

**Distributing Data Ownership: The Northwestern Geospatial Data Network
Donald G. Morgan¹ and Todd Glover²**

**¹Morgan, Donald G., Technical Manager Northwest Data Centre, Prince Rupert Forest Region,
Ministry of Forests, Bag 5000, Smithers BC, Canada, V0J 2N0**

**²Glover, T., Data Warehouse Architect, Information Systems Branch, Ministry of Environment,
Lands and Parks, P.O. Box 9346 Stn Prov Govt, Victoria, BC, Canada, V8W 9M1**

In 1999 Ministry of Forests (MOF) and Ministry of Environment, Lands and Parks (MELP) embarked on a joint ARC/INFO and Oracle data warehouse project to meet increasing requirements for delivery of the most current geospatial data to a variety of government, industry and non-governmental organizations (NGOs), web-based access to cartographic-quality maps, and a common analytical and standards environment for resource and planning analysis. The project, based in Smithers, British Columbia, incorporates a shared technical infrastructure leveraging MELP's existing infrastructure to deliver: distributed data custodianship; shared tools, best practices and standards; multi-agency warehousing of current data for analysis and distribution; user-driven, secure, online access to cartographic products; and end-user online access and analysis. From this project, strategic and operational resource GIS data was delivered to planning processes and MOF clients in November 1999. MOF's Northwest Data Centre together with MELP's Skeena GIS became operational as the first steps towards the Northwestern Geospatial Data Network. The key elements of the success of this project included management support, commitment to joint standards, commitment to honouring data ownership, a local vision of land information management, a functional multi-agency relationship, and a creative MOF/MELP project team. Further benefits include reduced data redundancy and broad seamless data sets that allow for consistent treatments of data and the application of data automation techniques.

1. Multi-Agency Land Management

In British Columbia first nations, forestry companies, non government organizations (NGOs), the Ministry of Mines, Energy and Petroleum Resources, the Ministry of Environment, Lands and Parks (MELP) and the Ministry of Forests (MoF) all have an interest in geospatial data that describes landscapes, natural resources, wildlife habitat, human and wildlife use. In addition, MELP and MoF are required, by legislation and policy, to cooperate on, or at least consult one another on, land management issues. This consultative relationship extends from large area (>500,000 ha) strategic initiatives, such as LRMPs (Land Resource Management Plans) and TSRs (Timber Supply Reviews), to tactical planning at the multiple watershed level, such as landscape unit planning (roughly 50,000 ha), down to the scale of cut block approval. Each agency also has its own separate business needs that require identical land and resource information.

2. Data Warehousing

A data warehouse is an infrastructure to collect, organize and distribute digital information. Its form is particular to the individual organization. Our implementation includes the collection of data such as forest cover and ecological inventories. The data is organized into a structured repository facilitating fast and efficient information retrieval and storage of end products of analysis, such as Timber Supply analyses and plot files. These structures must be in place to have a data warehouse.

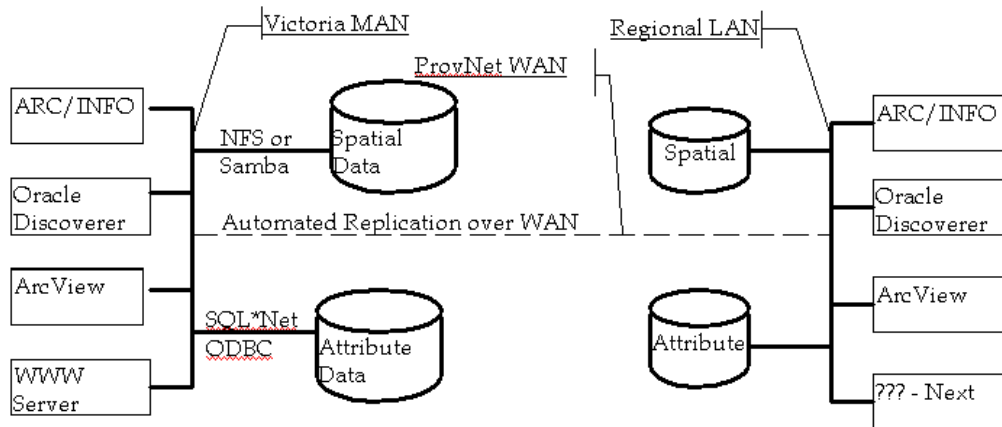
3. Data Warehousing at Ministry of Environment Lands and Parks

As implemented at BC MELP, it is as much a process as a store of data. Information is copied periodically from operational systems and external sources into a read-only database accessible to a wide variety of users.

