ISDN 1039-3218

# Data Storage and Use of GIS for the Port Phillip Bay Environmental Study

Robert Molloy[1] and Colin Arrowsmith[2]

[1]CSIRO, Environmental Projects Office

[2]RMIT, Department of Land Information

Technical Report No.45

Port Phillip Bay Environmental Study
CSIRO Environmental Projects Office
Canberra, ACT, Australia

GPO Box 1666, Canberra, ACT 2601, Australia

October 1998

# EXECUTIVE SUMMARY

The Port Phillip Bay Environmental Study conducted by CSIRO has produced an up-to-date description of the state of the Bay together with an explanation of the physical, chemical and ecological systems operating. In doing so a comprehensive collection of data has been compiled.

Adopting a spatial or geographic information system (GIS) to store, manipulate and access data collected as part of the Study, enabled that data to be mapped and interpolated for further scientific analyses. Spatial correlations between data sets could also be investigated.

Each of the Study tasks utilised varying sampling techniques, but generally each provided data that was spatially referenced (location identified by latitude and longitude or easting and northing). Spatial interpolation techniques were then used to produce area coverages on a baywide scale from the point sampling. The report outlines some of the limitations in the use of this interpolated data.

The handling of the various data sets created many problems. Each of the scientific tasks had unique data requirements and utilised numerous proprietary computer programs. This led to the adoption of dual data storage system, comprising the raw data collected in the field stored as ASCII text files, and the data that has been manipulated and uploaded into the GIS.

In this report the role of the Database Manager, the nature of the data collected and the use of the GIS for the Study is discussed, along with issues related to digital data storage.

# TABLE OF CONTENTS

**Page No.**

# 1. Introduction

The Port Phillip Bay Environmental Study (PPBES) conducted by CSIRO, commenced in 1992 and was completed in 1996. It produced an up-to-date description of the state of the Bay together with an explanation of the physical, chemical and biological systems operating.

The major output of the study was the development of integrated models of the Bay. The models provide a conceptual basis for synthesis and integration of the data sets and process knowledge provided by the diverse field surveys and experiments which made up the major part of the study. They also provide a basis for connecting the past and future management actions to the Bay's environmental health, allowing the development and analysis of management strategies and scenarios. In this role, models serve as a focus for the interface between environmental studies and management (page 158, Harris et al., 1996).

To enable the development of these models an extensive database of field measurements was obtained. Measurements such as tidal heights, wind, solar radiation and river flows, were used to drive the models. Other data included the results of a wide range of field sampling surveys carried out in the course of developing an understanding of the chemical and biological processes that occur within the Bay.

A direct result of this study has been the development and consolidation of an extensive database of marine and physical data for Port Phillip Bay.

## 1.1. Project Design

The CSIRO Project Office presented the Project Design to Melbourne Water in January 1992. The management-based Study objectives were designed to establish the essential scientific questions and the knowledge required to address the assumed nutrient and toxicant issues associated with Port Phillip Bay. Tasks were defined which would yield scientific answers, recognising that past and present work must be reviewed and integrated into the new research (CSIRO, 1992).

The design team believed that it was imperative to create a centralised comprehensive database for the storage of all scientific data collected for the study. This would include historical data previously stored on various systems throughout the state. Appointment of a permanent database manager was believed to be critical for the success of the project.

At the completion of the study the database would be able to be passed back to the user agencies to assist in the continued management of Port Phillip Bay.

# 2. DATA MANAGEMENT

The Project Design broke the study into four principle areas of interest.

1. Physical processes
2. Nutrient inputs and processes
3. Toxicant inputs and levels
4. Ecological processes

The nature of the data collected was reflected in the scope of the design. The result being a range of data sets that varied considerably in size, format, scale, complexity and quality.

## 2.1. Hierarchical file management system

The study was designed as a series of related tasks that were loosely defined by their principal interest - Physics, Nutrients, Toxicants, Ecology or General. General tasks were those that covered the mapping tasks, and other activities related to the organisation and management of the Study. These categories were then used in the task numbering system. For example the task looking at the nutrient concentrations in waters across the Bay, was referred to as Task N2.1. A full list of the tasks and their numbers is provided in Appendix 1.

Initially datasets were referred to by their Task Number and stored in appropriately named places in the file management system. This naming system however lost value as the study progressed and the data sets needed to be accessed based on their content not the task. The alternative was to create a file management system that adopted the principle interest and subdirectories based on the topic. Figure 1 provides a simplified version of this directory structure.

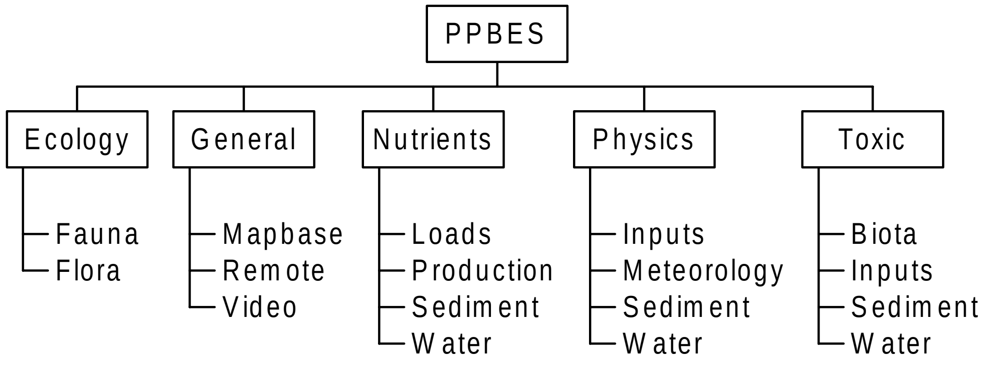


**Figure 1: Simplified picture of the directory structure**

As the project and database developed it transpired that there were two basic types of data. The original (sampled) data, and data that had been loaded and manipulated within the Geographic Information System (GIS).

A limiting factor of storing data within a GIS is that uncertainty is propagated and transformed each time a conceptual or physical surface is generated (Bedard, 1987). Data may be manipulated using the surface creation techniques within a GIS at the discretion of the user. The scientist who provides the field sampling data relinquishes control over the data integrity.

## 2.2. Data storage format

After discussion with members of the Technical Group supporting the Study, it was decided that data integrity could be maintained if the datasets were self-documenting. This could be achieved by including the supporting data descriptions or *metadata*, with the actual data. This information would include the source of the data, who collected it and when, along with a brief description of the methods used to collect or generate the data. Also most importantly the fieldnames, units and means for indicating missing data.

The design of the study had led to the separate tasks being conducted by numerous research and data collection agencies. This meant that there was no standard software being used, nor any standard format for data. The software and formats used were dependent on the nature of the task and the capabilities of the organisation completing the work. In this environment is became apparent that the simplest way to maintain usability of the data was to where possible leave the data as simple text files (ASCII), in a comma delimited format. Most software packages can import and export comma delimited files with minimal editing required.

This approach to data storage was not unique. The Konza Prairie Long term Ecological Research Program, which commenced in the United States in 1981, has utilised a similar approach to data storage (Briggs and Su, 1994). For this project all data, except for GIS coverages, satellite images and word-processing files were converted to ASCII format.

In order to separate the data from the header information, all lines within the header are preceded with a crosshatch. This allows the header information to be stripped from the file manually or by automated processes. The column descriptors (fieldnames etc.) are preceded by double crosshatches. Simple search function programs can be written to search all file headers and identify those files containing the fieldnames of interest.

As a further step in maintaining data integrity and usability a common list of field names was created for the Project.

Below is a sample header from one of the files from the PPBES database.

```
# PPBES_FILENAME: chla9095.dat
# PPBES_DIRECTORY: /ppbes/nutrient/water/
# SUPPLIED_BY: Andy Longmore, MAFRI Queenscliff, 29/2/96
#
# DESCRIPTION:
# Chlorophyll a results from MAFRI fixed sites water quality surveys
#
#     Site  East        North       Location    Date Range
#     1     305728      5752630     Blairgowrie 7/3/90 to 29/3/95
#     4     332219      5775051     Wooley Reef 7/3/90 to 29/3/95
#     6     322655      5798418     Sandringham 7/3/90 to 29/3/95
#     9     316444      5807219     Williamstown 7/3/90 to 29/3/95
#     11    298904      5792772     Werribee    7/3/90 to 29/3/95
#     13    284579      5775137     Clifton Springs 7/3/90 to 29/3/95
#     N2    284398      5780663     Pt Wilson   3/3/94 to 29/3/95
#     N3    309808      5774746     Portarlington 3/3/94 to 29/3/95
#
## COLUMNS 3
##
## COLUMN1.name site
## COLUMN1.long_name sample site identifier
## COLUMN1.units 1
## COLUMN1.missing_value -9999
##
## COLUMN2.name date
## COLUMN2.long_name date sample taken
## COLUMN2.units text (yyyy-mm-dd)
## COLUMN2.missing_value xxxx
##
## COLUMN3.name chla
## COLUMN3.long_name chlorophyll 'a'
## COLUMN3.units ug-1L
## COLUMN3.missing_value -9999
##
####
1,1990-02-21,0.610
4,1990-02-21,0.640
6,1990-02-21,0.540
11,1990-02-21,0.300
13,1990-02-21,0.770
1,1990-03-07,0.990
11,1990-03-07,0.560
…
```

## 2.3. Entity relationship model

In the early stages of the Study an Entity-Relationship (E-R) Model was developed to assist in conceptualising a centralised database (refer to Appendix 2). This model was to be used to load data into the Relational Database Management System (RDBMS) within the GIS.

Unfortunately with such a broad range of datasets, the entity relationship model became quite flat and relatively unworkable. Key relationships between many of the task datasets were non-existent. As a result of this the database manager did not pursue with the task of loading all data into a single RDMBS for the Study.

## 2.4. Sourcing data

Existing data fell into general classes - digital or paper based. Data in both classes was either in the form of maps or tables.

