Southern Finland Office
151/2013
Espoo

19.9.2013
updated 15.11.2022

Lithium Pegmatite Prospectivity Modelling in Somero-Tammela Area, Southern Finland

Hanna Leväniemi

GEOLOGIAN TUTKIMUSKESKUS • GEOLOGISKA FORSKNINGSCENTRALEN • GEOLOGICAL SURVEY OF FINLAND

PL / PB / P.O. Box 96
FI-02151 Espoo, Finland
Tel. +358 20 550 11
Fax +358 20 550 12

PL / PB / P.O. Box 1237
FI-70211 Kuopio, Finland
Tel. +358 20 550 11
Fax +358 20 550 13

PL / PB / P.O. Box 97
FI-67101 Kokkola, Finland
Tel. +358 20 550 11
Fax +358 20 550 5209

PL / PB / P.O. Box 77
FI-96101 Rovaniemi, Finland
Tel. +358 20 550 11
Fax +358 20 550 14

Y-tunnus / FO-nummer / Business ID: 0244680-7 • www.gtk.fi

GEOLOGICAL SURVEY OF FINLAND DOCUMENTATION PAGE

Date / Rec. no.

Authors Hanna Leväniemi	Type of report Archive report
	Commissioned by

Title of report
Lithium Pegmatite Prospectivity Modelling in Somero-Tammela Area, Southern Finland

Abstract
The Somero-Tammela area is one of the best known lithium pegmatite regions in Finland. This study investigates the possibility to model the prospectivity for the pegmatites using regional geophysical and topographic datasets.

As the pegmatites are usually non-responsive for geophysical methods and thus the nature of relationship between a dataset and the deposits is difficult to estimate, an empirical modelling method (weights of evidence) is applied that allows the modelling data to be selected based on statistical relationship parameters. This report describes the modelling datasets and working methods.

An update to the report in Nov 15, 2022, lists the training data point coordinates in Appendix 1.

Keywords
Lithium pegmatite, GIS, prospectivity modelling, weights of evidence

Geographical area
Somero, Tammela

Map sheet
2024, 2113

Other information

Report serial GTK archive report		Archive code 151/2013	
Total pages 15 p.	Language English	Price	Confidentiality Public
Unit and section ESY 211		Project code 2551005	
Signature/name Hanna Leväniemi		Signature/name	

Contents

Documentation page

1 INTRODUCTION

The purpose of this study is to model the prospectivity for lithium pegmatites on the pegmatite province of Somero-Tammela in southern Finland. The datasets used in modelling comprise of aerogeophysical datasets and topographic data. The report focuses on practical issues concerning data processing and modelling.

Lithium is one of the rare-element (RE) minerals that occur within pegmatite veins. There are several known pegmatite-related RE mineral deposit regions in mid- and southern Finland as shown in Figure 1. In each area, there may be several pegmatite groups with deviating mineral compositions and possibly also of different age (Alviola, 2003). RE pegmatites are commonly divided into two groups, namely LCT and NYF pegmatites; the first are associated with S-type, Al and quarz-rich granites and enriched in Li, Cs and Ta whereas the latter group, associated with A-type granites, is enriched in Nb, Y, F, U, Th, REE (Selway, 2006).

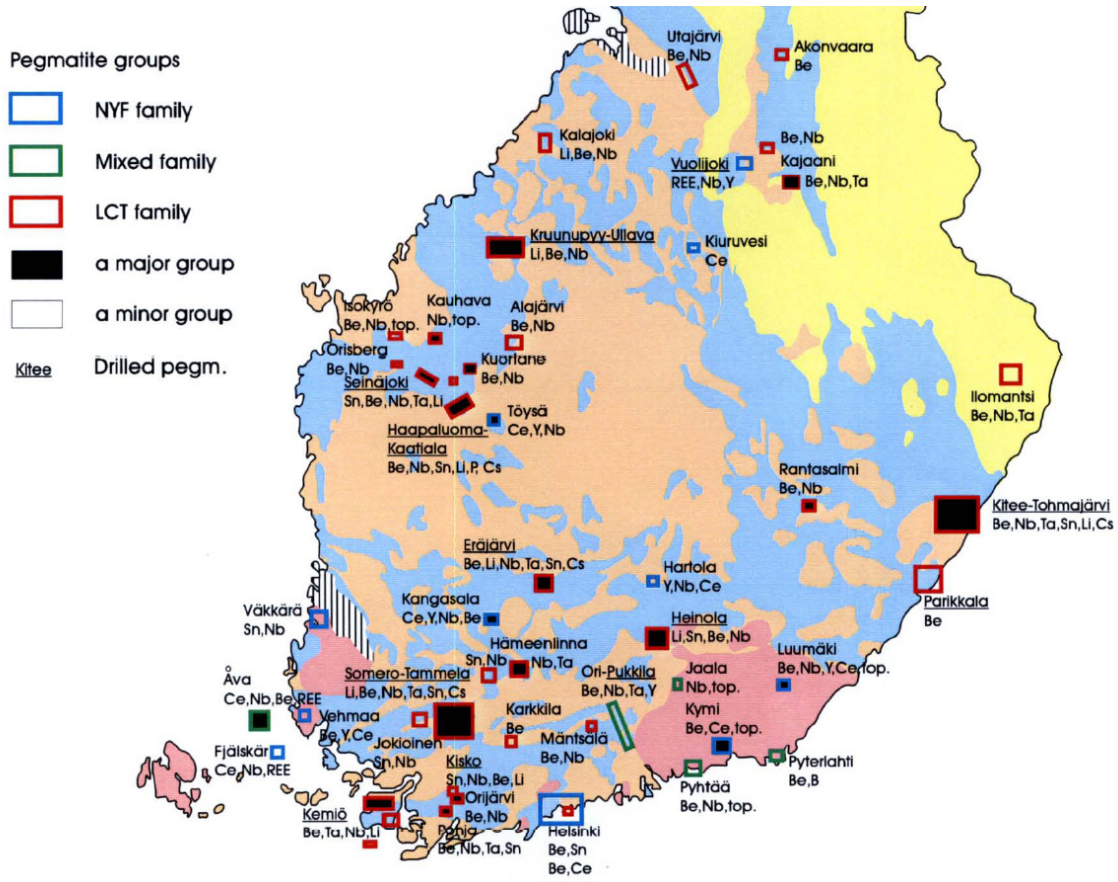

Figure 1. RE pegmatite groups in Southern Finland (from Alviola, 2003).

