A Web Service for the Dynamic Linkage and Visualisation of Multivariate Spatiotemporal Information

Simon Moncrieff *
Department of Spatial Sciences, Curtin University
GPO Box U 1987
Perth WA 6845, Australia
s.moncrieff@curtin.edu.au

Geoff West †
Department of Spatial Sciences, Curtin University
GPO Box U 1987
Perth WA 6845, Australia
g.west@curtin.edu.au

ABSTRACT

In spatial health research, it is necessary to not only consider the spatial and temporal distributions of diseases, but also external factors that influence the disease, such as environmental and socio-economic factors. In this paper, we propose a method for dynamically linking, and subsequently visualising multivariate spatiotemporal data. The aim of the method is to enable a user to create new analysis data sets by combining the output of multiple web services to facilitate the exploration of data relating to factors that may influence, or are related to, a disease or group of disease under consideration. By creating a generic layer store, termed a visualisation object, the approach utilises semantic web concepts in order to populate the layer store with vector data from map data layers, and the results of processing web services. This approach represents a dynamic, query based approach to web GIS that focuses on enabling access to data and data exploration. A number of visualisations, leveraging both the multivariate and temporal nature of the visualisation object, were developed to explore the information visualisation aspect of the visualisation object. While complex visualisations are possible using this approach, thematic maps for a single epoch can be generated and viewed within a standard web GIS using a previously proposed dynamic web map server, thus enabling the caching of the results of complex processing results for visualisation and data access.

Categories and Subject Descriptors

H.2.8 [**INFORMATION SYSTEMS**]: Database management—*Spatial databases and GIS*

*Cooperative Research Centre for Spatial Information.
†Cooperative Research Centre for Spatial Information.

HEALTHGIS '13, November 5-8 2013, Orlando, FL, USA

General Terms

Design, Algorithms

Keywords

Spatial web service, Geo-visualisation, Spatiotemporal analysis, Data linkage

1. INTRODUCTION

Information visualisation and visual analytics comprises the process of producing visual summaries of potentially large datasets, and the subsequent interaction with the visualisation in order to analyse and derive meaning. This process is also referred to as Geo-visual analytics in the case of applications that involve geographical space [2]. The aim of information visualisation is to enable data exploration by combining cognition science, data mining and human intelligence [1].

Health research is a field in which information visualisation is important, not only to understanding the spatial distribution of diseases, but also in understanding external factors that may influence the occurrence of a disease beyond the spatial. Understanding the links between diseases and such factors is particularly pertinent in the case of preventable diseases. Consequently, combining health related outcomes with associated socio-economic and environmental attributes, and incorporating temporal information for trend analysis, is becoming increasingly important in health data analysis [5].

The use of the semantic web enables the automated access of data over the web, both in terms of processing, sharing and visualising data across web services. Examples include data integration in the GIS domain [15], and the application of such approaches to health data, for example, through the use of a Web Processing Service (WPS[1]) [7]. Further, there is also recognition of the role of information visualisation in multivariate spatial analysis within web GIS, including the linked visualisations of multivariate spatiotemporal data [16, 8]

This paper combines these three factors through the introduction of a web service that combines aspects of a WPS, and web feature service (WFS[2]) The web service performs

[1] `http://www.opengeospatial.org/standards/wps`, accessed: 20 August 2013
[2] `http://www.opengeospatial.org/standards/wfs`, accessed: 1 September 2013

the dynamic linkage of web data sources to facilitate multivariate, spatiotemporal information visualisation, with a particular focus on Health data and aiding Health related research. To achieve the data linkage we leverage semantic web concepts in order to create a dynamic spatial dataset on the fly, through the integration of multiple geo-spatial web services, including geo-spatial web process services and vector based web data sources, such as WFS. The resulting dataset is extensible with respect to the number of attributes, facilitating multi-variate analysis, and incorporates a temporal dimension, enabling spatiotemporal visualisation. This process enables a user to rapidly generate a new dataset from dynamically linked information, generally comprising an analysis variable of interest, along with a set of potentially related variables. The user is then able to explore the relationship between the variables through visualisation and further analysis.

Attribute linkage is achieved by concatenating over a common spatial geometric resolution, for example, using common geometry boundaries, such as geo-political boundaries, or statistical areas used for statistical analysis. Consequently, while the spatial resolution is initially flexible, once the first attribute has been added to the dynamic dataset, the resolution is fixed. That is, the method generates a dataset derived from a number of linked data sources, a spatial geometry layer, and temporal information, which subsequently enables a user to visualise and interact with the new dataset. While this paper focuses on health, it is applicable to any datasets that are spatially aggregated for analysis, for example socio-economic and census data.

The creation of the dynamic dataset is implemented as a web service that consumes other web services in order to generate a composite data store. Input to the web service currently comprises a number of required parameters, and a specified data source, which is then used as the base attribute for the dynamic dataset. Compatible vector data sources include OGC Web Feature Service, ESRI rest services, or other vector input. For example, the use of a GeoJSON, or KML, file would enable a user to upload a dataset, which could then be linked with other web accessible data sources.

Due to the non-specific nature of the visualisation dataset, the output mechanism used for the encapsulated data can act as WFS or WPS dependent upon the required information output. For example, each map layer can be output individually with an associated year, or the map layers can be output as combined vector data incorporating year as an attribute for either data access to the entire visualisation dataset, or when required by complex spatiotemporal visualisations. Consequently, the output from the process can then be readily integrated into a Web GIS environment in order to facilitate the online access of the data. For example, a vector output of a map layer can be used as input to a dynamic map server that can then be used to render the corresponding thematic map as an WMS layer [12], or several output types can be used in conjunction with complex visualisations derived from the dynamic nature of the dataset [16, 8].

The layout of this paper is as follows, Section 2 outlines relevant approaches to web GIS, particularly in health, Section 3 details the method used to create an store the dynamic data layer, and Section 4 details examples of visualisations that can be achieved using the resulting linked information. This is followed by the conclusion.

