Archaeological Potential Modelling for the Capital Regional District

Prepared for:
Capital Regional District
and
The Archaeology Branch

By:
Alyssa Parker, BSc

Millennia Research Limited
510 Alpha Street
Victoria, BC
V8Z 1B2
(250) 360-0919

September 5, 2008
Project #MR0808

Acknowledgements

Millennia Research Limited would like to acknowledge the following persons or groups for their invaluable assistance in this project:

- Jeremy Bart and Amy Chen of the CRD GIS department, for providing the base mapping data
- James Hamly, Director of the South Island Aquatic Stewardship Society, Carol Hyland of Muze Creative, and Rob Struthers of Dennis & Struthers Visual Communications Inc. for their assistance in tracking down data for the "Lost Streams of Victoria" map (Sutherst 2003)
- Doug Glaum of the Archaeology Branch provided guidance and facilitated communications among the various parties involved
- Tracy Corbett of the CRD initiated the project

Table of Contents

List of Tables

List of Figures

Introduction

Millennia Research Limited (Millennia) was requested by the British Columbia Archaeology Branch and the Capital Regional District (CRD) to create a model of archaeological potential for the CRD. Similar models have been created in the Southern Vancouver Island region by Millennia (Eldridge and Parker 2006, 2007); knowledge gained during the creation of these prior models was used as the basis for creating a model which would cover the CRD.

The purpose of these models is to indicate areas on the landscape where it is more likely that an archaeological site might be located. Traditional use of the CRD landscape by the First Nation peoples was quite extensive – archaeological sites are not reflective of the full diversity and extent of these activities but are locations where there is tangible, physical evidence of this use. Unlike traditional use sites however, archaeological sites which pre-date 1846 are protected by the provincial *Heritage Conservation Act*. The intent of this model is to provide municipal planners within the CRD with a tool to guide future development in a manner which considers archaeological potential.

Archaeological sites in the CRD have been found to be highly correlated with certain landforms, including water features such as rivers and lakes, and terrain features such as the relatively flat tops of small knolls or ridges. Modelling therefore was focused on identifying these features and combinations of these features which would be higher potential for archaeological sites. Recorded sites in the region were used to analyze the model performance.

Study Area

The area modelled included most of the CRD, excluding the Gulf Islands and the Juan de Fuca Electoral Area (Figure 1). First Nations who would have traditionally used this region include the Tsawout, Tsartlip, Pauquachin, Tseycum, Songhees, Esquimalt, T'Souke, Scia'new, and Malahat Nations.