This study concentrates on the lithium-rich pegmatite veins on the Somero-Tammela area and thus the term ‘lithium pegmatite’ is used throughout the report instead of the more general ‘RE pegmatite’ when discussing the regions deposits.

2 MODELLING AREA

The study area is locate in the Häme volcanic belt that comprises volcanic rocks intercalated with minor greywackes and metapelitic units and is intruded by syntectonic plutonic rocks and late-tectonic K-granites and pegmatite dykes (Eilu et al., 2012). The area selected for modelling covers approximately 893 km^2 and includes all the known pegmatite veins (Figure 2).

The Somero-Tammela area is one of the most economically potential lithium pegmatite provinces in Finland. The ore-critical pegmatite dykes occur in groups/swarms; in this area at least seven groups of different mineral composition and possibly also of different age are known. There are few minor occurrences in the Somerniemi area (south-east corner of Figure 2) but the main deposits are located in the northern part of Somero municipality and in the southern part of Tammela municipality. The former of these are horizontal and thus appear visually larger in size than the latter, vertical deposits. The drilled deposits of Kietyönmäki and Hirvikallio are estimated to be of the same size (approximately 400 000 tons of ore at depths shallower than 100 m) with low Fe grades (i.e. good-quality Li ore). Although the mineralizations in the area are in general small in size, there may yet be deposits to be discovered in the area, the potential estimated to be highest around the known deposit of Kietyönmäki. (Alviola, 2003)

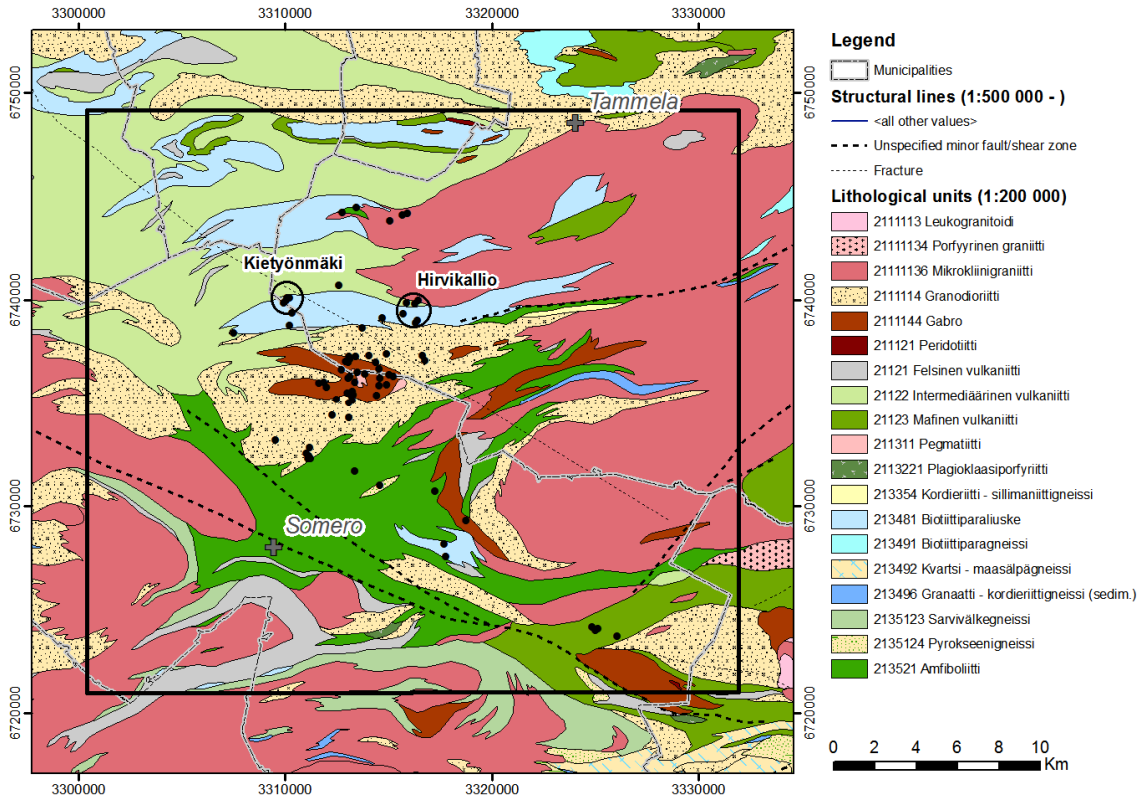

Figure 2. Somero-Tammela area, region selected for modelling shown with borders. Center points of known mineralized pegmatite dykes shown in black symbols (listed in Appendix 1).

3 MODELLING CONSIDERATIONS, OBJECTIVES AND METHODS

3.1 Applicable Characteristics of RE Pegmatites

There's little literature that can be applied on the the modelling problem at issue as the exploration nature of the pegmatites remains elusive - as Galeschuk and Vanstone (2007) formulate, the "nature of the granitic pegmatites limits the number of exploration tools that can be utilized" and to date, no definitive toolset has been identified. Galeschuk and Vanstone (2007) list the general geological characteristics of rare element pegmatites to include

- Dyke-like geometries
- Horizontally and vertically directed propagation
- Association with deep-seated structures and fractures
- Granitic mineralogy
- The pegmatites fractionate away from the intrusive source. The most economic concentrations for rarest minerals can be found in the most distal pegmatites.
- The host rock is geochemically and mineralogically altered (addition of apatite, tourmaline, carbonate, micas)

Most often geochemistry and mineral studies are the broadest-discussed options for pegmatite exploration in brownfield areas (Galeschuk and Vanstone, 2007; Černý and Trueman, 1982, Selway et al., 2006; Vieira, 2007). Regarding their geophysical properties the RE pegmatites might best be described as "non-responders" (Černý and Trueman, 1982). They are non-magnetic and resistive; the possible indicators might be the density contrast between the pegmatite and the host rocks or the pegmatites' gamma-radiation properties. Magnetic and EM methods are mostly used as secondary data to delineate host rock features that may act as structurally or stratigrafically controlling factors for the pegmatites.

As for the Finnish deposits, on the recently-studied Kaustinen pegmatite area detailed airborne geophysics and ground surveys fail to detect the pegmatites (Ahtola et al., 2010a, Ahtola et al., 2010b). For gravity density values of 2.7-2.8 g/cm^3 are reported. This would make the density contrast to the surrounding rock (mica schist in Kaustinen) negligible and thus the gravity method would be unsuitable for pegmatite detection.