Each data custodian is responsible for determining how and when their data is 'published' to the warehouse. This may occur periodically for constantly-changing business information, irregularly as data is updated, or only once for static data. This differs from traditional GIS data management, which gathers information as required for a specific project and then archives it.

The master warehouse in Victoria contains about 35 gigabytes (GB) of spatial data, and about 10 GB of attributes. Some of this information is loaded from MELP operational business systems by automated scripts. Some is loaded by GIS and attribute technicians who have acquired and translated it from other Ministries and outside agencies. MELP Regional GIS staff oversee capture and digitizing of many layers for local needs. Special projects such as the Watershed Atlas from BC Fisheries may take years to create, verify and load detailed information. Knowledge of what process was used for each data source is recorded in metadata in the warehouse, or in the MELP Data Registry.

Changes to spatial data are automatically copied to regional sites each night, so that all users have access to current information. Ancillary information such as symbology tables and fonts, and viewing software such as the MELP GOAT are stored in the data warehouse, and managed the same way, see **figure 1**.



Physical Structure of BC MELP Distributed Data Warehouse

Figure 1. Physical structure of BC MELP distributed data warehouse.

MELP's data warehouse architecture is designed to serve three classes of users:

1. MELP employees use spatial information directly on disk via the Local Area Networks in their offices. GIS technical staff log on to the local Unix server to run workstation ARC/INFO. Other employees run ArcView GOAT on a multi-user Windows NT Terminal server or their own desktop PC's. Client/server connections to the RDBMS reach attribute data using the same networks. Victoria staff in several offices access the warehouse over the high-speed (100 Mb) Metropolitan Area Network (MAN). Regional staff access local copies replicated by overnight copies. Staff in district offices have the same data needs, but lack fast network connections to MELP data servers. Procedures are being designed to make copies of warehouse data covering small geographic areas to distribute to them via CD-ROM.

2. External agencies and companies with their own computer infrastructure download files from the public FTP (File Transfer Protocol) site, decompress and convert them to use with their own GIS software. Some predefined queries are also available to fetch attributes from the data warehouse.

3. Public access via Web browser is still in the planning stages. As of late 1998, there is no affordable (to MELP) technology proven capable of serving the data in its current format to the world at large. Interest is increasing, so research will continue.

(adapted from Mackenzie, B. 1999. GIS '99 Conference Proceedings)

4. Data Warehousing in the Ministry of Forests

4.1 The Data Service Centre (DSC) Project

In 1997 the MoF introduced the concept of a Data Service Centre (DSC) to reduce operational costs and improve land management information efficiencies in its operations division. Some MoF business problems that were identified included; making land based decisions across a variety of administrative boundaries, the complexity of the decision making environment, recruitment of skilled staff, high costs of information technology, government downsizing and inconsistencies in organizational structure. The solution to these business problems was to consolidate skilled staff, consolidate land data storage and to consolidate infrastructure.

A DSC consists of a central spatial data warehouse and a set of decentralized client nodes. Data management services required by internal and external clients within a geographic area are organized into "working groups". A central warehouse with full GIS services is called a working group 1 (WG1) and a client node is called a working group 2 (WG2) or a working group 3 (WG3), depending on whether the

node performs inventory updates or not. The provincial design as of July 1999 is shown in **figure 2**. In the Prince Rupert Forest Region the central data warehouse is located in the Regional office (WG1) in Smithers and the client nodes are the Regional Forest Districts (WG2), collectively they form the Northwest Data Centre.

