the primary required forecast information on board a vessel. Experienced captains combine the latest wind forecast and water-level observations with their knowledge of prevailing ice conditions to extrapolate the current situation to the future. As a rule of thumb, the ice drift in the Baltic Sea is mainly wind forced, with the drift direction being offset by 20° compared with the wind direction because of the Coriolis force (Leppäranta and Yan, 1991). In addition to water-level observations, water-level predictions are also required for determining sea currents and their effect on ice drift.

Ice model data should be presented together with other information. Ice models predict the future ice conditions using weather forecasts as an input. They give statistical values in grids with a density of a few kilometres about the ice field properties, like thickness, ridging density, and ice drift. For a trained user, these parameters give additional information about the future ice conditions, especially when the model prediction correlates with the intuitive prediction made by the captain.

The resolution of the data products should be at the scale of the ship, i.e., around 100–200 m. The schedule of delivering the products and the latency between acquisition and delivery should be adapted to the user needs. For near-real-time users (1–6 h) the service should be available 24 h/day.

End-user requirements for auxiliary information

Ice information is not enough for icebreaker operations. The assistance planning also needs information about the ship traffic situation and scheduled departures and arrivals. To be able to assess the assistance need, basic information about ship parameters (type of ship, length, breadth, machine power, ice class, etc.) is also needed. For evaluating the performance of the ice-breaking service level, assistance statistics, including waiting times, are required.

Overall requirements for the system

There is a multitude of information sources available, and the amount of raw data produced by even a single data source is too large to be transferred to the ships as such. For example, the output of an ice model run is in the order of gigabytes. Therefore, data processing and filtering on the server side are required. Ships operate in different geographical sea areas and do not necessarily need the same data parameters. Data processing chains on the server side must retrieve the raw data, extract relevant information, tailor the information for different sea areas and users, and deliver the information to the end-users over a mobile communication link.

The real-time aspect in delivering the products is important because the ice conditions change over time. To provide the users with the latest available data, the processing chains must operate as automatically as possible. Whenever raw data are available, the system should start processing and delivering the processed products automatically without human intervention.

As there are multiple data providers, users may need to contact different organizations to obtain the data they need. The system should provide the users with data only via one or two services (one-stop shop), or at least the multitude of sources should cause no extra trouble for end-users. However, end-users need to be able to identify the data provider as they compare different products from different sources or when they want to assess the validity of the products.

The client application should present all the time-varying observational and forecast data parameters to the end-user in an understandable, intuitive, and easy-to-use way. An intuitive user interface is especially important for the icebreakers in the Baltic Sea, as they do not have a trained ice specialist on board. The end-user system should present a mosaic of satellite images from different areas and times together with other observational data and combine the observations with forecasts that are describing the same sea areas but at different times.

System used in the Baltic Sea

This section describes an operative system used within the Baltic Sea that fulfills some of the requirements defined in the previous section. The system can be considered as a distributed system that consists of an on-board end-user system called IceBreaker Plotter (IBPlott) and data processing and delivery systems on the server side ashore. The system is used in the Baltic Sea by Finnish and Swedish icebreakers for navigation and assistance activity planning.

End-user system

IBPlott is part of the IceBreaker Network (IBNet) system that is used to coordinate and communicate the daily assistance activities of the icebreakers. IBNet has a distributed database that contains information about the activities of the whole icebreaker fleet, and every icebreaker has a database node on board that is replicated to other icebreakers. In addition, IBNet provides data to produce statistics for evaluation of ice-breaking service performance (available from www.vtt.fi/palvelut/cluster1/topic1_5/ibnet.jsp?lang=en [accessed 21 May 2007]).

IBPlott is the geographical user interface of the IBNet system. It displays geolocated information on the screen as symbols and images and provides mechanisms for users to interact with the objects in the database. In addition to presenting relevant traffic-related information, IBPlott also displays geographical observational and forecast data to the user. Observations and forecasts about the prevailing conditions are a necessary addition to the information captains receive when they navigate in a dynamic and ever-changing environment.

Satellite images

IBPlott presents satellite images on the screen in geographically correct locations. It is possible to open several images at the same time as an image mosaic. This makes it easier to compare images from different sensors, in different resolutions, and from different times (**Figure 1**).

Ice charts and ice thickness charts can be opened as part of the image mosaics. They help captains estimate the thickness of the ice field. The drawing order of the images can be changed and one of the images can be set in a so-called "X-ray mode", in which the image is drawn only in a small window at the location of the mouse cursor. In **Figure 2**, a RADARSAT image is set in X-ray mode and drawn as a disc at the cursor location on top of a moderate resolution imaging spectroradiometer (MODIS) image. When the cursor is moved on the screen the user gets a feeling of seeing through the topmost MODIS image where the mouse cursor is located. This makes it possible to compare different images more easily.

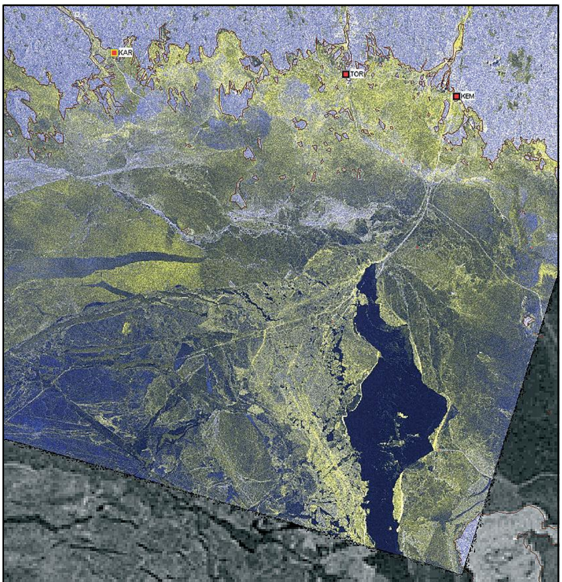

Figure 1. An image mosaic consisting of an Envisat false-color dual-polarized image at 100 m resolution presented on top of a RADARSAT image subsampled to 400 m resolution.

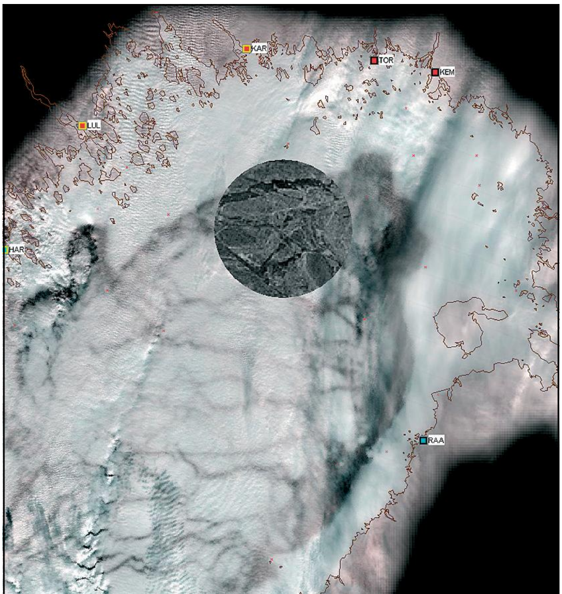

Figure 2. X-ray mode example. See text for explanation.

AIS trails

IBPlott is connected to the Automatic Identification System (AIS), and it has access to the latest position and speed observations and some history of the locations of all ships navigating in the Baltic Sea. This makes it possible to present trails of the ships colored as a function of the ship speed (**Figure 3**). Assuming that the ship speed correlates with the ice resistance, the colored trails drawn on the map help the user to estimate the navigational effect of the ice conditions within different areas, especially when the information is drawn on a satellite image, and the time of trails is adjusted to the time when the image was taken.

Weather information

IBPlott presents weather observations and forecasts on top of satellite images (**Figure 4**). The forecasted wind field is drawn on the map as wind barbs, and isobars show estimated paths of oncoming low pressure or storm centres. Wind observations tell about the current weather conditions elsewhere and help estimate the accuracy of the wind forecast, for example, the path and magnitude of a predicted storm. Water-level observations and predictions are used in determining sea currents. Time series presentation shows the weather and other parameters as a function of time at a selected point on the map or along a predetermined route.

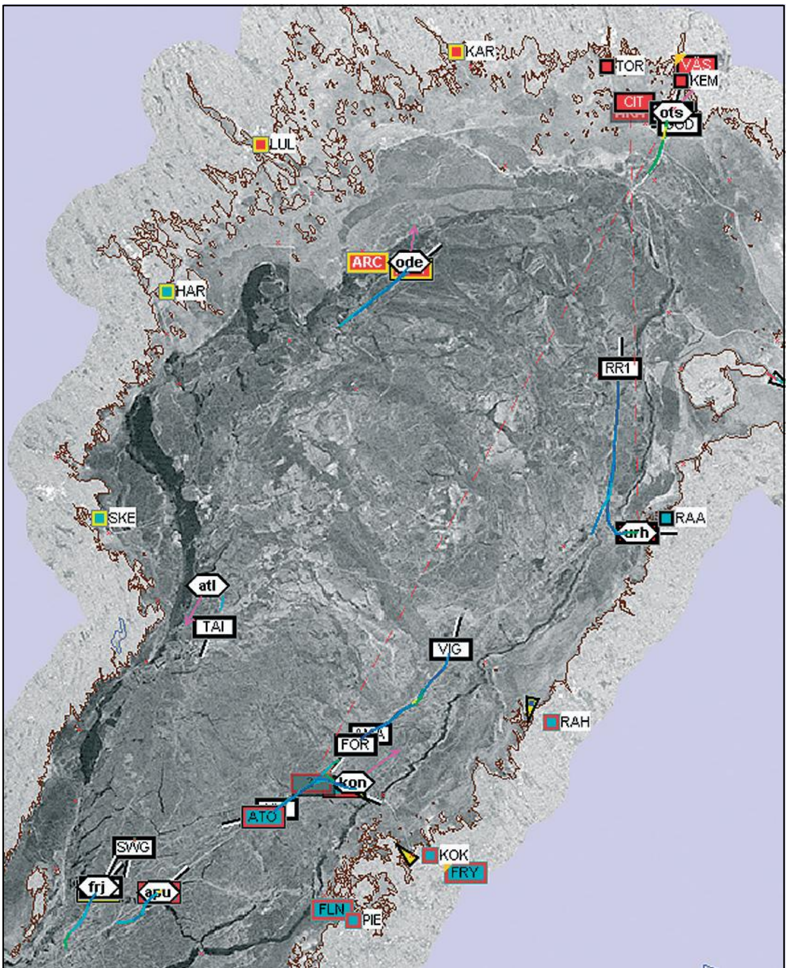

Figure 3. IBPlott presenting current traffic situation on a RADARSAT image. Colored lines attached to ship symbols (hexagons and rectangles) are the trails of the ships in a 1 h history, and colors indicate the speeds that the ships traveled.

Ice model forecasts

The newest feature to be used and tested during the winter 2006–2007 was to present ice forecasts on top of satellite images. A numerical ice model output grid is rendered as an image on the screen, with pixel colors determined by the values of the grid parameter the user has selected to be presented. Level and ridged ice thicknesses follow the color coding defined by the thickness chart produced by FIMR (Karvonen et al., 2003). The color coding used for other parameters is still to be adjusted based on user feedback. The transparency of the ice model presentation can be adjusted so that a satellite image in the background is mixed with the ice forecast colors. Ice drift is presented as arrows together with the divergence calculated from the ice drift vector field. The divergence indicates the local rate of change of the area in any location in the ice field. By visualizing the divergence, areas of compression (convergence) and decompression (divergence) are highlighted (**Figure 5**).

One way to present the multidimensional ice model data (where time can be considered as one of the dimensions) is to reduce the dimensions in the data presentation by calculating how the ship would perform in the predicted ice conditions. Ice model output, ship transit models, and an optimization framework were integrated into an end-user system prototype that finds an optimal route through the ice field given an ice model forecast (Berglund, 2006). A transit model predicts how a ship performs in predicted ice conditions, and the optimization framework calculates what variations of the initial route give the shortest transit times from one port to another.

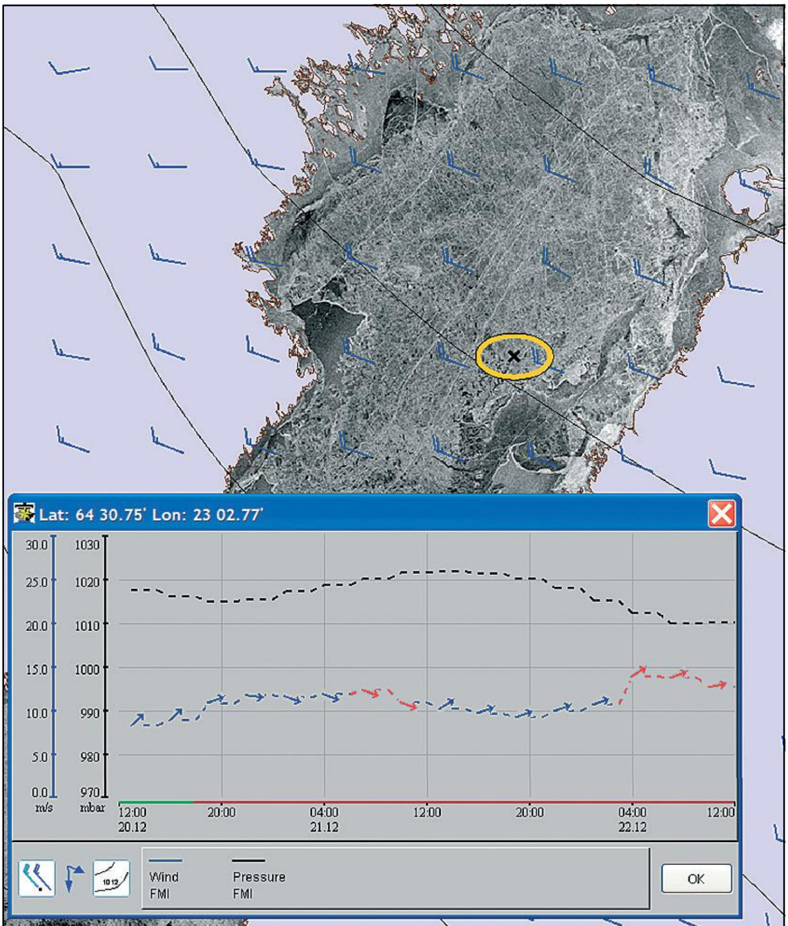

Figure 4. IBPlott presenting wind and pressure forecast on a RADARSAT image. A time-series presentation of wind and pressure at a selected point (×) is also shown.

The prototyped optimization features could be useful for icebreakers with experienced crew, not just for finding their own way in the ice, but to estimate how other ships they are assisting would perform in the ice between the waypoints icebreakers have ordered them to follow.

User interface

IBPlott has been designed to be intuitive and easy to use for ship personnel that are not computer and remote sensing experts. Whenever new data are delivered, IBPlott loads them automatically, and the latest observations and forecasts are always ready to be viewed. The user can easily select which data parameter layers and images are shown on the screen. Time-dependent data like weather data and ice model predictions are one challenge for the user interface (as the problem of handling time is introduced). The view on the screen in IBPlott is always looking at some time period, and all data products whose time extent intersects with the time window of the view are drawn on the map. Although satellite images also are valid for a specific time (i.e., the acquisition time), they are drawn on the screen regardless of the time the user has selected. Images are considered as background material on which forecast data are presented because in this way a user can more easily compare the satellite image with the forecasted weather conditions and predict the coming ice conditions. The user can browse back and forth in time using a time slider tool, and the situation on the screen is rendered accordingly.

Data processing and delivery

The aim of the system design is to provide the end-users with relevant information from a multitude of data sources in near

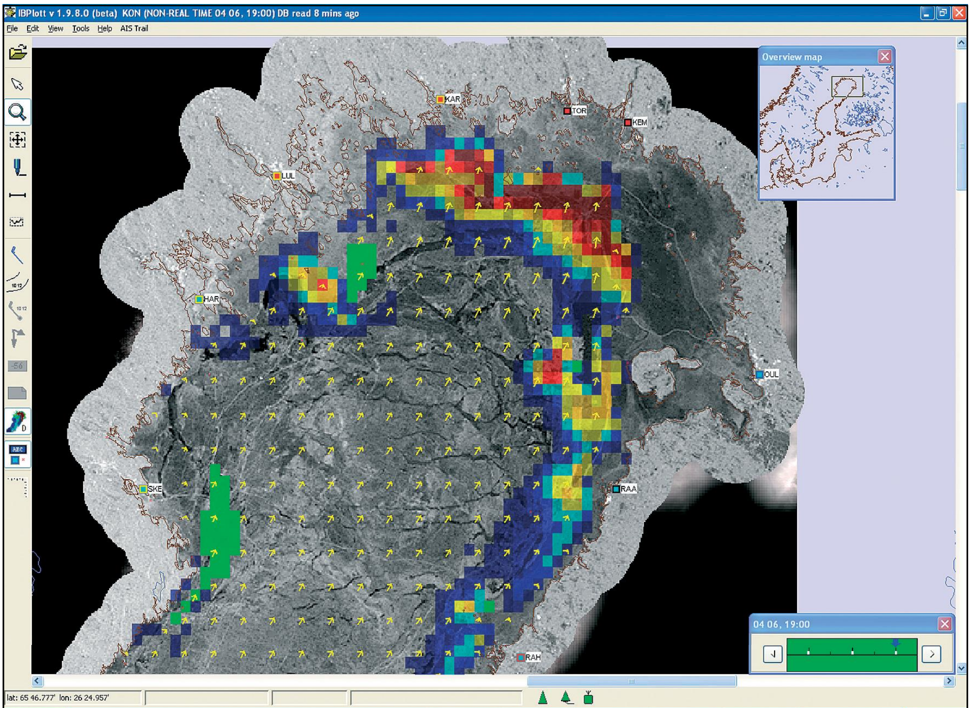

Figure 5. IBPlott presenting ice drift and the divergence on top of a RADARSAT image. The red color indicates that the ice is accumulating in the northern parts of the Bay of Bothnia, thus correlating with compression in the area, and the green color indicates divergence of the ice field near the coast of Sweden.

real time. The number of data products and providers is expected to increase when new satellites provide more images and also new imaging modes (such as dual polarization). Therefore, the data processing and delivery subsystems are an important part of the ice navigation support system. **Table 1** gives an example of currently used data products and their average data volumes per day. The amount of data delivered is optimized for current data transmission systems. As the systems develop, higher resolution products can be delivered with reasonable costs and without jamming the data transmission line. The amount can be adjusted simply by reducing the number of products delivered or by reducing the spatial or temporal resolution of images and model output grids.

The raw RADARSAT data are received and processed in Tromsö, Norway, by Kongsberg Satellite Services (KSAT), and the image data are then transferred using file transfer protocol (FTP) and the Internet to the Finnish Ice Service in Helsinki (**Figure 6**). Processing of the images consists of cutting the areas of interest from the remapped image, optional scaling of the image to eight bits (not needed for ScanSAR wide images), and then compressing it before sending to the icebreakers. It takes less than 2 h from overflight of a satellite to presenting a satellite image on the screen on board ships. The images transmitted are compressed to a few hundred kilobytes to save transmission bandwidth. During the last 2 years, Envisat advanced synthetic aperture radar (ASAR) images have also been used, and in the winter season 2005–2006 the ASAR images have been transferred from the European Space Agency (ESA) rolling archive at the ESA Centre for Earth Observation (ESRIN) in Italy. The use of MODIS images was prototyped in the winter 2005–2006, and MODIS images were delivered to operational use during the winter season 2006–2007.

During the ice season 2005–2006, 163 RADARSAT-1 and 82 Envisat ASAR images were processed and delivered to the Finnish icebreakers, and 60% of these were also delivered to the Swedish icebreakers. On average, 0.7 images per day were delivered during the most intensive month, February 2006. During the ice season 2006–2007, 143 RADARSAT-1 ScanSAR wide images and 72 Envisat ASAR images were delivered to the Finnish icebreakers.

Weather files are delivered by the Finnish Meteorological Institute (FMI) to the Finnish icebreakers and by the Swedish Meteorological and Hydrological Institute (SMHI) to the Swedish icebreakers. Ice model data coming from the FIMR and SMHI are delivered to end-users by the VTT Technical Research Centre of Finland using a system prototype called Façade. Facade reduces the amount of data in the ice model outputs by selecting the required data parameters from a larger

Table 1. Average data volumes per day per data product sent to the icebreakers during February 2006.

	No. of products per day	No. of kilobytes per product	Total no. of kilobytes
SAR	0.7	250	175
MODIS	0.3	350	105
High-resolution weather forecast	4.0	80	320
Weather observations	24.0	1.5	36
Textual weather forecast	4.0	1	4
Water-level observations and forecast	24.0	1.5	36
Ice chart	1.0	150	150
Thickness chart	0.7	50	35
Ice model forecast	1.0	350	350
Total			1211

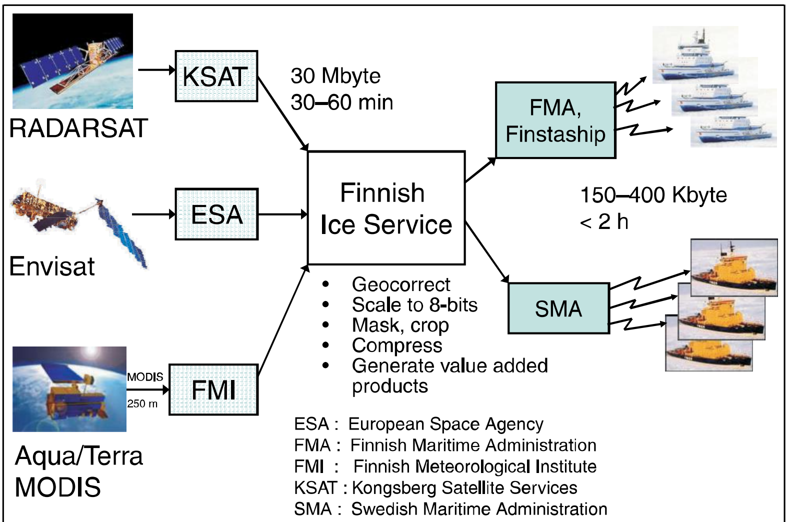

Figure 6. Delivery chain of satellite data to on-board users.

set of available parameters, cropping the original grid into smaller sea areas where the users are operating, and reducing the spatial and temporal resolution of the data grid.

Optimizing the amount of data delivered

The Façade concept has been introduced as a crucial part of the delivery system architecture (Kotovirta et al., 2003; 2006). Conceptually, Façade filters, tailors, and delivers only the relevant information to end-users and thus overcomes the main bottlenecks because of narrow bandwidth telecommunications and increased number of information products. Façade acts as a mediator between data providers and end-users by retrieving data products from the providers and tailoring them into an end-user-specific form (**Figure 7**). Tailoring may include cropping, filtering, compression, format conversion, remapping, and combining different data layers. Façade operates based on user profiles that describe different user needs, like what data parameters are needed, into what geographical area the files should be cropped, and how much data per day can be delivered over a mobile link.

The prototype implementation by the VTT Technical Research Centre of Finland has been developed and used within different research projects in which multiple products from multiple data providers have been delivered to ships for validation purposes. Façade delivers the data either directly by sending the files as e-mail attachments, or by creating a Web page from which the files can be retrieved. A prototype Web page that displays a map and thumbnail images of the full-size products located on the map has been implemented (**Figure 8**). The map presentation makes it easier for the user to compare different alternatives before starting to download the full-size

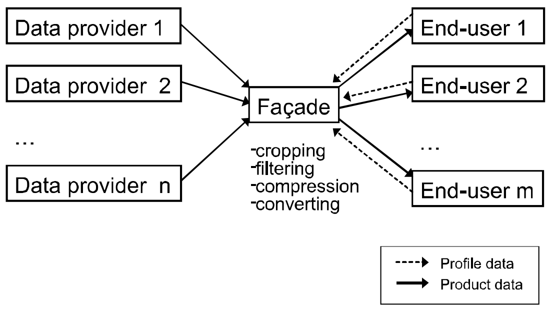

Figure 7. Schematic view of Façade as part of a data delivery chain.

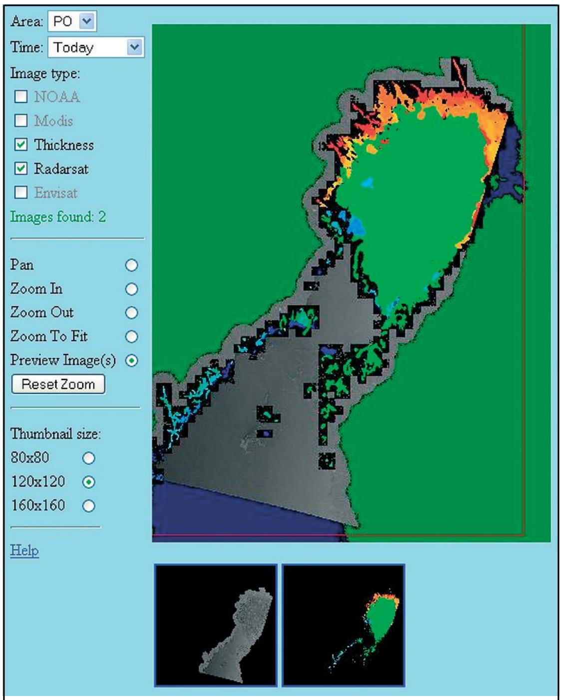

Figure 8. Web page that presents images for download according to user-selected criteria.

images. The usability of the interactive image selection and download mechanism is still to be verified when used on board ships over a low-bandwidth mobile link.

Conclusions and future work

Rapid growth of marine transportation requires faster and more user-friendly information products and distribution systems than have been available to date. Without proper information about the ice conditions, marine transportation becomes less smooth, causing considerable economical losses to society and industry.

In Finland and Sweden, digital satellite images are sent to all active icebreakers daily in the ice season and visualized on board together with ice charts, automatic classification maps, weather observations and forecasts, and ice model predictions. The use of SAR images on board has reduced the need for helicopter reconnaissance flights. The know-how of ice navigation and the systems developed, tested, and used operationally within the Baltic Sea could also be utilized in other ice-covered sea areas, like the Arctic.