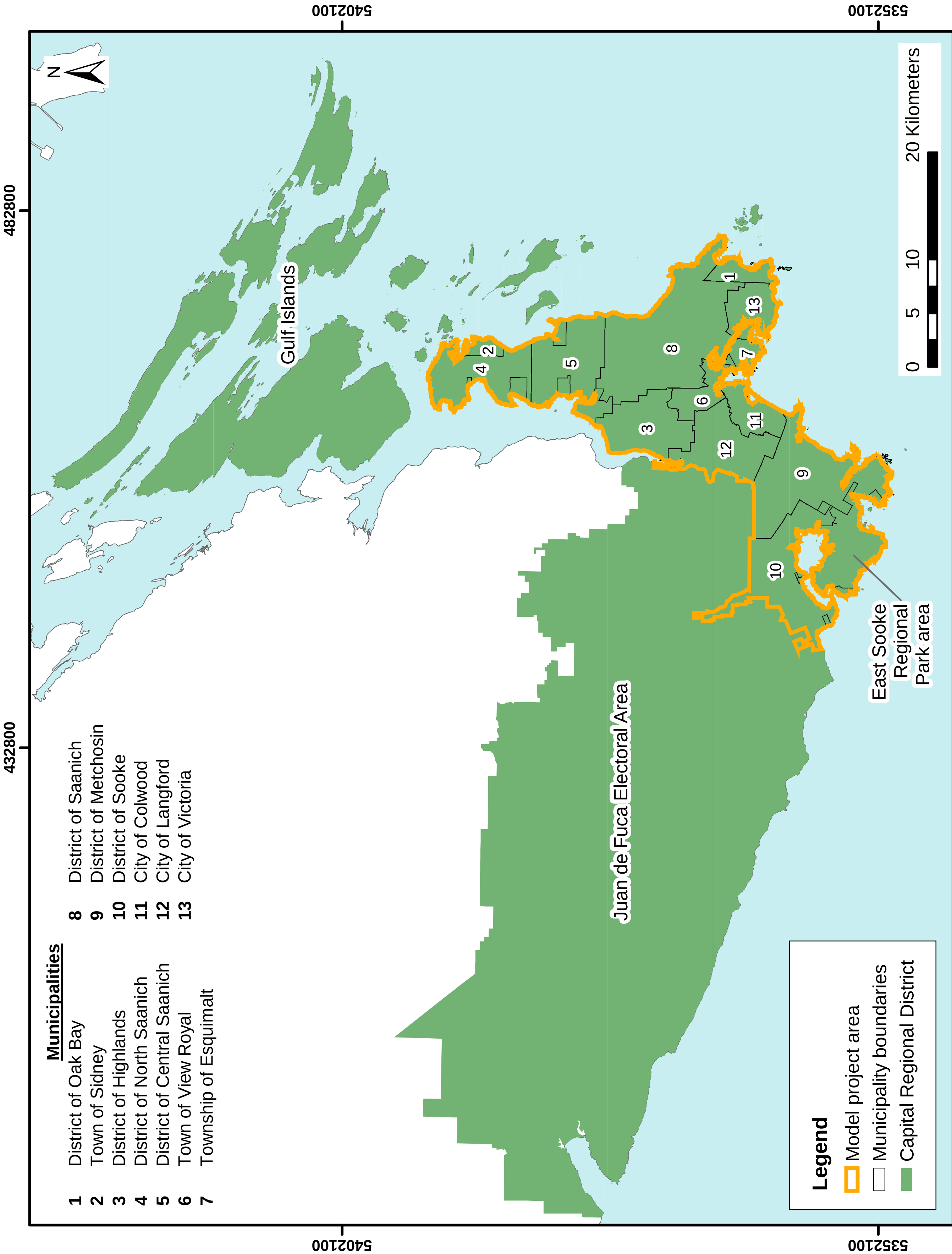
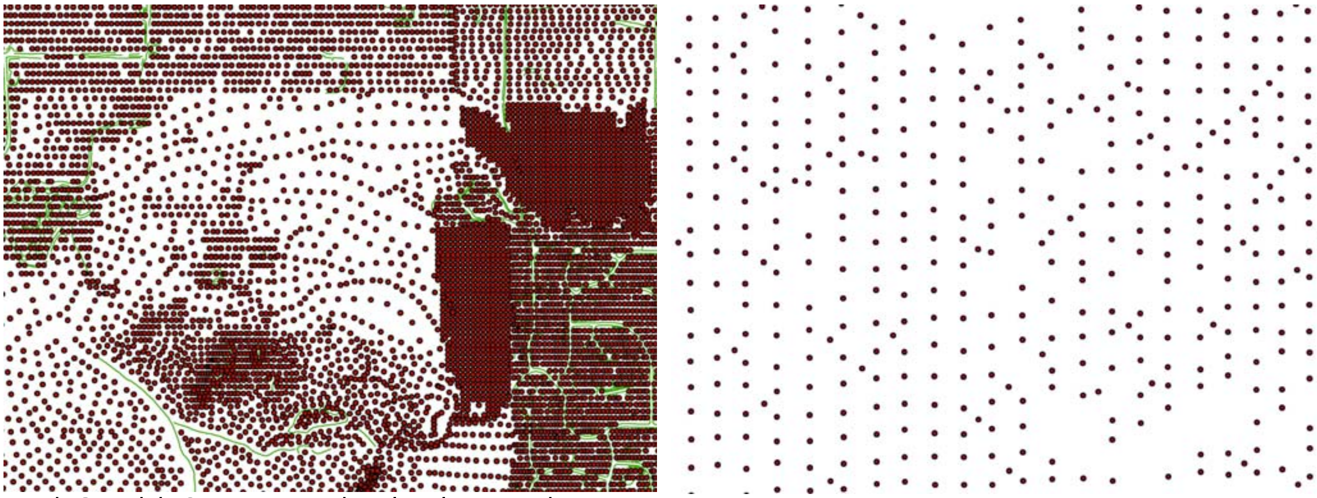


Figure 1. Study Area.