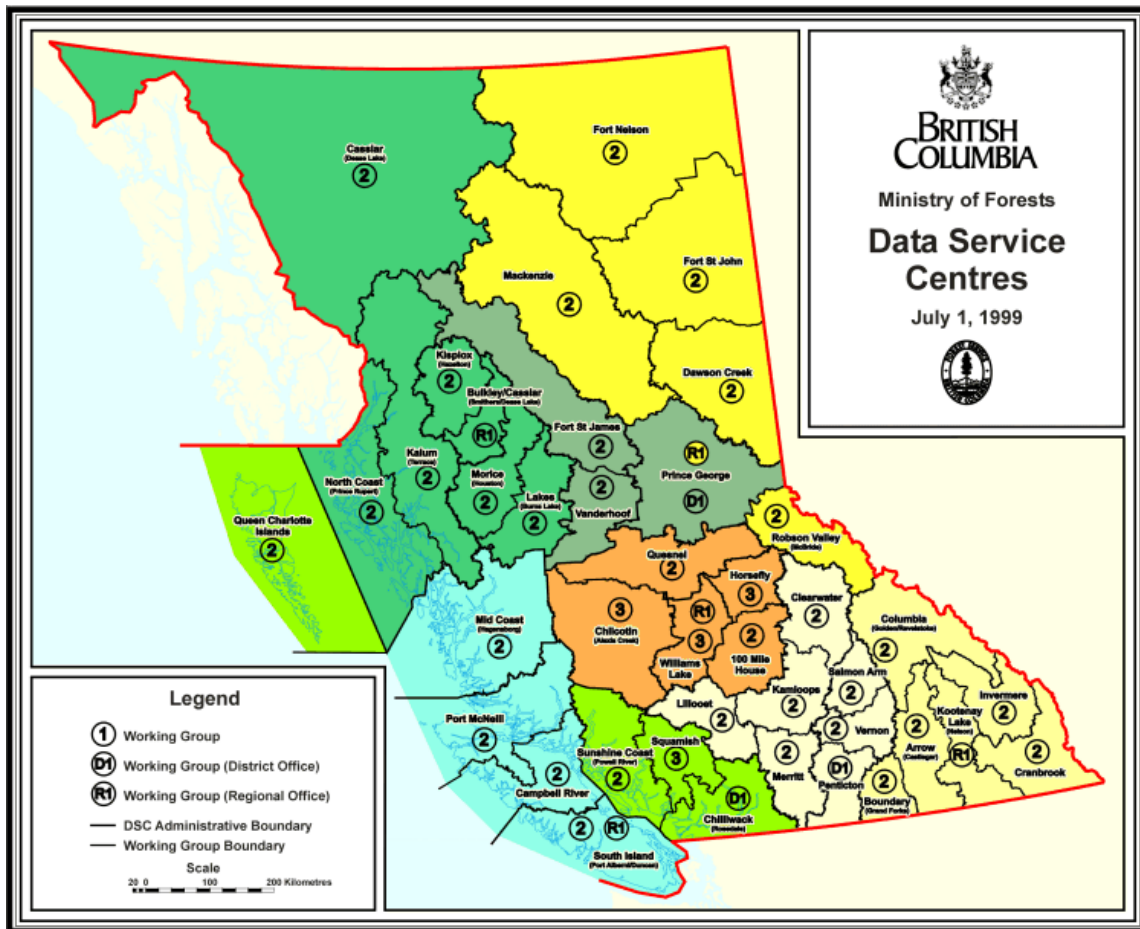


Figure 2. Geographic distribution of data service centres.

The primary functions of the “working groups” of a data centre is to do forest inventory updates, timber harvesting approvals and operational, tactical and strategic planning. The working group structure allows for an integrated approach to data management of MoF standard and non-standard data into a data warehouse, the capacity to develop data sharing agreements with other agencies and the ability to produce analysis ready data sets for end users.

4.2 Information Technology

Originally the intention of the DSC project was to use NT workstations running Arc/INFO and Microsoft Access in central locations for warehousing and analysis and the GeoMedia viewer as a GIS viewer and low-end analysis tool for clients. Currently, MoF Information Management Group in Victoria is investigating the consolidation of all GIS and data base infrastructure to one location in Victoria. Using Microsoft’s Terminal Server, provincial analysts could remotely access Arc/INFO and Oracle on centrally located servers. The terminal server, or multi-user NT, is still being evaluated and has not been fully tested or accepted, but is being considered as a component of a longer term solution to meeting GIS needs. In the short term GIS needs are being met on the desktop, on MoF Regionally based servers and by the centrally located terminal server.