Compared with the user requirements, the actual frequency of image delivery is not far from meeting the demand. Interpretation of the SAR images by experienced users is not considered difficult, as most of the end-users in the Baltic Sea have several years of experience handling and interpreting SAR images. To reduce data volumes, the images are presently subsampled to a 400 m resolution. Although many users are satisfied with the present resolution, there is a possibility to increase it owing to improved communication channels. Discussions with the users indicate that real-time information from the AIS system gives important add-on information about both the traffic situation and how the ice situation affects the performance of the ships.

New ways of presenting SAR-based products to the end-users have been piloted. The most recent products are numerical ice models and products based on dual-polarized images. End-user feedback through questionnaires and seminars shows that a rather long introduction time is needed to make the users comfortable with new products. The challenges for introducing new products are (*i*) developing proper algorithms and mechanisms for producing and transferring the products, (*ii*) validation of the products, (*iii*) usability of the products, and (*iv*) training. Validation is important for gaining user confidence about new products. Training is required to use the system and also to understand and interpret the data products correctly. As the number of data sources increases, the importance of reducing the dimensionality of the data delivered and presented to the end-users is emphasized.

Acknowledgments

The IBPlott system was originally designed as a joint assignment by the Finnish and Swedish maritime authorities. Pilot projects like POL-ICE, funded by TEKES (the Finnish Funding Agency for Technology and Innovation), and the PolarView project, funded by ESA, have enabled piloting of new functionality and services. Lastly, we would like to thank Mr. Jouni Vainio for providing information about the processing details within the Finnish Ice Service at FIMR.

References

Berglund, R. 2006. ViewIce – an example of an end user tool for route optimisation in ice covered waters. In *ARCOP Arctic Operational Platform, Technologies and Environment, Workshop 8*. Edited by L. Laiho,

P. Nordström, B. Jourio, and S. Sala. Ministry of Trade and Industry, Helsinki, Finland. pp. 32–34.

Chang, S.-J. 2001. Design and preliminary test on the integration of weather data in ECDIS for marine navigation. *Journal of Marine Science and Technology*, Vol. 9, No. 1, pp. 21–24.

Danduran, P., Mouchot, M.C., Garello, R., Fleury, D., and Thepaut, I. 1997. First realtime use of RADARSAT SAR imagery for ship navigation in Antarctica. In *IGARSS'97, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium: Remote Sensing — A Scientific Vision for Sustainable Development,* 3–8 August 1997, Singapore. IEEE, New York. Vol. 4, pp. 1678–1680.

Diarbakerly, S., El-Rabbany, A., and Coleman, D. 2002. On the integration of sea ice information into ECDIS. In *FIG XXII International Congress*, 19–26 April 2002, Washington, D.C. Available from www.fig.net/pub/fig_2002/Ts4–5/TS4_5_diarbakerly_etal.pdf [accessed 4 July 2007].

Drinkwater, M., and Rebhan, H. (Editors). 2003. *Cryosat science report*. ESA Publication Division, ESTEC, Noordwijk, The Netherlands. ESA Report SP-1272.

Grönvall, H., and Seinä, A. 2002. Satellite data use in Finnish winter navigation. In *Operational oceanography: implementation at the European and regional scales*. Edited by N.C. Flemming, S. Vallerga, N. Pinardi, H.W.A. Behrens, G. Manzella, D. Prindle, and J.H. Stel. Elsevier, Amsterdam. pp. 429–436.

Håkansson, B., Moberg, M., Thompson, T., and Gustavsson, A. 1995. Real-time use of ERS-1 SAR imagery for monitoring the ice in the Baltic Sea. In *IGARSS'95, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium: Quantitative Remote Sensing for Science and Applications*, 10–14 July 1995, Firenze, Italy. IEEE, New York. Vol. 2, pp. 856–858.

Karvonen, J., Similä, M., and Heiler, I. 2003. Ice thickness estimation using SAR data and ice thickness history. In *IGARSS'03, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 21–25 July 2003, Toulouse, France. IEEE, New York. pp. 74–76.

Kotovirta, V., Berglund, R., Haajanen, J., Mansner, J., and Tergujeff, R. 2003. Delivering near real-time met–ice–ocean observation and forecast data — the IWICOS Façade. In *Proceedings of the 3rd EuroGOOS Conference — Building the European Capacity in Operational Oceanography (EuroGOOS 2002)*, 3–6 December 2002, Athens. Elsevier, Amsterdam. pp. 639–641.

Kotovirta, V., Kanniainen, J., Veijonen, T., and Neuvonen, S. 2006. Building an application framework for monitoring the environment. In *IGARSS'06, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 31 July – 4 August 2006, Denver, Colo. IEEE, New York. pp. 2149–2152.

Leppäranta, M., and Yan, S. 1991. Use of ice velocities from SAR imagery in numerical sea ice modeling. In *IGARSS'91, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, 3–6 June 1991, Espoo, Finland. IEEE, New York. Vol. 3, pp. 1233–1237.

Pettersson, L.H., Sandven, S., Dalen, O., Melentyev, V.V., and Babich, N.I. 2000. Satellite radar ice monitoring for ice navigation of a tanker convoy in the Kara Sea. In *Proceedings of the 10th International Offshore and Polar Engineering Conference*, 28 May – 2 June 2000, Seattle, Wash. Edited by J.S. Chung, R.M.W. Frederking, H. Saeki, and W. Koterayama. International Society of Offshore and Polar Engineers (ISOPE), Cupertino, Calif. Vol. 1, pp. 691–696.

Sandven, S., Seinä, A., Berglund, R., Haas, C., and Jalonen, R. 2005. *Ice information system: review of ice services in the Northern Hemisphere*. Nansen Environmental and Remote Sensing Center (NERSC), Bergen, Norway. NERSC Technical Report 263. Available from www.arcop.fi/reports/D12.pdf [accessed 3 July 2007].

Seinä, A., Mälkki, P., and Grönvall, H. 2006. GSE programme benefits to the Baltic Sea ice navigation. In *European Operational Oceanography: Present and Future, Proceedings of the 4th International Conference on EuroGOOS*, 6–9 June 2005, Brest, France. Edited by H. Dalin, N.C. Flemming, P. Marchand, and S.E. Petterson. EuroGOOS Office and European Commission Research Directoriat-General, Luxemburg. pp. 667–672.

Smirnov, V. 2005. *Design of NSR ice information system — WP1.1a: Russian ice information system for supporting the NSR navigation*. Available from www.arcop.fi/reports/D11a.pdf [accessed 21 December 2006].

Swedish Maritime Administration. 2006. *Baltic maritime outlook 2006*. Swedish Maritime Administration, Norrköping, Sweden. Available from ec.europa.eu/transport/intermodality/motorways_sea/doc/potential/2006_03_baltic_maritime_outlook.pdf [accessed 21 December 2006].

Toudal Pedersen, L., and Saldo, R. 2004. Experience with near real time distribution of Envisat ASAR data to end-users. In *Proceedings of the 2004 Envisat and ERS Symposium*, 6–10 September 2004, Salzburg, Austria. European Space Agency (ESA), ESA Publication Division, ESTEC, Noordwijk, The Netherlands. pp. 733–738.

Vainio, J., Similä, M., and Grönvall, H. 2000. Operational use of RADARSAT SAR data as aid to winter navigation in the Baltic Sea. *Canadian Journal of Remote Sensing*, Vol. 26, No. 4, pp. 314–317.

Publication III

Kotovirta V., Jalonen R., Axell L., Riska K., Berglund R. 2009. A system for route optimization in ice-covered waters. Cold Regions Science and Technology 55(1): 52-62.

III

Cold Regions Science and Technology 55 (2009) 52–62

Contents lists available at ScienceDirect

Cold Regions Science and Technology

journal homepage: www.elsevier.com/locate/coldregions

A system for route optimization in ice-covered waters

Ville Kotovirta [a,*], Risto Jalonen [b,1], Lars Axell [c,2], Kaj Riska [d,3], Robin Berglund [a,4]

[a] VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044 VTT, Finland
[b] Helsinki University of Technology, Faculty of Mechanical Engineering and Architecture, Department of Applied Mechanics, Marine Technology, P.O. Box 5300, FI-TKK, Finland
[c] Swedish Meteorological and Hydrological Institute, SE-60176 Norrköping, Sweden
[d] ILS Oy, Heikkiläntie 8B 00210 Helsinki, Finland

A R T I C L E I N F O

Article history:
Received 14 April 2008
Accepted 1 July 2008

Keywords:
Sea ice
Ice fields
Optimization
Modeling
Mathematical models
Baltic Sea

A B S T R A C T

Information about ice is indispensable to navigation in ice-covered sea areas. For vessels traveling long distances in ice, it is worth planning routes that will reduce fuel consumption and travel time, as well as the risk of ending up in hazardous areas or getting stuck in the ice. In addition to observations on board, there is a multitude of data sources available for seafarers like satellite images, ice model data, weather observations and forecasts. However, it is difficult for a human to take into consideration all the time-varying data parameters when planning a route. In this paper, a prototype system for optimizing routes through the ice field is presented. The system integrates state-of-the-art ice modeling, ship transit modeling, and an end-user system as a route optimization tool for vessels navigating in ice-covered waters. The system has recently been validated on board merchant vessels in the Baltic Sea, and the system's performance has been analyzed statistically using AIS data. Based on the AIS data analysis the mean relative error of the estimated transit time was 0.144 [s/s] with a standard deviation of 0.147 [s/s] for long routes (90-650 km), and 0.018 [s/s] with standard deviation of 0.193 [s/s] for 50 km route segments.

1. Introduction

Information about ice is essential for navigation in ice-covered sea areas. For vessels traveling long distances in ice, it is worth designing routes that will reduce fuel consumption and travel time, as well as the risk of getting stuck in the ice or of ending up in dangerous areas. There is a multitude of information sources available for seafarers in ice-covered waters. For example near real-time delivery of satellite images to ice-going ships has been made for several years in different ice-covered sea areas like the Arctic (Pettersson et al., 2000; Smirnov, 2005), the Antarctic (Danduran et al., 1997; Toudal Pedersen et al., 2004) and the Baltic Sea (Håkansson et al., 1995; Vainio et al., 2000). State-of-the-art systems, such as IBPlott used in the Baltic Sea, present optical and radar satellite images, weather observations and forecasts, ice model forecasts and ice charts to the user on board (Berglund et al., 2007).

To take into consideration all the time-varying data in route planning on board is a difficult task even for an experienced navigator. Instead of a human a computer could plan the best route alternative given all the information available. For example, in weather routing the best route alternative is selected based on weather and oceanographic parameters. Computer-based optimization can also be applied to ice routing, i.e. selecting the best route alternative based on observed and predicted ice field properties. Weather routing and ice routing are parts of a broader problem of ship routing and scheduling. Ship routing considers the whole fleet, i.e. which ports are to be visited by the ships, and scheduling takes into consideration various events on ships' routes (Christiansen et al., 2004).

The use of desktop computers has greatly improved the possibilities to carry out ship transit modeling and transit simulations in ice, which have been described e.g. in Juurmaa, 1973; Gordin, 1978; German et al., 1981; Kämäräinen, 1986; La Praire et al., 1995, and Hannikainen, 2004. Transit simulation can be effectively used e.g. for calculating the mean speed of a vessel in variable ice conditions and evaluating the feasibility of a ship or even a fleet of vessels to transport goods on a certain area or route in ice-covered waters. The input data include sufficient knowledge of the ice conditions in the selected operation area, with regions of open water and level ice, ice channels, ice floes, ridged ice etc., along the route. Also, thrust produced by the ship propulsor(s), as a function of power, and total resistance of the ship, in various ice conditions, e.g. as a function of the level ice thickness, are key elements in the input data. This information can be acquired either by using mathematical models, as a result of model test series, or by a combination of these two methods. The reliability of the calculation methods must be ascertained by the support of full scale data. However, it is difficult to carry out a comparison with a

* Corresponding author. Tel.: +358 20 722111; fax: +358 20 7227024.
E-mail addresses: ville.kotovirta@vtt.fi (V. Kotovirta), risto.jalonen@tkk.fi (R. Jalonen), lars.axell@smhi.se (L. Axell), kaj.riska@ils.fi (K. Riska), robin.berglund@vtt.fi (R. Berglund).
1 Tel.: +358 9 451 3477; fax: +358 9 451 3493.
2 Tel.: +46 11 495 8602; fax: +46 11 495 8001.
3 Tel.: +358 9 4159 2440; fax: +358 9 4159 2444.
4 Tel.: +358 20 722111; fax: +358 20 7227024.

0165-232X/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.coldregions.2008.07.003

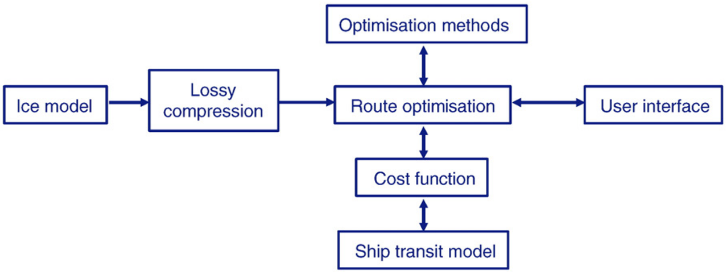

Fig. 1. Schematic view of the route optimization system.

sufficiently large database with a variety of ships and differing main parameters. The transit simulations are often completed with a comparison of the calculated results for several alternatives in an optimization process to find the best solution.

In this paper we describe the Ice Ridging Information for Decision Making in Shipping Operations (IRIS) system prototype. The IRIS system combines state-of-the-art ice modeling, ship transit modeling, and optimization methods as an operative on board route optimization system prototype for ice-covered waters. Some validation results based on on board application tests and position and speed data coming from the ship-borne Automatic Identification System (AIS) are presented.

2. Ice route optimization

A route optimization through ice requires ice model data, ship transit modeling and optimization methods. In addition, the results should be presented to end-users in a reasonable time, in a convenient format, and the amount of data delivered to mobile users must be adjusted to the capabilities of the data transmission line.

The basic principle of ice route optimization is presented in Fig. 1. The *ice model* calculates predictions of the ice conditions surrounding the ship on its route to the destination. Model output contains more data than is possible to be transferred to the optimization process on board a ship. Therefore, a *lossy compression* is required to reduce the amount of data, e.g. by lowering the spatial and temporal resolution and by selecting only the necessary parameters. The compression is lossy because by that way the amount of data can be reduced to the level that is reasonable to be transferred over a mobile link. However, the amount of information transferred must be sufficient to achieve the required quality of the optimization result. A *ship transit model* simulates the ship's performance in the prevailing ice conditions, and calculates the performance at a given location in space and time. A *cost function* calculates a cost for a route, and the *route optimization* compares different route alternatives in the process of finding a route with the minimum cost. The *user interface* allows the user to initialize the optimization process, view intermediate and final results, and alter the parameters of the optimization.

Ice model data are presented in a grid format in which every grid point contains parameters for describing the ice conditions, e.g. the probability density function (pdf) of the ice thickness. In the IRIS system, three parameters were used to describe the pdf: the level ice thickness h_i, the ridged ice thickness h_{eq}, and the ice concentration C. The equivalent ridged ice thickness is obtained from the ridge density D and the average ridge thickness H_r assuming the cross sectional areas of the actual ridged ice field and the simplified field to be the same. The ridge density gives the number of ridges per km along a straight line on the ice cover. The ridges are assumed to be of triangular shape, with the keel angle κ. Thus the (equivalent) ridged ice thickness is given as

$$h_{eq} = \mu \cdot \frac{1}{\tan\kappa} H_r^2. \quad (1)$$

These definitions are illustrated in Fig. 2.

The ice model predicts how ice conditions change over time, and therefore the ice model output contains multiple grids for different times (Fig. 3). Ice routing is treated as a continuous process, since the

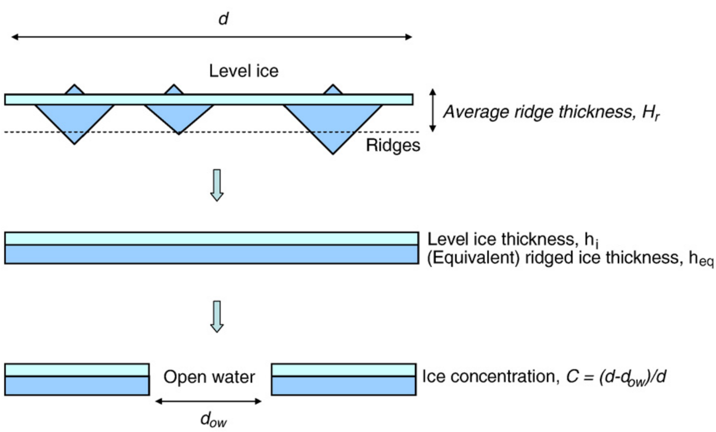

Fig. 2. Illustration of the level ice thickness, ridged ice thickness and ice concentration.

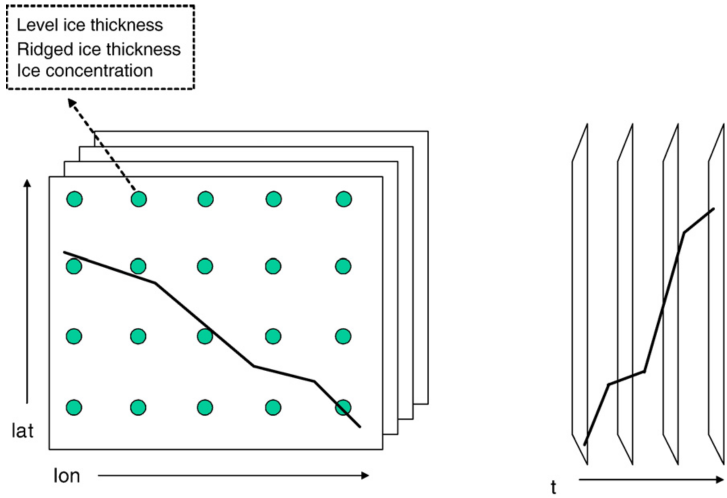

Fig. 3. Ice model data are represented in a grid format that contains grids for different times. The optimized route is presented as a black line going through space and time.

routes are not restricted to go via the grid points, but can vary continuously. The ice model data are interpolated in space and time in order to avoid discontinuities in the derivative of the cost function.

A route is considered as a point in $3n$ dimensional space, where n is the number of route points (waypoints) that have two spatial and one temporal coordinate. The task is to find a point in which the global cost function is minimized. This is a non-trivial task, as the cost function is complex and its derivative cannot be solved analytically. Three methods were implemented that do not need a derivative of the cost function: 1) Powell's method, which is a standard method in non-derivative optimization, 2) the polytope method, which is a reliable and simple optimization method, and 3) the simulated annealing, which is suitable for hard optimization problems. All these optimization methods were tested in the case of ice route optimization. In spite of the fact that the simulated annealing finds its way out from local minimums better, Powell's method provides faster results and was therefore selected to be used in the IRIS prototype.

The cost function includes restrictions, such as the land area, the pathways and corridors, and the areas of shallow water. The travel time is considered as one part of the local cost function, and it is computed as follows. The speed is a derivative of distance of time,

$$\mathrm{d}s/\mathrm{d}t = v(s,t) \tag{2}$$

where $v(s,t)$ is the ship speed at time t at distance s from the start of the route. The differential Eq. (2) is equivalent to:

$$\mathrm{d}t/\mathrm{d}s = 1/v(s,t) \tag{3}$$

As initial boundary conditions, s and t at the start point of the route segment are known, and $t(s_1)$ is solved, where s_1 is the distance at the end of the current route segment. In the IRIS system the numerical solution is computed by the fourth-order Runge–Kutta method, halving the length of the integration step and repeating until the result stabilizes. The ship speed, $v(s,t)$, is computed using the methods described later (Eqs. (23)–(26)).

3. The HIROMB ice model

This chapter describes the methods used by the HIROMB (High-Resolution Operational Model for the Baltic) model to predict the three ice parameters (level ice thickness, ridged ice thickness and ice concentration) used by the IRIS system. The HIROMB model is a fully baroclinic, thermodynamic three-dimensional model covering the North Sea, Skagerrak, Kattegat and the Baltic Sea, and it has been run operationally at the Swedish Meteorological and Hydrological Institute (SMHI) since 1995. It is a nested model in which the horizontal resolution increases from 12 nautical miles (nm) to 1nm going from the North Sea to the Baltic Sea, and the coarser-resolution grids supply the higher-resolution grids with boundary conditions. A recent addition to the grid hierarchy is the very high-resolution (60 m) grid for Brofjorden on the Swedish west coast. For the IRIS system, however, the resolution was limited to 3 nm. The model is further described in Wilhelmsson (2002).

The meteorological forcing for HIROMB is supplied by the atmospheric model HIRLAM (High-Resolution Limited-Area Model; run operationally at SMHI) for forecast lengths up to 48 h, or by the ECMWF (European Centre for Medium-Range Weather Forecasts) model for forecasts up to ten days ahead. The lateral boundary condition in terms of sea level is supplied by a separate storm-surge model (NOAMOD) with tidal components added at the open boundary in the North Sea. For other variables, climatology is used.

The ice model in HIROMB is viscous-viscoplastic and is a variation of the classic viscoplastic model proposed by Hibler (1979). In addition to the ice rheology, which is best described in Kleine and Sklyar (1996) and Wilhelmsson (2002), HIROMB solves equations of continuity for several ice variables. These were described and applied by Axell (2006) but will be summarized briefly here for convenience. The ice concentration C is solved according to[5]

$$\frac{\partial C}{\partial t} + \nabla \cdot (\overline{u}C) = R. \tag{4}$$

Here $\bar{u}$ is the ice drift and R is a ridging function parameterized as

$$R = \begin{cases} \nabla \cdot \underline{u} & ; \quad \nabla \cdot \underline{u} < 0 \ \& \ C \geq C_r, \\ 0 & ; \quad \text{otherwise}, \end{cases} \tag{5}$$

where $C_r = 1.0$ is the ice concentration at which ridging starts. Hence, R is non-zero and negative only during converging ice when the ice concentration is already 100%.

[5] Terms representing thermodynamical effects are included in the model, but have been left out in this paper for brevity.

Further, the total ice thickness h is calculated as

$$\frac{\partial(Ch)}{\partial t} + \nabla\cdot(\overline{u}Ch) = 0 \quad (6)$$

Here h is defined as the total ice thickness in the part of the grid square in which there is ice, and Ch is thus the mean total thickness (cubic meter ice per square meter) over the whole grid square.

The following ridging equations are mainly due to Lensu (2003), later modified by Lensu (2004, pers. comm.) and applied by Axell (2006). The ridge density D (number of ridges per km) is calculated according to

$$\frac{\partial(CD)}{\partial t} + \nabla\cdot(\overline{u}CD) = \frac{\beta R}{\varphi}; \quad (7)$$

cf. Lensu (2003, Eqs. 133–134) and our Eq. (4). Here we have also added β which is the fraction of deformation events that is due to ridging, the remaining part $(1-\beta)$ being due to rafting which is not taken into account in this model. β has been parameterized as

$$\beta = \begin{cases} 1 & ; \quad h_i \ge h_c \\ 0 & ; \quad h_i < h_c \end{cases} \quad (8)$$

where h_c=0.1 m and h_i is the level ice thickness (see below).

The function φ in Eq. (7) is the relative change in ice area per unit change in ridge density. It is based on Eq. (166) in Lensu (2003), but was later modified slightly to avoid a singularity and non-negative values for large values of D (Lensu, 2004, pers. comm.) The new formulation was

$$\varphi = \begin{cases} -\dfrac{315}{315CD+1000} & ; \quad p=1, \\ -\dfrac{315p\left(1-0.081\sqrt{CDH}\right)}{315CDp+1000} & ; \quad \text{otherwise}, \end{cases} \quad (9)$$

where H is the ridge sail height (see below) and the clustering variable p is given by

$$p = \min\left[1.24\exp\left(-0.16\sqrt{CDH}\right), 1.0\right] \quad (10)$$

Eq. (10) is based on Eq. (162) in Lensu (2003), but modified later by Lensu (2004, per. comm). One extra condition is applied: if $CDH \ge 68$ then

$$p = \frac{1}{3} \quad (11)$$

and

$$\varphi = -\frac{105}{315CD+3000} \quad (12)$$

The ridge sail height H is calculated with a prognostic equation as

$$\frac{\partial(CDH)}{\partial t} + \nabla\cdot(\overline{u}CDH) = \beta[\alpha_0\langle H|h_i\rangle + (\alpha-\alpha_0)H] \quad (13)$$

The expression within brackets (< >) in Eq. (13) is the mean ridge sail height formed from level ice thickness h_i, and is given by

$$\langle H|h_i\rangle = 3\sqrt{h_i} \quad (14)$$

Further, α in Eq. (13) is the appearing rate of ridge sails per distance unit, and is related to the ridging function as

$$\alpha = \frac{R}{C\varphi} \quad (15)$$

[cf. our Eq. (4) with Eqs. (131) and (136) in Lensu (2003)] and α_0 in Eq. (13) is given by

$$\alpha_0 = -\frac{3.17R}{p} \quad (16)$$

As stated above, the ridging equations above are due to Lensu (2003) and later unpublished modifications (Lensu, 2004, pers. comm.).

Finally, one extra prognostic equation is needed to account for level ice thickness h_l:

$$\frac{\partial(Ch_i)}{\partial t} + \nabla\cdot(\overline{u}Ch_i) = Ch_iR \quad (17)$$

As $R<0$ whenever ridging occurs, the term on the right-hand side of Eq. (17) is negative. The net effect is to subtract the effect of mechanical growth of ice, which thus affects the total ice thickness h [cf. Eq. (6)] but not level ice thickness h_i which only increases and decreases due to thermodynamical processes not discussed here.

The ridged ice thickness h_{eq} is here approximated as the deformed ice thickness, which is the difference between total and level ice thickness:

$$h_{eq} = h - h_i \quad (18)$$

The HIROMB ocean forecasting system includes data assimilation, to supply the model with as good initial conditions as possible for the benefit of the forecasts. The data assimilation method currently in use is the so-called Method of Successive Corrections (see e.g. Daley, 1991) which is applied for salinity, temperature, and ice variables. The latter are obtained from the daily operational ice charts from the Swedish Ice Service.

