Invited paper

The progress in knowledge of physical oceanography of the Gulf of Finland: a review for 1997–2007*

Tarmo Soomere1,*
Kai Myrberg2
Matti Leppärinta3
Alexei Nekrasov4

1 Centre of Nonlinear Studies, Institute of Cybernetics at Tallinn University of Technology, Akadeemia tee 21, EE–12618 Tallinn, Estonia;
e-mail: soomere@cs.ioc.ee
*corresponding author

2 Finnish Institute of Marine Research, Erik Palménin aukio 1, PO Box 2, FIN–00561 Helsinki, Finland

3 Department of Physics, University of Helsinki, Gustav Hällströmin katu 2, PO Box 64, FIN–00014 Helsinki, Finland

4 Russian State Hydrometeorological University, Malookhtinsky prospekt 98, RU–195196 St. Petersburg, Russia


* This study was supported by the Estonian Science Foundation (Grant 7413), the Marie Curie RTN SEAMOCS (MRTN-CT-2005-019374), and the Maj and Tor Nessling Foundation (project Oil spill model for winter conditions in the Gulf of Finland).

A large part of the paper was written during the visits of one of the authors (TS)

The complete text of the paper is available at http://www.iopan.gda.pl/oceanologia/
Abstract

The main findings of studies of the physical oceanography of the Gulf of Finland (GoF) during 1997–2007 are reviewed. The aim is to discuss relevant updates published in international peer-reviewed research papers and monographs, bearing in mind that a comprehensive overview of the studies up to the mid-1990s is available (Alenius et al. 1998). We start the discussion with updates on the basic hydrographical and stratification conditions, and progress in the understanding of atmospheric forcing and air-sea interaction. Advances in the knowledge of basin-scale and mesoscale dynamics are summarised next. Progress in circulation and water exchange dynamics has been achieved mostly by means of numerical studies. While the basic properties of circulation patterns in the gulf have been known for a century, new characteristics and tools such as water age, renewal index, and high-resolution simulations have substantially enriched our knowledge of processes in the Gulf of Finland during the last decade. We present the first overview of both status and advances in optical studies in this area. Awareness in this discipline has been significantly improved as a result of in situ measurements. Our understanding of the short- and long-term behaviour of the sea level as well as knowledge of the properties of both naturally and anthropogenically induced surface waves have expanded considerably during these ten years. Developments in understanding the ice conditions of the Gulf of Finland complete the overview, together with a short discussion of the gulf’s future, including the response to climate change. Suggestions for future work are outlined.

1. Introduction

Located at the north-eastern extremity of the Baltic, the Gulf of Finland (hereafter denoted as GoF) is the second largest sub-basin in this sea (Figure 1). This gulf is an elongated estuary (length c. 400 km, width from 48 to 135 km) with a mean depth of only 37 m and an extremely complex coastline. The western end of the gulf is a direct continuation of the Baltic Sea Proper, whereas the eastern end receives the largest single fresh water inflow to the Baltic Sea from the discharge of the River Neva.

The water body of this extremely active region is involved in a wide range of physical, dynamic and optical processes, some of which are still poorly understood. These are, for example, the vertical turbulent transport in strongly stable or unstable atmospheric marine boundary layers, the role of coherent structures in local dynamics, or the consequences of strong

---

to the Centre of Mathematics for Applications, University of Oslo within the framework of the MC TK project CENS-CMA (MC-TK-013909). We are grateful to Helgi Arst and her optics team from the Estonian Marine Institute for long-term collaborative research in the Estonian-Finnish project SUVI supported by the Estonian Academy of Sciences and the Academy of Finland.
west-east salinity gradients due to the voluminous river runoff in the east. On top of that, a multitude of local circulation cells at different scales occur in both the atmosphere and the sea. This results in extremely complex patterns of air-surface interactions, especially over the archipelago area and in the vicinity of the ragged coastline. Furthermore, the stratification is highly variable in both space and time, being forced by the large seasonal variability of the incoming solar radiation. Both buoyancy-driven and wind-driven currents play an important role in the formation of the circulation together with the dynamic consequences of frequently occurring large sea surface slopes. The resulting complexity of the dynamics in this water body resembles the multitude of processes in the open ocean and extends from small-scale turbulence right up to classical synoptic-scale structures and certain specific features of the overall circulation (Alenius et al. 1998).
The marine meteorological conditions of the GoF are characterised by a remarkable wind anisotropy (Soomere & Keevallik 2003) as well as by the non-homogeneous patterns of atmospheric temperature due to the very variable surface roughness in and around the gulf (Figure 1), and also the resulting spatio-temporal variability of the momentum and heat exchange between the sea and the atmosphere. The largest variation of sea level in the Baltic Sea area occurs in the eastern part of the gulf. The total range of historical extremes is from under $-1$ m to $+4.21$ m. Probably the most intense fast ferry traffic in the world crosses the central part of the gulf. Although the gulf is partially open to the prevailing winds and quite high and long waves may occur here, the wakes from the fast ferries make up a considerable part of the total wave activity in certain sub-basins.

Studies of physical oceanography of the GoF were started by Admiral Makarov already in the late 1890s and continued up to World War II with the objective of studying the entire GoF as a separate water body with its own internal dynamics (see e.g. Witting 1909, 1910, 1912, Palmén 1930, Palmén & Laurila 1938). Regrettably, no research covering the entire gulf was thereafter possible until the 1990s. Post-World War II studies were done mostly at a local scale; problems with getting access to certain waters made it difficult for research vessels to carry out investigations covering the entire gulf. The results from the early studies up to the mid-1990s were reviewed by Alenius et al. (1998).

This paper attempts to review and summarise the main advances in the physical oceanography of the GoF during 1997–2007. This time interval starts from the trilateral Estonian–Finnish–Russian Gulf of Finland Year 1996 (Sarkkula (ed.) 1997), a milestone for the joint studies of this area that has separated two worlds since the 1940s. This decade included intense studies on circulation, stratification, waves, sea-level, ice, optics and marine meteorological conditions as well as extensive analyses of historical and recently-collected data sets, the rapid development of numerical modelling tools, and the introduction of new theoretical concepts. The output of the research eventually contributed to another milestone – the declaration of the Baltic Sea as a particularly sensitive sea area by the International Maritime Organisation at the end of 2005, with the status becoming effective from July 2006.

The aim of this paper is to review the updates of the relevant information and basic changes of concepts, bearing in mind that (i) a thorough overview of the discussed subjects is available up to the mid-1990s (Alenius et al. 1998), (ii) progress in knowledge of Baltic Sea physics is comprehensively described by Omstedt et al. (2004), Leppäranta & Myrberg (2008) and (iii) various aspects of climate changes have been extensively discussed by
the BACC Author Team (2008). The introductions to most of the chapters have been kept as short as possible and the reader is frequently referred to these sources. A few detours to basic concepts and to historical data are necessary, however, in the discussions of items and processes (such as air-sea interaction, marine optics, ship waves, and sea-ice matters) that were not presented in the earlier reviews of the GoF.

Several processes in the GoF, such as the dynamics of water level, wind wave conditions in its entrance section, or water exchange with the Baltic Proper, are directly dictated by, and some others are intimately related to, the processes and forcing factors in the entire Baltic Sea. For this reason, a few parts of this overview contain necessary information about the physical oceanography of the Baltic Sea; nonetheless, we have tried to keep such excursions as infrequent and short as possible.

The paper starts with a review of the main results of the studies related to basin-scale features, such as the general hydrography and stratification conditions. This is followed by a short look at the determining features of the basic forcing factors – atmospheric forcing and air-sea interaction. The presentation continues with a description of studies of the more detailed features of the dynamics of the gulf, such as basin-scale and meso-scale circulation, and water exchange with the Baltic Proper. The chapter dealing with the optical properties of the sea has no analogue in the earlier reviews of the GoF; this is why it contains a relatively large amount of general material. The three following chapters consider the advances in understanding the phenomena at the surface of the gulf: sea-level studies, investigations into surface waves, and research on ice conditions in the GoF. These three chapters also contain somewhat more background and historical information. Aspects regarding the future of the GoF (including the response to climate change) are discussed very briefly in the bodies of the relevant chapters, because a thorough overview of these items for the entire Baltic Sea is already available (BACC 2008). In addition, an overview of the research highlights reflecting climate changes and climatic shifts in the gulf itself is presented in Soomere et al. (2009).

These studies comprise more than 200 peer-reviewed research papers and several monographs. Because of this substantial enrichment of knowledge, gathered within just a single decade, our primary aim in the present paper is to describe the key aspects of progress in awareness about the GoF. One particularly instructive feature of those studies is their complexity and large coverage of various disciplines, embracing all the major directions of classical and contemporary physical oceanography. Only a few places in the World can boast such a large coverage.
2. Hydrography

2.1. Observations

The Gulf of Finland can be interpreted as a large estuary forced by the typical features of such water bodies: a freshwater inflow at its landward end, a more saline wedge of water penetrating into the water body along the bottom from the seaward end combined with strong mixing and large gradients in the whole water body (Alenius et al. 1998, Myrberg 1998). The salinity has pronounced horizontal and vertical gradients in the entire gulf. The average salinity increases from east to west and from north to south. The surface salinity changes from 5–6.5 PSU in the west of the GoF to 0–3 PSU at its easternmost end. A specific feature of the gulf is that horizontal gradients of salinity and temperature can occasionally be extremely large as a result of local upwellings.

Part of this knowledge has been updated and made more detailed on the basis of the extensive measurement campaigns carried out during the Gulf of Finland Year 1996 (see e.g. Alenius et al. 2003). The overall hydrography of the GoF is indeed characterised by large variations in both salinity and temperature. The major feature of the vertical structure of the water masses is the quasi-permanent halocline separating the upper, relatively well-mixed, warmer and fresher water from the near-bottom, more saline and usually colder water, thus creating a characteristic two-layer structure. The halocline lies at a depth of 60–80 m in the western GoF. The bottom salinity can reach values up to 8–10 PSU due to the advection of more saline water masses from the Baltic Proper, but it has a significant natural variability that reflects irregular saline water intrusions, changes in river runoff and the precipitation-evaporation balance, and fluctuating meteorological forcing (Elken 2006, Elken et al. 2006).

The classical concept of the eastern GoF is that the permanent halocline may be missing there and that the salinity increases approximately linearly with depth. There is, however, an extremely rich variety of different phenomena in this area. Some areas in the eastern GoF have been studied very intensively during the last ten years. A number of expeditions were organised in these regions within the framework of a so-called ‘floating university’ (Nekrasov et al. 2002a,b, 2004). This region has revealed a handful of interesting features, such as unexpectedly sharp contrasts in summer oceanographic conditions in the Luga-Koporye region in 1997 and 1998 (Plink et al. 1998, Nekrasov 1999). These aspects are discussed in more detail in Section 4.5.

In the eastern GoF, to the east of the island of Gogland (Suursaari), the depth of the upper mixed layer does not usually exceed 20 m. As the
length of waves, the major mixing agent in this practically non-tidal area, decreases in areas affected by the coast, the typical depth of the mixed layer is 12–15 m. In front of the St. Petersburg flood barrier, the mixing depth even after strong storms is as small as c. 10 m. The maximum vertical density gradient (sometimes called the jump layer core) lies 1–3 m below the lower boundary of the mixed layer.


A specific region is Neva Bay – today, the area between the flood barrier and the mouth of the River Neva. Its mean depth is 4–5 m. The salinity is practically zero and the water masses there are vertically almost homogeneous, except for the bottom waters of the shipping channel at depths of 10–14 m, where at times estuary-type intrusions of slightly more saline water occur.

The seasonal cycle of the sea surface temperature (SST) is pronounced owing to the large variations in solar radiation at the latitudes of the gulf (see Section 3.2). Typically, maximum SSTs (usually 15–20°C, at most up to 23°C) are observed in July–August. The horizontal gradients can occasionally be large as a result of local upwelling events. The seasonal thermocline starts to develop in May and reaches a depth of 15–20 m in summer, when it is usually sharply delineated. It begins to erode in late August.

The seasonal evolution of the temperature below the surface layer is governed by the classical features of deep-sea dynamics, like the convection and horizontal advection of water masses. The time lag between the observed maxima of the temperature at the sea-surface and in the lower layers increases with depth: for example, the maximum temperature at a depth of 30 m is usually reached in September–October. Section 3.2 gives a more detailed picture of the seasonal temperature cycle.

The stratification of temperature is indirectly coupled with that of salinity. In spring, for example, when the river discharge into the GoF reaches its maximum, the salinity stratification is the strongest because the freshwater remains in the surface layer. Owing to the moderate transparency of GoF water (Fleming-Lehtinen et al. 2007), warming as a result of intensive solar radiation is confined to a thin surface layer. As the wave activity is relatively weak in late spring and early summer (Soomere 2005a, Broman et al. 2006), the warmer surface layer only weakly interacts with the water
masses below during these seasons. This process further strengthens the overall stratification.

As a result of strong gradients, the baroclinic (internal) Rossby radius $R_1$ (defining to a large extent the typical scale of meso-scale dynamic features such as eddies, fronts and local jets) is very small. Its typical values were established on the basis of nearly 2000 CTD observations made in 1993–99 (Alenius et al. 2003). Usually (in 2/3 of all observations) $R_1$ was in the 2–4 km range (Figure 2), which is less than half the corresponding values in the Gotland Deep. This quantity also has a large spatial and temporal variability. The highest values were observed in summer in the deep parts of the gulf. In the shallow coastal regions and in the eastern GoF, $R_1$ is even smaller (Nekrasov & Lebedeva 2002).

![Figure 2. Spatial variability of the internal Rossby radius (adapted from Alenius et al. 2003; based on original data)](image)

There has been discussion recently on whether the increase in atmospheric temperature is directly reflected by Baltic Sea temperatures. Although reports of the changes in SST in both the Baltic Sea as a whole and the GoF in particular (Fonselius & Valderrama 2003, Rönkkönen et al. 2004, Siegel et al. 2006) disagree to a certain extent in their estimates, they all agree that the SST of the GoF has risen in recent decades in parallel with the air temperature. The rise has been c. 0.5–0.8°C during the last 15–40 years. As in the entire Baltic Sea, however, this rise does not necessarily mean an overall rise in the temperature of the whole water column (Omstedt & Nohr 2004).

The water salinity in the upper layers of the GoF is showing a decreasing trend, apparently reflecting changes in the balance between the influence of the Baltic Proper water masses at the western end and the freshwater inflow into the eastern part of the gulf. The nature of the long-term changes in the salinity of the entire Baltic Sea is not well understood; for example, the
present salinity in the Baltic is about the same as at the beginning of the 1900s (Winsor et al. 2001). Climate models generally forecast an increase in precipitation in the north-eastern regions of the Baltic Sea catchment area and a consequent increase in river runoff into the GoF (Graham 2004). A further decrease in the surface-layer salinity of the GoF is therefore probable.

2.2. Modelling

The ability of models to simulate the general features of the GoF’s hydrography has improved considerably in recent years. One of the key issues is the refinement in the models’ resolution in both the horizontal and vertical directions. The reason for paying special attention to the horizontal resolution is that the baroclinic Rossby radius $R_1$ is very small. To properly describe the small-scale eddies, fronts and jets, it has been suggested that the grid size should be $1/2–1/3$ of $R_1$ (Drijfhout 1989, Lindow 1997). Fulfilling this requirement is therefore still a challenge in long-term simulations covering several years.

