Remote sensing to monitor interactions between aquaculture and the environment of Spencer Gulf, South Australia

Thesis submitted for the degree of

Doctor of Philosophy

Paul Erich Bierman
B.Sc. (Hons)

The University of Adelaide
School of Earth and Environmental Sciences

May 2010
Chapter 1: Introduction

1.1 Background

Worldwide, aquaculture was worth US$78.8 billion in 2006 and represented 47% of total fish food production (FAO 2009). In Australia, the value of aquaculture production in 2006-07 was $793 million, which accounted for approximately one third of total fisheries production (ABARE 2008). Approximately 26% of the national aquaculture production by value was from South Australia, with the southern bluefin tuna the most valuable species (ABARE 2008).

It is known that open-intensive aquaculture operations can potentially impact upon their surrounding environments through the release of waste from uneaten feed or fish excretions. Previous studies have shown that aquaculture wastes are high in nutrients and a significant amount of these nutrients are released to the environment (Gowen and Bradbury 1987; Hall et al. 1990; Holby and Hall 1991; Hall et al. 1992). Increases in nutrients have the potential to increase phytoplankton growth, which can possibly lead to harmful impacts upon the aquaculture and the surrounding marine environment (Anderson et al. 2002). Aquaculture industries are also susceptible to a changing environment and at risk from the effects of harmful algal blooms and other processes that can be potentially damaging to the industry. History has shown the impacts that harmful algal blooms can have on aquaculture operations (Underdal et al. 1989; Kent et al. 1995). As a result, it is necessary that the marine environment surrounding aquaculture is understood and monitored.

The southern bluefin tuna (SBT) aquaculture industry is based in the coastal waters of southwest Spencer Gulf, east of Port Lincoln, South Australia. The aquaculture industry was developed in 1990 through collaboration between the Australian Southern Bluefin Tuna Aquaculture Industry Association (ASBTIA), the Japanese Overseas Fisheries Cooperation Foundation and the South Australian Government. Juvenile SBT are caught from the eastern Great Australian Bight each summer and held in pens offshore from Port Lincoln where they are fed baitfish for a period of 3 – 5 months. Once the SBT have increased in size sufficiently they are harvested and exported, primarily to Japan.
In response to concerns expressed by both the aquaculture industry and the government regulators, and to increase the understanding of the marine environment of the SBT aquaculture region, the Cooperative Research Centre for the Sustainable Aquaculture of Finfish (Aquafin CRC) established a research program titled “Risk and Response – Understanding the Tuna Farming Environment” (Tanner and Volkman 2009). The primary objectives of this collaborative project included characterising the main oceanographic features of the tuna grow-out region and understanding the phytoplankton and nutrient characteristics through the development of three dimensional hydrodynamic and biogeochemical models. A pilot study by Treilibs and Lewis (2007), as part of a previous project by the Aquafin CRC (Tanner 2007), identified that satellite-based remote sensing imagery could be used to improve the understanding of surface water characteristics within and around Spencer Gulf. Thus by providing better information on the spatial and temporal variability of the surface water dynamics, they recommended that MODIS imagery could be used to improve our understanding of the aquaculture environment. This current project forms part of the Aquafin CRC’s Risk and Response project to increase the understanding of the SBT aquaculture environment through application of satellite-based remote sensing imagery.

1.2 Aims

The aim of this project was to increase the understanding of the waters of Spencer Gulf and the region around the SBT aquaculture industry through application of satellite-based imagery. This aim involved three objectives:

- To firstly confirm that MODIS satellite-based imagery is a suitable data source for use in the waters of Spencer Gulf and the SBT aquaculture region, in terms of the accuracy of the derived products and the possible limitations imposed by the coastal locality;

- To apply MODIS imagery to investigate the seasonal variations in chlorophyll-\(a\) concentration and sea surface temperature throughout Spencer Gulf; and

- To investigate the spatial and temporal variations of chlorophyll-\(a\) concentration and sea surface temperature within the SBT aquaculture region near Port Lincoln.
1.3 Scope

The application of satellite imagery in coastal oceans is known to be influenced by a number of factors that can reduce the accuracy of the final products, thus it is necessary for satellite-based imagery to be assessed before it can be applied with confidence in a region such as Spencer Gulf. A comprehensive validation of satellite imagery would involve a number of stages, including assessment of the sensor performance, the atmospheric correction, and the algorithms. However, since the purpose of this study is to confirm the applicability of the final satellite products and not to assess the accuracy of each intermediate stage in the satellite method, only the final MODIS output products are compared to field-based measurements.

The scope of this study is further limited to the assessment and application of MODIS imagery. Several other satellite-based instruments could have been applied in this study; but MODIS has been identified as the most appropriate based upon the high temporal frequency, large spatial coverage, good spectral and radiometric resolution, low cost, and ease of data access. Furthermore, only chlorophyll-\(a\) and sea surface temperature are investigated. Satellite imagery has the potential to be used to estimate other oceanic parameters, including suspended particulate matter and coloured dissolved organic matter; although these methods are mostly experimental and have not been developed or refined to the extent of the chlorophyll-\(a\) and sea surface temperature methods.

Only a few previous studies have applied satellite-based imagery to study the waters of South Australia. A couple of studies have applied SeaWiFS chlorophyll-\(a\) imagery (Kampf et al. 2004; Nieblas et al. 2009), and others have applied AVHRR or MODIS SST imagery (Petrusevics 1993; Herzfeld 1997; Kampf et al. 2004; McClatchie et al. 2006; Middleton et al. 2007; Nieblas et al. 2009) to investigate upwelling and coastal circulation features. Despite these studies, this present study is the first to apply satellite imagery to investigate aquaculture activities in South Australia, the first to apply MODIS ocean colour imagery in South Australia, and the first to validate the accuracy of satellite-based chlorophyll-\(a\) and SST estimates in the region.
1.4 Thesis Structure

The following thesis is divided into several sections. Background information relevant to the study is presented in Chapter 2. This chapter firstly describes the southern bluefin tuna aquaculture industry and the interactions between the aquaculture industry and its surrounding environment. This is followed by a description of the study area of Spencer Gulf, including a summary of the oceanography of the gulf and current knowledge of nutrients and primary productivity in the region. The concepts of satellite remote sensing for marine applications are then introduced and discussed, including the algorithms for calculation of chlorophyll-\(a\) and sea surface temperature and some previous applications. Chapter 3 is a review of analytical methods that can be used to investigate variability in coastal water quality data with an emphasis upon remote sensing imagery. The chapter also illustrates the application of these techniques to a sample dataset of MODIS chlorophyll-\(a\) imagery in South Australia. The material of Chapter 3 was presented at the EcoSummit conference in Beijing, China in 2007 and is to be published in the journal *Ecological Indicators* (Bierman et al. In Press).

Chapter 4 presents the results of the validation of MODIS chlorophyll-\(a\) and sea surface temperature imagery in South Australian waters. The applicability and limitations of the imagery are discussed in relation to their performance against field-based measurements. Chapter 5 then presents observations from the use of MODIS monthly composite chlorophyll-\(a\) and sea surface temperature imagery over 5 years to investigate seasonal variability throughout Spencer Gulf. Chapter 5 is also published, in a slightly modified form along with preliminary validation results from Chapter 4, in the final report for the Aquafin CRC Risk and Response project (Bierman et al. 2009; Tanner and Volkman 2009). Some of the results from both Chapter 4 and 5 were also presented at the International Geoscience and Remote Sensing Symposium (IGARSS) in Boston, USA in 2008 (Bierman et al. 2008). Chapter 6 then investigates fine-scale spatial and temporal variability within the SBT aquaculture region based upon daily MODIS imagery over the southwest Spencer Gulf area for a 6-year period. These results chapters are then followed by a discussion and conclusion chapter, where the implications of the results for the SBT aquaculture industry and the knowledge of the oceanography of the region are discussed. Limitations to the current study are also presented and suggestions presented for future research.
Chapter 2: Background

2.1 Aquaculture and the environment

2.1.1 Aquaculture

World capture fisheries production was worth an estimated US$91.2 billion in 2006, producing 92 million tonnes, of which 82 million tonnes was from marine waters (FAO 2009). In comparison to capture fisheries, which have been relatively stable over the past decade, aquaculture production is the fastest growing animal food-producing sector (FAO 2009). Aquaculture production in 2006 was 51.7 million tonnes, worth US$78.8 billion, and is growing at an annual rate of 7% (FAO 2009). Aquaculture accounted for less than 4% of seafood production in 1970, but as of 2006 accounted for 47% of the total food fish production by value (FAO 2009). Fisheries in Australia is a large industry producing 240,000 tonnes in 2006-07, worth $2.18 billion (ABARE 2008). South Australia”s share of the national fisheries production in 2006-07 was 18% by value, making South Australia the third largest fisheries producer in Australia behind Western Australia and Tasmania with 22% each (ABARE 2008). The gross value of Australian aquaculture production in 2006-07 was $793 million, accounting for approximately one third of the total fisheries production (ABARE 2008). The South Australian aquaculture production by value in 2006-07 was $208 million (ABARE 2008), approximately 26% of the national total. The major aquaculture species of South Australia can be seen in Table 2.1 along with their weight of production and farm-gate value in 2006-07.

It can be seen from Table 2.1 that the most important aquaculture species in South Australia is the southern bluefin tuna (SBT); it is also one of the most valuable aquaculture species in Australia. The fishing of SBT has been occurring in Australia since the 1950s, with the majority of the fish caught in the highly productive waters of the eastern Great Australian Bight during summer. A steady increase since the 1950s in the amount of SBT taken saw the catch peak at 21,500 tonnes in 1982. Due to depletion of stock from overfishing, Australia, Japan and New Zealand formed the international Commission for the Conservation of Southern Bluefin Tuna (CCSBT; www.ccsbt.org) in 1994, which led to the introduction of quotas to limit the total allowable catch. Australia”s annual quota has
since been set at 5,265 tonnes, of which 97% are farmed in Spencer Gulf near Port Lincoln (FRDC 2009). Port Lincoln is the only location in Australia where SBT are farmed.

The southern bluefin tuna aquaculture program at Port Lincoln, on the western shore of Spencer Gulf, comes under the definition of capture-based aquaculture; where fish are sourced from the wild, but grown in captivity using aquaculture techniques (FAO 2004). Schools of juvenile SBT are caught from the eastern Great Australian Bight between December and March each year using the purse-seine method, where a boat circles a school of SBT and lays out a net. The schools are then towed back to the farm site offshore from Boston Island near Port Lincoln, and held in circular pontoons 40 – 50 m in diameter. The SBT are fed sardines or other baitfish, one or two times per day six days per week, until they are harvested 3 – 5 months later. The industry can also be considered a form of open-intensive aquaculture (Lucas and Southgate 2003), where pens are placed in open water to utilise the natural currents to disperse wastes, rather than using artificial circulation, and the nutritional intake of the fish relies upon feed provided by the farmers and not sourced directly from the water column. Recent research by Clean Seas Tuna has enabled them to successfully close the life cycle of the SBT, and thus breed SBT in captivity (McGarry 2008; Austin 2009). As a result it is likely that the Port Lincoln based SBT aquaculture industry will expand rapidly in the near future.

Table 2.1: The major aquaculture species in South Australia and their production in 2006-07 from ABARE (2008).

<table>
<thead>
<tr>
<th>Species</th>
<th>Production (2006-07)</th>
</tr>
</thead>
<tbody>
<tr>
<td>* Other species</td>
<td></td>
</tr>
</tbody>
</table>

* Other species include snapper, microalgae, murray cod, yellowtail kingfish, golden perch and aquarium fish.
The SBT aquaculture activity takes place in the southwest of Spencer Gulf, offshore from Port Lincoln. The waters of southwest Spencer Gulf have been allocated into different aquaculture zones by the Department of Primary Industries and Resources, South Australia (PIRSA 2007). The main aquaculture zones in the Port Lincoln area are the Boston Bay and Lincoln Offshore aquaculture zone, Louth Bay aquaculture zone, Murray Point aquaculture zone, Proper Bay aquaculture zone, and the Todd River aquaculture zone. The Lincoln exclusion zone also exists amongst these aquaculture zones. The SBT aquaculture activities take place in the Boston Island East and the Lincoln Offshore sectors of the Boston Bay and Lincoln Offshore aquaculture zone (Figure 2.1).

Figure 2.1: The Boston Island East and the Lincoln Offshore sectors of the Boston Bay and Lincoln Offshore aquaculture zone, where the southern bluefin tuna aquaculture activities occur; coordinates from PIRSA (2007).
2.1.2 Effects of Aquaculture on the Environment

It is unavoidable that open intensive sea-cage aquaculture, such as the SBT aquaculture in Spencer Gulf, will release some waste into the surrounding environment. Waste products from aquaculture include uneaten food and faeces, which sink to the seafloor below the pens, and excretions that go directly into the water column. General waste products released from aquaculture may include organic carbon, organic nitrogen, ammonium, urea, bicarbonate, phosphate, vitamins, therapeutants (although not used for SBT) and pigments (Gowen and Bradbury 1987). It has been calculated from rainbow trout farms in Sweden that between 75 and 78% of carbon, 61 and 67% of nitrogen and 82% of phosphorus from the feed is lost to the environment (Hall et al. 1990; Holby and Hall 1991; Hall et al. 1992). The majority of nitrogen released to the environment was dissolved in the water and was able to directly stimulate primary production (Hall et al. 1992). It has been estimated for SBT aquaculture that between 76 and 86% of the nitrogen from the baitfish fed to the tuna is released into the water in dissolved form, 59 – 64% from soluble fish excretions and 17 – 22% from leaching of solid wastes (Fernandes et al. 2007).

Increased levels of dissolved nutrients in the water can often result in an increase in phytoplankton growth, particularly when the nutrient limiting growth increases in concentration. An increase in nutrients will not always result in higher levels of phytoplankton growth, since this may be limited by other factors such as the availability of light. In addition, increased growth of phytoplankton may occur without an increase in the abundance of phytoplankton, such as when the phytoplankton are grazed upon by other organisms or are flushed from the system by the currents. It has been observed for the SBT aquaculture region that nitrogen is the most important limiting nutrient, and changes to nitrogen concentrations are likely to be most important to phytoplankton ecology of the region (Thompson et al. 2009). This study of the nutrients of the tuna farming zone also indicated that the nutrient concentrations of the region are relatively low compared to similar coastal regions, and that nutrient inputs from the SBT aquaculture activities and recycling of nutrients through microbial processes in the sediments could lead to considerable changes to the availability of nutrients for phytoplankton growth.