Finally, the ice model in HIROMB has been validated in terms of ice drift, ice concentration and ice ridging in unpublished reports. The overall result was that HIROMB is able to simulate very realistic fields of ice ridging variables, but that the ice ridging is somewhat underestimated. The reason is probably a slight underestimation of the ice drift.

4. Ship transit model

This chapter describes the methods used for computing the ship speed in given ice conditions. In the IRIS system we are considering the effect of level ice thickness, ridged ice thickness and ice concentration on the ship speed.

The total ship ice resistance to be used to calculate the average ship speed (v) is calculated as:

$$R_{TOT}(v) = R_i(h_i, v) + R_r(h_{eq}, v), \quad (19)$$

where R_i is level ice resistance and R_r resistance from ridge rubble, h_i is the level ice thickness and h_{eq} is the mean thickness of the ridge rubble.

There are several different methods to calculate the level ice resistance. The approach applied here is basically based on the method derived by Lindqvist (1989), but in the present application the Lindqvist formula has been modified using also Ionov (1988) ice resistance formulation as it was done by Riska et al. (1997). Here the level ice resistance is assumed to be linear with speed

$$R_i = C_1 + C_2 v, \quad (20)$$

where the constants C_1 and C_2 are

$$C_1 = f_1\frac{1}{2\frac{T}{B}+1}BL_{par}h_i + (1+0.021\phi)(f_2Bh_i^2 + f_3L_{bow}h_i^2 + f_4BL_{bow}h_i)$$

$$C_2 = (1+0.063\phi)(g_1h_i^{1.5} + g_2Bh_i) + g_3h_i\left(1+1.2\frac{T}{B}\right)\frac{B^2}{\sqrt{L}} \quad (21)$$

Table 1
The factors in level ice resistance formula (21)

Constant	Value	Unit
f_1	0.23×10^3	N m^{-3}
f_2	4.58×10^3	N m^{-3}
f_3	1.47×10^3	N m^{-3}
f_4	0.29×10^3	N m^{-3}
g_1	18.9×10^3	N/(m/s*m$^{1.5}$)
g_2	0.67×10^3	N/(m/s*m^2)
g_3	1.55×10^3	N/(m/s*m$^{2.5}$)

The constant factors are given in Table 1.

The resistance in ridged ice is calculated with the formula

$$R_r(h_{eq}, v) = 4 \cdot C_3 \cdot h_{eq}^2 \cdot [B + 2C_\psi h_{eq}] \cdot (\mu_h \cos\phi_2 + \sin\psi \sin\alpha_2) + C_4 \cdot L_{par} \cdot h_{eq}^2 + C_5 \cdot \left[\frac{LT}{B^2}\right]^3 \cdot h_{eq} \cdot A_{WF} \cdot Fn^2 \quad (22)$$

where the constants are C_3=850 N/m^3, C_4=42 N/m^3 and C_5=1.3 kN/m^3 and the coefficient C_ψ=0.047ψ − 2.115 (min 0.0). The other symbols used are given in Appendix A, the list of symbols.

The ship speed in various ice thicknesses can be found by determining where the net thrust, T_{net}, i.e. the thrust available, is equal to the total ice resistance (for the speed). The net thrust, which is available to overcome the ice resistance, can be estimated by the following approximative formulas (23) and (24), which have been used previously e.g. by Kämäräinen (1986) and by Riska et al. (1997):

$$T_{net}(v) = \left(1 - \frac{1}{3}\frac{v}{v_{ow}} - \frac{2}{3}\left(\frac{v}{v_{ow}}\right)^2\right) \cdot T_{pull}, \quad (23)$$

where the bollard pull, T_{pull}, of the vessel is calculated using the formula

$$T_{pull} = K_e (P_s D_p)^{2/3}, \quad (24)$$

where the units are kN, kW and m and the denotations are P_S for propulsion power, D_p for propeller diameter, v_{ow} for open water speed and the coefficient K_e is the quality coefficient of the bollard pull. Its value for propellers with controllable pitch in a ship with one propeller is 0.78 and in a ship with two propellers, 0.98. According to Kämäräinen (1986), formula (24) is conservative.

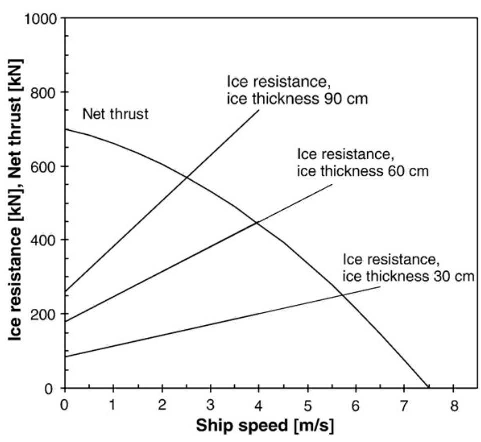

Fig. 4. Level ice resistance for three different ice thicknesses (30 cm, 60 cm and 90 cm) and the net thrust of an example vessel. The speed of the vessel can be found in the intersection of the linear resistance curve (separately at each ice thickness) and the net thrust curve.

Finally, the ship speed affected by level and ridged ice, $v_{i,eq}$, can be determined from the equation

$$T_{NET}(v) = R_{TOT}(h_i, h_{eq}, v_{i,eq}) \quad (25)$$

If the ice resistance is linear or at most quadratic polynomial versus speed, then Eq. (25) can be, for each ship and for each (h_i, h_{eq}) pair, solved explicitly. If the speed dependency of ice resistance is more complicated, then numerical methods must be applied. Fig. 4 illustrates how the ship speed depends on the net thrust and ice resistance.

A simple method was derived to take into account the effect of ice concentration. If the ice concentration C is equal or less than C_0 the ship is supposed to avoid all the ice and go at the open water speed, v_{ow}. When the ice concentration is equal or more than C_1 the ship goes at the speed affected by the level and ridged ice, $v_{i,eq}$. In between C_0 and C_1 the ship speed at time t and distance s from the start of the route is a linear combination of the open water speed and the ice speed, as shown in the Eq. (26) and illustrated in Fig. 5.

$$\hat{v} = \begin{cases} v_{ow} & ; \quad C \le C_0 \\ \dfrac{(C_1 - C)v_{ow} + (C - C_0)v_{i,eq}}{(C_1 - C_0)} & ; \quad C_0 < C < C_1 \\ v_{i,eq} & ; \quad C \ge C_1 \end{cases} \quad (26)$$

The estimated ship speed, $\hat{v}$, is calculated by first solving numerically the Eq. (25) to get $v_{i,eq}$ and then using the Eq. (26) with parameters C_0=70% and C_1=95%. Solving the Eq. (25) is time-consuming, and therefore to speed up the calculations a pre-calculated matrix of ship speeds is first calculated. Ship speeds are then interpolated from the matrix during the optimization to get faster results.

5. The IRIS system architecture

In this chapter the system architecture is described, i.e. how the components described earlier are integrated as the IRIS prototype.

5.1. Computational viewpoint

There is a multitude of ways to distribute the computational tasks of route optimization between on shore and on board computers. The distribution depends on how much data can be transmitted in a reasonable time over a mobile link to the ship. One extreme would be to run an ice model on board, and use the full resolution model output in the route optimization. This is probably not feasible as it would

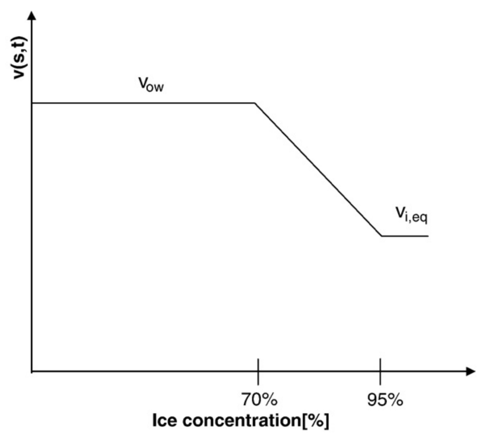

Fig. 5. Assumed relationship between ice concentration and ship speed.

require transmission of full resolution weather model data to the ship, and also a lot of computing power on board. The other extreme would be to minimize the amount of data transmitted by transferring only the waypoints of the initial guess to the server, carrying out all the calculations on the server side, and then returning the optimized route waypoints back to the ship. However, this would require an online connection when operating the system.

The IRIS system architecture runs the ice model and lossy compression on the server side ashore, and the route optimization with ship transit calculation on board. This appeared to be a good choice to distribute the computational tasks between on board and on shore computers, because the communication link did not allow continuous online connections to the ships. In addition, running the optimization on board enables faster responding interactive user interface.

The overall architecture of the IRIS system is described in Fig. 6. The HIROMB ice model uses weather forecast data coming from the atmospheric model HIRLAM. Ice model data are delivered to end-users via a dedicated component called Façade that handles different requests for ice data for different sea areas. Facade selects, filters, and delivers ice data to the users based on user-defined profiles (Kotovirta et al., 2003, 2006). During the processing the amount of data is reduced from order of gigabytes into order of hundred kilobytes. Façade thus takes care of the lossy compressing by cropping the original grid data files and reducing spatial and temporal resolution of the files.

5.2. End-user system

The end-user interface of the IRIS system is implemented in the ViewIce system. ViewIce is a route planning and decision support tool developed by the Technical Research Centre of Finland (VTT) especially for the needs of ships in winter traffic. It presents time-dependent weather, oceanographic, and ice data to the user, both observations and predictions, and optimizes ship routes given conditions data and a corresponding ship transit model. The transit model (described above) was implemented as part of ViewIce's optimization framework.

One of the aspects of ship route optimization system is how the results are presented to the user, and how the user interacts with the system. A typical user of the IRIS system would be a ship mate with only a little experience on computers, and therefore the user interface of the end product should be intuitive and easy to use. In the prototype, the optimization is started by drawing an initial guess on the map (Fig. 7). ViewIce starts the optimization and presents an animation about how the optimization proceeds, and how the optimization finds its way towards a local optimum. The optimization ends as the ending criteria are reached or the user stops the optimization. The user can stop the process at any time, and the best result so far is presented (Fig. 8). The optimization method, the parameters and ending criteria are defined so that the user gets a good enough result in minutes and does not have to wait for hours. The user can query the estimated transit time of any route drawn on the screen. It is also possible to give more than just one initial guess to ViewIce, and all the optimized result routes are drawn on the screen. This gives the user a possibility to compare different initial route alternatives and their optimized results.

In order to study the variability of the optimal route, a prototype method was developed that takes random samples of the routes that are within prescribed bounds of the optimum, and draws these routes together with the optimum (Fig. 8). Without any further analysis this presentation gives visual information about the sensitivity of the optimal route. If all the result routes close to the optimum go close to some point, there is a narrow corridor presented in the visual presentation. This suggests that one should not divert too much from the route at that point. On the other hand, if there are routes in some location going relatively far from the optimum route, there is probably a wider corridor where the user can more freely choose his route according to other preferences.

Although there are also other sources of data available, ice model forecasts are currently the only data source that is used in the automatic route optimization. When ice conditions are static, ice charts could in principle be used in optimizing routes. Also, manually derived ice situation forecasts that describe the future conditions could be used, but their spatial and temporal resolution is probably too coarse compared to ice model outputs. In addition, when there are no manual components in the data processing chain from ice models to ViewIce the users have always the latest data available for route optimization when they start using ViewIce.

6. Validation

The IRIS system was tested within the Baltic Sea on board merchant vessels, a research vessel and an icebreaker during the winter 2005. Some scientists of the project group participated in the test voyages, and installed and introduced the IRIS system to the ship crew. Afterwards, the crew was supposed to use the system by themselves and answer a questionnaire about the usage of the system. Because it was not feasible to organize more test voyages for validation on board the ships, also AIS (ship-borne Automatic Identification System) data were used for statistical validation of the transit times calculated by the system.

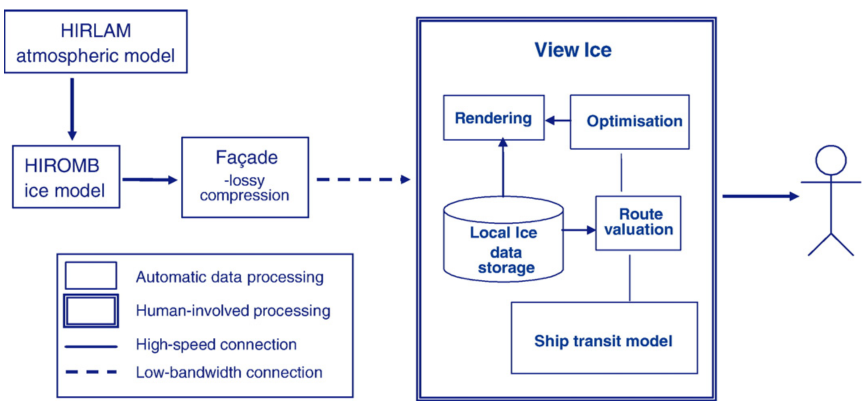

Fig. 6. The IRIS system architecture.

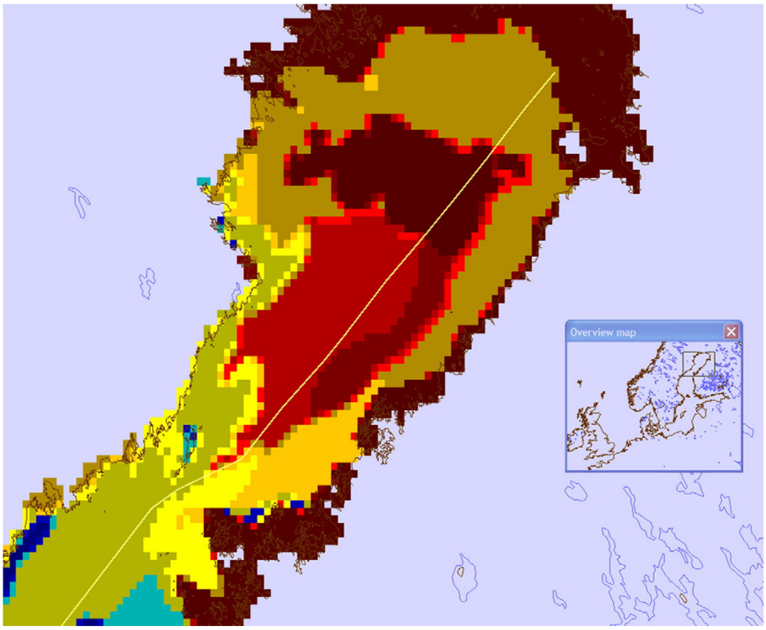

Fig. 7. ViewIce presenting an ice model output and a user-drawn initial route for optimization.

The quality of the optimization result depends on the quality and resolution of the ice conditions data, the amount of loss in the information because of lossy compression, the accuracy of the ship transit model, the properties of the optimization method used, and the calculation time reserved for the optimization process. An operational system should give results in a reasonable time in order for the results to be utilized within the time frame they apply. It is not feasible to spend too much time in fine-tuning a result that probably contains inaccuracies because of other components in the processing chain. Because the ice forecast model is stochastic in its nature, methodologies for analysis of stochastic system behavior could give improved knowledge about the influence of the different error sources on the end result. Validation of this would, however, require independent validation of the different components in the system and at this stage we do not have enough data to do that. We can just state that the total error (i.e. difference between true optimum and the

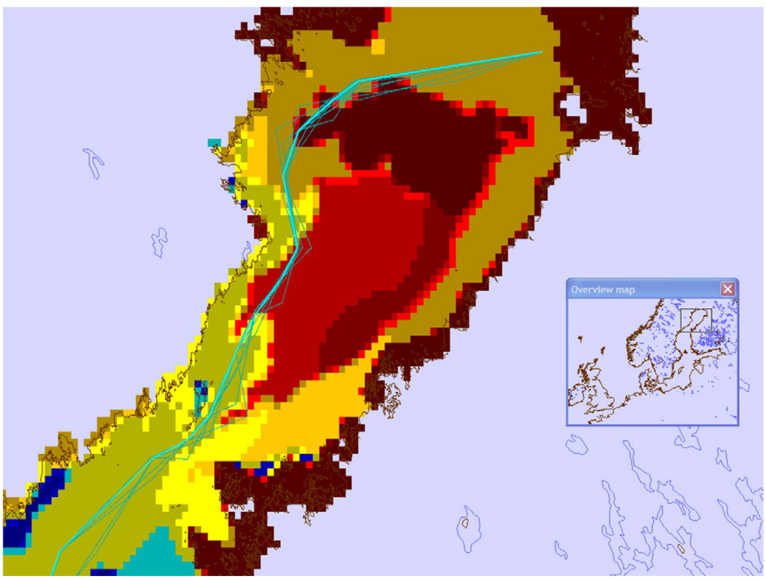

Fig. 8. ViewIce presenting the optimization result. A random sample of routes within bounds of 5% longer transit time than that of the optimal route are drawn together with the optimal route. This gives visual information about the sensitivity of the optimum.

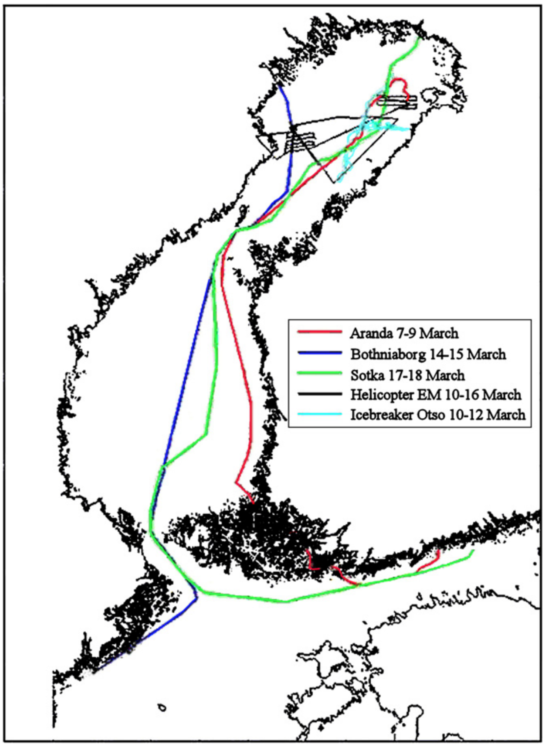

Fig. 9. IRIS validation voyages in the Baltic Sea during winter 2005 with RV Aranda, M/S Bothniaborg, M/T Sotka and IB Otso.

estimated one) is a stochastic variable which is determined by the behavior of the error elements of the different components of the processing chain:

$$e_{\text{result}} = f\left(e_{\text{ice model}}, e_{\text{compression}}, e_{\text{transit model}}, e_{\text{optimization}}\right) \quad (27)$$

where e_{result} is the total error in the optimized route, $e_{\text{ice model}}$ is the error produced by the ice model, $e_{\text{compression}}$ is the error due to lossy compression, $e_{\text{transit model}}$ is the error of the transit modeling, and $e_{\text{optimization}}$ is the error caused by the limitations of the selected optimization methods, e.g. the limited computing resources available for the optimization and the quality of the selected restrictions in the cost function, like the used land mask.

6.1. Test voyages

The test voyages to observe actual ice conditions and passage times through the Bothnia Bay were conducted on board M/S Bothniaborg owned by Wagenborg during 14–15 March, 2005, and on board M/T Sotka owned by Neste Shipping during 17–18 March 2005 (Fig. 9). While on board, the scientists used the system to estimate ship transit times which were then compared with the measured transit times. It was not possible to divert the ship from the original course just to test the routes the system had suggested. Therefore, the encountered ice conditions were most of the time different channels, closed tracks and varying ice conditions, and it was hard to make a comprehensive comparison between real performance and estimated performance based on the IRIS system. The predicted ice conditions matched however quite well with observed in the general level, and seemed to provide information about variability of the ice conditions from a statistical point-of-view. The system was also installed on board research vessel Aranda of the Finnish Institute of Marine Research (FIMR) that served as a base for in-situ and helicopter-based EM (Electro-magnetic) measurements at the time. During the field trial period scientists were also making observations and measurements on board IB Otso, where the IRIS system was installed. However, the system on board Aranda and Otso was used just to present ice model data, as these ships were not modeled in the transit model for transit time estimations.

A comparison between the estimated and true transit times for the voyages of Bothniaborg and M/T Sotka indicated that the error of the transit time estimate was below 7%. However, these voyages do not form a sufficiently solid basis for the assessment of the accuracy of estimated travel times. ViewIce as a navigational assistance tool was considered useful, but users were not optimistic about the optimization features. Their main argument was that the ships within the Baltic have to obey icebreaker orders in selecting waypoints and they do not have the freedom to navigate independently. This is a consequence of the current policy of the Finnish and Swedish winter navigation system in the Baltic Sea. If a ship deviates from the pre-defined routes and gets stuck it may have to wait a while for the assistance. However, in areas with less frequent traffic, different traffic patterns or non-existing icebreaker assistance, the optimization features may become more attractive on board merchant vessels.

6.2. Validation based on AIS data

The AIS system gives observed locations and speeds of ships, and that data were used to validate the estimated ship speeds and transit times. One downside of the data is that the AIS system does not give information about the engine power that would be an important parameter to understand whether a ship is slowing down because of reduced engine power or because of thicker ice in the area.

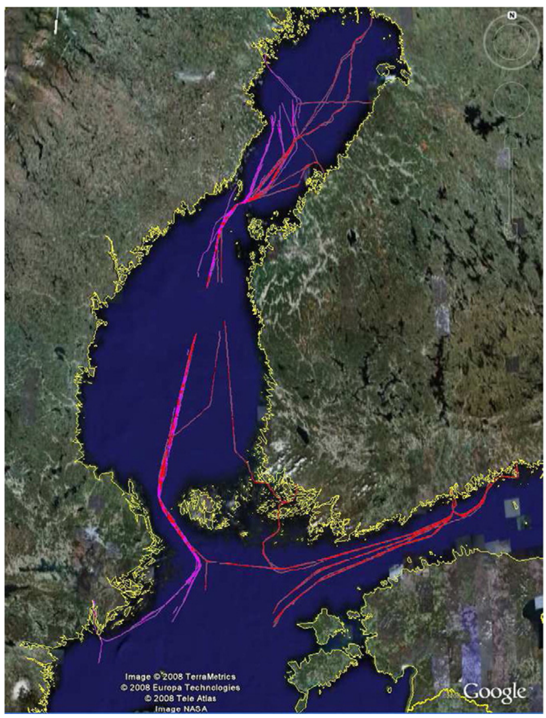

Fig. 10. AIS data tracks of M/S Bothniaborg, M/S Baltiaborg, (pink) and M/T Sotka (red) in the winter 2005. Some data are missing in the middle of the Bothnian Sea.

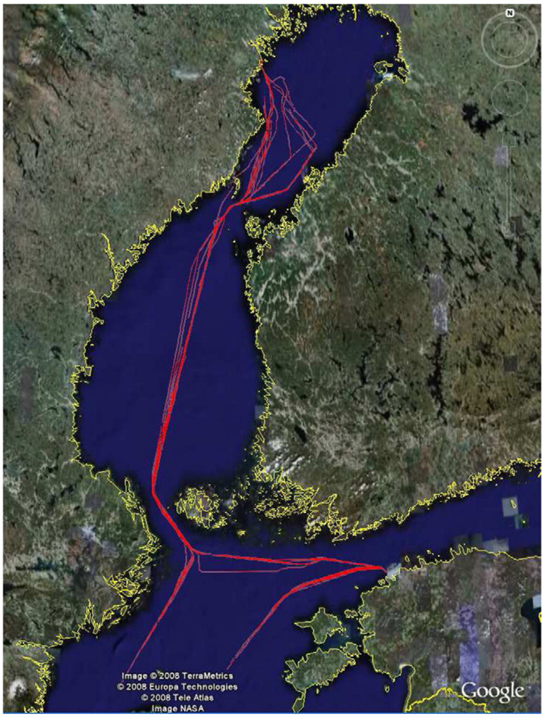

Fig. 11. AIS data tracks of M/S Bothniaborg and M/S Baltiaborg in the winter 2007 (M/T Sotka did not operate in the Baltic Sea that year).

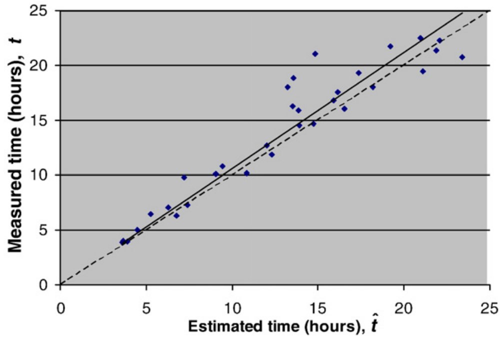

Fig. 12. The comparison between estimated and measured times for route segments whose lengths vary from about 50 nautical miles (90 km) to 350 nautical miles (650 km). The solid trend line calculated from the data is forced to intercept the origin ($t=1.06\hat{t}$). The dashed line indicates the optimal trend line ($t=\hat{t}$).

The ship speed and location observations of the three modeled ships, M/S Bothniaborg, M/S Baltiaborg and M/T Sotka, were collected from the AIS datasets for years 2005 and 2007 (Figs. 10 and 11). The dataset of the year 2006 was not used, because for technical reasons ice model data were not available for AIS validation. However, the winters 2005, 2006 and 2007 had all similar (mild) ice conditions, so the use of the 2006 dataset was not considered crucial for the validation. Table 2 gives a short excerpt of a data set that was used in the AIS validation.

Outliers, for example data samples where the ships were slowing down for an unknown reason, were removed from the data, and only observations where the ships were supposed to go at least partly in ice were selected. Route segments were formed from the observations by selecting sequences of observations were the spatial and temporal difference between the points was within a prescribed limit. The route segments were from about 50 nautical miles (90 km) to 350 nautical miles (650 km) in length. These segments were also chopped into shorter route segments of about 27 nautical miles (50 km) in length. The IRIS system was used to estimate transit times for full-length and chopped route segments, and the times were compared with the measured transit times. The idea of this comparison is to show how well, statistically, the IRIS system would estimate a transit time for route segments of different lengths given the route points and ice model forecast (Fig. 12).