Research models of the circulation in the entire gulf with a horizontal resolution down to $1 \times 1$ nautical miles have recently been implemented (Myrberg & Andrejev 2003, Andrejev et al. 2004a,b). Nested-grid approaches are often applied (Menshutkin 1997, Neelov et al. 2003, Tamsalu et al. 2003, Korpinen et al. 2004; see also the overview of nested operational models in Soomere et al. 2004) to ensure the best possible resolution for specific areas of interest. In such models the output from a large-scale low-resolution model is used to obtain the boundary conditions for a high-resolution sub-model. For limited areas (e.g. for forecasting floods in St. Petersburg) two-dimensional models using curvilinear grids have recently been applied (Klevanny et al. 2002, Klevanny & Mostamandy 2007, Averkiev & Klevanny 2007). This approach allows the use of high-resolution grids (down to 400 m in Neva Bay), which adequately trace the geometry of complex-shaped areas.

Simulations using high-resolution, three-dimensional versions of the classical circulation models reproduce the key features of the hydrography with reasonable accuracy. Surface salinity and vertical stratification are reproduced with an accuracy of 0.2–0.5 PSU, the sea-surface temperature with an accuracy of 0.5–1°C (Menshutkin 1997, Inkala & Myrberg 2001, Neelov et al. 2003, Myrberg & Andrejev 2003, Andrejev et al. 2004a,b).

The overall temperature stratification can also be modelled rather well by models based on somewhat different approaches to describe the physics. For example, the 13-box Baltic Sea model, which takes account of the barotropic and baroclinic dynamics of straits connecting 13 sub-basins of
the Baltic Sea (Omstedt & Axell 2003), can be tuned to provide accurate hindcasts for salinity and temperature distribution.

Abrupt changes due to upwelling are still a challenge for modellers (Myrberg & Andrejev 2003). This is due not only to shortcomings in the descriptions of vertical mixing processes (Meier 2003), but also to uncertainties in the available meteorological forcing. Another open question is whether the use of non-hydrostatic models in simulations of GoF dynamics would give better results than the models based on the hydrostatic assumption. The non-hydrostatic models are generally better at reproducing the basic hydrographical features and have certain advantages over the hydrostatic models in simulating the actual processes and hydrographic features (Zalesnyi et al. 2004, Myrberg et al. 2008, Passenko et al. 2008).

An additional reason for the improvements in model performance during recent years has been the refinement of the resolution of atmospheric models. Only recently has the resolution of these reached a satisfactory level of accuracy. For example, the resolutions of the FMI-HIRLAM (the version of the High Resolution Limited Area Model run by the Finnish Meteorological Institute) and of the WAM wave model have only since 2004 been extended to $0.08\times0.08$ (c. 9 km, see Soomere et al. 2008), as a result of which many local meteorological phenomena can now be described properly. The high resolution is genuinely necessary, because the GoF is very narrow and surrounded by land in all directions except to the west. Because of the resulting large gradients of key hydrometeorological fields, such as seashore roughness, heat and moisture exchange between the atmosphere and the sea, the patterns of atmospheric temperature and wind stress are at times very inhomogeneous. Certain generic features still cannot be described by the existing operational atmospheric models, for example, the sea breeze along the coasts of the GoF during summer and the accompanying low-level effects (Savijärvi et al. 2005).

3. Atmospheric forcing, freshwater and energy balance

3.1. Marine winds

The basic features of the winds in the GoF region are governed by the large-scale circulation patterns over the whole Baltic Sea. Estimated from the patterns of geostrophic winds, storminess was at a high level in most of this region from about 1881 to 1910, declining thereafter until around 1965. After that, the frequency of storms was found to have increased almost to the levels of the first decades of the 20th century (Alexandersson et al. 1998). These changes in forcing may be responsible for many of the trends and changes described below, such as the apparent change in the long-term
trend of relative mean sea level along the Finnish coast, or the increase in the annual sea level maxima (Section 6.3). They are also closely linked to variations in the NAO\textsuperscript{1} index.

The strongly anisotropic local wind climate is one of the main factors forming the local wave climate and sea level variations, and driving the associated sea-level-dependent coastal processes in the Baltic Sea (Hellemaa 1998, Andersson 2002). The prevailing W winds not only force water to enter the Baltic Sea through the Danish Straits, but also give rise to an eastward rising surface slope within the Baltic Sea. As this slope causes the largest water levels at the closed end of the Gulf of Finland, it is not unexpected that the annual means of the NAO index and the detrended annual mean sea level at all Finnish sea level stations have a high positive correlation with the significance level > 99.9\% (Johansson et al. 2001). The corresponding anisotropic wave regime (Jönsson et al. 2002, 2005, Soomere 2003) apparently contributes to the formation of specific features of the vertical structure of water masses at the entrance to the GoF (Elken 2006).

The wind data recorded along the northern coast of the gulf adequately represent the open-sea wind properties, but those recorded along the southern coast usually do not (Keevallik 2003a,b, Soomere & Keevallik 2003). This asymmetry reflects the pattern of dominant winds: onshore winds (for which the shore has less influence, Launiainen & Saarinen 1982) prevail on the northern coast and offshore winds on the southern coast. The tall cliffs along large sections of the latter is the probable cause of the mismatch between the observed and modelled wind properties (Ansper & Fortelius 2003).

The open-sea wind regime of the GoF reflects a superposition of frequently occurring SW and N–NW winds (predominating in the entire Baltic Sea basin) and more local but also relatively frequent W and E winds (E–NE winds on the southern coast) blowing along the gulf (Figure 3). SW winds are the most frequent. On the northern coast, moderate (6–10 m s\textsuperscript{-1}) and strong winds (> 10 m s\textsuperscript{-1}) come mostly from S or SW, but on the southern coast they more often blow from SW or W. SE winds are infrequent and relatively weak (Soomere & Keevallik 2003). The average wind speed may have opposite trends on the opposite coasts (Keevallik & Soomere 2004), a feature which may reflect long-term changes in the land

\textsuperscript{1}The NAO (North Atlantic Oscillation) index is the normalised sea-surface pressure difference between the Azores (Ponta Delgada) and Iceland (Stykkisholmur). It is positive when the atmospheric pressure is high in the south and low in the north. In this situation, strong W winds dominate in northern Europe and consequently, winters are mild in the Baltic Sea region. In the opposite case, usually cold but relatively weak E winds prevail in the GoF region.
surface roughness in the vicinity of some measurement sites. It may also be a reflection of the considerable changes in air-flow patterns, in particular, the variability of the mean wind direction in some seasons at the higher levels (Keevallik & Rajasalu 2001, Keevallik & Soomere 2008).

The mean wind speed during the most severe, long-lasting storms in the GoF (Table 1) is 2–3 m s$^{-1}$ less than in the Baltic Proper. S or SW winds are the strongest: the 3-hour mean wind speed may reach 24–25 m s$^{-1}$ once in a century. E winds are confined to a narrow direction span and may reach 23 m s$^{-1}$ (Soomere & Keevallik 2003). Further estimates of extreme

Table 1. The maximum 3-hour mean wind speeds [m s$^{-1}$] occurring once in 1–100 years. Based on Soomere & Keevallik (2003)

<table>
<thead>
<tr>
<th>Site</th>
<th>1 year</th>
<th>10 years</th>
<th>20 years</th>
<th>50 years</th>
<th>100 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>General</td>
<td>S-SW</td>
<td>E</td>
<td>N</td>
<td>NE, SE</td>
</tr>
<tr>
<td>Utö</td>
<td>20.9</td>
<td>23.7</td>
<td>24.5</td>
<td>25.4</td>
<td>26.2</td>
</tr>
<tr>
<td>Hanko</td>
<td>19.0</td>
<td>21.5</td>
<td>22.3</td>
<td>23.2</td>
<td>23.8</td>
</tr>
<tr>
<td>Isosaari</td>
<td>19.7</td>
<td>22.6</td>
<td>23.4</td>
<td>24.4</td>
<td>25.1</td>
</tr>
<tr>
<td>Kotka</td>
<td>18.4</td>
<td>21.2</td>
<td>22.0</td>
<td>23.1</td>
<td>23.8</td>
</tr>
</tbody>
</table>

*The island of Utö is located in the Baltic Proper, Hanko at the mouth of the GoF, Isosaari off Helsinki, and Kotka on the Finnish coast of the eastern GoF.
The estimates in Table 1 match the recorded wind properties in the strongest storms. The 6-hour mean wind speed reached 23 m s\(^{-1}\) near Ristinina harbour (Harju County, Estonia) in a very severe NNW storm on 15 November 2001 (Soomere 2005a) that caused the roughest wave conditions ever measured in the GoF (see Section 7.1). The maximum 10-minute wind speed was 20.5 m s\(^{-1}\) (with gusts of 30.3 m s\(^{-1}\)) at Pakri during windstorm Gudrun in January 2005 (Suursaar et al. 2006).

The wind regime of the GoF shows certain differences from that in the northern Baltic Proper. N to NW winds, the strongest in a large part of the northern Baltic Proper (Soomere 2001), are less frequent in the GoF. E winds, which are relatively weak in the Baltic Proper (Soomere & Keevallik 2001), may be nearly as strong as SW winds and may dominate during March-May in the GoF (Mietus 1998). The latter is the calmest period in the GoF (Niros et al. 2002); wave records suggest, however, that March and April may be fairly windy in the Baltic Proper (Broman et al. 2006). Unlike the wind regime in the Baltic Proper, the directional distribution of strong winds often agrees poorly with the analogous distribution for all winds (Figure 3). This means that the strongest winds may blow from directions from which winds are relatively infrequent (Soomere & Keevallik 2003).

The frequent, relatively strong afternoon surface winds observed along the northern coast of the GoF in typical summer conditions may reflect the interplay of the basic flow, seabreeze and the geometry of the gulf (Savijärvi et al. 2005). The meanders of the sea breeze, amplified by the unidirectional basic flow, may become evident as relatively strong winds, at times located just above the opposite coast. The anisotropy of the predominating winds suggests, however, that such strong coastal winds are more likely to occur near the Finnish coast.

Major progress in understanding the marine meteorological conditions in the GoF has become possible as a result of the availability of data from Kalbådgrund, a lighthouse located in the central part of the GoF (59°58’N, 25°37’E) at a distance of 29 km from the Finnish and 37 km from the Estonian coast. This is the only measurement point in the gulf where E and W winds are not affected by the coast; wind data from some other directions may contain some mainland-induced distortions. It provides key reference material when studying the accuracy of the wind forecast by meteorological models (Myrberg 1997). Another data set, more or less representative of the open sea wind regime, has recently been restored from the measurements carried out on the island of Naissaar (Keevallik 2003a).
The distribution of wind speeds at Kalbådagrund is close to the Weibull (Gnedenko) distribution (Niros et al. 2002) with the probability density function \( f(u) = ku^{k-1}b^{-k}\exp[-(u/b)^k] \), where \( u > 0 \) is the instantaneous wind speed, \( \bar{u} = b\Gamma(1 + 1/k) \), \( \bar{u}^2 = b^2\Gamma(1 + 2/k) \), \( \Gamma \) is the gamma function, and the bar implies a sample mean. The wind regime matches the general features of the north European wind climate, where the shape parameter \( k \approx 2.0 \) (Troen & Petersen 1989) and the wind speed distribution is close to the Rayleigh distribution.

Measurements from vessels and at lighthouses reveal that wind speeds on the open sea are considerably greater than those estimated from coastal data or from older lighthouse data (Niros et al. 2002). According to 1990s wind data, the probability of strong storms (wind speeds > 20 m s\(^{-1}\)) is dramatically greater than according to earlier data. Wind speeds exceed 10 m s\(^{-1}\) with a probability of 24% (13.6% for the period 1977–82) and 20 m s\(^{-1}\) with a probability of 0.2% (0.008%). More data concerning the general wind properties in the GoF can be found in Lopatukhin et al. (eds.) (2006a) and Lopatukhin et al. (2006b).

Considerable differences in the properties of air temperature and wind speed compared to the climatology of the Baltic Sea basin for the period 1961–90 (Mietus 1998) are reported by Niros et al. (2002) on the basis of original data from open-sea regions. The differences stem largely from the inadequacy of earlier data and are only partly related to the warm and windy conditions of the 1990s.

### 3.2. Air-sea interaction, fluxes, freshwater and energy balance

The volume of the GoF is 1100 km\(^3\). Annually, 112 km\(^3\) of freshwater are added to the basin by river runoff (Mikulski 1972, Bergström et al. 2001) plus the net atmospheric flux (precipitation minus evaporation). The freshwater flux is strongly dominated by the River Neva at the eastern end of the basin, which carries waters from the Finnish Saimaa Lake District and the Karelian lakes. Waters from the Finnish Central Lake District flow into the Gulf of Finland via the Kymijoki (River Kymi), while those from the Lake Peipsi catchment area are brought in by the River Narva. The monthly means peak at 50% above the annual mean level in May–June, whereas the winter level is c. 25% below the mean level (e.g. Bergström et al. 2001).

The atmospheric net contribution is fairly small – c. 7.6% of the total net contribution in this area (Table 2 in Omstedt & Axell 2003). Evaporation removes 450–500 mm water per year, whereas precipitation is approximately 600 mm. The compensating loss of water takes place through the western entrance of the basin, as discussed in Section 4.2.
In the GoF, the heat budget of the upper layer is driven mainly by solar and atmospheric forcing, whereas in the lower layer, horizontal advection of heat plays a very important role. The heat content of this basin shows an annual amplitude of $46 \times 10^{18}$ J, which corresponds to a mean temperature amplitude of 10.2°C, as shown already in the 1930s (Jurva 1937). The dominant part of the heat flux passes through the surface of the gulf. In summer the upper layer warms up, and on top of that, a surface layer forms with typical temperatures of 15–20°C. Beneath the surface layer the temperature decreases to a minimum of 0–2°C in the deeper basin halocline. Beneath the halocline the water is warmer due to horizontal advection. In winter the upper layer cools down to freezing point. In the eastern GoF the inflow of the River Neva has an influence on the temperature conditions and, technically, needs to be included in local studies.

Such a large annual cycle in the heat content means that the heating and cooling rates are of the order of 100 W m$^{-2}$ at the air-sea interface. It is convenient to take the zero reference of heat as the heat content of liquid water at 0°C; the ice represents ‘negative heat’ mainly with its latent heat. The heat content then becomes negative in normal winters in the GoF (Jurva 1937). The role of ice formation is not only to serve as a heat sink but it also serves as a solid interface protecting the underlying liquid water masses.

The surface heat exchange consists of incoming and outgoing solar radiation ($Q_s$ and $Q_r$, respectively), incoming thermal radiation from the atmosphere $Q_{La}$ and outgoing thermal radiation from the surface $Q_{Lo}$, sensible and latent heat fluxes ($Q_c$ and $Q_e$), and heat input from precipitation $Q_p$:

$$Q_n = Q_s - Q_r + Q_{La} - Q_{Lo} + Q_c + Q_e + Q_p.$$  

Here, $Q_n$ is the net gain at the surface. The sum $Q_R = Q_s - Q_r + Q_{La} - Q_{Lo}$ is the radiation balance, and the sum $Q_c + Q_e$ expresses the turbulent heat exchange.

In GoF latitudes there is a large annual cycle in the radiation. While the incoming solar radiation is close to zero in mid-winter, daily averages reach 200–300 W m$^{-2}$ in summer. The maximum level is c. 600 W m$^{-2}$. The Jokioinen station in southern Finland (60°49′N, 23°30′E, located slightly to the North of the GoF) well characterises the pattern of monthly solar radiation means in the GoF region: the average level there is from 10 W m$^{-2}$ in December–January to 250 W m$^{-2}$ in June.