Often as nutrients increase phytoplankton abundance increases also, and different species or groups of species will dominate (Anderson et al. 2002). Increased phytoplankton often has negative impacts on other marine life due to physical damage from spines or barbs of
the cells, or the production of toxins, which can be either released into the water column or transferred to other marine life as the cells are ingested (Hallegraeff 1993; Anderson et al. 2002). On the order of 5,000 species of marine phytoplankton have been identified throughout the world, but just 2% of these are known to be harmful or toxic (Landsberg 2002). Algal blooms can also be detrimental due to a reduction in dissolved oxygen in the water column that can occur as blooms break down (Hallegraeff 1993; Anderson et al. 2002). A number of studies have reviewed harmful algal blooms and their impacts upon the marine environment (Hallegraeff 1993; Smayda 1997; Smayda 1997; Anderson et al. 2002; Landsberg 2002).

Aquaculture can also potentially have other effects on the environment, such as benthic enrichment and fish attraction. Benthic enrichment from the accumulation of waste products below pens can lead to changes in the ecology of the benthos and the sediment chemistry. A number of studies have investigated aquaculture related benthic enrichment, such as Edgar et al. (2005) for Tasmanian salmon farms and Cheshire et al. (1996) for SBT aquaculture. Likewise, aquaculture has the potential to change the ecology and biodiversity by attracting wild fish and other marine life to the area (Carss 1990; Dempster et al. 2002; Boyra et al. 2004).

2.1.3 Environmental Affects on Aquaculture

Aquaculture can stimulate harmful algal blooms, but is also at risk from the influences of algal blooms whatever their cause. Nutrient increases resulting in algal blooms can occur due to natural processes such as upwelling, and also anthropogenic influences such as agricultural fertilizer runoff, sewage effluents, animal wastes, and atmospheric deposition of pollutants (Anderson et al. 2002). As has been discussed, harmful algal blooms affect other marine organisms through toxicity, mechanical and physical damage, and through reduced dissolved oxygen. These influences, however, are likely to be greater for captive fish, which are unable to move away from the dangers.

High concentrations of potentially harmful species have led to closures in shellfish aquaculture at a number of locations throughout South Australia (Marsh 2003). There are also a number of reported cases throughout the world where harmful algal blooms have significantly impacted upon finfish aquaculture operations. Along the Norwegian coast during May and June 1988, a toxic bloom of the phytoflagellate *Chrysochromulina*
polylepis resulted in the death of more than 500 tonnes of farmed fish from more than 120 fish farms, worth US$4.5 million (Underdal et al. 1989). *C. polylepis*, previously thought to be non-toxic, also caused significant damage to wild salmon, sea-urchin and starfish populations and completely eradicated a number of gastropod species (Underdal et al. 1989). Another instance of farmed Atlantic salmon being affected by a harmful algal bloom occurred in British Columbia, Canada, in April 1994. On this occasion a dense bloom of the diatoms *Skeletonema costatum*, *Thalassiosira aestivalis* and *T. rotula* resulted in the death of 16,000 fish. The diatoms, which were also previously thought to be non-dangerous to fish, irritated the gills of the salmon leading to excess mucus production and asphyxiation (Kent et al. 1995).

Aquaculture activities are also at risk from increases in suspended particulate matter in the water column. The disastrous southern bluefin tuna mortality event of 1996 in South Australia is an example of how a rapidly changing environment can impact severely on caged fish. On this occasion the suspension of fine sediments from the seafloor due to increased wave energy contributed to the death of 1700 tonnes of captive fish (Clarke 1996). Fernandes et al. (2006) studied the sediment characteristics around the southern bluefin tuna farming zone and discovered areas of fine sediments that can be easily mobilised off the seafloor. The calculations of Fernandes et al. (2006) indicate that these sediments may become resuspended with a current speed of only 20 cm s⁻¹, which is not uncommon in the area, and that they would remain in suspension for a period of at least 2 days. Thus these sediments could again pose a significant health risk for southern bluefin tuna farmed in the area, by irritating the gills of the fish leading to respiratory difficulties.
2.2 Spencer Gulf

2.2.1 Spencer Gulf

Spencer Gulf forms part of a unique gulf system along the coastline of South Australia (Figure 2.2). This triangular shaped inverse estuary is closed in the north but opens to the continental shelf in the south, and is bounded by Eyre Peninsula to the west and Yorke Peninsula to the east. The gulf covers an area of approximately 22,000 km² and spans a distance of over 300 km from the head to the mouth. Water depth within the gulf ranges from less than 10 m in the upper gulf, up to 50 m near the entrance. Along the shores of Spencer Gulf there are several towns whose major industries rely upon the gulf. In the southwest of the gulf near the southern tip of Eyre Peninsula is Port Lincoln, the unofficial capital of Eyre Peninsula. The most important and valuable industries of the area are fishing and aquaculture, with Port Lincoln supporting one of the largest commercial fishing fleets in Australia. Towards the north of the gulf are the towns of Whyalla, Port Augusta and Port Pirie, which support iron and steel manufacturing industries, power stations and lead smelting works respectively. On the eastern side of the gulf is the town of Wallaroo, a major commercial port and base for a large prawn trawling fleet. Along with these industries, the economy of Spencer Gulf also benefits from increasing tourism, and it is a popular destination for recreational fishing and water sports, especially during the summer months.

![Figure 2.2: South Australia and Spencer Gulf.](image-url)
2.2.2 Oceanography

Spencer Gulf belongs to a subclass of estuaries known as negative or inverse estuaries. This type of estuary is defined as a marine embayment where seawater is concentrated by the removal of freshwater, as opposed to the traditional positive estuary where seawater is diluted by the addition of freshwater (Pritchard 1952). In Spencer Gulf, evaporative losses exceed freshwater inputs, due to the semi-arid climate and lack of river input, leading to salinity within the gulf that is noticeably different from the nearby oceanic waters, and the establishment of a permanent salinity gradient. The salinity maximum occurs at the head and ranges between 43 psu (equivalent to parts per thousand) in late winter up to a maximum of 48 psu in late summer. The salinity decreases with distance from the head to reach 36 psu at the mouth (Nunes and Lennon 1986; Nunes Vas et al. 1990; Corlis et al. 2003).

Spencer Gulf shows a large seasonal temperature variation due to low thermal inertia from its shallow depth, with gulf waters warmer than the adjacent shelf waters in summer but cooler in winter (Nunes and Lennon 1986; Lennon et al. 1987; Nunes Vas et al. 1990). The difference in temperature between the gulf and ocean leads to considerable temperature gradients across the mouth of the gulf in both summer and winter (Nunes Vas et al. 1990; Petrusevics 1993). The water temperatures within the gulf range from a minimum of ~12 °C in winter up to ~24 °C in summer, with little variation throughout the majority of the gulf at any time during the year (Nunes and Lennon 1986; Nunes Vas et al. 1990). Petrusevics (1993) has described the formation and breakdown of the sea surface temperature (SST) front across the mouth of the gulf from early spring (September 1988) through to late autumn (May 1989). In early spring the temperature inside the gulf was observed to be 0.5 °C greater than the adjacent shelf. This difference rose to 0.8 °C by late November as the gulf warmed. The temperature difference continued to increase throughout summer from 1 °C in December to 3 °C in February and up to a maximum of 4 °C in April. This maximum temperature difference equates to a temperature gradient of 8x10^{-2} °C km^{-1}. The temperature difference across the front rapidly decreased to zero by May. Measurements of density across this front showed a density minimum at the front indicating a zone of convergence, and implying reduced exchange between the waters inside the gulf and on the shelf. Hence the SST front limits the exchange between the gulf and the nearby shelf during summer with no direct mixing between gulf and shelf waters; thus the gulf remains isolated during this period.
As mentioned, high evaporation in Spencer Gulf leads to higher salinity inside the gulf in relation to the nearby shelf. In summer, both the temperature and salinity inside the gulf are high, but as the water temperature drops in autumn a density gradient is established (Lennon et al. 1987). Lennon et al. (1987) have described how cooling of the high salinity gulf waters generates a gravity current that flows outwards from Spencer Gulf. The current follows the central channel and exits the gulf on the eastern side of the mouth, then flows across Investigator Strait and past the west coast of Kangaroo Island down the shelf near Du Couedic Canyon to a depth of 250 m (Lennon et al. 1987). Lennon et al. (1987) estimate a flow speed of 0.1 m s\(^{-1}\), and suggest that this outflow must continue for 3 months to remove all the salt accumulated within the gulf from the summer evaporation. Nunes Vas et al. (1990) suggest that the waters in the east of the gulf are longer isolated from contact with oceanic water than the water to the west, based on the higher salinity measurements in the east. Thus they imply that Spencer Gulf undergoes clockwise circulation. This finding is in agreement with Lennon et al. (1987), who suggest that the outflow of high salinity water occurs on the eastern side, and with Nunes and Lennon (1986), who use the alignment of the isohalines to suggest a cyclonic gyre bringing in ocean water in the west and removing gulf water in the east at a circulation velocity of 2 cm s\(^{-1}\). Although density exchange between the northern and southern gulf appears to occur year round, the out-flowing gravity current is restricted to the wintertime when the SST front no longer limits communication between the gulf and the shelf (Nunes and Lennon 1986; Nunes Vas et al. 1990). During winter the current is not believed to exit the gulf continuously, but is modulated by the strong spring-neap tidal cycle of the gulf (Lennon et al. 1987; Nunes Vas et al. 1990).

Spencer Gulf undergoes a 14-day cycle where the tidal amplitudes and current speeds range from maximum values at the spring tides to almost zero at the neap (dodge) tides 7 days later (Noye 1984). Spring tides produce considerable mixing and turbulence, preventing stratification from developing by vertically mixing the water column, while the weak neap tides result in the formation of salinity and density stratification. Hence during neap tides the high salinity and high density gravity current formed in the upper gulf can flow out of the gulf under an inflow of fresher and less dense water from the south (de Silva Samarasinghe and Lennon 1987; Lennon et al. 1987; Nunes and Lennon 1987; de Silva Samarasinghe 1989; Nunes Vas et al. 1989; Nunes Vas et al. 1990).
Outside the mouth of Spencer Gulf and to the west of Eyre Peninsula is the Great Australian Bight (GAB). The eastern GAB supports one of the most valuable pelagic ecosystems in Australian waters and is a point of congregation for juvenile southern bluefin tuna during summer months (Ward et al. 2006). One cause of this region’s productivity is the presence of coastal upwelling. The southern Eyre Peninsula is one of three locations in southern Australia where coastal upwelling is observed, along with southern Kangaroo Island and the Bonney Coast (Kampf et al. 2004). Coastal upwelling occurs when winds blow along the coastline, with the coast to the right in the Southern Hemisphere, consistently for a period of at least a few days. This leads to surface water being transported offshore, which is replaced by an inflow of cooler, higher nutrient water from deeper on the shelf. Kampf et al. (2004) observed such events occurring. They suggest that just 2 or 3 upwelling events occur each summer, with each event dropping local SST by 2 – 3 °C and increasing chlorophyll-\(a\) to greater than 4 \(\mu g\) L\(^{-1}\). Further study of the Eyre Peninsula upwelling was conducted by McClatchie et al. (2006), who show that it is not independent of the Kangaroo Island upwelling. Instead they suggest that the Kangaroo Island upwelling leads to the formation of a subsurface pool of nutrient rich water, which is drawn to the shallow coastal regions of the eastern GAB during subsequent upwelling events.

The GAB also contains a plume of warm water that originates in the shallow waters of the western bight during early spring and extends in a south-eastward direction throughout summer and autumn. Herzfeld (1997) and Herzfeld and Tomczak (1997) observed this feature and indicate that this warm water is produced by local influences and is independent of the Leeuwin Current. These studies do, however, suggest that the Leeuwin Current, which flows into the bight from the Western Australian coast, does connect with the GAB water in winter to produce a continuous band of warm water. The results of these studies suggest that this water can be 2 – 3 °C warmer than the surrounding water, and can extend as far east as 136°E, past the mouth of Spencer Gulf.

A recent study by Herzfeld et al. (2009) has investigated the circulation of Spencer Gulf and the tuna farming zone based upon hydrodynamic modelling. This study showed the presence of the clockwise gyre in the south of the gulf during winter that has also been observed in previous studies (Nunes and Lennon 1986; Lennon et al. 1987; Nunes Vas et al. 1990). The circulation is characterised by the outflow of high density water in the east, and an inflow on the western side of the mouth creating a northward flow through the tuna.
farming region. Herzfeld et al. (2009) showed that the seasonally averaged currents through the outer tuna farming zone for summer and winter are generally weak, at around 1 cm s\(^{-1}\), in a northwards direction. This is in agreement with Bierman (2005), who measured a net current of approximately 1 cm s\(^{-1}\), in a northeast direction, over 41 days during July and August 2005, resulting in a net displacement of 855 m day\(^{-1}\). Herzfeld et al. (2009) also calculated the flushing times of the region and showed that the tuna farming zone is flushed in 2 days, while the flushing time for the more sheltered Boston Bay is much longer at 8 days. The greater flushing in the tuna farming zone is likely due to its exposure to the larger gulf-wide circulation. Flushing plays an important role in dispersing excess nutrients from aquaculture activities and well flushed areas will be less likely to experience side effects from accumulation of nutrients such as increased phytoplankton abundance.

2.2.3 Nutrients and Phytoplankton

Little is known about the nutrient status of the entire Spencer Gulf. One study to examine nutrient concentrations on a gulf-scale was conducted by Smith and Veeh (1989). It was observed that phosphate concentrations increased with decreasing salinity, hence towards the mouth, but both nitrite and ammonia showed no spatial variation. Dissolved organic nitrogen, however, showed a linear increase in concentration with increasing salinity, towards the northern gulf. It was estimated that 90% of phosphate is removed from the system through primary production, and with no sources other than the mouth of the gulf, there is almost no phosphate remaining in the high salinity waters at the head of the gulf. However, unlike phosphorus, nitrogen is supplied from the atmosphere through nitrogen fixation. Thus Smith and Veeh (1989) believe the limiting nutrient in Spencer Gulf to be phosphorus, rather than nitrogen which is typically considered to be the limiting nutrient in marine environments.

The nutrients around the SBT tuna farming zone were investigated in a recent study by Thompson et al. (2009) as discussed earlier. The study measured surface and seafloor concentrations of nitrate, nitrite, ammonium, phosphate and silicate monthly around the tuna farming zone. No differences between surface and bottom samples were observed, and only nitrate showed any spatial variation with the lower concentrations observed furthest offshore. Some seasonal variations were observed, with maximum nitrate observed around September and October, and a peak in silicate between January and March. As
mentioned, nitrogen has the largest potential to limit production in the region and was limiting between January and August. However, phosphorus was also potentially limiting in September, as was silicon in December. Wild-Allen and Skerratt (2009) simulated the nutrient dynamics of the region using a biogeochemical model and showed depleted nitrogen and phosphorus offshore during summer due to phytoplankton assimilation. The model also showed higher nutrients around the farming areas possibly associated with fish farm inputs.