3.2 Modelling Method: Weights of Evidence

As the deposit characteristics are known to be ambiguous but there are still plenty of known locations for mineralizations, an empirical modelling method called Weights of Evidence (Bonham-Carter, 1994) was selected. The method is based on Bayesian statistics where certain conditions and additional knowledge can be used to adjust and improve the probability calculations in comparison to random event occurrence probability.

In terms of mineral exploration, for a certain size of area divided to small sub areas (unit cells) with known deposits (training points), the "random", i.e. *priori* probability for a mineral deposit occurring in a certain unit cell would be the number of all deposits divided by the total number of unit cells in the area. Now if we have some extra information (evidence), e.g. an EM dataset and we make an assumption that the deposits are related to high EM values, we can include this assumption in our probability model so that for cells with high EM values the probability ex-

ceeds the *priori* probability and for low EM value cells the probability is lower. These probability values are then called *posterior* probabilities. (Bonham-Carter, 1994)

For modelling, the data is first reclassified to an integer raster with a limited number of classes. The classes are tested in order (ascending or descending) or categorically (with no order) to see what type of correlation the data patterns show to training sites. The modelling procedure calculates a pair of weights (W^+ and W^-) for each class of the classified evidence layer. The two weights refer respectively to positive and negative evidence; if more points occur within a pattern that expected by chance then W^+ is positive and W^- negative. However, were there less training points occurring than would be due chance, the W^- value will be positive and W^+ negative. The arithmetic difference between the weights values is called the contrast; high contrasts indicate strong association between the evidence layer and the training point dataset. Based on the contrast and its derivative, the studentized contrast (also called confidence), the evidence layer is generalized into predictive binary pattern layer, the binary values corresponding to areas "favorable" and "non-favorable" to training point occurrence. When several of these predictor patterns are combined, we get a posterior probability map showing the areas favourable for training-point-like environments. (Nykänen & Salmirinne, 2007)

Combining the weights for the final model introduces the assumption of conditional independence (CI), that is, the evidence layers are assumed to be fully independent. This is not always the case with geodata, for example, the lithological map may be constructed based partly on airborne magnetics which would mean that the two datasets are mutually dependant. However, as Bonham-Carter (1994) notes, CI condition is probably always violated to some degree and CI tests should be applied while modelling in order to investigate the dependence issues and to reject or modify the input datasets.

Another important part of modelling is testing the model for validity, i.e. how well the model performs in case of deposits not used in model training. For this purpose, the training point dataset is usually divided in two, the first set is used for model training and the second for assessing the model performance in the validity tests.

The modelling was implemented using the ArcSDM toolset available for download at http://www.ige.unicamp.br/sdm/default_e.htm.

3.3 Modelling Considerations

In addition to the non-responsive nature of the pegmatites in respect to the available datasets, there's another, a dimensional consideration: at the study area the thickness of the pegmatite dykes is known to be approximately 5 – 20 meters which is small in comparison to spatial resolution of various datasets (Table 1). Thus it's debatable whether the distinct dykes would directly show in the regional datasets even if the respective property contrasts were high enough.

Dataset	**Raster data resolution**
Aerogeophysical data (magnetic, electromagnetic, radiometric)	50 m
Digital elevation data (National Land Survey)	25 m
Regional gravity dataset	2000 m
Till geochemistry	> 1000 m
Rock geochemistry	> 2000 m

Table 1. Spatial resolution of input datasets.

The non-responsive nature of the pegmatites and the dimensional issues certainly set a challenge for the modelling task at hand. However, there may be broader scale alterations and characteristics in the veins' environment that have enabled the generation of ore critical dykes; bearing that in mind it could be said that the modelling is also about finding the favourable environment if not the distinct dykes.

One should bear in mind that as the training points are rather clustered at the center of the study area, the type of environment present around this cluster will be likely to rule the model.

4 DATASETS AND MODELLING

The modelling is based on raster data processing. It can be readily seen when comparing the raster resolutions (Table 1) that only the aerogeophysical datasets (GTK's low-altitude airborne data) and the digital elevation data (by the National Land Survey) a have high enough a resolution for modelling.

Raster dataset resolution defines the modelling resolution as it's effectively no use applying a smaller unit cell size than the most common raster dataset resolution. Thus a unit cell size of 50 m was applied in the model.

4.1 Training sites

There are altogether 72 pegmatite vein observations in the study area (Figure 2). The data originates from a polygon/polyline dataset depicting the pegmatite veins (T. Ahtola, pers. comm.). As the modelling expects the training points to be of type point, this training point data set represents the mid-points of the vein features, listed in Appendix 1 (added Nov 11, 2022).

As modelling algorithm necessitates there be at maximum only one training point per each unit cell, this condition was first checked to ensure the validity of the training point set.

To validate the model after running the algorithms, the training point data set was divided by random selection into two: 70% of the training sites (50 points) were used for calculating the weights and model response and the remaining 30% were left for model validation. The *priori* probability with this training site set is 0.000140.

4.2 Evidence Dataset Descriptions

4.2.1 Magnetic data

Several derivatives of magnetic datasets were tested based on the notion that most LCT pegmatite bodies show some sort of structural control (Bradley and Cawley, 2013) and that such bedrock structures are often visible in the magnetic data. In the final model the possible structures were traced by calculating the local minima points in continues features in the TDR derivative (Verduzco et al., 2004) of the airborne magnetic dataset. The TDR derivative enhances the small features in the data and the local minima show the magnetic minima often related to faults and alteration zones.

The calculation algorithm for the local minima, based on surface curvature analysis (Phillips et al., 2007), produces a point set comprising of minima points of surface features that show directional elongation. These features may relate to faulting or alteration processes that have resulted in loss of magnetic material. Before the analysis the aeromagnetic data was upward continued by 70 m in order to eliminate any artificial or geological noise.

As the point data set cannot be used as evidence layer for the modelling algorithm, a new layer was formed by calculating the density of minima points. This would interpret as density of fractured/altered features, the implication being that with high density of points the rock shows more structural features. The density raster was reclassified to 20 classes using the Natural Neighbor classification algorithm of ArcGIS. The data was tested in descending order, that is, assuming that high values correlate to training sites.