The terrestrial radiation balance $Q_{La} - Q_{Lo}$ is negative except for very rare, extreme situations. When the temperatures of the atmosphere and the surface are equal ($T_o = T_a$), the balance equals c. –50 W m$^{-2}$ due to the lower emissivity of the atmosphere. The total radiation balance
(terrestrial + solar) turns positive from some time in March–April, when the albedo has become low enough, and later in ice-covered areas. In September–October the balance becomes negative for autumn and winter.

The air-sea surface temperature difference may have either sign. The sensible heat flux is usually within ±50 W m⁻². The direction of monthly median sensible heat flux was 0–5 W m⁻² upwards in April, July and August, according to measurements of air-sea exchange processes over the northern Baltic Sea based on data from three open-sea lighthouse meteorological stations, including Kalbådagrund, and the research vessel’s ‘Aranda’ automatic weather station in the period 1991–99 (Niros et al. 2002). Modelling investigations of the heat fluxes have been performed by Meier & Dösch (2002).

In May and June, a stable stratification and a downward sensible heat flux prevailed. The monthly mean latent heat fluxes were c. 25 W m⁻² and generally directed from sea to air during the whole study period (April to August), but downward latent heat fluxes did occur occasionally. Both the relative humidity and sensible heat flux showed pronounced diurnal variation, with maximum values in the morning (occasionally exceeding the means by one order of magnitude), while the latent heat flux reached its maximum in the afternoon.

The latent heat flux is mostly negative and in magnitude close to the sensible heat flux in the Baltic Sea. Therefore, the turbulent fluxes together are usually less than terrestrial radiation losses. Only when the air-sea temperature or humidity difference is large and the wind speed is high can the turbulent flux be a governing term. This is normally the case in late autumn. Heat transfer from precipitation takes place as sensible heat exchange and as phase changes. If the precipitation were solid, the phase change would take 30 W m⁻² heat from the surface layer, i.e. the solid precipitation would melt and the surface would cool down. Consequently, specific precipitation events (such as snowfall on open water, rain on snow, rain on an ice surface) are important heat transfer mechanisms.

The net heat flux \( Q_n \) through the sea surface is of the order of \( Q_n \sim 100 \text{ W m}^{-2} \). As expected, it is positive in summer and negative in winter (see also Meier & Dösch 2002). This value of \( Q_n \) corresponds to heating or cooling at the rate \( Q_n/(\rho c H) \), i.e. c. 0.04°C day⁻¹ for the whole upper layer (\( H \sim 50 \text{ m} \)) or 0.2°C day⁻¹ for the surface layer (\( H \sim 10 \text{ m} \)). Table 2 shows the magnitude of the radiational and turbulent fluxes throughout the year.

The occasional presence of ice considerably affects the air-sea fluxes in the GoF. In winter there is ice in the GoF, and the mixed layer extends to the halocline. The terrestrial radiation and turbulent losses are strong
Table 2. Mean course of the magnitudes of heat fluxes in the Gulf of Finland, assuming that the sea surface is ice-covered in winter. All data are given in W m$^{-2}$ (Leppäranta & Myrberg 2008)

<table>
<thead>
<tr>
<th>Season</th>
<th>Solar radiation</th>
<th>Terrestrial radiation</th>
<th>Turbulent fluxes</th>
<th>Net</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>incoming</td>
<td>outgoing</td>
<td>sensible heat</td>
<td>latent heat</td>
</tr>
<tr>
<td>winter</td>
<td>10</td>
<td>240</td>
<td>-300</td>
<td>-20</td>
</tr>
<tr>
<td>spring</td>
<td>130</td>
<td>280</td>
<td>-330</td>
<td>10</td>
</tr>
<tr>
<td>summer</td>
<td>200</td>
<td>320</td>
<td>-390</td>
<td>10</td>
</tr>
<tr>
<td>autumn</td>
<td>30</td>
<td>300</td>
<td>-360</td>
<td>-20</td>
</tr>
<tr>
<td>mean</td>
<td>360</td>
<td>1140</td>
<td>-1380</td>
<td>-20</td>
</tr>
</tbody>
</table>

when the sea-surface is open. Ice formation and snowfall with the resulting low surface temperature drastically weaken turbulent transfer. The level of solar radiation is also low during most of the ice season.

In spring the radiation balance turns positive, ice and snow melt, and the surface temperature increases. Turbulent transfer is weak due to low winds (as demonstrated in Section 3.1) and the stable stratification of the atmospheric surface layer. In this situation a relatively thin upper layer is formed, with the thermocline located at a depth around 10 m. In summer solar heating is strong and the surface temperature reaches a maximum in early August (15–23°C). The thermocline deepens to 15–20 m by the end of summer. The radiation balance turns negative in September. The basin cools as winter approaches, and terrestrial radiation losses, sensible heat flux and latent heat flux are all of equal magnitude. The thermocline extends down to the halocline.

Horizontal advective and diffusive heat transfer can be comparable to the upper layer atmospheric heating. If the horizontal current velocity and temperature gradient are $U \sim 5$ cm s$^{-1}$ and $\Delta T/L \sim 1°C$ (100 km)$^{-1}$, the advective change will be $U\Delta T/L \sim 0.04°C$ day$^{-1}$. Taking the horizontal diffusion coefficient as $D \sim 10^2$ m$^2$ s$^{-1}$ and temperature variations as 1°C over 10 km distances, the diffusive smoothing rate will be $D\Delta T/L^2 \sim 0.1°C$ day$^{-1}$. The lower layer is largely decoupled from the upper layer by the strong halocline, and its heating is taken care of by advection from the Gotland Sea.

Potential climate change estimates suggest that the annual mean temperature over Finland, including the GoF, will rise by a few degrees within a few decades (Jylhä et al. 2004). The latest results (IPCC 2007) suggest that the rise in air temperature may be c. 0.3°C per decade; the exact rate will depend strongly on the future emission of CO$_2$ to the atmosphere. Such changes may substantially modify the properties of
air-sea interaction processes and local meteorological conditions. Unfortunately, the relevant parameters were not estimated in recent local studies of climate change impact (Carter 2004).

4. Dynamics of water masses

4.1. Circulation and drift

A traditional but idealised view of the mean circulation of the GoF is that it is cyclonic with an average velocity of a few centimetres per second. These basic properties were identified nearly a century ago (Witting 1912, Pålmen 1930). These descriptions, however, should be regarded as educated guesses based on very sparse observational data. In fact, at long time-scales the circulation is intrinsically baroclinic due to the pronounced horizontal buoyancy gradients. At shorter time scales wind-stress at the sea-surface dominates. The sea-surface slope that results from the permanent water supply to the eastern part of the gulf also appreciably contributes to the existing circulation. Moreover, the gulf is large enough to experience the effects of the Earth’s rotation (Witting 1912, Pålmen 1930, Hela 1946).

The GoF circulation was modelled by Andrejev et al. (2004a) for 1988–92 with an accuracy basically sufficient for resolving meso-scale eddies in this area. The horizontal resolution of $1 \times 1$ nautical miles is almost enough to resolve the baroclinic Rossby radius (Alenius et al. 2003), yet the synoptic-scale phenomena occurring in places may be quite small and their dynamics belong rather to the sub-grid scales. Generally, there is a discernible cyclonic mean circulation, but the resulting patterns and the persistency of the currents deviate to some extent from the classical analyses of Witting and Pålmen. Both the mean and instantaneous circulation patterns in the Gulf of Finland contain numerous meso-scale eddies (identifiable as analogues to oceanic synoptic rings) with a typical size clearly exceeding the internal Rossby radius (Andrejev et al. 2004a). The modelled circulation patterns reveal certain nontrivial vertical structures that vary temporally and spatially.

The shallow Neva Bay is characterised by strong and persistent N to NW currents caused by the voluminous runoff from the River Neva.

The surface-layer (depths 0–2.5 m) flow pattern is characterised mainly by an Ekman-type drift in the rest of the GoF. There is an inflow near the southern coast at all depths with a high degree of persistency (up to 50%). A compensating outflow exists in the rest of the gulf. It becomes stronger and is highly persistent (up to 80%) in the sub-surface layers (2.5–7.5 m and downwards) slightly north of the gulf’s axis (c. 30 km off the Finnish coast) and weakens considerably near the bottom.
These features evidently play a major role in the functioning of the gulf. Results suggest that the external loading from the St. Petersburg area does not always directly affect the environmental state of the Finnish coastal areas and that the high nutrient concentrations in Finnish coastal waters are at times due to local sources.

The relatively persistent eastward flow in the sub-surface layer (depths 2.5–7.5 m; Figure 4), which exists along the whole gulf to the north of its axis, may dominate in quite a thick upper layer of the sea during certain seasons, as hypothesised by Soomere & Quak (2007). Multi-layer flows have also been detected for several bays on the southern coast of the GoF (Raudsepp 1998).

![Figure 4. The simulated mean circulation [cm s$^{-1}$] in the sub-surface layer between 2.5 and 7.5 m from September 1987 to August 1992 (upper panel, adapted from Andrejev et al. 2004b) and its schematic representation (lower panel, after Andrejev et al. 2004a); by kind permission of O. Andrejev](image-url)
Drifter experiments and corresponding modelling studies in the Gulf of Finland are reported in Gästgifvars et al. (2006). In two experiments in May 2003 the buoys moved with a velocity of 2% of the wind velocity and with a deviation angle of 0–10° to the right with respect to the local wind in moderate wind conditions. This tallies quite well with earlier observations, where wind-driven surface currents were c. 1–3% of the wind speed (Hela 1952). In contrast, the behaviour of drifters was quite different in weak wind conditions. The buoys indicated that the currents were directed some 60° to the left of the wind. This was apparently due to meso-scale eddies in the surface current field in which the dynamics of the lower layers overrode the local wind-induced drift. This result suggests that operational drift forecast models based solely on local wind information need to be further improved in order to reproduce accurately the actual hydrodynamic flow field and the corresponding drift.

4.2. Water exchange with the Baltic Proper

The net budget of water exchange between the GoF and the Baltic Proper across the Hanko-Osmussaar line was adequately estimated already by Witting (1910). But the magnitudes of the in- and outflows of 480 and 600 km$^3$ y$^{-1}$, respectively, would correspond to a mean flow of only 1 cm s$^{-1}$, which is clearly less than observed values (Mikhailov & Tshernyshova 1997). Moreover, by using numerical modelling tools, Lehmann & Hinrichsen (2000) found that the difference between the in- and outflows is around 130 km$^3$ y$^{-1}$. The estimated total in- and outflows, however, were much larger than those of Witting (1910).

The large deviations between these different estimates stem from the complexity of the water dynamics in the entrance area of the GoF. Leaving aside the small-scale dynamics, which is accurately reflected neither by measurements nor in numerical simulations, two basic processes occur in this region. Firstly, relatively persistent or quasi-periodic, meso-scale features (local jets, synoptic rings, inertial oscillations) induce relatively short-term transport of water across the entrance line. This water is usually not transported far from its original location and does not affect the dynamics of the interior of the gulf; however, its motion is resolved in contemporary research models. Secondly, the net in- and outflow reflects the amount of water entering the interior of the gulf, remaining there for a long time, and ultimately being transported out of it. Witting (1910) did make a rough estimate of this flow.

Estimates of the in- and outflow across the Hanko-Osmussaar section essentially depend on the time interval used in the averaging (Andrejev et al. 2004a). As expected, the largest values of the in- and outflow are obtained
if the instantaneous water dynamics at every time step is accounted for. A time step of c. 30 minutes appears to be short enough to properly account for the local meso-scale dynamics. According to the 5-year model run in 1987–92, the short-term meso-scale circulation system at the entrance of the GoF gives rise to average in- and outflows of 3154 and 3273 km$^3$ y$^{-1}$, respectively.

Another estimate was obtained for the water exchange using average velocities for the 5-year period. Such averaging damps out the short-term variability but accounts for the quasi-stationary features of the meso-scale circulation pattern at the western end of the GoF. Doing so results in average in- and outflows of 1417 and 1532 km$^3$ y$^{-1}$, respectively. This estimate of the net water exchange (119 and 115 km$^3$ y$^{-1}$ according to the two averaging schemes) is close to the total river runoff into the GoF.

The mean current field at the mouth of the GoF frequently shows a multi-layer velocity structure (Andrejev et al. 2004a) similar to that in the above-discussed interior dynamics. Typical velocities of the mean inflow adjacent to the Estonian coast over the entire water column are 1–4 cm s$^{-1}$. The most intense flow (7–10 cm s$^{-1}$) occurs near the surface. The northern side of the GoF is dominated by an offshore outflow (mean velocity of 8 cm s$^{-1}$) over a substantial range of depths down to 40–50 m. Somewhat counter-intuitively, a comparatively intense (also up to 8 cm s$^{-1}$) transport (apparently reflecting the Ekman drift) into the gulf takes place slightly to the north of its axis in the surface layer (see Meier 1999).

In contrast to classical estuarine dynamics, which is governed mostly by the balance between the freshwater flow and the impact of open sea water masses, the strongly anisotropic local wind dynamics in the GoF plays an important role in its water balance and in the formation of the vertical structure of the water masses at its entrance. Whereas the standard estuarine circulation is supported by most of the wind directions, long-lasting, strong SW winds push a large amount of fresher surface water into the gulf. The excess volume of water increases the hydrostatic pressure in the gulf and may lead to a gradual export of the salt wedge in the bottom layer of the gulf (Elken et al. 2003). A reversal may occur if SW wind speeds exceed a mean value of 4–5.5 m s$^{-1}$.

This effect mirrors the sporadic salt water inflow through the Danish Straits (e.g. Lass & Matthiäus 1996) into the Baltic Sea. Instead of a salt water inflow through shallow straits and the strengthening of stratification in the western Baltic Sea, export of saline bottom water accompanied by drastic variations in the halocline’s position or by its almost complete disappearance may take place here (Figure 5).
A detailed investigation of the position and properties of the halocline in the northern Baltic Proper and in the GoF, as well as of related meso- and basin-scale processes, suggests that the major consequence of this kind of counterestuarine transport is a weakening of the stratification of water masses at the entrance of the gulf, accompanied by an intensification of vertical mixing (Elken et al. 2006). The practical consequences of this finding for the functioning of the deep-water ecosystem are not understood yet; for example, potentially intense up- and downwelling may effectively supply oxygen to the deep layers of the northern Baltic Sea (Figure 6; Elken & Matthäus 2008).

Therefore, both the surface and near-bottom layers of the GoF respond rather actively to wind forcing. The basin-scale barotropic flows are frequently transformed into topographically-driven baroclinic meso-scale motions characterised by large isopycnal displacements (more than 20 m within a distance of 10–20 km) and high intra-halocline current speeds (>20 cm s⁻¹). This finding has led to a substantial revision of the traditional concept of mostly decoupled lower layer dynamics (Elken et al. 2006). The multitude of processes at the entrance to the GoF (Figure 6) certainly makes their modelling extremely difficult. Internal wave activity, a process scarcely reflected in recent studies, is apparently high in this
4.3. Water age

The calculated amounts of in- and outflowing water discussed in the previous section provide rather limited information on the factual water renewal rates and water quality in the GoF. The good water quality in semi-enclosed sea areas is intuitively associated with an intense water exchange and a low water age. The regions associated with the highest water ages and weak renewal are the most vulnerable to eutrophication, in particular to the
sedimentation of organic matter, and hence specific focus on the welfare of these areas may be advisable.