2. HEALTH WEB GIS

GIS is a useful, although under utilised, method for examining epidemiological data [10, 9]. Samarasundera et al. [14] provides an examination of the use of GIS in health. Exeter *et al.* [5] proposed that the linkage of multiple data sets would allow for the more efficient use of health and spatial data, and is especially salient in the case of health data analysis, as many factors influence health outcomes, including environmental and socio-economic factors. Thus, web GIS is of particular interest when considering access to health information. Web GIS approaches for health range from the mapping of layers depicting specific disease outcomes or health services [6], to providing access to specific spatial datasets in the form of interactive map layers [4]. Thematic maps are predominantly used in health based web GIS to present results, or visualise relatively static snapshots of a dataset [14]. Although such approaches are helpful for increasing the awareness of spatial information, the scope for data analysis is limited. Increased interactivity within a thematic map can be achieved through rendering the map on the client using vector data returned from the server, for example using SVG [10], KML [17], or GML [11]. Such data driven approaches incorporate the ability to perform basic data queries, for example, by including a year range, and can accommodate dynamic query driven thematic styling. However, as the output map layer remains unchanged, analysis using this approach is limited to the visualisation of the specified data subset. Consequently, Gao *et al.* [7] introduced an example of an WPS web service for determining a number of statistical properties of health data in order to make health information accessible over the web.

The methods outlined above form examples of web based access to health information including access to data layers [4], subsets of data [11], and processing services [7]. The first two sources comprise vector data sources, while the latter represents an example of appropriate WPS data sources. The approach proposed in this paper combines access to such information, in conjunction with dynamic web mapping [12] and multivariate visualisation [16, 8], to enable the dynamic linkage, and subsequent exploration, of data, representing an initial stage in spatial health data analysis.

3. VISUALISATION OBJECT

Figure 1 shows the workflow associated with the generation and access of a visualisation object data store. To handle the dynamic dataset generation web service and subsequent output, the concept of a visualisation object software agent is introduced, comprising a query engine, a data linkage module, a generic visualisation object data store, and an output mechanism for data visualisation.

3.1 Data Linkage

In order to generate a visualisation object, an initial attribute of interest is specified by a user, for example, a health summary statistic. Subsequent attributes are then appended onto the attribute list of the visualisation object using a common geometry resolution. This process is achieved using a combination of the query engine, and the data linkage module.

The query engine accesses web services that return vector

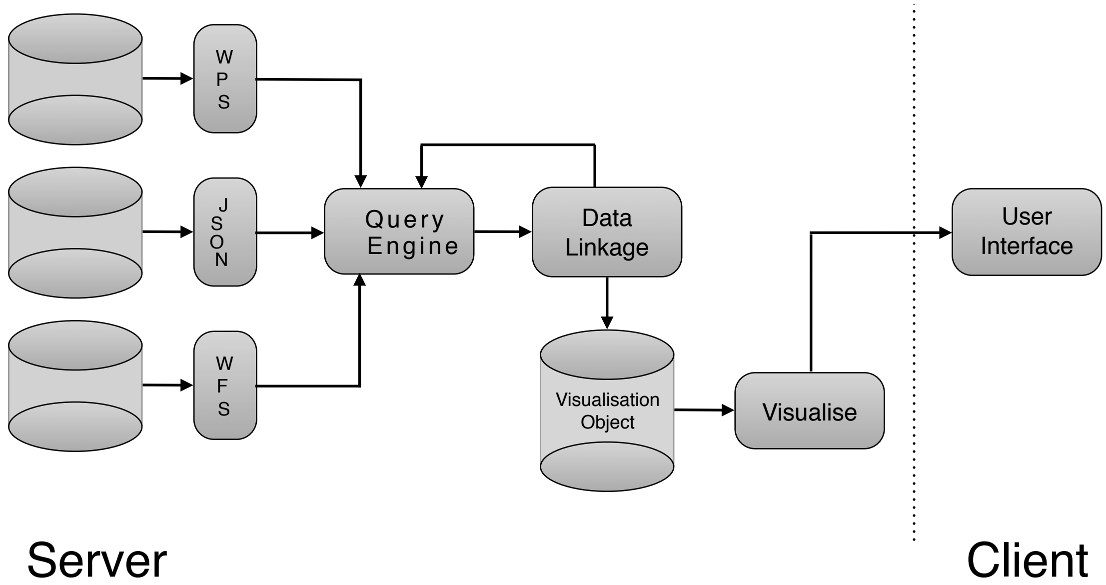


Figure 1: The web GIS workflow for the generation and access of a visualisation object.

data. The web services can comprise vector data sources, such as WFS map layers returning GeoJSON or GML vector data, or processing service that similarly return vector data, such as spatial health summary statistics. The query engine then transforms the data to generate output in a common data format that can be parsed by the data linkage module. In this manner, the data query engine can be extended to communicate with a number of data sources, including WFS and WPS, and with appropriate protocols in place, ESRI REST end points that return vector data. In order to facilitate the data integration process, semantic web concepts such as ontologies can be also be incorporated into the query engine.

The data linkage component then initiates a query engine process, and either appends the returned attribute(s) to an existing visualisation object, or initialises a new visualisation object, setting the geometric resolution for the object to that of the returned web service. The data linkage is performed using a programmatic join between the output of the query engine, and the existing geometries within the visualisation object, and can be achieved using either a unique identification code associated with each polygon in the geometry layer, or a spatial join method such as a geometry intersection. Consequently, the data linkage requires the data sources to be stored at common aggregation boundaries, such as geometries associated with census data collection and analysis. Therefore, to enable data linkage, the geometry resolution of the web service should be a parameter, or the data layer returned should correspond with a known geometry. In certain cases where the geometry of the web service and visualisation object do not match, data linkage can be achieved by aggregating up through a geometry hierarchy, or merging using geometric intersections accounting for population density, where appropriate. The latter is less desirable as it represents an approximation.

The output of the web service comprises a pointer to the newly generated data source corresponding to the created composite dataset, which can then be accessed through a REST interface returning vector data as output.