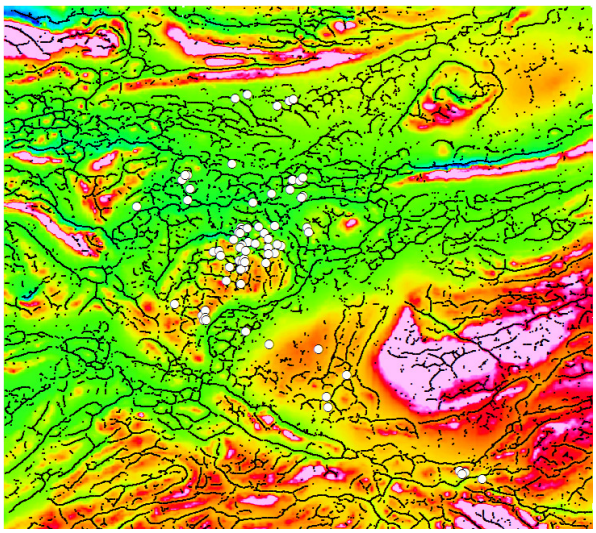

Figure 3. Aeromagnetic data with local minima points (in black). Pegmatite veins marked with white symbols.

4.2.2 Electromagnetic data

The pegmatites are expected to locate in resistive environment. The in-phase to quadrature component ratio (Re/Im) of GTK's frequency-domain airborne 3 kHz EM data was used to locate resistive areas: high ratio values correspond (as a rule of thumb) to good conductors and low values to poor conductors, i.e. resistive areas (e.g. Parasnis, 1979, pp. 143-144). The data was first masked to low cut-off value (±10 ppm for in-phase and 10 ppm for quadrature) to exclude near-zero values possibly resulting in erroneously interpretable ratio values. In order to get full coverage of data, the gaps resulting from the masking operation were filled with interpolation. The data is still prone to any "errors" caused by overburden conductors masking the bedrock resistivity.

The data was manually classified into 10 classes and an ascending-order relationship was assumed (i.e. small values were assumed to be related to training points).

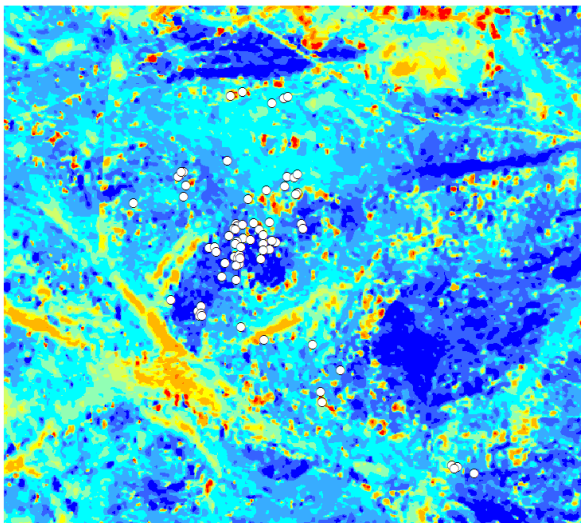

Figure 4. AEM Re/Im ratio (blue = low values, red = high values). Pegmatite veins marked with white symbols.

4.2.3 Radiometric data

Various ratio calculations of radiometric components K, eU and eTh were tested for modelling. As can be seen in Figure 5 the selected evidence layers (K/eU and U/eTh ratios) visually resemble each other and mutual dependency is likely to be an issue were the datasets used in the model simultaneously. Because of high visual resemblance the datasets were not combined to a single evidence layer as was done with the digital terrain model datasets (see next chapter); rather the datasets performance and independence was tested when combining the evidence layers to prospectivity models.

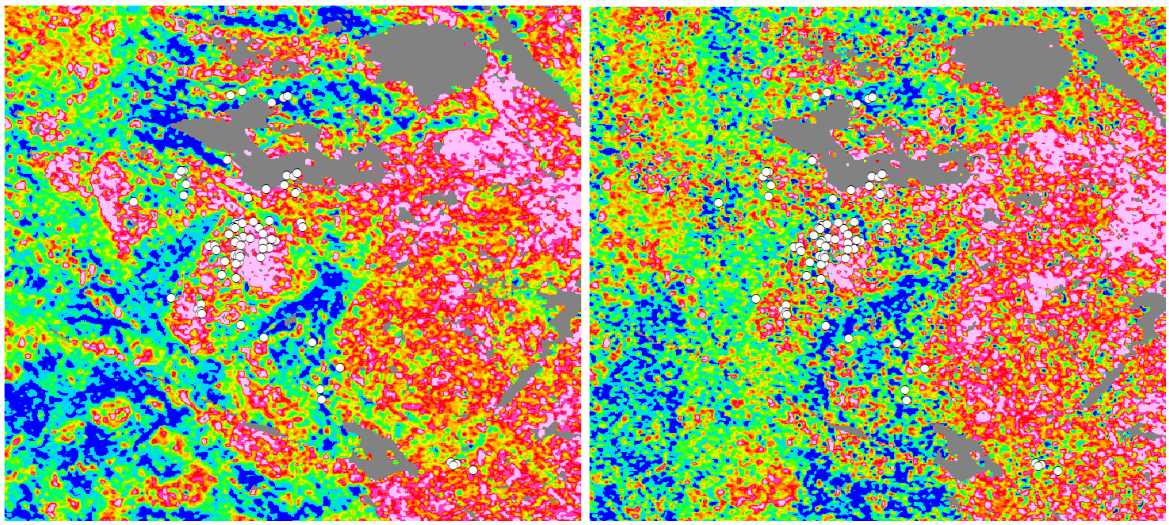

Figure 5. K/eTh ratio on the left; eU/eTh on the right. Blue = low values, magenta = high values. Dark gray areas represent the masked data. Pegmatite veins marked with white symbols.

The datasets were masked with a cut-off values (K: 0.5%, eU: 0.4ppm, eTh: 1.5 ppm) before processing, the values selected by visual inspection to match lakes and swamps that present low

value regions that may distort the modelling as their origin is certainly not from the bedrock. A moving window (3x3 cells) low-pass filter was also applied to reduce noise issues.

4.2.4 Topography

Granitic pegmatites are relatively hard-wearing and tend to stand above their surroundings (Bradley and Cawley, 2013). This seems by visual inspection like a valid assumption also in the Somero-Tammela region; the pegmatite veins seem to be located in higher ground and especially on the slopes as analyzed from the digital elevation model (DEM) with original resolution of 25 m (the data was transformed to 50 m cell size to match the resolution of other datasets).

The DTM data was preprocessed by subtracting a moving window (1250 m radius) mean from the original data to highlight local variation in the topography. The slopes layer was calculated using Spatial Analyst tools of ArcGIS. In order to avoid mutual dependence issues, the datasets were combined to a single evidence layer using logical operations (the Fuzzy Logic tools of ArcSDM); the datasets were combined with the fuzzy OR operator, resulting in a combination layer where areas of high elevation or high slope angle are given high values and low values represent flat areas of low elevation.