There have been a few studies that have investigated aspects of phytoplankton and chlorophyll-\(a\) variability of the SBT aquaculture region. Bierman (2005) analysed surface chlorophyll-\(a\) concentrations across the tuna farming zone as part of an investigation into the oceanographic conditions of the region. Water samples were collected on grids across the tuna farming zone in both March and June 2005. Surface chlorophyll-\(a\) concentrations in March ranged between 0.23 and 1.94 µg L\(^{-1}\) with a mean of 0.85 µg L\(^{-1}\). Spatial patterns in surface chlorophyll-\(a\) in March showed greater concentrations in the southwest of the study area in the waters just east of Boston Island. In June, surface chlorophyll-\(a\) concentration ranged from 0.63 up to 2.40 µg L\(^{-1}\), with a mean of 1.07 µg L\(^{-1}\). On this occasion greatest concentrations were observed in the north, in an area of greater numbers of tuna pens. However, a comparison of chlorophyll-\(a\) collected near tuna pens with control sites showed no increase in chlorophyll-\(a\) around the pens. Due to chlorophyll-\(a\) only being measured on two occasions, temporal variations could not be assessed.

Paxinos et al. (1996) studied the chlorophyll-\(a\) and phytoplankton throughout the region during 1995. Chlorophyll-\(a\) measurements ranged between 0.2 and 0.75 µg L\(^{-1}\), with the average chlorophyll-\(a\) inside Boston Bay slightly greater than outside the Bay due to the various terrestrial nutrient inputs. It was suggested that the phytoplankton dynamics of the Boston Bay region are similar to other regions of the Australian coastline (Paxinos et al. 1996). Paxinos (2007) also investigated the dynamics of phytoplankton around the tuna farms of Port Lincoln every 6 to 8 weeks between September 1997 and March 1999. Water samples were collected from 26 sites throughout the region, and included samples from around tuna cages and at control sites. Maximum surface chlorophyll-\(a\) concentrations occurred between March and June, with the range of values observed on the eastern side of Boston Island between 0.15 and 2.35 µg L\(^{-1}\). Averaging of the data by sampling date appeared to show increased chlorophyll-\(a\) at tuna cage sites compared to controls on a number of occasions, although statistical analysis showed no significant difference.
between cages and controls. There was, however, a statistically significant difference between the mean chlorophyll-a of an area of the lower Spencer Gulf region where no tuna farming activities occur, and the areas of the study where aquaculture activities are located. Diatoms were the dominant phytoplankton in the lower Spencer Gulf region, away from the aquaculture activity, while dinoflagellates were predominant east of Boston Island and diatoms and dinoflagellates were almost equally abundant inside Boston Bay.

Further investigation of phytoplankton dynamics around the SBT aquaculture region was conducted by the Aquafin CRC’’s Risk and Response project (Tanner and Volkman 2009). van Ruth et al. (2009) studied the temporal and spatial variability of phytoplankton in the tuna farming zone monthly between September 2005 and September 2006, in order to evaluate possible risks posed by harmful algae to the ecosystem and aquaculture industry. Water samples were collected on an east-west transect through the tuna farming zone. The majority of the chlorophyll-a occurred in the smaller size fraction of phytoplankton, < 5 µm. Surface chlorophyll-a concentrations peaked in May 2006 with minimum concentrations observed in November 2005. The mean chlorophyll-a of the region for each month ranged from approximately 0.2 µg L⁻¹ up to 1.5 µg L⁻¹, but there was large variability between locations with no clear spatial patterns. The total phytoplankton abundances followed the same seasonal trends as the chlorophyll-a concentration, with diatoms the most dominant phytoplankton group. Dinoflagellates and other phytoplankton groups were also present; although the temporal variation in diatom abundance strongly dominated the trends in total abundances and the chlorophyll-a concentrations. The diatom bloom in May 2006 appears to be related to a peak in silica concentrations in February that year.

van Ruth et al. (2009) investigated the primary production rates in the tuna farming zone. Productivity rates were measured at three locations around the Port Lincoln SBT aquaculture region during March, May, June and October 2007. Greatest primary production rates were observed during March, prior to the annual peak in phytoplankton abundance that occurs during autumn. Productivity decreased in May indicating the end of the annual rise in biomass, and further decreased in July and October. van Ruth et al. (2009) also suggest that the productivity of the aquaculture region is comparable to highly productive upwelling regions during the build-up to the autumn peak, but similar to oligotrophic regions during other times.
A number of potentially harmful phytoplankton species have been identified in the southwest of Spencer Gulf (Clarke 1996; Paxinos et al. 1996; Munday and Hallegraeff 1998; Paxinos 2007; van Ruth et al. 2009). Paxinos et al. (1996) identified *Pseudonitzschia*, which produces domoic acid causing the death of seabirds which consume affected fish (Garrison et al. 1992), and *Dinophysis acutae* and *D. fortii*, which both produce toxins that accumulate in shellfish, resulting in diarrhoeic shellfish poisoning when the shellfish are consumed (James et al. 1997; James et al. 1999). Clarke (1996) and Munday and Hallegraeff (1998) also identified *Chattonella marina* and *Heterosigma* in and around Boston Bay. *Chattonella marina* is known to affect fish by the production of brevetoxins and reactive oxygen radicals (Landsberg 2002), while some forms of *Heterosigma* have been shown to produce hydrogen peroxide (H$_2$O$_2$), which alters gill structure in fish and results in asphyxiation (Twiner and Trick 2000). Paxinos (2007) identified the potentially toxic dinoflagellates *Karenia mikimotoi*, which produces both haemolytic ichthyotoxins and is known to be responsible for a number of farmed fish kills, and *Karenia brevis*, which produces a brevetoxin responsible for neurotoxic shellfish poisoning (Paxinos 2007 and references therein). van Ruth et al. (2009) also identified the potentially dangerous diatoms *Chaetoceros*, that can cause physical damage due to its siliceous exoskeleton, and *Psuedonitzschia*, known to cause amnesic shellfish poisoning, in the tuna farming region.
2.3 Remote Sensing of Marine Environments

2.3.1 Ocean Colour

*Introduction to Ocean Colour*

It is the optical properties of dissolved and suspended matter in the upper ocean that determine ocean colour. Three main components, in addition to water itself, are responsible for the optical properties of natural waters. These are (1) phytoplankton, which includes other microscopic organisms, (2) inorganic suspended particulate matter (SPM), which includes particulate material from suspended sediments or river inputs, and (3) coloured dissolved organic matter (CDOM), from the degradation of plankton cells or organic matter input from terrestrial sources (IOCCG 2000).

Phytoplankton are the microscopic plants that convert sunlight to energy through photosynthesis in the oceans and are the base of the aquatic food web. They are generally smaller than 1 mm and cannot be seen with the naked eye. The nitrogen containing pigment called chlorophyll-\(\alpha\) is the main pigment responsible for photosynthesis and is found in all plants that photosynthesize. The concentration of chlorophyll-\(\alpha\) in the water column is commonly used as a proxy for the abundance of phytoplankton. Phytoplankton is the most important group of substances for causing changes to the optical properties of aquatic environments (IOCCG 2000) and chlorophyll-\(\alpha\) concentration is the most widely used product derived from ocean colour data (Carder et al. 2004).

Suspended particulate matter (SPM) is also known as total suspended solids (TSS), total suspended matter or total suspended material (TSM). This category comprises the inorganic matter that is derived from the suspension of sediments from wave and current action and the input of particles through river outputs, hence the influence of SPM is typically confined to the coastal zone (IOCCG 2000). Under the definition provided by the International Ocean-Colour Coordinating Group (IOCCG 2000), SPM includes only the inorganic particles, and not organic particulates. SPM is derived from a variety of sources each with different optical properties. For example, sands suspended from the seafloor will influence the ocean colour quite differently to clays introduced through river inputs (IOCCG 2000). The primary impact of these particles in the water is an increase in scattering and backscattering of light, resulting in an increased reflectance across the whole spectrum (Morel and Belanger 2006).
Coloured (chromophoric) dissolved organic matter, is also known as CDOM, yellow substance, gelbstoff or gilvin. CDOM is the fraction of the dissolved organic matter that absorbs light in the ultra-violet and visible regions of the spectrum, and can significantly modify the underwater light field in the oceans (Hoge et al. 1993). CDOM also includes organic particulate material not included in the SPM component (IOCCG 2000). CDOM is one of the strongest absorbing constituents in the coastal ocean and it can significantly reduce the photosynthetically active radiation available to phytoplankton (Hoge et al. 1993; Nieke et al. 1997). Waters high in CDOM appear darker due to an increase in absorption at shorter wavelengths (Morel and Belanger 2006). CDOM in estuaries and the coastal zone originates as fulvic and humic acids from the breakdown of plant matter and enters the waterway through river discharge and surface runoff (Carder et al. 1989). In oceanic regions the CDOM is produced from the bacterial degradation of plankton (Bricaud et al. 1981).

Marine waters are often segregated into two distinct classes depending on the factors that influence ocean colour. The two groups are known as Case 1 and Case 2 waters. Simply, Case 1 waters are clear ocean waters where the ocean colour is predominantly affected by phytoplankton, while Case 2 waters are coastal and estuarine waters where other constituents also affect the ocean colour. A formal definition is provided by the International Ocean-Colour Coordinating Group (IOCCG 2000) who define that „Case 1 waters are those waters in which phytoplankton (with their accompanying and covarying retinue of material of biological origin) are the principal agents responsible for variations in optical properties of the water“ while „Case 2 waters are influenced not just by phytoplankton and related particles, but also by other substances, that vary independently of phytoplankton, notably inorganic particles in suspension and yellow substance“. Therefore in Case 1 waters, the contributions of constituents other than phytoplankton are relatively small. In Case 2 waters, however, the substances other than phytoplankton, including SPM and CDOM, have to be treated as independent variables. Hence, Case 2 waters are more optically complex than Case 1 waters, and require new algorithms, since simplifying assumptions made for Case 1 waters do not hold for Case 2 waters (IOCCG 2000). Case 2 waters can be further divided to include waters primarily dominated by sediments, Case 2S, and by yellow substance (CDOM), Case 2Y (Morel and Belanger 2006).
Ocean Colour Sensors

The coastal zone colour scanner (CZCS) was the first satellite sensor to monitor ocean colour and was launched by NASA in 1978. The CZCS was a proof of concept mission to determine whether satellite-based instruments could be used to reliably measure information on ocean colour properties such as chlorophyll-\(a\) concentration (Gordon et al. 1980; Hovis et al. 1980). The CZCS experienced technical difficulties throughout its lifetime, including an intermittent infrared temperature sensor, degrading sensitivity and power supply problems which led to its retirement in 1986 (GSFC 1996). Despite this, the CZCS mission was considered a success (Barale and Schlittenhardt 1993) and led to a better understanding of the optical properties of aquatic substances (IOCCG 2000). The CZCS mission better identified the requirements for calibration, validation, atmospheric correction, bio-optical algorithms, data processing and data access for future missions such as the Sea-viewing Wide Field-of-view Sensor (SeaWiFS) (Hooker and McClain 2000). The characteristics of imagery from CZCS, SeaWiFS, MODIS and other sensors can be seen in Table 2.2.

SeaWiFS was NASA’s follow-on to the CZCS project and was launched in 1997. The main purpose of the SeaWiFS program is „to acquire data that is critical for the study of the role of ocean primary production in global biogeochemistry” (Hooker et al. 1992). Other goals of the program include, but are not limited to, determining the temporal and spatial extent of phytoplankton blooms, quantifying the ocean’s role in global carbon cycles and to advance scientific applications of ocean colour data (Hooker et al. 1992). The SeaWiFS instrument has been very successful and, at the time of writing, has been operating for almost 12 years.

NASA’s Moderate Resolution Imaging Spectroradiometer (MODIS) sensor builds upon experience derived from CZCS and SeaWiFS, with an emphasis on calibration and stability of the sensor through on-board calibrators (Esaias et al. 1998). Major differences in MODIS, compared to CZCS and SeaWiFS, include the fact that the MODIS sensor does not tilt, the ocean colour bands of MODIS are narrower than SeaWiFS to help with atmospheric correction and MODIS has more stringent signal-to-noise ratio specifications to reduce sensor-induced errors (Esaias et al. 1998). Another significant variation to the previous sensors is that MODIS records in 36 different bands, as opposed to just 6 and 8 bands for CZCS and SeaWiFS respectively (Table 2.2).
Table 2.2: The satellite sensors capable of ocean colour measurements.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Operation</th>
<th>Bands</th>
<th>Swath (km)</th>
<th>Spatial Resolution (m)</th>
<th>Radiometric Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>CZCS</td>
<td>1978 – 1986</td>
<td>6</td>
<td>1556</td>
<td>825</td>
<td>8 bit</td>
</tr>
<tr>
<td>SeaWiFS</td>
<td>1997 – present</td>
<td>8</td>
<td>1500</td>
<td>1100</td>
<td>10 bit</td>
</tr>
<tr>
<td>MODIS *</td>
<td>1999 – present</td>
<td>36</td>
<td>2330</td>
<td>250, 500, 1000</td>
<td>12 bit</td>
</tr>
<tr>
<td>OCTS</td>
<td>1996 – 1997</td>
<td>12</td>
<td>1400</td>
<td>700</td>
<td>10 bit</td>
</tr>
<tr>
<td>GLI</td>
<td>2002 – 2003</td>
<td>36</td>
<td>1600</td>
<td>250, 1000</td>
<td>12 bit</td>
</tr>
<tr>
<td>MERIS</td>
<td>2002 – present</td>
<td>15</td>
<td>1150</td>
<td>300</td>
<td>12 bit</td>
</tr>
<tr>
<td>Landsat ^</td>
<td>1984 – present</td>
<td>7</td>
<td>185</td>
<td>30</td>
<td>8 bit</td>
</tr>
</tbody>
</table>

^ Landsat TM and ETM+ sensors on Landsat 5 and 7.