3.2 Data Storage

As the analysis resolution, and incorporated datasets can be considered variables, storing the data as a map layer in a traditional relational database table is problematic. Thus, the visualisation object store is a generic container comprising meta data, including styling information, and $1 \ldots N$ map layers, each with an associated set of layer meta data. Each layer consists of $1 \ldots M$ attributes. The thematic styling information (Style Layer Descriptor - SLD) can be stored per layer, or a group of layers can share a single style. It should be noted that neither N, nor M, need be fixed for any visualisation object. The form of the visualisation object data store is shown in Figure 2. The visualisation object offers a number of analysis and visualisation options and properties depending upon the values of N and M:

- $N = 1$: In this case, the visualisation objects acts as a layer cache, storing a complex dynamic layer and the associated SLD. After the initial visualisation object creation, access time for the data can be greatly reduced. Consequently, the visualisation object can be used as a cache for the output of a compute intensive geo-spatial web processing service, for example, calculating the age standardised rate, or aggregating point data for analysis.

- $N > 1$: Multiple layers are grouped into a single visualisation object. For example, a temporal visualisation object, with each layer corresponding to a time epoch.

- $M = 1$: The visualisation object comprises a single attribute, the base attribute used to create the dynamic dataset. In this case, the visualisation object is akin to a cached vector representation of a processing result, that can be accesses using a web map service (WMS) thematic layer.

- $M > 1$: The visualisation object contains multiple attributes, enabling multivariate visualisation, analysis, and data access.

The visualisation object enables the generation of multiple layers by incrementing over N, for example, year. Additionally, multiple attributes, corresponding to the same geographic boundaries, can be appended to each layer within a visualisation object. This process enables the rapid generation of spatiotemporal and/or multivariate data sets, and thus visualisations. The visualisation objects can be temporarily stored based on the creation timestamp and a time to live, if the object is simply being used as a layer cache. Alternatively, the visualisation object can be persisted when extended data exploration is required, or the data and resulting visualisations are required to be permanent, for example, through sharing a visualisation via a permanent web page link. The visualisation object represents an object orientated approach to storing spatial data, and consequently, while a relational database such as Postgresql/PostGIS and be used, a hybrid relation and NoSQL, such as MongoDB, database approach can also be used; such an approach would be more efficient for larger data sets. In order to use a relational database approach, a generic object data store is required, in which attributes are stored as a string value, along with the original data type of the attribute; when extracted from the database, the attribute can be recast to the stored data type. The use of meta-data is also important both in forming, and subsequently accessing, the visualisation object. For example, the meta-data can used for automating further processing, and integrated new variables into the attribute set of the visualisation object.

The styling information input required to generate a thematic map associated with the dataset can also be created when generating the dynamic dataset. This is achieved by determining a style layer descriptor (SLD) for the dataset, which can be generated in two ways. If $N = 1$, then the SLD is created in the usual manner; thus, the dynamic dataset can be considered to be a dynamic thematic map layer. However, if a non-spatial axis, such as time, is included $(N > 1)$ in the data, a single SLD can be created for each epoch within the dataset, resulting in the visualisation object consisting of a group of dynamically created map layers; this approach enables spatial comparison within an epoch. Alternatively, each layer in the visualisation object can be styled according to a common SLD, created by pooling the values for a given styling attribute over all epochs to form a single feature vector. The styling classification can then be determined for this global feature vector. This methods is suited to information visualisation between epochs, for example, visualising spatial pattens in the data over time, such as shifts in the distribution of the population. The style layer descriptor is generated by a separate WPS that accepts the feature vector array, typically the base attribute for the dynamic dataset, the classification method, and the number of classes to be generated as input, and outputs the class boundaries, along with the number of entries within each class. A number of map classification methods are supported, including: natural breaks, quantiles, and equal intervals [13].

3.3 Data Access

The visualisation component of the visualisation object software agent was designed to make a number of data output mechanism available through the visualisation object REST interface. For example, the output can emulate a WFS, returning the attributes and geometries, which can then be combined with a previously proposed dynamic web map server to return a thematic map [12]. This method enables the dynamic generation of a thematic map given a number of input parameters specifying the map properties, such as the attribute to visualise, and filter parameters, such as the year, or epoch. This data access method enables the attributes in the visualisation object to be readily integrated into existing web GIS portals.

Alternatively, JSON or XML data can also be returned by the visualisation component, formatted in a manner that facilitates using the data output as the corresponding input data for non-spatial visualisations. For example, incorporating temporal information into the output information. This approach can be used to generate complex web based visualisations incorporating both multivariate and spatiotemporal aspects.

When generating the dynamic dataset, further processing can be specified within the dataset input generation query in the form of determining style information, and the processing of the attribute over a third, non-spatial axis. For the latter, if the dynamic dataset is generated with respect to the temporal axis, the first derivative can be calculated and used to determine the degree, or magnitude, of change that occurs at a spatial level per epoch, and the second derivative can be used to determine stability in the rate of this change. These processes are represented by the functions $f_i'(t)$ and $f_i''(t)$ respectively, where i is the i^{th} polygon in the geometry layer, and t is time. For the first and second derivative functions, the global styling pool approach is adopted in order to accurately reflect change over time.

4. VISUALISATION

A number of sample data sets, encompassing both health and related socio-economic were used to test the efficacy of both generating and viewing the data encapsulated by the visualisation objects.

Flexibility is paramount when displaying the visualisation object due to the dynamic nature of the object creation. The data contained in an object is not pre-determined, and consequently should facilitate multiple visualisation methods for the same object data sets in order for data exploration to be effective. A number of examples of such visualisations have been explored within a web GIS as case studies, including a spatiotemporal visualisation, a movie like map visualisation, and a simple table access method.

Both spatial and non-spatial visualisations were considered, along with linked visualisation. The output of the visualisation object comprises vector data, such as GeoJSON, which can then either be rendered as a map using a dynamic web mapping server [12], or rendered as a graph visualisation using a scalable vector graphics JavaScript API such as the one provided by the Data Driven Documents (D3) JavaScript libraries.[‡] For the latter, meta data is required specifying the format of the data as either temporal data, or data corresponding to a single epoch. Thematic map visualisations were generated using the previously proposed dynamic web mapping server [12], and visualised using the GeoExt [§] wrapper on the OpenLayers slippy map client.