Currently, forest cover inventory files are stored in IGDS format on the object distribution management (ODM) server in Victoria and replicated to MoF Regions and Districts. As the ministry rolls out its integrated corporate spatial and attribute database (INCOSADA) over the next few years this data will be

loaded from the INCOSADA object distribution management server (IODM) through the “INCOSADA Bridge” into a spatial database engine (SDE) warehouse in Victoria as ARC coverages and Oracle tables, as is shown in **figure 3**.

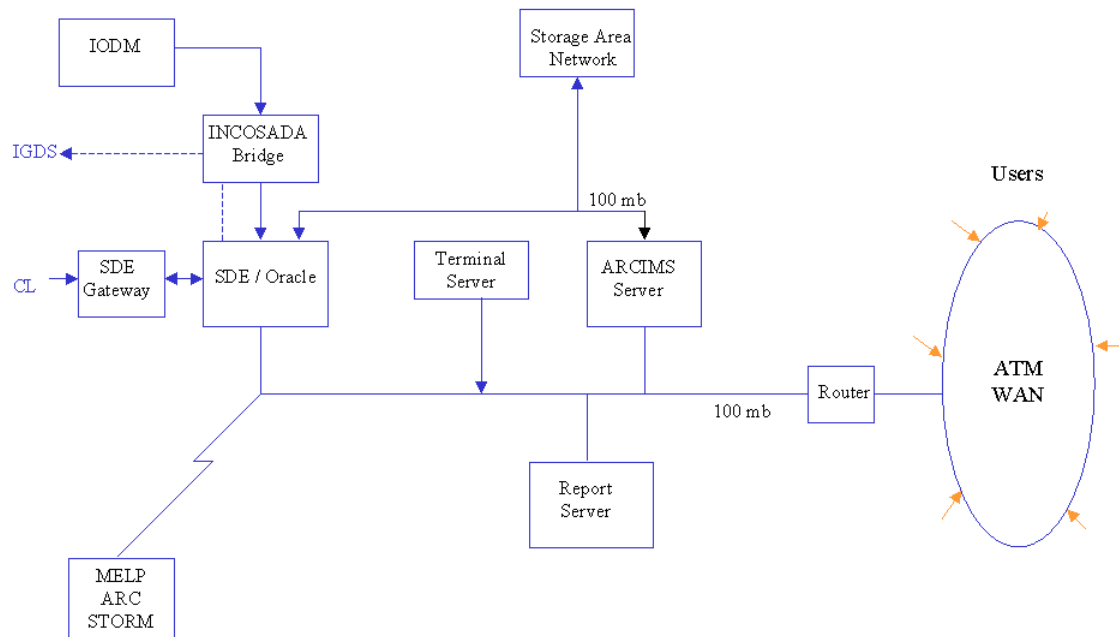


Figure 3. Data Warehousing at MoF: Victoria

A resource data management warehouse is being developed in Victoria and will serve as a gateway to forest cover data and other custodial data sets. Once online it will be a more sophisticated custodial data management system. However, unlike the automatic replication that is the current forest cover data distribution model, MoF Districts and Regions will have to manually access the resource data management warehouse to acquire data. The provincial MoF GIS community will continue to access central custodial data and synthesise and organize the data, as one of many data inputs, into useable information for statutory decision making.

4.3 Prince Rupert Forest Region DSC Implementation

To meet the urgency of their business needs the Prince Rupert Forest Region, facilitated by a supportive Regional management team, implemented its own data service centre vision. Prince Rupert Region has the advantage of being in the same building as the regional MELP office. In 1998 a \$2,500 fibre optic cable was installed that enabled MoF analysts to work off MELP’s UNIX server. This arrangement allowed analysts to provide GIS services to LRMPs, TSR and to Landscape Unit Planning and to become familiar with the structured warehousing done by MELP. By expanding upon the existing arrangement between MELP and MoF a local solution to DSC data warehouse design was implemented that increased the speed at which a data warehouse could be set up and services could be offered to the districts (WG2s), industry and NGOs. To formalise this shared infrastructure and software arrangement a memorandum of understanding was signed by the MoF and MELP Regional Managers. This arrangement continues to have significant cost savings to the provincial government by minimizing the replication of infrastructure and data management services.