A straightforward definition of the age of a water particle is the time elapsed since the particle left the sea surface (Deleersnijder et al. 2001). This definition is relevant in the GoF, where the salt wedge is usually quasi-stationary and water is supplied either as near-surface currents or through precipitation, or by river discharge. According to this definition, the oldest water is evidently located at the bottom and in the GoF is c. 8.3 years old (Meier 2005).

A more general definition of water age accounting for both horizontal and vertical displacements (Bolin & Rodhe 1973) and also a renewal index were used by Andrejev et al. (2004b) to characterise the water residence time. The former reflects the average time of residence of water particles in a selected area of the basin, the latter the amount of original water in the area at a given later time. The local maxima of these two quantities do not necessarily coincide, however; simultaneous analysis of these variables provides much better information about the water quality. This concept is similar to the one used by Meier (2007) for the entire Baltic Sea water renewal, which puts the age of GoF water at c. 20 years.

The water age in the GoF with respect to water exchange with the Baltic Proper is at most two years. Such small values confirm the very intense water exchange between the gulf and the Baltic Proper. The water age shows a pronounced spatial and temporal variability (Andrejev et al. 2004b). Its smallest values are found in the inflow regions and at river mouths (Figure 7), the highest ones in the outflow regions. The overall

Figure 7. Spatial distribution of water age in the Gulf of Finland; based on simulations for 1987–1992 (after Andrejev et al. 2004b, by kind permission of O. Andrejev)
distribution of the water age mimics the cyclonic mean circulation, and its mesoscale variability reflects the internal dynamics of the gulf, where a number of eddies, fronts and circulation cells are quasi-stationary; for example, in areas with persistent eddies the water age tends to be significant. During winter, convection decreases the vertical variability of the water age.

The highest water ages are found in the south-eastern parts of the gulf and are reflected by the environmental state of this region. The area of weakest water renewal combined with significant water age was in and to the north of Koporye Bay. This is consistent with recent observations of the oxygen deficit under the halocline during the summer of 1996 (Ljahkin et al. 1997) and can be explained by reduced ventilation through the halocline in conditions of extremely strong stratification (Section 2.1).

4.4. Mesoscale processes and frontal dynamics

The fine-structure of GoF dynamics has attracted several researchers during this last decade. Numerical studies of GoF circulation patterns reveal manifold meso-scale features that are probably jointly driven by the complex bottom topography and the small internal Rossby radius (Alenius et al. 2003). In the eastern part of the central gulf vortices are relatively small (Andrejev et al. 2004a); in addition, the Rossby radius there is particularly small, and this model was apparently unable to resolve certain features.

The above-described outflow to the north of the central axis can be interpreted as a persistent buoyant current triggered by the classical counterclockwise circulation in northern hemisphere basins due to the Coriolis effect and stabilised by both the presence of the coast and certain specific features of baroclinic instability (Stipa 2004). Generally, baroclinic instability leads to disintegration of flow, but the relatively gently sloping bottom off the Finnish coast causes an internal baroclinic adjustment of the flow to a quasi-stable state. Equivalently, the outflow can be interpreted as a (coastal) jet current along the bottom isolines. Such currents (analogous with nearly zonal flows on the beta-plane) are generally more pronounced in a baroclinic medium than in barotropic water bodies (although they are somewhat more apt to meander across the slope (Soomere 1995, 1996).

The structure of local currents in the relatively deep, semi-enclosed bays with large open boundaries with the rest of the GoF is expected to frequently follow the multi-layered flow structure of the entire gulf. For example, the general anticyclonic flow in the upper layer and the cyclonic flow in the lower layer of Muuga Bay are substantially affected by the local topography, which gives rise to horizontal and vertical circulation cells (Raudsepp 1998). The observed dynamics of the coastal jet was interpreted
in terms of travelling coastally trapped waves (as had been found in many other parts of the Baltic Sea earlier – Talpsepp 2006), although only a quarter of the eastward travelling wave period (around 40 days) was covered by the measurements. The wave supposedly reversed the cross-shore flow direction, while the longshore velocity was preserved (Raudsepp 1998). The velocity of the current displayed a periodic variability between 5 and 25 cm s\(^{-1}\) and a period of 3–4 days.

There have been very few observations of meso-scale eddies in the GoF. A rare description of such a structure at the entrance of the gulf was given by Pavelson (2005). The eddy was formed during the rapid splitting of the eastward downwelling jet into an offshore cyclonic and an onshore anticyclonic branch. The latter can be interpreted as an incipient eddy; the mechanism of formation is still unclear. It could be associated with the jet instability generated by the particular topographic features (see Stipa 2004). An anticyclonic velocity pattern was observed until the end of the experiment. The centre of the eddy lay within the seasonal thermocline at 30–35 m depth and had a diameter of c. 15–20 km (c. 4 times the internal Rossby radius). The maximum velocities in its core reached 0.35 m s\(^{-1}\), but its vertical structure was typical of that of a geostrophically balanced eddy.

A key process in the hydrography of the entire Baltic Sea is frontal development. Practically all its sub-basins are rich in fronts that separate water masses with different properties. The diversity of water masses follows from the estuarine character of the entire Baltic Sea and is enhanced by water exchange between the sub-basins and the vigorous modulation of patterns of currents by coastal processes. Fronts can be found all over the entire water body, but those occurring in the uppermost layer are studied in detail because of their important coupling with biochemical processes (Pavelson 2005).

The quasi-permanent salinity (density) front at the entrance to the GoF (Kononen et al. 1996, Pavelson et al. 1997, Laanemets et al. 1997) is the best known example of such dynamic features (Figure 8). The results of multidisciplinary field studies carried out onboard r/v ‘Aranda’ during the 1990s were summarised by Pavelson (2005). The persistence of the front can be explained by the existing mean cyclonic circulation in the GoF. The more saline water of the northern Baltic Proper enters the gulf along the Estonian coast, and the seaward flow of fresher water takes place along the opposite coast. The interface of these water masses of different salinity and/or temperature forms this quasi-permanent front in the entrance area of the GoF.
The front is typically oriented SW–NE and is positioned approximately parallel to the bottom slope. The frontal area responds to wind forcing in the following way. With E winds, the denser (more saline) water mass moves offshore. The front becomes sharp and is strongly inclined to the sea surface. When winds are from the W, the less dense water of the GoF forms a surface sublayer, giving rise to a secondary pycnocline approximately in the middle of the upper layer. The changes in the potential energy (stratification conditions) in the upper layer are coupled with differential advection induced by along-front wind stress and wind-generated vertical mixing (Pavelson et al. 1997).

4.5. Upwelling and downwelling as the major agents of mixing

Other key elements of meso-scale sea dynamics in the GoF are mechanically forced vertical motions of water masses called upwelling and downwelling. Upwelling usually results from the horizontal divergence of wind-driven motions in the surface layer and consists in the penetration of dense, cooler and usually nutrient-rich water towards the sea surface. Its complementary process – downwelling – takes warmer, usually nutrient-depleted surface water to the lower layers. Upwelling events play a key role in enhancing primary production and in phytoplankton dynamics over larger areas by replenishing the euphotic zone with the nutritional components that limit biological production for most of the growing season. It is of special importance for meeting the nutritional requirements of late-summer cyanobacterial blooms (Vahtera et al. 2005).
In the GoF upwelling is typically triggered by longshore winds (see e.g. Haapala 1994). To a limited extent it may also be caused by local currents and long waves excited by contemporary fast ferries (Madekivi 1993, Lindholm et al. 2001). Owing to the complex coastline and the many islands in the GoF, wind from virtually any direction may cause up- or downwelling near a certain coast with accompanying vertical mixing and displacement of water masses. The upwelling-driven vertical mixing in coastal waters is also dependent on freshwater and long-wave forcing (Laanearu 2003, Laanearu & Lips 2003). During summer and autumn, when the sea surface is warm, upwelling results in a local temperature drop of several degrees Celsius and is easily detected from thermal infrared satellite images. The major primary effects of upwellings are observed in a 5 to 10 km broad coastal zone (Lehmann et al. 2002, Myrberg & Andrejev 2003), but the nutrient-rich waters gradually spread offshore as a result of advection and horizontal diffusion in the form of so-called upwelling filaments. To describe their dynamics the 1-mile resolution models are still too coarse, and hence even more refined models have been developed to describe the fine-scale structure of upwellings (Zhurbas et al. 2006).

A remarkable upwelling event in terms of scale and intensity was observed on 17–25 July 1997 in Luga and Koporye Bays in the southern part of the eastern GoF (Nekrasov 1999). The uplifting of near-bottom waters destroyed the usual two-layer thermohaline structure, forming an extensive (covering more than $100 \times 30$ km) near-coastal patch with abnormally low surface temperature ($T_{\text{min}} < 4^\circ\text{C}$) and relatively high salinity.

![Figure 9. Maps of sea surface temperature (left) and salinity (right) in the eastern Gulf of Finland on 17–18 July 1997. Original images by A.V. Nekrasov, parts of which were published and discussed in Nekrasov (1999)](image-url)
(S_{\text{max}} > 5 \text{ PSU}) in Luga Bay (Figure 9). The horizontal gradients were exceptionally large, in excess of 3.5°C km\(^{-1}\) and 0.2 PSU km\(^{-1}\). The warm surface layer, with characteristics typical of this season (T \approx 18 - 20\°C, S \approx 2.5 - 3.5 \text{ PSU}), was found only to the north of latitude 60\°N (Figure 9). The speed of the prevailing wind, directionally favourable for such local upwelling, was quite low (generally below 5 m s\(^{-1}\), at times up to 7 m s\(^{-1}\)), considering the vigour of this event. This upwelling appeared to have been favoured by the very warm and calm weather in the preceding month: intense warming combined with fairly limited wave-induced mixing led to the formation of a very warm, uncommonly thin surface layer (1–3 m) on top of a ‘jump layer’ with extremely steep vertical gradients.

A 10-year simulation of the basic properties of the upwelling and downwelling events in the GoF was carried out by Myrberg & Andrejev (2003). The three-dimensional model was validated with the observations of Haapala (1994) off the Hanko Peninsula. It reproduces well the observed abrupt changes in temperature and salinity (10\°C and 0.5 PSU, respectively).

The scale and existence of GoF upwelling events, based on analyses of modelling results, surface temperature measurements and satellite observations, are as follows. The most important upwelling area is the Finnish coast of the GoF, where upwelling occurs as frequently as 30–50% of the

\[\text{Figure 10. Upwelling index (percentage of the time during which upward motion prevails) in the Gulf of Finland based on model results (after Myrberg & Andrejev 2003, by kind permission of O. Andrejev)}\]
The typical extent of a single upwelling event is 10–20 km (sometimes 40 km) offshore and c. 100 km alongshore. Its lifetime spans from several days to several weeks. It results in a drop of up to 10°C in the surface temperature between the upwelled water and the surrounding surface water, and the horizontal gradient may reach 1°C km\(^{-1}\). The vertical velocity can be as much as \(3 \times 10^{-5}\) m s\(^{-1}\) (Myrberg & Andrejev 2003). A wind event with a duration of c. 60 h is usually required for an upwelling event to develop, which corresponds to a wind impulse of the order of 4000–9000 kg m\(^{-1}\) s\(^{-1}\) during the stratified period (Haapala 1994).

The directional structure of winds over the GoF described above shows that sequences of relatively strong and long events of winds from different directions may occur in this area. Consequently, the average properties of the water masses across the gulf may be substantially modified by sequences of upwellings and downwellings of different extent. Evidence from long-term currents, temperature and salinity measurements at single points, and from repeated CTD-castings along transects perpendicular to the coastline of the gulf shows that patches of fresher water frequently occur in the middle of the GoF. The traditional explanation is that these patches reflect the voluminous runoff of the River Neva; however, an equally likely explanation is that more saline water has been formed along the coasts as a result of a sequence of upwellings (Talpsepp 2008).

The technology of remote sensing of many important parameters such as sea surface temperature or ice conditions within the range of visible

**Figure 11.** Sea surface temperature map during a strong upwelling event along the northern coast of Estonia on 7 August 2006. Courtesy of the Finnish Environmental Institute
or infrared light, or in the microwave range, is widely used in many investigations performed in the GoF area. A straightforward application is the detection of the spatial extent of upwellings. Figure 11 depicts the pattern of sea surface temperature that resulted from a long-lasting and spatially extensive upwelling along the southern GoF coast in August 2006. This upwelling was coupled with anticyclonic weather patterns and related E winds. With a longitudinal extension of 360 km and a cross-extent of 25 km, this event affected sea surface properties quite substantially for a month. It became perceptible as a result of a very abrupt drop in the coastal sea surface temperature from c. 21 to 5°C and was coupled with a pronounced longshore coastal jet with a speed of up to 60 cm s⁻¹. The amplitude of sea-surface salinity variations was 3.6 PSU. The subsequent offshore broadening of the upwelling zone and the development of filaments spread cold, more saline and more nutrient-rich water up to the central part of the GoF (Suursaar & Aps 2007).

5. Optics and light conditions

5.1. Light transfer and optically active substances

Solar radiation is one of the key forcing factors in physical oceanography. Distributed in the surface layer, it is the principal heat source for seawater and provides light quanta for photosynthesis. In the latitudes of the GoF, the solar radiation has an annual cycle with a range in the monthly means from 10 W m⁻² in December to 250 W m⁻² in June (FMI 1982 – see the description of the consequences of such variability in Section 3.2). To estimate the radiation level on the basis of solar altitude, atmospheric transparency and humidity, the Zillman (1972) formula has been found to give good results in the GoF (Ehn et al. 2002). The cloudiness correction is somewhat uncertain but several energy balance models have employed the Reed (1977) formula.

Solar radiation is subject to reflection and refraction when reaching the water surface. In the water solar beams are absorbed and scattered by water molecules and other optically active substances (OAS). Marine optics examines the transfer of solar radiation in seawater (so-called direct problems) and the use of optical signals for the identification of the OAS, which are examples of so-called inverse problems. A small part of the radiation is scattered back from the water body to the atmosphere. The task of optical remote sensing is to interpret the backscattered signal for the OAS in the water.

Only visible light and certain narrow parts of near-ultraviolet and near-infrared bands of the solar radiation may be transferred over significant
distances in liquid water. In natural waters the OAS consist of (i) coloured dissolved organic matter (CDOM, also called yellow substance), (ii) suspended matter, and (iii) chlorophyll. The Baltic Sea waters as well as all coastal seas belong (in terms of Jerlov’s optical classification, Jerlov 1976) to the Case III waters, or optically multicomponental water bodies (Arst 2003), where the concentrations of all the OAS are significant.

The CDOM strongly absorbs short wavelengths. Its contribution to the beam absorption coefficient is usually modelled with an exponential shape peaking at shorter wavelengths: 

$$a_y(\lambda) = a_y(\lambda_0) \exp[-\gamma(\lambda - \lambda_0)]$$

where $a_y(\lambda_0)$ is the damping at a reference wavelength $\lambda_0$ (often 380 nm) and $\gamma \approx 0.017 \text{ nm}^{-1}$ is the damping constant, one of the central optical characteristics of natural waters. Its inverse $\gamma^{-1}$ characterises the width of the CDOM absorption band. Chlorophyll $a$ has two absorption bands (430–440 nm and 660–690 nm). The absorption and scattering of particles of suspended matter depend on their size, shape and optical properties, and the resulting wavelength dependency is usually weak.