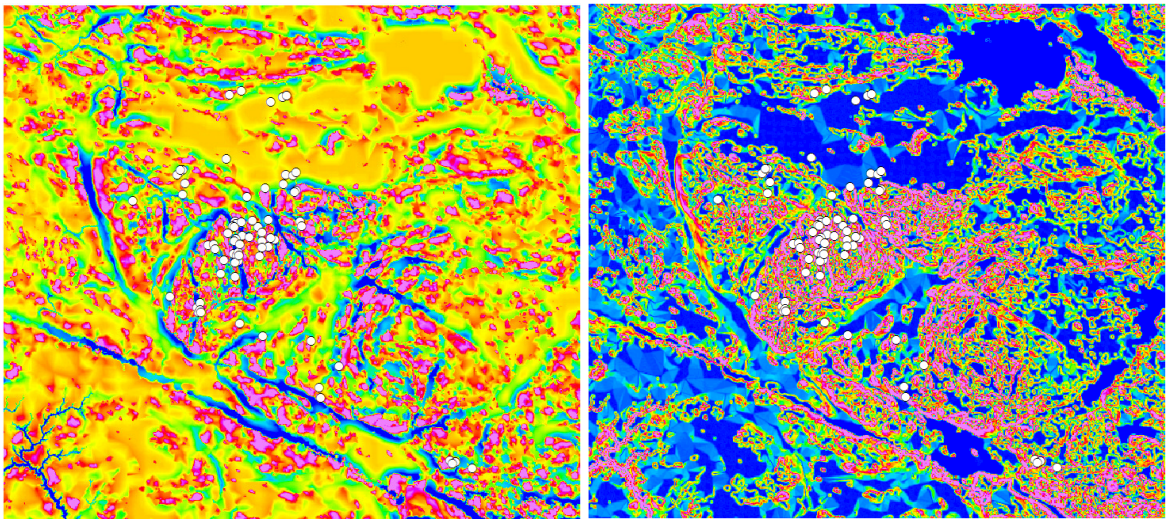

Figure 6. DTM mean residual on the left; DTM slope on the right. Blue = low values, magenta = high values. Pegmatite veins marked with white symbols.

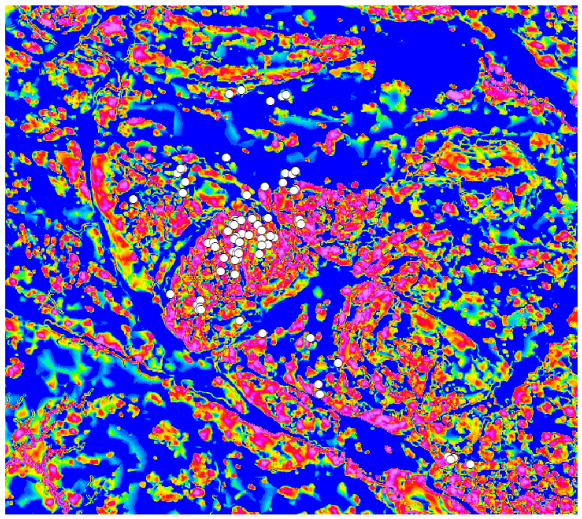

Figure 7. Processed DTM evidence layer on the left (blue = low values, magenta = high values). Pegmatite veins marked with white symbols.

4.2.5 Datasets not included in the model

Datasets that were considered and tested for modelling comprise regional gravimetric data, lithological unit polygons and geochemical data. These were not applied to model as the resolution of these datasets is too coarse.

5 MODELLING RESULTS AND VALIDATION

5.1 Weighting Parameters

Table 2 shows the weighting parameters relating the training point dataset to each input data layer. The default cut-off value of 2% confidence (studentized contrast) was used for breakpoint for binary class generalization, i.e. defining the regions favourable/non-favourable for training point occurrence.

The contrast values rank the evidence layer in order of modelling efficiency. Based on the weighting parameters the digital elevation model has highest contrast value indicating that it's the best predictor for the training points. There's no official criterion available for "good" or "bad" contrast but Nykänen & Salmirinne (2007) estimate a contrast value of 1.5 to be still quite high. Based on that criterion, all the contrast values here seem at least reasonable.

For AEM conductivity derivative (Re/Im) and radiometric eU/eTh ratio the positive weight value W^+ exceeds the negative value W^-. This indicates that these two layers predict the presence of training points better than their absence. For evidence layers with higher W^- than W^+ the case is opposite; both types of evidence layers are valuable for modelling.

The evidence layers show variability also in the "favourable" area size and number of training points falling within that area. The extreme ends are the two radiometric evidence layers: for the K/eTh layer the favourable pattern covers 72% of the total area and number of points is close to the the total number of training sites whereas the eU/eTh favourable pattern covers only 7% of the total area but still manages to catch a considerable number of training points.

Dataset	Area (km^2)	# Points	W^+	W^-	Contrast	Confidence
Aeromagnetic, density of local minima points	415	46/50	0.6814	-1.8994	2.5808	4.9508
AEM Re/Im ratio	211	26/50	0.7907	-0.465	1.2556	4.4354
Airborne radiometrics, K/eTh	646	47/50	0.1831	-1.5958	1.7789	2.4638
Airborne radiometrics, eU/eTh	65	12/50	1.1262	-0.1982	1.3243	3.9857
Digital elevation model, processed	461	49/50	0.6403	-3.1852	3.8255	3.787

Table 2. Weighting parameters for each input data layer.

5.2 Model Validations

In the model response calculation the evidence layers and their weighting parameters are combined into as single predictability layer. Three slightly different models were tested (mainly because of mutual dependency issues, see next chapter):

1. With 70% of points used for training and 30% for validation, both K/eTh and eU/eTh included.
2. With 70% of points used for training and 30% for validation, eU/eTh included and eK/eTh excluded.
3. With 70% of points used for training and 30% for validation, K/eTh included and eU/eTh excluded.

After the model response calculation the models may be validated with ArcSDM tools by studying the area frequency tables that estimate the efficiency of model classification and prediction. When the tables are produced with the validation point set (the part of the original deposit dataset that hasn't been used in the model training), we get an estimate of how well the model really predicts any deposits unknown to model. The PRC curves (Figure 8) provide a visual way for estimating the efficiency of prediction – were the model as good as a random guess, the curve would be a rising linear line $f(x)=y$ and the parameter Area Under Curve (AUC) would be 50%. When the PRC curve rises above the linear trend and the AUC parameter is above 50%, the model predicts the training points better than due chance. The higher the rise in the curve and the closer the AUC parameter is to 100%, the better the predictability. All three models behave similarly – this was to be expected as there's only little variation in the models. The AUC values are around 85% which indicates that the models do predict the validation points.