4.4 Data Translation and Warehousing

Forest cover inventory files, consisting of the FC1, a DGN graphic file, and an FIP attribute file, are stored on a MoF Regional server. The MoF Inventory Branch places copies that have been updated by Forest Districts using Intergraph's Microstation on Regional servers, after they have done their quality assurance checks. The Northwest Data Centre translates and quality assures all the regional FC1 and FIP files to single precision Arc coverages and Oracle tables. A system has been implemented that tracks the version and the date of the FC1/FIP file on the local network and those that have been translated to Arc and Oracle format.

Being coupled with MELPs Skeena Regional GIS the Northwest Data Centre has access to all of the Regional data replication provided by MELPs central data warehouse, in the future this will include data from the MoFs resource data management warehouse. In addition to the MELP warehouse data the data centre specifically manages 2,390 1:20,000 scale forest cover map sheets in an Arc library. Both Skeena GIS and the Northwest data centre develop value-added information from warehoused data, such as sensitive species habitat and areas of timber harvesting. This information is updated and distributed to forest districts, MELP field staff, industry, consultants, first nations and NGOs. With every data delivery improvements are made to the data sets so they are easier to use and meet the needs of a particular client.

5. Data and Analysis Delivery at the Ministry of Forests

Desktop GIS viewers are used by district staff with a varying degrees of success depending on commitment, training and ease of use. Intergraph's GeoMedia viewer tool was selected as the MoF standard viewer and is deployed throughout the Prince Rupert Forest Region. In addition, ESRI's ArcView 3.2 is also deployed at all locations. Data is prepared in the Regional office so that it is compatible with Arc/INFO, GeoMedia and with ArcView and is distributed by FTP or CD.

6. Multi-Agency Spatial Data Warehousing

6.1 MELP and MoF

Analysts access the MoF/MELP shared multi-processor HP UNIX machine from NT clients running Hummingbirds Exceed. A cluster of SUN UNIX workstations are used to complement the UNIX network, but, are used primarily for research and development. Eight to twelve users have access to 500 GBs of storage space that they use for projects and data warehousing. Each agency manages its own domain specific information, see **figure 4**.

Inter-agency security is maintained through the use of user account permissions. In addition warehoused data is stored in a read only state and, as a result, can not be compromised by warehouse users. Before data is posted to the warehouse metadata must be attached and the coverages must adhere to warehouse quality standards such as topological cleanliness.