The standard approach was to obtain where possible field data in its purest form. If a report contained both a table of data and a map depicting that data, then the table of data was selected for inclusion in the Study database.

This approach to data acquisition was to maintain data integrity and avoid propagating errors created through transferring interpolated data as depicted on maps. Maps while often seen as ideal means for presenting a great amount of data in a relatively simple method, are in there very nature a means for both generalising data and exaggerating through symbolisation the importance (accuracy and confidence) of the original field data.

In the early stages of the study, base information was required for general presentation purposes. Some topographic map data was purchased from the State, but much of the thematic data sets were digitised in-house, from maps provided by the supporting agencies.

One of these data sets was the Sediments coverage, which was digitised from a map (A1 size) provided by the Port of Melbourne Authority. The map in itself was a compilation of several sediment surveys that had been done since 1965. The sampling density varied considerably over the map. Some areas were interpolated based on a sampling density of 1 sample per 10 $km^2$, while in other areas it was 3 samples per $km^2$.

As the map was believed to be a good and fair indication of seabed sediment distribution within Port Phillip Bay, it was suitable for presentation of general information only and not for qualitative reference.

As the Study progressed preference was given to the generation of maps from source data rather than from digitising existing maps.

### 2.4.1. Purchasing data

User agencies involved in the Study were able to provide much of the base data in digital format at no cost. Other organisations such as the Bureau of Meteorology, Managed Information Technology Services (MITS) Ltd and the National Tidal Facility, which act as data brokers, charged a minimal service fee for the provision of data.

The Study had been designed such that individual tasks were sub-contracted to various research agencies. As a condition of the sub-contracts all data collected as part of the task was to be provided to the Study Office for inclusion in the database. In return the Study Office provided the contractors with data necessary to complete their task. This exchanging of data was the main function of the Database Manager during the field study period (1993-95).

In summary our approach was to avoid purchasing data unless it was considered essential for the carrying out of scientific tasks. However it should be realised that just about every task paid for as part of the study was a data collection exercise.

## 2.5. Model Data

The data generated by the numerous models developed as part of the Study was not envisaged to be stored as part of the database. The Study database was primarily a storage system for field data.

The size and number of files generated by the hydrodynamic model and the ecological model reduce the practicality of storage on normal computer systems.

The benefit of storing modelled data is also limited. The data itself is an indication of what is likely to occur based on the knowledge built into the process equations and algorithms. Therefore it is considered that storing actual field data is of more benefit for future users than the actual model output.

# 3. SYSTEM SELECTION AND MAINTENANCE

At the commencement of the study in 1992, the Database Manager carried out an investigation of available data storage and geographic information system (GIS) software and hardware. Based on the functional requirements, and the capacity needs of the study, a decision was made to purchase Arc/Info GIS software from Environmental Systems Research Institute (ESRI) Australia.

## 3.1. Functional requirements

The functional requirements for the system were identified as being:

- run an efficient computer operating system capable of handling large files and complex data manipulation;
- ability to handle spatially related data sets in both raster and vector format;
- extensive spatial modelling capabilities, including vector and raster analysis, surface generation and map composition;
- a relational database management (RDBM) system compatible;
- broad range of import/export functions.
- high speed and high resolution graphics workstation, with peripheral components - digitising table, colour plotter and tape drive;
- appropriate data backup and storage capabilities;
- provide overall reliability with strong local support for the software and hardware;
- able to increase in data storage capacity as required by the Study.

## 3.2. Capacity Planning

Prior to the purchase of the system, data storage requirements were scaled. Preliminary calculations estimated that 500 Mbytes would be initially required. The system was therefore purchased with a 1.2 Gb external hard disk.

As a requirement of the Study, historical data sets were downloaded onto the project database for inclusion in the initial analyses. As the study progressed it was necessary to upgrade the capacity of the system. A 4.2 Gb SCSI external hard disk was acquired in late 1994. The additional capacity was required for the considerable amount of data generated by the GIS, and for storage of remote sensing data.

## 3.3. Selected Software

A floating single user licence for Arc/Info, with system extensions TIN and GRID was purchased from ESRI Australia in October 1992. A single copy of PC ArcView was also purchased at this time. ESRI Australia acted as the Prime Vendors.

The Arc/Info software included a relational database, which meant a third party database (for example Oracle or Ingres) was not required.

The recommended platform for the Arc/Info system was a Sun Sparcstation 10 from Sun Microsystems, which has a proprietary Unix operating system.

The initial operating system was SunOS 4, but this was upgraded to Solaris 2.5 in June 1996. The upgraded operating system enabled the hydrodynamic and ecological models to be installed and run for demonstration purposes.

## 3.4. Hardware Configuration

The initial purchase in October 1992 was:

- Sun Sparcstation 10, desktop workstation, model 30 with a 19" colour monitor, 48 Mb RAM, 424 Mb internal hard disk and 1.44 Mb 3.5" internal floppy drive
- Sun 1.2 Gb external disk drive, 1/2 inch cartridge tape drive (150 Mb capacity), and a CDRom reader
- GTCO Digitising table (A0 size)
- HP Draftmaster RX colour pen plotter

Later purchases included:

- HP Paintjet XL300 postscript colour printer (A3 and A4 size paper)
- Sun 4.2 Gb external multi disk drive
- Sun 8 mm exabyte tape drive (2.3 Gb capacity)
- Internal hard disk replaced with a 1.04 Gb disk, as part of an operating system upgrade in August 1996.

## 3.5. Maintenance and Reliability

Annual maintenance from ESRI Australia was taken out for the Arc/Info software and extensions. Maintenance included regular software upgrades and application support over the phone.

Similarly annual maintenance was taken out for the workstation with Sun Microsystems. Software upgrades and telephone support was also included. However the system upgrades and installation of new components was provided as a fee for service.

The system has been extremely reliable with only one major breakdown occurring over the entire study period. One of the external hard disks malfunctioned during the relocation of the study office from West Melbourne to Collingwood. It was replaced as part of the maintenance contract, and the data was restored from the backup tape.

The system is relatively simple to maintain. The data files were routinely backed up to tape. Initially when only a cartridge tape was used for backup, an incremental backup strategy was adopted (ie. Only those files accessed since the last major backup were saved to file). Table 1 documents backup procedures. Level 5 backups were an incremental backup, and would only backup files modified or added to the system since the previous Level 5 backup. Level 3 backups would take place at the end of each week, therefore enabling an entire week's files to be saved. Once a month two Level 0 (ie. an entire dump of all files) would be undertaken, one tape was stored locally, the other stored off-site.

**Table 1: Backup schedule.**

| | Monday | Tuesday | Wednesday | Thursday | Friday |
|---|---|---|---|---|---|
| Week 1 | Level 5<br>Tape B | Level 5<br>Tape C | Level 5<br>Tape D | Level 5<br>Tape E | Level 3<br>Tape A |
| Week 2 | Level 5<br>Tape B | Level 5<br>Tape C | Level 5<br>Tape D | Level 5<br>Tape E | Level 3<br>Tape F |
| Week 3 | Level 5<br>Tape B | Level 5<br>Tape C | Level 5<br>Tape D | Level 5<br>Tape E | Level 3<br>Tape G |
| Week 4 | Level 5<br>Tape B | Level 5<br>Tape C | Level 5<br>Tape D | Level 5<br>Tape E | Level 3<br>Tape H |

Once a month a Level 0 (ie. an entire backup) would take place.

However after purchasing the exabyte tape drive, the incremental backup strategy was replaced by a whole of system backup strategy. That is the entire database was backed up on a regular basis.

## 3.6. System and Data Security

The Solaris operating system has standard password controlled access and a multi-level privilege access system. All general operations were conducted at base operating level. System maintenance and data transfer were conducted at the system manager (root) level. Passwords were kept secure and changed regularly.

For security, a monthly backup of the system was made and stored offsite.

Regular transfer of the full datasets to the CSIRO Division of Marine Research based in Hobart for use by the team developing the models, meant that another complete data set was stored off-site.

# 4. DATA SETS

## 4.1. Reference Data

### 4.1.1. Phase One data

The ‘Phase One Study”, 1968 to 1971, was carried out by the then Melbourne and Metropolitan Board of Works (MMBW) and the Fisheries and Wildlife Department (MMBW/FWD 1973). The study covered the physics, chemistry and biology of the Bay, and in doing so provided a base-line data set on which to understand the existing impacts of the Bay’s catchment. The data collected for the Phase One study provided an excellent reference point for documented change over time.

Copies of the Phase One report can be found in most major libraries in Australia, and Melbourne Water still have a supply of copies. Unfortunately the scientific data was not as easy to access. Much of the data collected for the Study had been converted to digital data on the computers of the time and had since been lost as a complete and accessible data set.

Melbourne Water through their information technology supplier, Managed Information Technology (MITS) Ltd, were able to translate the water sampling data into a suitable readable format, from archived magnetic tapes at a reasonable cost.

The data from the Phase One benthic surveys had been maintained by the Museum of Victoria, and was provided to the Study as part of Task E5.

The location of sampling sites was documented in the report. For the Bay-wide surveys, a regular grid of eighty-six stations was established to cover the whole of the Bay. The stations were at the intersections of two sets of parallel lines running 30° E and 30° W of true North. The origin of the grid (Station 917) was at 145° E 38° S. The sites were generated into map coverages by direct input of the coordinates. The attributes for each station were loaded from the digital files recovered from the Melbourne Water computer archives.

### 4.1.2. Historical water quality data

Some historical data was obtained from past publications by researchers such as D. J. Rochford who published in the Memoirs of the National Museum Victoria (Rochford 1966).

Rochford measured water temperature, salinity, oxygen, phosphate and nitrate in water samples collected at five stations within the Bay between 1947 and 1952. Station sites were digitised from an A4 sized map provided in the published papers. The values measured were manually keyed in for each station. Standard equations were used to convert the values into current units.

There are two issues faced when using historical data. The first is whether the original analytical methods used for measuring the various parameters are known. If the method is known, then its relevance/accuracy to current methods can be made. Without this knowledge comparison with current data may be improper. The other issue is related to positional accuracy and the method used for recording these positions. If longitudes and latitudes are used then question of which spheroid becomes relevant. If a Cartesian coordinate system is used, then details of the system used must be known.