[‡] `http://d3js.org`, accessed 16^{th} August 2013

[§] `http://geoext.org`, accessed 2^{nd} September 2013

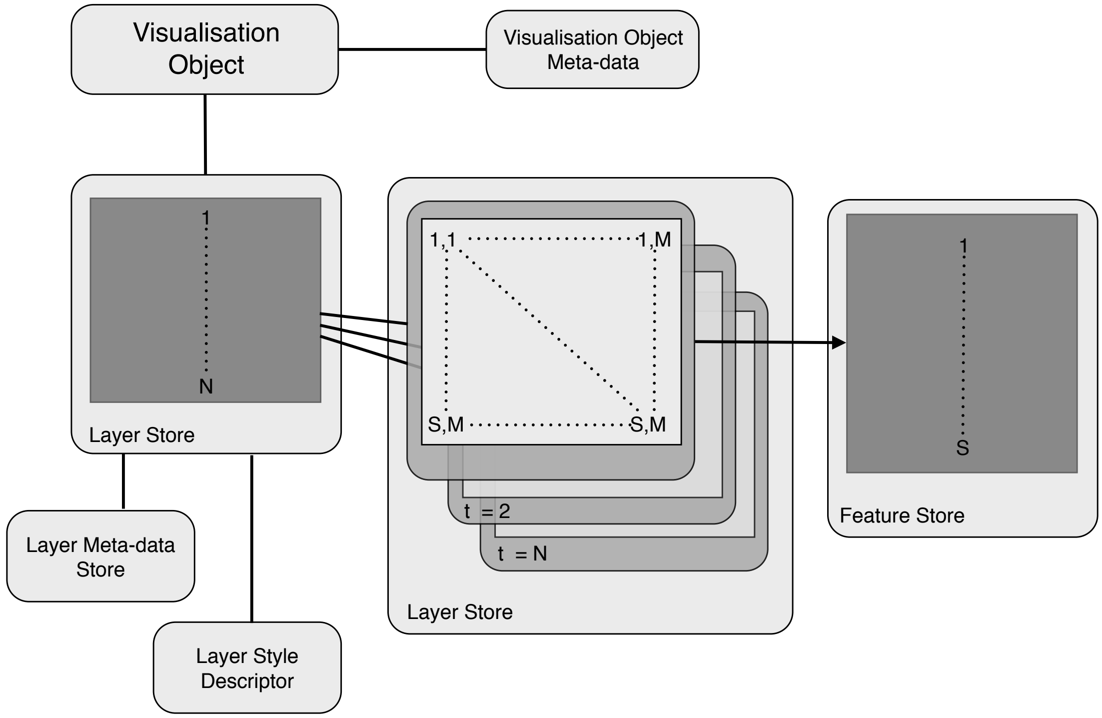


Figure 2: The visualisation object data store.

4.1 Data

To test the spatiotemporal model, a number of datasets were imported into a PostGIS database, using an extension to the the dynamic web map server proposed in [12] to enable access.

The geometry layers used comprise the Australian Bureau of Statistics (ABS) Statistical Areas (SA). The ABS reports over four different spatial resolutions, $SA1$ to $SA4$ corresponding to lowest and highest level of the hierarchy respectively, with groups of geometries within each level being encapsulated within a single geometry at a higher level. That is, there is a direct parent child relationship between SA geometry levels. The geometries represent spatial aggregation areas that are hierarchical, with the highest (lowest resolution) being $SA4$

The test datasets comprise ABS median data and ABS demographic data, extracted from ABS data packs, and a synthetic health data set. The ABS median data consisted of socio-economic attributes such as median individual, household and family incomes, monthly mortgage repayments, and median weekly rent. The demographic data corresponds to population counts per SA for a given age and gender. The health data set comprised a morbidity data set designed to simulate one year of hospitalisation events in Western Australia, comprising approximately $700,000$ events, stored as both point data, and unit record data spatially aggregated to the ABS $SA2$ statistical area.

To generate the composite data set, a number of default socio-economic indicators, such as income and household size, along with general demographic information, the population size and density, were concatenated onto the base attribute, accounting for the temporal axis where necessary. The default datasets are accessed through the dynamic map server that facilitates access to the Australian Bureau of Statistics (ABS) census data for the last three census years, 2001, 2006, 2011, at three geographic resolutions corresponding to the ABS statistical areas $SA2$, $SA3$, and $SA4$. It should be noted that in using a dynamic query driven approach to access the data, both the spatial resolution and time are input parameters. The process output and incorporated data sets can be adjusted through the inclusion of keywords in the input REST query string to trigger a search of the census data source, for example, education and employment data, and population density for subsets of the population can also be accessed.

4.2 Spatiotemporal Visualisation

Two visualisations were implemented, using D3, to explore different aspects of visualising spatiotemporal data.

The first example spatiotemporal visualisation method combines visualisations that considered the subject data both as individual layers and as a single data source; Figure 3 shows the outcome of the visualisation, and comprises a linked visualisation of a map view, and a gap minder[¶] style temporal plot using the $D3$ motion map API, with the map data reflecting the year displayed on the scatter plot. The map visualisation shows the thematic map corresponding to the median monthly mortgage repayment, while the graph graph shows the median monthly mortgage repayment (map and $y - axis$), the median weekly income ($x - axis$), and population size (radius) for Western Australia using the Australian Bureau of Statistics Statistical Area 2 geome-

[¶]Gapminder: `http://www.gapminder.org`, accessed: 29 June 2013 try layer boundaries. In this case, the visualisation object was initially created by selecting the median monthly mortgage repayment attribute, and subsequently appending the median individual income, and the total population counts per statistical area.

The visualisation was created by simultaneously generating a GeoExt map object and a D3 graph within the web portal. The thematic map visualisation was generated by adding a dynamic map layer for each epoch, or census year, in the data set. This was achieved by generating WMS requests and injecting a year filter parameter, and subsequently incrementing over census years. The data for the graph corresponds to JSON output from the visualisation object data agent accessed through an AJAX REST request; the JSON output returns the attributes stored by year. The map and graph visualisations are linked, the thematic map shown reflects the epoch displayed within the graph, with changes in the graph year triggering a changed in the displayed thematic map. In this case the map epoch correspond to the census years within the database, while the graph is interpolated between epochs to enable fluid animation within the graph. Smooth transitions within the slippy map client can be achieved by progressively alpha blending between map epochs, with the blending linked to changes in the year displayed on the graph.

