Development of a geographic information-driven real-time surveillance system for disease surveillance

Jiangping Shuai(1), Peter Buck(2), Caroline Chevalier(2) & Paul Sockett(2)

Summary
To respond to emerging public health threats such as West Nile virus, an advanced geographic information systems (GIS)-driven Web-based real-time surveillance system was developed to serve the National West Nile virus dead bird surveillance programme in Canada. The development of this system uses real-time Web GIS technologies and services to enhance conventional real-time surveillance systems based on real-time GIS requirements. The system has three modules: QuickTrack, QuickMap and QuickManage. QuickTrack is the real-time surveillance module that supports data collection, edit and transfer. QuickMap is the real-time Web GIS module that provides comprehensive real-time GIS supports and services in public health surveillance and information sharing. The QuickManage module is a Web-based system management package used to manage the entire system. This system offers an effective approach to enhance real-time public health surveillance systems by integrating real-time Web GIS technologies and services. The system demonstrates that real-time Web GIS technologies can play an important role in enhancing public health surveillance systems.

Keywords
Canada, Geographic information systems, Real-time, Surveillance, Web, West Nile virus.

Sviluppo di un sistema di sorveglianza in tempo reale delle malattie basato sulle informazioni geografiche

Riassunto
Per rispondere a minacce emergenti per la salute pubblica, come il virus West Nile, è stato sviluppato in Canada un sistema avanzato di sorveglianza in tempo reale via Web, basato su sistemi d’informazione geografici avanzati (GIS: geographic information systems) per gestire il programma di sorveglianza nazionale della West Nile col controllo dei volatili morti. Lo sviluppo di tale sistema utilizza tecnologie e servizi GIS via Web in tempo reale per intensificare i sistemi convenzionali di sorveglianza in tempo reale basati su requisiti GIS. Il sistema ha tre moduli: QuickTrack, QuickMap e QuickManage. QuickTrack è un modulo di sorveglianza in tempo reale che si occupa di raccogliere, pubblicare e trasferire i dati. QuickMap è il modulo GIS su Web in tempo reale che fornisce esaurienti supporti e servizi GIS in tempo reale nella sorveglianza della salute pubblica e nella condivisione delle informazioni. Il modulo QuickManage è un pacchetto di gestione di sistemi su Web utilizzato per gestire l’intero sistema. Tale sistema offre un approccio efficace per rafforzare i sistemi di sorveglianza della salute pubblica in tempo reale mediante l’integrazione di tecnologie e servizi GIS via Web in tempo reale. Inoltre esso dimostra che le tecnologie GIS via Web in tempo reale possono...
giocare un ruolo importante nell’incremento dei sistemi di sorveglianza della salute pubblica.

Parole chiave
Canada, Sistema informativo geografico, Sorveglianza, Tempo reale, Web, West Nile.

Introduction

In August 2001, West Nile virus activity was first reported in Canada, with the discovery of the virus in dead birds and mosquito pools in southern Ontario. In 2002, Canada reported its first confirmed human cases in areas of Ontario and Quebec. The virus was also found in birds, horses and mosquitoes in Manitoba, Nova Scotia, Ontario, Quebec and Saskatchewan. Two people in Alberta were infected and infection was thought to be related to travel. A total of 1 495 cases of West Nile virus and asymptomatic infections were reported in humans in 2003, 26 cases in 2004 and 238 cases in 2005 (11). In 2006, five provinces (Alberta, Manitoba, Ontario, Quebec and Saskatchewan) confirmed cases of human infection (13).

Dead bird surveillance plays an important role in West Nile virus surveillance as it helps explain the distribution of West Nile virus activity. Yaremych et al. found that 90.5% of the dead American crows were confirmed positive for West Nile virus from May to October 2002 in Illinois (16). In 2000, New York State, Bernard et al. reported that 47% of dead American crows were tested positive for West Nile virus (1). Watson et al. found a spatial association between early-season crow deaths and human West Nile virus infection cases (14). Mostashari et al. reported that dead bird data could be used for early warning of West Nile virus activity in small areas. Jurisdictions could also develop early larval control programmes and prioritise regions for surveillance based on spatial-temporal cluster analysis using dead bird data (9).

Geographic information systems (GIS) have been used widely in public health by the World Health Organization (WHO) that established a GIS for leprosy elimination (15) and by the Centers for Disease Control and Prevention (CDC) that developed a GIS and public health website (2). Open source GIS have also been used (3, 8). In Canada, GIS is used for a wide variety of public health surveillance systems, such as the disease surveillance online website which provides mapping and other services for cancer, cardiovascular diseases, major chronic diseases, notifiable diseases and injury surveillance (10).