If the sampling points have been recorded on a map, then the points can be digitised. However this will reduce the positional accuracy depending on the scale of the map.

For the Rochford data, the positional accuracy would have been about ± 5 km. For water sampling data, this accuracy would be appropriate. The water is a constantly moving mass, so future reference to the same location is only a general requirement.

With regard to analytical methods the Rochford data would be reasonably accurate for temperature, salinity, oxygen, and nitrate. It is now known that the techniques used for measuring phosphate would have provided values greater than the actual.

### 4.1.3. 1:50000 Strategy Planning Series

The MMBW 1:50000 Strategy Planning Series Maps which were converted to a digital format by Melbourne Water in the late 1980's were used as the base map product for the GIS. This data set provided coverages for all topographic features in the surrounding area of the Bay. It included: roads (major and minor), cultural features (jetties, railways, bridges, etc), hydrographic features (coastlines, rivers, lakes etc) and contours of the land.

The data was downloaded from Melbourne Water's Intergraph Mapping System in Microstation design file (*.DGN) format, which could be imported into the Arc/Info system. This transfer resulted in some loss of attribute information, but with no significant impact on the spatial accuracy.

Property map base data from Melbourne Waters Facility Information System was obtained, but this was not of a suitable scale for the general mapping requirements of the Study. Basically it was too detailed for Baywide mapping. It also only covered the metropolitan area, which was only part of the Study's area of interest.

## 4.2. Physical Data

### 4.2.1. Bathymetry

Bathymetric data was one of the data sets necessary to develop a complete physical computer model of Port Phillip Bay. The Royal Australian Naval Hydrographic Service supplied digital bathymetric data for Port Phillip Bay and an adjoining part of

Bass Strait, at no cost. However the conditions of supply were that the data be used for the purposes defined and that it not be forwarded on to other parties.

There were 3419 soundings provided. The unit of measure was metres, however several datums were used - mixture of chart datum and mean sea level depths. The positions were provided in geographic coordinates (latitude and longitude). The horizontal datum was the Australian Geographic Datum based on the 1966 Australian National Spheroid (AGD66). The data was also provided in the GeoVision GINA format.

The data was reformatted into Arc/Info format by staff from the Australian Survey and Land Information Group (AUSLIG). It was then loaded onto the PPBES Arc/Info GIS. The soundings were converted from chart datum to Australian Height Datum (AHD) and then projected from AGD66 onto the Australian Map Grid (AMG) Zone 55.

The soundings were used to generate a point coverage in Arc/Info, and from the point coverage a Triangulated Irregular Network (TIN) was generated. The TIN was then used to generate contours at fixed height intervals (Figure 2). These contour coverages have since been used in numerous maps as a base feature.

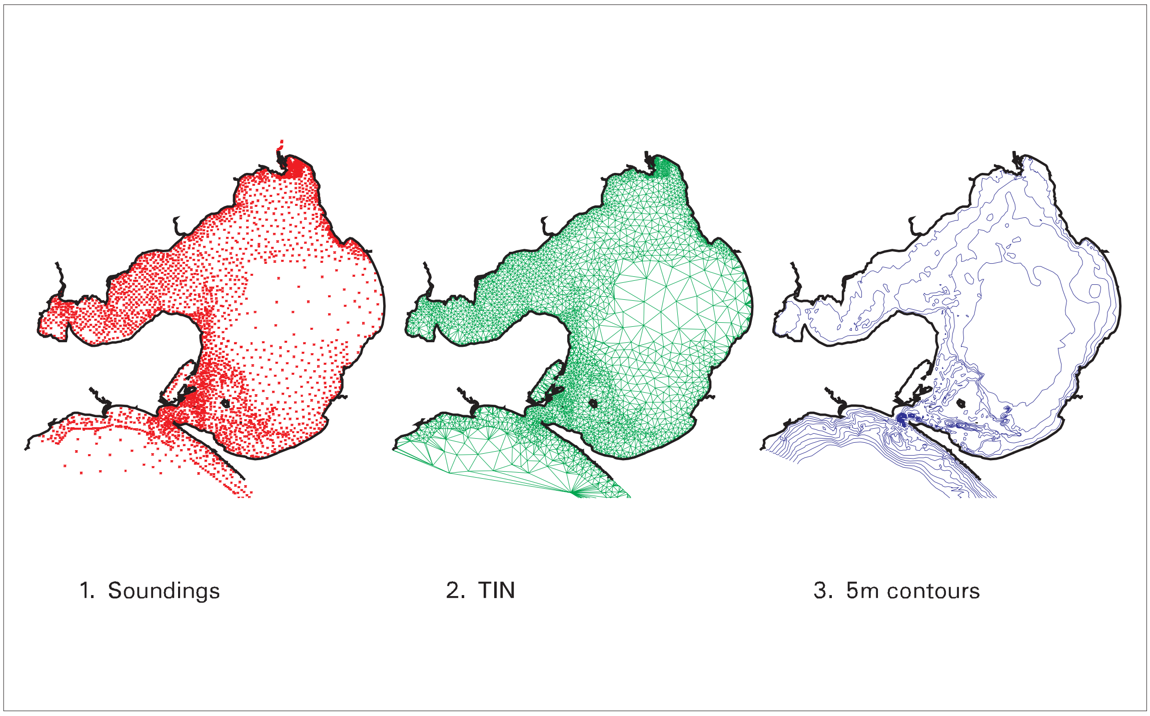


**Figure 2. – Three stage development of the bathymetry (soundings, TIN, and contours).**

### 4.2.2. Sea-level observations and tidal constituents

The National Tidal Facility and the Port of Melbourne Authority supplied sea-level observations from tide gauges around the Bay, and outside the entrance to the Bay. The data included hourly sea level observations filtered from six-minute seaframe sea levels for 8 locations for 1994 and 1995. Point Lonsdale hourly sea level observations were obtained for the period 1985 to 1995 and Lorne hourly observations for the period 1993 to 1995.

Data from Lorne were used to specify sea level in Bass Strait for the open boundary of the physical models of the Bay. For the hindcast modelling, the Lorne data was not available prior to 1993, so data from Point Lonsdale were used.

The tidal constituents for each location were also provided along with plots of the residual currents, which were used to indicate the quality of the data.

The National Tidal Facility provided the data for a small fee and subject to their standard conditions. These conditions included acknowledgement in any publications or report that utilised the data supplied, and that the data was supplied in good faith and is believed to be correct.

The Port Authority, being one of the managing agencies, provided data at no charge.

### 4.2.3. Currents

Current measurements included those obtained from moored current meters, coastal radar (COSRAD) and a vessel mounted acoustic doppler current profiler (ADCP). As well as these, currents were also measured in conjunction with the deployment of instrument mounted tripods (known as SEDCAM) on the seabed at eight sites around the Bay.

***Moored current meters***

Moored current meters were deployed at five sites in the Bay for two periods, each of about six weeks. Two current meters were located at each site, at about one quarter and three quarters of the water depth. The sites were all located in the main body of the Bay. The velocity and direction were measured continuously with a mean over a 10 minute period being recorded.

The first current meter deployment covered the period from 21 April 1994 to 2 June 1994, with meters located near the northern edge of the Great Sands, the Bay centre and Martha Point, Point Cook, and Red Bluff. Data were recovered from all meters during the first deployment period.

The second deployment covered the period from 6 September 1994 to 14 October 1994, with sites as above except that a site near Schnapper Point replaced the Great

Sands site. Data were not recovered from the lower meters at the central Bay and Red Bluff sites during the second period, due to instrument failure.

***COSRAD high frequency radar***

Continuous coastal radar measurements were made of two areas within the southern part of the Bay to determine surface water movement and velocities, using a COSRAD high frequency radar system developed by the Physics Department of the James Cook University. Measurements were taken in two phases over a period of two weeks in April and May 1994. The first phase covered the area just to the north of the Sands. The second covered the Sands region itself.

The data resulting from these deployments provided hourly surface currents over the coverage areas, with a spatial resolution of about 3km. At each point, the magnitude and direction of the current is provided. Figure 3, generated using the GIS shows two sets of current vectors 6 hours apart. The data is stored as both ASCII text files and as a point coverage within the GIS. Further information on COSRAD is available from PPBES Technical Report No. 19.

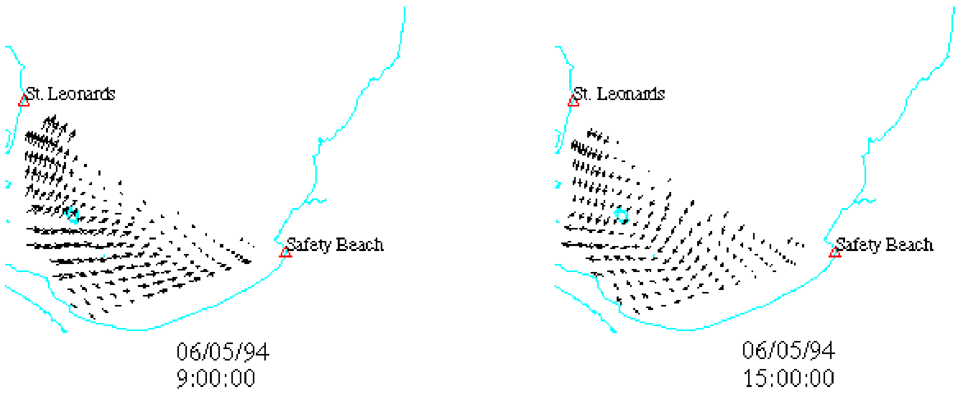


**Figure 3. Plots of current vectors for one of the areas surveyed using COSRAD. (Source: PPBES GIS).**

***Acoustic doppler current profiler (ADCP)***

Acoustic doppler current profiler (ADCP) underway data was collected in the Great Sands region over the period 7th to 12th May 1994. This work was done in conjunction with intensive conductivity, temperature and density (CTD) surveys and also at the same time as the COSRAD project. Due to the complex nature of this data, no attempt was made to convert it to the standard format adopted by the Study.