Table 2.3: The bands of the MODIS instrument and their primary application.
(Source: [http://modis.gsfc.nasa.gov/about/specifications.php](http://modis.gsfc.nasa.gov/about/specifications.php))

<table>
<thead>
<tr>
<th>Band</th>
<th>Bandwidth (nm)</th>
<th>Primary Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>620 – 670</td>
<td>Land/Cloud/Aerosols Boundaries</td>
</tr>
<tr>
<td>2</td>
<td>841 – 876</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>459 – 479</td>
<td>Land/Cloud/Aerosols Properties</td>
</tr>
<tr>
<td>4</td>
<td>545 – 565</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1230 – 1250</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1628 – 1652</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>2105 – 2155</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>405 – 420</td>
<td>Ocean Colour / Phytoplankton / Biogeochemistry</td>
</tr>
<tr>
<td>9</td>
<td>438 – 448</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>483 – 493</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>526 – 536</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>546 – 556</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>662 – 672</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>673 – 683</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>743 – 753</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>862 – 877</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>890 – 920</td>
<td>Atmospheric Water Vapour</td>
</tr>
<tr>
<td>18</td>
<td>931 – 941</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>915 – 965</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>3 660 – 3 840</td>
<td>Surface / Cloud Temperature</td>
</tr>
<tr>
<td>21</td>
<td>3 929 – 3 989</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>3 929 – 3 989</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4 020 – 4 080</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4 433 – 4 498</td>
<td>Atmospheric Temperature</td>
</tr>
<tr>
<td>25</td>
<td>4 482 – 4 549</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>1 360 – 1 390</td>
<td>Cirrus Clouds, Water Vapour</td>
</tr>
<tr>
<td>27</td>
<td>6 535 – 6 895</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>7 175 – 7 475</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>8 400 – 8 700</td>
<td>Cloud Properties</td>
</tr>
<tr>
<td>30</td>
<td>9 580 – 9 880</td>
<td>Ozone</td>
</tr>
<tr>
<td>31</td>
<td>10 780 – 11 280</td>
<td>Surface/Cloud Temperature</td>
</tr>
<tr>
<td>32</td>
<td>11 770 – 12 270</td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>13 185 – 13 485</td>
<td>Cloud Top Altitude</td>
</tr>
<tr>
<td>34</td>
<td>13 485 – 13 785</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>13 785 – 14 085</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>14 085 – 14 385</td>
<td></td>
</tr>
</tbody>
</table>
The MODIS spectral bands and their primary uses can be seen in Table 2.3. The MODIS sensor orbits the earth onboard two satellites: Terra, launched December 18th 1999, and Aqua, launched on May 4th 2002. Both satellites have near-polar orbits at an altitude of 705 km. The orbits are sun-synchronous, which means that the satellites pass over the same location at approximately the same time each day; Terra crosses the equator at approximately 10:30 am local time, while Aqua crosses the equator at approximately 1:30 pm. The Terra and Aqua satellites each have a true repeat cycle of 16 days, but due to the large swath width of the imagery MODIS provides global coverage every 1 – 2 days. Both MODIS instruments are identical, but the Terra instrument is affected by calibration issues that influence ocean colour estimates (Franz et al. 2007). Thus only MODIS Aqua imagery can be used for chlorophyll-a. The MODIS imagery covers a swath 2,330 km wide and has a pixel size of 250 m for bands 1 – 2, 500 m for bands 3 – 7 and 1000 m for bands 8 – 36, which includes the ocean-related bands. Another advantage of MODIS over previous sensors is the 12-bit radiometric resolution allowing for better resolution in radiance (see Barnes et al. 1998; Esaias et al. 1998; Savtchenko et al. 2004; NASA undated). As well as their application for ocean colour, the MODIS instruments are also designed for terrestrial and atmospheric applications.

There are also other satellite-based sensors designed to measure ocean colour such as the Japanese Aerospace Exploration Agency’s (JAXA) Ocean Colour and Temperature Scanner (OCTS) and Global Imager (GLI) and the European Space Agency’s (ESA) Medium Resolution Imaging Spectrometer (MERIS) instrument. The OCTS was launched in 1996 on the Advanced Earth Observing Satellite (ADEOS), but only operated for ten months due to an accident with the satellite (Kawamura 1998). Despite this, the OCTS mission provided a valuable dataset of satellite ocean colour measurements (Kawamura 1998). The GLI instrument was launched in December 2002 on ADEOS-II, but only operated until October 2003. MERIS was launched in 2002 and orbits the Earth on ESA’s ENVISAT satellite, providing global coverage every 3 days (ESA 2006). MERIS’’ primary objective is to measure ocean biology and water quality, but it is also able to contribute to atmosphere and land surface studies (ESA 2006).

Attempts have also been made to measure chlorophyll-a concentrations with the Landsat series of satellites. Currently both Landsat 5 and Landsat 7 are still operational. Landsat 5 carries the Thematic Mapper (TM) sensor and Landsat 7 the Enhanced Thematic Mapper (ETM+). Landsat TM / ETM+ imagery would have a great advantage over MODIS and
SeaWiFS for investigating chlorophyll-a variability on local scales due to its much higher spatial resolution; but the Landsat sensors were not designed for ocean colour applications. Numerous attempts have been made to derive empirical algorithms for the calculation of chlorophyll-a from Landsat TM and ETM+ imagery (Kim and Linebaugh 1985; Dwivedi and Narain 1987; Tassan 1987; Baban 1993; Lavery et al. 1993; Pattiaratchi et al. 1994; Mayo et al. 1995; Gabric et al. 1998; Han and Jordan 2005). These studies vary widely in approach to relate chlorophyll-a to Landsat TM/ETM+ reflectance or radiance. Despite this, it does appear that local empirical algorithms for Landsat TM/ETM+ chlorophyll-a can be developed, but the algorithms are restricted to the location, and possibly period of time, from which they were derived and are not transferable. Hence, there exists no globally-accepted Landsat chlorophyll-a method, and consequently Landsat application to study ocean and coastal chlorophyll-a variability is limited.

MODIS Ocean Colour Algorithms

A number of different algorithms have been developed to allow chlorophyll-a concentration to be calculated from MODIS satellite imagery. These algorithms are either empirical, based upon relating observed radiance measures to field-based chlorophyll-a measurements, or semi-analytical, based upon bio-optical modelling and an understanding of reflectance, absorption and scattering. The ocean biology processing group at NASA’s Goddard Space Flight Centre (GSFC) produce operational ocean colour products for the MODIS-Aqua instrument that are available from the ocean colour website, http://oceancolor.gsfc.nasa.gov/, and can be processed with the SeaWiFS Data Analysis System (SeaDAS) software, http://oceancolor.gsfc.nasa.gov/seadas/. The standard operational chlorophyll-a product distributed by the GSFC is derived via the empirical OC3M algorithm, which is based upon the SeaWiFS OC4v4 algorithm (O'Reilly et al. 2000). Alternative MODIS-Aqua chlorophyll-a products, which can be produced with SeaDAS although not distributed by the GSFC, are the Clark empirical chlorophyll-a (Clark 1997), the Garver-Siegel-Maritorena version 1 (GSM01) semi-analytical bio-optical model (Maritorena et al. 2002), and the Carder semi-analytical bio-optical model (Carder et al. 2003).

The SeaWiFS OC4v4 empirical chlorophyll-a algorithm is described in detail in O'Reilly et al. (2000). The MODIS OC3M is essentially the same algorithm as the SeaWiFS OC4v4 with modifications to account for the slight differences in band positions in MODIS

24
compared to SeaWiFS. The empirical algorithm relates \textit{in situ} chlorophyll-\(a\) measurements to the ratio of above-water reflectance at blue and green wavelengths. The \textit{in situ} dataset, consisting of 2,853 observations across a range of marine environments with different optical characteristics, is the largest ever assembled for algorithm refinement (O'Reilly et al. 2000). The relationship between chlorophyll-\(a\) concentration in mg m\(^{-3}\) (equivalent to \(\mu g\) L\(^{-1}\)) and remote sensing reflectance follows a fourth-order polynomial shown in Equation 2.1:

\[
Chl = 10^\gamma (a_0 + a_1 R_{3M} + a_2 R_{3M}^2 + a_3 R_{3M}^3 + a_4 R_{3M}^4)
\]  

(2.1)

where

\[
R_{3M} = \log_{10}\left[\frac{R_{443}}{R_{550}} \times \frac{R_{490}}{R_{550}}\right]
\]

The argument represents the maximum of the ratios of remote sensing reflectance at 443 nm (Band 9) to 550 nm (Band 12) and 490 nm (Band 10) to 550 nm (Band 12) (O'Reilly et al. 2000).

The Clark chlorophyll-\(a\) algorithm is also an empirically derived relationship between field-based chlorophyll-\(a\) concentrations and normalised water-leaving radiances, and follows from the CZCS empirical chlorophyll-\(a\) algorithms (Clark 1997). The basic form of the Clark algorithm is shown in Equation 2.2:

\[
Chl = 10^\gamma (a_0 + a_1 R)
\]  

(2.2)

where

\[
R = \frac{L_{\text{ww}}(443nm) + L_{\text{ww}}(530nm)}{L_{\text{ww}}(550nm)}
\]

and \(L_{\text{ww}}(443nm)\), \(L_{\text{ww}}(530nm)\) and \(L_{\text{ww}}(550nm)\) represent the normalised water-leaving radiances in MODIS bands 9, 11 and 12 respectively (Clark 1997; GSFC 2008).

Empirical algorithms, like Clark and OC3M/OC4v4, are best suited for application in Case 1 waters, defined as waters where phytoplankton are the principal agent responsible for variations in optical properties (IOCCG 2000), such as open ocean waters. In Case 2 or coastal waters, however, water constituents other than phytoplankton, such as suspended particulate matter (SPM) and coloured dissolved organic matter (CDOM), influence the optical properties of the water (IOCCG 2000). Thus the accuracy of empirically derived
algorithms in Case 2 waters is limited. Semi-analytical algorithms such as the Carder and GSM01 algorithms are designed to be applicable to coastal areas and Case 2 waters by taking these constituents into account (Maritorena et al. 2002; Carder et al. 2003).

Analytical algorithms incorporate radiative transfer modelling of the relationship between radiance and the inherent optical properties of backscattering and absorption. The term „semi-analytical“ is used to refer to the fact that aspects of the model are expressed by empirical relationships. The Carder model relates the observed remote sensing reflectance at a particular wavelength, $R_{rs}(\lambda)$, to the ratio of the backscatter term, $b_b(\lambda)$, and the absorption term, $a(\lambda)$. The backscatter term is broken down into the sum of the backscatter due to water, $b_{bw}(\lambda)$, and due to particles, $b_{bp}(\lambda)$, and the absorption term is broken down into the sum of the absorption due to water, $a_w(\lambda)$, phytoplankton, $a_{\phi}(\lambda)$, detritus, $a_d(\lambda)$, and CDOM, $a_g(\lambda)$. The model is inverted to calculate the absorption due to phytoplankton, $a_{\phi}(\lambda)$, which is then converted to a chlorophyll-$a$ concentration in mg m$^{-3}$ via Equation 2.3:

$$Chl = P_0 \ast [a_{\phi}(675\text{nm})]^P, \quad (2.3)$$

where $P_0$ and $P_1$ are empirical constants derived via regression with field-based datasets. Sometimes the model is unable to determine a value for $a_{\phi}(\lambda)$, and hence resorts to an empirical form and the chlorophyll-$a$ is determined from the band ratio of MODIS band 10 (488 nm) and 12 (551 nm) via Equation 2.4 (Carder et al. 2003):

$$\log[chl] = c_0 + c_1 \log\left(\frac{R_{488}}{R_{551}}\right) + c_2 \log\left(\frac{R_{488}}{R_{551}}\right) + c_3 \log\left(\frac{R_{488}}{R_{551}}\right), \quad (2.4)$$

Full details of the Carder algorithm are provided in the algorithm theoretical basis document of Carder et al. (2003).

The GSM01 model was developed by optimising the Garver and Siegel model (Garver and Siegel 1997), developed for the Sargasso Sea in the North Atlantic, for global application (Maritorena et al. 2002). The GSM01 model relates the normalised water-leaving radiance ($L_{wn}$), to the inherent optical properties (IOP) of backscatter and absorption and their components. The model is then inverted to determine the absorption due to phytoplankton, which is related to the concentration of chlorophyll-$a$ in a similar manner to the Carder
algorithm. Full details of the GSM01 model are provided by Garver and Siegel (1997) and Maritorena et al. (2002).

**Ocean Colour Applications**

There are a number of factors that can influence the estimation of chlorophyll-a concentrations from satellite observations. Therefore the resultant satellite-based chlorophyll-a products need to be assessed against field-based data in order to have confidence in their accuracy. Several studies have attempted this, employing a number of different methods to assess the different contributions to the accuracy of the output products. Bailey and Werdell (2006) discuss in detail methods for validation of the GSFC’s ocean colour products and show validation results for SeaWiFS measurements against *in situ* data from the SeaBASS dataset collected throughout the world’s oceans (Werdell et al. 2003). Bailey and Werdell (2006) discuss recommendations for validation, which include spatial averaging of pixels to account for errors in navigation and a window of 3 hours between field and satellite-based measurements to limit the influence of temporal variability. They then present results for SeaWiFS radiance measurements against *in situ* water-leaving radiance measurements, to assess the accuracy of the sensors calibration and the accuracy of the atmospheric correction, and show that the SeaWiFS water-leaving radiances perform within 6 – 12% for deep water pixels taking into account errors within the *in situ* measurements. Other studies have assessed the performance of different chlorophyll-a algorithms against *in situ* chlorophyll-a measurements. Pinkerton et al. (2005) for example investigated the performance of a number of band-ratio chlorophyll-a algorithms developed for different platforms in the waters around New Zealand. They considered algorithms for MODIS, SeaWiFS, CZCS, OCTS, MERIS and GLI and tested how well different algorithms for each of these satellite sensors compared to *in situ* chlorophyll-a data collected from a number of regions. They applied the ocean colour algorithms to measurements of surface reflectance to compare with the *in situ* HPLC-derived concentrations, rather than using satellite-derived chlorophyll-a estimates, thus they only assessed the performance of the algorithm, and not errors arising from sensor calibration or atmospheric correction. Overall it was considered that all the algorithms in question agreed well with the *in situ* measurements of chlorophyll-a, and the OC4v4 algorithm was the best of the 8 algorithms tested. The accuracy standards for ocean colour
imagery are generally considered by the ocean colour community to be ±5% for water-leaving radiances and ±35% for chlorophyll-α concentration (McClain 2009).

Over the past decade, satellite-based measurements have been very successful in generating high quality ocean colour data (McClain 2009). The application of satellite-based imagery for chlorophyll-α investigations has become more frequent, particularly in the last decade (McClain 2009), and as such the number of topics and disciplines in which chlorophyll-α imagery is utilised is too large to cover in detail in this review. Some common applications of chlorophyll-α imagery include observations of seasonal variations in phytoplankton, ocean currents and circulation, tracking of phytoplankton blooms, and quantifying primary productivity in the oceans.