Data Sources

The data used for modelling consisted of several datasets including water feature mapping, FISS (Fisheries Information Summary Systems) streams, Garry Oak mapping (1800s and 1990s), and DEMs (Digital Elevation Model). Two DEM sources were used to identify topographic features with archaeological potential: one created by photogrammetry of recent (2005) orthophotography for the CRD, and one from TRIM (Terrain Resource Information Management) mapping. The former DEM is much higher resolution than the TRIM data. The raw data consists of mass elevation points and breaklines from which a raster DEM was created. The mass elevation points for the CRD DEM range between five to 50 m spacing, whereas the TRIM DEM points range from 35 to 100 m spacing (Figure 2).



North Saanich CRD DEM – the closely spaced DEM points in the right-hand part of the image average between 10 and 20 m apart.

Highlands TRIM DEM – DEM points range from 35 to 100 m apart, with an average of 80 m between points.

Figure 2. These two images, of different locations but at the same scale, highlight the differences between the CRD DEM (on the left, with breaklines shown in green) and the TRIM DEM (on the right).

The CRD DEM allows much smaller topographic features to be modelled. However, this detailed DEM was unavailable for two regions within the study area: the District of Highlands, and a portion of East Sooke (most of which is within East Sooke Regional Park). For these regions, the TRIM DEM was used. The raster TRIM DEM was built from the raw elevation points. Though breaklines were provided, the layer did not contain the necessary elevation information, and could not be used in generating the raster DEM. Should more detailed DEM data (such as LiDAR) become available in the future, the model could be applied to those areas at a fairly low cost.

Several variables were then extracted from the raster DEMs. These include common terrain variable such as slope, aspect, and range (of elevation or slope). They also include less common variables, which were extracted by custom AML (Arc Macro Language) scripts which were designed by Millennia in conjunction with Timberline Natural Resource Group. These scripts consist of a series of moving windows which analyze the input DEM. The moving windows include a 25 by 1 cell horizontal window, a 1 by 25 cell vertical window, and a 9 by 9 cell square moving window. The results produced from the analysis include (Figure 3):

- "Positive" – this refers to the sum of the differences (in meters elevation) between the central cell in the window and all cells with lower elevations than the central cell. This is run using both the horizontal and vertical 25 by1 moving windows and the 9 by 9 window
- "Negative" – this is similar to the "Positive", except that it calculates the sum of the negative differences (i.e. between the central cell and all cells with higher elevations than the central cell). This is run with both 25 by 1 moving windows.
- "Positive count" or "Count" – this is the count of all cells with elevation values lower than the central cell. This is run using the 9 by 9 moving window for which there is an 81 cell maximum (9 x 9).