Optical investigations in natural waters largely address irradiance and its attenuation with depth, which leads to apparent optical properties (Arst 2003). The downwelling planar irradiance $E_d(z;\lambda)$ at a depth $z$ and wavelength $\lambda$ is the total radiation from the Sun arriving from the upper hemisphere onto a horizontal plane, while the upwelling planar irradiance $E_u(z;\lambda)$ accounts for the radiation coming from the lower hemisphere. The scalar irradiance is the total radiance falling on a sphere.

Two central optical characteristics are derived from the listed irradiances: the reflectance $r(z,\lambda) = E_u(z;\lambda)/E_d(z;\lambda)$ and the diffuse attenuation coefficient $K(z,\lambda) = -[E_d(z;\lambda)]^{-1}[dE_d(z;\lambda)/dz]$. The light diffusion coefficient $K_\Omega$ is defined correspondingly using the downwelling irradiance integrated over the whole optical or PAR (Photosynthetically Active Radiation) band $\Omega \approx 380–760 \text{ nm}$. Direct measurements of the inherent optical properties (see e.g. Arst 2003), the absorption and scattering coefficients $a(z,\lambda)$ and $b(z,\lambda)$, respectively, have been made by laboratory spectrometers and specific sounding devices, which use their own light source. Their sum $c(z,\lambda) = a(z,\lambda) + b(z,\lambda)$ is the beam attenuation coefficient.

5.2. Optical characteristics of Gulf of Finland waters

In situ investigations of marine optics in the Baltic Sea have largely concerned the Gotland Sea and southern basins (Dera 1992) (Figure 12); direct measurement data for the Gulf of Finland are quite limited. Optical remote sensing technology, however, has been extended over the whole Baltic Sea. Coastal waters exhibit a similarity of optical properties with
many lakes. For example, Arst (2003) presents results for Finnish and Estonian lakes, with many in her list belonging to the GoF drainage basin. Our knowledge of the optical properties of GoF water masses is based mainly on irradiance data, which are mostly from coastal areas and adjacent basins. For the central GoF only Secchi depth data are available. Also, a limited amount of information on the inherent optical properties is available, determined from water samples and in situ soundings.

The present average Secchi depth is around 5 m in the GoF, varying from 3 to 7 m depending on the time and place (Fleming-Lehtinen et al. 2007). Secchi depths are greater in the central basin than near coastal areas, and smaller during river runoff peaks, when large amounts of suspended matter are brought into the basin. In offshore waters much smaller Secchi depths are observed at times. In practice, the Secchi depth approximately equals twice the e-folding scale of light attenuation and can in

Figure 12. Spectra of light attenuation coefficient $c \text{ [m}^{-1}\text{]}$ in the Gulf of Gdańsk and the Gotland Basin compared with such spectra for general ocean coastal waters and pure water. Adapted from Dera (1992)
this way be linked with physical quantities. A large data base of the Secchi depth in the GoF extending over 100 years is described in Fleming-Lehtinen et al. (2007).

GoF water masses are characterised by high levels of CDOM and suspended matter from nearby land areas, strongly influencing the transfer of light. A large amount of yellow substance in the Baltic Sea turns the colour brown (Dera 1992). Although there are no rigorous data on such a colour change in the GoF, an analogous pattern evidently exists here as well. For such optically multi-componental waters it is a very difficult task to establish the relationship between the concentrations of OAS and the optical properties of the water mass.

Field measurements of several optical properties of GoF waters are available for coastal sites. Sipelgas et al. (2004) performed a field campaign off the north-western coast of Estonia. They obtained a regression formula $a(412 \text{ nm}) = 0.4288C_{\text{CDOM}} + 0.1603$ between the beam absorption coefficient $a$ (measured in m$^{-1}$) and CDOM concentration $C_{\text{CDOM}}$ (measured in mg l$^{-1}$); $a(412 \text{ nm})$ ranged from 0.5 m$^{-1}$ to 2.0 m$^{-1}$ and was thus dominated by $C_{\text{CDOM}}$. These authors also derived the following relationship between the concentration of suspended matter $C_{\text{SM}}$ (mg l$^{-1}$) and the beam scattering coefficient $b$ at 715 nm (m$^{-1}$): $C_{\text{SM}} = 0.538b(715 \text{ nm}) + 2.0683$.

The scattering coefficient was found to (i) be only weakly dependent on the wavelength $\lambda$, according to the power law $b \propto \lambda^n$ with $0 \geq n \geq -0.5$, and (ii) decrease with increasing water turbidity. Sipelgas et al. (2004) combined their results as an expression of the beam attenuation coefficient (in m$^{-1}$) over the optical band $c_{\Omega} - c_w = 0.3C_{\text{SM}} + 0.35C_{\text{chl}} + 0.86C_{\text{CDOM}}$, where $c_w$ is the attenuation coefficient of distilled water and $C_{\text{chl}}$ (mg m$^{-3}$) is the chlorophyll concentration. Although all three optically active substance groups are of significance, Sipelgas et al. (2004) concluded that, in their data, CDOM is the most important component as regards the variability of the attenuation coefficient. In general, it is clear that also chlorophyll has a large seasonal variability, which is reflected in the optical properties of the GoF waters.

As mentioned above, it is mostly near-shore sites that have been examined in terms of marine optics in the GoF. Optical properties of near-coastal waters are, however, frequently strongly site-specific and do not represent the entire GoF water body. Recent study sites in the GoF include Kunda Bay (59°30'N, 26°35'E) in the eastern part of the Estonian coast of the GoF (Arst et al. 1990), Munga Bay (Reinart et al. 2003) and Narva Bay (Kutser et al. 2007). Kunda Bay waters are quite turbid, definitely site specific, and do not represent well the general conditions in the GoF,
whereas the properties of Muuga Bay waters more reflect those of typical GoF waters.

Another study site representing coastal waters is Santala Bay (59°55′N 23°03′E) on the west coast of the Hanko Peninsula on the Finnish coast, where measurement data are available for 15–30 March 2000, during which period the bay was ice-covered. The apparent optical properties were different from those in open-water conditions because the incoming light is much more diffuse under ice cover than in open water conditions, a fact established by the measurements at this site (Leppäranta et al. 2003).

The waters of Santala Bay are oligo-mesotrophic. The diffuse attenuation coefficient over the PAR range is 0.5–0.9 m$^{-1}$ for the planar irradiance and 0.7–1.0 m$^{-1}$ for the scalar irradiance. The planar values would mean a Secchi depth of 2–4 m for these coastal waters. The OAS levels were: suspended matter 1.3–3.9 mg l$^{-1}$, chlorophyll 1.6–1.9 mg m$^{-1}$, and beam attenuation coefficient at 380 nm wavelength 1.2–3.9 m$^{-1}$. The latter value corresponds to 20–30% higher levels of CDOM than those obtained by Sipelgas et al. (2004) off the Estonian northwest coast. The beam attenuation coefficient was 0.6–1.4 m$^{-1}$ in the PAR band (Leppäranta et al. 2003, Arst et al. 2006).

Ice and snow cover have a major influence on light transfer, since the structure of the ice sheet and snow layer change (Granskog et al. 2004). The albedo of snow is high (0.5–0.9) and the optical depth of snow is small (10–15 cm), depending primarily on the wetness of the snow. The albedo of brackish ice is 0.2–0.5, depending on the wetness of the surface and the thickness of the ice, and the transparency of brackish ice is less than that of brackish water (in turbid lakes, the transparency of the ice is often better than that of the liquid water).

Reflectance measurements, with ice surface conditions varying from melting wet ice to refrozen snow patches, are reported by Leppäranta et al. (2003). The PAR-band reflectance ranged in 0.28–0.76, being lowest for ice covered by a thin water layer and highest for snow patches. The fresh snow reflectance was typically 0.8–0.9 and remained constant across the PAR-band, as reported earlier for Santala Bay (Rasmus et al. 2002). The reflectance of deteriorated snow had a weak maximum at 550 nm with the overall level above 0.4. The main optically active impurities in sea ice and brackish ice are gas bubbles and liquid brine pockets (which due to their size scatter all wavelengths of light equally well), and chlorophyll. The optical depth of GoF brackish ice is of the order of 1 m (Arst et al. 2006), and therefore light is transferred into the water body of the GoF when the snow layer is less than 5–10 cm thick.
5.3. Applications and long-term changes

Optical spectral measurements by satellites and airborne systems provide information on the OAS in the surface layer with a thickness roughly equal to the attenuation length. The remotely sensed signal, the surface reflectance, contains the surface reflection, which is a mirror reflection and independent of the OAS, and backscattering from the surface layer, which constitutes information on the OAS. So-called bio-optical models have been tuned to GoF conditions for the inversion of the remote sensing signal (e.g. Zhang 2005, Reinart & Kutser 2006). Figure 13 shows an example of a chlorophyll chart based on interpretation of satellite imagery.

![Chlorophyll chart of the entire Baltic Sea basin for July 2005](image)

**Figure 13.** Chlorophyll \(a\) chart of the entire Baltic Sea basin for July 2005, based on information from the MODIS Terra satellite. Map and graphics by Seppo Kaitala and Henrik Stutz, courtesy of the Finnish Institute of Marine Research.

The optical complexity of the waters, in particular the mixed influence of CDOM and suspended matter, limits the possibilities of identifying the OAS from remote sensing data in the GoF. These methods have therefore
been successful mainly in identifying algae blooms, owing to their strong signal (Zhang 2005, Kutser et al. 2006a,b). Optical methods are promising for identifying in the future different types of water in the basin, including river plumes. An effort was made by Sipelgas et al. (2006) to examine suspended particle transport in Pakri Bay (west of Tallinn) using MODIS images together with a hydrodynamic model; MODIS has been quite widely used in this region. Nonetheless, Reinart & Kutser (2006) concluded that MERIS leads to results of the same quality.

The mapping of harmful algal blooms is presently one of the most important applications of remote sensing in marine optics, especially in the Baltic Sea. For such purposes, the optical properties of cyanobacterial species have been examined in the laboratory (Kutser et al. 2006a). With regard to the evaluation of the background to long-term changes, however, it is not clear how the signal of climatic changes can be separated from the effect of local anthropogenic loads.

Chlorophyll absorption peaks can be utilised to estimate primary production in surface waters. In particular, in connection with algal blooms, the signal is so strong that it can be detected even from panchromatic images. For more detailed investigations and quantification, the potential of satellite remote sensing to detect cyanobacterial species has been examined by modelling (Kutser et al. 2006a,b). The 620–670 nm band of the MODIS satellite was found to be suitable for detecting suspended matter when dredging blooms were being monitored (Kutser et al. 2007).

Optical methods have been used to identify and quantify the influence of wakes from fast ferries on bottom sediments and the properties of water masses (Erm & Soomere 2004, 2006). Such an influence was observed at water depths in the 2–15 m range in a near-bottom layer with a thickness of c. 1 m. It usually lasted 6–15 minutes. Rough estimates suggest that c. 10,000 kg of sediments per metre of the affected sections of the coastline may be brought into motion by ship wakes annually.

Secchi depth measurements in the Baltic Sea date back more than one hundred years, which provides an insight into the changes in water quality (Sandén & Håkansson 1996, Fleming-Lehtinen et al. 2007). The mean Secchi depth has experienced an almost twofold decrease from 8 to 5 m in the GoF (Figure 14) over the last 100 years, which can be put down to anthropogenic influences. This change is obviously due to eutrophication, which has led to elevated levels of CDOM and chlorophyll, both in the Baltic Sea and in the lakes of the drainage basin.
6. Sea level

6.1. Spatial patterns and temporal variations

The basic components of sea level variations in the Baltic (incl. the GoF) were established well before the 1970s. The following factors substantially affect sea level at different time-scales: (i) changes in the rate of relative land uplift (ii) the wind regime, (iii) air pressure patterns and (iv) density of the sea water; factors (ii-iv) are partially interdependent (Alenius et al. 1998). Sea level variation due to tidal forces usually does not exceed a few cm in the GoF.

The northern coast of the GoF boasts three long sea-level time series: at Hanko since 1887, at Helsinki since 1904, and at Hamina since 1928; the readings at Turku (since 1922) represent to some extent the situation at the entrance of the GoF. Further evidence has become available in the last ten years regarding the changes mentioned in Alenius et al. (1998), namely, that the relative magnitude of land uplift with respect to the water level along the northern coast of the GoF apparently changed around 1960 (Johansson et al. 2001, 2003, 2004). For example, the relative uplift at Hanko was c. 3.3 mm y$^{-1}$ until 1960 and has been c. 1.6 mm y$^{-1}$ since
that year (Figure 15). Following this change, the first regression line for ‘fitted mean sea level’ (FMSL), used in many applications in Finland as the ‘true’ sea level, was fitted to the annual mean sea levels from the beginning of measurements up to 1960; the second line, fitted to sea levels since that date, had a clearly different slope. The rather uniform sea level rise along the southern coast of the GoF 2.1–2.8 cm y\(^{-1}\) in 1978–82 (Raudsepp et al. 1999) may have been due to the non-representative choice of the years in question within an increasing phase of interannual variation (see Figure 15).

![Figure 15. Annual mean (solid line and circles) and fitted mean sea level (dashed line) at Hanko. Based on Johansson et al. (2004). The data for 2003–06 were kindly provided by Milla Johansson (Finnish Institute of Marine Research)](image)

Long-period sea level variations in the GoF follow the fluctuations of the water surface of the entire Baltic Sea; a large proportion of these variations is imported from the North Sea (Vermeer et al. 1988). While the strongly anisotropic wind regime may contribute to long-term sea level changes by creating an eastward rising surface slope, it is definitely one of the main factors responsible for local sea level variations, e.g. by pumping water into and out of the Baltic Sea (Vermeer et al. 1988).

A remarkable aperiodic interannual sea level variability over 2 to 6 years was documented during the previous decade (Sjöberg & Fan 1986, Vermeer et al. 1988). This is best seen in the long sea level series at Stockholm (Ekman 1988), where the variability is 10–20 cm. An analysis based on a box model (each sub-basin constituting one box; one sea level series per
box accounted for) suggests that this variability is forced mostly by sea level variations in the Kattegat and by changes in the freshwater supply to the Baltic Sea (Samuelsson & Stigebrandt 1996). This argumentation, although inferred from an oversimplified model setup, is supported by the statistical analysis of the seasonal cycle of the winds and the sea level in the Baltic Sea (Lass & Matthäus 1996).

The most dominant component of the sea level variation is the annual cycle of the monthly mean water level, c. 50–80% of which is imported into the Baltic Sea from the North Sea (Samuelsson & Stigebrandt 1996). Its half-range (calculated from de-trended data over 20-year periods) was c. 10 cm at Helsinki and at Hamina during the first half of the 20th century, after which it increased rapidly to reach c. 17–18 cm in the 1970s–1980s, and decreased again to c. 13 cm at the turn of the century (Johansson et al. 2001). This is in accordance with the statistically significant increase in the amplitude of the annual variation of the whole Baltic Sea level for the period 1825–1984 (Ekman & Stigebrandt 1990). These changes only partially match the long-term changes in the wind regime described in Alexandersson et al. (1998).

The average range of the sea level variation on the southern coast of the GoF seems to be somewhat smaller (from 17 to 23 cm in 1978–82), with high water in October and low water in April (Raudsepp et al. 1999).

There are no significant peaks in time scales from hours up to several months of the water level variations in the Baltic Sea except for its eigen-oscillations (c. 0.5–1 cycles day\(^{-1}\)). The spectrum of the daily mean values is also rather white (Vermeer et al. 1988).