Satellite-based ocean colour imagery has become frequently used to assess seasonal variations in chlorophyll-α at many locations throughout the world’s oceans. For example Rivas et al. (2006) applied SeaWiFS chlorophyll-α imagery over six years to investigate annual and inter-annual variations in chlorophyll-α and phytoplankton for the Patagonian Shelf in South America. By applying SeaWiFS chlorophyll-α imagery they were able to develop a more comprehensive understanding of the seasonal variability of chlorophyll-α over the shelf than had previously been developed, and identified the importance of the frontal zones to the biology of the region. Likewise Navarro and Ruiz (2006) investigated spatial and temporal variability in phytoplankton in the Gulf of Cadiz in Spain using SeaWiFS chlorophyll-α images over a multi-year period. Unlike Rivas et al. (2006), who provided a mostly descriptive analysis of the seasonal chlorophyll-α variations, Navarro and Ruiz (2006) applied a more statistical analysis through the application of Empirical Orthogonal Functions (EOF). The EOF analysis, similar to principal components analysis (PCA), identified the primary modes of variability in the time-sequence of chlorophyll-α images and showed regions of differing chlorophyll-α and oceanographic characteristics within the study area.

Leterme and Pingree (2008) applied SeaWiFS chlorophyll-α imagery, along with satellite-based sea level data, to investigate the structure of Gulf Stream eddies in the North Atlantic. It was observed that cyclonic cold-core eddies were higher in chlorophyll-α than the anti-cyclonic warm-core eddies, as a result of localised upwelling and mixing of nutrients into the euphotic zone. The study also observed the propagation speed and direction of the eddies along the Gulf Stream and was able to infer the current flows
associated with the eddies based upon patterns of chlorophyll-a observed. Cipollini et al. (2001) used OCTS and SeaWiFS chlorophyll-a imagery to observe the propagation of Rossby waves across mid-latitude oceans. Rossby waves are a major feature associated with the global ocean circulation and are the dynamic response to atmospheric forcing upon the oceans that drives the ocean basin circulations (Chelton and Schlax 1996). Rossby waves have been previously described using SST and sea surface height by Chelton and Schlax (1996). Despite the fact that the chlorophyll-a changes associated with the Rossby waves were small in relation to the annual range of chlorophyll-a, they could still be detected in global ocean colour imagery, which suggests Rossby waves play a role in ocean biology (Cipollini et al. 2001).

Stumpf et al. (2003) have applied SeaWiFS ocean colour imagery to detect phytoplankton blooms in the Gulf of Mexico, and to identify and predict the trajectory of potentially harmful *Karenia brevis* blooms. Their investigations into methods to detect and predict harmful algal blooms along the Florida coast led to the development of a near real-time harmful algal bloom bulletin distributed to the local government to allow them to prepare for the impact of the blooms. Ahn et al. (2006) also used SeaWiFS imagery to investigate harmful algal blooms of *Cochlodinium polykrikoides* in Korean waters, while both Kutser et al. (2006) and Reinart and Kutser (2006) have investigated methods to detect cyanobacteria blooms in the Baltic Sea using a range of satellite sensors. A consistent theme in studies of harmful algal blooms is that chlorophyll-a imagery is capable of detecting areas of elevated concentration, but is not currently capable of identifying species or phytoplankton groups. Some species of interest have been shown to be detectable using location-specific algorithms, often only with imagery with higher spectral resolutions than the SeaWiFS or MODIS sensors. Despite the current limitations, the International Ocean Colour Coordinating Group (IOCCG) has a working group currently investigating the use of ocean colour imagery for identifying phytoplankton functional types (http://www.ioccg.org/), and research into the application of ocean colour imagery for phytoplankton functional type identification continues (Nair et al. 2008).

Several studies have utilised satellite chlorophyll-a imagery to understand primary productivity by phytoplankton in the oceans and to assess the uptake of carbon (Eppley et al. 1985; Antoine et al. 1996; Antoine and Morel 1996; Behrenfeld and Falkowski 1997). One of the first studies to investigate the use of satellite chlorophyll-a imagery to estimate ocean primary productivity was by Eppley et al. (1985) for the Southern California Bight;
field-based productivity and chlorophyll-$a$ concentrations were measured to investigate the non-linear and spatially and temporally inconsistent relationship between the two variables. They suggest that vertical changes in chlorophyll-$a$ will interfere with attempts to determine water column productivity. Behrenfeld and Falkowski (1997) developed a more comprehensive primary productivity model known as the vertically generalised production model (VGPM), which calculates depth-integrated phytoplankton carbon fixation primarily based upon satellite-based observations of chlorophyll-$a$. Falkowski et al. (1998) applied the algorithm of Behrenfeld and Falkowski (1997) based upon CZCS chlorophyll-$a$ imagery to investigate global ocean productivity, and thus were able to show the ocean primary productivity and carbon uptake on a global scale. The Ocean Productivity group from Oregon State University (www.science.oregonstate.edu/ocean.productivity/) provide data on the net primary productivity of the oceans based upon the VGPM of Behrenfeld and Falkowski (1997), as well as a modified form of the VGPM (Eppley-VGPM) that takes into account temperature dependence upon phytoplankton growth (Eppley 1972; Morel 1991) and a carbon-based production model (CbPM) (Behrenfeld et al. 2005).

Despite the number of international studies involving satellite-based chlorophyll-$a$ imagery, there have been few in the coastal waters of South Australia, and none validating the accuracy of the products. Kampf et al. (2004) and Nieblas et al. (2009) utilised SeaWiFS chlorophyll-$a$ imagery, along with other methods, to investigate upwelling along the South Australian coastline. The SeaWiFS imagery was used to locate anomalous surface chlorophyll-$a$ concentrations, thus identifying where and when coastal upwelling occurs. Used in conjunction with field-based surveys, SST imagery, and wind measurements, SeaWiFS chlorophyll-$a$ imagery has been shown to be beneficial to investigate the behaviour of coastal upwelling in South Australia.

van Ruth et al. (2009) applied 8-day averaged SeaWiFS chlorophyll-$a$ imagery for the southern Spencer Gulf and Port Lincoln region over a ten-year period. A seasonal cycle in chlorophyll-$a$, consistent with field-based measurements with a peak in May, was observed. This study was able to give one of the first indications of the annual and inter-annual variability of chlorophyll-$a$ in the region.
2.3.2 Sea Surface Temperature

Introduction to Sea Surface Temperature

Satellite-based measurements of sea surface temperature (SST) utilise Planck’s Law, which states that all objects above absolute zero emit electromagnetic radiation with the intensity of the emitted radiance at wavelength $\lambda$, $B(\lambda, T)$, related to the temperature of the surface, $T$. The temperature of the emitting surface can be calculated by inverting the Planck Function and using the measured Planck radiance at the given wavelength. Satellite sensors designed for surface temperature measurements in the infrared spectrum can record the radiance emitted by the surface at wavelengths near either 4 $\mu$m or 11 $\mu$m, where the absorption due to the atmosphere is at a minimum. Some satellite sensors can also measure SST in the microwave spectrum, which has the advantage that they can observe through clouds; but the spatial resolution is much coarser compared to infrared techniques.

SST Instruments

The first satellite-based sensors designed to determine sea surface temperatures were the CZCS and the Advanced Very High Resolution Radiometer (AVHRR) onboard NOAA’s TIROS-N satellite, launched in October 1978. The CZCS included a single band in the thermal infrared at 10.5 to 12.5 $\mu$m to measure surface temperatures. However, the CZCS thermal band never functioned satisfactorily due to losses of sensitivity and was not used after the first year of operation (GSFC 1996). The first AVHRR consisted of 4 bands, with both the bands from 3.55 – 3.93 $\mu$m and 10.50 – 11.50 $\mu$m capable of measuring surface temperature. The AVHRR instrument has since been launched on the satellites NOAA-6 through to NOAA-18 and was improved to include 5 bands in 1981 and 6 bands in 1998. The European Space Agency’s (ESA) first Along Track Scanning Radiometer (ATSR-1) was launched in 1991 onboard the European Remote Sensing satellite (ERS-1). Subsequently the ATSR-2 and Advanced ATSR (AATSR) were launched in 1995, onboard ERS-2, and in 2002, onboard ENVISAT, respectively. NOAA’s Geostationary Operational Environmental Satellites (GOES) have also been determining SST operationally since 1994. The SST functionality of MODIS builds upon the experience on the Advanced Very High Resolution Radiometer (AVHRR) and the Along Track Scanning Radiometer (ATSR) sensors (Brown and Minnett 1999).
**MODIS SST algorithms**

MODIS has the capability to measure radiances in both the 4 µm and 11 µm regions of the infrared spectrum. Both forms of the MODIS SST algorithm are described by Brown and Minnett (1999) and Franz (2006). The thermal infrared algorithm, known as the 11µm SST, utilises MODIS bands 31 and 32 at 11.030 and 12.020 µm respectively. The mid-range infrared algorithm, known as the 4 µm SST or SST4, utilises MODIS bands 22 and 23 at 3.959 and 4.050 µm. These two methods both have advantages and limitations. In the spectral region around 4 µm, the spectra of the earth’s emitted radiance and the reflected solar radiance overlap, thus measured radiances may include a combination of both. As a result the 4 µm SST measurements can only be determined at night when solar reflection does not occur. The level of solar radiation at 11 µm is not significant compared with the levels emitted from terrestrial sources, and thus solar reflection does not contribute significantly to the 11 µm bands. However the bands around 11 µm are more strongly influenced by water vapour absorption than the 4 µm bands. The 4 µm bands have greater temperature sensitivity than the 11 µm bands, thus are better able to resolve smaller differences in temperature.

The thermal infrared (11 µm) SST algorithm, as produced by the GSFC follows Equation 2.5:

\[
\text{sst}_{11} = a_0 + a_1 \cdot T_{31} + a_2 \cdot (T_{31} - T_{32}) \cdot \text{bsst} + a_3 \cdot (T_{31} - T_{32}) \cdot (\sec \theta - 1)
\]

(2.5)

where \(T_{31}\) and \(T_{32}\) are brightness temperatures in MODIS bands 31 and 32 at 11.030 and 12.020 µm, \(\theta\) is the sensor zenith angle and \(\text{bsst}\) is the baseline or reference SST (Franz 2006).

The mid-range infrared (4 µm) SST algorithm as produced by the GSFC follows Equation 2.6:

\[
\text{sst}_{4} = b_0 + b_1 \cdot T_{22} + b_2 \cdot (T_{22} - T_{23}) + b_3 \cdot (\sec \theta - 1)
\]

(2.6)

where \(T_{22}\) and \(T_{23}\) are the brightness temperatures in MODIS bands 22 and 23 at 3.959 and 4.050 µm and \(\theta\) is the sensor zenith angle (Franz 2006).
SST Applications

Satellite-based sea surface temperature imagery is widely used for many applications, for example in investigating local and global oceanography, fisheries, meteorology applications, and climate change analysis. On an ocean-basin scale, Kostianoy et al. (2004), for example, applied AVHRR SST imagery to investigate ocean fronts in the southern Indian Ocean, to the south of Australia. They used maps of SST gradients to identify and map the locations of oceanic fronts, thus increasing understanding of the temporal variability in their positions and intensity. On a more local scale, Vancamp et al. (1991) used AVHRR SST and CZCS chlorophyll-α imagery to investigate upwelling and currents along the northwest African coastline. Armstrong (2000) also applied AVHRR SST imagery to investigate the upwelling along the southern Californian coast, while Venegas et al. (2008) applied AVHRR SST imagery, as well as SeaWiFS chlorophyll-α and satellite-derived wind stress and altimetry, to investigate the northern Californian current system.

Myers and Hick (1990) investigated the relationship between catches of southern bluefin tuna and satellite-derived estimates of sea surface temperature from AVHRR along the Western Australian coastline. They only found a weak correlation between specific sea surface temperature and tuna catches, although it was found that high numbers of tuna concentrated in warm-core eddies that spin off from the Leeuwin Current. Thus they showed that identification and tracking of eddies by satellite SST imagery could be used to optimise fishing practices in the region. Reddy et al. (1995) also investigated the relationship between catches of southern bluefin and albacore tuna and satellite SST along the east coast of Tasmania. They showed that the tuna aggregated within eddies and along thermal fronts, thus they created an algorithm to locate thermal gradients within AVHRR SST imagery where tuna are likely to be located. Santos (2000) provides a review of the application of satellite and airborne remote sensing imagery for fisheries. Not only does the paper discuss the use of SST and ocean colour imagery for fisheries research, but many other passive and active remote sensing methods that have not been discussed here.

Satellite SST imagery also has applications in meteorology. Singh et al. (2006), for example, demonstrate how satellite imagery, including SST data, can be used to calculate surface latent and sensible heat fluxes over the ocean on a global scale. Satellite-based SST
imagery is also applied operationally in numerical weather prediction and ocean forecasting by the Australian Bureau of Meteorology (Beggs 2007).

Although time periods of at least 10 – 15 years are required for reliable long-term trends to be observed (Allen et al. 1994), several studies have utilised satellite-based SST observations to investigate changes in SST as a result of climate change. Lawrence et al. (2004) investigated long-term SST trends using AVHRR imagery from 1985 to 2000 and ATSR imagery from 1991 to 1999. They used monthly composite imagery for each sensor and calculated a global average SST for each month and observed trends in the global temperature over time. Seasonal cycles and El Nino signals were removed from the observations and the results indicate an increase in global SST of 0.09 ± 0.03 °C per decade for AVHRR and 0.13 ± 0.06 °C per decade for ATSR. Likewise Good et al. (2007) investigated long-term trends in global SST using AVHRR SST imagery between 1985 and 2004. They showed trends of increasing global SST of 0.18 ± 0.04 °C and 0.17 ± 0.05 °C per decade for day-time and night-time SST imagery respectively over the 20 year period. Hansen et al. (2006) also discussed global changes in temperature in relation to climate change. They compared observed sea, land and air temperatures to model simulations to ascertain accurate rates of global temperature increases. Along with model simulations, meteorological station data, and even paleoclimate temperatures, satellite SST data was discussed to assess long-term trends in temperature.