The ADCP measured the vertical structure of flows through the major channels in the Great Sands and hence gave a measure of the transport of water through these channels.

### 4.2.4. Meteorological data

Meteorological data sets collected for the Study included:

1. Air temperature and cloud coverage
2. Evaporation (gridded estimates covering the Bay.)
3. Humidity.
4. Light (total solar radiation and photosynthetic active radiation (PAR)).
5. Rainfall data. (rainfall measurements surrounding the Bay and gridded estimates covering the Bay.)
6. Wind
7. Meteorological data from weather station installed on Hovell Pile as part of Study.

***Air temperature and cloud coverage***

Air temperature data was obtained from the Victorian Environmental Protection Authority (EPA) for stations at Frankston and Point Wilson (Sept 1992 – Jan 1995). Air temperature was also obtained from the Bureau of Meteorology for Melbourne, Moorabbin and Laverton for the period 1985 - 1995. This data was provided with cloud coverage as well.

***Evaporation***

Hourly evaporation data was provided by the EPA, covering the years 1985 to 1994, and the first half of 1995. The data was located on a 10 km resolution grid, calculated from wind speed, air and water temperature and relative humidity. The model output was not provided in a format suitable for storage within the PPBES database.

***Humidity***

Humidity data were obtained from EPA stations at Frankston and Point Wilson (Sept 1992 – Dec 1994), and from measurements at Hovell Pile (March 1994 to May 1995).

***Light***

Solar radiation data comprising total solar radiation (TSR) and photosynthetic active radiation (PAR) were obtained from EPA sites at Brighton (TSR and PAR), Dandenong (TSR) and Geelong (PAR) and from the Bureau of Meteorology site at Laverton (TSR).

***Rainfall***

Rainfall data was obtained from a number of sources during the study. For the purposes of modelling, rainfall data was required on a uniform grid (10 km) covering the Bay. Due to the lack of data from gauges located on the Bay (only Hovell Pile, 1994-1995), rainfall data were interpolated from sites surrounding the Bay. Such data was obtained from both the EPA, and the Bureau of Meteorology. These two data sets are very similar, with small differences probably due to different choices of sites and interpolation techniques.

As well, rainfall over the Bay was estimated from weather radar data. This RAPIC radar derived data covered the period from March 1994 to February 1995.

### *Wind*

The aim of the study was to obtain wind data suitable for physical modelling for the period from 1985 to 1995. Winds over the Bay were interpolated from observations at three sites within the Bay (Hovell Pile, Point Wilson and Fawkner Beacon) and one site on land (Point Cook - a well exposed site with a long record of observations).

Correlations between points on the Bay and Point Cook were calculated for periods when all sites were available, and these correlations were used to calculate Bay-wide winds when only Point Cook data were available. The end results of this process were continuous hourly wind data on a 10 km resolution grid covering the entire Bay.

### *Hovell Pile weather station*

An automatic weather station was installed on Hovell Pile (navigational pylon located about 4 km offshore from Rosebud) for the period March 1994 to May 1995. Measurements included air temperature, humidity, precipitation, and PAR. Data was provided as smoothed hourly values based on readings taken every 10 minutes. In conjunction with these measurements, wind speed and wind direction (vector averages), water temperature and cumulative monthly rainfall were obtained from the PMA instruments also installed on the pylon.

Unfortunately the station was subject to fouling by seabirds despite the bird guards. This meant regular service was required throughout the period. Twice the rainfall filter was blocked and water had accumulated in the collecting funnel, suggesting that the data was invalid for some period prior to inspection. Consequently rainfall data in these periods was omitted from the data files. Other issues apart from the gaps in data were the concerns with the accuracy of data such as relative humidity, which was considered questionable and recommended not to be used.

An issue with any remote recording device, is data integrity. Especially when the performance of the instrument can be impaired but not totally stopped. For example if an optical sensor is only partially covered by bird excrement it will still appear to have been collecting data in the normal manner. Only when the instrument is checked, will the operator discover the problem, and then they are unlikely to be able to identify from what point in time the data is corrupted.

With the SEDCAM deployments a video camera was focussed on the optical sensors in an effort to identify when the sensor was inhibited. The other means for ensuring the quality of data is to have the instrument checked on a frequent basis. Then if the instrument is inhibited the data for the preceding period can be identified as corrupted.

### 4.2.5. Input flow data

For modelling purposes it was necessary to acquire data on the continuous inputs from point sources and natural watercourses for both the duration of the study and for a period of at least 10 years prior to the Study.

***Natural Water Courses***

An essential data set was flow data for the Yarra River. The Yarra has numerous flow recording stations, but the critical station for this study was the one immediately upstream from the Chandler Highway overpass at Kew. This point on the Yarra is above Dights Falls, which is the point where tidal effects are eliminated.

The flow recorder operates on a continuous basis, which means instantaneous flow measurements can be taken at any given time interval. Melbourne Water records this data as part of their telemetry network.

A ten year (1985 - 95) data set containing mean daily flow, measured in cubic metres per second ($m^2s^{-1}$) was provided. However for use in some of the Yarra River water quality studies flow data was obtained as mean flow per hour. Ten year data sets were also obtained where possible for other rivers, streams and drains. These data sets were either obtained from Melbourne Water or from Thiess Environmental Services.

The main issues in obtaining flow data are the selection of the timing interval and the location of the flow recorder. Generally flows were measured at points far enough upstream to remove the influence of tidal flows. As part of the Study models were developed to calculate loads to the Bay based on continuous upstream flow measurement and limited downstream sampling. (See Technical Report No. 33 for calculation of input loads for the Yarra River and Mordialloc/Patterson River System).

Often reference was made to the annual flow rates rather than daily flow rates. This resulted in several data sets being derived from the primary data. For consistency flowdata has been stored as mean daily flow in cubic meters per second ($m^3$ $s^{-1}$).

***Other discharges***

Flow data from Western Treatment Plant (WTP) for the period 1/1/1980 to the 1/1/1996 was obtained. The flow is provided as ML/day and is based on weekly level readings in the outlet drains. When this data is combined with the results of routine sampling for nutrients, heavy metals and other toxicants, load data can also be generated. Further information on outputs from WTP is provided in Technical Report No.15.

Several industrial plants draw water from rivers that discharge to the Bay. After use in the industrial process (generally cooling) the water is discharged to either the river or the Bay. An example being Newport Power Station. When the generator is operating the discharge is on average $18.5m^3/s$. However, this flow is included in calculations of flow for the Yarra River.

Similarly the flow from the Shell Refinery on Corio Bay was not obtained, because the water that is discharged is drawn from the Bay. Therefore the net load to the Bay is nil.

### 4.2.6. Temperature and Salinity Conductivity Temperature and Density (CTD)

Conductivity, temperature and density (CTD) surveys of the Bay were undertaken on four occasions - two in May 1994 and two in October 1994. As well as these surveys, more detailed surveys were carried out in the areas of the Sands, the Yarra plume and the region offshore from the Western Treatment Plant (Werribee). The CTD probe enables salinity, temperature and dissolved oxygen to be measured. Calibration of the instrument is provided through collection of reference samples, which are analysed, for salinity concentration under laboratory conditions. This work was reported on in Technical Report No. 34.

Salinity and temperature was also measured as part of the water quality surveys (Task N2.1). For this work the CTD probe was used in two ways. One for creating profiles at fixed sites and the other for continuous measurement as the boat traversed a path around the Bay. For the profiles the boat would remain stationary as the probe was lowered over the side, thus providing salinity and temperature against depth. For the underway analysis, the probe was placed in a container of water, which was continually filled by water being pumped from just below the surface. This enabled continuous water sampling as the boat moved along the set path.

In all there were 25 Baywide cruises which were carried out monthly between May 1993 and April 1995. There were also three special surveys done in the area offshore from WTP and Hobsons Bay. Details of this work are reported in Technical Report No. 24.

In addition to the work done as part of Task N2.1, data was also made available from MAFRI's five year program (1990-95), of fortnightly fixed site surface water quality sampling.

### 4.2.7. Sediment

***Sediment Characteristics***

Core samples were collected from the seabed at 80 sites throughout the Bay, and analysed for a variety of physical properties. These included: grain size, water content, void ratio, porosity, permeability and viscosity. Further details are provided in Technical Report No. 23.

***SEDCAM Deployments***

The Victorian Institute of Marine Science (now operating as part of the Marine and Freshwater Resources Institute) made use of their self designed and built SEDCAM instrument frame for the measurement of turbidity, currents, wave motion, and light levels at eight sites around the Bay. Bed sensors were also placed adjacent to the frame for the measurement of bed erosion and deposition levels. Deployments were for periods of 30 to 50 days, between June 1994 and September 1995. Further details are provided in Technical Report No. 28.

## 4.3. Nutrient data

### 4.3.1. Nutrient Sources

Nutrient inputs into the Bay come from Western Treatment Plant (WTP) (formerly known as the Werribee Treatment Complex), major and minor watercourses, groundwater and the atmosphere.

As part of the Study, WTP discharge records were obtained for the period 1980 to 1989. This data enabled annual loads of various nutrients and toxicants to be calculated (Refer Technical Report No. 15.). Melbourne Water measures flow and load on a weekly basis. This information is reported annually to the EPA as part of their discharge licence requirements. Other information on the discharges was provided in the form of Melbourne Water technical reports and previous studies.

Loads from the Yarra/Maribyrnong Rivers and from the Patterson/Mordialloc system were also calculated (Refer Technical Report No. 33). Information on other surface inputs was provided in Technical Report No. 12.

Information on the loads from groundwater and the atmosphere are provided in Technical Reports No. 13 and 5, respectively.

### 4.3.2. Water quality

As part of Task N2.1, water quality was measured as part of fixed site water quality monitoring program, and as part of underway sampling surveys. Both tasks were carried out by VFRI (now MAFRI), and were reported in Technical Report No. 24.