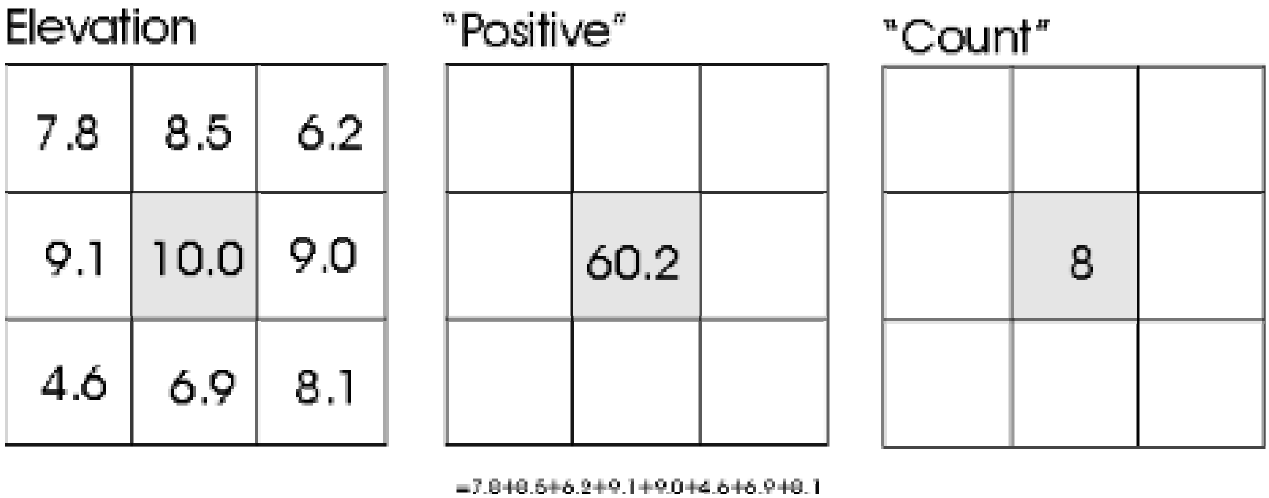


Figure 3. Simplified example of how "Positive" and "Count" values are determined for a DEM cell.

Modelling Methods

Modelling was initially based on findings from previous modelling projects in the region which had identified certain locations in the terrain where archaeological sites were more likely to be found (Eldridge and Parker 2006, 2007). These locations included: areas within a certain distance to water features, such as lakes, streams, and wetlands; terrain features, such as the tops and bases of steep slopes, the tops of knolls and ridges; areas within a certain distance of a fish stream; and areas within a certain distance to the coastline.

Water features, including fish streams and the coastline, were buffered to distances which had been found to be significant in the location of archaeological sites in previous modelling in the region. However, these original buffer distances were scaled down, as the better DEM for the CRD model allowed less reliance on water feature buffering to model for archaeological potential. The buffers were reduced to a scale that captured the significant water features without capturing too much land. The one exception was the coastal buffer, which was increased due the extremely high numbers of sites found on or near the coastline. Table 1 shows the actual buffers used for the water features.

Table 1. Water features buffer distances used in modelling.

Feature	Buffer distance (m)	Layer name
Coast (with original shoreline)	50	coastbuf50
Definite single rivers (with lost streams)	50	defsglriv50
Lakes	50	lake50
FISS stream	100	fiss100
Wetlands	50	wtlnd50
Double-line rivers	100	dblriv100

Terrain modelling was done in two parts – the first part was using the detailed DEM (referred to below as the main portion of the CRD), the second part using the TRIM for Highlands and a portion of East Sooke, where the detailed DEM was not available. This was necessitated by the difference in resolution of the two types of DEMs, as the values for terrain features from one DEM would not be transferable to the other. The models are discussed separately below.

A draft model was provided to the CRD based on the above methods. After initial modelling and the production of a draft model and report, further work was done to improve the model. The primary focus of this was obtaining the digital data for the "Lost Streams of Victoria" mapping project (Sutherst 2003). This map indicates a number of "lost" streams (streams which have been buried or diverted through culverts), wetlands, and original shorelines for the Greater Victoria area. This data was successfully obtained as digital line files, which were then correctly scaled and georeferenced. A new coastline layer was created by replacing the shoreline as currently mapped for the CRD with the original shoreline in the Inner Harbour/Gorge waterway areas and Esquimalt Harbour areas. This new layer was then buffered to 50m and replaced the coastline used in the initial version of the model. Lost streams were added to the single definite river layer, as were areas identified as former wetland. This layer was then buffered to 50m, and replaced the single definite river layer as it was used in the initial model (Figure 4).