Satellite-based SST imagery has been applied in South Australia on several occasions. The first application was by Petrusevics (1993), who used AVHRR SST imagery to investigate the SST front across the mouth of Spencer Gulf and showed how the SST front limits water movements across the mouth of the gulf during summer. Herzfeld (1997) also used AVHRR SST imagery and investigated the annual temperature cycles in the Great Australian Bight. More recently Kampf et al. (2004), McClatchie et al. (2006), Middleton et al. (2007), and Nieblas et al. (2009) all utilised satellite SST imagery, along with other data, in investigations into upwelling in South Australia.
Chapter 3: A review of methods for analysing spatial and temporal patterns in coastal water quality

3.1 Abstract

Coastal environments contain some of the marine world’s most important ecosystems and represent significant resources for human industry and recreation. Water quality in the coastal environment is extremely important for a number of reasons from the protection of marine organisms and the well being of marine ecosystems to the health of people in the region and the safety of industries such as aquaculture. As a result it is essential that environmental health in coastal environments is monitored. Traditional monitoring methods include assessment of biological indices or direct measurements of water quality, which are based on in situ data collection and hence are often spatially or temporally limited. Remote sensing imagery is increasingly used as a rich source of spatial information, providing more detailed coverage than other methods. But the complexity of information in the imagery requires new analysis techniques that allow us to identify the components and possible causes of spatial and temporal variability.

This chapter presents a review of methods to analyse spatial and temporal variations in remote sensing data of coastal water quality and discusses and compares these methods and the outcomes they achieve. Selected techniques are illustrated by using a sample dataset of MODIS chlorophyll-a imagery. We consider classification methods (cluster analysis, discriminant analysis) that may be used in exploratory, confirmatory and predictive ways, methods that summarize and identify patterns within complex datasets (factor analysis, principal components analysis, self-organizing maps), and techniques that explicitly analyse spatial relationships (the semivariogram and geographically weighted regression). Each technique has a different purpose and addresses different questions. This review identifies how these methods can be utilized to address water quality variability in order to foster a wider application of such techniques for coastal water quality assessment and monitoring.
3.2 Introduction

The majority of the world’s most productive marine ecosystems are found within coastal environments and owe their productivity, diversity and wealth of life to their terrestrial adjacency. Shelf regions are key areas for biological activity and generate the biological production that supports 90% of the world’s fish catches (Pauly et al. 2002). Coastal marine environments also contain greater biodiversity than open ocean regions (Gray 1997). These productive marine ecosystems are important habitats for many fish and other marine organisms that are not only a significant source of food for human consumption, but are also vital components of marine ecosystems. The proximity of these ecosystems to land often provides nutrients and favourable conditions that support primary producers, which in turn provide nourishment for higher levels of the food chain. However, the threat of rapid and often devastating changes to water quality through both anthropogenic and natural mechanisms is often increased as a result of the connection to the land. Changes in water quality can be potentially catastrophic for marine ecosystems as species are threatened by conditions which are no longer suitable for their survival. These changes in water quality also pose threats to humans through changes in waters utilised for recreation, fishing and industry. It is clear that human reliance on these precious resources requires regulatory interventions and that robust and reliable ecosystem condition indicators are needed for decision making (Pinto et al. 2009).

Coastal environments throughout the world are affected by eutrophication and harmful algal blooms, often as a result of anthropogenically driven increases in nutrients (Anderson et al. 2002). Harmful algal blooms can have significant consequences for marine ecosystems and the people dependant upon them (Hallegraeff 1993; Smayda 1997; Anderson et al. 2002). Reductions in water quality can also be caused by increases in concentration of pollutants or contaminants like oil, heavy metals, organic compounds etc. (Shahidul Islam and Tanaka 2004), increases in turbidity (Orpin et al. 2004), and changes in dissolved oxygen (Sanchez et al. 2007), which all have implications for the well-being of marine ecosystems and those reliant upon them.

As a result of the implications variations in water quality can have for both marine ecosystems and human well-being, it is necessary to monitor coastal environments. By monitoring changes in water quality we can observe, assess and correct long term trends in water quality degradation from anthropogenic sources such as industry and coastal
development. We can better understand changes in water quality due to natural sources. Finally we can possibly even reach an ability to predict changes in water quality to a point where the negative effects upon natural and economically important environments can be mitigated.

It is recognised that indicators need to be “applicable over a variety of spatial scales and conditions to support global as well as local comparisons” (Rees et al. 2008). But it is inherently difficult to develop and validate indicators, if spatio-temporal pattern are complex. This refers to indirect methods of using biological indicators (e.g. Borja and Dauer 2008; Weisberg et al. 2008), as well as for direct assessment of chemical and physical water quality properties. Statistical power in any underlying analysis can be improved if temporal and spatial data are used as a covariate or simply to stratify field collection points into similar areas hence allowing data to be pooled. This would reduce stochastic variability and thus act similar to a low pass filter for noisy data. The vast and growing amount of spatial information in the public domain has motivated us to examine the methodological potential for using satellite imagery in coastal marine water quality monitoring and assessment.

The methods of collecting water quality data depend upon the property or parameter to be measured, the scale of the study, and the questions to be addressed. Traditional techniques can include targeted campaigns of in situ field sampling, moored instruments, drifting instruments, and ship of opportunity measurements. These techniques often provide accurate and reliable water quality information, but the resultant data are frequently limited in space and time. Recently, the application of satellite and airborne remote sensing imagery to collect water quality information has become more frequent (e.g. Goetz et al. 2008, and references therein). Remote sensing datasets are generally more comprehensive than those collected using other techniques in that they provide greater spatial coverage with finer resolution and often increased temporal frequency and resolution. This makes remote sensing a rich source of data. However, the large amount of data presents challenges for the extraction of meaningful information of water quality parameters.

Several analytical techniques are available to extract spatial and temporal patterns and trends in order to provide enhanced understanding and resolve the full information hidden within water quality data. These techniques can be used to identify regions or periods of
time with different water quality characteristics, determine whether significant differences in water quality occur between different regions, and also to indicate the variables responsible for water quality variations. Remote sensing datasets, however, are generally spatially and temporally comprehensive, large in volume, and often contain internal correlation and redundancy: traditional statistical techniques are often inappropriate, but conversely, there is scope for development of new techniques to fully exploit the richness of the data.

The purpose of this chapter is to review analytical techniques that can be used to examine spatiotemporal variability in coastal water quality derived via remote sensing data. The differences between techniques, their advantages and limitations, as well as the questions they address and the outcomes they can achieve are reviewed in order to identify the most appropriate techniques for differing circumstances.

Along with reviewing these techniques and summarising some previous applications, several of the techniques will be illustrated with a sample dataset of MODIS imagery. This dataset consists of a one-year series of monthly MODIS chlorophyll-\(a\) images for 2006 over the coastal areas of South Australia, between 30 and 40 °S and 129 and 141 °E (Figure 3.1). The monthly MODIS images were obtained from the Goddard Space Flight Centre (oceancolor.gsfc.nasa.gov) and analysed in ENVI (RSI 2006). Along with the MODIS imagery, a set of field-based chlorophyll-\(a\) measurements collected in Spencer Gulf is also utilised when required.

Figure 3.1: South Australia.
3.3 Multivariate Analytical Techniques

Numerous multivariate methods are available to analyse spatial and temporal trends in water quality datasets. Some techniques are applicable only to more traditional data comprising measurements of discrete samples, while others are transferable to remote sensing data and some are exclusively applied to remote sensing data. These methods vary in the questions they address and the outcomes they reach. Table 3.1 summarises the techniques to be discussed in this chapter, their purposes and relationships to one another. Techniques such as cluster analysis (CA), discriminant analysis (DA), factor analysis (FA) and principal components analysis (PCA) are widely used, not just in marine applications but also in many other areas of science, to explore structure and relationships in multivariate data, and in some cases to predict responses. While many of these statistical methods have been widely used with point sample data, several have found new application to remotely sensed data. Some analyses explicitly address spatial variability and relationships (self-organising maps (SOM), the semivariogram and geographically weighted regression (GWR)) and naturally find most application with remotely sensed data. These techniques are reviewed to provide insight into how they can be utilised to derive conclusions from water quality data.
<table>
<thead>
<tr>
<th>Type</th>
<th>Technique</th>
<th>Purpose</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classification</strong></td>
<td>Cluster Analysis</td>
<td>Exploratory and Confirmatory</td>
<td>Classification of objects into groups displaying similar properties</td>
</tr>
<tr>
<td>Discriminant Analysis</td>
<td>Explanatory and Predictive</td>
<td></td>
<td>Determine variables responsible for segregation of objects into groups</td>
</tr>
<tr>
<td><strong>Summary</strong></td>
<td>Principal Components Analysis / Empirical Orthogonal Functions</td>
<td>Exploratory Data Reduction</td>
<td>Identify sources of variation in multivariate data</td>
</tr>
<tr>
<td>Factor Analysis</td>
<td>Explanatory</td>
<td></td>
<td>Determine underlying factors responsible for variations</td>
</tr>
<tr>
<td>Self-organising Maps</td>
<td>Predictive</td>
<td></td>
<td>Identify commonly encountered scenarios or patterns within spatial datasets</td>
</tr>
<tr>
<td><strong>Spatial Relationship Analysis</strong></td>
<td>Semivariogram Assessment</td>
<td></td>
<td>Understanding spatial relationships between measurements</td>
</tr>
<tr>
<td></td>
<td>Predictive</td>
<td></td>
<td>Interpolate between samples</td>
</tr>
<tr>
<td>Geographically Weighted Regression</td>
<td>Exploratory Predictive</td>
<td></td>
<td>Understand changes in relationships between variables in space</td>
</tr>
</tbody>
</table>
3.3.1 Cluster Analysis

The purpose of cluster analysis is to partition a set of objects into two or more groups based upon the similarity of the objects with respect to a chosen set of characteristics, so that similar objects are in the same class (Jackson 1983; Hair et al. 1995). Cluster analysis can be applied in both an exploratory and confirmatory sense: it can be used to either develop a new classification of the objects or to confirm a proposed grouping of the objects.

There are two distinct forms of cluster analysis: hierarchical and non-hierarchical. Hierarchical cluster analysis can be either agglomerative, where each object begins in its own cluster and then in subsequent steps the two clusters that are most similar are combined, or divisive, where all objects start in a single cluster and then in each step the objects that are most dissimilar are split off and made into smaller clusters. The hierarchical relationships between objects are commonly displayed graphically in a dendrogram or tree diagram. Non-hierarchical cluster analysis, also known as K-means clustering, differs in that it does not involve the construction of a dendrogram and requires the number of classes to be pre-specified. Initially a cluster seed is established, which is an initial guess at the cluster mean, and the objects within a predetermined threshold distance of the seed are included in the resulting cluster. Further cluster seeds are then chosen until all objects are assigned to a cluster. Unsupervised classification is a form of non-hierarchical cluster analysis applied to remote sensing imagery (Tou and Gonzalez 1974; Richardson 1993; Campbell 1996). The unsupervised classification procedure initially calculates initial class means and then iteratively classifies each pixel of the image into the nearest cluster. This procedure generates a classification map showing the regions of the image displaying similar properties. Technical details of the methodologies and mathematics are beyond the scope of this review, but for more information regarding the different forms of cluster analysis and their mathematical methods refer to Anderberg (1973), Tou and Gonzalez (1974), Hand (1981), Romesburg (1984), Hair et al. (1995) and Jain et al. (1999).

Cluster analysis is widely used in many disciplines to group objects/observations/samples that are described by a series of unrelated variables or measurements. In this particular case, we are concerned with the application of cluster analysis to differentiate water bodies, based upon multiple water quality parameters. Most commonly cluster analysis is
applied to discrete samples, relying on their distribution throughout a study region to infer spatial patterns.

Cluster analysis has been widely used in water quality studies in freshwater river systems. For example Singh et al. (2004) applied cluster analysis, as well as other techniques, to evaluate the water quality of the Gomti River in India, performing agglomerative hierarchical cluster analysis on a dataset consisting of 24 parameters measured at 8 locations monthly over a 5-year period, a total of 17,790 observations. The 8 regions of the river system were successfully grouped into 3 clusters representing low, moderate and high pollution. Many other studies have also applied hierarchical cluster analysis to river water and coastal water quality data: for example the Pisuerga River in Spain (Vega et al. 1998); Suquia River in Argentina (Alberto et al. 2001); the lagoon of Venice (Solidoro et al. 2004); Mahanadi River and estuary in India (Panda et al. 2006); and the Fuji River in Japan (Shrestha and Kazama 2007).

Application of hierarchical cluster analysis is illustrated in Figure 3.2. Hierarchical cluster analysis was performed on ten temporal chlorophyll-$a$ profiles extracted from the monthly MODIS imagery from 2006 for different areas of the South Australian coastal waters. Cluster analysis produced a dendrogram separating each location by the similarities in the chlorophyll-$a$ temporal profiles. As can be seen from Figure 3.2(b), stations 4, 6, 7, and 3 are all very different from the other stations. These stations are located within Spencer Gulf and Gulf St Vincent and in the upwelling region of the southeast. Thus from the results of the hierarchical cluster analysis we can conclude that these locations have different chlorophyll-$a$ characteristics from the rest of the study area, and therefore different phytoplankton and primary productivity properties, assuming the accuracy of the satellite-based chlorophyll-$a$ estimates.
Figure 3.2: (a) The location of the 10 stations at which chlorophyll-$a$ temporal profiles were extracted from MODIS monthly chlorophyll-$a$ images for 2006, and (b) an example dendrogram produced from a hierarchical cluster analysis of chlorophyll-$a$ temporal profiles at these locations.

An example of non-hierarchical cluster analysis is provided by McNeil et al. (2005) who applied k-means clustering to 30 years of surface water chemistry data from freshwater rivers, streams and lakes across Queensland, Australia. Approximately 34,000 measurements of many water chemistry variables were taken between the 1960s and 1990s covering the entire state. A two-stage k-means cluster analysis of this dataset first identified 347 groups of measurements, which were then further reduced to nine water quality types. These nine water quality provinces highlighted the natural processes of the regions influencing surface water chemistry.

Unsupervised classification is widely used to derive cover classes or groups from digital remote sensing data. Typically the individual wavebands of multispectral imagery provide the variables for the clustering, but images from several dates may also be used. For example Erkkila and Kalliola (2004) applied an unsupervised classification to six summertime Landsat images of the Archipelago Sea in Finland. Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper (ETM+) blue, green and red bands from the six images were used to create an 18-band image. The unsupervised classification of this image grouped together regions with similar spectral/temporal properties in the visible region of the spectrum, identifying groups which form zones with increasing distance from coastal waters to the open sea. These zones correlated well with measurements of chlorophyll-$a$ and Secchi disk depths obtained for the region, which showed greatest chlorophyll-$a$ concentrations and lowest Secchi depths in the inner archipelago closer to land.
Figure 3.3 illustrates the application of K-means unsupervised classification on one year of MODIS monthly chlorophyll-\(\alpha\) imagery over South Australia. The classification was performed in ENVI with a predetermined 5 classes. The classification groups together pixels within the image that have similar chlorophyll-\(\alpha\) properties, thus from the resultant image we can conclude that the offshore ocean waters are very dissimilar in the MODIS imagery to the coastal regions, similar to what is observed from hierarchical cluster analysis using discrete sample data.