Another test for PRC curves would be to test the model with a dataset consisting of point definitely not deposits ("not-sites"). The curve should preferably be below 50% for these points, indicating that in addition to managing to predict the true sites the model rejects the not-site locations. For our case there's no such no-sites dataset available but the test can be simulated with random points. (Theoretically there could be an unknown deposit located on a random point; however, as this is quite improbable this type of test can be used for practical reasons even though it's strictly speaking perhaps not scientifically correct). The Model #2 was tested for 22

random points (to equal number of validation points) and the results show that the model does indeed not predict the "not-sites" better than would be due chance.

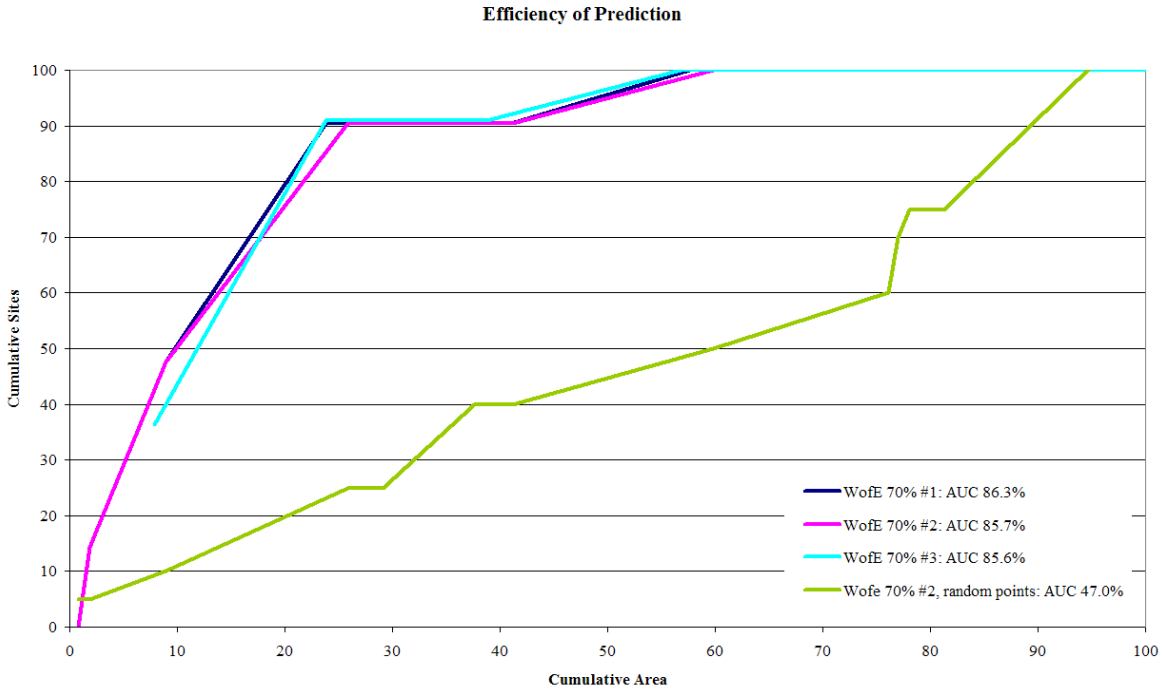

Figure 8. PRC curves for the three models.

5.3 Conditional Independences

Two tests were performed for models to check for possible conditional independence issues. As the model contents only vary in regard to radiometric evidence layers, this is really to check whether including all the radiometric evidence layers introduces independence issues to models. Pairwise tests between evidence layers (Bonham-Carter, 1994) were not performed although they would provide a more detailed insight into conditional independence within the models.

The two tests taken, the Overall CI Test (Bonham-Carter, 1994) and the Agterberg-Cheng CI test (Agterberg and Cheng, 2002), both dealing with the overall conditional independence of the model. The first of these, the Overall CI Test, compares the number of true training sites to number of points that are suggested to be on the area by the *posterior* probability result layer. In practice the predicted number of points is always larger than the true number of training sites, however, if the number of predicted points exceeds the number of training sites by 10-15% there's reason to consider the model to be conditionally dependant.

The Agterberg-Cheng test also compares the number of predicted points to the number of true sites by statistically testing the hypothesis that the two numbers are equal: values greater than 95-99% indicate that the hypothesis of conditional independence is not valid.

The test results (Table 3) clearly indicate that the Model #1 with both radiometric evidence layers included will not pass the conditional independency condition, as does not the Model #3 with K/eTh evidence layer included. The Model #2 with eU/eTh layer however shows acceptable levels of CI. Thus as the models validations show very similar results, this model is selected to be the final model based on the best CI test result.

Model	Overall CI test	Agterberg-Cheng CI test
Model #1	28%	85.6
Model #2	11%	69.9
Model #3	25%	80.1

Table 3. Conditional independence test results.

5.4 Prospectivity Map

The prospectivity map based on Model #2 is shown in Figure 9 together with the true mapped pegmatite vein polygons. The grey areas represent regions where the prospectivity values are lower than the *priori* prospectivity, that is, where the model prediction is as good as a random guess. The prospectivity values start to rise starting from the blue regions; magenta colors show the regions with highest prediction values. (The numeric prospectivity values are subjective to model and thus do not correspond to conventional probability range of 0-100%).