The relative error was computed as

$$e_i = \frac{t_i - \hat{t}_i}{\hat{t}_i}, \tag{28}$$

where t_i is the measured travel time for route segment i, and $\hat{t}_i$ is the estimated travel time for the same route segment. The mean of the relative error was 0.144 [s/s] with a standard deviation of 0.147 [s/s] for long routes, and 0.018 [s/s] with standard deviation of 0.193 [s/s] for 50 km route segments. The mean relative errors and standard deviations differ although the route segments are based on the same data samples. This is probably partly due to averaging of spatial and temporal errors in the ice model data along longer routes, and partly because the datasets differ (longer routes contain segments of open water that were filtered out when shorter segments were formed).

Estimated average speeds were also calculated, and these were compared with the measured average speeds (Figs. 13 and 14). In an unbiased case, the trend lines calculated from the data points in the

Table 2
Example data set that was used in the AIS data validation

Distance traveled (nautical miles)	Lon	Lat	Time	Measured speed (knots)	Estimated speed (knots)	Level ice thickness (m)	Ridged ice thickness	Ice concentration
1.55	22,936	64,393	0:05:42	14.82	12.46	0.24	0.19	93%
1.37	22,893	64,380	0:11:54	13.24	12.63	0.23	0.15	93%
1.46	22,837	64,377	0:17:54	14.56	12.77	0.22	0.12	94%
1.41	22,790	64,366	0:24:12	13.39	12.99	0.21	0.08	94%
1.25	22,748	64,356	0:30:44	11.42	13.17	0.20	0.04	95%
1.30	22,705	64,345	0:37:09	12.13	13.29	0.19	0.02	95%
1.38	22,657	64,334	0:43:20	13.40	13.39	0.19	0.01	96%
1.37	22,609	64,325	0:49:32	13.25	13.44	0.19	0.00	96%
1.33	22,562	64,316	0:55:45	12.84	13.43	0.19	0.00	96%
1.52	22,510	64,305	1:01:54	14.79	13.39	0.19	0.00	96%
1.34	22,463	64,296	1:08:18	12.51	13.36	0.19	0.01	96%
1.58	22,408	64,285	1:14:26	15.39	13.33	0.19	0.01	96%

The measured speed is the speed calculated from the location and time information coming from the AIS system. The estimated speed is the speed calculated by the IRIS system using the three ice parameters interpolated from the ice model data in the point in space and time.

scatter diagrams should have a slope of one as the estimated and measured speeds would correlate better.

More data and further studies are still needed to analyze the bias and variance of the error in more detail and find out which components affect the global error most. The winters 2005 and 2007 (and also 2006) were mild and thus no harsh ice conditions were met. Therefore, there were data samples only from speeds close to the open water speed of the ships. Also, only three ships were modeled for the transit calculations, and in addition, only two of them operated in 2007.

7. Discussion and further work

The ship route optimization through the ice field is a hard problem. Ice field modeling and ship transit modeling are both difficult tasks by themselves, so the challenges of building an integrated operative route optimization system are considerable. Besides that the system should model the future ice conditions and ship transit in ice correctly it should also take care of delivering and presenting the right amount of relevant information to the users on board. The IRIS system takes the first steps towards an operative end-user system that would integrate state-of-the-art ice modeling and state-of-the-art ship transit modeling and give trustworthy route alternatives through a ridged ice field.

The results show that the IRIS system could in principle be used in estimating ship transit times, but also that there is still lot of work to do in improving the different components of the system. The usability depends on the quality of separate components, i.e. ice models, transit models, data delivery mechanisms, optimization methods and the user interface.

In the current system, three parameters describing the ice field, the level ice thickness, the ridged ice thickness and the ice concentration, were used in calculating the ship speed and transit time. In the future, also other parameters could be taken into account, like ice drift, ice compression, ridging height and ridging density. As ships mainly follow ice channels, it would be interesting to model transit in old and new channels. Although, ice channels are a challenge for ice modeling, the channels could be estimated from AIS data, satellite data, or ground-based radar data. Additionally, more research on the effects of the ship main parameters and propulsion concept (conventional/double acting) are recommended. This future research should be carried out both by ridge transit tests in model scale as well as in full scale. The acquisition of comprehensive full scale data with detailed and even more extensive ice measurements along the ship route, and the utilization of such data for validation are also recommended.

More discussions between the end-users and the developers are required before the users can accept a route optimization system as a

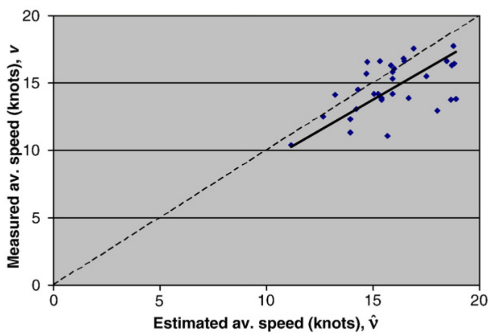

Fig. 13. The comparison between the estimated and measured average speeds along route segments of varying lengths from about 50 nautical miles (90 km) to 350 nautical miles (650 km). The solid trend line calculated from the data is forced to intercept the origin ($v=0.918\hat{v}$). The dashed line indicates the optimal trend line ($v=\hat{v}$).

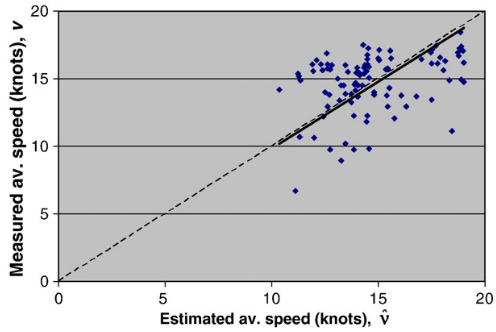

Fig. 14. The comparison between the estimated and measured average speeds for route segments of about 27 nautical miles (50 km) in length. The solid trend line calculated from the data is forced to intercept the origin ($v=0.986\hat{v}$). The dashed line indicates the optimal trend line ($v=\hat{v}$).

navigational support tool. The user interface and the presentation of the optimization results are important components of the whole integrated system. Currently, the IRIS system calculates optimal routes without any further analysis of the confidence of the result, with the exception of the visual presentation of the variability of the optimum. Ice models estimate the probability density functions (pdf) of the ice thickness, and these pdfs are converted into average ship speeds using the ship transit model. In the future, the pdfs of the ship speeds could be calculated instead of their plain mean values. Also, the pdf of the transit time of the whole route, or the probability of getting stuck, could be derived. This would require more computing power, but the results could be more useful from the user point-of-view.

Route optimization is useful when ships are crossing large sea areas where there are many possibilities to select a route. When considering relatively short distances within the Baltic Sea, there are only a limited number of possibilities to select a route. In addition, vessels need to follow icebreaker orders and do not have much freedom to choose route on their own. However, in early winter and late spring when the icebreaking assistance is reduced the ice information is valuable for independent navigation, and small savings due to improved ice navigation of individual ships could cause significant savings to the economy in view of the large number of vessel visits per year within the Baltic Sea. The route optimization could be used by icebreakers, not just for finding their own way, but to estimate how the ships they are assisting would perform in the ice between the waypoints icebreakers have ordered them to follow.

The know-how of ice navigation and the systems developed, tested and used operatively within the Baltic Sea could be utilized in other ice-covered sea areas also. Especially, when the economical exploitation of northern (the Arctic) energy resources realizes, proper understanding of ice conditions and tools facilitating ice navigation are needed.

Acknowledgements

The work described in this paper was mainly conducted in an EU funded 5th framework programme project called IRIS — Ice Ridging Information for Decision Making in Shipping Operations, during 2003–2005. AIS validation for 2007 was implemented within a project called ShipSensorNet funded by Finnish Funding Agency for Technology and Innovation.

Appendix A. The list of symbols

h_i	level ice thickness
h_{eq}	ridged ice thickness
C	ice concentration

D ridge density
κ keel angle
H_r average ridge thickness
H ridge sail height
h total ice thickness
β the fraction of deformation due to ridging
φ the ridge structural function
α the appearing rate of ridge sails per distance unit
μ_H friction coefficient between the hull and ice
ϕ bow angle
ϕ_2 the angle between buttock line at B/4 with the horizontal
α_2 waterline entrance angle at B/4
ψ frame normal angle at B/4
A_{WF} bow waterplane area
$F_n = v/\sqrt{gL}$ Froude number
g acceleration of gravity
L ship length
L_{bow} length of the ship bow
L_{par} length of the ship parallel midbody
B ship maximum breadth
T ship draught
v_{ow} open water speed
$v_{i,eq}$ ship speed affected by level and ridged ice
$\hat{v}$ estimated ship speed
t_i measured travel time for route segment i
$\hat{t}_i$ estimated travel time for route segment i

References

Axell, L., 2006. Extended ice forecast modeling for the Baltic Sea. The Fourth International Conference on EuroGOOS. EuroGOOS Office, SMHI, SE-60176 Norrköping, Sweden.

Berglund, R., Kotovirta, V., Seinä, A., 2007. A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea. Canadian Journal of Remote Sensing 33, 378–387.

Christiansen, M., Fagerholt, K., Ronen, D., 2004. Ship routing and scheduling: status and perspectives. Transportation Science 38, 1–18.

Daley, R., 1991. Atmospheric Data Analysis. Cambridge University Press, Cambridge.

Danduran, P., Mouchot, M.C., Garello, R., Fleury, D., Thepaut, I., 1997. First realtime use of RADARSAT SAR imagery for ship navigation in Antarctica. International Geoscience and Remote Sensing Symposium, IGARSS 1997. IEEE, New York, NY, USA, pp. 1678–1680.

German, J.G., Macpherson, M.D., Meakin, J., Parker, C.W., 1981. Marine transportation of oil and gas in the Alaskan Arctic. Sixth Ship Technology and Research (Star) Symposium, pp. 35–59.

Gordin, S., 1978. Full scale and model testing of ships operation in ice. The 3rd International Offshore Craft Conference, Day 2, Paper No. 4, London.

Hannikainen, T., 2004. Ship's Passage through a Ridged Ice Field. Master's thesis, Helsinki University of Technology, Espoo.

Hibler III, W.D., 1979. A dynamic thermodynamic sea ice model. J. Phys. Oceanogr. 9, 817–846.

Håkansson, B., Moberg, M., Thompson, T., Gustavsson, A., 1995. Real-time use of ERS-1 SAR imagery for monitoring the ice in the Baltic Sea. International Geoscience and Remote Sensing Symposium, IGARSS '95, Vol. 2. IEEE, New York, NY, USA, pp. 856–858.

Ionov, B., 1988. Ice Resistance and its Components. AARI, Leningrad. (in Russian).

Juurmaa, K., 1973. Feasibility Prediction of Winter Navigation for Vessels Designed for the Bay of Bothnia. Helsinki University of Technology, Masters Thesis. Espoo. (in Finnish).

Kleine, E., Sklyar, S., 1996. Mathematical features of Hibler's model of large-scale sea-ice dynamics. D. Hydrogr. Z. 47, 179–230.

Kotovirta, V., Berglund, R., Haajanen, J., Mansner, J., Tergujeff, R., 2003. Delivering near real-time met-ice-ocean observation and forecast data — the IWICOS Façade. 3rd EuroGOOS Conference — Building the European Capacity in Operational Oceanography (EuroGOOS 2002). Elsevier, Amsterdam, pp. 639–641.

Kotovirta, V., Kanniainen, J., Veijonen, T., Neuvonen, S., 2006. Building an application framework for monitoring the environment. International Geoscience and Remote Sensing Symposium, IGARSS 2006. IEEE, Denver, Colorado, US, pp. 2149–2152.

Kämäräinen, J., 1986. Icebreaking tanker transit simulation model. International Offshore and Navigation Conference and Exhibition, Polartech '86, Vol. 2. VTT, Espoo, Finland, pp. 1138–1159.

LaPrairie, D., Wilhelmson, M., Riska, K., 1995. A transit simulation model for ships in Baltic ice conditions. Helsinki University of Technology, Ship Laboratory Report M-200, Otaniemi, Espoo, Finland.

Lensu, M., 2003. The evolution of ridged ice fields. Dissertation for the Degree of Doctor of Philosophy. Helsinki University of Technology, Ship Laboratory, Finland.

Lindqvist, G., 1989. A straightforward method for calculation of ice resistance of ships. The 10th Conference on POAC, POAC'89. Luleå university of technology, Luleå, Sweden, pp. 722–735.

Pettersson, L.H., Sandven, S., Dalen, O., Melentyev, V.V., Babich, N.I., 2000. Satellite radar ice monitoring for ice navigation of a tanker convoy in the Kara Sea. The International Offshore and Polar Engineering Conference, Vol. 1. International Society of Offshore and Polar Engineers, California, USA, pp. 691–696.

Riska, K., Wilhelmson, M., Englund, K., Leiviskä, T., 1997. Performance of merchant vessels in ice in the Baltic. Winter Navigation Research Board, Report 52, Helsinki.

Smirnov, V., 2005. Design of NSR ice information system — WP1.1a: Russian ice information system for supporting the NSR navigation. Technical Report, Deliverable 1.1a, Arctic Operational Platform (ARCOP), EU. Available from < http://www.arcop.fi/reports/D11a.pdf > [25 Oct, 2007].

Toudal Pedersen, L., Saldo, R., 2004. Experience with near real time distribution of envisat asar data to end-users. Envisat and ERS Symposium. European Space Agency, pp. 733–738.

Vainio, J., Similä, M., Grönvall, H., 2000. Operational use of RADARSAT SAR data as aid to winter navigation in the Baltic Sea. Canadian Journal of Remote Sensing 26, 314–317.

Wilhelmsson, T., 2002. Parallellization of the HIROMB ocean model, Licentiate Thesis, Royal Institute of Technology, Department of Numerical and Computer Science, Stockholm, Sweden.

Publication IV

Kotovirta V., Karvonen J., von Bock und Polach R., Berglund R., Kujala P. 2011. Ships as a sensor network to observe ice field properties. Cold Regions Science and Technology 65(3): 359-371.

IV

Cold Regions Science and Technology 65 (2011) 359–371

Contents lists available at ScienceDirect

Cold Regions Science and Technology

journal homepage: www.elsevier.com/locate/coldregions

Ships as a sensor network to observe ice field properties

Ville Kotovirta [a,*], Juha Karvonen [b,1], Rüdiger von Bock und Polach [c,2], Robin Berglund [a,3], Pentti Kujala [c,4]

[a] VTT Technical Research Centre of Finland, PO Box 1000, FI-02044 VTT, Finland
[b] Finnish Meteorological Institute, PO Box 503, FI-00101 Helsinki, Finland
[c] Aalto University, School of Science and Technology, Department of Applied Mechanics, Marine Technology, PO Box 5300, FI-TKK, Finland

ARTICLE INFO

Article history:
Received 4 March 2010
Accepted 17 November 2010

Keywords:
Sea ice
Radar methods
Remote sensors
Navigation
Baltic Sea

ABSTRACT

This paper introduces a concept and a prototype system of using ships and coastal stations as a sensor network to obtain additional information about the ice field. The system collects marine radar images and ship performance observations, forms mosaics of images from multiples radars, calculates ice drift from subsequent radar images, analyzes trafficability in different sea areas using performance observations, and delivers processed images, trafficability estimation, and ice drift information to end-users. The prototype was developed and tested during the winters of 2008 and 2009 in the Baltic Sea. In this paper, we describe the prototype and discuss the usability of a ship sensor network. The concept appears to be feasible, and such a system would provide additional information about prevailing ice conditions.

1. Introduction

Information about prevailing ice and weather conditions is important for all operations in ice-infested sea areas. There are currently many data products available for seafarers, e.g., satellite images, ice charts, weather data, and ice model forecasts. Satellite images and ice charts go out of date quickly, however, especially when the ice field is moving. Models may also not predict the changing situation correctly. More frequent observations of the ice field are therefore needed to enhance and update the latest view based on satellite imagery and model results.

In this paper, we introduce a concept and a prototype system that utilize ships and coastal stations as sensors to obtain additional information about the ice field. The system collects near real-time ship performance and marine radar data automatically; processes, analyzes, and combines the data with other data sources on the server side; and delivers enhanced information back to ships and other users such as ice services and maritime authorities. Ship performance data give indirect information about trafficability of different sea areas, and marine radar images can be used for visual interpretation of the ice field and automatic ice drift computation. Mosaicing images from multiple radars and the collection of performance observations from many ships allow a broader view of the situation to be formed. The prototype was developed and tested in the Baltic Sea during the winters of 2008 and 2009, and it demonstrates the possibilities of using ships as sensors to acquire real-time information about the ice field in areas with frequent ship traffic.

2. The prototype system architecture

A reference architecture of ships as a sensor network is presented in Fig. 1. The system allows end-users to perceive prevailing ice conditions and estimate how the conditions will affect the performance of winter traffic. It automatically collects information about the ice field, sends it to the server side for processing, and delivers enhanced and combined information back to users. The system should deliver the latest updates as real time as possible as the ice field is constantly changing and the information becomes out of date quickly. While delivering the data, the system has to take into account the low bandwidth data communication channel to the ships.

The prototype system conforms to the reference architecture and can be considered to be a distributed near real-time environmental information system consisting of hardware and software components located on board and on server side computers (Fig. 2). The system collects marine radar images and ship performance observations, forms mosaics of images coming from multiples radars, calculates ice drift from subsequent radar images, analyzes trafficability in different sea areas, and delivers processed images and ice drift information to end-users. The prototype utilizes AIS (Automatic Identification System) and IBNet (Ice Breaker Network) system data for trafficability estimation. The information that the prototype system produces is

* Corresponding author. Tel.: +358 20 722111; fax: +358 20 7227024.
E-mail addresses: ville.kotovirta@vtt.fi (V. Kotovirta), juha.karvonen@fmi.fi (J. Karvonen), ruediger.vonbock@tkk.fi (R. von Bock und Polach), robin.berglund@vtt.fi (R. Berglund), pentti.kujala@tkk.fi (P. Kujala).
[1] Tel.: +358 9 192 91; fax: +358 9 179 581.
[2] Tel.: +358 9 45123478; fax: +358 9 451 23493.
[3] Tel.: +358 20 722111; fax: +358 20 7227024.
[4] Tel.: +358 400 878145; fax: +358 9 451 23493.

0165-232X/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.coldregions.2010.11.005

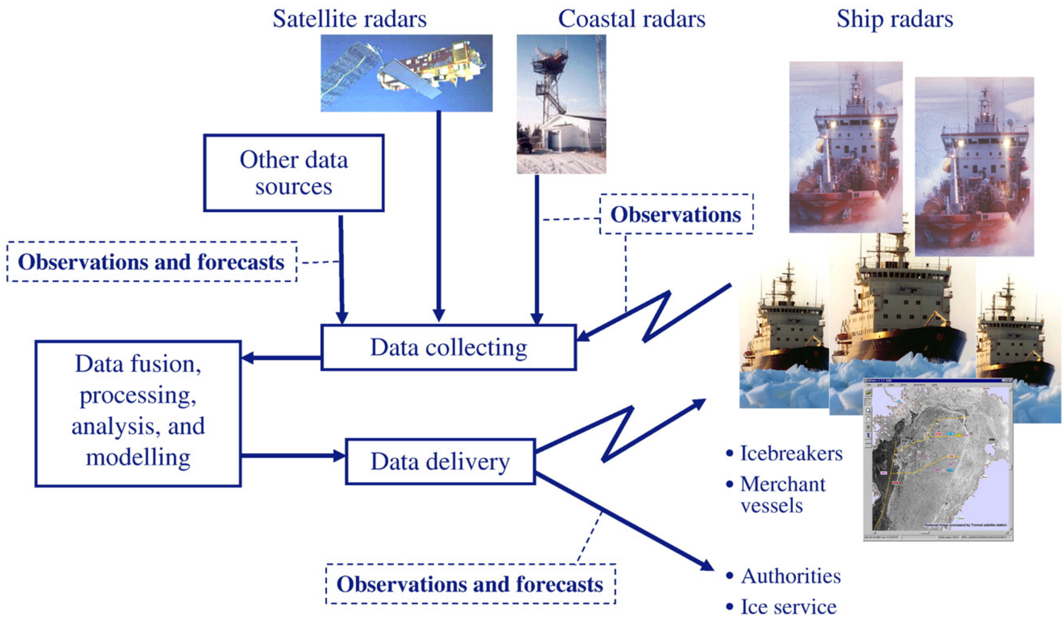

Fig. 1. Reference architecture for ships as a sensor network to observe ice field properties.

displayed to end-users using state-of-the-art decision support systems IBPlott (Berglund et al., 2007) and ViewIce that are in operative use in the Baltic Sea. Marine radar mosaics, ice drift observations, and trafficability estimations complement satellite images, ice model predictions, and weather data used in decision making on board. Ice drift observations calculated from the marine radar data can also be used to generate initial conditions for the ice model and to validate the modeled ice motion.

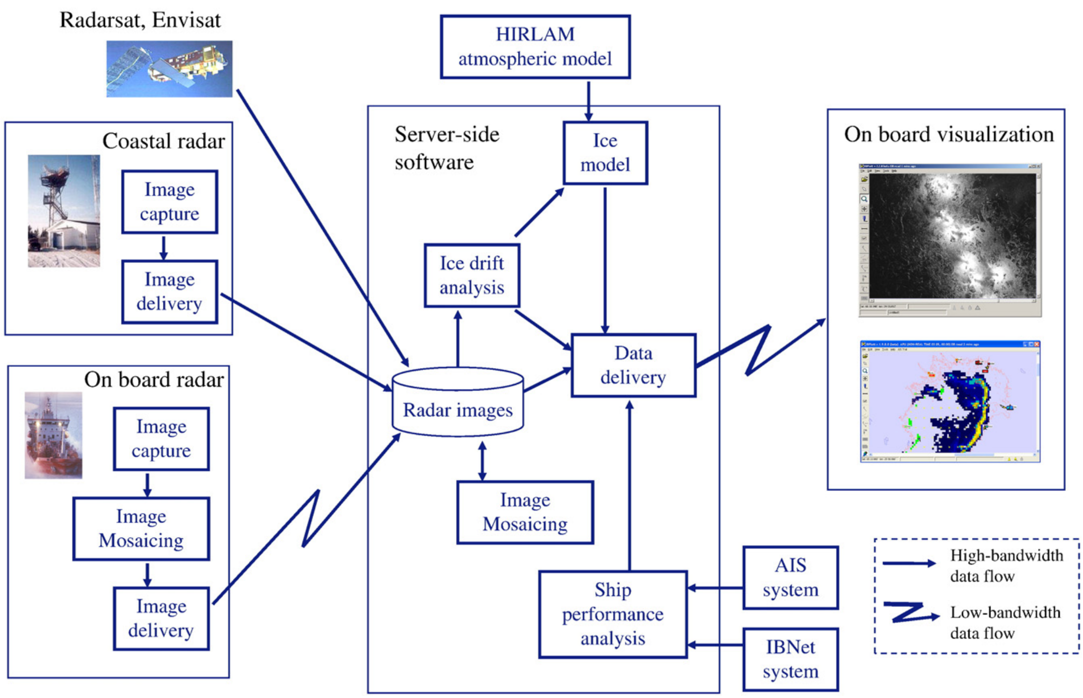

Fig. 2. Prototype system architecture.

The data delivery chain reduces the amount of data and aims to deliver only relevant information from sensors to end-users. Intelligent data delivery requires special software both on board and on the server side to deliver the right information to the right place at the right time. With a sampling rate of 20 MHz and a 12 bit A/D conversion stored in 2 bytes, one digitized radar signal stream is in the order of 40 Mbytes/s, which is not feasible to deliver from ships over a mobile communications link. The prototype system forms PPI (Plan Position Indicator) images from the radar signal and combines these into radar image mosaics that are then delivered. The size of the images and the delivery frequency, thus the amount of information, can be adjusted according to the data communication capabilities. One way to reduce the data stream volume is to deliver only some relevant features of the images and not the images themselves. For example, ice drift calculation from subsequent radar images can be seen as a lossy method to compress the information in the images. By delivering only the ice drift vectors, the transferred data volume is reduced significantly.

3. Ship performance analysis

Ships travelling through ice are affected by the resistance of the ice field. Their performance, i.e., the speed vs. thrust delivered by the propulsion system, can thus be used as an indicator of ice resistance. The speed of a ship in ice is subject to variations in the encountered ice thickness and mechanical ice properties such as bending strength and crushing strength. The thicker the ice, the more thrust a ship needs in order to maintain a certain speed. Ice compression (caused by a moving ice field) and different formations in the ice such as level ice, ridged ice, ice channels, and ice floes have different effects on the ship's travel.

The idea of using ships as sensors for trafficability in the ice field is based on findings presented earlier, e.g., Riska et al. (1997) have shown that the theoretical speeds of different types of commercial ships correlate as a function of the level ice thickness. It could therefore be possible to estimate the performance of one ship based on the performance of another. A ship transit model estimates how a ship performs in different types of ice conditions, i.e., what speed the ship can achieve with a given thrust. A transit model can be considered as a function f that associates ice conditions c, static ship parameters s (length, breadth, maximum engine power, etc.), non-static ship parameters d (draught, trim, etc.), and non-static ship propulsion system parameters p (used engine power, propeller pitch, etc.) with ship speed v.

$$v = f(\mathbf{c}, \mathbf{s}, \mathbf{d}, \mathbf{p}) \tag{1}$$

The method of using ship performance observations to estimate trafficability is based on an assumption that the transit model f is invertible, i.e., that the ice conditions c can be derived from the ship speed v and the parameters s, d, and p.

$$\mathbf{c} = f^{-1}(v, \mathbf{s}, \mathbf{d}, \mathbf{p}) \tag{2}$$

Let f_A be an invertible transit model of a ship A and let f_B be a transit model for a ship B. The speed that ship B would then achieve in the area in which ship A is travelling can be estimated using the observed parameters of ship A and the assumed propulsion system parameters of ship B as follows:

$$v_B = f_B\left[f_A^{-1}(v_A, \mathbf{s}_A, \mathbf{d}_A, \mathbf{p}_A), \mathbf{s}_B, \mathbf{d}_B, \mathbf{p}_B\right] \tag{3}$$

where v_B is the estimated speed of ship B, v_A is the measured speed of ship A, s_A and d_A are the static and non-static ship parameters of ship A, p_A is the measured propulsion system parameters of ship A, s_B and d_B are the static and non-static ship parameters of ship B, and p_B is the assumed propulsion system parameters of ship B. All measurements contain uncertainties, and the uncertainty of the speed estimate should be analyzed in order to assess the uncertainty of the trafficability estimate. The uncertainty of v_B can be derived as:

$$u(v_B) = \sqrt{\sum_{i=1}^{n} \left(u(x_i)\frac{\partial v_B}{\partial x_i}\right)^2} \tag{4}$$

where $u(v_B)$ is the uncertainty of v_B, x_i is all the variables in Eq. (3), and $u(x_i)$ is the uncertainties of variables x_i (ISO/IEC Guide 98-3, 2008). Depending on the transit model, an analytical solution to Eq. (4) may not be available and numerical approximations must be applied.