![Figure 3.3: Results of an unsupervised classification on 1 year of monthly MODIS chlorophyll-\(\alpha\) imagery for 2006, the colours show the grouping of regions with similar chlorophyll-\(\alpha\) characteristics.](image)

3.3.2 Discriminant Analysis

Discriminant analysis is a multivariate technique used to determine the variables responsible for the separation of objects into groups. Discriminant analysis can address a number of research questions including, but not limited to, determining whether statistically significant differences exist between two or more known groups, determining which independent variables account for the majority of the differences between groups, and establishing procedures for classifying objects into groups (Hair et al. 1995). The dependant variable may consist of just two groups, such as good or bad, or multiple groups, such as low, moderate and high pollution for example. The groups or categories must be mutually exclusive as objects may only be placed within one group. Discriminant analysis aims to determine whether these groups or classes can be separated using many independent metric variables, and to identify the variables contributing most to the classification. Discriminant analysis may be used in an explanatory sense (identifying the
variables contributing to the discrimination), and in a predictive sense (using the discriminant function to classify new samples).

The calculation of the discriminant function can be achieved via a simultaneous method, where all independent variables are considered at once, or a stepwise method, where variables are considered sequentially. The simultaneous method is used when the analyst wants to include all the independent variables and is not interested in intermediate results to determine the most discriminating variables. The stepwise method can be either forwards, starting with the best discriminating variable and including variables until all the variables which prove useful in discriminating between groups are chosen, or backwards, whereby variables are eliminated starting with the least significant first. The relative importance of each independent variable is determined by the standardised discriminant weights. The sign and magnitude of the discriminant weight assigned to each variable in the calculation of the discriminant function are used to show the discriminating power of each variable. The variables that contribute more to the discriminant function are those with the greater discriminant weights (Hair et al. 1995). The accuracy of the discriminant function may be assessed using independent data, such as a subset of the original data (the hold out group), and when used in a predictive sense, the discriminant function can be used to classify new samples. Further details of the methodology are once again beyond the scope of this review, but for more comprehensive details refer to Hand (1981), Gnanadesikan et al. (1989) and Hair et al. (1995).

Like cluster analysis, discriminant analysis has a wide range of applications in differing areas of research. However, studies involving the application of discriminant analysis to water quality are limited and are less common than cluster analysis. Similarly to cluster analysis, discriminant analysis has mostly been applied to analysis of river water quality. Singh et al. (2004) applied both a spatial and a temporal discriminant analysis to their dataset of 24 parameters collected monthly over 5 years in the Gomti River, India. For the temporal discriminant analysis the dataset was divided into three seasons: winter, summer, and monsoon. Standard, forward stepwise and backward stepwise methods were applied, with the standard and forward stepwise methods attaining an 89% accuracy using 23 and 16 parameters respectively. The backwards stepwise method however, achieved an 88% accuracy using just 5 of the independent variables. These five variables, which were therefore assumed to be the most significant parameters to discriminate between the seasons, were pH, temperature, conductivity, total alkalinity and magnesium. Singh et al.
(2004) also applied a spatial discriminant analysis to determine the variables responsible for variations between the regions of the river determined through cluster analysis. As for the temporal discriminant analysis the standard and forward stepwise methods obtained 92% accuracy using 23 and 17 parameters respectively, while the backwards stepwise method obtained 91% accuracy using just 9 parameters. The backward stepwise method showed that pH, temperature, alkalinity, calcium hardness, dissolved oxygen, biochemical oxygen demand, chloride, sulphate, and total kjeldahl nitrogen are responsible for the majority of variation between the study regions. Therefore discriminant analysis was able to contribute to a significant reduction in the multivariate dataset, while providing insight into the variables contributing to variations in water quality between periods and regions.

Discriminant analysis has also been used with remote sensing data. Andrefouet et al. (2004), for example, applied discriminant analysis to CASI hyperspectral airborne remote sensing imagery to determine the best wavelengths for discriminating between different species and communities in coral reefs in French Polynesia. Discriminant analysis highlighted the key non-redundant wavelengths capable of achieving good separation between the ecological groups. By measuring only the necessary wavelengths, rather than the full spectrum, the study was able to achieve greater spatial resolution and better signal-to-noise ratios.

A method with similarities to both cluster and discriminant analysis is a decision tree method known as classification and regression trees (CART; Breiman et al. 1984). CART uses a set of predictor variables to compute explanation schemes for a variable of interest, the target variable, which is assumed to be related to the predictor variables. Pesch et al. (In Press), for example, applied CART to predict the benthic organisms that would be expected to be found at different locations in the North Sea based upon measurements of water quality, such as bottom water salinity, temperature, phosphate and nitrate concentrations.

3.3.3 Factor Analysis and Principal Components Analysis

Factor analysis is a class of multivariate statistical methods which aim to determine the underlying structure of a multivariate dataset, to summarise and reduce the amount of data. The summary is achieved by obtaining the factors, or underlying dimensions, of the data which describe the data in terms of a smaller number of items than the original variables.
Data reduction is achieved by substituting the derived factors for the original variables in the dataset. Factor analysis has three main objectives: to identify the relationships among variables, to identify representative variables from a large set of variables, and create a new (smaller) set of variables replacing the original variables for future analyses (Hair et al. 1995). There are two methods which fall under the heading of factor analysis: common factor analysis (FA) and principal components analysis (PCA).

The purpose of common factor analysis is to describe the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called factors. Factor analysis reduces the number of variables by combining two or more variables into a single factor. The main applications of factor analysis are to reduce the number of variables, and also to detect structure in the relationships between variables. There are two forms of common factor analysis, R-type and Q-type. R-type factor analysis is used to derive the correlations between the variables whilst Q-type factor analysis is used to investigate the correlations between the objects or samples. Q-type factor analysis is similar to cluster analysis in that it groups the objects into classes, although it uses the correlations with other objects rather than the distance measures. For more details of factor analysis refer to Jackson (1983), Hair et al. (1995) and Manly (2005).

Principal Components Analysis is used to reduce the dimensionality of a dataset by explaining the variability of a large set of variables using linear combinations of the original variables, principal components. While many components are usually required to explain the total variability of a system, often the majority of the variability can be accounted for by a smaller number of principal components. Principal Components Analysis is an important technique in digital image analysis where it serves several functions. It is used to reduce the volume of large datasets, while retaining the majority of the image information; to extract a series of uncorrelated spectral components from multispectral and hyperspectral imagery; and to enhance image features for interpretation (Richardson 1993; Campbell 1996). For details regarding methods, procedures and applications of PCA refer to Jackson (1983), Richardson (1993), Hair et al. (1995), Campbell (1996), and Manly (2005).

One form of principal components analysis is empirical orthogonal functions (EOF) analysis. Although PCA and EOF are essentially the same, EOF is generally applied to a dataset consisting of a single parameter measured at different periods of time, such as a
temporal sequence of satellite images. In this situation, each image in time is considered a
different variable, and EOF uses combinations of these variables to identify the primary
modes of variability, in a similar manner that PCA uses combinations of original
parameters to identify the principal components. The aim of EOF is to simplify a complex
spatial and temporal dataset, such as remote sensing imagery, using a reduced set of
functions that represents the majority of the information of the original dataset (Vargas et
al. 2003).

Like the techniques considered earlier, FA, PCA and EOF are multivariate techniques that
are not confined to a particular field of research and are widely transferable between fields.
FA/PCA/EOF appears more widely used in the field of water quality research than either
cluster analysis or discriminant analysis, and hence has been applied more frequently to
marine and coastal water quality. Singh et al. (2004) applied both factor analysis and
principal components analysis to their water quality sampling data from the Gomti River in
India. Principal components analysis identified six principal components that accounted for
71% of the total variance in the water quality data. The first component was found to be
related to electrical conductivity, dissolved salts, alkalinity, chloride and sodium and
represented 27.9% of the total variance, while a second principal component was related to
dissolved oxygen and accounted for 17.3% of the total variance. The following four
components represented 8.5%, 7.0%, 5.5% and 4.7% of the variance respectively. Factor
analysis was applied to reduce the contribution of less significant variables from PCA.
Factor analysis revealed six factors, comprised of 14 of the original 24 parameters, which
explained 71% of the variance. The six factors from most significant to least significant
were found to be related to the mineral composition of the river water, anthropogenic
pollution sources, dissolution of soil constituents, mineral related hydrochemistry, fluoride,
and suspended sediments.

Reghunath et al. (2002) applied R-mode and Q-mode factor analysis to analyse
groundwater quality in Karnataka, India. R-mode factor analysis was used to show the
interactions between the variables, while Q-mode factor analysis reveals the relationships
between the sampling sites (Reghunath et al. 2002). Thirteen water quality parameters
were determined from 56 samples collected during May 1998. The most significant factor
derived from R-mode FA, which accounted for 45% of the variance, revealed high Na, Cl,
and electrical conductivity (EC) and moderate bicarbonate and total dissolved solids (TDS)
revealing that the electrical EC and TDS are mainly due to Na and Cl. The second factor,
representing 18% of the variability, was associated with high Ca and moderate bicarbonate, revealing the hardness of the water. The Q-mode factor analysis revealed 89.9% of the variance was explained by the first factor. The stations within this factor are located on either side of the main river, suggesting an exchange between river water and groundwater.

EOF analysis has been applied to satellite imagery in the Gulf of Cadiz, Spain. Vargas et al. (2003) used weekly composite SST imagery from NOAA-AVHRR over a 7 year period to analyse SST seasonality. Four significant modes were found, representing a total of 80.5% of the measured variance. The first mode, accounting for 60% of the variance, represented a semi-permanent north-south temperature gradient within the basin, while the second mode, accounting for 13% of the variance, accounted for the seasonal variations that occur upon this structure. Navarro and Ruiz (2006) applied EOF analysis to five years of weekly averaged SeaWiFS chlorophyll-a images. The analysis showed that the first mode, representing 20% of the variance, was related to the seasonal phytoplankton blooms, while the second mode, representing 10% of the variance, distinguished five different zones of the study region with different chlorophyll-a characteristics. Likewise Venegas et al. (2008) also used EOF analysis to interpret interannual variability in satellite-derived measurements of surface wind stress, sea surface temperature, chlorophyll-a and sea surface height off the coast of Oregon and Washington, USA.

Figure 3.4 illustrates the application of principal components analysis on the one year of monthly MODIS chlorophyll-a images of South Australia. PCA was performed in ENVI and identified a principal component that accounted for 81.6% of the variance, representing the overall high and low chlorophyll-a areas within the region. The second and third principal components identified seasonal influences within the dataset, and represented 5.6% and 4.7% of the variability respectively. The second PC appears to separate areas with relatively high chlorophyll-a in summer from those with relatively high chlorophyll-a in winter. Likewise the third PC contrasts areas with relatively high chlorophyll-a in spring from areas with relatively high chlorophyll-a in autumn. Figure 3.5 shows the contribution of each month to the first three principal components.
3.3.4 Self-organising Maps

The self-organising map (SOM) is a form of artificial neural network (ANN). ANNs are computer algorithms designed to extract interpretable patterns from large and complex datasets by simulating the information processing abilities of the brain (Dayhoff 1990). The SOM extracts patterns from large datasets by “learning” through an iterative process. In SOM, multidimensional input data are mapped onto a two-dimensional output space, thereby reducing the dimensions of the data. Details of the mathematics behind SOM can be found in Dayhoff (1990) and Kohonen (1997).

The initial step is to randomly distribute nodes onto the data space. The input data are then presented to the SOM in sequential steps. The vector of the first input sample is compared via the Euclidean distance to every node on the output space and the node that is most similar to, or the smallest Euclidean distance from, the input vector is the winning node.
and becomes the centre of an update neighbourhood. Within the update neighbourhood, all nodes and their weights are updated throughout the iterative procedure so that they converge on the input data. The winning node becomes similar to the input data and nodes compete to best represent the data. Sequentially the input vectors are added to the SOM and the SOM converges towards the input patterns. Convergence takes a large number of iterations ($10^5 - 10^6$) and it is considered achieved when the overall error or the sum of the squared Euclidean distances between all input and output data is a minimum (Richardson et al. 2002; Richardson et al. 2003).

Song et al. (2007) applied self-organising maps to classify benthic macroinvertebrate communities across differing levels of water quality and pollution in a Korean river system. Data from previous studies between 1984 and 2000 (Chon et al. 2002; cited in Song et al. 2007) were combined with data collected between 2000 and 2005. Initially the SOM was trained by arranging clusters based on the degree of pollution, then the SOM was used to map gradients in water quality based on observations of community data and derived indices.

Richardson et al. (2002) used a self-organising map to identify characteristic classes of chlorophyll-$a$ profiles and assign measured profiles to the representative classes. Fluorescence profiles, which were later calibrated to chlorophyll-$a$ concentrations, to a depth of 100 metres were collected as part of oceanographic cruises off the coast of southern Africa between 1993 and 1995. The SOM was trained to identify the underlying patterns or representative profiles or classes, and then the SOM was used to assess the frequency of occurrence of each of the classes in the profile datasets. The authors were then able to relate the seasonal patterns in the chlorophyll-$a$ profiles to physical and environmental processes such as upwelling.

Richardson et al. (2003) applied self-organising maps to identify patterns in satellite-derived SST imagery along the Atlantic coast of southern Africa. Eighteen years of weekly composite AVHRR SST images were used to extract synoptic scale SST patterns using the SOM technique, resulting in 15 maps representing the most frequent scenarios experienced in the data. The frequency of occurrence of each scenario was used to identify the most commonly occurring SST patterns. The seasonal characteristics were investigated by applying a seasonal frequency of the scenarios by month. This identified the most common patterns experienced during the different months of the year and also highlighted the intra-
month differences, with some months displaying a greater range of SST patterns than others.