Since the introduction of the national West Nile virus surveillance programme and subsequent requirement for GIS support in the Public Health Agency of Canada, the format of GIS applications has evolved constantly. In 2001, ArcView™ was used to create and export static image maps that were posted on the website. In 2002, html ImageMapper™ was introduced to transform ArcView™ maps and data into an interactive map website. In 2003 and 2004, a new interactive map website was developed using MapServer™ for which an automated ArcView™ application produced ten daily static maps. While this innovative approach reduces workload and file transfer volumes, the original static maps created in 2001 also remain functional. As the static and interactive maps operate independently, operation costs have increased. The preferred option is to streamline these applications and systems to automate any manual GIS operations.

Between 2001 and 2004, the GIS elements for West Nile virus surveillance have continued to play a crucial role. New surveillance methods and data sources have been introduced enabling a georeferencing solution that improves data quality and location accuracy. Funding opportunities, such as GeoConnections funding for sharing spatial data through the Canadian geospatial data infrastructure in the field of public health (4) have also arisen. The benefits of using GIS in other passive and active surveillance programmes have been discussed. To address these GIS needs and to streamline national West Nile virus surveillance systems, a Web-based integrated real-time surveillance system was proposed in 2003 and the development of this system was completed in 2004.
Materials and methods

The design of the real-time Web GIS technologies and the integration with a Web-based real-time surveillance system was based on the requirements of the national West Nile virus dead bird surveillance programme. During the development of this GIS-driven real-time surveillance system for disease surveillance, two important steps were taken in the design and integration process, namely: to determine the GIS requirements for a real-time surveillance system and to set the guidelines of GIS technologies development.

Determination of geographic information system requirements

From 2001 to 2004, there has been an increasing demand for GIS capacities as part of West Nile virus surveillance programmes. These needs included the quality improvement of spatial data, automation of spatial data generation, capacity to allow end-users to produce maps for any date, addition of security control in incident data, additional methods to generate coordinates using different location information, automatic generation of static maps, more interactive Web mapping functions, GIS support in data collection and editing and the integration of GIS operations in surveillance operations.

Based on the analysis of the objectives of West Nile virus surveillance and the increasing GIS needs to serve this surveillance programme, five major GIS requirements were identified for this system to serve national West Nile virus dead bird real-time surveillance, as follows:

- **Integration.** In the past, maintaining a separate real-time GIS system and surveillance system increased costs. In this system, GIS capacities and components needed to be fully integrated with real-time surveillance system to make the entire system more efficient and accessible for internal and external users. The real-time GIS components would also need to improve surveillance data collection and data editing.

- **Data quality improvement.** GIS was needed to improve the quality of spatial data in surveillance data collection. In dead bird surveillance, geo-referencing was required to identify bird location. A variety of location information (street address, postal code, GPS reading, mapbook and legal land description) was reported to the surveillance system based on different provincial reporting methods. All the information was used to generate coordinates. GIS was also used also for data validation when data recorded were not consistent.

- **Visualisation.** Visualisation is an important contribution of real-time Web GIS for a public health surveillance system. In this system, four types of visualisation were required, namely: the first is represented by maps concerning the spatial pattern of virus, the hotspots of virus activities and spatial clustering of virus activities in birds. These maps include bird location point maps and aggregated maps (tested birds and confirmed positive birds in each health unit). The second is the time series map (daily, weekly, monthly or yearly). The third is the ad hoc map that users can create from the website by defining any specific periods of time during the surveillance season. The fourth is a map that is suitable for reports, which include legends, statistics, logos, inserts and titles.

- **Data security and confidentiality.** A security mechanism was developed to protect the privacy of individuals or communities. Guarded by this mechanism, surveillance data in the map services could not be downloaded by any other GIS applications.

- **Spatial data sharing.** Sharing various spatial data between different surveillance programmes and other spatial data providers facilitates decision-making. The open geospatial consortium (OGC®) Web mapping service (WMS) standard has enabled GIS applications to share and distribute map-like views of spatial information in a distributed network (8). In Canada, the federal government has developed a Canadian Geospatial Data Infrastructure Programme (CGDI) to enable safe sharing of reliable geospatial data information across jurisdictions and enable collaborative and timely decision-making (4).
This system has developed the capacities to access the rich WMS data sources through the CGDI WMS technical standards (5). The system also created a WMS data source to provide national West Nile virus dead bird surveillance information to the public.