The fixed site program, was part of an ongoing series of programs that had been nearly continuous since the Phase One study. The number and location of sites within each program varied, along with the sampling frequency, which was either fortnightly or monthly. Variation of the program included taking profiles at each site (ie. water samples at various depths between the surface and bottom. Data from most of these programs is provided in the Study database, but is also available from MAFRI and the EPA.

For the study, a major variation to the fixed site program was the use of an automated flow analysis technique which enabled sampling to be carried out along a transect. The parameters sampled included ammonium, nitrate, nitrite, phosphate, silicate, fluorescence (chlorophyll a is derived from fluorescence) and the physical parameters salinity, temperature and dissolved oxygen. The location of the boat was recorded by use of Global Positioning System (GPS) (non- differential). The automatic analysis was logged against time. This then allowed the water quality to be related to boat position by use of time as the related feature.

The underway transects were carried out on 25 occasions at about monthly intervals between May 1993 and April 1995. The adjusted transect data provided water quality at intervals of 1 minute, which equated to about a distance of 100m.

In addition to the Baywide surveys, three smaller surveys were carried out. One in Hobsons Bay, and the other two offshore from Werribee.

The GIS was used to create interpolated spatial surfaces from the transect data. Surface snapshots (coverages) were created for each monthly survey for each parameter. These snapshots were then combined using techniques within the GRID module of the GIS to create average value maps, and also maps indicating the range of values that occurred throughout the course of the 25 surveys.

A further use of the interpolated surface coverages was the calculation of Baywide nutrient pool sizes. The volumes were calculated by multiplying the surface coverage by the depth coverage.

### 4.3.3. Production rates

#### *Microzooplankton Grazing Rates*

Between October 1994 and September 1995, on a monthly basis, samples of water were collected from mid-depth at four coastal and two midbay sites. Estimates of microzooplankton grazing rates were made for each sample. Further information is provided in Technical Report No. 31.

#### *Phytoplankton productivity*

Phytoplankton in three size fractions - pico (0.2 to 2 microns), nano (2 to 20 microns) and micro (> 20 microns) were collected from nine stations in Port Phillip Bay from August, 1993, to April, 1995, every four weeks. Collections were made at surface, mid-depth and bottom. Additional collections were also made to examine small-scale temporal and spatial variations in biomass and productivity.

The chlorophyll *a* content of each sample was determined by filtration, extraction and either spectrophotometry or fluorometry or both. The productivity (determination of carbon fixation rates) of each sample was measured by incubation with $H^{14}CO_3$ and subsequent scintillation counting.

Further information is provided in Technical Report No. 35.

#### *Microphytobenthos (benthic microalgal biomass)*

Various tasks and data collection exercises were carried out in order to determine the spatial and temporal variation in microphytobenthos biomass (benthic microalgal biomass) and primary productivity within the Bay.

Sediment associated plant pigment and associated data were determined at 65 subtidal sites throughout a six week period commencing in March 1994. Additional cores were collected at each of the sites for grain size analysis and determination of total organic matter. The water depth at each site was recorded with an acoustic echo sounder to the nearest 0.2 m. It was observed that on occasion the depth recorded was greater than that shown on the chart by over two metres (Brett Light, personal communication). The location of each sample site was recorded by Global Positioning System (GPS) incorporating differential correction (± 2.5 m).

To examine spatial variation in sediment associated chlorophyll *a* distribution, 72 sediment cores were collected in May 1994 from each of two sites (Werribee and Indented Heads) at a water depth of about 1 m.

Another section of this work was to determine the seasonal variation in microphytobenthos populations, their photosynthetic characteristics and in situ productivity. Such measurements would allow estimation of the overall contribution of microphytobenthos to primary productivity and, indirectly, to nutrient cycling in the Bay. Three sites were used, these being Indented Head, Carrum and Werribee. They were visited monthly throughout the period September 1994 to September 1995. Three water depths were examined at each site. Sediments from four cores were extruded and the 0-5 mm uppermost strata used to determine chlorophyll and pheophytin content.

The GIS was used to generate Baywide maps for each of the parameters measured. A further extension of the work was to calculate biomass volumes for the bay using the grid analysis functions within Arc/Info.

In a separate program conducted at one site (Werribee deep), PAR sensors and data loggers, incorporating wiper units to inhibit sensor fouling, were moored for approximately five week periods in winter 1995 and in summer 1996, to obtain high resolution data on water column attenuation. This was then used to calculate comparable *in situ* primary production for the two periods.

Further information is provided in Technical Report No. 30.

### 4.3.4. Sediment Quality

Nutrient fluxes from the sediments of Port Phillip Bay were estimated from the results of sediment coring and benthic chamber deployments between July 1993 and January 1996. Sediment coring estimated fluxes caused by diffusive processes only whereas benthic chambers measured fluxes caused by diffusion, advection and bio-irrigation. Benthic-associated primary production was found at three sites sampled. Chamber Work Sediment sampling, coring and chamber deployment and analysis of samples. Further information is provided in Technical Report No. 26.

#### *Biomarkers*

This study of organic biomarkers for sewage, primary production and bacteria in the sediments of Port Phillip Bay was designed to measure the relative and absolute contribution of these groups to the organic matter present. Forty-six sediment samples were collected from the Bay in June 1993 and analysed for biomarkers. In April 1994 a further 9 sediment samples and eight water-column samples from selected rivers, creeks and drains were collected to help gain a better understanding of inputs into Port Phillip Bay. Further information is provided in Technical Report No. 17.

## 4.4. Toxicants data

### 4.4.1. Loads

Loads of toxicants (metals and organics) were obtained for major freshwater inputs such as WTP, the Yarra and other minor rivers and streams. Information on these loads is provided in Technical Reports No.s 12, 15, and 33. The data obtained was derived from previous studies conducted by the EPA, Melbourne Water and other Drainage Authorities.

Depositional loads from the atmosphere were not measured as part of the Study, and subsequently know load data is stored on the database. However reference was made to earlier studies carried out by the EPA. Further information on atmospheric loads is provided in Technical Report No. 5.

### 4.4.2. Biota

Task T8 of the Study was devised to assess whether there had been any change in the concentrations of toxicants in various organisms (plants and animals) since previous studies and how the toxicant concentrations in commercially and recreationally important fish species compared with food quality standards. While the emphasis for this study was to be on the significance to health of consumers of seafood (CSIRO 1992), the data could also be used for a general assessment of the pollution status of the Bay.

Samples of common mussel (*Mytilus edulis planulatus* Lamarck); sand flathead (*Platycephalus bassensis*), and the dominant benthic flora species (one only at each site, except seagrasses) were collected and analysed from each of six sites within the Bay. Sites included: Geelong Arm; Hobsons Bay; Werribee (within 500m of shore); Mordialloc (within 500 m of shore); St Leonards (within 500 m of shore), and the southern part of the Bay between Mud Island and Martha Point. Further information is provided in Technical Report No. 22.

### 4.4.3. Water Quality

As part of the EPA's fixed site water quality monitoring program, a suite of metals and organics are sampled on a routine basis at six sites within the Bay. This program commenced in 1984 to ensure that the provisions of the State Environment Protection Plan (SEPP) were being met and to provide baseline data so that future changes in water quality could be assessed. Data from earlier programs conducted by the EPA dates back to 1975.

As part of Task T2, a Bay-wide survey of dissolved and suspended particulate metals and organics was carried out in 1993, covering 25 sites concentrated in the northern part of the Bay. In addition to the Bay-wide survey, a rain event survey within the plume of the Yarra River was conducted. Further information on both of these surveys is provided in Technical Report No. 18 and in EPA Publication No. 479.

### 4.4.4. Sediment

As part Task T7, sediment samples were obtained from 42 widely-distributed sites within the Bay, with three replicate sediment samples collected from each of 33 sites within the Bay and from nine widely-spaced input streams and drains. The sites within the Bay were primarily located in grid patterns near the WTP, Hobsons Bay and the centre of the Bay. Corio Bay was not sampled intensively because this had been done in recent studies. Surface sediments (0-10 cm) were collected at all sites, but at four sites, deeper cores were obtained. The samples were analysed for the metal concentrations and organics. Further information on both of these surveys is provided in MSL Report No. 211 (1994) and in EPA Publication No. 454.

## 4.5. Ecology Data

### 4.5.1. Fauna

***Benthic communities***

The objective of the benthic survey work for the Study was to assess change in benthic communities during the period 1969 to 1995. Data from three previous surveys was used in the comparison with the current survey. The data is stored and maintained by the Museum of Victoria.

Collection methods were the same in all surveys, each using Smith-McIntyre grabs deployed from 15–18 m research vessels. In the 1960s and 1970s navigation was by dead reckoning or horizontal triangulation. More recently, satellite navigation was used. Samples were washed in the field over 1.0 mm sieves, fixed in 5% formalin and sorted in the laboratory under stereomicroscopes. All specimens were transferred to 70% alcohol after identification for storage in the Museum of Victoria collections where most are now registered.

**Phase One (1969 – 73)**: A regular grid of 86 stations at 5 km intervals was established to cover the whole of the Bay. Five 0.1 $m^2$ samples were taken at one time from each station; collections were made irregularly over 3¼ years.

**Phase Two (1973 – 75)**: None of the surveys of the Phase One study took any account of temporal variability in community composition and structure. Quarterly sampling over three years at three stations (muddy sediments) redressed this shortcoming. The three stations were in similar environments in three different parts of Port Phillip Bay: in Corio Bay and in the centre of the Bay (both on predominantly silt-clay sediments, but differing in depth); and off Martha Point (on a sand-silt-clay sediment).

**Scallop Dredging Effects Survey (1991)**: The Victorian Fisheries Research Institute conducted a study of the effects of scallop dredging on benthic communities. Three locations in 12–15 m depth (off St Leonards, Portarlington and Dromana) were sampled intensively in an experimental design that employed dredged and control plots.