The colour scheme is also used to link the map and graph visualisation, in this case the colour is derived from the ColorBrewer [3] PurpleBlueGreen colour scheme, using the *Natural Breaks* map classification algorithm [13], and is applied to both the thematic map colouring, and the colouring of the graph elements. Alternatively, the graph can be coloured by spatial region. Consequently, the second visualisation example, depicted in Figure 4, shows the corresponding graph visualisation shown in Figure 3, but with colouring determined by spatial area, in this case achieved using the $SA4$ region classification. While the visualisation was generated from the same visualisation object, the resulting graph more clearly reveals the spatial distribution of the mortgage repayments and median income, with a main cluster depicting a positive correlation between mortgage repayments and income, while revealing two anomalous regions. The purple region, corresponding to the $SA4$ *Wheat Belt* area within Western Australia has a visibly negative correlation between income and mortgage repayments. Further, a number of outliers exist within the $SA4$ *Outback* region (pink), with high income, but relatively low mortgage rates compared to the rest of the state; using a temporal visualisation, the outliers become more distinct over time.

4.3 Multivariate Visualisation

A number of multivariate visualisations were explored. In order to examine potential aspects of data exploration using visualisations objects, the visualisations focussed on multiple visualisations of a single visualisation object, and incorporating interactivity leveraging elements of HTML 5 and scalable vector graphics. The resulting visualisation represent multiple views of the same information, with the ability to interactively filter the data displayed in within the visualisation.

The first example of a multivariate visualisation is shown in Figure 5, a parallel coordinate plot displaying a spatial health summary statistic, along with a number of socio-economic and demographic attributes, generated using the parallel coordinate API within D3. The direct age standardised rate, ASR, was used to determine the health summary statistic, and was calculated for Chapter K of the International Classification of Disease ($ICD10$) categorisation scheme, corresponding to diseases of the digestive system, with the at risk population being determined using the ABS demographic information. The spatial resolution used corresponds to the ABS $SA2$. In this case, the visualisation object was initialised using output from a processing web service determining the summary statistic for the hospitalisation event associated with Chapter K within the database. The ABS median data, and demographic data, the population count, were subsequently incorporated into the visualisation object. The colouring of the line plot corresponds to the red, yellow, green ColorBrewer [3] colour scheme, with map classification determined over the ASR using the *Natural Breaks* map classification algorithm [13]; using this approach, the parallel coordinate plot can be linked to a similarly styled thematic map layer. Figure 6 shows the same visualisation as Figure 5, with a filter applied to the ASR, resulting in only the statistical areas with an ASR within the filter band being displayed on the graph. Using this visualisation method, multiple filter bands can be applied simultaneously, facilitating the exploration of complex relationships between attribute dimensions. The visualisation methods also has the functionality to present the tabular representation of the data depicted within the parallel coordinate plot.

Figure 7 shows a segment of a larger scatter plot matrix generated using the same visualisation object created for Figure 5, generated using the scatter plot matrix API for D3. The colour scheme applied to the data points also corresponds to the colour scheme applied in Figure 5. The portion of the scatter displayed shows the ASR for $ICD10$ Chapter K, the associated confidence interval, or uncertainty, and the median weekly individual income and rent, calculated at an ABS $SA2$ resolution. The scatter plot matrix also enables filtering, with data point selected in one plot being highlighted in the remaining plots. The combination of presenting multiple scatter plots in the form of a matrix, and enabling filtering, facilities the rapid determination of the relationship that exists between attribute pairs. Thus, the parallel axis plot, and scatter represent different methods of viewing the same information, with each revealing different aspects of the data.

A number of additional visualisation and interface techniques were also explored to further determine the efficacy of the visualisation object model. One such interface technique consisted of the use of scalable vector graphics to embed further information within a visualisation. A further visualisation comprised a movie like thematic map layer presentation, generated using age as a non-spatial axis, rather than the usual temporal axis. The population information was divided into population density at 5 year increments, which resulted in 16 map layers that were incorporated into a thematic map layer visualisation cycling through the layers by applying alpha blending between adjacent layers, thus producing a movie like visualisation showing the change in the population density with respect to age. This technique can similarly be applied to any attribute data access that supports an incremental filter, for example, the increment can be also constructed over time.

The visualisation object was also explored as a means of accessing data. Consequently, a method for generating a tabular *visualisation* of the data contained within a visualisation object was developed. The method enables tabular access to the data, which can then be exported to CSV format. Using this method, a user can have offline access to the information with a visualisation object, and can conduct analysis in their preferred research environment. Thus, the visualisation object can be used to generate an analysis data set comprising health summary statistics, and related variables.

5. CONCLUSION

The visualisation object represents the generation of a set of related variables on the fly, and while the current variables integrated with the base attribute are broad, it represents the first stage of integrating health and related variables for multivariate analysis. For example, related disease codes, determined through semantic search techniques, can be automatically processed an integrated into a single analysis data set; additionally, specific disease, or condition, types can be matched to relevant sets of related variables, as determined by health experts. However, due to the degree of flexibility inherent in the approach, for example, the ability to persist a dynamically generated data set and link subsequent data or processing outputs, a number of usability issues necessarily arise. Relevant issues include developing methods for the specification of data sets to incorporate into the model, and determining the current components of the object in cases where the data is persisted. Further, standardisation of the communication protocols required to leverages the functionality of the proposed method is also required. A detailed discussion addressing such points is beyond the scope of this paper, but it is acknowledged that research into such usability and standardisation issues is required in order to incorporate the flexibility of the proposed methods into a web GIS portal.

ACKNOWLEDGEMENTS

This work has been supported by Curtin University of Technology, and the Cooperative Research Centre for Spatial Information, who activities are funded by the Australian Commonwealth's Cooperative Research Centres Programme.

6. REFERENCES

[1] Visual analytics: Definition, process, and challenges. In *Information Visualization: Human-Centered Issues and Perspectives*, pages 154–175, 2008.