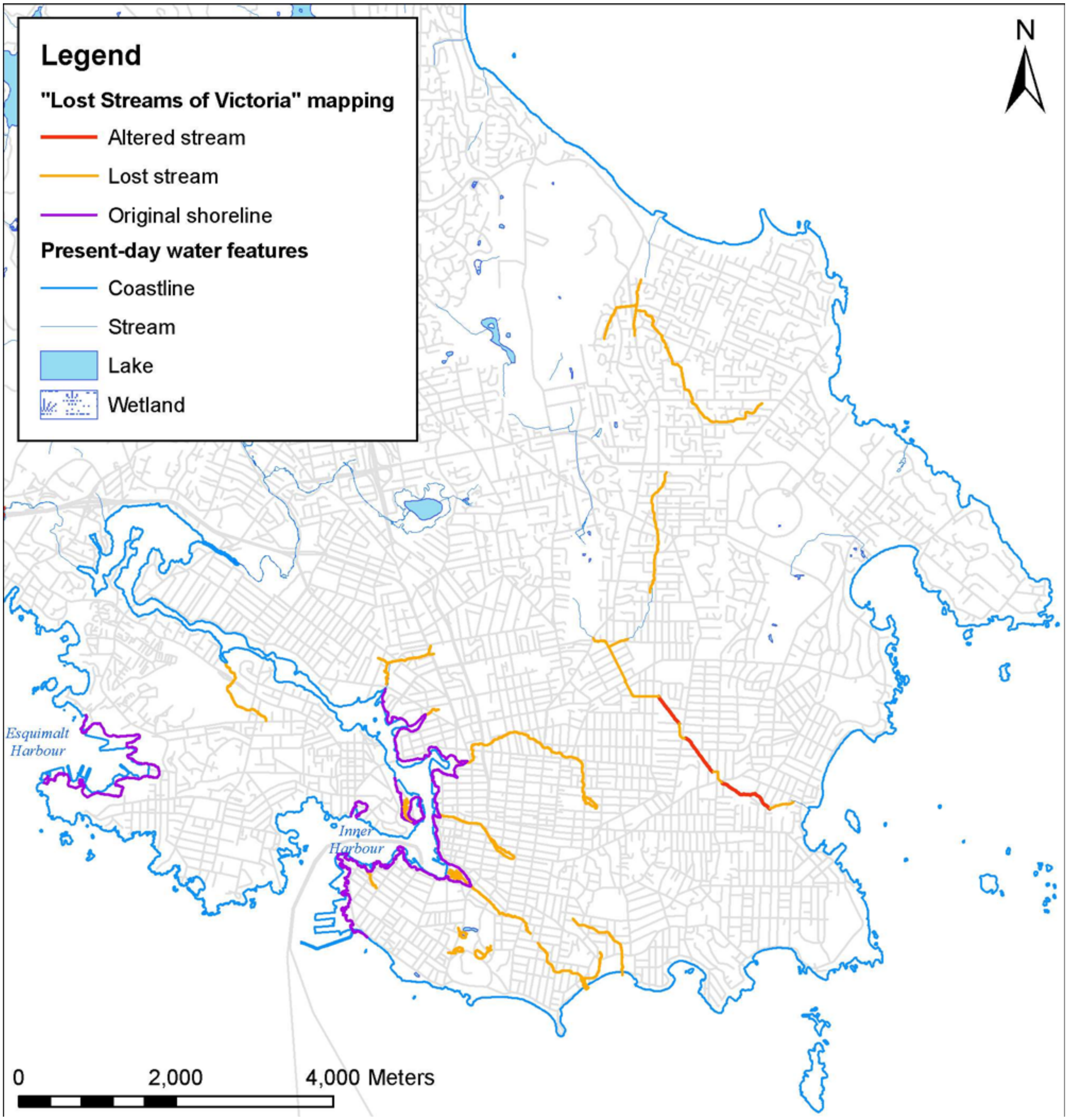


Figure 4. "Lost Streams of Victoria" mapping data incorporated into model.

Main portion of the CRD

Terrain features were identified using the layers which had been generated from the DEM. A variety of combinations of these layers was tested in an attempt to capture the types of landforms that were of interest as having archaeological potential. The resulting algorithms are shown in Table 2. Distribution of archaeological sites in the region suggests that confluences of water features have higher potential for archaeological sites; the model was therefore weighted to reflect this by combining the water feature buffers in two ways. The first was to identify areas where a stream entered a lake or the ocean, while the second was simply to identify all areas which were covered by at least one of the water feature buffers (Table 3).