In practice, the ice field is stochastic by nature, and the ice conditions as well as the ship speed can vary considerably over a relatively short distance. The ship's inertia also affects the speed, i.e., the ship decelerates when hitting, for example, a heavy ice ridge, and accelerates after penetrating the ridge. In that sense, a single momentary speed observation is not enough to determine trafficability in the area in which the ship is travelling, and the ship speed and ice condition estimates should be averaged over a sufficiently long distance. Speed variance should not be neglected, however, as it gives additional information about the ice field. Information about ships becoming stuck (i.e., $v=0$) is especially valuable to other ships approaching the area. In order to apply the trafficability information derived from observations collected from one ship to other ships in the area, we assume that the ice field consists of regions of similar ice conditions, i.e., an observation from one location represents the trafficability of a larger area. One challenge, however, is how we define optimally the regions of similar ice conditions using segmentation or clustering. The smaller the segmentation area, the more limited is the area of applicability based on observations from one ship. A larger area includes more statistical variations within the area, which lessens the correlation between different ships.

The prototype system utilizes AIS and IBNet data as input to trafficability estimation. The AIS enables the collection of ship speed and location information in near real time from all the ships using the AIS transponder, and IBNet provides information about icebreaker assistance activities, which helps with filtering of AIS data. The usage of AIS data has limitations. Although AIS data are available from almost all the ships, they do not contain propulsion system parameters and thus information about the true thrust produced by the propulsion system. The prototype system must assume that the ships are using full or almost full thrust, which introduces an additional error source in the trafficability estimation. A method to estimate the trafficability by calculating relative speeds from the AIS data is introduced in Section 3.1.

Due to the limitations of the available input data, we have to use a transit model, which does not necessarily require real-time observations of the propulsion system parameters. The prototype system utilizes a simple invertible transit model that can manage the deficiencies of the input data, called the *ICE–v* curve transit model. This model represents a simple mapping between the ship speed and a scalar value representing the ice conditions. Here, the scalar parameter is called the ice conditions equivalent (ICE) and represents all relevant ice field properties, such as level ice thickness, ice concentration, ice compression, etc., that affect the ship's travel using a scalar value. The determination of the *ICE–v* curve transit model by utilizing *h–v* curves is described in Section 3.2 and by statistical estimation based on AIS data in Section 3.3.

3.1. Trafficability estimation based on the relative speed

The prototype system allows the end-user to estimate the trafficability by using the relative speed of the ships. The relative speed is the ship speed divided by its designed open water speed (or maximum speed). The usability of the relative speed in trafficability estimation is based on the assumption that the ships use full or almost full constant thrust when going in ice and earlier theoretical findings by Riska et al. (1997) that the relative speeds of different types of ships correlate in similar ice conditions. Even though this method is very simple, it has some advantages. It only requires ship speed as a performance observation and can thus be applied to all ships using the AIS. The AIS data have to be preprocessed by filtering out ships that are not anticipated to use full throttle or that receive icebreaker assistance. The prototype system visualizes the relative speed to the user as colored ship trails (Fig. 3). By observing the colors and using his experience of ice-going ships, the user can estimate trafficability in different sea areas.

The method has some disadvantages. The ships do not use full throttle all the time but adjust the thrust according to the ice conditions. In addition, when the ships approach a port or receive icebreaker assistance, their speed does not correlate with the ice conditions. Real-time thrust information would give additional information but is not achievable with the current system. During special test campaigns, it was possible to record ship performance data in more detail, but it is not possible to access performance data of multiple ships in real time. Some ships have performance data logging systems, but these data are stored locally and are not available in real time over a mobile link.

3.2. Determining the *ICE–v* curve using *h–v* curves

The *h–v* curve presents the ice thickness (h) in relation to the ship speed (v), which is unique for every ship as it is depending on the ship's properties. The *h–v* curve represents the capability of a ship to obtain a particular speed in a particular level ice thickness and is determined from the estimated resistance of the ice and the thrust delivered by the propulsion system. In this section, we show how *h–v* curves are determined and utilized as *ICE–v* curves to estimate trafficability.

The procedure for determining the *h–v* curve is explained by way of an example on *R/V Aranda*. The thrust of *Aranda* is adjusted by the variation of the propeller pitch. The shaft RPM (revolutions per minute) is kept constant. It is important that the machinery system of the ship is understood, i.e., the power delivered to the propeller including the mechanical losses in gears, bearings, and other parts.

The open water performance was assessed by varying the propeller pitch in open water during a cruise from the Bay of Bothnia to the Gulf of Finland. The performance was assessed in the IM (ice mode), which means that both engines were connected to the propeller shaft. In open water, the data obtained for the vessel speed and power of the propeller shaft have been used to determine the net thrust curve (Kujala and Sundell, 1992).

$$T_{\mathrm{pull}} = K_e \left(D_p P_{\mathrm{sh}} \right)^{\frac{2}{3}} \tag{5}$$

$$T_{\mathrm{net}} = T_{\mathrm{pull}} \left(1 - \frac{1}{3} \frac{v}{v_{max}} - \frac{2}{3} \left(\frac{v}{v_{max}} \right)^2 \right) \tag{6}$$

where v is the actual ship speed, v_{max} is the maximum speed, T_{pull} is the bollard pull, K_e is a factor including the number of propellers (0.78 for single screw propulsion like *Aranda*), D_p is the propeller diameter, and P_{sh} is the total shaft power.

The net thrust curves for dedicated pitch positions of the *Aranda* are determined according to Eq. (6). The performance in ice is related to the performance in open water, and the operational limits in level ice are determined by the intersections between the net thrust curves and the theoretical determined ice resistance for various ice thicknesses.

The ice resistance in level ice is determined according to an empirical formula developed by Riska et al (1997), which is based on the findings by Ionov (1988) and Lindqvist (1989). The ice resistance is linear dependent on the ship speed.

$$\begin{aligned} R_i &= C_1 + C_2 v \\ C_1 &= f_1 \frac{1}{2\frac{T}{B} + 1} B L_{\mathrm{par}} h_i + (1 + 0.021\phi)\left(f_2 B h_i^2 + f_3 L_{\mathrm{bow}} h_i^2 + f_4 B L_{\mathrm{bow}} h_i \right) \\ C_2 &= (1 + 0.063\phi)\left(g_1 h_i^{1.5} + g_2 B h_i \right) + g_3 h_i \left(1 + 1.2 \frac{T}{B} \right) \frac{B^2}{\sqrt{L}} \end{aligned} \tag{7}$$

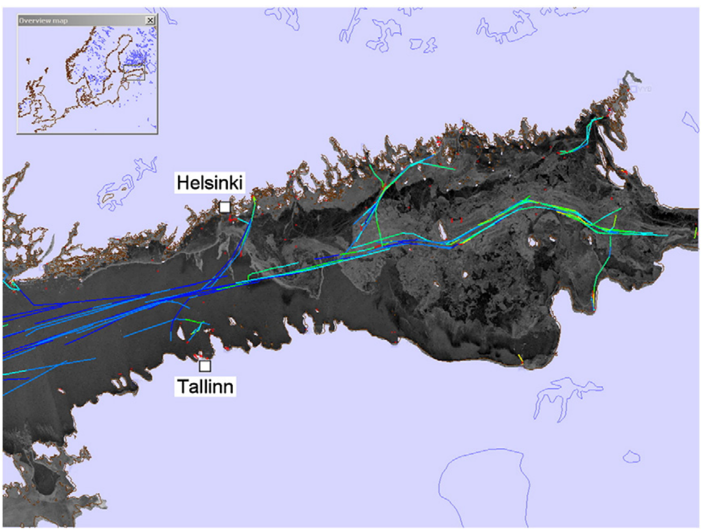

Fig. 3. Relative ship speeds visualized as colored ship trails on 3 March 2009 in the Gulf of Finland. Blue indicates easy going; green, that there are some difficulties; and yellow and red, that the ships are struggling to move onwards. The background image is taken by Radarsat satellite. The amount of traffic in the Gulf of Finland enables many ship speed observations for trafficability evaluation. The open water part on the left appears to be easy going, and the ice field closer to St. Petersburg on the right is harder to travel in. Note that the interpretation is based on ship speed only, so single observations may not be reliable, as ships could be slowing down because of other reasons than ice.

Table 1
Constants for ice resistance.

f_1	$0.23\frac{kN}{m^3}$
f_2	$4.58\frac{kN}{m^3}$
f_3	$1.47\frac{kN}{m^3}$
f_4	$0.29\frac{kN}{m^3}$
g_1	$18.9\frac{\frac{kN}{m}}{\left(\frac{m}{s \times m^{1.5}}\right)}$
g_2	$0.67\frac{\frac{kN}{m}}{\left(\frac{m}{s \times m^{2}}\right)}$
g_3	$1.55\frac{\frac{kN}{m}}{\left(\frac{m}{s \times m^{2.5}}\right)}$

The symbols used in Eq. (7) are explained in Appendix A, and the constants are given in Table 1. The constants have been developed by Riska et al. (1997) using the performance analysis of several different ships with the target to use the constants for determining the ice resistance of other ships. The plots in Fig. 4 represent the available net thrust, and, finally, the intersection between the net thrust curve and the resistance curve represents the limiting speed in the particular ice thickness (Fig. 5). The intersections between the resistance curve and the net thrust curve reflect the limiting ice thickness for a particular propeller pitch position. The intersections of the curves represent the *h–v* curve, which reflects for which particular ice thickness and speed the particular propeller pitch reaches its limit.

When *h–v* curves are used in trafficability estimation, the ice thickness is considered as the ice conditions equivalent value (ICE), i.e., phenomena such as ridges, rafted ice, and ice compression are taken into account as thicker ice. Compressions and ridges cause higher resistances in ice which either slow the vessel down or demand a higher power output from the engine to maintain a certain speed. If a ship is travelling in constant ice thickness with constant thrust and then encounters a speed reduction, the *h–v* curve would give a higher ice thickness since the *h–v* curve is only related to level ice. Since the thus determined ice thickness is not real but a combination of the real ice thickness and other phenomena, it is called ice conditions equivalent (ICE) and in this case presented in meters (m).

Fig. 6 shows the procedure, for example, of *RV Aranda*. The *ICE–v* curve is derived from *h–v* curve and plotted for the four different pitch positions, which are stated as percentages of the full pitch. At 87% propeller pitch, a speed of 4 m/s is obtained, and then the ship speed is reduced to 3.2 m/s although the real ice thickness does not change, but other phenomena such as compression might be encountered. According to the curve, the ICE would change from 0.21 m to 0.3 m.

Utilizing *h–v* curves *ICE–v* curve transit models has some disadvantages. The *h–v* curve has to be determined separately for each ship, and estimating different ice types as ice thickness introduces an additional error component in the calculations. An ice field rarely consists of only pure level ice but is a mixture of level ice, rafted ice, ridges, channels, open water, etc. In addition, the *h–v* curve assumes a constant RPM and

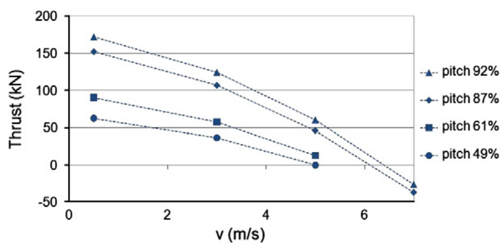

Fig. 4. A theoretical net thrust curve for *RV Aranda*.

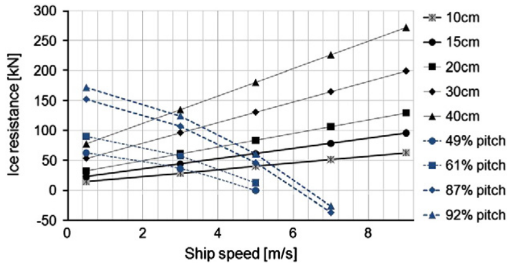

Fig. 5. *RV Aranda*'s ice resistance (straight lines according to the particular ice thickness) versus net thrust is dependent on the pitch position of the propeller blades in percentage.

is highly dependent on propeller pitch, as shown in Fig. 7. Thus, more accurate use would require observations of the propulsion system (RPM, pitch) to determine the thrust and the relevant *ICE–v* curve to be used.

3.3. Determining *ICE–v* curves statistically using AIS data

Preliminary work was done to determine *ICE–v* curves statistically using the AIS, the Ibnet system, and the ice chart data. A data set was collected that includes the ship speed, time, location, level ice thickness and ice concentration (in the ship's location in space and time), and information about the closest port and icebreaker activities. Ice thickness and concentration were taken from the ice chart, ship speed and location came from the AIS, and the icebreaker activity information was from the IBNet system. Locations in which the ships would probably not use full engine power were removed by selecting only the data points in which the ships were not assisted by an icebreaker and were not close to any port. From these data, the ship's relative speed and the ICE value at the same location were calculated for six cargo ships. The ICE value was derived from the level ice thickness and the ice concentration using the following equation:

$$ICE = \begin{cases} 0 & ; \quad C \le 70\% \\ \frac{(C-70\%)h_i}{(95\%-70\%)} & ; \quad 70\% < C < 95\% \\ h_i & ; \quad C \ge 95\% \end{cases} \tag{8}$$

where h_i is the level ice thickness and C is the ice concentration (the ratio of the areal extent of ice present and the total areal extent of ice and water).

A linear model was fitted to the data using least squares method. The model estimates the ICE value given the relative ship speed. We assume that the ICE value is 0 when the ship goes at its open water speed, i.e., when the relative speed is 1, thus the model becomes:

$$ICE = k(v_r - 1) \tag{9}$$

where ICE is the equivalent ice thickness, k is a ship-related coefficient, and v_r is the relative speed. The model is based on the

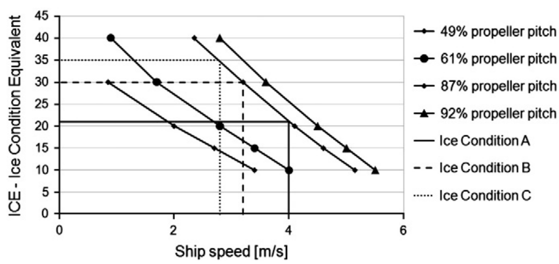

Fig. 6. *RV Aranda*'s *h–v* curves for different propeller pitches utilized as *ICE–v* curves.

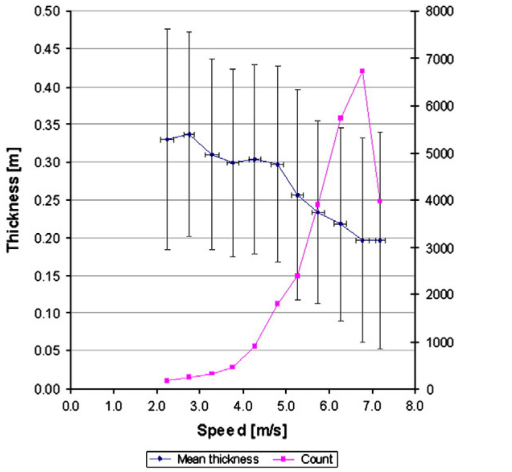

Fig. 7. The properties of an example data set derived from the AIS (ship speed and location), the Ibnet system (icebreaker assistance data for filtering), and ice charts (ice thickness) for cargo ship *Rautaruukki*. The error bars indicate the variability of the ice thickness with the same ship speed. The data point count indicates the amount of data samples as function of speed.

assumption that the ships are using maximum or almost maximum thrust to penetrate the ice, and the only factor for variations in the ship speed is the variation in ice field properties, i.e., the ice conditions equivalent (ICE). One-third of the data were used in the model fitting, and two-thirds were used to calculate the root–mean–squared error (RMSE) of a model. The ship-related coefficients k and the errors of the fitted models for the six ships are given in Table 2.

There are a number of uncertainties in the data set that explain the error of the models. The ships may not use full throttle all the time as assumed, the spatial and temporal resolution and accuracy of the ice charts are limited, and speed measurements contain an error component due to inaccuracies in the AIS data. There are also more samples for the higher speeds, which is due to the mild winters of recent years, i.e., there are fewer data samples representing ships moving in thick ice at slower speeds. To illustrate the amount of noise in the data, Fig. 7 shows the level ice thickness and its variation in relation to the ship speed for the cargo ship *Rautaruukki*.

Deriving *ICE–v* curves statistically from AIS data allows curves to be determined for all ships utilizing the AIS. However, to reduce the amount of uncertainty in the curves, this method would also require observations from non-static ship propulsion system parameters, such as the propeller pitch, RPM, and used engine power.

4. Radar images

Terrestrial marine radars as well as space-borne satellite radars are used to observe ice field properties for navigation, research, and long-term surveillance of ice field development. There are also on-board systems that utilize specialized algorithms to enhance ice features on the radar image to improve visual interpretation.

The prototype system digitizes marine radar images, forms an image mosaic from subsequent images if the radar is moving, delivers the images to the server side for additional processing, calculates ice drift from the images, and visualizes them to end-users. The aim is to take advantage of the whole radar sensor network consisting of coastal and ship-borne radars to gain more information about the ice field.

Table 2
Results of the model fitting (Eq. (9)) for six cargo ships. The ship name, the MMSI number (Maritime Mobile Service Identity), the width and length of the ship, the ship-related coefficient (k in Eq. (9)), and the root–mean–squared error (RMSE) are given.

Ship name	MMSI number	Width	Length	k	RMSE
Steel	230202000	27 m	167 m	−0.70	0.16 m
Rautaruukki	230358000	27 m	167 m	−0.84	0.16 m
Birka Express	230366000	22 m	154 m	−0.90	0.18 m
Birka carrier	230367000	22 m	154 m	−0.85	0.17 m
Birka trader	230368000	22 m	154 m	−0.90	0.18 m
Nemuna	304475000	14 m	90 m	−0.66	0.17 m

4.1. Image capture

The prototype system utilizes a radar video server developed by the Finnish company Image Soft Oy. The radar display used on board does not necessarily contain all the information available in the radar signal for detecting ice features. The radar video server captures the radar image before the signal is filtered by the display system. In theory, this should provide the maximal amount of information available in the signal, but in practice, the sampling frequency limits the information bandwidth. The radar video server was operating at a 20-MHz sampling rate.

The video server is based on PC technology and forms PPI (Plan Position Indicator) images from the radar signal, the triggering pulse, and the antenna pulse (Fig. 8). The radar signal contains information about echo intensities and distances, while triggering and antenna pulses indicate when and in which direction the radar signal was sent. The signal interface unit is required to adjust the signal levels between the radar and the video server. Images are stored on a local hard disk and delivered onward using a TCP/IP connection. The raw signals can also be stored for advanced analysis later on, although this requires a large amount of disk space (40 Mbytes/s) and, in practice, it is only feasible to store a few minute-long radar signal snapshots.

The prototype was tested on the Raahe coastal radar station in the winter of 2008 and on board the research vessel *Aranda* and icebreaker *Otso* in the winter of 2009. The radars were X-band radars, and their detailed properties are given in Table 3.

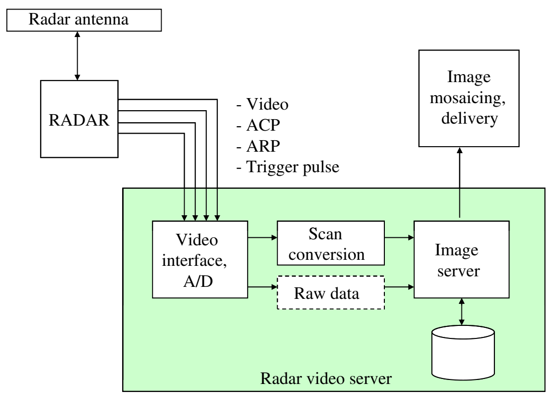

Fig. 8. Radar video server components. The radar signal consists of the analog radar video and three additional pulsed signals: ACP (Azimuth Change pulse), which registers the antenna rotation angle; the ARP (Azimuth Reset pulse), which is sent once per revolution; and the Trigger pulse, which is synchronized with the pulse start. The Video interface combines the digitized video signal with the pulse signals into 16-bit words at a 20 MHz rate. These samples are scan converted into a rectangular coordinate system before storing as image files in the image server. From the image server, the images are delivered to mosaicing and then delivered to server side and viewed locally on board. Optionally the raw digitized video can be stored as a file, but the data rate for this storage mode is 40 Mbytes/s. The raw digitized video data are used mainly for debugging purposes.

Table 3
Radar properties.

Radar	Antenna	Height over water level	Pulse power	Frequency [MHz]	Pulse width [ns] (short, medium, long pulse)	Pulse repetition frequency (PRF) [Hz]	Rotation period [s]
Raahe pilot station—coastal radar	Parabolic antenna (Aspo AES404), beamwidth –horizontal: 0.5 degrees, –vertical 4 degrees	45 m	25 kW	9410 ± 30 9375 ± 30	50, 250, 600	800–2000	2.6
RV ARANDA	XN-24AF, 8′, beamwidth –horizontal: 0.95 degrees, –vertical 20 degrees	20 m	25 kW	9410 ± 30	70, 150, 300	1000–3000	2.47
IB OTSO	12′, beamwidth –horizontal: 0.6 degrees, –vertical 12 degrees	30 m	25 kW	9345–9405	60, 250, 800	3000 (SP), 1500 (MP), 750 (LP)	2.75

4.2. Image mosaicing

Individual radar images contain noise and only cover a limited spatial area within the radar range (Fig. 9). When the ship is moving, however, images are available from a larger area, and if radar images are captured from multiple moving ships, an even larger area can be covered. Mosaicing is one image processing techniques that is useful for combining images from overlapping spatial areas, resulting in one large image in which the boundaries between the original images are not visible. The prototype forms image mosaics from multiple ice radar images (Fig. 10). When the ice field is stationary, a mosaic formed along a ship track of many hours can be useful for navigational purposes.

A high-quality mosaic requires the individual source images to be visually homogeneous. Considering this requirement, the main difficulties when building mosaics from ship radar images are (1) strong range dependence of the signal, (2) signal noise, (3) sensitivity to rasterization errors due to erroneous position or orientation information, and (4) location errors due to non-stationary ice causing multiple contours or blurring in the final mosaic.

The range dependence comes from the fact that the power of the radar echo decreases proportionally to $1/R^4$ of a point target, or as $1/R^3$ for the backscatter from a distributed target like sea clutter or ice, where R is the distance to the target. Furthermore, the backscatter coefficient σ^0 varies as a function of the incidence angle, approximately $1/R$ or $1/R^2$, causing the signal power to have a net range dependency of $1/R^4$ or $1/R^5$ (Marton, 1999). Depending on the gain settings, the range dependency may also cause saturation of the signal from objects close to the ship, thus hiding interesting ice structure information. The signal is usually acquired via a logarithmic amplifier, offering a higher dynamic range of the digitized signal. The ship radars also contain other signal filters and gain adjustments, which make it difficult, in practice, to calculate any calibrated backscatter coefficient σ^0 for the target.

The second problem, signal noise, can be overcome by filtering several images before adding them to the mosaic. In the systems we used, the most annoying noise was caused by interference from other ship radars on board, causing very strong arc-formed lines on the images. A median filter applied to a sequence of images is a very efficient way of enhancing the signal (backscatter from rough ice surface) and suppressing spurious noise. Rasterization errors due to missing or faulty ship heading information or a missing antenna pulse cause the orientation of the whole or part of the image to change rapidly, making, for example, ice drift calculations impossible. If it only applies to a minority of the images, this problem can also be significantly reduced by a median filtering technique.

The last problem – drifting ice – determines the maximum time span for gathering overlapping images, although the images are properly shifted according to the position of the ship. A typical ice drift speed of 0.3 knots would cause the position to change 9 m in 1 min, i.e., for high-resolution radar images (7.5 m/pixel), the time span for individual images to be included in a mosaic for a sharp image mosaic is of the order of a minute.

The mosaicing method utilized in the prototype is depicted in Fig. 11. The images were captured in bursts of 8 to 10 images at 10-min intervals. The images in one burst were combined using median

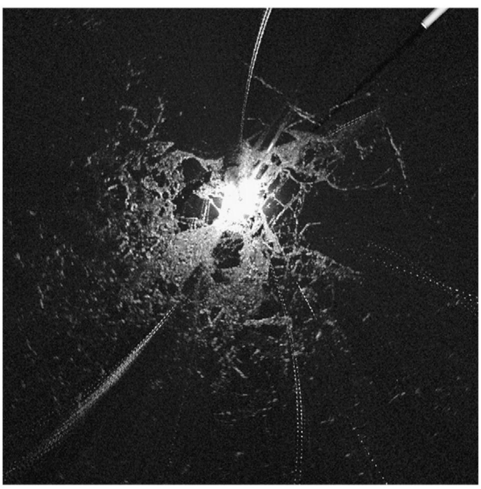

Fig. 9. A radar image (9 km × 9 km) rendered from a single radar antenna rotation. Some artefact signals and noise from other radars and electric devices on board are visible.