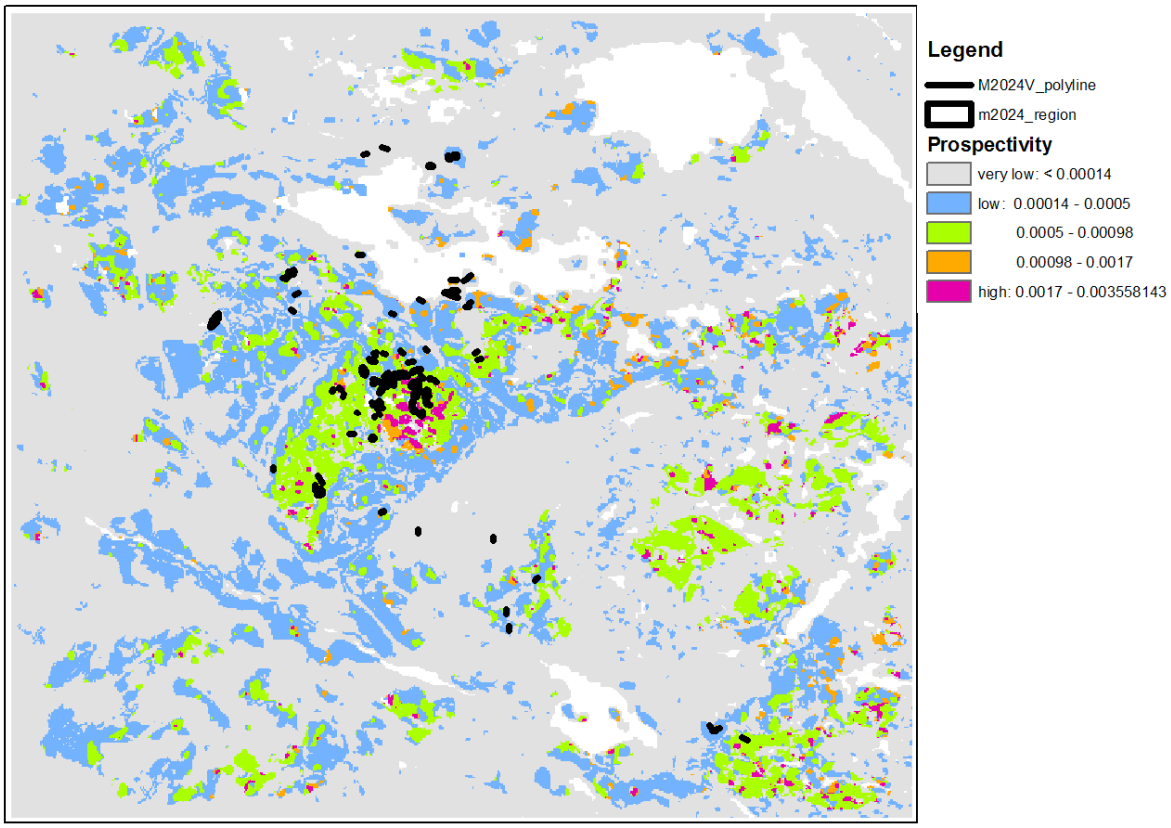

Figure 9. Final prospectivity model result.

Despite the challenges in modelling the model performs rather well in the central regions of the area where most of the veins are clustered. In a detail view (Figure 10) many of the known veins are located on regions with high predictability; especially promising is the fact that the major vein swarm of Kietyönmäki is located in the highest predictability class.

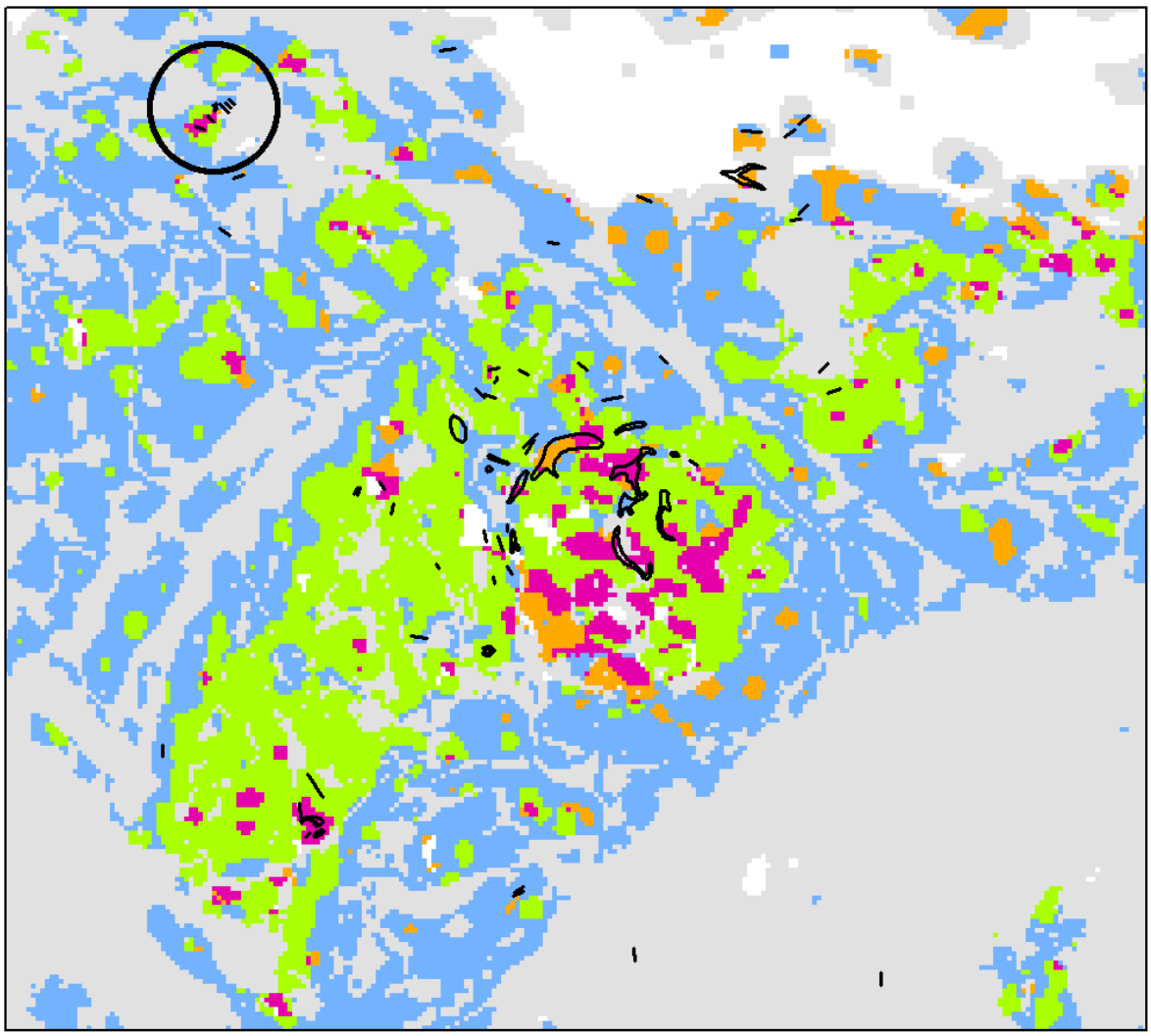

Figure 10. Model details at the center of the study area. Kietyönmäki pegmatite veins inside the black circle.