[2] G. Andrienko, N. Andrienko, D. Keim, A. M. MacEachren, and S. Wrobel. Editorial: Challenging problems of geospatial visual analytics. *J. Vis. Lang. Comput.*, 22(4):251–256, Aug. 2011.

[3] C. Brewer. Colorbrewer, www.colorbrewer.org, 2012.

[4] J. Cinnamon, C. Rinner, M. D. Cusimano, S. Marshall, T. Bekele, T. Hernandez, R. H. Glazier, and M. L. Chipman. Online map design for public-health decision makers. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 44(4):289–300, 01 2009.

[5] D. J. Exeter, S. Rodgers, and C. E. Sabel. "whose data is it anyway?", the implications of putting small area-level health and social data online. *Health Policy*, 2013.

[6] D. H. Foley, R. C. Wilkerson, I. Birney, S. Harrison, J. Christensen, and L. M. Rueda. Mosquitomap and the mal-area calculator: new web tools to relate mosquito species distribution with vector borne disease. *Int J Health Geogr*, 9(11), 2010.

[7] S. Gao, D. Mioc, X. Yi, F. Anton, E. Oldfield, and D. Coleman. Towards web-based representation and processing of health information. *International Journal of Health Geographics*, 8(1):3, 2009.

[8] M. Jern. Collaborative web-enabled geoanalytics applied to oecd regional data. In *Proceedings of the 6th international conference on Cooperative design, visualization, and engineering*, CDVE'09, pages 32–43, Berlin, Heidelberg, 2009. Springer-Verlag.

[9] K. Joyce. "to me it's just another tool to help understand the evidence": Public health decision-makers' perceptions of the value of geographical information systems (gis). *Health & Place*, 15(3):831–840, 2009.

[10] R. Kamadjeu and H. Tolentino. Web-based public health geographic information systems for resources-constrained environment using scalable vector graphics technology: a proof of concept applied to the expanded program on immunization data. *International Journal of Health Geographics*, 5(1):24, 2006.

[11] A. M. MacEachren, S. Crawford, M. Akella, and G. Lengerich. Design and implementation of a model, web-based, gis-enabled cancer atlas. *Cartographic Journal, The*, 45(4):246–260, 2008.

[12] S. Moncrieff, G. West, J. Cosford, N. Mullan, and A. Jardine. An open source, server-side framework for analytical web mapping and its application to health. *International Journal of Digital Earth*, 0(0):1–22, 2013.

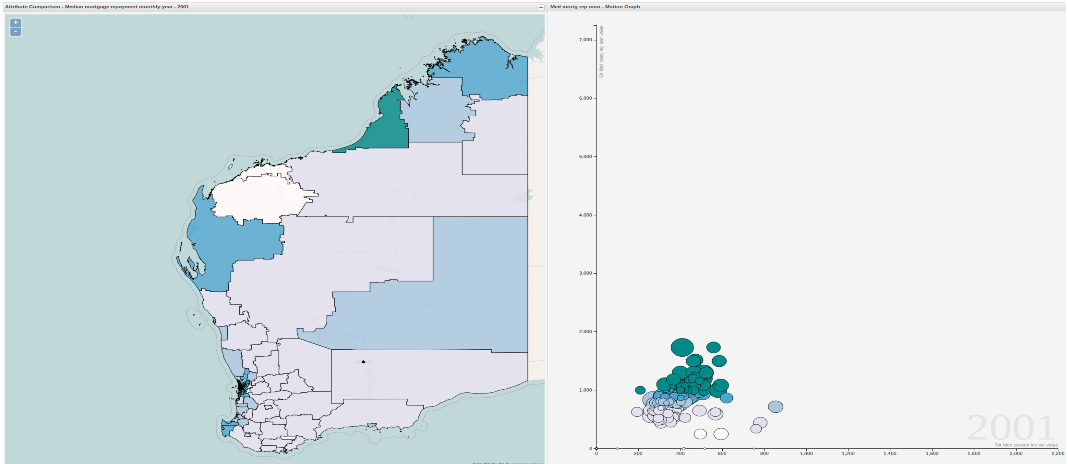
[13] S. Rey and L. Anselin. Pysal: A python library of spatial analytical methods. In M. M. Fischer and A. Getis, editors, *Handbook of Applied Spatial Analysis*, pages 175–193. Springer Berlin Heidelberg, 2010.

[14] E. Samarasundera, T. Walsh, T. Cheng, A. Koenig, K. Jattansingh, A. Dawe, and M. Soljak. Methods and tools for geographical mapping and analysis in primary health care. *Primary Health Care Research & Development*, 13(01):10–21, 2012.

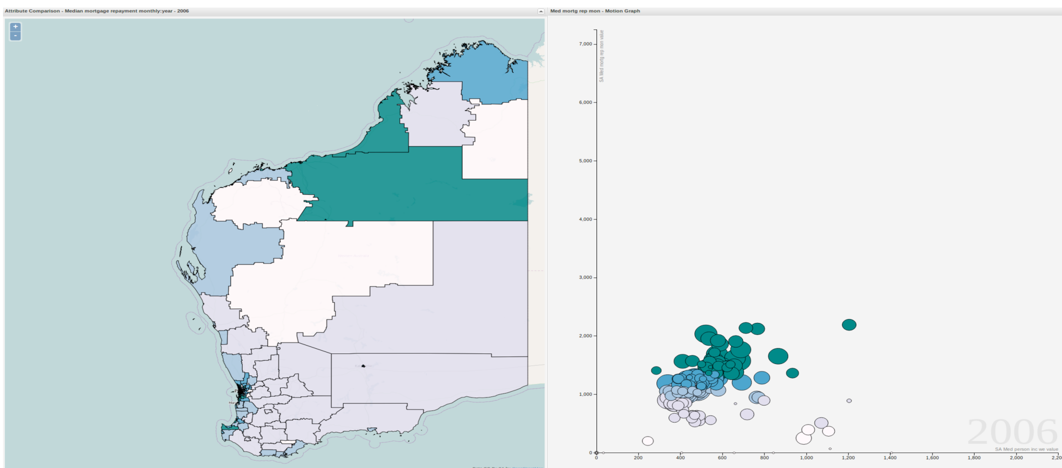
[15] X. Shi. *The dynamic integration of distributed gis through semantic web services*. PhD thesis, Morgantown, WV, USA, 2007. AAI3298738.