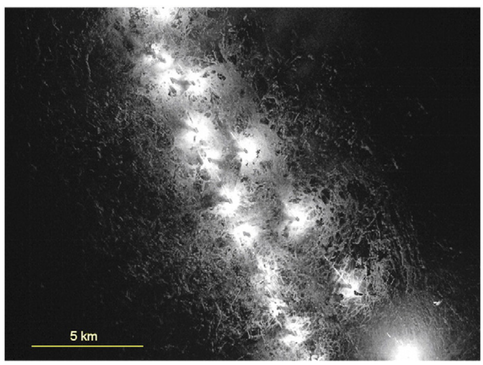

Fig. 10. An example image mosaic formed from radar images captured on board icebreaker *Otso* in the Gulf of Bothnia. Two tracks are merged in the image as *Otso* first moved south and later travelled north again using an almost parallel route. The bright spot in the lower right-hand corner is an artefact created by the mosaicing algorithm when *Otso* was stationary for some hours.

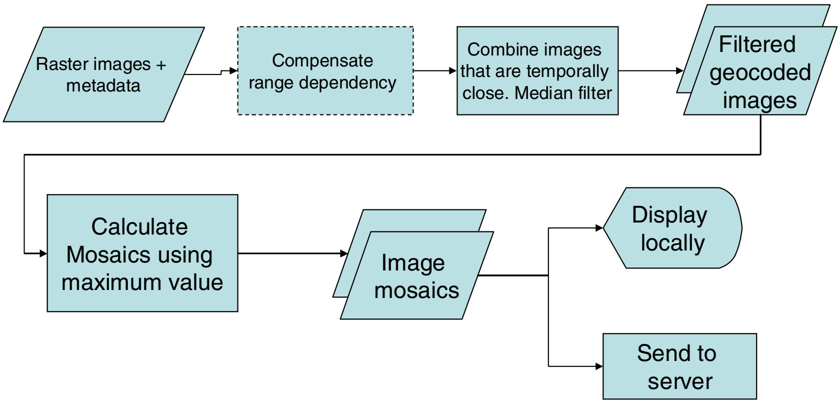

Fig. 11. Image mosaicing method. The prototype implementation does not use range compensation (marked as dashed box in the figure).

filtering, and the resulting filtered image was added to the total mosaic by selecting the maximum value for each corresponding pixel. The problem of intensity saturation of near-range pixels was solved by defining a threshold above which the pixel was considered transparent when combining the image with the rest of the mosaic. Visualization experiments using artificial colors were performed to illustrate changes in the ice field—changes caused either by ice drift or by different viewing angles.

4.3. Ice drift estimation

Ice thickness is not the only factor affecting ship performance in ice. Ice compression caused by a moving ice field is also a major factor. In order to estimate or predict the ice compression, observations about ice field movements are required. Different methods have been used to observe ice drift, such as buoys (Heil et al., 2001, Inoue et al., 2009, Rampal et al., 2008), marine coastal radars (Sun et al., 2004, Mahoney, et al., 2007), camera images (Leisti et al., 2009), and successive satellite images (Karvonen et al., 2008, Gutierrez and Long, 2003, Thomas et al., 2009). The prototype system uses both visual interpretation and numerical computation to determine ice drift from marine radar images.

4.3.1. Visual interpretation of ice drift

The user can view an animation of successive radar images from a stationary ship or coastal radar, or visually compare marine radar image mosaics with each other and satellite radar images. An experimental color-coded image mosaic is available. The colored mosaic is formed by rendering different individual radar images from different times to different RGB channels (Fig. 12). As a result, static parts in the image are rendered in black and white while moving parts are colored. As the time difference between the RGB channels and the projection of the image are known, ice drift can be estimated by calculating the distance between the red, blue, and green images of a moving ice feature.

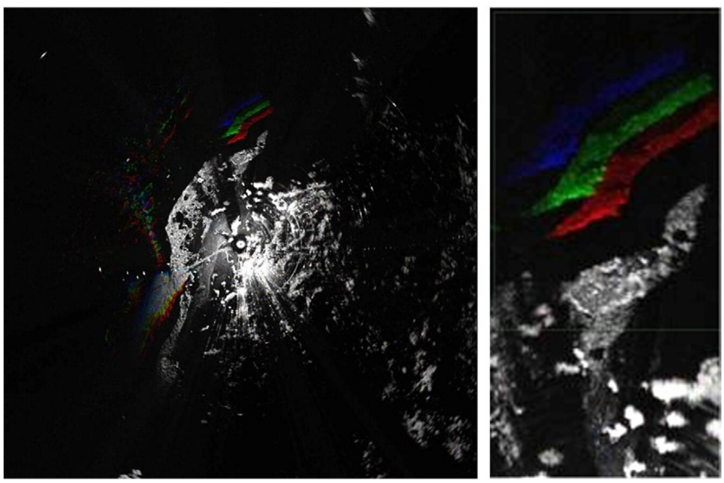

Fig. 12. An experimental color-coded radar image mosaic formed from images taken from the Raahe coastal station (64.667 N, 24.407E). Image on the left is the whole image mosaic (30 km × 30 km), while the image on the right (5 km × 10 km) is a magnification of a moving ice floe. The land is located on the right side of the radar, while ice is detected on the left side. Three successive radar images taken at 1-h intervals are combined using red, green, and blue channels in a sequence. As a result, static parts of the image appear in black and white while moving ice floes show up in colors.

4.3.2. Numerical computation of ice drift

In addition to the visual interpretation, numerical ice drift computation is also performed. Numerical values can be presented to users as vector fields or utilized in the assimilation or validation of ice models. Methods to determine ice drift from subsequent marine radar images (Sun et al., 2004) and satellite radar images (Karvonen et al., 2008) have been presented earlier. The prototype system utilizes a method developed at the Finnish Meteorological Institute. The method consists of preprocessing and ice drift detection based on a phase correlation method (Karvonen et al., 2008).

The preprocessing includes ship motion compensation, temporal median filtering, noise reduction, and edge detection. In the case that the radar is moving, i.e., if it is on board a moving ship, the ship motion must be compensated for. Temporal median filtering is applied to remove random noise caused by other radars and electrical devices on board. Anisotropic median filtering (Karvonen, 2009) is also applied to remove small bright spots, which are probably echoes from ships that could be moving and could lead to misinterpretations. Finally, edge detection is performed using the Canny edge detection algorithm (Canny, 1986). The Canny algorithm is based on a gradient operator and hysteresis thresholding utilizing two thresholds. The algorithm detects all the locations containing edges in a radar image. Because the motion detection method used here is sensitive to edges, it is not reasonable to try to locate drift for areas with no edges or very few edges. Fig. 13 shows an unfiltered image, median-filtered image, and a motion-compensated temporal-median-filtered image.

In the ice drift estimation, data windows of fixed size are extracted from the two images between which the motion is to be estimated. The data windows are arrays of pixel values of a given size (N by N pixels, a typical value of N is 16), and they are sampled by sliding a window over the images. The data windows of corresponding image location in the two images are used in the drift estimation. It can be assumed that the drift is within the data window by selecting a large enough window size, which is dependent on the temporal difference of the two radar images.

Phase correlation [Kuglin & Hines, 1975] is a more robust approach than using cross correlation between the data windows. Cross correlation can be computed by applying a Fast Fourier Transfrom (FFT) to the two data windows of the two images, performing multiplication between the transformed data windows, and then performing Inverse FFT (IFFT) to yield the cross correlation array. In the phase correlation approach, the FFT transform coefficients are normalized to unit magnitude before the correlation in the frequency domain, thus the correlation is based only on the phase information and is less sensitive to differences in image intensities.

Phase correlation is applied to preprocessed images but only to areas containing detected edges. This makes the search faster and more reliable, as phase correlation works best in edged areas. Phase correlation is less sensitive to intensity variation than the commonly used cross correlation, and it is sensitive to the edges present in the images. The 2D fast Fourier transform (FFT) is applied to the data window, the FFT coefficients of the two data windows are normalized by their magnitudes, and the FFT coefficients of the two data windows are then multiplied and the inverse 2D FFT applied. is the phase correlation image computed from the normalized cross-power spectrum, and the 2D motion vector $D = (\mathrm{d}x, \mathrm{d}y)$ can be estimated as:

$$\begin{aligned} D &= (\mathrm{d}x, \mathrm{d}y) = \arg\max_{(x,y)}\left\{I_p(x,y)\right\} \\ &= \arg\max_{(x,y)}\left\{FFT^{-1}\left(\frac{(X_1^*(k,l)X_2(k,l))}{|X_1^*(k,l)X_2(k,l)|}\right)\right\} \end{aligned} \quad (10)$$

As the FFT assumes that the data are periodic, a Gaussian window is applied to the data windows before the transformation. The best matching displacement is then defined by the maximum of the phase correlation. The search for the best local phase correlation is performed in a multiresolution image pyramid in a recursive manner starting from the lowest resolution level to the highest. For each level, two branches (two I_ps) are studied recursively: the I_p computed when the other image is moved according to the highest correlation at the previous resolution level and the I_p, assuming no motion has occurred. Windows with a lower phase correlation than a given threshold are omitted. Finally, vector median filtering is performed with a given radius R_m ($R_m = 3$ in this study) to obtain the magnitude and direction of the motion. Fig. 14 shows an ice drift vector field calculated from radar images taken on board a stationary ship, and Fig. 15 shows an ice vector drift field when the ship is moving.

4.3.3. Computing divergence and curl from ice drift

The compression in ice caused by a moving ice field is a challenge for ships travelling in ice. The amount of compression can be estimated by calculating the divergence of the ice drift vector field. The divergence div(D) of a vector field $D = (\mathrm{d}r, \mathrm{d}c)$ is defined as

$$div(D) = \nabla \cdot D \quad (11)$$

where dr and dc are the motion in row and column directions, respectively. A simple discrete estimate for divergence is

$$div(D) = \frac{1}{2}[\mathrm{d}r(i+1,j) - \mathrm{d}r(i-1,j) + \mathrm{d}c(i,j+1) - \mathrm{d}c(i,j-1)] \quad (12)$$

The ice field converges when div(D) is negative, i.e., compression is presented.

The curl for a vector field can be computed as

$$curl(D) = \nabla \times D \quad (13)$$

and the corresponding discrete estimate for curl magnitude is

$$|curl(D)| = \frac{1}{2}[\mathrm{d}c(i+1,j) - \mathrm{d}c(i-1,j) + \mathrm{d}r(i,j+1) - \mathrm{d}r(i,j-1)] \quad (14)$$

The direction of the curl vector is the direction of the normal of the vector field. An example of calculated divergence and curl is given in Fig. 16.

5. Discussion and further work

In this paper, we have described the concept of utilizing ice-going ships as a sensor network to collect relevant information automatically from the ice field to improve winter traffic. A prototype system has been built that collects ship performance observations and ship radar images to analyze trafficability in different sea areas, create radar image mosaics, derive ice drift from subsequent images, and present these new sources of information to end-users in combination with traditional ice products such as satellite images, ice charts, and weather and ice model forecasts. Based on the prototype test results, it seems that such a system would provide additional information on prevailing ice conditions.

Ship performance observations and ship transit models can be used to determine trafficability in different sea areas to some extent. In the prototype system, trafficability is estimated based on near real-time access to ship speed information of all ships via the AIS, and the icebreaker assistance activity information via the IBNet system. *H–v* curves can be used as transit models to determine trafficability. They are determined by taking into account the ship's physical properties. While using *h–v* curves, the current system needs to make assumptions about the used engine power and propeller pitch. In the future, there should be real-time access to non-static propulsion system parameters in order to determine the thrust delivered by the propellers. This would enable more accurate estimations of the

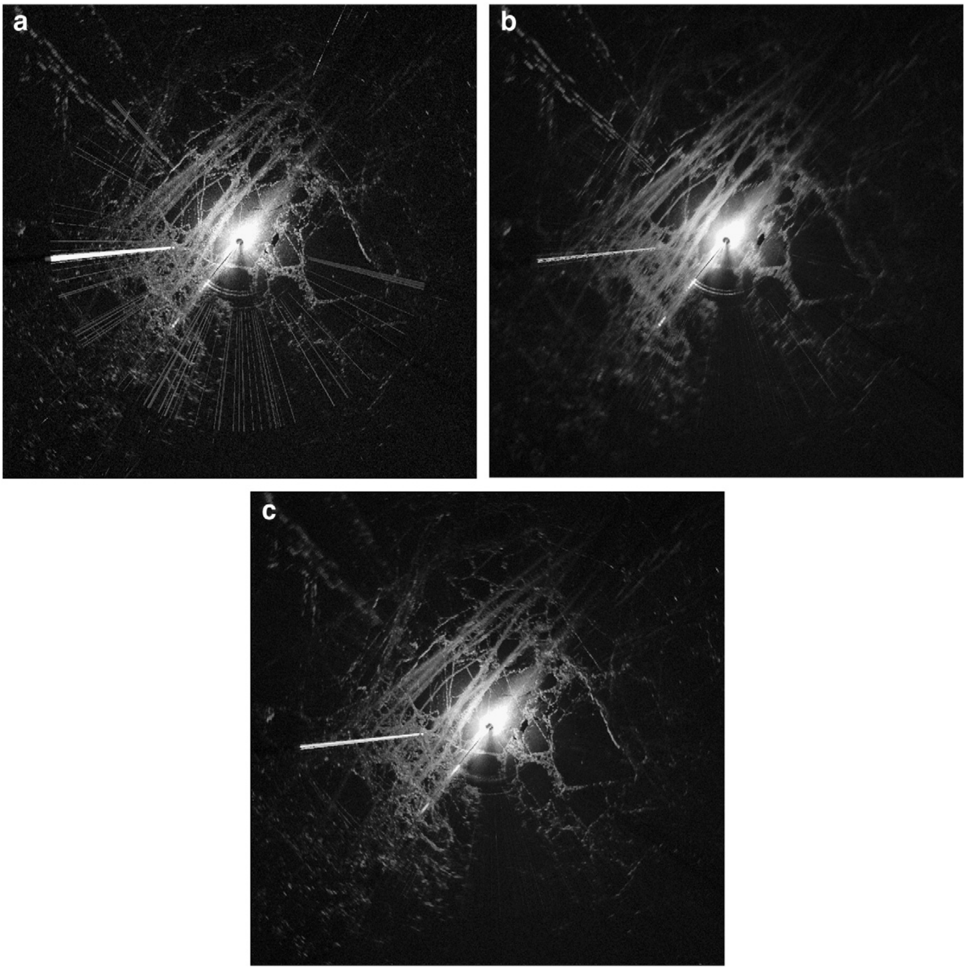

Fig. 13. An unfiltered radar image (a), temporal median-filtered image without motion compensation (b), and temporal median-filtered image with motion compensation (c). The image pixel resolution is 7.5 m, and the sizes of the images are 9 km × 9 km.

resistance caused by the ice field. The propulsion system parameters would also be helpful in statistical determination of *ICE–v* curves. Preliminary work was conducted by fitting a linear model to speed observations and ice chart data, but additional information about the propulsion system would reduce some of the uncertainties and improve the model usability in determining the trafficability.

Coastal and ship-borne radar images can be used to complement satellite images in limited areas when the ice field is changing rapidly. Radar images are not an alternative to satellite images but they can be utilized in data fusion with these. Image mosaics formed from ship-borne radar images from many ships cover a larger area and update the information in the satellite images. Further work should concentrate on improving the mosaicing algorithm to obtain as homogeneous a mosaic as possible, determining optimal ways of detection and compensation of faulty heading information before mosaicing, determining ice field parameters from the ship radar image (deformed ice, ridges, etc.) and optimal ways of presenting ice drift vectors in combination with satellite images, and assimilating ice drift observations in ice drift models to obtain improved short-term forecasts.

Automatic ice drift monitoring is possible using marine radar imagery collected from both stationary coastal radars and moving ship radars. Information about ice drift is important to mariners as drift opens leads in the ice, causes compression in the ice, and forms ice ridges. Experiments were performed using artificial colors to visualize ice drift. Using colors in this way is not intuitive to the mariner, however, and a combination of a gray-scale image and ice drift vectors therefore seems to be the optimal way of presenting the information from marine radars. The ice motion algorithm has been validated for SAR data by buoy measurements (Karvonen et al., 2008), but ice drift calculated from radar data has not yet been validated numerically. Only comparisons to visual interpretation have been made by animating a series of radar images and showing the estimated ice motion field in parallel. However, although automatic methods would not produce exact numerical values, additional visual information about the ice drift direction and magnitude was valuable for estimating the compression. In the future, the ice drift should be validated in more detail with, e.g., ice buoy measurements, ice drift derived from some independent source, and ice model predications.

The prototype system was tested using a single marine radar capture device. The system should also be tested with multiple ships equipped with the image capturing and mosaic processing system. This would allow mosaicing of images from the same time but

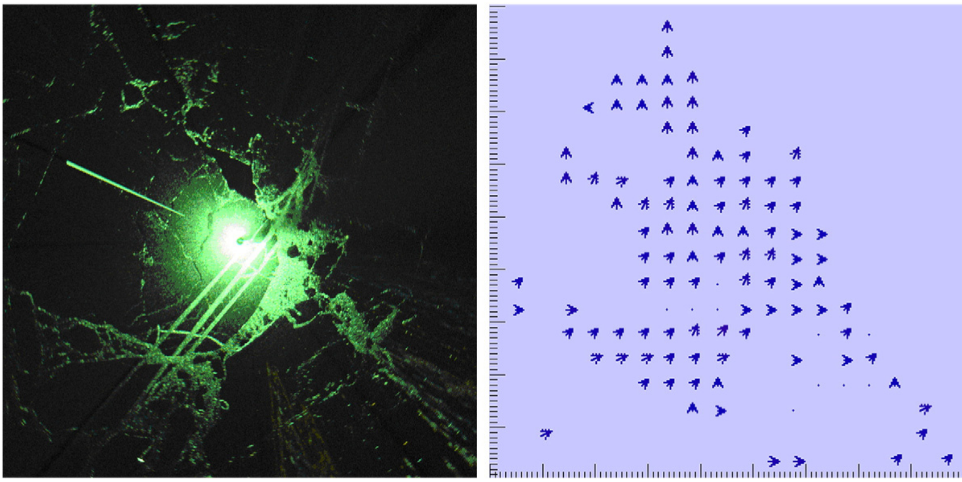

Fig. 14. The image on the left is a composite of two radar images, the earlier of which is rendered in a cyan tone and the latter is in a yellow tone. The radar images were taken on 11 May 2009 at 12:40 UTC and at 12:50 UTC, and the ship location (center of the image) at 12:40 UTC was 65.542 N, 24.413E. The image on the right shows the 10-min ice drift vector field calculated from the radar images. The ship was not moving by engine. The pixel size is 7.5 m, and in the scale bars shown in the image on the right, the distance between the shorter tick marks is 100 m and 1 km between the longer tick marks. The drift vector lengths are scaled by four, i.e., the length of the arrow corresponds to four times the drift within 10 min. The detected zero motion vectors (no direction) are denoted by dots.

different spatial locations, and ice drift calculation could be performed by comparing radar images taken from different ships operating in the same area. Ice drift could also be calculated by comparing images from multiple ships and satellite images. Fig. 10 presents a mosaiced image combining two trails from the same ship and, at the same time, illustrating what the mosaic would look like if the image were formed using two image sets from two different ships going on parallel trails. The combined image mosaic from two ships shows a larger area than the images from a single ship.

At this stage, the prototype system's technical feasibility is proven, but intensive user tests have not yet been carried out. Preliminary discussions with the end-users indicate that radar images delivered from other ships and coastal stations would help decision making in certain situations, and also, the relative ship speed information (Fig. 3) has been taken into operative use. However, in the future, work end-users should be involved in the development so that the additional information coming from the ship sensor network is used in decision making in an optimal way.

Acknowledgment

The work described in this paper was mostly conducted in a project called ShipSensorNet funded by the Finnish Funding Agency for Technology and Innovation.

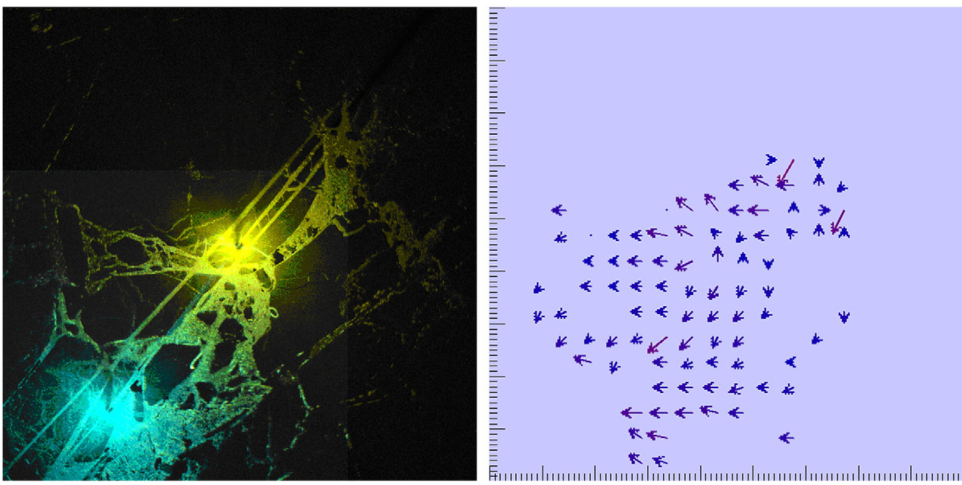

Fig. 15. The image on the left is a composite of two images taken on 11 May 2009 at 06:00 UTC and 11 May 2009 at 06:10 UTC, and the ship location (centre of the image) at 06:00 UTC was 65.525 N, 24.392E. The earlier of the images is rendered in a cyan tone and the latter is in a yellow tone and as the ship is in motion, it has moved from the bright yellow area to the bright cyan area during the 10-min interval. The image on the right shows the calculated ice drift vector field between the two radar images. The pixel size is 7.5 m, and in the scale bars shown in the image on the right, the distance between the shorter tick marks is 100 m and 1 km between the longer tick marks. The drift vector lengths are scaled by four, i.e., the length of the arrow corresponds to four times the drift within 10 min.

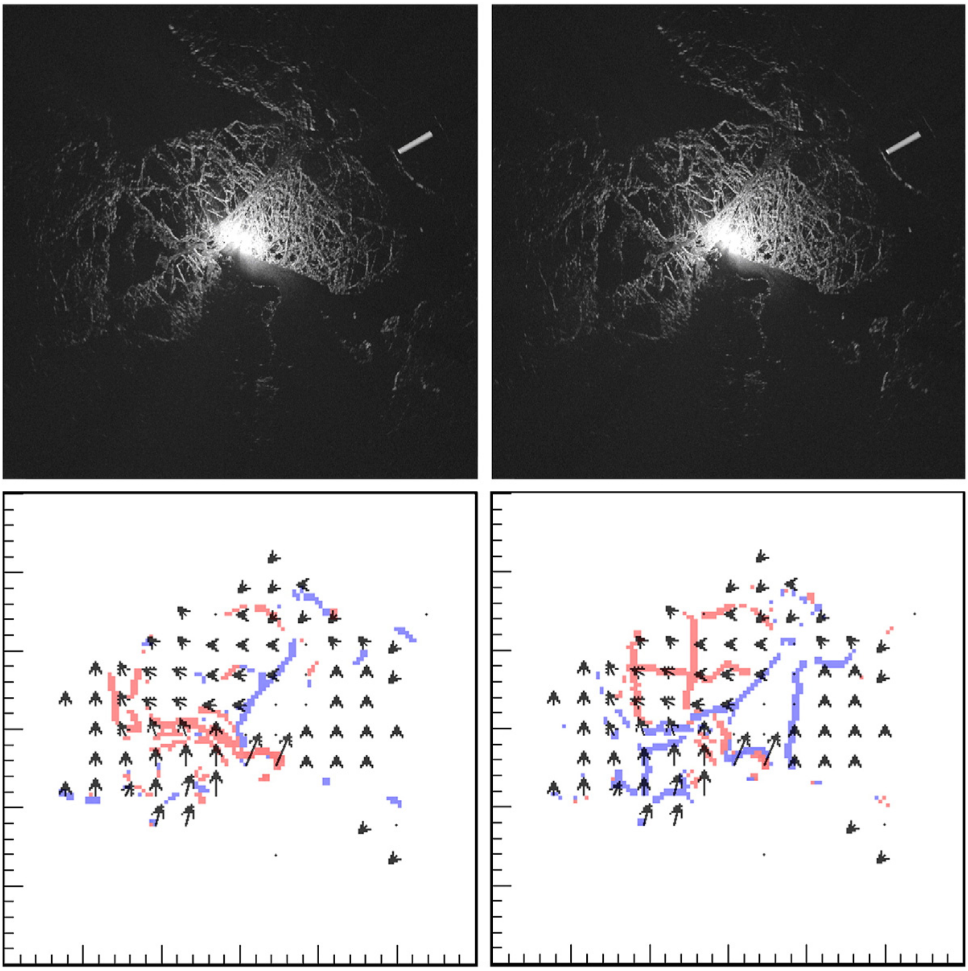

Fig. 16. These four images show a pair of radar images, taken on 9 April 2009 at 16:00 (upper left) and at 16:30 (upper right), and the computed divergence (lower left) and curl (lower right) plotted on the computed ice drift vector field. The ship location at 16:00 UTC was 65.335 N, 23.881E. The pixel size is 30 m, and in the scale bars shown in the lower images, the distance between the shorter tick marks is 500 m and 2.5 km between the longer tick marks. The drift vector lengths are scaled by four, i.e., the length of an arrow corresponds to four times the drift within 10 min. Blue divergence means diverging ice, and red means converging ice, and the tone becomes darker as the absolute value increases. A red curl means a clockwise curl and blue a counter-clockwise curl.

Appendix A. List of symbols

v	ship speed
v_r	relative ship speed ($v_{ice/}v_{ow}$)
v_{max}	maximum ship speed
v_{ow}	open water speed
v_{ice}	ship speed in ice
c	ice condition parameters (ice thickness, concentration, etc.)
s	static ship parameters (length, breadth, maximum machine power, etc.)
d	non-static ship parameters (draught, trim, etc.)
p	non-static ship propulsion parameters (used engine power, propeller pitch, etc.)
f	ship transit model
T_{pull}	bollard pull
K_e	factor including the number of propellers, which is 0.78 for single screw propulsion like *Aranda*
D_p	propeller diameter
P_{sh}	total shaft power
ICE	ice conditions equivalent
g	acceleration of gravity
h_i	level ice thickness
C	ice concentration
L	ship length
L_{bow}	length of the ship bow
L_{par}	length of the ship parallel midbody
B	ship maximum breadth
T_{par}	ship draught
R	distance from radar to target
σ^0	backscatter coefficient
I_p	phase correlation image
$D=(\mathrm{d}r, \mathrm{d}c)$	vector field describing ice drift (ice motion)
dr	ice motion in row direction
dc	ice motion in column direction
div(D)	divergence of a vector field D
curl(D)	curl of a vector field D

References

Berglund, R., Kotovirta, V., Seina, A., 2007. A system for icebreaker navigation and assistance planning using spaceborne SAR information in the Baltic Sea. Canadian Journal of Remote Sensing 33, 378–387.