**Development guidelines for the QuickMap module**

In this system, all the GIS components have been encapsulated in the QuickMap module. QuickMap has been developed to meet immediate GIS needs in national West Nile virus dead bird surveillance as well as to serve future zoonotic disease surveillance programmes. Hence, four guidelines have been identified for the development of QuickMap, as follows:

- **Real-time.** The GIS capacities should provide real-time support for recording and editing of surveillance data. Whenever new data is entered into the system or existing data is edited, the geo-referencing and spatial validation processes commence automatically. Visualisation of surveillance data should also be real-time. Once the surveillance data is recorded into the database, all maps and spatial queries are updated immediately. Static maps and WMS data reflect the most recent data at the time of updates. New technologies have been developed for ad hoc maps in which date spans are defined by the users on the fly.

- **Web-based configuration.** In this system, the GIS surveillance links and interactions can be edited from a Web-based configuration page. This has provided a convenient way to update system parameters, such as static map production schedules, ArcIMS™ service names, colour of the thematic maps, geocoding services and MWS resources.

- **Automation of GIS functions.** Geo-reference and spatial validation functions have been developed to run automatically. If users provide multiple location information, the system will automatically follow a predefined order to generate coordinates and check if it is spatially valid. Static maps, point shapefiles and WMS updates are operated automatically by the server. This design has reduced operation and maintenance costs.

- **Reusability and scalability.** The GIS functions and Java classes of the systems should be able to serve other surveillance programmes in the future, as needs arise. The links between the real-time surveillance module and the real-time GIS module should be not hard-coded, so that system administrators can adjust Web mapping and modify queries for other surveillance programmes. The system is also scalable and can be deployed to one server (for a small user group), or to three servers (Web GIS servers, database server and Web/application server) for larger user groups.

**Results**

As mentioned earlier, the integrated real-time West Nile virus dead bird surveillance system has three modules: QuickTrack, QuickMap and QuickManage. The three modules have been integrated seamlessly (Fig. 1). The major components of these three modules are listed in Table I. The system is a J2EE application. The functionality of the application is separated along different tiers to provide a separation of responsibilities. The Web tier uses the Struts framework. The business tier is based on the Spring Framework™ to manage business objects because it can effectively organise middle tier objects. Spring framework takes care of plumbing and thus helps to reduce development and maintenance efforts. The data tier is developed using iBatistm™ because of the simplicity of iBatistm™ data mapper over object relational mapping tools.

The system uses ArcIMS™ as the geospatial processing and mapping engine, MapServer™ as the WMS engine, and Oracle® as the database engine. Apache™ and Tomcat™ are used as Web application servers because they are open-source and reliable. The system uses Struts to implement internationalisation. Currently, the public website supports English and French. However it could support other languages as required.

This system has a public-internal infrastructure so one system serves both the public
QuickManage

QuickManage is the system configuration module of this system. It provides real-time management of system configuration, static map schedules, group management, user management, session management, embargo management, WMS management, and geo-coding management.

QuickMap

QuickMap provides real-time georeferencing, spatial validation, real-time interactive maps, real-time GIS and real-time surveillance data link, web GIS clients, GIS in QuickTrack, static map generation, storage and retrieval, time-series maps, OGC WMS access, OGC WMS services, and point shapefile generation.

QuickTrack

QuickTrack is the real-time surveillance module of this system. It provides real-time incident data reporting, editing, and search, data transfer ETL, aggregated data query, and connection to the database. It has replaced the old desktop GIS-based West Nile virus dead bird surveillance system and has been established to serve national West Nile virus dead bird real-time surveillance since May 2005 (12).

Table I
System modules and components list

| Modules      | Component list                                                                 :
|--------------|-------------------------------------------------------------------------------:
| QuickTrack   | Incident data reporting, Incident data editing, Incident data search, Data transfer ETL, Aggregated data query                      |
| QuickMap     | Geo-referencing, Spatial validation, Real-time interactive map, Real-time GIS and real-time surveillance data link, Web GIS clients, GIS in QuickTrack, Static map generation, storage and retrieval, Time-series maps, OGC WMS access, OGC WMS services, Point shapefile generation |
| QuickManage  | System configuration, Static map schedules, Group management, User management, Session management, Embargo management, WMS management, Geo-coding management |

ETL: extraction, transformation and loading
OGC: open geospatial consortium
WMS: Web mapping service

This approach has streamlined system management and reduced maintenance cost. Public users can access the public section. Registered users can access the internal part after logging in. The real-time map and data query pages use the same page for both the public and internal parts. However, if a user is not logged in, a portion of these pages is neither visible nor accessible. The system is flexible using this mechanism. Currently it is deployed with integrated public and internal sites, but can also be hosted as a public system only or as an internal system on its own by changing the setting in an xml file. This might be useful if an organisation has already set up a surveillance system and just needs real-time GIS functions. The GIS functions of this system can be added to its surveillance system, connected to the database and real-time maps can be generated.