**Phase Three Survey (1994–1995):** Fourteen stations were located within each of the three strata (eastern sandy sediments, western sandy sediments, muddy central basin plus Corio Bay). Within each stratum, stations were depth stratified, 7 and 12 m depths each with seven stations in the shallower sandy zones, and 17 and 22 m depths each with seven stations in the deeper muddy zones. Two grab samples were taken at each station, but only a single grab was sorted from most stations. Twenty-three of the 42 stations were placed at the location of Phase One stations. Five replicate grab samples were taken from each of the three muddy Phase Two stations and at the Portarlington site of the 1991 study, thus repeating the sampling design of those studies. In all, 58 samples were taken in October 1994; 60 in April 1995; and 50 in October 1995.

Further information is provided in Technical Report No. 29.

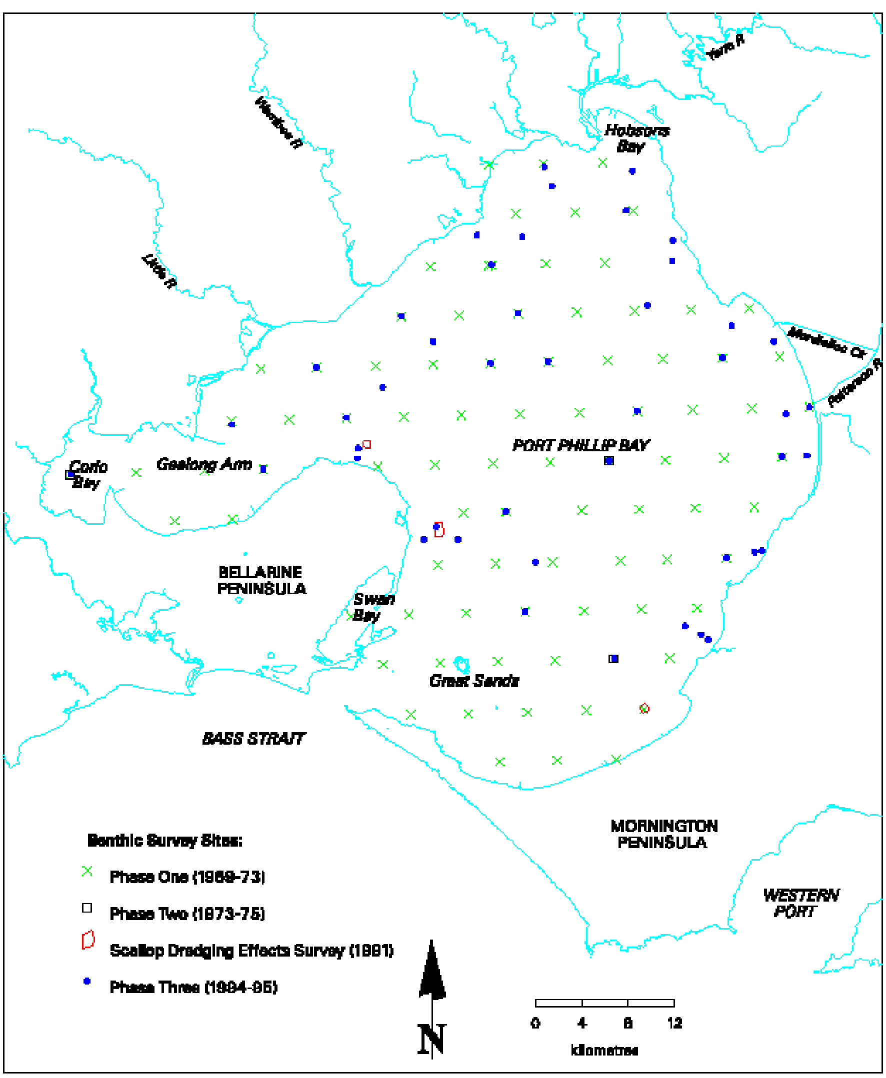


**Figure 4. Map of benthic infauna sampling sites. (Map produced using the GIS, stored as an encapsulated postscript file then converted to a TIF format file.)**

### *Fish*

The structure of fish communities in the Bay was analysed by trawling at 22 depth-stratified stations between 1990 and 1994. Trawling was undertaken at depths of 7, 12, 17 and 22 m. As part of the same task the diets of 35 species of fish were recorded in the four main depth regions of the Bay.

Changes in abundance over a longer period was assessed using data from similar surveys carried out in 1972-75 and 1990-91.

The full data set from this task and earlier studies was not loaded onto the Study database, but is available from MAFRI. Further information is provided in Technical Reports No. 21 and 36.

### *Shell Fish*

Various data sets on the abundance and size of scallops, derived from dredge surveys in 1982, 1989, 1990, 1991, 1992, 1993 and 1994, was provided to the Study by MAFRI. However the investigation of shellfish were not considered to be necessary in meeting the objectives of the Study.

### *Sabella*

Data on the distribution and abundance of the European Fan worm *Sabella spallanzanii* was not collected as part of the Study, however reports prepared by MAFRI from their 1994, 1995 and 1996 dive surveys were made available to the Study. The full data set from this task was not loaded onto the Study database, but is available from MAFRI.

### *Birds and Mammals*

In the early stages of the study it was decided that the study of birds or marine mammals was not a necessary component in meeting the objectives of the study. The preference was to select indicators that were more easily monitored and for which appropriate historical data sets were available (Harris et al. 1996, p152).

## 4.5.2. Flora

### *Phytoplankton*

Task E10.4 provided a comprehensive description of the composition, distribution and abundance of phytoplankton in Port Phillip Bay. Water samples were collected from seven sites (Figure 5) every two weeks from March 1990 to February 1995, and chlorophyll *a* and phytoplankton cell concentrations were determined from 876 and 884 samples respectively. Cell concentrations were calculated for all nano- and microplankton species or taxa (the size fractions 2-20 μm and > 20 μm respectively), but only for those picoplankton (size range < 2 μm) identifiable under a light microscope.

The full data set from this task was not loaded onto the Study database, but is available from MAFRI. Further information is provided in Technical Reports No. 27 and 40.

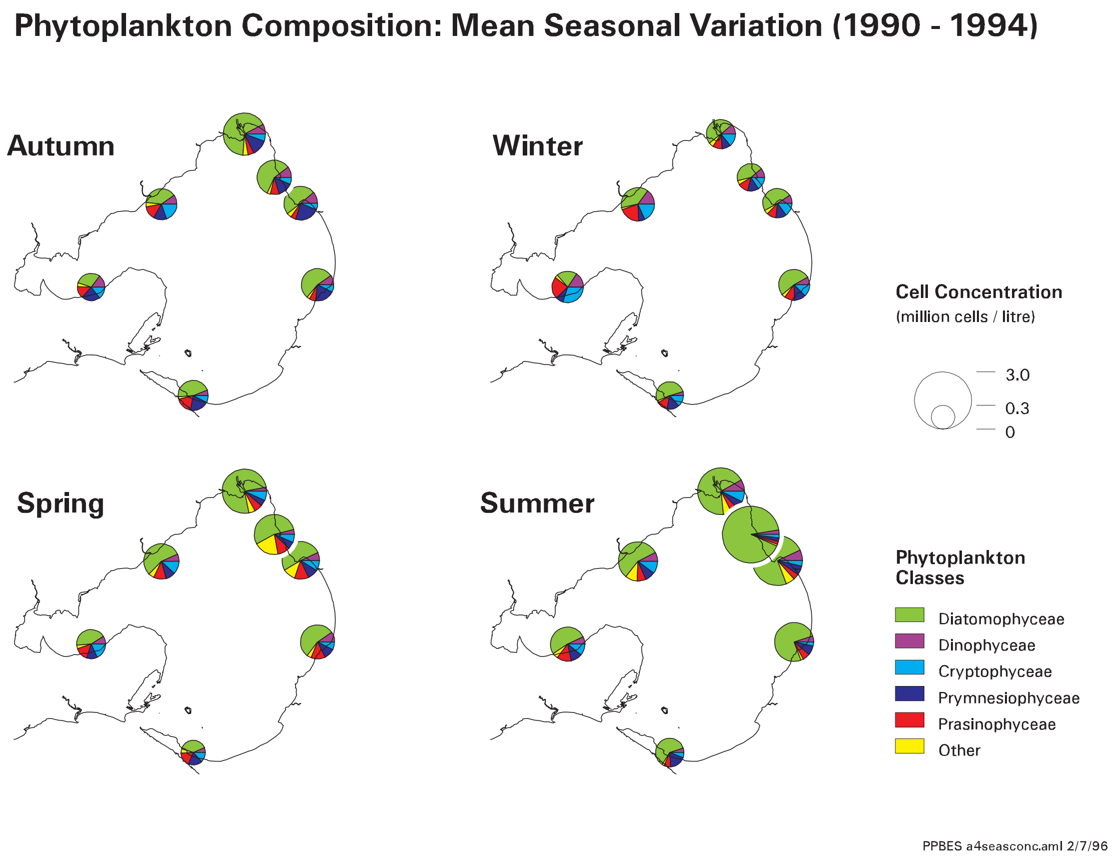


**Figure 5. Map of phytoplankton sampling sites with seasonal concentrations for the dominant species. (Map produced using the GIS, and stored as an encapsulated postscript file (EPS)).**

### *Benthic vegetation*

A pilot study determined that remote sensing of benthic vegetation using airborne multispectral scanners was capable of addressing the issues of spatial resolution, accuracy and the ability to distinguish between different types of submerged vegetation. Task G2.2 was subsequently commissioned to produce habitat maps of the Bay showing the distribution of macrophytes to a depth of at least 10 m using the Compact Airborne Spectrographic Imager (CASI).