Table 2. Terrain features and algorithms used to identify them.

Landform feature	DEM variables used to identify the feature
Tops and bases of steep slopes (rockshelters and lookouts)	Areas with a slope less than or equal to 15 degrees, which were within 25m of a slope greater than or equal to 30 degrees
Knolls and ridge tops	A positive count greater than or equal to 55 was combined with areas where the 9 by 9 elevation range was between (inclusively) 3 and 9 m, and also combined with slopes less than or equal to 6 degrees
Small wetlands (buffered to 20 m → smwtlnd2_20)	A positive count of less than or equal to 15 was combined with a 50 by 50 elevation range >= 20 m. This was further limited by combination with a 4 by 4 slope range > 4 (i.e. very flat areas removed), and by limiting the layer to areas with a slope of <= 15 degrees. This identified streams, wetlands, and small water bodies, not all of which appear in TRIM map features.

Table 3. Algorithms used to build water components of the model.

Model	Algorithm
Rivers/lakes/coast confluences model (This adds potential where a river enters a lake or the ocean.)	([defsglriv50] = 1 OR [dblriv100] = 1 OR [fiss100] = 1) & ([lake50] = 1 OR [coastbuf50] = 1)
Water model	[coastbuf50] = 1 OR [dblriv100] = 1 OR [defsglriv50] = 1 OR [fiss100] = 1 OR [lake50] = 1 OR [smwtlnd2_20] = 1 OR [wtlnd50] = 1

These layers were then added up using map algebra, with values assigned to each layer to determine their weighting in the final model (Table 4).

Table 4. Model components and their model values.

Model	Final Model Value
Confluences model	+1
Water model	+1
20m buffer small wetlands	+1
Top/base slope	+2
Knoll/ridge	+2
Aspect 120° to 330°	+1

Areas with a model value of two or higher were included in the final model. This means that water feature buffers and aspect were only included where they overlapped with another modelled feature, whereas the terrain models would be definitely included. Areas with a slope of greater than

15° were excluded. The final model was then clipped to exclude areas which would have false potential – for example the docks at CFB Esquimalt, buffered highways and roadways (to remove ditches and embankments, on/off ramps etc.), artificial ponds and reservoirs, and other modified areas. This was done in part by buffering roads, using a 30 m (on each side) buffer of freeways, 20 m buffer of highways, 10 m buffer of intermediate roads and ramps, and a 3 m buffer of lanes. Manual digitizing was done of other areas, such as the docks, and anything missed in the road buffers (particularly at interchanges). Single pixels were also filtered out.

East Sooke and Highlands

The Highlands and East Sooke Park area DEM models were composed a bit differently, though the water models were the same, and the DEM models were based on capturing the same types of terrain features as captured by the other model. The combination of the variables used for these two regions are listed in Table 5.

Table 5. Variables, algorithms, and model values for the Highlands and East Sooke models.

Variable	Algorithm	Final Model Value
Confluences model	Same as in Table 3	+1
Water model	Same as in Table 3	+3
50m buffer small wetlands	Positive count ≤ 15 & 25 by 25 elevation range ≥ 20 & 3 by 3 slope range > 4 & slope ≤ 15°	+1
Elevation ≤ 10		+1
9 by 9 range of elevation ≤ 20		+1
Aspect 120° to 330°		+1
Positive 25 ≥ 50		+1
Positive 9 ≥ 100		+1
Count 9 ≥ 40		+1
Negative 25 ≥ -200		+1
	Total value possible	*12*
Top/base slope	30 m buffer of slopes ≥ 20°, & slope ≤ 15°	Added as high potential

The final model for East Sooke included values of eight or higher as high potential, and for the Highlands values of seven or higher were included. The top/base slope model was then added as high potential, then the models were limited to areas with slopes of ≤ 15°, and single pixels were filtered out.