Canny, J., 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8, 679–698.

Gutierrez, S., Long, D.G., 2003. Application of optical flow and scale space methods to sea–ice motion in Antarctica. Ocean Remote Sensing and Imaging II, August 5, 2003–August 6. SPIE, San Diego, CA, United states, pp. 249–256.

Heil, P., Fowler, C.W., Maslanik, J.A., Emery, W.J., Allison, I., 2001. A comparison of East Antarctic sea–ice motion derived using drifting buoys and remote sensing. International Symposium on Sea Ice and its Interactions with the Ocean, Atmosphere and Biosphere, Int. Glaciological Soc, UK, pp. 139–144.

Inoue, J., Enomoto, T., Miyoshi, T., Yamane, S., 2009. Impact of observations from Arctic drifting buoys on the reanalysis of surface fields. Geophysical Research Letters 36 (8), 1–5.

Ionov, B., 1988. Ice resistance and its components (in Russian). AARI, Leningrad.

ISO/IEC Guide 98-3:2008, Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement (GUM:1995). Geneva, Switzerland.

Karvonen, J., Simila, M., Lehtiranta, J., 2008. SAR-based estimation of the Baltic Sea ice motion. 2007 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2007. Institute of Electrical and Electronics Engineers Inc, Barcelona, Spain, pp. 2605–2608.

Karvonen, J., 2009. C-band sea ice SAR classification based on segmentwise edge features. To Appear in the Book Geoscience and Remote Sensing. InTech Education and Publishing. 2009.

Kuglin, C., Hines, D., 1975. The phase correlation image alignment method. Proceedings of IEEE International Conference on Cybernetics and Society, pp. 163–165.

Kujala, P., Sundell, T., 1992. Performance of Ice-Strengthened Ships in the Northern Baltic Sea in Winter 1991. Finland.

Leisti, H., Riska, K., Heiler, I., Eriksson, P., Haapala, J., 2009. A method for observing compression in sea ice fields using IceCam. Cold Regions Science and Technology 59, 65–77.

Lindqvist, G., 1989. A straightforward method for calculation of ice resistance of ships. The 10th Conference on POAC, POAC '89. Luleå University of Technology, Luleå, Sweden, pp. 722–735.

Mahoney, A., Eicken, H., Shapiro, L., 2007. How fast is landfast sea ice? A study of the attachment and detachment of nearshore ice at Barrow, Alaska. Cold Regions Science and Technology 47, 233–255.

Marton, C.G., 1999. Ice navigation simulation: Shipboard Radar Simulation Module (SRSM). Canada.

Rampal, P., Weiss, J., Marsan, D., Lindsay, R., Stern, H., 2008. Scaling properties of sea ice deformation from buoy dispersion analysis. Journal of Geophysical Research, C: Oceans 113, C03002–C (12 pp.).

Riska, K., Wilhelmson, M., Englund, K., Leiviskä, T., 1997. Performance of merchant vessels in ice in the Baltic. Winter Navigation Research Board, report 52, Helsinki.

Sun, H., Shen, Y., Qiu, D., 2004. Measuring sea ice drift via cross-correlation of radar ice images. Journal of Hydrodynamics 16, 716–720.

Thomas, M., Kambhamettu, C., Geiger, C.A., 2009. Mapping of large magnitude discontinuous sea ice motion. SIGSPATIAL Special 1, 45–50.

Publication V

Kotovirta V., Toivanen T., Järvinen M., Lindholm M., Kallio K. 2014. Participatory surface algal bloom monitoring in Finland in 2011–2013. Environmental Systems Research 3(24).

V

Kotovirta et al. Environmental Systems Research 2014, 3:24
http://www.environmentalsystemsresearch.com/content/3/1/24

Environmental Systems Research
a SpringerOpen Journal

RESEARCH ARTICLE Open Access

Participatory surface algal bloom monitoring in Finland in 2011–2013

Ville Kotovirta[1*], Timo Toivanen[1], Marko Järvinen[2], Matti Lindholm[3] and Kari Kallio[3]

Abstract

Background: Algal mass occurrences are one of the most distinguishing effects of eutrophication in lakes and the coastal waters of the Baltic Sea. Algal bloom occurrence in water bodies varies greatly in terms of both space and time, even during short periods, which makes reliable monitoring of blooms difficult. In this paper, we explore the possibilities to extend the sensor network both spatially and temporally by applying participatory sensing to surface algal bloom monitoring in Finnish lakes and the coastal areas of the Baltic Sea.

Results: Two participatory sensing systems were used to collect visual algae observations by citizens: the mobile phone application Levävahti (Algae Watch) and the collaborative web service Järviwiki (Lake wiki), during the summers of 2011–2013. Citizen observations were compared with the visual observations performed by trained expert observers, and mean correlations between citizen and expert observations were calculated using the bootstrapping method: *0.72, 95%* confidence interval (CI) *[0.53 0.86]; 0.65, 95%* CI *[0.35 0.86];* and *0.56, 95%* CI *[0.29 0.76]* for the years 2011, 2012 and 2013.

Conclusions: Surface algal bloom monitoring is needed to obtain data on algal bloom frequency and intensity, in particular in lakes where the use of satellite remote sensing has limitations and/or phytoplankton monitoring is infrequent or totally lacking. The correlations between expert and citizen observations suggest that citizen observers can provide additional information to support algal bloom monitoring of inland and coastal waters.

Keywords: Algal bloom monitoring; Cyanobacteria; Lakes; the Baltic Sea; Mobile phone; Participatory sensing; Citizen science

Background

Algal bloom monitoring

Algal mass occurrences, in particular cyanobacterial surface blooms, are one of the most distinguishing effects of eutrophication in lakes and the coastal waters of the Baltic Sea (Solimini et al. 2006). It is therefore logical that the frequency and intensity of cyanobacterial blooms are used in the assessment of the ecological status of surface water bodies in Europe under the European Water Framework Directive (European Commission 2000; Carvalho et al. 2013). Algal bloom occurrence in water bodies varies greatly in terms of both space and time, even during short periods, which makes reliable monitoring of blooms difficult. Traditional manual monitoring with biweekly to monthly or longer sampling frequency easily leads to situations in which not all bloom events can be detected. Remote sensing of algal mass occurrences by satellites can provide high temporal and spatial resolution bloom information of sea areas (e.g. Reinart and Kutser 2006). Monitoring of small water bodies by satellite remote sensing requires images with a good spatial resolution (<30 *m*). At present, such images are not operationally available on a daily basis. One possible way to increase information on cyanobacterial surface bloom situations is to conduct visual observations of blooms by trained observers at fixed observation sites and periods, a method that has been used in Finland since 1998 to estimate average weekly surface bloom situations (Rapala et al. 2012). Although these results cannot at present be used as a bloom metric to assess ecological status, due to a lack of reference conditions and class boundary values, they can be used to support classification decisions by expert judgement (Aroviita et al. 2012). Due to the large number of lakes in Finland (>*188 000* with a surface area of >*5* acres), visual

* Correspondence: ville.kotovirta@vtt.fi
[1]VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044, VTT Espoo, Finland
Full list of author information is available at the end of the article

observations by trained observers are restricted to a rather limited set of lakes.

One possible way to increase both temporal and spatial coverage of surface bloom visual observations in a larger number of lakes and the coastal regions of the Baltic Sea is to include citizens as observers. This has already started in Finland, but the applicability of the observations made by citizens has not been verified before.

Citizen science and participatory sensing
Citizen contributions to environmental monitoring are increasing (Conrad and Hilchey 2011). This is part of a broader emerging field of citizen science in which citizens produce scientifically meaningful observations or analyses (Haklay 2012). Citizen science has been successfully applied to ecological research, e.g. for monitoring birds, insects and invasive plants (Dickinson et al. 2012), and citizens have been involved in environmental monitoring, to some extent, for over a hundred years (e.g. the Christmas Bird Watch, started by ornithologists of North America, has been ongoing since 1900, Haklay 2012).

Recent advances in information and communication technology (ICT) and increased awareness of the status of the environment, in particular global climate change, have activated people even more to participate in monitoring (Burke et al. 2006). Mobile devices with GPS receivers provide a useful platform for collecting data about the environment. Citizens can provide observations actively themselves (i.e. participatory sensing), enable their mobile devices to collect data automatically (i.e. opportunistic sensing, Lane et al. 2008) and analyse or discuss the data or the environmental conditions using proprietary software or social media tools such as Facebook, Google+ and Twitter.

Methods for participatory and opportunistic citizen sensing have been presented for various applications and in various forms. Eiman (2010) presented a method for mapping noise levels in a city area, while Paxton and Benford (2009) described a study in which CO_2 level measurements were observed with a handheld device and combined with written observations and video clips. Mednis et al. (2011) developed a system for road irregularity detection using data from accelerometers of mobile phones.

Citizen sensing has been applied to hydrology, and, e.g., Olmanson et al. (2008) used citizen Secchi depth measurements as in-situ data source for satellite image calibration, Sunyoung et al. (2011) developed a mobile application (Creek Watch) for citizens to monitor waterways (amount of water, rate of flow and amount of litter), Lowry and Fienen (2013) presented a method for stream stage monitoring in which citizen passers-by make observation using fixed measuring devices, Toivanen et al. (2013) presented a method for observing Secchi depth and turbidity using an inexpensive measurement device and a mobile phone camera, and Leeuw and Boss (2014) developed the HydroColor app which estimates the concentration of total suspended matter and the backscattering coefficient from water reflectance measured by a mobile phone camera. The European Environment Agency published the Marine LitterWatch mobile phone-based app (http://www.eea.europa.eu/themes/coast_sea/marine-litterwatch) in 2013 to involve citizens in the monitoring of marine litter distribution and composition. The app was specifically developed for the needs of the EU's Marine Strategy Framework Directive.

Suggested architectures and platforms for more generic participatory sensing include G-Sense (Perez et al. 2010), PRISM (Platform for Remote Sensing using Smartphones) (Das et al. 2010), the personal environmental impact report PEIR (Mun et al. 2009) and the EnviObserver participatory sensing system (Kotovirta et al. 2012).

Focus of this paper
In this paper we discuss the results of applying participatory sensing to surface algal bloom monitoring in Finnish lakes and the coastal areas of the Baltic Sea during the summers of 2011, 2012 and 2013. The hypothesis is that voluntary citizen monitoring can extend the existing sensor and observation networks both spatially and temporally and provide useful information about cyanobacterial bloom coverage. We compared the citizen observations with weekly observations made by trained experts and found a clear correlation between the bloom intensity averages, supporting our hypothesis.

Results
Number of observations
A total of *4572* trained expert observations and *872* volunteer citizen observations, of which *269* were made using the mobile phone application, were received during the weeks *24* to *38* in the summer of 2011. In the summer of 2012, *4427* expert observations and *319* citizen observations (*156* mobile) were received, and in the summer of 2013, *4150* expert observations and *465* citizen observations (*134* mobile) were performed (Table 1). Figure 1 presents a comparison of the locations of the expert observation sites with the citizen observations for the summer of 2011. To analyse the spatial distribution of the citizen and expert observations the numbers of observations were calculated for rectangles sized one degree (latitude) by two degrees (longitude) covering the whole dataset. The most observations were done in south and south-west Finland for the year 2011, and a similar pattern was also observed for the years 2012 and 2013.

The number of volunteer citizen observations was highest in the summer of 2011. This probably resulted

Table 1 Yearly counts and correlations

Year	Area	Count			'No algae'		Correlation of original datasets	Mean of bootstrap correlations	95% confidence	Citizen spatial correlation	Citizen count correlation
		Experts	Citizens	Mobile	Experts	Citizens					
2011	Whole	4572	872	31%	80%	60%	0.82	0.72	[0.53 0.86]	−0.20	−0.50
2012	Whole	4427	319	49%	86%	52%	0.88	0.65	[0.35 0.86]	−0.27	−0.59
2013	Whole	4150	465	29%	83%	42%	0.69	0.56	[0.29 0.76]	−0.32	−0.68
2011	Area1	283	217	20%	76%	79%	0.64	0.47	[0.13 0.79]	−0.24	0.16
2011	Area2	1257	433	35%	73%	61%	0.84	0.71	[0.49 0.88]	−0.04	−0.03
2012	Area2	1242	196	45%	78%	54%	0.55	0.40	[0.05 0.71]	0.68	−0.00
2013	Area2	1108	182	35%	73%	54%	0.80	0.57	[0.21 0.82]	−0.45	−0.47

Year, area (*Whole = 59.5N 20E, 67.5N 32E, Area1 = 59.5N 22E, 60.5N 24E, Area2 = 59.5N 22E, 61.5N 26E*), the count of observations by experts and citizens, the percentage of citizen mobile observations, percentages of 'no algae detected' observations, the correlation of expert and citizen weekly averages using the original datasets, mean of correlations calculated using bootstrapping method, *95%* confidence interval of the bootstrap correlations, the correlation of weekly citizen observations' variation and the weekly spatial distribution (citizen spatial correlation), and the correlation of the weekly citizen observations' variation and the weekly count of observations (citizen count correlation).

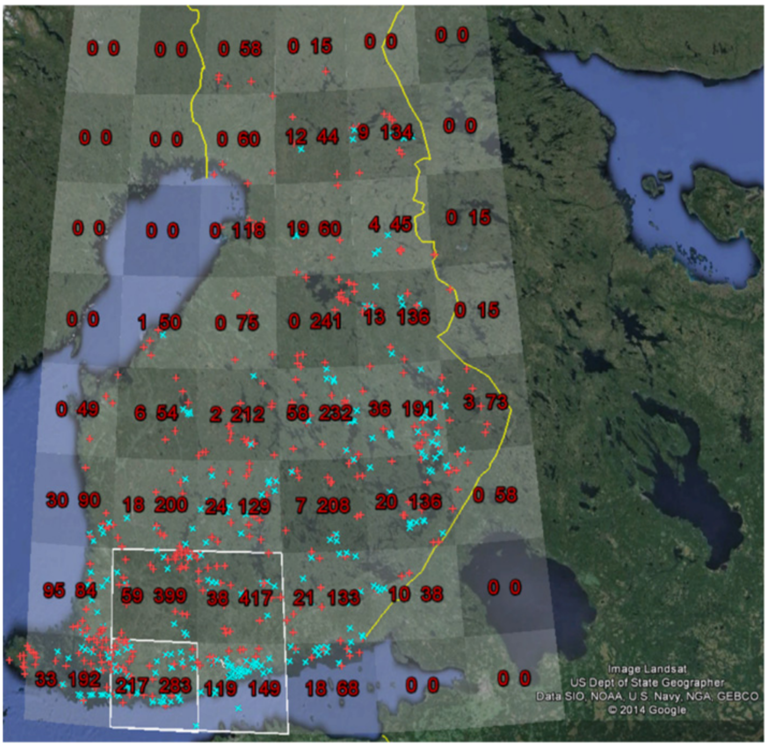

Figure 1 Comparison of the geographical locations of the citizen and expert observations in the summer of 2011. The blue xs indicate the locations of the citizen observations and the red plus signs ('+') indicate the pre-determined locations of the expert observation sites. The numbers of citizen and expert observations (in this order) are presented for each grey rectangle (dimension of one degree latitude and two degrees longitude) which together cover the whole dataset (approx. *59.5N 20E, 67.5N 32E*). The white rectangles indicate the smaller areas used in the analysis (*Area1 = 59.5N 22E, 60.5N 24E* and *Area2 = 59.5N 22E, 61.5N 26E*).

from a worse surface algal bloom situation (i.e. more surface blooms) in 2011 than in subsequent years, which probably made citizens more eager to make observations. The overall algal bloom intensity was estimated by taking the average value of all the expert observations, which was *0.21* for the summer of 2011, *0.15* for the summer of 2012 and *0.19* for the summer of 2013. These correlate with the total number of citizen observations (*831*, *303* and *443*).

Weekly averages

The weekly averages of both citizen and expert observations and the *95%* confidence intervals calculated using the bootstrapping method are presented in Figures 2, 3 and 4 for the years 2011, 2012 and 2013, respectively. The best accuracy for citizen observation averages was achieved for the year 2011 and the weakest for 2012. This may be again explained by the overall algae situation and the amount of citizen observations collected yearly. The percentage of citizen mobile observations and the percentages of 'no algae detected' observations are given in Table 1 for each year. The percentage of citizen mobile observations was around *30%* for the years 2011 and 2013, and around *50%* for the year 2012, although the total number of citizen observations was the lowest for 2012. The citizen averages are biased towards higher bloom intensity values than the averages of the expert observations, which can be seen in Figures 2, 3 and 4. One possible explanation, but not necessarily the only one, is revealed by comparing the percentages of 'no algae detected' (i.e. observation value '*0*') by citizen and expert observers. The portion of 'no algae' observations is lower for citizens than for experts, and therefore the averages of citizen observations are biased towards higher values (see also discussion about data quality).

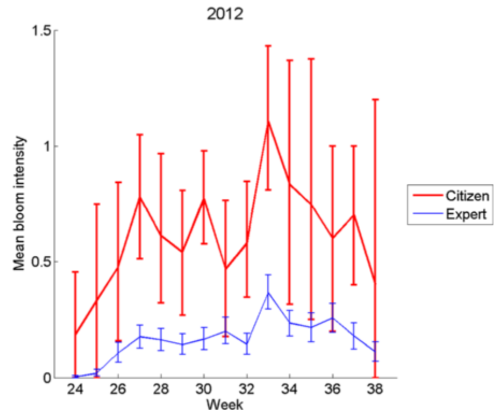

Figure 3 **Weekly averages of citizen and expert observations and 95% confidence intervals in the summer of 2012.**

Correlations

The correlations of expert and citizen weekly averages and their *95%* confidence intervals for each year are given in Table 1. We show both the correlations calculated from the original datasets and the mean of correlations calculated using the bootstrapping method. The correlation histograms are given in Figures 5, 6 and 7 and they represent graphically the correlation distributions. It can be seen that the distributions are left-tailed or left-skewed, and the corresponding skewness values are *−0.45*, *−0.80*, and *−0.56* for the years 2011, 2012 and 2013. It appeared that the correlation calculated from the original datasets for the year 2012 was above the

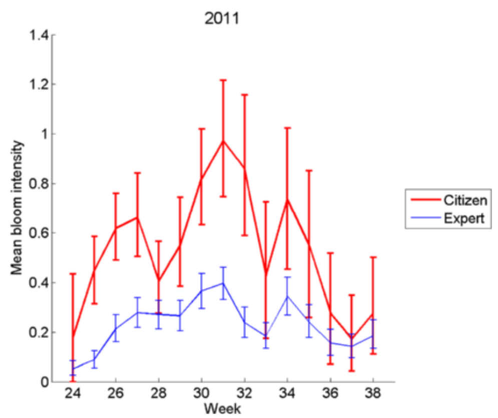

Figure 2 **Weekly averages of citizen and expert observations and *95%* confidence intervals in the summer of 2011.**

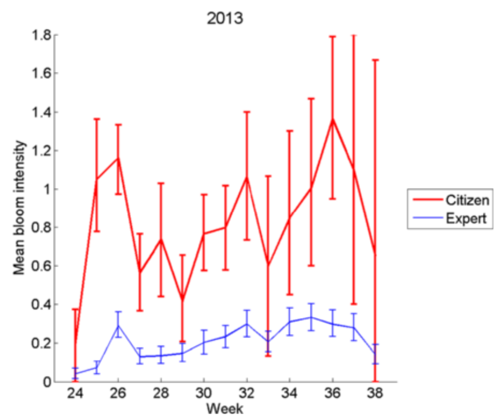

Figure 4 **Weekly averages of citizen and expert observations and *95%* confidence intervals in the summer of 2013.**

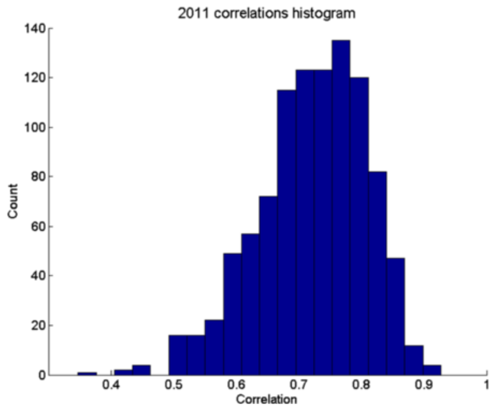

Figure 5 Histogram of the correlations of the citizen and expert weekly averages in the summer of 2011.

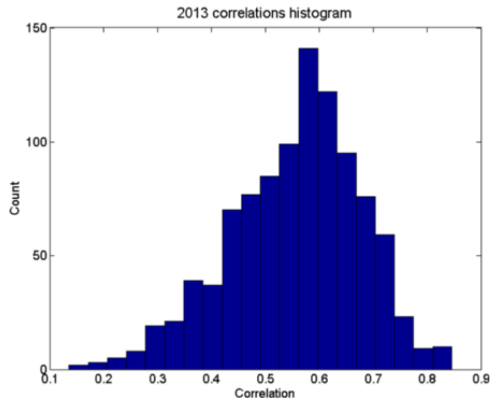

Figure 7 Histogram of the correlations of the citizen and expert weekly averages in the summer of 2013.

calculated *95%* confidence interval, which was anticipated to be due to pure chance, but it also shows how the original correlation would not probably be a reliable estimator of the true correlation. The correlation between the weekly citizen *95%* confidence interval width, i.e. the variation of the citizen weekly averages, and the citizen observations' spatial distribution, and the correlation between the variation of the citizen weekly averages and the number of citizen observations are also given in Table 1. The correlations indicate that the spatial distribution of the citizen observations does not clearly explain the variation of weekly averages, but the negative correlations of the number of citizen observations and the variation (for the whole geographic area) indicate that the more observations there are for one week the narrower is the confidence interval of the average, which shows consistency in the data.

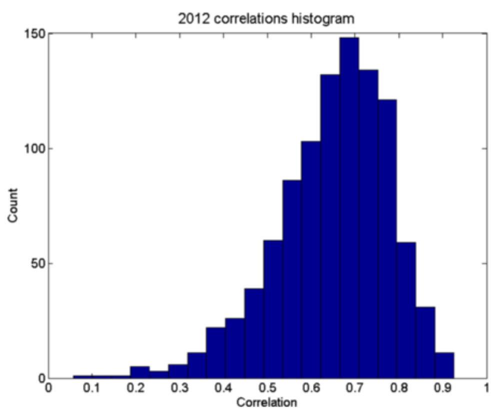

Figure 6 Histogram of the correlations of the citizen and expert weekly averages in the summer of 2012.

Discussion

Citizen activity

Every year during the test period, the citizen activity tends to decrease towards the end of the summer, although the bloom events have not yet decreased markedly. This drop in observation activity may be explained by the summer holiday season in Finland, which commonly starts in mid-summer and lasts until the beginning of August, i.e. weeks *25* to *31*. There are more potential citizen observers near waters (e.g. through boating, water sports, active use of summer cottages) in the middle of the holiday season.

In addition to the surface bloom situation and holiday season, other factors are likely to influence the observer activity. We published invitations to participate in the pilot study in local and national news services and at water-related events and exhibitions at the beginning of the summer algal growth season in Finland, but we did not follow up how well the message was received by potential users. We did not require any registration for the mobile part of the system, which makes the adoption of the system easier, but, at the same time, it makes the analysis in detail of the user activity and motivation of making observations more difficult.

Different frameworks have been suggested to motivate citizen observers. Reddy et al. (2010a) presented a recruitment framework for identifying potential participants for data collections, and Juong-Sik and Hoh (2010) discussed an incentive mechanism for stimulating participatory sensing applications. Micro-payments as an incentive mechanism are explored by Reddy et al. (2010b).

We used a simple motivation mechanism in which users received notifications about the algae situation and reminders to contribute to observations. In the future, more efforts are needed to recruit suitable observers and to motivate them to continue making observations. Sunyoung et al. (2011) suggest that creating a successful citizen sensing application requires the application to be designed together with various stakeholders and ensuring that the gathered data can be put to use. Currently, as part of the MARMONI EU Life+ project (http://marmoni.balticseaportal.net/wp/), representatives of non-governmental organizations and school teachers are trained to make algae (cyanobacterial bloom and bladderwrack occurrence) and Secchi depth observations. The idea is that the representatives will act as trainers in their organizations and schools.

In general, the citizens are motivated to participate in environmental science projects (Roy et al. 2012). However, one challenge of harnessing citizen observers is to find an economically sustainable solution that motivates not just the citizens and researchers but also the system developers to innovate, update and maintain state-of-the-art, easy-to-use tools for making observations. Citizen sensing activities and campaigns can be funded in the context of research projects, but to fully empower citizens as observers of various environmental parameters also requires new commercial innovations.

Data quality

The quality of the collected information is an important concern of citizen sensing (Haklay et al. 2010, Comber et al. 2013). In many cases, the quality varies, it is not documented at all, it fails to follow scientific principles of sampling design, and its coverage is incomplete (Goodchild and Li 2012).

When looking at the data and graphs of this study, we note that the citizen observation averages are biased towards higher bloom intensity values than the expert observation averages. This can be explained by noting that the citizens made fewer 'no algae detected' observations than the experts. Trained observers are instructed to make observations whether there are algae or not, but citizens make observations whenever they find it useful. Citizens probably make observations more often when algae are visible and tend to omit 'no algae' reports.

When the observations are based on human senses, the measurements are not of uniform quality but vary according to individual capabilities. In this study the citizen observers had also the possibility to submit photos along with their classification of algal blooms. The quality of these photos varies depending on the mobile phone model, the camera resolution, the distance, angle of view and degree of possible surface reflection, and therefore automatic classification of algal blooms is not feasible. However, the photos can be used visually as an additional method to verify the citizen observations.