A series of CASI flights over the Bay were conducted in March and April 1994. Some problems with atmospheric and in-water optical models due to variable environmental conditions within the Bay, precluded the identification of all features mapped in the pilot study. Although not all benthic features could be reliably identified from the survey, classes based on spectral characteristics were generated. A series of spectral class maps were produced showing significant patterns that may correspond to benthic features. However, due to time overruns, these classifications were not sufficiently verified against ground data, and some spectral classifications were inconsistent for the same area in images where overlap occurred. Thus, these maps should be considered only as interim "benthic habitat" maps and should not be used as a definite baseline on which to assess future changes in benthic vegetation. Opportunities exist for further data processing and verification by interested agencies in conjunction with CSIRO. For future reference, remote sensing should be considered for shallow water benthic mapping, provided suitable ground truthing exercises are carried out simultaneously.

To assist with mapping benthic vegetation, video surveys were undertaken at 29 locations in the Bay. The survey path was an asterisk-shaped design, 1 km in diameter. Macroalgal coverage was assessed at 20 m intervals along each transect. Maps of macroalgal cover were then interpolated from this data using kriging techniques. Estimates of the reliability of the resulting maps could be generated using coefficients of variation. The location of sites was obtained by use of differential GPS. Copies of these videos are held by MAFRI and the CSIRO Environmental Projects Office in Canberra.

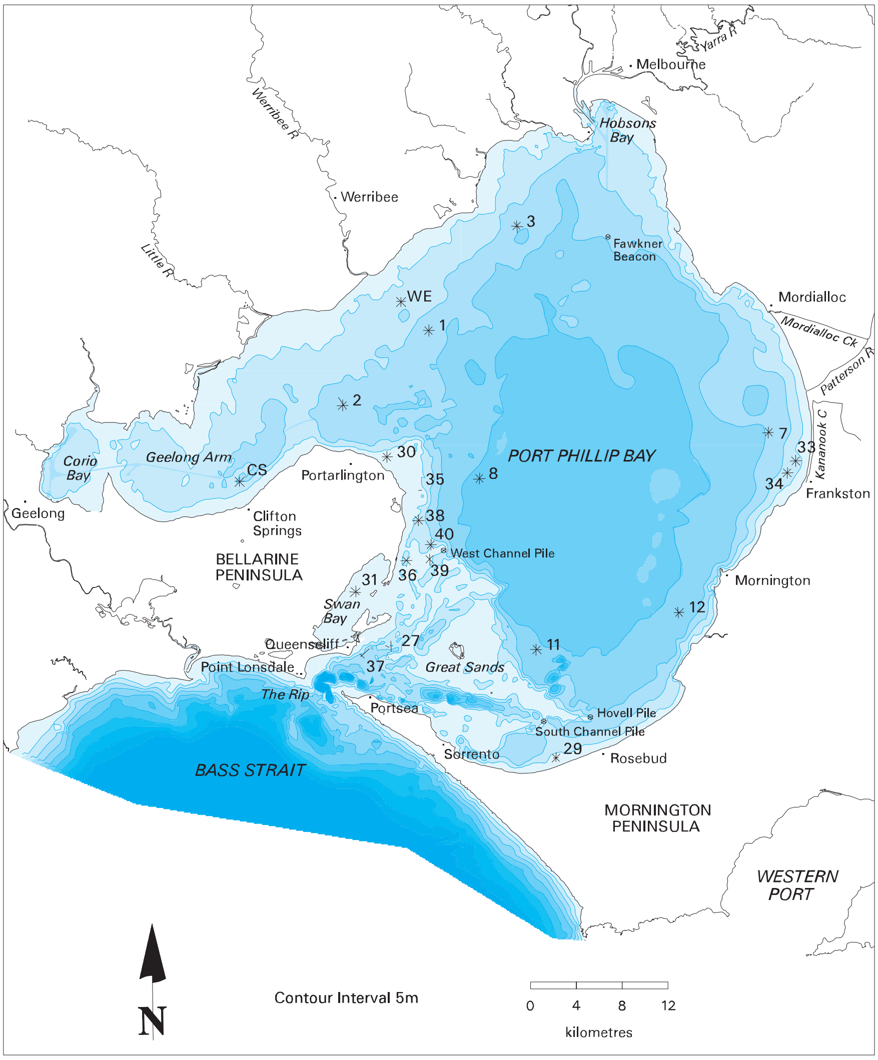


**Figure 6. Map of benthic video survey transects. Incomplete asterisks indicate that the survey was only partially completed. (Map produced using the GIS, and stored as an encapsulated postscript file (EPS)).**

### *Benthic macroalgae*

As part of Task N2.3 an investigation of benthic macroalgae was carried out by CEE Pty Ltd, to determine the spatial and temporal variation in taxonomic composition and biomass of major macroalgal assemblages in the north part of the Bay, and to document the nitrogen content of abundant macroalgae. In addition to measurement of biomass and nutrient content, the relationship between macroalgal biomass and seabed cover was examined. The standing crop and associated nutrient pool of macroalgae in the north of Port Phillip Bay was estimated.

Macroalgal assemblages were surveyed at five locations on four occasions at 4, 6, 8 and 10 m depths from 1995 to 1996. The five locations were Kirk Point and Wedge Point adjacent to the Western Treatment Plant at Werribee, Altona Bay southwest of the mouth of the Yarra River, Green Point at Brighton, and Aspendale two kilometres southeast of Mordialloc Creek.

At each site, a 100m transect was deployed on the seabed along the depth contour (parallel to the shore), and a video record of the macroalgae along the transect line was taken by divers. Point-intercept estimates of algal cover were made using a 3 m long pole marked with 15 points at 20 cm intervals. The pole was placed transversely across the transect at 5m intervals. The algal group or animal under each point was recorded.

A 1 $m^2$ quadrat frame was placed beside the transect line at 10 m intervals, and photographed. All algae within each quadrat were harvested and placed in mesh bags for subsequent weighing, taxonomic examination and nutrient content analysis.

The 35 mm slides of photographed quadrats were provided to CSIRO at the completion of the task. These slides were forwarded to the records officer at Melbourne Water's Western Treatment Plant for storage and future reference. CEE Pty Ltd holds the video records of each dive. Further information is provided in Technical Report No. 32.

# 5. USE OF GIS

## 5.1. Creating coverages

Most data was supplied to the Project Office in digital format. The occasional historic data set (for example sediments throughout the Bay, temporal sea grass movement and some of the Phase One Study sample stations) could only be obtained in hardcopy format. This necessitated the digitising of these hardcopy maps, transformation onto AMG coordinates, and then topological editing linework via ArcEdit.

Once spatial data were correctly registered (on AMG), and linework edited, polygonal data were labelled and point features (such as sample locations) were tagged so that attribute information could be entered into the database and linked to the graphical elements.

Where data was provided in ASCII format (such as the nutrient transect data), each record in the file (ie row) was assigned a unique identifying number. The file was then split into two, one containing the identifying number and latitude and longitude, the other also containing the identifying number along with the relevant nutrient and physical attributes. The file containing the coordinate pairs was then used to generate a spatial coverage and the attribute data uploaded into a data file. These two files were then joined within Arc/Info.

## 5.2. Programming with Arc/Info

Arc/Info GIS provides command level processing. However many routine tasks can be automated, and menus developed for novice or high level uses through the development of programs written in the Arc/Info Macro Language (AML$^{TM}$). AML is a procedural (or interpreted) language (translates and acts on each command line within the program). AML programs enable interactive menus, combinations of Arc/Info tools, and new commands to be developed. Because the intent of the PPBES database was not for open access to large numbers of users, menus development was kept to a minimum. However a number of routine tasks particularly relating to plan generation, graphing of data, and transformation of coordinate data were automated using AML.

## 5.3. Surface Creation

Surfaces were either created based on grids (also known as lattices) or triangulated irregular networks (TIN).

Arc/Info allows the creation of two types of surfaces – integer or floating point grids. We generally used floating point grids, where the value of the centre of the cell is calculated, and values for surrounding points within that cell are interpolated from the central points of the surrounding cells (neighbors).

The GRID module within Arc/Info provides four surface generating functions - Inverse distance weighting (IDW), Trend surface interpolation, Kriging, and Spline. For the study we made extensive use of IDW and Kriging.

Simply put, a grid of fixed resolution (eg. cells 1 km x 1 km) is placed over the survey area. The sample points are used to generate a value for the cells they appear in. The values for the remaining cells are generated based on the known values and the application of specific mathematical formula based upon the known values and the nature of variation between points.

The accuracy of the surface is primarily dependent on the number and location of sample points. However the varying techniques for establishing the values between points can also affect the accuracy of the generated surface.

Increasing the number of sample points generally increases the accuracy of the generated surface. However, the location of sample points may be just as important. Ideally the sample points should be spaced such that they indicate any ridges or valleys in the surface. To achieve this sampling must occur either side of the change point, and at the change point if its location is relevant. Unfortunately this is good in theory, but not as easy in the field. For example when taking grab samples from a boat, you are unlikely to see any difference between one sample point and the next. Also the number of sample points is usually related to the size of the budget.

***Inverse Distance Weighting (IDW)***

Inverse Distance Weighted interpolation determines cell values using a linearly weighted combination of a set of sample points. The weight is a function of inverse distance. Put simply the closer a cell is to a known point, the greater the influence on generating the value of the cell. IDW is an averaging function, so unless ridges and valleys are depicted in the sampling design, they will not be generated by the IDW function. IDW is also unable to account for noise in the input data, unlike some of the more complex surface generators, which actually create a modelled surface based upon the points provided. In summary IDW is a relatively simple method of interpretation, and surfaces generated using it, should be interpreted with caution.

Due to its simplicity and ease of operation, IDW was used extensively throughout the study. One example was the surface water quality maps generated from data obtained from MAFRI’s underway water analysis system. This system produced a water quality measurement (ammonia, nitrate, nitrite, phosphate, silicate, chlorophyll a, salinity, temperature and dissolved oxygen) every minute as the boat moved along a transect. This equated to a value every 200 metres. The location was referenced using GPS.