The system has replaced the old desktop GIS-based West Nile virus dead bird surveillance system and has been established to serve national West Nile virus dead bird real-time surveillance since May 2005 (12).

QuickTrack module

QuickTrack is the real-time surveillance module of this system. It provides real-time
data reporting, editing and query functions. Dead bird submissions and testing data are recorded at the case level. In data reporting and editing, the system handles over 100 data fields that are divided into eight sections. In addition, many data validation rules have been implemented. Users can create new records in QuickTrack, edit existing records or search specific records. QuickTrack also has an embedded real-time map with interactive zoom, pan and coordinate tracking functions (Fig. 2).

Based on the quality of the spatial data elements, records are classified as complete and incomplete. When the geographic information is correct (coordinates, health unit, province/territory) the record is classified as ‘complete’. If coordinates are missing or incorrect, it is classified as ‘incomplete’. Internal users will see a red warning when they open incomplete records for edit. Moreover, a record is classified as ‘valid’ if the reported health unit is within the reported province; otherwise the record is classified as ‘invalid’. These invalid records are still stored in the database but, although available, are only available for internal users to search and edit.

All aggregated queries are created on the fly in order to protect data confidentiality and accomplish real-time performance. Four aggregated data queries are available, as follows:

- date of first and last cases of positive dead birds and total confirmed positive dead birds during the period
- monthly record of dead bird testing results (new and cumulative results)
- daily records of dead bird tested numbers and confirmed positive numbers
- dead bird testing statistics by species.

These queries can be aggregated at three regional levels: national, provincial/territorial and health unit, and the query results can be saved. Surveillance data reporting, editing, managing, and record searches are internal functions: only registered users can log on to the internal site and can operate the system based on their roles. Provincial/territorial users can only report, edit and manage data records in their own jurisdiction.

![Figure 2](http://www.izs.it/vet_italiana/)

**Figure 2**

Data reporting and editing with an embedded real-time map in QuickTrack.
A special component (extraction, transformation and loading or ETL) has been developed to transfer data from external systems to this system automatically. This ETL component has been particularly useful in that it has enabled data transfer from other systems to this system and has played a significant role in data quality assurance. If a number of records need to be corrected manually using external GIS applications (such as ArcView™ or ArcMap™), all the corrected data can be transferred to the system using this ETL component, thus avoiding the manual update of large datasets through the Web interface. In the future, ETL may become a key tool if provinces/territories cease to report data to this system on a daily basis. Using this ETL component, provincial users could export their data to a predefined format and automatically transfer the export to the national system.

**QuickMap module**

QuickMap is the real-time Web GIS module. It consists of a real-time Web GIS interface and a number of Web GIS functions and capacities that are important for public health surveillance. This system allows users to process, visualise, query and link available data for any dates. All incident and aggregated maps are created on the fly to protect data confidentiality and accomplish real-time performance.

When the Web page is launched, a real-time map is automatically generated, based on the predefined date. As a major function of real-time GIS, users can also define any dates to generate new maps and all these real-time maps can be queried or selected. Users can zoom to different provinces by changing to the desired province from a combo box. The map page has common Web interactive mapping tools. The internal map also has a WMS data access tool.

On the public site, the real-time map displays the surveillance information at health unit level (Fig. 3). On the internal map, users can see the dead bird points within the permission assigned to them, based on confidentiality requirements. However, all internal users can...
visualise the aggregated map that shows the dead bird submitting and testing information by health units (Fig. 4). Users can select incidents in the internal map by drawing a rectangle on the map. The selection results will be listed in a table below the map. If users click an incident ID in the table, the system will switch to the data editing page with loaded data.

QuickMap also provides real-time geo-referencing and spatial validation. Six different methods were developed and integrated to produce correct latitudes and longitudes: on-screen location pinpoint, GPS coordinate data, address geo-coding and address intersection geo-coding, postal code lookup, map grid (for Nova Scotia) lookup and legal land description lookup (for Alberta, Manitoba and Saskatchewan). A process of validation will go through the six methods if coordinates fall outside the boundary of the health unit reported.