[16] Q. Van Ho, P. Lundblad, T. Astrom, and M. Jern. A web-enabled visualization toolkit for geovisual analytics. *Information Visualization*, 11(1):22–42, Jan. 2012.

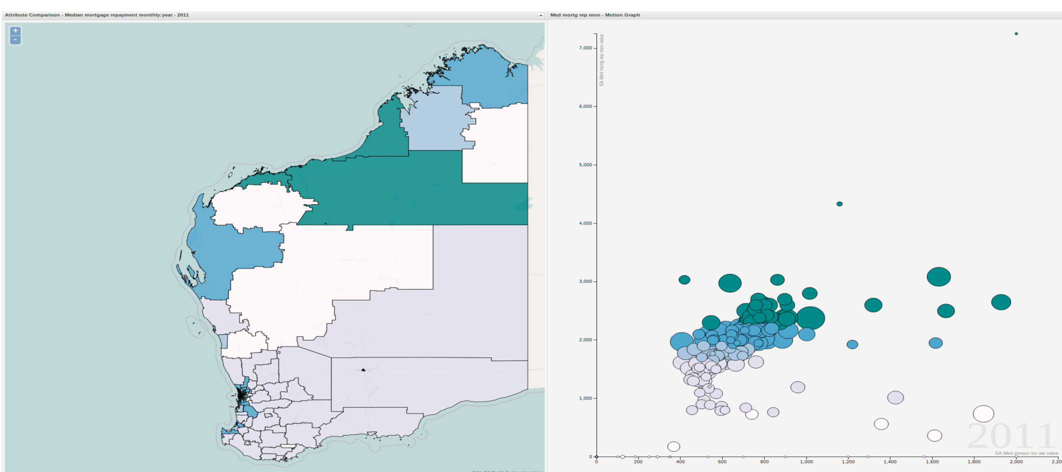
[17] Q. Yi, R. E. Hoskins, E. A. Hillringhouse, S. S. Sorensen, M. W. Oberle, S. S. Fuller, and J. C. Wallace. Integrating open-source technologies to build low-cost information systems for improved access to public health data. *International Journal of Health Geographics*, 7(1):29, 2008.



(a)



(b)



(c)

Figure 3: Map and Gapminder visualisation depicting the median monthly mortgage repayment (map and $y - axis$), the median weekly income ($x - axis$), and population size (radius) for Western Australia using the ABS $SA2$ geometry layer boundaries, for (a) 2001, (b) 2006 and (c) 2011. The thematic map shows the median monthly mortgage repayment (base map: OpenStreetMap – `http://www.openstreetmap.org/copyright`).

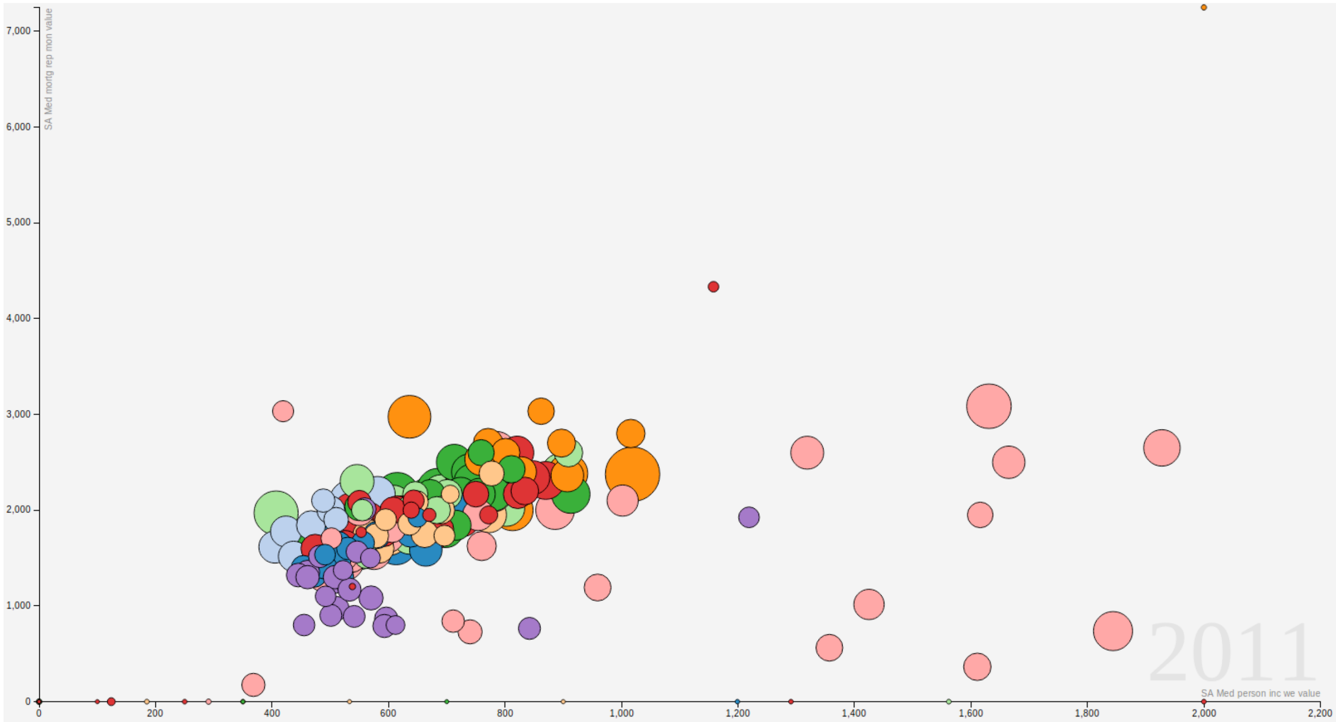


Figure 4: Gapminder like visualisation, D3 motion graph, depicting the median monthly rental payment ($y - axis$), the median weekly income ($x - axis$), and population size (radius) for Western Australia using the ABS $SA2$ geometry layer boundaries, with the colour scheme determined by spatial region, corresponding to $SA4$.