In order to visualise the changing water quality across the Bay, and to calculate approximate nutrient bay-wide pool sizes, the data was processed using IDW to create surface coverages. A hypothetical grid (cells 250m x 250m) was laid across the Bay. At the centre of each grid cell a search was made for valid data within a specified radius (this defaulted to the nearest three points when no valid points were found in the specified radius). Data from all valid points within each search are weighted by the inverse of their distance from the central point to produce an interpolated value for the cell central point. The horizontal scale of discrimination can be varied arbitrarily, but using a search radius of 5km to produce interpolations on the 250m grid provided the researchers sufficient detail and suitable accuracy to make some broad assessments of the variation of water quality across the Bay. A further extension was to multiply the surface grid by the water volume (equivalent size grid generated from bathymetry) to obtain a nutrient poolsize for the Bay.

***Kriging***

Kriging is based on the regionalised variable theory that assumes that spatial variation of the attribute is statistically homogeneous throughout the surface. Kriging estimates the covariance of measurements between locations as a function of their separation and orientation. This characterisation of spatial dependency is then used to determine an ‘optimal’ set of weights applied to the sample data in the neighbourhood of an unsampled location (A further discussion of Kriging is provided in Technical Report No. 42). Kriging offers the use of several mathematical functions to model the spatial variation in the measured attribute. For reliable results the operator is required to graph the semi-variance against the distance between the pairs of sample points (a separate routine within Arc/Info), and assess the benefits of varying mathematical models.

In summary Kriging offers superior surface creation to IDW, in that it has the ability to allow for noise within the data, and that mathematical models of best fit can be chosen. However its disadvantage is its complexity in selecting the appropriate mathematical model and settings.

***Spline***

A surface created using the Spline function within the GRID module of Arc/Info uses a two-dimensional minimum curvature interpolation resulting in a smooth surface that passes exactly through the input points. The routine allows for weight function to control the curvature of the output surface. The higher the weight the coarser the surface.

***Triangulated Irregular Network (TIN)***

The Triangulated Irregular Network (TIN) module within Arc/Info was another tool used to create surface coverages from point data. TIN creates a digital elevation model by joining points of known value and location, via a network of irregular triangles. The line between two points can be interpolated with reference to the value at each end. From the TIN, contours (lines of equal value) can be generated to provide depiction of the surface. An additional function of the program, is the ability to shade each triangle based on its aspect and slope. A TIN was used to generate the bathymetry for the Bay.

# 6. COMPLETING THE PROJECT

## 6.1. Meta Data

Throughout this report, stress has been placed on how the various data sets were acquired, stored, managed and distributed to the various scientists. However, possibly more important than the data itself, is the data about that data, the meta data. Data regarding detection limits, persons or agencies collecting the data, units used, methods of analyses used, original sources of data, and spatial, temporal and attribute resolution can all be considered to be meta data. It is imperative that such data be supplied with the original data sets in order to provide end users the necessary information in order for them to ascertain what limitations and therefore what applications can be expected for those data sets.

## 6.2. Use of CDRom

In order to preserve the Study data collection as an indefinite data record, multiple copies of the data, will be supplied to the user agencies on CDRom. The data to be stored will generally be the raw data that was collected as part of the field program for the Study.

Large data sets such as the CASI data or output from the models will not be included. Data of this nature is impractical to include as it is not supported by available technical reference.

Other data not to be included will be data that was purchased for use as part of the Study and would be in breach of copyright to distribute. Data such as the tidal data supplied by the National Tidal Facility and the map base data supplied by MITS fits into this latter category.

## 6.3. Conclusion

The activities of the Database Manager were critical in completing the Study. As the Study progressed, the objective of building a centralised comprehensive database became less important as emphasis was placed on facilitating the transfer of data between researchers and data collection agencies.

In theory an accessible database would ensure data could be transferred on a needs basis and at the inception of the researcher. However this accessibility, which for the Study was limited to the CSIRO modellers proved to be inefficient. The researchers, who had supplied the original data at interim periods throughout their programs, were frequently updating (amending or deleting invalid results). It was therefore essential that these revisions were likewise made to the Study database, and passed on to the end users. If the system had had uncontrolled access, it would have been impossible to know who had actually used the data and in what capacity.

Another of the key roles of the Database Manager was to search for data within organisations that was not freely available. In general only data that was the basis of published reports was known about, and subsequently requested by researchers. In order to get copies of relevant unpublished data the Database Manager had to develop a strong communication network with local research organisations along with developing an understanding of the needs of researchers involved with the Study.

The Port Phillip Bay Environmental Study has demonstrated quite clearly, that one of the key factors to the success of the research program was having the services of a dedicated Database Manager. The use of a Geographic Information System to present, analyse and manipulate the numerous spatially related data sets proved to be another essential factor in the Study's success.

## 7. ACKNOWLEDGEMENT

We would like to acknowledge the support of the PPBES Technical Group, who provided advise and support in defining the requirements and format of the data required for the Study. We also acknowledge the contributors of data, namely the researchers who carried out the scientific data collection, and for the supply of related data that had been acquired previously by their organisations.

Melbourne Water Corporation through the Port Phillip Bay Environmental Study provided funding for this work.

A special mention goes to a number of Cartography students from the RMIT Land Information Department who were employed as vacational scholars to assist in the data management and use of the GIS for the Study.

# 8. REFERENCES

Bedard, Y. (1987). Uncertainties in land information system databases, Proceedings, Auto-Carto, 8, 175-84.

Briggs, J.M. and Su, H. (1994). Development and refinement of the Konza Prairie LTER Research Information Management Program. In 'Environmental Information Management and Analysis: Ecosystem to Global Scales'. (Eds W.K. Michener, J.W. Brunt and S.G. Strafford.) pp87-100. (Taylor and Francis: London).

CSIRO 1992 Port Phillip Bay Environment Study: Project Design. A report to Melbourne Water by CSIRO INRE Project Office, CSIRO Australia

Harris, G., Batley, G., Fox, D., Hall, D., Jernakoff, P., Molloy, R., Murray, A., Newell, B., Parslow, J., Skyring, G. and Walker, S. (1996) Port Phillip Bay Environmental Study Final Report, CSIRO, Canberra, Australia

MMBW/FWD. (1973). Environmental study of Port Phillip Bay: Report on Phase One 1968-1971. Technical report. Melbourne and Metropolitan Board of Works and Fisheries and Wildlife Department of Victoria.

# 9. APPENDICES

## Appendix 1. PPBES Tasks and nature of data collected

| Task number and name | | | Nature of data collected |
|---|---|---|---|
| **PHYSICAL PROCESSES** | | | |
| P1A | Measurement of basic physical parameters | | sea level observations, currents, salinity, temperature and dissolved oxygen |
| P1B | Current measurements by COSRAD | | surface currents |
| P1C | Provision of meteorological data | | Wind, rain, pressure, temperature, humidity, solar radiation, PAR |
| P2 | Mixing processes | | CTD and microstructure surveys |
| P3 | Inputs from WTP and other industrial sources | | flow data |
| P4 | Inputs from rivers | | flow data |
| P5 | Inputs from land (see N1.2) | | |
| P6 | Inputs from groundwater (see N1.4) | | |
| P7 | Suspended and bed sediments | | sediment cores, sedcam deployments |
| P8 | Hydrodynamic and transport models | | |
| | | | |
| **NUTRIENTS** | | | |
| N1 | Nutrient Inputs/Loads | | |
| | Nl.l | WTP & point sources. | |
| | N1.2 | urban drains | |
| | N1.3 | rivers and streams | |
| | N1.4 | groundwater | |
| | N1.5 | aeolian and atmosphere | |
| N2 | Nutrient Pathways/ Budgets (water) | | |
| | N2.1 | Nutrient status of waters | |
| | N2.2 | Phytoplankton productivity and biomass | |
| | N2.3 | Mapping and sampling of macroalgae | |
| N3 | Nutrient Dynamics in Sediments | | |
| | N3.1 | Sediments nutrient chemistry | |
| | N3.2 | nutrient efflux | |
| | N3.3 | process rates | |
| | N3.4 | Biomarkers in sediments | |
| **TOXICANTS** | | | |
| T1 | Input water quality | | |
| T2 | Toxicants in bay waters | | |
| T3 | Stormwater (see T1) | | |
| T4 | Groundwater | | |
| T5 | Atmosphere | | |
| T6 | Toxicants aspects of shipping | | |
| T7 | Toxicants in bay sediments | | |
| T8 | Toxicants in biota | | |
| T9 | Fate of sediment-bound toxicants | | |
| T10 | Toxicants aspects of dredging | | |

| | | | |
|---|---|---|---|
| **ECOLOGY** | | | |
| E1 | Phytoplankton (see N2.2) | | |
| E2 | Light attenuation (see P7) | | |
| E3 | Macroalgae and seagrasses (see N2.3) | | |
| E4 | Benthic microalgae (see N2.4) | | |
| E5 | Benthic infauna | | |
| | E5.1 | Analysis of historic data on zoobenthos | |
| | E5.2 | Survey of zoobenthos 1994-95 | |
| E6 | Valued molluscs | | |
| E7 | Reef invertebrates (see E6) | | |
| E8 | Distribution, abundance and diet of demersal fish | | |
| E9 | Toxic algae blooms (see N2.2) | | |
| E10 | Other biota and communities | | |
| | E10.1 | Role of benthic invertebrates in nutrient cycle | |
| | E10.2 | Role of zooplankton in nutrient cycle | |
| | E10.3 | Predcition of E.coli on beaches | |
| | E10.4 | Distribution and abundance of phytoplankton | |
| | E10.5 | Zooplankton grazing rates | |
| **GENERAL TASKS** | | | |
| G1 | Data management | | |
| G2 | Resource inventory/mapping of habitats | | |
| | G2.1 | Habitat mapping by remote sensing I | Fieldchecks |
| | G2.2 | Habitat mapping by remote sensing II | CASI data |
| | G2.3 | Habitat mapping by underwater video | Videos (29No.) |
| G4 | Statistical evaluation of task designs and data | | |
| G5 | Bibliography and data inventory | | |
| G6 | Review of PPB (Literature Review) | | |
| G7 | Review of climate change and ozone depletion | | |
| G8/9 | Integrated ecological modelling | | |

**Appendix 2. Entity-Relationship diagram for the proposed centralised database**