The mid-range Baltic Sea level variability, with periods between 2 days and several years, has a significant spatial coherence (Samuelsson & Stigebrandt 1996). In particular, all sea level variations from periods of 6 months up to 10 years (incl. the ‘pole tide’ with a period of 14.3 months) have an almost identical geographical pattern and, most probably, have a common origin (Ekman 1996). Yet it is debatable whether the whole Baltic Sea level acts as a quarter-wavelength (with the node at the Danish Straits and the largest water level variations in the inner parts of the GoF and the Gulf of Bothnia) or a half-wavelength oscillator (Figure 16) for these periods (Samuelsson & Stigebrandt 1996, Ekman 1996). The quarter-wavelength regime seems to be universal for shorter periods, from a few days up to one month, whereas for periods longer than 1 month and up to 9 years the sea level regime may resemble a half-wavelength oscillator (Samuelsson & Stigebrandt 1996). In the GoF the appearance of both regimes is the same.
Horizontal long-term variations in the Baltic Sea level are caused mostly by variations in the local wind climate (Andersson 2002) and by the basin-scale variation in density (Ekman & Mäkinen 1996) that usually follows salinity changes. The latter is driven mostly by the circulation and mixing of water masses, and by river inflow, which has a large annual variability (Bergström & Carlsson 1994; see also Laanearu & Lips 2003). The sea level in the large river mouths, e.g. the River Narva, is frequently higher than in the rest of the sea (Raudsepp et al. 1999); yet typically, the difference is caused by the particular location of the tide gauge landwards of a relatively high sand bar (Laanearu et al. 2007), and there is no great difference in the correlations between the stations located in large river mouths and at other sites (Raudsepp et al. 1999).

The temporal and spatial fluctuations of the water level along the southern coast of the GoF from Narva to Dirhami are well correlated. The correlation is somewhat higher than, for example, in the Gulf of Riga, but is slightly smaller than in Väinameri (Moonsund). The spatial correlation radius is 200–400 nautical miles, that is, far beyond the cluster of measurement sites. The temporal correlation radius (defined as the e-folding scale) is c. 10 days. Particularly synchronised behaviour is found in the western part of the GoF. The water level is relatively easily predictable in the entire gulf: the dominant annual harmonic describes c. 40–45% of the total variability in the monthly mean sea level values, and simple stochastic models correctly reproduce > 90% of the sea level variability along the Estonian coast (Raudsepp et al. 1999).

6.2. Trends and extremes

A statistically significant increase (at the 99% level) in the entire Baltic Sea level in December–January and an equivalent decrease in February–March at Stockholm (Ekman 1998) was poorly represented in the GoF except for the decrease in February–March at Hanko (significance 97%, Johansson et al. 2001). There exist several highly interesting albeit statistically non-significant trends. The monthly means during April–October have decreased and those in November–March have increased, but
neither trend is significant. In terms of monthly means the seasonal sea level maximum shifted from the end of September to a point two weeks later before the 1980s, and occurred a full month later by the end of the century (Johansson et al. 2001; see also Raudsepp et al. 1999).

The annual standard deviation of the sea level in the GoF correlates with the NAO-index at the 97–99% level of significance. It has an apparently increasing trend with a small minimum in the 1950s and the most prominent increase in the 1960s–1970s. The largest values of this trend occur, as expected, in the innermost part of the GoF (Johansson et al. 2001). The probability distributions of the sea levels over a long time period were asymmetric (high sea levels being more probable than low levels as a common feature at all the sites along the Finnish coast) and generally became broader (in particular, the probabilities for extremely high sea levels have increased) during the 20th century. Variability is the most pronounced during the winter period (November to January), but is much less obvious in the summer (May to July). The probability distributions have different shapes during different seasons and in different sub-basins of the Baltic Sea (Johansson et al. 2001). The origin of this feature is apparently the highly anisotropic wind regime in the area.

The annual maximum values of the sea level have increased significantly on Finnish coasts during the last 70 years (at the 99% level), mostly during the latter half of the 20th century (Johansson et al. 2001). The trend is essentially asymmetric: the annual minimum values do not show any significant changes. The increase is the most pronounced in the Baltic Sea nodal area, where maxima have increased by c. 10 cm in half a century. This feature indicates that the overall variability in the water level in

Table 3. Water levels during windstorm Gudrun (8–9 January 2005) and historical maxima at the measurement sites along the coast of the Gulf of Finland (Suursaar et al. 2006, Johansson et al. 2001; data from the FIMR, www.fimr.fi). The sites where new records were established in 2005 are indicated in bold.

<table>
<thead>
<tr>
<th>Location</th>
<th>Maxima on 09 January 2005</th>
<th>Highest prior to 2005</th>
<th>Measurements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dirhami</td>
<td>134</td>
<td>148</td>
<td>18 October 1967</td>
</tr>
<tr>
<td>Tallinn</td>
<td>152</td>
<td>135</td>
<td>15 November 2001</td>
</tr>
<tr>
<td>Kunda</td>
<td>139</td>
<td>157</td>
<td>06 January 1975</td>
</tr>
<tr>
<td>Narva-Jõesuu</td>
<td>194</td>
<td>202</td>
<td>23 September 1924</td>
</tr>
<tr>
<td>Turku</td>
<td>130</td>
<td>127</td>
<td>09 January 1975</td>
</tr>
<tr>
<td>Hanko</td>
<td>132</td>
<td>123</td>
<td>09 January 1975</td>
</tr>
<tr>
<td>Helsinki</td>
<td>151</td>
<td>136</td>
<td>27 Jan 1990</td>
</tr>
<tr>
<td>Hamina</td>
<td>197</td>
<td>166</td>
<td>07 December 1986</td>
</tr>
</tbody>
</table>
the Baltic Sea has increased even more than its local variations. One can thus conjecture that large-scale meteorological and hydrological phenomena rather than local storms have caused these changes.

Extremely high water levels occurred in the GoF in January 2005 during windstorm Gudrun (Table 3). As a result of strong cyclonic activity, the Baltic Sea level was very high (+70 cm) already before the storm (Suursaar et al. 2006). New sea level records were established at Tallinn and at Toila (whereas at the other sites along the southern coast the water level was close to the historical maxima), and at all four stations in Finland close to and in the GoF (Suursaar et al. 2006).

The sea level reached a relatively modest value of 230 cm in St. Petersburg, where the historical maximum (+421 cm) was recorded in 1824. A thorough overview of the climatology of water levels in the eastern part of the GoF and the methods for forecasting exceptional coastal inundations in this area is presented in Antonov (2003). Klevanny (2004) describes an alternative way of calculating extreme water levels in the eastern Gulf of Finland.

6.3. Implications of potential sea level changes

The mean sea level in the Gulf of Finland is basically controlled by the sea level in the World Ocean. Current scenarios for the coming 100 years given by the Intergovernmental Panel on Climate Change (IPCC 2007) predict an increase in global mean sea level in the range of 0.3–0.9 m, mostly due to thermal sea water expansion and the potential gradual melting of the overland cryosphere. Climate change is likely to affect the factors controlling the difference in water level between the Baltic Sea and the North Atlantic, and consequently, also the salt water inflow conditions into the Baltic Sea.

The potential effects of global climate change on the sea level on the Finnish coast may be relatively large (Johansson et al. 2004). The project FIGARE/FINSKEN developed scenarios up to the year 2100 for four key environmental attributes (climate, sea level, surface ozone, and sulphur and nitrogen deposition) as well as for future socio-economic developments in Finland (Carter et al. 2004). Although the uncertainties in the scenarios are large, in most cases the rise in water level is expected to balance the land uplift along the northern coast of the GoF, and the past declining trend of the relative sea level is not expected to continue.

The southern and eastern coasts of the GoF may experience a relative water level rise. A relative water level rise of 1 m, which is larger than the maximum increase by the year 2100 according to the scenarios of the IPCC (2007), would cause serious consequences in certain areas (Kont et al. 2003). Land loss in the Tallinn area would probably be insignificant,
but in the easternmost part of Estonia, an important recreational area adjacent to Narva-Jõesuu with excellent sandy beaches would be under severe pressure. Another site of great risk is the Sillamäe dumping site of the former uranium enrichment plant. Although it is currently protected by new harbour constructions and their access road, possible damage to this site poses one of the greatest threats to the GoF environment.

Orviku et al. (2003) suggest that the seemingly increasing storminess (expressed as a statistically significant increasing trend of the number of storm days over the last half-century) in the eastern Baltic Sea has already caused extensive erosion and alteration of depositional coasts, such as sandy beaches. They express the opinion that the destruction of beaches due to the more frequent occurrence of high water levels and intense waves, as well as the lengthening of the ice-free period, have overridden the stable development of several sections of Baltic Sea coasts.

7. Surface waves

7.1. Wind waves

The wave regime of the GoF depends to some extent on the wave conditions in the northern Baltic Proper. Overviews of wave studies performed for different parts of this basin are presented in Soomere (2005a, 2008) and Broman et al. (2006). The overall pattern of wave activity, simulated with the use of the second-generation HYPAS spectral wave model (Jönsson et al. 2002, 2005), follows the anisotropy of the wind regime in the Baltic Proper (Soomere & Keevallik 2001). The regions with the largest wave activity are found along the eastern coasts of the Baltic Proper, in particular, in the entrance area to the GoF. The qualitative features of these studies match well the results of recent wave measurements in the northern Baltic Proper (Kahma et al. 2003). Their quantitative features, however, differ from the factual ones, because of the ice cover in some areas of the northern Baltic Proper and the GoF during the season that statistically contains the highest wind speeds.

Although earlier studies of the possible modification of the wave climate caused by the apparent increase in storminess (WASA Group 1995) identified no substantial changes in the wave regime, dramatic variations of the annual mean wave height in the northern Baltic Proper have been detected from long-term instrumental measurements (Broman et al. 2006) and visual observations (Soomere & Zaitseva 2007).

The basic advances in surface wave studies in the GoF area during the last decade are based on (i) the implementation of modern spectral wave models (e.g. Komen et al. 1994) that adequately represent the sea state even
in such a challenging place for wave modellers as the northern Baltic Sea (e.g. Tuomi et al. 1999) and the Gulf of Finland (Soomere 2005a, Soomere et al. 2008), (ii) a radical improvement in the quality of atmospheric models accompanied by the availability of high-quality marine wind data from Kalbadagrund and from Bogskär in the northern Baltic Proper, and (iii) the directional wave data from the central part of the GoF (Pettersson 2001). Also, a new wave atlas has recently been published, nearly 40 years after the previous one (Lopatukhin et al. (eds.) 2006a).

Wave measurements in the GoF were performed during c. 10 months in different seasons at the beginning of the 1990s and have been carried out since November 2001 during the ice-free period (Pettersson 2001). For the coastal areas of Estonia there are only visual observations from the coast and short-term measurements made using pressure-based sensors (Soomere 2005a, Soomere et al. 2008).

Directional wave statistics are available for two sites (off Helsinki, 59°57.9’N, 25°14.1’E, 4.5 months in 1990–92, and off Porkkala, 59°44.5’N, 24°18.5’E, 6.5 months in 1994). The significant wave height of 4.0 m (November 1991) was the highest in this area until 2001. The maximum significant wave height $H_s$ occurring once in 100 years in the GoF is estimated to be 3.8 m and the corresponding single wave height 7.1 m (Alenius et al. 1998, Pettersson 2001).

The influence of waves originating from the Baltic Proper becomes evident in some cases when S or W winds prevail. Turning wind conditions frequently cause bi-modal wave patterns (Pettersson 2001). The average wave directions are often concentrated in narrow sectors along the gulf axis, although the wind directions are more evenly spread (Kahma & Pettersson 1994, Alenius et al. 1998, Pettersson 2004). Although meandering of the geostrophic flow (Savijärvi et al. 2005) may contribute to this feature to some extent, it is usually attached to the slanting fetch conditions in which the wind direction is oblique to the gulf axis and the different components of the (nearly) saturated wave fields often propagate in different directions. Shorter waves are usually aligned with the wind, while somewhat longer and higher waves propagate along the gulf axis.

On 15 November 2001, $H_s = 5.2$ was recorded off Helsinki (Pettersson & Boman 2002). Wave heights exceeding 4 m apparently occurred along a large part of the Estonian coast during this storm (Soomere 2005a). Unusually long and high waves, with peak periods up to 12 s, occurred in the central area of the GoF on 9 January 2005 (Figure 17). The wave height reached 4.5 m near Naissaar and was close to 4 m near Helsinki (Soomere et al. 2008).
As the regular wave measurement network in the GoF consists of one waverider, both scientific and operational wave models provide important information about instantaneous and long-term wave properties. The spatial resolution of operational models is 5–9 nautical miles (c. 9–15 km) in the Baltic Sea (Soomere et al. 2004). The scientific models use resolutions of a few miles (e.g. Soomere 2003). In order to adequately account for the basic geometry of the GoF, and to resolve the effects occurring in the slanted fetch conditions, the grid step should not exceed 1–2 miles (Pettersson 2004, Soomere 2005a). It is predicted that such models will be used in the operational regime based on different nesting and coupling schemes (Soomere et al. 2004).

A resolution of the order of a few hundred metres is frequently necessary to resolve the effects of complex geometry and bathymetry in local applications (Soomere 2005a, Laanearu et al. 2007). The basic features of wave fields calculated for Tallinn Bay with a resolution of 1/4 mile were reliable to a depth of c. 5 m and as close to the coast as 200–300 m (Soomere 2005a). The reason for such a favourable feature, which is generally not true on open ocean coasts, is the shortness of predominating waves in the GoF.

The fetch length in most storms is relatively short in the GoF and the changes in wind properties are rapidly reflected in the wave field properties. The wave fields in sub-basins of the GoF such as Tallinn Bay or Narva Bay in many cases mimic the changes in the local wind properties. This feature allows the local wave climate to be estimated with a greatly simplified dependence of the wave field on the one-point marine wind, which still
adequately represents wave conditions in more than 99.5% of cases (Soomere 2005a). The numerically simulated wave regime based on Kalbadagrud wind data matches well the data set collected in 1974–80 (Orlenko et al. (eds.) 1984) for the relatively calm spring and summer seasons. Wind waves have moderate dominating periods (normally < 5–6 s) in Tallinn Bay. The significant wave height exceeds 0.5–0.75 m with a probability of 10% and 1.0–1.5 m with a probability of 1% (Soomere 2005a).

Significant seasonal and spatial variability in the wave properties occurs in semi-sheltered bays of the GoF such as Tallinn Bay. This is an expected feature and reflects an analogous variability in the GoF (Alenius et al. 1998). Some areas of Tallinn Bay are, however, unexpectedly calm, mainly because the mainland and the surrounding islands together with numerous shallow areas shelter the bay from waves coming from the typical strong wind directions (Soomere 2005a).

7.2. Anthropogenic waves

The intense fast ferry traffic between Tallinn and Helsinki has created concerns about its potential influence on the marine environment. High-speed ships cross the gulf nearly 70 times daily during the high season. They frequently sail in the transcritical velocity range (i.e. within ±15% of the maximum phase speed $\sqrt{\frac{gH}{h}}$ of surface waves at finite depths). The wave systems excited at these speeds may contain solitonic waves, which are fundamentally different from the waves produced by conventional ships (Soomere 2005b, 2006, 2007).