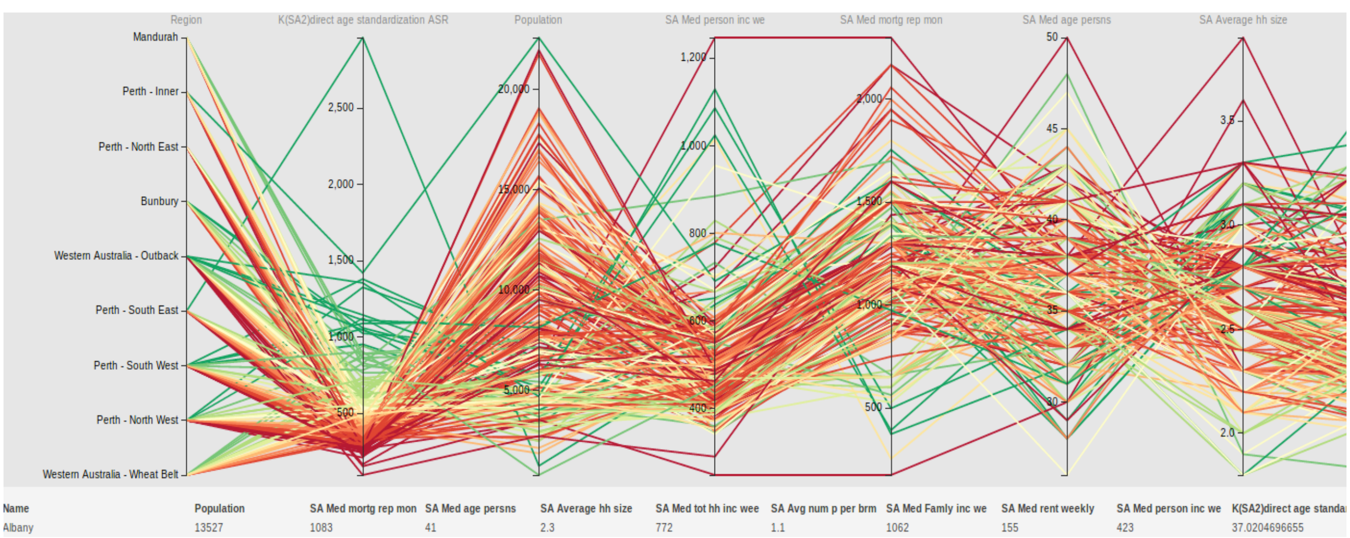


Figure 5: Parallel coordinate plot depicting a statistical summary of health data, the direct age standardised rate for the International Classification of Disease ($ICD10$) Chapter K (diseases of the digestive system), along with a number of socio economic variables. The top portion of a tabular representation of the data is also shown.

Figure 6: Parallel coordinate plot depicted in Figure 5, with a filter applied to the direct age standardised rate.

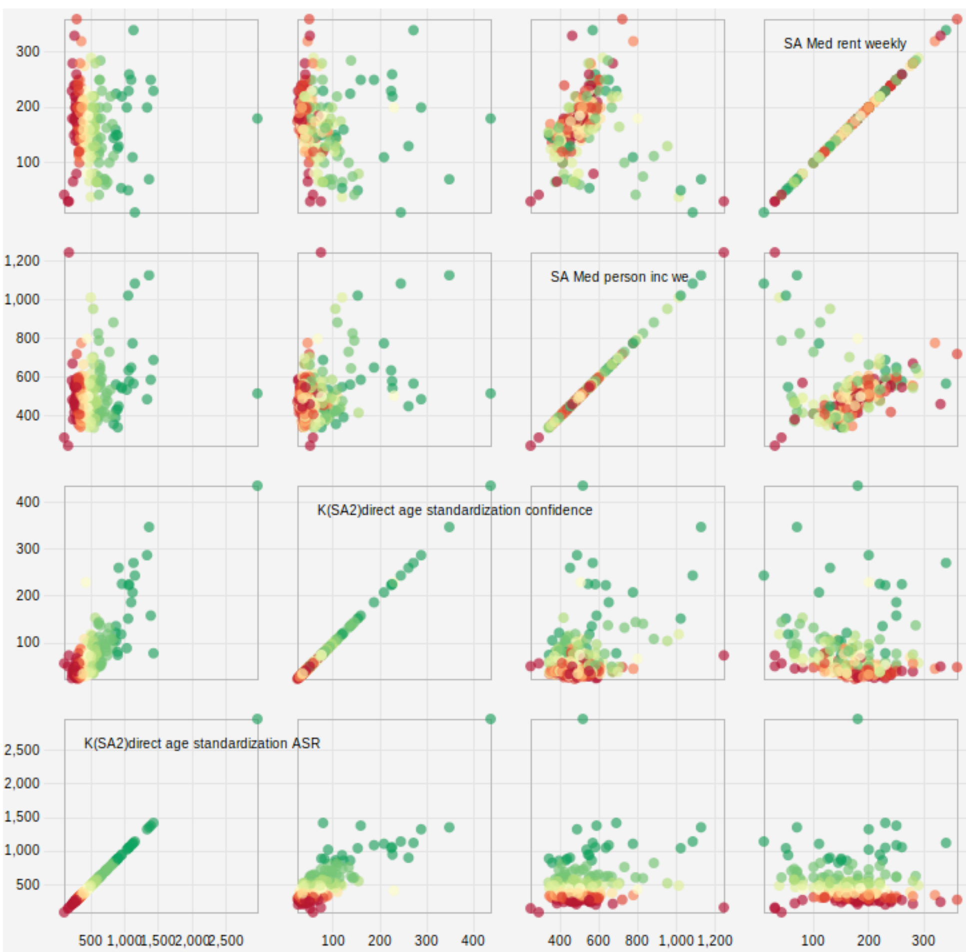


Figure 7: Scatter plot matrix showing the ASR for $ICD10$ Chapter K (diseases of the digestive system), the associated confidence interval, and the median weekly income and rent, at ABS $SA2$ resolution.