After assessing the model performance, as described below, the final step is to combine all the models into one then add all the recorded archaeological sites into the model as high potential. The sites are simply converted to a raster with a cell size of 5 m, and then an “OR” statement is executed between the high potential in the model and the sites raster. The result then becomes the final model.

Model Results and Discussion

Model performance was analyzed by comparing the percentage of sites captured to the percentage of land area captured, as described by Kvamme's Gain Statistic (Kv Gain) (Kvamme 1988). Kv gain is a useful indicator of the performance of archaeological potential models.

$$\text{Kv Gain} = 1 - \left(\frac{\%\ \text{land captured}}{\%\ \text{sites captured}}\right)$$

The maximum Kv value possible is 1.0 which would indicate all sites captured in land limited only to the area of the sites themselves. A high Kv gain is therefore indicative that the model captures a majority of sites, while minimizing the land area modelled as high potential.

Previous modelling projects have found that sites are often found to be mis-mapped by a few to several hundred meters (Eldridge and Parker 2006; Eldridge, et al. 2007). This may be due to errors in translating from a sketched site map to the computer, GPS errors (if a GPS is used), or poorly-drawn site maps. An error of as little as five meters can make the difference between a site being captured by the model or not. Based on this, sites were buffered to 10 m (which also helps to account for sites which are mapped as very small polygons to represent a point, either because the site is very small or the actual extent of the site is unknown). The buffered sites were then intersected with the model to determine how many sites were captured.

Main portion of CRD model results

Of a total of 572 sites in the main CRD model area which are recorded as being at least partially intact, 411 (72%) are captured by the model (or are within 10 m of a modelled area). A total of 13% of the land area modelled is captured as having potential. This results in an overall Kv gain of 0.82.

East Sooke and Highlands model results

The Highlands and East Sooke models, though based on a less detailed DEM, are working quite well – though this may be in part due to the low numbers of sites with which to check the model performance. Both the Highlands and East Sooke models captured only 9% of the land area, while capturing a majority of the sites in those areas (85% of the sites in East Sooke, and 5 out of the 6 sites in Highlands). This works out to a Kv gain of 0.89 for both East Sooke and Highlands, though the few sites in these areas makes using the Kv gain less appropriate. Ground-truthing would be required to more accurately determine the performance of these models.

Use of the model

The final model is a representation of archaeological potential – areas included in the model are considered to have high potential, whereas all other areas are considered low potential. This means there is strong potential for finding an archaeological site in areas with high potential, and low, but not impossible, in areas with low potential. Not every area modelled as high potential will include an archaeological site, and archaeological sites may be found in areas modelled as low potential.

Throughout history, aboriginal groups made use of many resources all over the landscape. This model is not intended to reflect traditional use of the landscape; rather, it is intended as a means to determine areas where the potential for finding tangible evidence of this land use is high.

Model Limitations

There are some limitations to the model that it is important to take into consideration. These may be influencing the results of the modelling, in that 28% of the sites were not captured. In some areas, such as Esquimalt and Victoria Harbours, the coastline has been significantly modified by building of dockyards, wharves, piers, and breakwaters. The mapped modern coastline includes these features, and thus not only does the model capture them as high potential (based on the coastal buffer), but it also misses the true area of potential – the original coastline. A similar effect, related to stream modification, occurs in many areas throughout the CRD. Ditching, dredging, and redirecting of streams means that the current mapped course of the stream may not necessarily reflect the original location. This was at least partially accounted for by the inclusion of the "Lost Streams of Victoria Mapping" (Sutherst 2003) as discussed above, though it is likely that this data source is not comprehensive, and that there are other areas of modified terrain which are not accounted for in this mapping. In addition, only the Greater Victoria area was encompassed in this mapping, and other areas of the CRD likely have similar (though probably to a lesser extent) stream and shoreline modifications. An extensive search of all available historic map records would be required to continue this mapping. To further address the issue of stream modification, where the stream was classified in the mapping as "ditch" and it seemed to connect the natural "stream" portions of a watercourse, it was assumed that it was likely close to the original course, and was therefore included in the model. Of course, this means that in some cases there may be inaccurate potential along these sections, while the true potential along the original course of the stream would not be captured.