The damaging potential of shipping in the vicinity of fairways is widely recognised. Ship wakes can cause extensive shoreline erosion, resuspension of bottom sediments, trigger ecological disturbance, and harm the aquatic wildlife (Schoellhamer 1996, Bourne 2000). Enhancement of vertical mixing along the fairway may intensify eutrophication effects and influence harmful algae blooms by causing the transport of nutrients from sediments into the euphotic layer (Lindholm et al. 2001). These aspects have been extensively studied for the archipelagos adjacent to the GoF (e.g. Rönberg 1975, 1981, Rönberg et al. 1991, Madekivi 1993).

The share of ship waves in the total wave activity may be remarkably high in Tallinn Bay, although many of its coasts are open to high wind waves (Soomere et al. 2003a,b, Soomere 2005b). The properties of ship waves were measured in different coastal areas with a depth of 5–7 m and located at distances of 2–10 km from the fairway (Soomere & Rannat 2003) at sites that were apparently the most vulnerable with respect to the ship waves (Kask et al. 2003).
Waves from conventional passenger and cargo ships, and from hydrofoils have typical heights of 20–30 cm and periods of 3–4.5 s, and are almost indistinguishable from the natural background (Soomere & Rannat 2003).

The leading wake waves from large high-speed car-carrying catamarans and monohull vessels (called fast ferries below) frequently have heights of c. 1 m and periods of 10–15 s (Figure 18); such waves occur extremely seldom under natural conditions in Tallinn Bay. The wakes typically consist of two or three groups with periods of 15–9 s (the longer waves arriving first), 9–7 s, and c. 3 s, respectively. The highest waves are usually the leading ones; yet the third group may contain the highest waves in remote areas (c. 70 cm as far as 8–10 km from the ship lane; Soomere & Rannat 2003). The intensity of the ship wave field appears to be highly variable in different parts of Tallinn Bay (Torsvik & Soomere 2008).

Owing to the very limited numbers of wind wave measurements in Tallinn Bay (see above), comparisons of the impact of waves of different origin rely on the modelled wind wave climate (Soomere 2005a). The presence of sea ice (which statistically takes place mostly during the windy winter season; Mietus 1998) is ignored, so the parameters of wind waves are somewhat overestimated. Although very high waves occasionally occur in Tallinn Bay, the daily highest ship waves mostly belong to the annual highest 1–5% of wind waves. The annual mean energy of ship-generated waves in the coastal area of this bay forms c. 5–8% of the total wave energy (6–12% during the spring and summer seasons) and c. 18–35% (27–54% during the...
Leading waves from fast ferries may cause unusually high near-bottom velocities at depths of 10–30 m (Soomere & Kask 2003). The near-bottom orbital velocity created by a ship wave with a height of 1 m and a period of 12 s is c. 45 cm s\(^{-1}\) in a 10 m deep sea area and c. 20 cm s\(^{-1}\) in a 30 m deep area. These velocities considerably exceed the typical speeds of local near-bottom coastal currents. The fast ferry traffic is thus a qualitatively new forcing component of vital impact on the marine ecosystem in Tallinn Bay as well in other non-tidal or micro-tidal semi-sheltered areas (Soomere 2005b, 2006).

A large proportion of the waves from fast ferries in a coastal area with a depth < 10 m matches the shape of the periodic solutions of the Korteweg-de Vries (KdV) equation called cnoidal waves; for this reason, the ship-wave-induced near-bottom orbital velocities are apparently very much larger (by a few tens of % at a depth of c. 10 m) than expected from linear wave theory (Soomere et al. 2005). The shape of the highest ship waves practically coincides with that of a KdV solitary wave. These waves are not necessarily solitons, because it is not clear whether they preserve their shape in time and in collisions with similar entities.

The primary reaction of the seabed to large near-bottom velocities consists in an intensive (re)suspension of the benthic layer and bottom sediments (PIANC 2003) which, in Tallinn Bay conditions, may be accompanied by an enhancement of sediment transport through a multi-frequency wave system, where waves of different origin transport sediments at different water depths (Soomere & Kask 2003). This reaction can be identified and roughly estimated by optical measurements (Bauer et al. 2002, Erm & Soomere 2004). The clearly detectable influence of ship wakes extends to a depth of at least 15 m and usually lasts some 6–15 minutes (Figure 19). It is limited to a water layer with a thickness of c. 1 m near the seabed. The largest changes in the optical properties of the water masses occur at relatively small depths of 2.1–4.8 m. Since most of the fast ferry sailings take place in the biologically active season and during the daytime, an overall deterioration in the underwater light conditions due to this traffic is likely. About 10 000 kg of sediments per metre of the affected sections of the coastline is brought into motion by ship wakes annually. The total loss of sediments could be several hundred litres per metre of coastline (Érm & Soomere 2004, 2006).

Wave groups excited by fast ferries are often practically non-dispersive entities that may remain compact at large distances (Soomere 2006). The remote influence of wakes is thought to be responsible for the abrupt thermal
Figure 19. The surface and underwater light field and the diffusion attenuation coefficient $K_d$ near the Island of Aegna on 20 July 2003 in a sea area with a depth of 4.8 m. Wakes from fast ferries arrived 2–3 minutes before the maximum values of $K_d$ were reached. The instantaneous values of the measured parameters (averaged over 5 s) and the moving average over 5 subsequent measurements are shown. The plots of irradiance $E_0(z_l)$, downwelling irradiance at the upper sensor $E_d(z_u)$ and at the lower sensor $E_d(z_l)$ were normalised using the downwelling irradiance values $E_d(+0)$ measured by a plane sensor above the sea surface (from Erm & Soomere 2006, by kind permission of A. Erm).

changes in shallow inlets several kilometres away from the fairway (Lindholm 1997) and even on the open sea (Fagerholm et al. 1991). It has been suggested that fast ship traffic should be treated as a source of energy pollution in the form of unusual waves (Soomere 2005b).

Since both the length and the height of waves generally increase with ship speed (PIANC 2003, Soomere 2007, Torsvik & Soomere 2008), the environmental impact of fast ferry traffic can be alleviated if the speeds of certain types of ships are reduced. Starting from 1997, several countries and communities have imposed relevant regulations primarily based on wave height criteria in the nearshore environment (Parnell & Kofoed-Hansen 2001, PIANC 2003). Restrictions for Tallinn Bay have been under discussion since 2001 (Soomere et al. 2003a,b).

8. Ice structure, growth, decay and ice dynamics

Sea ice has a remarkable influence on the oceanography of the GoF. Ice formation and melting influence the stratification of the waters; particularly
important is the freshening of the surface layer in spring. Even more
dramatic is the nonlinear dynamic behaviour of drift ice. Thick ice reduces
the transfer of momentum from the wind to the water body. The circulation
becomes weak as it is forced only by boundary fluxes at the ice edge
(Leppäraanta 2005). The specific features of circulation in the subsurface
layer (Andrejev et al. 2004a) may thus dominate in quite a thick upper
layer of the sea, as hypothesised in Soomere & Quak (2007).

The presence of ice also plays a significant role in ecology and in
environmental protection. Various substances derived from precipitation are
accumulated in the ice sheet during the winter season. Moreover, impurities
are captured from the water body and the sea bottom during ice formation
(Leppäraanta et al. 1998). These substances are transported by the ice and
released to the sea water as the ice melts, and the surface layer may become
enriched in pollutants. This transport also affects sedimentation at the
sea bottom. Other geological ice effects are bottom scouring and shore
modification. The brackish ice of the Baltic serves as a growth habitat for
ice algae, as is the case with normal sea ice. Ice cover is also an important
part of the living environment of Baltic seals.

For humans, the ice cover has for long provided possibilities of over-ice
traffic, fishing and seal hunting. On the other hand, the ice is a barrier
to shipping, which was forced to come to a halt in winter before the
development of icebreakers in the late 1800s. The forces generated by ice
have been of concern to ships and offshore construction activities: ice
pressure has destroyed several ships in the past, and in the Gulf of Finland,
a lighthouse was damaged in this way in the 1960s.

8.1. General ice conditions

The ice cover of the Baltic brackish basin is essentially a sea ice cover
(Leppäraanta & Myrberg 2008). This means that the crystal structure is
similar to that of sea ice rather than that of freshwater ice; in addition, the
ice is saline, and at larger scales the ice cover appears as pack ice or drift ice
consisting of leads, undeformed spots, and broken ice fields with slush belts,
rafted ice, ridges etc. (Figure 20). The GoF is one of the three Baltic Sea
basins – in addition to the Gulf of Bothnia and the Gulf of Riga – where
significant amounts of ice form every winter.

On average, sea ice is present in the GoF for five months each winter,
usually from December to April. The average freezing date is 1 December
in Neva Bay, and the last drift ice floes, observed off Vyborg, have typically
melted by 1 May (SMHI & FIMR 1982). The range of the freezing and break-
up dates is 15 November–15 January and 15 April–15 May, respectively.
At the beginning of winter the freezing front progresses from Neva Bay westwards along the northern coast. Consequently, the ice season is more severe towards the east and the north of the gulf. Normally, the GoF becomes fully ice covered, but in mild winters only the part east of the Narva–Kotka line freezes. The main part of the basin freezes during January and the ice melts in April, giving an average length of three months for the period of major ice coverage. The ice reaches its maximum extent in February or March. The maximum annual thickness of coastal fast ice varies from 30 to 80 cm in Vyborg and Neva Bays.
The heat inflow due to the easterly coastal current from the Baltic Proper and the dominance of S to SW winds helps to keep the Estonian coastal area ice-free for some time. The asymmetry of the ice conditions between the southern and northern coasts is also influenced by the coastal morphology, which supports a broad, landfast ice zone along the coast of Finland but leaves almost no fast ice on the Estonian coast. Outside the landfast ice zone there is drift ice with a highly non-linear mechanical behaviour. Ice thinner than 20 cm normally drifts well with the wind and currents. S and N winds frequently open wide leads on the lee side. E winds push the ice out of the GoF, while W winds pack the ice toward the eastern end. In the presence of compressive drift, pressure ridges form.

The landfast ice zone is found in coastal areas where the ice is stationary for most of the ice season. Its width depends on the bottom topography: islands and grounded sea ice ridges serve as support points to stabilise the ice sheet. The size of the ridges is such that the edge of the landfast ice in the GoF is in the neighbourhood of the 10 m isobath (Leppäranta 1981).

The basic features in a drift ice ‘landscape’ are ice floes with open water between them. Floes consist of patches of smooth undeformed ice separated by deformed ice. A floe may be from 10 m to 10 km in size. The thickness of the ice varies because of differences in the age of the ice and the different mechanisms by which various ice types are formed. Open water formations are formed mechanically and are categorised as leads. The polynyas typical of some regions in polar seas do not exist in the Baltic Sea.

A good description of the ice conditions is provided by the operational ice charts, published daily for the Baltic Sea. The ice information is presented in accordance with international standards, i.e. based on surface reports from ships and coastal stations, and satellite remote sensing (mainly NOAA and satellite SARs). The traditions of ice reporting and observational techniques are such that ice charts tend to describe ice cover as it appears to an observer from an aircraft or satellite.

8.2. Structure and properties of ice

The brackish ice of the Baltic has a fine-scale structure similar to that of sea ice (Palosuo 1961, Kawamura et al. 2001). The ice crystals have irregular boundaries and there is significant brine entrainment, the salinity of the ice being 0.5–2 PSU in winter. According to observations in the GoF, the freshwater ice type occurs only in estuaries, where the salinity of the surface water is < c. 1.5 PSU (Kawamura et al. 2002).

Ice formation begins in slightly supercooled water. In calm, ‘quiet conditions’ small ice spheres grow into thin discs and dendritic forms and finally join into a solid ice sheet. The ice grows horizontally on the surface
with vertical crystal \( c \)-axes; the horizontal crystal size ranges from 1 mm up to 10 cm but the vertical extent is less than 1 mm. With increasing intensity of the thermal forcing or other disturbances of the ‘quietness’, the orientation of the crystal axes is more random and the crystal size is smaller. The resulting primary ice layer is very thin. In turbulent conditions small frazil ice crystals form in the surface layer. At first, they move freely in the surface layer, but then build up a solid sheet when the buoyancy of the ice crystals overcomes the turbulence. The thickness of this solid ice layer is then a few centimetres, and the crystals are small, less than 1 mm, with random \( c \)-axis orientation. Quiet conditions are typical for landfast ice formation, but in the drift ice zone both types are found.

After the formation of the primary ice, the ice may grow down from the ice bottom as congelation ice. Through a 5–20 cm thick transition zone the congelation ice crystals become vertical columns, the \( c \)-axes realign themselves horizontally, and the size of the columns increases with depth. The congelation ice is the dominant type in the fast ice zone (Kawamura et al. 2001). Frazil ice crystals may be produced in leads or at the ice edge and later adhere to the bottom of an ice sheet formed earlier; then the ice sheet will consist of alternating columnar and frazil ice layers. It is not well known how much frazil ice there really is in the GoF ice sheet, probably much less than there is congelation ice. On the top of the primary ice, superimposed ice forms from slush or a mixture of snow and water (due to flooding of the ice, melting of snow or liquid precipitation) or water on ice. Frozen slush, also called snow-ice, is a common feature; its crystals are as small as frazil ice crystals. The proportion of snow-ice in the GoF is typically 10–50\% (Palosuo 1963, Granskog et al. 2004). In an ice sample, the snow-ice layer is normally taken to be the opaque topmost layer, since the large volume of gas bubbles in snow-ice scatters light strongly.

In drift ice fields greater variety in the crystal structure results from the interplay between the mechanical and thermal evolution of the ice conditions. Lead opening gives way to frazil ice formation. Because of rafting, which occurs mainly in thin ice, duplicated or even multiple layering at times results in an intricate stratification of the ice sheet crystal structure. The ice cover may break into floes and small blocks, which then freeze together into a continuous ice sheet with the ice pieces randomly oriented, and then the three structural types are found within the individual blocks. Areally averaged, the mechanically deformed ice accounts for 10–50\% of the ice volume (SMHI & FIMR 1982, Kankaanpää 1997).

The salinity of new ice is 20–40\% of the salinity of the parent water or 1–2 PSU in the GoF (Palosuo 1963, Granskog et al. 2004). With time the salinity decreases, mainly due to gravity drainage. When the ice warms up,
the brine pocket system expands into a drainage network, out of which the brine flows under gravity. In spring the salinity drops almost to zero; recorded residual levels have been < 0.1 PSU. The vertical salinity profile shows a rather homogeneous shape with a weak maximum inside the ice (Figure 21), typical of warm ice. With the increased flooding of snow-ice and the loss of brine, the remaining slush may become very salty.

**Figure 21.** Evolution of salinity in western landfast ice in the Gulf of Finland. From Granskog et al. (2004), printed by permission of the American Geophysical Union
Light transmission through ice and the light conditions underneath ice have been investigated in Santala Bay (Arst et al. 2006). For bare and snow-covered ice the PAR band albedo was 0.28–0.76 and the light transmittance was 1–52%. The light field below the ice was much more diffuse than in open water conditions. The euphotic depth was 0.1–5.5 m. Brackish ice was found to be much less transparent than lake ice because of the brine pockets with their algae population.

8.3. Ice growth, decay and dynamics

Undeformed thermally grown sea ice in the GoF can be up to 90 cm thick; this is about the same as in nearby lakes. The thinnest ice is found close to the ice edge. Ice thinner than 5 cm breaks easily and is not a stable feature. Table 4 shows statistics for Utö at the western edge of the GoF and for Kotka on the eastern Finnish coast. The latter freezes every winter while near Utö the probability of ice formation is 80%.