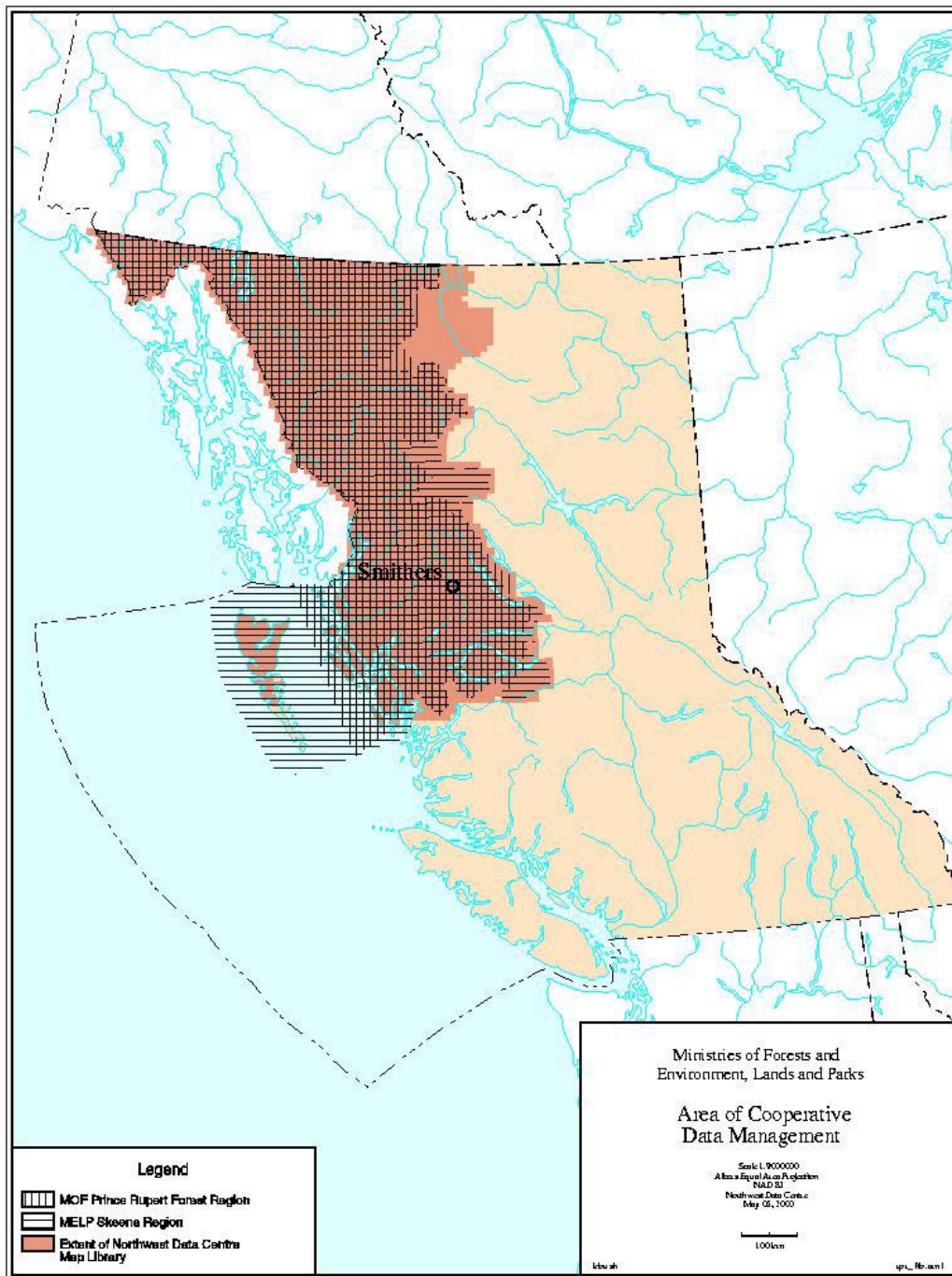


Figure 4. Area of Co-operative Data Management.

As well as supplying warehouse data cut to a client’s geographic interest, the warehouse data is available on MELPs GIS Oracle Access Tool (GOAT). This widely implemented GIS viewer tool allows distributed MELP and MoF clients access to warehoused themes through the MELP terminal server.

6.2 Northwestern Geospatial Data Network

The human dimension is critical to the success of a shared geospatial data network. The technology required to implement a shared system is not a barrier, it is the relationships between the people that are the most important. The shared infrastructure and co-operative relationship between MoF and MELP in the Northwest, for example, has had positive benefits to other inter-ministry activities, building bridges through GIS. The initial participants interested in the data network are the Wet'suwet'en first nation, McGregor Analysis Group, Canadian Forest Products(Canfor), Skeena Cellulose Industries, Sierra Club, MoF and MELP. We plan on holding several workshops and invite other land information agencies to discuss how the Northwestern Geospatial Data Network should evolve. This direction is also supported by Forest Renewal BC .

To facilitate the sharing of data to the wider community of Northwestern BC geospatial data is distributed on CD's or by posting to an ftp sites. We are standardizing a variety of coverages, including forestry data, across the region on a voluntary basis. For example, the data centre has mapped all non-MoF forest data models to the MoF data model so that seamless coverages are available in the Prince Rupert Forest Region.

We are in the process of setting up a intranet group made up of Northwestern geospatial data users. An intranet group will allow all Northwestern Geospatial Data Network participants to share agency specific web pages. Also we are setting up an ARC Internet Map Server (IMS) to publish MoF/MELP land information on our web page so that other agencies can access our published warehouse data. With time other agencies will hopefully put their information on line so that we can easily share data, thus reducing information management costs and increase geospatial data efficiency in Northwestern BC.

7. Biography

Don Morgan is the Technical Manager of the BC Forest Service's Northwest Data Centre in Smithers BC and is the Planning Systems Biologist with the BC Forest Service Research Branch. His main focus is on planning systems for ecosystem management, including GIS systems. He received a B.Sc. in wildlife biology and computational mathematics from Trent University, Peterborough, Ontario in 1984 and a B.Sc.(honours) from Carleton University, Ottawa, Ontario in Quantitative Ecology and Computer Science in 1991. He has been with the BC Forest Service since 1992.