Self-organising maps are also demonstrated by the case study using MODIS monthly chlorophyll-\(a\) images for 2006. The SOM was performed in ENVI using a series of IDL routines developed by Morton Canty ([http://mcanty.homepage.t-online.de/software.html](http://mcanty.homepage.t-online.de/software.html)) (Canty 2006). The SOM procedure generated 3 classification images from the 12 monthly average chlorophyll-\(a\) images for 2006 (Figure 3.6). Each of the 3 resulting classification images separates the region of South Australia into five groups based upon their MODIS chlorophyll-\(a\). The SOM assigns each pixel of the scene into a class, grouping together areas with similar chlorophyll-\(a\) properties. Figure 3.6 shows three output SOM images, each image represents a different pattern observed in the image dataset with the different shades of grey separating the classes. The SOM produces a classification similar to unsupervised classification, but instead of producing one classification image that represents the entire dataset SOM generates multiple images to represent different scenarios or patterns in the imagery.

![Figure 3.6: The three output images of SOM analysis on the 2006 monthly MODIS chlorophyll-\(a\) images. The three images represent common patterns in the chlorophyll-\(a\) imagery, while the shades of grey in each image show the grouping of similar regions.](image)

3.3.5 Semivariogram Analysis

The semivariogram is a very different technique to those described above, in that its primary objective is to obtain a measure of the variability between measurements as their spatial separation increases. Consequently the method explicitly analyses spatial relationships between measurements, whether they are discrete samples, or from pixel values in remotely sensed images. The semivariogram is a description of the spatial variance and aims to show at what distance the measurements become less similar. The semivariogram is a plot of the semivariance against the distance between the
measurements. Equation 3.1 (from Webster (1985), cited in Eleveld and van der Woerd (2006)), is used to calculate the semivariance:

$$\gamma(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \{z(x_i) - z(x_i + h)\}^2$$

(3.1)

where $\gamma$ is the semivariance, $h$ is the lag distance, $n$ is the number of observations, $z$ is the value and $x_i$ is the location. As the lag distance becomes small and approaches zero, the semivariance decreases but never reaches zero, due to an offset caused by unexplained error. The non-zero variance when the lag distance is very small is called the nugget variance. As the distance becomes large, the semivariogram levels out at a point where there is no longer any correlation between the measurements. The variance at this point is known as the sill, and the distance at which this occurs is the range. Figure 3.7 shows an idealised example of a semivariogram. After analysis of the spatial variance between samples or measurements, the resultant model may be used to interpolate and predict values in unsampled regions. More details regarding the semivariogram methods can be found in Isaaks and Srivastava (1989) and Atkinson and Lewis (2000).

Figure 3.7: An idealised example of a semivariogram adapted from Eleveld and van der Woerd (2006).

Chandrasekharan et al. (2008) used a semivariogram in their investigation into soil and groundwater water quality in India following the tsunami of 2004. The semivariogram was used to investigate the spatio-temporal variability of electrical conductivity (EC) and pH in groundwater samples collected from areas of the coast at three time periods following the tsunami and compared with results prior to the tsunami. The semivariogram determined the variance between measurements at different locations and the change in variance with changing separation. The semivariogram model was then used to interpolate EC and pH.
measurements at unmeasurable locations to show the water quality along the entire affected coastline.

Anttila et al. (2008) applied semivariogram analysis to investigate patchiness of water quality in relation to sampling methodology. The study aimed to investigate spatial variability to estimate the representativeness of discrete chlorophyll-a measurements in the Enonselka basin of Lake Vesijarvi in southern Finland. The semivariogram graphed the variance between measurements against the distance between them to determine the separation at which the measurements are no longer spatially dependant. The results identified that discrete chlorophyll-a samples taken in this study area were only valid in the immediate area of the measurement, with the representativeness of measurements decreasing to just 50% within a distance of 250 metres (Anttila et al. 2008).

Eleved and van der Woerd (2006) applied semivariogram analysis to satellite-based remote sensing imagery. SeaWiFS chlorophyll-a and suspended particulate matter (SPM) images were obtained for the North Sea to investigate variability of these parameters in the region. Semivariogram analysis was applied to both chlorophyll-a and SPM images from multiple dates to investigate the spatial variability in water quality, and the differences in the spatial patterns between seasons. The semivariograms obtained showed how the measurements of the water quality parameters relate to each other as the separation between them increases, and locates the separation at which the measurements become independent. However, the exact range where the measurements become independent is subjective and depends upon the model used to fit the variogram (Eleved and van der Woerd 2006).

The application of a semivariogram is illustrated by a set of field-based chlorophyll-a measurements collected in a coastal region of South Australia in May 2007, approximately 50 m apart along a 5 km transect. The semivariance was calculated in Microsoft Excel using the geostatistics add-in from the Soil Science department of North Dakota State University (http://www.soilsci.ndsu.nodak.edu/w188/W188LinksT.html). The resultant semivariogram depends strongly upon the minimum lag distance selected. Figure 3.8 shows the semivariogram for the field-based measurements with a minimum lag distance of 100 m. There is a large nugget effect with the field-based chlorophyll-a showing large semivariance at very small separation distances due to the highly variable nature of chlorophyll-a over small distances. The graph appears to level off at around 1200m, the
distance at which the measurements are no longer correlated. There is some inconsistency in the variance beyond this distance, although this is likely to be an artefact of the sampling with an unequal number of measurements for each separation distance.

A second illustration comes from MODIS chlorophyll-\(a\) over the same region of coastal South Australia in May 2007. In this example the chlorophyll-\(a\) values were extracted from pixels along a 40 km transect and the semivariogram calculated. Figure 3.9 shows the semivariogram from the MODIS satellite chlorophyll-\(a\) measurements using a minimum lag distance of 1000 m. The semivariogram shows a lag distance of approximately 17 km at which the measurements are no longer correlated. It can be observed that the two semivariogram images are very different, resulting from the high spatial variability in the field-based measurements over short distances and the effective averaging of chlorophyll-\(a\) in the MODIS imagery. The field-based measurements could be considered independent from each other at distances as small as 50 m, due to the large nugget offset, but the satellite based measurements are not independent of one another until a separation distance of 17 km, due to the spatial averaging and overlap of the field-of-view of the satellite sensor.

Figure 3.8: Semivariogram for field-based chlorophyll-\(a\) measurements collected 50 m apart along a transect in coastal waters of South Australia in May 2007.
Figure 3.9: Semivariogram from MODIS chlorophyll-α satellite imagery in South Australian coastal waters in May 2007.

3.3.6 Geographically Weighted Regression

Geographically weighted regression (GWR) is a statistical technique used to investigate spatial variations in the relationships between variables (Tate and Atkinson 2001). Classical regression methods are used to identify relationships between variables, but they rely on the assumption that the relationships are constant across the study area. The application of an average relationship over a region can mask interesting variations in the relationships (Tate and Atkinson 2001). GWR overcomes the problem of varying relationships by calibrating the regression model at each point with the data weighted according to how far away from the point it was recorded (Tate and Atkinson 2001). GWR is an extension of the traditional regression technique which allows local variations in variable relationships to be incorporated by estimating coefficients that are specific to each location (Brunsdon et al. 1996). GWR allows the regression model to incorporate different relationships between variables at different points in space in order to capture spatial variation (Wooldridge et al. 2006). For details of the GWR methods and mathematical models refer to Brunsdon et al. (1996) and Tate and Atkinson (2001).

Geographically weighted regression was applied by Wooldridge et al. (2006) to investigate the relationships between nutrient enriched runoff and chlorophyll-α concentrations in the Great Barrier Reef Lagoon, Queensland. It was proposed that post-European changes to land use have resulted in increased nutrient discharge to the lagoon and consequently
increased phytoplankton blooms in the lagoon. GWR was thus used to explore regional differences in the relationship between freshwater runoff from rivers and chlorophyll-\(a\) in the lagoon between the undisturbed northern catchments and the human impacted central and southern catchments. The results indicated that for regions outside the influence of runoff the chlorophyll-\(a\) concentrations were uniform, while within the mixing zone there was a spatially variable response of chlorophyll-\(a\) to the runoff.

Geographically weighted regression also has applications to remote sensing imagery (Foody 2003). Regression techniques are commonly used with remote sensing imagery to derive relationships between surface variables and remote sensing variables. However, as has been mentioned, such relationships may not be constant in space. Therefore GWR allows parameters of the regression model to vary throughout the study area, producing a more accurate relationship than a global regression model. Foody (2003) applied GWR to investigate the relationship between rainfall and remote sensing derived normalised difference vegetation index (NDVI). It was shown that a global regression between rainfall and NDVI showed a strong relationship between the variables, although GWR identified that there were local variations in the relationship. GWR therefore derived a more accurate relationship between rainfall and NDVI than did the global regression which assumed spatial consistency.

### 3.4 Overview

Cluster Analysis has been used by many authors to study the similarities in water quality measured at differing locations and is one of the most frequently applied of the techniques reviewed in this paper (Vega et al. 1998; Alberto et al. 2001; Erkkila and Kalliola 2004; Singh et al. 2004; Solidoro et al. 2004; McNeil et al. 2005; Panda et al. 2006; Shrestha and Kazama 2007). The raw measurements from multiple parameters are used to indicate how similar these locations are in terms of their water quality characteristics. The locations displaying close similarities are grouped together based upon their water quality information alone, without prior knowledge of the positions of the locations relative to each other. Widely used in river studies, cluster analysis is under-utilised in coastal and marine applications where it could easily be applied. Cluster analysis could be used to show the similarities between neighbouring embayments with differing levels of anthropogenic influence or to determine the influences upon a water body using the similarities with other water bodies to interpret the factors responsible for changes in water
quality. Hierarchical cluster analysis enables the relationships between samples to be displayed in a graphical manner, when the objective is to observe similarities among a relatively small number of samples. Non-hierarchical techniques are more applicable to larger datasets whereby the relationships between individual samples are not as important as the segregation of samples into classes. Unsupervised classification is the form of non-hierarchical cluster analysis most commonly applied to remote sensing imagery. It groups image pixels into classes, using measurements across different wavelengths or at different periods of time, producing a classification map that depicts regions of similar water quality characteristics. Although clustering of pixels is generally based on radiance recorded by different wavebands of multispectral sensors, the same technique can be applied to multi-temporal image sets.

The application of discriminant analysis to water quality studies has been limited and even more so in marine remote sensing applications, although it has been used successfully to relate vegetation cover measurements to satellite imagery in terrestrial environments, for example Lewis (1998). Studies discussed above have shown how discriminant analysis can be applied to determine the variables responsible for the variations in water quality between different seasons and different locations throughout freshwater river systems (Singh et al. 2004), and how discriminant analysis can be applied in applications of remote sensing imagery (Andrefouet et al. 2004). Despite this, discriminant analysis, although a common technique in other areas of research, has not yet been widely applied to water quality studies or studies involving remote sensing imagery.

Factor analysis and principal components analysis are both commonly applied to coastal and marine water quality datasets. Factor analysis is used to describe the relationships amongst variables and to reduce their number by combining similar variables, which represent the same underlying dimension of variability, into factors. Principal components analysis is similar, and is used to summarise the variance of the original variables in terms of a smaller number of components, formed from combinations of the original variables. PCA describes the variability in terms of the measured or observable parameters, while FA explains the variability in terms of unobservable or hidden factors, which are driven or influenced by the original parameters. While PCA may be able to assist the analyst to conclude that the nutrients nitrate, nitrite and ammonia are responsible for the majority of the water quality variability in a given area for example, the results of FA are often able to be interpreted to further lead the analyst to conclude that the water quality variations are
driven by the agricultural processes used on the neighbouring land. EOF analysis, a specialised form of PCA applied to temporal sequences of remote sensing images, is becoming more frequently applied to investigate modes of variability in ocean properties such as chlorophyll-\(a\) and sea surface temperature (Vargas et al. 2003; Navarro and Ruiz 2006; Venegas et al. 2008) and is likely to become more common as the application of remote sensing imagery grows.

The semivariogram is a technique that can be used to measure water quality variations and determine a measure of the rate of change of the variance. The semivariogram can be used to show the separation between measurements at which they become independent and are no longer related to each other, and provide a basis for interpolation between dispersed samples. The semivariogram can also be applied to temporal variability where the variance of measurements taken at different times is graphed against the time interval to show the period of time at which measurements become independent. The semivariogram is another underutilised technique that can be applied to coastal water quality studies. One specific application of the semivariogram to coastal studies may be to use the range, the separation at which measurements are independent, in the planning of future studies when determining the separation of \textit{in situ} sampling.

The geographically weighted regression technique can be used in two ways: firstly as a more accurate method of classical regression, by incorporating spatial changes in the relationships between variables into the model; secondly, to understand spatial variability, but not in the same way as other techniques. GWR assesses the changes in the way variables relate to one another at different geographic locations. Thus GWR can increase the understanding of the spatial variability of a variable by increasing the knowledge of the spatial variability of its relationship to other variables.

Water quality in coastal waters can have considerable consequences for marine ecosystems and the people dependent upon them. Monitoring techniques can provide useful data regarding water quality, but analysis is required to fully understand patterns of spatial and temporal variation. Different techniques are available and can be applied to answer a variety of questions. Techniques such as cluster analysis and discriminant analysis can be used to group regions according to their water quality characteristics and to determine parameters controlling variations in water quality. Techniques such as factor analysis, principal components analysis and empirical orthogonal function analysis can reveal the
underlying factors behind variability, identify the components responsible for the majority of variability, and illustrate the different modes of variability within complex datasets. Self-organising maps are able to extract otherwise unobservable patterns from within datasets and to show the frequency of occurrence of each pattern. The semivariogram is a method of measuring the rate of change of water quality and determining the separation at which measurements become independent. Finally, geographically weighted regression allows the changing relationships between variables to be examined.

Water quality data collection has historically involved discrete sampling, and therefore analysis of the data has been conducted by techniques that were designed to be applied to such spatially separated measurements. As we have seen, techniques like cluster analysis, discriminant analysis, factor analysis and principal components analysis are all very capable of deriving meaningful conclusions from such data. Increasingly the application of more spatially comprehensive data sources, such as remote sensing imagery, is being used for water quality monitoring (e.g. Goetz et al. 2008; Gons et al. 2008). Thus the techniques used for analysis need to evolve also. Traditional multivariate techniques designed for discrete data sampling can be transferred to remote sensing imagery, as we have seen with cluster analysis, discriminant analysis, factor analysis and principal components analysis. As the application of remote sensing imagery has become more frequent we have also seen the development of techniques designed to suit these spatially comprehensive datasets. Analyses such as self-organising maps, semivariogram analysis and geographically weighted regression, while perhaps not been designed exclusively for use with remote sensing imagery, are well suited for application to these comprehensive datasets. As the availability of global and regional data from many satellites spanning longer time frames increases, we are likely to see an increase in the application of remote sensing imagery for water quality monitoring and the development of an improved spatial representation of marine ecosystem health indicators. Arising from this we will also see an increase in the application of techniques, such as those mentioned above, in the analysis of these water quality datasets.