Table 4. Statistics of the annual maximum ice thickness 1961–1990 (Seinä & Peltola 1991). Ice-free cases are included with zero ice thicknesses

<table>
<thead>
<tr>
<th>Station</th>
<th>Total ice thickness [cm]</th>
<th>Snow ice thickness [cm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>min average max std. dev.</td>
<td>no. of winters</td>
</tr>
<tr>
<td>Utö</td>
<td>0 21 69 22 19</td>
<td>8 50 17 10</td>
</tr>
<tr>
<td>Kotka</td>
<td>16 50 90 18 26</td>
<td>4 18 35 9 19</td>
</tr>
</tbody>
</table>

The thermal growth of sea ice is treated as a one-dimensional problem motivated by the large vertical temperature gradients, which represent much greater heat fluxes vertically than horizontally (e.g. Leppärinta 1993). In the growth of congelation ice and snow-ice the latent heat released due to freezing is conducted through the ice and snow to the atmosphere by radiation and turbulent heat fluxes. The latent heat released in frazil ice formation is transported away by the turbulence to the free surface and out to the atmosphere. The general Baltic Sea thermodynamics models are well applicable to the GoF and provide quite good results (Leppärinta 1983, Saloranta 2000). Thin ice, however, is difficult to model well.

Apart from the landfast ice zone, the GoF ice forms a dynamic drift ice cover. This is a long and narrow field with dynamics sensitive to the compactness and thickness of the ice and to the wind direction.

The mechanical behaviour of drift ice changes qualitatively with increasing compactness, from a frictionless medium to a viscous fluid and further to an elastic-plastic plate (Leppärinta 2005). Drift ice is a compressible
The packing density (compactness) of ice floes may lie anywhere between zero and one and changes easily. The ice is driven by winds and currents. Wind is a purely external force, but the dynamics of ice and water are coupled. The interfacial ice-water stress depends first of all on the ice-water velocity difference. The faster-moving medium drives the other while the slower medium acts as a retarding frictional force to the faster one. Sometimes the drift can be as large as 20–30 km per day. In open ice fields the floes drift independently and, like any floating drifters on the sea surface, with a velocity close to that of the surface current. At the other extreme, a compact ice field may stand stationary even under considerable forcing because the internal friction of the ice overcomes the external forcing. Indeed, the whole GoF can be covered with landfast ice.

Plastic rheology is considered to represent best the physical behaviour of compact drift ice. The basic motivation for the plastic behaviour is (i) that the internal ice stress in the ice ridging process (assumed to be the main process in the dissipation of kinetic energy) appears to be strain-rate independent, and (ii) that there exists, for a compact drift ice field, a nonzero yield strength. The drift ice material shows strain hardening: as the thickness increases by compression, so does the yield strength. And in the case of drift ice, stress is allowed only during compression since there is no tensile strength.

Several drift ice experiments have been performed in the Gulf of Bothnia (e.g. Leppäranta 1981, Leppäranta et al. 2001) but far fewer in the GoF. Basically, the ice dynamics in the GoF is quite similar to that in the Bay of Bothnia. Baltic Sea ice dynamics models include all the main basins with the same physical descriptions (Haapala & Leppäranta 1996). Specific fine-resolution GoF ice models have been developed as well (Leppäranta & Wang 2002, Wang et al. 2007, 2008). The evolution of the ice conditions can be predicted rather well, the main problems being the ice mechanics in the near-fast ice boundary zone. Also, there is a large change in the ice quality across the basin, from thin pancake ice in the west to heavily ridged pack ice in the east, which requires good advection schemes from numerical models.

The theoretical free drift solution (ice speed 2% of the wind speed, direction 30° to the right of the wind direction) is a useful approximation for open drift ice fields (Leppäranta 2005). But for compact drift ice there is no general simple solution. The mobility of a compact ice cover, however, has been evaluated from experimental data from cold winters. An ice thickness of 50 cm in the drift ice is usually strong enough to resist wind forcing, except when the wind is just pushing ice out of the GoF into the northern Baltic Proper. When the ice thickness is less than 10 cm, it offers little
resistance to wind, and such ice can be easily pushed to the eastern end of the basin, as has been observed.

Sea ice ridges are a major feature of drift ice fields. They present the largest loads to ships. Also, the scouring of the sea bottom by ridge keels needs to be considered when pipelines or cables are laid on the sea floor. A field programme of Baltic sea ice ridges was undertaken in 1987–91. Detailed field investigations of the geometry, internal structure and strength of individual ice ridges, including those in the GoF, are presented in Kankaanpää (1997) and Leppäranta & Hakala (1992). The keel and the sail have triangular cross-sections. A representative slope angle for Baltic ridges is 25°. For the internal structure, the upper part (up to 1.5 m) of the keel or rubble is consolidated ice, while the lower keel as well as the sail are porous (porosity around 0.3). The role of the sail is mainly to serve as a very good feature for ridge identification. The GoF ridges studied were typical ones rather than maximum cases. The largest mapped ridges were 6–8 m thick (Leppäranta & Hakala 1992). It is likely that larger ones exist, in particular at the eastern fast ice boundary between Kotka and Vyborg (Leppäranta & Wang 2002).

High rates of ridged ice production may substantially increase the probability of shipping disasters (Palosuo 1975, Pärn et al. 2007). Since this production rate is sensitive to changes in the ice properties and meteorological conditions, the relevant climate changes may have a large

Figure 22. Resultant ice thickness field after five days of W 10 m s\(^{-1}\) winds. Left: initial ice thickness 20 cm; right: initial ice thickness 50 cm (adapted from Leppäranta & Wang 2002; original color images created by K. Wang)
influence on navigational safety. In addition, the abrupt increase in oil transport along the GoF means increasing risk levels, which is critical in wintertime, when monitoring, forecasting and combing oil spills is particularly difficult (Wang et al. 2008).

A model study of GoF ice dynamics demonstrated that the mobility of ice with a thickness of c. 20 cm is already restricted (Leppäranta & Wang 2002), and that of 50 cm-thick ice is even lower (Figure 22). In the east, 20 cm-thick ice forms ridges up at the mouths of Neva Bay and Vyborg Bay. For thicker ice the channels on either side of Gogland are too narrow, and ridging is blocked out further east. This is what happens in cold winters such as 1987. When the wind forcing can no longer overcome the yield stress of the ice, the GoF ice cover may remain stationary in the mid-winter for up to two months. Displacements are typically smaller, and the length scale of the motion increases so that islands and shoals lock the drift into the fast ice zone.

8.4. Ice climatology

The collection of time series records of the dates of freezing and ice break-up as well as ice thickness commenced in the Gulf of Finland in the 1800s (Leppäranta & Seinä 1985, Jevrejeva et al. 2004). The longest of these time series is the break-up of ice at Helsinki, recorded since 1829. Along with the increase in winter sea traffic a large set of coastal ice observation stations were founded (Palosuo 1953). Statistical ice charts produced by SMHI and FIMR (1982) have been available since the beginning of the 1960s.

During the 20th century, the variability in the dates of freezing and ice break-up has been about two months. Only the latter shows a significant trend toward shorter ice seasons, as much as 15–20 days per hundred years (Leppäranta & Seinä 1985, Jevrejeva & Leppäranta, 2002, Jevrejeva et al. 2004). A decreasing trend has been found in the maximum annual ice thickness. The decreasing trend of the total ice period (understood as the period from the first appearance of ice up to its total disappearance) in the eastern part of the GoF is caused largely by the earlier disappearance of ice (Figure 23). A possible climate warming would move the climatological ice edge further north, and its impact would probably be to increase ice season variability.

Jevrejeva et al. (2004) examined the evolution of ice seasons in the Baltic Sea during the 20th century. The statistical question of combining data from sites with different ice probabilities is solved by using the fractiles of the distributions. These 100-year long time series provide evidence of a general trend toward easier ice conditions. As above, this is best seen in the length of the ice season, which is decreasing by 14–44 days per century. The trends
Figure 23. The dates of the first appearance and disappearance of ice at Loviisa (dotted line) and Narva (solid line) in the Gulf of Finland in 1950–2005, and at Kihnu (dashed line) in the Gulf of Riga. Adapted from Sooäär & Jaagus (2007), with permission from Estonian Academy Publishers. The data for Loviisa are from the Finnish Institute of Marine Research.

of c. 8–20 days per century to earlier ice break-up are in a good agreement with the warming trend in winter air temperatures over Europe.

Numerical models have produced scenarios for future ice conditions in the GoF under climate change (Haapala & Leppäraanta 1997, Haapala et al. 2001). In an average ice season with the year 2100 horizon, only the eastern part of the basin would freeze over, as is the case in very mild winters at present. The length of the ice season is expected to shorten by 1.5–2 months, and the ice thickness would decrease by 20–30 cm.

9. Closing remarks

The Gulf of Finland is a unique sea region: it is a combination of the remote influence of the open sea with significant freshwater impact on the hydrography, biogeochemistry and ecosystem dynamics of the area. Although termed a gulf, this sea-area possesses many characteristics typical of numerous estuaries in the World Ocean: strong salinity gradients and variable current fields. The investigations described have considerably increased our knowledge of GoF dynamics during the last 10 years. Among a handful of highlights, this decade has confirmed that, in spite of the relatively small size of the basin, small values of the internal Rossby radius make the gulf dynamically extremely complex, whereas many of its features are characteristic of those of the open ocean.
The small size plays a two-fold role here. On the one hand, it is the basis of the high susceptibility of the processes in the GoF to external forcing factors. For example, potential changes in river discharge caused by past and future variations in precipitation (that do not necessarily become evident or measurable in many coastal areas of the ocean) are easily detectable here. The dramatic changes in water transparency should be interpreted as a warning sign in this context. On the other hand, the moderate size of the GoF enables basin-wide studies to be performed with a high temporal and spatial resolution within a reasonable budget and allows conclusions to be drawn with confidence about the functioning of the whole gulf.

Many important results obtained for this area either mimic or mirror analogous findings in other parts of the ocean. The combination of various processes and forcing factors in particular highlights the phenomena occurring at the entrance area of the gulf. For example, the potential of the water ‘chimney’ in this area (Elken 2006) has as yet been poorly recognised, but it may play a role comparable to that of salt water inflow in forming the bottom waters of the entire Baltic Sea and in driving a version of a ‘conveyor belt’ in this basin.

The variety of processes and their frequent interactions make this region a unique testing ground for the identification of climate changes and/or shifts in climatic regime, and their consequences. The new findings are thus expected to guide the marine science community to integrated information about relevant factors affecting the dynamics in other parts of the Baltic Sea. They will also have global importance as a guideline for future ecosystem modelling of water bodies with similar estuarine characteristics (e.g. semi-enclosed shelf seas, estuaries/deltas of the major rivers).

Apart from the many well-known effects of the future climate, several much more subtle aspects may change in the GoF. The forecast climate changes are likely to affect factors controlling not only the volume of the water body and its mean temperature, but also the conditions of salt water inflow into the entire Baltic Sea, the overall transport scheme of waters within it, the distribution of upwelling and downwelling patterns, the location of areas of the largest wave intensity and wave-induced mixing, and therefore the vertical and horizontal distribution of salinity, temperature, oxygen and nutrient fluxes, and other decisive background constituents of the local ecosystem (Soomere et al. 2009). For example, an increased sea surface temperature and/or salinity, which may easily happen if the structure of up- and downwellings changes (see Talpsepp 2008), may lead to a reduction in the ice cover, to a shortening of the ice period in the GoF, and to manifold related effects. Wind stress at the sea-surface is evidently increasing during the winter months; this will stretch the extremes in wave
and sea-level heights. The timely detection of such changes is a major challenge for scientists, and launching adaptation measures for society is an accompanying challenge for decision-makers.

Last but not least, the GoF is an area where inspired applications, such as the practical use of the intrinsic properties of current patterns for reducing coastal pollution by means of dynamic ship routing (Soomere & Quak 2007), appear to be technically feasible and economically acceptable. The stability of the major features of the underlying multi-layer flows with respect to changing climatic conditions and the potential of the exploitation of the subsurface current for the prospective location of the major fairway are important issues for further study.

Acknowledgements

The authors are extremely grateful to Torsten Seifert for providing the original, high-resolution image for Figure 1a, Oleg Andrejev for supplying original modelling results for Figures 4, 7 and 10, Jüri Elken for providing the original graphics for Figures 5 and 6, Ants Erm for composing the colour version of Figure 19, and Keguang Wang for producing the colour version of Figure 22.

Obituary

Professor Alexei Vsevolodovich Nekrasov (1933–2008) passed away during the final stages of the work on this paper. Even though he was suffering from a severe illness, he kept up his positive attitude to life and his scientific activities until the very end. He was a world-famous expert in the field of ocean tides. In Baltic Sea research, his work serves as an outstanding contribution to the better understanding of the physics of the Gulf of Finland. Several generations of oceanographers were educated on his various university courses, as well as the expeditions of the Baltic Floating University, of which he was one of the chief initiators.

References


T. Soomere, K. Myrberg, M. Leppäranta, A. Nekrasov


Haapala J., Marnela M., Juottonen A., Leppäranta M., Tuomenvirta H., 2001, 
Modelling the variability of the sea-ice conditions in the Baltic Sea under 

Hela I., 1946, Coriolis-voiman vaikutuksesta Suomenlahden hydrografisiin oloihin, 
Terra, 58 (2), 52–59.

Hela I., 1952, Drift currents and permanent flow, Commentat. Phys.-Math. XVI. 

Hellemee P., 1998, The development of coastal dunes and their vegetation in 

Inkala A., Myrberg K., 2001, Comparison of hydrodynamical models of the Gulf of 

IPCC, 2007, Climate change 2007: The physical science basis, Contribution of 
Working Group I to the Fourth Assessment Report of the Intergovernmental 
Panel on Climate Change, S. Solomon, D. Qin, M. Manning, Z. Chen, 
M. Marquis, K. B. Avert, M. Tignor, & H. L. Miller (eds.), Cambridge Univ. 


Jevrejeva S., Drabkin V. V., Kostjukov J., Lebedev A. A., Leppäranta M., Mironov 
Ye. U., Schmelzer N., Sztoobryn M., 2004, Baltic Sea ice seasons in the twentieth 

Jevrejeva S., Leppäranta M., 2002, Ice conditions along the Estonian coast in 


Johansson M. M., Kahma K. K., Boman H., 2003, An improved estimate for the 
long-term mean sea level on the Finnish coast, Geophysica, 39 (1–2), 51–73.


Jönsson A., Broman B., Rahm L., 2002, Variations in the Baltic Sea wave fields, 
Ocean Eng., 30 (1), 107–126.

Jönsson A., Danielsson A., Rahm L., 2005, Bottom type distribution based on wave 


Jylhä K., Tuomenvirta H., Ruosteenoja K., 2004, Climate change projections for 

Kahma K. K., Pettersson H., 1994, Wave growth in a narrow fetch geometry, Global 
Atmos. Ocean Syst., 2, 253–263.

Kahma K. K., Pettersson H., Tuomi L., 2003, Scatter diagram wave statistics from 

Kankaanpää P., 1997, Distribution, morphology and structure of sea ice pressure 
ridges in the Baltic Sea, Fennia, 175 (2), 139–240.


Physical oceanography of the Gulf of Finland


Talpsepp L., 2006, *Periodic variability of currents induced by topographically trapped waves in the coastal zone in the Gulf of Finland*, Oceanologia, 48(S), 75–90.