As in any data gathering task relying on volunteer user contributions, there is a risk of faulty input by human errors or even service misuse. The risk of misuse may be even higher when no registration to the service is required and the abusers cannot be tracked. It is important to determine the quality of user observations, especially when using the data for evaluating and validating predictive models or as ground-truth data for reference. During the performed pilot trials, no service misuse was detected, and no actions to remedy that kind of activity were necessary. Obvious test uses of the system were detected manually, but in the future some automatic identification of faulty or accidental observations should be implemented to improve the data quality. For example, Alabri and Hunter (2010) describe a framework combining data quality control and trust metrics to enhance the reliability of citizen science data, and Kuan et al. (2010) as well as Yang et al. (2011) propose reputation management systems to evaluate the trustworthiness of gathered data by co-observers and data end-users.

Data privacy is one concern that must be considered, and it is discussed in several studies, e.g. Christin et al. (2011) conducted a survey on privacy in mobile participatory sensing applications and showed that almost all applications capture location and time information. Methods for protecting privacy have been presented, e.g. by Kazemi and Shahabi (2011). In our case the mobile citizen observers were anonymous so the privacy was not the primary concern, however, with the missing contact details we could not ask for feedback about the mobile system. Even though the names of the observers were not public (or not even available), individual identities could be inferred from the location information, e.g. if a summer cottage is used as a regular observation station. If data privacy becomes an issue later on the exact location could be hidden in the published data and stored unaltered for research purposes.

Our study was based on realistic data collected from experts making observations regularly in stationary locations and citizens making observations at more or less sporadic locations and times. The dataset did not enable very accurate comparison of expert and citizen observations in terms of space and time. We compared average values of large geographic area, which introduced errors in the data and therefore uncertainties in the results, as the algal bloom may develop differently in different parts of the averaged area. To study the quality of citizen observations in more detail, a special campaign could be organised to ensure that citizens and experts observe the same algae situation in the same region at the same time.

Usefulness of the data collected

Based on our analysis, the systematic positive correlations between the expert and citizen algae observations in consequent years suggest that citizen observers can extend the current observation network spatially and temporally and provide additional information that supports the monitoring of the algal bloom situation. Already now the visual algal bloom observations by experts are used as supporting information for the Water Framework Directive (WFD) ecological classification of surface waters. All lakes in Finland cannot be monitored by authorities with available resources, thus citizens can provide additional information from areas that are not currently monitored and additional sampling of water quality can be carried out in areas where citizens have reported algal blooms.

The reliability and usefulness of citizen observations for monitoring terrestrial and marine environments have been analysed in several studies. Obrecht et al. (1998) conclude that citizen Secchi depth measurements are nearly identical with the measurements made by professionals. Delaney et al. (2008) identified obstacles in citizen monitoring and concluded that, with proper training, citizens can provide reliable aid in collecting knowledge about both native and invasive crabs. A study by Gallo and Waitt (2011) concludes that citizen scientists are able to detect and report invasive plants in their local areas, and the data can be used by professional scientists. D'Hondt et al. (2012) created noise maps based on citizen observations and concluded that they are comparable with official simulation-based noise maps. Lottig et al. (2014) analysed over 140 000 Secchi observations from 3251 lakes in the USA and demonstrated that citizen science can provide the critical monitoring data needed to improve spatial and temporal scales.

Conclusions

Surface algal bloom monitoring is needed to obtain data on algal bloom frequency and intensity, in particular in lakes where the use of satellite remote sensing has limitations and/or phytoplankton monitoring is infrequent or totally lacking. In this paper we present how citizens can also take part in the algal bloom monitoring in Finland and thus accumulate additional information on bloom occurrences. Observations by untrained citizen observers were collected in the summers of 2011, 2012 and 2013 and compared with the trained expert observations. Two systems for citizen sensing were developed and applied in the study: a mobile phone application called Levävahti (Algae Watch) and a collaborative web service about Finnish lakes called Järviwiki (Lake wiki). A clear correlation between the expert and citizen observations was found in the analysis, which suggests that citizen observers can provide additional information for algae monitoring.

Europe is facing increasing monitoring requirements to meet obligations under, for example, the Water Framework Directive and the UN Convention on Biological Diversity (CBD). This opens up opportunities to develop and test new innovations for citizen science and community-based environmental monitoring. These innovations should not only provide new ways of gathering data but also engage and encourage the community in sustainable management of the environment.

Methods

Visual observations of algal bloom situation

In the visual algal bloom monitoring of Finnish waters, cyanobacteria bloom intensity is evaluated both by expert and citizen observers using four classes: 0 = not detected, 1 = detected, 2 = high amount and 3 = very high amount (Rapala et al. 2012). Experts observe from June to September in the fixed shore observation sites, and citizens make additional observations for all water bodies at any time of year as frequently as they wish. The visual surface bloom observations focus on cyanobacteria (blue-green algae), as many bloom-forming cyanobacteria can form dense surface scums that can be toxic to humans as well as to other biota:

0: Not detected. No algae on the water surface or on the shore line. The Secchi depth visibility is normal.
1: Detected. Greenish flakes (cyanobacteria colonies) detected in the water or when taken into a transparent container, or narrow stripes on the shore. The Secchi depth is reduced by algae.
2: High amount. The water is clearly coloured by algae, small surface scums or cyanobacterial mass on the beach are detected.
3: Very high amount. Wide and heavy surface scums or thick aggregates of cyanobacteria are detected on the shore.

Trained expert observations

National surface algal bloom monitoring in Finland has been carried out since 1998. It is coordinated by the Finnish Environment Institute (SYKE) and carried out in co-operation with local environmental authorities, municipalities and private trained persons (Rapala et al. 2012). The aim of monitoring is to provide an up-to-date overview of the cyanobacterial situation and information about the spatial and temporal variation during the summer.

The visual observations are made weekly by expert observers, i.e. authorities and trained volunteers, in approximately *320* pre-determined locations in Finland, of which approximately *260* represent lakes and approximately *60* coastal areas of the Baltic Sea. Observation sites have been selected to represent various types of waters with differing trophy, water colour (humic content), size and geographical location. Many observation sites are located near public beaches or cities. Many expert observers represent regional environment authorities from municipalities who are responsible for the monitoring of the status of recreational or drinking waters, incl. beaches, and who have been trained to detect cyanobacteria blooms by the phytoplankton and monitoring experts of the Centres of Economic Development, Transport and Environment or the Finnish Environment Institute (SYKE).

In this study, we concentrate on visual observations, as these can be compared with the citizen observations. However, other means are also used in algal monitoring, e.g. for the Finnish coastal areas of the Baltic Sea the information is obtained from satellite imagery, commercial ships with ship-of-opportunity devices and the Border Guard. In addition, if the amount of algae is visually estimated to be high or very high a sample is taken for the qualitative analysis of the bloom taxa using microscopy.

Citizen observations

In addition to observations by trained experts during the weekly national bloom situation monitoring, visual observations by untrained citizen observers were collected in 2011–2013. The architecture of the citizen observation system is depicted in Figure 8. Two applications for citizen sensing were developed and applied in the study to receive both mobile observations from ad-hoc locations and observations from stationary observation sites defined by citizens: the mobile phone application Levävahti (Algae Watch) and the web-based lake information system called Järviwiki (Lake wiki, http://www.jarviwiki.fi/wiki/J%C3%A4rviwiki:About, Rapala et al. 2012). Citizens evaluated occurrences of algae with the same scale as trained observers irrespective of the observation application. The different algae classes and their criteria were described with example photographs in the applications' help pages to provide a tutorial about algae observation.

The Levävahti application (Algae Watch) was implemented on a participatory sensing platform called EnviObserver (Kotovirta et al. 2012) and was made available for Nokia (Java ME) and Android-based mobile phones in the

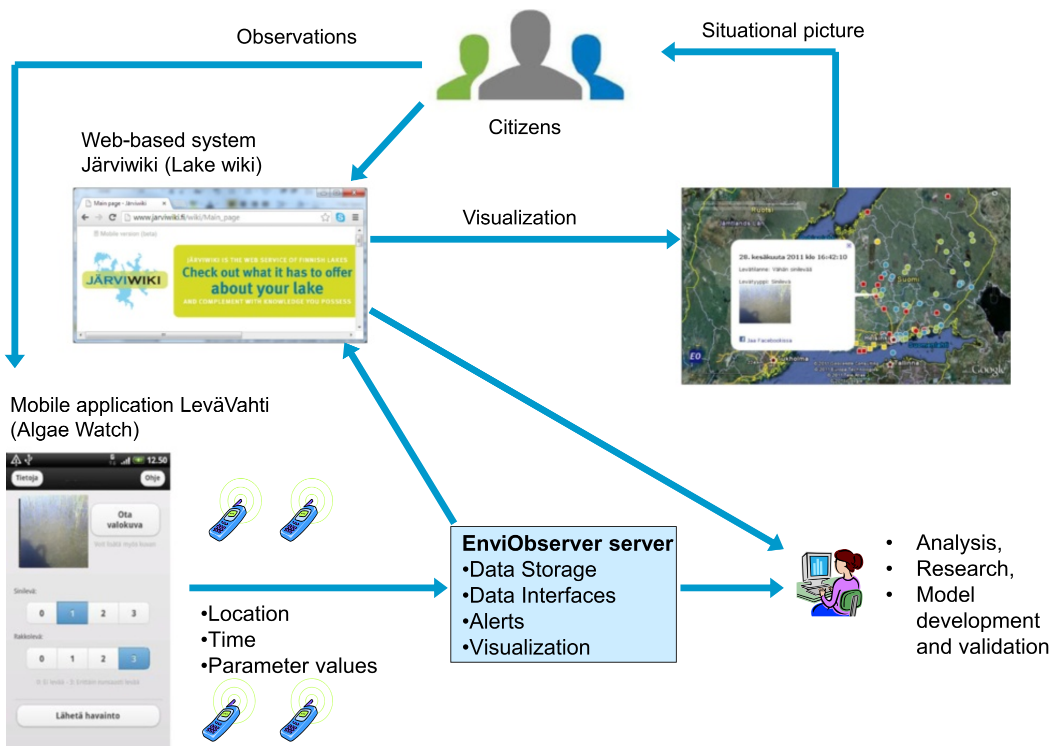

Figure 8 Overall architecture of the participatory algae monitoring system. Two participatory sensing systems were used to collect citizen algae observations: the mobile phone application Levävahti (Algae Watch) and the collaborative web service Järviwiki (Lake wiki).

summer of 2011; for Nokia (Java ME), Android and iPhone phones in 2012 and 2013; and for Windows phones in 2013. The on-line observations were stored together with a GPS location, time and voluntary image taken by the user.

Järviwiki (Lake wiki) is a collaborative web service for sharing information about lakes in Finland, raising awareness and promoting protection of waters. The users of Järviwiki can write about or discuss a lake and set up their own observation sites and report on water parameters such as algal blooms, surface water temperature and ice cover. The visualizations of observations received from the Levävahti application and the observation sites in Järviwiki were implemented in the Järviwiki web service.

Citizen observers for the study were recruited by publishing invitations in local and national news services and water-related events and exhibitions at the beginning of the summer season in Finland. Cyanobacterial blooms are typically observed in Finnish waters from late June until the end of August. The mobile phone application was open to everyone without registration, requiring only downloading of the software. The Järviwiki web service was also open to everyone, but registration was needed to set up an observation site.

Comparison

To evaluate the usefulness of citizen observations, they were compared with the visual observations by experts. Due to the nature of this observation campaign, the locations and timing of the mobile phone citizen observations were quite sporadic, compared with the regular weekly observations by trained observers in pre-determined locations. The differences in location and timing did not enable direct comparisons between individual observations. Instead, we used averaging of both space and time and compared the averaged values. We calculated weekly averages for the whole geographic area of the dataset in Finland and for two smaller geographical areas (see Figure 1) to study the geographical variability of the observations. The sizes of the smaller areas were selected heuristically so that there were enough expert and citizen observations to calculate weekly averages. The smallest area was used only for the year 2011 dataset which contained enough observations for the analysis. The time range was restricted to weeks *24* to *38* (approx. early June to mid-Sept) in order to have enough citizen observations for weekly averaging. The algal blooming occurs during the summer months in Finland and the citizens appeared to be more active observers during the summer time.

We calculated the correlations between the weekly averages of citizen and expert observations for each year. In order to estimate the accuracy of the correlations we used the bootstrapping method, in which the observed data are used as an approximating distribution of the real distribution. We resampled random samples from each week's dataset and constructed a sampled observation dataset for each week that was equal in size to the original dataset. We then calculated new sampled averages for each week, and used those in calculating the correlations between citizen and expert observations. We repeated the process *1000* times and got distributions for weekly averages and also for the correlations of the weekly averages, and from the distributions we calculated the confidence intervals both for the weekly averages and the correlations of averages.

To study the variation of weekly averages of citizen observations we used the width of the *95%* confidence interval as a measure of variation and calculated its correlation with the spatial distribution of citizen observations. The spatial distance between observations might (partly) explain the variation, as the algae situation may progress differently in different latitudes, or even in different lakes in the same latitude. The spatial distribution was calculated as a mean distance of the observations from their geographical centre point. We also calculated the correlation of the variation with the number of citizen observations.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

VK designed and conducted the research, analysed the results, drafted and polished the manuscript; VK and TT designed the EnviObserver and Levävahti system architectures and participated in the mobile application design; TT participated in the implementation of the mobile phone application and contributed to the manuscript; ML designed and implemented the Järviwiki web service, MJ coordinated the national algal bloom monitoring at the Finnish Environment Institute and contributed to the manuscript; KK participated in the mobile phone application design and contributed to the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The work has been conducted in the Measurement, Monitoring and Environmental Assessment research programme funded by Tekes, the Finnish Funding Agency for Technology and Innovation, in the Pathway project funded by the Academy of Finland, and in the EU Life+ funded project GisBloom (Life09 ENV/FI/000569). We want to thank Markku Huttunen and Renne Tergujeff from VTT for their participation in the design and development of the Levävahti (Algae Watch) mobile application; Olli Saarela from VTT for his advice on data analysis; Seppo Kaitala, Maria Kämäri, Jari Silander and Juha-Markku Leppänen from the Finnish Environment Institute for participating in the design of Levävahti and the recruitment of participants; and Ville Peltola and Asko Relas from IBM Finland for participating in the mobile application design and developing the first iPhone version of LeväVahti for the summer of 2011.

Author details

[1]VTT Technical Research Centre of Finland, P.O. Box 1000, FI-02044, VTT Espoo, Finland. [2]Finnish Environment Institute (SYKE), SYKE Jyväskylä Office,

Survontie 9A, 40500 Jyväskylä, Finland. [3]Finnish Environment Institute (SYKE), P.O. Box 140, 00251 Helsinki, Finland.

Received: 20 July 2014 Accepted: 19 November 2014
Published online: 02 December 2014

References

Alabri A, Hunter J (2010) Enhancing the quality and trust of citizen science data. In: Proceedings of 2010 6th IEEE International Conference on e-Science, eScience 2010, Brisbane, Australia, 7–10 Dec 2010

Aroviita J, Hellsten S, Jyväsjärvi J, Järvenpää L, Järvinen M, Karjalainen S-M, Kauppila P, Keto A, Kuoppala M, Manni K, Mannio J, Mitikka S, Olin M, Perus J, Pilke A, Rask M, Riihimäki J, Ruuskanen A, Siimes K, Sutela T, Vehanen T, Vuori K-M (2012) Guidelines for the ecological and chemical status classification of surface waters for 2012–2013 – updated assessment criteria and their application. Environmental Administration Guidelines 7/2012, Finnish Environment Institute (SYKE)., p 144, URN:SBN 978-952-11-4114-0. ISBN 978-952-11-4114-0

Burke J, Estrin D, Hansen M, Parker A, Ramanathan N, Reddy S, Srivastava M B (2006) Participatory Sensing. In: Proceedings of the International Workshop on World-Sensor-Web (WSW'2006), ACM, Boulder, CO, USA, 31 Oct 2006

Carvalho L, Poikane S, Lyche Solheim A, Phillips G, Borics G, Catalan J, De Hoyos C, Drakare S, Dudley BJ, Järvinen M, Laplace-Treyture C, Maileht K, McDonald C, Mischke U, Moe J, Morabito G, Nõges P, Nõges T, Ott I, Pasztaleniec A, Skjelbred B, Thackeray SJ (2013) Strength and uncertainty of lake phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704:127–140, doi:10.1007/s10750-012-1344-1

Christin D, Reinhardt A, Kanhere SS, Hollick M (2011) A survey on privacy in mobile participatory sensing applications. J Syst Softw 84(11):1928–1946

Comber A, See L, Fritz S, Van der Velde M, Perger C, Foody G (2013) Using control data to determine the reliability of volunteered geographic information about land cover. Int J Appl Earth Obs Geoinf 23:37–48, ISSN 0303–2434, Available online http://dx.doi.org/10.1016/j.jag.2012.11.002 Accessed 8 Oct 2014

Commission E (2000) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Off J Eur Communities 327:1–72

Conrad C, Hilchey K (2011) A review of citizen science and community-based environmental monitoring: issues and opportunities. Environ Monit Assess 176:273–291, doi:10.1007/s10661-010-1582-5

D'Hondt E, Stevens M, Jacobs A (2013) Participatory noise mapping works! An evaluation of participatory sensing as an alternative to standard techniques for environmental monitoring. Pervasive Mob Comput 9(5):681–694, ISSN 1574–1192, 10.1016/j.pmcj.2012.09.002

Das T, Mohan P, Padmanabhan V, Ramjee R, Sharma A (2010) PRISM: Platform for Remote Sensing using Smartphones. In: Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services (MobiSys '10). ACM, New York, NY, USA

Delaney DG, Sperling CD, Adams CS, Leung B (2008) Marine invasive species: validation of citizen science and implications for national monitoring networks. Biol Invasions 10:1, doi:10.1007/s10530-007-9114-0

Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, Phillips T, Purcell K (2012) The current state of citizen science as a tool for ecological research and public engagement. Front Ecol Environ 10(6):291–297

Eiman K (2010) NoiseSPY: a real-time mobile phone platform for urban noise monitoring and mapping. Mob Netw Appl 15(4):562–574

Gallo T, Waitt D (2011) Creating a successful citizen science model to detect and report invasive species. Bioscience 61(6):459–465, doi:10.1525/bio.2011.61.6.8

Goodchild MF, Li L (2012) Assuring the quality of volunteered geographic information. Spat Stat 1:110–120, ISSN 2211–6753, available online http://dx.doi.org/10.1016/j.spasta.2012.03.002 Accessed 8 Oct 2014

Haklay M (2012) Citizen Science and Volunteered Geographic Information – Overview and Typology of Participation. In: Sui DZ, Elwood S, Goodchild MF (eds) Volunteered Geographic Information, Public Participation and Crowdsourced Production of Geographic Knowledge. Springer, Berlin

Haklay M, Basiouka S, Antoniou V, Ather A (2010) How many volunteers does it take to map an area well? The validity of Linus' Law to volunteered geographic information. Cartogr J 47(4):315–322, 10.1179/000870410X12911304958827

Juong-Sik L, Hoh B (2010) Dynamic pricing incentive for participatory sensing. Pervasive Mob Comput 6(6):693–708

Kazemi L, Shahabi C (2011) Towards Preserving Privacy in Participatory Sensing. In: 2011 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). IEEE Computer Society, Seattle, WA, United States

Kotovirta V, Toivanen T, Tergujeff R, Huttunen M (2012) Participatory Sensing in Environmental Monitoring – Experiences. In: Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS), 2012 Sixth International Conference, Palermo, Italy 4–6 July 2012., pp 155–162, doi:10.1109/IMIS.2012.70

Kuan LH, Salil SK, Wen H (2010) Are You Contributing Trustworthy Data?: the Case for a Reputation System in Participatory Sensing. In: Proceedings of the 13th ACM international conference on Modeling, analysis, and simulation of wireless and mobile systems (MSWIM '10). ACM, New York, NY, USA

Lane ND, Eisenman SB, Musolesi M, Miluzzo E, Campbell A (2008) Urban sensing systems: opportunistic or participatory? In: Proceedings of the 9th workshop on Mobile computing systems and applications, CA, USA, 25–26 Feb 2008

Leeuw T, Boss E (2014) Crowdsourcing water quality data using the iPhone camera. Poster presented at the 2014 Ocean Science Meeting, Honolulu, Hawaii, 22 Feb-3 March 2014. Available online http://www.eposters.net/pdfs/crowdsourcing-water-quality-data-using-the-iphone-camera.pdf Accessed 18 July 2014

Lottig NR, Wagner T, Norton Henry E, Spence Cheruvelil K, Webster KE, Downing JA, Stow CA (2014) Long-term citizen-collected data reveal geographical patterns and temporal trends in lake water clarity. PLoS ONE 9(4):e95769, doi:10.1371/journal.pone.0095769

Lowry CS, Fienen MN (2013) CrowdHydrology: crowdsourcing hydrologic data and engaging citizen scientists. GroundWater 51(1):151–156

Mednis A, Strazdins G, Zviedris R, Kanonirs G (2011) Real time pothole detection using Android smartphones with accelerometers. In: Proceedings of the 7th IEEE International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS'11), Barcelona, Spain 27–29 June 2011, pp 1–6

Mun M, Reddy S, Shilton K, Yau N, Burke J, Estrin D, Hansen M, Howard E, West R, Boda P (2009) PEIR, the personal environmental impact report, as a platform for participatory sensing systems research. In: MobiSys '09 Proceedings of the 7th international conference on Mobile systems, applications, and services, Kraków, Poland, 22–25 June 2009

Obrecht DV, Milanick M, Perkins BD, Ready D, Jones JR (1998) Evaluation of data generated from lake samples collected by volunteers. Lake Reserv Manag 14:21–27, doi:10.1080/07438149809354106

Olmanson LG, Bauer ME, Brezonik PL (2008) A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. Remote Sense Environ 112:4086–4097, doi:10.1016/j.rse.2007.12.013

Paxton M, Benford S (2009) Experiences of Participatory Sensing in the Wild. In: Proceedings of the 11th International Conference on Ubiquitous Computing (UbiComp '09). ACM, New York, NY, USA, pp 265–274

Perez A, Labrador M, Barbeau J (2010) G-sense: a scalable architecture for global sensing and monitoring. IEEE Netw 24(4):57–64

Rapala J, Kilponen J, Järvinen M, Lahti K (2012) Finland: Guidelines for Monitoring of Cyanobacteria and their Toxins. In: Chorus I (ed) Current Approaches to Cyanotoxin Risk Assessment, Risk Management and Regulations in Different Countries. Federal Environment Agency (Umweltbundesamt), Germany, pp 54–62, http://www.uba.de/uba-info-medien-e/4390.html. Accessed 18 July 2014

Reddy S, Estrin D, Srivastava M (2010a) Recruitment framework for participatory sensing data collections. In: Proceedings of the 8th International Conference on Pervasive Computing, Helsinki, Finland, pp 138–155

Reddy S, Estrin D, Hansen M, Srivastava M (2010b) Examining Micro-Payments for Participatory Sensing Data Collections. In: Proceedings of the 12th ACM International Conference on Ubiquitous Computing (Ubicomp '10). ACM, New York, NY, USA, pp 33–36

Reinart A, Kutser T (2006) Comparison of different satellite sensors in detecting cyanobacterial bloom events in the Baltic Sea. Remote Sens Environ 102:74–85

Roy HE, Pocock MJO, Preston CD, Roy DB, Savage J (2012) Understanding citizen science & environmental monitoring. Final report on behalf of UK Environmental Observation Framework. NERC Centre for Ecology & Hydrology and Natural History Museum. http://www.ceh.ac.uk/products/publications/documents/citizensciencereview.pdf

Solimini AG, Cardoso AC, Heiskanen AS (eds) (2006) Indicators and methods for the ecological status assessment under the Water Framework Directive, Linkages between chemical and biological quality of surface waters. European Commission, Luxembourg

Sunyoung K, Robson C, Zimmerman T, Pierce J, Haber E (2011) Creek Watch: Pairing Usefulness and Usability for Successful Citizen Science. In:

Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems (CHI '11). ACM, New York, NY, USA, pp 2125–2134
Toivanen T, Koponen S, Kotovirta V, Molinier M, Chengyuan P (2013) Water quality analysis using an inexpensive device and a mobile phone. Environ Syst Res 2(9) doi:10.1186/2193-2697-2-9
Yang H, Zhang J, Roe P (2011) Using reputation management in participatory sensing for data classification. Procedia Comput Sci 5:190–197

doi:10.1186/s40068-014-0024-8
Cite this article as: Kotovirta *et al.*: Participatory surface algal bloom monitoring in Finland in 2011–2013. *Environmental Systems Research* 2014 **3**:24.

Aalto-DD 252/2016
VTT SCIENCE 143

From time to time we are surprised by the environment. We get caught in the rain or may be disappointed when going for a swim in an algae-choked lake. Companies and businesses are also affected by the dynamics of the environment, and on a larger scale, societies are sometimes surprised by environmental hazards such as storms, fires, landslides, earthquakes and tsunamis. This thesis studied improvements in the timely delivery of relevant near real-time environmental information to support decision making in dynamic environments. The findings are not limited to ice navigation and water quality monitoring studied in this thesis, but can be applied to other domains of environmental monitoring as well, for example air quality, disaster and built environment monitoring. New methods are needed to utilise the increasing amount of environmental data in society.

ISBN 978-952-60-7168-8 (printed)
ISBN 978-952-60-7167-1 (pdf)
ISSN-L 1799-4934
ISSN 1799-4934 (printed)
ISSN 1799-4942 (pdf)

978-951-38-8480-2 (printed)
978-951-38-8479-6 (pdf)
2242-119X
2242-119X (printed)
2242-1203 (pdf)

Aalto University
School of Science
Department of Computer Science
www.aalto.fi

BUSINESS +
ECONOMY

ART +
DESIGN +
ARCHITECTURE

SCIENCE +
TECHNOLOGY

CROSSOVER

DOCTORAL
DISSERTATIONS