Model errors may result also from other types of landscape modifications – blasting forms cliffs that may not be natural, ponds are dug for farmland irrigation, wetlands are drained to plant fields, coastal waterbodies are filled (for example, James Bay), etc. While attempts have been made (through manual digitizing) to remove some of the erroneous potential, it would be difficult and very time-consuming to capture all of it, and also very difficult if not impossible to add potential where features have been modified.

Sites used to analyze the model did not include historic sites or sites which were noted as "destroyed". However, it was later noted that some sites had in fact been destroyed, but did not use that term, so were included in the analysis. A further review of the model and sites may improve the overall performance, as the destroyed sites cannot be reliably used to analyze the model performance due to landscape modification. In some cases, sites were noted to be mapped in the incorrect location. Though buffering the sites when analyzing model performance should reduce the effects of minor (10 m or less) site location errors, some sites may be in error by much larger distances. This should be taken into consideration with respect to the Kv gain, as it is likely that some site locations were in fact captured by the model though the available mapping suggests that they weren't (though of course, the reverse could also be true). For this reason, follow-up field work and ground-truthing of the model would be very useful to better test the model's performance and to provide recommendations for refinement and improvement of the model.

There was difficulty in modelling for certain sites – these were primarily sites which were located inland, in areas where the terrain was relatively flat and featureless or where small landforms were not evident in the DEM. They appeared to be randomly located – that is, they did not follow any defined landforms or mapped features such as rivers. An attempt was made to associate the sites with the location of Garry Oak meadows; when visually inspecting the map, there appeared to be a correlation with the edges of Garry Oak meadows. The sites appeared to be primarily located at the edges or just outside of (on the inland side) the meadows. It was thought that this might be an indication of people going further inland, through the Garry Oak, to the stands of cedar, in order to harvest the cedar. However, the visual correlation was not confirmed by quantitative analysis – excluding the sites within 1km of the coast, there was a very similar density of sites within Garry Oak meadows and outside the meadows. This changed very little even when the meadows were buffered into by 150m, and this distance of Garry Oak meadow was considered "outside" (based on the idea that people may have camped near the cedar harvesting location, but within the Garry Oak meadow).

A final factor to consider when looking at the model and its results is the DEM. While the main portion of the CRD had a fairly detailed DEM, it is still somewhat generalized. For instance, a site that is known to be at the base of a steep slope is missed by the model, as the slope is shown much shallower in the DEM due to the spacing of the elevation data points from which the DEM is interpolated. Therefore, it is impossible to capture that location in the model without using a fairly low slope in the "top/base slope" algorithm, which then captures too much area. A more detailed DEM (such as LiDAR) would undoubtedly increase the accuracy of the model.

Recommendations

This model is performing quite well considering the limitations described above, however, as with all such projects, it would benefit from ground-truthing.

Ways to improve the model include:

- Incorporating input from the First Nations. Meetings with knowledgeable First Nations members could be arranged, and their input could be used to revise the model. For example, local community members are known to have information about former wetland locations and unrecorded sites.

- Conducting research of historic mapping records to identify areas of landscape modification, in particular original locations of streams and wetlands. Alternately should the CRD undertake such research in the future, it could easily be incorporated into the model.

- Incorporation of LiDAR data should it become available.

References

Eldridge, M. and A. Parker
2006 Archaeological potential model of Hul'quminum Group Core territory.

2007 *City of Langford Archaeological Overview Assessment – Phase 1 Baseline Mapping*. Millennia Research Limited.

Eldridge, M., A. Parker, J. Brunsden and I. McKechnie
2007 *Campbell River Forest District Archaeological Overview Assessment*. Millennia Research Limited.

Kvamme, K. L.
1988 Development and testing of quantitative models. . In *Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling*, edited by W. J. Judge and L. Sebastian, pp. 325-428. US Department of the Interior, Bureau of Land Management Service Center, Denver, Co.

Sutherst, J.
2003 Lost Streams of Victoria. (map with text)