6 CONCLUSIONS

The work models lithium pegmatite veins using regional geophysics and topographic datasets. Although the veins are apparently too small to show in regional datasets and are also non-responsive for geophysical methods, the model relies on the vein environment to show as anomalous and thus it would be more accurate to say that the study models the suitable environment for the pegmatites rather than the actual veins locations.

The major part of the deposits is located in a cluster in the central regions of the study area. This is where the model also shows highest predictability values. If the veins' location is due to the fact that the central parts of the area are just mapped in more detail but there could be more veins outside the central area, then the model should be considered slightly biased and more valid in the central regions.

There are relatively few evidence layers available for modelling. This is partly due to the dimensional issues mentioned earlier (for many possible evidence layers the resolution is too coarse in

comparison to veins' dimensions) but also because the lithium pegmatites are non-responsive for many datasets. However, it was possible to extract evidence layers from all the three airborne geophysical data (magnetic, electromagnetic and radiometric), showing that where conceptual models based on expert opinion are difficult to construct, the empirical methods like the Weights of Evidence employed in this study can still work and provide useful insight into the relationships between the evidence layers and deposits.

The resulting prospectivity model performs reasonably well based on the validity tests and also on visual inspection. Were there more datasets available in the future (e.g. a more detailed geochemical mapping) the model can be updated and improved with the new datasets.

References

Agterberg, F.P. and Cheng, Q., 2002. Conditional independence test for weights-of-evidence modelling. Natural Resources Research 11, 249-255.

Ahtola, T., Kuusela, J., Koistinen, E., Seppänen, H., Hatakka, T. and Lohva, J., 2010a. Report of investigations on Leviäkangas lithium pegmatite deposit in Kaustinen, Western Finland. Geological Survey of Finland, report M19/2323/2010/32.

Ahtola, T., Kuusela, J., Koistinen, E., Seppänen, H., Hatakka, T. and Lohva, J., 2010b. Report of investigations on Syväjärvi lithium pegmatite deposit in Kaustinen, Western Finland. Geological Survey of Finland, report M19/2323/2010/44.

Alviola, R., 2003. Pegmatiittien malmipotentiaalista Suomessa. Geological Survey of Finland, report M10/03/85. (in Finnish)

Bonham-Carter, G.F., 1994. Geographic Information Systems for Geoscientists - modelling with GIS. Pergamin, New York, 398 p.

Bradley, D. and Cowley, A., 2013. A Preliminary Deposit Model for Lithium-Cesium-Tantalum (LCT) Pegmatites. U.S. Geological Survey, Open-File Report 2013-1008.

Černý, P. and Trueman, D.L., 1982. Exploration For Rare-element Granitic Pegmatites. In: Černý, P. (ed): Short Course In Granitic Pegmatites In Science And Industry. Mineralogical Association of Canada, Short Course Handbook, 8, 463-493.

Eilu, P., Ahtola, T., Äikäs, O., Halkoaho, T., Heikura, P., Hulkki, H., Iljina, M., Juopperi, H., Karinen, T., Kärkkäinen, N., Konnunaho, J., Kontinen, A., Kontoniemi, O., Korkiakoski, E., Korsakova, M., Kuivasaari, T., Kyläkoski, M., Makkonen, H., Niiranen, T., Nikander, J., Nykänen, V., Perdahl, J.-A., Pohjolainen, E., Räsänen, J., Sorjonen-Ward, P., Tiainen, M., Tontti, M., Torppa, A. and Västi, K., 2012. Metallogenic Areas in Finland. In: Eilu, P. (ed): Mineral Deposits and Metallogeny of Fennoscandia. Geological Survey of Finland, Special Paper 53, 207-324.

Galenschuk, C. and Vanstone, P., 2007. Exploration Techniques for Rare-Element Pegmatite in the Bird River Greenstone Belt, Southeastern Manitoba.

Nykänen, V. and Salmirinne, H., 2007. Prospectivity Analysis of Gold Using Regional Petrophysical And Geochemical Data From the Central Lapland Greenstone Belt, Finland. In: Ojala, V.J. (ed): Gold in the Central Lapland Greenstone Belt. Geological Survey of Finland, Special Paper 44, 251-269.

Parasnis, D.S., 1979. Principles of Applied Geophysics. Chapman and Hall, London, Third Edition.

Phillips, J.D., Hansen, R.O. and Blakely, R.J., 2007. The use of curvature in potential-field interpretation. Exploration Geophysics, 38, 111-119.

Selway, J.B., Breaks, F.W. and Tindle, A., 2006. A Review of Rare-Element (Li-Cs-Ta) Pegmatite Exploration Techniques for the Superior Province, Canada, and Large Worldwide Tantalum Deposits. Exploration and Mining Geology, 14, 1-30.

Verduzco, B., Fairhead, J.D., Green, C.M. and MacKenzie, C., 2004. New insights into magnetic derivatives for structural mapping. The Leading Edge, February 2004, 116-119.

Vieira, R., 2007. Lithium Pegmatites Exploration in Eastern Portugal using Statistical and Geostatistical Analysis. The Association of Applied Geochemists: 23rd International Applied Geochemistry Symposium, June 2007, Oviedo, Spain.

Appendix 1. List of training dataset coordinates (added Nov 15, 2022). Coordinates in EUREF-FIN.

Easting	Northing
314966	6741010
315601	6741290
315833	6741360
315688	6736540
307395	6735570
312616	6733810
313043	6733480
312948	6733390
313763	6733590
313268	6733200
314485	6733380
314331	6732509
314944	6733540
314467	6733840
313214	6732600
314458	6733020
311061	6729610
311118	6729470
312947	6731430
312672	6741410
313350	6741630
310256	6736530
312500	6737890
313638	6735810
314622	6736290
315773	6737010
316183	6737000
313005	6734460
313031	6734340
313321	6734420
311514	6733150
311797	6733210
311910	6732960
312397	6732340
312912	6732660
313070	6732580
313147	6732750
313172	6732290
312999	6732190

313395	6733690
314270	6734140
314829	6734560
315143	6733450
310142	6737290
309957	6737150
309837	6737050
310107	6735930
310005	6737270
312849	6734200
312957	6734170
313952	6734490
316336	6737140
316332	6736170
316242	6736060
316545	6734470
316659	6734220
317164	6727910
318671	6726490
317612	6725350
317701	6724750
324763	6721330
325048	6721240
324909	6721170
325967	6720890
313270	6728870
314512	6728170
309440	6730350
312194	6731570
310917	6729720
311082	6729980
311002	6729450
314833	6733045