Using the WMS capacities, internal users can add WMS data layers from both public and secure WMS sources to the real-time map. For example, users can add vegetation index, weather conditions, land-use, or landscape layers to the real-time map, explore the spatial association between West Nile virus activities and different risk factors and generate hypotheses.

QuickManage module
QuickManage is a group of tools and services that enable administrators to manage the different parts of the system. The configuration component handles 38 parameters. These define how maps are drawn, what fields are used for colours, what fields are used for labels, and the server names, image types and ArcIMS™ service. Using these parameters, the links between QuickTrack and QuickMap modules and mapping methods are not hard-coded. If a new ArcIMS™ image map service is set up for a surveillance programme,
administrators can change a configuration parameter from the Web to utilise it. If this system is used to serve other surveillance programmes, all the parameters can be changed from the Web with ease and there is no need to change the source code of the GIS components.

**Discussion**

GIS has been used widely in public health mapping. Maged and Boulos provide a comprehensive review of these practices (7). However, Web GIS is less commonly applied in Web-based public health surveillance systems, especially Web-based real-time GIS for real-time surveillance systems.

Driven by innovative real-time Web GIS technologies, this system has been designed and developed as an integrated system to track, update, manage, map, query and deliver West Nile virus dead bird surveillance data. Both public health professionals and the general public can access it at different levels. This system has become an important resource for monitoring and understanding West Nile virus activities in Canada (6).

This system demonstrates that Web GIS may enhance public health systems in two ways. One is to enhance surveillance data management. This is achieved by applying geo-referencing, spatial validation, GIS functions in data reporting tools and link GIS and surveillance data reporting and editing. The other is to enhance surveillance information services, such as the provision of a variety of visualisation, spatial query, WMS data overlay and WMS data serving (Fig. 5).

This system demonstrates that real-time Web GIS can play four important roles in enhancing public health surveillance: data quality improvement, visualisation, spatial data sharing and spatial data generation. These roles are equally important and are interactive. The use of this system in production has shown that geo-referencing and spatial validation are particularly useful as they have generated reliable coordinates and validated spatial relationships among coordinates, health units and provinces/territories.

In this system, multiple visualisation methods have been developed to display health events, or to assist exploratory spatial analysis. The real-time interactive mapping function allows

![Diagram](https://via.placeholder.com/150)

**Figure 5**
Enhancement of real-time Web GIS to real-time surveillance
users to create maps for any period of time, which allows public health professionals to examine health events over a specific period. Static maps are generated by the server and are stored in the database. This method significantly reduces server loads. In the internal system, users can also visualise individual dead birds on the map, thereby facilitating the identification of hotspots.

There is a specific risk in using Web GIS. In certain conditions, external GIS applications can access ArcIMS™ services and review data in shapefiles. Consequently, storing data in shapefile format might not meet the strict security requirements of this system. New technologies have been developed to meet security requirements. Using these new technologies, this system stores surveillance data in an Oracle® database, not in shapefile. By taking this approach, layers in ArcIMS™ services do not have surveillance data. Accessing the ArcIMS™ service from outside this system will not reveal any surveillance data. These technologies also reduce development costs. These secure real-time Web GIS technologies can be recommended in public health surveillance systems.

The development of this system demonstrated that there were multiple benefits to integrate Web real-time GIS systems into surveillance systems. It eliminated the need to design, develop and maintain separate GIS systems and surveillance systems. This integrated system has enhanced surveillance data quality by providing geo-referencing and spatial validation. It has also reduced operation costs since there is also no need to transfer data between GIS components and surveillance components. There are no data discrepancy issues among the GIS and surveillance components in this integrated system, while data discrepancy is a challenging issue if GIS system and surveillance system are separate. The integrated approach has also made architecture design and system development of real-time surveillance and real-time GIS easier, as the two modules share the same database.

Acknowledgements

The development of this system was sponsored and supported by the Foodborne, Waterborne and Zoonotic Infections Division, Public Health Agency of Canada; Canadian Network for Public Health Intelligence; GeoConnections; Canadian Cooperative Wildlife Health Centre; GeoTango; Office of Public Health Practice, Public Health Agency of Canada; Epidemiology Services, British Columbia Centre for Disease Control; Nova Scotia Department of Natural Resources; and York University.

Grant support

The development of this system was funded by GeoConnections and Foodborne, Waterborne and Zoonotic Infections Division of the Public Health Agency of Canada. We would like to extend our special thanks to GeoConnections for its funding support.

References